



**Electrical & Computer Engineering Department**

**EE 6733: Simulation and Digital Analysis of Signals**

**Comparing Fitness Functions in a Genetic Algorithm designed to  
match recorded and simulated EMG signals.**

**Project Report**

**Tolulope Olugbenga (Student ID: 3643581)**

**December 10th, 2019.**

## **Table of Contents**

1 The Main Objective.....	3
1.1 Rationale .....	3
2 Dataset / Existing Implementation .....	3
3 Methods.....	4
3.1 The Genetic Algorithm .....	4
3.1.1 The GA key parameters .....	4
3.2 The Fitness Functions .....	5
3.2.1 The Power Spectrum .....	5
3.2.2 The Amplitude Histogram.....	6
4 Results.....	7
4.1 Comparing the Power Spectrum and Amplitude Histogram.....	8
5 Conclusions and Future Work.....	11
References .....	12

## **1 The Main Objective**

The main objective was to compare fitness functions in a genetic algorithm that was designed to match simulated EMG signals with recorded EMG signals. The two fitness functions to be compared are the amplitude histogram and the power spectrum of signal.

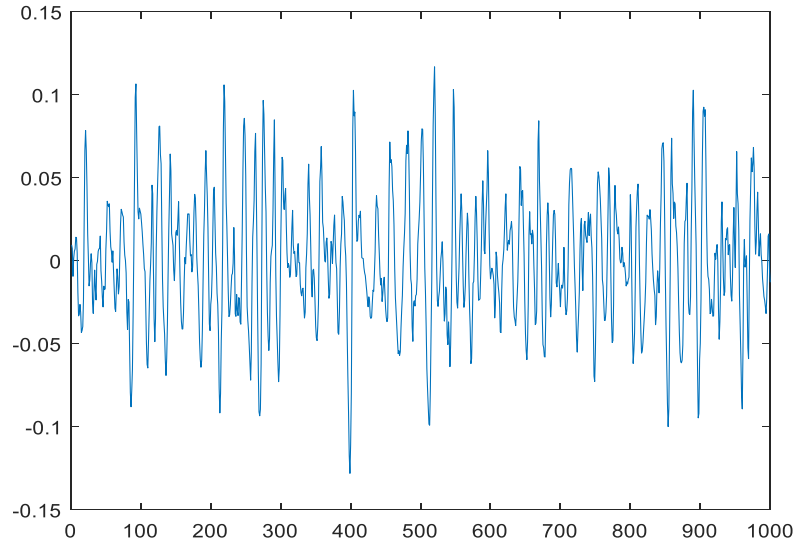
### **1.1 Rationale**

One of the major reasons of this experiment is to show a pure validation of the myosim tool; the aim is to show that the tool can generate realistic signals that match the recorded ones. And if we ever wanted clean samples of EMG data and we don't have enough; the tool can be used to generate more.

## **2 Dataset / Existing Implementation**

Shriram Raghu; who is a PhD student at the University of New Brunswick has worked with myosim earlier in his career; he worked on matching the simulated and recorded signals based on their power spectrums. The signals were collected as part of another SEMG study that focused on investigating variability in SEMG signal features. The signals were measured from the biceps brachii, with electrodes in bipolar configuration, and amplifiers set to bandpass the signal from 20-500Hz. An instrumentation amplifier with CMRR>100dB was used.

Shriram supplied me with the recorded data, his genetic algorithm, and his myosim codes. The record contained a total of 22 recordings taken from 11 subjects at 2 contraction levels; they were 5 secs long signals sampled at 5000Hz, which I later down sampled to 1000Hz for easier analysis of the data. The figure below shows a plot of the first record (1sec).



*Figure 1:- The figure shows a plot of the first second of the first EMG record*

### **3 Methods**

Myosim was used to generate simulated signals using a set of parameters gotten randomly within the range of a normally recorded signal. This was now passed into the genetic algorithm.

#### **3.1 The Genetic Algorithm**

Genetic Algorithms (GA) were first proposed by John Holland in 1975; they are good for optimizing wide range of problems and finding acceptable solutions in a short time. The algorithm begins by arranging the parameters in a chromosome like structure. A typical GA consists of three main steps repeatedly performed in a loop namely: fitness proportionate selection, crossover, and mutation[1]. The GA was used to match a total of 22 recordings taken from 11 subjects at 2 contraction levels.

##### **3.1.1 The GA key parameters**

The aim was focused on four key parameters; the number of fibres per motor unit (N), the number of motor units (M), the depth of the motor unit (D), and the average conduction velocity (CV). The table below shows a summary of the ranges of the parameters used, the population size, crossover rate, and the mutation rate.

Parameter	Value
M	50 – 200
N	100 – 400
D	1 – 15 mm
CV	3 – 6 m/s
Population size	100
Crossover rate	10%
Mutation rate	5%

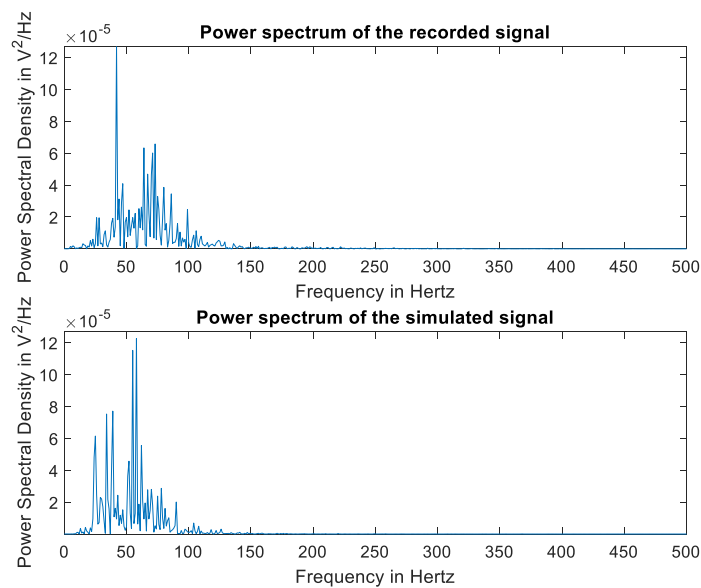
*Table 1:- The table shows the GA parameters*

The GA ran for 5 generations, with a population size of 100, crossover rate of 10%, and a mutation rate of 5%.

## **3.2 The Fitness Functions**

### **3.2.1 The Power Spectrum**

The first fitness function which was originally created by Shriram compared the Power Spectrum of the recorded and simulated signal. A match was found when they both have similar power spectral densities (PSDs). To measure the similarity of two spectra, the mean absolute error (MAE) was computed and normalized by the total power of the record. The fitness score was derived from the MAE between the PSDs. If the fitness score was below a specified threshold; the simulated signal was saved as a possible solution. More readings can be found here [1].

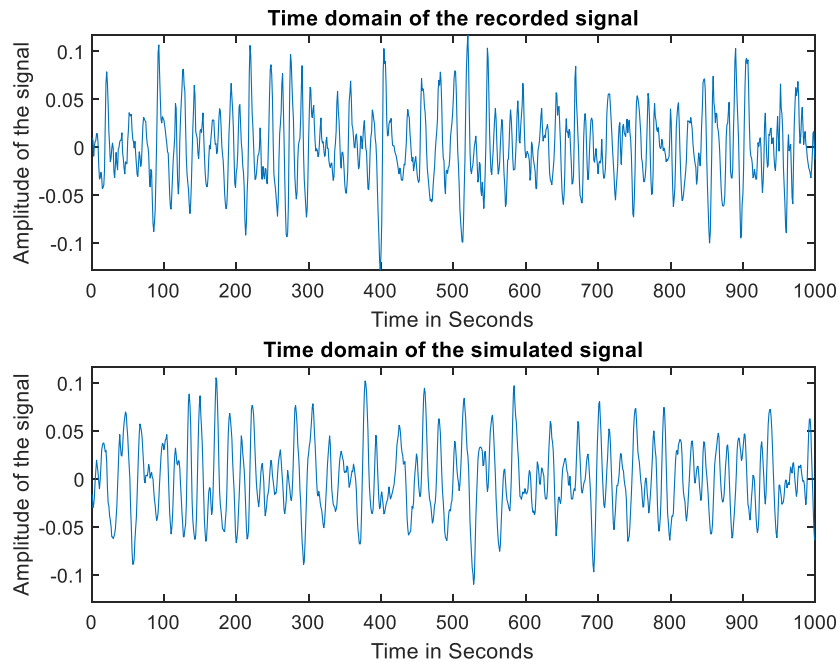


*Figure 2:- The figure shows the power spectrum of the first record (1sec) and a simulated solution*

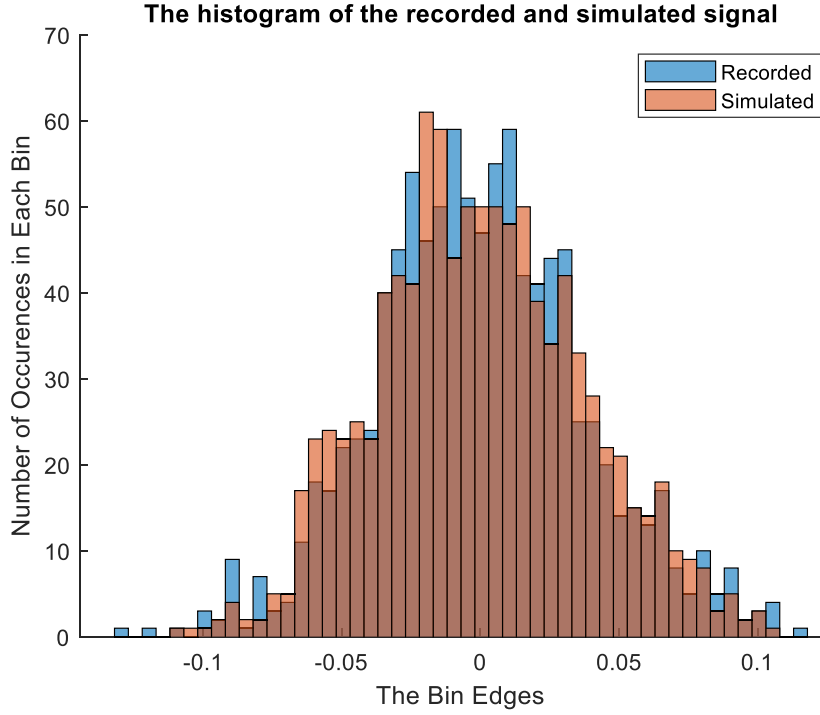
### **3.2.2 The Amplitude Histogram**

The second fitness function compares the histogram of the amplitudes of the recorded and the simulated signals. The comparison was done with one second segments of the signals and a bin count of 50. I got the probability of each bin in the recorded signal, and the probability in each bin of the simulated signal based on the bin edges of the recorded signal; this makes the comparison more efficient. I used the Bhattacharyya Coefficient (BC) to get a value between 0 and 1 based on how close their histogram probabilities were, 0 means not similar, and 1 means very similar. BC is calculated by taking the sum of the square root of the product of the probabilities, i.e.  $\sum (\sqrt{P \cdot Q})$ , where P and Q are the probabilities of the recorded and simulated signals respectively.

The value gotten from BC was defined as the fitness score; if the fitness score was greater than a particular threshold, the signal is saved in a particular directory and the loop continues. A threshold of 0.975 were used for all records except record 18, which was 0.97. When the GA finishes with a record; the simulated signal of the record with the highest threshold is saved in a different directory. The figures below shows the plot and the histogram of the first record and its best solution; the solution had a fitness score of 0.9909, and the parameters were; Fibres, MUAP, Offset, and CV of, 360, 180, 10.25, and 4.4 respectively.



*Figure 3:- The figure shows the plots of a recorded and simulated signal (1sec)*



*Figure 4:- The figure shows the histogram distributions of a recorded and simulated signals (1sec)*

To verify the results gotten from the matching process, four time domain features were computed for both the recorded and the simulated signals; the Mean Absolute Value (MAV), Slope Sign Changes (SSC), Waveform Length (WL), and Zero Crossings (ZC). The Mean Frequency (MF) of the signals were also computed.

## **4 Results**

The histogram function was used on a total of 22 records with a sampling rate of 1000Hz, threshold of 0.975 except for record 18 which was 0.97; the GA ran for 5 generations, with a population size of 100, crossover rate of 10%, and a mutation rate of 5%. Sometimes the GA had to run more than once to get a better solution for the records. The signals with the best fitness scores were saved in a separate directory; the table below gives a summary of the best solutions gotten for all the 22 records; it shows the fitness score and the key parameters.

Record No.	Fitness Score	Fibres	MUAP	Offset	CV
1	0.9909	360	180	10.25	4.4
2	0.9869	300	90	6.25	5.6
3	0.9835	380	200	5.25	4
4	0.9832	140	160	2.5	5.4
5	0.9855	360	180	6.5	4.4
6	0.9824	380	200	7	5.8
7	0.9867	320	160	9.75	5
8	0.9854	140	170	5.25	5.8
9	0.9840	360	60	4.75	5
10	0.9906	140	160	3.5	6
11	0.9834	120	190	11	6
12	0.9843	280	60	9	5.8
13	0.9859	400	50	6	4
14	0.9819	400	170	7.5	5.2
15	0.9863	280	70	3.25	3
16	0.9875	260	150	8	5
17	0.9781	340	130	10.75	4.2
18	0.9748	180	170	7.75	5.8
19	0.9890	120	100	3.25	4.2
20	0.9874	380	80	3.75	3.2
21	0.9843	100	150	7.25	5.8
22	0.9808	340	180	7.5	5.4

*Table 2:- The table shows the parameters of the best solutions gotten for each record*

#### **4.1 Comparing the Power Spectrum and Amplitude Histogram**

The solutions for the power spectrum fitness function were gotten from Shriram. Using the time domain features and the mean frequency that were listed in the last section. I made plots based on the recorded signal, and the solutions gotten from the power spectrum and amplitude histogram fitness functions. The figures below shows the Mean Absolute Value (MAV), Mean Frequency, Slope Sign Changes (SSC), Waveform Length (WL), and Zero Crossings (ZC) for all records respectively. The features were calculated at one sec segments which sums up to five features values; the black circles, the blue crosses, and the red diamonds, are from the recorded signal, the amplitude histogram solutions, and the power spectrum solutions respectively.



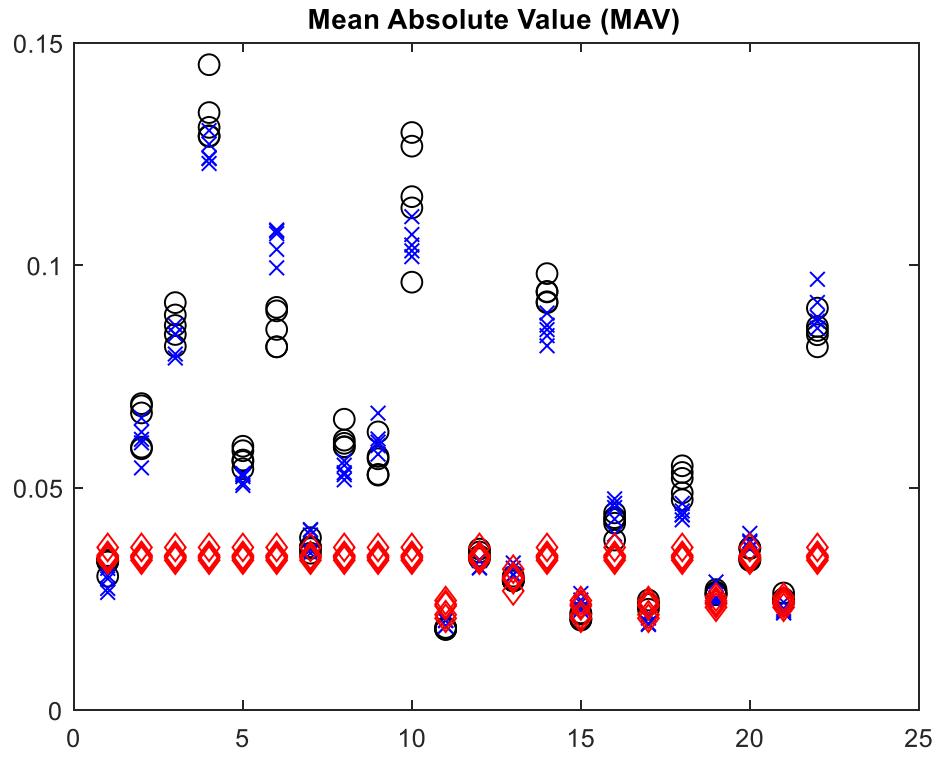


Figure 5:- The figures shows a plot of the mean absolute value.

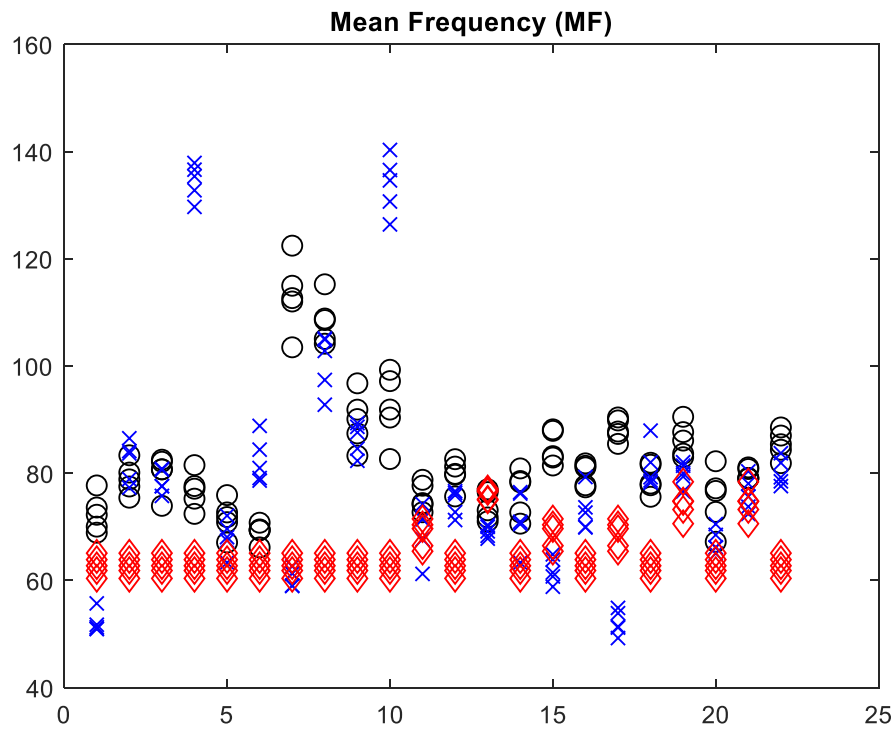


Figure 6:- The figure shows a plot of the mean frequency.

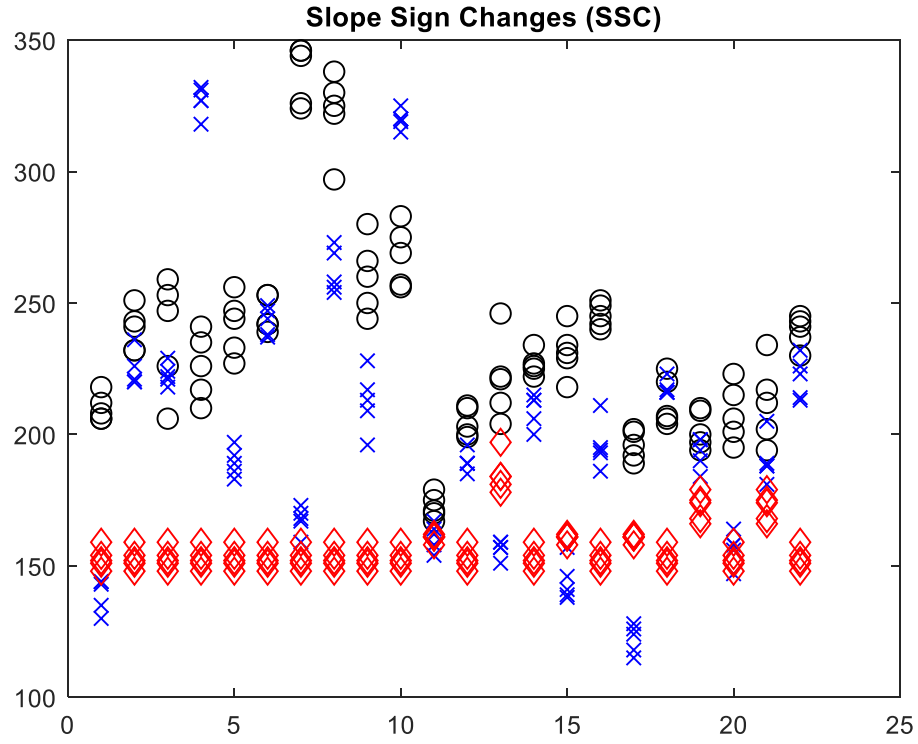


Figure 7:- The figure shows a plot of the slope sign changes.

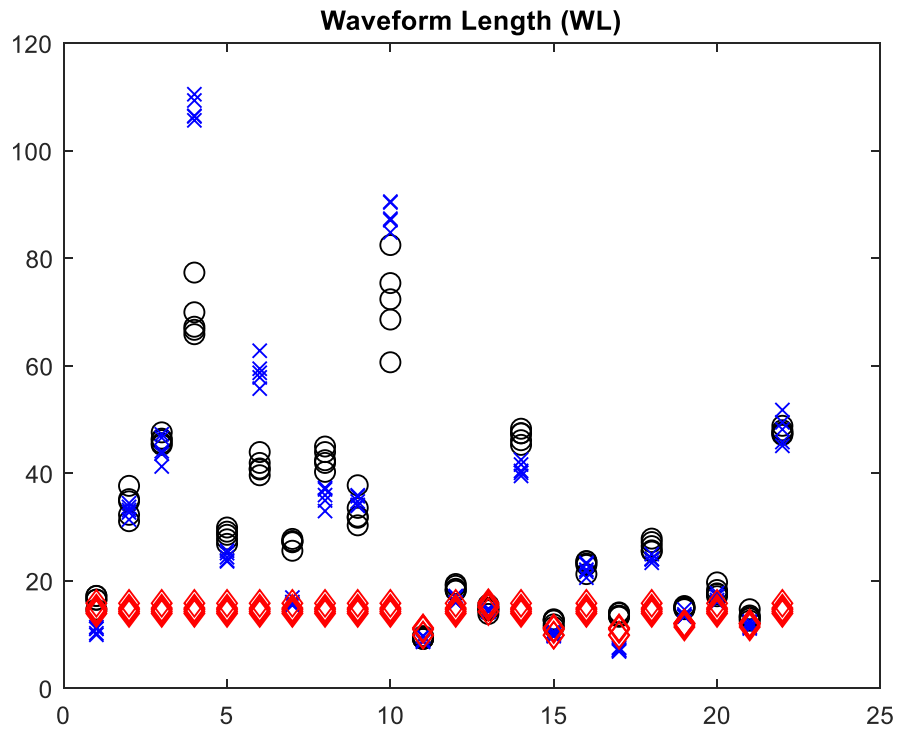


Figure 8:- The figure shows the plot of the waveform length.

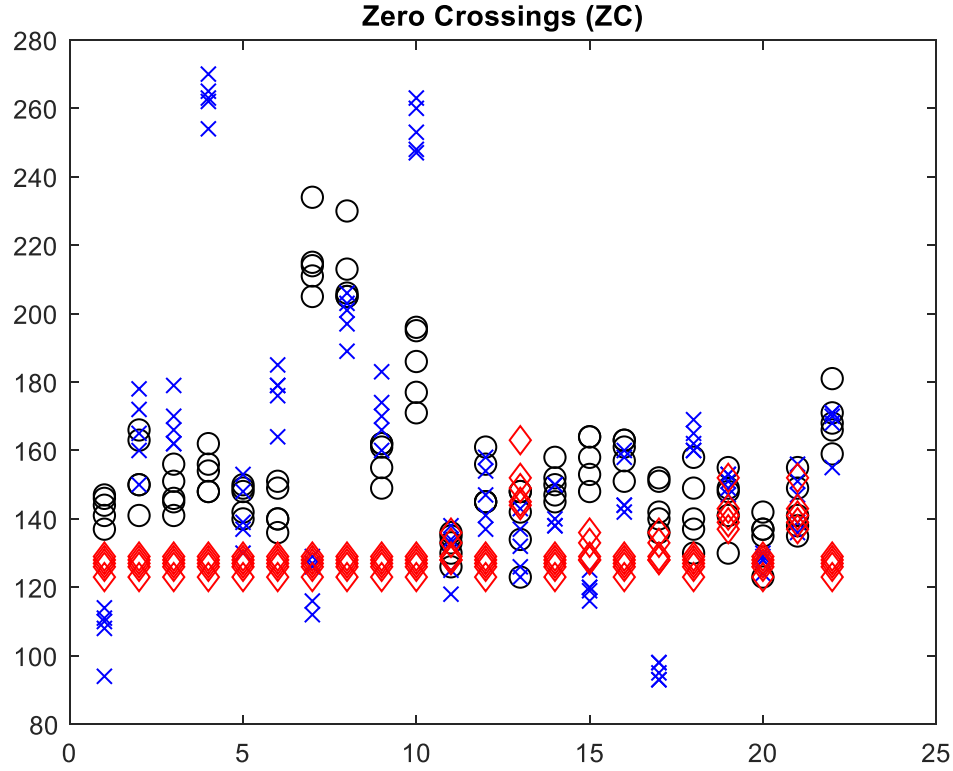


Figure 9:- The figure shows a plot of the zero crossings.

## **5 Conclusions and Future Work**

On the plots above; the black circles, the blue crosses, and the red diamonds, are for the recorded signal, the amplitude histogram solutions, and the power spectrum solutions respectively. They both had some problems in matching some features of some of the records. The power spectrum solutions matched all the features of record 11 and 13 correctly, which is about 9% of the available records; while the amplitude histogram solutions matched all the features of record 2, 11, 12, 14, 18, 19, 21, and 22 correctly, which is more than 36% of the available records. Therefore, we can say that the amplitude histogram solutions did a better job at matching the records features than the power spectrum solutions.

For future work, the GA could be modified a little, to calculate the features values of the best solution and if it doesn't fall in an acceptable range. A population could be created with the best solution and next few best solutions, and manipulated until we get a better solution in the range.

## **References**

- [1] S. T. P. Raghu and Y. Shi, “Validating an Electromyography Simulation Tool using a Genetic Algorithm to Match Simulation Parameters with Recorded Data,” 2017.
- [2] D. Hermawanto, “Genetic Algorithm for Solving Simple Mathematical Equality Problem,” 2013.
- [3] U. Bodenhofer and Q. Ai, “Genetic Algorithms : Theory and Applications Genetic Algorithms : Theory and Applications,” May, 2014.