Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 December 2017 doi:10.20944/preprints201712.0001.v1

A Sequential Algorithm for Signal Segmentation

Paulo Hubert (phubert@ime.usp.br) *[1], Linilson Padovese (lrpadovese@usp.br) [2] and Julio Stern (jstern@ime.usp.br) [1]

[1] Instituto de Matematica e Estatistica - University of Sao Paulo (IME-USP)
[2] Escola Politecnica - University of Sao Paulo (EP-USP)
* Corresponding author

Abstract

The problem of event detection in general noisy signals arises in many applications; usually, either a functional form for the event is available, or a previous annotated sample with instances of the event that can be used to train a classification algorithm. There are situations, however, where neither functional forms nor annotated samples are available; then it is necessary to apply other strategies to separate and characterize events. In this work, we analyze an acoustic signal obtained from a hydrophone, and are interested in separating sections, or segments, of the signal which are likely to contain significative events. For that, we apply a sequential algorithm with the only assumption that an event alters the average power of the signal. The algorithm is entirely based on bayesian methods, and shows a very promising performance in detecting either short or long events.

Keywords signal processing; bayesian methods; subaquatic audio; hydrophone; unsupervised learning

1 Introduction

Signal processing is a field of intense research, both in engineering, physics and medicine[1], and in the statistical and MAXENT literature [Bretthorst (1988)] [Jaynes (1987)] [Ruanaidh (1996)]. The most recurrent problems in the field are signal estimation, model comparison, and signal detection[Bretthorst (1990A)] [Bretthorst (1990B)] [Bretthorst (1990C)]: in signal estimation, given a functional form for the signal (for instance, an exponentially decaying tonal model) and a (discretized) sample, the researcher wants to estimate the functional form's parameters (the rate of decay, the fundamental frequency). In model comparison, there are different possible functional forms for the signal, and given one or more samples, one is interested in selecting the most adequate model. In signal detection, the researcher is given a sample of the signal, and must decide if a given functional form (which we might call an *event*) is or is not present.

These situations arise tipically when the researcher has some control over the process of data acquisition. Namely, she usually is well aware of the presence of the event of interest in the sample, or knows precisely the kind of event she is looking for (and / or its duration). In these situations, Jaynes and Bretthorst have succesfully applied maximum entropy and bayesian methods to solve the basic problems of signal estimation, model comparison and signal detection.

On the other hand, there are situations in which the researcher has very little information about the times of occurrence or precise nature of events. This is the case in this work: we are given a subaquatic audio recording, obtained from a hydrophone, with a total length of 3 months. We are told that, during these three months, many phenomena were possibly captured by the hydrophone, including fish vocalization, the chanting of whales, motor boats and what-not, and we are asked to separate from the signal a number of samples where we believe one of these things is happening.

Put in this way, this is a very general endeavour, and we propose to treat it like so, using a minimum of assumptions. The maximum entropy + bayesian inference framework is thus the theoretical toolbox of choice, since from maximum entropy we learn how to avoid insidious and implicit assumptions to sneak into the model, and from bayesian inference we learn how to make the best use of all prior available information.

[1] For a quick and interesting exposition on the applications of signal processing, see for instance `https://signalprocessingsociety.org/our-story/signal-processing-101`. For subacquatic signal processing, see [Etter (2013)].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 December 2017 doi:10.20944/preprints201712.0001.v1

Our approach is as follows: first we propose a very simple model, which we call a model for *power regime switch*, that allows us to divide a given signal in two sections with different average powers. We then utilize some measure of statistical evidence to guarantee that indeed the signal power is different between the two samples. If this is the case, we proceed recursively, dividing each new segment into two, measuring the evidence for power difference, and so on. This approach assumes only that whenever a signal is added to background noise, the average total power of the signal changes.

We show that the algorithm thus defined is able to efficiently separate a long signal into segments that capture the occurrence of events of variable lengths.

2 Event detection in subacquatic acoustic data

The Laboratory of Acoustics and Environment (LACMAM), at EP-USP, has developed an equipment called OceanPod: it is a hydrophone with an autonomy for 5 months of continuous recording, with a frequency band of 5Hz to 40kHz, and a sampling frequency that can be set from 5kHz to 24kHz[2].

In 2015, one OceanPod was installed in the region of *Parque da Laje*, a marine preservation park in the brazilian coast, in the region of Santos, SP. The hydrophone was installed at a 20m depth, and recorded 3 months of subacquatic sounds before its retrieval by the laboratory. The equipment was configured to sample the signal with a frequency of 11025 Hz.

Once retrieved, the signal was submitted to analysis, with a first goal of understanding the behavior of small fishing vessels that (ilegally) visit the park. With a better understanding of the patterns of behavior of these vessels, it is hoped that better fiscalization policies can be built and implemented.

There are difficulties, however, with the analysis of these data. Most of all, there is the problem of characterizing the sound of boat engines, in order to separate it from background noise and other events that are also present in the time series. One approach has been made[Hubert (2017)], based on a chirp model for the boat engine; and even though it showed promising results, there was a problem of computation efficiency that prevented its immediate application to the entire signal data.

Apart from the efficiency issues, this method is not easily generalized for the detection of other types of events, since it strongly depends on a specific functional form for the event's acoustic signature. Since there are many different events probably occurring during these 3 months, the need for a more general approach was readily evident.

The basic idea was then to, first and foremost, separate from the signal the sections that are likely to contain some event (as opposed to sections containing noise only). The underwater oceanic environment is very economic in this sense: there can be periods of many minutes (even hours) where nothing but the waves (and the omnipresent sound of shrimps and barnacles clapping and clicking) can be heard. This indicates that, rather than processing the entire signal, it would be much better to first obtain sections where we believe something is happening, and then try to figure out exactly what it is.

In this paper we present a maximum entropy + bayesian approach to this problem, and show that it can efficiently separate these eventful situations from the raw signal. To do that, we begin by considering the model for a single phenomenom, and to keep the model as general as possible, we assume solely that *the beginning of an event induces a change in the signal's average power.*

3 The power regime switch model

Suppose we are given a discretely sampled signal $y \in \Re^N$. Given a sampling frequency rate f_s, this signal corresponds to a total duration of $T = N \cdot f_s$ seconds. We assume that $E(y_i) = 0$, and also assume that $Var(y_i) = \sigma_0^2 < \infty$.

Now let's imagine that, at some instant $\bar{t} \in [0, T]$, some event has started. We assume that, whatever this event is, it induces a change in the total average power of the signal; this is a rather reasonable assumption, since for the combination of two random variables to have the same average power as one of its components, it is necessary that the variables have an exact covariance of $\sigma_1^2/2$, where σ_1^2 is the average power of either one of the variables.

In this situation, and again assuming that all signals have 0 mean amplitude and finite power, the maximum entropy principle leads us to choose the following probabilistic model for y:

[2] More information about the OceanPod can be found at `http://lacmam.poli.usp.br/Submarina.html`

 doi:10.20944/preprints201712.0001.v1

$$y_i \sim \begin{cases} \mathcal{N}(0, \sigma_0^2) & \text{if } i \leq \bar{t} \\ \mathcal{N}(0, \delta\sigma_0^2) & \text{if } i > \bar{t} \end{cases}$$

The parameter δ represents the ratio between the average powers of the two signal sections or segments.

Our assumption that any event will cause the total power to change translates to the hypothesis $\delta \neq 1$. Assuming this hypothesis as true, our goal is to estimate the value of $\bar{t}$, the time when the event starts.

With a gaussian model, we arrive at the likelihood

$$\mathcal{L}(\bar{t}, \sigma_0^2, \delta | y) = (2\pi\sigma_0^2)^{-\frac{N}{2}} \delta^{-\frac{N-\bar{t}}{2}} exp \left[-\frac{\sum_{i=1}^{\bar{t}} y_i^2}{2\sigma_0^2} - \frac{\sum_{i=\bar{t}+1}^{N} y_i^2}{2\delta\sigma_0^2} \right] \tag{1}$$

Since our main interest lies in the estimation of $\bar{t}$, we start by adopting a non-informative uniform and improper prior for δ; this choice allows us to analitically integrate equation 1 to obtain

$$\mathcal{L}(\bar{t}, \sigma_0^2 | y) = \int_0^\infty \mathcal{L}(\bar{t}, \sigma_0^2, \delta | y) d\delta \tag{2}$$

$$\propto \left[\frac{\sum_{i=\bar{t}+1}^{N} y_i^2}{2\sigma_0^2} \right]^{-\frac{(N-\bar{t}-6)}{2}} \Gamma\left(\frac{N-\bar{t}-2}{2}\right) exp \left[-\frac{\sum_{i=1}^{\bar{t}} y_i^2}{2\sigma_0^2} \right] \tag{3}$$

If the average power of the first section (i.e., the average power of noise, if we assume that the signal starts with noise only) is known, the above equation can be used directly to obtain the posterior distribution for $\bar{t}$. If σ_0^2 is unknown, we can adopt the Jeffreys prior $\pi(\sigma_0) = 1/\sigma_0$, and again integrate the above equation analitically to arrive at

$$\mathcal{L}(\bar{t}|y) = \int_0^\infty \mathcal{L}(\bar{t}, \sigma_0^2 | y) \pi(\sigma_0) d\sigma_0 \tag{4}$$

$$\propto \left(\sum_{i=\bar{t}+1}^{N} y_i^2 \right)^{-\frac{(N-\bar{t}-6)}{2}} \Gamma\left(\frac{N-\bar{t}-2}{2}\right) \left(\sum_{i=1}^{\bar{t}} y_i^2 \right)^{-\frac{(\bar{t}+6)}{2}} \Gamma\left(\frac{\bar{t}+6}{2}\right) \tag{5}$$

From this point, it suffices to pick a prior $\pi_t(t)$ for $\bar{t}$, multiply it by the above quation, and obtain (the kernel of) the posterior distribution for $\bar{t}$:

$$P(\bar{t}|y) \propto \pi(t) \cdot \left(\sum_{i=1}^{\bar{t}} y_i^2 \right)^{-\frac{(\bar{t}+6)}{2}} \left(\sum_{i=\bar{t}+1}^{N} y_i^2 \right)^{-\frac{(N-\bar{t}-6)}{2}} \times \tag{6}$$

$$\Gamma\left(\frac{\bar{t}+6}{2}\right) \Gamma\left(\frac{N-\bar{t}-2}{2}\right) \tag{7}$$

This gives us a discrete distribution with support over $1 < \bar{t} < N-1$; this means that it is in principle possible to calculate exactly the normalization constant that would make 6 a proper probability distribution. If we need to estimate $\bar{t}$, on the other hand, we can optimize equation 6 by inspection to find the MAP estimate.

In the next figures we present the distribution in 6 calculated over a simulated (gaussian) signal of size $N = 3000$, with $\delta \in \{1.1, 1.5, 2\}$, $\bar{t} \in \{N/3, N/2, 2N/3\}$ and $\sigma_0 = 1$. We notice a few important features of the above method; first, it works very well for higher values of δ, as would be expected. Second, when it does not work well, it often attributes high posterior probability to points next to the extremes of the interval, which is an artifact of the high variation of variance estimatives in small samples. This articact, however, could be easily eliminated by the use of an informative prior on $\bar{t}$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 December 2017 doi:10.20944/preprints201712.0001.v1

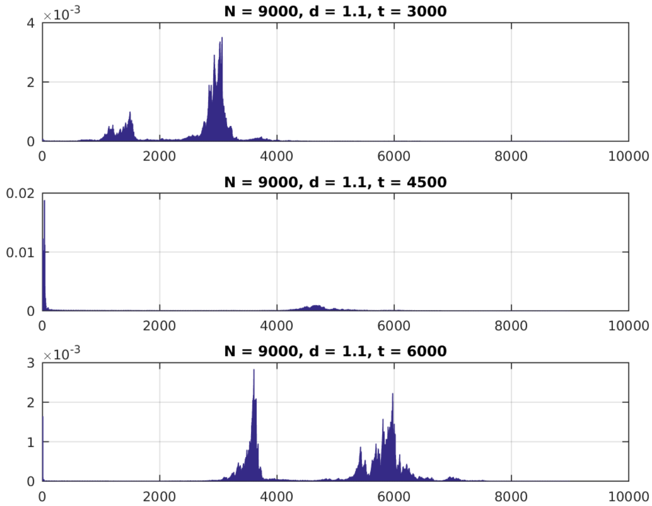

Figure 1: Posterior distribution for $\bar{t}$, $\delta = 1.1$

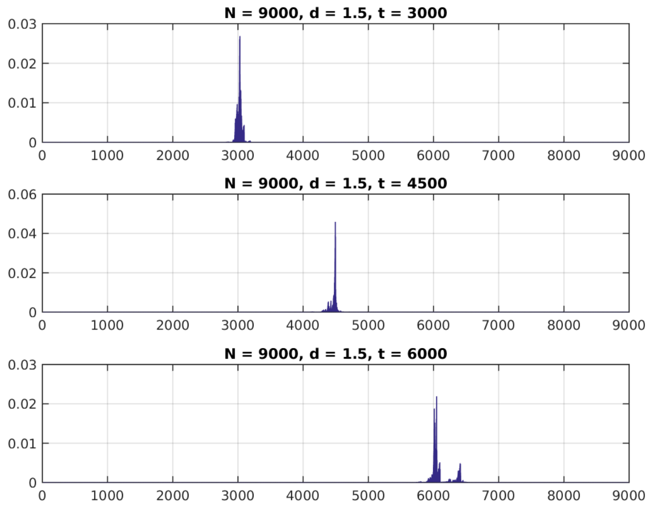

Figure 2: Posterior distribution for $\bar{t}$, $\delta = 1.5$

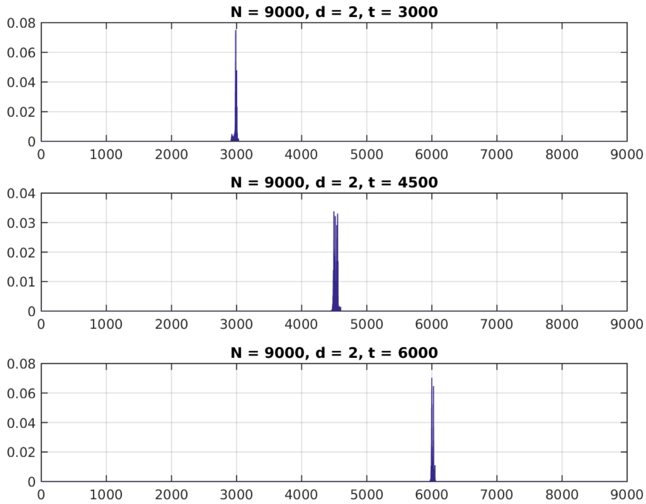

Figure 3: Posterior distribution for $\bar{t}$, $\delta = 2$

4 The sequential segmentation algorithm

It is straightforward to extend the regime switch model to signals with two or more power changes. The complexity of the model, however, would increase exponentially on the number of changes, and also there would be the question of how many power changes one expects in a signal. To try and keep things simple, we propose then an alternative based on the recursive application of the segmentation model of last section.

The idea is the following: given a signal $y \in \Re^N$, we use the model to estimate the point $\bar{t}$ which divides the signal into two smaller windows with maximally different average powers. This will give us two new signals, y_1 and y_2, which we can in turn divide again by using the same model. This process can continue until some stopping criterion is met. This is the sequential segmentation algorithm, defined below:

(Sequential segmentation algorithm) Define the function seg($y \in \Re^N$,n_{min}) as:

1. **If** $N < n_{min}$, stop, returning the empty vector $t = []$;
2. **Else**
   (a) Obtain the MAP estimate $\bar{t}$;
   (b) Define $y^1 = y_{1,\dots,\bar{t}}$ and $y^2 = y_{\bar{t}+1,\dots,N}$;
   (c) (Stopping criterion) **If** $var(y^1) = var(y^2)$, stop, returning the empty vector $t = []$;
   (d) **Else** return the concatenated vector $[seg(y^1, n_{min}) \quad \bar{t} \quad \bar{t} + seg(y^2, n_{min})]$.

This (recursive) algorithm will produce an ordered vector $\tau \in [1, \dots, N-1]^k$ with the starting points of k signal segments, where the segments have different average power. The condition $N \geq n_{min}$ guarantees that the algorithm stops and is well defined; the main question is how to decide if $var(y^1) = var(y^2)$, i.e., to define the stopping criterion.

5 Stop criteria for the segmentation algorithm

As it is clear from the definition, the matter is one of testing the hypothesis of equality of variances, given two samples with mean 0. Or, using our parametrization, to test the hypothesis $H_0 : \delta = 1$.

Our model is defined in the parametric space $\Theta = \Re_+ \times \Re$, where $\theta = (\sigma_0^2, \delta)$; under H_0, we have $\Theta_0 = \Re_+$. Our hypothesis thus lies in a subspace of Θ, i.e., it is a *sharp* hypothesis.

The traditional statistical literature proposes a few different ways to test equality of variances, the most known being possibly the F test, and the likelihood ratio test. However, it is well reported that the traditional, frequentist tests, have a few drawbacks, particulary some problems related to the definition of the alternative hypothesis[Good (1992)], or with the choice of an appropriate significance level for the decision function[Perez (2014)][Pereira (1993)]. The classical bayesian alternative would then be a bayes factor test, which in turn would meet some difficulties with the fact that our null hypothesis induces a lower dimensional parametric space[Pereira (1999)].

A full bayesian procedure is available, however, that is well suited to the test of sharp hypothesis such as $H_0 : \delta = 1$; this is the now well-known *full bayesian significance test* (FBST) of Pereira and Stern[Pereira (1999)]. This procedure works in the full parametric space, defining an evidence measure based on the *surprise set* of points having higher posterior density than the supremum posterior under H_0. The test avoids altogether the imposition of positive probabilities over null measure sets such as $\Theta_0 : \{\theta \in \Theta : \delta = 1\}$, and has been tested many times with very good results (for a few examples see [Chakrabarty (2017)][Hubert (2009)][Lauretto (2009)][Stern (2002)]).

Of course, even with the application of the FBST, there remains the problem of defining a decision function over the evidence value for H_0, $ev(H_0, y)$; as we will see, however, this decision function will become trivial when we calibrate our testing procedure with the use of appropriate priors.

Recalling the probabilistic model of our signal, given y and $\bar{t}$, and defining y^1 and y^2 as in the above algorithm, we write the full posterior

$$P(d, s|y^1, y^2) \propto \pi_\delta(d)\pi_\sigma(s)(2\pi s^2)^{-\frac{n_1+n_2}{2}} d^{-\frac{n_2}{2}} exp\left[-\frac{\sum_{i=1}^{n_1} y_{1,i}^2}{2s^2} - \frac{\sum_{i=1}^{n_2} y_{2,i}^2}{2ds^2}\right] \tag{8}$$

where n_1 and n_2 are the corresponding dimensions of y^1 and y^2, π_δ is the prior for δ and π_σ the prior for σ_0.

To incorporate the lack of knowledge about the base signal variance σ_0, we adopt a Jeffreys prior $\pi_\sigma(s) = 1/s$. For δ, however, we would like to choose an informative prior, to model our knowledge about the signal characteristics.

But what do we know about δ? Let's think in the following terms: suppose we are to pick at random two contiguous sections of our signal, with sizes n_1 and n_2; suppose that these sections have s_1 and s_2 as their respective average powers. What can we say about $\delta = \frac{s_2}{s_1}$?

Well, unless we happen to pick by chance two segments that include a change of regime (in our terms, the beginning or end of an event), we expect δ to be very close to 1. But how strong is this belief?

It depends, of course, in our perceived probability of finding an event at random in our signal. As we have mentioned before, the ocean's subacquatic landscape is a rather minimalist environment, with long periods of very low or no activity. So we believe that, in our thought experiment above, we are very likely to pick sections with δ very close to 1.

However, if we happen to pick a segment with an event, then we can expect to find $\delta >> 1$, since most events in our signal have large SNR values (for instance, the already mentioned boat engines, running at a small distance from the hidrophone). It is then reasonable to believe that, even though δ is likely close to 1, it can sometimes differ significantly and assume high values, close to $\delta = 2$ or even higher.

All of these considerations lead us to pick a prior distribution on δ that is: (a) centered around 1; (b) high-peaked around this same value; but (c) with larger tails than the gaussian. Further on, to keep matters as simples as possible, we would like our prior to have few parameters (since we might use these parameters for the calibration of our algorithm). Combining all of these objectives, we end up with a Laplace prior for δ:

$$\pi_\delta(d) = \frac{1}{\beta} e^{-\frac{|d-1|}{\beta}} \tag{9}$$

In practical terms, this prior will tend to favor $H_0 : \delta = 1$, inversely with the value of β. This value can be used as a calibration parameter for the detection algorithm. Also, picking a sufficiently low value for β will guarantee a minimum prior probability for the meaningless event $\delta \leq 0$.

Again, we must stress that in working with acoustic signals with a sampling rate as high as 11kHz, we will be dealing with large sample sizes; tipically we will define a smallest detectable event as a segment with a duration of around 1s, which means that we will be comparing the variances of samples with sizes $N = 11025$ each. On the other extreme, the algorithm will start with a signal of duration in the order of minutes (the files obtained from the hydrophone are configured as 15 min long for default), which translates to sample sizes in the order of millions. Our numerical tests indicate that the value of β must be set correspondingly; our best results used $\beta \propto 10^{-5}$, as we will see in the results section.

With the model completely specified, the evidence value for H_0 is calculated as

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 December 2017 doi:10.20944/preprints201712.0001.v1

$$ev(H_0, y) = \int_{T(y)} P(\sigma_0, \delta|y) d\sigma_0 d\delta \tag{10}$$

where

$$T(y) = \{(\sigma_0, \delta) \in \Theta : P(\sigma_0, \delta|y) > sup_{\Theta_0} P(\sigma_0, \delta|y)\} \tag{11}$$

To define the stopping criterion for the algorithm, it suffices to set a minimal evidence for H_0, α_{min}; then, the criterion will be:

(Stopping criterion) Given y^1, y^2 and α_{min}:

1. Obtain $s_0 = sup_{\Theta_0} P(\sigma_0, \delta|y)$;
2. Obtain the evidence $ev(H_0, y) = \int_{T(y)} P(\sigma_0, \delta|y) d\sigma_0 d\delta$;
3. **If** $ev(H_0, y) < \alpha_{min}$, return 1; **else** return 0.

We calculate the above integral using a simple Metropolis-Hastings algorithm, with 10000 points after a burnin of another 10000 points.

6 Results

For the numerical tests, we have picked three distinct sections of our original signal.

The first is a recording of 2015-30-01, saturday, from 02h02m56s to 02h17m56s. On this day there is no activity (at least visually perceivable) beyond background noise (concentrated around 5kHz) in this signal (see spectrograms below).

The second instance is another 15 minute long recording, now of date 2015-02-02, monday, from 07h50m49s on. In this instance, the spectrogram shows an event of long duration (approximately 10 minutes). Hearing this instance we can classify this event as the passage of a large sized vessel, from a long distance and with low speed.

The third instance is a recording from 2015-08-02, sunday, from 11h26m39s to 11h41m39s. In this section the spectrogram shows intense activiy; hearing the sample we identify the sounds of engines, from smaller boats at a short distance from the hydrophone.

Now we apply the sequential segmentation algorithm to each of this samples; we pick $\beta \in \{3e^{-5}, 1e^{-5}\}$. Higher values for β result in an excess of segments being returned.

Each 15 minutes section took on average 2 minutes to process on a Pentium notebook with i5 processor of 1.6GHz, and 8Gb RAM.

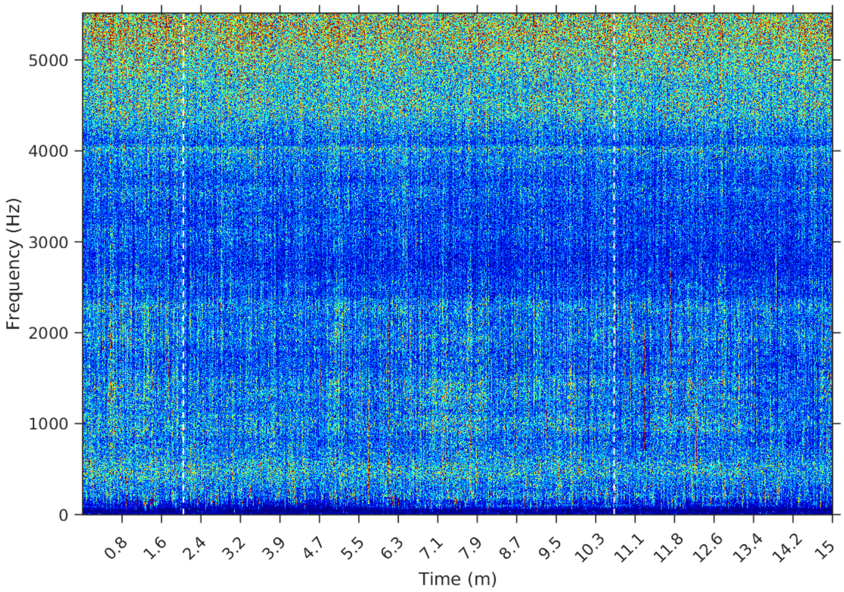

Figure 4: Algorithm results, $\beta = 3e^{-5}$, 2015-30-01 sample

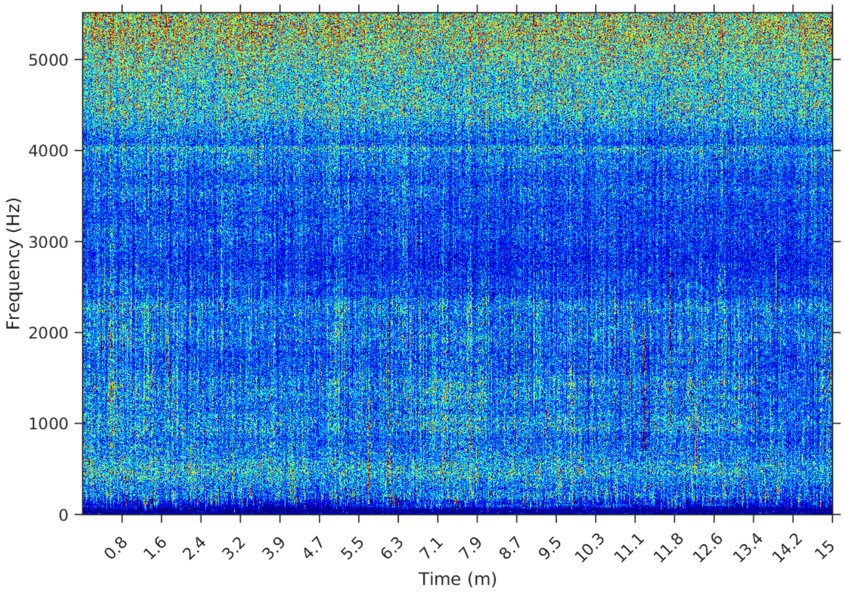

Figure 5: Algorithm results, $\beta = 1e^{-5}$, 2015-30-01 sample

 doi:10.20944/preprints201712.0001.v1

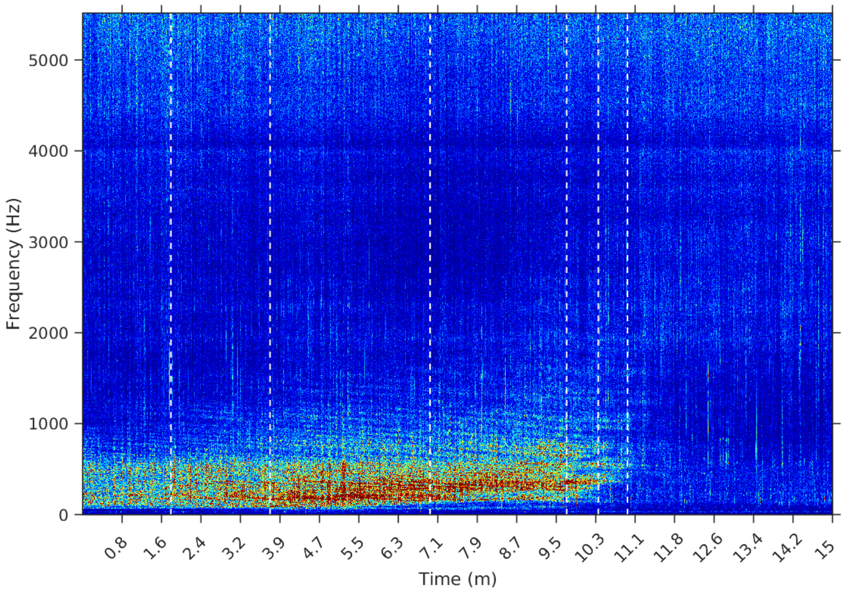

Figure 6: Algorithm results, $\beta = 3e^{-5}$, 2015-02-02 sample

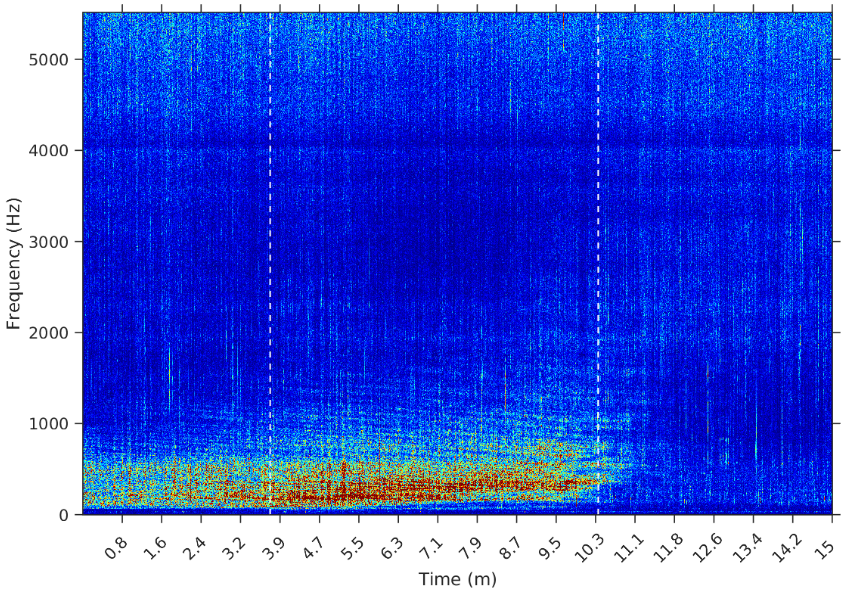

Figure 7: Algorithm results, $\beta = 1e^{-5}$, 2015-02-02 sample

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 December 2017 doi:10.20944/preprints201712.0001.v1

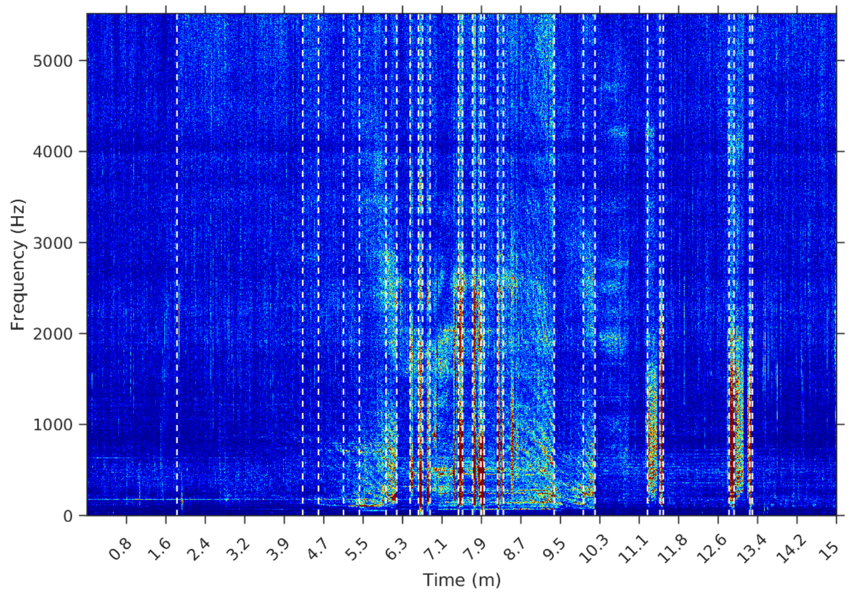

Figure 8: Algorithm results, $\beta = 3e^{-5}$, 2015-08-02 sample

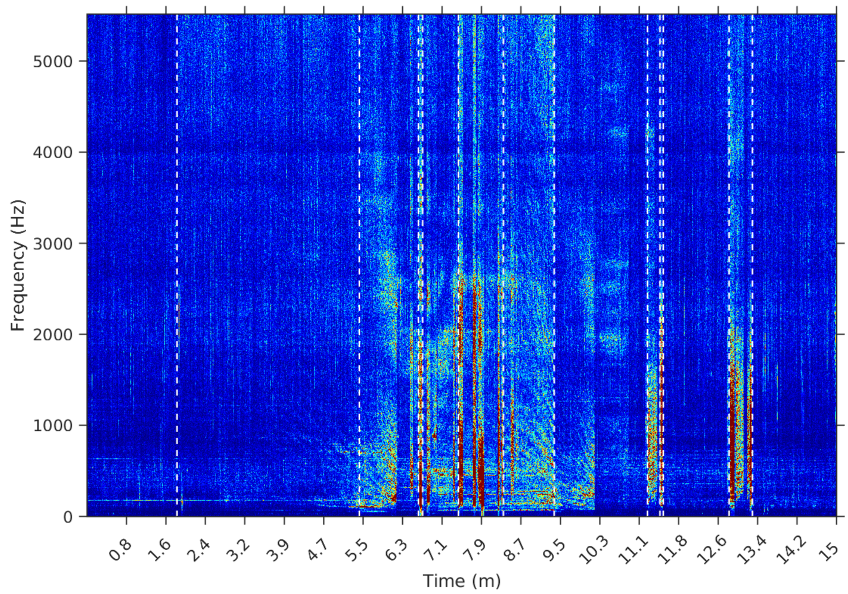

Figure 9: Algorithm results, $\beta = 1e^{-5}$, 2015-08-02 sample

As we can see, the algorithm was able to capture precisely the quick events of 2015-08-02, while at the same time extracting most of the long duration event of 2015-02-02, and not segmentating altogether the quiet sample from 2015-30-01 (at least when $\beta = 1e^{-5}$).

The value of β altered significantly the behavior of the algorithm; when $\beta = 1e^{-5}$ many segments are dropped, in comparison with the case $\beta = 3e^{-5}$. This indicates the need for careful calibration of this parameter when applying this algorithm for the entire data set.

Finally, it is interesting to note that, when we provided the model with a highly informative prior such as the one we have picked for δ, the value of α_{min} became irrelevant, since the resulting evidence values would quickly converge to either 0 or 1 (at least for the 10000 MCMC sample we are using; i.e., evidence was less than $1e^{-4}$, or greater than 0.9999). In this way, we believe that the use of the FBST measure overcame the need for a minimum significance level, and balanced errors of both types (false positives vs false negatives) naturally by the use of an strongly informative prior distribution over the parameter of interest.

7 Conclusions

Our goal in this paper was to define an algorithm that could be as general as possible, in order to prepare our signal data for further analysis. We wanted it to be general specially because we do not know in advance exactly what kinds of patterns (events) there might be in the data. In this sense, this algorithm can be described as an unsupervised learning method, and could be applied to the analysis of any signal or time series where the assumptions hold.

Future work now involves the application of our algorithm through the entire time series of the signal. After the segmentation, we will investigate methods to group similar events together, where "similar" can be defined in terms of duration and initial time of the day of the events, but should also include information about the signal itself: its spectrum, for instance.

Another line for future work is in the calibration of β, the hyperparameter of the model. As we have shown, the choice of this parameter is determinant to the behavior of the algorithm. To calibrate it, we might use a few known samples to pick an optimal value, or develop a scheme that allows β to vary during the execution of the algorithm (as a function of the sample size, for example).

The MATLAB code for the segmentation algorithm is available upon request.

The authors would like to acknowledge professors Carlos Pereira and Victor Fossaluza for very helpful insights in this research project. In particular, they were the ones first suggesting that we should follow a general approach before trying to model specific events.

References

[Bretthorst (1988)] Bretthorst, L. *Bayesian Spectrum Analysis and Parameter Estimation*, Springer-Verlag: New York, USA, 1988; ISBN 978-0-387-96871-1

[Bretthorst (1990A)] Bretthorst, L. Bayesian Analysis. I. Parameter Estimation Using Quadrature NMR Models, *Journal of Magnetic Resonance* **1990**, *88*, 533-551

[Bretthorst (1990B)] Bretthorst, L. Bayesian Analysis. II. Signal Detection and Model Selection, *Journal of Magnetic Resonance* **1990** *88*, 552-570

[Bretthorst (1990C)] Bretthorst, L. Bayesian Analysis. III. Applications to NMR Signal Detection, Model Selection and Parameter Estimation, *Journal of Magnetic Resonance* **1990**, *88*, 571-595

[Chakrabarty (2017)] Chakrabarty, D. A New Bayesian Test to Test for the Intractability-Countering Hypothesis, *JASA* **2017**, *112 (518)*, 561-577

[Etter (2013)] Etter, P.C. *Underwater Acoustic Modeling and Simulation*, CRC Press, Boca Raton, USA, 2013; ISBN 978-1-4665-6494-7

[Good (1992)] Good, I.J. The Bayes/Non-bayes compromise: a Brief Review, *Journal of the American Statistical Association* **1992**, *87 (419)* 597-606

[Hubert (2009)] Hubert, P., Lauretto, M., Stern, J.M. FBST for Generalized Poisson Distributions *AIP Conference Proceedings* **2009**, *1193 (1)* 210-217 DOI: 10.1063/1.3275617

[Hubert (2017)] Hubert, P., Padovese, L.R., Stern, J.M. Full bayesian approach for signal detection with an application to boat detection on underwater soundscape data, to appear in *Maximum Entropy Methods in Science and Engineering*, Springer-Verlag, New York, USA

[Jaynes (1987)] Jaynes, E. Bayesian Spectrum and Chirp Analysis, In *Maximum-Entropy and Bayesian Spectral Analysis and Estimation Problems*, C.R. Smith, G.J. Erickson, Eds.; D. Reidel Publishing Co., Dordrecht, Holland, 1987; ISBN 978-94-009-3961-5

[Lauretto (2009)] Lauretto, M.S., Nakano, F., Faria Jr., S.R., Pereira, C.A.B., Stern, J.M. A straightforward multiallelic significance test for the Hardy-Weinberg equilibrium law, *Genetics and Molecular Biology* **2009**, *32 (3)* http://dx.doi.org/10.1590/S1415-47572009000300028

[Perez (2014)] Perez, M., Pericchi, L.R. Changing Statistical Significance with the Amount of Information: The Adaptative α Significance Level, *Statistics & Probability Letters* **2014**, *85*, 20-24 DOI: 10.1016/j.spl.2013.10.018

[Pereira (1993)] Pereira, C.A.B., Wechsler, S. On the Concept of P-value, *Revista Brasileira de Probabilidade e Estatistica* **1993**, *7* 159-177

[Pereira (1999)] Pereira, C.A.B., Stern, J.M. Evidence and credibility: full Bayesian significance test for precise hypotheses, *Entropy* **1999**, *1* 99-110

[Ruanaidh (1996)] Ruanaidh, J.J.K., Fitzgerald, W.J. *Numerical Bayesian Methods Applied to Signal Processing*, Springer, New York, USA, 1996; ISBN 978-0387946290

[Stern (2002)] Stern, J.M., Zacks, S. Testing the independence of Poisson variates under the Holgate bivariate distribution: the power of a new evidence test, *Statistics & Probability Letters* **2002**, *60 (3)*, 313-320