Automated Machine Learning-based Extraction of Shallow Water Bathymetry from LiDAR Point Clouds: Developing an Operational Workflow Via Accuracy Analysis

Kim Lowell (Corresponding Author, klowell@ccom.unh.edu), Brian Calder (brc@ccom.unh.edu), Brian Miles (bmiles@ccom.unh.edu)
Centre for Coastal and Ocean Mapping (CCOM) and NOAA-UNH Joint Hydrographic Centre (JHC), University of New Hampshire, Durham, NH 03824, UNITED STATES

Abstract

Lowell and Calder (2022) have produced a two-stage automated method for identifying bathymetric soundings ("*Bathy*") in 500m-by-500m "tiles" of airborne LiDAR data that requires a minimum of human interaction. The first-stage applies the completely automated CHRT algorithm (Calder and Rice 2017) that is widely used operationally for sonar data to an individual tile. The second stage uses machine learning (ML) to cluster CHRT results for the tile and applies a rule-based decision tree to the clusters to determine the depth range of *Bathy* soundings. This depth range is used to identify individual *Bathy* soundings. Of note is that this "CHRT-ML" process is completely unsupervised meaning that it does not necessitate an initial training set with individual soundings having been identified as *Bathy* or *NotBathy*.

Global agreement with *Bathy* soundings extracted by National Atmospheric and Oceanic Administration (NOAA) procedures that require considerably more human effort was generally above 80% for 123 tiles in the Florida Keys. Moreover, the false negative rate (i.e., percent of NOAA *Bathy* soundings identified as *NotBathy* by CHRT-ML) was generally below 5%. Accuracy rates were further examined by tile "type" – e.g., tiles covered by substantial amounts of coral reefs, tiles with depths approaching the limits of light penetration – and found to vary by tile type.

To further the goal of enabling operational use of CHRT-ML, the proof-of-concept code ML portion of CHRT-ML has been software-engineered to conform to current Python programming standards. Moreover, two accuracy (relative to NOAA *Bathy/NotBathy* sounding classifications) analyses are being conducted:

1. Spatially and statisticall comparing bathymetric depth maps produced by NOAA and CHRT-ML at multiple spatial resolutions.
2. Examining accuracies relative to which of the rules in the CHRT-ML decision tree was applied to each tile to identify *Bathy* soundings.

The ultimate goal is to enable CHRT-ML to act as a first-pass classifier such as Figure 1. In addition to the CHRT-ML classification produced, accuracy estimates will be produced based on characterization of the uncertainty associated with each tile type/CHRT-ML rule combination.

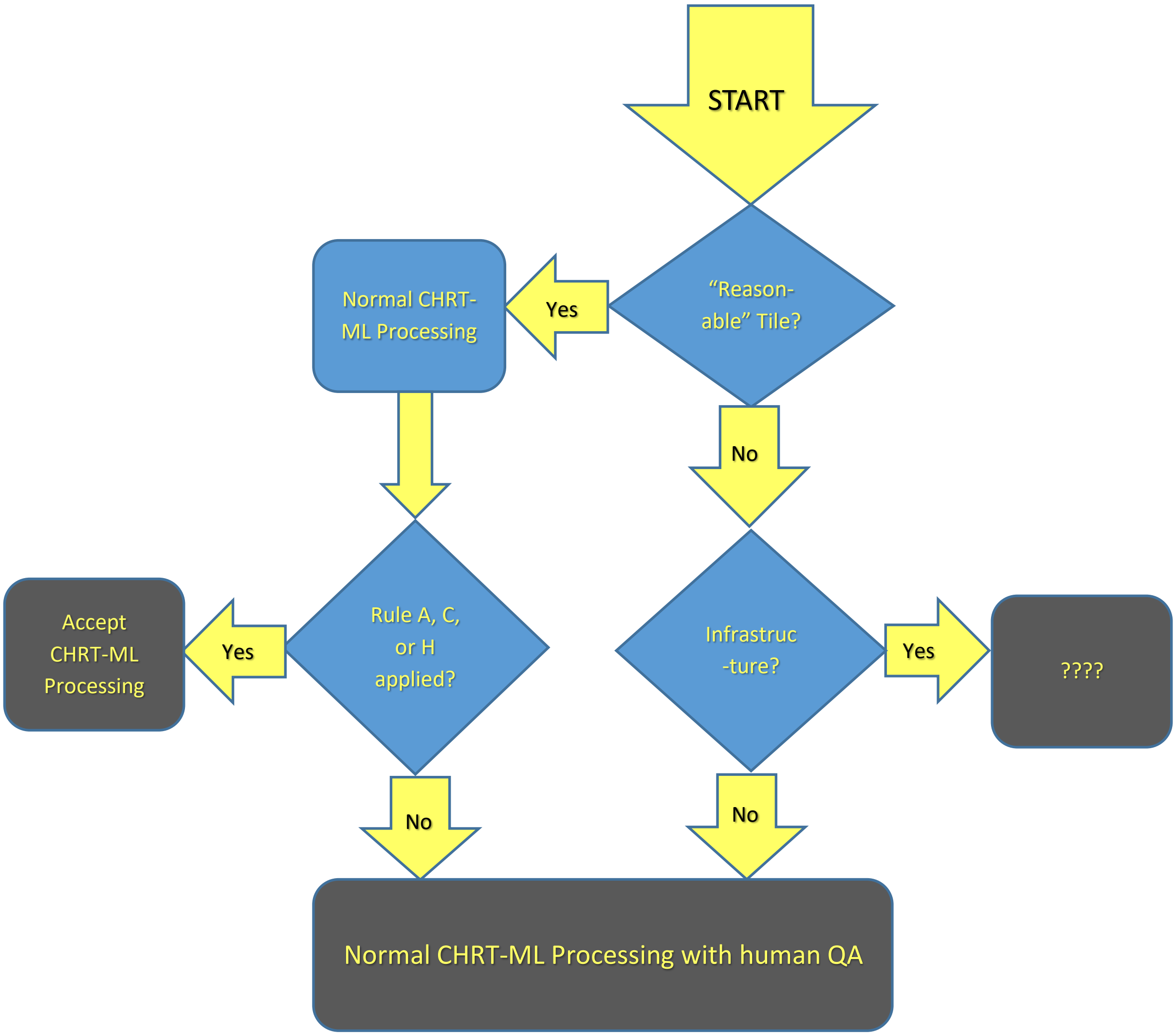

Figure 1. Hypothetical processing flow of operational use of CHRT-ML to process airborne LiDAR point clouds for shallow-water bathymetry.

Bibliography

Lowell, K., Calder, B., 2022. Operational performance of a combined density- and clustering-based approach to extract bathymetry returns from LiDAR point clouds. *International Journal of Applied Earth Observation and Geoinformation (Special Issue: Recent Advances in Geocomputation and GeoAI for Mapping)*. 107: 102699. DOI: https://doi.org/10.1016/j.jag.2022.102699.

Calder, B., Rice, G., 2017. Computationally efficient variable resolution depth estimation. *Computers & Geosciences* **106**: 49-59. DOI: dx.doi.org/10.1016/j.cageeo.2017.05.013.