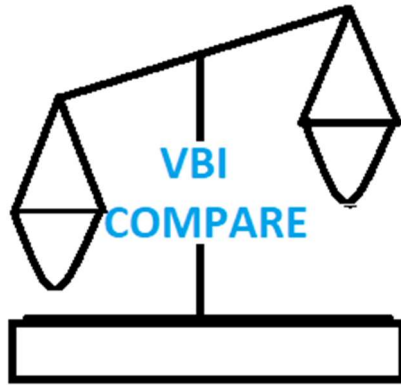


# VBI Compare

V1.0



CCOM/JHC, UNH & NOAA

**User's Manual**

## Contents

In Brief.....	3
How to Use.....	4
User Options .....	5
User Data Inputs .....	6
Operation .....	10
Reputation Calculation.....	12
Early Termination.....	13
How Does It Work? .....	13
NBS Data .....	13
VBI Data .....	14
Secondary Data .....	14
Data Outputs.....	15
Credits .....	15

## In Brief

VBI Compare was created by the Center for Coastal and Ocean Mapping (CCOM) to compare Volunteered Bathymetric Information (VBI) to authoritative data for reputation calculations. Currently, it is functional for US waters only using the National Bathymetric Source (NBS) as its authoritative data source. The goal of this program is to quickly collect and compare VBI data to NBS data to determine the quality of the VBI source. Based on user inputs, the program collects the desired data for comparison from the Amazon Web Service S3 bucket holdings of the NBS and the Crowd Sourced Bathymetry holdings of the Data Centre for Digital Bathymetry (DCDB). Once the data is collected, a batch file is created to initiate the reputation calculation. The batch file can be run immediately or stored for later and processed via a command prompt.

## How to Use

Upon starting VBI Compare, the user is presented with a graphic user interface (GUI) which allows them to provide the required inputs. The inputs selected will drive which files are obtained from the various AWS S3 holdings that VBI Compare scrapes. The possible selections and their meanings are described below. The user is required to select: one primary data source, the desired data search method, the need for secondary data, the output they need, and the output directory. Based on some of these required inputs, others become available to the user to further direct the program. These include the desired area for geographic searches and the specific charts or ships if the user is interested in specific chart or vessel data. Each of these options is further explained below.

The screenshot shows the VBI Compare v. 1.0 GUI. It features a 'User Inputs' section with several sub-sections. 'Select your primary data source' (A) has radio buttons for 'National Bathymetry Source' and 'DCDB Crowdsourced Bathymetry'. 'Data Search Method' (B) has radio buttons for 'Specific Vessels', 'Specific Charts', and 'Area Search'. 'Get secondary data based on primary?' (C) has radio buttons for 'Yes' and 'No'. 'What data collection method is preferred?' (D) has radio buttons for 'Compile S3 URLs' and 'Download Resources Locally'. Below these, there is a 'Select Area' button and a 'Or Type: NW Corner' field with coordinates (lat: 45.00, -65.00) and a 'SE Corner' field with coordinates (lat: 41.00, -67.00) (E). Below this, there are fields for 'List Desired Vessels' (F) with the value 'all: Copper Star, Tapestry' and 'List Desired Charts' with the value 'all: USCGA04C, USCGA05F'. Below these is an 'Output Directory' field (G) with a browse button. At the bottom, there is a 'Run Reputation Calculation?' section (H) with radio buttons for 'Yes' and 'No' (selected), and 'Clear Data' and 'Run!' buttons.

Figure 1: VBI Compare Main GUI

## User Options

- Select your primary data source. Figure 1 A  
National Bathymetric Source: Directs the program to first collect NBS data from the NBS AWS S3 holdings

Volunteered Bathymetric Information: Directs the program to first collect VBI data from the NCEI DCDB AWS S3 holdings.

This section allows the user to select the source they wish to base their search on. For example, a user may have a specific vessel or vessels in mind and wish to compare it to all charts that it crosses. Alternatively, a user may have a specific chart or charts of interest that they wish to compare to all vessels that cross it. In the case of an area search, selecting NBS results in all charts that intersect the search area. For VBI, an area search would result in all vessel track lines that intersect the search area. There are two primary source options:

Note: For an area search, it is recommended that the user select National Bathy Source as their primary. If VBI is selected, all vessel track lines that cross the area will be collected, some of which can continue for long distances beyond the area. This can cause the program to require the collection of a large number charts to cover the full extent of the track lines.

- Data Search Method. Figure 1 B  
Selections in this section allow the user to designate the search method the program should perform for the primary data source. The area search option is always available, but the specific vessels and specific charts options will only become available if VBI or NBS are selected in the primary data source section, respectively.

Specific Vessels: When VBI is the primary source, this selection instructs the program to search the DCDB CSB API (<https://gis.ngdc.noaa.gov/arcgis/rest/services/csb/MapServer>) for file names using the specific vessel names provided by the user. Selection of this option makes the “List Desired Vessels” input box available to the user to type in the name of their desired vessel(s).

Specific Charts: When NBS is the primary source, this selection instructs the program to search the NBS AWS S3 holdings for NBS tile files associated with the specific chart numbers provided by the user. Selection of this option makes the “List Desired Charts” input box available to the user to type in the number(s) of their desired charts

Area Search: This option is available regardless of the primary data source. Selecting it will make the Select Area button clickable to the user. This option allows the user to designate a specific geographic area of interest. Alternatively, the user may manually type in the north west and south east corners of their bounding box of interest. The program will then search for the primary source and return files that intersect that area. Typed coordinates must be in un-projected WGS 84 coordinates.

- Get secondary data based on primary? Figure 1 C

This section allows the user to tell the program whether they want the secondary data type collected based on the results of the primary data search. Secondary data is the other data type not selected as the primary source. For instance, if a user has a specific chart of interest and wants to find ship tracks that intersect it, they would select NBS as their primary source and then 'yes' to get secondary data. Alternatively, if a user only wants to collect the NBS data for that chart of interest, they would select 'no' for the secondary data.

Yes: Directs the program to collect secondary data source based on the primary. Selecting yes will also make running a reputation calculation available to the user if desired.

No: The program will only collect the primary data source. With 'no' selected, a reputation calculation cannot be performed. This option can be used simply as a download/ data search tool, or for future self- made batch files or other uses.

- What data collection method is preferred? Figure 1 D

This section allows the user to designate the type of data output the program generates. These options are based on the method by which the user would like to process the collected data.

Local Download: Selecting this method will cause the program to download the AWS files for all associated queries to a folder structure created in a designated local folder set by the user. The program will create subfolders in this location for data type, vessel name (DCDB data), and tiles (NBS data). Additionally, a text file will be created for each data type. These text files hold a list of file paths for each file to be used in a reputation calculation. This method is generally slower.

S3 URL Compilation: If this method is selected, the program will collect the AWS S3 URLs for the files desired by the query. A text file will be created with a list of these URLs to access the data directly in the cloud. Up to two text files will be created, one for each data type. This option allows for comparison processing to be done via direct communication with the AWS S3 buckets avoiding egress costs. This method should be faster but requires internet connectivity for the entire comparison process.

## User Data Inputs

### Geographic area Search

If the user elects to use an area search in the Data Search Method section, the Select Area button becomes available. Clicking the 'Select Area' button opens a pop-up window where the user can click and drag an area box over a map. The corner points of the selected box are collected by the program and the north west and south east corners are pulled and displayed in the main GUI. These coordinates are used as the search parameters for any requested queries.

The user can zoom in and out on the map using the wheel of their mouse, a two finger stretch on their mouse pad, or the zoom in and out buttons on the GUI with a left click. The user can also pan by clicking and holding the left mouse button or clicking down on the center wheel of their mouse. To draw a selection box, click the 'Draw Box' button and then left click and drag across the map. Clicking 'Apply!' sends the coordinates to the main GUI for use in the desired searches and closes the area box pop-up.

Alternatively, if the user knows the northwest and southeast corner of their desired box, they may type the un-projected WGS 84 coordinates of those points into the main GUI or the Select Area pop-up. In this instance, the select area pop-up tool will not be necessary, but may be used to confirm the box location. Coordinates must be in the form: Lat, Long with negative signs (-) denoting south latitude and west longitude. Example 43.00, -69.00

If the NW/ SE boxes are already populated in the User Inputs GUI, it will be shown in the area search pop-up in black once opened. This area will be overwritten if the user drags a new box, upon closing the area search GUI.

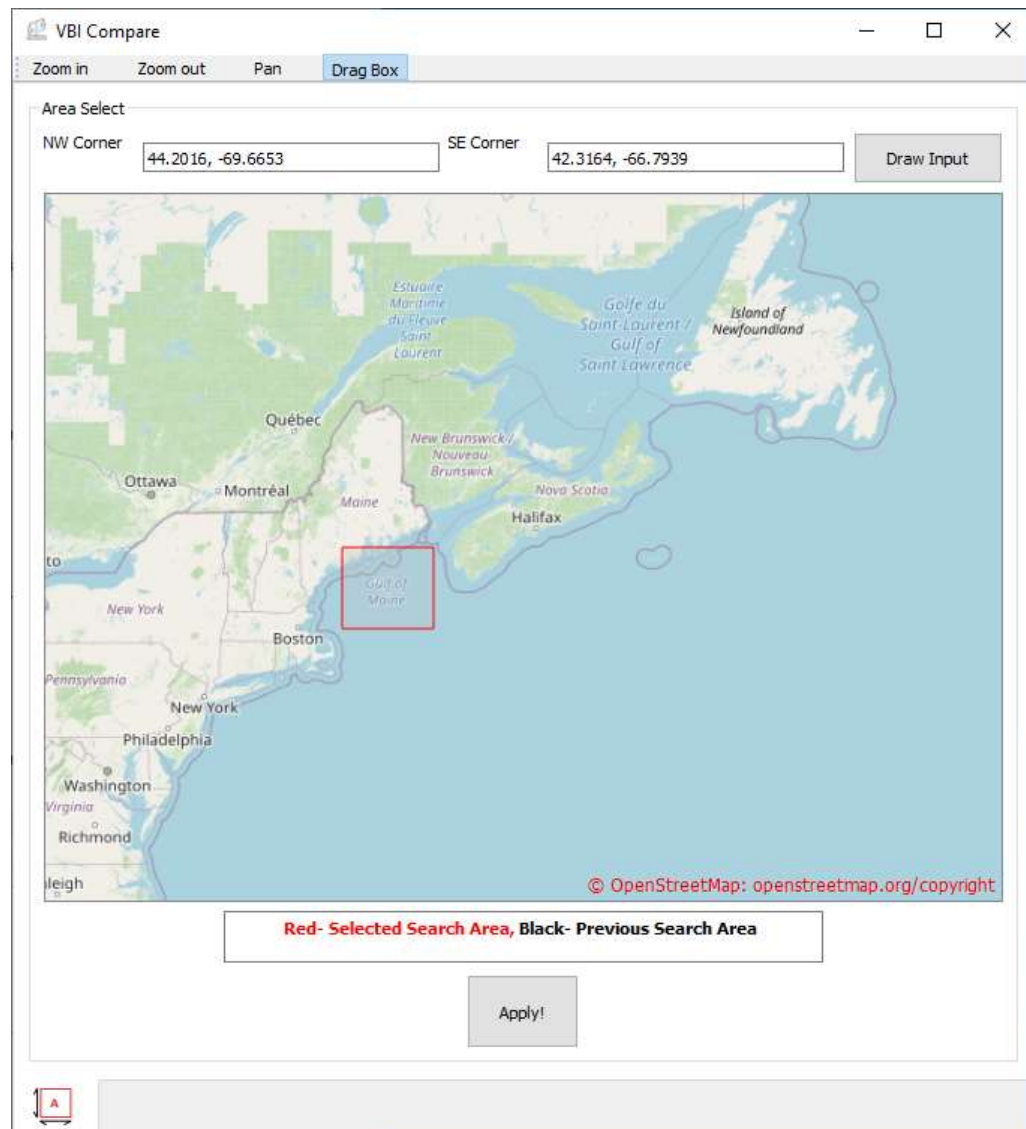


Figure 2: Area Search Pop-up

### Specific Vessels Box


This text box becomes active when the user selects the 'Specific Vessels' option in the Data Search Method section. This box is where the user types the names of the vessel, or vessels, they are interested in. For multiple vessels, the names of those vessels must be separated by a comma. For example: Copper Star, Tapestry



### Specific Charts Box

This text box becomes active when the user selects the 'Specific Charts' option in the Data Search Method section. This box is where the user types the number of the navigational chart, or charts, they are interested in. For multiple charts, the numbers of each chart must be separated by a comma. For example: US3MA1BF, US3MA1AC. VBI compare uses the Coast Survey ENC rescheme API ([https://gis.charttools.noaa.gov/arcgis/rest/services/MarineChart\\_Services/Status\\_New\\_NOAA\\_ENCs/MapServer](https://gis.charttools.noaa.gov/arcgis/rest/services/MarineChart_Services/Status_New_NOAA_ENCs/MapServer)), therefore the chart numbers must be those associated to rescheme tiles. The number of a desired chart can be found here: <https://distribution.charts.noaa.gov/ENC/rescheme/>

### Output Directory

This text box is required input. The user clicks on the  to open a pop-up window where the user selects the parent folder to which they would like their data to be downloaded to. For local downloads, this parent folder will be populated with subfolders as shown below. One branch of folders is made per data type (NBS/ VBI). If the user wants both primary and secondary data, then both branches will be created. **Note: Any folder names for the desired location must not have any spaces or hyphens in their name. An \_ may be used.**

For S3 URL compilations, the parent folder is populated with subfolders for each data type (NBS/ VBI).

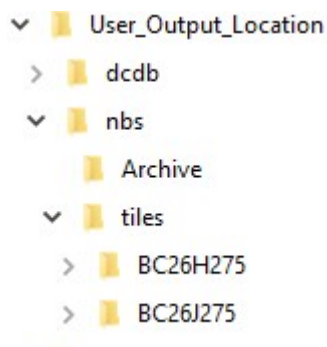


Figure 4: Example NBS Local Download Structure

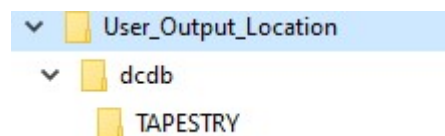


Figure 3: Example VBI Local File Structure

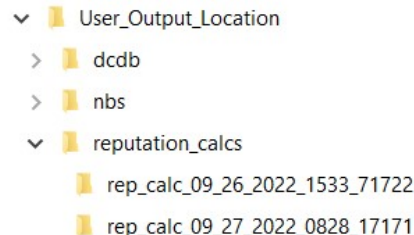


Figure 5: Example Reputation Calculation File Structure

No additional folders are created. The text files created by the program are placed in the appropriate folder based on the file type.

If reputation calculations are made by the program as described below, an additional folder structure for the batch files and associated calculation files will be made. A "reputation\_calcs" folder will be added to the user's selected output folder. Within it, a folder for each individual calculation run will be made. The

folder will house the batch file created by VBI Compare as well as any files created by the reputation program.

**NOTE: If the user must move the files and folders created by VBI compare, move the output folder with the substructure unchanged. If the substructure is changed, the ability of the program to find needed files will breakdown.**

### Run Reputation

This allows the user to decide if they want to run the reputation calculation immediately or not. If the user did not select to get secondary data, this option is automatically set to 'No.' If the user selected 'yes' to collecting secondary data, the reputation calculation can be completed. This process will build a folder structure within the user selected output folder. A folder named "reputation\_calcs" will be created which contains a sub-folder for each reputation calculation run. These folders are named "rep\_calc\_DATE(MM\_DD\_YY\_hhmm)\_RANDOMNUMBER." The date and time in the file name are the date and time the reputation calculation began. The random number is used to associate all of the files created during the calculation process. The batch file and all other files made during each run will contain this same random number.

Yes: Selecting Yes tells the program to build a batch file using the data collected based on all the previous inputs. The program then conducts the reputation calculation immediately.

No: Selecting No tells the program to build a batch file using the data collected based on all the previous inputs and saves the .bat for later use.

## Operation

Once the user has filled out the GUI with their desired information, clicking the "Run!" button will start the data gathering process. Alternatively, the user can click the Clear Data Button to remove their inputs to the data gathering sections of the GUI. Upon clicking "Run!", the program does an initial validation check. This ensures that the user has provided all of the necessary information given their inputs in the top four sections of the GUI. If any information is missing, a pop-up will appear that details what is missing.

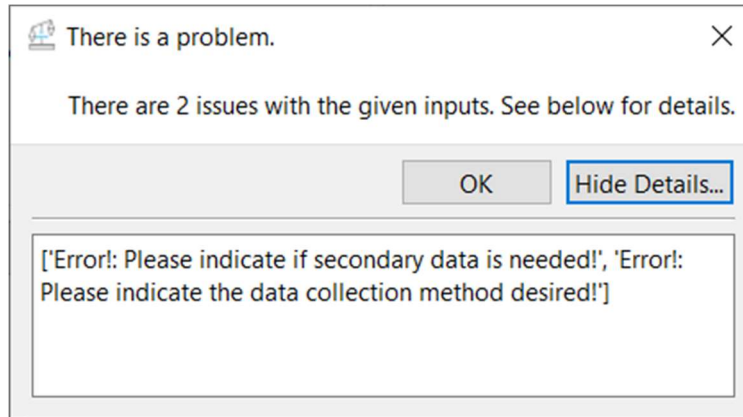


Figure 6: Example Warning Pop Up

If validation passes, then the program will begin scraping the various desired API and AWS S3 holdings. A monitor pop-up will also appear which will give the user information about the status of their request.

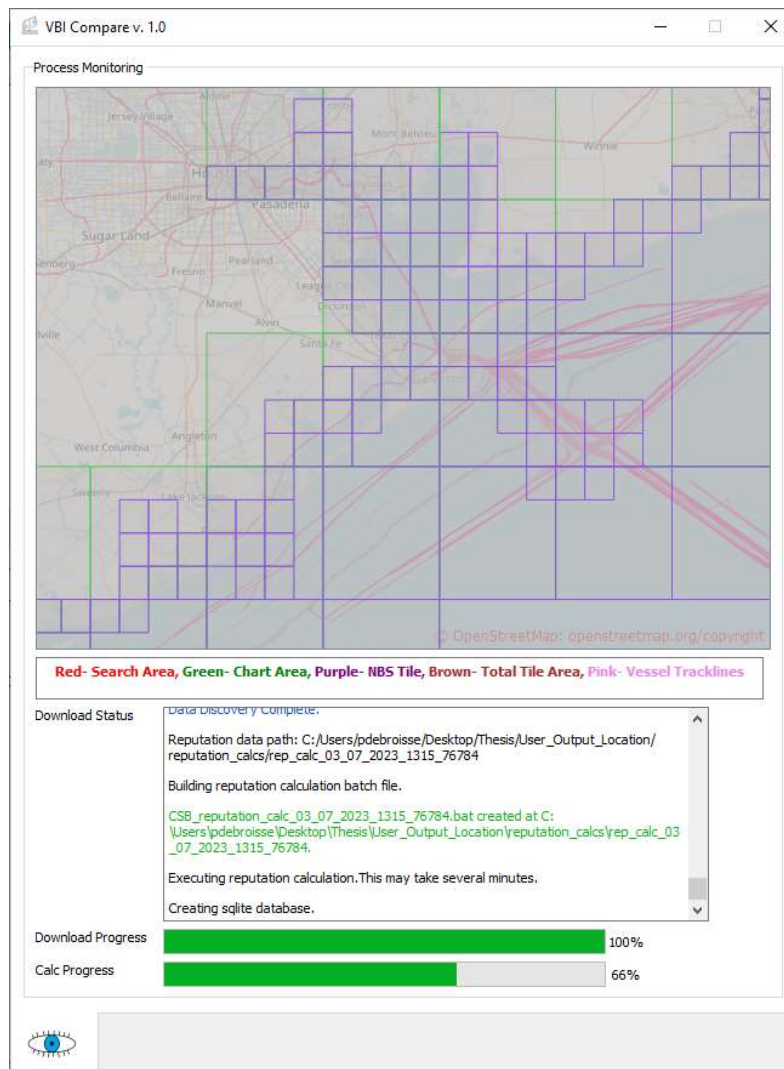


Figure 7: VBI Compare Monitoring Window

The map window will populate initially in one of two ways. For a specific chart or vessel search, the map will populate showing the whole earth. For an area search, the map will populate zoomed in on the requested area, bounded in red. As NBS tiles are downloaded, their boundaries will populate the map in purple. Only tiles that are available in the NBS AWS S3 holdings will populate. Any MCD charts that may be affected by these NBS tiles will be shown outlined in green. A list of potentially affected charts will also be displayed in the Download Status window. Additionally, as VBI track lines are collected, they are displayed on the map window in pink. Finally, once the process is complete, if NBS tiles were collected, the map will zoom out to show the total area covered by the collected tiles in brown, if that area is larger than the original search area. If the user completed a process with NBS as their primary data source and elected to collect the secondary VBI, this total area is the same as the bounding box that was given to the DCDB CSB API to search for VBI data.

The Download Status window shows the steps the program is performing textually. Messages in black indicate a normal status change. A message in yellow is a warning while red indicates a failure. Green messages show that data has been collected and is available either as a downloaded file or a text file containing S3 URLs. It will also show the location of those files. Blue messages indicate a process is complete. Common warnings include: a download is being skipped because the file already exists locally, and a file name was listed in an API but is not available in an AWS bucket. Common failures include: the inability to reach an AWS bucket or an API. This usually only occurs with NCEI, NBS, or MCD is performing maintenance on a server.

The download progress bar shows the percentage of completion of the data discovery process graphically.

The Calc Progress bar shows the percentage of completion of the reputation calculation, if performed.

Once a process is complete, the Download Status box will state ‘Requested Process Complete’ in blue. The user can then click the X in the upper right corner to close the monitor and either start a new process on the main GUI or close the program all together.

## Reputation Calculation

If the user has elected to collect primary and secondary data, a batch file is created for reputation calculation. The user can select whether or not to conduct that calculation immediately or at a later time.<sup>1</sup>

If the user decides to run the calculation later, the batch file is stored in the user output folder. To run that batch file, the user must use the command prompt. To do this, open command prompt from the start menu. Then copy and paste or type the location and file name of the desired batch file to be run. For example:

C:\Users\Desktop\Thesis\User\_Output\_Location\reputation\_calcs\rep\_calc\_09\_22\_2022\_1202\_99710\VBI\_reputation\_calc\_09\_22\_2022\_1202.bat. Then hit enter. The .bat will be run in the same manner as

---

<sup>1</sup> For information about how the reputation is calculated, see this article: <https://ihr.iho.int/articles/estimating-observer-and-data-reputation-in-mariner-volunteered-bathymetry/>.

if VBI Compare had run it and the associated outputs and files will be created in the same folder as the batch file.

## Early Termination

If the user notices an error in their request, they have two options: let the program finish or end the process early. Ending the process early can be achieved by clicking the X on the monitor window. If the process is incomplete, a pop- up will appear asking the user to confirm they want to terminate the process. Terminating the process results in an immediate stop. This can result in files being corrupt or incomplete if a local download was selected. It should be avoided if possible. In the case of a URL compilation, the result could be an incomplete list of URLs. If the program is re-run, newly downloaded data files are compared to previously downloaded files and will replace any corrupt files. Or, in the case of a URL compilation, a complete new .txt file is created.

## How Does It Work?

There are 16 potential setup cases for the GUI (Table 1).

Table 1: VBI Compare Use Cases

Primary Data	Collect Secondary Data?	Search Type	Output Method
NBS	NO	Specific Charts	Local Download
NBS	NO	Specific Charts	URL List
NBS	NO	Area	Local Download
NBS	NO	Area	URL List
NBS	YES	Specific Charts	Local Download
NBS	YES	Specific Charts	URL List
NBS	YES	Area	Local Download
NBS	YES	Area	URL List
VBI	NO	Specific Charts	Local Download
VBI	NO	Specific Charts	URL List
VBI	NO	Area	Local Download
VBI	NO	Area	URL List
VBI	YES	Specific Charts	Local Download
VBI	YES	Specific Charts	URL List
VBI	YES	Area	Local Download
VBI	YES	Area	URL List

## NBS Data

VBI Compare uses various NOAA data products to collect and furnish desired data to the user. For National Bathymetric Source authoritative data, the program utilizes the [NBS AWS S3 bucket](#). This bucket contains the current [Blue Topo](#) holdings of the NBS. In order to determine which files to collect, the program completes a search of the NBS tile GeoPackage using a geographic area. If NBS is the primary data source, this area is either the one provided by the user or is the bounding box of the MCD chart(s) named by the user. If VBI was the primary data set, the program will do an intersection search against the [NOAA Marine Chart Division's \(MCD\) Status of New NOAA ENC's Map Service](#) using the track

line geometry and then search the GeoPackage using the area of each chart found. Searching the GeoPackage results in the direct URL paths to the data within the bucket, if it exists. If data exists for that tile, it is collected and then output in whichever way the user has selected (local download or S3 URL compilation).

### VBI Data

For volunteered bathymetric data, the program searches the National Centers for Environmental Information's (NCEI) hosted [Data Centre for Digital Bathymetry AWS S3 bucket](#) for data based on the users input. For both geographic area and specific vessel searches, the program uses the [IHO Data Centre for Digital Bathymetry \(DCDB\)](#) Crowd Sourced Bathymetry (CSB) API to search for file names associated with search. Once a list of files is collected from the API, the program searches for those files in the DCDB AWS S3 Bucket. If a requested file exists, it collects it then outputs it in whichever way the user selected.

### Secondary Data

If both a primary and a secondary data type are to be collected, the program runs the normal process for the primary data source based on the user inputs, then uses data from that process to implement the data query for the secondary data source. For instance, if a user conducts an area search with NBS data as primary and VBI as secondary, the program will search the NBS GeoPackage for tile numbers and then collect the data for those tiles from the NBS AWS bucket. This is the normal NBS process. Once complete, the program then creates a bounding box containing all of the tiles that were collected. It then conducts an area search using this new bounding box for the VBI data collection. Table 2 below provides a basic explanation of each type of primary/ secondary collection.

*Table 2: Primary/ Secondary Decision Matrix*

Primary Data	Secondary Data	Search Type	Explanation
NBS	VBI	Specific Charts	NBS data for desired charts is collected, bounding box for all collected charts created, area search for VBI using bounding area.
NBS	VBI	Area	NBS data is collected for the area, bounding box for all collected charts created, area search for VBI using bounding area.
VBI	NBS	Specific Vessels	VBI data collected for specific vessels, track line waypoints are collected, intersection search for NBS using VBI track lines.

VBI	NBS	Area	VBI data is collected for the area, track line waypoints are collected, intersection search for NBS using VBI track lines.
-----	-----	------	--

## Data Outputs

Depending upon the user's needs, the program will provide the requested data in two ways: local download or S3 URL list in a text file. The local download allows the user to collect the data from the various AWS holdings for use locally. In addition to the locally downloaded files, a text file for each datatype is created containing a list of file locations. The URL list is provided for those who wish to work directly with the data in its AWS bucket. This method is much faster but requires internet connectivity to complete the entire reputation calculation. In the case of a local download, the user is provided with .tiff and .tiff.aux files for NBS data and .csv files for VBI data. In the case of S3 URL list, the program provides a .txt file, one for each data type, with URLs listed for each .tiff (NBS) and .csv (VBI) data file.

**NOTICE:** For local downloads, the program only collects files which are newer than those already existing in the user's output folder for NBS data. If old files exist in the user's output folder, those files are moved to an archive folder within that file's parent folder. Since VBI files are not updated in the DCDB, file age does not need to be compared. If a file exists in the output directory, the new download is skipped. Otherwise, files will be downloaded as requested. For S3 URL lists, a text file is created for each run with a time stamp in the file name.

Any files or data created during the reputation calculation process will be located in the same folder as the batch file created by VBI compare. Any output plots will be saved here as well.

## Credits

VBI Compare was developed at the Joint Hydrographic Center/ Center for Coastal and Ocean Mapping at the University of New Hampshire.

For bugs and/ or informational requests, contact:

- Patrick Debrosse ([patrick.j.debrosse@noaa.gov](mailto:patrick.j.debrosse@noaa.gov))
- Brian Miles([bmiles@ccom.unh.edu](mailto:bmiles@ccom.unh.edu))
- Brian Calder([brc@ccom.unh.edu](mailto:brc@ccom.unh.edu))

Maps shown in the GUI are curtesy of OpenStreetMap. Please go to [openstreetmap.org/copyright](https://openstreetmap.org/copyright) for more information.

