CPOD.exe : a guide for users

C-PODs:

- are self-contained ultrasound monitors that select tonal clicks and record the time, duration and other features of each click to 5microseconds resolution. A period of sound is 'tonal' when a narrow band of frequencies within the range 20kHz – 160kHz contains more energy than the rest of that frequency range.

- By selecting which sounds to log C-PODs can provide highly sensitive monitoring for porpoises and other cetaceans within a data volume of around 8GB per year, where continuous wideband recording would be over 30TB per year.

CPOD.exe is a dedicated program that:

- carries out objective automated data analysis to find click trains in the data and identify those produced by dolphins, porpoises, and all echo-locating cetaceans except sperm whales.

- provides fast display and export of data.

- enables you to change C-POD settings

Quick start

1. Run the POD (see the User Guide for instructions) and read the SD card. This gives you a .CP1 file of raw data.

2. Run CPOD.exe and click the large green text, upper screen, to show the menu.

3. Click the 'Detect cetacean clicks train in CP1 files ..' button on the 'Trains' page of the menu.

4. A visual progress display follows, and useful text is put onto the Results page of the menu summarising findings.

5. When completed, copy the Results text to clipboard (button at bottom) and paste it into a spreadsheet. May be useful later.

6. Click the Open Files button, top left, and select either the CP1 file, or the newly created CP3 file. Both will be opened.

7. Click the 'Show from start' button and then use the 'Show next screen' button to move forward.

8. The default display will show the amplitude of 'NBHF' (porpoise–type) and 'Other cet' (all other cetaceans, including dolphins) clicks in Hi and Mod quality as lines colour-coded by frequency.

9. Experiment with different displays by going to the Files page of the menu and selecting different parameters and display scales.

10. Move to any time of particular interest by dragging the slider in the bar above the display area.

11. Select 'NBHF' or 'Other cet' on the Files page; Got To Export; Select data: N of clicks, period: 1min, omit negative time periods, and Export from file 2. You will find a list of minutes with detections on the Results page.

12. Read this manual for details on all this, and much, much more!

Display basics

In CPOD.exe:

The main MENU is shown or hidden when the mouse pointer enters the information strip at the top of the screen, (which carries information on the item currently under the pointer) or when the large green text ('Sp class' in this illustration) is clicked:
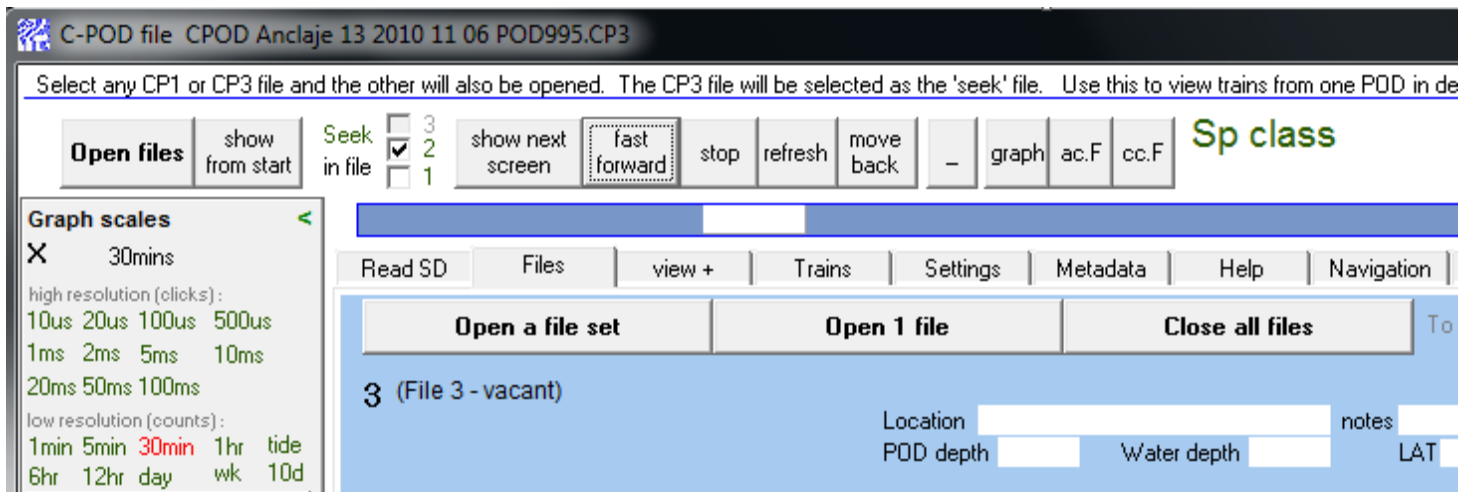
In this manual 'view+ page' etc. refers to the MENU page of that name as shown above.

Display options are shown in a panel at the right of the main screen, as above.

This panel can be hidden and restored by clicking < or >

Colour coding: blue text shows data, green text is clickable, grey text is information notes or disabled active text:
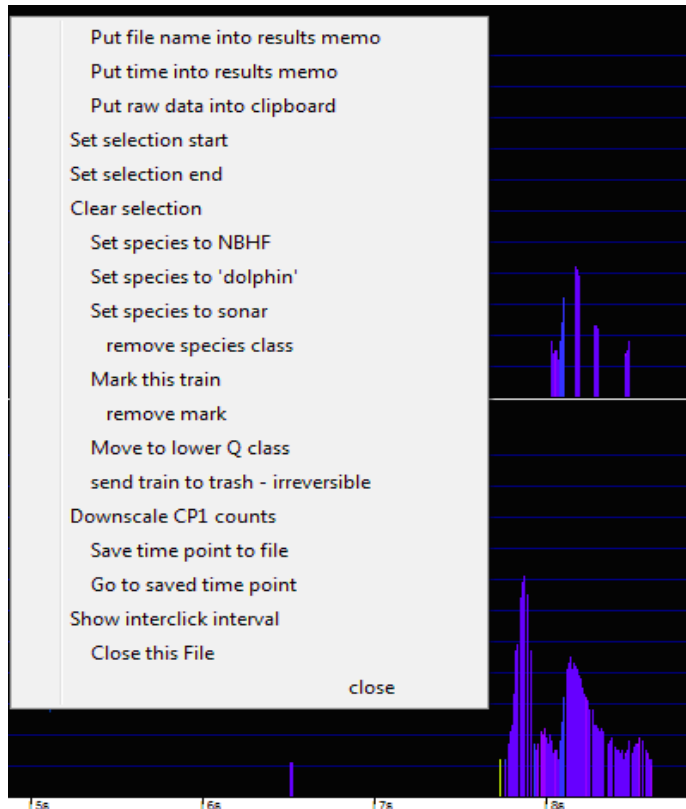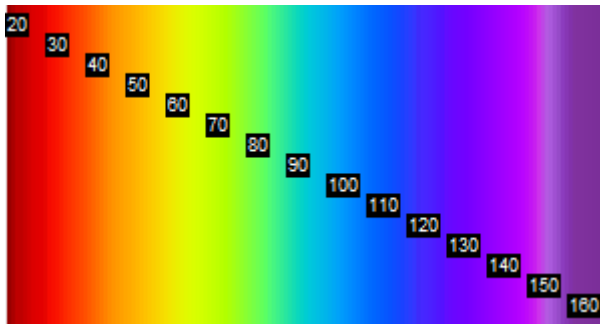
To: 31/1/2011 13:43
clear selection

save endpoints   delete saved endpoints

Pop-up menu: right-clicking in the display area of the screen shows a pop-up menu with items that mostly use time information derived from the position of the mouse pointer:

Put file name into results memo
Put time into results memo
Put raw data into clipboard
Set selection start
Set selection end
Clear selection
Set species to NBHF
Set species to 'dolphin'
Set species to sonar
   remove species class
Mark this train
   remove mark
Move to lower Q class
send train to trash - irreversible
Downscale CP1 counts
Save time point to file
Go to saved time point
Show interclick interval
Close this File
                              close

Frequencies (kHz) of sounds are colour coded:

System requirements

CPOD.exe:

- runs on any version of Windows.
- is portable – it requires no installation program and can be directly run from any drive, including a CD or memory stick.
- can be simply deleted along with any other POD-related files.
- requires an external USB SD-card reader (supplied with C-PODs) on most PCs, as the built-in readers are often not fully compatible at present (2011).
- reads data from the C-POD's SD card only into a directory with the word 'CPOD' (no space or dash) somewhere in the path or folder name. It's easiest to create a folder called CPOD data and put all data files in that, as the files have unique names.

For brief information on versions go to 'program versions' or read the 'CPOD software version history.doc' on the website for more detail than most people will want! If you need to use the old (first) version of the classifier, use CPOD.exe v1.054. It remains available on the website, but it gives a higher level of false positives.

File types, and cropping files

On the SD card:

C-POD SD cards are specially formatted and contain 4 files DATA0.CHE, DATA1.CHE etc that are empty shells that sit in exactly the same place on each card, so that the POD can write data without opening or closing files, and intact files are present even if the power was suddenly lost. SD cards wear out with repeated writing so they have a wear levelling algorithm, 'WLA', that moves the card directory frequently, and anything that triggers the WLA makes the card unusable by the POD. So avoid any action that might write a file to an SD card, however small the file is.

- SETUP.CHE – this holds the operational settings for the POD. The settings can be safely read from this file via the Settings page.

- DATA0.CHE, DATA1.CHE etc. – these are 'shell files' that occupy nearly the whole of the card. They contain as much data as has been written and the rest is 'blanks'. Old data is only erased just ahead of the writing process. Old data can be erased using the button, lower left, on the Read SD page, but it's slow.

- Other files – if any other file is present the card has been corrupted and will need re-formatting using special software that Chelonia can provide at no cost.

On the PC:

- **.CP1* – these are generated when the data is read from the SD card. They contain all the clicks logged, plus the temperature and angle of the POD to vertical that is read at the start of each minute.

- **.CP3* – these are generated when the CP1 file is processed via the button on the 'Trains' page. They contain only the clicks that have been identified by the KERNO classifier as belonging to a train, plus a record of temps, angles, and the number of clicks in each minute of the CP1 file. They package train information with each click.

- a '*File set*' is a CP1 file and the corresponding CP3 file, plus any map files associated with them (see below).

- **.c1m, *.c3m* – these are map files that allow the program to move rapidly to any time even in large files. They must be in the same directory as the CP1 or CP3 file. If deleted the file will be recreated and nothing has been lost except a little time.

- **.txt* – many export functions create text files that can be imported very easily into databases, spreadsheets and word-processors.

- **.cpe* – export settings files. You can store and recall your export settings.

- *CPODcodes.txt* – a file of defined format that holds user-defined codes that can be read from the file 'Notes', if present, and exported. This allows files to be classified easily in later processing.

- *CPOD sightings entry.xls* – an Excel spreadsheet that can be used to enter species sightings and then write these into a CP3 file, where they can be displayed as text annotations in high resolution displays.

- **.arr – array files*. These are large files generated from sets of CP3 files that have been named (or renamed) with their Y,X position on a rectangular grid, so that detections across the whole array can be displayed at speed.

- *Others* – *CP2 files are created and normally removed during the process of train detection and classification. They are similar to *CP3 files, but hold every click and are consequently very large (up to about 5GB), and can be safely deleted.

Opening data files

- The 'Open files' button (see first graphic) opens a file pair – a CP1 file as File 1 at the bottom of the screen and the corresponding CP3 file in the upper panel. This is the normal viewing mode that allows you to see the relationship between the raw and processed data. The 'Open file pair' button on the 'Files' page does the same.

- From the 'Files' page you can open any combination of CP1 or CP3 files instead, so you can, for example, view files from a vertical array one above the other.

- If no files are open right-clicking the display area clicks the 'Open files' button, or left-clicking opens a single file.

- To set CPOD.exe as the default program to open these files: right-click on a .CP1 file; select 'Open with'; find CPOD.exe; Select 'always open this file type with ..'. Repeat for .CP3. Then when you double click a file CPOD.exe will start and the file name will appear in the 'Open file' dialog when activated.

- Array files are opened by the 'Start array viewer' button on the 'view +' page.

Closing data files

- via the right-click menu

- by clicking the file name on the 'Files' page.

- by clicking the 'Open files' button on the main screen or the 'Open file pair' button on the 'Files' page.

Cropping data files

It may save time later to remove the ends of the file where the POD was not in the water, or the boat sonar was still being heard. This cropping is also useful for selecting small parts of a large file to send to others.

To crop a file:

1. Open a CP1 file or a file pair.

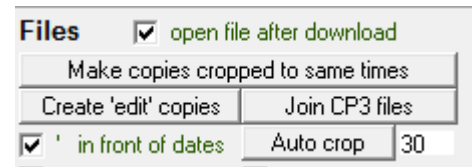2. Set a time selection enclosing the part you wish to retain.

3. Click 'Copy selected time period into new file(s)' and both the CP1 and CP3 file, if open, will be copied producing files with the length of the selection added, here it's 'A0b 912 PART 1m.CP1'

Note: the CP3 file holds some data on the number of clicks in the CP1 file, but these will now refer to the original CP1 file. If you re-process the cropped CP1, then the CP3 record refers to that shorter file.

Auto-crop. In the process above you looked at the file to see when the POD was in the water and the boat noise had died away. If you have many PODs you might prefer to use 'auto-cropping'.

A batch of CP1 files can be selected and those longer than one day will be cropped to the time period within the earliest and latest continuous *n* minute periods of logging. *n* is taken from the adjacent box with 30minutes as the default. New cropped files are created. You need to make new CP3 files from these. Check that your new cp1 files have been correctly cropped. For example, if the POD(s) switched on briefly during an otherwise off period, this might cause Auto crop to keep more of the cp1 file than you thought. This would affect some values on the Analysis page.

Data volume

You can assess the *SD card fill time* from the data from a short deployment: go to the 'analysis' page and analyse the CP1 file. The output includes, e.g. '1GB lasts 320d 0h 29m' so you can calculate the time for a 4GB card. In sites with sediment transport noise you should divide the estimate by two but the very worst scenario is where every minute reached the limit of 4096 clicks per minute, giving about 60 days logging from 4GB.

Analysis of wideband hydrophone data files can be so time-consuming that large amounts of acoustic data sometimes remain un-analysed after expensive collection. To lighten this burden:

C-PODs:

- select data – only tonal events are logged
- compress data – storing a summary of each tonal event

so the data volume is typically reduced by a factor of 8000 compared with wideband data.

CPOD.exe:

- identifies trains, further reducing the data volume by a factor of 10 – 5000.
- provides powerful filters to isolate subsets of interest.
- 'seeks' rapidly through data skipping everything excluded by filters in force.

Processing raw data to find cetaceans
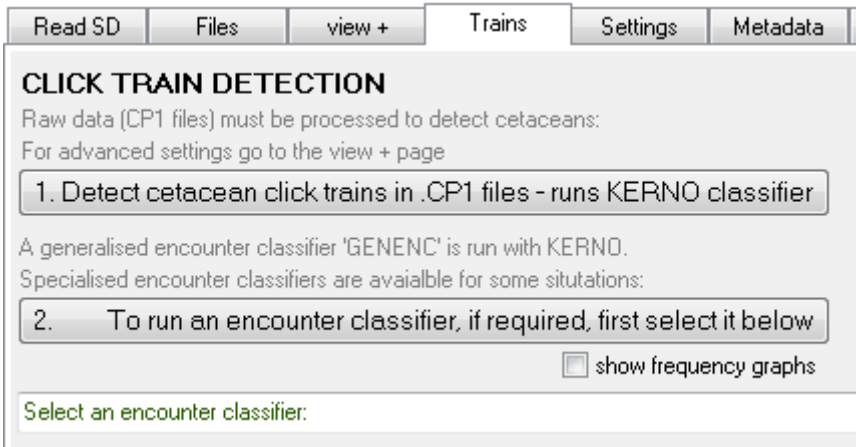
The normal workflow goes:

Read SD card – generates CP1 files

Look at these and crop them to the limits of the good data.

Process the CP1 files to get CP3 files. This is done on the 'Trains' page via the 'Detect … KERNO classifier' button. An encounter classifier 'GENENC' is also run automatically. Other encounter classifiers can be used subsequently on the file if required.

CLICK TRAIN DETECTION

Raw data (CP1 files) must be processed to detect cetaceans:
For advanced settings go to the view + page

| 1. Detect cetacean click trains in .CP1 files - runs KERNO classifier |

A generalised encounter classifier 'GENENC' is run with KERNO.
Specialised encounter classifiers are avaialble for some situations:

| 2.   To run an encounter classifier, if required, first select it below |

☐ show frequency graphs

Select an encounter classifier:

If you have years of data this process may take many hours – an evolving estimate of the time required is shown. The process can be stopped using the 'stop' button at any time.

If you want to process all CP1 files in a directory and all of its subdirectories you can opt to do this by saying yes to "Do you wish to select all files in a directory and its sub-directories?" You are given this option when clicking either the KERNO or Encounter classifier buttons.

Viewing data: high resolution

High resolution displays show individual clicks and range from 10microseconds per screen pixel to 100millisconds per pixel.

Selection is at the right of the main screen:

or by zooming.

At 10μs resolution a continuous display of 6 month's data would require a computer screen (mostly empty) about 50,000km wide, so CPOD.exe moves to (seeks) the next click as you go forwards.

The SPL (the maximum peak-to-peak pressure within the click), duration, frequency, frequency trend, bandwidth and some envelope parameters of each click can be viewed.

You can select which parameter to display:
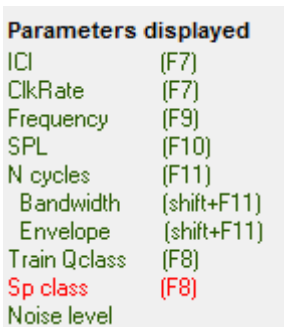

ICI, Click rate, Train Q class, and Sp class are only available from CP3 files, so any CP1 file on show willshow the last parameter selected that is possible from a CP1 file. This can be useful for understanidng sources of marine sound.

At all resolutions except the two highest (10μs and 20μs) it is possible for more than one click to start within the time covered by one pixel-width of the screen. In those cases both clicks are drawn, but some may not be seen. In SPL displays this produce 'multi-coloured clicks':
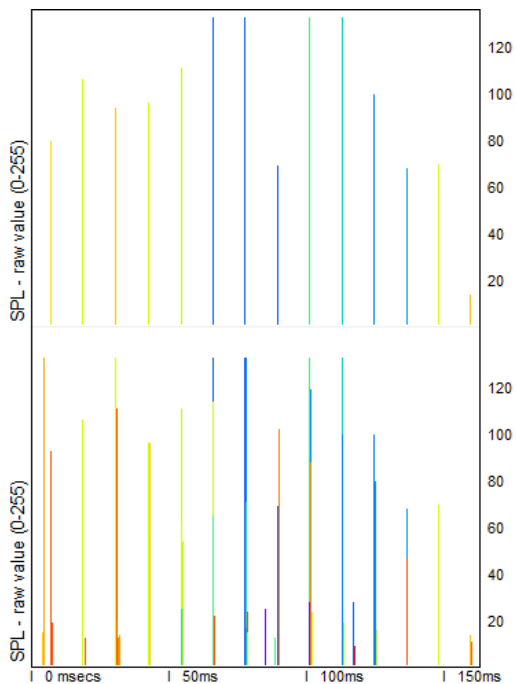

The total number of clicks within the screen is shown at the bottom right of the display screen, for each file (f1, f2).

This shows a dolphin click at high resolution:
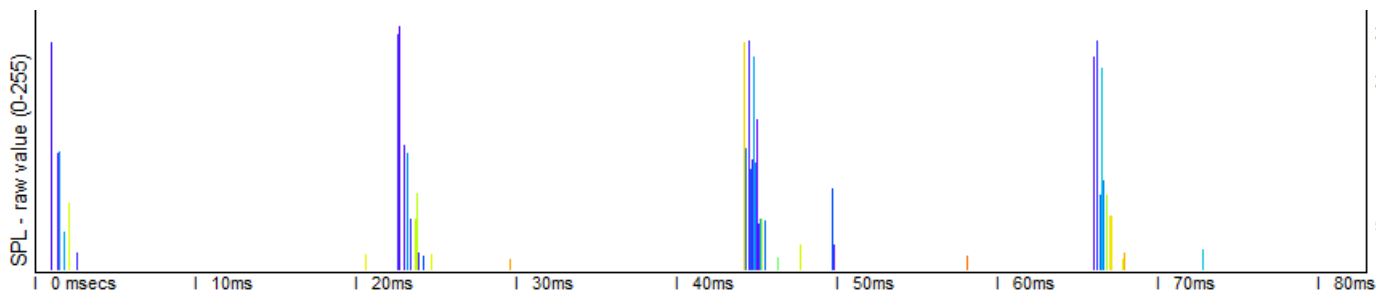
In the CP1 file the clicks arrived as multipath-replicates, that show variation in frequency (colour-coded in Sound Pressure Level displays), and great variation in SPL (height). These multi-path replicates are seen in many cetacean and boat sonar click trains and arise from sound reflections (echoes), and possibly from sound refraction as well.

The CP3 shows a single click selected as representing the multipath cluster:
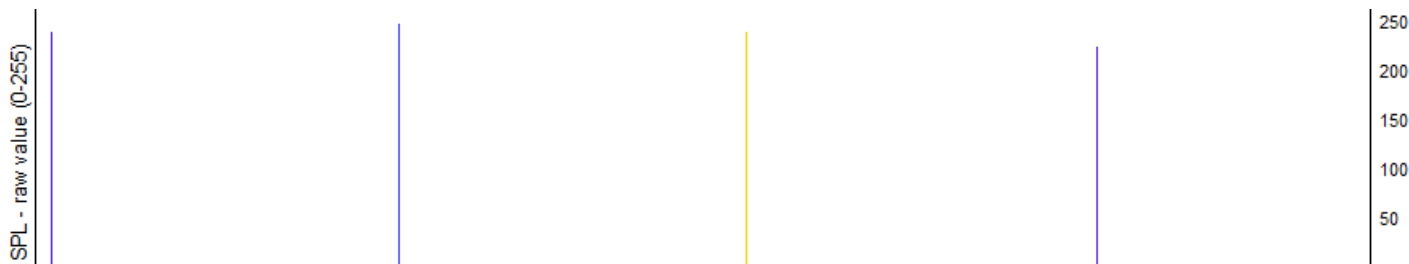
Viewing data: low resolution

Low resolution displays show counts of clicks over periods ranging from 1 minute to 10 days, as bar charts:
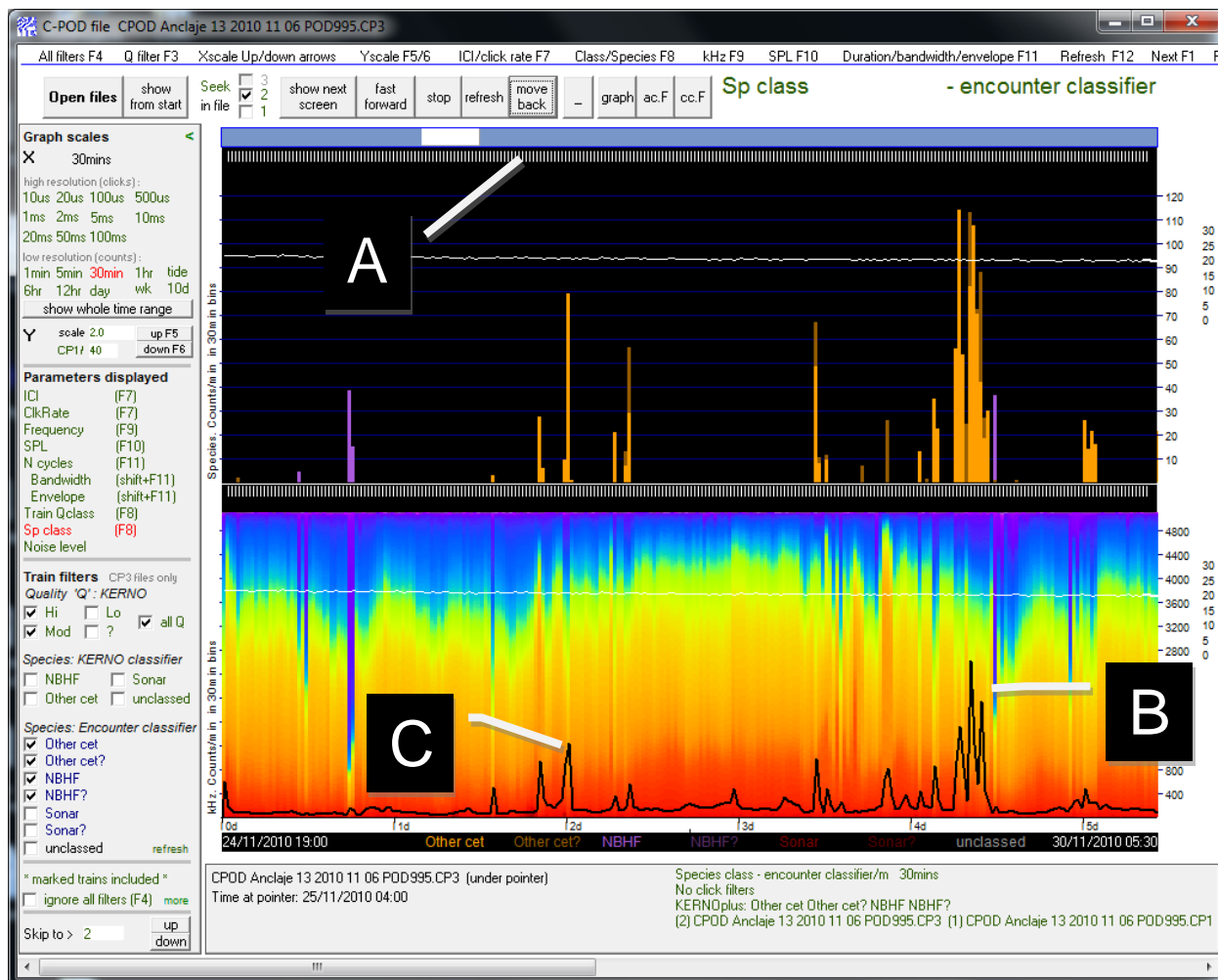
A: vertical lines show the angle of the POD to vertical. The line is white if the POD is ON, red if the POD is OFF.

B: stacked bars showing frequency distribution, coded by colour. In this location bottlenose dolphins (orange in the CP3 file shown in the upper panel) are producing lower frequency noise and porpoises (purple in the CP3) are producing higher frequency noise. Sonars (maroon in the CP3) produce blocks of single colour when their narrowband emissions dominate the sound spectrum.

C: the black line shows the mean number of clicks logged per minute. This may be limited to 4096 clicks per minute (to avoid memory overload), and if that limit is being reached in most minutes in the interval displayed (6 hours in this case) the clipping shows as a flat top level to this line – see sediment noise.

Selection of time resolutions is on the 'Files' page, or see quicker ways to select time resolution.

The zoom range from 10 microseconds per pixel to 10 days/3 pixels is about 1: 30 billion and it's all useful!

Low resolution displays can be slow because of the time taken to read a large file, and you can stop a display, changed the settings and start it again using the 'stop' and 'show next screen' buttons.

Time selections

A period of time within a file can be selected for:

• making a new shorter cropped file created from the original. The original file is retained unchanged.

• analysis of cetacean encounters, or periods of noise.

• extracting data for export from the same time period in many files, e.g. in a calibration study or impact assessment. A *quick alternative* for some purposes is to click the '*Summary of displayed minutes*' button at the bottom of the 'Export' page.

Time selections:

TO SET:

- To set a selection running from the start you only need to mark the end, and vice-versa.

- Right-click on the display area with the pointer at the time required, when a file is open, and select 'Set selection start' or 'Set selection end'.

or

- Enter the time required in the 'Time POD STARTED' boxes at the top of the 'Read SD' page, and then use 'Times' on the 'view +' page:

Times:
Set selection start
Set selection end
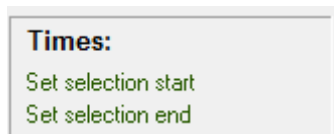
or

- Recall saved endpoints by clicking the green file duration label at the right:

1 Mousehole pinger f2 POD495 file01.CP1

POD 495

| Location | | notes 21 h cycle |
| POD depth 0 | Water depth 0 | LAT | LONG |
Start 12/8/2012 12:46   End 28/9/2012 15:50   47d 3h 4m/46d 23h 58m
whole file/stored endpoints

Range 46d 23h 59m (selection)   From: 12/8/2012 14:48 (selection)   To: 28/9/2012 14:46 (selection)

TO CLEAR:

- Click 'clear selection' on the 'Files' page – bottom left in the graphic above.

- Or select 'Clear selection' from the right-click menu in the display area.

- To remove the endpoints saved in the file click 'delete saved endpoints'.

TO SAVE:

- Click 'save endpoints' on the 'Files' page – bottom left in the graphic above. This saves the endpoints to all open files.

Time selections cover whole minutes.

When no time selection is set, the maximum range covered by all the open files is used.

Click and Train Parameters

Clicks

The following parameter are stored for clicks:

- Time of start, to 5 microseconds resolution. The minute, and the number of microseconds in the minute, are shown separately.

- Duration: represented as the number of cycles at the dominant frequency. This is a more useful measure than absolute duration as, across frequencies, it correlates better with other useful measures.

- Frequency: a single value representing the dominant frequency based on zero-crossing intervals. The ZC method used gives no weighting to louder cycles in the click, in contrast to an FFT spectrum. For narrow band clicks they give virtually identical values, but for broadband clicks the difference is great. Multipath replicates at different frequencies are often logged from broadband clicks.

- SPL: sound pressure level = 'Pmax' the maximum peak-to-peak range of the click based on point values of pressure. Measurement of SPL is limited to a small range at the lower end of the huge dynamic range of the system, and is logged on an 8-bit (0-255) scale. This scale is not corrected for the frequency response of the system. Corrected values in Pascals can be viewed or exported by selecting the 'show SPL as Pa' option on the 'view +' page.

- The weakest detected porpoise clicks have a peak to peak SPL of around 1 Pascal. No conversion to decibels is made.

- Bandwidth: a measure on an arbitrary scale of bandwidth. Zero bandwidth sounds are pure tones. High bandwidth sounds resemble white noise, cracking, crunching and other such sounds that have no easily

recognisable pitch. This parameter, and the next 3 are of low value because the analysis window from which they are derived during logging may include some non-click time.

- Envelope measures, slope 1 and slope 2, representing the slope from the max SPL of the first cycle to the peak SPL of the click, and from the 5th cycle to the peak.

- End Frequency: a single value derived from the last zero crossing interval.

The more important of these can be captured from a screen of CP1 data via the right-click menu, and pasted directly into a spreadsheet, for example:

time	µs	avF	endF	nCyc	Pmax
26/11/10 08:35	49064540	103	125	5	249
26/11/10 08:35	49064690	47	60	5	13
26/11/10 08:35	49070855	104	121	5	248
26/11/10 08:35	49078340	103	121	5	248
26/11/10 08:35	49086335	105	125	5	248
26/11/10 08:35	49094935	106	128	5	251
26/11/10 08:35	49095020	109	125	6	29

… or they can be exported via the Export page.

Trains

Physical properties:

- Modal kHz: This can be unstable when the number of clicks in the train is low.

- Number of clicks in the train: trains of less than 5 clicks are not recognized by the KERNO classifier in CPOD.exe.

- Clicks/second: = 1/mean inter-click interval, rounded to the nearest integer.

- Mean SPL: this uses the raw SPL value, not the corrected value in Pascals.

time	uS	avF	endF	nCyc	Pmax	Train
26/11/10 08:35	49070855	104	121	5	248	96
26/11/10 08:35	49078340	103	121	5	248	96
26/11/10 08:35	49086335	105	125	5	248	96
26/11/10 08:35	49094935	106	128	5	251	96
26/11/10 08:35	49104290	118	128	9	247	96
26/11/10 08:35	49113830	119	128	9	246	96
26/11/10 08:35	49123870	116	108	9	245	96
26/11/10 08:35	49134695	109	121	10	250	96

These click data are captured to the clipboard by right-clicking the CP3 file and then pasting them into a spreadsheet. They include fewer clicks than the corresponding period of the CP1 file, and show a train identification number (right column) that can be used to graph overlapping trains separately.

Classified properties from the **KERNO** classifier:

Four properties are produced by the KERNO classifier and have more or less independent scales:

- Q class: this is an assessment of the chance of the train arising from a non-train source, such as shrimp clicks, rain etc. The classes – Hi, Mod, Low, ? (- 'doubtful') are defined by different detection thresholds, so the sensitivity and risk of false positives fall as you move up the classification. – more below in Data Filters.

- Species class – more below in 'Species' and in 'Species identification.

- Species classification quality: High or not.

- ICI classification quality. A train may be reliably identified as coming from a particular source, e.g. NBHF cetaceans, but the actual selection of clicks in the train may be unreliable due to surface echoes or missing clicks due to spikes of ambient noise.

- WUTS risk: some trains have features that are also seen in 'Weak Unknown Train Sources' – see below. Any large set of trains from porpoises or dolphins will contain some trains with these features so WUTS are not classified as a separate species, but a high incidences of WUTS calls for visual validation of the CP3 data.

Classified properties from **users**:

- Marked trains: you can mark any trains via the right-click menu, and then exclude these trains from analysis or exclude all other trains. This is a useful tool for exploring the properties of some set of trains that you think you can recognise by eye.

CP3 files can also be edited to change the classification of trains, and marks can be cleared from the whole file.

Note:

Frequency filters can be set:

on trains, where they apply to the modal frequency of the train, *or*

on clicks where they apply to the frequency of individual clicks.

SPL filters can also be set

on trains, where they apply to the mean SPL of the train, *or*

on clicks where they apply to the SPL of individual clicks.

Species

When CP3 files are generated from CP1 files (via the 'Trains' page) the initial classification of species is into:

- NBHF: species producing Narrow-Band High Frequency clicks. NBHF species include all porpoises, *Phocoenidae*, all dolphins in the genus *Cephalorhynchus,* some dolphins in the genus *Lagenorhynchus,* the dwarf and pygmy sperm whales in the genus *Kogia,* and *Pontoporia.*

- Other cetaceans: 'other cet' refers to all non-NBHF toothed whales, *Odontocetes,* except the Sperm Whales which clicks at too low a frequency. This class is referred to as 'dolphins' or 'dols' in places in the software where short words are required.

- Sonar, pingers etc.: depth finders, fish finders, some pingers, VEMCO fish tags, and, surprisingly, some Acoustic Doppler Current Profilers. VEMCO fish tags are typically logged at close to 70kHz. ADCPs are usually operating at much higher frequencies than the POD can detect, but commonly do also produce undocumented lower frequency sounds. All these sources can significantly interfere with the function of a POD that is near such a sound source.

- Unclassified source: could be any of the above, or could be a spurious train arising by chance from unrelated sources like shrimps, sediment noise, or WUTS, that occur in some seas, and resemble cetacean click trains.

WUTS – weak unknown train sources – are found in some seas and are *not* classified as a species because they overlap with the characteristics of NBHF or other cetaceans too much. Instead, each train carries a 'WUTS risk' evaluation which has a high rate of false positives (i.e. a cetacean train is given a WUTS risk), but is still useful as a basis for creating a warning when the fraction of cetacean trains with a WUTS risk is unusually high and visual validation is needed. WUTS may be some small crustacean doing something a bit noisy.

WUTS features are:

- linear increase in ICI

- shortest ICI may be <2ms

- short clicks

- no multipath

- SPL envelope of train often flat

- may show frequency modulation through the train

- frequency often around 120kHz or 80kHz but a range of values is seen

- usually isolated events against a quiet background.

'My Species' – you can *mark* a subset of trains that you see as distinctive and the filters on the 'view +' page allow you to select or exclude only this subset, so that you can export its characteristics.

Encounter classifiers

These do not alter the KERNO output, but add power to the classification by taking a wider time-span into account. They take discrete periods of time of visits by individuals or groups of animals to the vicinity of the POD and use acoustic information from the whole period.

Data Filters and which to use

In CP3 files all clicks are in trains, and the train filters, Quality, Species, click rate, modal frequency, N in Train, mean SPL, can be applied very quickly during display and export. Train filters are ignored if display or export is from a CP1 file.

Standard filtering

When a CP3 file is open, the Metadata page shows warnings if there are any. In the absence of warnings, false positives are likely to be below 15%, and are typically very much lower. If this is an acceptable rate of false positives you can select Hi and Mod Quality trains, plus whichever of NBHF, 'Other cetacean', or 'Sonar, pingers etc.' you require, or select a Species Group (when this feature is available). If you need to be sure the rate of false positives is lower than 15% you should visually validate a sample of detections, say 100 trains, to get an idea of error rates. In most case this will be very low and you can proceed with standard filtering, but if the rate is too high you can either:

- edit the data to remove false detections or mark true detections, or

- visually assess Hi trains to see if this sub-set is sufficiently accurate, and use the analysis page to see if there are enough of these detections to meet your statistical power target.

For advanced users: In many data files from quieter locations, analysis of NBHF trains could safely include Lo and ? (doubtful) quality classes, but this always requires visual validation.

This means:

- Quality: Hi or Moderate
- Species: NBHF (from KERNO classifier)

For advanced users: To explore the data in depth click filters can also be set. They are on the view + page. The default values can be restored by clicking in each number box that has been changed.

If an encounter classifier has been used on the CP3 file there will be text against the check boxes in the 'species from encounter classifier' section:

Up to six species, including sonars, identified by any encounter classifier may be shown. Whenever a file that has been processed by an encounter classifier is opened the new species potentially identified by that classifier are added to the list and can be filtered in any batch process.

If a species from an encounter classifier is checked the Quality filters are automatically all selected via the 'all Q (F3)' as the encounter classifier will have taken these into account. However you can still make a narrower selection from the Quality classes, but you will need to evaluate the effects yourself. The KERNO classifier species are not available as filters if an encounter species is selected. From the encounter classifier 'Other Cet' will only give you highly reliable dolphin detections, while 'Other Cet?' will add some that are less reliable. and some visual validation is a good idea.

For more information see 'Encounter classifiers'

If **click filters** are set when exporting or viewing train data from a CP3 file the results can be puzzling, as the *train* may have individual *clicks* excluded from it and be represented by only one or two clicks that are far apart in time.

The F4 function key provides a quick way of seeing all clicks or trains that have been excluded.

The 'Files' page has settings for some other filters:

Most users will not need to use any of these options.

- *ICI assessment quality* is only relevant to studies of behaviour as shown by ICIs.

- *Species classification quality*: this is just a higher bar set for the species classification, and may prove useful as part of validation but it has not been evaluated yet, and the relevant norms are not known. However it is clear that it will remove a much larger fraction of 'other cet' trains than 'NBHF' trains for the reason discussed in 'Species identification'.

- *WUTS risk*: see Species identification and Metadata.

- '*Marks*' are added by users editing files.

Navigation and Editing

Numerous tools exist in CPOD.exe to enable you to find and view whatever interests you in the data.

In high resolution displays a 'Seek' function moves through the files until some clicks or trains are found that meet the filter criteria. The file you 'seek' in is normally the CP3 file opened as file2, but it can be any open file, that you can select here:

If no file is selected as the 'seek file' then long series of blank screens may exist during periods of silence, as if the program has stuck. You can define how many clicks must be shown, so that screens with less will be skipped. This value is set at bottom left of the screen, and the actual number displayed is also shown:
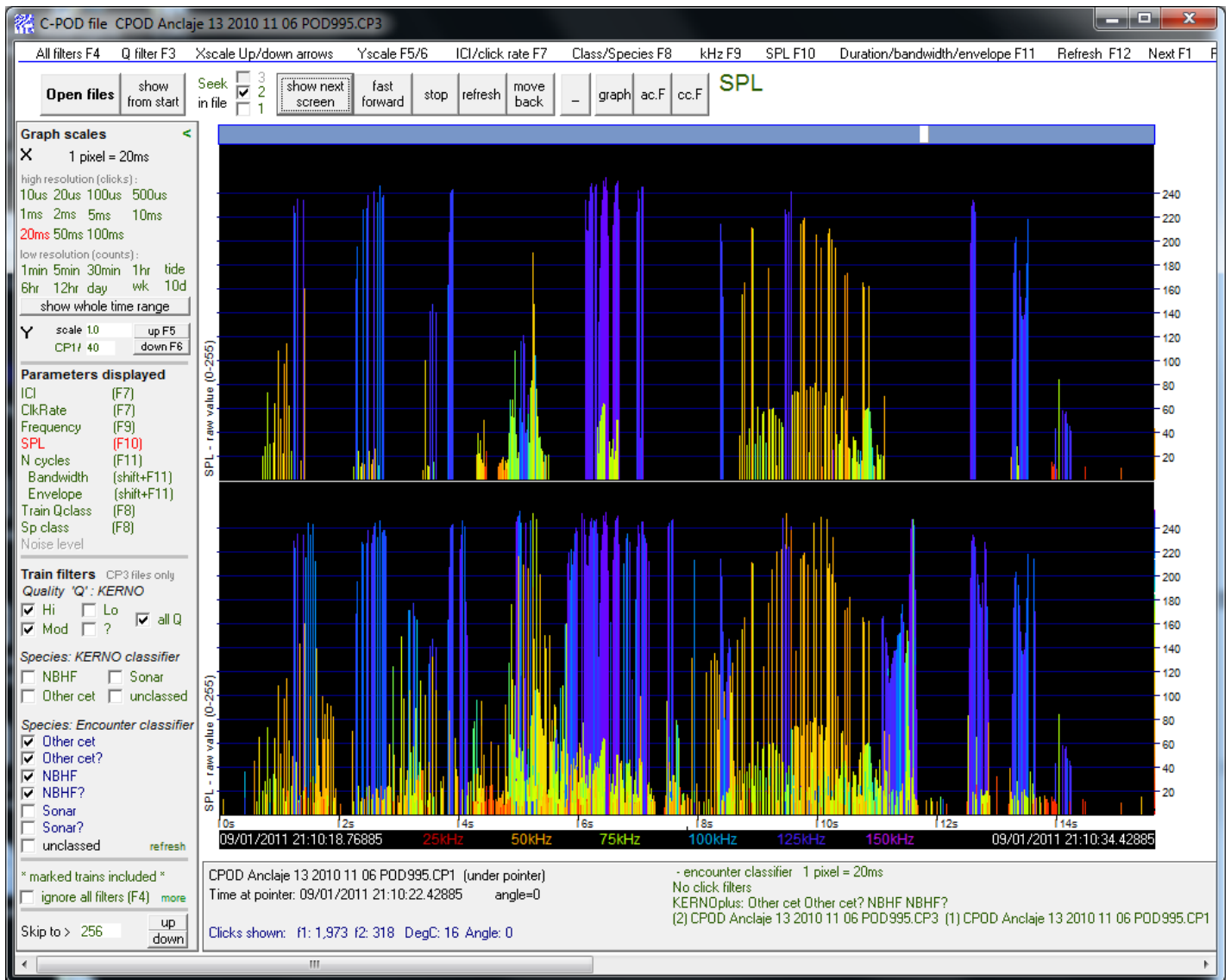
Navigational short-cuts:

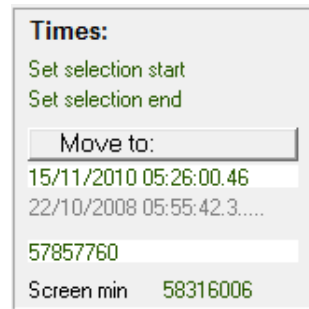| Task	Key
Show next screen	F1 (Function key 1)
Move earlier	Left arrow shows data that is to the left of the pointer
Move back	Shift F1
Move later	Right arrow shows data that is to the right of the pointer
Move to file start	Home
Move to file end	End
Scroll forwards / backwards	Mouse wheel – moves forwards or backwards ½ screen at a time without skipping
Move to a time	Drag the time bar slider to move to a date and time shown on the slider.
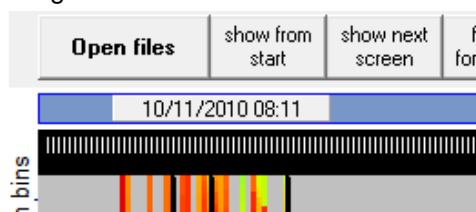
In high resolution displays you should stop before the time you want, as the view will

	skip to the next train.
Move to a time	Enter a time in the Move to box on the 'view+' page:

Times:
Set selection start
Set selection end

Move to:
15/11/2010 05:26:00.46
22/10/2008 05:55:42.3.....

57857760
Screen min   58316006

	The time format is critical and is shown below the edit box.
Move to last high resolution screen	Ctrl Z – move back through the last 18 high resolution screens.
Zoom in	Down arrow. The new view has the time at the mouse pointer position at the centre of the screen.
Zoom in	Left drag to right – zooms in to highest resolution that can show the time period marked by the yellow line drawn at the top of the display area. The step from low resolution to high resolution may be too big for this method.
Zoom out	Up arrow. The new view has the time at the mouse pointer position at the centre of the screen.
Move to next screen with a difference between files.	Ctrl C – useful to compare edited versions of a CP3 file.

Display short-cuts:

show/hide results page	F2 (Function key 2)
return to graphical display	F12
ignore/apply Quality filters	F3
ignore/apply all filters	F 4
Yscale up/down	F 5/F 6. Adding a digit ahead of the value in the Yscale edit box is a quicker way to make a big change:

Y scale 2.0   up F5 / down F6   CP1/40

Downscale CP1 counts	You can downscale the graph of a CP1 file by replacing 40, in the screen grab above, with a larger value. This way you can get comparable scaling to help show the contribution of train sources to overall noise.
	The right-click pop-up menu also allows you to downscale CP1 counts.
Put time into results memo	Ctrl T *or* right-click menu. These are in a format that can be pasted into the 'Move to:' box to review any time of interest.
Put a white-background version of the display onto the clipboard	Ctrl G. Graphics for documents.

Parameter selection

|	
train click rate/ICI	F7
train quality/species class	F8
click frequency	F9
Sound Pressure Level	F10 Can be raw values or Pascals – selected on view + page.
Duration	F11
bandwidth / envelope	Shift F11
Step through kHz filters	Ctrl F. Then key F steps through filters in 20kHz bands. 'Stop' button stops it.

The right-click-in-the-display-area pop-up menu has some shortcuts too, and in these the time, or the file number, is taken from the exact position of the mouse pointer:
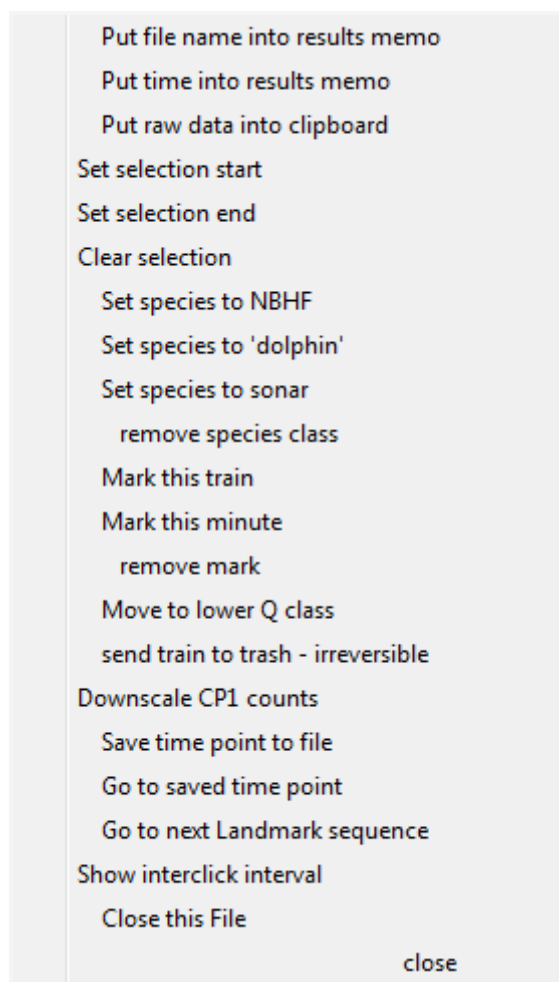
Editing tools

When is editing needed? For most purposes you would look at a sample of trains – say 100 – from different parts of the file and see how many look wrong. If the result is well below the level of errors in your method you don't need to spend a lot of time hand-weeding 0.3% of trains that are errors. But if you are working in a very low density area – say, studying Baltic porpoises – you may have very few true positives and these could be outnumbered by 1 false positive per month. Then editing is vital. (For more detail see the document 'Validation of porpoise detections' on www.chelonia.co.uk)

The fortunate thing about this is that it is in low density areas that editing proves necessary, but in those areas there are few positives to look at, so it's not too time consuming. A team of 7 people, some unfamiliar with CPOD.exe, recently edited a data set containing 2.6 billion clicks and 65 POD years of data from Puck Bay, in Poland in 2 days.

Editing or marking

Two methods are possible:

1. Changing or removing the species (NBHF, other cet, sonar) class allocated to a train. This leaves no trace of the change, but you can carry out the process on a copy of the original CP3 file, and compare them, train by train – see navigation above. You can make 'edit copies' of a set of files by clicking a button on the 'view +' page.

2. Marking trains. This is more powerful. You can filter out marked trains from their species classes, or view only marked trains. You can clear all marks. The original classification is still present.
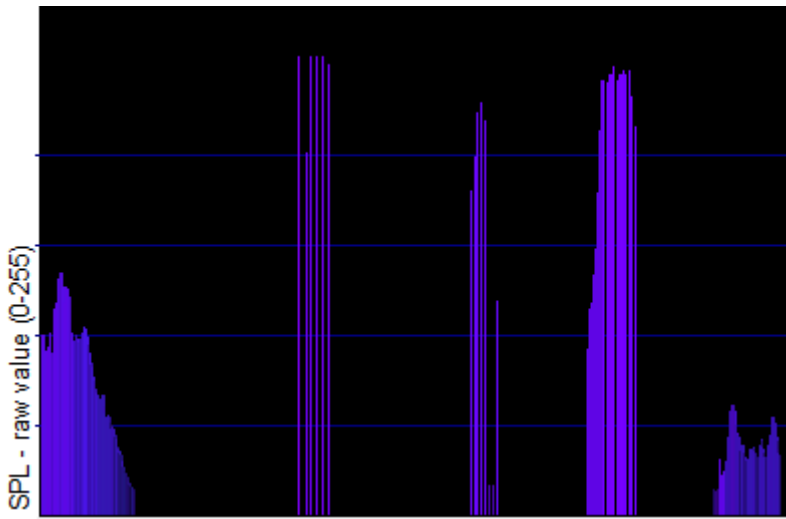
The main drawback of method 2 is that you have to remember to apply the filter you require for the marked trains, as well as any other filters.

Enhanced view

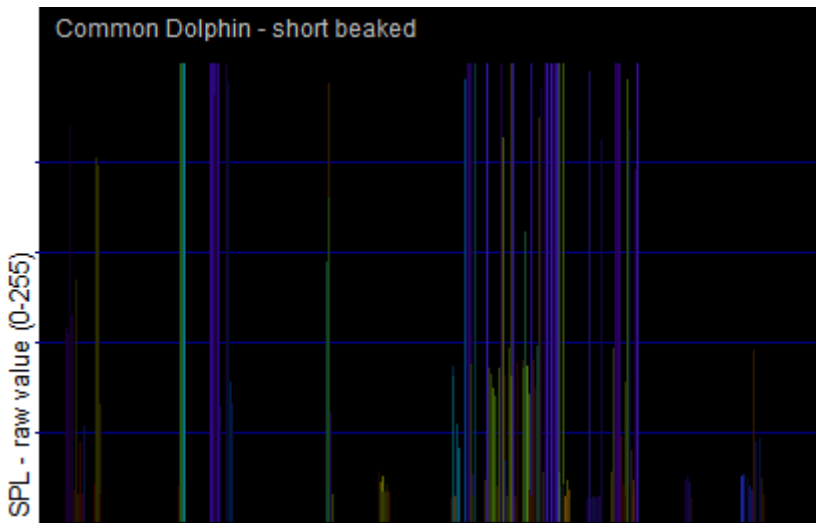Two tools are available via the 'view+' page:
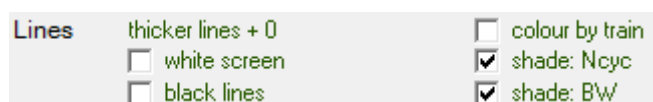
This shows porpoise clicks – they are not much affected by the shading because they are long and low bandwidth:

… but you can see the effect on the weakest clicks, as they are typically shorter and more broadband.

While these non-NBHF clicks look much duller:

The shading by click duration (Ncyc) is more powerful and useful.

In effect the view can be tuned to show 4 parameters – time, amplitude, frequency and duration, in one view.

Scrolling

The mouse wheel allows you to move backwards and forwards through the data around a train without automatically skipping to the next screen with a train – that might be days away.

Retrace

Ctrl Z loops back through the last 18 high resolution screens. These screens are 'live' – you can move on from, or modify them in the normal way.

Dual filtering

Mostly you will want to view trains of Hi and Mod Q class, as these are the ones you will use in statistical analysis. But you will want to see trains of Lo and ? class around them to help evaluate them. To avoid constantly switching between them using F3 you can set the filters to include all Q classes and then click 'Seek in' and it will change to:

Now the program will move to the next Hi or Mod train, but will show all Q classes selected around it.

Step through frequency bands

Ctrl F filters the lowest 20kHz. Pressing F loops through successive 20kHz bands and then restores no filters. The 'stop' button – you guessed - stops this.

Metadata

The Metadata page of the menu can show four sets of data from each file. Warning of possible adverse features are shown only if it was processed with the KERNO classifier.

Files & POD … self-explanatory

Notes and warnings:

This version of the error reporting is provisional. It compares actual results with typical values for 'good' data where errors in classification are low. The warnings (in grey) are:

This CP3 file comes from a longer CP3 file so some warnings may not apply. Reprocess the cropped CP1 file to restore that information. … CP3 files carry click counts from CP1 files, and use these in some warnings, but the values are not updated when a CP3 file is cropped.

Files processed with the KERNO classifier are compared with some norms to identify unusual patterns and these warnings may be given:

The proportion of Hi quality NBHF trains is very low. Validation of NBHF trains required

Classification rate of high frequency clicks is low. Validation of NBHF trains required

High proportion of NBHF trains resemble WUTS. Validation of NBHF trains required. This does not mean that any WUTS are present, as the characteristics of WUTS overlap those of cetacean trains.

Sonars very frequent. Validation of NBHF trains required

The NBHF:Sonars ratio is low. Validation of NBHF trains required …because false positives from sonars may be significant.

If small dolphins are present the NBHF:"dolphin" ratio is low. Validation of NBHF trains required …because false positives from dolphins may be a significant part of the NBHF category.

If Tursiops or larger cetaceans are present the NBHF:"dolphin" ratio is low. Validation of NBHF trains required … as above.

Modal frequency of NBHF trains is atypical … kHz. Validation of NBHF trains required …the modal frequency of NBHF trains is usually above 125kHz. ADCPs and sonars may generate some false positive NBHF trains with lower modal frequencies.

Proportion of NBHF trains logged during sediment transport noise is very high. Validation of NBHF trains required

Proportion of NBHF trains logged during sediment transport noise is excessive. Validation of NBHF trains required

The proportion of Hi quality "dolphin" trains is very low. Validation of "dolphin" trains required …data with only 'Mod' Q class trains is suspect.

High proportion of "dolphin" trains resemble WUTS. Validation of "dolphin" trains required …as for WUTS above.

Sonars very frequent. Validation of "dolphin" trains required

The "dolphin":Sonars ratio is low. Validation of "dolphin" trains required

The "dolphin":NBHF cetaceans ratio is low. Validation of "dolphin" trains required

Environment

Average daily temperature range = … degC …where a POD has spent time on deck or in a bucket this value is higher!

Average angle of POD to vertical = … deg …an indication of deployment in a high current area, or where a POD was not correctly orientated.

… % of minutes have possible sediment transport noise

… minutes per day have possible boat sonars

… % of minutes have more than 2000 clicks logged …an indication of high noise levels.

Settings … self-explanatory. POD gain values are POD-specific.

SD cards

SD cards formatted for PODs have only two kinds of file on them:

- SETUP.CHE files store POD-specific operating values.
- DATA0.CHE, DATA1.CHE etc. are shell files that can hold data. Each file is 1GB except for the last. The POD writes to fixed positions that fall within these file without reference to their position on the SD card which it assumes.

If any other file is present, or was present, then the card will not work in a C-POD, and any data will be unreadable. Chelonia currently supplies 4GB, class 10, cards made in Japan and marketed by Transcend. These cards have yellow locking tabs.

Some notes on SD cards:

SD cards are complicated! They have a 'wear levelling algorithm' (WLA) to prevent one area of the card from being worn out early, error maps, 4 write modes some of which may crash independently, delayed erasure, etc. Writing any new file to the SD card can trigger the WLA and make the card unreadable by the C-POD.

Windows does not format or read SD cards well (it fails to align sector boundaries efficiently), and many laptop SD card readers do not work correctly with the SDHC cards used in PODs. The USB card readers supplied with PODs should be used instead.

Corrupted cards can be reformatted if they can be seen by Windows. The program to do this operates under Windows XP only, and can be supplied on request. It can re-format any USB drive on a PC or network, and must be used with care, so it is not on the website! Re-formatting does not erase any data, which can then be downloaded normally.

Reading POD data

A directory (folder) with 'CPOD' in its name, or path (the address of the folder, e.g.: C\Users\Nk\CPOD\...) is needed. You cannot save data to a directory without this. This avoids accidentally saving data to the SD card which would instantly corrupt it. Keep all your raw data files from a project in one directory so that they can be analysed in large batches.

To read the SD card follow the steps 1,2,3…

The file name from this upload will be 'tank 2011 10 08 POD892 file01.CP1' Such names can be sorted alphabetically by Windows to give a list of files ordered by location, then date, then POD number, then file type (CP1 or CP3).

The upload results are listed, and can be copied and pasted into a spreadsheet. The columns are:

Event	either creation of a file, or an error.
Log_mins, CP_mins	the number of logging minutes (ON) recorded in the summary data on the SD card by the POD, and the number of minutes in the CP1 file (all = ON and OFF), which is usually larger.
Log_MB, CP_MB	the data volume size recorded in the summary data on the SD card by the POD, and the size of the CP1 file.
Delay	a user controlled delay before the POD starts. This has not been available to users in more recent versions of CPOD.exe as it did not function correctly above 255 minutes.
DutyCyc	the setting for number of minutes OFF between minutes ON recorded in the summary data on the SD card by the POD.
Summary	whether summary data written by the POD was found on the SD card.

The list of files includes all errors. Most of these are not significant and are explained here:

	Event	POD	Log_Mins	Log_MB	Delay	DutyCyc	Summary	CPfile	CP_mins	CP_MB			
line 1	file	448	30d 5h 4m	18	0	no OFF	YES	C:\Users\Nk\Desktop\CPOD d	43504	18.31 MB			
line 2	file	474	47d 20h 41m	13	0	no OFF	YES	C:\Users\Nk\Desktop\CPOD d	68921	13.7 MB			
line 3	file	985	not known	not known	0	no OFF	NO	C:\Users\Nk\Desktop\CPOD d	2	80.38 KB			
line 4	error	460	not known	not known	0	no OFF	NO	Minute counter jumped by	-22	at	28/1/11 00:06	to	0
line 5	error	460	1 TimeSeqErrors from	91				NQ 2011 07 09 POD460 file01	#########				
line 6	error	460	not known	not known	0	no OFF	NO	Minute counter jumped by	27	at	28/1/11 00:07	to	28
line 7	file	460	not known	not known	0	no OFF	NO	C:\Users\Nk\Desktop\CPOD d	226549	16.86 MB			

Line 1,2: normal upload.

Line 3: no summary, but this very brief test file is OK.

Line 4: reports a jump in the sequence of minute numbers logged by the POD, and, in the next line:

Line 5: in the same minute (shown in ####) 1 click has a time preceding a click that has already been logged, and:

Line 6: the minute counter is shown as out of sequence, as it returns to a regular sequence.

… this file may have had 1 minute added by this data error, so the file should be examined. This particular error appears as an isolated minute ON when the POD was otherwise OFF. Where such an error occurs during an ON period it is easy to compare the characteristics of the minutes before and after to see if the minute is real or spurious.

Line 7: details of the file uploaded.

Significant errors:

- Single time sequence errors in minutes may be associated with 1 minute shifts in the time scale, and this may be important if numerous or if detailed comparisons with visual or other acoustic observations are required.

- Large numbers of time sequence errors in multiple minutes. This may indicate extensive data corruption, and has

Skip	0	bytes
Stop after	12	blanks

  been seen in a few files. The 'skip …. Bytes' option can recover these files completely. Contact Chelonia for assistance.

- One large counter jump. Check that the file duration is as expected. The time of the error can be viewed in the file and may show that the upload process has read into old data that exists on the card. This is easily recognised as the temperature usually shows a sudden step, and the character of ambient noise abruptly changes at the same time.

- No data obtained, or data shorter than expected.

- PreStart clicks are not known to be significant, but might be associated with a 1 minute timing error throughout the subsequent file. This can be corrected easily if it can be identified.

NO data

If no data is obtained use the utility on the 'Read SD' page to read Block 2 of DATA0.CHE on the SD card. This should show varied values, as below:

Show summary from DATA0.CHE		clear
500	91	
501	35	
502	140	
503	11	
504	9	
505	126	
506	15	
507	16	
508	6	
509	11	
510	91	
511	133	
Show block n of data from any file:	2	

But if all the values in the second column are 0 or 255, then no data are present.

Block 1 of DATA0.CHE normally shows many zeroes plus some values that hold the summary and logging data. If Block 1 is blank the card has never been in a running C-POD.

Existing data on SD cards can be erased using the 'Erase data on SD card' button on the 'Read SD' page, but it is slow.

Short data

If the file duration is much less than expected, and that shown in the summary (if present) on the SD card, then try increasing the `Stop after 12     blanks` 12 to 10000. If more data is read check that it is not old data (see 'significant errors' above).

If the duration is short and many errors are shown in the upload report the card should not be re-used until it has been be reformatted and tested – see SD cards above.

Archiving SD data

You may wish to keep a copy of the data as stored on the SD card, to allow for any later discovery of any data corruption problem. If you are able to view and assess CP1 and CP3 files, checking that they run for the correct length of time, and show a normal graph of temperature, realistic values for the angle of the POD to vertical then this is not needed. Archiving requires keeping copies of all the DATAn.CHE files (about 1GB each), although they may contain no data. To do this:

- Set up a folder/directory to hold the copies of the SD card, ideally on an external hard drive.
- Save an empty text file called 'POD archive.txt' to this folder.

After you have read an SD card:

- Click the 'Archive SD card' button on the 'Read SD' page.
- You will be asked to 'Select the POD archive.txt file' Select it and clicks 'Save'
- The program will create a folder within your folder with the POD number as the name, plus a day number, so that subsequent back-ups of the same POD's SD card will each have its own folder name, and they will sort by POD number and date.
- Copies of SETUP.CHE and all the DATAn.CHE will be saved to this folder, and an entry made in the POD archive text file, and shown on the Results page.

You can back-up an SD card in this way at any time, but it's best to do it immediately after reading the card as the POD archive.txt file then holds a copy of the start time you entered on the Read SD page. That information is not on the SD card itself.

Exporting data

You can export almost anything you want in a format that can be read into spreadsheets and databases.

Results can go to the Results page of the menu. From there you can copy them and paste them directly into a spreadsheet, or into a text editor. (Get EditPad from www.just-great-software.com – it's a text editor that is vastly better than the microbial NotePad.) From there they can go into a file that can be imported into Access, or you can save the results directly into such a file.

For large data volumes, export to a *text file* is very much faster than to Results. The text file can be opened directly by spreadsheets or imported into databases.

If you want to export all appropriate files in a directory and all of its subdirectories you can opt to do this by saying yes to "Do you wish to select all files in a directory and its sub-directories?" You are given this option when clicking any of the export buttons i.e. Export from batch of files, Train details, Full train details, Click details, Encounter details, Batch detection and noise summaries, List files, List settings and List Landmark Sequences.

Useful functions, that apply to most forms of export:

'*use time selection*' – a time selection can be set on the 'view +' page, or otherwise, and will then limit the scope of export from a batch of files.

'*omit headers*' is useful when a batch of files are processed and you wish to move the data into a spreadsheet or database.

'*include annotations*' – annotations are species identifications added to minutes by the user.

'*export SPL as Pa*' gives the value of the clicks maximum sound pressure level in Pascals. This true SPL is not generally useful for visual displays, because it conceals where the click reaches the ceiling of the SPL scale, but it is more useful when exported.

Various export functions include '*Nall*' – the total (unfiltered) number of clicks in the CP1 file in the period in question – in the export as a useful index of noise levels.

Saving / recalling export settings via *.cpe files

There are enough options to give a real chance of error when repeating export operations, so you can save the export settings you currently have to a file which will have the extension *.cpe* and load these settings on subsequent occasions. The buttons are low down on the Export page. The settings file contents can be viewed with a text editor.

Times

Times are exported as:

1. dd/mm/yyyy hh:mm This format can be selected in Excel: Format > cells > Custom > 'dd/mm/yyyy hh:mm'. You may need to format the cells in the column before pasting in the data.

    1. minute number. This is the number of minutes since the start of 1900. Excel and Access store dates as the number of days since the same starting point, i.e. minute n / 1440.

Within minutes, times are expressed as either:

- microseconds. The time resolution of C-PODs is 200 nanoseconds (0.2µs), but the smallest unit expressed in CP1 files is 5 µs. Maximum value 12,000,000.

- milliseconds. Maximum value 60,000.

To calculate a precise Excel-type date/time value: = (minute + (microseconds / 12000000)) / 1440

There is some temperature dependent drift of the CPOD clock. The difference between PODs is least where PODs are subjected to the same constant low temperatures.

DPM, DP10M, DPH

Detection Positive Minutes, Detection Positive 10 Minutes, Detection Positive Hours.

These count periods of 1, 10, 60 minutes which have trains meeting whatever filter criteria are set. The periods are aligned with the clock So DP10M will go from 3.20-3.40-3.50 etc. rather than in 10 minute periods counted from the start of the file.

DPM is a standard measure of how much time animals are present, which can be expected to vary with the density of groups of animals.

Waiting time

Mean waiting times are not directly exportable from CPOD.exe because it can be a badly behaved statistic. e.g.
- sites with strong tidal patterns of presence can show long waiting times despite very high mean densities.
- if animals move faster waiting times (and encounter durations) fall, although habitat use duration and density did not change.
- if the sensitivity of the logger falls long encounters can be split into several shorter ones and waiting time can fall which is going the wrong way. This can be demonstrated in some data by excluding Mod Q trains.
- it becomes non-linear with density above lower density values than DPM.

Mean waiting times came from a power analysis of some early Danish data with few detections, and was the most powerful statistic, but that has not proved generally true, and DPM is linear over a wider range.

Specific waiting time after some impact is a useful measure.

Time present / train duration total

The sum of the in-area time for each animal that visits an area taken over some observational period is exactly equivalent to the average density (of animals) in that area over the same period of time. In POD data the detected train is time in which an animal is known to be present. It is only a small fraction of the total time that any animal is within detection range, but it can be expected to have a fixed relationship to the total time. As density rises the detected trains become progressively more likely to coincide so the relationship will gradually become deviate from a linear one as this metric 'sum of train durations' starts to saturate.

To use this relationship in relating C-POD data to density, 'Time present' can be exported directly, and is the sum of the duration of trains logged, in milliseconds. This will be far less than the actual time present because of the directionality of the sonar beam, and the area monitored is not exactly known. Determination of the scaling factors that relate density to time present and to the species, depth, behaviour, etc. are a major research priority for static acoustic monitoring.

Time present export gives:

File	ChunkEnd	Millisecs	Nfiltered	Nall	MinsOn
CPOD Anclaje 13 2010 11 06 POD995.CP3	22/11/10 23:59	0	0	0	0

CPOD Anclaje 13 2010 11 06 POD995.CP3	23/11/10 23:59	0	0	14705	57
CPOD Anclaje 13 2010 11 06 POD995.CP3	24/11/10 23:59	0	0	102800	488
CPOD Anclaje 13 2010 11 06 POD995.CP3	25/11/10 23:59	20455	1357	127754	1440
CPOD Anclaje 13 2010 11 06 POD995.CP3	26/11/10 23:59	140	34	248314	1440
CPOD Anclaje 13 2010 11 06 POD995.CP3	27/11/10 23:59	1921	113	255218	1440
CPOD Anclaje 13 2010 11 06 POD995.CP3	28/11/10 23:59	287	10	346943	1440
CPOD Anclaje 13 2010 11 06 POD995.CP3	29/11/10 23:59	32531	2096	485173	1440
CPOD Anclaje 13 2010 11 06 POD995.CP3	30/11/10 23:59	0	0	229744	1440

'Nall' is the total (unfiltered) number of clicks in the CP1 file for each day (as the time unit here is days).

N of encounters

This is no longer exportable directly. As a statistic it was originally proposed as a way of counting porpoises, but it does not work in this way as an encounter with many trains could be due to one animal moving slowly over the ground or to many moving faster. The concept of *time present* (see above) provides a much better basis for viewing the relationship between detections and porpoise density.

N of clicks

Note: if the SPL is filtered, this will use the raw value, not the corrected value in Pascals.

Inter-click intervals

For studies of behaviour. Check the number of clicks via the 'Analysis' page before exporting. If the number is large, export to a text file to save time. It may be useful to set these two filters on the 'view +' page:

More filters:
ICI assessment quality: ☑ high only
Species classification quality: ☑ high only
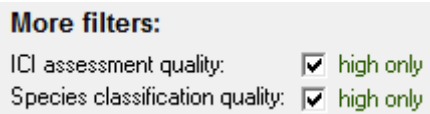
Noise levels

There are several sources of potentially relevant measurements:

1. 'Nall' – the number of clicks in periods that have low train detection rates for Hi or Mod trains.

2. Identification of sediment transport noise.

3. MMM – the mean multipath minima exported with train details.

The first two are exported with 'Detections and Environment'.

The third is a novel approach: If noise levels are high enough to push the detection threshold up, the average SPL of the weakest clicks in the multipath clusters in trains should be louder. The multipath minima have been used as there are relatively few direct-path clicks in any train that are close to the detection threshold. Including the echo replicates in the multipath clusters increases the useful data volume. The 'Noise levels' export option gives values for the number of all clicks, the number of multipath minima and their mean value (' Nclx, nMMM, MMM'). This approach needs investigation and validation.

Detection and environment

A generally useful export format. For example:

File	ChunkEnd	Minute	Temp	Angle	MinutesON	DPM	Nfiltered/m	Nall/m	%TimeLost	SonarRisk	kHz_continuous_noise
Anclaje 13	6/11/10 10:59	58303379	0	0	0	0					
Anclaje 13	6/11/10 11:59	58303439	0	0	0	0					
Anclaje 13	29/11/10 03:59	58336079	19.8	0	60	0	0	1928	4	0	0
Anclaje 13	29/11/10 04:59	58336139	20	0	60	0	0	1251	7	0	0
Anclaje 13	29/11/10 05:59	58336199	19.8	0	60	0	0	341	0	0	0
Anclaje 13	29/11/10 06:59	58336259	20	0	60	7	12.08	125	0	0	0

Nfiltered/min is the number of click in trains passing by your filters.

'Nall' is the total (unfiltered) number of clicks in the CP1 file for the same time period.

TimeLost is the % of time in periods when the POD had 'maxed out' – reached the logging limit for a minute.

Sonar Risk is 0 or 1 if some risk was identified.

kHz_continuous_noise give the frequency of any such noise identified.

Here are hourly data from a highly tidal site:
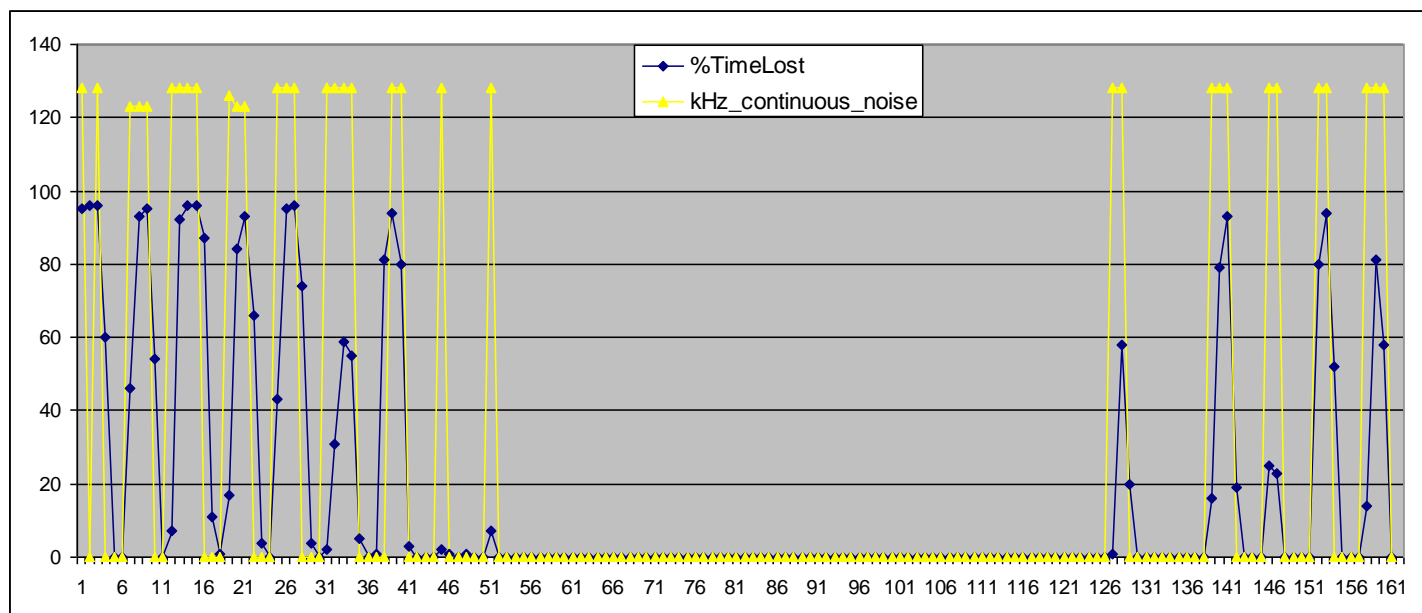
The sediment at this site generates noise at the mid to low end of the NBHF frequency range, and is so intense that many hours are maxing out – reaching the 4096 clicks limit in many minutes except during the neap phase of the tidal cycle.

Train details

These data are exported to a text file that can be read into a database, spreadsheet or text editor. All filters, and any time selection, are used. Fields are separated by tabs.

File	TrnID	Time	Min	SpClass	TrClass	Start	TrDur_us	NofClx	Clx/s	modalKHz	avSPL	MaxICI_us	MinICI_us	MMM	TimeLost	ClksThisMin
	108	14/9/10 17:52	58227472	NBHF	High	27751750	768450	17	21	137	112	51565	38890	0	0	479
	109	14/9/10 17:52	58227472	NBHF	High	29169230	1147635	27	23	137	87	89200	38950	16	0	479
	64	15/9/10 04:01	58228081	NBHF	High	26582670	1596215	31	19	128	76	103675	46375	12	0	1231
	83	15/9/10 04:05	58228085	NBHF	High	36509015	1225250	24	19	125	123	101025	48720	0	0	1025
	98	15/9/10 04:06	58228086	NBHF	High	30830515	1219105	46	37	121	50	52020	23295	15	0	777
	106	15/9/10 06:10	58228210	NBHF	High	37567585	1775050	40	22	129	122	77595	42180	14	0	621
	63	15/9/10 18:53	58228973	NBHF	High	15817845	88655	19	203	126	48	9630	4440	0	0	372
	20	16/9/10 19:45	58230465	NBHF	High	50459080	62645	15	223	127	25	6600	3230	0	0	38

The fields are:

TrnID   An identifier in the range 5..127 for the train. The only use of this is to match this summary to details of each click.

Time   Start time as dd/mm/yyyy hh:mm See Times above

Min   See Times above

SpClass   'NBHF' etc.

TrClass   'Qclass' Hi, Mod, Low, ?

Start   Start time within minute, in microseconds. See Times above

TrDur_us   Total duration of train in microseconds. Can be used to calculate overall average ICI = duration / (NofClx – 1)

NofClx   This includes 'virtual clicks' that are inferred by the train detector where the real clicks may have been 'knocked out' by a noise spike. Including these makes the calculation above valid.

Clx/s   Average click rate, rounded to whole number of clicks per second

modalKHz   modal frequency of train clicks in kHz.

avSPL   Average SPL of clicks in train in raw units, not corrected for frequency.

MaxICI_us   The first and last ICIs (inter-click-intervals) are excluded as these are more often incorrect.

MinICI_us	As above
MMM	Mean-multipath-minima = the mean SPL of the weakest clicks in all the multipath clusters in the train.
TimeLost	% of logging time lost due to minutes maxing out.
ClksThisMin	Total number of clicks in the CP1 file for the minute

Full train details adds these fields:

LastICI_us	
ICI_rising	% of ICIs that are longer than the previous ICI. This figure is high where trains come from WUTS.
avEndF	Average end frequency of clicks.
MinF	Lowest click frequency in train.
MaxF	Highest click frequency in train.
avNcyc	Average number of cycles in all clicks in the train.
MaxNcyc	Highest number of cycles in any click in the train.
MaxSPL	Highest raw SPL value of any click in the train.
NofClstrs	Number of train clicks that are part of a multipath cluster.
avClstrN*10	Average number of multipath replicates in clusters * 10.
Frange	Max click kHz – Min click kHz.
avSlope1, 2 avBW	Average of these parameters. These are difficult to use as many values are unrepresentative due to the analysis time window used by the POD not being optimally positioned.

Click details

To get details of the clicks on a screen use the right-click 'put raw data into clipboard' option, and then paste them into a spreadsheet.

Few users will wish to delve this deeply, but you can export click details in bulk. They are only exported to a text file.

Check the number on the 'Analysis' page first or you may create a file that is four times the size of the source file.

time	microsec	cycles	SPL_Pa	kHz	Bandwidth	end kHz	Qn	TrN
2/1/10 03:48	25197620	10	22	132	1	142	2	5
2/1/10 03:48	25227195	10	21	129	1	116	2	5

Some common tasks

For porpoises or dolphins, use Q classes Hi + Mod and NBHF or Other cetacean for these tasks.

Larger cetaceans need a more sophisticated assessment of appropriate filtering.

Required:	from	period	Omit negative time periods	to
A neat list of files, dates, and click rates	List: Files	-	-	Results
Were my settings the same?	List: Settings	-	-	Results
List of minutes with detections	A data: N of clicks	1min	yes	Results
Measure of habitat use	A data: DPM	days	No	Results
Porpoise positive days	"	"	"	Results
Comparison of PODs set 1m apart	A data: Train duration	days	No	Results
Train structure	A data: Inter-click intervals	-		text
Detections and noise etc. for detailed analysis	A data: Detections and environment	hours	No	Results
Diel or tidal patterns of detection	B: train details	-	-	Results

Are there distinct types of encounter?	C: encounter details	-	-	Results
Which of DPM, DP10M, DPH should I use?	D: Batch detection…	-	-	Results

Encounters

Encounters are not useful as measures of cetacean density, but are a very useful tool for the discrimination of species and behaviours. The 'export encounter details' function lets you examine the data more deeply or develop your own species classifiers. It produces some very interesting results and is worth playing with!

C   Encounter details   Gap 10   minutes   ☐ graph
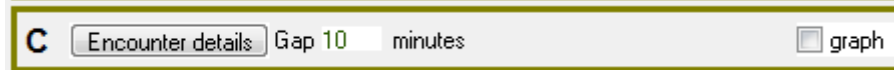
The 'Encounter details' button requires a file pair (CP3 and CP1) to be open. It runs a process which:

1. Finds encounters separated by 10 (or other) minutes without any train that passes the filters set.

    1. Collects various measures of the encounter from the train in the CP3 file.

    2. Draws graphs of frequency distribution if the graph box is checked.
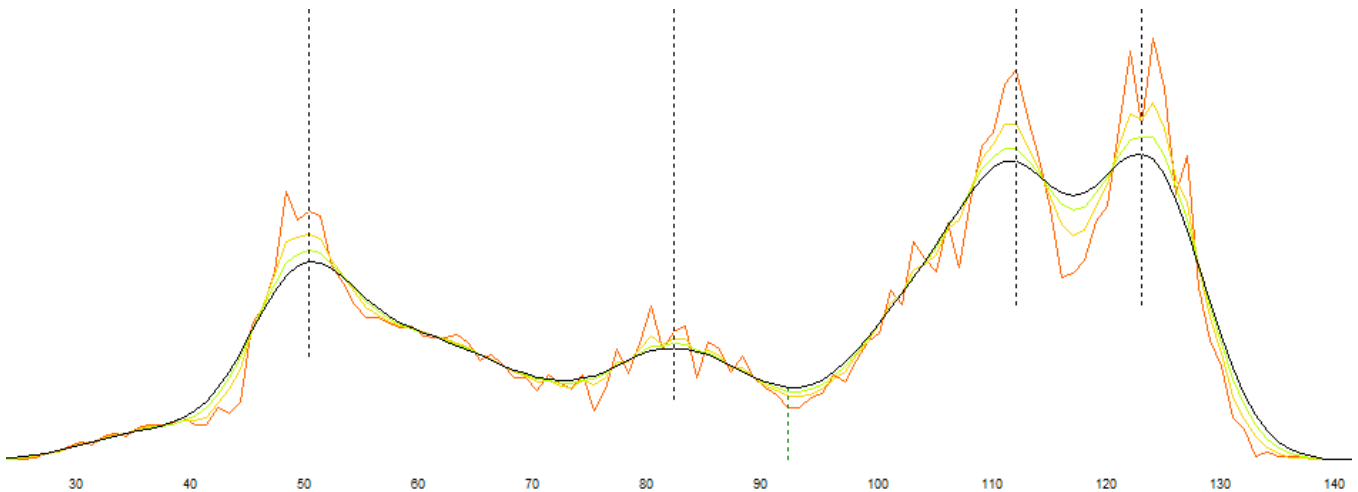
    3. The graph shows amplitudes by kHz (the X axis). Black line = smoothed data, Coloured lines show successive smoothing stages of raw data, dashed black = maxima selected, dashed green = minima selected.

    4. Uses the start and end minutes defined in the CP3 file (step 1) and collects measures from the same time period of the CP1 file.

    5. Places these in two boxes (not previously visible) on the 'Trains' page:

Field names ending in .T are values from trains:

File	
EncN	
Start	Text date of start
StartM	Minute number of start
Minutes	Duration of encounter
NClx.T	Number of clicks in trains in encounter
TPsum	Time present (sum of train durations)
avNcyc.T	
avSPL.T	
kHz12.5ile.T	12.5th centile of click kHz distribution
kHzMedian.T	etc
kHz87.5ile.T	
kHzMode.T	
%up.T	% of clicks with an end frequency higher than the average frequency
avUp.T	Average rise in frequency within clicks
%down.T	etc
avDown.T	
Clks/s_25ile	25th centile of click rate
Clks/s_median	etc
Clks/s_75ile	
Clks/s_mode	
PkHt	Height of first peak on frequency graph
PkKHZ	kHz of first peak
PkHt	etc
PkKHZ	
PkHt	
PkKHZ	
PkHt	
PkKHZ	
KhzMin	First minimum on frequency graph
KhzMin	etc
KhzMin	
KhzMin	

The corresponding values from the CP1 file have no data on click rates as no trains are defined in CP1 files.

This may look rather daunting, but try this:

- Paste the CP3 data into a spreadsheet, and paste the CP1 data to the right of it, so that they can be sorted together.
- Graph kHzMode.T, kHzMedian.T, Clks/s_25ile, Clks/s_mode, Clks/s_75ile
- Graph the kHz values from the CP1 file
- Sort on kHzMode.T
- Sort on Clks/s median

The sorted graphs may show distinctly different subsets of encounters. Are they different species, or different behaviours? Gets interesting here!

Handling exported data

Times as the number of the minute since the start of 1900 can be handled easily, and can be simpler than trying to second-guess Excel's determination to automatically format date/time values.

Hour of Day

For example, in Excel: Hour of Day = INT((Minute – INT(Minute /1440)*1440)/60)

ChunkEnd	Minute	HourOfDay
24/11/10 21:59	58329959	21
24/11/10 22:59	58330019	22
24/11/10 23:59	58330079	23
25/11/10 00:59	58330139	0
25/11/10 01:59	58330199	1
25/11/10 02:59	58330259	2

If 60 in the formula is reduced to 10, the formula gives numbered 10 minute time periods.

Tidal Phase

For example, in Excel: Tidal phase = ((C2+$G$1)-INT((C2+$G$1)/744)*744)/744 … with 'minute' in column C, and a fixed single value for the 'tidal offset' in minutes in cell G1.

Tidal phase is given in the range 0-1. This simple calculation does not take the variation within the lunar cycle into account. The tidal offset should be set to give the correct tidal phase at some point in the data, preferably near the middle as this formula is an approximation and will go out of phase over time.

The calculated time of the cycle should be checked against proper tide tables to ensure that the level of errors is acceptable.

Importing train data into Access

This is easy and does not require you to set up a table first.

Open a blank database and go to:

- File>Get External Data>Import> set 'files of type' to text files .txt > Import
- Then select:
- Delimited
- ... next...
- Tab
- First Row contains filed names
- ...next...
- Finish (this skips several pages)

A template database for CPOD data will be added to the Chelonia website, and will contain queries for a range of useful functions.

Graphs

The 'graph' button shows a graph of click parameters within the period displayed on screen, which can be the whole file.

Mode = 144 (kHz weighted by corrected SPL)

Mode = 144 (kHz weighted by corrected SPL)
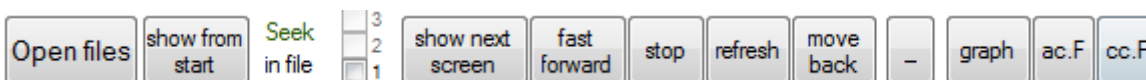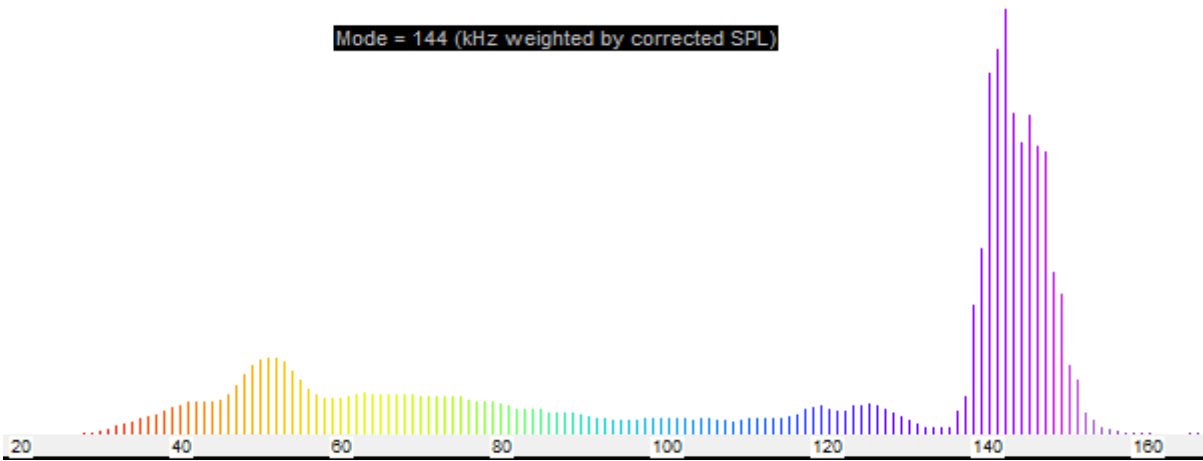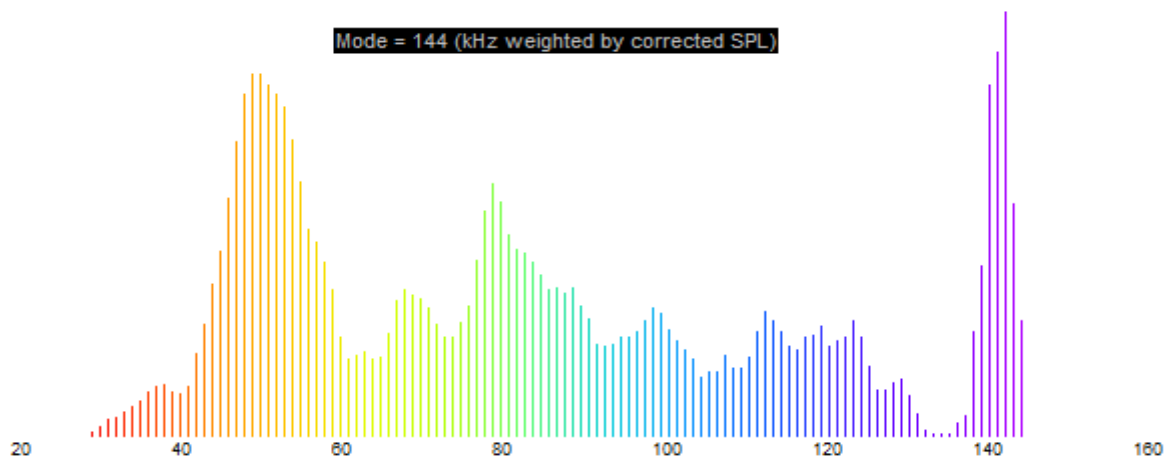
The Y axis is adjusted to the highest value, so there may appear to be more clicks in the CP3 file than in the CP1.

All click parameters and the click rate in trains can be graphed in this way. This is a graph of clicks/s in NBHF trains. The mode is at 23clicks/s.

To select which parameter to graph (kHz is the default) go the view + page.

Mode = 23 (Clicks/s)
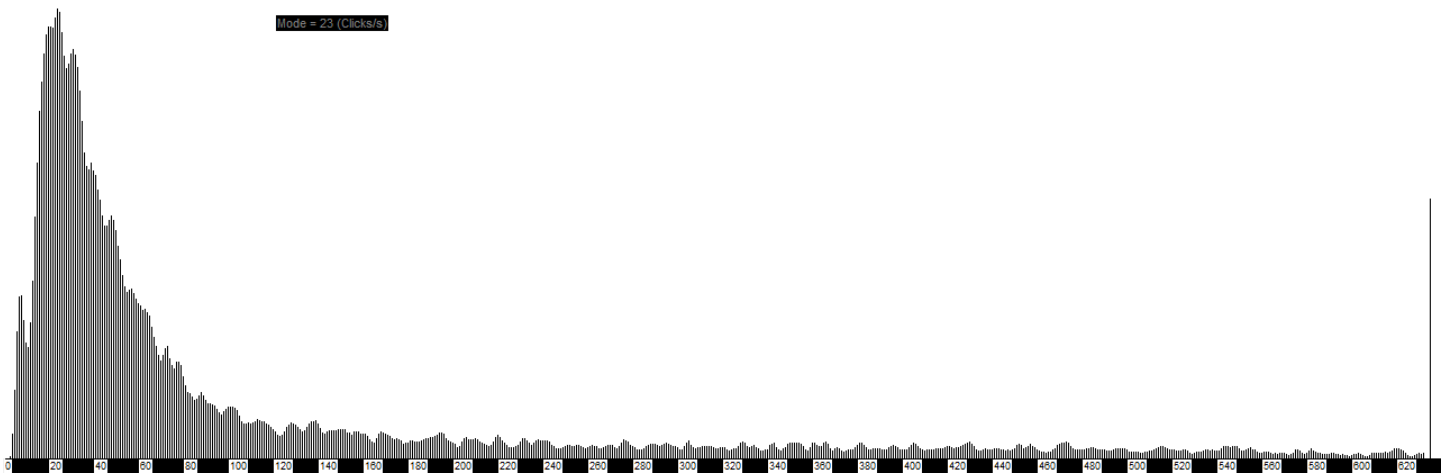
While the graph below comes from deep water, and has a modal click rate of 6/second:

Mode = 6 (Clicks/s)
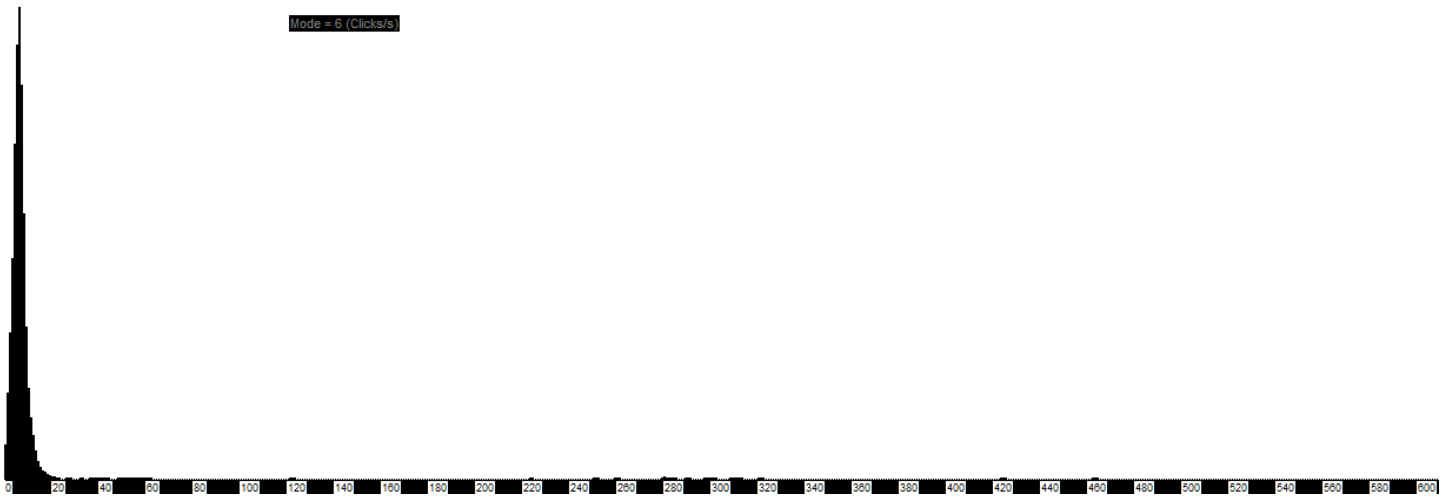
Diel graphs

Graphs can be plotted to show mean 24 hour patterns for various click parameters if "Show mean values for hour of day" is checked:

If 'Show mean values for each hour' is not checked then the data will be put in 10minute bins, if there are sufficient data points.

If a file set is open the upper graph will show the CP3 train clicks as filtered.

In the case shown below the number of clicks is graphed and there is a very strong diel pattern that is exactly parallel to the overall number of clicks in the CP1 file.

The data come from a deep-water C-POD deployed in the Ligurian Sea and show activity of deep-diving cetaceans with very few clicks from any other source being logged:

Max value 21192

N = 1,145,540 — Hour of day — M6 Buoy 2009 2009 08 28 POD439 file001 PART 208d 15h 20m.CP3

Max value 155178

N = 9,552,748 — Hour of day — M6 Buoy 2009 2009 08 28 POD439 file001 PART 208d 15h 20m.CP1
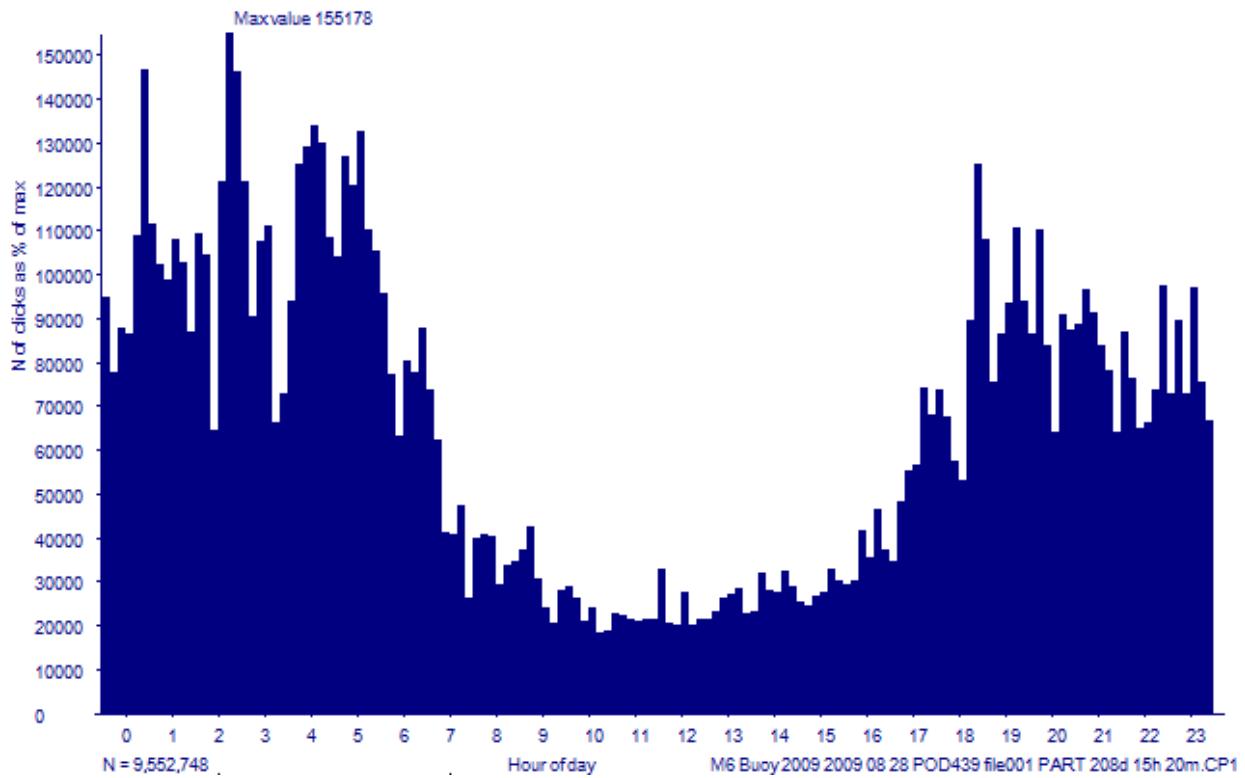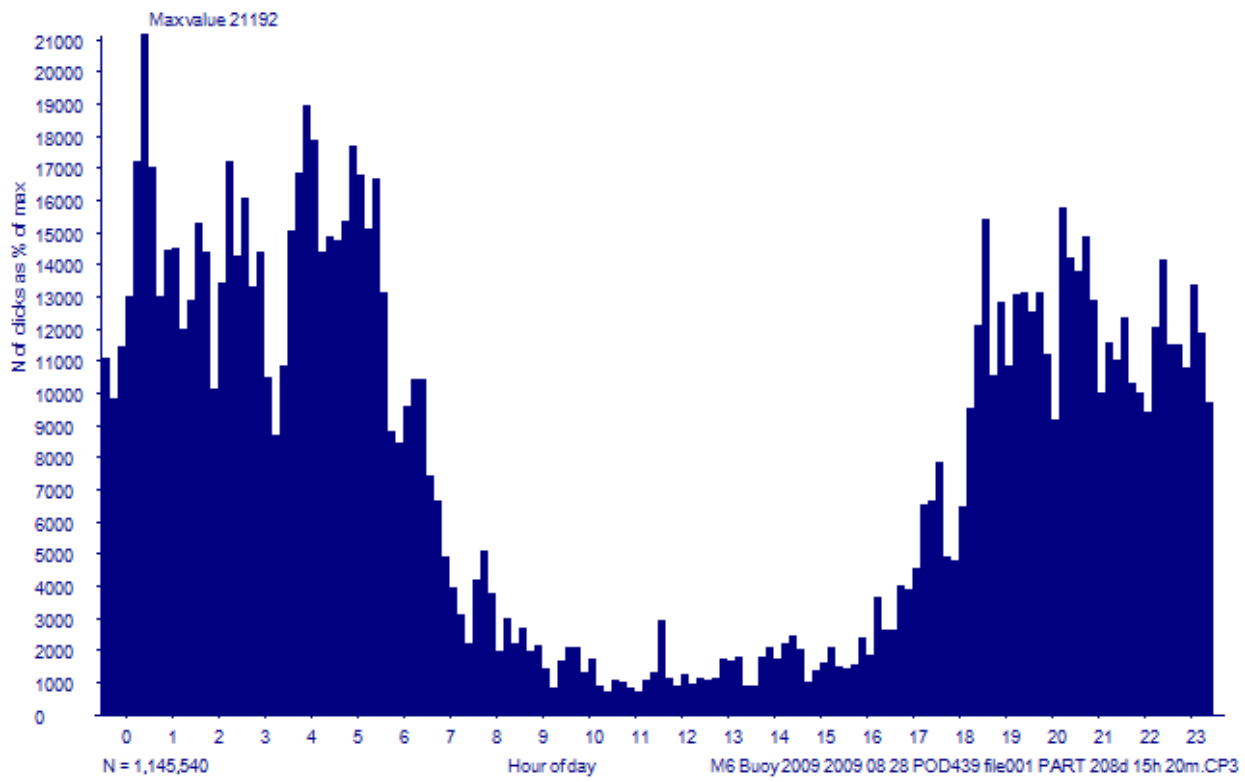
A diel graph of frequencies shows a very interesting shift in parallel with the diel pattern:

32

N = 1,145,540   Hour of day   M6 Buoy 2009 2009 08 28 POD439 file001 PART 208d 15h 20m.CP3

N = 9,552,748   Hour of day   M6 Buoy 2009 2009 08 28 POD439 file001 PART 208d 15h 20m.CP1
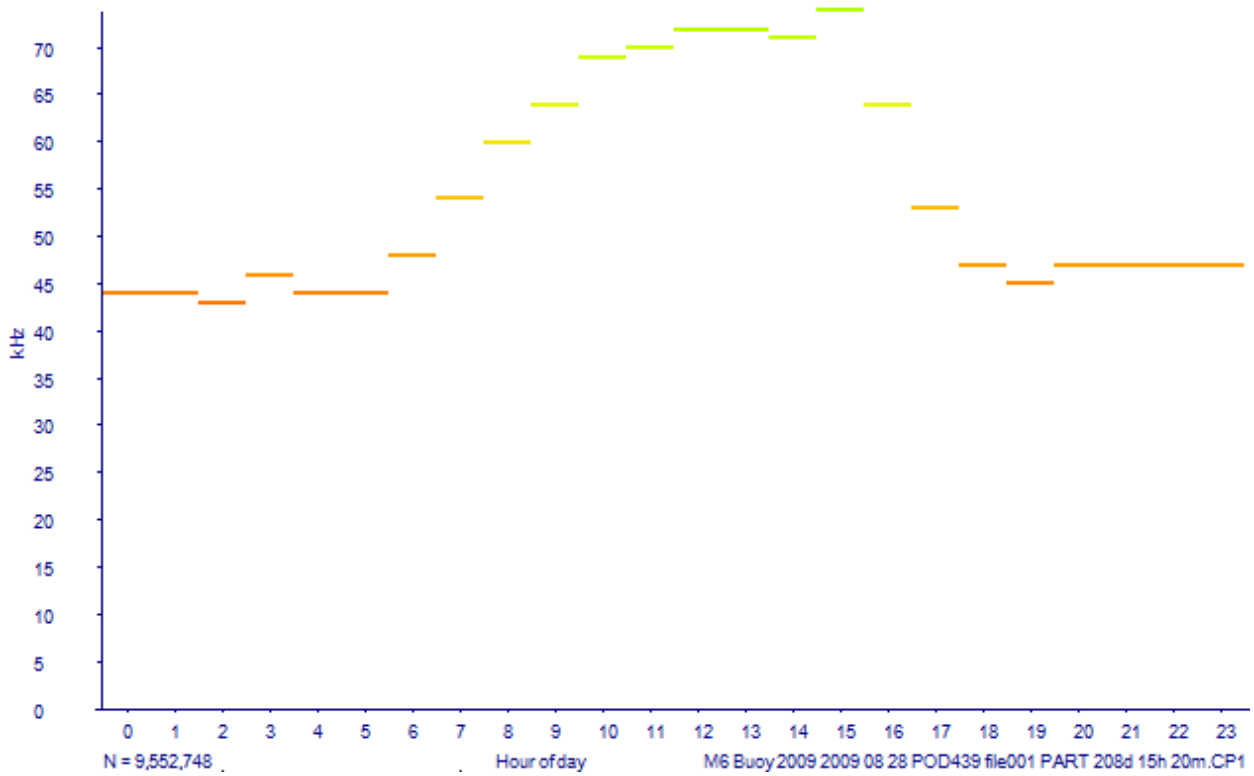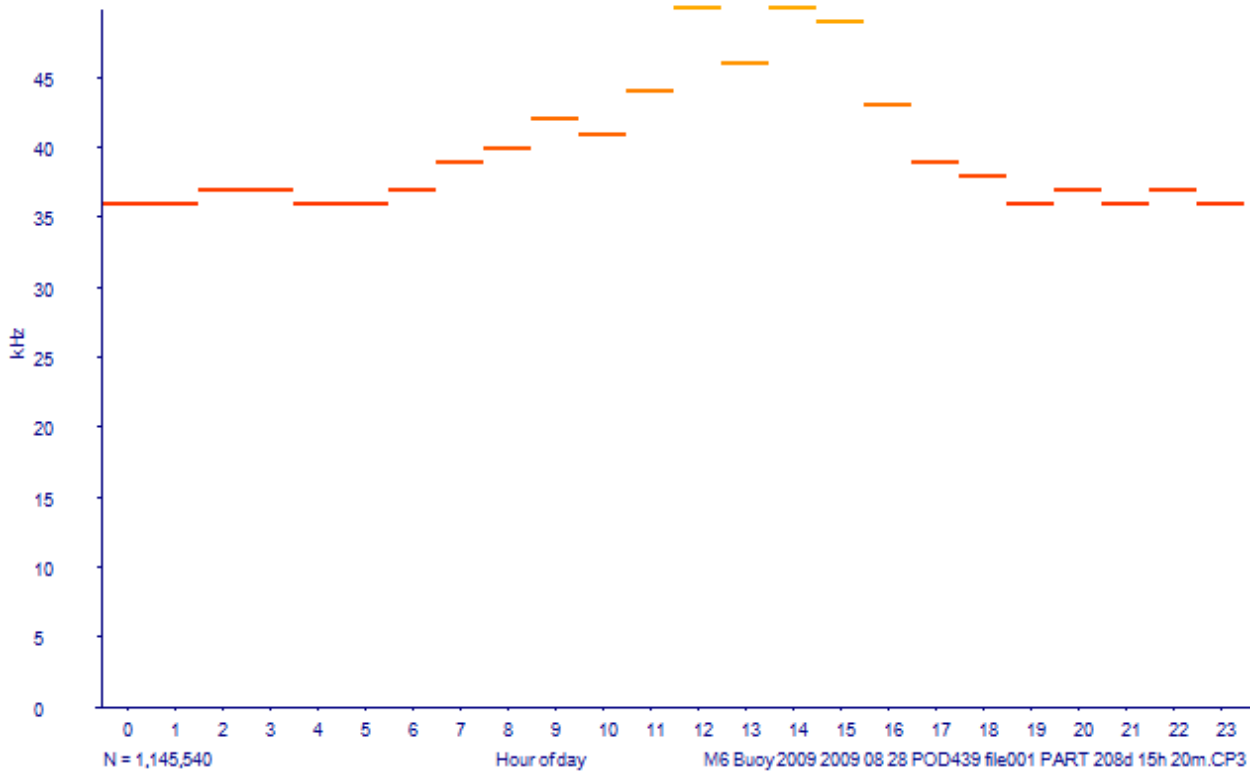
Such diel patterns have been ascribed to 'chorusing' of some part of the biota, and are often seen. To test whether such a pattern might be due to boats you can graph 'sonars' to see if they correspond, and you can also view the diel pattern of the frequency of the clicks as above.

33

Quick analysis

The 'Analysis' page of the menu gives a very quick overview of data files. Try varying the species or other filters and see the effects. In this example looking at the output of the GENENC encounter classifier to sonars are visible at around 50kHz, 100kHz and porpoises at 130kHz.

A1b 914.CP3

Train filters: GENENC: NBHF NBHF? Sonar Sonar?   No click filters
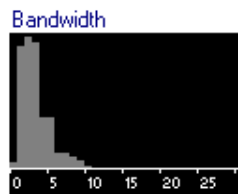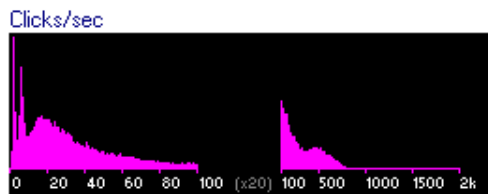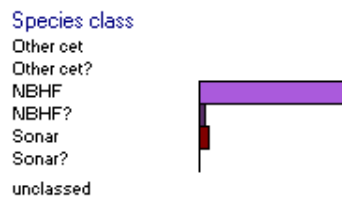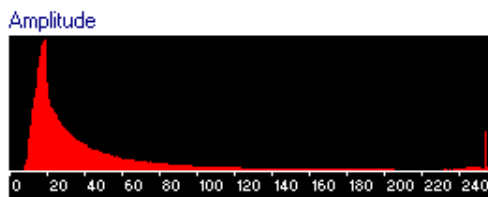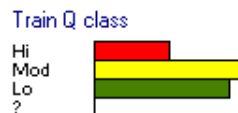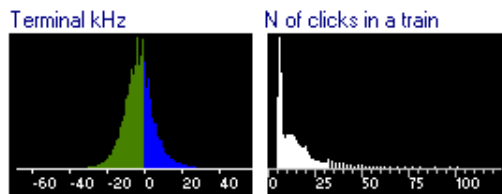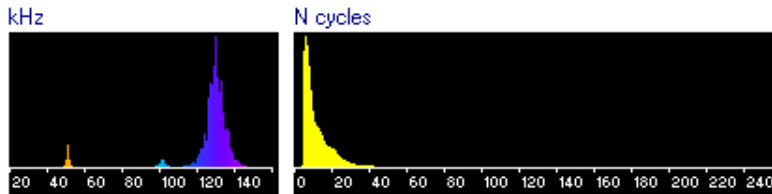
kHz   N cycles

file
3
2
1
Analyze

step through species classes
put class totals in Results
ignore all filters

... go to filters page ...

Terminal kHz   N of clicks in a train   Train Q class

Hi
Mod
Lo
?

454,292 clicks found

5,096,508 clicks in CP1 file
3.03 clicks per min
182 clicks per hour
4,368 clicks per day
DPM per day 86.43 (8989)

Amplitude

Species class

Other cet
Other cet?
NBHF
NBHF?
Sonar
Sonar?

unclassed

File duration analysed: 103d 23h 59m
Acoustic duration: 103d 23h 58m (99%)
0% truncated in CP1 (411 @ 4096)

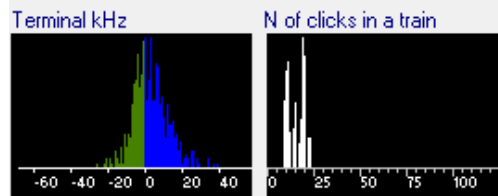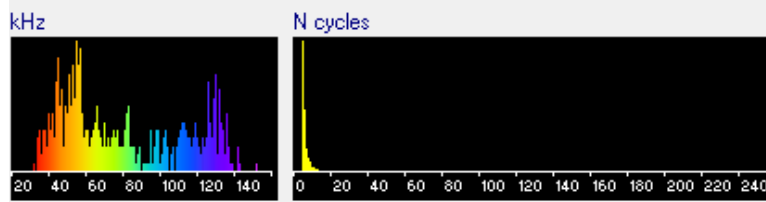Clicks/sec   Bandwidth

Mean temp 4C
Mean angle 8

TIC value: 4.9 N= 45787

F12 restores data display

Looking at a CP3 and *stepping through the (KERNO) species class* (check box beside 'Analyze' button) we see:

NQ3 0 00 00 POD414 file001.CP3

Train filters: All Q classes. (good ICIs only) Species: 'Dolphins'.

No click filters

kHz

N cycles

Terminal kHz

N of clicks in a train

Train Q class
Hi
Mod
Lo
?

Amplitude

Species class
NBHF
'Dolphin'
Cet group
unclassed
Sonar

Clicks/sec

Bandwidth
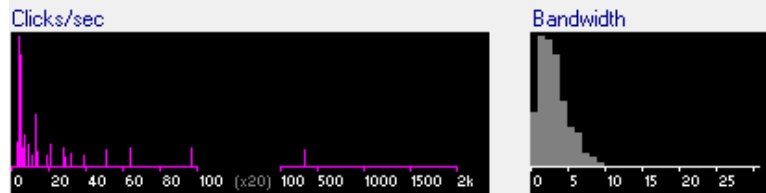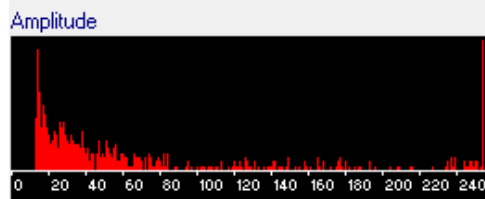
file
3
2
1

Analyze

... go to filters page ...

☑ step through species classes
☐ put class totals in Results
☐ ignore all filters

529 clicks found

1,079,610 clicks in CP1 file
0.01 clicks per min
0.71 clicks per hour
16.93 clicks per day
DPM per day 1.00 (35)

File duration analysed: 34d 10h 8m
Acoustic duration: 31d 6h 6m (90%)
0% truncated (0 @ 65536)

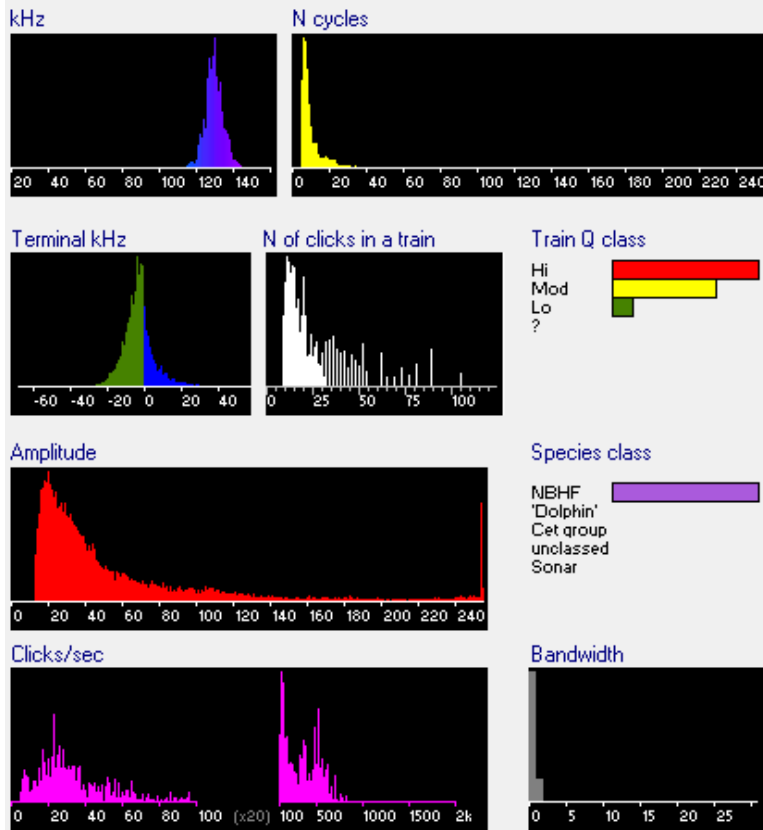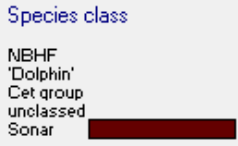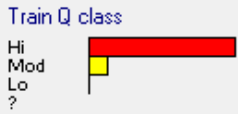Mean temp 13C
Mean angle 8

TIC value: 4.6 N= 39

F12 restores data display

In this file 'Other cetaceans' shows mainly short, broader band, clicks with a wide bimodal frequency spread, and mostly very low click rates. These would fit bottle-nose dolphins that are found at this site, but there may be some mis-identification of porpoises as bottlenoses in view of the frequency peak around 130kHz.

In this file 'NBHF' trains show:

NQ3 0 00 00 POD414 file001.CP3

Train filters: All Q classes. (good ICIs only) Species: NBHF.
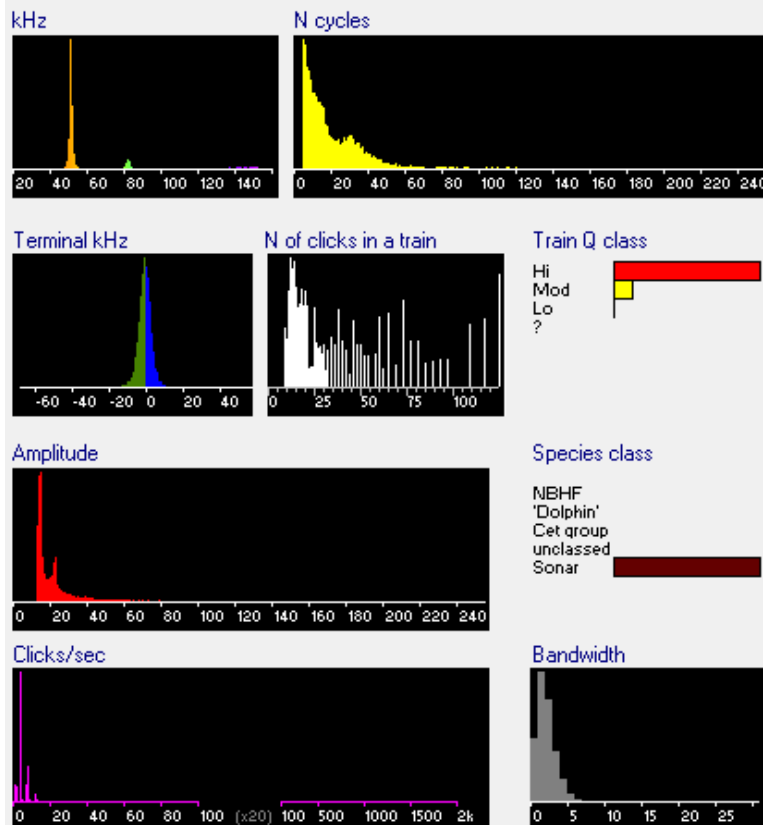
As expected: around 130kHz, very narrowband, longer, with faster trains and consequently more clicks per train. Some feeding buzzes are likely in view of the number of fast trains around 500/s.

In this file 'Sonars' show:

NQ3 0 00 00 POD414 file001.CP3

Train filters: All Q classes. (good ICIs only) Species: sonar.

As expected: very long, very slow, very narrow frequency bands.

In this file 'unclassed' shows a mixture of all of these so it will include false negatives.

Conclusion: looks like good data from an NBHF + dolphins + boats site.

Array viewer

If POD positions can be approximated to a rectangular grid, with or without gaps, then their contemporary detection data can be summarised into an array file, and viewed at speed in the array viewer. The array viewer controls are on the 'view+' page.

In the array view, grey squares show missing or dead loggers. Where a surface and bottom logger are present, the data is shown over dark-blue or black bars.

The bars in each square represent successive 10 second slices of the minute. Frequency in kHz is shown as the standard colour code, and click rate by the bar height.
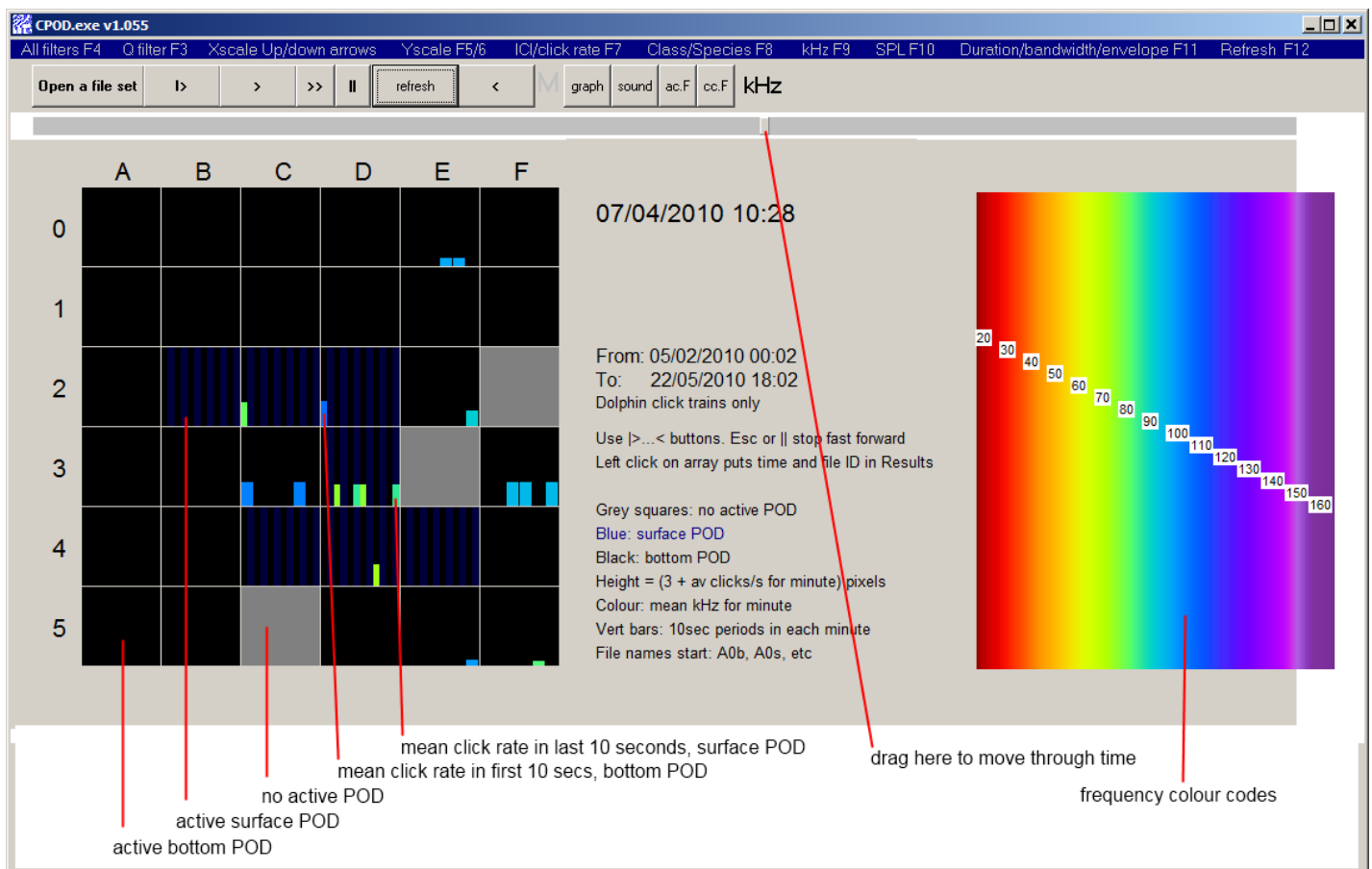
To create an array file:

Name all the files you wish to use by their array position. Only the CP3 files will be used, so you could copy all of those in Windows Explorer and rename the copies so that they start with the column letter, followed by the row number, followed by s for surface or b for bottom.

The array file can hold two 'species', and two positions (surface and bottom), and 6 time slices, for each minute. To make the array file:
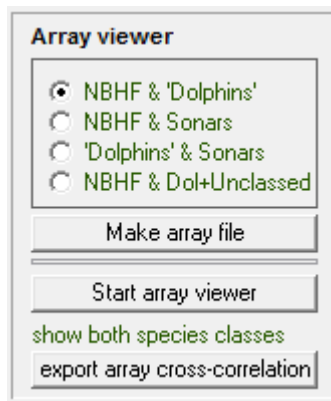
- select species pair
- click 'Make array file'
- name the array file e.g. xxxxx.arr'
- select all the files from the array
- wait until the process finishes.

To view an array file:

Click 'Start array view' and select your array file.

Click 'show both species classes' to loop through species display selections.

No filters apply within the array viewer.

Array file display

See the graphic above. For each 10 seconds, if one or more click trains are present, the display shows a bar whose height represents the mean click rate /s within the trains (1/mean inter-click-interval), and the colour represents the mean frequency for all the clicks in trains for that species group for the whole minute.

Navigation

Minutes with no clicks are automatically skipped, and the number of minutes skipped is shown below the large label giving the current minute.

The buttons *Show from start, show next screen, fast forward, stop, refresh screen,* and *go back* work on the array file, and so does dragging the time slider.

You can put a time, in the format 22/10/2009 14:22, into the 'move to screen before' box on the 'view+' page and click 'go' Keep a list of interesting date.times and paste them into the box as required.

Stopping

Click 'Cancel array view' on the 'view+' page. If you open a file (or up to 3), select a display unit of less than 1minute and click 'refresh' you will see the minute before the one you were viewing as the time position is remembered.

The 'export array cross-correlation' button automates cross-correlation for a set of files using a finer time resolution than the c.c. described below.

Auto-correlation

The ac.f button shows a long autocorrelation on one of the open files.

The file is selected in advance by double-clicking in its display area.

The time resolution must be 1min or higher and is used to define the bin size.

The auto-correlation is the correlation between a time series and itself. Initially the correlation has to be 1.0 – a perfect correlation. Then the comparison is with itself offset by 1 minute and the correlation falls.

Successive values are obtained by making bigger time offsets. If a species typically makes brief visits to the logger site then the ac.f will fall below significance when the offset is bigger than the duration of most visits, so this is a neat way of characterising encounter durations.

Auto-correlation function values that are more than *2/SqRt(Number of values)* from zero represent a *p* value of <0.05. Significant autocorrelations are common, but may be coincidental, so an intelligent view is essential. A file with a big bunch of trains one day, and another 30 days later, and not much else will show a very significant correlation at 30 days,

but it was based on just one pair of encounters and really tells you nothing new. When many encounters are involved it becomes more useful, as shown in the example below.

Here is the autocorrelation function for NBHF clicks at the same site as the analysis above:
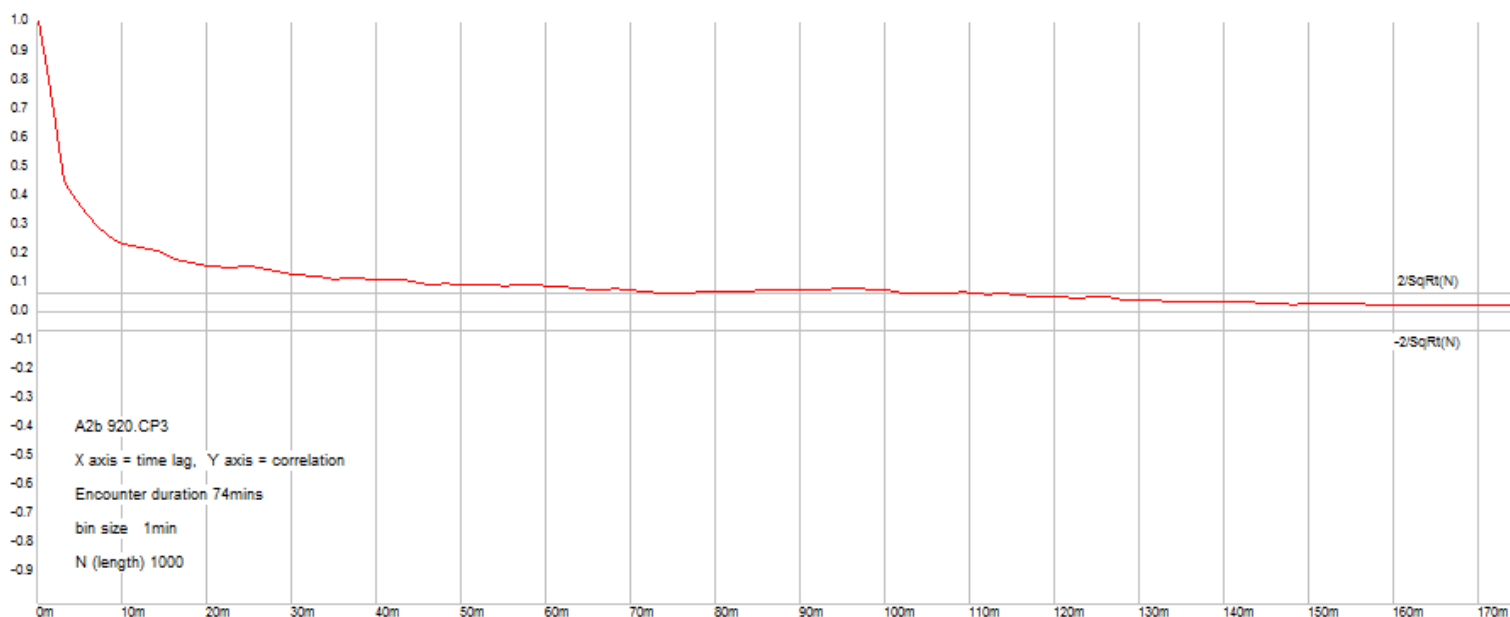
This shows that porpoises are still likely to be around 2 hours after they first appear. It could be that they typically stay for that long, or that a shoal of fish stays that long and porpoises make multiple visits to it over that time, etc.

Dolphin clicks show a much shorter encounter duration of about 5 minutes, so they seem to be doing something different at this place:
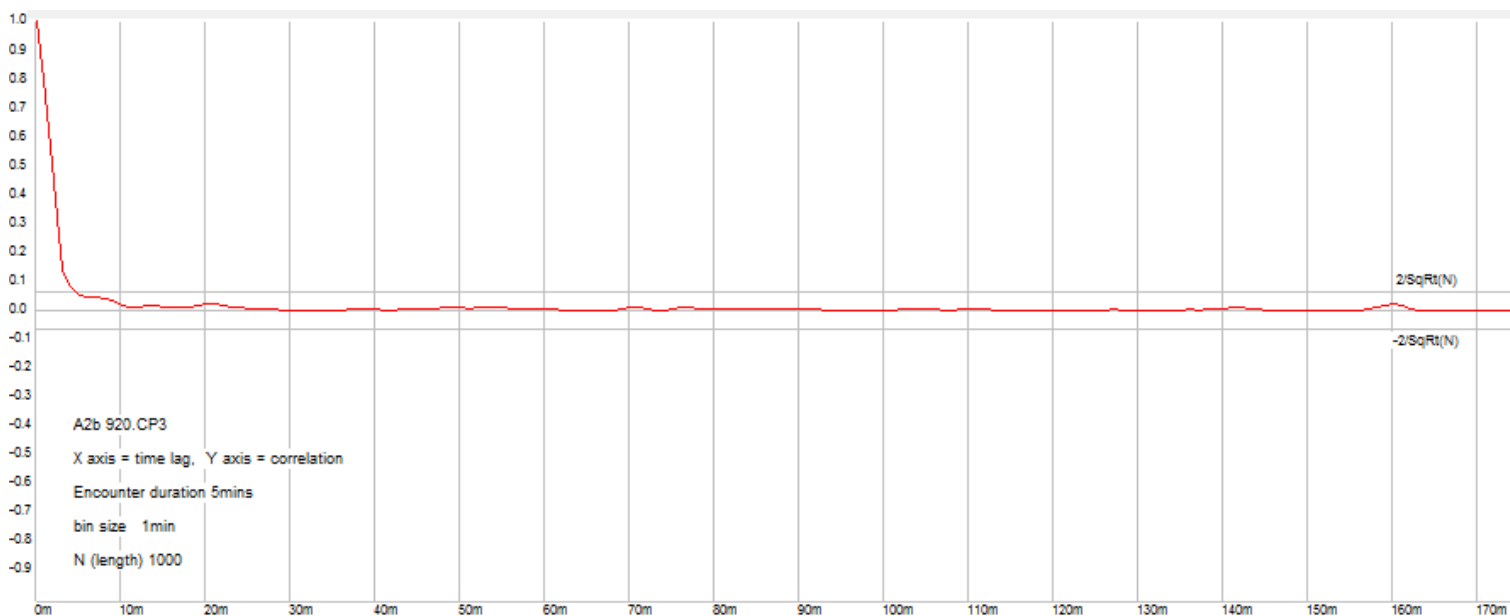
Boats show a 3 hour correlation period, which could be hauling pots, or it could be the duration of the morning migration of this species out of the harbour nearby and back again in the evening to roost.

The ac.f of the CP1 file can tell you things too:

AA 2010 05 20 POD838.CP1

X axis = time lag, Y axis = correlation

Encounter duration 180mins

bin size   30min

N (length) 1000
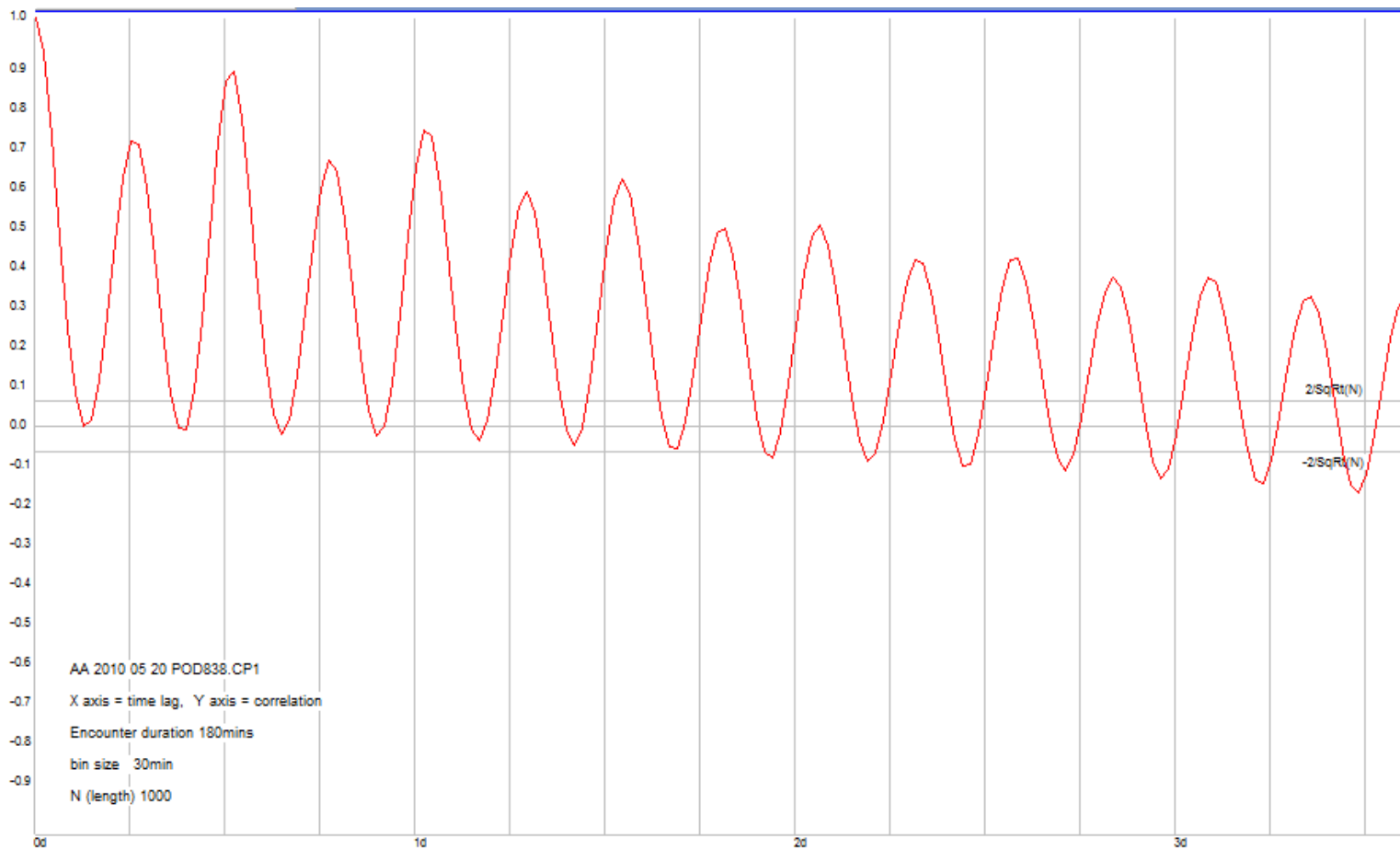
This is the ac.f of all clicks from a macrotidal estuary with sediment transport noise.

Ebb:flood asymmetry is seen, and so is the lunar cycle, which is more obvious if the time unit selected is increased to ¼ tide cycle:

AA 2010 05 20 POD838.CP1

X axis = time lag, Y axis = correlation

Encounter duration 3348mins

bin size   93min

N (length) 420
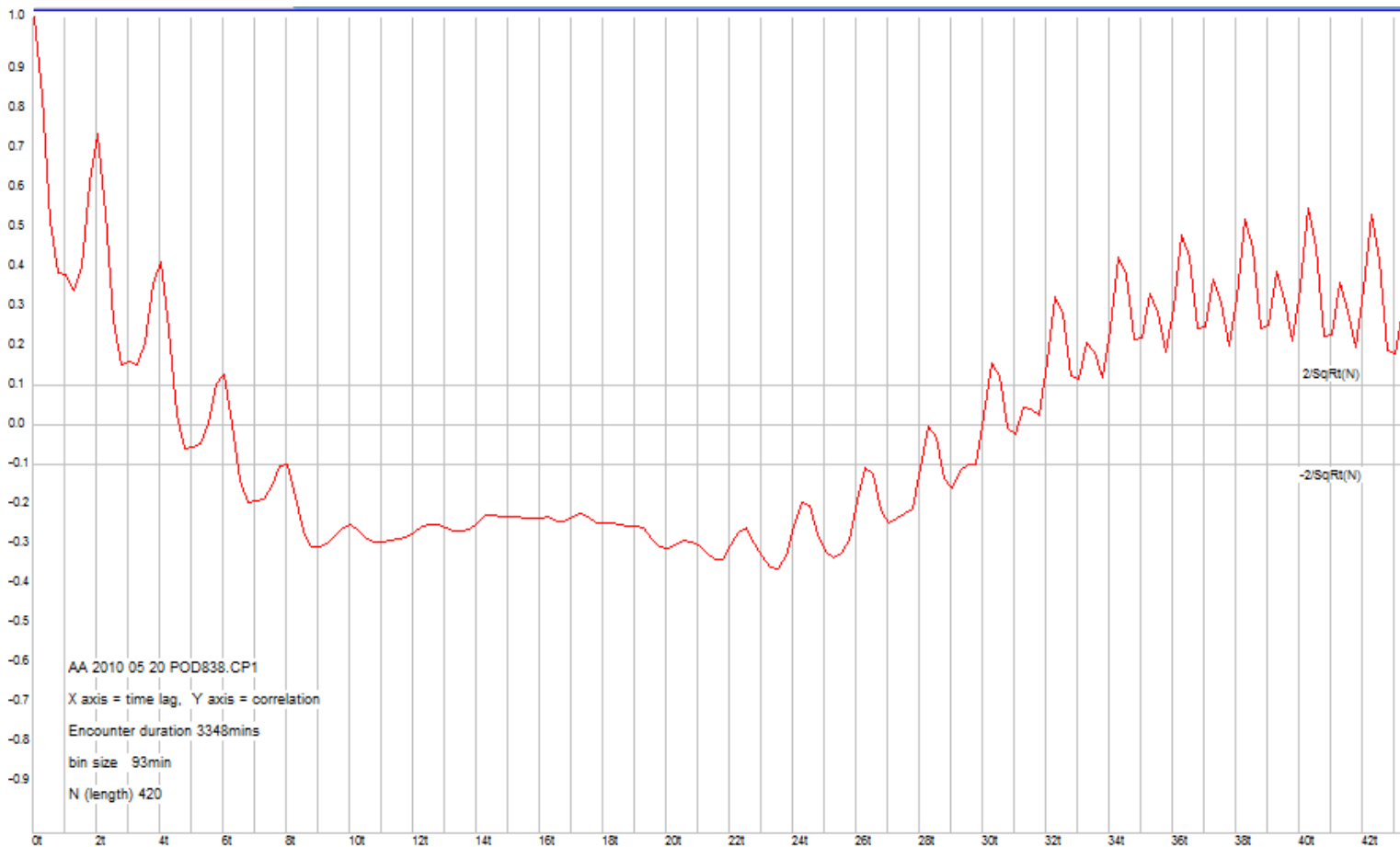
Now the correlation becomes significantly negative as the time increases until the offset data is at the opposite tidal phase.

The ac.f shows diurnal cycles in noise from crustaceans, weather patterns, etc. It's fun, but tricky!

The smallest time bin is 1 minute.

Tidal patterns will match the tidal cycle grid lines without drifting and this helps to distinguish tidal cycles from diurnal cycles which is often difficult as cetaceans are often reacting to both and the cycles are of similar length x2.

The formula used is taken from 'The Analysis of Time Series' by Chris Chatfield, 6th edition, published by Chapman and Hall, 2004. p23:

Given N observations $x_1, \ldots, x_N$, on a time series, we can form $N-1$ pairs of observations, namely, $(x_1, x_2), (x_2, x_3), \ldots, (x_{N-1}, x_N)$, where each pair of observations is separated by one time interval. Regarding the first observation in each pair as one variable, and the second observation in each pair as a second variable, then, by analogy with Equation (2.2), we can measure the correlation coefficient between adjacent observations, x_t and x_{t+1}, using the formula

$$r_1 = \frac{\sum_{t=1}^{N-1} (x_t - \bar{x}_{(1)})(x_{t+1} - \bar{x}_{(2)})}{\sqrt{\left[\sum_{t=1}^{N-1} (x_t - \bar{x}_{(1)})^2 \sum_{t=1}^{N-1} (x_{t+1} - \bar{x}_{(2)})^2\right]}} \qquad (2.3)$$

where

$$\bar{x}_{(1)} = \sum_{t=1}^{N-1} x_t / (N-1)$$

is the mean of the first observation in each of the $(N-1)$ pairs and so is the mean of the first $N-1$ observations, while

$$\bar{x}_{(2)} = \sum_{t=2}^{N} x_t / (N-1)$$

Formula 2.3 is used in CPOD.exe. Various modifications of this formula are often used for the same purpose, and generally represent minor changes that simplify the calculation with little effect on the output when the number of values is high. But in POD data the number of values can be low when larger bin sizes are used, so the full version of the formula has been used here.

The formula, for r1, gives the correlation between each minute and the next one (except that it is not every minute, as the last must be omitted as there is no subsequent minute).
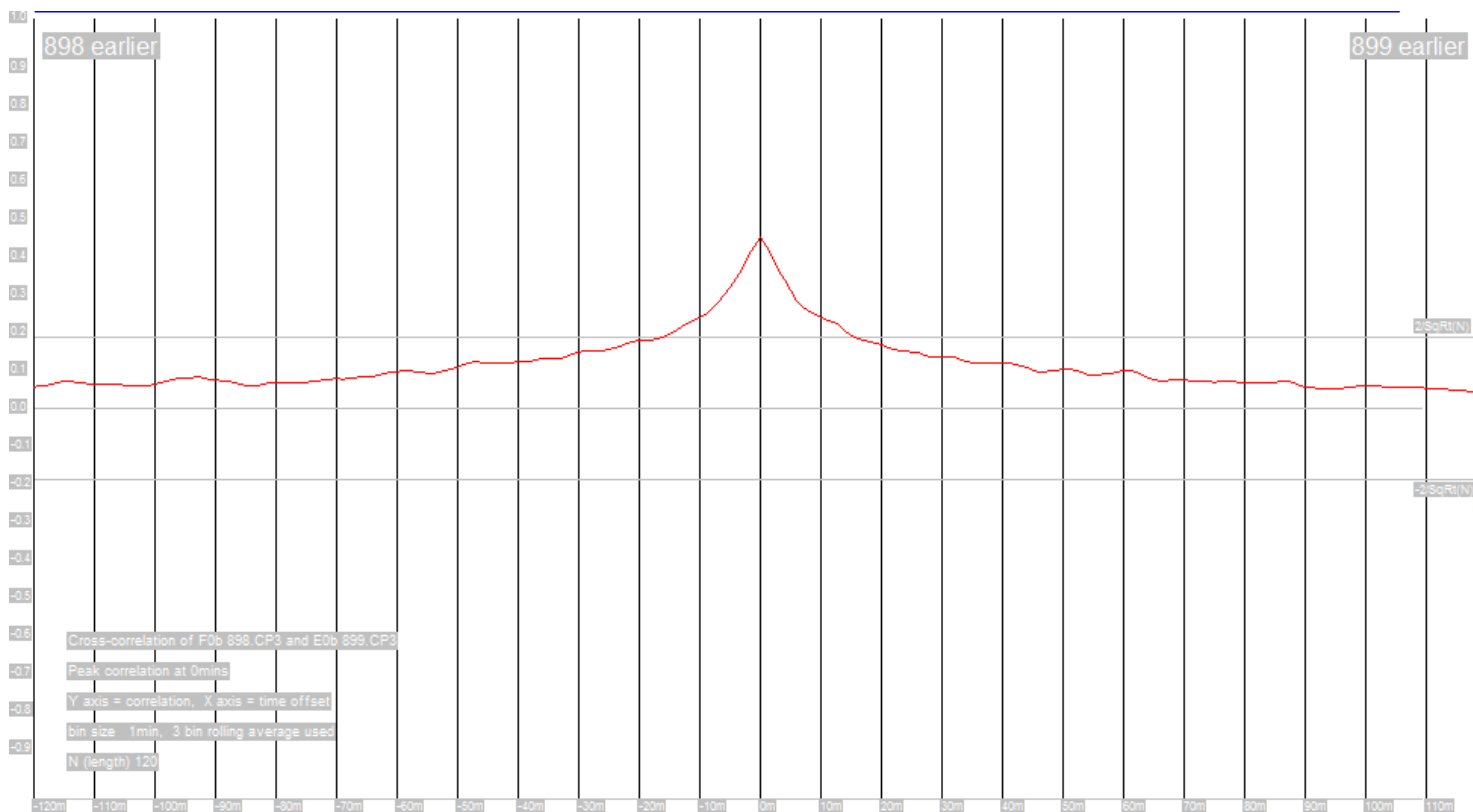
r2 is the correlation between each time unit and two time units later (in this case the last two must be omitted). In CPOD.exe the number of values in the series of lag values r1,r2... is limited to the lower of 1000 or 20% of the number of bins. The larger bin sizes smooth the output where data is sparse, but as it reduces the number of data points the length of the autocorrelation may fall to 80% of the length of the data file. The same length of autocorrelation is used for all lag values.

The graph drawn is the series of values of r1,r2,... etc and is called the correlogram. The horizontal limits marked *2/SqRt(N)* represent approximate 5% *p* values – points outside them are 95% likely to indicate a real temporal correlation between values separated by that time difference. Values commonly exceed these levels because the data is usually non-random as there are days, tides and storms as well as animal migrations, patchily distributed prey etc.

Cross-correlation

Cross-correlation works in a similar way to auto-correlation, but uses two different files and could be used on CP3 files to test for direction of travel between two POD locations.

If animals mainly went one way, the peak correlation would not be at zero minutes as it is here, comparing two PODs, numbers 898 and 899:

Cross-correlation of F0b 898.CP3 and E0b 899.CP3
Peak correlation at 0mins
Y axis = correlation, X axis = time offset
bin size   1min, 3 bin rolling average used
N (length) 120

Changing POD settings

The operational settings carried in the SETUP.CHE file can be changed via CPOD.exe.

Put the SD card into the USB SD card reader, and follow the instructions on the 'Settings' page. As elsewhere, green text can be clicked and clicking 'logging continuously', '4096' and '20kHz' will loop through available settings.

Switch angle

PODs switch ON or OFF depending on their angle to vertical when the minute rolls over. The switch angles can be changed to invert the ON/OFF orientation. The diagram shows the ON/OFF sectors as the transducer end of the POD is rotated in a vertical plane around the lid end in the centre.

These are settings for a POD to be ON facing downwards, but not ON when horizontal. Not ON at horizontal is useful to avoid using memory and power by logging while in storage or transit before immersion, but it is not an essential setting.
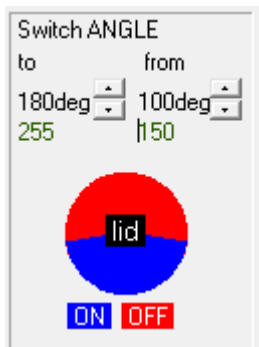
The gain, and various other standard POD settings, are not modifiable as any changes would alter the sensitivity which has been calibrated.

You can read the settings in the SETUP.CHE file (or a DAT0.CHE, CP1 or CP2 file) using the 'Show settings from a C-POD file' button. For C-PODs before number 125 (except 104) you will need to check 'show as setup version 0' to display the settings correctly.

Low sensitivity

These settings can be useful to stop the POD from logging weaker clicks which can fill the memory very quickly, especially in locations in which wind-driven or tidal currents produce sediment transport noise. The sound pressure level scale values will still be correct.

Continuous / intermittent logging

C-PODs can run intermittently to increase running time. Any number of minutes OFF can be inserted between minutes ON, but with the C-POD this is not worthwhile above 1 minute OFF as the power consumption only falls by about half when OFF. The next version of the C-POD will have a deep-sleep mode with minimal power use, so that intermittent logging can run for a number of years.

Limit on number of clicks logged in any minute

In normal use a limit of 4096 clicks per minute is used to avoid filling the memory when, for example, 120,000 clicks per minute would be logged during sediment transport events. If cetaceans are present they are most often detected within the first 4096 clicks. More investigation of this, and other noise issues, will be made.

Input high-pass filter

This filter can be set to 10,20,40,80kHz. 20kHz is the normal setting. 10kHz often shows effects (variable reduction in sensitivity) of ambient noise.

Each step up increases sensitivity to NBHF clicks but levels above 20kHz are not as useful as you might expect. The reasons are:

- In quiet locations there is not much noise to filter out.

- Train detection becomes less sensitive when click rates in a frequency band become more numerous and trains are consequently less clean.

- The rate of mis-classification of dolphins as NBHF species may increase.

Studies of NBHF species in waters with many shrimp clicks may be a good case for using the 40kHz filter.

Studies of beaked whales using deep-water deployments may be improved by using the 10kHz filter.

Handy Toolkit

Explore the items on the 'view+' page, looking at the menu bar at the top of the screen for brief notes on each. With a file open, right-click on the screen and look at the menu items. A lot of these are self-explanatory and are not mentioned here. Few users will required the advanced train detection settings.

If you want to auto crop all the files in a directory and all of its subdirectories you can opt to do this by saying yes to "Do you wish to select all files in a directory and its sub-directories?" which pops up when clicking the "Auto crop" button.

Move the mouse ponter over an item and look for the notes at the top of the screeen – e.g. for the 'Auto crop' button

Landmark sequences

These are sequences of trains showing a more or less linear decrease in ICI, and occur as porpoises approach a landmark, which in this case is either the POD or other structures nearby or beyond the POD.
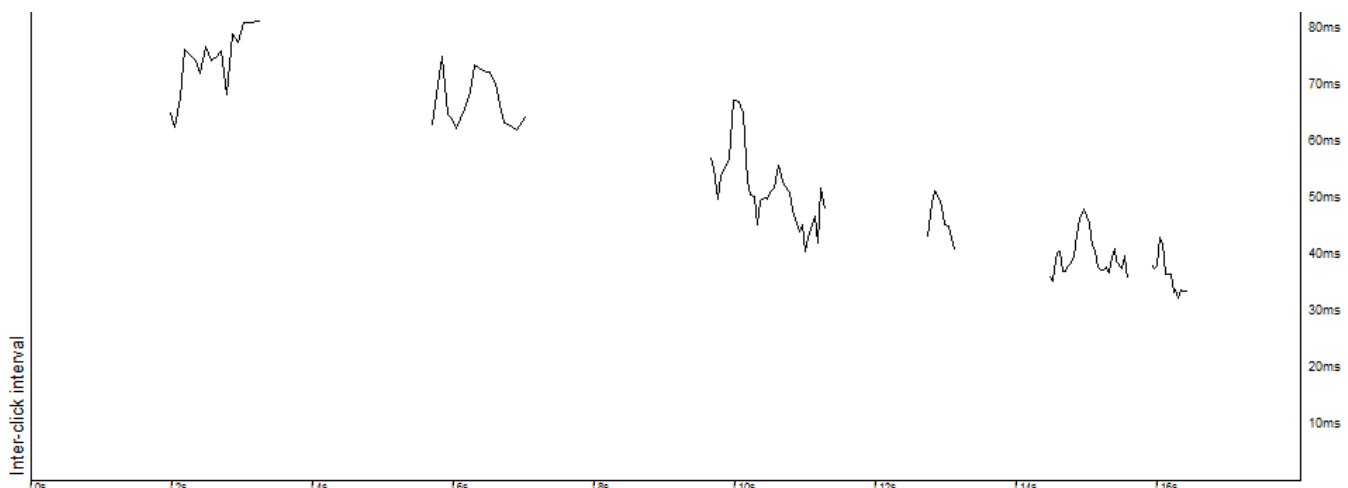
In high resolution mode you can move to the next landmark sequence by right clicking on the display and selecting 'Go to next Landmark sequence'.

Landmark sequences allow you to calculate the likely distance from the POD using the two-way travel time of a click. In an ICI of 75ms sound will travel about 1.5m * 75 = 112.5m. Two-ways gives a distance between the animal and its target of 56m. Some 'processing time', maybe 10ms, should be subtracted from the ICI.

Using the same assumptions, the swimming speed of the animal can be calculated, and they come out OK!

A summary of all the landmark sequences in a file or a batch of files can be exported using the 'List Landmark Sequences' button. The data in this export includes the above calculations.

Sonar detection

One output of the KERNO classifier is an assessment of whether a boat sonar may be present in the data. This is based on an autocorrelation that identifies repeating patterns of clicks (pulses) in the range 1/s to 40/s. Sonar risks are shown as coloured bars in low resolution displays, if 'sonar risk' is checked on the 'view+' page.

In this case a burst of clicks at porpoise frequencies coincides with sonar detection, and is identified as sonar trains.

Continuous noise source detection

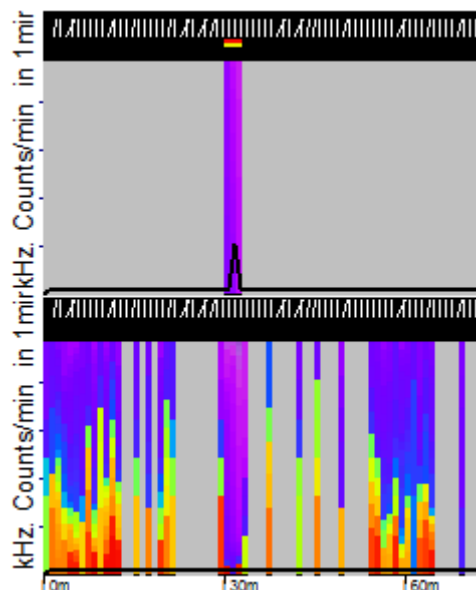Another output of the KERNO classifier is an assessment of the presence of continuous noise sources. These include:

- sediment transport noise – the one most commonly detected
- rain
- some man-made noise sources such as pingers
- unidentified sources

The dominant frequency is colour coded in the bars above the CP3 file display if 'sediment transport noise' is checked on the 'view+' page.

Export of 'detections and environment' includes the frequency of continuous noise sources, if identified. Rain may be identified as such, but the spectrum of rain is often too broad to be logged as tonal events by the C-POD.

Noise display and export

The effects of noise on echo-location and communication of odontocetes are not well understood. Low frequency sound (say, below 250kHz) propagates over very long distances so the local levels are stabilised by being averages from a many sources at different distances. High frequency sound (say, over 20kHz) is significantly absorbed during transmission so local sources are more conspicuous. Cetaceans, crustaceans, sediment in suspension, rain and breaking bubbles, cavitating propellers, boat sonars are sometimes major contributors.

Conventional sound metrics average spectral power levels over time and are not available from POD data. However their relevance to cetaceans is doubtful, as averaging the spiky amplitude profile of shrimp clicks over time obscures the fact that the brief lulls between clicks may mean they are less significant than, say, transport of fine sediment that produces a more uniform background that will interfere with every returning echo.

To provide some insight into nose patterns a noise display is available for minutes and larger units:
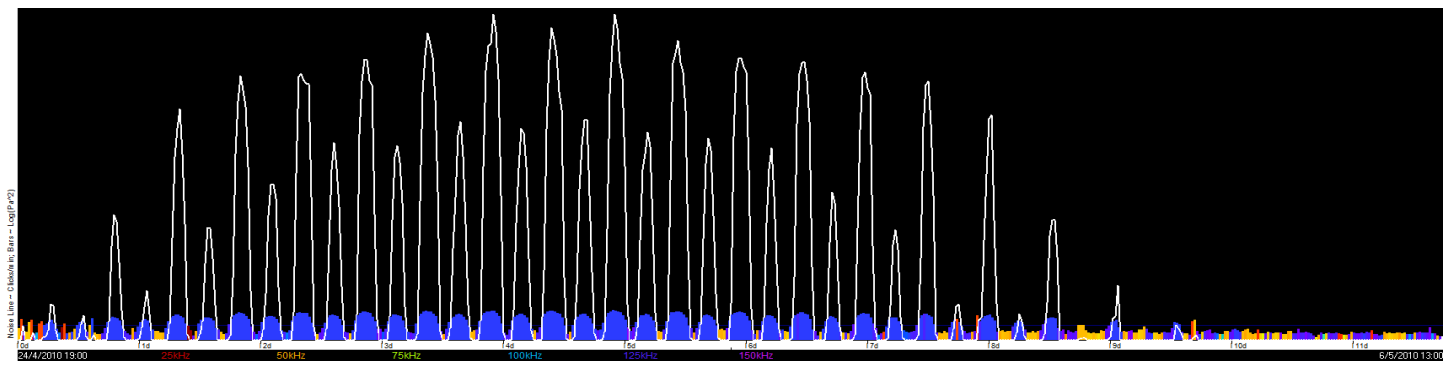
The white line shows the estimated number of clicks per minute. In noisy situations the limit on the number of clicks per minute may be reached well before the end of the minute, so the noise display presents an estimate of how many clicks would be logged if there were no limit. The graph above shows a period of spring tides in a macrotidal estuary, with the asymmetry between flood and ebb shown very clearly.

The graphic below shows the frequency distribution and click counts from the same location, where high levels of tonal high frequency ultrasound max out the logging during periods of rapid tidal flow.
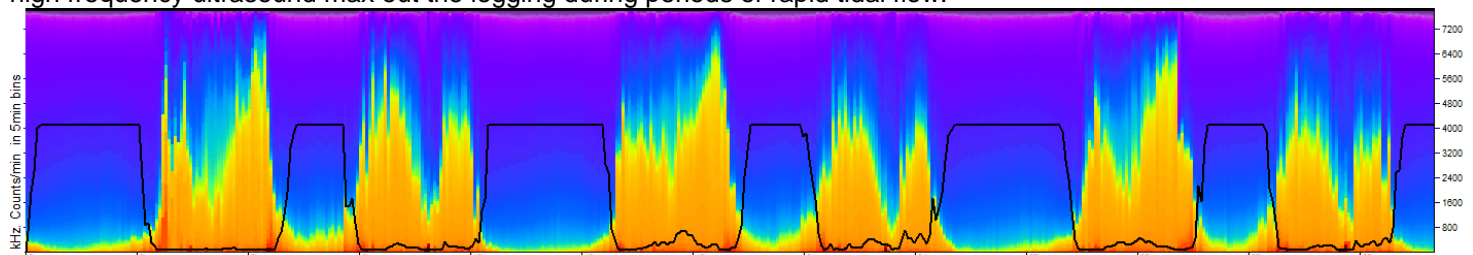
The noise display also shows vertical coloured bars. The colour represents the modal frequency (kHz) of the clicks logged. The height represents the Log of the sum of the squared sound pressure level (Pa) of logged clicks i.e. it accounts for the frequency response of the hydrophone.
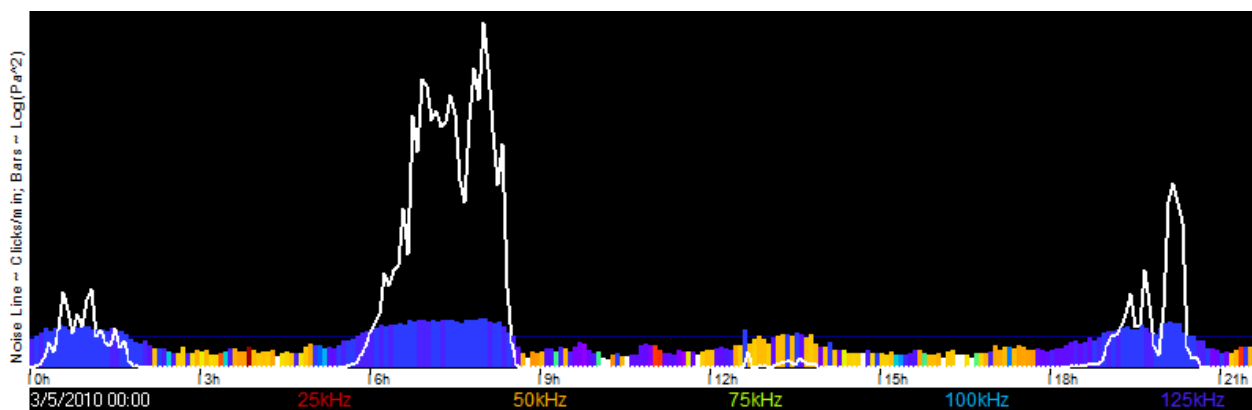
The relevant values can be exported via the export page. A novel metric, MMM, that is exported as well is explained in the section on exporting data.

All noise metrics require a lot of work to evaluate their significance for cetaceans and for cetacean detection. This is not easy! Increasing noise impairs echo-location, impairs the detection of their clicks and the performance of train detectors.

Temperatures

These can be viewed with or without grid-lines. Changes are usually gradual but there is +-2 degree C of noise on the temperature reading, which is smoothed out when intervals longer than 30 minutes are viewed. The temperature scale of C-PODs is not been calibrated.

Big temperature swings (more than 1 degree per day, averaged over hours) often indicate mixing of water bodies unless the POD is in very shallow water.

Changing the start time of a file – the 'Files' section of the 'view+' page

The 'Change start time' will change all open files to the new start time which is taken from the 'Read SD' page, so you need to put the correct start time there before you click this button. If a CP3 file exists you need to change that along with the CP1 or they will not match up any more.
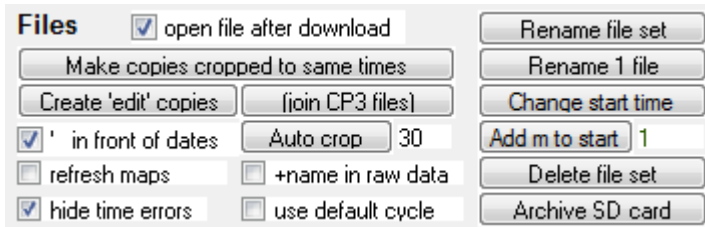
For small time changes it's quicker to use 'Add m to start' m = minutes and if you want to move the start one hour earlier enter -60 - actually you need to enter 60 and then put the '–' in front. A day later is 1440.

Changing the name

If you need to rename a file it is best to use the 'rename file set' button because it will rename the CP1, CP3 and their map files all in one go, so that they work together.

This button will ask you if you want to construct the new file name in the standard way by taking the location and upload date from the Read SD page. If you choose no you can construct your own new name, or escape from the process.

Deleting map files

If in doubt about whether to delete a map file (*.c1m or *.c2m) that you think may belong to a data file you have deleted, do delete it – the program will always re-create this file if it is needed. These files speed up movement through large files but contain no acoustic data at all.

Join CP3 files

This button allows you to join CP3 files together so that where you had four CP1 files and four corresponding CP3 files the first of the CP3 files will now cover the whole data set. They are joined in alphabetical name sequence which will be correct if they were the names given by CPOD.exe, but might not be if you have changed them. The first CP3 file will now 'know' the total number of clicks in the CP1 files, and in each minute of them.

Capture graphics

Ctrl G captures the display area of the screen, so that you can paste it directly into your document, like this:
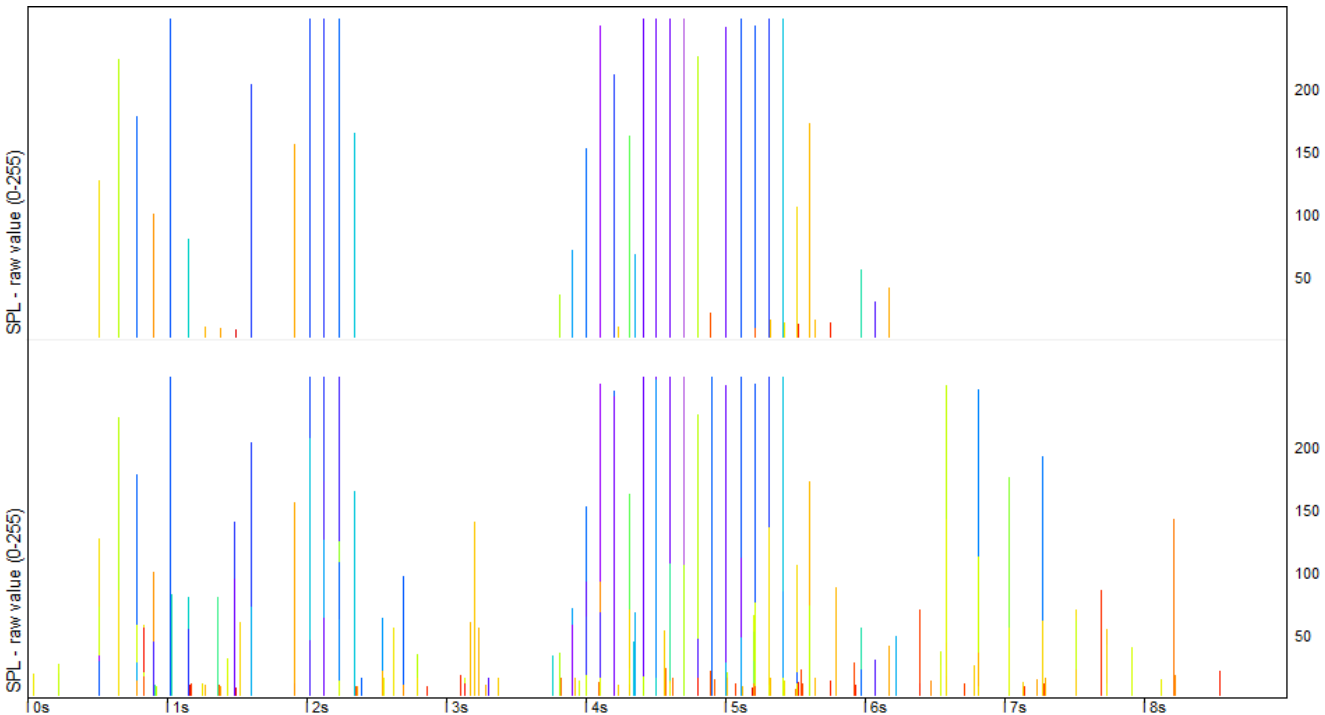
You are offered the option of having the lines in black, but the backgrounds are always changed to white to make printing easier.
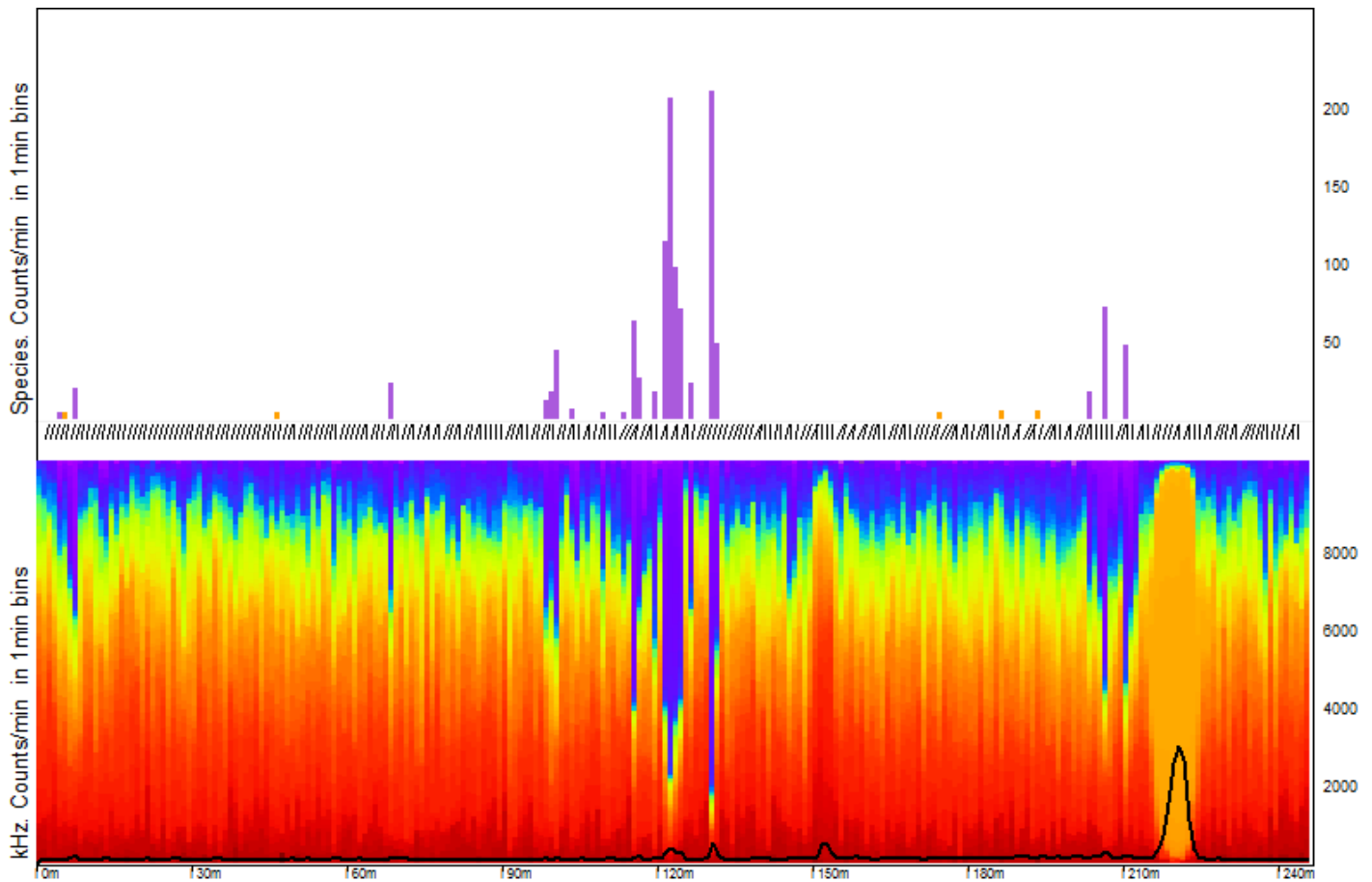
(Here you see a quiet location with most of the high frequency noise events coming from NBHF species, and two lower frequency events that will be boat sonars, as they show a narrow frequency range.)

Annotations

You can add up to 6 annotations (= text notes identifying species, etc) to any minute. To view the annotations, check the 'annotations' box on the 'view+' page.
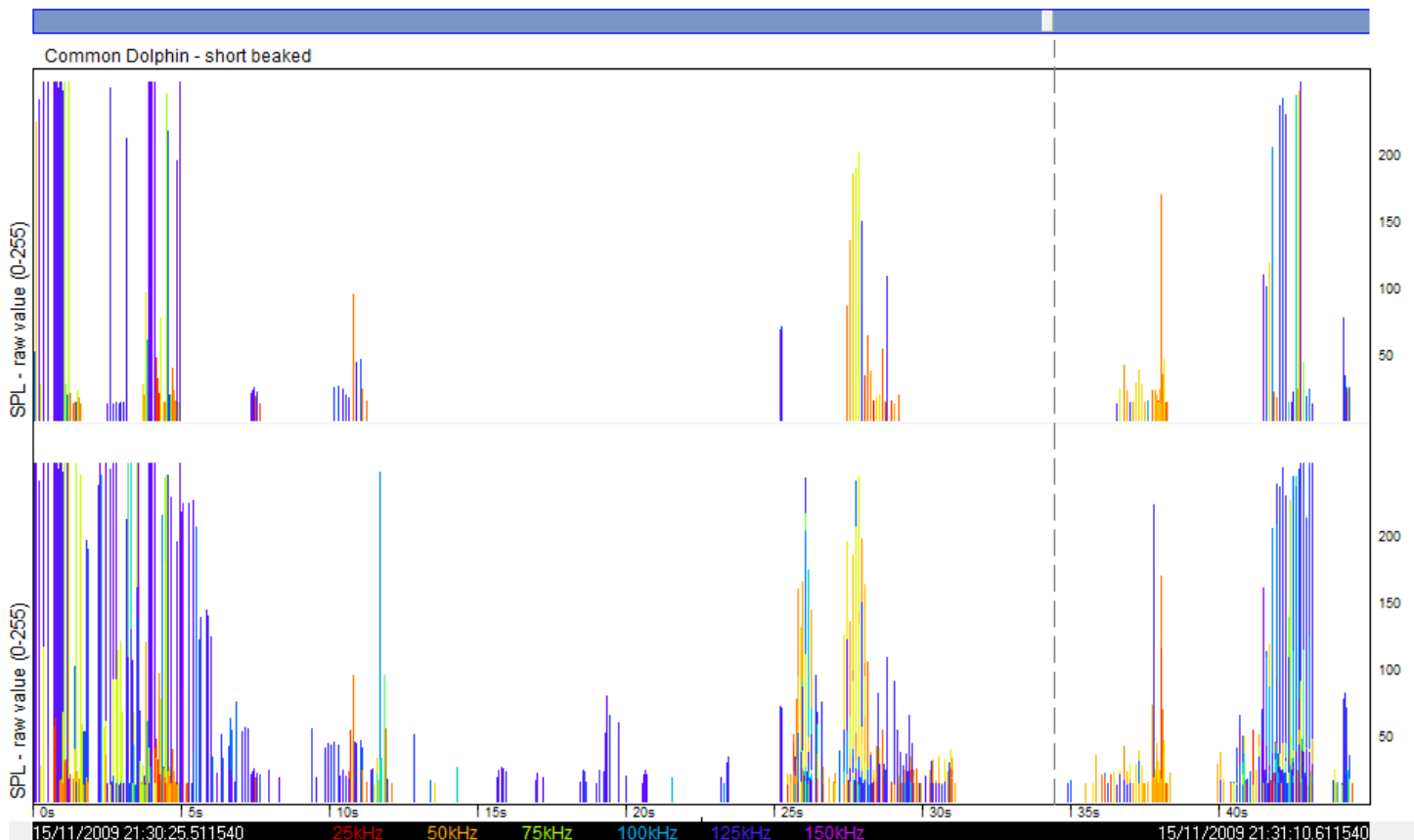
Enter your sightings into the spreadsheet 'CPOD sightings entry.xls' and follow the instructions given there. 'Species' can include sea state, boats, etc! You will see that you can invent up to 10 'species' of your own. They could be behaviours, presence of birds, etc. The text for those is written into the file so they can be displayed like the 'species' provided.

Deployment codes

You can define codes to go into the notes field of each file when the SD card is read. Instructions on this are found in the file 'CPODcodes.txt' which is itself a template.

If 'CPODcodes.txt' is found in the directory you are exporting from when you use the 'List Files' button, CPOD.exe will find any codes that are in 'CPODcodes.txt' and are also in the file 'Notes' field, and will export the code to the file listing. That list can be read into a spreadsheet or database, and used in analysis.

Time interval clustering

This handy measure of detector performance is really a development tool, but you may find it useful.

The idea is that because cetacean encounters typically create a bunch of detections, and then there is a silent period with none at all, you can assess this temporal clustering as a way of detecting scattered false positives. They will break up the silent periods.

So the 'TIC value' is calculated by:

- summing the longer of the inter-train intervals before or after a train and separately summing the shorter of those intervals and then dividing the total of longer values by the total of shorter values.

Easy! High values are good, low are bad.

Random trains give a TIC value of 2.0 if there are enough of them (i.e. the value is subject to sampling error).

The 'TIC value' comes with some health warnings:

- If you make the detector tougher (say you go to just Hi Q trains, or you raise the minimum mean train SPL) you will get fewer detections and some encounters are lost. This may improve the TIC value even though the detector actually had no false positives. To limit this the TIC value has an upper limit on 'longer' periods (set in development tools on the 'view+' page). And the number of trains is reported to give you a clue.

- If the source of false positives is as strongly clustered in time as cetaceans are, then it will not reflect them as false positives. Boat sonars are an example. So it's really an assessment of performance on clustered train sources, not just cetaceans.

The TIC value is reported routinely on the 'Analysis' page and you can get it any time with Ctrl Q, which gives you:

That's a moderate TIC value on a file you've never looked at before, and you immediately know there's some good, or interesting bad, stuff in it!

Drop all filters and it falls to 2.9. That's not hopeless, but it tells you there are noise trains in the file.

If you visually marked bad trains, you can, in just a minute, copy the CP3, remove the marks from the copy, open both, filter out marked trains, press Ctrl Q and see the TIC on both.
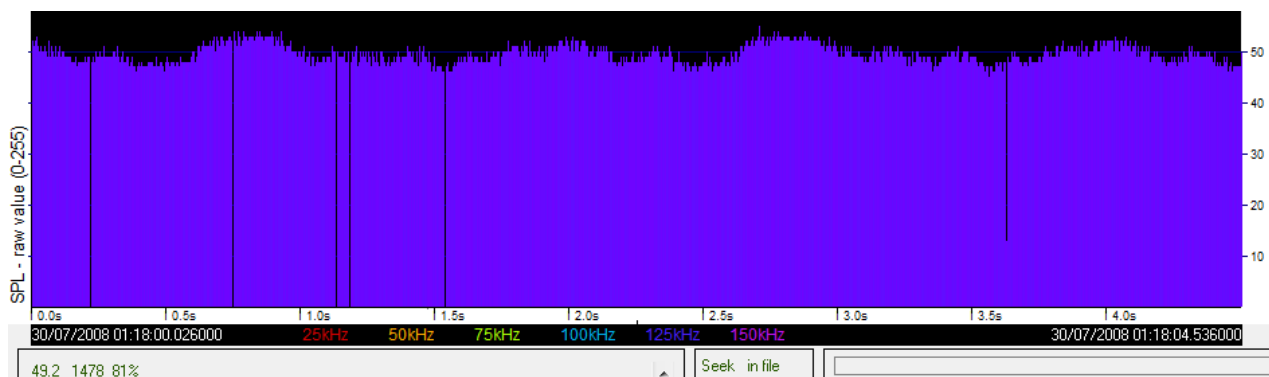
For detector developers, Nix Tic can save an age doing manual assessment and annotations.

Development tools – avSPL

This is a tool for calibrators:

Check this box and the average SPL of clicks louder than 25 in the time on display is shown in the information box, bottom left of the screen.

These data come from a rotational test of a POD. The mean SPL is 49.2, and the minimum is 81% of the maximum. 1478 clicks were used in the calculation.

Cetacean detection

Download 'CPOD – How it works' from www.chelonia.co.uk.

Species identification

This is a theoretical diagram of the problem of classifying the species origin of a train using the mean value of the bandwidth and frequency of the clicks in a train:
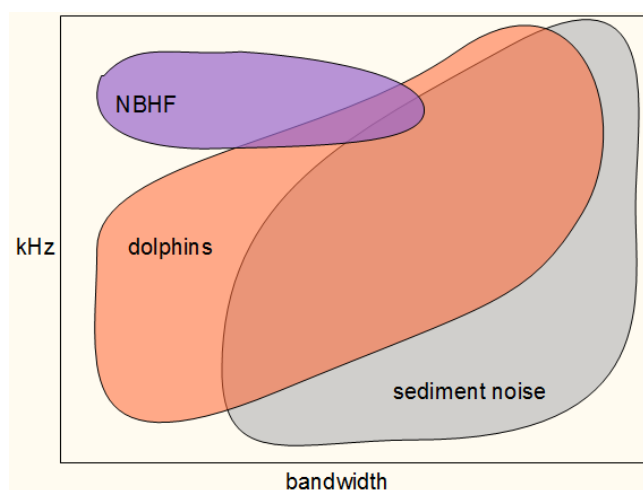
The overlap between noise and dolphin clicks is great, and this is why you see many dolphin trains that are 'unclassified'. The overlap between NBHF and noise is smaller, but still significant. The most surprising overlap is between dolphins and NBHF species, as the published literature might lead you to expect almost none. This is because most detailed investigations of clicks are based, for good reasons, on attempts to capture the centre of the click beam at fairly short range. Relatively small (= a few hundred to a few thousand) samples of clicks are captured. As a result the whole spectrum of acoustic behaviour of the animals, and the directions off the centre of the beam, are not adequately represented, but they are frequently seen in the data captured by static loggers that run for long periods.

('Sediment noise' is most often sand in suspension. It emits noise pulses at high rates, that depend on the density of the particles (n/m^3) at frequencies that depend on the diameter of the particles, and it is a major contributor to background noise in shallow waters.)

The KERNO classifier uses more numerical descriptors of trains than the two shown above, and prioritises distinguishing non-train sources (sediment noise, shrimps, propellers, etc) from train sources and represents this as the Q value. Trains in the areas overlapped by noise, as above, may be given low Q values, even when there is little noise present, but still have correct species identification.

The limitations of the KERNO species classification are:

- Dolphins produce some narrowband high frequency trains. This is seen more commonly from the small delphinids (<2m long) than from *Tursiops truncatus*, but is seen very often from *Tursiops aduncus.*

- NBHF trains are occasionally misclassified as 'other cetacean' (rather than 'unclassified'). This happens most often with very loud clicks with a great deal of multipath propagation, which causes a spread of frequencies.

- Beaked whales are particularly likely to be classified by KERNO as sonars because they produce long clicks at unusually regular inter-click-intervals. Trains not classed as sonars are likely to remain unclassified because of

the sonar risk. To analyse beaked whale data it is useful to remove all filters and export encounter details. More on this in encounters.

- Orcas, Pilot whales, and many other species have not been assessed and will often not be correctly classified by KERNO. As large enough data sets become available they will be very interesting, and a range of classifiers will be developed for them.

Encounter classifiers:

Perhaps the biggest problem in acoustic monitoring of cetaceans is the task of analysing the acoustic data to arrive at a useful classification. Cetaceans can be detected in a variety of ways:

1. Detecting individual clicks with characteristics that match those of particular cetacean types.

2. Detecting the click trains made by cetaceans.

3. Detecting encounters with cetaceans as periods when either or both of the methods above are giving multiple detections.

Click characterisation methods (1) are useful for porpoises, although weak. For dolphins they are both time consuming and inefficient – i.e. high sensitivity cannot be achieved without high levels of false positives except in very quiet locations.

 The KERNO classifier works at the level of individual trains (2) and has some weak positive feedback over a period of a minute. Its main limitations are:
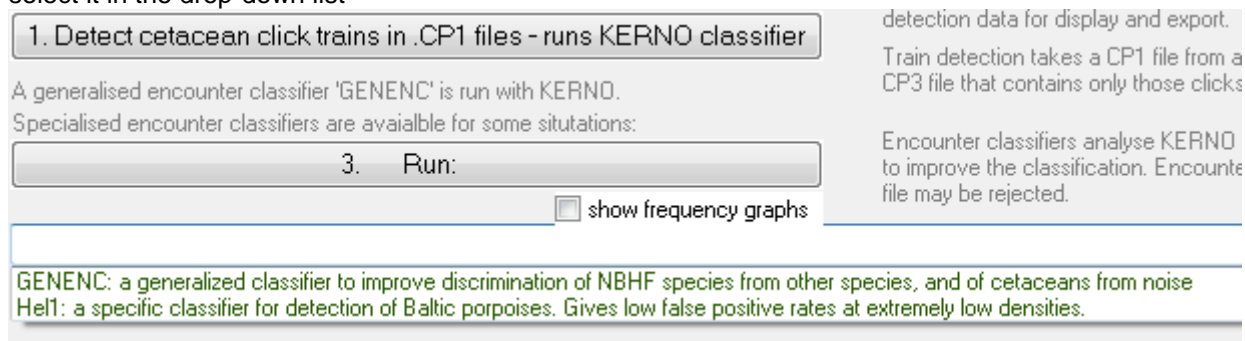
- It classifies dolphins as 'NBHF' when they produce long high frequency clicks, as they then resemble porpoises more closely than the currently published studies generally indicate.

- It classifies some porpoises (NBHF species) as dolphins when there is a strong noise background, or when they are very close.

- It fails to classify a large fraction of dolphin trains as they resemble beach noise and other sources. It has been viewed as too conservative by some users. The reason for that is that false positives are generally much more damaging than false negatives.

Encounter classifiers (3) take in a wider time span and achieve much better results in these difficult situations. In CPOD.exe they run quickly as they only look at the output of the KERNO classifier in parts of the data demarcated by Mod or Hi quality trains.
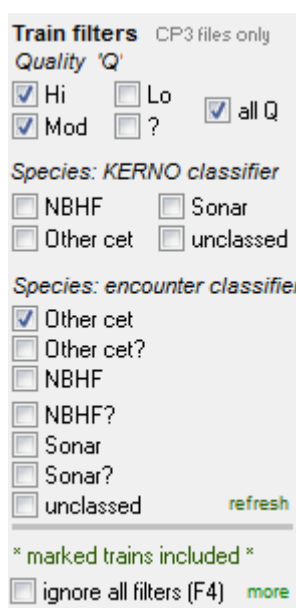
- Available classifiers are listed in the drop-down list on the 'Trains' page – the 'Hel1' classifier was the first.

- These classifiers will, like the KERNO classifier, *not* be gradually improved. Better versions will be added to the list with a different name.

- The GENENC classifier is automatically run after the KERNO classifier. It, and other encounter classifiers, can be run on more recent CP3 files without re-running KERNO and a message is placed on the Results page if the CP3 file is not suitable. In that case run the CP1 file through the KERNO classifier and it will produce a CP3 file that can be handled by the encounter classifier.

- They will give trains an 'encounter species' (or leave them unclassified). These species do not replace the KERNO species, and are only filterable independently of those species.

- They erase the results of any previous encounter classifier. If you wanted to compare results in copies of the same file this could be done my 'marking' the trains, as those marks are not erased.

- The last classifier used on a CP3 file, if any, is shown on the Files & POD section of the Metadata page.

To *run and use the classifier (note… GENENC is automatically run after the KERNO classifier in version 3.032 and higher, but can be run on CP3 files created with earlier versions)*:

- select it in the drop-down list

- the button, right, will show 'RUN: GENEC'. Click this.

- select either the CP3 or the CP1 files that you wish to process. Both must be present in the same directory. If you select both the process is run twice on the same pair of files.

- the normal settings for using the results of this classifier for dolphins are as below:

The classes 'Other cet?' and 'NBHF?' correspond to selecting lower quality classes when using the KERNO classifier output i.e. the risk of false positives is higher and visual validation is necessary.

The *GENENC* classifier is useful in these situations:

- For NBHF species occurring with other cetaceans that are using high frequencies (false NBHF positives) This has been a significant problem in Peru and in the Upper Gulf of California.

- For NBHF species occurring with sediment transport noise (false NBHF positives). This has been a problem in the River Weser, Germany

- For any broadband species, as GENENC reduces false negatives that were previously unclassified trains.

- Where a broadband species of interest is rare and an NBHF species is common (NBHF sometimes appeared as false positive 'other cet')

The main limitation of *GENENC* is:

- Where there are significant densities of WUTS (weak unknown train sources) GENENC moves many, but not all, of these into the broadband or 'unclassed' classes, so visual data validation is necessary to assess their incidence or remove them.

GENENC is not required in:

- NBHF studies in generally quiet locations with no non-NBHF cetaceans. There may be a minimal gain in NBHF detection positive minutes, or no significant change, from using GENENC as there are no sources of false positives and false negatives are low in these situations.

- Studies in which true NBHF detections are both numerous and far in excess of any source of false positives. GENENC might increase the detection rate significantly, but that will merely be equivalent to raising the sensitivity of the instruments when they are already far in excess of the minimum sensitivity required to answer the questions of the study.

The *Hel1* classifier:

.. aims to remove false positives arising mainly from Baltic WUTS and noise 'mini-bursts'. It was developed on the basis of work done at the Hel Marine Station in Poland in 2010 and it achieves very low levels of false positives at the cost of losing about 15% of true positives. Its efficiency depends on the assumption that there are no dolphins present or other sources of false positives not seen in Polish waters, so its application outside those areas would require visual validation.

Using encounter classifications:

Select any of the species in the lower section of 'Train filters'. The KERNO classification is suppressed and only those trains given one of the species classes selected are shown.

The available options show up to 6 species identified by whatever encounter classifier has been used on the open CP3 files.

When the encounter classifier species are in use the Hi, Mod, Lo, ? quality classes are no longer operative. Instead there may be a choice of quality levels provided by the encounter classifier e.g. NBHF and NBHF? from GENENC.

To *see the clicks rejected by the encounter classifier:*

If you want to examine, for example, NBHF trains rejected by the encounter classifier, set the selection set as above and go to Files page and click 'mark current selection of trains'

-
- set 'exclude marked trains'
- set 'NBHF'
- set whatever Q filters are of interest

Why use encounter species classifications? They can:

- greatly reduce the incidence of false positives.
- improve discrimination between species.

Program versions

Version 2.000 replaced the original (provisional) train detection with the KERNO classifier which greatly improves the results, typically increasing true positives while greatly reducing false positives. Subsequent versions of CPOD.exe add

display, analysis or export features and fix bugs, but try to avoid any changes, including improvements, affecting the results of processing a file as this would create incompatible results. However there have been errors:

The latest version is preferred because all known bugs have been fixed.

Version 2.020 introduces an encounter classifier. Earlier versions cannot 'see' the classifications produced by this.

Files processed with 2.016 and 2.017 require re-processing and a warning is given by subsequent versions when any CP3 file created with these versions is opened.

Versions to avoid:

Three, briefly released versions affected the performance of the classifier: 2.007, 2.016, 2.017.

Early versions did not always read the whole of long data sets (>1GB) These were version 2.007 and earlier.

Files processed in earlier versions can always be shown and data can be exported from them. Files processed with the KERNO classifier can be filtered by mean click rate (1/ICI), mean click amplitude, modal frequency, or number of clicks in the train, but these filters do not work on CP3 files made pre-KERNO.

Version 2.n also introduced a sediment transport noise detector, which is useful in recognising soundscapes, and an error estimator, giving instant qualitative reports on potential risks in the analysis of files, based on noise levels, sonar prevalence and WUT risk. These may be developed further.

Hardware notes

C-PODs log tonal ultrasounds including those from all odontocetes (cetaceans with teeth: toothed whales, dolphins and porpoises) except the sperm whale. All these species produce clicks in trains at rates that can vary from 1 per second to over 1500 per second.

Simple testing

To check your POD is working:

Start it.

Put it upright – transducer up – and leave it for 1 minute

Make light quick strokes of the transducer housing with one finger at about 1 per second in some distinctive pattern – say 3 strokes - 2sec pause – 2 strokes – 2 sec pause – 3 strokes.

Stop the POD and read the file. Do not process it to find trains.

View it at 10ms resolution and you should see your barcode 3 - 2 – 3.

Acoustic performance and standardisation

Transducer: Piezo-ceramic transducer sensitive to all frequencies below 160kHz. The transducer housing is a highly abrasion resistant plastic, and the transducer mounting is shock-resistant. No cases of broken or damaged C-POD transducers have been found so far.

Incoming sound is high-pass filtered, with options of 20kHz, 40kHz, 80kHz.

Standardisation:

- The complete instrument is rotated in a sound field and adjusted to give a radially averaged, temperature corrected, pressure reading within 5% of the standard at 130kHz (±0.5dB).

- Radial variation < ±3dB at 130kHz.

Hydrophone frequency response
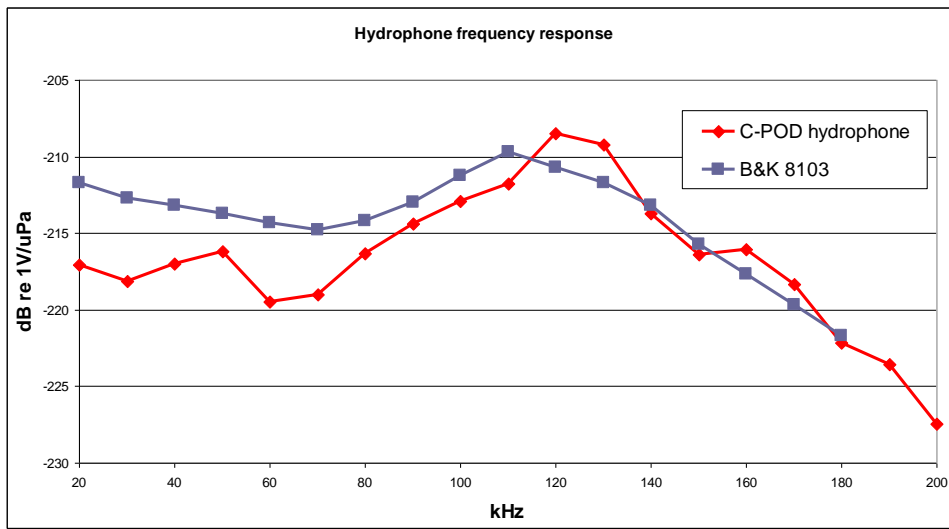
The frequency response (from tests by the National Physical Laboratory in the UK) of the C-POD hydrophone is shown above in comparison with the widely used B&K 8103.
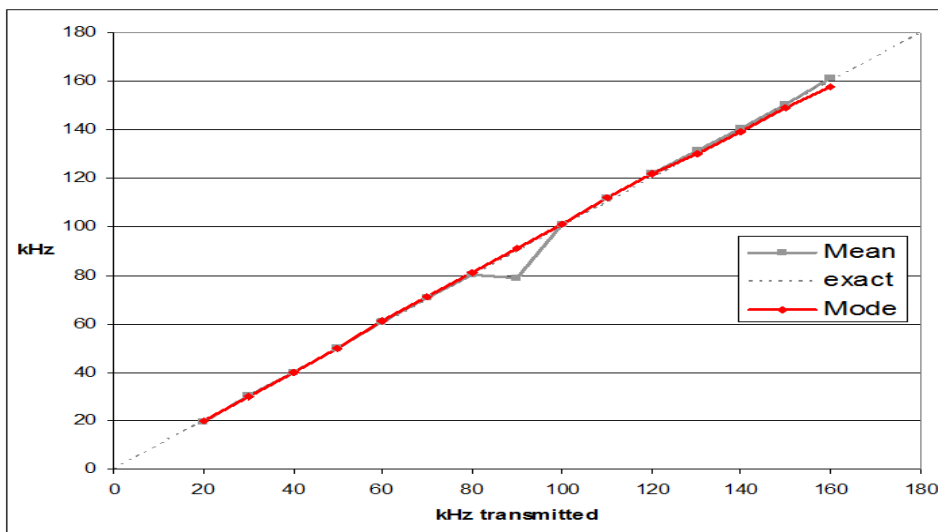
The frequency readout for pure tones of 10 cycles duration is shown above from tests by the National Physical Laboratory in the UK. At 90kHz some low values arise from an internal resonance.
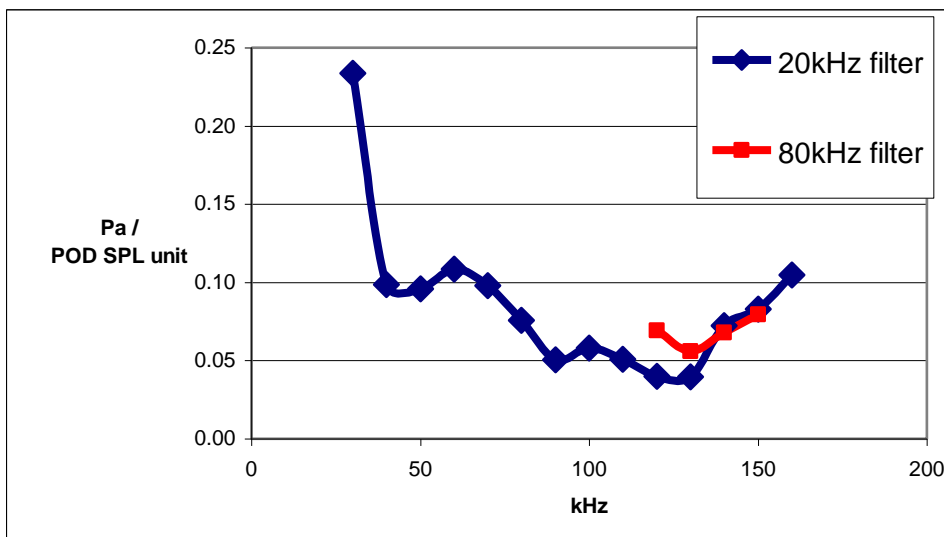
The frequency response scale of a C-POD is shown above. The C-POD software allows sound pressure levels to be viewed as the raw value or, corrected for frequency, as Pascals.

Porpoise clicks are detected down to a minimum of approximately 0.5Pa peak to peak, with maximum detection range of approximately 300m. Dolphins are detected at much longer distances.

Batteries

10 alkaline D-cells. The cells are in two stacks and the lower powered stack is selected and exhausted first. This allows the second stack to be re-used unless maximum running times are required.

Running time

Running time: 5 – 6 months. 4GB is adequate in most locations for this, but on sandy substrates with movement 4GB may be inadequate.

Housing

- The housing is buoyant when fully loaded. This allows recovery of instruments that have been lost at sea. Polypropylene tube 90mm outside diameter, 660mm long, with mooring line 1.5m long.

- A stainless steel mooring ring at the mid-point of the housing holds a polypropylene mooring line with abrasion shield and spliced attachment loop. The mid-point attachment minimises leaning downstream in a current.

- Weight: 2.1kg without batteries. Batteries 1.6kg. Buoyancy loaded = 0.5kg in freshwater.

- Depth: 100m. A single POD has been tested to destruction and failed at 220m.

- Two external rubber buffer rings to reduce impacts.

Deep C-POD

A deep-water version is available that differs only having a heavy anodised aluminium housing that is capable to immersion to over 3000m. The dimensions are:

- Length: 686mm

- Maximum diameter: 112mm (includes rubber buffer rings)

- Weight in air, no batteries: 6.7kg Batteries 1.6kg

- Safety: a pressure bleed screw is present for use if water ingress is detected. No instances of this have yet occurred.

FAQ, and notes:

Decibels. Where are the decibels?

C-PODs sense pressure. The basic physical unit of pressure is the Pascal.

Decibels are multiples of logarithmic ratios of power, not pressure.

Hydrophones respond linearly to pressure.

To convert a pressure to any measure of power (energy/time) requires taking the root mean square of the pressure through a period of time.

To convert a single point pressure to any measure of power involves assuming:

- the waveform is sinusoidal – usually false.

- the waveform continues uniformly – often seriously false, and always false in cetacean acoustics.

- the pressure is the peak of the wave – true for C-POD data.

But isn't the decibel the correct standard that we should use?

- The decibel is rightly deprecated by the International Committee for Weights and Measures, because it is unnecessarily complicated and does not facilitate learning and communication.

- Progressives in underwater acoustics advocate its abandonment – see Medwin's superb text book 'Sound in the Sea: from Ocean Acoustics to Acoustical Oceanography'.

- The decibel is grossly complicated by being expressed as a logarithm that is then multiplied by 10 (if the measurement is a power-like unit) or 20 (if the measurement is pressure-like), but the reader may not be clear about when or whether the author took the root mean square of the pressure etc.

- The SI units of energy and power are the Joule and the Watt. Joules are familiar to biologists, and are used to measure the energy content of the diet of animals. The ability to think easily across such diverse physical aspects as sound output and diet is a great gift from physics to biology that is obstructed by use of the decibel.

- For most users, except the very talented, it takes time to become conversant with such units and time is a crucial finite resource for many mortals. Biologists will also be needing time to struggle with multivariate statistics, information criteria, linkage disequilibrium, etc. etc.

- The decibel is associated with an exceptionally high number of errors in published papers. These errors have all got past multiple authors and referees.

- The decibel is used in electronics, but it causes trouble there too. For example, the gain of an amplifier in dB is misleading if the impedance has changed, where the gain expressed as a simple ratio of voltage out/voltage in would have been exactly correct.

- In an attempt to accommodate the decibel two definitions of 'sound pressure level' have been created. (ISO uses the point value of acoustic pressure, which is used here. ANSI uses the RMS acoustic pressure.) This dramatically illustrates the need to move to SI units, because you often do not know if authors were using this definition or not.

- Units that cause confusion should never be tolerated on the supposition that they're basically easy and those who are confused are really lacking intelligence. (Does a reluctance to remove confusion indicate intelligence?) One of the three countries that do not accept SI units lost the $0.3billion Mars Climate Observer mission through unit confusion. The other two countries are Liberia and Gabon.

But, surely, the decibel has these advantages:

- The decibel enables a wide range of values to be neatly expressed: true, but powers of 10 do that neatly for far wider scales of length etc. (Did you know: the diameter of Earth is 61.05dB re 1m while the diameter of a human hair is -42dB re 1m).

- The decibel avoids the need to multiply and divide: true – that did aid acoustic thinking in the pre-calculator era, and was its main value then. Now we even use tools to add and subtract!

- Sound intensity falls in proportion to the square of the distance from the source (in the far field, apart from absorption) so decibels represent that more easily: true, but sound pressure fall linearly with distance from the source (in the far field, apart from absorption).

- The decibel corresponds to perceptible differences in intensity: partly true – for part of the human hearing scale that is roughly true, but it's no way to define a physical unit and is totally irrelevant to most uses. A log scale for frequency would actually fit perceptible pitch differences rather well, but fortunately we still use a linear scale for this.

- Those who've mastered decibels don't need to struggle with a change in practice: true, but they impose indefinite confusion and complexity on those that follow them, and obstruct some interesting cross-discipline physics.

Only one of these advantages of the decibel, the last, is actually substantial.

Consider a culture in which the sensitivity of a hydrophone is described as 26.0µV / Pascal

… a proposition is made that this would be better expressed as -211.7dB re 1V/µPa

If you support that proposition, which is current practice, please send your arguments to Nick.Tregenza@chelonia.co.uk and, if appropriate, they will be added to this document.

How much of the time are cetaceans clicking?

Porpoises click for a very large fraction of time – probably more than 85%, and other cetaceans in darkness or turbid water probably also click a great deal of the time. A more thorough answer to this important question is needed!

Could you stop changing the software, please?

Yes/No. The number of changes is slowing down! Apologies for the occasional new error!

Over some months and longer there will be major additions, especially to the secondary classifiers where a series of new classifiers will be introduced.

Which version should I use?

The latest. You do not need to re-process your files to detect trains.

Do C-PODs detect whistles?

Whistles above 20kHz could be detected. They would appear as a series of maximum length (255 cycles) clicks with minimal silent gaps between, or as shorter events. Most whistle content is below 20kHz, some whistles show significant spectral complexity, so the C-POD is not a suitable instrument for recording whistles.

The original TPOD.exe train classifier was a 'positive' system that aimed to identify trains like porpoise trains.

This was changed to a 'negative' system that aimed to identify trains that were unlike noise trains. This had the advantage that it did not prejudge acceptable train descriptors and improved performance on a difficult species – the Dusky Dolphin. It also showed up some previously unknown trains sources – e.g. WUTS – more clearly.

The v1 CPOD.exe classifier was based on the classifier developed for the T-POD, but gained discriminatory power from the better click descriptions (T-PODs only stored the click duration!).

The KERNO classifier improved the performance by changing part of the analysis to non-parametric methods to reduce the effect of outliers (clicks that appear to be part of a mathematically defined train, but actually come from a different source). The most useful *not-noise / train sources* boundary is set on the Q class scale at *Mod / Low* but in quiet locations the *Low* and *doubtful* trains may also be cetacean trains, while in the toughest locations it may be necessary to exclude the *Mod* trains. The Q classes can be viewed as different thresholds along a Receiver Operating Characteristic (ROC) curve, but these curves do not apply correctly to this system – a different curve would be needed for a different pattern of ambient noise.

Should I use PODs or hydrophones?

C-PODs contain hydrophones but do not retain digitised sound because it requires too much memory and battery power. Instead they do initial analysis in real time to find tonal events and store summary data on each.

If you wish to study the acoustic details of sounds made by cetaceans you need a hydrophone, amplifier, fast analogue-to-digital converter (ADC) that samples at least three times the highest frequency of interest and can pass data to your computer, and an anti-aliasing filter. Then you need appropriate tools for navigating the data and analysing it.

If you wish to study the presence and behaviour of cetaceans you could use:

Hydrophone, ADC etc as above, or in a complete data capture package. Then you need to find a way of analysing those data. No software currently available gives false positives as low as the C-POD without substantial loss of sensitivity. Data volumes are huge for systems that record continuously – for porpoise studies around 1TB per 10days. Analysis, when achieved, is typically very time-consuming, but access to full waveforms is nice. They are not as valuable as you might hope because weaker clicks are less typical as they contain more noise.

Hydrophone array: this approach requires more hardware but can help with the false positive/sensitivity trade-off by searching for groups of clicks coming from a common bearing by using Pamguard / Rainbow Click software. Autonomous systems for deployment over months are not yet demonstrated.

C-PODs: Currently the easiest tool to use, which also gives high sensitivity + very low false positive rates, but no waveforms are available for inspection or analysis. Extensive testing in many situations has shown reliable results.

C-POD-F: This development from the C-POD will capture 'voucher' specimens of louder waveforms, have various other improvements and lower power processing of cetacean clicks.

CPOD.exe program notes

CPOD.exe is written in Delphi 7, a version of Pascal. Logical constructs in Pascal match spoken English very well (unlike Visual Basic), so the code is easy to read for a self-taught programmer.

CPOD.exe's memory footprint is modest, and small when not processing files to detect cetaceans. The train detection process may be slowed if available system RAM is less than 300MB.

The program is optimised for speed in handling large files, but it does not use multiple threads so that it does not consume too much of your system processing power. However multiple copies of CPOD.exe can be run so that you can process files in the background. In effect you have user controlled multi-threading. The speed of viewing and exporting data may be limited either by processor speed or hard drive reading speeds.

Various parameters are stored and processed as integers. This does produce curious patterns in some X/Y scatter graphs.

Acoustic test tank notes

Tank testing is tricky! Inter-POD comparisons at sea are generally more powerful, reliable, informative and useful, but require deployments long enough to capture many hundreds of cetacean trains. Tank testing is quicker. It can be accurate or can just provide assurance that instruments are still working normally. There have been two studies that have shown that tank tests do broadly correspond to performance at sea for porpoises.

Tank testing notes:

- tanks are mostly subject to significant reverberation and testing has to deal with that by confining the analysis to the period before the first echo arrives. Echoes can be louder than the direct path signal and if this is the case the tank is not suitable for quantitative testing.

- sound sources vary in their radial output and an accurate measurement of the sound field must be made at the POD hydrophone position and must take into account the radial variation in sensitivity of the calibration hydrophone as this may be comparable with the level of standardisation of the POD.

- tests of reproducibility of results are essential.

- PODs must be soaked for 24 hours before testing.

- The temperature of the POD hydrophone must be known.

- The radial variation of the POD must be tested using at least 8 positions, preferably 32.

Clean signal time

Any tank has a 'clean signal time' that is the difference in arrival time at the receiver between a signal arriving by a direct path and the earliest echoes (reflections) of that signal.

For a tank 2.5m in diameter with a source 25 cm below the surface and the receiver 1m away at the same depth, the clean signal time is 42 microseconds. Then the surface echo starts to arrive and interfere with the directly received signal if it is continuing. If the signal is a pulse at 130kHz it will be continuing if the pulse produced by the acoustic transmitter is more than 5 cycles long. If the horizontal separation is reduced to 10cm the clean signal time increases to 30 cycles at 130kHz.

At the end of the clean signal time the receiver is in an interference field of echoes that evolves over time in a complex pattern determined by all the acoustic surfaces and properties of the tank and its environment. It is extremely difficult to work with this interference or echo field because -

- at points during the time evolution of a signal it may be more intense than the directly received signal.

- at any time the intensity varies strongly between points that may be only half a wavelength apart. For a 130kHz signal the wavelength in water is 12mm. The accuracy of positioning of the POD transducer in the small cylinder that contains it is plus or minus 1mm, so this could be positioned consistently within the sound field, but the field itself will only be reproducible if the water depth, source position etc are comparably consistent, and that is much harder to achieve.

To maximise clean signal time, put the source near the centre of the tank, and put the receiver near to it, but outside the near field – say >10x the diameter of the source. In a good tank the C-POD will log 4000 clicks per second correctly. From electronic sources (bypassing the transducer) the rate is much higher.

If any echo is louder than the direct signal, which is a common situation, this will be the last detected signal as the transmitted intensity is reduced to find the threshold. In that case the clean signal time is irrelevant because detection is being made later in the echo field. At this threshold level the POD data will show a nice clean record of single detections at the transmitted rate, but do not be deceived!

C-PODs, unlike T-PODs, are not calibrated to produce a standard threshold, but to give an accurate measurement of the sound pressure level of a 130kHz signal.

Electromagnetic and electrostatic interference

RF interference can be huge nuisance – it may appear as anything from big bursts of clicks to variability of the duration of logged clicks, and it varies over time. The commonest source is laptop screens. RF interference penetrates water to tank depths but is reduced if the water is saline. Screening the POD within a steel cylinder produces striking improvements in the evenness of the durations of detected acoustic signals.

Electrostatic coupling between the source and receiver can occur over surprisingly long distances because of the dielectric properties of water and may cause puzzling effects when you are exploring the sounds in your tank. A hydrophone then shows a perfect replica of the transmitted signal without any detectable time delay, and without all the distortions that are found in the acoustic transmission route. Move the receiver or transmitter around and no change in timing occurs (this assumes you have the source drive signal as a trigger for capturing the waveform).

Test signal

Two types of signal are in use: square envelope and raised cosine envelope.

Many signal generators can produce square envelope (simple bursts) pulses, and they are satisfactory, but the ends of the signal will show high levels of distortion with high frequency components emerging. So the duration of the detectable signal will be lower than the length of the burst. An 8cycle bursts is as short as you should go.

Raised cosine envelopes are more like a porpoise, and avoid the distortion problem, but a smaller fraction of the burst will be above the threshold when a weak signal is used.

Setup process:

1. Use a 10 cycle square burst of sine waves at 130kHz repeated at intervals of 0.1s and adjust the output power downwards until they are logged at around 50 on the SPL scale. This will be your test signal amplitude.

2. Adjust the vertical position of the POD or source to find the level giving the maximum recorded amplitude. This will be in the horizontal region at the top of the curve of response against vertical height and will be the position most tolerant of small positioning errors.

3. Use an output power that is 4x this and look at the data at the highest time resolution to check that the first detected click from each burst is louder than the subsequent echoes (if any). It not you are working in a badly echoic sound field and will not get good results until that is remedied, which may not be easy.

4. Revert to the test signal amplitude and reduce the pulse duration until the resulting detections become uneven. Use 8 cycles if this is above that level, or more if necessary. This is the test signal duration.

5. Use higher repetition rates in steps (doubling) until the logged amplitude starts to change – echoes are now interfering with the following signal. Go back to a slower rate (say 4x interval) where this does not occur. This is the test signal repetition rate.

6. Test each POD at a minimum of 8 positions radially.

Acknowledgments