Some of the challenges that are normally encountered in floodplain studies have been covered in the article “Challenges in Aerial LiDAR Processing in Flood Plain Studies” which has been published previously in LiDAR News Vol3, No 6. This article will expand on the challenges that were initially presented in the floodplain related LiDAR studies. My prior article can be characterized as “Single Data Collection—Incremental Data Processing”.

If you recall, most work relating to floodplain mapping includes high-accuracy digital elevation data that is required for hydraulic modeling; but, the floodplain area may comprise just a fraction of a total project’s full extent (administrative or watershed area). Normally, it is not realistic or advisable to acquire new digital topographic data of only the meandering floodplains, since it will be costly to develop and collect the necessary twisting flight lines while also meeting the control point requirements. Instead, it is typically more practical to obtain digital topographic data, or LiDAR, of the entire administrative or watershed area.

Then the engineer will utilize the most rigorous post-processing procedures specifically within the expected floodplain area for study.

Collecting data over the larger areas and rigorously processing the data within the selected areas is considered a valid approach; but, such an approach also poses several potential technical issues. For example, it could be likely that the flood plain area: 1) may actually increase at a later time through natural or man-made influences, 2) requires “incremental data processing” to develop a new surface for the
floodplain and 3) involves merging any initial and less rigorously processed data into the rigorous processing extent. When extending the rigorous extent, it is not permissible to alter any of the raw data in any way. Methods for aerial LiDAR data collection and processing workflows are generally well-established; however, there are several challenges faced in each LiDAR data processing activity.

For example, if we look at linear projects for infrastructure or corridors, such as processing an in situ pipeline alignment, we must take a somewhat different approach to project planning and processing than that associated with a floodplain. Case studies for pipelines present unique challenges, especially while using and integrating a newly acquired LiDAR dataset along the alignment. It is common that existing pipeline corridors will already have LiDAR data sets that have been collected and processed multiple times, in various non-contiguous collection projects, over the pipeline’s existence. The work may also be collected and processed by multiple vendors.

This paper addresses the issues of LiDAR preprocessing related challenges and the solutions to remedy the problems with temporal datasets and with potentially multiple specifications for collections over the same area.

**“Managing Multiple LiDAR Datasets”—Challenges in Pipeline Related Studies:**

LiDAR data is used for generating DEM and contours for engineering or other studies relating to a pipeline alignment. Consider that when a pipeline alignment changes, the corridor also undergoes alterations. As a result, this requires the acquisition of new LiDAR data, specifically information relating to noncontiguous areas where there is perhaps no existing LiDAR coverage. A substantial expense for small and noncontiguous acquisitions is associated with the initial mobilization cost of the aircraft to fly over the corridor site. Mobilization can be considered a fixed-cost for any project that then fractionally increases per square mile as the size of the project area decreases. For example, in a long pipeline corridor project that extends hundreds of linear miles, the cost of data collection for small noncontiguous areas will escalate drastically. To balance the cost and the data under these circumstances, the project owner may seek alternative sources for elevation information which may include the use of existing LiDAR data collected over a period of time by multiple vendors for various applications.

The “single data collection-incremental data processing” is primarily a processing related effort after the project’s data is collected using prescribed methods, whereas “managing multiple datasets” requires a data preparation effort where the LiDAR data was collected using various prescriptions (specifications) and from different sources. When managing multiple datasets it is imperative that appropriate processes are defined, prepared and understood prior to the actual merging of any data which will include its filtering and classification.

**Methodology**

Michael Baker Jr., Inc. (Baker) developed a comprehensive approach for meeting the requirements of our customer’s terrain development projects that involve managing datasets from various sources acquired over a period of time.

The process flow is comprehensive and starts with the review of the following parameters:

- Data-acquisition parameters like pulse density (calculated as pulses per unit area commonly known as pulses per square meter);
- Spatial accuracy of the collected LiDAR from different sources;
- Completeness of the data set (type of breaklines collected like hydro, transportation, etc. and 2D/3D breaklines);
- Spatial reference framework (datum, projections, units);
- Industry standard guidelines (see References) and specifications followed;
- Reports on systematic errors in the sensors used in the collection;
- Environmental conditions;
- Temporal elevation changes, natural or manmade, between the multiple data acquisition dates.

The initial inspection of the various reports and metadata associated with any data ensures the parameters used in data collection are consistent in all the datasets and they are suitable for the development of DEMs and contours needed for the project. This review is followed by a visual examination of the datasets to assess their overall quality and to evaluate visible differences in elevations in the areas where more than one dataset overlap. To make the process more systematic, each noncontiguous area is assigned a name in the geo-file/shapefile of the coverage area created for this purpose.

As shown in Figure 1, the existing LiDAR dataset is already processed and available. The newly collected LiDAR will account for changes in the pipeline...
route. The job is to seamlessly merge the newly collected LiDAR with the already processed data. The newly collected LiDAR is a noncontiguous area spread over the entire pipeline corridor.

The initial assessment and visual inspection of the input terrain datasets is an essential step in deciding whether any vertical shift should be applied to the new dataset so that it is merged with the new data without any anomalies or artifacts. Locations along the boundary containing flat and viewable surfaces are compared and the average elevation difference is calculated area-by-area and updated in the earlier created shape file. By comparing the ground surface created with the existing data to the surface produced with new data, a “best-fit” vertical offset is calculated. It is obvious that a single vertical shift will not adjust the entire dataset. It is necessary to determine more than one vertical shift that can be applied to a group of areas.

At the end of the process, vertical shift is applied to the different areas and shown, for example, in the following Table:

<table>
<thead>
<tr>
<th>Group</th>
<th>Area</th>
<th>Vertical Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Area 1, 2, 3, 4</td>
<td>0.3 ft US</td>
</tr>
<tr>
<td>Group 2</td>
<td>Area 5, 7, 9, 10</td>
<td>0.28 ft US</td>
</tr>
<tr>
<td>Group 3</td>
<td>Area 6, 8, 11</td>
<td>0.2 ft US</td>
</tr>
</tbody>
</table>

After the vertical shift has been applied, the next step is to ensure that the overlapping data is tied together sufficiently in the horizontal. In order to make any horizontal shift distinct features such as buildings are examined. If no significant offset is observed then no horizontal shift will be applied to the dataset.

After the data is shifted, the overlapping areas are inspected again for misalignment using the shape file as reference. In some cases stepping occurs between datasets. Stepping is an elevation seam that is visible at the transition line between data sources when the terrain is merged. Stepping occurs as a result of differences in vertical and horizontal accuracies, as well as resolution variances between two LiDAR datasets. To address stepping while maintaining the fidelity of the data to its original accuracy, we generate a temporary ground class from the new data and a DEM is created and analyzed along with DEM from the old data. The visual inspection of the DEM will help to determine a ridging along the boundary or other indicators of a vertical or horizontal misalignment between overlapping datasets.

If the end product from the LiDAR dataset is the creation of Triangulated Irregular Network (TINs), contours with use of breaklines, then this is reviewed from the multiple datasets. Breaklines are supplemental 3D linear features used to define edges and control smoothness in surface models. It is likely that all the breaklines are not created in same manner. Some of the breaklines may have been created in 2D using LiDAR intensity images or some of them may have been created using LiADRGrammetry in 3D. Some breaklines will cover only select hydro features like lakes and ponds while others may have hydro and transportation breaklines. A visual inspection and potential cleanup of breaklines may be necessary to bring them to a common level. A TIN is then produced from LiDAR multipoint data and the adjusted breaklines are incorporated.

The process developed by Baker will enable creation of a single merged dataset from multiple LiDAR datasets. The process also takes care of other issues like artifacts, dataavoids, and elevation seam leading.

**Conclusion**

Merging different datasets in pipeline corridor studies are unique. A need for a reliable process is required to create seamless data for the project coverage from existing and newly acquired LiDAR data while maintaining the vertical accuracy requirements. Baker has developed a methodology for producing merged terrain datasets that focuses on maintaining terrain fidelity and elevation accuracy inherited from the input datasets. The developed process encapsulates added-value by providing the ability to continually mine the disparate LiDAR datasets to provide continuous elevation data coverage for future modification to pipeline corridor studies. The ability to develop quality LiDAR datasets from different LiDAR datasets provides an opportunity to leverage existing valuable products in a cost-effective and timely manner.

"Defining appropriate processes is key towards successfully managing temporal LiDAR datasets."

**References**


Dr. Srinivasan “Srini” Dharmapuri has over 27 years of extensive, wide-ranging experience within the Geospatial industry; most notably with LiDAR, Photogrammetry, and GIS. He has worked in both the private and public sectors, as well as internationally.