

MSE 491 - Lab 1

Using LDA to classify EMG sensor values

Harkirat Gill

301219830, hsg12@sfu.ca

Table of Contents

[Feature extraction 3](#_Toc517372889)

[RMS 3](#_Toc517372890)

[Waveform Length 3](#_Toc517372891)

[AR coefficients 4](#_Toc517372892)

[Classifier 4](#_Toc517372893)

[Results 4](#_Toc517372894)

[Accuracy 4](#_Toc517372895)

[Plots 6](#_Toc517372896)

[Raw Data 6](#_Toc517372897)

[Window Size = 100 7](#_Toc517372898)

[Window Size = 1000 11](#_Toc517372899)

[Window Size = 50 15](#_Toc517372900)

[Window size = 200 20](#_Toc517372901)

Table of Figures

[Figure 1 Raw Values 6](#_Toc517372863)

[Figure 2 RMS, window size 100 7](#_Toc517372864)

[Figure 3 Waveform length, window size 100 8](#_Toc517372865)

[Figure 4 AR, window size 100 9](#_Toc517372866)

[Figure 5 Classified Train vs Labeled, window size 100 10](#_Toc517372867)

[Figure 6 Classified Test vs Labeled, window size 100 10](#_Toc517372868)

[Figure 7 Accuracy of training and testing dataset, window size 100 11](#_Toc517372869)

[Figure 8 RMS, window size 1000 11](#_Toc517372870)

[Figure 9 Waveform, window size 1000 12](#_Toc517372871)

[Figure 10 AR, window size 1000 13](#_Toc517372872)

[Figure 11 Classified test vs Labeled data, window size 1000 14](#_Toc517372873)

[Figure 12 Classified train vs Labeled data, window size 1000 14](#_Toc517372874)

[Figure 13 Accuracy of training and test dataset, window size 1000 15](#_Toc517372875)

[Figure 14 RMS, window size 50 15](#_Toc517372876)

[Figure 15 Waveform, window size 50 16](#_Toc517372877)

[Figure 16 AR, window size 50 17](#_Toc517372878)

[Figure 17 Classified test data vs Labeled, window size 50 18](#_Toc517372879)

[Figure 18 Classified train data vs Labeled, window 50 19](#_Toc517372880)

[Figure 19 Accuracy of training and testing dataset, window size 50 19](#_Toc517372881)

[Figure 20 RMS, window size 200 20](#_Toc517372882)

[Figure 21 Waveform, window size 200 21](#_Toc517372883)

[Figure 22 AR, window size 200 22](#_Toc517372884)

[Figure 23 Classified train data vs Labeled data, window size 200 23](#_Toc517372885)

[Figure 24 Classified test data vs Labeled data, window size 200 23](#_Toc517372886)

[Figure 25 Accuracy of training and testing dataset, window set 200 24](#_Toc517372887)

[Figure 26 Accuracy after discarding half of training dataset, window size 200 24](#_Toc517372888)

# Feature extraction

## RMS

function result = rms(data, window\_size, moving\_window)

result = [];

for i = 1:window\_size:size(data,1)

a = sqrt(1/window\_size\*sum(data(i:i+window\_size-1, :).^2));

result = [result; repmat(a,window\_size,1)];

end

We traverse through the training data and calculate RMS of all 8 channels over a window size. Variable ‘a’ takes in the RMS value of all 8 channels. Then we repeat the vector ‘a’ upto the window size and concat it to the ‘result’ matrix.

At the end of for-loop, the result contains the rms values of entire data over static windows with size of window\_size.

## Waveform Length

for i = 2:window\_size:size(data,1)

a = sum(abs( data(i:i+window\_size-2,:) - data(i-1:i+window\_size-3,:) ));

result = [result; repmat(a,window\_size,1)];

end

Similar format applies to waveform length. ‘a’ takes the waveform of a fixed static window and ‘result’ returns the entire waveform of all static windows with no overlap.

## AR coefficients

result = [];

for i = 1:window\_size:size(data,1)

for sensor = 1:1:8

ar\_data = data(i:i+window\_size-1, sensor);

AR\_coeff = aryule(ar\_data , p);

a(sensor) = -AR\_coeff(2:end) \* ar\_data(end:-1:end-p+1);

end

result = [result; repmat(a,window\_size,1)];

end

‘ar\_data represents the data over which we want to calculate the AR coefficients. ‘AR\_coeff’ represents the AR coeffiecients. AR\_coeff is 5x8 long (1 normalized vector + 4 order x 8 sensors)

‘a’ intakes the xi+1 by utilizing the AR coefficients and last 4 data points from ar\_data.

# Classifier

trainer = fitcdiscr([xTrain\_rms], yTrain);

classified\_train = predict(trainer, [xTrain\_rms]);

accuracy\_train = sum(yTrain==classified\_train)/size(yTrain,1)\*100.0

classified\_test = predict(trainer, [xTest\_rms]);

accuracy\_test = sum(yTest==classified\_test)/size(yTest,1)\*100.0

The trainer that yields the highest accuracy is trained using only rms values. Using waveforms or my perception of AR values decreases accuracy significantly.

# Results

## Accuracy

|  |  |  |
| --- | --- | --- |
| **Window Size** | **Accuracy on Training Set** | **Accuracy on Testing Set** |
| 100 | 86.5% | 68% |
| 1000 | 44% | 20% |
| 50 | 83.25% | 59% |
| 200 | 90% | 80% |

As we note, window size 200 gives the highest accuracy. This is because the labels, i.e., labels in yTrain and yTest, are 200 ms long.

Additionally, I was able to achieve higher accuracy by discarding half of the training dataset.

xTrain = xTrain(1:10000,:);

yTrain = yTrain(1:10000,:);

By utilizing the above code, I discarded half of the training dataset and the accuracy after doing so was the following:

|  |  |  |
| --- | --- | --- |
| **Window Size** | **Accuracy on Training Set** | **Accuracy on Testing Set** |
| 200 | 86% | 84% |

This might be because dataset of 20,000 might be overfitting our classifier. The more data is introduced to our classifier, the more variance will exist within classes. This variance might yield a different plane on which the variance within classes might overlap between classes.

## Plots

### Raw Data

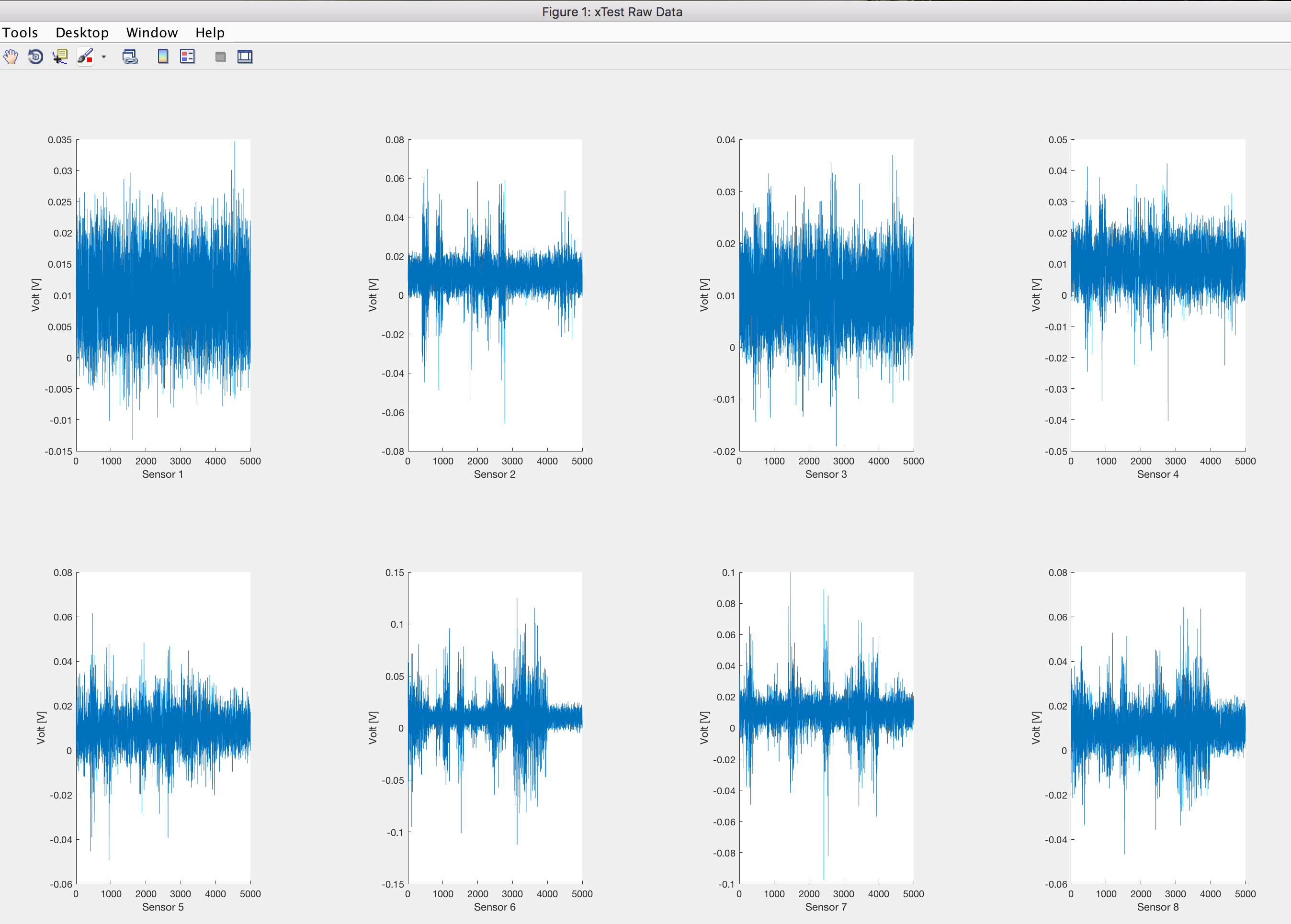


Figure 1 Raw Values

### Window Size = 100

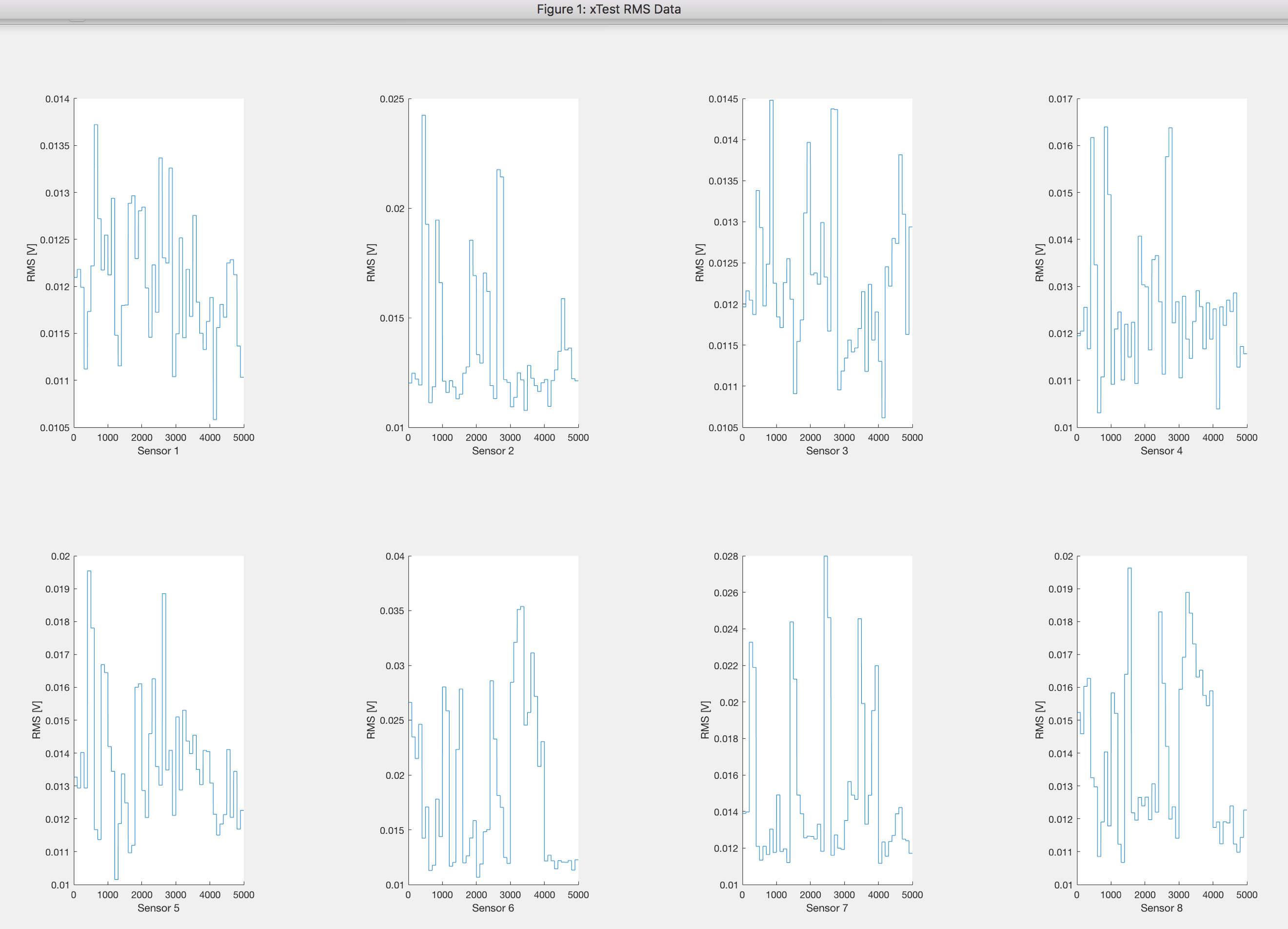


Figure 2 RMS, window size 100

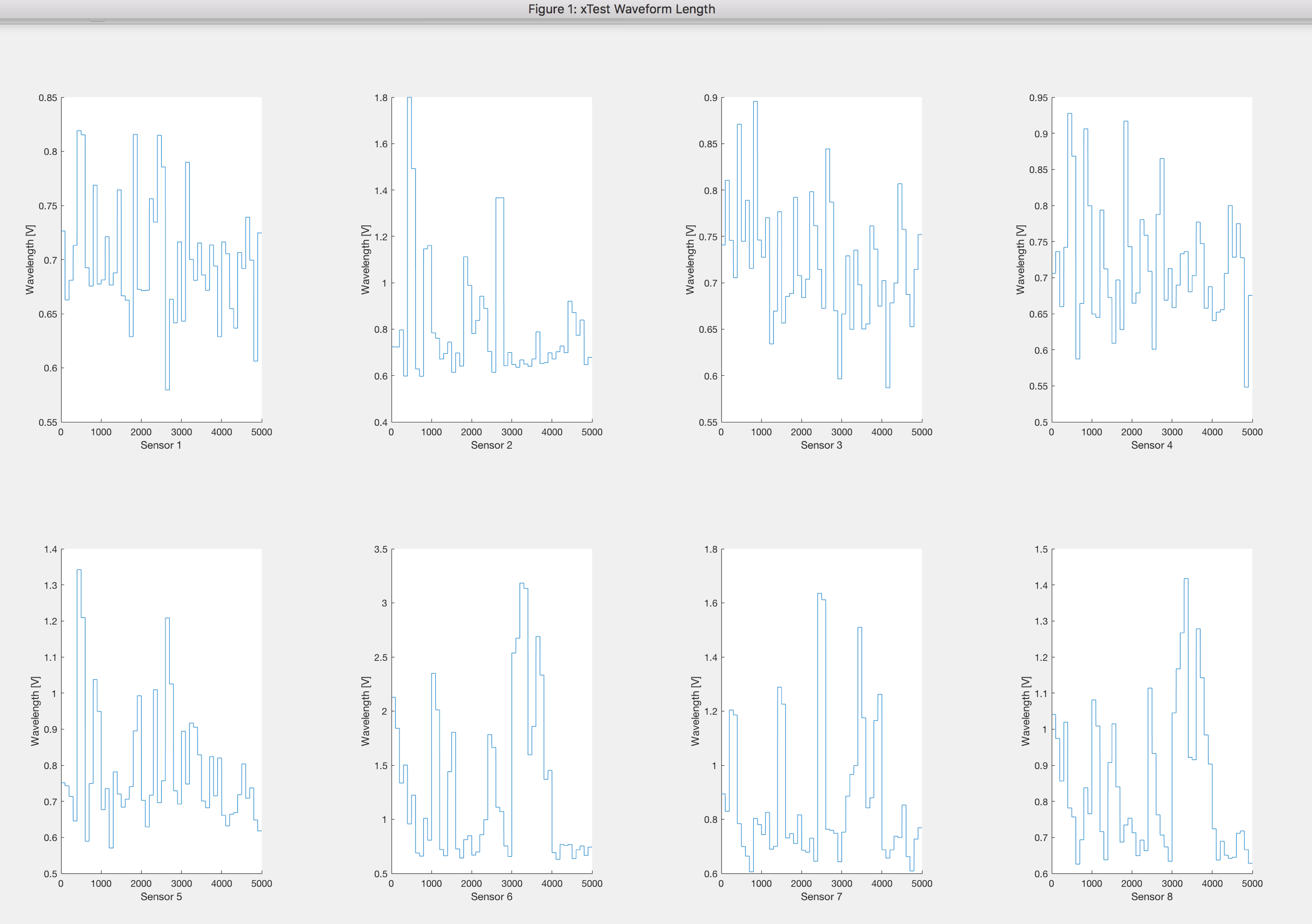


Figure 3 Waveform length, window size 100



Figure 4 AR, window size 100

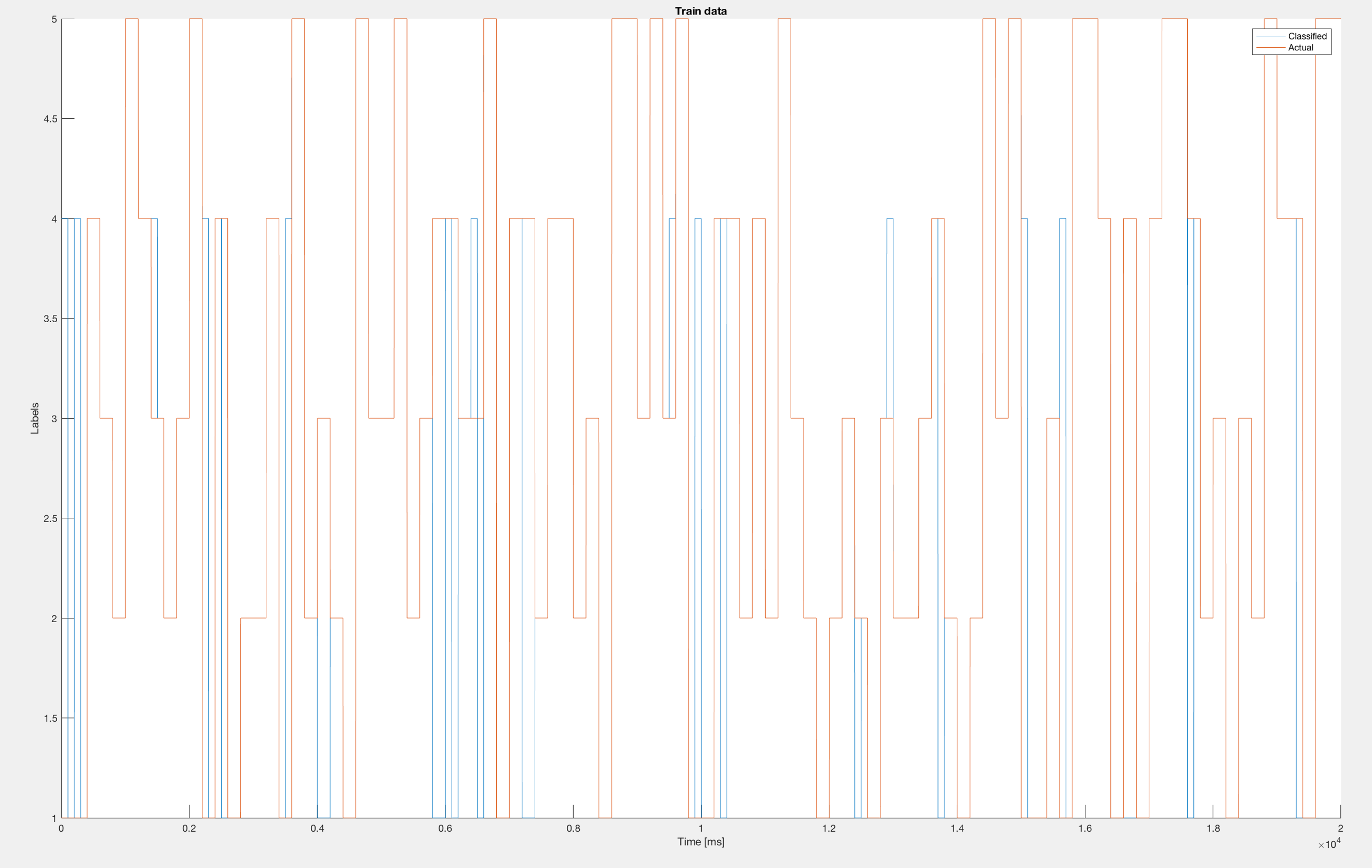


Figure 5 Classified Train vs Labeled, window size 100

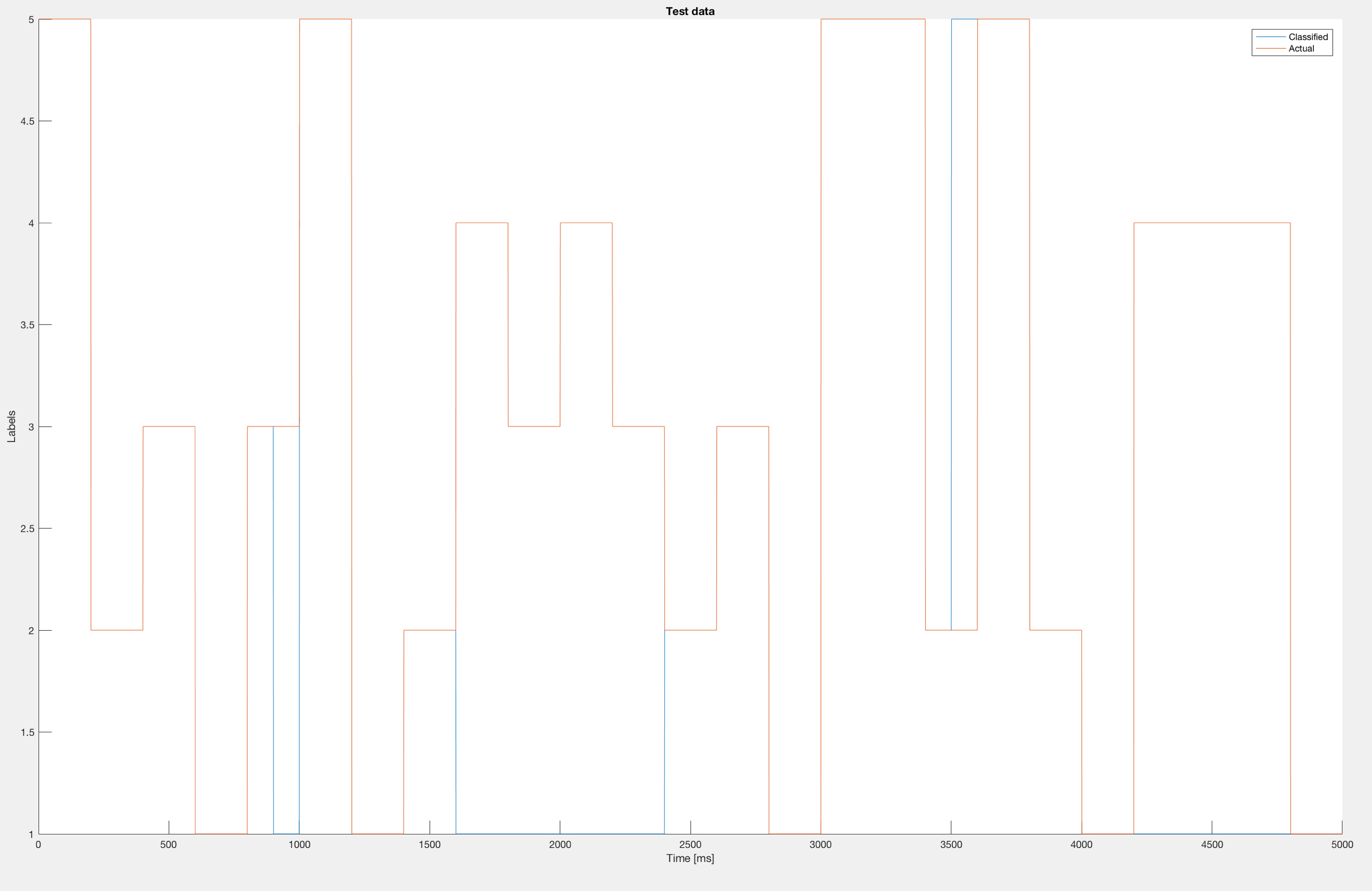


Figure 6 Classified Test vs Labeled, window size 100

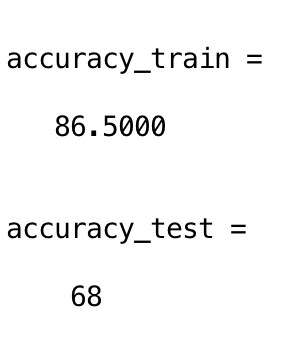


Figure 7 Accuracy of training and testing dataset, window size 100

### Window Size = 1000

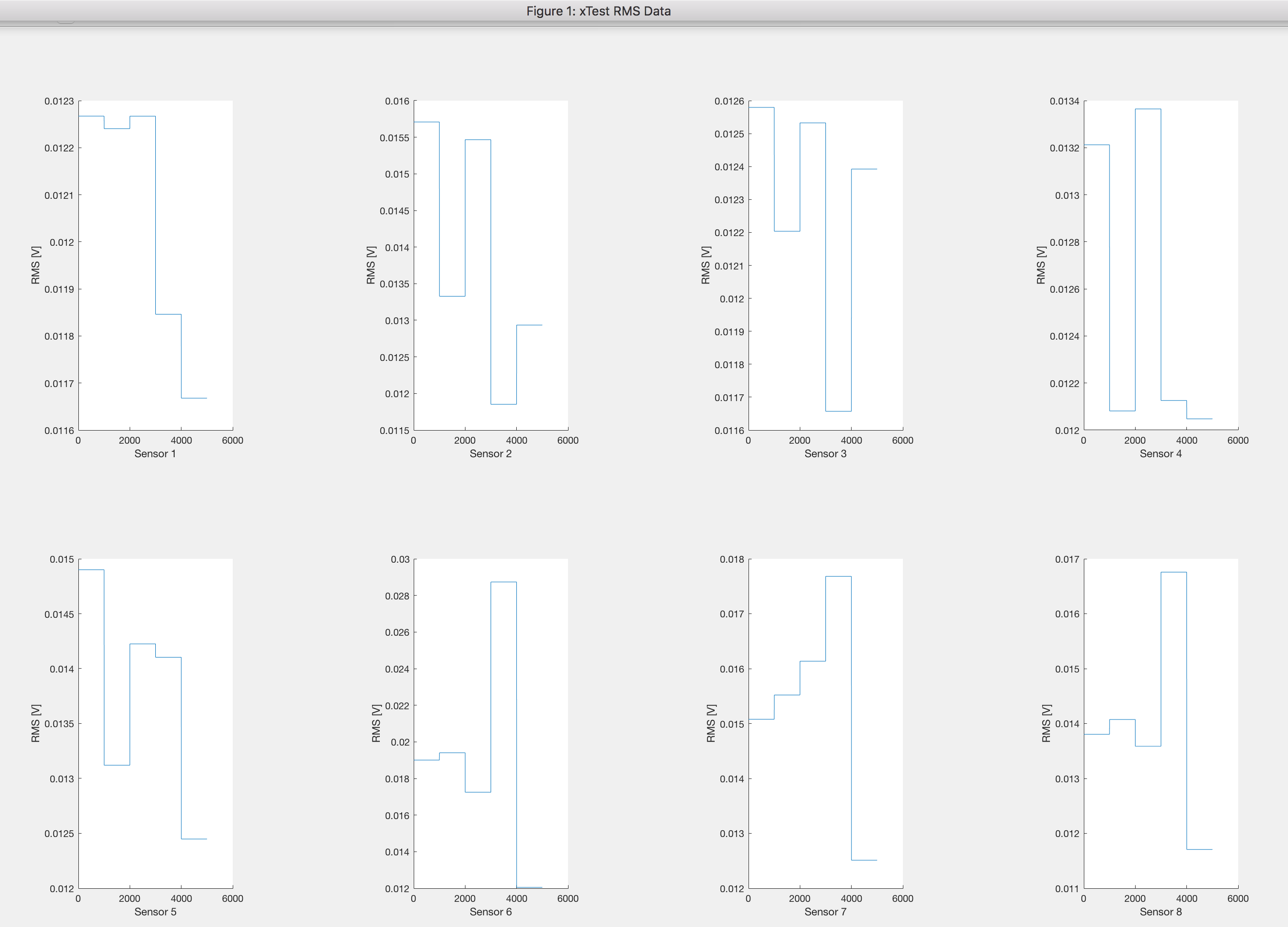


Figure 8 RMS, window size 1000

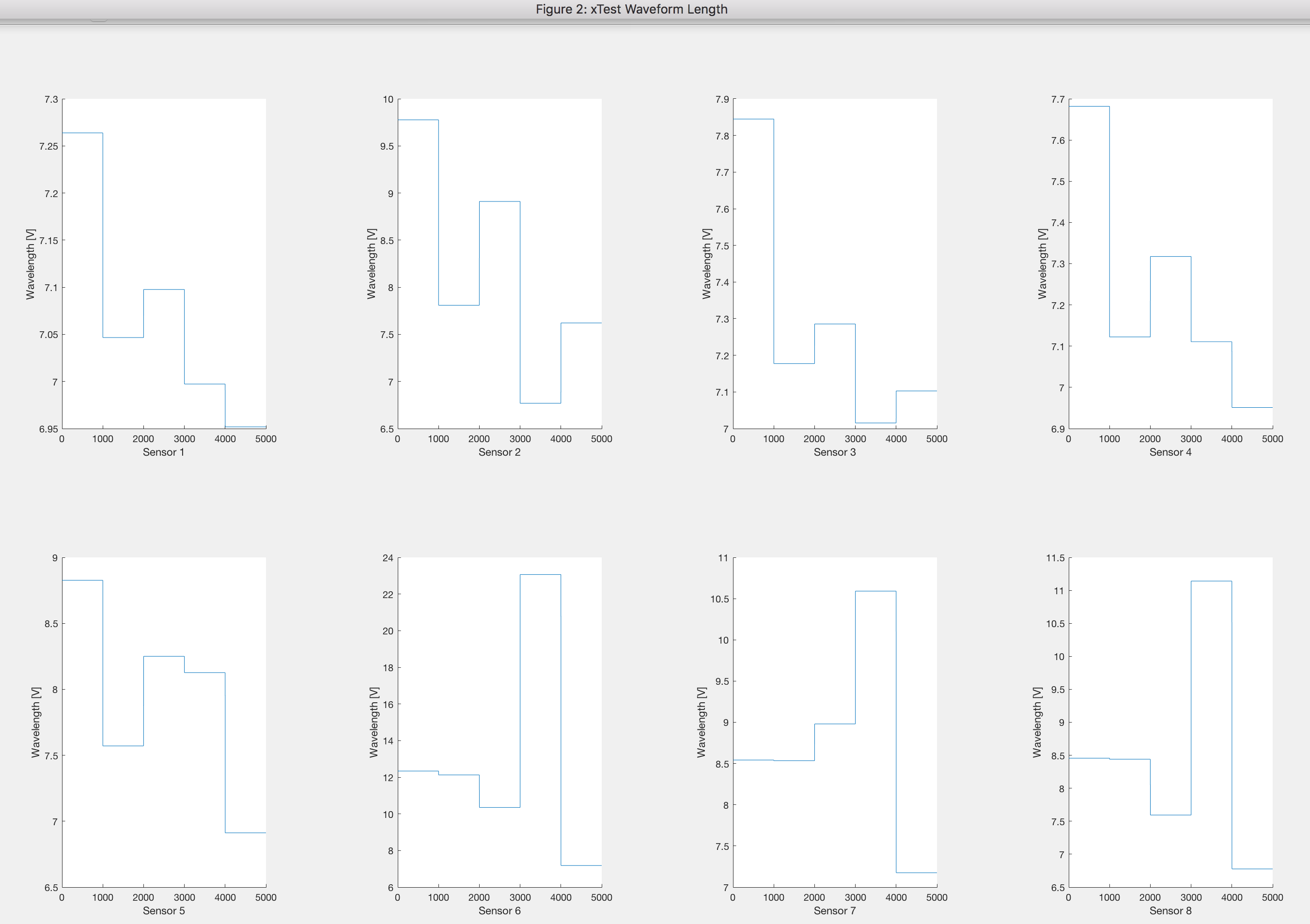


Figure 9 Waveform, window size 1000

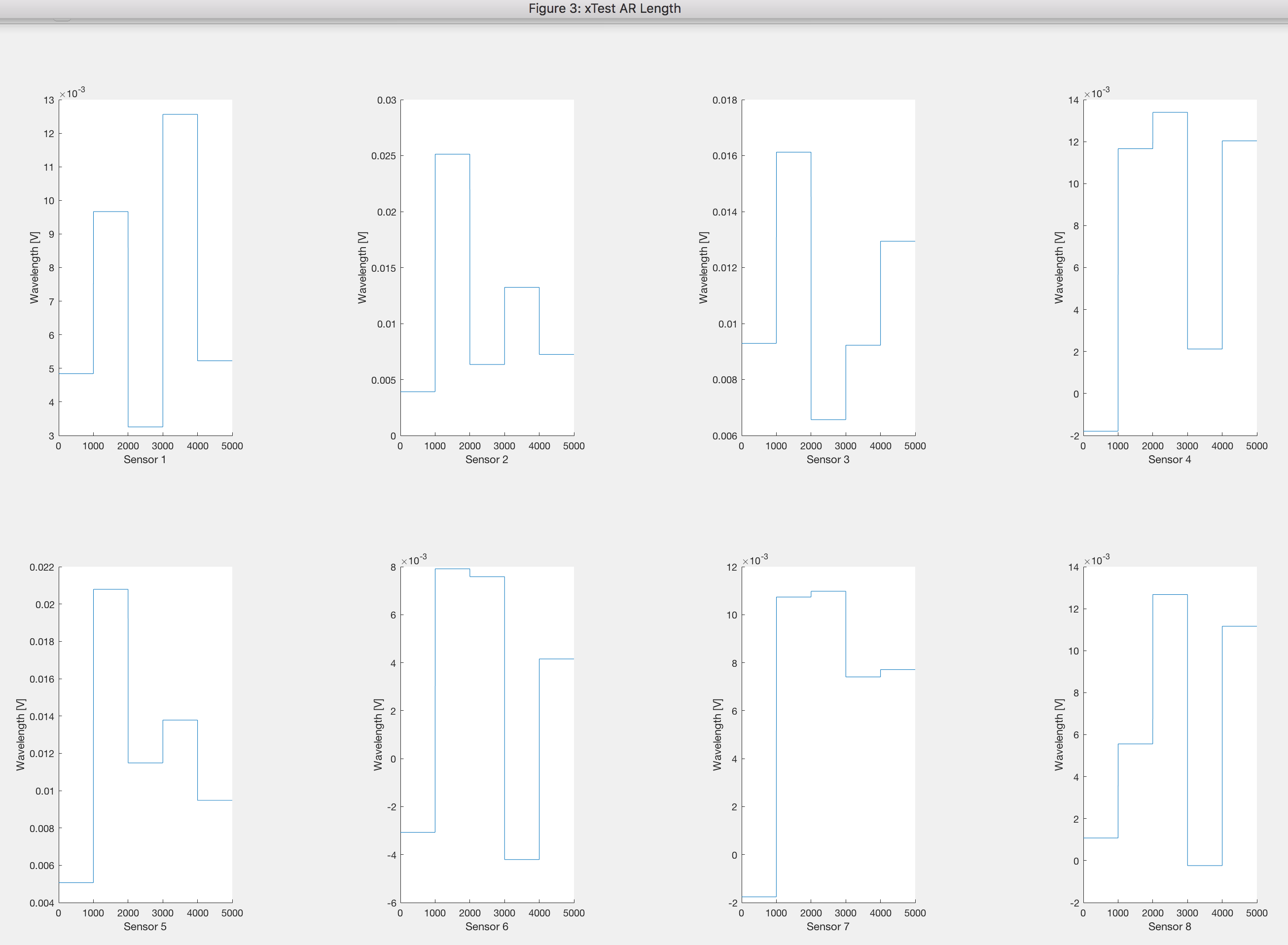


Figure 10 AR, window size 1000

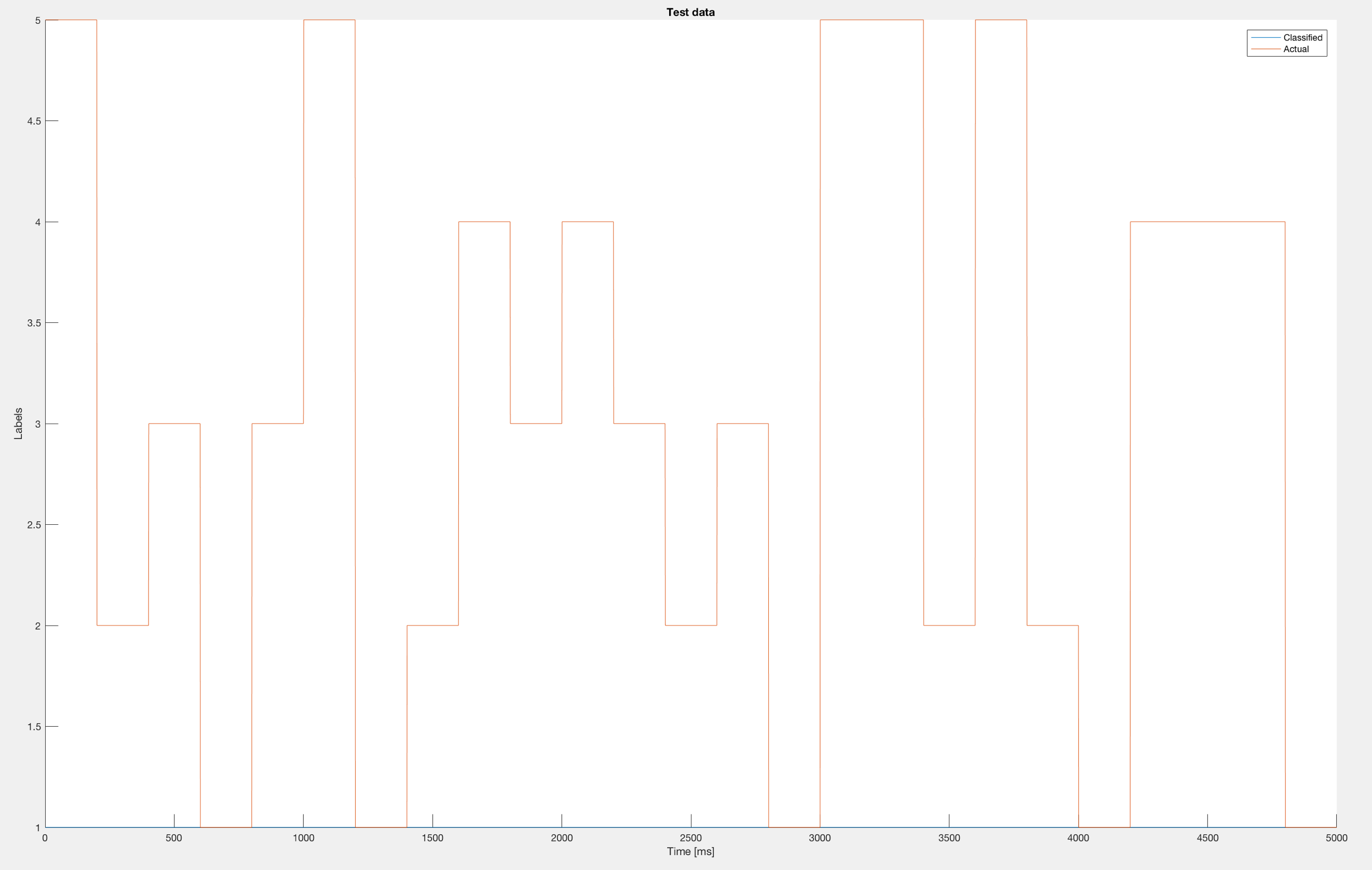


Figure 11 Classified test vs Labeled data, window size 1000

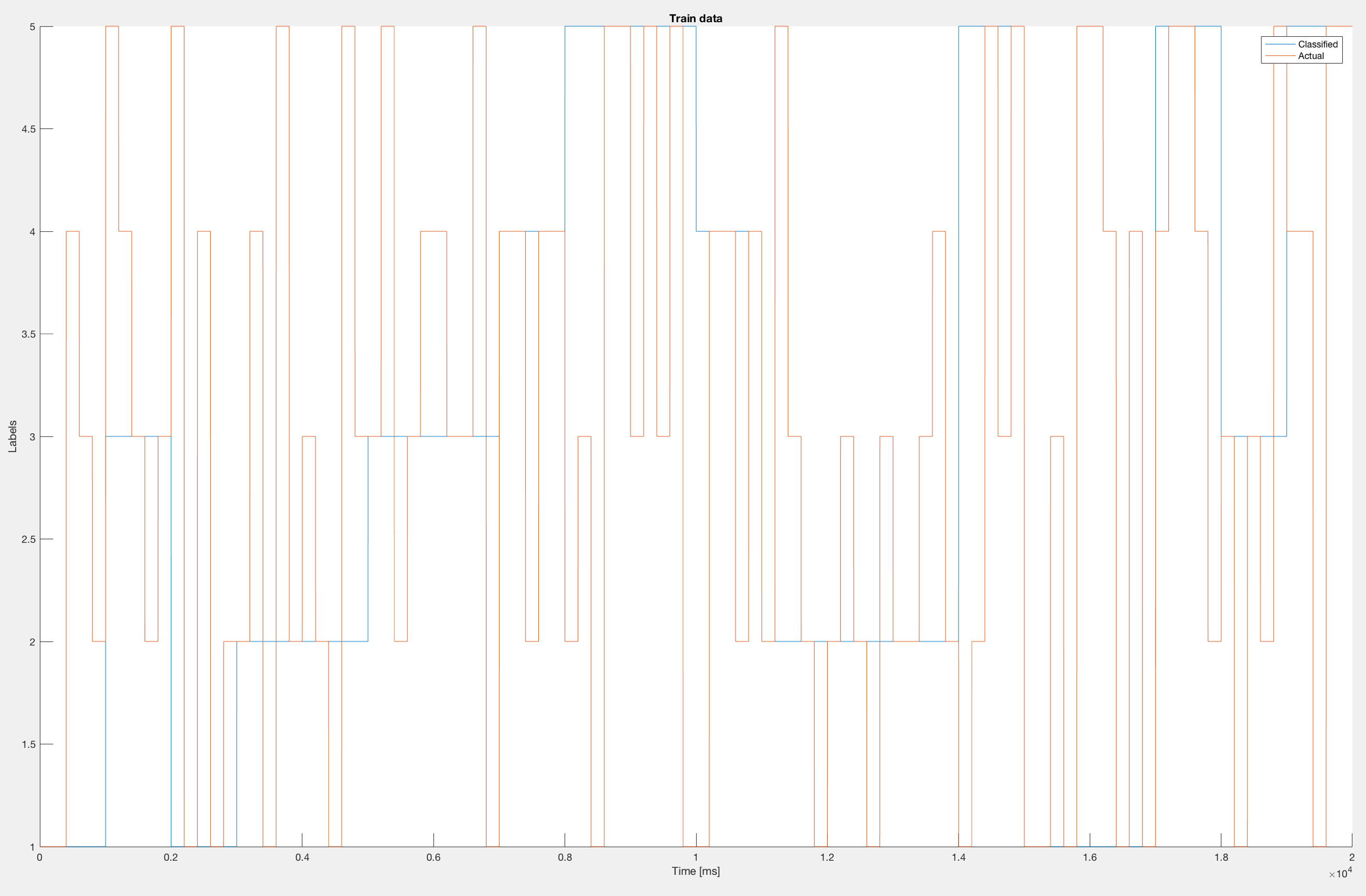


Figure 12 Classified train vs Labeled data, window size 1000

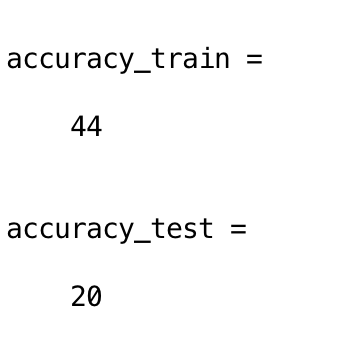


Figure 13 Accuracy of training and test dataset, window size 1000

### Window Size = 50

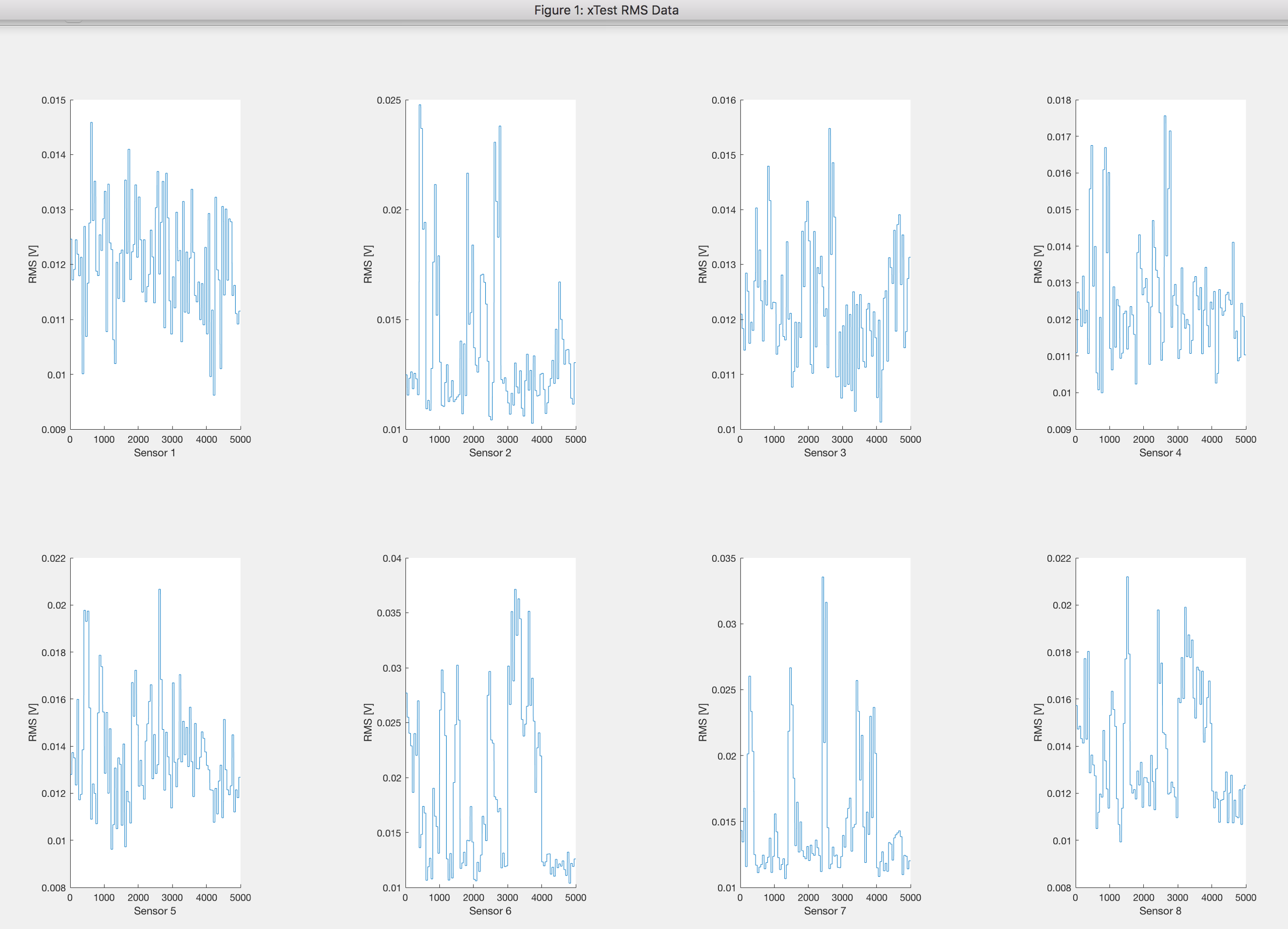


Figure 14 RMS, window size 50

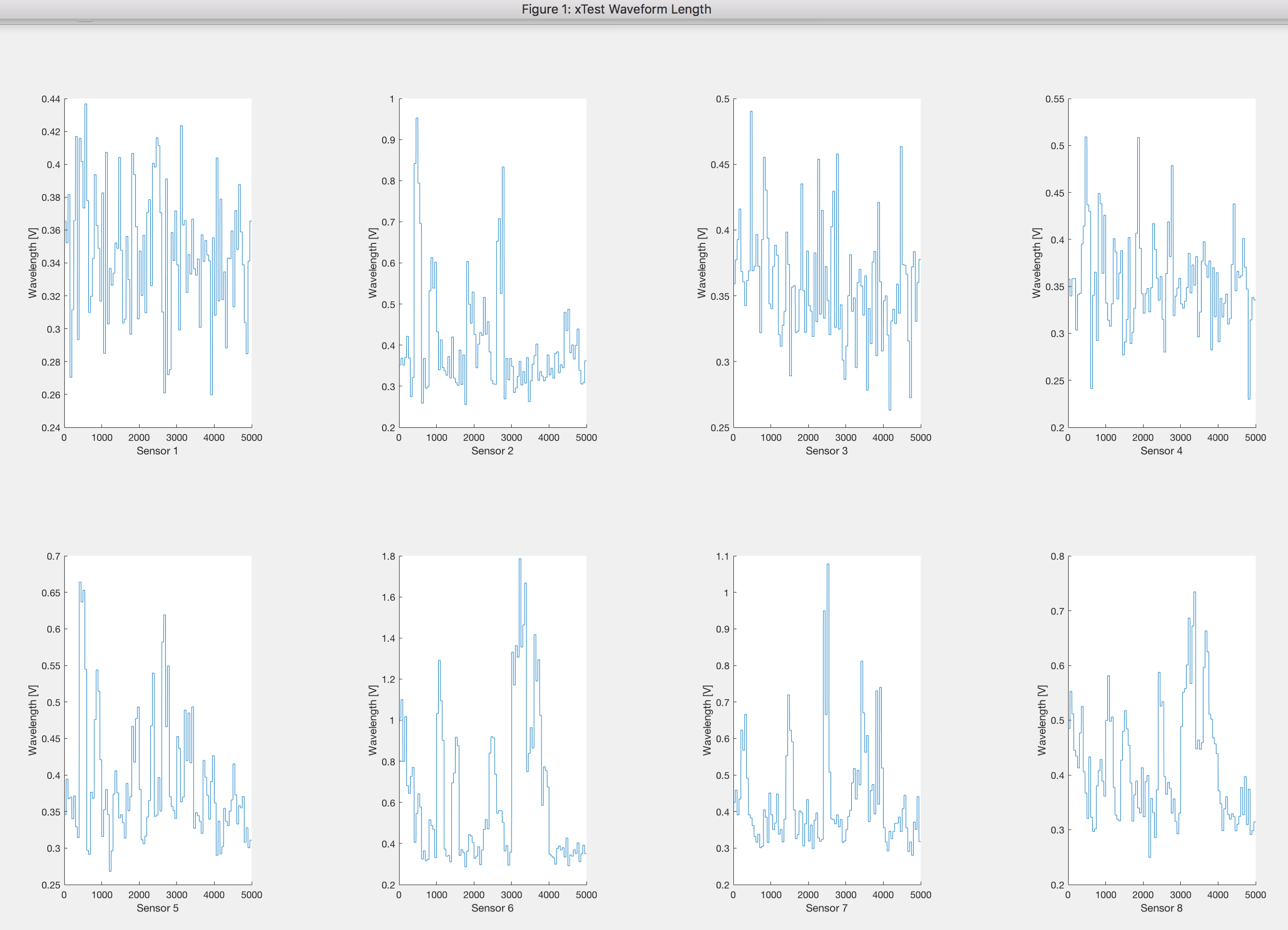


Figure 15 Waveform, window size 50

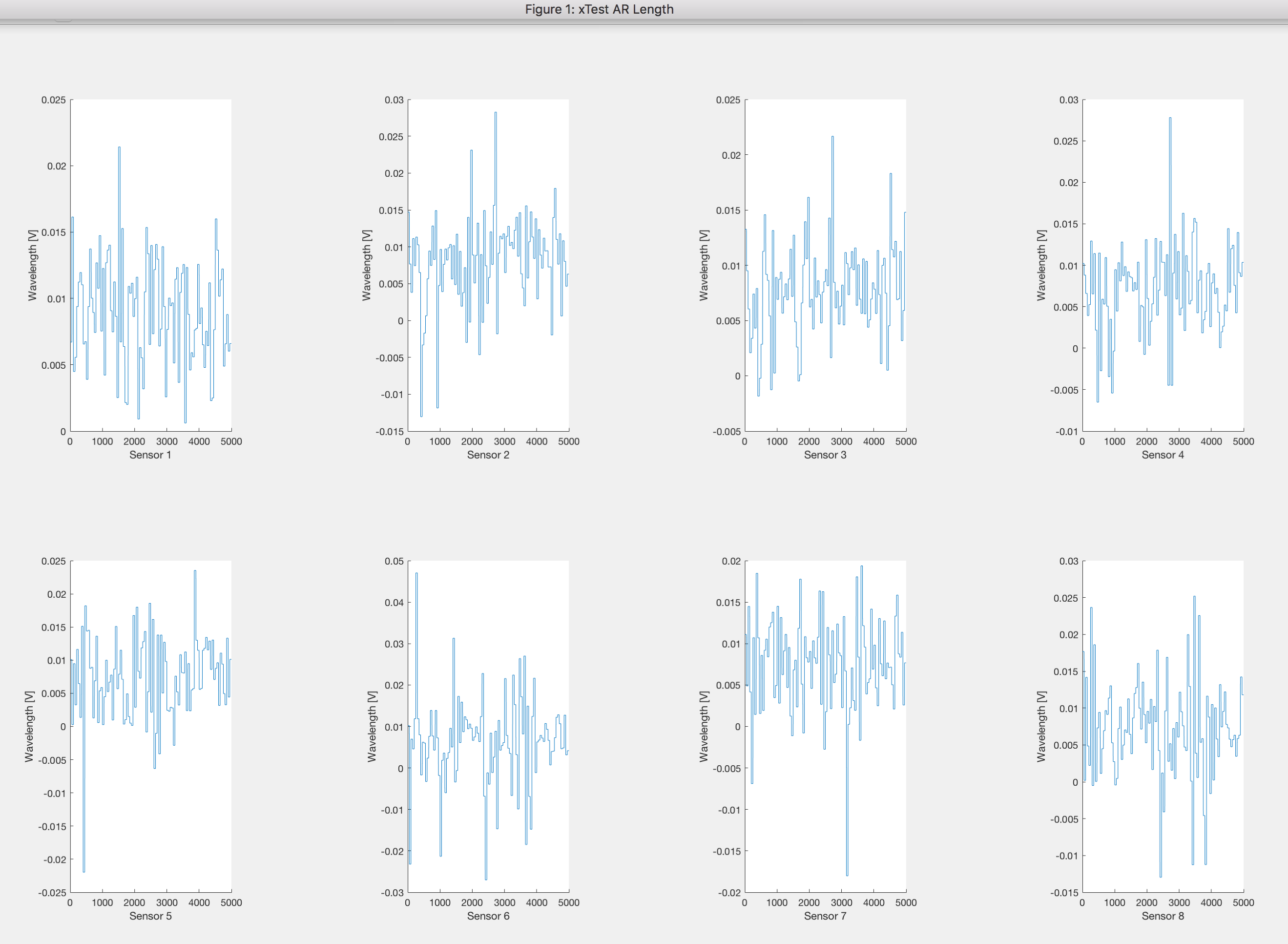


Figure 16 AR, window size 50



Figure 17 Classified test data vs Labeled, window size 50

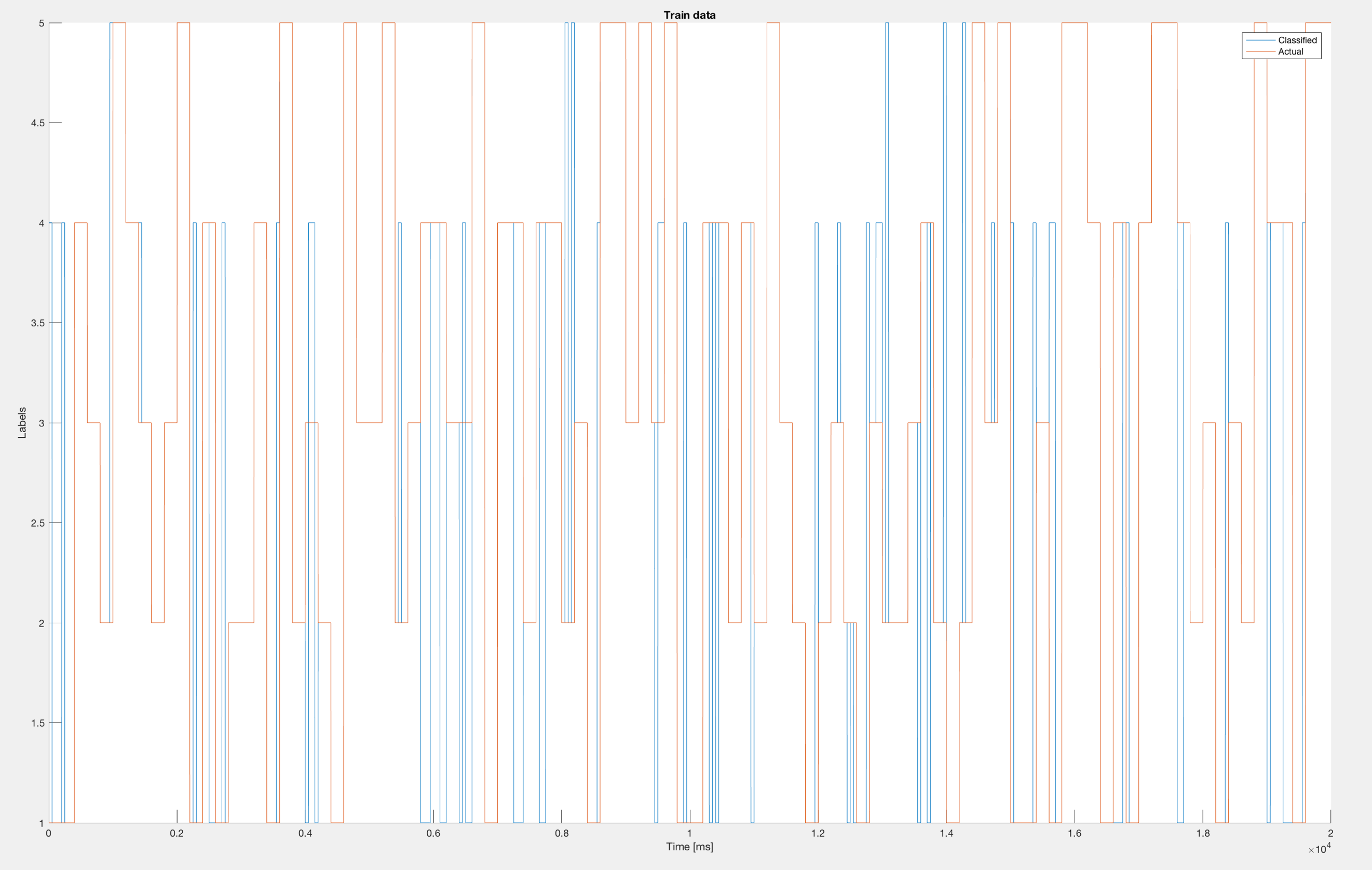


Figure 18 Classified train data vs Labeled, window 50

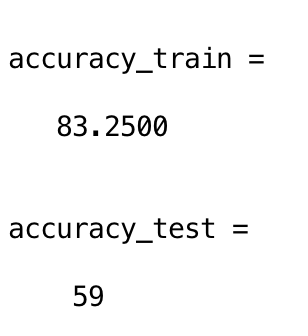


Figure 19 Accuracy of training and testing dataset, window size 50

### Window size = 200

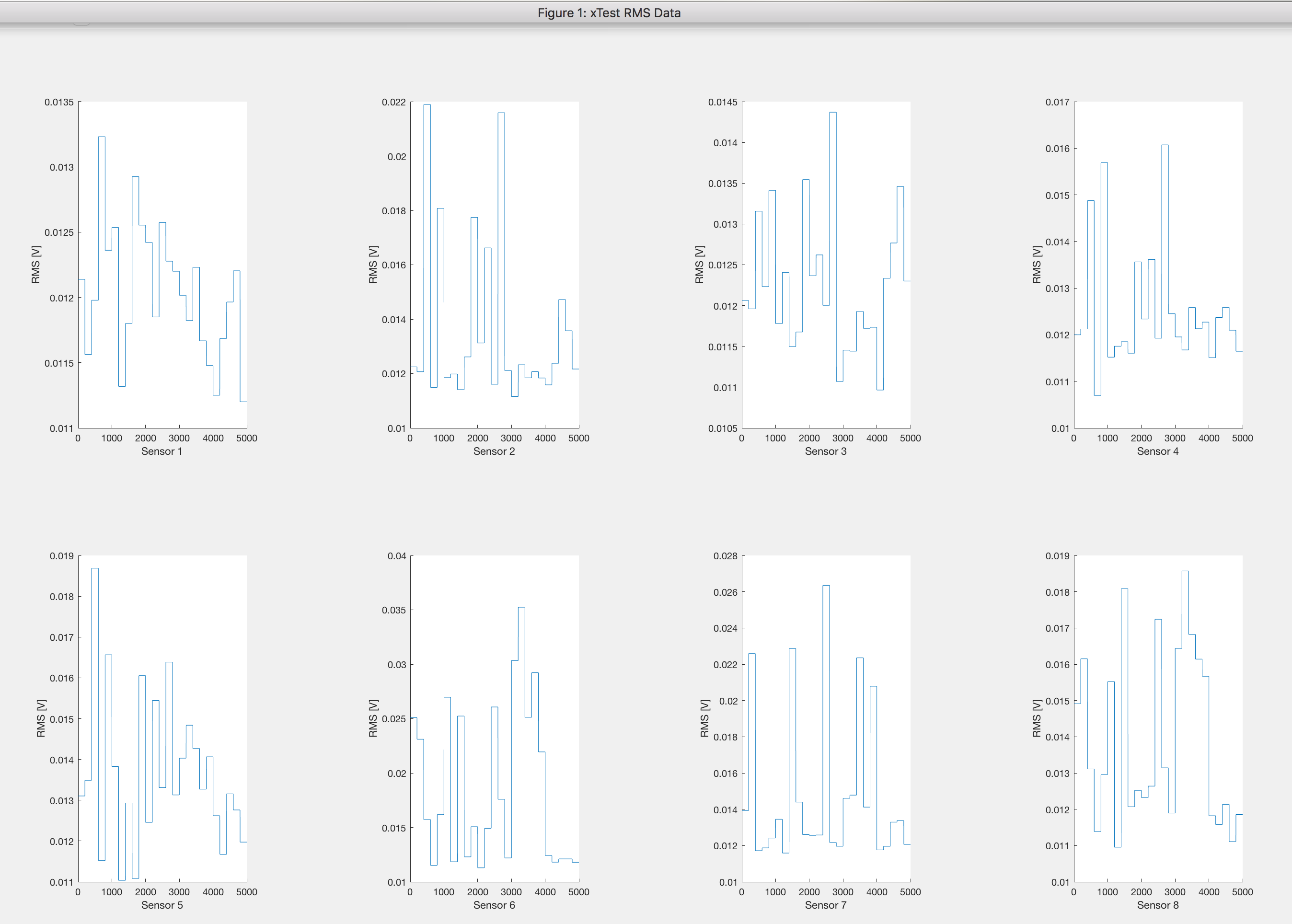


Figure 20 RMS, window size 200



Figure 21 Waveform, window size 200

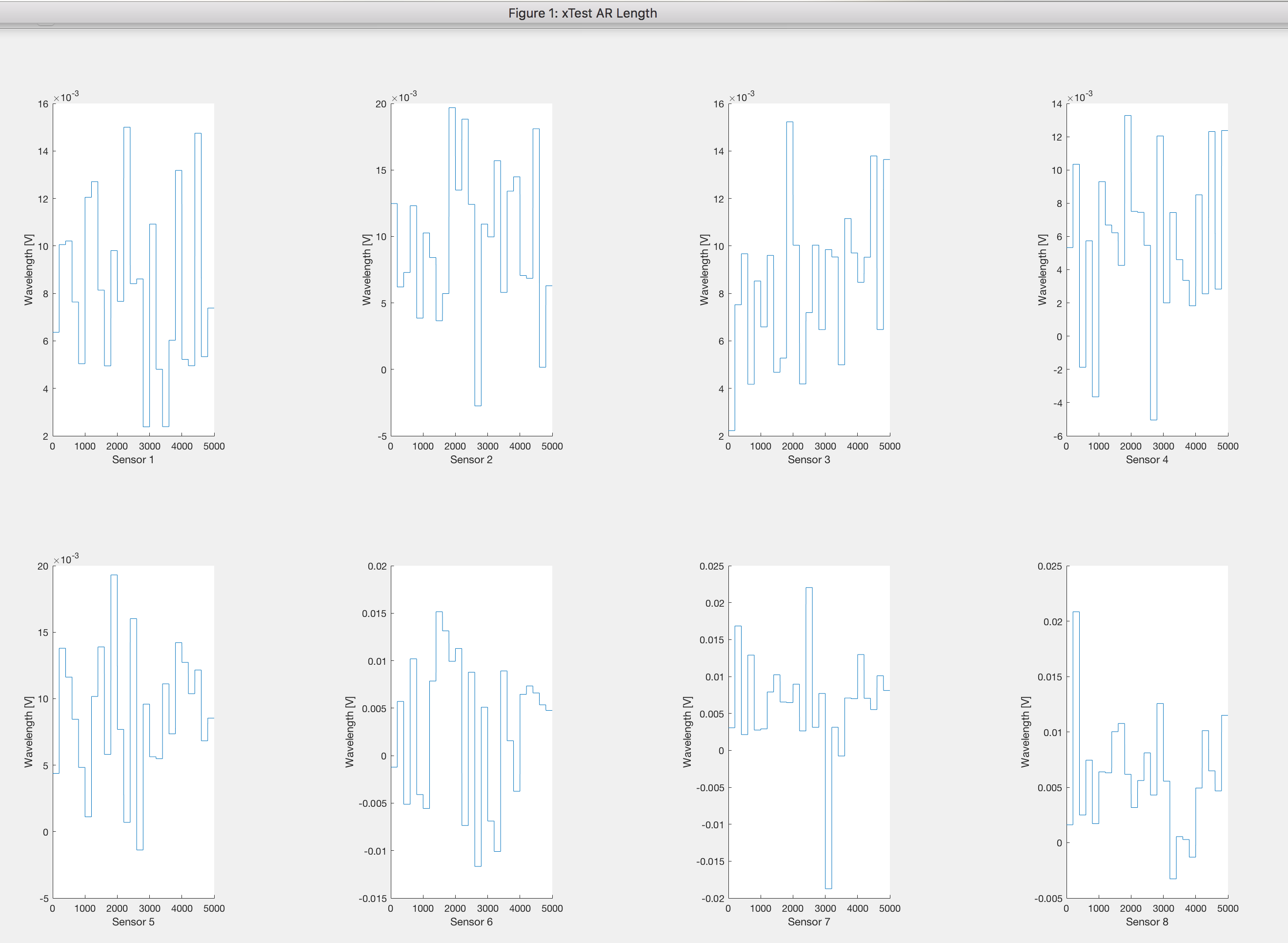


Figure 22 AR, window size 200

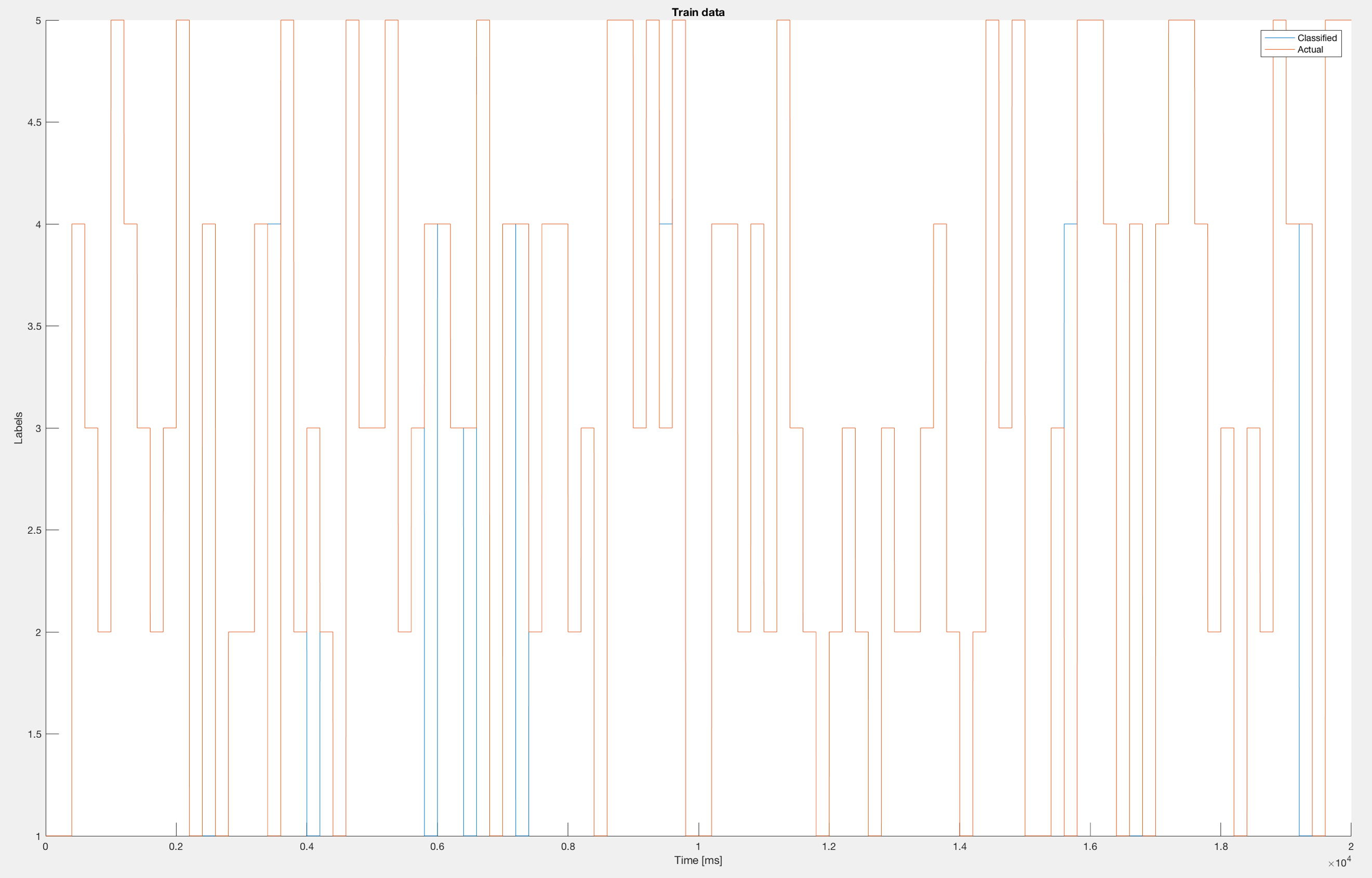


Figure 23 Classified train data vs Labeled data, window size 200

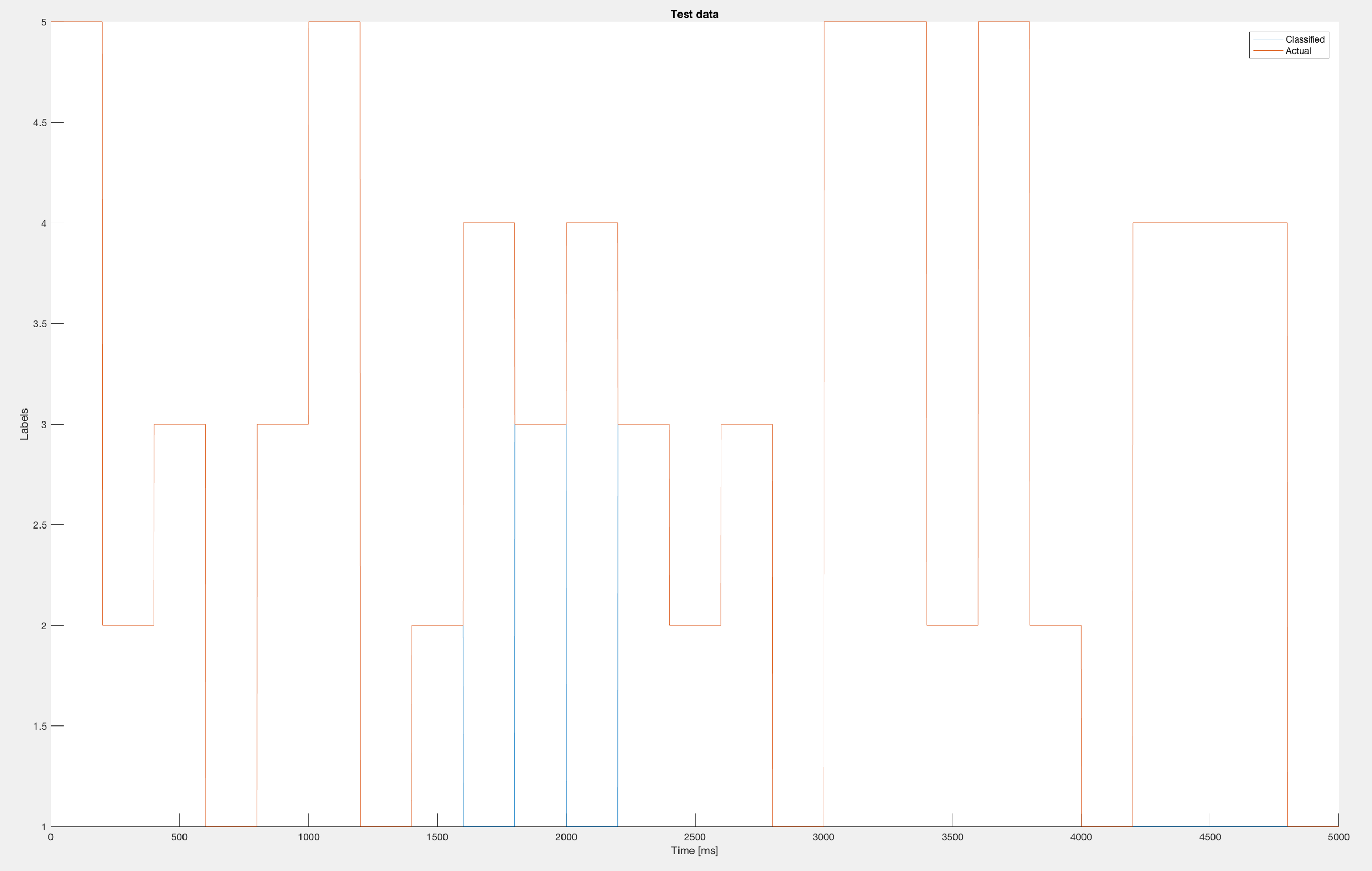


Figure 24 Classified test data vs Labeled data, window size 200

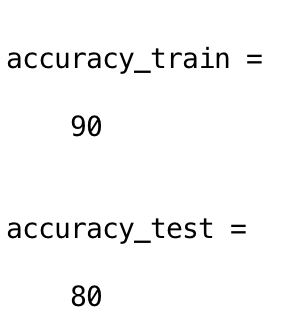


Figure 25 Accuracy of training and testing dataset, window set 200



Figure 26 Accuracy after discarding half of training dataset, window size 200