UHDAS and CODAS

Dr. Julia M Hummon

University of Hawaii

hummon@hawaii.edu

Revision History

May 12, 2011	draft

Table of Contents

1 Introduction

UHDAS refers to a suite of programs and processes developed at the University of Hawaii that perform data acquisition, data processing, and monitoring, at sea. In addition, access to documentation and code are provided on the ship's network. We have tried to make a system that is useful and reliable, easy to operate, and which provides as close to a final dataset as is reasonably automatable while maintaining the fundamentals necessary to reprocess the data from scratch if necessary.

Documentation for UHDAS and CODAS (the processing component) are housed

- at sea (http://currents)
- on land (http://currents.soest.hawaii.edu/docs/adcp_doc/index.html)

2 UHDAS Overview

This section is available in the CODAS documentation, referenced above.

UHDAS has four components at sea:

- Acquisition
- Processing
- Monitoring
- Access (to data and figures)

2.1 Data Acquisition component

Data acquisition programs are written in C, and the gui and supporting code are written in C and Python.

Data acquisition includes

- a dialog with each of the RDI ADCPs to set parameters and start pinging
- acquisition and timestamping of passive serial inputs
- data collected are
  - binary records (from ADCP ensemble)
  - NMEA strings (from serial inputs)
- NMEA data recorded usually comprise
  - GGA messages (gps) from two sources if possible
  - gyro heading
  - accurate heading (POSMV, Ashtech, Seapath, Mahrs, Phins,... if available)
- files roll over every two hours
- timestamps are zero-based decimal day (Jan 1, 12:00 UTC is 0.5, not 1.5)
- a parsed version of each NMEA string is added to a set of intermediate files to stage information for the processing component ("rbin" files)

2.2 Processing component

Processing code is written in C and Python. Final processed output are written as Matlab and netCDF files on a regular basis. Processing is done using a CODAS database (Common Ocean Data Access System) as storage and retrieval system. The suite of programs designed to extract from, manipulate, and write to the database is known as "CODAS ADCP Processing" and has been free, maintained, and in use since the late 1980's.

In a batch mode, CODAS processing can be applied to single-ping data gathered by UHDAS (or the commercial RDI software "VmDAS"), or averaged data collected by VmDAS or the original DAS2.48 (used with Narrowband ADCPs in the late 1980's and through the 1990's).

At sea, a UHDAS installation acquires data and uses CODAS processing to calculate ocean velocities from ADCP measured velocities, position, and heading (gyro, corrected to accurate heading if one is available). The following three levels of processing combined are called CODAS Processing:

(1) CODAS steps performed on single-ping data

- make sure every ADCP ping has a position and a heading
- gather the next T seconds of data (eg. 300 seconds)
- screen the ADCP data to eliminate bad values (eg. acoustic interference)
- average in earth coordinates
- write to the disk

(2) CODAS steps performed on averaged data

- load measured velocities into the database
- add navigation to the database

The following are steps automated on a ship with UHDAS, but can be done afterwards with human intervention

(3) CODAS Post-processing (on averaged data)

- correct the gyro heading to the accurate heading device (if there is one)
- apply scale factor if specified (eg. NB150, BB150, WH300)
- apply additional fixed rotation if specified
- edit out bad bins or profiles (eg. data below the bottom)

It should be noted that Phased Array ADCPs ("Ocean Surveyor" series) are capable of pinging in either narrowband or broadband mode. They are also capable of pinging in interleaved mode, where narrowband and broadband pings alternate. This mode of operation is possible with both UHDAS or VmDAS, but with VmDAS it is much more difficult to access the second ping type. If interleaved mode is chosen with UHDAS acquisition, CODAS processing treats each pingtype as a virtual instrument, processing the data from each ping type in a separate directory.

2.2.1 UHDAS Enhancements to CODAS Processing

UHDAS adds steps to the basic processing at sea by extracting (on a regular basis) processed, corrected, edited data for scientists to use during the cruise. These data and figures that are generated from them, are available on the ship's web.

- every 5 minutes
    - get the last 5 minutes of new data

- rotate to earth coordinates
  - in general, this is done using a gyro as the primary heading device and correcting to the "accurate heading device" (if one exists)
  - on Healy, the POSMV is (at present) the sole attitude device
- edit single-ping data (for this 5-minute chunk)
- average, write to disk (staging for addition to the codas database)
- save the 5-minute chunk of data, specifically for plotting

- every 30 minutes
  - the CODAS database is updated with the staged averages
  - scale factor and fixed rotation are applied if specified
  - the averages in the database are also edited (to look for bad bins or bad averaged profiles, and the bottom)
- after the codas database is updated
  - the data are extracted and averaged (for plotting)
  - the data are extracted with "every bin, every profile"
  - data are stored as Matlab and netCDF files, accessible via ship's web site or via windows shares [samba] or nfs.
  - vector and contour plots of the last 3 days of data are updated, available via ship's web

2.3 Monitoring component

Monitoring programs are written in Python and make use of Linux system calls.

Monitoring includes

- daily email sent to land with information about
  - processes running, disk space, error messages
  - data processing status
  - daily email includes a message about heading correction
  - the last 3 days of heavily averaged (vector plot)
- **on land**, when the email is received, crude vector and contour plots are made of the data snippet that was sent
- **at sea**, the following are available on the ship's web:
  - most recent 5-minute profiles of all instruments that are pinging
  - the last 3 days of data shown as contour and vector plots
  - the last half-day of gyro and "accurate heading"

2.4 Access to data and figures

Two fundamental access mechanisms exist

(1) ship's web (usually http://currents)

- figures

        - most recent ocean velocity profile for each instrument
        - 3-day tail with surface velocity vectors
        - 3-day tail contour plot (vs time, longitude, or latitude)
- data
    - all data so far, (thick layers, eg 50m) over 1-hour, for vector plots
    - all data so far, (thinner layers, eg 10m) over 15 minutes, for contour plots

(2) network shares

- in the adcp home directory, `www/figures/png_archive` contains copies of the most recent 3-day figures made every day so far in the cruise (also available via `http://currents`)
- in the **data** directory, `current_cruise/proc` contains processing directories for all instruments logging data.
    - the same data files that are available on the web (15-minute and 1-hour) are available in the vector and contour subdirectories.
    - In the "contour" subdirectory of each processing directory are the following types of files, each extracted at the native resolution of the CODAS database, i.e. binsize is bin size of the instrument, and the averaging interval is the same as in the CODAS database:
        - a collection of Matlab files grouped as `allbins_*.mat`
        - one netCDF file (suffix is "`nc`")
- in the data directory, `current_cruise/raw` contains all the data logged by UHDAS, including the single-ping ADCP data.

3 Codas Processing Overview

This section is available in the CODAS documentation, referenced above.

CODAS (Common Ocean Data Access System) is more than a database. The word has come to be associated with a suite of open-source programs for processing ADCP data. CODAS consists of C, Matlab, and Python scripts that will run on Windows, Linux, SunOS, or Mac OSX, and can process pingdata from a Narrowband ADCP, data collected from a Broadband or Ocean Surveyor data by VmDAS, or data collected by any of those instruments using UHDAS (open source acquisition software that runs RDI ADCPs). Older processing scripts were written in Perl, but we have shifted to Python for the newer versions of these scripts.

Some kind of data treatment is necessary because the acquisition programs write binary files to the disk that are not readable by commercial plotting packages. In fact, there are not actually any ocean velocities stored in the files. A shipboard ADCP reports currents measured along each of its beams. These currents must be transformed into earth coordinates, and the motion of the ship taken out. Ancillary data such as heading and position are used to determine the ocean velocity from the measured velocities.

There are at least four necessary processing steps which are performed by (or made possible) by the CODAS routines. First, an ocean reference layer is used to remove the ship's speed from the measured velocities. By assuming the ocean reference layer is relatively smooth, positions can be nudged to smooth the ship's velocity, which directly results in the smooth reference layer velocity. (This was more important when fixes were rare or jumpy (such as with LORAN) or dithered (such as SA GPS signals prior to 2001).

Second, calibration routines are available to estimate the heading misalignment from either “bottom track” or “water track” data. Watertrack calibration routines use sudden accelerations (such as stopping and starting of the ship when doing station-work) to derive a heading misalignment. For a ship traveling at 10 kts, a 1-degree heading error results in a 10 cm/s cross-track velocity error. It is critical that the misalignment be accounted for if one is to avoid cross-track biases in the velocities.

Third, a GPS-derived heading source (such as Ashtech, POSMV, or Seapath) may provide a more accurate (though often less reliable) heading source than a gyro. Routines are in place for pingdata and UHDAS data to correct the gyro heading with the GPS-derived heading, using a quality-controlled difference in headings. An example is available for VmDAS data. Gyro headings may be reliable but they can vary with oscillations of several degrees over several hours, thus creating spurious fluctuations in the ocean velocity that resemble “eddies” , but which are solely the result of cross-track velocity errors (from the associated gyro heading errors).

Fourth, it is crucial that bad data be edited out prior to use. Traditionally, the data available from the DAS2.48 narrowband data was averaged in 5 minute groups. VmDAS and UHDAS also output time-averaged data,. which can be loaded into the CODAS database for further processing. With CODAS processing, a graphical interface allows identification of the bottom and selection of bad profiles or bad bins based on a variety of criteria. To some extent this can be automated, but for final processing, a person must visually inspect all the averages from a cruise. The graphical interface vastly speeds up editing to the point where it can take only a few minutes of user time per day of data for a typical cruise. On some ships, Healy included, this ideal may only rarely be reached.

CODAS processing has moved beyond averaged data to the realm of single-ping data. Whatever acquisition program was used to record the data also averages it. Prior to averaging, some attempt is made to eliminate bad pings. CODAS processing includes routines that average single-ping data collected by VmDAS or UHDAS. These routines allow the single-ping data to be screened more extensively prior to averaging. Under certain conditions, this may be necessary to avoid subtle underway biases caused by bubbles or ice near the transducer. CODAS processing includes the ability to read single-ping data files and look at the characteristics of the instrument (such as acoustic backscatter or beam velocities) one ping at a time.

3.1 Processed Data Conventions:

1. CODAS database times are all UTC.

2. The **decimal day** is a zero-based value. Noon January 1 UTC is 0.5, not 1.5. This is the best convention when dealing with data that may cross over a year boundary. (if your cruise does cross over a year boundary, the decimal days will just keep increasing past 365). This convention is used in naming the raw data files as well.

3. Data averaged over a period of time have an associated time and position. These represent the values at the end of the ensemble. Other scalars associated with a given ensemble (temperature, heading) are presented with two choices: 'averaged' (over the interval) and 'last' (to match the timestamps)

4. Pitch and roll (if available) are recorded if available, but are not used in the data processing.
5. Processing parameters (paths, transducer depth, transducer alignment angle, and serial messages used for position and heading) are contained in two MATLAB files in each processing directory.

   - `config/HLY10TC_cfg.m`
   - `config/HLY10TC_proc.m`

4 ADCP processed data

Errors in ADCP data (ocean velocities) can come from a variety of sources. Two of the most common are:

1. errors in the original ADCP velocity estimates;
2. errors in ship's position and attitude measurements used to transform the measured velocity of the water relative to the ship, in instrument coordinates, into the velocity of the water over the ground, in geographic coordinates.

Relative velocity errors can come from acoustic noise, electrical noise, bubbles, interference from other sonars, and failure of the ADCP itself. Future data sets may be optimized by minimizing sources of noise, etc., but once the data have been collected, the only option is to detect and edit out whatever errors are present. More often than not, such errors appear as a bias towards zero velocity relative to the ship; an example will be described in 3.1.

The most critical part of the position and attitude measurement is the heading; specifically, the heading of the transducer, which requires an accurate measurement of the ship's heading together with an accurate estimate of the (constant) alignment angle between the transducer and the ship's heading reference, nominally the keel. On the Healy the alignment angle is well-known, and the headings are obtained from the POSMV, which is stable, reliable, and closely monitored.

4.1 Introduction to “bias towards zero”

An ADCP measures the Doppler frequency shift in the beams, translating that into a velocity component along each beam. Using the orientation of the transducer relative to the ship's keel and the beam geometry, velocities measured along the beams can be translated into flow of the water past the ship (measured velocity in ship coordinates).

If the ocean had no flow, the sum of the measured velocity and the ship's velocity should be zero (Figure 1). Bias towards zero usually occurs in all 4 beams; the estimated velocity component along each beam is smaller than it should be, so the estimated horizontal velocity components of the water relative to the ship are also smaller than they should be. When the ship's velocity is added, the estimated velocity of the water over the ground is therefore biased in the direction of the ship's motion (Figure 1). This bias in the direction of motion is the “red flag” in processed ADCP data. The bias is proportional to the ship's speed; when the bias is small, it may not become apparent unless there are large changes in the ship's speed or heading.

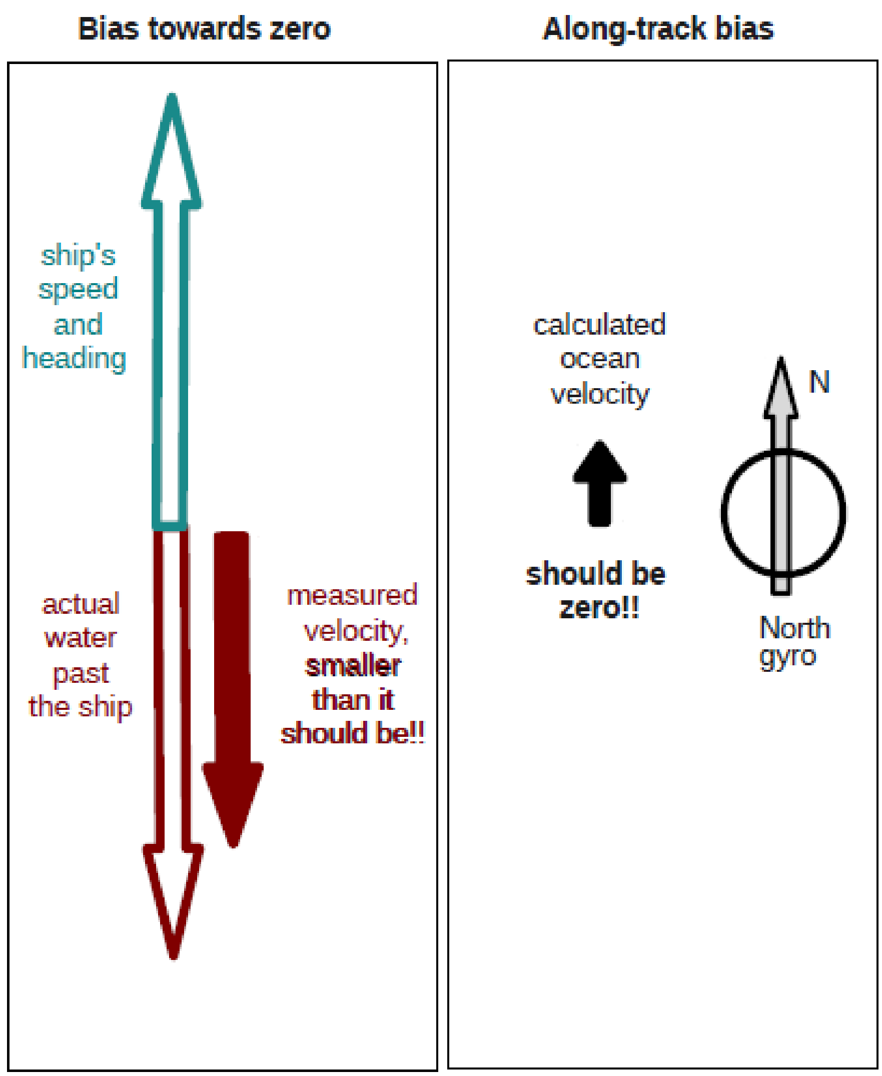


Figure 1: Cartoon illustrating "bias towards zero"

A common cause of bias towards zero is a blanking interval that is inadequate for the installation and the scattering conditions. In that case, the received sound is dominated or biased by the reverberation of the outgoing ping; because the reverberation is at the transmitted frequency, the measured Doppler shift biased towards zero. This is known as "ringing", and typically affects only the top of the velocity profile. Figure 2 shows the biased beam velocity components, and the correspondingly biased ocean velocity estimates, from a ship with a particularly bad case of ringing caused by poor window and well design. Although the Healy has not shown problems with ringing, it is affected by other sorts of bias, with magnitude increasing with depth rather than being concentrated at the top of the profile.

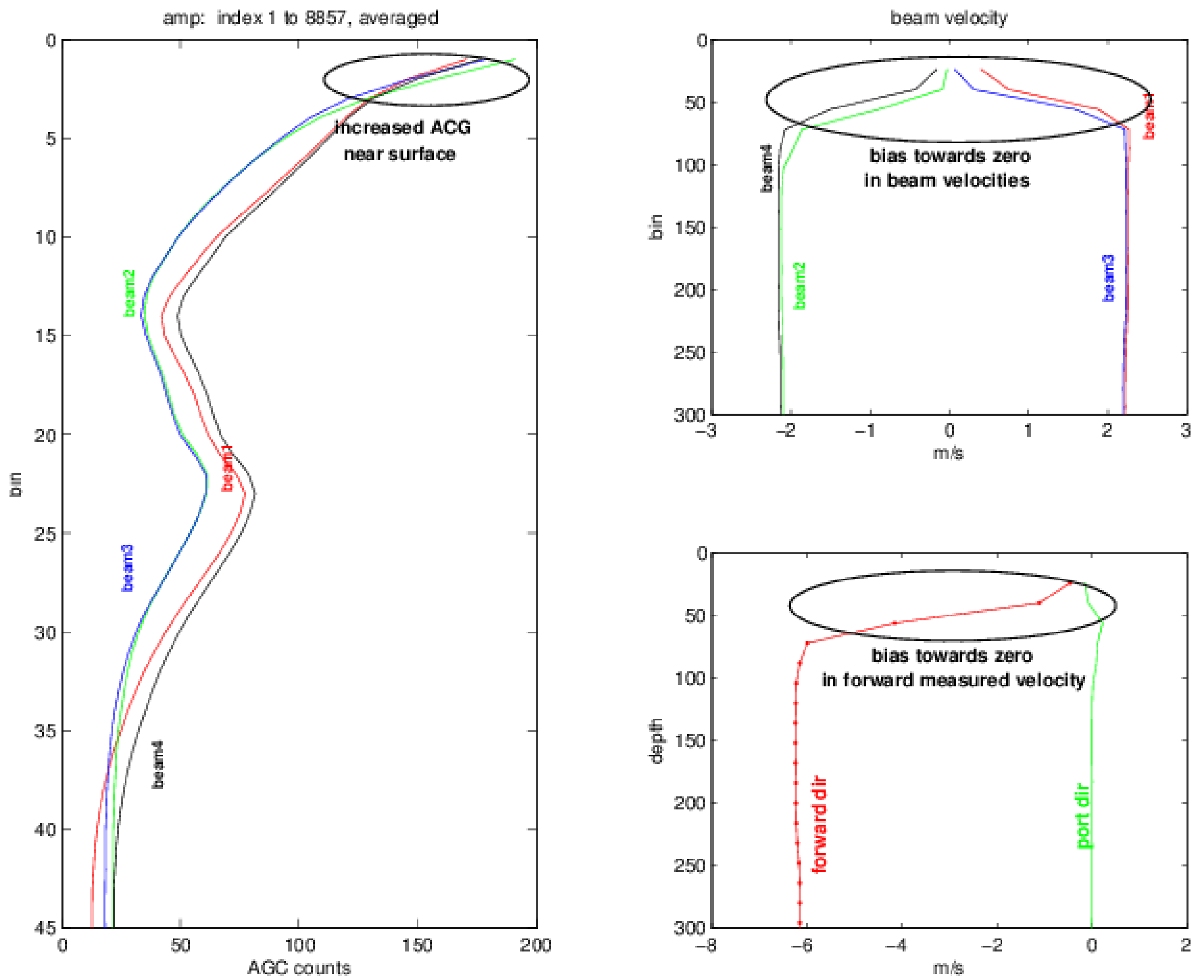


Figure 2: Example of ringing, showing bias towards zero in beam velocities and in measured velocities (fwd direction)

4.2 ADCP data processed using averages (LTA) and single-ping

ADCP data from HLY10TB (Seattle to Honolulu) were collected with both the OS150 and OS75 set to "ping as fast as possible" and to acquire data in narrowband mode only, with 8-m and 16-m bins, respectively. Data were then processed in two ways: first, using the VmDAS pre-averaged files (*.LTA, or "long term averages"), with a 5-minute averaging period; and second, using CODAS to generate the 5-minute averages from the single-ping *.ENX files recorded by VmDAS. The difference is in the single-ping editing provided by CODAS processing.

4.2.1 Bias towards Zero and CODAS processing

The effect of acoustic interference is obvious in the OS150 ocean velocity data processed using the LTA (unedited, averaged) data. With CODAS single-ping editing, the interference is effectively removed (Figure 3).

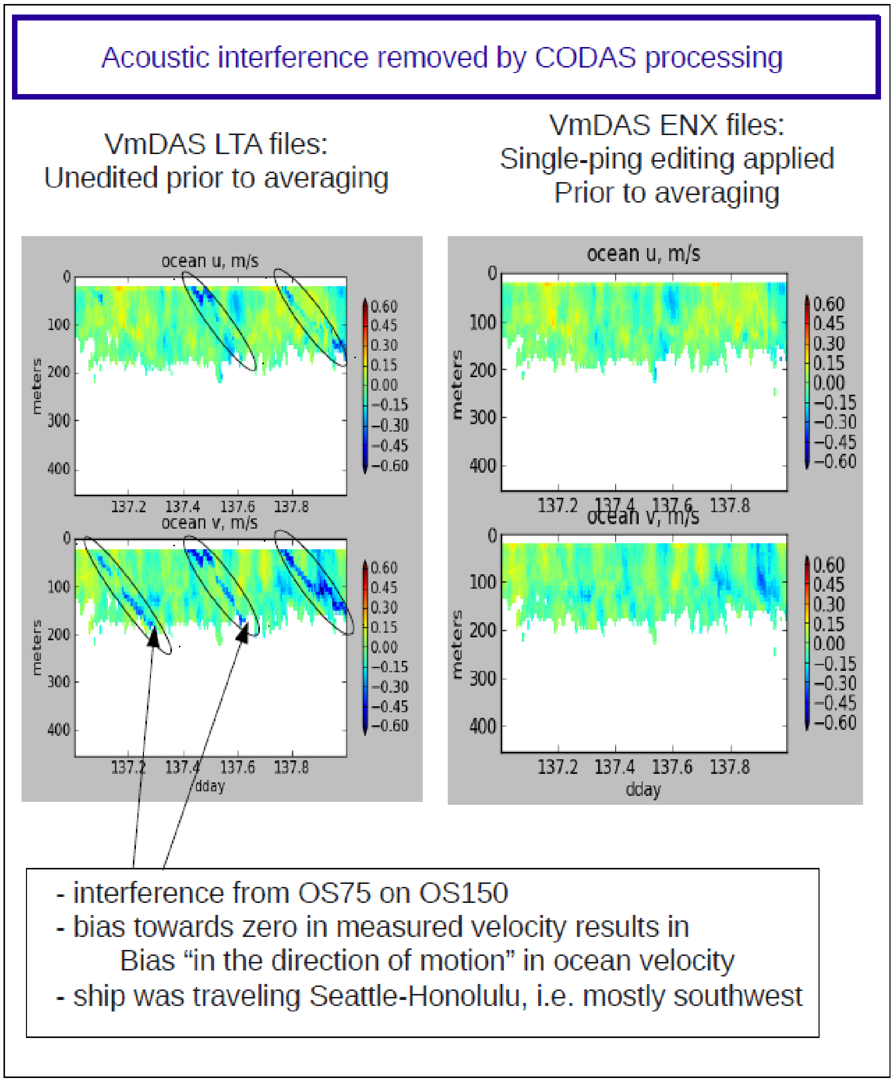
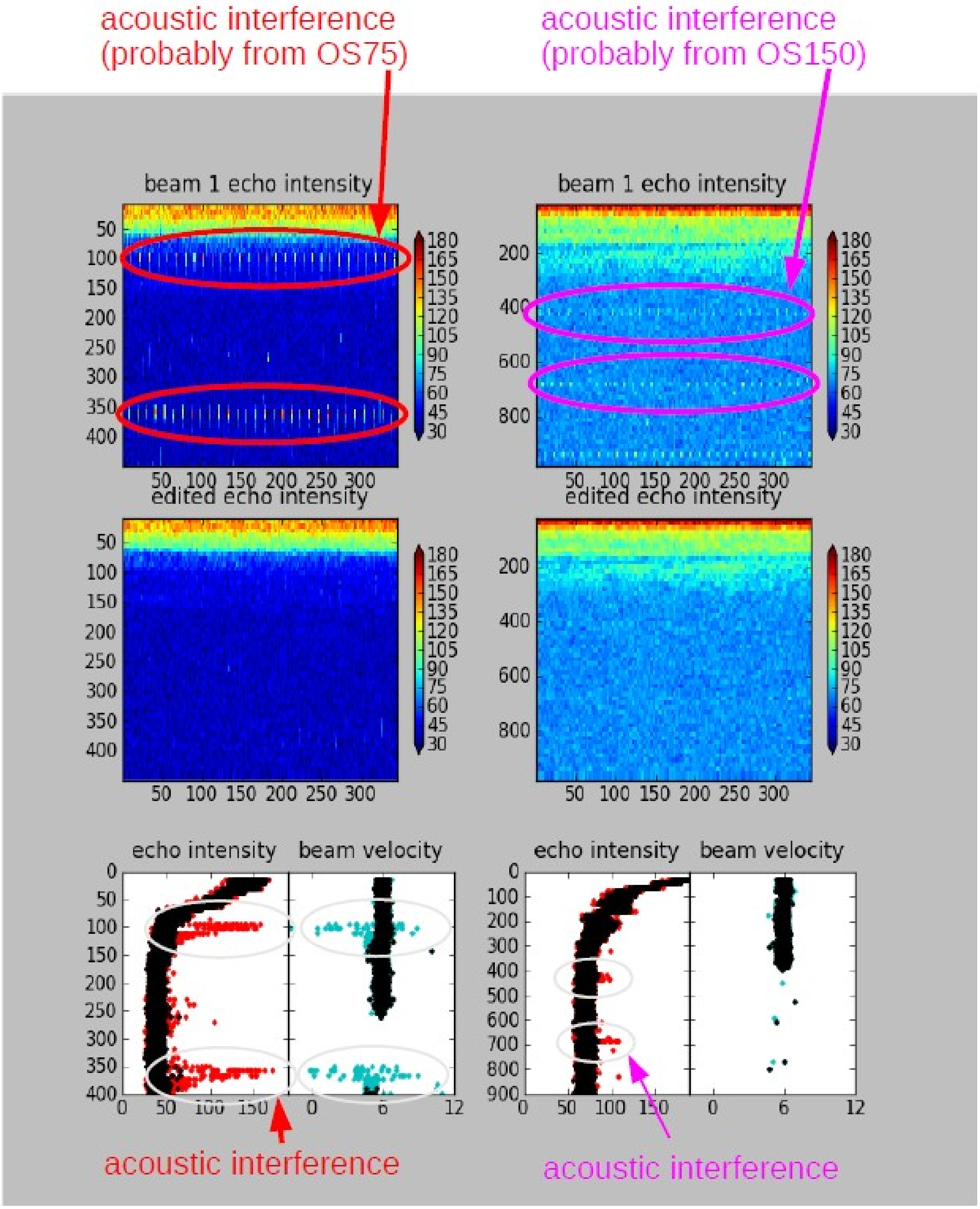


Figure 3: Effect of CODAS single-ping editing

Figure 4 shows the algorithm used to identify and remove the interference from the single-ping data. Note the obvious spikes in the OS150 Echo Intensity (red) and the corresponding contamination of the beam velocity (cyan), The black dots represent the remaining Intensity and beam velocity after the offending bins have been flagged as bad. This single-ping editing is what cleans up the blue stripes in Figure 3.



Another kind of “bias towards zero” can occur when the weather is bad (sea state is high) and bubbles are entrained under the hull. The bubbles absorb, reflect, and scatter the sound, as well as generating noise, so range is reduced. A bias towards zero at the tops of the profiles, especially pronounced in profiles with severely shortened range, is common, although its precise cause is not clear (Figures 5 and 6). CODAS single-ping editing can eliminate most of the bias, but an additional stage of manual editing of the averaged data is often needed. It can be a difficult judgment call.

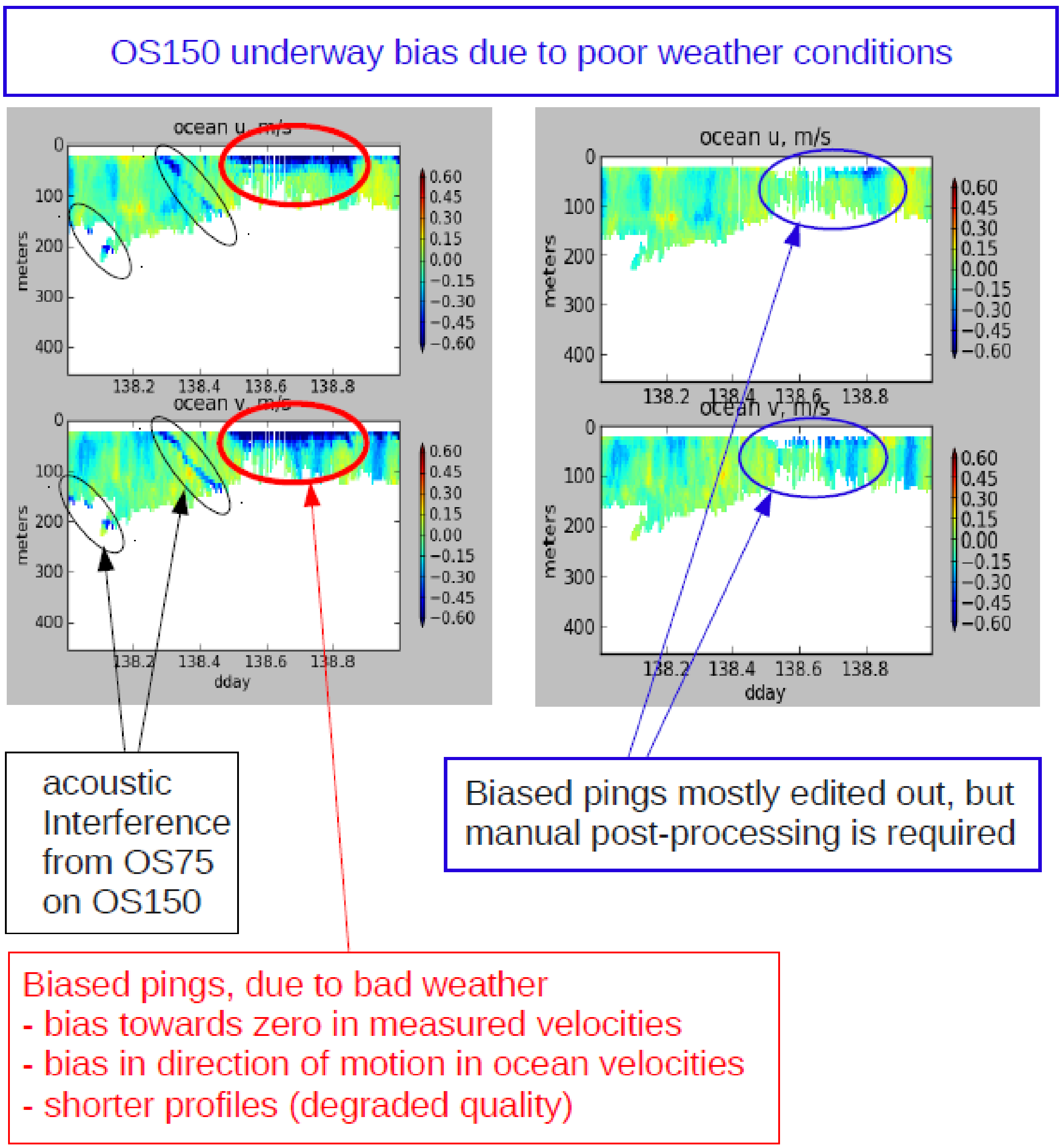


Figure 5: OS150 underway bias from bad weather (left), and the same data after CODAS singleping editing and averaging (right).

Although the OS150 was able to capture more bins (Figure 5) than the OS75 (Figure 6), velocity data around decimal day 138.6 are clearly biased (left). The OS75 did a better job of rejecting bad data, but some highly biased pings were not rejected. Reprocessing using single-ping editing (right) resulted in several hours of data are deemed unrecoverable.

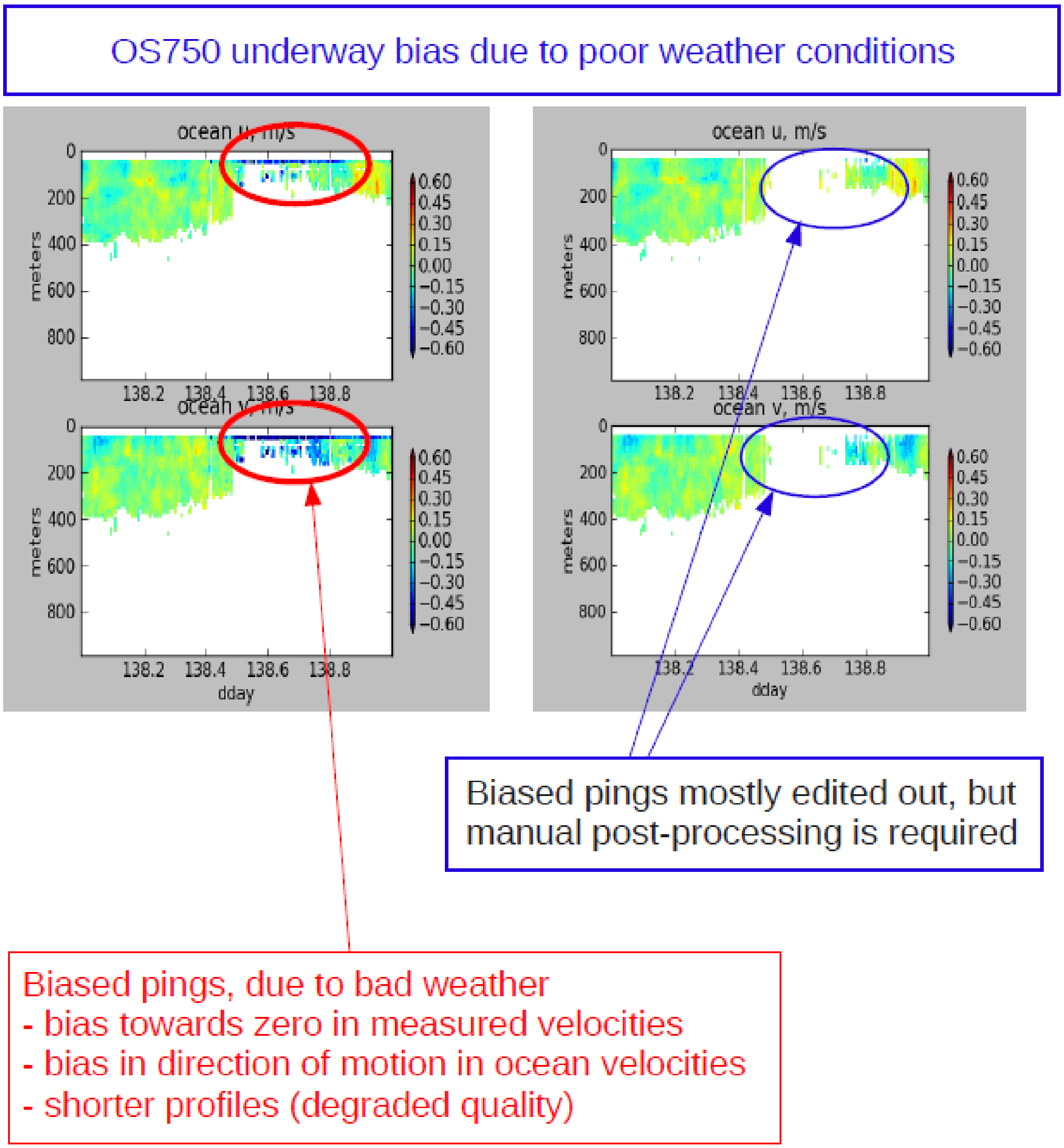


Figure 6: OS75 underway bias due to bad weather (left) and the same data after CODAS singleping editing and averaging (right).

5 Summary

CODAS processing provides tools to counteract several common problems in ADCP data. Examples were shown in which the effects of acoustic interference can be removed from the final dataset by editing at the single-ping level prior to averaging. See CODAS documentation for discussions about calibration and heading corrections.