

COSC2669 | Case Studies in Data Science

Sugarcane crop and soil analysis using satellite imagery

10th Oct, 2019

Master of Data Science, RMIT University

Group - 08

**Vishwa Gandhi – s3714805, Jigar Mangukiya - s3715807, Vikas Virani – s3715555,
Salina Bharthu – s3736867, Ria Talwar – s3729618, Sarthak Sirari – s3766477**

Contents

1.	Introduction	3
	1.1 Soil Suitability Evaluation.....	3
	1.2 Yield forecasting.....	4
	1.3 Crop health monitoring.....	4
	1.4 High-level Architecture	5
2.	Soil suitability evaluation	6
	2.1 Prologue	6
	2.2 Data source	7
	2.3 Methodology.....	7
3.	Yield Estimation Forecasting.....	10
	3.1 Prologue	10
	3.2 Calculating NDVI values	11
	3.3 Removing cloud cover.....	11
	3.4 Methodology.....	13
4.	Crop health monitoring (Under development)	15
	4.1 Prologue.....	15
	4.2 Challenges	16
5.	User Interface	17
	5.1 Prologue	17
	5.2 Technology Stack	17
	5.3 Home Page Dashboard - Screen 1.....	17
	5.3.1 Crop Health Meter (Functionality in Pipeline)	17
	5.3.2 Yield Forecating.....	17
	5.3.3 Satellite and corresponding NDVI Images	17
	5.4 Soil Suitability Evaluation - Screen 2	18
6.	Conclusion	20
7.	Team Management.....	21
	References.....	23

1. Introduction

The report is written to supplement the study undertaken as part of a dataathon project “satellite intelligence”. High level goal of the project was to utilize remotely sensed data to make a positive impact on sugarcane industry in Australia. Our project is oriented towards adding value to sugarcane farmers in planning, expanding & maintaining their crops & business. We formulated 3 major use-cases to harness the power of machine learning & AI along with satellite imagery to help them make more informed and data driven decisions for stakeholders in sugarcane industry. These use-cases are, 1) Yield forecasting (planning), 2) Crop health monitoring (maintaining) & 3) Soil suitability evaluation (expanding). We made a consolidated user interface, which brings together all 3 functionalities and facilitates the decision making and analysis.

1.1 Soil Suitability Evaluation

- Sugarcane, like all other crops, require essential nutrients which are acquired from soil. Nutrients like Nitrogen, phosphorous & potassium are very basic for growth of any crop. Other factors of soil that affect the crop quality and growing potential are slit, water exchange capacity, mineral exchange capacity etc.
- Every crop has minimum level of required these nutrients as well as optimum levels of these nutrients. Some nutrients like earth salts are not suitable for certain types of plants or trees while they are necessary for cultivation of palm trees & coconut trees etc.
- When sugarcane farmers want to extend their business by buying new land, it becomes necessary for them to assess the suitability of land they are purchasing. As land is a substantial investment and return on this investment depends on land's ability to support cultivation, soil suitability surveys are some of the most essential considerations apart from logistical and other aspects of new land.
- Traditional approach towards this assessment is random sampling & laboratory based analysis. This approach is often exhaustive and hence expensive, as testers have to manually gather and assess samples from multiple locations. Even after these surveys, information can be less accurate than expected as it is not possible to sample entire farm to assess the quality.
- Using rasters of soil nutrient level data and sugarcane mask geometries, we have trained a machine learning model to identify potentially suitable soil for sugarcane cultivation. This model can narrow down the area of manual search and improve accuracy of such efforts.

1.2 Yield forecasting

- Forecasting approximation of yield is a valuable information for farmers as well as other stakeholders like government bodies, insurance agencies and sugar mills.
- Farmers can utilize forecasted yields to plan their finances and compare the data with previously recorded yields to assess the quality of crops.
- Similarly government bodies, insurance agencies & sugar mills can utilize forecasts to plan their logistical & financial strategies.
- We utilized sentinel 2A imagery from various bands to estimate current yield and then utilize this estimation to forecast the yields for multiple harvest cycles.

1.3 Crop health monitoring

- Sugarcane is commodity produce and hence farming for sugarcane has undergone industrialization. As a result of which, sugarcane farms cover large areas.
- Knowing the health of the sugarcane is an essential information for farmers. Like most of the other crops, sugarcane farming is heavily utilizing hybrid seeds, which have greater immunity towards disease and pests. However, crops are still prone to fungus and some pests which have, over the time, developed immunity to the hybrid seed culture.
- Keeping track of crops spread across large geographical area in real time, is challenging and resource consuming.
- However, recent advances in satellite sensor technology, affordability of launching satellites, minimization of size of electronics has paved the way for organizations like planet Ltd., who have capabilities of imaging entire earth every single day using high resolution sensors.
- High frequency and easy availability of satellite images, commoditization of computational resources & advances in fields of ML & AI makes a complete package for monitoring and analyzing large geographical areas using satellite imagery and ML models. These implementations have been utilized by numerous governments as well as private companies to monitor disease in crops, flood water, ocean levels & volume of Antarctic ice.
- Utilizing crop health index, we are in process of deploying a model that can detect sudden anomalies in sugarcane crop and can flag the harmful regions that require further attention and action from farmers.

1.4 High-level Architecture

- Following image shows the high level architecture to show how the solution is deployed and how all modules are unified with a consolidated node-js UI.

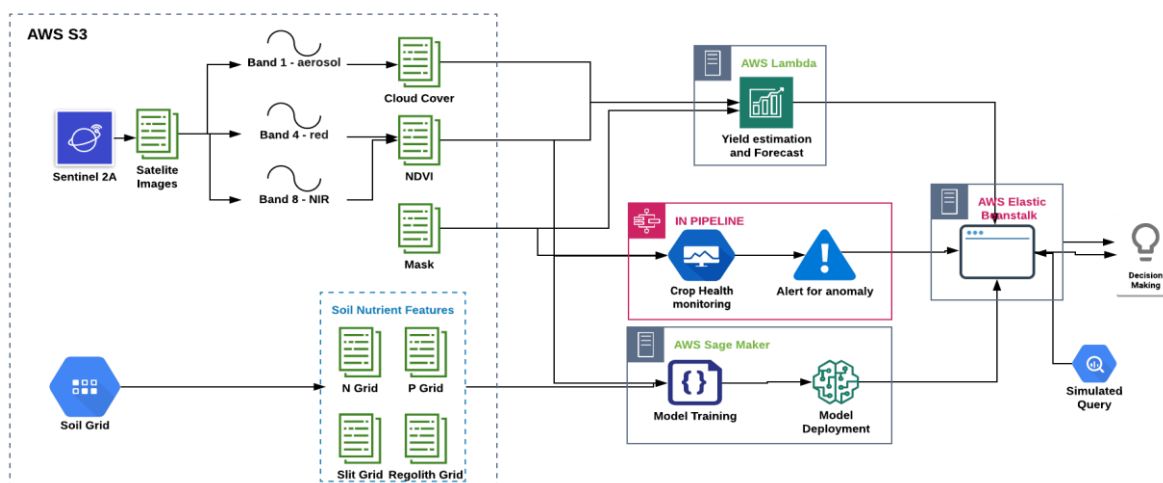


Figure 1- High Level Architecture

- Dataset is in image or GeoTiff format and hosted on AWS Simple storage service (S3) buckets, optimized for parallel reads.
- 3 modules described above are main computational services. Yield forecasting & crop health monitoring models are deployed on AWS Lambda Serverless applications, which read data from S3 buckets in form of default tiles, mask images & NDVI images in various bands. These modules populate respective components on node-js UI upon user selection.
- Soil suitability evaluation module is a state preserving ML model, which is deployed on AWS SageMaker with exposed api endpoint to receive user selection for area to evaluate and find suitable farmland from the area.
- Node-js UI is deployed on AWS elastic beanstalk application which facilitates insights mining and decision making.
- There is simulation API which will simulate live data by periodically feeding new time-stamped satellite images to application. This service allows for visualizing and demonstrating expected performance of application.
- The application currently has limited live time as entire application runs on AWS Free-tier for the purpose of demonstration.

2. Soil suitability evaluation

2.1 Prologue

- Soil is the most valuable natural asset required for crop's growth. It is composed of various components like organic material, mineral material, water and air. Each plant species require specific composition of these soil components. Soil suitability evaluation is important practice to ensure plant specific soil composition in particular land. Manual evaluation is time consuming and resource exhaustive task. It involves human resources to physically visit the area, collect samples and test it in laboratories. Land suitability is the fitness measure of a land for crop cultivation. This is the primary step for any farming body planning for crop cultivation in a new land.

- Considering complexity involved in manual inspection, we have worked on alternative

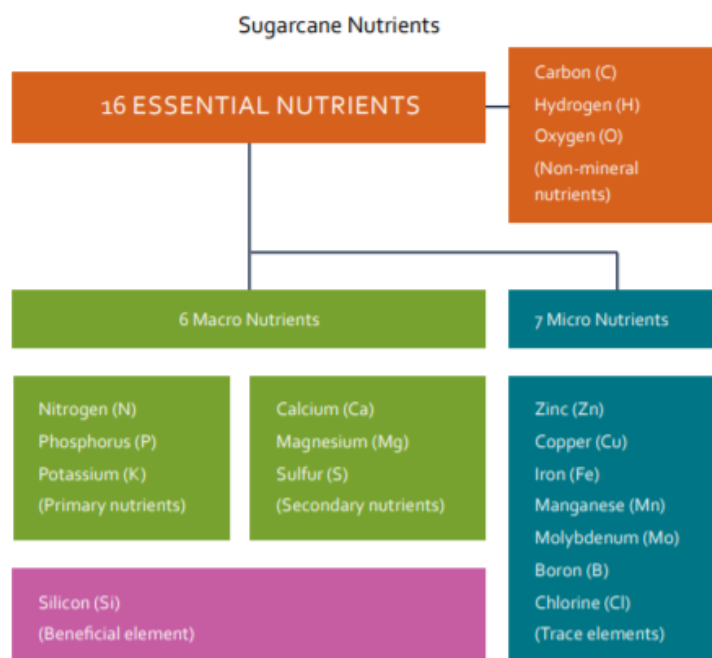


Figure 2- Sugarcane Nutrients Chart

approach using satellite data and soil grid data. Moreover, cost associated with manual approach is significantly reduced. With growing capabilities of crop harvesting and manufacturing, this sort of analysis approach enables farming bodies to prepare action plans in advance. Moreover, it can be further developed to help government for land usage planning. Objective of the development is to determine inherent capability of the land for specific crop cultivation using remote sensing data.

- All essential nutrients should be present in adequate proportion in soil crop system to maintain balances nutrients. Balance nutrients term means that all essential nutrients should be present in adequate concentration in the soil for potential crop yield.
- Effective depth in context of crop growth is the depth of the soil which is exploited by plant roots for nutrients and water for physical support. Sugarcane roots are majorly found in the top 25cm depth of soil.

2.2 Data source

- Data used for our application is collected from ASRIS (Australian Soil Resource Information System) for entire Australia. Data is in the geotif format which needs to be processed further to integrate with existing methods. These soil grids are available at various depths ranging from 0-30 cm at interval of 5/10cm. We have acquired data for soil characteristics like slit, regolith, water capacity, pH, Nitrogen, Phosphorous etc.

Modelling steps



Figure 3- Steps for soil suitability analysis

Following are the major implementation steps for this module:

- i. Acquiring enough data from reliable sources to support model performance
- ii. Trimming down the soil grid GeoTiff to fit with SENTINEL satellite imagery
- iii. Feature engineering model parameters to enhance performance measures
- iv. Building machine learning classification model to identify potential land based on sugarcane suitability

2.3 Methodology

- Various soil nutrients at different levels like 00-05 cm, 05-15 cm, 15-30cm, 30-60 cm are downloaded as a grid for whole Australia, which is quite a large file. This each file is then trimmed down to extract our area of interest only from the whole grid. For this purpose, QGIS software is used. Each of these file is loaded into QGIS software & JSON file with bounding box co-ordinates is provided to do raster extraction of particular location from whole file. The resulting tile of our area of interest is in the dimensions of 1197 by 1197 pixels. Since the soil grid data acquired from ASRIS is in the resolution of 90 meters per pixels & our image dataset is in the resolution of 10 meters per pixel. So, we've further resized these grid data to match our actual image dataset using SGRR (simple grid repetitive resampling) technique for grid word. Once a whole image for each nutrient is extracted, we've written an automated script which will divide these whole image into a set number of tiles same as our original data tiles.
- Labeling the current sugarcane region manually is difficult task so, based on the Geojson file which contains different shape polygons of current sugarcane growing regions, we have checked if those regions co-exist with our image data & based on if it exist or not, we've created a sugarcane masks which are essentially an identification of which part of particular image/area is a current sugarcane growing region for our image dataset. After getting all of the soil nutrients for each tile as our features & sugarcane masks of each tile as our target, we then overlaid these feature onto sugarcane masks to make our final

processed data set to train our model on. Below figure is an explanation of the combined dataset

Binary Classifier Model

Features Soil Nutrients Grid

+

Training Data Labels - Mask

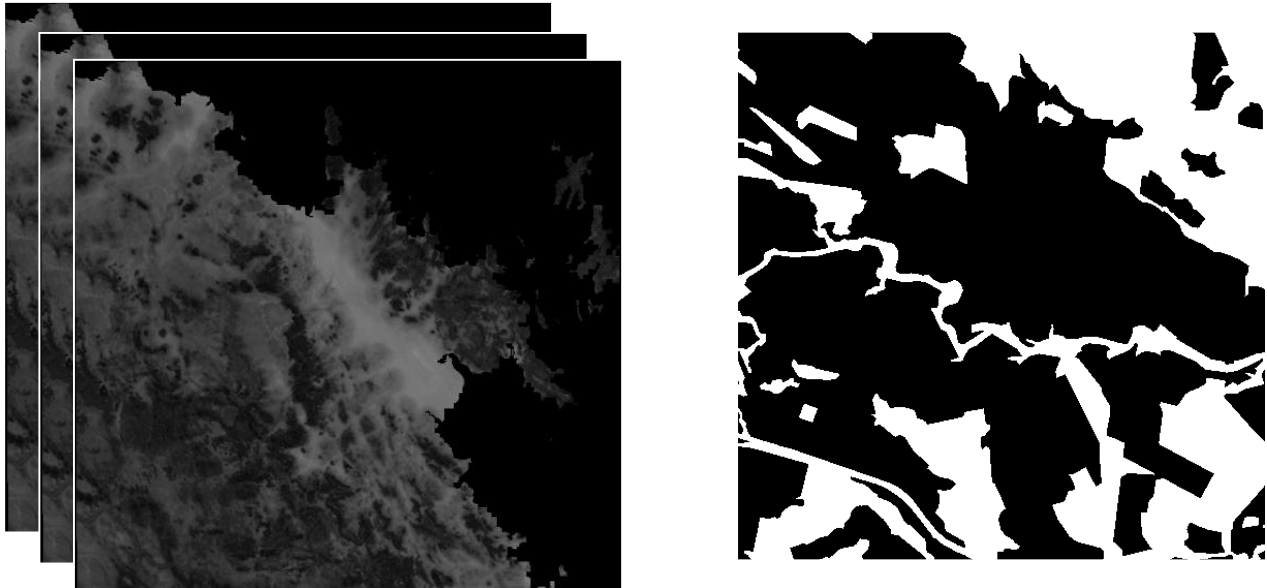


Figure 4- Soil grid and mask

- This makes for our final full dataset which consists of 65 image tiles with sugarcane masks values & corresponding soil nutrients grid values for each tile.
- By doing some feature engineering on these different nutrient values at different depth; this is essential as different nutrients at different depth will affect the growth of sugarcane differently, we've identified some important features; along with the sugarcane masks as our label, we've trained our XGBoosting algorithm to predict the potential region for sugarcane growth. Currently our model has an average f1-score of 0.51, as we've used only 6 of the features/nutrients to train our model for initial stage. We are tuning some parameters as well as adding more features like average monthly temperature of the area, average rainfall and other essential nutrients required in sugarcane growth. Below is a screenshot of accuracy at the time this report was written.

	precision	recall	f1-score	support
0	0.51	0.65	0.57	582394
1	0.51	0.37	0.43	582394
accuracy			0.51	1164788
macro avg	0.51	0.51	0.50	1164788
weighted avg	0.51	0.51	0.50	1164788

Figure 5- snapshot of model accuracy parameters

- To remove the effect of imbalance in sugarcane masks values, we've under sampled our dataset and trained it on 3.2 Million pixels. We've tested our model on approx. 1.1 Million pixels for which the overall average accuracy is 0.51. We then tested the model on whole image & the identified predicted area is as per the below image,

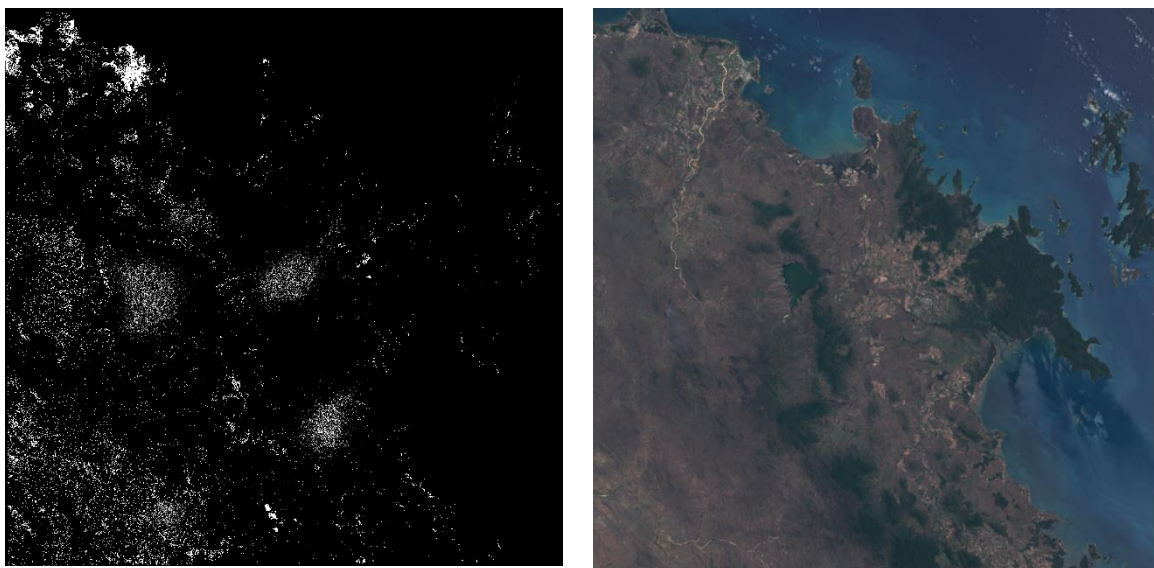


Figure 6- Map of sugarcane potential area marked

- In the above image, pixels marked white are the potential sugarcane area except from current one predicted by our model. As can be seen, this predictions are on pixel level, so instead of identifying soil suitability at each pixel, a farmer can narrow down a search based on group of pixel which are more suitable for sugarcane as potential growth area.
- As mentioned earlier, by using these inferences, a government or farmer can narrow down his/her search for future sugarcane growth to a particular region/area & from there further analysis can be done for soil suitability of sugarcane cultivation.
- Soil suitability may not have a direct lucrative benefits as ROI or profit from yield prediction but it an essential part to start with when growing a sugarcane in new areas & on a larger scale, it can potentially add much value to economy. From perspective of government, if you want to improve your sugarcane export, identifying soil that is most suitable for sugarcane will give you high mass of sugarcane from minimal soil use. Even a 5-10% increase compared to regular soil in which sugarcane is grown without testing nutrients can add value in millions to export industry.

3. Yield Estimation Forecasting

3.1 Prologue

- Predicting the crop yield is an important aspect from the farmer's perspective. There are various stakeholders who can utilize the yield prediction including farmers, banks and industry stakeholders.
- Yield estimation helps farmers to plan finances, forward selling and various crop management practices. Understanding probable yield outcomes over the course of the season has significant benefits when it comes to making forward crop sales.
- To help insure the farmer against the fall in prices in the future, the farmer can take the path of forward selling. The price and quantity are decided beforehand and the farmer must deliver on a specified future date. This helps in getting an estimate of the crop sold even before plantation.
- Combining yield outputs with good market analysis can significantly improve returns.
- If good prices exist and market is falling, having to sell in advance is invaluable.
- Various agribusiness companies already offer services for forward selling and have predetermined rates. A farmer who has an estimate of the yield can decide to forward sell.

2019 - In Season Pools					
Pricing Pool	Priced - Average AUD / Tonne Actual		Mark to Market - AUD / Tonne IPS		
	Percentage	Average	Gross	Shared Pool	Net
2019 MSF Guaranteed Floor	100%	\$403.50	\$389.20	-\$2.22	\$386.98
2019 MSF Actively Managed	55%	\$422.44	\$387.93	-\$2.22	\$385.71
2019 MSF Late Season Pool	27%	\$425.03	\$395.01	-\$2.22	\$392.79

2019 - Forward Season Pools					
Pricing Pool	Priced - Average AUD / Tonne Actual		Mark to Market - AUD / Tonne IPS		
	Percentage	Average	Gross	Shared Pool	Net
2019 Maryborough Collective	100%	\$422.03	\$406.94	-\$2.22	\$404.72
2019 Mulgrave Collective	100%	\$450.15	\$434.09	-\$2.22	\$431.87

Figure 7- Forward Selling sugarcane to mills

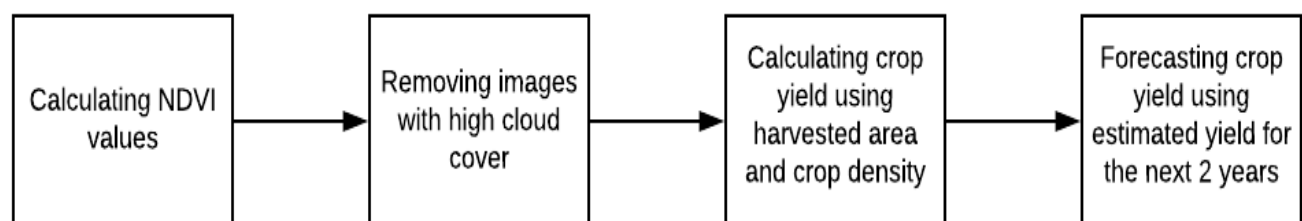


Figure 8- Steps for yield estimation and forecasting

- The steps followed for yield estimation and forecasting are as follows:
- We see a sample image for NDVI and Cloud Cover below:

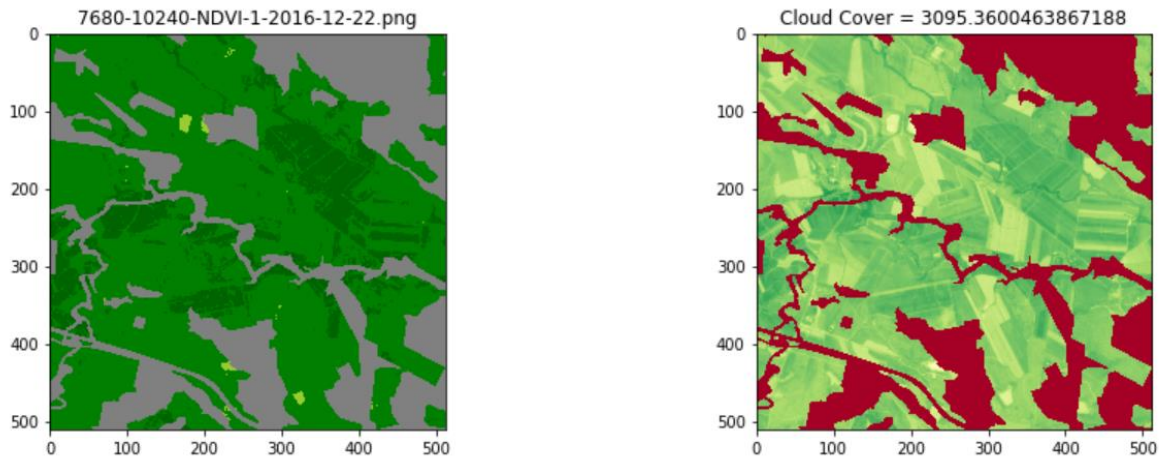


Figure 9- NDVI images with masks

3.2 Calculating NDVI values

- First, we calculate the NDVI values from the satellite images using Band 4 and Band 8. NDVI values are indicators of green-ness of cover surface reflectance and can be used to identify relative state of crops in an area. Greener crops have higher NDVI values than brown or unhealthy crops. Following formula is used to calculate NDVI value of each pixel in a raster.

NDVI is calculated in accordance with the formula:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

NIR - reflection in the near-infrared spectrum

RED - reflection in the red range of the spectrum

3.3 Removing cloud cover

- The cloud cover can be seen in the band 1 images. Band 1 images are generated from sensors to detect coastal aerosol on the cover surface. Clouds formed due to condensed

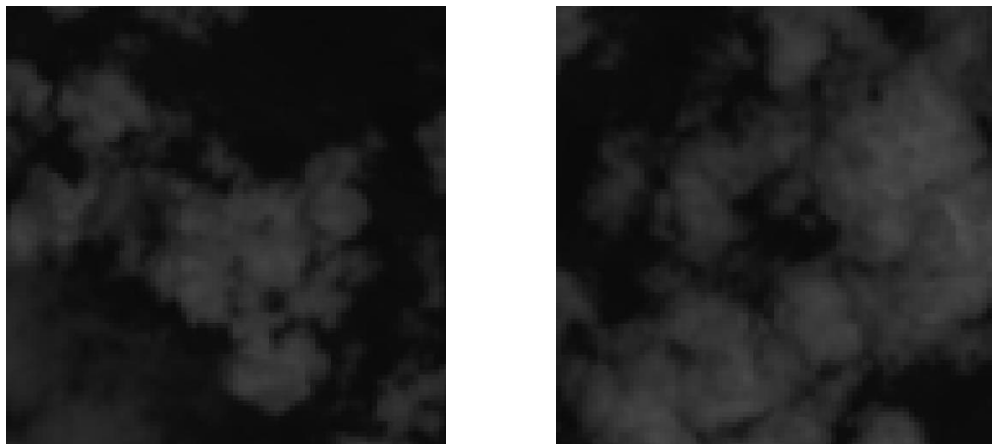


Figure 10- Band-1 Images with cloud cover

water vapour, can be easily detected using this sensor. Here we are utilizing the ability of relative cloud cover in a time stamped image. Once we have identified that image has

higher than a threshold cloud cover, we deal with those images treating them as special case.

- There are images with cloud cover, and it is a crucial problem because it introduces error and produce misleading results at later stage of analysis. Therefore, to deal with cloud cover of images, we have considered band-1 images, which is the data of coastal/aerosol band. This band is sensitive to cloud and other atmospheric effects and it is used to estimate the concentration of aerosols in the atmosphere, which can be used to refine the atmospheric correction procedures such as dark object subtraction.
- To handle this cloud covered images, we have first calculated the cloud cover index of images considering band-1 images.
- Now, we have NDVI values and cloud cover index values for each image.
- Below plot shows the cloud cover and NDVI value over the period of approx. 2 years:

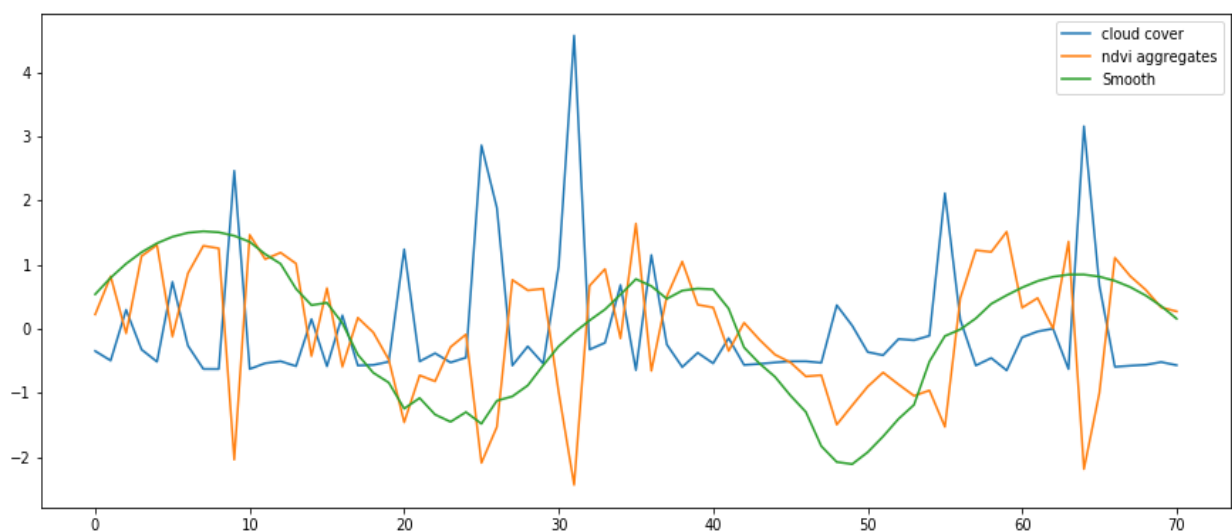


Figure 11- Cloud cover and NDVI for 2 years

- There are evident peaks in the plot of cloud cover and at the same point, negative peak in the plot of NDVI. This depicts the images having very high cloud coverage and due to that, very low vegetation index for that particular image. These images are filtered out.
- After this, to remedy the effect of this filtering, the NDVI data is smoothened using **Savgol function**. Savgol function is used with window length 23 and very small polyorder 2, as high window value and small polyorder value can make line smoother as shown in the above plot.
- Now, considering pixel by pixel NDVI value, the harvested area of all the images are marked followed by finding difference of harvested area of two consecutive images. This harvested area difference is further in-corporate with the data of crop density data of respective area during that period. The data of crop density is acquired from Australian Sugar Mill Council.

3.4 Methodology

- Now, in-order to forecast the crop yield data, the data is first converted to time series object and monthly average crop yield data is plotted in the below graph:

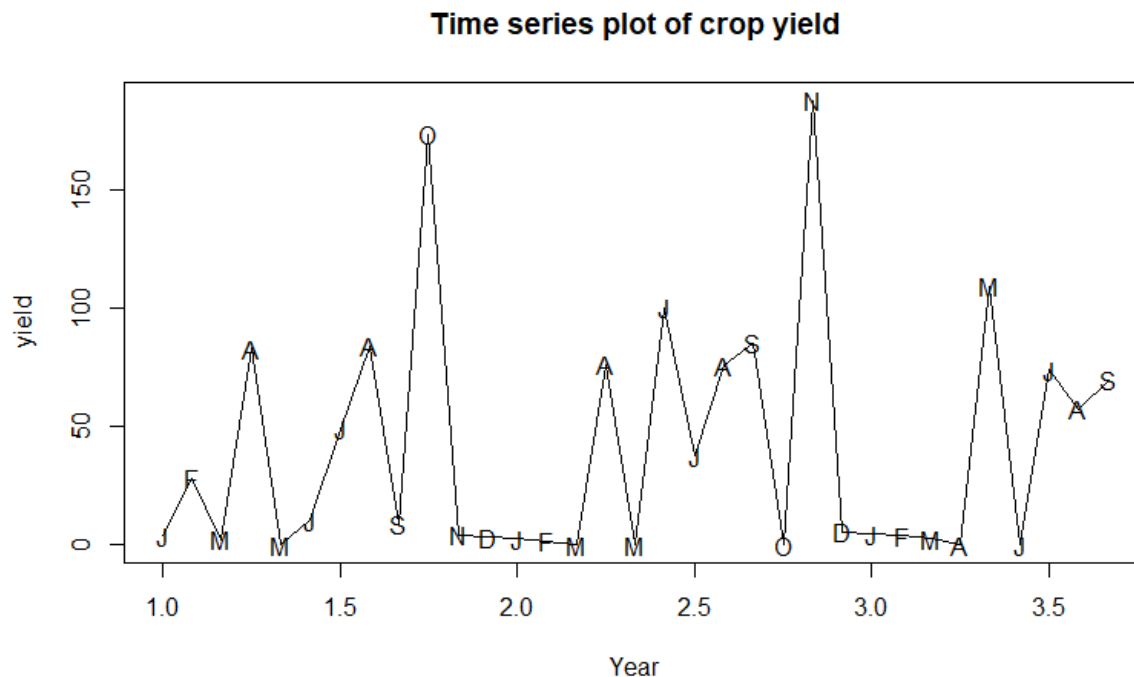


Figure 12- Estimated Yield Time series plot

- As we can see from the above plot, there is no evident trend in the crop yield data, on the contrary, seasonality is very apparent. Also, other components (such as stationarity, intervention etc.) of time series data are explored in-order to select the model for forecasting. For instance, to check stationarity of data, adf test is applied. Below snapshot shows the result of adf test:

```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression none  
  
Call:  
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
  
Residuals:  
    Min       1Q   Median       3Q      Max  
-90.398 -12.752  -0.005   52.550  145.526  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
z.lag.1         -0.3733    0.2085  -1.791  0.0838 .  
z.diff.lag      -0.4995    0.1657  -3.015  0.0053 **  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 58.77 on 29 degrees of freedom  
Multiple R-squared:  0.528,    Adjusted R-squared:  0.4955  
F-statistic: 16.22 on 2 and 29 DF,  p-value: 1.87e-05
```

Figure 13- ADF test summary to check stationarity

Null Hypothesis of ADF Test: series is non-stationary at a 5% level of significance.

Alternate Hypothesis: Series is stationary at 5% level of significance.

- As the p-value is less than 0.05, the crop yield time series data is stationary at 5% level of significance.

- Considering these characteristics of crop yield data, it is fitted to various forecast models in order to select the forecast model.
- There are two different types of models used in-order to forecast the yield data that are, state-space-models and exponential smoothing models. Based upon the value of error measures, the best fit model is selected.
- Below snapshot shows various error measures of two different types of models.
- Error measures of state-space-model with Additive error, no trend and additive

```

      AIC      AICc      BIC
380.3353 381.1629 384.8248

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -0.009865333 50.57679 42.27601 -Inf   Inf  0.9374269 -0.2738163
> |
seasonality:

```

- Error Measures of Holt winters' additive seasonality Exponential Smoothing Model:

```

      AIC      AICc      BIC
415.4776 464.3347 442.4147

Error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -1.1313 54.675 30.68245 -Inf   Inf  0.6803516 -0.6011228

```

- From the above snapshots, we can see values of various error measures such as AIC, BIC, AICc, MASE etc. We have considered holt winters' exponential smoothing models as value of Mean absolute scaled error (MASE) is considered to select the model.
- Below is the plot showing 2 years forecast data for crop yield with 95% of confidence interval data.

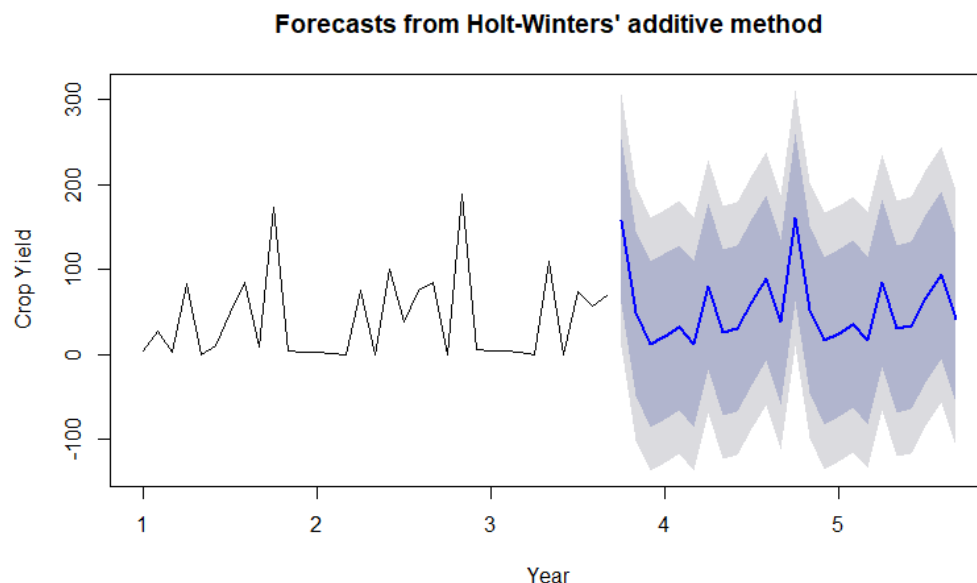


Figure 14- Crop Yield Forecast for 2 years

4. Crop health monitoring (Under development)

- This module is included in the functionalities statement after first two modules were implemented due to its practical usefulness and viability. However, this module is still under refinement and final implementation details are not determined yet. Following section outline the approach & challenges currently faced in implementation.

4.1 Prologue

- Lifecycle of sugarcane crop has a decisive curve which stays roughly the same year on year. We have utilized this behaviour of sugarcane to measure crop health using monthly average NDVI score. Using historical running average of NDVI values we formed a standard curve, which will become a reference dataset for us.
- As and when new data is available from satellites in form of time-stamped images, a new running average is calculated for that month. This becomes the observation point for the model to assess the R-Squared error from the reference plot.
- Following image from Thayse A., Dourado Hernandes from their study “Sugarcane land use and water resources assessment in the expansion area in Brazil” is an indicative plot for standard Normalized NDVI value for sugarcane life cycle.

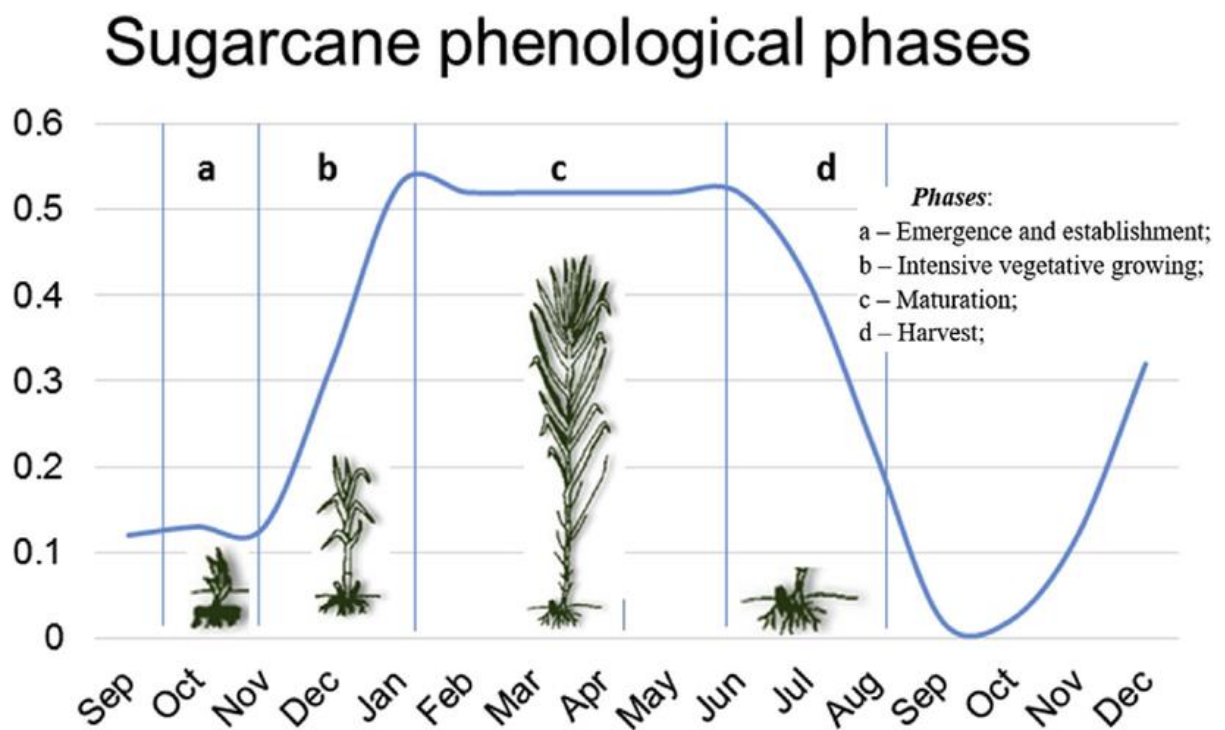


Figure 75- Sugarcane phonological phases

- Empirically determining threshold of normalized R-squared error measure to map to a particular category of crop health.
- For a very high fluctuation from standard plot, appropriate flags will be raised to let concerned stakeholders know about abnormalities using notifications on UI.

4.2 Challenges

- **Data:** Finding images with unhealthy crop cycles to better model and calibrate the R-Squared threshold
- **Cloud Cover:** If all images for a month has a cloud cover, no inferences about NDVI averages can be made without inducing large errors.
- **Validation:** it is inherently difficult to validate the accuracy of such models due to limited scope of the project. Validation requires on-field measurements which is not possible at the moment considering the timeline and overall scope.

5. User Interface

5.1 Prologue

- Predicting the crop yield is an important aspect from the farmer's perspective. There are various stakeholders who can utilize the yield prediction including farmers, banks and industry stakeholders.

5.2 Technology Stack

Following are the technologies used develop the application.

- Node.js
- HTML
- CSS
- JavaScript
- jQuery

5.3 Home Page Dashboard - Screen 1

Following image shows current implementation of prototype dashboard in home page.

5.3.1 Crop Health Meter (Functionality in Pipeline)

- We have developed a Health Crop Meter which shows the NDVI score for the current month. We use the monthly average NDVI scores from previous years and compare it with the current month's score. If any anomalies if found in the scores, i.e. the NDVI score drops significantly low, then the authorities could be alerted for possible threat of crop diseases in the region. There meter is divided into following three section: -
 1. GREEN Zone: This is the safe zone, which indicates the normal accumulated NDVI score for the current month.
 2. YELLOW Zone: This zone indicates below accumulated NDVI score for the month.
 3. RED Zone: If the NDVI score for the month lies in this zone, the authorities will be informed for a manual health check for crop diseases in the region.

5.3.2 Yield Forecatng

- In order to show the yield forecast for the upcoming months, we have used an Interactive Line Chart. This chart shows the monthly yield forecast for the next two years (in tonne per hectare). The user can hover over the chart to check the yield predictions for each month. Also, the user can drag and select a region in the chart to check the total yield and the average yield forecast of the selected time period.

5.3.3 Satellite and corresponding NDVI Images

- At the bottom of our dashboard, we have added a couple of Image Slider containing the timeline of the Satellite Images and their corresponding NDVI Images, along with their timestamp.

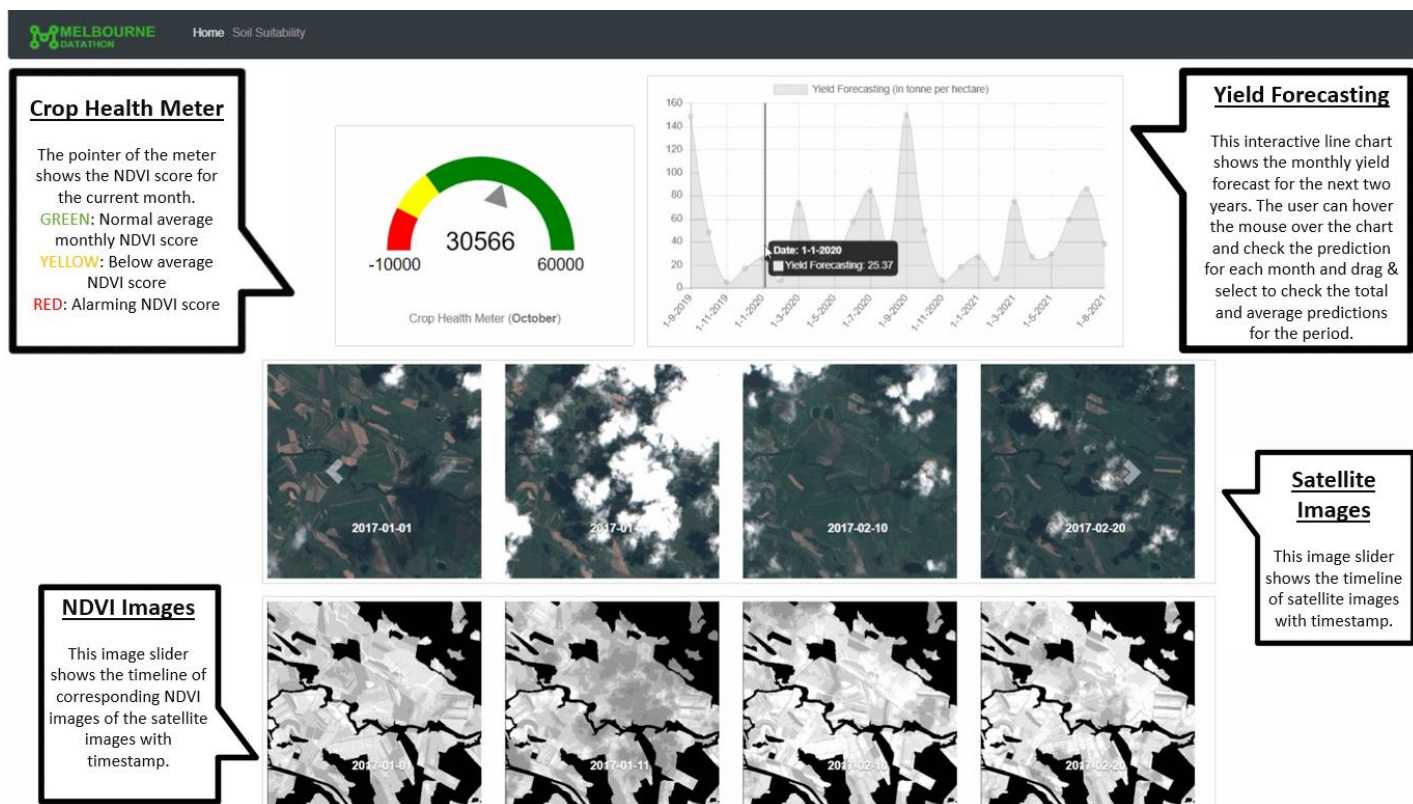


Figure 86- GUI snapshot Screen 1

5.4 Soil Suitability Evaluation - Screen 2

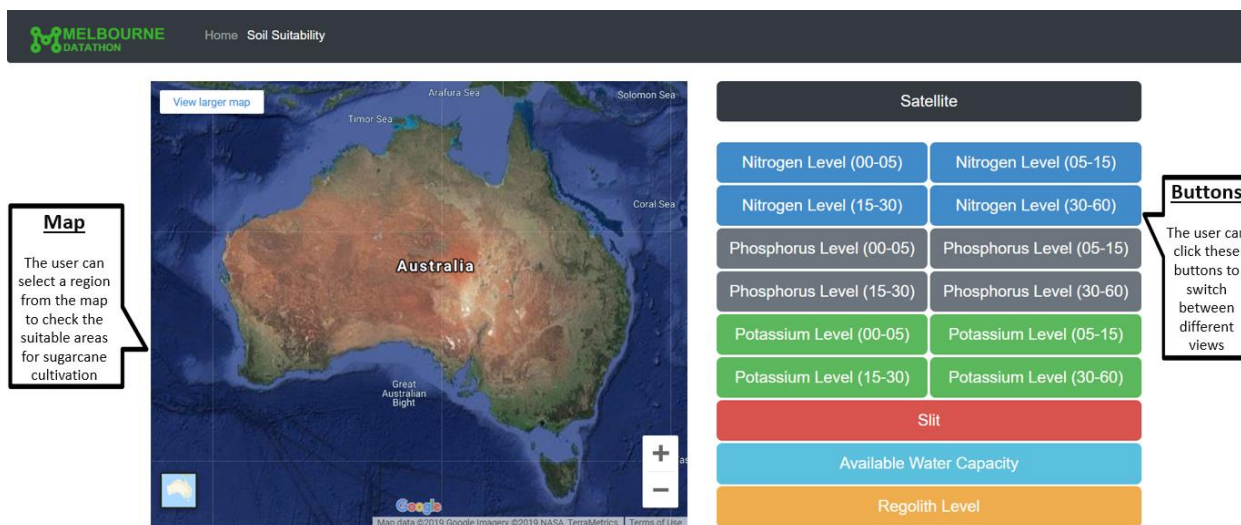


Figure 17- GUI snapshot Screen 2

- In the other screen shown below, we have a map in which the user can select any region and the suitable areas for sugarcane crop cultivation will be highlighted in the selected region, using our trained model. Apart from this, the user can view the Nitrogen,

Phosphorous and Potassium levels in the soil for any selected region at different depths, along with the along with Slit, Available Water Capacity and Regolith Level.

- Currently our model is limited to Australia for soil suitability evaluation, but we will expand our areas covered as we get their corresponding soil data.

6. Conclusion

- All 3 use-cases combined makes up for a good package for allowing various stake-holders in sugarcane farming and sugar industry in making data driven decision from insights derived from satellite data and processed using machine learning algorithms. The final solution that will included fully implemented use-cases with polished UI can serve as a platform to demonstrate viability of utilizing AI & ML models in sugarcane and in general agricultural sector. It also demonstrates the ability of cloud platform in hosting various type of computational resources to support latest trends in data analysis and modelling technologies.
- There are however, some limitations which can't be overcome due to limited time and scope of the project. Most of the data that is used for soil suitability evaluation is used as given without considering the temporal aspect of changing soil conditions. Models lack certain key validation attributes as it requires on-field observations and interactions which are not possible at this time. Soil nutrient data is offered by planet Ltd., but it is a paid resource cost of which falls beyond the scope of this demonstrative project
- Some temporary limitations which will be resolved in the final solution includes a more sophisticated version of crop health monitoring system and more functional and informative user interface which can facilitate detailed insight mining.

7. Team Management

This Project management report is for WIL project titled **Sugarcane crop and soil analysis**. Following table contains our work breakdown and task management details:

Task	Member
Initial data processing: Trimming down satellite image into tiles, code to mask sugarcane regions	All
Studying various stakeholders associated with sugarcane crop chain to validate uses cases of application	All
Data Acquisition: Soil grid data <ul style="list-style-type: none">- From reliable platforms, collecting relevant data for soil suitability analysis for sugarcane- Preprocessing soil grids to integrate with existing data	Vishwa, Vikas
Configuring to use SARA API to fetch real time data of SENTINEL satellite images	Salina, Sarthak
NDVI calculation and smoothing its value	Vikas
Removing cloud cover & calculating crop yield using NDVI based harvested area	Jigar
Yield estimation and forecasting	Salina, Ria
Validating forecasting model accuracy and enhancing performance	Salina, Ria
Building CICD pipeline for machine learning classification model to identify potential suitable soil <ul style="list-style-type: none">- Data preparation & preprocessing- Feature engineering to find effective features- Selecting optimal model and parameters	Jigar, Vikas , Vishwa
Testing against unknown data to model <ul style="list-style-type: none">- Performance evaluation of different models- Integrate with real time simulation module in UI	Vishwa, Sarthak
Crop health monitoring initial analysis <ul style="list-style-type: none">- Studying effect of monthly average NDVI on crop health- Uncovering parameters affecting sugarcane crop health	All
Application UI designing <ul style="list-style-type: none">- Home page dashboard: Crop health meter, yield forecasting interactive chart, satellite and corresponding NDVI image view panel- Soil suitability evaluation functionality page: Detailed view of soil characteristics data, Highlighting regions reflecting suitability result generated by model	Sarthak, Jigar, Ria
Presentation	All
Report Writing	All

Tools used:

- Trello: For tracking tasks, maintaining update status (<https://trello.com>)
- GitHub: Version controlling of data files and code (<https://github.com>)

The division of work was roughly equal and everyone has equally put in effort during entire project. Please consider it to be ~ 16.667 % per individual team member.

References

- [1] Molijn, R., Iannini, L., Vieira Rocha, J. and Hanssen, R. (2019). Sugarcane Productivity Mapping through C-Band and L-Band SAR and Optical Satellite Imagery. *Remote Sensing*, 11(9), p.1109.
- [2] Friedrich Theodor Rudorff, B.; Alves Aguiar, D.; Adami, M.; Pereira Galvao Salgado, M. Frost damage detection in sugarcane crop using MODIS images and SRTM data. In *Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium*, Munich, Germany, 22–27 July 2012; pp. 5709–5712.
- [3] Clw.csiro.au. (2019). Soil and Landscape Grid of Australia. [online] Available at: <https://www.clw.csiro.au/aclep/soilandlandscapegrid/> [Accessed 22 Oct. 2019].
- [4] Asris.csiro.au. (2019). A S R I S. [online] Available at: <http://www.asris.csiro.au/> [Accessed 22 Oct. 2019].
- [5] Scihub.copernicus.eu. (2019). Open Access Hub. [online] Available at: <https://scihub.copernicus.eu/> [Accessed 22 Oct. 2019].
- [6] Wiedenfeld, R. Water stress during different sugarcane growth periods on yield and response to N fertilization. *Agric. Water Manag.* 2000, 43, 173–182.
- [7] De Souza, C.H.W.; Lamparelli, R.A.C.; Rocha, J.V.; Magalhães, P.S.G. Height estimation of sugarcane using an unmanned aerial system (UAS) based on structure from motion (SfM) point clouds. *Int. J. Remote Sens.* 2017, 38, 2218–2230.
- [8] Johnson, R.; Viator, R.; Veremis, J.; Richard, E.; Zimba, P. Discrimination of sugarcane varieties with pigment profiles and high resolution, hyperspectral leaf reflectance data. *J. Assoc. Sugar Cane Technol.* 2008, 28, 63–75.
- [9] uveiller, G.; López-Lozano, R.; Baruth, B. Enhanced Processing of 1-km Spatial Resolution fAPAR Time Series for Sugarcane Yield Forecasting and Monitoring. *Remote Sens.* 2013, 5, 1091–1116.
- [10] Australian Sugar Milling Council. (2019). Australian Sugar Milling Council | The Peak Body for Raw Sugar Producers and Exporters. [online] Available at: <https://asmc.com.au/> [Accessed 22 Oct. 2019].
- [11] Association, S. and Industry, S. (2019). Sugar Mills - Australian Sugar Cane Farmers Association. [online] Australian Sugar Cane Farmers Association. Available at: <http://www.acfa.com.au/sugar-industry/sugar-milling/> [Accessed 22 Oct. 2019].