Error Analysis on Grid-Based Slope and Aspect Algorithms

Qiming Zhou and Xuejun Liu

Abstract
Slope and aspect are the most frequently used surface geomorphic parameters in terrain analysis. While derived from grid DEM, the parameters often display noticeable errors due to errors (a) in data, (b) inherent in data structure, and (c) created by algorithms. It has been observed that some controversial results were reported in evaluating the results by various slope and aspect algorithms, largely because of the variety in assessment methodology and the difficulties in separating errors in data and those generated by the algorithms. This paper reports the study that assesses and compares the results from numerous grid-based slope and aspect algorithms using an analytical approach. Tests were made based on artificial polynomial surfaces which can be defined by mathematical formulae, with controllable “added” data errors. By this approach, different algorithms were quantitatively tested and their error components were analyzed. Thus, their suitability and tolerance related to DEM data characteristics can be described.

Introduction
Slope and aspect have been regarded as two of the most important geomorphic parameters, as they not only efficiently describe the relief and structure of the land surface, but are also widely applied as vital parameters in hydrological models (Band, 1986; Moore et al., 1988; Quinn et al., 1991; 1995; Jenson, 1994), landslide monitoring and analysis (Duan and Grant, 2000), mass movement and soil erosion studies (Dietrich et al., 1993; Desmet and Govers, 1996a; Mitaseva et al., 1996; Biesemans et al., 2000) and landuse planning (Desmet and Govers, 1996b; Stephen and Irvin, 2000).

In most applications today, slope and aspect are typically derived from DEM that is based on raster data structure. To date, there have been numerous mathematical models and algorithms that calculate slope and aspect from elevation data (e.g., Sharpeack and Akin, 1969; Fleming and Hoffer, 1979; Horn, 1981; Unwin, 1981; O’Callaghan and Mark, 1984; Zavenbergen and Thorn, 1987; Wood, 1996). Although the mathematical definition of slope and aspect is quite clear, its implementation based on grid-based DEM may vary, since some assumptions must be made on how the continuous surface is approximated by discrete sample points (i.e., grid cells). The variation in such implementation would only present minor problems in applications such as surface visualization and classification, but its impact on terrain analysis that is based on quantitative models could be very significant (Zhou et al., 1998; Tang, 2000; Zhou and Liu, 2002). It was pointed out that selection of algorithms could be a critical factor that might create a great impact on the analytical results (Moore, 1996; Burrough and McDonnell, 1998).

Studies have been conducted to analyze the errors created by the slope and aspect algorithms with a variety of approaches and methodologies. One approach emphasizes errors in DEM itself while paying little attention on models. For example, Skidmore (1989) and Florinsky (1998) analyzed slope and aspect errors using the actual survey data. Another approach focused on mathematical models while ignoring DEM errors (e.g., Hodgson, 1995; Jones, 1998). The conclusions from these studies have been quite different, sometimes controversial. For example, Skidmore (1989) concluded that third-order finite difference method (Horn 1981) derived better results than the second-order finite difference method. His finding was also confirmed by Florinsky (1998). On the other hand, Hodgson (1995) and Jones (1998) stated that the latter performed better than the former in computing slope and aspect values.

To make a fair comparison between different slope and aspect algorithms, three pre-conditions must be satisfied:

1. The source of error should be identified;
2. The kinds of errors must be independent, identifiable and controllable in the test so that their impact can be quantified; and
3. A reference (or true value) must be established to make the results comparable.

Zhou and Liu (2002) reported a method that is based on mathematical surfaces to evaluate the errors generated by flow routing algorithms in digital terrain analysis. Using this method, errors created by an algorithm can be isolated from DEM errors so that the quantitative measures can be made to compare a derived, specific catchment area with the true value defined by the polynomial surfaces.

Using an analytical approach, Florinsky (1998) developed test for the precision of four methods for computing partial derivatives of elevations, and produced formulae for Root Mean Square Error (RMSE) of four local topographic variables including slope and aspect as focused by this paper. Florinsky also used a real-world DEM to map the error distribution for visual analysis, but the contribution from the DEM data error largely remained unknown.

This paper reports the study that attempts to extend the above methods into the test of slope and aspect calculation. In this study, we extended Florinsky’s method to general terms and designed tests on artificial polynomial surfaces with controllable data errors. Two representative surfaces, namely, the inverse ellipsoid and Gauss synthetic surface, were used to test six commonly-employed slope and aspect algorithms.
found in literature and GIS software. The derived slope and aspect values were then compared with the true value derived by mathematical inference. Quantitative analysis can then be conducted for the comparison of the algorithm accuracies.

### Slope and Aspect Algorithms

At a given point on a surface \( z = f(x, y) \), the slope (S) and aspect (A) is defined as a function of gradients at \( X \) and \( Y \) (i.e., W-E and N-S) directions, i.e.,

\[
S = \arctan \sqrt{f_x^2 + f_y^2}
\]

\[
A = 270^\circ + \arctan \left( \frac{f_y}{f_x} \right) - 90^\circ \left( \frac{f_x}{|f_x|} \right)
\]

where \( f_x \) and \( f_y \) are the gradients at W-E and N-S directions, respectively.

From the above equations, it is clear that the key for slope and aspect computation is the calculation of \( f_x \) and \( f_y \). Using a grid-based DEM, the common approach is to use a moving \( 3 \times 3 \) window to derive finite differential or local surface fitting polynomial for the calculation. Figure 1 shows various methods found in the literature and GIS software.

Considering the popularity and the use of different algorithms, we have selected six commonly-employed algorithms for test (Table 1):

- Second-order Finite Difference (2FD) (Fleming and Hoffer, 1979; Zevenbergen and Thorne, 1987; Ritter, 1987).
- Third-order Finite Difference Weighted by Reciprocal of Squared Distance (3FDWSRD) (Horn, 1981).
- Third-order Finite Difference Weighted by Reciprocal of Distance (3FDWDD) (Unwin, 1981).
- Frame Finite Difference (FFD) (Chu and Tai, 1995), and
- Simple Difference (SIMPLE-D) (Jones, 1998).

### Error Components

The accuracy of slope and aspect computation is directly related to the partial derivatives at \( X \) and \( Y \) direction, \( f_x \) and \( f_y \), which are estimated by numerous proposed methods. Taking the second-order finite difference as an example, let \((x, y)\) denote the coordinates of the center cell in a \( 3 \times 3 \) window and \( g \) denote the DEM spatial resolution (i.e., grid cell size), the partial differential can be expressed as (King, 2000):

\[
df_x = \frac{g^2}{6} \left( \frac{f_x(x+g, y) + f_x(x, y-g) - 2f_x(x, y)}{2} \right) + \frac{dz_x - dz_y}{2g}
\]

\[
df_y = \frac{g^2}{6} \left( \frac{f_y(x, y+g) + f_y(x, y-g) - 2f_y(x, y)}{2} \right) + \frac{dz_x - dz_y}{2g}
\]

The first term represents the errors caused by the uncertainty of mathematical model implementation and \( f_x \), \( f_y \), \( g \), and \( y \) are dependences of \( f_x \), \( f_y \), \( x \), \( y \), \( x \), \( y \), \( x \), \( y \), \( x \), \( y \), \( x \), \( y \), \( x \), \( y \), \( x \), \( y \), \( x \), \( y \), \( x \), \( y \). Since the relationships between these variables and \( x \) and \( y \) are usually not clear, it is difficult to define their values in application. Thus it is common to set the upper limits for \( f_x \) instead. Let \( M_x \) and \( M_y \) as the upper limits of \( f_x \) in terms of \( x \) and \( y \), respectively; Equation 3.1 can be altered as:

\[
df_x = \frac{g^2}{6} M_x + \frac{dz_x - dz_y}{2g}
\]

\[
df_y = \frac{g^2}{6} M_y + \frac{dz_x - dz_y}{2g}
\]

The second term of the equation represents DEM data error (including data precision).

<table>
<thead>
<tr>
<th>( f_x )</th>
<th>( f_y )</th>
<th>( f_z )</th>
<th>( f_{xy} )</th>
<th>( f_{xz} )</th>
<th>( f_{yz} )</th>
<th>( f_{xyz} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (x - z) / 2g )</td>
<td>( (z - x) / 2g )</td>
<td>( (z - y) / 2g )</td>
<td>( (x - z) / 2g )</td>
<td>( (z - x) / 2g )</td>
<td>( (z - y) / 2g )</td>
<td>( (x - z) / 2g )</td>
</tr>
</tbody>
</table>

**Table 1. Mathematical Models of Slope and Aspect Computation from DEM**

- **Simple Difference (SIMPLE-D)**
- **Weighted constrained quadratic surface**
- **Frame Finite Difference (FFD)**
- **Second-order Finite Difference (2FD)**
- **Third-order Finite Difference (3FD)**
- **Third-order Finite Difference Weighted by Reciprocal of Squared Distance (3FDWSRD)**
- **Third-order Finite Difference Weighted by Reciprocal of Distance (3FDWDD)**
In Equation 3.2, \( M_r \) and \( M_z \) are estimated based on the ‘worst scenario’, usually represent much larger estimates than the actual case, with a given distribution of probability. Let the RMSE of \( M_r \) and \( M_z \) equal to \( M \), and \( m \) denote the RMSE of the DEM, the result is:

\[
m_r^2 = m_z^2 = \left( \frac{g^2}{6} \right) M^2 + \frac{m^2}{2g^2} \tag{3.3}
\]

Deriving the differentials of slope and aspect equations (Equations 2.1 and 2.2), and considering \( S = \arctan \sqrt{f_x^2 + f_y^2} \) and \( \tan S = f_x^2 + f_y^2 \), the result is:

\[
dS = \frac{f_x df_x + f_y df_y}{(1 + \tan^2 S) \tan S} \quad \text{and} \quad dA = \frac{f_x df_x + f_y df_y}{\tan^2 S} \tag{3.4}
\]

Referring to Equation 3.3, the RMSE of slope \((m_s)\) and aspect \((m_A)\) can therefore be expressed as:

\[
\begin{align*}
m_s^2 &= \left[ \left( \frac{g^2}{6} \right) + \frac{m^2}{2g^2} \cos^2 S \right] \sin^2 S \\
m_A^2 &= \frac{1}{\sin^2 S} \left[ \left( \frac{g^2}{6} \right) M^2 + \frac{m^2}{2g^2} \sin^2 S \right] \tag{3.5}
\end{align*}
\]

Note that Equation 3.5 is for the second-order finite difference method. Let \( a = \frac{g}{6} \) and \( b = \frac{1}{\sqrt{2}} \); Equation 3.5 can then be expressed in a general form:

\[
\begin{align*}
m_s^2 &= \sqrt{a^2 M^2 + b^2 m^2} \cos^2 S \\
m_A^2 &= \frac{1}{\sin^2 S} \sqrt{a^2 M^2 + b^2 m^2} \sin^2 S \tag{3.6}
\end{align*}
\]

Similar to the above procedure and referring to Table 1, we can derive the RMSE for each selected algorithm as shown in Table 2.

Equation 3.6 and Table 2 show that the overall errors of derived slope and aspect come from three sources:

1. **Algorithm errors**: caused by approximation and sampling of continuous surfaces (Variable \( M \) in Equation 3.6),
2. **DEM data errors**: caused by DEM data capture and generation (Variable \( m \) in Equation 3.6), and
3. **DEM spatial resolution** (i.e., the grid cell size—Variables \( a \) and \( b \) in Equation 3.6, defined in Table 2).

(Note that the coefficient \( b \) of \( m \) for 2FD and 3FD confirm the results reported by Florinsky (1998), which represent special cases for the selected algorithms.)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Coefficient ( a ) of ( M )</th>
<th>Coefficient ( b ) of ( m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2\textsuperscript{nd} order Finite Difference</td>
<td>( \frac{1}{6g^2} )</td>
<td>( \frac{1}{\sqrt{2}g^2} )</td>
</tr>
<tr>
<td>3\textsuperscript{rd} order Finite Difference</td>
<td>( \frac{1}{6g^2} )</td>
<td>( \frac{1}{6g^2} )</td>
</tr>
<tr>
<td>3\textsuperscript{rd} order Finite Difference Weighted by Reciprocal of Squared Distance</td>
<td>( \frac{1}{6g^2} )</td>
<td>( \frac{1}{\sqrt{5.33}g^2} )</td>
</tr>
<tr>
<td>3\textsuperscript{rd} order Finite Difference Weighted by Reciprocal Distance</td>
<td>( \frac{1}{6g^2} )</td>
<td>( \frac{1}{5.83g^2} )</td>
</tr>
<tr>
<td>Frame Finite Difference</td>
<td>( \frac{1}{6g^2} )</td>
<td>( \frac{1}{2g} )</td>
</tr>
<tr>
<td>Simple Difference</td>
<td>( \frac{1}{8} )</td>
<td>( \frac{1}{2} )</td>
</tr>
</tbody>
</table>

**Test Design**

By examining Equations 3.5 and 3.6, we understand that for a given DEM spatial resolution \( g \), the accuracy of derived slope and aspect is related to the errors by the algorithm that derives \( f_x \) and \( f_y \) (\( M \)), and DEM error \( m \). The overall accuracy of slope and aspect is dependent upon which of the \( M \) and \( m \) dominates the analysis. While analyzing the algorithms using the real-world DEM (e.g., Skidmore, 1989; Chang and Tsai, 1991; Florinsky, 1998; Bolstad and Stove, 1994), the DEM error \( m \) would cause much greater error than that by the algorithm, so that the method would be more appropriate for analyzing the impact of data error on derived slope and aspect values. On the other hand, using DEM defined by mathematical surfaces (Hodgson, 1995; Jones, 1998; Carter, 1992) would eliminate data error thus the observed errors would only be caused by algorithms. Since the error sources could not be defined, the results of the studies appeared inconclusive.

Our approach is to employ mathematical surfaces, in a similar way to Hodgson (1995) and Jones (1998), but with a complexity that represents a closer approximation to the actual land surface. For this purpose, we have selected inverse ellipsoid (Equation 4.1) and Gauss synthetic surface (Equation 4.2) to define the surfaces and generated DEM for a given resolution (Figure 2 and Figure 3).

\[
\frac{x^2}{A^2} + \frac{y^2}{B^2} + \frac{z^2}{C^2} = 1 \quad (z < 0) \tag{4.1}
\]

\[
z = A \left[ 1 - \frac{x}{m} \right] e^{-\left( \frac{x}{m} \right)^2} - B \left[ 0.5 \frac{x}{m} - \left( \frac{x}{m} \right)^3 \right] - C e^{-\left( \frac{x}{n} \right)^{0.4}} - C e^{-\left( \frac{x}{n} \right)^{0.4}} \tag{4.2}
\]

where \( A, B \) and \( C \) are parameters determining surface relief, and \( m, n \) in Equation 4.2 are the parameters controlling the
The true values of slope and aspect can be computed using the above equations and Equations 2.1 and 2.2.

On the DEM generated from the above surfaces, the selected algorithms have been applied to compute slope and aspect values and statistics were then generated to compare the RMSE between results derived by different algorithms. In order to analyze the impact of DEM data error, we added some noise (i.e., random errors) into the generated DEM to simulate DEM data error.

### Results and Discussion

Tables 3 and 4 summarize the RMSE comparison results between the selected six algorithms on two surfaces (inverse ellipsoid and Gauss synthetic surface) with no added error (i.e., DEM data error = 0). Figures 4 and 5 show the RMSE of derived slope and aspect on DEM with various added data errors.

By analyzing Tables 3 and 4, it is shown that on a DEM with high accuracy, the error of derived slope and aspect is sourced from the estimates of partial derivatives $f_x$ and $f_y$ and sampling errors. In this case the second-order finite difference method provides a better result than the third-order finite difference methods. For the tested six algorithms, from the best to worst the order is Second-order Finite Difference (2FD), Third-order Finite Difference Weighted by Reciprocal of Squared Distance (3FDWRS), Third-order Finite Difference Weighted by Reciprocal of Distance (3FDWRD), Third-order Finite Difference (3FD), Frame Finite Difference (FFD), and Simple Difference (SIMPLE D).

In Equation 3.6, if we ignore the algorithm error (i.e., the first term), we then derive the Florinsky’s (1998) RMSE formulae in general terms:

$$m_S = bm \cos^2 S \quad m_A = \frac{bm}{\tan S}$$

Equation 5.1 suggests that the influence of DEM data error ($m$) relates to slope ($S$) and coefficient $b$, which is determined by grid cell size $g$. Referring to Table 2, the third-order finite difference methods appear to be less sensitive to data errors for given $g$ and $S$.

Algorithm error and DEM data error differently influence the accuracy of derived slope and aspect. In general, all the third-order finite difference methods have applied some smoothing effect on the local data window, in order to avoid local relief extremes (Burrough and McDonnell, 1998) for better computation results of surface parameters. On the other hand, the second-order finite difference and simple difference

### Table 3. Statistics of Derived Slope with No Data Error (Units: Degrees)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>2FD</td>
<td>0.216</td>
<td>0.003</td>
<td>0.189</td>
<td>0.002</td>
<td>0.104</td>
<td>-0.002</td>
<td>100/0</td>
<td>15/85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FD</td>
<td>0.358</td>
<td>0.003</td>
<td>0.320</td>
<td>0.002</td>
<td>0.160</td>
<td>-0.003</td>
<td>100/0</td>
<td>8/92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FDWRS</td>
<td>0.384</td>
<td>0.004</td>
<td>0.345</td>
<td>0.003</td>
<td>0.170</td>
<td>-0.003</td>
<td>100/0</td>
<td>9/91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FDWRD</td>
<td>0.507</td>
<td>0.004</td>
<td>0.459</td>
<td>0.003</td>
<td>0.214</td>
<td>-0.003</td>
<td>100/0</td>
<td>10/90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FFD</td>
<td>1.295</td>
<td>0.046</td>
<td>1.292</td>
<td>0.046</td>
<td>0.078</td>
<td>-0.002</td>
<td>51/49</td>
<td>49/51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4. Statistics of Derived Aspect on with No Data Errors (Units: Degrees)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
<th>Ellipsoid</th>
<th>Gauss Synthetic Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>2FD</td>
<td>0.133</td>
<td>0.117</td>
<td>0.133</td>
<td>0.117</td>
<td>0.000</td>
<td>-0.008</td>
<td>52/48</td>
<td>45/55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FD</td>
<td>0.197</td>
<td>0.130</td>
<td>0.197</td>
<td>0.130</td>
<td>-0.001</td>
<td>0.000</td>
<td>52/48</td>
<td>51/49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FDWRS</td>
<td>0.122</td>
<td>0.118</td>
<td>0.122</td>
<td>0.118</td>
<td>-0.001</td>
<td>-0.002</td>
<td>52/48</td>
<td>48/52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FDWRD</td>
<td>0.160</td>
<td>0.124</td>
<td>0.160</td>
<td>0.124</td>
<td>-0.001</td>
<td>-0.001</td>
<td>52/48</td>
<td>50/50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FFD</td>
<td>0.342</td>
<td>0.167</td>
<td>0.342</td>
<td>0.167</td>
<td>2.347</td>
<td>-0.340</td>
<td>50/50</td>
<td>52/48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIMPLED</td>
<td>20.76</td>
<td>15.13</td>
<td>20.63</td>
<td>15.128</td>
<td>20.63</td>
<td>15.128</td>
<td>20.63</td>
<td>15.128</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Comparison between accuracy of derived slope and aspect on inverse ellipsoid with increasing data errors.
methods only utilize a part of samples in the local window, so that they are more sensitive to the data error. This fact is also confirmed by Figures 4 and 5.

The results also confirm the test results reported by Jones (1998). While using an error-free synthetic, trigonometrically defined surface (Morrison’s surface III), the 2FD method gave the best results for both gradient and aspect. Using a real-world DEM, where data errors were unavoidable, the 3FD performed best.

**Conclusion**

Slope and aspect are two most frequently utilized variables in GIS-based terrain analysis and geographical modeling. There have been numerous analyses on the accuracy of these variables derived from grid-based DEM. The reported findings, however, did not always agree, and sometimes they are controversial. This study attempts to evaluate the issues and establish a fair quantitative measure for assessing various slope and aspect algorithms. With the findings of this study, we can draw the following conclusions:

1. It is important to identify the sources and nature of errors of derived slope and aspect for evaluating algorithms and mathematical models.
2. Evaluation of algorithms and models must be based on an objective, data-independent methodology. Our approach is based on DEM defined by mathematical surfaces, thus the DEM data error can be controlled to make the fair comparison between selected algorithms.
3. On a DEM with high accuracy, the error of derived slope and aspect is sourced from the estimates of partial derivatives \( f_x \) and \( f_y \) and sampling errors. In this case the second-order finite difference method provides a better result than the third-order finite difference methods.
4. In reality, the influence of DEM data error is in general much larger than the algorithm errors, thus it suggests that the third-order finite difference method would be more appropriate for applications since it is least sensitive to the DEM data error.

Further study will be focused on the impact of grid data structure and other related parameters such as resolution, precision, orientation, and surface complexity on the results of digital terrain analysis. Mathematical models other than slope and aspect will also be analyzed for their algorithm accuracy and sensitivity to data errors. The real-world tests will be needed to compare with the findings by the theoretical analysis. Based on these analysis, the ultimate goal is to set a conclusive guideline for deriving geomorphic parameters from DEM for a given application project.

**Acknowledgments**

This study is supported by Hong Kong Baptist University Faculty Research Grant FRG/98-99/I-II-35, “3-Dimensional Hydrological Modeling.” The constructive comments and suggestions for improving the quality of this paper from anonymous referees are also acknowledged.

**References**


