



HYPACK
a xylem brand

Sounding Better!

Updates to the DATA SPLITTER-JOINER

By Connor MacDonald

In 2020, several updates have been made to the DATA SPLITTER-JOINER.

FASTER FILE LOADING

The performance of the method used to load file data into the program has been increased by over 300%. In the 2019 version of the program, a 356 MB file would load in 27 seconds on my machine. In the latest update, the same file now loads in 8 seconds on my machine.

FASTER DRAW TIMES IN THE DISPLAY WINDOW

Previously, the display would run quite poorly because of multiple factors. Files with large amounts of position data, multiple background files that need to be drawn, and creating a large number of lines were all contributing factors to the slow performance. Zooming in with a mouse scroll, panning the screen, and drawing lines were not smooth operations and could take a long time. These actions are significantly smoother and faster in the latest update, though the same factors still affect performance.

SPLIT BY FEATURE UPGRADES

The final updates are both related to the Split By feature that was added in 2018. The Split By feature automatically creates the lines to split the loaded files instead of requiring manual line creation. Previously, there were two methods—Split By Time and Split By File Size—that allowed the program to split files into chunks based on one of those parameters.

In the latest update, there is a new method called Split By Depth Gaps.

SPLIT BY DEPTH GAPS

This method is a bit different than the others. The other methods are to help split datasets into manageable, organized chunks. This method is more of a filter; it cuts out data without any depth, side scan, or LiDAR data.

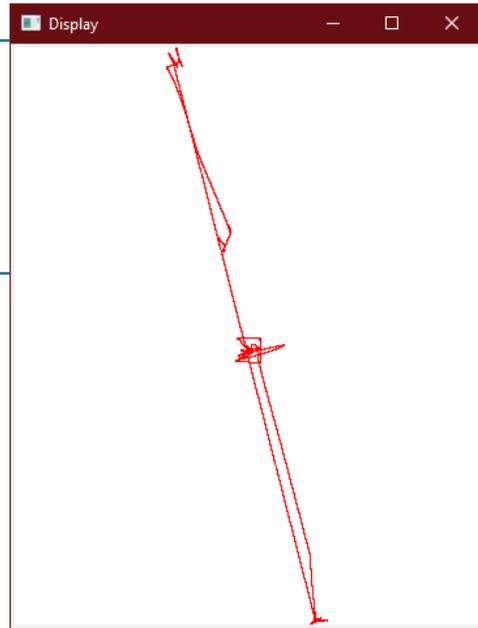
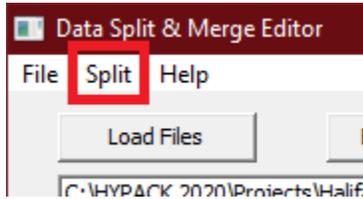
For example, I have a large file loaded into the program with multiple days of position and depth data (Figure 1). There are large portions of the file that do not contain depth data, and are unnecessary to keep for the sake of processing the data in SBMAX64 (the 64-bit SINGLE BEAM EDITOR).

Figure 1 Full Data Set—Large Segments without Depth Data

To get remove the data without depth information:

1. Click **[Split]** in the main window of the program (Figure 2). This opens a new window containing the Split By parameters **feature**.

Figure 2



2. Open the **Depth Gaps** tab to tell the program you want to remove gaps in depth data (Figure 3).

Figure 3: Defining the Parameters for a Depth Gaps Split

3. Define the split parameters and click **[Apply]**.

I am going to use the default time of 60 minutes for the maximum gap size in Minutes, because I know there are multiple hours of position data without any depths in this file.

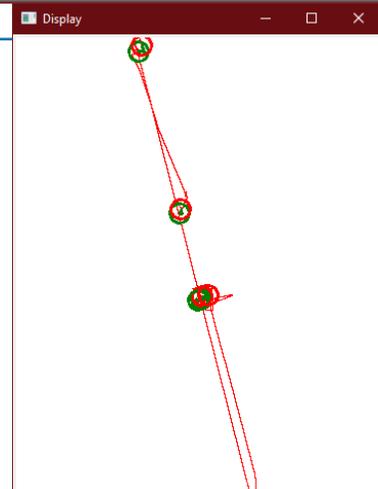
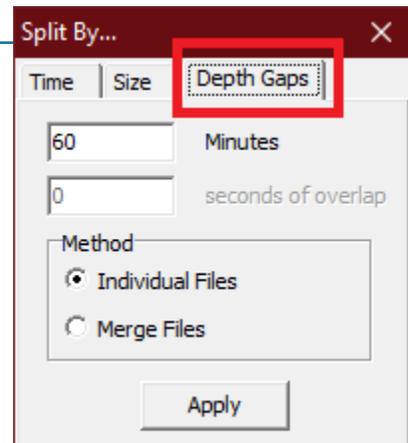
Once I click [Apply], I see the sections of data with no gaps in depth over 60 minutes have been selected (Figure 4).

Figure 4: Result Preview

Here, the Split By Depth Gaps has found 5 small sections containing depth information in this nearly half a gigabyte of data. That is a serious amount of useless data in regards to processing with SBMAX64. The total file size of the split files was brought down just over 68% in this case.

SPLIT BY DEPTH GAPS AND OVERLAPPING DATA

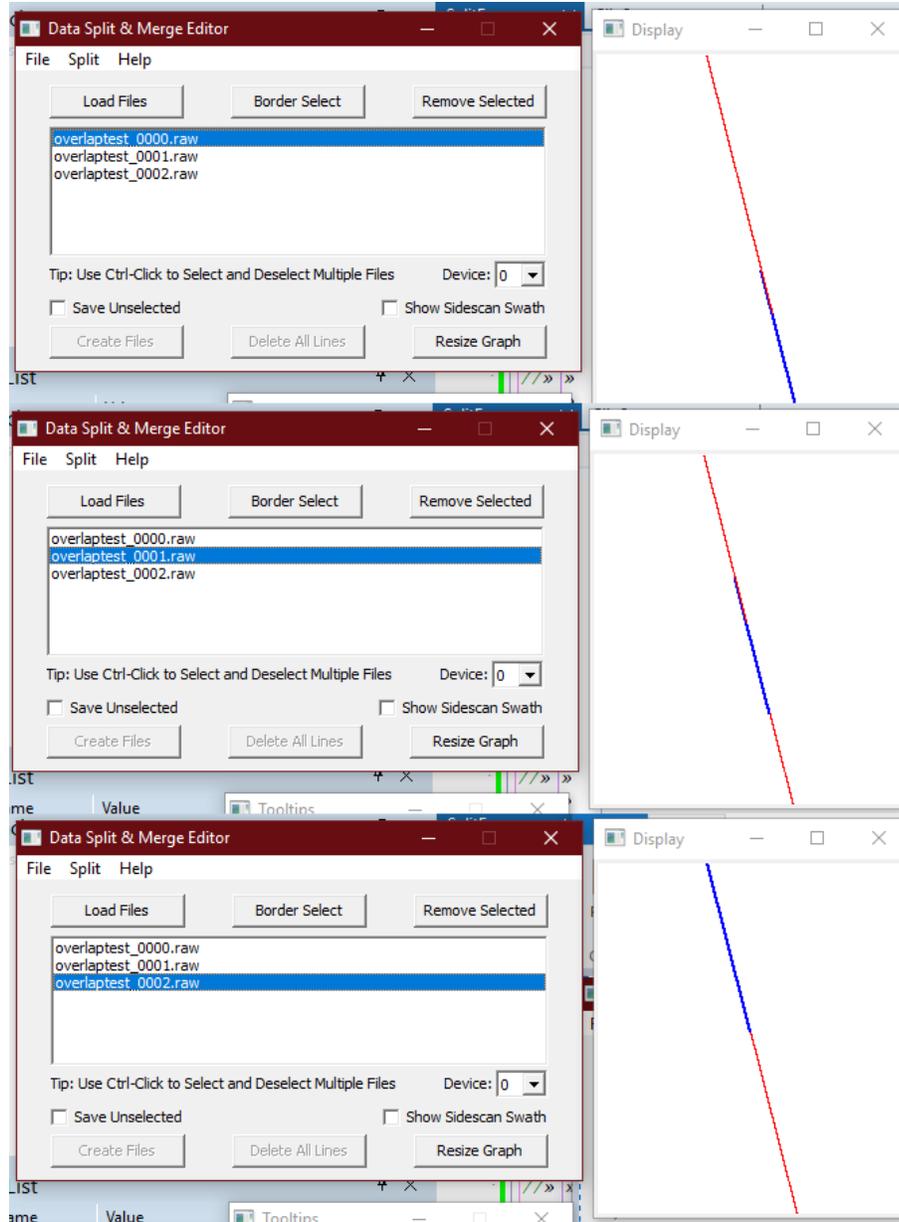
You may have noticed that, in Figure 3, there is a new box with radio buttons titled Method. This brings us to the latest update to the DATA SPLITTER-JOINER.



Previously, when splitting files with the Split By feature, the program would perform the splitting function to each file individually. This led to a messy result when using overlapping data. Let me explain.

In the following example, I have loaded three files with overlapping data into the program (Figure 5). You can see each one individually highlighted to show where the overlaps occur.

Figure 5: Data Files Overlapping End-to-End

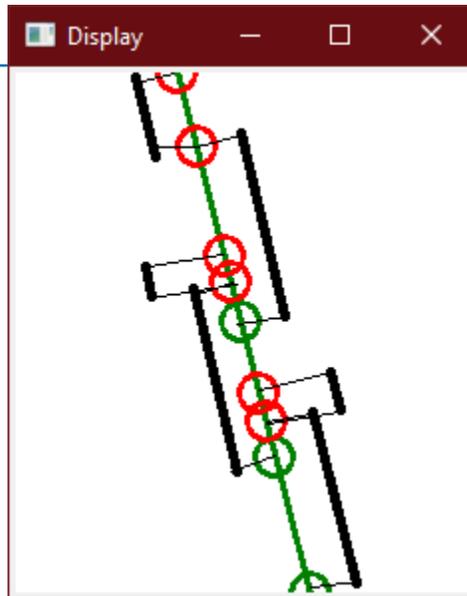


Individual Method with Overlapping Data

Let's try doing a Split By Time with 60 minutes using the original Individual Files method.

Figure 6 Individual Split Method from Three Line Files that Overlap End-to-End

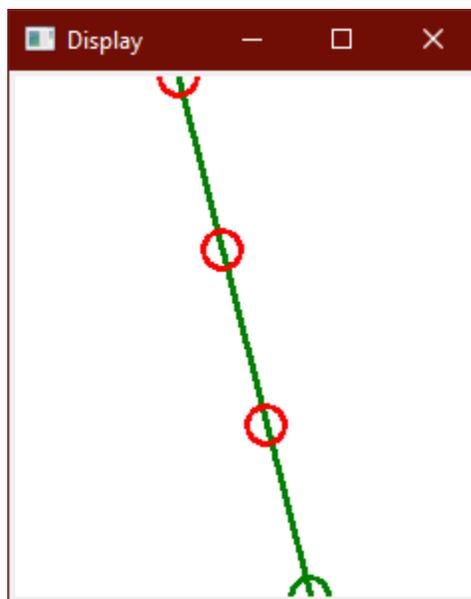
I have drawn black lines to show in more detail how the lines were created by the Split By feature (Figure 6). It is quite a mess. At the beginning of each file, a line has been created. The rest of the file that didn't fit into the 60 minute split has been turned into a short line that ends at the end of the file. This created 6 lines. One for the first 60 minutes of each of the 3 files loaded into the program, and one for the data past the first 60 minutes until the end of each of the 3 files.



Merge Files with Overlapping Data

Now let's try again with the Merge Files method instead (Figure 7). No black lines necessary here. The result is exactly what we would want to see when the files contain overlapping data.

Figure 7 Merged Files Split Method from Three Line Files that Overlap End-to-End



Choosing a Method

If you are unsure which method to use, it is as easy as comparing the results in the display window as you switch between the two methods and click [Apply]. If you need to split files individually, use the Individual Files method. If you need to split over merged files as a whole, use the Merge Files method. If you don't need to split files in a particular way, I would recommend trying both and choosing the one that is more visually appealing. The method choice simply limits the scope of the splitting method to individual files or the entire dataset that is loaded in the program. And after clicking [Apply] in the Split By window, you can continue to manually move the lines in the display window as you please.

SPLIT BY DEPTH GAPS AND NON-OVERLAPPING DATA

Here's a quick example where you might want to choose the Individual Files method instead:

Figure 8: Merged Files Method and Non-overlapping Data

I have two files loaded without any overlap. There is a gap in the data between them, but they are close enough that when you use the Merge Files method, it jumps into the next file instead of ending at the end of the first file (Figure 8). I've included the black lines again to show you what the program is doing.

Let's try again with the Individual Files method (Figure 9).

Figure 9: Individual Files Method and Non-overlapping Data

As you can see, in this case the Individual Files method makes more sense, and appears more visually appealing.

So, as I said, if you aren't sure which method to choose, just use trial and error with your datasets and choose what looks better. It is as quick as choosing the other method and clicking [Apply].

