GEOMORPHOMETRY PERUGIA ITALY 2025 JUNE 9-13

Enriching LiDAR data with partial derivatives and its uncertainty estimates

Carlos López-Vázquez
LatinGEO Lab IGM+ORT
Universidad ORT Uruguay
Montevideo, URUGUAY
carlos.lopez@pedeciba.edu.uy,
carloslopez@fi365.ort.edu.uy

F. J. Ariza-López
Dpto. Ingeniería Cartográfica,
Geodésica y Fotogrametría.
Universidad de Jaén
Jaén, España
fjariza@ujaen.es

J. F. Reinoso-Gordo
Dpto. Expresión Gráfica
Arquitectónica y en la Ingeniería
Universidad de Granada
Granada, España
jreinoso@ugr.es

Abstract— **After some manipulations, raw LiDAR data can be filtered so as to describe the terrain surface or bare ground. The resulting set is composed of (x,y,z) triplets, with some uncertainty in all three dimensions. We want to describe a procedure able to enrich this 3D cloud adding also estimates of partial derivatives with its uncertainty, thus becoming a 7D point. Our theoretical framework assumes that the point cloud is itself error free, and that the uncertainty in the partial derivatives estimate will arise because of the finite spacing between points. Computing the enriched cloud will be a one-time operation. The resulting dataset could be used in many ways, including improving elevation estimates by using higher order Hermite interpolants intended to exploit the derivative values. The uncertainty information might be propagated to the interpolated elevation. The envisioned procedure has contacts with numerical methods for solving partial differential equations. Finite Difference methods rely on a regular grid like the one provided by standard DEMs to produce estimates of the partial derivatives which are functions of the grid size. Finite element methods use a non-regular grid, and its derivative estimates can also be related to some norm of the elements, defined through elaborated topological relationships. Meshless methods use cloud of points not linked in elements, an approach which offers enough flexibility to deal with selected problems in fluid mechanics, where moving boundaries and free surface precludes using rigid meshes. Tools already developed to estimate partial derivatives and its uncertainties can be conveniently reused with the LiDAR cloud dataset, even without a partial differential equation problem involved.**

I. INTRODUCTION

Slope information is used extensively for a wide range of applications, from terrain analysis and hydrological modeling to urban planning, agriculture, engineering, environmental protection, and recreation. By understanding how steep or flat the terrain is, professionals can make more informed decisions about land use, risk management, and infrastructure design. Additionally, slope is often combined with other spatial data to provide a more comprehensive understanding of the landscape. Aspect is also a critical terrain attribute of the terrain and has a wide range of applications across various domains. It is used to model and analyze solar exposure, climate variation, vegetation, agriculture, erosion, hydrology, landslide risk, wildlife habitats, and more. Aspect, in combination with other environmental factors, helps in decision-making processes for resource management, environmental conservation, urban planning, infrastructure design, and outdoor recreation.

Both slope and aspect are examples of functions defined from the first partial derivatives of the terrain surface, which in turn depends on a model of the terrain. The description of the terrain surface (in all its variants: surface, terrain, bathymetry, etc.) has become instrumental in many application fields. The dominant format is the Digital Elevation Model (DEM), which can be loosely defined as a set of elevation values defined over a regular grid. Other possibilities certainly exist: contour lines, triangular irregular networks (TIN), spectral models, etc. All of them have advantages and disadvantages. LiDAR point clouds and TIN data sets have some formal similarities, but we will focus in the LiDAR case. In traditional TIN models each point was measured and selected carefully in the terrain, reflecting the choices of the surveyor. On the other hand, LiDAR data were produced by fully automated means, with little or no chances to select the point location. Thus, there is room for a more refined treatment. The popularity of LiDAR datasets is growing [1]. Lower instrument costs, affordable and better tools to process the huge number of points, use of UAV and steady increase in computer performance paved the way for this.

There are a number of options to estimate the partial derivatives, but its uncertainty is not very well known. When slope or aspect values are involved in a decision, their uncertainty might need to be known to ascertain their impact on the uncertainty of the decision. The procedure is named sensitivity analysis, and can lead to a non reliable decision if the uncertainty is too high.

Let's assume that the filtered LiDAR data define a cloud of points located over the terrain, while any other early hits for vegetation, power lines, bridges and so were removed. In addition we can restrict ourselves to the case where such points have themselves no positional error. Some interpolation procedure will be required in order to estimate the elevation in other locations. A number of options exist. They are more or less sophisticated, leading to different accuracy estimates for the surface. The simplest one is the nearest neighbor, only exact for constant (horizontal) terrain. A second possibility is the bilinear, which exactly matches elevation and slope provided the terrain is planar. They are examples of first and second order methods (p=1 and p=2), because they are exact interpolators of surfaces defined by polynomials up to degree p-1 but not degree p.

There exist many other methods (bicubic, radial base functions, inverse distance weighting, splines, natural neighbors, kriging, etc.) where the concept of order is not meaningful. The Hermite family is a particular one, and has the property to interpolate not just elevation values but also partial derivatives when they are available. The Shepard method (also known as IDW) is one of them, but there exist also bivariate Hermite, cubic Hermite, Hermite splines, Hermite Radial Base Functions, etc. They produce more accurate estimates of the elevation and slope at the price of using more data and computing time. Once established that there are interest in this we will describe how we plan to enrich a given LiDAR cloud with two first partial derivative estimates and its uncertainty. They comprise 4 extra attributes attached to a 3D point cloud adding up to a 7D point dataset, which might have other attributes as well.

II. Connection to Partial Derivative Equations

All numerical methods for solving Partial Derivative Equations (PDE) follow a similar procedure. Given a set of (x_i,y_i), i=1,...,N locations where the solution value is requested, the first step is to assume that the z_i values are known. Then, some expressions relate the (x_j,y_j,z_j) values of points in the neighborhood of the i-th one to estimates of the partial derivatives at the i-th point. Given such partial derivative estimates, if we impose that they should satisfy the PDE then a set of (usually nonlinear) equations in z_i, i=1,...,N can be established, and the problem turns out to be an algebraic one. All the procedure is well established, carefully analyzed and documented in books [2]. There exist many computer libraries tailored to solving general problems, as well as others devoted to specific PDE e.g. OpenFOAM (https://github.com/OpenFOAM/OpenFOAM-dev), SU2 (https://github.com/su2code/SU2) or deal.II (https://github.com/dealii/dealii). One mandatory requirement for any estimate of the partial derivatives is that their uncertainty can be computed, and can be adjusted to a given tolerance. In a PDE context, this can be achieved by refining the grid, or adding extra locations. The uncertainty estimate of the partial derivative is thus integral part of the procedure, and it is not connected with the specific PDE to solve.

In our case we do not have a PDE to solve, but on the other hand we have no unknowns because we already have the elevation values. So, the partial derivatives estimates can be obtained straightforwardly using such routines. Following common practice, such values come jointly with its (deterministic) uncertainty estimate. In general such values are not deemed to be an error bound, but just a reasonable estimate of its size.

A. *Case of regular grids (DEM)*

This is the best known case. We are used to Finite Difference approaches, where (for example) a polynomial of second degree is fitted to elevations of points located over a 3x3 window. Depending on which of the 9 points we want to use, we can recover the Evans-Young formula [3], the Horn's one [4], or others intended for estimate first and second partial derivatives. If we need first, second and third partial derivatives we need to use at least a 5x5 window to use the formulae proposed by [5]. The location of the elevations is always over a regular grid, a situation far from the LiDAR case.

Recently [6] show that in this case it is possible to extract not only partial derivative estimates but also its uncertainty. To achieve so he proposed as an option to estimate it following standard practice ([7], [8]) by using two methods of different order. For example this can be done for first partial derivatives with the pair {Evans-Young, Florinsky}, which are examples of methods of order 2 and 4. The uncertainty of the higher accuracy Florinsky's method can be estimated as the absolute value of the difference of the Evans-Young's and Florinsky's values. Similar results hold for the second partial derivatives despite not considered in [9]. However, since Evans-Young's cannot produce estimate of the third partial derivatives, to estimate its uncertainty Florinsky's method needs to be used jointly with an even higher order method when available.

Despite not widely considered in geosciences, it is possible to use the uncertainty value to further refine the method. If at the j-th point the pair {Evans-Young, Florinsky} produces an uncertainty estimate which is too large, it is is possible to switch locally to an even higher order method pair to reduce it at the price of involving a larger neighborhood and more sophisticated methods. This will be explained in the following section.

B. Case of irregularly placed points (LiDAR and TIN)

Neither Florinsky nor Evans-Young can be used with irregularly placed points. None of them produces uncertainty estimates, which is the novelty here. Thus, we propose to reformulate the problem from the beginning. Instead of using Finite Difference methods, we will resort to another less known PDE approach named Meshless methods. They are almost the single alternative to certain PDE problems. For example, problems in fluid mechanics with moving boundaries like free surface flows, or crack propagation in solids. They are very difficult to tackle with rigid meshes, so Meshless methods were designed.

Among other alternatives, we will sketch a basic one denoted strong p-version. Given a neighborhood J={i, j_2, ..., j_k} of the i-th point, for k=3 (k = number of points in the neighborhood including the i-th) we can use points in J to fit a plane that goes to the three points. Then, an estimate of both first partial derivatives are at hand so we have an estimate with p=2. We can then request k=6. With the set J we can estimate a second order polynomial that goes to the six points. New estimates of the first partial derivatives are available, now with p=3. If the absolute difference between $f_{x,(p=2)}$ and $f_{x,(p=3)}$ are below a given tolerance tol we are done. Otherwise, we can request k=9, discard the estimate of p=2 and analyze the difference between $f_{x,(p=3)}$ and $f_{x,(p=4)}$. In general, to attain degree p we will need k=(p+1)(p+2)/2 points in the J set. After specifying a tolerance tol, the method increases the smallest p until the uncertainty is below tolerance.

As described the process is very straightforward, and we omit some practical details here. The set J must be unisolvent, i.e. the corresponding Vandermonde matrix should be invertible. Otherwise, a point can be discarded and substituted by another neighbor. The use of polynomials is indeed not the single option, and Radial Basis Functions are also an alternative [8]. Many of these possibilities are considered, however, in mathematical libraries, e.g. SciPy, NumPy, Scikit-Learn for the Phyton environment include specific modules for interpolation interpolations and fitting, and we need not to describe it further here. The overall procedure is denoted as the strong form of the p-version of the Meshless methods. Describing other alternatives (weak form, h-version and r-version, etc.) are outside our goal here.

III. ILLUSTRATING THE OPERATION

Fig.1 illustrates the Meshless approach, using a PDE solver of the heat equation. Points are located irregularly in the domain, and its color denotes the order of the approximation locally used to estimate the partial derivatives. At the beginning the approximation is of 2^{nd} order everywhere. The presence of a heat source (red cross in the figure) introduces an irregularity in the problem. This is conceptually similar to a pit in the terrain, or a sharp mountain. Our goal here is to illustrate the procedure by looking at the final solution in Fig. 2. Notice that most of the points, far from the heat source, remain of 2^{nd} order approximation.

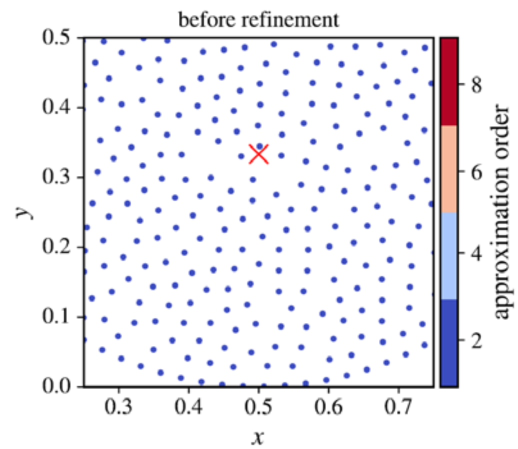

Figure 1. Example of a derivative approximation of variable order. The red cross denotes an irregularity (a heat source) while solving a PDE. Notice that all points are initially blue, corresponding to a derivative approximation with p=2. (from [10])

However, in the vicinity of the source the approximation order was increased attaining even p=8, in order to keep uncertainty below a given tolerance.

The sketched procedure is a simplified version of what needs to be implemented. Many details are omitted, and future works will offer a throughout description. Presently we are considering the use of the MEDUSA package [10] which has libraries suited for the purpose. It implements the strong form of the Meshless method.

The abovementioned procedure can be extended even to DEM. Present day procedures are of fixed order for the whole domain, typically p=2. There is nothing between the choice of the 9 points of a 3x3 window and 25 points of a 5x5 window. Such limitation does not apply to a Meshless method, which can add one point at a time thus using more compact schema. As described, we considered only explicit schemas; there are also implicit ones, able to attain (over regular grids) substantially larger values for p (of the order of p=100).

IV. CONCLUSIONS

Despite we present here work in progress, we feel that the connection with well developed numerical procedures for solving PDE equations is already an interesting contribution. For a PDE is mandatory to keep the partially derivative error estimates under control, so there are provisions to alter either the grid size (for the case of Finite Difference methods over regular grids), or the order of the approximation for the Finite Element method. The former is known as the h-version while the latter is known as the p-version. For the case of LiDAR data sets, the geometry can be appropriately handled through the so-called Meshless method, which is itself an even more flexible approach than the Finite

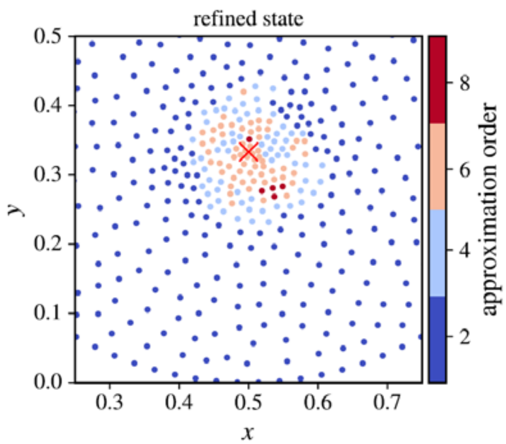

Figure 2. Same set of points than in Fig. 1, but now its color illustrates the local approximation order used in order to keep the uncertainty below a prescribed tolerance. In the neighborhood of the irregularity the order increases up to p=8. (from [10])

Element method. Our assumptions are that the filtered LiDAR data representing base ground is free from positional errors, thus attaching the entire uncertainty estimate to the effect of finite distances between points. Two reasons prevent using the h-version. The first one is that we assume that the LiDAR information is already captured. The second one is that adjusting for the density might introduce scale effects, because closer points describe information not available at other scale. So we plan to compute the partial derivatives using as many as necessary points in the neighborhood chosen to be separated at least a prescribed h distance. The approach is fully deterministic, establishing an interval for the estimates independent of any stochastic assumption of the elevation error surface. On the other hand, it does not provide any cue of the distribution of the likely random error of the derived variables. We plan to use the MEDUSA package [10] on a LiDAR dataset of Andalusia as a case test.

V. Acknowledgments

This research has been partially funded by the Project: "Point-to-point quality in LiDAR data for Digital Terrain Elevation Models in Engineering" (PID2022-138835NB-I00) funded by MICIU/AEI/10.13039/501100011033/ and FEDER, UE. More information is available online: https://coello.ujaen.es/investigacion/web_giic/P2PQuaLiDEN/ (accessed on 8 January 2025).

References

[1] Guth, P. L., and T.M. Geoffroy, 2021. LiDAR point cloud and ICESat-2 evaluation of 1 second global digital elevation models: Copernicus wins. Transactions in GIS, 25, 2245–2261. https://doi.org/10.1111/tgis.12825

[2] Pinder, G. F., 2019. Numerical Methods for Solving Partial Differential Equations: A Comprehensive Introduction for Scientists and Engineers. Wiley. ISBN: 978-1-119-31611-4

[3] Young, M. and Evans, I.S., 1978. Statistical characterization of altitude matrices. Report No. 5, Grant DA-ERO-591-73-G0040. Dept. of Geography, University of Durham, England, 26 pp.

[4] Horn, B.K., 1981. Hill shading and the reflectance map. Proceedings of the IEEE, 69(1), pp.14-47. https://doi.org/10.1109/PROC.1981.11918

[5] Florinsky, I. V., 2009. Computation of the third-order partial derivatives from a digital elevation model. International Journal of Geographical Information Science, 23 (2), 213–231. https://doi.org/10.1080/13658810802527499

[6] López-Vázquez, C., 2022. Uncertainty interval estimates for computing slope and aspect from a gridded digital elevation model. International Journal of Geographical Information Science, 36(8), pp.1601-1628. https://doi.org/10.1080/13658816.2022.2063294

[7] Zienkiewicz, O.C. and Zhu, J.Z., 1987. A simple error estimator and adaptive procedure for practical engineerng analysis. International journal for numerical methods in engineering, 24(2), pp.337-357. https://doi.org/10.1002/nme.1620240206

[8] Jančič, M., Strniša, F. and Kosec, G., 2022, July. Implicit-explicit error indicator based on approximation order. In 2022 7th International Conference on Smart and Sustainable Technologies (SpliTech) (pp. 01-04). IEEE. https://doi.org/10.23919/SpliTech55088.2022.9854342

[9] Jančič, M. and Kosec, G., 2024. Strong form mesh-free hp-adaptive solution of linear elasticity problem. Engineering with Computers, 40(2), pp.1027-1047. https://doi.org/10.1007/s00366-023-01843-6

[10] Slak, J. and Kosec, G., 2021. Medusa: A C++ library for solving PDEs using strong form mesh-free methods. ACM Transactions on Mathematical Software (TOMS), 47(3), pp.1-25. https://doi.org/10.1145/3450966