AgriSense

A Project Report

Submitted by

Vidhi Agrawal
Darpan Bendre
Sadhika Bhatia
Darshan Daliya

Under the Guidance of

Prof. Kamal Mistry

In partial fulfilment for the award of the degree of

BACHELORS OF TECHNOLOGY

COMPUTER ENGINEERING

At

MUKESH PATEL SCHOOL OF TECHNOLOGY
MANAGEMENT AND ENGINEERING

MARCH 2021

DECLARATION

We, Vidhi Agrawal, Darpan Bendre, Sadhika Bhatia, Darshan Daliya, Roll No. B003, B008, B010, B018 B.Tech (Computer Engineering), VIII semester understand that plagiarism is defined as anyone or combination of the following:

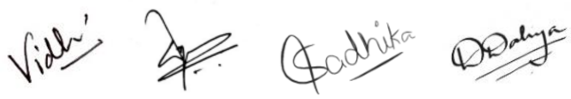
1. Un-credited verbatim copying of individual sentences, paragraphs or illustration (such as graphs, diagrams, etc.) from any source, published or unpublished, including the internet.

2. Un-credited improper paraphrasing of pages paragraphs (changing a few words phrases, or rearranging the original sentence order)

3. Credited verbatim copying of a major portion of a paper (or thesis chapter) without clear delineation of who did wrote what. (Source:IEEE, The institute, Dec. 2004)

4. I have made sure that all the ideas, expressions, graphs, diagrams, etc., that are not a result of my work, are properly credited. Long phrases or sentences that had to be used verbatim from published literature have been clearly identified using quotation marks.

5. I affirm that no portion of my work can be considered as plagiarism and I take full responsibility if such a complaint occurs. I understand fully well that the guide of the seminar/ project report may not be in a position to check for the possibility of such incidences of plagiarism in this body of work.

Signature of the Student: 

Name: Vidhi Agrawal, Darpan Bendre, Sadhika Bhatia, Darshan Daliya

Roll No. B003, B008, B010, B018

Place: Mumbai

Date: 27-10-2020

CERTIFICATE

This is to certify that the project entitled "AgriSense" is the bonafide work carried out by Vidhi Agrawal (B003), Darpan Bendre (B008), Sadhika Bhatia (B010), Darshan Daliya (B018) of B.Tech (Computer Engineering), MPSTME (NMIMS), Mumbai, during the VIII semester of the academic year 2020-2021, in partial fulfillment of the requirements for the award of the Degree of Bachelors of Engineering as per the norms prescribed by NMIMS. The project work has been assessed and found to be satisfactory.

Prof. Kamal Mistry

Internal Mentor

Examiner 1

Examiner 2

Dean

ABSTRACT

Agriculture alone contributes to over 17% of the Indian economy yet, it is one field in which technology is used the least. Farmers have been growing the same crops in the same regions for decades. In regions where the conditions are traditionally suited to grow high profit cash crops farmers have been flourishing but this can only be said for a very small number of farmers. In most regions farming is not profitable and is full of uncertainties of the market. Whilst farmers know that historically a certain crop is suited for the region but the information, they lack is that over the years due to the changes in temperature, moisture and elevation of the land there is a chance that a certain highly profitable crop which has never been grown in a particular region can be grown on their land. We are developing a Region Compatibility based Crop Recommendation system using Remote Sensing and Market Analysis to give the farmers all the information they require to find out if a particular cash crop can be grown on a particular piece of land. We will be studying data about a particular region, namely its important factors like Precipitation, Land Surface Temperature and elevation and matching that area with a similar area where a highly profitable crop is grown. We will be extracting this data using remote sensing. We will finally give a comparison in percentage to show how similar the two regions are, this will help the farmers decide if they should be planting a certain crop. We will also give a market analysis to show how much the farmer stand to earning he chooses to grow a certain crop. We are bridging the gap between the farmers and technology.

TABLE OF CONTENTS

List of Tables

List of Figures

Abbreviations

Abbreviation	Description
GEE	Google Earth Engine
LST	Land Surface Temperature
DSM	Digital Soil Mapping
EVI	Enhanced Vegetation Index
VOD	Vegetation Optical Depth
MODIS	Moderate Resolution Imaging Spectro Diameter
SMAP	Soil Moil Moisture Active Passive
SAR	Synthetic Aperture Radar
PA	Precision Agriculture
NPK	Nitrogen Phosphorus Potassium
CHIRPS Pentad	Climate Hazard Group Infrared Precipitation with Station Data
SRTM	Shutter Radar Topography Mission
ML	Machine Learning
ROI	Region of Interest

Chapter 1
Introduction

1.1 Motivation

Agriculture remains to be a great player in the generation of revenue and a source of food for many people all over the world, especially India. Over 70 per cent of the rural households depend on agriculture. Agriculture contributes about 17% to the total GDP and provides employment to over 60% of the population. Over the past years, this sector has seen a lot of changes and advancement in the different farming approaches and techniques. For example, nowadays, there is the use of organic fertilizer, the consumption of reduced amounts of pesticides, the use of different tractors and machinery. The availability of such inputs has seen the need for the use of natural resources and process with aim of improving agricultural output and reducing costs. Even with the economy depending so much on Agriculture it is one of the fields in which the use of technology is negligible. The use of modern technology in agriculture comes with a lot of benefits. They don't have knowledge about other cash crops that can be grown as they don't have enough information related to it.

The lack of information accessibility and knowledge that farmers have pertaining to the kind of crops that lead to a higher yield and more profitability. Over generations farmers belonging to a particular region continue to grow the same crops over and over again. There are two reasons for this.

1. They don't have knowledge about other cash crops that can be grown as they don't have enough information related to it.
2. They aren't sure of the price they will be getting and have no idea if it will be more profitable to grow the alternate crops and hence aren't willing to change as they think it is a big risk.

Due to these lakhs of farmers are struggling to make ends meet. Agriculture being the backbone of the Indian Economy, there is a need to empower our farmers with information obtained through technology.

Thus, to address this problem faced by farmers will be developing "Region Compatibility based Crop Recommendation system using Remote Sensing and Market Analysis" through which where farmers will be able to deduce the possibility of growing alternate crops based on satellite imaging and also know the increase in profit that they can generate based on market price prediction. We will be using Google Earth Engine to extract data, Power BI to visualize our results, MS Excel to compile data and Python for coding.

We have done a review of literature through which we researched and studied the different cash crops that are grown in India and the conditions required to grow them. We then selected the crops which will give the highest profitability based on historic data collected through various reputable sources. We have then implemented the system by running this data against the specific regions chosen by the user i.e. a farmer in this case to give a percentage match.

1.2 Problem Definition

"Region Compatibility based Crop Recommendation system using Remote Sensing and Market Analysis"

This is a platform which where farmers will be able to deduce the possibility of growing alternate crops based on satellite imaging and also know the increase in profit that they can generate based on market price prediction. Region compatibility as the name suggest will help farmers decide if a particular land is suitable to grow an alternate crop. It will predict this based on the data we extract from google earth engine. We will be studying historical data (12 years) of crops and the optimum conditions that they have been growing in for years. We will then use remote sensing to map these particular regions and then compare how similar they are to a region on which a farmer is wanting to grow that crop. We will also be basing the recommendation based on the higher profit that he stands to gain. We will be doing a thorough price analysis to make sure that the farmers make a well-informed decision. We will be showing a percentage comparison where that particular crop is growing and how similar the land being compared too is in terms of elevation, temperature, precipitation and soil quality

We will be using Google Earth Engine to extract data, Power BI to visualise our results, MS Excel to compile data and Python for coding.

1.3 Project Scope

1.3.1 Farmer Awareness

Farmers don't have enough knowledge about ways in which they can maximize output from their fields. They also don't have enough information as to which alternate crops can grow in their region. There is always a possibility that a particular crop could be grown in a region which would give a higher yield and would also be more profitable.
Using satellite imaging our platform will help them study their geographical region, compare it with alternate regions and find out how feasible it is to grow another crop.

1.3.2 Comparison

Even if a farmer is told that a particular crop can be grown in a particular region there is a chance that the crop fails. It is a big risk which a farmer can't afford to take. Using our platform a farmer will be able to able to compare another location/region where a particular crop is growing and the factors such as soil and weather conditions of both regions will be compared to give an exact percentage showing how similar the two regions are. This process will be done for all cash crops. So a farmer can compare crops grow in different regions and he will know which one has the least chance of failure.

1.3.3 Market Prediction

The main concern for a farmer remains if he will receive a fair price for the produce grown by him. Through our platform he will be able to know the predicted prices of the market for the alternate crop being grown depending on the month of harvest. We will be analyzing the data over the last 5 years to predict trends for the same.

We have done a review of literature through which we researched and studied the different cash crops that are grown in India and the conditions required to grow them. We then selected the crops which will give the highest profitability based on historic data collected through various reputable sources. We have then implemented the system by running this data against the specific regions chosen by the user i.e a farmer in this case to give a percentage match.

Chapter 2
Review of Literature

More and more researchers have begun to identify the problem in Indian agriculture and are increasingly dedicating their time and efforts to help alleviate the issue. Our main target was to understand the working of the google earth engine and find out how well it can be used to extract relevant data.

Will researching we came across **[1]** in which comparison of pixel-based approaches to crop mapping in Ukraine was done and explored efficiency of the Google Earth Engine (GEE)cloud platform for solving “Big Data “problem and providing high resolution crop classification map for large territory. We found that Google Earth Engine (GEE) provided very good performance in enabling access to remote sensing products through the cloud platform they went hand in hand with **[3]** which the various applications of Google earth engine in the Agriculture sector were explored. Various aspects such as vegetation mapping and monitoring, landcover mapping and other remote sensing agricultural applications along with a little insight into the different satellites that we can refer to calculate the stated aspects. When we looked more into the various ways in which remote sensing can be used to measure vegetation depth we came across **[2]** The proposed combination of multisensor (optical and microwave) remote sensing data for crop yield estimation and forecasting used two novel approaches. We the lag between Enhanced Vegetation Index (EVI) derived from MODIS and Vegetation Optical Depth (VOD) derived from SMAP was used as a new joint metric combining the information from the two satellite sensors in a unique feature or descriptor. The second approach used avoided summarizing statistics and uses machine learning to combine full time series of EVI and VOD. Results confirmed the value of using both EVI and VOD at the same time, and the advantage of using automatic machine learning models for crop yield/production estimation. In line with this in **[4]** we studied that nine spectral and temperature bands from relatively low-resolution satellite images were used as their feature for predicting county-level corn and soybean yields in US and they demonstrated that it is possible to achieve better crop yield prediction accuracy using MODIS satellite imagery by employing more complex models.
This gave us a sense of direction as to the usefulness of remote sensing.

We had gained information regarding how remote sensing can be used for various purposes but to study how to actually map the crops based on multiple sources or satellites we looked into [5].

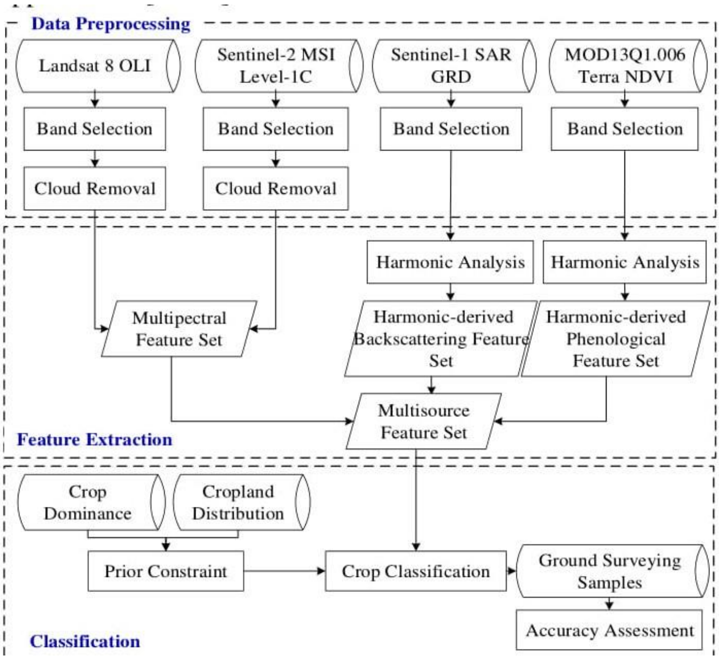


Fig 2.1: Flow diagram for large scale crop mapping [1]

The study proposed a method for large scale crop mapping based on multi-source remote sensing images. To be specific,(1) harmonic analysis was conducted on NDVI time-series derived from MODIS images and SAR backscattering coefficient time-series derived from Sentinel-1 data, respectively extracting harmonic-derived phenological features and harmonic-derived backscattering features, and then combined with spectral features from Landsat-8 and Sentinel-2 images to construct the final multi- source feature set for crop classification; (2) it employed prior constraints of crop dominance and cropland distribution to reduce misclassifications in large scale crop mapping; and (3) the whole process was conducted on the Google Earth Engine (GEE) online platform. This was further explored in **[7]** in which an approach was presented demonstrating the use of GEE in rapid prototyping and testing of crop mapping and monitoring applications which have the potential to scale up global activities in this domain to the 10-30 m resolution range. This is likely to have a crucial impact on our capacity to enumerate crop production statistics that are relevant in our food security monitoring work. To further explore how nutrients can be mapped using remote sensing to get more details in **[8]** a fine DSM method of soil nutrient content using high resolution remote sensing images and multi-scale auxiliary data for PA application. We designed different automatic extraction methods based on high resolution remote sensing images for agricultural production units in plains and mountainous areas. Finally, machine learning methods were used to map the spatial distribution of soil nutrients. Based on this we understood that when we can to potentially map the crops to certain region soil plays a very important role in that aspect. To explore Mapping Cropping Practices on a National we assessed the use of annual, quarterly, and eight-day temporal features for mapping five cropping practices on annual croplands across Turkey in **[10]**. Five cropping practices: Spring and winter cropping, summer cropping, semi-aquatic cropping, double cropping, and greenhouse cultivation were studied. Our study presents an open and readily available framework for detailed cropland mapping over large areas, which bears the potential to inform assessments of land use intensity, as well as land and water resource demands. This helped us in the mapping of regions and extraction of data.

Recommendation was going to play a key part of what we are doing. To understand the recommendation systems that have been used, we studied **[6]**. Where development of an ontology-based recommendation system for crop suitability and fertilizers recommendation. The system predicts suitable crop for the field under consideration based on region in Maharashtra state of India and type of soil. It provides proper recommendation of fertilizers to the farmers. Fertilizer recommendation is done based on nitrogen, phosphorus, and potassium (NPK) contents of soil. Along with fertilizer recommendation system also provides suggestions about crop suitability region. Recommendation system uses random forest algorithm and k-means clustering algorithm. To further understand the use of Machine learning algorithms we looked into **[11]** in which an intelligent system was proposed, called AgroConsultant, which intends to assist the Indian farmers in making an informed decision about which crop to grow depending on the sowing season, his farm's geographical location, soil characteristics as well as environmental factors such as temperature and rainfall. This helped us narrow down the algorithms to be used.

Through this we gained through knowledge about the use of Google Earth Engine, how mapping could be done efficiently and effectively using remote sensing. We further gained an understanding of the Machine learning algorithms used to develop a robust recommendation system. The next task was to study which factors affected cash crops and to know which regions they grow in.

We did a comprehensive study on crops to understand the growing conditions of various crops and then finalize our target crops. We studied various factors like temperature, rainfall, altitude, harvest time and elevation for crops like Mushroom, Avocado, Coffee, Tea, Jute, Rapeseed, Rubber, etc. We also analysed the potential target regions they grow in. Our findings have been summarized in the table given below:

Table 2.1: Crop Based Research

Crop/Factor	Temp (ºC)	Soil Type	Rainfall/Humidity	Elevation	Time	Region
Mushroom	25-30	Clayey and Loamy	55-70%	900m	14 weeks	Orissa, Karnataka, Maharashtra, North East states
Cashew	25-49	Laterite, Red and Coastal Sandy	50cm-250cm	700m	18 weeks	Gujarat, Maharashtra, Karnataka, Tamil Nadu, Karnataka
Rapeseed	10-30	Clayey and Loamy	350-550mm	650m	24 weeks	Rajasthan, Gujarat, West Bengal
Jute	24-37	Alluvial	1000mm	500-1800m	24 weeks	Bihar, Orissa, Andhra Pradesh, Meghalaya
Avocado	12-30	Range of Soils	2000mm	700-900m	5-6 years	Karnataka, Kerala, Tamil Nadu
Rubber	20-35	Range of soil	2000mm	600m	6-7 years	Kerala, North East states
Tea	20-30	Well drained, deep, friable loams	1500mm	2000m	8 months	North East, Kerala, Uttarakhand

Coffee	17-25	Rich moist, well drained, loamy	1000-2000mm	200-300m-Robusta 1000-2000-Arabica	75 days to sprout, 2 years for final use	Karnataka, Kerala, Tamil Nadu, Orissa, Andhra Pradesh, Telangana, North East

Table 2.2: Algorithm Based Research

Paper	**Input**	**Output**	**Algorithm**	**Accuracy/Error Rate**
AgroConsultant	Soil Type, Aquifer, Soil pH, Top soil Characteristics, Season	Recommends a cash crop and a cereal	• Decision Tree • KNN • Random Forest • Neural networks	90.20% 89.78% 90.43% 91%
Synergistic integration of optical and microwave satellite data for crop yield estimation	Growing and senescence stages of crops: corn, soybean and wheat.	Estimate the county-level surveyed total production, as well as individual yields of the major crops grown in the region	• Regularized Linear Regression(RLR) • Kernel Ridge Regression(KRR)	61% 71%
Understanding Satellite-	Multispectral Images	Corn and Soyabean	CNN	5.24 RMSE

Imagery-Based Crop Yield Predictions				
Crop Suitability and Fertilizers Recommendation Using Data Mining Techniques	• District, State, Season • Soil Characteristics	Crop Suitability Fertilizer Recommendation	• Random Forest • K means	68%
Large Scale Crop Mapping from Multi-Source Remote Sensing Images in Google Earth Engine	Multispectral Images	Growth Trends of 3 crops in a particular region	CART- Classification and Regression Tree	84.25%
Evaluation of Arable Land Yield Potential Through Remote Sensing Monitoring	Multi-temporal remote sensing images of Landsat5 TM and BJ-1	wheat yield potential for cropland of Beijing suburb. understand the wheat medium- and low-yield fields of Beijing area	Principal Component Analysis (PCA) and Multi-linear regression (MLR) analysis	RMSE < 517 kg/ha
Fine mapping of key soil nutrient content using high	high resolution remote sensing images and multi-scale auxiliary data for PA	DSM method of soil nutrient content	Random Forest (Grid based and Oarcel based)	RMSE between 0.18 to 45.39

resolution remote sensing image to support precision agriculture in Northwest China	application.			
Using Google's cloud-based platform for digital soil mapping	29,784 soil profiles from the NASIS database (USDA–NRCS, 2014) which contains laboratory measurements of topsoil organic carbon content	prediction of a categorical property (soil class) based on a classification algorithm, and prediction of a continuous property (soil organic carbon content) via a regression technique.	Serial Classifier algorithm (similar to random forests)	84.25%

Chapter 3

Analysis & Design

3.1 Architecture

The architecture of the system is divided into two section the client side and the server-side. On the client-side the user interacts with the web interface whereas on the server-side actual computations are done.

Client-side: Input parameters are taken from the user and corresponding output for compatibility percentage and price analysis is displayed on the web interface of the system.

Server-side: From the client-side request is sent to the Webapp Server and feature extraction process is done with the help of the ML Model and a compatibility percentage along with price analysis is calculated. The calculated result is then passed to the web interface where it is displayed to the user as output.

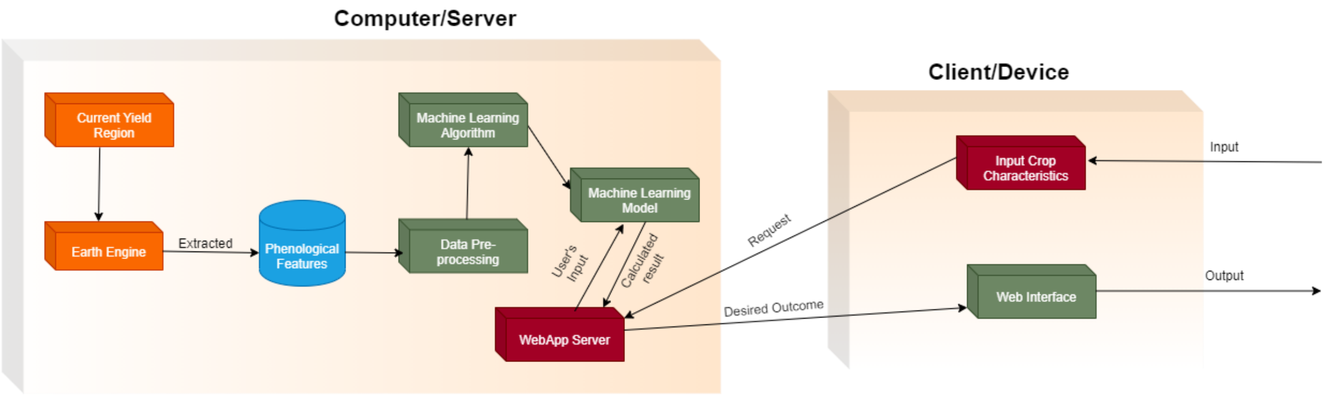


Fig 3.1: Architecture Diagram of the Proposed System

3.2 Use Case Diagram

The use case is essentially a primary example of how the proposed software application or system is meant to be used, from the user's point of view. A use case diagram will typically show system 'actors' (humans or other entities external to the system) and how they interact with the system. Technically, each action such a system actor can perform with the application or system is considered to be a separate use case. In our model we have three actors:

i)User-Client

ii)Administrator (has unrestricted access)

iii)Crop wise parameter information (derived database)

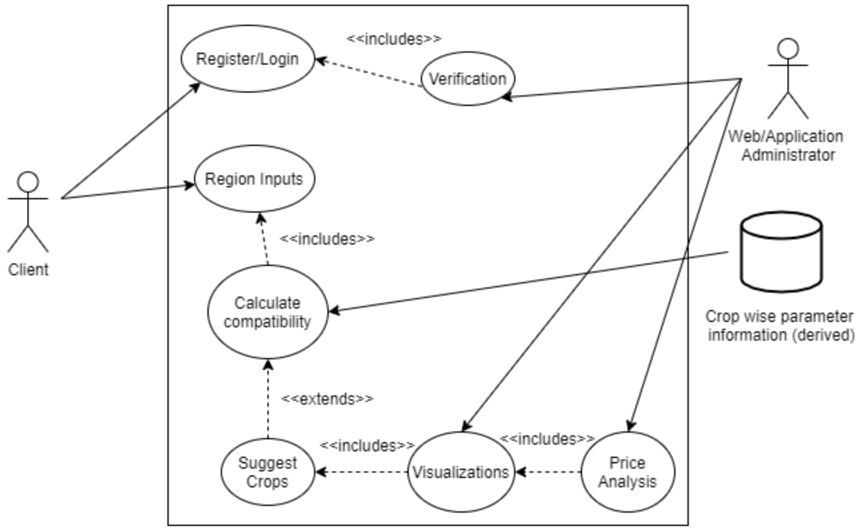


Fig 3.2.1: Use Case Diagram

Three main actors Client, Web/Application Administrator and Crop wise parameter information (derived dataset). The Admin has unrestricted access whereas the client's access is limited to front-end interaction with the system as shown above in the figure.

Once the Client is registered or verified in case of existing user, he/she can proceed to the maps interface where he/she can input the required co-ordinates. Once the co-ordinates are validated and submitted the backend computation is done. A compatibility percentage along with the price analysis is shown to the user on the frontend.

3.3 Sequence Diagram

Sequence Diagrams are interaction diagrams that detail how operations are carried out. They capture the interaction between objects in the context of a collaboration. Sequence Diagrams are time focus and they show the order of the interaction visually by using the vertical axis of the diagram to represent time what messages are sent and when. Below is the Sequence Diagram for our system.

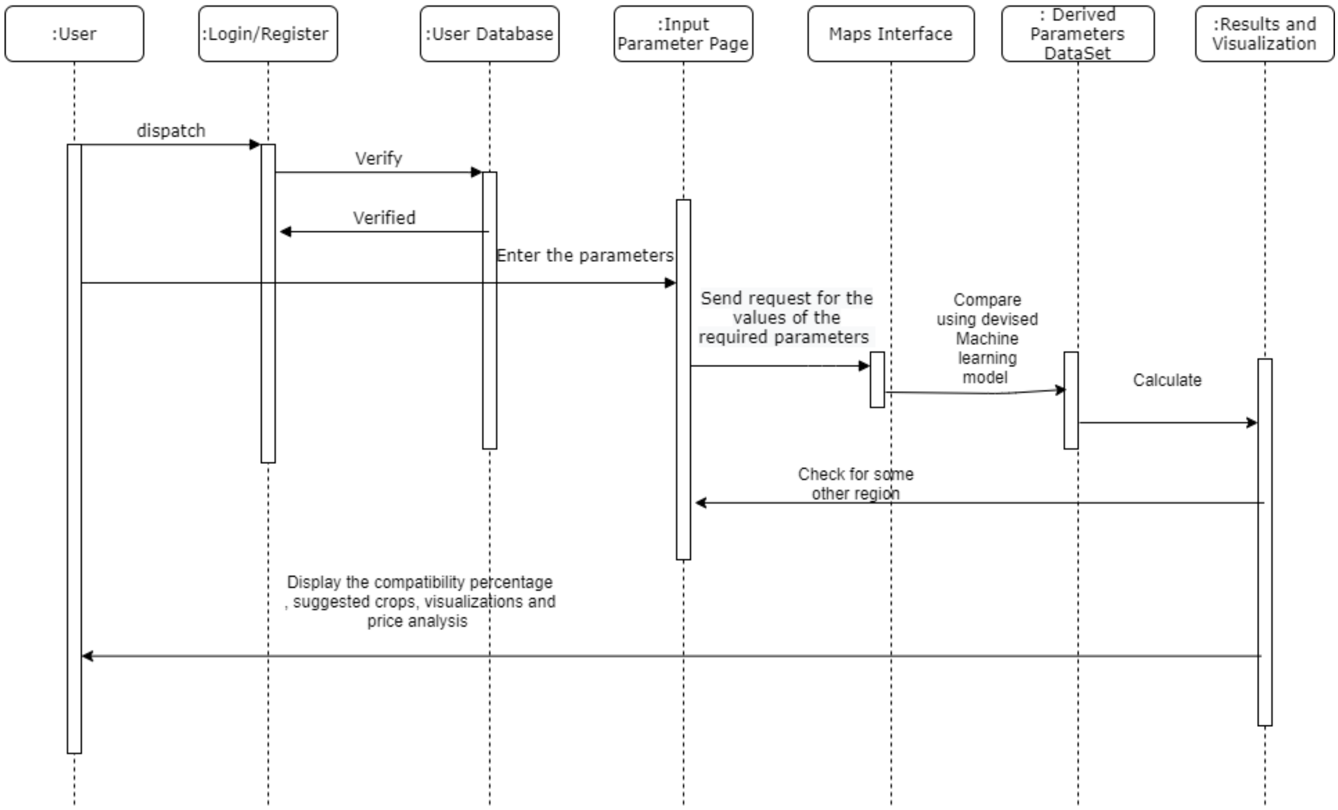


Fig 3.2.2: Sequence Diagram

The above diagram shows sequential interaction at every stage within the system. The arrows represent the flow of data amongst the different modules within the system, it shows what input is given to the module and what output is generated accordingly.

Chapter 4
Implementation

4.1 Environment

4.1.1 Google Earth Engine:

Google Earth Engine is a planetary-scale platform for Earth science data & analysis. It combines a multi-petabyte catalogue of satellite imagery and geospatial datasets with planetary-scale analysis capabilities and makes it available for scientists, researchers, and developers to detect changes, map trends, and quantify differences on the Earth's surface.

4.1.2 Jupyter Notebook:

The *Jupyter Notebook* is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and much more.

4.1.3 Visual Studio Code:

Visual Studio Code is a free source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

4.1.4 Website Development Tool

4.1.5 Power BI

Power BI is a business analytics service by Microsoft. It aims to provide interactive visualizations and business intelligence capabilities with an interface simple enough for end users to create their own reports and dashboards.

4.2 Implementation Approach

GEE enabled us to get a global-scale insight which helped us to prepare our database by remote sensing. We used interactive time-lapse viewer to travel back in time and analyse the change in environmental patterns over the past years.

After a rigorous literature survey, we finalised on the following:

Parameters

1. Land Surface Temperature
2. Precipitation
3. Elevation

Crops

1. Coffee
2. Avocado
3. Mulberry

Districts

1. Chikmagaluru (Karnataka)
2. Madikeri (Karnataka)
3. Nilgiris (Tamil Nadu)
4. Tirap (Arunachal Pradesh)
5. Changlang (Arunachal Pradesh)
6. Wayanad (Kerala)
7. Munnar (Sikkim)
8. Hassan (Karnataka)
9. Yercaud (Karnataka)
10. Kodaikanal (Tamil Nadu)

Duration

12 Years (2007-2019)

4.2.1 Importing Rasters

After selecting the factors, we want to analyse (precipitation, land-surface temperature and elevation) our first step was to select the target satellites for the remote sensing of this data.

The three rasters that we are working with:

1. **CHIRPS Pentad: Climate Hazards Group InfraRed Precipitation with Station Data (version 2.0 final)-** Analyzing long term trends in precipitation data for a specific region of interest (ROI)

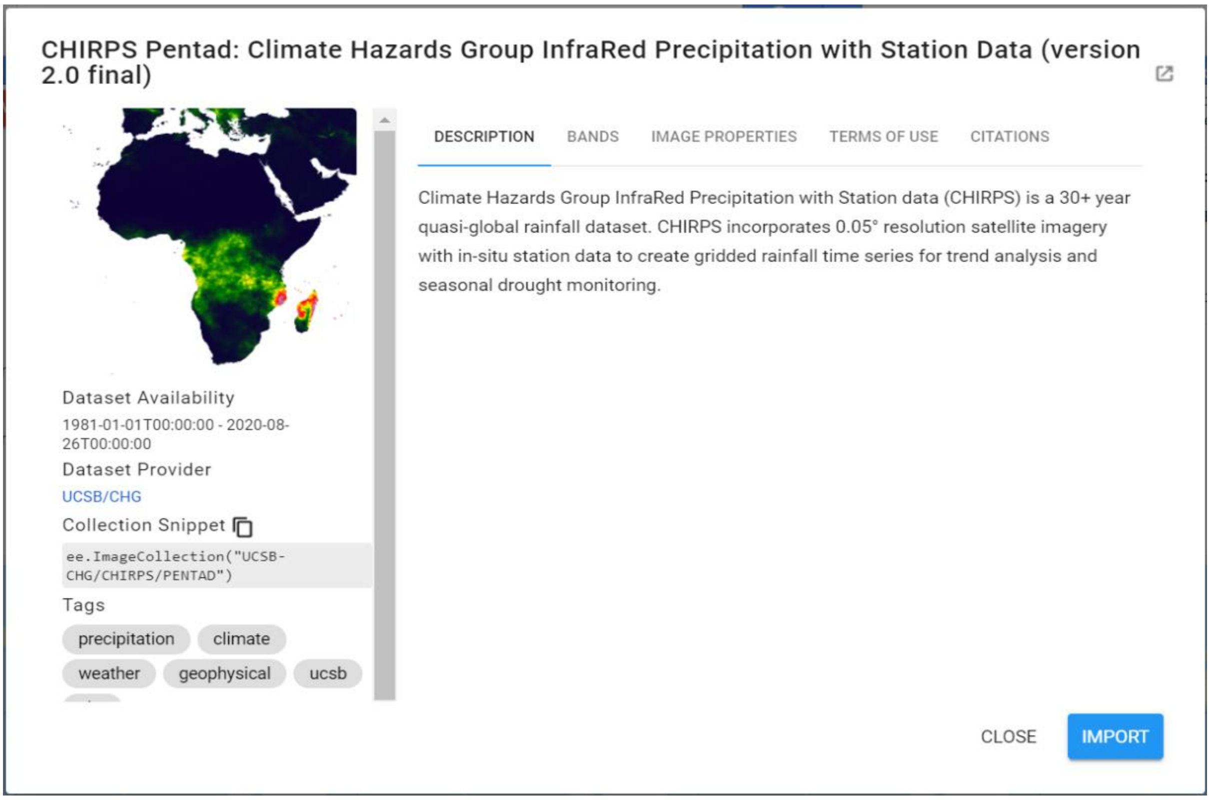


Fig 4.2.1.1: Precipitation Raster

2. **MOD11A2.006 Terra Land Surface Temperature and Emissivity 8-Day Global 1km-** Analyzing long term trends in land surface temperature data for a specific region of interest (ROI)

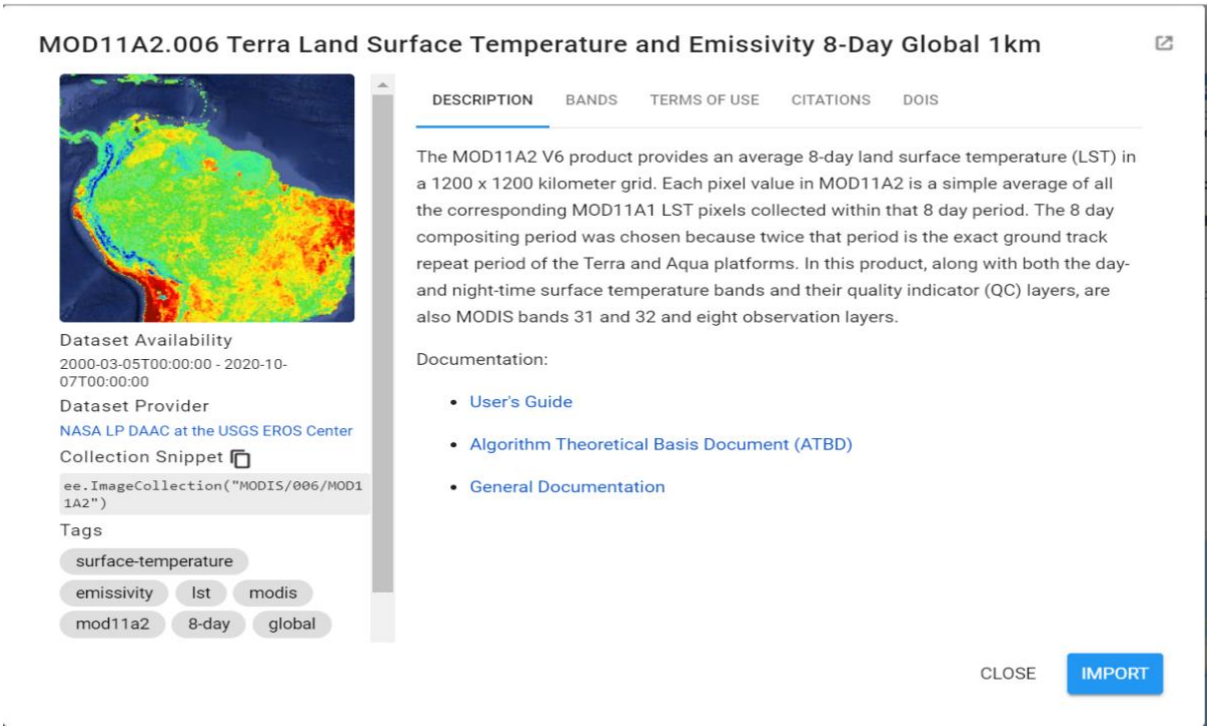


Fig 4.2.1.2: LST Raster

3. **NASA SRTM Digital Elevation 30m-** Analyzing elevation data for a specific region of interest (ROI)

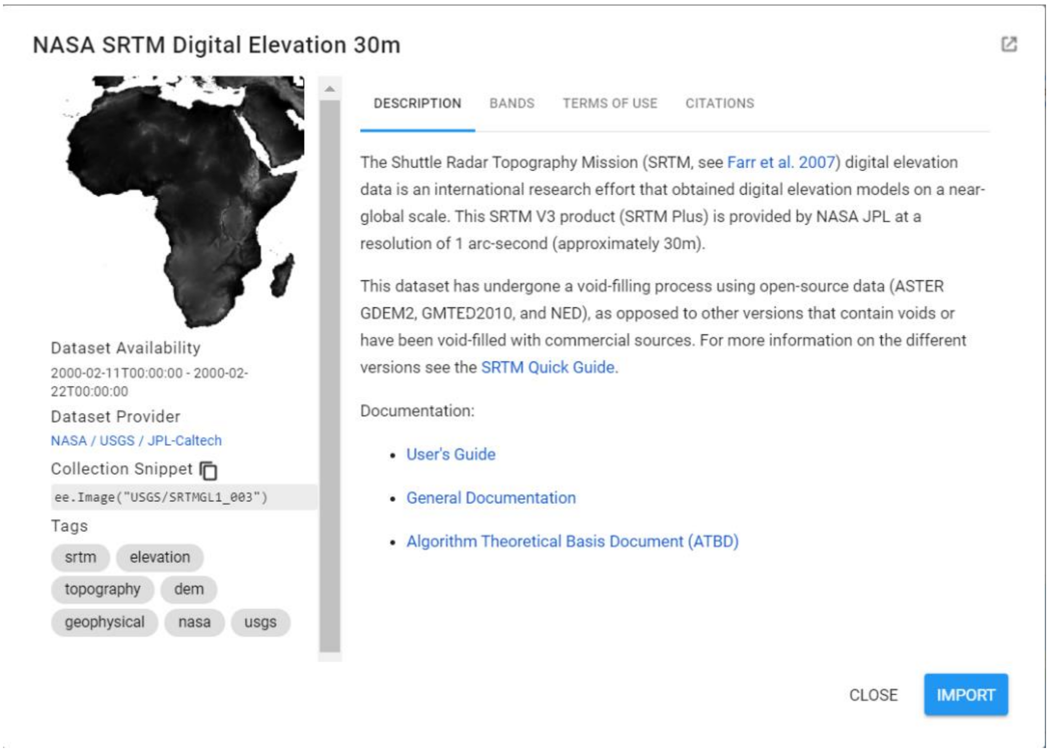


Fig 4.2.1.3: Elevation Raster

4.2.2 Defining ROI

Our next step is to select a particular region of interest for which we want to gather the data. We have analysed over 7 districts for each crop over a span of 12 years. ROI is defined by making a polygon from geometry imports option as shown in the image below.

Fig 4.2.2: ROI Selection

4.2.3 Code to extract Data

Once we have selected the desired ROI, we need to extract data (precipitation, land-surface temperature and elevation) for that particular region. This is achieved by writing specific code for the particular parameter in the script box of GEE using either JavaScript or Python.

Code for different parameters is as follows:

1. **Land-Surface Temperature:** Through this code snippet we are extracting the land-surface temperature in degree Celsius for the specified ROI for a span of 12 years. We have plotted a time-series chart for the same to understand it better.

```
Imports (4 entries)
  var CHIRPS: ImageCollection "CHIRPS Pentad: Climate Hazards Group InfraRed Precipitation wi…
  var modis: ImageCollection "MOD11A2.006 Terra Land Surface Temperature and Emissivity 8-Day…
  var srtm: Image "NASA SRTM Digital Elevation 30m" (1 band)
  var roi: Polygon, 5 vertices

// Import image collection


// A start date is defined and the end date is determined by advancing 1 year from the start date.
var start = ee.Date('2007-01-01');
var dateRange = ee.DateRange(start, start.advance(12, 'year'));

// Filter the LST collection to include only images from time frame and select day time temperature band

var modLSTday = modis.filterDate(dateRange).select('LST_Day_1km');

// Scale to Kelvin and convert to Celsius, set image acquisition time.
var modC = modLSTday.map(function(image) {
  return image
    .multiply(0.02)
    .subtract(273.15)
    .copyProperties(image, ['system:time_start']);
});
```

```
// Chart the time-series
var temp_trend = ui.Chart.image.series({
  imageCollection: modC,
  region: roi,
  reducer: ee.Reducer.median(),
  scale: 1000,
  xProperty: 'system:time_start'})
  .setOptions({
    lineWidth: 1,
    pointSize: 3,
    trendlines: {0: {
        color: 'CC0000'
     }},
     title: 'LST  Time Series',
     vAxis: {title: 'LST Celsius'}});
print(temp_trend);

//Clip to roi
var LSTclip = modC.mean().clip(roi);

// Add clipped image layer to the map.
Map.addLayer(LSTclip, {
  min: 0, max: 40,
  palette: ['blue', 'limegreen', 'yellow', 'darkorange', 'red']},
  'Mean temperature');
```

Fig 4.2.3.1: Code Snippet LST

2. **Precipitation-** Through this code snippet we are extracting the annual mean precipitation in mm for the specified ROI for a span of 12 years. We have plotted a time-series chart for the same to understand it better.

```
//Define date range of interest
var precip = CHIRPS.filterDate('2007-01-01','2019-12-31');

var TS5 = ui.Chart.image.series(precip, roi, ee.Reducer.mean(),1000, 'system:time_start').se
title: 'Precipitation Full Time Series',
vAxis: {title: 'mm/pentad'}, });
print(TS5);

var precip1year=CHIRPS.filterDate('2018-01-01','2018-12-13');
var TS1 = ui.Chart.image.series(precip1year, roi, ee.Reducer.mean(),1000, 'system:time_start
title: 'Precipitaon 1-Year Time Series',
vAxis: {title: 'mm/pentad'}, });
print(TS1);

var yearPrecip = precip1year.mean().clip(roi);

var meanPrecip = precip.mean().clip(roi);
Map.addLayer(yearPrecip, {min: 0, max: 40,
  palette:['lightblue','blue','darkblue']}, 'Year Precipitation');
Map.addLayer(meanPrecip, {min: 0, max: 40,
  palette:['lightblue','blue','darkblue']}, 'Mean Precipitation');
```

Fig 4.2.3.2: Code Snippet Precipitation

3. **Elevation-** Through this code snippet we are extracting the mean elevation in meters for the specified ROI for a span of 12 years.

```
Map.addLayer(srtm, {min: 0, max: 1000}, 'DEM');
Map.addLayer(srtm, {min: 0, max: 300, palette: ['blue', 'yellow', 'red']}, 'Elevation ab
var hillshade = ee.Terrain.hillshade(srtm);
Map.addLayer(hillshade, {min: 150, max:255}, 'Hillshade');
var high = srtm.gt(200);
Map.addLayer(high, {}, 'Above 200m');
var slope = ee.Terrain.slope(srtm);
Map.addLayer(slope, {min: 0, max: 60}, 'Slope');
Map.addLayer(srtm, {min: 0, max: 1000}, 'DEM');
Map.addLayer(roi);
var dict = srtm.reduceRegion({
        reducer: 'mean',
        geometry: roi,
        scale: 90
});
print('Mean elevation', dict);
```

Fig 4.2.3.3: Code Snippet Elevation

4.2.4 Output

After the successful execution of the codes we generated the following outputs:

1. **Land-surface temperature (LST) time series chart-** The chart shows graphical representation of LST in degree Celsius for the particular ROI over a span of 12 years.

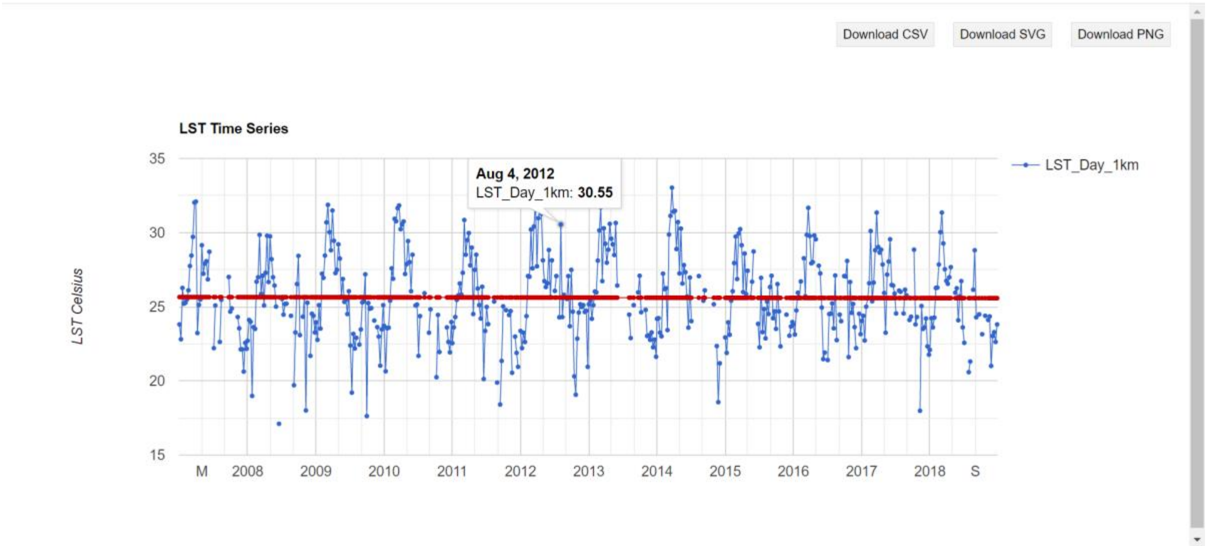


Fig 4.2.4.1: Graphical Representation LST

2. **Precipitation full time series chart-** The chart shows graphical representation of precipitation in mm for the particular ROI over a span of 12 years.

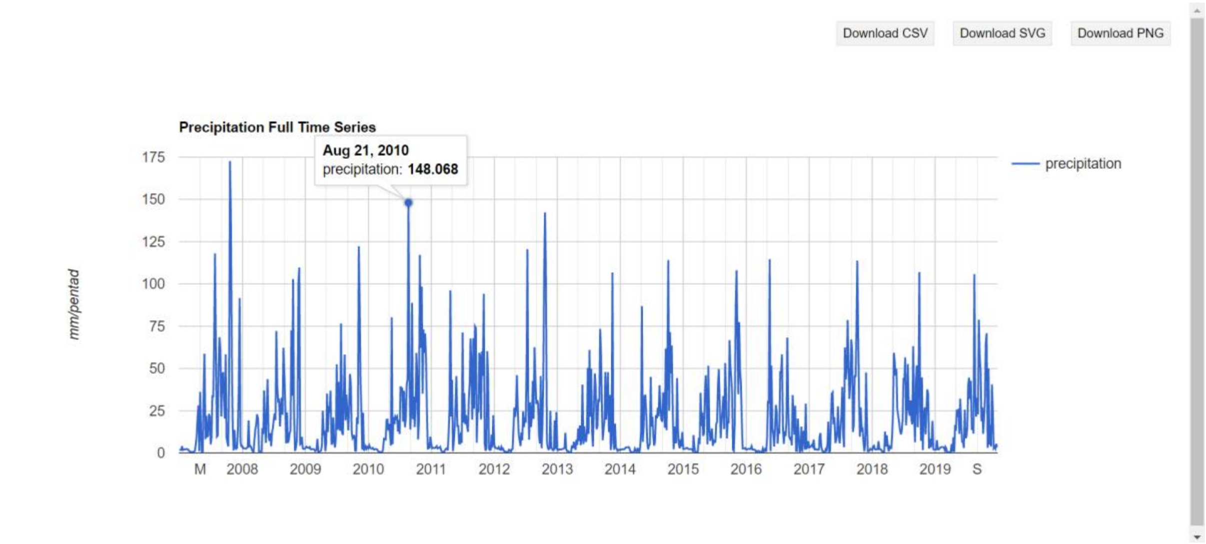


Fig 4.2.4.2: Graphical Representation Precipitation

3. **Elevation-** Mean elevation is obtained in meters for the particular ROI.


```
Mean elevation                    JSON
▾Object (1 property)              JSON
  elevation: 1412.8020456998074
```


Fig 4.2.4.3: Elevation Analysed

4.2.5 Data Collected

The following data is collected:

Temperature

	A	B	C
14	Apr 7, 2014	39.07	
15	Apr 15, 2014	36.61	
16	Apr 23, 2014	33.17	
17	May 1, 2014	34.69	
18	May 9, 2014	34.69	
19	May 17, 2014	33.65	
20	May 25, 2014	34.21	
21	Jun 2, 2014	31.95	
22	Jun 10, 2014	28.45	
23	Jun 18, 2014	27.13	
24	Jun 26, 2014	35.63	
25	Jul 4, 2014		
26	Jul 12, 2014		
27	Jul 20, 2014		
28	Jul 28, 2014	26.05	
29	Aug 5, 2014		
30	Aug 13, 2014		
31	Aug 21, 2014	28.21	
32	Aug 29, 2014		
33	Sep 6, 2014	27.97	
34	Sep 14, 2014	29.81	
35	Sep 22, 2014	30.77	

Hassan_Karnataka_t1

Precipitation

	A	B
21	Apr 6, 2014	8.949
22	Apr 11, 2014	23.702
23	Apr 16, 2014	15.032
24	Apr 21, 2014	2.55
25	Apr 26, 2014	22.051
26	May 1, 2014	27.183
27	May 6, 2014	38.282
28	May 11, 2014	1.179
29	May 16, 2014	20.347
30	May 21, 2014	19.464
31	May 26, 2014	61.447
32	Jun 1, 2014	28.355
33	Jun 6, 2014	12.774
34	Jun 11, 2014	11.601
35	Jun 16, 2014	9.37
36	Jun 21, 2014	6.519
37	Jun 26, 2014	36.495
38	Jul 1, 2014	14.122
39	Jul 6, 2014	16.756
40	Jul 11, 2014	17.043
41	Jul 16, 2014	11.164
42	Jul 21, 2014	17.988

Hassan_Karnataka_p1

Fig 4.2.5: Dataset Generated

Similarly, we have collected the data for all the districts for each crop.

4.3 Machine Learning Model

We plan on designing a Machine Learning Model that will be trained to recognize various patterns over a set of training data, providing it an algorithm that it can use to reason over and learn from the phenological dataset. Once we have trained the model, we can use it to reason over data that it hasn't seen before and make sound predictions generating a compatibility score so that we can achieve our goal of mapping crops that are highly profitable but sparsely grown.

4.3.1 Data Pre-processing/Preparation

Data pre-processing is a data mining technique which is used to transform the raw data in a useful and efficient format. It is mainly done to make the data consistent and ready for computation.

Main steps in data pre-processing are:

1. **Dataset Compilation:** In this step all the individual datasets for precipitation, LST and elevation for all the different districts is compiled together into a single dataset for each crop.
2. **Data Cleaning:** The data can have many irrelevant and missing parts. To handle this part, data cleaning is done. It involves handling of missing data, noisy data etc. In our case we had missing values as shown in the dataset obtained previously. To overcome this, we used the interpolate function of pandas which is a data-centric python package. Interpolate() was used to fill missing values within the data frame.
3. **Data Consistency:** The data being still inconsistent even after cleaning as the revisit period for MOD11A2.006 Terra Land Surface Temperature and Emissivity 8-Day Global 1km is 8 days and CHIRPS Pentad: Climate Hazards Group InfraRed Precipitation with Station Data (version 2.0 final) is 6 days there was a difference in number of values. We used the moving averages technique to smoothen this inconsistency and obtain a consistent dataset.
4. **Creation of Target Variable:** Crop_class target variable has been created which holds values 0/1/2 depending on the crop

Chapter 5
Conclusion & Future Work

5.1 Conclusion

We have implemented the Region-Compatibility based Crop Recommendation System using Remote Sensing and Market Analysis” by studying the past research. Conducted and through this we gained thorough knowledge about the use of Google Earth Engine, how mapping could be done efficiently and effectively using remote sensing. We further gained an understanding of the Machine learning algorithms used to develop a robust recommendation system.

We then did a comprehensive study on crops to understand the growing conditions of various crops to finalize our target crops. We studied various factors like temperature, rainfall, altitude, harvest time and elevation for a range of crops before selecting Coffee, Avocado and Mulberry.

We used Google Earth Engine, Visual Studio, Power BI, Jupyter Notebook to implement our system. GEE enabled us to get a global-scale insight which helped us to prepare our database by remote sensing. We used interactive time-lapse viewer to travel back in time and analyse the change in environmental patterns over the past years.

We then studied the regions in which Coffee, Avocado and Mulberry are grown and how they have reacted to changes in factors such as precipitation, elevation and temperature over 12 years. We selected the target satellites for the remote sensing of this data.

We then selected a particular region of interest for which we wanted to gather the data. Once we selected the desired ROI, we needed to extract data (precipitation, land-surface temperature and elevation) for that particular region. This was done by writing specific code for the particular parameter in the script box of GEE using either JavaScript or Python.

After the successful execution of the codes we generated

1. Land-surface temperature (LST) time series chart
2. Precipitation full time series chart
3. Elevation

We collected the data for all the districts for each crop. We are working on a ML Model to realise our crop compatibility recommender and price analysis system

5.2 Future Work

1. Creating a user-friendly front-end system.
2. Creating a robust ML model
3. We will be expanding the regions to cover the entire country
4. We will be adding data regarding more crops
5. Creating IOS and android app in Native languages of Farmers
6. Using more parameters to get an accurate prediction
7. Using real time data regarding prices from APMC markets
8. Predicting pricing in the future based on historical data

REFERENCES

[1] Liu, X., Zhai, H., Shen, Y., Lou, B., Jiang, C., Li, T., ... & Shen, G. (2020). Large-Scale Crop Mapping From Multisource Remote Sensing Images in Google Earth Engine. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 414-427.

[2] Lemoine, G., & Léo, O. (2015, July). Crop mapping applications at scale: Using Google Earth Engine to enable global crop area and status monitoring using free and open data sources. In 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 1496-1499). IEEE.

[3] Soria-Ruiz, J., & Fernández-Ordoñez, Y. M. (2017, July). Crop discrimination using remote sensing data in a region of high marginalization. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3031-3034). IEEE.

[4] Lemoine, G., & Léo, O. (2015, July). Crop mapping applications at scale: Using Google Earth Engine to enable global crop area and status monitoring using free and open data sources. In 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 1496-1499). IEEE.

[5] Banta, S. J. (1983). Symposium on Potential Productivity of Field Crops under Different Environments (Vol. 15). Int. Rice Res. Inst..

[6] Mutanga, O., & Kumar, L. (2019). Google Earth Engine Applications.

[7] Doshi, Z., Nadkarni, S., Agrawal, R., & Shah, N. (2018, August). AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-6). IEEE.

[8] Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017, July). Large scale crop classification using Google earth engine platform. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3696-3699). IEEE.

[9] https://www.agrifarming.in/cost-of-cultivation-of-avocado-propagation-yield-in-india

[10] Riha, S. J., Wilks, D. S., & Simoens, P. (1996). Impact of temperature and precipitation variability on crop model predictions. Climatic change, 32(3), 293-311.

[11] Lina, L. (2011, March). Rsearch on National Spatial Temporal Changes and Correlation of Temperature and Precipitation Basing on GIS. In 2011 Fourth International Conference on Intelligent Computation Technology and Automation (Vol. 2, pp. 166-169). IEEE.

[12] https://www.agrifarming.in/coffee-farming-project-report-cost-profits

[13] Ferrari, E., Coscarelli, R., & Sirangelo, B. (2018). Correlation analysis of seasonal temperature and precipitation in a region of Southern Italy. Geosciences, 8(5), 160.

Acknowledgement

There are many individuals who have helped us in planning and executing this project by giving us innovative ideas, motivations and opportunities.

We would like to express our gratitude towards our mentor, Professor Kamal Mistry who guided and supported us throughout the course of this project. His invaluable insights and guidance provided us with constant support, motivation and the ability to improve upon the project in every way possible. His push for excellence has helped us achieve the goals and milestones that the project was planned with.

We would also like to extend our gratitude to the Head of Department of Computer Engineering, Dr. Pravin Srinath and Dean of Mukesh Patel School of Technology and Management for providing us with the opportunity and the resources to pursue this project and accomplish our goals.

We would like to thank our project coordinators and faculty members who helped us by scheduling all the project presentations and by providing us with constant feedback which helped us improve our project at every step of the way.

Signature:

Vidhi Agrawal	Darpan Bendre	Sadhika Bhatia	Darshan Daliya
B003	B008	B010	B018