

**IMPACT OF FLOOD ON CROPPING PATTERN AND
PRODUCTION OF MAIZE USING GEOSPATIAL
TECHNOLOGY IN NORTH BIHAR**

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

Geospatial Information Science and Technology

By

Himanshu Kumar

Registration Number 42000592

Supervised by

Dr. Rohan Kumar (25322)

Geospatial Information Science and
Technology (Assistant Professor)
Lovely Professional University, Punjab

Co-Supervised by

Dr. Sujay Dutta

Space Application Centre (Scientist 'SG')
Indian Space Research Organisation
(ISRO), Ahmedabad, Gujarat, India



LOVELY PROFESSIONAL UNIVERSITY, PUNJAB

2024

DECLARATION

I, hereby declared that the presented work in the thesis entitled “**Impact of Flood on Cropping Pattern and Production of Maize using Geospatial Technology in North Bihar**” in fulfilment of degree of **Doctor of Philosophy (Ph.D.) in Geospatial Information Science and Technology** is outcome of research work carried out by me under the supervision and guidance of Dr. Rohan Kumar, working as Assistant Professor in the Geospatial Information Science and Technology of Lovely Professional University, Punjab, India. The research work was co-supervised by Dr. Sujay Dutta, working as Scientist ‘SG’, in the Space Application Centre, Indian Space Research Organisation (ISRO), Ahmedabad, Gujarat, India. In keeping with general practice of reporting scientific observations, due acknowledgements have been made whenever work described here has been based on findings of other investigators. This work has not been submitted in part or full to any other University or Institute for the award of any degree.



(Signature of Scholar)

Name of the scholar: Himanshu Kumar

Registration No.: 42000592

Domain: Geospatial Information Science and Technology

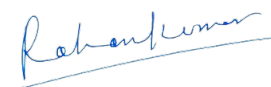
School of Electronics and Electrical Engineering,

Lovely Faculty of Technology and Sciences,

Lovely Professional University, Phagwara (Punjab), India

CERTIFICATE

This is to certify that the work reported in the Ph. D. thesis entitled “**Impact of Flood on Cropping Pattern and Production of Maize using Geospatial Technology in North Bihar**” which is submitted by **Mr. Himanshu Kumar** (Registration No. 42000592) for the award of degree of Doctor of Philosophy in **Geospatial Information Science and Technology** from Lovely Faculty of Technology and Sciences of Lovely Professional University, Punjab, India. The work presented in the thesis is bonafide record of his original work carried out under my supervision and no part of thesis has been submitted for any other degree, diploma or equivalent course.



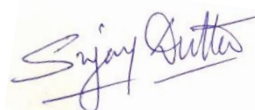
(Signature of Supervisor)

Dr. Rohan Kumar

Assistant Professor,

Geospatial Information Science and
Technology,

Lovely Professional University,
Phagwara, Punjab, India



(Signature of Co-Supervisor)

Dr. Sujay Dutta

Scientist ‘SG’

Space Application Centre (SAC),

Indian Space Research Organisation
(ISRO), Ahmedabad, Gujarat, India



Dedicated
to
My Family Specially
Maa and Papa



ACKNOWLEDGEMENT

First of all, I would like to express my gratitude and sincere thanks to my family for their sincere love, understanding and support to me. This thesis would never have been completed without their patience and support.

Apart from my efforts, the success of this study depended mainly on the encouragement and guidelines of many others. At this juncture, I would like to take an opportunity to thank all those who have directly or indirectly helped me in the completion of my research work.

First and foremost, I would like to express my profound gratitude to my supervisors, **Dr. Rohan Kumar** (Assistant professor, Lovely professional University, Phagwara) and **Dr. Sujay Dutta** (Scientist “SG”, Space Application Centre, ISRO, Ahmedabad) for their precious guidance and productive criticism during the tenure of my research work. Professional expertise, open discussions, endless advice and constant unflinching support rendered by Dr. Rohan Kumar have shaped this thesis to its present form. I'm immensely thankful to him for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. To him, I remain, professionally and emotionally indebted. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Dr. Sujay Dutta is gratefully acknowledged for providing me the necessary research guidance and facilities to complete this work.

I'm grateful to the committee members for sparing their valuable time in monitoring the work progress and providing useful tips, critical comments and precious suggestions from time to time.

I express my special gratitude to **Dr. Magan Singh** (Senior Scientist, ICAR- National Dairy Research Institute, Karnal) for their guidance, encouragement and support. He is a great teacher, educator and advisor.

I am deeply thankful to **Dr. Sateesh Karwariya** for their suggestion, encouragement and moral support during my research work.

I'm thankful to my younger brother **Mr. Virendra Kumar** for helping during my field work.

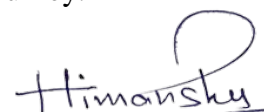
I am fortunate to express my heartfelt gratitude to **Ms. Ankita Agarwal** for her support in formatting and making minor corrections in my thesis.

I would like to thank all the teaching and non-teaching staff members of the School of Electronics and Electrical Engineering, Lovely Professional University for their support during my research work and also for providing the necessary facilities to improve the quality of research work.

I feel lucky to be blessed by the company of wonderful friends who have shared my failures and triumphs. For help at various stages, special thanks go to **Mr. Akhilesh Kumar Yadav**, **Mr. Ranjeet Kumar** and **Mr. Chandan Kumar**.

I thank my parents for their faith in me and for allowing me to be as ambitious as I wanted and shared their support, morally and physically.

Above all, to the Great Almighty, the author of knowledge and wisdom, for His countless love and strength that He bestowed upon me throughout this journey.



(HIMANSHU KUMAR)

Dated:

CONTENTS

	Page No
DECLARATION	I
CERTIFICATE	II
ACKNOWLEDGEMENT	III
CONTENTS	V
LIST OF FIGURES	IX
LIST OF TABLES	XII
LIST OF ABBREVIATIONS	XIII
ABSTRACT	XV
Chapter 1: Introduction	1
1.1 General	1
1.2 Brief description of Flood	2
1.2.1 Global Flood Scenario	2
1.2.2 Flood Scenario of India	7
1.2.3 Flood Scenario of Bihar	9
1.3 Application of Remote Sensing and Geospatial Technology in Flood Management	11
1.4 Study area	14
1.5 Motivation	15
1.6 Research Objectives	16
1.7 Brief Overview of Adopted Methodology	17
1.8 Organization of Thesis	18
Chapter 2: Review of literature	20
2.1 Crop Mapping Method	21
2.1.1 Ground-based surveys	22
2.1.2 Crop inventory data:	22
2.1.3 Remote sensing	22
2.1.3.1 Steps /Overview of crop mapping using remotely sensed satellite data	22
2.1.4 Machine learning	23
2.2 Concluding remarks for best method/ suitable method for crop mapping	24
2.2.1 Use of Optical datasets	25
2.2.2 Use of Synthetic Aperture Radar (SAR) datasets	26

2.2.3	Combining satellite data and Data fusion techniques	26
2.2.3.1	Approaches/methods used for combining multi-sensor datasets	27
2.3	Image Classification	31
2.3.1	Types of Image Classification Techniques	32
2.3.2	Two different ways of Classification	32
2.3.3	Comparison and Applications methods	34
2.3.4	Challenges and Opportunities:	34
2.4	Area/ Acreage Estimation	36
2.5	Crop Yield and Production Estimation	39
2.6	Flood	42
2.6.1	Remote Sensing Technology for flood management	42
2.7	Research Gap	43
 Chapter 3: Data Used and Methodology		45
3.1	Data Used	45
3.1.1	Sentinel-1 SAR	47
3.1.2	Sentinel-2 Optical	48
3.1.3	PLANET-NICFI (PlanetScope)	50
3.1.4	Shuttle Radar Topography Mission (SRTM)	51
3.1.5	Other Datasets	51
3.1.6	Ground Truth (GT) data collection	51
3.1.6.1	Field photographs	53
3.1.7	Crop Cutting Experiment (CCE) Data	56
3.1.7.1	Field Photograph of Crop Cutting Experiment	57
3.2	Software Used:	61
3.2.1	ArcGIS 10.5	62
3.2.2	QGIS (Quantum GIS)	62
3.2.3	Google Earth Engine	62
3.3	General Methodology	63
3.4	Concluding Remarks	66
 Chapter 4: Flood Extent and Affected Area Mapping		67
4.1:	Introduction	67
4.2	Materials and Methods	69
4.2.1	Methodology	70
4.2.1.1	Harmonization of PlanetScope and Sentinel-2A/B MSI	70
4.2.1.2	Google Earth Engine (GEE)	71
4.2.1.3	Random Forest Classifier	73

4.3	Results and Discussion	74
4.3.1	Backscatter intensity of Flooded Pixels	74
4.3.2	Flood Impact Analysis	74
4.3.2.1	Flood Progression evaluation using Sentinel-1 SAR Data	78
4.3.2.2	Pre-Flood Land Use Map	80
4.3.2.3	Flood Progression Assessment (2020)	81
4.3.2.4	Flood Progression Assessment (2021)	85
4.4	Accuracy Assessment and Result Validation	87
4.5	Impact of Flood on Maize	88
4.6	Conclusion	93
 Chapter 5: Maize (<i>Zea Mays L.</i>) Crop Identification/ Discrimination and Acreage Estimation		94
5.1	Introduction	94
5.2	Materials and Method	97
5.2.1	NDVI Generation	98
5.2.2	Methodology	98
5.2.2.1	Google Earth Engine (GEE)	99
5.2.2.2	Machine Learning Algorithm Used for Crop mapping	100
	A. Random Forest	100
	B. Support Vector Machine (SVM)	101
	C. Classification and Regression Tree (CART)	102
5.3	Results and discussion	104
5.3.1	Crop Identification	104
5.3.2	Acreage Estimation of Maize Crop	106
5.3.3	Performance of Machine Classifiers with various datasets	109
5.3.4	Prediction Accuracy	110
5.3.5	Validation	111
5.4	Discussion	112
5.5	Conclusion	114
 Chapter 6: Crop Yield and Production Estimation		115
6.1	Introduction	115
6.2	Materials and Method	118
6.2.1	Experimental Site Selection	118
6.2.2	Methodology	119
6.2.3	Cloud Computing Platform	121
6.2.4	Acreage Estimation	121
6.2.4.1	Multi-date NDVI Extraction	122

6.2.4.2 Use of Random Forest Algorithms for Acreage Estimation	124
6.2.5 Yield Estimation	126
6.2.5.1 Input datasets for Model	126
6.2.5.2 Semi Physical Model for Yield Estimation	127
6.2.5.3 Relationship between maize crop yield/CCE data and indices values	127
6.2.5.4 Production Estimation	128
6.3 Result and Discussion	128
6.3.1 Crop Acreage Estimation	129
6.3.2 Crop Yield Estimation	131
6.3.3 Maize crop production estimation	137
6.3.4 Validation of estimated yield with CCE and Government Report	137
6.4 Conclusion	139
Chapter 7: Summary and Conclusion	141
7.1 Findings	142
7.2 Recommendation and Future perspectives	145
List of Publications	146
List of Conferences	147
Bibliography	148
Appendices	170

LIST OF FIGURES

Figure No.	Title	Page No.
1.1	Number of deaths by flood disaster: 2022 compared to the 2002-2021 annual average	3
1.2	Occurrence by Flood disaster: 2022 compared to the 2002-2021 annual average	3
1.3	Top 5 countries in economic losses and mortality by flood	4
1.4	FAO hunger map	5
1.5	Year wise (2008–2018) total loss in crop and livestock production	6
1.6	Impact of flood on SDG Goal-1&2	7
1.7	Flood prone areas in India	8
1.8	Flood hazard map of Bihar	9
1.9	Location map of study area	14
3.1	Wavelengths and spatial resolutions of Sentinel-2 satellite (MSI instruments)	50
3.2	GT data collection for crop identification	52
3.3	Sample of GT data collection	54
3.4	Field photographs of study area (3A)	55
3.5	Field photographs of study area (3B)	56
3.6	Crop Cutting Experiment	57
3.7	Field Photograph of Crop Cutting Experiment (FP-1)	59
3.8	Field Photograph of Crop Cutting Experiment (FP-2)	60
3.9	The yield of the extracted crop is being weighed to estimate the total yield of the entire area	61
3.10	Overall flowchart of research methodology	66
4.1	Methodology for flood mapping	71
4.2	The Interface of GEE's Cloud Computing Platform	72
4.3	Random forest classifier flowchart for land use map preparation	73
4.4	Backscatter intensity of flooded pixel	76
4.5	District wise flood area of statistics	78

4.6	Backscatter response of VV polarization in 2020 and 2021	79
4.7	Land use map of pre-flood area	80
4.8	Backscatter response of different polarization	81
4.9	Flood progression of 2020	82-83
4.10	Flood affected land use map of 2020	84
4.11	Flood progression of 2021	86
4.12	Flood affected land use map of 2021	87
4.13	Comparison of average yield of flood and non- flooded years	92
5.1	Harmonization, rescaling and fusion of PlanetScope and Sentinel-2A/B MSI data (1B)	97
5.2	NDVI Profile presenting the growth profile of Maize and Wheat.	98
5.3	Methodology for crop acreage estimation	99
5.4	Crop mapping through GEE platform	100
5.5	Random Forest classifier based maize identification	101
5.6	Support Vector Machines Classifier based maize identification	102
5.7	Classification and Regression Tree Classifier based maize identification	103
5.8	Training accuracy in the present study	104
5.9	Maize crop mapping from Sentinel-2A/B, PlanetScope and integration of Sentinel-2A/B and PlanetScope sensors using SVM, CART and RF Classifiers	105
5.10	Maize crop map of 2022	107
5.11	Maize crop map of 2023	107
5.12	Performance of classifiers in the GEE cloud computing platform	109
6.1	Flowchart illustrating the steps of acreage and yield Estimation	120
6.2	The GEE based maize acreage estimation for input parameter for yield estimation	121
6.3	Multi-date stacked NDVI from experiment site	123

6.4	NDVI values are depicting the growth profile of Maize, Wheat, Forest and Water	123
6.5	Flowchart of RF algorithms to estimate acreage for input parameter for yield estimation	125
6.6	Map showing maize crop extent of study area	130
6.7	Sensor wise classification accuracy	130
6.8	District wise maize crop area of study area	131
6.9	Indices-wise relationship between Maize crop yield and Indices values	132
6.10	District wise maize crop yield (Kg/ha ⁻¹) of study area	133
6.11	Choropleth map of maize crop's yield: (A) 2021-22 & (B) 2022-23	134
6.12	District-wise maize crop production	138
6.13	Graph showing variation between actual and predicted yield	139

LIST OF TABLES

Table No.	Title	Page No.
1.1	Flood related damage in Bihar in last 5 decade	10
2.1	The advantages and disadvantages of crop mapping method	21
2.2	Different studies on crop mapping and approach used	29
3.1	Datasets used in the study	46
3.2	Band-wise resolution scale and description	50
4.1	District-wise flood-affected area was extracted using Sentinel-2 and Planetscope data	77
4.2	Districts wise maize crop acreage, yield and production estimation during 2021-22 and 2022-23	89-90
4.3	Districts wise maize crop acreage, yield and production as per GoB Report	91
5.1	District-wise acreage estimation	108
5.2	Performance of classifiers in terms of accuracy	110
5.3	Year-wise available acreage data of rabi/winter maize crop as per the report of Directorate of Economics and Statistics, Government of Bihar.	111
6.1	Districts wise maize crop acreage, yield and production during 2021-22 and 2022-23	135-136

LIST OF ABBREVIATIONS

Abbreviations	Description
AFS	Area Frame Sampling
ANN	Artificial Neural Network
AVHRR	Advanced Very High Resolution Radiometer
AWEI	Automated Water Extraction Index
CART	Classification and Regression Trees
CCE	Crop Cutting Experiment
CCP	Cloud Computing Platforms
CITARS	Crop Identification Technology Assessment for Remote Sensing
CNN	Convolutional Neural Network
CRED	Centre for Research on the Epidemiology of Disasters
DAC&FW	Department of Agriculture, Cooperation and Farmers Welfare
DB	Decibel
DEM	Digital Elevation Model
ESA	European Space Agency
EVI	Enhanced Vegetation Index
EW	Extra Wide Swath
FAO	Food and Agriculture Organization
FAOSTAT	Food and Agriculture Organization Statistics
FASAL	Forecasting Agricultural Output Using Space, Agro-Meteorology and Land-Based Observations
FMIS	Flood Management Information System
FPAR	Fraction of absorbed photosynthetically active radiation
GEE	Google Earth Engine
GHSL	Global Human Settlement Layer
GIS	Geographic Information System
GNDVI	Green Normalized Difference Vegetation Index
GPS	Global Positioning System
GPU	Graphics Processing Unit
GRD	Ground Range Detected
GT	Ground Truth
IRS	Indian Remote Sensing
ISRO	Indian Space Research Organization
IW	Interferometric Wide Swath
LACIE	Large Area Crop Inventory Experiment

LSWI	Land Surface Water Index
LULC	Land Use and Land Cover
ML	Machine Learning
MNDWI	Modified Normalized Difference Water Index
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	Multispectral Imager
MSL	Mean Sea Level
NCRB	National Crime Records Bureau
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NICFI	Norway's International Climate & Forests Initiative
NIR	Near-Infrared
NOAA	National Oceanic and Atmospheric Administration
NRSC	National Remote Sensing Centre
OA	Overall Accuracy
PCA	Principal Component Analysis
REI	Red Edge Index
RF	Random Forest
RFC	Random Forest Classifier
RS	Remote Sensing
RSEO	Remotely Sensed Earth Observation
SAR	Synthetic Aperture Radar
SDG	Sustainable Development Goal
SDMA	State Disaster Management Authority
SLC	Single Look Complex
SM	Strip map
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
TM	Thematic Mapper
TOA	Top-Of-Atmosphere
UAV	Unmanned Aerial Vehicle
UNFPA	United Nations Population Fund Annual
USA	United States Of America
USAID	United States Agency for International Development
USDA	United States Department of Agriculture
USGS	United States Geological Survey
WV	Wave Mode

ABSTRACT

Flood is the major cause of fatalities associated with natural disasters in the world. In India especially in the state of Bihar, where about half of the area (North Bihar) face severe flood frequently due to overflow of major rivers during monsoon. The Kosi River flood is known for its devastation in north Bihar and frequent changes in its course. Therefore, overall objective of this study is to investigate the impact of floods on cropping patterns, production, and productivity in north Bihar using geospatial technology.

Floods in Bihar have a significant impact on the agricultural sector, especially on planting patterns and the state's crop production. The floods affect the cropping patterns in various ways. Firstly, the flood submerges the standing crops in water, causing to damage or complete loss. Secondly, it can delay the sowing season by making fields too wet, thereby shortening the growing season. Farmers may have to wait until the water has drained out from their fields and the soil has dried before planting their crops. Because, waterlogging after floods may create unfavourable conditions for crop growing, leading to changes in cropping patterns. However, flooding has many positive effects on agricultural land such as supplying nutrients needed for plant growth. When floods occur, they bring with them large amounts of sediment that is rich in organic matter and nutrients like nitrogen, phosphorus and potassium, which are essential for plant growth. These nutrients help in improving soil fertility and increasing crop yields (specially for rabi crops) in North Bihar, which is beneficial for agriculture dependent farmers.

Therefore, it is observed that a rapid and robust flood extent demarcation and affected area mapping system is needed for early warning systems and to prioritise the relief, rescue and subsidy. Nowadays, the Remotely Sensed Earth Observation (RSEO) datasets are being commonly used for monitoring and mapping of flood events, which is freely available for researchers. However, the unavailability of resources to download, store and process satellite data is a challenging task for the users. In this context, JavaScript code for processing huge (big data) datasets hosted on cloud computing platforms (CCP) such

as Google Earth Engine (GEE) is edited, tested and developed in the present study. This JavaScript code is capable to robust flood mapping and monitoring using microwave synthetic aperture radar (SAR) remotely sensed satellite datasets at large scale within a short time of period. In this study, we observed that about ~ 12.63% (701967 ha) areas are flooded in 2020. About ~ 17.20% (955897 ha) areas are also flooded in 2021. In the floods of 2021, about 4% more area of North Bihar has been flooded as compared to the floods of 2020.

Apart from this, the agricultural sector is facing multiple challenges that threaten sustainable food production. Factors such as climate change, soil degradation, water scarcity, and the overuse of chemicals and fertilizers are posing a threat to the agriculture sector. Achieving the SDG 2030 goal of Zero Hunger status requires a holistic and adaptive food ecosystem that addresses these challenges. Therefore, there is a need for a system that can enable monitoring of the agricultural sector on a near real-time basis. In this regard, estimation of crop acreage, yield and production are necessary to address these challenges.

Acreage Estimation: Image classification is an essential factor for crop mapping and identification. In recent years, many researchers focused on improving data mining/machine learning algorithms to more accurately deal with image classification and predictive problems. The publicly availability of geospatial datasets and free access to cloud-based geo-computing platforms such as GEE are widely being used for mapping and monitoring of crop phenology, acreage estimation and crop yield forecasting. In the present study, the maize (*Zea mays* L.) crop has been identified and acreage estimated using ground truth (GT)/samples data collected from study area, machine learning algorithms and integrated Sentinel-2A/B and PlanetScope satellite data. In which, the study assessed and compared the performance of Classification and Regression Trees (CART), Support Vector Machine (SVM) and Random Forest (RF) algorithms of Machine Learning (ML) for acreage estimation of maize crops using Google's GEE cloud computing platform. Wherein, we found that RF outperforms CART and SVM algorithms in the GEE platform

with PlanetScope data (90.17 % Overall Accuracy (OA) with Kappa 0.89) and also with the integration of PlanetScope and Sentinel-2A/B data (OA=95.53%, Kappa 0.91). But, CART outperforms RF and SVM algorithms with Sentinel-2A/B data (OA=88.59 %, Kappa 0.85). Based on these approaches, it is found that approximately 321252 hectares of land are devoted to maize cultivation during 2021-22 and about 329504 hectares during 2022-23 in North Bihar, India, which is about 2.5% higher as compared to 2021-22.

Yield Estimation: Timely information about crop conditions is essential for decision-makers to make real-time decisions to maximize yield and production, import and export decisions, ensuring food security, crop insurance, maintaining the development of sustainable agriculture and the prompt construction of appropriate warehouses. Traditional methods of crop yield estimation are time-consuming, costly and have low efficiency. Although various spectral indices were developed using remotely sensed satellite data for mapping and monitoring of vegetation health and productivity such as Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation Index (GNDVI), Enhanced Vegetation Index (EVI), Land Surface Water Index (LSWI). These spectral indices intensely evaluated the relationship with maize crop yield using a regression model and demonstrated the methodology for maize crop yield estimation. The obtained result was also validated with Crop cutting experiment (CCE) data of study area. In the present study, a new index 'REI (Red Edge Index)' has been evolved for the estimation of maize crop yield, which was found to be performing well in terms of prediction accuracy than previously developed indices. Wherein, the obtained correlation with maize yield was $R^2 = 0.84$ with REI, $R^2 = 0.79$ with NDVI, $R^2 = 0.76$, with EVI, $R^2 = 0.70$, with GNDVI and $R^2 = 0.50$ with LSWI.

Here, it was estimated that about 24,90,242 tonnes of maize has been produced in 2021-22 and 22,66,837 tonnes in 2022-23. The total production in 2021-22 was higher by about 9.85% as compared to 2022-23 due to positive post-flood effect on agricultural land (nutritional soil accumulation). The estimated yield was found to be aligned with the maize

crop yield production report of the Bihar state government. However, the accuracy of the methodology can be further validated by using it in different locations and crops with large to small datasets of CCE.

It is expected that this tested algorithms and developed web-based JavaScript code can be applied anywhere in the world for flood extent demarcation, precisely crop mapping and impact of flood on acreage and yield within a short time of period. In addition, crop yield predictive model has been developed for district/field-wise maize crop yield and production estimation, which is simple and cost-effective for developing countries like India and will improve accuracy by using them in different crops and locations with large-scale datasets.

It is crucial for policymakers and farmers to develop strategies to mitigate the floods impact on agriculture and promote crop diversification to reduce the risk of crop loss due to floods. It is also expected that this study will be helpful for crop cultivation management, precision agriculture, crop insurance and for making a decision support system to prioritise the input subsidy for farmers.

1.1 General

Maize is among most vital cereal crops of the world, which is termed the ‘Queen of Cereals’. It has the highest genetic yield potential among the cereals crops, which offers raw material for industrial products, an essential food, and livelihood for most of the population in developing countries. This crop is adaptable to a wide range of elevations, up to 3000 m from mean sea level (MSL) under different agro-climatic conditions (iimr.icar.gov.in; farmer.gov.in). Methods of maize processing and consumption significantly differs geographically. Among the many maize products; maize flour and meal are the mostly consumed products (USAID,2013; Ranum *et al.*, 2014). It is a principal carbohydrate sources of human nourishments in developing nations. In developed world maize is used for animal feed (Ranum *et al.*, 2014). The largest maize production (nearly 35% of the world) is reported in the USA followed by China, Brazil, India, Argentina and Ukraine. Maize is an important cereal crop in India too, which ranks third after rice and wheat.

In India, leading maize producing states are “Karnataka, Madhya Pradesh, Maharashtra, Rajasthan, Bihar, Uttar Pradesh” (Singh *et al.*, 2018). It is a vital crop that engages 15 million Indian farmers in cultivation. Maize crop has potential to double the income of Indian farmers. It has been realised for creating improved income to farmers and enabling profitable employment. India possess great possibility in growth of Maize value chain.

Bihar is the 6th largest maize producer among states in India. Its geographical area is about 93.60 lakh hectares. Total sown area in the state is 78.82 lakh hectares, while net sown area is 57.12 lakh hectares. Bihar is divided by Ganges River into North Bihar (area: 53.3 thousand km²) and South Bihar (area: 40.9 thousand km²). North Bihar is highest producer of maize crop in comparison of South Bihar (Singh *et al.*, 2018). Which gets submerged almost every year by floods. Yet it still positioned itself among top maize producing state. In recent years, Bihar is leading in high-yielding maize crop that envisages high possibility in agricultural force (<http://krishi.bih.nic.in>).

In the present study, the northern part of Bihar (North Bihar) has been chosen as a study area because about 76% population is dependent on agriculture in the region that are severely affected by concurrent flooding (Anonymous, 2020). This study is intended to contribute to the existing knowledge about the impact of floods on agriculture by providing a detailed assessment of the impact of floods on maize crop's production and productivity in North Bihar. Wherein, production focuses on total output of goods during a specific period, productivity emphasizes the efficiency of producing that output. The findings of the study will be of interest to policymakers, farmers and other stakeholders working to reduce the impact of floods on agriculture.

1.2 Brief description of Flood

1.2.1 Global Flood Scenario

According to Centre for Research on the Epidemiology of Disasters (CRED) 2023 report, globally 387 catastrophic events are observed in 2022, which is slightly higher than the average number of catastrophic events between 2002 and 2021(370). Which caused the deaths of 30,704 lives, an economic loss of US\$223.8 billion, and affected 185 million people (Fig. 1.1). Wherein, an average of 176 flood events per year occurred worldwide between 2002 and 2022, causing property damage of approximately 44.9 billion United States dollars (US\$) (Fig. 1.2).

In 2022, a total of 7954 people will die due to flood disasters, which is 53% more than the average from 2002 to 2021(CRED 2023). Pakistan was deeply affected by the floods that occurred between June and September, affecting about 33 million people and resulting in 1,739 lives. The economic damage caused by these floods was a staggering \$15 billion (Fig. 1.3). Similarly, flash floods wreaked havoc in India as well, claiming 2,035 lives and an economic loss of \$4.2 billion. About 7.2 million people were affected by floods in Bangladesh, while an economic loss of \$5 billion was reported in China.

Nigeria also experienced devastating flooding, causing 603 deaths and an economic cost of \$4.2 billion. In South Africa, 544 people lost their lives due to flood-related incidents. In Brazil, floods in February killed 272 people.

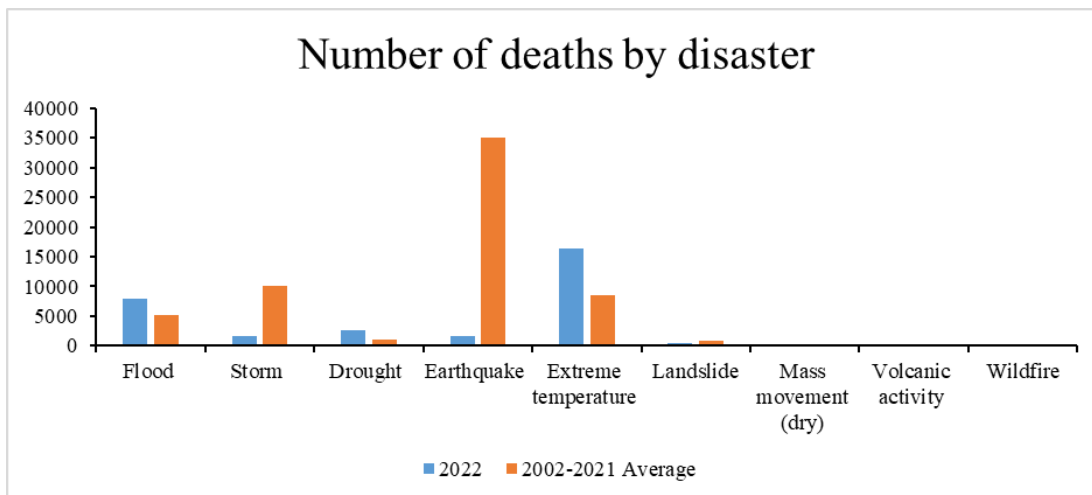


Figure 1.1: Number of deaths by flood disaster: 2022 compared to the 2002-2021 annual average (Source: CRED. 2022 Disasters in numbers. Brussels: CRED; 2023)

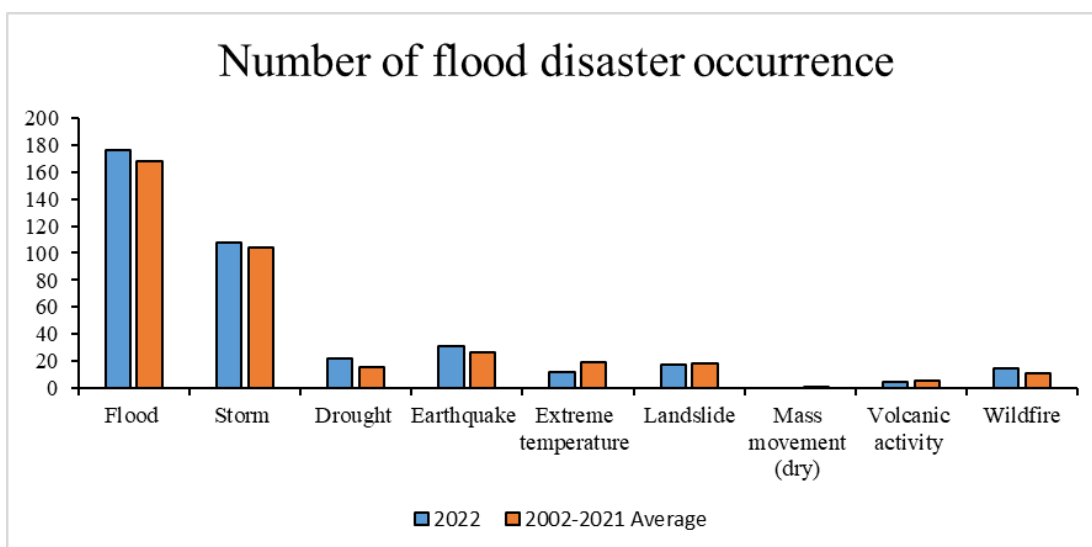


Figure 1.2: Occurrence by Flood disaster: 2022 compared to the 2002-2021 annual average (Source: CRED. 2022 Disasters in numbers. Brussels: CRED; 2023)

In addition, the floods in eastern Australia during February and March caused massive economic damage with an estimated economic cost of A\$6.6 billion. These catastrophic flood events in various regions highlight the urgent need for comprehensive strategies and investments in disaster risk reduction and management to minimize the devastating impacts on human life, the economy and infrastructure.

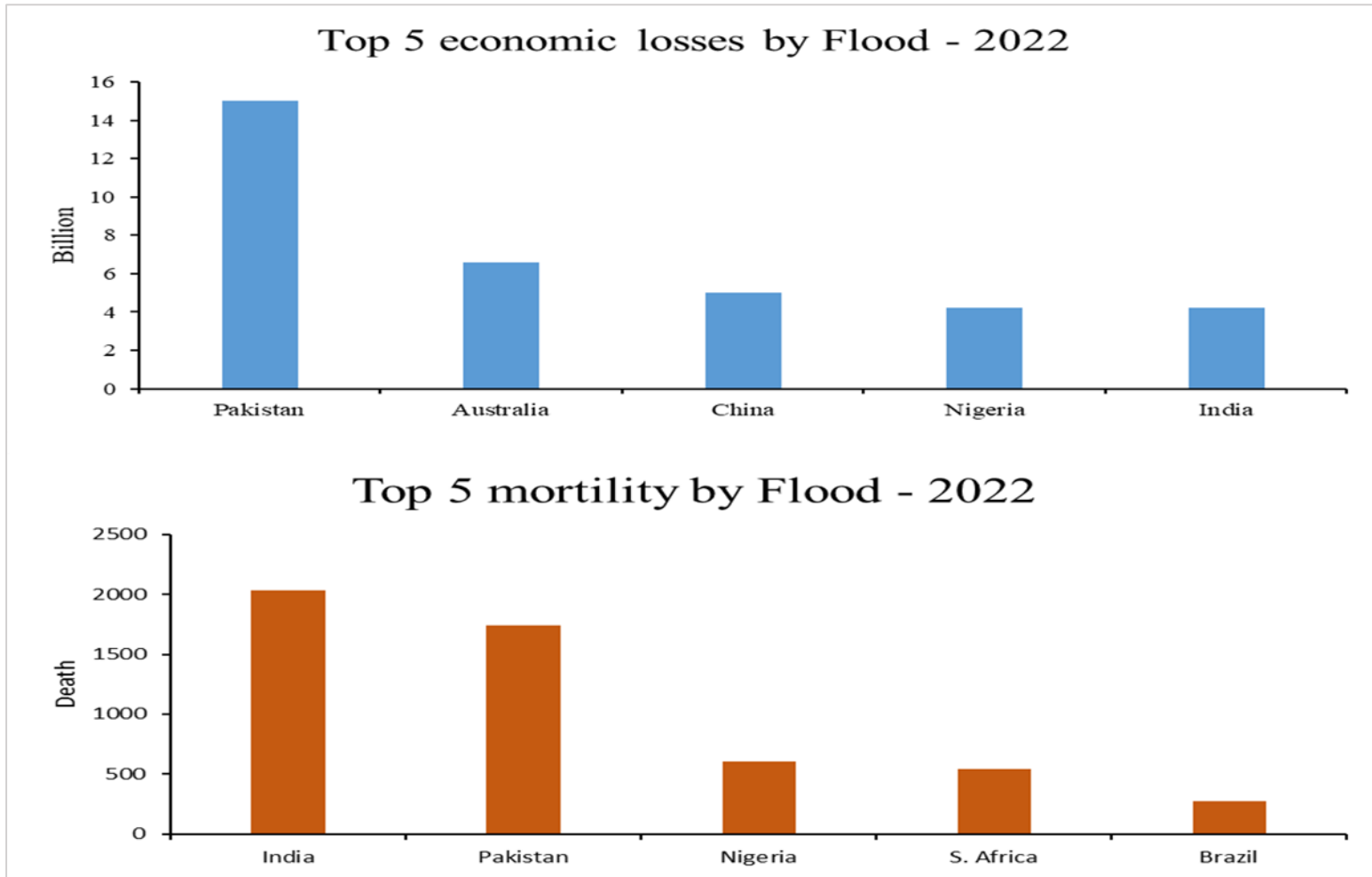


Figure 1.3: Top 5 countries in economic losses and mortality by flood (Source: CRED. 2022 Disasters in numbers.

Brussels: CRED; 2023)

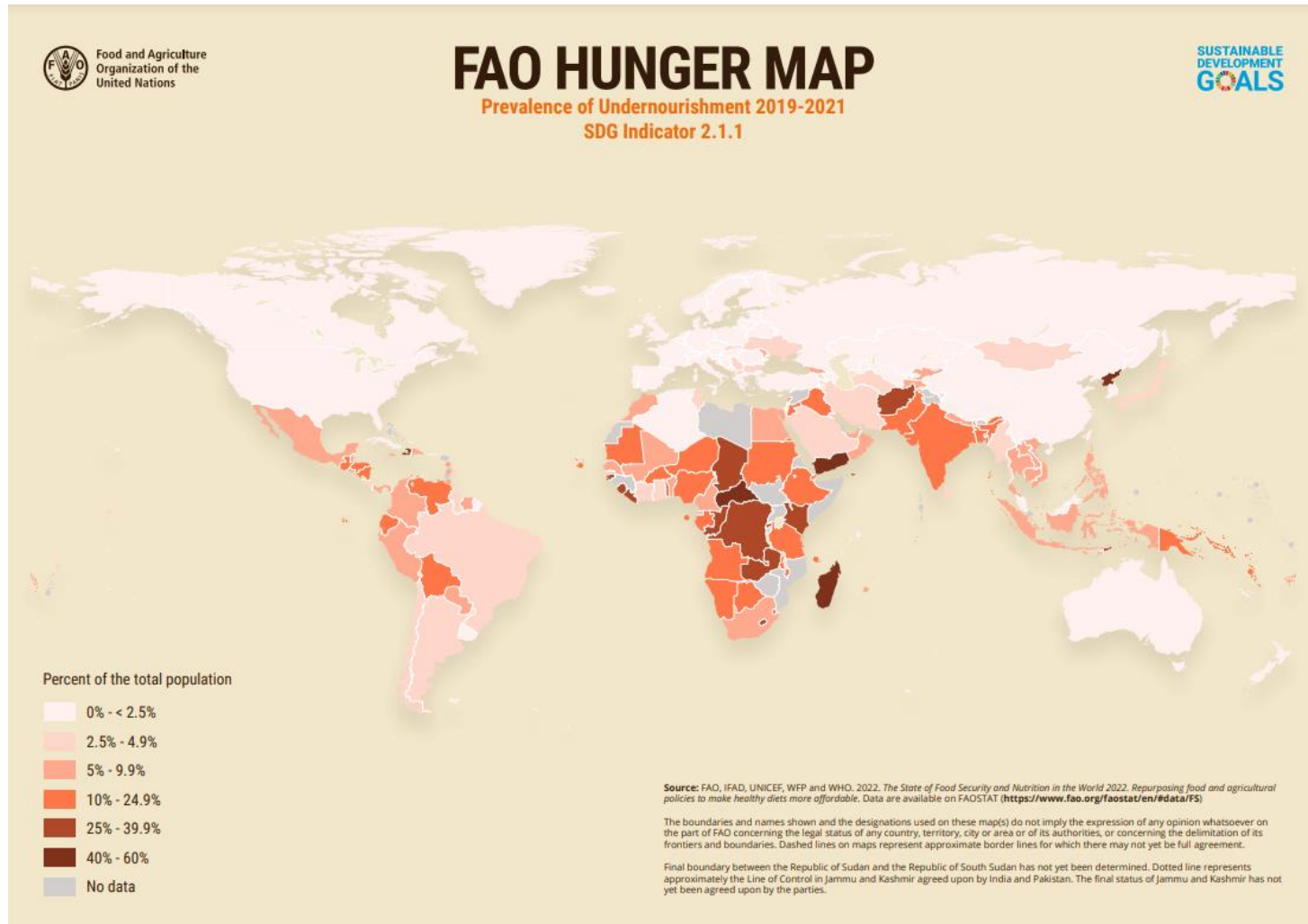


Figure 1.4: FAO Hunger Map (Source: FAO, IFAD, UNICEF, WFP and WHO. 2022)

The FAO of the United Nations disclosed that floods constituted the second most significant agricultural calamity, following droughts, with a detrimental impact amounting to \$21 billion (Fig. 1.4 & 1.5) in terms of crop and livestock losses across the developing countries between 2008 and 2018 (FAO 2021). This amount is 19% of total loss. In Asia, agricultural production has experienced significant losses, especially in 2015. “This spike in losses was due to several major disasters in the region, including the Nepal earthquake, monsoon flooding in Myanmar, Bangladesh, and India, and widespread flooding in Chennai, India” according to (FAO 2021).

Flooding is a major threat to achieving the SDG goals (Fig. 1.6), especially "no poverty (Goal 1) and zero hunger (Goal 2)" (<https://sdgs.un.org>; Kim *et al.*, 2023). They can destroy crops, homes and infrastructure, lead to food shortages and price increases, which can have a significant impact on farmers and other rural communities.



Figure 1.5: Year wise (2008–2018) total loss in crop and livestock production
(Source: FAO, 2021)



Figure 1.6: Impact of flood on SDG Goal-1&2 (After; <https://sdgs.un.org/goals>)

1.2.2 Flood Scenario of India

In India, floods are a major natural disaster. The main reason for this is heavy rains, which overflow the rivers and the surrounding areas get submerged. India receives mostly rainfall during the monsoon season, from June to September. However, the distribution of rainfall is irregular, with some areas receiving much more rain than others. This variation can lead to floods in areas that are not traditionally prone to them. The Ganga-Yamuna basin states “Haryana, Himachal Pradesh, Uttar Pradesh, Bihar, and West Bengal” are prone to floods. Flood-prone areas in India include Assam, where the Brahmaputra River causes frequent floods, and Odisha, which is at risk from the Mahanadi River. In Andhra Pradesh, the Godavari, Krishna, and Pennar rivers pose flood threats. Gujarat experiences floods in the basins of the Narmada, Sabarmati, and Tapti rivers. In recent years, the Krishna and Godavari river basins have also caused severe floods in Karnataka and Maharashtra.

The map labelled as figure 1.6 provides a visual representation of the flood-prone and liable-to-flood areas in India. The map depicts the regions across the country that are susceptible to flooding due to various factors such as heavy rainfall, proximity

to rivers, and topographical features. This map is included here to enable the viewer to quickly understand the extent and distribution of flood-prone areas in India (Fig. 1.7).

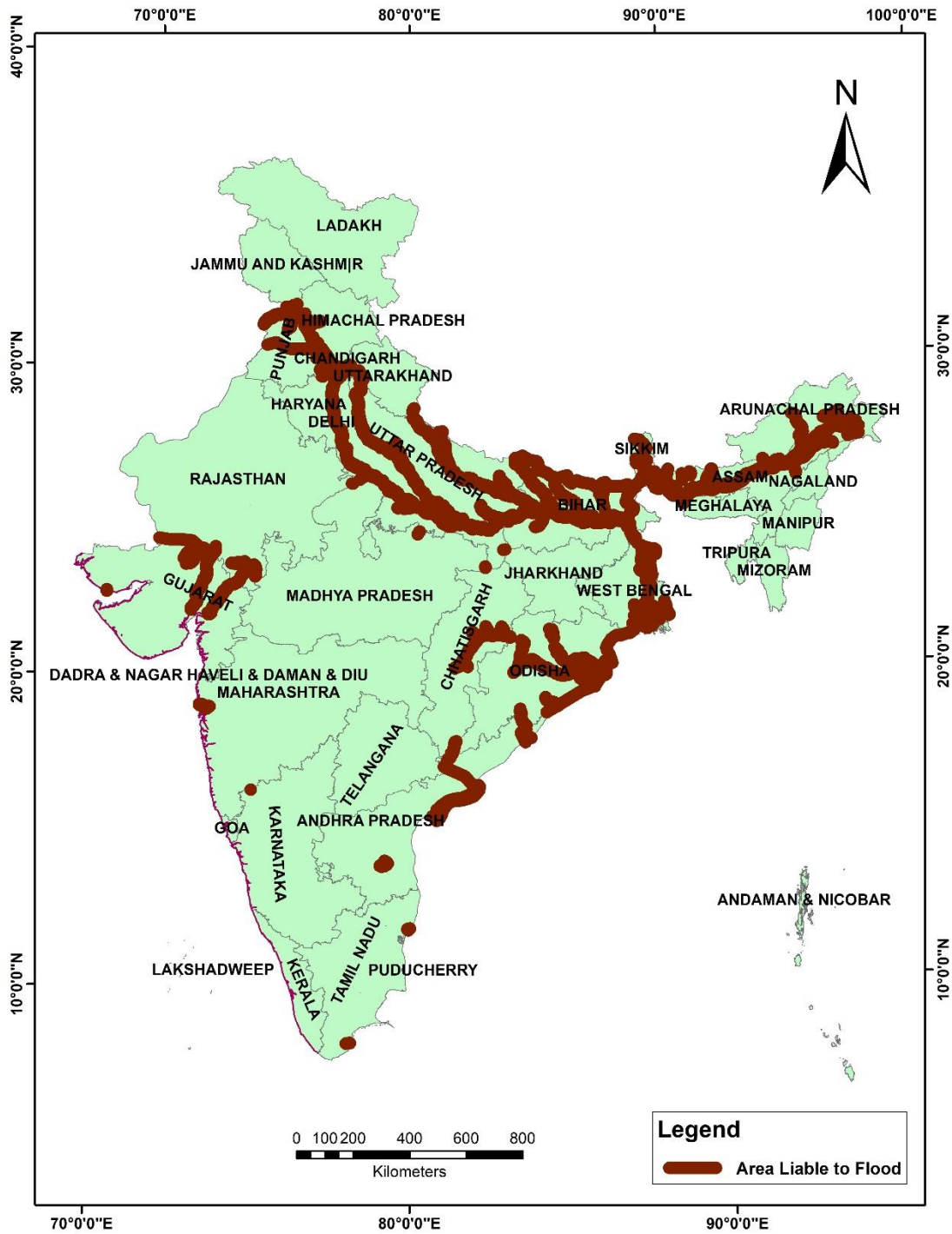


Figure 1.7: Flood prone areas in India (Source: NITI Aayog (2021)/
<https://ndma.gov.in/Natural-Hazards/Floods>)

1.2.3 Flood Scenario of Bihar

Bihar is characterized by a diverse topography which includes many perennial and non-perennial rivers (Fig. 1.8). These rivers, which originate from Nepal, carry a large amount of silt, which eventually gets deposited in the plains of Bihar. The region experiences most of its rainfall during the monsoon season, usually lasting 3 to 4 months. The flow of the river increases significantly, sometimes up to 50 times its normal rate, causing frequent floods in Bihar. About 73% of the total land of Bihar is vulnerable to floods. The state bears a significant burden of flood damages in India, contributing about 30–40% of the country's annual flood-related losses. This region is home to 22.1% of India's flood-affected population, with 28 districts falling in the category of most flood-prone areas (<http://bsdma.org>).

Flood is a major natural disaster in Bihar and has a severe impact in the form of loss of life, property, infrastructure and agriculture every year. Climate change is increasing the intensity and frequency of floods around the globe (Sadiq *et al.*, 2022).

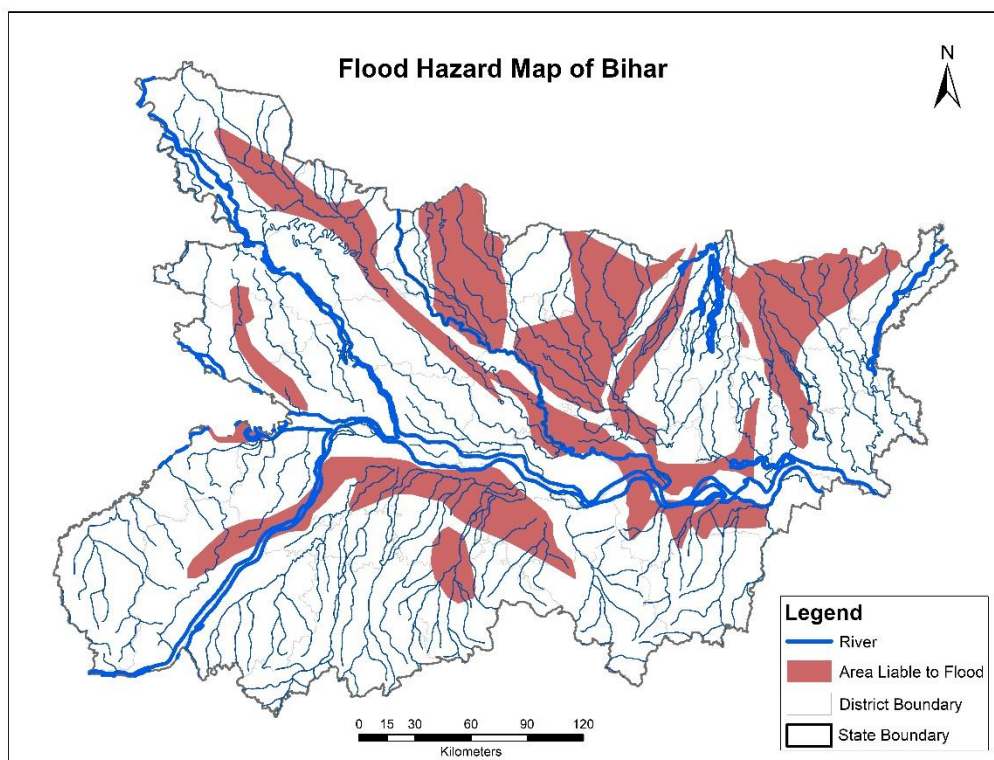


Figure 1.8: Flood hazard map of Bihar (Source: BMTPC)

Many houses, roads, bridges and public facilities are washed away during floods. This damage disrupts the normal life of the affected population, as many people are forced to flee their homes and seek temporary shelter in relief camps. The floods also cause significant damage to crops and agricultural land, leading to a loss of livelihood for farmers (Table 1.1). Rapid growth of urban, deforestation, unplanned development and erratic rainfall are the main causes of frequent floods in north part of Bihar (North Bihar). Bihar is a state in India where about 76 percent of the population (Table 1.1) is dependent on agriculture, which is severely affected by frequent floods (Anonymous, 2023a, b, c).

Table-1.1: Flood related damage in Bihar in last 5 decades

STATEMENT SHOWING FLOOD DAMAGE DURING 1980 TO 2019 in BIHAR										
Year	Area affected in (mha)	Population affected in (million)	Damage to Crops		Damage to Houses		Cattle lost nos.	Human live lost nos.	Damage to public utilities in Rs Crore	Total damage crops, houses & public utilities in Rs Crore
			Area (m.ha)	value (Rs. crore)	Nos.	Value in (Rs. Crore)				
1980	1.9	7.6	1	Neg	81016	3.7	42	67	54	57.7
1981	1.3	7	0.6	72.1	76776	4	11	18	3	79.1
1982	0.9	4.7	0.3	97	68242	6.9	14	25	9.6	113.4
1983	1.5	5.1	0.4	23.4	24160	1.2	19	6	1.7	26.3
1984	3.1	13.6	1.6	185.4	334285	2.2	90	143	27.2	234.6
1985	0.8	4.9	0.4	30.3	69168	4.3	9	69	1.9	36.5
1986	1.8	7.3	0.6	103.7	76756	4.4	491	94	200	308.1
1987	4.2	24.5	2.2	601.3	1305484	217.7	4116	1116	490.6	1309.6
1988	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1989	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1990	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1991	0.98	4.823	0.4	23.61	27324	6.14	84	56	1.4	31.15
1992	0.08	0.55	0.02	0.58	1281	0.16	0	4	0.0075	0.75
1993	1.56	5.35	1.14	130.5	219826	88.14	420	105	30.41	249.05
1994	0.65	4.01	0.35	56.16	34876	4.95	35	91	1.52	62.63
1995	0.92	6.63	0.42	195.14	297765	75.1	3742	291	21.83	292.07
1996	1.19	8.73	0.73	71.69	116194	14.95	171	222	10.35	96.99
1997	1.47	6.96	0.65	57.37	174379	30.56	151	163	20.38	108.31
1998	2.51	13.47	1.28	366.96	NA	NA	NA	NA	NA	NA
1999	0.84	6.56	0.3	242.04	91813	53.85	136	243	54.1	349.99
2000	0.78	8.24	0.42	75.53	227897	148.56	1702	308	28.55	252.64
2001	1.19	9.1	0.65	267.22	222074	173.58	565	231	183.54	624.34
2002	1.97	16.02	0.94	511.5	419014	526.21	1450	489	408.92	1446.63
2003	1.82	8.16	0.79	94.45	84424	37.57	131	297	11.92	143.94
2004	2.7	21.3	1.4	522.05	929773	758.09	3272	885	1030.5	2310.64
2005	0.46	2.1	0.13	11.64	5538	3.83	4	58	3.05	18.52
2006	17.83	1.16	0.15	14.32	28115	25.97	31	44	85.27	125.56
2007	1.88	24.44	1.06	768.38	784328	831.45	2423	1287	642.41	2242.24
2008	0.64	5.09	0.26	167.3	256447	314.93	32026	626	467.91	950.14
2009	1.11	2.2	0.05	21.82	7674	5.28	2	125	5.3	32.4
2010	0.2	0.72	0.01	3.12	15170	7.05	0	32	1.69	11.86
2011	3.82	7.14	0.34	102.96	85182	69.07	183	249	1.54	173.57
2012	0.11	0.24	0.03	3	2261	1.6	0	15	1.62	6.22
2013	2.36	7.23	0.71	222.8	169501	35.34	6480	253	18.23	276.37
2014	3.9	3	0.5	176.4	13662	10.78	34	158	5.4	192.58
2015	0.01	0.49	0	0	518	0.69	1	27	0.03	0.72
2016	4.33	8.79	0.71	519.77	69102	40.03	538	458	116.49	676.29
2017	3	18.5	0.3	685.87	263848	778.79	472	815	103.37	1568.03
2018	0.03	0.15	0	5.14	1074	0.41	0	1	0.01	5.56
2019	1.06	14.94	0.36	446.64	45161	27.96	183	300	167.5	642.1

(Source: Disaster Management Department, Govt. of Bihar; Flood Hazard Atlas of Bihar, 2019)

However, floods are considered a regular concern in North Bihar during the monsoon season. Due to overflow of rivers in neighboring country Nepal, North Bihar has experienced severe floods during the last three decades (Flood Hazard Atlas of Bihar, 2019 and Kumar *et al.*, 2022). During COVID-19 pandemic, it has worsened the situation, where natural disasters and other factors have disrupted the economic and social stability of communities. Floods created significant impact on agricultural sector, which is a major source of livelihood for many people in Bihar. When floods occur, crops are completely destroyed or damaged, making it difficult for farmers to earn a living. Additionally, floods have damaged homes and other infrastructure, making it unsafe or impossible for people to stay in their communities. As a result, many people in North Bihar are forced to migrate to other areas in search of work and income during floods to sustain their lives. It is a difficult and often traumatic experience, as families are separated and individuals are forced to leave their homes and communities behind. Efforts to address the root causes of migration during floods, such as investing in flood-resistant infrastructure and promoting economic diversification, can help to reduce the need for people to leave their homes in search of livelihoods during times of crisis in north Bihar.

To mitigate the impact of floods, an early warning system is also highly needed. Because timely monitoring of flood extents helps authorities identify the severity and plan relief operations. But, there is a lack of an effective flood monitoring and early warning system due to the poor availability of resources in developing countries like India (Wu *et al.*, 2012). Flood intensity has been increasing over the last three decades (Freer *et al.*, 2013), therefore the role of remote sensing (RS) is crucial for flood mapping, monitoring and model development to monitor the impact of floods.

1.3 Application of Remote Sensing and Geospatial Technology in Flood Management:

Over the years, remote sensing satellite data has been capable of monitoring flood extent, intensity, progression and deterioration on a real-time basis. Remote Sensing satellite has capability to view earth surface in synoptic way from space based of aerial satellite sensors. Aerial or satellite imagery can be used to estimate surface

water extent (Groeve *et al.*, 2010). Hence, the role of remote sensing and geospatial technologies for flood extent mapping and monitoring are important for mitigating the impact of floods. This technology is capable to monitor the flood extent and to provide real-time information of flood impact. Because, it allows for the collection of data on rainfall, water levels, and soil moisture, which can provide valuable information to farmers and local authorities to prepare for floods and make decisions about when to plant, harvest, and evacuate the population in the event of a flood. For flood-related mapping, synthetic aperture radar (SAR) data having upper edge than multispectral optical data, because of its all-weather and day-night sensing capability. The European Space Agency (ESA) has launched “Sentinel-1A satellite on 3 April 2014 and Sentinel-1B was on 25 April 2016”. Its revisit time is six days at equator with two satellites (Sentinel-1A/B) (Torres *et al.*, 2012). It is the first-ever global SAR mission, whose datasets are open access for the global public and researchers.

Nowadays, the Remotely Sensed Earth Observation (EO) datasets are being commonly used for disaster management purposes (Schumann *et al.*, 2018), which is freely available for researchers. Due to spatial and temporal characteristics of remote sensing data, it can be used to acquire the essential information from geomorphological features of rivers. This information can be beneficial for mitigation measures in the time of disaster. For different analysis, multiple sources of satellite data are freely available. However, due to resource unavailability, downloading, storing and processing of satellite datasets are big task for users.

Hence to overcome these problems, Google launched the most advanced cloud-based geospatial processing platform “Google Earth Engine (GEE)”. “It enables to access high-performance computational resources to process satellite data without the requirement of local storage, in addition to allowing up-to-date remote sensing databases for scientific and academic purposes” stated (Gorelick *et al.*, 2017; Schumann *et al.*, 2018). It enables to share the developed codes of different analysis to multiple users and researchers. The GEE, introduced by Google, Inc., as a new computing platform for large-scale data processing such as the time series data analysis of Landsat archive (Gorelick *et al.*, 2017). GEE platform hosted a complete, up-to-date and ready SAR data archive of Sentinel-1A/B Ground Range Detected (GRD) data.

In the present study, Sentinel-2 MSI and PLANET-NICFI (PlanetScope) datasets have been utilized for LU/LC mapping and Sentinel-1 for flood extent mapping and monitoring. The ability of SAR sensors to detect the extent of flooding depends on different scattering mechanisms. For the identification of inundated pixels, several SAR-based flood identification methods have been used scattering mechanisms by applying backscatter thresholds to satellite imagery (Chini *et al.*, 2017). Usually, the change detection technique is being utilize for the identification of flooded pixels using SAR datasets. Likewise, various methods and indices exist for the extraction of waterbodies using optical and microwave satellite datasets.

The Normalized Difference Water Index (NDWI) was first proposed by McFeeters in 1996 and was shown to be a robust index for detecting waterbodies, especially in areas with high vegetation cover. The Modified Normalized Difference Water Index (MNDWI) was proposed by Xu in 2006 as a modification of NDWI, specifically designed for urban areas. It is known to be effective for extracting waterbodies in urban areas with high reflectance from built-up structures. Recently, the Automated Water Extraction Index (AWEI) was developed by Feyisa *et al.* in 2014 and is designed to overcome the limitations of NDWI and MNDWI in areas with mixed pixels, such as rivers or wetlands. AWEI is based on a combination of the green, blue, and red bands, as well as the shortwave infrared and thermal infrared bands, which allows for more accurate detection of waterbodies. These indices have been widely used and validated for detecting waterbodies and assessing the impact of floods.

Apart from this, thresholding technique are also being used for identification of flood extent using JavaScript code. Therefore, it is always necessary to have accurate data of flooded areas to make an accurate assessment of the damage to make a viable decision on prioritizing relief. Numerous research have been undertaken over the years to reduce the consequences of frequent floods in Bihar, India, using optical and SAR satellite datasets. In the present work, a cloud computing platform (CCP) has been utilized to rapidly demarcate the flood-affected region during flood events and develop an algorithm for tracking flood procession of North Bihar.

1.4 Study area

The present work is carried out for North Bihar, which is located between 24.33611 to 27.52083 N latitude and 83.33055 to 88.29444 E longitude as shown in Figure 1.9. As per census 2011, Bihar has total 10.41 crores populations. Bihar gets about 1205 mm of rain falls in a year on average and has the World's most fertile alluvial plains of Gangetic Valley. Its soil distribution is loam, clay, clay loam and sandy loam (DAC&FW, GoI, 2020). The main agricultural products of Bihar include maize, wheat, and rice. North Bihar is highest producer of maize crop in comparison of South Bihar (Singh *et al.*, 2018).

The present work is consisting 22 districts (North Bihar) of Bihar state of India. North Bihar is home to several rivers that have been responsible for recurrent flooding in the region. The major rivers in Bihar include the Kosi, Gandak, Budhi Gandak, Bagmati, and Mahananda. These rivers have caused devastating floods in the state in the past, with north Bihar being the worst affected.

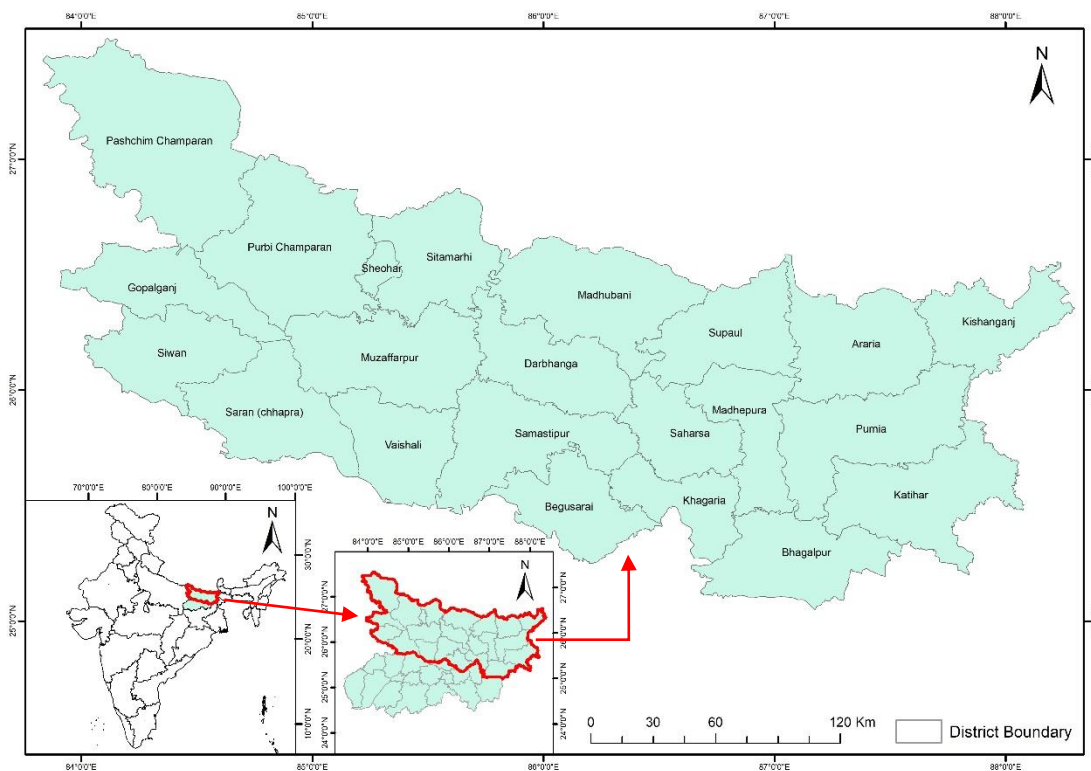


Figure 1.9: Location map of study area

The Kosi River, which runs through Bihar, is known as the "Sorrow of Bihar" owing to its unpredictable and destructive floods. The study area experiences frequent floods due to riverine floods induced by excessive rainfall, and some of the major flood events in the study area have occurred in 1987, 1998, 2000, 2001, 2003, 2004, 2007, 2008, 2010, 2013, 2017, 2018 and 2020 (Flood Hazard Atlas of Bihar, 2019; Kumar *et al.*, 2022).

In this region, maize, wheat and rice are the major crops grown, which provide food, fodder and industrial raw material (Ethanol Industry) to the state and contribute significantly to the economy of the state. The region has the most fertile land and 76 per cent of the population is dependent on agriculture (Kumar *et al.*, 2022; Spera *et al.*, 2016). Agricultural land is degrading and shrinking day by day due to rapid urbanization and the growing population (Singh *et al.*, 2014). There are two main agricultural operations seasons in the region; viz. In the monsoon (Kharif) season mostly rice crops are grown, and in the winter (Rabi) season (November to April) mostly maize and wheat crops are grown. Because there is a less uncertain climate in Rabi season as compared to Kharif.

Hence, it is easier and less risky to grow maize crops in Rabi as compared to the Kharif season in Bihar (Bihar Study Report:2016). Most of the farmers in this region are smallholder farms. Therefore, this region is ideal for investigating how well micro-satellites such as PlanetScope can be able to map the field levels crop (Jain *et al.*, 2016) and impact of flood on crop too.

1.5 Motivation

Maize (*Zea mays* L.) is globally recognized as one of the most important cereal crops with significant economic and nutritional importance. Its versatile nature and adaptability to different agro-climatic conditions make it a staple crop for many smallholding farmers in developing countries like India. This crop provides both food and fodder for millions of people and animals, thus playing an important role in global food security (FAOSTAT, accessed on 10/May/2023).

In India, maize has a double yield potential compared to other cereal crops like wheat and rice. This yield potential has been made possible by the adoption of improved technologies, such as hybrid varieties, which have resulted in significant yield gains in maize cultivation. As a result, maize has become a highly productive and profitable crop for farmers in India, contributing significantly to the agricultural and economic development of the country. Bihar is the sixth largest producer of maize among the states of India, with about 76% of its population dependent on agriculture. Which gets submerged almost every year by floods. Yet it still positioned itself among top maize producing state.

With the development of remote sensing technology, it is now possible to track the extent and intensity of floods in near real-time basis using satellite and aerial photographs. This approach will enable the authorities to take timely and appropriate decisions to minimize the potential impact of flood on life, property and agriculture sector. With the increasing availability of freely accessible remotely sensed satellite data such as Landsat, Sentinel-1, Sentinel-2, and NIFCI-Planet, crop acreage and yield estimation has become more accurate and efficient. This technology enables monitoring of crop phenology and damage due to natural calamities (flood, drought, heat wave etc.) on near real-time basis, by improving the accuracy of yield forecasts and facilitating quick and effective decision-making. The use of remote sensing technology for crop monitoring has also significantly reduced the cost and time associated with traditional methods such as ground surveys.

Nowadays, various techniques are being used for flood extent demarcation, crop mapping and monitoring such as thresholding method, classification, change detection and data fusion. Research is still needed to develop more effective algorithms for flood extent mapping and classifiers for crop area and yield estimation.

1.6 Research Objectives

The aim of this study is to assess the impact of flood on cropping pattern and Production of maize using geospatial technology in north Bihar. Hence, the study is divided into three objectives to accomplish the desired outcome.

Thus, the major objectives are as follows:

- (i) To estimate the flood extent and its impact on maize productivity
- (ii) To discriminate maize from other crops
- (iii) To estimate the acreage and yield of maize.

1.7 Brief Overview of Adopted Methodology

In the present study, remotely sensed satellite datasets and the Otsu model on GEE platform are employed for flood extent mapping. Ground Truth (GT) data and machine learning approach are utilized to discriminate maize crops from other crops and estimate acreage. The evolved spectral indices, machine learning algorithms, semi-physical models and climate datasets has been used for maize crop yield and production estimation.

1.8 Organization of Thesis

The thesis is organized into eight chapters. The summary of the chapters is provided as follows:

Chapter 1: Introduction

This chapter is focused as an introductory section of the thesis, providing a broad overview of the research topic, definition of the problem, selected site, motivation and objectives of the research. The main focus of this thesis is to assess the impact of flood on maize crops in North Bihar, India using remote sensing satellite imagery.

Chapter 2: Review of literature

It provides a brief overview of technology implementation for crop mapping and monitoring, techniques for image classification, a comparison of methods for flood extent mapping and assessing its impact on crops, and the identification of research gaps. Additionally, the review highlights areas where further research is needed to address current gaps in knowledge and methodology.

Chapter 3: Methodology and Data used

This section introduces advanced methods and techniques employed to achieve

the desired results of the study. It begins by providing a brief overview of the data used and the processing approaches adopted for handling large datasets, such as SAR data from Sentinel-1 for flood mapping and Optical datasets from Sentinel-2 and PlanetScope for crop acreage, yield, and production estimation. Furthermore, the section outlines the methodology for acquiring Ground Truth (GT) data and conducting Crop Cutting Experiments (CCEs) to accomplish the study's objectives. These processes play a vital role in validating and refining the results obtained through remote sensing techniques.

Moreover, this section delves into the complexity of data processing and emphasizes the utilization of the cloud computing platform GEE. By incorporating these methods, the study aims to comprehend the intricacies involved in data processing and harness their potential for successful implementation in achieving the study's objectives.

Chapter 4: Flood Extent and Flood-Affected Area Mapping

This section provides a comprehensive description of the application of remote sensing techniques for flood-affected area mapping, monitoring and managing in North Bihar. It specifically focuses on two key methods: the use of the random forest algorithm for land use classification and thresholding techniques for rapid flood extent mapping, both implemented on the GEE's cloud computing platform.

Chapter 5: Maize crop identification/ discrimination and acreage estimation using remote sensing satellite data, machine learning algorithms and cloud computing platform.

It offers a detailed explanation of how remote sensing satellite data, machine learning algorithms, and cloud computing platforms are employed for crop identification, discrimination, and acreage estimation. The chapter further discusses the significance of cloud computing platforms, like GEE, which offer the computational infrastructure and resources required to process large-scale remote sensing datasets. These platforms enable efficient data processing, allowing for crop identification and estimation of crop acreage at various scales.

Chapter 6: Model development for maize (*Zea mays* L) crop yield and production estimation using machine learning algorithm, evolved spectral indices and cloud computing platform.

This chapter focuses on the methodology of estimating maize crop yield and production using machine learning algorithms and cloud computing platforms. It outlines how these techniques/methods are employed to analyse relevant data and make accurate predictions for maize crop yield and production. Additionally, the chapter introduces a newly developed spectral index that has been evolved as part of the present study. This spectral index is designed to capture specific characteristics and parameters related to crops, enhancing the accuracy and effectiveness of yield and production estimation.

Chapter 7: Conclusion and recommendations

In this section, the study concludes by addressing the research questions and presenting the key research findings. It provides clear and concise answers to the research inquiries, consolidating the knowledge gained throughout the study. Additionally, the present study summarizes the most effective techniques utilized, highlighting their significance in addressing real-world challenges. It also underscores the societal relevance of these techniques and their potential impact on future directions, emphasizing their value in advancing agricultural practices, disaster management, and decision-making processes.

Maize cultivation has emerged as an important source of income and employment for millions of Indian farmers. It has several advantages over other crops, including its high productivity, low input requirement and ability to adapt to different agro-climatic conditions. The crop is being used for both food and non-food purposes, such as animal feed, industrial applications, and biofuel production. It has the potential to make a significant contribution to India's efforts to double farmers' income.

Therefore, innovative agricultural research and better management practices are essential to increase productivity and reduce food insecurity. As the global population continues to grow, the demand for food is also increasing. To meet this demand, we need to produce more food with fewer resources, including land, water, and energy. Agricultural research and development can help us achieve this goal by developing new crop varieties that are more resistant to pests and diseases, more tolerant to environmental stresses, and more efficient in nutrient uptake and utilization.

In addition to research, better management practices are also crucial to increase agricultural productivity. This includes the adoption of modern technologies, such as precision agriculture, smart irrigation, and integrated pest management, which can help farmers optimize the use of resources and minimize waste. Effective management practices also involve better farm planning, soil and water conservation, and climate-smart agriculture, which can help mitigate the negative impacts of climate change on crop yields and production.

In this regard, accurate and timely information on crop type, acreage, and yield is crucial for effective agricultural planning, policy-making, and resource allocation (Jamal *et al.*, 2023). This information can be helpful for farmers, policymakers, and other stakeholders to make quick decisions about agricultural investments, crop management, and market opportunities (Niftiyev *et al.*, 2023). Farmers and other stakeholders can anticipate market trends and adjust their production and marketing strategies accordingly. Concerned authorities and stakeholders can also allocate

resources more effectively and efficiently to promote sustainable development and profitable agricultural systems.

2.1 Crop Mapping Method

Crop mapping is a technique/method used to identify and classify the different types of crops in a target area (Table 2.1). It provides valuable information for agricultural planning, resource allocation and monitoring of crop health. Nowadays, several methods are being employed for crop mapping such as remote sensing (Deb *et al.*, 2021), machine learning (Han *et al.*, 2019; Rukhovich *et al.*, 2021), ground-based surveys and crop inventory data (Darji *et al.*, 2023).

Table 2.1: The advantages and disadvantages of crop mapping method

Method	Advantages	Disadvantages
Remote sensing	Cost-effective, efficient, can cover large areas	Less accurate than ground-based surveys or crop inventory data
Machine learning	Can be more accurate than remote sensing, can be used to identify crops that are not easily visible in remote sensing imagery	Requires a large amount of training data, can be expensive to develop and deploy
Ground-based surveys	Most accurate, can collect data on other factors such as crop health and pest infestation	Time-consuming and expensive, can only cover small areas
Crop inventory data	Very accurate, can be used to track crop production over time	May not be available for all areas, can be outdated

2.1.1 Ground-based surveys

Ground-based surveys involve physically field visiting and inspecting the crops in a specific area. Agronomists and field experts collect data on crop characteristics such as height, leaf shape, color, and growth stage. This information is used to identify and map different crop types. While this method is more time-consuming and labour-intensive, it can provide accurate results when combined with other techniques.

2.1.2 Crop inventory data

This method provides detailed information about the crop types grown in specific regions and their acreage and yield. Which can be collected by governments, agricultural organizations, or private companies.

2.1.3 Remote sensing

Remote sensing is a valuable technique that involves the use of satellite or aerial imagery to assess and monitor agricultural land and its associated crops. It allows for identification, classification, and mapping of different types of crops over large areas. In this process, the imagery is typically analyzed using computer algorithms to identify the spectral signatures of different crops.

2.1.3.1 Steps /Overview of crop mapping using remotely sensed satellite data

- a) **Satellite or Aerial Imagery Acquisition:** Remote sensing data is collected using satellites or airborne platforms equipped with sensors such as multi-spectral or hyperspectral cameras. These sensors capture images of the Earth's surface in different wavelengths of the electromagnetic spectrum.
- b) **Pre-processing:** The acquired imagery undergoes preprocessing steps to correct for atmospheric interference, geometric distortions, and other image artifacts. This ensures that the data is in a usable form for subsequent analysis.
- c) **Image Classification:** The preprocessed imagery is subjected to image classification techniques. This involves assigning each pixel or group of pixels in the image to specific land cover/crop types. Several classification algorithms can be employed, including supervised and unsupervised methods.

- I. Supervised Classification: In supervised classification, training samples of known land cover types are selected on the image, and the algorithm learns to classify the rest of the pixels based on their spectral properties. It requires the user to identify representative training areas for different crop types.
 - II. Unsupervised Classification: It involves grouping pixels with similar spectral characteristics without prior knowledge of the land cover types. The algorithm automatically identifies clusters in the image based on statistical properties, and the user assigns land cover labels to the resulting clusters.
- d) **Crop Mapping and Analysis**: Once the classification is completed and validated, the resulting maps can be used to extract relevant information about the crops. This may include the distribution, extent, and spatial patterns of different crop types within the study area. Crop mapping can assist in monitoring changes in land use, estimating crop yield, identifying areas of crop stress or disease, and supporting decision-making in agriculture, such as resource allocation and crop management strategies.
- e) **Accuracy Assessment**: To ensure the reliability of classification results, accuracy evaluation is necessary. This approach comparing the classified results with GT dataset, which may include field surveys or high-resolution imagery. It also helps in evaluating the performance of classification algorithm and provides the overall classification accuracy. It's worth noting that the accuracy of crop mapping using RS technology depends on various factors, including image quality, classification methods, availability of ground truth data, and the expertise of the analyst. Advances in remote sensing technologies and machine learning algorithms continue to enhance the accuracy and efficiency of crop mapping, making it essential tool for precision agriculture and land management.

2.1.4 Machine learning

Machine learning (ML) is a powerful method for crop mapping that utilizes algorithms to analyze data and classify different crop types. In this method, models are trained by RS imagery and ground-based survey data to identify patterns and

relationships between features and crop types. With accurate labeling and pre-processing of the data, the model can make predictions on unlabeled data to map and classify crops in a given area. Continuous monitoring and evaluation of the model's performance ensure reliable and up-to-date results for crop mapping.

Nowadays, Machine learning algorithms, such as Random forest, Support vector machine, deep learning and convolutional neural networks, are used to process and analyze remote sensing data for crop mapping (Lary *et al.*, 2016; Inglada *et al.*, 2015; Prins *et al.*, 2021). These algorithms are designed to learn from large datasets (collected from study area such as ground truth data) and can automatically identify patterns and features in satellite images that correspond to different crop types. Machine learning approach can also be used to integrate other data sources, such as weather data, soil data (Cravero *et al.*, 2022), and farm management practices, to improve the accuracy of crop mapping and monitoring (Goldstein *et al.*, 2018; Filippi *et al.*, 2019).

2.2 Concluding remarks for best method/ suitable method for crop mapping

Remote sensing is often the most cost-effective and efficient method for large-area mapping, but it can be less accurate than ground-based surveys or crop inventory data. Ground-based surveys are more accurate, but they can be more time-consuming and expensive. Crop inventory data can be very accurate, but it may not be available for all areas.

Nowadays, remote sensing satellite data and machine learning algorithms have become powerful tools for crop mapping and monitoring (Abdi *et al.*, 2020). But, accurate crop-type mapping over a large area is still a challenging task (Weiss *et al.*, 2020). The quality of RS images, including their spatial, spectral, radiometric and temporal resolutions, is the main constraint (Gomez *et al.*, 2016). To improve the accuracy of crop type mapping using remote sensing, it is necessary to focus on improving data quality and developing more sophisticated image interpretation and information extraction algorithms (Wang *et al.*, 2023). In this context, ML algorithms provide a powerful and emerging toolset for satellite data/imageries interpretation and information extraction for large-area mapping, facilitating accurate and efficient

analysis of vast geographic areas. One of the reasons for the emergence of ML is the availability of large and diverse datasets, which allow algorithms to be trained more effectively. Additionally, advancements in computing power, such as the use of cloud computing platform (CCP) and powerful GPUs (Graphics Processing Units) have accelerated the training and deployment of ML models.

In addition, the advent of high-resolution satellite imagery, such as Sentinel-1&2, Landsat-8 and NICFI-PlanetScope has revolutionized the way of crop mapping and monitoring. The availability of these free and open-source remote sensing satellite datasets has facilitated the use of these tools for agriculture and food security monitoring. These satellites capture images with high spatial and temporal resolutions, allowing detailed observations of land surface features and changes (Panjala *et al.*, 2022). Optical sensors and Synthetic Aperture Radar (SAR) are two different types of remote sensing technologies that provide unique capabilities for data acquisition and analysis.

2.2.1 Use of Optical datasets

Optical sensors collect information in the visible and near-infrared (NIR) portions of the electromagnetic spectrum. This means that they can see the Earth's surface in the same way that human eye does, which makes them good for applications such as land cover mapping, agriculture monitoring, and disaster response. However, optical sensors are not able to see through clouds or darkness, so they cannot be used in all conditions. Landsat-8 is a long-standing optical satellite sensor with a legacy of providing multispectral imagery for more than 40 years. It captures images with a resolution of 30 meters and has a revisit time of 16 days. The data from Landsat-8 can be used for various applications, including land cover classification, forest monitoring, and agriculture management.

Sentinel-2 is also an ESA satellite that provides high-resolution multispectral data for monitoring land use and land cover (LULC) changes, including crop mapping. Sentinel-2 has a 10-meter spatial resolution, which allows for the identification of individual crops and the detection of subtle changes in vegetation growth and health.

NICFI-PlanetScope is a satellite imagery program that provides global, high-resolution satellite imagery for environmental monitoring, disaster response, and agricultural management. NICFI-PlanetScope offers daily imagery coverage of entire planet at spatial resolution of 3 to 5 meters, making it an ideal tool for crop monitoring and mapping. These remote sensing satellite data provide high spatial and temporal resolution imagery that can capture detailed information on crop type, growth stage, and health status. This information is valuable for crop mapping, which is the process of identifying and delineating the extent and location of different crops in the study area. It can be helpful in improving farm management practices, monitoring crop productivity and predicting crop yields.

2.2.2 Use of Synthetic Aperture Radar (SAR) datasets

Synthetic Aperture Radar (SAR) sensors emit microwave signals towards the target area and then receive and measure the backscattered signals, which can penetrate through various atmospheric conditions, including clouds, fog, and haze. SAR datasets can be used effectively in all weather conditions and at any time of day or night.

Nowadays, Sentinel-1 SAR satellite of the European Space Agency (ESA) that provides high-resolution images regardless of weather and light conditions, which is being utilized to detect changes in land use, such as deforestation, urbanization, and agricultural practices.

2.2.3 Combining satellite data and Data fusion techniques

SAR and optical data capture different aspects of the Earth's surface. Optical sensors capture reflected sunlight and provide information about the spectral properties of crops, such as their color and vegetation indices. On the other hand, SAR sensors transmit microwave signals, which offer valuable data for analysis with capability to operate regardless of weather conditions or time of day. By fusing SAR and optical data, the strengths of both datasets can be leveraged to improve crop mapping accuracy. Various fusion techniques, such as data fusion algorithms or machine learning approaches, can be applied to integrate SAR and optical information effectively.

These days, combining satellite data from multiple sensors is an effective approach to improve the accuracy of crop mapping. Different sensors have different spatial and spectral resolutions, which can complement each other and provide more comprehensive and accurate information on crop distribution, growth, and health status.

Data fusion techniques are an approach to combining satellite data from multiple sensors. This technique involves integrating data from different sensors to produce a single composite image that combines the advantages of each sensor. The fused image can provide more detailed and accurate information on crop type, growth, and crop health status than each individual sensor alone. A study conducted by Q. Wang *et al.*, (2017), presented a method for data fusion of Landsat-8 and Sentinel-2 satellites for crop mapping and monitoring. Which is more accurate in terms of accuracy in comparison with individual satellites. The study concluded that the fusion of multispectral data from Sentinel-2 and Landsat-8 can improve the accuracy and reliability of crop mapping (Xiong *et al.*, 2017; Blickensdorfer *et al.*, 2022). The development of these methods has led to significant improvements in the quality of satellite data. Due to the emergence of complementary satellite sensors, such as Sentinel-1, Sentinel-2, and PlanetScope, has allowed for better data coverage and reduced the issues of data quality. The combination of these sensors provides a more comprehensive view of the Earth's surface, enabling more accurate and reliable data for various applications, including LULC mapping, urban planning, disaster management and crop mapping (Yan *et al.*, 2021).

2.2.3.1 Approaches/methods used for combining multi-sensor datasets

There are several methods commonly used for combining multi-sensor datasets, enabling the integration of information from different sensors. Here are some widely employed approaches:

- **Statistical Fusion:** This method utilizes statistical techniques such as regression models, principal component analysis (PCA), or multivariate analysis to combine data from multiple sensors. The statistical models help in identifying correlations and dependencies among the datasets, allowing for the creation of a unified representation.

- **Sensor-specific Fusion:** This method takes into account the characteristics and limitations of each sensor. It involves applying sensor-specific algorithms or calibration techniques to harmonize the datasets and reduce inconsistencies or biases between them. This ensures compatibility and improves the accuracy of the combined information.
- **Fusion at Feature Level:** This approach involves extracting relevant features from each sensor's data and combining them to create a feature vector that represents the combined information. Feature extraction techniques, such as wavelet transforms or Fourier analysis, are commonly used for this purpose.
- **Data-driven Fusion:** This method utilizes machine learning algorithms to integrate multi-sensor datasets. Techniques like ensemble methods, neural networks, or deep learning models can learn patterns and relationships in the data and produce fused outputs that leverage the strengths of each sensor (Table 2.2).

Nowadays, the use of ML algorithms to combine data from multi-source satellite data for crop mapping is becoming increasingly common. Several ML algorithms were applied on “multi-spectral and multi-temporal satellite images to derive crop classification models” (Viskovic *et al.*, 2019). Recently, LiDAR, Sentinel-2 and aerial imageries used to map crop types using machine learning, where “the combination of all three datasets proved to be the most effective at differentiating between the crop types, with RF algorithm providing the highest overall accuracy of 94.4%” stated Viskovic *et al.*, 2019. Moumni *et al.*, (2021) used “high spatiotemporal resolution Sentinel-1 and Sentinel-2 satellite images and three machine learning classifier algorithms artificial neural network (ANN), support vector machine (SVM), and maximum likelihood (ML)” to overcome the challenges of heterogeneity for crop types mapping. The results showed that combining images from different sensors improved crop-type classification performance (accuracy) compared to using optical or SAR data alone, with an overall accuracy of 89% and Kappa of 0.85. Another study proposed a method for large-scale crop mapping using discrete grids and machine learning to integrate GaoFen-1 and Sentinel-2 imagery. That method achieved good performance with high accuracy, and discrete grid technology can improve processing efficiency for large-scale remote sensing data (Yan *et al.*, 2021).

Now, it is observed that machine learning algorithms are an innovative approach to combining satellite data from multiple sensors, which can be used to analyze and integrate data from different sensors and extract features that are relevant to crop mapping.

Table 2.2: Different studies on crop mapping and approach used

Author	Paper title	Approaches	Satellite image	Location	Accuracy
Tufail <i>et al.</i> , 2021	“A machine learning approach for accurate crop type mapping using combined SAR and optical time series data”	Random forest algorithm	Sentinel-1 and Sentinel-2	Punjab, Pakistan	97% overall accuracy
Fatholouloumi <i>et al.</i> , 2022	“An Innovative Fusion-Based Scenario for Improving Land Crop Mapping Accuracy”	RF, SVM and ANN algorithm with fusion techniques	Sentinel 1, Sentinel 2, and Landsat-8 imagery	Ontario, Canada	91% overall accuracy
Lin, Chenxi, <i>et al.</i> , 2022	“Early- and in-season crop type mapping without current-year ground truth: Generating labels from historical information via a	RF classifier	Landsat-8 and Sentinel-2 data	NE China	85% overall accuracy

Author	Paper title	Approaches	Satellite image	Location	Accuracy
	topology-based approach”				
Feng Gao and Xiaoyang Zhang 2021	“Mapping Crop Phenology in Near Real-Time Using Satellite Remote Sensing: Challenges and Opportunities ”	Curve-based and trend-based methods	Harmonized Landsat and Sentinel-2 (HLS)	Global	Not reported
Yan <i>et al.</i> , 2021	“Large-scale crop mapping from multi-source optical satellite imageries using machine learning with discrete grids”	RF and SVM algorithm	Sentinel-2A/B MSI (10 m resolution), Gaofen-1 WFV (16 m resolution), Gaofen-6 WFV (16 m resolution), Ziyuan-3A MUX (5.8 m resolution), Ziyuan-3B MUX (5.8 m resolution)	Sanjiang Plain in China	Overall accuracy achieved- 86 % in 2017 and 88 % in 2018
Moumni <i>et al.</i> , 2021	“Machine Learning-Based Classification for Crop-Type Mapping Using the	Artificial Neural Network (ANN), Support Vector Machine (SVM), and	Fusion of high spatiotemporal resolution Sentinel-1 and Sentinel-2 satellite images	semi-arid area of Morocco	Overall accuracy of 89%

Author	Paper title	Approaches	Satellite image	Location	Accuracy
	Fusion of High-Resolution Satellite Imagery in a Semiarid Area”	Maximum Likelihood (ML)			
Meng <i>et al.</i> , 2021	“Deep learning-based crop mapping in the cloudy season using one-shot hyperspectral satellite imagery”	Three Convolutional Neural Network (CNN) Model	hyperspectral data, cloud-free S2 images	Hubei province of China	Convolutional neural network algorithms with one-shot hyperspectral imaging and data augmentation techniques achieved an accuracy of more than 94%

2.3 Image Classification

Remotely sensed satellite image classification is the process of assigning land cover categories (or classes) to image pixels based on their spectral, spatial and temporal properties. Remote sensing images are acquired by sensors mounted on satellites, aircrafts or drones that capture the electromagnetic radiation reflected or emitted by the Earth’s surface and atmosphere. It is useful for various applications such as LULC mapping, environmental monitoring, disaster management, urban planning, resource exploration and crop mapping.

2.3.1 Types of Image Classification Techniques

There are two main types of image classification techniques of RS technology such as unsupervised and supervised.

a) Unsupervised Classification

Unsupervised classification is a technique that does not require any prior knowledge or training samples from the user. It groups pixels into clusters based on their spectral similarity using algorithms such as K-means or ISODATA (Kumar *et al.*, 2022). The user then assigns meaningful labels to each cluster based on their visual interpretation or ancillary data. Unsupervised classification is an easy and fast way to segment an image, but it may not produce accurate or consistent results due to the variability of spectral signatures within and between classes (Yan *et al.*, 2021).

b) Supervised Classification

Supervised classification is a technique that requires the user to provide representative samples (or training areas) for each LULC class of interest. The algorithm then uses these samples to learn the spectral signatures of each class and applies them to the entire image using a classifier such as maximum likelihood, minimum distance or ANN, RF, SVM etc. Supervised classification can produce more accurate and reliable results than unsupervised classification, but it requires more user input and may be affected by the quality and quantity of training samples (Gao & Liu, 2014).

2.3.2 Two different ways of Classification

A. Pixel Based Classification

B. Object Based Classification

A. Pixel Based Classification

Pixel based classification is a technique that uses only the spectral information of each pixel, such as the intensity or color, to assign it to a class. It does not consider the spatial relationship between pixels or the shape and size of the objects. Pixel based classification is simple and fast, but it may produce noisy or inaccurate results due to

mixed pixels, spectral variability and spatial heterogeneity (Whiteside & Ahmad, 2005; Duro *et al.*, 2012).

Pixel based classification can be performed using various algorithms, such as maximum likelihood, minimum distance, SVM and ANN. These algorithms use different criteria to measure the similarity or distance between pixels and classes. Pixel based classification can be applied to any type of remote sensing images, regardless of their spatial resolution or number of bands. However, pixel based classification may not be able to capture the semantic meaning or context of the land cover classes, especially for high resolution images that contain complex and diverse features (Yan *et al.*, 2021).

B. Object Based Classification

Object based classification is a technique that uses both the spectral and spatial information of a group of pixels, called objects or image objects, to assign them to a class. It considers the shape, size, texture and context of the objects, as well as their spectral characteristics. Object based classification is more complex and time-consuming, but it may produce more accurate and consistent results by overcoming some of the limitations of pixel based classification (Duro *et al.*, 2012; Kotaridis & Lazaridou, 2021).

Object based classification can be performed using various algorithms, such as segmentation, watershed or convolutional neural networks. These algorithms use different criteria to group pixels into objects based on their homogeneity or heterogeneity. Object based classification can also use different features to describe the objects, such as mean, standard deviation, compactness or entropy.

Object based classification can be applied to high-resolution images that contain rich spatial information and require more detailed analysis. However, object based classification may not be suitable for low resolution images that have poor spatial information and require more general analysis (Gao & Liu, 2014).

2.3.3 Comparison and Applications methods

Pixel and object based classification have different advantages and limitations depending on the characteristics of the remote sensing images and the objectives of the analysis. Pixel based classification is more suitable for low-resolution images that have simple and homogeneous features and require less user input and processing time. Object based classification is more suitable for high-resolution images that have complex and heterogeneous features and require more user input and processing time.

Pixel and object based classification have different applications depending on the type and scale of the land cover classes. Pixel based classification is more suitable for mapping broad-scale classes that have distinct spectral signatures, such as water, vegetation or urban areas. Object based classification is more suitable for mapping fine-scale classes that have similar spectral signatures but different spatial patterns, such as crops, forests or buildings.

2.3.4 Challenges and Opportunities

Remote sensing image classification faces several challenges due to the complexity and diversity of the Earth's surface and atmosphere, as well as the limitations and variations of the sensors and platforms. Some of the common challenges are:

- The presence of noise, clouds, shadows, haze and atmospheric effects that degrade the quality and consistency of the images.
- The high dimensionality, redundancy and correlation of the spectral bands that increase the computational cost and complexity of the classification algorithms.
- The lack of ground truth data or reference maps that are necessary to train and validate the classification results.
- The dynamic nature of the land cover classes that changes over time due to natural or human-induced factors.
- The scale mismatch between the spatial resolution of the images and the size of the land cover objects or phenomena.

To address these challenges, remote sensing image classification can benefit from some of the recent advances and opportunities in the field, such as:

- The development of new sensors and platforms that provide higher spatial, spectral and temporal resolution, as well as wider coverage and accessibility of the images.
- The application of machine and deep learning techniques that can learn complex and nonlinear features from large amounts of data without requiring much user intervention or domain knowledge.
- The integration of multisource and multitemporal data can capture complementary and dynamic information from different sensors and periods.
- The use of semantic web technologies and ontologies can facilitate the interoperability and standardization of the data and metadata across different sources and domains.
- The involvement of crowdsourcing and citizen science can leverage the collective intelligence and participation of the public for data collection, annotation and validation.

Several studies have compared pixel and object based classification for crop mapping using various types of remote sensing images. For example, Castillejo-Gonzalez *et al.* (2009) found that “an object-based method out-performed five pixel-based supervised classification algorithms (parallelepiped, minimum distance, Mahalanobis Distance Classifier, Spectral Angle Mapper, and MLC) for mapping crops and agro-environmental associated measures using Landsat TM images”.

Belgiu *et al.* (2018) compared “pixel-based and object-based time-weighted dynamic time warping analysis for mapping crops using Sentinel-2 images”. They concluded that object-based analysis achieved higher accuracy than pixel-based analysis by exploiting both spectral and spatial information.

2.4 Area/ Acreage Estimation

Crop yield and production estimation is an essential component of economic planning in agricultural sector (Gallego *et al.*, 2014). It involves two main branches: crop area and yield estimation. Whereas crop area estimation is considered relatively easier than crop yield estimation (Craig and Atkinson, 2013). Many factors create complications in crop area computation such as small scattered and diversified cropping patterns, complex physiography, extended sowing process, field size, changes in cropping pattern, mixed cropping system with phenological differences, and short-duration crops.

A detailed review about the techniques that are used for the calculation of crop area was carried out by Craig and Atkinson (2013). Crop area estimation is conventionally carried out by comprehensively recording all farms or using samples that are collecting in the field. Following sampling methods namely, Area Frame Sampling (AFS), farm list sampling and multiple frame sampling is combination of first two sampling methods. In few instances, area estimation can be done based on professional judgement which results in qualitative/ voluntary crop reporters. Among other sources, “crop area estimate can be learnt from administrative surveys, crop processing units (for example, cotton or jute mills) and markets” (Ray, S. & Neetu. 2017).

Since conventional method uses professional expertise, it may induce human bias in final area estimate. Further, conventional area estimation methods are also time-consuming, costly and tedious. For mountainous terrains, conventional method have proven many difficulty.

Remote sensing tools can be used to negate many issues that are arising in conventional area estimation techniques. Since last few decades, crop area estimate using remote sensing tools have either been carried out directly or indirectly utilizing offering support to area sampling schemes. Some of the important advantages of remote sensing includes temporal capability, synoptic coverage, multispectral and multiresolution imageries pertaining to LULC in addition to crop differentiation.

Crop area estimation using remotely sensed satellite data began in the early 1970s. In 1971, “the Corn Blight Watch Experiment was jointly carried out by the U.S. Department of Agriculture (USDA), the National Aeronautics and Space Administration (NASA), and several universities” (Sharples, 1973; Ray, S. & Neetu. 2017). The objective of this experiment was to develop a method for detection and monitoring crop using satellite datasets. The experiment involved the use of Landsat satellite imagery to monitor the growth and health of crops in the United States. Some studies conducted in the 19th century to assess the capabilities of remotely sensed satellite datasets for crops inventory (MacDonald, 1984) such as the “Crop Identification Technology Assessment for Remote Sensing (CITARS) and the Large Area Crop Inventory Experiment (LACIE)” (MacDonald, 1984; Ray, S. & Neetu. 2017). The LACIE experiment helped pave the way for further research and development in use of RS data for a variety of applications in agriculture.

In recent years, advancements in remote sensing technology have led to the development of new sensors and platforms like unmanned aerial vehicles (UAVs) and hyperspectral sensors. These new technologies have further improved the accuracy and resolution of RS dataset for crop area estimation.

Accurate crop area estimation is an essential first step in determining crop yields. Several methods are used to estimate crop area, including remote sensing, ground surveys, and statistical sampling. Remote sensing has emerged as a powerful approach for crop area estimation, with satellite-based platforms providing valuable information on LULC.

Land cover area estimation is essential for monitoring and managing natural resources, understanding environmental changes, and assessing the impacts of human activities on ecosystems. Traditionally, land cover areas have been estimated/obtained through design-based approaches that use probability or random reference sampling. However, this method can be costly and requires an effective sampling design to ensure that the estimates are representative of entire study area (Zhang *et al.*, 2019; Pengra *et al.*, 2020).

In recent years, Earth-Observation-based mapping approaches have become increasingly popular for land cover area estimation. These approaches use remote sensing data, such as satellite imagery, to classify land cover types and estimate their areas. This approach does not require a huge reference sample, which can save time and resources too. Although, the lack of a reference sample can lead to bias in the estimates, particularly if the remote sensing data do not capture all the land cover types or if they are affected by atmospheric conditions or other factors that affect the accuracy of the classification (Kleinewillinghöfer *et al.*, 2022). However, there are limitations to using remote sensing for crop area estimation. These include issues related to cloud cover, sensor resolution, and data processing. In addition, differences in crop phenology, planting practices, and cultural practices can lead to difficulties in accurately identifying and mapping crop types (Kumar *et al.*, 2022).

To overcome these limitations, the use of remote sensing products in combination with reference samples is becoming popular nowadays (Stehman & Foody, 2019; Centre *et al.*, 2011). These hybrid methods can reduce sampling efforts and provide a more effective method to estimate land cover areas. The reference samples can be also used to calibrate the RS data, to validate the classification results, and estimate the accuracy of the estimates (Singh *et al.*, 2020). This hybrid method can use different types of reference samples, including ground-based surveys, aerial photography, or other remote sensing data with higher spatial resolution.

Nowadays, land cover is often extracted from remotely sensed data using a classification analysis, which involves grouping pixels or image segments with similar spectral characteristics into distinct land cover classes (Stehman & Foody, 2019). The resulting land cover map is a thematic map that shows the spatial distribution of LULC in the study area. This approach is widely used for land cover mapping and has been applied in various applications.

Several studies have been conducted to map and estimate crop acreage using remotely sensed data at varying spatial resolutions (coarse to fine spatial resolution data). The Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat,

Sentinel-1, and Sentinel-2 satellites have been widely used in these studies (Mondal *et al.*, 2014; Peng *et al.*, 2011). In some studies, the MODIS data have been used to estimate crop acreage at a regional scale, while Landsat data have been used to estimate crop acreage at a field scale. Sentinel-1 and Sentinel-2 data have been used to estimate crop acreage at a regional to field scale. These studies have demonstrated the utility of RS dataset for crop mapping and acreage estimation.

In India, the FASAL (Forecasting Agricultural output using Space, Agro-meteorology and Land-based observations) program is a joint initiative by the Indian Space Research Organization (ISRO) and the Ministry of Agriculture and Farmers Welfare. The program aims to provide timely and accurate information on crop area, crop production, and yield estimation using remote sensing, satellite-based observations, and ground-based data. The program uses a combination of remote sensing data from satellites such as MODIS, Landsat, and Sentinel-2, as well as ground-based weather data, to provide crop forecasts and yield estimations. The FASAL program has been implemented in several states in India and has been successful in improving crop planning and management (Arumugam *et al.*, 2021; Latwal *et al.*, 2019).

Nowadays, remote sensing satellite data-based studies have become gradually popular for crop studies to provide accurate and timely information on crop area, yields, and other crop-related parameters. However, studies that use both remote sensing and machine learning approaches for crop area estimation and mapping are still limited.

2.5 Crop Yield and Production Estimation

Crop yield estimation is a challenging task, due to complexity of factors (climate variability, soil conditions, pests and diseases, and cultural practices) that affect crop health, growth and production. Consequently, it requires a combination of field measurements, crop modeling and remote sensing data. Crop models can be used to simulate the impacts of different climatic variables on crop production, while field measurements such as crop biomass and leaf area index provide crucial information on crop health, growth and production.

The increasing demand for crop insurance at the plot-level is increasing the demand for crop area and yield estimation too. But traditional methods such as field surveys and crop cutting experiments (CCE) are still extensively used to estimate crop area and yield, which are costly and time-consuming. Field surveys require significant resources, including personnel, equipment, and transportation, while crop-cutting experiments can only provide estimates for small, localized areas. Therefore, there is a need for more efficient and accurate methods of estimating crop acreage and yield, which can be provided by advancements in remote sensing technology (Hudait & Patel, 2022).

Nowadays, remote sensing technology has emerged as a valuable tool, which is commonly used to track changes in crop growth and development, as well as to identify stress factors that may impact yield. Over the years, advancements in remote sensing technology have enabled more accurate and periodic data collection on crop area and yield. These technologies allow us to timely monitor crop health, growth and production and identify patterns and trends that may not be possible with other methods.

The method of crop yield estimation based on the analysis of the relationship between the vegetation index value obtained from RS data and the actual crop yield (CCE data) on the ground is becoming popular nowadays (Mehdaoui & Anane, 2020; Iniyana *et al.*, 2022). Vegetation indices are popular tools for monitoring the vegetation, greenness, productivity and health of crops.

Some previous studies suggest that various spectral indices have been developed using remotely sensed satellite data to estimate vegetation properties such as NDVI (Rouse Jr *et al.*, 1973), GNDVI (Gitelson *et al.*, 1996), LSWI (Xiao *et al.*, 2002), EVI (Huete *et al.*, 2002), which have been utilized in many studies for crop mapping and yield estimation. Among these indices, NDVI is the most popular vegetation index, which is widely used in empirical regression models to predict crop yields (Franch *et al.*, 2019). Prasad *et al.* (2006) conducted a study to assess and predict crop yields using a piecewise linear regression method with breakpoint. In their analysis, they considered several important factors including NDVI, soil moisture, surface temperature, and

rainfall data. Ferencz *et al.* (2010) used “two methods for yield estimation of different crops in this study using remote sensing satellite data. Wherein, first method was developed for field-level estimation and involved the selection of reference crop fields using Landsat Thematic Mapper (TM) data for classification.

The second method presented used only NOAA - AVHRR and officially reported county-level yield data. The results showed that the developed method was stable and accurate for operational use in county-level yield estimation.” Ren *et al.* (2011) demonstrate the potential of using remote sensing approaches for monitoring agricultural systems and assessing their vulnerability to drought. The results of study indicate a significant negative correlation between agricultural production and occurrence of meteorological droughts, highlighting the impact of drought as a stressor to agricultural production systems.

A semi-automated approach has been developed using Indian Remote Sensing (IRS) satellite data and vegetation indices to identify and monitor annual cropping patterns (Mondal *et al.*, 2014). Recently, Rao *et al.* (2021) demonstrated the potential of high-resolution satellite imagery and SVM for identifying major crops in eastern India, achieving 85% classification accuracy. A study examines the relationship between satellite data from various sensors, soil/relief parameters and crop yield data. Wherein, the correlation is higher when the crop and field are spatially heterogeneous and delineated at the correct phenological timing. They suggest that the higher resolution and additional red edge spectral band are advantages (Vallentin *et al.*, 2022).

Remote sensing based modern yield estimation techniques offer the advantages of accuracy, efficiency and dynamic estimation compared to the traditional manual field survey method. With the development of Internet, cloud computing platform, big data, high resolution remotely sensed satellite imageries, and geospatial datasets, it has been widely utilized for crop yield estimation (Wei *et al.*, 2018). Over the years, multiple techniques have been devised to estimate crop yield using remotely sensed satellite data (Jiang *et al.*, 2014 and Singh *et al.*, 2020). However, some methods have encountered

challenges in accurately predicting crop yields due to the unavailability of satellite datasets caused by climatic conditions like cloud cover (Awad *et al.*, 2019).

2.6 Flood

Floods are a recurrent phenomenon in many parts of the world (Jain *et al.*, 2023). India is a country which is prone to frequent floods, causing massive socio-economic and loss of lives. Despite significant investment and continued flood control efforts, the problem of flooding remains a major challenge for the country, especially for the Indian state of Bihar. Where, nearly half of the region (North Bihar) gets flooded every year due to the overflow of major rivers during the rainy season. This region is prone to severe flooding during the monsoon season due to the overflow of several rivers that pass through the region. These rivers are Mahananda River, Koshi River, Bagmati River, Budhi Gandak River and Gandak River, which all originate in Nepal. which makes it highly susceptible to flooding every year. The region receives heavy rainfall during the monsoon season, which causes the rivers to overflow and inundate the surrounding areas. The Kosi, Gandak, and Bagmati rivers are the major rivers that cause flooding in the state. Floods have significant socio-economic impacts in the region.

This region (North Bihar) faces severe devastation in the form of crops, livestock, and infrastructure, which leads to significant economic losses. Moreover, floods result in the displacement of people, loss of property, and loss of life. According to the National Crime Records Bureau (NCRB) report 2021, Bihar has recorded the highest number of flood-related deaths in the country in past decade. In 2021, the state witnessed 351 fatalities, which is the highest among all Indian states. To address the problem of floods in Bihar, there is a need for comprehensive and integrated approach that incorporates climate change adaptation strategies, improved early warning systems, modern technology and effective community participation.

2.6.1 Remote Sensing Technology for flood management

In this context, Remote sensing satellite-based approaches can provide critical information to help authorities take preventive measures during flood events. Remote

sensing technology can be used to monitor the water levels of rivers and reservoirs, detect the extent of floods, and identify areas that are at the highest risk of flooding. This information can help prioritize relief and rescue operations and ensure that resources are directed to the areas that need them the most.

Determining the extent of flooding is crucial for effective disaster management. In-situ and remote sensing methods are being widely used for flood mapping. In-situ method involves collecting data directly from the flood water level on the ground. This can be done using water level gauges, measurement tapes and other instruments. In this context, remote sensing methods rely on data obtained from aerial or satellite sensors, which can cover larger areas (synoptic view) in a shorter amount of time.

According to Jensen (2013), remote sensing is a valuable tool for flood mapping because it allows for data collection in a timely and cost-effective manner. Remote sensing data can be used to detect and monitor floodwater levels, as well as identify flood-prone areas. The data can also be used to estimate the volume of water in flooded areas and track changes over time. Satellite imagery is capable of estimating the extent of flooding over vast geographic areas, including areas that are difficult to access. Additionally, it allows continuous monitoring of flood events at regular intervals, as noted by Bhatt and Rao (2016). Tanguy *et al.* (2017) have also highlighted that satellite-based imagery is the ideal approach for determining the extent of flood-affected regions due to its synoptic view. But, it may not be possible to avoid flood risks or prevent their occurrence. However, it is entirely possible to limit their impact and minimize the damage (Elkhrachy, 2015).

2.7 Research Gap

In recent years, several studies have been conducted for crop mapping. Wherein, remote sensing satellite data, geospatial technology, and machine learning approaches have transformed the agricultural sector, which provides new opportunities for sustainable and efficient crop area, yield and production mapping and monitoring. Moreover, some of the major gaps are as follows;

- There is no any specific study conducted to identify the maize (🌽) crop acreage with integration of high-resolution satellite (📡) data (Sentinel-2 and Planet-NICFI) in Bihar.
- Traditional methods of crop yield estimation are costly and time-consuming. Therefore, there is a need of modern approaches for maize crop yield estimation, which require remote sensing satellite data, CCE data, statistical and machine learning approaches.
- Recurring flood is a common phenomenon in Bihar, which damage people's lives, infrastructure and agriculture sector every year. Despite this, limited studies have been conducted on flood's impact on agriculture.
- Synthetic aperture radar (SAR) data have more potential than optical datasets for flood mapping because of their all-weather sensing capability. However, timely processing of large amounts of SAR datasets is a major challenge for flood mapping and monitoring.
- Numerous studies have been carried out on flood disasters in Bihar, which require high-end computer infrastructure such as local storage and commercial GIS software to analyse the complex data sets (optical and SAR remote sensing satellite datasets) involved in flood studies. But, no any robust model has been developed to identify the flooded area within a short time of period.
- In recent years, several studies have been carried out to studies crop phenology by using satellite data. But, no specific studies have been conducted in India to assess the impact of floods on crop acreage, yield, pattern and production.

Floods are a major cause of huge damage in the form of loss of life, property, agriculture and other required natural resources all over the world. In India, about half of Bihar (North Bihar) face severe flood frequently due to overflow of major rivers during monsoon season. The Kosi river flood is well known for devastating and frequent changes in its course. On the other hand, it is also providing nutrients for the agricultural land (specially for rabi crops). Therefore, a rapid and robust flood extent demarcation and affected area mapping system is needed in North Bihar, India to prioritise the relief and subsidy.

In this context, JavaScript code has been edited, tested and developed for the processing of enormous (big data) datasets hosted on the GEE cloud computer platform. This JavaScript code is capable to robust flood mapping and monitoring using Microwave (SAR) satellite datasets at large scale within a short time of period. Integrated satellite imagery of Sentinel-2 & PlanetScope, and Machine learning (ML) approaches have been used for maize crop mapping and acreage estimation in North Bihar, India.

In this study, crop yield predictive model has been also developed using Crop Cutting Experiment (CEE) & remotely sensed datasets and district-wise maize crop yield and production have been estimated. We expected that this tested algorithms, edited and developed web-based JavaScript code can be applied anywhere in the world for flood extent demarcation, precisely crop mapping and the impact of flood on acreage and yield within a short time of period. In addition, crop yield predictive model has been also developed for district/field-wise maize crop yield and production estimation.

3.1 Data Used

In this study, the major emphasis is on north Bihar which is severely affected by flood during the monsoon season. Therefore, various datasets were required to perform the analysis as shown in Table 3.1.

Table 3.1: Datasets used in the study

Data	Resolution	Source	Duration	Aim
Sentinel-1 SAR	5 x 20 Metres	ESA	June to October (2020 & 2021)	<ul style="list-style-type: none"> • Flood Extent mapping
Sentinel-2 (MSI) Optical	10 Metres		March (2020 & 2021)	<ul style="list-style-type: none"> • Land Use Mapping • Extract REI, NDVI, GNDVI, LSWI and EVI for Yield estimation Crop Yield estimation
PLANET-NICFI (PlanetScope)	4.77 Metres	NICFI	January to May (2022 & 2023)	<ul style="list-style-type: none"> • Maize Crop Acreage estimation
Handheld GPS/ GPS enabled Smart Phone	Ground Truth (GT) data collection	Field Survey	February to March (2022 & 2023)	Sample/ input/ training data for <ul style="list-style-type: none"> • crop differentiation • acreage estimation and validation
Handheld GPS/ GPS enabled Smart Phone	Crop Cutting Experiment (CCE)	Field Survey	April to May (2022 & 2023)	Input data or training data for <ul style="list-style-type: none"> • yield estimation • predictive model development
Shuttle Radar Topography Mission (SRTM)	30 Metres	NGA and NASA	2000	<ul style="list-style-type: none"> • Terrain correction

Note: MSI- Multispectral Instrument, NICFI- Norway's International Climate and Forests Initiative, and ESA- European Space Agency

3.1.1 Sentinel-1 SAR

SENTINEL-1 (S-1) is a radar imaging mission developed by the European Space Agency (ESA) as part of the Copernicus programme. Launched in 2014, S-1 is equipped with a C-band (instrument at 5.405GHz) Synthetic Aperture Radar (SAR) that offers high-resolution imagery of the Earth's surface in all-weather (can penetrate clouds), day and night conditions with a wide swath of 250 km. The mission consists of two identical satellites that orbit the Earth, covering the entire planet. Its revisit period is twelve days with single satellite and six days with double satellite. SENTINEL-1's capabilities have proved particularly useful in monitoring and responding to natural disasters such as floods, forest fires, and oil spills, as well as in mapping sea ice, tracking shipping, and monitoring crop phenology. The data collected by SENTINEL-1 is freely available to scientists, researchers, and the public, enabling a better understanding of our planet and its changing environment.

The four unique modes of Sentinel-1 are as follows;

- i. Interferometric Wide Swath (IW), provides a broad coverage of the Earth's surface with a swath width of up to 250 km and having 5 x 20 m spatial resolution.
- ii. Extra Wide Swath (EW), provides an even wider coverage with a swath width of up to 400 km and 25 x 100 m spatial resolution (3-looks).
- iii. Stripmap (SM), provides high-resolution imagery with a swath width of up to 80 km and 5x5 m spatial resolution.
- iv. Wave mode (WV) is designed specifically for monitoring ocean waves, currents, and wind direction, and has a swath width of up to 20 km and 5 x 20 m spatial resolution.

The combination of these four modes allows Sentinel-1 to capture a wide range of imagery for variety of applications, including disaster response, land and ocean monitoring, and scientific research (Access ESA Website/ SciHub; <https://scihub.copernicus.eu>).

Sentinel-1 data products

- **Raw Level-0 data:** It is the original unprocessed data, and are primarily intended for specialized scientific usage.
- **Level-1 Single Look Complex (SLC) data:** It comprises complex imagery with amplitude and phase information and is systematically distributed but limited to specific relevant areas.
- **The Ground Range Detected (GRD) Level-1 data:** It is systematically distributed and includes multi-looked intensity-only data, making it suitable for various applications like land cover mapping, disaster management, and forest monitoring.
- **Level-2 Ocean (OCN) data:** It provides geophysical parameters of the ocean, such as ocean currents and wind speed. The OCN data products are systematically distributed and are used for a variety of applications, including coastal monitoring and maritime traffic management.

In the present study, GRD data is used to map flood extent, which consists of four band combinations. Depending on the instrument's polarization settings, each scene contains either one or two of the four possible polarization bands. In which, the Interferometric Wide swath (IW) mode has been used due to its conflict-free nature and the availability of both vertical transmit, vertical receive (VV) and vertical transmit, horizontal receive (VH) polarization. The four possible combinations are single band VV or HH, and dual-band VV+VH or HH+HV. SAR is a powerful imaging technology that uses microwave signals to create high-resolution images of the Earth's surface, even in adverse weather and lighting conditions. The IW mode of Sentinel-1 SAR provides a wide swath, which enables large area coverage with high temporal resolution. Additionally, combination of VV and VH polarization enhances the interpretation of the radar signals, especially in detecting changes in the target's surface.

3.1.2 Sentinel-2 Optical

Sentinel-2 is a satellite mission developed by the European Space Agency (ESA) as part of the European Union's Copernicus Programme. The primary aim of this

mission is to provide high-resolution, multispectral imagery of Earth's land and coastal regions. These satellites are equipped with multispectral imager (MSI) that can capture images in 13 spectral bands ranging from visible to shortwave infrared with a resolution of 10, 20, or 60 meters, depending on the band (Fig. 3.1).



Figure 3.1: Wavelengths and spatial resolutions of Sentinel-2 satellite (MSI instruments) (Source: <https://sentinels.copernicus.eu>)

This enables the satellites to observe the Earth's surface in high detail, making them useful for wide range of applications, including agriculture, forestry, and land-use monitoring. The revisit frequency of this satellite is ten days with each single satellite and five days with combined/dual satellites. This frequent coverage is particularly useful for monitoring changes over time, such as crop growth and land-use changes.

The Sentinel-2 mission currently consists of two satellites in orbit, Sentinel-2A and Sentinel-2B. Both were launched in 2015 and 2017, respectively. The mission is expected to continue until at least 2027. The data from the mission are freely available to the public through the Copernicus Open Access Hub. Its swath is 290 kilometres and

its orbit is sun-synchronous at an altitude of 786 km with worldwide coverage. This satellite has 10 m spatial resolution (band: 2, 3, 4 and 8), 20 m (band: 5, 6, 7, 8a, 11 and 12) and 60 m (band: 1, 9 and 10). Which is currently being used for land use mapping, crop area, yield and production estimation.

3.1.3 PLANET-NICFI (PlanetScope)

PLANET-NICFI is a collaboration between Planet and Norway's International Climate and Forest Initiative (NICFI) that aims to improve monitoring of tropical forests to combat deforestation and reduce greenhouse gas emissions.

Table 3.2: Band-wise resolution scale and description

Band	Min	Max	Scale	Resolution	Description
B	0	10000	0.0001	4.77 meters	Blue
G	0	10000	0.0001	4.77 meters	Green
R	0	10000	0.0001	4.77 meters	Red
N	0	10000	0.0001	4.77 meters	Near-infrared

(Source: <https://developers.google.com>)

The initiative provides high-resolution satellite imagery or visual basemaps to support NICFI's forest monitoring efforts. It is a constellation of small satellites that offer high-resolution remotely sensed satellite data of Earth's surface. These datasets have spatial resolution of about 5 (4.77m) metres, which can be used to monitor land use change and deforestation in tropical regions (Gorelick *et al.*, 2017). In addition, it is an 8-bit time series mosaic dataset, optimized to reduce the effects of clouds, haze, and other image variability, providing a clear and consistent view of Earth's surface.

This dataset can be freely accessed for non-commercial purposes by researchers and scientists through the GEE, a cloud computing platform. In collaboration with Google and the Norway's International Climate and Forest Initiative (NICFI), PlanetScope's high-resolution composite basemaps for tropical regions have recently become accessible within GEE (Vizzari, 2022). In GEE platform, it is available in four bands viz. Blue, Green, Red, and Near-infrared. Each band has a spatial resolution of

4.77 meters, which is the same as the original imagery captured by Planet's satellites (Planet Team, 2017). This dataset is available as a monthly composite since September 2020. Now, users can access it by signing up and accepting the terms at planet.com/nicfi. In the present study, PlanetScope's monthly mosaic basemaps from January to April of 2022 & 2023 have been used. Presently, the PlanetScope dataset has been used for accurately acreage, yield and production estimation of study area.

3.1.4 Shuttle Radar Topography Mission (SRTM)

The Shuttle Radar Topography Mission (SRTM) data is a high-resolution digital elevation model (DEM) generated from radar data collected by Space Shuttle Endeavour in February 2000. The data provides a highly accurate and detailed representation of the Earth's surface topography, including features such as mountains, valleys, and rivers. The SRTM data has a spatial resolution of approximately 30 meters and covers most of Earth's surface between 60°N and 56°S latitude. The data is available in two versions: the original SRTM dataset, which has a resolution of 1 arcsecond (approximately 30 meters), and the SRTM-3 dataset, which has a coarser resolution of 3 arcseconds (approximately 90 meters). The SRTM data is freely available to the public through a number of online sources, including the United States Geological Survey (USGS) EarthExplorer website (<http://earthexplorer.usgs.gov/>), NASA's Earthdata website, and the CGIAR-CSI SRTM website. The data can be downloaded in a variety of formats, including GeoTIFF, ASCII, and DEM.

3.1.5 Other Datasets

The study have also used IMD (IMD, 2020/2021)/India-WRIS (2020/2021) data for rainfall observations, Population data from Census of India (Census of India, 2011), Global Human Settlement Layer (GHSL) by European Commission (JRC, 2015) and Fatalities data from State Disaster Management Department, Bihar (Anonymous, 2020b).

3.1.6 Ground Truth (GT) data / Reference data Collection

Ground truth data collection is an essential aspect of remote sensing-based research, as it provides accurate and reliable reference data for image analysis. Road

network and LULC images of the study area are prepared before going to the field for ground truth collections, which speeds up the data collection process and helps in saving time. GT data were collected randomly via handheld GPS and smartphone-based applications for identification of crop types and specific crops such as maize, wheat etc. in the study area (Table 3.3; Fig. 3.2 & 3.3; Appendix-A). About 1451 GT points in 2022 and 1327 in 2023 were collected from the study area (Fig. 3.2), which included information like crop name, district name, farm size, geographical location, field photographs and names of adjacent crops. After the collection of GT data, these points (GT) were used to train machine learning algorithms and validate the data during classification/ identification and accuracy validation. About 25 per cent of GT data has been used for accuracy assessment.



Figure 3.2: GT data collection for crop identification

Ground Truth (GT) Data Collection

SI No.	Information collection	Detail
1.	Date of Observation	17/02/2022
2.	Crop Name	Maize
3.	Field ID	269
4.	Size of Crop Field	1 ha
5.	Date of Sowing	Last week of Nov. 2021
6.	Place with District Name	Adampur Mal, Dist-Katihar
7.	Latitude	25.750474
8.	Longitude	87.861726
9.	Altitude	40m
10.	Accuracy	1.8m
11.	Field Photograph No.	37
12.	Data Collector	Himanshu Kumar
13.	Remarks	

Figure 3.3: Sample of GT data collection

3.1.6.1 Field photographs:

Field photographs were also taken in the study area (Fig. 3.4 & 3.5). This photograph helps in visual documentation, capturing the physical characteristics of the study area such as the surrounding environment, land use patterns and other notable observations. It can also be used as visual evidence that can support the findings of the study. Researchers can record and preserve its current status for future comparison, analysis and better understanding of the study area.



Figure 3.4: Field photographs of study area (3A)



Figure 3.5: Field photographs of study area (3B)

3.1.7 Crop Cutting Experiment (CCE) Data

The crop cutting experiment (CCE) data is widely used in agriculture for crop yield estimation, which is a traditional and accurate method to assess crop yields. It is based on the principle of random stratified sampling (so-called smart sampling), where the total area under study is divided into strata, and a sample of plots is randomly selected from each stratum (MNCFC,2020; Chaudhari *et al.* 2019; Tripathy *et al.*, 2022). Wherein, randomly selecting a sample of fields from a large area and physically harvesting the crops from the selected fields to determine the average yield and the total production. The data collected from CCE was used to calculate total crop yield over a large area. Hence, the CCE was performed in the experimental site (Fig. 3.6). Wherein, about 110 to 122 locations of the CCE plots for each year were randomly selected based on vegetation condition (NDVI values), length of growing period and maize acreage extent map of North Bihar.

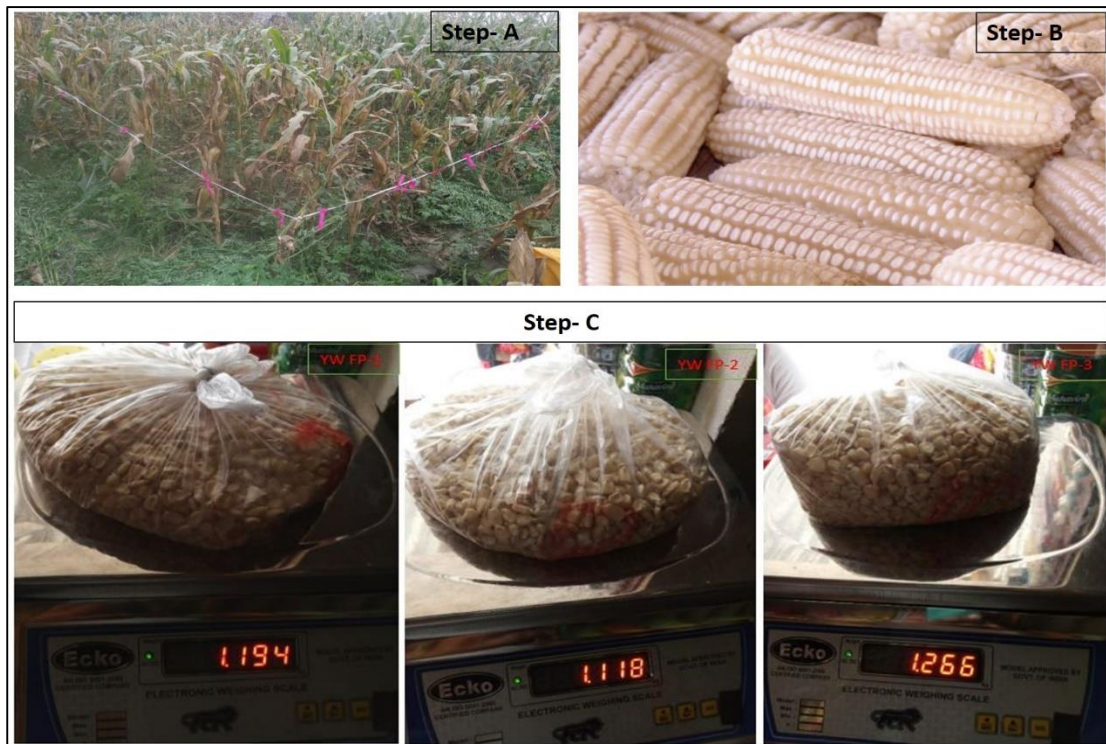


Figure 3.6: Crop Cutting Experiment

The CCE data is often used to validate and calibrate remote sensing datasets for crop yield estimation. It is also used to ground truthing the remote sensing data, which helps in improving the accuracy of the estimation. Apart from that, this CCE dataset

has been utilized as training data for the development of a yield predictive model, average maize crop yield estimation and accuracy testing or model validation. This dataset can be used for various purposes, such as monitoring the productivity of crops, estimating the supply and demand of crops, and making informed decisions about farming practices.

3.1.7.1 Field Photograph of Crop Cutting Experiment

The field photographs of CCEs were taken from the study area, as depicted in Figures 3.7 and 3.8. The CCE involved gathering data on crop yield and productivity by physically measuring and harvesting samples of crops from a selected area within the study site. The collected field photograph can be used as visual evidence of the experiment, showcasing the specific crops, their growth stage, and the sampling process such as experiment's setup and execution. It adds transparency, credibility, and a visual context to the findings and conclusions drawn from the study.

Therefore, some field photographs of the study area have been included in the thesis to clearly understand the geographical context and ground reality of the study site. Figure - CCE FP-1 and CCE FP-2 represent the mature or harvesting stage of maize crops, which is the ideal stage for CCE data collection from the field (Fig. 3.7 & 3.8). Whereas, Figures – YW FP-1, YW FP-2 and YW FP-3 represent the maize crop yield, which is being weighed to estimate the total yield of the entire field (Fig. 3.9).

CCE FP-1



Figure 3.7: Field Photograph of Crop Cutting Experiment (FP-1)

CCE FP-2



Figure 3.8: Field Photograph of Crop Cutting Experiment (FP-2)

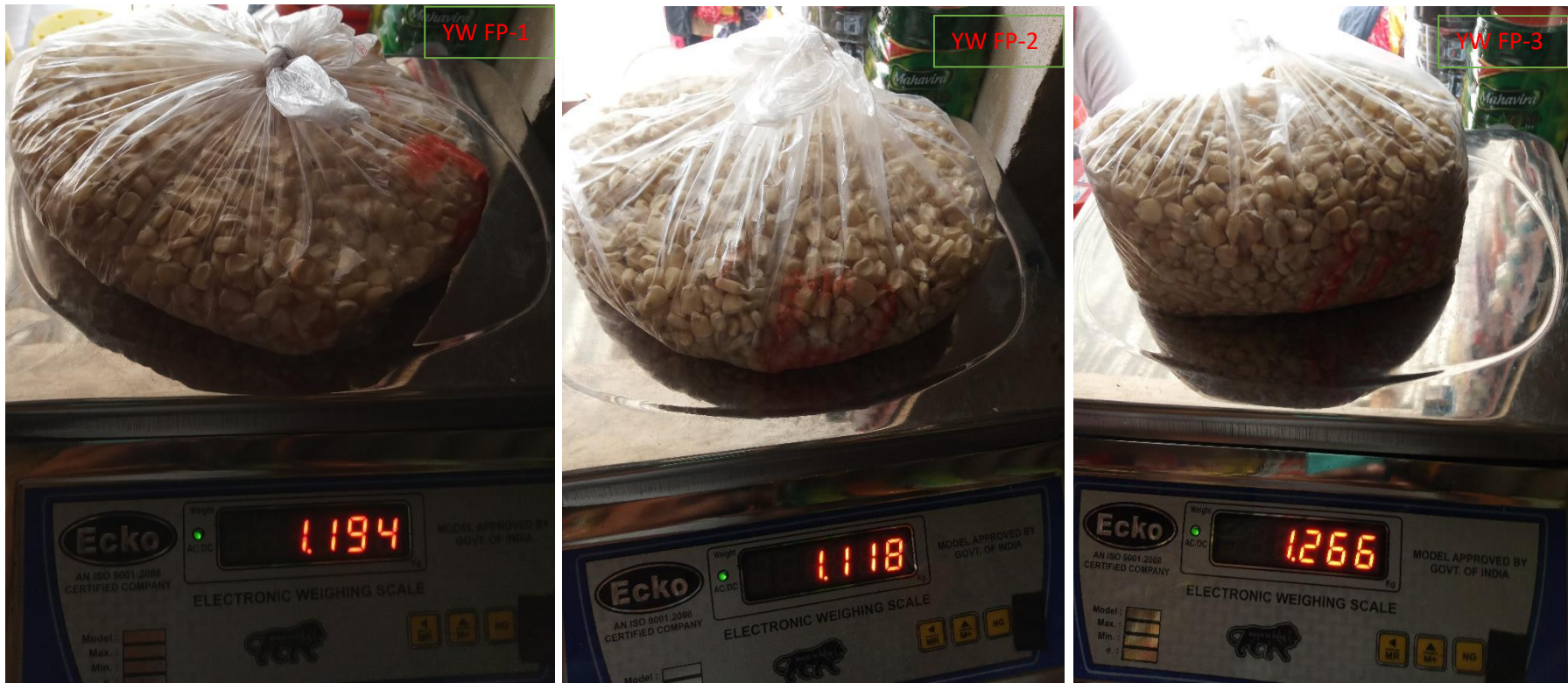


Figure 3.9: The yield of the extracted crop is being weighed to estimate the total yield of the entire area.

Here, CCE was done by randomly selecting several plots. Once the plots were selected, the crops were harvested (Fig. 3.6). After this the crops were extracted, dried and weighed (Ray et. al., 2021). The average yield of the crops harvested was then used to estimate yield of entire field. In the present study, remote sensing satellite was also used to select representative samples for CCEs. By using satellite data to identify areas with similar crop conditions, it is possible to select samples that are more likely to be representative of the overall crop population. This can help to improve the accuracy of CCEs.

Nowadays, CCE is a valuable approach for estimating crop yields. It is being used by governments, farmers, and businesses to make decisions about crop production and marketing. CCE is also used to track changes in crop yields over time.

Here are some of the benefits of using CCE;

- It is a relatively accurate method of estimating crop yields.
- It is a cost-effective method of estimating crop yields.
- It is a quick and easy method of estimating crop yields.
- It can be used to estimate yield of any crop.

CCE is a valuable tool for anyone who needs to estimate crop yields. It is an accurate, reliable and cost-effective method of estimating crop yields.

3.2 Software Used

GIS software allowed for the effective management and integration of spatial data, including maps, satellite imagery, and aerial photographs. It facilitated the creation of detailed and accurate geospatial datasets, which were essential for mapping and visualization purposes. By overlaying various layers of datasets, researchers can be able to identify patterns, relationships, and trends in the geographic context.

In the present study, GIS and digital image processing software have been used to collect, analyze and interpret geographic data and images. The use of the software is as follows.

3.2.1 ArcGIS 10.5

It is a highly used commercial GIS software, which is widely used in both academic and industry. Its wide range of capabilities for data management, analysis and visualization make it a popular choice among professionals in various fields. One of the key strengths of ArcGIS is its robust data management capabilities. It allows users to efficiently organize, store, and manipulate diverse types of geographic data, including maps, satellite imagery, and multi-sensor data. The software provides tools for data integration, cleansing, and quality assurance, ensuring that the information utilized is accurate and reliable. Hence, the software has been used to complete the necessary process and prepare the map layout for the study.

3.2.2 QGIS (Quantum GIS)

The QGIS is an open-source GIS software that is gaining popularity due to its free and open-source license. It offers many of the same features as ArcGIS, but with a different user interface. This software is widely used in various fields, including academia, industry, and government agencies. It can be customized and extended by the user community as per their choice. This fosters innovation and allows for the incorporation of new functionalities and plugins to enhance the software's capabilities. In the present study, this software has been used for data accessing and map preparation purposes.

3.2.3 Google Earth Engine

Google Earth Engine is a cloud-based platform and geospatial analysis tool developed by Google that enables users to access and analyze vast amounts of satellite imagery and geospatial data. It offers a range of features and capabilities that make it a valuable resource for research, monitoring, and analysis. It offers vast collection of satellite imagery and geospatial datasets. Users can access a diverse array of high-resolution imagery from various satellite missions, including Landsat, Sentinel, and MODIS. Additionally, the platform provides access to climate, environmental, and socioeconomic datasets, allowing for comprehensive analysis. It utilizes Google's cloud computing platform to handle large datasets and complex analytical tasks efficiently without the requirement of local storage.

Nowadays, GEE platform is being commonly used for explore the Earth surface, track changes over time, map trends and identify patterns. User can build custom applications using GEE's API to visualize data, analyze trends and share their findings with others. Here, the GEE platform was used to access SAR and optical data and estimate flood extent, land use/land cover mapping and acreage estimation.

3.3 General Methodology

In the present study, a three-step methodology is employed to achieve the objectives of the study. The first step is specifically designed for flood extent mapping and monitoring, while the second step is intended to discriminate maize from other crops and estimate the acreage of maize. The third step focuses on the estimation of maize crop yield and production. A flowchart of the methodology is shown in Fig. 3.10.

1st Steps

Sentinel-1 SAR data has been utilised for flood extent mapping. The LULC map has prepared using sentinel-2 MSI and PlanetScope datasets for extracting flooded agriculture land. All the necessary analysis, such as the pre-processing of Sentinel-1 SAR datasets, was carried out on the GEE cloud platform using the SNAP software package. The GEE platform provides a robust and efficient way to manage and process massive amounts of remotely sensed dataset, while the SNAP software package offers numerous features for SAR data processing and analysis.

The combination of GEE and SNAP allowed for the seamless execution of pre-processing steps and the generation of high-quality SAR images for accurate flood extent mapping. Wherein, thresholding technique has been utilised for identification of inundated pixels. An automated thresholding approaches can automatically detects threshold value to distinguish the water pixels from other pixels without having any training sample datasets (Liang, J et. al., 2020). After that, the pre-flood layer of water bodies was deducted from the obtained flood extent to get the final result. The detailed methodology is given in Chapter-4.

2nd Steps

Remotely sensed satellite datasets have been used to differentiate maize from other crops and estimate the acreage of maize. Here, we used various classification Machine learning algorithms such as CART, RF and SVM classifiers to classify maize from other crops using GT data. Then, the performance of utilized approaches was compared and analysed in terms of overall classification accuracy. This work has been done to provide accurate and timely information on maize acreage and its distribution, which can be useful for crop management and food security planning. The detailed methodology is given in Chapter-5.

3rd Steps

For crop yield estimation, a freely accessible Google's GEE platform has been used to process voluminous dataset of multiple dates NICFI-PlanetScope and Sentinel-2 remote sensing satellite imagery. Because, this platform has the capability to analyse, visualize and access massive datasets such as Landsat, Sentinel, MODIS, PlanetScope etc. achieve without the requirement of a local storage system (Amani *et al.*, 2020) and GIS Software. The Sentinel-2 imagery was harmonized and resampled to the exact resolution as PlanetScope (4.77 m) using the bicubic interpolation function (Keys, 1981) in the GEE platform. A fusion algorithm have also applied to merge Sentinel-2 and NICFI data using Web-based GEE JavaScript code `'var fusion = ee.Image.cat(sentinel2, nicfi);'`. The detailed methodology is given in Chapter-6.

Here, the cumulative flowchart of step-1, 2 and 3 is given for clear understanding of the present study (Fig. 3.10). Wherein, Step-1 has been developed for estimating the extent of flood and its impact, Step-2 for crop acreage estimation and Step-3 for district-wise crop yield and production estimation. In the flowchart, different colors have been used to represent various processing steps. Yellow indicates the processing steps involving GT and CCE data collected using handheld GPS devices. Blue is used for processing steps carried out on the GEE platform, specifically for SAR data processing used in flood extent mapping. Violet represents processing steps related to optical data used for crop acreage and yield estimation. Finally, green is used to denote the final output generated by the adopted method.

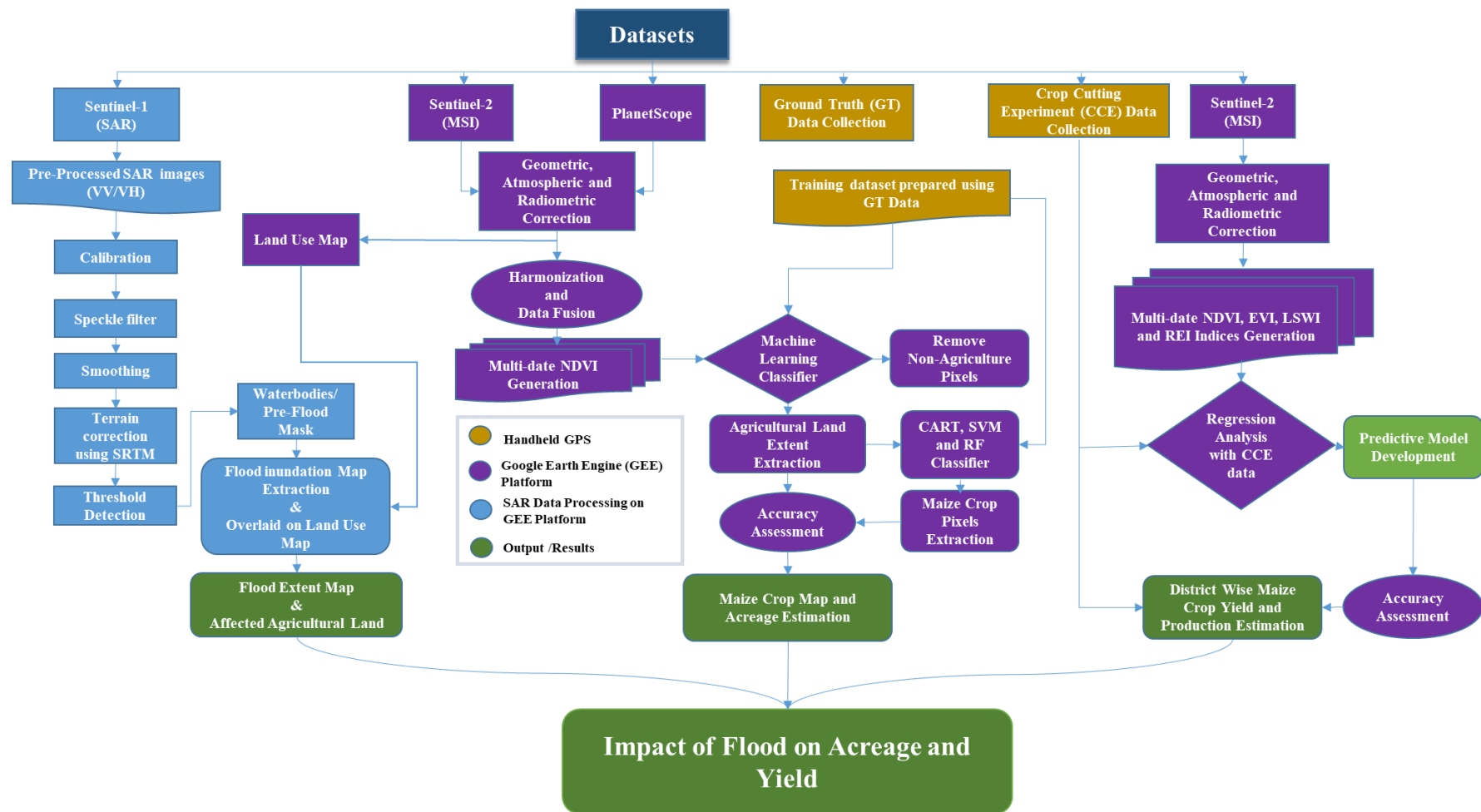


Figure 3.10: Overall flowchart of research methodology

3.4 Concluding Remarks

This chapter explains the dataset used and the methodology adopted for the study. Wherein, A cumulative flowchart of step 1, 2 and 3 is given for a clear understanding of the present study. Here, step 1 is developed for estimating the extent of flood and its impact, step 2 for crop acreage estimation and step 3 for district wise crop yield and production estimation. In the next chapter, flood extent and affected area mapping of the study area is explained in detail.

4.1 Introduction

Flood is a major natural disaster in the Indian state of Bihar. Which has a severe impact on the loss of properties, infrastructures and agriculture every year. Rapid Growth of urban, deforestation, unplanned development and erratic rainfall are the main causes of frequent floods in North Bihar. Bihar is a state in India where about 76 percent of the population is dependent on agriculture, which is severely affected by frequent floods (Kumar *et al.*, 2022). However, floods are considered a regular concern mostly in North Bihar during the monsoon season. Due to overflow of rivers in neighbouring country Nepal, North Bihar has experienced severe floods during the three decades (Freer *et al.*, 2013 and Kumar *et al.*, 2022). But the current COVID-19 pandemic has worsened the situation, where natural disasters and other factors have disrupted the economic and social stability of communities. Floods created significant impact on agricultural sector, which is a major source of livelihood in Bihar. When floods occur, crops are completely destroyed or damaged, making it difficult for farmers to earn a living. Additionally, floods have damaged homes and other infrastructure, making it unsafe or impossible for people to stay in their communities. As a result, many people in North Bihar are forced to migrate to other areas in search of work and income during floods to sustain their lives. It is a difficult and often traumatic experience, as families are separated and individuals are forced to leave their homes and communities behind. Efforts to address the root causes of migration during floods, such as investing in flood-resistant infrastructure and promoting economic diversification, can help to reduce the need for people to leave their homes in search of livelihoods during times of crisis in north Bihar.

To mitigate the impact of floods, an early warning system for floods in Bihar is also highly needed. Hence, the role of geospatial technology for flood extent mapping and monitoring is important for mitigating the impact of floods in Bihar. This technology is capable to monitor the flood extent and to provide real-time information of flood impact. Because, it allows for the collection of data on rainfall, water levels,

and soil moisture, which can provide valuable information to farmers and local authorities to prepare for floods and make decisions about when to plant, harvest, and evacuate the population in the event of a flood. Synthetic aperture radar (SAR) data has an upper edge than multi-spectral optical datasets for flood mapping and monitoring due to its all-weather sensing capability (in cloud conditions).

The European Space Agency (ESA) has launched “Sentinel-1A satellite on 3 April 2014 and Sentinel-1B was on 25 April 2016”. Its revisit time is six days at equator with two satellite (Torres *et al.*, 2012). This satellite is openly available to all. Nowadays, the Remotely Sensed Earth Observation (EO) datasets are being commonly used for disaster management purposes (Schumann *et al.*, 2018), which is freely available for researchers. However, due to unavailability of resources to download, store and process satellite data is a huge task for the users.

Hence to resolve these complications, Google launched the most advanced cloud-based geo-computing platform “GEE”. This platform enables to processing of huge satellite datasets without requiring local storage (Ghosh *et al.*, 2022). In the present study, Sentinel-2 MSI and PLANET-NICFI (PlanetScope) datasets have been utilized for LU/LC mapping and Sentinel-1 for flood extent mapping and monitoring.

The ability of SAR sensors to detect extent of flooding depends on different scattering mechanisms. “For the identification of inundated pixels, several SAR-based flood identification methods being used scattering mechanism by applying backscatter thresholds to satellite imagery” (Chini *et al.*, 2017). Usually, the change detection technique is being utilized for the identification of flooded pixels using SAR datasets.

Likewise, various methods and indices exist for the extraction of waterbodies using optical and microwave satellite datasets. The NDWI was proposed by McFeeters in 1996 and was shown to be a robust index for detecting waterbodies, especially in areas with high vegetation cover. The MNDWI was proposed by Xu in 2006 as a modification of NDWI, specifically designed for urban areas. It is known to be effective

for extracting waterbodies in urban areas with high reflectance from built-up structures. Recently, the AWEI was developed by Feyisa *et al.* in 2014 and is designed to overcome the limitations of NDWI and MNDWI in areas with mixed pixels, such as rivers or wetlands. AWEI is based on a combination of the green, blue, and red bands, as well as the shortwave infrared and thermal infrared bands, which allows for more accurate detection of waterbodies. These indices have been widely used and validated for detecting waterbodies and assessing the impact of floods.

Apart from this, thresholding techniques are also being used for identification of flood extent using JavaScript code. Therefore, it is always necessary to have accurate data of flooded areas to make an accurate assessment of the damage to make a viable decision on prioritizing relief.

Many research has been undertaken over the years to reduce the consequences of frequent floods in Bihar, India, using optical and SAR satellite datasets. But limited studies have been conducted, that focus on the effects of floods on agriculture that utilizing cutting-edge technologies like machine learning algorithms, high-resolution satellite data (such as Sentinel-1&2 and PlanetScope) and cloud computing platforms like GEE.

The present study aims to address the existing research gap by incorporating these advanced approaches, which offer faster flood extent demarcation and map generation capabilities. Additionally, this study leverages the combined power (data fusion) of Sentinel-2 and PlanetScope satellite datasets for precise land use classification using machine learning approaches.

4.2 Materials and Methods

In this study, the major emphases on the states Bihar which is severely affected during the monsoon season. Therefore, various datasets were required to perform the analysis such as Sentinel-1 SAR dataset for flood extent mapping (Table 3.1). Wherein, the GEE cloud computing platform has been used to process the Sentinel-1 dataset, which enables access to Sentinel-1 data with the SNAP software tool package. This

SNAP package is helpful to process SAR satellite data such as Noise remove, radiometric, orbital and terrain corrections by SRTM datasets and converted backscatter intensity to decibels (dB) according to (i)

$$\text{Eq. } \sigma^{\circ} = 10 * \log_{10} \sigma^{\circ} \quad (i)$$

In the work, all the accessible Sentinel-1 SAR datasets have been utilized for mapping and monitoring purposes of flood.

In the present work, bands: Blue, Green, Red, and NIR utilised of NICFI (PlanetScope) and Sentinel-2 satellite datasets for Land Use mapping. The least cloud-covered (< 10%) imageries of March 2020 and March 2021 have been used with the help of available GEE Tools like “CLOUDY_PIXEL_PERCENTAGE”. Moreover, the QA band of Sentinel-2A/B has been utilized to eliminate the cloud cover (Singha *et al.*, 2020; Kumar *et al.*, 2022).

4.2.1 Methodology

4.2.1.1 Harmonization of PlanetScope and Sentinel-2A/B MSI

Bilinear interpolation is a widely used method for resampling images (Xia *et al.*, 2023; Otsu, 1979). In this study, bilinear interpolation method has been applied in the GEE platform to adjust the resolution of Sentinel-2 datasets and match it to the 4.77-meter resolution of PlanetScope (Script-I). We utilised band: Blue, Green, Red, and NIR of PlanetScope and Sentinel-2A/B MSI for land use mapping. The land use map has been prepared using harmonized sentinel-2 MSI and PlanetScope datasets for extracting flooded agriculture land of 2020 and 2021.

In the present work, Sentinel-1 SAR data has been utilised for the identification flood extent. Wherein, Otsu automatic thresholding technique has been utilised for identification of inundated pixels (Kordelas *et al.*, 2018; Moharrami *et al.*, 2021). An automated thresholding approaches can detects threshold value to distinguish the water pixels from other pixels without having any training sample datasets. Presently, the threshold value is equal to or less than -3 dB used as flood. The pre-flood layer of water

bodies was deducted from the obtained flood extent to get the final result. Figure 4.1 shows the flowchart of the method that was used in this study.

Script-I

```

Resample the Sentinel-2 10m bands to PlanetScope 4.77m using bilinear interpolation method

var resampled
=image.select(bands).resample('bilinear').reproject(proj4.77m);
    
```

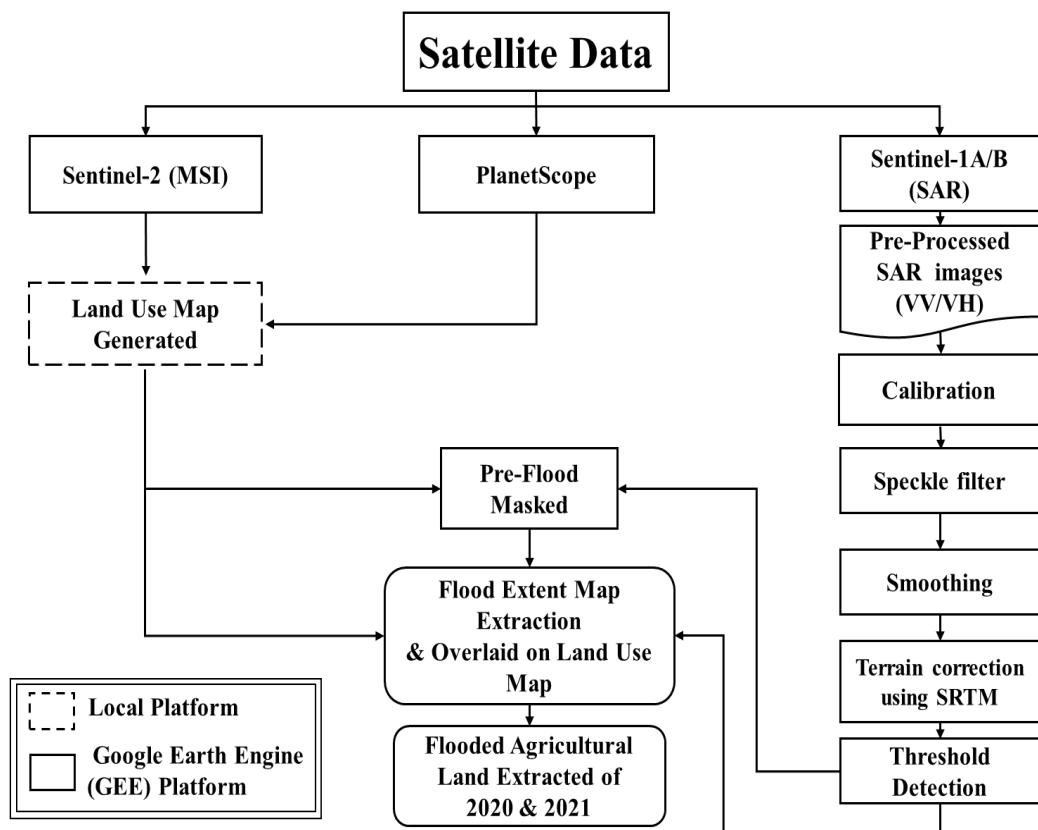


Figure 4.1: Methodology for flood mapping

4.2.1.2 Google Earth Engine (GEE)

Synthetic Aperture Radar (SAR) data processing is a complex task that requires separate computational systems and storage space along with specialized software.

SAR data is acquired through the use of radar instruments that transmit microwave signals to Earth surface and then record reflected signals. This data can be used to generate images of the Earth surface, but the processing of SAR data is complicated by its large size, complexity, and the need for specialized software and hardware.

However, the GEE platform provides a unique solution to these challenges by providing a CCP that allows users; “to process different types of datasets through improved algorithms and JavaScript codes without requiring local storage or downloading raw imagery” (Gorelick *et al.*, 2017).

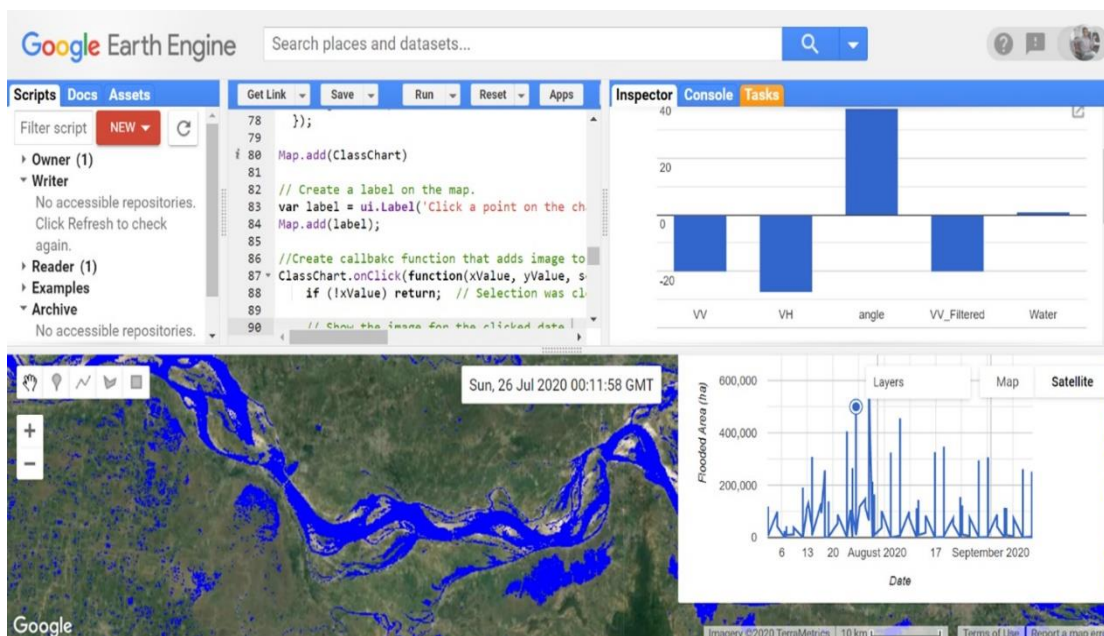


Figure 4.2: The Interface of GEE’s Cloud Computing Platform

The GEE platform provides access to a vast array of satellite imagery, including SAR data, which can be processed in real-time or near-real-time.

The GEE cloud computing platform have used for the entire analysis work and a web-based IDE code was developed (<https://code.earthengine.google.com>) for estimation of flood extents and impacted agricultural lands (Fig. 4.2).

4.2.1.3 Random Forest Classifier

Random Forest Classifier (RFC) is a machine learning algorithm that is commonly used for image classification tasks, including land use map preparation using remote sensing data (Fig. 4.3). In this context, Sentinel-2 and PlanetScope data are often used as inputs to the algorithm. RFC works by creating multiple decision trees, each of which predicts the land use class of a pixel based on its spectral properties. The algorithm then combines the results of these decision trees to produce a final classification map. This classifier is commonly used due to its capability to handle complex spectral interactions and nonlinear relationships between the spectral bands and land use classes, making it a powerful tool for image classification.

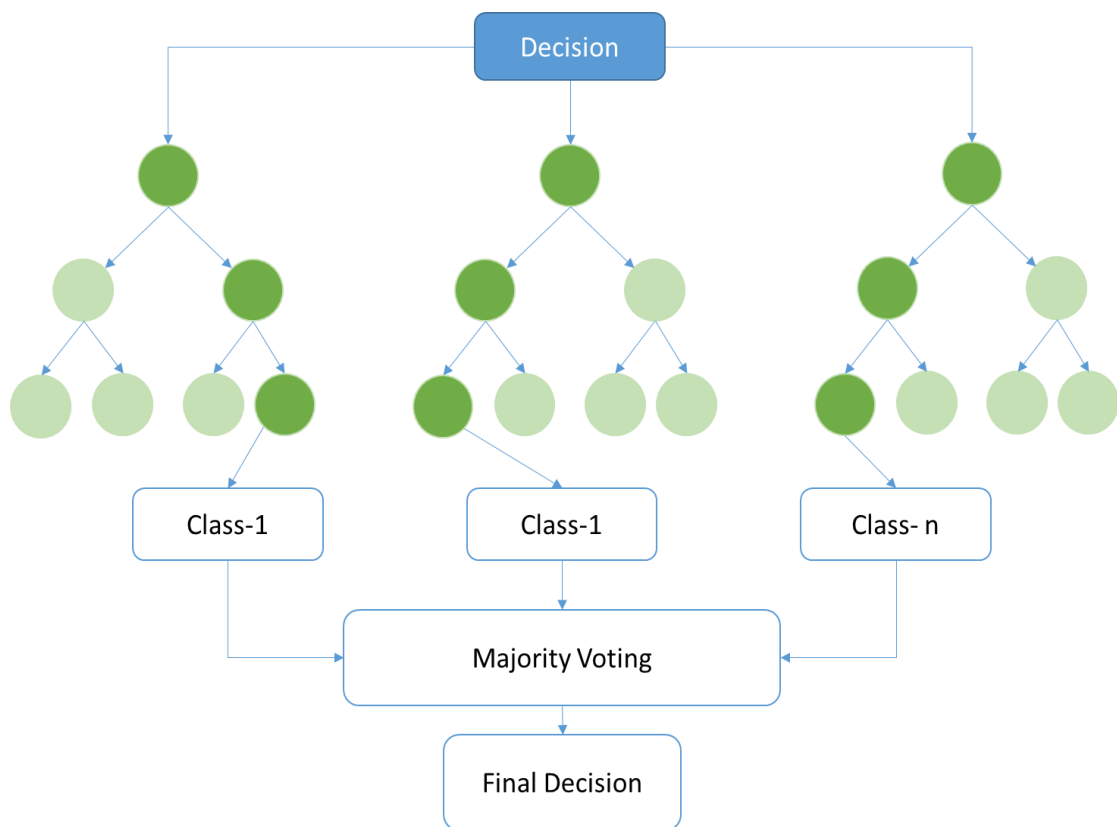


Figure 4.3: Random forest classifier flowchart for land use map preparation

In the present study, Sentinel-2 and PlanetScope data are used together for land use mapping that allows a more comprehensive view of the Earth's surface. Sentinel-2 data provides broad spectral coverage, allowing for identification of a wide range of

land cover types. While, PlanetScope data provides high-resolution imagery, which allows for the identification of smaller features on the Earth's surface. The land use map generated by RFC was utilized in present study for flood impact assessment.

4.3 Results and Discussion

4.3.1 Backscatter intensity of Flooded Pixels

The intensity of backscatter pattern of the processed microwave (SAR) sentinel-1 dataset is influenced by several factors, such as the roughness, slope, moisture content, and vegetation cover of the surface. For instance, a smooth surface like water reflects less radar energy and produces a dark backscatter pattern, while a rough surface like mountainous terrain reflects more radar energy and produces a bright backscatter pattern. Thus, the intensity of the backscatter pattern can give an inference about the types of features present on the Earth's surface, such as water bodies, forests, urban areas, and agricultural land.

In the present study, “the pixels having low grayscale intensity represents waterbodies features while high-intensity pixels resemble man-made or non-water features” (Ghosh *et al.*, 2022). VV polarization has used in this work (Fig. 4.4). The backscatter intensity has been measured in the dB (decibel) scales as shown in figure 5. This information can be used to monitor changes in the land surface to plan and implement appropriate management and conservation strategies.

4.3.2 Flood Impact Analysis

An analysis was conducted to assess the impact of floods in North Bihar. For this analysis, satellite images from March 2020 were used to identify pre-flood conditions, while images from June to October in both 2020 and 2021 were used to identify peak flood events. By subtracting the pre-flood layer from the peak-flood layer, the actual flood extent was derived. To calculate the affected area and population in the region (Table 4.1), the boundaries of villages, blocks, and districts provided by the Survey of India (having population data in the attribute table) were overlaid on the actual flood extent map.

Wherein, it is found that about 22 districts of Bihar were on high alert in monsoon season every year (CWC, 2020; CWC, 2021). Because, these districts are situated in the basins of Kosi, Ganga, Gandak and Mahanadi, which overflow every year due to heavy rainfall. The analysis indicates that almost 701942 hectares (about 21 districts) area in 2020 and 955897.00 hectares (About 22 districts) in 2021 of North Bihar was flooded (Table 4.1). The worst affected districts in 2020 are Darbhanga, Muzaffarpur, West Champaran, Saran and Siwan. In 2021, Bhagalpur, Darbhanga, Muzaffarpur, West Champaran, Katihar, Vaishali, and Khagaria are severely affected by flood (Fig. 4.5).

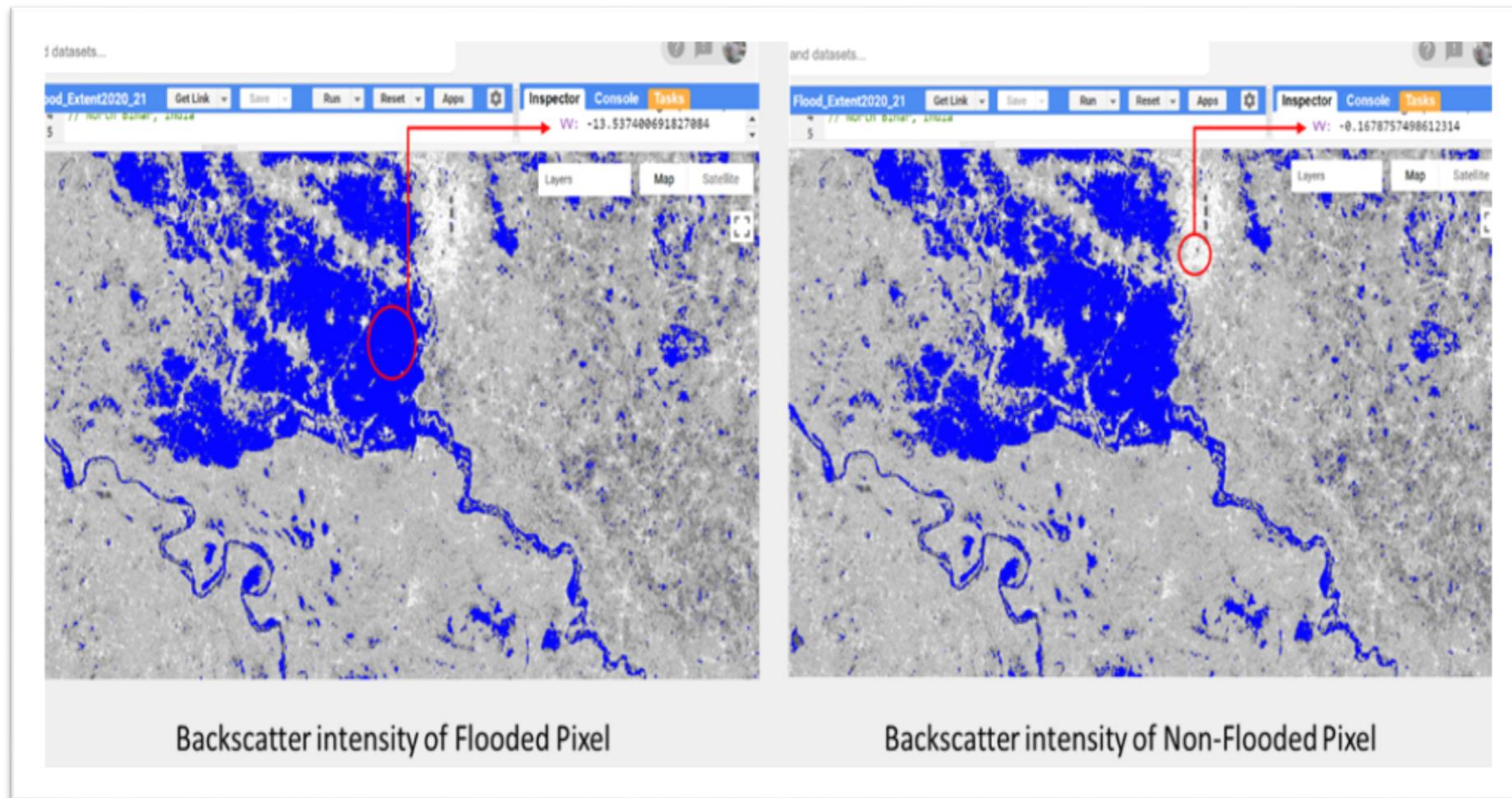


Figure 4.4: Backscatter intensity of pixel

Table 4.1: District-wise flood affected area was extracted using Sentinel-2 and PlanetScope data

Sl.No	Flooded districts	Geographic area in ha.	Total Flood affected area (ha-1)		Flood affected area in %		Change %	Total Flood affected Agricultural land (ha-1)		Total Population	Flood affected Population	
			2020	2021	2020	2021		2020	2021		2020	2021
1	Arariya	279704.17	3947.74	44483.00	1.41	15.90	14.49	3441.74	43340.00	2806200.00	39607	446287
2	Begusarai	193309.07	25.00	40811.00	1.29	21.11	19.82	0.00	37915.00	2954367.00	38111	623720
3	Bhagalpur	255331.10	42244.77	86790.00	16.55	33.99	17.45	34112.77	76829.00	3032226.00	501685	1030689
4	Darbhanga	250777.93	88383.58	78003.00	35.24	31.10	-4.14	84302.58	73628.00	3921971.00	1382250	1219906
5	Gopalganj	204110.19	26793.86	11290.00	13.13	5.53	-7.60	22462.86	8867.00	2558037.00	335798	141493
6	Katihar	303561.68	48795.21	81420.00	16.07	26.82	10.75	40255.21	68816.00	3068149.00	493181	822926
7	Khagaria	149187.39	37009.35	51667.00	24.81	34.63	9.82	35361.35	49774.00	1657599.00	411205	574064
8	Kishanganj	198830.24	3597.59	21873.00	1.81	11.00	9.19	2903.59	20594.00	1690948.00	30596	186019
9	Madhepura	180027.66	27470.43	29620.00	15.26	16.45	1.19	23587.43	26133.00	1994618.00	304359	328175
10	Madubani	350145.84	31718.79	49118.00	9.06	14.03	4.97	30583.79	48013.00	4476044.00	405473	627894
11	Muzaffarpur	317792.18	70620.01	68351.00	22.22	21.51	-0.71	62583.01	60480.00	4778610.00	1061906	1027787
12	Pachim Champan	523841.15	35093.54	28461.00	6.70	5.43	-1.27	26583.54	19256.00	3922780.00	262798	213130
13	Purba Champan	397076.47	76045.16	56278.00	19.15	14.17	-4.98	67483.16	49819.00	5082868.00	973433	720399
14	Purnia	321126.64	24183.87	57396.00	7.53	17.87	10.34	20368.87	51827.00	3273127.00	246497	585017
15	Saharsa	166419.72	38150.54	29208.00	22.92	17.55	-5.37	33951.54	25706.00	1897102.00	434897	332957
16	Samastipur	268545.22	25867.16	61176.00	9.63	22.78	13.15	23895.16	58048.00	4254782.00	409835	969262
17	Saran	267895.35	32303.37	40353.00	12.06	15.06	3.00	27047.37	34639.00	3943098.00	475467	593948
18	Sheohar	44185.79	4845.21	2970.00	10.97	6.72	-4.24	4770.21	2903.00	656916.00	72034	44155
19	Sitamarhi	218858.48	32173.55	23434.00	14.70	10.71	-3.99	30813.55	22306.00	3419622.00	502705	366152
20	Siwan	221989.66	31425.75	8163.00	14.16	3.68	-10.48	26767.75	5312.00	3318176.00	469734	122016
21	Supaul	241658.20	8536.27	35178.00	3.53	14.56	11.02	5149.27	28788.00	2228397.00	78715	324386
22	Vaishali	202102.12	12736.51	49854.00	6.30	24.67	18.37	8281.51	38670.00	3495249.00	220271	862198
		5556476.26	701967.26	955897.00	12.63	17.20	4.57	614706.26	851663.00	68430886.00	9150559	11772367

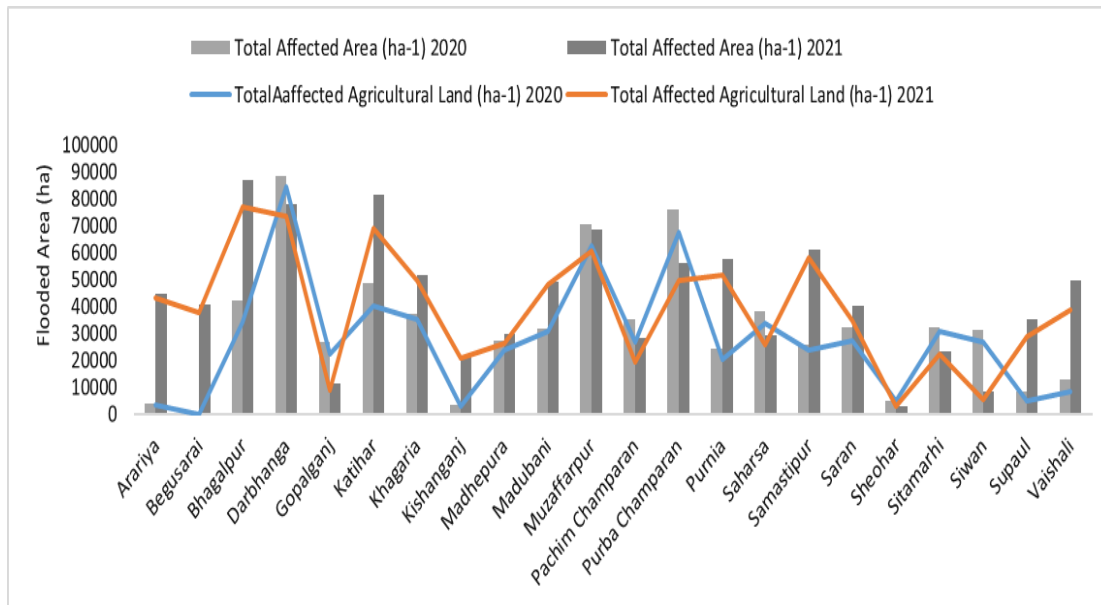


Figure 4.5: District wise flood statistics

In 2020, 614706.26 hectares and in 2021 851663.00 hectares of agricultural lands were submerged due to floods. In respect to submerged area; The floods of 2020 have submerged more than 70,000 hectares in 3 districts, around 40,000 to 50,000 hectares in 2 districts, 30,000 to 40,000 hectares in 7 districts, 20,000 - 30,000 hectares in 4 districts and less than 13,000 hectares in 5 districts (Fig. 4.5).

The floods of 2021; More than 70,000 hectares in three districts, 50000 to 60000 hectares in five districts, 30000 to 40000 hectares in six districts, 20000 to 30000 hectares in five districts and about 3000 to 11000 hectares in 3 districts have been submerged. The majority of Bihar's population relies on agriculture for their livelihood, which is the sector that suffers the most damage from the recurrent floods. In detailed affected areas and affected population are shown in Table 4.2.

4.3.2.1 Flood Progression evaluation using Sentinel-1 SAR Data

Presently, VV polarization has been utilized for flood extent mapping. Wherein, backscatter response of VV ranges between -8 and -16 dB (Fig. 4.6). In this study, we found that high values of VV represent non-water features and water bodies shown by a low value of backscatter response.

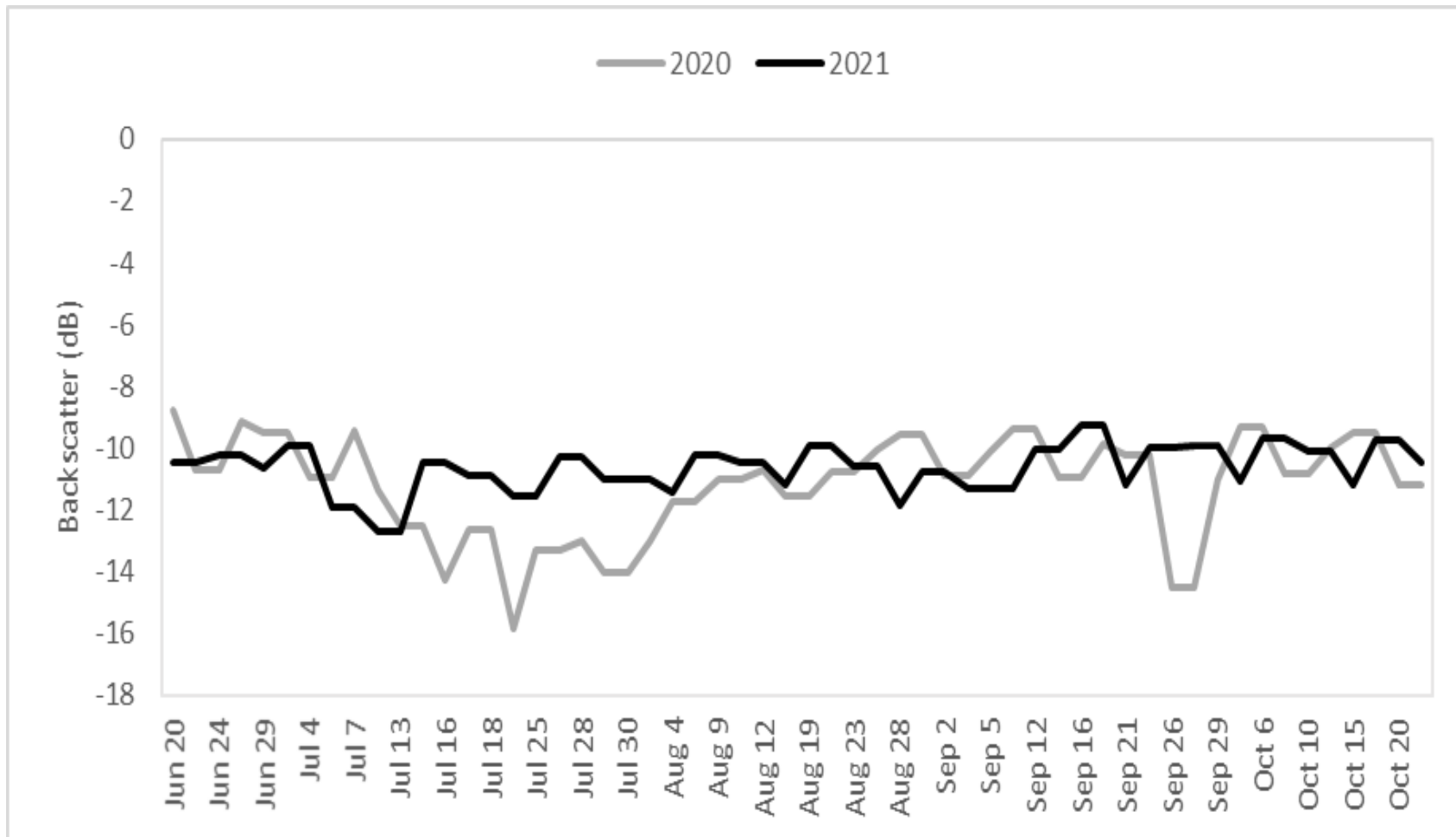


Figure 4.6: Backscatter response of VV polarization in 2020 and 2021

4.3.2.2 Pre Flood Land Use Map

Land use mapping involves using remote sensing data to identify and classify the different types of land cover such as forests, croplands, urban areas, and water bodies. This information is often used by environmental managers, urban planners, and policymakers to make quick decisions about how to use and manage the land. Remote sensing is very important tool for land use mapping, as it provides a comprehensive and accurate view of the Earth surface from distance.

In this study, Sentinel-2A/B (< 10 % clouded data) and PlanetScope images were used to extract landuse maps using Random forest methods on GEE cloud computing platform. A pre-flood land use map (Fig. 4.7) has been used and the flood extent has been overlaid on it to find the land use impacted by floods 2020 and 2021.

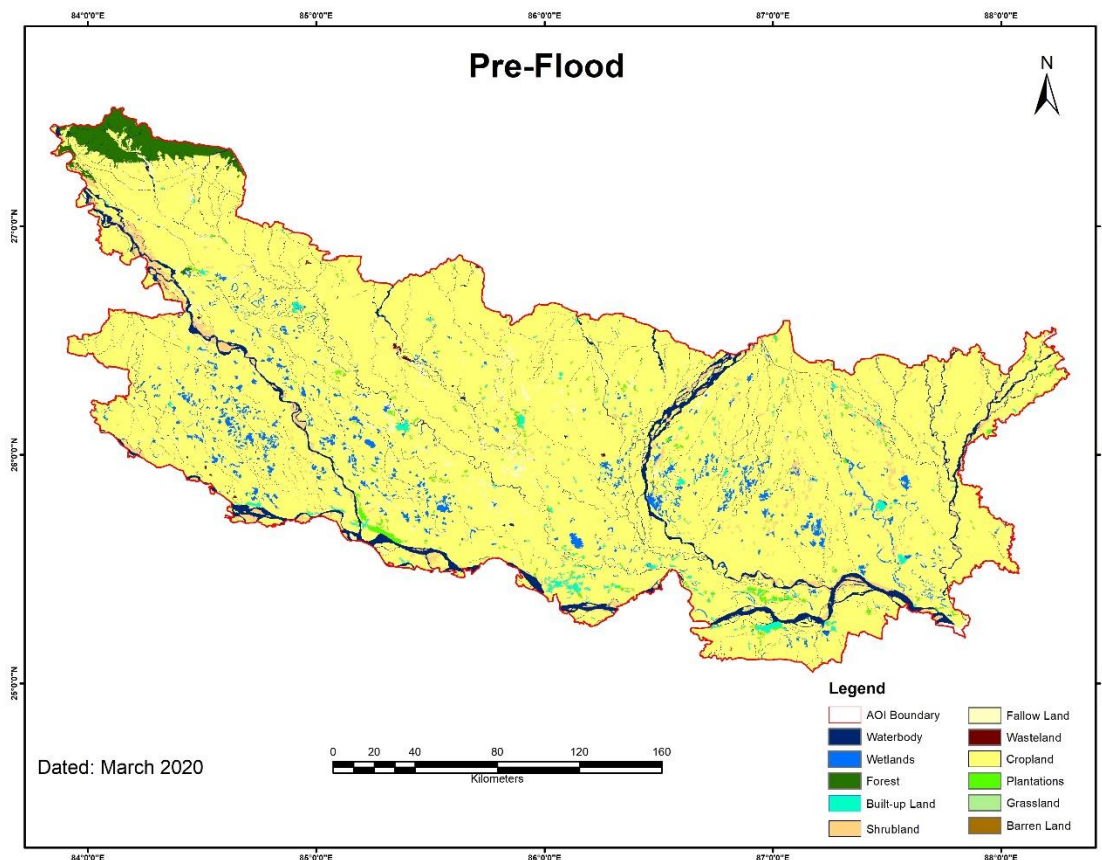


Figure 4.7: Land use map of pre-flood area

4.3.2.3 Flood Progression Assessment (2020)

Flood progression in North Bihar based on Sentinel-1 SAR data from June to October 2020 were shown in Figure 10. We used VH and VV polarization for flood delineation. The backscatter response of VV polarization falls within the range of -8 to -16 dB, whereas for HV polarization, it ranges from -16 to -23 dB. This variation in backscatter response between VV and HV polarizations indicates the different scattering properties of the target features being observed by the radar system (Fig. 4.8).

Severe rainfall was started from June to till September and developed flood-like conditions for North Bihar in 2020. The extent of the flood shows the effects of rain-induced flooding in Fig. 4.9. Major areas were submerged from 07th to 29th July 2020. And again some parts of North Bihar were inundated during 21st to 30th September due to heavy rain in the river basin of North Bihar (Fig. 4.9).

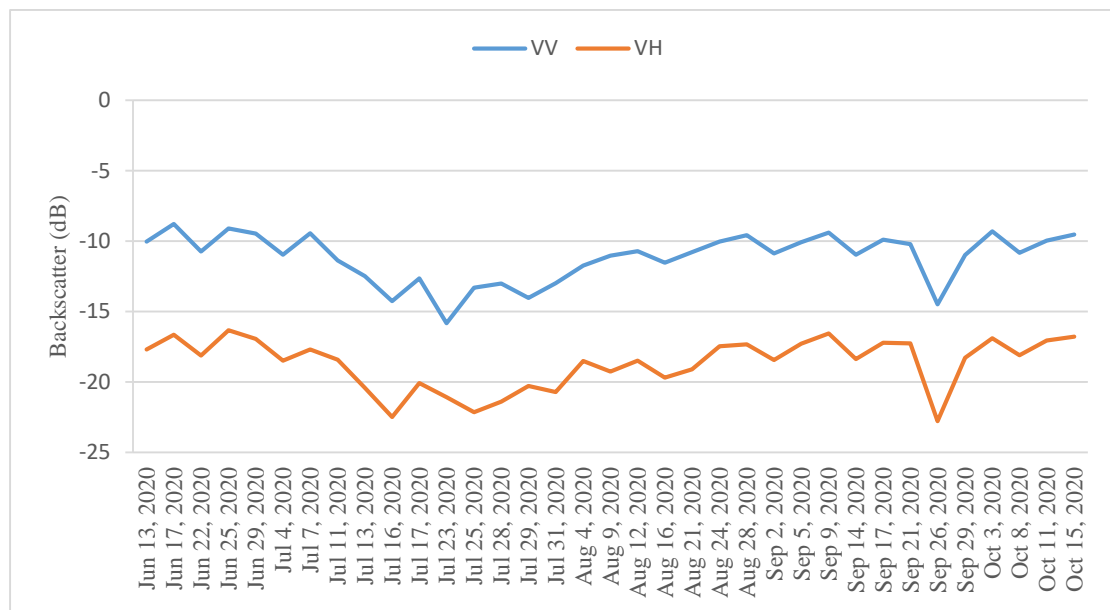
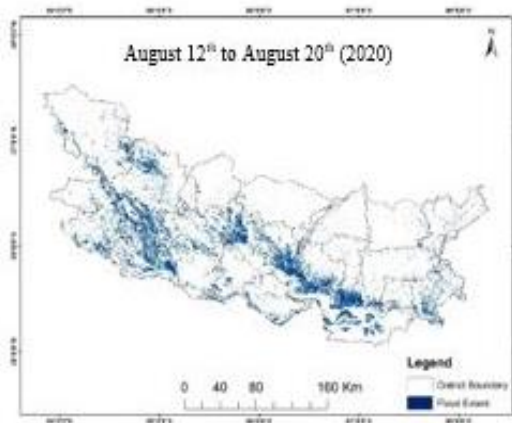
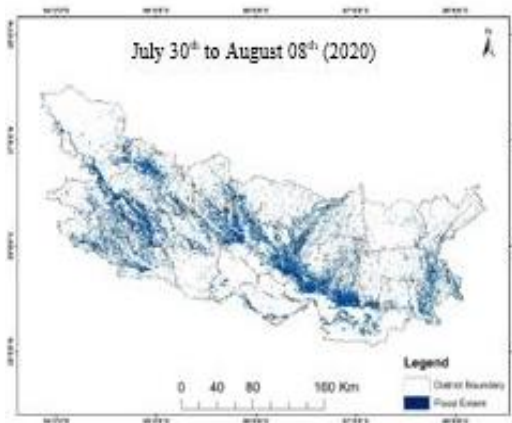
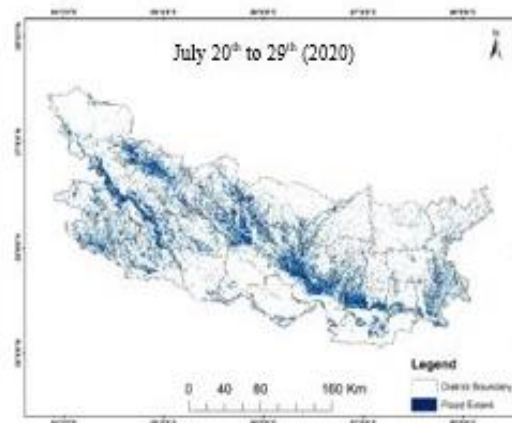
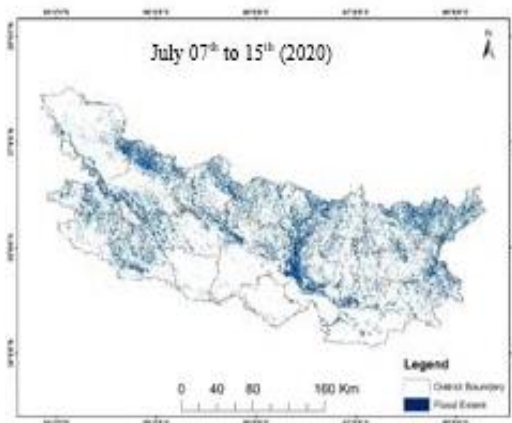
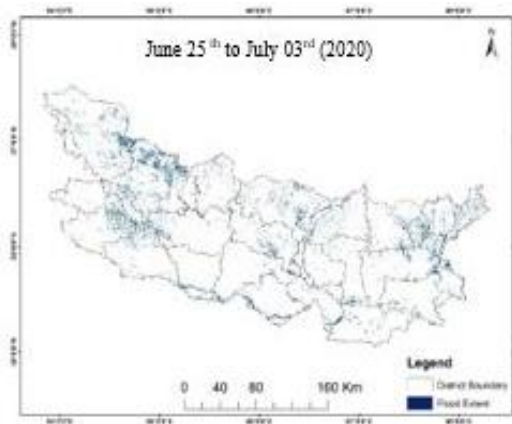


Figure 4.8: Backscatter response of different polarization



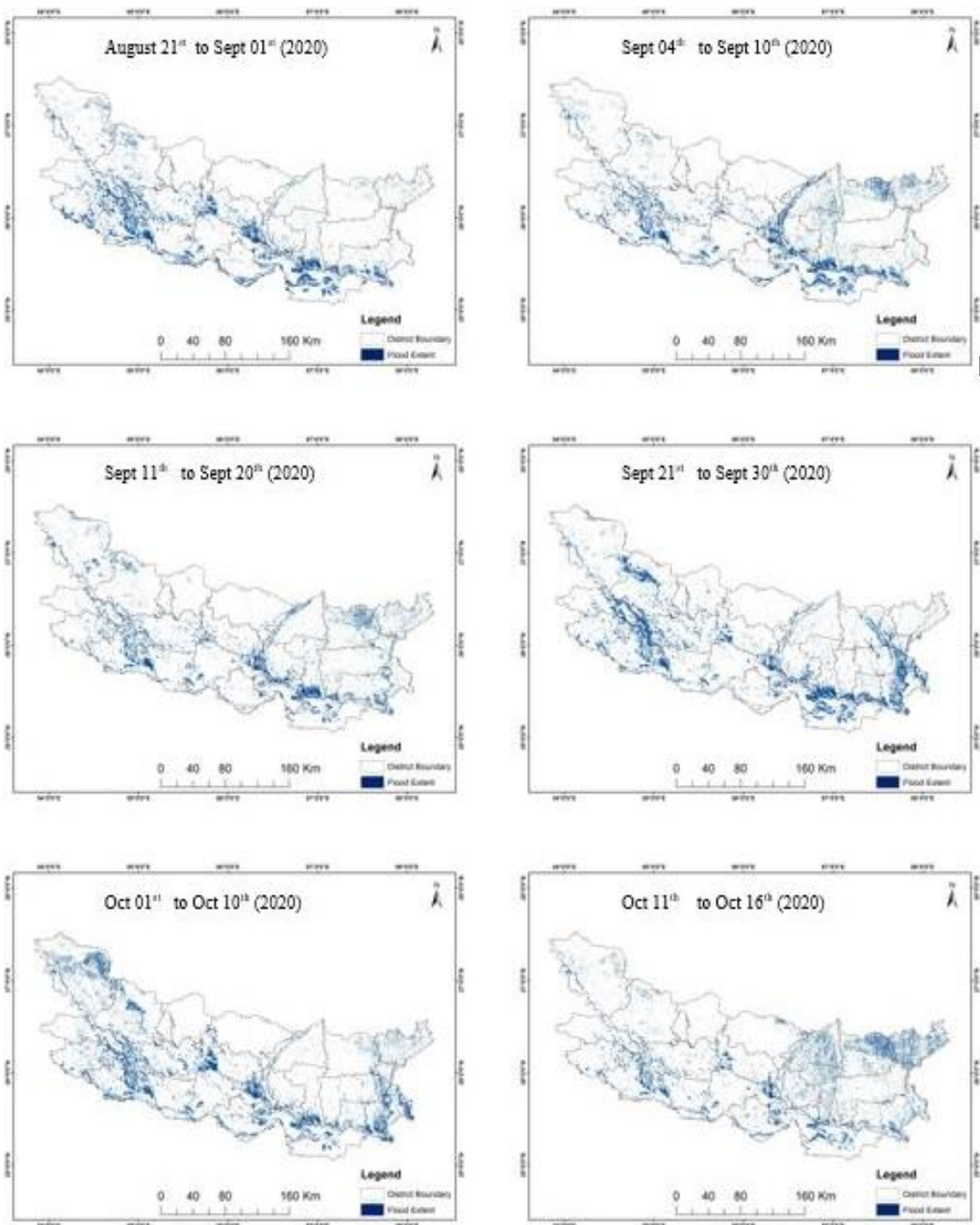


Figure 4.9: Flood Progression of 2020

- **Flood 2020**

More than 8000 hectares of croplands have been affected by the 2020 floods in Darbhanga, Purbi Champaran and Muzaffarpur districts of Bihar and between 30,000 to 40,000 hectares of cropland in Katihar, Khagaria, Bhagalpur, Saharsa, Sitamarhi and Madubani districts. Around 20,000 to 30,000 hectares of crop land has been submerged in the districts of Saran, Siwan, Pachim Champaran, Samastipur, Madhepura, Gopalganj and Purnia due to the 2020 floods. In Bihar's Saran, Siwan, Paschim Champaran, Samastipur, Madhepura, Gopalganj and Purnia districts, 20,000 to 30,000 hectares of crop land has been submerged. About 8000 hectares of crops have been affected in Vaishali district and 2000 to 5000 hectares in Supaul, Sheohar, Arariya, Kishanganj and Begusarai districts (Fig. 4.10).

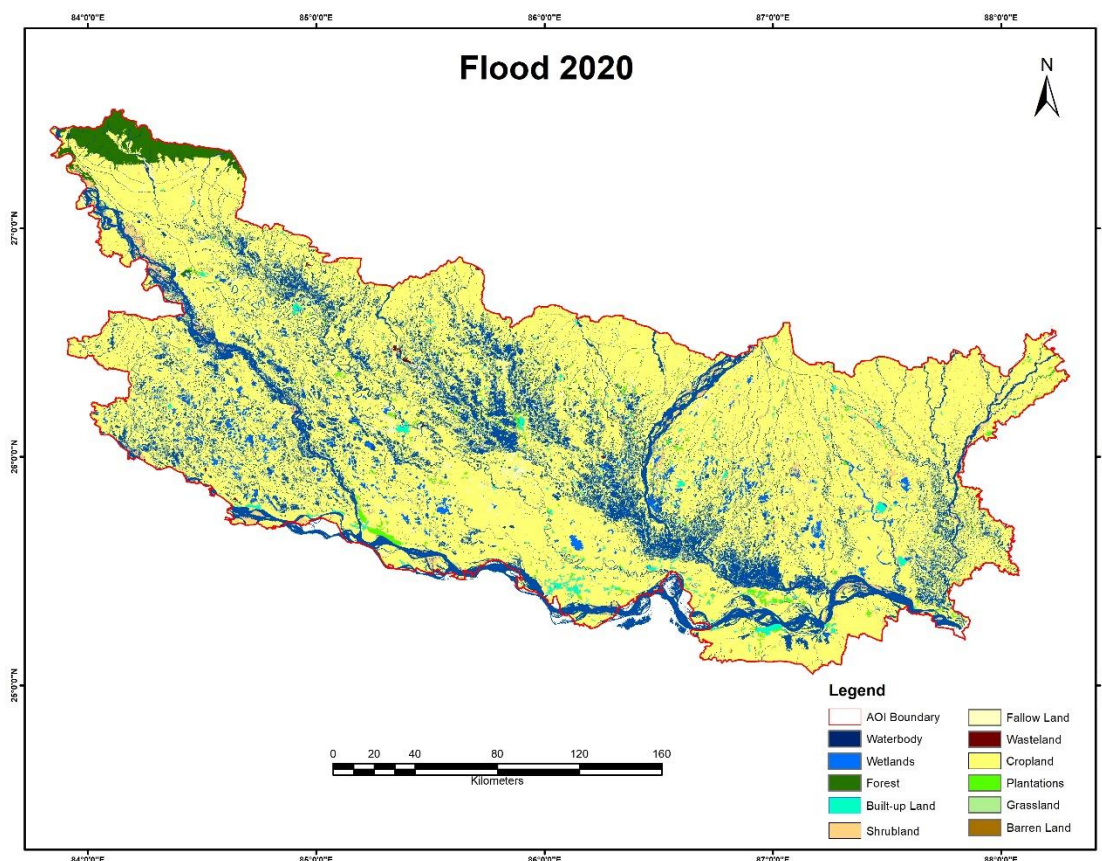


Figure 4.10: Flood affected Land Use Map of 2020

4.3.2.4 Flood Progression Assessment (2021)

In 2021, Sentinel-1 SAR datasets and VH and VV polarization have been used for Flood progression assessment in North Bihar, which is prolonged from June to October 2021 as representation in Figure 4.11. Flood 2021 was analysed to compare the severity between floods 2020 and 2021. The comparison of the floods of 2021 to the floods of 2020 can help to identify factors that contributed to severity of floods. This information can be used to develop strategies to mitigate the effects of future floods.

- **Flood 2021**

The Analysis shows that flood 2021 events arisen in four phases. The first phase comes in the last week of June 2021, the second phase in the first week of July 2021, the third phase in the first week of August 2021 and the fourth phase in the last week of August 2021. The second phase of flood has had more impact on the region.

Wherein, more than 60,000 hectares cropland in four districts (Bhagalpur, Darbhanga, Katihar, Muzaffarpur), 40000 to 6000 hectares cropland in six districts (Samastipur, Purnia, Purba Champaran, Khagaria, Madubani, and Arariya), 30,000 to 40,000 hectares in four districts (Vaishali, Begusarai, Saran and Supaul), 20,000 to 25,000 hectares in five districts (Madhepura, Saharsa, Sitamarhi, Kishanganj and Pachim Champaran) and about 3000 to 9000 hectares in 3 districts (Gopalganj, Siwan and Sheohar) have been submerged (Fig. 4.12).

Agricultural lands have been most affected by floods in Bhagalpur, Darbhanga, Katihar and Muzaffarpur districts of Bihar. These districts are vulnerable to floods due to their geographical location and proximity to major rivers like Ganga, Gandak and Kosi.

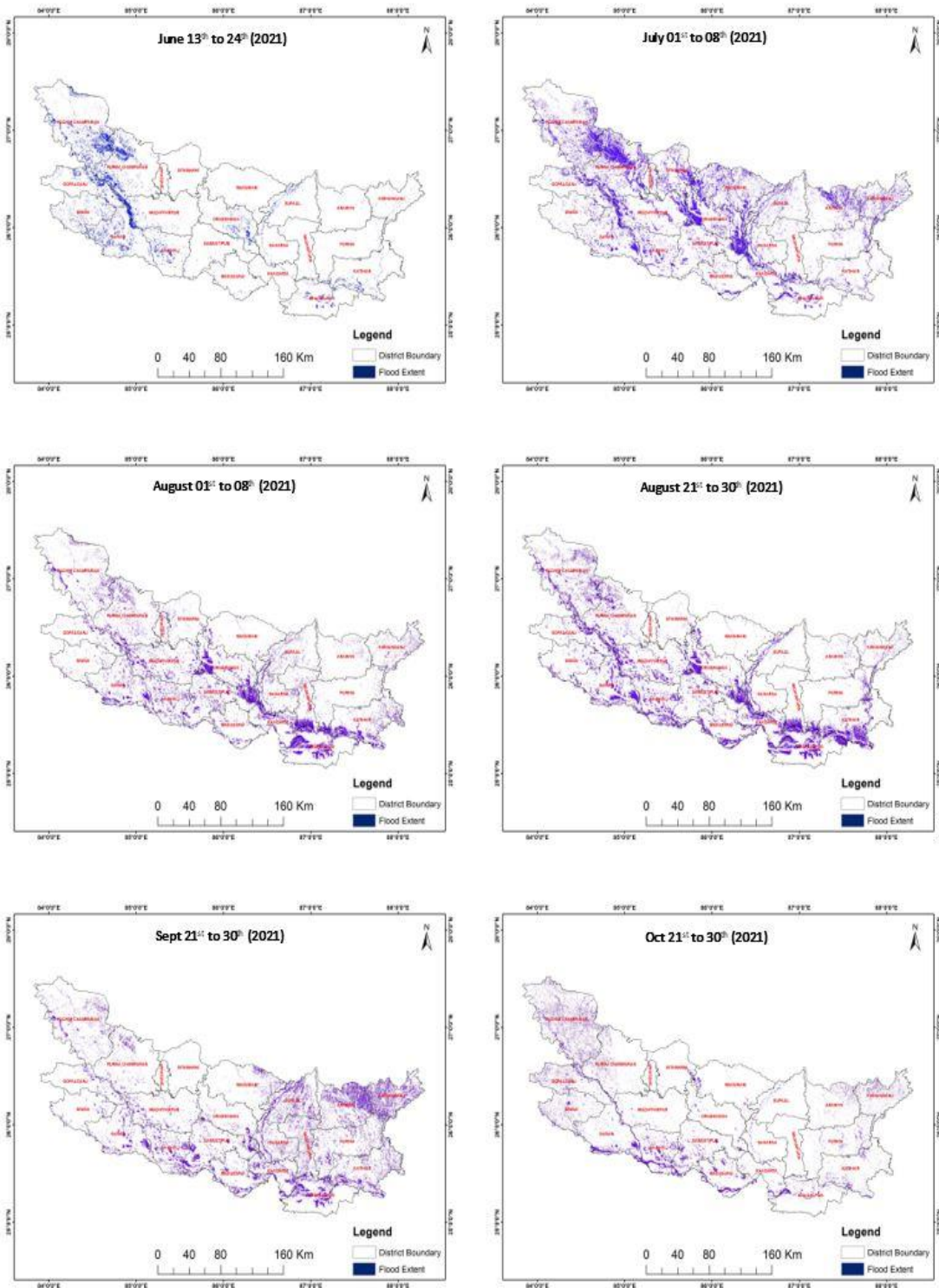


Figure 4.11: Flood Progression of 2021

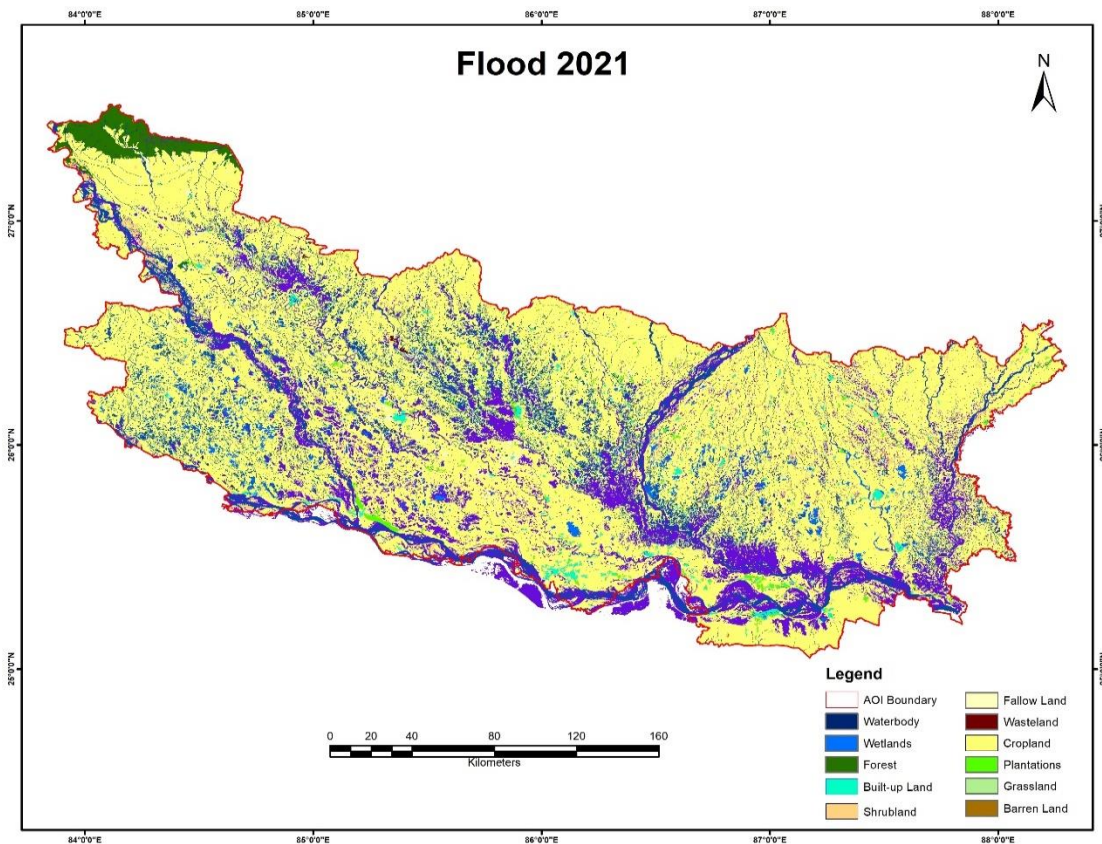


Figure 4.12: Flood affected Land Use Map of 2021

4.4 Accuracy Assessment and Result Validation

The accuracy of the extracted flood extent areas was validated independently using datasets that included high-resolution satellite images from Google Earth and the Flood extent layer of NRSC (Report NRSC, 2019). Validation points were carefully selected to ensure representation across the entire images. Approximately 260 validation points were chosen within the study area. Subsequently, confusion matrices were generated to assess the classification performance, commonly employed method for multi-class classification evaluations (Olofsson *et al.*, 2014). The overall accuracies obtained exceeded 89%, thereby confirming the suitability of the Otsu automatic thresholding method for rapidly and efficient flood mapping (Kordelas *et al.*, 2018). In addition to the aforementioned validation datasets, advisory from State Disaster Management Authority (SDMA), the Flood Management Information System (FMIS) Bihar and Crop area affected report of NRSC (Report NRSC, 2020) were employed to further validate the study. This comprehensive validation approach, which involved multiple independent sources, strengthens the credibility of the study's results.

4.5 Impact of Flood on Maize

Flooding has a significant impact on the agricultural sector of Bihar as well as many other flood-prone regions around the world, especially on cropping patterns, productivity and production. Here are some findings on how floods are affecting crop patterns, productivity and production in Bihar. Firstly, the standing crops get submerged due to flood water, leading to their damage or complete loss. Paddy is the main crop of rainy season and is cultivated in almost all the districts of Bihar, which is usually destroyed due to severe floods or long-time submergence (Anonymous, 2020a, b, c). Secondly, fields may remain waterlogged for several months or even years. This renders them uncultivable and unsuitable for growing crops. Farmers cannot plant their usual crops on such land until it has been properly drained, dried, and rehabilitated. Which lead to shift in cropping patterns in some places, with farmers choosing to grow less flood-prone crops, such as maize and pulses, instead of paddy.

However, floods have many positive impacts on agricultural land such as supply of essential nutrients that are necessary for plant growth. When floods occur, they carry with them a large amount of sediment that is rich in organic matter and nutrients such as nitrogen, phosphorus, and potassium, which are necessary for plant growth. These nutrients help to improve soil fertility and increase crop yield (specially for *rabi* crops *viz*; maize) in North Bihar, which is beneficial for farmers who depend on agriculture for their livelihood.

In the present study, it is observed that area under maize is approximately 2.5% higher in 2022-23 (Non-flooded years) as compared to 2021-22 (flooded years), but there is approximately 10.50% lower yield and 8.97% lower production in 2022-23 as compared to 2021-22 (Table 4.2). Hence, the study shows that when there is flood, the yield of maize crop is higher than in non-flood years (Fig. 4.13). Due to unavailability of current data, estimated acreage, yield and production have been compared with already available government-reported data (Table 4.3), where a positive correlation has been found.

Table 4.2: Districts wise maize crop acreage, yield and production estimation during 2021-22 and 2022-23

Sl. No.	District	2021-22					2022-23				
		Flooded Year					Non-Flooded Year				
		Acreage		Yield	Production		Acreage		Yield	Production	
		(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)	(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)
1	Arariya	30784	30.78	9987.74	307458916	307.46	31205	31.21	8839.15	275827266	275.83
2	Begusarai	40052	40.05	7463.73	298938807	298.94	41613	41.61	6881.56	286362442	286.36
3	Bhagalpur	17050	17.05	6508.76	110973577	110.97	17373	17.37	5740.73	99733124	99.73
4	Darbhangha	5425	5.42	4407.44	23908731	23.91	5610	5.61	4019.59	22549437	22.55
5	Gopalganj	3907	3.91	4332.24	16925239	16.93	3974	3.97	3691.07	14670126	14.67
6	Katihar	38141	38.14	9949.47	379480991	379.48	38335	38.33	8805.28	337546300	337.55
7	Khagaria	33754	33.75	7171.20	242055466	242.06	35601	35.60	6540.13	232835379	232.84
8	Kishanganj	1118	1.12	9133.85	10215566	10.22	1180	1.18	7809.44	9211613	9.21
9	Madhepura	17734	17.73	8846.16	156879040	156.88	17868	17.87	8005.77	143048369	143.05
10	Madubani	43	0.04	3968.00	170068	0.17	45	0.05	3579.14	161256	0.16
11	Muzaffarpur	10537	10.54	5383.30	56724801	56.72	10902	10.90	5340.23	58218353	58.22
12	W. Champaran	1811	1.81	5501.88	9966491	9.97	1830	1.83	5402.85	9889631	9.89
13	E. Champaran	4946	4.95	3811.50	18852632	18.85	5139	5.14	3895.35	20016496	20.02

Sl. No.	District	2021-22					2022-23				
		Flooded Year					Non-Flooded Year				
		Acreage		Yield	Production		Acreage		Yield	Production	
		(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)	(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)
14	Purnia	47939	47.94	9597.42	460086686	460.09	49514	49.51	8109.82	401549894	401.55
15	Saharsa	12425	12.43	6904.32	85786659	85.79	12540	12.54	6089.61	76364715	76.36
16	Samastipur	28005	28.01	5740.58	160765804	160.77	28771	28.77	4890.97	140716613	140.72
17	Saran	5644	5.64	5047.92	28490208	28.49	5734	5.73	4674.37	26802406	26.80
18	Sheohar	1502	1.50	5437.40	8169150	8.17	1530	1.53	4600.04	7037718	7.04
19	Sitamarhi	3144	3.14	5829.30	18325221	18.33	3109	3.11	4698.42	14605795	14.61
20	Siwan	5277	5.28	3690.00	19472204	19.47	5388	5.39	3623.58	19525580	19.53
21	Supaul	4711	4.71	9804.45	46191290	46.19	4703	4.70	8363.19	39336174	39.34
22	Vaishali	7303	7.30	4163.39	30404862	30.40	7540	7.54	4088.45	30828940	30.83
Total (AOI)		321252	321.25	7751.68	2490242408	2490.24	329504	329.50	6879.54	2266837629	2266.84

Note: AOI (Area of Interest), ha (Hectare), '000 (Thousand), Kg (Kilometre), Sl (Serial), No (Number)

Table 4.3: Districts wise maize crop acreage, yield and production as per GoB Report

District	GoB report 2020-21 (Rabi Maize)			GoB report 2019-20 (Rabi Maize)			GoB report 2018-19 (Rabi Maize)		
	Area (ha)	Yield (Kg/ha)	Production (tonne)	Area (ha)	Yield (Kg/ha)	Production (tonne)	Area (ha)	Yield (Kg/ha)	Production (tonne)
Arariya	27543.00	9740	268269	28265.00	9242	261225	26499.00	9932	263188
Begusarai	36309.00	7110	258157	29516.00	7928	234003	19881.00	6869	136563
Bhagalpur	17324.00	6124	106092	15686.00	7306	114602	15629.00	8133	127111
Darbhangha	5611.00	3364	18875	5478.00	4438	24311	6175.00	6185	38192
Gopalganj	3689.00	4687	17290	4120.00	4782	19702	3242.00	3831	12420
Katihar	36011.00	9642	347218	30965.00	10243	317174	29745.00	10074	299651
Khagaria	32871.00	7573	248932	32942.00	7021	231286	32965.00	6530	215261
Kishanganj	1124.00	8956	10067	1299.00	10243	13306	1268.00	8744	11087
Madhepura	16307.00	8342	136033	18757.00	8994	168700	18713.00	9258	173245
Madubani	0.00	0	0	0.00	0	0	53.00	5484	291
Muzaffarpur	10201.00	5105	52076	10091.00	5043	50889	9838.00	4566	44920
West Champaran	1707.00	5558	9488	1834.00	5767	10577	1873.00	4596	8608
East Champaran	4484.00	4297	19268	4457.00	4100	18274	4544.00	3830	17404
Purnia	47464.00	9269	439944	45267.00	8834	399889	35786.00	9434	337605
Saharsa	12264.00	6551	80341	11733.00	6117	71771	11471.00	6826	78301
Samastipur	27207.00	5681	154563	27385.00	5984	163872	31389.00	5346	167806
Saran	5472.00	4998	27349	5318.00	4476	23803	4543.00	4502	20453
Sheohar	1459.00	4775	6967	1427.00	5036	7186	1314.00	4837	6356
Sitamarhi	3155.00	6240	19687	3335.00	5583	18619	3464.00	4788	16586
Siwan	4970.00	3761	18692	4970.00	3560	17693	4970.00	3456	17176
Supaul	4604.00	8922	41077	4599.00	11139	51228	4595.00	1009	46364
Vaishali	7030.00	3795	26679	6960.00	4060	28258	7046.00	3695	26035
Total (AOI)	306806.00	7520	2307064	294404.00	7630	2246368	275003.00	7508	2064623

Note: AOI (Area of Interest), GoB (Government of Bihar) ha (Hectare), Kg (Kilometre)

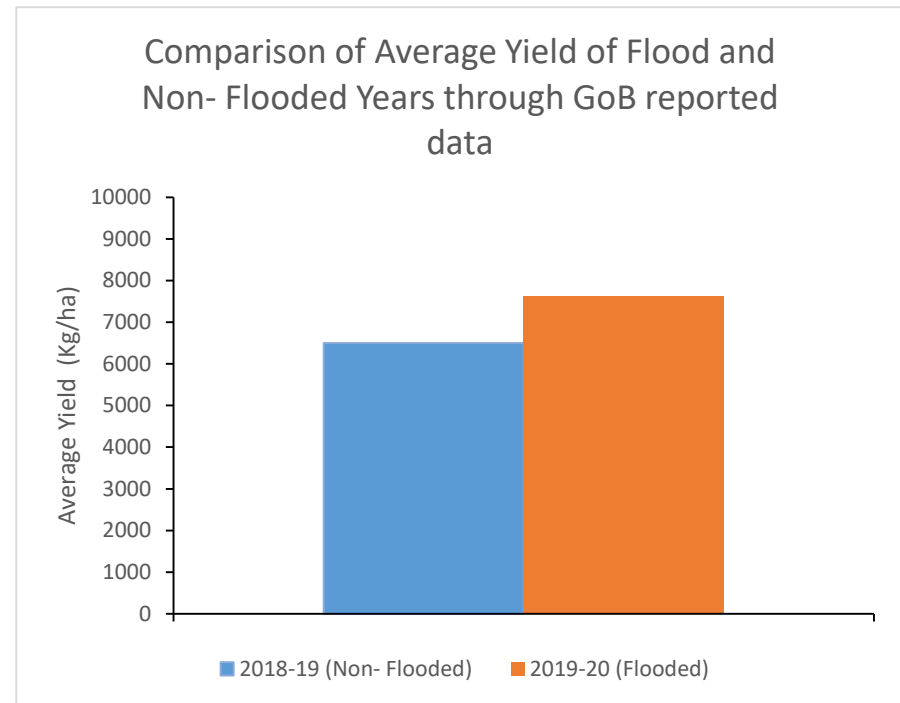
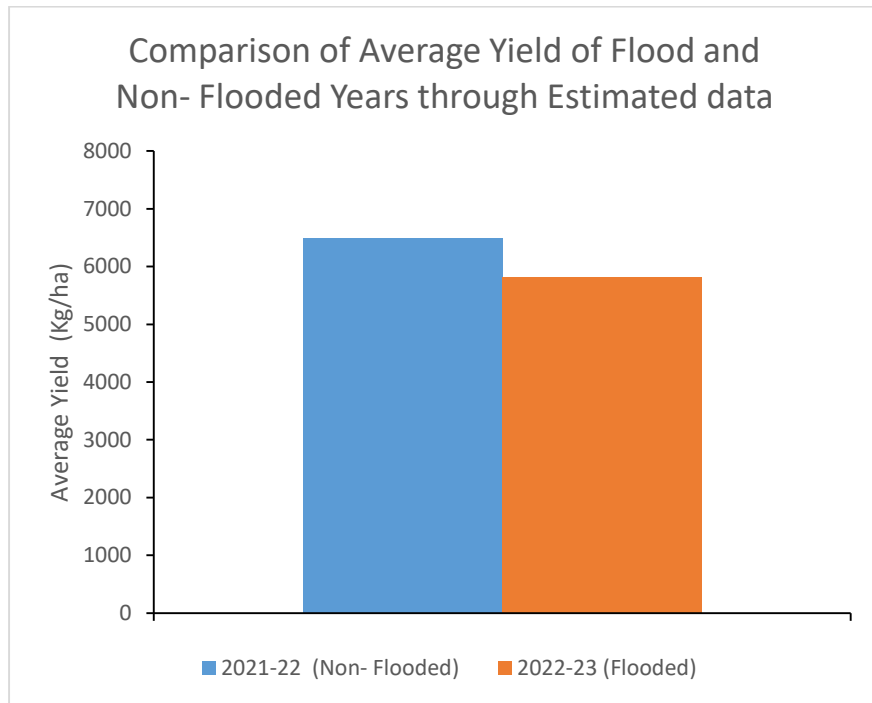


Figure 4.13: Comparison of average yield of flood and non- flooded years

4.6 Conclusion

In this study, the Impact of flood on agriculture for North Bihar has been assessed. Optical remote sensing data; Sentinel-2 and PlanetScope hosted on GEE cloud platform have been used to delineate pre and post-flood land use maps for the year 2020 and 2021. Sentinel-1 SAR data has been used to identify flood pixels using automated thresholding technique on GEE. A JavaScript code has been developed for the processing of enormous datasets hosted on the GEE cloud computer platform within a short time of period. This JavaScript code is capable to robust flood mapping and monitoring using microwave (SAR) satellite datasets at large scale. It has been observed that about ~ 12.63% (701967 ha) of study area were flooded in 2020. For 2021 floods, about ~ 17.20% (955897 ha) of study area were found to be flooded. Districts most affected by floods were Bhagalpur, Darbhanga, Katihar, Muzaffarpur and Gopalganj in 2021. In the floods of 2021, about 4% more area of North Bihar was flooded as compared to the floods of 2020. Results were validated using flood extent layer generated by NRSC and advisory document of State Disaster Management Authority of Bihar. It is expected that the generated maps and its area of statistics of flood extents and flooded agriculture land will be beneficial for policy makers for future planning. Furthermore, the next chapter elaborates on maize crop identification and acreage estimation to assess the impact of flood on the crop.

CHAPTER-5 MAIZE (*ZEA MAYS L.*) CROP IDENTIFICATION/ DISCRIMINATION AND ACREAGE ESTIMATION

5.1 Introduction

The 21st century is facing many challenges like climate change, rapid urbanization, population growth, food shortage, poverty and hunger (Mondal *et al.*, 2014; FAO. 2021). The world population is expected to grow from 7 billion (1.2% growth per year) to around 10 billion by 2050 (UNFPA, 2012, FAO. 2017), which will accelerate the high demand for food consumption such as meat, fruits and vegetables, and increase pressure on natural resources. According to FAO. 2021, “it is projected that between 720 to 811 million people in the world faced hunger in 2020 and around 118 million more people were facing hunger in 2020 as compared to 2019. More than half of the world’s undernourished are found in Asia (418 million) and more than one-third in Africa (282 million)”.

Therefore, innovative agricultural research and better management practices are necessary to increase productivity to reduce food insecurity. In this context, Accurate and timely information on crop type, acreage and yield is essential for estimating crop production to reduce food insecurity, loss assessment for crop insurance schemes and developing adaptation strategies at local, national and global levels (Arumugam *et al.*, 2021a; Kumar *et al.*, 2022). Hence, an innovative agriculture management plan for better agriculture practices is urgently needed for countries like India, where about 58% of the population is engaged in the agriculture sector for their livelihood, and more than 75% of the country's farmers are smallholders (Gulati *et al.*, 2018; Rai *et al.*, 2019; Arumugam *et al.*, 2021).

In this context, Maize (*Zea mays*) crop has the potential to reduce food insecurity, generate better income for the farmers and also qualifies as the potential of doubling the farmer's income (Maize Vision 2022). “Maize is the 2nd most important cereal crop in the world in terms of acreage. Global maize production touched approx. 1040 million MT in 2016-17, wherein, the US has been the leading producer, followed by China, accounting for about 38% and 23% respectively” stated (Singh *et al.*, 2018).

Among the maize growing countries, “India ranks 4th in the area and 7th in production, representing around 4% of the world maize area and 2% of total (Global) production with a quantum of 26 million MT in 2016-17, with about 15 million Indian farmers are engaged in Maize cultivation” (DACNET, 2020; Maize Vision 2022). This crop is the 3rd most important food crop after rice and wheat, grown in the Indian states of Karnataka, Madhya Pradesh, Maharashtra, Rajasthan, Bihar and Uttar Pradesh (https://farmer.gov.in/m_cropstaticsmaize.aspx, accessed dated 21.06.2022). Bihar is a leading producer of Rabi maize in India, contributing about 10% of India’s total production (Maize Vision 2022).

Nowadays, remote sensing satellite and ML approach-based crop mapping is growing and expanding from small to large scale due to the free accessibility of high-resolution satellite dataset, which is a cost-effective solution for farm-level mapping (Chlingaryan *et al.*, 2018; Griffiths *et al.*, 2019; Htitiou *et al.*, 2022; Pande, 2022). In this regard, several studies have been carried out for crop mapping using coarse to fine spatial resolution data based on MODIS, Landsat, Sentinel-1 and Sentinel-2 at regional to global levels (Mondal *et al.*, 2014; Peng *et al.*, 2011). The FASAL program is going on for crop forecasting, acreage estimation and crop yields prediction at the national scale in India. In the FASAL program, various approaches such as weather-yield models, crop simulation models and remote sensing-driven statistical models have been used (Arumugam *et al.*, 2021b; Latwal *et al.*, 2019). But, remote sensing satellite and machine learning approach-based crop studies are very limited (Arumugam *et al.*, 2021a). Therefore, remote sensing and machine learning approach has been employed for the estimation of maize crop in North Bihar in the present study.

These days, combining satellite data from multi-sensor source is a common approach to improve the accuracy of crop mapping (Costa *et al.*, 2021; Mizuochi *et al.*, 2021). Several studies have combined multi-sensors source satellite data such as MODIS with Sentinel-2A/B (Zhu *et al.*, 2018), Sentinel-1A/B with Sentinel-2A/B (Costa *et al.*, 2021; Luca *et al.*, 2022; Tavares *et al.*, 2019), and Sentinel-1A/B, Sentinel-2A/B with Landsat to improve accuracy of the classification and crop type mapping. Yan *et al.* (2021) have evaluated the performance of a combination of

GaoFen-1 and Sentinel-2A/B data for crop mapping. Recently, high-resolution satellite imagery of PlanetScope combine Sentinel-1 with Sentinel-2 for LULC Classification of Central Brazil (Vizzari, 2022) and crop type mapping of Smallholder Farms using the SVM algorithm in eastern India (Rao *et al.*, 2021).

However, the integration of Sentinel-2A/B and PlanetScope data will significantly increase classification accuracy. But, it requires enormous volumes of computing space to combine multi-source satellite data (data fusion), download, store and process the multi-source satellite imagery for large-scale crop classification is a challenging task. The GEE platform provides a high-performance computational cloud platform for big data analysis like satellite imagery without need for local storage (Gorelick *et al.*, 2017). Furthermore, this platform offers up-to-date and ready-to-use remotely sensed (Landsat, Sentinel-1, Sentinel-2 and PlanetScope archives) geospatial and ancillary datasets for researchers and academicians. It also allows to share the developed JavaScript codes of various analyses within many users.

The GEE platform is being widely used due to supporting some of the most popular machine learning classifiers; CART (Breiman *et al.*, 1984; Mather & Tso, 2016) is “a binary decision tree classifier that employs a predetermined threshold, RF is an ensemble classifier that uses multiple CART” (Breiman, 2001) and SVM is also an excellent classifier that uses input parameters as the kernel type (Burges, 1998; Mountrakis *et al.*, 2011). But, the RF algorithm of machine learning is being extensively used for the classification of remotely sensed datasets in the GEE Platform, due to its non-parametric nature (Amani *et al.*, 2020; Praticò *et al.*, 2021).

Therefore, integrated Sentinel-2A/B and PlanetScope data and machine learning approach have been employed for the estimation of maize crop in North Bihar, India and the performance of various machine learning algorithms are compared and analysed in present study. In this study, overall objective was to develop a web-based IDE code by JavaScript API for robust mapping of crops using satellite imagery and machine learning algorithms. In addition, the performance of multiple classifiers of machine learning was to be assessed. This study will be helpful for crop cultivation

management, crop insurance and for making a decision support system to prioritise the input subsidy to farmers. The developed JavaScript code will also be helpful for the application of crop monitoring tools.

5.2 Materials and Method

The present study is conducted in North Bihar, India for the winter crop growing season of 2022 and 2023. Wherein, PlanetScope, Sentinel-2 satellite data from January to April of 2022 & 2023 and GT data (Table 3.1 & Fig. 3.2) were used for maize crop identification and acreage estimation. In this process, the Sentinel-2 imagery was harmonized and rescaled to the exact resolution as PlanetScope (4.77 m) using the bicubic interpolation function (Keys, 1981) in the GEE cloud platform (Fig. 5.1). After each processing step, Normalized Differential Vegetation Index (NDVI) was calculated from the harmonized imagery.

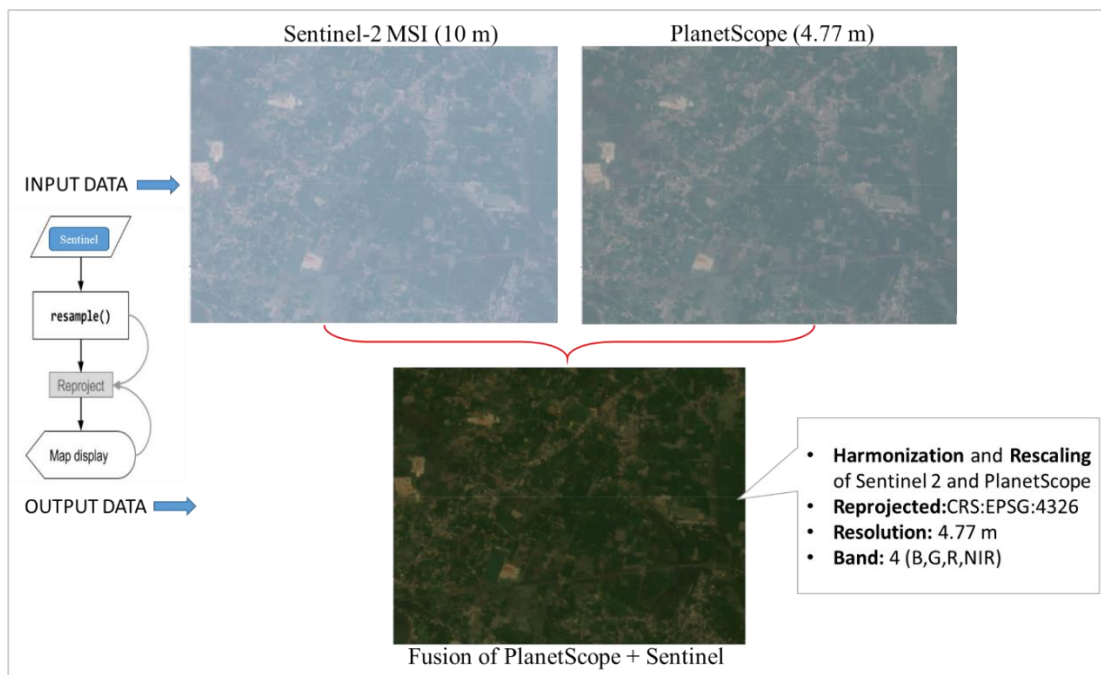


Figure 5.1: Harmonization, rescaling and fusion of PlanetScope and Sentinel-2A/B MSI data (1B)

5.2.1 NDVI Generation

NDVI is commonly utilized for crop mapping and vegetation indices (Rouse Jr *et al.*, 1973). Because the value of NDVI for each pixel varies with different stages of crop growth (Fig. 5.2). “These values remain low at the time of sowing, rise by maturity and fall again at the time of harvesting. These variations help to discriminate the maize crop from other crops” (Kumar *et al.*, 2022).

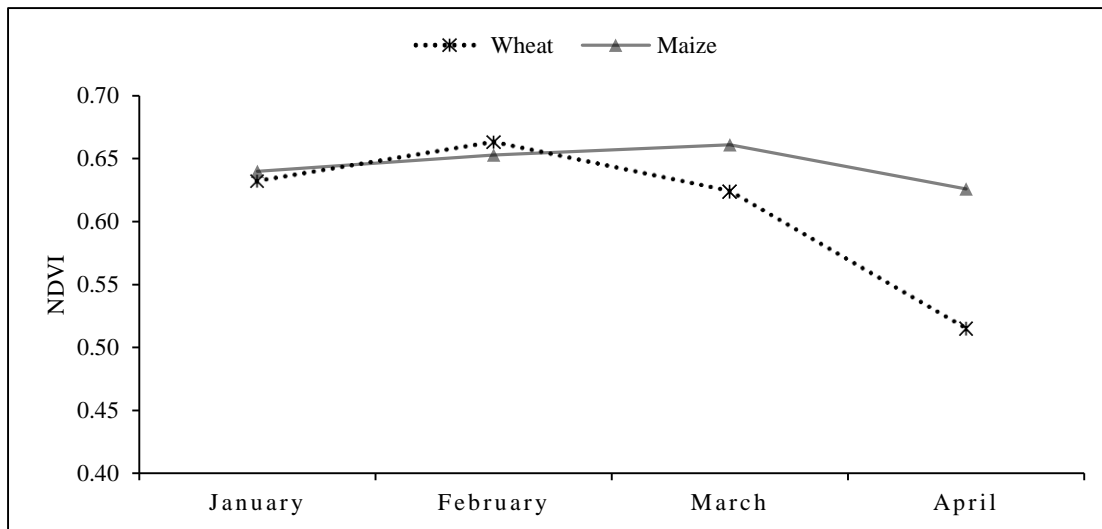


Figure 5.2: NDVI Profile presenting the growth profile of Maize and Wheat.

NDVI for four dates (January, February, March and April) has been generated from each satellite data i.e. Sentinel-2A/B and PlanetScope. So, a total of 8 band images were generated for each year (2022 & 2023). In North Bihar, maize and wheat are the major crops grown in the Rabi season (November to April). Its growth profile is depicted in figure 5.2 using NDVI values.

5.2.2 Methodology

In the present study, an integrated Sentinel-2 and PlanetScope data, machine learning approach and GEE cloud computing platform are employed to improve crop classification, prediction and mapping efficiency and accuracy of various ML models. CART, RF and SVM models have been used for identification and mapping of maize crops using GT data as collected from the study area. Then, the performance of utilized

approaches was compared and analysed in terms of overall classification accuracy. A flowchart of the methodology is shown in Fig. 5.3.

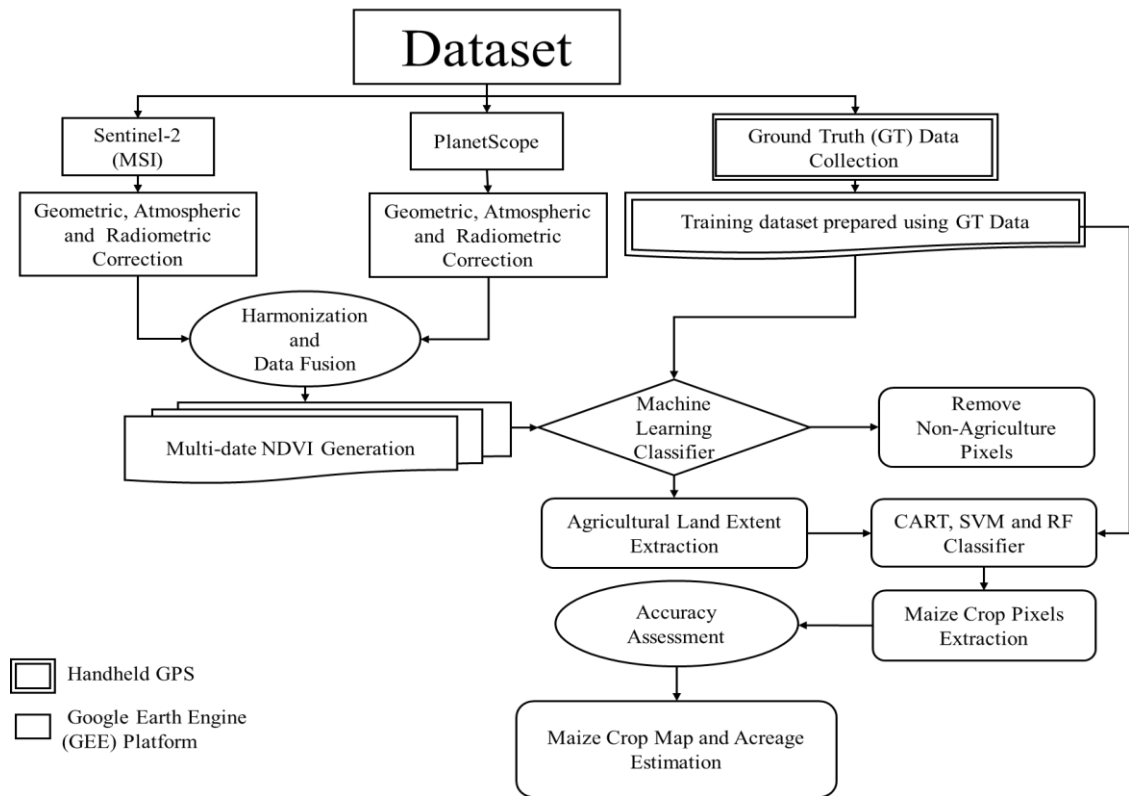


Figure 5.3: Methodology for crop acreage estimation

5.2.2.1 Google Earth Engine (GEE)

Satellite data processing is a challenging task due to the requirement of separate computational systems, storage and local space along with the software. To overcome this, Google launched the efficient and advanced cloud computing platform "Google Earth Engine " in 2010.

It provides a unique cloud computing platform to process different types of big datasets through improved algorithms and JavaScript codes without requiring local storage or downloading raw imagery, especially on a nearly real-time basis (Kumar *et al.*, 2022). It also allows to share the developed JavaScript codes of various analyses within many users.

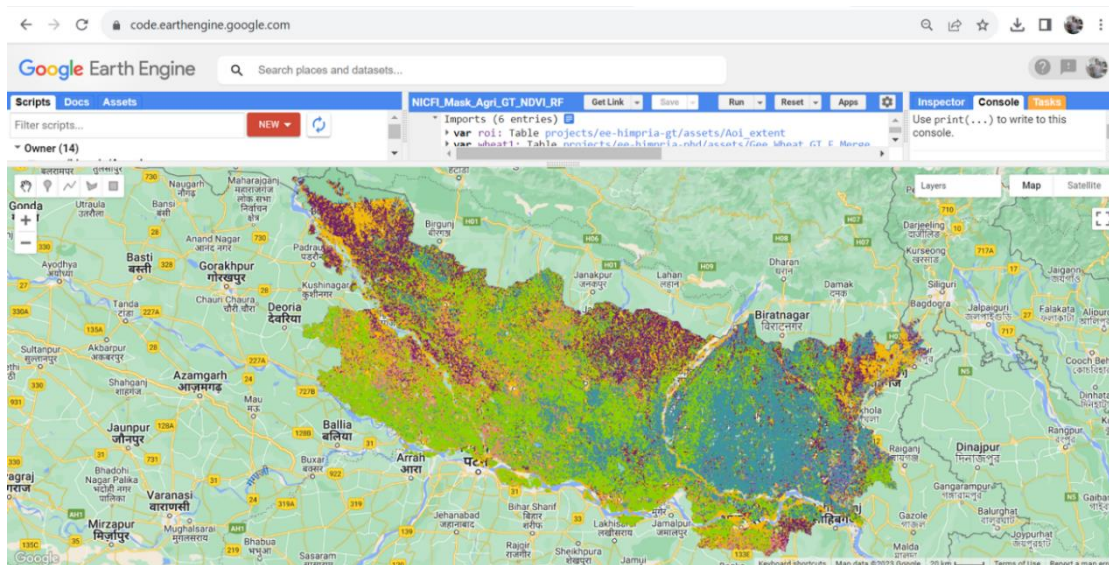


Figure 5.4: Crop mapping through GEE platform.

The GEE cloud computing platform has used for the entire analysis work and JavaScript codes have been developed ([https://code.earthengine.google.com/users %2Fhimpria%2Fconference%3ACropland_Extent](https://code.earthengine.google.com/users%2Fhimpria%2Fconference%3ACropland_Extent)) for robust mapping and monitoring of maize crop (Fig. 5.4).

5.2.2.2 Machine Learning Algorithm Used for Crop mapping

A. Random Forest

Random Forest (RF) algorithm of machine learning is a popular supervised classification classifier technique, which includes multiple decision trees and takes averages to improve the accuracy of predictions of any datasets (Breiman, 2001) (Fig. 5.5). This classifier can learn the characteristics of the target features based on the training dataset and identify these learned features in the unclassified data. (Belgiu & Drăguț, 2016). So, it is commonly used by the remote sensing and GIS community (Shelestov *et al.*, 2017a).

Hence, the random forest algorithm has been utilized in this study for maize identification and mapping using the GEE cloud-based geo-computing platform (Oliphant *et al.*, 2019). Whereas, we found that the RF algorithm performed better with PlanetScope data and also with the integration of PlanetScope and Sentinel-2A/B data. Their overall accuracy is 84.31% (kappa 0.79) with Sentinel-2 data, 90.17% (kappa

0.89) with PlanetScope data, and 95.53% (Kappa 0.91) with the integration of PlanetScope and Sentinel-2A/B data.

In addition, it is observed that the integration of PlanetScope and Sentinel-2A/B data is increasing the accuracy of classification and prediction. In the GEE platform, this classifier is relatively fast and capable of handling big data such as the entire satellite data of the present study.

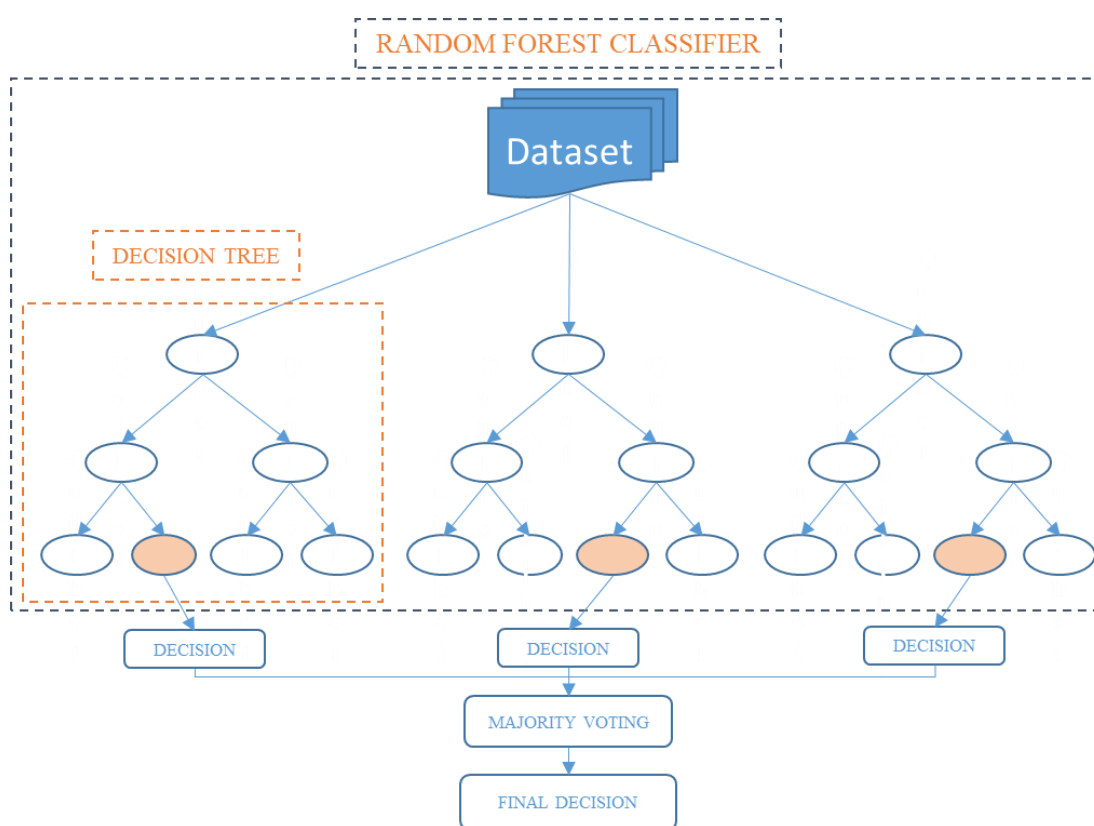


Figure 5.5: Random Forest classifier based maize identification

B. Support Vector Machine (SVM)

SVM classifier is popular a non-parametric supervised method to use for classification problems, regression problems, and uniqueness detection (Shelestov *et al.*, 2017a). It works on the concept of margin, this margin is the shortest distance between the decision boundary and the training dataset. It has also the capability to fit an optimal hyperplane to separate the dataset into separate number of pre-defined

classes using a training sample (Fig. 5.6). Which helps in getting a good classification result from complex datasets. So, this algorithm is being commonly used for crop type classification, mapping and monitoring. In the present study, SVM classifier has been used for maize crop mapping using the GEE cloud computing platform. Whereas, we found that the SVM classifier performed well with big datasets. Their overall accuracy is 76.99% (kappa 0.69) with PlanetScope data, 77.48% (kappa 0.69) with Sentinel-2 data, and 83.13% (Kappa 0.77) with the integration of PlanetScope and Sentinel-2A/B data.

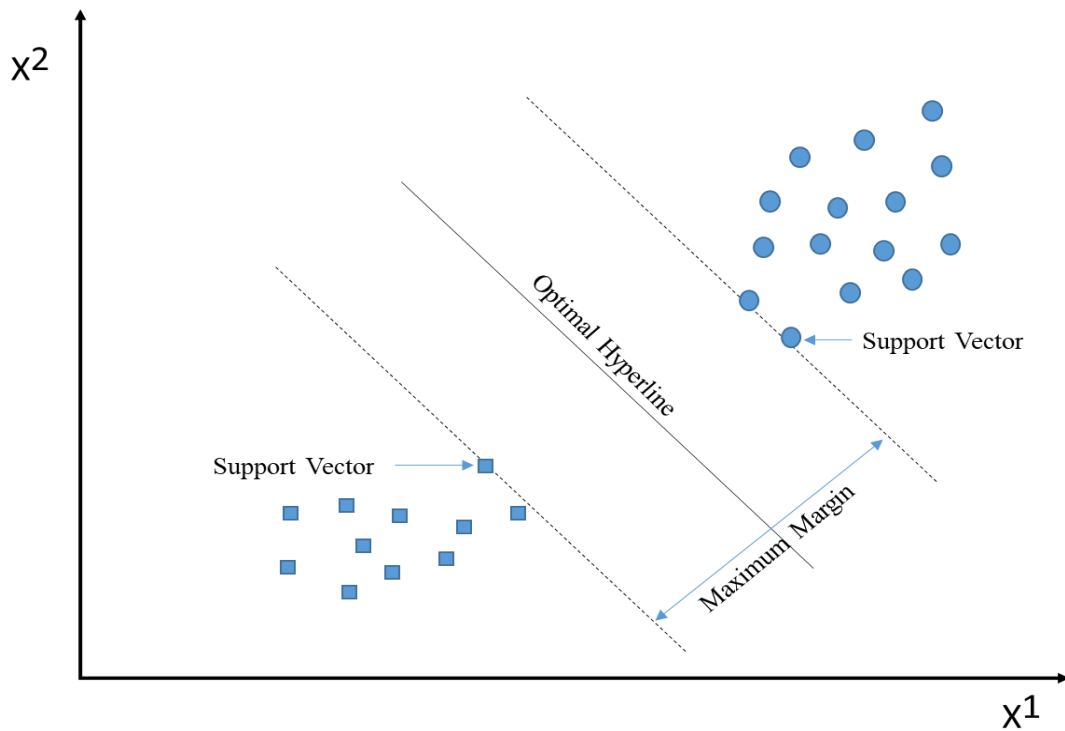


Figure 5.6: Support Vector Machines Classifier based maize identification

C. Classification and Regression Tree (CART)

The CART is a non-parametric decision tree (DT) classifier that employs a predetermined threshold (Breiman *et al.*, 1984). This algorithm builds regression trees using Gini's impurity index and “entropy” to obtain the information that can be expressed mathematically as equations (ii & iii)

$$\text{Gini}(E) = 1 - \sum_{i=1}^c p_i^2 \quad (\text{ii})$$

$$\text{Entropy: } H(x) = 1 - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (\text{iii})$$

“At each internal node, the tree’s leaf nodes in corresponding division regions are decided by related splitting rules” (Luo *et al.*, 2022). Which is sorting down the tree from the root to leaf nodes, as depicted in Fig. 5.7. So, this classifier can be used for classification, regression and predictive modelling problems.

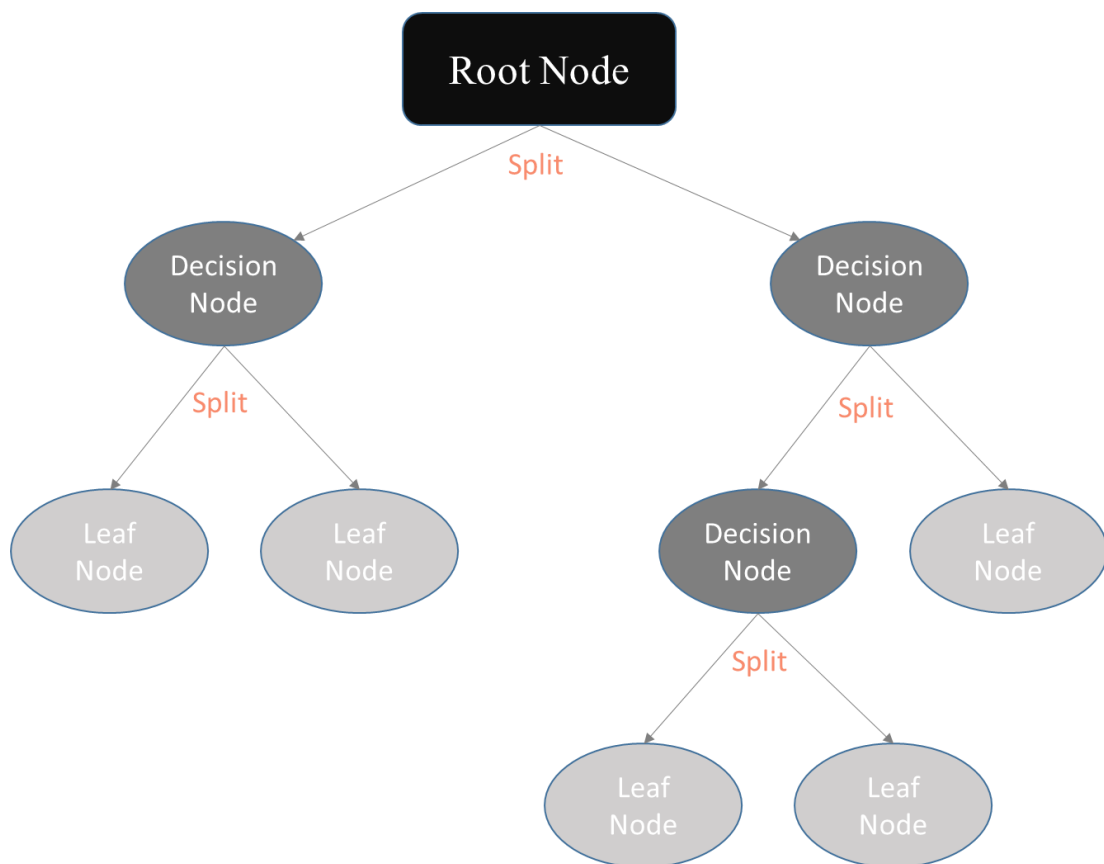


Figure 5.7: Classification and Regression Tree Classifier based maize identification

In this study, the CART classifier was used for maize crop mapping using the GEE cloud computing platform. Whereas, we observed that the CART classifier performed well with big data such as time-series datasets of remote sensing. Their overall accuracy is 78.06% (kappa 0.71) with PlanetScope data, 88.59% (kappa 0.85)

with Sentinel-2 data, and 90.54% (Kappa 0.87) with the integration of PlanetScope and Sentinel-2A/B data.

5.3 Results and discussion

5.3.1 Crop Identification

In the present study, about 1016 GT data were used to train the ML algorithms (SVM, CART and RF) for classification and identification of maize crop using the GEE platform. In addition, training accuracy has been also assessed and observed that the training accuracy of CART, RF and SVM is 99%, 98% and 83% respectively (Fig. 5.8).

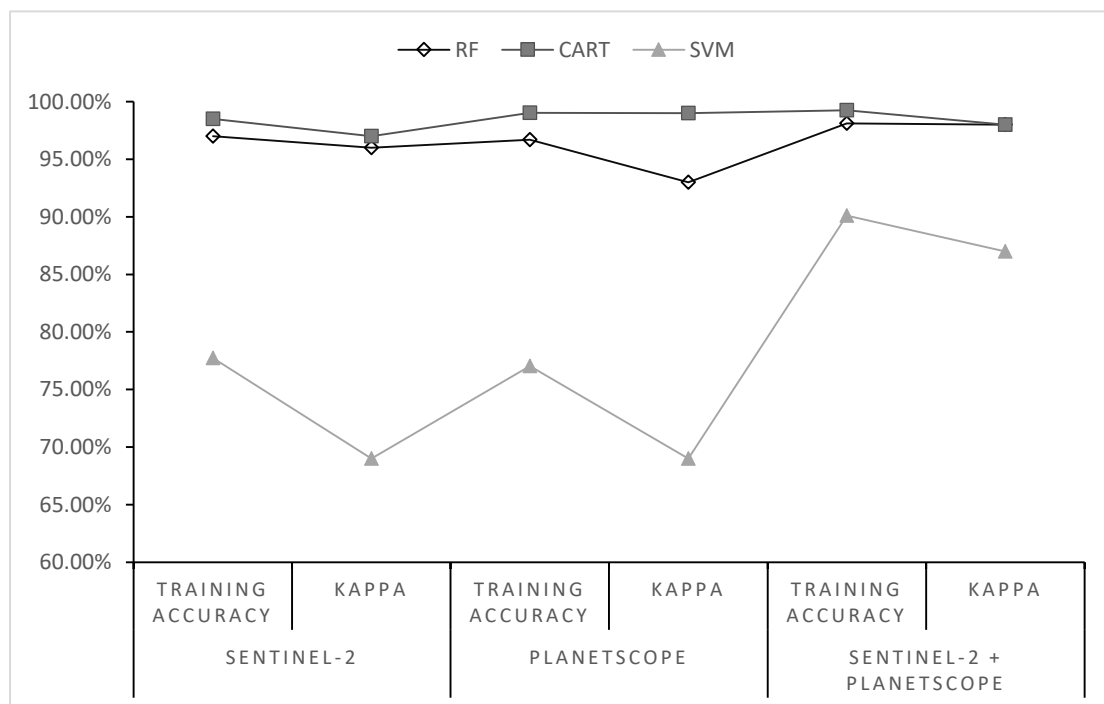


Figure 5.8: Training accuracy in the present study

Then SVM, CART and RF algorithms have been applied to each dataset such as Sentinel-2, PlanetScope, and the combined satellite datasets (Sentinel-2 and PlanetScope) to extract the pixels of maize crop. The satellite data and ML algorithms-wise extracted pixels/ patches of maize crop have been represented in the Fig. 5.9. The red polygons (Fig. 5.9) have been used on the classified images to identify differences in detection according to satellite data and ML algorithms.

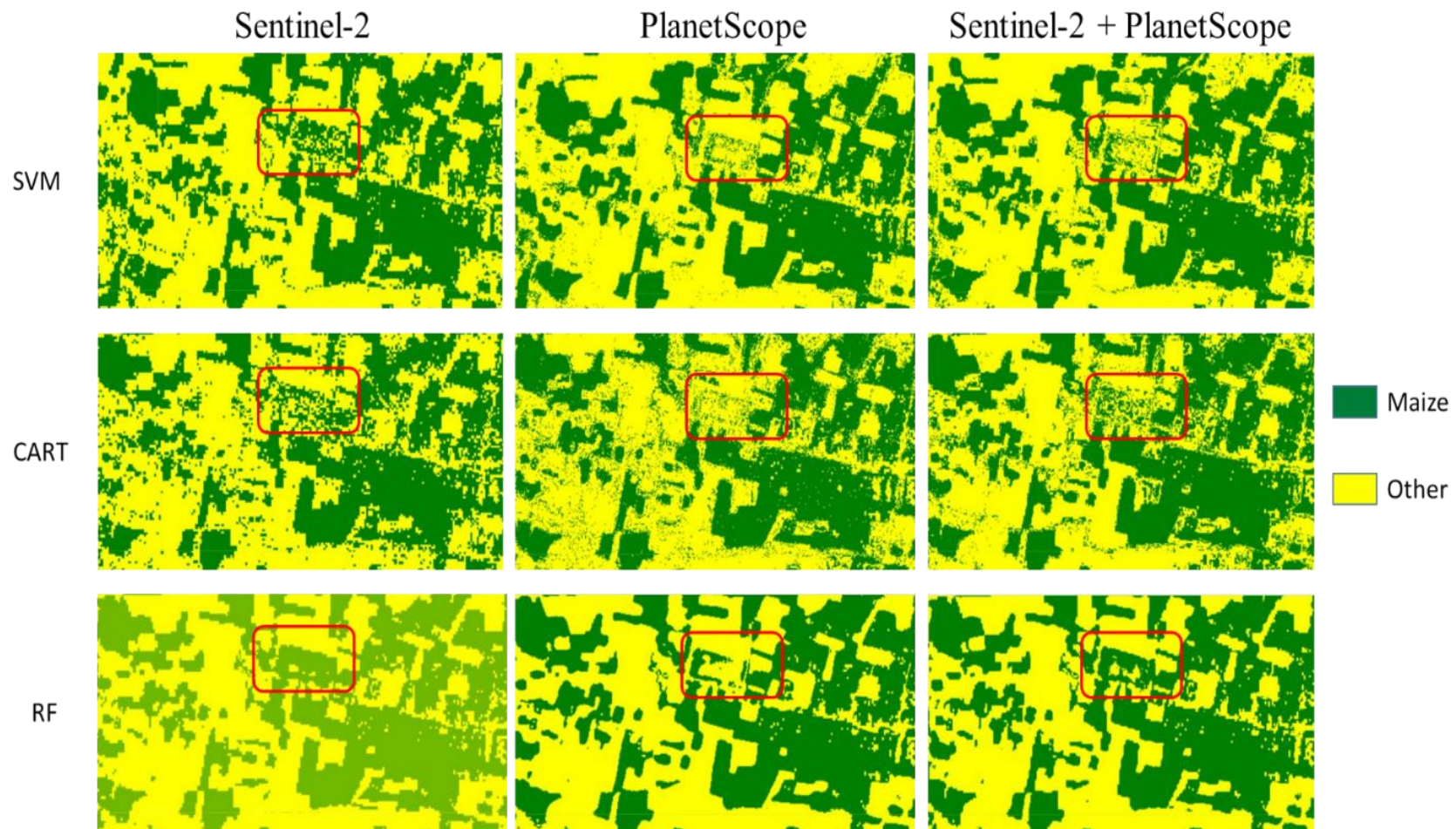


Figure 5.9: Maize crop mapping from Sentinel-2A/B, PlanetScope and integration of Sentinel-2A/B and PlanetScope sensors using SVM, CART and RF Classifiers.

In this study, it is observed that the 10-meter spatial resolution data from the Sentinel-2 satellite sensor makes it easier to depict large patches of maize crops. The PlanetScope sensor can more accurately extract the small patches of maize crops (\geq about 5 m²), which will be helpful for mapping small-holding farms as the size of land holdings is small in most areas of South Asia (<2 ha) (Jain *et al.*, 2016).

In addition, we found that the integration of PlanetScope and Sentinel-2A/B data yielded better performance and helped in the accurate demarcation of maize crops from other crops. Hence, the combination of PlanetScope and Sentinel-2A/B data is a feasible and accurate approach for crop identification.

5.3.2 Acreage Estimation of Maize Crop

A combination of PlanetScope and Sentinel-2A/B satellite imageries have been used for maize crop area estimation. The district-wise (Table 5.1) area was calculated based on identified pixels of the maize crop using the JavaScript code (*ee.Geometry.Polygon.area*) in the GEE platform. In this study, the analysis shows that about 321252 ha in 2021-22 and 329504 ha in 2022-23 of land are devoted to the cultivation of maize in North Bihar, India. Finally, a maize crop map has been generated using the GEE platform and QGIS Software for pictorial/ synoptic representation (Fig. 5.10 & 5.11).

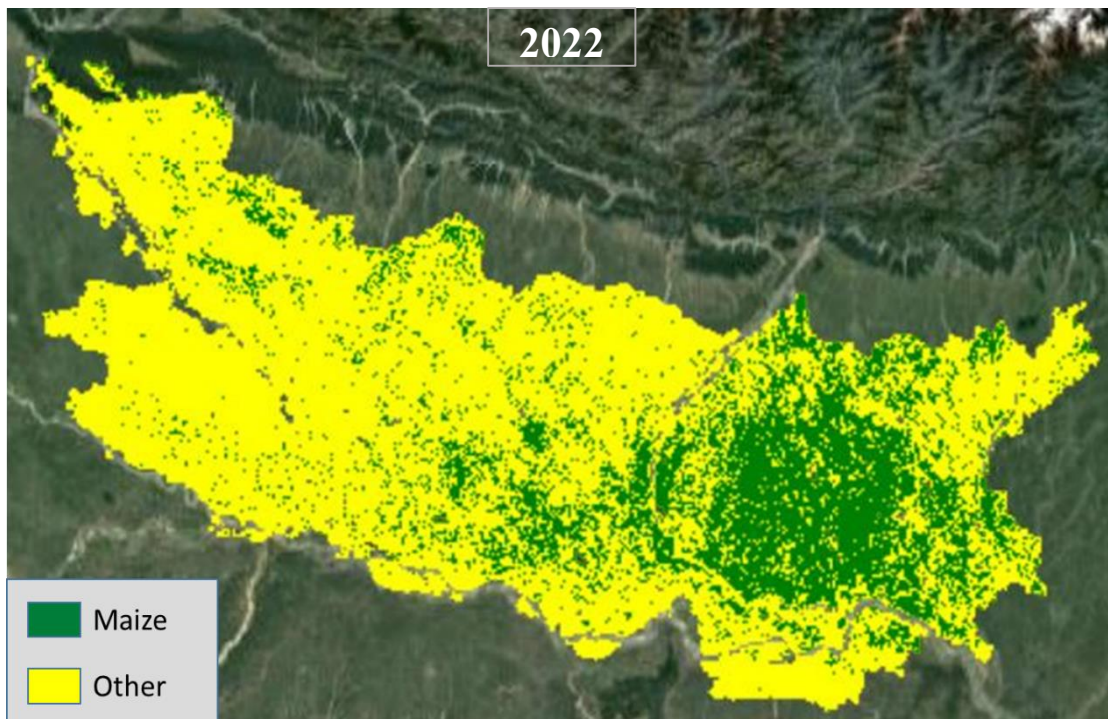


Figure 5.10: Maize crop map of 2022.

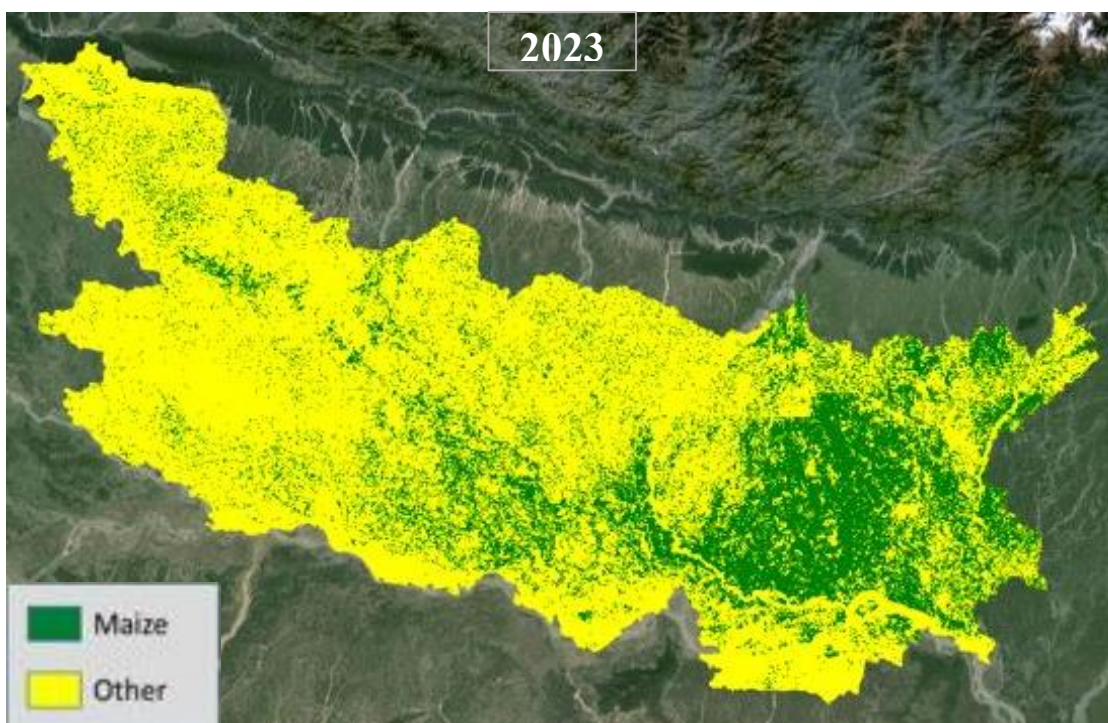


Figure 5.11: Maize crop map of 2023.

Table 5.1: District-wise acreage estimation

Sl. No.	District	2021-22		2022-23	
		Acreage		Acreage	
		(ha)	('000 ha)	(ha)	('000 ha)
1	Arariya	30783.62	30.78	31205.17	31.21
2	Begusarai	40052.20	40.05	41613.02	41.61
3	Bhagalpur	17049.88	17.05	17372.91	17.37
4	Darbhanga	5424.63	5.42	5609.89	5.61
5	Gopalganj	3906.81	3.91	3974.49	3.97
6	Katihar	38140.84	38.14	38334.54	38.33
7	Khagaria	33753.83	33.75	35601.01	35.60
8	Kishanganj	1118.43	1.12	1179.55	1.18
9	Madhepura	17734.14	17.73	17868.15	17.87
10	Madubani	42.86	0.04	45.05	0.05
11	Muzaffarpur	10537.18	10.54	10901.84	10.90
12	W. Champaran	1811.47	1.81	1830.45	1.83
13	E. Champaran	4946.25	4.95	5138.56	5.14
14	Purnia	47938.58	47.94	49514.03	49.51
15	Saharsa	12425.07	12.43	12540.16	12.54
16	Samastipur	28005.15	28.01	28770.67	28.77
17	Saran	5643.95	5.64	5733.90	5.73
18	Sheohar	1502.40	1.50	1529.93	1.53
19	Sitamarhi	3143.64	3.14	3108.66	3.11
20	Siwan	5277.02	5.28	5388.48	5.39
21	Supaul	4711.26	4.71	4703.49	4.70
22	Vaishali	7302.91	7.30	7540.50	7.54
Total (AOI)		321252.12	321.25	329504.45	329.50

Note: ha- hectare, AOI- Area of Interest, '000- Thousand

5.3.3 Performance of Machine Classifiers with various datasets

In the present study, the performance of machine classifiers has been evaluated within the GEE cloud computing platform with different datasets. Wherein, it is observed that RF outperforms CART and SVM algorithms in the GEE platform with PlanetScope data and also with the integration of PlanetScope and Sentinel-2A/B data (Fig. 5.11). With the integration of Sentinel-2A/B and PlanetScope data, the overall accuracy of crop mapping improved by about 6%, indicating that the integration of Sentinel-2A/B and PlanetScope data is more suitable for crop identification.

Although, CART outperforms RF and SVM algorithms with Sentinel-2A/B data. In the GEE platform, RF and CART classifiers are relatively overfitted, fast and capable of handling big data such as the entire satellite data of the present study (Fig. 5.12). But, the performance of the SVM classifier in the GEE platform was not fast, well and adequate (Shelestov *et al.*, 2017b).

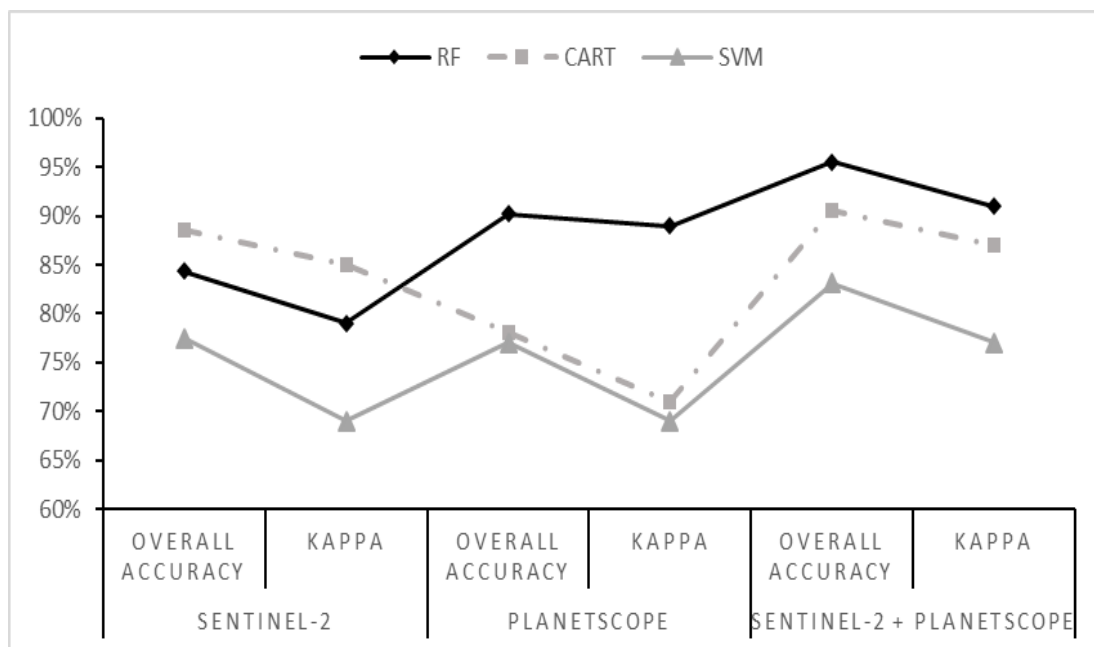


Figure 5.12: Performance of Classifiers in the GEE Cloud Computing Platform

5.3.4. Prediction Accuracy

The GT dataset is randomly split into two sections; viz. 70% for classifier training and 30% for accuracy and validation purposes. In the GEE platform, the Confusion Matrix method (Stehman, 1997) is commonly used for accuracy assessment, which has also been used in the present study.

Whereas, we found that RF outperforms CART and SVM algorithms in the GEE platform with PlanetScope data (OA = 90.17 %, Kappa 0.89) and also with the integration of PlanetScope and Sentinel-2A/B data (OA = 95.53%, Kappa 0.91). But, CART outperforms RF and SVM algorithms with Sentinel-2A/B data (OA = 88.59 %, Kappa 0.85). In addition, it is observed that the integration of PlanetScope and Sentinel-2A/B data is increasing the accuracy of classification and prediction (Table 5.2).

Table 5.2: Performance of classifiers in terms of accuracy

Performance of Machine Learning Classifiers						
Classifier	Sentinel-2		PlanetScope		Sentinel-2 + PlanetScope	
	OA	Kappa	OA	Kappa	OA	Kappa
RF	84.31%	0.79	90.17%	0.89	95.53%	0.91
CART	88.59%	0.85	78.06%	0.71	90.54%	0.87
SVM	77.48%	0.69	76.99%	0.69	83.13%	0.77

Note: RF- Random Forest, CART- Classification and Regression Trees, SVM- Support Vector Machine, and OA- Overall Accuracy

Table 5.3: Year-wise available acreage data of *rabi*/winter maize crop as per the report of Directorate of Economics and Statistics, Government of Bihar.

District	Rabi Maize (2020-21)	Rabi Maize (2019-20)	Rabi Maize (2018-19)
	Area (ha)	Area (ha)	Area (ha)
Arariya	27543	28265	26499
Begusarai	36309	29516	19881
Bhagalpur	17324	15686	15629
Darbhanga	5611	5478	6175
Gopalganj	3689	4120	3242
Katihar	36011	30965	29745
Khagaria	32871	32942	32965
Kishanganj	1124	1299	1268
Madhepura	16307	18757	18713
Madubani	0	0	53
Muzaffarpur	10201	10091	9838
Pashchim Champaran	1707	1834	1873
Purbi Champaran	4484	4457	4544
Purnia	47464	45267	35786
Saharsa	12264	11733	11471
Samastipur	27207	27385	31389
Saran	5472	5318	4543
Sheohar	1459	1427	1314
Sitamarhi	3155	3335	3464
Siwan	4970	4970	4970
Supaul	4604	4599	4595
Vaishali	7030	6960	7046
Total (AOI)	306806	294404	275003

5.3.5 Validation

Due to the unavailability of the latest data, a direct comparison between the estimated crop area and the latest data is not feasible (Hudait *et al.*, 2022). However, we present the rabi maize acreage of different years from 2018-19 to 2020-21 (Table 5.3),

as reported by the Directorate of Economics and Statistics, Government of Bihar. The analysis shows that some districts are relatively close match between the extracted area and the government data of 2020-21, such as Bhagalpur, Darbhanga, Gopalganj, Kishanganj, Madhepura, Pashchim Champaran, Saharsa, Samastipur, Sheohar, and Siwan (Table 5.3). However, a few districts show a difference between the extracted area and the government data. For example, Arariya, Begusarai, Katihar, Khagaria, Muzaffarpur, Purnia, Supaul, and Vaishali have higher extracted areas (about 1 to 10 %) compared to the government data of 2020-21. Furthermore, the obtained results have been validated with GT data, which is collected in the form of polygons of the plot of the study area.

5.4 Discussion

Recently, high-resolution PlanetScope data (4.77m) has become available in the GEE platform in collaboration with the NICFI program, which can be utilized for crop identification, mapping and monitoring in tropical regions. But, ESA's Sentinel-2 (10m) data can be used on a global scale. Hence, this is an excellent opportunity for researchers and the scientific community to improve accuracy of crop mapping and monitoring. Which can be helpful in precision agriculture.

Nowadays, some recent literature is suggesting that multi-source satellite data is more reliable for crop mapping with state-of-art ML algorithms (Prins & Niekerk, 2021; Vizzari, 2022; Yan *et al.*, 2021a). Hence, the present study is conducted using multi-source data, which demonstrated that maize crops can be accurately mapped using machine learning and a combination of PlanetScope and Sentinel-2A/B data. In which, it is necessary to integrate multisource satellite data to accurately map the crop. Thus, it requires a powerful computing system, large storage and GIS software.

To overcome this, The freely available GEE cloud computing platform used to access and integrate multi-source remote sensing satellite data. Because the GEE platform offers powerful computing capabilities to handle big datasets of remote sensing that can be utilized for crop mapping from small to large scales. It can also deal

with weather phenomena like clouds and shadows with the help of GEE's tool 'CLOUDY PIXEL PERCENTAGE' (Kumar *et al.*, 2022; Shelestov *et al.*, 2017a).

To integrate the satellite data, the Sentinel-2 imagery was re-projected, harmonized and rescaled to the same resolution with PlanetScope (4.77 m) using the bicubic interpolation function (Keys, 1981) in the GEE platform. Then, the PlanetScope and Sentinel-2A/B data are integrated with the help of available tools in the GEE platform. This approach is a novel and cost-effective approach that can automatically demarcate maize crop extent using machine learning and integrated satellite data. Which can be applied anywhere for large-scale precision agriculture.

In the present study, it is observed that the 10-meter spatial resolution data from the Sentinel-2 satellite sensor makes it easier to depict large patches of maize crops. The PlanetScope sensor can more accurately extract the small patches of maize crops (\geq about 5 m²), which will be helpful for mapping small-holding farms as the size of land holdings is small in most areas of South Asia (<2 ha) (Jain *et al.*, 2016). In addition, we found that the integration of PlanetScope and Sentinel-2A/B data yielded better performance and helped in the accurate demarcation of maize crops from other crops (Fig. 5.10). Hence, the combination of PlanetScope and Sentinel-2 data is a feasible and accurate approach for crop identification.

Furthermore, this study also assessed the performance of machine classifiers for the precise mapping of maize crops. Whereas, it is observed that RF outperforms CART and SVM algorithms in the GEE platform with PlanetScope data and also with the integration of PlanetScope and Sentinel-2A/B data. Although, CART outperforms RF and SVM algorithms with Sentinel-2A/B data. It is also observed that RF and CART algorithms are overfitted for accurate mapping of crops. Likewise, the performance of the SVM algorithm within the GEE platform was not good enough for crop mapping. After extensive analysis, i have extracted the acreage of each district of North Bihar of 2022 and 2023. This is made possible through the use of a sophisticated JavaScript code, ML classifier and integrated dataset from PlanetScope and Sentinel-2A/B satellites.

5.5 Conclusion

In the present study, Integrated satellite imagery of Sentinel-2 and PlanetScope, and Machine learning approaches have been used for maize crop mapping in North Bihar, India. In addition, the Performance of CART, Random Forest and SVM Algorithms of Machine Learning has been evaluated for crop mapping using the GEE Cloud Computing Platform. Whereas, we found that RF outperforms CART and SVM algorithms with PlanetScope data and also with the integration of PlanetScope and Sentinel-2A/B data. The use of integrated satellite imagery and RF algorithms improved the accuracy of crop mapping by about 6%. This shows that the integration of Sentinel-2A/B and PlanetScope imagery is more suitable for crop identification. But, CART outperforms RF and SVM algorithms with Sentinel-2A/B data. Hence, the CART algorithm can be implemented for accurately crop mapping with Sentinel-2 data. In this study, a maize crop map has been generated using the GEE platform and QGIS software, and their overall accuracy is 95.53%. It is also found that approximately 321252 ha in 2021-22 and 329504 ha in 2022-23 are devoted to maize cultivation in North Bihar, India. This tested algorithms, edited and developed web-based JavaScript code can be applied anywhere in the world for crop mapping. This study is expected to be helpful in making decision support systems for crop cultivation management, crop insurance and prioritizing input subsidies to farmers. In addition, the maize crop area map/maize extent layer will be used for crop yield and production estimation, which is explained in the next chapter.

CHAPTER-6 CROP YIELD AND PRODUCTION ESTIMATION

6.1 Introduction

The rise of greenhouse gas emissions, deforestation, rapid urbanization and environmental degradation has led to global warming and weather pattern changes. This has altered the growing season, soil moisture levels and availability of crops, it especially affects developing countries. These changes have severe consequences on agriculture, food security, economics and human health globally. Addressing these issues is vital for ensuring a sustainable future for agriculture and those dependent on it.

Therefore, crop acreage estimation and determining their yield is crucial in monitoring and managing agriculture at a field level. This will help in making quick decisions in case of crop failure and obtaining crop insurance. The increasing demand of crop insurance in agriculture has led to a demand for plot-level acreage and yield estimation data. This requires detailed information, but traditional methods like field surveys and crop cutting experiment are costly and time-consuming (Hudait and Patel, 2022). So, there is a requirement of a systems or methods for cost-effective, time-efficient and precise crop statistics. which will be able to support micro-level agricultural planning and ensure accurate crop insurance coverage.

However, accurately mapping and monitoring of smallholding agricultural plots are challenging due to the diversity in crops, farming systems and limited plot size (Lowder *et al.*, 2016; Paliwal and Jain, 2020). In this context, remote sensing (RS) is a technology that allows us to gather information about the Earth's surface and atmosphere from a distance, typically from satellites or aircraft at regional and global scales (Paliwal and Jain, 2020). In recent years, advancements in remote sensing technology have enabled us to make more accurate, recurrent and timely estimation of crop acreage and yield, making it an essential tool for supporting agricultural decision-support systems.

Nowadays, remotely sensed Earth observation satellite data based precisely mapping and monitoring of agricultural land (Singh *et al.*, 2020) is being widely used due to publicly accessible high-resolution Sentinel-2 and PlanetScope satellite data (Kumar *et al.*, 2022). These datasets are being extensively used by researchers, government agencies and agricultural companies for crop area/ acreage estimation, crop type delineation, crop conditions and growth patterns mapping, crop insurance and yield prediction to provide farmers with valuable insights to improve crop productivity and efficiency (Neetu and Ray, 2019; Paliwal and Jain, 2020). In addition, the integration of remote sensing techniques and machine learning approaches has revolutionized the agriculture sector, which is now commonly used for crop type mapping, distribution and acreage estimation (Prins *et al.*, 2020; Mizuochi *et al.*, 2021; Hudait and Patel, 2022). Several studies have been conducted for crop acreage estimation using remote sensing and machine learning approaches. Wherein, multi-sensor data and ML algorithms were utilized to estimate crop area and high accuracy was demonstrated in mapping and monitoring different types of crops (Costa *et al.*, 2021; Yan *et al.*, 2021; Veerabhadraswamy *et al.*, 2021; Kumar *et al.*, 2023).

Traditional methods of crop yield estimation are biophysical crop-simulation models, agronomic techniques, crop-growth models and agrometeorological statistical techniques, which are time-consuming and costly (Betbeder *et al.*, 2016). The statistical regression method (Rembold *et al.* 2013) is a widely used method for crop yield estimation because it has fewer requirements for input data compared to other methods such as machine learning and artificial intelligence. This method is based on “empirical relationships between measured in-situ crop yield and vegetation indices values” (Mehdaoui and Anane, 2020). The vegetation indices used in this method provide information about the greenness, productivity, and health of crops.

According to previous studies, various spectral indices were developed using remotely sensed satellite data such as NDVI (Rouse Jr *et al.* in 1973), GNDVI (Gitelson *et al.*, 1996), LSWI (Xiao *et al.* 2002), EVI (Huete *et al.*, 2002), which have been utilized in many studies for crop mapping and yield estimation. Among these indices,

NDVI is the most popular vegetation index, which is widely used in empirical regression models to predict crop yields (Franch *et al.* 2019).

Some previous studies have shown that extensive crop simulation (Jin *et al.*, 2019) and crop-cutting experiment datasets were used for crop yield estimation, which is an accurate and reliable method. However, this method is time-consuming and expensive, which limit the scope of remote sensing models. Because there is challenging to obtain reliable input data from agricultural fields for calibrating and validating remote sensing models for crop yield estimation.

Nowadays, remote sensing technology has emerged as a promising alternative to traditional methods of crop yield assessment (Awad *al.*, 2019). The use of remotely sensed satellite dataset is a cost-effective, less time consumption and efficient way to collect input data for crop yield estimation. With the use of remote sensing satellites, aircraft, or unmanned aerial vehicles (UAVs) can provide wide coverage of agricultural lands and gather information about crop health, growth, and yield potential (Wang *et al.*, 2020).

In recent years, MODIS, Landsat, LISS-3&4, Sentinel and PlanetScope datasets are increasingly being used for crop yield forecasting and monitoring (Mehdaoui and Anane, 2020). These datasets have wide area coverage, and near real-time information, allowing for better prediction of future yields. The use of satellite data has significantly improved accuracy and efficiency of yield forecasting. “Despite its increasing importance, remote sensing faces several significant technical challenges in its operational application such as low spatial resolution” (Imanni *et al.*, 2022).

Low spatial resolution means that small features, such as individual crops or trees, may not be visible in the image. In north Bihar, most farmers have an average landholding size of less than 2 hectares, which can result in significant errors in crop production estimation (Jain *et al.*, 2016). The publicly accessibility of Sentinel-2 (10m spatial resolution) of the European Space Agency (ESA) and PlanetScope (4.77m spatial resolution) through the NICFI program provides high spatial and temporal

resolution satellite data (Planet, 2017), which have minimized the drawbacks and improved its capabilities to identify the crop in small holding size plot too. Apart from high-resolution imagery, the sentinel-2 satellite has specific spectral 3 red-edge bands (704, 740 and 783 nm), which is more sensitive to monitoring the vegetation conditions (Dong *et al.* 2019) and yield estimation (Cao *et al.* 2017) than red-infrared indices (Delegido *et al.* 2013).

In many parts of the world, the reliance on traditional statistical methods for estimating agricultural production often results in unreliable and delayed reports regarding expected crop yields and production (Ali *et al.* 2022). However, the advent of remote sensing techniques and machine learning approaches provide a remarkable alternative for the estimation of crop production (Khaki *et al.* 2021).

Presently, the overall object of the study is to develop the model/ algorithms for yield and production estimation of maize crop using remotely sensed satellite data, an evolved spectral index and machine learning algorithm. Wherein, a combination PlanetScope and Sentinel-2 data has been utilized for maize crop acreage estimation and Sentinel-2 satellite data used to extract the spectral indices such as REI, NDVI, GNDVI, LSWI and EVI for maize crop yield estimation of north Bihar, India. The production is determined by combining information on the amount of land (acreage) dedicated to the maize crop and the average yield data.

6.2 Materials and Method

6.2.1 Experimental Site Selection

Agriculture is a significant contributor to the economy of India. However, it is facing various challenges such as flood, climate change, long-term droughts, and rainfall variability. Which have significant implications for crop production and food security. Therefore, the development of a monitoring system for agricultural production in India has become essential to plan and predict crop yields accurately. The northern part of Bihar is chosen for the experimental site (Fig. 1.9) and model validation in the present study due to its unique position as a major contributor to rabi maize production in India (Singh *et al.*, 2018).

A reliable and comprehensive satellite datasets (PlanetScope and Sentinel-2A/B) and CCE data (Fig. 3.6) are required for the estimation of maize crop yield and production. In this study, Integrated satellite datasets of PlanetScope and Sentinel-2A/B were used for crop classification and identification. Wherein, Sentinel-2A/B satellite data was used to extract the spectral indices such as REI, NDVI, GNDVI, LSWI and EVI for maize crop yield estimation.

6.2.2 Methodology

In the present study, a freely accessible Google's GEE cloud computing platform (CCP) has been used to process voluminous dataset of multiple dates NICFI-Planet Scope remote sensing satellite imagery. Because, this platform has the capability to analyse, visualize and access massive datasets such as Landsat, Sentinel, MODIS, PlanetScope etc. achieve without the requirement of local storage system (Amani *et al.*, 2020) and GIS Software. In the present study, a visual representation of the methodology is presented in Figure 6.1.

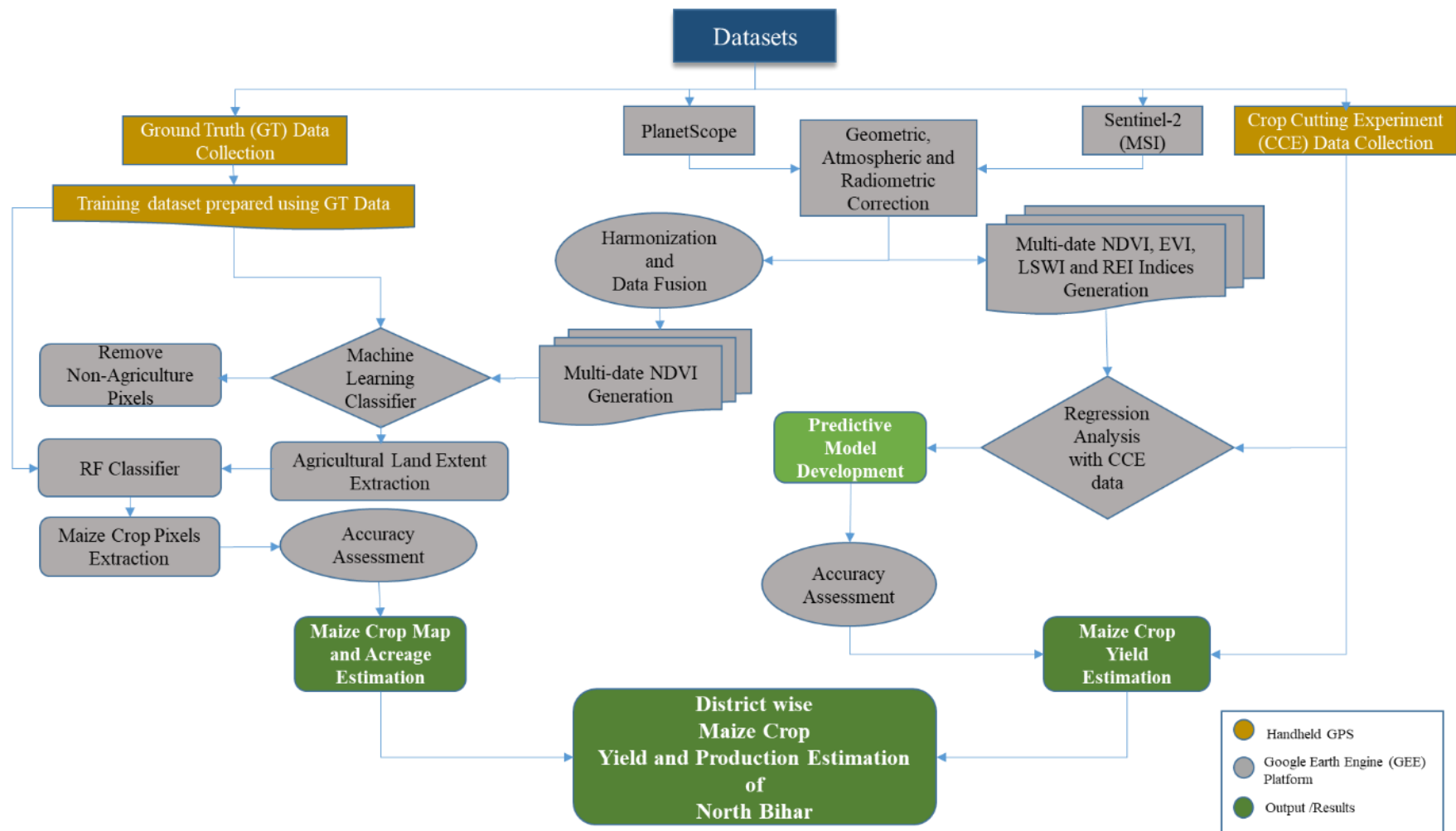


Figure 6.1: Flowchart illustrating the steps of acreage and yield estimation

6.2.3 Cloud Computing Platform for data processing

Satellite data processing is essential for crop mapping, monitoring and many other applications. However, processing satellite data requires vast amounts of computational resources, storage capacity, and specialized software. This could be prohibitively expensive, time-consuming and challenging for researchers and organizations working with satellite data, especially those who do not have access to expensive hardware and software.

In the present study, the GEE cloud geo-computing platform have been used to provide solutions to these challenges. Hence, this platform has been used for satellite data processing and analysis. The JavaScript codes have been also developed (<https://code.earthengine.google.com/?scriptPath>) for maize crop acreage estimation (Fig. 6.2). to prepared the input data of yield estimation model.

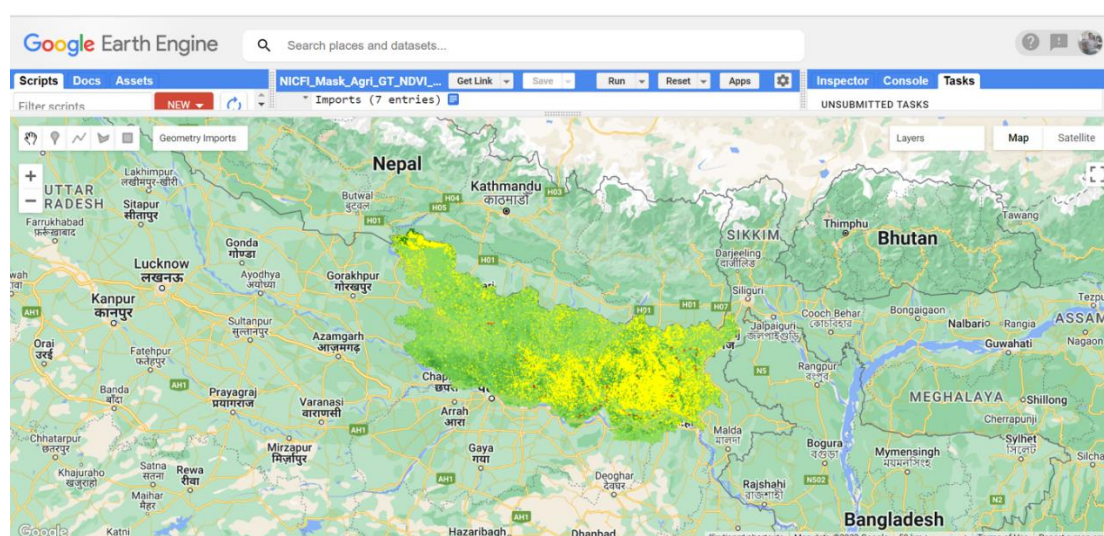


Figure 6.2: The GEE based maize acreage estimation for input parameter for yield estimation

6.2.4 Acreage Estimation

The Sentinel-2 imagery was harmonized and resampled (**Script-1**) to the exact resolution as PlanetScope (4.77 m) using the bicubic interpolation function (Keys, 1981) in the GEE platform. Fusion algorithm have also applied a to merge Sentinel-2 and NICFI data using Web-based GEE JavaScript code “ *var fusion = ee.Image.cat (sentinel2, nicfi);*” .

```

var sentinel2_resampled = filteredSentinel2.map(function(image){
  return image.resample('bilinear').reproject({
    crs: nicfiBasemap.projection(),
    scale: 4.77
  });
});

```

After each processing step, Multi-date NDVI was generated from the harmonized imagery and stacked in a single image (Fig. 6.3). Then, Agricultural land was extracted using Random Forest (RF) algorithms (*ee.Classifier.smileRandomForest*). Again, Random Forest (RF) algorithms of machine learning was used in the present study for crop classification, prediction and mapping of maize crops using GT datasets. Accuracy assessment has been also performed for classified datasets to validate the output/result.

6.2.4.1 Multi-date NDVI Extraction

The NDVI is being widely used as a tool for measuring vegetation, mapping crops, and monitoring changes in vegetation growth over time (Rouse Jr *et al.* in 1973). The NDVI is working based on the principle that healthy plants absorb light in the visible spectrum for photosynthesis, and reflect light in near-infrared (NIR) spectrum. However, stressed or unhealthy plants reflect less NIR light, resulting in a lower NDVI value.

The NDVI can calculate by the formula:

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$$

where NIR is the reflectance in the near-infrared spectrum, and Red is the reflectance in the red spectrum. The resulting NDVI values range from -1 to +1.

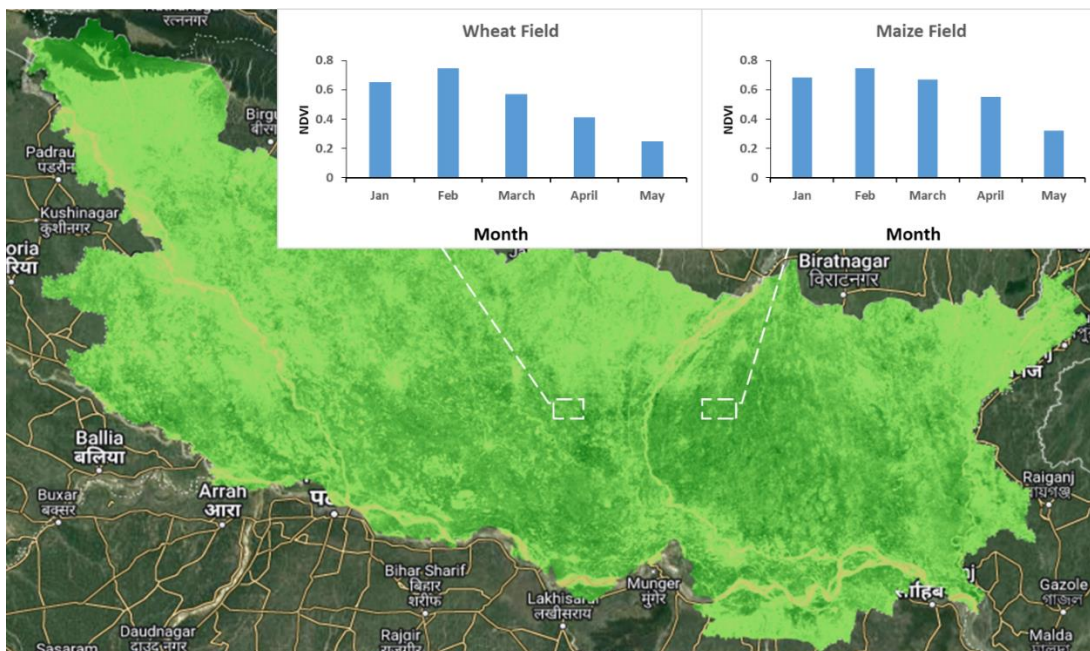


Figure 6.3: Multi-date stacked NDVI from experiment site.

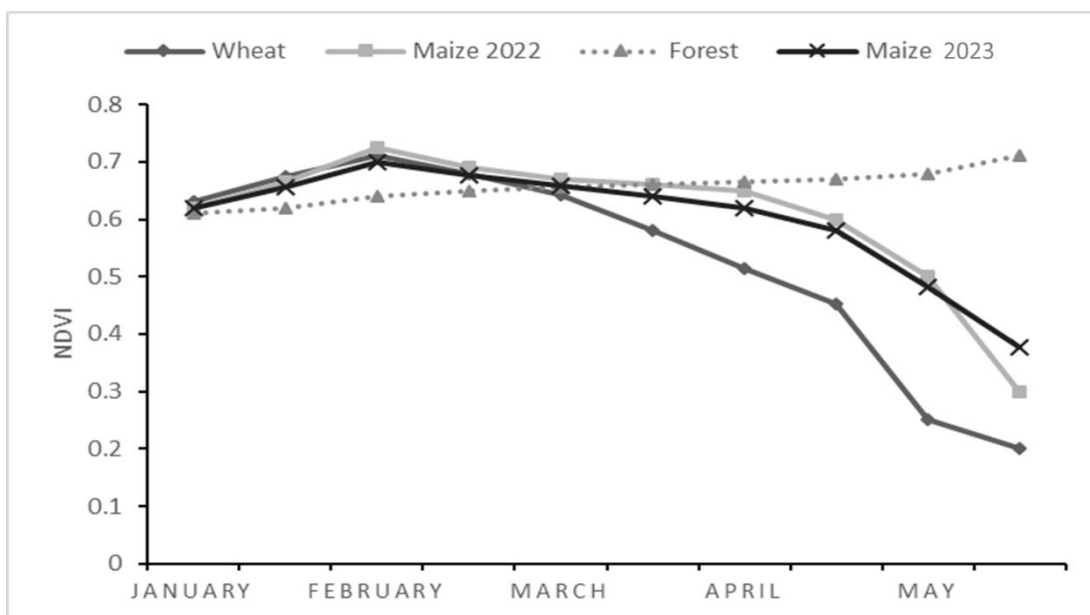


Figure 6.4: NDVI values are depicting the growth profile of maize, wheat, forest and water.

In crop mapping, NDVI is extensively used to distinguish between different crop types, estimate crop acreage, yield, and monitor crop health. For instance, researchers have used this index to know the impact of drought on corn and soybean

yields in the United States (Wang *et al.*, 2019). In vegetation mapping, NDVI has been used to monitor vegetation extent in the Amazon rainforest (Asner *et al.*, 2005).

In this work, five dates (January to May 2022) NDVIs were obtained from each satellite i.e. Sentinel-2 and Planet Scope of North Bihar and stacked (Fig. 6.3) it for estimation of maize crop. It observed that NDVI values of maize crops remain low at the time of sowing, rise by maturity and fall again at the time of harvesting (Fig. 6.4). Maize and wheat are the main crops grown in the Rabi season in North Bihar.

In the present study, a supervised classification approach has been used, which involves the use of training samples to classify an image into various land cover classes, including different crop types. This method works on training samples, which can be collected from ground truth data or existing land cover maps, and then used to train a classification algorithm, such as RF to classify the image.

6.2.4.2 Use of Random Forest Algorithms for Acreage Estimation

Random Forest (RF) is a popular ML algorithm that is widely used in supervised classification problems. It is a type of ensemble learning method that combines multiple decision trees to make more accurate and robust predictions. The RF algorithm was first introduced by Leo Breiman and Adele Cutler in 2001 and has since become one of the most widely used ML algorithms in various fields *viz*; finance, healthcare, agriculture and remote sensing.

In the RF algorithm, multiple decision trees are built using randomly selected subsets of the training data and features. The trees are then combined through a process called bagging, which averages their predictions to produce the final output. The use of multiple trees and feature subsets helps to reduce overfitting and improve the accuracy and generalization of the model (Fig. 6.5).

One of the key advantages of the RF algorithm is its ability to handle high-dimensional data and complex interactions between features. It can also handle missing values and noisy data, making it a robust and versatile method for various applications. Several studies have shown the effectiveness of the RF algorithm in various

classification tasks. Cutler *et al.* (2007) demonstrated that “the RF algorithm outperformed other popular machine learning algorithms such as SVM and ANN” in classifying remote sensing data.

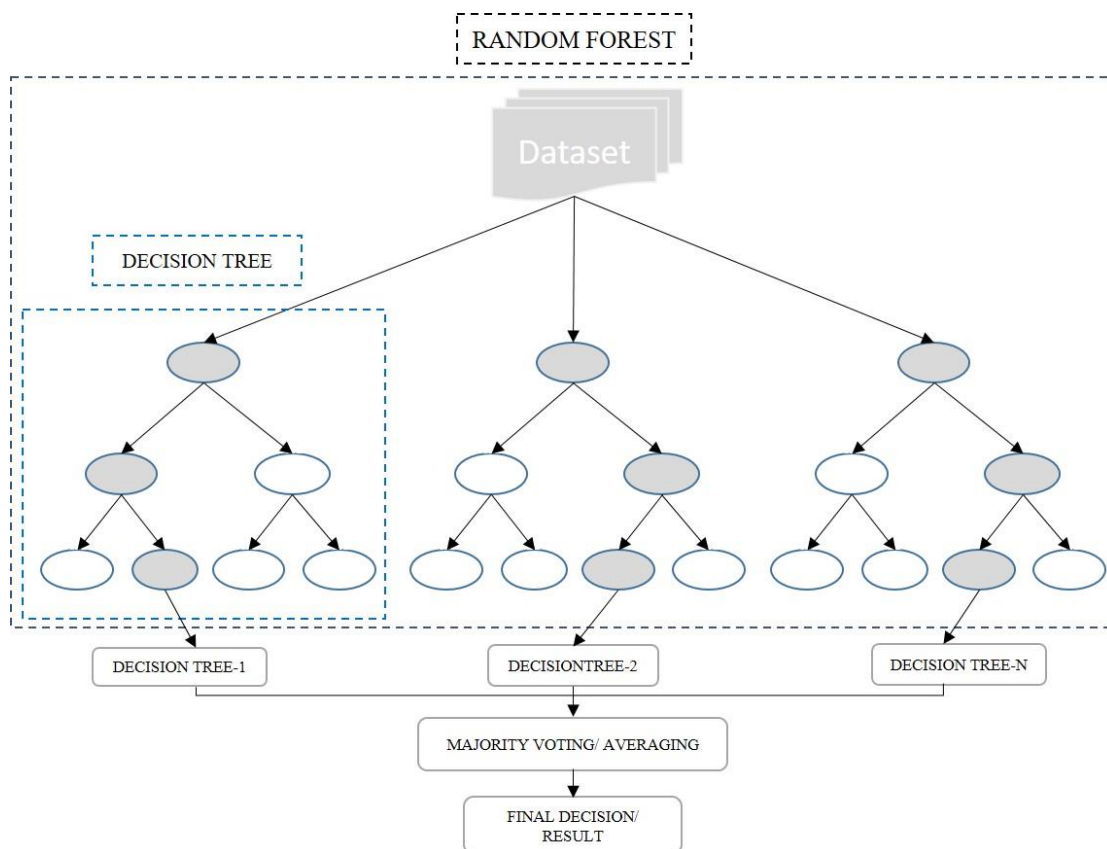


Figure 6.5: Flowchart of RF algorithms to estimate acreage for input parameter for yield estimation

In another study, RF algorithm achieved high accuracy in predicting cancer prognosis based on gene expression data (Chen *et al.* (2014). In crop mapping, RF can be used to classify remote sensing data into different crop types. Nowadays, the RF approach is being employed to map crops in various regions around the world, including North America, Europe, Africa, and Asia. Zhang *et al.* (2018) used this approach to map major crop types in China using Landsat and MODIS data. In another study by Yu *et al.* (2020), RF was used to map soybean fields in the United States using Landsat data.

Nowadays, RF and GEE platform is being commonly used for crop mapping and monitoring using high-resolution Sentinel-1 & 2 datasets and ML algorithms (Richards *et al.*,2021; Mandal *et al.*, 2021). Hence, the RF algorithm has been used to

map the acreage of maize crops of 2022 & 2023 using the GEE platform. Wherein, agricultural land was extracted from multi-date NDVI using Random Forest (RF) algorithm (ee.Classifier.smileRandomForest).

6.2.5 Yield Estimation

6.2.5.1 Input datasets for Model

- **Insolation and Fraction of absorbed photosynthetically active radiation (FPAR) data:**

Daily insolation datasets were obtained from the MOSDAC portal and converted to 8-day products. Eight-day FPAR datasets are also obtained from the NASA LP DAAC. Both datasets have been resampled to 10 m spatial resolution.

- **Radiation-use efficiency (RUE):**

RUE (ϵ_0) is the process of quantifying the photosynthetic conversion of solar radiation into plant biomass, which is a key parameter in agricultural science to assess the productivity of plants and crops (Kiniry *et al.* 1989; Shi *et al.* 2022). The formula for calculating RUE is expressed as:

$$\text{RUE} = \frac{\text{Biomass produced (g or kg)}}{\text{Solar radiation intercepted (MJ or } \frac{\text{MJ}}{\text{m}^2}\text{)}}$$

In the present study, the RUE of maize crop is taken from the literature, which is $3.5 \pm 0.7 \text{ g MJ}^{-1}$ (Kiniry *et al.* 1989; Hatfield 2014; Torres *et al.* 2017).

- **Indices Extraction and Evolution**

Nowadays, NDVI, GNDVI, EVI and LSWI etc. spectral indices are commonly used in Earth observation applications (Zeng *et al.*, 2022). Wherein NDVI, GNDVI, and EVI are vegetation indices that are used to monitor vegetation health and changes, LSWI is used to detect the presence of water on land surfaces (Chandrasekar *et al.*, 2010; Pan *et al.*, 2021). The choice of the index to use depends on the specific application and the environmental conditions being monitored. I have used the GEE cloud computing platform to reduce processing time, generate and export the required indices, and re-projected them to the World Geodetic System 84 (WGS-84) for further analysis.

In the present study, five date's (January to May) NDVI, GNDVI, LSWI, EVI and REI were obtained from Sentinel-2 satellite data of each year (2022 & 2023) for maize crop yield estimation. I have used the GEE cloud computing platform to reduce processing time, generate and export the required indices, and re-projected them to the World Geodetic System 84 (WGS-84) for further analysis. Recently developed "Red Edge Index (REI)" index performed well when estimating crop production. REI is calculated as the ratio between the top-of-atmosphere (TOA) reflectance of a red band around 665 nm and a Vegetation Red Edge (RE) band around 704-785 nm. In other words, REI computes vegetation status or photosynthetic activity by measuring the difference between RE (vegetation strongly reflects) and Red reflectance (absorption of vegetation). Its value varies from -1 to +1 and its formula is $(RE - Red) / (RE + Red)$.

$$REI = (RE - Red) / (RE + Red)$$

It can be calculated by band-4 (Red) and band-6&7 (RE) of Sentinel-2 satellite imagery. Densely vegetated areas will typically have positive or higher values, while water, urban, and non-vegetated (moisture-stressed vegetation) areas will have values that are close to zero or negative. Healthy vegetation (chlorophyll) reflects more RE and green reflectance compared to other wavelengths. In the present study, five date's (January to May) REI were obtained from Sentinel-2 satellite data of North Bihar.

6.2.5.2 Semi Physical Model for Yield Estimation

Presently, semi-physical model has been used for computation of maize crop biomass using Vegetation Index, RUE, PAR, FPAR, Rainfall and Temperature (Tripathy *et al.* 2021). The biomass of maize crop was calculated using equation (iv).

$$\text{Biomass Index} = \text{Indices} * \epsilon_0 * \text{PAR} * \text{FPAR} * \text{Rain} * \text{Temp} \quad (\text{iv})$$

6.2.5.3 Relationship between maize crop yield/CCE data and indices values

In the present study, a regression analysis was carried out (Fig. 6.9) and correlated the sum of multi-date biomass indices with CCE data averaged at district level. The relationship between crop yield and various vegetation indices is represented by the respective R-squared (R^2) values. R^2 is a statistical measure that indicates the

strength of the relationship between two variables, with a value of 1 indicating a perfect relationship and 0 indicating no relationship.

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{total sum of squares (SST)}}$$

$$= 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}.$$

Here, Red Edge Index has been used for cluster/district-wise maize crop yield estimation, as it is performing well for maize crop yield estimation. The formula for yield biomass (YB) estimation/Prediction based on REI index is as follows:

$$YB = (\text{Biomass Index} * \text{Intercept}) + \text{Coefficients} \quad (\text{v})$$

Or

$$REI \text{ YB} = (\text{Biomass} * REI * 7001) + 0.36050$$

This is cost-effective, efficient and less time-consuming method for yield estimation and predictions.

6.2.5.4 Production Estimation

The production of maize is determined by a combination of two key factors: the area of land (acreage) dedicated to growing maize and the average yield data. By integrating these two information, total maize production has been estimated of the study, which is vital for assessing supply, market trends, and planning agricultural strategies.

6.3 Result and Discussion

This study aimed to develop a method to estimate maize yield and production using remotely sensed satellite datasets such as Sentinel-2 and PlanetScope and the GEE cloud computing platform. This study has evaluated five vegetation indices (NDVI, EVI, GNDVI, LSWI, and REI) to develop the model. The results showed that

the GEE cloud computing platform is capable to fastly processing the data, demonstrating the potential for rapid and efficient processing of large remote sensing datasets. On the other hand, this developed red edge index performed well in terms of accurate estimation of maize crop yield.

6.3.1 Crop Acreage Estimation

Crop identification, discrimination, and acreage estimation using remote sensing satellite data can be achieved through several methods, including supervised and unsupervised classification, object-based image analysis, and machine learning algorithms.

In the present study, supervised classification approach has been used, which involves the use of training samples to classify an image into various land cover classes, including different crop types. This method works on training samples, which can be collected from ground truth data or existing land cover maps, and then used to train a classification algorithm, such as maximum likelihood, SVM, RF and CNN, to classify the image.

The PlanetScope provides high-resolution imageries at a spatial resolution of 4.77 meters, while Sentinel-2 offers imagery at 10 meters. The combination of PlanetScope and Sentinel-2 satellite imagery has proven to be an effective method for crop mapping (Rao *et al.*, 2021). Hence, the integrated satellite data of PlanetScope and Sentinel-2A/B were used for maize crop mapping and identification in North Bihar using RF algorithms of machine learning (ML) and the GEE platform, which is an effective and efficient method for crop mapping and monitoring (Fig. 6.6). Where, the study achieved overall 84 percent of accuracy (kappa 0.79) with Sentinel-2A/B datasets, 90 percent (kappa 0.89) with NICFI-PlanetScope datasets, and 95 percent (Kappa 0.91) with the combination of PlanetScope and Sentinel-2A/B datasets (Fig. 6.7). The area of target crop was estimated by identifying and mapping its spatial extent and then calculating the area using the scale and resolution of the image (Fig. 6.6).

Acreage map of maize crops has been also generated of North Bihar using the GEE platform and QGIS Software. Furthermore, cluster/district-wise maize crop area of 2022 and 2023 were estimated by using identified pixels of the maize crop (Table 6.1).

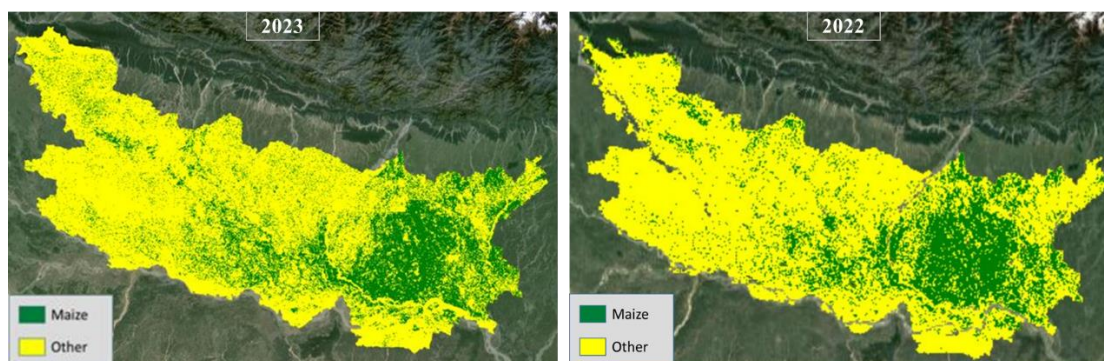


Figure 6.6: Map showing maize crop extent of study area.

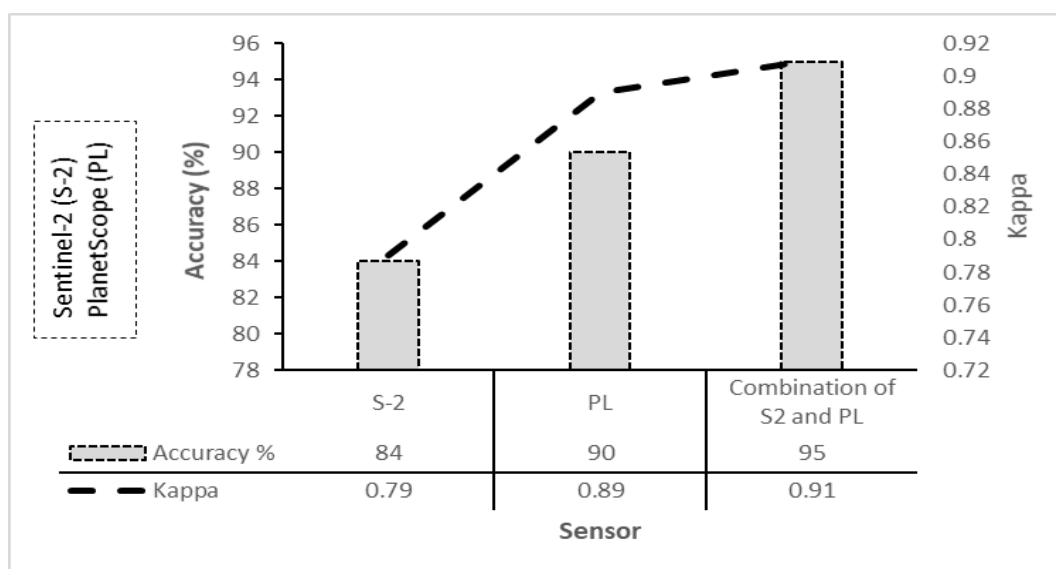


Figure 6.7: Sensor wise classification accuracy

Wherein, analysis shows that the district with the largest area under maize crop is Purnia with an area of 47,938 ha in 2021-22 and 49,514 ha in 2022-23. Begusarai is the second-largest district in both estimation years, with an area of 40,052 ha in 2021-22 and 41,613 ha in 2022-23. Katihar follows in third place with an area of 38,140 ha in 2021-22 and 38,335 ha in 2022-23. Khagaria and Arariya complete the list of the top

five largest districts, with areas of 33,753 and 30,783 ha, respectively. The remaining districts have areas ranging from 43 to 28,005 ha.

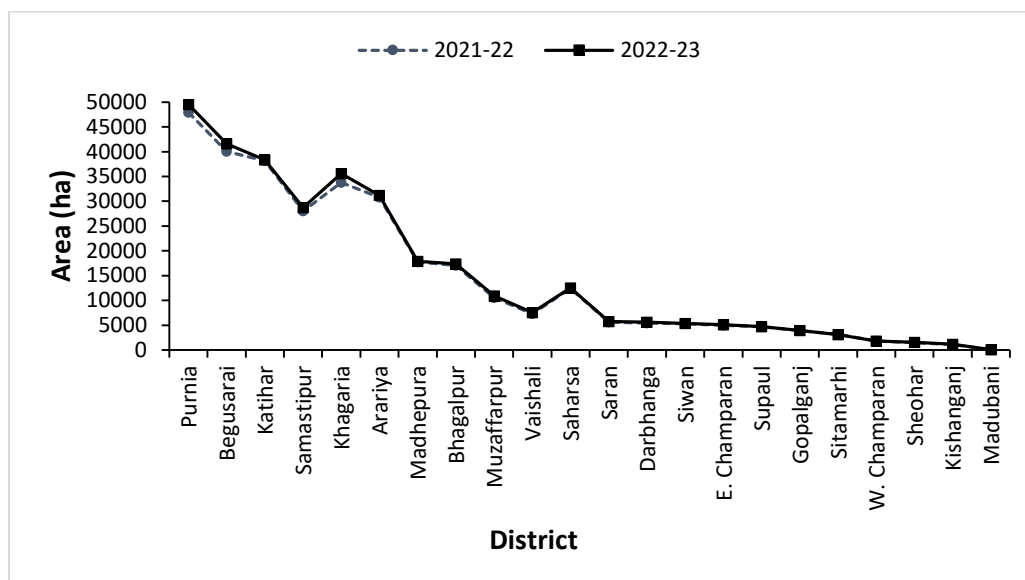


Figure 6.8: District wise maize crop area of study area.

In which, Saran, Siwan, Supaul, Gopalganj, Darbhanga, Purbi Champaran, Sitamarhi, Kishanganj, Sheohar and Madhubani are the smallest districts in terms of maize growing, each with an area of less than 6,000 hectares in both estimation years. The bottom 5 districts with the smallest area are Sitamarhi, Kishanganj, Sheohar, Madubani, and Pashchim Champaran (Fig. 6.8).

6.3.2 Crop Yield Estimation

Remote sensing satellite data with high spatial resolution is a powerful tool for crop monitoring and assessment of crop yield and production. This technique allows for frequent and rapid assessment of crop growth and yields over large areas, which is critical for decision-making in agriculture. The accuracy of crop classification and the use of appropriate vegetation indices are essential for accurate yield and production estimation.

In the present work, we found that the relationship between CCE yield and index is $R^2 = 0.84$ of REI, $R^2 = 0.79$ of NDVI, $R^2 = 0.76$ of EVI, $R^2 = 0.70$ of GNDVI and

$R^2 = 0.50$ of LSWI (Fig. 6.9). REI shows highly positive correlation with maize crop yield, higher REI values represent higher crop yield and lower REI values represent lower crop yield and NDVI is the second highest correlated index with maize crop yield. The EVI index is indicating a moderate relationship. The GNDVI index and LSWI index have indicated a weak relationship between this index and crop yield.

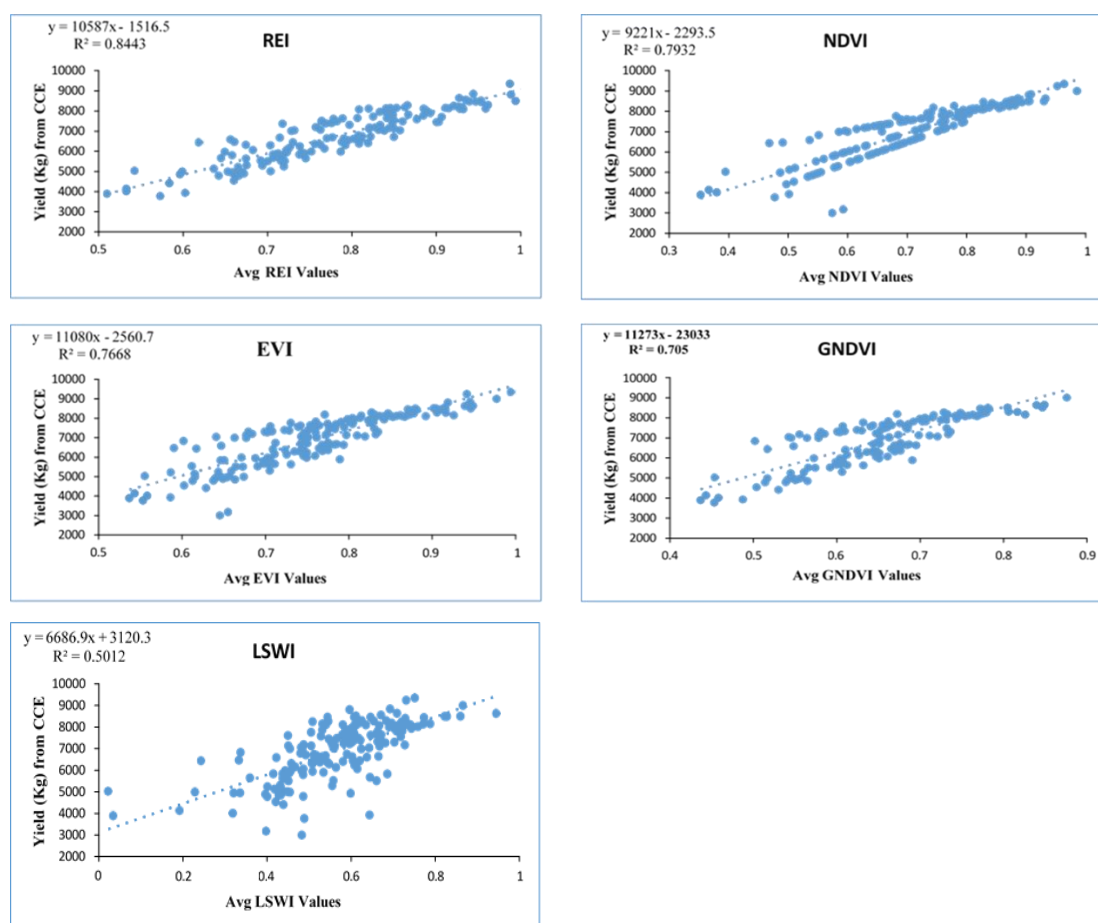


Figure 6.9: Indices-wise relationship between CCE maize yield and Indices values.

In the present study, district-wise average yield of maize crop was estimated of 2021-22 and 2022-23 using evolved REI index based formula (i). Overall highest yield was observed in 2021-22. Wherein, the highest yield is observed in Arariya district, with 9.98 tonnes per hectare, while the lowest yield is in Siwan district, with 3.69 tonnes per hectare (Fig. 6.10). It is interesting to note that there is a substantial difference in yield among the districts, with a difference of more than 5 tonnes per hectare between Begusarai and Samastipur. This variation may be attributed to various factors, including soil quality, weather conditions, and agricultural practices in each district.

The top five districts in terms of yield, Arariya, Katihar, Supaul, Purnia and Kishanganj are all located in the north-eastern part of Bihar (Simanchal Area), near the border with Nepal in both estimation years. These districts have a favourable climatic conditions, fertile soil and sufficient rainfall, which helps in higher yield. On the other hand, the districts with lower yields are mostly located in the western and central parts of north Bihar, such as Muzaffarpur, Gopalganj, Saran, and Vaishali in both estimation years.

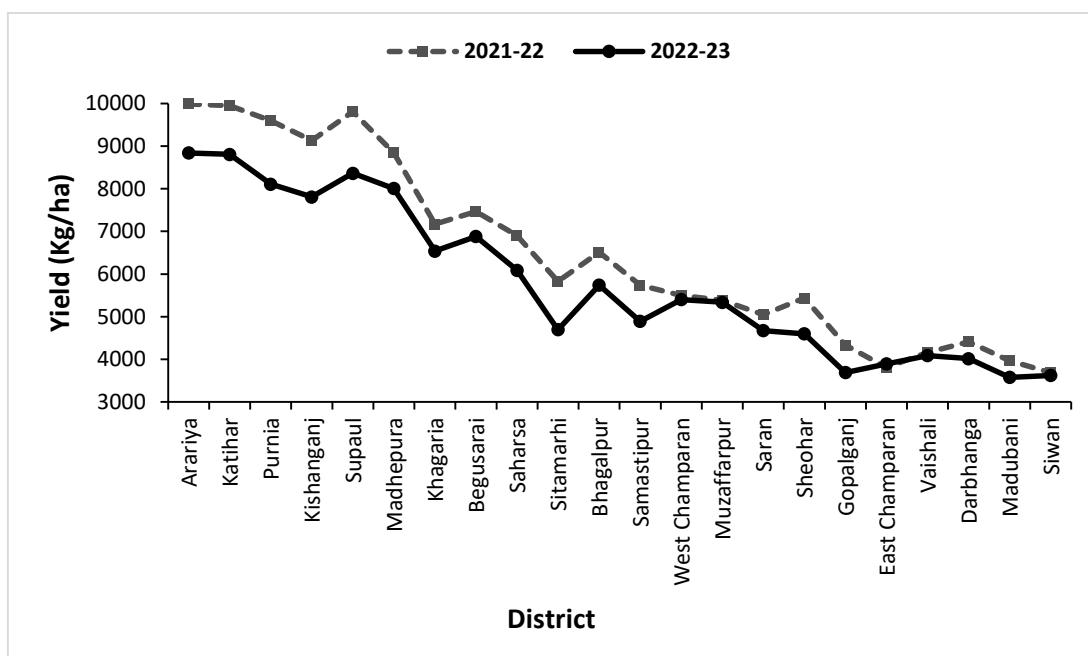


Figure 6.10: District wise maize crop yield (Kg/ha^{-1}) of study area

In addition, choropleth map of maize crop yield has been also prepared for better representation (Fig. 6.11 A & B). In which, red colour indicates higher yield and light orange colour indicates lower yield.

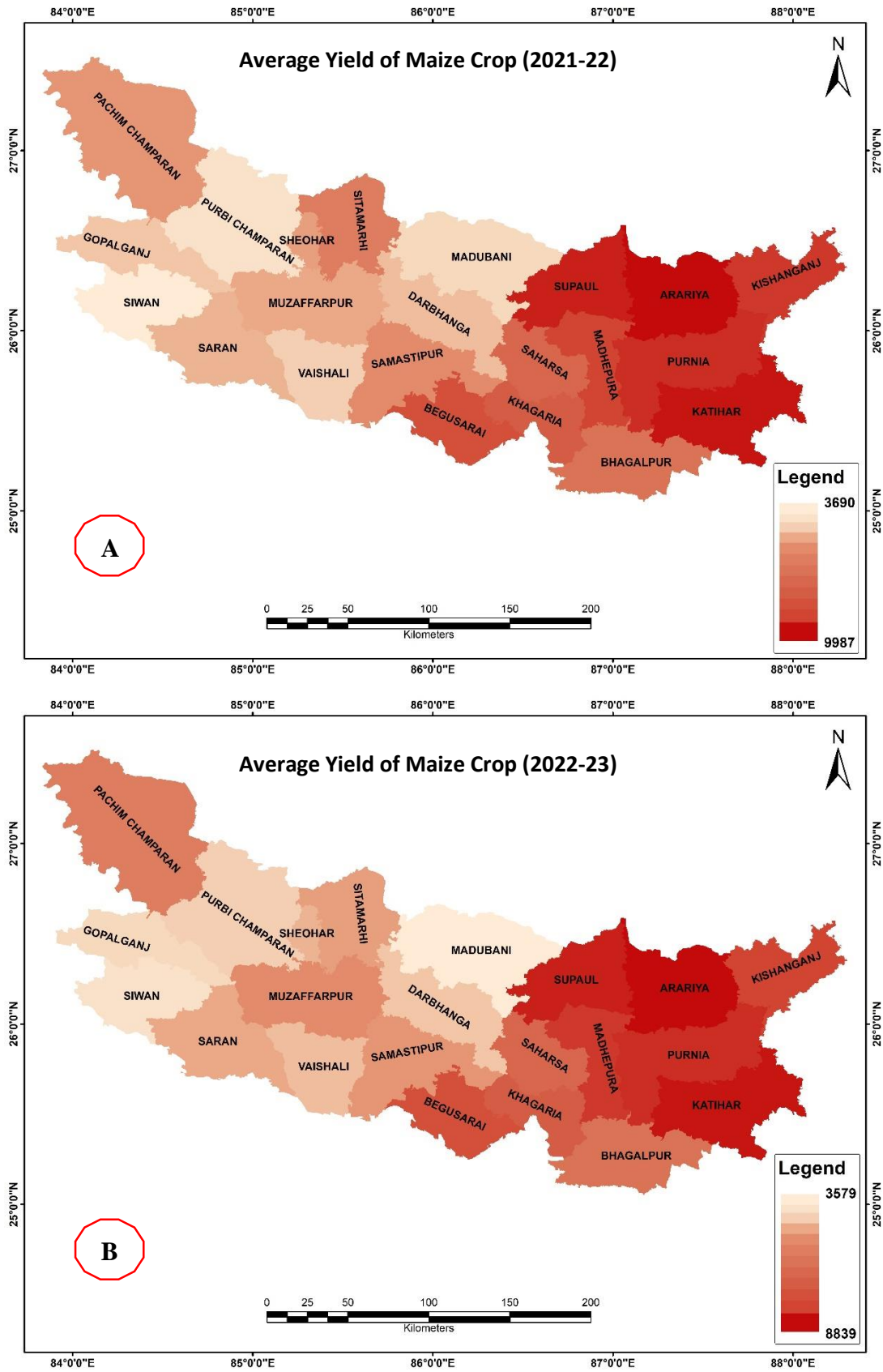


Figure 6.11 Choropleth map of maize crop's yield: (A) 2021-22 & (B) 2022-23

Table 6.1: Districts wise maize crop acreage, yield and production during 2021-22 and 2022-23

Sl. No.	District	2021-22					2022-23				
		Acreage		Yield	Production		Acreage		Yield	Production	
		(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)	(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)
1	Arariya	30784	30.78	9987.74	307458916	307.46	31205	31.21	8839.15	275827266	275.83
2	Begusarai	40052	40.05	7463.73	298938807	298.94	41613	41.61	6881.56	286362442	286.36
3	Bhagalpur	17050	17.05	6508.76	110973577	110.97	17373	17.37	5740.73	99733124	99.73
4	Darbhangha	5425	5.42	4407.44	23908731	23.91	5610	5.61	4019.59	22549437	22.55
5	Gopalganj	3907	3.91	4332.24	16925239	16.93	3974	3.97	3691.07	14670126	14.67
6	Katihar	38141	38.14	9949.47	379480991	379.48	38335	38.33	8805.28	337546300	337.55
7	Khagaria	33754	33.75	7171.20	242055466	242.06	35601	35.60	6540.13	232835379	232.84
8	Kishanganj	1118	1.12	9133.85	10215566	10.22	1180	1.18	7809.44	9211613	9.21
9	Madhepura	17734	17.73	8846.16	156879040	156.88	17868	17.87	8005.77	143048369	143.05
10	Madubani	43	0.04	3968.00	170068	0.17	45	0.05	3579.14	161256	0.16
11	Muzaffarpur	10537	10.54	5383.30	56724801	56.72	10902	10.90	5340.23	58218353	58.22
12	W. Champaran	1811	1.81	5501.88	9966491	9.97	1830	1.83	5402.85	9889631	9.89
13	E. Champaran	4946	4.95	3811.50	18852632	18.85	5139	5.14	3895.35	20016496	20.02

Sl. No.	District	2021-22					2022-23				
		Acreage		Yield	Production		Acreage		Yield	Production	
		(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)	(ha)	('000 ha)	(Kg/ha)	(Kg)	('000 tonnes)
14	Purnia	47939	47.94	9597.42	460086686	460.09	49514	49.51	8109.82	401549894	401.55
15	Saharsa	12425	12.43	6904.32	85786659	85.79	12540	12.54	6089.61	76364715	76.36
16	Samastipur	28005	28.01	5740.58	160765804	160.77	28771	28.77	4890.97	140716613	140.72
17	Saran	5644	5.64	5047.92	28490208	28.49	5734	5.73	4674.37	26802406	26.80
18	Sheohar	1502	1.50	5437.40	8169150	8.17	1530	1.53	4600.04	7037718	7.04
19	Sitamarhi	3144	3.14	5829.30	18325221	18.33	3109	3.11	4698.42	14605795	14.61
20	Siwan	5277	5.28	3690.00	19472204	19.47	5388	5.39	3623.58	19525580	19.53
21	Supaul	4711	4.71	9804.45	46191290	46.19	4703	4.70	8363.19	39336174	39.34
22	Vaishali	7303	7.30	4163.39	30404862	30.40	7540	7.54	4088.45	30828940	30.83
Total (AOI)		321252	321.25	7751.68	2490242408	2490.24	329504	329.50	6879.54	2266837629	2266.84

Note: AOI (Area of Interest), ha (Hectare), '000 (Thousand), Kg (Kilometre), Sl (Serial), No (Number)

6.3.3 Maize crop production estimation

The highest maize production was recorded in Purnia district in both estimation years, which produced about 460 thousand tonnes in 2021-22 and about 401 thousand tonnes in 2022-23. Katihar followed closely with a production of 379 thousand tonnes in 2021-22 and 337 thousand tonnes in 2022-23, while Arariya and Begusarai came in third with 307 thousand tonnes in 2021-22 and 286 thousand tonnes in 2022-23 respectively. The fourth and fifth-highest producing districts were Begusarai in 2021-22 and Khagaria in 2021-22, respectively, producing about 299 and 242 thousand tonnes. In 2022-23 estimation years, Arariya and Khagaria come at fourth and fifth position in production.

Furthermore, we observed that Samastipur district produced 160.77 thousand tonnes of maize, while Madhepura and Bhagalpur produced about 157 and 111 thousand tonnes of maize in 2021-22 respectively. The remaining districts produced maize in smaller quantities, with Madubani producing the lowest at only 170 tonnes in 2021-22. Whereas, Madhubani, Sheohar, Kishanganj and West Champaran districts recorded the lowest production in the year 2022-23 (Fig. 6.11).

6.3.4 Validation of estimated yield with CCE and Government Report

Presently, i have used 25% CCEs data for validation of extracted yield from the study area. In this context, i have plotted the data on a scatter plot to visualize the relationship between actual yield and predicted yield. Wherein, it is observed that there is a general positive correlation between the two variables, but there are some samples that deviate significantly from the general trend. The correlation coefficient between the Actual/CCE Yield and the Predicted Yield is 0.79, indicating a strong positive correlation. Wherein, about 15% difference is observed in estimated yield (Fig. 6.12).

In addition, i have also verified and compared the estimated yield with the reported data from the Government of Bihar. Where we found that a variation of around 18% is observed in the expected/estimated yield. It is expected that the accuracy of estimation can be improved with the use of high-resolution spatial and temporal satellite datasets (Tripathy *et al.*, 2021).

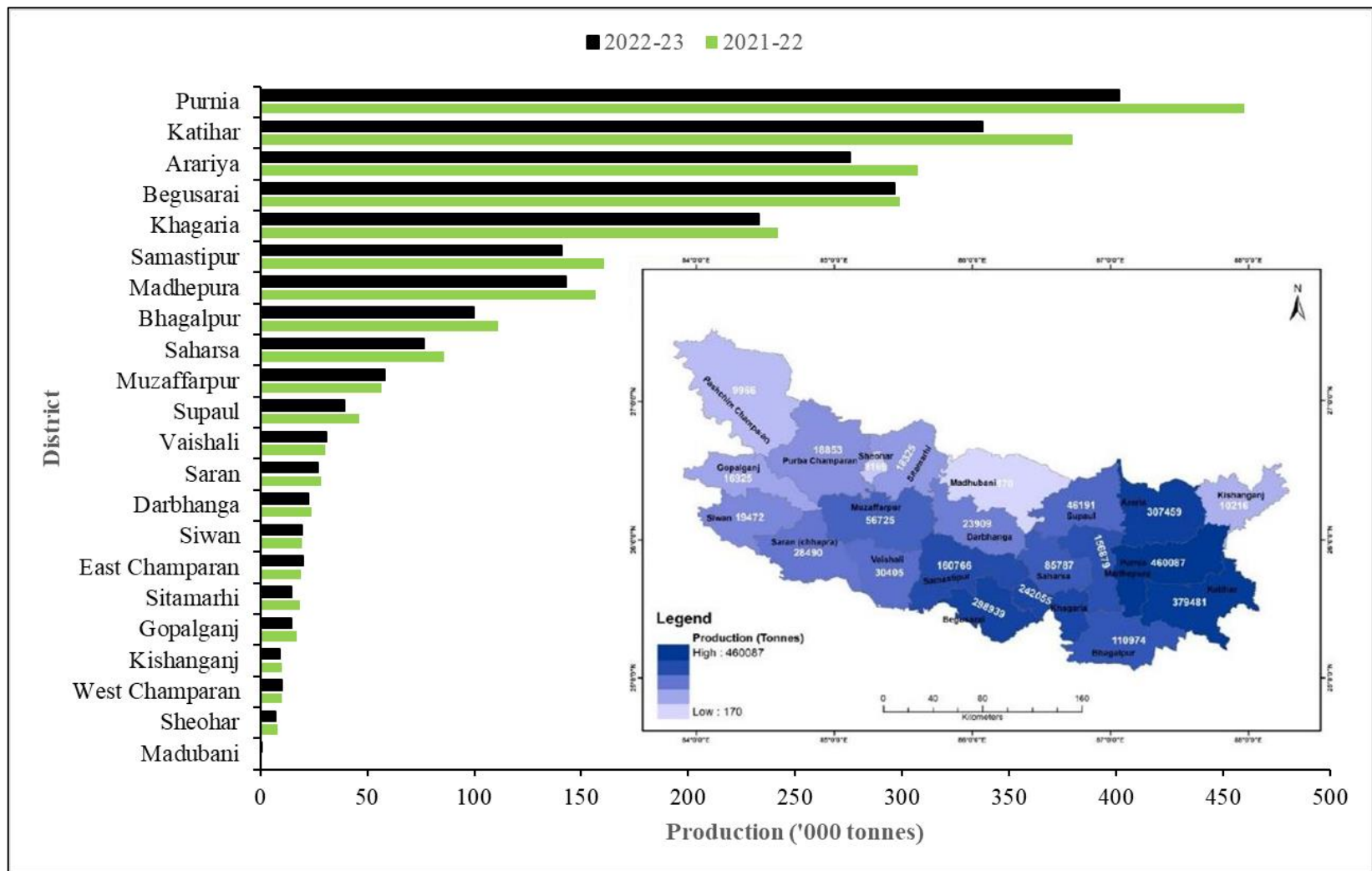


Figure 6.12: District-wise maize crop production

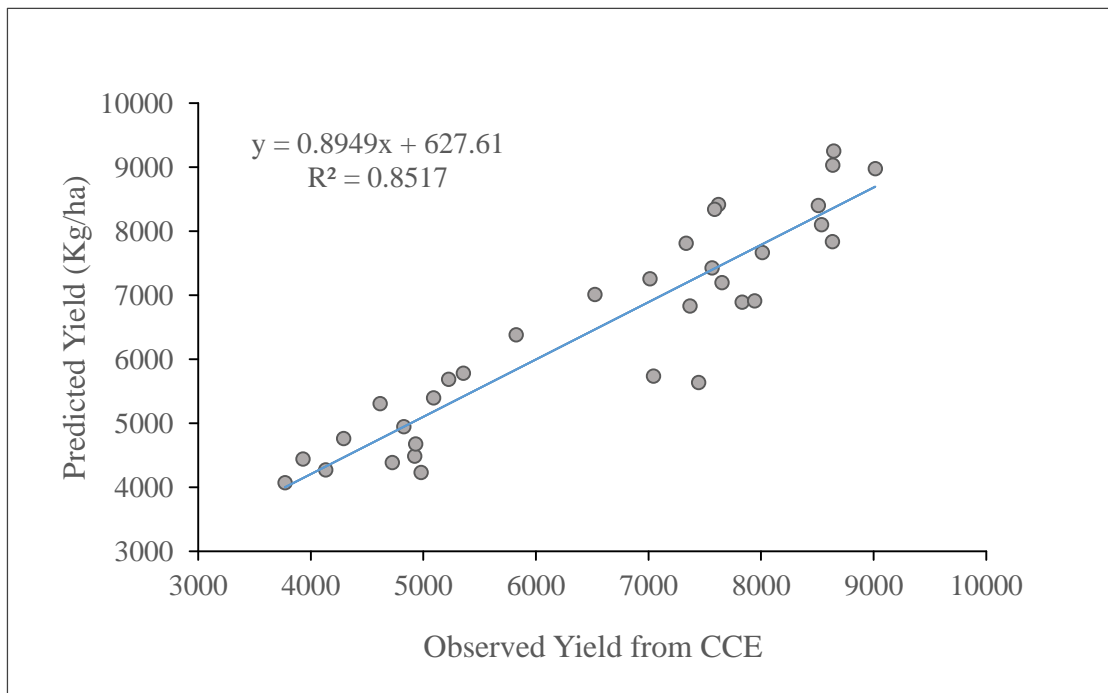


Figure 6.13: Graph showing variation between actual and predicted yield

6.4 Conclusion

The study was carried out for the development of accurate maize crop acreage and yield estimation model using remote sensing satellite imagery, GT and CCE data. For acreage estimation, i have achieved overall 84 percent of accuracy (kappa 0.79) with Sentinel-2A/B datasets, 90 percent (kappa 0.89) with NICFI-PlanetScope datasets, and 95 percent (Kappa 0.91) with the combination of PlanetScope and Sentinel-2A/B datasets.

All five spectral indices intensely evaluated the relationship with maize crop yield using a regression model and demonstrated the methodology for maize crop yield estimation. The obtained result was also validated with CCE data of study area. Wherein, newly developed Red Edge index is performing well ($R^2 = 0.84$) in terms of correlation between vegetation indices and yield than previously developed indices. Therefore, it is expected that demonstrated method will be helpful for precisely crop growth, phenology, stress, yield estimation and prediction.

This approach relies on publicly accessible Sentinel-2, PlanetScope datasets and the GEE Platform for data processing, which minimizes financial constraints and maximizes the replicability of the used approach. However, the accuracy of the methodology can be further validated by using it in different locations and crops with large to small datasets of CCE. All data processing JavaScript GEE code and datasets utilized in the study will be made available as per user request for further use.

In recent years, several studies were conducted to study crop phenology. Wherein, advanced computation and geospatial technologies were used for crop area, yield and production mapping and monitoring. But frequent floods are common phenomena in the study area (North Bihar), which damage to people's lives, infrastructure and agricultural sector every year in Bihar. Despite this, no specific studies have been conducted in North Bihar, India to assess the impact of flood on crop acreage, yield and production.

To address this gap, a systematic approach has been employed. Firstly, the flood extent has been precisely delineated by harnessing Sentinel-1 SAR remote sensing satellite data. The processing of SAR datasets was efficiently executed through the GEE cloud computing platform, which reduce processing time and expedited the identification and mapping of flood-affected regions. This initial step laid the groundwork for comprehending the spatial dimensions of the issue.

Moreover, the estimation of maize crop acreage has been precisely conducted through the amalgamation of various data sources. The combination of PlanetScope and Sentinel-2 remote sensing satellite data has used for a detailed assessment of maize crop. Wherein, GT data of the study area has been collected to train the ML algorithms such as RF, SVM, CART, and Naive Bayes for acreage estimation. These algorithms have ability to understand complex patterns and relationships within the satellite data, which enhanced the precision of maize crop acreage estimation.

Traditional methods of crop yield and production estimation are costly and time consuming. Therefore, modern approaches have been used for maize crop yield and production estimation, using remote sensing satellite data, CCE data, statistical and machine learning approaches.

By combining these multi-source datasets and harnessing the capabilities of advanced machine learning techniques, a comprehensive and meticulous evaluation of flood impact on maize crops has been assessed. It is expected that this study will be helpful for disaster management, agriculture policy, and food security in the region.

In this work, the North Bihar which includes 22 districts; has been chosen as a study area because about 76% of the population (number) is dependent on agriculture in the region that are severely affected by flood every year. This study is intended to contribute to the existing knowledge about the impact of floods on agriculture by providing a detailed assessment of the impact of floods on maize crops acreage, yield and production through integration of state-of-the-art techniques such as remote sensing and machine learning.

7.1 Findings

In Chapter 4, flood extents and flood-affected areas of North Bihar for the year 2020 and 2021 using all accessible Sentinel-1 SAR, Sentinel-2 MSI and PlanetScope images with additional supporting datasets available on the GEE cloud computing platform have been demarcated. In addition, i have developed JavaScript code for processing the massive dataset hosted on the GEE cloud computer platform for robust flood mapping and monitoring at large scale using Microwave (SAR) satellite dataset within a short time of period.

Here, the present study showed that floods severely impacted a large portion of North Bihar during the monsoon season of 2020 and 2021. About 701967 hectares (614706 ha agricultural) land in 2020 and 955897 hectares (851663 ha agricultural) in 2021 have been severely flooded. It is observed that about ~ 12.63% (701967 ha) in 2020 and about ~ 17.20% (955897 ha) of study area were flooded in 2021. In the floods of 2021, about 4% more area of North Bihar were flooded as compared to the floods of 2020.

District-wise flood-affected areas were extracted and it was found that the most flood-affected districts in 2021 were Bhagalpur, Darbhanga, Katihar, Muzaffarpur and

Gopalganj. Results were validated using flood extent layer generated by NRSC and advisory document of State Disaster Management Authority of Bihar.

In Chapter 5, I have proposed a method to discriminate /identify the maize crop from other crops using machine learning algorithms (RF, SVM and CART), publicly available remote sensing satellite datasets and cloud computing platforms such as GEE. In the present study, Integrated satellite imagery of Sentinel-2 and PlanetScope, and Machine learning approaches have been used for maize crop mapping of 2022 and 2023 in North Bihar, India.

In addition, I have assessed and compared the Performance of CART, Random Forest and SVM Algorithms of ML for crop identification and mapping using the GEE Cloud Computing Platform. Wherein, we found that RF outperforms CART and SVM algorithms in the GEE platform with PlanetScope data (90.17 % Overall Accuracy (OA) with Kappa 0.89) and also with the integration of PlanetScope and Sentinel-2A/B data (OA = 95.53%, Kappa 0.91). But, CART outperforms RF and SVM algorithms with Sentinel-2A/B data (OA = 88.59 %, Kappa 0.85). The use of integrated satellite imagery with RF algorithms improved the accuracy of crop mapping by about 6%. Which shows that the integration of Sentinel-2A/B and PlanetScope imagery is more suitable for crop identification with RF algorithm. But, the CART algorithm can be suitable for accurately crop mapping with Sentinel-2 data.

In this study, a maize crop map has been generated using the GEE platform and QGIS software, and their overall accuracy is 95.53%. It is also found that approximately 321252 hectares of land in 2021-22 and 329504 hectares of land in 2022-23 are devoted to maize cultivation in North Bihar, India. I have also developed a web-based JavaScript code that can be tested anywhere in the world for robust mapping of crops under various climatic conditions. We expected that this study will be helpful for crop cultivation management, precision agriculture, crop insurance and for making a decision support system to prioritise the input subsidy for farmers.

In Chapter 6, we intended to develop a method to timely monitor crop growth, estimate acreage, yield and production using the publicly accessible high-resolution remotely sensed satellite data (Sentinel-2 and PlanetScope) and cloud-based geo-computing platforms such as GEE. Traditional methods of crop yield estimation are time-consuming, costly and having low efficiency. But, the remote sensing technology based approach is cost effective as compared to the traditional method. With the development of RS technology, various spectral indices such as NDVI, GNDVI, EVI and LSWI were developed using remotely sensed satellite data for mapping and monitoring vegetation health and productivity.

In the present study, the newly developed 'REI (Red Edge Index)' has been used for the estimation of maize crop yield, which is performing well in terms of prediction accuracy than previously developed indices. Wherein, the obtained correlation with maize yield was $R^2 = 0.84$ with REI, $R^2 = 0.79$ with NDVI, $R^2 = 0.76$, with EVI, $R^2 = 0.70$, with GNDVI and $R^2 = 0.50$ with LSWI. Acreage has also been estimated using high-resolution satellite imagery and a random forest machine learning algorithm as input parameters for the yield model and cluster/district-wise production estimation. It is expected that newly developed crop yield estimation model is simple and cost-effective for developing countries like India and will improve accuracy by using them in different crops and locations with large-scale datasets.

In short, by employing advanced geo-computation approaches and geospatial technologies, the study was conducted to analyze the spatial extent of flood affected areas to assess the impact of flood on maize productivity. Wherein, it was estimated that about 24,90,242 tonnes of maize has been produced in 2021-22 and 22,66,837 tonnes in 2022-23. The total production in 2021-22 was higher by about 9.85% as compared to 2022-23 due to positive post-flood effect on agricultural land (nutritional soil accumulation).

7.2 Recommendation and Future perspectives

In recent years, geospatial technology has gained significant attention due to publicly availability of high-resolution remotely sensed satellite data, the advancements in computational methods and machine learning algorithms. While these approaches have yielded satisfactory results for assessment of flood extent and its impact on crop acreage, yield and production. There is still scope for further refinement and improvement.

1. Despite advancements in technology, data quality and availability still remains a challenge, especially in developing regions. Low resolution geospatial data can hinder the accuracy of analyses.
2. The Scalability of geospatial datasets and technologies remains a concern, especially when working with large datasets. Geospatial applications require large computational storage (high end computer) to handle big data.
3. Machine learning algorithms have robustness in crop mapping and require further development to handle various large and complex datasets. In this context, deep learning approaches can be adaptable and reliable.
4. Many applications require geospatial data analysis on a real-time basis, such as disaster management, forest fire mapping and monitoring. Developing real-time processing capabilities for geospatial data remains a challenge. This can be improved by the development of cloud computing platforms with public accessibility.
5. This developed, edited and tested JavaScript code can be used for processing voluminous datasets. It can be tested at other locations for its robustness for demarcating flood extents.
6. The methodology adopted in the present study for crop mapping can be used to test it on large scale to establish its utility and robustness.

LIST OF PUBLICATIONS

1. **Himanshu Kumar**, Sateesh Kumar Karwariya, Rohan Kumar (2022). Google Earth Engine-Based Identification of Flood Extent and Flood-Affected Paddy Rice Fields Using Sentinel-2 MSI and Sentinel-1 SAR Data in Bihar State, India. *Journal of Indian Society Remote Sensing*, 50, 791–803. Springer Nature
2. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta, Magan Singh and Sateesh Kumar Karwariya (2023). Utilizing Machine Learning Algorithm, Cloud Computing Platform and Remote Sensing satellite data for Impact Assessment of Flood on Agriculture Land. *Current Science*. 125 (8), 886-895.
3. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta and Magan Singh (2022). Impact assessment of flood on agricultural land using cloud based computing platform in Kosi River Basin, North Bihar, India. In International Conference on *River Corridor Research and Management*, 376, 293-305. Singapore: Springer Nature
4. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta and Magan Singh (2023). Google's Cloud Computing Platform based Performance Assessment of Machine Learning Algorithms for Precisely Maize Crop Mapping using Integrated Satellite Data of Sentinel-2A/B and PlanetScope. *Journal of Indian Society Remote Sensing*, 51, 2599–2613. Springer Nature
5. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta, Magan Singh and Sateesh Kumar Karwariya (2023). Rapid Cropland Mapping with Sentinel-2 Satellite data and Machine Learning algorithms on Google Earth Engine: Improving Efficiency and Accuracy. *Indian Journal of Agricultural Research*. **Accepted**
6. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta, Magan Singh and Sateesh Kumar (2023). Model Development for Acreage, Yield and Production Estimation of Maize (Zea mays L) Crop using an evolved Spectral Index and Machine learning algorithm. *Advances in Space Research*. **Under Review**

LIST OF CONFRENCES

1. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta, Magan Singh and Sateesh Kumar Karwariya (2022). *Assessment of flood impact on Agricultural land using cloud-based geo-computing platform in North Bihar, India*. Virtual International Conference on “Challenges to disaster risk reduction and resilient habitat” during April, 5-6, 2022. Organized by Centre for Disaster Management Studies, Shaheed Bhagat Singh College, University of Delhi, New Delhi, India, In Association with National Institute of Disaster Management, MoHA, Govt. of India, International Geographical Union Commission on Hazard and Risk & United nations office for disaster risk reduction.

2. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta and Magan Singh (2022). *Google Earth engine based rapid cropland extent and change mapping using cloud based computing platform in Bihar, India*. International Conference on “Geography and the digital intersections and conflations” from 23-25 May, 2022. Organized by Department of Geography, University of Delhi.

3. **Himanshu Kumar**, Rohan Kumar, Sujay Dutta and Magan Singh (2022). *The impact assessment of flood on agricultural land using cloud based computing platform in Kosi River Basin, North Bihar, India*. 2nd International Conference on River Corridor Research and Management from 30th May-1st June, 2022. Organized by Indian Institute of Technology Guwahati and Indian Institute of Technology Jammu.

BIBLIOGRAPHY

1. Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, 57(1), 1-20.
2. Tobón-Marín, A., & Canon Barriga, J. (2020). Analysis of changes in rivers planforms using google earth engine. *International Journal of Remote Sensing*, 41(22), 8654-8681. DOI: 10.1080/01431161.2020.1792575
3. Ali, A. M., Abouelghar, M., Belal, A. A., Saleh, N., Yones, M., Selim, A. I., & Savin, I. (2022). Crop Yield Prediction Using Multi Sensors Remote Sensing (Review Article). *The Egyptian Journal of Remote Sensing and Space Science*, 25(3), 711-716, ISSN 1110-9823. DOI: 10.1016/j.ejrs.2022.04.006.
4. Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., Moghaddam, S. H. A., Mahdavi, S., Ghahremanloo, M., Parsian, S., Wu, Q., & Brisco, B. (2020). Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5326–5350. DOI: 10.1109/JSTARS.2020.3021052
5. Anonymous (2022a). Retrieved from <https://developers.planet.com/docs/data/visual-basemaps/>, accessed on 24 July 2022).
6. Anonymous (2023a). Retrieved from <http://drdpat.bih.nic.in/PA-Table-04-Bihar.htm>, 2020.
7. Anonymous (2023b). State Disaster Management Department, Bihar. Retrieved from <http://disastermgmt.bih.nic.in/cumulative%20flood%20report%202020/cum05092020.pdf>
8. Anonymous (2023c). Retrieved from <https://farmech.dac.gov.in/FarmerGuide/BI/index1.html>, 2020.
9. Anonymous (2020). Retrieved from <https://farmech.dac.gov.in/FarmerGuide/BI/index1.html>, 2020.
10. Anonymous (2020a). Retrieved from <http://drdpat.bih.nic.in/PA-Table-04-Bihar.htm>, 2020.
11. Anonymous (2020b). State Disaster Management Department, Bihar. Retrieved from

<http://disastermgmt.bih.nic.in/cumulative%20flood%20report%202020/cum05092020.pdf>

12. Arumugam, P., Chemura, A., Schauburger, B., & Gornott, C. (2021a). Remote sensing based yield estimation of rice (*Oryza sativa* L.) using gradient boosted regression in India. *Remote Sensing*, 13(12), 1–18. DOI: 10.3390/rs13122379
13. Arumugam, P., Chemura, A., Schauburger, B., & Gornott, C. (2021b). Remote Sensing Based Yield Estimation of Rice (*Oryza Sativa* L.) Using Gradient Boosted Regression in India. *Remote Sensing*, 13(12). DOI: 10.3390/rs13122379
14. Asner, G. P., Knapp, D. E., Broadbent, E. N., Oliveira, P. J., Keller, M., & Silva, J. N. (2005). Selective logging in the Brazilian Amazon. *Science*, 310(5747), 480–482.
15. Awad, M. M. (2019). Toward precision in crop yield estimation using remote sensing and optimization techniques. *Agriculture*, 9(3), 54. DOI: 10.3390/agriculture9030054
16. Band-wise resolution scale and description, https://developers.google.com/earth-engine/datasets/catalog/projects_planet-nicfi_assets_basemaps_africa
17. Belgiu M., Csillik O., Asner G.P., Kelly M., White J.C., Vaughan R., Wulder M.A. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment*, 204:509–523.
18. Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. DOI: 10.1016/j.isprsjprs.2016.01.011
19. Betbeder, J., Fieuzal, R., & Baup, F. (2016). Assimilation of LAI and dry biomass data from optical and SAR images into an agro-meteorological model to estimate soybean yield. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(6), 2540–2553. DOI: 10.1109/JSTARS.2016.2541169.
20. Bhatt, C. M., & Rao, G. S. (2016). Ganga floods of 2010 in Uttar Pradesh, north India: a perspective analysis using satellite remote sensing data. *Geomatics, Natural Hazards and Risk*, 7(2), 747–763.
21. Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., & Hostert, P. (2022). Mapping of crop types and crop sequences with combined

- time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote sensing of environment*, 269, 112831.
22. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
 23. Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). *Classification And Regression Trees* (1st ed.). Routledge. DOI: 10.1201/9781315139470
 24. Building Materials and Technology Promotion Council (BMTPC): Vulnerability Atlas - 3rd Edition; Peer Group, MoHUA; Map is Based on digitised data of SOI, GOI; Census of India 2011; Flood Atlas (1987), Task Force Report (2004), C.W.C., G.O.I. Houses/Population as per Census 2011. <https://vai.bmtpc.org/map/floodmap/flood-bihar.pdf>
 25. Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2), 121–167. DOI: 10.1023/A:1009715923555
 26. Cao, Z., Cheng, T., Ma, X., Tian, Y., Zhu, Y., Yao, X., ... & Li, X. (2017). A new three-band spectral index for mitigating the saturation in the estimation of leaf area index in wheat. *International Journal of Remote Sensing*, 38(13), 3865-3885. DOI: 10.1080/01431161.2017.1306141
 27. Castillejo-González, I.L., López-Granados, F., García-Ferrer, A., Peñá-Barragán, J.M., Jurado-Expósito, M., Orden, M.S., & González-Audícana, M. (2009). Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. *Computers and Electronics in Agriculture*, 68, 207-215.
 28. Census of India (2011). (https://censusindia.gov.in/2011-prov-results/data_files/bihar/Provisional%20Population%20Totals%202011-Bihar.pdf).
 29. Central Water Commission (2020). Daily Flood Situation Report cum Advisories, Government of India, New Delhi. Retrieved from <http://cwc.gov.in/fmo/dfsra, 2020>.
 30. Central Water Commission (2021). Daily Flood Situation Report cum Advisories, Government of India, New Delhi. Retrieved from <http://cwc.gov.in/fmo/dfsra, 2021>.

31. Centre for Research on the Epidemiology of Disasters (CRED). Disasters in numbers. Brussels: CRED; 2023.
32. Centre, J. R., for the Protection, I., of the Citizen, S., Fritz, S., Craig, M., Bossyns, B., Gallego, J., & Michaelsen, J. (2011). Best practices for crop area estimation with remote sensing : *GEOSS community of practice Ag 0703a*. Publications Office. DOI: 10.2788/31835
33. Chandrasekar, K., Sessa Sai, M. V. R., Roy, P. S., & Dwevedi, R. S. (2010). Land Surface Water Index (LSWI) response to rainfall and NDVI using the MODIS Vegetation Index product. *International Journal of Remote Sensing*, 31(15), 3987-4005. DOI: 10.1080/01431160802575653
34. Chaudhari, K. N. (2019). Final report on Development of Standard Operationalized Procedure using Remote Sensing Technology for Optimization of CCEs /BPSG/PMFBY-02/201819.
35. Chini, M., Hostache, R., Giustarini, L., & Matgen, P. (2017). A hierarchical split-based approach for parametric thresholding of SAR images: Flood inundation as a test case. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 6975-6988. DOI: 10.1109/TGRS.2017.2737664
36. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, 151, 61–69. DOI: 10.1016/j.compag.2018.05.012
37. Costa, J. D. S., Liesenberg, V., Schimalski, M. B., Sousa, R. V. D., Biffi, L. J., Gomes, A. R., ... & Bispo, P. D. C. (2021). Benefits of combining ALOS/PALSAR-2 and Sentinel-2A data in the classification of land cover classes in the Santa Catarina southern Plateau. *Remote Sensing*, 13(2), 229.
38. Craig, M., & Atkinson, D. (2013). A literature review of crop area estimation. Accessed July, 2, 2023. https://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/GS_SAC_2013/Improving_methods_for_crops_estimates/Crop_Area_Estimation_Lit_review.pdf

39. Cravero, A., Pardo, S., Sepúlveda, S., & Muñoz, L. (2022). Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review. *Agronomy*, 12(3), 748. DOI: 10.3390/agronomy12030748
40. Cs. Ferencz , P. Bognár , J. Lichtenberger , D. Hamar , Gy. Tarcsai , G. Timár , G. Molnár , SZ. Pásztor , P. Steinbach , B. Székely , O. E. Ferencz & I. Ferencz-Árkos (2004) Crop yield estimation by satellite remote sensing. *International Journal of Remote Sensing*, 25:20, 4113-4149, DOI: 10.1080/01431160410001698870
41. DACNET (2020). India maize scenario. www.eands.dacnet.nic.in
42. Darji, P., Desai, N., Bhavsar, D., & Pandya, H. (2023). A Review: Applications of Remote Sensing In Agriculture. *International Association of Biologicals and Computational Digest*, 2(1), 108-117.
43. De Luca, G., Silva, J. M., & Modica, G. (2022). Short-term temporal and spatial analysis for post-fire vegetation regrowth characterization and mapping in a Mediterranean ecosystem using optical and SAR image time-series. *Geocarto International*, 37(27), 15428-15462.
44. Deb, D., Mandal, S., Deb, S., Choudhury, A., & Hembram, S. (2021). Crop Production Estimation Using Remote Sensing. *Geospatial Technologies for Crops and Soils*, 229-243.
45. Delegido, J., Verrelst, J., Meza, C. M., Rivera, J. P., Alonso, L., & Moreno, J. (2013). A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *European Journal of Agronomy*, 46, 42-52. doi:10.1016/j.eja.2012.12.001
46. Department of Agriculture, Cooperation and Farmers Welfare (DAC&FW) (2020), (<https://farmech.dac.gov.in/FarmerGuide/BI/1.htm>).
47. Dong, T., Liu, J., Shang, J., Qian, B., Ma, B., Kovacs, J. M., & Shi, Y. (2019). Assessment of red-edge vegetation indices for crop leaf area index estimation. *Remote Sensing of Environment*, 222, 133-143. <https://doi.org/10.1016/j.rse.2018.12.032>
48. Duro, D. C., Franklin, S. E., & Dubé, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for

the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote sensing of environment*, 118, 259-272.

49. Dusabe-Richards, P., Ngoc Ha, T., Robinson, T.P., Choummanivong, M., Conchedda, G., Barendse, J., Linard, C. (2021). Mapping Cropping Systems in the Lao PDR Using Time Series Sentinel-1 and Sentinel-2 Data and Machine Learning Algorithms. *Remote Sensing*, 13(5), 827.
50. Elkhachy, I. (2015). Flash flood hazard mapping using satellite images and GIS tools: a case study of Najran City, Kingdom of Saudi Arabia (KSA). *The Egyptian Journal of Remote Sensing and Space Science*, 18(2), 261-278.
51. European Commission, Joint Research Centre (JRC); Columbi Global Human Settlement Layer (GHSL) a University, Center for International Earth Science Information Network - CIESIN (2015): GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015). European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/jrc-ghsl-ghs_pop_gpw4_globe_r2015a
52. European Space Agency (ESA). User Guides—Sentinel-2—Sentinel Online. Available online: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/overview> (accessed on 1 July 2022).
53. European Space Agency. (2022, March 23). Sentinel-2B. Retrieved May 11, 2023, from https://www.esa.int/Enabling_Support/Space_Engineering_Technology/Sentinel-2B
54. European Space Agency. (2022, March 23). Sentinel-2B. Retrieved May 11, 2023, from https://sentinel.esa.int/documents/247904/349490/s2_sp-1322_2.pdf
55. European Space Agency. (n.d.). Sentinel-2. Retrieved May 11, 2023, from https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2
56. FAO, IFAD, UNICEF, WFP and WHO (2021). The State of Food Security and Nutrition in the World, Transforming food systems for food security, improved nutrition and affordable healthy diets for all. Rome, FAO. DOI: 10.4060/cb4474en
57. FAO, IFAD, UNICEF, WFP and WHO (2022). The State of Food Security and Nutrition in the World, Repurposing food and agricultural policies to make

- healthy diets more affordable. Data are available on FAOSTAT (<https://www.fao.org/faostat/en/#data/FS>)
58. FAO (2017). The future of food and agriculture – Trends and challenges. Rome.
 59. Farmer’s portal, https://farmer.gov.in/m_cropstaticsmaize.aspx, accessed dated 21.06.2022.
 60. Farr, T.G., E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, P. Rosen, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner (2007), The Shuttle Radar Topography Mission. *Reviews of Geophysics*, volume 45, RG2004, DOI:10.1029/2005RG000183.
 61. Fatholouloumi, S., Karimi Firozjaei, M., & Biswas, A. (2022). An Innovative Fusion-Based Scenario for Improving Land Crop Mapping Accuracy. *Sensors*, 22(19), 7428. DOI:10.3390/s22197428
 62. Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, 140, 23-35, DOI: 10.1016/j.rse.2013.08.029
 63. Filippi, P., Jones, E. J., Wimalathunge, N. S., Somarathna, P. D., Pozza, L. E., Ugbaje, S. U., ... & Bishop, T. F. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture*, 20, 1015-1029. DOI: 10.1007/s11119-018-09628-4
 64. Flood Hazard Atlas of Bihar- A Geospatial Approach (2019). Prepared by National Remote Sensing Centre, Indian Space Research Organisation, Dept. of Space, Govt. of India, In Association with National Disaster Management Authority, Ministry of Home Affairs, Govt. of India, “Disaster Management Department, Govt. of Bihar, Patna” & Bihar State Disaster Management Authority, Govt. of Bihar, Patna. <https://bhuvan.nrsc.gov.in/pdf/Flood-Hazard-Atlas-Bihar.pdf>
 65. Flood Management Information System, Water Resources Department, Government of Bihar, https://www.fmiscwrdbihar.gov.in/Load_FMIS_Site/indexold.htm. Accessed on dated 2020.
 66. Flood prone areas map of India. Retrieved from NITI Aayog (2021)/ <https://ndma.gov.in/Natural-Hazards/Floods>

67. Food and Agriculture Organization (FAO) (2021). The impact of disasters and crises on agriculture and food security, DOI: 10.4060/cb3673en
68. Franch, B., Vermote, E. F., Skakun, S., Roger, J. C., Becker-Reshef, I., Murphy, E., & Justice, C. (2019). Remote sensing based yield monitoring: Application to winter wheat in United States and Ukraine. *International Journal of Applied Earth Observation and Geoinformation*, 76, 112-127. DOI: 10.1016/j.jag.2018.11.012
69. Freer, J., Beven, K., Neal, J., Schumann, G., Hall, J., & Bates, P. (2013). Flood risk and uncertainty. In J. Rougier, S. Sparks, & L. Hill (Eds.), *Risk and Uncertainty Assessment for Natural Hazards* (pp. 190-233). Cambridge: Cambridge University Press. DOI:10.1017/CBO9781139047562.008
70. Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O.M., Shelestov, A., & Kussul, O. (2014). Efficiency assessment of using satellite data for crop area estimation in Ukraine. *Int. J. Appl. Earth Obs. Geoinformation*, 29, 22-30. DOI:10.1016/j.jag.2013.12.013
71. Gao, F., & Zhang, X. (2021). Mapping Crop Phenology in Near Real-Time Using Satellite Remote Sensing: Challenges and Opportunities. *Journal of remote sensing*, 2021, 1-14. DOI:10.34133/2021/8379391
72. Gao, Z., & Liu, X. (2014). "Support vector machine and object-oriented classification for urban impervious surface extraction from satellite imagery," 2014, *The Third International Conference on Agro-Geoinformatics*, Beijing, China, 2014, pp. 1-5, DOI: 10.1109/Agro-Geoinformatics.2014.6910661.
73. Ghosh, S., Kumar, D. and Kumari, R. (2022). Evaluating the impact of flood inundation with the cloud computing platform over vegetation cover of Ganga Basin during COVID-19. *Spat. Inf. Res*, 30, 291–308. DOI: 10.1007/s41324-022-00430-z
74. Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58(3), 289–298. DOI:10.1016/S0034-4257(96)00072-7
75. Goldstein, A., Fink, L., Meitin, A., Bohadana, S., Lutenberg, O., & Ravid, G. (2018). Applying machine learning on sensor data for irrigation

- recommendations: revealing the agronomist's tacit knowledge. *Precision agriculture*, 19, 421-444. DOI: 10.1007/s11119-017-9527-4
76. Gomez, C., White, J.C., & Wulder, M.A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, 55-72.
 77. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202, 18-27. DOI: 10.1016/j.rse.2017.06.031.
 78. Griffiths, P., Nendel, C., & Hostert, P. (2019). Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sensing of Environment*, 220, 135–151. DOI: 10.1016/j.rse.2018.10.031
 79. Groeve, T.D. (2010). Flood monitoring and mapping using passive microwave remote sensing in Namibia. *Geomatics, Natural Hazards and Risk*, 1, 19-35. DOI: 10.1080/19475701003648085
 80. Grohman, G., G. Kroenung, and J. Strebeck (2006). Filling SRTM voids: The Delta Surface Fill method. *Photogrammetric Engineering and Remote Sensing*, v. 72, no. 3, p. 213-216.
 81. Gulati, A., Terway, P., & Hussain, S. (2018). Crop insurance in India: Key issues and way forward. ICRIER, Working Paper 352,
 82. Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., ... & Yang, X. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant methods*, 15(1), 1-19.
 83. Hanqiu Xu (2006). Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery, *International Journal of Remote Sensing*, 27:14, 3025-3033, DOI: 10.1080/01431160600589179
 84. Htitiou, A., Boudhar, A., Lebrini, Y., Hadria, R., Lionboui, H., & Benabdelouahab, T. (2022). A comparative analysis of different phenological information retrieved from Sentinel-2 time series images to improve crop classification: a machine learning approach. *Geocarto International*, 37(5), 1426–1449. DOI:10.1080/10106049.2020.1768593

85. Hudait, M., & Patel, P. P. (2022). Crop-type mapping and acreage estimation in smallholding plots using Sentinel-2 images and machine learning algorithms: Some comparisons. *The Egyptian Journal of Remote Sensing and Space Science*, 25(1), 147–156. DOI:10.1016/j.ejrs.2022.01.004
86. Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, 83(1-2), 195-213.
87. Indian Metrological Department (IMD) report and advisory, Bihar (2020)
88. India-WRIS. Retrieved from (<https://indiawris.gov.in/wris/#/rainfall>), 2020.
89. Inglada, J., M. Arias, B. Tardy, O. Hagolle, S. Valero, D. Morin & G, D. (2015). Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery, *Remote Sensing*, 7 (9): 12356–12379. DOI:10.3390/rs70912356.
90. Iniyan, S., Jebakumar, R. (2022). Mutual Information Feature Selection (MIFS), Based Crop Yield Prediction on Corn and Soybean Crops Using Multilayer Stacked Ensemble Regression (MSER). *Wireless Pers Commun*, 126, 1935–1964. DOI:10.1007/s11277-021-08712-9
91. Jain, M., Srivastava, A. K., Balwinder-Singh, Joon, R. K., McDonald, A., Royal, K., Lisaius, M. C., & Lobell, D. B. (2016). Mapping smallholder wheat yields and sowing dates using micro-satellite data. *Remote Sensing*, 8(10), 1–18. DOI: 10.3390/rs8100860
92. Jain, S.K., & Singh, V.P. (2022). Strategies for flood risk reduction in India. *ISH Journal of Hydraulic Engineering*, 29:2, 165-174, DOI: 10.1080/09715010.2021.2019136
93. Jamal, M. R., Kristiansen, P., Kabir, M. J., & Lobry de Bruyn, L. (2023). Challenges and Adaptations for Resilient Rice Production under Changing Environments in Bangladesh. *Land*, 12(6), 1217.
94. Jensen, R.J. (2013). *Remote Sensing of the Environment: An Earth Resource Perspective*. Prentice-Hall, Inc., 2nd Edition. pp.1-10.
95. Jiang Z, Chen Z, Chen J, Ren J, Li Z, Sun L. (2014) The Estimation of Regional Crop Yield Using Ensemble-Based Four-Dimensional Variational Data Assimilation. *Remote Sensing*; 6(4):2664-2681. DOI: 10.3390/rs6042664

96. Jin, Z., Azzari, G., You, C., Di Tommaso, S., Aston, S., Burke, M., & Lobell, D. B. (2019). Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote Sensing of Environment*, 228, 115-128.
97. Keys, R. (1981). Cubic convolution interpolation for digital image processing. *IEEE transactions on acoustics, speech, and signal processing*, 29(6), 1153-1160.
98. Khaki, S., Pham, H., & Wang, L. (2021). Simultaneous corn and soybean yield prediction from remote sensing data using deep transfer learning. *Scientific Reports*, 11(1), 11132. DOI: 10.1038/s41598-021-89779-z
99. Kim, W.S., Iizumi, T., Hosokawa, N., Tanoue, M., & Hirabayashi, Y. (2023). Flood impacts on global crop production: advances and limitations. *Environmental Research Letters*, 18.
100. Kleinewillinghöfer, L., Olofsson, P., Pebesma, E., Meyer, H., Buck, O., Haub, C., & Eiselt, B. (2022). Unbiased Area Estimation Using Copernicus High Resolution Layers and Reference Data. *Remote Sensing*, 14(19). DOI: 10.3390/rs14194903
101. Kordelas, G., Manakos, I., Aragonés, D.G., Díaz-Delgado, R., and Bustamante, J. (2018) Fast and Automatic Data-Driven Thresholding for Inundation Mapping with Sentinel-2 Data. *Remote Sensing*, 10, 910.
102. Kotaridis, I., & Lazaridou, M. (2021). Remote sensing image segmentation advances: A meta-analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 173, 309–322. DOI:10.1016/j.isprsjprs
103. Kumar, H., Karwariya, S. K., & Kumar, R. (2022). Google Earth Engine-Based Identification of Flood Extent and Flood-Affected Paddy Rice Fields Using Sentinel-2 MSI and Sentinel-1 SAR Data in Bihar State, India. *Journal of the Indian Society of Remote Sensing*, 50(5), 791–803. DOI: 10.1007/s12524-021-01487-3
104. Kumar, H., Kumar, R., Dutta, S., & Singh, M. (2023). Google's Cloud Computing Platform-Based Performance Assessment of Machine Learning Algorithms for Precisely Maize Crop Mapping Using Integrated Satellite Data of Sentinel-2A/B and PlanetScope. *Journal of the Indian Society of Remote Sensing*. DOI: 10.1007/s12524-023-01764-3

105. Kumar, H., Singh, M., Dutta, S., Karwariya, S. K., & Kumar, S. (2022). Fodder crop estimation using Sentinel-2A/B satellite data for West Bengal, India. *Indian Journal of Agricultural Sciences*, 92(6), 716-20.
106. Lary, D. J., A. H. Alavi, A. H. Gandomi, and A. L. Walker. (2016). Machine Learning in Geosciences and Remote Sensing. *Geoscience Frontiers*, 7 (1): 3–10. doi:10.1016/j.gsf.2015.07.003.
107. Latwal, A., Saxena, S., Dubey, S. K., Choudhary, K., Sehgal, S., Neetu, & Ray, S. S. (2019). Evaluation of Pre-Harvest Production Forecasting of Mustard Crop in Major Producing States of India, Under Fasal Project. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-3/W6, 115–122. DOI: 10.5194/isprs-archives-XLII-3-W6-115-2019
108. Liang, J., & Liu, D. (2020). A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159, 53-62. DOI: 10.1016/J.ISPRSJPRS.2019.10.017
109. Lin, C., Zhong, L., Song, X. P., Dong, J., Lobell, D. B., & Jin, Z. (2022). Early- and in-season crop type mapping without current-year ground truth: Generating labels from historical information via a topology-based approach. *Remote Sensing of Environment*, 274, 112994. DOI: 10.1016/j.rse.2022.112994
110. Lowder, S. K., Skoet, J., & Raney, T. (2016). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development*, 87, 16–29. DOI: 10.1016/j.worlddev.2015.10.041
111. Luo, H., Li, M., Dai, S., Li, H., Li, Y., Hu, Y., Zheng, Q., Yu, X., & Fang, J. (2022). Combinations of Feature Selection and Machine Learning Algorithms for Object-Oriented Betel Palms and Mango Plantations Classification Based on Gaofen-2 Imagery. *Remote Sensing*, 14(7). DOI: 10.3390/rs14071757
112. Macdonald, R. B. (1984). A summary of the history of the development of automated remote sensing for agricultural applications. *IEEE Transactions on Geoscience and Remote Sensing*, (6), 473-482. DOI: 10.1109/TGRS.1984.6499157

113. Mahalanobis National Crop Forecast Centre (MNCFC) (2020). Gram Panchayat Level yield estimation using Technology under PMFBY: A summary of the Pilot Studies conducted during Kharif (2019-20).
114. Maize Vision (2022): FICCI PwC knowledge Report, 5th edition of India's Maize Summit.
115. Maize, agro-climatic conditions information of maize, retrieved from <https://iimr.icar.gov.in/>
116. Mandal, B., Sreenivas, K., Jain, R., (2021). Mapping irrigated rice fields using Sentinel-2 and Google Earth Engine: A Random Forest model approach. Geocarto International, 1-16.
117. Mather, P., & Tso, B. (2016). Classification methods for remotely sensed data. CRC press.
118. McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, 17(7), 1425-1432, DOI: 10.1080/01431169608948714
119. Mehdaoui, R., & Anane, M. (2020). Exploitation of the red-edge bands of Sentinel 2 to improve the estimation of durum wheat yield in Grombalia region (Northeastern Tunisia). *International Journal of Remote Sensing*, 41(23), 8986–9008. DOI: 10.1080/01431161.2020.1797217
120. Meng, S., Wang, X., Hu, X., Luo, C., & Zhong, Y. (2021). Deep learning-based crop mapping in the cloudy season using one-shot hyperspectral satellite imagery. *Comput. Electron. Agric.*, 186, 106188. DOI: 10.1016/j.compag.2021.106188
121. Mizuochi, H., Iijima, Y., Nagano, H., Kotani, A., & Hiyama, T. (2021). Dynamic Mapping of Subarctic Surface Water by Fusion of Microwave and Optical Satellite Data Using Conditional Adversarial Networks. *Remote Sensing*, 13(2). DOI: 10.3390/rs13020175
122. Moharrami, M., Javanbakht, M. and Attarchi, S. (2021). Automatic flood detection using sentinel-1 images on the google earth engine. *Environmental Monitoring and Assessment*, 193, 248. DOI: 10.1007/s10661-021-09037-7
123. Mondal, S., Jeganathan, C., Sinha, N. K., Rajan, H., Roy, T., & Kumar, P. (2014). Extracting seasonal cropping patterns using multi-temporal vegetation indices from IRS LISS-III data in Muzaffarpur District of Bihar, India. *Egyptian Journal*

- of Remote Sensing and Space Science*, 17(2), 123–134. DOI: 10.1016/j.ejrs.2014.09.002
124. Moumni, A., & Lahrouni, A. (2021). Machine Learning-Based Classification for Crop-Type Mapping Using the Fusion of High-Resolution Satellite Imagery in a Semiarid Area. *Scientifica*, 2021. DOI:10.1155/2021/8810279
 125. Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259. DOI: 10.1016/j.isprsjprs.2010.11.001
 126. NASA Jet Propulsion Laboratory, "Shuttle Radar Topography Mission," last updated December 18, 2019, <https://www.jpl.nasa.gov/spaceimages/details.php?id=PIA03361>.
 127. National Crime Records Bureau (NCRB) report 2021, Ministry of Home Affairs, Government of India. Accidental deaths and suicides in India.
 128. Neetu, & Ray, S. S. (2019). Exploring machine learning classification algorithms for crop classification using Sentinel 2 data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 573–578. DOI: 10.5194/isprs-archives-XLII-3-W6-573-2019
 129. NICFI Tropical Forest Basemaps Now Available in Google Earth Engine. Available online: <https://www.planet.com/pulse/nicfi-tropical-forest-basemaps-now-available-in-google-earth-engine/> (accessed on 24 July 2022).
 130. Niftiyev, I., & Ibadoghlu, G. (2023). Longitudinal Principal Component and Cluster Analysis of Azerbaijan's Agricultural Productivity in Crop Commodities. *Commodities*, 2(2), 147-167.
 131. NITI Aayog (2021): Report of the Committee constituted for formulation of strategy for Flood Management Works in entire country and River Management Activities and works related to Border Areas (2021– 26).
 132. NRSC (2019) - Cumulative Flood Inundated areas of Bihar State (9 to 23 July 2020), dated 24.07.2020 DSC/NDEM Map no :2020/22, NRSC/ISRO, Hyderabad.
 133. NRSC (2020) - Cropped area affected due to flooding in Bihar state (based on flood layer from Jul 3 to 7 Aug 2020), dated 19.08.2020 Map no: 2020/92, NRSC/ISRO, Hyderabad

134. Oliphant, A. J., Thenkabail, P. S., Teluguntla, P., Xiong, J., Gumma, M. K., Congalton, R. G., & Yadav, K. (2019). Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30-m data using a random forest classifier on the Google Earth Engine Cloud. *International Journal of Applied Earth Observation and Geoinformation*, 81, 110–124. DOI: 10.1016/j.jag.2018.11.014
135. Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., and Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. 2014, 148, 42–57
136. Otsu, N., (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on systems, man, and cybernetics*, 1979, 9(1), 62–66.
137. Paliwal, A., & Jain, M. (2020). The Accuracy of Self-Reported Crop Yield Estimates and Their Ability to Train Remote Sensing Algorithms. *Frontiers in Sustainable Food Systems*, 4(March), 1–10. DOI: 10.3389/fsufs.2020.00025
138. Paliwal, A., & Jain, M. (2020). The accuracy of self-reported crop yield estimates and their ability to train remote sensing algorithms. *Frontiers in Sustainable Food Systems*, 4, 1-10. DOI: 10.3389/fsufs.2020.00025
139. Pan, L., Xia, H., Yang, J., Niu, W., Wang, R., Song, H., ... & Qin, Y. (2021). Mapping cropping intensity in Huaihe basin using phenology algorithm, all Sentinel-2 and Landsat images in Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation*, 102, 102376. DOI: 10.1016/j.jag.2021.102376
140. Pande, C. B. (2022). Land use/land cover and change detection mapping in Rahuri watershed area (MS), India using the google earth engine and machine learning approach. *Geocarto International*, 0(0), 1–21. DOI: 10.1080/10106049.2022.2086622
141. Panjala, P., Gumma, M.K., Teluguntla, P. (2022). Machine Learning Approaches and Sentinel-2 Data in Crop Type Mapping. In: Reddy, G.P.O., Raval, M.S., Adinarayana, J., Chaudhary, S. (eds) *Data Science in Agriculture and Natural Resource Management. Studies in Big Data*, vol 96. Springer, Singapore. DOI:10.1007/978-981-16-5847-1_8

142. Peng, D., Huete, A. R., Huang, J., Wang, F., & Sun, H. (2011). Detection and estimation of mixed paddy rice cropping patterns with MODIS data. *International Journal of Applied Earth Observation and Geoinformation*, 13(1), 13–23. <https://doi.org/https://doi.org/10.1016/j.jag.2010.06.001>
143. Pengra, B. W., Stehman, S. V., Horton, J. A., Dockter, D. J., Schroeder, T. A., Yang, Z., ... & Loveland, T. R. (2020). Quality control and assessment of interpreter consistency of annual land cover reference data in an operational national monitoring program. *Remote Sensing of Environment*, 238, 111261.
144. Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. <https://api.planet.com>
145. Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. <https://api.planet.com>
146. Planet. (n.d.). Planet and Norway's International Climate and Forest Initiative (NICFI). Retrieved May 11, 2023, from <https://www.planet.com/solutions/international-climate-and-forest-initiative/>
147. Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied earth observation and geoinformation*, 8(1), 26-33.
148. Praticò, S., Solano, F., Di Fazio, S., & Modica, G. (2021). Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. *Remote Sensing*, 13(4). DOI:10.3390/rs13040586
149. Prins, A. J., & Van Niekerk, A. (2021). Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms. *Geo-Spatial Information Science*, 24(2), 215-227, DOI: 10.1080/10095020.2020.1782776
150. Wang, Q., Blackburn, G.A., Onojeghuo, A.O., Dash, J., Zhou, L., Zhang, Y., & Atkinson, P.M. (2017). Fusion of Landsat 8 OLI and Sentinel-2 MSI Data. *IEEE Transactions on Geoscience and Remote Sensing*, 55, 3885-3899. doi: 10.1109/TGRS.2017.2683444.
151. Rai, R. (2019). Pradhan Mantri Fasal Bima Yojana: An Assessment of India's Crop Insurance Scheme. ORF issue brief, 16, 296.

152. Ranum, P.M., Peña-Rosas, J.P., & Garcia-Casal, M.N. (2014). Global maize production, utilization, and consumption. *Annals of the New York Academy of Sciences*, 1312. DOI: 10.1111/nyas.12396
153. Rao, P., Zhou, W., Bhattarai, N., Srivastava, A. K., Singh, B., Poonia, S., Lobell, D. B., & Jain, M. (2021). Using sentinel-1, sentinel-2, and planet imagery to map crop type of smallholder farms. *Remote Sensing*, 13(10), 1–13. DOI: 10.3390/rs13101870
154. Ray, S. & Neetu (2017). Crop area estimation with Remote Sensing. In: J. Delincé (ed.), *Handbook on Remote Sensing for Agricultural Statistics (Chapter 5)*. *Handbook of the Global Strategy to improve Agricultural and Rural Statistics (GSARS)*: Rome.
155. Ray, S.S. (2021). How India is using space technology to ‘smart sample’ crops. *Nature India*. DOI: 10.1038/nindia.2021.35
156. Rembold, F., Atzberger, C., Savin, I., & Rojas, O. (2013). Using low resolution satellite imagery for yield prediction and yield anomaly detection. *Remote Sensing*, 5(4), 1704-1733. DOI: 10.3390/rs5041704
157. Ren J, Chen Z, Tang H, Zhou Q, Qin J. (2011). Regional crop yield simulation based on crop growth model and remote sensing data. *Transactions of the Chinese Society of Agricultural Engineering*, 27, 257–264.
158. Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. In *Third Earth Resources Technology Satellite-1 Symposium*, Vol. 1 (pp. 309-317).
159. Rukhovich, D. I., Koroleva, P. V., Rukhovich, D. D., & Kalinina, N. V. (2021). The use of deep machine learning for the automated selection of remote sensing data for the determination of areas of arable land degradation processes distribution. *Remote Sensing*, 13(1), 155.
160. Saad El Imanni, H., El Harti, A., & El Iysaouy, L. (2022). Wheat Yield Estimation Using Remote Sensing Indices Derived from Sentinel-2 Time Series and Google Earth Engine in a Highly Fragmented and Heterogeneous Agricultural Region. *Agronomy*, 12(11), 2853. DOI: 10.3390/agronomy12112853

161. Sadiq, R., Akhtar, Z., Imran, M., & Ofli, F. (2022). Integrating remote sensing and social sensing for flood mapping. *Remote Sensing Applications: Society and Environment*, 25, 100697. DOI: 10.1016/j.rsase.2022.100697
162. Schumann, G. J., Brakenridge, G. R., Kettner, A. J., Kashif, R., & Niebuhr, E. (2018). Assisting flood disaster response with earth observation data and products: A critical assessment. *Remote Sensing*, 10(8), 1230. DOI: 10.3390/rs10081230
163. SciHub; <https://scihub.copernicus.eu>
164. Sharples, J.A. (1973). The Corn Blight Watch Experiment: Economic Implications for Use of Remote Sensing for Collecting Data on Major Crops. LARS Technical Report, Paper 121. Purdue University Publication: West Lafayette, IN, USA.
165. Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017a). Exploring Google earth engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science*, 5(February), 1–10. DOI: 10.3389/feart.2017.00017
166. Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017b). Large scale crop classification using Google earth engine platform. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017-July, 3696–3699. DOI: 10.1109/IGARSS.2017.8127801
167. Shuttle Radar Topography Mission 1 Arc-Second Global. <https://doi.org/10.5066/F7PR7TFT>
168. Singh S.K., Chandel V., Kumar H. and Gupta H. (2014). RS & GIS based urban land use change and site suitability analysis for future urban expansion of Parwanoo planning area, Solan, Himachal Pradesh (India). *International Journal of Development Research*, 4 (8): 1491-1503. Online available at: <http://www.journalijdr.com/sites/default/files/1933.pdf>
169. Singh, M., Dutta, S., Kala, S., Dwivedi, S., Meena, R. K., Meena, V. K., & Onte, S. (2020). Fodder crops assessment using multi-temporal Landsat-8 data by NDVI based classification in Haryana state of India. *Range Management and Agroforestry*, 41(1), 67-73.

170. Singh, S. B., Kasana, R. K., & Singh, S. P. (2018). Status of Corn Cultivation in Bihar: opportunities and future Challenges. https://krishi.icar.gov.in/jspui/bitstream/123456789/21351/2/Corn_Status_in_Bihar_SBSINGH.pdf
171. Singha, M., Dong, J., Sarmah, S., You, N., Zhou, Y., Zhang, G., Doughty, R., & Xiao, X. (2020). Identifying floods and flood-affected paddy rice fields in Bangladesh based on Sentinel-1 imagery and Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166, 278–293. DOI: 10.1016/j.isprsjprs.2020.06.011
172. Slater, J. A., G. Garvey, C. Johnston, J. Haase, B. Heady, G. Kroenung, and J. Little (2006). The SRTM data ‘finishing’ process and products, *Photogramm. Eng. Remote Sens.*, 72, 237–247.
173. Spera, S. A., Galford, G. L., Coe, M. T., Macedo, M. N., & Mustard, J. F. (2016). Land-use change affects water recycling in Brazil's last agricultural frontier. *Global change biology*, 22(10), 3405-3413. DOI: 10.1111/gcb.13298
174. State Disaster Management Department, Bihar. Retrieved from <http://disastermgmt.bih.nic.in/cumulative%20flood%20report%202020/cum05092020.pdf>
175. Stehman, S. V, & Foody, G. M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, 231, 111199. DOI: 10.1016/j.rse.2019.05.018
176. Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1), 77–89. DOI: 10.1016/S0034-4257(97)00083-7
177. Sustainable Development Goal (SDG), NO Poverty (Goal 1) and Zero Hunger (Goal 2). Retrieved from <https://sdgs.un.org>
178. Tanguy, M., Chokmani, K., Bernier, M., Poulin, J., & Raymond, S. (2017). River flood mapping in urban areas combining Radarsat-2 data and flood return period data. *Remote Sensing of Environment*, 198, 442-459.
179. Tavares, P. A., Beltrão, N. E. S., Guimarães, U. S., & Teodoro, A. C. (2019). Integration of sentinel-1 and sentinel-2 for classification and LULC mapping in the urban area of Belém, eastern Brazilian Amazon. *Sensors*, 19(5), 1140.

180. The Food and Agriculture Organization Corporate Statistical Database (FAOSTAT), accessed on 10/May/2023.
181. Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., ... & Traver, I. N. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, 120, 9-24, DOI:10.1016/j.rse.2011.05.028
182. Tripathy, R., Chaudhari, K.N., Bairagi, G.D., Pal, O., Das, R., & Bhattacharya, B.K. (2021). Towards Fine-Scale Yield Prediction of Three Major Crops of India Using Data from Multiple Satellite. *Journal of the Indian Society of Remote Sensing*, 50, 271 - 284. DOI: 10.1007/s12524-021-01361-2
183. Tufail, R., Ahmad, A., Javed, M.A., & Ahmad, S.R. (2021). A machine learning approach for accurate crop type mapping using combined SAR and optical time series data. *Advances in Space Research*. DOI: 10.1016/j.asr.2021.09.019
184. United Nations Population Fund Annual (UNFPA) Report 2012, Page- 1-52
185. United States Agency for International Development (USAID), Micronutrient Programs and DSM Nutritional Products. 2013. Fortification basics, Maize Flour/Meal.
186. Vallentin, C., Harfenmeister, K., Itzerott, S., Kleinschmit, B., Conrad, C., & Spengler, D. (2022). Suitability of satellite remote sensing data for yield estimation in northeast Germany. *Precision Agriculture*, 23(1), 52-82.
187. Veerabhadraswamy, N., Devagiri, G. M., & Khaple, A. K. (2021). Fusion of complementary information of SAR and optical data for forest cover mapping using random forest algorithm. *Curr. Sci*, 120, 193-199.
188. Viskovic, L., Kosovic, I. N., & Mastelic, T. (2019). Crop Classification using Multi-spectral and Multitemporal Satellite Imagery with Machine Learning. *International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1–5. DOI: 10.23919/SOFTCOM.2019.8903738
189. Vizzari, M. (2022). PlanetScope, Sentinel-2, and Sentinel-1 Data Integration for Object-Based Land Cover Classification in Google Earth Engine. *Remote Sensing*, 14(11). DOI: 10.3390/rs14112628
190. Wang, Q., Wang, J., Zhang, H., Li, M., Wang, Q., Zhang, H., Li, M. (2019). Predicting corn and soybean yield response to drought with satellite remote sensing data. *Remote Sensing*, 11(2), 129.

191. Wang, X., Fang, S., Yang, Y., Du, J., & Wu, H. (2023). A New Method for Crop Type Mapping at the Regional Scale Using Multi-Source and Multi-Temporal Sentinel Imagery. *Remote. Sens.*, 15, 2466.
192. Wang, X., Huang, J., Feng, Q., & Yin, D. (2020). Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. *Remote Sensing*, 12(11), 1744. DOI: 10.3390/rs12111744
193. Wei, M., Qiao, B., Zhao, J., & Zuo, X. (2018). Application of remote sensing technology in crop estimation. In 2018 IEEE 4th International Conference on Big Data Security on Cloud (Big Data Security). IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS) (pp. 252-257). IEEE.
194. Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*, 236, 111402.
195. Whiteside, T.G., & Ahmad, W. (2005). A Comparison of Object-Oriented and Pixel-Based Classification Methods for Mapping Land Cover in Northern Australia. Proceedings of SSC2005 Spatial intelligence, innovation and praxis: The national biennial Conference of the Spatial Sciences Institute, September 2005. Melbourne: Spatial Sciences Institute. ISBN 0-9581366-2-9
196. Wu, H.; Adler, R.F.; Hong, Y.; Tian, Y.; Policelli, F. (2012). Evaluation of Global Flood Detection Using Satellite-Based Rainfall and a Hydrologic Model. *J. Hydrometeorol.* 13, 1268–1284. DOI: 10.1175/JHM-D-11-087.1
197. Xia, M., Li, S., Chen, W., and Yang, G., Perceptual image hashing using rotation invariant uniform local binary patterns and color feature. *In Advances in Computers*, 2023,130,163-205. Elsevier. DOI: 10.1016/bs.adcom.2022.12.001
198. Xiao, X., Boles, S., Liu, J., Zhuang, D., & Liu, M. (2002). Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. *Remote Sensing of Environment*, 82(2-3), 335-348.
199. Xiong, J., Thenkabail, P. S., Tilton, J. C., Gumma, M. K., Teluguntla, P., Oliphant, A., ... & Gorelick, N. (2017). Nominal 30-m cropland extent map of continental Africa by integrating pixel-based and object-based algorithms using

- Sentinel-2 and Landsat-8 data on Google Earth Engine. *Remote Sensing*, 9(10), 1065.
200. Xu, H. (2006). Modification of normalized difference water index (MNDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. DOI: 10.1080/01431160600589179
 201. Yan, P., Tan, Y., Tai, Y., Wu, D., Luo, H., & Hao, X. (2021). Unsupervised learning framework for interest point detection and description via properties optimization. *Pattern Recognition*, 112, 107808. DOI: 10.1016/j.patcog.2020.107808
 202. Yan, S., Yao, X., Zhu, D., Liu, D., Zhang, L., Yu, G., Gao, B., Yang, J., & Yun, W. (2021). Large-scale crop mapping from multi-source optical satellite imageries using machine learning with discrete grids. *International Journal of Applied Earth Observation and Geoinformation*, 103, 102485. DOI: 10.1016/j.jag.2021.102485
 203. Yan, Y., & Ryu, Y. (2021a). Exploring Google Street View with deep learning for crop type mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171, 278-296. DOI: 10.1016/j.isprsjprs.2020.11.022
 204. Yu, Q., Wang, J., Wang, L., & Huang, D. (2020). Mapping soybean fields in the United States using Landsat imagery and machine learning approaches. *Remote Sensing*, 12(6), 1056.
 205. Zeng, Y., Hao, D., Huete, A., Dechant, B., Berry, J., Chen, J. M., ... & Chen, M. (2022). Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nature Reviews Earth & Environment*, 3(7), 477-493. DOI: 10.1038/s43017-022-00298-5
 206. Zhang, H., Li, J., Li, J., Huang, X., (2018). Mapping cropland extent and major crop types across China using Sentinel-1 data. *Remote Sensing*, 10(9), 1389.
 207. Zhang, H., Liu, L., He, W., & Zhang, L. (2019). Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 58(5), 3071-3084.
 208. Zhu, X., Cai, F., Tian, J., & Williams, T. K. A. (2018). Spatiotemporal fusion of multisource remote sensing data: Literature survey, taxonomy, principles, applications, and future directions. *Remote Sensing*, 10(4), 527.

APPENDIX - A

GT Data collection form sample

SI No.	Information collection	Detail
1.	Date of Observation	
2.	Crop Name	
3.	Field iD	
4.	Size of Crop Field	
5.	Date of Sowing	
6.	Place with District Name	
7.	Latitude	
8.	Longitude	
9.	Altitude	
10.	Accuracy	
11.	Field Photograph No.	
12.	Data Collector	
13.	Remarks	