Flood induced crop loss assessment based on satellite derived flood parameters and vegetation indices in North-Eastern Region of India

A Report submitted in partial fulfilment

for the award of the Degree of

Master of Science
in
Remote Sensing & Geographic Information System

by

Yengkhom Palmol Singh

pursued in

MSc. (Remote Sensing & Geographic Information System)

Amity University Noida UP

NORTH EASTERN SPACE APPLICATIONS CENTRE

27th May 201

Bonafide Certificate

This is to certify that the project report entitled "**Estimation of the crop loss base on satellite derived flood parameters and vegetation indices**" submitted by Yengkhom Palmol Singh to the North Eastern Space Applications Centre, Umiam, Shillong and Amity University, Noida, UP in partial fulfilment for the award of the degree M.Sc. (RS & GIS) in Amity Institute of Geo-informatics& Remote Sensing (AIGIRS), is a bonafide record of the project work carried out by him under my supervision from 7th January 2019 to 31st May 2019.

Shri Nilay Nishant
Sci/Engr SD
RS & GIS Division
North Eastern Space Applications Centre
Umiam, Shillong

Place: Umiam, Shillong

27th May, 2019

Declaration by Author

This is to declare that this report entitled "Estimation of the crop loss base on satellite derived flood parameters and vegetation indices" has been written by me. No part of the report is plagiarized from other sources. All information included from other sources have been duly acknowledged. I aver that if any part of the report is found to be plagiarized, I shall take full responsibility for it.

Yengkhom Palmol Singh
TR2019034
M.Sc. (RS & GIS)
Amity University Noida, UP

Place: Umiam, Shillong

Date: 27/05/2019

Acknowledgments

First of all, I would like to extend my heartfelt gratitude to Dr. Shri P.L.N. Raju, Director, North Eastern Space Application Centre (NESAC), for providing me the opportunity to work under this prestige organization and to explore new techniques.

I owe a very deep sense of gratitude to Shri Nilay Nishant, Scientist/Engineer (SD), NESAC, for all his precious and most valuable guidance, support, time and help during my entire period, without which this work would have been impossible. I would also like to thank Dibya Jyoti Chutia, senior scientist NESAC for allowing me to choose my area of interest for this project. I would also like to thank Ritu Anil Kumar scientist (SC) for her valuable support from the selection process till the end of project.

My sincere thanks also goes to all staffs (technical and non-technical) of NESAC for their help and encouragement during my work.

Lastly but not the least, I am also thankful to all my friends who help me directly or indirectly during my work.

Abstract

Flood one of the most devastating natural disaster often causes damage to the agricultural crop in India. In the study, crop loss is analysed for the Kharif season particularly in North-eastern states of India. To understand the event of a flood and its impacts on agriculture is a very important component, but it is a very complicated process at the same time. It is important to understand the start date of the flood, duration, crop growth stage and degree of crop damage, for crop monitoring and risk management in agricultural decision making. At present, the concerned agencies mostly depend on field surveys to obtain crop loss information and compensate farmers' loss claim, but this method is expensive, labour-intensive, and time-consuming. So such methods are applicable only at a small scale. Recent studies use the moderate resolution imaging spectroradiometer (MODIS) derived normalized difference vegetation index (NDVI) and MODIS NRT flood data, to find out the relationship between NDVI and flood event. The analysis is based on time-series difference and comparison of weekly NDVI between the flood years against the normal year (average/sum between years 2013-2018). The land cover data (LULC 50K first cycle) were used to isolate the input dataset pixels as Kharif crops. This study also describe the development of an Earth Observation (EO) base flood crop loss assessment, for supporting flood-related crop statistics and insurance decision-making. The studies proved that MODIS data can be useful to determine the crop loss impact due to flood.

Contents

List of Figures, Images and Tables

1. Introduction

Floods are the major natural hazard faced by many of the North-Eastern states every year during the monsoon season, which leads to a huge loss to crops. Agriculture is the major occupation and one amongst the prominent sources of livelihoods of the majority of people living in the North-Eastern states of India. However, the output from it is not sufficient to feed the growing population. In the year 2015, Manipur has experienced the worst flood in the last 200 years. All the major rivers were overflowed causing havoc and washing away connecting bridges, breaching of embankments, cutting off many villages from the mainland. Also, all the major canals are also overflowing which leads to inundate the paddy fields. In the year 2017, nearly 600 hectares of crop loss due to flood (https://www.ifp.co.in/page/items/38660/nearly-600-hectares-of-crop-lost-in-flood/).

Assam is one of the worst flood-affected the state in India. Every year during the monsoon season the river Brahmaputra and its tributaries has been causing a flood to the Brahmaputra Valley, which leads to huge loss of cultivated lands in terms of lakhs of hectares. According to Rastriya Barh Ayog (RBA), the flood-prone area of the state is 31.05 lakh hectares (about 39.58% of the total land of Assam) (Karuna Phukan Dept. of Economics, Dibrugarh University). During the year 2014, the total crop area (in hectares) affected due to flood is 372178 (Source: NE data bank (NEDFi)).

The traditional methods to assess the crop loss is through field visit and survey the local farm, but this method is pertinent only when the study area is an easily assessable and small area, however, this method is extremely time-consuming as well as cost. In order to overcome these obstacles scientists use remote sensing techniques. It has been proved that satellite images and their products can be used in crop yield assessment efficiently because of their wide area coverage, frequent revisit, rapid data distribution, relatively low data cost, and strong crop/land discrimination.

Over the past few decades, flood events and the accurate assessment of their impact on crops have been significant research components of food security and agricultural planning (Smith 1997; Tholey et al. 1997; Shrestha et al. 2013; Yu et al. 2013; Kang et al. 2014). Assessing the direct impact or loss as well as understanding the indirect economic impact of floods could be vital information for decision-makers for precise cost-benefit analysis and appropriate prevention plans. Assessing flood damage is not only important to understand the direct loss of crop production, but also necessary to recognize significances towards the overall food market (Del Ninno et al. 2003). Although it might seem to be a direct task, it is an extremely difficult and delicate method (Frazier and Page 2000).

Vegetation indices (VI's) based crop loss assessment methods are much more vigorous compared to the image classification based methods. By transforming the daily products into weekly/bi-weekly composite products, the VI is able to reduce data anomalies such as cloud cover effect. VI's had been used to monitor crop growth through remotely sensed techniques (Ren et al. 2008; Liu et al. 2012). The majority of these VI's are reflective indicators providing a quantified measure on crop conditions. Among these various VI's, NDVI is considered to be the primary index exercised to monitor crop conditions (Quarmby et al.

1993; Kang et al. 2013; Yu et al. 2013; Johnson 2014). Moreover, the study showed that with regression analysis NDVI residuals were able to reflect crop stress, especially during the crop-growing period. As healthy vegetation is highly reflective to the infrared part of the spectrum, the plant produces high NDVI value and stressed plant shows lower infrared reflection, resulting in low NDVI. The image 1.1 shows the land use of the region.

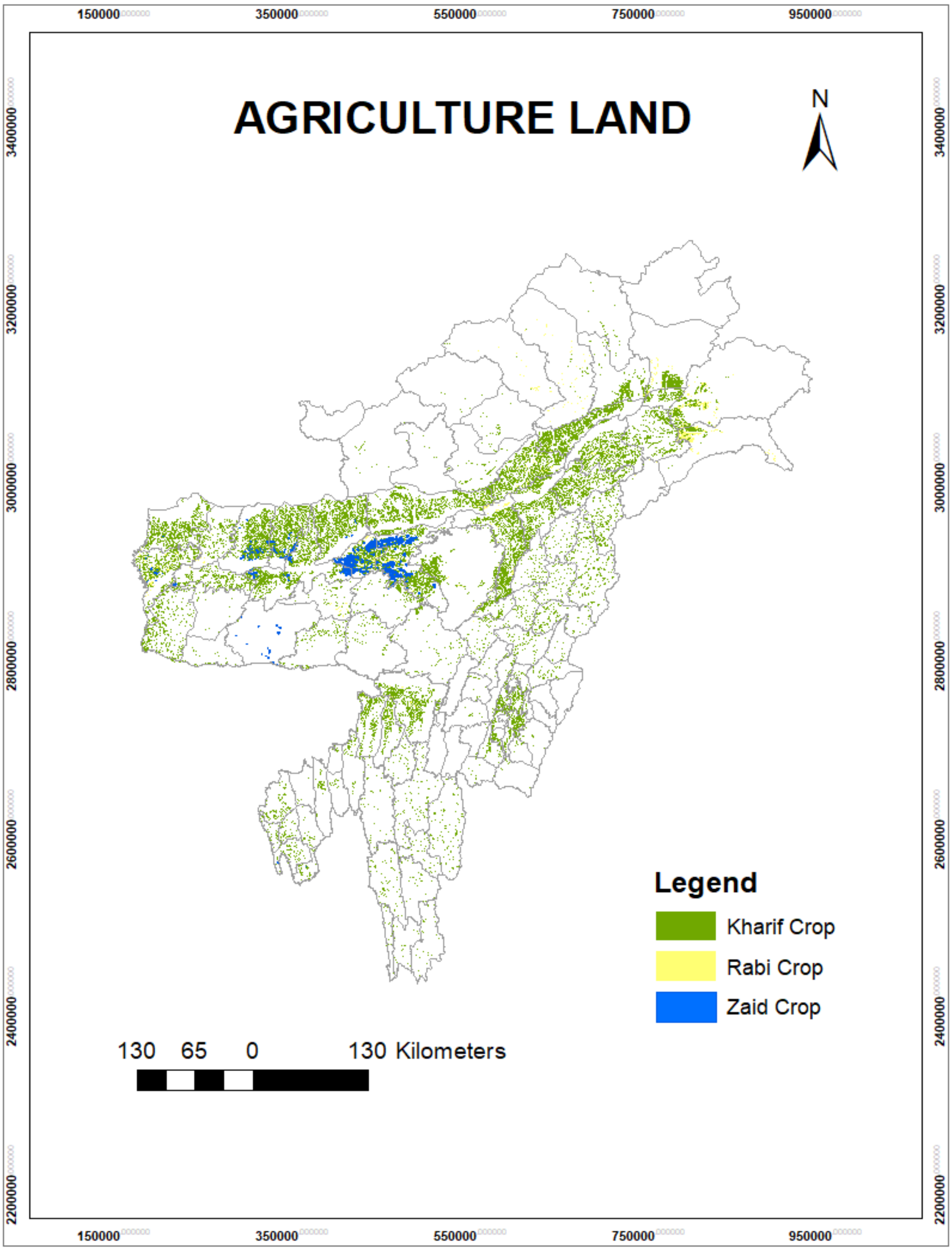


Image 1.1 Land Use showing kharif, rabi and zaid crop.

After the event of a flood, it is important to know the extent of flood water to crops and its duration for crops loss estimation. The high temporal resolution of MODIS is a very important factor to be considered in early recovery assessment since timeliness is the most important factor to be considered at this stage. It is well known that optical remote sensing is affected by clouds, which commonly present during and shortly after a flood event. The high temporal resolution significantly improves the probability of acquiring cloud-free data. The delay of data acquisition could be just days with MODIS instead of two or more weeks in the case of ETM+. Therefore, MODIS and its derived standard products from National Aeronautics and Space Administration (NASA), such as NDVI, are chosen as the primary Earth Observation (EO) data source for producing the flood damage assessment information.

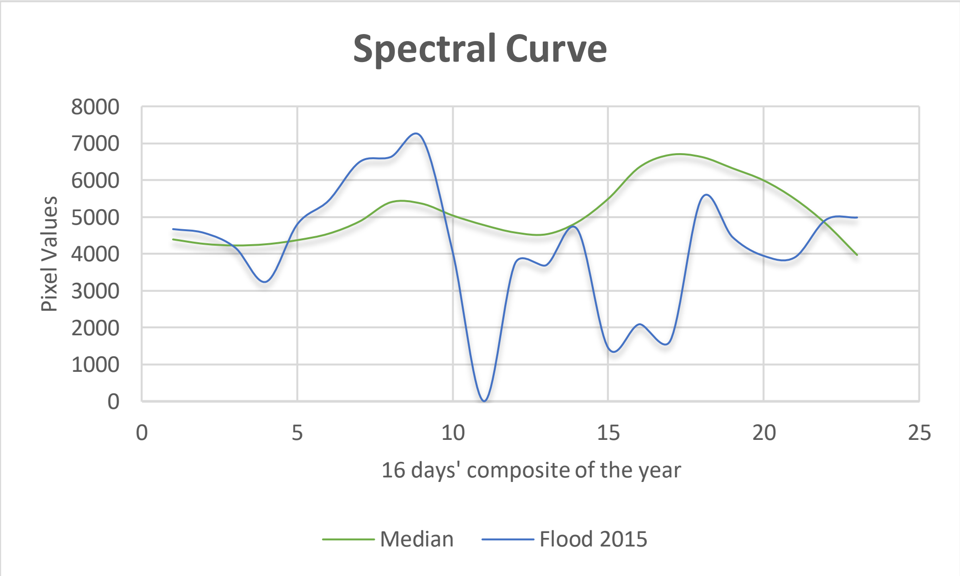


Fig. 1.1: Spectral Curve showing normal year and flood year.

The spectral reflectance shown in fig. 1.1 indicates two spectral curve one in normal year and another one in flood year. The median is calculated from year 2001-2017 and the flood year 2015, it is clearly seen in the image that the curve is falling sharp in the month of June where the flood happen.

The aim of the study will be to utilize daily MODIS NDVI product and MODIS NRT to detect the relation of NDVI with crop yield using linear regression and find out how the flood water affect the crop yield, finally examine whether these NDVI indicators can actually be utilized to assess the crop yield as well as crop yield loss due to flood.

1.1 Literature Review

(Ranjay Shrestha *et al.* 2017) conducted research where it was mentioned the importance of studying the flood event in order to understand their impact on crop and its damage is directly related to crop yield, so require an accurate assessment to quantify the damages. The moderate resolution imaging spectroradiometer (MODIS) weekly normalized difference vegetation index (NDVI) is used to detect and analyse the flood damage to corn during the growing season. The NDVI for the year 2000-2014 including both the flooded year as well as the normal year was used to detect flood occurrence. The regression analysis was performed between change in NDVI and change in corn yield as independent and dependent variables respectively for further analysis on the impact of the flood on corn yield. The result shows that the linear regression approach is a good indicator to quantify the yield loss due to flood. Additionally, with the 250 m MODIS-based NDVI, these yield losses can be estimated up to field level.

Flood creates a drastic change to crop condition profile and this change relates to crop damage caused by the flood. Using satellite derive different vegetation indices crop condition profile and its damage due to flood can be estimated. Vegetation index such as normalized difference vegetation index (NDVI), vegetation condition index (VCI), mean vegetation condition index (MVCI), and ratio to median vegetation condition index (RMVCI) were used in the study but MODIS derived product is used mainly due to its high temporal resolution that is very important for the constructing crop condition profile. The study shows that crop condition profiles can effectively detect flood damage and estimate the damage due to flood (Yu G E *et al*).

(Youngjoo Kwak *et al*. 2015) conducted research to estimate crop loss due to flood in Cambodia, using the integrated approach to detect and monitor flood areas with flood depth and duration for near real-time rice crop damage estimation using MODIS time-series data. The data used in the study include MLSWI, EVI from MODIS, and new FID from DEM, land use, and simplified empirical damage curves. The study foundthattheMLSWIthresholdcanbemosteffectiveat0.75to detect water bodies. The rapid flood damage assessment showed the possibility to produce distribution maps for the NRT estimation of rice crop damage by using the MODIS-derived MLSWI and EVI data before and after flooding.

(Youngjoo Kwak *et al*. 2016) conducted an experiment for rapid damage assessment due to flood and disaster risk reduction. They have mentioned that it is important to identify and characterize flood area, location and duration of flood for the purpose of flood risk reduction. To identify flood inundation area for the whole nation is a big challenge but in this research, they use MODIS NRT data. It has been finding out that this data is useful not only to detect flood inundation area but also for flood mapping and monitoring. The study shows that for flood risk assessment it highly depends on the temporal and spatial dynamics of exposure such as distributed population for the Bangladesh region.

(Kwak Y *et al*. 2014) conducted research on surface water detection on the base of the Indus river flood 2010 which affected entire Pakistan. A modified surface water index derived from near-real-time Moderate Resolution Imaging Spectrometer (MODIS) images and digital elevation model (DEM) was used. The detection of flood waters and the estimation of flood volumes are important to determine a hazard in flood risk. The results found that the MODIS-DEM combined approach was possible for automatic, instant flood detection.

(Shrestha R *et al.* 2016) studied regression base corn yield assessment using MODIS base daily NDVI in Lowa State, USA. They have mentioned that the traditional method of visiting field and surveying farmers to estimate crop yield has been considered inefficient and impractical especially in cases when fields are not easily accessible. Remote sensing techniques, therefore, has been utilizing to overcome these obstacles with good success. Normalize Difference Vegetation Index (NDVI) based models are considered to be the most effective and utilized technique in crop yield assessment and can provide up to field level assessment. MODIS based 250m daily NDVI is used in the study in order to estimate the corn yield in 4 Agricultural Statistics Districts (ASD) in Iowa State. A linear regression model was derived between the NDVI curve and corn yield using all counties within the 4 ASD between years 2000 to 2014. The regression model showed a statistically significant relation between NDVI curve and corn yield with a coefficient of deterministic (R-square) over 0.80 in all 4 ASD.

(Gumma M K *et al.*2014) studied on a mapping of rice cropland extent and area in high cropping intensity environment of Bangladesh using MODIS data. Rice is mostly consumed by Asian where the product intensity is very high, still not sufficient for the people of Asia. One of the reasons for the difference in product intensity is due to natural disasters as well as human-induced disaster. It is important to understand the spatial and temporal information seasonally for the rice crop extent in order to understand the agricultural productivity and sustainable use of limited resources. MODIS derived NDVI and seasonal field plot information used to map rice crop extent and area of three seasons. The result of the study shows that moderate spatial and high temporal resolution can be used effectively for capturing the seasonal variability in rice crop extent and area.

(Di L *et. al.*2017) mentioned in their studies that information on crop loss, such as flooded acreage and degree of crop damage on time is very crucial for crop monitoring and risk management in agricultural and any disaster which affects the crop. Field survey is the traditional method to obtain the information on crop loss, but it is time consuming, tedious and expensive. It is applicable only in a small area. Current studies try to prove that flood crop loss can be estimated using different data from Earth Observation (EO) such as MODIS derived NDVI and NRT data, other than this they have also used geospatial products which include Cropland Data Layer (CDL), Common Land Unit (CLU) and administrative data. The studies show that remote sensing base flood crop loss service system (RF-CLASS) is useful in flood monitoring and assessment, finely generate the flooded crop loss products and proved the possibility of using such products to improve the agricultural decision-making.

(Del Ninno C *et. al.* 2003) conducted a study on one of the worst flood in the history of Bangladesh which inundated two-third of whole country in the year 1998, causing severe damage on rice crop and threatening the food security for the millions of households. The flood peak was happen in early September, due to this damage on rice crop is severe as the harvesting season is just about to arrive. This paper highlights the contribution by government, including liberalization to stabilization of rice markets during and after the flood. In total they cover 750 households in three rounds over a 13-month period then, they analyse the impacts of the flood on household assets, consumption and nutritional outcomes. In this paper they also highlight the 1974 flood where thousands of people were dead due to famine in 1975, the main cause was the increase of food price during and after the flood. Finally, they highlight the role of private markets and government investment to maintain the food availability, limiting price increases and supplementing household access to food, thereby helping to avoid a major food crisis.

(Quarmby N A *et. al.* 1993) uses multi-temporal NDVI measurement of AVHRR data to estimate and predict crop yield. The main parameter used in the study to monitor and predict the crop yield is normalized difference vegetation index (NDVI). The major problem of using AVHRR data for crop monitoring is large area covered by one pixel in ground is comparable to field size. Yield has been estimated for the crop such as rice, cotton, maize and wheat with the high degree of accuracy using linear regression between NDVI and yield. The study illustrates that the crop yield information would be available prior to 50-100 days of harvest which enable early assessment.

(Smith L C. 1997) studied on river inundation area, stage and discharge using satellite remote sensing. In the study he had mentioned the use of active and passive remote sensing for estimating inundation area and delineate flood boundaries. One of the active remote sensing radar altimeters is used for directly measuring stage variation in large rivers. It also possible to obtain estimates of river discharge from space, using ground measurements and satellite data to construct observed curves that relate water surface area to discharge. Synthetic aperture radar (SAR) sensor has been used in the study as it can penetrate the cloud, forest canopies and also can detect the standing water through standing aquatic plants. He also suggest that the single polarization, fixed frequency SARs are not sufficient for mapping inundation area in all environment. However, by utilizing both visible/infrared and SAR data may be effective way to monitor inundation area in vegetated riverine environment. Study found out that the Radar altimetry has the potential for measuring stage in large rivers to within 10 cm.

(Ren J *et. al*. 2007) conducted a research on yield estimation for winter wheat using MODIS-NDVI data in Shandong, China. The importance of yield estimation is well known in agricultural management and policy development at regional and national levels. They have used the MODIS-NDVI data, with a 250 m resolution to estimate wheat yield. The Savitzky-Golary filter was applied in the MODIS data to remove the cloud and improve the quality of image. NDVI at country level was derived in order to find its relationship with winter wheat production and linear regression between them was developed. The results were validated using the ground survey data it has been found out that the good predicted yield data could get 40 days before the harvesting time. Finally the author suggested the method used in the study is good in predicting winter wheat production and yield estimation.

(Kang L et. al. 2014) mentioned in their studies that NDVI had been used in many studies related to precipitation correlation but the NDVI correlation is affected by many factors such as soil background. NDVI is widely used for monitoring plant growth status, also in agriculture application such as drought monitoring, phonological estimation, crop yield prediction, land cover change detection and biomass estimation. Here, they applied a new approach geographically weighted regression model to analyse the NDVI correlation on three different types of land use (grass land, fallow land and winter wheat land). The result shows that this model has better than the previous models. The R^2 value and proportion of residual with spatial autocorrelation is showing better result in geographically weighted regression model as compare to global regression model.

(Liu J *et. al.* 2012) conducted a research on crop green leaf area index (LAI) estimation from Landsat images for multiple growing seasons to assess the vegetation indices. They have used the regression model to estimate crop LAI from a vegetation index derived from satellite data. To estimate the green LAI, a semi-empirical equation was applied from Landsat-5/7 data using a few vegetation indices, such as normalized difference vegetation index (NDVI), optimized soil adjusted vegetation index (OSAVI), the two band enhanced vegetation index (EVI2) and the modified triangular vegetation index (MTVI2). To remove the uncertainty in LAI estimation the PROSPECT leaf model coupled to the SAIL canopy model along with the 6S atmospheric transmission model was applied. The sensitivity analysis shows that NDVI is most swayed by leaf chlorophyll but the least affected by leaf inclination, OSAVI and MTV2 are more efficient in reducing soil effects, and EVI2 has a better performance in decreasing aerosol perturbation. The final result shows, the ability of vegetation indices in generating regression model base on green LAI product for seasonal growth monitoring.

(Shrestha R *et. al.* 2013) studies on land/water detection and delineation using Landsat data. The accurate information on water detection and delineation is vital for research area such as flood prediction, monitoring and relief, wetland inventories and water resource evaluation. Remote sensing technology is implemented in the study as the traditional method of surveying the actual field site and visiting consumes time as well as difficult to access the area in case flood, cyclone, etc. Landsat provides multiple bands however, it is important to understand how to use the image and which spectral band and classification methods to use for the best hydrological classification. They have use the Matlab and ENVI software for this paper and concluded that band 5 i.e. mid-infrared in Landsat TM is best suitable for land/water delineation. In addition, single-band density slicing with only 2 ranges (water and non-water) appeared to be an outstanding method to discriminate water from other land features using Landsat TM band 5.

(Johnson D M. 2014) studied on country-level corn and soybean yield forecast for Corn Belt region of United State (US) using timely available remotely sensed datasets. The datasets used in the study include- normalized difference vegetation index (NDVI), day time and night time land surface temperature (LST) and precipitation from the National Weather Service (NWS). The study was conducted for the growing season of 2006-2011, the land cover data were used to isolate the input dataset pixels as to corn and soybeans. It has been found that there is a positive correlation for both corn and soybean yield with NDVI particularly in the middle of summer and there is negatively correlated to daytime LST at the same time. However, night-time LST and precipitation showed no correlations to yield, regardless of the time prior or during the growing season. Consider only NDVI and day time LST as input from 2006-2011 datasets, regression model were applied in both crops and determined R^2 coefficient is found to be 0.93. At last, the available models were used to predict the sample for the year 2012.

(Frazier P S *et. al.* 2000) conducted a simple research on detecting and delineating the water bodies using Landsat TM data. The main concern of the study was to determine the accuracy of using simple digital image processing techniques to map riverine water bodies of flood plain areas. The outcome of the study is compared with 6-band maximum likelihood classification, also the water boundaries delineated by digital image processing were compared with water boundaries delineated from colour aerial photography. Density slicing of the single mid-infrared band 5 proved as effective as multispectral classification succeeding an overall accuracy of 96.9%, a producer's accuracy for water bodies of 81.7% and a user's accuracy for water bodies of 64.5%.

1.2 Problem Statement

A flood is an unstoppable event in the NE regions particularly in Brahmaputra valleys, Barak valleys and Manipur Valleys. Agriculture is the main source of income as well as the economy in this region but due to flood, it's hampering. One of the main reason for unable to handle the problem is lack of availability of information on flood inundation area, flood duration, crop acreage, etc., to get that information on time is a challenging task for a large region.

The Earth Observation (EO) data could be used to get that information to determine the post-flood crop loss assessment. MODIS derived vegetation indexes (VI's) can be to use to get the health condition of the crop and generate the statistics along with flood area and yield. Such information could be useful to decision makers.

1.3 Objectives

(i) To evaluate the possibility of using MODIS derived NDVI to forecast crop yield on the NE region.

(ii) Crop loss assessment base on MODIS NRT flood data.

1.4 Study Area

For the study area, the North-eastern states were selected except Sikkim, the selected states are Assam, Arunachal Pradesh, Manipur, Mizoram, Nagaland, Tripura, and Meghalaya. Northeast India is located between latitude 25.5736° N and longitude 93.2473° E and covers a total area of 255168 km^2 (Image 1.1). The region has a population of 39 million and the geographical area of 26.2 million hectares, which is 3.85% and 8% of the population and area of the country, respectively. In the NE states, there are roughly two different types of agricultural practices- (i) settled farming practiced in the plains, valleys, foothills and terraced slopes, and (ii) shifting cultivation (Jhum) practiced on the hill slopes.

The region is rich in biodiversity and major areas are under sustenance agriculture. Rice is the major annual crop in this region, while tea, jute, cotton, potato, sugarcane, and oilseeds are the major cash and annual crops (S. Naresh Kumar et al. 2011). Northeast (NE) India has 64% of the total geographical area under forest cover. The entire region is a part of Indo-Burma and Himalayan hotspots, 2 of 25 such hotspots in the world (S. K. Jain et al. 2012).

Agriculture is the dominant source of economy in this region any changes in the spatial and temporal pattern of rainfall directly affects the monsoon dependent agriculture ecosystem. NE India is vulnerable to water-induced disasters such as flood due to its location in the eastern Himalayan periphery, fragile geo-environmental setting, and economic underdevelopment. The most vulnerable area in this region flood is the Brahmaputra valley and Barak valley and its tributaries. Manipur valleys are also one of the most flood-prone areas due to its oval shape. The total geographical area of the valley is 1900 sq.km and the total catchment area of Manipur river system is 6332 sq.km. The topography of Nagaland is much divided, full of hill ranges, which break into a wide disorder of spurs and ridges. Even though 90% of the area is mountainous Nagaland face the worst flood in the year 2018 which is followed by landslides, which leads to loss of property, livelihood and even loss of life.

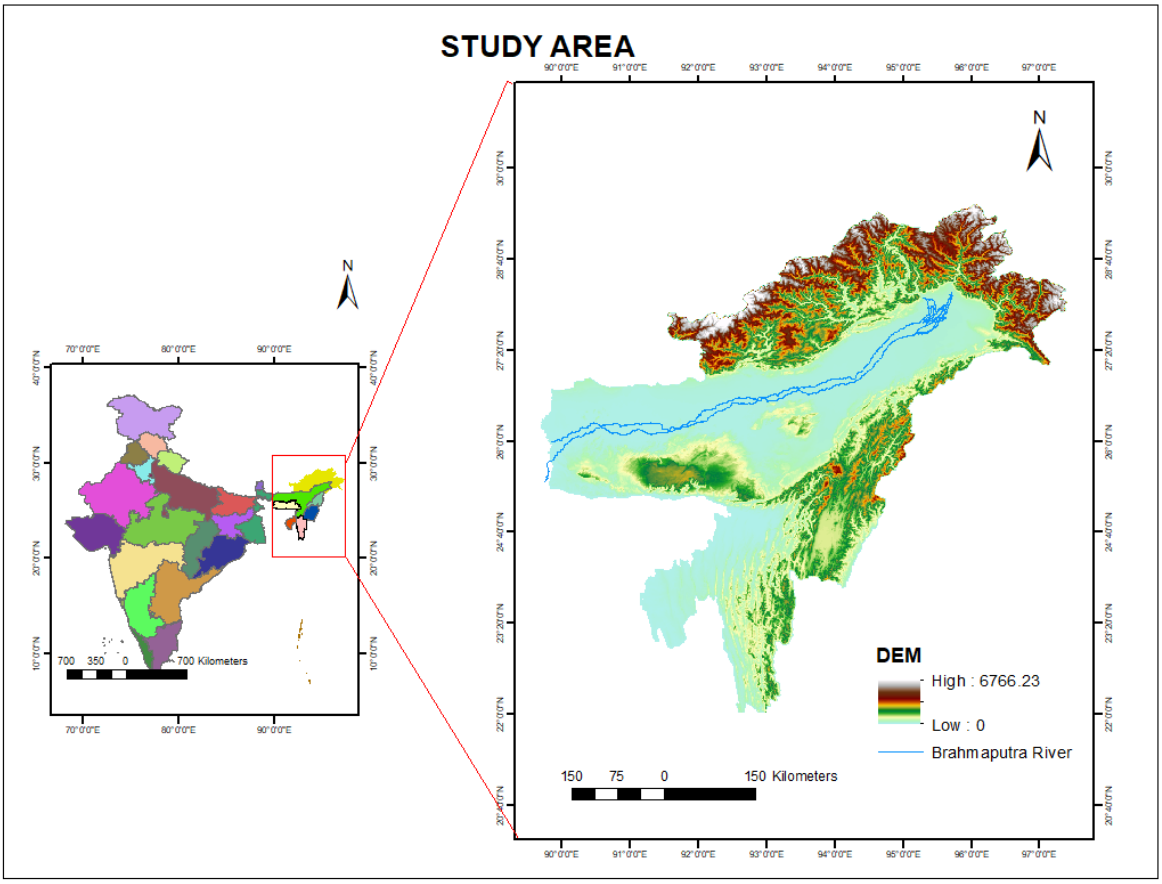


Image 1.2: Study Area

2. Data & Methods

2.1 Data Use

The primary data used for this study include MODIS daily NDVI 16 day's composite, MODIS NRT flood data 14 days' composite. For the crop type identification, 50k LULC of the first cycle for the North East was used. The shapefile of the study area with district wise boundary was provided by NESAC. Secondary data include the crop yield for the year 2013-2017, which is downloaded from Data.gov.in (https://www.data.gov.in). Others data used in this study include- SRTM DEM for the whole NE region.

Moderate-Resolution Imaging Spectroradiometer (MODIS) is an Earth Observing System (EOS) instrument on board the Terra and Aqua platforms, launched in December 1999 and May 2002. MOD13A2 16 day's composite is downloaded from Land Processes Distributed Active Archive Centre (LP DAAC) (https://www.lpdaac.usgs.gov/lpdaac/get_data/data_pool). MOD13A2 has twelve products at 1km spectral resolution in this study consider only the Normalized Difference Vegetation Index (NDVI) product. MODIS NRT Flood data was downloaded NRT Global Flood Mapping (https://floodobservatory.colorado.edu).

2.2 Methods

The pre-processing methods need to be carried out before the work process starts. MODIS NRT Data came in a zip file so we need to extract the file. Extracting hundreds of data and arranging is time taking so we use the python to unzip the file and arrange in the proper folder. Now the unzip shapefile has to clip as per the study area, add the field in attribute table as "Value" and using calculate geometry give the value as "1". All this process is performed using a model maker. The clip shapefile with attribute value has to be converted to raster in order to find the frequency of the flood. To do this, we use batch processing from feature to raster tool in ArcGIS. The converted raster files have the only one value i.e. 1 for the flooded area but we don't have the value for non- flooded area for that we need to make the null area as '0'. In order to do this, we need to make another model.MOD13A2 Data coordinate system came in sinusoidal so we need to re-project to UTM WGS 1984. To do this we used the Batch Processing tool from Arc toolbox such that we can project the multiple numbers of raster file at once. The shapefile of the study area is converted to grid into 20 zones which correspond to 0.3 degrees, using create fishnet tool in ArcGIS. Yield data pre-processing include the arranging of data for the year 2013-2017 according to district wise for the Kharif season only, to do this Microsoft Excel is used.

2.2.1 MODIS-NDVI Data Processing

Time series of MODIS MOD13A2 product (Vegetation Indices 16-Day L3 Global 1km) over the period from July 2013 to October 2018 have been used in this study. The dataset provides two VIs layers: Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI). This study was focused on NDVI because it is widely used in phonological works, e.g. Boschetti *et al*. (2009), Colombo *et al*. (2011), Soudani *et al*. (2008), Hmimina *et al*. (2013) and because it is known to be more sensitive to small increases in the amount of photosynthetic vegetation (Soudani *et al*., 2006; Soudani *et al*., 2008; Sesnie *et al*., 2012).

MODIS derived NDVI came in Hierarchal Data Format (HDF) file so it first needs to convert to image format using ERDAS Imagine. At first, the original NDVI layers were stack for all the years selected, there is a total of 23 layers and then crop cover particularly Kharif crop is masked that delineated only those areas under Kharif crop. A copy of the crop cover LULC was obtained North Eastern Space Application Centre (NESAC) and used to eliminate the influence of non-agricultural and non-annual crops on the NDVI signal. Consequently, all areas with non-agricultural land were masked out and NDVI values were extracted for only those areas having Kharif crops. Then calculate the mean of for Kharif season, as the MOD13 NDVI value range from -2000 to 10000, it needs to correct the scale divided by 10000.

Table: 2.1 - 1: Composite periods of the MOD13 year, 2: first day of each compositing period, 3, 4: Julian day to Gregorian dates, 5: corresponding season

(1) Composite	**(2) CDOY**	**(3) Non-Leap year**	**(4) Leap year**	**(5) Season**
1	1	01-Jan	01-Jan	Winter
2	17	17-Jan	17-Jan	
3	33	02-Feb	02-Feb	
4	49	18-Feb	18-Feb	
5	65	06-Mar	05-Mar	
6	81	22-Mar	21-Mar	Spring
7	97	07-Apr	06-Apr	
8	113	23-Apr	22-Apr	
9	129	09-May	08-May	
10	145	25-May	24-May	
11	161	10-Jun	09-Jun	
12	177	26-Jun	25-Jun	Summer
13	193	12-Jul	11-Jul	
14	209	28-Jul	27-Jul	
15	225	13-Aug	12-Aug	
16	241	29-Aug	28-Aug	
17	257	14-Sep	13-Sep	
18	273	30-Sep	29-Sep	Autumn
19	289	16-Oct	15-Oct	
20	305	01-Nov	31-Oct	
21	321	17-Nov	16-Nov	
22	337	03-Dec	02-Dec	
23	353	19-Dec	18-Dec	

2.2.2 MODIS NRT Flood Data Processing

The MODIS Near-Real-Time Global Flood Mapping product contains global daily surface and flood water maps at approximately 250 m resolution, in 10x10 degree tiles. There are four different types of products: (i) MFW: MODIS Flood Water, (ii) MSW: MODIS Surface Water (e.g. MFW before subtracting the reference water), (iii) MWP: MODIS Water Product (combines both MFW and MSW, raster only), and (iv) MFM: MODIS Flood Map = annotated 10x10 degree map/graphics product (png format). For the study, MFW for the period from July 2013 to Oct 2018 with 14 days' composite is used to avoid cloud cover issue. The layers for the selected months are stacked, the stack image is in percentage so we need to convert it to binary, to find out the flood peak day and then sum the layers using the stack sum function in Erdas Imagine. Finely calculated the percentage of the flood occurrence. But there is some error in the image, the snow region in Arunachal Pradesh is shown as flood water. To eliminate the snows misclassified as water, digital elevation model is used to eliminate the water pixels on slopes. Also, find out the maximum days of flood from the stack image using stack (maximum) tool from ERDAS Imagine.

2.2.3 Yield Modelling

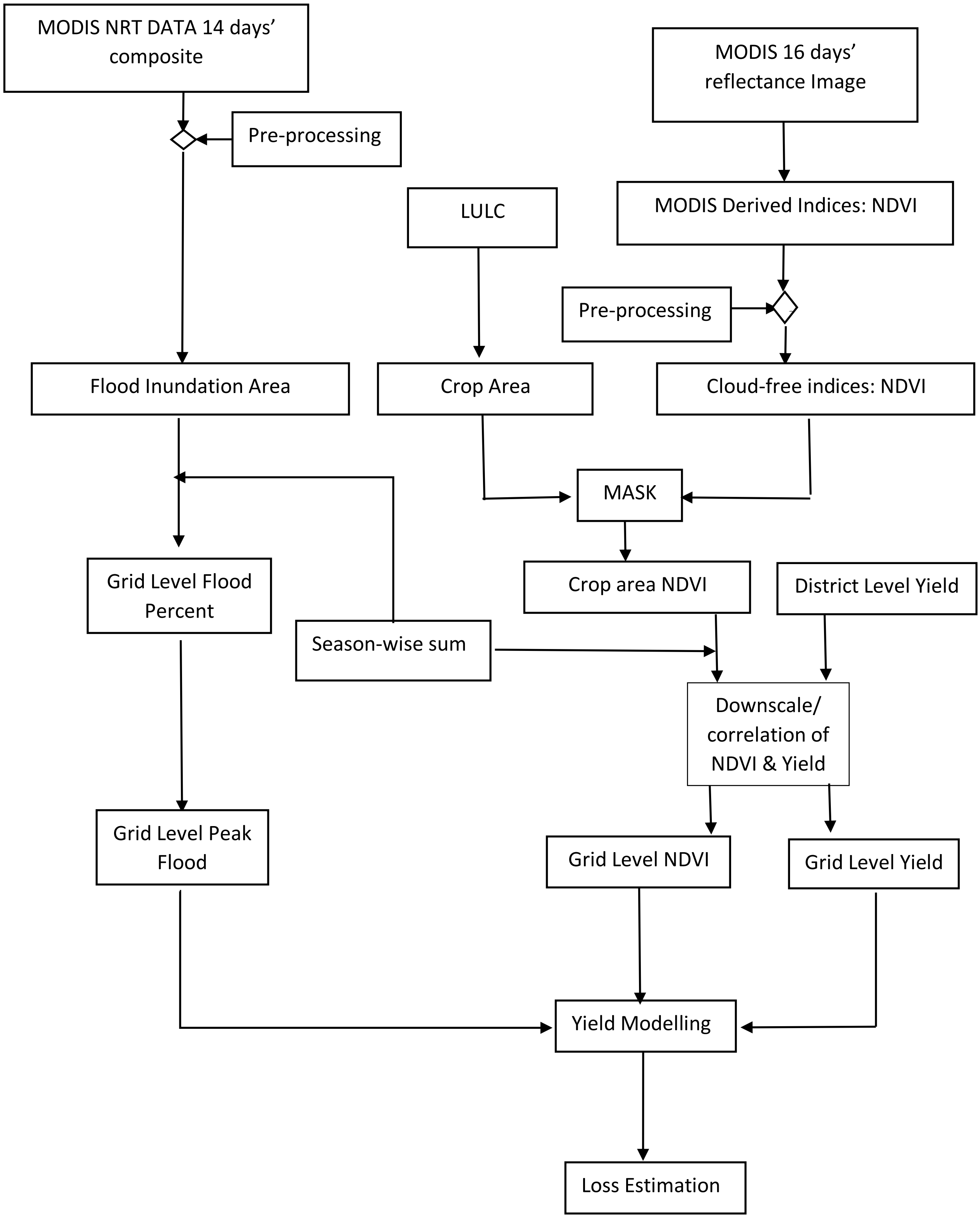
Before processing the yield modelling, first student t-test is needed to run in order to find out the significant districts such that it will give a positive correlation with the NDVI. The t-test is the hypothesis test, it analysis if the mean of two data sets is greatly differenced from each other, i.e. whether the population means is equal to or different from the standard mean. The following is the t-test formula:

$$t = \frac{X - \bar{X}}{\frac{SD}{\sqrt{n}}}$$

Where, X = sample mean, $\bar{X}$ = specified mean, SD = standard deviation, n = number of counts

$$SD = \sqrt{\frac{\sum(X^2 - \overline{X^2})}{n}}$$

Figure 2.1: Methodology Flow Chart



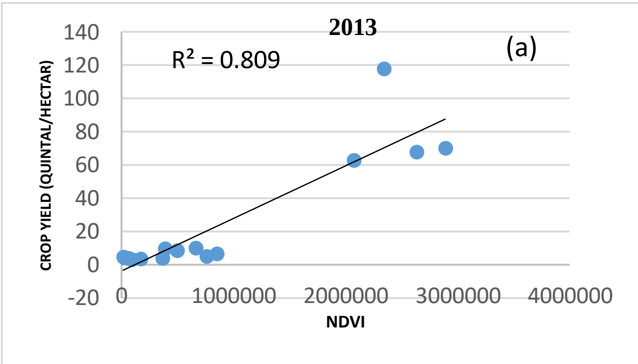
3. Results and Discussion

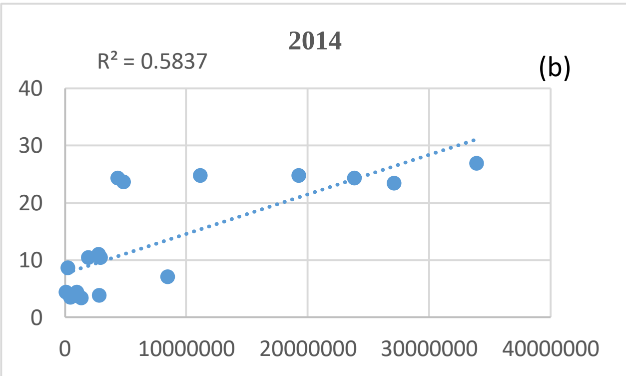
3.1 Relationship between NDVI and crop yield

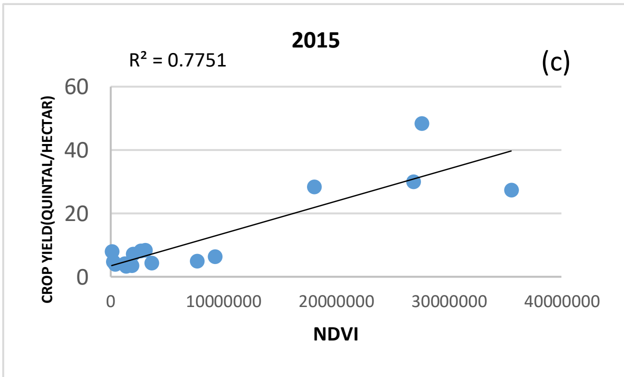
The linear relationship between yields data and MOD13A2 for the 4 years (2013-2016) showed the fit between NDVI and yield as positive for all years, the highest relationship found in the year 2013 which show R^2 =0.809 and lowest relation show in the year 2017 ($R^2 = 0.3216$). The following table 2 shows the simple regression between NDVI and yield.

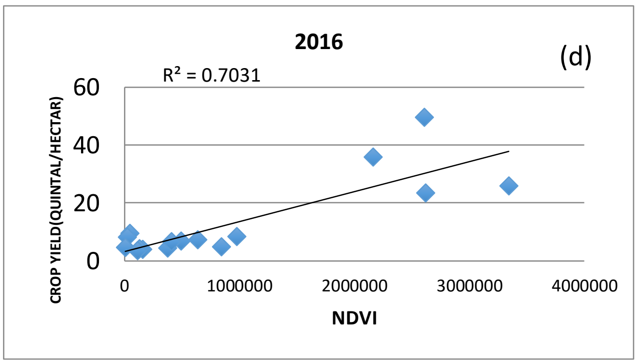
Table 3.1: Simple regression between NDVI and yield

Years	R^2
2013	0.809
2014	0.583
2015	0.775
2016	0.703
2017	0.3216









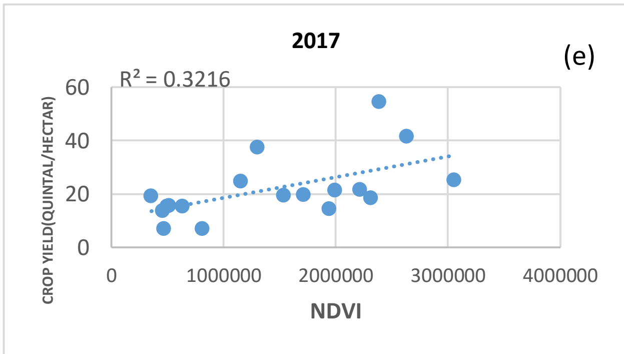
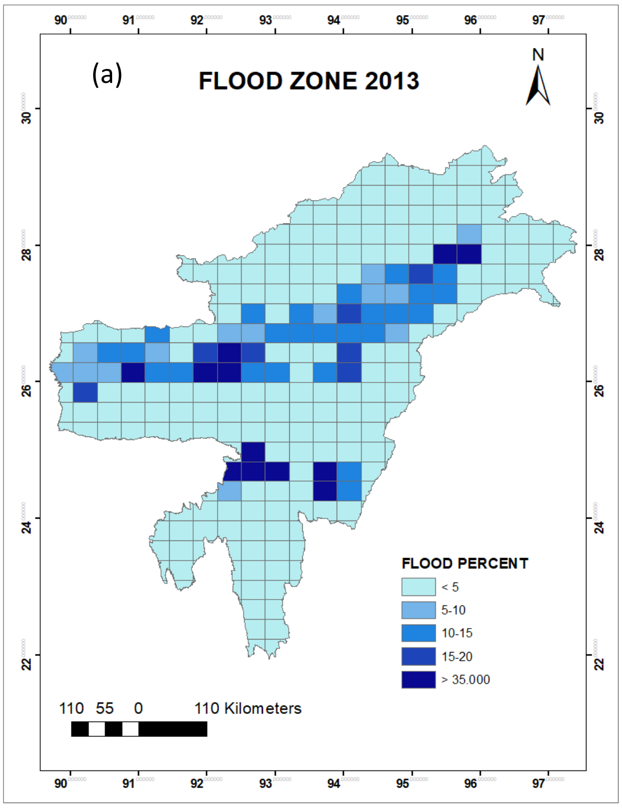
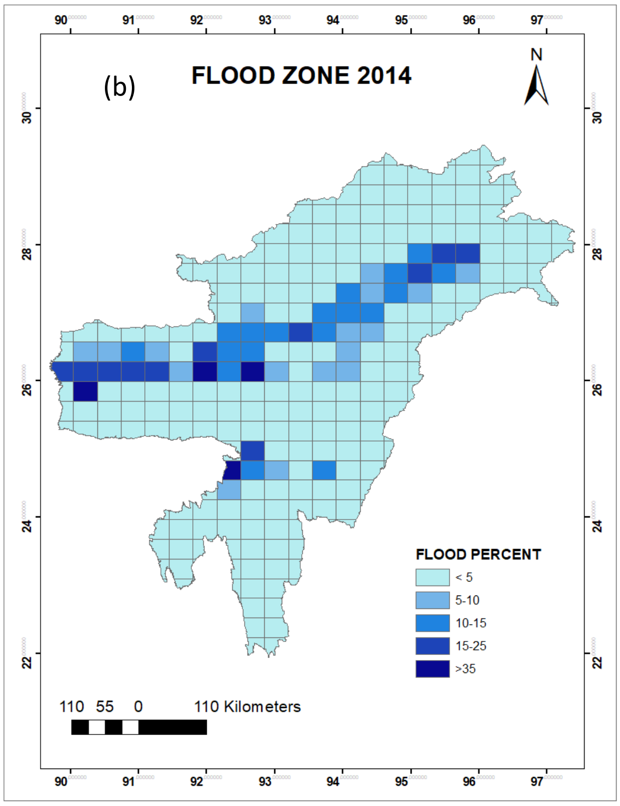


Figure: 3.1 NDVI and crop yield correlation (a) 2013, (b) 2014, (c) 2015, (d) 2016 & (e) 2017.

3.2 Flood Trend Analysis

The flood record for the year 2013-2018 is studied in this study, the analysis is carried out only the significant zones of the study area. The study shows that flood is concentrated mainly in the Brahmaputra Valleys, Barak Valleys and Manipur Valleys as shown in image 3.1 (a), (b), (c), (d), (e) and (f), for the year 2013, 2014, 2015, 2016, 2017 and 2018 respectively. Flood severity level is shown in five classes: < 5 indicates no flood or very less flood, 5-10 indicates less flood, 10-15 indicates moderate flood, 15-25 indicates severe flood and above 35 indicates very severe flood. Much of the severe flood is happen in the Cachar, Karimganj, Hailakandi, Karbi Anglong, Bishnupur, Imphal West, Thoubal. Manipur valley shows high flood percent in the year 2013, 2015, 2016 & 2017, among them the year 2015 shows the highest flood percent. In Assam, Barak valley and Brahmaputra Valley is worst affected area, almost all the years is show high flood percent in the Barak Valley.





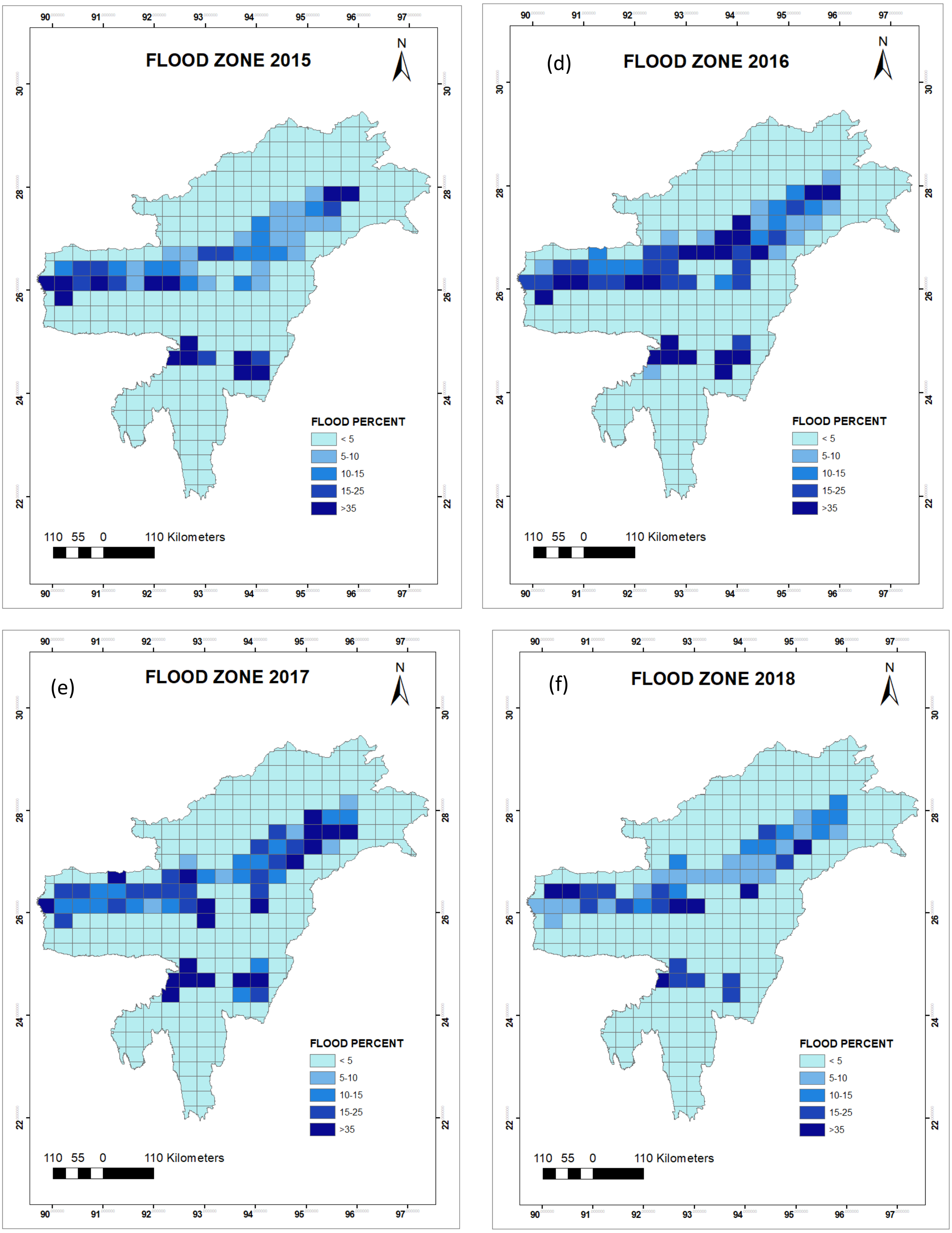
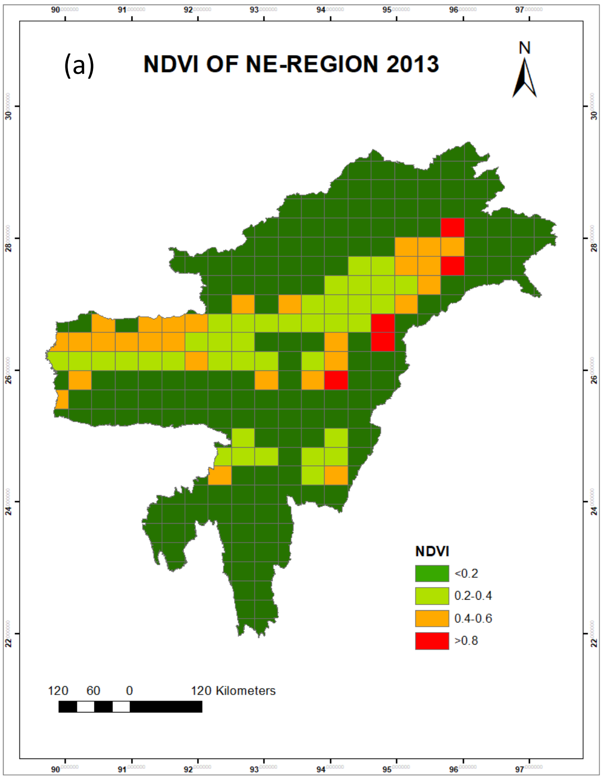
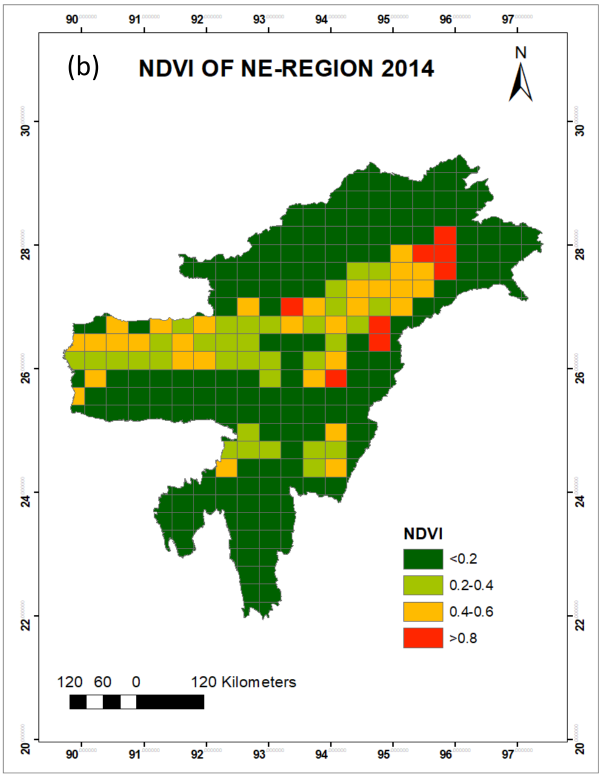


Image 3.1 Flood Zone (a) 2013, (b) 2014, (c) 2015, (d) 2016, (e) 2017 & (f) 2018

3.3 NDVI Trend Analysis

NDVI for the year 2013-2018 is analysed, NDVI range is classified into four classes such as less than 0.2, 0.2-0.4, 0.4-0.6 and above 0.8. The class below 0.2 indicates very low NDVI that means there is less vegetation cover in that particular cover, as the analysis is carried out only in significant zone, the dark green part is excluded from the analysis. In general the NDVI value from 0.2-0.3 indicates the shrub and grassland, crops NDVI value is particularly from 0.4-0.5 and 0.6-0.8 NDVI value indicates temperate and tropical rain forest. In the year 2013 much of the region is showing NDVI value range from 0.2-0.4 and 0.4-0.6, this indicates that the region mostly covered with crops, only few part is showing high vegetation value above 0.8 which is shown in red colour. The most common NDVI value range for all the year is 0.4-0.6, some of the northern region shows high NDVI value, those are the region where flood percent is low. The NDVI trend is shown in image 3.2 (a), (b), (c), (d), (e) and (f), for the year 2013, 2014, 2015, 2016, 2017 and 2018 respectively





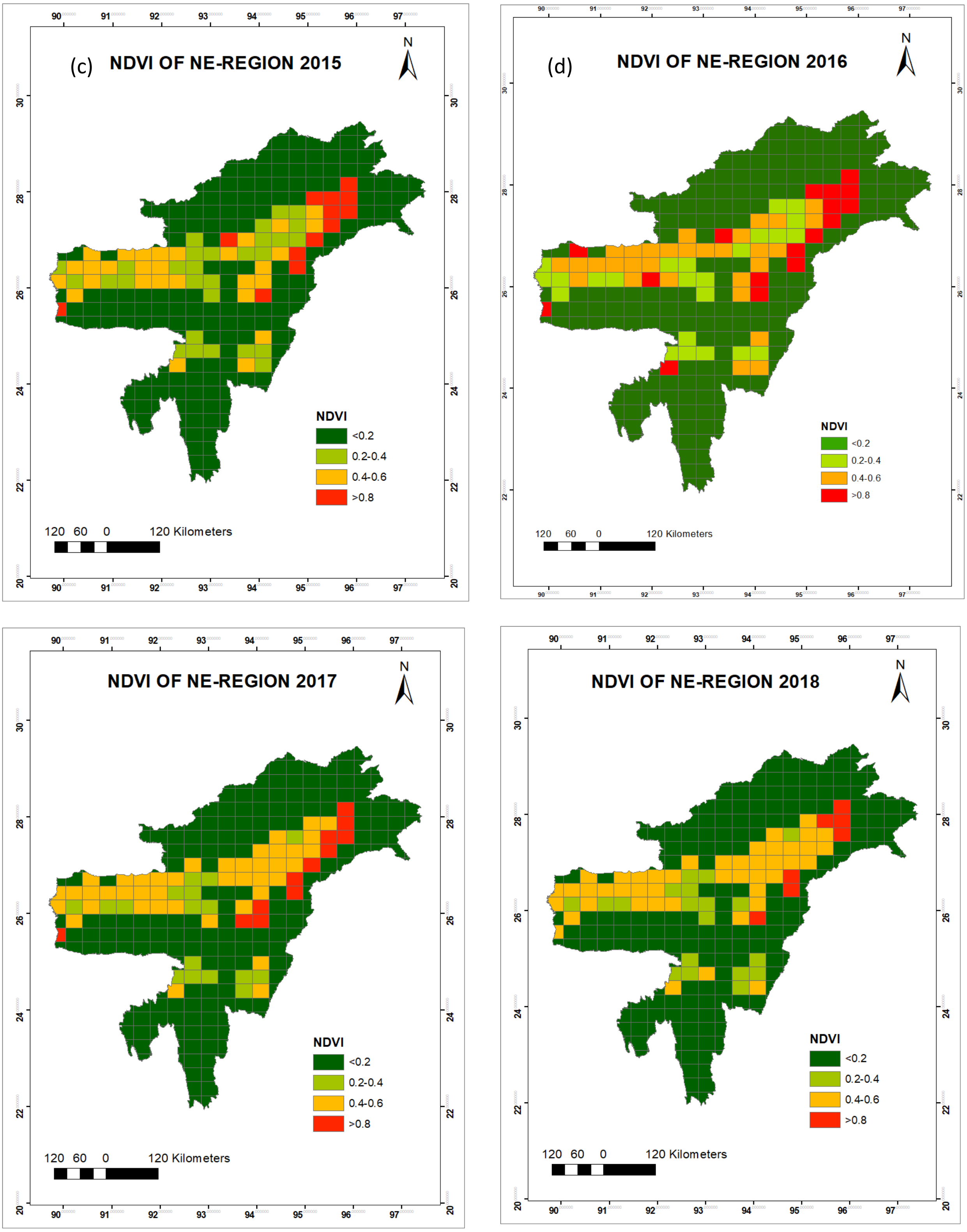
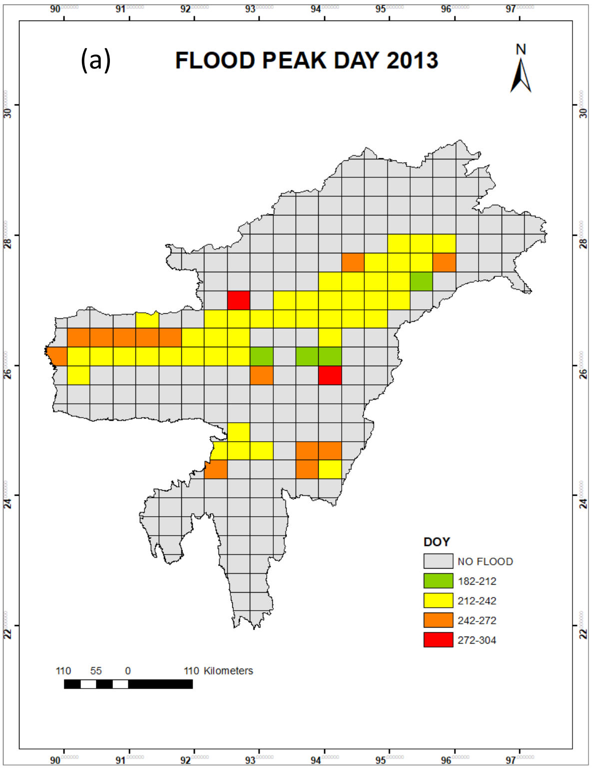
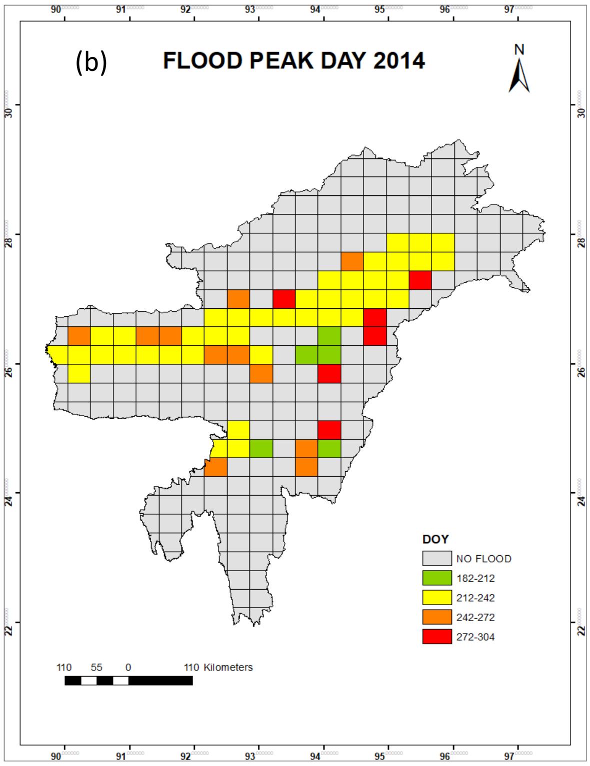


Image 3.2 NDVI (a) 2013, (b) 2014, (c) 2015, (d) 2016, (e) 2017 and (f) 2018.

3.4 Peak Flood Day

Flood peak day is one of the important parameter to be consider in the study, because it tell us the day where the flood is most severe. The flood peak day is shown in image 3.3 Flood Peak Day (a) 2013, (b) 2014, (c) 2015, (d) 2016, (e) 2017 and (f) 2018. The days are indicate in day of year (DOY), for the Kharif season the day start from 182 and ends at 304 day. The Brahmaputra valleys shows the maximum flood peak between 212-242 days for the year 2013, 2014, 2015 & 2018 but in the year 2016 & 2017 the flood peak is showing between 182-212 days. In Barak valley the year 2013, 2014, 2016 & 2018 shows flood peak days between 212-242 days but for the year 2015 and 2017 the flood peak day occurred between 242-272 days. In Manipur valley the maximum flood peak point occurred between the days 242-272. The year 2015, 2016 & 2018 shows that flood occurred in end of the season i.e. between the days 272-304.





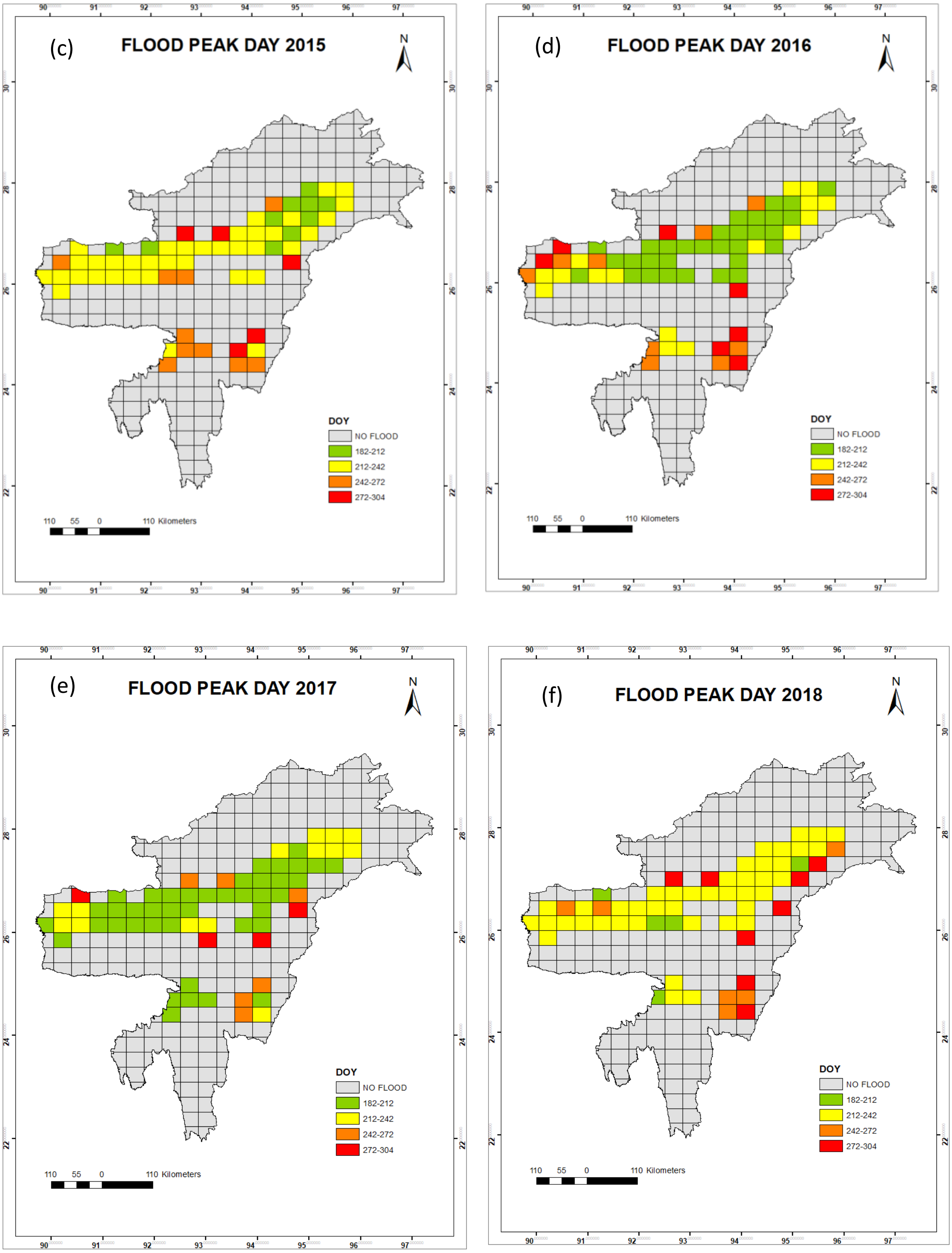
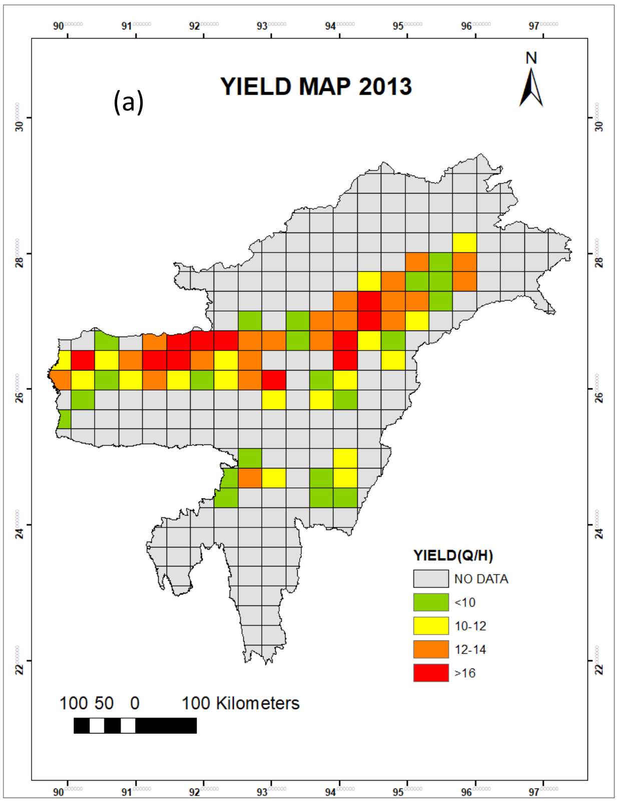
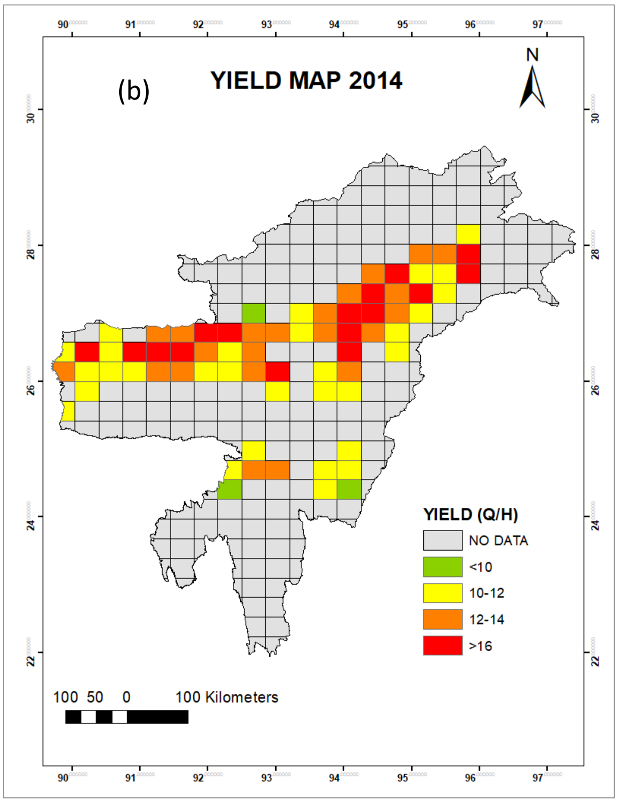


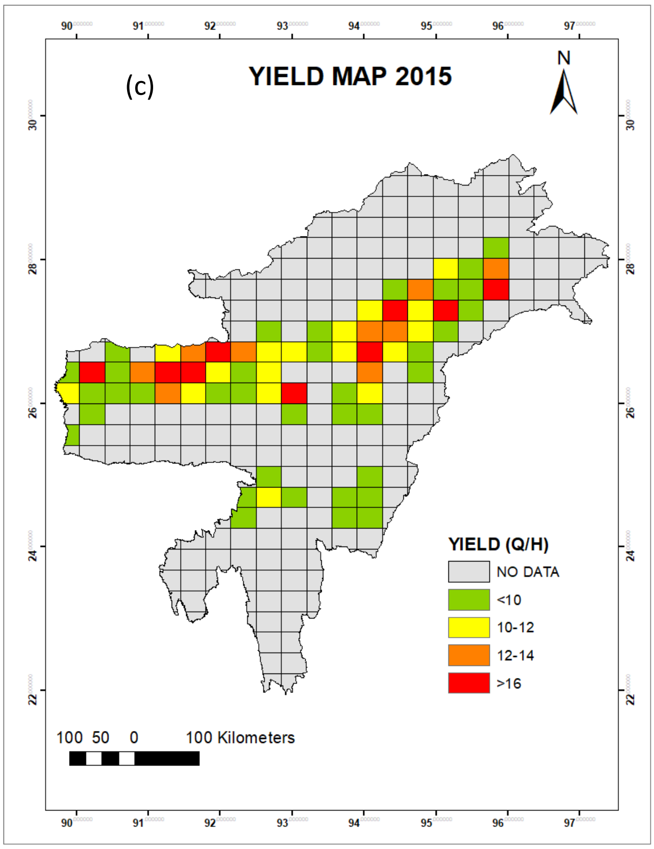
Image 3.3 Flood Peak Day (a) 2013, (b) 2014, (c) 2015, (d) 2016, (e) 2017 and (f) 2018.

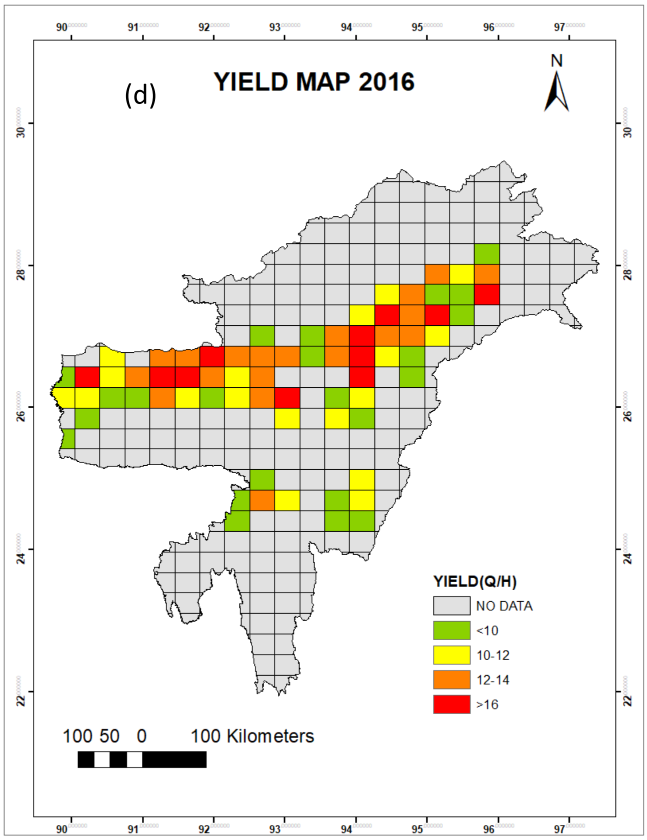
3.5 Yield Analysis

Crop yield can be described as the measurement of the amount of a crop that was harvested per unit of land area, it can also term as the actual seed generated from the plant. Crop yield plays a very important role in economy particularly in developing countries, as farming is the most important source of income. The grid level yield is shown in the image 3.5. It has been found that the yield had decreased drastically in the year 2014 and 2015, the main reason was due to flood and followed by drought year particularly in north-eastern region.









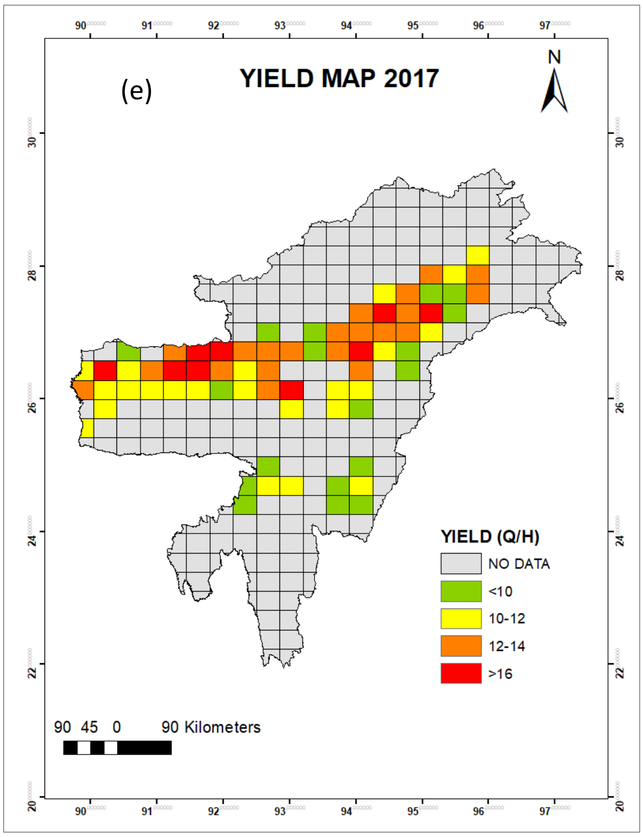


Image 3.5 Yield Map (a) 2013, (b) 2014, (c) 2015, (d) 2016, and (e) 2017.

4. Conclusion

The research conducted in this paper concluded that the 16 days' composite MODIS NDVI product has significant relationship with the yield data. The median NDVI for the entire year is compare against the 16 days' composite NDVI for the same area shows the impact on NDVI due to flood i.e. during the flood month of the year 2015. MODIS NRT 14 days' composite is useful in detecting flood percentage as well as peak flood days which is a good parameter for yield loss estimation. The linear regression analysis between NDVI and yield shows the high positive relationship between them R^2 =0.809. But for the year 2017 shows less relationship this is due to lack of availability of yield data for all the selected districts. Flood trend analysis indicates that most of the flood is concentrated in the Brahmaputra Valleys, Barak Valleys and Manipur Valleys. Even though the year 2014 and 2015 is declared as flood year for the whole country, one of the worst flood happen in Manipur Valleys, later on followed by drought. The highest value of NDVI is seen in the region where there are less flood and vice-versa. The yield trend also shows the negative rate of change for the year 2014 and 2015.

References

Shrestha R, Di Liping, Yu Eugene G, Kang L, Yuan-zheng S and Yu-qi B. 2017. Regression model to estimate flood impact on corn yield using MODIS NDVI and USDA cropland data layer, Journal of Integrative Agriculture 2017, 16(2): 398–407.

Gumma M K, Thenkabail P S, Maunahan A, Islam S and Nelson A. 2014. Mapping seasonal rice cropland extent and area in the high cropping intensity environment of Bangladesh using MODIS 500 m data for the year 2010, ISPRS Journal of Photogrammetry and Remote Sensing 91 (2014) 98–113.

Shrestha R, Di Liping, Yu Genong, Shao Y, Kang L and Zhang B. 2013. Detection of Flood and Its Impact on Crops Using NDVI - Corn Case, Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics)

Kwak Y and Iwami Y. 2014. Nationwide flood inundation mapping in Bangladesh by using modified land surface water index, ASPRS 2014 Annual Conference Louisville, Kentucky

Kwak Y, Park J and Fukami K. 2014. Near Real-Time Flood Volume Estimation from MODIS Time-Series Imagery in the Indus River Basin, IEEE Journal of selected topics in applied earth observations and remote sensing, vol. 7, no. 2.

Kwak Y, Shrestha B B, Yorozuya A and Sawano H. 2015. Rapid Damage Assessment of Rice Crop AfterLarge-Scale Flood in the Cambodian Floodplain Using Temporal Spatial Data, IEEE journal of selected topics in applied earth observations and remote sensing, vol. 8, no. 7.

Kwak Y, Park J, Arifuzzaman B, Iwami Y, AmirulMd and Kondoh A. 2016. Rapid exposure assessment of nationwide river flood for disaster risk reduction, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B8, 2016 XXIII ISPRS Congress.

Di Liping, Yu Eugene G, Kang L, Shrestha R and Yu-qi B. 2017. RF-CLASS: A remote-sensing-based flood crop loss assessment cyber-service system for supporting crop statistics and insurance decision-making, Journal of Integrative Agriculture 2017, 16(2): 408–423.

Yu Genong, Di Liping, Zhang Bei, Shao Y, Shrestha R and Kang L.2018. Remote-Sensing-Based Flood Damage Estimation Using Crop Condition Profiles, IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON).

Shrestha R, Di Liping, Yu Eugene G, Kang L, Li Lin, Rahman Md. S, Deng M and Yang Z. 2016. Regression based Corn Yield Assessment using MODIS Based Daily NDVI in Iowa State, The Fifth International Conference on Agro-Geoinformatics.

Del Ninno C, Dorosh P A, Smith L C. 2003. Public policy, markets and household coping strategies in Bangladesh: Avoiding a food security crisis following the 1998 floods. World Development, 31, 1221–1238.

Quarmby N A, Milnes M, Hindle T L, Silleos N. 1993. The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. International Journal of Remote Sensing, 14, 199–210.

Ren J, Chen Z, Zhou Q, Tang H. 2008. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. International Journal of Applied Earth Observation and Geoinformation, 10, 403–413.

Shrestha R, Di L P. 2013. Land/water detection and delineation with Landsat data using Matlab/ENVI. In: Proceeding of 2013 Second International Conference on AgroGeoinformatics. IEEEE, USA. pp. 211–214.

Yu G, Di L, Zhang B, Shao Y, Shrestha R, Kang L. 2013. Remote-sensing-based flood damage estimation using crop condition profiles. In: Proceeding of 2013 Second International Conference on Agro-Geoinformatics. IEEEE, USA. pp. 205–210.

Liu J, Pattey E, Jégo G. 2012. Assessment of vegetation indices for regional crop green LAI estimation from Landsat images over multiple growing seasons. Remote Sensing of Environment, 123, 347–358.

Kang L J, Di L P, Deng M D, Shao Y Z, Yu G N, Shrestha R. 2014. Use of geographically weighted regression model for exploring spatial patterns and local factors behind NDVIprecipitation correlation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 4530–4538.

Johnson D M. 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sensing of Environment, 141, 116–128.

Frazier P S, Page K J. 2000. Water body detection and delineation with Landsat TM data. Photogrammetric Engineering and Remote Sensing, 66, 1461–1468.

Tholey N, Clandilloni S, De Fraipont P. 1997. The contribution of spaceborne SAR and optical data in monitoring flood events: Examples in northern and southern France. Hydrological Processes, 11, 1409–1413.

Smith L C. 1997. Satellite remote sensing of river inundation area, stage, and discharge: A review. Hydrological Processes, 11, 1427–1439.

Appendix

Python code to unzip the files

```
import glob
importzipfile

for filename in glob.iglob('F:/PALMOL/MODIS_NRT/MODIS_14/*/*/*.zip'):
print(filename)

zip_ref = zipfile.ZipFile(filename, 'r')

zip_ref.extractall(filename.split('\\')[0]+'/'+filename.split('\\')[1]+'/'+filename.split('\\')[2]+'/')
zip_ref.close()
```

Python code to move folder

```
Import glob

Import os

For filename in glob.iglob('E:/palmol/MODIS_NRT/MODIS_14/*/*/*.shp'):

    Print(filename)

    Os.rename(filename,'E:/palmol/MODIS_NRT/MODIS_14/MOD_14/'+filename.split
('\\')[3])
```

Python code for flood peak

```
Import gdal

Import numpy as np

    ds = gdal.Open("D:/Palmol/project/nrt/2014/stack_nrt_14.tif")

data = ds.ReadAsArray()

#d=np.zeros((121))

max_flood_index = np.argmax(data, axis=0)

# create the output image

driver = ds.GetDriver()

#print driver

outDs = driver.Create("D:/Palmol/project/nrt/2014/out_14.tif", 4552, 4552, 1,
gdal.GDT_Float32)

outBand = outDs.GetRasterBand(1)

#outData = numpy.zeros(colors)
```

```
        # write the data
        outBand.WriteArray(max_flood_index)

        # georeference the image and set the projection
        outDs.SetGeoTransform(ds.GetGeoTransform())
        outDs.SetProjection(ds.GetProjection())
        deloutDs

import numpy as np
import matplotlib.pyplot as plt
from copulalib.copulalib import Copula
plt.style.use('ggplot')

def generateData():
    global x,y
    x = np.random.normal(size=250)
    y = 2.5*x + np.random.normal(size=250)

# Data and histograms
def plotData():
    global x,y
    fig = plt.figure()
    fig.add_subplot(2,2,1)
    plt.hist(x,bins=20,color='green',alpha=0.8,align='mid')
    plt.title('X variable distribution')
    fig.add_subplot(2,2,3)
    plt.scatter(x,y,marker="o",alpha=0.8)
    fig.add_subplot(2,2,4)
    plt.title('Joint X,Y')
```

```
    plt.hist(y,bins=20,orientation='horizontal',color='red',alpha=0.8,align='mid')
    plt.title('Y variable distribution')
    plt.show()

def generateCopulas():
    global x,y
    fig = plt.figure()

    frank = Copula(x,y,family='frank')
    uf,vf = frank.generate_uv(1000)
    fig.add_subplot(2,2,1)
    plt.scatter(uf,vf,marker='.',color='blue')
    plt.ylim(0,1)
    plt.xlim(0,1)
    plt.title('Frank copula')

    clayton = Copula(x,y,family='clayton')
    uc,vc = clayton.generate_uv(1000)
    fig.add_subplot(2,2,2)
    plt.scatter(uc,vc,marker='.',color='red')
    plt.ylim(0,1)
    plt.xlim(0,1)
    plt.title('Clayton copula')

    gumbel = Copula(x,y,family='gumbel')
    ug,vg = gumbel.generate_uv(1000)
    fig.add_subplot(2,2,3)
    plt.scatter(ug,vg,marker='.',color='green')
    plt.ylim(0,1)
    plt.xlim(0,1)
    plt.title('Gumbel copula')
```

```
    plt.show()

#-------------------------------------------------------------------------------
# Run the functions


generateData()
plotData()
generateCopulas()
```