Stream guiding algorithm for deriving flow direction from DEM and location of main streams

JIAHU WANG[1], LI LI[1], ZHENCHUN HAO[1] & JONATHAN J. GOURLEY[2]

1 *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China*
tigerlly@126.com

2 *NOAA/National Severe Storm Laboratory, Norman, Oklahoma 73072, USA*

Abstract The drainage paths and directions within the drainage basin are important for analyses of the interactions between human and nature. The *stream burning* algorithm is a popular D8-based method and can be effective in the digital reproduction of a known and generally accepted stream network. The *stream guiding* algorithm has been developed in this paper to overcome the stream burning algorithm's disadvantage of locally altering elevation in order to provide the consistency between existing vector hydrography and the DEM. In the new algorithm, flow direction of LMS (location of main streams) grids will be determined first; then possible outlets in non-LMS area will be found; and finally, the flow direction of undetermined area will be calculated by a "filling up" technique. Evaluations for Taiwan Island show that the new algorithm has a similar performance to that of the stream burning algorithm in river network reproduction. The new algorithm obeys the "steepest decent rule" and DEM data more strictly than the stream burning algorithm, especially around the LMS grids.

Key words flow direction derivation; DEM; location of main stream

1 INTRODUCTION

The drainage paths and directions within the drainage basin are important for analyses of the interactions between human and nature. Many methods of judging flow direction and extracting channel network appeared after the 1980s, including the D8 method (O'Ccallaghan & Mark, 1984), the D∞ method (Tarboton, 1997), the multi-direction method (Tarboton, 1997), and other methods to do with pits – the local elevation minima (Hutchinson, 1989). In general, a drainage direction map (DDM) can be derived from a digital elevation map (DEM) by applying standardized and automated procedures. Many software packages, in particular GIS, provide tools to derive the drainage direction for each raster cell of a DEM, by comparing the elevation of the cell to the elevation of its neighbouring cells.

The well-known D8 algorithm (O'Callaghan & Mark, 1984) is the most commonly used method for approximating flow directions on a topographic surface, and this method tracks "flow" from each pixel to one of its eight neighbour pixels. However, it is based on two simplifying assumptions: 1. the use of eight discrete flow angles; and 2. each pixel has a single flow direction (SFD), that does not capture the geometry of divergent flow over hillslopes. That is, all flow leaving any given pixel (or the area contributing flow to the pixel) is assumed to flow into a single downstream neighbour pixel. The D8 method is well-suited to the identification of individual channels, channel networks, and basin boundaries.

Deriving drainage direction from a DEM by the D8 method is a straightforward approach; however, poor quality or simply the inherent generalization of a DEM may cause derived drainage lines to differ from reality (Döll & Lehner, 2002). Thus, a number of methods for DEM improvements have been suggested, such as the removal of spurious sinks (Jenson & Domingue, 1988; Soille *et al.*, 2003), incorporation of vector stream data for stream burning (Maidment, 1996; Mizgalewicz & Maidment, 1996; Saunders, 1999) or surface reconditioning (Hutchinson, 1989, 2004; Hellweger, 1996).

Several studies have addressed this problem through a DEM post-creation modification technique utilizing the vector hydrology layer, a process commonly referred to as the *stream burning* (Mizgalewicz & Maidment, 1996; Saunders & Maidment, 1996; Hellweger, 1997). Where vector hydrography information exists, it can be integrated into the DEM prior to the actual analysis. This process is referred to as the *stream burning* and can be effective in the digital reproduction of a known and generally accepted stream network. However, it has the disadvantage of locally altering elevation in order to provide the consistency between existing vector hydrography and the DEM. Several methods exist (Hutchinson, 1989; Saunders & Maidment, 1996) but greatly differ in their success of improving watershed delineation (Saunders, 1999). The pre-processing of the vector information required often represents an intensive effort.

In this paper we present our new *stream guiding* algorithm to derive the DDM by an interactive analysis of the DEM and the location of main streams (LMS) without altering any DEM data. It inherits the advantage of the *stream burning* algorithm: bringing additional information into the DEM to position main rivers correctly; and overcomes the disadvantage of locally altering elevation. Evaluations are given by both manual inspection and a statistical comparison after detailed description of the algorithm.

2 STREAM GUIDING ALGORITHM

In this algorithm, flow direction of LMS grids will be determined first; then possible outlets in the non-LMS area will be found; and finally, flow direction of the undetermined area will be calculated by a "filling up" technique. The flow diagram in Fig. 1 shows its mechanism. The whole strategy is toward extending indirect outlets to upstream from estuary by LMS data, and making it easier for inland grids, and finding a way to all the outlets. We thus named it the *stream guiding* algorithm.

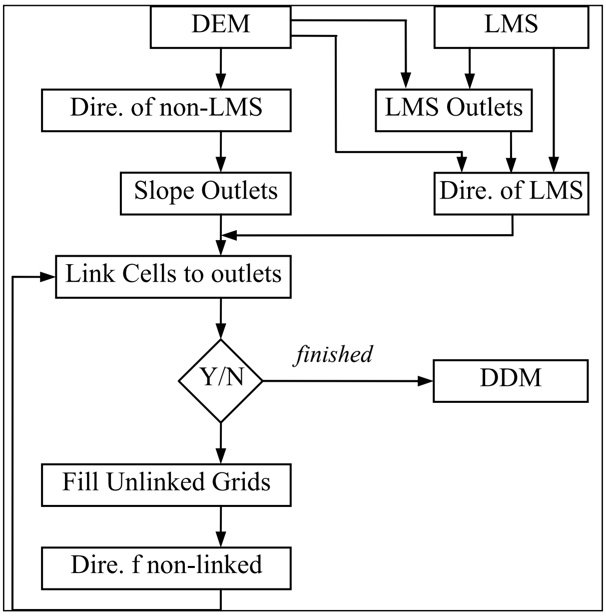
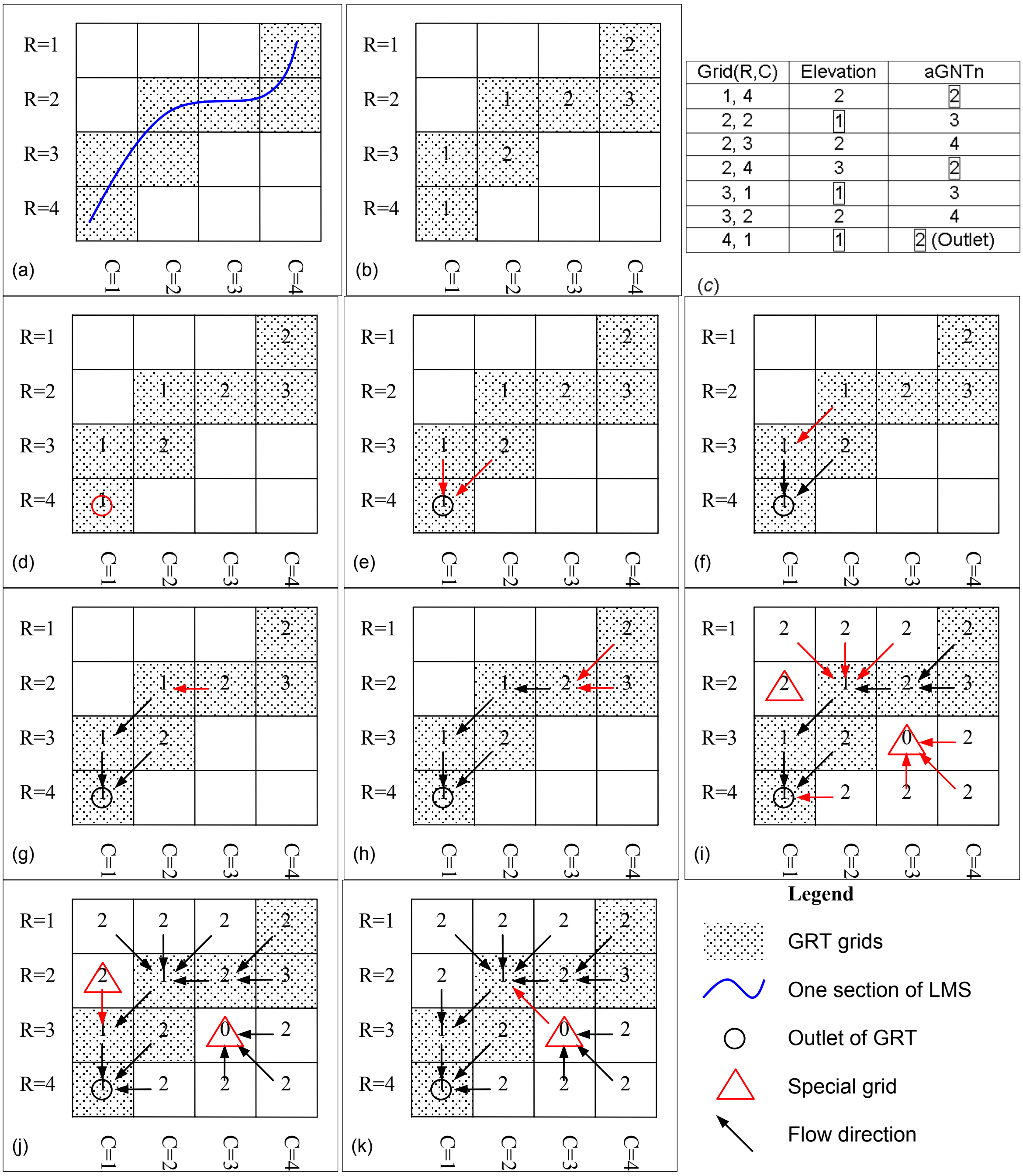


Fig. 1 Flow diagram of stream guiding algorithm (DEM is digital elevation model; LMS is location of main streams; Dire. refers to flow direction; DDM is drainage direction map).

2.1 Flow direction determination

The location of the main stream (LMS) can be derived from maps or from available data sets (e.g. the ArcWorld database) (see Fig. 2(a)). It should be converted into gridded values, as the elevation of LMS grids from DEM data is necessary for interactive analysis (see Fig. 2(b)).

To each LMS grid, the number of its adjoining LMS grids from all eight possible directions (LMS-a8n) needs to be counted first (see table in Fig. 2(c)). Then we can obtain the estuary grid by picking up the one with both minimum elevation and GNT-a8n. This is done based on a



Grid(R,C)	Elevation	aGNTn
1, 4	2	2
2, 2	1	3
2, 3	2	4
2, 4	3	2
3, 1	1	3
3, 2	2	4
4, 1	1	2 (Outlet)

Fig. 2 Theoretical description of deriving flow direction from the DEM and LMS (location of main streams).

comparison of grids in the middle of a LMS to the grid at the end of the LMS (one end of the estuary or several ends of each branch's upstream should have smaller LMS-a8n). In addition, the estuary grid is lower than the other grids. Estuary here means river outlets to both oceans and inland lakes (see Fig. 2(d)). We need to do this repeatedly to find each dendritic river network's estuary in a region.

The flow direction of all adjoining LMS grids of an estuary can be set directly by pointing to it (see Fig. 2(e)). The lower elevation grid has priority to receive water from the upper grid if they have the same upper adjoining LMS grids. For example: $grid_{(3,1)}$ and $grid_{(3,2)}$ have the same upper adjoining LMS grids $grid_{(2,2)}$ in Fig. 2(f), $grid_{(2,2)}$ should flow into $grid_{(3,1)}$ because $grid_{(3,1)}$ is lower than grid$_{(3,2)}$ (elevation 1 m *vs* 2 m). This "steepest descents" processing is repeated so that the flow direction of all the LMS grids can be determined (see Fig. 2(f)–(h)).

2.2 Flow direction determination

Our method scans all non-LMS grids with a 3 × 3 moving window, as used in the other regular D8 method. If their flow direction is pointing to any determined grids, either directly or indirectly, we upgrade them to determined grids. These grids can be found in Fig. 1(i) as grid$_{(1,1)}$, grid$_{(1,2)}$, grid$_{(1,3)}$ and grid$_{(4,2)}$.

Ambiguous flow directions (the same minimum downslope gradient is found in two cells) are usually resolved by an arbitrary assignment. Grid$_{(2,1)}$ in Fig. 1(h) is a typical example; it can flow into grid$_{(2,1)}$ and$_{(2,2)}$ according to elevation. In our method, we set it to flow into the downstream one because the downstream grid should be a bit lower than upstream one, although we can not establish that in the DEM. This means flow direction of grid$_{(2,1)}$ should point to grid$_{(3,1)}$ in Fig. 2(j).

The flow direction of local elevation minima (pits) grids can not be determined directly either. If a pit grid adjoins the ocean grid in any of the eight possible directions, it is considered as a small estuary; otherwise it needs to be filled-up and scanned again. An example of such a pit, grid$_{(3,3)}$, can be found in Fig. 2(j), its flow direction can be determined by just one cycle of filling-up and scan (see Fig. 2(k)). Actually, the process of filling-up and scan often need to be repeated hundreds or thousands times in order to determine all the grids' flow directions.

2.3 Overview of method

The stream guiding algorithm can be summarized by 10 steps as in Fig. 2: (a) determine the LMS and grid with same resolution and span as the DEM; (b) get elevation of LMS grids from the DEM data; (c) compare elevation and LMS-a8n (number of its adjoining LMS grids from eight possible directions); (d) pick up the grid with minimum elevation and LMS-a8n as outlet; (e) set flow direction of all adjoining grids of outlet pointing to outlet; (f) compare elevation of newly determined grids, pick up one with minimum elevation, set flow direction of other adjoining grids pointing to the picked one in step f; (g) repeat step f; (h) repeat step f until the flow directions of all LMS grids have been determined; (i) scan flow direction of the non-LMS grids with a 3 × 3 moving window, and determine the grids either directly or indirectly pointing to determined grids; (j) determine the flow direction of ambiguous grids (the same minimum downslope gradient is found in two cells) by pointing to the downstream grid; and (k) deal with pit(s) by the filling-up method or set it as possible small outlet.

2.4 Example

A real case illustrated in Fig. 3 shows the working of the stream guiding algorithm from a zoom out view. The region is a stochastic rectangular area from Hydro1k (Team, 2003) and the LMS used here was digitized from DEM maps. The processing of this real area is the same as described above and the grid number is the only difference. We obtain reasonable results, such as: (1) LMS map (Fig. 3(a)); (2) determined LMS map (Fig. 3(b)); (3) scan of slope outlets (Fig. 3(c)), and (4) final drainage direction map (Fig. 3(d)).

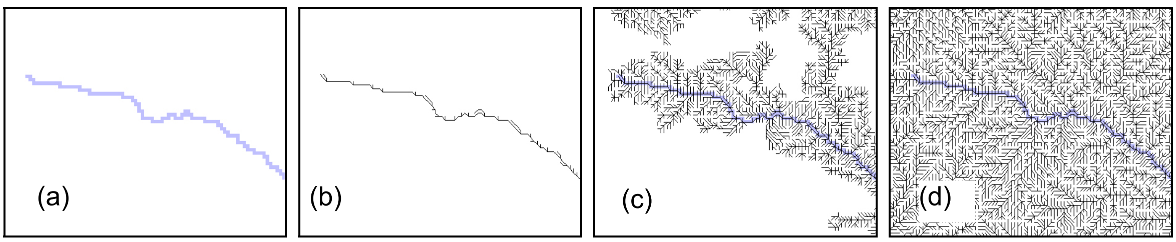


Fig. 3 An example of deriving flow direction from DEM and LMS in a stochastic rectangle area from Hydro1k; and (d) final DDM (drainage direction map).

3 EVALUATION AT AN ISLAND SCALE

A continent is a perfect region to use for an evaluation, but an island is more convenient for testing the results as it can be considered as a shrunken continent because (1) it is surrounded by ocean, and (2) has both mountain and plain areas. Taiwan Island was chosen for this study (Fig. 4(a)). Its land area is approximately 36 000 km^2 with mountains reaching 3952 m in the middle. It is an interesting island for hydrologists because in this relatively small area, the annual mean rainfall is about 2510 mm, about 2.6 times the world annual mean rainfall, but nearly 78% of the rain occurs during the wet months from May to October.

The DEM data for the island are from Hydro1k (Team 2003), and the river network is from the ArcWorld database (ESRI, 1992); 6.1% of total DEM grids were marked as LMS. Deriving flow direction according to the *stream burning* algorithm was executed by ARCGIS (ESRI, 2006); deriving flow direction according to the *stream guiding* algorithm was completed using the Channel Network Tool (Wang *et al*., 2005), a small software developed by us.

3.1 Qualitative evaluation

The performance of our algorithm is difficult to assess directly from drainage direction maps. We therefore checked it by generating a flow accumulation map (FAM). A FAM represents the number of upstream grids, i.e. the number of those grids that drain through a given grid, and is comparable to LMS.

From the FAM (Fig. 5) we can find that both stream burning (see Fig. 4(b)) and stream guiding (see Fig. 4(c)) algorithms generated perfect river networks, and especially the location of main streams. The difference between them is hard to identify visually and the two figures in Fig. 5 are so alike that one could be a duplicate of the other. This shows that the stream guiding algorithm has the same performance in deriving the river network as the stream burning algorithm.

But, their detail is really different because of the different mechanisms used in their determination. The stream burning algorithm gave a regular flow direction around the LMS grids (Fig. 5(a)), but the flow direction in such areas has no relationship with the elevation. However, flow direction generated by the stream guiding algorithm is closely related to the elevation values even that of the LMS grid (see Fig. 5(b)).

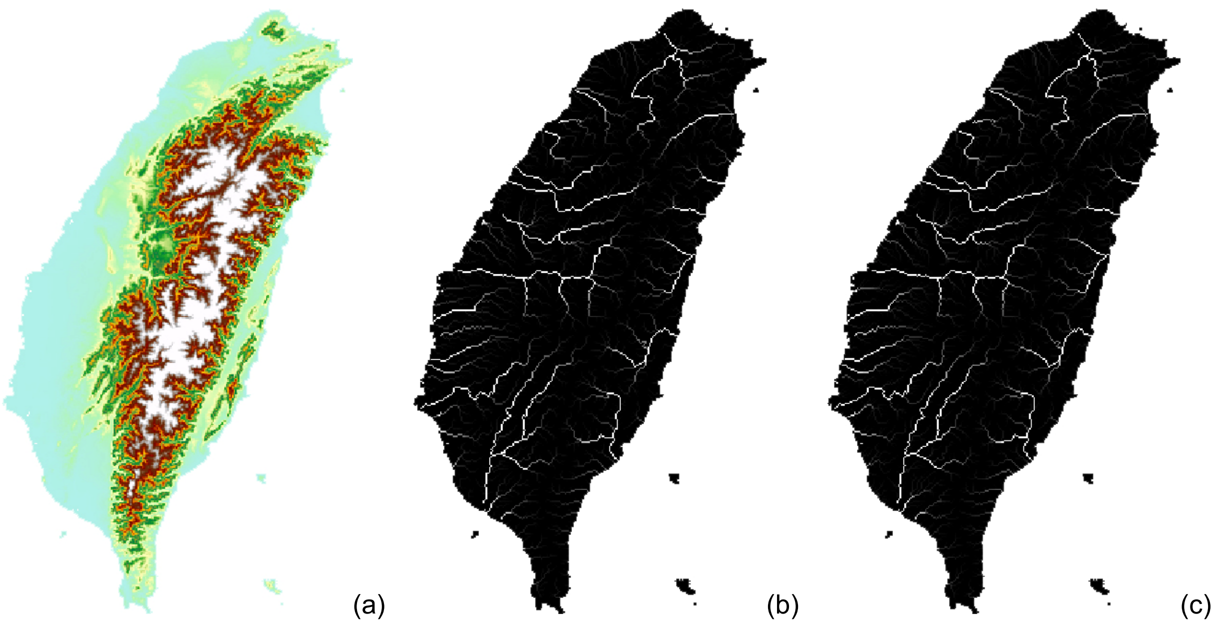


Fig. 4 (a) Study area. Overview of two FAMs (flow accumulation maps): (b) according to stream burning algorithm, and (c) according to stream guiding algorithm.

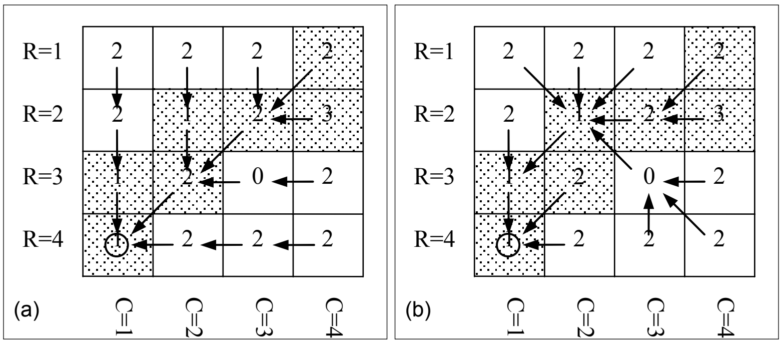


Fig. 5 Detailed view of two DDMs (drainage direction map): (a) flow direction near LMS according to stream burning algorithm has nothing to do with the elevation, and (b) flow direction generated by stream guiding algorithm sticking to the elevation values and "steepest decent rule".

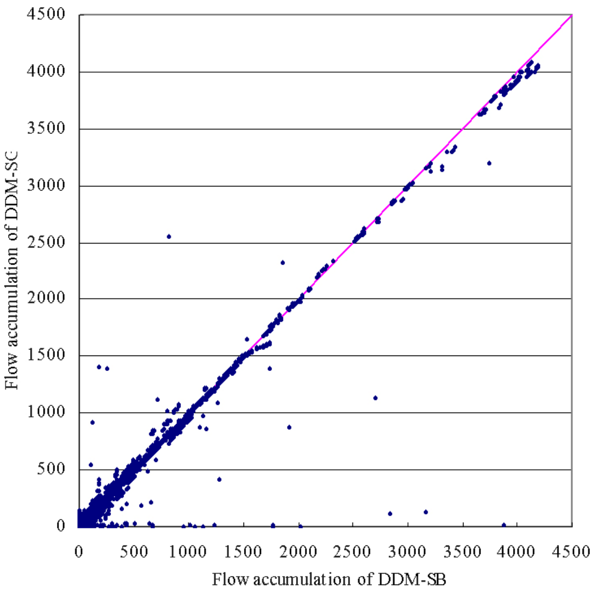


Fig. 6 Grid based comparison of FAMs between stream guiding and stream burning algorithms.

3.2 Quantitative evaluation

Grid based comparison of flow accumulation by the stream burning and stream guiding algorithms in Fig. 6 shows a strong cluster around the 1:1 line. The degree of correspondence is quantified by the modelling efficiency *ME* (Jansen & Heuberger, 1995), which is equivalent to the Nash-Sutcliffe coefficient and measures the goodness-of-fit to the line-of-perfect-fit (the 1:1 line):

$$ME = \frac{\sum_{i=1}^{J}\left(N_i^B - \overline{N}_i^B\right)^2 - \sum_{i=1}^{J}\left(N_i^G - \overline{N}_i^B\right)^2}{\sum_{i=1}^{J}\left(N_i^B - \overline{N}_i^B\right)^2} \quad (1)$$

where N_i^B is the flow accumulation of cell i according to the stream burning algorithm; $\overline{N}_i^B$ is average of all N_i^B; N_i^G is the flow accumulation of cell i according to the stream guiding algorithm; J is the total number of grids covering Taiwan Island, i.e. 46 937.

The ME here is 0.971. If we take flow accumulation according to the stream burning algorithm as the benchmark, about 75.7% grids of flow accumulation determined by the stream guiding algorithm have a bias ratio of less than ±1%, and 86.2% grids have a bias ratio of less than ±5%. This means that the river networks derived by these two algorithms are similar.

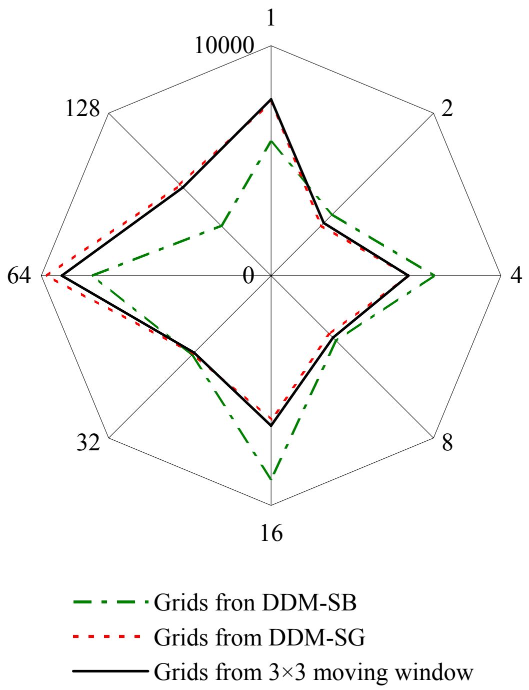


Fig. 7 Grid based comparison of DDM showing the DEM alternation in the stream burning (SB) algorithm and lost information from DEM around LMS grids.

The difference in detail between the two DDMs is also remarkable. The numbers of grids inclining to each direction are shown in Fig. 7. The flow direction of 93.2% grids according to the stream guiding algorithm follow the "steepest decent rule", when compared to the results of a 3 × 3 moving window. Only 75.1% of grids determined by the stream burning algorithm follow this rule, because the elevation of many grids has been lowered through a DEM post-creation modification technique. This result shows that the stream guiding algorithm respects the "steepest decent rule" more strictly than the stream burning algorithm.

4 CONCLUSION

The D8 method is widely used and implemented in many GIS software packages, despite its limitations. It is useful for a number of applications, such as extracting river network maps, longitudinal profiles, and basin boundaries. The s*tream burning* algorithm is a good method to obtain accuracy for dendritic river networks by importing the location of main streams into a DEM. But the alteration of the DEM leads to the loss of some information from the DEM around LMS grids; hence, we develop the *stream guiding* algorithm. Evaluations show that the new algorithm has a similar and good performance to the stream burning algorithm in reproducing river networks. The new algorithm follows the "steepest decent rule" and DEM data more strictly than the stream burning algorithm, and so the detail, especially around the LMS grids, is better.

Acknowledgements The financial support from the National Natural Science Foundation of China (50879016, 40830639, 40801012) is gratefully acknowledged. The authors would like to extend their appreciation to the Fundamental Research Funds for the Central Universities and the Special Fund of the State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering (1069-50985512).

REFERENCES

Döll, P. & Lehner, B. (2002) Validation of a new global 30-min drainage direction map. *J. Hydrol.* **258**(1-4), 214–231.

ESRI (1992) Global topographic database ArcWorld 1:3M on CDROM, Environmental Systems Research Institute, Inc. (ESRI).

ESRI, I. (2006) ESRI ArcMap 9.2.Redlands, USA.

Hao, Z., Wang, J. *et al*. (2005) Principium and function of Channel Network Tool-I (in Chinese). *Hydrology (China)* **25**(2), 15–19.

Hellweger, F. L. (1997) AGREE – DEM Surface Reconditioning System. http://www.ce.utexas.edu/prof/maidment/gishydro/ferdi/research/agree/agree.html as of July 1999.

Hellweger, R. (1996) Agree.aml. Center for Research in Water Resources, The University of Texas at Austin, Austin, Texas, USA.

Hutchinson, M. (1989) A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *J. Hydrol.* **106**(3-4), 211–232.

Hutchinson, M. F. (2004) ANUDEM Version 5.1 User Guide. http://cres.anu.edu.au/outputs/software.php. The Australian National University, Centre for Resource and Environmental Studies, Canberra, Australia.

Jenson, S. & Domingue, J. (1988) Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogram. Engng Remote Sens.* **54**(11), 1593–1600.

Maidment, D. R. (1996) GIS and hydrological modeling: an assessment of progress. Presented at the Third International conference on GIS and Enviromental Modeling, Santa Fe, New Mexico, January 1996.

Mizgalewicz, P. J. & Maidment, D. R. (1996). Modeling agrichemical transport in midwest rivers using geographic information systems. Center for Research in Water Resources Online Report 96-6. University of Texas, Austin, Texas, USA.

O'Callaghan, J. &. Mark, D. (1984) The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics, and Image Processing* **28**(3), 323–344.

Saunders, W. (1999) Preparation of DEMs for use in environmental modeling analysis. ESRI User Conference, 24–30 July 1999, San Diego, California (http://www.esri.com/library/userconf/proc99/proceed/papers/pap802/p802.htm).

Saunders, W. & Maidment, D. (1996) A GIS assessment of nonpoint source pollution in the San Antonio-Nueces Coastal Basin. Center for Research in Water Resources Online Report: 96–91.

Soille, P., Vogt, J. V., *et al*. (2003) Carving and adapting drainage enforcement of grid digital elevation models. *Water Resour. Res.* **39**(12), 1366–1378.

Tarboton, D. (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.

Team, H. (2003) Web page of the HYDRO1k Elevation Derivative Database. Technical CLAVR-1 global clear/cloud classification algorithm for the Advanced Very High Resolution.

Wang, J., Hao, Z., *et al*. (2005) Extraction of drainage structure from digital elevation model by combination with raster river network (in Chinese). *J. Hohai University (Natural Sciences)* **33**(2), 119–122.