Computers & Geosciences 35 (2009) 289–300

Contents lists available at ScienceDirect

Computers & Geosciences

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

Evaluating error associated with lidar-derived DEM interpolation

Christopher W. Bater *, Nicholas C. Coops

Integrated Remote Sensing Studio, Department of Forest Resources Management, University of British Columbia, 2310–2424 Main Mall, Vancouver, British Columbia, Canada V6T 1Z4

ARTICLE INFO

Article history:
Received 21 March 2007
Received in revised form
10 July 2008
Accepted 9 September 2008

Keywords:
Remote sensing
DEM validation
Prediction uncertainty
CART analysis

ABSTRACT

Light detection and ranging (lidar) technology is capable of precisely measuring a variety of vegetation metrics, the estimates of which are usually based on relative heights above a digital elevation model (DEM). As a result, the development of these elevation models is a critical step when processing lidar observations. A number of different algorithms exist to interpolate lidar ground hits into a terrain surface. We tested seven interpolation routines, using small footprint lidar data, collected over a range of vegetation classes on Vancouver Island, British Columbia, Canada. The lidar data were randomly subsetted into a prediction dataset and a validation dataset. A suite of DEMs were then generated using linear, quintic, natural neighbour, regularized spline, spline with tension, a finite difference approach (ANUDEM), and inverse distance weighted interpolation routines, at spatial resolutions of 0.5, 1.0 and 1.5 m. In order to examine the effects of terrain and ground cover on interpolation accuracies, the study area was stratified by terrain slope, vegetation structural class, lidar ground return density, and normalized difference vegetation indices (NDVI) derived from Quickbird and Landsat7 ETM+ imagery. The root mean square (RMS) and mean absolute errors of the residuals between the surfaces and the validation points indicated that the 0.5 m DEMs were the most accurate. Of the tested approaches, the regularized spline and IDW algorithms produced the most extreme outliers, sometimes in excess of ± 6 m in sloping terrain. Overall, the natural neighbour algorithm provided the best results with a minimum of effort. Finally, a method to create prediction uncertainty maps using classification and regression tree (CART) analysis is proposed.

1. Introduction

Light detection and ranging (lidar) technology is an active remote sensing technique capable of simultaneously mapping the Earth's surface and overlying features, including vegetation and buildings, with sub-metre vertical accuracy. Small footprint, discrete return airborne lidar typically employs a laser scanner slaved to an inertial measurement unit (IMU) and a global positioning system (GPS) for accurate measurements of aircraft orientation and position. The laser scanner emits discrete pulses to the ground from which multiple reflections or returns can be detected, allowing for the simultaneous mapping of the ground, vegetation, and other features. The distance estimates are based on the time between pulse emission and return detection.

Lidar data has been applied across a variety of disciplines, including archaeology, structural geology, geomorphology, engineering, resource management, and disaster assessment and planning. Recently, lidar has been gaining recognition in forestry and ecology as an effective tool for estimating a variety of vegetation metrics, including tree heights, biomass, crown size, leaf area index, and vertical canopy structure (e.g. Næsset, 1997;

* Corresponding author. Tel.: +1 604 822 4148; fax: +1 604 822 9106.
E-mail address: cbater@interchange.ubc.ca (C.W. Bater).

0098-3004/$ - see front matter
doi:10.1016/j.cageo.2008.09.001

Lefsky et al., 1999; Dubayah and Drake, 2000; Lim et al., 2003; Coops et al., 2004). Regardless of the variable under investigation, estimates are typically based on their height above a continuous digital elevation model (DEM) representing the Earth's bare surface. Airborne lidar surveys are typically designed to have a dense and evenly distributed point spacing. In higher leaf area canopies, however, ground visibility is reduced, resulting in datasets containing a large number of vegetation returns and a relative paucity of terrain information. This will have implications for not only the quality of derived DEMs and their representation of terrain morphology, but also for the accurate estimation of vegetation metrics.

Previous research has shown that the accuracy of DEMs varies with changes in terrain and land cover type (e.g. Adams and Chandler, 2002; Hodgson and Bresnahan, 2004; Hodgson et al., 2005; Su and Bork, 2006). By surveying ground returns at their x and y locations, Hodgson and Bresnahan (2004) decomposed lidar error into four components. Specifically and in decreasing order of importance, these included: lidar system measurements, interpolation error, horizontal displacement error, and survey error.

Adams and Chandler (2002) assessed the accuracy of a lidar-derived DEM of the Black Ven mudslide in Dorset, United Kingdom. The DEM had a spatial resolution of 2 m, although no information on the density of ground returns was available. Validation data consisted of a DEM generated from survey-grade points. Adams and Chandler (2002) reported an overall root mean square (RMS) error of 0.26 m, and found that the lidar data tended to increasingly underestimate terrain elevation as slope increased.

Hodgson et al. (2005) examined the effects of land cover and slope on DEM accuracy. Located in a watershed in North Carolina, USA, their study area consisted of gently rolling terrain. Land cover classes included grass and scrub/shrub, and pine, deciduous, and mixed forests. Lidar data were collected in leaf-off conditions, with an average ground return posting distance of one point every 31.1 m^2 (corresponding to density of 0.03 points/m). Slope was then modelled by linear interpolation of a triangulated irregular network (TIN). Reference data consisted of 1225 survey-grade points collected along 23 transects, and reference slope was calculated as the average slope of adjacent segments along survey transects. Hodgson et al. (2005) reported RMS errors of 0.145–0.361 m for the different land cover classes, with higher errors occurring in areas with tall canopy vegetation. The scrub/shrub class, however, exhibited the largest RMS error. Little evidence was found for increased elevation errors in areas with slopes from 0° to 10°, but lidar-derived slope was generally under-predicted as terrain slope increased.

The selection of an appropriate algorithm and spatial resolution for DEM interpolation may become an important decision, especially in uneven terrain, as differences in terrain model heights may directly affect the estimates of vegetation metrics. Interpolation methods can be broadly defined as being deterministic or probabilistic (Maune et al., 2001). Deterministic methods are based only on surrounding values, with algorithms using mathematical formulae to determine the influence of immediate neighbour values. Probabilistic geostatistical methods rely on spatial autocorrelation, and account for distance and direction when determining the importance of surrounding values (Maune et al., 2001).

Interpolation algorithms vary widely in their complexity, ease of use, and computational expense. Previous authors have investigated DEM interpolation routines with varying results. Using lidar data collected over smoothly varying hillslopes near Humberside, United Kingdom, Lloyd and Atkinson (2002) employed cross-validation and a jack-knife approach (e.g. subsets of the lidar data) to test inverse distance weighted (IDW) interpolation and two types of kriging. Their original dataset consisted of 1,39,694 returns within a 500 m by 500 m area (corresponding to a mean point density of 0.56 pulses/m^2). To assess the ability of the interpolation algorithms to predict heights where point densities were low, Lloyd and Atkinson (2002) systematically decimated their data, validating surfaces created using 50% (0.28 pulses/m^2), 25% (0.14 pulses/m^2), and 5% (0.03 pulses/m^2) of the original dataset. Lloyd and Atkinson (2002) found that kriging was the more accurate method when point densities were low, but concluded that no benefit was derived from using the more sophisticated geostatistical approach when large amounts of data were available.

Yanalak (2003) examined 1 m resolution DEMs created using linear, natural neighbour, nearest neighbour, weighted average polynomial, minimum curvature and multiquadratic routines. The gridding methods were applied to five 1 ha test datasets, each containing 150 scattered reference points digitized from five theoretical test surfaces; the true heights of the test profile data were calculated using surface equations. Yanalak (2003) found that the minimum curvature and multiquadratic routines were the most accurate based on test profiles across the surfaces.

Abramov and McEwan (2004) tested linear interpolation, splining, nearest neighbour, and natural neighbour routines using regularly spaced Mars Orbiter Laser Altimeter (MOLA) data collected over the Martian Korolev crater, and a simulated MOLA dataset collected in Iceland. While the former analysis was strictly qualitative, the latter employed a DEM produced by the Icelandic Geodetic Survey, the heights of which were sampled to simulate a typical MOLA survey. The point data were interpolated at three spatial resolutions using the four algorithms, and then the new DEMs were compared to the original. Abramov and McEwan (2004) concluded that natural neighbour interpolation was not only the most accurate algorithm, but also generated the fewest visual artifacts. The most accurate spatial resolution varied by algorithm, with the high resolution surfaces being most accurate for natural neighbour and nearest neighbour, the medium resolution DEM for linear, and the low resolution DEM for splining.

Su and Bork (2006) tested DEMs generated from densely arranged lidar returns (densities averaged 0.75 points/m^2). Using a jack-knife approach similar to that of Lloyd and Atkinson (2002) and Su and Bork (2006)