Facets of Uncertainty in Digital Elevation and Slope Modeling

ZHANG Jingxiong  LI Deren

ABSTRACT This paper investigates the differences that result from applying different approaches to uncertainty modeling and reports an experimental examining error estimation and propagation in elevation and slope, with the latter derived from the former. It is confirmed that significant differences exist between uncertainty descriptors, and propagation of uncertainty to end products is immensely affected by the specification of source uncertainty.

KEYWORDS uncertainty; accuracy assessment; error surfaces; geostatistics; stochastic simulation; realizations; digital elevation models (DEMs); slope

CLC NUMBER P231.5; TP751

Introduction

With sophistication of information technology such as global positioning system and remote sensing, an increasing quantity of digital terrain data is produced from various sources, contributing to accurate mapping and dynamic monitoring of the natural and built landscapes\(^1\). The value of spatial information, however, depends heavily on a good understanding and proper handling of uncertainty, which occurs due to the inability of any information systems to represent the real world as numeric-alphabets with absolute accuracy, and imperfection of the processes of data acquisition and geo-processing. Relative researches can be see References [4-24].

However, it is never practical or possible to measure elevation at every location to the highest precision, except for at a few locations. It is often at such individual locations that errors are estimated, provided both the data and reference are available or can be retrieved accurately therein. The process of interpolation in such occasions implies that there will be extra uncertainty in the interpolated error surfaces, except at locations where both the data and reference are well defined and available.

Real-world situations become less straightforward due to the disparity in localization of data and reference sources. In other words, when data and reference locations do not coincide originally, values of a field variable at unsampled locations are often estimated from an interpolation process, casting uncertainty onto error at these estimated locations. As a matter of fact, maps are usually not constructed with uniform accuracy, but typically obtained from interpolation or extrapolation based on a fraction of samples, with certainty increasing near data locations. Various GIS technical terms such as coverages and lattices are likely to obscure the distinctions between reality, models, measurements and maps. Thus, it is imperative to expose the spatially varying inaccuracies prevailing in the fundamental process of measuring errors over space.

It becomes clear that there is a hierarchy of
methods for evaluating errors. The most accurate one is that if both data and reference are completely specified, though cautions are required over whether the data or the reference is really mapped with comparable supports. The next most accurate scenario is when both the data and the reference are readily specified or easily retrievable at locations being investigated. Less than that, one may be able to estimate unknown values of the underlying variable at unsampled locations in the data layer where reference values are available, or the other way round, to infer unknown reference values at data locations. The most inaccurate situations occur at locations where both data and reference have to be inferred. For estimating errors at unknown locations, there are two apparent options: kriging or stochastic simulation. Stochastic simulation is preferred over kriging as the former is set to reproduce the spatial structure observed or inferred from sample data.

1 Conceptualization of reality, variability, and error

It is customary to denote the variable such as elevation as \( Z(x) \), its structural component as \( z(x) \), and error as \( \varepsilon(x) \), resulting in the following model:

\[
Z(x) = z(x) + \varepsilon(x) \quad (1)
\]

A common numeric approximation slope \((i,j)\) to the tangent of slope for grid location \(x=(i,j)\) is estimated via:

\[
\tan(\text{slope}(i,j)) = \left[ (\partial Z/\partial X)^2 + (\partial Z/\partial Y)^2 \right]^{1/2} \quad (2)
\]

where partial derivatives are estimated from:

\[
\partial Z/\partial X = (Z(i,j+1) - Z(i,j-1))/2 \quad (3a)
\]

\[
\partial Z/\partial Y = (Z(i+1,j) - Z(i-1,j))/2 \quad (3b)
\]

and where the central grid cell \((i,j)\) is shown in relation to its 8 immediate neighbors in a 3 by 3 local window.

One may generate fields of random error (e.g., using root mean squared errors (RMSEs) as standard deviation), and imposed onto the specified surface at hand. However, such non-spatial error models are limited in error propagation. Hunter et al. investigated error propagation of elevation into slope and aspects, for which DEM errors are modelled as spatially autoregressive random fields:

\[
\varepsilon(x) = \rho W_{\varepsilon}(x) + N(0,1) \quad (4)
\]

and are added as a disturbance term to elevations.

Assuming that one source of reference data is available which can be taken as constituting the true variable \( Z_t(x) \). Let \( z_t(x) \) denote results of measurement or calculation. Since error is defined as the difference between reference and measurement, it might be possible to view the underlying field \( Z_t(x) \) as the sum of its approximation \( z_t(x) \) and an error term \( \varepsilon(x) \), as shown below:

\[
\varepsilon(x) = Z_t(x) - z_t(x) \quad (5a)
\]

\[
Z_t(x) = z_t(x) + \varepsilon(x) \quad (5b)
\]

References investigated stochastic simulation based on samples specific errors, which is then added to the original DEMs data to facilitate simulation of DEMs of higher accuracy, since quality data are incorporated in determining spatial errors.

Usually, error is defined as the difference between measured or computed results and references. Let \( Z_t(x) \) denote, again, the original variable, \( z_t(x) \) measured or calculated, and \( \varepsilon(x) \) error. It is easy to obtain that:

\[
\varepsilon(x) = z_t(x) - Z_t(x) \quad (6a)
\]

\[
z_t(x) = Z_t(x) + \varepsilon(x) \quad (6b)
\]

Above conceptualization means that the simulation will generate equally probable realizations of surfaces which conform to the level and structures of errors found in existent data set or, more realistically, at sampled locations.

Kyriakidis et al. investigated a conflation method for combining primary data of higher accuracy with densely sampled secondary data of lower accuracy. The mainstream technique is co-kriging that exploits spatial and cross-variable correlations. From conflated data set, one will be able to derive error surfaces by subtracting the generated reference data from the data set being analyzed. With recognition of uncertainty in primary data source due to sparse sam-
pling, one can further apply stochastic simulation for generating versions of reference data to acknowledge uncertainty in reference sampling. While it may appear straightforward to compute error, issues do exist. First, there is an issue about how error is defined. Then, comes the issue of error estimation. As defined in Eqs. (5) and (6), errors are generally computed on the assumptions that both data and reference are available at a location and that they are unambiguously defined. Difficulties occur when either the data being tested or the reference data are not available at a particular location. It becomes necessary to activate a process of interpolation so that values at unknown locations are estimated from known or sampled locations. It is possible to infer that the greatest uncertainty occurs at locations where neither the data being tested nor the reference is available, while certainty in error estimation increases at locations where both the data and reference exist. When estimating errors at locations where reference data are available, then uncertainty in error data is affected only by the uncertainty with which unsampled locations are predicted from the available data.

2 A test with photogrammetric data

A test site in an Edinburgh suburb was chosen for experimentation. The test site is to the north of Blackford Hill, where the Royal Observatory is situated. The test site and its vicinity contain a variety of thematic and topographic features: a wooded valley, mostly residential buildings, roads and footpaths, recreational areas, a small lake, agricultural fields and worked allotments, hills and flat ground. It is obviously appreciable that the fabric of urban and suburban areas is often highly varied and compressed, where human beings have made intensive use of every possible space, which requires that terrain data should be recorded at large scale to permit more accuracy.

Two sets of aerial photographs were acquired over the test site, one at a scale of 1:5 000 and the other 1:24 000. Both were flown in natural color, and the larger scale aerial photographs were used for generating reference data in the form of 142 irregularly spaced points, while the smaller scale aerial photographs for data being tested (74 irregularly spaced points). All the data collected were later analyzed using the public-domain geostatistical software system GSLIB.

Experimental semi-variograms for the reference and test data sets are shown in Fig. 1(a) and 1(b), respectively, while the correlogram between these two data sets is shown in Fig. 1(c). Based on the experimental variograms in Fig. 1, Gaussian variogram models are fitted to the data with an equal range of 84 m and sills of 200.0, 139.2, and 166.9 for the 1:5 000 scale data, 1:24 000 scale data, and their cross-variogram.

Upon completion of variogram model fitting, it is possible to perform kriging. Fig. 2 shows the surfaces of estimation (Fig. 2(a)-2(c)) and standard error (Fig. 2(d)-2(f)) for kriging based on the primary data set, the secondary data set, and both, respectively. It is easy to interpret that kriging using the primary data set of higher accuracy and sampling density produced Z estimation with lower variance than using the secondary source. Further, as can be anticipated, co-kriging produced the most accurate estimation of Z values locally and globally, as it called in knowledge embodied in both the primary and secondary data sources. As the spot heights collected from the 1:24 000 scale aerial photographs are known to suffer inaccuracy at a RMSE of about 0.6 m, it is necessary to incorporate a nugget variance of 0.36 in the variogram models concerned.

As expected, kriging tends to produce smoothed surfaces as shown in Fig. 2(a)-2(c). To retain realistic spatial variability, stochastic simulation is preferred. Stochastic simulation aims for reproducing spatial statistics, such as histograms, and structures, such as variograms, as derived from sampled data and will honor samples at data locations, in addition, if the conditional approach is taken. Gaussian sequential approach is often implemented for generating
realizations meeting specified structures and honoring observed values\cite{10}. This requires a process known as normal score transfer by which the original data-based cumulative density function is transferred into normal scores to facilitate modeling with Gaussian parameters of mean and standard deviate.

Results are shown in Fig. 3 where semi-variograms are depicted for normal scores derived from 1:5 000 (Fig. 3(a)) and 1:24 000 (Fig. 3(b)) scale spot heights, respectively. In Fig. 3, experimental variograms are illustrated with diamonds while modes with curves.

Fig. 1 Semi-variograms derived from photogrammetric elevation data

Fig. 2 Kriged surfaces of elevation based on 1:5 000 scale, 1:24 000 scale, and their combination, respectively, with their corresponding standard errors

Fig. 3 Semi-variograms for normal score transformed data

Results from conditional simulation are presented as maps in Fig. 4, where Fig. 4(a)-4(c) show the mean surfaces of simulated elevation surfaces (of 100 realizations) for reference, test data, and co-simulation, respectively, while Fig. 4(d)-4(f) representing their respective surfaces of standard error. It is clearly the case that simulated surfaces appear more varied than their kriged counterparts, and should therefore be closer to the reality.

As both co-kriging and co-simulation were implemented for this study, some discussion is advised. If two sources of data are combined via a geostatistical approach, the resultant data set is supposed to possess accuracy superior to any component data set. If one source of data is measured with precision far greater than the other, it may still make sense to tune the end data product with the source of lower accuracy. However, the issues of support or scale seem to play a role or set a limit to co-kriging and co-simulation, because interaction of the spot height data based on the 1:5 000 and 1:24 000 scale aerial photographs seems not smooth, as shown in Fig. 2(c) for co-kriged and Fig. 4(c)
for co-simulated surfaces. There is certainly room for improvements in terms of the optimal methods for combining spatial data sources.

One may derive error surfaces through subtracting the kriged or simulated surfaces based on the 1:24 000 scale data by the kriged or simulated reference surfaces based on the 1:5 000 scale data. Based on the kriged and simulated surfaces shown in Fig. 2 and Fig. 4, error surfaces were derived, which depict the discrepancies over space between the test data-based surfaces and the reference surfaces, as are shown in Fig. 5(a) and 5(b) while standard errors in error estimates in Fig. 5(d) and 5(e). Clearly, simulation tends to expose spatial variability as opposed to kriging. It may be sensible to use co-kriged and co-simulated surfaces as references to derive error surfaces $E_s^2$, although this was not covered in this paper.

To facilitate further handling of error surfaces, it is necessary to narrow down locations where either test spot heights or reference spot heights are available. For this, kriged and simulated test surfaces were checked with the reference spot heights, resulting in error data whose semivariograms are shown in Fig. 6(a) and 6(b), respectively. Alternatively, one may choose to retain test spot heights but simulate reference surfaces from a set of existing reference spot heights. The semivariograms of error data derived from testing the 1:24 000 scale spot heights against the kriged and simulated reference surfaces are shown in Fig. 6(c) and 6(d), respectively.

Obviously, errors estimated by kriging look smoother than the simulated versions. It is also interesting to note that error statistics derived from retaining the original reference spot heights (estimating the test elevation surfaces) tend to be less varied than that if retaining the test spot heights instead (estimating the reference surfaces). One reason for this is that the reference surface is sampled more densely than the test surface, as is implied by the typical looks of surfaces of standard errors in Fig. 2(d) and 2(e).

Consider visualizing inaccuracy of the test data layer as an error surface $e(x)$. Surface $e(x)$ will be dependent on a variety of factors including measurement precision and proximity to sampled locations. When more than one data source are involved, as is the case for analysis of error surfaces, inaccuracy of the resultant data set will be spatially varying and rely on configuration of a location to the samples of all data sources concerned. From this, it is possible to infer that uncertainty is minimized at locations where both data sets are sampled, while the maximum uncertainty occurs at locations where both data sets are unsampled, as is clearly shown in Fig. 5(d) and 5(e).

It seems more natural to use the statistics in Fig. 6(a) and 6(b) than those in Fig. 6(c) and 6(d) to explore error propagation from elevation data to slope data products. Towards this end, normal scores were derived from error estimates and their variograms are shown in Fig. 7 with re-
results for kriging and simulated-based error data shown in Fig. 7(a) and 7(b), respectively. Again, experimental semi-variograms are indicated by diamonds while models by curves.

Simulation was then carried out to generate 100 realizations of error surfaces, which allowed for estimation of test elevation surfaces by adding a realized error surface to the kriged reference surface shown in Fig. 2(a). Results are shown in Fig. 8, where Fig. 8(a) and 8(c) correspond to the error surfaces of mean elevation and standard errors based on the kriged data, while Fig. 8(b) and 8(d) for the simulated data.

It is straightforward to understand the differences of the surfaces being tested by comparing test elevation surfaces shown in Fig. 2(b) and Fig. 4(b) with those shown in Fig. 8(a) and 8(b), respectively. Moving onto the topic of error propagation from elevation to slope variables, it will be further seen that different approaches in pursuing uncertainty modeling are likely to obtain different results in terms of error quantity and distributions.

Based on simulated elevation surfaces, it is an extension of the calculation using the formulas in Eqs. (2) and (3), slope surfaces based on individual realizations of elevation surfaces and to summarize them in terms of means and standard errors. Depending on the approaches taken to generate elevation surfaces, the resulting spatial statistics about uncertainty and errors in slope will be different.

Take a line similar to those of Fig. 5(a), 5(b), 5(d), and 5(e), it is possible to use simulated elevation surfaces summarized in Fig. 4(a) and 4(b) to calculate slope surfaces, and to derive error surfaces in slope. Results are shown in Fig. 5 where surfaces of means and standard errors in slope errors illustrated in 5(c) and 5(f),
Facets of Uncertainty in Digital Terrain Models: Comparison of Stochastic Simulation and Kriging

ZHANG Jingxiong, et al.

Using the elevation error surfaces summarized in Fig. 8 (a) and 8(b), 100 slope surfaces were calculated for elevation surfaces that were simulated by imposing kriged and simulated error surfaces onto the reference elevation surface, respectively. The resulting slope surfaces were summarized, as are shown in Fig. 9 as surfaces of means and standard errors for slope originating from kriged and simulated errors.

Conclusions

This paper has confirmed that great differences exist in terms of quantities and distributions of uncertainty and error in elevation and slope surfaces, depending on whether uncertainty is modeled directly from stochastic simulation of the original properties or by addition of simulated error surfaces to assumed reference surfaces. It is important that environmental modelers and geospatial analysts pay due attention to the seemingly trivial task of defining errors in geospatial context.

Geostatistics has become increasingly popular among geographers as well as other earth scientists. As measurement inaccuracy or error has been widely known to image interpreters, map digitizing operators, and land surveyors alike, it is important to differentiate stochastic variations from measurement errors. With the understandings that huge quantity of spatial data is now cheaply collected from various platforms and remote sensors, terrain sampling and modeling often require that GIS practitioners take an integrated view about uncertainty and error in this special setting.

REFERENCES

1 Longley P A, Goodchild M F, Maguire D J, et al. (2001) Geographic Information Systems and Science.
2 Wang G X, Gertner G, Parysow P et al. (2001) Spatial prediction and uncertainty assessment of topographic factor for revised universal soil loss equation using digital elevation models. ISPRS Journal of Photogrammetry and Remote Sensing, 56(1): 65-80
3 Hoersch B, Braun G, Schmidt U (2002) Relation between landform and vegetation in alpine regions of Wallis, Switzerland. Computers, Environment and Urban Systems, 26(2,3): 113-139
4 Unwin l)J (1995) Geographical information systems and the problem of "error and uncertainty". Progress in Human Geography, 19(4): 549-558
5 Veregin H (1997) The effects of vertical error in digital elevation models on the determination of flow-path direction. Cartography and Geographic Information Systems, 24(2): 67-79
6 Goovaerts P (2001) Geostatistical modelling of uncertainty in soil science. Geoderma, 103:3-26
7 Vitek J D, Richards D G (1978) Incorporating inherent map error into flood-hazard analysis. Professional Geographer, 30(2): 168-173
8 Lee J, Snyder P K, Fisher P F (1992) Modeling the effect of data errors on feature extraction from digital elevation models. Photogrammetric Engineering and Remote Sensing, 58(10): 1461-1467
9 Hunter G J, Goodchild M F (1997) Modeling the uncertainty of slope and aspect estimates derived from spatial databases. Geographical Analysis, 29(1): 35-49
10 Fisher P F (1998) Improved modelling of elevation error with geostatistics. Geoinformatica, (2): 215-233
11 Csislag F, Kummert A, Kertesz M (1992) Resolution, accuracy and attributes; approaches for environmental geographical information systems. Computers, Environment and Urban Systems, 16(4): 289-297
12 Journel A G (1996) Modelling uncertainty and spatial dependence; stochastic imaging. International Journal of Geographical Information Systems, 10: 517-522
13 Bishop T F A, McBratney A B (2001) A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma, 103(1-2): 149-160
14 Heuvelink G B M, Pebesma E J (1999) Spatial aggregation and soil process modeling. Geoderma, 89(1-2): 47-65
15 Arbia G, Griffith D, Haining R (1999) Error propagation modeling in raster GIS; adding and updating operations. Cartography and Geographic Information Science, 26(4): 297-315
16 Kyriakidis P C (2001) Geostatistical models of uncertainty for spatial data. In: Hunsaker C T, Goodchild M F, Friedl M A, et al. eds. Spatial Uncertainty in Ecology. New York: Springer. 175-213
17 Zhang J, Goodchild M F (2002) Uncertainty in geographical information. London; Taylor & Francis.
18 Holmes K W, Chadwick O A, Kyriakidis P C (2000) Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. Journal of Hydrology, 233(1-4): 154-173
19 Atkinson P M (1997) On estimating measurement error in remotely-sensed images with the variogram. International Journal of Remote Sensing, 18 (14), 3 075-3 084
20 Curran P J, Dungan J L (1989) Estimation of signal to noise; a new procedure applied to AVIRIS data. IEEE Transactions on Geoscience and Remote Sensing, 17: 620-628
21 Wilcox D J (1984) Proposed map accuracy standards for a multipurpose cadastre. Computers, Environment and Urban Systems, 9(2,3): 203-207
22 Adkins K F, Merry C J (1994) Accuracy assessment of elevation data sets using the global positioning system. Photogrammetric Engineering and Remote Sensing, 60(2): 195-202
23 Ruiz M O (1997) A causal analysis of error in viewsheds from USGS digital elevation models. Transactions in GIS, 2(1): 85-94
24 Endreny T A, Wood E F, Hsu A (2000) Correction of errors in SPOT-Derived DEM's using GTOP30 data. IEEE Transactions on Geoscience and Remote Sensing, 38(3):1 234-1 241
25 Kyriakidis P C, Shortridge A M, Goodchild M F (1999) Geostatistics for conflation and accuracy assessment of digital elevation models. International Journal of Geographical Information Science, 13: 677-707
26 Goovaerts P (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology, 228(1-2): 113-129
27 Zhang J (1996) A surface-based approach to the handling of uncertainties in an urban-orientated spatial database; [Ph. D Dissertation]. Edinburgh: The University of Edinburgh.
28 Deutsch C V, Journel A G (1998) GSLIB: Geostatistical software library and user’s guide. New York; Oxford University Press.