EE5111: Industrial Control & Instrumentation
Fault Diagnosis and Control (II)

Sunan Huang

Temasek Laboratories

National University of Singapore

— Aug. 2021—

Learning Objective

- Understand the fundamentals of fault detection scheme
- You can do a simple fault detection

Topics To Be Covered

- **Hardware diagnosis method**
- **Soft diagnosis method**

NUS National University of Singapore

Review 1: fault modelling

$$f(t) = \alpha(t - T_0)F(t)$$

$$\alpha(t - T_0) = \begin{cases} 0, & t < T_0 \\ 1, & t \geq T_0 \end{cases}$$

$$\alpha(t - T_0) = \begin{cases} 0, & t < T_0 \\ 1 - e^{-\rho(t - T_0)}, & t \geq T_0 \end{cases}$$

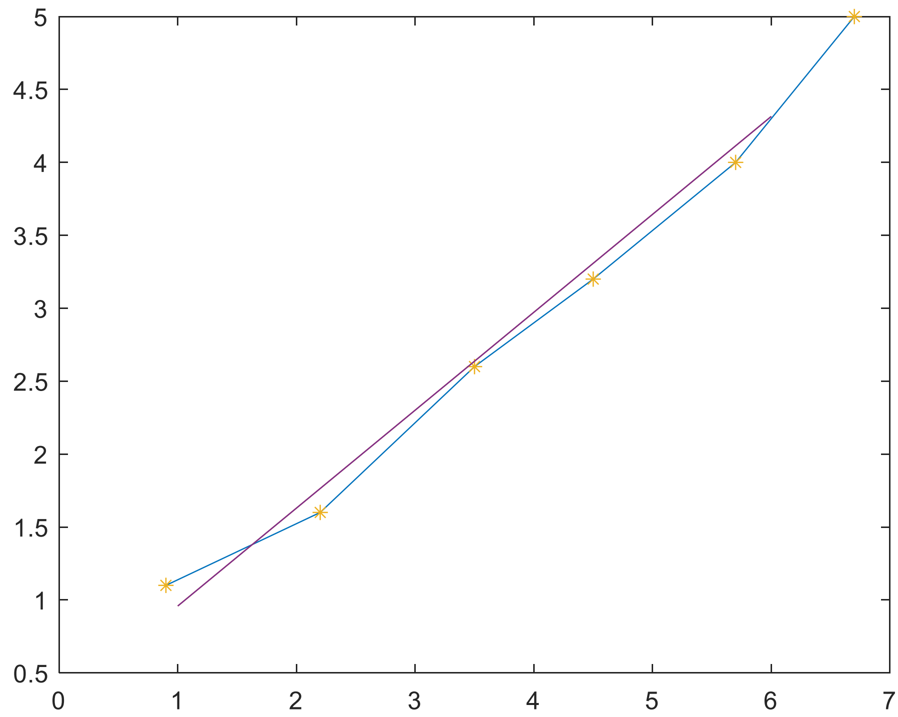
Review 2:fault function modelling

We have fault data collected.

Why we need to use a black-box model to approximate the fault behavior?

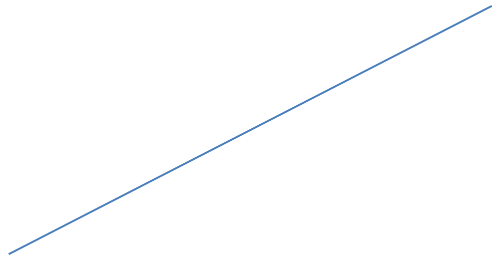
input	0.9	2.2	3.5	4.5	5.7	6.7
output	1.1	1.6	2.6	3.2	4.0	5.0

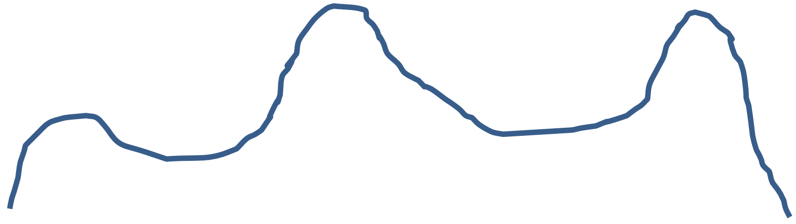
```
y1=0.6713*x1+0.2873;
```



We have two kinds of models to approximate fault data.

1) Linear model to find the linear relationship between input and output data
2) Nonlinear model to find the nonlinear relationship between input and output data





Review 3:fault function modelling

Linear model: y= $\boldsymbol{b_0} + \boldsymbol{b_1 x_1} + \cdots + \boldsymbol{b_n x_n}$

We have a definite answer for linear model

Nonlinear model: y=$\boldsymbol{b_0} + \boldsymbol{b_1 x} + \boldsymbol{b_2 x^2} + \boldsymbol{b_3 x^3} + \cdots + \boldsymbol{b_p x^p}$

We can use this model to try, but not guarantee to get an accurate model. This is only for <u>one variable x.</u>

Nonlinear model differs from linear one. It is necessary to choose the order number *p*.

$$y=\boldsymbol{b_0}+\boldsymbol{b_1 x}+\boldsymbol{b_2 x^2}+\boldsymbol{b_3 x^3}+\cdots+\boldsymbol{b_p x^p}$$

$$X=\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^p \\ 1 & x_2 & x_2^2 & \dots & x_2^p \\ & \vdots & & & \\ 1 & x_n & x_n^2 & \dots & x_n^p \end{bmatrix} \qquad Y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Let $\hat{b}$ be the estimated vector of $b=[b_0, b_1, \dots, b_p]^T$

Estimated coefficients $\hat{b}=(X^T X)^{-1} X^T Y$

Estimated output $\hat{Y}=X\hat{b}$

How do we choose the number *p*?

$$\boldsymbol{e}=\boldsymbol{Y}-\boldsymbol{\hat{Y}},$$

How do we choose the number *p*?

$$e = Y - \hat{Y},$$

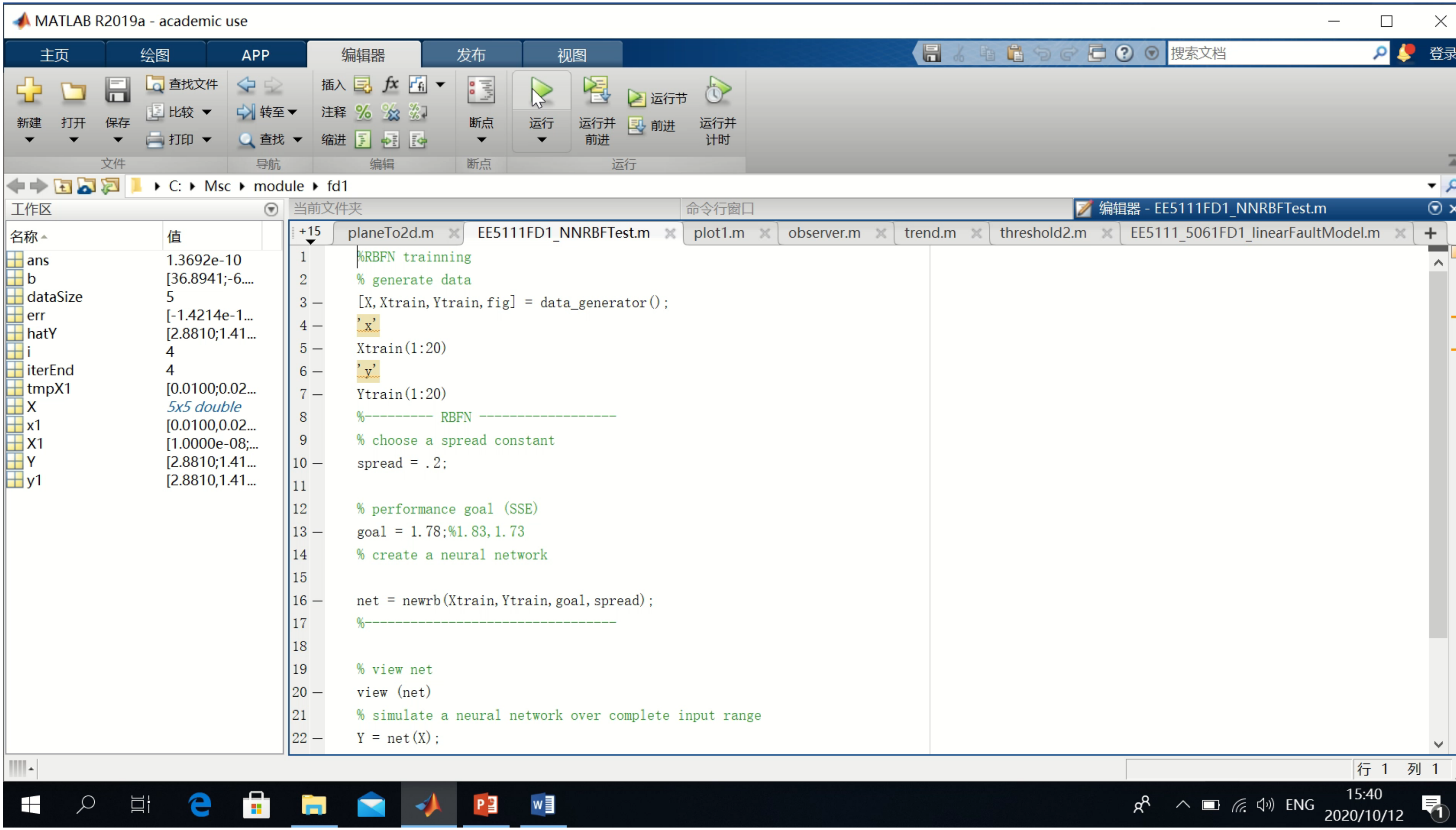
-

If |e|< tolerance number (user defined), *p* is ok; otherwise, *p*=*p*+1.

Continue to do

NUS National University of Singapore

For any fault, we can use neural network to build a model



1.Automatic fault detection

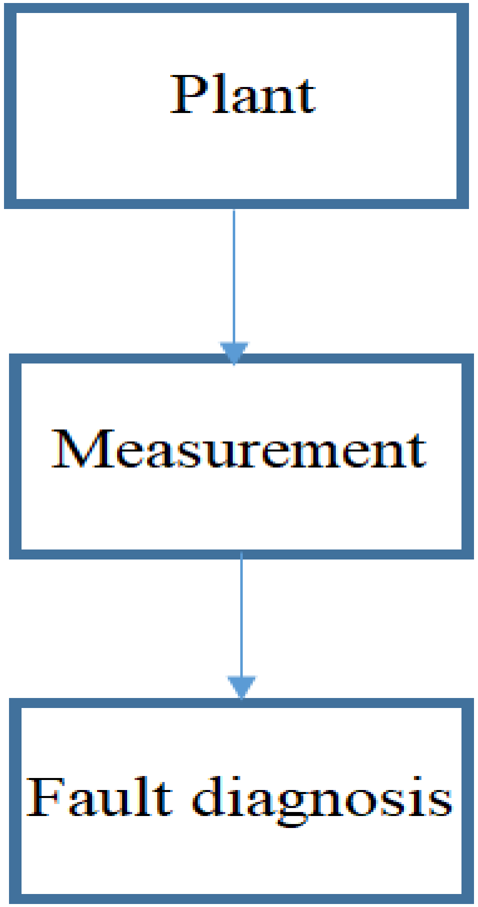
- ***Automatic fault detection (AFT)***: 1)determination of the faults present in a system and 2)the time of fault detection.

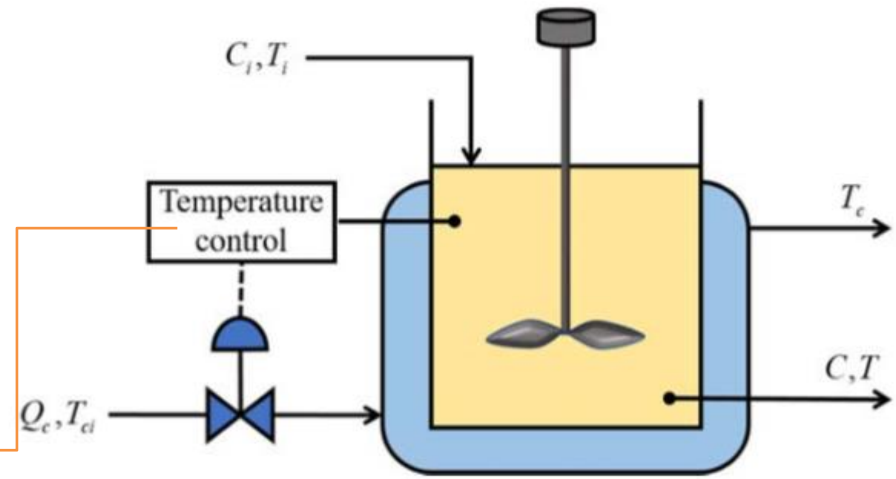
AFT should detect malfunctions in real time, as soon as possible

2 Fault diagnosis methods

- Hardware method
- Software method (Model-based method)

2.1. Hardware method---sensor is used to monitor device or system directly

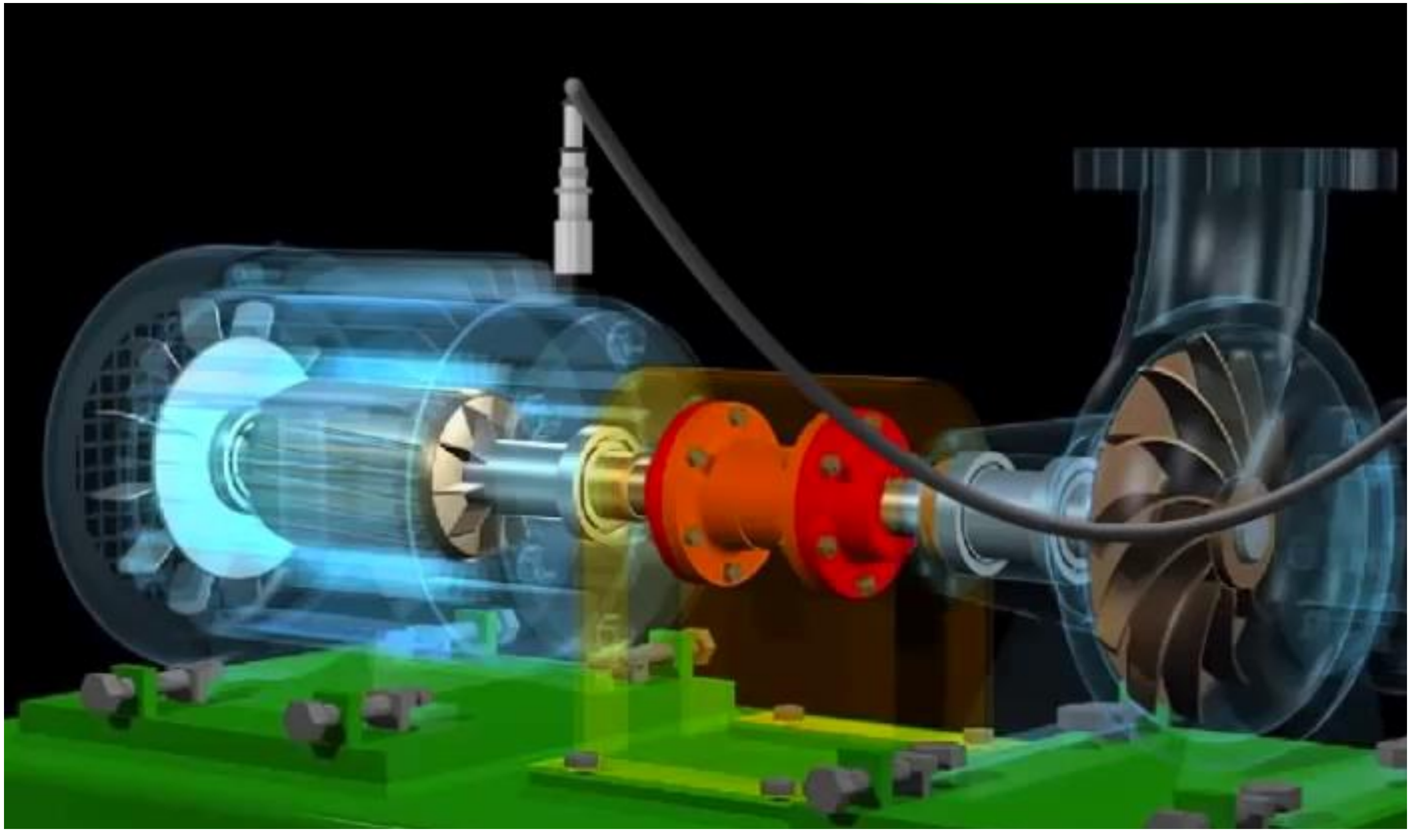




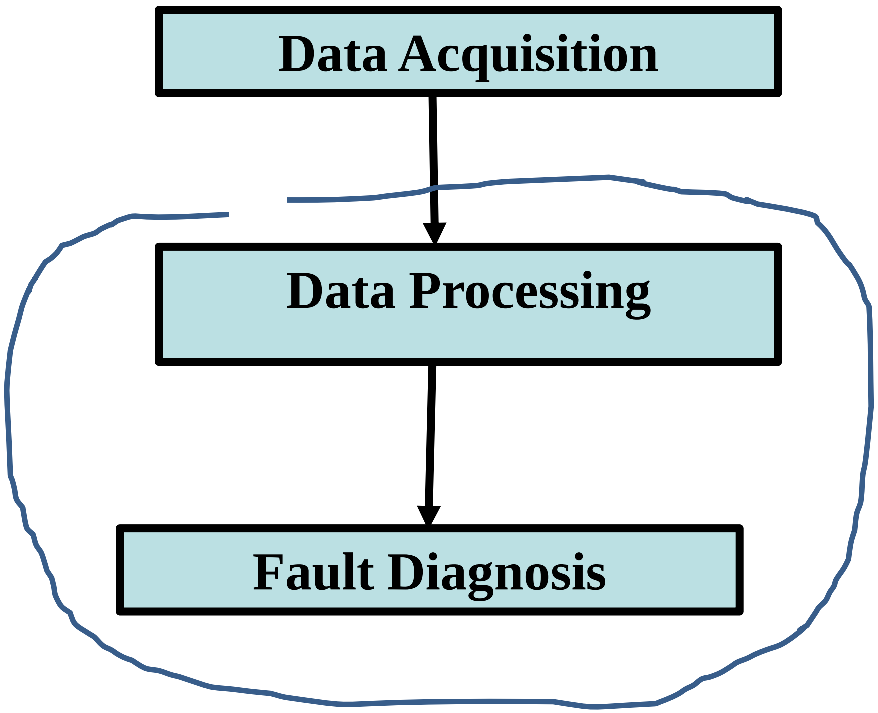
Schematic of the continuous stirred tank reactor (CSTR)

In general, hardware method is to use sensors to check somethings which reflect faults. In this case, your sensors should be placed to the measurement point closely

For example,



Three-step process of hardware fault diagnosis method



Collection and storage of raw data

Extract features from the data-2.1.1

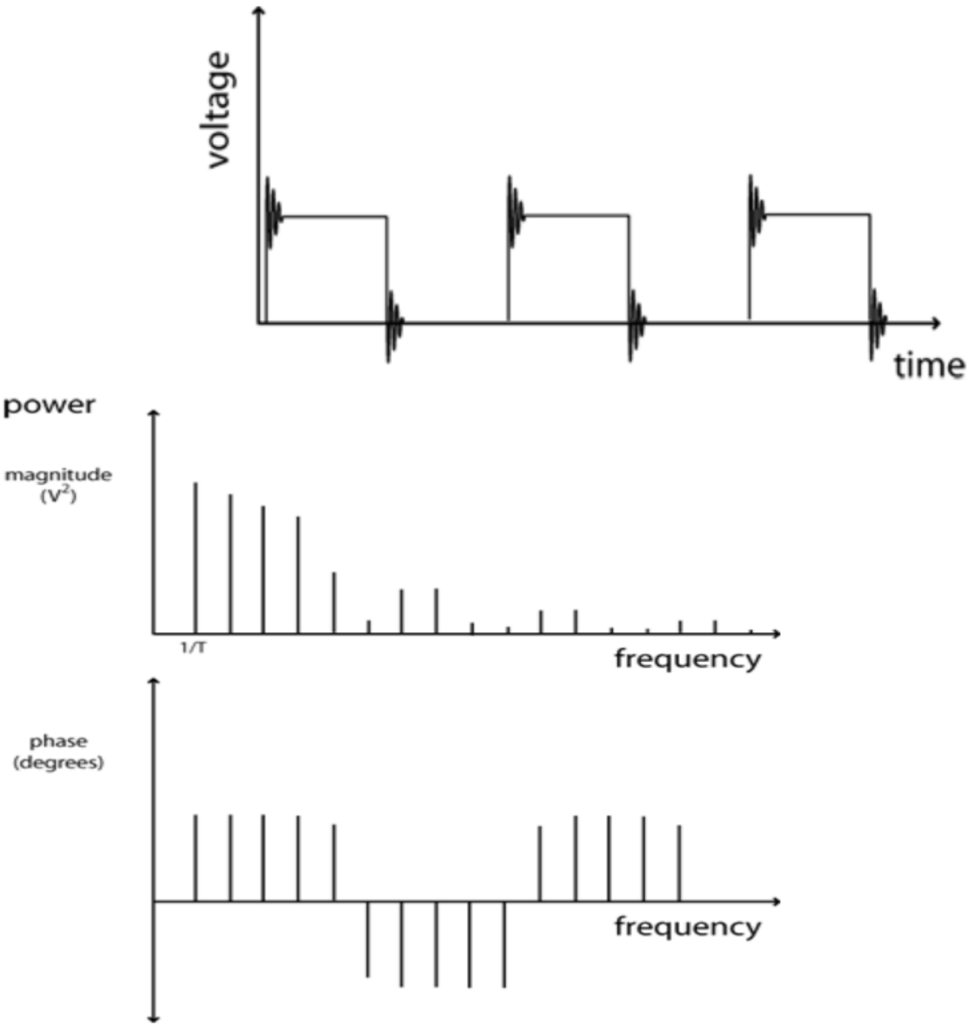
Fault decision -2.1.2

2.1.1 Data processing

It has two methods:

1)time-domain and

2) frequency-domain



2.1.1.1 Time-domain method:

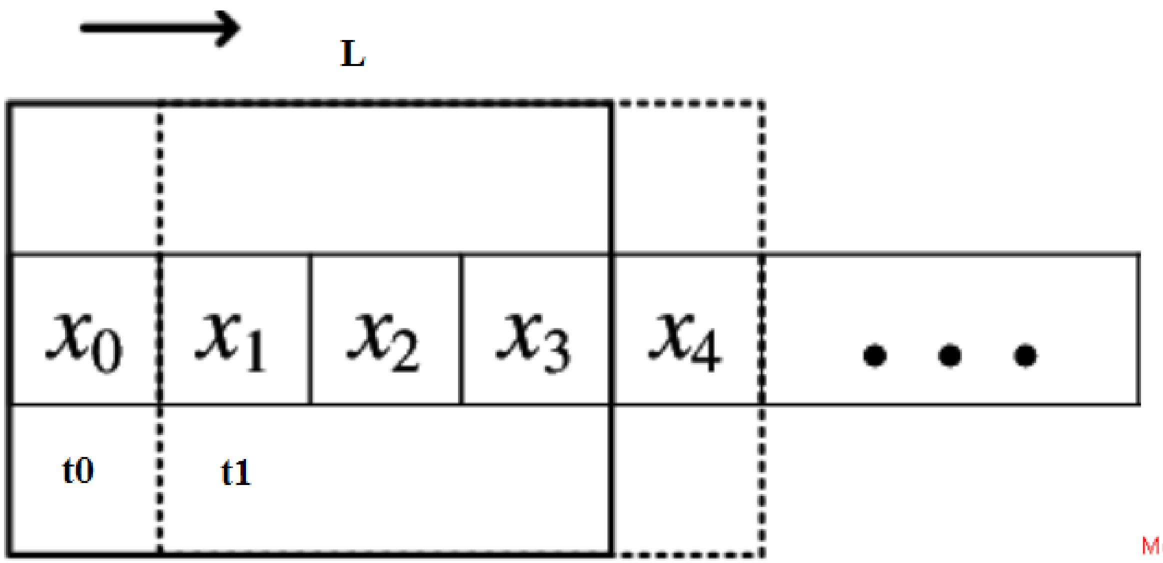
A time domain analysis is an analysis of physical signals, mathematical functions, or time series of economic or environmental data, in reference to time

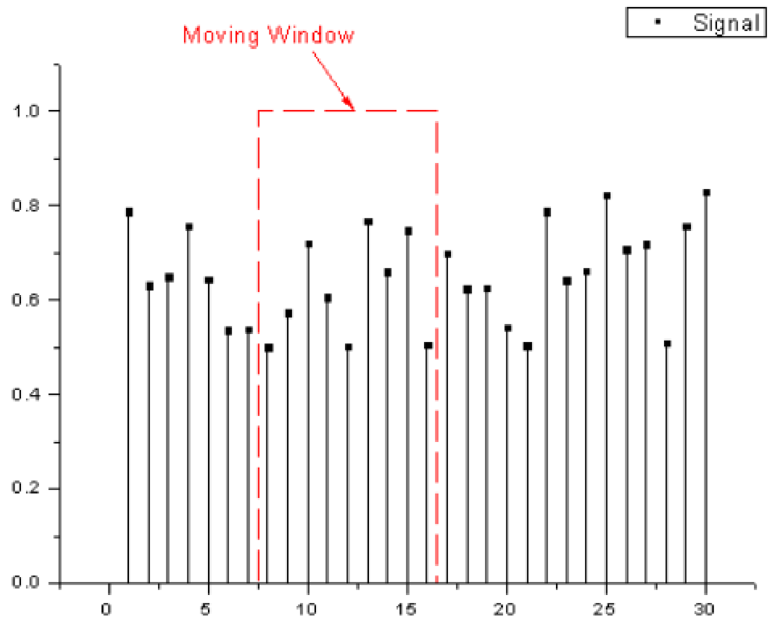
You define a time window length *L*. And then use statistical methods to do the assessment of statistical moments

- **mean,**
- **variance,** → $$\sigma^2 = \frac{\Sigma(X_i - \mu)^2}{L}$$
- **kurtosis, or**
- **parameters such as e.g. minimum and maximum value, peak-to-peak value, or root mean square (RMS).**

NUS National University of Singapore

Schematic of the sliding-window for time series





The <u>kurtosis</u> is the fourth central moment for a <u>standard normal distribution</u> .

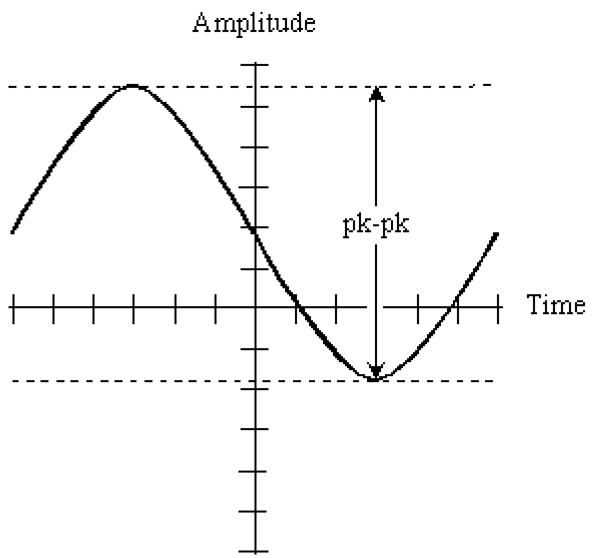
For data Y_1, Y_2, ..., Y_N, the formula for kurtosis is:

$$\text{kurtosis} = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^4/N}{s^4}$$

where $\bar{Y}$ is the mean, s is the standard deviation, and N is the number of data points.

NUS National University of Singapore

<u>Peak-to-peak</u> (pk-pk) is the difference between the maximum positive and the maximum negative amplitudes of a waveform

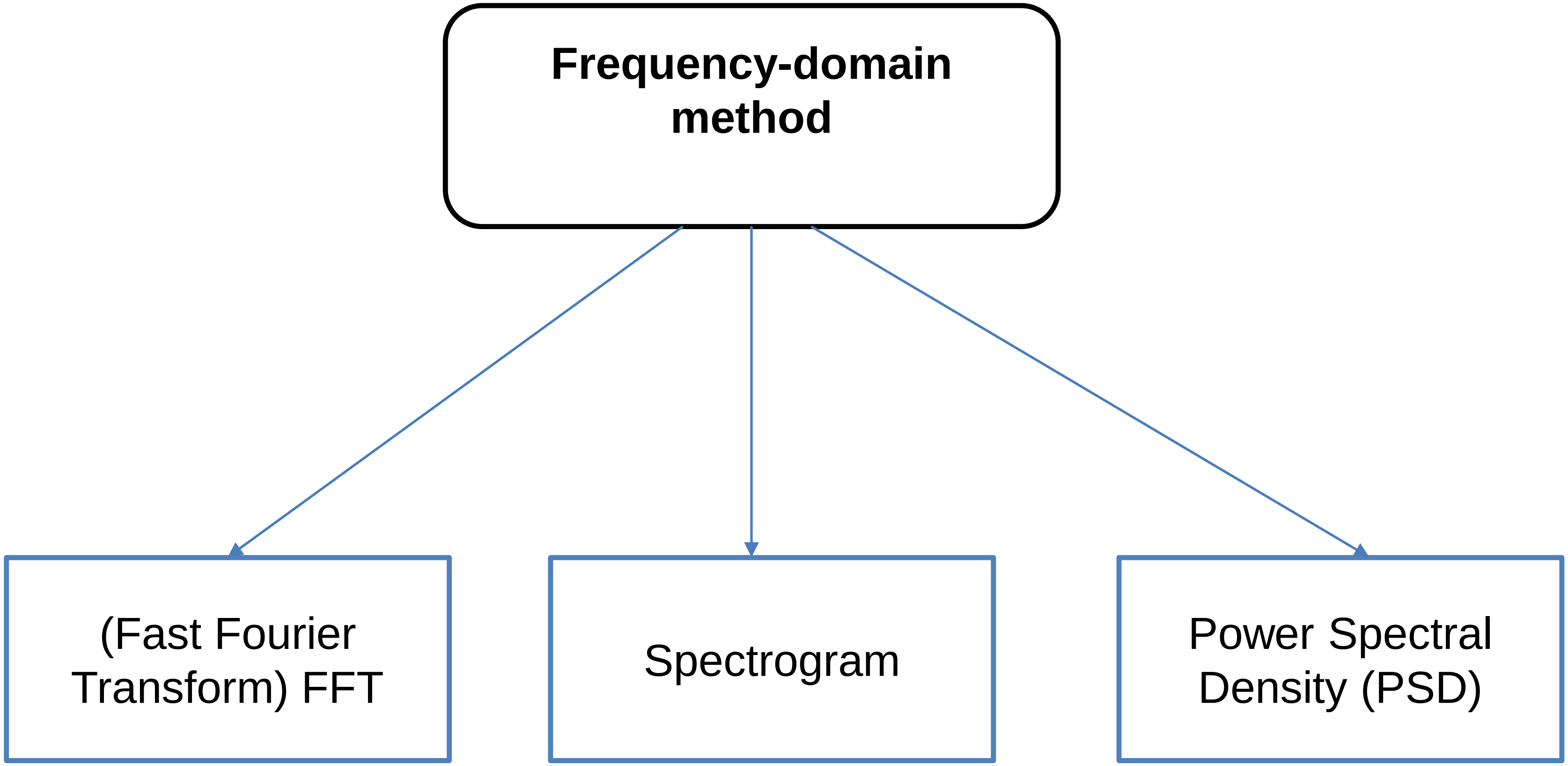


2.1.1.2. Frequency-domain method

Frequency domain is an analysis of signals or mathematical functions, in reference to frequency, instead of time.

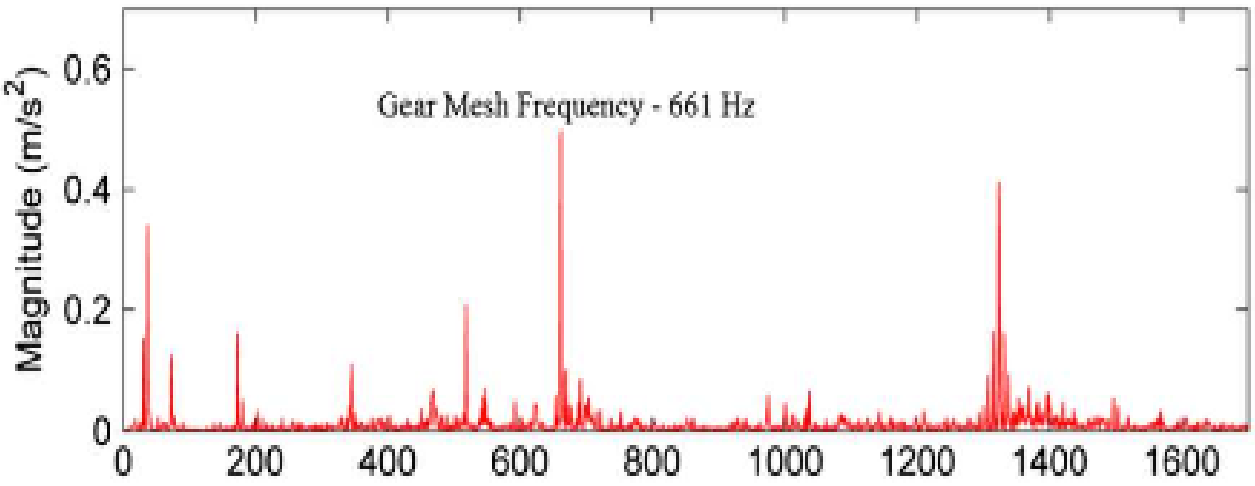
As stated earlier, the time-domain representation gives the amplitudes of the signal at the instants of time during which it was sampled. However, in many cases you need to know the frequency content of a signal rather than the amplitudes of the individual samples.

For simple sine waves, the vibration frequency could be determined from looking at the waveform in the time domain; however, as different frequency components and noise are added, spectrum analysis is necessary to obtain a clearer picture of the vibration frequency.


Frequency-domain method
(Fast Fourier Transform) FFT
Spectrogram
Power Spectral Density (PSD)

A. Fast-Fourier Transform (FFT)

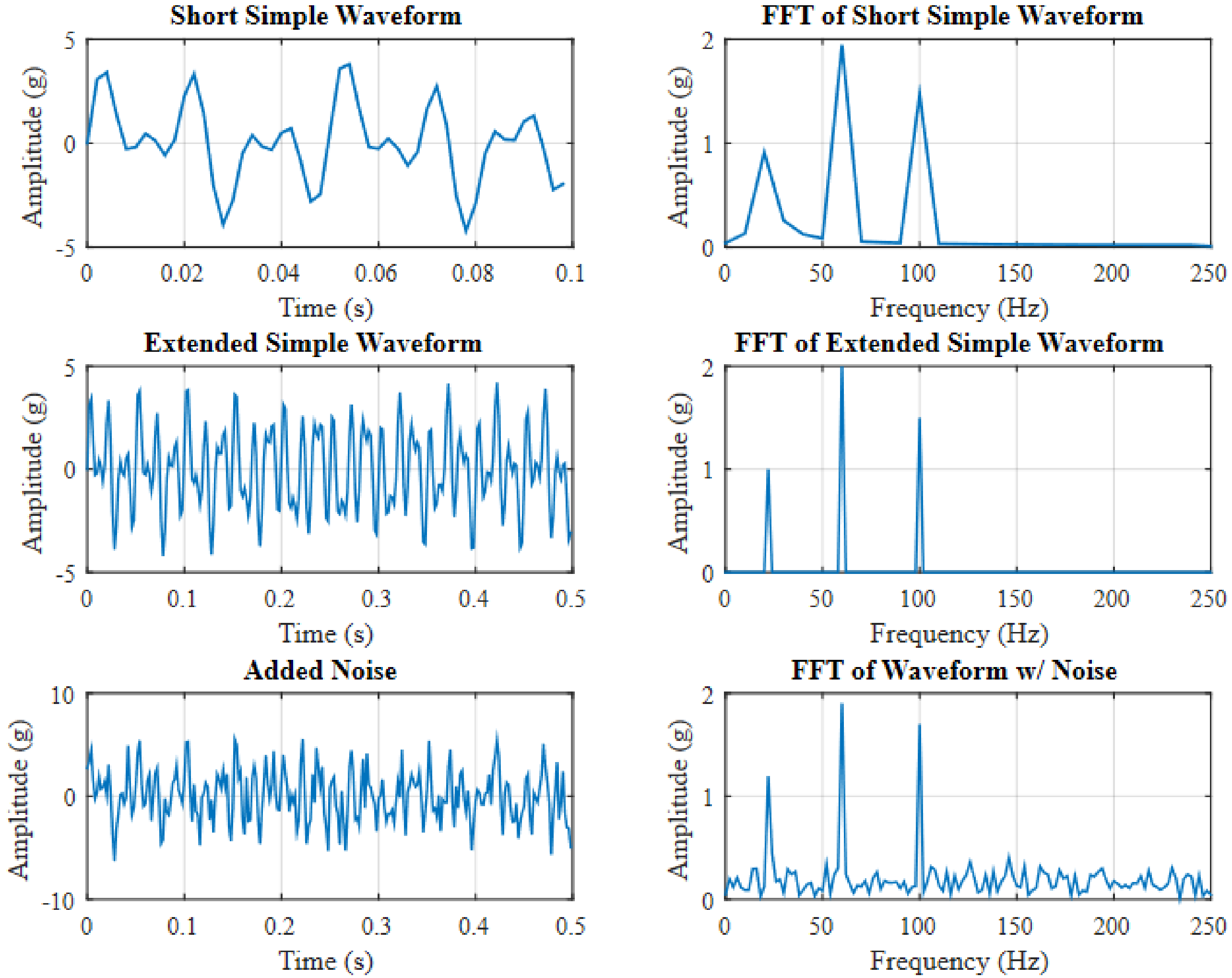
A FFT is an algorithm that computes the discrete Fourier transform (DFT) of a sequence. It provides spectrum analysis. In general, some physical components' frequencies and amplitudes in healthy condition are almost fixed.



Frequency spectrum from a HSS gear in a wind-turbine gearbox in healthy condition

Time domain vs FFT

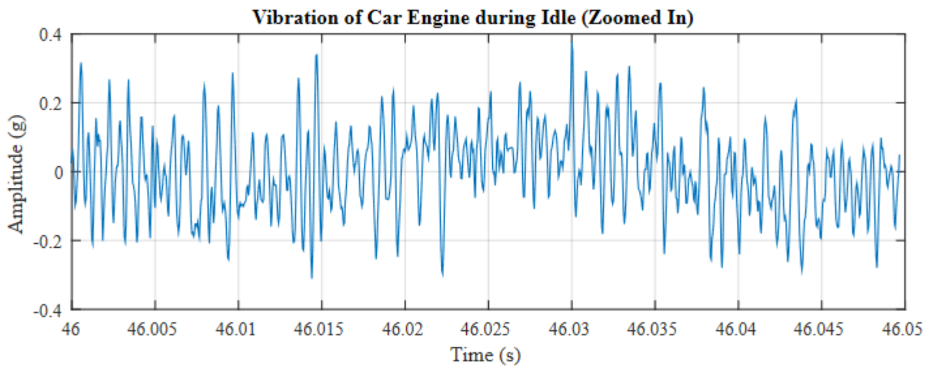
The signal constructed waveform will consist of three different frequency components: 22 Hz, 60 Hz, and 100 Hz. These frequencies will have an amplitude of 1g, 2g, and 1.5g respectively.

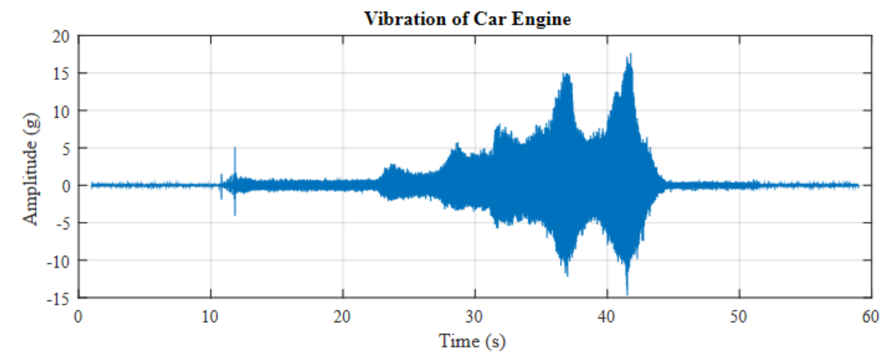


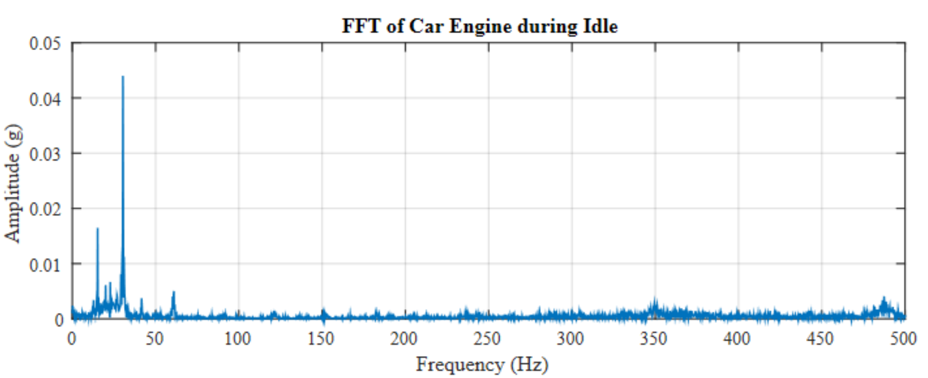
B. Spectrogram

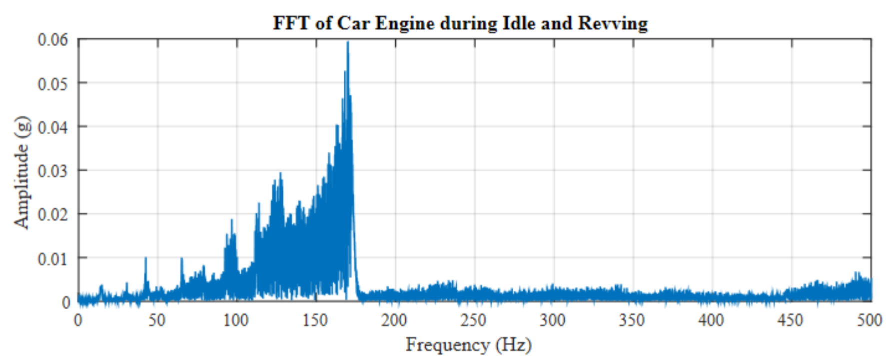
A ***spectrogram*** takes a series of FFTs and overlaps them to illustrate how the spectrum (frequency domain) changes with time. If vibration analysis is being done on a changing environment, a spectrogram can be a powerful tool to illustrate exactly how that spectrum of the vibration changes.

In many applications the vibration frequency will change with time and you can run into trouble if you only look at the FFT. Let's zoom out of the area where the car engine is running at a relatively fixed rate, and compute an FFT of the entire signal.

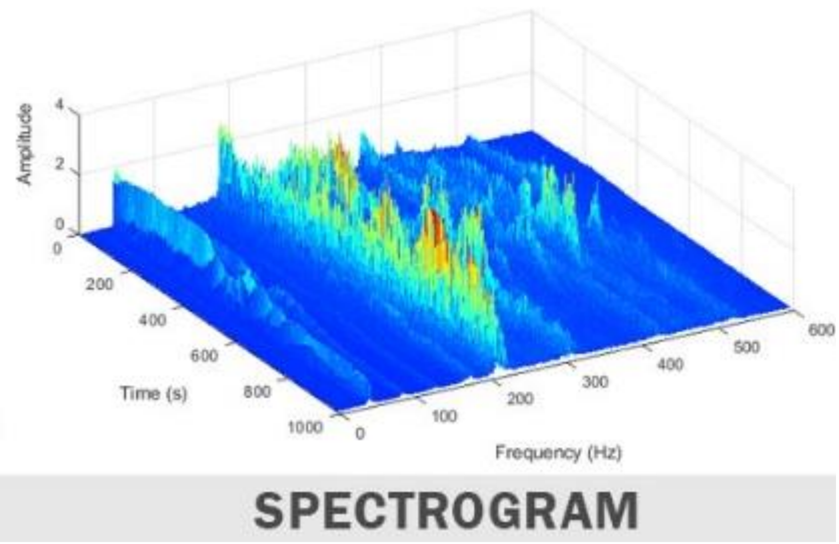


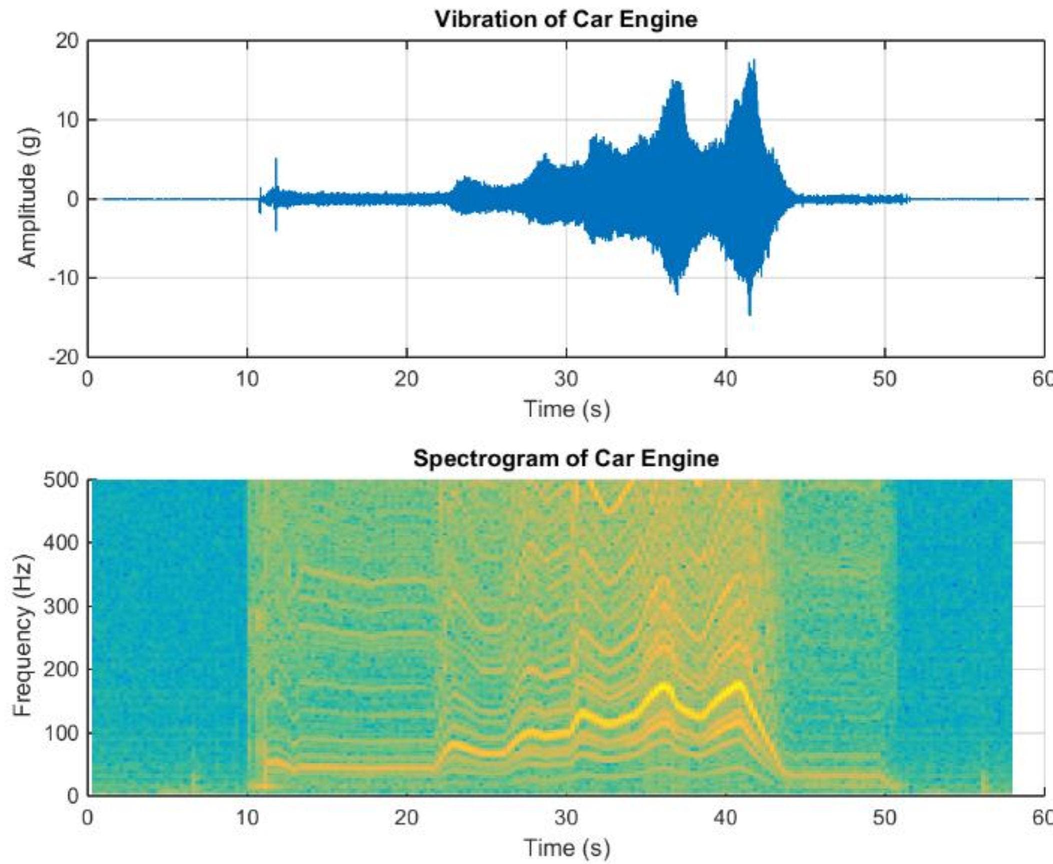






Spectrogram analysis



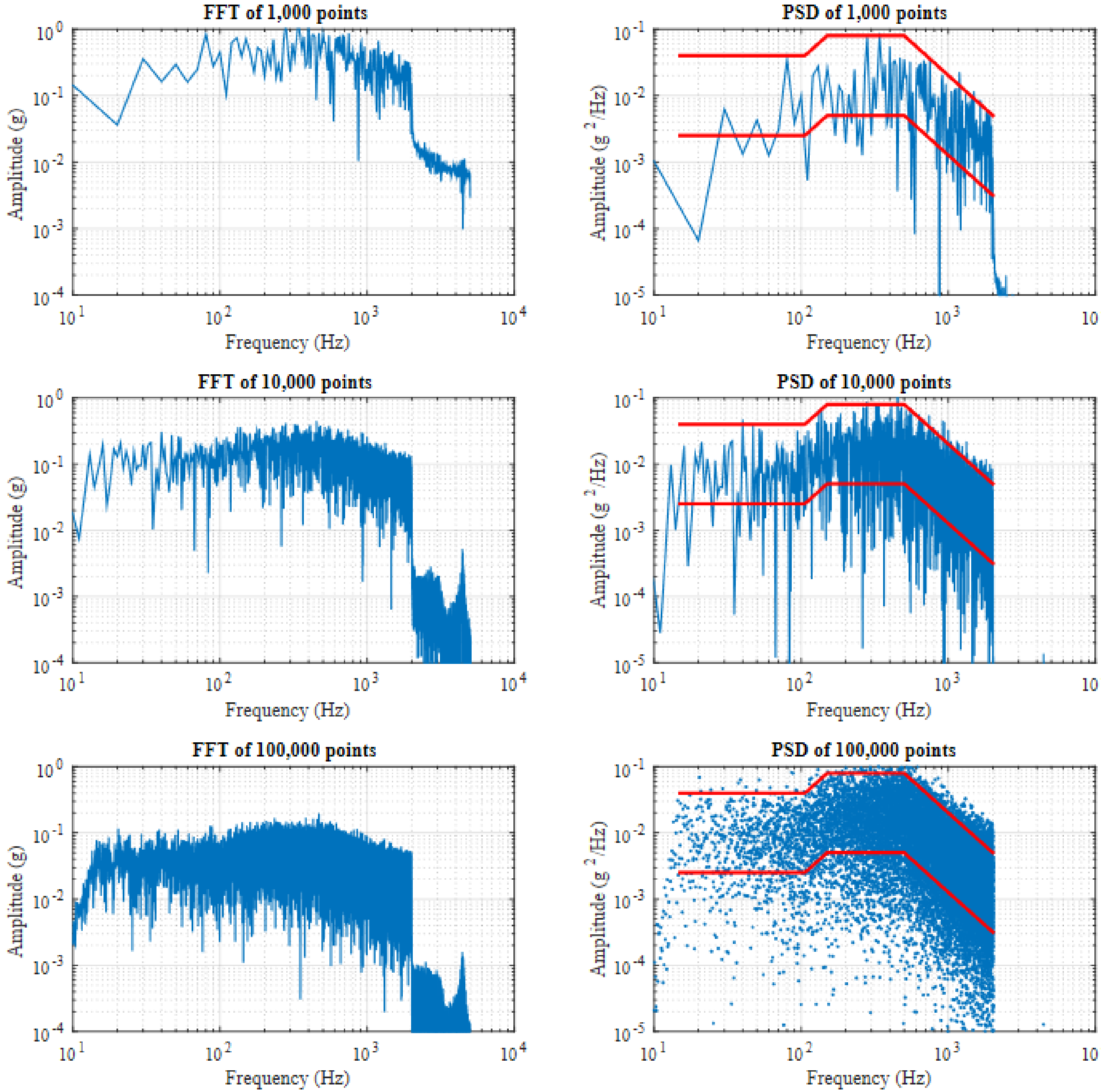


https://blog.endaq.com/vibration-analysis-fft-psd-and-spectrogram

C. Power Spectral Density (PSD)

A PSD is computed by multiplying each frequency bin in an FFT by its complex conjugate which results in the real only spectrum of amplitude in power square [2].

A PSD is typically used to characterize broadband random signals.

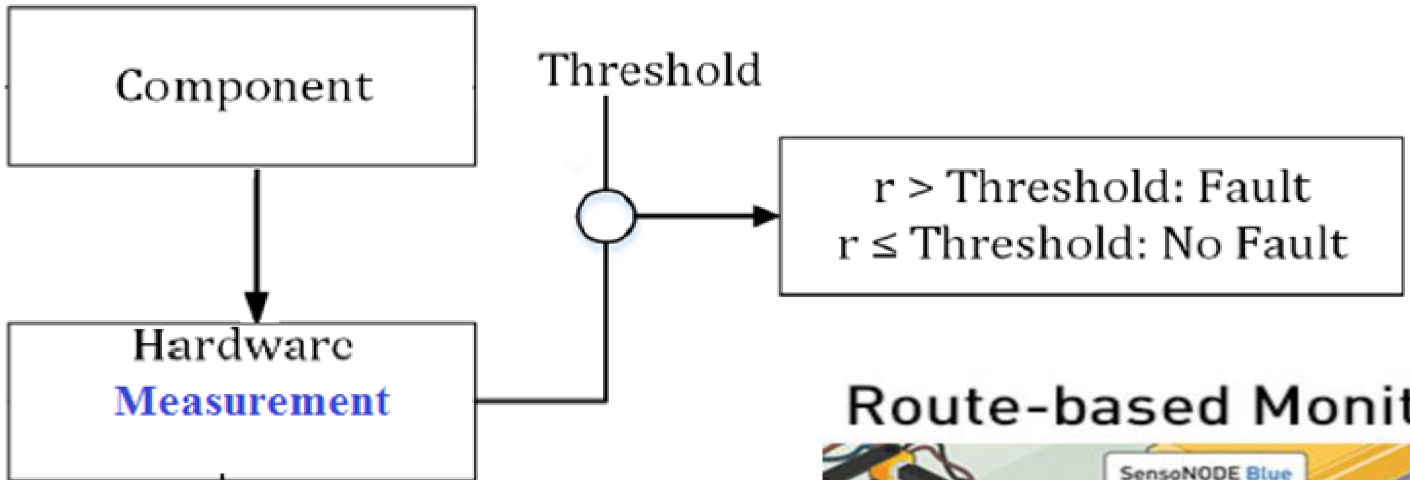

FFT of 1,000 points
PSD of 1,000 points
FFT of 10,000 points
PSD of 10,000 points
FFT of 100,000 points
PSD of 100,000 points
Amplitude (g)
Amplitude (g²/Hz)
Frequency (Hz)

2.1.2. Hardware sensor fault diagnosis method ---time domain

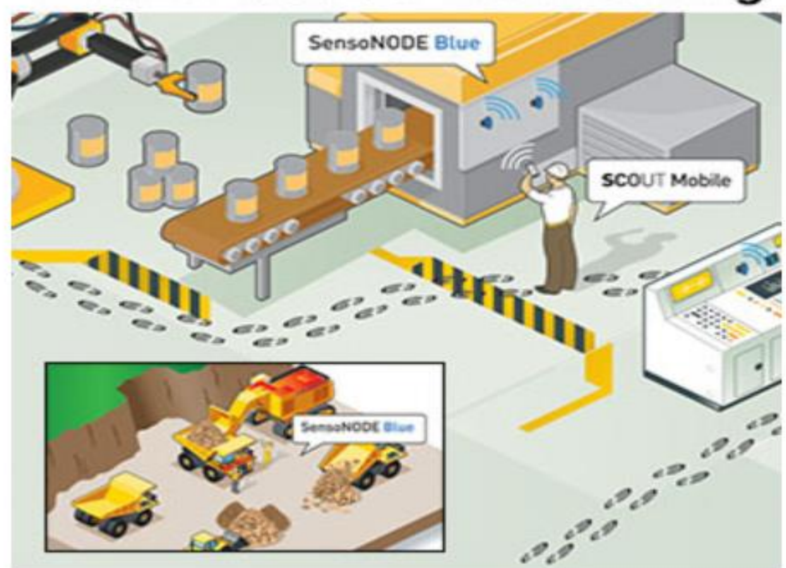
- **Threshold decision (compared with threshold)**

- **Trend analysis (sensitivity)**

A. Threshold decision method (compared with threshold)



Route-based Monitoring



Threshold---

1. Collect healthy data
2. Compute a statistical moment (mean)
3. Compute maximum error (sensor reading compared with mean)
4. Given a tolerance number (user defined from experiment)
5. threshold=

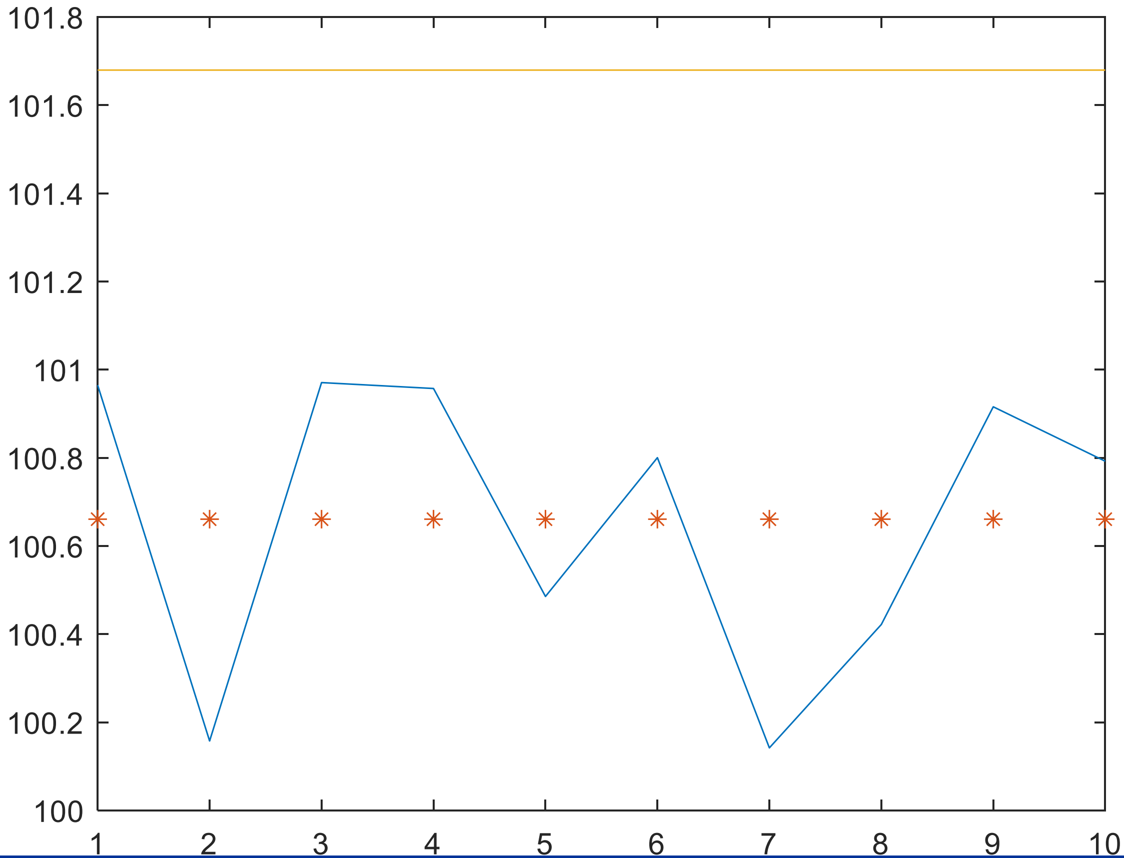
statistical moment+maximum error+tolerance number

Why we need to have a tolerance number?

For example, we have the following healthy data.

100.9649 100.1576 100.9706 100.9572 100.4854 100.8003 100.1419 100.4218 100.9157 100.7922

Average= 100.6608, maximum error=0.5189, tolerance number=0.5 (1.5). Threshold=101.6796



- **Logic rule of fault detection:**

If $|measured\ signal| > threshold$, fault;

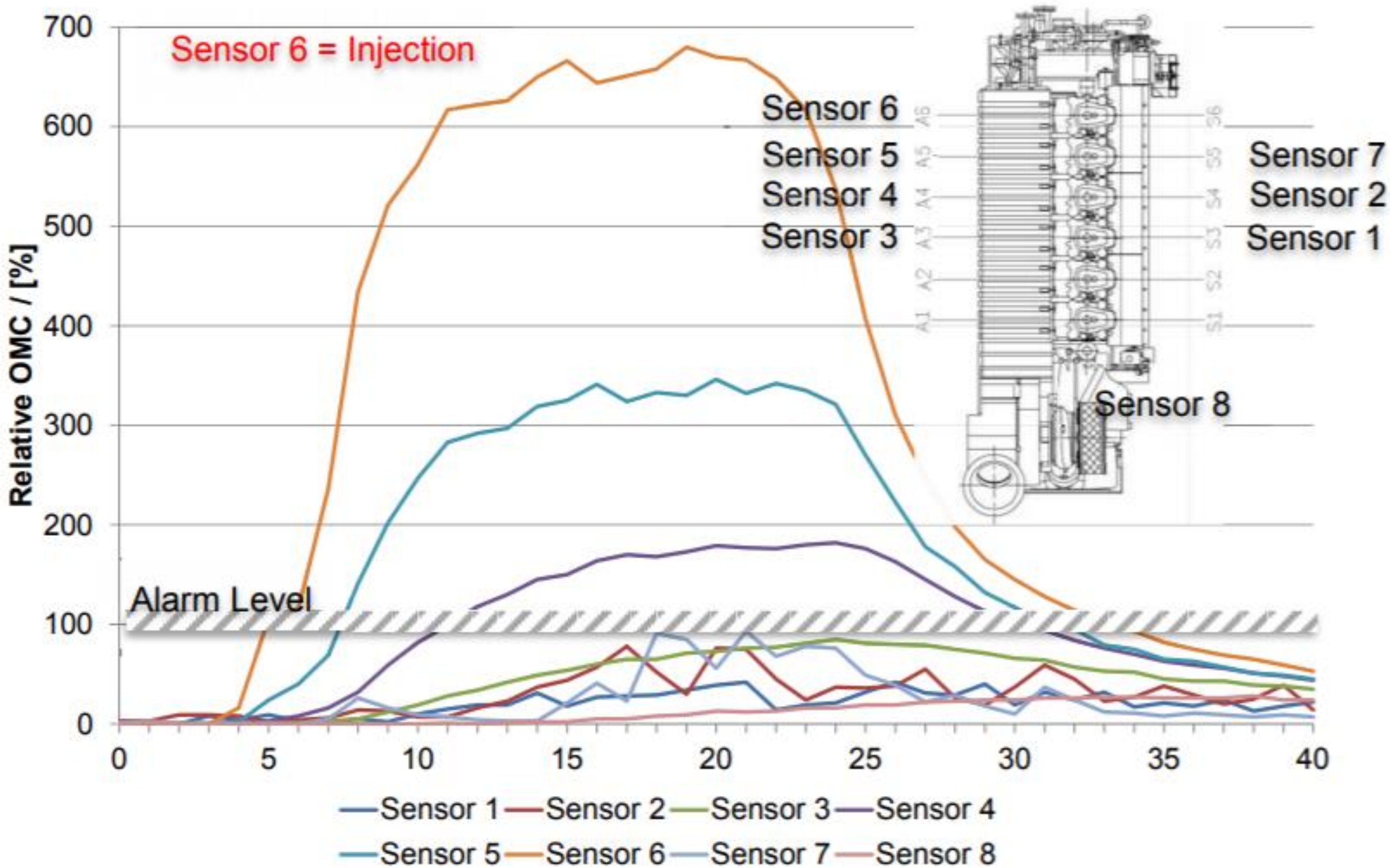
otherwise, no fault

$|\cdot|$ is the absolute value

Detection time:

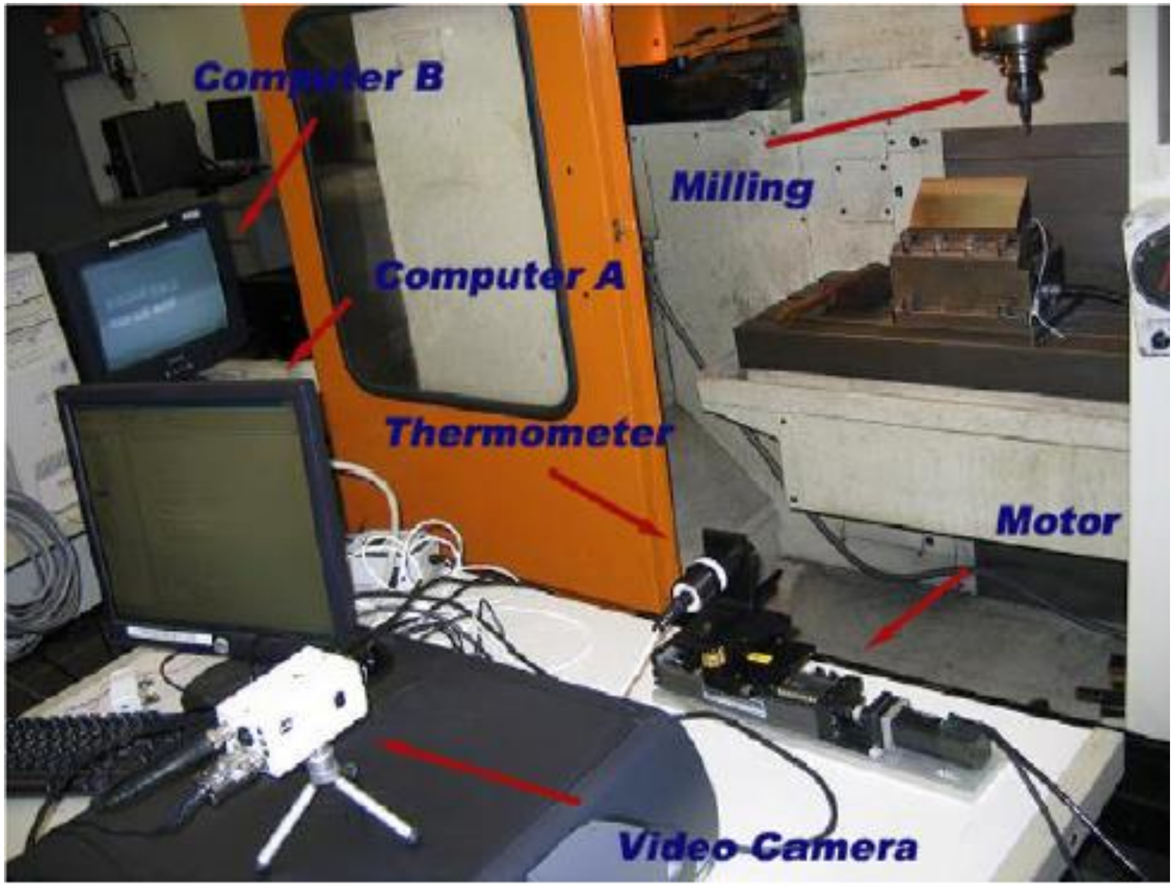
When |*measured signal*| > *threshold*, please record time immediately and this is the detection time.

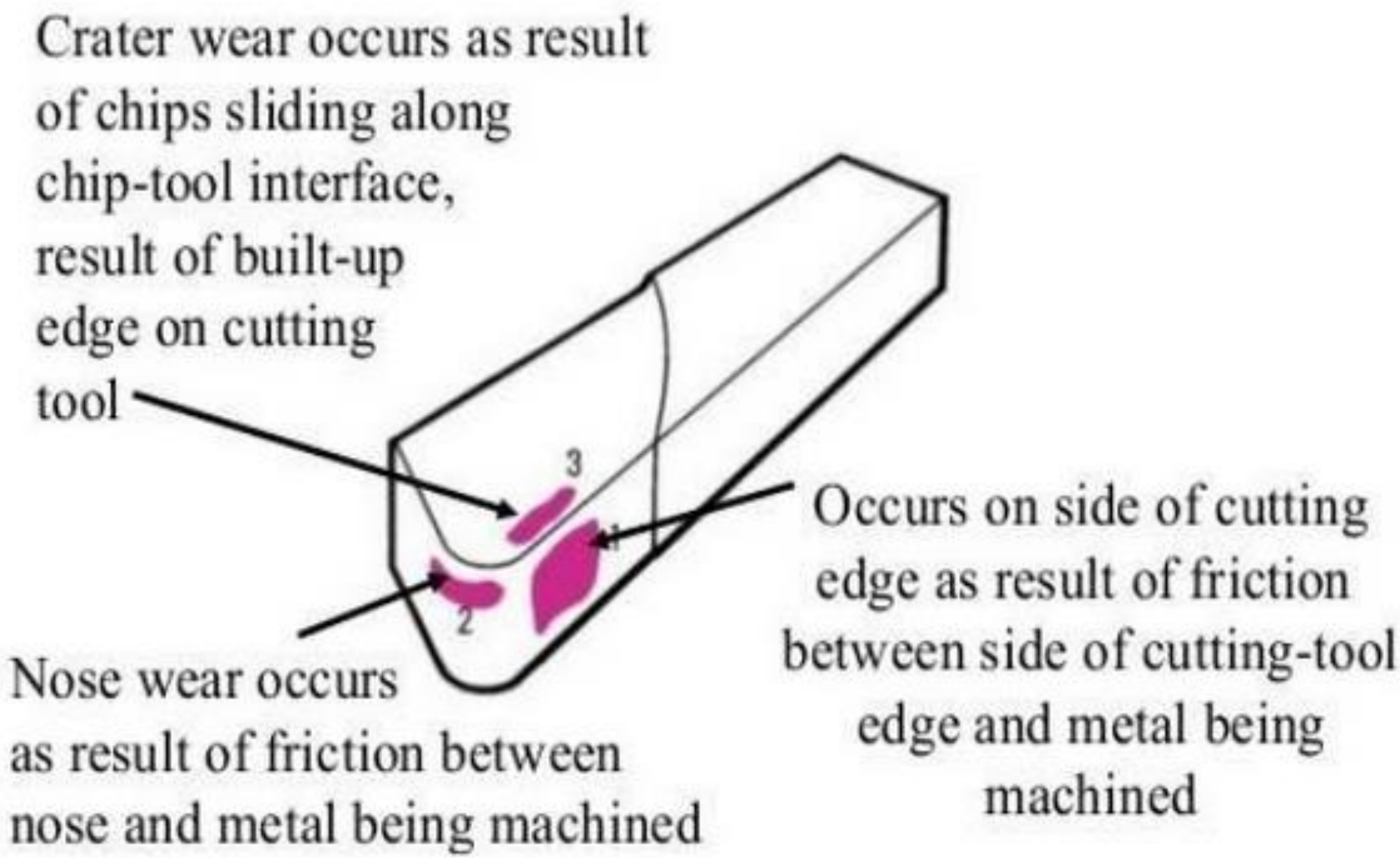
The relative oil mist concentration (OMC) monitor oil mist detection



One example: monitoring system of cutting tool

We use thermometer (contactless infrared) to measure the temperature of cutting tool

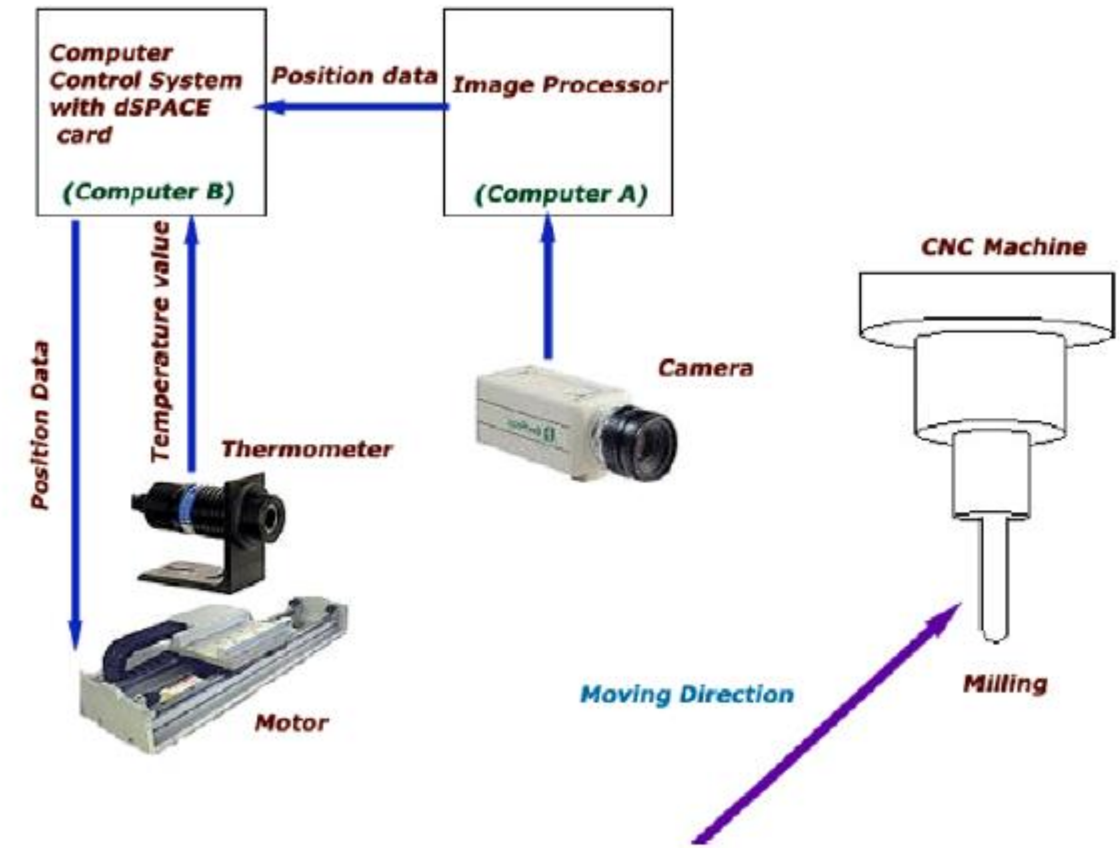


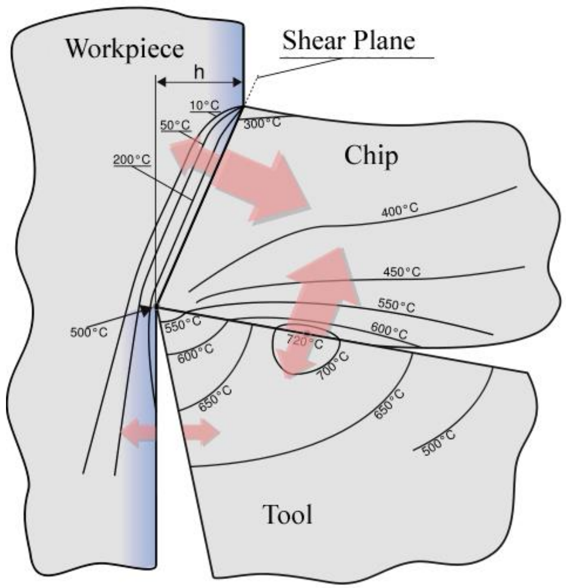


[Akshay Arvind], Tool Wear and Tool life of single point cutting tool , https://www.slideshare.net/akshayarvind3/tool-wear-and-tool-life-of-single-point

Cutting tool is moving to one direction. We have to follow it and do measurement

Cutting tool wear monitoring system

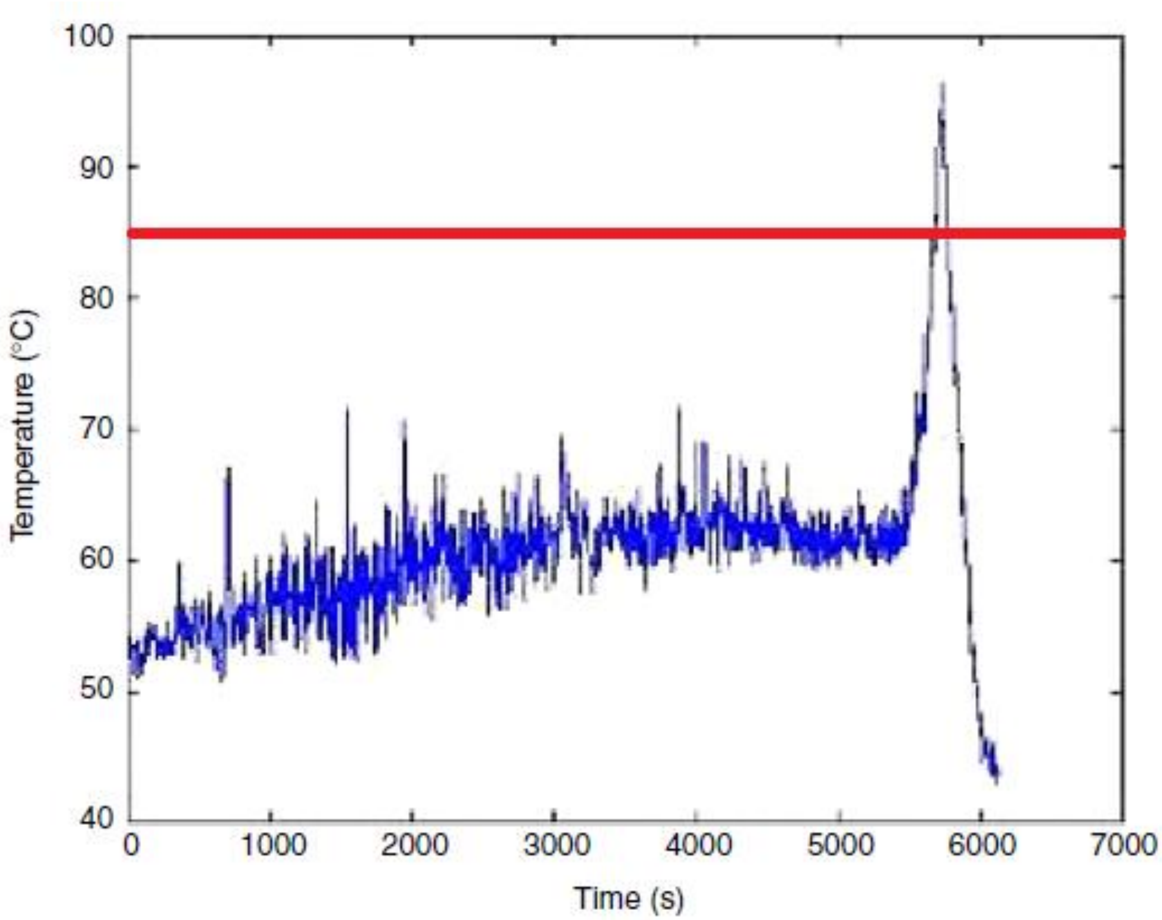




NUS National University of Singapore

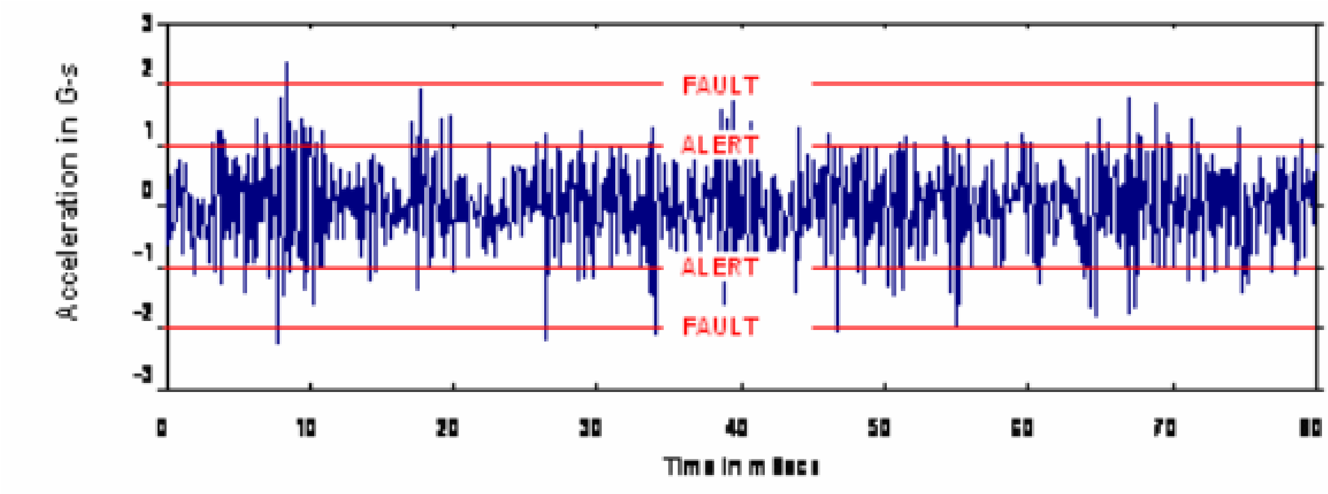
Our project

The threshold is 85.



Another example:

Monitoring Motor & Fan Vibrations



Time Waveform

B.Trend analysis

This method uses a sensitivity analysis.

It depends on the calculation of sensitivity.

NUS National University of Singapore

Sensitivity

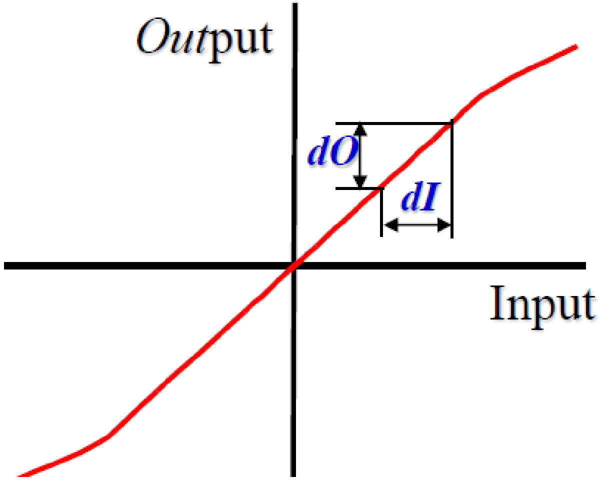
- Sensitivity : the rate of change in output corresponding to the rate of change in input *dO/dI. Also we use ΔO/ΔI where*

  Δ represents the changes of the variable.i.e,

$$\frac{\Delta O}{\Delta I} = \frac{O_i - O_{i-1}}{I_i - I_{i-1}}$$

The derivative of a function of a real variable measures the sensitivity to change of the function value (output value) with respect to a change in its argument (input value).

- At different range, the sensitivity may differ.



An example of sensitivity

- Input x =[0.0100 0.0200 0.0300 0.0400 0.0500];
- output y =[2.8810 3.01 3.86 4.71 4.75];

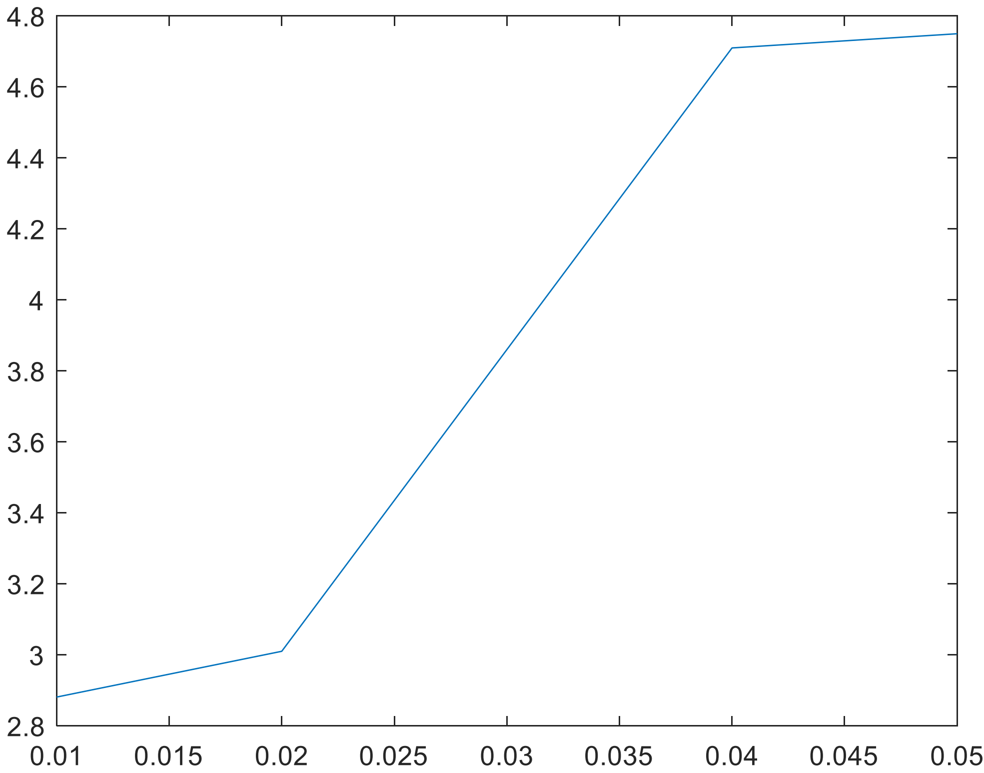
Sensitivity :

0.01--0.02 ΔO/ΔI=(3.01-2.8810)/0.01=12.9;

0.02-0.03 ΔO/ΔI=(3.86-3.01)/0.01=85

0.03-0.04 ΔO/ΔI=(4.71-3.86)/0.01=85

0.04-0.05 ΔO/ΔI=(4.75-4.71)/0.01=4.0



There are mainly two approaches to analyze sensitivity:

- Local Sensitivity Analysis
- Global Sensitivity Analysis

Local sensitivity analysis is derivative based (numerical or analytical). The term local indicates that the derivatives are taken at a single point.

At the sampling time Δt, ΔI, ΔO, Gain=$\frac{\Delta \boldsymbol{O}}{\Delta \boldsymbol{I}}$ or $\frac{\Delta \boldsymbol{O}}{\Delta \boldsymbol{t}}$

Global sensitivity analysis uses a global set of samples to explore the design space.

For trend analysis, we have to determine the range of the sensitivity.

Method:

- Collect the data from healthy state
- Compute all the sensitivities
- Find the maximum and minimum values which are the range of the sensitivity.

Trend analysis for fault detection

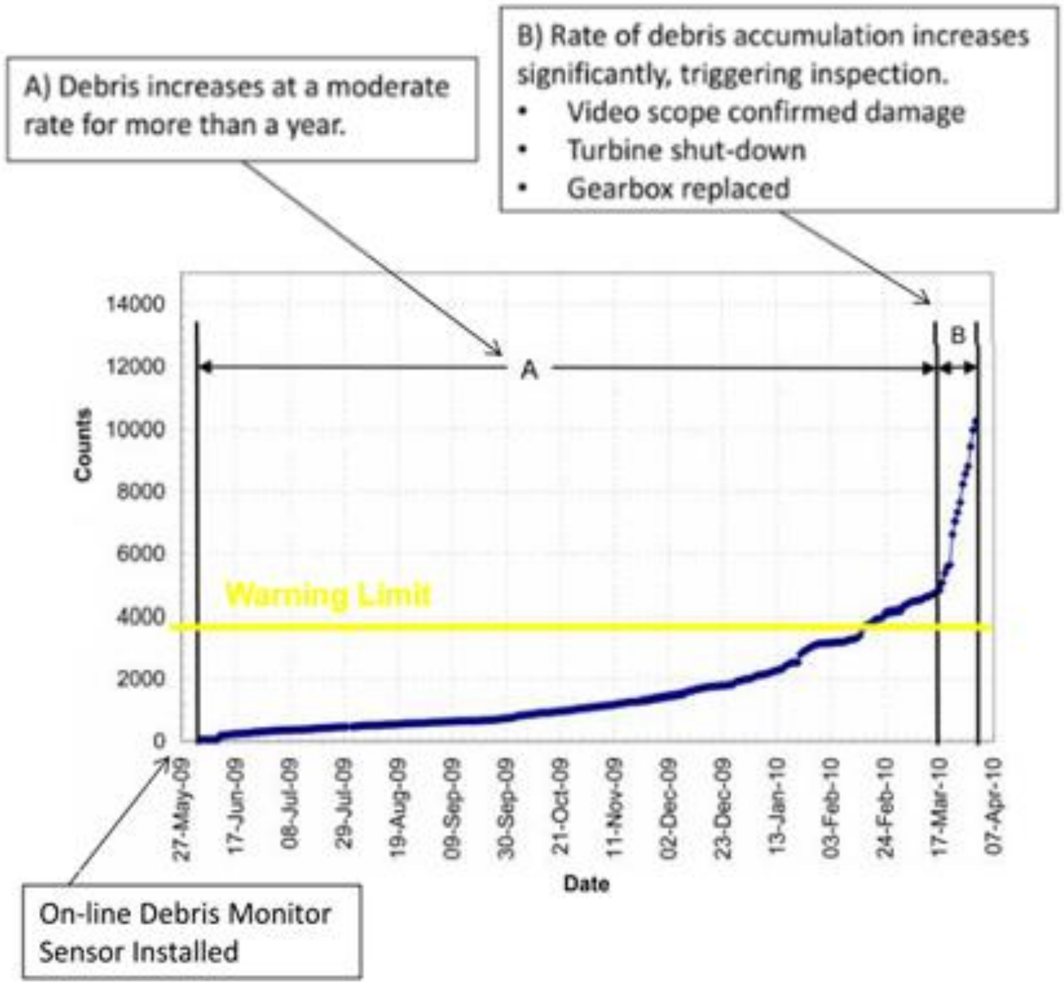
If $s_m \leq$sensitivity$\leq s_M$, healthy system;

Otherwise, fault occurrence.

where s_m is the minimum value of the sensitivities of all the collected data, while s_M is the maximum value of the sensitivities of all collected data

Fault detection time: Once the sensitivity exceeds the pre-defined range of the sensitivity, please record time which is the detection time.

One example: Gearbox fault detection

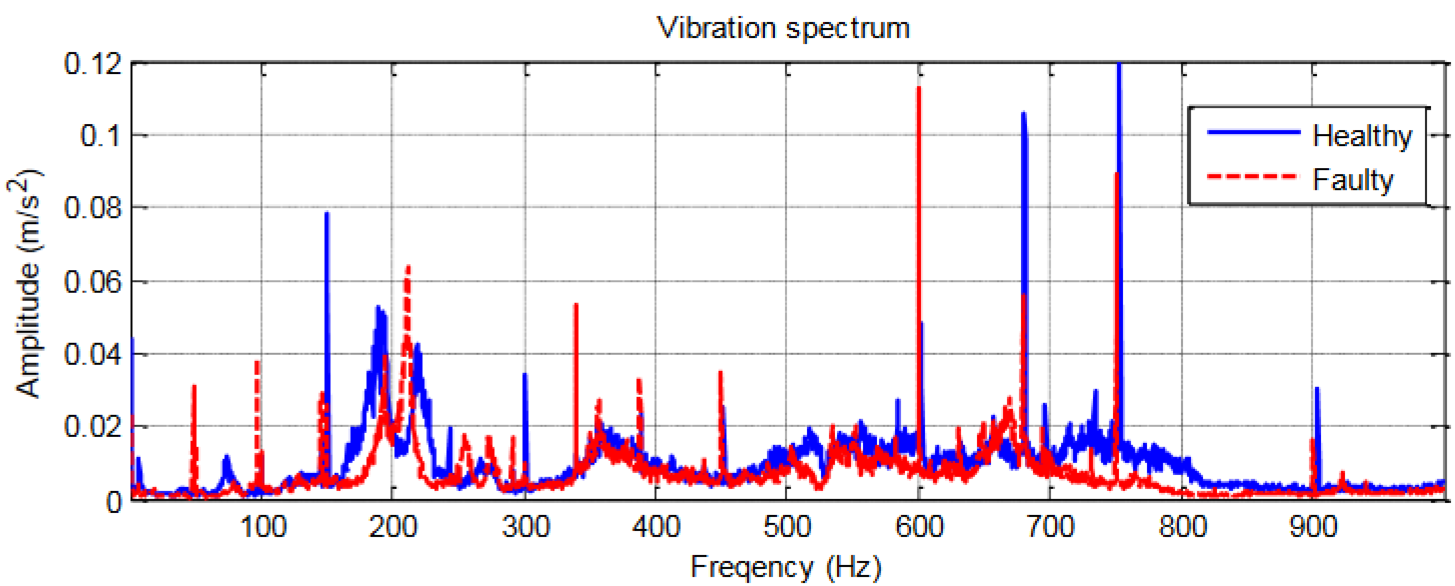


Example of gearbox fault detection by means of a wear debris sensor

2.1.3. Hardware sensor fault diagnosis method ---frequency domain!!!

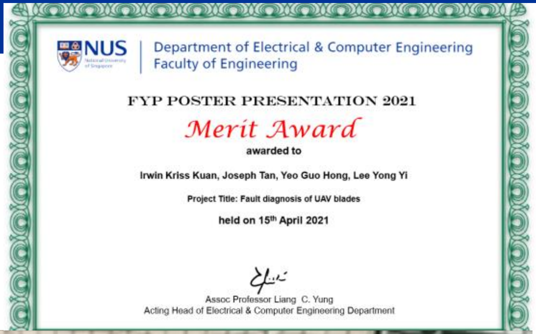
Taking into account the specific geometries and kinematics of the machine components, such as e.g. bearings or gearwheels, it is possible to distinguish between normal rotation and characteristic faulty frequencies. In this way, the vibration spectrum can reveal detailed diagnostic information about the location of a fault.

Fault diagnosis method is to check the FFT in healthy condition. And this FFT is regarded as a reference. Later, all FFT obtained will compare with the reference. If the new FFT exceeds the amplitudes or frequencies of the reference, the fault occurs. Different frequencies may represent different devices. For example, rotation speed, bearing , or gear box.



I cited this picture from OSAMA HASAN INBAIA HAMOMD's Ph.D thesis 2017, University of Huddersfield

One FYP project uses this method to detect propeller's fault


NUS
Department of Electrical & Computer Engineering
Faculty of Engineering

FYP POSTER PRESENTATION 2021

Merit Award

awarded to

Irwin Kriss Kuan, Joseph Tan, Yeo Guo Hong, Lee Yong Yi

Project Title: Fault diagnosis of UAV blades

held on 15th April 2021

Assoc Professor Liang C. Yung
Acting Head of Electrical & Computer Engineering Department

2.1.4 Hardware sensor fault diagnosis--condition monitoring

Condition monitoring is based on being able to monitor the current condition and predict the future condition of machines while in operation. Thus it means that information must be obtained externally about internal effects while the machines are in operation.

Working principle and goal of condition monitoring

The basic principle of condition monitoring is to select a physical measurement which indicates that deterioration is occurring, and then to take readings at regular intervals. Condition monitoring is used for detecting changes or trends in controlling parameters or in the normal operating conditions which indicate the onset of failure.

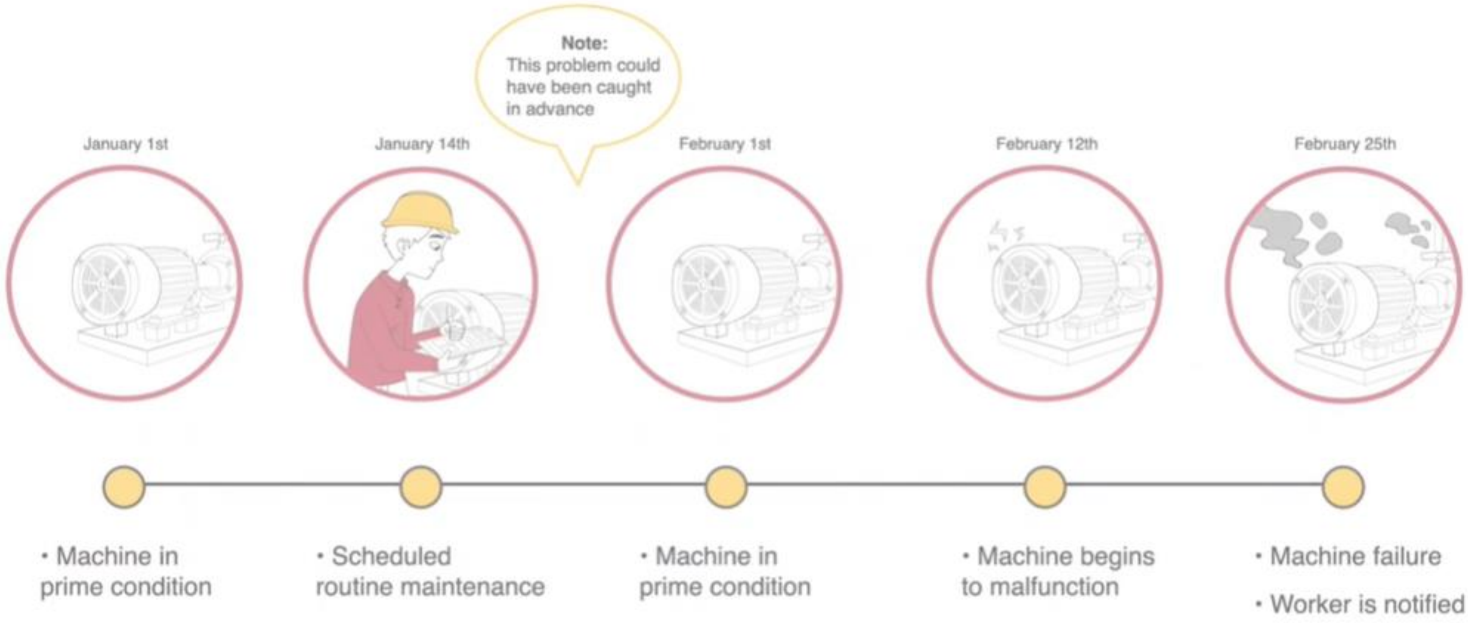
The goal of condition monitoring is to examine its functional health.

Faults always develop in machinery and even the most thorough and comprehensive routine maintenance programme cannot stop that. Therefore, it's important to have condition monitoring put in place as it puts you in the driving seat to actively prevent breakdowns and optimize maintenance resources where and when they are needed.

http://sensor-works.com/why-condition-monitoring-is-important/

Some benefits include:

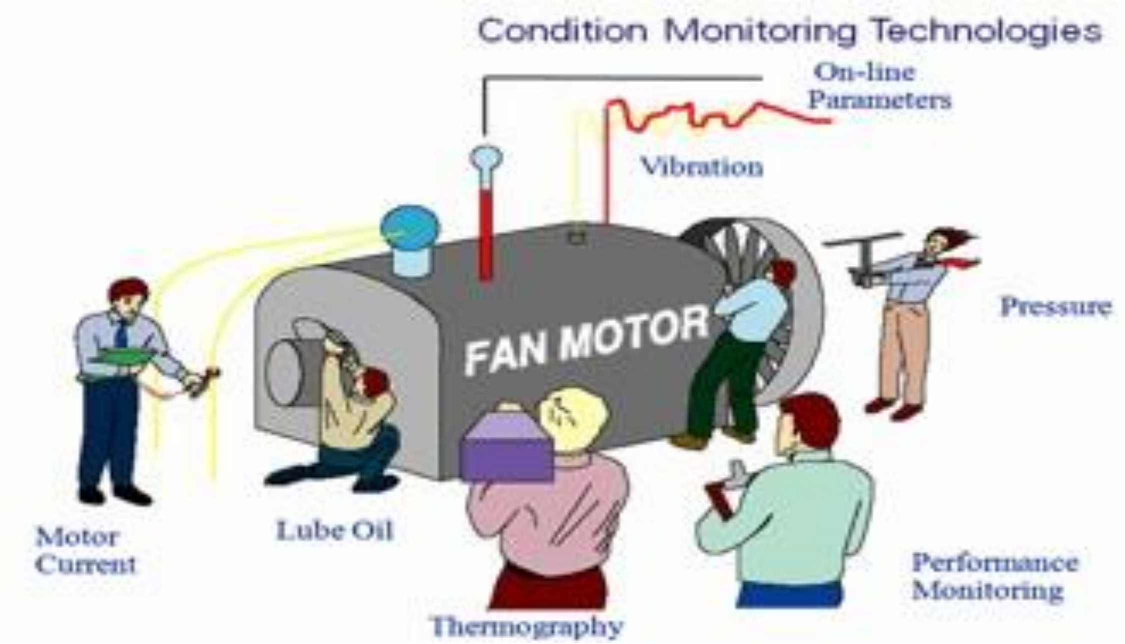
- Reducing machine failures
- Extend asset life
- Increase the rate of production
- Cut maintenance costs



Measurement Technologies

Many sensors detect changes in equipment components.

The following figure shows a variety of condition monitoring technologies which help understand health.

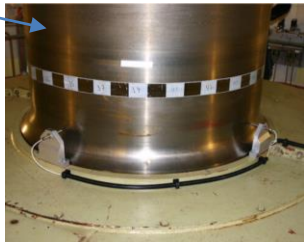


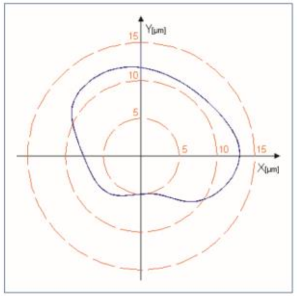
Sensors used in condition monitoring

Resistance Temperature Detector (RTD)

Measurement	Sensor	Frequency range
Vibration	Accelerometer	>100Hz
Vibration	Proximity probe (displacement)	<300Hz
Speed	Proximity probe	<300Hz
Motor current	Current shunt Current clamp	Up to 50kHz
temperature	RTD Thermocouple Infrared camera	Up to 10 Hz Multiple frames per sec
Pressure	Dynamic pressure	>100Hz
High-Frequency “Noise”	Ultrasonic	>20 kHz

Unlike accelerometers, which measure acceleration to determine vibration, proximity probes are noncontacting transducers that measure distance to a target. These sensors are almost exclusively used in rotating machinery to measure the vibration of a shaft.





Shaft Orbit: the path of a shaft rotation.

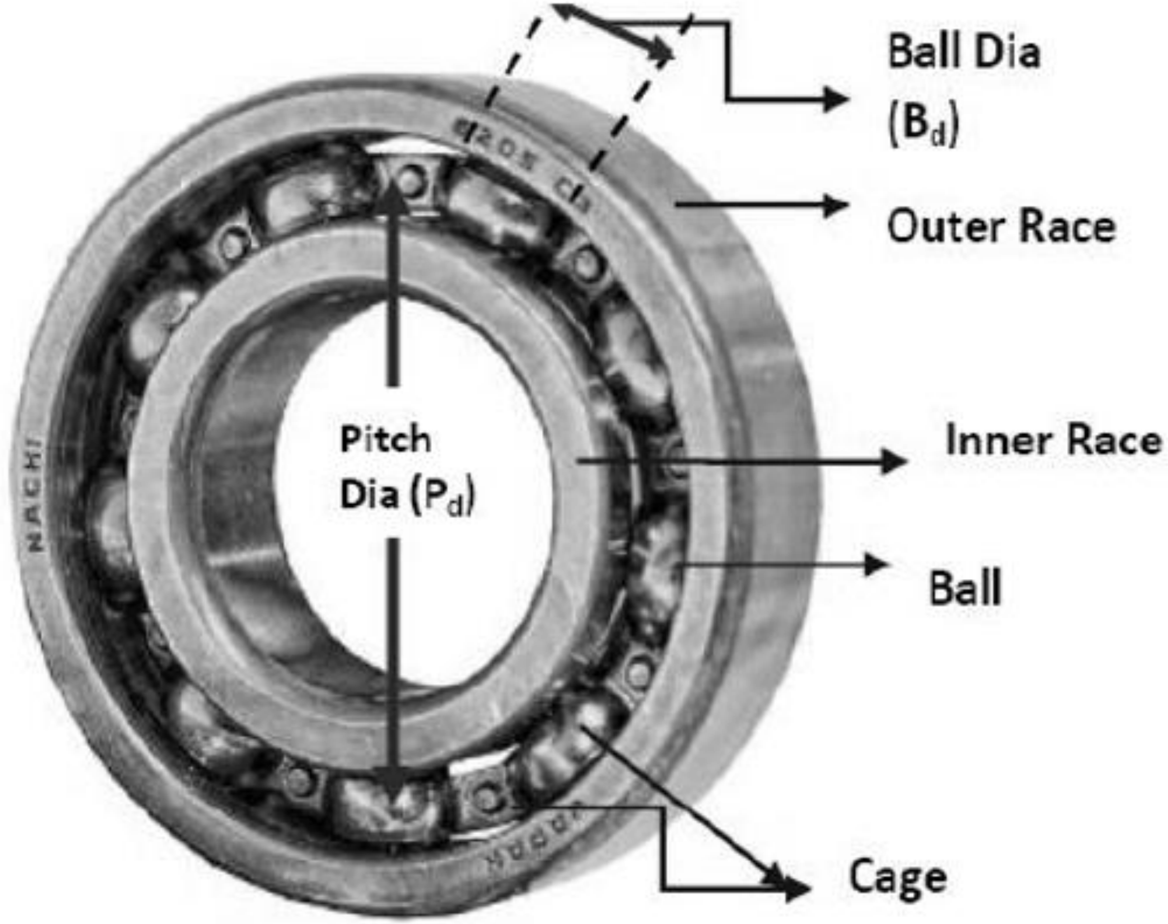
Each sensor offers the ability to monitor the degradation of mechanical and electrical components

- Vibration sensors are used to detect roller bearing wear, gearbox wear, shaft misalignment, unbalance, and mechanical looseness.
- Speed sensors work with vibration sensors to correlate vibrations to rotating speed and shaft angular position.
- Motor current sensors are commonly placed at the motor control center. They can detect eccentric rotors, loose windings, rotor bar degradation, and electrical supply unbalance.
- Dynamic pressure sensors are used for combustion dynamics, flow turbulence, and cavitations.
- Temperature sensors are typically used to detect heat caused by friction. They often accompany vibration sensors to collaborate vibration-detected degradation.
- Thermal imaging detects hundreds of temperatures within the camera’s field of view.
- Ultrasonic sensors can detect electrical problems including corona, arcing, and tracking. They can also be used to detect early signs of roller bearing wear.

There are pros and cons to each monitoring system and the selection of the correct techniques will depend on the machine type, mode of operation and failure modes

For example, vibration monitoring is very good at detecting a range of faults in machinery that rotates at a fixed speed, but may fail to detect problems in a noisy environment.

An example of condition monitoring ----vibration measurement

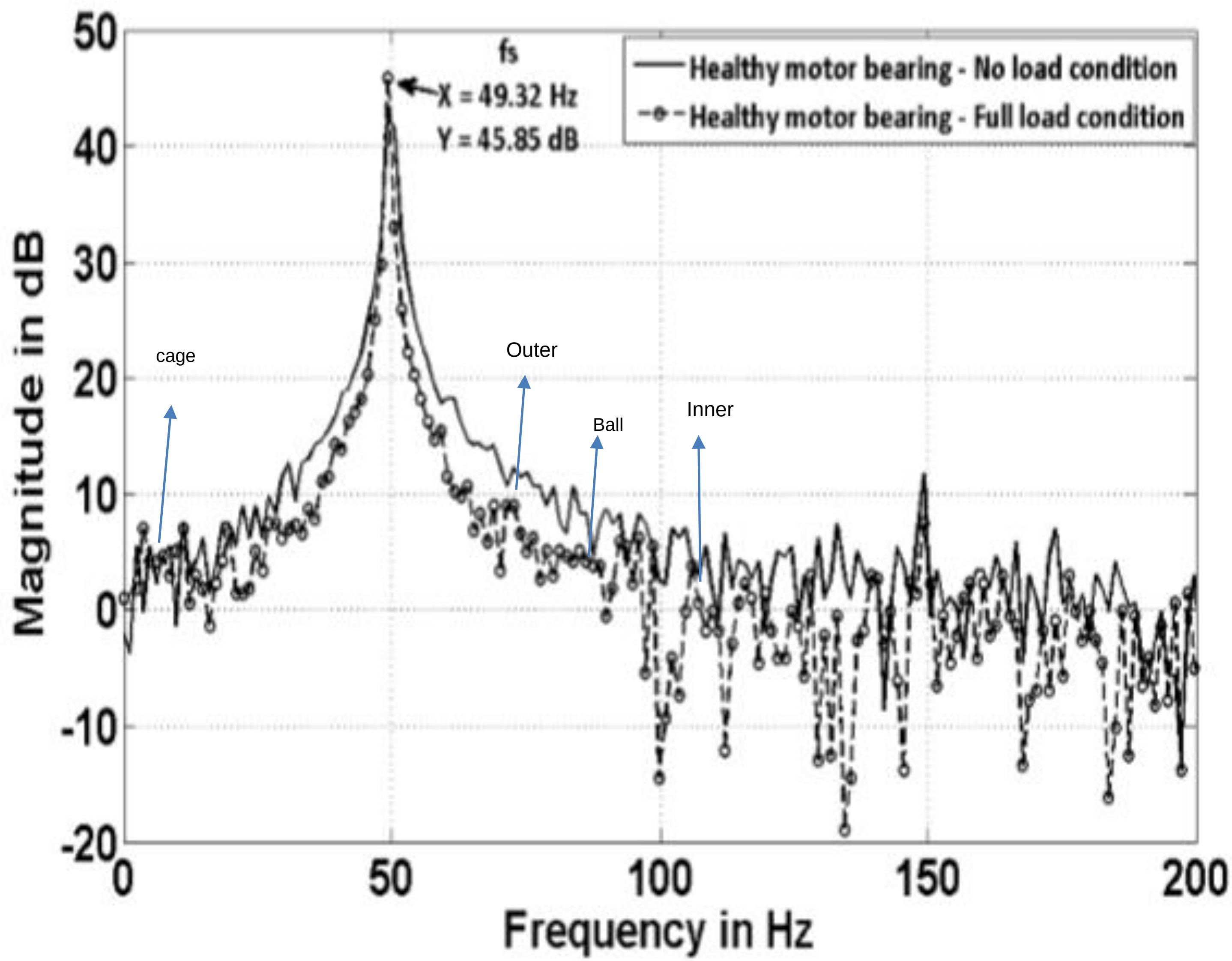


Application of MEMS Accelerometer for Detection and Diagnosis of Multiple Faults in the Roller Element Bearings of Three Phase Induction Motor,2015

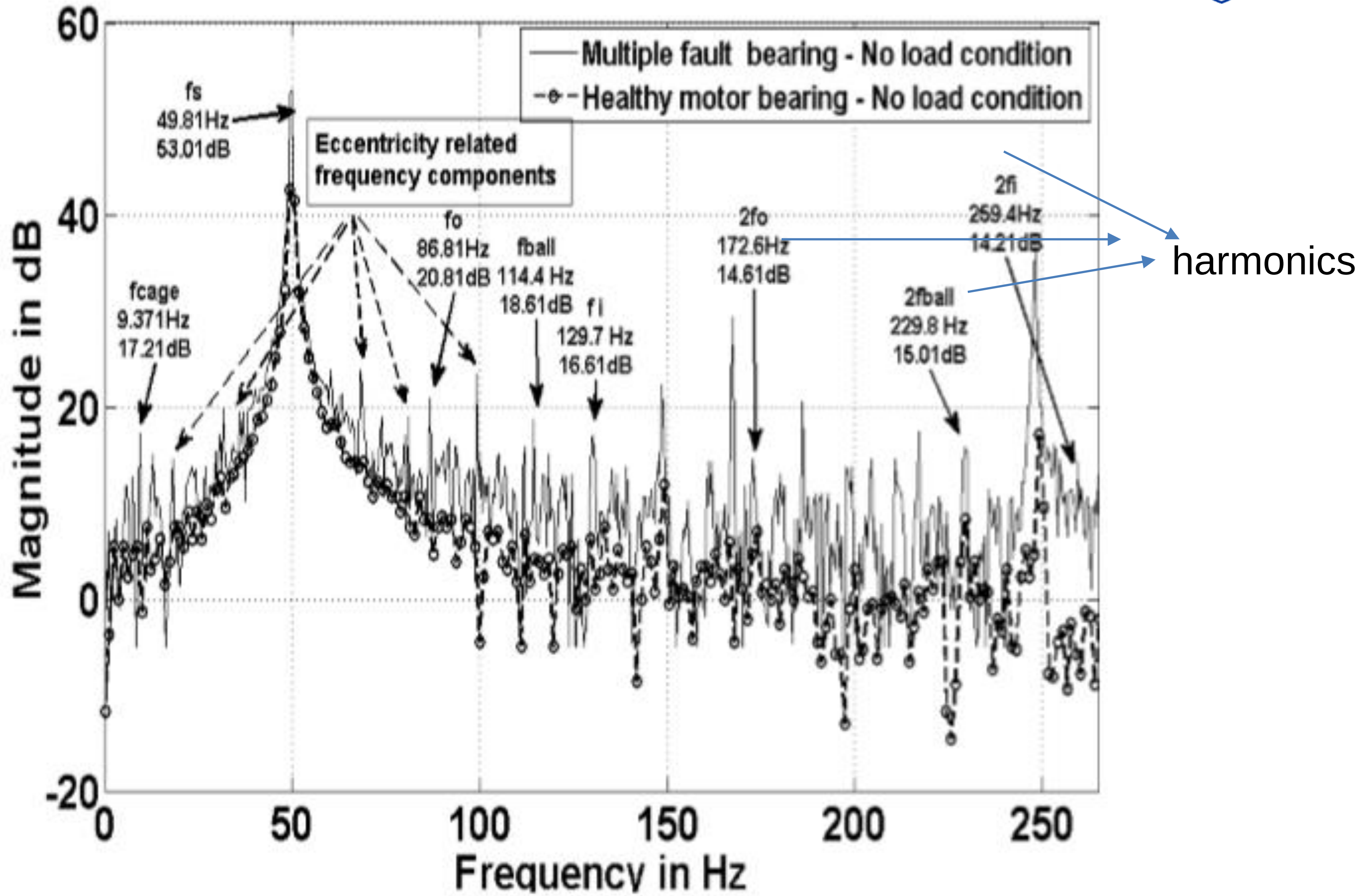
There are four possible faults:

inner race f_i; outer race f_o ;ball spin f_{ball} ;Cage defect f_{cage}

Photograph of the defective inner race from bearing removed from Gearbox



Comparison of healthy bearing under no load and full load condition.



Comparison of multiple faults bearing with healthy bearing – no load condition.

Advantages of hardware diagnosis:

- Frequency, threshold or sensitivity can be obtained easily
- Approach is simple

Disadvantages of hardware diagnosis:

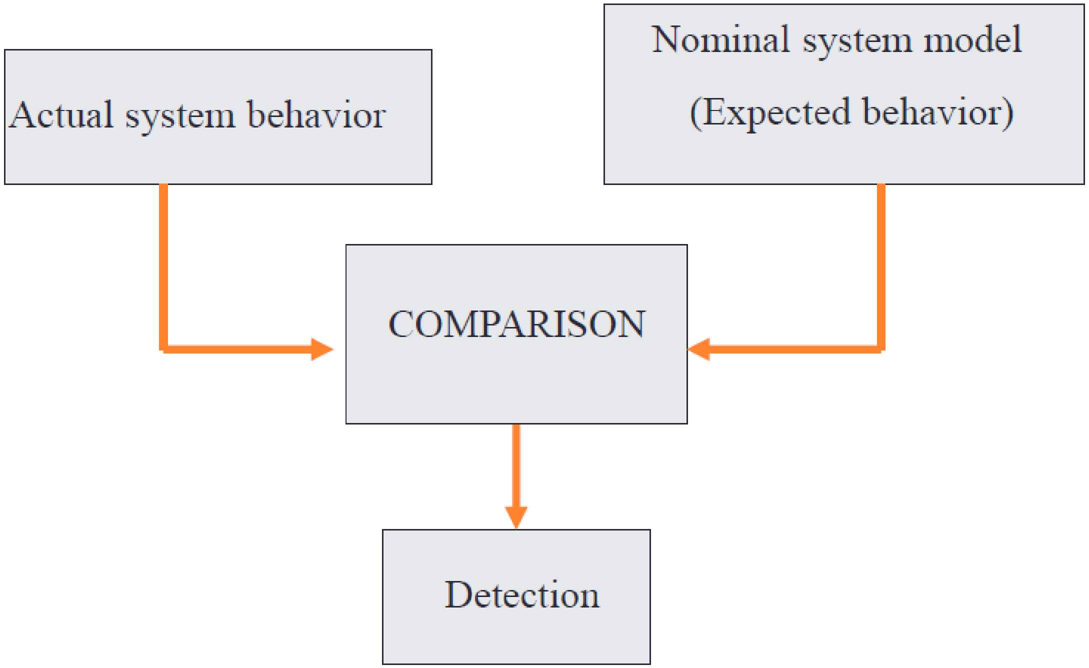
- Size
- It is difficult to maintain the hardware systems
- Sometimes we have to use multiple sensors to detect faults.

Now we focus on software method---model-based fault detection, where the model represents the system， even we cannot measure some variables.

3. Model-based fault detection method

Working principle of model based approaches:

- Compare actual system with a nominal model system

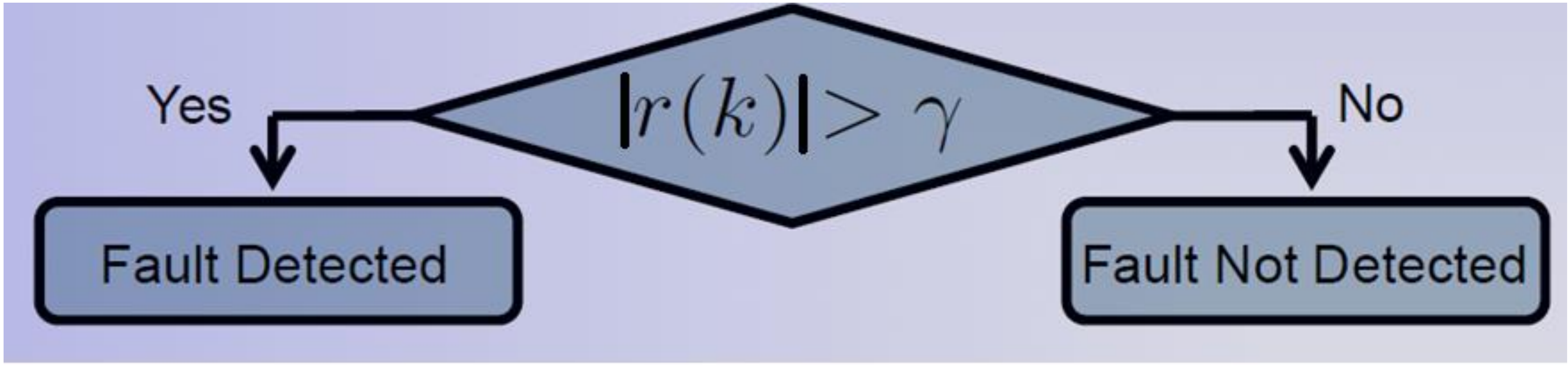


Standard model-based approaches for fault detection (I)

- Compute estimated output
- Generate a residual using the actual output

$$r(k) = y_{\text{true}}(k) - y_{\text{estimated}}(k)$$

- Compare it with a given threshold



Threshold calculation

Mean+deviation+small tolerance number

Mean +α ×deviation $(\alpha > 1)$

Solve the equation to derive the threshold. For example, linear system can be expressed by

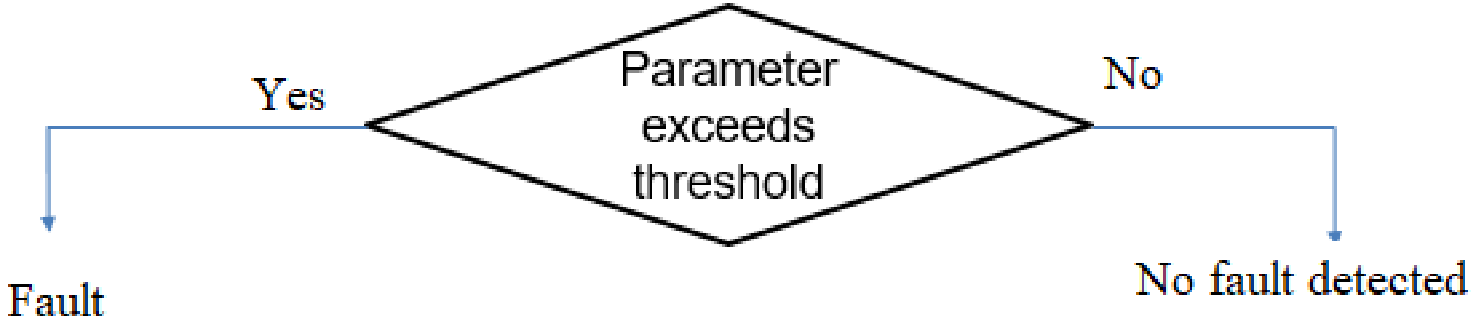
$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$\boxed{y(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t)}$$

Standard model-based approaches for fault detection (II)

$$\dot{x}(t) = Ax(t) + Bu(t)$$

- Compute estimated parameters (healthy state) and their threshold
- Compare estimated parameters (current time) with fixed healthy parameter thresholds

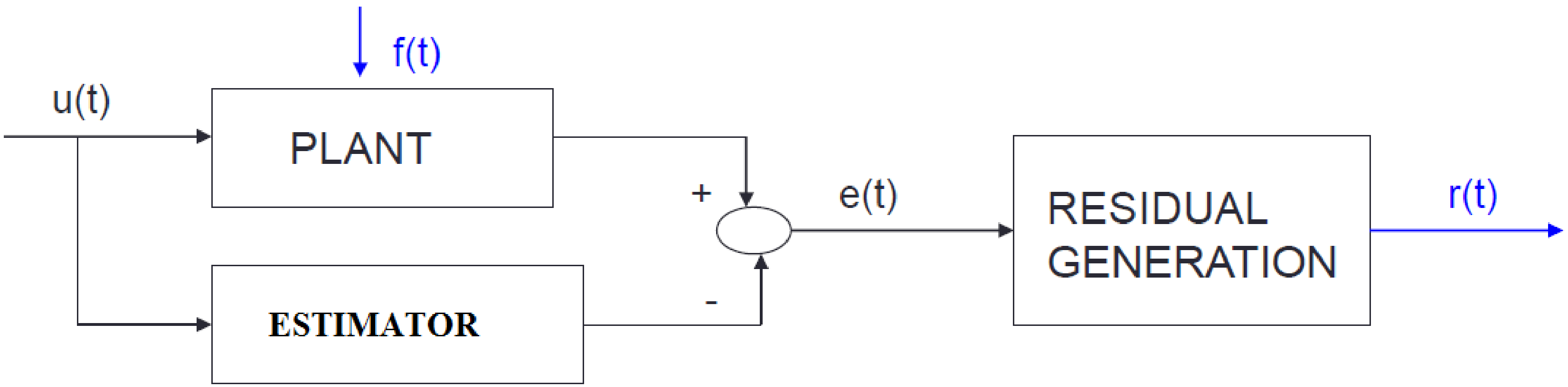


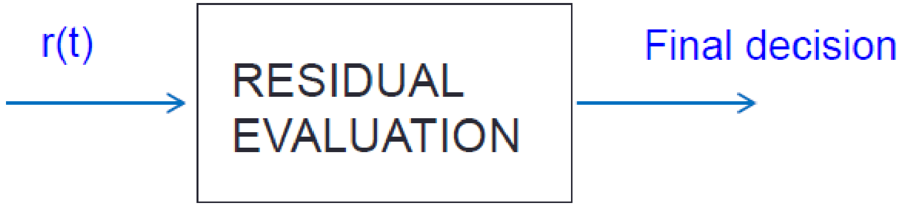
Model includes

- Estimator (linear or nonlinear model observer)
- Filter (linear or nonlinear *Kalman* filters)
- Nonlinear neural network model (nonlinear systems)

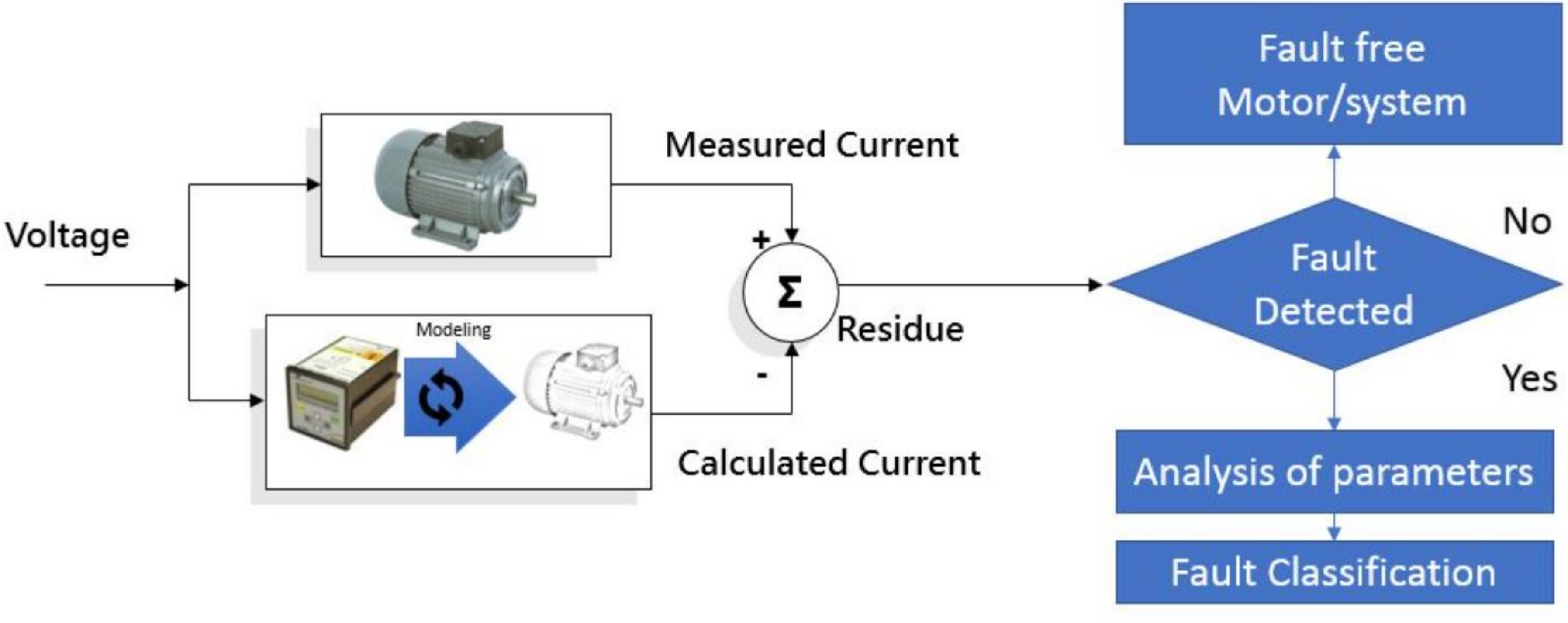
How do we get a model?

Model-based method

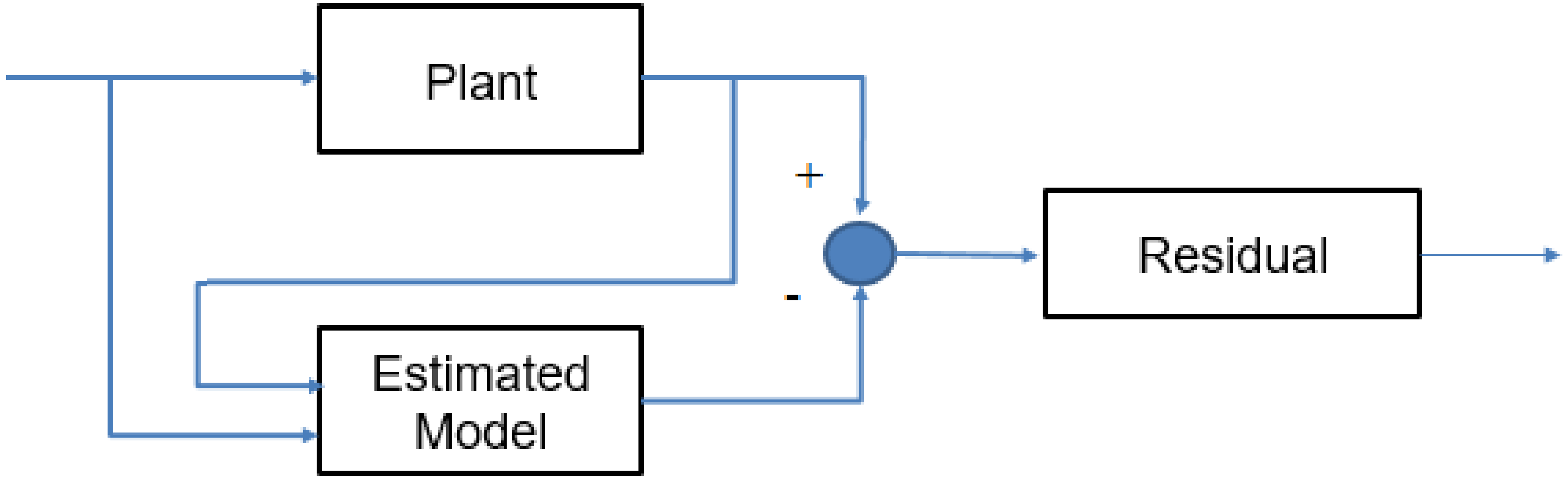


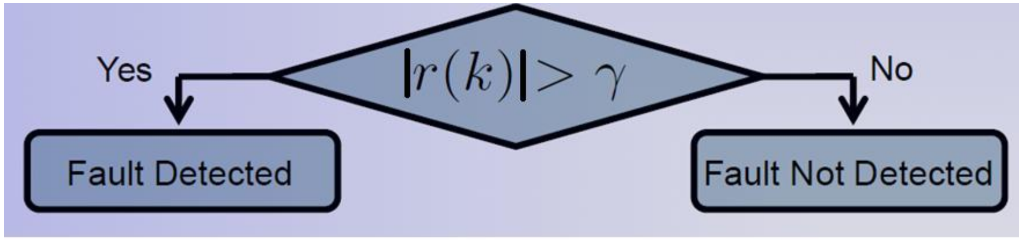


An example of model-based detection

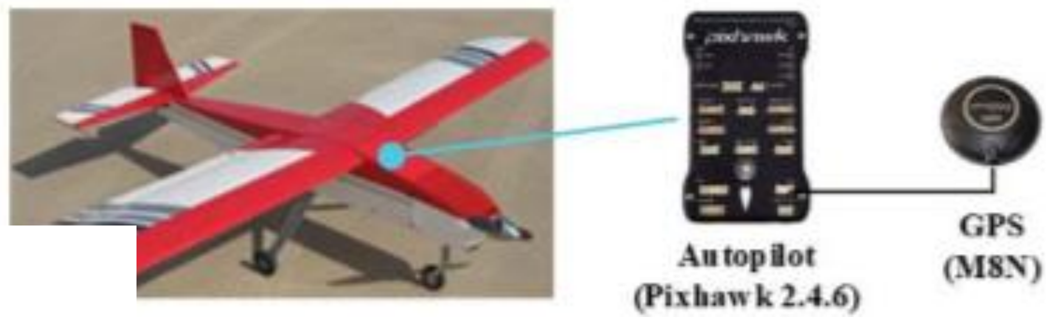


Another type of model –based method





Fault detection of unmanned aerial vehicle (UAV)

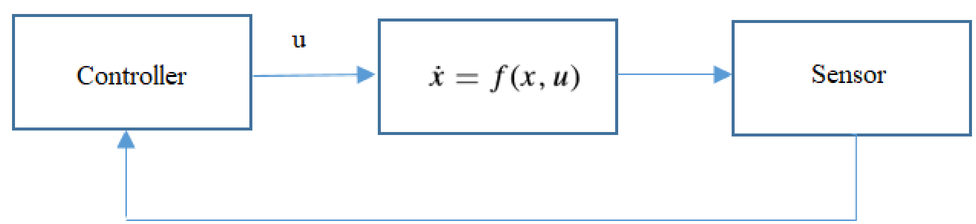


We can use a nonlinear model to represent it.

$$\dot{x} = f(x, u)$$

$$x = [p_x \, p_y \, p_z \; u \;\; v \;\; w \;\; \gamma \;\; \theta \;\; \psi \;\; \omega_x \, \omega_y \, \omega_z]^T$$

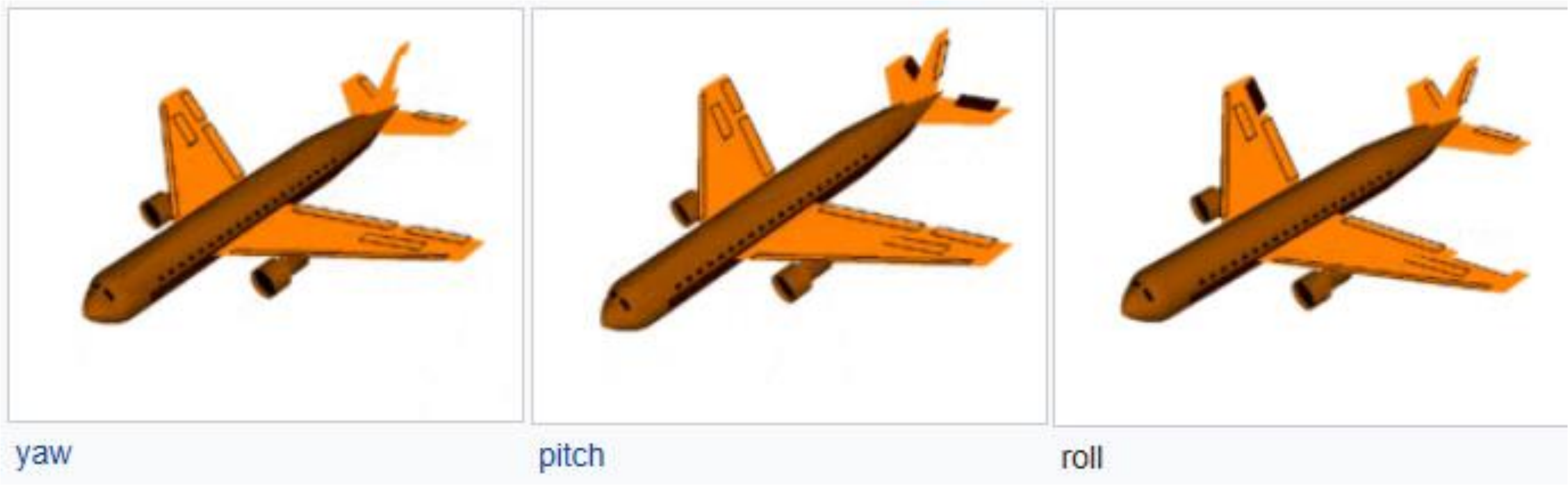
$$u = [\delta_a \, \delta_e \, \delta_r \, \delta_t]^T$$



longitude (p_x), latitude (p_y) and altitude (p_z) for position; u, v and w for velocity; roll (γ), pitch (θ) and yaw (ψ) for attitude; ω_x, ω_y and ω_z for angular rate

aileron (δ_a), elevator (δ_e), rudder (δ_r) and throttle (δ_t).

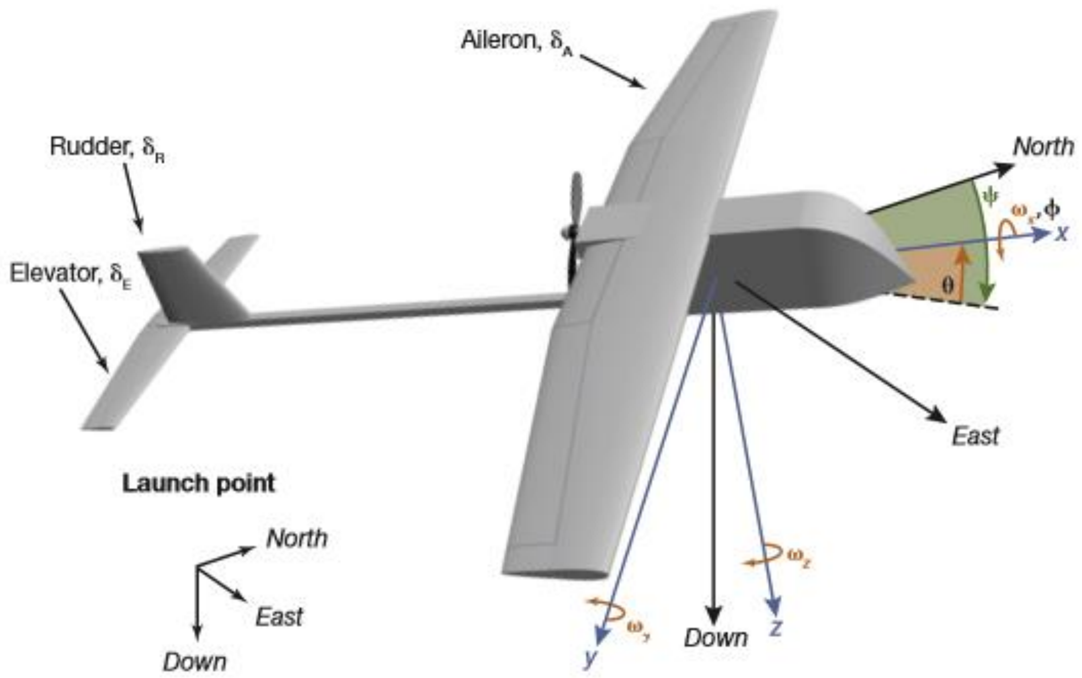
What are yaw,pitch, and roll?



$$\dot{x} = f(x, u)$$

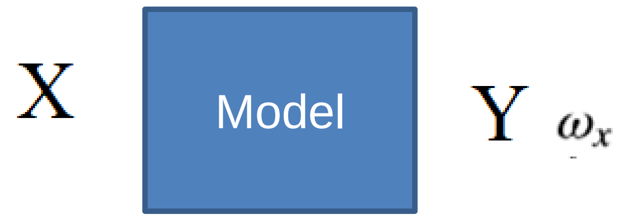
$$x = [p_x\ p_y\ p_z\ u\ v\ w\ \gamma\ \theta\ \psi\ \omega_x\ \omega_y\ \omega_z]^T$$

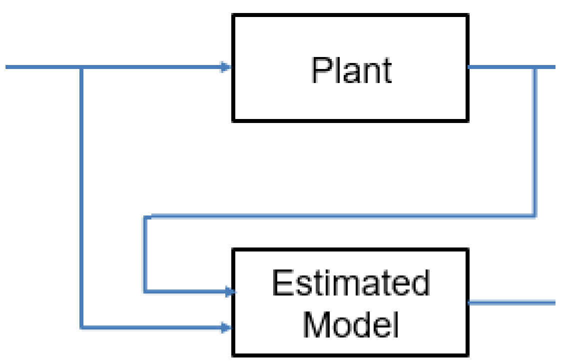
$$u = [\delta_a\ \delta_e\ \delta_r\ \delta_t]^T$$



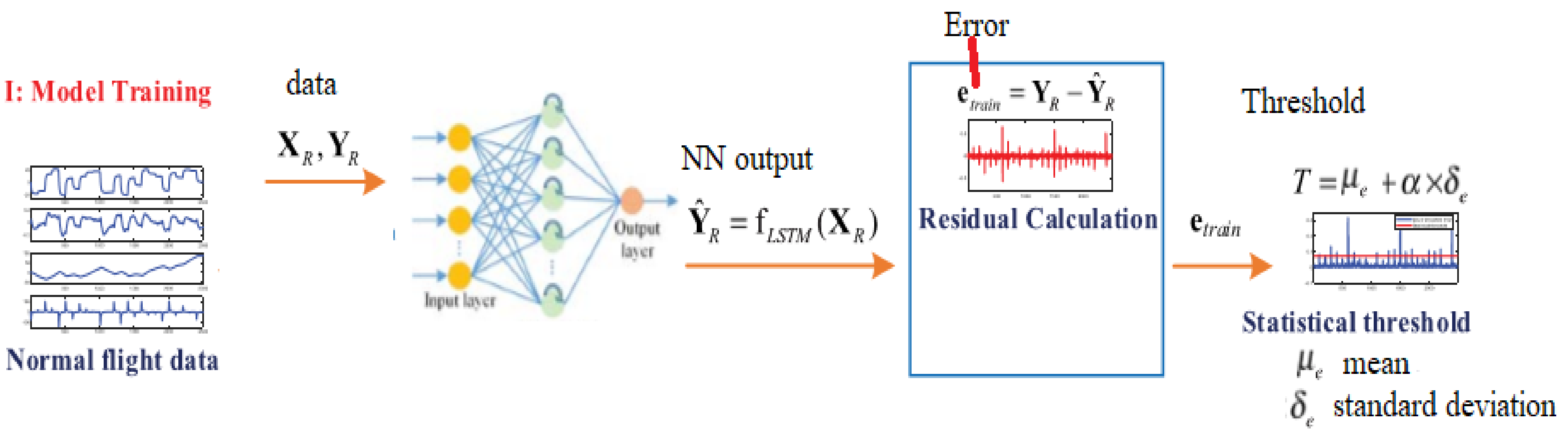
One issue is roll rate which is measured by gyroscope. For example, we want to detect if the roll rate has fault. A nonlinear model is built based on input/output data from healthy state.

Symbol	Parameter	Unit
γ	Roll angle	deg
θ	Pitch angle	deg
ψ	Yaw angle	deg
ω_x	Roll rate	deg/s
ω_y	Pitch rate	deg/s
ω_z	Yaw rate	deg/s
δ_a	Aileron deflection	deg
δ_e	Elevator deflection	deg
δ_r	Rudder deflection	deg
V	Indicated airspeed	m/s
h	Altitude	m

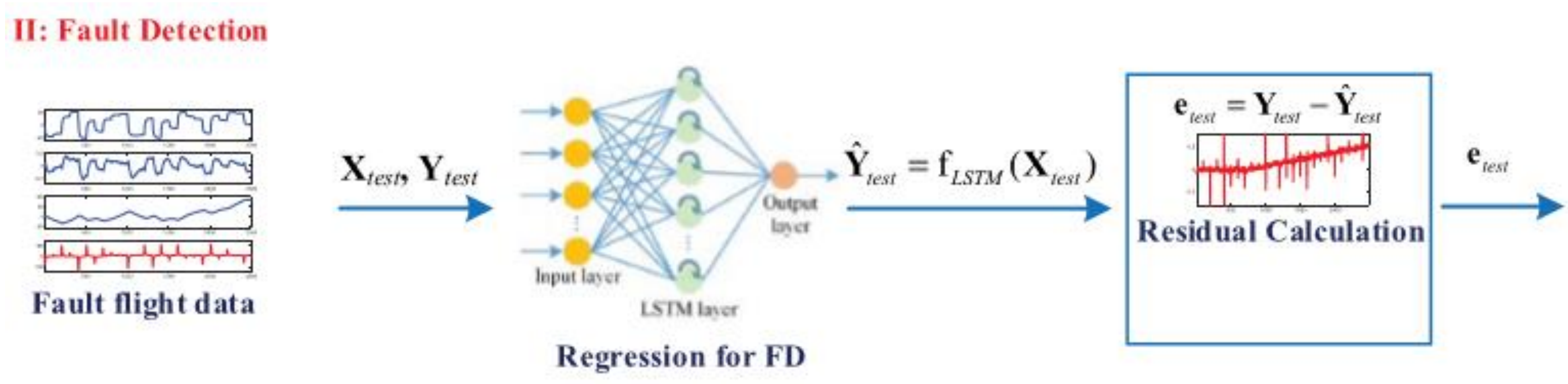




Neural network is used to train the model from the healthy data



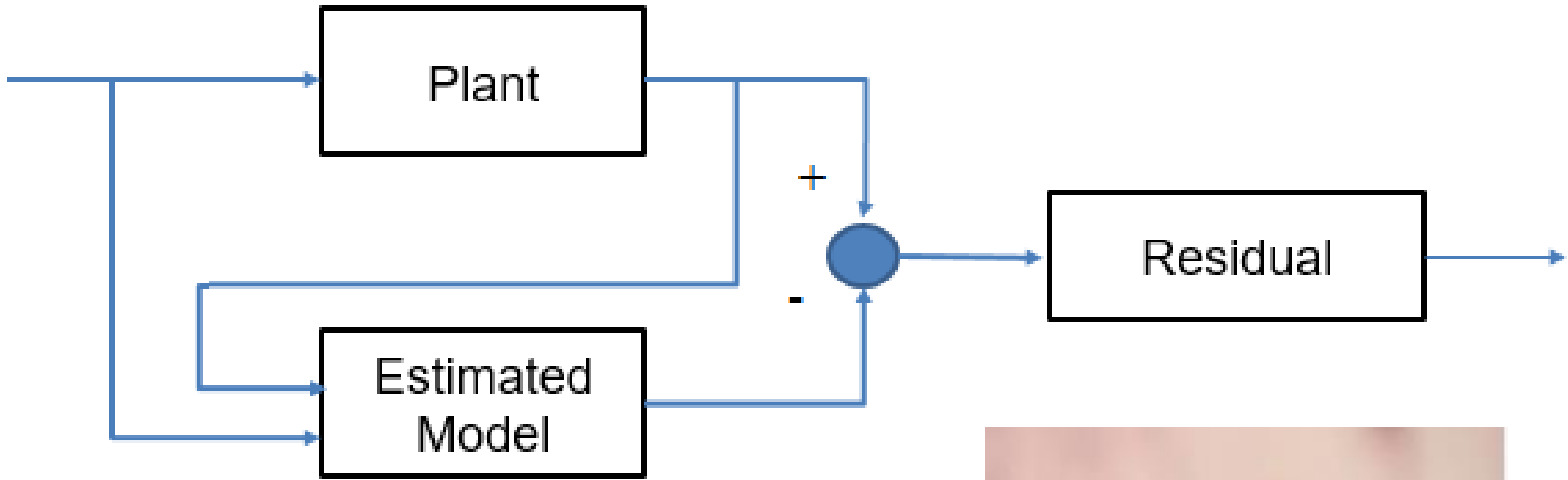
Fault detection



If $|e_{test}| > T, fault;$
$otherwise$ no fault detected.

Detection of hack attack

This method can be used in detecting the hack attack to your system, for example, spoofing unmanned aerial vehicle (UAV).



We use a linear observer

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t), \quad x(0) = x_0$$

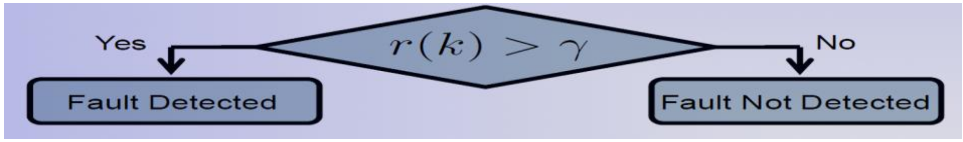
$$y(t) = Cx(t)$$

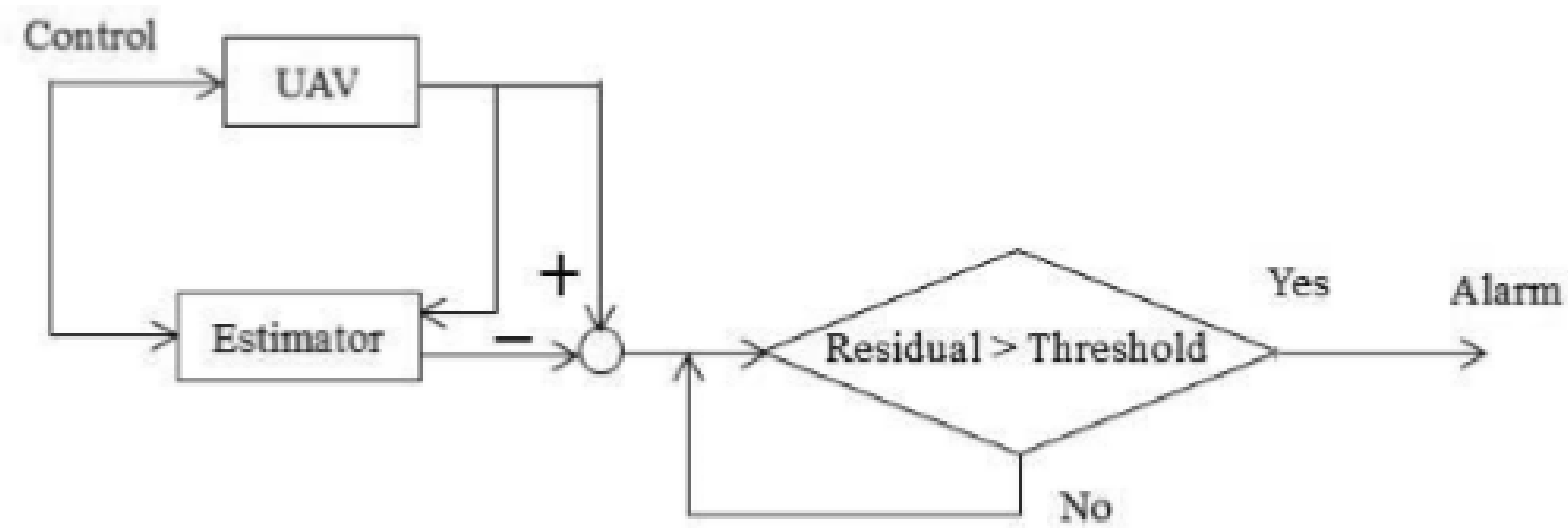
Observer

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(y(t) - \hat{y}(t))$$

$$\hat{y}(t) = C\hat{x}(t)$$

Residual $\quad \mathrm{r(t)} = \tilde{y} = y - \hat{y},$





$$\begin{cases} \|\tilde{z}\| \le \varpi(t), & \text{the UAV controlled is in healthy state} \\ \|\tilde{z}\| > \varpi(t), & \text{the UAV controlled is attacked.} \end{cases}$$

For the x-axis, the model is given by

$$X_1 = \begin{bmatrix} 0 & 0 \\ 1 & -75.92 \end{bmatrix} X_1 + \begin{bmatrix} 9.78 \\ -3.28 \end{bmatrix} u_1 + w_1$$

$$z1 = [0\ 1]X_1$$

$$\dot{\hat{X}}_1 = \begin{bmatrix} 0 & 0 \\ 1 & -75.92 \end{bmatrix} \hat{X}_1 + \begin{bmatrix} 9.78 \\ -3.28 \end{bmatrix} u_1 + L_1(z_1 - \hat{z}_1),$$

$$\hat{z}_1 = [0\ 1]X_1$$

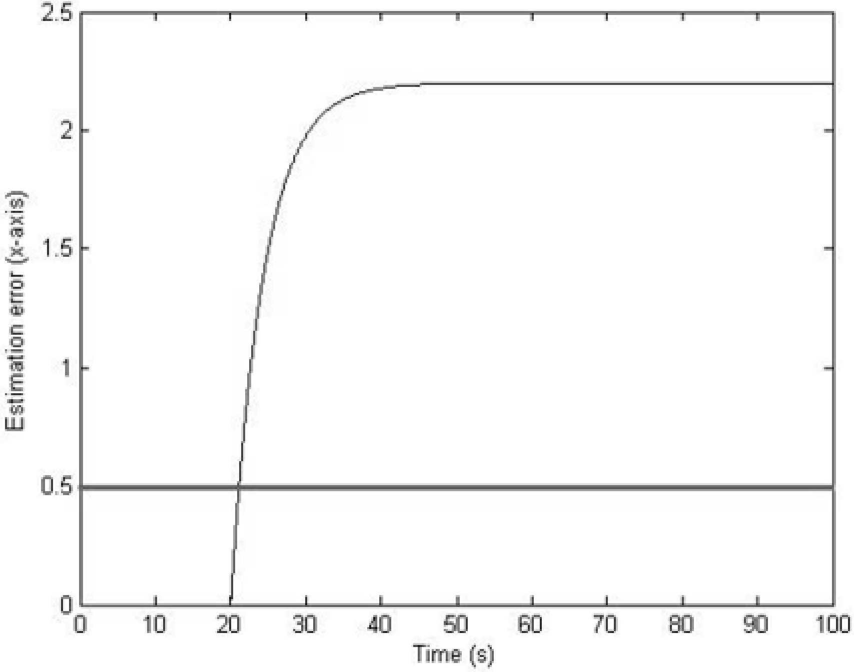
For the y-axis, the model is given by

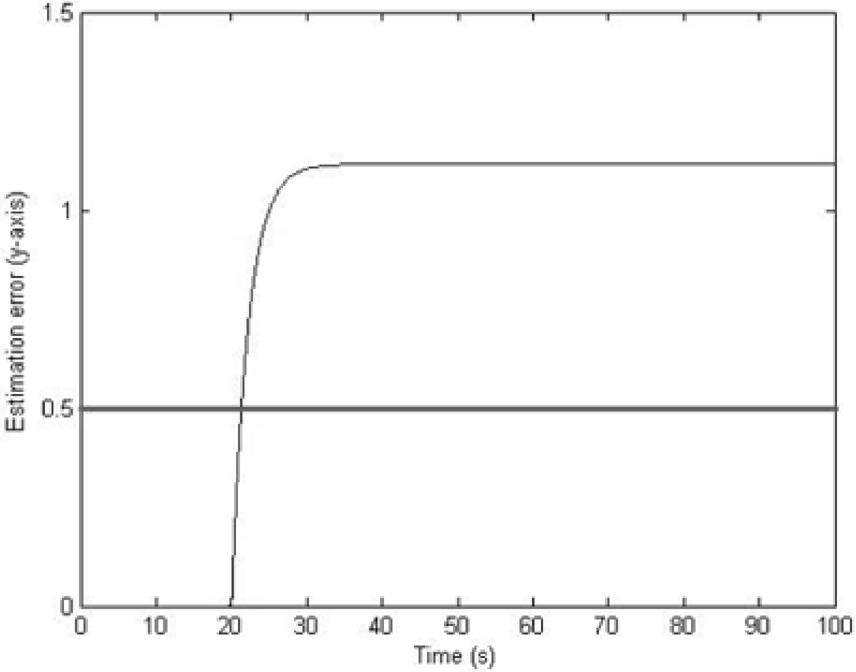
$$z_2 = G(s) = \frac{-276.4s + 743.5}{61s^2 + 4463s} u_2$$

$$\dot{\hat{X}}_2 = \begin{bmatrix} 0 & 0 \\ 1 & -73.16 \end{bmatrix} \hat{X}_2 + \begin{bmatrix} 12.18 \\ -4.53 \end{bmatrix} u_2 + L_2(z_2 - \hat{z}_2),$$

$$\hat{z}_2 = [0\ 1]X_2$$

Simulation result:





Design includes

- Estimator
- Residual
- Threshold
- Fault detection decision

Advantages of model-based fault detection

- It is software and easy to maintain
- The approach is standard one.

Disadvantages of model-based fault detection

- Model must be known
- Control must be known

Summary

There are two fault diagnosis methods

Hardware sensor-based scheme:

Time-domain: threshold decision and trend analysis.

Frequency-domain: FFT

Model-based schemes:

Estimator, residual , threshold, fault decision.

Reference

- J. Gertler(1998), Fault detection and diagnosis in Engineering Systems, Marcel Dekker, New York
- J. Chen and R.J. Patton (1999), Robust model-based fault diagnosis for dynamic systems, Kluwer Academic Publishers
- S.Huang,K.K.Tan,P.V.Er,T.H.Lee, Intelligent Fault Diagnosis and Accommodation Control, CRC Press,2020

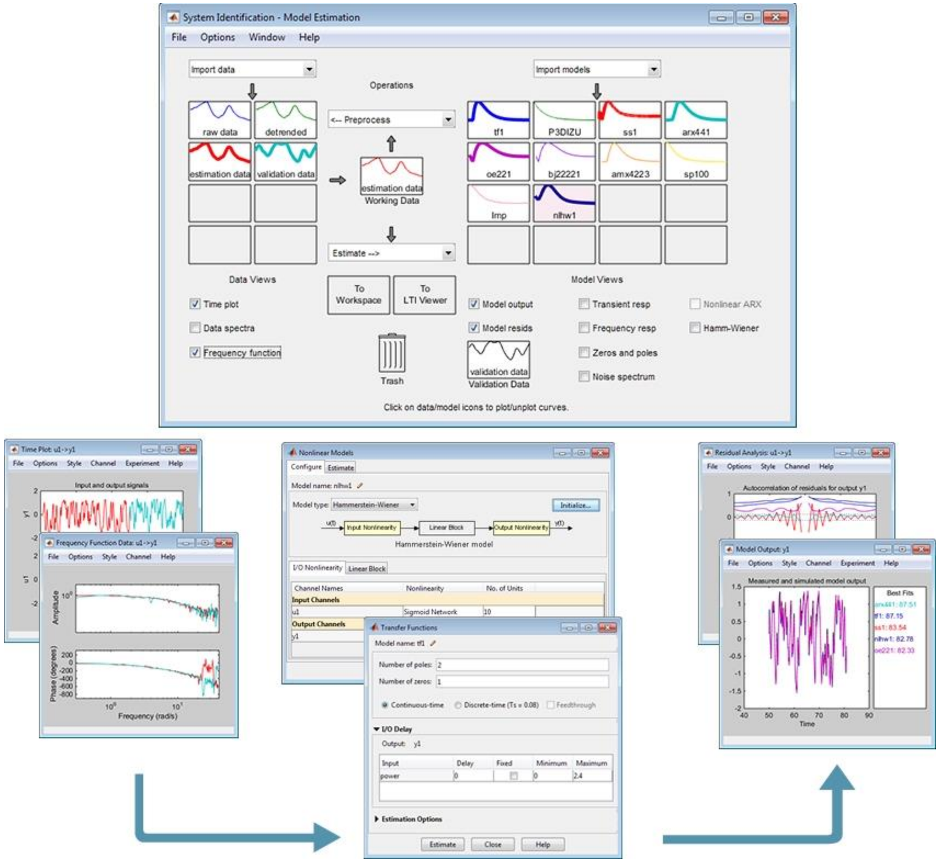
NUS National University of Singapore

Black-box model

- Step response
- Square wave

Use MATLAB toolbox:

System Identification Toolbox



Thank you!

Recess time (15 mins)