A comparison of geographical information systems–based algorithms for computing the TOPMODEL topographic index

Feifei Pan
Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

Christa D. Peters-Lidard
Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Michael J. Sale and Anthony W. King
Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

Received 2 February 2004; revised 1 March 2004; accepted 20 April 2004; published 24 June 2004.

[1] The performance of six geographical information systems (GIS)-based topographic index algorithms is evaluated by computing root-mean-square errors of the computed and the theoretical topographic indices of three idealized hillslopes: planar, convergent, and divergent. In addition to these three idealized cases, two divergent hillslopes with varying slopes, i.e., concave (slopes decrease from top to bottom) and convex (slopes increase from top to bottom) are also tested. The six GIS-based topographic index algorithms are combinations of flow direction and slope algorithms: i.e., single flow direction (SFD), biflow direction (BFD), and multiple flow direction (MFD) plus methods that determine slope values in flat areas, e.g., W-M method [Wolock and McCabe, 1995] and tracking flow direction (TFD) method. Two combinations of horizontal resolution and vertical resolution of the idealized terrain data are used to evaluate those methods. Among those algorithms the MFD algorithm is the most accurate followed by the BFD algorithm and the SFD algorithm. As the vertical resolution increases, the errors in the computed topographic index for all algorithms decrease. We found that the orientation of the contour lines of planar hillslopes significantly influences the SFD’s computed topographic index. If the contour lines are not parallel to one of eight possible flow directions, the errors in the SFD’s computed topographic index are significant. If mean slope is small, TFD becomes more accurate because slope values in flat areas are better estimated.

INDEX TERMS: 1899 Hydrology: General or miscellaneous; 1824 Hydrology: Geomorphology (1625); 1832 Hydrology: Groundwater transport; KEYWORDS: GIS, TOPMODEL, topographic index, single flow direction algorithm, biflow direction algorithm, multiple flow direction algorithm


1. Introduction

[2] TOPMODEL is a topography-based concept for watershed hydrology modeling. Since the TOPMODEL was first proposed in 1979 [Beven and Kirkby, 1979], it has been widely used to study the effects of topography on hydraulic processes including flood frequency, streamflow generation, flow paths, geomorphic characteristics, and water quality [Wolock and McCabe, 1995]. In addition to the success of the TOPMODEL concept in traditional hydrologic modeling, this concept has also been successfully incorporated into several ecosystem-atmosphere models including the Regional Hydro-Ecological Simulation System (RHESSys) [e.g., Ford et al., 1994]; the Land Ecosystem-Atmosphere Feedback model (LEAF-2) [Walko et al., 2000]; the TOPMODEL-based Land Atmosphere Transfer Scheme (TOPLATS) [Famiglietti and Wood, 1994; Peters-Lidard et al., 1997]; the Catchment Model [Koster et al., 2000]; and the Common Land Model (CLM) [Dai et al., 2003].

[3] To apply the TOPMODEL, a modeled catchment is partitioned with a regular grid or lattice. The so-called "topographic index" is then calculated for each cell in the catchment. The topographic index, ln(a/tan b), is the natural logarithm of the ratio of the specific flow accumulation area a to the ground surface slope tan b. The surface slope can be evaluated from digital elevation model (DEM) data. The specific flow accumulation area is the total flow accumulation area (or upslope area) A through a unit contour length L. To compute the total flow accumulation area A, flow directions are tracked upslope, starting from the cell of interest to the upstream divide of the watershed, and then tracked downslope accumulating cells contributing to the drainage area of the cell of interest. Here we note an uncertainty associated with the definition of the flow accumulation area A. From the presentations of Beven and

Copyright 2004 by the American Geophysical Union.
0043-1397/04/2004WR003069S09.00
Wood [1983, Figure 2] and Kirkby [1997, Figure 1], one can reasonably conclude that the flow accumulation area is defined along the ground surface. However, using DEMs in geographical information systems (GIS), the computed flow accumulation area is generally the area projected to x-y plane, and this calculation of A has become standard practice. The difference between these two areas is negligible if the slope is less than 0.5 (m/m), and most of the slopes in the watersheds to which the TOPMODEL is applied are less than 0.5 (m/m) [e.g., Montgomery and Dietrich, 1992]. Therefore, for consistency with standard practice, we adopt the convention of calculating the flow accumulation area or upslope area as the area projected to x-y plane.

[4] The slope term (\(\tan(\beta)\)) in the topographic index arises from the assumption that the surface of the water table is parallel to the ground surface. Thus the local hydraulic gradient is assumed to be equal to the slope of the ground. Because flow direction depends on hydraulic gradient, or ground surface slope, flow direction and the calculation of the upslope accumulation area should be consistent with the local slope value that is used to compute \(\ln(\alpha / \tan(\beta))\). Thus the computed topographic index is dependent upon the calculation of both slope and flow direction.

[5] Although many researchers have investigated algorithms for calculation of slope and flow direction, those studies tend to focus on either slope [e.g., Jones, 1998; Zhang et al., 1999] or flow accumulation area [e.g., Tarboton, 1997; Rieger, 1998]. Few studies have examined the combined effects of slope and flow direction algorithms on the topographic index. The works of Quinn et al. [1991], Wolock and McCabe [1995], and Mendicino and Sole [1997] are exceptions. However, these studies only investigated the difference in the statistical moments or distributions of the computed topographic index; no comparisons of errors between the “true,” i.e., analytically solved, and the numerically computed topographic indices were carried out. Comparing against “truth” is difficult because an analytical expression for the real terrain does not normally exist. Therefore one cannot usually determine which topographic index algorithm is more accurate. This paper seeks to resolve that issue and provide objective evaluation of the appropriate numerical algorithm for calculating the topographic index.

2. Topographic Index Algorithms

[6] A topographic index algorithm is actually a combination of two algorithms, one to calculate flow direction and another to calculate slope. Here we describe three commonly used flow direction algorithms that allow flow in one, two, and more than two directions. We identify each topographic index algorithm by the name of the associated flow direction algorithm (e.g., single flow direction, biflow direction, and multiple flow direction), because each flow direction algorithm uses its own algorithm to determine slope.

2.1. Single Flow Direction (SFD) Algorithm

[7] The single flow direction (SFD) algorithm [O’Callaghan and Mark, 1984] calculates flow direction as the steepest slope direction, which is determined by the Maximum Downward Gradient (MDG). This SFD algorithm, also known as the D8 algorithm, is widely used in DEM data analysis [e.g., O’Callaghan and Mark, 1984; Band, 1986; Greenlee, 1987; Mark, 1988; Jenson and Domingue, 1988] and GIS software (e.g., the “FLOWDIRECTION” function in ARC/INFO GRID).

[8] MDG computes the downhill elevation gradients of a 3 x 3 cell window along eight directions (i.e., elevation of the center cell minus elevation of each of its eight neighbors divided by the distance between those two cells). The slope of the central cell is calculated as the largest of the eight directions. Steepest slope direction is the direction from the central cell to the neighbor generating the largest downhill elevation gradient.

[9] If the central cell has a lower elevation than one of its neighbors, the downhill gradient along this direction is negative. If the calculated slope (i.e., the largest one among eight directions) is less than zero, this cell is called a sink or a pit. At a sink, there is only inflow, and no outflow. However, it is often computationally required to force watersheds to have outlets. Therefore sinks in DEM data are usually filled before watersheds can be delineated and other hydrologic parameters estimated. One common filling approach for sinks generated by MDG is to raise the elevation of a sink to the lowest elevation among its neighbors [e.g., Jenson and Domingue, 1988; Tarboton et al., 1991]. This algorithm is used in the ARC/INFO GRID (i.e., the “FILL” function). For example, the recalculated slope value of the “filled” cell is now zero; the cell and its neighbors form a flat area. To determine flow direction in a flat area, the method suggested by Jenson and Domingue [1988] (e.g., the ARC/INFO FLOWDIRECTION function) is used.

[10] Once the single flow direction is determined for each cell, flow accumulation area (or upslope area) (A) is calculated using a recursive procedure [e.g., Tarboton, 1997]. The specific flow accumulation area (a) is (A) divided by a contour length, which is equal to the grid size or horizontal resolution of the DEM. The slope is set to be the maximum downward elevation gradient.

2.2. Biflow Direction (BFD) Algorithm

[11] The biflow direction algorithm (BFD), also known as D8, was proposed by Tarboton [1997]. In this algorithm the 3 x 3 cell window is divided into 8 triangular facets. The slope direction and magnitude of each facet are compared. The steepest downward direction is chosen and divided into two directions along the edges forming that facet. The proportion of flow along each edge is inversely proportional to the angle between the steepest downward direction and the edge. Therefore at most two flow directions can be assigned to each cell. The contour length is defined as the grid cell size (DEM resolution), and the slope is set to be the largest slope of 8 facets [Tarboton, 1997].

[12] Like the SFD algorithm, the BFD algorithm has trouble defining flow direction in flat areas. A GIS-based algorithm for returning flow directions [Jenson and Domingue, 1988], such as the ARC/INFO “FLOWDIRECTION” function, can be used to determine the flow direction in flat areas, such that only one flow direction is assigned to each cell in a flat area.

2.3. Multiple Flow Direction (MFD) Algorithm

[13] Quinn et al. [1991] first suggested the multiple flow direction (MFD) algorithm to improve representation of the convergence or divergence of flow. Wolock and McCabe...
respectively, where HR is the horizontal resolution, 0
McCabe
flow directions [the MFD algorithm, a GIS-based algorithm for returning
path between X and Y, which is \( \frac{2}{(2^{\sqrt{HR}} + HR)} \), where
HR
is the cell size.

[1995] showed how to implement this algorithm using
ARC/INFO GRID functions. Unlike the SFD and BFD
algorithms, the MFD algorithm allows flow in more than
one or two directions. In the MFD algorithm, each flow
direction is weighted by the downward elevation gradient
(i.e., from the central cell to each of its 8 neighbors)
multiplied by a “contour length”. There are two ways to
set the contour length: i.e., \( \frac{HR}{2} \) and \( \sqrt{2HR} \) [Quinn
et al., 1991], or 0.6HR and 0.4HR [Wolock and
McCabe, 1995], for cardinal and diagonal flow directions,
respectively, where HR is the horizontal resolution, 0° grid
cell size of the DEM. Here, we assume the HR in easting
and northing directions are the same (which is true for most
DEM data). To differentiate those two algorithms for
assigning the contour lengths, we call Quinn’s method the
MFD algorithm, and Wolock and McCabe’s algorithm the
MFD* algorithm. The slope is the weighted summation of
all positive downward gradients [Quinn et al., 1991]. As for
the MFD algorithm, a GIS-based algorithm for returning
flow directions [Jensen and Domingue, 1988], such as
ARC/INFO flow direction function can used to determine
the flow directions in flat areas.

2.4. Recalculation of Slopes in Flat Areas

[14] Because the topographic index cannot be calculated
directly in flat areas, estimating slopes in flat areas is a
critical issue. Wolock and McCabe [1995] defined the
minimum slope value to be equal to \((0.5 \times \text{vertical resolution})/(\text{horizontal resolution})\). We refer to this as the
W-M method.

[15] To maintain consistency between flow direction and
slope in flat areas, we propose a simple algorithm called the
tracking flow direction (TFD) method. This algorithm starts
from cell X and following the flow direction determined by
ARC/INFO “FLOWDIRECTION” function that searches
for the nearest cell Y with a lower elevation than cell X. The
slope at cell X is then set to be the ratio of elevation
difference to length of flow path between cells X and Y
(see Figure 1). If a nonzero slope value cannot be found for
the cell over the whole domain, the value of −1 is assigned
to this cell, indicating that slope is undefined. The minimum
of all defined slopes is then assigned to each undefined cell.

3. Idealized Hillslopes

[16] To compare the computed and theoretical “true”
topographic indices, we apply six GIS-based topographic
index algorithms to three idealized hillslopes, i.e., planar,
divergent and convergent, for which we can calculate the
theoretical topographic index analytically. Although those
hillslopes are ideal simplifications, they represent three
types of real terrain: terrace, divergent and convergent.

[17] For the planar and convergent hillslopes, we use one
slope value for each case, i.e., slope is spatially uniform. For
the divergent hillslopes, in addition to the constant slope
cases, we also test two divergent hillslopes with varying
slopes, i.e., concave (slope decreases from top to bottom),
and convex (slope increases from top to bottom). Although
slopes are not spatially uniform for those concave and
convex divergent hillslopes, to simplify the problem, we
set slope constant along each contour line. Under such
conditions, we can theoretically compute the specific flow
accumulation area of each cell along a contour line as the
total upslope area divided by the length of the contour line.
This calculation is valid only when the contour line is
closed for the divergent or convergent hillslopes, or the
contour line is infinitely long for the planar hillslopes.

[18] Prior to introducing each idealized case, we first
define three domains. The computation domain is defined
as the region where we use a recursive algorithm to compute
the flow accumulation area based on flow directions deter-
mined by the SFD, BFD or MFD algorithms. We can
think of the computational domain as a part of a watershed.
However, this part is not just an arbitrary domain. It starts
from a divide or a local peak and covers a part of the
downstream area, since we will set the top boundary parallel
to the contour lines. Therefore, when we compute the flow
accumulation area, we do not miss any contribution from
upslope crossing the top boundary.

[19] The influence domain is defined as the area inside
the computational domain that receives the flow contribu-
tion from the lateral boundaries. The comparison domain is
inside the computational domain but with the influence
domain excluded. Our theoretical calculation of the topo-
graphic index is valid only if there is no influence from the
lateral boundaries. Therefore, to compare the computed and
theoretical topographic indices properly, we need to
carry out such comparison inside the comparison domain,
because only inside the comparison domain are the com-
puted flow accumulation area and thus the computed
topographic index not influenced by the lateral boundaries.

3.1. Planar Hillslopes

[20] A three-dimensional, idealized planar hillslope is
illustrated in Figure 2. In a NROW × NCOL domain
(NROW and NCOL are numbers of rows and columns),
the coordinates of each cell are given by:

\[
\begin{align*}
x &= \left( j - 0.5 \right) \cdot NCOL - 1 \\
y &= 0.5 \cdot NROW - 1 - i \cdot HR_i, \quad 0 = i, \ldots, NCOL - 1 \quad (1)
\end{align*}
\]

where \((i, j)\) are row and column indices of each cell, HR is
grid cell size or horizontal resolution of DEM data. In this
study, we set NROW = 2001, and NCOL = 4001.
The distance from \((x_0, y_0)\) to the contour line passing cell \((x, y)\) given by
\[ D = \frac{(x-x_0) \tan \alpha + (y-y_0)}{\sqrt{(\tan \alpha)^2 + 1}} \]
where \(\alpha\) is the angle between each contour line and x axis (from the contour line to x axis in the counterclockwise direction) and we call it contour angle (Figure 2).

The angles between these flow directions and x axis are 0 \(^\circ\), 45 \(^\circ\), 90 \(^\circ\), 135 \(^\circ\), 180 \(^\circ\), 225 \(^\circ\), 270 \(^\circ\), and 315 \(^\circ\).

The angles formed by LB and AI (i.e., \(\alpha\)), and the comparison domain, because it is in both SFD’s and BFD’s influence domain, we need to follow flow directions flowing into four boundaries: LB, RB, and BB.

The angle \(\beta\) is derived from the contour line to x axis in the counterclockwise direction and is called contour angle (Figure 2). 

\[
\beta = \tan^{-1}\left(\frac{\text{increment of elevation in y direction}}{\text{increment of elevation in x direction}}\right)
\]

The computational domain for the planar hillslopes is confined by four boundaries (Figure 2), i.e., LB, RB, TB, and BB. TB is called the top boundary and BB is the bottom boundary. They are parallel to the contour lines and associated with the highest and lowest elevations in the computational domain, respectively. LB and RB are two lateral boundaries. Facing the direction of downward gradient of elevation (shown as a block arrow in Figure 2), LB is on the left side, and RB is on the right side. We set the computational domain for the planar hillslopes as follows:

\[
\begin{align*}
\text{Condition} & \quad \text{Distance} \\
\text{LB} & \quad d_1 \\ 
\text{RB} & \quad d_2 \\ 
\text{TB} & \quad d_3 \\ 
\text{BB} & \quad d_4
\end{align*}
\]

where \(d_1, d_2, d_3, d_4\) are distances from origin \((0, 0)\) to four boundaries: LB, RB, BB, and BB.

The computational domain, the theoretical topographic index is given by:

\[
\ln\left(\frac{a}{\tan \beta}\right) = \ln \left(\frac{\text{upslope area / contour length}}{\text{slope}}\right) = \ln \left(\frac{D*L}{\text{upslope area}/\text{contour length}}\right) = \ln \left(\frac{D}{\tan \beta}\right)
\]

where \(D\) is the distance from cell \((x, y)\) to TB, and \(L\) is the contour length. As we mentioned in section 1, the flow accumulation area of the upslope area is the area projected to x-y plane, i.e., DL. Although \(L\) does not appear in the final expression, equation (6) is valid only as \(L\) is infinite, or there is no influence due to the lateral boundaries. However, \(L\) cannot be infinite. Therefore a proper comparison between the computed and theoretical topographic index can only be carried out in a comparison domain where the lateral boundaries have no influence.

Our first step is to remove the influence domain from the computational domain. In order to determine the influence domain, we need to follow flow directions flowing into the computational domain from two corner cells on the top boundary (TB). The MFD algorithm has a larger influence region than the SFD algorithm or the BFD algorithm, because it allows more flow dispersion. Therefore we choose the MFD’s comparison domain as the common comparison domain, because it is in both SFD’s and BFD’s comparison domains.

Figure 3a shows 8 possible flow directions at a cell. The angles between these flow directions and x axis are \(0^\circ\) (direction 1), \(45^\circ\) (direction 2), \(90^\circ\) (direction 3), \(135^\circ\) (direction 4), \(180^\circ\) (direction 5), \(225^\circ\) (direction 6), \(270^\circ\) (direction 7), and \(315^\circ\) (direction 8).

Figures 3b, 3c, and 3d illustrate the computational, influence and comparison domains for three cases corresponding to three different contour angles, i.e., \(0^\circ\), \(45^\circ\) and \(22.5^\circ\). The area between LB and AI is the influence domain due to the lateral boundary LB. The angle formed by LB and AI (i.e., \(aa\)) is the maximum angle among the angles formed by LB and each flow direction at point A which points to the computational domain. Similarly, we can determine the
influence domain due to the lateral boundary RB, which is formed by RB and BI. The angle between RB and BI is \( \beta \).

In Figure 3b, both points A and B have three flow directions, i.e., 6, 7, and 8. However, only direction 8 at point A and direction 6 at B point into the computational domain. Therefore angles \( \alpha_a \) and \( \alpha_b \) are equal to 45°.

Similarly, in Figure 3c, both A and B have three flow directions, i.e., 5, 6, and 7. Among those directions, only direction 7 at point A and direction 5 at point B flow into the computational domain. Therefore angles \( \alpha_a \) and \( \alpha_b \) are equal to 45°.

In Figure 3d, both A and B have four flow directions, i.e., 5, 6, 7 and 8. Among those directions, only directions 7 and 8 at A, and directions 5 and 6 at B flow into the computational domain. Comparing the angles, we can find that angle \( \alpha_a \) is formed by LB and direction 8, and angle \( \alpha_b \) is formed by RB and direction 5. Therefore \( \alpha_a \) and \( \alpha_b \) are equal to 67.5°. Once \( \alpha_a \) and \( \alpha_b \) are determined, we can determine the influence domain and thus the comparison domain.

### 3.2. Divergent Hillslopes

#### 3.2.1. Divergent Hillslopes With Constant Slopes

[27] A top and a side view of a divergent hillslope are illustrated in Figure 4. In a \( NCOL \times NROW \) domain, the coordinates of each cell are the same as the planar hillslopes (i.e., equation (1)). The theoretical elevation \( Z_t \) is given by:

\[
Z_t = f_1*HR - r*\text{slope} = f_1*HR - r*\tan\beta
\]

where \( \tan\beta \) is slope, \( \beta \) is slope angle (\( 0^\circ < \beta < 90^\circ \)), \( r = \sqrt{x^2 + y^2} \) is radius of the contour line passing \((x, y)\), and \( f_1*HR \) is the elevation at the center of the domain. In this study, we set \( NCOL = 4001, NROW = 2001, f_1 = 3000 \), and the computational domain inside the circle (exclude the origin \((0, 0)\)) with a radius equal to \( R \) (Figure 4), which satisfies the following conditions:

\[
0 < r \leq R
\]

where we set \( R = 500HR \) in this paper.

[28] In the computational domain, the theoretical topographic index is given by:

\[
\ln(a/\tan\beta) = \ln\left[\frac{\text{upslope area}}{\text{contour length}}\right] / \text{slope} = \ln\left[\frac{\pi r^2}{2\pi r \tan\beta}\right] = \ln\left[\frac{r}{2 \tan\beta}\right]
\]

The comparison domain is the same as the computational domain, since the contour lines are closed.

#### 3.2.2. Divergent Hillslopes With Varying Slopes

[29] Figure 5 illustrates two divergent hillslopes with varying slopes: slope decreases from to top bottom (concave, Figure 5a) and slope increases from top to bottom (convex, Figure 5b). The theoretical elevation \( Z_t \) is given by:

\[
\text{Concave : } Z_t = c_1(r - c_2)^2
\]

\[
\text{Convex : } Z_t = c_3 - (r/c_4)^2
\]
where \(c_1, c_2, c_3\) and \(c_4\) are coefficients describing the shapes of the hillslopes. In this study, we set \(c_1 = 10^{-5}\), \(c_2 = 2249.49HR\), \(c_3 = 3000HR\), \(c_4 = 500\). We set the computational domain to be the same as the divergent cases with constant slopes. In the computational domain, the theoretical topographic index is given by:

\[
\ln(a + \tan \beta) = \ln \left[ \frac{\text{upslope area}}{\text{contour length}} \right] \frac{\text{slope}}{c_0}
\]

\[
= \ln \left[ \frac{\pi r^2}{2 \pi r \text{slope}} \right] = \ln \left[ \frac{r}{2 \text{slope}} \right] \tag{11}
\]

Slope is computed as:

\[
\text{Concave} \quad \text{slope} = \frac{\partial Z_i}{\partial r} = 2c_1(c_2 - r) \tag{12}
\]

\[
\text{Convex} \quad \text{slope} = \frac{\partial Z_i}{\partial r} = 2r/c_4^2
\]

As with the divergent cases with constant slopes, the comparison domain is the same as the computational domain, since the contour lines are closed.

### 3.3. Convergent Hillslopes

[30] A top view and a side view of a convergent hillslope are illustrated in Figure 6. In a \(N\) \(x\) \(N\) domain, the coordinates of each cell are given by:

\[
x = [j - 0.5(NCOL - 1)]*HR, \quad j = 0, \ldots, NCOL - 1
\]

\[
y = [-i*HR, \quad i = 0, \ldots, NROW - 1] \tag{13}
\]

In this paper, we set \(NCOL = 4001\), and \(NROW = 2001\). Unlike the divergent hillslopes, in the x-y plane we set the origin \((0, 0)\) at the upper boundary, rather than the center, of the domain. Otherwise, the convergent hillslope would be filled and becomes a flat area if we set the origin at the center of the domain, and apply ARC/INFO FILL function. The theoretical elevation of each cell is given by:

\[
Z_i = f_2 + rslope = f_2 + r \tan \beta \tag{14}
\]

where \(\tan \beta\) is slope and \(\beta\) is slope angle \((0^\circ < \beta < 90^\circ)\), and \(f_2\) is the elevation of the origin \((0, 0)\). In this paper, we set

---

**Figure 4.** Top and side views of a divergent hillslope with a constant slope. Both the computational domain and the comparison domain are inside the circle (exclude the origin \((0, 0)\)) with a radius equal to \(R\), and \(\beta\) is slope angle.

**Figure 5.** Side view of two divergent hillslopes with varying slopes: (a) slope decreases from top to bottom (concave) and (b) slope increases from top to bottom (convex).
f_2 = 200, and the computational domain as the region confined by two thick curves and two lateral boundaries shown in Figure 6, which satisfies the following conditions:

\[ R_1 / C_2 < r < R_2, \quad q / C_2 < F(15) \]

where \( r = \sqrt{x^2 + y^2}, \quad \theta = |\arctan(x/y)| \). In this study, we set \( R_1 = 1400HR, \quad R_2 = 1900HR \) and \( \Phi = 45^\circ \). The theoretical topographic index is given by:

\[
\ln(a / \tan \beta) = \ln \left( \frac{\text{upslope area}}{\text{contour length}} \right) / \text{slope} = \ln \left( \frac{\pi R_2^2 - \pi r^2}{2 \pi r \tan \beta} \right) = \ln \left( \frac{R_2^2 - r^2}{2r \tan \beta} \right)
\]

The above expression is valid only if the contour lines are closed. However, the contour lines in the computational domain are open (Figure 6). Therefore the comparison domain is not the same as the computational domain. As we did for the planar hillslopes, we first identify the influence region where the flow accumulation area is influenced by the lateral boundaries based on flow directions given by the MFD algorithm.

The procedure to determine the influence and comparison domains for the convergent hillslope is shown in Figure 7. The computational domain is formed by two lateral boundaries: AD and BC, and two arcs: AEB and CFD. Point A has three flow directions, i.e., 3, 4, and 5. Among those directions, only direction 5 flows into the computational domain. Therefore the influence region due to the lateral boundary AD is the region between AD and AG, where the angle formed by those two lines is 45°. Similarly, the influence domain due to the lateral boundary BC is the region between BC and BG, and the angle formed by those two lines is also 45°. Excluding the influence domain (shaded area) from the computational domain, we obtain the comparison domain (unshaded area inside the computational domain).

4. Evaluation Method

To compute the topographic index of each idealized hillslope we first develop an idealized DEM data set based on equations in section 3. The ARC/INFO FILL and FLOW-DIRECTION functions are applied to determine flow direction at each cell over the whole domain. The ARC/INFO-derived flow directions are used only in flat areas where flow directions cannot be resolved by the SFD, BFD, or MFD algorithms. Those flow directions are also utilized by the TFD algorithm for recalculating slope values in flat areas.

We use a recursive algorithm to compute flow accumulation area inside the computational domain. The slope, contour length, and topographic index at each cell are determined first. Next, the root-mean-square error (RMSE) of the calculated topographic index is computed as follows:

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left( \ln(a / \tan b)_{Ti} - \ln(a / \tan b)_{Ci} \right)^2}{n}},
\]

where subscripts \( Ti \) and \( Ci \) stand for theoretical and computed topographic indices at cell i and \( n \) is the number of cells.
of total cells inside the comparison domain. The relative RMSE is:

\[
\text{Relative RMSE} = \frac{\text{RMSE}}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} \ln\left(\frac{a}{\tan^3(\theta)}\right)}}
\]  \hspace{1cm} (18)

5. Results and Discussions
5.1. Comparison of Relative RMSEs

[34] A total of 324 tests were made: three types of idealized hillslopes (i.e., planar, divergent with constant slopes, and convergent) times two sets of data resolutions (i.e., (HR = 10 units, VR = 1 unit), (HR = 10 units, VR = 0.1 unit)) times six combinations of flow direction and slope algorithms (i.e., SFD+W-M, SFD+TFD, BFD+W-M, BFD+TFD, MFD+W-M, MFD+TFD) times nine mean slope values (i.e., 0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5). The contour angle is 45° for the planar hillslopes. The relative RMSEs of the computed topographic index are shown in Figures 8a, 8b, and 8c, for planar, divergent, and convergent hillslopes, respectively.

[35] No matter whether the hillslope is planar, divergent, or convergent, we observed five common error characteristics for the computed topographic index (as illustrated in Figure 8): (1) The MFD + TFD algorithm yielded the smallest errors for all cases. (2) As the slope increases, the overall errors decrease. (3) As the vertical resolution of the data increases, the overall errors decrease. (4) If the mean slope is less than the threshold, i.e., VR/HR (dotted line on Figure 8), there is no difference between the BFD algorithm and the SFD algorithm. If the mean slope is greater than VR/HR, the BFD algorithm is better than the SFD algorithm for divergent and convergent hillslopes, and a little worse than the SFD algorithm for planar hillslopes with a 45° angle between the contour line and x axis. The effect of the angle between the contour line and x axis on the computed topographic index of planar hillslopes will be discussed in section 5.2. (5) Error in the computed topographic index due to W-M method and TFD method for recalculating slope values in flat areas appears when the mean slope is less than 0.5VR/HR. The errors associated with TFD are less than those associated with W-M, because W-M only assign 0.5VR/HR as the min-
minimum slope value whereas TFD assigns a smaller value than 0.5VR/HR for replacing zero-slope values in flat areas. On the other hand, if mean slope is greater than 0.5VR/HR, TFD can also replace zero slopes with a larger value (see Table 1).

Another set of 40 tests were also carried out to investigate the effects of the definition of the contour length, the variation of slopes and the vertical resolution on the errors of the computed topographic index. Those 40 tests are combinations of two divergent hillslopes (i.e., concave and convex) times four algorithms (i.e., SFD+TFD, BFD+TFD, MFD+TFD, and MFD*+TFD) times five vertical resolutions (i.e., 0.01, 0.1, 1, 5, and 10 units). The results are shown in Figure 9. These results indicate that the MFD+TFD algorithm is the best for all cases. The difference due to the different algorithms to assign the contour lengths (i.e., MFD and MFD*, as discussed in section 2.3) is negligible. The errors decrease and finally become stable for all cases as the vertical resolution becomes finer and finer. However, a fine vertical resolution cannot eliminate the difference among the errors due to different algorithms, i.e., MFD is better than BFD, and BFD is better than SFD.

5.2. Effect of Contour Line Orientation

In section 5.1 we only tested the planar hillslopes with a 45° angle between the contour line and x axis. To investigate the effect of the contour line orientation, two contour angles, i.e., 0° and 22.5° were examined. The vertical and horizontal resolutions are 0.1 unit and 10 units respectively. Figure 10a shows that there is no difference among the SFD, the BFD, and the MFD algorithms for 0°, and the SFD algorithm gives the largest error for 22.5° (Figure 10b). To further understand the effect of contour line orientation, we also tested 22.5° at HR = 10 units, and VR = 0.0001 unit, and the results were plotted in Figure 10c. Figure 10c shows that only when the planar hillslope is relatively steep (e.g., slope >0.1) will using a finer vertical resolution DEM reduce the errors for the SFD algorithm. For a relatively flat planar hillslope, even using a finer resolution DEM, errors remain in the SFD algorithm. However, this is

Table 1. Minimum Slope Values Determined by the Tracking Flow Direction (TFD) and Wolock and McCabe’s [1995] (W-M) Methods

<table>
<thead>
<tr>
<th>Theoretical Slope Value</th>
<th>Planar Divergent Convergent</th>
<th>Planar Divergent Convergent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TFD W-M</td>
<td>TFD W-M</td>
</tr>
<tr>
<td>0.001</td>
<td>0.0010 0.05</td>
<td>0.0009 0.05</td>
</tr>
<tr>
<td>0.005</td>
<td>0.0048 0.05</td>
<td>0.0050 0.05</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0092 0.05</td>
<td>0.0088 0.05</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0414 0.05</td>
<td>0.0414 0.05</td>
</tr>
</tbody>
</table>

Figure 9. Plots of relative RMSE versus vertical resolution for two divergent hillslopes: (a) concave (slope decreases from top to bottom) and (b) convex (slope increases from top to bottom).
not the case for either the BFD algorithm or the MFD algorithm.

Figure 10 indicates that the errors in the SFD’s computed topographic index of a planar hillslope are not only a function of vertical resolution of DEM data, but also the orientation of the contour lines. When the contour angle is an integer multiple of 45° (e.g., 0°, 45°, 90°...), the theoretical flow direction (i.e., normal direction of the contour lines) is parallel to one of 8 possible flow directions in the SFD algorithm. Thus the SFD algorithm can resolve the flow directions without any approximation. However, if the normal direction of the contour lines is not parallel to one of eight possible flow directions (e.g., contour angle = 22.5° and the angle between the normal direction of the contour lines and x axis is 67.5°), the SFD algorithm must approximate the flow direction (e.g., either 45° or 90°), and the difference between theoretical flow direction and the SFD’s flow direction is 22.5°.

6. Conclusions

By comparing the numerically computed and theoretical (analytically calculated) topographic indices for 526 cases, we found that the MFD algorithm is best in terms of accuracy for all idealized hill slope cases. Although the BFD algorithm is a little worse than the SFD algorithm when the angle between the contour line and x axis is 45° for planar hillslopes, overall the BFD algorithm is better than the SFD algorithm. The difference in the errors of the computed topographic index due to the different algorithms for assigning the contour lengths is negligible (i.e., MFD [Quinn et al., 1991] and MFD* [Wolock and McCabe, 1995]). We found that the orientation of the contour lines of the planar hillslopes significantly influences the SFD’s computed topographic index. If the contour lines are not parallel to one of eight possible flow directions, the errors in the SFD’s computed topographic index are significant.

Recalculating nonzero slopes in flat areas increases errors only when the mean slope value is less than 0.5VR/HR. The tracking flow direction (TFD) method is more accurate than the W-M method [Wolock and McCabe, 1995] because TFD can give a minimum slope value less than 0.5VR/HR. Therefore, to achieve the highest accuracy, we recommend using a combination of the MFD algorithm and the TFD method to compute topographic index, especially when the vertical resolution of DEM data is low.
Acknowledgments. The authors wish to thank H. I. Jager for her valuable comments and suggestions. Useful comments were provided by the Associate Editor and two anonymous reviewers. Research was supported by an appointment to the ORNL Postdoctoral Research Associates Program (F. Pan), which is sponsored by Oak Ridge National Laboratory and administered jointly by Oak Ridge National Laboratory and by the Oak Ridge Institute for Science and Education under contract numbers DE-AC05-84OR21400 and DE-AC05-76OR00033, respectively. This research was also supported by the U.S. Department of Energy’s Office of Science and performed at Oak Ridge National Laboratory (ORNL). ORNL is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725. Additional research support was provided by NASA grant NAG5-8698 (Peters-Lidard) and the United States Environmental Protection Agency through Grants R825210 (Peters-Lidard) and R825211 (Peters-Lidard). Although the research described in this article has been funded wholly or in part by the United States Environmental Protection Agency through Grants R825210 and R825211, it has not been subjected to the Agency’s required peer review and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

References


