Crowdsourced Bathymetry Data

Data Processing and Quality Assessment

Anthony Klemm, NOAA
Coast Survey Development Lab

IHO CSB Workshop – Wellington, NZ – March 2025

How to get from this:

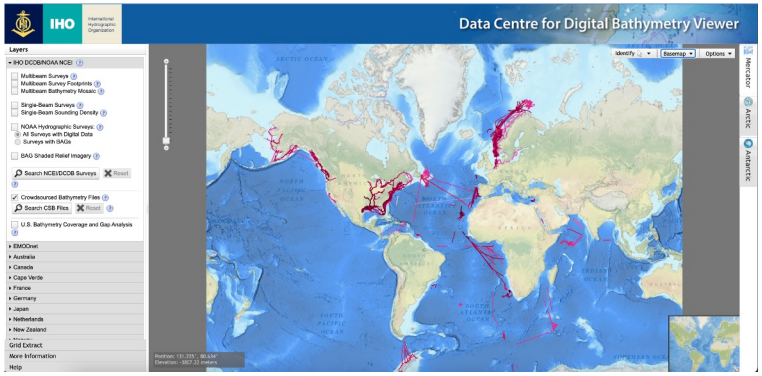

```
{
    "crs": {
        "horizontal": {
            "type": "EPSG",
            "value": 4326
        },
        "vertical": "Transducer"
    },
    "providerContactPoint": {
        "orgName": "Example Cruises Inc",
        "email": "support@example.com",
        "logger": "Rose Point ECS",
        "loggerVersion": "1.0"
    },
    "convention": "XYZ CSB 3.0",
    "dataLicense": "CC0 1.0",
    "platform": {
        "uniqueID": "EXAMPLE-f8c469f8-df38-11e5-b86d-9a79f06e9478",
        "correctors": {
            "positionReferencePoint": "GNSS"
        }
    }
}
```


CSB data log file (with JSON metadata string)


```
LON, LAT, DEPTH, TIME
68.499214, 15.832683, 59.3, 2020-02-25T01:08:06Z
68.499151, 15.832738, 59.3, 2020-02-25T01:08:07Z
68.498965, 15.832905, 61.3, 2020-02-25T01:08:11Z
68.498965, 15.832905, 61.3, 2020-02-25T01:08:11Z
68.498655, 15.833184, 61.3, 2020-02-25T01:08:15Z
68.498592, 15.833239, 61.3, 2020-02-25T01:08:16Z
68.498213, 15.833567, 55.3, 2020-02-25T01:08:23Z
68.49815, 15.833622, 55.3, 2020-02-25T01:08:24Z
68.49815, 15.833622, 55.3, 2020-02-25T01:08:24Z
68.497713, 15.83401, 54.3, 2020-02-25T01:08:30Z
68.497399, 15.834287, 53.3, 2020-02-25T01:08:35Z
68.497399, 15.834287, 53.3, 2020-02-25T01:08:36Z
68.497336, 15.834341, 53.3, 2020-02-25T01:08:36Z
68.497147, 15.834506, 59.3, 2020-02-25T01:08:39Z
68.497147, 15.834506, 59.3, 2020-02-25T01:08:40Z
68.497084, 15.83456, 59.3, 2020-02-25T01:08:40Z
68.496959, 15.83467, 59.3, 2020-02-25T01:08:43Z
68.496897, 15.834725, 59.3, 2020-02-25T01:08:44Z
68.496897, 15.834725, 59.3, 2020-02-25T01:08:44Z
68.496708, 15.83489, 54.3, 2020-02-25T01:08:47Z
68.496708, 15.83489, 54.3, 2020-02-25T01:08:47Z
68.496646, 15.834946, 54.3, 2020-02-25T01:08:48Z
68.496457, 15.835112, 49.3, 2020-02-25T01:08:50Z
68.496457, 15.835112, 49.3, 2020-02-25T01:08:51Z
68.496205, 15.835332, 53.3, 2020-02-25T01:08:55Z
68.496143, 15.835387, 53.3, 2020-02-25T01:08:55Z
```


Data and identifying token are submitted to DCDB via HTTPS post

Data discovery and access via map viewer.

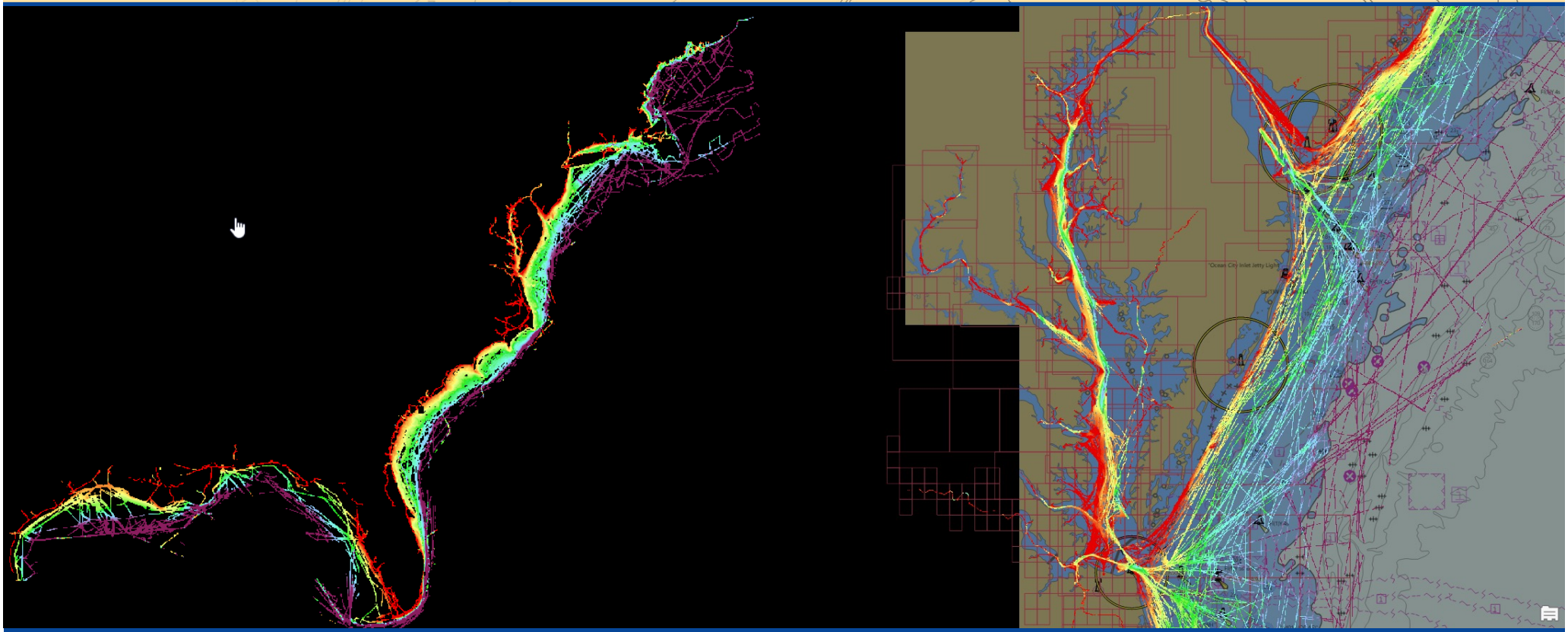


Frequent update of viewer

Funded by NOS OCS

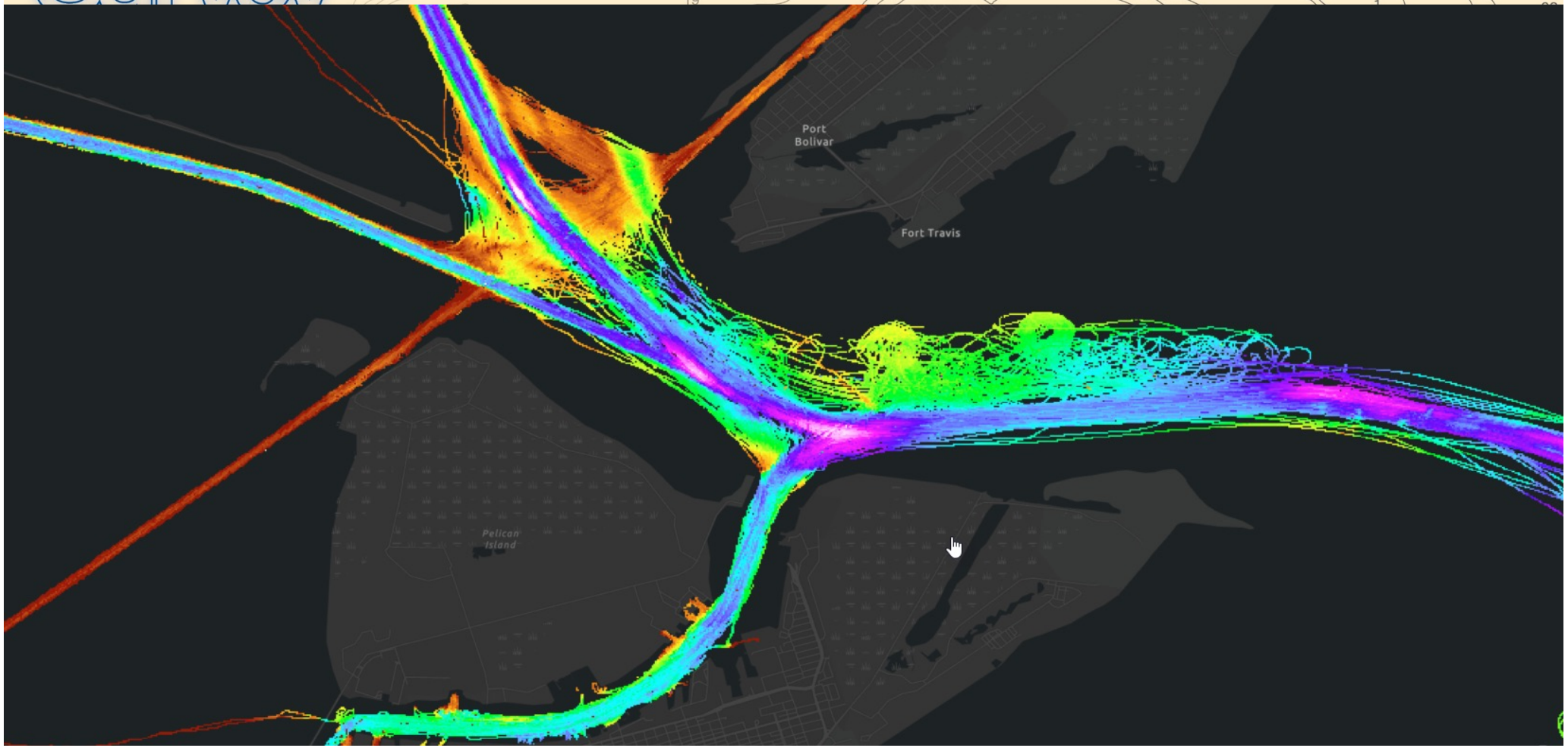
Coast Survey

To this:



Coast Survey

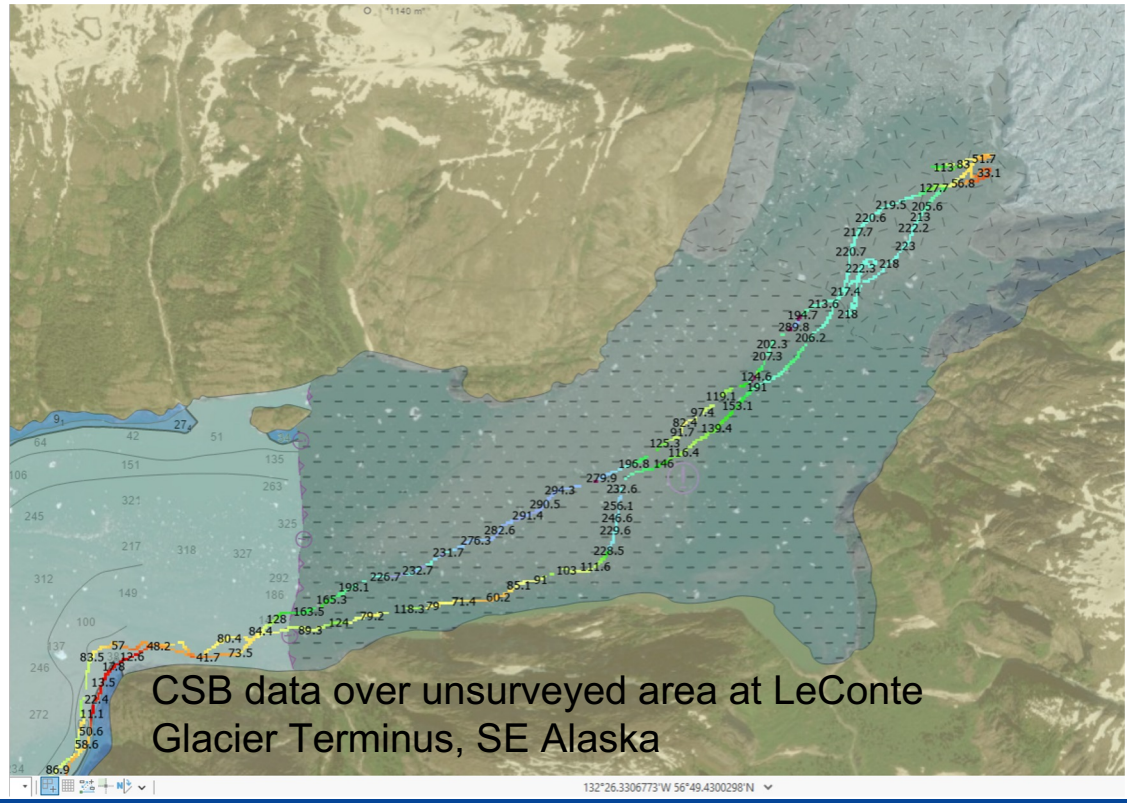
To this:



Why we are doing this?

KEY USE CASES FOR CSB WITHIN NOAA:

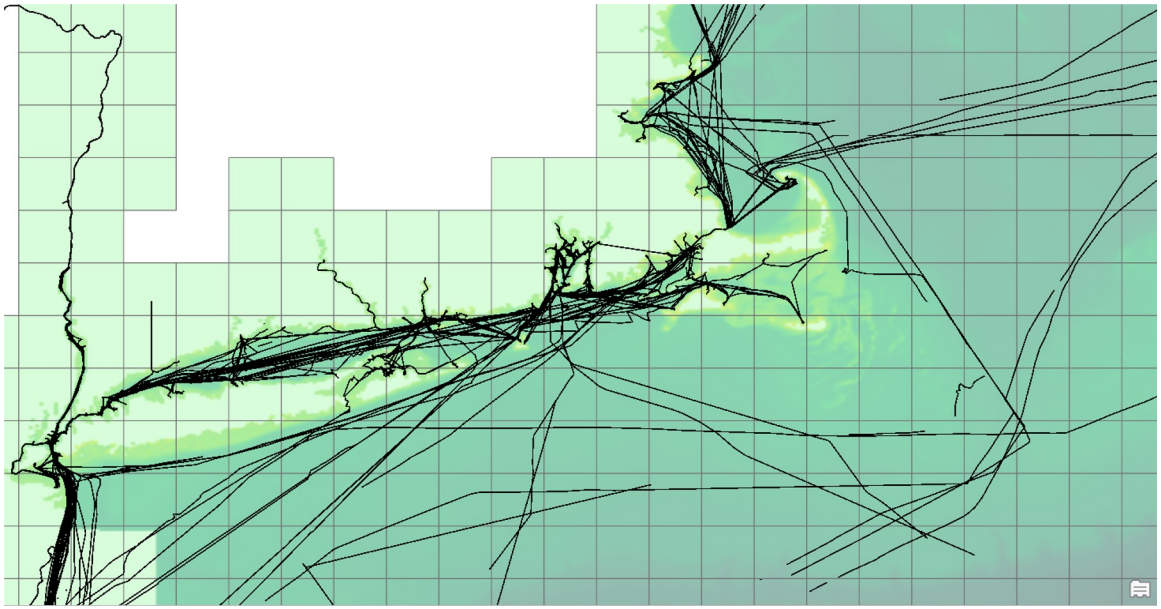
- Filling in gaps and improving our bathymetric record (i.e. crowd-to-chart)
- Detection of chart discrepancies and change detection
- Reconnaissance to increase efficiency and safety of field hydrography



The Value of CSB Data
A full Crowd-to-Chart Data Pipeline

Automated data pipeline currently in operational beta testing

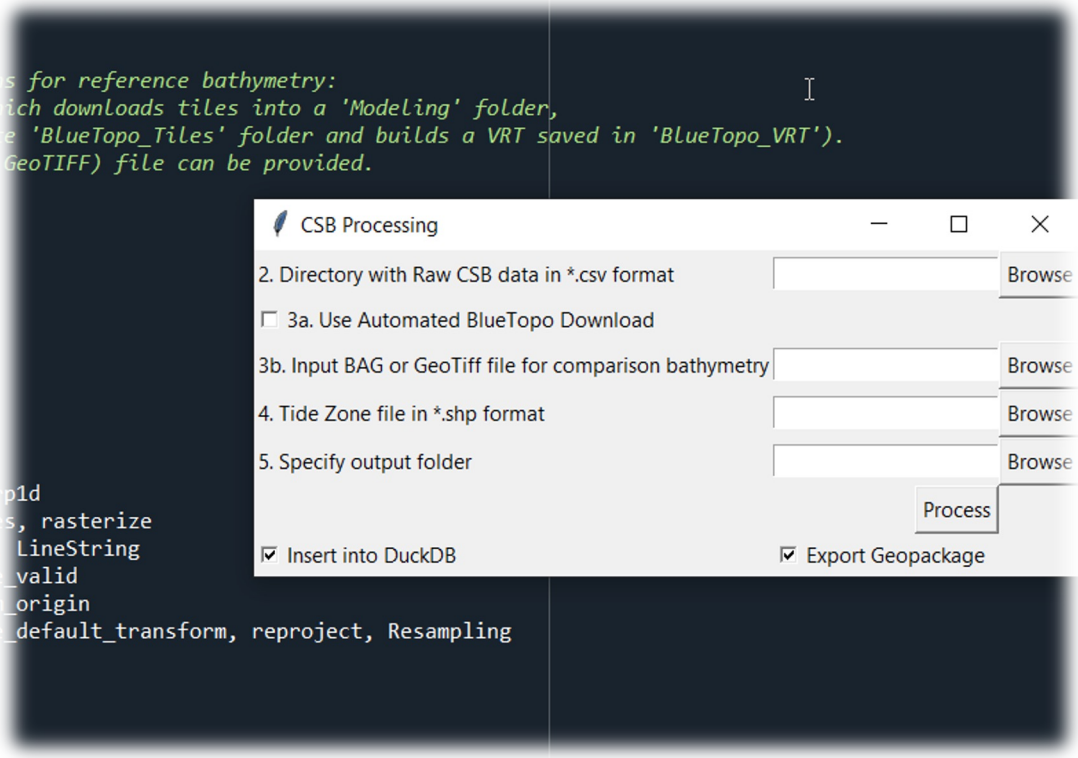
- Scraper developed to programmatically extract CSB data from DCDB via API
- Raw data scraped and processed tile-by-tile



The Value of CSB Data
A full Crowd-to-Chart Data Pipeline

Automated data pipeline currently in operational beta testing

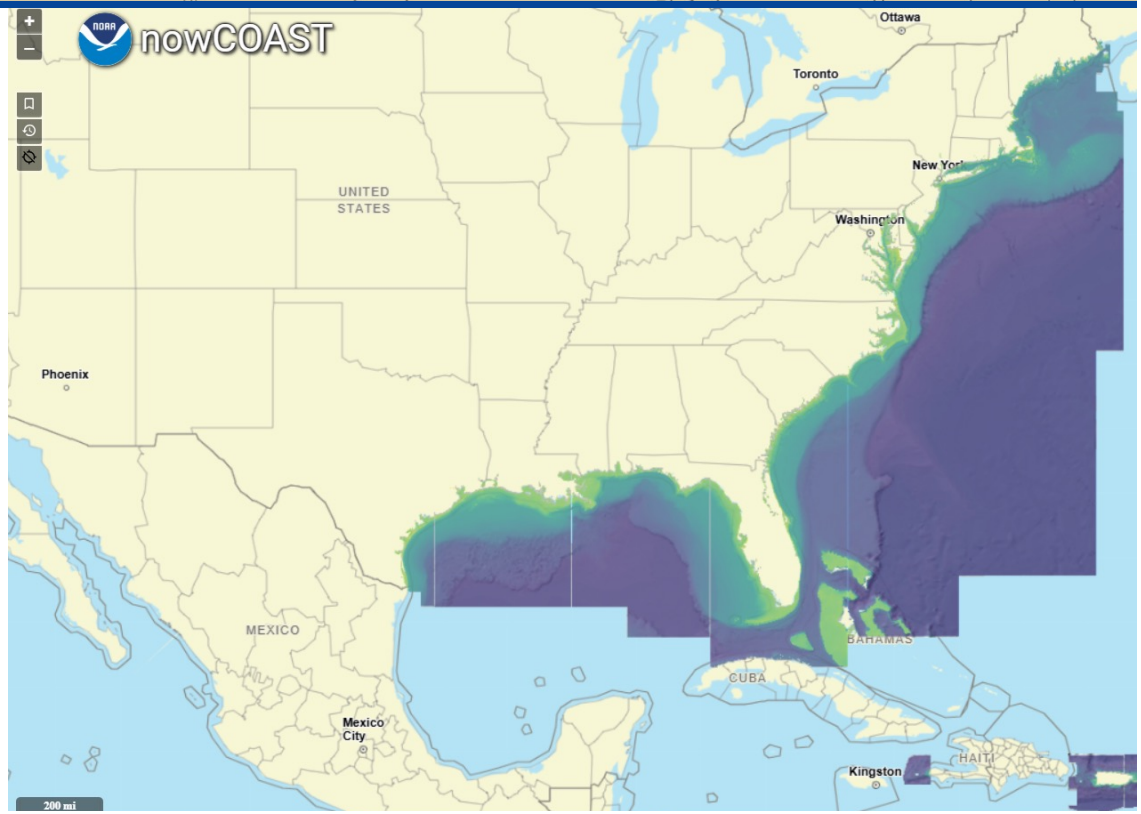
- Processing software from opensource libraries
- Initial ETL process filters obviously erroneous data (less than 0.5m, greater than 1000m), and dates outside the possible submission range
- High-speed vessels filtered later



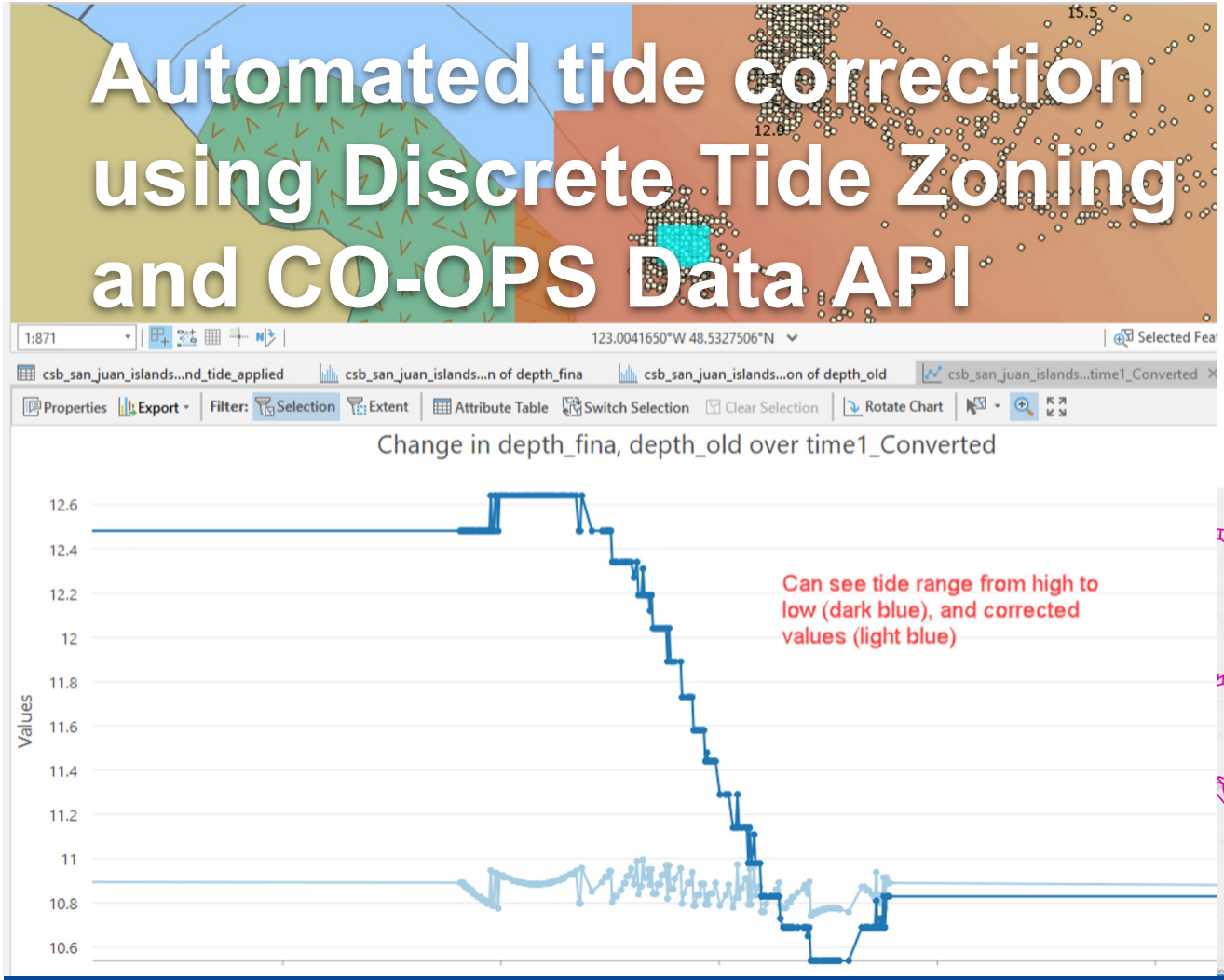
Reference Bathymetry Critical to Processing Pipeline

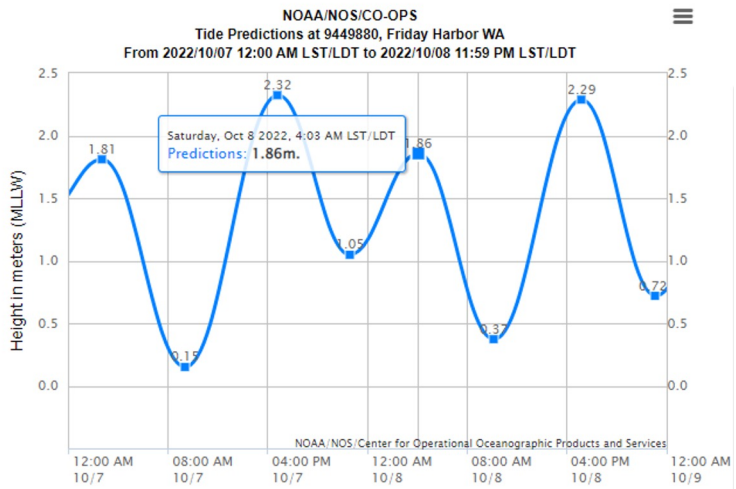
NOAA BlueTopo Reference Bathymetry

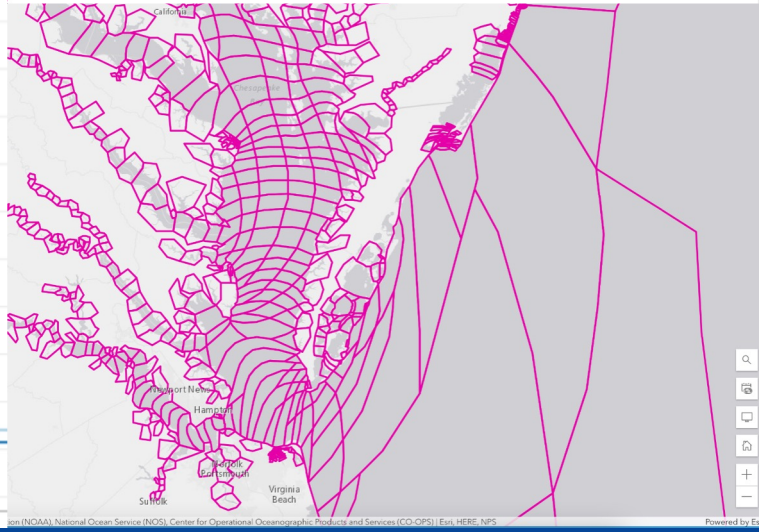
- High quality reference bathymetry fetched via spatial query API
- Optionally could use individual BAG files
- Critical for estimating transducer drafts and assessing data quality



Automated tide correction using Discrete Tide Zoning and CO-OPS Data API

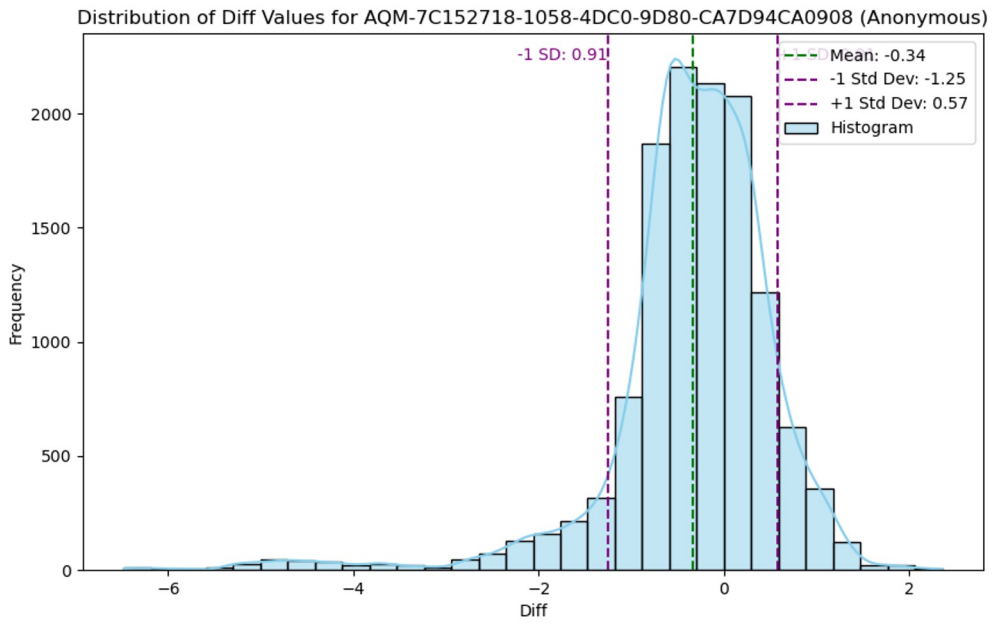






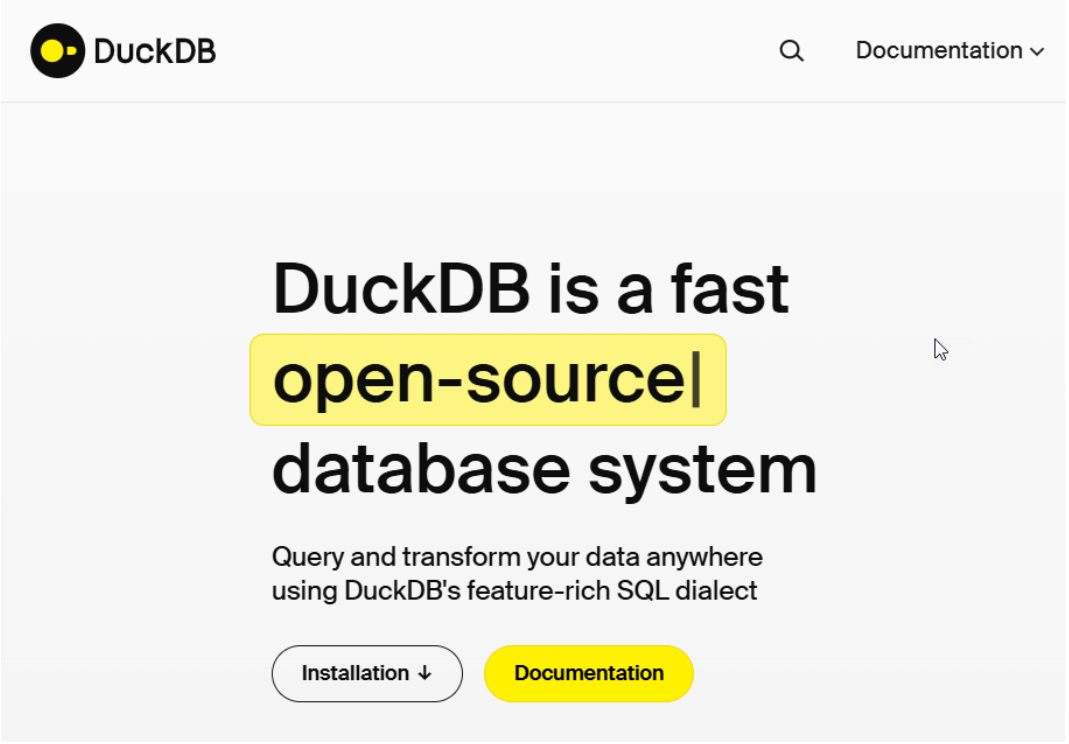
Transducer Draft Estimation

- Vertical bias detection and comparative analysis against reference bathymetry of known accuracy
- **Initial** transducer draft estimates done on tile-by-tile basis



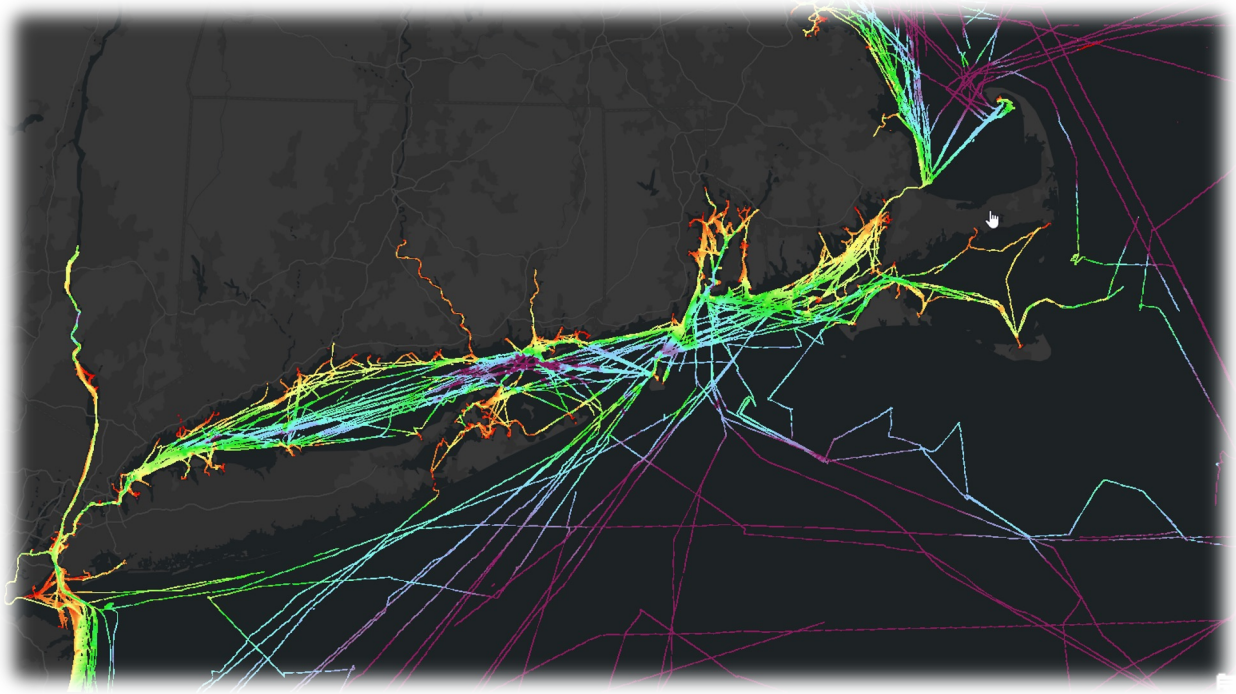
Ingestion of Spatial Database

- Ingestion of Spatial Database of processed points from each Tile
- Uses DuckDB architecture for post-processing data analysis and deliverable export
  - Quality and contributor analytics dashboard
  - Point data and geotiff from **individual transits** exported from DuckDB and delivered to NOAA's External Source Data Team



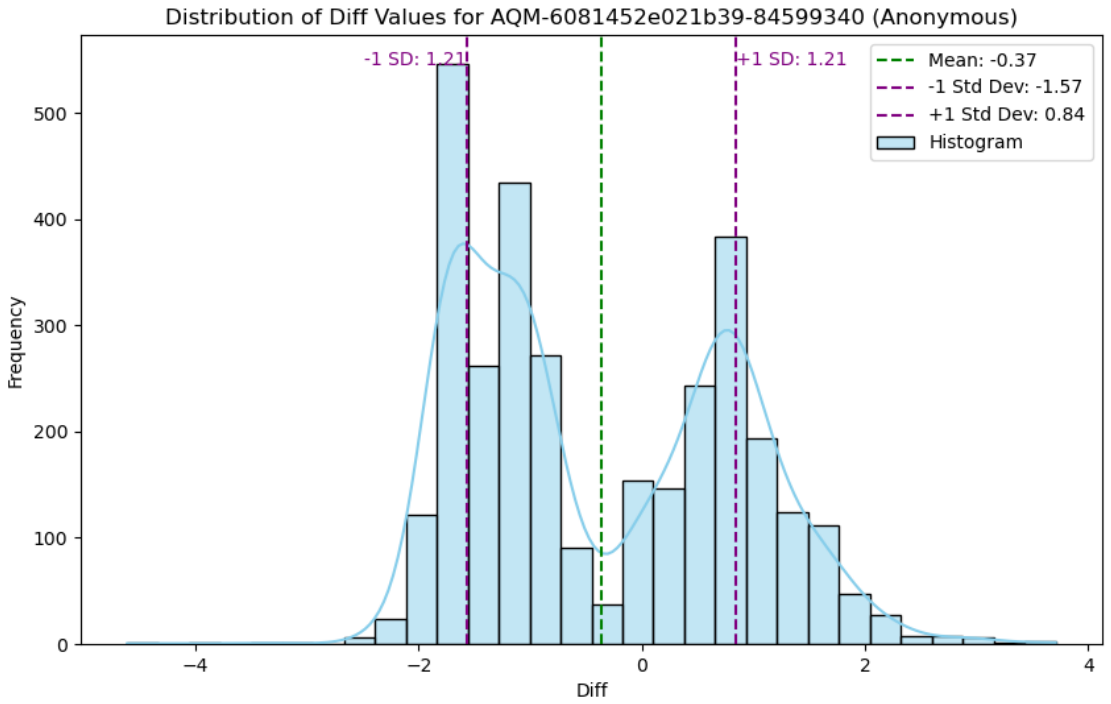
Additional Transducer Offset Calculations

- Via SQL commands, tide-corrected data is compared to reference bathymetry.
- New transducer offsets are **calculated based on data from all tiles,** (not restricted to tile boundaries as before)
- This improves transducer draft estimates from off-shore



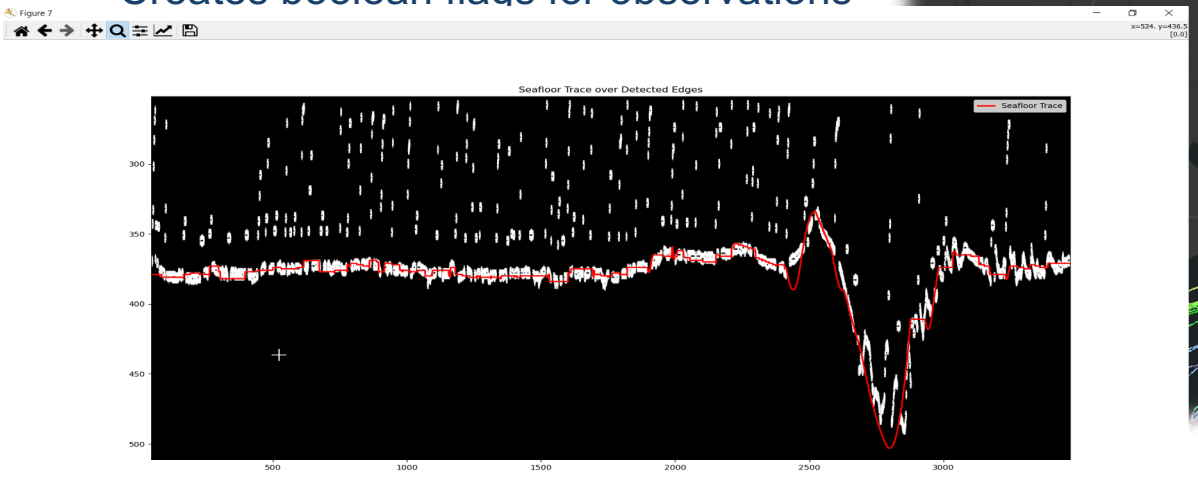
Future Development in Transducer Offset Estimations

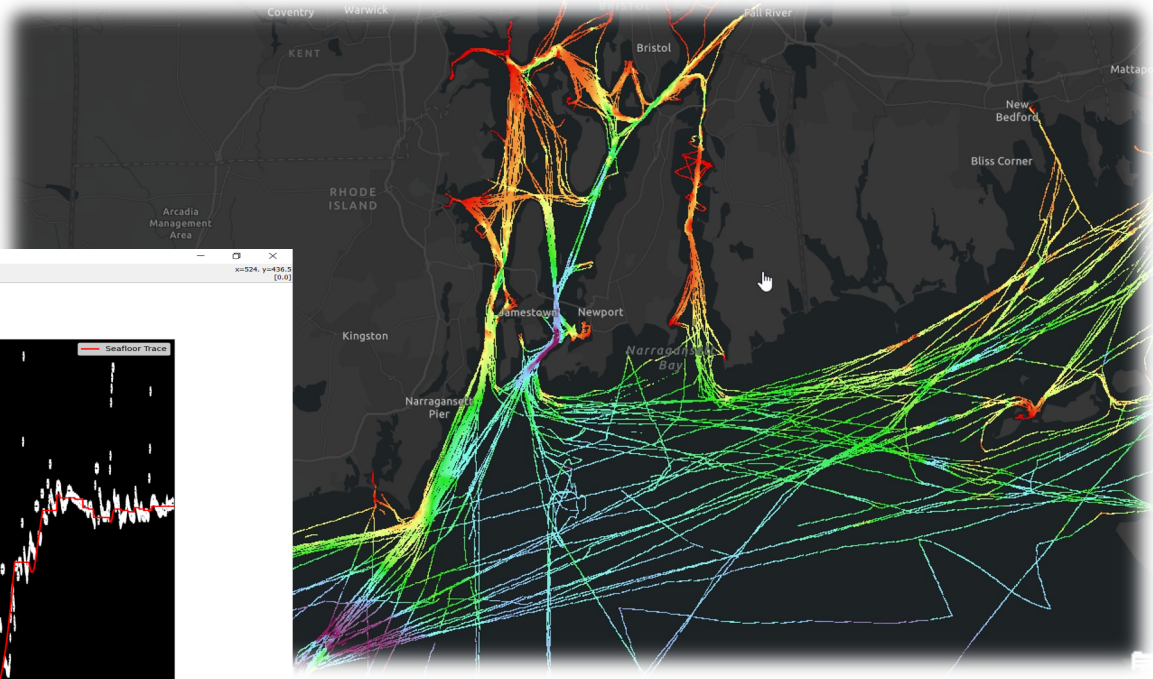
- To account for larger changes in vessel draft due to loading/unloading
- From analysis of bimodal distribution of difference histogram, transducer drafts can be apportioned more accurately based on the time of those measurements.



Testing Automated Outlier Detection Strategies

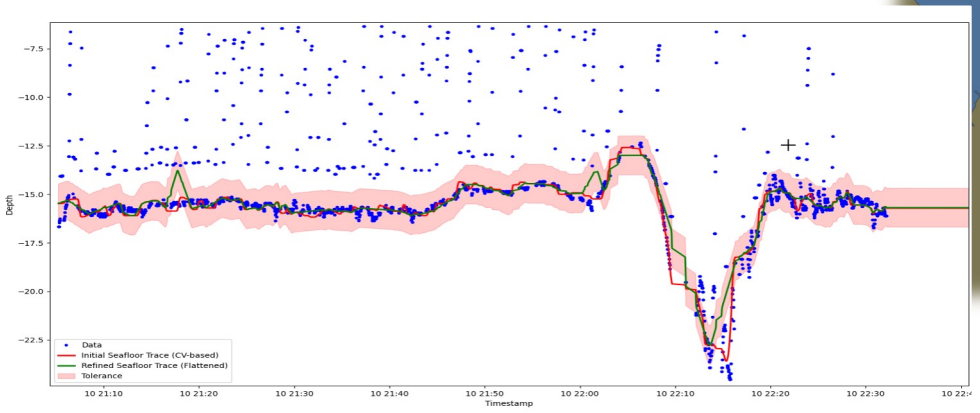
- Outlier detection algorithm on timeseries dimension utilizing Computer Vision (OpenCV) and statistical methods
- Creates boolean flags for observations

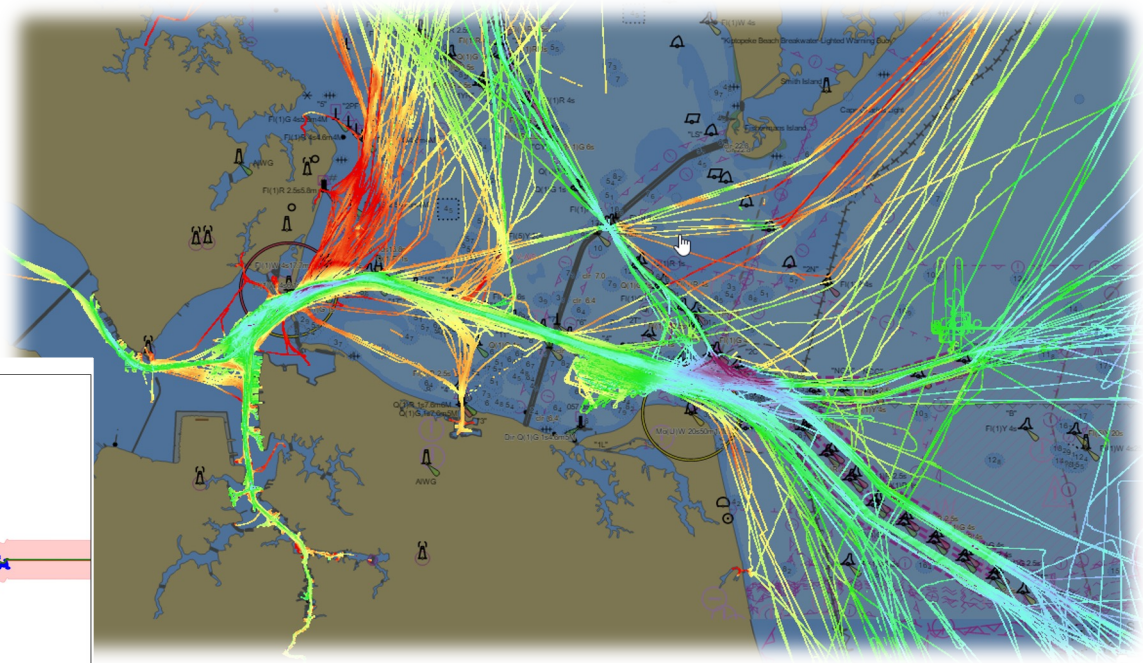




Testing Automated Outlier Detection Strategies

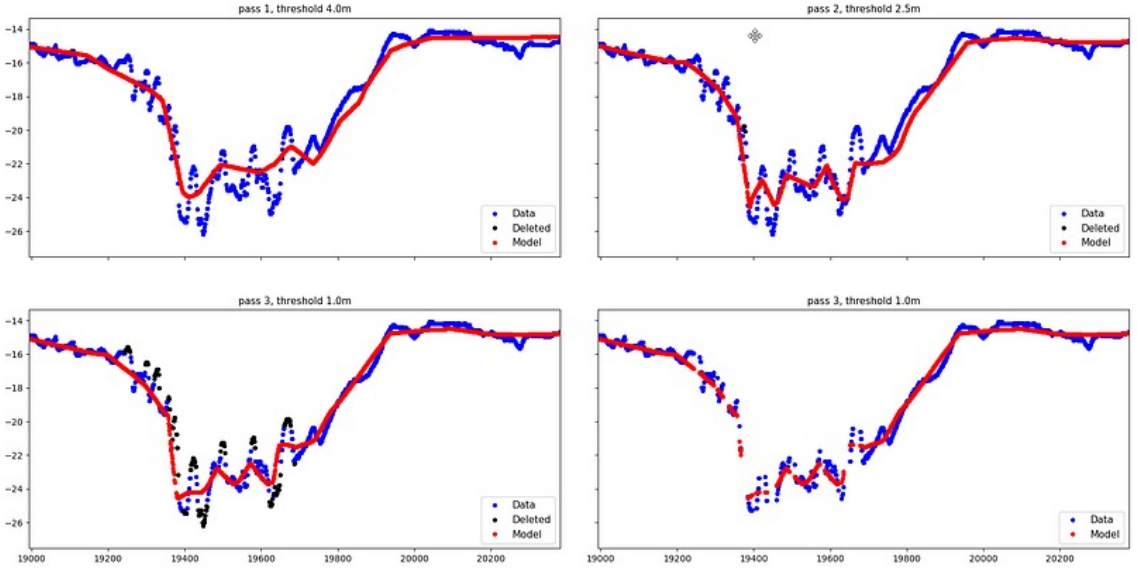
- Outlier detection algorithm on timeseries dimension utilizing Computer Vision (OpenCV) and statistical methods
- Creates boolean flags for observations that may be outliers

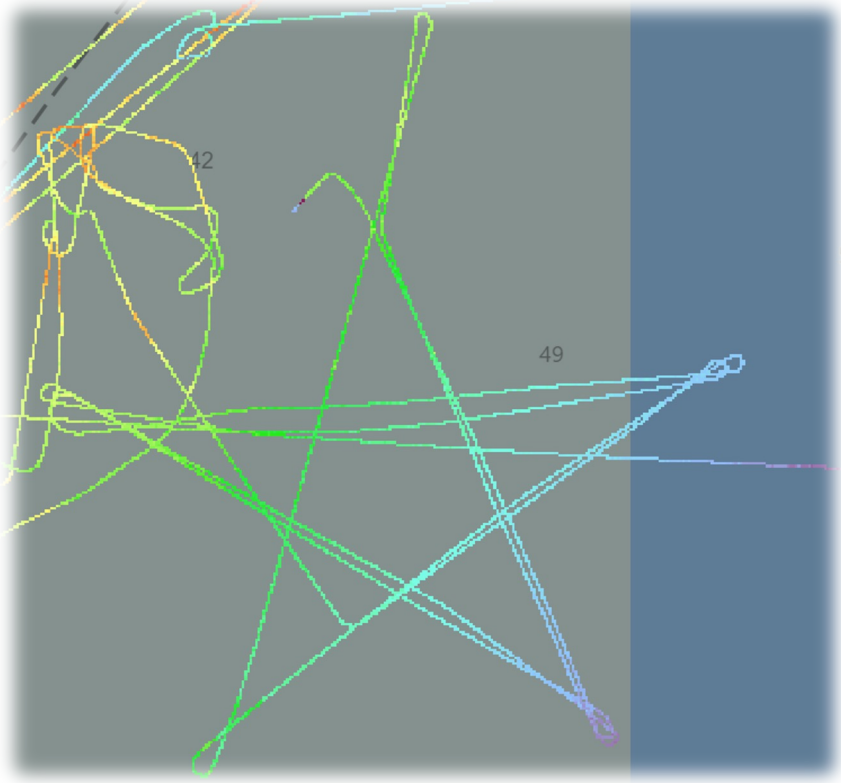




Testing Automated Outlier Detection Strategies

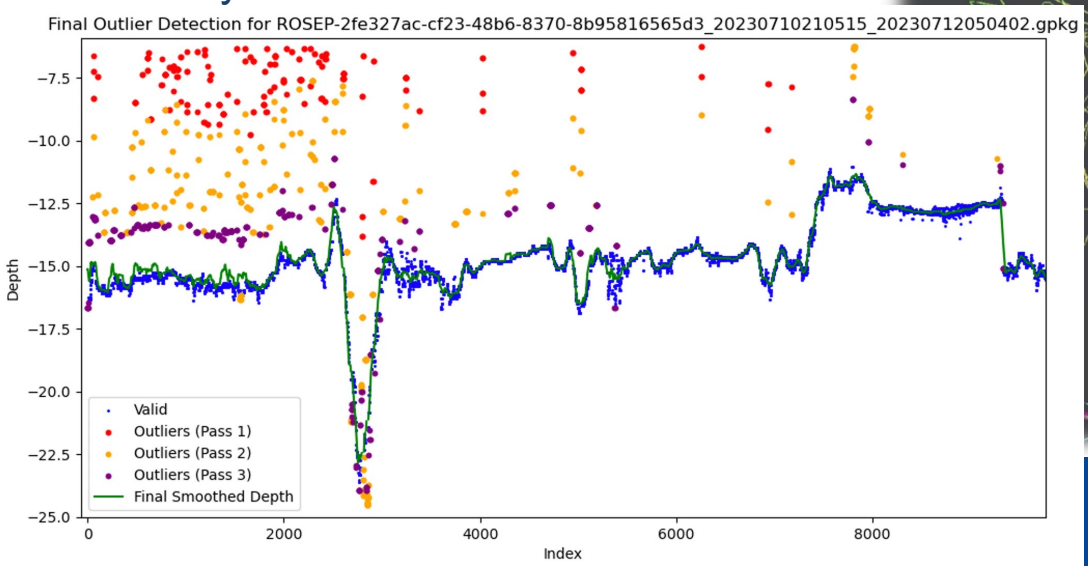
- Sequential Neural Network Outlier Detection Model – more accurate than CV method, but very slow to compute

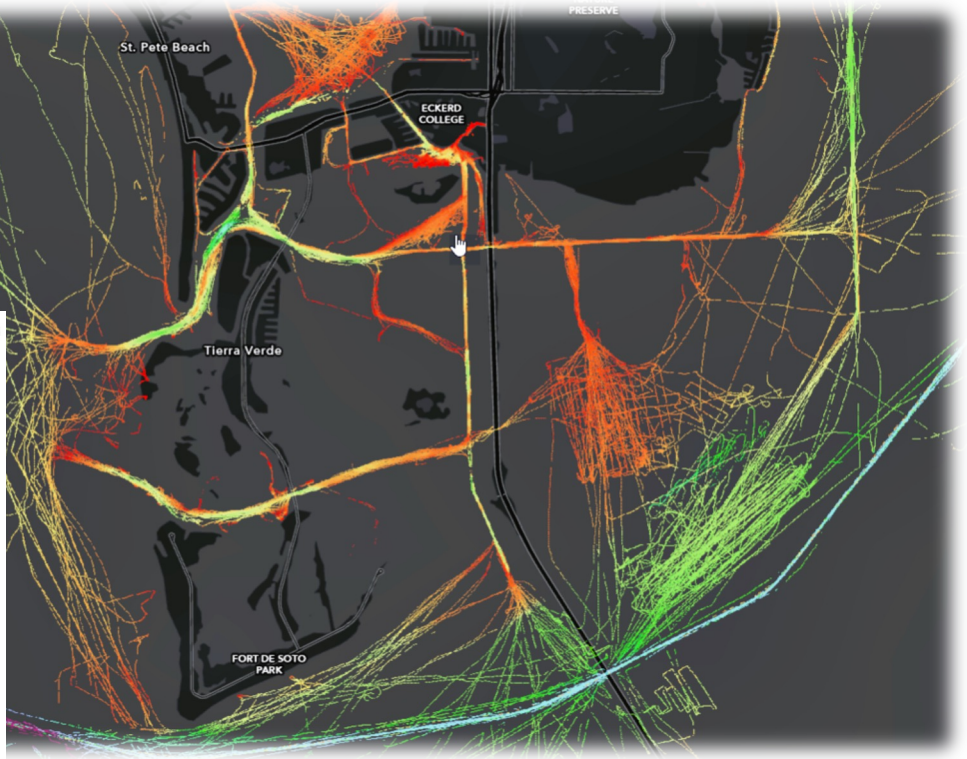




Testing Automated Outlier Detection Strategies

- Outlier detection algorithm on timeseries dimension utilizing **Predictive Mean Matching Imputation** algorithm thresholding – 1000x faster than SNN
- Creates boolean flags for observations that may be outliers

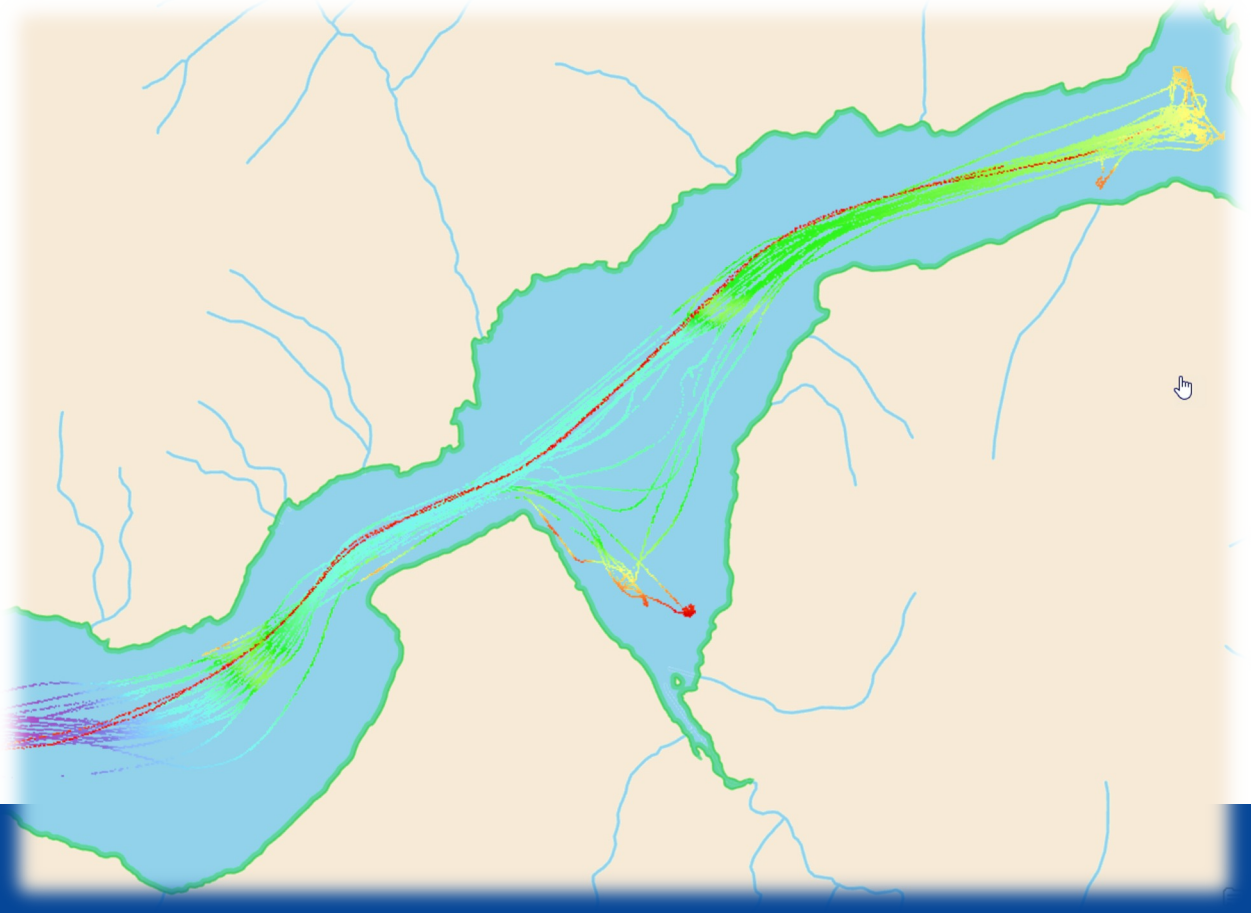




Testing Automated Outlier Detection Strategies

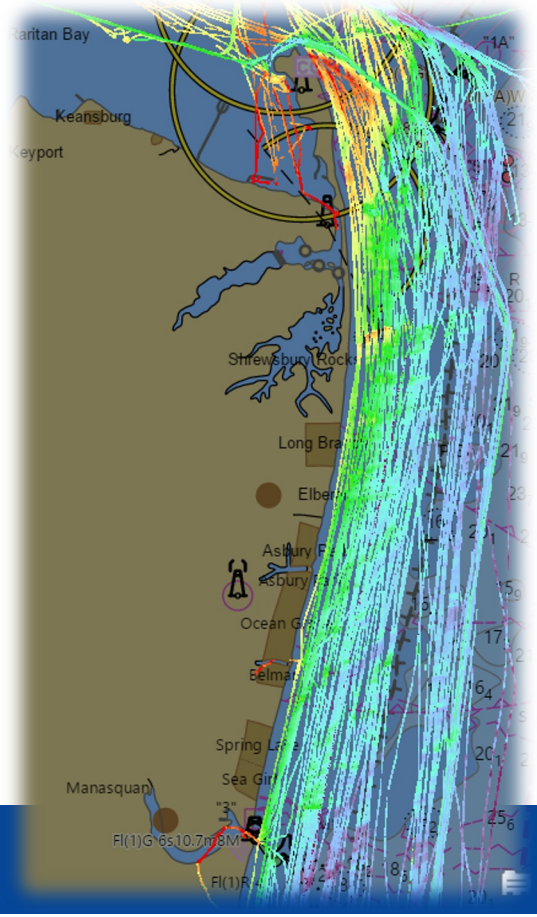
- Outliers still exist - but detection algorithms are getting better
- Future development:

  detecting and removing contiguous segments of lost bottom tracking via reference bathymetry comparisons

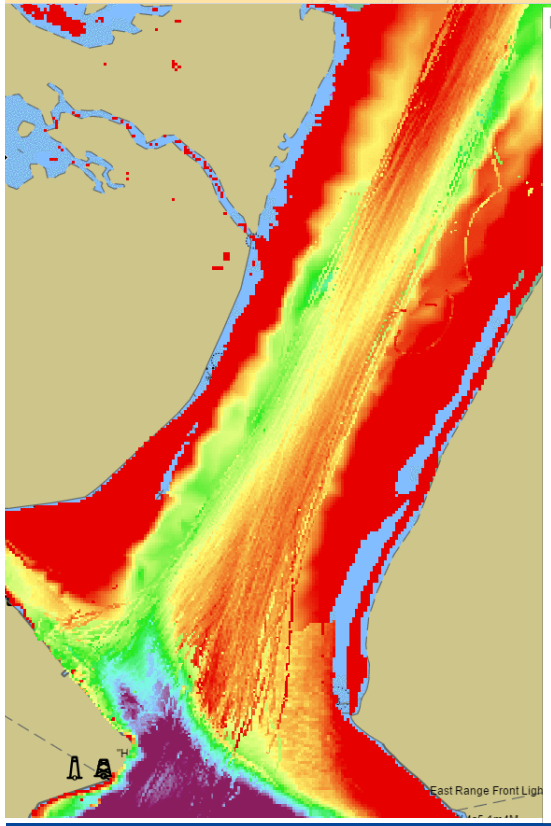


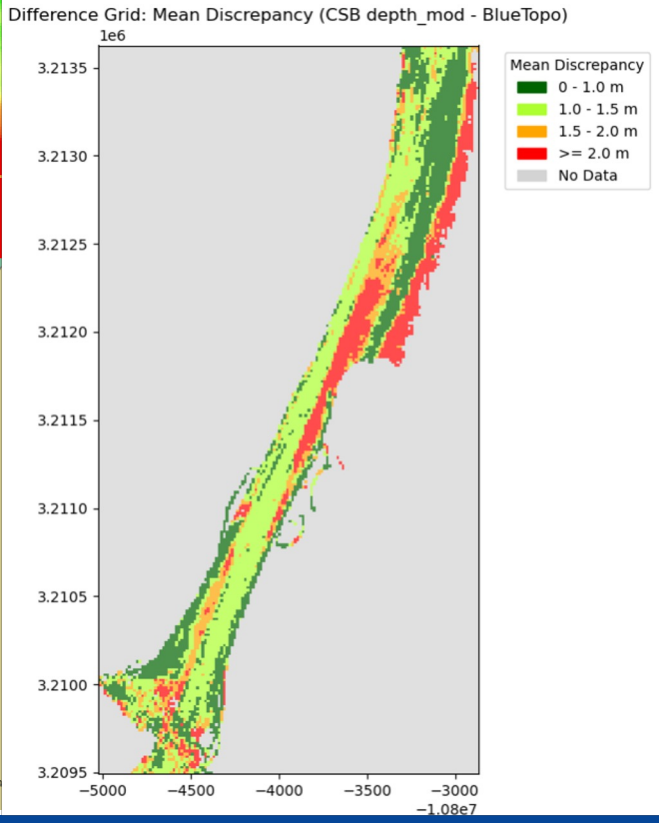
Final Product Export from DuckDB Spatial Database

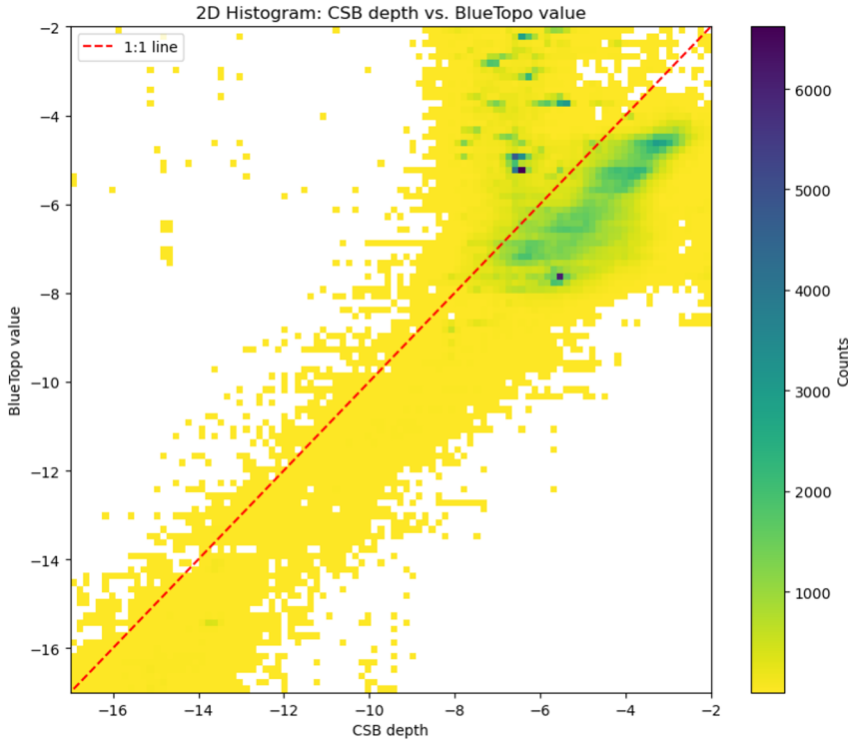
- Once outliers are detected, final products are exported from the DuckDB spatial database
- Geotiffs and Geopackages with point data are filtered for outliers and vessel speeds above 20 knots (is that the right number?)
- Intend to create WMTS of processed CSB rasterized to 10m resolution with option to spatially query and fetch processed point data from DuckDB spatial database



Automated Comparison Products
Discrepancy Modelling



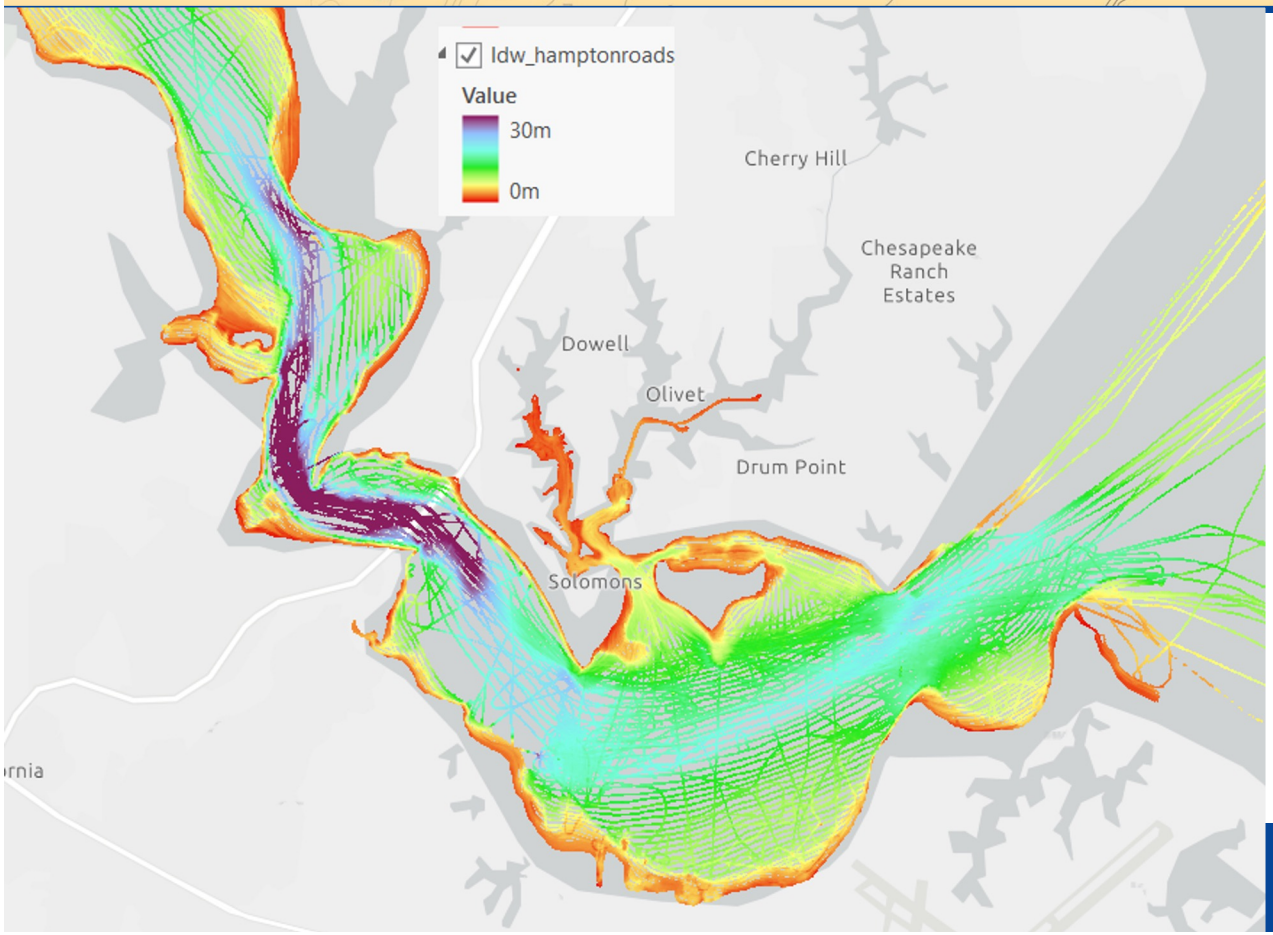


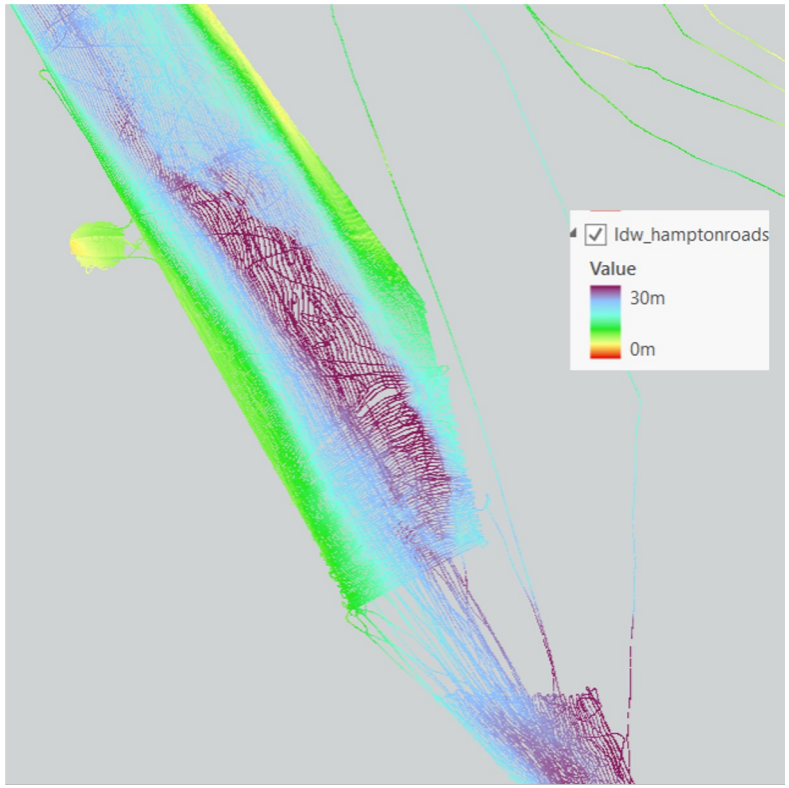


(Texas) Gulf Intracoastal Waterway

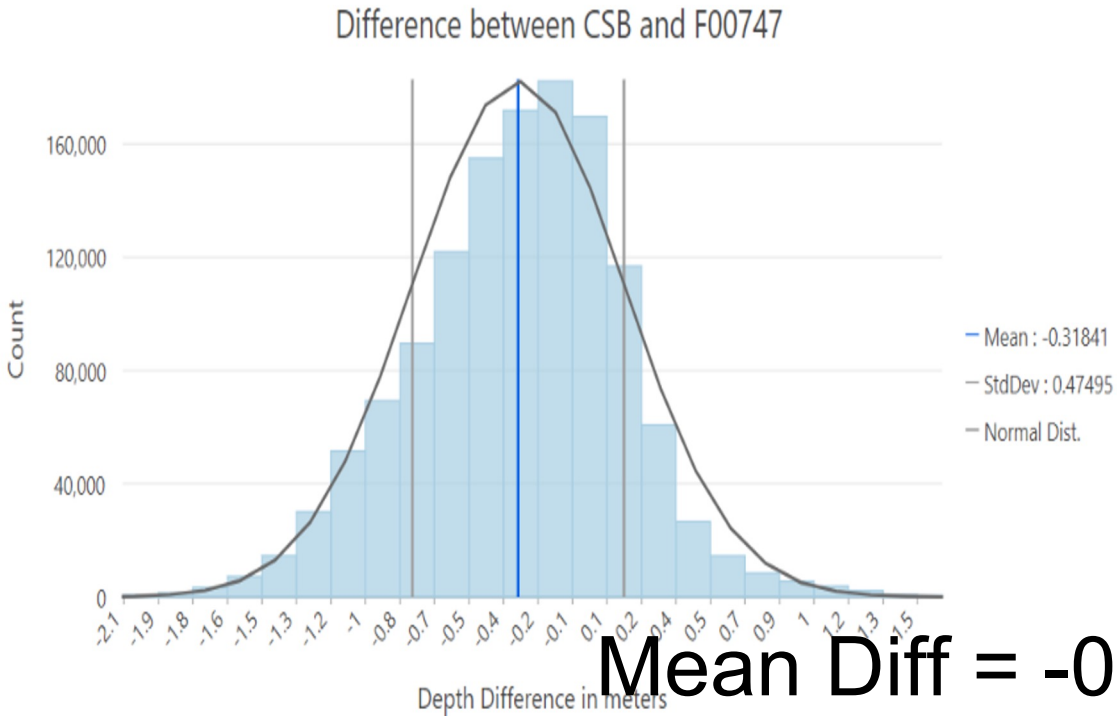
CSB Quality Assessment - Bay Hydro II

collected CSB during hydrographic surveys



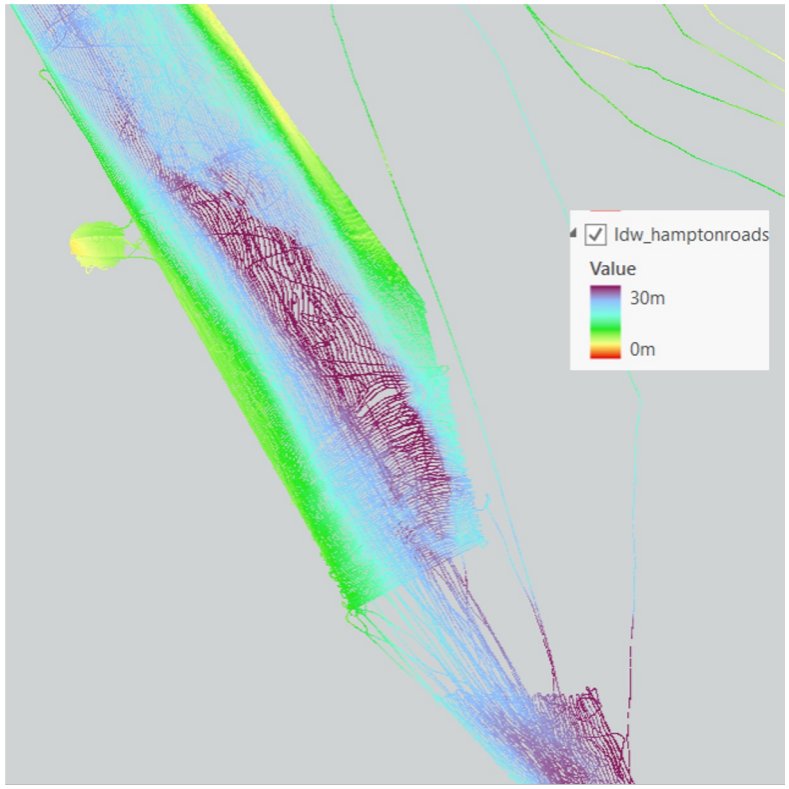


CSB Quality Assessment - Bay Hydro II collected CSB during hydrographic surveys

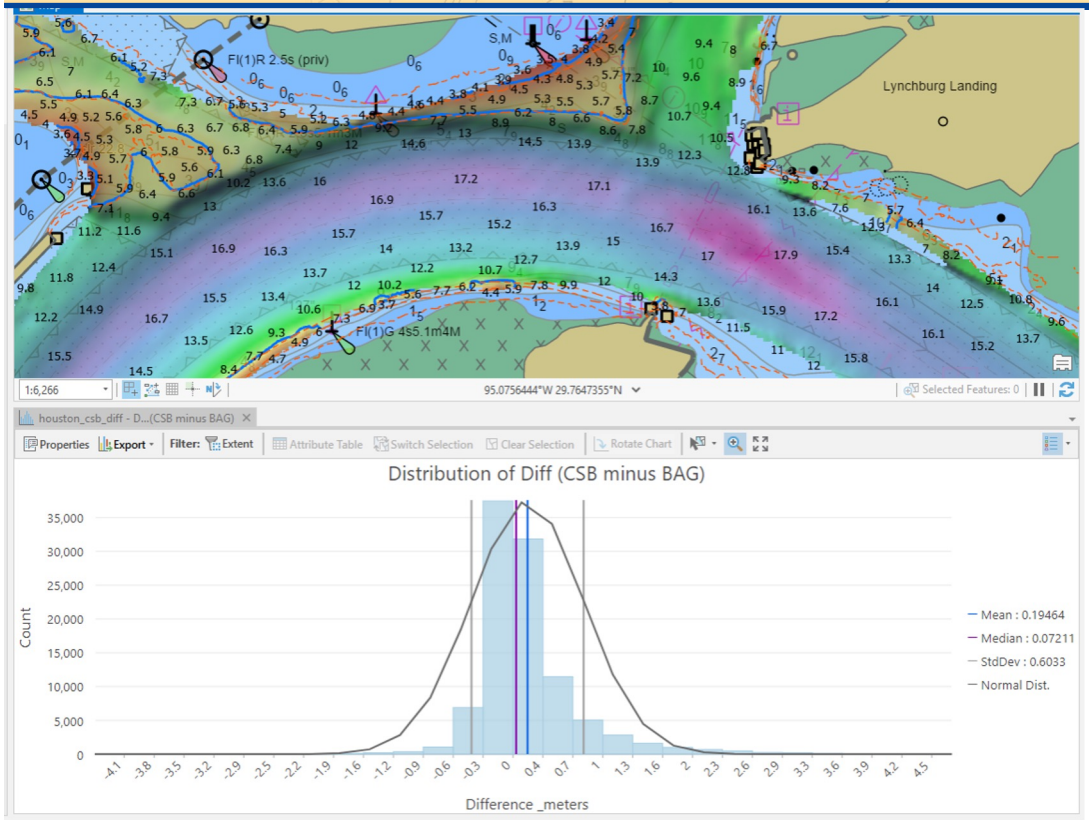


Mean Diff = -0.31

Std dev = 0.47m



Preliminary results are promising



Houston, TX

Comparison of CSB to recent survey:

Mean difference: 0.19 m
Standard deviation: 0.60 m

Processed CSB Quality Characterization:

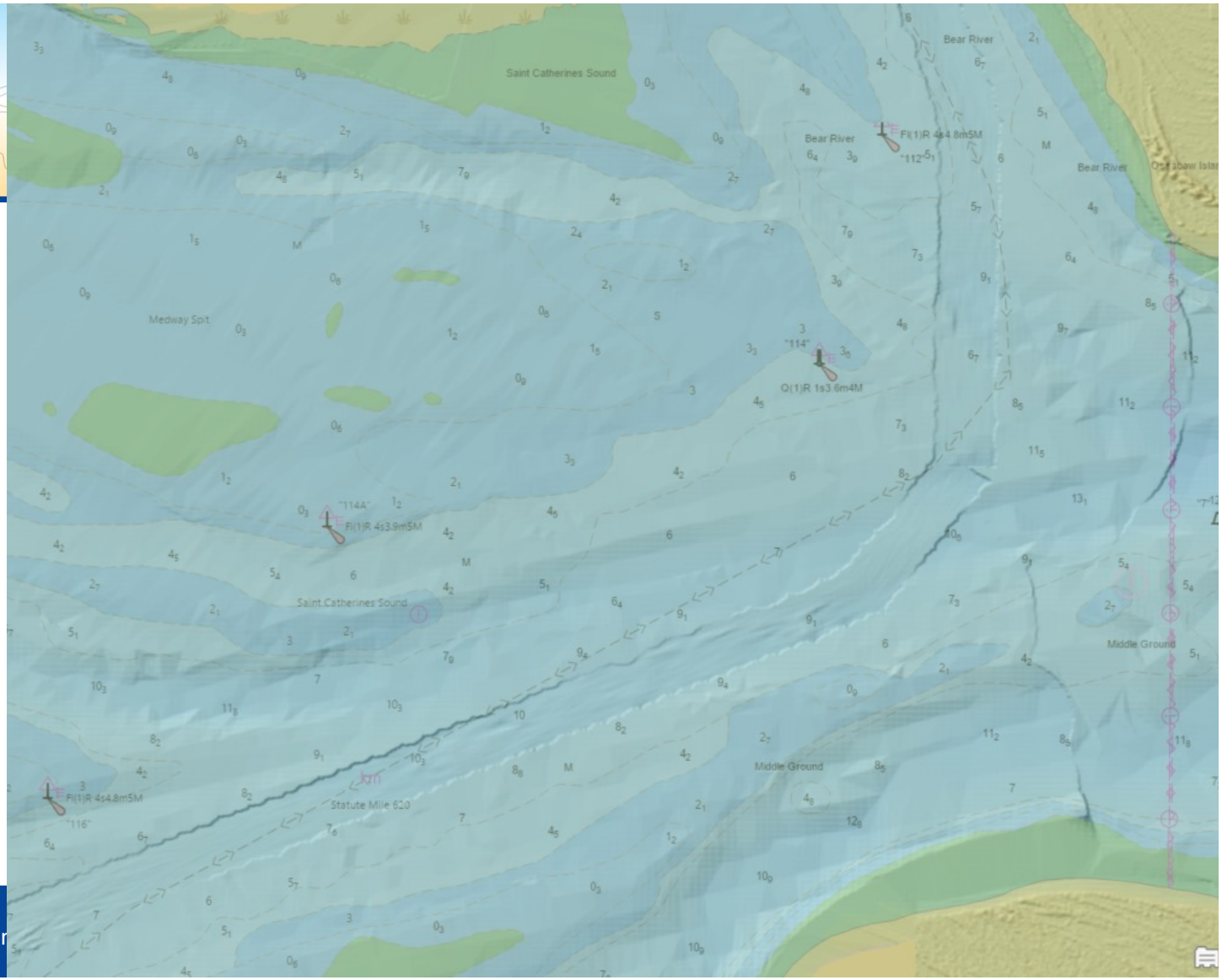
Processed CSB accuracy is generally CATZOC C capable (or better)

Table 4-1 – ZOC Categories

ZOC	Position accuracy	Depth accuracy	Seafloor coverage
A1	± 5 m + 5% depth	0.50 m + 1% depth	Full area search undertaken. Significant seafloor features detected and depths measured.
A2	± 20 m	1.00 m + 2% depth	Full area search undertaken. Significant seafloor features detected and depths measured.
B	± 50 m	1.00 m + 2% depth	Full area search not achieved; uncharted features hazardous surface navigation are not expected but may exist
C	± 500 m	2.00 m + 5% depth	Full area search not achieved, depth anomalies may be expected.
D	Worse than ZOC C	Worse than ZOC C	Full area search not achieved, large depth anomalies may be expected.
U	Unassessed – The quality of the depth data has yet to be assessed.		

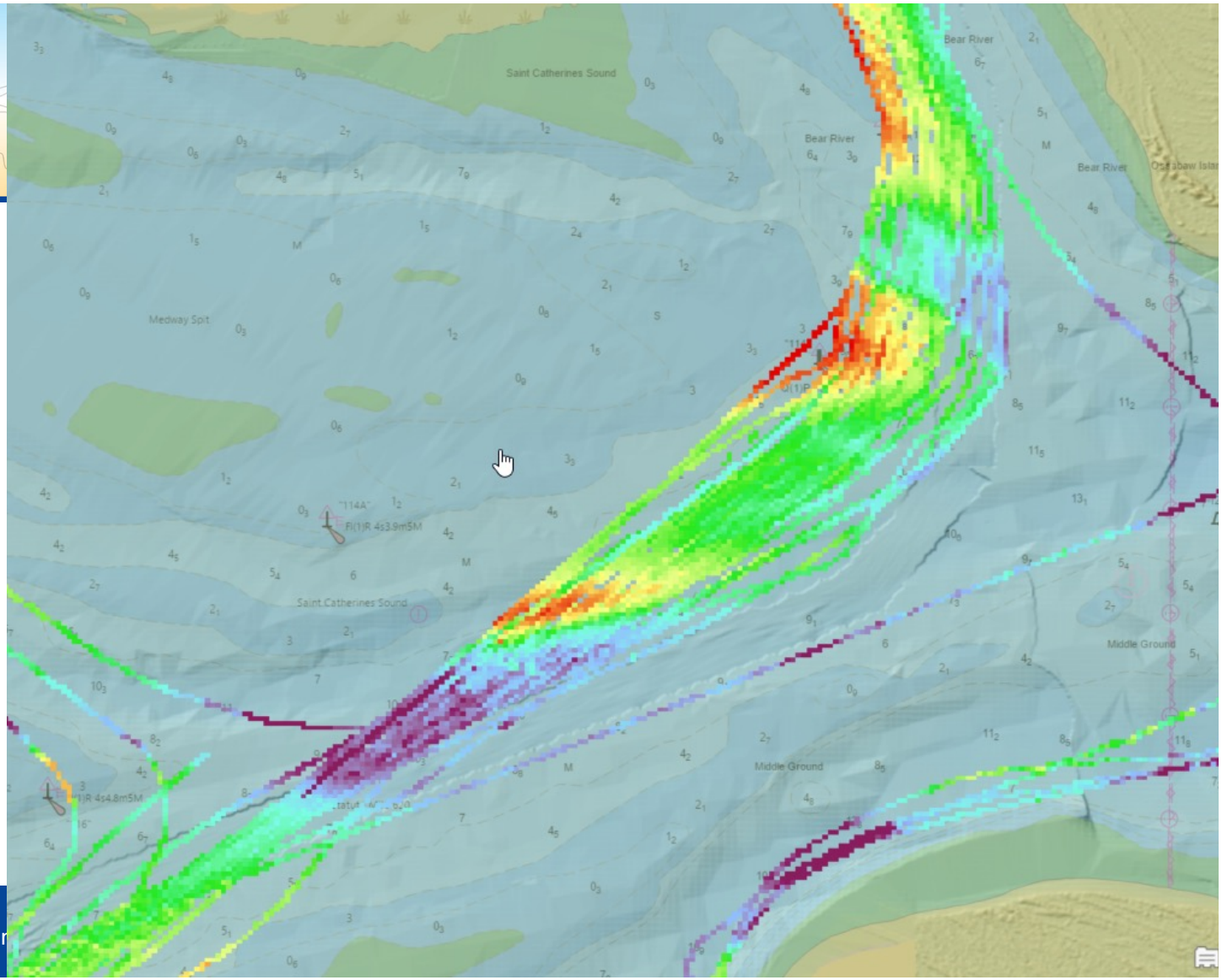
Coast Survey

- Detecting coastal change over time and bathymetric discrepancies in NBS
- Automated change - detection product updated as new data is processed



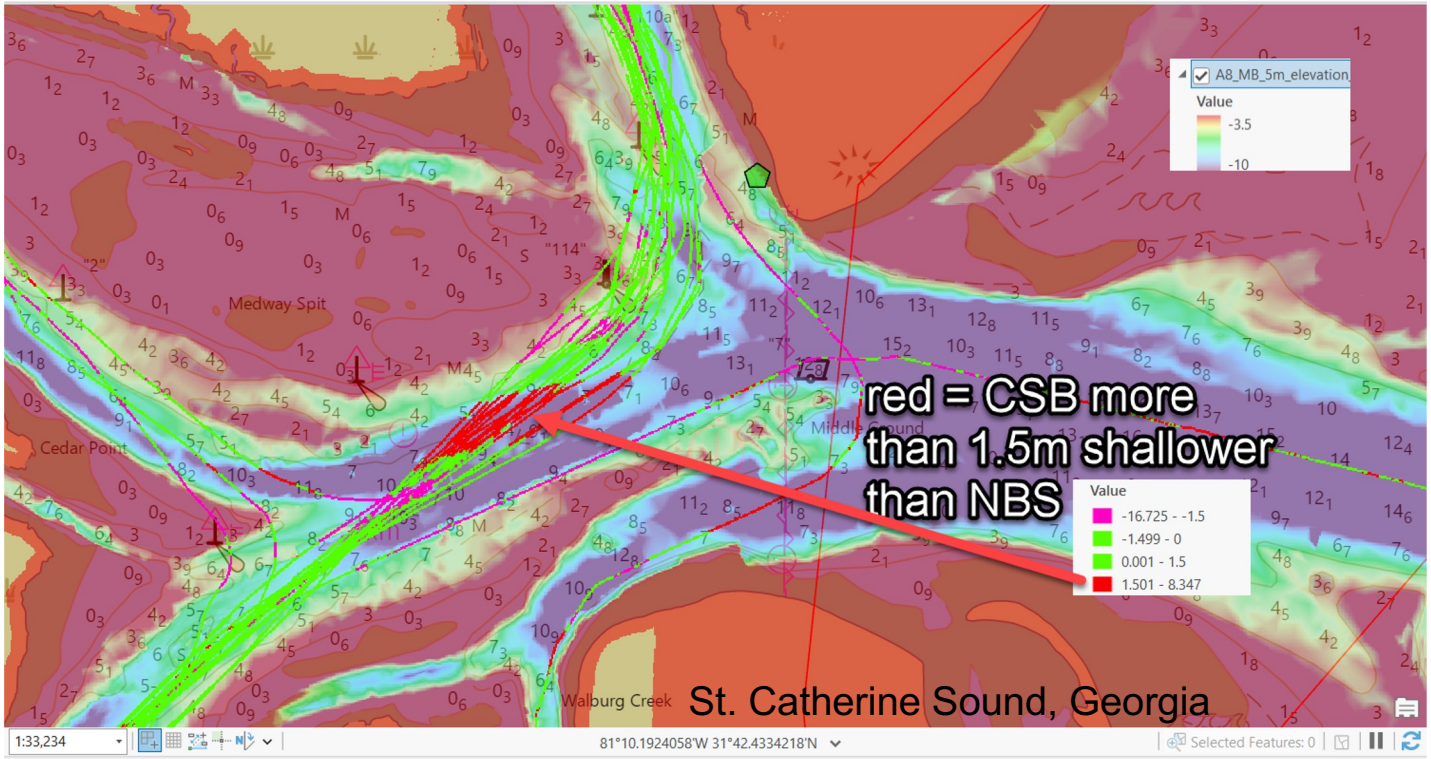
Coast Survey

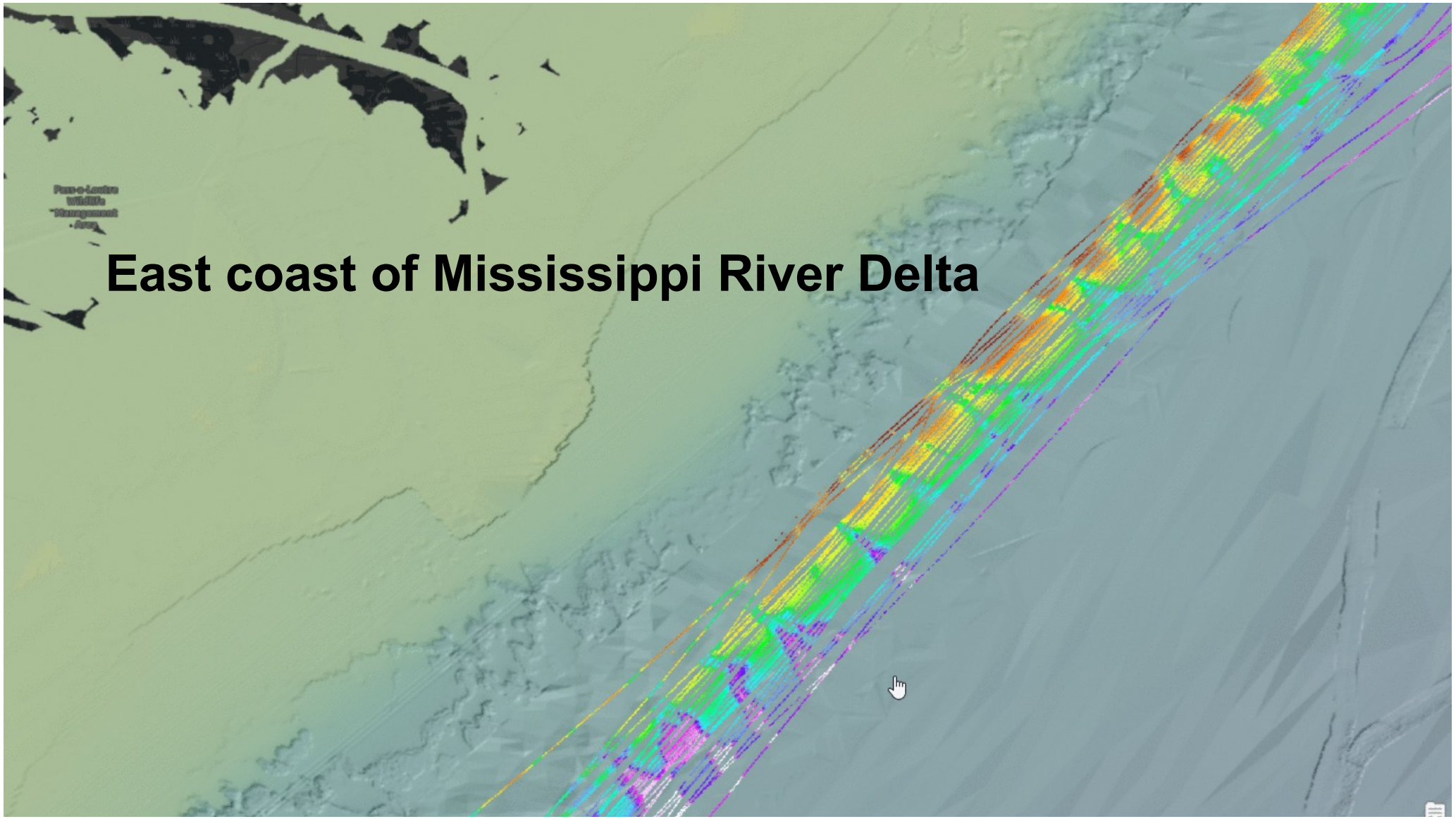
- Detecting coastal change over time and bathymetric discrepancies in NBS
- Automated change - detection product updated as new data is processed



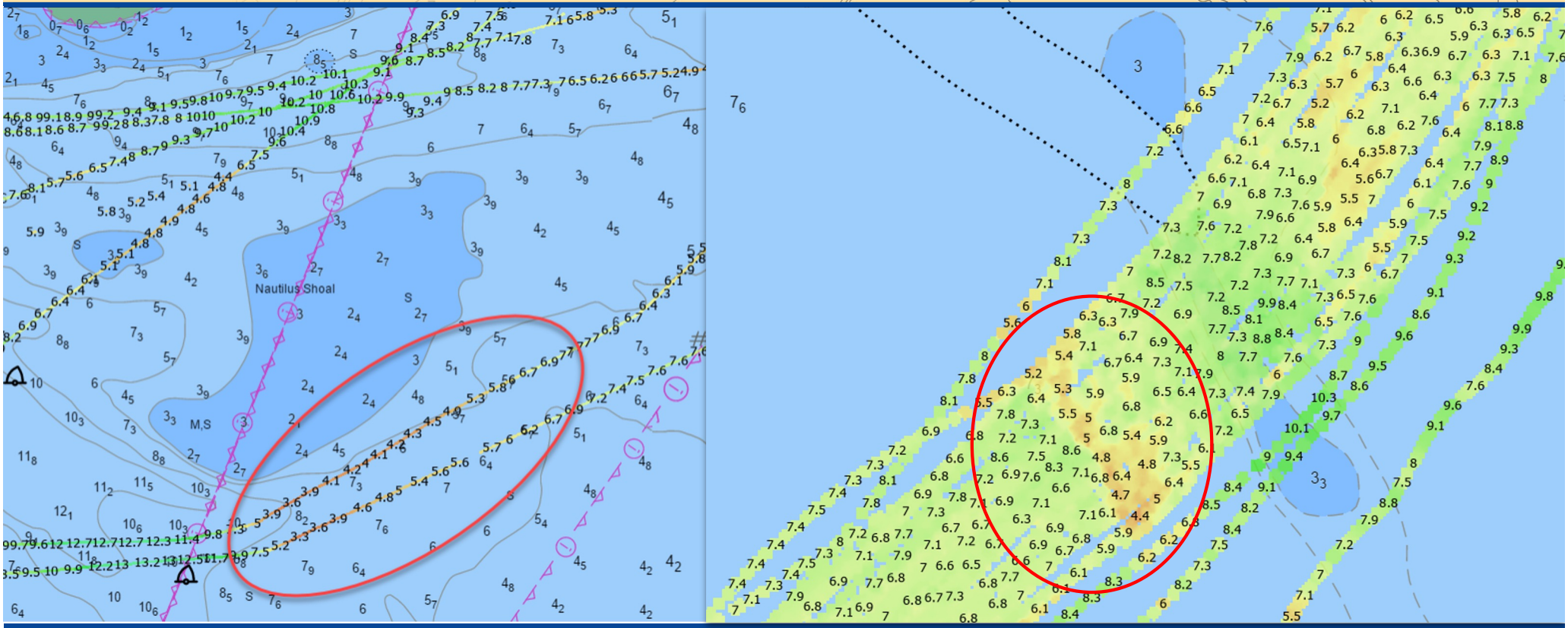
The Value of CSB Data
NBS Discrepancy Modelling

- Detecting coastal change over time and bathymetric discrepancies in NBS
- Automated change - detection product updated as new data is processed




East coast of Mississippi River Delta

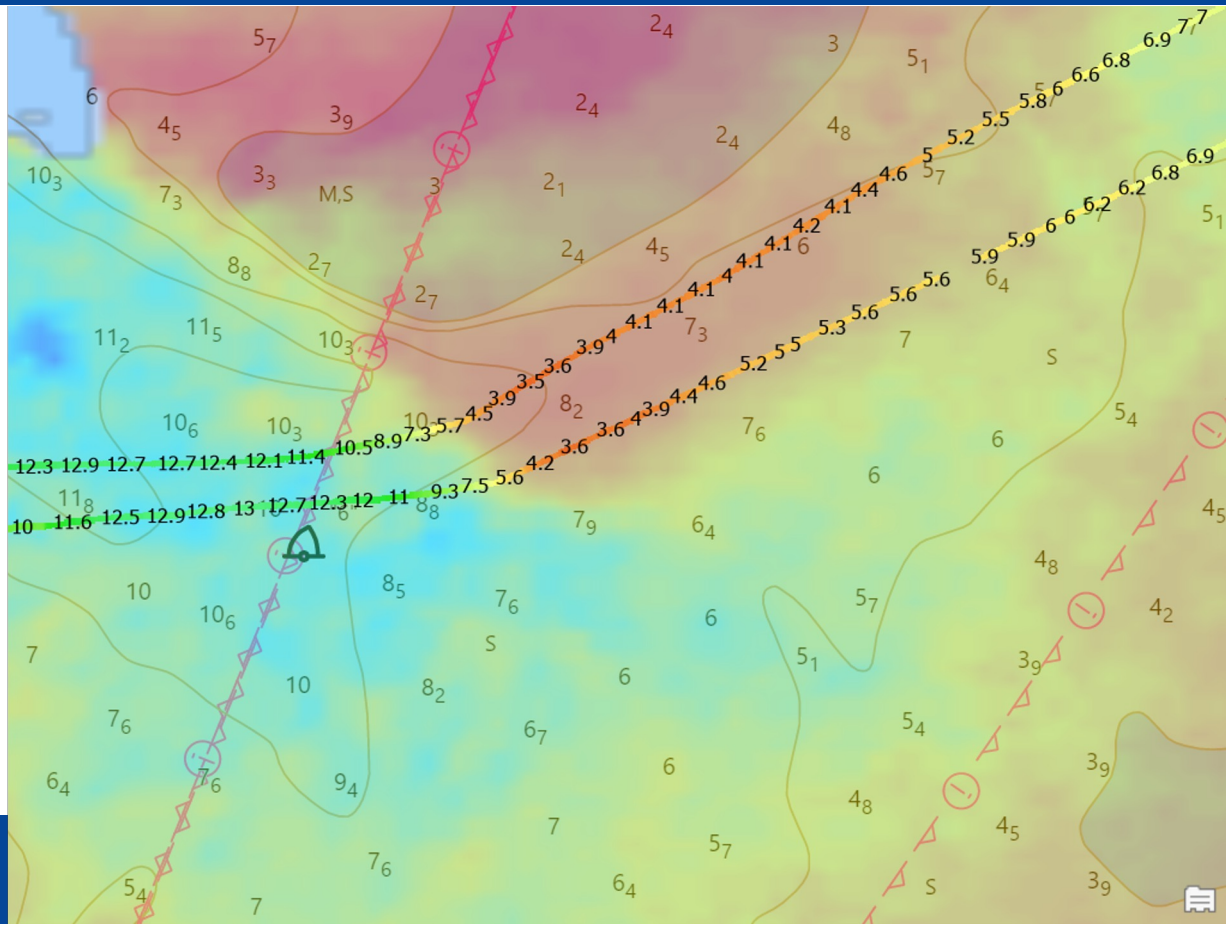
Mischarted Shoals Detected in Chesapeake and Delaware Bays



Office of Coast Survey
National Oceanic and Atmospheric Administration

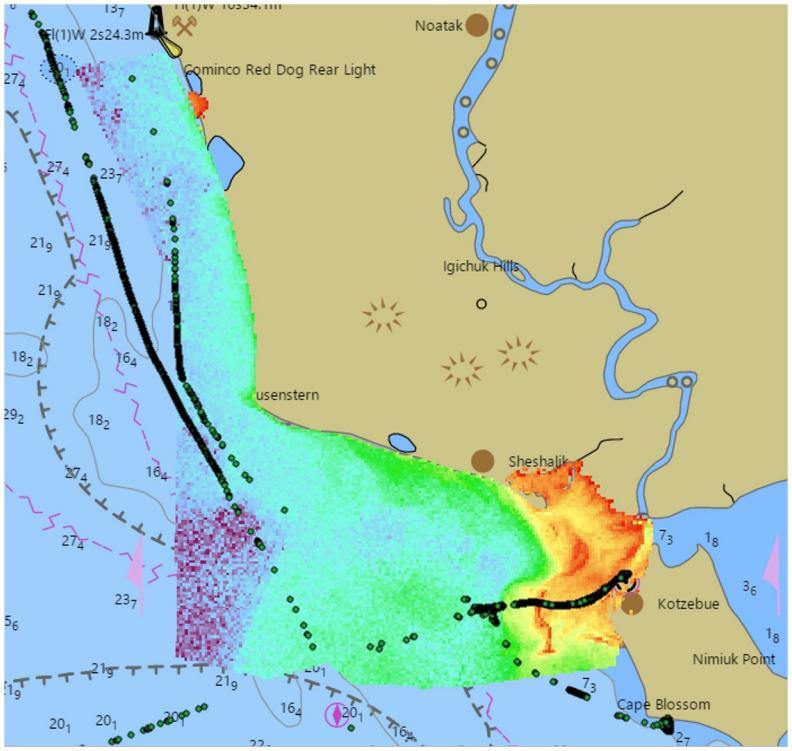
The Value of CSB Data
- SDB correlation and groundtruthing

- CSB detected and SDB confirmed shift of Nautilus Shoal in Mouth of Chesapeake Bay



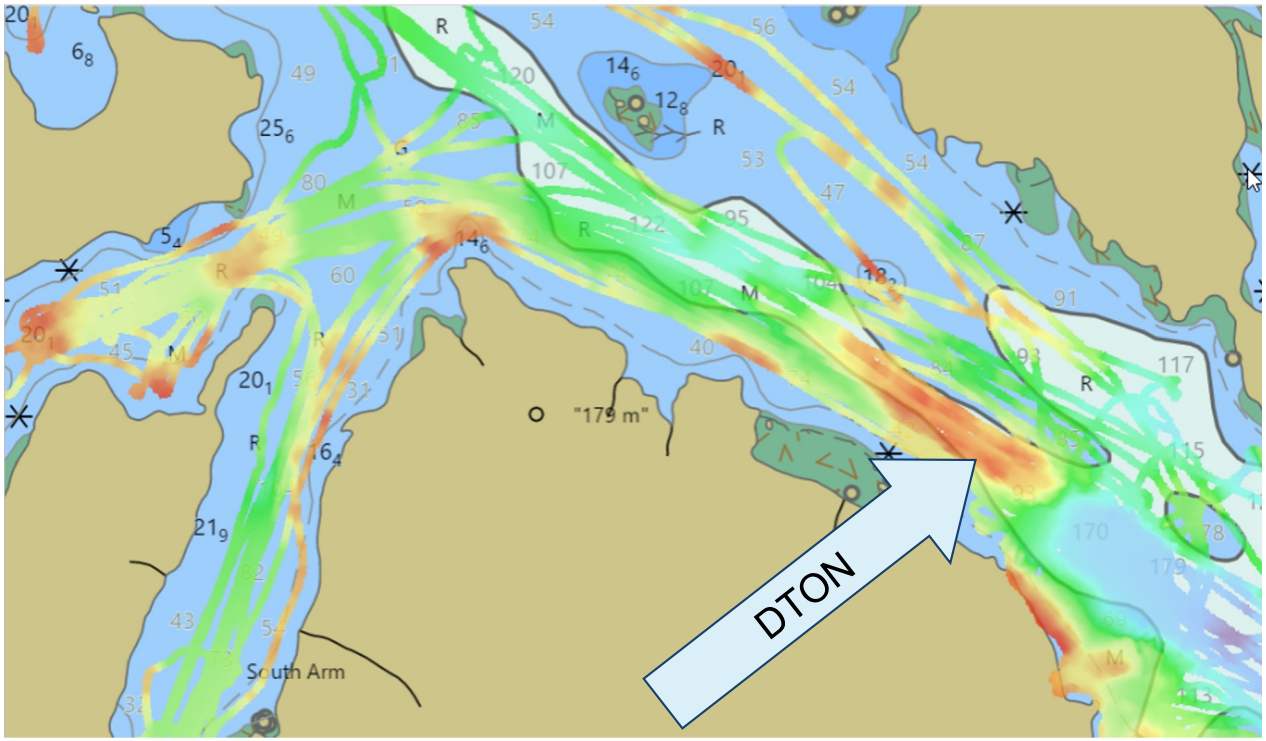
The Value of CSB Data
- SDB correlation and groundtruthing

CSB used in analysis of Satellite-Derived Bathymetry Products in Remote Alaskan Arctic



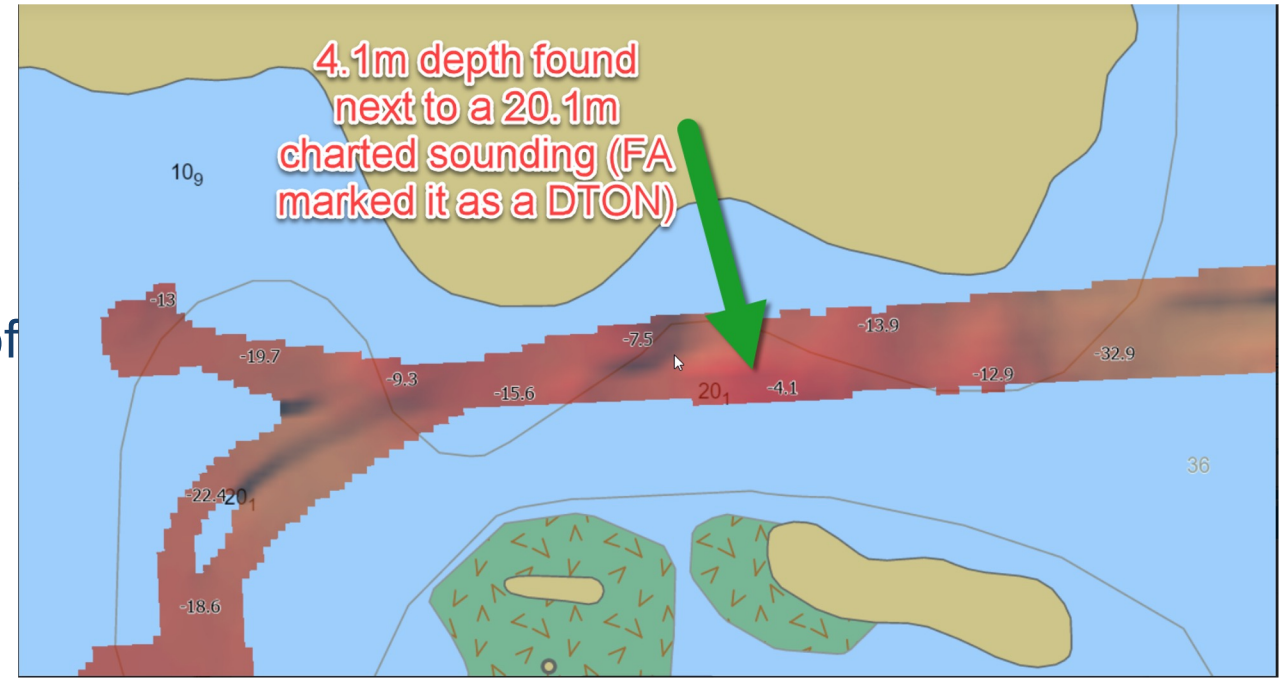
Reconnaissance - Detecting Dangers to Navigation before deploying field hydrographers

- Fairweather 2023 Dixon Entrance Project - CSB identified over half of field-submitted DTONs ahead of time



Reconnaissance - Detecting Dangers to Navigation before deploying field hydrographers

- Fairweather 2023 Dixon Entrance Project - CSB identified over half of field-submitted DTONs ahead of time



We need a larger crowd

Part 4: Questions / Next steps / Vision

The available CSB data is a drop in the bucket compared to available AIS data.

We need to support the adoption, contribution, publicization, and use of CSB data.


Dallas
MISSISSIPPI
ALABAMA
GEORGIA
Jackson
Montgomery
LOUISIANA
Baton Rouge
Tallahassee
Jacksonville
Houston
New Orleans
Orlando
Tampa
FLORIDA
Gulf of America
Miami
Straits of Florida
Thank you!
Questions/comments:
anthony.r.klemm@noaa.gov