Oil Palm Tree Yield Analysis

Scenario:

A team of plantation planners are concerned about the yield of oil palm trees, which seems to fluctuate. They have collected a set of data and needed help in analysing on how external factors influence fresh fruit bunch (FFB) yield. Some experts are of opinion that the flowering of oil palm tree determines the FFB yield, and are linked to the external factors. Perform the analysis, which requires some study on the background of oil palm tree physiology.

Domain Background Study: Oil Palm Tree Physiology

The oil palm is monoecious (can bear both male and female flowers) in an alternating cycle to prevent self-pollination. The tree matures after 30 months of field planting and will be productive all year long for the next 20 to 30 years. The younger trees have rougher trunks, older trees have smoother trunks. The tree may grow up to sixty feet and more in height. A typical oil palm's fruit bunch weighs between 20 to 25 kilograms, containing an average of 1,000 to 3,000 fruitlets. Peak crop yields are achieved from the age of 9 to 18, and gradually decline thereafter.

Oil Palm thrives in wet and humid environment with suitable amount of sunshine, warm temperatures with a bound of 22 to 33 °C, and a high rainfall rate of 2000 to 2500 mm per annual. Some known events that influence palm oil production are:

- El Niño-Southern Oscillation phenomenon with prolonged dry spells that may lead to bunch failure and floral abortion. This can last for a few months.
- La Niña brings heavier rainfall, inducing poor pollination and fruit-sets of the trees, and may disrupts harvesting and logistics management.
- Production of oil palm can be naturally disrupted for 1 to 2 years due to sex differentiation.
- In water deficit, floral initiation can be delayed, male/female ratio disrupted, and the rate of flower abortion becomes higher.

Climate change is expected to bring about stronger El Niño and La Niña events. This may have negative impact on oil palm production.

Apart from weather conditions, oil palm also requires soil with good texture allows reasonable drainage while retaining plenty of exchangeable cations and contains a good amount of organic matter. The trees should be planted on lands no higher than 20° of slope. Soil water shortage can reduce the number and mass of FFB. See visualisation in Figure 1.

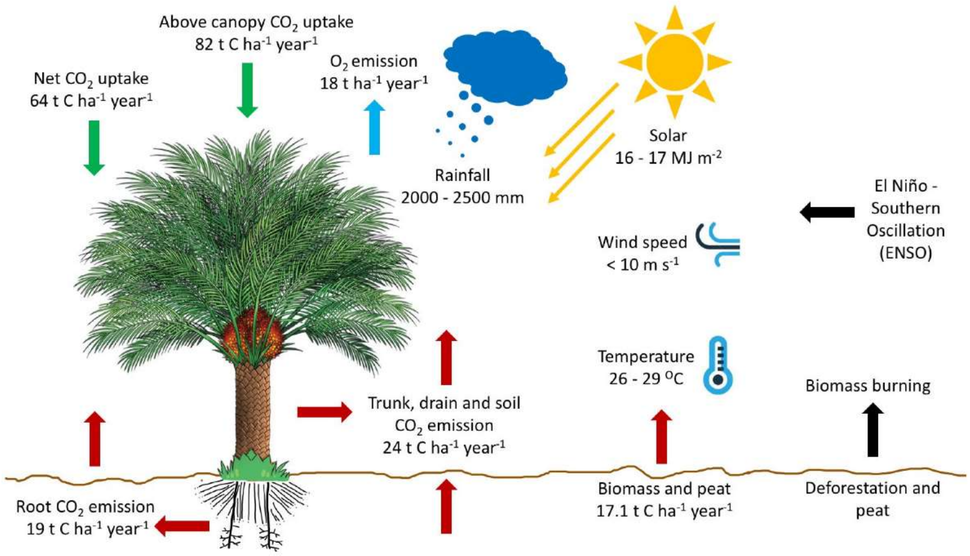

Figure. 1. Environment required for productive oil palm growth. Sourced (Uning et al, 2020)

[1] Uning R, Latif MT, Othman M, Juneng L, Mohd Hanif N, Nadzir MSM, Abdul Maulud KN, Jaafar WSWM, Said NFS, Ahamad F, Takriff MS. A Review of Southeast Asian Oil Palm and Its CO2 Fluxes. Sustainability. 2020; 12(12):5077. https://doi.org/10.3390/su12125077

[2] Filho WRLL, Rodrigues FHS, Ferreira IVL, Correa LO, Cunha RL, Pinheiro HA. Physiological responses of young oil palm (Elaeis guineensis Jacq.) plants to repetitive water deficit events. Industrial Crops and Products. 2021; 172:4052. ISSN 0926-6690. https://doi.org/10.1016/j.indcrop.2021.114052.

Summary:

i. Oil palms have no seasons but have natural sex-changing cycle that may disrupt FFB yield.
ii. Extreme temperatures negatively impact FFB yield, consistent optimal temperature is between 22–29 °C.
iii. Water deficit and high-water content negatively impacts FFB yield by reduction in FFB production, mainly from El Niño and La Niña events.

Analysis:

Table 1 shows the properties of the dataset that is received. A time series analysis is performed on the dataset. More detailed information on the individual variables can be found at the interactive .HTML report under Variables tag.

< profiling_report_for_oil_palm_yield_variables.html>

S.No.	Variable	Unit	Description	Frequency
1	SoilMoisture	Millimetres in top 0-1.6m soil	How much water presents in top 0-1.6 meter soil per unit area	Monthly
2	Average Temp	Celsius degree	Average temperature throughout the specified duration	Monthly
3	Min Temp	Celsius degree	Minimum temperature throughout the specified duration	Monthly
4	Max Temp	Celsius degree	Maximum temperature throughout the specified duration	Monthly
5	Precipitation	Millimetres	Accumulative precipitation in a month	Monthly
6	Working days	Days	Number of working days of harvesting	Monthly
7	HA Harvested	Hectare	Hectare of land covered for harvest	Monthly
8	FFB Yield	Tonnes/Hectare	Fresh Fruit Bunch Yield (Tonnes/Hectare)	Monthly
9	Date	DD/MM/YYYY	Day Month Year of record	Monthly

Table. 1. The dataset properties.

A Shapiro-Wilk normality test shows that the variables have a mix of normal and non-normal distribution. Therefore, Spearman's Rank Correlation will be used to assess the relationship between two variables.

The evaluation metric for Spearman's Coefficient (adapted from Dancey and Reidy, 2004) is as follows:

Spearman R	Correlation	Observed pairs
>=0.7	Very strong positive relationship	“Average_Temp” and “Max_Temp”
0.4 to 0.69	Strong positive relationship	“SoilMoisture” and “Precipitation”, “Average_Temp” and “HA_Harvested”
0.3 to 0.39	Moderate positive relationship	“Min_Temp” and “Precipitation” “Max_Temp” and “HA_Harvested”
0.2 to 0.29	Weak positive relationship	none
-0.19 to 0.19	No relationship	“SoilMoisture” and “Min_Temp”, “SoilMoisture” and “Working_days”, “SoilMoisture” and “FFB_Yield”, “Average_Temp” and “Min_Temp” “Average_Temp” and “Working_days”, **“Average_Temp” and “FFB_Yield”,** “Min_Temp” and “Max_Temp”, “Min_Temp” and “Working_days”, “Min_Temp” and “HA_Harvested, **“Min_Temp” and “FFB_Yield”,** “Max_Temp” and “Working_days”, **“Max_Temp” and “FFB_Yield”,** “Precipitation” and “Working_days”, “Working_days” and “HA_Harvested”, **“Working_days” and “FFB_Yield”**
-0.2 to -0.29	Weak negative relationship	“Precipitation” and “HA_Harvested”
-0.3 to -0.39	Moderate negative relationship	“SoilMoisture” and “HA_Harvested”, “Average_Temp” and “Precipitation”, **“Precipitation” and “FFB_Yield”,** **“HA_Harvested” and “FFB_Yield”**
-0.4 to -0.69	Strong negative relationship	“SoilMoisture” and “Average_Temp”, “SoilMoisture” and “Max_Temp”,

		"Max_Temp" and "Precipitation"
<=-0.7	Very strong negative relationship	none

Table. 2. The correlation between each observed variables and the strength of the correlation. Relationships with FFB yield is highlighted in bold.

From Table 2 the following relationships can be compiled:

(i) Higher maximum temperature results in higher average temperature
(ii) Higher precipitation results in higher soil moisture (soil is wet when rain is present)
(iii) Higher minimum temperature results in higher precipitation (humid conditions)
(iv) Higher precipitation and soil moisture results in lower hector area harvested and vice versa (harvesting process presumably disrupted due to raining condition)
(v) Higher average temperature results in lower precipitation and vice versa (heat and dryness)
(vi) Higher precipitation results in lower FFB Yield and vice versa (flowers died and less successful in bearing fruits)
(vii) Higher soil moisture results in lower average temperature and maximum temperature and vice versa (Heat and dryness)
(viii) Higher maximum temperature results in lower precipitation and vice versa (heat and dryness)

However, there is a moderate negative relationship between higher hector area harvested and FFB yield. Intuitively, more land area covered should results in higher FFB yield. It could be due to lower FFB growing on oil palm trees, therefore requires plantation workers to collect from more areas, or unknown external factors. It is also observed from Table 2 that working days is not correlated to any other variables.

Below shows the visualisations of the correlation of variables and FFB yield over a span of 10 years, plotted on monthly data.

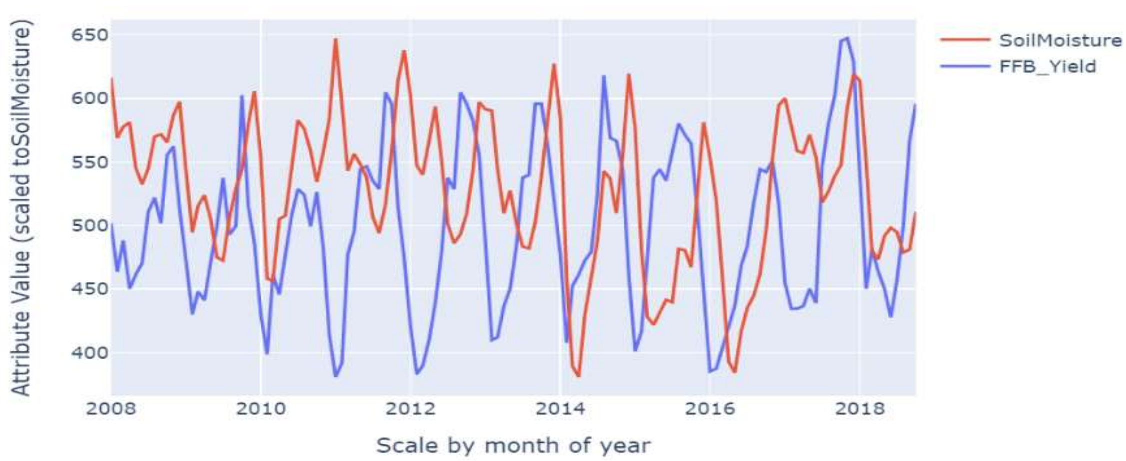

Figure. 2. Correlation plot between FFB yield with soil moisture, scaled for visualisation.

Overall, in general, FFB yield decreases as soil moisture decreases with 1-2 months lag.

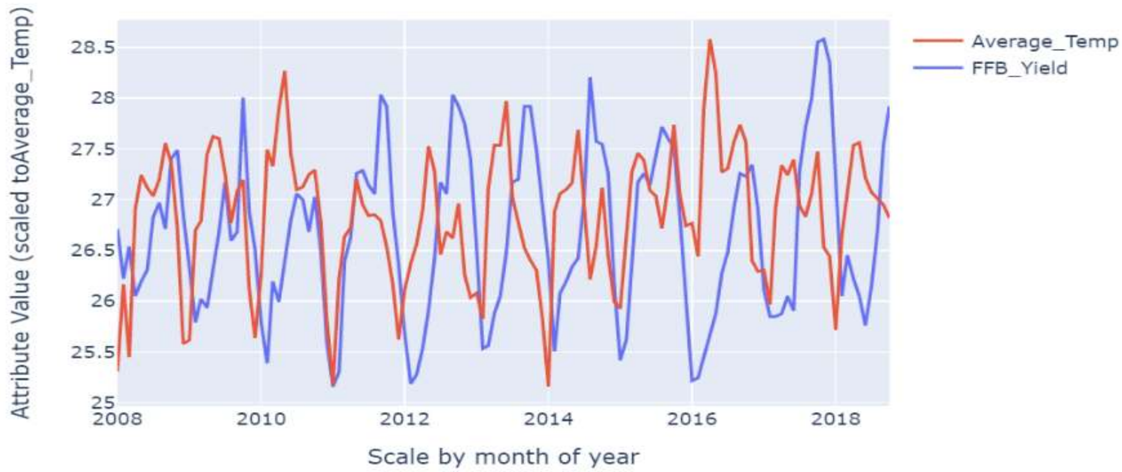

Figure. 3. Correlation plot between FFB yield with average temperature, scaled for visualisation.

Similarly, FFB yield increases with average temperature increases except for extreme temperatures above dotted line such as event between 2010 to 2011, 2013 to 2014, and 2016 to 2017.

Figure. 4. Correlation plot between FFB yield with precipitation, scaled for visualisation.

From the Spearman's correlation test in (vi) states there is a moderately negative relationship between FFB yield and precipitation. However, from Figure 4 it can be observed FFB yield generally benefitted with the increase of precipitation except during extreme cases such as 2011 events or sharp spikes of precipitation within such amount of time span of 1 to 3 months.

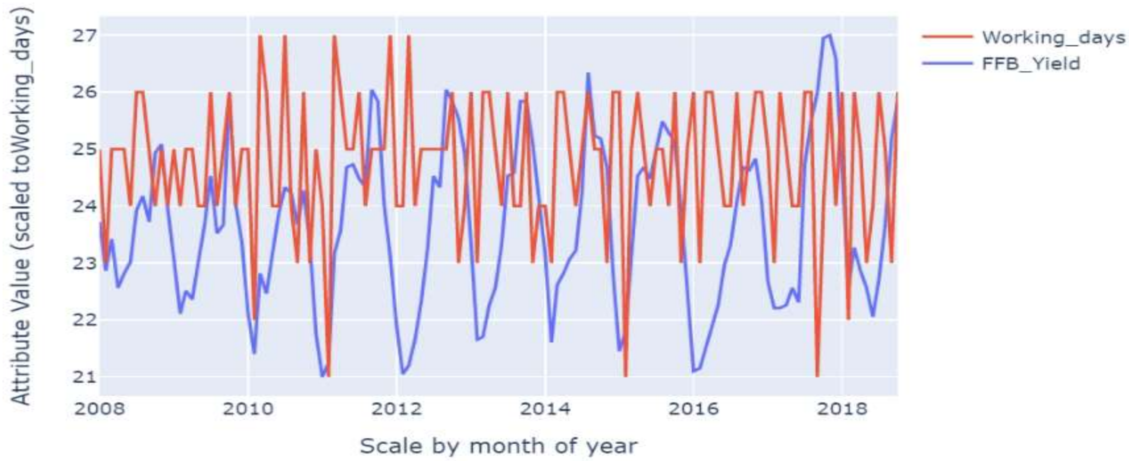

Figure. 5. Correlation plot between FFB yield with working days, scaled for visualisation.

There are no obvious trends or patterns observed in Figure 5. Number of working days in a month has no impact on the FFB yield.

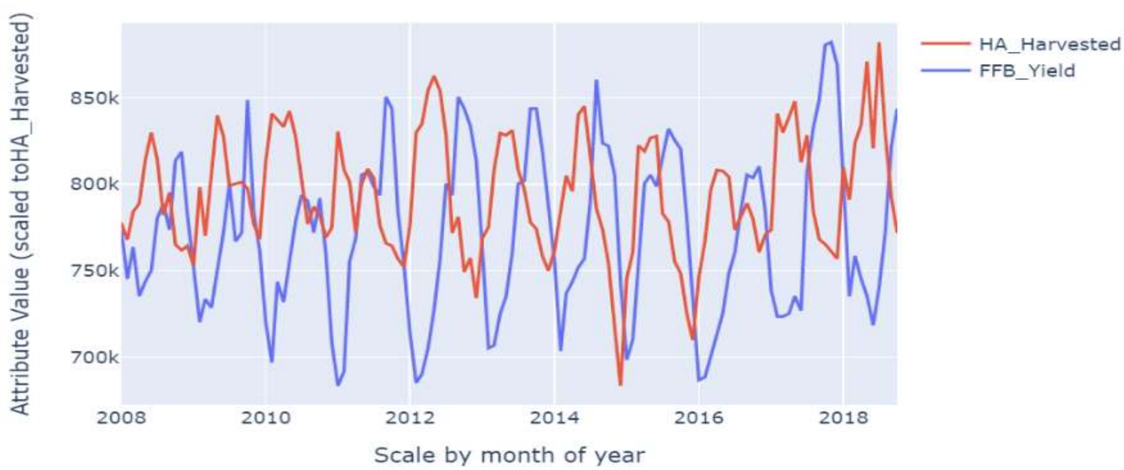

Figure. 6. Correlation plot between FFB yield with hector area harvested, scaled for visualisation.

There are no obvious trends or patterns observed in Figure 6. Hector area harvested has no impact on FFB Yield.

Augmented Dickey-Fuller (ADF) test is done on each variable to determine their stationarity. “SoilMoisture”, “Average_Temp”, “Max_Temp”, “Precipitation”, “Working_days” and “HA_Harvested” are non-stationary, meaning that they are seasonal and fluctuate across time. “Min_Temp” and “FFB_Yield” are stationary, meaning that they are consistent across time.

To visualise the seasonality, the variables are plotted on time series plot over 10 years.

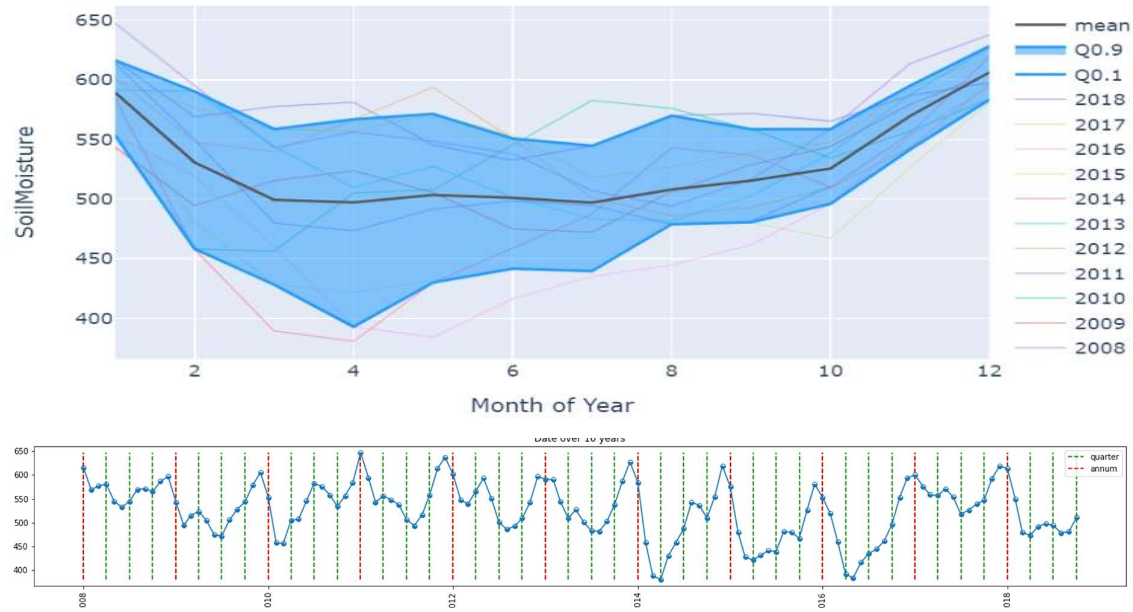

Figure. 7. Time series plot on soil moisture monthly over 10 years.

The trend observed from Figure 7 is the soil starts with lower moisture at the beginning of Q1 of each year and ends with increased moisture at throughout year. Exception observed in 2013 and 2018.

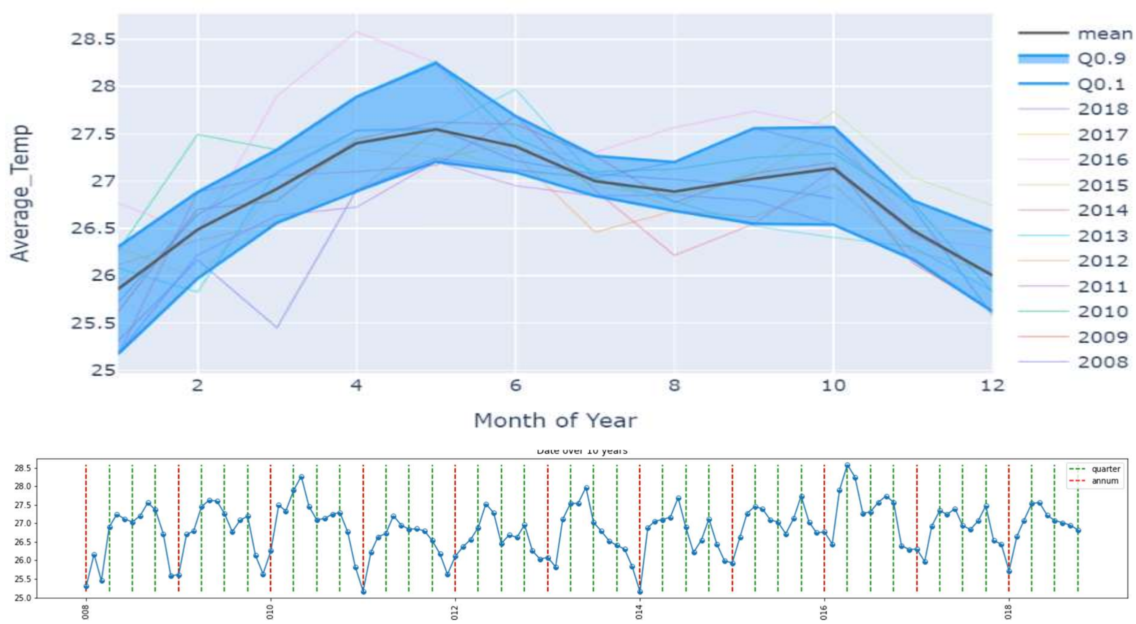

Figure. 8. Time series plot on average temperature monthly over 10 years.

The trend observed from Figure 8 is end of Q1 and Q2 have higher temperatures. Temperature drops every Q4.

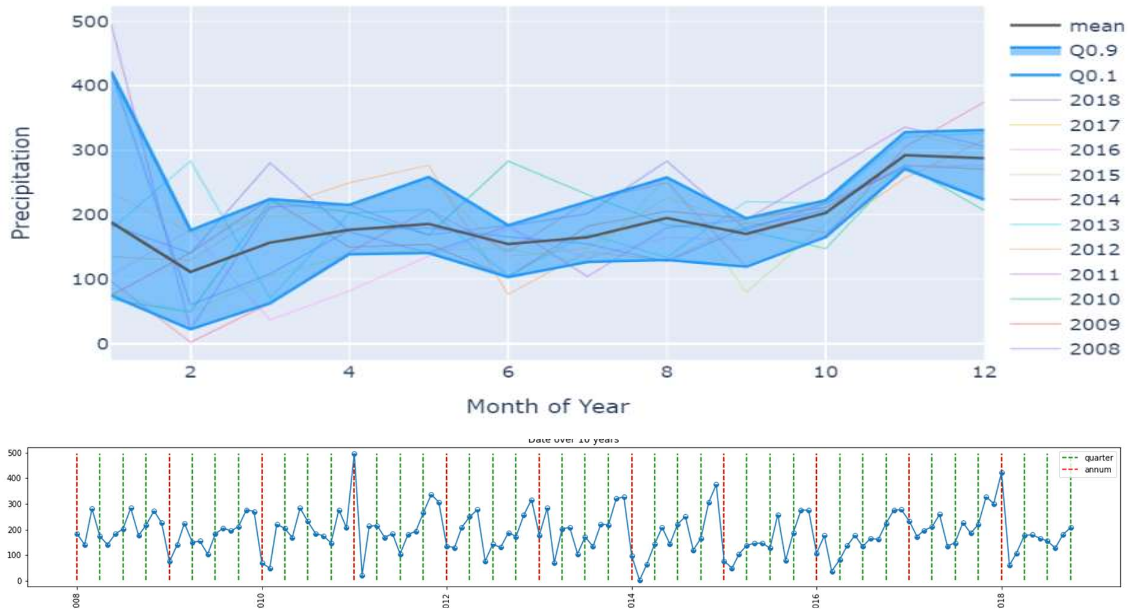

Figure. 9. Time series plot on precipitation monthly over 10 years.

The trend observed from Figure 9 is each year starts with lower precipitation and ends with increased precipitation at throughout year. More extreme precipitation at Q4 of 2010 and 2017 observed.

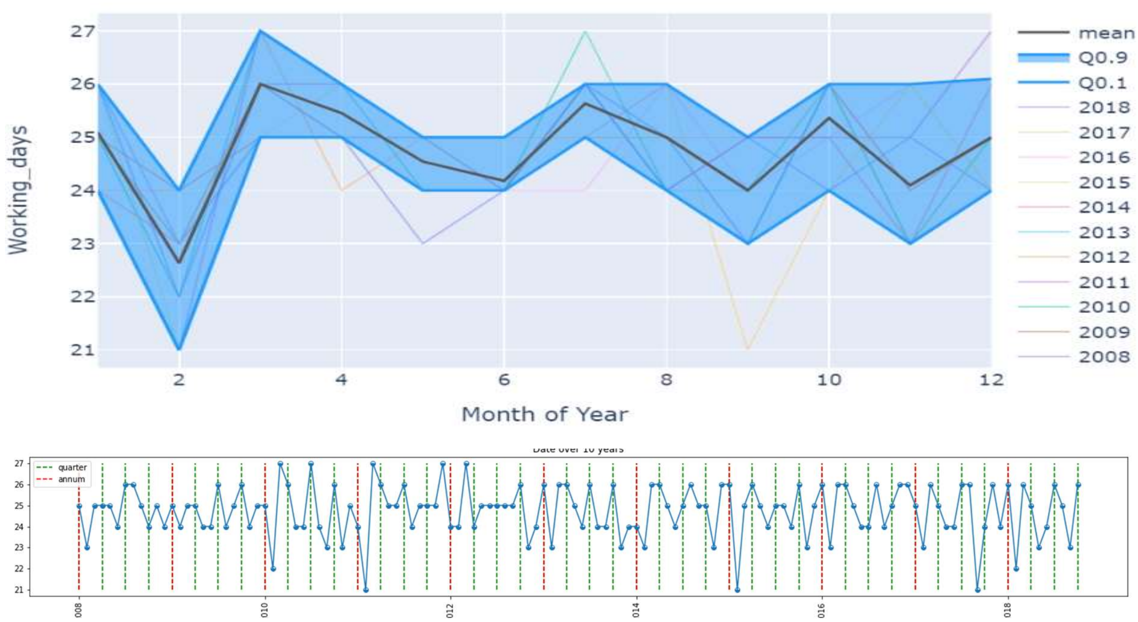

Figure. 10. Time series plot on working days monthly over 10 years.

No significant trend observed for working days in Figure 10.

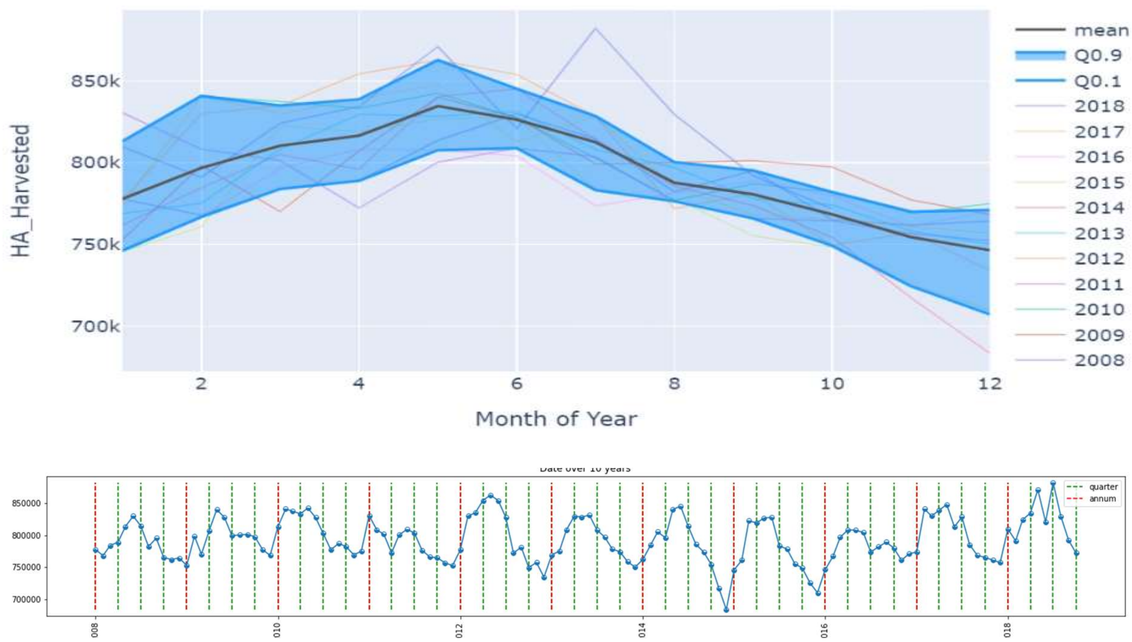

Figure. 11. Time series plot on hector area harvested monthly over 10 years.

The trend observed from Figure 11 is each year starts with on hector area harvested, peaks at Q2 then decline again except 2018.

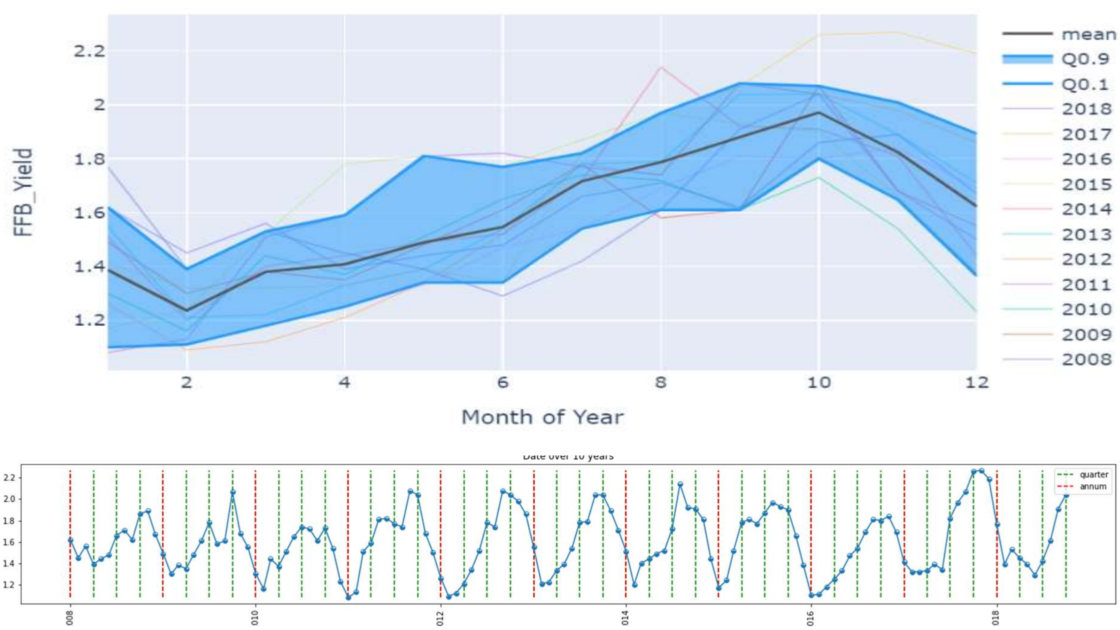

Figure. 12. Time series plot on FFB yield monthly over 10 years.

Although statistically proven stationary, there is a trend observed from Figure 12 is each year starts with FFB yield and ends with increased precipitation at throughout year where the yield peaks by end of Q3 and decline at start to mid of Q4.

The distributions of three weather-dependent variables are analysed to find trends.

Available at https://github.com/aplatyps/data_science_demo

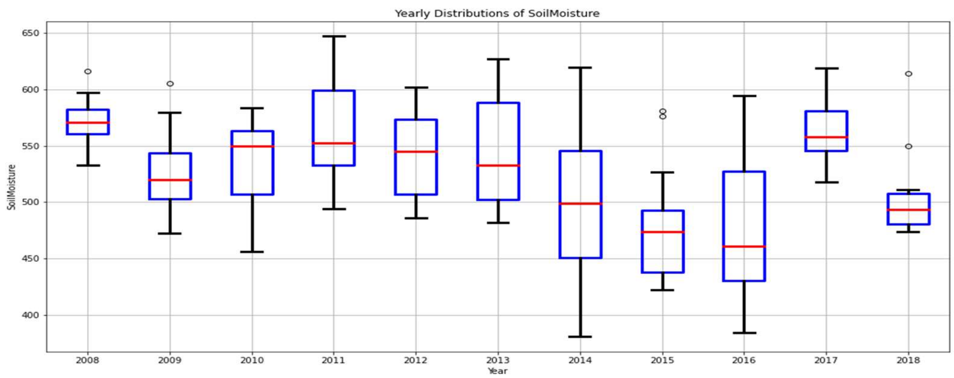

Figure. 13. Box plot showing the distribution of soil moisture over 10 years.

Figure 13 shows that 2011 has highest soil moisture; 2014 has the most fluctuations in soil moisture; 2008, 2009, 2015 and 2018 have some abnormal spike of soil moisture. Overall, soil moisture fluctuation follows the trend of El Niño and La Niña events (refer to El Niño and La Niña Years and Intensities table at https://ggweather.com/enso/oni.htm) .

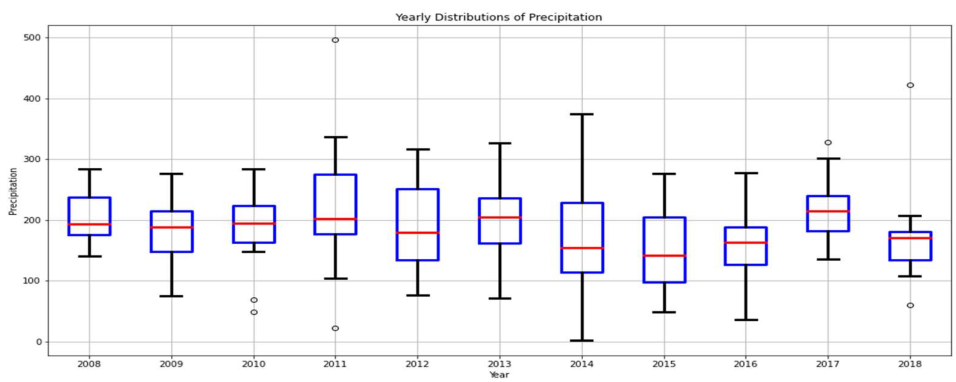

Figure. 14. Box plot showing the distribution of precipitation over 10 years.

Figure 14 shows very similar trend as soil moisture albeit with less fluctuations. 2014 has the most fluctuations in precipitation within a year. Abnormally high amount of precipitation can be observed in 2011, 2017 and 2018, while abnormally low precipitation can be observed in 2010, 2011 and 2018.

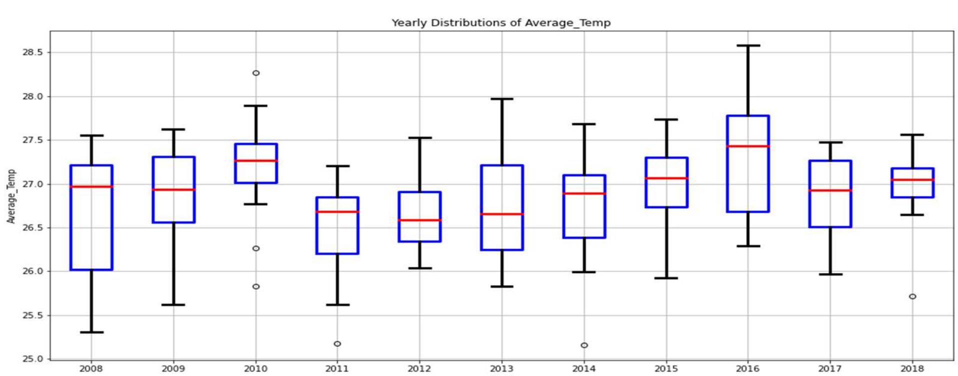

Figure. 15. Box plot showing the distribution of average temperature over 10 years.

Overall, the average temperature still fluctuates well within the range where oil palm trees thrive. From Figure 15 it is observed that 2016 has the highest increase in average temperature. Average temperature spikes can be observed in 2010 while dips can be observed in 2011, 2014 and 2018.

Interpretation of data analysis:

Low of season FFB yield usually lasts from November until February; moderate season is usually from March until August and the peak season is either in September or October of every year. The decline of FFB yield between November and February is likely linked to the higher precipitation observed in the same period, where monsoon rains occur. Impacts in the long term include El Niño event that occurs every 3 years on average, with strong El Niño events typically occur every 6 to 10 years, while La Niña events occur every 3 to 5 but on occasion can occur over successive years.

In the domain background study, it is mentioned that oil palms are very sensitive to water content and low moisture is the most common stress condition that can negatively impact FFB production. The data also shows that FFB yield reacts faster to soil moisture increase or decrease. The most critical moisture stress periods are 24 months, 18 months, and 5 months prior to the maturation of the fruit bunches. Within these critical periods, if the soil moisture and were to have serious fluctuations such as higher soil moisture than usual or sudden low soil moisture, FFB yield will experience drastic decline.

Precipitation has been noticed to have a negative correlation to FFB yield. Sex determination of the flowers occurs during 24 months before fruit ripening. When subjected to water stress such as high precipitation during sex determination, a higher proportion of flowers become male, therefore unable to develop into fruit bunches for the next 18 months.

From the data, 2017 appears to have the best conditions of soil moisture and precipitation to achieve best FFB yield. However, tree age and maturity need to be taken in consideration as well given that oil palm tree peaks in productivity at age 9 to 18.

Summary of causes of FFB yield fluctuations observed from data analysis:

- Too little soil moisture during El Niño and too much soil moisture during La Niña and monsoon cycle.
- Plant natural sex-changing cycle disrupted by weather conditions.
- Age/maturity factor.

=== END OF REPORT ===

Available at https://github.com/aplatyps/data_science_demo