Big Data in Plantation Monitoring

Presented by

C.M. Toh and Prof. Ir. Dr. Ewe Hong Tat

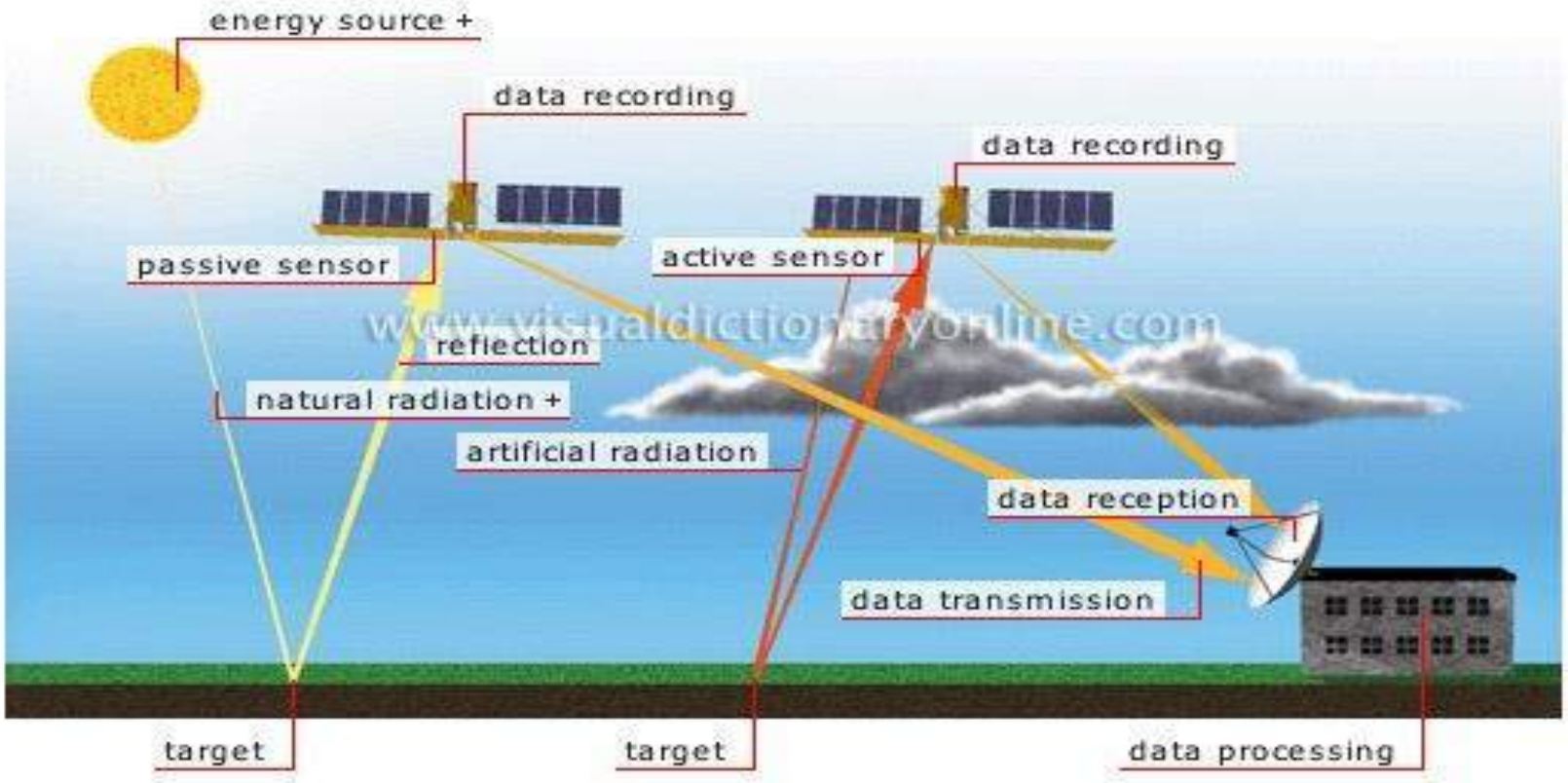
Universiti Tunku Abdul Rahman (UTAR)

In collaboration with MPOB and Applied Agricultural Resources Sdn Bhd (AARSB)

Remote Sensing of Environment

Active Sensing – Utilizes own illumination/radiation source and receives the reflected signals

Passive Sensing – Detects natural radiation that is emitted or reflected by the object or surrounding area being observed



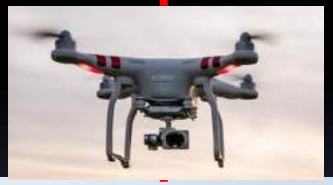
Remote Sensing:

- to derive information of the earth's surface from images acquired from a distance above the ground
- Different types of instrument: e.g.: Optical, Microwave and LiDAR

600 – 1000 km
Space

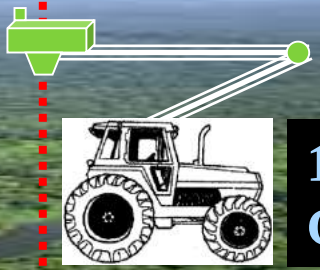
3 – 10 km
Sky

300 m – 3 km
Sky



100 – 300 m
Sky

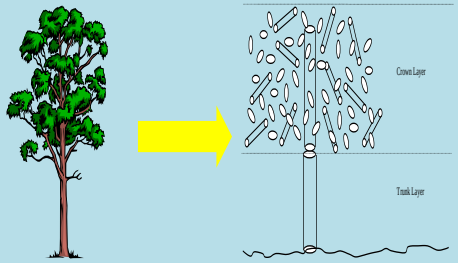
SHTey



1m – 5 m
Ground

Image Acquisition

Satellite Radar Returns and Image Processing



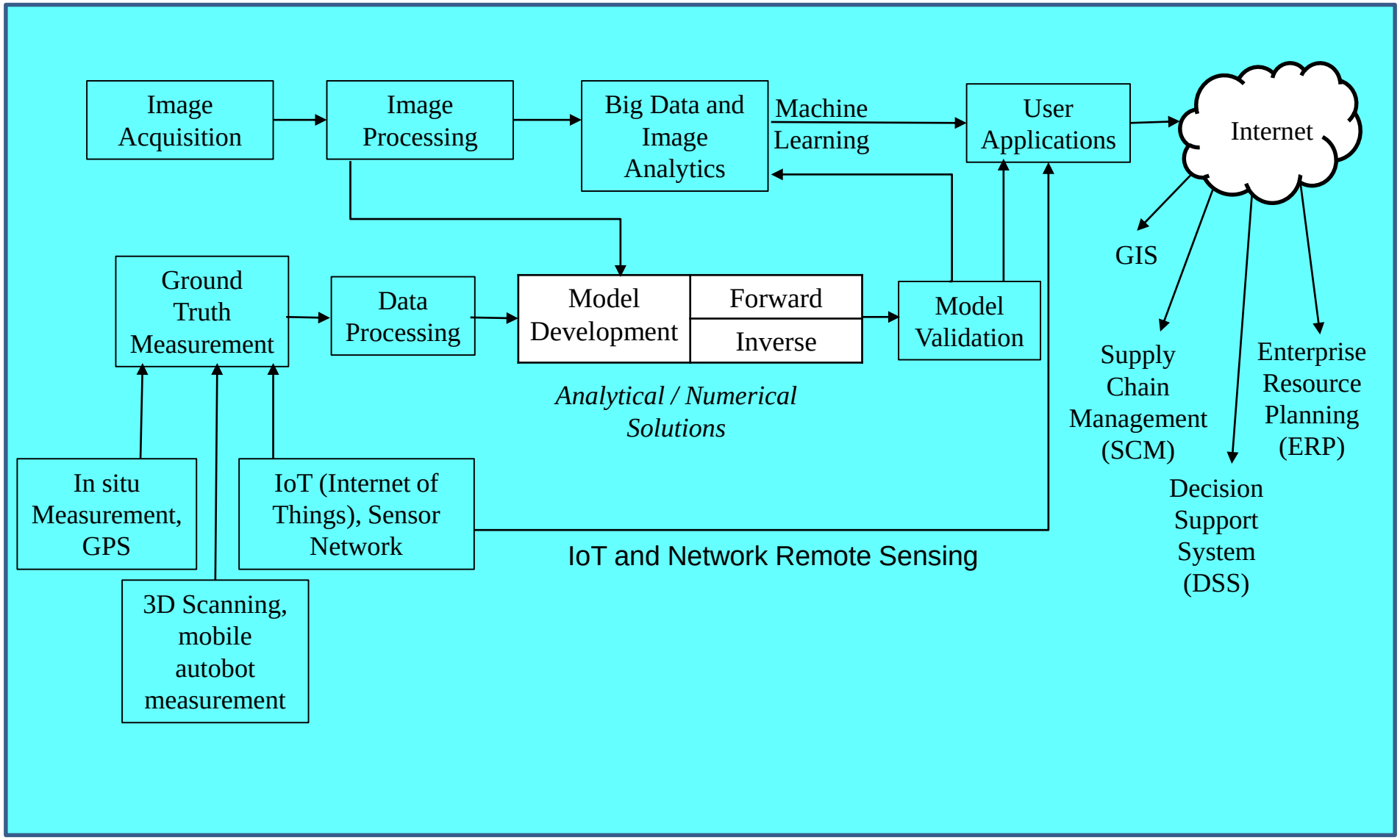
Theoretical Modelling

Ground Truth Measurement / Sensor Network and Intelligent Computing

Remote Sensing Applications

- Land Use Classification
- Paddy Yield Precition
- Sea Ice Thickness Retrieval
- Oil Palm Growth Monitoring

Long Term Remote Sensing Modelling Framework (Convergence of Technologies)



Connection to Remote Sensing

- **Big Data Analytics and Deep Learning**
  - **Intelligent analysis (AI)** of huge remote sensing data and images
  - **Machine Learning/ Deep Learning** based SAR image analysis
  - **Data mining** and **knowledge discovery**

- **IoT and Network Remote Sensing**
  - Sensing of environment anywhere and anytime through **IoT**
  - **Electromagnetic sensing + IoT & network sensing**

Technological Singularity in Year 2045

- Ray Kurzweil, Futurist and Google Engineering Director

Microwave Remote Sensing

- **Benefits**
  - Global and wide coverage
  - All time coverage (day and night) and penetration through cloud and rain
  - Additional information (penetration below surface, wide range of frequency, multi-polarization)
- **Problems/Challenges**
  - How to know what contributes to the received signal?
  - How to interpret the microwave images?
  - How to utilize the information for real applications?

Theoretical Models for Vegetation

- Developed the extension of dense medium effect to vegetation medium (RT-PACT for vegetation)
- Dense Medium that considers
  - Near field interaction effect
  - Coherent effect

- Suitable for multiple configurations
  - Temperate Coniferous Forest
  - Tropical Deciduous Forest
  - Tropical Evergreen Forest
  - Paddy Field, Oil Palm Plantation

- Study of radar returns for various frequencies, observation angles, types of vegetation, growth stages

Role in Big Data

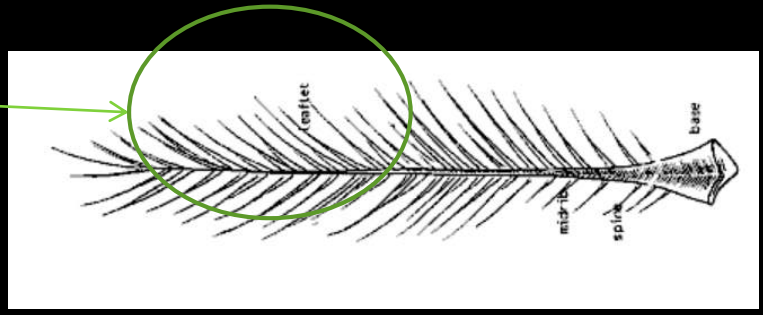
- Covers huge area; good amount of information per scene.
- Fusion possible between optical (multispectral/hyperspectral) + Microwave.
- We can use the information to understand more on how oil palm interacts with microwave.
- With analysis, we can classify and monitor oil palm condition via these images.

Potential of Plantation Monitoring Using L Band SAR Image

Oil Palm Gallery

Oil Palm Anatomy

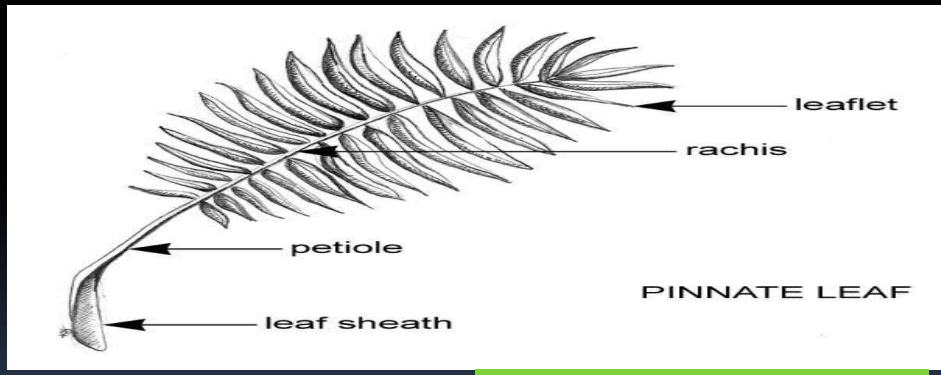
- Pinnae
- The leaflets of oil palm.
- Leaf Area Index (LAI)



- Frond
- Considered as the

"branch" of oil palm"

Used for biomass estimation

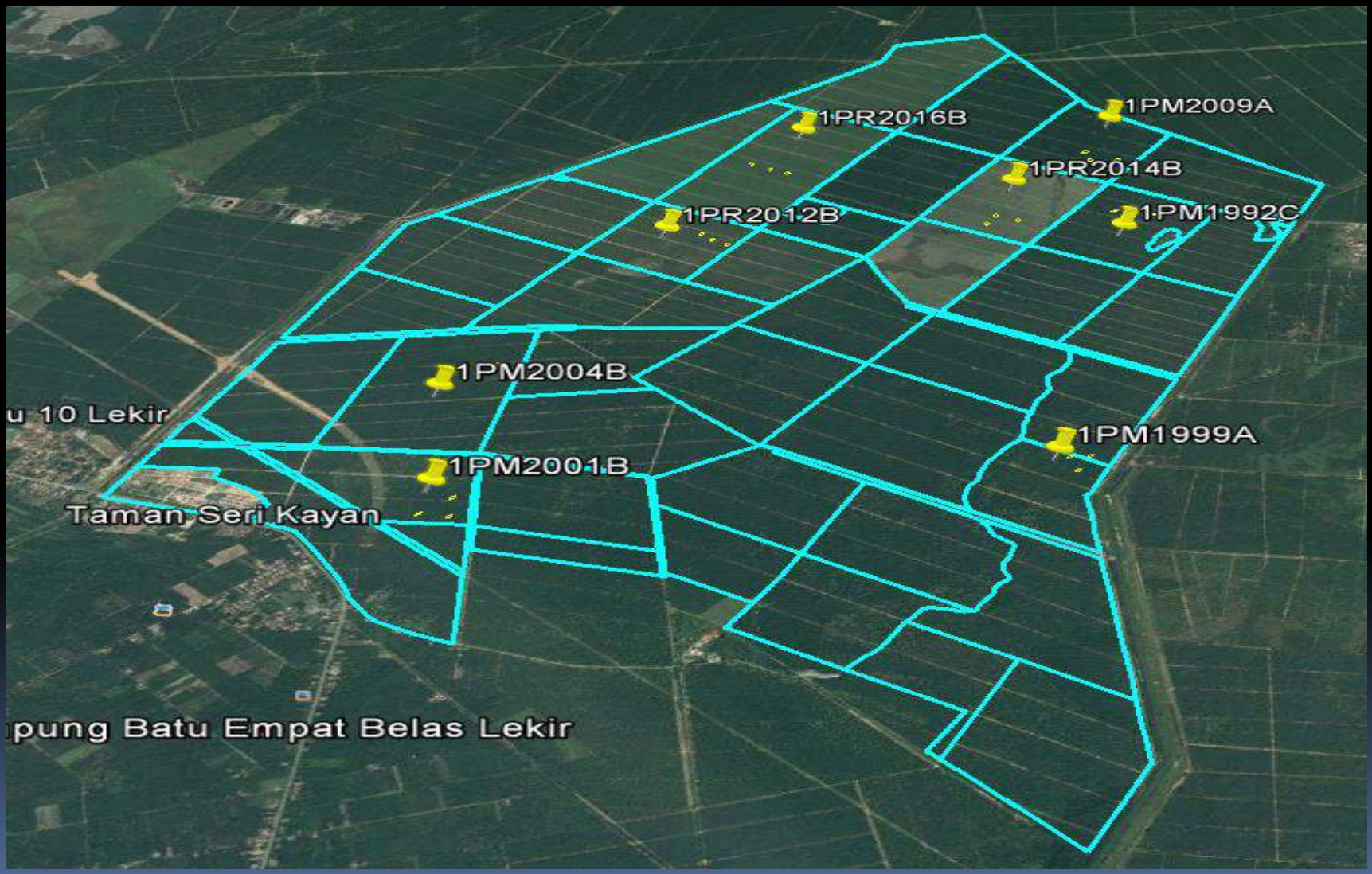


Frond of the oil palm

Images taken from http://www.fao.org/docrep/t0681E/t0681e02.htm

A Microwave Satellite Image (HH Polarization) at Setiawan/Manjung

Study Site on Google Earth



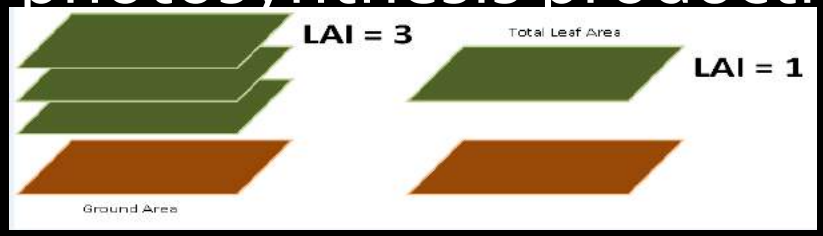
Derivative Parameters Calculated from collected data.

- LAI.
- Frond Cross Section Area.

Oil Palm Leaf Area Index (LAI)

- Directly translate to canopy density.
- Prediction of photosynthesis production.

$$LAI = \frac{Total\ Leaf\ Area}{Projected\ Ground\ Area}$$



- Oil palm optimum LAI is around 5.5 to 6.5 for maximizing fruit production.
- Too high of LAI has a lot of unused leaves (pinnae) for photosynthesis.

LAI Calculation

- LAI Equation is as following:

$$LAI = 0.55FN\left(\frac{L}{100}\right)\left(\frac{W}{100}\right)\left(\frac{PPH}{10000}\right)$$

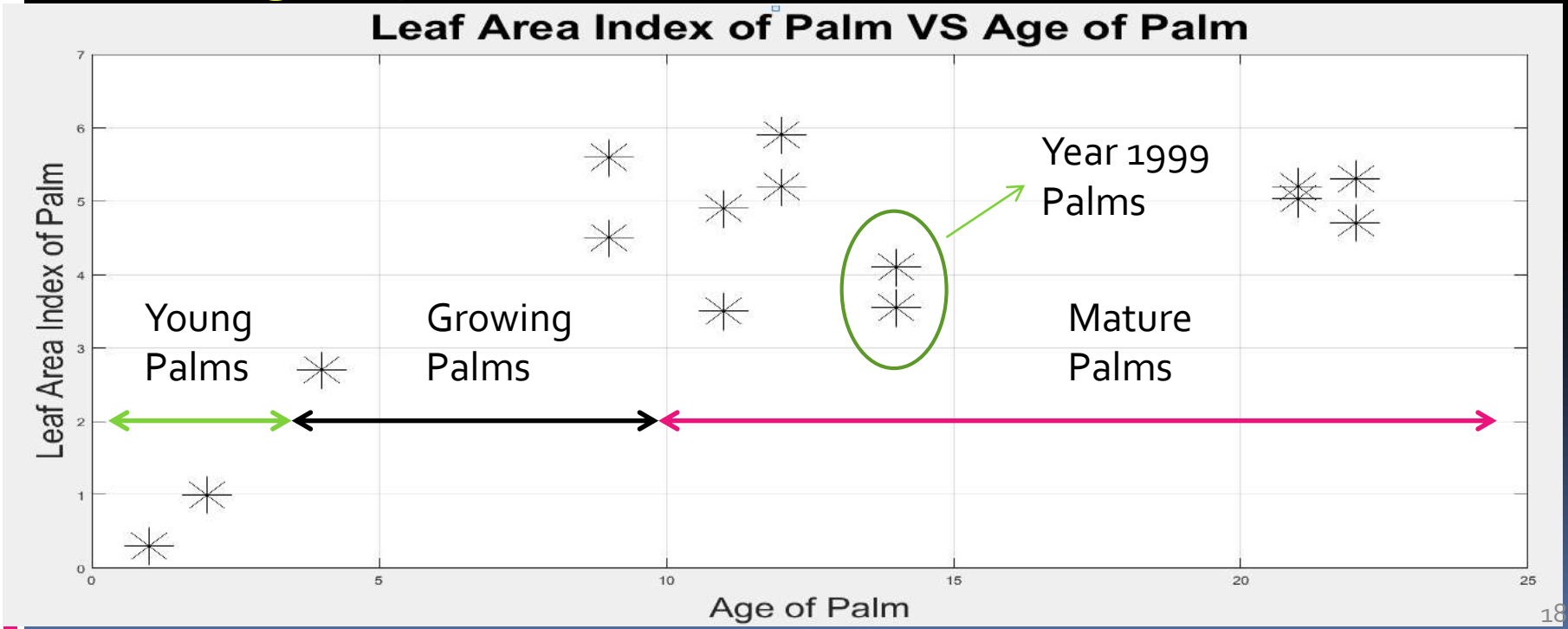
- Where:

1) L is the length of pinnae.
2) W is the width of pinnae.
3) N is number of pinnae per frond.
4) F is the number of frond per palm.
5) PPH is the palms per hector for the block.
6) 0.55 is the correction factor

* Frond 17th was chosen as the sample for calculation for a particular plot of oil palms.

LAI Data versus Palm Age

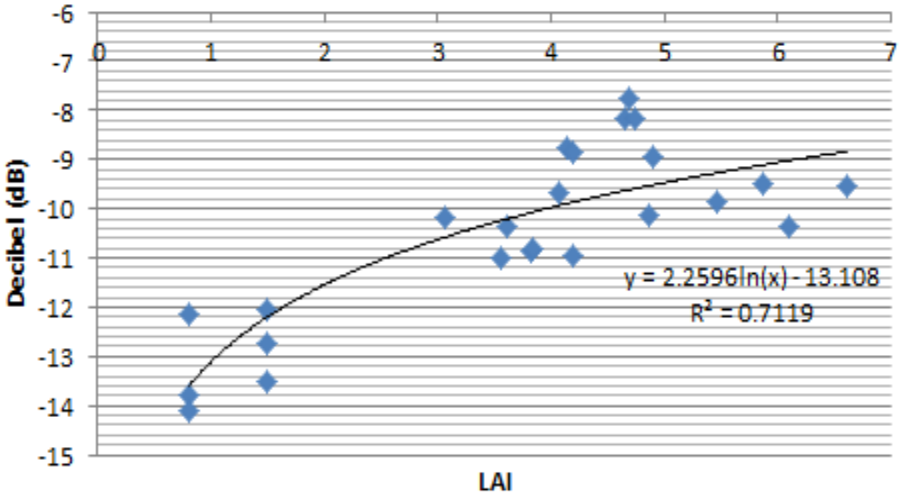
- LAI calculated from collected data and census data.
- 3 groups of interest.



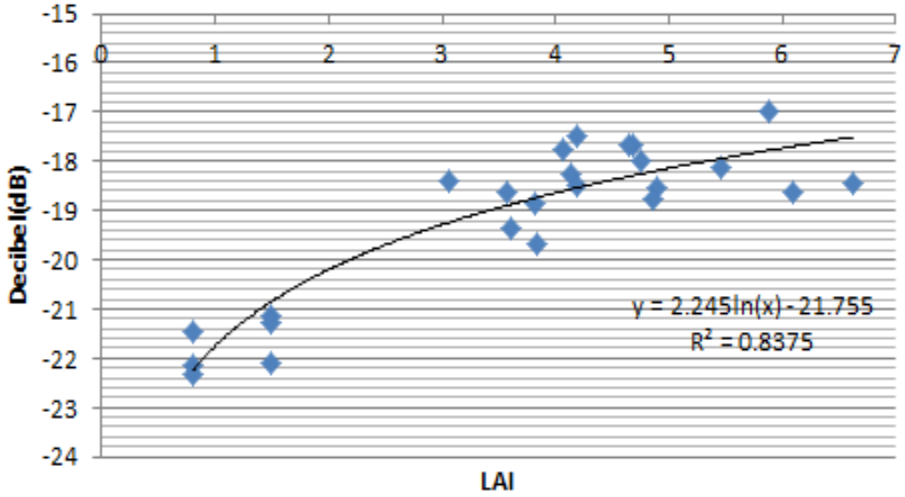
Observation

- 3 stages of oil palm growth. Young, Growing and Mature.
- Discrepancy on 1999 palms due to lower LAI value to others.
- Found that 1999 palms were also shorter in height.
- Verification with AARSB is that this block of palms are under attack by Basal Stem Rot (Ganoderma)

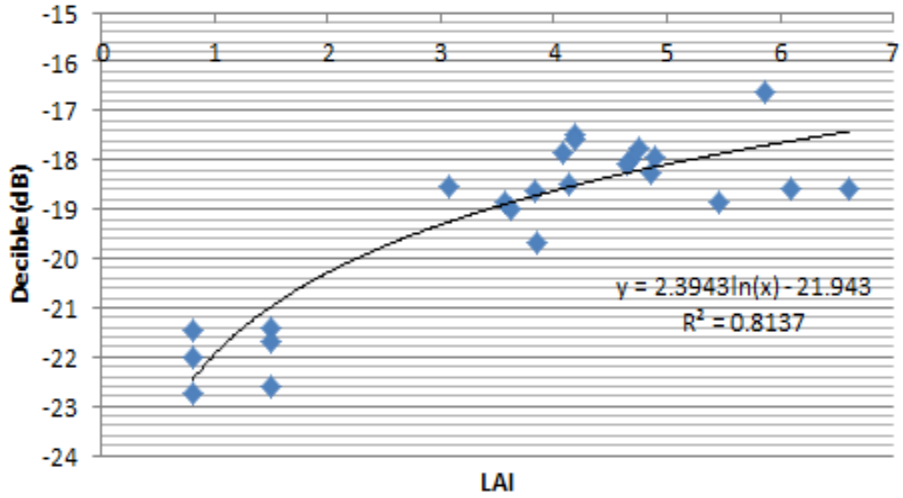
Backscatter Coefficient VS LAI (HH)



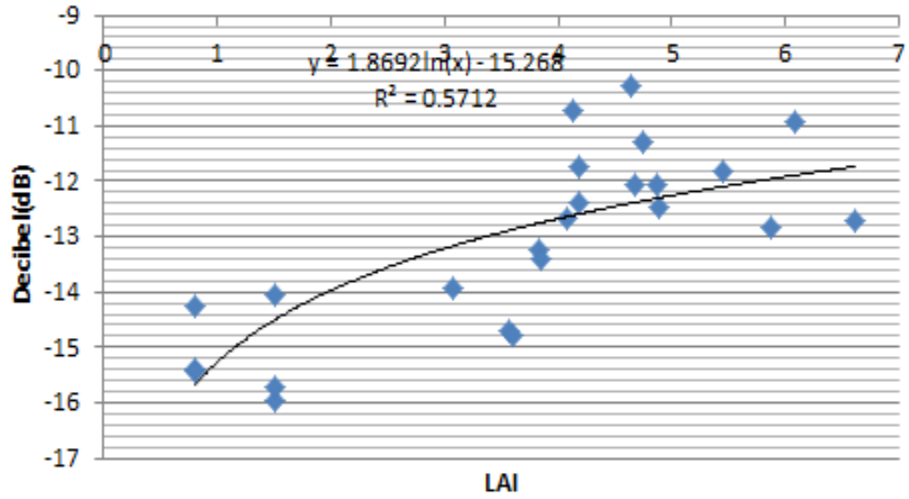
Backscatter Coefficient VS LAI (VH)



Backscatter Coefficient VS LAI (HV)

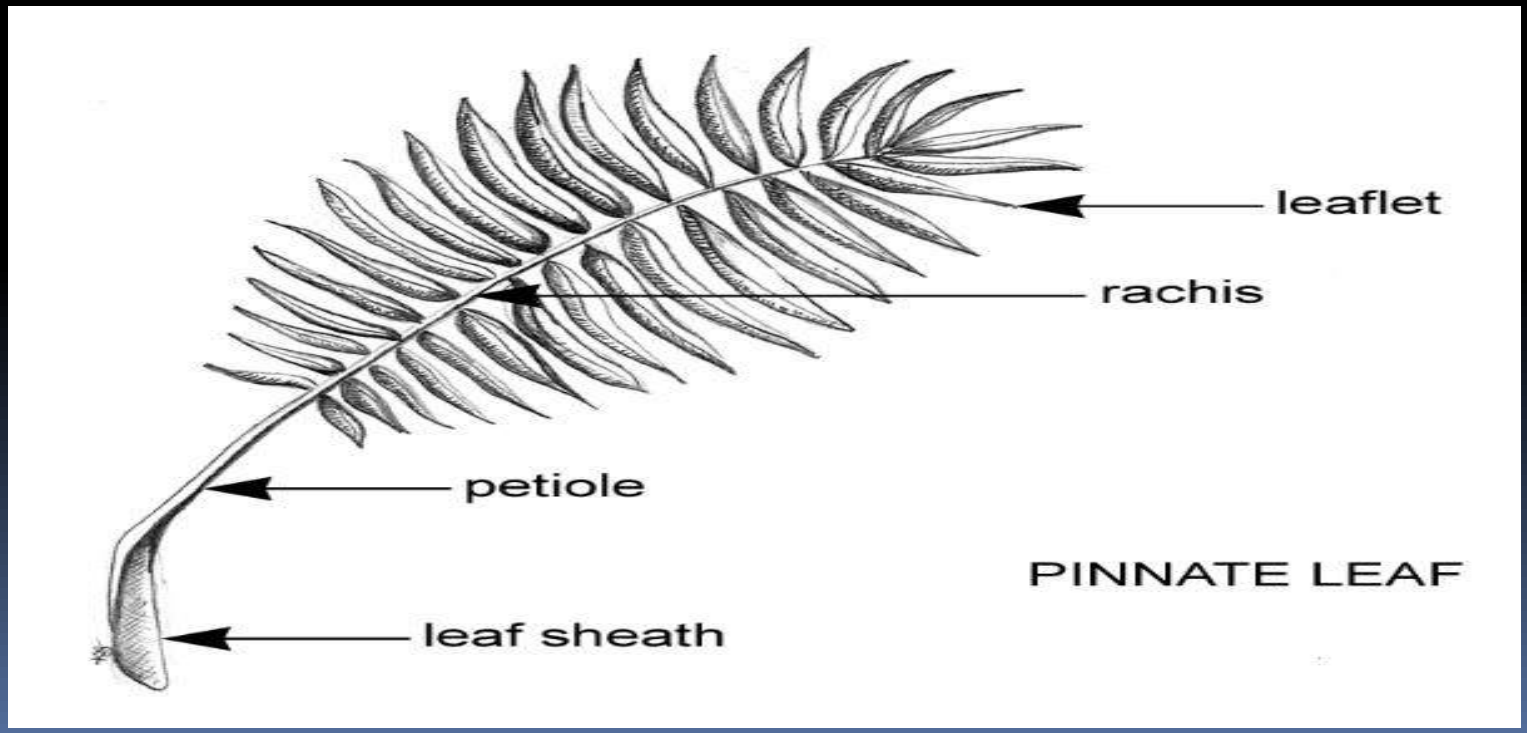


Backscatter Coefficient VS LAI (VV)

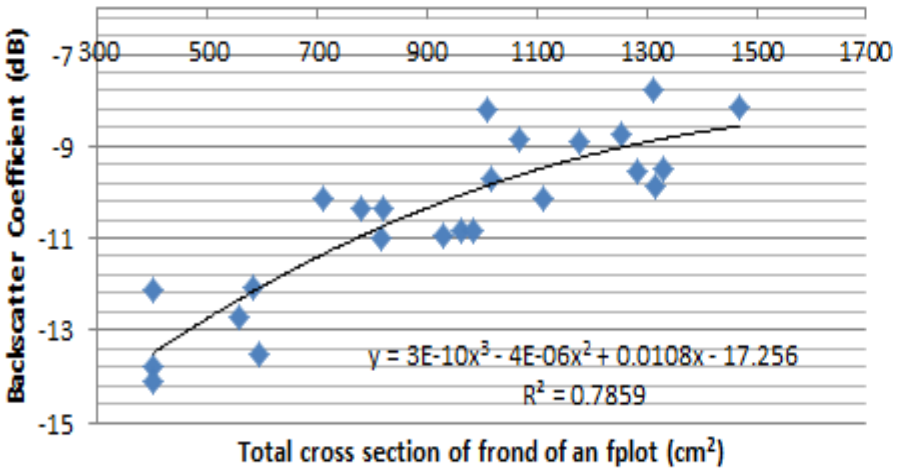


Frond Cross Section Area

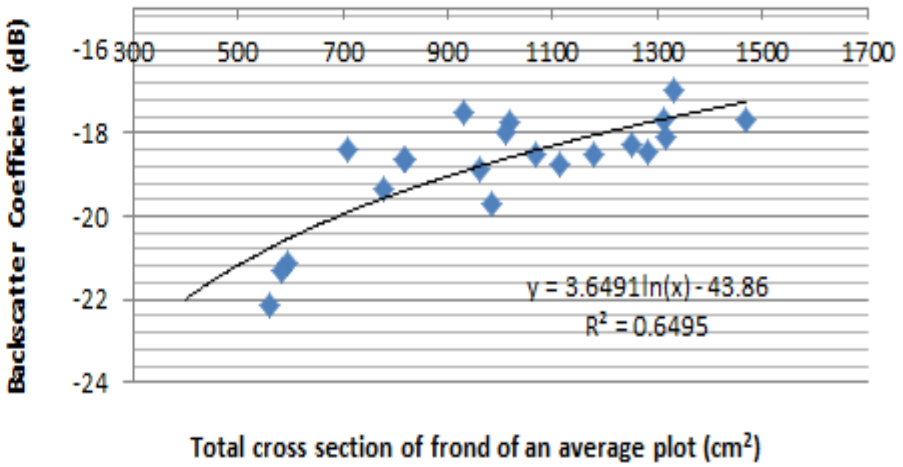
- Used to estimate the above ground biomass of oil palm canopy by Corley et al..



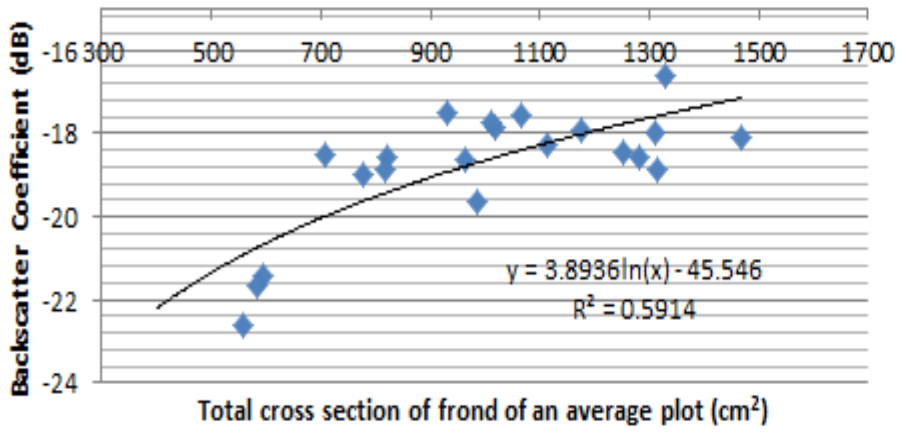
Backscatter Coefficient VS Frond Cross Section Area(HH)



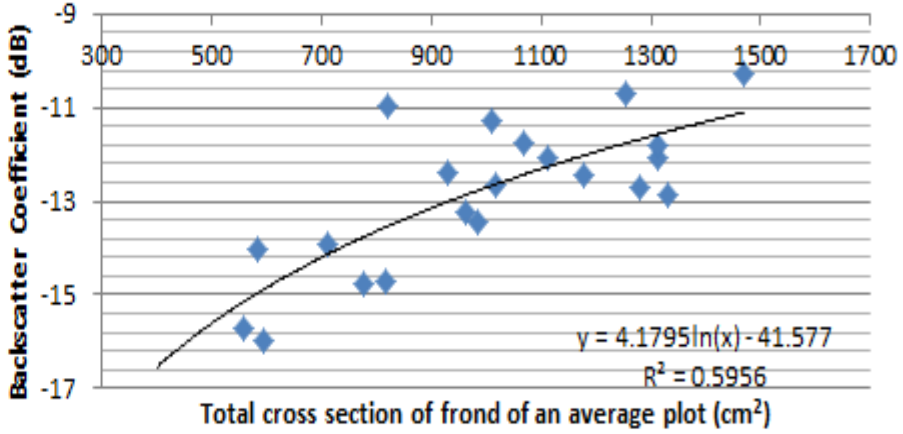
Backscatter Coefficient VS Frond Cross Section Area(VH)



Backscatter Coefficient VS Frond Cross Section Area(HV)

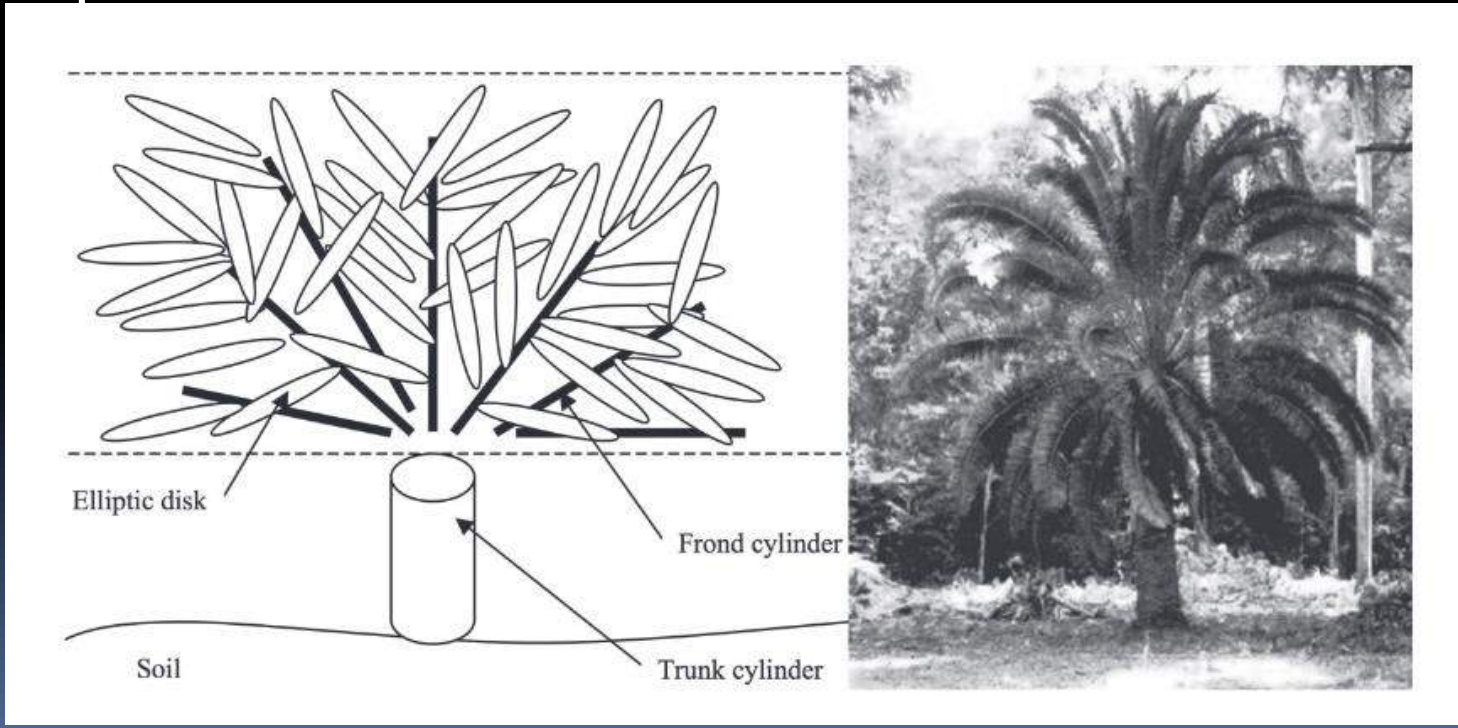


Backscatter Coefficient VS Frond Cross Section Area(VV)

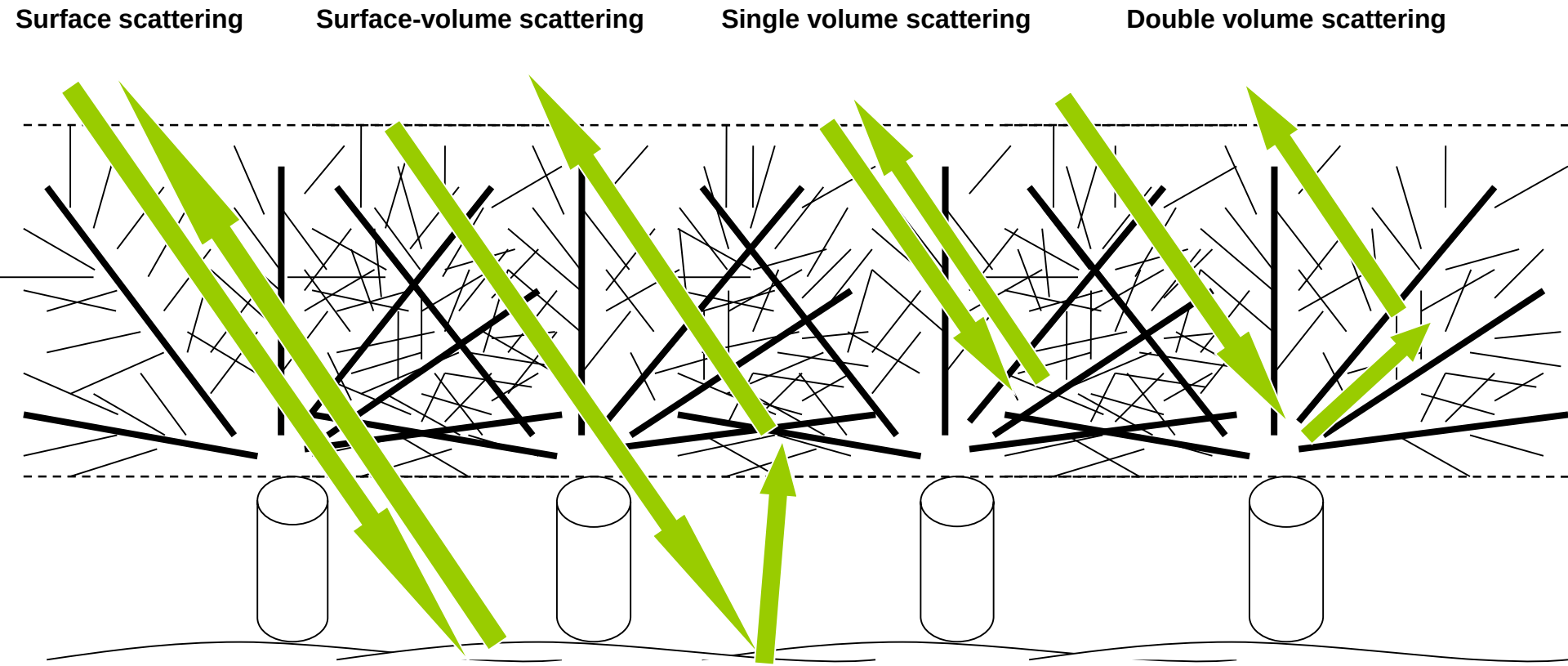


Theoretical Model

- Based on work done by Teng et al. and Koay et al. with principle from Radiative Transfer equation.



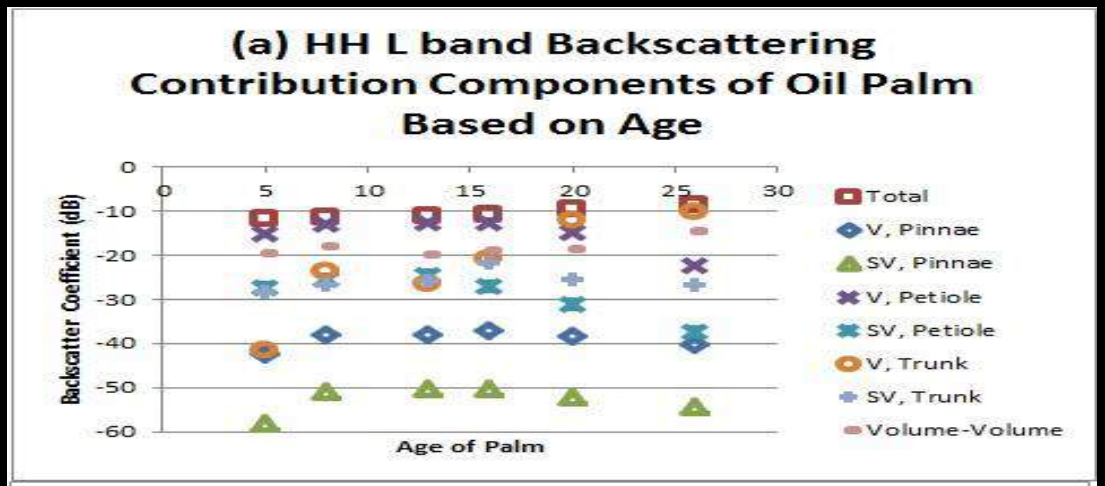
DEVELOPMENT OF MICROWAVE SCATTERING MODEL

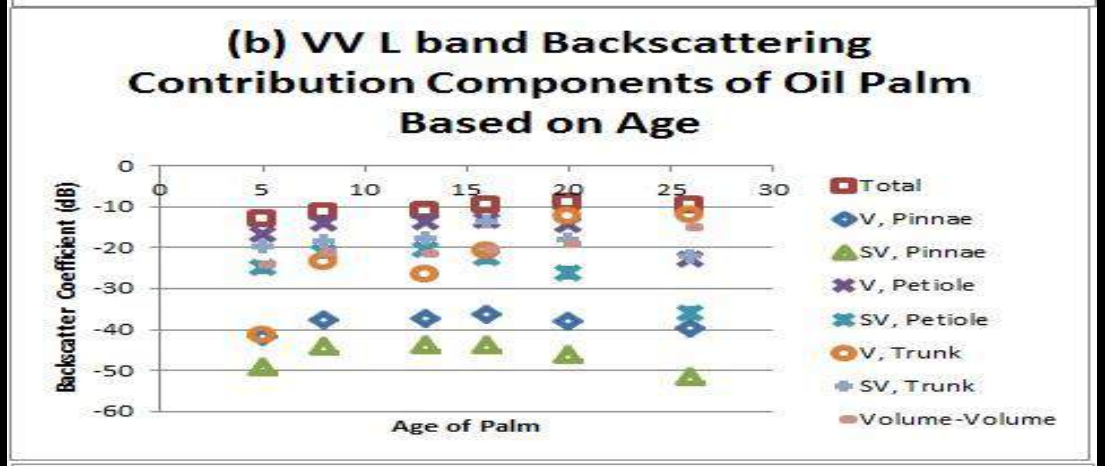


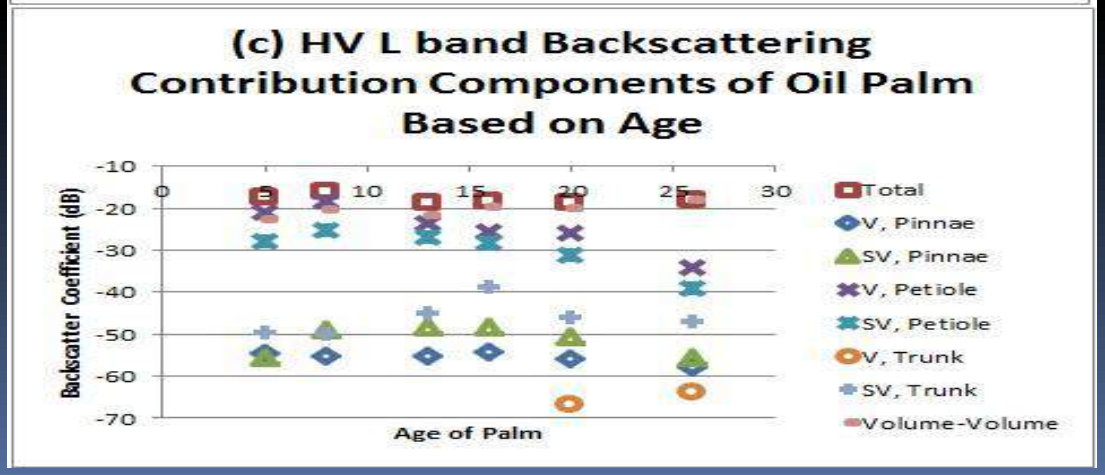
Backscattering components based on the iterative solutions of the radiative transfer equations

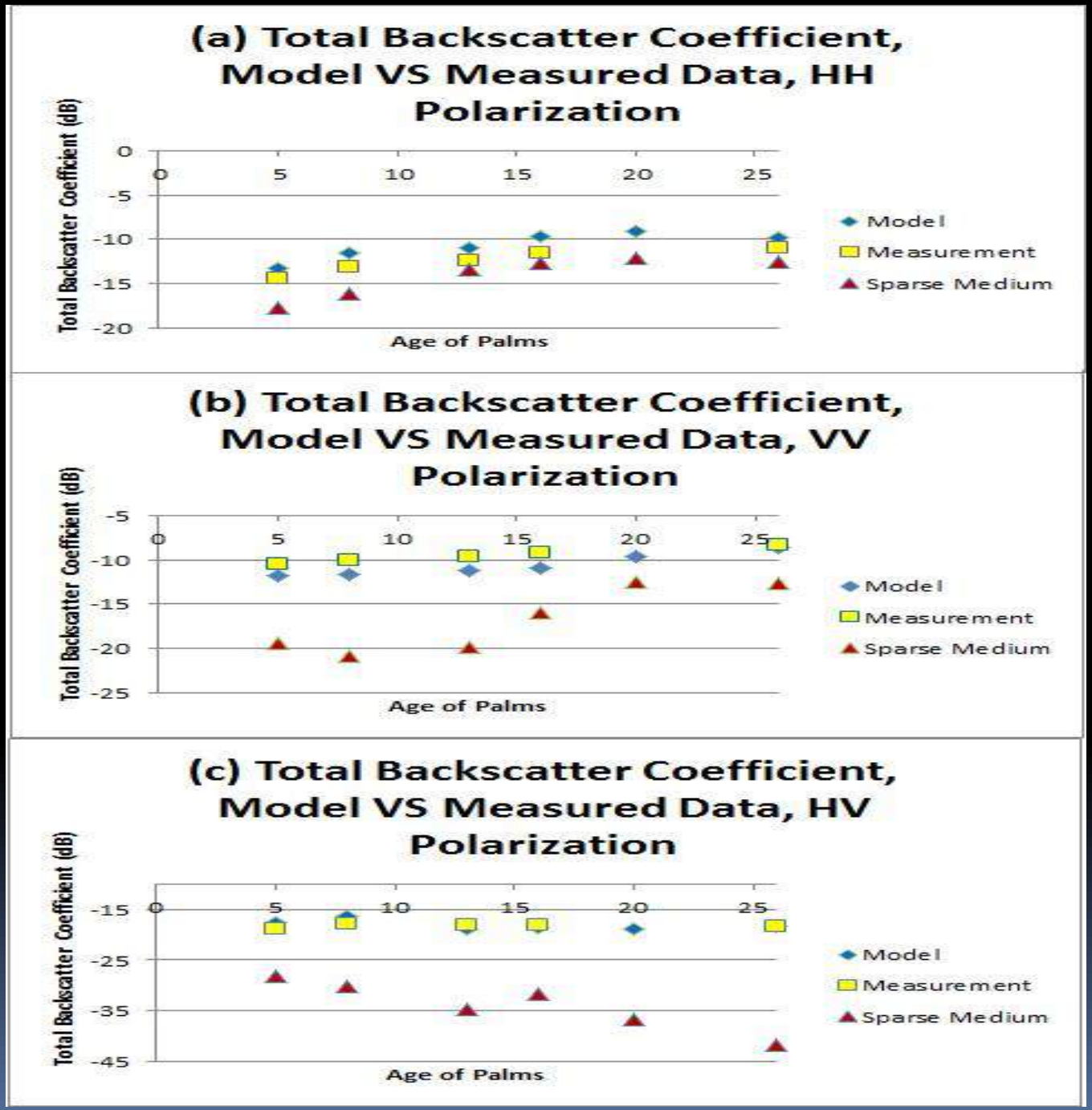
Purpose of Modeling

- Simulation of oil palm backscatter and compared to satellite data.
- Breakdown of backscatter contribution from structure of oil palm.
- Provides better understanding of results
- Total Backscatter Contribution and scattering mechanisms


(a) HH L band Backscattering Contribution Components of Oil Palm Based on Age
Backscatter Coefficient (dB)
0
-10
-20
-30
-40
-50
-60
0
5
10
15
20
25
30
Age of Palm
Total
V, Pinnae
SV, Pinnae
V, Petiole
SV, Petiole
V, Trunk
SV, Trunk
Volume-Volume


(b) VV L band Backscattering Contribution Components of Oil Palm Based on Age
Backscatter Coefficient (dB)
0
-10
-20
-30
-40
-50
-60
0
5
10
15
20
25
30
Age of Palm
Total
V, Pinnae
SV, Pinnae
V, Petiole
SV, Petiole
V, Trunk
SV, Trunk
Volume-Volume


(c) HV L band Backscattering Contribution Components of Oil Palm Based on Age
Backscatter Coefficient (dB)
-10
-20
-30
-40
-50
-60
-70
0
10
15
20
25
30
Age of Palm
Total
V, Pinnae
SV, Pinnae
V, Petiole
SV, Petiole
V, Trunk
SV, Trunk
Volume-Volume


(a) Total Backscatter Coefficient, Model VS Measured Data, HH Polarization
Total Backscatter Coefficient (dB)
0
-5
-10
-15
-20
0
5
10
15
20
25
Age of Palms
Model
Measurement
Sparse Medium
(b) Total Backscatter Coefficient, Model VS Measured Data, VV Polarization
Total Backscatter Coefficient (dB)
-5
-10
-15
-20
-25
0
5
10
15
20
25
Age of Palms
Model
Measurement
Sparse Medium
(c) Total Backscatter Coefficient, Model VS Measured Data, HV Polarization
Total Backscatter Coefficient (dB)
-15
-25
-35
-45
0
5
10
15
20
25
Age of Palms
Model
Measurement
Sparse Medium

Study on L Band SAR Image and Basal Stem Rot (Garnoderma)

- Disease affects yield of oil palm, shorten the economic lifespan, incur losses – Basal Stem Rot (BSR).
- Early detection is key, allows countermeasures and treatment by using of SAR remote sensing.
- Here we attempt to study how BSR affects microwave backscattering coefficient in L band SAR.

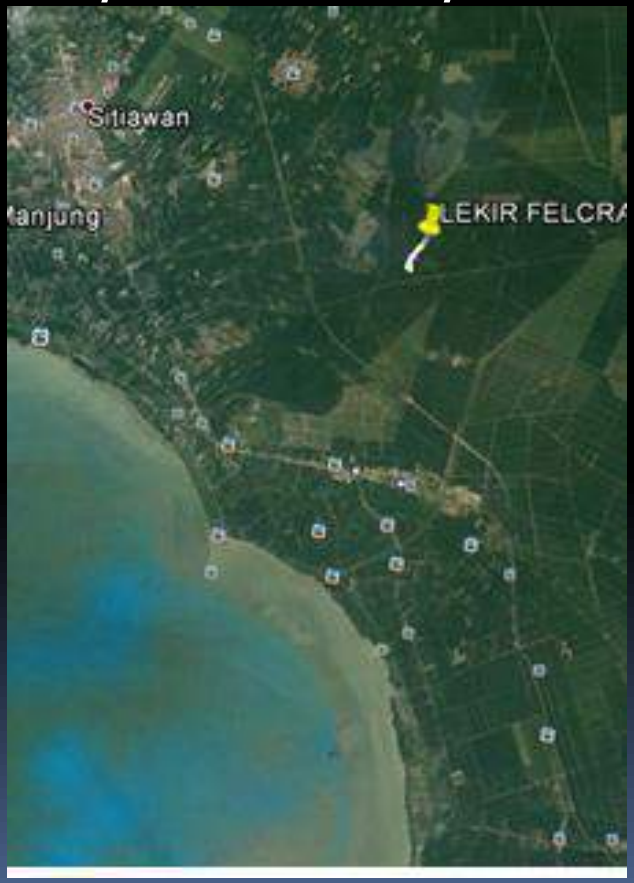
Basal Stem Rot

- Caused by *Ganoderma* fungus.
- Disease kills 80% of the stands when they are only halfway through their economic lifespan.

Garnoderma fungus on base of oil palm trunk.

Site Study Location

- FELCRA Lekir, Setiawan, Perak, Malaysia



Site Description

- Relative flat terrain.
- GPS Coordinates: 4.201483333N, 100.78445E.
- Has identified *Garnoderma* infested sites.
- Easier access due to near proximity with Setiawan.

Disease Severity Index

- An indicator on how severe the oil palm *Garnoderma* infection is. [3] Data collection based on the table, To to T2.

Tier	Symptoms/Observations
To	Healthy palms, absence of white button (Ganoderma fungus) at the stem base.
T1	Shrinking fronds, white button present at stem base.
T2	Rotting trunk body, depressed foliage, and white button present at stem base.

T0 Oil Palm (8 year old)

T1 Oil Palm (16 year old)

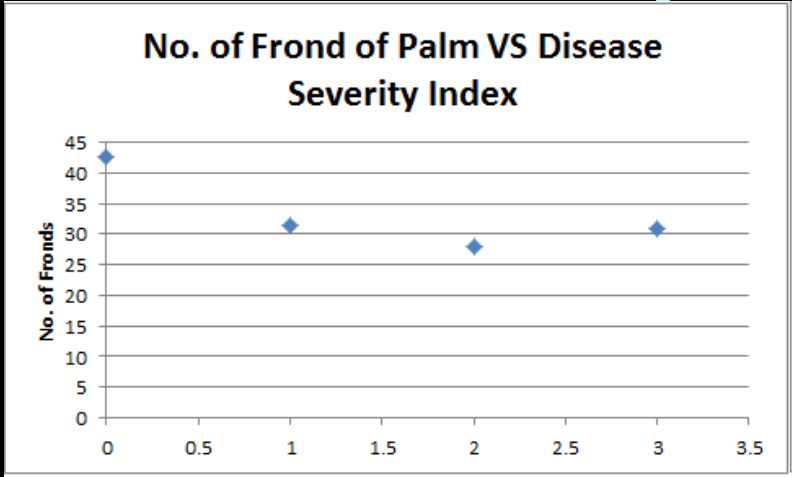
T2 Oil Palm (16 year old)

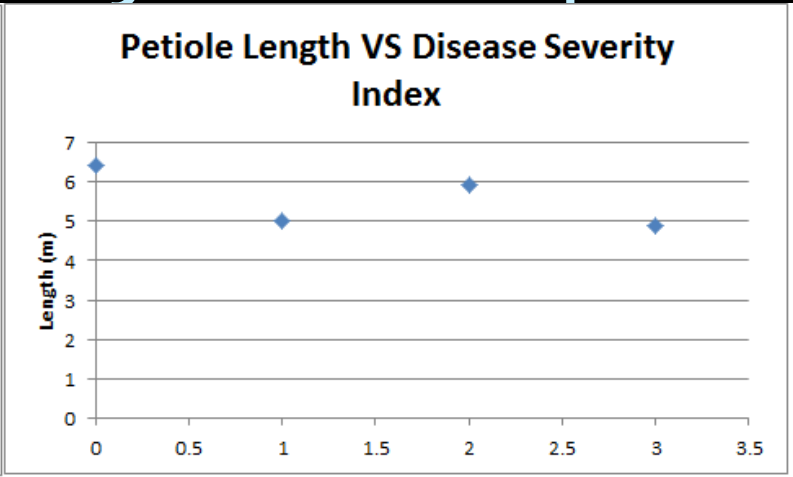
Fallen palm due to *Garnoderma* (T3, dead palm)

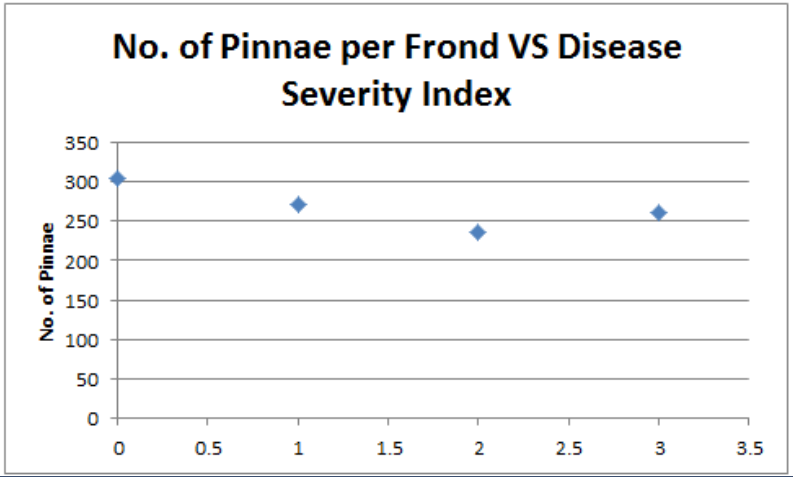
Oil Palm Data Collection

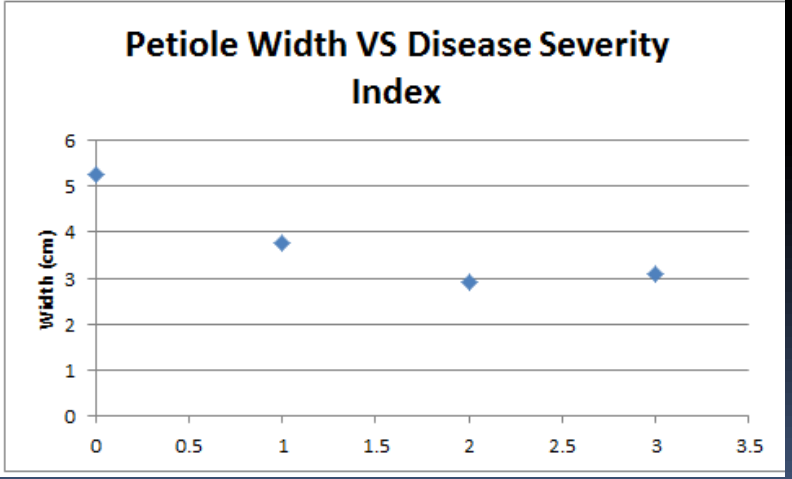
- Palms of 8 years old and 16 years old were studied in this trip.
- Difficulty in obtaining samples due to requirement – palms infected purely by *Garnoderma*, untreated palm.

Visible Changes of parameters to Ganoderma, 16 year old palms

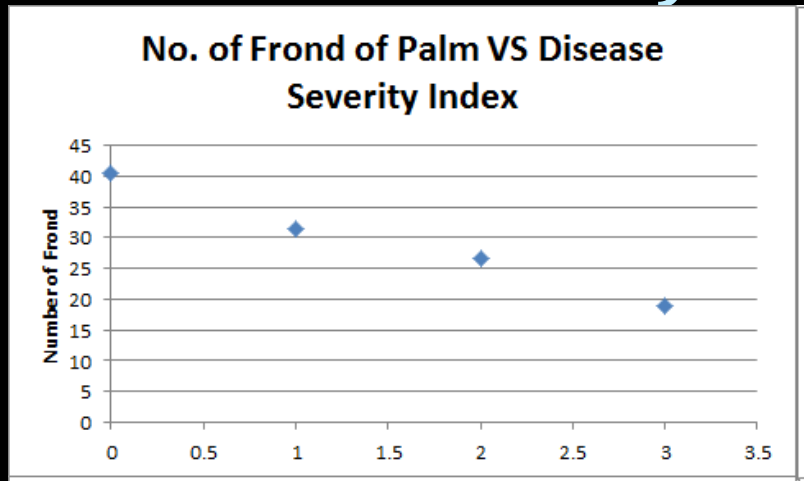


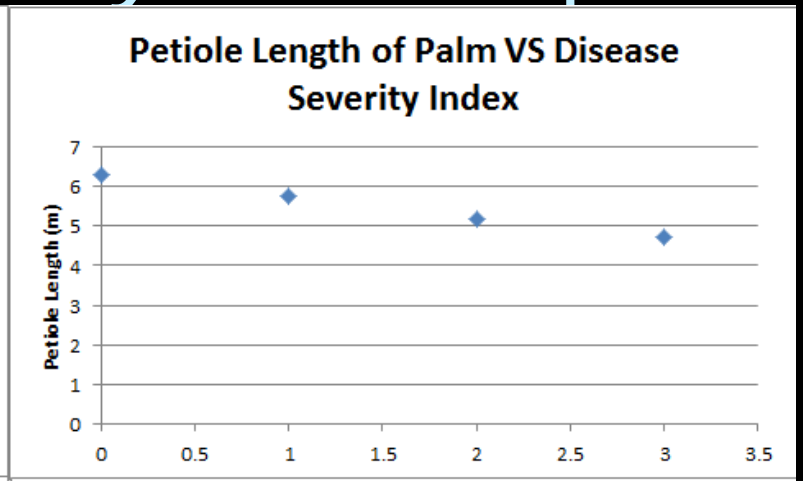


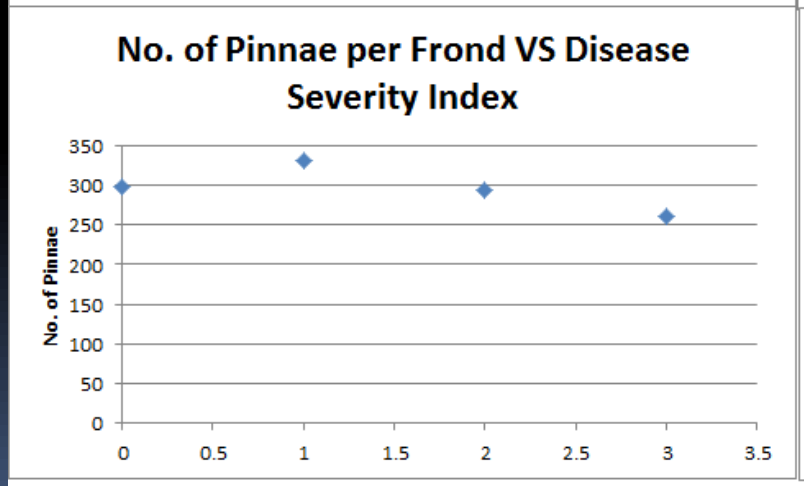


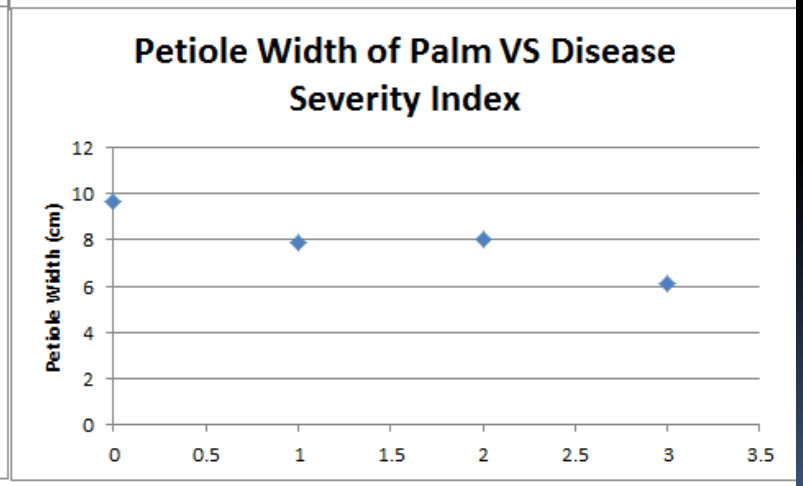


Visible Changes of parameters to Ganoderma, 16 year old palms



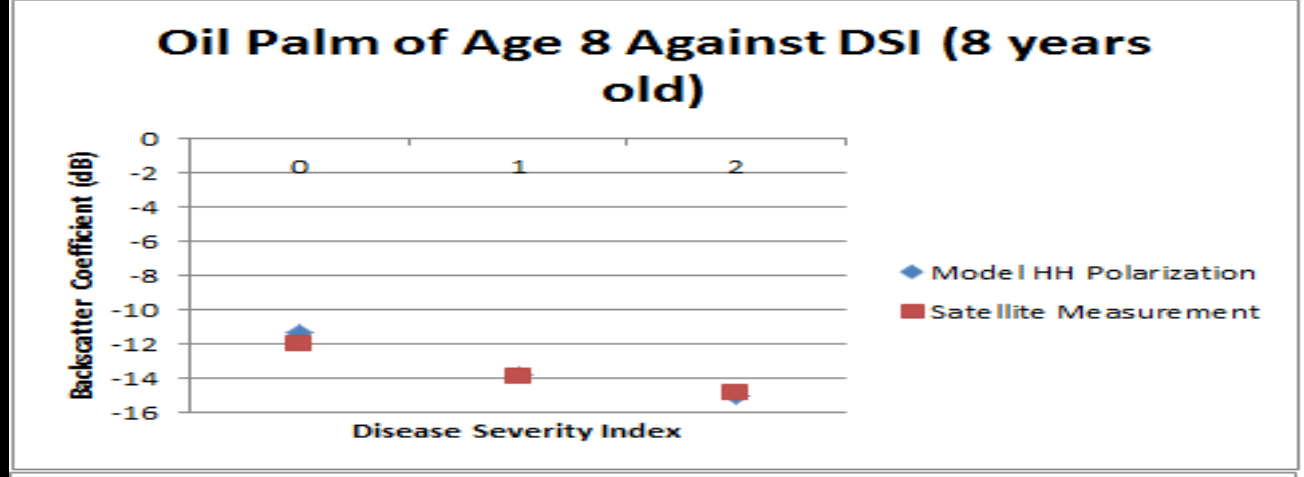


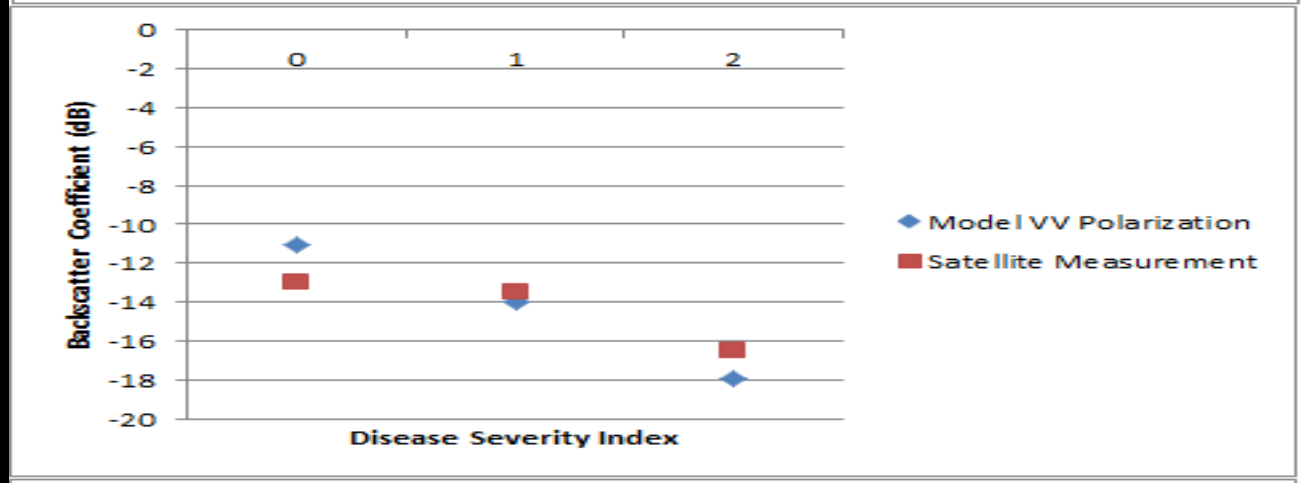


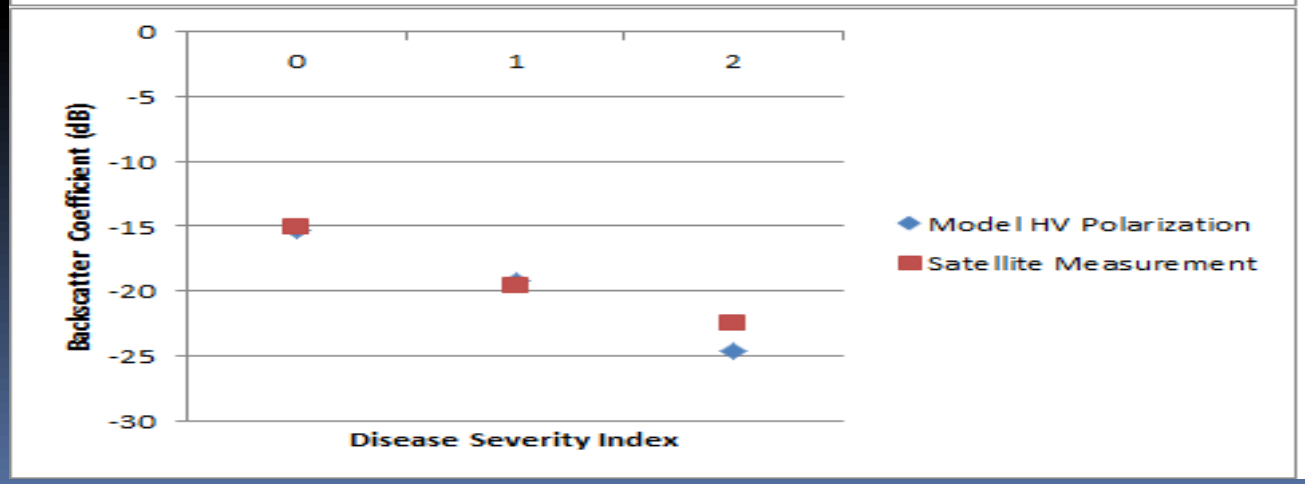


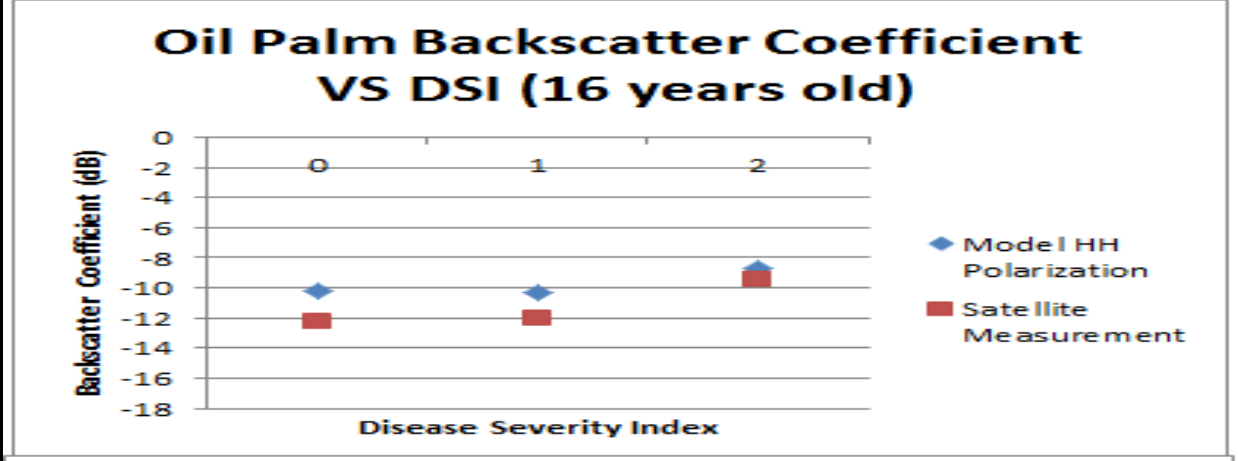
Garnoderma palms from Census Data

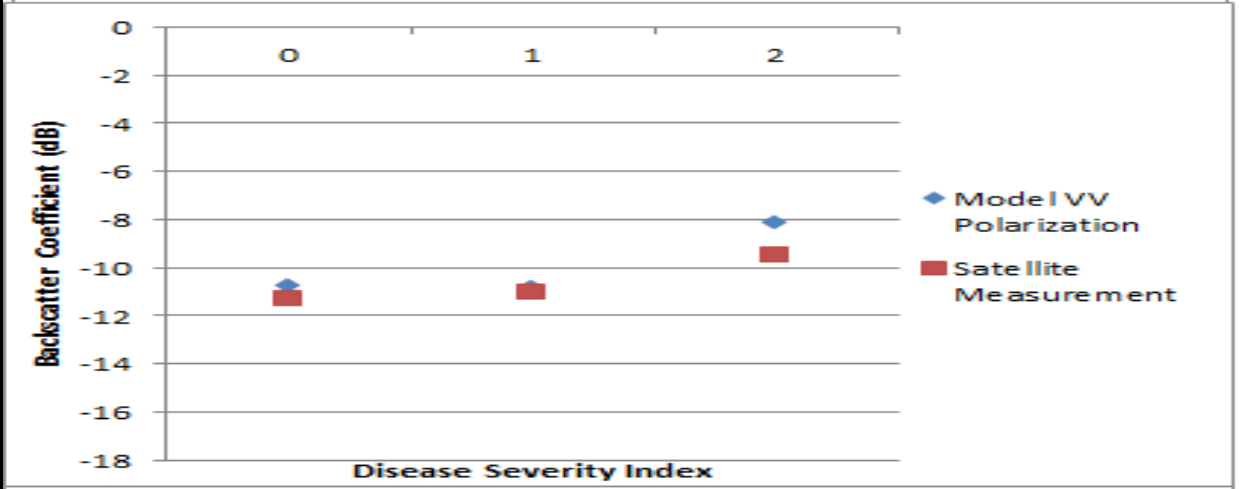
- Garnoderma census data obtained from Malaysia Oil Palm Board. This allows us to extract backscattering coefficient.
- 20 palms for tier 0, 1 and 2 palms were extracted for both age 8 and, 16 respectively. For each palm pixel, the average of 3x3 surrounding pixels was taken to address possible GPS location error.
- The values of the backscattering coefficient was averaged and showed in the following graphs.
- Theoretical modeling was applied with collected parameter .

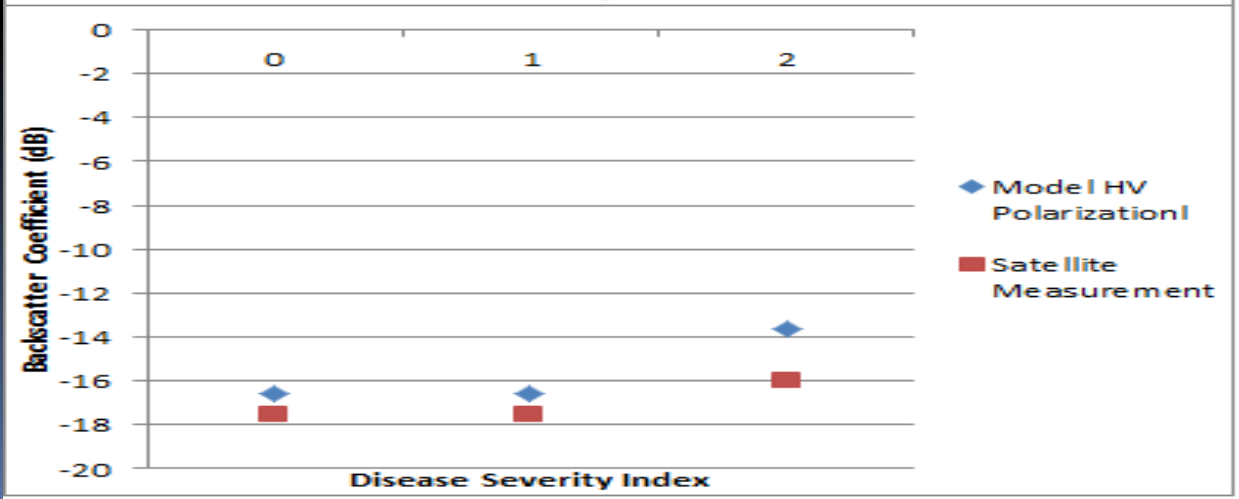

Oil Palm of Age 8 Against DSI (8 years old)
Backscatter Coefficient (dB)
0
-2
-4
-6
-8
-10
-12
-14
-16
0
1
2
Disease Severity Index
Model HH Polarization
Satellite Measurement


Backscatter Coefficient (dB)
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
0
1
2
Disease Severity Index
Model VV Polarization
Satellite Measurement


Backscatter Coefficient (dB)
0
-5
-10
-15
-20
-25
-30
0
1
2
Disease Severity Index
Model HV Polarization
Satellite Measurement


Oil Palm Backscatter Coefficient VS DSI (16 years old)
Backscatter Coefficient (dB)
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
0
1
2
Disease Severity Index
Model HH Polarization
Satellite Measurement


Backscatter Coefficient (dB)
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
0
1
2
Disease Severity Index
Model VV Polarization
Satellite Measurement


Backscatter Coefficient (dB)
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
0
1
2
Disease Severity Index
Model HV Polarizationl
Satellite Measurement

Classification of Oil Palm using L band SAR Image

Machine Learning on Oil Palm using SAR Image

- Overview and classification criteria
- Classification using SVM (Support Vector Machine)
- Use of CNN (Convolutional Neural Network)
- Hybrid SVM CNN Classifier

Machine Learning on Oil Palm using SAR Image

- From the data and ground truth, we split the oil palms into several categories.
- Based on the knowledge that backscatter is sensitive to frond then height of trunk at later stage.
- Separated into 5 categories as follows.

Categories/Classes

- Class 1, Red – Mature palms having petiole cross sectional area > 900cm^2 and tall in height > 7 meters.
- Class 2, Green – Growing palms having petiole cross sectional area > 900cm^2
- Class 3, Grey – Young palms having petiole cross sectional area between 600 to 900cm^2
- Class 4, White – Young palms younger than age of 5 and petiole cross sectional area < 600cm^2
- Class 5, Blue – Ground surface

Classification using SVM

- We first use SVM (Support Vector Machine) to classify the data, 50 data for per class, where 40 is used as training data and the rest as testing data, accuracy of 81.8%.


1PR2016B
1PM2009A
1PR2014B
1PR2012B
1PM1992C
1PM2004B
1PM2001B
1PM1999A
Taman Seri Kayan

CNN Classification of L band SAR Image

- Convolutional Neural Network (CNN)
- Similar dataset with SVM, 50 points of data per class.
- 40 points for training, 10 points for testing.
- Takes 3x3 data patch with respect of data point at center.
- Data augmentation: Mix of Horizontal flip and vertical flip, Rotation of 90, 180, and 270 degrees. Increases training data 7 fold.
- If data augmentation was not used, training accuracy was not good.

Results, Average Accuracy of 42%

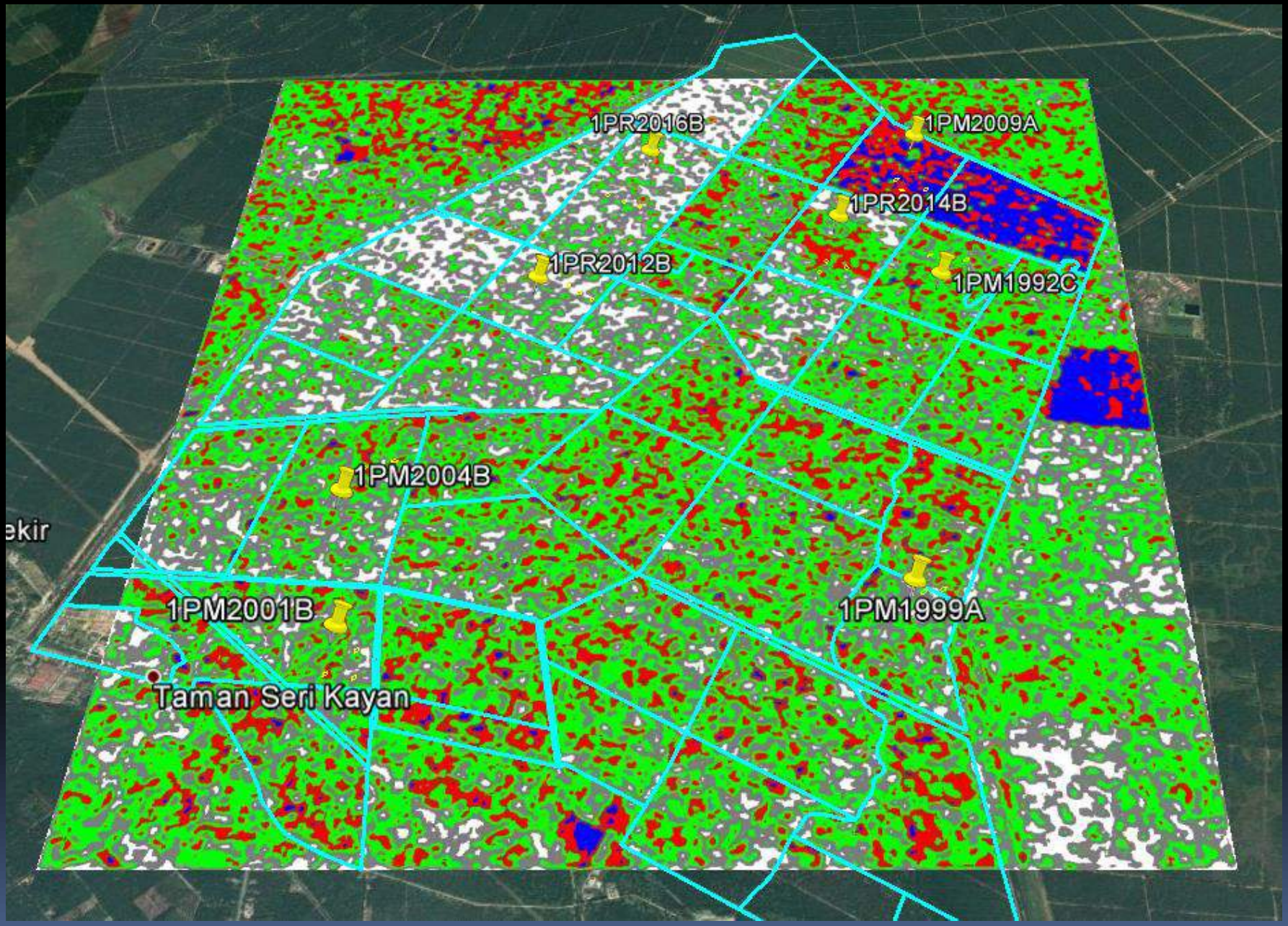
			Predicted Class			
	Class	1	2	3	4	5
True	1	60%	40%	0%	0%	0%
Class	2	20%	40%	0%	0%	40%
	3	20%	20%	30%	20%	10%
	4	10%	10%	30%	50%	20%
	5	0%	0%	50%	0%	50%

CNN + Model Simulation

- Previous results was below average.
- Lack of data is the problem.
- Compliment with simulation data generated from the model.
- 50 + 50 simulated points of data in the form of 3x3 patches.
- 80 for training, 20 for validation.
- Increases training data with augmentation also.

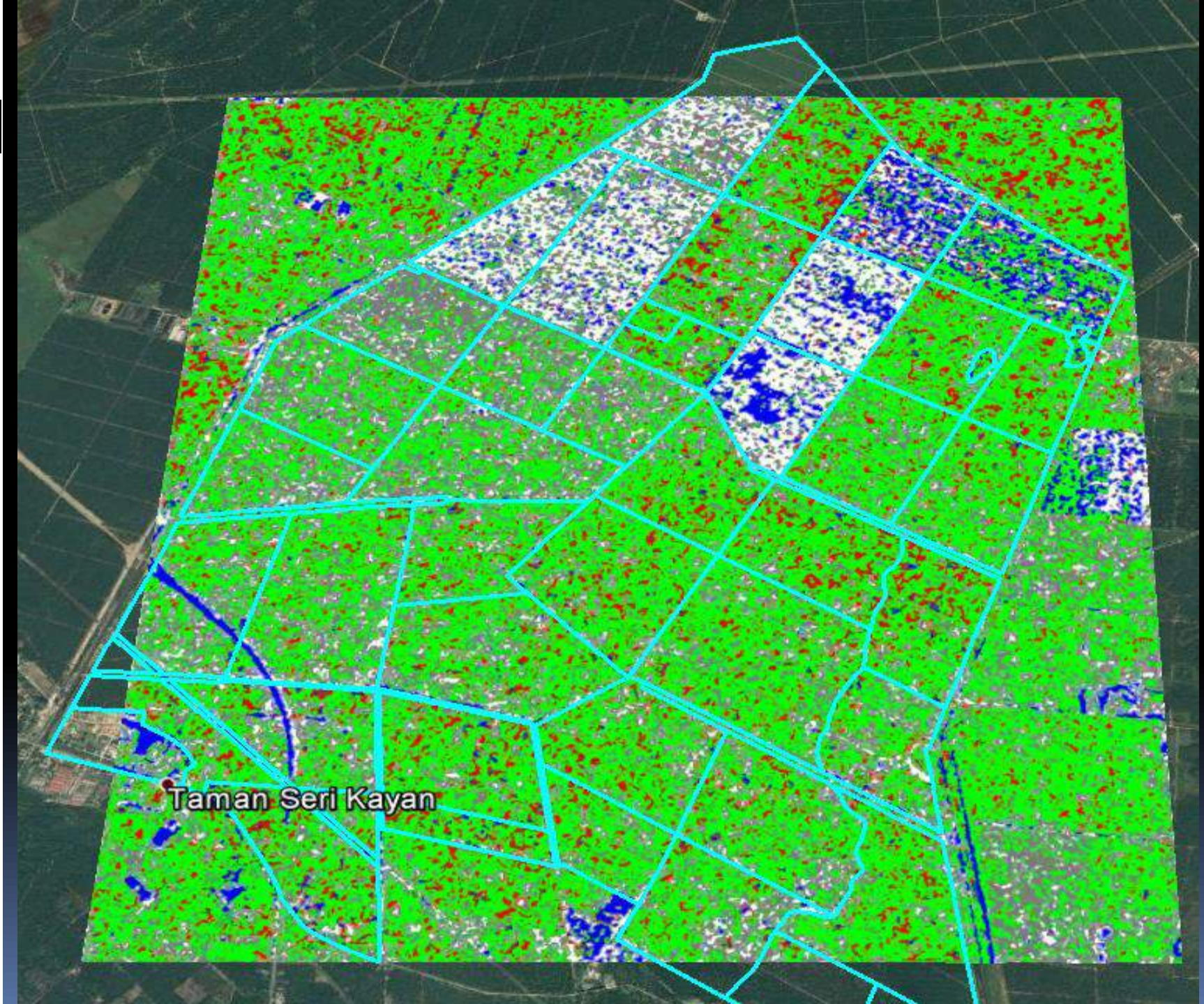
Results,65% average accuracy

			Predicted Class			
	Class	1	2	3	4	5
True	1	85%	10%	5%	5%	0%
Class	2	15%	65%	10%	5%	5%
	3	15%	25%	60%	0%	0%
	4	25%	5%	10%	55%	5%
	5	10%	25%	5%	0%	60%


1PR2016B
1PM2009A
1PR2014B
1PR2012B
1PM1992C
1PM2004B
ekir
1PM2001B
1PM1999A
Taman Seri Kayan

Discussion

- Good separation between class 1 with other classes, average on the other classes.
- Different from SVM Results.
- We can do simple hybrid classification by using SVM and CNN with weighted confidence from both techniques.
- $\mathrm{R} = w \times P(class = 1|using\ SVM) + (1 - w) \times P(class = 1|using\ CNN)$
- 87.4% Accuracy Achieved


Taman Seri Kayan

Summary

- Potential of Oil Palm monitoring using Big Data.
- Advancement of technology available to us to collect huge amount of data in short time span.
- Big data bringing in new knowledge and understanding.
- Classification using big data shows promise.

International Geoscience and Remote Sensing Symposium

IGARSS 2022

Kuala Lumpur, Malaysia

Hosted by:

Supported by:

Thank you and Welcome to IGARSS 2022 Malaysia!