

An exploratory analysis of non-Poisson temporal behaviour in snapping shrimp noise

Legg, M. W. (1)*, Duncan, A. J. (2), Zaknich, A. (3) and Greening, M. V. (4)

(1) MOD, Defence Science and Technology Organisation, HMAS Stirling, Australia

(2) CMST, Curtin University of Technology, Bentley, Australia

(3) School of Engineering Science, Murdoch University, Rockingham, Australia

(4) MOD, Defence Science and Technology Organisation, Edinburgh, Australia

ABSTRACT

Snapping shrimp are a well known interference source for underwater sonar and communication systems, particularly in shallow and harbour waters. The noise produced by snapping shrimp is highly impulsive and the amplitude statistics are non-Gaussian. Impulsive noise is most often modelled in a way that implicitly assumes that the temporal statistics are Poisson. The Poisson assumption implies that a snap from any shrimp is completely independent of snaps from other shrimp. This paper reports on an exploratory analysis of non-Poisson temporal behaviour in snapping shrimp noise using real acoustic data from different geographic locations in Australian coastal waters. The analysis makes use of various statistical techniques applied to snaps detected in high-pass filtered data using a threshold technique. Attempts are made to eliminate multi-path effects, which can introduce correlations between snap arrivals, from other possible effects such as interactions between shrimp. The results are compared and contrasted between different geographic locations.

INTRODUCTION

In shallow and harbour waters one of the most dominant non-Gaussian noise sources that contributes to the persistent ambient acoustic noise is the snapping shrimp (Cato and Bell, 1992). Snapping shrimp are small shrimp, usually only a few centimetres long, characterised by a single enlarged claw used for producing a loud impulsive snap. The action of closing the claw shoots a jet of water out of the claw and a cavitation bubble is produced. It is the collapse of this cavitation bubble that results in a loud snap (Versluis et al., 2000). Individual peak-to-peak sound pressure levels (source levels) of snapping shrimp have been reported in excess of 180 dB *re*: 1 μ Pa @ 1 m in a controlled environment (Au and Banks, 1998) and in excess of 187 dB *re*: 1 μ Pa @ 1 m from *in situ* measurements in an underwater environment (Ferguson and Cleary, 2001). In addition to very high source levels, the snap from snapping shrimp is also extremely broadband extending from below 1 kHz (Miklovic and Bird, 2001) to above 200 kHz (Cato and Bell, 1992). The high source level and broad spectrum of snapping shrimp noise is particularly important to underwater acoustic applications that are expected to operate in shallow waters. The effect of such noise on spread-spectrum acoustic communications systems provides the motivation for the work presented in this paper.

It has been shown that the amplitude statistics of snapping shrimp noise are non-Gaussian (Bertilone and Killeen, 2001). Models used to describe non-Gaussian amplitude statistics often assume the temporal statistics to be Poisson, however there is anecdotal evidence to suggest that the temporal statistics of snapping shrimp noise may be non-Poisson. Cato and Bell (1992) report bursting of the higher level shrimp snaps, and hypothesise that shrimp may snap in response to other snaps as a survival mechanism. Such interaction between shrimp would result in non-Poisson temporal statistics. Potter and Chitre (1999) provide evidence of structure in both time and space, but point out that it is difficult to isolate if it is due to spatial or temporal effects (Potter, Lim and Chitre, 1997). Spatial anisotropy and

clumping are reported from studies with directional hydrophone fields (Potter and Koay, 2000) but the question of non-Poisson temporal behaviour of snapping shrimp is not concluded and remains a challenging problem.

Shrimp snaps are highly localised events occurring apparently at random within the ambient noise therefore displaying the characteristics of a random point process (Snyder and Miller, 1991). Other examples of point processes are emissions from radioactive decay, neuron firings, shot noise in optics, seismic events and lightning strikes. Processes having intervals between events that are independent and identically distributed are termed *renewal processes* (Cox and Lewis, 1966). The Poisson process is a renewal process with the added constraint that there is no trend in the series. Real world processes are unlikely to be Poisson processes but may still be renewal processes. This paper reports on a histogram test for Poisson statistics applied to three real acoustic data sets and one simulated data set.

DATA AND PRE-PROCESSING

The acoustic data used for the analysis were sourced from recordings in Spencer Gulf (SGF) South Australia, Seal Island (SEAL) Western Australia, and Feather Reef (FR) Queensland. The Spencer Gulf data was recorded using a hydrophone placed mid-water in a water column of unknown depth. The Seal Island data was recorded using a hydrophone near the surface in a very shallow water column of between 3 and 5 metres, near to a set of reefs. The sea state was very low and winds calm. The Feather Reef data was recorded using a bottom mounted hydrophone in 24 metres of water. The sea state and wind conditions are not known for the Spencer Gulf and Feather Reef data. It was assumed in all cases that the shrimp were located on the sea floor.

A simulated data set (SIM) was created by randomly distributing zero mean, high amplitude Gaussian distributed points to simulate "snaps", and then adding zero mean Gaussian background noise. The amplitude of the snaps had 50 times the standard deviation of the background noise, and

were randomly distributed so that the actual snap locations were drawn from a Poisson process. The simulated data set was used to provide a baseline for guarding against artefacts introduced by the data processing.

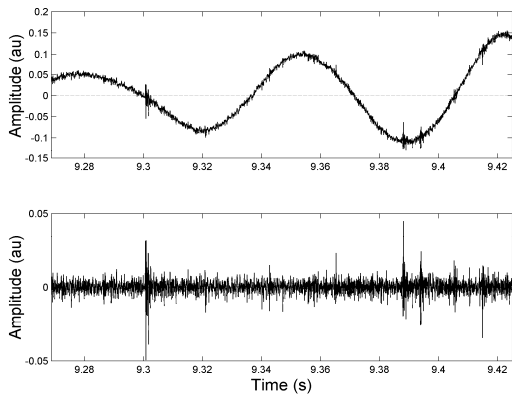


Figure 1. Raw (upper) and high-pass filtered (lower) versions of the same section of Spencer Gulf data. The high-pass filter was configured to remove energy below 1kHz.

All of the acoustic data sets were recorded at sonic frequencies. The Spencer Gulf and Feather Reef data sets were recorded onto DAT using a sample rate of 32kHz and long play option. The Seal Island data was recorded onto DAT using a sample rate of 41kHz and standard play option. The two data sets recorded at 32kHz contained substantial amounts of energy below 1kHz, which is below the frequency that snapping shrimp energy becomes significant. A 2^{14} point FFT-FIR high-pass filter, with stop band at 400Hz and pass band at 900Hz, was applied to all data sets, including the simulated data, to remove the unnecessary low frequency components. Figure 1 shows the difference between raw and high pass filtered data for a representative section of the Spencer Gulf data set.

Shrimp snap detection

A thresholding technique was used for the detection of shrimp snaps. An estimate of the standard deviation (σ) was calculated using the first ten thousand timeseries samples. Threshold levels were set as multiples of σ and detections declared when the amplitude of the timeseries exceeded the value of the threshold. Modified thresholding techniques can be used to overcome difficulties when processing real data. Common modifications include adaptive estimation of the background noise level, see for example Watkins et al. (2004), and insertion of dead time following detection to allow for relaxation.

Visual inspection of the filtered timeseries showed that the need for adaptive estimation of the background noise level was minimal for the filtered timeseries. Adaptive estimation of background noise levels was not used for snap detection. A dead time after detection was used for snap detection because each snap was followed by a significant amount of oscillation. Figure 2 shows a large amplitude snap after which there is a period of oscillation that extends at least 3 ms beyond the main impulse of the snap. An exponential decay curve was fitted to the envelope of the snap and the relaxation time was calculated as 1.2ms. The exponential decay and relaxation time are also shown on Figure 2. High sample rate recordings of snapping shrimp clicks by Au and Banks (1998) show relaxation within 100 μ s of the main impulse, which is an order of magnitude shorter than the relaxation time observed in the data. The long relaxation times observed are therefore attributed to effects other than

the original source snap, such as the impulse response of the environment and the recording system.

In the analysis, described later, the dead time was chosen by visual inspection of large snaps. At the time of relaxation after the large snaps, the oscillations had settled to a significantly lower level than the anticipated detection threshold of 3 standard deviations. The high detection threshold allowed a dead time to be chosen that was slightly shorter than the longest relaxation time found in all of the acoustic data sets.

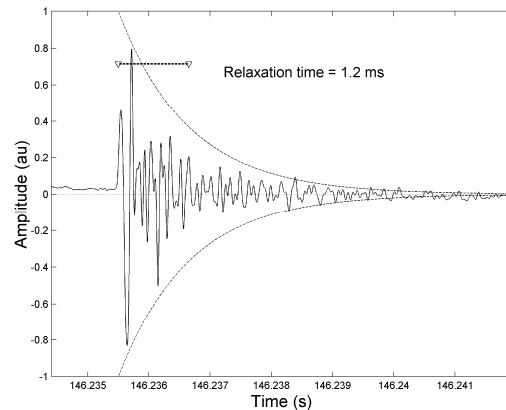


Figure 2. A large amplitude shrimp snap and the exponential decay of the envelope. The relaxation time calculated from the exponential decay curve is 1.2ms.

Figure 3 shows a representative filtered timeseries with threshold set at three standard deviations and the points where snap detections have been declared. This example has been specifically chosen to demonstrate the effect of the dead time. The first snap detection contains a nearby snap that was not detected because it lies within the dead time.

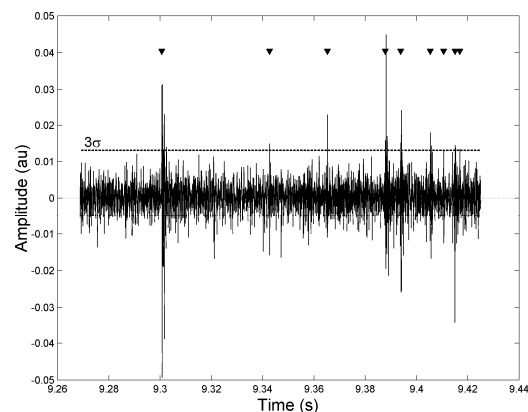


Figure 3. The representative filtered timeseries with a detection threshold set at three standard deviations (dashed line). Each detection is shown as an inverted triangle.

A combination of thresholding and dead time was used to significantly reduce the probability of detecting multiple copies of a single snap arriving as a result of multi-path propagation. The shrimp snaps most likely to produce significant multiple copies were found to be the high power, close proximity snaps. Distant shrimp snaps had direct and surface reflected pulse arrivals at approximately the same time but were only declared as one snap due to the dead time. For the largest nearby snaps only the direct path and first surface reflected path produced pulses greater than 1 standard deviation of the background noise, and only the direct path produced pulses greater than 3 standard deviations. When the threshold was set at 3 or more standard deviations the surface

reflected replicas were not detected even if they lay outside the dead time of the direct path pulse.

INTER-SNAP INTERVAL HISTOGRAM

The inter-snap interval histogram (IIH) is a computation of the frequency of occurrence of time intervals between consecutive shrimp snaps. The IIH allows comparisons to be made between empirical data and theoretical models, and is the basis for the exploratory analysis presented.

An empirical IIH was calculated for each data set by first detecting shrimp snaps using threshold $T\sigma$ and dead time τ , giving a vector of snap detection times $t_s(k)$. Using detection times $t_s(k)$, the inter-snap intervals were calculated using

$$t_{\Delta}(k) = t_s(k + 1) - t_s(k) \quad (2)$$

where $t_{\Delta}(k)$ are the inter-snap intervals. The inter-snap intervals must start using the first snap rather than an arbitrary point to avoid biasing the series (Cox and Lewis, 1966). The empirical IIH was then calculated as

$$\bar{p}(t_{\Delta}) = \frac{\zeta(t_{\Delta}, \Delta t)}{N\Delta t} \quad (3)$$

with N the number of bins used for the histogram, a bin width of Δt , and $\zeta(t_{\Delta}, \Delta t)$ the total number of inter-snap intervals within the bin of width Δt centred at t_{Δ} . Figure 4 shows a semi-logarithmic plot of an empirical IIH (circles) computed from the simulated data set. Snaps were detected using a threshold at 3σ and a dead time extending 937 μs (30 samples) past the snap detection point. The empirical points lie on a straight line when plotted on the semi-logarithmic scale suggesting exponential distribution of the inter-snap intervals. Exponentially distributed inter-snap intervals are expected if the snaps are drawn from a Poisson process (Cox and Lewis, 1966). The difference between the empirical IIH and the exponential distribution forms a zero'th order test for Poisson statistics (Lowen and Teich, 1992).

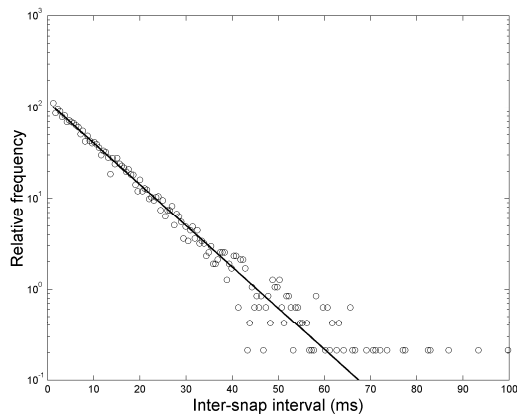


Figure 4. The empirical (circles) and theoretical (solid line) inter-snap interval histograms for simulated data. The detection threshold was set at 3σ .

Given a set of events that are Poisson distributed in time t , the probability density $p_T(t_{\Delta})$ of times between consecutive events t_{Δ} has the form

$$p_T(t_{\Delta}) = \lambda e^{-\lambda t_{\Delta}} \quad (4)$$

where λ is a rate parameter (Cox and Lewis, 1966). For snapping shrimp the events are snaps from an ensemble of

shrimp that contribute to the background noise, and the rate λ will vary with the number of shrimp that make up the ensemble. Detection of the shrimp snaps required a fixed amount of dead time following each snap, which impacts on the probability density given in Equation 4. The dead time changes the probability density because the probability of an event occurring within the dead time τ is exactly zero. Ricciardi and Esposito (1966) modify Equation 4 to allow for dead time. For a constant dead time τ , that must follow every detection, the probability density becomes

$$p_T(t_{\Delta}, \tau) = \lambda e^{-\lambda(t_{\Delta} - \tau)} \quad (5)$$

Inter-snap intervals having the dead time modified exponential distribution of Equation 5 arise from a process called the *Dead Time Modified Poisson (DTMP)* point process (Lowen and Teich, 1992). In the dead time modified exponential distribution the parameters τ and λ are required. Parameter τ is known, it is the dead time, and λ can be estimated using the method of moments

$$\lambda = \frac{1}{\bar{t}_{\Delta} - \tau} \quad (6)$$

where

$$\bar{t}_{\Delta} = \int_{\tau}^{\infty} t_{\Delta} p_T(t_{\Delta}, \tau) dt_{\Delta} \quad (7)$$

Equation 7 is an expression for the first raw moment of the empirical data. Figure 5 shows the empirical IIH for three acoustic data sets and one simulated data set. A dead time modified exponential distribution fit was conducted for each empirical data set with rate λ estimated using Equation 6.

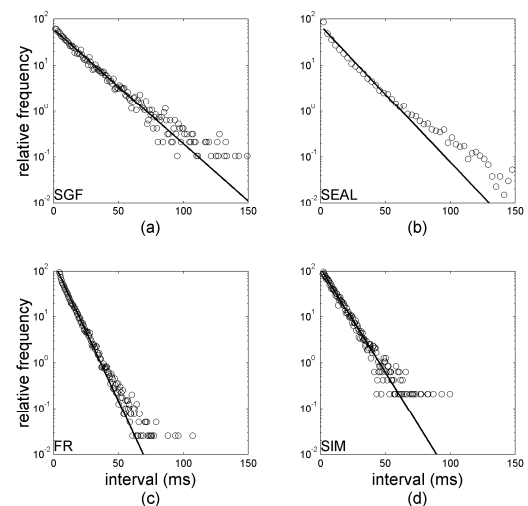


Figure 5. Empirical (circles) and theoretical (solid line) inter-snap interval histograms for the Spencer Gulf (a), Seal Island (b), Feather Reef (c), and Simulated (d) data sets. The detection threshold was set at 3σ for all four plots.

There is good agreement between the distribution of empirical IIH points and the theoretical fit for the Spencer Gulf and Simulated data sets, marginal disagreement for the Feather Reef data and strong disagreement for the Seal Island data. Good agreement is expected for the simulated data because the inter-snap intervals were drawn from a dead-time modified exponential distribution. For all data sets the agreement between empirical and theoretical results is better toward short intervals, with deviations occurring at longer

intervals. Banding appears on all plots toward longer intervals due to low counts at these interval lengths, and is exaggerated by the semi-logarithmic scale.

A two-sided Kolmogorov-Smirnov (K-S) single sample test (Dudewicz and Mishra, 1988) was used to test a null hypothesis that the empirical data was drawn from a dead-time modified exponential distribution. The test statistic used in the K-S test was

$$D_N = \sup_{t_\Delta} |\bar{c}(t_\Delta) - c_T(t_\Delta)| \quad (8)$$

Where sup is the supremum function (Weisstein, 1998),

$$\bar{c}(t_\Delta) = \int_{\tau}^{t_\Delta} \bar{p}_T(t_\Delta) dt_\Delta \quad (9)$$

is the empirical cumulative distribution function and

$$c_T(t_\Delta) = 1 - e^{-\lambda(t_\Delta - \tau)} \quad (10)$$

is a dead time modified exponential cumulative distribution function.

The D_N statistic was computed using Equation 8 for a cross-section of N values and plotted along with 0.05 and 0.01 significance levels in Figure 6. Using this construction of the K-S test, the null hypothesis was rejected if D_N was greater than either significance level for any value of N. The null hypothesis was strongly rejected if the 0.05 significance level was exceeded for any value of N greater than the suggested minimum reliable value of 40 (Dudewicz and Mishra, 1988).

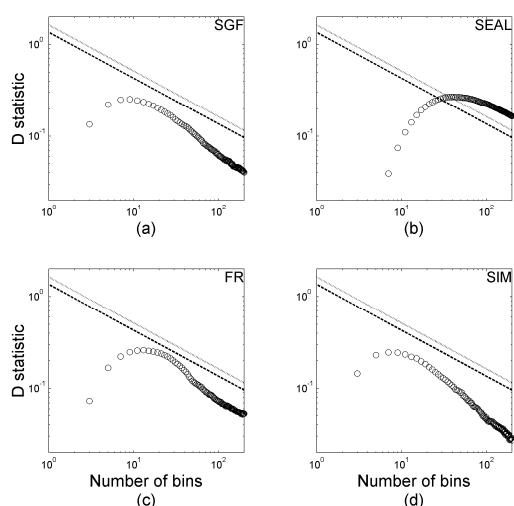


Figure 6. Kolmogorov-Smirnov D_N statistic (circles) as a function of the number of bins N for the Spencer Gulf (a), Seal Island (b), Feather Reef (c) and Simulated (d) data sets.

The significance levels are $\alpha=0.05$ (upper line) and $\alpha=0.01$ (lower dashed line).

The test results were compared with the agreement between theoretical and empirical histograms shown in Figure 5. Instances when the theoretical histogram divides the empirical histogram evenly, such as in Figure 5d, the null hypothesis is expected to be accepted. Conversely for instances when the theoretical histogram deviates from the empirical histogram the null hypothesis is expected to be rejected. Test results shown in Figure 6 reveal that only the Seal Island data crossed either of the significance levels toward rejection of the null hypothesis, with both the 0.05

and 0.01 levels exceeded for N values greater than 40. This result is in agreement with the departure of the empirical IH data from the theoretical fit in Figure 5b, especially toward the longer intervals. Acceptance of the hypothesis is expected for the simulated data, and consistent with the histogram fit for the Spencer Gulf data shown in Figure 5a. The acceptance of the Feather Reef data does not agree with the departure of the empirical histogram from the theoretical fit, shown in Figure 5c.

DISCUSSION

A simple threshold detection technique, combined with dead time after detection, has been used to locate shrimp snaps within real and simulated acoustic data. It was found that the combination of thresholding and dead time after detection was able to eliminate multi-path propagated replicas of individual snaps from being included in the analysis thereby removing the possibility of non-Poisson behaviour being attributed to propagation effects. An exploratory analysis of empirical and theoretical inter-snap interval histograms was conducted using a dead time modified exponential distribution model to account for missed detections during the dead time. The exploratory analysis showed that two out of three acoustic data sets provide evidence of non-Poisson statistics but overall the analysis was inconclusive because the Spencer Gulf data set was consistent with Poisson statistics. A two-sided Kolmogorov-Smirnov single sample test was conducted to complement the exploratory analysis but it was found that only the Seal Island data set rejected the null hypothesis, that the inter-snap intervals were drawn from a dead-time modified exponential distribution, at either the 0.01 or 0.05 significance levels.

CONCLUSIONS

Evidence has been provided that supports the notion of non-Poisson temporal behaviour in snapping shrimp noise. Of three acoustic data sets only one collected at Seal Island displayed conclusive evidence of non-Poisson behaviour in both the exploratory and statistical analysis. Further research will be conducted to conclude if the Seal Island data is unique or if the analysis presented is of too low statistical power to reveal more subtle non-Poisson behaviour.

ACKNOWLEDGEMENTS

The authors would like to thank D. Cato, D. Matthews and R. McCauley for providing respectively the Spencer Gulf, Seal Island and Feather Reef acoustic data, and also D. Bertilone for helpful discussions.

REFERENCES

- Au, W.L.L. & Banks, K. 1998, 'The acoustics of the snapping shrimp *Synalpheus parneomeris* in Kaneohe Bay', *J. Acoust. Soc. Am.*, vol. 103, no. 1, pp. 41-47.
- Bertilone, D.C. & Killeen, D.S. 2001, 'Statistics of biological noise and performance of generalized energy detectors for passive detection', *IEEE J. Oceanic Eng.*, vol. 26, no. 2, pp. 285-294.
- Cato, D.H. & Bell, M.J. 1992, 'Ultrasonic ambient noise in Australian shallow waters at frequencies up to 200kHz', *Materials Research Laboratory Technical Report MRL-TR-91-23*, Defence Science and Technology Organisation, Sydney, Australia.
- Cox, D.R. & Lewis, P.A.W. 1966, *The statistical analysis of series of events*, Methuen & Co., London.
- Dudewicz, E.J. & Mishra, S.N. 1988, *Modern mathematical statistics*, John Wiley & Sons, New York.

- Ferguson, B.G. & Cleary, J.L. 2001, 'In situ source level and source position estimates of biological transient signals produced by snapping shrimp in an underwater environment', *J. Acoust. Soc. Am.*, vol. 109, no. 6, pp. 3031-3037.
- Lowen, S.B. & Teich, M.C. 1992, 'Auditory-nerve action potentials form a nonrenewal point process over short as well as long time scales', *J. Acoust. Soc. Am.*, vol. 92, no. 2, pp. 803-806.
- Miklovic, D.W. & Bird, M.T. 2001, 'Snapping shrimp: measuring their natural distribution in space and time using low frequency arrays', *MTS/IEEE Conference and Exhibition OCEANS, 2001*, Honolulu, Hawaii.
- Potter, J.R. & Chitre, M. 1999, 'Ambient noise imaging in warm shallow seas; second-order moment and model-based imaging algorithms', *J. Acoust. Soc. Am.*, vol. 106, no. 6, pp. 3201-3210.
- Potter, J.R. & Koay, T.B. 2000, 'Do snapping shrimp chorus in time or cluster in space? Temporal-spatial studies of high-frequency ambient noise in Singapore waters', *Proceedings of the Fifth European Conference on Underwater Acoustics, ECUA 2000*, Lyon, France.
- Potter, J.R., Lim, T.W. & Chitre, M. 1997, 'Ambient noise environments in shallow tropical seas and the implications for acoustic sensing', *Oceanology International '97*, Singapore.
- Ricciardi, L.M. & Esposito, F. 1966, 'On some distribution functions for non-linear switching elements with finite dead time', *Kybernetik*, vol. 3, pp. 148-152.
- Snyder, D.L. & Miller, M.I. 1991, *Random point processes in time and space*, 2nd edn, Springer-Verlag, New York.
- Watkins, P.T., Santhanam, G., Shenoy, K.V. & Harrison, R.R. 2004, 'Validation of adaptive threshold spike detector for neural recording', *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, San Francisco, California.
- Weisstein, E.W. 1998, *The CRC concise encyclopedia of mathematics*, CRC Press, Boca Raton, Florida.
- Versluis, M. Schmitz, B. von der Heydt, A. & Lohse, D. 2000, 'How snapping shrimp snap: through cavitation bubbles', *Science*, vol. 289, pp. 2114-2117.