Survey Planning Tool: From Concept to Reality

Kent Malone

Canadian Hydrographic Service

Abstract

The Canadian Hydrographic Service (CHS) has the challenge of producing nautical charts and other navigation publications for parts of three oceans, the St. Lawrence River, the Great Lakes and a myriad of other rivers and waterways. Having limited resources and at the high cost of mobilizing surveys, it is imperative for the CHS to deploy those resources in the most meaningful and effective manner. The Data Acquisition and Analysis Committee were tasked with devising a means of addressing these issues on a National scale. The committee proposed an analysis formula and initiated a project to build a Survey Planning Tool using Geographic Information System (GIS) technologies. Information was gathered, integrated and the analysis was performed. This paper looks at the development of the tool, from the conceptualization of the analysis model through the collection and integration of data to final analysis. Areas highlighted are the results, challenges, the blossoming of ideas and the next steps in the evolution of the planning toolset.

Introduction

Canada, bordering on three oceans, has the longest coastline in the world and has numerous navigable inland lakes and rivers. In total, there are over 5,800,000 km^2 of navigable waterways in Canada. The Canadian Hydrographic Service (CHS) has a limited amount of resources to survey these vast stretches of water. Over the past three decades technologies have changed dramatically. Advances in Multibeam, LIDAR and GPS have led to a new definition of 'modern standards'. With the exception of LIDAR, data is still collected by going to sea in ships and boats.

In 2012, the CHS Data Acquisition and Analysis Committee (DAAC) was tasked with developing a methodology to determine how to prioritize hydrographic surveys – where is it best to send personnel and ships? The chosen methodology needed to be national in nature such that survey requirements can be equally treated across all regional offices. The methodology outlined in the document 'NOAA National Survey Plan 2010' was used as a starting point for discussion. Through much deliberation, discussion, trials and analysis, a model was developed.

This model was brought to life using GIS technologies. All necessary data was collected, assimilated and brought into the GIS environment. The processing of the model components were conducted using ArcGIS geoprocessing tools. Statistical information was gathered using ArcGIS measurement techniques.

Survey Planning Analysis Model

The criteria for a suitable model were that the model should be simple and can be applied consistently across the country. The expectation going into the analysis is that there are huge amounts of Canadian waterways that will be found to require new surveys. The challenge was to narrow these vast amounts to something that is practical to deal with. The model that was developed has 3 stages as described in the following sections.

Stage 1 - Navigationally Significant Areas

Charts have many uses including: navigation, engineering and environmental studies, fishing, and offshore resource exploration and exploitation. The scope of this project was narrowed to surface navigation. The primary consideration is: what is the risk of grounding? As such, all waters that are deep and wide are not of concern. Considering the draft of current vessels and expected drafts over the coming decades [CHINAMAX 24m], it was decided that only waters shallower than 50 meters are significant for surface navigation. A polygon built from the shore to the 50 metre contour describes the *Navigationally Significant* area.

CHS charts cover 5,800,000 km^2 of Canadian waters. Selecting 50 metres as navigationally significant reduces the area of interest down to 1,300,000 km^2.

Stage 2 - Survey Deficiencies

The second stage of this model is to determine which areas within the navigationally significant areas as described above are deficient in surveys. This stage of the model, in its simplest form is:

Survey deficiency = desired survey quality – actual survey quality

To evaluate survey quality the S-57 attribute CATegory Zone Of Confidence (CATZOC) was used. CATZOC describes the quality of bathymetric surveys using three parameters: horizontal position accuracy, depth accuracy and seafloor coverage. This changes the model to:

Survey deficiency = Desired CATZOC – Surveyed CATZOC

The first question to be answered is: how is the desired survey quality or desired CATZOC to be determined?

It was deemed that all navigationally significant areas require better than CATZOC C surveys. [CATZOC C characteristics: Low accuracy survey or data collected on an opportunity basis such as on passage.]

The risk of grounding increase with three parameters: an irregular seafloor with a high propensity of rocks, reefs and shoals; a seafloor subject to change through dynamic processes such as silting; and a complex waterway being narrow, restricted and/or sinuous, requiring a high planimetric accuracy. Any area having one or more of the above attributes should be surveyed to

CATZOC A standards. [CATZOC A characteristics: Controlled, systematic, high accuracy survey on WGS 84 datum; using DGPS or three lines of position (LOP) with multibeam, channel or mechanical sweep system.]

Areas that are regular, basin-like, wide, with unchanging seafloor pose a very low grounding risk. Surveys classified as CATZOC B are sufficient for these areas. [CATZOC B characteristics: Controlled, systematic survey to standard accuracy.]

The areas as described in the two preceding paragraphs are intersected with the *Navigationally Significant* areas from Stage 1 to generate the *Desired CATZOC A* and *B* polygons.

The limits of all existing CATZOC A and B surveys are gathered from existing databases. A description of how the existing surveys were evaluated for CATZOC is contained in the section 'Collection of Data'. The survey limits are intersected with the *Navigationally Significant* polygon to develop the *Surveyed CATZOC A* and *B* polygons.

The survey deficient areas are then generated as:

Survey Deficiency CATZOC A = intersection of *Desired CATZOC A* and *Surveyed CATZOC A*

Survey Deficiency CATZOC B = intersection of *Desired CATZOC B* and (*Surveyed CATZOC A* or *Surveyed CATZOC B*)

Stage 3 - Ranking Deficient Areas

It was expected, and found to be true, that the stage two result would be a staggering amount of territory. The total deficient area is 1,200,000 km^2. A method to rank these areas from most to least critical is needed. Two criteria were introduced: marine traffic and search and rescue capabilities.

In the interest of narrowing the potential survey areas, only commercial traffic was considered. All other traffic (recreational, fishing, government fleet, light commercial) were not considered. As a further refinement, initially only two classes of commercial vessels were considered: ships carrying hazardous cargo and ships carrying passengers. In the event of a marine incident, these two categories of vessels are considered to the produce the most dire consequences: potentially heavy loss of life and extreme damage to the environment.

The definition of ships carrying hazardous cargo was limited to tankers carrying oil, oil products and LNG (Liquefied Natural Gas). This includes small tankers serving coastal communities. There are many other types of hazardous materials which are transported by other means, for example as containerized cargo. These were not considered.

The definition of passenger vessels was restricted to vessels carrying more than 50 people. This excludes recreational and excursion vessels.

During the first implementation of data into the database, it was determined that this model did not hold up in the Arctic. There are communities for which no explicit tanker nor passenger traffic could be identified and yet standard freight traffic was evident. Therefore a 3rd category of general cargo had to be included.

The marine traffic was identified as routes only. There was no attempt to classify routes by volume of traffic or gross tonnage or any other parameter. As a consequence, the route to Corner Brook has the same weight as the route to Vancouver.

The routes are developed as polygons. A description of how these polygons were developed is given in the section 'Collection of Data'.

Search and Rescue seems out of context when discussing where to conduct hydrographic surveys. This was viewed as taking proactive action. The objective of this survey planning exercise is to determine where the need for surveys is the greatest with respect to preventing groundings. Clearly the need for surveys is greatest where the infrastructure to respond to groundings is the weakest. Conducting surveys and providing accurate navigation products will help to reduce the risk of marine incidents in remote areas. Search and Rescue polygons were developed depicting a relative scale of good, fair and poor response capabilities.

The deficient survey areas were ranked following the rules outlines in the table below:

Rank	Traffic	Deficient	SAR
1	Hazardous+Passenger	A or B	Poor
2	Hazardous+Passenger	A or B	Fair
3	Hazardous+Passenger	A or B	Good
4	Hazardous	A or B	Poor
5	Hazardous	A or B	Fair
6	Hazardous	A or B	Good
7	Passenger	A or B	Poor
8	Passenger	A or B	Fair
9	Passenger	A or B	Good
10	Commercial	A or B	Poor
11	Commercial	A or B	Fair
12	Commercial	A or B	Good
13	Other	A or B	Poor
14	Other	A or B	Fair
15	Other	A or B	Good

The result of this analysis is that most of the areas deficient in surveys fall in the ranking 13, 14 and 15. Considering only the top 12 rankings, this model has reduced the area of most concern from 1,200,00 km^2 down to 118,000 km^2.

Collection of Data

The Survey Planning model uses a number of varied data themes. Each theme has its own sources and methods of compilation.

Shoreline and land polygon

In designing the database, there were two items that impacted the selection of data. One was that the geoprocessing would be done by comparing polygons. It was therefore necessary to have land polygons from which to construct the other polygons. The second was an objective that the planning tool be equally useful on a national level and on a very local level. To satisfy both requirements, it was decided to use National Topographic Data Base (NTDB) 1:250,000 data to construct the land polygons. Over 400 NTDB datasets were merged together to create the land polygons.

There are issues with the storage size of a single polygon having a resolution corresponding to 1:250,000 and covering the entire Canadian landmass. Initially the country was divided into 15 areas. While this worked well when working with the data as local shape files, performance was extremely poor when accessing the same data from the database across the network. The land polygon was then reduced in resolution by a factor of 24. Total size dropped from 50G to 2.4G. Performance improved but only marginally. A raster version of the land was tried. Again there was only a limited gain in performance. This continues to be an issue.

Navigationally Significant area

The definition in this model of navigationally significant waters is all waters less than 50 metres depth. The CHS is the holder of bathymetric data and one would assume that this data is readily accessible. Such is not the case. For the most part, the data resides in individual datasets each of which corresponds to a field survey and covers a small geographic area. There are on-going efforts to load the data into the Bathymetric Data Base (BDB). At the time of the initial loading of the Survey Planning database, this effort was far from complete. As a consequence, the 50 metre contour was obtained from various sources.

Quebec Region's BDB was fully loaded and the 50 metre contour was extracted from this database. The contour is as accurate as the data contained in the database. For the most part the underlying data is good to very good with certain areas meeting only CATZOC C standards. Overall, the results were good and the resulting polygon can be considered to be reasonably accurate. In one area there were some very small polygons in off-lying areas which are probably artifacts of the process.

Pacific Region provided a digital polygon describing the 0 to 50 depth area. This polygon was constructed as a merge of polygons from various digital chart files. The resolution of the data was very high, higher than needed. The agreement of the provided polygon to the land file constructed from NTDB data ranged from excellent to poor. Intertidal areas were captured as holes in the depth polygon. These were unwanted.

Central and Arctic Region has no readily accessible data so provided NOAA GEODAS data for the Great Lakes and GEBCO data for the Arctic. In the Arctic there exists spotty bathymetric data, therefore the data provided is, a best, an interpretation.

The Atlantic Region 50 metre contour was digitized from medium scale charts (~1:250,000). The quality is suitable for the survey planning tool.

Waterway Characteristics

There are no known databases which capture the three waterway characteristics: irregular seafloor; seafloor subject to change; and complex waterway. A suggestion was made to develop mathematical metrics for these three categories. This was abandoned as it was recognized that there was not the necessary data to evaluate any such metrics. Instead, each regional office was asked to develop polygons representing each of the three categories based on the knowledge of experienced hydrographers. Descriptive guidelines were written for each category. Polygons were delivered and subsequently loaded into the database.

Existing Survey Limits

The Bathymetric Data Base (BDB), once fully populated, should provide all necessary data for this category. At the outset of this project the BDB was not fully populated. Subsequently, the data was collected in a variety of fashions.

Quebec Region data was fully supplied directly from the BDB. The limits of surveys categorized as CATZOC A and CATZOC B were provided.

Pacific Region was partially supplied from the BDB. The remainder was collected by extraction from the CHS metadata database CHSDIR as described below.

Atlantic Region and Central & Arctic Region data was extracted entirely from CHSDIR.

The database CHSDIR contains information about surveys including survey limits. Metadata of surveys for a particular area were extracted into a spreadsheet. CATZOC has only recently begun to be captured in CHSDIR and as such is nowhere close to complete. To assess the CATZOC of surveys the following fields were used: depth method, position method, scale and year. It was not possible to assess line-spacing (seafloor coverage). Guidelines were developed for each of the four fields. All sources designated as CATZOC A or B were retained, all others were removed. The limits of the surveys were extracted from the database as kml (Keyhole Markup Language) files. In many instances the polygon limits were not the limits of the surveys but were the four corners of the physical sheet. As a consequence, this did not always give a good representation of the area surveyed.

Search and Rescue

It was desired to have a representation of the relative classification of search and rescue across the country. To accomplish this, only the standing Search and Rescue resources across the country were evaluated. This included SAR helicopter resources, life boat stations, inshore rescue boat stations and Canadian Coast Guard (CCG) vessel area of responsibility. A simple three-tier classification was developed:

Good: Within range of helicopter support without refueling and within range of an inshore rescue boat or a life boat or within 12 hours of CCG vessel.

Fair: Within range of helicopter support with 1 refueling.

Poor: Less than fair.

Helicopter range was calculated as 2/3 maximum flight distance. This allows for some time on site before refuelling. CCG vessel range was based on 20 knots. This analysis does not take into consideration other resources that may be available such as CG Auxiliary, RCMP or regional police resources.

The result was not surprising, SAR response capabilities are better in the southern areas of Canada and diminish farther north.

Marine Traffic Routes

Collecting marine traffic information proved to be challenging. The MARIN research group at Dalhousie University has compiled a large amount of information in database form. They were contracted to provide marine routes for all Canada, classified by the original two categories, hazardous and passenger cargoes. This information was ameliorated by local knowledge, information from ferry companies, internet research on specific ports, information captured in CHS Sailing Directions and information obtained from INNAV, the acronym given to the Vessel Traffic Management Information System of CCG. At a later stage in the project it was realized that information would be required for general commercial vessel traffic as well hazardous and passenger traffic. This information was researched using INNAV for one shipping season for the Arctic only.

All traffic routes were developed as polygons. Charted information was used to establish some of these polygons. This included: routes normally travelled, maintained channels, and traffic separation schemes. In most cases, a linear route was converted to a polygon by applying a nominal width. The width applied varied with the body of water being traversed. In some instances it was very difficult to determine routes actually travelled.

Data Integration and Processing

The GIS system chosen was ArcGIS. The data was gathered locally as shape files and later imported into a SDE (Spatial Database Engine) database. All of the geoprocessing was done as set manipulation (intersection, union, merge, dissolve) of polygons. As a consequence, all of the thematic layers specific to the model had to be constructed as polygons.

Data Integration

Data was received in many formats. The versatile Global Mapper software package proved to be invaluable. It was used to read the variety of native formats and export the data into shape files.

In some instances, the datasets contained more data than required. For example, the hydrographic layer of the NTDB was used to develop the shoreline and land features. This hydrographic layer contains many unwanted features: lakes, ponds and rivers. All of these had to be identified and removed by hand.

In ArcGIS, the data can be displayed in any projection desired, changing on the fly. However, the data is stored in relation to a projection or georeferencing system. For ease of data transferability, data georeferenced to WGS84 is preferred. This georeferencing system is rectangular and as such there is a large amount of distortion, particularly in the Arctic. Initially, Polyconic was selected as the reference projection. Polyconic provides the most realistic view of the large extent of the Canadian territory, especially the Arctic areas. During implementation, performance became an issue. Web displays are based on World Mercator. The data was transformed to World Mercator to eliminate the need of the software to transform the data on the fly and hopefully, but with limited success, to improve performance.

With the ability to use various projections on the fly and having datasets covering large geographic areas, it is necessary to ensure that there are sufficient vertices to maintain the correct geometry for any object in any projection. With an insufficient number of vertices, not only will the shape change from projection to projection, but the georeferencing processes will yield different results. For many of the data layers, it was necessary to introduce vertices. As a minimum, vertices were added at 15 minutes in the east-west direction and 1 degree in the north-south direction.

By contrast to objects described in the previous paragraphs, there are some objects that had too many vertices and required generalization. The most flagrant of these are the polygons output from BDB. These polygons appear to follow a rasterized version of the survey area. The result is that straight lines are represented by an unnecessary saw-tooth pattern.

A certain amount of attribution was desired for every data theme. A small amount of metadata was collected as attribution, this included year of data collection, relative accuracy of the data and data source. Whenever the data relates to other data sources or databases, a reference number belonging to the native source was retained to allow cross-reference. To be able to provide more detailed information, other attribution was captured. For example, the survey layer contains the following attribution: depth method, position method, scale, year of survey and collection agency.

Desired attribution did not always transfer directly with the files. In some instances supplementary information in spreadsheet form was imported as tables. The tables were joined to the shape files and the attribution transferred as attribution to the shape files. In extreme cases, information was researched and entered one record at a time.

Certain information types required the transformation of point and/or line data into polygons. In some cases this was done purely by interactive techniques. In other cases, ArcGIS tools such as polyline to polygon and drawing a buffer around a line were used.

ArcGIS Tools

The ArcGIS environment proved to be an excellent platform on which to integrate and process the data. There is a wide variety of tools, many were taken advantage of, the great majority not yet explored. Building and modifying objects is extremely easy and dexterous. There are limits, however. The entire country of Canada at a resolution of 1:250,000 is somewhat taxing, but there are tricks to make even this reasonably easy.

Many of the ArcGIS tools were used in the integration of the source materials. As previously mentioned these included polyline to polygon and building a buffer around a line.

The tools used for evaluating the Survey Planning Tool model itself are the geoprocessing tools. These tools perform set functions (merge, intersect, union and dissolve) on polygons. Despite the very large polygons being processed, these processes ran in a matter of minutes. As previously discussed in ‘Data Integration’, care has to be taken to use the appropriate projection or unexpected results will occur. Sometimes there will be some clean-up work required. Small polygons and slivers may be created.

Area (kilometer^2) was calculated for many of the layers. Again a simple tool was employed: geometry calculation. One has to be careful to switch to an equal area projection prior to performing the calculations. A trial was conducted where areas were calculated based on different projections: equal area, Polyconic, world Mercator and georeferenced. Calculations based on polyconic were within 5% of equal area while the other two proved to be totally erroneous.

To make the database available on a national level, the data was installed in a SDE database (Spatial Database Engine). The data and results are accessible but the performance was so slow as to be painful to use. Attempts have been made to improve the performance. Drastic strides are required but to date results have been less than satisfactory. Most of the attempts have been with the data itself. The relatively high resolution of the land polygons appears to be one of the factors. The resolution was drastically reduced and some improvement was realised.

Survey Planning Tool Results

The results obtained can only be taken as preliminary due to the less than perfect quality of some the data, however, the results are indicative of reality and the methodology has been proven to give results.

Canada has an enormous amount of territorial waters and navigable inland lands and waters. The methodology employed provided a method to narrow the focus to where surveys are truly required. The following table displays this narrowing of focus. From a potential of 5,800,000 km^2 of area to survey, the methodology has reduced the immediate concerns to 118,000 km^2. In terms of survey resources, this drops an unimaginable 6,000 vessel seasons down to a manageable 100 vessel seasons (approximations).

Region	EEZ	Navigationally Significant	Deficient	Top 12 Rankings
Atlantic	1,601,919	131,568	108,650	31,612
Quebec	141,352	28,193	22,195	2,338
Central	137,243	88,568	68,393	465
Arctic	3,537,689	1,048,326	1,003,686	83,070
Pacific	455,869	26,550	20,955	1,154
Grand Total	5,874,072	1,323,205	1,223,878	118,639

Area expressed in km^2

Much of the country is deficient in surveys. The deficiencies are most prevalent in the northern waters, the Arctic and Labrador. However, there are large areas in the south which are deficient as well and these southern areas have much more commercial and other types of traffic. Cases in point are the Northumberland Strait and Lake Superior, both of which have charts based on data dating from the 1950's.

Through this development process it became evident that a GIS is indeed the right tool for this job. Notwithstanding the performance issues related to the database, the GIS provides all the necessary integration and geoprocessing tools, presents a vivid graphical image of the situation and can quickly deliver good statistical information.

The demonstration of the tools has spawned a number of other ideas:

- Identify potential areas for LIDAR surveys by identifying depth areas less than 20 metres.

- Determine what type of survey is best suited where: ship-based, shore-based, LIDAR, through-ice.
- Develop survey plans for a specific site. For this survey, identify what can be surveyed by LIDAR, where CATZOC A techniques will be required, and where CATZOC B techniques are sufficient. Calculate the square kilometers of survey and thereby determine the level of effort required for each. Determine if the survey local is best suited for ship or shore based deployment.

Challenges

There have been many challenges faced in the building of the Survey Planning Tool and many more challenges to be faced and resolved. The following is a brief summary of some of the challenges.

Much of the desired data does not reside in a form that is easily used. This takes time, the data has to be painstakingly put together, piece by piece.

Canada is huge. Whenever an information type has to be chased down every bay, fiord, river and lake, the level of effort to incorporate the data increases dramatically.

The information on the existing surveys is to date incomplete, sometimes not properly categorized by CATZOC, and often the polygon limits are either too general or too complex. There are concerted efforts to populate each region's BDB. Once this is fulfilled, the survey information will be totally reliable.

The tool must be available to all managers across the country. The performance issues must be resolved. The toolset must be completed. Training and support must be delivered to all users.

The database must be kept up-to-date. Efforts are being made to update the information directly from other databases, specifically CHSDIR and BDB.

Next Steps

Items identified under the previous section, Challenges, are actively being worked on and will continue to be worked on until desired results are achieved.

Currently there is an effort underway to extend the model to include products. The intent is to use the one tool for all planning: surveys and production. The expected outcome is to be able to provide and maintain a medium term (5 year) plan. At the moment of writing, this model is in the conceptualization stage. Once an acceptable model has been agreed upon, the necessary supporting information will be researched and integrated, the model run, results evaluated and adjustments to the model made as necessary.

References

NOAA, *NOAA Hydrographic Survey Priorities*, 2012 Edition

IHO Special Publication S-57, *IHO Transfer Standard for Digital Hydrographic Data*, Appendix A, Chapter 2 - Attributes