

ITD Report: Remote Sensing Project: Chennai Floods 2015

Indu Ilanchezian (IMT2014024)
Tummala Armitha (IMT2014060)
Nikunj Gupta (IMT2014037)

May 6, 2018

Contents

0.1	Introduction	2
0.2	Data	2
0.2.1	Radar Data	2
0.2.2	Sentinel-1 Data Overview	3
0.2.3	Data Format: Naming Convention	5
0.2.4	Sentinel-1 Toolbox S1TBX	6
0.2.5	Data Used	6
0.3	Methodology	8
0.3.1	Multi-looking	10
0.3.2	Calibration	10
0.3.3	Terrain Correction	10
0.3.4	Cropping	10
0.3.5	RGB Composite	10
0.3.6	Mask Extraction for classes	11
0.3.7	Classification	12
0.4	Results	13
0.4.1	Classified maps	13
0.4.2	Accuracy, Precision, Recall and Kappa Score	16
0.4.3	OOB and importance of Bands (RF)	18
0.4.4	Distribution of Pixels	18
0.4.5	User's Accuracy	19
0.4.6	Producer's Accuracy	21
0.5	Conclusion	22
0.6	References	22

0.1 Introduction

Land Use Land Change (LULC) analysis is one of the primary applications addressed using geo-spatial data analysis. A broad range of problems fall under this category. Some examples include flood monitoring, landslide monitoring, urban growth patterns, change in vegetation patterns and forest cover analysis. In this report, we describe methods for analyzing the impact of flood on an urban area using remote sensing data.

High and medium resolution radar data are made available by sources such as the Sentinel-1 satellite mission. Radar data is useful for flood monitoring since the radar signals (microwaves) can penetrate through clouds and hence provide a clear view of the scene, unlike optical images where the view is obstructed by cloud cover.

India being predominantly a tropical wet/humid region experience 3-4 months of monsoon (from July-September) in most parts. The southern regions of India (parts of Tamil Nadu and Andhra Pradesh) witness heavy rainfalls in the post monsoon period (October-November) [1]. Given the vast extent of the country and the long duration of monsoon, floods are a frequent occurrence in the country. For efficient disaster management, it is essential to have methods in place to assess the damage caused by floods.

In this report, we analyze the damages and impacts of one such flood event which occurred in Chennai in the year 2015 using remote sensing data. For our analysis, we use the radar data from Sentinel-1 satellite mission. We perform image pre-processing such as multi-looking, calibration and geo-referencing (terrain correction) followed by feature extraction and classification using machine learning. The rest of this report discusses our methods in detail.

0.2 Data

0.2.1 Radar Data

Sentinel-1 is the first of the Copernicus Programme satellite mission conducted by the European Space Agency. This space mission is composed of two satellites, Sentinel-1A and Sentinel-1B, that carry a C-band Synthetic-Aperture Radar (SAR) instrument which provides a collection of data in all-weather, day or night [2].

Sentinel-1 data is fundamentally different from Landsat data. Landsat data is optical and Sentinel-1 is radar. The grey levels of the scene are related to the relative strength of the microwave energy back-scattered by the landscape elements. Different surface features exhibit different scattering characteristics:

- Urban areas: very strong back-scatter
- Forest: medium back-scatter
- Calm water: smooth surface, low back-scatter
- Rough sea: increased back-scatter due to wind and current effects

0.2.2 Sentinel-1 Data Overview

The SENTINEL-1 Synthetic Aperture Radar (SAR) instrument may acquire data in four exclusive modes:

- Stripmap (SM): Stripmap (SM) mode acquires data with an 80 km swath at slightly better than 5 m by 5 m spatial resolution (single look). The ground swath is illuminated by a continuous sequence of pulses while the antenna beam is pointing to a fixed azimuth angle and an approximately fixed off-nadir angle (this is subject to small variations because of roll steering). SM images have continuous along track image quality at an approximately constant incidence angle [3].
- Interferometric Wide swath (IW): The Interferometric Wide (IW) swath mode is the main acquisition mode over land and satisfies the majority of service requirements. It acquires data with a 250 km swath at 5 m by 20 m spatial resolution (single look). IW mode captures three sub-swaths using Terrain Observation with Progressive Scans SAR (TOPSAR). With the TOPSAR technique, in addition to steering the beam in range as in ScanSAR, the beam is also electronically steered from backward to forward in the azimuth direction for each burst, avoiding scalloping and resulting in homogeneous image quality throughout the swath [3].
- Extra Wide swath (EW): Similar to the IW mode, the Extra Wide (EW) swath mode employs the TOPSAR technique to acquire data over a wider area than for IW mode using five sub-swaths. EW mode acquires data over a 400 km swath at 20 m by 40 m spatial resolution [3].
- Wave (WV): Sentinel-1 Wave mode is similar to ERS and Evnisat wave mode imaging but with improved spatial resolution, larger vignettes and a 'leap frog' acquisition pattern as illustrated in the figure below. WV acquisitions consist of several vignettes exclusively in either VV or HH polarisation, with each vignette processed as a separate image. WV mode products can contain any number of vignettes, potentially amounting to an entire data-take. Each vignette is contained in an independent image within the product [3].

Sentinel has the following Levels of data:

- Raw Level-0 data.
- Processed Level-1 Single Look Complex (SLC) data comprising complex imagery with amplitude and phase
- Ground Range Detected (GRD) Level-1 data with multi-looked intensity only
- Level-2 Ocean (OCN) data for retrieved geophysical parameters of the ocean

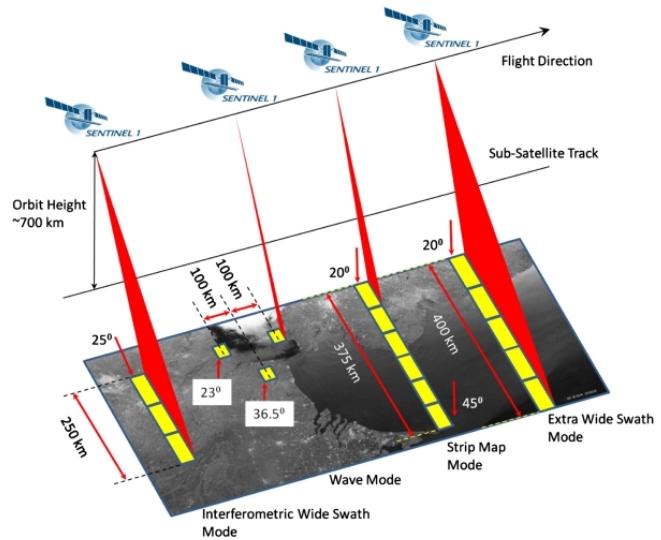


Figure 1: Modes of data acquisition

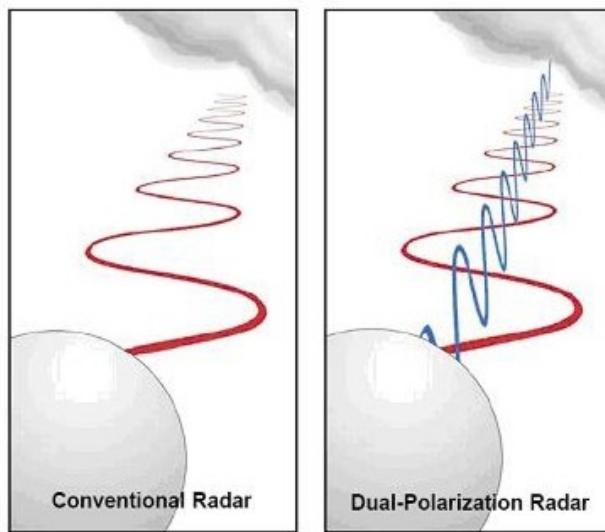


Figure 2: Difference between Polarisations

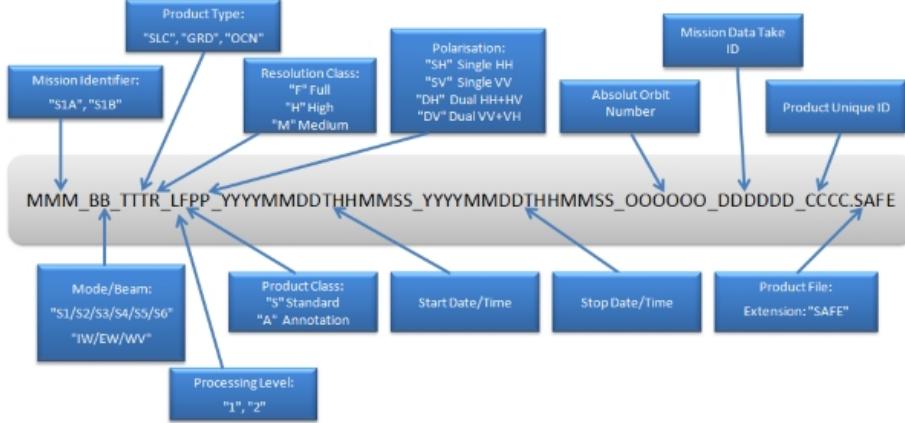


Figure 3: Nomenclature of data

Sentinel-1 has the following polarisations:

- Horizontal Transmit Horizontal Receive (HH)
- Horizontal Transmit Vertical Receive (HV)
- Vertical Transmit Vertical Receive (VV)
- Vertical Transmit Horizontal Receive (VH)

0.2.3 Data Format: Naming Convention

The top-level SENTINEL-1 product folder name is composed of upper-case alphanumeric characters separated by an underscore. The Mission Identifier (MMM) denotes the satellite and will be either S1A for the SENTINEL-1A instrument or S1B for the SENTINEL-1B instrument. The Mode/Beam (BB) identifies the S1-S6 beams for SM products and IW, EW and WV for products from the respective modes. Product Type (TTT) can be RAW, SLC, GRD or OCN. Resolution Class (R) can be F (Full resolution), H (High resolution), M (Medium resolution), or underscore (not applicable to the current product type). Resolution Class is used for SLC and OCN only. The Processing Level (L) can be 0, 1 or 2. The Product Class can be Standard (S) or Annotation (A). Annotation products are only used internally by the PDGS and are not distributed. Polarisation (PP) can be one of: SH, SV, DH, DV. The product start and stop date and times are shown as 14 digits representing the date and time, separated by the character "T". The absolute orbit number at product start time (OOOOOO) will be in the range 000001-999999. The mission data-take identifier (DDDDDDD) will be in the range 000001-FFFFFF. The product unique identifier (CCCC) is a hexadecimal string generated by computing CRC-16 on the manifest file using CRC-CCITT. The folder extension is always "SAFE" [3].

0.2.4 Sentinel-1 Toolbox S1TBX

The Sentinel-1 Toolbox (S1TBX) consists of a collection of processing tools, data product readers and writers and a display and analysis application to support the large archive of data from ESA SAR missions.

0.2.5 Data Used

We used data of 3 dates:

- 31/10/2015: before crisis
- 6/11/2015: on the day of crisis
- 24/11/2015: after crisis

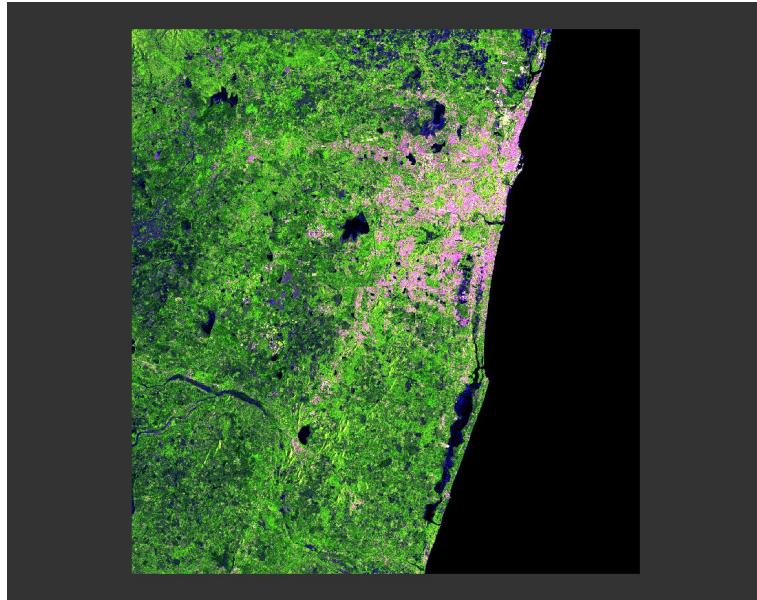


Figure 4: RGB image for 31/10/2015

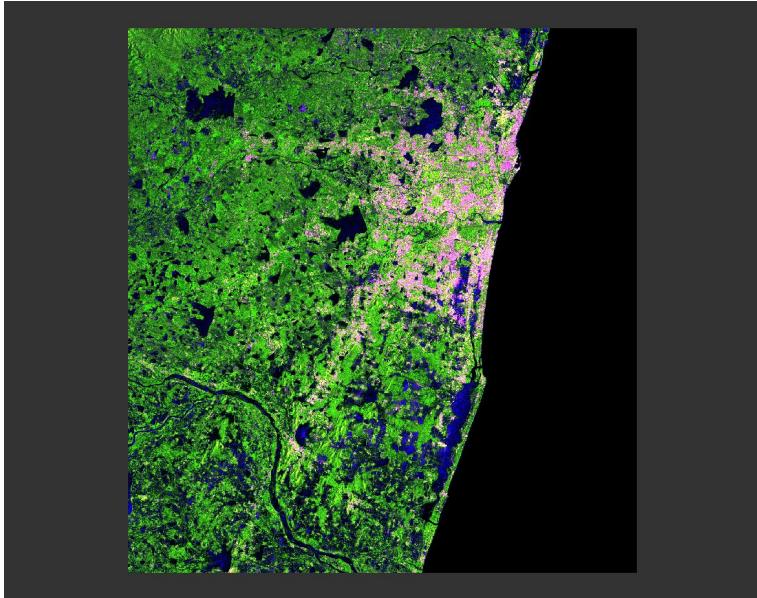


Figure 5: RGB image for 6/11/2015

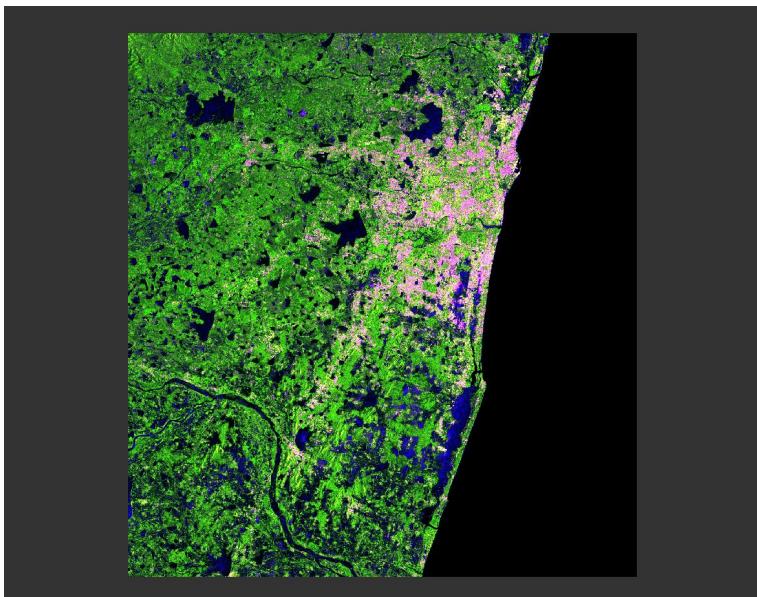


Figure 6: RGB image for 24/11/2015

0.3 Methodology

The flowchart below depicts the procedure we have followed for estimating damage occurred in Chennai in 2015 due to floods

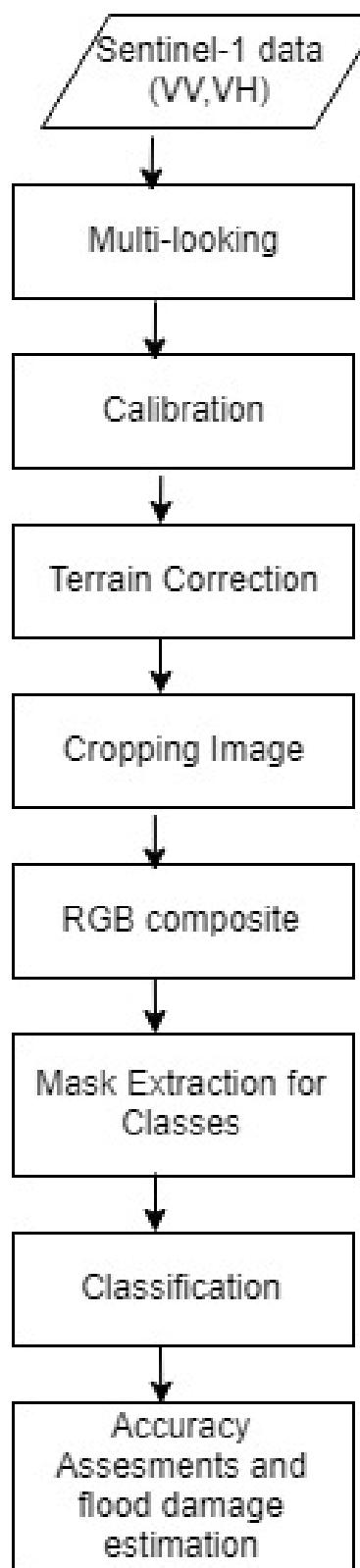


Figure 7: Flowchart

0.3.1 Multi-looking

It reduces the speckle or noise. There are different types of multi-looking algorithms. The basic multi-looking algorithm consists of a window and it reduces the speckle by applying the mathematical calculation on the pixels under the window and replacing the centre of the window with the new value. The users can define $N \times N$ window over which the averaging occurs[4]. In our project, we have used 3×3 window for reducing the noise by averaging the adjacent pixels.

0.3.2 Calibration

It is essential to compare two images and also to remove measurement errors. In our project, we have used Radiometric Calibration. “*Radiometric Calibration refers to the ability to convert the digital numbers recorded by satellite imaging systems into physical units.*” [5]. The results of calibration are the Sigma nought, Beta nought and Gamma nought of VV and VH image.

Sigma nought is the scattering coefficient, also the conventional measure of the strength of radar signals reflected by a distributed scatterer. It is measured in dB. It is a normalized dimensionless number, comparing the strength observed to that expected from an area of one square meter. Sigma nought is defined with respect to the nominally horizontal plane, and in general has a significant variation with incidence angle, wavelength, and polarization, as well as with properties of the scattering surface itself [3].

Beta nought is the radar brightness coefficient. It is the reflectivity per unit area in slant range which is dimensionless [3].

0.3.3 Terrain Correction

To project the image onto the map system and to correct the distortions (layover and foreshortening) in the terrain [7].It eliminates the side looking geometry effects of radar images [6]. In our project, we have used Range Doppler Terrain Correction.

0.3.4 Cropping

We have cropped the area of interest (Chennai) for the project.

0.3.5 RGB Composite

We have produced a RGB composite using the sigma nought of VV image and sigma nought of VH image. The RGB composite is created by assigning sigma nought of VV polarization to red channel, sigma nought of VH polarization to green channel and the ratio of sigma nought VV to sigma nought VH to the blue channel.

0.3.6 Mask Extraction for classes

From the RGB composite, which is similar to False Color Composite for optical images, we can identify the various features on the Earth's surface. We have extracted masks for four classes (water, urban, vegetation and open land). The masks are extracted for all classes for both training and testing.

0.3.7 Classification

Automated classification of water areas is required for estimating the flooded area and to measure the retraction of water post the flooding period. The accuracy of the estimation of the flood inundation area depends on the accuracy of the classifier.

For optical images, the different reflectance bands and the derived properties such as Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and so on are used as features for classification. In radar data the only available information are the intensities of the reflected waves. Since we use the dual polarized data in our methods, we have both VV and VH polarization intensities available. From these intensities, other features such as Sigma nought, Beta nought and Gamma nought values are derived, for each of the polarizations. Other approaches for classification using radar data include use of Gray Level Co-occurrence Matrix features for classification [8]. The GLCM computes image features such as contrast texture, intensity, energy. Any combination of these features can be used for classification. For our analysis, we restrict the features to the six features obtained as result of calibration, since these features can efficiently classify the pixels to the categories of our interest. In radar images, water can be clearly distinguished from any other class (such as barren land, urban areas and vegetation) since water completely absorbs the radar signal and appears completely black. Thus, the binary classification of water and non-water areas is a relatively simple task and there is no need for computation of indices like NDWI (as in optical data) for deriving a water mask. We perform classification of pixels into four classes: water, urban, vegetation and barren lands as this allows us to answer a wider range of questions.

With the above four-class classification, we can estimate the following:

1. the extent of flooding area
2. Built up area which was inundated in floods
3. extent of vegetation affected during floods
4. filling up of barren lands (dry lake/river areas) with water during floods

We assess the performance of three class of classification models:

1. Random Forest Classifier
2. Support Vector Machines
3. Two layer neural nets (also Multi-layer perceptrons, used interchangeably)with 100 hidden units in the first layer and 50 hidden units in the second.

0.4 Results

0.4.1 Classified maps

Random Forest (RF)

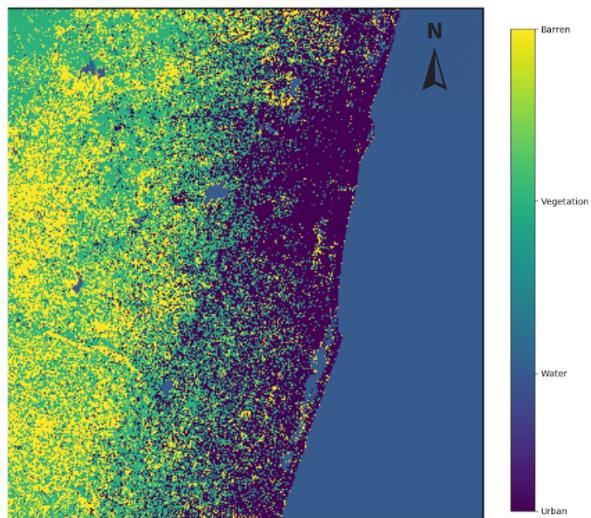


Figure 8: 31/10/2015

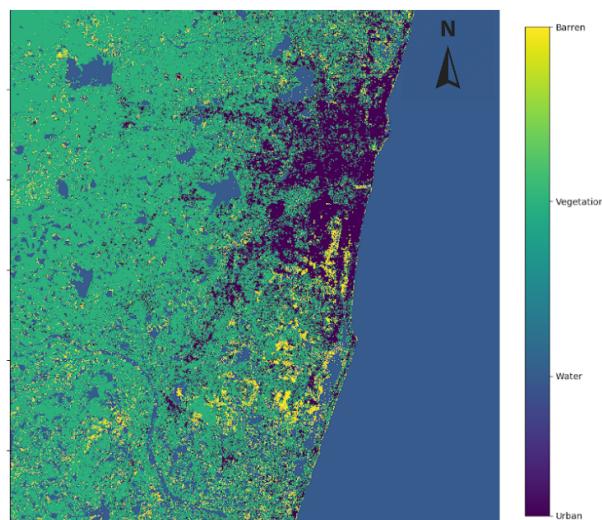


Figure 9: 06/11/2015

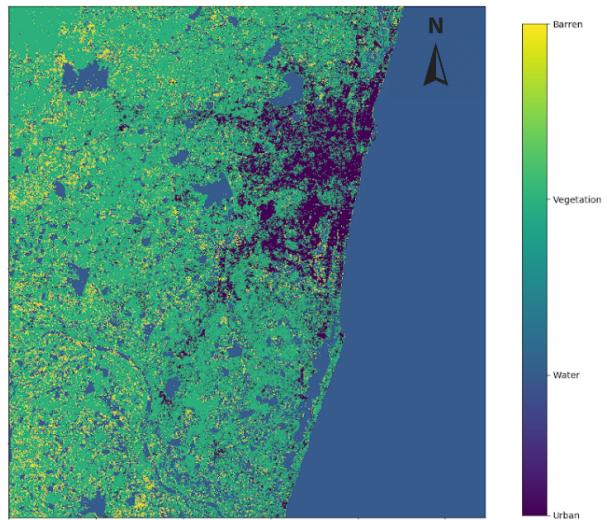


Figure 10: 24/11/2015

Two Layer Neural Nets/Multi Layer Perceptron (MLP)

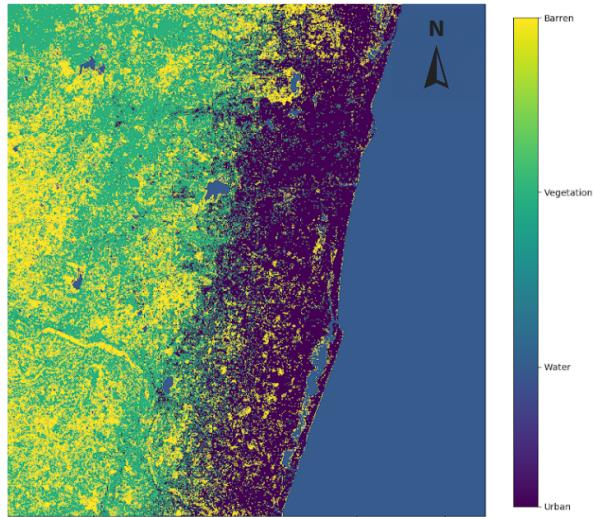


Figure 11: 31/10/2015

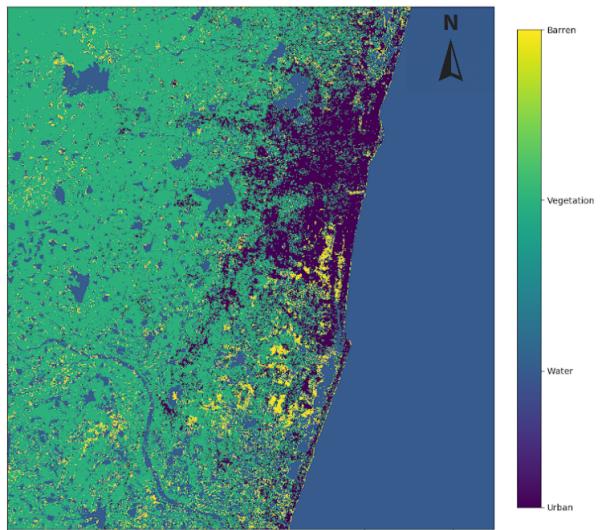


Figure 12: 06/11/2015

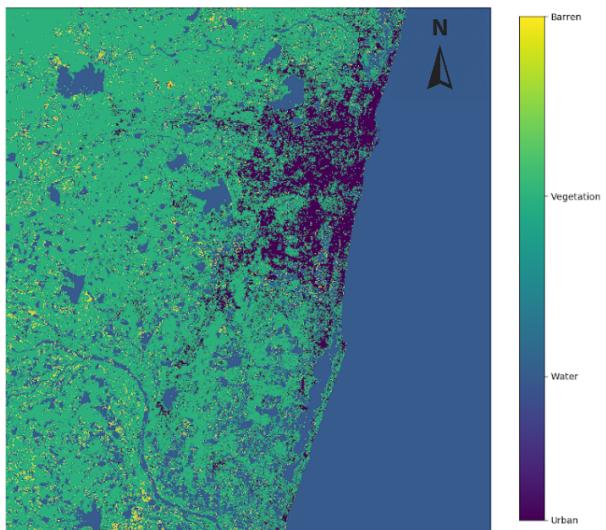


Figure 13: 24/11/2015

Support Vector Machine (SVM)

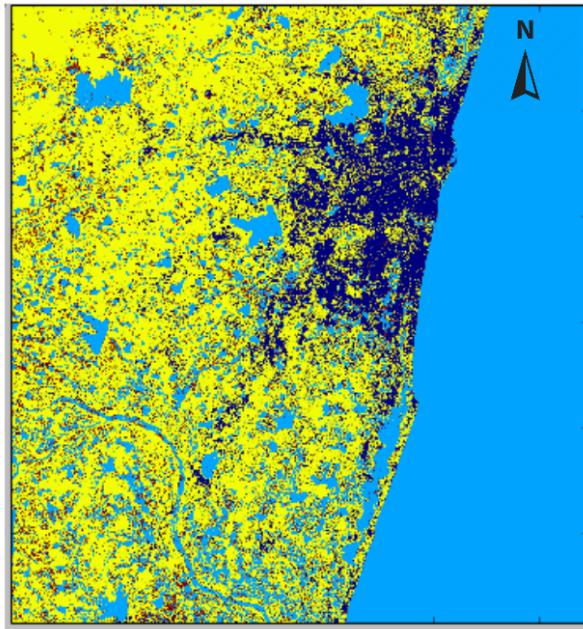


Figure 14: 24/11/2015

0.4.2 Accuracy, Precision, Recall and Kappa Score

	31-10-2015	24-11-2015	06-12-2015
Training Set	131699	122810	131699
Test Set	137946	144611	137946

Figure 15: Training data and Test data

	31-10-2015	24-11-2015	06-12-2015
SVM	0.9344	0.9064	0.9349
MLP	0.9178	0.9029	0.9503
RandomForest	0.9122	0.8972	0.9327

Figure 16: Accuracy assessment - overall accuracy

	31-10-2015	24-11-2015	06-12-2015
SVM	0.9344	0.9064	0.9384
MLP	0.9178	0.9029	0.9503
RandomForest	0.9122	0.8972	0.9327

Figure 17: Precision

	31-10-2015	24-11-2015	06-12-2015
SVM	0.9344	0.9064	0.9384
MLP	0.9178	0.9029	0.9503
RandomForest	0.9122	0.8972	0.9327

Figure 18: Recall

	31-10-2015	24-11-2015	06-12-2015
SVM	0.8689	0.8380	0.8970
MLP	0.8373	0.8331	0.9173
RandomForest	0.8279	0.8232	0.8879

Figure 19: Kappa score

0.4.3 OOB and importance of Bands (RF)

	31-10-2015	24-11-2015	06-12-2015
OOB	92.2041%	94.1495%	95.6522%
Band Importance	Band 0 importance: 0.108451731987 Band 1 importance: 0.146252843402 Band 2 importance: 0.127333094343 Band 3 importance: 0.172086444633 Band 4 importance: 0.191412233899 Band 5 importance: 0.254463651735	Band 0 importance: 0.0511213887536 Band 1 importance: 0.0729701476495 Band 2 importance: 0.0793171008437 Band 3 importance: 0.270990538248 Band 4 importance: 0.25841721172 Band 5 importance: 0.267183612785	Band 0 importance: 0.0666043633464 Band 1 importance: 0.0564836015953 Band 2 importance: 0.107757597493 Band 3 importance: 0.259653704732 Band 4 importance: 0.215748590113 Band 5 importance: 0.293752142721

Figure 20: OOB and importance of Bands (RF)

0.4.4 Distribution of Pixels

Water	2496504	Water	3366244	Water	3190314
Urban	1809231	Urban	892004	Urban	1204257
Vegetation	2273948	Vegetation	3290429	Vegetation	3288477
Barren	1411223	Barren	436767	Barren	302396

Difference in water pixels(before rains and during rains):869740
Difference in water pixels(during rains and after rains):175930

Difference in area (in sq. km) (before and during floods):86.9740
Difference in area (in sq. km) (during and after floods):17.5930

Figure 21: Distribution of pixels (RF)

Water	2496053	Water	3484529	Water	3224278
Urban	1805209	Urban	769140	Urban	1124064
Vegetation	2278672	Vegetation	3560815	Vegetation	3351629
Barren	1410972	Barren	170960	Barren	285473

Difference in water pixels(before rains and during rains): 988476
 Difference in water pixels(during rains and after rains): 260251

Difference in area (in sq. km) (before and during floods): 98.8476
 Difference in area (in sq. km) (during and after floods): 26.0251

Figure 22: Distribution of pixels (MLP)

Water	3317316
Urban	905471
Vegetation	341869
Barren	344008

Figure 23: Distribution of pixels (SVM)

0.4.5 User's Accuracy

	31-10-2015	24-11-2015	06-12-2015
SVM	0.8931	0.9899	0.9280
MLP	0.9005	0.9922	0.9320
RandomForest	0.8927	0.9899	0.9270

Figure 24: Water

	31-10-2015	24-11-2015	06-12-2015
SVM	0.9831	0.8850	0.9539
MLP	0.9651	0.8697	0.9558
RandomForest	0.9570	0.8734	0.9390

Figure 25: Urban

	31-10-2015	24-11-2015	06-12-2015
SVM	0.8324	0.9567	0.9292
MLP	0.8150	0.9774	0.9602
RandomForest	0.8176	0.9466	0.9370

Figure 26: Vegetation

	31-10-2015	24-11-2015	06-12-2015
SVM	0.8402	0.3508	0.7263
MLP	0.8194	0.2842	0.7636
RandomForest	0.8056	0.3790	0.7258

Figure 27: Open Lands

0.4.6 Producer's Accuracy

	31-10-2015	24-11-2015	06-12-2015
SVM	0.9031	0.9024	0.9706
MLP	0.9005	0.8764	0.9670
RandomForest	0.8763	0.8876	0.9670

Figure 28: Water

	31-10-2015	24-11-2015	06-12-2015
SVM	0.9522	0.9768	0.9563
MLP	0.9485	0.9892	0.9760
RandomForest	0.9567	0.9787	0.9609

Figure 29: Urban

	31-10-2015	24-11-2015	06-12-2015
SVM	0.9492	0.9139	0.9223
MLP	0.9038	0.7978	0.9295
RandomForest	0.8788	0.7998	0.9014

Figure 30: Vegetation

	31-10-2015	24-11-2015	06-12-2015
SVM	0.6825	0.6556	0.5800
MLP	0.6427	0.7576	0.5896
RandomForest	0.5877	0.5305	0.5648

Figure 31: Open Lands

0.5 Conclusion

Thus, we have used radar data to classify remote sensing data and estimate the extent of flooding in Chennai in the 2015 floods. Radar data gives us the advantage of cloud-free view of the Earth. With sigma nought, beta nought and gamma nought values for two different polarization, we are able to get classification accuracies of over 90% using standard classifiers: Random Forests, SVMs and two-layer Neural Nets (MLPs). Thus, radar data can be effectively used for flood damage assessment.

0.6 References

- [1] Kiladis, G. N., Sinha, S. K., 1991: ENSO, monsoon and droughts in India. In: Glantz, M. H., Kalz, R. W., Nicholls, N. Teleconnections Linking Worldwide Climate Anomalies — Scientific Basis and Social Impact. New York: Cambridge University Press, Chapter 14, pp. 431–458.
- [2] Sentinel-1. Retrieved from: https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-1/Introducing_Sentinel-1
- [3] Sentinel-1 SAR User Guide. Retrieved from: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/>
- [4] Radar Image Properties. Natural Resources Canada. Retrieved from: <http://www.nrcan.gc.ca/node/9299>
- [5] Radiometric Calibration. Retrieved from: <https://www.sdsstate.edu/radiometric-calibration>
- [6] http://www2.gi.alaska.edu/rgens/teaching/geos639/terrain_correction.pdf
- [7] <http://geoinformaticstutorial.blogspot.in/2012/03/terrain-correction-of-sar-images-part-1.html>
- [8] Shokr, M. E. (1991), Evaluation of second-order texture parameters for sea ice classification from radar images, *J. Geophys. Res.*, 96(C6), 10625–10640, doi: 10.1029/91JC00693.