The influence of elevation uncertainty on derivation of topographic indices

Hebeler, F; Purves, R S
The influence of elevation uncertainty on derivation of topographic indices

Abstract

Digital elevation models at a variety of resolutions are increasingly being used in geomorphology, for example in comparing the hypsometric properties of multiple catchments. A considerable body of research has investigated the sensitivity of topographic indices to resolution and algorithms, but little work has been done to address the impact of DEM uncertainty and elevation value error on derived products. By using higher resolution data from the Shuttle Radar Topography Mission - of supposed higher accuracy - for comparison with the widely used GLOBE 1km data set, error surfaces for three mountainous regions were calculated. Correlation analysis showed that error surfaces related to a range of topographic variables for all three regions, namely roughness, minimum and mean extremity and aspect. This correlation of error with local topography was used to develop a model of uncertainty including a stochastic component, permitting Monte Carlo Simulations. These suggest that global statistics for a range of topographic indices are robust to the introduction of uncertainty. However, the derivation of watersheds and related statistics per watershed (e.g. hypsometry) is shown to vary significantly as a result of the introduced uncertainty.
The influence of elevation uncertainty on derivation of topographic indices

Felix Hebeler and Ross S. Purves

University of Zurich, Department of Geography
Winterthurer Str. 190, 8057 Zurich
Switzerland

Geomorpholgy, accepted for publication June 21, 2007

Abstract

Digital elevation models at a variety of resolutions are increasingly being used in geomorphology, for example in comparing the hypsometric properties of multiple catchments. A considerable body of research has investigated the sensitivity of topographic indices to resolution and algorithms, but little work has been done to address the impact of DEM uncertainty and elevation value error on derived products. By using higher resolution data from the Shuttle Radar Topography Mission - of supposed higher accuracy - for comparison with the widely used GLOBE 1km data set, error surfaces for three mountainous regions were calculated. Correlation analysis showed that error surfaces related to a range of topographic variables for all three regions, namely roughness, minimum and mean extremity and aspect. This correlation of error with local topography was used to develop a model of uncertainty including a stochastic component, permitting Monte Carlo Simulations. These suggest that global statistics for a range of topographic indices are robust to the introduction of uncertainty. However, the derivation of watersheds and related statistics per watershed (e.g. hypsometry) is shown to vary significantly as a result of the introduced uncertainty.

1 Introduction

1.1 Digital elevation models

Representing the face of the earth as a continuous surface is a key task in research characterising the form of the earth's surface, modelling processes occurring on this surface or developing a better understanding of links explaining the relationship between process and form. Digital models of terrain can generally be categorised as either regular or irregular tessellations of point data, sometimes with additional ancillary information representing structural features such as breaks in slope or drainage divides (Weibel and Heller, 1991; Burrough and McDonnel, 1998). Since these models are usually taken to represent a continuous differentiable surface on which values of elevation and its derivatives (for example, slope and curvature) can be calculated at any point, some implicit or explicit scheme for interpolating or approximating values is required. Examples include linear
interpolation, quadratic patches, thin-plate splines, Bezier splines and Coons patches. Within the GIScience community much attention has focussed on the veracity of the representation gained through different approaches (e.g. Schneider, 1998; Wood, 1998; Wise, 1998, 2000; Hutchinson and Gallant, 2000; Hugentobler et al., 2004; Chaplot et al., 2006); however, the wide availability of data which are regularly tesselated and the resulting ease of computation means that most research in the field of geomorphology using terrain data has been carried out on regular tesselations commonly known as Digital Elevation Models (DEMs), with some significant exceptions, (e.g. Tucker et al., 2001). In recent years the availability, resolution and coverage of digital data representing the surface of the earth has rapidly increased (Jarvis et al., 2004). So-called seamless datasets representing large parts of the earth’s surface are available at resolutions ranging from approximately 1 km (GLOBE/GTOPO30) (USGS, 1996; GLOBE Task Team & others, 1999) down to 90 m and (for the contiguous USA) 30 m (SRTM, Rodriguez et al., 2005). Higher resolution data are collected by national mapping agencies in many countries as a component of national topographic databases, generally with resolutions of 10 - 50 m. Finally, in recent years very high resolution datasets consisting of LIDAR data with nominal resolutions of the order of meters have become available for specific locations (e.g. Staley et al., 2006). This increase in data availability and relative ease with which a wide range of topographic parameters may be derived has resulted in a steady increase in research within geomorphology which directly applies DEMs (Fig. 1). DEMs are, for instance, commonly used for extraction of a variety of forms (Herrington and Pellegrini, 2000; Fisher et al., 2004), fluvial geomorphic analysis at scales ranging from small catchments, through regional analysis to continental scales (Finlayson and Montgomery, 2003; Oksanen and Sarjakoski, 2005a), the modelling and validation of surface process models of long-term landscape evolution (Codilean et al., 2006) and as input data for ice sheet modelling at a range of scales (Bamber and Bindschadler, 1997; Lythe et al., 2001).

A relatively small number of topographic indices form the core of most descriptive use of DEMs within geomorphology. These include basic descriptive statistics of elevation (e.g. maximum, mean, minimum, standard deviation and hypsometry), local relief, gradient and aspect, curvature and

Figure 1: Number of papers mentioning “DEM” or “DTM” in their title, abstract or keywords in the journal Geomorphology, between Jan. 1994 and Dec. 2006 (Source: Web of Science Index)
hydrological products such as flow direction, flow accumulation and watershed boundaries. As an input to numerical models, DEMs form a set of initial and boundary conditions for which numerical solutions to a set of driving equations are derived. For example, in ice sheet modelling elevation firstly determines locations at which positive mass balance will allow accumulation (Pollard, 1983; Marshall, 2002) and secondly surface slope and ice sheet thickness combine to allow derivation of the driving stress of the ice sheet through, for example, the Shallow Ice Approximation (Nye, 1957). It is therefore clear that DEMs, and topographic indices derived from them, are an increasingly important tool for both analysing forms and modelling processes in geomorphology.

1.2 Error and uncertainty in DEM

DEMs can be derived from a variety of sources requiring different processing methods, including the digitisation of contour maps, interpolation from spot height measurements collected in the field and processing of radar or laser measurement data. As such, any resultant DEM is subject to both the precision and accuracy of the measurement sensor together with the quality of digitising or interpolation methods (Heuvelink, 1998), and thus is subject to error, from these multiple sources. In this paper the term error implies the deviation of a measurement from its true value, implying that the elevation error of a DEM can only be determined if a set of more accurate reference data is available (Fisher and Tate, 2006). Error is therefore implicitly associated with any DEM, but usually both the magnitude and spatial distribution of the error at any particular location are unknown. Error thus creates uncertainty, which can be approximated through the use of geostatistics or uncertainty models (Holmes et al., 2000). Here, the term uncertainty is used where a (modelled) value is expected to deviate from its true value, but it is uncertain to what extent, and is often associated with confidence intervals (Kyriakidis et al., 1999; Endreny and Wood, 2001; Shortridge, 2001). Error and accuracy of DEM products and production methods have been the subject of much research and methods to both describe and reduce error have been developed (Wechsler, 2006). Different types of error are often listed, with Wise (2000) describing blunders, systematic and random errors as being typical in DEMs. Blunders are gross errors which occur less frequently in DEM products and can be the results of failing measuring equipment or digitising errors. Systematic errors show a common trend or dependency, and can be the results of processing or recording procedures, such as radar shadow effects (Shortridge, 2006) and terracing in poorly interpolated contour-derived DEMs (Wood, 1996). When known, systematic error can often be computationally reduced or eliminated by, for example, detrending data. Random error originates from a variety of sources, and no trend can be observed. Fisher and Tate (2006) summarise the three main sources of such DEM error as:

1. measurement and generation of source data;
2. data processing and DEM generation from source data; and
3. the properties of the terrain surface being modelled with respect to its representation in a DEM.

The third of these sources of error is particularly important since it emphasises a fundamental and often neglected consideration when working with DEMs: the resolution and representation of a terrain surface and the derivation of products from that terrain surface in themselves introduce ambiguities and can thus best be categorised as uncertainty (Schneider, 2001; Fisher and Tate,
Error surfaces can be generated for DEMs by using data which are assumed to have a higher accuracy. Subtracting this higher accuracy data from the DEM under study creates a surface that can be analysed to better understand the nature and sources of error. In such an analysis three components of error can be identified that are important for geostatistical modelling. The first is a random component, where no spatial dependency can be resolved (at the DEM’s resolution). This component is also known as noise or nugget in semivariograms. The second component is related to the fact that terrain attributes (e.g. elevation, slope and roughness) typically change gradually over space, and this gradual change may also be reflected in the associated error surface. Error surfaces often exhibit characteristic ranges of spatial correlation, resulting in error patterns that stem from correlations with the underlying terrain (Hunter and Goodchild, 1997). A third component may be global trends superimposed on local patterns. These can often be detected using directional variograms, or profile plots (Liu et al., 1999; Holmes et al., 2000). Despite a recognition of the necessity for detailed error models and the suggestion that these be distributed with digital elevation data (Ehlschlaeger and Goodchild, 1994; Fisher, 1998), DEMs are still commonly distributed with at best global error or accuracy figures (Fisher, 1998), usually stating root mean squared error (RMSE) or standard deviation for vertical and horizontal accuracy (e.g. GLOBE, SRTM and most data from higher resolution DEMs provided by national mapping agencies). These global accuracy figures are of limited use since they contain no information on the spatial distribution of error, which, as discussed above, is often spatially correlated with topographic attributes such as altitude, slope or roughness (Holmes et al., 2000; Oksanen and Sarjakoski, 2005b). In order to geostatistically model the uncertainty inherent in DEMs, all three of the components of error observed in error surfaces should be taken into account.

Where no spatial dependency of error can be identified, e.g. because of the lack of higher accuracy reference data, assumptions about the spatial correlation of error must be made, or all of the error must be modelled as random noise. When modelling error using the RMSE supplied by the producers, it is often assumed to be normally distributed (Fisher, 1998; Oksanen and Sarjakoski, 2006; Wechsler, 2006). However, it has been suggested that this assumption is not generally valid (Holmes et al., 2000) and that more complex distributions of random errors should be modelled. When information on the spatial autocorrelation of the error is known, it can be incorporated into an uncertainty model (Wechsler, 1999). Common methods include the use of global and local spatial correlation measures such as Moran’s I that have been used, for example, in measuring the effect of random swapping of values (Fisher, 1998), or spatial moving averages that include correlation measures from variograms (Kyriakidis et al., 1999; Oksanen and Sarjakoski, 2006). While these methods introduce spatial correlation to the modelled uncertainty surfaces, correlation with the properties of the underlying surface is not explicitly accounted for.

### 1.3 Aims

The problem of uncertainty in results derived from digital elevation models has long been recognised and addressed in a variety of work. A number of experiments examine the robustness of descriptive indices of geomorphometric measures, such as slope and aspect (Evans, 1980; Burrough and McDonnel, 1998; Hodgson, 1998; Jones, 1998; Zhang et al., 1999), hypsometry (Strahler, 1952; Hurtrez et al., 1999), and hydrological catchment areas (Walker and Willgoose, 1999; Gallant and Wilson, 2000). While these studies generally focus on the effects of data models, resolutions and algorithms, the uncertainty stemming from elevation data itself, and the propagation and impact
on model results of this uncertainty compared to that introduced through data processing and model calculations has been largely ignored (Fisher and Tate, 2006). Despite their widespread use, there are few studies examining the accuracy of either GTOPO30 or GLOBE DEMs (Harding et al., 1999) beyond the global figures provided by the data producers (GLOBE Task Team & others, 1999), but with the availability of data from the Shuttle Radar Topography Mission (SRTM), a number of studies have looked at the uncertainties inherent in different versions of SRTM data products. Accuracy studies have focused on technical issues of data acquisition (Heipke et al., 2002) or on comparison of SRTM data with higher resolution data such as spot height measurements, ICESat or LIDAR data (Sun et al., 2003; Carabajal and Harding, 2005). SRTM data appear to be sensitive to overestimating terrain height in densely vegetated areas due to scattered first returns from canopies (Shortridge, 2006). Incomplete data for a variety of reasons can lead to uncertainties resulting from the interpolation of “data holes” (Jarvis et al., 2004). However, in general SRTM data have a much higher accuracy than GLOBE DEM, and it has been implied that SRTM data may be used as ground truth for accuracy studies of GLOBE DEMs (Jarvis et al., 2004), when the focus is on generating a valid large-scale error model that can be generalised, rather than detailed analysis of regional uncertainties.

While SRTM data are available for large regions of the Earth’s surface, for a number of reasons lower resolution DEM data (of supposed lower accuracy) is still being used in many applications. Large scale environmental models, such as global climate models and ice sheet models, run on resolutions of 1 to 20 km, and the use of higher resolution data is not sensible or even unpractical because of the demands on computational and memory capacity. Furthermore SRTM data are not available at latitudes above 60°N: models focussing on higher latitudes, such as permafrost, snowcover and ice sheet models are therefore still dependent on lower resolution data.

This paper therefore has two key aims: firstly, to develop a robust model of the error and/or uncertainty in GLOBE data and its relationship (if any) to the underlying terrain surface, and secondly, to illustrate the potential uses of this uncertainty model in geomorphometry through a simple case study. We first set out methods for deriving error surfaces of GLOBE DEM using SRTM as ground truth in regions where both data sets are available. These error surfaces are then used in the analysis of possible dependencies of GLOBE uncertainty on terrain properties. Where robust and generalisable dependencies are derived, this information can be incorporated using regression modelling to approximate error or uncertainty of GLOBE data for regions where no SRTM data is available, and to assess the impacts of uncertainty on previously completed studies. Finally, a case study assesses the impact of the modelled uncertainty on a standard set of geomorphometric analyses through Monte Carlo simulation.

2 Developing an uncertainty model

The development of a generalisable model of uncertainty in GLOBE elevation values was carried out through the following steps:

- Resampling and registration of higher accuracy data to allow a comparison with GLOBE data.
- Qualitative exploration of the variation in error values.
- Exploration of correlations of terrain parameters with error.
Table 1: Descriptive statistics of the GLOBE data used for the three study areas. (StDev = standard deviation; Skew = skewness, Kurt = kurtosis.

- Development of a model of terrain uncertainty.
- Generation of multiple uncertainty surfaces including a stochastic component.

2.1 Study areas and data sets

This work emerged from the need for a uncertainty model for GLOBE DEM data for Fennoscandia, to be used as input in Monte Carlo Simulations (MCS) for sensitivity testing of ice sheet model results (Hebeler and Purves, 2004). GLOBE data are produced using a variety of data sources and techniques, with the major contributors being USGS DEM data for the US, and digital terrain elevation (DTED) data and digital chart of the world (DCW) data for much of the rest of the northern hemisphere. While these sources are essentially the same as used for GTOPO30, the GLOBE data set has been refined over GTOPO30 by using higher accuracy data, where available (Hastings and Dunbar, 1998). For example, GLOBE has been refined using the local DEM of Italy provided by the Servizio Geologico Nazionale, which showed a slight improvement over DTED data for high altitudes, but much better agreement with actual terrain for lower elevations (GLOBE Task Team & others, 1999). The accuracy for GLOBE data sourced from DTED data is reported in meta-data as being around 18 m RMSE (USGS, 1996), decreasing to around 97 m RMSE for DCW data. For other areas, especially in South America much larger RMSE values have been reported. Fig. 2 shows the variation in source data for GTOPO30 data in different regions and an error surface calculated for South America showing the strong correlation of error magnitudes with different data sources. As the aim of this study was to develop and test a robust uncertainty model that is of general use for lower resolution data over large areas, GLOBE DEM data were selected over GTOPO30 because of their supposed higher accuracy and more recent production. Given this dependency of error on source data, study areas were chosen for their comparability with Fennoscandia - that is to say regions where GLOBE data were mostly derived from DTED data, with mountainous topography and spanning a similar range of altitudes to those found in Fennoscandia (between sea level and around 2500 m). The study areas selected are shown in Fig. 2 and include the European Alps, the Pyrenees and Turkey. Table 1 shows the descriptive statistics for these regions.

2.2 Data preparation

For our analysis, SRTM3 data provided by CGIAR-CIAT (Jarvis et al., 2006) were used. This data resembles the post-processed NASA SRTM data with voids filled using generated contours and auxiliary data, where available (Jarvis et al., 2004). In order to compare these SRTM data at a resolution of ~100 m with GLOBE at ~1 km, SRTM data were downscaled to fit the extent and resolution of the corresponding GLOBE data. While for some operations such as gradient calculation, it is necessary to work with projected data, any interpolation applied in projecting
or resampling data alters the original data and introduces additional uncertainty (Montgomery, 2001). Our experiments indicated that error dependencies on terrain attributes varied relatively little between projected and unprojected data. Therefore bearing in mind that the values of some attributes calculated in the unprojected data were misleading, we carried out error analysis using the unprojected WGS84 spatial reference in which both SRTM and GLOBE data are distributed. To allow the derivation of error surfaces SRTM data were downsampled following the approach of Jarvis et al. (2004) by calculating the mean of all SRTM cells within the bounds of each GLOBE data cell. This approach is based on the assumption that GLOBE data represents the average altitude within each corresponding DEM cell, while SRTM data represents the maximum height at each posting, thus averaging of SRTM data should be used for downsampling when comparing with GLOBE. Both SRTM and GLOBE DEMs were then clipped to eliminate waterbodies before the calculation of error surfaces. Finally an error surface representing the deviation of GLOBE from the averaged SRTM data set was calculated by subtraction. For the purposes of our analysis, the higher accuracy SRTM data are assumed to be error free and thus a difference surface is assumed to completely describe the error in GLOBE data.

### 2.3 Error surfaces

Fig. 3 shows an error surface calculated for the Alps. Through visual examination a number of features become apparent. Most strikingly, it is clear that error values are strongly spatially correlated, following the boundaries of prominent features within the data set. In areas of low
relief, relatively small errors of $\pm 20$ m with a short wavelength of variation are apparent. Finally, an observation can be made regarding the dependency of error with DEM source, as the magnitude of error appears to be less in the Italian part of the Alps - where different source data are used in the production of the DEM - than for the French or Swiss regions. This first examination of the error surface suggests a correlation of error with terrain features and attributes, but also indicates inhomogeneity in the error values and spatial correlation depending on both terrain type (high versus low relief) and data sources. Fig. 4 shows the semivariogram map for the error surface derived for the Alps. A general SW-NE trend is visible with high spatial autocorrelation indicated by low variance within a range of around three cells.

### 2.4 Correlating error with terrain parameters

Having carried out a visual analysis of error surfaces and considered previous work which has examined errors in terrain values at a range of scales, we could commence development of a model of error. The first step in developing this model was to examine the correlation of error with a range of terrain parameters. Table 2 shows the complete set of terrain parameters tested for correlation with GLOBE error. In practice, given the very large number of data points, all terrain parameters were statistically significantly correlated with the magnitude of error. In order to develop a useful error model, factor analysis was carried out to reduce the number of potential variables in the error model. Results from the factor analysis and the breakdown of intercorrelation of the variables suggested the variation in derived error to be best reproduced by two or three terrain parameters. Table 3 shows the correlation of a range of terrain parameters with the magnitude of error for the different study regions. Although parameters exhibited highest correlation coefficients with the
magnitude of error, a relationship between aspect and the sign of the error is also visible in the error surfaces (Fig. 3) and was also used in the development of the uncertainty model described in the following section.
### Attributes
- **Altitude**: Value of GLOBE cell
- **Error, absolute Error**: (absolute) Deviation of GLOBE from mean SRTM value

### Derivatives
- **Aspect**: Direction of first derivative of elevation
- **Slope**: Magnitude of first derivative of elevation
- **Plan Curvature**: 2nd derivative orthogonal to direction of steepest slope
- **Profile Curvature**: 2nd derivative in direction of steepest slope
- **Total Curvature**: Compound curvature index

### Indices
- **Maximum/ Mean/ Minimum Extremity**: Deviation of center cell from the maximum/ mean/ minimum of its 3x3 neighbourhood
- **Roughness (Altitude)**: Standard deviation of altitude in a 3x3 neighbourhood
- **Roughness (Slope)**: Standard deviation of slope in a 3x3 neighbourhood

Table 2: Attributes and indices calculated and tested for correlation with GLOBE error. Extremity calculated following Carlisle (2000).

### Table 3: Correlation coefficients of calculated parameters with GLOBE error and magnitude of error (AbsError) for the three study sites.
Table 4: Constants used in the regression of error (a,b,c), the regression of the residuals (d,e) and the binary logistic regression of the error sign (f,g,h).

<table>
<thead>
<tr>
<th>Constant</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>roughness of altitude (magnitude of error)</td>
</tr>
<tr>
<td>b</td>
<td>minimum extremity (magnitude of error)</td>
</tr>
<tr>
<td>c</td>
<td>constant (magnitude of error)</td>
</tr>
<tr>
<td>d</td>
<td>magnitude of error (residual regression)</td>
</tr>
<tr>
<td>e</td>
<td>constant (residual regression)</td>
</tr>
<tr>
<td>f</td>
<td>mean extremity (sign of error)</td>
</tr>
<tr>
<td>g</td>
<td>aspect (sign of error)</td>
</tr>
<tr>
<td>h</td>
<td>constant (sign of error)</td>
</tr>
</tbody>
</table>

Table 5: Example of fitting the regression model for the Alps study area. The dependent variable is error magnitude; $r^2$ values and coefficients for different combinations of independent variables shown. Coefficients according to Table 4 where not explicitly given.

<table>
<thead>
<tr>
<th>Variables</th>
<th>$r^2$</th>
<th>Coefficients</th>
<th>Constant c</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.453</td>
<td>0.753\times a</td>
<td>0.034\times b</td>
</tr>
<tr>
<td>3</td>
<td>0.444</td>
<td>0.444\times a</td>
<td>0.024\times b</td>
</tr>
<tr>
<td>2</td>
<td>0.441</td>
<td>0.468\times a</td>
<td>0.033\times b</td>
</tr>
<tr>
<td>1</td>
<td>0.437</td>
<td>0.515\times a</td>
<td>7.388</td>
</tr>
</tbody>
</table>

2.5 Building an uncertainty model

A simple linear regression model of error with terrain parameters was developed through the stepwise substitution of different combinations of parameters. Table 4 lists the parameters selected for the three regressions used in the uncertainty model. Table 5 shows an exemplary model fitting process with different combinations of parameters, and illustrates that most of the variation in the surface is explained by two terrain parameters, roughness and minimum extremity, with only a slight improvement with the use of three or four parameters. It was decided to include the second parameter into the regression, which improved $r^2$ by around 5-10 %, despite some intercorrelation of roughness and minimum extremity. By including a logistic regression the sign of the error was also simulated, using aspect and mean extremity. Both variables are uncorrelated, and it was found that correlation of sign of error with aspect could be slightly improved when transformed using a sine function. However, across the three test areas, no global transfer function could be found that resulted in an improvement for all three data sets. If the regression analysis had shown that the error surface was completely correlated with terrain parameters derived from the GLOBE data, it would have been possible to eliminate this error from the GLOBE data, and thus potentially reduce the uncertainty of terrain parameters derived from the GLOBE data. However, as shown in Table 6, whilst regression analysis showed strong correlations with the error surface, a significant quantity of the error is not explained. Thus, in order to model the uncertainty surface, stochastic elements were included, providing a suitable input for Monte Carlo simulation as every uncertainty surface produced will be different. In developing the model we used three qualitative and quantitative observations from the development of the error surfaces described above:

- Error surfaces are spatially autocorrelated with a range of approximately three grid cells.
Error magnitude can be described through terrain parameters, with all terrain parameters significantly correlated, and factor analysis suggesting that two to three parameters should adequately describe most of the variation.

The sign of the error can be approximated from terrain parameters using a logistic regression. The final error model consists of three parts, and took the following form. Regression of the modelled magnitude of error:

\[ \text{abs}(\varepsilon') = a \times \text{roughness} + b \times \text{extremity}_{\text{min}} + c \]  

where \( \text{abs}(\varepsilon') \) is the magnitude of the derived error and \( a, b, \) and \( c \) are constants given in Table 4. Fig. 5 shows the residuals of this regression for the Alps. These residuals are correlated with the magnitude of the calculated error and can be described by the following linear equation:

\[ \text{res}_1 = d \times \text{abs}(\varepsilon) + e \]  

where \( \text{res}_1 \) is the residual of Eq. 1, and \( d \) and \( e \) are constants (Table 4). The residuals \( (\text{res}_2) \) of Eq. 2 are essentially random and can best be simulated by a transformed normal distribution \( \text{res}_2 = N(0,1) \Rightarrow \mathcal{D}(\mu = 0, \sigma = 30, \gamma_1 = 2, \gamma_2 = 6.8) \) which are then randomly added to \( \text{res}_1 \), introducing a first stochastic component to the model.

Finally, the sign \( s \) of the error is calculated as follows:

\[ s = f \times \text{extremity}_{\text{mean}} + g \times \text{aspect} + h \]  

where \(-1 \leq s \leq 1\) and \( f, g \) and \( h \) are constants (Table 4). A random number \( r \) is then selected, where \( 0 \leq r \leq 1 \), and applied to the following equation

\[ s' = \begin{cases} -1 & \text{if } r \leq |s| \text{ and } s < 0 \quad \text{or} \\ 1 & \text{if } r \leq |s| \text{ and } s \geq 0 \quad \text{or} \\ -1 & \text{if } r > |s| \text{ and } s \geq 0 \quad \text{or} \\ 1 & \text{if } r > |s| \text{ and } s < 0 \end{cases} \]  

(4)

to introduce a second, constrained random component to the model. The value of uncertainty in elevation at each cell \( U_{\text{tot}} \) is thus derived through the following equation:

\[ U_{\text{tot}} = (\text{abs}(\varepsilon) + \text{res}_1 + \text{res}_2) \times s' \]  

(5)

Table 6: Results of regression of the magnitude of error, the sign of error and the residuals of the error, with their corresponding constants (Table 4) for all study areas.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>magnitude</th>
<th>sign</th>
<th>residual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r^2 )</td>
<td>( a )</td>
<td>( b )</td>
</tr>
<tr>
<td>Alps</td>
<td>0.441</td>
<td>0.468</td>
<td>0.033</td>
</tr>
<tr>
<td>Pyrenees</td>
<td>0.405</td>
<td>0.617</td>
<td>0.032</td>
</tr>
<tr>
<td>Turkey</td>
<td>0.423</td>
<td>0.502</td>
<td>0.028</td>
</tr>
<tr>
<td>Global</td>
<td>0.512</td>
<td>0.032</td>
<td>7.62</td>
</tr>
</tbody>
</table>
Figure 5: Plot of the residuals $res_1$ from the regression of error magnitude with roughness and $extremity_{min}$ against the magnitude of error, showing a good fit with a single factor linear regression.

Figure 6: Error surface derived from GLOBE data for an approximately 200x300 km subset of the Alps study area (top left) together with five example uncertainty surfaces for the same area. While the main landscape features are detectable in both the error surface and all five uncertainty surfaces, the stochastic elements in the uncertainty surface, and the globally applied (low) omnidirectional spatial correlation result in a noisier structure of the uncertainty surfaces, when compared to the GLOBE error.

In order to take account of the spatial autocorrelation of the error surface shown in Fig. 3, the values of the error surface derived are transformed to a normal distribution and a convolution filter (Oksanen and Sarjakoski, 2005a) using a range derived from Fig. 4 is applied, in order to adjust spatial correlation of the modelled uncertainty to that of the original measured error, before the uncertainty surface is finally transformed back to its original distribution.

By adding the uncertainty surface to the original GLOBE data it is possible to examine the influence of this uncertainty on derived terrain parameters. Fig. 6 shows a set of five example uncertainty surfaces for a subset of the Alps together with the derived error surface. Distributions of modelled error against derived error for all three study regions are shown in Fig. 7.
3 Example Application

Having developed an uncertainty model, the next obvious step is to apply the model and determine whether it has any significant influence when deriving common topographic indices using GLOBE DEM data. In this case study we choose two approaches to examining the impact of uncertainty through the use of Monte Carlo Simulation (MCS). In the first experiment we derive a set of global parameters for the GLOBE DEM with 40 realisations of the uncertainty model added. In a second set of experiments the boundaries of individual watersheds were first derived before geomorphological indices per watershed were calculated for comparison. Finally, we examine in detail the differences within and between two large catchments defined through the MCS methods. For this example, all analysis was carried out on the Alps region shown in Fig. 3.

3.1 Methodology

A set of 40 uncertainty surfaces was created using the regressions described in section 2.5 and added to the original GLOBE DEM. As discussed earlier, to calculate meaningful topographic parameters these DEMs were projected, using bilinear interpolation, into an Albers Equal Area projection with the central meridian running through the centre of the DEM and two standard
Areal unit | Parameters calculated
--- | ---
per DEM | elevation, slope, number of watersheds, watershed delineation, Strahler order, stream length, hypsometric integral
per Strahler order | elevation, slope, stream length, relative area
per watershed | elevation, slope, stream length, area, hypsometric integral
per cell | watershed membership likelihood

Table 7: Geomorphological parameters calculated for the Alps using Monte Carlo Simulation.

parallels dividing the area into even thirds, thus minimising distortion. WGS84 was used as the reference ellipsoid and all further calculations were carried out using the projected DEM. To calculate meaningful hydrological indices, all sinks in the resulting DEMs were filled using ArcGIS. TARDEM (Tarboton, 1999) was used to calculate a set of parameters for every realisation (see Table 7). To delineate watersheds a pour point must be defined indicating a point on a stream network which accumulates all upstream flow. This process was automated by defining pour points to be the first cell encountered, of a given stream order, when travelling down the stream network of a given realisation, where first order streams are defined as those with no other cells draining into them and stream order is increased when two streams of equal order meet (Tarboton, 1999). Cells were then assigned a Strahler order based on the lowest order basin in which they were contained.

3.2 Results

3.2.1 Summary global statistics

Global statistics for the 40 different realisations of the Alps, as generated by Monte Carlo Simulation are shown in Fig. 8. The box and whisker plots divide the data into four quartiles with the median data value being displayed in the middle of the box plot (if the data are not skewed), the first and third quartiles the ends of the box and the minimum and maximum values displayed as the ends of the whiskers.

Fig. 8A,E show that global statistics of altitude and the distribution of altitude (shown through hypsometry) change very little for any of the MCS surfaces generated. However, some variation, particularly in the upper quantile, is visible in the global statistics of slope (Fig. 8B). Fig. 8C,D,F also show global statistics calculated per Strahler order as calculated for the network derived from the MCS realisations. Once again little variation is visible in elevation, but it is now revealed that most variation in slope occurs in cells which are assigned to the 1st order of the Strahler network or the headwaters of streams. Fig. 8F suggests that the number of cells assigned to each order of the Strahler network is relatively robust for the MCS.

3.2.2 Watershed derivation and derived statistics

The second set of experiments examined the robustness of watershed boundaries as a function of the uncertainty in elevation values in GLOBE. For these experiments, 6th order watersheds were delineated, where a watershed’s pour point was defined as the first downstream point belonging to a 7th order stream. Thus each watershed contains cells with Strahler orders of 1 to 6. Fig. 9A shows the total number of watersheds derived, with all other results in Fig. 9B-F sorted according to this total number of watersheds. The notches in the center of the box plots indicate an estimate of the
uncertainty around the medians. Boxes whose notches do not overlap indicate that the medians differ with a confidence of $p = 0.05$ (MATLAB, 2006). By examining these notched box plots it is clear, firstly, that most of the derived statistics are robust for all realisations. However, the box plots for slope show that on some occasions the median slope values are statistically significantly different (for example the 4th and 19th realisations of slope in Fig. 9B).

3.2.3 Comparison of two watersheds

Fig. 10 shows the membership likelihood of DEM cells belonging to one of two large, neighbouring watersheds (defined upstream of 7th order streams) when different uncertainty surfaces are applied. The variation in area is large with a variation of over 320% and 290% for W1 and W2, respectively. For both W1 and W2, most variation in watershed area is the result of the stream network capturing large areas of “flatlands”, though in the case of W1 some instability is also vis-
Figure 9: Watershed statistics for 40 MCS realisations using the uncertainty model. Number of 6th order watersheds (A), sorted by size. Absolute number (solid black) of delineated 6th order watersheds are plotted together with the number of watersheds with an area smaller than 2000 cells (solid red) and number of watersheds with areas larger than 5000 cells (solid blue). Mean across all 40 DEMs plotted in dotted light gray. Mean watershed slope (B), mean area (C), mean elevation (D) and mean maximum stream length (E, as given by plen using TARDEM) plotted across all 40 DEMs using box plots. For better visibility outliers are not drawn.

ible in the mountainous part of the catchment. Both catchments agree well with national borders, which in this region lie along the main ridge of the Alps, demonstrating the stability of catchment boundary definition in areas of high relief. Derivation of the elevation and slope statistics for the different watershed realisations (Fig. 11) show considerably more variability than that for the global statistics, since the region over which these statistics is calculated has itself considerable variation. Although more variation in catchment area was shown for W1, both elevation and slope vary more for W2.
Figure 10: Frequency of cells belonging to two selected watersheds W1 (left) and W2 (right) across all 40 MCS runs (top). Most variation in watershed size is detectable in the lower regions of the catchment area, but some variation is also evident in the high mountain regions of both watersheds. Hypsometric curve (bottom) across 40 MCS runs for the two selected watersheds, showing a considerable amount of variation in form, depending on the size of lowland area ‘captured’ by a MCS simulation run.
4 Discussion

4.1 Uncertainty model

4.1.1 Data quality

On deriving the initial error surface for the Alps, a number of patterns presented themselves on a first visual inspection. Error appeared to be correlated with the underlying terrain, with large error values following the main ridge of the Alps. Error also appeared to be correlated with aspect, as different error signs appeared to be distributed on the either side of the major valleys. Comparable patterns were found for the other two test areas. This hints at luff side overestimation and lee side underestimation of elevation values by radar sensors, suggesting that our assumption of error free SRTM data may have been wrong. An equally likely cause would be a systematic misregistration between GLOBE and SRTM data. Tests confirmed that correlation of GLOBE and SRTM improves when their locations are systematically shifted, and consequently standard deviation of derived error decreases. However, the necessary shift is inconsistent across the three test areas and even across different subsets of the same test area, and the improvement is only minimal. In the Alps, the differences in GLOBE DEM production from DTED - median values from USGS/GTOPO in France versus spot sampling in the southwestern corner of the GLOBE
cell for the rest of the Alps excluding the Italian part (GLOBE Task Team & others, 1999) - become apparent as they result in different shifts of the data subsets in relation to SRTM. This heterogeneity, amongst other data quality issues, is likely to be the reason why error distributions showed only weak correlation with aspect when analysed, due to a significant scattering of the values, despite obvious visual dependency of error with aspect (Fig. 3). Furthermore, although correlation can be improved when using sinus transfer functions for aspect, the necessary functions vary considerably for the three areas, and global values that improve aspect-error correlation for all data sets could not be found. This again suggests a possible misregistration of GLOBE DEM data during the combination of different local sources. However, by using aspect only to predict the sign of error values, it proved to be a valuable parameter in locally autocorrelating error surfaces across areas of similar aspect.

Taking the above into consideration, it is unclear what causes the observable SW - NE trend in the spatial configuration of autocorrelation depicted in Fig. 4. While it corresponds with the general flight path of the Shuttle Radar Topography Mission (Eineder et al., 2001), Guth (2006) reported diamond shaped patterns for error encountered in 1° SRTM data for the United States. At this point it is difficult to determine whether SRTM error is large enough to influence the regression analysis, as the overall vertical accuracy of SRTM for the selected study areas is reported to be less than 5 m (Rodriguez et al., 2005) which is much less than the reported RMSE values for GLOBE. By using CGIAR STRM data, voids in the original SRTM data have been filled, and interpolated values have thus been integrated in the regression analysis. Especially for the Alps data set, voids are common at very high altitudes. However, as these voids have been carefully filled and the respective altitude values are less important for the Fennoscandian study area, void-filled DEM data have been used for ease of processing and analysis. The use of SRTM30 for comparison with GLOBE would have been an alternative to the use of SRTM3, which would have avoided the problem of void-filling and resampling. Tests have shown that SRTM30-derived error surface show similar, yet slightly smaller correlations. Furthermore, a key aim in our experiments was to start with a dataset which had undergone as little preprocessing as possible and thus could be considered to be “ground truth” and SRTM3 was judged to be optimal for this task.

A further effect visible in the error surface for the Alps, is the pronounced change in error magnitude visible along the Italian border (Fig. 3), which confirms the suggested better agreement of the data source used for the compilation of the Italian part of the GLOBE data (Hastings and Dunbar, 1998) with SRTM data. This fact is not explicitly accounted for in the uncertainty model but was discussed as a possible error source in section 1.2.

4.1.2 Quality of the uncertainty model

While analysis suggests an overestimation of uncertainty by our model for the Italian part of the Alps, it suggests that our method of deriving error from averaged SRTM data is valid: the better accordance with terrain characteristics of the Italian data at 1km resolution is well captured by the derived error surfaces. At the same time, the global uncertainty model fits all three test areas equally well (Table. 6), suggesting the impact of this data heterogeneity to be tolerable. The correlation coefficients were broadly similar for the three datasets (Alps, Pyrenees and Turkey) over which they were derived and, based on this result, an uncertainty model was used in which it was assumed that the mean regression parameters for these three areas could be used for any areas with similar descriptive statistics and terrain types, given that the nominal data sources are the same. The distribution of the derived and modelled error align relatively well (Fig. 7), with
proportions of positive and negative error, mean and standard deviation and skewness of both distributions agreeing. However, the modelled errors have greater magnitudes (positive and negative) than the original error surface as suggested by its lower kurtosis. This is mainly due to the random proportion of error that can not be correlated with terrain attributes and constitutes the uncertainty of this unknown error proportion. Although it may be possible to tune the model to better replicate this feature of the distribution, a method to perform this tuning whilst preserving the dependence of error on terrain features in space proved difficult. This combination of deterministic and stochastic model components makes it difficult to definitively assess the quality of the uncertainty model compared to the derived error, particularly because of both the heterogeneity of the error and the need to consider both aspatial and spatial error distribution characteristics.

4.1.3 Improving the error model

Semivariogram maps of the derived error surfaces all show a characteristic directional trend, but with a very short range. This is due to the heterogeneity of the observed error, with large areas of small scale, gradually changing error, contrasting with the large error values that abruptly change sign, following the main valleys and ridges. Although directional effects were detected and directional variograms of the error surfaces shown to feature varying range values, the range of the convolution filter in the error model was derived from an omnidirectional variogram map. Using a directional variogram map to derive a convolution filter could improve simulation of correlation in spatial error. Alternatively, analysis of subregions and the application of local semivariograms to the modelled uncertainty data could also improve the simulation of the local amount of spatial autocorrelation. A correlation analysis of the range of autocorrelation of the error with respect to attributes of the underlying terrain would be the next step towards an advanced uncertainty model, e.g. following the approach of Kyriakidis et al. (1999) who used stochastic simulation with varying local uncertainty models. The model presented here does not calculate mean errors of zero for every grid cell when error surfaces are averaged over a large number of runs, as information on the likely sign and magnitude of error is included in the model. This suggests that such errors could be treated as systematic or trends as defined in section 1.2 and should be eliminated prior to analysing and modelling uncertainty. However, this result only holds for certain terrain types within the study data, and correlations were not good enough to clearly identify systematic errors and trends. The model presented aims to provide a means of estimating uncertainty for larger scales, and for the whole of the study area, the simulated uncertainty, averaged over a larger number of realisations, sums to zero for all tested DEMs.

If the model is to be transferred to areas without available SRTM reference data, such as Fennoscandia, the modelling of residuals from error magnitude regression (Eq. 2) can not be done using correlation with derived error. As the residual distribution is symmetric, a simulation using transferred random normal distribution might be possible.

Alternative concepts for modelling GLOBE error using artificial neural networks (ANN) (Behrens et al., 2005) have been explored for the Alps test area using the same set of derived terrain parameters. While the general, systematic error patterns were reproduced well after filtering of the training data, sensible integration of the filtered random error proved to be difficult. However, the use of ANNs might be a viable approach when trying to reproduce systematic, correlated DEM error.
4.2 Example application

The purpose of providing an example application to illustrate the use of an uncertainty model focussed on improving our understanding of the influence of uncertainty in elevation values of widely used digital elevation products on a range of popular topographic indices commonly calculated in the literature. We illustrate these examples on a DEM in the Alps, but we deliberately do not draw any conclusions on the geomorphological implications of these results - rather we restrict ourselves to commenting on differences between parameters for different realisations.

4.2.1 Summary global statistics

When calculating global statistics, it is clear that most of the illustrated terrain parameters are very robust to the uncertainty in elevation modelled. The single exception is the slope of the first order streams. This result makes sense, since first order streams are likely to be found in the roughest areas (where there is the least terrain convergence) and thus are also most susceptible to uncertainty in elevation values. These results do not however suggest that uncertainty is unimportant in considering terrain derivatives, but rather that when globally averaged the effects of uncertainty tend to cancel themselves out.

4.2.2 Watershed derivation and derived statistics

The results for the derivation of watersheds have some important differences from the global statistics calculated. Firstly, they show that the number of watersheds is sensitive to terrain uncertainty, with a considerable variation in the total number of watersheds derived (Fig. 9). Since not every pixel in the DEM is assigned to a watershed, the number of watersheds can grow whilst the number of larger watersheds remains roughly constant and the number of smaller watersheds increases. This result suggests that terrain uncertainties lead to a migration of pour points upstream and that edge effects are of potential importance in limiting the possible number of watersheds derived. Calculation of descriptive statistics per watershed show that, once again, these results are relatively robust with the exception of slope. In this case, statistically significant differences exist between median values of slope at the $p = 0.05$ level. Importantly, slope is no longer averaged according to Strahler order, but rather according to watersheds. However, since the majority of cells are assigned to first order streams, it is likely that the variability in these values strongly influences this statistic.

4.2.3 Comparison of two watersheds

The final set of comparisons between the two catchments show the greatest variability, and thus is perhaps of the most immediate importance to geomorphology. The spatial extents of W1 and W2 vary by over 320% and 290% respectively between their smallest and largest extents. This sensitivity in watershed area to elevation uncertainty is much larger than that reported by Jamieson et al. (2004) for changes in watershed delineation as a function of resolution. Changes in watershed area also have a significant impact on all other terrain variables summarised per basin. For example, hypsometric curves show considerable difference in form varying from convex to S-shaped (Fig. 10). These results are important, since comparison and interpretation of hypsometric curves are typical geomorphological tasks, and our result suggests that uncertainties in elevation may have a significant influence on the hypsometric curves. Where hypsometric curves and integrals are automatically derived, care is required to ensure that the results are robust to potential uncertainties.
in elevation. Furthermore, larger variations in topographic parameters occur not for W1, which varies the more in area, but W2 where the overall elevation is higher and thus the terrain surface is rougher and uncertainty is correspondingly greater. Furthermore, since most of the uncertainty in catchment area in W2 is found in relative “flatlands”, the capture of small areas of relatively constant elevation can significantly change the distribution of both elevation and slope, especially those assigned to lower order streams. These values also vary for W1, but less markedly, presumably because uncertainty in elevation here results in the capture of areas of both high and low relief.

A proviso is required here - GLOBE data appear to have been fitted to a river network algorithmically. Thus, derivation of watershed areas with GLOBE tends to yield few surprises and authors have reported good agreement between GLOBE and other, higher resolution datasets (Jamieson et al., 2004). The modelled uncertainty surfaces result in considerable deviations from the watersheds derived by the hydrologically corrected GLOBE data. However, they provide a realistic estimate of the uncertainties present in a lower accuracy DEM product and, we believe, suggest a transferrable methodology for estimating the impacts of errors in elevation on the derivation of geomorphological indices.

5 Conclusions

This paper has introduced a methodology for, firstly, calculating an error surface and, secondly, based on regression with topographic indices, applying such error surfaces in developing uncertainty models which may be used in assessing the impact of uncertainty in elevation on specific geomorphological studies. The derivation of error surfaces is an important first step in assessing the likely sources of error, and where these errors are predominantly systematic, in suggesting ways to detrend data and thus reduce the influence of error. Where errors contain a significant stochastic component, as is the case here, the use of Monte Carlo Simulation provides a tractable tool for investigating the implications of uncertainty in elevation on both the derivation of geomorphological indices and, as is our intention in future work, the effects of uncertainty on more complex process models.

Application of the model to a case study showed that global statistics describing elevation, hypsometry and mean relative catchment area were relatively robust to uncertainty, though slope, particularly for low order streams showed large values of uncertainty. When comparative statistics for 6th order watersheds were calculated, some statistically significant differences in mean watershed slope were found, though most parameters were once again robust to uncertainty. The most striking results concerned the impact of uncertainty on two large watersheds, where hypsometric curves, catchment area, elevation and slope were all shown to have considerable uncertainty as a function of uncertainty in elevation values.

Acknowledgements

This research is funded by the Swiss National Science Foundation (SNF Project Number 200020-109449). We are grateful for useful discussions with Jo Wood and Pete Fisher (City University London), Phaedon Kyriakidis (Department of Geography, UCSB), Adrian Roellin (Institute for Mathematics, University of Zurich) and Oliver Korup (WSL, Davos) in formulating some of the
ideas in this paper. We would also like to thank Ian S. Evans and Peter L. Guth for their constructive reviews which helped us to considerably improve this paper.

References


Tarboton, D. G., 1999. TARDEM, a Suite of Programs for the Analysis of Digital Elevation Data. Utah State University, 8200 Old Main Hill, Logan, UT 84322-8200, USA.


