



---

## Condition Based Monitoring Statistical Functions

*Ver 0.0.1*

---

Created:	1/7/2019
Author:	Sharkey, Tom
Last Modified:	7/1/2021
Modified by:	Sharkey, Tom

---

## Modes:

### Terms used:

**Data frame** – Used to describe a set of 512 elements of “raw” time domain data. This is the amount of data seen on the time series plot and so a data frame can be considered a basic unit, to which each statistical formula is applied.

**Data point** – One element of “raw” time domain data, sent from a mote.

**RPM** – Revolutions per minute. Describes the rotational speed of the motor or vibration source being measured. Converted to Hertz by dividing by 60.

**PDF** – Probability Density Function. A function of a continuous random variable whose value at any given point can be interpreted as providing a relative likelihood that the value would equal that sample. Useful visualisation of the spread of time series data values – also necessary to describe the effects of kurtosis and skew factor on the data. See Figure 6.

**g** – Measurement of acceleration used in applications describing impact and vibration of mechanical systems.  $1g = 9.8m/s^2$ .

### Mathematical notation:

- $x_i$  - Current data point
- $\bar{x} / \mu$  - Mean
- $\sigma$  - Standard deviation.
- **N** - Number of data points in calculation
- $||$  - Modulus or positive value.
- $\mu_3$  - Third central moment.
- $\mu_4$  - Fourth central moment

## DFT – Discrete Fourier Transform

Explanation	<p>DFT provides spectral analysis of the time domain signal. It contains the same information as the time domain plot but provides a much clearer visualisation of the motor / imbalanced load acting on the mote. Clear spikes in frequency are seen at the frequency of motor rotation (RPM/60).</p> <p>Measures the dominant frequency of the repetitive impulse period of certain faults due to the contact between rolling elements and defective spot. Therefore, the type of bearing defects such as outer race, inner race or ball defect can be detected [1].</p> <p>“Discrete” simply describes a computationally-efficient algorithm which involves breaking up the time data into equally-spaced samples</p>
Calculation used	The Fourier transform used is a fast Fourier transform performed on the embedded system i.e. the mote (C code).
Typical values	Figure 1 below shows typical values for DFT for an imbalanced motor at 2000 RPM.
Application	<p>The DFT plot is the most effective method of showing what source(s) of vibration are acting on the mote in question.</p> <p>The rotation of the motor is seen at lower frequencies, at the fundamental frequency of the motor and integer multiples of this frequency.</p> <p>Bearing faults constitute small spikes at high frequencies, which should also be visible on this graph, if any are present.</p>
Plot shown	In Figure 1, the GUI plots the time and frequency data for a motor acting on an imbalanced load at 2000 RPM. This evaluates to about 33.3Hz, which is the first large spike seen in the figure. Additionally, the imbalance creates small frequency spikes at integer multiples of this fundamental frequency.
Code	<p>See ARM math for extra info</p> <p><a href="https://www.keil.com/pack/doc/CMSIS/DSP/html/group__RealFFT.html#ga180d8b764d59cbb85d37a2d5f7cd9799">https://www.keil.com/pack/doc/CMSIS/DSP/html/group__RealFFT.html#ga180d8b764d59cbb85d37a2d5f7cd9799</a></p>

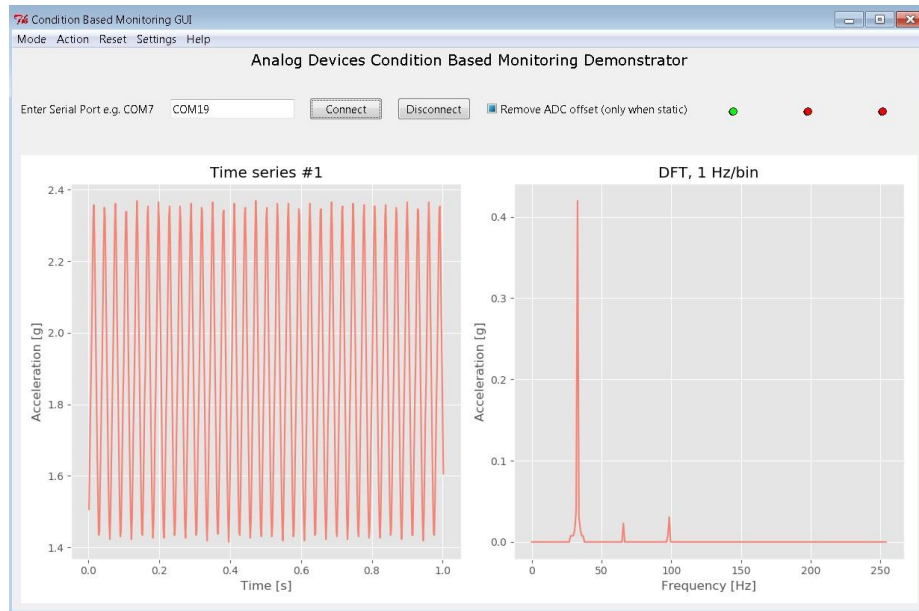


Figure 1: DFT data for 2000 RPM motor with imbalance

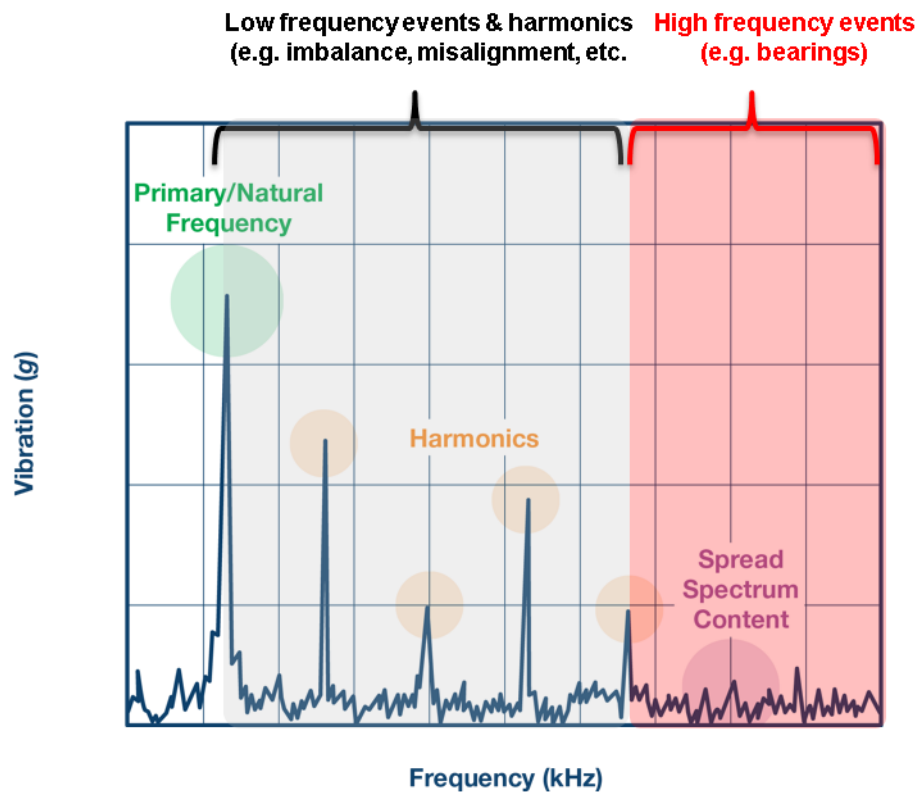


Figure 2: Explanation of DFT data

## Peak

Explanation	<p>Provides useful indication for the max value of the time domain amplitude of the signal. Not used in characterisation of motor or bearing faults, but can be used to calculate some other useful statistics for bearing fault characterisation, such as crest factor.</p> <p>Included in program to give user more flexibility with calculation of statistics, and a means of setting an alarm on the max value of acceleration that a mote should reach.</p>
Calculation used	Compares largest value in each raw data frame to the current max. If it is larger, sets a new max.
Typical values	Vibration source dependent. Test motor did not exceed 0.6 g.
Application	For user visualisation and setting of alarms.
Plot shown	Figure 3 shows peak values for almost 40 data frames, quite steady around the 0.53g value. Label on peak graph shows max value achieved over all data frames collected.
Code	<pre>def peak(dataList):     '''Takes in list and returns float furthest     deviation from 0'''     max = 0     for data in dataList:         if math.fabs(data) &gt; max:             max = data     return max</pre>

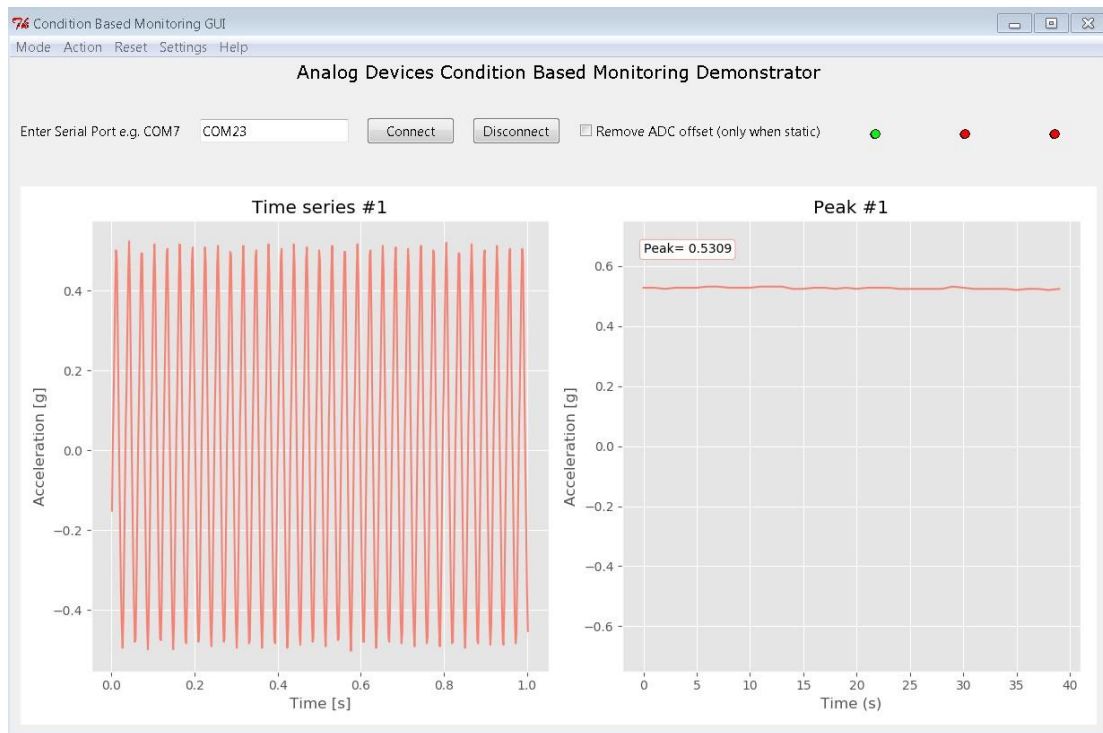


Figure 3: Peak

## P2P - Peak-to-Peak

Explanation	Similar to peak value, but calculates amplitude from max to min, rather than from zero value to max.
Calculation used	Subtracts max time series value from min time series value achieved.
Typical values	~1g when measuring an imbalanced motor at 2000 RPM – will vary based on motor used – should be approximately double peak value.
Application	Not used in further statistical analysis. Useful for setting alarms on mote values i.e. Set alarm to trigger if P2P exceeds 1g acceleration.
Plot shown	Figure 4 shows peak-to-peak values for 25 data frames. Approximately 1.020 in value.
Code	<pre>def peak_to_peak(dataList):     '''Takes in list and return float max -min value'''     max=0     min=0     for data in dataList:         if data&gt;max:             max=data         elif data&lt;min:             min=data     return max-min</pre>



Figure 4: Peak to Peak

## RMS – Root Mean Square

Explanation	Root mean square (RMS) of the time series data. Denotes the average value of the signal if all values have been made positive.
Calculation used	<p>Sums the squares of all data points in data frame, divide by length of frame, get square root of answer. Basically the mean if all values are first made positive.</p> $x_{rms} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \dots + x_n^2)}$
Typical values	For a sin wave, RMS is given by $\frac{Peak}{\sqrt{2}}$ or 0.707 of the peak value.
Application	The RMS value increases gradually as a bearing fault develops. It can be used to characterise when a bearing fault might have occurred <a href="#">[1]</a> .
Plot shown	Figure 5 shows RMS value for over 20 data frames. It is a very steady value for the waveform shown, at 0.34 acceleration.
Code	<pre>def rms(dataList):     '''Takes in list and returns float Root mean square value'''     ms=0     for data in dataList:         ms+=(data**2)     ms/=float(len(dataList))     return(float(math.sqrt(ms)))</pre>

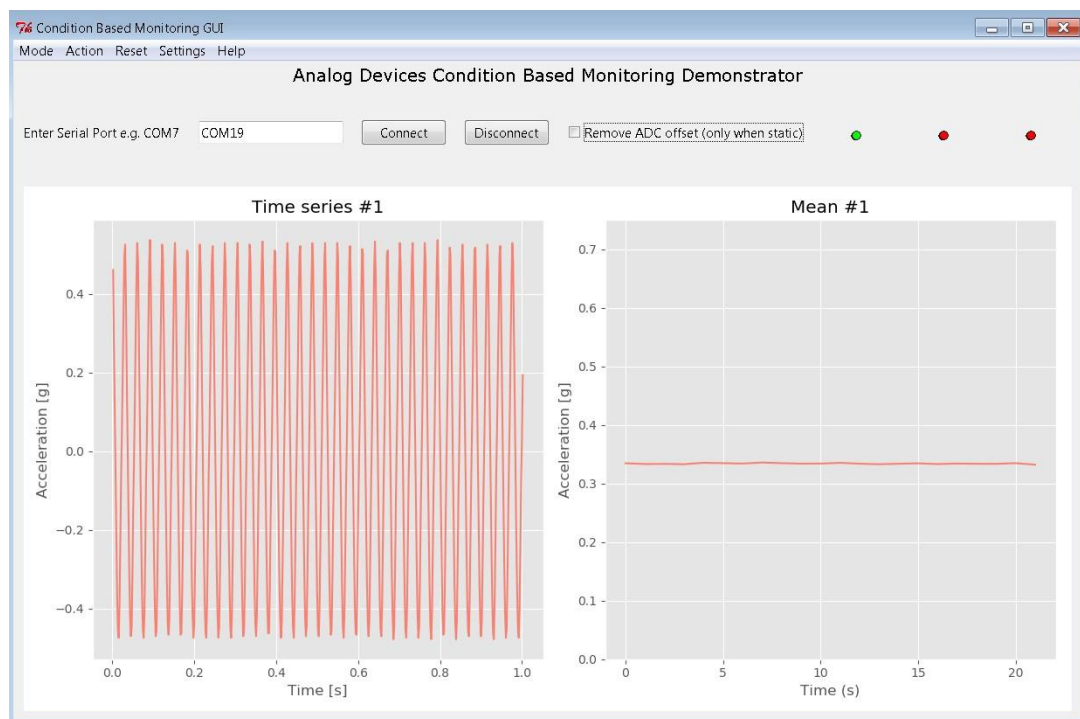


Figure 5: Root Mean Square

## Standard Deviation

Explanation	Standard Deviation is a measure of the average distance a single data point is from the mean value (0). A large standard deviation indicates that all data points are spread over a wide range, not clustered around the mean value.
Calculation used	<p>Sums the square of the difference between each data point in a frame and the mean. Calculates the square root of this sum averaged.</p> $s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$ <p>Where <math>x_i</math> is the current data point, <math>\bar{x}</math> is the mean, N is the number of data points summed.</p>
Typical values	From testing – standard deviation values for our specific motor were around 0.335. This implies that most of the data lies between $\pm 0.67$ of the mean (0). See “ <a href="#">Chebyshev’s Inequality</a> ” for greater detail on this.
Application	Standard deviation is used in conjunction with kurtosis to assess bearing faults. Kurtosis is known to increase as a bearing fault develops, but this only tells half of the story. Standard deviation describes how far points spread out from the mean, and kurtosis indicates whether outliers are infrequent and extreme (high kurtosis), or frequent and moderately sized (low kurtosis).
Plot shown	<p>Figure 6 describes the relationship between standard deviation in the normal distribution and the percentage of data contained within each section. This gives an approximation of the percentage of data contained within 1, 2 and 3 standard deviations.</p> <p>Figure 7 is a typical example of standard deviations taken from an unbalanced motor rotating at 2000 RPM. See “Typical Values” for relevance.</p>
Code	<pre>def standard_deviation(dataList):     '''Takes in a list and returns float standard deviation'''     n = len(dataList)     xBar = mean(dataList)     total=0     for x in dataList:         total+= (x-xBar)**2     return math.sqrt((total/float(n-1)))</pre>



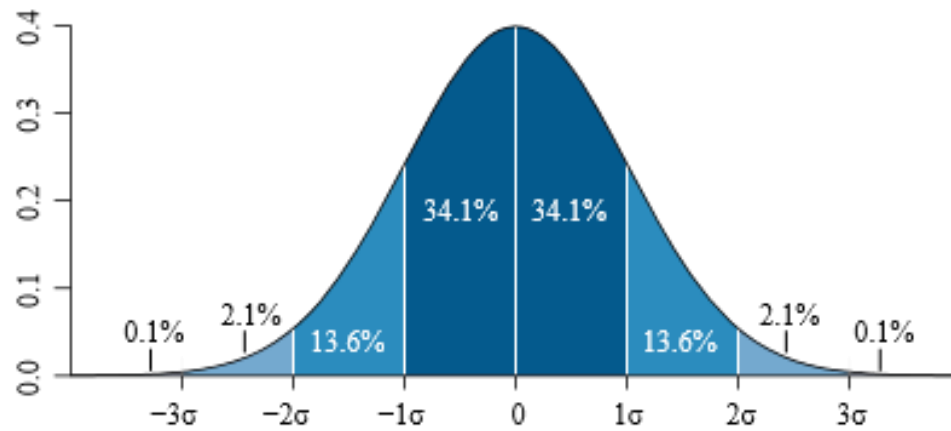


Figure 6: Standard Deviation for Normal Distribution [2]

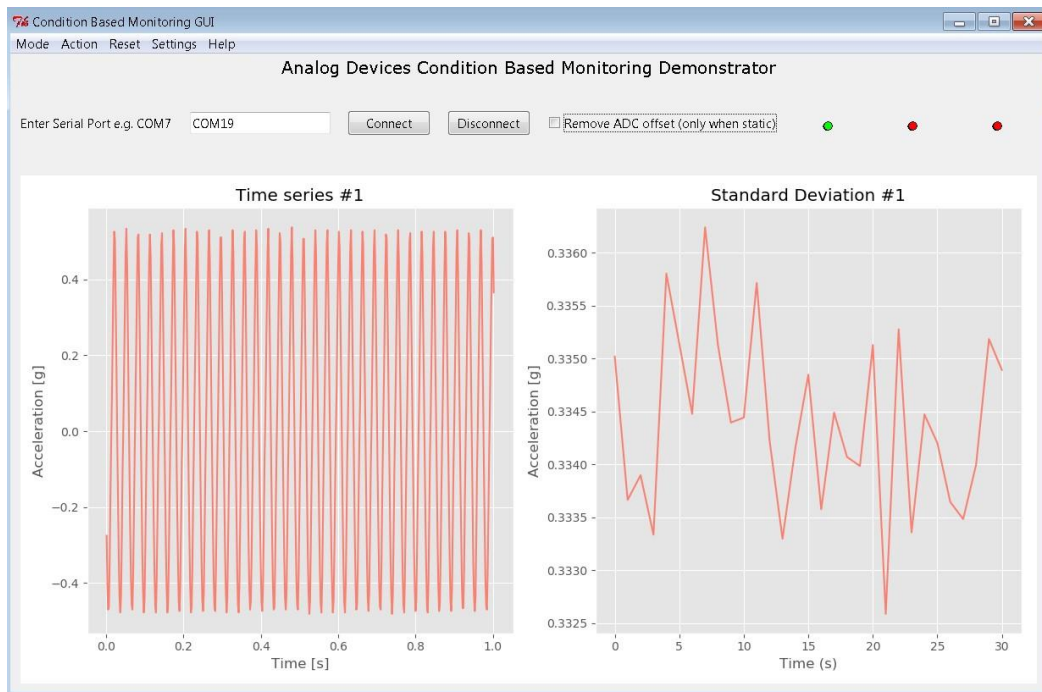


Figure 7: Standard Deviation

Note: The standard deviation values shown in Fig.7 correspond to the width of the  $\sigma$  intervals in Fig.6. A larger value of standard deviation would indicate a data distribution with values which are more spread out from the mean.

## Kurtosis

Explanation	<p>The “tailedness” of the data, kurtosis describes the “mass” of data contained in the tails of the probability density function. Kurtosis indicates whether outliers are infrequent and extreme (high kurtosis), or frequent and moderately sized (low kurtosis).</p> <p>The kurtosis displayed is the “excess kurtosis” which is the actual kurtosis value minus 3. This causes the kurtosis value of the normal distribution to be 0.0. A distribution with larger and more infrequent outliers will have a higher kurtosis. Such a distribution is “leptokurtic” – see Figure 8.</p>
Calculation used	<p>Scipy.kurtosis given here:  <a href="https://github.com/scipy/scipy/blob/v1.3.0/scipy/stats/stats.py#L1137-L1214">https://github.com/scipy/scipy/blob/v1.3.0/scipy/stats/stats.py#L1137-L1214</a>            Uses the second and fourth moments to calculate kurtosis</p> $Kurt[X] = E \left[ \left( \frac{X - \mu}{\sigma} \right)^4 \right] = \frac{\mu_4}{\sigma^4}$ <p>Where <math>\mu_4</math> is the fourth central moment and <math>\sigma</math> is the standard deviation.</p>
Typical values	Kurtosis’ value for normal rolling element bearing (using the Fisher definition) is 0.
Application	Deviations from a zero value for kurtosis can be used to indicate a bearing fault.
Plot shown	<p>Figure 8 shows the difference between positive and negative kurtosis, as compared to a normal distribution which has 0 excess kurtosis</p> <p>Figure 9 shows 30 data frames of typical values for excess kurtosis.</p>
Code	Function taken from scipy.stats, see “Calculation used”.

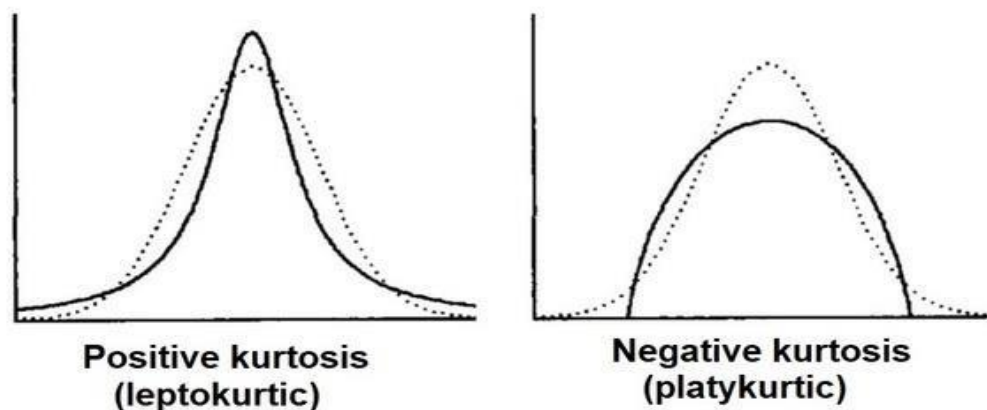


Figure 8: Kurtosis Explanation [3]

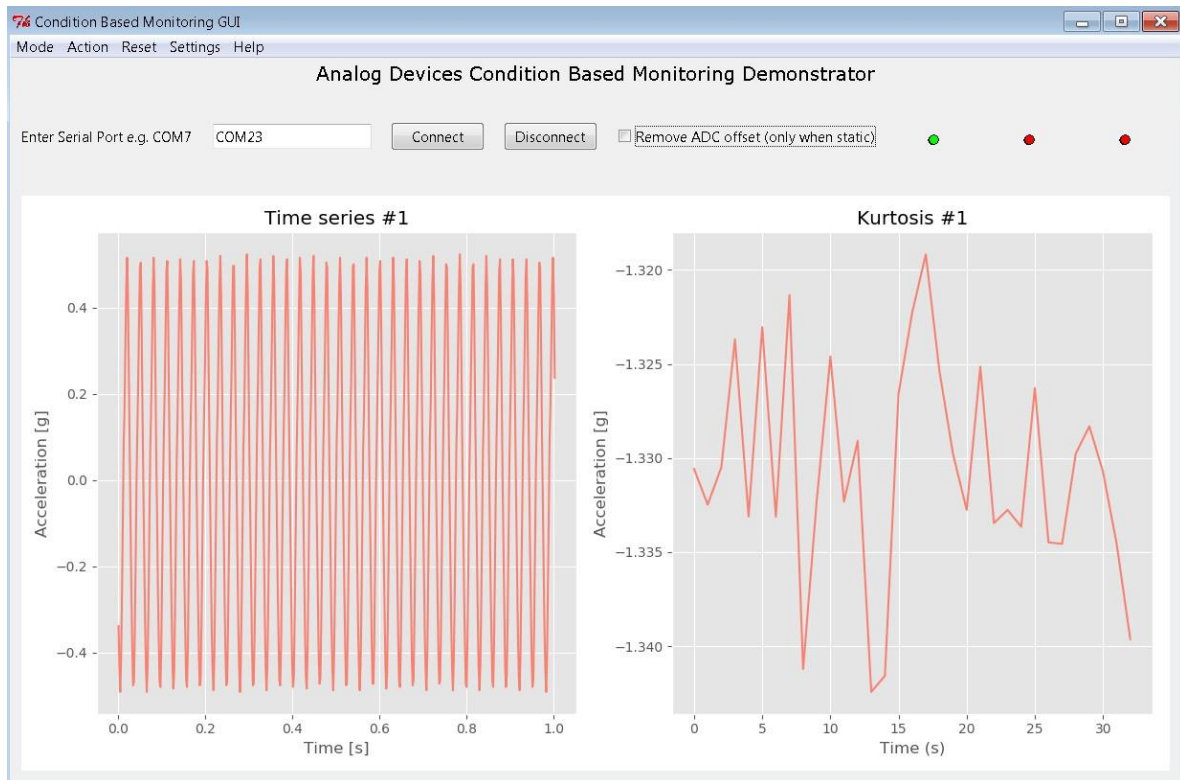


Figure 9: Kurtosis

## Skew Factor

Explanation	Quantifies asymmetry behaviour of vibration signal through its probability density function. A function which is negatively skewed will have a long tail to the left, the data is more concentrated to the right. A function with a positive skew has a long tail to the right.
Calculation used	<p>The method employed is the Fisher Pearson coefficient of skewness.</p> $\gamma_1 = E \left[ \left( \frac{X - \mu}{\sigma} \right)^3 \right] = \frac{\mu_3}{\sigma^3}$ <p>Where <math>\mu</math> is the mean, <math>\sigma</math> is the standard deviation, <math>\mu_3</math> is the third central moment.</p>
Typical values	<p>Normal distribution – 0</p> <p>Exponential distribution – 2</p>
Application	It is known the skewness value for a normal bearing is approximately zero. This will shift either negative or positive as a bearing fault develops.
Plot shown	Figure 10 shows the difference between positive and negative skewness as compared to the normal distribution.
Code	<p>Function used is from scipy stats:</p> <p><a href="https://github.com/scipy/scipy/blob/v1.3.0/scipy/stats/stats.py#L1137-L1214">https://github.com/scipy/scipy/blob/v1.3.0/scipy/stats/stats.py#L1137-L1214</a></p>

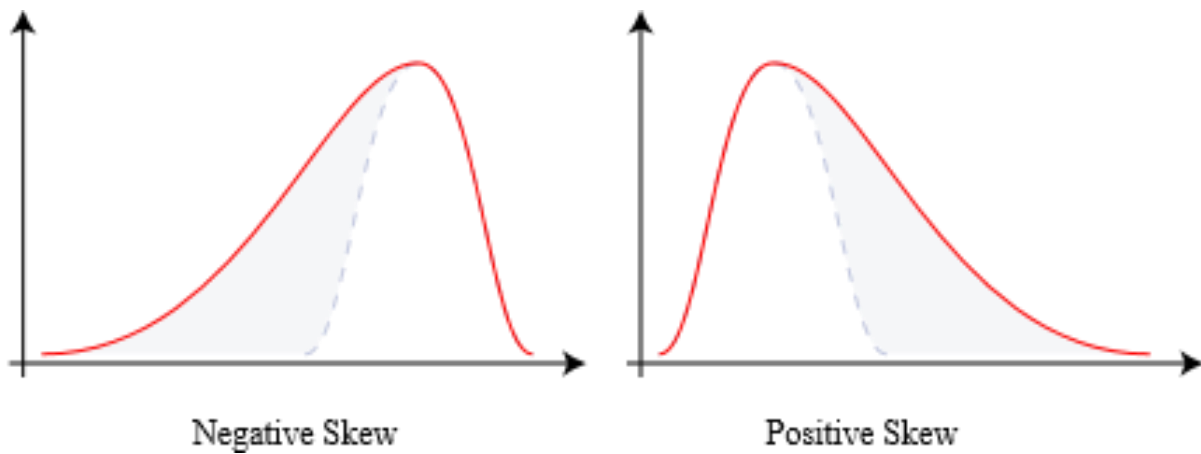


Figure 10: Skew Explanation [4]

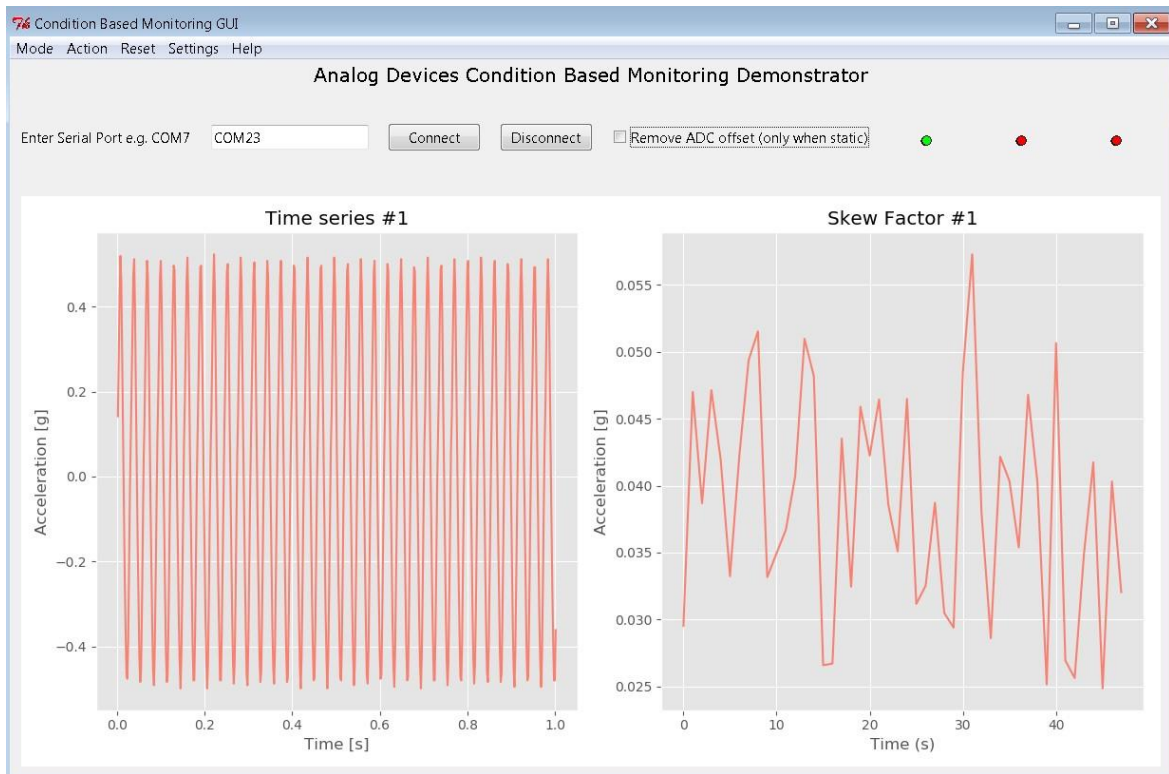


Figure 11: Skew Factor

## Crest Factor

Explanation	Indicates how extreme the peaks are in a waveform. Describes the ratio of the peak of a waveform to the effective value (rms).
Calculation used	$C = \frac{ x_{peak} }{x_{rms}}$
Typical values	<p>Some default values of crest factor are</p> <p>Sin wave – 1.414</p> <p>Triangle wave - 1.732</p> <p>Gaussian noise - <math>\infty</math></p> <p>The values for an unbalanced motor at 2000 RPM were ~1.63.</p>
Reason used	The greater the crest factor, the greater the energy contained in higher frequency harmonics. As higher frequency harmonics are often used to mark bearing issues, the crest factor can be applied to find these faults.
Plot shown	Figure 12 shows typical values for crest factor for an unbalanced motor at 2000 RPM.
Code	<pre>def crest(dataList):     '''Takes in list and returns float crest factor'''     absPeak = float(peak(dataList))     xRms = float(rms(dataList))     return absPeak/xRms</pre>

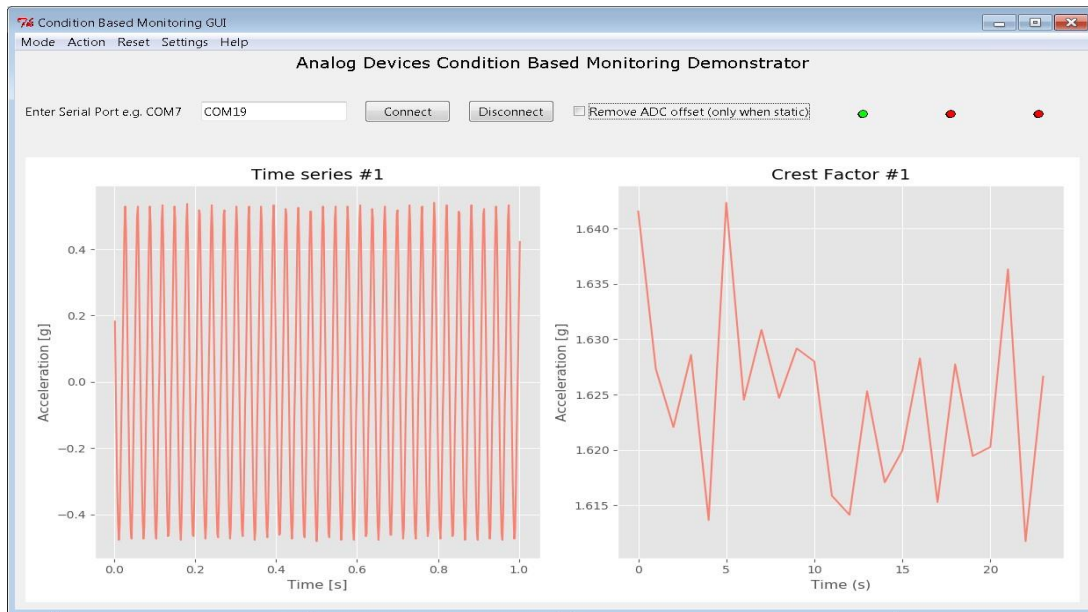


Figure 12: Crest Factor

## References

- [1] [https://www.researchgate.net/publication/320074126\\_A\\_Review\\_of\\_Feature\\_Extraction\\_Methods\\_in\\_Vibration-Based\\_Condition\\_Monitoring\\_and\\_Its\\_Application\\_for\\_Degradation\\_Trend\\_Estimation\\_of\\_LowSpeed Slow Bearing](https://www.researchgate.net/publication/320074126_A_Review_of_Feature_Extraction_Methods_in_Vibration-Based_Condition_Monitoring_and_Its_Application_for_Degradation_Trend_Estimation_of_LowSpeed_Slow_Bearing)
- [2] [https://commons.wikimedia.org/wiki/File:Standard\\_deviation\\_diagram.svg](https://commons.wikimedia.org/wiki/File:Standard_deviation_diagram.svg)
- [3] <http://www.columbia.edu/~ld208/psymeth97.pdf>
- [4] [https://commons.wikimedia.org/wiki/File:Negative\\_and\\_positive\\_skew\\_diagrams\\_\(English\).svg](https://commons.wikimedia.org/wiki/File:Negative_and_positive_skew_diagrams_(English).svg)