Geoderma 170 (2012) 217–226

Contents lists available at SciVerse ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

A comparison of LiDAR-based DEMs and USGS-sourced DEMs in terrain analysis for knowledge-based digital soil mapping

Xun Shi [a,*], Lyndsey Girod [a], Robert Long [b], Roger DeKett [b], Jessica Philippe [b], Tom Burke [b]

[a] Department of Geography, Dartmouth College, 6017 Fairchild, Hanover, NH 03755, USA
[b] USDA-NRCS, 481 Summer Street, Suite 202, St., Johnsbury, VT 05819, USA

ARTICLE INFO

Article history:
Received 21 April 2010
Received in revised form 4 October 2011
Accepted 21 November 2011
Available online 24 December 2011

Keywords:
LiDAR, digital elevation model
Digital soil mapping
Terrain analysis
Slope gradient

ABSTRACT

While LiDAR-based digital elevation models (DEM) are more accurate and precise than the USGS-sourced DEM that are widely used in soil mapping in the US, their high cost and other problems prohibit an easy decision of adopting them in service-oriented soil mapping conducted by a government agency like USDA-NRCS. This study compares the performances of LiDAR-based DEM and the USGS-sourced DEM in calculating slope gradient as an input for knowledge-based digital soil mapping (KBDSM), aiming to provide scientific evidence and more importantly, propose a scientific approach to evaluating the two types of DEM for KBDSM. We conducted the comparison by evaluating how closely the DEM-based slope gradient values match the field-measured values. For a small watershed in northern Vermont, US, we prepared three DEM, including a 10-m DEM interpolated from the 7.5-minute USGS topographic map, a 1-m DEM based on LiDAR points, and a 5-m DEM resampled from the 1-m DEM. When calculating slope gradient, we applied two neighborhood sizes (10 m and 30 m), two neighborhood shapes (square and circular), and three slope gradient algorithms (Evans–Young, Horn, and modified Zevenbergen–Thorne) to the three DEM. We then compared the calculated slope gradient values with the values measured by soil scientists at 159 sample locations in the study area. Statistics show that across all the tested settings, the LiDAR-based DEM perform significantly better than the USGS-sourced DEM. We conclude that LiDAR-based DEM may considerably improve the quality of inputs for KBDSM. We also find that the results from the 1-m LiDAR-based DEM and the resampled 5-m DEM do not show considerable and consistent differences.

1. Introduction

Terrain attributes such as slope gradient can be major elements in a soil-landscape model, the basis of soil mapping. In digital soil mapping (DSM), terrain attributes are usually derived from gridded digital elevation models (DEM). Thus the quality of the DEM may have a direct impact on the quality of the resulting soil maps, making the choice of DEM critical to the success of any DSM project.

In the US, The Natural Resources Conservation Service (NRCS) of the United States Department of Agriculture (USDA) conducts routine soil survey and mapping to provide soil information to farmers, planners, engineers, and others. Herein we call this type of soil mapping *service-oriented* soil mapping to distinguish it from *research-oriented* soil mapping that serves academic purposes. While numerous *computation-based* techniques relying on dense sampling have been explored in research-oriented soil mapping (see McBratney et al., 2003 for a comprehensive review), today, most (if not all) service-oriented soil mapping is still *knowledge-based*, i.e., the soil-landscape models are built by soil scientists based on their knowledge of local soils, particularly that gained through years of fieldwork. Hudson (1992) considers this process to be the current *paradigm* of soil survey and mapping. It has been argued that the knowledge-based approach is efficient and economical when the mapping area is large and the knowledge is available, which is often the case in a service-oriented soil mapping project (Hudson, 1992; MacMillan et al., 2007; Shi et al., 2009).

Then, in knowledge-based digital soil mapping (KBDSM), on the one hand, soil scientists use field measurements to build soil-landscape models; on the other hand, the terrain attribute values sent into the soil inference are calculated from DEM. Hence how well the calculated values match the field measured values may greatly impact the quality of the resulting soil maps, because if they do not well match, we are eventually feeding the model with values it was not built for. Therefore in KBDSM this match is a critical criterion in choosing DEM. This also implies that a DEM with higher accuracy and precision does not automatically justify its use in KBDSM (Smith et al. 2006). For example, when measuring slope gradient in the field, the soil scientist may subjectively decide to ignore a small bump on the ground, and in this case a DEM with higher accuracy and precision is not necessarily better in generating values that match the field measurement.

* Corresponding author at: 6017 Fairchild, Department of Geography, Dartmouth College, Hanover, NH 03755, USA. Tel.: +1 603 646 0884; fax: +1 603 646 1401.
E-mail address: xun.shi@dartmouth.edu (X. Shi).

0016-7061/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2011.11.020

The DEM provided by the US Geological Survey (USGS) or customarily created based on the USGS topographic maps are currently the main DEM products used in the service-oriented soil mapping in the US. In this paper we call these DEM *USGS-sourced DEM*. Major justifications for using the USGS-sourced DEM include their complete national coverage and low cost. On the other hand, the inherent error in these DEM resulting from the data source and the creation process has been well studied (e.g., Guth, 1999; Holmes et al., 2000; USGS, 2000), and the impact of the coarse resolutions of these DEM on environmental modeling has drawn considerable and continuous attention, particularly from the soil science area (e.g., Erskine, et al., 2007; Pike, et al., 2006; Shi, et al., 2007; Smith et al. 2006; Thompson et al., 2001; Venteris and Slater, 2005).

An alternative to the USGS-sourced DEM is the DEM created from the Light Detection and Ranging (LiDAR) data. LiDAR data are gathered by emitting fast pulses of a focused infrared laser from a remote object, oftentimes a plane flying over a specified flight path. Through measuring the elapsed time between the transmitted and received signal, the distance between the ground and the plane is determined and in turn the elevation of the ground is derived (US Army Corps of Engineers, 2002). Various interpolation methods can then be applied to LiDAR points to generate gridded DEM (Bater and Coops, 2009; Liu, 2008; Raber, et al. 2002; US Army Corps of Engineers, 2002).

The LiDAR process can cover a large area quickly, and more importantly, the data collected have extremely fine resolution, both horizontally and vertically, with a high level of accuracy. In fact, LiDAR-based DEM are among the ones with the finest resolution and accuracy available today (Anderson et al., 2006; Chow and Hodgson, 2009; Hodgson, et al. 2003; Liu, 2008), and have been used in many environmental applications (e.g., Akay et al., 2009; Barber and Shortridge, 2005; Kasai et al., 2009; Murphy et al., 2008; Rothwell and Lindsay, 2007; Tenenbaum et al., 2006). However, the fine resolution and high accuracy come with costs. First, LiDAR data are much more expensive, compared with the USGS-sourced DEM. Second, the data size of LiDAR raw data is huge and requires extensive labor and computation to produce DEM (Anderson et al., 2006; Liu, 2008). Third, LiDAR-based DEM, even generated from the LiDAR points that have gone through the post-process to remove non-ground noise, may still pick up surface details that are not important or even become noise to the soil scientist when assessing soil-landscape relationships, like tree canopies, roads and associated berms and ditches, furrows, and surface stoniness. Automatically identifying and removing those artifacts and noises from the LiDAR-based DEM remain a challenge (Liu, 2008).

Because of their fine resolution and accuracy, LiDAR-based DEM naturally come under consideration for DSM. However, the high cost and lack of direct evidence that they can significantly improve the quality of KBDSM prohibit an easy decision of adopting this type of DEM in service-oriented soil mapping conducted by a government agency. Comparisons between LiDAR-based DEM and the USGS-sourced DEM can indeed be found in the literature (e.g., Hodgson, et al., 2003; Wang and Zheng, 2005; Zhang et al., 2009), but none of those studies are within the context of soil mapping, and thus their results are hardly applicable to KBDSM. More generally, while numerous studies claiming that they are evaluating "accuracy" of DEM-based terrain attributes such as slope gradient, aspect, and curvatures, few directly serve the purpose of KBDSM. In fact, we consider that for a real-world land surface those terrain attributes do not have objective and universal accuracy in the first place, because in the real world those attributes do not have objective and universal *true values*. The value of, e.g., slope gradient, is fractal and also method- or model-based. As Florinsky (1998) points out, the "accuracy" is reference-dependent and people have used widely different references in the "accuracy" evaluation, including hand measurements from topographic maps (Evans, 1980; Skidmore 1989), calculated values from a high-resolution DEM (Chang and Tsai, 1991), calculated values from a mathematically created smooth (derivable) surface (Hodgson, 1995; Jones, 1998a; 1998b; Zhou and Liu, 2004), and field measurements (Bolstad and Stowe, 1994; Giles and Franklin 1996; Shi, et al. 2007). Apparently, the "accuracy" evaluated using these different references do not mean the same thing. Florinsky (1998) proposed a way to get around this problem, but unfortunately his method does not really address the problem. He implied that his mathematical method for estimating the root mean square errors (RMSE) associated with the different algorithms for calculating certain terrain attributes might be a better way to evaluate the accuracy of the calculated values, as it does not rely on reference values. What the method evaluates, eventually, is not the accuracy of the values, but the sensitivity of the algorithm to noise in the DEM. His finding that the Evans–Young method (Evans, 1979; 1980; Young, 1978) is the best among the several considered methods is consistent with the empirical study conducted by Jones (1998b). His statement that "(i)t should be stressed that we examine the fundamental error in the algorithms rather than an error associated with how well polynomials are used within those methods … to model the real elevation distribution" indicates his understanding of the limitation (Florinsky, 1998, page 52). But strictly speaking, the "error" in this quotation is not real error if we are talking about, e.g., the difference between the calculated slope gradient value and the true slope gradient value; instead, it is the difference between the slope gradient value calculated using elevations with noise in them and the slope gradient value calculated using the true elevations; in other words, it is the sensitivity to elevation error.

Since no true value of certain terrain attributes like slope gradient can be defined for a real-world land surface, the quality or usefulness of the calculated values of such attributes is then entirely application- and situation-specific, and the evaluation has to be based on what is used in that specific application as the reference value. In KBDMS, this reference value is the field measurement.

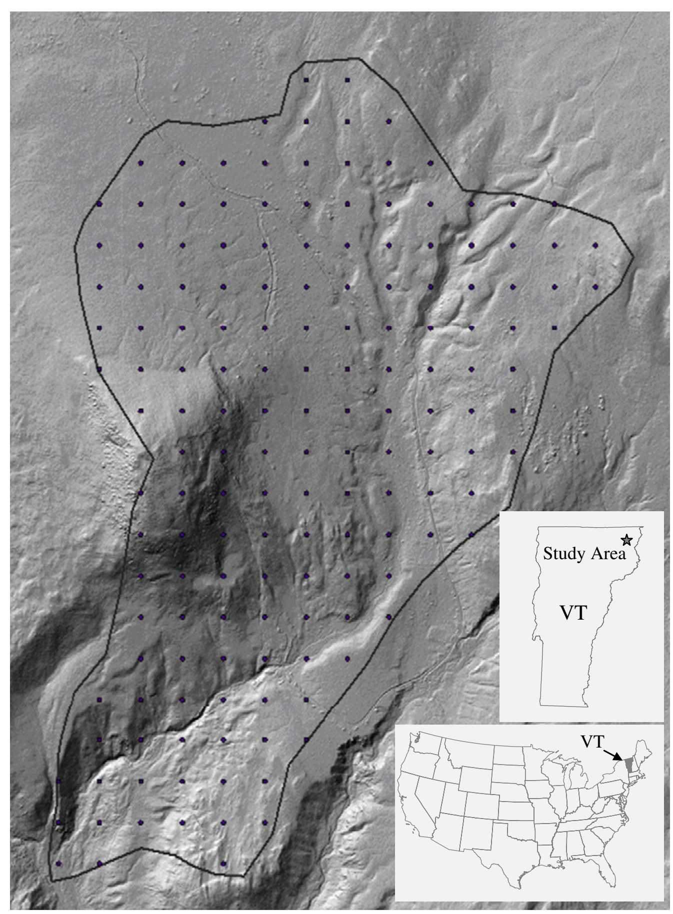

Fig. 1. The study area and sample locations.