Guidelines for coastal lidar

by Melanie Lück-Vogel, CSIR, Chris Macon, US Army, and Lauren Lyn Williams, DEA

This article provides guidance to technical and non-technical coastal practitioner regarding the required precision of airborne lidar data for the typical coastal applications, output formats and value-added products. It also touches on issues of flight campaign planning and a national coastal lidar inventory.

Highly accurate and detailed topographic information has been identified as being crucial for addressing a wide range of coastal management issues at all levels of government, from environmental management to coastal engineering. These include, for example, assessment of coastal vulnerability to sea level rise and sea storm surge hazards, the establishment of coastal management lines and/or development setback lines, development and maintenance of ports and shipping routes as well as informing coastal planning.

Internationally, airborne light detection and ranging (lidar) technology has been identified as a cost-effective solution for obtaining high-resolution topographical information and is widely employed in the coastal zone on a routine basis. However, in the South African context, while the overall value of lidar is widely acknowledged and recognised, the acquisition costs are still high. Therefore, the application of the data is largely limited to geospatial practitioners and/or engineers and has not yet been fully explored for large scale operational coastal management.

Lack of funding frequently limits lidar's application in the coastal zone to once-off lidar coverages of relatively small areas and for use by well-resourced clients and/or authorities. However, the coastal zone would benefit immensely from repeat surveys, particularly considering the dynamic physical processes which can change the coastal landscape literally overnight, such as storm events. While lidar's initial investment might be more costly than aerial photography, for example, it might be cheaper in the long term as it allows for a host of additional analyses that more conventional datasets do not cater for, e.g. volumetric analyses to detect the extent of (illegal) sand mining.

The uncoordinated acquisition thus resulted in a patchwork of topographic coastal lidar coverages with large gaps in-between and in certain cases probably unintentional duplication of effort. Further, analysis of the existing data showed that the technical specifications of the various datasets frequently differ, which makes it difficult to compare products and to draw conclusions from disparate datasets. The non-uniformity also poses challenges on the derivation of nation-wide lidar products and services, such as the related Coastal Flood Hazard Online Decision Support Tool which is being developed by the Department of Environmental Affairs (DEA) for its Oceans and Coastal Information Management System (OCIMS) (https://ocims-dev.dhcp.meraka.csir.co.za/ocims-coastal-hazard/).

Fig. 1: Participants during the third Coastal Lidar Workshop held at the CSIR, Stellenbosch in December 2017.

Discussions with stakeholders and lidar data providers showed that the reasons for the large variability of the specifications, coverage and precision of the existing lidar data are usually vague terms of reference due to limited technical expertise on the client's side and limited understanding on the data provider side of the specifications needed for the client's intended applications. This leads to frustration on both sides and waste of money if unsuitable data resulted from this mismatch.

In order to overcome these limitations, the coastal stakeholder community established the Coastal Lidar Community of Practise (CoP) in 2014, coordinated by the CSIR Coastal Systems Research Group and consisting of coastal practitioners, representatives from the three spheres of government, research institutions and various lidar consulting companies. The primary objective for establishing the CoP was twofold, namely to provide the users with better understanding of lidar technology and constraints and to provide lidar providers with

Field of application	Feature/purpose	Absolute vertical accuracy (cm)	Sigma	Repetition
Coastal zone management	Inform the establishment of coastal management lines	5	1	Once off
	Assist in identifying the dimensions of coastal structures, e.g. building height **	20	1	
	Assessing coastal structures before and after storm events	20	1	After events
	Land use planning **	20	1	
	Beach erosion/accretion	10	1	
	(Illegal) sand mining	20	1	Biannual
Estuary management	Berm height	10	1	Biannual for two years, then every 3 years
	Volume of estuary and floodplains	5	1	
	Mouth and channel dynamics	20	1	
	Biomass assessment	10	1	5 to 10 years
	Shifts in critical habitats **	20	1	
Risk and vulnerability modelling (coastal engineering)	Wave run up and coastal flooding	10	1	
	Foredune sizing and integrity assessment	20	1	
	Vegetation structure	20	1	
Disaster management	Topography and surface (including infrastructure height)	5	1	
	Erosion and inundation areas	10	1	After events
Conservation	Detection of certain alien vegetation **	20	1	
	Habitat mapping **	20	1	
	Coastal vegetation resilience **	20	1	
	Vegetation structure	20	1	
Navigation	General charting *	50	2	
	Ports and harbour construction, e.g. breakwaters	10	1	
Design	Port and coastal structures	10	1	
	Harbour bathymetry *	50	1	Quarterly for 3 years
Geology	Tectonics	50	1	
	Submarine sediment dynamics *	50	1	
	Morphology of reefs and rocky areas	50	1	

*Table 1: Technical minimum specification with regards to accuracy and resolution. (*Features derived from bathymetric lidar sensors; **Applications combining lidar and other datasets.)*

a better understanding of the user needs. A further aim was to identify for which coastal areas lidar data exist. To date, three CoP workshops have been held (25 to 26 March 2014, 21 April 2015 and 5 to 6 December 2017, attended by 46, 22 and 17 participants respectively).

In order to grow the knowledge beyond what is currently available nationally, international coastal lidar experts from the US Army's Corps of Engineers' Engineering, Research and Development Center (USACE ERDC) and their Joint Airborne Lidar Bathymetry Technical Center of Expertise (JALBTCX) participated in the workshops. Further international coastal and nearshore lidar experts participated during the 2015 Workshop which was held back-to-back with the International CoastGIS Symposium in Cape Town.

The latest workshop held in December 2017, attended by participants from the DEA, Department of Rural Development and Land Reform's Eastern Cape Surveyor General and Chief Directorate: National Geospatial Information, Western Cape Government: Department of Environmental Affairs and Development Planning, Stellenbosch University, Nelson Mandela University, Council for Geoscience, Santam, ERDC and CSIR, revealed that there is still great insecurity amongst lidar users and government entities on defining the specifications in the Terms of Reference for coastal lidar acquisition. The participants therefore decided to review and expand the specifications which were drafted during the 2014 workshop and to share them with the wider geospatial community.

This article thus aims to provide guidance to technical and non-technical coastal practitioner regarding the required precision of airborne lidar data for typical coastal applications, output formats and value-added products that should ideally form part of coastal lidar deliverables as well as metadata standards that should be considered for compliance with national and international ISO data standards. Issues around flight campaign planning and a national coastal lidar inventory are mentioned as well.

Application, accuracy and precision

During the workshops, minimum requirements in terms of absolute vertical data accuracy, precision in terms of sigma as well as repetition frequency for the most common coastal applications were specified The results are summarised in Table 1. Features which are usually derived from bathymetric lidar sensors (which, to our knowledge, are not yet available in South Africa) are indicated with an asterisk (*). In some cases, e.g. for most conservation applications, lidar is not seen as a stand-alone technology for the mapping, but usually used in combination with other data such as multispectral satellite or airborne data. Those applications are indicated with (**).

All specifications listed above are optimal-case values. It was however recognised that vertical accuracies better than 20 cm might be difficult to achieve. It was agreed that 20 cm vertical accuracy would be acceptable in most of those cases where currently lower (i.e. more accurate) values are indicated.

With regards to the envisaged repetition rates, i.e. how frequently the dataset should be updated for the respective applications, detailed guidance is given only for a few applications. For most of the other applications, annual replications would be desirable, if feasible.

Output formats and value-added products

Recent experiences with projects that required the ingestion and analysis of lidar-derived elevation data from different sources showed that there is a great variance in terms of format of the point cloud data and the coding of the actual elevation data. Table 2 gives an overview of the most important recommendations to standardise future acquisitions.

In many cases the original lidar-derived point clouds are of little value to the user, with the files being too big and requiring special software for viewing/editing. Therefore, value-added products such as digital surface models (DSMs) in raster format or in the form of topographic contours are frequently required, too. Further, from the LAS format the additional extraction of (basic) land cover information is possible. We recommend the user specifies required standard value-added products in the Terms of Reference accordingly. Typical value-added products are listed in Table 3.

Those products will suffice for most coastal management application requirements. However, certain coastal applications might benefit from additional value-added products, e.g. information on the vegetation structure, a normalised surface model (which will give building and/or vegetation height) and/or the estimated surface area of estuaries at 10 cm intervals. These non-standard products are currently either generated by the original data provider on request or, where capability exists, by the data owners themselves. Further, usage of lidar intensity values is becoming more and more interesting in the research domain (which will require additional documentation on flight geometries from the data provider).

General acquisition considerations

Acquisition time

Several coastal applications require topographic lidar data acquisition of the intertidal area as well (see Table 1), especially since bathymetric lidar systems dedicated to submerged sea-floor mapping are currently not available in (southern) Africa. It is therefore recommended to time acquisition flights to be done at low tide, if possible at spring-low tide.

<table>
<tr><th>Criterion</th><th>Recommendation</th><th>Comments</th></tr>
<tr><td>Data format</td><td>Current approved LAS format</td><td>LAS format is superior to XYZ ASCII format, as the LAS format contains additional information on the received points, such as the intensity (which can provide additional information on the surface type), number of returns per point and RGB colour. In cases where XYZ format is required explicitly by the user, please provide both.</td></tr>
<tr><td>Coding of the elevation values</td><td>In meters, i.e. if a point is 1,35 m high, give the value as 1,35</td><td>This will require the elevation to be coded in float (or 32 bit) format. Avoid forcing data to integer values, e.g. by multiplying with 100 (i.e. 1,35 m would appear as 135). This stretching will cause confusion when different datasets are used in one application.</td></tr>
<tr><td>Vertical datum of the elevation data</td><td>Preferably give both: mean sea level and land levelling datum/ellipsoid</td><td>This will require either calibration with synchronous ground survey or linkage to trigonometric beacons.</td></tr>
<tr><td>Naming of the individual LAS tiles</td><td>Please do not name and number the tiles randomly but stick to a clear continuous numbering scheme.</td><td>Either number tiles in reading mode, i.e. start with 001 for the upper left tile and continue from there:
<table>
<tr><td>001</td><td>002</td><td>003</td></tr>
<tr><td>004</td><td>005</td><td>006</td></tr>
<tr><td>007</td><td>008</td><td>009</td></tr>
</table>
or use consistent row-and-column systems:
<table>
<tr><td>1001</td><td>1002</td><td>1003</td></tr>
<tr><td></td><td>2002</td><td>2003</td></tr>
<tr><td>3001</td><td>3002</td><td></td></tr>
</table>
This will help the user to more quickly identify tiles.</td></tr>
<tr><td>GPS timestamps</td><td>Acquisition date and time to be added as extra column to LAS data.</td><td>This will help in later data analysis (e.g. related to vegetation cover/phenology or tidal conditions), especially if data are flown over a longer time period.</td></tr>
</table>

Table 2: Output format of the point cloud data.

Name	Description	Comments
Elevation contours	Polyline shapefile (.shp) (vector)	Contour interval according to user specs (for coastal applications usually 10 cm)
Digital surface model	This includes the top-of-canopy/rooftop elevations	Typical output raster format would be GeoTIFF or geocoded JPEG2000. The user needs to specify the output pixel size. As a rule of thumb, the pixel size should be half the size of the smallest object of interest, i.e. if you are interested in tracking coastal access of 2 m width, aim for a pixel size of 1 m. NB: the minimum possible pixel size might be limited by the point density of the point cloud. If the desired pixel size is smaller than the distance between the points, there will be no-data pixels in the resulting raster image, i.e. data gaps. Consider this in the Terms of Reference in terms of minimum point density per m^2.
Digital terrain model	This includes ground level elevation only, i.e. buildings and vegetation canopy elevations are removed	As for digital surface models
Land cover classification	The LAS format allows the embedding of an additional data "column" for land cover classification	Classification of the data by the data provider in at least broad land cover classes (e.g. water, bare ground, vegetation) might help with the later interpretation of the data.
Footprint of the individual LAS tiles with respective LAS tile names attached	Polygon (.shp) and/or Google Earth KML polygons	Lidar datasets of one flight campaign can consist of thousands of individual image tiles. A footprint that can be overlaid with other GIS data helps to locate individual tiles much quicker.
Ground truth comparison report	Provide number of ground reference points and RSME (individual and overall)	This provides a good overview of horizontal and vertical accuracy of the data and potential local variance.

Table 3: Standard products that can complement the raw lidar point cloud data.

Defining area of interest

In order to avoid important areas being left out, such as pieces of the estuarine functional zone or coastal infrastructure, it is highly recommended that the client provides the data provider with detailed information on the exact area that should be covered. This information should be provided in form of either an Esri shapefile or a Google Earth KML polygon. It is advisable for the data provider to review and confirm the area of interest to be flown. If necessary, several iterations might be required.

It needs to be noted that the service providers usually collect more data than just inside the polygon to confirm that the polygon is fully covered and for the best logistic fit for flying the area. These "extra" data are paid for in the collection but are generally not delivered. It costs extra to process this data, but in a data starved environment such as South Africa it might be worth negotiating with the data provider and have this data processed and delivered as well. It might be sufficient to have these areas classified with only automated algorithms to reduce the amount of additional costs related to manual classification.

Removal of noise

Some of the existing coastal lidar datasets contain a lot of noise, either in form of exaggerated peak or trough values on the terrestrial side, or through the presence of data points on wavy ocean surfaces. Noisy peak or trough values should be removed by the data provider.

Data points on the wavy ocean surface should also be removed by the data provider – if they are not the explicit target of the data acquisition, as they do not provide any valuable information for coastal management but rather unnecessarily increase the size of the point cloud and subsequent derivatives (such as contours).

Safety standards

Besides the technical specification standards, safety standards relating to the actual acquisition have to be considered as well. These standards are governed by aviation administrations and laser safety which the service providers handle. These standards entail weather conditions under which it is safe to fly and other such aspects. It is important for the client to understand these constraints and make allowance in their project timeframes to accommodate possible delays in acquisition.

Another safety risk for consideration in the coastal and wetland context are circumstances which might impede proper ground referencing, thus resulting in lower absolute accuracies. Examples are physical inaccessibility due to remote and steep terrain such as cliffs or secluded beaches, in some cases the occurrence of hippos and crocodiles, or property access restrictions such as in mining areas. Many risks and disappointments resulting from those issues can be avoided if proper pre-scoping with the client takes place.

Dynamic standards and developments

All recommendations above are based on existing expertise and experience with coastal lidar. We are aware that these recommendations might change and expand with growing experience and field of application and the continuous progress of lidar technologies. It is envisaged to review them from time-to-time, should need arise.

	Suggested field name	Explanation
1	Title	Full title for citation
2	Alternate Title	Dataset short title
3	Publication/reference date	Publication date
4	Topic category, keywords	Broad discipline, e.g. lidar, digital elevation model, topography and so on.
5	Abstract, description	Provide an abstract for the data, analogous to a publication abstract.
6	Data quality	E.g. draft; final
7	Status	One of: Not for release, In prep, Submitted to SAEON, Published, or any notes indicating the status of the accession.
8	Publisher	E.g. DEA, if CSIR produced data for a DEA project.
9	Geographic identifier/ location/name	Geographical location, e.g. South Africa, East London
10	West bounding coordinate	Decimal degrees, western hemisphere locations are negative
11	South bounding coordinate	Decimal degrees, southern hemisphere locations are negative
12	East bounding coordinate	Decimal degrees, western hemisphere locations are negative
13	North bounding coordinate	Decimal degrees, southern hemisphere locations are negative
14	Spatial reference system and projection	Name of reference system, e.g: Projection: Geographic Spheroid: Clarke1880 or WGS84 Datum: Hartebeesthoek94
15	Begin sampling date	Sampling or acquisition data (for remote sensing data) in YYYY-MM-DD format
16	End sampling date	In YYYY-MM-DD format
17	Subject	Subject keywords, e.g. lidar, digital surface model, topographic contours, etc.
18	Language	Language(s) used within the dataset, e.g. English, Afrikaans, etc.
19	Scale/resolution	Factor which provides a general understanding of the density of spatial data in the dataset or point density per m^2.
20	Data/spatial representation type, format	E.g. point cloud, polygon, polyline, raster (depending on post-processing level)
21	Resource maintenance	Dataset last updated in YYYY-MM-DD format
22	Maintenance and update frequency	Information regarding how often the data is intended to be updated, e.g. five-year cycles, after storm events; alternatively an indication if the data acquisition was a once-off exercise.
23	Provenance/lineage statement	Information about the events or source data used in constructing the data specified by the scope or lack of knowledge about lineage; Where does the data set come from? References to data already assimilated on which this is based.
24	Distribution format(s)	E.g. LAS/*XYZ* point cloud, polygon, polyline, raster (depending on post-processing level)
25	License and user rights	Any descriptive constraints, e.g. data not for distribution, only printed version etc. We recommend you choose from Creative Commons licenses. For details: https://en.wikipedia.org/wiki/Creative_Commons_license. Include link to text of chosen license.
26	Resource online URL	Link to the data
27	Author/creator responsible party/ individual	Data provider/vendor who collected, cleaned up, processed the data? • Last-name, First-name MI • Last-name, First-name MI
28	Author institution	E.g. CSIR
29	Author contact details	E.g. email address, phone, postal address
30	Contributor to the data creation	Other parties involved in data generation
31	Contributor institution	Other parties involved in data generation
32	Contributor role	E.g. ground validation, data cleaning
33	Metadata file identifier	Unique identifier/file name for this metadata file, e.g. Landcover.meta
34	Metadata type	Name of the metadata standard (including profile name) used, e.g. SANS 1878
35	Metadata standard version	Version (profile) of the metadata standard used. Identification of the version of the metadata standard used to document the data set, e.g. FGDC-STD-001-1998. Name of the metadata standard (including profile name) used, e.g. SANS 1878.
36	Metadata language	Language used for documenting, e.g. English Version (profile) of the metadata standard used
37	Metadata character set	E.g. US-Ascii, UTF-8
38	Metadata date stamp	Date that the metadata were created or last updated in YYYY-MM-DD format
39	Metadata author	Name of the author and organisation
40	Contact person	Details of the point of contact/person/organisation/unit responsible for metadata
41	Organisation	Details of the point of contact responsible for metadata
42	Postal address	Details of the point of contact responsible for metadata
43	Physical address	Details of the point of contact responsible for metadata
44	Phone	Details of the point of contact responsible for metadata
45	Fax	Details of the point of contact responsible for metadata
46	E-mail	Details of the point of contact responsible for metadata

Table 4: Suggested minimum information to be included in the metadata for lidar data for a national inventory. Highlighted fields should be mandatory.

Suggested metadata fields

During stocktaking, the technical specifications or exact acquisition dates and times (which might be relevant for tidal considerations) were not available for some existing datasets, because it was either not provided by the data providers or were considered irrelevant by the original clients. While nobody likes to put metadata together, information on the origin of the data and the details of its acquisition and pre-processing are important for data sharing and sophisticated lidar applications. Further, lidar data are increasingly being recognised as geospatial datasets, and will therefore need to conform to the metadata standards as defined by the Spatial Data Infrastructure Act No. 54 of 2003. The ISO standard 19115 was adopted by the South African Bureau of Standards (SABS) as the official metadata standard for South Africa (SANS 1878).

Based on the SANS and ISO standards, Table 4 lists and explains metadata fields that should accompany raw lidar data as well as the value-added products derived therefrom. The metadata should preferably be compiled in XML format, as it is both machine- and human readable. However, as this format might appear quite cryptic to most users, we suggest that the metadata fields should be completed in any other tabular format (e.g. Microsoft Word or Excel), but adhering to the standard field names in order to allow for later conversion into XML, if necessary. Please note that this metadata format was not strictly developed for lidar point clouds and it is expected that some fields are not applicable for lidar and need to be left blank, while other additional fields relevant to lidar might need to be added.

Coastal lidar custodian or repository

Currently, the existing lidar datasets in South Africa are hosted and made available (if applicable) by the individual institutions and entities who commissioned the acquisition of the respective data. As mentioned above, this practice bears the risk of double acquisition for overlapping areas, inaccessibility of datasets and various data and data quality standards.

As the acquisition for lidar data is still very expensive in South Africa, the Community of Practice members indicated an interest in the creation of a central South African coastal lidar data repository, as exists for other types of geospatial information in other countries and in South Africa.

> Vague terms of reference can waste money and cause frustration

Here, a national custodian has been defined as a central agency where different types of geospatial data from different sources are curated and distributed. With regards to the role of such an agency for lidar data, it should be the curation and provision of the original point clouds as well as the basic value-added standard products. Further, it has to be discussed if the generation of non-standard products, such as vegetation structure for example, should reside with the custodian or if this shall remain within the responsibility of the original data provider, owner, or data requestor. Given the unique nature of the lidar point cloud, there needs to be discussion regarding the feasibility of embedding the function of the lidar custodianship in an existing institution or if it would be more appropriate to create a new agency for this purpose. However, the expected massive amount of data will require a host that is a prepared and an experienced data custodian, as well as decent metadata provision on the data provider side.

Identifying, creating or assigning national lidar custodianship still needs further discussion. A quicker and less costly option might be the creation of a national lidar data inventory, similar to the US Interagency Elevation Inventory where footprints and metadata of existing lidar data are held. This inventory lets people know where data exist and how to access it. This solution also enables data owners to indicate existence of data which might be access-restricted. Embedding a coastal lidar inventory into existing coastal-related geospatial structures, such as OCIMS (either as a layer in the Coastal Viewer (http://mapservice.environment.gov.za/Coastal%20Viewer/) or in the Data section of OCIMS (www.ocims.gov.za/dataset) can be achieved relatively easily. A pre-condition for creating a meaningful inventory would be, the provision of meaningful metadata.

Either solution – the custodian/repository or the inventory of footprints and metadata – would potentially make new acquisitions more cost-effective by targeting data gaps better. A central repository/inventory also supports the creation of strategic funding consortia for future acquisitions, similar to the multispectral SPOT mosaics for South African government entities. Most of the present coastal lidar data owners were comfortable making their data – or at least their metadata and footprints – available for this purpose. It is anticipated that discussions around this topic will be the focus of upcoming Community of Practice meetings in the near future.

Looking to the future

This guideline document was produced based on the authors' and the community's existing knowledge and experience. Acknowledging that there might be activities overlooked in the above, the authors encourage the coastal user community to contact them for sharing further experiences and recommendations to help improve this first version of coastal lidar guidelines.

Acknowledgements

This article would not have been possible without the input from and discussions with all the people contributing to the lidar workshops, including national and international participants. Further, a lot of learning came from hands-on processing of large amounts of data for the DEA OCIMS project.

Notes

Note 1: For explanation of Sigma refer to Appendix 3 of: http://sageo.org.za/wp-content/uploads/2014/08/Report_Coastal-LiDAR-WS_25-26March_Stellenbosch-final-draft_29Aug2014.pdf

Note 2: T Hengl, (2006). Finding the right pixel size. Computers & Geosciences 32, 1283 – 1298.

Contact Melanie Lück-Vogel, CSIR, mluckvogel@csir.co.za