

Advances in Hydrographic Data Processing: Time for a Paradigm Shift

Jonathan Beaudoin

Chief Scientist, QPS Canada Inc., Fredericton, NB, Canada, beaudoin@qps.nl

Moe Doucet

Chief of Software Research, QPS USA, Portsmouth, NH, USA, doucet@qps-us.com

Abstract

Hydrographic data processing can be challenging, even for experienced users. The problem is often the human operator. Humans make errors when transcribing vessel configurations from one software application to another or from one coordinate frame convention to another. Humans make errors when importing ancillary data and then failing to associate it with the correct data files. Humans make errors when changing processing configurations and not then performing the appropriate reprocessing. Any error along the way leads to poor results in the final product and many wasted hours troubleshooting the source of the error.

A paradigm shift is a fundamental change in approach or underlying assumptions. With the release of QPS Qimera, we are striving for a paradigm shift in that we are automating the mundane and error prone tasks for which computers are well suited but humans are not, for example data transcription, processing state management and job scheduling. Qimera isolates the stages for which a human brings value to the process, for example processing configuration management and data validation. In this paper, we explore the methods used by Qimera to guide the operator through data processing, simplifying what has traditionally been a convoluted process.

Biography

Jonathan Beaudoin has a PhD (2010) in Geodesy and Geomatics Engineering from the University of New Brunswick. He is the Chief Scientist of QPS in Fredericton, NB, Canada and is an Affiliate Assistant Professor at UNH-CCOM and an Honorary Research Associate at UNB.

Maurice Doucet has a B.S. in Computer Science from Chapman University. He is the Chief of Software Research of QPS and was the lead developer on a number of products for the hydrographic community including Qimera, FMMidwater, and FMGT. He has lead efforts to produce over 15 commercial software products in the 2D/3D/4D graphics industry.

Introduction

Hydrographic Processing can be hard. It's true. Even for experienced users. Mistakes happen. Safeguards catch them. Mistakes are fixed, sometimes at great cost. But they happen again and again, perhaps with different permutations, project after project. Many of the frustrations are due to the fact that the human operator must connect all the pieces together to come up with the final processing solution.

What kinds of errors are humans making to cause these problems? Some examples include:

- Typos when transcribing vessel configurations from one software application to another or from one coordinate frame convention to another.
- Failing to import necessary ancillary files.
- Importing ancillary data and then failing to apply it to the correct data files.
- Changing processing configurations and not understanding the scope of reprocessing that is required (i.e., which files must now be reprocessed) or simply forgetting to trigger the appropriate reprocessing for the files that need it.
- Changing processing configurations and not appreciating the dependencies of intermediate computed values that must now also be reprocessed, or not triggering the reprocessing of intermediate values in the correct order.

Why are humans not good at these? Data entry is always prone to error, this is a well-established fact in the field of human computer interaction (HCI). For complex operations that require multiple steps in a very specific order, humans are prone to error especially in high stress environments, or in situations where fatigue is a factor or when working in a highly distractive environment.

In the hydrographic surveying industry, the concept of Total Propagated Uncertainty (TPU) helps us understand sources of uncertainty from particular sensor measurements. It also allows us to appreciate the mathematical model used to combine and propagate their total combined effect to allow for an estimation of the uncertainty for the soundings that we compute from them. Each sensor contributes its own share towards an uncertainty budget and understanding TPU allows us to make quantitative assessments of the quality of the data that we collect. It also allows us to identify the leading sources of uncertainty and to make good decisions on where to focus our attention on making improvements, either in equipment or in survey practices, to achieve a desired result. The concept of TPU is well established and forms a significant part of the basis for internationally recognized survey standards such as IHO S-44 (IHO, 2008).

What these types of standards don't address is the effect of random human error. If sensor measurements must be combined mathematically by humans to achieve a result, then surely the inevitable human error in effecting the combination will have an effect on the TPU of the resultant soundings. Human error, in these situations, can take the form of "blunder" type errors that can easily exceed the uncertainty levels of the individual measurements. For example, applying the incorrect tide file would result in an uncertainty on the order of the tide magnitude itself, as opposed to the much smaller uncertainty in the tidal measurement itself had the correct tidal corrector been applied.

Just like total propagated uncertainty (TPU), human error can compound with each mistake. Imagine a data set from the last project you worked on and giving it out to twenty different data processors. What is the likelihood of getting twenty sets of results that are all consistent and correct? Slim. Now imagine a slightly different scenario. Imagine taking the same data set and using your traditional software and workflows and over the course of a twenty day period processing the same data set, from scratch, every day. Would you get the same result every day? We would like to think so, but it's not likely. Humans are as much as source of uncertainty in our work as the measurements themselves. We must acknowledge the role of "Human TPU". If you don't recognize the role of Human TPU, you may have great measurements but your output deliverables may not reflect that. Even worse, you may not even realize that you have poor results. A good many of us spend early parts of our careers in this first situation. There's a reason you send seasoned personnel on important jobs.

What does it take to beat the problem of Human TPU? You can hire great people and train them well and keep them trained well. You can develop Best-Practices, you can institute Standard Operating Procedures (SOPs). You can use small and simple ideas like check-lists and "cookbooks" and "cheat sheets". You can put together great onboarding and training procedures for new staff. You can designate staff to focus solely on Quality Assurance and have them institute procedures and mechanisms to ensure high quality outcomes. Perhaps the same staff are tasked with compliance monitoring.

You can document workflows using mechanisms like Standard Operating Procedures (SOPs) and instituting training protocols. These often include check-lists, screen shots, explanatory documentation, etc. These are effective but they place a burden on the data processing effort. They must be maintained and kept up to date. This incurs cost. If a process is worth documenting, then it is best supported with training. More costs. If it is worth documenting and training, then it is worth monitoring for compliance. More costs. The personnel who are assigned to compliance monitoring have their own training and SOPs, often with the same software tools that the data processors use. More costs.

Once you have enough documents, you run into the problem of version tracking, collaborative updating and dissemination. Updating a document is only one part. The staff who actually use the document must be trained on the new methods. More costs. Some go as far as to include the staff using the SOPs as stakeholders to review or even advise the updating procedure. More costs.

There is a large and often unacknowledged hidden cost to complex workflows in that a significant amount of human effort must be expended just to make sure that other humans don't make mistakes. For every person that you hire to process data, you need to pay what amounts to a tax on their time, perhaps best expressed as an overhead cost. For every unit of data processing effort, you may need 0.5 units of QA effort, for example. You hire humans to conduct an activity, then you hire more humans to check their work. You may have a great result, but it took a long time, a lot of people or a lot of money to get to it.

These are all great practices to mitigate the impact of human error, but really this is a case of *working harder* when we could be instead *working smarter*. How can we work smarter? We

need to recognize that human error is a big part of the problem. People don't set out to make errors but they happen nonetheless, despite our best efforts. They are inevitable and software should be designed to eliminate or substantially reduce human error.

Improving Workflows and Outcomes: In Theory

As an industry, we have routinely solved technical problems with better instruments with new capabilities, new processing routines and sometimes even better algorithms. But nobody has really looked at the problem of reducing human error. Until recently, it has been accepted that it takes highly trained personnel to achieve results, and even with highly trained personnel, it is far too easy to make mistakes. Handling mistakes takes significant QA procedures and resources, as was pointed out in the introduction.

What can we do differently? We can begin by recognizing that some things are best done by computers and other things are best done by humans. Of all the work required to get from a set of raw sensor measurements to arrive at an accurate georeferenced sounding, in which parts of this process do humans bring value?

There are several obvious examples where humans win out, at least for the moment:

- Data integration troubleshooting: Identifying causes of errors and taking appropriate remedial action.
- Processing configuration management: Managing the “recipe”, not the process, to get the results you want. For example, knowing that you'd like to use GPS for vertical referencing instead of tides.
- Data validation: Knowing good data from bad. There are many automated routines to remove echosounder blunder detections, however, a human is almost always required to vet this work.

Looking at the other side, computers handily outperform human operators on more mundane tasks such as:

- Data entry and transcription: Numerical data entry and copying from one format to another is a particular weakness for humans.
- Unit conversion: Converting survey measurement units, for example from meters to feet, can lead to a significant source of error.
- Coordinate frame transposition: The coordinate frame of a ship reference survey may not be in the desired coordinate frame that the survey software requires. The same can be said for the various reference frame conventions for the sensors deployed on a vessel. Matching the x-axis measurements from one to the y-axis of the other, for example, is an error prone process. Opposing sign conventions further compounds this problem, for example the z-axis being positive upwards in one reference frame and positive downward in another.
- Processing state management and Job Scheduling: Importing ancillary data and/or making a change to a desired processing configuration change can lead to error if a human is responsible for triggering a set of processing in a particular sequence.

With these exact points in mind, we have built QPS Qimera from the ground up to provide a fully functional multibeam sonar processing application that is both easy to use and also specifically designed to reduce human error. The design vision of Qimera is to make hydrographic data processing intuitive and as simple as possible, all the while offering powerful capabilities to those that need them without cluttering the workflow for those that do not.

This vision is achieved by Four Design Pillars that touch on the points considered earlier where humans are prone to error. The Design Pillars are listed below in the order that a new user would experience them in Qimera, with a discussion immediately in the following sections, along with implementation examples from Qimera.

- **Pillar 1: Deep and Intelligent Data Extraction.** Qimera extracts as much as possible from raw sonar files so that you do not need to enter vessel configuration information, sound velocity profiles, etc.
- **Pillar 2: Guided Workflow.** After data is imported, Qimera guides you through the necessary stages to arrive at a map and to begin validating your results.
- **Pillar 3: Processing State Management.** Once you have a map in hand, you may choose to make processing configuration changes or edit some data. Qimera manages which files are affected by the changes and documents the type of reprocessing that is required and for which files it is required.
- **Pillar 4: Dynamic Workflow.** Any reprocessing or data validation automatically updates your soundings for you as well as any maps that include those soundings.

Pillar 1: Deep and Intelligent Data Extraction

Many modern multibeam data file formats contain the raw sensor records from the multibeam itself (range and angle measurements, primarily) as well as the records from ancillary sensors required to accurately georeference the sonar measurements. This includes positioning measurements and orientation measurements. Some formats include the sound velocity profile used for acoustic ray path calculations in real-time. Some also include vessel configuration information to capture the location of the sensors in the vessel reference frame. Some file formats accommodate vertical referencing information as well, such as tide or GPS/GNSS height. Files that contain all of this information have the benefit of providing a complete record of all information required to geo-reference echosounder measurements in the same file. In Qimera lingo, we refer to these types of files as “rich” files, in that they are rich in their information content. With rich data files, Qimera can determine what to do for processing and the user does not need to intervene.

Vessel Information and Setup

With rich data file formats, a post-processing experience can be built to take complete advantage of the fact that the incoming data file records everything that is required to process the data. Qimera can extract the vessel name and sonar serial number and can create a vessel configuration file for the incoming data. File formats should contain enough metadata about the survey platform and sensor suite to enable automatic recognition of files coming from the same

platform versus files coming from another platform. The image below shows the result of adding an entire directory of QINSy .db files and Kongsberg .all files from a variety of platforms loaded into a single project in Qimera and then grouped by vessel. These two file formats are rich and provide Qimera the information it needs to easily identify what platform any given file derives from.

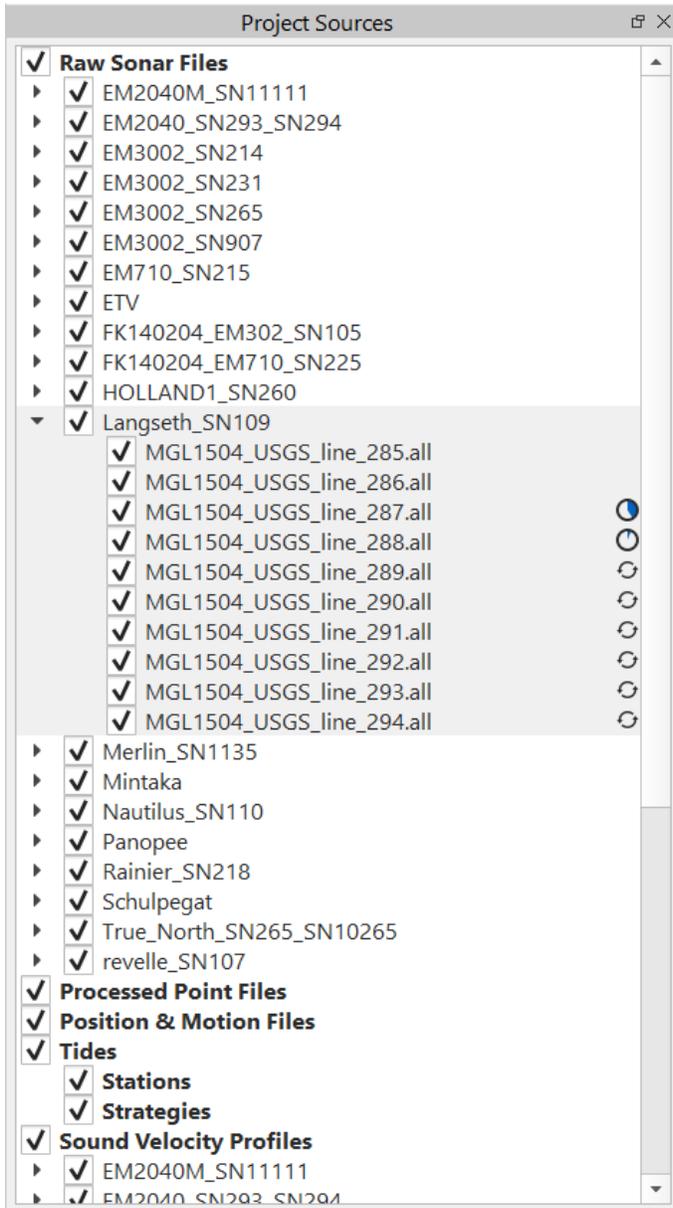


Figure 1. Qimera Project Source view showing raw sonar file listing after import of several files from multiple platforms. Files are grouped by vessel.

The .db and .all format also provides the vessel configuration that captures the location and orientation of all sensors that are represented in the file, as shown in the Vessel Editor figure below for the *Langseth* file grouping from Fig. 1. Note that the angular and linear offsets for all systems are retrieved from the incoming file. At this stage also, Qimera takes care of the

differing coordinate frame conventions and measurement units and converts them to a common frame and unit system on import. This completely eliminates the human error associated with creating a vessel configuration file.

All vessel configuration information is stored internally in Qimera on a file-by-file basis. If, for some reason, the vessel configuration changes throughout the course of a survey, the vessel editor will show each change as a unique and new time entry associated with the change. If, on the other hand, all vessel configurations are the same, then the Vessel Editor will only show a single entry that is reported at the time of the first occurrence of the configuration. If you want to make a change in configuration, you can easily add a new time entry. If you have a survey where the configuration changes routinely, you will see multiple entries in the Vessel Editor.

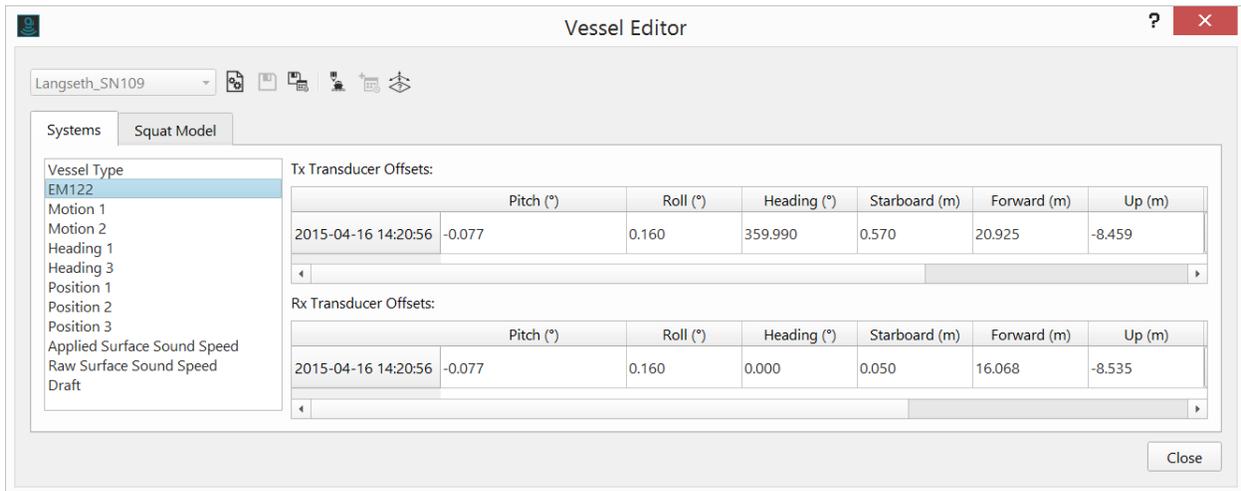


Figure 2. Qimera Vessel Editor showing sensor listing and properties, as decoded from the incoming data file.

Previous to the release of Qimera, data processors using competing software packages were responsible for not only creating and configuring the vessel configuration file in the post-processing application, but they were also responsible for correctly associating every incoming data file with the correct vessel file every time they import data. The opportunities for human error to be introduced are numerous. With rich file formats and the use of Qimera, this source of error is completely eliminated. Surveyors who take the time to “get it right” during acquisition will be rewarded with an error free data import and vessel configuration experience when post-processing in Qimera.

Sound Velocity Profiles

Qimera will also extract and import sound velocity profiles (SVP) from rich data files that have them, as well as metadata about the SVP like the time of acquisition and the position of the SVP if this information is available in the file format (note that almost all formats are missing one or the other). Qimera will extract the record(s) from each incoming multibeam file but it does a second pass to determine if the SVP has already been included in the project by comparing the SVP to those from a previous import session. Only unique instances of SVPs will show up in the project. For a survey with 100 multibeam files, for example, users do not want to manage 100 instances of the same SVP; this second step of determining if an SVP is unique or not is an

important one. Once imported, the SVP casts can be grouped by vessel to ease management, as shown in Fig. 3. The SVP files can be edited and their metadata updated as well, as shown in Fig. 4. Having the SVP automatically imported into the project is another case where Qimera completely eliminates the human error associated with finding the SVP file, converting it to the native format required for post-processing and then loading it into the project. This depends, of course, on the surveyor taking the time to load the SVP information into the acquisition system.

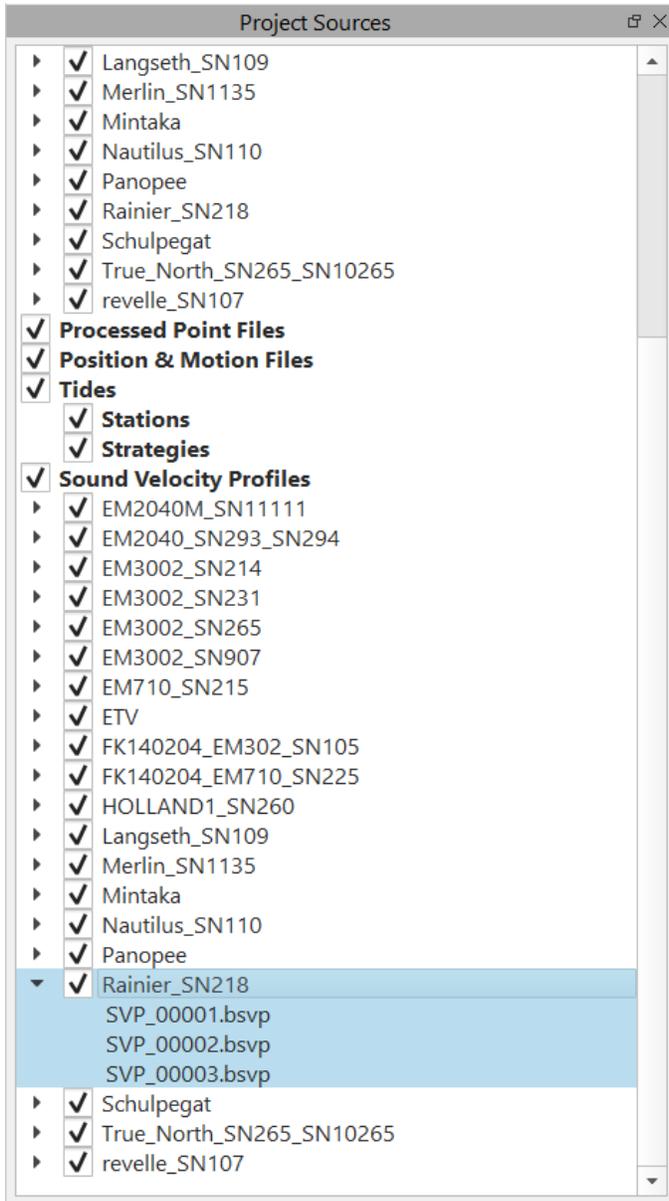


Figure 3. Qimera Project Source listing of SVP casts associated with a particular vessel.

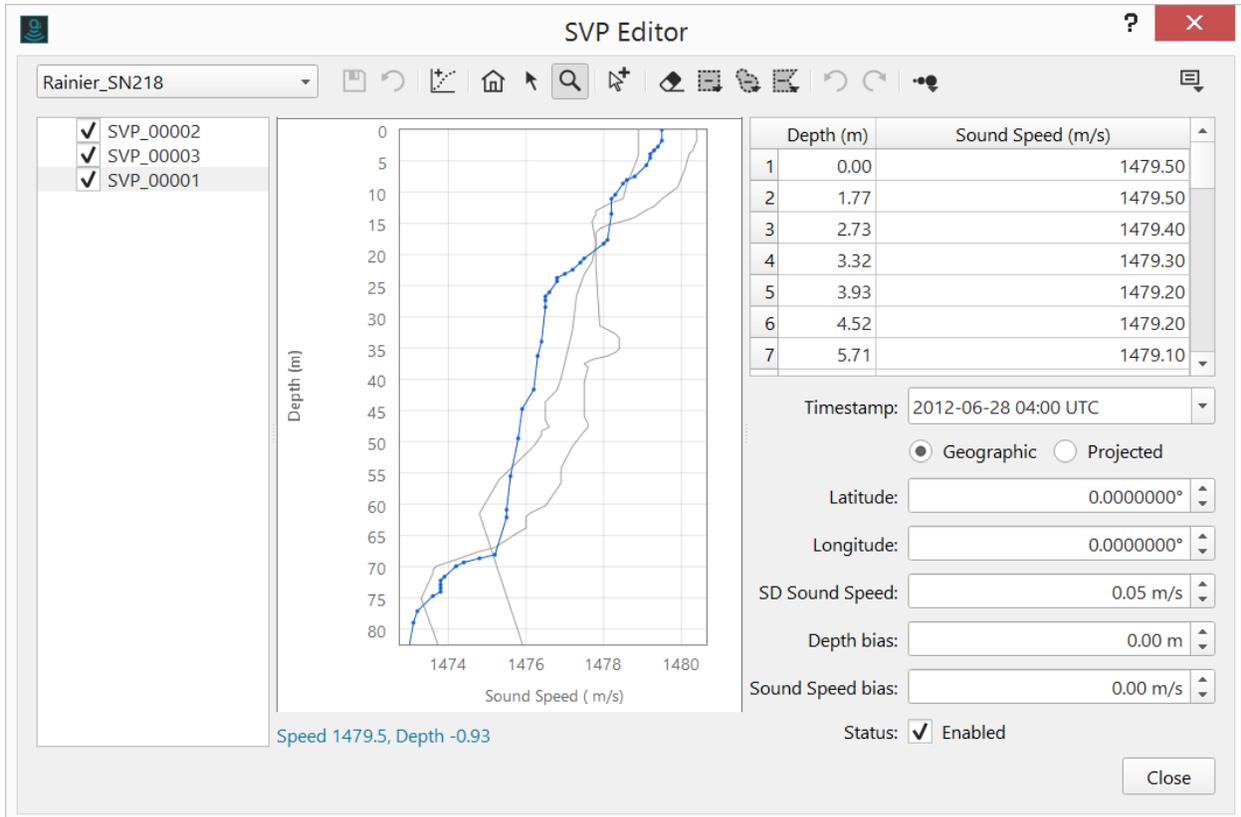


Figure 4. SVP Editor view of the SVP files highlighted in Fig. 3. On extraction of the SVP, as many of the metadata fields in the lower right panel are filled out as possible. The metadata provided for SVP varies quite a bit from one file format to another and it is worth exploring what your preferred file format supports.

To complete the elimination of human error on import of SVP, Qimera goes one step further and encodes the dependency of any given multibeam file on a particular SVP cast in the Processing Configuration for the file, as can be seen in Fig. 5. After import, Qimera is aware of the new SVPs and it is aware that particular multibeam data files depend on this SVP for processing. No further human interaction is required for SVP unless a change in configuration is deemed necessary.

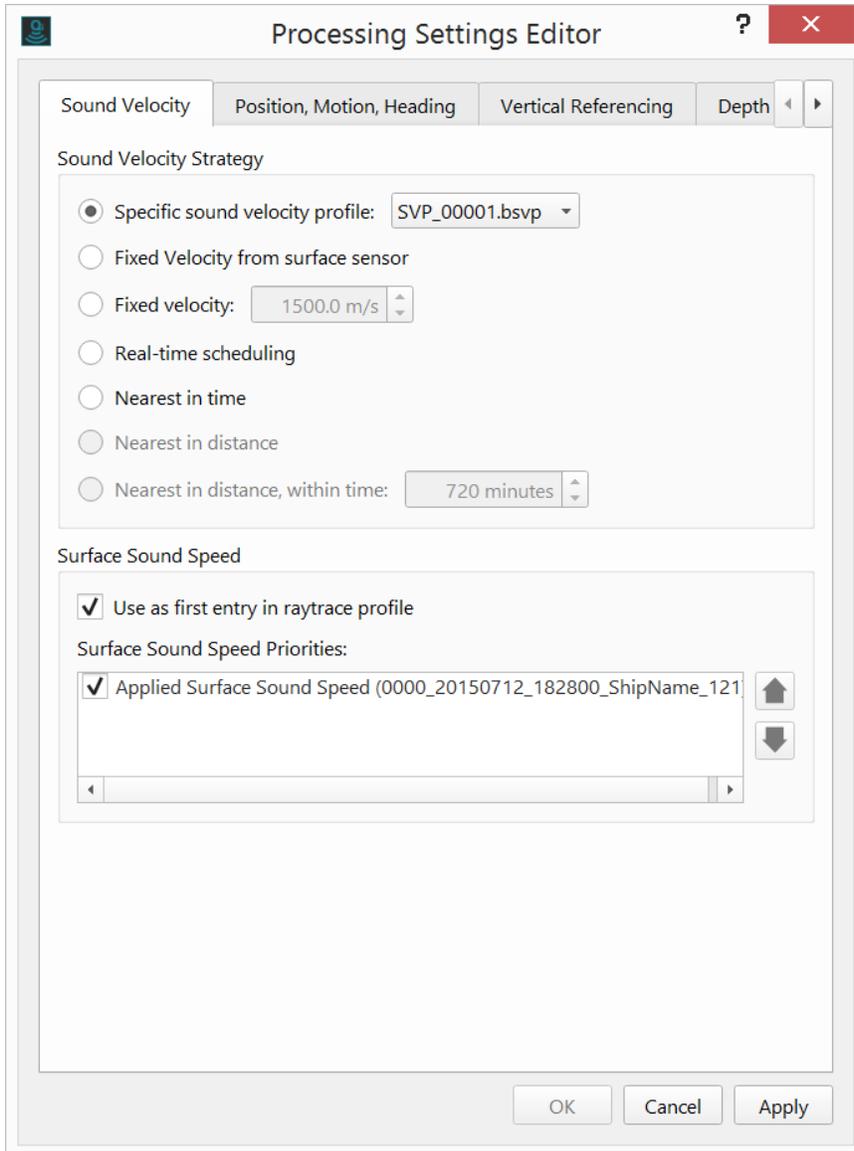


Figure 5. Processing configuration setup for SVP for a particular multibeam survey line that had a single SVP file decoded from it during import. For files without SVP, Qimera defaults to use the velocity reported in the multibeam packets, this is typically from a surface probe and is used primarily for beamforming and beam steering.

Ancillary Data Streams

Qimera automatically decodes and imports all ancillary data streams that are available in a file. For rich file formats that support multiple data streams, Qimera also decodes the preferred observation system that was configured in real-time by the surveyor. An example is shown in Fig. 6. where Position system #3 was chosen as the priority system by the operator. Qimera lists all sources of ancillary information, and it encodes the real-time prioritization selected by the surveyor. This completely eliminates human error in making these selections during the import stage or during post-processing. Users may of course choose to swap priorities after the fact, but the default always matches the surveyor's configuration such that no action is required from the user after the import stage.

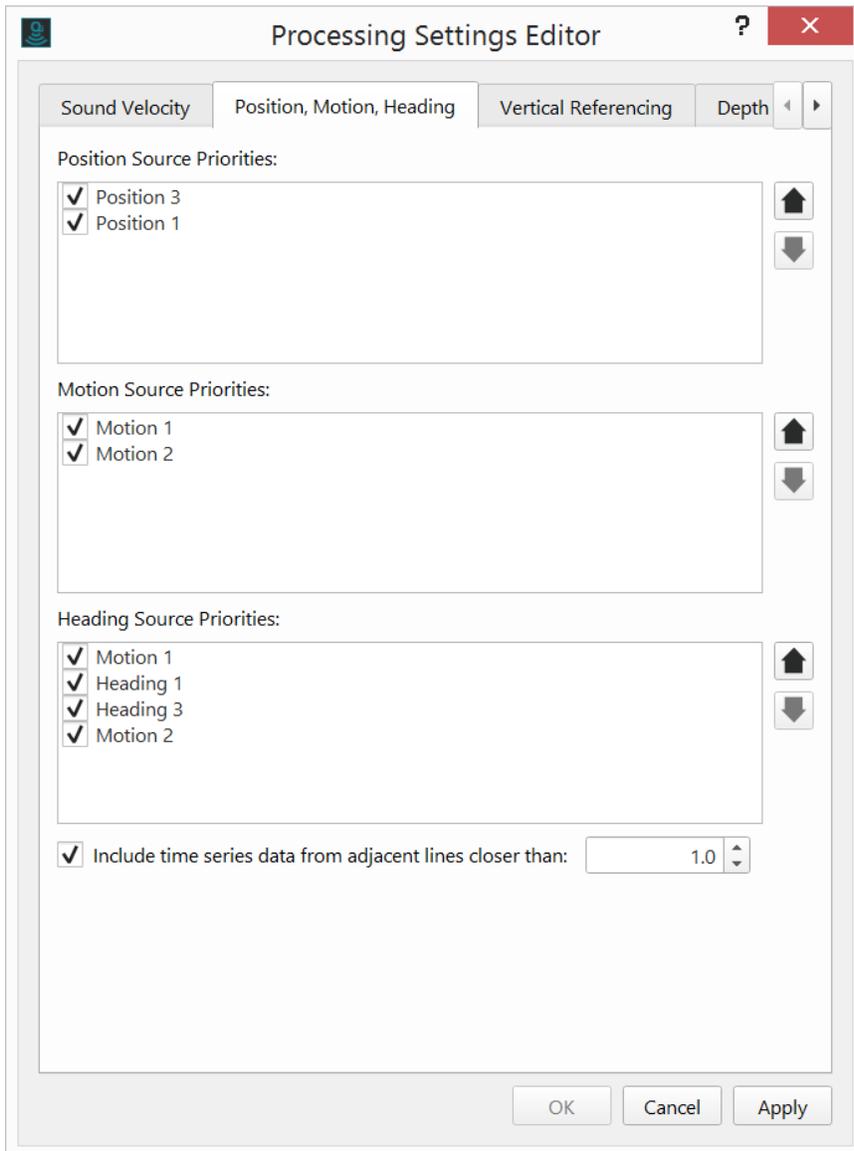


Figure 6. Listing of ancillary sensor streams for position, motion and heading for a Kongsberg .all file. Note the priority ordering in each list window, as determined from decoding the sensor priorities from the incoming data file. In this case, the tertiary position system (Position 3) was chosen as the priority system over the primary one (Position 1).

Vertical Referencing

Very few file formats provide sufficient measurements or metadata to effect a rigorous vertical referencing to the user's desired vertical datum. Some formats report tide, e.g. QINSy .db and Hypack .HSX, but the vast majority do not. For both rich and lean files, Qimera will configure the vertical referencing to be "None" when no vertical correctors are found and the user can process their data and arrive at a map. They must then import tide at a later stage as an additional corrector. For files that contain tide, the vertical referencing mode is configured appropriately and the tide is used during processing. The only two formats that provide sufficient information to use GPS/GNSS heights for vertical referencing are QINSy .db and

Hypack .HSX formats (provided that both acquisition systems are configured to generate these results). In these cases, Qimera will detect and configure the processing to use the GPS/GNSS antenna height. Most file formats record the GPS/GNSS height, but do not provide enough information to detect if this data stream should be used for vertical referencing. In these cases, the height is always extracted anyway, but the user must choose to use this for vertical referencing.

The Cost of Lean Files

It's worth examining the opposite of rich file formats. There are "lean" files that contain very little information beyond the basic echosounder measurements themselves and perhaps the primary ancillary position and motion sensor data streams. In these cases, there is little information for Qimera to extract and the user must update the vessel configuration information correctly and then also choose the correct configuration. The cost of lean file formats is potential human error.

Of course, some rich files may end up being lean files if users do not take advantage of an acquisition system's capability to record items like sound speed profiles or vessel sensor geometry, or even something as simple as the vessel name. For example, some surveyors may not be in the habit of encoding patch test correctors and sensor locations in their acquisition systems because they are used to encoding this information in their post-processing applications. A rich format only becomes rich if the information is encoded.

Qimera can handle lean files as well, but more human interaction is required. For files without a vessel name, the user must choose one during import of the first file and then must correctly choose the correct vessel on every subsequent import. Qimera reduces error by limiting the choice to the vessel names that already exist in the project.

For file formats that do not encode vessel sensor geometries, Qimera will create an empty default vessel configuration, but it will warn the user after extraction that they have an empty vessel configuration. The user must then enter the offsets manually in the Vessel Editor. This configuration can be saved and re-used during subsequent import sessions and Qimera reduces error by limiting the choice of vessel configuration to only those in the present project. Qimera is also designed to reduce human error during data entry in the Vessel Editor by avoiding the use of arbitrary axis names like X, Y and Z. Instead the terms *Forward*, *Starboard* and *Up* are used to denote the positive directions of each axis and the English names help avoid the errors that occur when matching coordinate frames from another software application or report (see the column headings in Fig. 2).

For files that do not provide an SVP, Qimera will default to use the surface sound speed probe data and will model the water mass to have that sound speed from top to bottom. This is typically suboptimal but a reasonable recourse when no SVP is present. In these cases, the user is warned about the odd configuration, but they are not stopped from processing using this simple SVP model. The Qimera philosophy is to let users get to a map with little interference

and to let them bring the additional correctors that they require. Similarly, a user can process data and create a map without tide correctors.

The Benefit of Rich Files

As explained above, Qimera extracts as much information as possible and configures the appropriate processing configuration for the user. The techniques used address the sources of human error associated with data entry, data transcription, unit conversion and coordinate frame transposition. They also circumvent the file management errors that humans make, like choosing the incorrect SVP file during import or picking last year's vessel file instead of this year's or an entirely incorrect vessel file altogether.

With a correctly configured acquisition system, post-processing is very simple and Qimera will prompt the user to process the data immediately after import. Provided that the rich file format gave everything that was needed, the user can get straight to work doing data validation. Users who invest in acquisition systems and/or sonar hardware that allows them to encode all necessary ancillary data and survey geometry information will reap the rewards for doing so in post processing. Having to record ancillary data and survey configuration information in separate documents can lead to substantial human error.

Pillar 2: Guided Workflow

The First Pillar gave an incoming user with rich data files an excellent chance at importing all their measurements and metadata while eliminating the potential for human error. The next design pillar addresses the need to get quickly to a bathymetric product, typically a gridded terrain model, quickly, easily and with little room for error. We achieve this in Qimera with a Guided Workflow, a simple approach with unobtrusively prompts the user into the next step in the process through a series of questions. This is achieved with a series of drop down questions that float over the main 3D display, prompting the user to the most likely next action, as shown in Fig. 7.

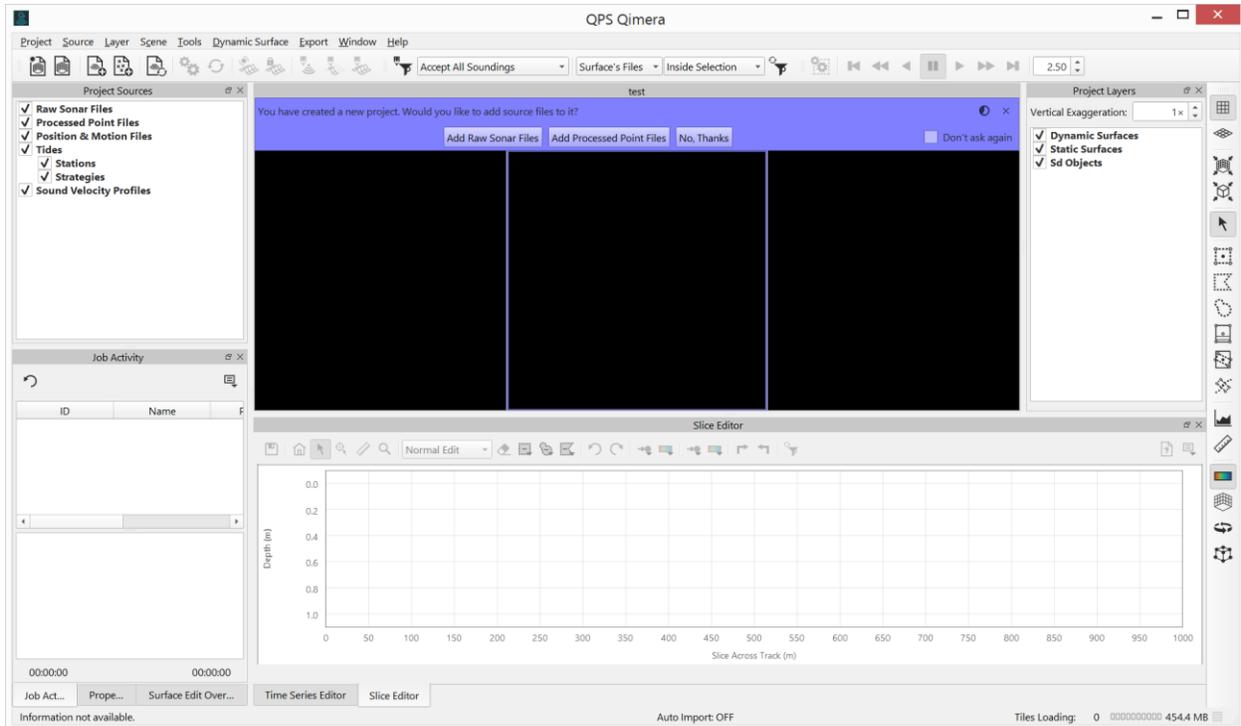


Figure 7. Guided Workflow prompt (in purple), prompting the user if the user would like to add data to their newly created project.

The startup sequence after starting a new project would be:

- You have created a new project. Would you like to add source files to it?
 - Answering YES launches the file import dialog. The user then navigates using their file browser to the location of their files. They can also specify the coordinate reference frame in which the incoming positioning data is reported. The display adapts based on the file type, for example, it will prompt for a vessel name for files that do not provide this. See Fig. 8. for an example of importing a Hypack .hsx file.
- You have added raw sonar files, but they need to be processed before they can be used. Would you like to process them now?
 - Answering YES will launch the processing engine, which is a background task. The processing engine will use all of the ancillary data and processing configuration information that was detected in the file. The outcome of this is a set of QPS QPD files, one for each incoming raw data file. The QPD file contains the georeferenced soundings as well as their computed TPU.
- Some lines became ready for cleaning. Would you like to add them to a Dynamic Surface?
 - Answering YES will launch the grid creation dialog box. By default, all of the files that were imported are considered for inclusion in the grid. The user is presented with a suggested cell size that is based on the mean depth of the data that was computed in the processing stage.

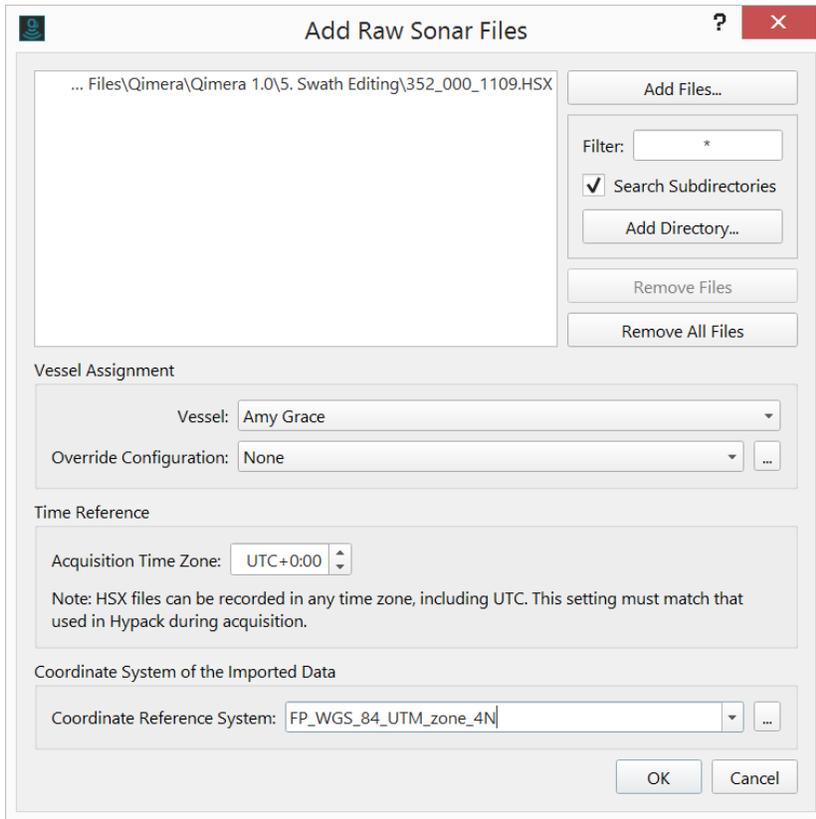


Figure 8. File import dialog for a Hypack HSX file. The dialog adapts to allow specification of the vessel name (which is sometimes not encoded in Hypack). It also allows specification of the time reference frame that the data was recorded in since Hypack allows users to record in a time frame of their choice. Lastly, the coordinate system of the imported data will default to WGS84 but will prompt the user to choose a projected coordinate frame since Hypack records position data in projected units in HSX files.

For a user that's unfamiliar with the software, this mechanism guides you through the processing stages to your desired product. Not only does it introduce a new user to Qimera, it also lets non-expert users arrive at typical bathymetric deliverables with little training or expert knowledge. This touches on the Job Scheduling aspects discussed earlier and it builds on the simple idea that the user should not have to guess or be trained on what is required to create a basic map. There are many other Guided Workflow questions that can be presented to the user in a variety of situations, these are covered in the next section on Processing State Management.

All of the actions that the Guided Workflow prompts users towards can be achieved manually through buttons on the main GUI toolbar. The Guided Workflow can be completely ignored once a user is familiar with the software in that the prompts will slide away on their own if they are ignored (note the small timer icon in the upper right corner of the question prompt in Fig. 7).

Another feature of the Guided Workflow is the ability to have it recall your decisions. Looking at the lower right corner of the question prompt in Fig. 7, you will notice a small tick box with the text "Don't Ask Again". These appear on all Guided Workflow prompts and they allow the user to craft their own personalized experience when they reach particular stages of the process. The user's decisions are stored as an application preference, see Fig. 9. If a user finds that they

are always creating a project, clicking a button to import data, then processing the data, then creating a map, then answering “Yes” to all of the Guided Workflow prompts and choosing to record their decision will result in a very streamlined experience the next time they create a new project. On their next new project, they will be immediately prompted to import new files and then the application will import the data, process it, and prompt the user for the creation of a grid. For users with rich file formats, this can result in a repeatedly streamlined experience. For users with leaner file formats, they may need to import ancillary data, thus some stages of the procedure may be less desirable for them to automate. This can be dealt with in the Guided Workflow by answering NO at the points where the user needs to diverge to a manual mode, and these answers can be stored for reuse in the next new project session.

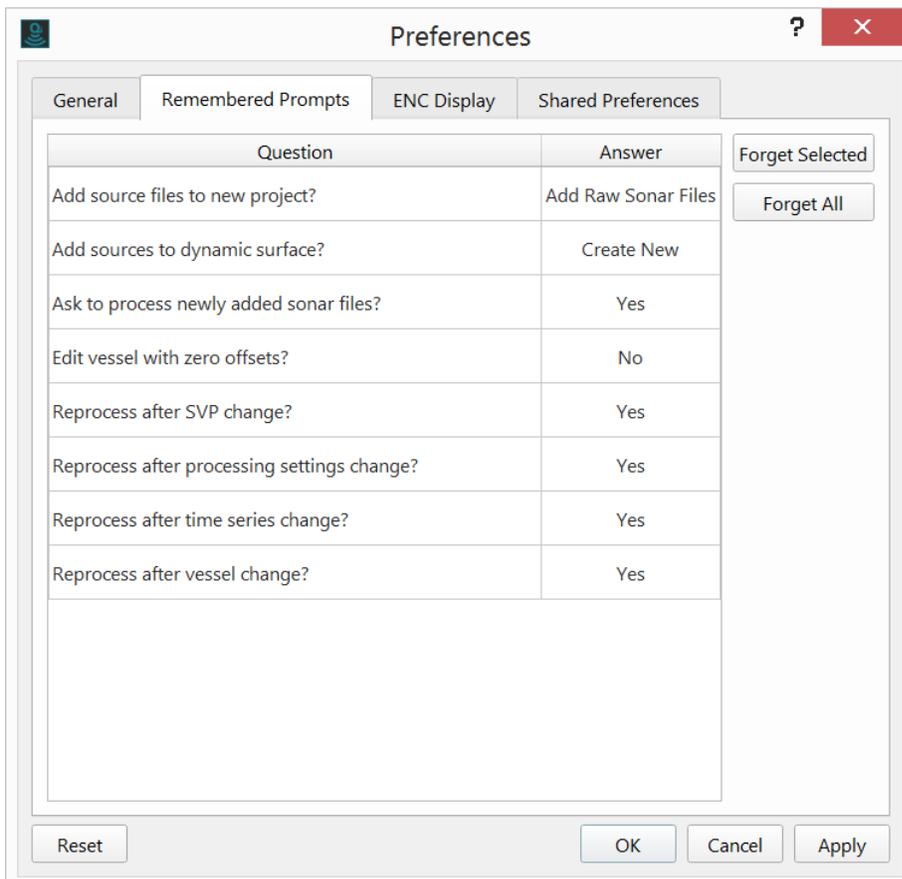


Figure 9. The user's Guided Workflow preferences can be viewed and cleared out from the application preferences widget.

One final feature of interest is that the Guided Workflow is also used to warn users about unusual situations, for example, having a vessel configuration where all the sensor offsets are all zero. In this particular situation, the user is prompted as follows after their data import stage: “Your vessel offsets are all zero. Would you like to edit your vessel settings now?” This achieves the purpose of warning the user about an unusual situation, and it optionally brings the user to the correct interface in the application to take corrective action, this being the Vessel Editor. This particular situation always arises with lean file formats that do not store vessel

information. It can also occur with rich file formats in which the surveyor did not bother to input the vessel configuration into the acquisition system.

Pillar 3: Processing State Management

The First and Second Pillars get a user to a stage where they have imported data and have generated a map. What about the case where they need to perform some editing on ancillary data streams, or perhaps make a change in configuration in how ancillary data are prioritized or used? The Third Pillar addresses this by having Qimera manage the processing state of data files for the user. What does this mean? It means that Qimera can capture and record what type of reprocessing must be done for any activity and can then take the appropriate action in response.

Some practical examples are considered below:

- Editing ancillary data: A spike is removed from a position system. Any soundings need to now have their geo-referencing updated.
- Processing configuration change: A backup motion sensor is chosen over the primary motion sensor due to failure of the primary sensor. Recalculation of beam launch vectors must be done and the soundings must be raytraced. All geo-referencing must also be updated.
- Import of ancillary data: Tide data are imported. The vertical referencing of soundings must be updated with the new corrector.
- Vessel configuration change: A patch test is evaluated for a small set of lines and is saved in the Vessel Editor for application to all survey lines. Recalculation of beam launch vectors must be done and the soundings must be raytraced. All geo-referencing must also be updated.

Qimera deals with the scenarios described above with Processing State Management, which is to say that it codifies, maintains, and manages the relationships between observations and results. The desire is for Qimera to handle all of the reprocessing complexities that making a change may entail. This is accomplished by three techniques: (1) change scoping, (2) change tracking, and (3) processing abstraction. Briefly, change scoping determines which files are affected by any user action. Change tracking maintains a log of what kind of change was done to a particular file identified in the first step. Processing abstraction lets the user update their results with a single action instead of needing to understand what type of reprocessing is necessary for any given change.

The first technique identifies which multibeam files are affected by any given user change, i.e. Qimera determines the scope of work that will result from the user change. For example, if you edit an SVP file, only those files that actually use that edited SVP file should require reprocessing. If you edit a vessel configuration time entry, only the files that fall within the affect time frame are marked as being modified. This particular innovation limits reprocessing to only the files that need it and saves processing time. More importantly, it completely eliminates the human error that is prevalent in other software applications where a user must

decide which files need to be processed after a change. Once the scope of work has been determined, Qimera updates the Project Sources dock with symbolization to indicate which files need reprocessing, see Fig. 10.

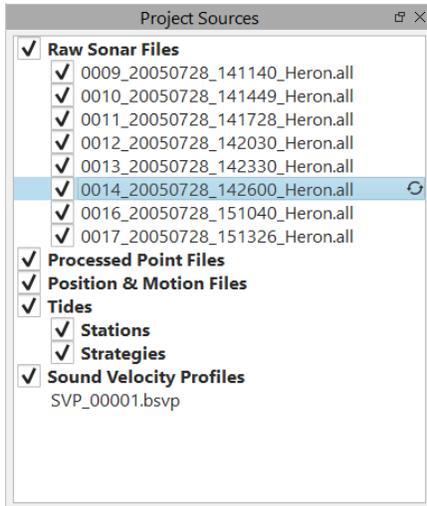


Figure 10. Raw sonar files that need to be reprocessed after a user initiated change in the project are marked up with the circular arrows update symbol found on the right side of the file listing.

The second technique, that of change tracking, has Qimera track what modifications were done on a line-by-line basis for the affected files. This is accomplished by maintaining a small metadata file that captures the nature of the change for a given file. Note that this does not indicate what type of reprocessing must be done, it instead captures what type of modification was done that may affect the results of a particular file. The user can peek behind the scenes to determine what type of modifications affected a file via a right-click context menu option, the dialog that launches is shown in Fig. 11.

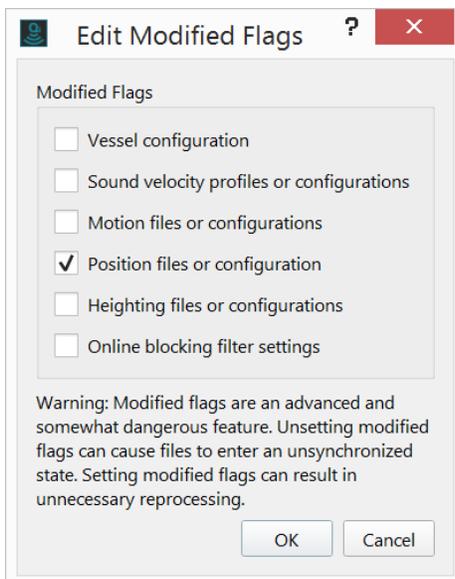


Figure 11. Dialog used to show what recent modifications were done that would affect the required reprocessing for a particular line. This dialog is provided as a view behind the scenes and as an escape hatch for when users inadvertently make a change that might trigger a large reprocessing job.

The third and final component is the processing abstraction of the mathematical steps required to get from the raw sensor measurements to a georeferenced sounding or from any partial computation stage in between. Looking back at the processing examples at the beginning of this section, it is clear that very different reprocessing actions are required for any given change. Though there are several reprocessing actions available, in Qimera we have decided to hide this complexity from the user completely.

We did not do this to provide a black box processing experience, instead we have simply encoded the dependencies between measurements and results and we have tabularized the appropriate action to take for any given change. Qimera is doing the same kind of reprocessing that users are familiar with having to complete in other software packages: merging, computing TPU, ray tracing, etc. All that is done differently is removing the human from the loop with the goal of eliminating human error.

Users no longer need to worry about what kind of reprocessing needs to get done, they only need to be aware that some reprocessing needs to be done in order to update their results. The user does not need to know which processing routines or subsystems are triggered in the same way that a car driver does not need to understand what happens when they turn the key in the ignition of their car or when they shift the transmission from Park to Drive. This completely eliminates the human error in scheduling the appropriate kind of reprocessing for any given change to the inputs. If you want a particular change to take effect, you should not have to worry about what files need to be reprocessed, nor what kind of processing needs to be done.

This experience is augmented with the Guided Workflow that prompts the user to reprocess data after any user initiated change. Some examples are listed below.

- After time-series edit: “You have modified time series data. Would you like to reprocess the affected file now?”
- After SVP edit: “You have modified SVP data. Would you like to reprocess the affected files now?”
- After Processing Setting Edit: “You have modified processing settings. Would you like to reprocess the affected files now?”
- After vessel editor edit: “You have modified vessel configuration data. Would you like to reprocess the affected files now?”

As with the Guided Workflow prompts that appear after creating a new project, the user’s decision can be encoded as a preference. By answering YES to the above questions and choosing to preserve this decision, a Dynamic Workflow can be experienced which is discussed in the next section.

One particularly power feature is to have the Processing Configuration file update automatically with import of new ancillary information. For example, if you had data previously without tidal corrections, and you imported tide, then the Processing Configuration updates itself automatically for the files for which the tide overlaps in time. Any subsequent processing of multibeam files, whether it’s immediately done after import of the tide or delayed, will use the tide correction. This strategy bundles the act of import, processing configuration, and processing

into a single task. Once a tide file is imported in the project, it will be matched up to and applied to any multibeam files that are imported at a later stage, as long as their times overlap. The same principle applies for externally sourced height or navigation data like an Applanix SBET file. After import, the Processing Configuration switches to GNSS heighting mode automatically. If you bring in any more bathymetric data from the same vessel that happens to overlap in time with the SBET data, it is automatically configured to be used. Essentially, you import the ancillary time-series data and it is used automatically from that point forward for existing files in the project and for all future files that come into the project.

How does this combine into a user experience? Effectively, from a user's point of view, they have raw sensor data, they have the recipe card (in the form of the Processing Configuration dialog) to get from raw data to results, and they have results. As they bring new information into the project, the "recipe card" updates itself automatically. Though there are complex operations happening when data is updated or reprocessed, it is not necessary for the user to understand what is happening behind the scenes. They simply need to know if their results are up to date or not.

Effectively, the user cannot decouple changes in processing inputs and their configuration from the outputs. For example, if a user edits the tide, the tide correction for pings will be recomputed for them. If they edit an SVP, the soundings will be re-raytraced. If they import a post-processed navigation file, the navigation will be updated. The appropriate action is taken for the user. This design couples Action and Effect into an atomic and indivisible task. The user cannot make changes without updating output. They can delay the update, but Qimera's change tracking will ensure that they don't forget. Users do not need to remember what processing needs to be done, just that some processing must be done.

Pillar 4: Dynamic Workflow

The Fourth Pillar is a culmination of the first three pillars working together to build a "live" processing and validation environment where gridded data products can be quickly and easily updated after processing and/or validation. This is done to provide feedback to the user as quickly as possible to give an interactive feel to the data processing experience.

The key enabling technology in this particular case is the Dynamic Surface, a QPS grid storage format that, coupled with the sounding results files (QPD files), allows for quick updates to the surface when any given source file is updated. The updating is automatic after any processing or editing activity, though it can be delayed if desired. When you update processing for a line, the Dynamic Surface only updates itself for the area that encompasses the line. If you edit a small amount of sounding data in a particular area, only that area is updated in the Dynamic Surface.

The Dynamic Workflow comes to life when the Guided Workflow and Processing State Management is enabled in full automatic mode. As discussed in the previous sections, the Guided Workflow prompts that a user is presented with can be configured to apply automatically when the same actions are repeated in the future. If a user chooses to have the application react dynamically to their changes, i.e. automatically trigger the appropriate re-processing following a

user action, then the application will respond to changes and all soundings will be updated automatically at the completion of the user action. Combined with the Dynamic Surfaces, the gridded data products that are derived from the soundings will update as well. From the user perspective, they commit an action and with a bit of time, they see the result.

What can a Dynamic Workflow do to improve post-processing outcomes?

- It allows for “live” processing state management. You can make a configuration change for a single file, see if it improves the resulting surface. If it does, you can make the change for all files. If not, you can undo the change and the surface re-grids the affected area to return to its original state. Trial and error for problematic processing can be quick and simple with much less room for human error.
- It allows for “live” data validation of soundings to remove outlier detections. For data validation to be most effective, you need immediate feedback on what you’re doing. Dynamic Workflow connects validation to grids and lets you validate with confidence. When you edit soundings, the grid can update nearly immediately.

In essence, the Dynamic Workflow makes it easy to make processing configuration adjustments or to perform data validation and to immediately assess the impacts of changes. The shortened feedback cycle between cause and effect promotes causal reasoning, a key ingredient for natural human learning processes. In effect, it allows users to train themselves.

Improving Workflows and Outcomes: In Practice

We set out to design a modern bathymetric data processing application with the primary design goal of reducing human error while also making it user-friendly. Have we met the goal that we set for ourselves? We can confidently answer YES and we have evidence of this in two forms. Firstly, we have the results of a client web-based survey conducted in late 2016 where we asked clients about all of our software products, including their impressions of Qimera. Secondly, we have a small case study of three clients who share a seabed mapping philosophy and staffing strategy that Qimera uniquely enables them to achieve. Both of these are explored further in the sections below.

QPS 2016 client survey

In late 2016, we sent out an email to clients asking them to give us feedback on all aspects of QPS, including software, customer services, etc. At the time of the survey, Qimera had been on the market for just over one year. We had plenty of anecdotal evidence from client site visits and trade shows that Qimera was easy to use, but this was our first concentrated feedback across a wide range of clients and user bases. In the survey, respondents were asked to complete the sentence “The best thing about Qimera is ...”

What did they say?

- Ease-of-use: 45%
- Processing and Analysis Capabilities: 17%
- Performance and Scalability: 15%
- Modern and Intuitive User Interface: 14%
- Support for Data Formats: 6%
- Misc.: 3%

Qimera's ease-of-use was the most frequent response (45%), with nearly as many comments as the next three categories combined (46%). The responses for this particular question were a free form text entry in the web seminar, so some categorization was required for the summary numbers presented above.

Client Case Study

We decided to build a workflow that makes processing easy. We then asked people if it was easy, and they overwhelmingly said yes, as evidenced in the client survey results discussed above. But does "easy" actually bring tangible value to the survey world? What does "easy" enable you to do?

To answer this, we have the kind permission of three QPS clients to include their shared experience in this discussion. The three clients have among themselves three deep-water survey ships. They all use Qimera and are listed in order of adoption:

- R/V *Falkor*, Schmidt Ocean Institute
- E/V *Nautilus*, Ocean Exploration Trust
- NOAA Ship *Okeanos Explorer*, NOAA Office of Ocean Exploration and Research

What do they do?

These three institutions have many things in common. They all more-or-less share a similar mandate in that they all promote scientific exploration at sea. They also all recognize seafloor mapping as being a core capability of their platforms (all platforms are outfitted with multibeam echosounders). They are all multidisciplinary science platforms and can accommodate a variety of scientific missions. Mapping is not always the main mission, however, the mapping capabilities of the platforms are often used in supporting roles when scientific objectives are focused on seabed observation and/or sampling, for example ROV dives.

What hardware tools do they use?

All three ships use deep water Kongsberg Maritime multibeam echosounders (EM302), however *Falkor* is additionally equipped with an EM710. Kongsberg echosounders use a vendor-specific (but openly described, .all) file format and this format is a rich format in that it allows the capture of the vessel's geometric configuration, the sound velocity profile, as well as a real-time integrated solution as performed by the Kongsberg software itself. The format also allows

recording of multiple position and motion sensor data streams, with the operator's priority preference being encoded in the file as well.

How do they work?

The three platforms share staffing models as well in that they are only lightly staffed with permanent personnel that are highly trained in sonar operations and data processing. These staff are occasionally or regularly augmented with a small number of interns that join cruises on a rotational basis. On some platforms, visiting science parties are expected to provide the bulk of the personnel required to staff a particular mission. In these cases, the scientific personnel and interns joining the ships who are participating in mapping missions do not have consistent experience, training or domain knowledge for seabed mapping operations. They need to be quickly trained to process data and achieve quality results. Sometimes those mapping results feed into seabed sampling operations conducted that same day.

What is their common challenge?

What do these platforms have in common? They all recognized that they have a uniquely challenging situation at sea: They need to map and deliver timely products that feed into planning and operations almost immediately after the mapping. The vessels also need to do this with a small staff who have varying levels of skill and experience in the seabed mapping field. Each organization needs to do this continuously mission after mission, with a largely fresh set of personnel on each mission besides the small number of permanent personnel that are assigned to the vessels.

What has been their common solution to their shared challenge?

All three organizations recognized the value of Qimera as being the key enabling technology in this scenario. What makes their situations particularly efficient with Qimera is that they have the benefit of a rich file format from their echosounder's acquisition system. All three platforms conduct a yearly "shakedown" type assessment, in some cases expert personnel join the ship to verify that all mapping systems are functioning as expected and that they are still correctly configured. With the yearly verification complete, you no longer need a highly trained surveyor to conduct post-processing since the surveyor's knowledge of the configuration is largely encoded in the raw data files. This allows Qimera to automatically set up the post-processing experience without further human interaction. This results in turn-key, rapid mapping workflow since the files can be imported and processed to a map immediately with repeatable and consistent results with a few mouse clicks. With the staffing scenarios outlined above, the training that a new processing watch stander needs is largely limited to the human intensive portions of processing, that being data validation using the various sounding editors in Qimera.

In conclusion, to answer our initial question, what does "easy" allow you to do: In this particular example, with a correctly configured acquisition system that provides a rich file format, the ease-of-use provided by Qimera allows the user to proceed almost immediately to the task of data validation. The ease-of-use of Qimera allows non-expert personnel to be trained in a relatively short period of time, such as on the transit out to their site, and it allows them to produce high-

quality mapping products in near real-time that support seabed observation and sampling campaigns.

What Can You Do?

We have built a modern, easy-to-use, and intuitive hydrographic data processing software application that users confirm that (a) it is easy and (b) the design features we have focused on are indeed bringing value to post-processing efforts for early adopters. As we've pointed out, Qimera is powered by rich file formats. To get the most out of using Qimera, there are a few points that you should consider to get the most out of it.

Are you a project manager? To get the most out of our work to eliminate opportunities for human error, you should insist on hardware/software combinations that provide rich file formats when you set about choosing the right tools for the job. If you have an existing combination that you trust, you should examine it closely to see what elements require human data entry or configuration in post-processing to see if you can make operational changes to improve outcomes.

Are you a system integrator, patch test specialist or a shake down guru? You should insist that sensor locations and alignment angles are entered in the acquisition system. Best-practice dictates that you should verify that the entry is correct by implementing quality assurance procedures to ensure that the system is correctly integrated. The attitude of "we'll just fix it in post-processing" exposes post-processing data efforts to the risk of human error. Get your system correctly integrated before surveying, evaluate your integration to verify it and "bake it in" to your files. Once that's done, document what you've done and inform the survey crew why these changes are important so that they understand why the acquisition system is configured the way that it is.

Are you a surveyor? You should also insist on rich file formats. You should have Qimera running alongside your acquisition system, it can be configured to automatically import data at the end of a survey line, process it, and add it to a grid. This gives you a near real-time second opinion on the quality of your real-time integration in the acquisition system.

Are you a hardware vendor? You should look at what users can store in your native file format. Do they allow for a complete record of the survey data and configuration such that it could be automatically processed? Don't think of just your users, think of your own engineering, customer support and sales staff. Qimera has become popular for hardware vendors doing installations, demos and sales. Help them out by consider the following:

- Vessel Name and sonar serial number. Data processing software should be able to distinguish between files from one platform or another. Having this capability allows users to add large collections of files and have them get sorted out by vessel with ease.
- Multi-head configurations: Be smart about this. Software should be able to tell what sonar a particular ping comes from. Furthermore, a sonar head should be unambiguously matched to a sensor location in the vessel configuration.

- Sensor locations and alignment angles. This is absolutely necessary. If this is not complete, then users must enter this information manually into a vessel configuration file, and from that point onward, they will always need to remember to choose the correct configuration for every one of their files that is ever imported into a project.
- SVP record, with full metadata. Allowing surveyors to encode an SVP into the data files is another important step towards allowing automatic post-processing. Full metadata such as the latitude and longitude of the cast and the time of acquisition are important as well. Additionally, the time that the cast is applied in the data stream is very useful as well.
- Channel prioritization. If multiple data streams are supported for any given data stream (for example, primary GPS and backup GPS), you should allow the user to choose which is the primary during acquisition and you should encode that decision in the file.
- Draft. For users who are using tides for vertical referencing, you should allow for entry of the draft. This completely eliminates an error prone step in post-processing.
- Clearly and unambiguously document your sign conventions and coordinate reference frame in an easy-to-find and single spot in your documentation. Don't forget sign conventions for sensors as well. This helps your own staff, it helps your users and it helps software vendors to clearly identify what to do.

Conclusion

With Qimera, we set out to make hydrographic data processing easy and less prone to human error. Early in this paper, we stated that a paradigm shift is a fundamental change in approach or underlying assumptions. How is making something easier a paradigm shift?

Change in approach: As an industry, we've been solving the wrong problem for too many years. The problem to solve has been human error. A lot of thought put into workflow analysis, human psychology, and the nature of human error pays off in building intuitive user interfaces and workflows that provide quality outcomes with a high probability of success and repeatability from one user to another.

Change in underlying assumptions: Many users are not surveyors. We've changed the underlying assumption that you need to be an expert surveyor with years of experience to properly process multibeam bathymetry data.

References

IHO (2008). "IHO Standards for Hydrographic Surveys, Publication No. 44, 5th Ed." International Hydrographic Organization, Monaco.