
The F-POD
What it does,
and how…
Nick Tregenza
Exeter University
Chelonia Limited
Feb 2020

What it does

- The F-POD is an applied acoustics tool that aims to make acoustic monitoring of cetaceans sensitive, accurate, efficient and cheap.
- It logs cetacean clicks above 17kHz, so its use is confined to odontocetes smaller than the Sperm Whale.
- To monitor all cetacean sounds a combination of an F-POD and a wav-file recorder can be good – the wav file recorder can then be set to a relatively low sampling rate (e.g. <50k samples/s) to extend its running time, while the F-POD runs at 1m samples/s and can do so for 4months.

The monitoring challenge

Date from monitoring for more than a few weeks requires high-performance automated analysis to avoid the need for human analysis, because human analysis is :

1. very costly in time

2. dangerously subjective, to the point where the highest detection rate periods from some analysts are entirely rejected as false positives by others.

Click data is inherently difficult because:

- Perfect examples of the most distinctive cetacean clicks sometimes come from non-cetacean sources, and sometimes in large numbers.
- The least distinctive cetacean clicks – short dolphin clicks – have no features that distinguish them from clicks from a myriad of sources in the sea, except being very loud at the dolphin that produced them.
- Continuous recording uses a lot of memory e.g. 1TB / 10 days

There are several potential solutions. All involve the evaluation of groups of clicks:

The T-POD, C-POD, and F-POD seek to identify the *trains of clicks* made by cetaceans.

In trains there is a strong resemblance of clicks to the click before and after in the same train.

Train identification also has challenges:

- **Click train identification:**
- **There are more 5-click sequences in 10000 clicks than there are seconds since the origin of the universe**
- **so many chance sequences of features will occur.**

The F-POD approach is:

Select clicks that are either long and tonal, or loud, or meet some intermediate criteria. A long click is defined as one with 5 cycles of similar wavelength (approx. +- 25%). If the click is only 4 cycles long it must meet a weak amplitude criterion that can be controlled by the user, (although making changes without extensive study of their effects is unwise). Shorter clicks must meet higher amplitude criteria.

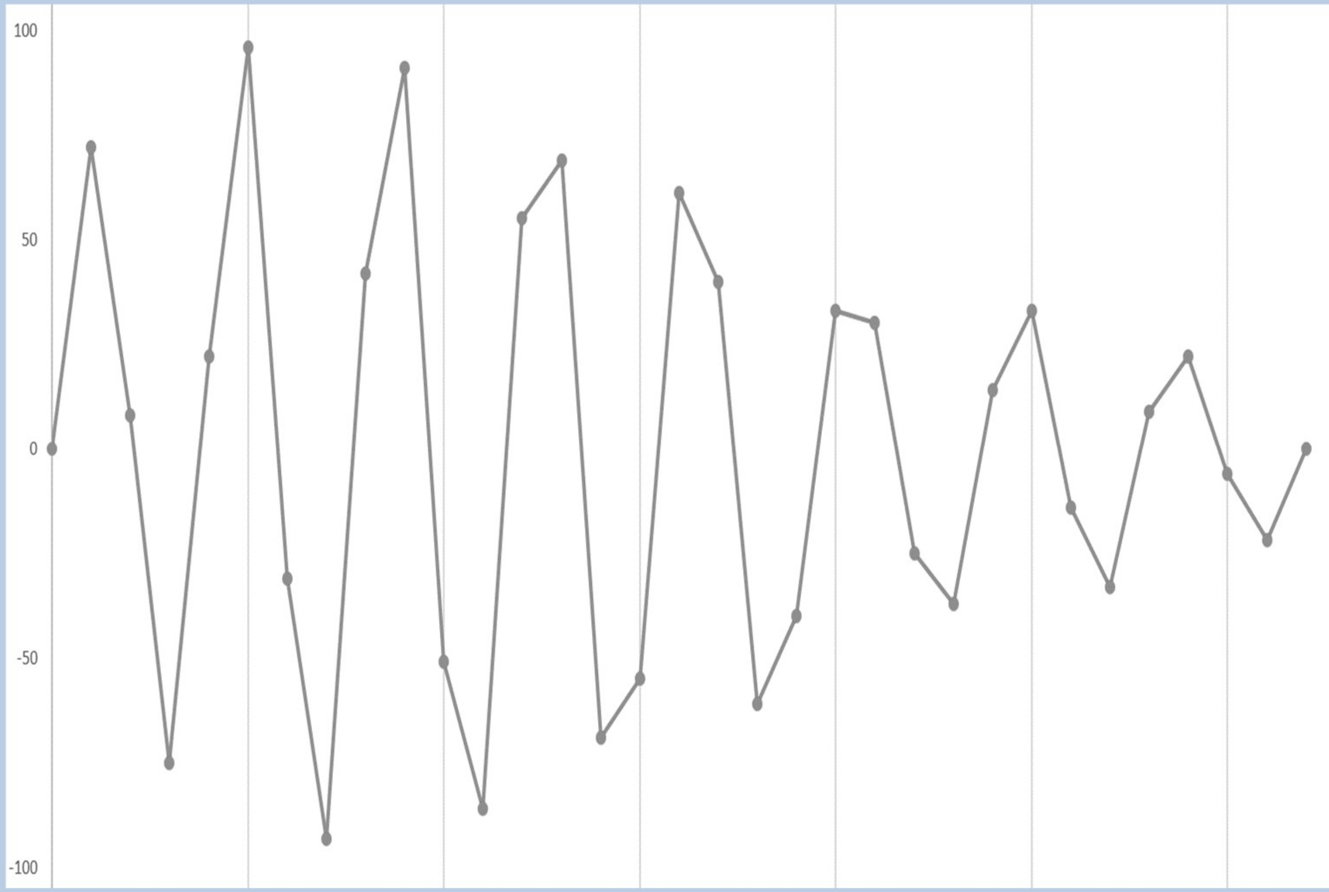
- These selection criteria keep data volumes down.
- Other weaker criteria reject very early reverberation and very high bandwidth clicks.

Capture ultra-high resolution time domain data. Sound is sampled at 1 m/s, then upsampled to 4 m/s, then low pass filtered using a binomial filter, then characterised in real time at the new virtual sampling interval of 250 ns.

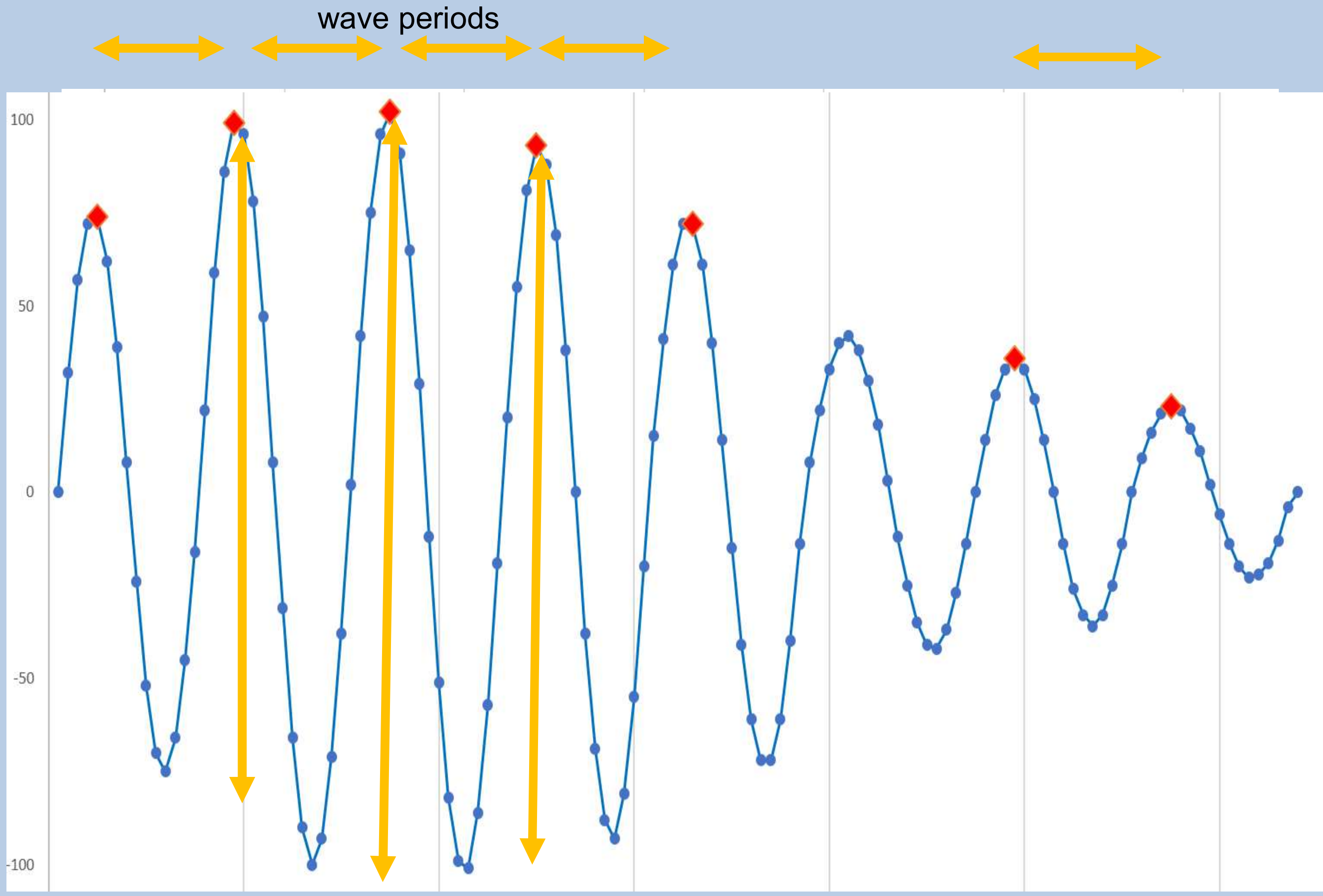
These data enhance the power to detect trains.

the F-POD a waveform recorder :

Data is sampled at 1 microsecond intervals:



… and up-sampled in the sea to 4 million samples per second. This gives:



◆ new more accurate peak positions

12 features logged:

Duration

N of cycles

Five wave periods

Three amplitudes

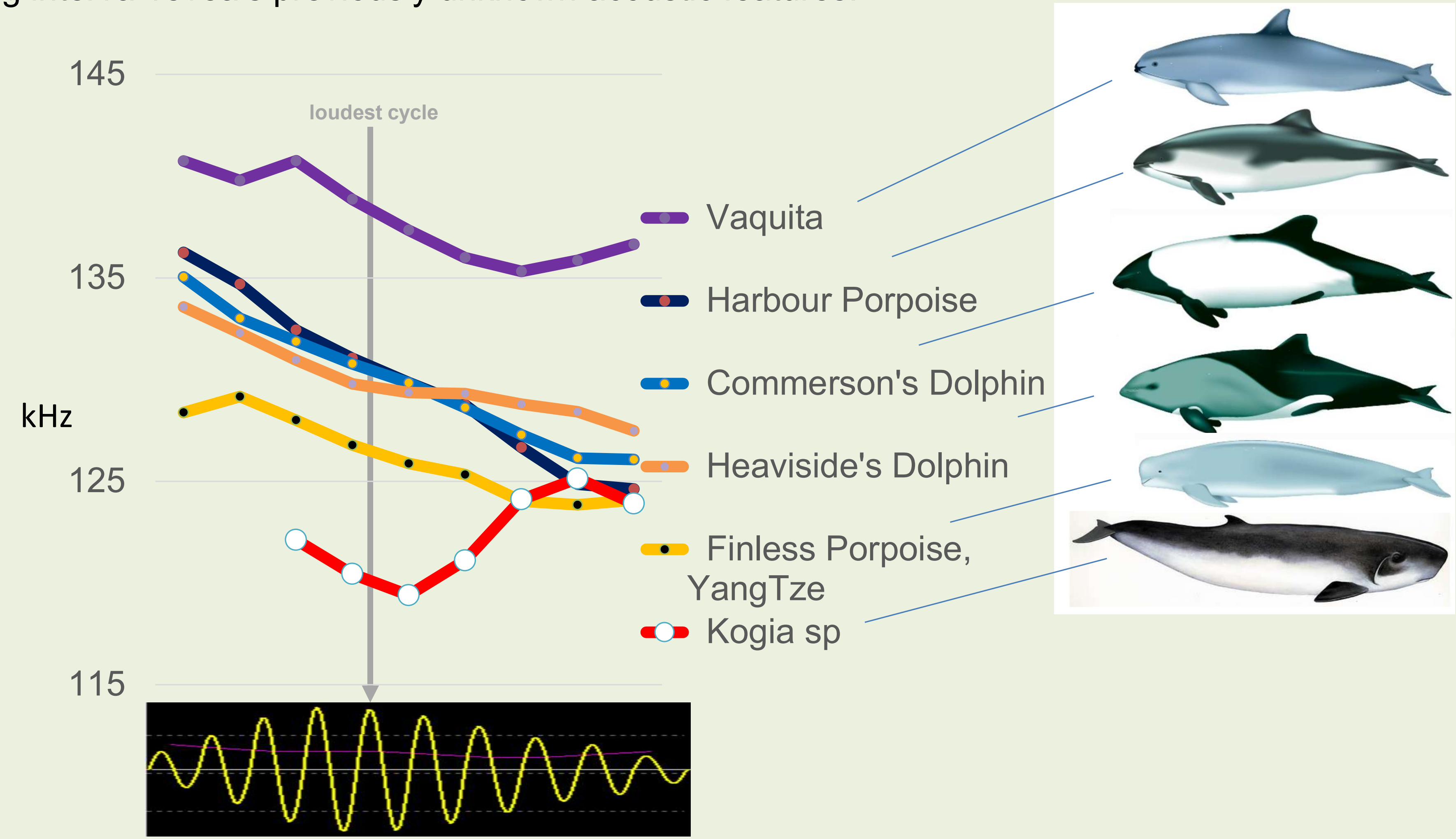
N of amplitude inflections

Period range

+ time of click

Upsampling and low-pass filtering has corrected the wave number of the loudest cycle, and the number of amplitude inflections.

More detail is captured from some selected clicks.
The short sampling interval reveals previously unknown acoustic features:



NBHF clicks are downsweeps *Kogia* is acoustically different from other NBHF species

Post-processing on the PC – the KERNO-F classifier

- **‘Chance trains’ that are purely chance sequences of clicks that happen to match the characteristics of true trains are the biggest problem – they occur often because there are so many opportunities.**

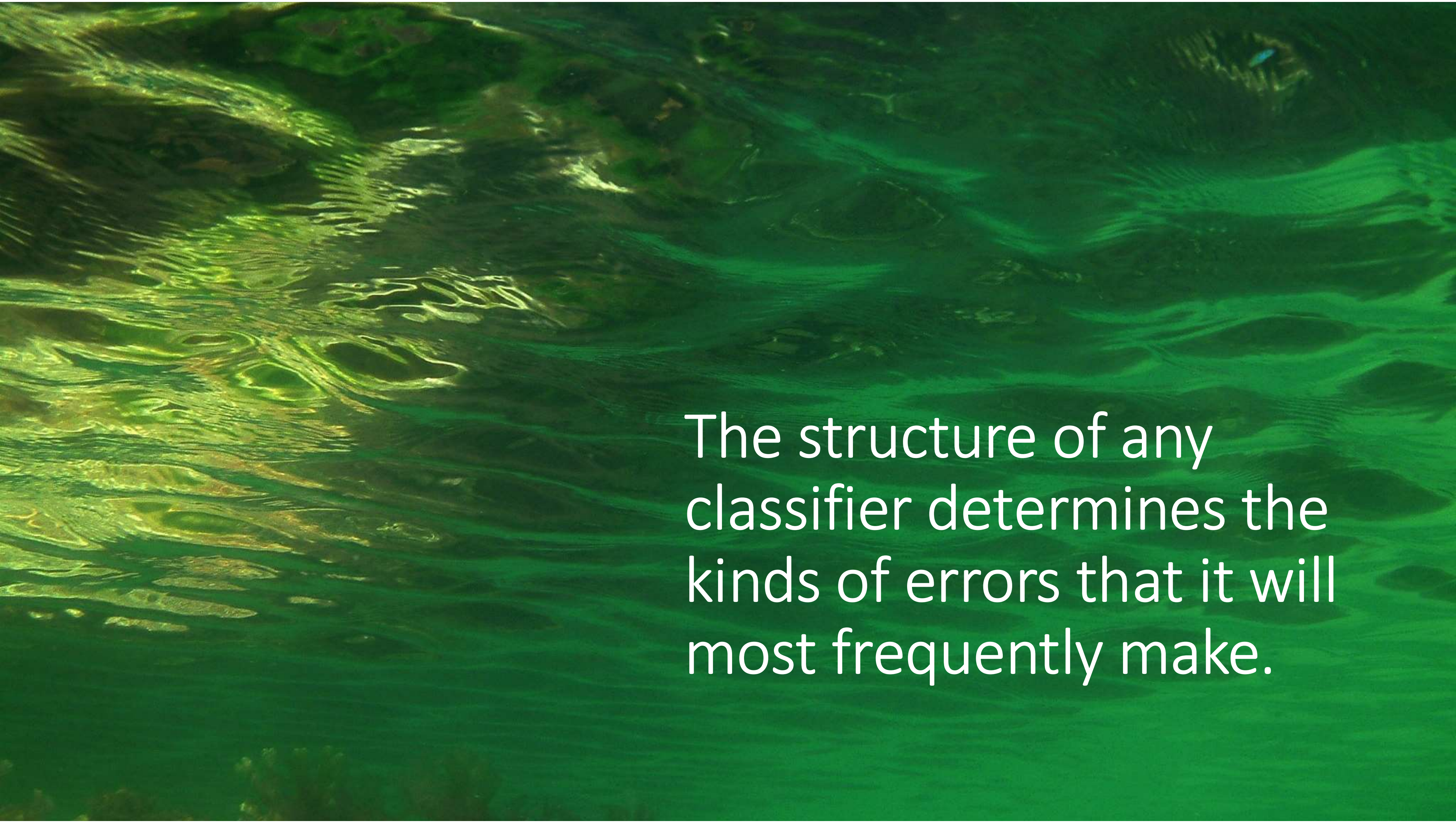
KERNO-F process:

Steps:

- 1. Finds trains of similar clicks at more or less regular intervals, but without applying all the possible tests of similarity that can be used given the high-resolution data captured. This uses nested multiple hypothesis testing.
- The reason for that limitation is that it allows ‘wrong’ trains to be selected from the multiple options by not forcing selection of the 'best mimic’, and still leaves some criteria with which to test whether that train is a true train. This can be viewed as a defence against ‘overfitting’ by a classifier. Is this a novel development in case-by-case classification?
- 2. Measure the ‘coherence’ of the train. Coherence is the similarity of successive clicks and intervals in the train. It is typically high in true trains (sequences from the same source) and low in chance trains (sequences from varied sources) .
- 3. Apply threshold criteria on train coherence. This gives the high, moderate, low and doubtful 'quality' classes in the software. They address the question: How confident are we that this train came from a train source? So boat sonars and trains from weak unknown train sources could and should also get high Q values.
- 4. Apply weak feedback within single clock minutes. Isolated weak trains are down-classed, and vice-versa. This process is not carried across minute boundaries as that would mean that a part of the file could give different results if analysed in isolation. A difficult choice: stronger feedback could be valuable, but it can be applied later in the form of an encounter classifier if needed, like the very efficient Hel1 classifier for Baltic C-POD data.

Design implications of train detection:

- **The train detection process determines what data is useful as an input. That becomes the basis for the click selection. The design has been recursive because of this. We have found that:**
- **Fine time-domain click features, like the wave-number of the loudest cycle of the click, may have little value in determining the species or source, but are valuable in measuring train coherence.**
- **Weak clicks have a proportionately larger content of ambient noise. This cannot be removed because the click is too short to do that in any useful way. So the coherence measures go down, and weak trains are lost if the background is not very quiet. Having such a sensitivity to ambient noise is not useful in a monitoring role. Further, in very quiet conditions more weak unknown train sources are picked up and reduce the overall reliability.**
- **So the most sensitive instrument settings may not be worth using, even if they do give higher true positive rates.**
- **The long NBHF (narrow band high frequency) clicks made by porpoises and some dolphins contain more information in each click and are consequently easier to work with when weak. So this makes the lack of any amplitude criterion for long clicks appropriate – the lower detection limit is then more often the system noise level and is consequently stable.**
- **Bringing enough data back to shore to enable an understanding of the processes to be developed is valuable, and possibly essential, but an alternative system that uses embedded artificial intelligence running in real time with transmission of only the output is currently in development by Newcastle University.**


The structure of any classifier determines the kinds of errors that it will most frequently make.

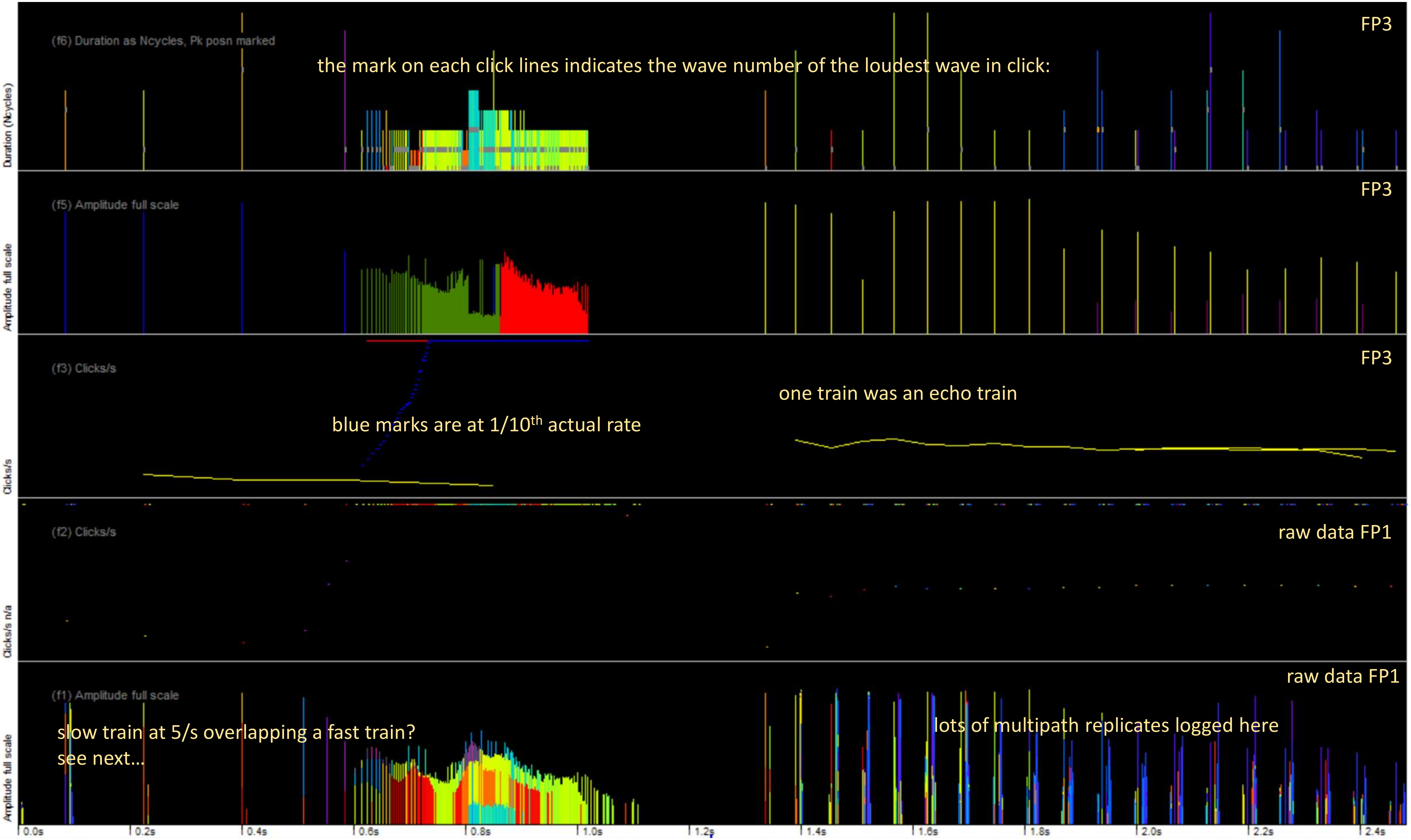
The F-POD + KERNO-F error susceptibilities are:

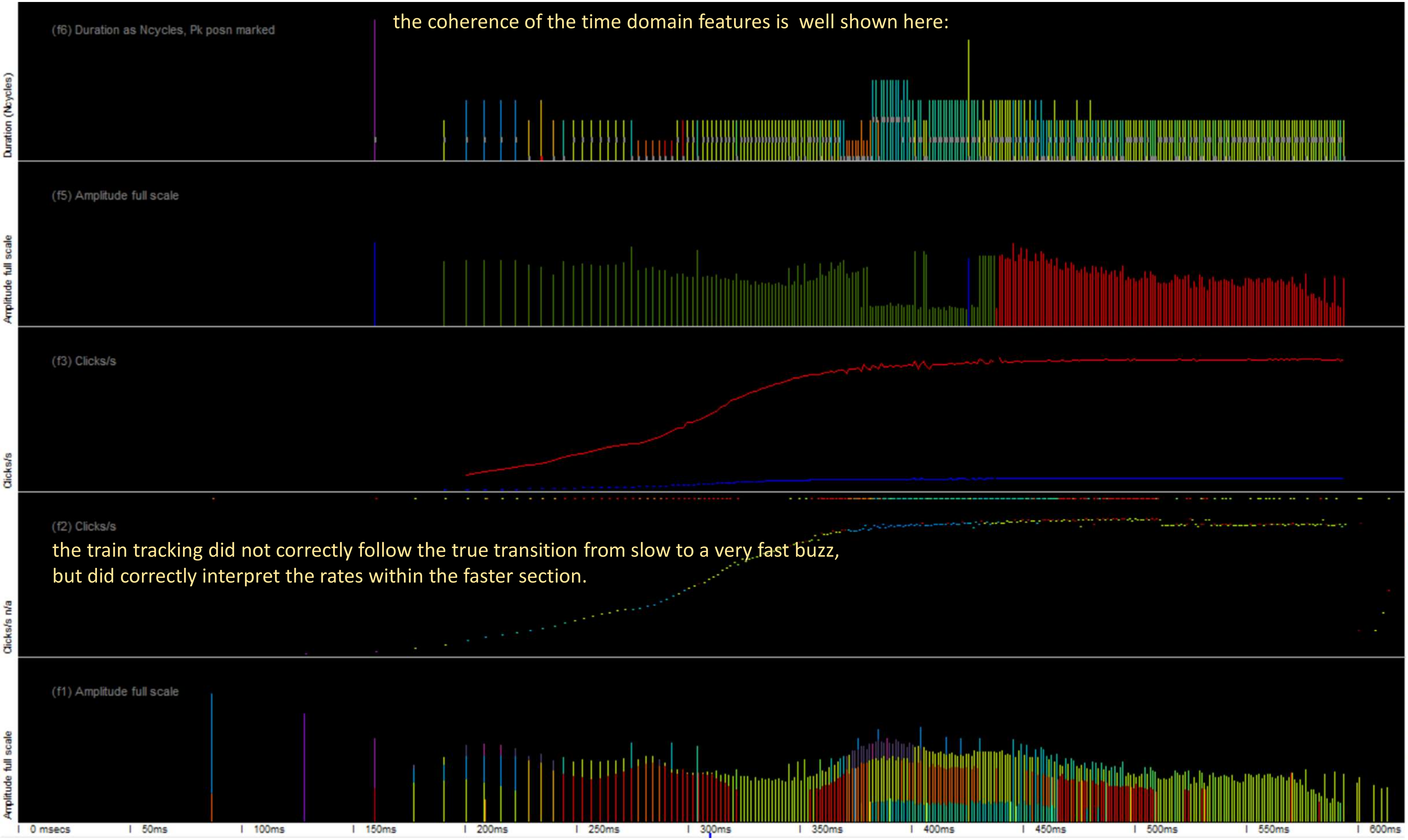
- Trains that are more irregular in time or acoustic features are more likely to be missed. For example, if Orcas produced click trains with irregular intervals to make them less recognisable to prey, they would also be less recognisable to the F-POD.
- Trains that have very long inter-click intervals are more likely to be missed because many more chance trains will overlap them.
- Weak trains i.e. low amplitude, are more likely to be missed (see above).
- Increasing ambient click rates (all sources) will reduce the detectability of weak clicks and the extraction of true click trains.
- Because there is overlap between the species guilds there is a rate of false positives in any class from other classes. This is reduced by a feedback processat the soct of a negative interaction with other train sources.
- Silent animals will be missed.

Validation

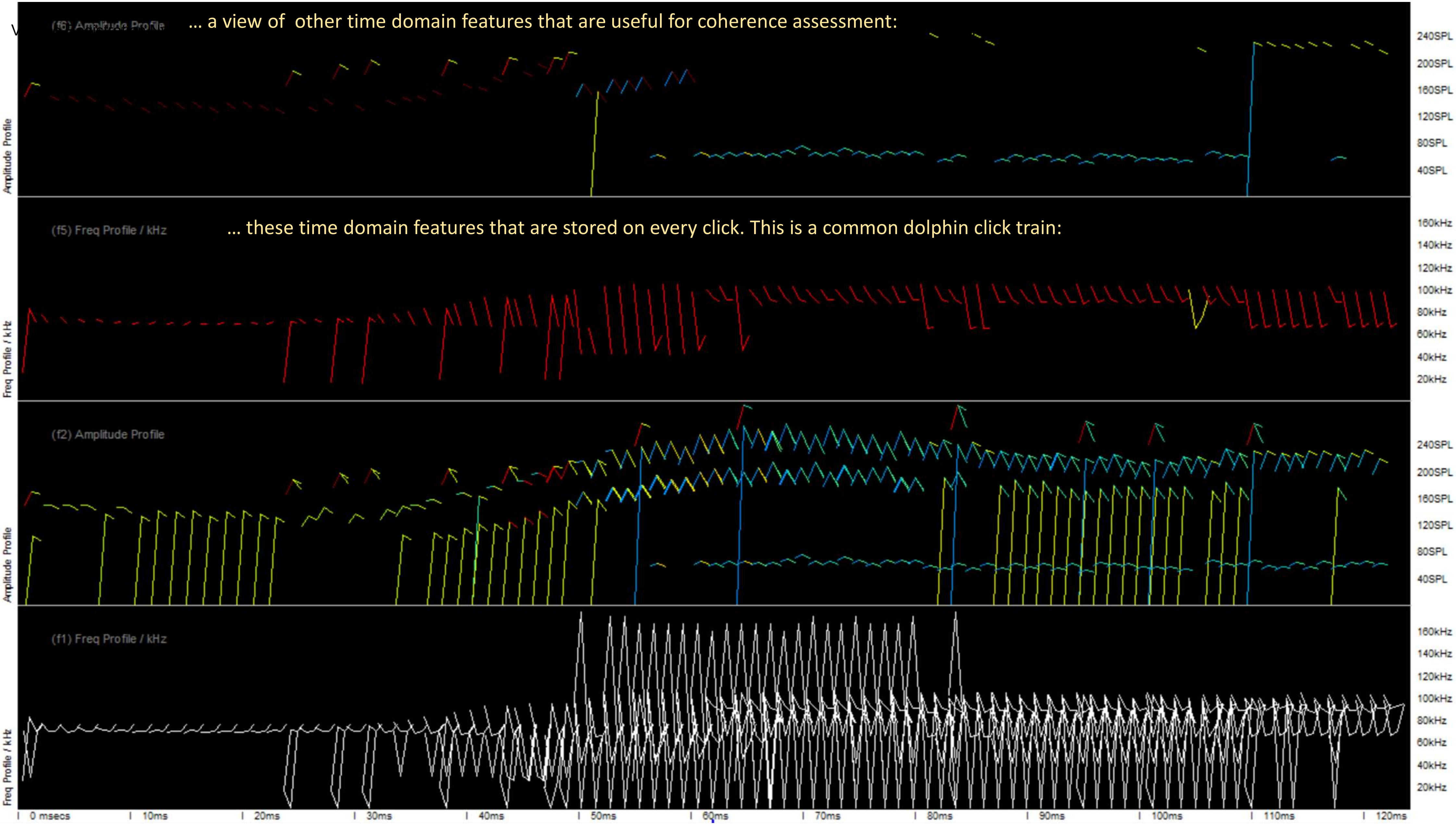
- Complex algorithms cannot be validated by inspecting the code. Empirical tests are essential.
- The 'gold standard' is comparisons of acoustic detections with other evidence of the presence of cetaceans e.g. sightings.
- that's hard, and intermediate validation is useful:

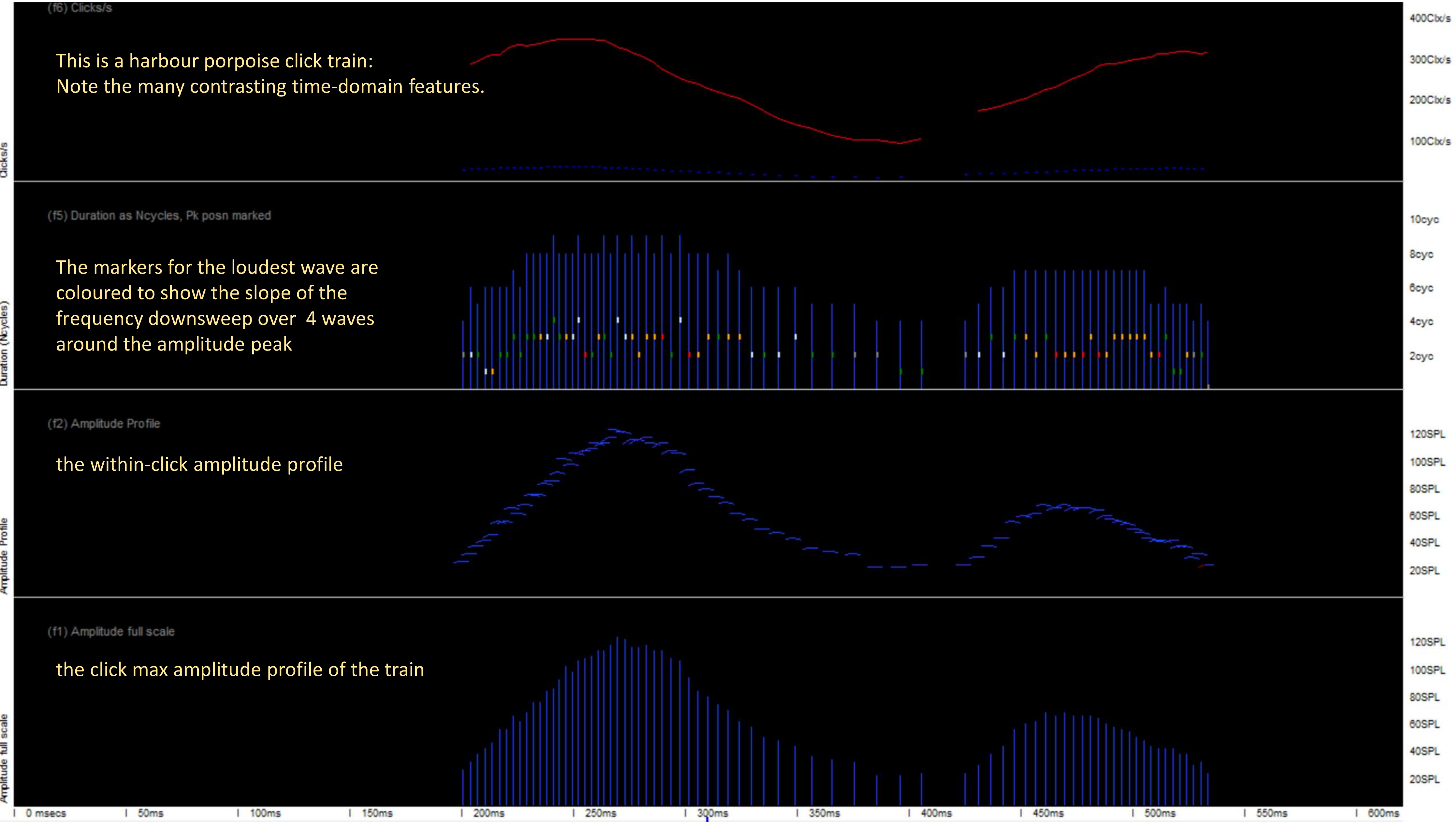
- A measure of the clustering of the Q classes of trains is useful and is incorporated in the POD software.
- Visual checking of the accuracy of identification of the sequence of clicks attributed by the algorithm to the same train is enabled:


FP3
(f6) Duration as Ncycles, Pk posn marked
the mark on each click lines indicates the wave number of the loudest wave in click:
Duration (Ncycles)
6cyc
4cyc
2cyc
FP3
(f5) Amplitude full scale
Amplitude full scale
400SPL
300SPL
200SPL
100SPL
FP3
(f3) Clicks/s
one train was an echo train
blue marks are at 1/10th actual rate
Clicks/s
40Clx/s
30Clx/s
20Clx/s
10Clx/s
(f2) Clicks/s
raw data FP1
Clicks/s n/a
40Clx/s
30Clx/s
20Clx/s
10Clx/s
raw data FP1
(f1) Amplitude full scale
slow train at 5/s overlapping a fast train?
see next…
lots of multipath replicates logged here
Amplitude full scale
400SPL
300SPL
200SPL
100SPL
0.0s
0.2s
0.4s
0.6s
0.8s
1.0s
1.2s
1.4s
1.6s
1.8s
2.0s
2.2s
2.4s



the coherence of the time domain features is well shown here:
(f6) Duration as Ncycles, Pk posn marked
Duration (Ncycles)
6cyc
4cyc
2cyc
(f5) Amplitude full scale
Amplitude full scale
400SPL
300SPL
200SPL
100SPL
(f3) Clicks/s
Clicks/s
800Clx/s
600Clx/s
400Clx/s
200Clx/s
(f2) Clicks/s
the train tracking did not correctly follow the true transition from slow to a very fast buzz,
but did correctly interpret the rates within the faster section.
Clicks/s n/a
800Clx/s
600Clx/s
400Clx/s
200Clx/s
(f1) Amplitude full scale
Amplitude full scale
400SPL
300SPL
200SPL
100SPL
0 msecs
50ms
100ms
150ms
200ms
250ms
300ms
350ms
400ms
450ms
500ms
550ms
600ms


(f6) Amplitude Profile
… a view of other time domain features that are useful for coherence assessment:
240SPL
200SPL
160SPL
120SPL
80SPL
40SPL
Amplitude Profile
(f5) Freq Profile / kHz
… these time domain features that are stored on every click. This is a common dolphin click train:
160kHz
140kHz
120kHz
100kHz
80kHz
60kHz
40kHz
20kHz
Freq Profile / kHz
(f2) Amplitude Profile
240SPL
200SPL
160SPL
120SPL
80SPL
40SPL
Amplitude Profile
(f1) Freq Profile / kHz
160kHz
140kHz
120kHz
100kHz
80kHz
60kHz
40kHz
20kHz
Freq Profile / kHz
0 msecs
10ms
20ms
30ms
40ms
50ms
60ms
70ms
80ms
90ms
100ms
110ms
120ms


(f6) Clicks/s
This is a harbour porpoise click train:
Note the many contrasting time-domain features.
400Clx/s
300Clx/s
200Clx/s
100Clx/s
Clicks/s
(f5) Duration as Ncycles, Pk posn marked
The markers for the loudest wave are coloured to show the slope of the frequency downsweep over 4 waves around the amplitude peak
10cyc
8cyc
6cyc
4cyc
2cyc
Duration (Ncycles)
(f2) Amplitude Profile
the within-click amplitude profile
120SPL
100SPL
80SPL
60SPL
40SPL
20SPL
Amplitude Profile
(f1) Amplitude full scale
the click max amplitude profile of the train
120SPL
100SPL
80SPL
60SPL
40SPL
20SPL
Amplitude full scale
0 msecs
50ms
100ms
150ms
200ms
250ms
300ms
350ms
400ms
450ms
500ms
550ms
600ms

Species identification

- At present the F-POD software classifies click trains to **guilds:**
- 'NBHF' = porpoises and some dolphins that make NBHF-type clicks.
- ‘Other cetaceans’ = all those species that produce short clicks ‘broad-band transients’.
- Boat sonars.
- Unclassed – these mostly belong to one of the other guilds but cannot be identified as such with sufficient confidence.

WUTS

- WUTS may be small crustaceans, molluscs grazing on the hydrophone housing, stress noises from wooden structures, …or …?
- It is clear there are varied Weak Unknown Train Sources, and they have a very wide range of train and click features. So WUTS are handled as a species-reliability-risk-factor for both individual trains and whole minutes, but they are not identified as such by KENRO-F version 1.
- Their presence in large numbers has been seen a few times and these files show an excessive proportion of 'unclassed' species trains, and of very fast click trains.
- Files with a significant proportion of WUTS generally have more than 3 classification warnings i.e. a risks sum of 4 or more.

Feedback / interaction

- Many dolphins sometimes produce NBHF clicks, and boat sonars can, after multiple reflections can resemble cetacean clicks, as can sediment transport, WUTS and other noise soruces.
- So a feedback stage can sometimes reclassify clicks in a moving window of 11 minutes if trains fitting multiple sources are present.
- This interaction reduces the false positive problem that could, in effect, prevent the extinction of a species subject to such feature overlap from ever being acoustically detected by fully independent classifiers.
- The Hel1 classifier used this approach to obtain a false postive rate of < 1 false second per year for porpoise detections.


Thank you for
your attention
Comments and questions are very
welcome. Please contact
nick.tregenza@chelonia.co.uk
All pictures from Mount's Bay, SW Britain