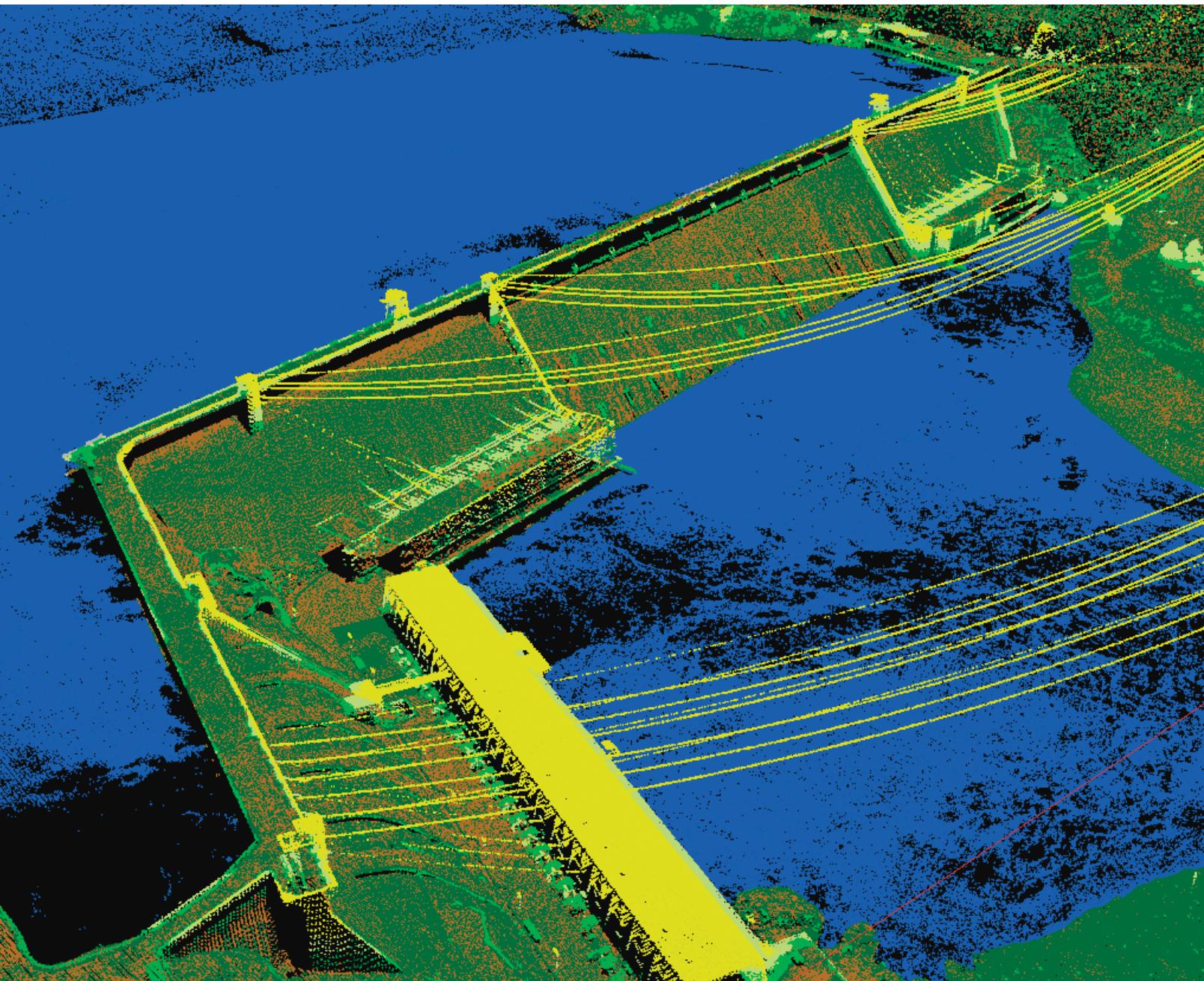


# 5 Ways to Use Lidar More Efficiently

By Clayton Crawford, Esri Lidar and 3D Analysis Product Lead, and Raghav Vermula, Esri Lidar and 3D Analysis Product Engineer

▼ Point cloud of the Grand Coulee Dam symbolized by LAS class code. Source lidar made available by the National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center.



The use of lidar has exploded in recent years, and for good reason. The technology can produce higher-quality results than traditional photogrammetric techniques for lower cost. This is accomplished, to a large degree, by the automated collection of measurements that are sampled very densely. The progress in laser scanning hardware has been astounding. A major side effect of switching to lidar-based technology has been the challenges associated with increasing data volume and the expansion of software processing capabilities needed. Fortunately, ArcGIS offers many tools for managing lidar point clouds and deriving useful products from them to aid scientific research and decision making. Here are five tips that enable ArcGIS to take best advantage of this data.

## 1 Have Lidar in LAS Format

Having lidar in LAS format may be obvious to the initiated but not to those new to using lidar data. LAS, short for LASer, is the industry standard format for lidar. The specification is maintained and published by the American Society of Photogrammetry and Remote Sensing (ASPRS). It was intended primarily for airborne applications but is also commonly used for terrestrial and mobile lidar. It's binary, efficient, widely supported, and the format ArcGIS works best with. See the Additional Resources section at the end of this article for information on currently supported versions. Note that ArcGIS works with LAS format lidar of all kinds: airborne, terrestrial, and mobile. The latter two are most useful when viewed in 3D, while airborne is useful in both 2D and 3D and can be processed with numerous surface analysis tools.

## 2 Make Sure the LAS Files Are "Baked" for Use in GIS

There are many flavors of LAS. Some are better than others for use in GIS. LAS was originally intended as an exchange format for laser hardware vendors. A lot goes on between initial data collection of a raw LAS file and its delivery to a client as a ready-to-use file. A few critical items in LAS processing are projection, tiling, and classification.

All the LAS files for a project should be placed into a projected coordinate system (PCS). The PCS should be the dominant one needed by most of the intended users of the data so on-the-fly projection is not required when it is used. On-the-fly projection is expensive in terms of performance and should be avoided. Note: It's not uncommon for LAS files to have been projected but to be missing the projection metadata that's supposed to be included in their header records. Files missing projection metadata are noncompliant with the specification and should be rejected or repaired. ArcGIS allows use of .prj files that can remedy this situation easily if going back to the data vendor is not an option.

Tiling should be performed on the LAS files. This avoids having relatively few swath-based files with overlapping extents that can be gigabytes in size. It's better to have many smaller files that don't overlap. Huge files are hard to manage, period. Smaller files are better. Also, LAS has no inherent spatial indexing, so retrieving points for subareas requires scanning the entire file to locate them. Scanning a 3 GB file for every spatial query is not workable. Files of 200 MB or less are more appropriate. (Note that spatial indexing support for LAS will be added to ArcGIS 10.2 through the addition of ancillary files. This will allow more efficient use of larger files and access to files on a network, though the non-GIS related practical constraints of huge files remain.)

Classified lidar is more useful. The majority of GIS applications related to lidar have at least some need for bare earth elevation models, which require properly classified data. Classification is nontrivial and usually performed by the data provider.

Some users believe that last returns (i.e., the last strike of a laser pulse) are sufficient to isolate the ground. This is incorrect. Last returns can occur on rooftops and in tree canopies. At a minimum, airborne lidar should be classified into ground versus nonground. Often, model key (thinned ground), water, noise, and overlap points are also categorized. There are other possible classes such as buildings and vegetation height. The greater the degree of classification (generally), the more useful the data. However, this can become prohibitively expensive because more classification means more processing and more human intervention. For a comprehensive list of guidelines, see the National Geospatial Program Lidar Base Specification 1.0, listed under the Additional Resources section.

## 3 Consider Your Options

ArcGIS provides several complimentary options for accessing lidar. There are three primary data access mechanisms: the LAS dataset, the terrain dataset, and the mosaic dataset. Knowing about these data types will let you determine which type to use.

The LAS dataset, introduced in ArcGIS 10.1, provides a simple way to access LAS files directly without importing or converting to some other format so you can start working with lidar data immediately. Using a simple toggle on a toolbar seamlessly switches between points and surfaces in both 2D and 3D viewing environments. Points can be symbolized using standard LAS attributes such as class code and return number. Points can be queried and used as a backdrop for measurements. Point class codes can be edited to fix misclassified points (which always manage to sneak through and get discovered when using the data). Surface analysis, with support for breakline constraints, and point metrics can be performed via geoprocessing tools.

The terrain dataset is a geodatabase-based solution for airborne lidar. Terrains can efficiently store and retrieve lidar surfaces from a database based on area of interest and level of detail queries. If only the lidar point geometry is needed—without the other attributes—bringing points into a terrain and shelving the LAS files can save a lot of storage space.



↑ A digital surface model (DSM) near Plant City, Florida, made from first return airborne lidar, which includes the building roofs and treetops. Hydro-flattening of the water features is accomplished through the addition of breaklines that are incorporated into the surface model.

Along with other GIS data layers, terrain datasets can be stored in a geodatabase and benefit from support for multiuser access and versioned editing. Because they are spatially indexed and pyramided into multiple levels of detail, they are also efficient and network-friendly.

The mosaic dataset is used to catalog, analyze, display, and serve massive image collections. In ArcGIS 10.1, it was enhanced to support LAS files, LAS datasets, and terrain datasets as imagery. The mosaic dataset performs on-demand rasterization, presents a map-like view of the lidar, and can be used as input to analytic functions as well as be the basis for sharing via elevation services. Essentially, the benefits mosaic datasets offer for imagery have been extended to include lidar.

## 4 Stage Data Appropriately

Lidar data is notoriously large. Careful planning is required to avoid bringing a network to its knees or making users wait too long for data to display. To determine the best overall approach, identify workflows by asking questions such as How big is the dataset? and Will the entire lidar collection be processed in order or will it be subject to ad hoc queries?

For example, look at a large statewide lidar program. Ultimately it may provide the public with ad hoc access to the data, but

initially, all holdings will go through a standard process of review, cleanup, and derivative creation. Consequently, it could make sense to house all data on a large central server and bring pieces of it (in ordered sequence) to a local machine for review and processing. Moderate-size solid-state drives are now affordable so the local machine, where many reads and writes will take place during processing, can work off a fast solid-state drive. Once the work is done, the processed data can be moved back to the server.

Data moves off and back onto the server once but allows local processing of the data, which is very fast. Depending on workflows, there are many options. The moral of this story is that with lidar, I/O tends to be very expensive, so minimize it to keep that cost down.

## 5 Pick the Right Points for the Job

The expression “lidar paints the surface with measurements” is another way of saying it is super dense. This density can be beneficial for capturing the detail of a rough or complex topography or creating a decent bare earth model for an area covered by forest. However, for open ground that’s gently sloped, the data is invariably oversampled.

Fortunately, point filtering can help. The filtering process includes just the points needed while excluding the others. The LAS specification has support for a point type called model key,

which is a subset of ground points. This thinned set will create a surface within a given vertical accuracy of the full resolution point set. Using just model key points to construct a ground surface may reduce the point count significantly. An 80 percent reduction rate is not uncommon. This benefit comes with just a small hit in vertical accuracy. Often, the accuracy is still sufficient for many engineering applications. The presence of these points requires the data to have been explicitly processed to flag or code them. Fortunately, it's common practice.

People often make the mistake of including all lidar return points when constructing a digital surface model (DSM). This kind of elevation model, which includes tree tops and building roofs, is also called a highest hit surface. Modern lidar is capable of processing multiple returns from individual laser pulses. In vegetation, returns greater than one represent either intercanopy points or ground beneath the vegetation. Including these points is unnecessary and

wasteful when making a DSM. Include them and the results will tend to look correct, but these unnecessary points can skew the results and will add to the cost of processing. All ArcGIS tools offer a way to filter on return. Use the first return, which will be the highest.

## Additional Resources

ArcGIS 10.1 supports versions 1.0–1.3 of LAS plus subversions of 1.4 that are 1.3 compliant. See the ASPRS LAS 1.4 Specification at [asprs.org](http://asprs.org). The US Geological Survey's National Geospatial Program Lidar Base Specification 1.0 (the official published version of the v13 draft) is available at [pubs.usgs.gov/tm/11b4/](http://pubs.usgs.gov/tm/11b4/). For specific information on using lidar with ArcGIS, see ArcGIS 10.1 Help under the LAS format topic and visit the ArcGIS 3D Resource Center at [resources.arcgis.com/en/communities/3d/](http://resources.arcgis.com/en/communities/3d/).

▼ A high-resolution DEM of Rock Island, located along the Georgia coast next to the Deboy sound, created in ArcGIS from airborne lidar and collected as part of the 2010 Coastal Georgia Elevation Project. Source lidar made available by the NOAA Coastal Services Center.

