ELSEVIER

An evaluation of methods to determine slope using digital elevation data

S.D. Warren[a,*], M.G. Hohmann[b], K. Auerswald[c], H. Mitasova[d]

[a] College of Natural Resources, Colorado State University, Fort Collins, CO 80523-1490, USA
[b] US Army Engineering Research and Development Center, PO Box 9005, Champaign, IL 61826-9005, USA
[c] Chair of Grassland Science, Technische Universität München, D-85350 Freising-Weihenstephan, Germany
[d] Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695-8208, USA

Abstract

Variation in the computation of slope from digital elevation data can result in significantly different slope values and can, in turn, lead to widely varying estimates of environmental phenomena such as soil erosion that are highly dependent on slope. Ten methods of computing slope from distributed elevation data, utilizing capabilities inherent in five different geographic information systems (GIS), were compared with field measurements of slope. The methods were compared based on (1) overall estimation performance, (2) estimation accuracy, (3) estimation precision, and (4) independence of estimation errors and the magnitude of field measured slopes. A method utilizing a very high resolution digital elevation model (DEM) (1 m) produced slightly better estimates of slope than approaches utilizing somewhat lower resolution DEMs (2–5.2 m), and significantly better estimates than a method utilizing a 12.5 m DEM. The more accurate method was significantly biased, however, frequently underestimating actual slope. Methods that averaged or smoothed high resolution DEMs over larger areas also produced good estimates of slope, but these were somewhat less accurate in areas of shallow slopes. Methods utilizing differential geometry to compute percent slope from DEMs outperformed methods utilizing trigonometric functions. Errors in slope computation are exaggerated in soil erosion prediction models because erosion typically increases as a power function of slope.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Digital elevation model; Geographic information system; Slope; Soil erosion

* Corresponding author.
E-mail address: SWarren@cemml.colostate.edu (S.D. Warren).

1. Introduction

Numerous empirical and process-based models have been developed to predict environmental phenomena such as soil erosion and sediment deposition. In order for the models to be applied with maximum effectiveness in large areas of complex terrain, integration with geographic information systems (GIS) is requisite. Slope steepness is a fundamental parameter in most soil erosion models. In a GIS environment, the most efficient method to determine slope is through the use of digital elevation models (DEMs). DEMs represent topography either by a series of regular grid points with assigned elevation values or as a triangulated irregular network (TIN) where each point is stored by its coordinates and the surface is represented by triangular facets.

Various techniques have been devised to compute slope from DEMs. Different computational methods, however, may produce significantly different slope values from the same DEM (Snyder, 1983; Skidmore, 1989; Srinivasan and Engel, 1991). Variation in the computation of slope can, in turn, lead to widely varying estimates of soil erosion (Srinivasan and Engel, 1991). Soil erosion models are especially sensitive to errors in slope because runoff-induced erosion increases as a power function of slope (Zingg, 1940; Nearing, 1997). Most soil erosion models, while operable within GIS, were not developed using DEM-generated slopes. Rather, they were developed based on slopes obtained from field or plot measurements. Hence, when applying soil erosion models within GIS, it is imperative that DEM-generated slopes are compatible with measurements made in the field. Few studies, however, have compared computed estimates of slope with actual field measurements (e.g., Bolstad and Stowe, 1994; Giles and Franklin, 1996), and none has evaluated this with respect to soil erosion prediction.

In this paper, we contrast 10 methods of computing slope from distributed elevation data and compare all with field measurements of slope. The comparison of the slope estimation methods is based on four criteria including (1) overall estimation performance, (2) estimation accuracy (i.e., amount of systematic error or bias in the estimates), (3) precision (i.e., random error), and (4) independence of the estimation errors (i.e., residuals) and the magnitude of the field-measured slopes. The effects of errors in slope computation are evaluated in light of their impact on soil erosion estimation.

2. Methods

To evaluate the capabilities of various GIS-based methods to provide reliable estimates of slope using the elevation data obtained by geodetic survey, we compared computed slopes with slopes measured directly in the field. When comparing the relative capabilities of different GIS-based methods, two basic approaches are possible. A single operator may perform all operations, which will minimize differences due to operator bias. Alternatively, operators with long experience with particular systems may be selected, thus taking maximum advantage of system capabilities. A basic assumption of erosion modeling is that the model results are independent of the user so long as the person has profound knowledge of the methods used. Hence, we followed the second approach. Each operator was selected based on expertise with a particular GIS system. All operators were aware that it

was important to create a DEM and resulting estimations of slope which would closely agree with field measurements. This lead to different approaches by the different GIS operators based on the capabilities of the systems used, but reflects the reality of modeling with GIS. As derivation of slope with GIS is not normally optimized by feedback mechanisms based on field validation, and as some GIS allow for different approaches, one of the assigned operators tried multiple approaches with their assigned GIS (SPANS). To further highlight personal preferences, which may be important but are rarely documented in modeling papers, one GIS (ArcInfo) was operated by two different people who opted for contrasting approaches.

2.1. Field measurements

The Scheyern Experimental Station of the Forschungsverbund Agrarökosysteme München (Munich Agroecosystems Research Alliance) was used as the field study site. The 143-ha study area is a hilly landscape that has been under cultivation for several centuries. Arable slopes range up to 30%. An intense tachymetric survey was conducted at the research station. Precise elevation was recorded for 500 locations distributed as a 50 × 50 m grid (Fig. 1). Approximately 4000 additional positions were surveyed in areas of complex topography in order to represent the landscape in more detail. The use of additional survey points helps eliminate the Gibbs phenomenon which creates large errors in DEMs in areas of jump discontinuity (Florinsky, 2002). All points were available for the computation of a DEM.

Slope measurements were made at 57 randomly distributed points throughout the central part of the study area after harvest when the soils had settled. Slope was measured over a 5 m distance in the direction of the perceived maximum slope. At each point, measurements were made with a water level in three replicates over a width of 1.5 m. Accuracy of the measurements was at least 0.1%. The average 95% confidence interval for the three replicates was 0.3% and contains measurement error as well as variation in slope over the 1.5 m width. Slopes ranged in steepness from 1.6% to 23.6%.

2.2. Methods of slope estimation

Five different GIS were used to process the survey data. In each case, data processing was conducted by a person with several years of specialized experience with the particular system, thus minimizing the impact of user capabilities on potential differences in outcomes. None of the operators were allowed access to the measured slope data. With algorithms inherent in each system, the operators computed their respective DEMs and their best estimates of slope at the 57 aforementioned randomly distributed points.

Two basic approaches were used to compute the slope in percent; the first uses trigonometry and the second is based on differential geometry. The first approach computes percent slope as a change in elevation Δz over a certain distance Δs

$$\text{slope}(\%) = 100(\Delta z/\Delta s) \tag{1}$$

where Δz is the difference in elevation (m) between the given grid point and the lowest grid point in its neighborhood, and Δs is the distance (m) between those two points
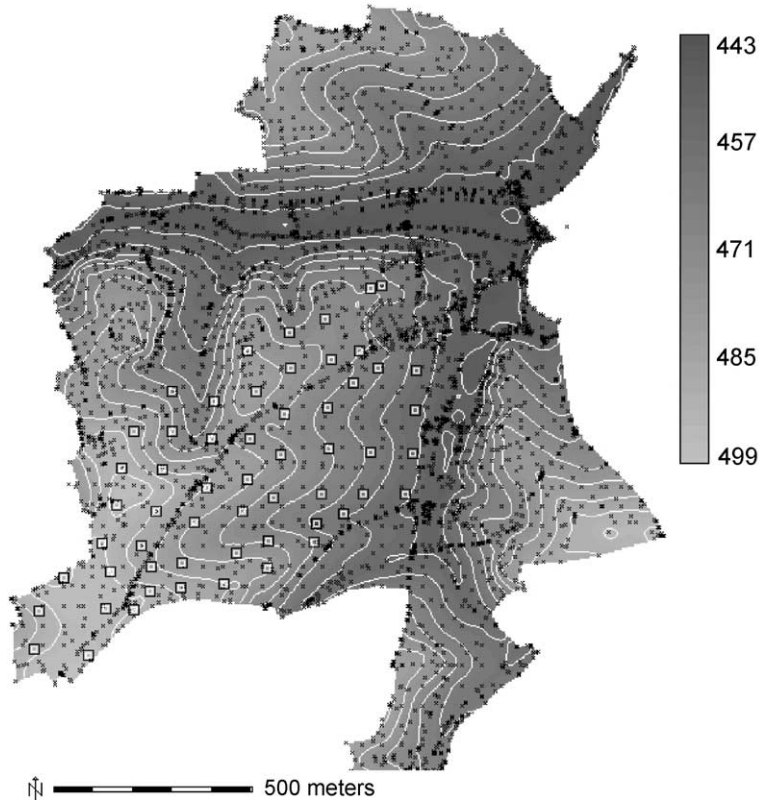
Fig. 1. An elevation map of the Scheyern Experimental Station near Freising, Germany. The small x's represent locations where elevation measurements were taken to allow the development of a digital elevation model. The boxes represent the points where precise field measurements of slope were made. Elevation is depicted in meters.

(usually a grid cell side or a diagonal). It is important to note that this approach considers only eight possible directions of steepest slope.

The second approach is based on differential geometry and computes percent slope as a magnitude of the gradient vector (tangent vector of the surface pointing in the direction of steepest slope). With this approach, slope is computed at a grid point rather than over a distance as it is in the previous case. The general equation is

$$\text{slope}(\%) = 100\sqrt{f_x^2 + f_y^2} \tag{2}$$

where f_x^2 and f_y^2 are partial derivatives of a bivariate function $z = f(x, y)$ which represents the elevation surface. Various methods use different approximation functions such as spline or bivariate polynomials. The most common approximation for regular grid DEMs is a bivariate second order polynomial which leads to a very simple estimate of the derivatives as weighted averages of the elevation differences between the given point and all points within its 3×3 neighborhood (Horn, 1981; Neteler and Mitasova, 2002).

2.2.1. HIFI88

The HIFI88 system was developed to create topographic maps from geodetic data (Ebner et al., 1987) and does not include full GIS capabilities. Landscape surfaces are represented in a regular grid format, with triangulated irregular networks (TIN) located within the grid cells to more accurately describe break lines in the topography. This method allows rapid and efficient processing of the data. A finite elements method is used to estimate elevation at grid points based on interpolation from the given elevation values. A least squares solution for the surface is calculated assuming that (a) the surface fits as close as possible to the input data points, and (b) the surface is as smooth as possible.

For this analysis, a 10 m resolution grid was initially calculated for the entire study site using the data from the field survey points. This size was chosen both for computational efficiency and to meet the requirement that no more than nine grid cells were represented by each data point. In order to create a satisfactory grid density around the points where slope was to be estimated, a 4 m resolution grid was generated from the original 10 m DEM. A subsequent DEM was interpolated for a 1 m resolution grid within a 20×20 m area surrounding each point of interest. A standard local polynomial interpolation was used to compute slopes on a regular grid using the differential geometry approach (Eq. (2)). Slope was recorded for the grid points closest to the test points and an average was taken. This average was used for comparison with field slope measurements.

2.2.2. PCRaster

We utilized an older shareware version of PCRaster, a raster GIS developed by the University of Utrecht for educational purposes (van Deursen and Wesseling, 1993). The operator of this system was less-experienced than the operators of the other GIS. Interpolation among the scattered point data was conducted by ordinary point kriging utilizing the included Gstat package (http://www.gstat.org). The result was a 12.5 m resolution DEM. The slope of each grid cell was then calculated as the maximum change in elevation over the distance between the given cell and the lowest one in a 3×3 cell window using the trigonometric approach (Eq. (1)).

2.2.3. Arc/Info

With ArcInfo version 6.0 (Environmental Systems Research Institute, 1991), a triangulated irregular network (TIN) surface was first created from the scattered elevation data using the Delaunay method of triangulation. From the TIN, a surface grid was created using two different interpolation methods. With the linear interpolation method, each triangle was considered planar and autonomous, thus creating a multifaceted surface. The elevation of any point falling within a triangle was calculated based solely on the elevation values for the triangle nodes. Grids with 2 m (Arc/Info1) and 2.5 m (Arc/Info2) resolution were created with this method.

With quintic interpolation, the geometry of neighboring triangles was considered, thus allowing for undulations on individual triangle facets and creating a smooth, continuous surface overall. Quintic interpolation uses a bivariate fifth-degree polynomial. Break lines were not included in the calculation of the quintic interpolation. This method was used to compute a 2.5 m resolution DEM (Arc/Info3).

Slope was determined for the resulting DEM on a cell by cell basis using the trigonometric approach (Eq. (1)), calculating the maximum rate of change between each cell and its surrounding 3 × 3 neighborhood.

2.2.4. SPANS

Version 5.3 of the Spatial Analysis System (SPANS) (http://www.pcigeomatics.com) was used to generate DEMs at resolutions of 2.59 and 5.20 m. First, all elevation points were connected with their nearest neighbors to form a TIN surface. Each triangle in the network formed a plane. Using a linear interpolation, the TIN was converted to a grid DEM which was then smoothed using a 3 × 3 filter. The differential geometry approach (Eq. (2)) using the standard polynomial approximation applied to a 3 × 3 matrix of elevation values was used to compute the slope in percent. Finally, the resulting slope was smoothed using 3 × 3 (SPANS3), 7 × 7 (SPANS7) and 15 × 15 (SPANS15) grid matrices at the 2.59 m resolution, and a 3 × 3 (SPANS3b) matrix at the 5.20 m resolution.

2.2.5. GRASS

The Geographic Resources Analysis Support System (GRASS) version 4.1 (http://www.grass.itc.it/) was used to derive a 2 m resolution DEM from the scattered elevation points. The model was computed by interpolating between the given data points using a regularized smoothing spline function with tension (Mitasova and Mitas, 1993). The method assumes that the approximation function should pass as closely as possible to the given data points and should be as smooth as possible. A tension parameter controls the distance over which a given point influences the resulting surface model and enables the user to tune the character of the resulting surface from a membrane to a thin plate. A smoothing parameter controls the deviation of the resulting surface from the individual data points. The parameters can be selected empirically by visual analysis of the interpolated surface or they can be optimized by minimizing the cross-validation error (Mitasova et al., 1995; Hofierka et al., 2002). In this application, the parameters were adjusted empirically using visual analysis to minimize overshoots and deviations. The selected tension parameter was 40 and the smoothing parameter was set to 0.1. Slope was determined for each grid cell based on the differential geometry approach (Eq. (2)) using the partial derivatives of regularized spline with tension (GRASSrst).

2.3. Statistical analyses

Because the design of the study allowed individual researchers some professional discretion, analysis of data and interpretation of findings were complicated by several uncontrolled but realistic sources of variation. Therefore, any observed variation in estimation ability among methods does not necessarily result solely from differences in the methods per se. Differences in the scale of estimation, due to differences in DEM grid-size and estimation neighborhood, will also affect estimates. Therefore, in addition to the unconstrained comparison of methods, we also examined subsets of methods that differed by a single parameter (e.g., grid size, estimation neighborhood, or algorithm) in an attempt to identify the influence of specific parameters on the estimates.

2.3.1. Exploratory data analysis

We tested the normality of the field-measured slopes, estimation errors (i.e., difference between the measured and estimated slopes), and relative errors (i.e., ratio of estimation errors and field-measured slopes) using the methods of D'Agostino et al. (1990). We also examined the spatial structure of the field-measured slopes and estimation errors using correlograms. It is critical that data are assessed for spatial independence prior to statistical analyses, since spatially autocorrelated data violate the assumption of sample independence, thereby biasing the interpretation of inferential tests (Legendre, 1993). Spatial autocorrelation in the field-measured slopes and estimation errors was assessed with correlograms (Sokal and Oden, 1978). Autocorrelation coefficients (Moran's I) were tested against the null hypothesis H_o: $I = 0$. Moran's I ranges between -1 and $+1$; positive coefficients indicate aggregation of similar values and negative coefficients indicate segregation. Significant autocorrelation indicates that the value of the variable at a given location depends on the values at neighboring locations. Critical values were calculated from the standard error of I, and the hypothesis of spatial dependence was tested with t. The overall significance of the correlograms was evaluated at a conservative Bonferroni-corrected critical level prior to evaluating the significance of coefficients at individual distance lags. This corrected probability ($\alpha = 0.006$) was calculated as the quotient of the critical probability level ($\alpha = 0.05$) and the number of distance lags ($n = 8$) for which Moran's I was calculated (Oden, 1984). The autocorrelation analysis was performed using GS^+ (Version 2.3b, Gamma Design Software, 1994).

2.3.2. Overall estimation performance

We measured the overall estimation performance of the various methods with the standard error of prediction or root-mean-square error (RMSE), which is calculated as the square root of the mean of squared differences between estimated (x_i) and measured slopes (t_i) for n sample locations:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - t_i)^2}{n - 1}} \tag{3}$$

Methods with low RMSE have greater predictive ability than methods with large RMSE (i.e., estimates deviate less from the field-measured slopes). However, because the RMSE is based on the squared differences between the estimates and the measured slopes, large errors will have a disproportionate influence on the value of the index.

To assess the magnitude of the estimation errors relative to the steepness of the field-measured slopes, we also calculated the RMSE for the relative errors $[(x_i - t_i)/t_i]$, which we henceforth refer to as the relative root-mean-square error (RRMSE). When multiplied by 100, this index represents the error of the slope estimates as a percentage of the field-measured slopes. The RRMSE is particularly sensitive to estimation errors in flat areas, as large estimation errors at locations having shallow slopes will assert a greater effect on the value of the index than equivalent errors at locations with steep slopes. We estimated the variance of the RMSE and RRMSE with the jackknife procedure (Manly, 1991). Confidence intervals for the indices were calculated as $t_{n-1}\text{SE}$, where SE is the standard

error of the jackknifed distribution and t_{n-1} is the critical value of a two-sided 95% confidence interval from a t distribution with $n-1$ degrees of freedom. When making multiple comparisons, confidence intervals can provide more insight than significance tests as they reveal the degree of uncertainty in each comparison.

2.3.3. Estimation accuracy

We assessed the accuracy of the different methods by comparing the slope estimates and field-measured slopes. We used Williams' corrected G-test to determine whether the frequency of over- versus underestimated slopes differed (Sokal and Rohlf, 1994). For an unbiased estimation method, the number of under- and overestimated slopes is expected to be equal. For this test, a P-value of 1.00 indicates perfect fit to expectation, and a P-value of 0.000 indicates a complete lack of fit. Tests were evaluated at a sequential Bonferroni adjusted critical value ($P < 0.005$) to account for multiple comparisons. Because information about the magnitude of the errors is not included in the G-test, we also tested whether the mean difference between the estimated and measured slopes was significantly different from zero with paired-comparisons t-tests. For an unbiased estimation method, the mean difference should not differ from zero.

2.3.4. Estimation precision

We assessed the precision of the different estimation methods by first centering the error distributions to have a zero median (i.e., for each method, we subtracted the median from each residual) and then calculating the RMSE of the centered distributions. Confidence intervals (CI) for the RMSE indices were calculated as $\pm t_{n-1}SE$, where SE is the standard error of the jackknifed distribution and t_{n-1} is the critical value of a two-sided 95% confidence interval from a t-distribution with $n-1$ degrees of freedom (Manly, 1991).

2.3.5. Independence of estimation errors and field-measured slopes

To test whether estimation errors were independent of slope, we correlated field-measured slopes and estimation errors. Because data distributions were skewed and unimproved by transformations (see Section 3.1), correlations were tested with nonparametric and randomization methods. Outliers were identified and removed prior to analysis. For each estimation method, linear associations between estimation errors and slope were tested with Pearson's product-moment correlation coefficient (r). The significance of the correlation coefficients was evaluated with P-values representing the likelihood of obtaining a correlation coefficient larger than the observed simply by chance (Manly, 1991). In order to test for nonlinear associations, we calculated Spearman's rank-order correlations (r_s) (Sokal and Rohlf, 1994). We also calculated (r_s) for the absolute values of the estimation errors and the field-measured slopes to assess associations between the slope steepness and magnitude of the estimation errors, without regard to whether slopes were over- or underestimated. In an attempt to account for the increased likelihood of committing type I errors, due to spatial autocorrelation (see Section 3.1), we evaluated correlations at a conservative level of significance ($P = 0.01$), and with a sequential Bonferroni adjustment for multiple comparisons (Rice, 1989). Although more exact methods of correcting for spatial dependence in variables have been

developed for correlation analysis, they are limited to normally distributed data (e.g., Dutilleul, 1993).

2.4. Erosion modeling analysis

We assessed the effect of slope estimation method on soil loss prediction using the Universal Soil Loss Equation. The equation has the form

$$A = RKLSCP \qquad (4)$$

where A is average annual soil loss, R represents the erosivity of average rainfall and runoff, K is a measure of soil erodibility, L represents the effect of slope length, S is a measure of the effect of slope steepness, C represents the influence of plant cover, and P represents conservation support practices (Wischmeier and Smith, 1978). We calculated S using an equation by Nearing (1997) to account for the effect of slope. To avoid introducing additional sources of variation, the product of the remaining parameters was held constant at 3.38 t ha^{-1} year^{-1} which represents the average conditions at the site (Fiener and Auerswald, 2003).

3. Results

3.1. Exploratory analyses

The distribution of field-measured slopes was positively skewed ($P<0.01$) and leptokurtic ($P<0.05$). Similarly, the distributions of the estimation and relative errors were positively skewed for all 10 methods ($P<0.01$), and leptokurtic for all methods except Arc/Info1, PCRaster, and SPANS15 ($P<0.05$). Data transformations did not improve the normality of the distributions.

The field-measured slopes exhibited significant spatial autocorrelation (Moran's I) at all distance lags examined (Table 1). At distances less than 150 m and greater than 330 m, field-measured slopes were negatively autocorrelated; at intermediate distances, they were positively autocorrelated. The positive and negative autocorrelations of the field measurements are a result of the asymmetric valley shapes with positive correlation for distances in locations on some slopes and negative correlations of sites on opposite slopes. For the estimation errors, most methods exhibited positive autocorrelation to distances of approximately 150–210 m. The PCRaster and HIFI88 methods were exceptions, as their residuals were not autocorrelated at any distance lags less than 210 m. At average lag distances greater than 210 m, most estimation methods either did not have autocorrelated errors, or exhibited alternating positive and negative spatial autocorrelation.

3.2. Overall estimation performance

Comparisons of the root-mean-square errors (RMSE) and 95% confidence intervals indicate that the HIFI88, GRASSrst, and SPANS15 methods have a significantly lower

Table 1
Positive (+) and negative (−) autocorrelation coefficients (Moran's I) significant at a Bonferroni adjusted critical value of $P < 0.006$ and lag distances for the field-measured slopes and estimation errors

Estimation method	Distance lag (m)							
	90	150	210	270	330	390	450	510
Field-measured slopes	−	+	+	+	+	−	−	−
Arc/Info1	+	+			−	+	−	−
Arc/Info2	+		+					
Arc/Info3	+	+	+		+	−	+	
GRASSrst	+	+				−	+	−
HIFI88			+	−	+	−	+	−
PCRaster			−	−				
SPANS3	+		+					−
SPANS3b	+		+					−
SPANS7	+	+						−
SPANS15	−	+						−

estimation error than the PCRaster and Arc/Info1 methods (Fig. 2). None of the other methods differed significantly from one another. Errors relative to the field-measured slopes were lower for the HIFI88 method (28%) than for the Arc/Info1, Arc/Info2, PCRaster, and SPANS3 methods (>49%, Fig. 3). None of the other estimation methods differed from each another. Unlike the results for the RMSE, the GRASSrst and SPANS15 methods, as well as all other slope computation methods, had comparatively higher RRMSEs than the HIFI88 method.
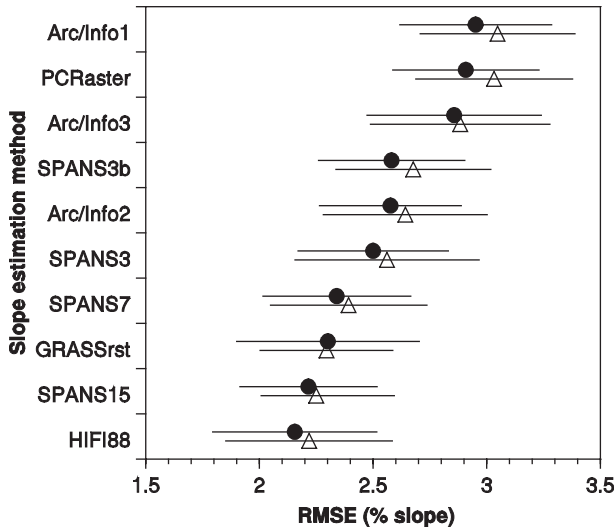

Fig. 2. The root-mean-square error (RMSE) as a measure of overall estimation performance (open triangles) and estimation precision (closed circles) of the 10 methods of estimating slope. Lower values are indicative of better performance. The horizontal bars represent the confidence intervals.
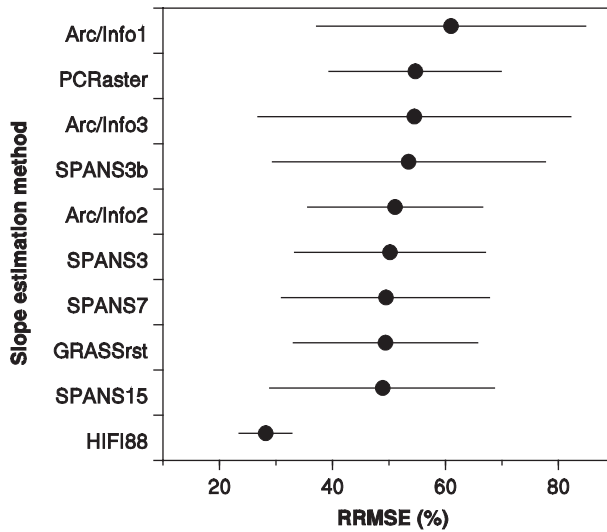
Fig. 3. The relative root-mean-square error (RRMSE) multiplied by 100, as shown here, represents the error of slope estimates as a percentage of field measured slopes. Lower values are indicative of better performance. The horizontal bars represent confidence intervals.

3.3. Estimation accuracy

When evaluated at a conservative sequential Bonferroni critical value ($P < 0.005$), the results of the G-tests for goodness-of-fit indicate that none of the estimation methods are significantly inaccurate (Table 2). Paired-comparison t-tests gave similar results (Table 3). However, if these two analyses are evaluated without the conservative Bonferroni adjustment (i.e., at $P < 0.05$), then the Arc/Info1, HIFI88, and SPANS3b methods are

Table 2
Number of over- versus underestimated slopes using Williams' corrected G-tests for goodness-of-fit[a]

Estimation method	# Underestimated	# Overestimated	G	P
Arc/Info1	34	16	5.61	0.018
Arc/Info2	31	19	1.89	0.170
Arc/Info3	30	20	0.99	0.319
GRASSrst	29	21	0.27	0.606
HIFI88	34	16	5.61	0.018
PCRaster	20	30	0.99	0.319
SPANS3	29	21	0.27	0.606
SPANS3b	33	17	4.19	0.041
SPANS7	31	19	1.89	0.170
SPANS15	30	20	0.99	0.319

[a] Bonferroni critical value for tablewide significance is $P < 0.005$. Estimation methods in bold text are significantly biased when evaluated without the sequential Bonferroni adjustment in critical value (i.e., $P < 0.05$).

Table 3
Paired-sample *t*-tests for mean difference between estimated and field-measured slopes[a]

Estimation method	Mean difference	S.D.	SE of mean	95% CI	*t*-value	*P* (two-tailed test)
Arc/Info1	− 0.8449	2.663	0.380	− 1.610, − 0.080	− 2.22	0.031
Arc/Info2	− 0.6327	2.282	0.326	− 1.288, 0.023	− 1.94	0.058
Arc/Info3	− 0.3143	2.533	0.362	− 1.042, 0.413	− 0.87	0.389
GRASSrst	− 0.2265	2.019	0.288	− 0.806, 0.353	− 0.79	0.436
HIFI88	− 0.6143	1.742	0.249	− 1.115, − 0.114	− 2.47	0.017
PCRaster	0.7143	2.658	0.380	− 0.049, 1.478	1.88	0.066
SPANS3	− 0.5898	2.193	0.313	− 1.220, 0.040	− 1.88	0.066
SPANS3b	− 0.7673	2.288	0.327	− 1.424, − 0.110	− 2.35	0.023
SPANS7	− 0.5347	2.000	0.286	− 1.109, 0.040	− 1.87	0.067
SPANS15	− 0.4020	2.006	0.287	− 0.978, 0.174	− 1.40	0.167

Mean differences are in units of % slope.
  [a] Bonferroni critical value for tablewide significance is *P* < 0.005. Estimation methods in bold text are significantly biased when evaluated without the sequential Bonferroni adjustment in critical value (i.e., *P* < 0.05).

biased estimators, as they more frequently underestimated than overestimated slopes, and had mean differences less than zero (Tables 2 and 3).

3.4. Estimation precision

The HIFI88 and SPANS15 methods had greater precision (i.e., smaller RMSEs for median centered estimation errors) than the PCRaster and Arc/Info1 methods; no other methods differed in precision (Fig. 2).

3.5. Independence of estimation errors and field-measured slopes

Spearman rank-correlation analyses using the conservative Bonferroni adjustment indicate that the errors of the GRASSrst, PCRaster, SPANS7, and SPANS15 methods

Table 4
Spearman's rank-order correlations and Pearson's product-moment correlations between the field-measured slopes and estimation errors

	Spearman's rank-order correlation		Pearson's product-moment correlation	
	Measured slopes	*P*	Measured slopes	*P*
Arc/Info1	− 0.4115	0.004*	− 0.441	0.0013*
Arc/Info2	− 0.4155	0.004*	− 0.438	0.0007**
Arc/Info3	− 0.3068	0.038	− 0.298	0.0225
GRASSrst	− 0.5132	0.000**	− 0.491	0.0003**
HIFI88	− 0.3207	0.030	− 0.351	0.0085*
PCRaster	− 0.4572	0.001**	− 0.451	0.0004**
SPANS3	− 0.3558	0.016	− 0.332	0.0124
SPANS3b	− 0.4280	0.004*	− 0.381	0.0046*
SPANS7	− 0.5592	0.000**	− 0.534	0.0000**
SPANS15	− 0.6748	0.000**	− 0.605	0.0000**

  *Significant at *P* < 0.01.
  **Significant at the sequential Bonferroni adjusted level of *P* < 0.001.

correlated negatively with the field-measured slopes (Table 4), suggesting that over-estimates are associated with shallow slopes and/or underestimates are associated with steep slopes. Pearson's correlation analyses gave similar results. Neither the rank, nor randomization correlation tests identified any significant association between the absolute estimation errors and the field-measured slopes even when evaluated at a less conservative P-value without the sequential Bonferroni adjustment.

3.6. Erosion modeling analysis

Estimated soil loss averaged 3.6 t ha^{-1} year^{-1} when using measured slopes. Deviations ranged from 0 to 4.5 t ha^{-1} year^{-1}, with a mean of 0.8 t ha^{-1} year^{-1} when using computed slopes (Fig. 4). The ranking between methods was similar to the ranking for precision. Some locations exhibited a strong deviation in one direction for all methods (note especially outlying data points where the estimated soil loss from measured slope was approximately 3.5 t ha^{-1} year^{-1}) indicating that this problem was sometimes independent of the method of
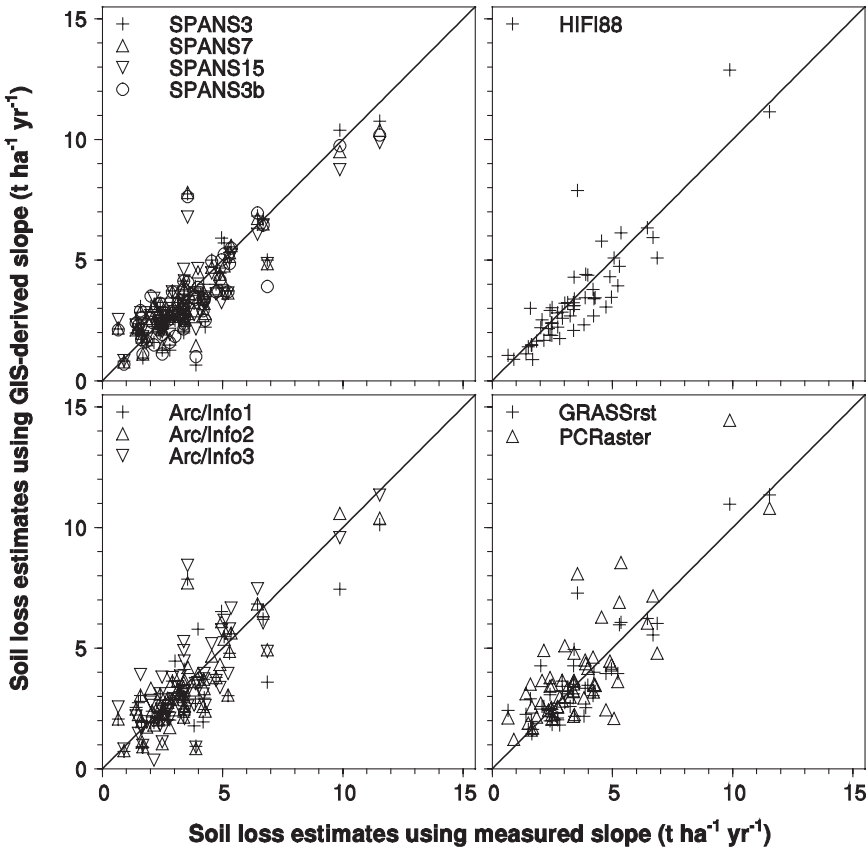

Fig. 4. Soil loss estimates predicted by the Universal Soil Loss Equation from measured and GIS-derived slopes. All other parameters of the USLE were held constant with average conditions at the site.

slope computation. In other cases, strong positive or negative deviations were the result of slope computation methodology. For example, soil loss estimates using computed slopes ranged from 4.6 t ha^{-1} year^{-1} greater to 2.4 t ha^{-1} year^{-1} less than the soil loss estimate of 9.9 t ha^{-1} year^{-1} when measured slope was used in the computation of soil loss (Fig. 4).

4. Discussion

In our analysis, the HIFI88 method appeared to outperform most other methods of slope estimation. It produced the lowest overall RMSE and RRMSE and exhibited the highest precision of all the methods tested. However, it did show some degree of estimation bias, frequently underestimating the measured slope. The GRASSrst and SPANS15 methods also produced low RMSEs, but were unbiased in their estimations of slope. However, their RRMSEs were marginally higher than the HIFI88 method suggesting that their slope estimates were somewhat less accurate at sites with shallow slopes, a phenomenon that is not uncommon when deriving slope estimates from DEMs (Florinsky, 1998). This observation is confirmed by the highly significant correlation of their errors with the field-measured slopes.

Among the methods tested, the poorest performers were Arc/Info1 and PCRaster. These methods produced the highest overall RMSEs and were the least precise. In addition, the Arc/Info1 method had a tendency to underestimate the measured slopes, while the errors of PCRaster were negatively correlated to the measured slope, indicating a tendency for greater inaccuracy at the high and/or low extremes of the range of slopes measured in this study.

4.1. Impact of resolution

The Nyquist frequency, which is twice the grid resolution, describes the lower size limit of the terrain features that a grid-based DEM is able to delineate (Table 5). When using the trigonometric approach, slope is measured as a change in elevation (*z*) over some distance (*s*), and the minimum distance over which slope can be measured in a raster-based DEM is between two adjacent grid cells. Consistent with the Kotelnikov-Shannon sampling

Table 5
Grid size and minimum scale of precision (Nyquist frequency) for the different methods of slope estimation

Estimation method	Grid size (m)	Nyquist frequency (m)
Arc/Info1	2 × 2	4
Arc/Info2	2.5 × 2.5	5
Arc/Info3	2.5 × 2.5	5
GRASSrst	2 × 2	4
HIFI88	1 × 1	2
PCRaster	12.5 × 12.5	25
SPANS3	2.59 × 2.59	5.18
SPANS3b	5.2 × 5.2	10.4
SPANS7	2.59 × 2.59	5.18
SPANS15	2.59 × 2.59	5.18

theorem (Benedetto and Ferreira, 2001), the grid cell size of the DEM must be less than one half the size of the smallest geographic unit to be investigated. Therefore, to estimate slope over a 5 m hillslope segment, the DEM resolution should be at least 2.5 m. When using the differential geometry approach, slope is computed at a point. To properly capture the geometry of a 5 m long slope by the approximation function, at least 2.5 m resolution is again needed. Most of the GIS operators selected DEM resolutions between 1.0 and 2.6 m (Table 5).

ArcInfo and SPANS use a TIN to create a DEM, which is then interpolated to a grid. HIFI88 uses a raster DEM which is further divided into triangles to better capture break lines. PCRaster and GRASSrst compute raster DEMs directly. All approaches except GRASSrst use grids to estimate slope; GRASSrst computes slope directly from given points. Therefore, the minimum precision of four of the estimation methods is directly constrained by grid resolution. If differences in performance among methods are primarily an artifact of DEM resolution, then one might expect a positive relationship between grid size and estimation accuracy (i.e., small grid size results in low RMSE). Consequently, the relatively poor performance of the PCRaster method may be at least partially related to the lower DEM resolution (12.5 m) used to calculated slopes, rather than the accuracy of the interpolation and slope estimation algorithms. By the same token, the somewhat better performance of the HIFI88 method may be a consequence of a higher resolution DEM. However, the SPANS3b method, utilizing a DEM resolution of 5.2 m, or slightly more than double the recommended minimum, was statistically similar to all other methods in terms of accuracy (Fig. 2). More importantly, the performance of the SPANS3b method was very similar to that of the SPANS3 method. The two methods were identical except for the resolution of the DEM they generated, indicating that, within limits, DEM resolution may be less important than other factors in determining methodological accuracy.

4.2. Impact of smoothing and slope averaging

In addition to differences in DEM interpolation method and resolution, the accuracy of the various slope estimates is influenced by the degree of smoothing applied during the computation of the DEM as well as by the size of the area that is considered when computing the slope for a given point (either in terms of the DEM resolution or the area from which the slopes are averaged). If one assumes that slopes computed with the least smoothing using the smallest area will provide the most accurate estimates, then the relationship of RMSEs among the various methods should be ordered as PCRaster and SPANS15>SPANS3b and SPANS7>Arc/Info1, Arc/Info2, Arc/Info3, HIFI88, and SPANS3. Note that GRASSrst was not included in this comparison, because it estimates slopes over variable areas depending on the spatial distribution of data (see Section 2) and a tension parameter. Based on this assumption, the PCRaster and SPANS15 would be expected to have poor predictive ability since they computed slopes using the elevation data over larger areas (PCRaster using elevation matrix with $3 \times 12.5 = 37.5$ m length) or averaged slopes over larger areas (SPANS15 averaging slopes from matrix with $15 \times 2.59 = 38.9$ m length). Yet, the SPANS15 method had a comparatively low RMSE, while PCRaster method had a high RMSE. The apparent difference in RMSEs may be

attributed to the finer resolution of the SPANS15 DEM (2.59 m) versus the PCRaster DEM (12.5 m).

The SPANS3 and SPANS7 methods, which utilized the same resolution as the SPANS15 method but averaged slopes from smaller areas, had higher RMSEs than the SPANS15 method. This indicates that averaging slopes computed at higher resolution from relatively large neighborhood can lower the RMSE. Indeed, although not readily apparent from the method descriptions, the better performing methods such as GRASSrst and HIFI88 used elevations from larger areas to compute the high resolution DEMs and slopes. The GRASSrst method uses approximately 200 points to compute the interpolation function, its derivatives, and slope; the HIFI88 method initially computed a 10×10 m grid which was further refined to 4 and 1 m resolution DEMs. At the same time, the better performing approaches used various levels of smoothing when computing the DEM as well as averaged slopes (e.g., the SPANS15 and HIFI88 methods). While it may seem counterintuitive that smoothing and averaging data from larger areas result in more accurate slope estimates than a more local, less smoothed computation, this is entirely possible if slopes were negatively autocorrelated at small distances, and positively autocorrelated at larger distances. It requires only that one imagine a rough surface angled at a gradual incline. Over small distances, the surface would appear highly variable, but at larger distances, measured slopes would be very similar regardless of where they were measured. Some evidence of this pattern of variation was found for the field-measured slopes, which showed significant negative autocorrelation at the first lag distance (90 m), and positive autocorrelation at greater distances (Table 1). However, the minimum scale over which we could examine the spatial variation in slope was larger than the scale used to estimate slopes (e.g., minimum lag of 90 m versus estimation scales of < 30 m), thereby precluding an unequivocal interpretation of this observation.

4.3. Impact of method for slope computation

As explained in Section 2, two basic approaches for computation of slope were used. The methods which used the differential geometry approach performed better than the methods based on simple trigonometry. This result is in agreement with the previously published tests (Skidmore, 1989; Jones, 1998). Lower accuracy of the trigonometric slope estimates usually results from the fact that only eight directions of steepest slope are considered and only two points are used for the estimate; the differential geometry approach uses all nine points within the 3×3 elevation matrix and any direction of the steepest slope is possible.

4.4. Impact on soil erosion prediction

Slope is a critical parameter controlling most translocation processes within landscapes. When attempting to model such processes in a GIS environment, it is imperative that algorithms used to derive slope from DEMs represent actual slope as closely as possible. This is especially critical when attempting to model soil erosion, as slope has a superproportional influence on its estimation (Zingg, 1940; Nearing, 1997).

It is less critical for parameters such as runoff velocity which is subproportionally influenced by slope (Manning, 1889). With physically based erosion models, errors in slope calculation at any point on a hillslope will be propagated downhill, as the erosion processes in downslope grid cells are affected by inputs from all of the upslope contributing cells.

Typically, deviations between measured soil loss and modeled predictions are attributed to deficiencies in the models themselves. However, this study illustrates that parameter determination may play a significant role in modeling errors. The various methods used for slope computation produced a large scatter in soil erosion predictions (Fig. 4). This study used only a single, simple soil loss prediction model, but similar results could be expected from other models, as all incorporate slope as an important parameter. For example, EUROSEM is a dynamic single-event model with a completely different approach to erosion prediction (Morgan et al., 1998). Nevertheless, in cases where soil erosion is limited by rill transport, the influence of slope is raised to the power of 1.5, much like the USLE. Where soil erosion is limited by interrill transport, the effect of slope may be raised to the power of 10 or more, depending on Manning's roughness coefficient. In either case, the effects of errors in slope computation are exacerbated due to the superproportional consideration of slope in the calculation of soil erosion.

5. Conclusions and practical application

Because estimation methods may vary in performance under different terrain conditions or estimation scenarios, the results of our analyses are specific to the particular scales, parameters, and landscape examined. Therefore, we caution against broad speculation on the general performance of the various methods. Despite this caveat, evidence that the accuracy of slope estimates can vary among methods and that a portion of the variation is attributable to DEM resolution is still pertinent. Our findings suggest that greater attention be given to these potential sources of variation when using slopes estimated from DEMs in soil erosion prediction models.

Ultimately, the size and spatial distribution of the original elevation data limit the minimum precision of a DEM, which subsequently limits the minimum precision of the slope estimates. The most effective DEM resolution will ensure that precision is not lost due to choice of an overly large grid size as was the case in the PCRaster approach in this study. Similarly, a grid size that is too small may result in an estimate of slope variation at a much higher level of detail than is relevant for the process being modeled.

Probably the most informative result of our study is that the best slope estimates can be obtained by computing the DEM at high resolution, but with sufficient smoothing, using elevation data from larger areas. Computing slope at the given point as an average of values estimated in its neighborhood can also improve the results. Our study has confirmed that the differential geometry approach based on local polynomial approximation (using weighted averages of all points within the 3 × 3 neighborhood) provides better results than the trigonometric approach.

Acknowledgements

References

Benedetto, J.J., Ferreira, P.J.S.G. (Eds.), 2001. Modern Sampling Theory: Mathematics and Applications. Birkhäuser, Boston.

Bolstad, P.V., Stowe, T., 1994. An evaluation of DEM accuracy: elevation, slope and aspect. Photogramm. Eng. Remote Sensing 60, 1327–1332.

D'Agostino, R.B., Belanger, A., D'Agostino Jr., R.B., 1990. A suggestion for using powerful and informative tests of normality. Am. Stat. 44, 316–321.

Dutilleul, P., 1993. Modifying the *t*-test for assessing the correlation between two spatial processes. Biometrics 49, 305–314.

Ebner, H., Reinhardt, W., Hößler, R., 1987. Generation, management and utilization of high fidelity digital terrain models. Int. Arch. Photogramm. Remote Sens. 27, 556–566.

Environmental Systems Research Institute, 1991. Arc/Info 6.0. Redlands, California, USA.

Fiener, P., Auerswald, K., 2003. Effectiveness of grassed waterways in reducing runoff and sediment delivery from agricultural watersheds. J. Environ. Qual. 32, 927–936.

Florinsky, I.V., 1998. Accuracy of local topographic variables derived from digital elevation models. Int. J. Geogr. Inf. Sci. 12, 47–61.

Florinsky, I.V., 2002. Errors in signal processing in digital terrain modeling. Int. J. Geogr. Inf. Sci. 16, 475–501.

Gamma Design Software, 1994. GS+: Geostatistics for the Environmental Sciences, Version 2.3. Gamma Design Software. Plainwell, Michigan, USA.

Giles, P.T., Franklin, S.E., 1996. Comparison of derivative topographic surfaces of a DEM generated from stereoscopic SPOT images with field measurements. Photogramm. Eng. Remote Sensing 62, 1165–1171.

Hofierka, J., Parajka, J., Mitasova, H., Mitas, L., 2002. Modeling impact of terrain on precipitation using 3-D spatial interpolation. Trans, 135–150.

Horn, B.K.P., 1981. Hill shading and the reflectance map. Proc. Inst. Electric. Electron. Eng. 69, 14–47.

Jones, K.H., 1998. A comparison of algorithms used to compute hill slope as a property of the DEM. Comput. Geosci. 24, 315–323.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659–1673.

Manly, B.F.J., 1991. Randomization and Monte Carlo Methods in Biology. Chapman & Hall, New York, NY.

Manning, R., 1889. On the flow of water in open channels and pipes. Trans. Inst. Civ. Eng. Irel. 20, 161–207.

Mitasova, H., Mitas, L., 1993. Interpolation by regularized spline with tension: I. Theory and implementation. Math. Geol. 25, 641–655.

Mitasova, H., Mitas, L., Brown, W.M., Gerdes, D.P., Kosinovsky, I., Baker, T., 1995. Modelling spatially and temporally distributed phenomena: new methods and tools for GRASS GIS. Int. J. Geogr. Inf. Syst. 9, 433–446.

Morgan, R.P.C., Quinton, J.N., Smith, R.E., Govers, G., Poesen, J.W.A., Auerswald, K., Chisci, G., Torri, D., Styczen, M.E., 1998. The European soil erosion model (EUROSEM): a dynamic approach for predicting sediment transport from fields and small catchments. Earth Surf. Process. Landf. 23, 527–544.

Nearing, M.A., 1997. A single, continuous function for slope steepness influence on soil loss. Soil Sci. Soc. Am. J. 61, 917–919.

Neteler, M., Mitasova, H., 2002. Open Source GIS: A GRASS GIS Approach. Kluwer Academic Publishing, Boston.

Oden, N.L., 1984. Assessing the significance of a spatial correlogram. Geogr. Anal. 16, 1–16.

Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution 43, 223–225.

Skidmore, A.K., 1989. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. Int. J. Geogr. Inf. Sci. 3, 323–334.

Snyder, M.W., 1983. A comparison of four techniques for the calculation of slope and aspect from digital terrain matrices. M.S. Thesis. University of Illinois, Urbana. 85 pp.

Sokal, R.R., Oden, N.L., 1978. Spatial autocorrelation in biology: 1. Methodology. Biol. J. Linn. Soc. 10, 199–228.

Sokal, R.R., Rohlf, F.J., 1994. Biometry, 3rd edition. Freeman, San Francisco, CA, USA.

Srinivasan, R., Engel, B.A., 1991. Effect of slope prediction methods on slope and erosion estimates. Appl. Eng. Agric. 7, 779–783.

van Deursen, W.P.A., Wesseling, C.G., 1993. The PCRaster package. Dept. Physical Geography, Univ. Utrecht, The Netherlands.

Wischmeier, W.H., Smith, D.D., 1978. Predicting rainfall erosion losses—a guide to conservation planning. U.S. Dept. Agric., Agric. Handb. No. 537. Washington, DC.

Zingg, A.W., 1940. Degree and length of land slope as it affects soil loss in runoff. Agric. Eng. 21, 59–64.