Automated Processing of Arctic Crowd-Sourced Hydrographic Data While Improving Bathymetric Accuracy and Uncertainty Assessment

Khaleel Arfeen

BES, Geomatics & Computer Science, University of Waterloo, 2014

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of

Master of Science in Engineering

In the Graduate Academic Unit of Geodesy and Geomatics Engineering

Supervisor: Church Ian, Ph.D., Geodesy and Geomatics Engineering

Examining Board: Marcelo C. Santos, Ph.D., Geodesy and Geomatics Engineering
Suprio Ray, Ph.D., Computer Science

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

December 2019

© Khaleel Arfeen, 2019

Dedication

I would like to dedicate this thesis to all my friends, family and loved ones – your support matters.

Abstract

Melting sea ice has led to an increase in navigation in Canadian Arctic waters. However, these waters are sparsely surveyed and pose a risk to mariners. Recognizing this issue, the government of Canada has granted funds towards the development of a pilot program to begin collecting bathymetric data through a novel crowd-sourced approach. The project is a coalition between four Canadian partners from across the country; The University of New Brunswick's Ocean Mapping Group is tasked with the processing of the collected data and this thesis will focus on this aspect. Through an automated approach the data has been processed with the end-product being a final depth measurement with the associated uncertainty. The software is Python based and has been broken down into several modules to complete the task at hand.

Utilizing specialized hydrographic equipment, designed to be low-cost and simple to operate, participating communities in the Canadian Arctic have been given the opportunity to collect bathymetric data while traversing their local waterways. As the pilot phase of the project is done, this thesis delves into the steps taken to fulfill the processing goals. The primary motivation surrounds how the processing workflow was completed through automation while mitigating errors and achieving transparency in the uncertainty assessment in the crowd-sourced bathymetric (CSB) data.

Particular emphasis is placed upon the issues of collecting valuable hydrographic data from the Arctic with analysis of different methods to process the data with efficiency in mind.

These challenges include obtaining a reliable GNSS signal through post-processing, qualification of the GNSS data for vertical reference, utilizing the HYCOM hydrodynamic model to collect sound velocity profiles and the identification and quantification of uncertainty as part of the Total Propagated Uncertainty (TPU) model.

Several case study type examples are given where an investigation is conducted using processed collected and/or model data. Discussions surround the results of multi-constellation vs. single-constellation GNSS in the Arctic and the effects on the qualification rate for use as vertical referencing. Similarly, work towards comparing the model used to collect SVP data with equivalent real-world data collected by the Canadian Coast Guard is discussed. Finally, uncertainty has been quantified and assessed for the collected data and the results of the uncertainty assessment are provided using CHS/IHO survey standards as a benchmark.

Acknowledgements

I would like to express my gratitude to all those that assisted with the completion of this Thesis. Firstly, I would like to thank my supervisor, Dr. Church for his guidance, support and providing me with the opportunity to study Ocean Mapping – this really changed my life.

Secondly, to all those that have lent their support and feedback throughout the process. It is not always easy to understand all the concepts and theory and I often needed to validate my understanding and bounce ideas around. My friends in the Ocean Mapping Group, CIDCO and fellow GGE faculty members, patiently and generously lent me their time to do so – I am appreciative of this. A special thanks to Anthony, Thalia, Heather, Ahmadreza, Mark and Andrea.

Thirdly, those that provided me with the data needed to make this project a reality, without your contributions this project would not have been a reality. Thank you to Elaine Tessier & Jean Allard of Natural Resources Canada, Julien Desrochers from CIDCO and to the Amundsen Science team.

Lastly, I would like to express my gratitude to the indigenous communities that made this project come to life. Without your participation in this project, nothing would have happened.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

ARAIM	Advanced Receiver Autonomous Integrity Monitoring
CCGS	Canadian Coast Guard Ship
CD	Chart Datum
CHS	Canadian Hydrographic Service
CIDCO	Centre Interdisciplinaire pour le Développement de la Catographie des Océans
CSB	Crowd-Sourced Bathymetry
CTD	Conductivity, Temperature, Pressure/Depth
DFO	Department of Fisheries and Oceans
DGNSS	Differential GNSS
ENS	Electronic Navigation Systems
ERS	Ellipsoidal Referenced Survey
GDOP	Geometrical Dilution of Precision
GLONASS	Global Navigation Satellite System
GNSS	Global Navigation Satellite System
GOFS	Global Ocean Forecasting System
GPS	Global Positioning System
HyVSEPs	Hydrographic Vertical Separation Surfaces
IGS	International GNSS Service
IHO	International Hydrographic Organization
IOC	Intergovernmental Oceanographic Commission
MBES	Multi-Beam Echosounder
MSL	Mean Sea Level
MSL	Mean Sea Level
NAD83	North American Datum of 1983
NOAA	National Oceanic and Atmospheric Administration
NRCAN	Natural Resources Canada
NSV	Number of Satellite Vehicles
PPK	Post Processed Kinematic
PPP	Precise Point Positioning
PPP	Precise Point Positioning
RTK	Real Time Kinematic
RTOFS	Real Time Operational Global Ocean Forecast System
SBAS	Satellite Based Augmentation Systems
SBES	Single-Beam Echosounder
SDHGT	Standard Deviation Height
SDLAT	Standard Deviation Latitude
SDLON	Standard Deviation Longitude
SSH	Sea Surface Height
SVP	Sound-Velocity Profile
TEOS-10	Thermodynamic Equation of Seawater 2010
THREDDS	Thematic Real-time Environmental Distributed Data Services
THU	Total Horizontal Uncertainty

TPU   Total Propagated Uncertainty
TVU   Total Vertical Uncertainty
TWTT   Two-Way Travel Time
WOA   World Ocean Atlas (01,18)

1 Introduction

1.1 Background

The Canadian Arctic is vast covering more than 4.4 million km^2, stretching across Northern Canada from 60° northern latitude to the Beaufort Sea and Baffin Bay. With 47% being underwater and over 36,000 islands, the area is characterized by an abundance of inlets, channels, fjords and bays (Government of Canada, 2018). The Canadian Hydrographic Service (CHS) is the governmental body which is federally mandated to produce and deliver navigational charts, publication and services for the immense extent of the navigable Canadian waters.

Climate change can be blamed for causing the melting of Arctic sea ice. As the Arctic ocean frees itself of this ice, some parties will find new opportunities towards the navigation in these waters. However, vast areas of the Arctic remain unchartered to this day and this presents a liability towards navigational safety. In 2014 the Office of the Auditor General of Canada released a report indicating that the CHS estimates that about 1% of the Canadian Arctic has been surveyed to modern standards (Office of the Auditor General of Canada, 2014), see Figure 1. Traditional survey methods would be far too costly and timely to undertake for a project of this magnitude. That being said, the Government of Canada has granted funding towards an initiative to obtain a greater understanding of the Arctic waterways using a nontraditional approach.
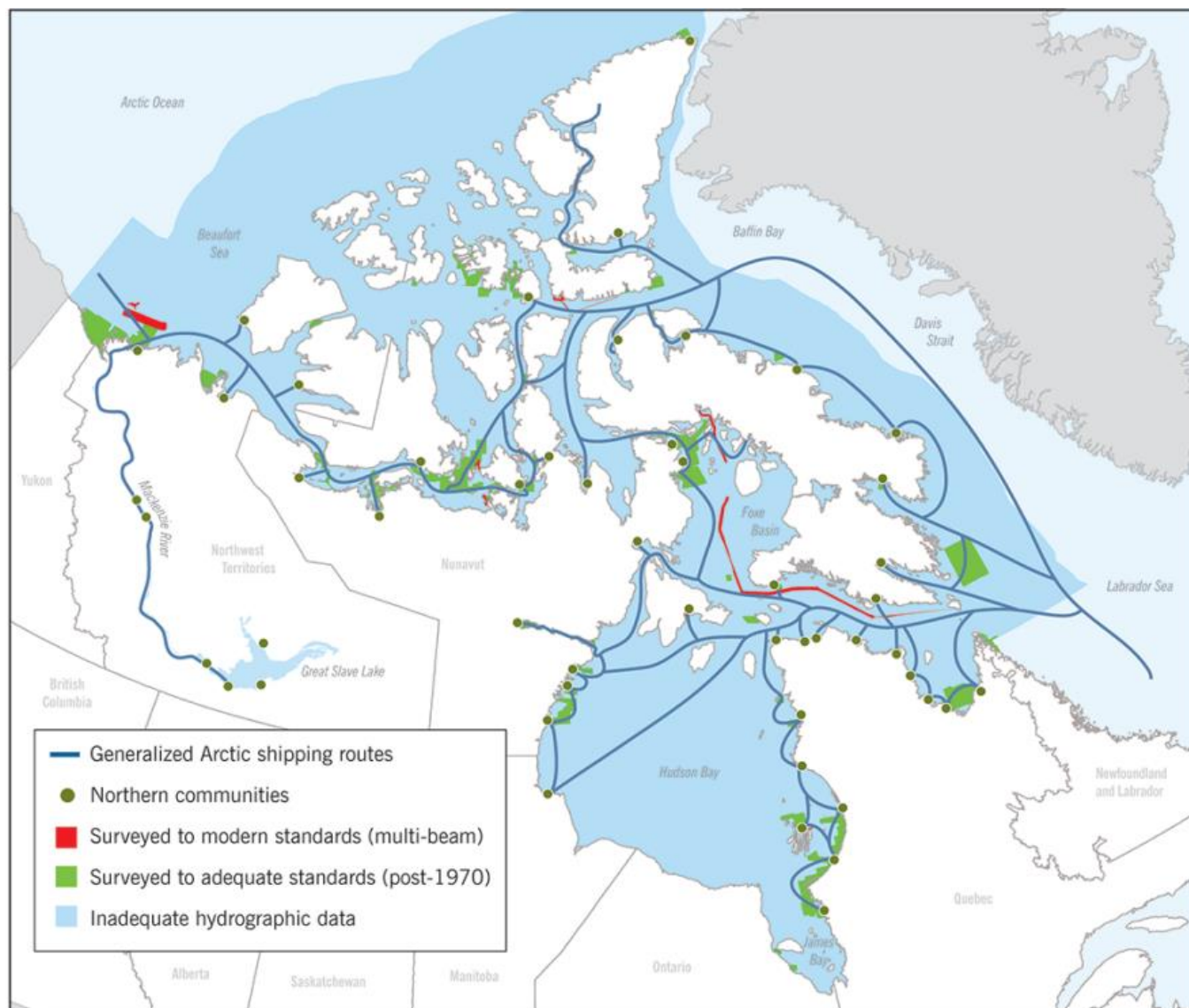
Figure 1. Many Arctic shipping routes have inadequate surveys and charts (src: oag-bvg.gc.ca)

1.2 CSB Initiative

This initiative is a Crowd-Sourced approach where participating local communities have been

empowered with the choice of collecting crowd-sourced bathymetric (CSB) data while traversing

their local waterways. Along with a coalition of diversified hydrographic partners from the

University of New Brunswick, York University, Memorial University and Centre

Interdisciplinaire pour le Développement de la Catographie des Océans (CIDCO), local community participation was required in order to make this pilot crowd-sourced project a reality. Participants from across the Canadian Arctic, including Iqaluit, Nu; Gjoa Haven, Nu; and Quaqtaq, QC; as shown in Figure 2, were given specialized hardware which could be towed from a watercraft while simultaneously collecting GNSS, motion and sonar data.

1.	Gjoa Haven, Nunavut
2.	Iqaluit, Nunavut
3.	Quaqtaq, Quebec

Figure 2. Participating Communities in the CSB Project

The International Hydrographic Organization (IHO) has long been an advocate for the collection of crowd-sourced bathymetric data, stating "[CSB] helps improve mankind's understanding of the shape and depth of the seafloor" (International Hydrographic Organization, 2018). Despite the multitude of collected bathymetric data, less than 15% of the world's oceans have been surveyed.

Mapping of coastal waters is marginally better with about 50% of the world's coastal waters of less than 200m being surveyed (International Hydrographic Organization, 2018).

With local hydrographic offices cognizant of their limited resources and looking for solutions, many are aware that they cannot rely solely on their own capabilities. Instead hydrographic offices must engage the stakeholders and the public to build the capacity needed to fulfill their missions (Cooper & Hersey, 2014). The IHO has listed crowd-sourced bathymetry as a priority solution to mapping the world's waters (Reed, 2017), with CSB said to play a role in helping to map the world's oceans.

While there are certainly many benefits associated with CSB, there are some limitations. There are no current standards for CSB provided by the CHS which leads to potential errors and uncertainty in the data collected. For example, the equipment used, sensor offsets and calibrations and water conditions at acquisition are unknown which contribute to unknown data quality. In all cases, CSB data is of lesser quality when compared to a professional survey. The Ocean Conservancy has presented that the challenge with CSB is "to ensure the reliability of crowdsourced data by managing and structuring the process to ensure that it is reliable, useable and accurate" (Ocean Conservancy, 2013). Taking this into account, CSB still presents many unmatched benefits towards hydrographic surveying. In addition to reducing time and cost expenditures and being potentially quicker to detect hazards, CSB data helps to supplement gaps where bathymetric data is scarce such as in difficult to access and unexplored areas, such as the Arctic (Smith S. , 2018). Regardless of the situation, some data is almost always better than no data and ultimately helps to increase navigational safety in uncharted waters.

Without the help of the local communities this project would not be possible. With the appropriate training these communities were able collect bathymetric data in areas where there has traditionally been absolutely no hydrographic data. Their help leads to better understanding of Canada's Arctic waters through exploration and monitoring of these waterways.

1.3 Hardware

A simple and low-cost device was created by the not-for-profit partner CIDCO in order to fulfil the surveying needs. This device, called the HydroBall is a simple Single-Beam Echosounder (SBES) which features a GNSS receiver, Inclinometer and Processor to obtain bathymetric data. The HydroBall itself measures 40cm in diameter, 28lbs and features a 24hr battery operation. The HydroBall is shown in Figure 3.

This device can easily be towed by a fishing boat or canoe. Floating on its own or stabilized with a catamaran mount, the device stays flush to the water surface. To measure tilt a Honeywell HMR3400 inclinometer is used. For location tracking a GNSS Hemisphere Eclipse P320 receiver was chosen with L1/L2 and GLONASS capabilities. The sonar system used is an Imagenex 852 single-beam echosounder. This SBES operates as a frequency of 675 kHz, has a beam width of 12°, a range up to 100m and a 10Hz ping rate.

Figure 3. The HydroBall with Catamaran Mount

1.4 Workflow & Automation

1.4.1 Workflow Overview:

Once data is collected by the project participant(s), the raw files consisting of GNSS, IMU and sonar data are uploaded via a web form and sent to a receiving server where it is processed and visualized on an open-access web platform. The GNSS data records are stored as a .bin format file using recordings from the Hemisphere GNSS sensor. The IMU data records are stored in a proprietary .dev file which also includes other parameters such as equipment status, NMEA etc. The IMU data is found under the heading $PTNTHPR, giving heading, pitch and roll. The sonar data is recorded as the proprietary .852 Imagenex format.

Apart from acquisition, to make this project feasible a system needed to be in place to efficiently process the data through automation while assessing uncertainty and improving accuracy. Labour costs, human-error and time constraints are mitigated through automation. Additionally, automation allows for quicker times to create the final corrected data and to distribute this data to

the associated stakeholders. This grants the greater ease to scale this project, allowing for a larger amount of communities to join the initiative and contribute in the future.

To process the data efficiently required a comprehensive workflow which spans from acquisition to dissemination. To make the processing more robust, the specific steps of the process have been modularized to do one thing and to do it well. These modules are visualized in the technical flowchart in Figure 4. In the technical flowchart we can see the processing steps end-to-end. After the data is uploaded on the Raw Server, the processing of the raw data commences. Firstly, the GNSS data undergoes post-processing via the Precise Point Positioning (PPP) method to improve the raw GNSS solution, explained in GNSS Processing. Afterward, each processed GNSS record is qualified and sorted based on a threshold criteria, explained in GNSS Qualification.
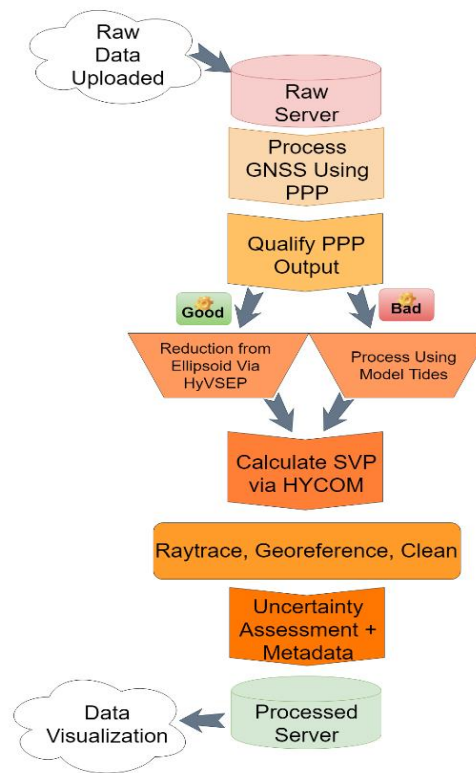
Figure 4. Technical Flowchart

From here the data is deemed either to have an acceptable or non-acceptable GNSS quality. If acceptable, the processed GNSS record can be used for the vertical reference and is either kept at the ellipsoidal reference or reduced to Mean Sea Level (MSL) or Chart Datum as described in Vertical Reduction. If non-acceptable, the processed GNSS record cannot be used for the vertical reference and thus a tidal reference must be used. During phase 1 of this project, all data is to be kept at the ellipsoidal reference. This is done to ensure transparency as tidal and reduction methods can vary depending on the method chosen. This allows users of the data to do the necessary conversions themselves, creating more transparency in the data stream.

Subsequently, the salinity and temperature vs depth array data must be given to calculate the sound speed in water. Traditionally, Conductivity, Temperature and Pressure/Depth (CTD) casts are used in-situ to gather this data. However, to simplify participant tasks this is not done in the field and instead this data is queried from the Hybrid Coordinate Ocean Model (HYCOM) freely-available oceanographic model. When the temperature and salinity vs depth data has been queried from the model, it is then used as the input to calculate the sound speed in water, further explained in SVP.

Once the sound speed is known, raytracing is done to determine the depth of the water from the SBES. Additionally, georeferencing and data cleaning are done in this step to obtain a final depth reading. At this point, the uncertainty is calculated, explained in Uncertainty, and all processing steps and necessary information are included in the metadata, explained in Metadata. The processing is then marked as complete and data is sent to a final server where it will be shared and visualized.

This thesis encompasses the work done towards the data processing, specifically towards the tasks of the GNSS PPP processing, GNSS qualification, salinity and temperature model query, sound speed calculation, metadata creation and uncertainty assessment. The tasks of reduction and tidal reference have also be conducted, though these will be implemented at a future time as the project advances. The other tasks of raytracing, georeferencing, cleaning and visualization are to be handled by the other project partners. Only the modules undertaken by the University of New Brunswick will be discussed in this thesis, with the following paragraphs giving a high-level summary of each module providing a primer before a more in-depth analysis in the Methodology.

1.4.2   GNSS Role:

The location information provided by the GNSS receiver plays a crucial role in positioning the sonar data onto an ellipsoidal referenced coordinate system. Unfortunately, obtaining a strong GNSS signal at high latitudes is more challenging than at mid-latitude regions. However, several methods exist to improve upon the accuracy of the GNSS reading, including PPP. After the GNSS data has been processed using the PPP method, qualification of the readings on a per-line basis is done to sort the lines on quality. If the quality meets threshold requirements, it may then be used for vertical referencing. The threshold values used are modeled partially after CHS/IHO level 1a and 1b survey requirements for the static portions of the horizontal and vertical components.

1.4.3   SVP Role:

When conducting a hydrographic survey, depth is never directly measured it is only inferred. We directly measure the time it takes for the sonar signal to propagate to the seafloor and back, known as the Two-Way Travel Time (TWTT). Combining the ½(TWTT) with sound speed, we are able to calculate depth (Snellen, 2008). The speed of sound calculation is derived from temperature,

salinity and pressure/depth variables and is given as a function of depth (Beaudoin, Transit Sounding Reclamation, 2009). This array of sound velocity vs depth data is known as the Sound Velocity Profile (SVP).

The SVP plays a crucial role in obtaining an accurate bottom depth calculation and it is important to have the input variables represented accurately. The objective of the SVP module is to utilize an appropriate oceanographic model which can provide the inputs to the speed of sound calculation for the Arctic survey region. Once the model has been selected, the scripts must then query the hydrodynamic model for the oceanographic data per input line while considering date, time and location. After the inputs to the speed of sound calculation have been retrieved, the conversion to sound velocity is made using a sound velocity algorithm followed by creation of the final formatted SVP.

1.4.4 Uncertainty & Metadata Role:

It is necessary to communicate any errors and uncertainties associated with the data to the users. This ensures communication is made that even though the data has value, there still does exist some limitations influencing how much faith one puts into the data and into the derived uses. The IHO Standards for Hydrographic Surveys, S-44 5th Edition (International Hydrographic Organization, 2008) set guidelines for the definition of *Error* and *Uncertainty*.

Error is defined as "The difference between an observed or computed value of a quantity and the true value of that quantity" (International Hydrographic Organization, 2018). The true error can never be known as it is not possible to know the true value with full certainty.

Uncertainty is defined as the upper and lower boundaries surrounding a value that will contain the true value at a specific confidence level (International Hydrographic Organization, 2008). Typically, the confidence level is given at the 95% (2σ) using a normal statistical distribution model.

Error sources are identified and discussed, however the values associated with these errors and analysis are technically uncertainty sources and not errors (International Hydrographic Organization, 2018). The IHO S-44 document provides a comprehensive list of uncertainties to be included when used on a navigational chart. Not all the requirements for navigational charts apply to CSB data, due to the nature of acquisition, hardware utilized, less strict onus for safety of navigation etc. A visualization of the general sources of errors when conducting a hydrographic survey can be seen in Figure 5.
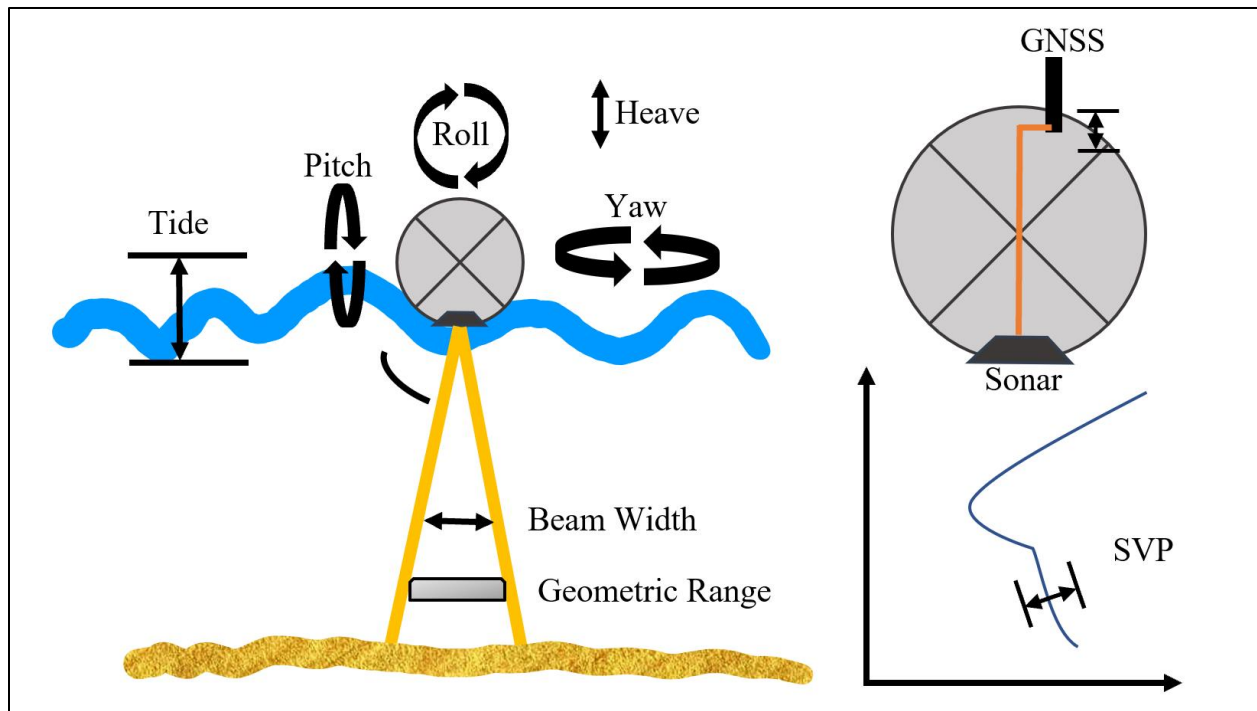
Figure 5. Error model for a surface survey (src: (Naankeu Wati & Seube, 2016)) [modified]

The uncertainties for each of the component sources of error are combined together to give a final uncertainty values, creating a sum of each component interacting in the complete survey system. The Total Propagation of Uncertainty (TPU) is the recognized standard of quantifying uncertainties by the IHO. The TPU model is well documented for bathymetric surveys by Hare et. Al (1995).

The TPU is the result of the uncertainty propagation when all contributing measurement uncertainties, both random and bias based, have been accounted for. Propagated uncertainties are provided in both the vertical and horizontal planes (International Hydrographic Organization, 2008). The Total Vertical Uncertainty (TVU) is the component of the TPU which covers the vertical plane and is a 1-dimensional quantity. The Total Horizontal Uncertainty (THU) is the

component of the TPU which covers the horizontal plane and is a 2-dimensional quantity though reported as a single radial value (International Hydrographic Organization, 2018). This uncertainty is most often reported in meters derived from the variance from the assumed statistical model distribution. Ultimately, the final depth accuracies strived for in this project will be based on CHS/IHO level 1a/1b or better TVU and THU components, described in Appendix B.

1.5 Problem Statement & Contribution

The burdensome task of surveying and mapping Arctic waters cannot be completed through traditional methods - it is just too great of an ordeal to overcome in a world of limited resources. Instead, non-traditional methods, such as CSB, will help to pave the way in creating a scalable method of surveying these waters. Pilot projects such as these are necessary to begin to implement the theory and methods which will be used and then to improve on these as the project scales.

It is the hope of all parties involved with this pilot that their contribution put forth will help to shape the future of CSB in Canada. CSB will be a necessary tool towards the efforts of mapping Canada's Arctic Waters, along with helping to contribute towards mapping the world's waters. This thesis will help to contribute towards the body of knowledge surrounding CSB as well as providing a case-study on the feasibility of crowd-sourced mapping initiatives specifically in the Northern Arctic regions.

1.6 Research Questions

Given the nature of the project, the question remains on how to *efficiently* process crowd-sourced hydrographic data through *automation* while improving bathymetric *accuracy* and assessment of *uncertainty*.

Given the somewhat vague nature of some terms used in this research question, the terms in italics can be defined to give the reader a better understanding:

Efficiency: In terms of human labour, processing time and the associated costs of processing.

Accuracy: Selecting the correct methods and tools that improve the raw data and minimize human and systematic (environmental, equipment, process) errors to ultimately minimize final depth error.

Automation: Minimizing human labour and time, highly related with efficiency.

Uncertainty: To estimate the errors and identify the limitations of the methods and tools and transparently convey this to the end user.

1.7 Research Objectives

Using a variety of existing tools and systems, we can incorporate them together to form a cohesive automated environment which minimizes human labour, processing time and the associated costs while obtaining a higher degree of accuracy in the data.

Given this premise, the following were the research objectives encountered during this project surrounding the processing of the data:

- Select an appropriate means to post process raw GNSS data to obtain higher accuracy

- Establish a filter on GNSS readings for quality in order sort whether an ellipsoidal or tidal vertical reference should be used

- Give guidance towards finding tidal model(s) appropriate for Arctic data

- Give guidance towards reducing ellipsoidal referenced data to a common hydrographic datum using HyVSEPs

- Select an appropriate model to provide SVPs for the corresponding Arctic locations

- Select an appropriate algorithm for calculating the speed of sound in water

- Identify and quantify sources of uncertainty on the final depth measurement

- Document uncertainty, processing methods and limitations in the metadata

- Complete data processing in an automated workflow

2 Literature Review

2.1 Hydrographic Survey Standards

Given the serious nature of providing navigational charts and the associated risks to mariners, understandably the hydrographic community typically is not the most receptive to drastic change. The proposed project challenges many of the current status-quos surrounding hydrographic data as we move towards an automated workflow; This is in opposition to a more rigorous process with many checks and balances as traditionally seen. The checks and balances involve manually going through each step of the processing, the collection and use of in-situ sound casts, installation or use of tidal gauges and manual cleaning of the data etc. The primary bodies regulating bathymetric data in North America are the Canadian Hydrographic Service (CHS, Canada) and the National Oceanic and Atmospheric Administration (NOAA, United States). The International Hydrographic Organization (IHO) is an international body of the same mandate. The organizations produce their own documents regulating surveying standards for their respective domain(s). These documents help establish methods and best practices for hydrographic surveying and can be accessed online. In regards to crowd-sourced bathymetric data, the IHO has come up with specific standards and guidance outlined in a 2018 document titled Guidance on Crowd-sourced Bathymetry (International Hydrographic Organization, 2018).

2.2 Crowd-Sourced Bathymetry (CSB)

Following with the theme of CSB surveys, Reed et al (2017) has described the importance as well as provided updates in terms of regulations around CSB. This thesis outlines the results of a CSB project using Electronic Navigation Systems (ENS) which is a pre-installed system on many ships. The thesis outlined methods on recording and managing metadata as well as discussed difficulties

and best practices of CSB project. The difficulties surround costs of running these programs such as equipment and labour as well as barriers to entry such as agreements by the crowd. The best practices outlined apply specifically towards NOAA projects and state that the data must be un-corrected for tide to preserve the raw data quality and that offsets and time synchronization must be shared. The program was largely considered a success. However, it was met with significant limitations which included no post-processing methods discussed and recognition that uncertainty assessment is extremely difficult. Also, the project in this thesis receives a high quantity of data providing a greater ability to validate data through statistical redundancy, something that is not possible for the project discussed in this thesis.

In a separate study conducted by Novaczek et al (2019), the authors used crowd-sourced Olex data to create regional bathymetry and geomorphology maps of the Newfoundland and Labrador shelves of Eastern Canada. Olex provides a commercial charting system with use of their equipment which ranges from survey-grade multi-beam to single-beam fish finders. Olex users are encouraged to share their bathymetric data and in exchange these users will get access to a database with bathymetric data from other Olex participants across the globe. The authors state that to date, Olex has compiled over 8.6 billion depth measurements from approximately 10,000 vessels globally which makes it the largest crowd-sourced bathymetric data initiative to date. The authors used Bayesian Kriging to generate a continuous bathymetric surface from incomplete and sparse Olex coverage of the study region. The final product resulted in a 75m bathymetric grid, completed at a fraction of the cost when compared to multi-beam data collected over the same area. The authors claimed the final product provided over 100x finer spatial resolution than what was previously available from General Bathymetric Chart of the Oceans (GEBCO) for majority of the 672,900 km^2 area. The interpolated bathymetry was tested against independent data provided by

Fisheries and Oceans Canada for accuracy and was found to have a Spearman correlation of 0.99. The significance of this study is that it shows how crowd-sourced data can be used to increase bathymetric coverage in areas where there is limited coverage, all while retaining accuracy.

2.3 Challenges Surveying in the Arctic: GNSS

Regardless of the method used to acquire hydrographic data, particular challenges occur when surveying in the Arctic. The primary challenges common across literature are obtaining valuable GNSS data and issues with obtaining accurate tidal heights and sound velocity profiles. Two great starting points towards understanding GNSS positioning and issues in the Arctic are Guide to GPS Positioning (Wells, et al., 1999) and GNSS Use in the High Arctic: Issues and Solutions from a Canadian Perspective (Langley, GNSS Use in the High Arctic: Issues and Solutions from a Canadian Perspective, 2009). The major issues of using GNSS in the Arctic compromises primarily of ionosphere and geometry errors. Jensen et al (2010) discussed these issues with intention to raise awareness and foster thoughts around the issues. The authors concluded that geometry issues can only be corrected by improving navigation system coverage in this region. Further conclusions stated that using multiple frequencies basically eliminated ionosphere delays, although high order effects still present an issue. Jayachandran et al (2009) reiterated the issues with ionosphere errors associated with GNSS in the Arctic and taking it further they described the concept of scintillation which has large temporal and spatial variance across the Arctic. The scientific research objectives of this paper were to assess the drivers and variability of polar cap convection and the generation and dynamics of ionization structures for which they concluded that more research needed to be completed. Finally, Reid et al (2015) explains GNSS integrity in the Arctic with particular emphasis on using Advanced Receiver Autonomous Integrity Monitoring

(ARAIM) and Satellite Based Augmentation Systems (SBAS) as potential aid to navigation and mapping in the Arctic. Reid explains that this is just currently not feasible because of limitations in the GNSS constellations covering the Arctic. However, planned expansion is expected in the next decade and opens the possibility to further research including uses in vertical referencing for bathymetric surveys.

2.4 Challenges Surveying In The Arctic: Tidal Heights, Sound Velocity Profile (SVP)

Information on tidal referencing and obtaining SVPs are given by Beaudoin et al (2006). The authors were interested in raytracing the water column from Arctic survey data from the CCGS Amundsen. As no CTD casts were available, a proposal to use climatology models to collect SVPs were explored. The authors chose to collect this data from World Ocean Atlas (WOA01) and concluded that it is acceptable to use this data in absence of CTD casts as final sounding accuracy was not significantly affected, but cautiously warned not all model domains were equal. The best performing region was the Western Arctic domain and the worst performance came from the Hudson Bay domain. The authors finally suggested improvements to the WOA01 Arctic coverage with a focus on problematic areas. Since then, many of the proposed changes have been implemented and are available in a recently released WOA18. Taking his work further Beaudoin (2009) discussed using Canadian Coast Guard vessels as a source for bathymetric data. However, to use this data the author needed to overcome what he identified as two main issues, varying water level heights in space and time and variations of sound speed in the water column. In order to mitigate the former, the author used the WebTide service to predict tidal height which is valid across the Canadian Arctic. The author discussed the process of working with the system and stated

that it was the best source of vertical control in the Arctic, although he did not discuss GNSS controls. For sound speed, he built upon the research from (2006) using WOA01for SVPs. The conclusion was that the model provided reasonable results in relatively stable waters and when compared to CTD casts, achieved at most a few decimeters in difference for single-beam data.

It is always best practice to collect temperature and salinity values through in-situ casts while conducting a survey, though this is not always possible for varying reasons. In the case of this project, having the participants collect this data would add undue complexity as the participants are not hydrographers and would add considerable expense towards equipment and training. As a result, during processing of the data from this project, no observed sound velocity information will be obtained in-situ and data from hydrodynamic models be used in lieu. When presented with a similar problem, Beaudoin et al (2009) has shown that the use of the World Open Atlas (WOA) model could be used appropriately to supplement the absence of observed SVPs without seriously affecting sounding accuracy in the Arctic. Church et al (2012) has shown that model output could be used as a substitute for actual observations with little effect on ray-tracing uncertainty. Furthermore, the authors stated that a model can be used to learn about the temporal distribution of salinity and temperature in comparison to a single point as when taking a traditional cast. This aids in survey planning and the use of a model would improve survey efficiency and uncertainty. To further this thought, Calder et al (2004), has shown that using models could bring down operational costs of conducting surveys by avoiding taking casts. In their study, these authors have shown that a model-based SVP can be used within the budget error of International Hydrographic Organization S-44 standards. However, the authors did specifically mention that more studies are necessary when working in shallow environments (0-30m) to determine if a model is suitable.

Once a model has been chosen with appropriate spatial and temporal coverage, the temperature, salinity and depth variables are able to be retrieved. At this point a conversion of these inputs to a final sound velocity in water is needed. There exist multiple equations to perform this conversion, some more specialized to certain applications than others. These are available in literature from authors such as   (Chen & Millero, 1980), (Grosso, 1974), (Mackenzie, 1981), (Wilson, 1960) and (Medwin, 1975). Pike and Beiboer (1993) have made a comparative study of several different algorithms along with recommendations. For water with depths less than 1000m the Chen and Millero equation is recommended by these authors. While the Chen and Millero equation has been adopted for many years as standard, in 2009 the Intergovernmental Oceanographic Commission (IOC) adopted the Thermodynamic Equation of SeaWater 2010 (TEOS-10) equation. This equation now stands as the official descriptor of seawater and ice properties in marine science. Significant changes introduced in this formula are the use of Absolute Salinity along with introduction of units to Ocean Salinity (g/kg) (Intergovernmental Oceanographic Commission, 2010). It is commonly accepted that the greater the accuracy in determining the sound-velocity, the greater that accuracy carries forward into the final depth measurement (Alkan, Kalkan, & Turkiye, 2006) and the appropriate selection of the correct equation is important.

In order to assess the validity of the SVP, Beaudoin et al (2006) has laid out some methods on the comparison between observed SVPs and model derived SVPs. In this case it is necessary to compare the model-derived data to a 'true' dataset. The two profiles can be directly compared by plotting both into one depth-velocity graph. However, it can be taken further by completing the raytracing using both profiles independently to calculate depth. When plotting these results, the real-world implication of the associated error can be made when comparing to a known bottom depth.

2.5 Challenges Surveying in The Arctic: Combining the Issues

Combining GNSS and Tidal limitations, Church et al (2009) discussed issues relating to obtaining a reliable GNSS signal in a fjord setting in Oliver Sound aboard the Heron. With lack of tidal information for the body of water and issues obtaining a reliable GNSS signal, tidal models were compared to GNSS augmented with PPP and PPK. The conclusions were that PPK is likely best option for vertical positioning if within base station limits, otherwise PPP outperforms in long range surveys. Tidal models provide the best solution when working in areas where there is no reliable GNSS signal.

2.6 Errors and Uncertainty Assessment

An excellent resource for learning about uncertainty in science is written by Fischhoff et al (2014). In their article 'Communicating Scientific Uncertainty', the authors state that when communicating uncertainty scientists are required to identify relevant facts, characterize relevant uncertainties, access their magnitude, draft possible messages and evaluate the success. The article presents a great starting point to learning about uncertainty and its effects on research.

Narrowing down to a more hydrographical focused view on uncertainty, the IHO Standards for Hydrographic Surveys, S-44 5th Edition (International Hydrographic Organization, 2008) document provides necessary standards surrounding the acquisition and assessment of bathymetric and LIDAR data. Specifically, this documentation set these standards towards data which is collected and implemented in navigational charts used for safely navigating waters and marine habitat protection. The standards include the well-recognized IHO S-44 Minimum Standards for

Hydrographic Surveys, which outline the minimum requirements to meet different survey orders (Special, 1a, 1b, 2). The calculations towards Total Horizontal Uncertainty (THU) and Total Vertical Uncertainty (TVU) are also given. Additionally, necessary definitions around Uncertainty, Error, Confidence Levels, Accuracy and other relevant terms are given.

Bongiovanni et al. (2016) presented data from the USGS survey of Buzzard Bay, Ma. The sonar used was a noisy legacy system which did not meet the hydrographic standards set by NOAA. The article discusses the process for assessing the data for hydrographic purposes and a proposal of methods for assessing uncertainty to allow the data to be used for chart purposes. Ultimately, the authors found the data did not meet IHO Total Vertical Uncertainty (TVU) for Order 1a. Primary cause of failing to reach this standard is due to a roll bias found in this particular dataset which was not able to be corrected. The conclusion is that the uncertainty calculation remains an estimation at best and is very difficult to estimate. The recommended mitigation is to conduct crossline analysis at time of survey to help reduce systematic errors and also through post-processing using aggressive filters.

Finally, Hare (1995) outlines a system of measuring bathymetric uncertainty using error equations derived from using the method of propagation of errors. Hare presents a thorough analysis for defining and measuring error budgets and uncertainties. He discusses the errors given from the sonar system and auxiliary sensors in order to compute the depth and position uncertainties for each sounding. This method aids in helping to provide an indication of where errors are highest and can be used to implement improvements to increase depth and position accuracy.

2.7   Data Processing Automation

A paper completed by Tantra (2010) describes the author's experience building a Python automation framework, specifically focusing on documenting the experience. The ultimate goal of the said framework was to reduce the verification time and potential defects through the automation of test cases for their application. As stated by the author, Python was chosen for this project for the following reasons:

"No compilation is necessary in Python and thus ensures that the latest script is run when combined with a source code repository system, such as SVN. Additionally, management of many small compiled executables and the respective code is very a very difficult task, which the use of Python eliminates."

Python is easy to read and flexible, meaning many people can start writing scripts by easily learning from examples and existing scripts. The author found that once momentum for test scripts had started, additional script were easily created and started accumulating quickly. Additionally, Python does not limit users to a specific programming paradigm such as object oriented and thus also speeding up development time, especially by new programmers.

Python is mature and has a significant amount of support behind it. With the author's project, development time was limited and the final product needed to be delivered quickly. Python has many libraries available for use right away which helps to speed up development time without having to 'reinvent the wheel'.

3 Methodology

3.1 Introduction

As previously stated, as part of the crowd-sourced initiative to collect bathymetric data, surveys have been conducted by non-specialized operators. The logistics of finding interested communities, the training of operators and meeting the challenges of providing these operators with simple to operate, inexpensive, standalone system has been completed by the project partners (Leighton, 2019). After data acquisition, the operators then upload the collected data to a raw data server. Metadata will accompany the raw data with majority of the metadata being auto-generated with minimal user-input.

Upon arrival of the raw data onto the server, initial file conversions and implied metadata fields are calculated with the help of automation scripts polling the server for incoming data. The same automated scripts then mark the data as ready for processing and then pushes the data to a processing server where it will follow a systematic approach to obtain final depth soundings along with uncertainty, see Table 1. Upon completion of processing, the data is flagged as complete. From here the data is then pushed to a final server with help of an automation script, where it will undergo a quality check and ultimately the final bathymetric data will be disseminated for visualization.

This thesis will focus on the contribution of UNB which is to complete steps regarding the automated data processing. This is to be completed through a systematic approach using new and

existing tools brought together in a cohesive environment. The steps to be completed are shown in the following table.

	Task
1.	Process GNSS data via Precise Point Positioning (PPP)
2.	Qualify the GNSS PPP output via a threshold filter
3.	Reduction of the vertical datum from Ellipsoid to Mean Sea Level (GNSS Only)
4.	Obtain tidal heights via a hydrographical tidal model for rejected GNSS readings
5.	Collect Sound Velocity Profiles (SVP) for each reading
6.	Document metadata and provide a transparent explanation of uncertainties

Each of the tasks shown in Table 1 will be explained in detail in the following sections.

3.2   GNSS Processing

The HydroBall is able to capture both L1 and L2 GPS frequencies as well as GLONASS signals. However, there are significant challenges surrounding obtaining accurate readings in a predominately Arctic survey setting when compared to readings obtained at mid-latitudes. The main challenges in resolving an accurate GNSS reading occur primarily from poor satellite-receiver geometry and ionospheric effects (Jensen & Sicard, 2010).

Due to low inclination angles of GPS and GLONASS constellations, the Satellite geometry is never favorable in the Arctic and satellites will reside close to the horizon. The GPS constellation has an inclination rate of 55^0 relative to the equator (European Space Agency, 2018) and the GLONASS constellation has a slightly more favourable inclination of 64.8^0 relative to the equator (European Space Agency, 2018), see Figure 6 and Figure 7. The factor known as the Geometrical Dilution of Precision (GDOP) is a term used to capture the GNSS geometry in 3D position and time. A higher GDOP reading indicates a poorer satellite geometry and in turn negatively influences the ultimate positioning accuracy (Wells, et al., 1999), see Figure 8. The GDOP acts as a multiplier towards measurement accuracy, meaning we want to minimize these effects as much as possible to obtain the most accurate position (Wells, et al., 1999).
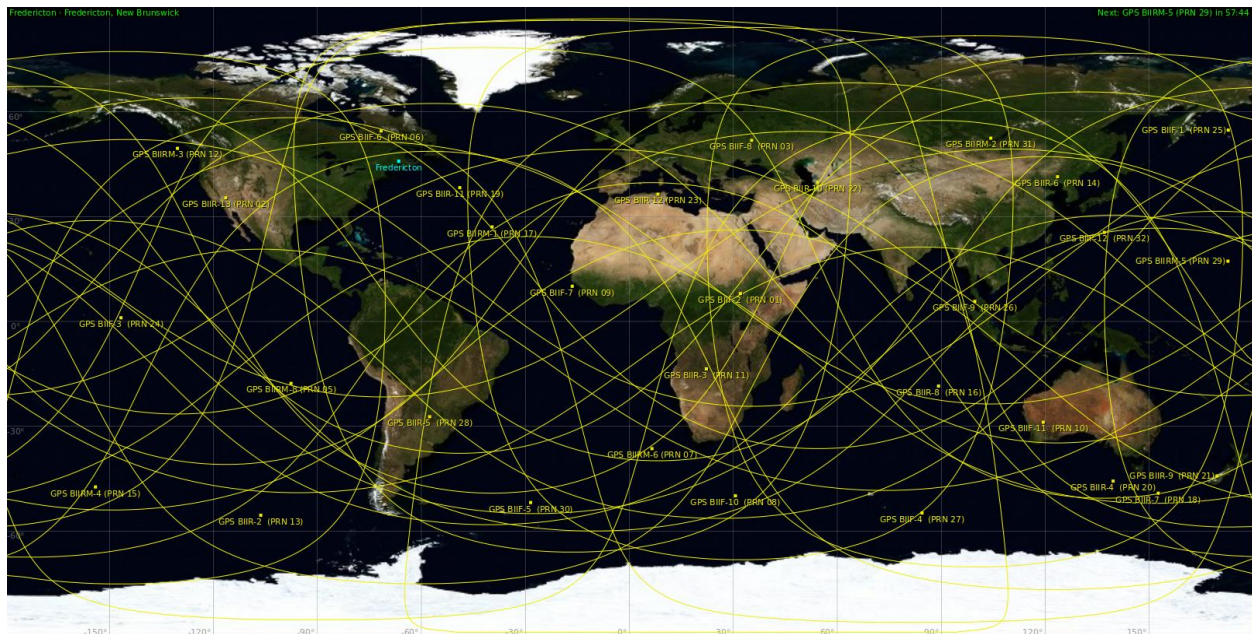

Figure 6. GPS satellite Orbit (55 degree inclination) (src: courtesy of Gpredict)
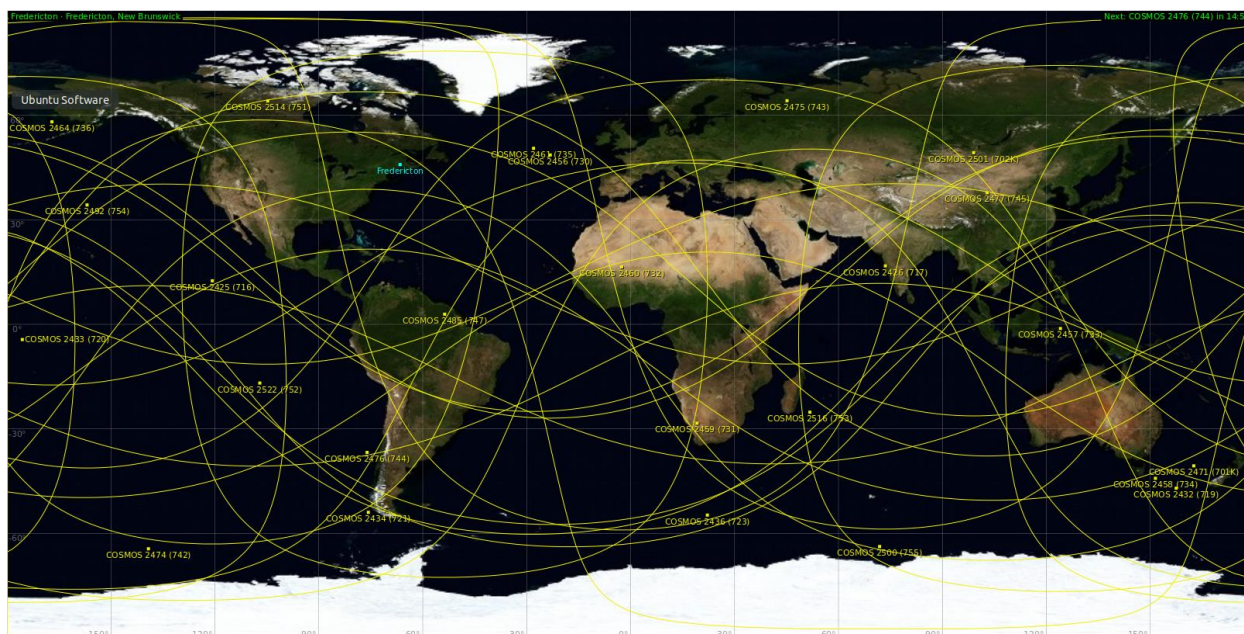
Figure 7.GLONASS Satellite Orbit (64.8 degree inclination) (src: courtesy of Gpredict)
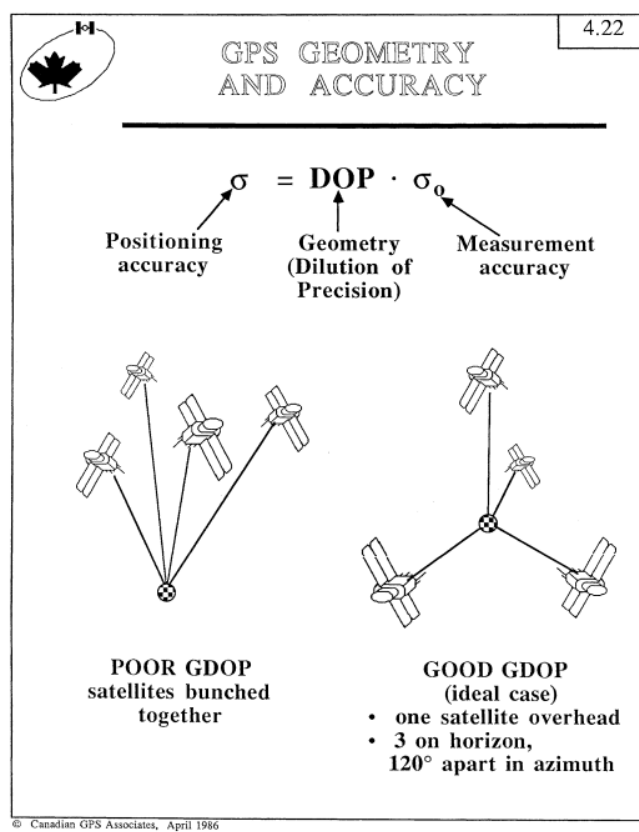

Figure 8. GNSS Geometry (DOP) (src: Wells et al, 1999)

28

Ionospheric errors also play a crucial role at higher latitudes when we have low satellite elevation (Monge, Radicella, Clara de Lacy, & Herraiz, 2011). While it is true that dual-frequency receivers correct first-order ionospheric effects almost completely, higher order effects are not handled (Jensen & Sicard, 2010). These higher order ionospheric errors are a function of the electron density distribution and the magnetic field vector along the GNSS signal propagation path (Matteo, Morton, Chandrasekaran, & Van Graas, 2009). These higher order ionospheric effects can cause errors of more than 10cm.

Several methods exist to help obtain a greater positional accuracy such as Differential GNSS (DGNSS), Real Time Kinematic (RTK), Post Processed Kinematic (PPK) and Precise Point Positioning (PPP). Due to the increased complexity of setup (base station needed), distance restrictions and associated costs involved with DGNSS, RTK and PPK solutions, PPP was chosen as the post-processing method. It can be shown that PPP provides an excellent performance especially when using final orbit solutions, available 12-18 days after collection (Table 2), to generate the final solution (Malinowski & Kwiecien, 2016). It may be argued that one need not wait for final orbit solutions and that a similar high accuracy PPP solution may be derived from rapid orbit solutions (Malinowski & Kwiecien, 2016). In either case, PPP obtains similar benefits of using a differential system as well as providing simplicity for the operator. PPP is able to provide a sub decimeter accuracy with dual frequency kinematic readings (Huber, et al., 2010).

Type		Accuracy	Latency	Updates	Sample Interval
GPS Satellite Ephemerides / Satellite & Station Clocks					
Ultra-Rapid (observed half)	orbits	~3 cm	3 - 9 hours	at 03, 09, 15, 21 UTC	15 min
	Sat. clocks	~150 ps RMS			
		~50 ps SDev			
Rapid	orbits	~2.5 cm	17 - 41 hours	at 17 UTC daily	15 min
	Sat. & Stn. clocks	~75 ps RMS			5 min
		~25 ps SDev			
Final	orbits	~2.5 cm	12 - 18 days	every Thursday	15 min
	Sat. & Stn. clocks	~75 ps RMS			Sat.: 30s
		~20 ps SDev			Stn.: 5 min
GLONASS Satellite Ephemerides					
Final		~3 cm	12 - 18 days	every Thursday	15 min

There are several choices of systems available to run the PPP processing including RTKLIB, UNB GAPS and Natural Resources Canada (NRCAN) PPP. Each of these systems provides roughly the same output, however NRCAN PPP was ultimately chosen. This is because GAPS does not support GLONASS which is deemed necessary for the project as GPS inclination is limited for the project region. Additionally, RTKLIB was deemed more difficult to run in automation and requires manual fetching of the IGS correction products. The NRCAN PPP system allows for several different methods of submission. Keeping automation goals in mind a Python script utilizing a WGET method to interact with the NRCAN webserver was created. The script allows for submission of the Receiver Independent Exchange Format (RINEX) file to the NRCAN webserver, polling of the webserver checking for completion and final fetching of the completed

NRCAN PPP output. It could be argued that the NRCAN webservice is a tried and tested system providing the most up to date PPP processing, especially with the recent 2018 update.

The final output of the NRCAN PPP service is provided as a .zip file containing of importance a .pdf summary of the processing, a .sum of the processing parameters summary, an errors.txt file outlining errors and warnings which occurred during processing and a .pos file which contains the actual position information. The contents of the final output are explained in Appendix A.

While PPP no doubt improves GNSS data it does come with limitations. The primary limitation is the long convergence time needed to obtain a solution. A (2017) paper by Yu and Goa found that multi-constellation PPP convergence times were primarily distributed between 10 to 40 minutes. However, the convergence time will vary based on the number of satellites available, satellite geometry, atmospheric conditions, receiver multipath environment and the quality of IGS correction products (see Table 2 above) (Novatel, 2018). Surveying in the Arctic does present its own limitation, particularly on higher degree ionospheric errors, satellite geometry and topographic zenith delay which add to uncertainty and diminish accuracy in the GNSS data. Additionally, in cases where we cannot obtain an accurate GNSS signal, the PPP processing will reject these readings and are documented in the .sum file. As shown, there are limitations when using GNSS for some of which we can help to mitigate, others we cannot. Either way, it is important to document these limitations in the metadata so that the final user obtains a transparent explanation of the limitations and challenges with this data.

3.3   GNSS Qualification

Once the location data has been processed using PPP, the processed GNSS readings are then placed under a threshold filter to sort acceptable vs. non-acceptable readings on a per line basis. This step is crucial since vertical height is of utmost importance in a hydrographic survey and poor GNSS readings will hinder the accuracy of the final depth sounding. If the GNSS reading is acceptable the ellipsoidal vertical height may be used, otherwise estimation of the vertical height through a hydrodynamic tidal model and a traditional water level reference approach must be done.

Table 3. DOP Ratings (src: (Langley, 1999)

DOP Value	Ratings
1	ideal
2-4	excellent
4-6	good
6-8	Moderate
8-20	Fair
20-50	poor

To filter the results a combination of five threshold variables were chosen. These variables consist of the number of satellite vehicles (NSV), geometric dilution of precision (GDOP), standard deviation of latitude (SDLAT), standard deviation of longitude (SDLON) and standard deviation of height (SDHGT). The minimum standard deviation values chosen for latitude, longitude and height are 5.0m, 5.0m and 0.5m respectively. These values were chosen to maintain the requirements for the static portion of the CHS level 1a and 1b survey standard. The maximum GDOP value selected is 3.5 which according to Langley (1999) is considered an excellent rating, see Table 3. The GDOP threshold value may be changed later to reflect a good value (4-6), which would be sufficient for the project. A minimum NSV value of 6 was selected. When surveying in the North Arctic, satellite counts and geometry are often unfavourable as they are often limited in

count and clustered towards the southern hemisphere. By including the GDOP and NSV along with the standard deviations, we are better able to assess the favorability of GNSS conditions. The NSV provides the satellite count, where a minimum of 4 satellites are required to get a 3-dimensional location. However, satellite geometry may be at a tight acute angle relative to the horizon and thus GDOP is necessary to quantify the satellite geometry as the NSV would not take this into account. These parameters combine to assess the quality of the GNSS readings on a per data-line basis and all values can be adjusted to be more or less lenient depending on requirements.

Using these threshold values, each reading is sorted on a line-by-line basis and then put into a .srt file which tags the GNSS readings as acceptable or not-acceptable. Some limitations of this step are that it is necessary for the system to be able to keep track of which readings obtained the vertical reference from GNSS or from tidal models and include this in the metadata. This tracking and documenting adds a layer of complexity to the system. Also, it is difficult to come up with a one-size fits all approach to threshold values, however using the outlined values should provide a consistent criteria for filtering. The .srt file serves as a way to track ellipsoidal vs. tidal referenced data. In this phase of the project, data will be referenced using ellipsoidal reference only (explained in the following paragraph) and thus all GNSS readings deemed 'unacceptable' will be archived for future processing.

3.4 Vertical Reduction

In order to provide a meaningful and consistent vertical reference the accepted GNSS data must be reduced to a common datum. Currently the GNSS data would be referenced to ellipsoidal height (NAD83), however this value is not very meaningful to a mariner. For example, depending on

33

location, the ellipsoidal height may be negative or positive with respect to the waterline, adding confusion for those navigating the waters. In order to have a meaningful height the data is converted to a more pragmatic hydrographic vertical datum, with the two most commonly used references being Chart Datum (CD) and Mean Sea Level (MSL). In Canada, the CHS licenses and restricts the use of Chart Datum for safety concerns and as a result MSL would likely be the vertical reference used.

To convert ellipsoid height to MSL a separation surface connecting MSL to a geodetic ellipsoid must be used. The separation surface used for this project is provided by the CHS with support from the Canadian Geodetic Survey and labeled as Hydrographic Vertical Separation Surfaces (HyVSEPs). This model extends across all tidal waters of Canada to provide continuous coverage of more than 7 million square kilometers of ocean and more than 200 thousand square kilometers of shoreline (Robin, Nudds, MacAulay, De Lange Boom, & Bartlett, 2016). The model tested has been implemented as a MatLab script with the user input to obtain the separation value being model region, reference datum, output datum and location information.

The limitations associated with reduction include errors related to the separation/conversion values themselves as well as approximations made by gridding locations to find the associated separation values. During the initial phases of this project, data is not converted and kept referenced to the ellipsoid for dissemination. This is done to ensure the raw data quality is preserved, recognizing that the interested parties may keep the ellipsoidal reference or convert to hydrographic vertical reference using a method of their choice.

3.5   Tidal Heights

When an accurate GNSS signal cannot be resolved we are unable to reliably use the vertical ellipsoid height as a reference for our sounding data. As a result, we need to accurately measure the fluctuations in the body of water due to tidal constituents in order to obtain the precise height of the vessel. Ideally, we would be able to use a tidal gauge or setup a GNSS buoy to obtain tidal records for the survey region. Unfortunately, tidal gauges are sparse in the Arctic and would not be available in the survey locations of this project. This limitation is overcome by using a hydrodynamic tidal model.

The tidal model which would best fit the needs of this project was determined to be the Arctic9 model provided Department of Fisheries and Oceans (DFO). The software built upon Arctic9 is called WebTide. This barotropic ocean circulation model covers the entire Arctic region and provides a good estimate of predicted tides (Church, Hughes, & Haigh, 2007). When averaged over the entire model domain, the rms error of the tidal constituents is about 17 cm for M2, 4 cm for N2, 7 cm for S2, 6 cm for K1 and 3 cm for O1. The large M2 error is dominated by a large tidal range occurring in Frobisher Bay (Collins, Hannah, & Greenberg, 2011). WebTide provides discrete model domains covering Canada's waters shown in Figure 9.
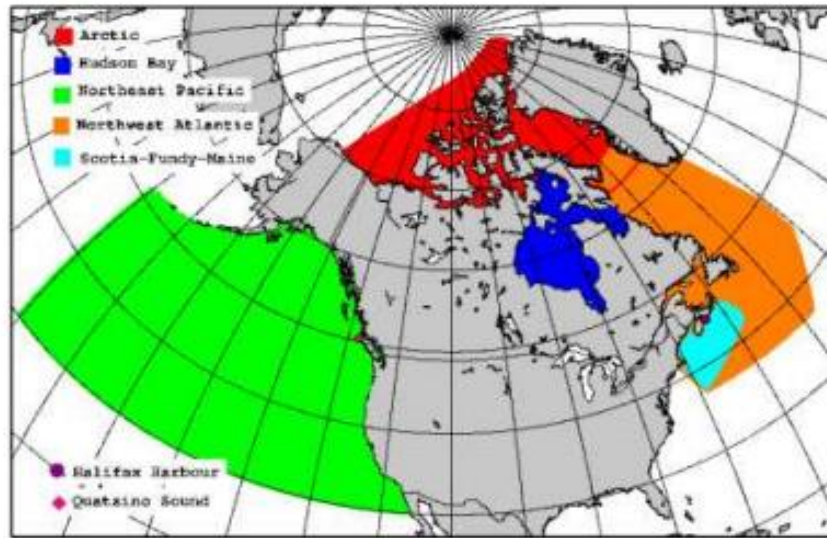
Tidecor is software which provides backend processing for WebTide. A custom implementation of Tidecor was created by OMG UNB to query WebTide which takes the GNSS latitude and longitude, year and time information are used as inputs to the model. The encompassing domain is selected with a model lookup map and then a tidal height is returned corresponding the appropriate average of the nearest three-vertices mesh element (Beaudoin, 2009).

Due to the nature of this project, the data collected will predominately be from regions close to shore. As a result, majority of the data will likely fall out of model bounds and when queried in the model will not return a tidal height value. To circumvent this issue, the nearest available node must be found and using this node the tidal height value is obtained. This method is outlined by Beaudoin (2009). The final tidal height is relative to WebTide vertical reference, believed to approximate MSL. The WebTide vertical datum is zero height above (or below) which the tides oscillate. This datum accounts for the tidal signal but not long-term or short-term non-tidal sea-

level effects. When ignoring short-term sea-level effects, the water line should reduce to MSL. However, longer-term effects may be present which adds uncertainty to the vertical measurement.

The limitations of using tidal models for vertical referencing are that it is difficult to verify the accuracy of the tidal heights from the model and thus it is difficult to truth model height compared to actual sea surface height. Collins et al (2011) provide a validation of the Arctic9 model in certain areas and could at the very least be a source to help understand the model uncertainty. Another limitation is that we must now account for draft and dynamic draft of the vessel. In the case of the HydroBall, draft will remain relatively constant depending on the towing option selected. However, there are no heave sensors and thus we cannot measure and account for this. Additionally, we are not able to measure non-tidal variance towards Sea Surface Height (SSH) (inverse barometer effect, weather patterns, irregular phenomenon) (Church, Hughes, & Haigh, 2007). We can augment the WebTide model with an operational ocean model which can help to approximate these variances, such as the oceanographic forecast model HYCOM. To transparently assess uncertainty, we must include the model used and associated model uncertainties as well as communicate the lack of heave and non-tidal variations measurements that factor into the final depth readings.

This Tidal Height step will be omitted during the pilot phase of the project as the project partners have agreed to keep final depth soundings vertically referenced to the ellipsoid. However, the purpose of the tidal section is to obtain a vertical height when the GNSS height is unreliable. Thus, the data which does not meet the GNSS quality threshold in the GNSS Qualification phase will be archived and processed during later phases of the project.

3.6 SVP

When conducting a hydrographic survey, depth is never directly measured it is only inferred. We directly measure the time it takes for the sonar signal to propagate to the seafloor and back, known as the Two-Way Travel Time (TWTT). Combining the TWTT with incidence angle and sound speed, we can calculate depth (Simons & Snellen, n.d). The speed of sound calculation is derived from temperature, salinity and pressure/depth variables and is given as a function of depth (Beaudoin, 2009). We must then have a depth related array of sound speeds to satisfy Snell's law and take account of refraction.

Due to project constraints towards operational simplicity, in-situ casts will not be taken and thus sound speed is derived from oceanographic models. When selecting a model, we must be able to have accurate temperature and salinity readings throughout the water column and the model coverage must extend to the surveying locations of this project at an acceptable temporal and spatial resolution.

Three model candidates that have Arctic coverage were selected, being the World Ocean Atlas 18 (WOA18), Real Time Operational Global Ocean Forecast System (RTOFS) and Hybrid Coordinate Ocean Model (HYCOM). While HYCOM and RTOFS are numerical baroclinic hydrodynamic ocean models, the WOA18 is an oceanographic climatology derived from a collection of in-situ oceanographic data. As a result, the WOA18 data does not provide the same temporal resolution of the HYCOM and RTOFS data except on an annual, seasonal and monthly scale, where values are interpolated across the grid (NOAA, 2019). After careful contemplation,

the HYCOM model was ultimately chosen for its sufficient Arctic coverage, archive access and webserver support, as summarized in   Table 4. The HYCOM model can provide archived data, nowcasts (current state) and forecasts of oceanic parameters, this includes three-dimensional data for temperature, salinity and depth (Metzger, et al., 2014).

Table 4. Summary of Comparisons between the three candidate Ocean Models

	HYCOM (GOFS 3.1)	WOA18	RTOFS
Temperature, Salinity, Depth Available	Yes	Yes	Yes
Horizontal Resolution	0.08 deg lon x 0.08 deg lat (40S-40N); 0.08 deg lon x 0.04 deg lat (poleward)	0.25 deg lon x 0.25 deg lat (60N-90N)	0.08 deg lon x 0.08 deg lat
Vertical Resolution	0m-5000m	0m-4000m	0m-4000m
Temporal Resolution	3 hours	Annual, Seasonal, Monthly	3 hours
Access	Thredds, FTP, OpenDAP, NetCDF, HTTPS	Thredds, HTTPS, FTP, GeoPortal, NetCDF, Live Access Server	FTP, NetCDF, OpenDAP, HTTPS
Archive Access	Yes	Yes, limited	No
Arctic Support	Yes	Yes	Yes
Automation Support	Yes	Yes, limited	Yes

Once the model was selected, Python scripts were then written on interacting with the model through the Thematic Real-time Environmental Distributed Data Services (THREDDS) server. HYCOM has several different model domains to choose from each covering different areas with differing resolution. The chosen domain was the Global Ocean Forecasting System (GOFS 3.1) specified with GLBv0.08 resolution. The specifics towards this domain are provided in Table 5.

To collect the necessary temperature and salinity data arrays needed to calculate the sound velocity, the HYCOM model requires inputs of date, time, latitude, longitude and depth for the area of interest.

These inputs are obtained from the .srt file created in the GNSS Processing module upstream. The SVP Python script will load the .srt into memory and iterate line by line on the qualified GNSS readings to extract the necessary data inputs needed to formulate the HYCOM query before connecting to the model. The HYCOM model does not directly accept the real values obtained from the .srt file, instead requiring conversion to their proprietary index to obtain results. A summary of the conversions necessary for date, time, latitude and longitude are provided in Table 6.

Date & Time: The model indexes date and time as hours from an epoch beginning at Jan 01, 2000; 00:00:00. Thus, the date and time given in the .srt must first be converted to hours since the epoch, done with assistance from the Python DateTime library. This value is then used to obtain the corresponding HYCOM time index value. Unfortunately, there exists gaps in the HYCOM model with respect to time and as a result the time index does not linearly relate to the hours since epoch. To overcome this hurdle, a lookup table must be used to obtain the correct index value for the given hours since epoch.

A separate function has been created to facilitate this logic. As the time index value max is always increasing, the lookup table must be periodically updated to include latest time values. The lookup table is stored as a Python dictionary in a text file and when the function is called during class initialization of the main SVP script, the lookup table will be loaded into memory and be updated

or created if non-existent. The Python dictionary data structure is used for quick look-up and insertion times when compared with a python list. On average the performance of a Python dictionary is O(1) for lookup and insertion with the worst case being O(n). Comparatively, the lookup and insertion times for a list are O(n) and O(log n) in worst case scenarios (Ozvald & Gorelick, 2018). Once the lookup table is created or updated, it is then queried using the hours since epoch and returns the HYCOM time index value.

Latitude: The latitude is stored in the degree, minutes, seconds format and must be converted to decimal degrees to obtain the HYCOM latitude index. After conversion, the decimal degree value is then sent to a helper function which returns the correct index value based on the input latitude with respect to the latitude ranges and the index step size described in Table 6. More specifically, the helper function determines which range in the HYCOM grid latitude (80^0S-90^0N) the input latitude falls within. Once the range is determined, a calculation involving the grid step (0.08^0 between 40^0S-40^0N; 0.04^0 poleward of 40^0N and 40^0S) is done to determine and return the HYCOM latitude index value.

Longitude: Similarly to the latitude, the longitude must first be converted to decimal degrees. After conversion, this value is sent to a helper function which returns the correct index value based on the parameters shown in Table 6. Longitude is recorded as degrees East (0^0 - 360^0) with a constant grid step of 0.08^0 in the HYCOM model used. This makes the calculation very straightforward of dividing the longitude by the grid step to obtain the HYCOM longitude index value.

Depth: The index values of 0:39 are consistently used in all queries. This will return data for the entire HYCOM depth range which corresponds to 0m-5000m.

The .srt file contains potentially hundreds to thousands of lines with unique data values. Due to the model resolution limits, identical queries may be produced for input data close in space and time proximity. It would be burdensome to processing efficiency and the THREDDS server to run each query individually. To minimize queries being sent to the HYCOM server, it was necessary to create a query cache and thus allowing only unique queries to be sent to the model.

Table 5. Summary of GOFS 3.1 Specifications (src: https://www.hycom.org/dataserver/gofs-3pt1/analysis)

HYCOM (GOFS 3.1)	
Title	Global Ocean Forecasting System (GOFS 3.1)
Grid	GBLv0.08
Institution	Naval Research Laboratory: Ocean Dynamics and Prediction Branch
Span	80^0S – 90^0N
Data Range	July,2014 – Present
Temporal Frequency	3 hours
Resolution	90^0N – 40^0N: 0.08 deg lon x 0.04 deg lat 40^0N – 40^0S: 0.08 deg lon x 0.08 deg lat 40^0S – 80^0S: 0.08 deg lon x 0.04 deg lat

Latitude Hycom Map

index range: [0:3250]

index	value (deg)	step (deg)
0-1000	(-80) - (-40)	0.04
1000-1500	(-40) - (0)	0.08
1500	0	0.08
1500-2000	(0) - (40)	0.08
2000-3250	(40) - (90)	0.04

Depth Hycom Map

index range: [0:39]

index	value (m)	step (m)
0-6	0-12	2
7-14	15-50	5
15-19	60-100	10
20-21	125-150	25
22-26	200-400	50
27-32	500-1000	100
33-34	1250-1500	250
35-37	2000-3000	500
38-39	4000-5000	1000

Longitude Hycom Map

*longitude formated in degrees East

index range: [0:4499]

index	value (deg)	step (deg)
0	0	0.08
4499	359.92	0.08

Time Hycom Map

* Hours since epoch (2000-02-02 00:00:00)

index range: [0:T]; T updates every hour

index	value (h)	step (h)
0	157812	3.06

After the HYCOM index values have been obtained for date/time, latitude, longitude and depth, the model is then connected to and the queries for temperature and salinity are made. The resulting temperature-depth and salinity-depth arrays are stored temporarily in memory.

In the case where null data is returned by the model, a radial search function is executed to find the nearest-neighbour with non-null data. The search function works by using a counter and index to iteratively search a radius from given location by one step. The function starts off going up from the center by the step size. If this fails it returns to the center and goes right by one step size. This is repeated going down and to the left respectively. If no data is returned, the function will then search the diagonals, beginning with going up and to the right of center by one step size, down and

43

to the left of center by one step size, up and to the left of center by one step size and down and to

the right of center by one step size. If these queries fail, the step size is increased by one and the

search is repeated until data is found. Note that the radial search only applied to the location, date

and time are held constant during the process.

After retrieval of the temperature and salinity arrays, the data must be converted from HYCOM

values to real world values to calculate the sound velocity. The conversion factor is provided in

equations 1 and 2.
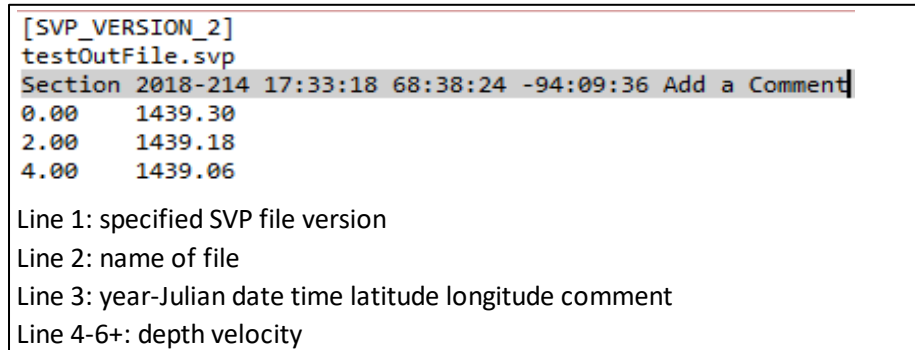
$$sal = (int(sal) * 0.001) + 20 \tag{1}$$

$$temp = (int(temp) * 0.001) + 20 \tag{2}$$

Once converted, the temperature and salinity arrays are combined with latitude and longitude

values and the sound velocity is calculated using the TEOS-10 algorithm. The significant change

that occurs with TEOS-10 when compared the traditional methods is that the use of Absolute

Salinity is adopted over Practical Salinity when describing the content of salt in seawater. Absolute

Salinity is the mass fraction of salt in seawater and has units of g/kg whereas Practical Salinity is

the measure of conductivity of seawater and is unitless. This has numerous advantages which can

be found in the TEOS-10 manual (Intergovernmental Oceanographic Commission, 2010), with the

primary benefit being increased accuracy from the use of Practical Salinity. The main disadvantage

to the TEOS-10 algorithm is the complexity of the algorithm, making it unfeasible to derive our

own implementation. Fortunately, a Python implementation is freely available in the Gibbs-

SeaWater (GSW) Oceanographic Toolbox Libraries, which provided the functions to convert

Practical Salinity to Absolute Salinity and to perform the TEOS-10 sound velocity calculation. The final sound velocities are formatted as velocity-depth arrays in the Caris format (Teledyne Caris, 2018).

Finally, the Caris SVP file is saved in the working directory. This file format is described in Figure 10. Under the comment section, the lines which the SVP are applicable to in the .srt are given as well as the query used to generate the data from HYCOM. This is done to ensure transparency in processing and reproducibility. A single SVP file is created for each unique sound velocity profile which is derived from each unique HYCOM query.

```
[SVP_VERSION_2]
testOutFile.svp
Section 2018-214 17:33:18 68:38:24 -94:09:36 Add a Comment
0.00     1439.30
2.00     1439.18
4.00     1439.06

Line 1: specified SVP file version
Line 2: name of file
Line 3: year-Julian date time latitude longitude comment
Line 4-6+: depth velocity
```

Figure 10. Caris SVP File Format Specifications Example

The challenges associated with collecting accurate SVP involve model limitations and sound velocity conversion. Resolutions for the model grid are not ideal in both location and time, leading to inaccuracies. More so, truthing the model data is difficult as it requires in-situ casts which are difficult and costly to obtain. Limitations when converting to sound velocity are that we must be trusting of the algorithm chosen. The sound velocity algorithm and chosen model should be the best selection for the area of interest and ideally would be applicable universally across survey

45

regions. To communicate transparency to the final user, metadata should include the uncertainty associated with the chosen SVP model obtained through comparison with observed data, limitations of the search and query functions and the sound speed calculation used.

3.7 Uncertainty

Uncertainty must be calculated throughout processing to provide an accurate estimation of the uncertainty in the data. The following paragraphs will explain each of the sources of uncertainty, the validity of the source of uncertainty for this project, and how to measure it. Finally, the applicable uncertainties will be combined using the Total Propagated Uncertainty (TPU) method outlined by Hare (1995) to create the final uncertainty calculation for both the horizontal and vertical components.

It will be assumed when calculating uncertainty that raytracing will always be done for a roll angle of +/- 6^0, to ensure the sonar beam is always at nadir. Also, when applying the PPP processing it is assumed that the IGS correction products at the Rapid or Final level will be applied.

3.7.1 GNSS Uncertainty:

All GNSS systems are composed of many satellites orbiting Earth. At a given time there may be several satellites above the horizon with a direct path to a GNSS receiver. This allows reception of the signals transmitted by the satellites in order to compute a position. However, there exist errors which arise between the satellite and receiver on the signal which degrades positioning accuracy. These errors present themselves as clock errors, satellite ephemeris errors, ionospheric delay errors, tropospheric delay errors, multipath errors and GNSS hardware errors. While there

are potentially many sources of errors, the main contributors on positioning error will be discussed, and if quantification of these errors is possible, a method would be proposed.

Clock Errors: These errors occur at the satellite and receiver level.

Satellite Clock Errors: The atomic clocks present in GNSS satellites are known to be extremely precise. However, these atomic clocks are still prone to drift a small amount, and unfortunately, even a small inaccuracy can result in a significant error in the calculated position by the receiver. For example, a 10ns clock error can result in a 3m position error (Novatel, 2018). Fortunately, the satellite clock error can largely be resolved using precise clock corrections provided by the International GNSS Service (IGS). Through the use of Rapid or Final clock corrections, this error can mostly be compensated with PPP processing (International GNSS Service, 2019). While this error is largely mitigated by using PPP combined with the correction products, such as those provided by the IGS, these correction products do still have some inherent errors present. The IGS states an average of 75ps RMS with a standard deviation of 25ps when using rapid and final clock correction products. This error is accounted as part of the PPP processing. However, the error of the correction products are omitted from the final uncertainty calculation as it is difficult to convert this uncertainty to a ground level location error, however it is still important to know that it exists.

Receiver Clock Errors: While the GNSS satellites have precise atomic clocks aboard, GNSS receivers tend to have less precise and more inexpensive clocks. This results in the receiver clock drifting from the satellite clock, which causes range errors in the position calculation (Karaim, Elsheikh, & Noureldin, 2018). Fortunately, this error is largely removed through estimation as an additional unknown variable in the position estimation process (Karaim, Elsheikh, & Noureldin,

2018). As this error is largely accounted for by the receiver, it is omitted from the uncertainty calculation.

Satellite Ephemeris Errors: Satellites travel in very precise and well-known orbits, but just as the atomic clock drifts by a small amount, so does the satellite in orbit. As a result, the ephemeris error is caused by a deviation between the true satellite position and the computed position by the receiver. Fortunately, through the use of ephemeris corrections provided by the IGS, the satellite ephemeris error can be largely mitigated through the use of PPP along with correction products. Similarly to the satellite clock correction products, the orbit correction products do have some inherent errors as well. The IGS states an average of 2.5cm uncertainty of the satellite tract when using Rapid and Final correction (International GNSS Service, 2019). This error is accounted as part of the PPP processing. However, the errors of the correction products are omitted from the final uncertainty calculation again due to complexity in converting this to a ground-based location uncertainty, however it is still important to know that it exists.

Ionospheric Delay Errors: The ionosphere is a layer of atmosphere between 75-1000km above Earth (Stanford, 2018). This layer of the atmosphere contains electrically charged ions, which can cause a delay in the GNSS signal travelling through (Hegarty, 2013). The level of ionization is primarily affected by the intensity of solar activity but also the season and time of day. Through the use of dual-frequency receivers and PPP, this error can be mostly eliminated at about 99% (Karaim, Elsheikh, & Noureldin, 2018). As the dual-frequency bands of L1 and L2 are accounted for during the PPP processing, the ionospheric delay error is omitted from the uncertainty calculation.

Tropospheric Delay Errors: The troposphere is the layer of atmosphere closest to the Earth. Humidity, temperature and atmospheric pressure cause the GNSS signal to become delayed as it travels through the dry gasses and water vapour (Hegarty, 2013). The wet portion of this error accounts for approximately 10% of the delay with the dry portion accounting for the other 90% (Karaim, Elsheikh, & Noureldin, 2018). While the atmospheric pressure and air temperature of the dry portion can be modelled and accounted for, the humidity (wet portion) is largely variable and localized, making it difficult to mitigate (Karaim, Elsheikh, & Noureldin, 2018). The simplest way to take into account the wet portion of the tropospheric delay error is to use differential GNSS or RTK, both of which are not utilized as a part of this project. Fortunately, the NRCAN PPP systems does account for both the dry and the wet portions of the tropospheric delay through the use of modelling. The combined estimated error from the dry and wet portion of the tropospheric delay is estimated in the standard PPP output as a standard deviation (95%) at zenith for each data line in the file. Using sample data from this project, on average, this error is relatively small at sub-centimeter level, though on occasion it approaches decimeter level. This error is accounted for during the PPP processing and thus omitted from the final uncertainty calculation.

Multipath Errors: As the satellites send the signal to the receiver, there may be instances where the signal reflects of an object before reaching the receiver antenna. As the reflected signal travels farther to reach the receiver antenna, this causes the receiver to calculate an incorrect position. This error cannot be mitigated through post-processing or differential positioning and can limit positioning accuracy even when all other error sources have been eliminated (Karaim, Elsheikh, & Noureldin, 2018). A simple way to avoid this error is to keep the GNSS receiver in an area free

of obstructions to prevent the satellite signal from reflecting. However, this is not always practical as this project requires the GNSS receiver to be kinematic.

Fortunately, multipath errors do not occur significantly when conducting a hydrographic survey. As the hardware in this project is being towed, this eliminates any multipath errors caused by human interference. Additionally, the NRCAN PPP processing implements a cutoff rate of 7.50, which helps to mitigate any multipath errors reflecting off the water surface. As a result, the quantification of multipath errors will be omitted from the final uncertainty calculation.

GNSS Hardware Errors: There exists certain errors introduced from the receiver while measuring satellite signals. This error is complex and is generated from the receiver hardware and/or software. These errors include, but are not limited to, noise introduced by system components such as cables, antenna and amplifiers; microwave radiation sensed by the receiver antenna unrelated to the GNSS signal; and signal quantization noise (Karaim, Elsheikh, & Noureldin, 2018). Receiver noise cannot be avoided entirely and can vary between GNSS receivers. As this error is difficult to quantify, can vary between each receiver and typically is very small (mm level), it is omitted from the final uncertainty calculation.

3.7.2   PPP Uncertainty:

After processing, the NRCAN PPP provides an estimation of positioning error as a standard deviation (95%). There is an uncertainty introduced into the final horizontal and vertical positioning as a result of error introduced during the PPP processing unaccounted for with this error estimate. These errors in general, are a function of the implementation and accuracy of the

models used in the software and also dependent on the quality and content of the submitted observation files (Yves, Pierre, Francois, Pierre, & Jan, 2008), (Guo, 2015), (Malinowski & Kwiecien, 2016), (Abdelazeem & Celik, 2014). Both the inherent PPP engine errors and the estimated errors as part of the system output will be discussed.

PPP Engine Errors: The NRCAN PPP processing engine changed in August 2018. As a result, quantification of the processing error is unknown as studies on the new engine have not been released in the academic community, and older papers quantify errors on the legacy system, such as (Yves, Pierre, Francois, Pierre, & Jan, 2008). As a result, the PPP engine errors will largely be omitted. It is important to note that the new NRCAN PPP engine does take into account Ocean Tide Loading (OTL) and Solid Earth Tides, however the Atmospheric Tides are unaccounted for and contribute to the Engine errors.

In general, it is widely accepted that the NRCAN PPP engine is one of the best, if not the best, PPP processing engines currently available. The engine employs up to date methods which have been tried and tested and considered the best. For example, the tropospheric method used complies with the IERS VMF1 standard and is considered to be the most accurate function. Overall, the NRCAN PPP processing will yield sub-decimeter results for dual-frequency kinematic GNSS, which can drastically improve non-aided GNSS readings. At this moment there exists no study on the error of the new NRCAN PPP engine itself, so the inherent engine errors will be omitted from the final uncertainty calculation. However, it is important to note that there exist errors corresponding to the products used while calculating the least squares solution such as the IGS correction products for both satellite clock and orbit. Errors also are formed due to the omission

of Atmospheric Tides from the PPP processing. This will propagate an uncertainty into the final location measurements provided by the engine.

PPP Estimated Errors: The NRCAN PPP system estimated the positional errors as a standard deviation at 95% in meters. The corresponding errors are estimated throughout the NRCAN PPP processing as a product of the least squares calculation. These positional errors reported as North, East, Up components (StDev_North, StDev_East, StDev_Up respectively) can be added to the uncertainty calculation directly for the vertical component and almost directly for the horizontal component. The North and East components are converted to a single value using the distance root mean squared (DRMS) calculation (Novatel, 2003). The formulas for each of the error components are below in equation 3 and 4.

$$\text{PPP\_Horizontal\_Error}^2 = \text{StDev\_North}^2 + \text{StDev\_East}^2 \qquad (3)$$

$$\text{PPP\_Vertical\_ Error}^2 = \text{StDev\_Up}^2 \qquad (4)$$

These estimated errors can be used to calculate the uncertainty for each GNSS data line and are used for the final positioning uncertainty.

3.7.3   Sonar Sounder Uncertainty:

Geometric Range Error: The geometric range is the radial distance between the sonar and the target, see Figure 11. As a reminder, to calculate the geometric range a short sonar pulse of duration

52

τ is transmitted towards the target. The sonar receiver records the time until the signal is returned (Hansen, 2012). The time delay, t, is then used to calculate the range using the following formula:

$$\textbf{R = ct/2} \tag{5}$$

where R = range(m); c = speed of sound in water(m/s); t = time delay(s)

Note that a division of 2 is necessary to account for the two-way travel time (TWTT) and the sound speed must be known to calculate the range.
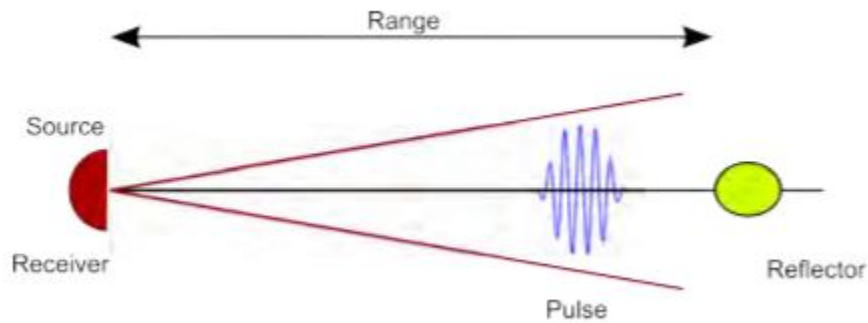
Figure 11. Estimation of the Geometric Range (Hansen, 2012)

The geometric range error, also known as the range resolution, can be calculated if the sound speed is known as well as the pulse length, τ, of the sonar. The pulse length is the duration of time a given frequency is emitted by the sonar transducer and is measured in seconds. The range resolution dictates the minimum spacing two sonar pulses can be separated and still be detected. Having a shorter pulse length will grant a better range resolution, however this is at the expense

of range as shorter pulses have less transmitted energy (Hansen, 2012). The geometric range error is shown in Figure 12.
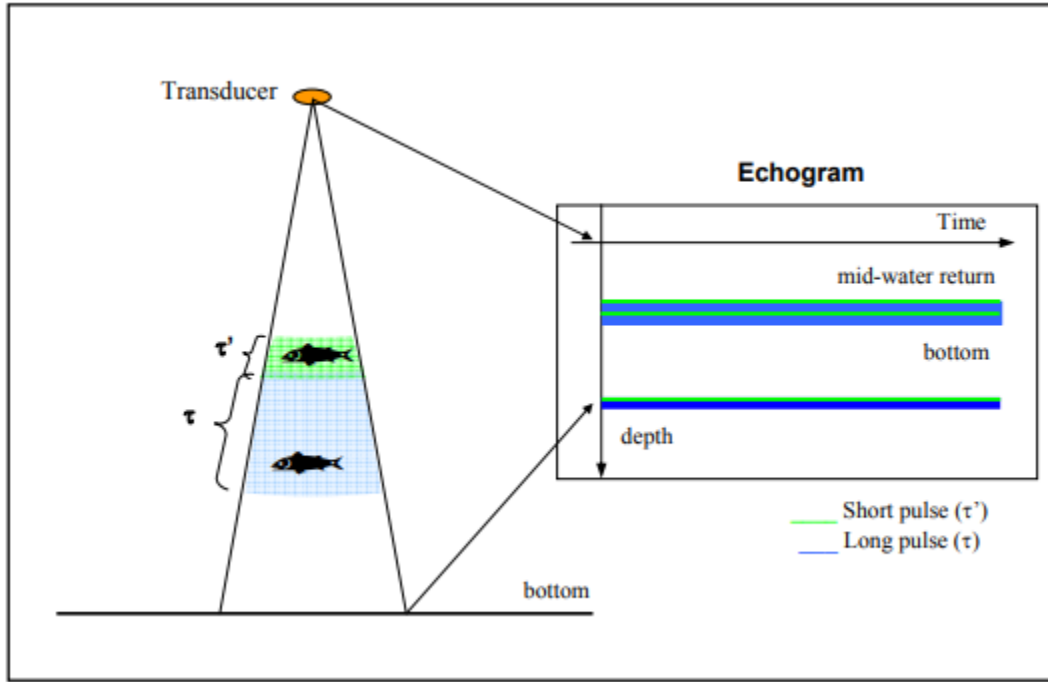
Figure 12. Geometric Range Error (International Hydrographic Organization, 2018)

The equation for a continuous width pulse geometric range error is the following:

$$\sigma_R = c * \tau/2 \tag{6}$$

where σ_R = range error(m); c = speed of sound in water(m/s); τ = pulse length(s)

Note that again a TWTT must be accounted for, so the result is divided by 2 and that the sound speed must also be known to calculate the range error.

The HydroBall features an adjustable pulse length which is typically set to 100-150 μs and thus the maximum geometrical range can be computed as follows:

$$\text{Range Error}_{100} = 1460\text{m/s} * (100\ \mu s\ /\ 10^6)\ /2 = 0.073\text{m or } 7.3\text{cm} \qquad (7)$$

$$\text{Range Error}_{150} = 1460\text{m/s} * (150\ \mu s\ /\ 10^6)\ /2 = 0.1095\text{m or } 10.95\text{cm} \qquad (8)$$

Note that the sound speed used for the calculation is 1460m/s which is the approximate average sound speed of all 17 casts utilized later in the results section.

From the calculation above we can see the range error to be between 7.3cm and 10.95cm. This error is added to the uncertainty calculation as a vertical component. The value of 10.95cm is used to be conservative during final position error estimation.

Raytracing Error: The errors in the raytracing process occur due to two separate errors; 1) a sound speed profile error, and 2) a beam measurement error.

Sound Speed Profile Error: When conducting the beam raytracing, a sound speed profile is used to compute the refraction angle at each interface and thus the final beam angle. However, when supplied an incorrect sound speed profile, the result is an incorrectly calculated beam angle at the sound speed interface, as a result of the calculation from Snell's Law. Thus, the sound speed profile error is a deviance between the true geometric beam angle from the calculated geometric beam angle (Hare, 1995). This error is related to the Beam Angle Measurement Error explained below.

Beam Angle Measurement Error: As previously mentioned, as a beam travels through a body of water, the beam experiences bending with changes in the sound speed gradient, as defined by Snell's Law (Hare, 1995). Figure 13 depicts the sound speed interface where bending occurs. There are many of these interfaces throughout the water column. In reality, these interfaces are continuous and not as discrete as pictured, however since the sampling is discrete, it is depicted as such. The discrete calculation of the beam angle will yield an ever so slight error compared to if it a continuous interface was used for the calculation. Thus, the beam measurement error is then a result of the deviance from the calculated beam angle from the true beam angle.
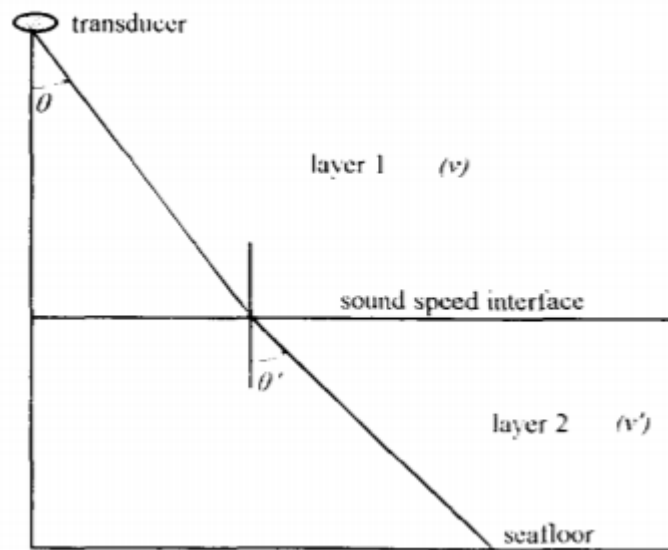
Figure 13. Sound Speed Interface [Snell's Law] (src: (Hare, 1995))

This particular error will consist of both a horizontal and vertical components when the beam exits the sonar at an angle. As a result, this error can be significant when using a multi-beam sonar as this system relies on sending beams at various angles. In a single-beam system, such as the one

56

used in this project, the beam will theoretically travel straight down through the water column at nadir. This eliminates the bending in the beam due to Snell's Law as there is no need to calculate a beam angle and subsequently removes the beam measurement and sound speed profile errors.

However, in reality, the sonar system experiences a combination of roll and pitch, which may prevent the sonar from firing at nadir. However, if these rotations are smaller than ½ the beam width, they can be ignored (International Hydrographic Organization, 2017). Figure 14 depicts this using a 12^0 beam width at various roll angles between +/-6^0. From the figure it is shown that if the tilt angle is equal to half the beam angle (+/- 6^0), there will be an ensonified area at nadir.
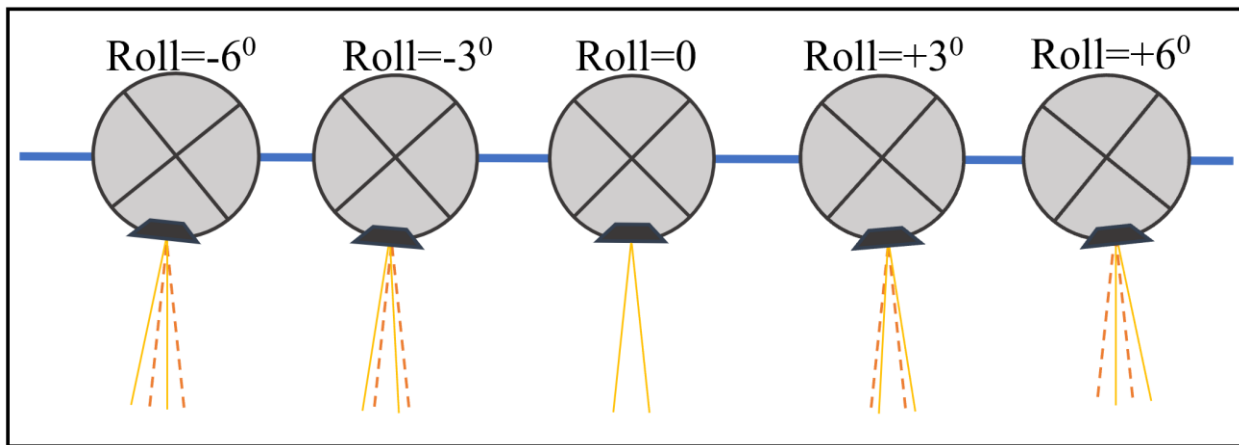
Figure 14. Rotations within ½ beam width can be ignored with a SBES

In the case of the HydroBall, the beam width is 12^0 meaning that any rotation less than +/-6^0 need not be accounted for as it will not influence the geometric beam angle error and sound speed profile error. When analyzing the tilt sensor readings aboard the HydroBall for 11 separate surveys, it was calculated that the rotation was less than +/-6^0 on average 94.18% of the time, see Table 7. When

57

analyzing the readings which were above the $+/-6^0$ threshold, it was found that most of the corresponding GNSS readings would not have passed the qualification threshold due to high values in the standard deviations of the positioning.

Table 7. Eleven surveys were qualified to see how many instances the IMU recorded a greater than 6-degree roll (1/2 beam width)

	FALSE	TRUE	TOTAL	% TRUE
1	3310	11420	14730	77.53%
2	90	3515	3605	97.50%
3	386	14767	15153	97.45%
4	0	154	154	100.00%
5	18	46	64	71.88%
6	268	12939	13207	97.97%
7	51	7882	7933	99.36%
8	281	7223	7504	96.26%
9	0	1264	1264	100.00%
10	1411	117601	119012	98.81%
11	514	69852	70366	99.27%
	Average TRUE %:			**94.18%**

Thus, it can be shown that the beam width from the single-beam echosounder is wide enough to absorb most of the positioning errors caused by rotations of the vessel. In the cases where the threshold was exceeded, the GNSS data typically is not favourable and would not be utilized. For these reasons, the raytracing error is not accounted for during the uncertainty process.

Beam Width Error: The beam width is a function of the transducer dimensions and the frequency, a higher frequency with a large transducer will yield a more narrow beam when compared to a lower frequency and smaller transducer (International Hydrographic Organization, 2017). As the sonar beam propagates through the water column, the ensonified area will expand spherically as a function of beam width, pulse width and depth (Smith W. , 2008), see Figure 15.
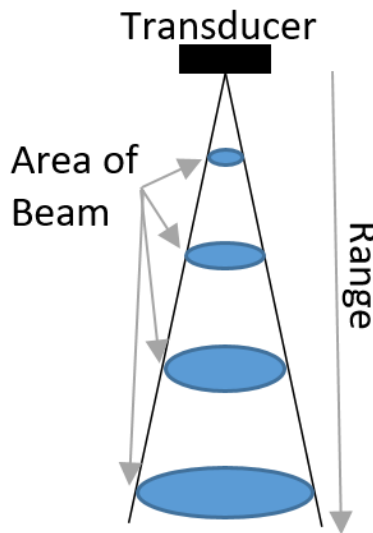
Figure 15. Ensonified Area as a result of beam width and depth (src: (Simmonds and MacLennan, 2006 (modified)))

Therefore, as the range increases for a given sonar beam, the ensonified area will also increase. As single-beam sonar bases bottom detection on the first return, it is commonly assumed that the single-beam will base bottom detection on the most shallow point within the ensonified area (Smith W. , 2008). However, this first return can occur anywhere within the ensonified area, and this leads to a horizontal error of +/- the ensonified width. In Figure 16, the first return is coming from an object closer to the surface than the seafloor. The depth for this sonar return is then recorded as the height of the anomaly which occurs somewhere within the ensonified area. However, it is unknown exactly where this anomaly lies within this area, and thus the horizontal error is equal to the ensonified width at this depth.
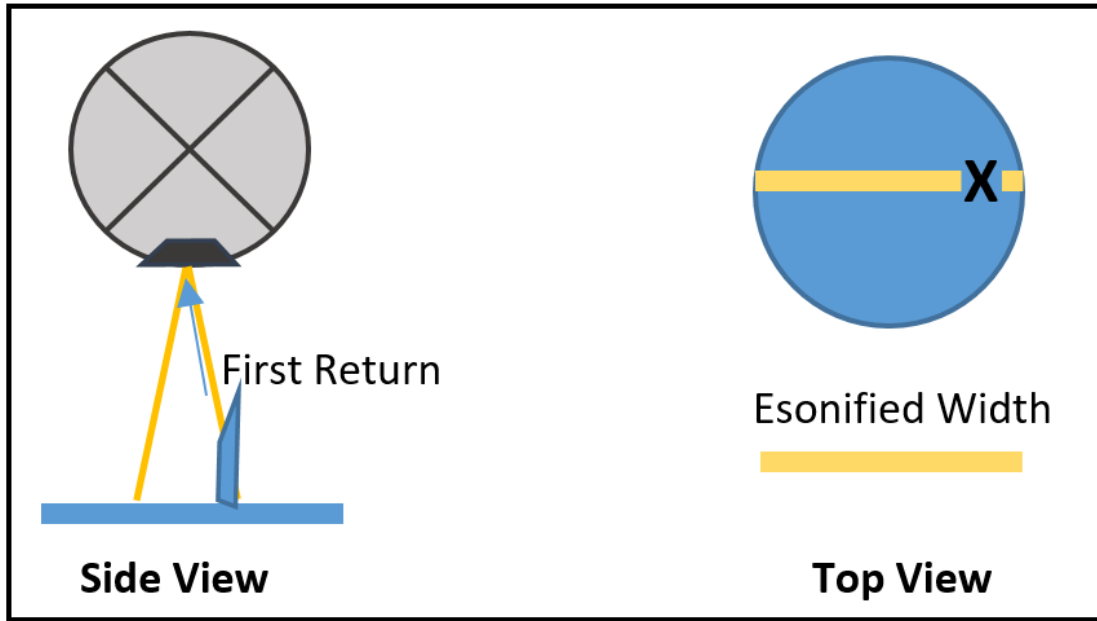
Figure 16. Beam Width can influence the positioning of the first return

Creating 10m intervals in the depth range of the sonar system used in this project and the associated beam width, an ensonified width in meters has been calculated, see Table 8. Here we can see that as depth increases, the ensonified width also increases. The IHO maximum THU for the depth interval at level 1a/1b and 2 are also given. Cells highlighted in green are within the THU for that level of standard and cell highlighted red exceed the threshold for the standard. For the sonar used in this project, the beam width error for any depths over the 20 to 29 meter interval will exceed the entire horizontal uncertainty for level 1a and 1b.

Table 8. Calculated beam width and influence on S44 error budget by depth interval (m)

Beam Width = 12 degrees			
Interval	Ensonified Width (m)	S44 1a/1b	S44 2
0-9	1.89	5.45	20.9
10-19	3.99	5.95	21.9
20-29	6.1	6.45	22.9
30-39	8.2	6.95	23.9
40-49	10.3	7.45	24.9
50-59	12.4	7.95	25.9
60-69	14.5	8.45	26.9
70-79	16.61	8.95	27.9
80-89	18.71	9.45	28.9
90-100	21.02	10	30

In regard to safety of navigation, when using a single beam sonar the wide beam will always detect the shoal, although with a potential at portraying the shoal larger than it really is. In this sense, a hazard to navigation will be correctly placed however may just appear larger on the sonar return. The issue arises when needing to represent the sea floor accurately as the single beam may smear the object beyond its true proportions and may completely miss holes in the sea floor. As a result, this error is not commonly incorporated into the uncertainty calculation when the primary goal is safety of navigation. On one hand, if this error is included, it can quickly consume the entire horizontal uncertainty budget and may limit the usefulness of the data as the final positioning accuracy would be overly conservative. However, on the other hand, it is an error that exists and should be accounted for especially when a concern for anything other than safety of navigation is of importance, such as representing the real sea floor.

In this project, safety of navigation is the greatest concern, so this error will not be included in the final uncertainty calculation. However, users of the data should be aware of the beam width error

and know that any recorded anomaly will have a degree of horizontal uncertainty and so a separate calculation including this error will be given in this report to illustrate the effects of this error on the final uncertainty.

3.7.4   Lever Arm Uncertainty:

Lever arms or sensor offsets are measured as part of a reference frame which is used to compute the ship's coordinate system (Clarke, 2003). Lever arms exist between the GNSS receiver, Sonar and IMU sensors. The lever arm values typically need to be calculated using surveying methods. However, in a standalone manufactured device such as the HydroBall, these lever arms are known as part of the design. The uncertainty arises from manufacturer errors which cause the actual lever arm values to deviate from the value laid out in the design. As the HydroBall is a compact device, lever arms are kept as short as possible and thus their effects on the total uncertainty is minimized. While the manufacturer of the HydroBall does not explicitly provide estimations for the lever arm uncertainty as a whole, manufacturing tolerances are provided. Based on these measurement tolerances and personal communication with the manufacturer, the estimated level arm uncertainty can be conservatively approximated to 0.005m at 95%. Thus, the level arm error of 5mm will be included in both the vertical and horizontal uncertainty calculation.

3.7.5   Reduction Uncertainty:

When conducting a hydrographic survey, typically positioning is referenced to the ellipsoid via a GNSS receiver, known as an Ellipsoidal Referenced Survey (ERS). Unfortunately, ellipsoidal referenced positions are not practical to mariners, specifically when considering height/depth of water. Typically, these positions are converted (reduced) to a more practical hydrographic datum, such as Mean Sea Level (MSL) or Chart Datum (CD). This conversion is not perfect, and these

errors create vertical uncertainties. To perform this conversion, a hydrographic separation surface has been released in 2016 allowing conversion to and from ERS to hydrographic datums in all Canadian waters. This separation surface, named Hydrographic Vertical Separation Surfaces (HyVSEPs), captures the spatial variability of the tidal datum using semi-empirical models, tidal station observations, sea level rise estimates, dynamic ocean model results, satellite altimetry and a geoid model (Robin, Nudds, MacAulay, De Lange Boom, & Bartlett, 2016).

The accuracy targets of sub decimeter levels were believed to be achieved with the development of HyVSEPs, however the approach used for validation relied on biasing to published benchmarks and has not yet been validated. (Renoud, 2018) has demonstrated that this accuracy target may be underestimated, especially in the Arctic region, and further validation of the target is needed. Robin et. Al. in their study validating the HyVSEPs grids, found error estimates of 6.6cm and 17.7cm for the CANNORTH and CANHUD grids respectively (Robin, Nudds, MacAulay, De Lange Boom, & Bartlett, 2016).

This being said, reduction is not necessary at this phase in the project as data will be kept referenced to the ellipsoid. As a result, the reduction error will not be applied in the final uncertainty calculation. In the case where this changes and reduction will take place using HyVSEPs, a vertical uncertainty of 6.6cm or 17.7cm depending on region may be applied to compensate for this error until further validation is made.

3.7.6   Dynamic Draft Uncertainty:

Dynamic draft is described as the instantaneous depth of the sonar transducer below the reference water line. The dynamic draft is comprised of three components of which the relation is shown in the equation 9 below (Hare, 1995).

Dynamic Draft = Draft – Squat – Load (9)

Draft is the vertical distance between the referenced water level and the center of the sonar transducer when the vessel is at rest. Squat is the dynamic component of the draft and is a result of hydrodynamic forces acting on a vessel as the water moves across it. Specifically, as water is displaced it creates an increase in the current velocity moving across the hull and causes a general lowering of the vessel into the water from a decrease in water pressure (Church I. , 2017). Thus, squat is a function of the depth of water and the speed of the vessel (Hare, 1995). The load is the change in the draft over time, typically due to reduction in fuel and supplies aboard the vessel. If the draft was measured before and after the survey, a linear function can be derived and used to approximate the load as a function of time (Hare, 1995).

As this project will keep data referenced to the ellipsoid, the need for draft, squat and load calculations is bypassed (Clarke, Dare, Beaudoin, & Bartlett, 2005). As a result, the Dynamic Draft need not be accounted for in the final uncertainty calculation. If a hydrographic based datum is used as a reference in the future, estimates for draft and squat must be made, load can be ignored as the HydroBall does not experience any loading.

3.7.7   Roll and Pitch Uncertainty:

There exists an error due to the uncertainty in the pitch and roll accuracy of the IMU. This error can arise in both the horizontal and vertical directions. The Honeywell HMR3400 used in the HydroBall has an accuracy of 0.5^0 between 0^0 and $+/-30^0$ and 1.2^0 between $+/-30^0$ and $+-60^0$ (Honeywell, 2005). When a roll occurs, the position of the GNSS receiver has an uncertainty in the horizontal and vertical directions due to changes in the GNSS receiver's location in the frame, see Figure 17. When the roll angle is zero, there is only a horizontal error and when the roll angle is not equal to zero, there exists both horizontal and vertical errors:

Horizontal Error: $\sigma_H = 2*R*Sin(\sigma_{IMU}/2)*Cos\theta$ (10)

Where θ = roll angle; σ_{IMU} = IMU error; R = distance between the reference point and the GNSS receiver

Vertical Error = $\sigma_V = 2*R*Sin(\sigma_{IMU}/2)*Sin\theta$ (11)

Where θ = roll angle; σ_{IMU} = IMU error; R = distance between the reference point and the GNSS receiver
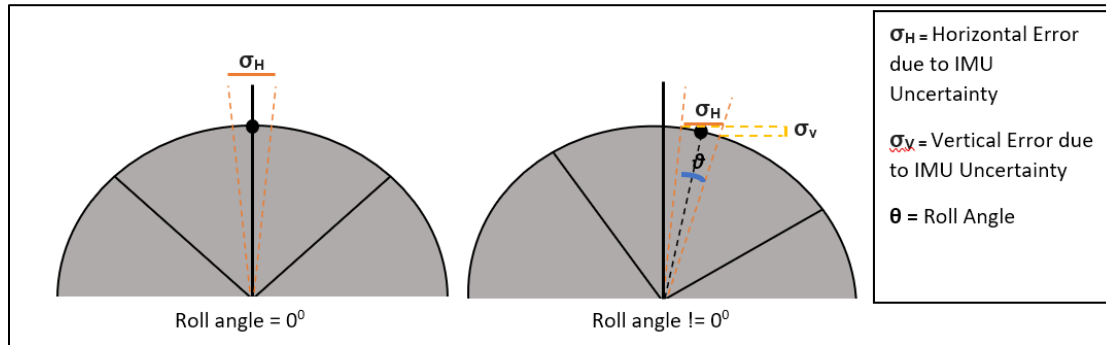
Figure 17. Horizontal and vertical components of roll and pitch uncertainty

65

This error comes with the assumption that the maximum tilt of the HydroBall will be +/-6^0, fitting with the beam width error and explanation. Precise estimations for the roll and pitch errors may be made by analyzing the exact roll angle and determining the appropriate IMU uncertainty. However, for simplicity the maximum roll and pitch error will be calculated using a constant roll angle of +/-6^0. This will help speed up processing by preventing cross-referencing of IMU and GNSS data. Using this given roll angle, the maximum horizontal and vertical roll and pitch errors may be 0.174cm and 0.018cm respectively. These values will be used in the final uncertainty calculations for both the horizontal and vertical components.

3.7.8 Latency Uncertainty:

When conducting a survey, the data collected by the individual sensors in the survey system are not synchronous. The data from these various sensors are not acquired at exactly the same time and thus a latency error must be considered (Naankeu Wati & Seube, 2016). For example, the GNSS samples at a frequency typically higher than that of the sonar. If the timestamp of the GNSS is used as the base system time, the sonar may not truly have recorded data at that specific time. The difference in timing between the timestamp (blue) and the true measurement time (black) is known as the latency uncertainty, see Figure 18.
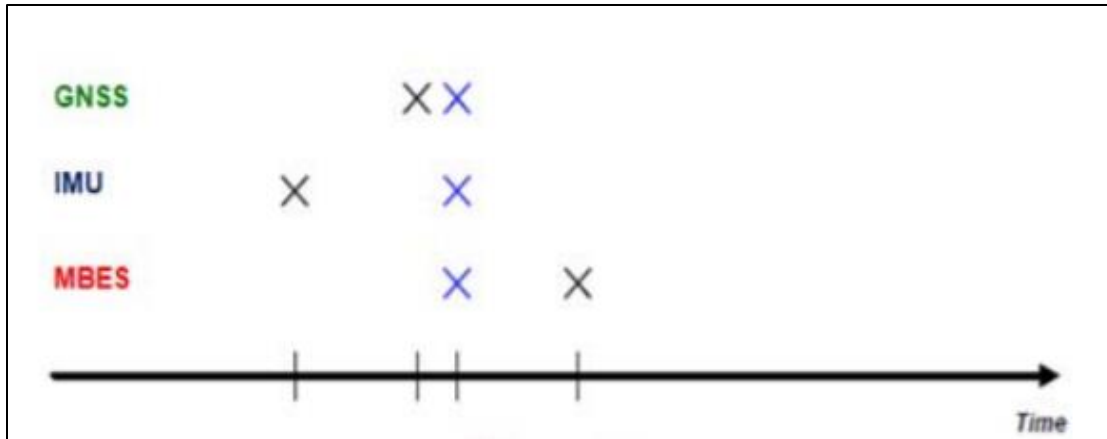
Figure 18. Example of true measurement time vs. recorded time (src: (Naankeu Wati & Seube, 2016))

Unfortunately, latency errors have not been calculated for the HydroBall, however it is expected to be rather insignificant, especially considering this is a system where components have been carefully examined for cohesion. As a result, latency errors will not be considered in the final uncertainty calculation.

In the future, it will be recommended to determine the latency of the system using the methods outlined (International Hydrographic Organization, 2017). The procedure involves running two pairs of survey lines over a sloping seabed at different speeds. The steeper the gradient of the seabed the higher the resolution of the parameter.

3.7.9 SVP Model Uncertainty:

There is an uncertainty introduced into the final horizontal and vertical positioning as a result of sound-velocity profile errors. These occur as a result of inconsistent model derived salinity-depth and model derived temperature-depth profiles when compared to the true profiles. The error between the model derived sound-velocity and the true sound velocity results in an uncertainty

which propagates into an uncertainty of the sound-speed profile. This results in a vertical range uncertainty and if the beam exits on an angle, a horizontal positioning uncertainty.

SVP uncertainties can be classified as an environmental error, which can devour a large portion of the error-budget if not properly managed. Since the sound-velocity profile is needed to calculate distance, both vertically and horizontally, improper measurement of the SVP can result in an improper measurement of both the vertical horizontal positioning, as shown in Figure 19. Vertically, this would mean incorrect placing of the seafloor either falsely too shallow or too deep. Horizontal errors would only occur if the sonar beam has exited on an angle and would result in false horizontal placement of the seafloor due to errors while raytracing with the incorrect sound-velocity.
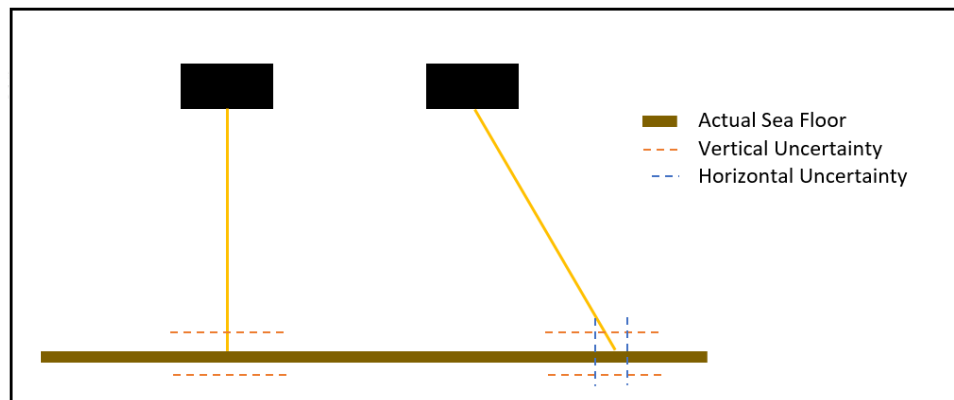
Figure 19. Induces uncertainty from the SVP on both Horizontal and Vertical Components

To take this uncertainty into account, we must compute an associated model error through comparison from a 'truth' (observed) dataset. This dataset will consist of 17 in-situ CTD casts from locations in the Arctic taken by members aboard the CCGS Amundsen in 2017. This data has been filtered to only include depths which match the depth range of the HydroBall (100m). To

make the comparison, HYCOM salinity and temperature data will be queried for the depth ranges at each CTD cast location. The error is then calculated based on the comparison, resulting in an uncertainty as a function of depth, time and space.

While a full HYCOM sound speed uncertainty model is out of the scope of this project, the foundation has been laid for continuation towards this cause by the Ocean Mapping Group of the University of New Brunswick. Currently the model is limited to shallow water depths of 100m or less, however this can be expanded to include all depths and by using a greater amount of data. In this preliminary model, the Canadian Arctic has been partitioned into areas which will serve to provide areas of uncertainty, rather than having a blanketing uncertainty for the entire Canadian Arctic. A regional breakdown of the Arctic can approximate the WebTide model domains for future ease with tidal predictions, see Figure 20. When using this reference, the Canadian Arctic will then be separated into three main locations: Northwest Atlantic, Hudson Bay and Arctic. While it will be important to account for year and seasonality in the HYCOM SVP uncertainty model, for this analysis due to limited data, the time consideration will just be at the yearly level.
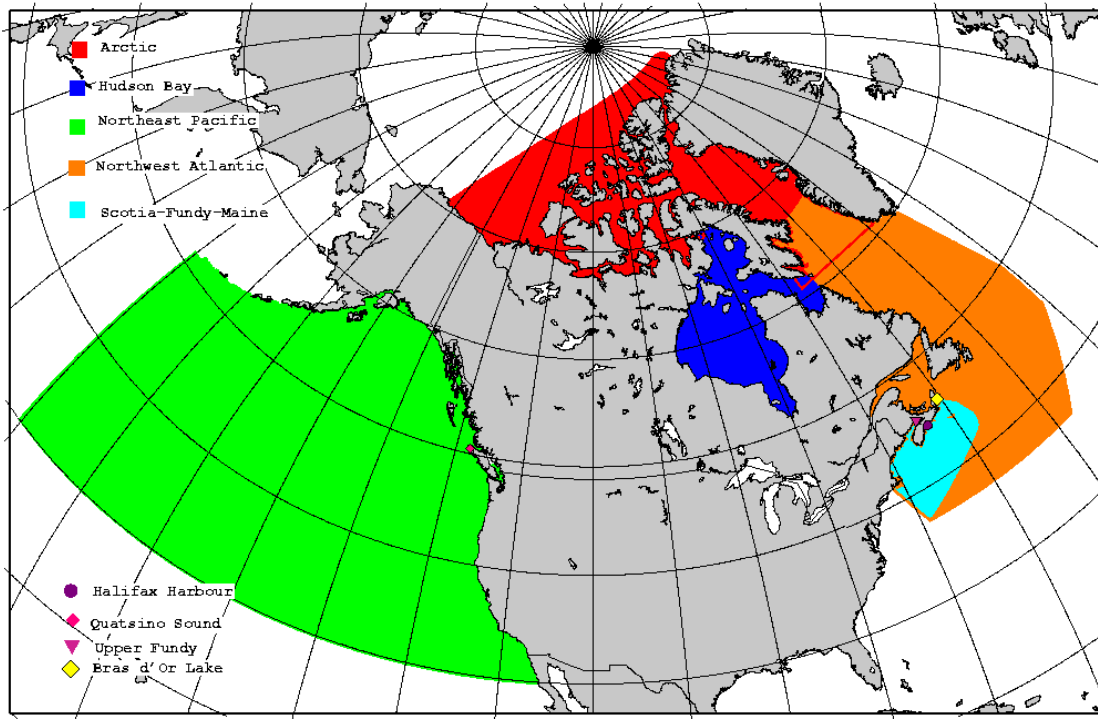
Figure 20. Webtide Model Domain Regions

Following this, the depth uncertainty is calculated through a direct depth-dependent comparison of the Model and 'truth' SVP data for 10m intervals. To do this, the cumulative sound speed is calculated at each depth interval for both model and observed data. The sound speed is then used to raytrace a vertical depth and the error is then calculated based on the depth interval. An average is made for each depth interval of the depth errors of all casts in the interval. This average error across the casts and by depth interval is then used in the final uncertainty calculation as a vertical component.

3.7.10 Sound Speed Uncertainty:

After obtaining salinity, depth/pressure and temperature values either from observations or models, these parameters are used to calculate the sound speed in water. The current UNESCO standard to conduct this conversion is the Thermodynamic Equation of Seawater – 2010 (TEOS10). This equation is based on a Gibbs function formula which is a function of absolute salinity, temperature and pressure (Intergovernmental Oceanographic Commission, 2010).

The maximum error in the equation has been shown to range between +/- 0.3m/s and +/- 0.45m/s, see Figure 21. Using this figure along with error estimated provided by (Allen, Keen, Gardiner, Quartley, & Quartley, 2017), an error rate of +/- 0.35m/s can be safely assumed for the Canadian Arctic. When considering the depth range of the sonar used in this project (100m), an error rate of +/- 0.35m/s would translate into an error of 0.0046m (4.59mm) when using a sound speed of 1460m/s (realistic for the Arctic).
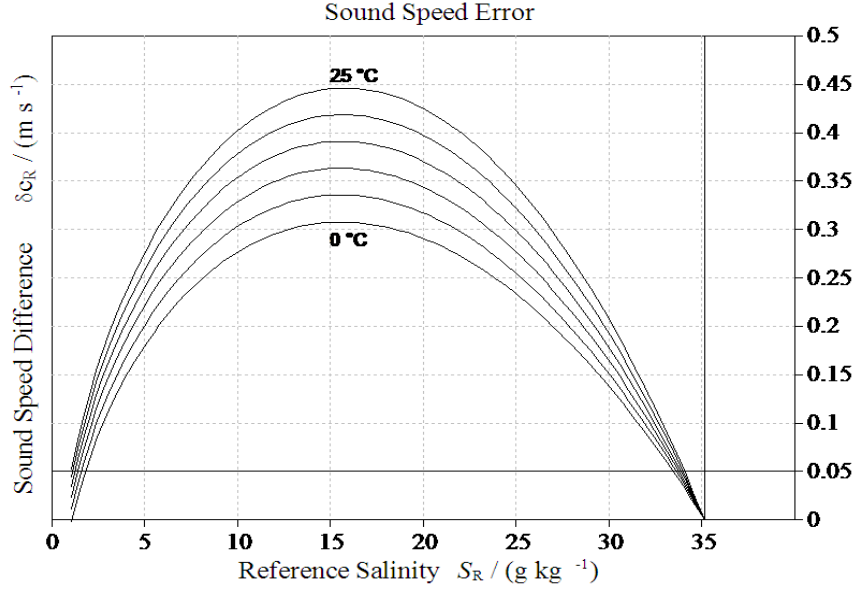
Figure 21. Error in sound speed, if computed from measured Reference Salinity using the Gibbs function for SSW. Results are shown for temperatures between 0 and 25 °C and at atmospheric pressure.(src: (Allen, Keen, Gardiner, Quartley, & Quartley, 2017)

The maximum error comprises a small fraction of the total error budget but would still be used in the final uncertainty calculation in the vertical component.

3.7.11 Total Propagated Uncertainty:

From the sources of error above, the comprehensive total propagated uncertainty calculation for both the horizontal and vertical components would be as follows:

$$\sigma^2_{Horizontal} = \sigma^2_{SatClk} + \sigma^2_{RecClk} + \sigma^2_{SatEph} + \sigma^2_{Ion} + \sigma^2_{Trop} + \sigma^2_{Multi} + \sigma^2_{GNSSHW} + \quad (12)$$
$$\sigma^2_{PPPeng} + \sigma^2_{PPPest} + \sigma^2_{BW} + \sigma^2_{LA} + \sigma^2_{RP} + \sigma^2_{Lat}$$

$$\sigma^2_{Vertical} = \sigma^2_{SatClk} + \sigma^2_{RecClk} + \sigma^2_{SatEph} + \sigma^2_{Ion} + \sigma^2_{Trop} + \sigma^2_{Multi} + \sigma^2_{GNSSHW} + \quad (13)$$
$$\sigma^2_{PPPeng} + \sigma^2_{PPPest} + \sigma^2_{GeoRan} + \sigma^2_{RayTr} + \sigma^2_{LA} + \sigma^2_{Red} + \sigma^2_{DynDr} + \sigma^2_{RP} + \sigma^2_{SVP}$$
$$+ \sigma^2_{SS}$$

Where:

σ^2_{SatClk} = satellite clock uncertainty σ^2_{LA} = lever arm uncertainty

σ^2_{RecClk} = receiver clock uncertainty σ^2_{RP} = roll and pitch uncertainty

σ^2_{SatEph} = satellite ephemeris uncertainty σ^2_{Lat} = latency uncertainty

σ^2_{Ion} = ionospheric uncertainty σ^2_{GeoRan} = geometric range uncertainty

σ^2_{Ion} = tropospheric uncertainty σ^2_{RayTr} = raytracing uncertainty

σ^2_{Multi} = multipath uncertainty σ^2_{Red} = reduction uncertainty

σ^2_{GNSSHW} = GNSS hardware uncertainty σ^2_{DynDr} = dynamic draft

σ^2_{PPPeng} = PPP engine uncertainty σ^2_{SVP} = sound-velocity profile uncertainty

σ^2_{PPPest} = PPP estimated uncertainty σ^2_{SS} = sound speed uncertainty

σ^2_{BW} = beam width uncertainty

The sources of uncertainty mitigated during other processes or omitted are shown below in strikethrough:

$$\sigma^2_{Horizontal} = \cancel{\sigma^2_{SatClk}} + \cancel{\sigma^2_{RecClk}} + \cancel{\sigma^2_{SatEph}} + \cancel{\sigma^2_{Ion}} + \cancel{\sigma^2_{Trop}} + \cancel{\sigma^2_{Multi}} + \cancel{\sigma^2_{GNSSHW}} + \quad (14)$$
$$\cancel{\sigma^2_{PPPeng}} + \sigma^2_{PPPest} + \sigma^2_{BW}* + \sigma^2_{LA} + \sigma^2_{RP} + \cancel{\sigma^2_{Lat}}$$

$$\sigma^2_{Vertical} = \cancel{\sigma^2_{SatClk}} + \cancel{\sigma^2_{RecClk}} + \cancel{\sigma^2_{SatEph}} + \cancel{\sigma^2_{Ion}} + \cancel{\sigma^2_{Trop}} + \cancel{\sigma^2_{Multi}} + \cancel{\sigma^2_{GNSSHW}} + \quad (15)$$
$$\cancel{\sigma^2_{PPPeng}} + \sigma^2_{PPPest} + \sigma^2_{GeoRan} + \cancel{\sigma^2_{RayTr}} + \sigma^2_{LA} + \cancel{\sigma^2_{Red}} + \cancel{\sigma^2_{DynDr}} + \sigma^2_{RP} + \sigma^2_{SVP}$$
$$+ \sigma^2_{SS}$$

This cleans up to the following final uncertainty calculation for both the horizontal and vertical components:

$$\sigma^2_{Horizontal} = \sigma^2_{PPPest} + \sigma^2_{BW}* + \sigma^2_{LA} + \sigma^2_{RP} \qquad (16)$$

$$\sigma^2_{Vertical} = \sigma^2_{PPPest} + \sigma^2_{GeoRan} + \sigma^2_{LA} + \sigma^2_{RP} + \sigma^2_{SVP} + \sigma^2_{SS} \qquad (17)$$

Note that the beam width uncertainty is marked by an Asterix. This is because this uncertainty will be included for demonstration purposes in this thesis only.

3.8   Metadata

To achieve transparency in the data, metadata is included with the final depth soundings. This ensures that interested parties are able to view information regarding the hardware used, data acquisition and data processing. The metadata also serves as a platform to communicate the data uncertainties and error at each step of processing. Certain fields are user-generated upon data submission such as the information of the participant including the type of vessel used to collect data, usage of catamaran or not and the participant's ID. Upon data submission certain fields are auto-generated such as the unique identifier for the HydroBall, the make and models of sensors, sensor offsets and date and time information. Additionally, some metadata is completed during processing such as the processing status, methods of processing, reference datums and derived uncertainties. The metadata that UNB is responsible for is included in Table 9 below.

Table 9. UNB Responsible Metadata Fields

STATUS	Status of data post-processing
Geodetic reference system	Reference system used
GNSS uncertainty (acquisition)	Asses if GNSS quality meets requirements
GNSS uncertainty (PPP)	Asses the quality of GNSS processing
Tide (model)	Chosen tide model
Tide (Uncertainty)	Asses the uncertainty of the tide model

Reduction method	ERS/WLRS
SVP (model)	Chosen model
SVP (uncertainty)	Uncertainty of SVP model
Ray tracing method	Method used
Data cleaning	Method used
Final georeferencing	sounding uncertainty, explanation of uncertainty, variance per depth

4   Results

The results and related discussions for the GNSS processing, GNSS qualification, SVP and Uncertainty modules are outlined below.

4.1   GNSS Module

While many studies exist on the benefits of additional processing on GNSS readings (McHugh, Church, Kim, & Maggio, 2015), (Rizos, Janssen, Roberts, & Grinter, PPP versus DGNSS, 2012), (Rizos, Janssen, Roberts, & Grinter, 2012) something that had not been investigated thoroughly was the benefits or drawbacks on using a single constellation (GPS) versus using multiple constellations (GPS + GLONASS) in the Arctic. To make this assessment, 22 raw GNSS datasets were obtained from on-site recordings in the Arctic collected by the HydroBall. The locations ranged from Gjoa Haven, Baffin Island to Greenland, as shown in Figure 22. The data was collected between September 2017 and October 2018 and includes only kinematic GNSS data, see Table 10. The data was processed according to the discussed workflow with the use of Natural Resources Canada's PPP engine and then sorted using the qualification scripts.

1.	Gjoa Haven, Nunavut
2.	Davis Straight, Nunavut
3.	Davis Straight, Greenland
4.	Iqaluit, Nunavut
5.	Quaqtaq, Quebec

Figure 22. Location of GNSS Dataset

Table 10. Arctic Dataset Metadata

File:	Start	Finish	Duration	Region	Latitude	Longitude	Static/Kinematic	IGS Product
13_43_52-2018_10_11-gps	11/10/2018 13:44	11/10/2018 13:54	0:10:12	Iqaluit, NU	63.7257	-68.6516	Kinematic	Final
13_46_34-2017_09_22-gps	22/09/2017 13:46:55	22/09/2017 19:42:13	5:55:18	Quaqtaq, QU	61.0176	-68.6516	Kinematic	Final
13_59_50-2018_10_11-gps	11/10/2018 14:00	11/10/2018 15:48	1:48:31	Iqaluit, NU	63.7214	-68.6516	Kinematic	Final
14_07_23-2018_10_12-gps	12/10/2018 14:07	12/10/2018 19:45	5:37:47	Iqaluit, NU	63.7071	-68.6516	Kinematic	Final
14_41_53-2018_10_15-gps	15/10/2018 14:42:14	15/10/2018 14:45:01	0:02:47	Iqaluit, NU	63.7641	-68.6516	Kinematic	Final
14_45_26-2018_10_15-gps	15/10/2018 14:45:45	15/10/2018 18:59:55	4:14:10	Iqaluit, NU	63.7535	-68.6516	Kinematic	Final
15_11_52-2018_10_30-gps	30/10/2018 15:12:15	30/10/2018 21:01:00	5:48:45	Iqaluit, NU	63.7058	-68.4219	Kinematic	Final
15_16_59-2018_10_17-gps	17/10/2018 15:17:20	17/10/2018 15:32:30	0:15:10	Iqaluit, NU	63.7368	-68.6785	Kinematic	Final
15_32_51-2018_10_17-gps	17/10/2018 15:33:10	17/10/2018 19:04:35	3:31:25	Iqaluit, NU	63.7074	-68.7022	Kinematic	Final
15_40_04-2018_10_14-gps	14/10/2018 15:40:27	14/10/2018 18:38:22	2:57:55	Davis Strait, NU	68.6240	-68.6516	Kinematic	Final
15_41_04-2018_10_21-gps	21/10/2018 15:41:27	21/10/2018 15:44:59	0:03:32	Iqaluit, NU	63.7178	-68.4433	Kinematic	Final
15_45_32-2018_10_21-gps	21/10/2018 15:45:52	21/10/2018 15:49:55	0:04:03	Iqaluit, NU	63.7172	-68.4421	Kinematic	Final
15_55_37-2018_10_21-gps	21/10/2018 15:55:57	21/10/2018 21:00:19	5:04:22	Iqaluit, NU	63.7056	-68.4328	Kinematic	Final
16_32_34-2018_08_02-gps	2/8/2018 16:33	2/8/2018 17:04	0:31:52	Gjoa Haven, NU	68.6240	-95.8763	Kinematic	Final
17_05_13-2018_08_02-gps	2/8/2018 17:05	2/8/2018 17:13	0:07:55	Gjoa Haven, NU	68.6277	-95.8763	Kinematic	Final
17_24_50-2018_08_02-gps	2/8/2018 17:25	2/8/2018 17:58	0:33:05	Gjoa Haven, NU	68.6277	-95.8854	Kinematic	Final
18_37_04-2017_12_11-gps	11/12/2017 18:37	11/12/2017 18:53	0:16:20	Davis Strait, Greenland	68.6277	-53.1342	Kinematic	Final
18_58_37-2017_12_11-gps	11/12/2017 18:58	11/12/2017 19:09	0:10:02	Davis Strait, Greenland	68.6277	-53.1330	Kinematic	Final
19_20_43-2018_08_02-gps	2/8/2018 19:21	2/8/2018 19:49	0:28:42	Gjoa Haven, NU	68.6277	-95.8876	Kinematic	Final
20_00_26-2018_08_02-gps	8/8/2018 20:00	8/8/2018 20:17	0:17:06	Gjoa Haven, NU	68.6277	-95.8778	Kinematic	Final
20_28_02-2018_08_02-gps	8/8/2018 20:28	8/8/2018 20:45	0:16:37	Gjoa Haven, NU	68.6282	-95.8778	Kinematic	Final
20_39_25-2018_10_15-gps	15/10/2018 20:39:45	15/10/2018 23:12:21	2:32:36	Iqaluit, NU	63.7403	-68.6769	Kinematic	Final

Once processed, each of the 22 datasets were summarized into a mean and standard deviation for each file. The mean and standard deviations of each of these 22 datasets were then used to create a single mean and standard deviation for each of the qualification metrics. The findings are shown below in Table 11. We can see from the table that on average the NSV and GDOP values are more favourable when using both constellations, as expected. For the most part the rest of the variables are consistent amongst both options, with the exception of the clock standard deviation, showing variance.

Table 11. Summarized Qualification Metrics

	Number Satellites Visible (NSV)		Geometric Dilution of Precision		Standard Deviation LAT(95%) (m)	
	GPS+GLONASS	GPS Only	GPS+GLONASS	GPS Only	GPS+GLONASS	GPS Only
Mean	12.54	8.84	1.90	2.45	0.17	0.19
Stdev	1.28	0.89	0.25	0.45	0.13	0.14
Max	15.39	10.74	2.50	4.04	0.50	0.52
Min	10.42	7.32	1.55	1.82	0.05	0.04

	Standard Deviation LON(95%) (m)		Standard Deviation HGT(95%) (m)		Standard Deviation CLK(95%) (ns)	
	GPS+GLONASS	GPS Only	GPS+GLONASS	GPS Only	GPS+GLONASS	GPS Only
Mean	0.17	0.17	0.32	0.34	1.22	0.94
Stdev	0.11	0.11	0.24	0.27	1.28	0.67
Max	0.36	0.40	0.87	1.00	5.15	2.38
Min	0.04	0.03	0.08	0.08	0.24	0.22

A chart was created to summarize the qualification parameters of location (standard deviation of latitude, standard deviation of longitude and standard deviation of height) for all datasets in Figure 23. We can see again that the data is relatively the same amongst the two. We do see a higher maximum for the standard deviation of height with the GPS only measurement.
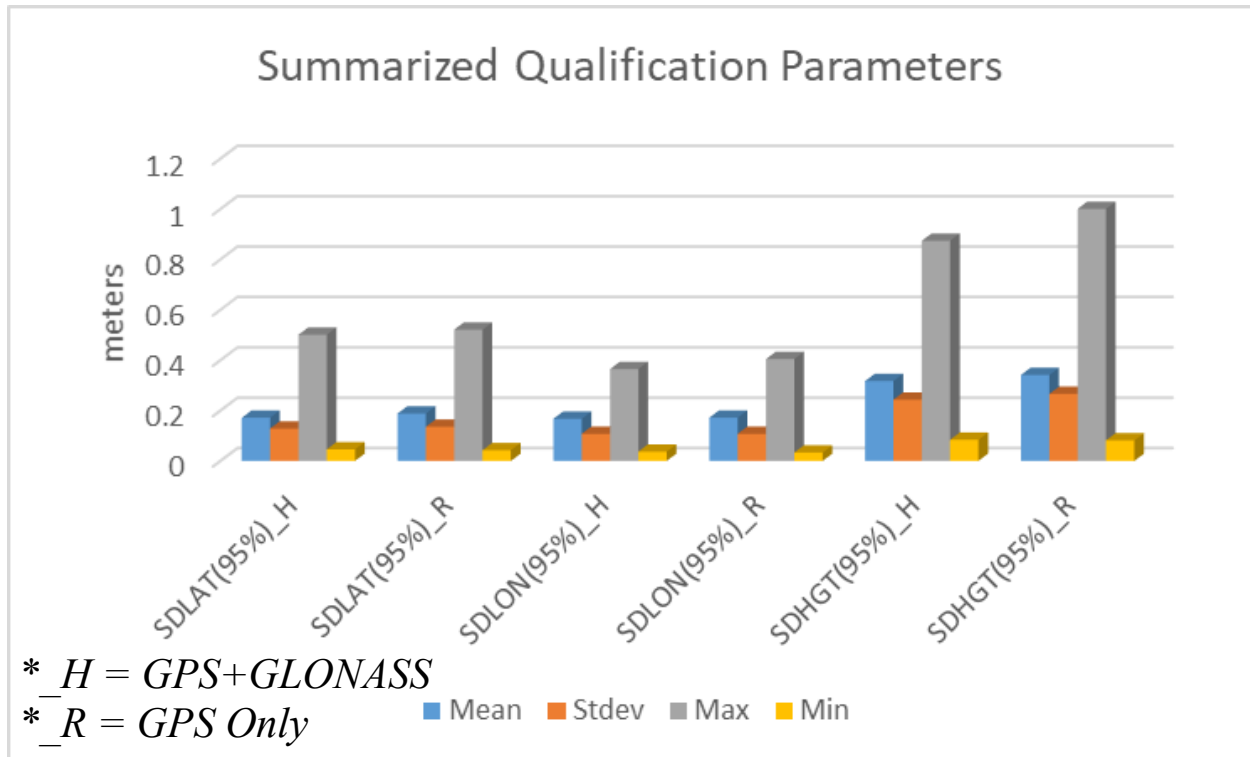
Figure 23. Summarized Qualification Parameters

Both datasets were then taken and processed and analyzed using the qualification scripts and a final pass rate for all the lines in the set was created. This is shown in Figure 24, where a number of interesting findings can be inferred. Overall all location values are favorable with the use of GLONASS in addition to GPS. This did come at the expense of an increase in the clock standard deviations. This is important because the error of the receiver's clock will be reflected in the computed range and may negatively influence the final location values (Misra, 1996). Any possible gain in the latitude, longitude and height measurements may be negated by a poor clock measurement, in some cases. A slightly higher pass rate was seen when using both constellations but not by a significant amount. On days where the GPS signal was poor (ie. 15_41_04-2018_10_21-gps &15_45_32-2018_10_21-gps), the qualification rate was 0% in the GPS only

data. However, when adding GLONASS support observed the overall qualification rate increased to 88% and 96% respectively, allowing us to use this data. For this factor alone, it may be a better choice to use both constellations. The final average pass rate for all data was 84.6% for GPS and GLONASS observations and 74.7% for GPS Only.
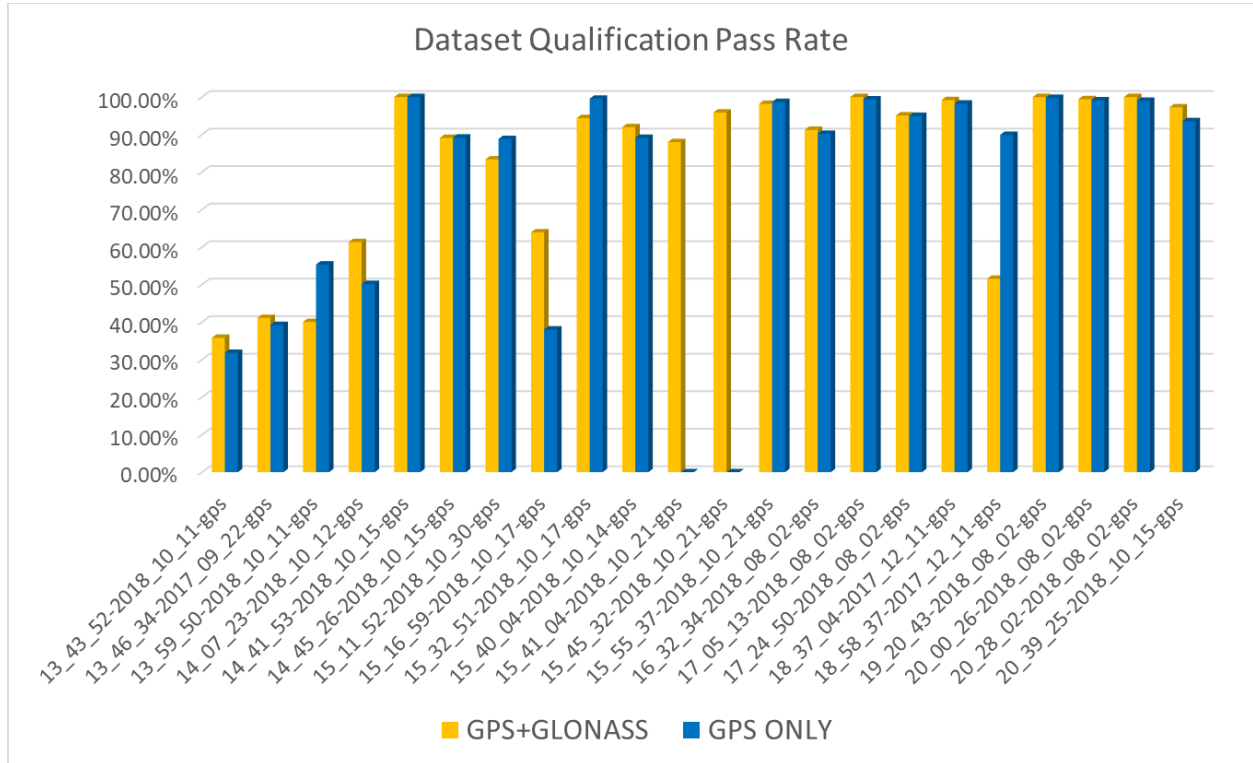
Figure 24. Dataset Qualification Pass Rate

4.2   Sound Velocity Profile Module

To help assess the accuracy of the HYCOM model data, real CTD casts were gathered from the Canadian Coast Guard vessel, the Amundsen. The data used was collected in various locations across the Canadian Arctic in the year 2017. The observed CTD casts were filtered to include only data with depths of 100m or less as this is the vertical limit of the hardware employed in this

project. In total there are 17 CTD casts which serve as the 'truth'. The locations of these casts are visualized in Figure 25 along with metadata information. The collective temperature and salinity graphs are shown in Figure 26, where each cast is colour coded and the location shown. Comparable HYCOM data was collected based on the time, date and location of the CTD cast. All data has been formatted into three fields: depth (m), temperature (^{o}C) and salinity (ppt). A comparison of the Salinity, Temperature and Depth data was made through correlation, graph visualization and differential comparison. Following this, the data was converted to SVPs and a further comparison was made on the depth error when used to raytrace.

#	Cruise+Cast	Date	Depth	Latitude	Longitude
1	1702-086	8/9/2017 0:31	55	68.4862	-103.429
2	1702-087	8/9/2017 5:23	41	68.3279	-102.94
3	1702-085	8/8/2017 12:48	91	68.2471	-101.811
4	1702-088	8/9/2017 12:09	95	68.2426	-101.795
5	1702-083	8/8/2017 6:06	100	68.3028	-101.745
6	1702-084	8/8/2017 9:08	97	68.3027	-101.738
7	1702-090	8/9/2017 22:14	81	68.3061	-100.802
8	1702-091	8/10/2017 5:45	53	69.1707	-100.705
9	1702-092	8/10/2017 9:22	52	69.1691	-100.695
10	1702-089	8/9/2017 17:03	28	68.4906	-99.8955
11	1702-096	8/15/2017 7:17	89	65.451	-83.2584
12	1702-095	8/13/2017 19:14	16	68.7668	-80.841
13	1702-094	8/13/2017 16:17	42	69.2716	-80.6078
14	1702-002	7/8/2017 22:03	76	58.4258	-78.3035
15	1702-005	7/10/2017 21:00	92	62.3669	-74.662
16	1702-006	7/11/2017 9:09	92	61.7868	-71.9121
17	1702-007	7/11/2017 19:31	94	61.0419	-69.7163

Figure 25. Locations of CTD Casts from the Amundsen

82
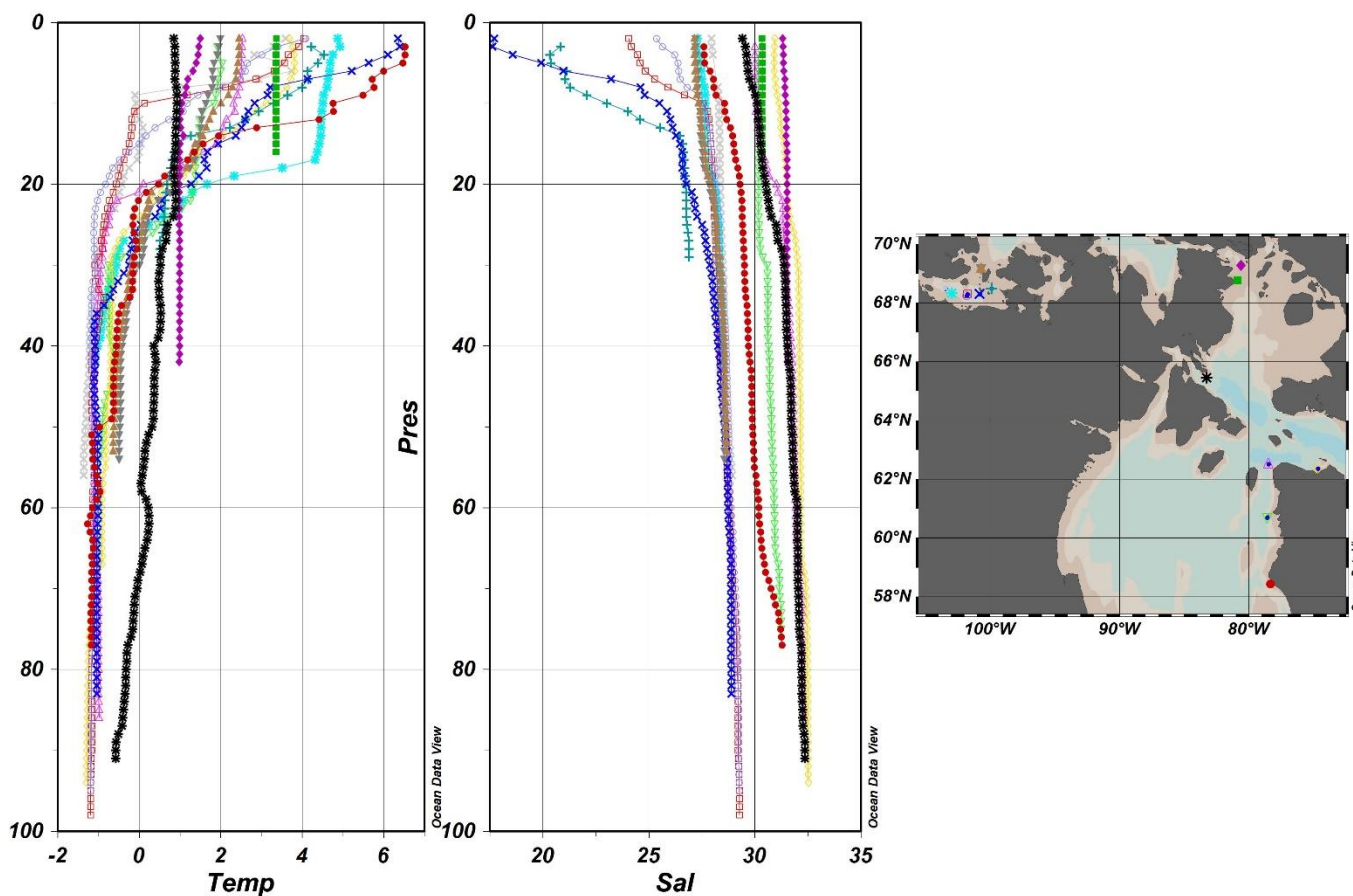
Figure 26. Combined Temperature-Pressure and Salinity-Pressure graphs

After the 17 casts of observed data had been collected and formatted, HYCOM model queries were created to obtain model casts for the same locations and time. To be able to compare the two datasets, some modifications needed to be made. The observed casts have a vertical resolution of 1m while the model cast's vertical resolution is more course at a range of 2m up to 10m intervals. To make a fair comparison, the model data was converted to the 1m scale keeping the value constant for the range. Additionally, NaN values in the observed dataset were ignored if occurring as the first or last value in the casts. In the 1702_086 dataset a series of NaN values were observed

in the salinity readings between 5-7 meters. In this case the NaN values were replaced with a linear average between the 4 and 8 meter values. Finally, when the length of the dataset did not agree between the observed and model datasets, the last model value was held constant and extended to the length of the observed cast. In the case where the model cast length was extended beyond the model, the extra data was ignored.

Now the two datasets could be compared and to do so a Pearson Product-Moment Correlation (r) value was calculated to compare the salinity and temperature values across each dataset, as shown in Table 12 and Figure 27. From the table we can see that the maximum correlation for temperature and salinity is 0.94 and 0.96, the minimum correlation is -0.90 and -0.96 and the average correlation is 0.44 and 0.57 respectively. We can see that salinity tends to be better correlated between the observed and model data than temperature. There were three temperature datasets with negative correlations, two salinity datasets with negative correlation and one salinity dataset with no correlation.

It is expected that the correlation between the two datasets will have a weak to strong positive correlation and the average values meet these predictions. What was not predicted was the negative correlation between datasets, especially present in the 1702-002 and 1702-089 datasets. When examining these values, it appears that the model values for these two datasets extend straight down, when compared to a wider ranged value for the observed data. A comparison of the salinity and temperature for both datasets comprehensively are shown in Figure 28. In either case the correlation values did not show any noticeable pattern in location or depth, as both shallow and deep casts had a mix of high and low correlation values.

Table 12. Observed vs. Model Temperature and Salinity Correlation Values for Each Dataset

Dataset	Correlation Temperature	Correlation Salinity
1702-002	-0.551	-0.815
1702-005	0.870	0.928
1702-006	0.709	0.783
1702-007	0.867	0.881
1702-083	0.605	0.770
1702-084	0.553	0.758
1702-085	0.546	0.805
1702-086	0.583	0.905
1702-087	0.805	0.964
1702-088	0.561	0.800
1702-089	-0.895	-0.957
1702-090	0.723	0.672
1702-091	0.900	0.944
1702-092	0.856	0.934
1702-094	-0.403	0.447
1702-095	-0.231	-0.010
1702-096	0.942	0.943
Max	0.942	0.964
Min	-0.895	-0.957
Mean	0.438	0.574
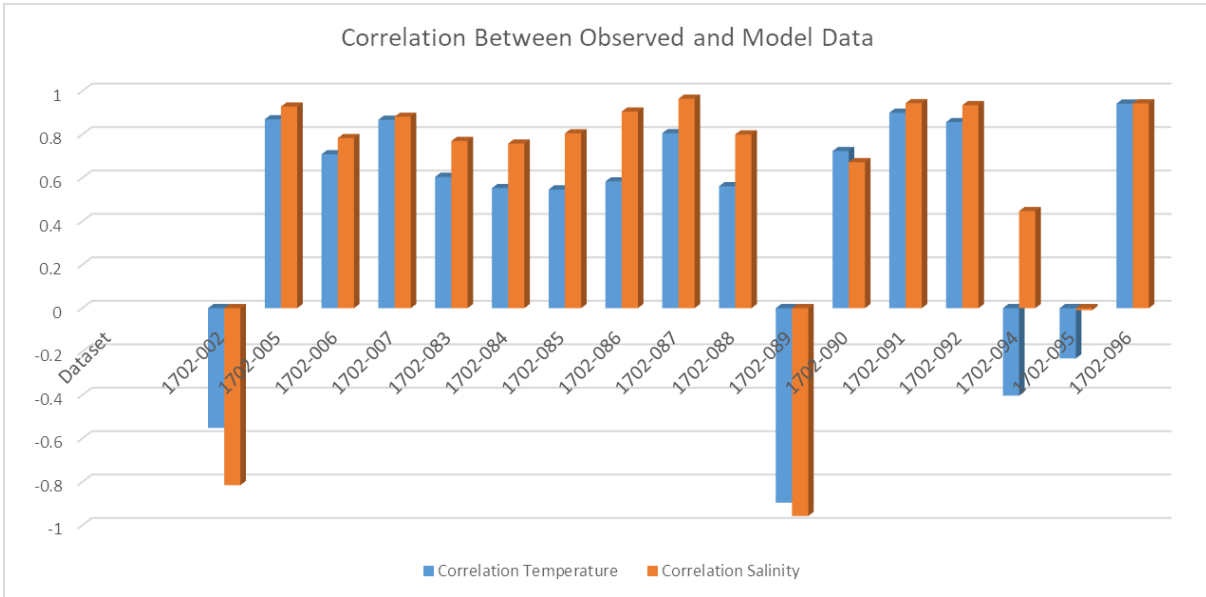
Figure 27. Correlation between Observed and Model Temperature and Salinity Values
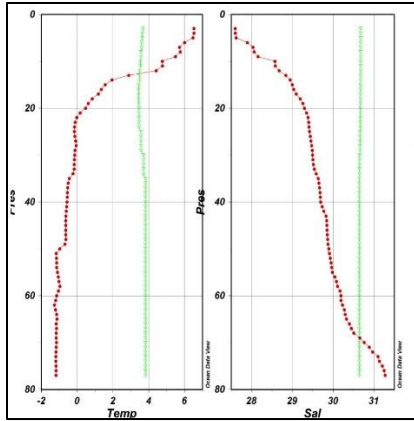
1702-002
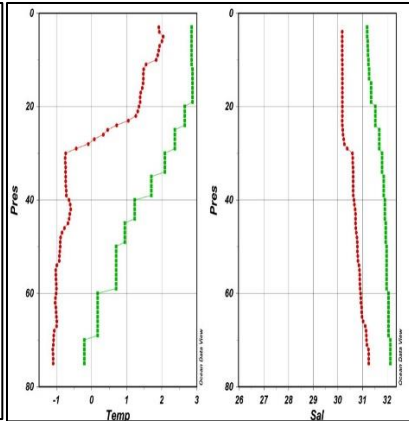
1702-005
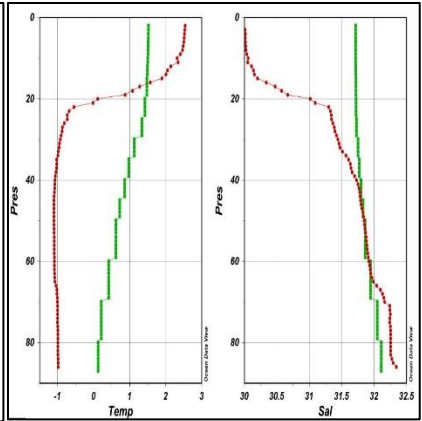
1702-006
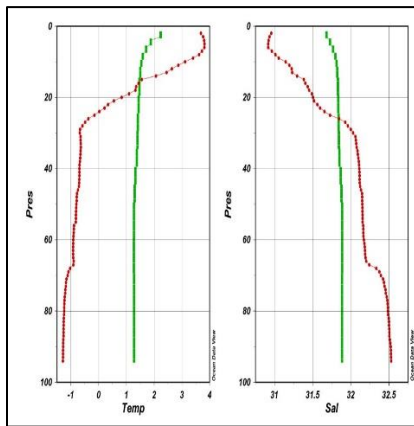
1702-007
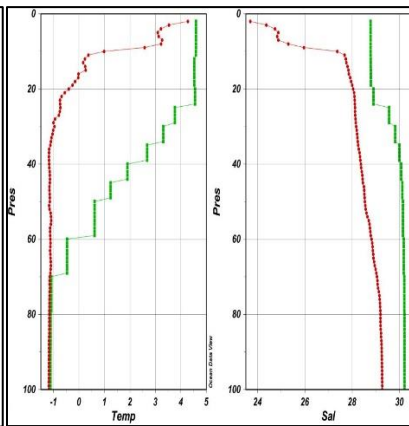
1702-083
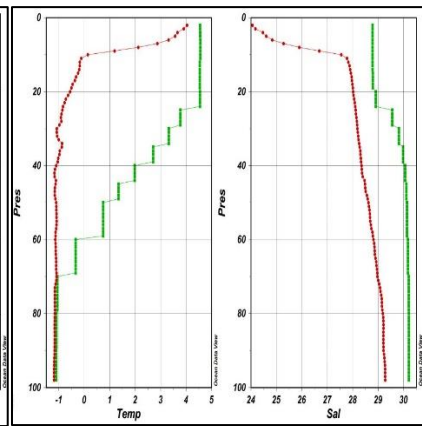
1702-084
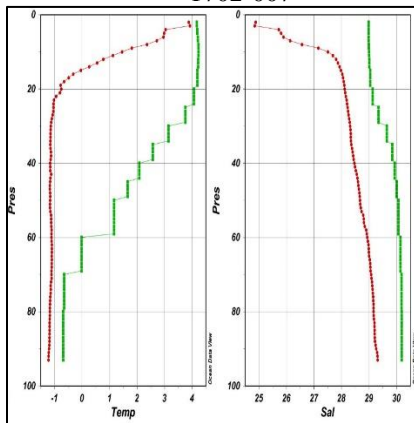
1702-085

1702-086

1702-087

1702-088 1702-089 1702-090

1702-091 1702-092 1702-094

1702-095 1702-096
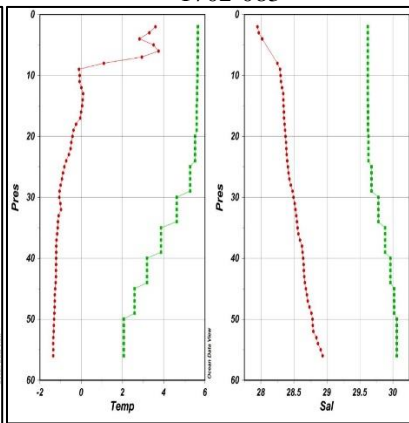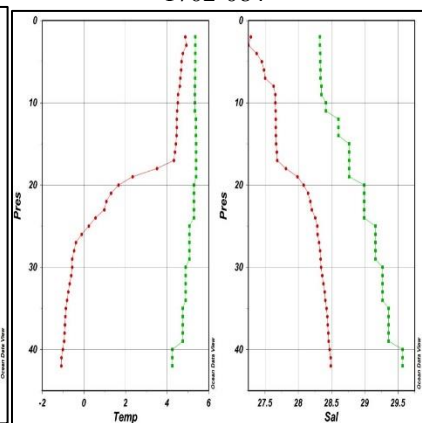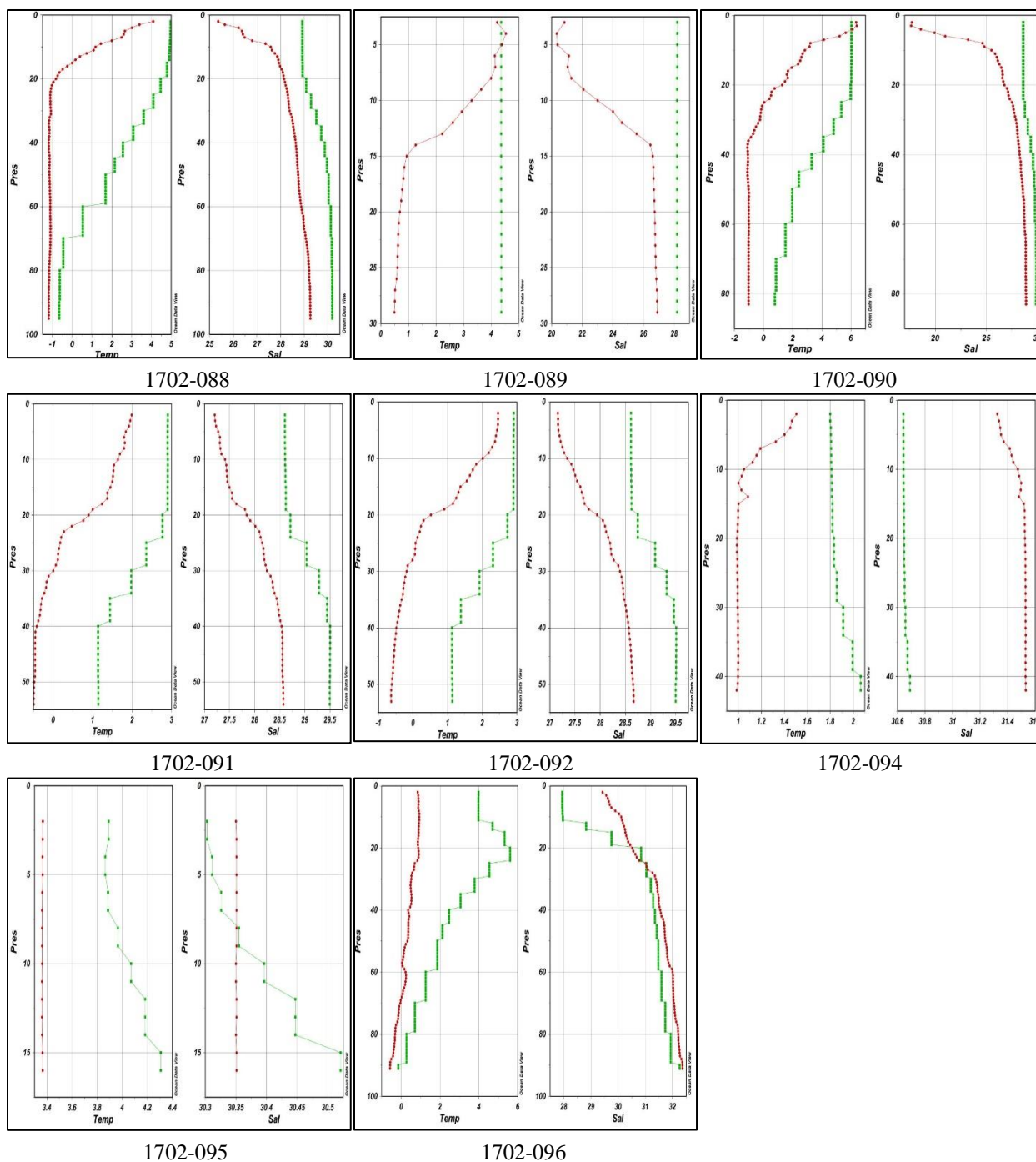
Figure 28. Temperature and Salinity Graphs for CTD Casts (Red = Cast; Green = HYCOM Model)

To further the analysis, the differences between the observed values and model values were calculated. Table 13 shows the results per dataset. Here we can see that the mean temperature difference across all 17 datasets is -2.24°C with a standard deviation of 1.74 °C and the mean salinity difference across all datasets is -0.91 ppt with a standard deviation of 1.06 ppt.

Table 13. Temperature and Salinity Differences between Observed and Model Salinity and Temperature Values

Dataset	Δ Temperature (oC)		Δ Salinity (ppu)	
	Mean	Stdev	Mean	Stdev
1702-002	-3.374	2.340	-0.996	0.855
1702-005	-1.613	0.574	-1.147	0.137
1702-006	-1.186	1.004	-0.344	0.645
1702-007	-1.529	1.390	0.147	0.445
1702-083	-1.992	1.845	-1.461	0.814
1702-084	-2.126	1.899	-1.447	0.793
1702-085	-2.224	1.669	-1.311	0.607
1702-086	-4.755	1.269	-1.302	0.106
1702-087	-3.350	2.210	-0.903	0.110
1702-088	-2.748	1.754	-1.198	0.476
1702-089	-2.463	1.559	-3.301	2.537
1702-090	-2.915	1.798	-1.680	2.097
1702-091	-1.682	0.404	-1.003	0.184
1702-092	-1.651	0.605	-0.980	0.222
1702-094	-0.809	0.195	0.839	0.058
1702-095	-0.672	0.163	-0.034	0.077
1702-096	-2.231	1.326	0.499	0.548
Mean(P)	**-2.244**		**-0.915**	
Stdev(P)	**1.737**		**1.057**	

To get a depth dependent difference in temperature and salinity, values for all casts were grouped together in 10m intervals. Since not all casts were equal in depth, a frequency distribution graph

88

was created to show the number of values per depth range in Figure 29. The box and whisker plot

was created to depict the differences between the observed and model temperature and salinity

values were then calculated and shown in Figure 30 and Figure 31. Here we can see the largest

disparity occurring between 20-39m depth range for temperature and 0-9m depth range for salinity.

Both salinity and temperature have a large spread of values with many outliers, probably as surface

water is difficult to predict due to large amounts variable environmental conditions causing

mixing.
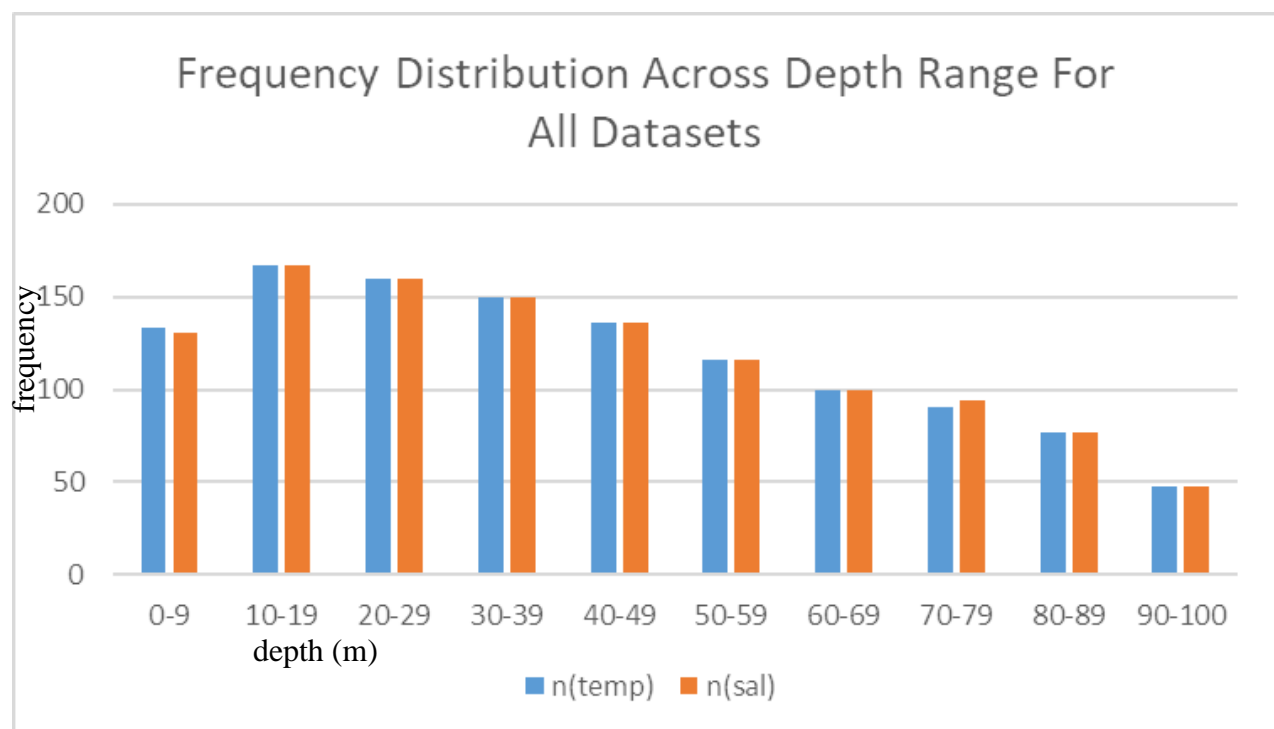
Figure 29. Frequency Distribution Across Depth Range for All Datasets
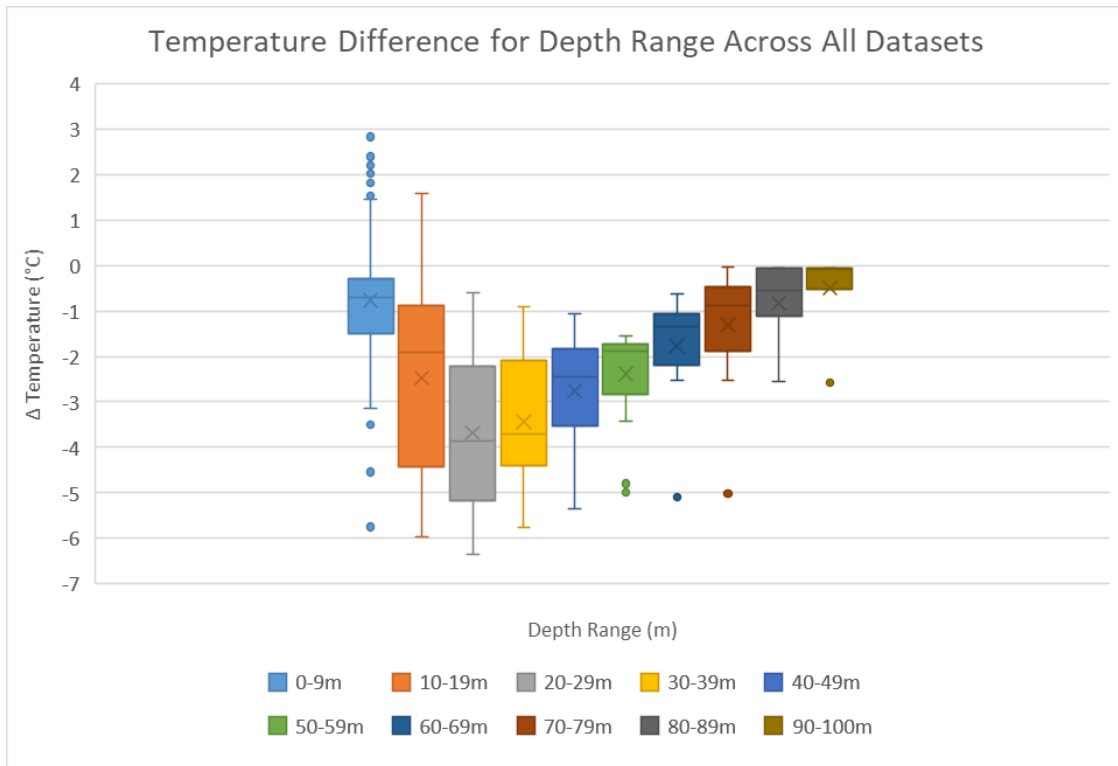
Figure 30. Box and Whisker Plot Depicting Temperature Difference Between Observed and Model Data at Depth Range Interval
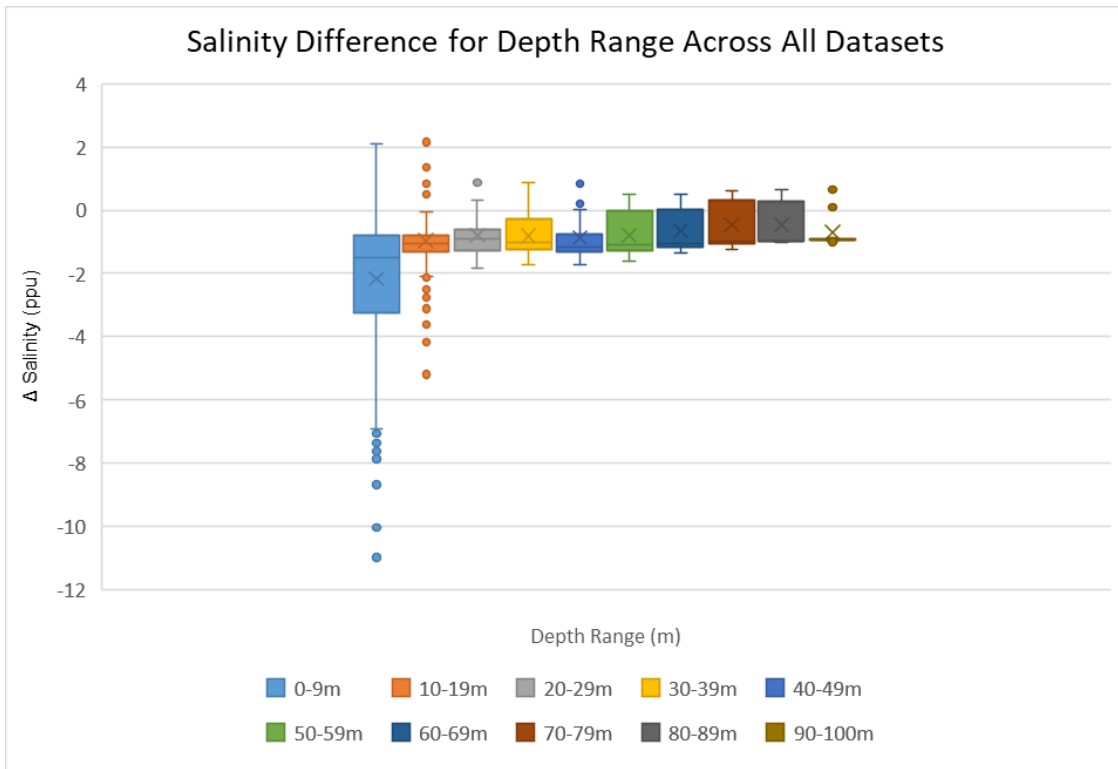
Figure 31. Box and Whisker Plot Depicting Salinity Difference Between Observed and Model Data at Depth Range Interval

The salinity and temperature values were then used to create an SVP. The TEOS-10 sound speed algorithm is the accepted standard, however, there exists several other methods to compute sound velocity (ie. Mackenzie, Del-Grosso, Chen&Millero). A quick comparison between the algorithms is given in Figure 32 and Figure 33, depicting the differences in m/s between TEOS-10 for both observed and model datasets. We can see that overall the Mackenzie calculation yields a consistent lower sound speed velocity and the Chen & Millero yields a higher sound speed velocity when compared to TEOS-10. Overall, the difference is mostly negligible as the largest difference is 0.32 m/s.
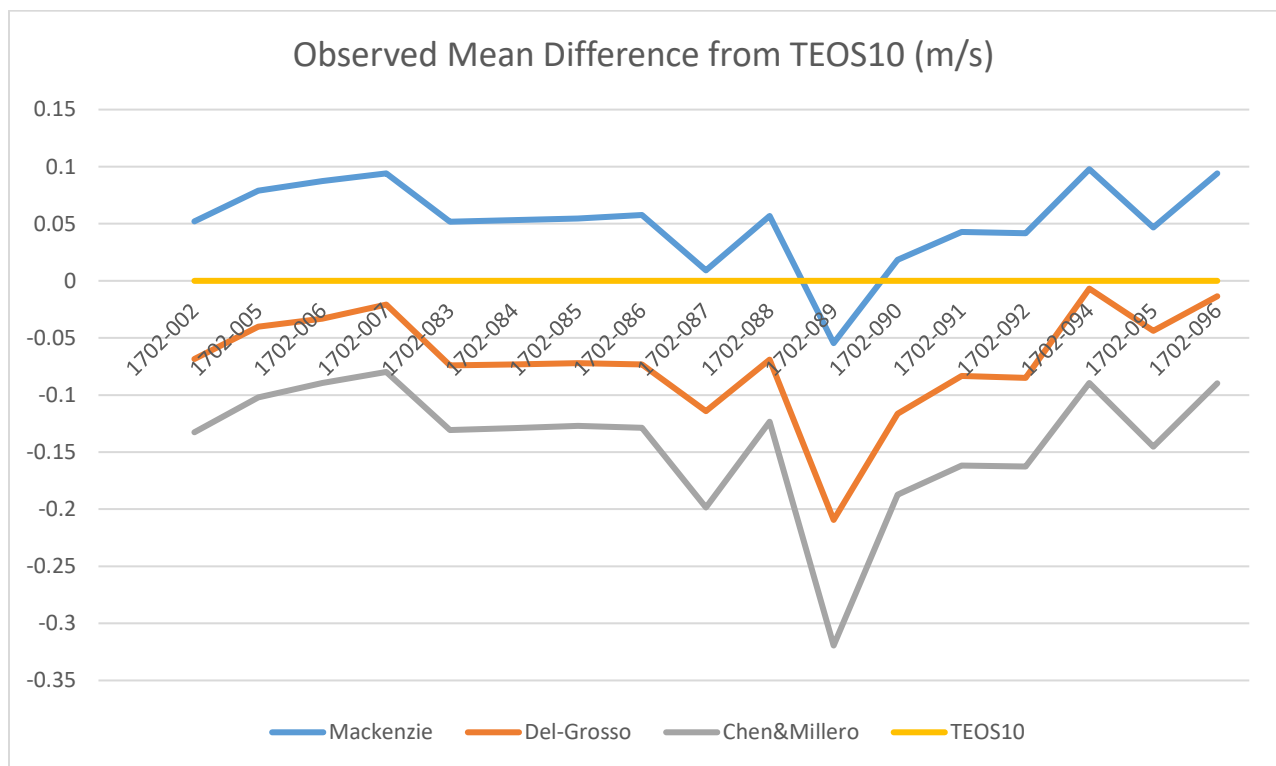
Figure 32. Observed Mean Sound Velocity Difference From TEOS-10
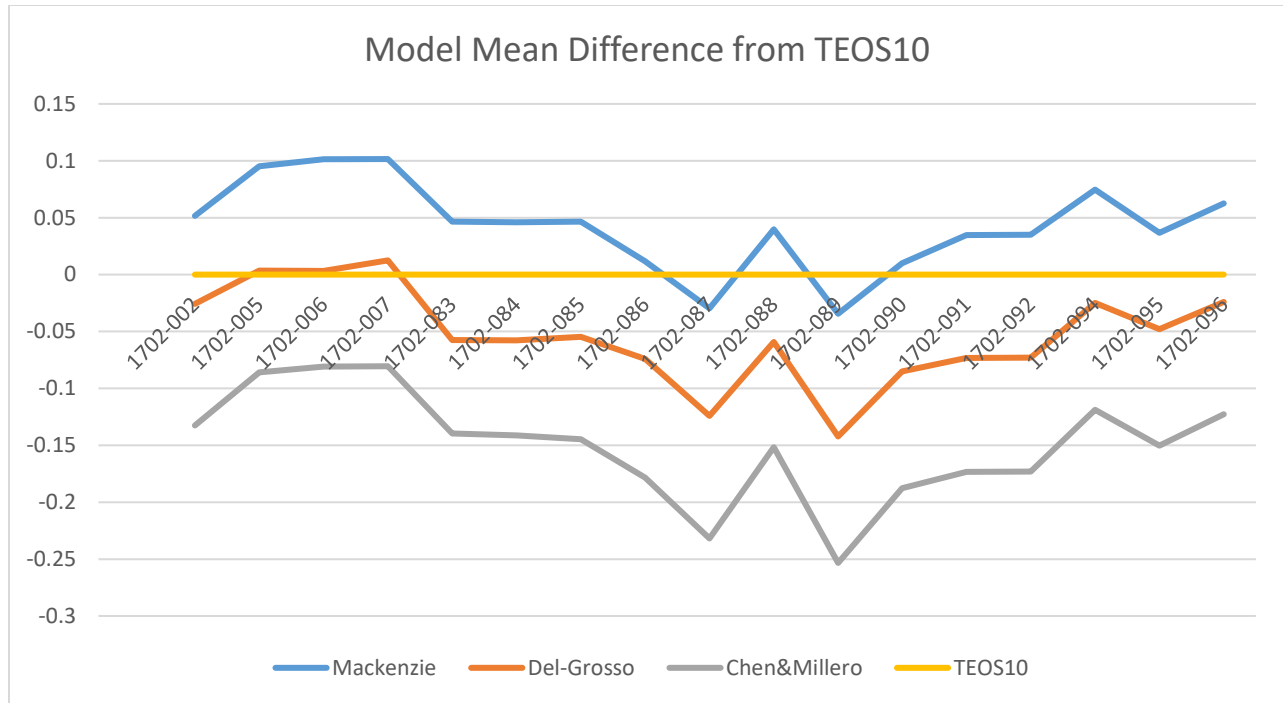
Figure 33. Model Mean Sound Velocity Difference From TEOS-10

To then finalize the comparison, raytracing was performed between the datasets to calculate the depth error that would occur between the two datasets. Since a single-beam echosounder is used in this project, raytracing was done at nadir using the TEOS-10 SVP. The travel-time was computed for the observed SVP and then using that travel time a depth was computed using the model SVP. The results of this comparison are shown in Table 14. It is shown that the maximum depth difference between the datasets is 1m, the minimum depth difference is 0.03m and the average depth difference is 0.55m. This translates into a maximum, minimum and average depth error percentage of 1.55%, 0.168% and 0.76% respectively.

Table 14. Raytracing Results Comparison Between Observed and Model Depth Differences

| Dataset | Depth Error (Obs vs. Model) | | | | | | |
	obs_depth (m)	obs_owtt (s)	sim_depth (m)	depth_diff (m)	error%	lat	long
1702-002	77	0.053	77.850	0.850	1.092	58.4258	-78.3035
1702-005	75	0.052	75.457	0.457	0.606	60.6961	-78.5643
1702-006	86	0.060	86.353	0.353	0.409	62.5129	-78.4889
1702-007	94	0.065	94.442	0.442	0.468	62.3669	-74.662
1702-083	100	0.070	100.758	0.758	0.753	68.3028	-101.7446
1702-084	98	0.068	98.784	0.784	0.793	68.3027	-101.7378
1702-085	93	0.065	93.798	0.798	0.851	68.2471	-101.8109
1702-086	56	0.039	56.879	0.879	1.546	68.4862	-103.4292
1702-087	42	0.029	42.453	0.453	1.067	68.3279	-102.9396
1702-088	95	0.066	95.921	0.921	0.961	68.2426	-101.795
1702-089	29	0.020	29.297	0.297	1.015	68.4906	-99.8955
1702-090	83	0.058	84.001	1.001	1.192	68.3061	-100.8015
1702-091	54	0.037	54.333	0.333	0.613	69.1707	-100.7054
1702-092	53	0.037	53.320	0.320	0.600	69.1691	-100.6952
1702-094	42	0.029	42.071	0.071	0.168	69.2716	-80.6078
1702-095	16	0.011	16.031	0.031	0.195	68.7668	-80.841
1702-096	91	0.063	91.579	0.579	0.632	65.451	-83.2584
Max				1.001	1.546		
Min				0.031	0.168		
Avg				0.549	0.762		

Overall, the observed maximum depth error of 1.546% in dataset 1702-086 is not significant. According to the International Hydrographic Office (IHO) Total Vertical Uncertainty (TVU) guidelines, this falls within level 1A survey standards. When calculating the error for the corresponding depth of 56m, the allowable TVU is 0.883m. The observed error is 0.879m and thus meets the requirements of a level 1A survey, not taking other sources of error into account. The minimum depth error of 0.168% in the 1702-094 dataset, actually meets Special Order survey standards for TVU with an allowable error of 0.402m while we have calculated an error of 0.071m, again not taking other errors into account. The IHO Standard Table can be found in Appendix B

and more information on the survey standards can be found in the IHO S44 manual (International Hydrographic Organization, 2008).

Similarly, the IHO has released a white paper titled 'Regulations of the IHO for international charts and chart specification of the IHO', where the organization describes Category Zone of Confidence (CATZOC). These values are used to highlight the accuracy of data presented on charts (International Hydrographic Organization, 2017). According to the set-out specifications, all datasets meet the highest CATZOC ratings of A1 for depth accuracy, without taking other sources of error into account. The CATZOC table is presented in Appendix C, and further information on these standards can be found in the IHO S-44 manual (International Hydrographic Organization, 2008).

To calculate a depth dependent error, the 17 CTD casts were analyzed further. Casts 1-10 would fall within the 'Arctic' domain and casts 11-17 would fall under the 'Hudson Bay' domain, no casts would fall within the 'NorthWest Atlantic' domain, see Figure 25. The matching HYCOM data and the CTD casts were then raytraced to give respective depths. The depths were then compared for each cast for each depth range and a % error calculated along with standard deviation for each region. The results are shown in Figure 34.

On average, the 'Hudson Bay' domain has a lower average error when compared to the 'Arctic' domain. This trend may become more or less apparent as more data is added to this comparison model. Overall, the largest error occurs between 30m-50m in the 'Arctic' domain, with an error rate of 1.16%.

Hudson Bay	Model Error by Cast and Water Depth (% Error Observed)								
Depth (m)	002	005	006	007	094	095	096	Average	Stdev
0-9	0.46	0.37	0.14	0.47	0.07	0.15	0.77	0.35	0.23
10-19	0.04	0.45	0.04	0.27	0.12	0.19	0.93	0.29	0.29
20-29	0.44	0.54	0.21	0.02	0.14		1.05	0.40	0.34
30-39	0.67	0.64	0.33	0.15	0.16		1.00	0.49	0.31
40-49	0.83	0.65	0.38	0.25	0.17		0.91	0.53	0.28
50-59	0.96	0.65	0.41	0.31			0.84	0.64	0.24
60-69	1.04	0.63	0.42	0.37			0.77	0.64	0.25
70-79	1.09	0.61	0.41	0.41			0.70	0.65	0.25
80-89			0.41	0.45			0.64	0.50	0.10
90-100				0.47			0.63	0.55	0.08

Arctic	Model Error by Cast and Water Depth (% Error Observed)											
Depth (m)	083	084	085	086	087	088	089	090	091	092	Average	Stdev
0-9	0.72	0.75	0.63	0.98	0.26	0.85	0.71	0.97	0.44	0.29	0.66	0.25
10-19	1.08	1.16	0.99	1.41	0.35	1.20	0.88	1.16	0.50	0.42	0.92	0.35
20-29	1.26	1.32	1.20	1.60	0.75	1.38	1.02	1.35	0.59	0.55	1.10	0.34
30-39	1.30	1.34	1.25	1.64	1.02	1.41	1.02	1.44	0.62	0.59	1.16	0.33
40-49	1.25	1.28	1.22	1.60	1.07	1.37		1.42	0.61	0.60	1.16	0.33
50-59	1.15	1.19	1.16	1.55		1.31		1.36	0.61	0.60	1.12	0.32
60-69	1.04	1.07	1.06			1.21		1.29			1.13	0.10
70-79	0.92	0.95	0.96			1.10		1.22			1.03	0.11
80-89	0.83	0.86	0.88			1.01		1.19			0.95	0.13
90-100	0.75	0.79	0.85			0.96					0.84	0.08

Figure 34. HYCOM Model Error % vs. In-Situ Data

The discrepancies between the observed and model depths would be a result of differing input temperature and salinity values. A larger difference in salinity and temperatures would lead to a higher calculated error % between the observed and model depths. The sound velocity algorithm should play no significant role as the algorithm was kept the same throughout the study.

4.3 Uncertainty Module

From the previously mentioned final uncertainty calculation , Table 15 is made to identify the individual components. The uncertainty values which are created dynamically are indicated as well as any uncertainty values which are static along with the values. The total uncertainty that is static for $\sigma^2_{Horizontal}$ is +/- 0.0067m and for $\sigma^2_{Vertical}$ it is +/- 0.1192m.

Table 15. Final Uncertainty Calculation Components

$\sigma^2_{Horizontal} = \sigma^2_{PPPest} + \sigma^2_{BW}* + \sigma^2_{LA} + \sigma^2_{RP}$	
σ^2_{PPPest}	Dynamic
$\sigma^2_{BW}*$	Dynamic
σ^2_{LA}	+/- 0.005m @95%
σ^2_{RP}	+/- 0.00174m @95%
$\sigma^2_{Vertical} = \sigma^2_{PPPest} + \sigma^2_{GeoRan} + \sigma^2_{LA} + \sigma^2_{RP} + \sigma^2_{SVP} + \sigma^2_{SS}$	
σ^2_{PPPest}	Dynamic
σ^2_{GeoRan}	+/- 0.1095
σ^2_{LA}	+/- 0.005m @95%
σ^2_{RP}	+/- 0.00018m @95%
σ^2_{SVP}	Dynamic
σ^2_{SS}	+/- 0.0045m

To calculate the dynamic components of the uncertainty, several differing methods were used. The σ^2_{PPPest} uncertainty was calculated through directly obtaining the SDLAT, SDLON and SDHGT from the NRCAN PPP output (.pos). The $\sigma^2_{BW}*$ is a function of depth and the corresponding value for each depth value is assigned as laid out in Table 8. The σ^2_{SVP} is a function of depth and location, as outlined in Sound Speed Uncertainty: and Figure 34.

To conduct the final uncertainty assessment, the same 22 raw GNSS datasets used in the GNSS Module section were analyzed, see Figure 22 for metadata. Please note, two of the datasets needed to be omitted from the uncertainty processing due to corrupt sonar data (14_41_53-2018_10_15 & 15_16_59-2018_10_17). For the uncertainty analysis the GNSS data was accompanied by the sonar and IMU data and georeferenced. The GNSS data was processed using NRCAN PPP with dual-constellation (GPS+GLONASS) and then qualified using the GNSS Qualification script. The Sonar, IMU and GNSS data were then georeferenced using the DepthStar proprietary software developed by CIDCO for the HydroBall. Once georeferenced, python scripts were developed to help match the georeferenced data line to the appropriate GNSS PPP data line in order to extract

the σ^2_{PPPest} . Additionally, the scripts also calculated the σ^2_{BW} and σ^2_{SVP} as a function of sonar depth and in the case of σ^2_{SVP}, also for location. The final THU and TVU were then calculated in two sets for each of the data lines; One set with the σ^2_{BW}* uncertainty and one without. Following this, summary data was made for the uncertainty per file with the maximum, minimum, average and standard deviations being calculated for THU and TVU for each file. The results of this analysis are displayed below in Table 16 (omitting beam width) and Table 17 (including beam width).

Table 16. Results of the Uncertainty Analysis for the Sample Datasets (Omitting Beam Width, all values in meters)

File:	SVP Region	THU_Max	THU_Min	THU_Avg	THU_StDev	TVU_Max	TVU_Min	TVU_Avg	TVU_StDev
13_43_52-2018_10_11	CANARC	0.351	0.325	0.329	0.006	0.640	0.574	0.599	0.019
13_46_34-2017_09_22	CANHUD	0.756	0.116	0.236	0.107	1.035	0.264	0.517	0.170
13_59_50-2018_10_11	CANARC	0.384	0.100	0.158	0.068	1.090	0.254	0.375	0.093
14_07_23-2018_10_12	CANARC	0.774	0.066	0.163	0.079	1.119	0.263	0.520	0.175
14_41_53-2018_10_15	--	--	--	--	--	--	--	--	--
14_45_26-2018_10_15	CANARC	0.711	0.043	0.095	0.107	1.438	0.218	0.446	0.204
15_11_52-2018_10_30	CANARC	0.831	0.039	0.091	0.070	1.220	0.220	0.566	0.189
15_16_59-2018_10_17	--	--	--	--	--	--	--	--	--
15_32_51-2018_10_17	CANARC	0.398	0.041	0.087	0.075	1.105	0.228	0.546	0.192
15_40_04-2018_10_14	CANARC	0.385	0.040	0.063	0.033	0.893	0.226	0.368	0.109
15_41_04-2018_10_21	CANARC	0.365	0.350	0.352	0.002	0.758	0.728	0.739	0.008
15_45_32-2018_10_21	CANARC	0.361	0.341	0.344	0.003	0.821	0.739	0.786	0.028
15_55_37-2018_10_21	CANARC	0.538	0.035	0.062	0.059	1.068	0.233	0.525	0.154
16_32_34-2018_08_02	CANARC	0.353	0.103	0.106	0.015	1.027	0.574	0.660	0.089
17_05_13-2018_08_02	CANARC	0.208	0.203	0.205	0.001	0.986	0.593	0.695	0.081
17_24_50-2018_08_02	CANARC	0.164	0.055	0.060	0.005	1.016	0.555	0.639	0.061
18_37_04-2017_12_11	CANARC	0.149	0.121	0.124	0.002	0.310	0.207	0.252	0.025
18_58_37-2017_12_11	CANARC	0.259	0.251	0.253	0.002	0.485	0.389	0.412	0.020
19_20_43-2018_08_02	CANARC	0.159	0.114	0.124	0.007	1.054	0.600	0.689	0.050
20_00_26-2018_08_02	CANARC	0.236	0.130	0.153	0.021	1.121	0.591	0.720	0.061
20_28_02-2018_08_02	CANARC	0.080	0.077	0.077	0.000	0.912	0.577	0.703	0.055
20_39_25-2018_10_15	CANARC	0.265	0.035	0.062	0.043	0.895	0.285	0.602	0.119

Table 17. Results of the Uncertainty Analysis for the Sample Datasets (Including Beam Width, all values in meters)

File:	SVP Region	THU_BW_Max	THU_BW_Min	THU_BW_Avg	THU_BW_StDev	TVU_Max	TVU_Min	TVU_Avg	TVU_StDev
13_43_52-2018_10_11	CANARC	4.341	2.215	3.325	1.051	0.640	0.574	0.599	0.019
13_46_34-2017_09_22	CANHUD	16.528	4.106	7.606	2.195	1.035	0.264	0.517	0.170
13_59_50-2018_10_11	CANARC	12.590	1.991	4.108	0.803	1.090	0.254	0.375	0.093
14_07_23-2018_10_12	CANARC	14.620	2.030	5.799	2.269	1.119	0.263	0.520	0.175
14_41_53-2018_10_15	CANARC	--	--	--	--	--	--	--	--
14_45_26-2018_10_15	CANARC	13.009	1.934	5.214	1.508	1.438	0.218	0.446	0.204
15_11_52-2018_10_30	CANARC	14.730	1.932	7.333	2.863	1.220	0.220	0.566	0.189
15_16_59-2018_10_17	CANARC	--	--	--	--	--	--	--	--
15_32_51-2018_10_17	CANARC	12.684	1.931	7.148	3.028	1.105	0.228	0.546	0.192
15_40_04-2018_10_14	CANARC	12.529	1.940	4.780	1.545	0.893	0.226	0.368	0.109
15_41_04-2018_10_21	CANARC	4.355	4.340	4.342	0.002	0.758	0.728	0.739	0.008
15_45_32-2018_10_21	CANARC	6.444	4.333	5.636	1.024	0.821	0.739	0.786	0.028
15_55_37-2018_10_21	CANARC	12.469	1.929	7.157	2.321	1.068	0.233	0.525	0.154
16_32_34-2018_08_02	CANARC	18.816	8.303	9.258	1.621	1.027	0.574	0.660	0.089
17_05_13-2018_08_02	CANARC	16.365	8.404	9.373	1.094	0.986	0.593	0.695	0.081
17_24_50-2018_08_02	CANARC	16.220	8.255	9.059	1.094	1.016	0.555	0.639	0.061
18_37_04-2017_12_11	CANARC	4.116	2.011	2.289	0.708	0.310	0.207	0.252	0.025
18_58_37-2017_12_11	CANARC	4.249	2.141	2.295	0.547	0.485	0.389	0.412	0.020
19_20_43-2018_08_02	CANARC	18.824	8.314	9.029	1.058	1.054	0.600	0.689	0.050
20_00_26-2018_08_02	CANARC	16.396	8.330	9.582	1.257	1.121	0.591	0.720	0.061
20_28_02-2018_08_02	CANARC	14.578	8.277	9.851	0.926	0.912	0.577	0.703	0.055
20_39_25-2018_10_15	CANARC	14.549	4.025	8.683	2.194	0.895	0.285	0.602	0.119

As the IHO S-44 minimum requirements for survey 1a and 1b both have a depth dependent calculation (see Appendix B), the datasets were analyzed on a per line basis. The THU and TVU values calculated incorporating the sonar depth and then qualified as a pass/fail for each line. The results per file are summarized in Table 18 below, where THU_BW is the total horizontal uncertainty with the beam width uncertainty.

Table 18. IHO S44 1a 1b PASS/FAIL Rate

		TVU		THU		THU_BW	
File:	**SVP Region**	**PASS**	**FAIL**	**PASS**	**FAIL**	**PASS**	**FAIL**
13_43_52-2018_10_11	CANARC	0.00%	100.00%	100.00%	0.00%	100.00%	0.00%
13_46_34-2017_09_22	CANHUD	79.69%	20.31%	100.00%	0.00%	19.87%	80.13%
13_59_50-2018_10_11	CANARC	92.19%	7.81%	100.00%	0.00%	98.95%	1.05%
14_07_23-2018_10_12	CANARC	73.10%	26.90%	100.00%	0.00%	65.68%	34.32%
14_41_53-2018_10_15	CANARC	--	--	--	--	--	--
14_45_26-2018_10_15	CANARC	89.34%	10.66%	100.00%	0.00%	78.68%	21.32%
15_11_52-2018_10_30	CANARC	84.04%	15.96%	100.00%	0.00%	41.31%	58.69%
15_16_59-2018_10_17	CANARC	--	--	--	--	--	--
15_32_51-2018_10_17	CANARC	91.47%	8.52%	99.99%	0.00%	41.71%	58.29%
15_40_04-2018_10_14	CANARC	99.77%	0.23%	100.00%	0.00%	80.64%	19.36%
15_41_04-2018_10_21	CANARC	0.00%	100.00%	100.00%	0.00%	100.00%	0.00%
15_45_32-2018_10_21	CANARC	0.00%	100.00%	100.00%	0.00%	38.26%	61.74%
15_55_37-2018_10_21	CANARC	92.38%	7.62%	100.00%	0.00%	43.33%	56.67%
16_32_34-2018_08_02	CANARC	99.66%	0.34%	100.00%	0.00%	0.00%	100.00%
17_05_13-2018_08_02	CANARC	98.46%	1.54%	100.00%	0.00%	0.00%	100.00%
17_24_50-2018_08_02	CANARC	100.00%	0.00%	100.00%	0.00%	0.00%	100.00%
18_37_04-2017_12_11	CANARC	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
18_58_37-2017_12_11	CANARC	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
19_20_43-2018_08_02	CANARC	89.20%	10.80%	100.00%	0.00%	0.06%	99.94%
20_00_26-2018_08_02	CANARC	71.07%	28.93%	100.00%	0.00%	0.14%	99.86%
20_28_02-2018_08_02	CANARC	100.00%	0.00%	100.00%	0.00%	0.11%	99.89%
20_39_25-2018_10_15	CANARC	99.77%	0.23%	100.00%	0.00%	17.61%	82.39%

4.3.1 Total Horizontal Uncertainty (THU)

When analyzing the THU without beam width (Table 16), we can see that the average THU (THU_Avg) is 0.157m when averaged between the datasets, with a maximum of 0.352m and minimum of 0.060m. The standard deviation is relatively low with the average between the datasets being +/- 0.035m and the maximum value of +/- 0.107m occurring in the 13_46_34-2017_09_22 dataset. The maximum THU was found in the 15_11_52-2018_10_30 dataset with a value of 0.831m and the minimum THU was found in the 15_55_37-2018_10_21 dataset with a value of 0.035m.

When introducing beam width uncertainty these values change rather drastically (Table 17). The average THU (THU_BW_Avg) increases to 6.593m on average between the datasets, with a

maximum of 9.851m and minimum of 2.289m. The standard deviation also increases throughout the datasets with the average between the datasets being +/- 1.455m and the maximum value of +/- 3.028m occurring in the 15_32_51-2018_10_17 dataset. The maximum THU was found in the 19_20_43-2018_08_02 dataset with a value of 18.824m and the minimum THU was found in the 15_55_37-2018_10_21 with a value of 1.929m.

When viewing Table 18, all datasets meet IHO S-44 level 1a and 1b requirements for THU. When the beam width uncertainty is added, this is not the case. While four of the datasets maintain the 100% pass rate (13_43_52-2018_10_11, 15_41_04-2018_10_21, 18_37_04-2017_12_11 & 18_58_37-2017_12_11), an equal number and more have basically a 0% pass rate (16_32_34-2018_08_02, 17_05_13-2018_08_02, 17_24_50-2018_08_02, 19_20_43-2018_08_02, 20_28_02-2018_08_02 & 20_00_26-2018_08_02). The rest of the datasets range between a 17% and 99% pass rate.

4.3.2   Total Vertical Uncertainty (TVU)

When analyzing the TVU (Table 16, Table 17), we can see that the average TVU (TVU_Avg) is 0.568m on average between the datasets, with a maximum of 0.786m and minimum of 0.252m. Standard deviations for the TVU are relatively low at decimeter and sub-decimeter levels, with the average being +/- 0.095m and the maximum standard deviation being +/- 0.204m in the 14_45_26-2018_10_15 dataset. The maximum TVU recorded was 1.438m in the 14_45_26-2018_10_15 dataset and the minimum TVU recorded was 0.207m in the 18_37_04-2017_12_11 dataset.

When viewing Table 18, there are mixed results when analyzing the pass/fail rate for IHO S-44 TVU level 1a and 1b requirements. There are 7 datasets with an approximate 100% pass rate (15_40_04-2018_10_14, 16_32_34-2018_08_02, 17_24_50-2018_08_02, 18_37_04-2017_12_11, 18_58_37-2017_12_11, 20_28_02-2018_08_02, 20_39_25-2018_10_15), while there are three datasets which have a 0% pass rate (13_43_52-2018_10_11, 15_41_04-2018_10_21 & 15_45_32-2018_10_21), primarily resulting from poor GNSS quality. The rest of the datasets ranges between 70% and 99% pass rate.

5   Discussion

5.1   GNSS Module

Overall, it is recommended that when surveying in the Canadian Arctic or any location above 55^0 North, to use dual-constellation GNSS. While the addition of GLONASS added an increase in standard deviation of the clock, the overall benefits are greater than the drawbacks. Perhaps most importantly, in datasets where the qualification pass rate was 0% due to poor quality GPS data, the addition of GLONASS allowed this data to be converted to a high qualification pass rate. In other words, in cases where GPS data is not good enough, the addition of GLONASS provides excellent supplementation to drastically improve qualification pass rates.

5.2   Sound Velocity Profile Module

It can be shown from the results that using the HYCOM model would be an acceptable alternative to obtain temperature and salinity values in the absence of observed CTD casts when in the shallow waters of the Canadian Arctic. When converted to a SVP and used to calculate depth through the water column, the maximum depth error seen was 1.55% for shallow depth (<100m) Canadian Arctic water casts. While it is not always cost-effective or possible to obtain observed CTD casts when off-shore, these are very favorable results and show that the HYCOM model can be used and thus reduce time and costs during a survey.

5.3   Uncertainty Module

Overall, the final uncertainty values for the 20 datasets were rather favourable. While it is true that some uncertainties were omitted from the calculation either due to difficulties in calculating an

applicable value (ie. σ^2_{PPPeng}) or from not being fully applicable to the scope of the project (ie. $\sigma^2_{BW}*$), it can be argued that for the nature of the project, the uncertainty assessment was thorough.

When the beam width uncertainty ($\sigma^2_{BW}*$) is neglected from the THU, we can see that the average and even the maximum THU values, 0.157m & 0.831m respectively, always meet the IHO 1a and 1b survey standards. As a result, all datasets meet the requirements of the maximum allowable THU and unexpectedly also within the maximum allowable THU for special order surveys as well. This shows that when conducting surveys of this type that horizontal accuracies are not the largest of concerns as they can easily be met for all IHO survey orders. Additionally, it can be shown that a simple single-beam sonar paired with a dual-frequency GNSS sensor is capable of meeting these horizontal standards, the lowest recorded THU was a small 0.035m. This potentially opens up the possibilities for other similar platforms to begin contributing to the CSB initiative.

Of course, the case is not the same when factoring in the beam width uncertainty ($\sigma^2_{BW}*$). We can see that only 4 of the 20 surveys met IHO 1a and 1b requirement for THU at the 100% pass rate, half of the surveys fell within a 17-99% pass rate and 6 surveys had a 0% pass rate. The beam width uncertain certainly does increase the THU by a large amount. For example, the maximum recorded THU in all the datasets is 18.824m; This value approaches the 20m allowable THU for IHO level 2 surveys, the most relaxed standard. This is rather shocking as just by omitting this uncertainty constituent, we have met IHO special order requirements for THU, the most rigorous standard. However, as we have discussed in **Sonar Sounder Uncertainty: Beam Width Error**, it is acceptable to omit this uncertainty as the nature of this project is primarily with the concern of safety of navigation and not necessarily with accurate representation of the seafloor. The largest

source of uncertainty to the THU beyond beam the width uncertainty is the horizontal PPP estimated uncertainty (σ^2_{PPPest}).

The TVU does have an impressive pass rate, especially when considering the nature of this project. We can see the 7 of the 20 surveys had a 100% pass rate, half of the surveys fell within a 70-99% pass rate and 3 surveys has a 0% pass rate. The average TVU between all the datasets is 0.568m, falling just above the depth independent portion of the IHO 1a and 1b standards. Nonetheless, the results were still favourable with almost all surveys having a 70% or greater IHO level 1a and 1b pass rate. For the 3 surveys which had a 0% pass rate, it was found that the largest contributor to the TVU is the vertical PPP estimated uncertainty (σ^2_{PPPest}) with values averaging around 0.4m-0.5m, often consuming the entire TVU budget alone. The second greatest contributor to the 0% pass rate on the 3 surveys fluctuates between the geometrical range uncertainty (σ^2_{GeoRan}) and the SVP uncertainty (σ^2_{SVP}), with values ranging on average between 0.1m-0.2m each. As vertical accuracy is typically the primary concern with bathymetric surveys, it can be shown that the results were quite favourable. Most of the data met the IHO level 1a and 1b maximum allowable TVU levels, affirming the possibility of using simple hardware to conduct surveys in the Arctic while meeting surveying standards.

The biggest contributor to the THU and TVU is the PPP estimated uncertainty (σ^2_{PPPest}), making the GNSS component the largest uncertainty contributor overall. This is somewhat expected given the nature of GNSS coverage in the Arctic. Overall, the results were quite favourable as most of the data fell within the IHO level 1a and 1b maximum total uncertainty components for horizontal and vertical. It should be noted that while a reference to IHO level 1a and Special Order surveys

have been made, in reality these particular standard require a full seafloor coverage not possible with a single-beam sonar. Thus, these comparisons are used primarily for reference and not an indication that the final data will meet the full extent of these standards. The IHO level 1b standard does not require full seafloor coverage and thus the derived data from this project would fit in this category.

5.4 Future Work

Although, this pilot program did achieve quite a good amount in a short period of time, it was far from perfect and there are always ways in which to improve.

5.4.1 GNSS Module:

It would be recommended that the comparison of dual-constellation vs single-constellation GNSS data be furthered through an increase in the number of datasets as well and an incorporation of a location component. This will help to better assess the impacts of low inclination of GNSS on Arctic surveys and how GLONASS can help mitigate these issues.

5.4.2 Model Tides & Reduction

It would be recommended that a deeper investigation be made into examining the effects of utilizing model tides along with tidal reduction on data which the GNSS qualification does not pass. This data currently archived in the current system as only ellipsoidal referenced data is used. However, with the addition of the tidal route of processing, this data may be used to help revive data that may have been unusable.

5.4.3   SVP Module:

Similarly to the GNSS module, it would be recommended that the HYCOM model comparison to in-situ data be extended through the use of more data covering a larger range of times, locations and depths. While a full HYCOM sound speed uncertainty model is out of the scope of this project, the foundation has been laid for continuation towards this cause by the Ocean Mapping Group of the University of New Brunswick. Currently the model is limited to shallow water depths of 100m or less, however this can be expanded to include all depths and by using a greater amount of data.

5.4.4   Uncertainty Module:

It would be recommended that future work be done to improve on the uncertainty model created for this project for shallow water Arctic surveys. To do this, one may help identify other sources of uncertainty or help to improve on the quantification of other uncertainties. For example, a systematic examination of the NRCAN PPP accuracy would help to identify any errors within the PPP engine itself. Furthermore, an uncertainty comparison using modelled tides vs. GNSS referenced data would be helpful to identify best practices and potentially improve poor uncertainty results through augmentation.

5.4.5   Crowd-Sourced Bathymetry:

The benefits of CSB data is becoming well-known and project like this prove their potential effectiveness towards global bathymetric coverage. It would be recommended that project like this continue and build upon the work that has been done, not just with this project but also others. While hardware was kept relatively inexpensive and simple, the potential for greater complexity arises in future iterations. The next logical step being the incorporation of a multi-beam sonar to

provide greater seafloor coverage, increased knowledge of the seafloor and improvements towards

safety of navigation.

6 Conclusion

In conclusion, we have covered a wide array of topics while attempting to answer one question; how to *efficiently* process crowd-sourced hydrographic data through *automation* while improving bathymetric *accuracy* and assessment of *uncertainty*. Challenges associated with conducting hydrographic surveys in the Canadian Arctic have been explored, including methods on improving GNSS data, qualification of GNSS data for use as the vertical survey reference, proposed methodology for reducing ERS surveys to MSL and obtaining tidal data for offsets, usage of hydrographic models to supplement the lack of in-situ CTD casts and a thorough uncertainty identification and analysis. We have shown how dual-constellation GNSS appears to be the better choice when surveying in the Canadian Arctic, that the HYCOM model can be a perfectly acceptable solution when in-situ CTD casts cannot be taken and tackled the identification and assignment of uncertainty to the project data and sources of data, while explaining the research objectives. Overall, the research objectives of the project have been met and in large the project has been considered a success. There are still more ways the project can be improved, however the foundation has now been laid and future work can expand on the work done by all the project partners with help of the indigenous communities.

It is safe to say that crowd-sourced initiatives such as this one, help to get a better understanding of the Canadian Arctic waterways. This initiative hopes to help propel the understanding of the Canadian Arctic waterways through a unique crowd-sourced approach. It is our belief that projects like these help to provide a win-win situation for all stakeholders; The indigenous communities involved receive bathymetric maps for their involvement and the larger community in general receives data which help propel bathymetric awareness in these remote areas.

We have demonstrated a full end-to-end workflow on how to process crowd-sourced bathymetric data in this thesis. We sincerely hope that the work done as part of this project will go on to help pave the way for future expansion to existing project and to new projects towards the future of CSB in Canada. As beautifully expressed:

"Ocean exploration gives mankind a sense of human progress and heritage. It provides the experience and knowledge necessary to undertake stewardship of the ocean and its resources, and thus sets a course for future generations to navigate. What lies ahead is still unknown. Whatever it is, however, will be influenced by what is found through tomorrow's exploration – and, will likely be different than today's predictions!"
Office of the Chief Scientist, NOAA

Bibliography

Fischhoff, B., & Davis, A. (2014). *Communicating scientific uncertainty.* PNAS.

Abdelazeem, M., & Celik, R. (2014). Accuracy and repeatability investigation of CSRS-PPP online processing service. *Coordinates*.

ADMIRALTY. (2017). *Zones of Confidence (ZOC) Table.* Retrieved from admiralty.co.uk.

Alkan, R., Kalkan, Y., & Turkiye, A. (2006). *Sound Velocity Determination with Empirical Formulas and Bar Check.* Munich: Shaping the Change.

Allen, J., Keen, P., Gardiner, J., Quartley, M., & Quartley, C. (2017). A new salinity equation for sound speed instruments. *Authors Limnology and Oceanography*.

Arfeen, K. (2019, 03 19). Automated Processing Of Arctic Crowd-Sourced Hydrographic Data While Improving Bathymetric Accuracy And Uncertainty Assessment. Biloxi, Mississippi, USA: THE HYDROGRAPHIC SOCIETY OF AMERICA (THSOA).

Arfeen, K., & Church, I. (2019). Automated Processing Of Arctic Crowd-Sourced Hydrographic Data While Improving Bathymetric Accuracy And Uncertainty Assessment. *USHYDRO 2019.* Biloxi: THE HYDROGRAPHIC SOCIETY OF AMERICA (THSOA).

Beaudoin , J., Clarke, H., & Bartlett. (2006). USAGE OF OCEANOGRAPHIC DATABASES IN SUPPORT OF MULTIBEAM MAPPING OPERATIONS ONBOARD THE CCGS AMUNDSEN. *Lighthouse*.

Beaudoin, J. (2009). *Transit Sounding Reclamation.* Fredericton: HydroOctave Consulting.

Bongiovanni, C., & Schmidt, V. (2016). Evaluating Outside Source Interferometric Data for Chart Updates. *Canadian Hydrographic Conference.* Halifax: CHC2016.

Calder, B., Kraft, B., Moustier, C., Lewis, J., & Stein, P. (2004). MODEL-BASED REFRACTION CORRECTION IN INTERMEDIATE DEPTH MULTIBEAM ECHOSOUNDER SURVEY. *European Conference on Underwater Acoustics (ECUA)* (pp. 795-800). Delft: ECUA.

Chen, C.-T., & Millero, F. (1980). A new high pressure equation of state for seawater. *Deep Sea Research*.

Church, I. (2017). *GGE3353 Heave Lecture.* Fredericton: University of New Brunswick.

Church, I., Brucker, S., Hughes, J., Haigh, S., Bartlett, J., & Janzen, T. (2009). Developing Strategies to Facilitate Long Term Seabed Monitoring in the Canadian Arctic using Post Processed GPS and Tidal Models. *US Hydro*.

Church, I., Hughes Clarke, J., Haigh, S., & Roodesh, R. (2012). Modelling the estuarine circulation of the Port of Saint John: Visualizing complex sound speed distribution. *Canadian Hydrographic Conference.* Niagara Falls: CHC2012.

Church, I., Hughes, J., & Haigh, S. (2007). Use of nested finite0element hydrodynamic model to predict phase and amplitude modification of tide within narrow fjords. *US HYDRO.* US HYDRO.

Clarke, H. (2003). A reassessment of vessel coordinate systems: what is it that we are really aligning? . *US Hydrographic Conference.* Biloxi.

Clarke, H., Dare, P., Beaudoin, J., & Bartlett, J. (2005). A stable vertical reference for bathymetric surveying and tidal analysis in the high Arctic. *US Hydrographic Conference.* San Diego.

Collins, K., Hannah, C., & Greenberg, D. (2011). *Validation of a High Resolution Modelling System for Tides in the Canadian Arctic Archipelago.* DFO.

Cooper, P., & Hersey, J. (2014). Crowdsourcing Enhances Navigation Awareness. *Hydro International*.

European Space Agency. (2018, 06 22). *GLONASS Space Segment*. Retrieved from ESA Navipedia: https://gssc.esa.int/navipedia/index.php/GLONASS_Space_Segment

European Space Agency. (2018, 06 19). *GPS Space Segment*. Retrieved from ESA Navipedia: https://gssc.esa.int/navipedia/index.php/GPS_Space_Segment

Government of Canada. (2018, 12 14). *Arctic Charting*. Retrieved from Fisheries and Oceans Canada: http://www.charts.gc.ca/arctic-arctique/index-eng.asp

Grosso, V. (1974). New equation for the speed of sound in natural waters (with comparisons to other equations). *The Journal of the Acoustical Society of America*.

Guo, Q. (2015). Precision comparison and analysis of four online free PPP services in static positioning and tropospheric delay estimation. *GPS Solutions*.

Hansen, R. (2012). *Introduction to sonar.* Oslo: University of Oslo.

Hare, R. (1995). DEPTH AND POSITION ERROR BUDGETS FOR MULTIBEAM ECHOSOUNDING. *International Hydrographic Review*.

Hegarty, C. (2013). *Workshop on GNSS Data Application to Low Latitude Ionospheric Research* . Bedford: International Centre for Theoretical Physics.

Honeywell. (2005). *Digital Compass Solution HMR3400.* Retrieved from honeywell.com : digchip.com/datasheets/parts/datasheet/2592/HMR3400-pdf.php

Huber, K., Heuberger, F., Abert, C., Karabatic, A., Weber, R., & Berglez, P. (2010). PPP: Precise Point Positioning – Constraints and Opportunities.

Intergovernmental Oceanographic Commission. (2010). *The International Thermodynamic Equation of Seawater -2010: calculation and use of thermodynamic properties.* UNESCO.

International GNSS Service. (2019). *IGS Products.* Retrieved from igs.org: https://www.igs.org/products

International Hydrographic Organization. (2008). *IHO Standards for Hydrographic Surveys S-44 5th Ed. .* Monaco: International Hydrographic Bureau.

International Hydrographic Organization. (2017). *Regulations of the IHO for International (INT) Charts and Chart Specifications of the IHO.* Monaco: International Hydrographic Organization.

International Hydrographic Organization. (2018). *Guidance on Crowdsourced Bathymetry (Draft).* Monaco: International Hydrographic Organization 2018.

Jayachandran, P., Langley, R., MacDougall, J., Mushini, S., Pokhotelov, D., Hamza, A., & Mann, I. (2009). Canadian High Arctic Ionospheric Network (CHAIN). *Radio Science.*

Jensen, A., & Sicard, J. (2010). Challenges for positioning and navigation in the Arctic. *Coordinates.*

Jensen, A., & Sicard, J.-P. (2010). Challenges for Positioning and Navigation in the Arctic. *Coordinates.*

Karaim, M., Elsheikh, M., & Noureldin, A. (2018). GNSS Error Sources. *intechopen.*

Langley, R. (1999). Dilution of Precision. *Innovation.*

Langley, R. (2009, 06 24). GNSS Use in the High Arctic: Issues and Solutions from a Canadian Perspective. Tromso, Norway.

Leighton, S. (2019). The Hydroball Project: Crowd Source Bathymetry in Northern Canada. *The Journal of Ocean Technology.*

Mackenzie, K. (1981). Nine-term equation for sound speed in the oceans. *The Journal of the Acoustical Society of America.*

Malinowski, M., & Kwiecien, J. (2016). A Comparative Study of Precise Point Positioning (PPP) Accuracy Using Online Services. *Sciendo.*

Matteo, N., Morton, Y., Chandrasekaran, P., & Van Graas, F. (2009). Higher Order Ionosphere Errors at Arecibo, Millstone, and Jicamarca. *International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2009).*

McHugh, C., Church, I., Kim, M., & Maggio, D. (2015). Comparison Of Horizontal And Vertical Resolvable Resolution Between Repetitive Multi-Beam Surveys Using Different Kinematic Gnss Methods . *INTERNATIONAL HYDROGRAPHIC REVIEW .*

Medwin, H. (1975). Speed of sound in water: A simple equation for realistic parameters. *The Journal of the Acoustical Society of America*.

Metzger, Smedstad, Thoppil, Hurlburt, Cummings, Wallcraft, & Franklin. (2014). US Navy Operational Global Ocean and Arctic Ice Prediction Systems. *Oceanography*.

Misra, P. (1996). The Role of the Clock in a GPS Receiver. *GPS World*.

Monge, B., Radicella, S., Clara de Lacy, M., & Herraiz, M. (2011). On the effects of the ionospheric disturbances on precise point positioning at equatorial latitudes. *GPS Solutions*.

Naankeu Wati, G., & Seube, N. (2016). ERROR BUDGET ANALYSIS FOR SURFACE AND UNDERWATER SURVEY SYSTEM. *International Hydrographic Review*.

NOAA. (2019, 08 05). *WORLD OCEAN ATLAS 2018 (WOA18)*. Retrieved from nodc.noaa.gov: https://www.nodc.noaa.gov/OC5/woa18/

Novaczek, E., Devillers, R., & Edinger, E. (2019). Generating higher resolution regional seafloor maps from crowd-sourced bathymetry. *PLOS ONE*.

Novatel. (2003). *GPS Position Accuracy Measures* . Calgary.

Novatel. (2018). *An introduction to GNSS*. novatel.com. Retrieved from novatel.com.

Novatel. (2019). *An Introduction to GNSS*. Retrieved from novatel.com: https://www.novatel.com/an-introduction-to-gnss/chapter-4-gnss-error-sources/error-sources/

Ocean Conservancy. (2013, 02 21). *Crowdsourcing the Ocean Floor: How Mariners Can Gather Valuable Information for Better Decision-Making*. Retrieved from oceanconservancy.org: https://oceanconservancy.org/blog/2013/02/21/crowdsourcing-the-ocean-floor-how-mariners-can-gather-valuable-information-for-better-decision-making/

Office of the Auditor General of Canada. (2014, 07 18). *Chapter 3—Marine Navigation in the Canadian Arctic*. Retrieved from oag-bvg.gc.ca: http://www.oag-bvg.gc.ca/internet/English/parl_cesd_201410_03_e_39850.html#ex4

Ozvald, I., & Gorelick, M. (2018). *High Performance Python; Chapter 4 - Dictionaries and Sets*. O'REILLY.

Pike, & Beiboer. (1993). A comparison between algorithms for the speed of sound in seawater. *Hydrographic Society*.

Reed, A. (2017). Steps Towards a True Crowd, Collecting Bathymetry via Electronic Navigation Systems. *U.S. HYDRO 2017*. Galveston: U.S. HYDRO.

Reid, T., Walter, T., Blanch, J., & Enge, P. (2015). GNSS Integrity in the Arctic. *The 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*. Tampa.

Renoud, W. (2018). *COMPARISON OF REGIONAL GRAVIMETRICGEOIDS IN THE CANADIAN ARCTIC FOR THEASSESSMENT OF CVDCW HYVSEPS ACCURACY.* Fredericton: University of New Brunswick.

Rizos, C., Janssen, V., Roberts, C., & Grinter, T. (2012). PPP versus DGNSS. *Geomatics World*.

Rizos, C., Janssen, V., Roberts, C., & Grinter, T. (2012). Precise Point Positioning: Is the Era of Differential GNSS Positioning Drawing to an End? *Precise Point Positioning*.

Robin, C., Nudds, S., MacAulay, A., De Lange Boom, B., & Bartlett, J. (2016). Hydrographic Versital Separation Surfaces (HyVSEPs) for the Tidal Waters of Canada. *Marine Geodesy*.

Simons, D., & Snellen, I. (n.d). *Lecture notes "Seafloor Mapping" part2.* Delft Institute of Earth Observation and Space Systems (DEOS).

Smith, S. (2018, 06 18). *From seaports to the deep blue sea, bathymetry matters on many scales*. Retrieved from oaacoastsurvey.wordpress.com: https://noaacoastsurvey.wordpress.com/category/crowdsourced-bathymetry/

Smith, W. (2008). An uncertainty model for deep ocean single beam and multibeam echo sounder data. *Marine Geophysical Researches*.

Snellen, M. J. (2008). Correcting bathymetry measurements for water sound speed effects using inversion theory. *The Journal of the Acoustical Society of America*.

Stanford. (2018). *The Earth's Ionosphere*. Retrieved from solar-center.stanford.edu: http://solar-center.stanford.edu/SID/activities/ionosphere.html

Tantra, J. (2010). Experiences in Building Python Automation Framework for Verification and Data Collections. *PyCon Asia-Pacific.* PyCon Asia-Pacific 2010.

Teledyne Caris. (2018). *SVP Files.* Retrieved from teledynecaris.com: http://www.teledynecaris.com/docs/4.4.11/hips%20and%20sips/CARIS%20HIPS%20and%20SIPS%20Help/SVP-Editor.20.11.html

Wells, D., Beck, N., Delikaraoglou, D., Kleusberg, A., Langley, R., & Vanicek, P. (1999). *GUIDE TO GPS POSITIONING.* New Brunswick: Department of Geodesy and Geomatics Engineering, University of New Brunswick.

Wilson, W. (1960). Equation for the Speed of Sound in Sea Water. *The Journal of the Acoustical Society of America*.

Yu, X., & Gao, J. (2017). Kinematic Precise Point Positioning Using Multi-Constellation Global Navigation Satellite System (GNSS) Observations. *International Journal of Geo-Information*.

Yves, M., Pierre, T., Francois, L., Pierre, H., & Jan, K. (2008). online Precise Point Positionin. *GPS World*.

Appendix

6.1   Appendix A

Contents of the NRCAN PPP output ZIP file:

Filename	Type	Description
yyy.csv	CSV	TABLE OF: [latitude_decimal_degree, longitude_decimal_degree, ellipsoidal_height_m, decimal_hour, day_of_year, year, ortho_height_m_cgvd28, rcvr_clk_ns]
yyy.pdf	PDF	NRCAN GENERATED REPORT
yyy.pos	POS (TEXT)	LOCATION INFORMATION AND STATISTICS (SEE OUTPUT DESCRIPTION BELOW)
yyy.sum	SUM (TEXT)	SUMMARY DATA OF PROCESSING (SEE OUTPUT DESCRIPTION BELOW)
errors.txt	TEXT	COLLECTION OF ERRORS THAT OCCURRED DURING PROCESSING
output_descriptions.txt	TEXT	DESCRIPTION OF OUTPUT FIELDS OF DATA (SEE BELOW)

OUTPUT_DESCRIPTIONS.TXT CONTENTS (TAKEN DIRECTLY FROM NATURAL RESOURCES CANADA)

```
----------------------------------------------------------------------------
CANADIAN GEODETIC SURVEY, SURVEYOR GENERAL BRANCH, NATURAL RESOURCES CANADA
GOVERNMENT OF CANADA, 588 BOOTH STREET ROOM 334, OTTAWA ONTARIO K1A 0Y7
Phone: 343-292-6617
Email: nrcan.geodeticinformationservices.rncan@canada.ca
----------------------------------------------------------------------------

/**********************************************************************
 **********************************************************************
                    ~~~ Disclaimer ~~~
 Natural Resources Canada does not assume any liability deemed to have been
```

caused directly or indirectly by any content of its CSRS-PPP positioning
service.
```
*************************************************************************
*************************************************************************/


/*************************************************************************
*      -- This file describes the contents of CSRS-PPP outputs.                              *
*                                                                                              *
*    1- SUM file :    contains the parameters and the results of the PPP                 *
*                            processing.                                                       *
*                                                                                              *
*    2- POS file :    contains the positioning information for each epoch                *
*                            processed.                                                        *
*                                                                                              *
*    3- CSV file :    A comma separated (.csv) format text file containing              *
*                            the positioning and clock information for each epoch          *
*                            processed.                                                        *
*                                                                                              *
* Note: For more CSRS-PPP documentation and for PDF report description             *
*            please visit: https://webapp.geod.nrcan.gc.ca/geod/tools-outils/ppp.php *
*************************************************************************/


/*********************
 1- SUM file description
 *********************/
```
HDR Header
VER Software version
NOW UTC time at which the data set was processed
RNX Name of RINEX file processed
MOD Processing mode: static or kinematic
BEG Beginning of processed data in RINEX file
END End of processed data in RINEX file
INT Interval of data processing
EPO #epochs processed #epochs tried to be processed #epochs expected in the RINEX file in the time
interval
EXE Processing time
OBS Observables processed for each system
PAR Parameters estimated in the PPP filter
        TYPE   Type of parameter
        SYS   GNSS for which this parameter applies (blank means all GNSS)
        SIG   Signal for which this parameter applies (blank means all GNSS)
        SAT   Satellite for which this parameter applies (ALL means all satellites for this GNSS and signal)
        INIT   Initial constraint put on the parameter in the PPP filter
        NOISE Process noise put on the parameter in the PPP filter
SP3 SP3 (precise satellite orbits) files
CLK CLK (precise satellite clocks) files
ERP ERP (Earth rotation parameters) files
ATX ANTEX (phase-center corrections) file

DCB Differential code bias (DCB) or Bias SINEX files
GIM Global Ionospheric Map (GIM) files
VMF Vienna Mapping Function 1 (VMF1) grid files (tropospheric correction)
ELV Elevation mask (degrees)
TZD A priori tropospheric model applied
REC Receiver type (as read from the RINEX file)
ANT Antenna type (as read from the RINEX file)
ARP Antenna reference point (north, east, up) (as read from the RINEX file)
PCO Phase center offsets for the user antenna for each frequency (as read from the ANTEX file)
OTL Ocean Tide Loading coefficients
YAW Periods where satellites are in eclipse: PRN START_TIME END_TIME TYPE_OF_ECLIPSE BETA_ANGLE SATELLITE_TYPE
IAR Integer Ambiguity Resolution: percentage of ambiguities fixed
AVF A posteriori Variance Factor used to scale the covariance matrix
APR Source of the a priori coordinates
ELL Ellipsoid used for (x,y,z) to (lat,lon,h) transformation
POS Information on the a priori and estimated coordinates
      CRD   Coordinate type
      SYST   Reference system
      EPOCH   Epoch of the coordinates
      A_PRIORI   A priori coordinates
      ESTIMATED*   Estimated coordinates
      DIFF*   Differences (m) between the estimated and a priori coordinates
      SIGMA*   Standard deviation (95%) (m) of the estimated coordinates
      CORRELATIONS* Correlations matrix of the estimated coordinates
PRJ UTM* and MTM** projections
OHT Orthometric height***
VLM Velocity model used for epoch transformation
OFF Receiver clock offset at the last epoch and its standard deviation
DRI Approximate clock drift (ns/day)
RES Root-mean-square of the residuals for each signal, and classification per elevation (ELV)
FLG Additional information on each satellite
      PRN Satellite PRN
      EPO Number of epochs where the satellite is used in the PPP solution
      TRK Number of epochs where the satellite is rejected due to missing observations (tracking issues)
      CLK Number of epochs where the satellite is rejected due to missing satellite clock corrections
      EPH Number of epochs where the satellite is rejected due to missing ephemeris (orbit) corrections
      ELV Number of epochs where the satellite is rejected due to elevation angle below cutoff angle
      YAW Number of epochs where the satellite is rejected due to satellite yaw manoeuver (eclipse)
      DCB Number of epochs where an observation is rejected due to missing DCB corrections
      SLP Number of cycle slips detected
      MIS Number of epochs where an observation is rejected due to blunders detected in the misclosure vector
      RES Number of epochs where an observation is rejected due to a large residual

(*) Only for Static mode
(**) Only for Static if user selects NAD83(CSRS) as the reference frame and the estimated position is in Canada

(***) Only for Static if the estimated position is within in the Geoid model

```
/**********************
  2- POS file description
 **********************/
```
DIR   Processing Direction:
        FWD : Forward
        BWD : Backward smoothed
        SCA : final estimated position with scaled sigmas (Only for Static mode)
FRAME   Reference Frame
STN   Station name
DAYofYEAR   DAY of YEAR
YEAR-MM-DD   Observation Date
HR:MN:SS.SS Observation Time
NSV   Number of satellite
GDOP   Geometric Dilution of Precision
RMSC(m)   stds/RMS of Carrier Phase residuals
RMSP(m)   stds/RMS of Pseudo-Range residuals
DLAT(m)   Difference between estimated and a priori positions- North dir.
DLON(m)   Difference between estimated and a priori positions- East dir.
DHGT(m)   Difference between estimated and a priori positions- Up dir.
CLK(ns)   Clock offset
TZD(m)   Estimated Total Zenith Delay for each epoch
SDLAT(95%) Standard deviation(95%) of estimated positions - North dir.
SDLON(95%) Standard deviation(95%) of estimated positions - East dir.
SDHGT(95%) Standard deviation(95%) of estimated positions - UP dir.
SDCLK(95%) Standard deviation(95%) of clock offset
SDTZD(95%) Formal uncertainty (sigma 95%) of the estimated Total Zenith Delay
LATDD   Estimated Position - Latitude Degrees
LATMN   Estimated Position - Latitude Minutes
LATSS   Estimated Position - Latitude Seconds
LONDD   Estimated Position - Longitude Degrees
LONMN   Estimated Position - Longitude Minutes
LONSS   Estimated Position - Longitude Seconds
HGT(m)   Estimated Position - Ellipsoidal Height
UTMZONE*   Estimated Position - UTM Zone
UTM_EASTING* Estimated Position - UTM Easting
UTM_NORTHING* Estimated Position - UTM Northing
UTM_SCLPNT* Estimated Position - UTM Scale
UTM_SCLCBN* Estimated Position - UTM Combined Scale
MTMZONE**   Estimated Position - MTM Zone
MTM_EASTING** Estimated Position - MTM Easting
MTM_NORTHING** Estimated Position - MTM Northing
MTM_SCLPNT** Estimated Position - MTM Scale
MTM_SCLCBN** Estimated Position - MTM Combined Scale
ORTHGT(m)*** Estimated orthometric Height

(*) Only for Kinematic mode

(**) Only for Kinematic if user selects NAD83(CSRS) as the reference frame and all positions is in Canada
(***) Only for Kinematic if all positions within the Geoid model

/*********************
 3- CSV file description
 *********************/
latitude_decimal_degree
longitude_decimal_degree
ellipsoidal_height_m
decimal_hour
day_of_year
year
ortho_height_m_CGVD?? : Orthometric height (only for kinematic mode if all positions within the Geoid model)
rcvr_clk_ns : receiver clock offset (ns)

120

6.2 Appendix B

IHO S-44 5TH EDITION MINUMUM STANDARDS FOR SURVEYS (International Hydrographic Organization, 2008)

IHO STANDARDS FOR HYDROGRAPHIC SURVEYS (S-44)
5th Edition February 2008

TABLE 1
Minimum Standards for Hydrographic Surveys
(*To be read in conjunction with the full text set out in this document.*)

Reference	Order	Special	1a	1b	2
Chapter 1	**Description of areas.**	Areas where under-keel clearance is critical	Areas shallower than 100 metres where under-keel clearance is less critical but *features* of concern to surface shipping may exist.	Areas shallower than 100 metres where under-keel clearance is not considered to be an issue for the type of surface shipping expected to transit the area.	Areas generally deeper than 100 metres where a general description of the sea floor is considered adequate.
Chapter 2	**Maximum allowable THU 95% *Confidence level***	2 metres	5 metres + 5% of depth	5 metres + 5% of depth	20 metres + 10% of depth
Para 3.2 and note 1	**Maximum allowable TVU 95% *Confidence level***	a = 0.25 metre b = 0.0075	a = 0.5 metre b = 0.013	a = 0.5 metre b = 0.013	a = 1.0 metre b = 0.023
Glossary and note 2	***Full Sea floor Search***	Required	Required	Not required	Not required
Para 2.1 Para 3.4 Para 3.5 and note 3	***Feature Detection***	Cubic *features* > 1 metre	Cubic *features* > 2 metres, in depths up to 40 metres; 10% of depth beyond 40 metres	Not Applicable	Not Applicable
Para 3.6 and note 4	**Recommended maximum Line Spacing**	Not defined as *full sea floor search* is required	Not defined as *full sea floor search* is required	3 x average depth or 25 metres, whichever is greater For bathymetric lidar a spot spacing of 5 x 5 metres	4 x average depth
Chapter 2 and note 5	**Positioning of fixed aids to navigation and topography significant to navigation.** (95% *Confidence level*)	2 metres	2 metres	2 metres	5 metres
Chapter 2 and note 5	**Positioning of the Coastline and topography less significant to navigation** (95% *Confidence level*)	10 metres	20 metres	20 metres	20 metres
Chapter 2 and note 5	**Mean position of floating aids to navigation (95% *Confidence level*)**	10 metres	10 metres	10 metres	20 metres

15

121

6.3 Appendix C

ZONES OF CONFIDENCE (ZOC) TABLE (ADMIRALTY, 2017)

ZOC [1]	Position Accuracy [2]	Depth Accuracy [3]		Seafloor Coverage	Typical Survey Characteristics [5]
A1	± 5 m + 5% depth	=0.50 + 1%d **Depth (m)** / **Accuracy (m)** 10 / ± 0.6 30 / ± 0.8 100 / ± 1.5 1000 / ± 10.5		Full area search undertaken. Significant seafloor features detected [4] and depths measured.	Controlled, systematic survey [6] high position and depth accuracy achieved using DGPS or a minimum three high quality lines of position (LOP) and a multibeam, channel or mechanical sweep system.
A2	± 20 m	= 1.00 + 2%d **Depth (m)** / **Accuracy (m)** 10 / ± 1.2 30 / ± 1.6 100 / ± 3.0 1000 / ± 21.0		Full area search undertaken. Significant seafloor features detected [4] and depths measured.	Controlled, systematic survey [6] achieving position and depth accuracy less than ZOC A1 and using a modern survey echosounder [7] and a sonar or mechanical sweep system.
B	± 50 m	= 1.00 + 2%d **Depth (m)** / **Accuracy (m)** 10 / ± 1.2 30 / ± 1.6 100 / ± 3.0 1000 / ± 21.0		Full area search not achieved; uncharted features, hazardous to surface navigation are not expected but may exist.	Controlled, systematic survey achieving similar depth but lesser position accuracies than ZOC A2, using a modern survey echosounder [5], but no sonar or mechanical sweep system.
C	± 500 m	= 2.00 + 5%d **Depth (m)** / **Accuracy (m)** 10 / ± 2.5 30 / ± 3.5 100 / ± 7.0 1000 / ± 52.0		Full area search not achieved, depth anomalies may be expected.	Low accuracy survey or data collected on an opportunity basis such as soundings on passage.
D	*Worse than ZOC C*	Worse than ZOC C		Full search not achieved, large depth anomalies expected.	Poor quality data or data that cannot be quality assessed due to lack of information.
U	Unassessed - The quality of the bathymetric data has yet to be assessed				

7 Curriculum Vitae

Candidate's full name: Khaleel Arfeen

Universities attended:

BES, Honours Geomatics (major) & Computer Science (minor) [Co-op] (2009-2014). University of Waterloo, Waterloo, Ontario.

Publications:

Arfeen, K., & Church, I. (2019). Automated Processing Of Arctic Crowd-Sourced Hydrographic Data While Improving Bathymetric Accuracy And Uncertainty Assessment. *USHYDRO 2019.* Biloxi: THE HYDROGRAPHIC SOCIETY OF AMERICA (THSOA).

Conference Presentations:

Arfeen, K. (2019, 03 19). Automated Processing Of Arctic Crowd-Sourced Hydrographic Data While Improving Bathymetric Accuracy And Uncertainty Assessment. Biloxi, Mississippi, USA: THE HYDROGRAPHIC SOCIETY OF AMERICA (THSOA).