Manatee call detection using Recursive Least Squares Predictors

Akhil Perincherry

Department of Electrical and Computer Engineering, University of Florida

***Abstract*—The problem considered is of detecting signal in a signal plus noise scenario. The paper addresses the problem of detecting Manatee calls by using the RLS online algorithm to obtain a trained model for the Manatee call and for the noise. These models are applied in parallel and comparing their prediction performance sheds light on the Manatee call regions. The filter parameters were obtained by a 3 fold cross validation approach. Multiple Manatee models and noise models were investigated to obtain the best model. The underlying role of the RLS is discussed pertaining to this problem. The final model is evaluated using receiver operating characteristics. The effect of the model on the quasi stationary nature of the signal is studied.**

***Index Terms*—RLS, Adaptive filter, Detection, Prediction**

I. Introduction

THE The Recursive least squares (RLS) is an adaptive filter which recursively finds the coefficients that minimize a weighted linear least squares cost function relating to the input signals. It has it's rate of convergence an order faster than the LMS filter because the RLS filter whitens the data by using the inverse correlation matrix of the data. This also results in increase of computational complexity though. The μ parameter in LMS is replaced by the inverse of the correlation matrix of the input vector which whitens the tap inputs. Also, the rate of convergence of the RLS algorithm does not vary with condition number of the ensemble average correlation matrix of the input vector. The forgetting factor in RLS helps tune the filter with respect to the environment being stationary or non stationary. The excess mean square in RLS asymptotically converges to zero.

A. RLS Equations

The RLS algorithm is a system of initialize-update equations that solves the Wiener Hopf equations in a computationally efficient manner. The cost function that it is considered is the Least square error of the estimate. Let $order$ be the filter order.

1) Initialize the weight vector of length $order$ to zeros and the inverse autocorrelation matrix P to a diagonally loaded matrix - loading factor of 1000 was chosen here.
2) Consider the first $order$ number of elements of the input.
3) Find the instantaneous output using the current weight vector and the input considered in the above step.
4) Compute the instantaneous error.
5) Use the above result to update the weight vector as in (2).
6) Slide the input used in step 2 by one sample and repeat the whole procedure.

Initialization equations:

$$w_{est}(0) = 0$$

$$P(0) = largeNumber * IdentityMatrix$$

The update equations are as follows:

$$t(n) = P(n-1)u(n)$$

$$k(n) = \frac{t(n)}{\alpha + u^H(n)t(n)}$$

$$e(n) = d(n) - W_{est}^H(n-1)u(n)$$

$$w_{est}(n) = w_{est}(n-1) + k(n)e(n)$$

$$P(n) = \alpha^{-1}P(n-1) - \alpha^{-1}k(n)u^H(n)P(n-1)$$

The comparison metrics used are the Normalized MSE as shown in (3).

$$NMSE = \frac{mean(abs(Output - estimatedOutput)^2)}{mean(abs(Input)^2)} \quad (1)$$

The forgetting factors α implies the memory of the learning algorithm, it implies how the past input samples are weighted. Here, the αvalues are generated in such a way that the exponential window is halved at previous L samples i.e.

$$\alpha = 0.5^{\frac{1}{halfSample}}$$

If window is to be designed such that the exponential window is half at the 500th sample, $\alpha = 0.5^{\frac{1}{500}} = 0.9986$ and so on.

RLS is basically solving Wiener Hopf system of equations recursively and with an exponential window.

II. Implementation Algorithm:

A Manatee train signal, noise signal and test signal are provided. The steps employed are discussed in brief below. Detailed discussions can be found in the results section.

- Optimum set of filter parameters was obtained by a 3 fold cross validation using RLS.
- The training signal was segmented into 10 calls and 10 models were arrived at for the Manatee train signal.
- Similarly, 4 noise models were developed.
- Filter weights were extracted with respect to NMSE values and the smoothness of the filter tap trajectories.

- The prediction errors (NMSE) were computed for all the models which were then smoothed by a moving average filter.
- By comparing the NMSE values of the two models, a decision was arrived at for the Manatee calls.
- The model is evaluated with respect to the ground truth by means of it's Receiver operating characteristics.
- Band pass filter was implemented to note the effect of filtering on the model's performance.

III. Results

The algorithms were implemented in MATLAB. The environment used is Windows 8.1 Intel i7 processor with 8GB RAM. A forgetting factor of around 0.9999 and an order of 3 was seen to generate the best results. A training signal consisting of 10 Manatee calls is provided and a noise signal of about 2 seconds length is provided to train the respective Manatee and noise models. It is to be noted that the training signal of duration 30 seconds is sampled at 48000kHz while the noise signal is sampled at 44.1kHz. Therefore, it is imperative to re sample the noise signal to 48kHz using multirate sampling techniques. Special care must be taken to avoid folding of higher frequencies while this is done aka aliasing. This can be achieved by simple low pass filtering.

A. *Choice of filter parameters*

A three fold cross validation technique was implemented to obtain the optimum set of filter parameters. The training signal was segmented into two sets where one set was used for training the filter while the other set was used as testing or a validation tool. Here, 7 out of the 10 Manatee calls were segmented out of the original training signal to obtain the cross validation training set. The remaining 3 were used as the cross validation testing set. 8 training models were considered by taking random combinations of the 7 calls out of the total 10. Figure 1 shows the Manatee training signal.

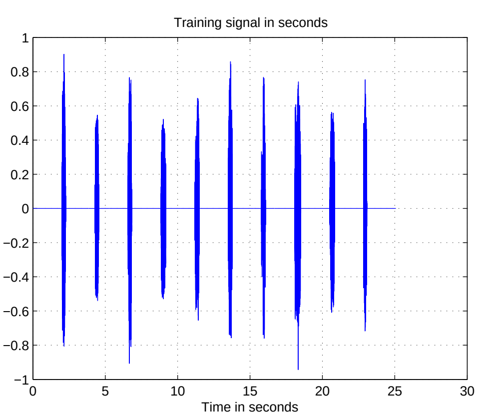


Figure 1. Manatee training signal

RLS predictor model was implemented on each of the models and appropriate filter weights were chosen. Filter weights were chosen based on two conditions, they correspond to the least NMSE values and their trajectories were largely smooth in the considered region. The predictor outputs were obtained for the corresponding test sets with the chosen filter weights and compared to the true validation test set. The models were assessed by evaluating the normalized mean square error values.

Here, filter orders of 3,4,5,6 and forgetting factors of 0.99,0.999,0.9999 and 0.99999 were tested. Filter tracks corresponding to forgetting factor 0.99 were jagged while 0.99999 uses a lot of previous samples to arrive at the next predicted estimate - uses approximately 1.44 seconds of previous data (samples corresponding to the half value of the exponential window),which is pretty high for an audio signal of this nature (quasi-stationary).

Cross validation model 8 was seen to perform the best whose training set included Manatee calls 2,3,4,5,6,8,9,10 and test set included Manatee calls 1,7,9 out of the original training set. The reason for this is revealed by looking at the spectrogram of the corresponding calls and it is observed that calls 8 and 10 capture a lot of frequencies of the Manatee calls while other calls under fit the data. This is explained in depth in the following sections. The variation of NMSE with respect to filter order and forgetting factor for this model is shown in Figure 2. It was observed by looking at the NMSE matrix values that forgetting factors 0.999 and 0.9999 performed pretty much identically but the filter tracks were much smoother for the latter. Forgetting factor of 0.9999 implies that the previous 0.14 seconds of the data is the point where the exponential window is halved which is a reasonable choice. In order comparison, orders 3 and 5 were seen to be the better ones with order 5 model out performing the order model only marginally. Since the difference was not much, there was not enough motivation to pick order 5 over order 3 as smaller orders are preferred for a non stationary signal and is also computationally cheaper. Order 6 was seen to give pretty bad results compared to the others. The matrix values are shown for reference. This was seen to be the general trend for all the cross validation models.

B. *Training*

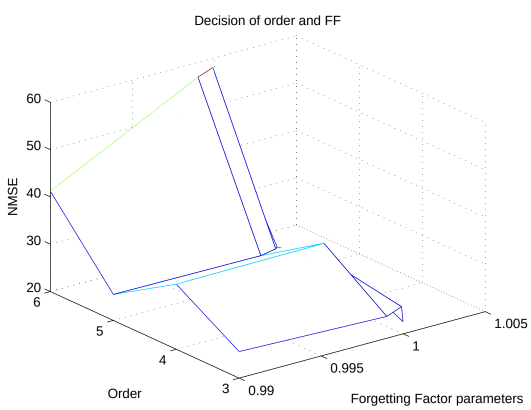
Initially, all the 10 Manatee calls were segmented out individually to arrive at the 10 filter training models. Order 3 and forgetting factor 0.9999 was considered. RLS predictor was run on all the ten calls to arrive at the ten Manatee training models. Similarly, 4 models of noise was implemented. The noise signal was seen to exhibit 3 different frequency patterns. These were used to arrive at 3 noise models. The fourth noise model was implemented by considering the entire noise sequence without segmentation. Therefore, in total there are 10 different Manatee training models while there are 4 different noise models. Again, those filter weights were chosen which correspond to a largely smooth region while resulting in low NMSE values. In this paper, Manatee models 5,8 and 10 will be referred to time and again for comparison as it was observed that Manatee models 5 performed the worst while 8 and 10 generated pretty accurate detections.

In the plot captions, Training models 1 - 10 refers to 10 Manatee calls while Noise model 11 refers to the entire sequence without segmenting and Noise models 12 - 14 are the segmented noise variations i.e In the Figure 3, noise model 11 refers to the whole sequence, noise model 12 refers from start to 0.313 seconds, noise model 13 refers to 0.32 to 1.113 seconds and noise model 14 refers to 1.117 to 1.348 seconds.

For completion, the method used to select filter weights is shown in Figure 4 for call 8 or training model 8. The method used is identical for all the models. Generally, filter weights are not reliable at the beginning of a sequence. Here, the filter weights at 0.4964 seconds were considered since the tracks are smoothly varying at that point and also it corresponds to the second lowest NMSE value. The global minima wasn't considered since it corresponds to the beginning of the sequence and the filter has mostly not adapted yet by that time. The NMSE values are smoothed using a 200 order moving average filter.

C. Testing

The testing data is a 44.1kHz sampled 30 second duration signal. Therefore, the signal is first re sampled to 48kHz using the multirate sampling technique discussed earlier. The test signal is filtered through all combinations of Manatee training model and noise training model arrived at earlier. The normalized mean squared error is computed by normalizing the error with a smoothed version of the test signal squared values at all samples. For this, smoothing was done by using a moving average filter or order 3 (same as the order of the RLS model). The NMSE is of high frequency and consequently it was smoothed using a 3000 order moving average filter. This is done for the noise model as well as the Manatee model. The NMSE difference of these two signals are compared against a threshold to decide if the signal is detected or not. The test signal and it's spectrogram is shown in Figure 5. The spectrogram is generated by dividing the signal into 200 windows using a hamming window and then taking the corresponding next power radix 2 FFT to ensure oversampling. This is the method employed to generate all the spectrograms shown ahead as well. The ground truth of the test signal is arrived at by listening to the audio signal and observing the magnified spectrogram at corresponding location. There were 16 Manatee calls detected and they are show in Table 1.



NMSE values	Ord 3	Ord 4	Ord 5	Ord 6
FF 0.99	25.6273	33.7840	25.4878	41.1998
FF0.999	24.6187	33.9570	25.2456	56.8924
FF 0.9999	25.8331	28.1545	25.8663	58.0706
FF 0.99999	22.5807	28.8702	26.1527	51.5103

Figure 2. Decision of order and forgetting factor by a 3 fold cross validation

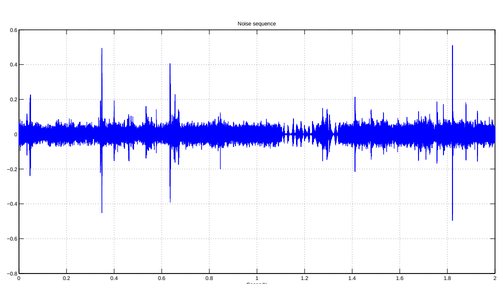


Figure 3. Noise sequence

Call	Start time (s)	End time(s)
1	0.975	1.2
2	3.45	3.675
3	4.65	4.875
4	5.025	5.325
5	5.775	6
6	7.8	8.25
7	8.55	8.85
8	9.75	9.9
9	11.63	11.85
10	14.93	15.38
11	15.6	15.82
12	18.45	18.75
13	19.41	19.88
14	20.63	20.93
15	25.05	25.27
16	25.73	25.95

Table I
GROUND TRUTH FOR THE TEST SIGNAL

D. Model Evaluation

The model is evaluated by plotting the ROC curves which is a variation of probability of detection with probability of false alarm. The thresholds were generated by considering a linearly

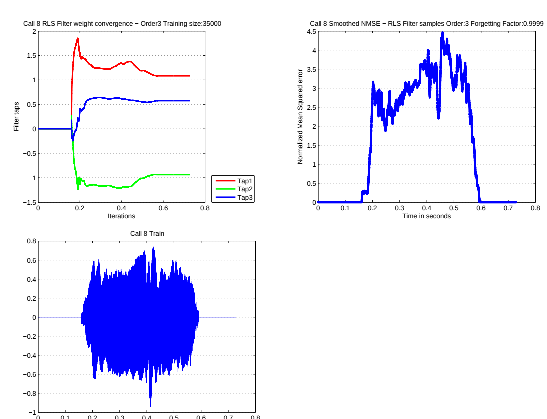


Figure 4. Selection of filter weights

spaced values from the minimum of the error difference signal to the maximum value of the error difference. The error difference refers to the error of the noise model - error of the Manatee model.

1) ROC Curve: As stated earlier, Manatee models 8 and 10 were seen to perform the best while noise models 13 and 14 were seen to perform the best amongst the noise models. Manatee model 5 was seen to perform the worst. Hence the results will be discussed with respect to these models.

The ROC curves for Manatee models 5,8 and 10 for filter order 3 are shown in Figure 6. It is seen that model 8 approximates the ideal ROC curve pretty well.

The reason for the superior performances of models 8 and 10 over 5 can be by looking at the frequency response of the corresponding filter coefficient models. This is shown in Figure 7. The frequency response was generated by evaluating the power spectral densities of the filter taps i.e. taking a 64 point FFT, take it’s absolute squared and evaluating the PSD on a dB scale. It is pretty direct to see that the filter is essentially trying to suppress the low frequency components (noise) and boosting the high frequency portions (Manatee calls). Model 8 attenuates low frequencies upto frequency a little greater than 1kHz while model 10 shows attenuation to frequencies little lower than 1kHz. Model 5 on the other hand attenuates low frequencies only upto 0.5kHz which results in a lower SNR and detection is then a lot harder.

This can also be seen from their corresponding spectrograms as shown in Figure 8. It is noticed that call 5 is much ’neater’ than the other 2 calls i.e. it shows a simple variation of frequency and does not capture all the features required to model the test signal. This leads to under fitting by the model. Call 8 on the other hand shows much more ’dirtier’ variation, the calls are distributed over a lot more frequency bins which is the case with most of the calls in the test signal as seen from the test signal’s spectrogram. Therefore, model 8 is a more able candidate to model the test signal and detect the Manatee calls.

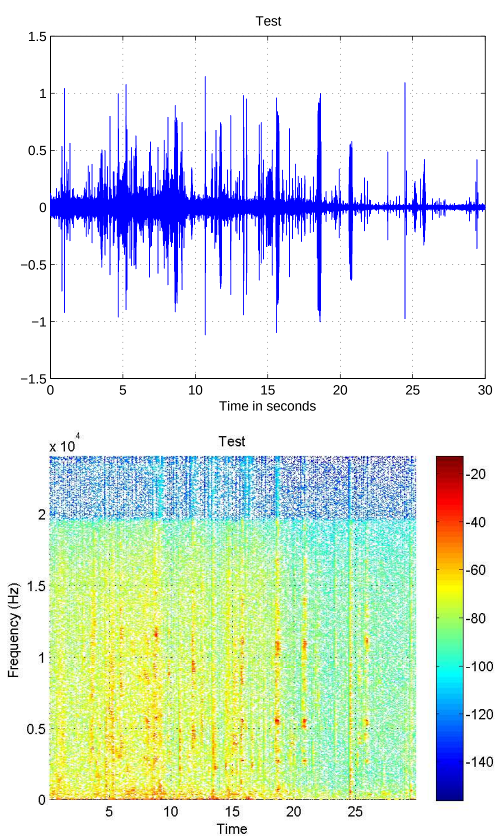


Figure 5. Test signal and its spectrogram

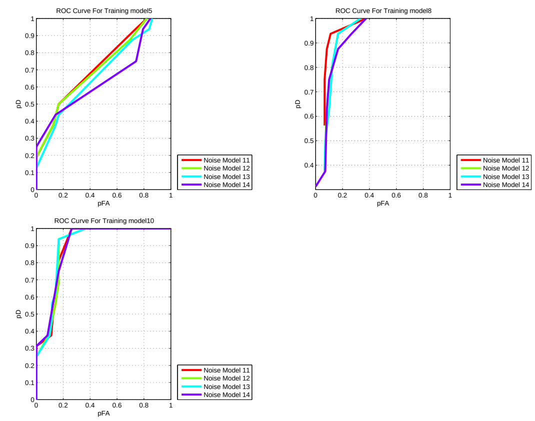


Figure 6. ROC curves for order 3

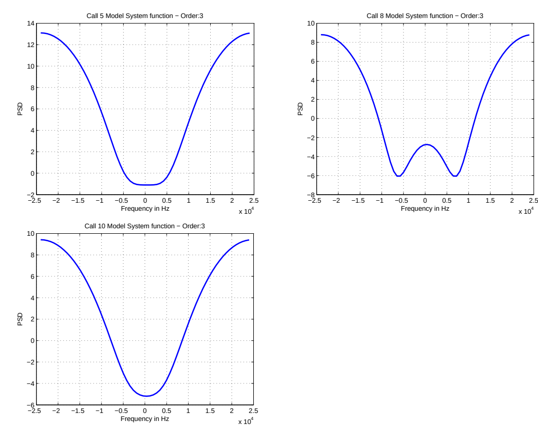


Figure 7. Frequency response of Models 5,8 and 10 in dB

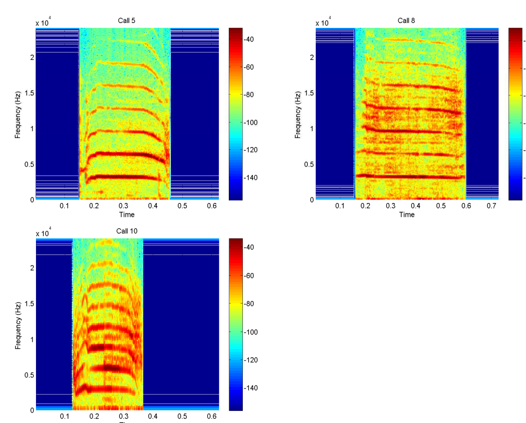


Figure 8. Spectrograms of calls 5,8 and 10

2) Order justification using ROC: It was earlier stated that RLS predictor of order 4 and order 5 does not offer much more than what order 3 offers while order 6 is sub-par comparatively. This can easily be verified from looking at their ROC plots in Figure 9. Only 2 noise models were considered for orders 5 and 6 as these were observed to be the best performing noise models.

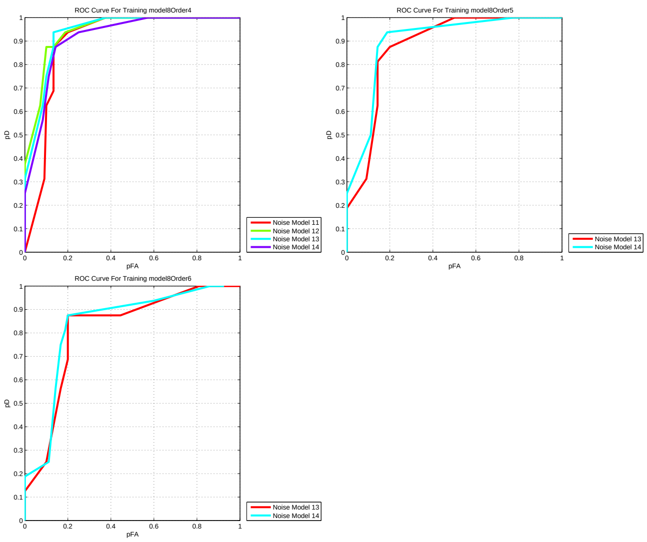


Figure 9. Order justification

3) Detection square wave and error difference signal: Figure 10 shows the Manatee call detection by model 8 with noise model 13 when threshold of the error difference signal is taken to be -0.08. The truth signal is shown for comparison. The error difference and the error signal signal is shown for comparison in Figure 11.

E. Band-pass filtering

It was learnt from [3] that, Manatee calls are concentrated between 1.2kHz and 20kHz. Therefore, bandpass filtering the noise model before training and the test signal could help improve the performance. It should be noted that, the RLS model is exactly trying to do this very thing, so this is just boosting the filter or pushing your filter to go that extra mile. A 10th order bandpass Butterworth filter of cutoff frequencies 1.2kHz and 20kHz was employed here for comparison purposes. Note that in the results shown before this section, no bandpass filter was used.

Figure 12 shows the change in spectrogram of the test signal before and after bandpass filtering. It can be observed that the low frequencies are wiped out after bandpass filtering and exhibits more contrast.

Figure 13 shows the influence of band pass filtering on the detection performance. There is not much difference in the performance because the RLS filter is already doing a good job of suppressing low frequencies.

F. Multiple models consideration

we can consider multiple models for Manatee signal and multiple models for noise for better performance. Then, the integration of the models has to be considered. One way to do this is to compare the performance of each Manatee model sample by sample and use the model which results in better prediction, similarly for noise. The models has to be chosen in such a way that they complement each other i.e. they model different frequencies. This can be done by looking at each model's spectrograms and look for the models that together model the frequency ranges present in the test signal. For noise, it is not as important as the Manatee model. But here, one faces the problem of over fitting the model.

IV. Conclusion

In conclusion, the detection problem considered here is all about overcoming the SNR problems. The filter adapts and models itself to perform band pass filtering of some kind to attenuate the noise and boost the Manatee calls. Manatee Models 8 and 10 were seen to perform the best on account of it's superior low frequency suppressing while Manatee model

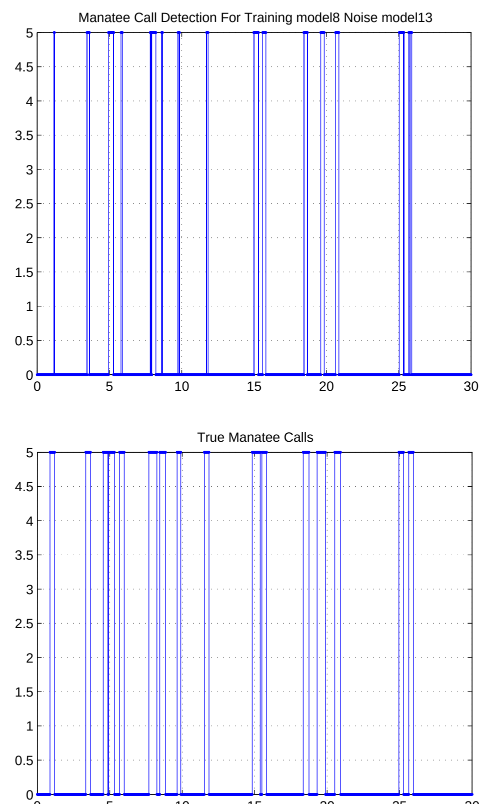


Figure 10. Detection result

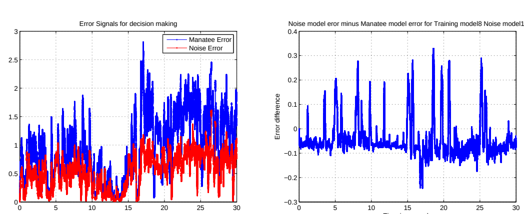


Figure 11. Error signals

5 does not do a very good job of suppressing background noise. Filter parameters of order 3 and forgetting factor 0.9999 were seen to be the optimal parameters. Band pass filtering did not make much difference to the detection performance. Best performances were obtained with Manatee model 8 and Noise model 13.

REFERENCES

[1] Simon Haykin. 1996. Adaptive Filter Theory (3rd Ed.). Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[2] .M. H. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley & Sons, Inc

[3] Zheng Yan, Background Noise Cancellation for Acoustic Detection of Manatee Vocalizations

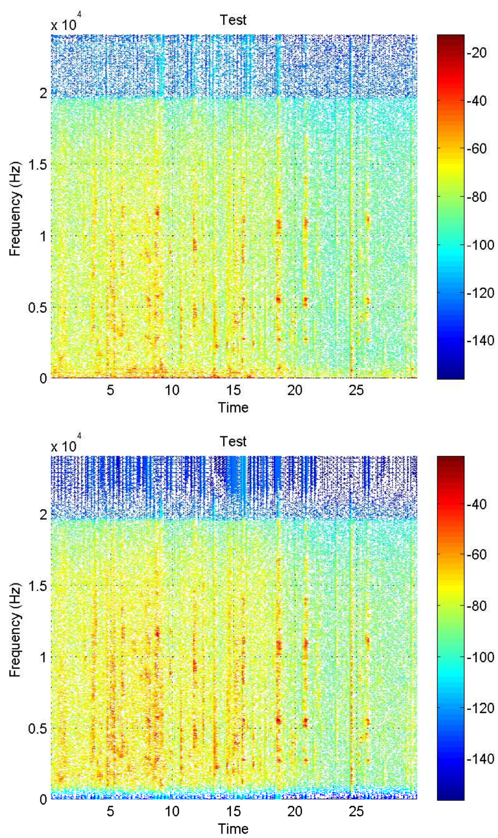


Figure 12. Change in spectrogram of test signal after filtering

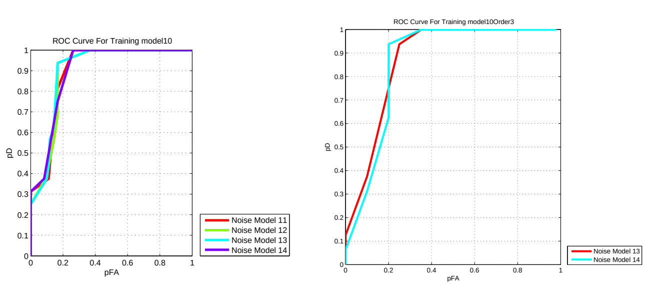


Figure 13. Effect of bandpass filtering on detection performance - Left figure is without band pass filtering and right figure is with bandpass filtering