ELSEVIER

Journal of Hydrology 240 (2001) 225–242

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network

R. Turcotte*, J.-P. Fortin, A.N. Rousseau, S. Massicotte, J.-P. Villeneuve

Institut National de la Recherche Scientifique — Eau, Université du Québec, 2800, rue Einstein, C.P. 7500, Sainte-Foy, Québec, Canada G1V 4C7

Received 18 June 1999; revised 11 July 2000; accepted 19 September 2000

Abstract

Distributed hydrological models require a detailed definition of a watershed's internal drainage structure. The conventional approach to obtain this drainage structure is to use an eight flow direction matrix (D8) which is derived from a raster digital elevation model (DEM). However, this approach leads to a rather coarse drainage structure when monitoring or gauging stations need to be accurately located within a watershed. This is largely due to limitations of the D8 approach and the lack of information over flat areas and pits. The D8 approach alone is also unable to differentiate lakes from plain areas.

To avoid these problems a new approach, using a digital river and lake network (DRLN) as input in addition to the DEM, has been developed. This new approach allows for an accurate fit between the DRLN and the modelled drainage structure, which is represented by a flow direction matrix and a modelled watercourse network. More importantly, the identification of lakes within the modelled network is now possible. The proposed approach, which is largely rooted in the D8 approach, uses the DRLN to correct modelled flow directions and network calculations. For DEM cells overlapped by the DRLN, flow directions are determined using DRLN connections only. The flow directions of the other DEM cells are evaluated with the D8 approach which uses a DEM that has been modified as a function of distance to the DRLN.

The proposed approach has been tested on the Chaudière River watershed in southern Québec, Canada. The modelled watershed drainage structure showed a high level of coherence with the DRLN. A comparison between the results obtained with the D8 approach and those obtained by the proposed approach clearly demonstrated an improvement over the conventionally modelled drainage structure. The proposed approach will benefit hydrological models which require data such as a flow direction matrix, a river and lake network and sub-watersheds for drainage structure information. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Digital elevation model; Digital river and lake network; Hydrological modelling; Watershed delineation

1. Introduction

The use of distributed hydrological models for the simulation of river flows has become widely used in hydrological engineering. Distributed models account for the spatial variability of physical properties and allow for the spatial assessment of modelled hydrological variables. To achieve this type of representation, distributed models require data that define the internal drainage structure of the watershed. Most of the available algorithms (Tribe, 1992) automatically extract the drainage information, such as flow directions from cell to cell, river network segments and associated sub-watersheds, from a digital elevation model (DEM).

---

* Corresponding author. Fax: +1-418-654-2600.
E-mail address: richard_turcotte@inrs-eau.uquebec.ca (R. Turcotte).

PII: S0022-1694(00)00342-5

For example, the Olivera and Maidment (1999) rainfall–runoff model for routing spatially distributed excess precipitation requires a flow direction matrix for the calculation of unit hydrographs over sub-watersheds. The ANSWERS model (Beasely et al., 1980) uses a flow direction matrix and a river segment network as inputs, although the procedure to obtain a network that matches the flow directions is not completely automated (Joao and Walsh, 1992). The TOPMODEL (Beven and Kirkby, 1979) uses flow directions to compute the topographic index. In the case of TOPMODEL the extraction of flow directions from a DEM has been studied by various researchers (Quinn et al., 1991; Wolock and McCabe, 1995). The Cabral et al. (1991) rainfall–runoff model utilises a river network obtained from a DEM. Other models such as SLURP (Kite, 1995) and PRMS (Leavesley and Stannard, 1995), to name a few, use sub-watersheds as calculation units. Finally, HYDROTEL (Fortin et al., 1995, 2001) directly uses a direction matrix, a river segment network and associated sub-watersheds as inputs. All these models, and others, could benefit at various levels from a fully automated and accurate determination of a watershed's internal drainage structure.

From a hydrological point of view, the current widespread DEM-based approaches have two major limitations. First, on a macroscopic scale, data modelled using a DEM are representative of the actual drainage structure of a watershed, but on a smaller scale, a perfect fit between these data and the actual terrain features is never obtained (Tribe, 1992). For example, the watershed drainage structure modelled using a DEM does not fit with the actual drainage structure in flat areas. Secondly, by definition, a DEM does not contain information about lakes. It is impossible to determine if a given area of equal elevation is either a lake or a flat area, where a river possibly flows. When drainage structure data are used in a hydrological simulation context, the difference between these two cases needs to be assessed because water flows differently in a lake than in a river. These two limitations implies there exists a gap between data modelled from a DEM and actual drainage data required by some distributed hydrological models.

One way to address the limitation related to errors in the location of the modelled watercourse network is to use a so-called "stream burning" approach (Saunders, 1999). This kind of approach, introduced by Hutchinson (1989), proposes the use of ancillary information regarding watercourse network, namely the stream lines of digital maps, into the DEM to force flow through cells corresponding to the stream line network. Although the algorithms presented in Saunders (1999), and reviewed briefly later in this paper, are subject to some problems and do not account for lakes, the idea to use conjointly ancillary information about the watercourse network and the DEM is promising.

This paper, which is organised in four sections, presents the development of a new approach that forces a watershed's internal drainage structure to fit accurately with a given ancillary watercourse network with an emphasis on differentiation of lakes from flat areas. The first section presents an overview of the conventionally used approach for the modelling of watershed drainage structures. The second section provides some information about digital river and lake network (DRLN) data sources. The third section describes the proposed approach and the fourth section describes an application on a southern Québec watershed.

To avoid any confusion, the vector and raster-based representations of the ancillary network will be referred to as the DRLN (vector-DRLN and raster-DRLN). The watercourse network obtained using the conventional DEM approach or the proposed approach will be referred to as the modelled network with a direct reference to the approach used.

Finally, the reader should note that, despite the fact that the proposed algorithm is a general one, this algorithm has been integrated to HYDROTEL (Fortin et al., 1995, 2000). HYDROTEL is a distributed hydrological model designed to take advantage of remote sensing data and GIS. Note that HYDROTEL can be used as a standalone version or as part of GIBSI, an integrated modelling system for watershed management (Villeneuve et al., 1998; Rousseau et al., 2000). Simulations are used to assess the impact of watershed management scenarios on water quality and yield, both in time and space, where water pollutants from agricultural nonpoint sources as well as from municipal and industrial point sources are considered. For the latter, an accurate link between point load locations and their actual location within the modelled network must be clearly established. Also, for

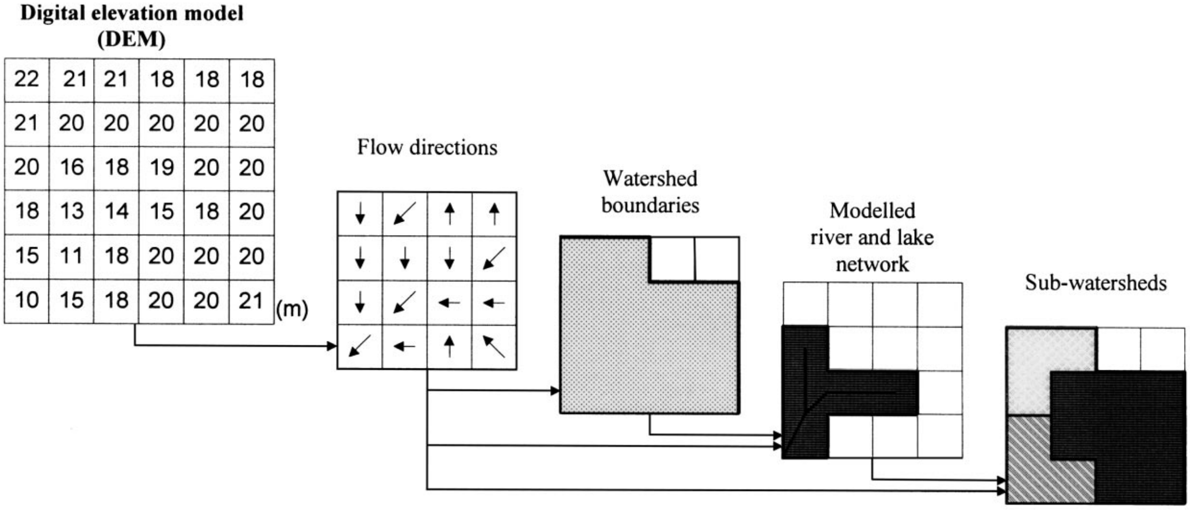


Fig. 1. Flow structure data of a watershed determined using a digital elevation model.

accurate model calibration, streamflow gauges and water quality stations need to be properly located on the network. Finally, when consulting distributed simulated results, users of GIBSI must be able to easily link these segments with those actually surveyed.

2. Computerized extraction of watershed drainage structure using a DEM

2.1. Conventional algorithm

For distributed hydrological models which require a square grid to support physiographic data and rely on a continuous drainage network where each grid cell must be directly connected to one, and only one, of its neighbouring cells, the eight flow directions (D8) is a valid approach to model the watershed drainage structure (Tribe, 1992). Implementation of the D8 approach allows for a coherent link between flow directions, river and lake network, and sub-watersheds associated with each segment of the watercourse network (e.g. Martz and Garbrecht, 1992). Numerous algorithms based on the D8 approach have been reported in the literature (Martz and Garbrecht, 1993, 1992; Jenson and Domingue, 1988; Martz and De Jong, 1988; Morris and Heerdegen, 1988; O'Callaghan and Mark, 1984). These algorithms, which are conceptually similar, are mainly distinguishable by the way they treat problematic cases (e.g. definition of flow directions over flat areas and pits).

The implementation of the D8 approach requires four steps. These steps are illustrated in Fig. 1. In the first step, the flow direction of each cell of the DEM matrix is computed by the steepest downstream slope among the eight slopes obtained. Since the D8 approach cannot determine the flow direction of a cell for all possible cases — one of those cases is that of a cell surrounded by cells with the same elevations — an arbitrary criterion is used to assign that flow directions (Garbrecht and Martz, 1997). Examples of strategies to assign flow directions are presented in Jenson and Domingue (1988) and Garbrecht and Martz (1997). In the second step, the watershed boundary is defined by finding all cells for which there is a path to the watershed outlet, starting at the cell which includes the watershed outlet and following flow directions in the opposite direction. The third step involves the determination of the modelled river network. The drainage area upstream of any cell can be computed by counting cells that flow through this given cell. Those cells having a drainage area greater than a threshold, that can be varied depending on the approach used (Tarboton et al., 1991), are considered to be part of the river network. Note that cells which are part of the river network are grouped by segment where a segment is defined for each part of the network located between either: (1)

two confluence cells, (2) one confluence cell and one network upstream cell or (3) one confluence cell and the outlet cell. In the fourth and final step, using the computed flow directions, the sub-watershed associated with each modelled river network segment is determined for a given segment by finding all the cells flowing through one of the cells of the network segment.

2.2. Limitations of the D8 approach

The D8 approach does not allow for an exact match between modelled and actual river networks. In addition to the basic assumption error (Martz and Garbrecht, 1995; Carrara, 1986), three sources of error can explain discrepancies between the networks: (1) the inherent limitations associated with the D8 approach, (2) the presence of flat areas and pits, and (3) the lack of information on the locations of lakes.

2.2.1. Eight flow directions error

The representation of flow directions using only eight possible directions implies that only these eight directions are used to approximate the continuous flow direction field. Thus, a given discrete direction is used to replace all flow directions included within an angular interval of $\pi/4$ and centred on the discrete direction resulting in a loss of information about the actual flow path of the terrain.

This loss of information, associated with the more generally admitted effect of vertical and horizontal resolutions of DEM, generates parallel flow paths (Fairfield and Leymarie, 1991). Hence, for a group of cells, when the variation of the surface aspect is included inside an angular interval covered by one of the eight discrete directions, a unique flow direction value is assigned to all cells in the group. Consequently, two basic problems may occur: (1) the absence of existing river segments, and (2) the presence of parallel river segments where only one segment actually exists. In the first case, since the flow paths are parallel and do not converge, the threshold for deciding if a given cell is part of the network is not exceeded. In the second case, cells will have an upstream area larger than the threshold area, and the slope in the direction of the main flow becomes much larger than the perpendicular slope.

2.2.2. Presence of flat areas and pits

Because of a lack of data on flow paths over flat areas and pits in the DEM, determination of flow directions over these areas requires an assessment of artificial flow paths, defined arbitrarily. Thus, an accurate modelling of the river network over these areas is impossible. This is particularly true in large valleys with gentle slopes where it is practically impossible to follow meanders. Problems related to the treatment of flat areas and pits have been studied extensively by Martz and Garbrecht (1998).

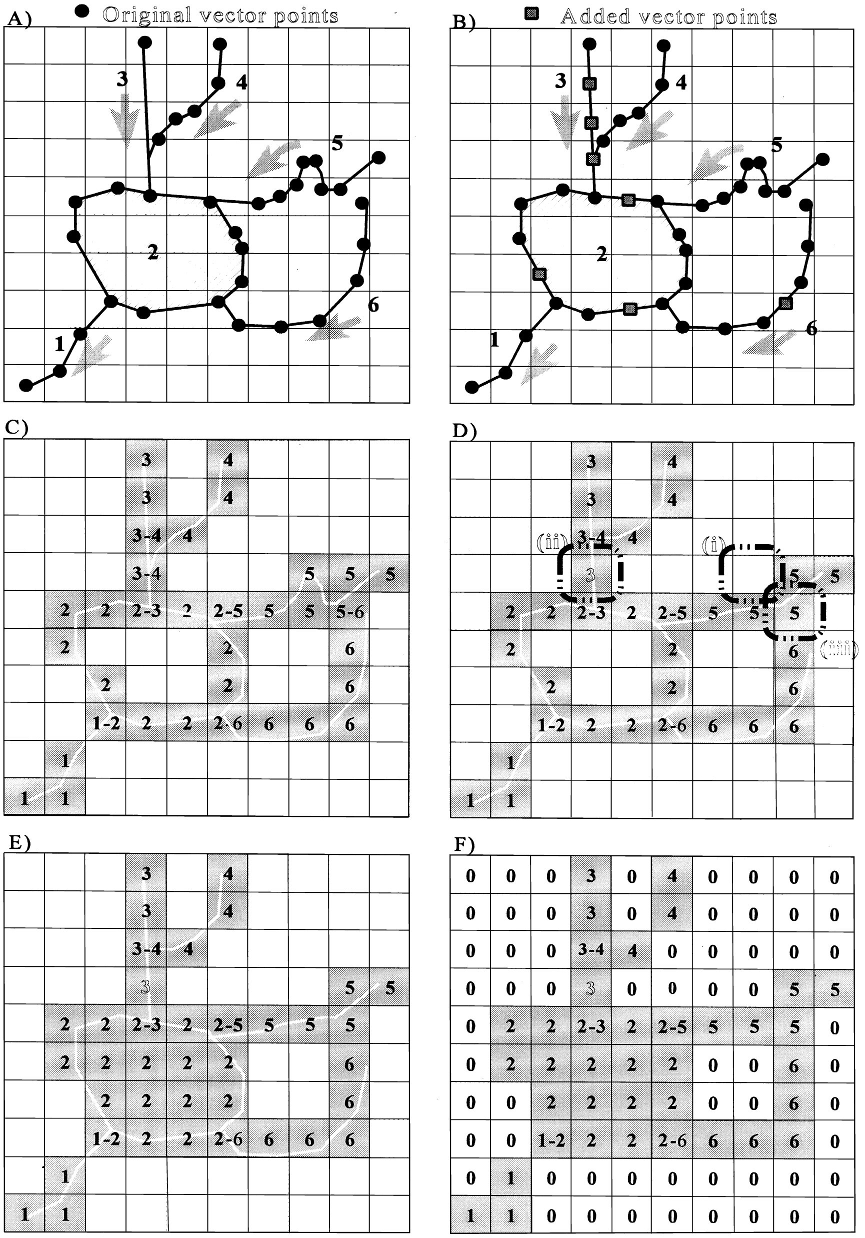
2.2.3. Absence of information about lake locations

With the D8 approach, the upstream area of a cell cannot be split between neighbouring cells in the case where the watercourse is wider than one cell. Furthermore, the D8 approach cannot estimate whether or not a watercourse has a width greater than one cell. In fact, the D8 approach is inappropriate for the identification of wide segments of a river network, such as those that include lakes. It is important to point out that the DEM, which is the only input datum required by the D8 approach, does not include information about lake locations. Therefore, the exclusive use of a DEM is not sufficient to delineate whether a constant elevation is either a lake or a flat area.

2.2.4. Discussion

The aforementioned sources of errors resulting from the D8 limitation, and the incapacity to determine lakes are intrinsic to the D8 approach. This means that these errors cannot be eliminated by a better DEM or by a decrease in the underlying grid size. To circumvent these sources of fundamental errors, it is necessary to use a different approach. Note that the error associated with flat areas and artificial pits is not, in principle, fundamental since it can be eliminated with a more accurate DEM. Obviously, in an operational context, such a DEM is not normally available and flat areas and pits represent problematic surfaces.

Some authors have attempted to bypass the D8 limitations. Fairfield and Leymarie (1991) worked on problems related to parallel flow lines. They inserted a random alteration in the evaluation procedure of the flow direction. The aspects resulting from this approach have a realistic appearance but the resulting values do not correspond to the actual



			3		4				
			3		4				
			3-4	4					
			3-4				5	5	5
	2	2	2-3	2	2-5	5	5	5-6	
	2				2			6	
		2			2			6	
		1-2	2	2	2-6	6	6	6	
	1								
1	1								

			3		4				
			3		4				
			3-4	4					
			3					5	5
	2	2	2-3	2	2-5	5	5	5	
	2				2			6	
		2			2			6	
		1-2	2	2	2-6	6	6	6	
	1								
1	1								

			3		4				
			3		4				
			3-4	4					
			3					5	5
	2	2	2-3	2	2-5	5	5	5	
	2	2	2	2	2			6	
		2	2	2	2			6	
		1-2	2	2	2-6	6	6	6	
	1								
1	1								

0	0	0	3	0	4	0	0	0	0
0	0	0	3	0	4	0	0	0	0
0	0	0	3-4	4	0	0	0	0	0
0	0	0	3	0	0	0	0	5	5
0	2	2	2-3	2	2-5	5	5	5	0
0	2	2	2	2	2	0	0	6	0
0	0	2	2	2	2	0	0	6	0
0	0	1-2	2	2	2-6	6	6	6	0
0	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0

Fig. 2. An example of DRLN conversion from vector to raster-based format: (A) initial digital network in vector format; (B) re-sampling of the vector network, (C) conversion of vector points in row and column coordinates; (D) elimination of problematic cases — (i) cells which need a double-direction; (ii) artificial confluences; (iii) repeated confluences; (E) lakes filling; and (F) cells which are not included in the network are assigned a value of zero.

features of the terrain. Other authors, including Tarboton (1997), Costa-Cabral and Burges (1994) and Quinn et al. (1991), used non-discrete flow directions and/or multiple flow directions from one cell to resolve this issue. These strategies have generally produced better results compared with the standard D8 approach but have the disadvantage of eliminating the unimodal link between flow directions and river network location. Note that these strategies only remove errors linked to flow directions and other sources of errors have remained.

The use of ancillary data related to the locations of rivers and lakes appears to be an inevitable way to produce a better watershed drainage structure. This kind of additional data may then be used to obtain information on both the location of rivers within flat areas and pits and the shape of lakes.

3. Digital river and lake network

3.1. Data sources on network location

For this study, the DRLN used for the determination of the drainage structure of the study watershed comes from digital maps stored in a vector format. The location of the network is then defined from groups of points, each group defines a line that represents a part of the network. The link between groups of points can be found easily by using the extreme points of each group that are included in more than one group (Fig. 2A).

The advantage of using a vectorized network as basic input comes from the ease of recognising the continuous nature of the network. The potential to establish a permanent link between the network used for hydrological modelling and the corresponding maps is another advantage. However, several authors (e.g. Gandolfi and Bischetti, 1997; Coffman et al., 1972) acknowledged that some discrepancies may exist between the field network and the network extracted from maps (e.g. in terms of location for first- and second-order segments as well). Nevertheless, the use of the network extracted from maps is a step in the right direction since the accuracy of the network extracted from maps, in terms of horizontal location, is generally better than that of the conventional, raster-based, grid representation. Furthermore, the use of a threshold approach, variable or not, for the upstream points of the network is independent of the classification order.

3.2. Network conversion from vector to raster

The D8 approach is organised around a raster representation of the data. This means that it is necessary to convert the vector-DRLN into a raster format. Before doing that, vector-DRLN generally needs a pre-treatment to ensure (1) that each river segment is represented by a single line, this implies that rivers defined by a right and a left bank must be redefined by a single vector, and (2) that lakes are represented by closed polygons. For the second case, loops occurring in the vector layer must be identified as lakes or as natural loops by using alternative information, like paper maps or remote sensing images. A close polygon must be made for each loop representing a lake. For natural loops, that are actually bifurcation around islands, one main flow path must be conserved entirely while the other flow path must be disconnected at the point where the bifurcation occurs (see Saunders, 1999).

Fig. 2 presents an example of a transformation of an already pretreated vector-DRLN to raster-DRLN. The artificial vector-DRLN shown in Fig. 2A illustrates the most often encountered problematic cases. The steps involved are:

(1) re-sampling of all vector points in such a way that at least one point on each cell crossed by the initial vector exists (Fig. 2B);
(2) conversion of vector points from spatial (x, y) coordinates to raster (row, column) coordinates (Fig. 2C);
(3) identification and correction of cases where the use of a raster representation instead of a vector representation causes problems (Fig. 2D);
  (i) removal of points which are incompatible with the definition of only one flow direction over a cell;
  (ii) removal of consecutive confluences;
  (iii) removal of artificial confluences caused by the crossing of a given cell by two different vectors;
(4) filling of lake areas (Fig. 2E).

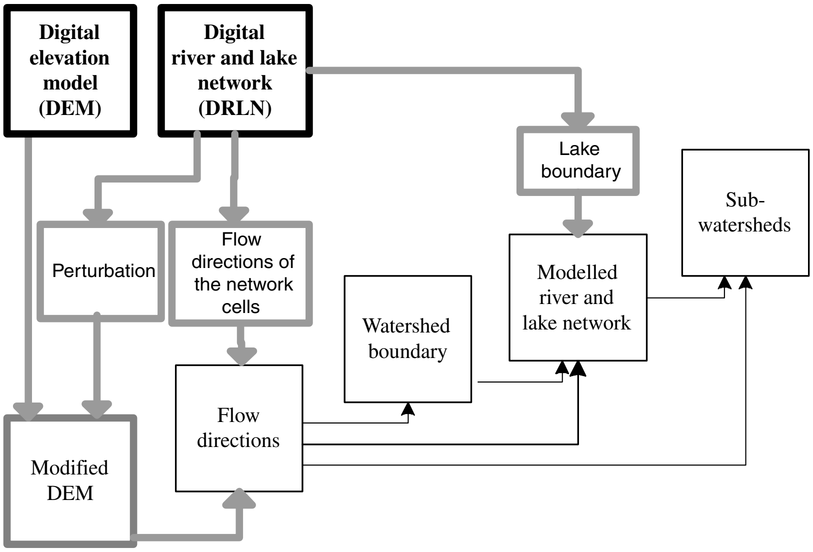


Fig. 3. Relations between data sets for the proposed approach. Quadrilaterals with wide and narrow lines are, respectively, used for input data and modelled data. Arrows represent algorithmic links between data sets. Narrow lines are used for the D8 approach algorithm and wide lines, associated with boxes, are used for the algorithms of the proposed approach.

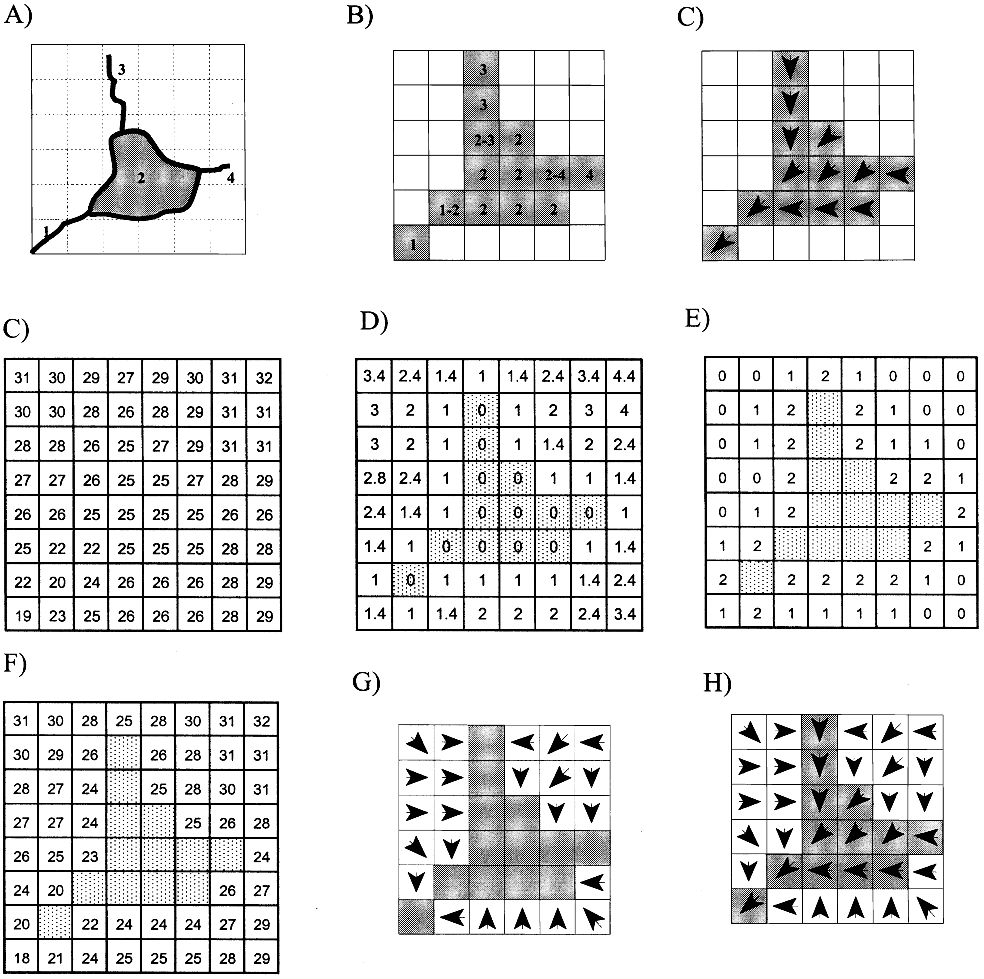
The values taken by the raster-DRLN are: (1) zero for cells which are not part of the network, (2) the identification number of a vector for cells which are included in only one vector, and (3) a list of identification numbers for cells included in more than one vector — confluence cells (Fig. 2F).

Note that in the third step, the majority of cases can be corrected automatically. For each case, it is necessary to examine each line of the vector-DRLN starting upstream. For case 3-i, the strategy is to remove from the raster-DRLN each cell that needs to flow back to its own upstream cell. As a matter of fact, this approach eliminates all meanders that are smaller than the size of a grid cell. For case 3-ii, the strategy is to keep the most upstream confluence and to keep the identifier of only one network segment for following confluences by removing the other(s) arbitrarily. Finally, for case 3-iii, confluences not located at the downstream cell of at least one network segment must be identified as artificial confluences and a list of these must be given to the user in order to force a manual correction. In one particular circumstance, when the artificial confluence corresponds to the upstream-end of a head network segment (see Fig. 2D), it is easy to remove automatically this cell from the raster-DRLN.

4. Proposed approach

The goal of this study is the development of an automated approach to determine the flow structure data of a watershed making use of a DRLN. The determination of drainage structure data with the D8 approach uses a sequence of operations where data are linked. In this context, the insertion of information about network location, coming from an ancillary source (the raster-DRLN), cannot be done directly without modifying the sequence of operations. This means that if the modelled network is directly replaced by the raster-DRLN, a lot of discrepancies between this network and flow directions will occur. This implies that the raster-DRLN and the DEM must be used together for the determination of flow directions (see Fig. 3).

As mentioned earlier, the use of raster-DRLN with DEM to improve the automated extraction of watershed's internal structure has already been introduced by Hutchinson (1989). Saunders (1999) also reviewed and proposed some algorithms that used the original D8 approach with a modified DEM. Thus, in these algorithms, the raster-DRLN is used only to perform the modification of the DEM. To give an idea, the simplest approach allows for an addition of 5 m for



C)

31	30	29	27	29	30	31	32
30	30	28	26	28	29	31	31
28	28	26	25	27	29	31	31
27	27	26	25	25	27	28	29
26	26	25	25	25	25	26	26
25	22	22	25	25	25	28	28
22	20	24	26	26	26	28	29
19	23	25	26	26	26	28	29

D)

3.4	2.4	1.4	1	1.4	2.4	3.4	4.4
3	2	1	0	1	2	3	4
3	2	1	0	1	1.4	2	2.4
2.8	2.4	1	0	0	1	1	1.4
2.4	1.4	1	0	0	0	0	1
1.4	1	0	0	0	0	1	1.4
1	0	1	1	1	1	1.4	2.4
1.4	1	1.4	2	2	2	2.4	3.4

E)

0	0	1	2	1	0	0	0
0	1	2		2	1	0	0
0	1	2		2	1	1	0
0	0	2			2	2	1
0	1	2					2
1	2					2	1
2		2	2	2	2	1	0
1	2	1	1	1	1	0	0

F)

31	30	28	25	28	30	31	32
30	29	26		26	28	31	31
28	27	24		25	28	30	31
27	27	24			25	26	28
26	25	23					24
24	20					26	27
20		22	24	24	24	27	29
18	21	24	25	25	25	28	29

Fig. 4. Example of flow directions determination: (A) vector view of the DRLN, (B) raster view of the DRLN, (C) flow directions for cells overlapped by a DRLN, (D) vector-DRLN, (E) raster-DRLN, (F) DEM, (G) distance, in cell units, between a given cell and the vector-DRLN, (E) example of perturbation coefficient calculated with $R_m = 2$ and rounded to the nearest integer, (F) modified DEM, (G) flow direction of cells not overlapped by the DRLN. Numbers (1–4) are used to identify network segments.

all cells that are not included in the raster-DRLN (Saunders and Maidment, 1995). Another approach, called the Tribburn approach, modifies the elevation of the cells included in the raster-DRLN in such way that the resulting elevations are coherent with the flow directions within the watercourse network. Thus, the cells of the raster-DRLN are offset from the land surface cells by two units.

The idea of using a modified DEM has been adopted for the proposed approach. The main limitations of modified DEM approaches are that (1) they are unable to cope with lakes and (2) they will produce parallel flow paths when the raster-DRLN differs for more than one cell with the modelled network (Saunders, 1999). Considering these limitations, the approach proposed in the present paper must include both the raster-DRLN and the DEM. It also appears that modifications to the DEM must by done in a such way that coherent modifications must be performed on a buffer zone around the raster-DRLN in order to smooth out any adjustments.

Consequently, the predominance of the DEM over

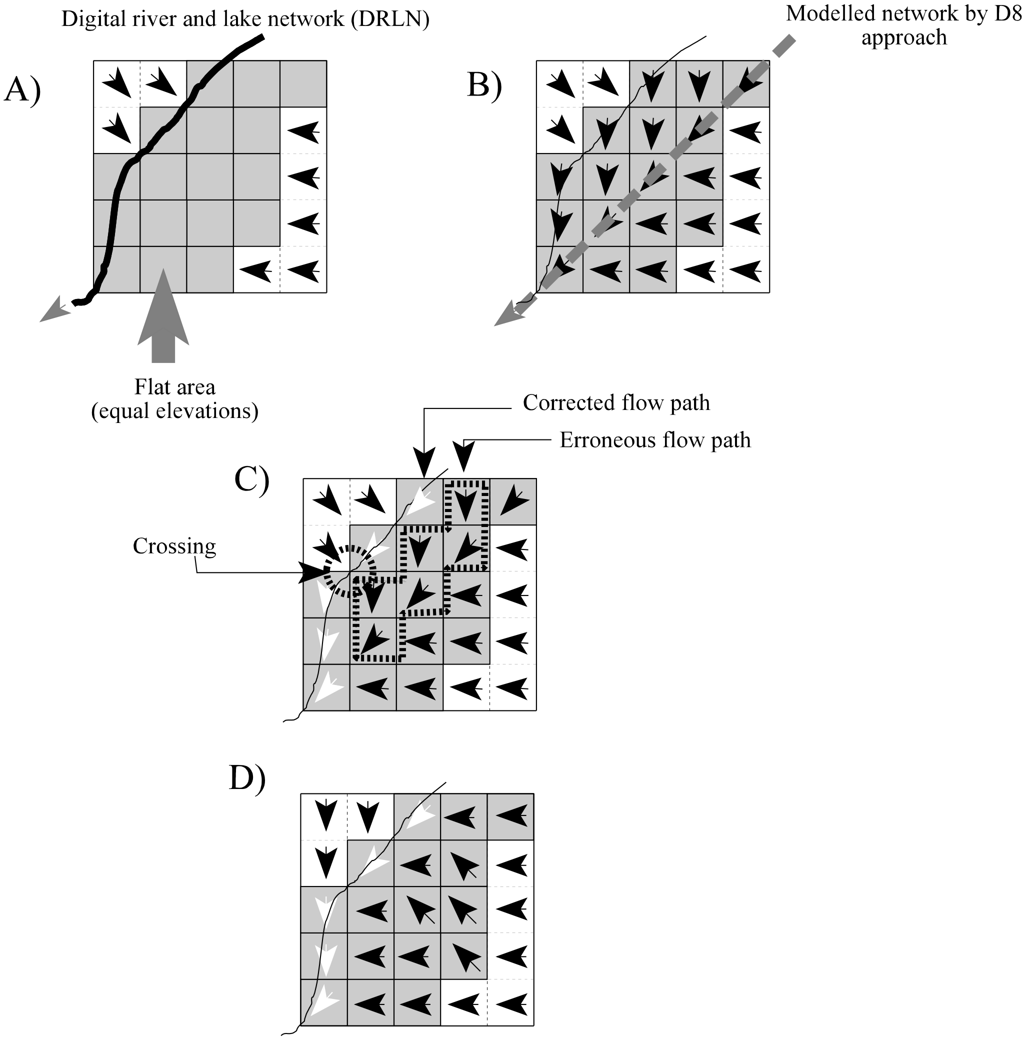


Fig. 5. An example of discrepancies generated by the unique adjustment of the flow direction of cells overlapped by the DRLN: (A) initial case, (B) modelled network and flow directions obtained using the D8 approach, (C) example of discrepancies in the direction field, and (D) example of expected flow directions after the use of the proposed approach (white arrows represent modified flow directions after using the DRLN).

the raster-DRLN must depend on the proximity of the cell over which the flow direction is being evaluated with respect to the location of the DRLN. Hence, in the proposed approach cells that are crossed by the DRLN have their flow directions determined solely by the DRLN. Other cells have their flow directions influenced increasingly by the DEM as their distance from the DRLN increases. According to these postulates, the flow directions matrix is built in two distinct phases. First, the flow directions of cells that are overlapped by the DRLN are determined. Second, flow directions of cells that have not been determined yet are processed.

More information about the drainage structure of the watershed can be obtained by using algorithms based on the D8 approach. Note also that each time that values, such as slopes and other geometrical variables, need to be computed by using elevations, the original DEM is used. The modified DEM serves only as a computational strategy to obtain flow directions. However, the modelled network cannot be completely obtained with the D8 approach. The latter is improved by combining the D8 approach with an algorithm which accounts for the DRLN and provides information about the locations of lakes. This algorithm will be presented later.

4.1. Flow direction determination of cells overlapped by the DRLN

A raster-based representation of the DRLN is used to locate those cells which are overlapped by the DRLN. Starting at the outlet, the algorithm goes back upstream following the outline of the DRLN and assigns the flow directions of visited cells so that each cell flows into the cell located directly downstream. Only the flow direction determination of cells which are part of lakes is done differently.

Since the flow directions of lake cells cannot be determined with a downstream cell, they are grouped together and flow directions are determined to minimise the distance between any lake cell and the lake outlet. Note that flow directions on lake cells are arbitrary and are used only to create a flow directions matrix that does not contain any undetermined cells. Fig. 4A–C illustrates an example of flow directions determination for the cells which are part of a DRLN.

4.2. Flow direction determination of cells not overlapped by the DRLN

The flow direction determination of cells overlapped by the DRLN ensures local coherence between the modelled network and the DRLN. However, it does not have a global effect on the watershed drainage structure. Local corrections of the flow direction of cells that belong to the DRLN can create discrepancies in the flow direction field (see Fig. 5). Indeed, it can convert the initial problem of network fitting to a problem of coherence between modelled sub-watersheds and the DRLN.

For example, flow directions in the neighbourhood of the network must be determined in such a way that they are consistent with the topographical interpretation, which is implicitly embedded in the path of the DRLN. To achieve such consistency, a modified DEM is derived from the original one by modifying the elevations using the distance between each grid cell and the location of the DRLN and by determining flow directions with the D8 approach.

The modified DEM is obtained by remodelling the surface of the terrain in the vicinity of the DRLN. This task is done by subtracting a perturbation coefficient from the original elevations of the DEM using the following equation:

$$E'(i,j) = E(i,j) - P(i,j) \tag{1}$$

where i,j are the row and column coordinates of a given cell; $E(i,j)$, the initial elevation, given by the original DEM, of the cell located at (i,j); $E'(i,j)$, the modified elevation of the cell located at (i,j); and $P(i,j)$, the perturbation coefficient of the cell located at (i,j).

The perturbation coefficient should increase as we get closer to the DRLN. Inversely, it should vanish gradually as we move away from the DRLN. We choose to express the perturbation coefficient using an inverse power-law function.

$$P(i,j) = \frac{1}{2}\left(\frac{R_{\mathrm{m}}}{R(i,j)}\right)^{1/\alpha} \tag{2}$$

where R_{m} is a maximum radial influence; α, a flaring coefficient; and $R(i,j)$, the distance, in number of cells, between the cell located in (i,j) and the nearest DRLN cell.

The inverse power-law was chosen for the following reasons: (1) it takes larger values for cells close to the DRLN and it tends toward zero for cells far from the DRLN, (2) it increases or decreases asymptotically, and (3) it can be fully expressed in terms of the maximum radial influence (see explanation below for the flaring coefficient). Note that any type of function having the same properties can be used instead of this inverse power-law.

In Eq. (2), the value of the maximum radial influence must be large enough so the overall modification in elevations eliminates all discrepancies in the flow direction field shown in Fig. 5. On the other hand, if this value is very large then the perturbation negates the information originally provided by the DEM. That is, the DEM no longer affects the determination of the flow directions. As a result, the drainage structure can become quite artificial since the flow directions may be more affected by the remodelled terrain than by the local variation in elevations. To balance these two constraints, the maximum radial influence must be exactly equal to the maximum gap between the

DRLN and the modelled network provided by the D8 approach.

Although the new generation of GIS can handle floating point values, computations are done using values rounded to the nearest integer. This means that when the distance between a given cell and the nearest DRLN cell is equal to the maximum radial influence, the inverse power-law function gives a perturbation coefficient equal to 0.5 which is then rounded to 1. The use of integers should be seen as an approach that is for computer codes which are limited by an integer computation. Note that the proposed approach can be easily adapted to computer codes designed for handling floating point values.

The flaring coefficient has to be adjusted to allow for differentiation in the perturbation of neighbouring cells. This is best illustrated with the following example. Using a maximum radial influence equal to 5 ($R_m = 5$) implies that the inverse power-law function gives a perturbation coefficient equal to 0.5, and rounded to 1, when $R = R_m$. ($E(R = R_m) = 0.51$). Cells with $R = 4$ must have a perturbation coefficient at least equal to 1.5, if, when rounded, this value has to be at least equal to 2. Knowing that the slopes of the inverse power-law function increase when R increases, we can adjust the flaring coefficient by forcing E to be equal to 1.5 when $R = R_m - 1$. This approach gives the following expression for the flaring equation:

$$\alpha = \frac{\ln(R_m) - \ln(R_m - 1)}{\ln(3)} \tag{3}$$

Note that the inverse power-law equation takes an infinite value when $R = 0$. This implies that cells located on the DRLN take an infinite negative elevation value in the modified DEM. Hence, the flow direction of DRLN cells cannot be evaluated using the modified DEM. They are, in fact, determined with the algorithm introduced in Section 4. It also implies that cells which are in the vicinity of the DRLN flow directly into the network.

Fig. 4D–G shows, in detail, different steps of the algorithm. It is noteworthy that the flow directions determined with the proposed approach (Fig. 4H) are the superposition of directions of Fig. 4C and G.

4.3. Modelled river and lake network determination

When each cell has been assigned a flow direction, cells included in the modelled watershed can be directly identified using the D8 approach. However, a different approach is adopted to determine the modelled river and lake network. For this purpose, it is necessary to develop a new algorithm which will use both the flow direction matrix and the DRLN. For this step, the latter is needed to locate lakes.

In the first step, which is comparable to the D8 approach, the proposed algorithm includes cells that drain an upstream area greater than a given threshold area in the modelled network. This first step uses only the flow directions as input data. Note that at this step, contrary to the standard D8 approach, the set of cells that exceed the given threshold may not create a continuous network. This can be explained by the fact that lakes can be wider than one cell and that the upstream area of a given lake can be spread over more than one cell. Therefore, if the upstream areas of some cells of a given lake are smaller than the threshold area, this does not imply that the whole lake does not drain an area greater than the threshold area.

Accordingly, the second step accounts for the DRLN. Therefore, cells located on areas where a lake exists are considered potentially part of the modelled network. They are actually included in the modelled network if the outlet cell drains an area greater than the threshold area. To illustrate that, let us just imagine a 6-cell-threshold for the network presented on Fig. 4H.

It is interesting to analyse the previous approach with respect to the drainage density of the DRLN. If, for example, an extremely dense DRLN is used as input for the flow structure determination, this network will have a major impact on the flow directions. These flow directions, along with the DRLN and a threshold area, which can be variable (Tarboton et al., 1991, 1992), will be used together to determine the modelled network. If the drainage density of the DRLN is too large for the hydrological modelling requirement, then the chosen threshold area will be defined so that the modelled network will be less dense than the DRLN. Inversely, if a perfect fit between the modelled network and the main watercourse of a watershed is desired while an

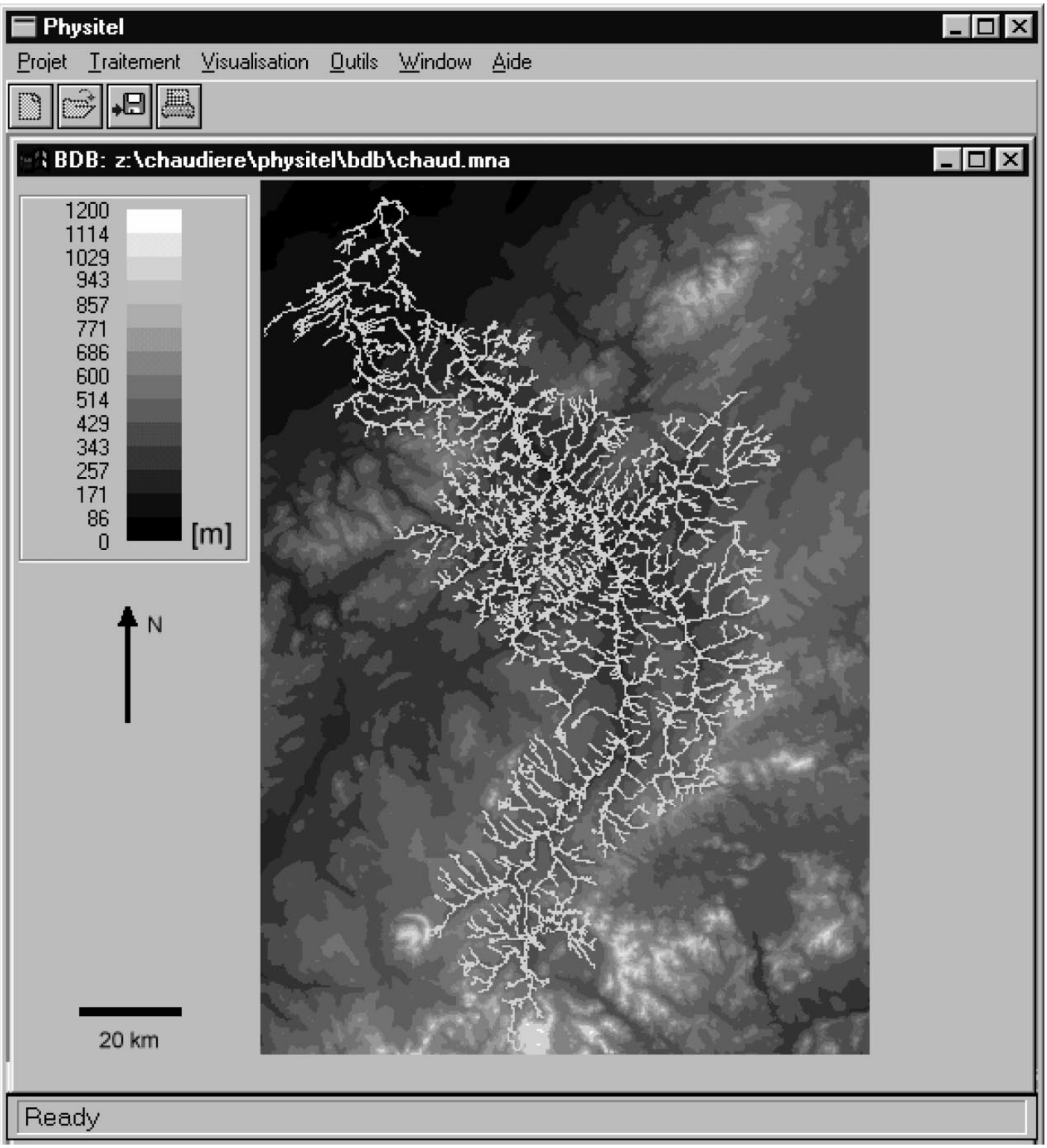


Fig. 6. DEM and DRLN for the Chaudière River watershed, Québec, Canada.

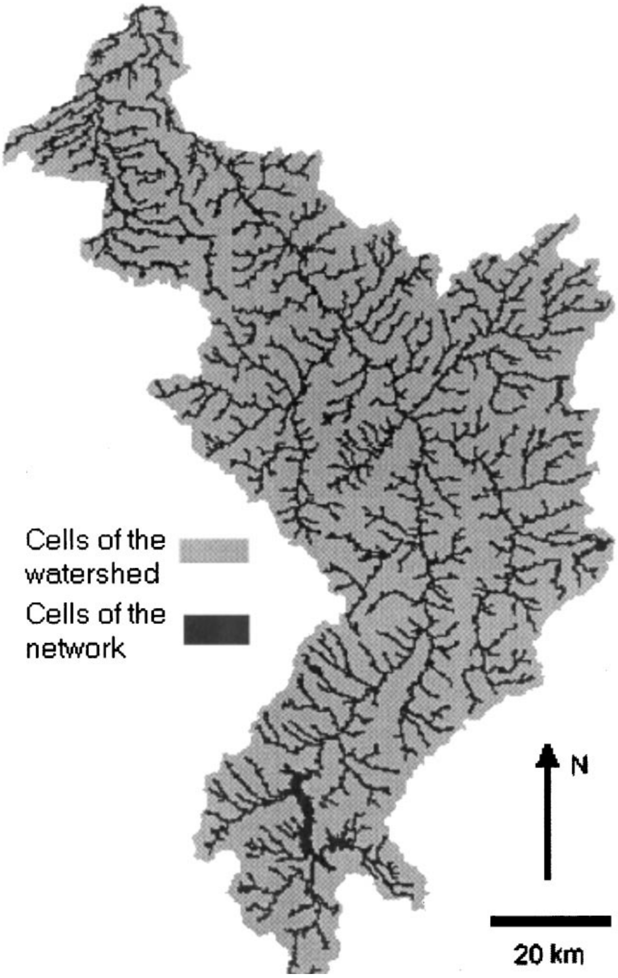


Fig. 7. Chaudière River watershed and modelled network.

approximation of the rest of the network is sufficient, then a less dense DRLN could be taken. The use of an appropriate approach for the threshold area will lead to a more dense modelled network than the DRLN. In fact, the proposed approach allows for a modelled network with a drainage density which does not depend on the drainage density of the DRLN. Inversely, it also allows the use of the drainage density of the DRLN without using any threshold area.

5. Application

This section presents a sample application of the proposed approach on the Chaudière River watershed. This 6680-km^2 watershed is located south of Quebec City, Canada, near the Maine, USA, border. The DEM used for this application was extracted from 1:250 000 contour line maps. This DEM, based on a 100 m square-grid mesh, is shown in Fig. 6. This figure also displays the DRLN obtained from a 1:50 000 digitised map.

Fig. 7 introduces the modelled boundaries of the Chaudière River watershed. The modelled area of the watershed is 6746 km^2. The difference between this value and the official Quebec Government value of 6680 km^2 (M.R.N., 1969) is less than 1%. Note that the official area was evaluated manually using

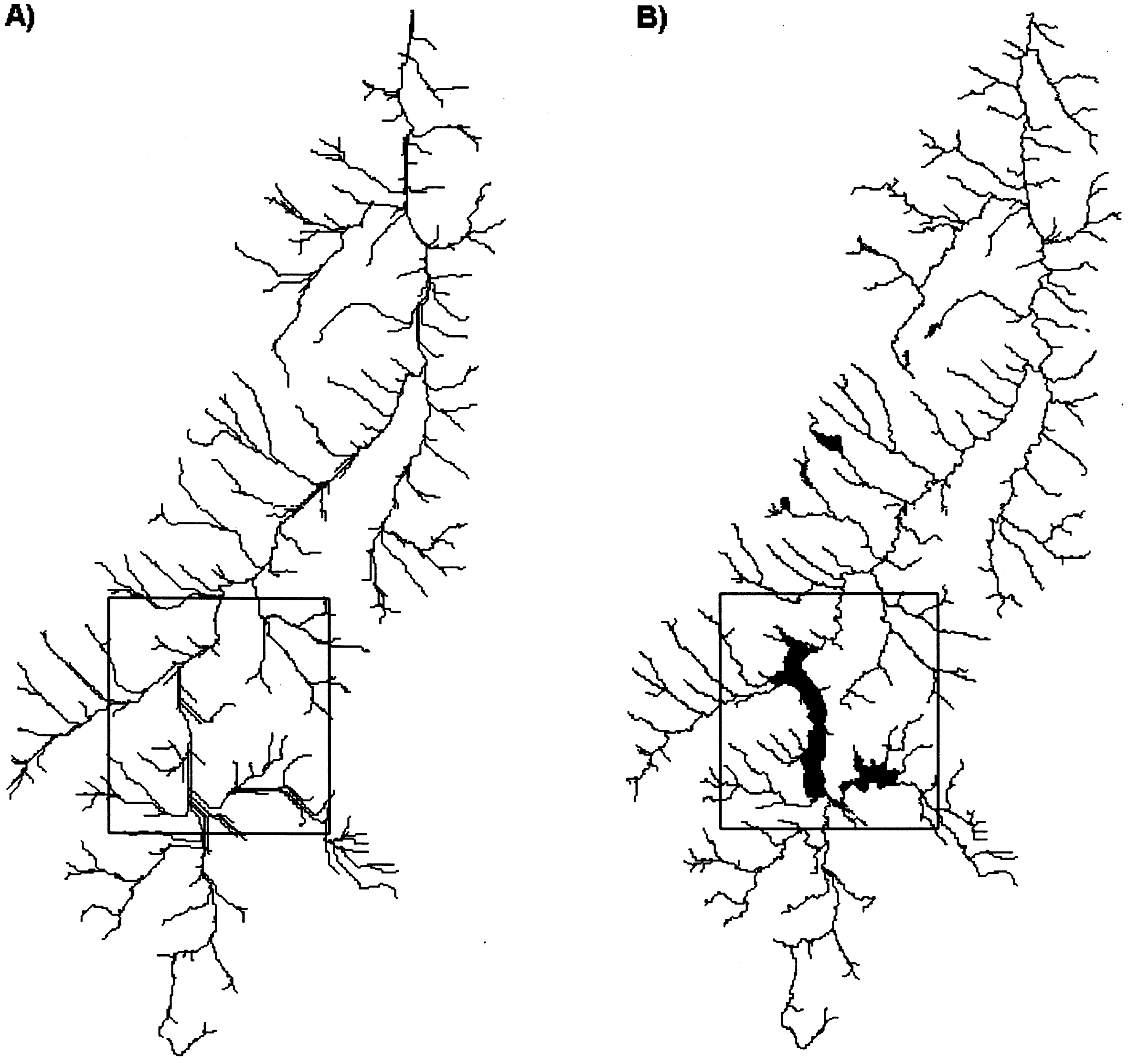


Fig. 8. Modelled network of a problematic area using: (A) the D8 approach, and (B) the proposed approach (Upper-Chaudière River sub-watershed).

1:50 000 topographic maps and, thus, is likely to be inaccurate considering the typical errors associated with this approach. Fig. 7 also presents a macroscopic view of the modelled network. Since the maximum distance gap between the DRLN and the watercourse network modelled with the D8 approach was four cells (400 m), a maximum radial influence (R_m) equal to 4 was used to obtain this result. A 200-cell threshold (2 km^2) was chosen for the upstream points of the network. In fact, the only major differences between the DRLN (Fig. 6) and the modelled network came from the threshold value which led to a less ramified modelled network. This is especially true in the centre part of the watershed where the drainage density of the DRLN is much greater than that for other parts of the watershed. Again, the use of a variable threshold approach is possible with the proposed algorithm. It is also possible to use no

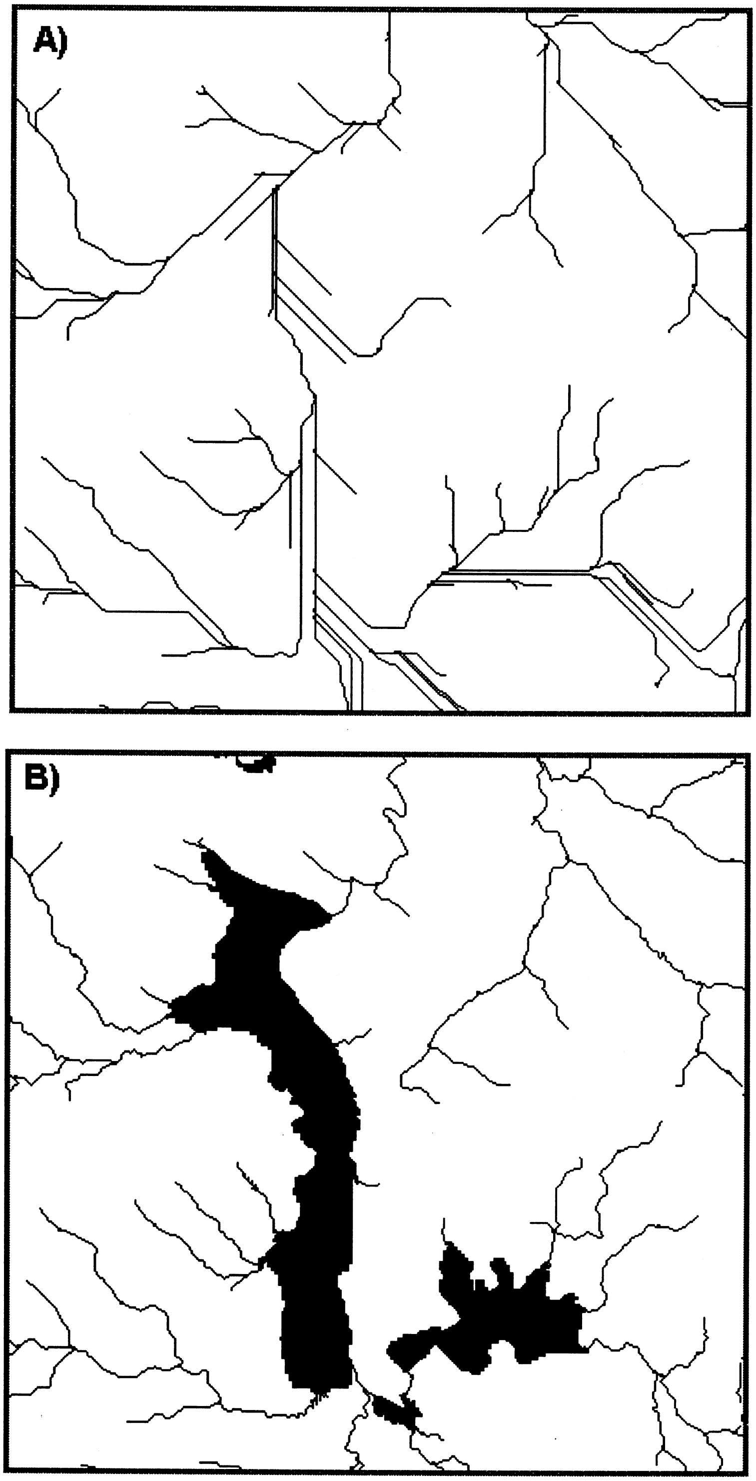


Fig. 9. Modelled network of a problematic area using: (A) the D8 approach, and (B) the proposed approach (Lake Mégantic region).

threshold area at all and then use directly the drainage density of the DRLN.

Fig. 8 shows a comparison between modelled networks generated with the D8 approach and those obtained with the proposed approach in the southern part of the Chaudière River watershed. Flow paths from the two approaches are close, but they indicate that the network modelled with the D8 approach does not always match the DRLN. Fig. 8 reveals that the D8 approach produced a lot of parallel river segments.

The most notable difference between the two modelled networks is associated with the lake areas. Fig. 9 introduces a view of the Lake Mégantic region which clearly illustrates the inadequacy of the D8 approach to delineate lakes. On the other hand, this figure shows that the proposed approach can be used easily to identify lake contours.

Fig. 10A focuses on an area where elevations are more or less constant and where the D8 approach generated parallel flow paths (Fig. 10B). On the other hand, Fig. 10C reveals that the proposed approach completely solved this problem.

Fig. 11 presents flow directions determined using the two approaches. It appears that it is impossible to determine flow paths in flat areas using only a DEM. Again, the proposed approach overcame this problem

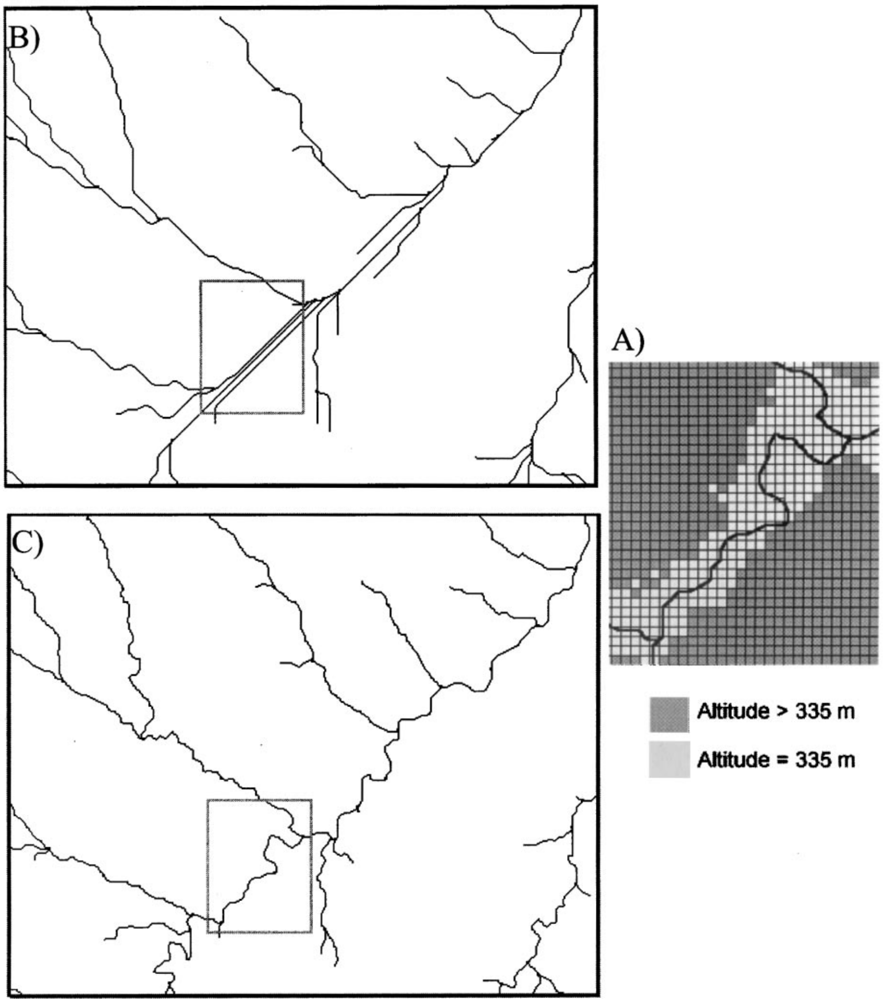


Fig. 10. Modelled network of a wide flat area (see A) using: (B) the D8 approach and (C) the proposed approach.

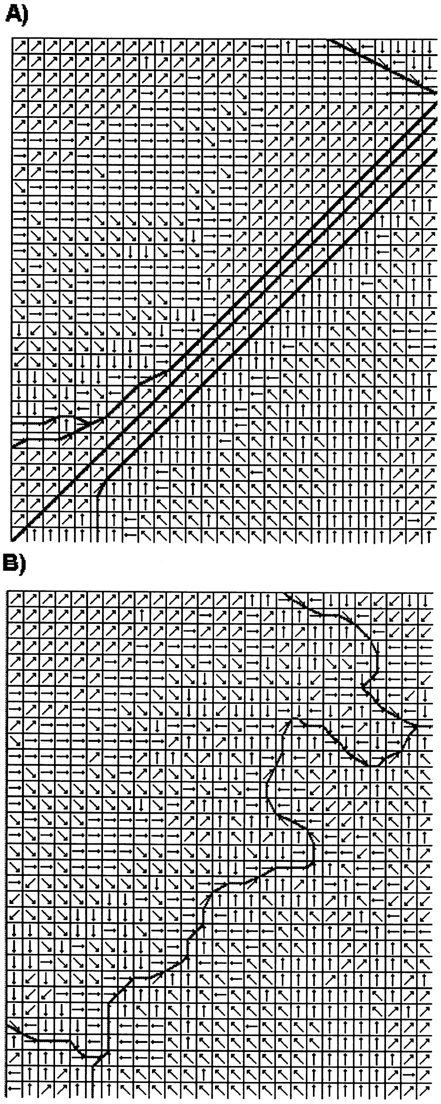


Fig. 11. Modelled flow directions of a flat area using: (A) the D8 approach, and (B) the proposed approach.

and indeed, was required to obtain a well-modelled river and lake network.

6. Conclusions

This paper introduced an original approach for the automatic determination of the drainage structure of a watershed. The proposed approach used both a DEM and a DRLN as input. This allowed for the definition of a drainage structure which was in agreement with the DRLN. It also led to a better match between observed and modelled flow structure. This paper showed numerous limitations of the widely used D8 approach and highlighted that these limitations could be overcome by the proposed approach.

From an algorithmic point of view, the proposed approach is largely based on the standard algorithms of the D8 approach. However, the main differences between the two approaches reside in the determination of flow directions and in the derivation of the modelled river network. For cells which are part of the DRLN, flow directions are determined using network connections only. Flow directions of remote watershed cells were evaluated using the D8 approach with a DEM modified according to the distance away from the DRLN. Moreover, the resulting modelled river network clearly depicts the locations of lakes.

The proposed approach was applied to the Chaudière River watershed, Québec, Canada. Modelled data of the drainage structure showed a high level of coherence. A comparison of results obtained with the D8 approach and those obtained with the proposed approach clearly demonstrated the superiority of the proposed approach over the former.

Future work will involve a study of the approach on several watersheds. Finally, the use of remote sensing as a source of data for the DRLN will be tested.

References

Beasely, D.B., Huggins, L.F., Monke, E.J., 1980. ANSWERS: a model for watershed planning. Trans. ASAE 23 (4), 938–944.

Beven, K., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 24, 43–69.

Cabral, M., Bras, R.L., Tarboton, D., Entekhabi, D., 1991. A distributed, physically-based rainfall–runoff model incorporating topography for real-time flood forecasting. Hydrology and Water Resources Systems, Report Number 332, Ralph M. Parsons Laboratory, M.I.T. Department of Civil Engineering.

Carrara, A., 1986. Drainage and divide networks derived from high-fidelity digital terrain models. NATO Advance Study Symposium.

Coffman, D.M., Keller, E.A., Melhorn, W.N., 1972. New topologic relationship as an indicator of drainage network evolution. Water Resour. Res. 8, 1497–1505.

Costa-Cabral, M.C., Burges, S.J., 1994. Digital elevation model networks (DEMON): a model of flow over hillslopes for computation of contributing and dispersal areas. Water Resour. Res. 30 (6), 1681–1692.

Fairfield, J., Leymarie, P., 1991. Drainage networks from digital elevation models. Water Resour. Res. 27 (5), 709–717.

Fortin, J.P., Moussa, R., Bocquillon, C., Villeneuve, J.P., 1995. HYDROTEL, un modèle hydrologique distribué pouvant bénéficier des données fournies par la télédétection et les systèmes d'information géographique. Rev. Sci. Eau 8, 97–124.

Fortin, J.P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J., 2001. A distributed watershed model compatible with remote sensing and GIS data, part 1: description of the model. J. Hydrol. Engng, ASCE 6 (2).

Gandolfi, C., Bischetti, G.B., 1997. Influence of the drainage network identification method on geomorphological properties and hydrological response. Hydrol. Process. 11, 353–375.

Garbrecht, J., Martz, L.W., 1997. The assignment of drainage direction over flat surfaces in raster digital elevation models. J. Hydrol. 193, 204–213.

Hutchinson, M.F., 1989. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. J. Hydrol. 106, 211–232.

Jenson, S.K., Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographical information system analysis. Photogrametric Engng Remote Sensing 54, 1593–1600.

Joao, E.M., Walsh, J.S., 1992. GIS implications for hydrologic modeling: simulation of nonpoint pollution generated as a consequence of watershed development scenarios. Comput. Environ. Urban Systems 16, 43–63.

Kite, G.W., 1995. The SLURP model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology, Water Resources Publications, pp. 521–562.

Leavesley, G.H., Stannard, L.G., 1995. The precipitation–runoff modeling system — PRMS. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology, Water Resources Publications, pp. 281–310.

Martz, L.W., De Jong, E., 1988. CATCH: a Fortran program for measuring catchment area from digital elevation models. Comput. Geosci. 14 (5), 627–640.

Martz, L.W., Garbrecht, J., 1992. Numerical definition of drainage network and subcatchment areas from digital elevation models. Comput. Geosci. 18 (6), 747–761.

Martz, L.W., Garbrecht, J., 1993. Automated extraction of drainage network and watershed data from digital elevation models. Water Resour. Bull. 29 (6), 901–908.

Martz, L.W., Garbrecht, J., 1995. Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method — comment. J. Hydrol. 167, 393–396.

Martz, L.W., Garbrecht, J., 1998. The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models. Hydrol. Process. 12, 843–855.

Morris, D.G., Heerdegen, R.G., 1988. Automatically derived catchment boundaries and channel networks and their hydrological applications. Geomorphology 1 (2), 131–141.

M.R.N., 1969. Superficie des bassins versants du Québec. Ministère des Richesses Naturelles, Direction générale des eaux, Service de l'hydrographie, Rapport H-1, 60 pp.

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

Olivera, F., Maidment, D., 1999. Geographic information systems (GIS)-based spatially distributed model for runoff routing. Water Resour. Res. 35 (4), 1155–1164.

Quinn, P., Beven, K., Chevalier, P., Planchon, O., 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrol. Process. 5, 59–79.

Rousseau, A.N., Mailhot, A., Turcotte, R., Duchemin, M., Blanchette, C., Roux, M., Etong, N., Dupont, J., Villeneuve, J.-P., 2000. GIBSI — an integrated modelling prototype for river basin management. Hydrobiologia 422/423, 465–475.

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

Saunders, W., Maidment, D.R., 1995. Grid-based watershed and stream network delineation for San Antonio-Nueces coastal basin. Proceedings of Texas Water '95: A Component Conference of the First International Conference of Water Resources Engineering, August 16–17, 1995, ASCE, San Antonio, Texas.

Tarboton, D.G., 1997. A new method for the determination of flows 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. Process. 5, 81–100.

Tarboton, D.G., Bras, R.L., Rodriguez-Iturbe, I., 1992. A physical basis for drainage density. Geomorphology 5 (1/2), 59–76.

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

Villeneuve, J.-P., Blanchette, C., Duchemin, M., Gagnon, J.-F., Mailhot, A., Rousseau, A.N., Roux, M., Tremblay, J.-F., Turcotte, R., 1998. Rapport final du projet GIBSI, mars 1998, Rapport No. R-462. INRS-Eau, Sainte-Foy, Québec.

Wolock, D.M., McCabe Jr., G.J., 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. Water Resour. Res. 31 (5), 1315–1324.