Reanalyzing environmental lidar data for archaeology: Mesoamerican applications and implications

Charles Golden a,⁎, Timothy Murtha b, Bruce Cook c, Derek S. Shaffer d, Whittaker Schroder e, Elijah J. Hermitt f, Omar Alcover Firpi g, Andrew K. Scherer g

⁎ Corresponding author.
E-mail address: cgolden@brandeis.edu (C. Golden).

Available online xxxx

1. Introduction

In this paper, we present a preliminary archaeological assessment of extensive transects of lidar collected by environmental scientists over southern Mexico in 2013. These data are publically accessible, but to date have not otherwise been widely used for research by archaeologists. In particular, we report on the completion of a first phase of research, consisting of: 1) characterization and classification of the cultural and ecological context of the samples, and 2) bare earth processing and visual inspection of a sample of the flight paths for identification of probable anthropogenic Precolumbian features. These initial results demonstrate that significant contributions to understanding variations in Precolumbian land-use and settlement patterns and change is possible with truly multi-regional lidar surveys not originally captured for archaeological prospecting. We point to future directions for the development of archaeological applications of this robust data set. Finally, we highlight the potential for enriching archaeological research through tightly coupled collaborations with environmental science and monitoring. Archaeologists in the neotropics can acquire more data, better realize the full potential of lidar surveys, and better contribute to interdisciplinary studies of human-environmental dynamic systems through regionally focused and collaborative scientific research.

2. Data collection

Hank Margolis of Laval University, Quebec, Canada (currently at NASA Headquarters) collected Lidar data over extensive areas of Mexico in April and May 2013, using NASA-Goddard’s lidar, Hyperspectral & Thermal Imager (C-LiHT) system (Cook et al., 2013; Hernández-Stefanoni et al., 2015). The primary objective of the mission was to refine measurements of aboveground forest carbon stocks in Mexico (Fig. 1). In total, these data include 610 lidar samples ranging in size from 3 ha to 4100 ha (Fig. 2). These the flight paths offer remarkable transects of a wide variety of modern population centers and land use regimes in Mesoamerica. Of critical interest for archaeologists, they also broadly cover significant regions of Precolumbian settlement,
including unprotected and protected archaeological zones, such as the Montes Azules Biosphere Reserve in Chiapas, the Calakmul Biosphere Reserve in Campeche, and the Kiuic Reserve in Yucatan.

Although focused on the measurement of modern biomass, agricultural societies have shaped the landscapes crisscrossed by the flight paths for at least four millennia. The point-cloud data gathered includes regions in which Precolumbian settlement has not previously been documented, as well as detailed samples of areas with well-documented pre-modern occupation ranging in scale from small villages with evidence for ancient agricultural modifications and major political capitals. Here, we provide the preliminary results from a contextual analysis of the flight paths and visual inspection of a sample of the flight paths, relying on bare earth processing using the LIDAR ANALYST (Textron Systems, 2016) extension for ArcGIS 10.3. Our primary

Fig. 1. Above ground carbon density in tC/ha (Cartus et al., 2014; WHRC, 2016) and locations of LIDAR flight paths.

Fig. 2. Frequency distribution of the LIDAR flight paths. Samples range in size from 3 ha to >4000 ha. The majority of the samples (n = 490) cover areas of 500 ha or more.
purpose for this first set of visual evaluations was to document the diversity, scale and intensity of anthropogenic landscape features as well as potential challenges for using data not initially collected for such purposes for archaeological research. We conclude this paper with an overview of further research directions. We emphasize that this is a statement of preliminary research and future research opportunities. In coming years we will develop more fully and intensively the potential of this data set and others collected by environmental researchers for use in archaeology.

2.1. LIDAR applications in Mesoamerican archaeology

Over the past decade, investigators have transformed the analysis and use of lidar information in archaeology from an experimental approach to a well-proven methodology. Lidar provides an unparalleled capacity to document cultural landscapes from the air, accurately and precisely delimiting many anthropogenic features in a diversity of lowland settings, even beneath the tropical forest canopy (e.g., Carson et al., 2014; Chase et al., 2011, 2014a, 2014b; Doneus et al., 2008; Evans and Fletcher, 2015; Fernandez-Diaz et al., 2014; Hightower et al., 2014; von Schwerin et al., 2016). The impact of LIDAR has been particularly profound in the neotropics, where such data offer the promise for achieving the long-running archaeological goal of full-coverage survey, even in densely forested zones where traditional ground-based survey is impractical in terms of labor cost and time (e.g., Balkansky et al., 2000; Falconer and Savage, 1995; Kolb and Snead, 1997; Kowalewski, 1990; Parsons, 1990; Plog, 1990; Sanders, 1999; Sanders and Santley, 1983; Terrenato and Ammerman, 1996; Underhill et al., 2002; Wilkinson, 2000). The costs of an extensive lidar survey are undoubtedly significantly lower than decades of ground-based reconnaissance (Chase et al., 2011), and collaborative efforts by groups of archaeologists can reduce the per-site cost of data collection for each project. Nonetheless, given the current state of research funding opportunities mounting a lidar campaign and subsequent data processing remain prohibitively expensive for most archaeologists outside those countries where national data sets and robust systems of dissemination provide ready access.

Beyond constraining the broader application of lidar in the tropical Americas, limits on funding data collection has shaped sample selection of regions surveyed. Even though one clear benefit of lidar lies in the identification of archaeological sites and features in previously undocumented areas, data collection undertaken with archaeological questions in mind has typically centered research on improving the resolution and accuracy of maps of known sites—. In part, this situation is shaped by research questions, but it is also likely that lidar survey over previously undocumented areas may be seen as a potentially high-risk and low-return investment of limited research support by reviewers and funding organizations. As a result, despite the spectacular results of LIDAR surveys at research zones from Southern Chiapas (Rosenwig et al., 2013, 2014), to northern Yucatan (Hare et al., 2014; Hutson, 2015), to central Belize (Chase et al., 2011, 2014a and 2014b), the actual spatial coverage of lidar data produced by dedicated archaeological projects in southern Mesoamerica is regionally clustered. While collaborative conferences, publications and workshops that have been undertaken offer important opportunities for inter-regional comparisons, no lidar database collected for archaeological research in Mesoamerica offers inter-regional samples.

2.2. Opportunities and challenges for interdisciplinary collaborations

Of course, archaeologists are not the only investigators to mount lidar campaigns in the neotropics, and abundant lidar data are being collected by governmental agencies and environmental scientists as part of academic investigations, infrastructural development, and international conservation efforts focused on land use and land cover change (LULC), including the United Nations REDD + programs (Corp et al., 2015). The scale and intensity of these surveys vary considerably. Whether it is hazard modeling or forest structure measurements the usefulness of lidar for monitoring, documenting and modeling environmental systems is expanding, but with a small number of exceptions (see Weishampel et al., 2012), archaeologists have not fully leveraged the collaborative potential of such research relationships. These surveys are commonly coupled to other forms of environmental remote sensing, while macro-regional archaeological data is often not suited to rapidly compare biodiversity, biomass and LULC, and sustainability of modern ecosystems data. Data such as the G-LiHT material considered here present the opportunity to broaden the coverage of remote prospection of Precolumbian landscapes at much reduced cost, and better integrate anthropological and archaeological questions within environmental science (Challis et al., 2011). Our suggestion is that perhaps archaeology can be better coupled to the development of truly interdisciplinary remote sensing missions that concern human-natural systems dynamics, such as Weishampel and colleagues (Weishampel et al., 2012), instead of finding ways for archaeologists to use remote sensing for cultural site and feature identification exclusively.

Because investigators did not collect the lidar data discussed in this article with archaeological goals in mind, these and similar data sets present some challenges, some discussed in Section 4, for studying Precolumbian features of the landscape. Yet, we suggest that these data may also open up an opportunity to evaluate different approaches to archaeological sampling and transformative information about previously undocumented archaeological sites and regions. Because funding for environmentally-focused lidar coverage of areas not traditionally surveyed by archaeologists (well beyond known site centers) is not dependent on the surety of identifying ancient anthropogenic features, such data can provide important surveys of areas that are not otherwise known to present archaeological features. The sample areas surveyed exhibit selection bias determined by the research questions of the environmental researchers, and are thus not a statistically random sample of the landscape per se. Nonetheless, because sampled areas were not selected with known archaeological sites, or likely site locations, as part of the selection criteria, such sampling provides an opportunity to investigate areas not based on the known presence or prior documentation of archaeological sites.

While this is not a statistically random sample, it does offer a categorically different type of sample in terms of scale and location traditionally acquired by archaeologists (cf. Ammerman, 1981; Binford, 1964; Dixon, 1989; Gallant, 1986; Kellogg, 1987; Mueller, 1974; Orton, 2000; Plog et al., 1978; Redman, 1987). In this, it shares much with the implementation of the “site-less survey,” which has more typically been focused on the surface distribution of artifacts across the landscape (e.g., Dunning and Daniel, 1983; Dunning, 1992; Gallant, 1986; Rhoads, 1992; Wandsnider and Camilli, 1992). Here, too, the limits of the lidar survey were not predetermined by the problematic definition of archaeological sites as discrete and bounded cultural phenomenon. Instead of a focus on artifacts like traditional site-less survey, however, the sample collected may better reflect the distribution across the landscape of cultural features not otherwise encompassed by recognized sites; this is a distribution that may challenge and transform our understanding of settlement patterns and landscapes.

Indeed, the pros and cons of a non-archaeological focused survey are evident in these data. On the one hand, many flight paths missed known large archaeological sites (in some cases by less than 500 m). Further, in most cases where flight paths did intersect known large sites those they intersected those sites only in part. On the other hand, our preliminary analyses have already revealed important, and previously undocumented, anthropogenic features and variations in settlement within and between regions that would likely not otherwise have been evident if sample selection processes had favored the inclusion of known, and strongly suspected, site loci. These inspections also reveal extensive multi-site or regional features and critical perspectives on differences in inter-site settlement patterns, that would not be revealed by site
focused intensive archaeological sampling. In the Maya lowlands, such perspectives are rare (though see Canuto, 2002; Golden et al., 2008; Ford, 1986; Ricketson and Ricketson, 1937; Smith, 2000).

3. Mission background

The data analyzed and discussed in this article were collected in April of 2013 as part of a multi-institutional, bi-national study of above-ground biomass (AGB) and species-richness that covered large swaths of Mexico (Hernández-Stefanoni et al., 2015). The data were collected as part of deforestation reduction strategies including the United Nations REDD + (Reducing Emissions from Deforestation and forest Degradation, plus conservation, sustainable management of forests and enhancement of forest carbon stocks), and to aid in the design of effective strategies for selecting natural protected areas (see http://www.un-redd.org/aboutredd). Among the instruments employed for this mission was NASA-Goddard’s lidar, Hyperspectral and Thermal Imager (G-LiHT).

Bruce Cook leads the G-LiHT team at NASA-Goddard. G-LiHT was designed as a relatively inexpensive, robust and portable research tool for evaluating the potential benefits of data fusion for studies of terrestrial ecosystems (see Cook et al., 2013 for full details). G-LiHT is a multi-sensor airborne imaging system that includes LiDAR, Imaging Spectrometer and Thermal instrument intended to simultaneously map the composition, structure and function of terrestrial ecosystems. For the purpose of the present paper we are concerned with the data collected by the system’s lidar sensors.

The airborne laser scanning (ALS) instrument of G-LiHT is a VQ-480 (Riegl USA, Orlando, FL, USA) which includes a high-performance laser rangefinder and a rotating polygon mirror with three facets to deflect a 1550 nm Class 1 laser beam onto the ground. A user-selectable pulse repetition rate up to 300 kHz provided an effective measurement rate of up to 150 kHz along a 60° swath perpendicular to the flight direction, resulting in a pulse density of about 6 pulses m−2. A laser beam divergence of 0.3 mrad produced a 10 cm diameter footprint at the nominal operating altitude of 335 m. The small footprint laser beam allows detection of small gaps in the canopy and the ability to characterize fine scale disturbances, which are difficult to deconvolve from large footprint lidar waveforms. The mirror speed was set to a maximum of 100 rotations s−1 during G-LiHT acquisitions, whose points are spaced 0.23 m apart within a line perpendicular to the flight direction and 0.57 m between lines with a nominal aircraft speed of 110 km. Up to eight discrete ranging returns were identified and recorded for a given pulse. Three-dimensional lidar returns and user-friendly data products (see Cook et al., 2013) are openly distributed through the G-LiHT Data Center Webmap (http://gliht.gsfc.nasa.gov).

4. Contextual analysis of the flight paths

Six hundred and ten (610) samples were captured between April and May 2013 and processed by the NASA Goddard space center. The samples are unique in that even though individual sets of point-cloud data vary in size and distribution, as an interconnected group they transect key ecological and physiographic regions in the lowlands of Mexico (Fig. 1). While in the longer term we will investigate each sample set for archaeological details, we recognized that a project of this scale would first benefit from a macro analysis of the flight path data and preliminary analysis of a sub-set of data to build a foundation for more collaborative and detailed analysis in our second phase of research. Using global and national scale data, we have analyzed the flight paths according to three key categories:

1. **Known Archaeological Context**, including: Proximity to all known sites and Proximity to documented ‘large sites’.

2. **Core Ecological Context**, including: Forested Area, Soil Type and Use, Soil Erosion, and Proximity to rivers, streams and water.

3. **Modern landcover and administrative context**, including: 2010 Land Use and Land Cover (250 m), Protected Areas and Urbanized Areas.

The categories and the descriptive analysis are designed to examine the distribution and diversity of sampled regions.

4.1. Known archaeological context

The core purpose of this paper is to document the general characteristics and qualities of a publically distributed and free to the end-user, but non-archaeologically focused, lidar survey that spans large swaths of Mesoamerica. In future analyses and publications we will focus on improving archaeological site databases and developing a more robust landscape perspective that considers the inter-regional distribution of anthropogenic features beyond the scale of the site. Here, though, we limit our discussion to a clear description of how this sample overlaps with regional archaeological context. To evaluate the archaeological context of the flight paths, we first overlaid the flight path locations with information downloaded from Witschey and Brown (2010). While these locations have not been universally ground-truthed and verified, as a whole they provide the most accessible and reliable macro-regional database of known archaeological sites in the lowlands. Site ranking and categories are provided in these data, which was used for only cartographic purposes in this paper (Witschey and Brown, 2010). Our intent is not to evaluate these data using the lidar samples or to test their ranking, but to use these data to characterize the known archaeological context of these lidar samples. We performed four types of simple spatial measurements using these data in order to qualitatively characterize the archaeological context of each flight path.

First, all archaeological sites were ‘buffered’ using standard geoprocessing tools at 5 and 10 km. We chose these distances to quickly identify whether lidar samples were at least regionally covering areas with known archaeological context. For each flight path, any buffer that intersected a part of the flight path counts as a potential known site. Table 1 illustrates the results of our buffer-based analysis. Next, we identified the centroid of each flight path, identified the closest site, measuring its straight-line distance (in meters) (Fig. 3). Finally, we measured the distance from each flight path centroid to sites classified by Witschey and Brown (2010) as Type I and Type II (Fig. 4).

Five-kilometer buffers and ten-kilometer buffers exhibit similar frequency distributions. The majority of the flight paths have five or fewer known sites within 5 and 10 km. Some flight paths covered densely ‘known’ archaeological regions, but more than thirty percent (30%) of the flight paths do not have a single known archaeological site within 5 km of any section of the flight path. The large distances recorded from flight paths to known site centers highlights that the majority of these samples come from previously unstudied and understudied archaeological regions.

One hundred and eighteen (118) of the six hundred and ten (610) flight paths have a known archaeological site within 2 km of the flight path centroid. This number drops to under sixty or <10% of the entire sample with the intensely surveyed areas surrounding the site of Kiuic removed. Moreover, there is a clear and abundant majority of flight paths that are >5 km from a site recorded in the Witschey and Brown (2010) database. When the proximity and buffers are compared an even clearer clustering pattern emerges. Flight paths, with a previously documented site <2 km distant have a higher probability of multiple sites within 5 and 10 km buffers. This pattern may be the result of archaeological research intensity near documented sites and not necessarily a reflection of past site distributions; thus known sites tend to cluster, leaving large swaths of the landscape unsurveyed.

1 Witschey and Brown largely compiled site locations from a variety of readily available bibliographic sources (available here: http://mayagis.smv.org/MayaSites_Bibliography.pdf).
When flight paths are compared to Witschey and Brown’s (2010) type I and II sites, forty-three are located anywhere from 1.25 km to more than 200 km from each flight path centroid. Less than 10% are within 10 km of a type I and II site, while 209 flight paths or thirty-four percent (34%) have a type I and II site within 20 km. The majority of flight paths have a type I and II archaeological site within an average of fifty (50) km, but beyond 20 km. Perhaps such an approach is not as useful as a simple buffer analysis, because these metrics do not account for the fact that the lidar flights were restricted to Mexico, even though the closest large site may be in Belize or outside the lowland Maya region where fewer large sites have been classified or sites beyond the lowland Maya region. For example, several lidar samples cover areas near the Olmec site of San Lorenzo, which is not included in the Witschey and Brown (2010) database.

In general, the LIDAR flight paths exhibit a rather wide range of distributions in relation to known and unknown archaeological regions, but show a greater tendency to cover unknown or understudied regions of known and unknown sites. It is also reasonable to suggest that these data were collected in a way that avoids major sites and their associated site centers. Simply, the majority of the samples do come from previously un-surveyed and understudied areas and certainly regions that are lesser known from an archaeological perspective. In our estimation, at least 50% of the sampled flight paths have not been studied by archaeologists, even generally. Perhaps as many as 90% cover areas that have never been studied in even a mildly intensive manner.

One set of samples unique in this respect includes significant coverage of the Maya site of Kiuic, in Yucatan. These data were collected specifically to assess lidar data sets under leaf-on and leaf-off forest conditions for the purposes of estimating biomass and species richness, for characterizing these vegetation attributes in tropical dry forests. Kiuic lies within the Kaxil Kiuic Biocultural Reserve (http://www.kiuic.org/), which was the site of this study due to existing ground data and ongoing ground-based research in the reserve (Hernández-Stefanoni et al., 2015). Investigators acquired sixty-nine small samples adjacent to the site and reserve of Kiuic. These samples do not provide the extensive reach of the other flight paths, but further analyses of these data will offer another unique sample from this set for archaeologists to investigate how overlapping samples influence accuracy and precision of archaeological site identification.

### 4.2. Core ecological context

In addition to the archaeological context, we used several macro-environmental data sets of to try to characterize the variability of flight path contexts. Importantly, in this first step, we are not using these data to offer any critical interpretations about lowland demography, settlement or political ecology. Nonetheless, we recognize that these data offer incredible opportunities to address these questions and those questions will be at the foundation of our next phase of research. However, for that work we will need to acquire environmental data and develop models that are scaled appropriately to studies of demography,

<table>
<thead>
<tr>
<th>Frequency of LIDAR samples</th>
<th># of sites within 10 km</th>
<th># of sites within 5 km</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>101</td>
<td>226</td>
</tr>
<tr>
<td>5</td>
<td>239</td>
<td>254</td>
</tr>
<tr>
<td>10</td>
<td>66</td>
<td>89</td>
</tr>
<tr>
<td>15</td>
<td>34</td>
<td>27</td>
</tr>
<tr>
<td>20</td>
<td>66</td>
<td>6</td>
</tr>
<tr>
<td>25</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>30</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>35</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>40</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>55</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>60</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>65</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>75</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Fig. 3. Chart illustrating a count of the number of ‘known sites’ within a 5 km buffer of the flight path (y) compared to the linear distance (in meters) from the flight path to the closest known site. Chart illustrates that when samples are close to known sites there is a greater likelihood of additional sites recorded.
settlement patterns and political organization. Here, we chose macro scale variables to characterize the variability of samples from two perspectives: 1) macro data that may reflect the potential to influence past activities (prehistoric activities), and 2) data that may reflect visibility challenges and the processing of these data for archaeological documentation.

The data we selected are low resolution but regionally reliable and comparable throughout the sampled regions. In many respects these data are the environmental equivalent of the Witschey and Brown data; they are regionally accurate though they lack the precision that will be needed in future research focused on sites or sub-regional analyses of the lidar data. The variables are primarily focused on key aspects of forest structure, soil, topographic position and proximity to water. These core environmental characteristics have not only been shown to have a clear relationship with the form and distribution of past settlements in the lowlands (Gómez-Pompa et al., 2003; Sanders 1962, 1963, 1973 and 1977) but also should reflect the key variations in ecoregions in the neotropics.

4.3. Forested type

Using the Global EZ Level 2 – Global Ecological Zone data (FAO, 2012) we summarized forest type by flight path polygon the proportion of each flight path and forest type (Fig. 5). Five types of forests are found in the entire sample area and four are found in different proportions for each flight path. The five observed are: 1) Tropical Rain Forest; 2) Tropical Moist Deciduous Forest; 3) Tropical Dry Forest; and 4) Tropical Mountain System. Ninety-five (95) of the samples contain a majority of Tropical Rain Forest, while ninety-six (96) of the samples have at least some tropical rain forest in the sampled polygon area. Four hundred and forty (440) have a majority proportion of Tropical Moist Deciduous Forest, while an additional four samples include some Tropical Moist Deciduous Forest. Fifty-seven (57) samples have >50% Tropical Dry Forest, with four additional samples having Tropical Dry Forest presence. Seventeen (17) samples are covered by Tropical Mountain System while ten more have presence. The distributions are ‘gross’ considerations of forest type, but clearly transect important macro-regional differences in forest vegetation (Fig. 5). Importantly, all of the samples exhibit low diversity of classified forest types, thereby reflecting more of a thematic distribution of forest types. One forest type dominated most samples. Two or fewer forest types dominate all samples.

4.4. Soil properties

Key soil properties offer a different perspective from forest type. We looked at several factors, primarily related to basic soil type, erosion and soil humidity to characterize the sampled flight paths (see Fig. 7; INEGI, 2004). The flight paths broadly sample fourteen major soil types, with the majority from Rendzina, Litosol and Luvisol orders, but a broad distribution to representative orders from throughout the lowlands (Fig. 6). From the perspective of erosion, the flight paths are distributed across three general areas. Ten of the sample majority area covers regions that are unclassified; however, of those ten cases, six have some percentage of coverage without appreciable erosion. Only five of the sample areas contain a majority of soils with appreciable erosion. Areas without appreciable erosion cover the majority of sample areas (595).

4.5. Water

Using INEGI 1:1,000,000 data (INEGI, 2000) we measured the distance from each flight path centroid to nearest inland water feature.
These data are illustrated in Fig. 9. Flight paths vary in distance to inland water feature from those that are directly adjacent to key inland water features to (90) km distant. Most importantly, these data show a rather normally distributed sample with roughly 50% < 20 km from an inland water source and beyond 20 km. These data and the scale especially, capture the drastic differences in water availability and access.
Fig. 7. Distribution of lidar flight paths and streams and rivers (source INEGI, 2000). While many of the samples in the northern Yucatan are located far from surface water, samples in the southwest are located adjacent to streams and rivers.

Fig. 8. 2010 Land cover (source data NALCMS, 2010): comparison between forested areas, shrubland and cropland. Lidar sample areas are ranked in descending order from samples with the highest percentage of forested area to the lowest percentage of forested area covered by the flight path. The Fig. illustrates the inverse relationship between samples covered by forested area and those covered by crop and shrubland.
Fig. 9. 2010 Land cover (source data NALCMS, 2010): comparison of forest types. Lidar sample areas are ranked in descending order from samples with the highest percentage of tropical or sub-tropical broadleaf evergreen forest to the lowest percentage covered by the flight path.

Fig. 10. Lidar flight path distribution as compared to urban areas (INEGI, 2014) and Protected Natural Areas (CONANP, 2016). A small percentage of samples were captured in protected natural areas.
throughout the sampled region. Surface water is scarce throughout the Yucatan, while it is abundant throughout Chiapas, Vera Cruz and parts of Campeche.

4.6. Modern landcover and political and administrative areas

The purpose for this analysis was again to simply characterize or qualify the context of each sample, recognizing that modern and recent variables will have a significant impact on the processing and inspection of these LIDAR data. Similar to the environmental variables described above, we acknowledge that while these factors are critically important in our next phase of research, we will need more precise and accurate data to fully investigate the methodological implications of these factors for archaeological site and feature visibility. Three landcover types dominate the sample areas.

The data are summarized using the ‘isectpolytorast’ tool in the Geospatial Modeling Environment (Beyer, 2012). For each sample, a proportional measure of each landcover type (NALCMS, 2010) is calculated for each flight path polygon. The following land cover summary types are: 1. Temperate or sub-polar needleleaf forest, 2. Tropical or sub-tropical broadleaf evergreen forest, 3. Tropical or sub-tropical broadleaf deciduous forest, 4. Temperate or sub-polar broadleaf deciduous forest, 5. Mixed forest, 6. Tropical or sub-tropical shrubland, 7. Temperate or sub-polar shrubland, 8. Tropical or sub-tropical grassland, 9. Temperate or sub-polar grassland, 10. Wetland, 11. Cropland, 12. Barren land, 13. Urban and built-up, and 14. Water. To characterize landcover context, we combined all forest types (categories 1–5), Shrublands and Wetlands (categories 6–10) and Developed areas (categories 12–13). These three major categories exhibited the most significant distributions in the sampled areas. Fig. 8 illustrates these differences. Although the majority of the sampled areas have a significant percentage of forested areas, a great deal of diversity of coverage is evident in these data. Croplands and Shrublands exhibit a diverse proportional presence for more than two-thirds of the sample areas. Moreover, Shrublands and Croplands are the majority land cover for one-third of the 610 sampled areas. In Fig. 9, we illustrate the proportional distribution of forest type as characterized in the 2010 LULC data (NALCMS, 2010). These data demonstrate a similar pattern as that observed for the Global forest data, reinforcing the observation that tropical evergreen and deciduous forests cover the majority of the sample areas.

4.7. Protected areas and urbanized areas

Only fifty-five of the sampled areas fall within the boundaries of areas defined by CONANP as Protected Natural Areas (Fig. 10). Forty-seven (47) have overlap with the subcategory of Biosphere Reserve, while three are in National Parks, and five are located in an ecological conservation area. > 90% of the flight paths fall in areas outside of protected zones in the lowlands; however, many are near and adjacent to these important protected areas. All of the sampled areas avoided areas characterized as local urban zones, which is not surprising. What is interesting is the proximity of some flight paths coupled to the entirety of the sample as a useful means for evaluating the role of population growth and urbanization pressures on archaeological site preservation in our next phase of research. Many of these sampled areas are undoubtedly impacted by their proximity to urbanized zones and transportation development that has been rapidly occurring in the lowlands, especially on the Yucatan Peninsula. Similarly, the integrity of these forested areas as above ground carbon stocks may face challenges due to urbanization and population growth.
5. Visual inspection of flight paths

For this initial evaluation we also systematically investigated thirty-one (31) flight paths from a variety of ecological, archaeological and cultural contexts in two select regions. We selected these opportunistically based upon our field-work based knowledge of the regions, and ability to conduct brief site visits to verify the visual inspection. We focused our analysis on:

1. The Central Yucatan

2. Chiapas

For each of the flight paths, we derived sub-meter pixel resolution bare earth digital elevation models using a batch bare earth process in lidar Analyst, and processed all 610 of the samples following identical algorithms. Estimated point spacing and DEM/Hillshade resolution are reported in Table 2. Because the model used to measure above ground carbon storage separates lidar points into two categories: 1) Above ground (trees and vegetation) and 2) Bare Earth, there is a significant benefit to repurposing these environmental data here. In the automated
process, estimated point density for each sample is calculated and a bare earth digital elevation model was calculated (see Table 2 for sample point spacing and dem resolution). Using ArcGIS 10.3, hillshades for DEM were batch calculated using a default azimuth of 315° and a default altitude of 45°. In future studies, we will systematically investigate how different processing techniques influence outcomes, but here we were focused solely on exploring the potential of these samples. Using the hillshade and draped DEM, potential archaeological features were recorded and reviewed by at least two members of the research team.

For all of the digital elevation models, we used the last returns and filtered bare earth points. A number of papers have recently described new techniques for processing and visually inspecting lidar data for archaeological purposes. For this phase of research, we wanted to develop an efficient process for interpolating bare earth surfaces and calculating a comparable hillshade for each sample. Clearly, there are future benefits for using these data to investigate how a variety of ecological factors can influence and potentially inform the processing of DEMs and hillshades for archaeological purposes. Here, we aimed only to provide

Fig. 13. Digital elevation model and hillshade of lidar sample s450, near the site El Kinel. Several structures and features were visible in the sample corresponding to field observations (AH – Aguada (historic); D – Depression; M – Mound; MG – Mound group; P – Patio group; T – Terrace; X – Unknown feature).
a first view and comparable view of the potential usefulness of these data in two regions. We summarize some general archaeological observations about the flight paths in Table 2. Archaeological features particularly those we infer to be masonry platforms and terraces, are abundant, visible and widely distributed in both regions. However, the density, form and distribution of these features varies considerably and reflects differences in ecological and archaeological context.

5.1. Central Yucatan description

The fourteen samples in the central Yucatan are located near sites, like Becan, Xpujil and Kohunlich and a little more than 50 km northeast of the Calakmul Biosphere Reserve (Fig. 11). All of the samples contain evidence of archaeological mounds, most of the samples contain evidence of formal plaza groups, and two exhibit evidence of monumental architecture (Fig. 12). Generally, house mounds and plaza groups are evenly dispersed throughout the uplands. In several of the samples, the upland settlements extend into the more marginal lowlands (Fig. 12). While isolated house mounds are more prevalent than plaza groups, substantial landesque or landscape features accompany both forms of pre-Columbian architecture. Perhaps the most striking observation recorded is the density and distribution of terraces and field boundaries in the uplands (Fig. 12). While these observations are consistent with Turner (1974, 1983) and more recent studies (Lemonnier and Vannière, 2013), the extent, scale and intensity of the terraces and field boundaries suggest an extensive regional distribution, accompanied by variable densities. The striking intensity matches known terraced regions from western Belize, despite very different patterns of form and construction (Chase and Chase, 1998; Murtha, 2015). Field boundaries are widespread and positively correlate with terrace construction. Another significant observation from these samples is the widespread presence of household or small-scale aguadas. At least 50% of the sampled areas exhibit some evidence of water management in the form of retention.

5.2. Chiapas description

The seventeen samples in Chiapas are located west and south of the Usumacinta River, in and around the Montes Azules Biosphere Reserve. These data encompass several previously documented archaeological sites including El Kinel (Guatemala), Nuevo Jalisco, Benemerito de las Americas, and El Palma (Golden and Scherer, 2006; Scherer et al., 2007; Tovalín Ahumada and Ortiz Villareal, 2005; Tovalín Ahumada et al., 2004; Velazquez Valadez, 1986). Although sixteen out of seventeen samples contain evidence of archaeological mounds, the majority of monumental architecture and plazas are restricted to known sites in the area, including the aforementioned Benemerito de las Americas and El Palma. Additional monumental architecture, previously undocumented, is also evident to the west of El Kinel, to the south of modern Frontera Corozal, and in two isolated portions of the Montes Azules Biosphere Reserve, though the latter settlement may be associated with sites registered by Blom and Healey (Ekholm, 1992).

Archaeological mounds are predominantly located on elevated terrain, evidenced on hilltops or at the base of hills. In areas of higher settlement, these mounds tend to form discernible patio groups of three to four structures. A significant observation in this area is the amount of terracing, both in proximity to large centers as well as dispersed regionally. Rectilinear landscape modifications also accompany dense settlement in some areas. To the south, in the Montes Azules, terraces are mostly concentrated on the gentle slopes of hills, leaving a subtle

### Table 2

**Summary of visual inspection.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>s412</td>
<td>0.33</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s413</td>
<td>0.336</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s414</td>
<td>0.348</td>
<td>111</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s415</td>
<td>0.335</td>
<td>30</td>
<td>N</td>
<td>N</td>
<td>y</td>
<td>N</td>
</tr>
<tr>
<td>s431</td>
<td>0.345</td>
<td>11</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s432</td>
<td>0.375</td>
<td>113</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s433</td>
<td>0.353</td>
<td>80</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s435</td>
<td>0.255</td>
<td>165</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s436</td>
<td>0.287</td>
<td>63</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s452</td>
<td>0.334</td>
<td>197</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s453</td>
<td>0.299</td>
<td>111</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s454</td>
<td>0.309</td>
<td>142</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s455</td>
<td>0.328</td>
<td>220</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s456</td>
<td>0.366</td>
<td>133</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s441</td>
<td>0.329</td>
<td>8</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>s442</td>
<td>0.331</td>
<td>5</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s443</td>
<td>0.344</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s444</td>
<td>0.312</td>
<td>14</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s445</td>
<td>0.321</td>
<td>16</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s446</td>
<td>0.312</td>
<td>11</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>s447</td>
<td>0.349</td>
<td>6</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>s448</td>
<td>0.341</td>
<td>3</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s449</td>
<td>0.367</td>
<td>106</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s450</td>
<td>0.39</td>
<td>102</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s457</td>
<td>0.359</td>
<td>183</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s458</td>
<td>0.351</td>
<td>91</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s459</td>
<td>0.341</td>
<td>20</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s460</td>
<td>0.304</td>
<td>7</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s478</td>
<td>0.271</td>
<td>26</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>s479</td>
<td>0.324</td>
<td>10</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>s469</td>
<td>0.417</td>
<td>139</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>
mark picked up by the lidar. These terraces are likely to be anthropogenic features as they occur primarily on such gentle slopes, follow the orientation of landforms, and in spite of consistent forest cover do not occur in all tiles as would be expected if they resulted from processing errors. One of the surprising characteristics of these terraces is that though each individual construction appears to be small in scale, collectively the terraces cover vast areas between settlements. In fact, many of these terraces have few to no archaeological mounds directly associated with them.

Further research will address why the extent of terracing is much greater than that number of mounds expected in this area given comparison with the data from Yucatan and other regions covered by this LIDAR mission. This may reflect real inter-regional differences in Precolombian population levels, or a briefier and/or more sporadic history of occupation than in the areas of Yucatan covered. Another explanation, however, may be that of sampling error, if flight paths consistently missed clustered settlement adjacent to the terracing. Alternatively, if structures in this region were largely unmounded, pole and thatch they would not appear in the data. So to, if in the more broken topography of the Usumacinta River basin the ancient residents reserved gentle slopes for terracing, preferring to place their houses in less arable spaces atop hills or cliffs (as Golden and Scherer have recently observed at the regional capital of Piedras Negras, Guatemala) these may have been missed or lost in data collection due to problems data shadows resulting from a single, relatively low-altitude pass of the G-LiHT instrument. Perhaps the most parsimonious explanation is that the dense canopy in terraced areas, and particularly the dense underbrush of secondary forest, obscured the documentation of low house platforms.

Additional features include evidence of water retention, which though limited, takes the form of aguadas, as well as a dam and water catchment area southeast of El Palma. Unlike the Central Yucatan samples, the Chiapas data suggest that settlement density decreased at the periphery of larger centers. This of course could be a sampling issue, because little is known east or west of these G-LiHT transects. Yet the abundance of landscape modifications, such as terraces and water management features, suggest that despite the lack of archaeological mounds visible in the data, the landscape was extensively modified and populated in antiquity.

Site Visits and Verification Golden and Scherer have conducted long-term, regional archaeological reconnaissances in the Usumacinta River Basin of Chiapas with the participation of Schroder and Alcover in recent years. This work, much of it carried out prior to the collection of G-LiHT data in April of 2013, provides a ground-based assessment of some of the lidar data. In 2006 and 2007, Golden, Scherer and colleagues made a topographic map using a Total Station, which was georeferenced using a Magellan Mobile Mapper 6, of much of the site of El Kinel, Guatemala (Golden and Scherer, 2006; Scherer et al., 2007). El Kinel appears in the tile identified as AMIGACarb_Chiaps_1_GLAS_Apr2013_s450_dtm (Fig. 13) located in largely denuded agricultural fields currently under a mixed regime of relatively low-height crops including maize, beans, and squash. Alignment of the Total Station map with lidar data confirm the accuracy and precision of the G-LiHT instruments under these conditions. All mapped structures and other ground features, including “borrow pits” that were excavated in antiquity to provide the building materials for mound platforms, are evident in the lidar data. Because the lidar data extend to areas that were not mapped on the ground, these newly acquired data reveal additional settlement and extend the map of the site. Given the existing data set, it is reasonable to assume that the features identified as residential mounds in the lidar data do indeed represent anthropogenic landscape modifications.

Brief site visits to the site of El Palma by Roberto Velazquez Valadez (1986), and Benemerito Primera Sección and Nuevo Jalisco by (Alejandro Tovalin and Víctor Ortiz 2005; Tovalin Ahumada et al., 2004) provide measured tape-and-compas maps of major architecture at those sites. The architectural core of El Palma conforms to the architectural complex visible in AMIGACarb_Chiaps_1_NFL_Apr2013_s469_dtm, while Benemerito Primera Sección appears in AMIGACarb_Chiaps_1_GLAS_Apr2013_s457_dtm, and Nuevo Jalisco in AMIGACarb_Chiap_Campeche_NFL_Apr2013_i0473_dtm. In 2015 and 2016, Golden and Scherer made brief visits to El Palma and Benemerito Primera Sección. While these visits accompanied by local landowners did not involve formal mapping, they did involve preliminary confirmation of anthropogenic features evident in the lidar data and the confirmation of some feature locations with GPS.

These visits did not confirm on the ground all of the anthropogenic structures identified in the analysis of lidar data. However, we argue they are sufficiently to strongly support the hypothesis that all anthropogenic features identified in the lidar data are in fact Precolombian constructions, and that false positives are minimal. The lidar data provide more complete, accurate and precise maps of site cores at these three centers than was provided by preliminary on-site visits, as well as providing important new data concerning the distribution of ancient settlement surrounding the monumental architecture. In addition to residential structures, terraces and water management features are also evident.

However, in the Chiapas, some samples, it seems as if we cannot identify some features from the lidar data alone, suggesting that in some samples false negatives may be a more pressing concern than false positives when identifying anthropogenic features from select samples. Of particular significance when considering the use of these data, the site of Nuevo Jalisco is currently under high canopy forest, but with significant undergrowth. This highlights one of the challenges of working with data not collected with archaeological site and feature identification in mind. Although the flight line was sufficient to address the biological and environmental questions of the researchers involved in the 2013 mission, the lidar system was challenged in this single pass to acquire enough ground points to fully resolve the major architecture at Nuevo Jalisco, though what structures are evident conform closely to Tovalin and Ortiz’s ground map (2005). More problematic is that we cannot identify smaller structures on the landscape surrounding Nuevo Jalisco in the lidar data due to lack of sufficient ground points. Perhaps, as Keith Prufer and colleagues have demonstrated, landscape features are a more effective means to document settlement patterns in areas with low and dense vegetation (Prufer et al., 2015).

6. Summary and conclusions

These data are remarkable in offering the most extensive lidar transect of modern population distribution and land use zones available to Mesoamerican archaeologists offering an unparalleled comparative data set. The paths transect critically important and under-documented cultural landscapes, potentially shaped by four millennia of land use since the first sedentary, agricultural communities were established offering a window on human decision-making, agricultural intensification and political evolution in a wide range of ecological, topographic, and hydrological zones. Beyond mere site prospection, the reanalysis of these data over the long-term as a collaborative effort among archaeologists, remote sensing specialists, and ecological scientists, can offer: 1) important collaborative insight into the evolution of these landscapes, and 2) the coupled human and natural dynamics influencing their form, distribution and ecosystem health in the past and today, at a uniquely macro regional scale.

6.1. Flight paths and archaeological data

This process actually revealed a compelling issue and opportunity for archaeologists. Witschey and Brown’s (2010) efforts to compile site location data should be recognized for its value to modern archaeological research. We are suggesting that as a discipline we need to try to find secure, standardized and more integrated methods for site location archives. As more remote sensing data becomes available from public and private sources in areas with sensitive archaeological data, even
as software for interpreting these data become more user-friendly, high precision data concerning site locations will inevitably become more widely accessible. We will need rapid and reliable methods for comparing site location data with high-resolution environmental data. These data reflect a broadly distributed sample from the perspective of archaeological context. Samples are adjacent to known sites as well as in areas not previously surveyed. A clear challenge emerged from this macro level analysis. While Witschey and Brown (2010) offer the most comprehensive database compiled for known sites in the lowlands, their efforts relied on data from a variety of sources and there are clear issues of reliability and precision. The broader issue of reliable and accurate archaeological data management needs to be addressed so that these data can be integrated into remote sensing efforts and flight plans with precision. Further, just as scholars have called for and shown the benefit of site-less surveys based on the surface distribution of artifacts across the landscape, remote-sensing (and particularly LIDAR) data at the inter-regional scale offer the opportunity to move beyond the constraints and problems of understanding human activities within the bounded notion of the site. They offer the opportunity to re-focus instead on features distributed without such predetermined boundaries.

6.2. Next steps

The data analyzed to this point represent the results of a preliminary stage of research focused only on a small subset of the data collected by Cook and colleagues. These results thus merely hint at the research potential of these lidar transects. Because of the scale and the scope of these data, the next phase of research will we undertake requires a more expansive effort among a larger group of scholars that includes archaeologists and natural systems scientists expert in each bio- and cultural zone encompassed by the flight paths. Only such a collective effort can document the full sample for evidence of prehistoric, historic and modern land use patterns, households and archaeological features.

6.3. Final thoughts

Our primary conclusions in this paper based on our basic context analysis and limited visual analysis, are:

1) There is an important anthropological and archaeological opportunity to collaborate with environmental science and earth observations, even if or perhaps especially if the samples are not tied to specific sites and documented regions (Challis et al., 2011);

2) Existing lidar data, even when not acquired with archaeological purposes in mind, can provide important different perspectives on archaeological sites and regions at low cost, particularly when sampling understudied regions;

3) As valuable as they are, these data align site-focused archaeological perspectives and even the process of archaeological permitting, thereby necessitating a uniquely collaborative and regional archaeological research design at multiple scales; and

4) As more sophisticated and rapid earth observation technologies are developed for monitoring the health and well-being of the planet, anthropology and archaeology as disciplines can and should contribute to the analysis and interpretation of these data.

5) While cultural and ecological regionalism has a strong tradition in archaeology and anthropology, site specific sampling has limited our ability to contribute to broader regional interpretations and archaeological theory. In the lowlands, a clearer regional settlement ecology has emerged even from preliminary re-analysis of these data.

In the second phase of research we will further develop and establish data and processing standards to be used by each member of the research group, along with standardization in annotation that take into account the variable contexts of the samples. This larger project will provide a unique perspective window for addressing long-standing questions about the settlement and cultural ecology of the lowland Maya. The prospects of our multi-year multi-disciplinary project will address broader questions of long-term sustainability and landscape resilience. While much work needs to be done, this preliminary analysis (this pilot study) has already provided important new information and perspectives on the regional settlement ecology of the lowlands.

Acknowledgements

Field research at El Kinel, Guatemala was conducted by the Sierra del Lacandón Regional Archaeology Project, directed by Charles Golden, Andrew Scherer, and Rosaura Vasquez, with support from the Heinz Foundation, The National Geographic Society, Wagner College, Baylor University, the Theodore and Jane Norman Fund and the Jane’s Fund from the Latin American and Latino Studies program at Brandeis University. All research was conducted with the permission of the community of La Tecnica and the Instituto de Antropología e Historia of Guatemala. Field research in Chiapas, Mexico described in this article conducted by Golden and Scherer takes place thanks to a permit from the Instituto Nacional de Antropología e Historia of Mexico, with support from Brandeis University and Brown University. Portions of this research reported here was supported by The Hamer Center at Penn State and The National Science Foundation (CNH1211809). Lidar data were collected by Dr. Hank Margolis (NASA Headquarters, Program Manager, NASA Terrestrial Ecology Program) with support from a NASA Carbon Cycle Science award to Ross Nelson (Program Announcement Number NNH10DD001N-CARBON).

References


