GEOSPATIAL BIG DATA ANALYTICS AND SPATIAL DATA INFRASTRUCTURE IN DISASTER MANAGEMENT:
A CASE OF PLUVIAL FLOODING IN CHENNAI CITY

Thesis submitted in partial fulfilment of the requirements for the award of the degree of

Masters in Planning (Environment Planning)

By
Madhukar Kuchavaram
2018MEP007

Under the Guidance of
Prof. Dr. N. Sridharan

SCHOOL OF PLANNING AND ARCHITECTURE, BHOPAL
NEELBAD ROAD, BHAURI BHOPAL (MP)-462 030

June 2020

Declaration

I **Madhukar kuchavaram**, Scholar No. **2018MEP007** hereby declare that the thesis titled Geospatial Big Data Analytics and Spatial Data Infrastructure in Disaster Management: A Case of Pluvial Flooding in Chennai City, submitted by me in partial fulfilment for the award of Master of Planning (Environmental Planning), at School of Planning and Architecture, Bhopal, India, is a record of bonafide work carried out by me. The matter/result embodied in this thesis has not been submitted to any other University or Institute for the award of any degree or diploma.

K. Madhukar

Signature of the Student

Date: 10-07-2020

Certificate

This is to certify that the declaration of **Madhukar kuchavaram** is true to the best of my knowledge and that the student has worked under my guidance for one semester in preparing this thesis.

RECOMMENDED

N. Cdhalay 10/07/2020

Signature of the Guide

ACCEPTED

N. Cdhalay 10/07/20

Signature of the Head,

Department of Environmental Planning

Date: 10/07/2020

Place: Bhopal

DEDICATIONS

My parents

Mr. k.Mallesh Chary & Mrs. K.Anuradha

My sister

Ms. Manasa.k

My Friends

Saikrishna. K

Mahendar Reddy

ACKNOWLEDGMENT

Undertaking Master's thesis at School of Planning and Architecture, Bhopal has been a true life to change experience for me and it would not have been possible without the support and guidance that I received from many faculty members and people.

I would like to thank first and foremost to my thesis Guide Professor, **Dr. N Sridharan**, who has supported me with his complete interest throughout my thesis research patiently guided me through the past six months and supported me whenever support was needed. His constant inspiration when I was down solely with some weaknesses helped me in successfully completing this project.

I thank the Head of Department of Planning, **Professor Dr. N Sridharan**, other Faculty members of the department and External jury members for their valuable suggestion, precious time for the discussion and their support in completion of my thesis.

I thank **Dr. Rama Panday**, who is my Thesis Coordinator, for her constant reminders and for guiding me in topic discussions.

My deep appreciation goes to the members of Chennai metropolitan development authority, Tamil Nadu State Planning Development Society, Tamil Nadu Remote sensing Application Centre. Further credits to United States Geological Survey (USGS) & ISRO Bhuvan-India, for accessing the satellite images and GITHUB for accessing Python codes.

Finally, I offer my sincerest gratitude to my parents and my elder brother **K. Srikanth chary** for love and their support which helped me to cross all the hurdles coming in my successful education life.

SPECIAL THANKS

VENU MADHAV ANNAVARAPU

SAI PAVAN RELANGI

FAIZA JAMAL

Mr. AMIT SRIVASTAVA

Mr. JAGADES VIGNESVAR

Mr. BALAJI

Mr. SHEKHAR

Mr. JAYAPRAKESH GANDHI

VISHAL KUMAR

MANPREET VARMA

K. SAISREE REDDY

ABSTRACT

Social media generates the post- disaster scenario. There is no effective utilization of this data sets in disaster management (urban flooding), which is valuable and available in real time compared to conventional data sources. Though spatial data infrastructure (SDI) was proven successful in monitoring and supporting disaster resilience, had delays and data gaps in open mapping portals. These Crowdsourced data generated in social media can be integrated with SDIs to prepare near real time flood vulnerability assessment. The study reveals the importance of social media through geospatial Big data analytics to manage urban flooding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

This chapter Introduces the Importance of social media and data analytics data sets in Pluvial Flooding and Contribution of bigdata analytics in sustainable development goals for development and humanitarian action. Approaches of various national disaster management system are discussed where data analytics is utilized in real time monitoring of disasters. In contest need for study was established targeting the effective utilization of analytics and micro sensors in post disaster scenario.

………………………………………………………………………………………..

1.1 Background Study

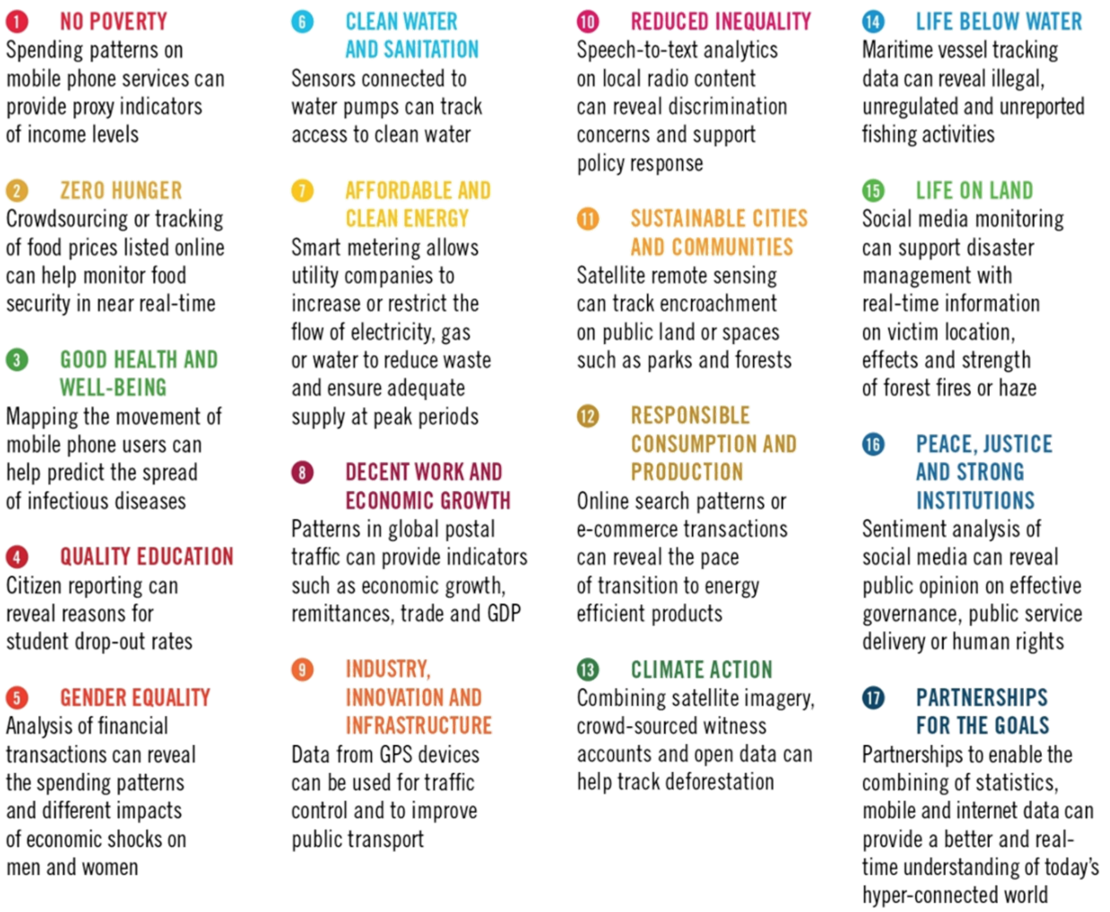
Natural disasters are a suddenly disrupted event with high magnitude which changes the pattern of life. Every year worldwide 160 million people are affected and around 90,000 people are killed in effect of natural disaster(WHO, 2012). These events cannot be completely minimized but the risk associated with this disaster can be reduced to some extent and most of the conventional disaster management systems lack the support in risk reduction (Rottach, 2004). Disaster management cycle comprises of two phases (Risk management, crisis management) and four stages (Mitigation, Preparedness, Response and recovery) with challenging situations. Effective utilization of modern analytics and volunteered geographical information can address these complex situations(Senaratne et al., 2017).

In today's world Crowdsourced data is a vital communication platform and is defined by its connectedness, interactivity and content which is generated by an individual. Twitter, Facebook, LinkedIn, blogger and google etc. are some of the most influential social networking tools generating over 2.5 Exabytes of information every single day. Almost 90% of the worlds data has been generated in the last few years(*Reliance on Social Media in Today's Society - Digital Marketing Blog*, n.d.).

1.2 UNDP SDG's 2015 and Big Data Analytics

There is still inadequacy in critical data for global, regional and national development policymaking. Governments fall short of access to data on entire populations, especially the poorest and most marginalised. This is applicable even at the international level who are unable to support the most vulnerable and marginalized citizens, lest they invest in the contemporary data gathering techniques. Big data, when applied affectively, leads to alterations in the decision-making process that conventionally depend on traditional statistics. A more receptive mechanism needs to be set up that enables the government to handle information and give a prompt response, considering its high frequency. (*Big Data for Sustainable Development* | *United Nations*, n.d.)

Figure 1 Data Analytics contribution in Sustainable Development



Source :(Big Data for Sustainable Development | *United Nations*, n.d.)

It has, however, been found that Big data has been playing a very non-crucial role in policy making and is majorly used for agenda setting and policy making because of ineffective functioning of the ecosystem and absence of standards and frameworks. Big data has only recently been welcomed by the national governments and other policy makers for policy making, where data is very crucial.

Big data analysis has become usual for the private sector in the recent times with consumer profiling, personalised services, and predictive analysis being utilised in marketing, advertising and management. A peek into people's welfare could be achieved with these techniques and the like, including satellite data. Decision-making becomes quick, efficient and rational with the use of newer technologies and progress, as measured by Sustainable Development Goals (SDGs), is comprehensive and just.(*Big Data for Sustainable Development* | *United Nations*, n.d.)

1.2.1 Goal no.15: Life on Land and Big data

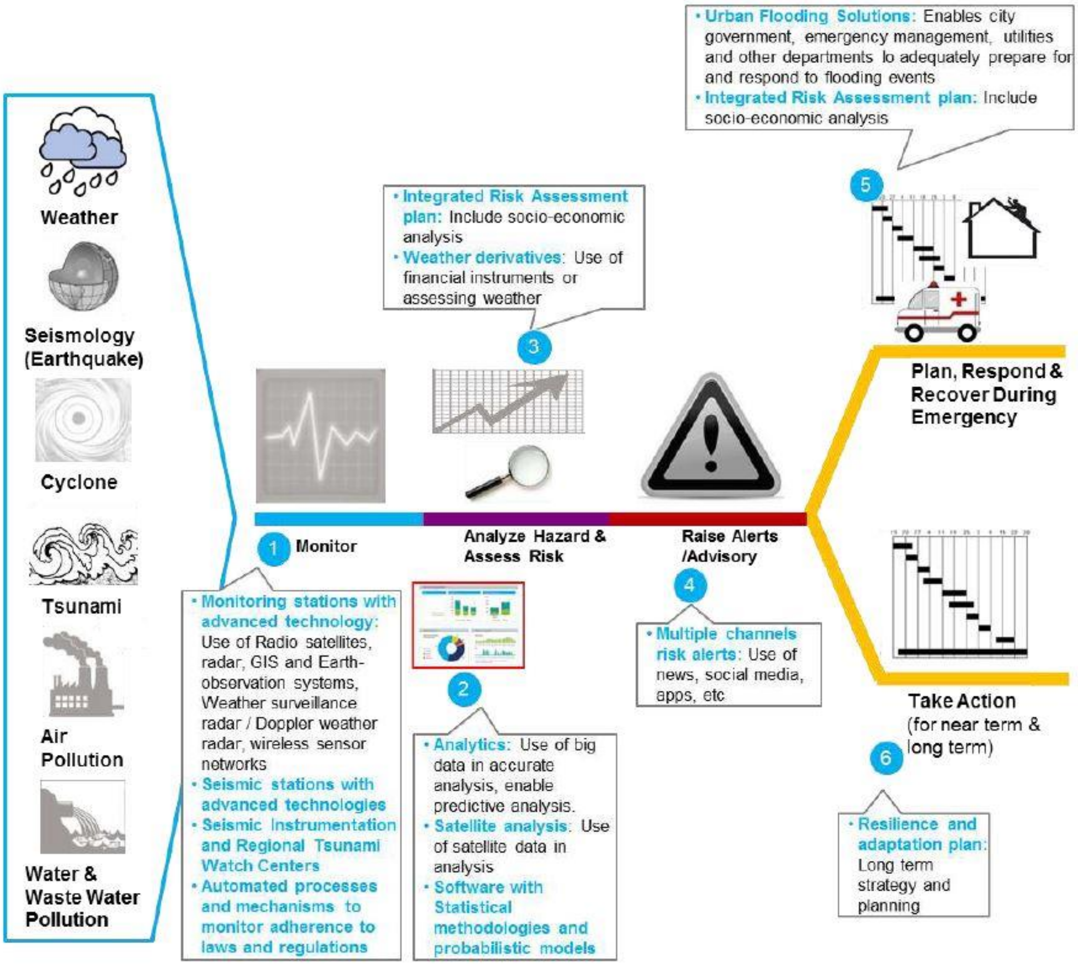
Figure 2 Life on Land (Goal -15)

Globally accurate data is delivered timely by satellite-based monitoring and local sensors (mobile phones) give instantaneous precision. Both short term and long term trends are taken care of by Big data and it can easily analyse the impact of extreme weather conditions and movement of displaced population so as to pilot disaster mitigation enabling supply of targeted relief.(UN, 2015) There is an ongoing research that uses mobile phone data to understand human movement post calamities. (Lokanathan et al., 2017)

1.3 Smart environment management – NASSCOM

There has been a huge fluctuation in India's temperature and rainfall pattern, consequences of which being the Uttaranchal floods and Cyclone Phailin that resulted in loss of economy. A major loophole being the ineffective communication between central, national and local bodies. There is a need of city-specific, current and exhaustive database with details on climate change consequences so as to plan spatial adaptation strategies .(NASSCOM, 2011).

Figure 3 Data Analytics in Smart climate change



Source :(NASSCOM, 2011)

1.3.1 components of the ecosystem

1. Monitoring: The base to even city-functioning is formed by City governance that includes daily works of the local government and collaboration with other organisations to serve all citizens.

2. Analyse hazard and assess risk: This includes risk evaluation from the hazard and action plan for predicament.

3. Raise alerts/advisory: Several communication modes give alerts via conventional media and social media.

4. Respond to emergency: Response to emergency includes medical, fire and police services along with disaster management services.

5. Take long-term action: Long-term action plan for precaution is produced after thorough investigation and analysis of the impact.

1.3.2 Issues and challenges

Table 1 Issues and challenges in climate change and pollution mitigation

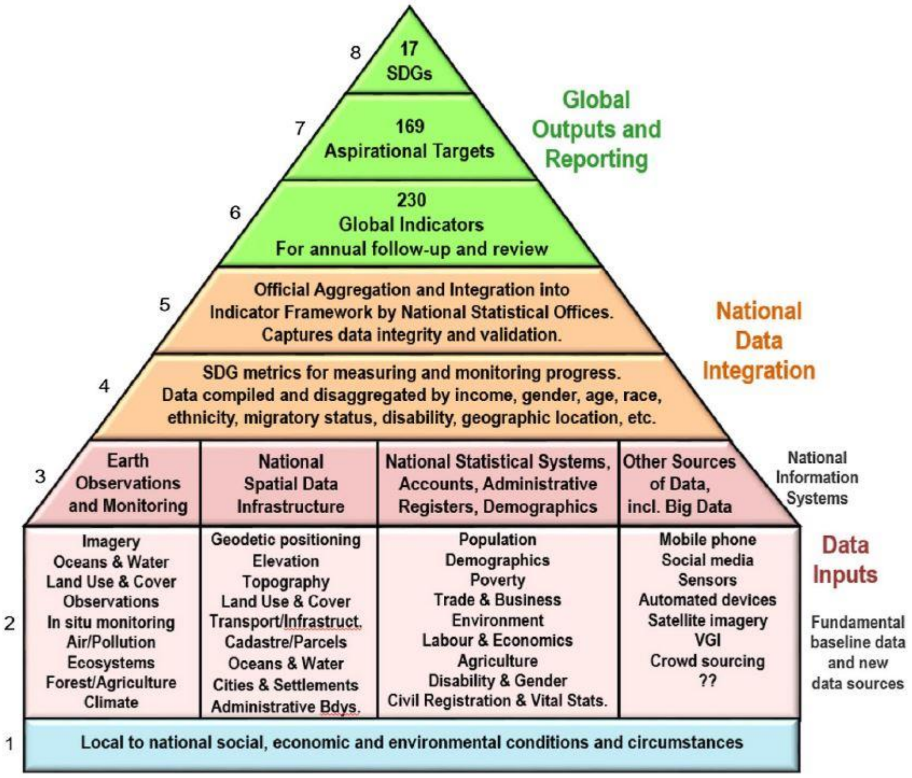
ISSUE	Description
a) Monitoring	Various organizations are involved in monitoring and hence, there is no central source of monitored data hampering lack of integrated monitoring.
b) Lack of coordination amongst various organizations	Lack of coordination amongst various organizations.
c) Lack of coordination amongst various organizations	Missed alerts leading to inefficient response and longer recovery.
d) Emergency response	Multiple agencies are involved in emergency response. Data unavailability and inadequate capacity for live monitoring of situation hinder the response.
e) Long-term action	Air/water pollution data is only used to get a view on pollution but no real-time actions are taken to reduce the pollution.

Source : (NASSCOM, 2011)

1.4 Sustainable development and Spatial Data Infrastructure

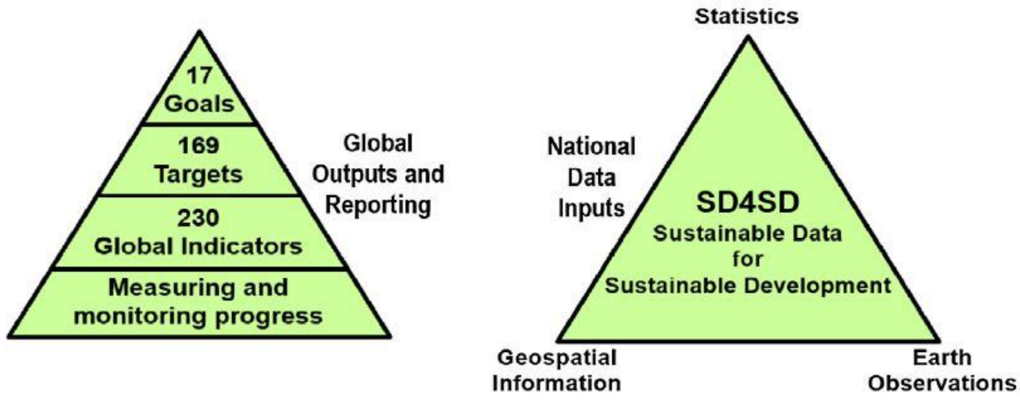
The 2030 Agenda for Sustainable Development guides nations collectively to run and alter the social, economic, and environmental dimensions of people and the planet over the next decade and a half in form of a global policy. In order to achieve sustainable development, all nations and the global policy community have to overcome various geographic development challenges. Several issues affecting sustainable development can be analysed and modelled spatially that gives the comprehensive framework important for global partnership, consensus and rational decision-making. Inspire of the progress in geospatial information technologies; awareness, understanding and uptake at policy and decision making level are still missing i.e. the vital and integrative role of geospatial information and related enabling architectures such as National Spatial Data Infrastructures.(Scott & Rajabifard, 2017)

Figure 4 SDGs from real-world conditions.



Source: (Scott & Rajabifard, 2017)

Figure 5 indicators, targets and goals in SDG's

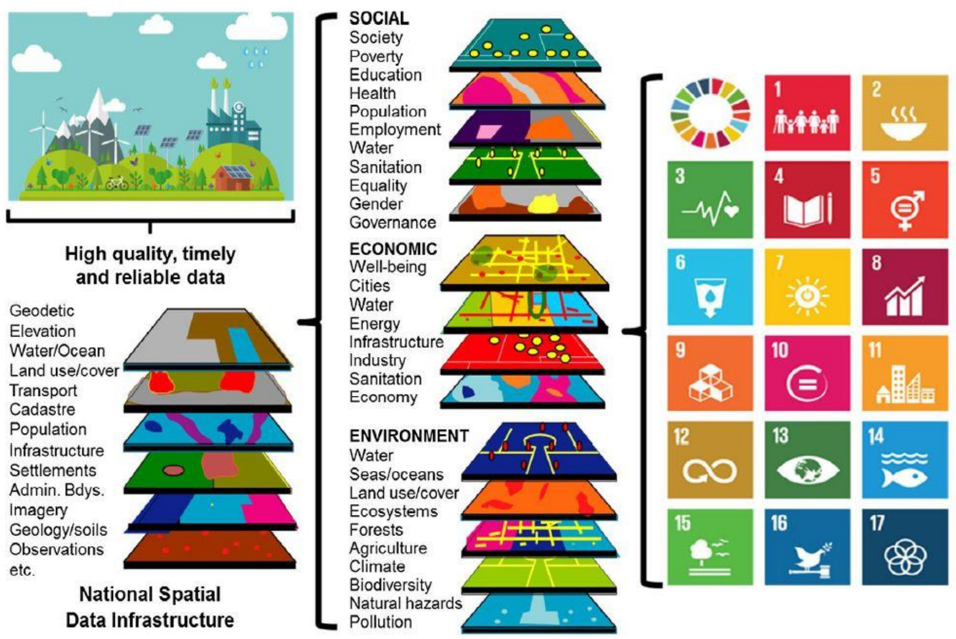


Source :(Scott & Rajabifard, 2017)

1.4.1 Geographical Data Infrastructure

Spatial data have emerged so as to refine sharing, integration and access of spatial services and information for rational policy making and decision-making. Getting better results out of spatially related economic, social and environmental decision-making is its principle objective and its design needs knowledge on the type of concept, contribution components, global drivers and their impacts, and the consumer needs.(Scott & Rajabifard, 2017)

Figure 6 Geospatial data themes within the (NSDI) to accommodate the SDG's



Source:(Koswatte et al., 2015)

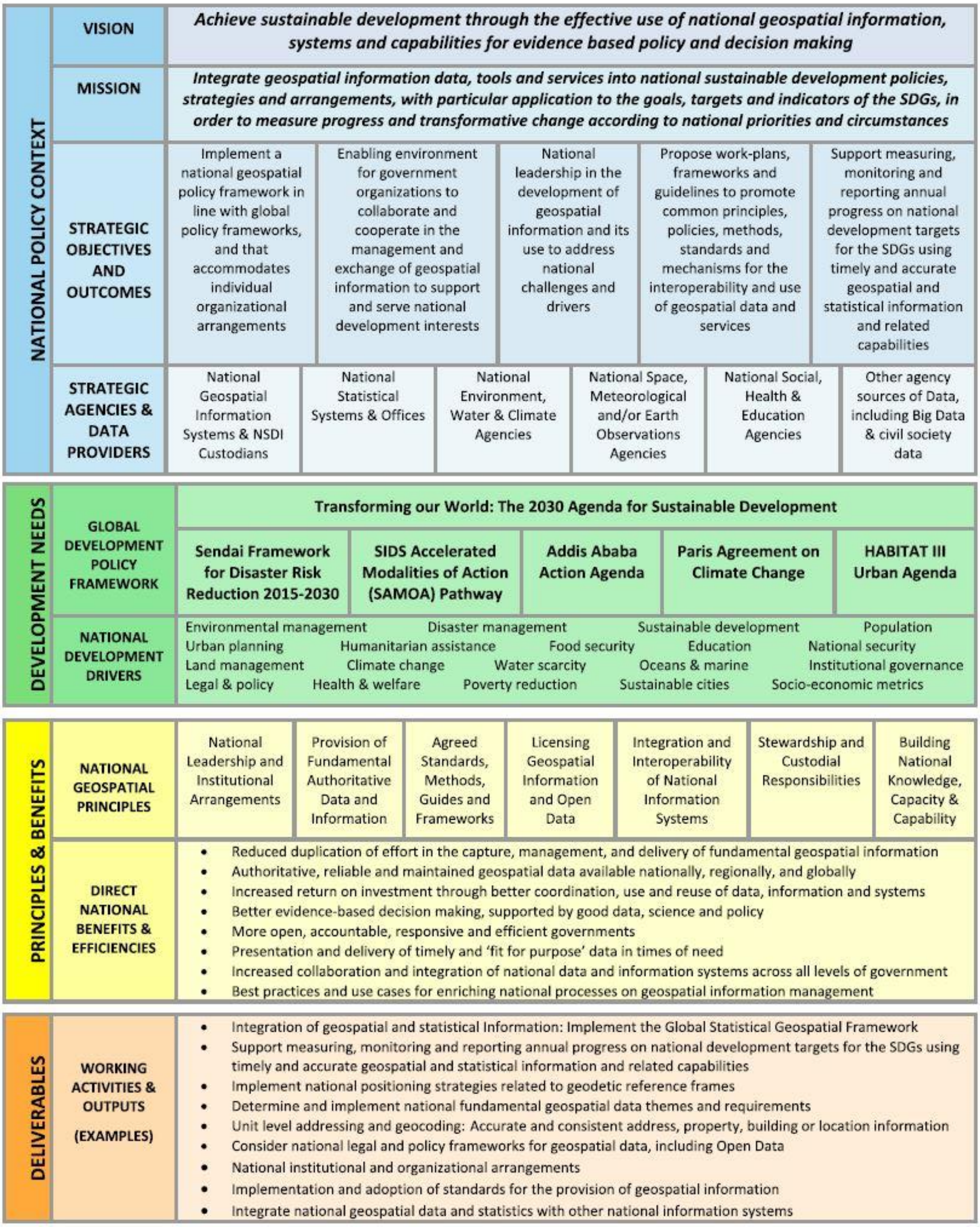
1.4.2 Strategic framework for National geospatial Information

Finally, now-a-days there is increase in perception on implementing SDGs, measuring and monitoring their progress requires a lot of fresh data and extensive modelling, better data management and analysis. Apart from that, it requires alterations and collaborative approach to corelate data–statistical, demographic , environmental , earth observations, and other geospatial data to geographic location.(Scott & Rajabifard, 2017)

“sustainable data for sustainable development” requires compatible earth observation data, geospatial and statistical as the vital information system which inputs into the global framework, and data by enabling technologies, robust methodological approaches, good science, and sound policy. These Database in combination with new information, including social media Data, ought to grant the basic baseline data inputs, in addition to new required data collections, to facilitate countries to monitor and measure progress and alteration within their individual circumstances.(UN, 2015)

Currently, there is no prominent framework or national policy for authorities to decide the implementation and integration of geospatial information into their development plans and this results in corresponding vital gaps and related points with national geospatial frameworks and methods. The answer to this resides in policy-making is considered that economic and social value of geospatial data. Additionally, sustainable production, regular and dissemination of geospatial data reflects the world that needs to be measured and monitored. These complementing strategies are vital to invest in, support, and proceed the strategic plan for developing countries.(Scott & Rajabifard, 2017)

Figure 7 National sustainable development. Framework for spatial information



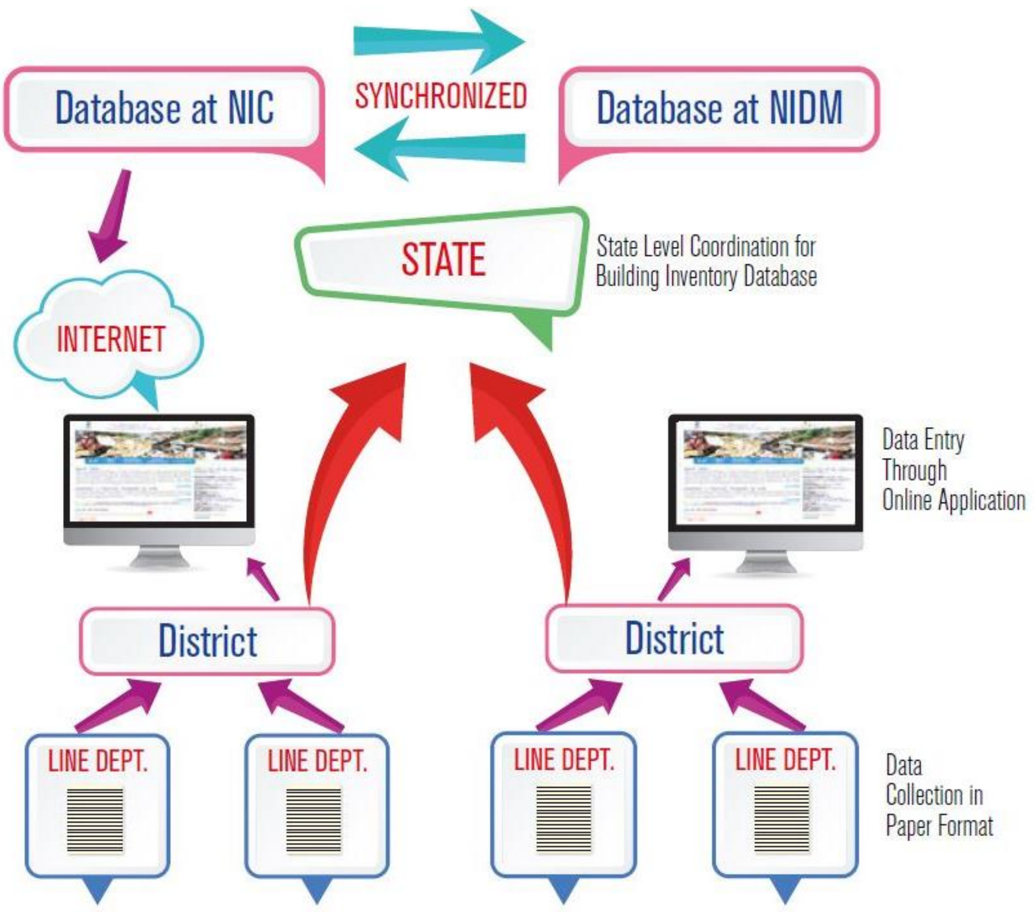
Source:(Scott & Rajabifard, 2017)

1.5 Indian disaster resource network – NIDM

IDRN is a tool that gives important equipment information, HR and critical supplies, existing at district level with the line departments and organisations. IDRN can be used by the state and district authorities during precadiment to have a quick access to information about existing resources that prevent loss of life and damage to property and infrastructure.

NIDM, that monitors and maintains IDRN portal, mainly does user administration. Data is updated in IDRN portal by State Department of Relief/Revenue/Disaster Management and SDMA in coordination with District Magistrate/District Collectors, who are authorised for enabling data collection and upgradation at district level. Forms of data collection are given by District Collector to the line departments and it is to be done at the district level with technical aid from District Informatics Officers of NIC.(Management, 2017)

Figure 8 Schematic diagram depicting functioning of IDRN



Source:(Management, 2017)

1.5.1 National Disaster Management Information & Communication

NDMICS groups geospatial data requirements into core data sets and hazard specific geospatial data (that are analysed from different sources and clubbed at a common base in accordance with standards and appropriate format) along with dynamic data, both spatial and non-spatial that are linked for efficient decision making. Data sets that are required for management of all disasters are core data sets (e.g. topographical, hydrological, socio-economic, infrastructural details etc) In development of a multi-layered GIS for Disaster Management, the main components of GIS are Input Maps (Survey of India topographical sheets, Soil taxonomy map etc.), GPS enabled Field Data and Satellite Data (Indian Remote Sensing (IRS) data). Spatially referred data are modelled as Vector data model (for discrete features) or Raster data model (for continuous features).
The following three types of data constitutes a GIS-platform and are used at varying levels in all GIS applications, including those deployed in EOCs.

- The Geo-database view: It is a view of the geographic database containing spatial datasets representing location information as a vital GIS model (raster, topologies, features, networks, etc).
- The Geo-visualization view: It is a set of intuitive maps and visualization that displays the feature relationships on the earth surface.
- The Geo-processing view: It is a bunch of information alteration techniques that derive new spatial datasets from pre-existing ones by function application.

1.6 Need for study

Most of the disaster management resource network systems are limited to user administration where access is denied to public usage. Modern technology provides unprecedented new opportunities for fulfilling the needs of decision makers. Some information systems that are already in operation demonstrate the potential for integrating real-time data with archival information in dynamic models to arm decision makers with powerful tools.

Chapter 2 Research Framework

This Chapter includes the research framework stating the research problem. It describes the research aim- to assess the "Integrating Big Data analytics using Spatial data infrastructure with disaster management to facilitate and coordinate the exchange and sharing of static and dynamic spatial data between all the emergency forces for better Decision-making". with detail methodology, research design, and the quantitative and deductive approaches to fulfill the objectives having scope and limitations.

...

2.1 Problem statement

Most of the disaster management systems lack the support for near-real time flood inundation mapping and situational awareness for decision making.

Social media generates the post- disaster scenario. There is no effective utilization of this data sets in disaster management (urban flooding), which is valuable and available in real time compared to conventional data sources., which can be explored more effectively.

2.2 Research Question

Can social media data sets be used as information source for near real time flood inundation mapping?

2.3 Research Aim

"Integrating Geospatial Big Data analytics with Spatial data infrastructure and creating social media-based flood vulnerability mapping to validate with conventional vulnerability mapping techniques."

2.4 Research Objectives

OBJECTIVE 1: To Identify available Open Geospatial Consortium (OGC) and Application programming interface of social media platforms for data mining.

OBJECTIVE 2: To prepare urban flood vulnerability and inundation maps using Analytical hierarchy process and multi criteria parametric analytics.

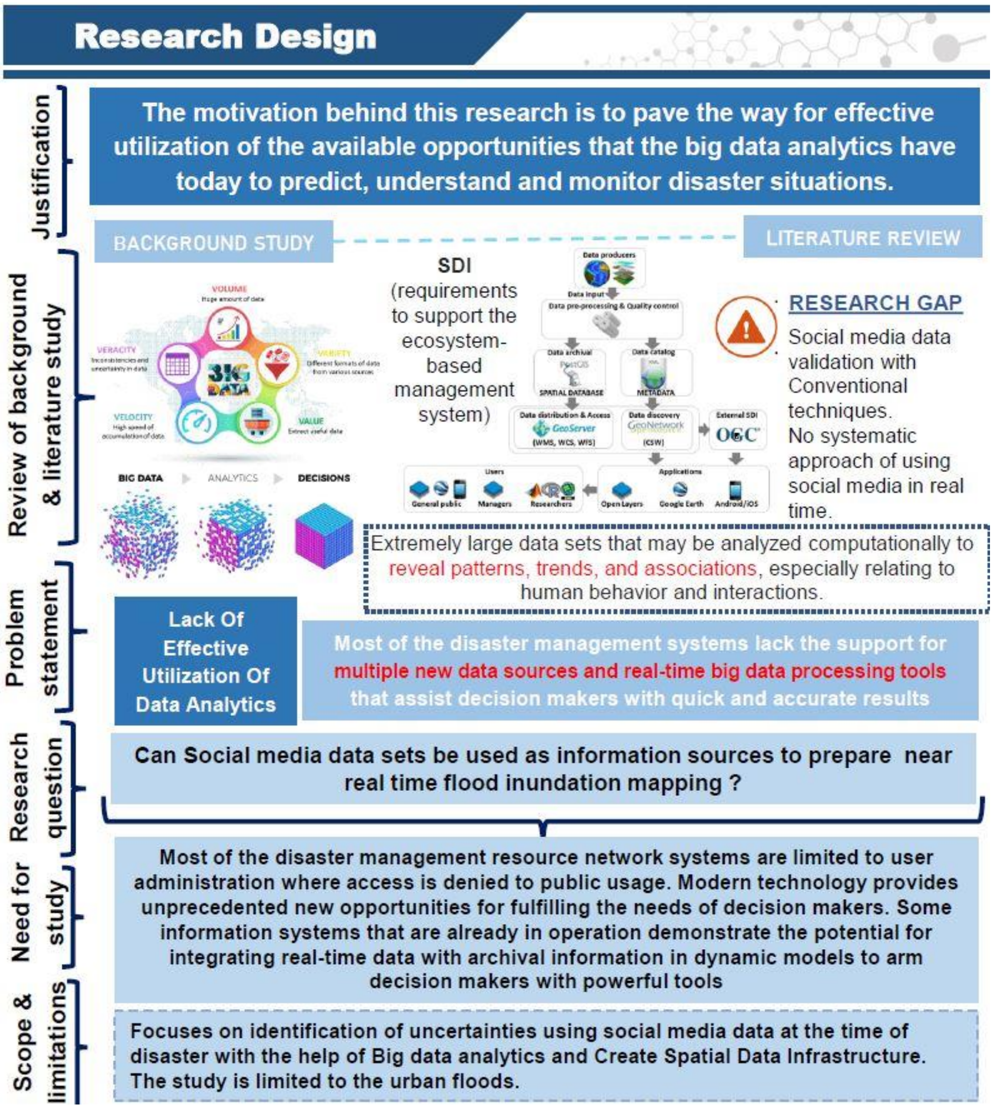
OBJECTIVE 3: To prepare and review the potential of urban flood inundation map using crowdsourced spatial information.

OBJECTIVE 4: To validate social media-based flood vulnerability map with conventional mapping techniques.

2.5 Research methodology

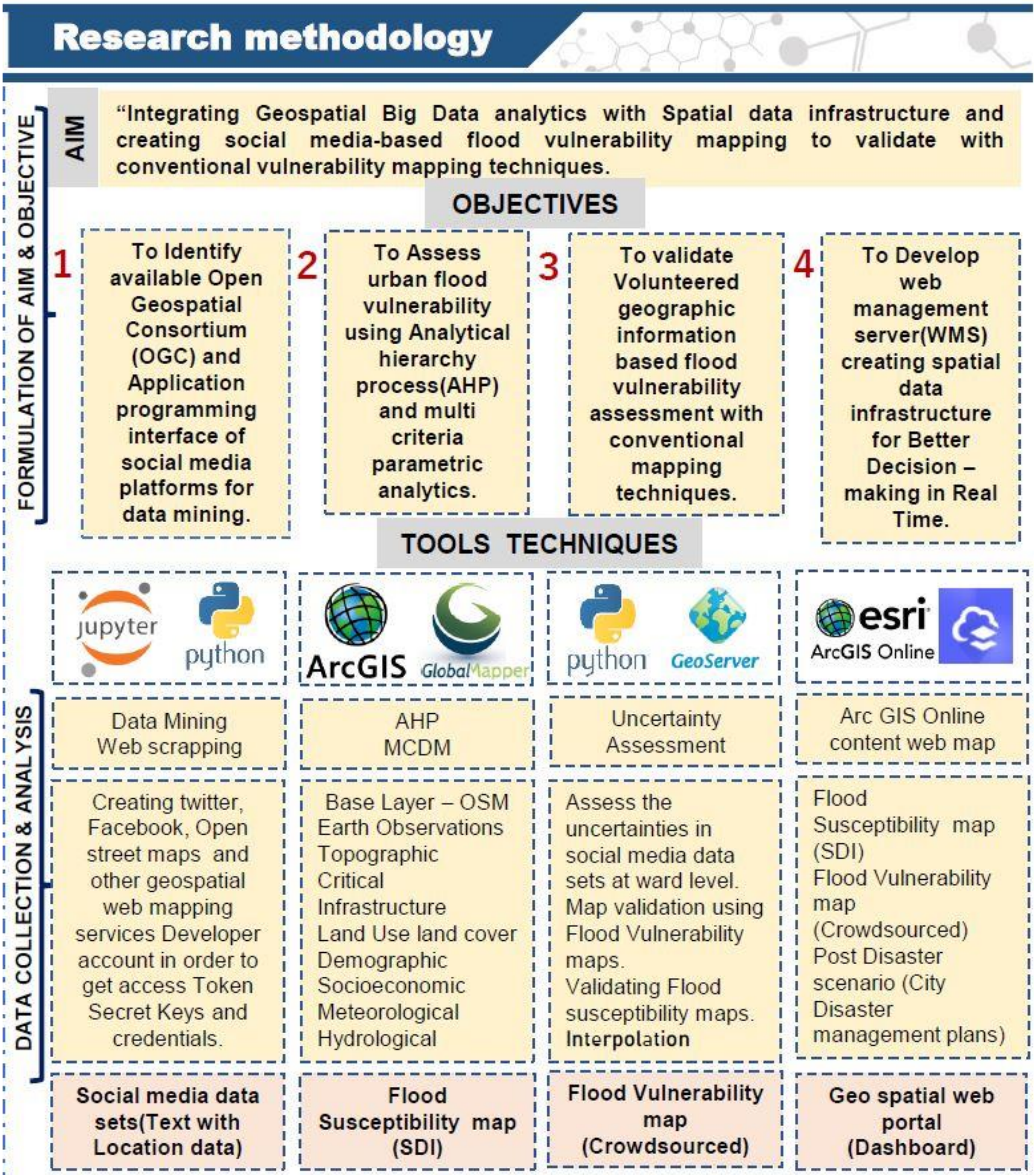
The methodology of the study starts with the background study related to the importance of the Big data analytics and the presently available open data sources and the sustainable development goals of UNDP, 2015 integrating Big data Analytics. Later the study is established with the recent literature studies and review done on the Spatial Data Infrastructure.

Figure 9 Research Design



The study primary focus on the parameters of Spatial Data Infrastructure and Integration of Big data analytics which are drawn from the literature review and aims to Develop the Emergency management tool during Risk Management Phase.

Figure 10 Research Methodology



Source: Author

2.6 Research Design

Two approaches are used to analyze the study areas in context of urban flooding. A Stastical approach for the Prediction model is used for Flood susceptibility mapping which gives the Flood extent based on affecting parameters. Social media data mining approach is to deal assessing the concept of Crowdsourcing.

2.6.1 Single sensor approach

In this approach it seeks to create numerical research or data which then can be converted into usable statistics and single sensor is a satellite. To get numerical data so then, can turn this into facts and statistics and find the Flood prediction. It identifies outcomes through statistical analysis.

2.6.2 Multiple sensor approach

In this approach it seeks to assumptions is developed from Real time generate social media data. Mobile phones are considered as multiple sensors.

2.7 Scope and limitations

In this research, the study is limited to the Urban Floods. Focuses on identification of uncertainties using social media data at the time of disaster with the help of Big data analytic.

Chapter 3 Research Data & Methods

This Chapter illustrates the data to be collected and brief about the techniques adopted for the preparation of Real time Analytical Dashboard. The detailed process for mapping the parameters to create a model based on occurrence of flood. Data mining real time data using Python programming Language. The validation techniques to assess the uncertainties of social media data sets using API's (Application programming Interface) Access Token keys.

..

3.1 Techniques for the research

Perspective to analyse the Vulnerability associated with pluvial flooding using social media data sets where python Programming language and jupyter notebook are used for data scrapping and data visualisation. Mining data from social media sites requires access to API of that particular Application. Few statistical and analytical methods were adopted from the relevant researchers which were mentioned from the literature study.

In the whole approach there are three techniques used to analyze the flood inundation. Firstly, with prediction model where Occurrence ratio method is used and flood susceptibility is analysed using GIS and remote sensing Techniques. In the second Technique Social media data sets are used to create a Flood inundation at street level. Application programming interface of twitter is used for data mining and Arc GIS API for data Visualization's

3.1.1 Occurrence Ratio Flood Modelling

The Occurrence ratio is the Correlation between Flood Occurrence and each Flood influencing parameter. The Occurrence ratio of a particular factor in a causative parameter can be calculated when compared with Flood Inundation. In the correlation analysis, one is considered as an average value and value greater than one is considered as a higher correlation, while a value less than one implies a lower correlation. The Occurrence ratio of each factor was obtained using Eq. (1) and then the Flood susceptibility model (LSM) was calculated using Eq. (4).(Sahana & Patel, 2019)

$$\text{Occurrence Ratio (OR)} = \frac{\text{Flood Occurrence}}{\text{Factor Ratio}} \tag{1}$$

Where,

$$\text{Flood Occurrence} = \frac{\text{Flood Occurrence in individual factor}}{\text{Total Flood points}}$$

$$\text{Factor Ratio} = \frac{\text{No. of Flood points in individual class}}{\text{Total number of points in whole classes}}$$

Weighting of parameters

Individually each Flood parameters can have different impact of geographical associations with Flood. Anyways, as Flood is a dependent factor of multiple parameters, predictive model analytics of Floods requires analysis Influencing factors weights. To determine the correlation between each parameter, parameter rating (PR) were derived for every geospatial parameter based on their influence of geospatial association with each Flood factor i.e. Frequency ratio, thus:

$$PR = [RF_{max} - RF_{min}] / [RF_{max} - RF_{min}]_{min} \tag{2}$$

Where,

$$RF = \frac{\text{Factor class FR}}{\sum \text{Factor classes FR}} \tag{3}$$

Note: PR = Predictor Rate
RF = Relative Frequency

Table 2 FR Model Data Sources

DATA SOURCE	RESOLUTION	LAYERS
Bhuvan	WMS Imagery	Geomorphology
		LULC
ASTER-DEM	30m	Slope gradient
		Stream order
		Elevation
SLUSI	1:50,000	Soil
(GSI 2013-2014).	1:50,000	Geology

Source: Author

Database creation

The influencing parameters of flood and its database creation require knowledge about the occurrence of flood in that particular factor. Generation and identification of spatial database for flood occurrence models plays crucial role in the flood susceptibility mapping. Although flood susceptibility mapping had no regulations for calculating the influencing factors. The influencing factors on flood susceptibility parameters includes, geology, geomorphology, topography, lulc are considered most influencing compared with rest of the parameters. In the research also considered DEM, drainage density etc. based on availability of data in required format. Based on literature study conducted, six influencing factors are considered as major parameters where weighted overlay analysis is conducted considering flood occurrence in those particular parameters. influencing parameters raster's are created and projected using UTM (Universal Transverse Mercator) system. causative parameters where applied with occurrence ratio model.(Sahana & Patel, 2019)

3.1.2 Crowdsourced Inundation mapping

Now a days the citizen science which is popular in disaster management to collect data. Social media and mobile phone technologies have completely changed disaster management and emergency response by enabling public affected to create real time and local information on emergency events. Crowdsourcing is a model which refers to technique that collects data from public (the crowd) provide response through open web portals which are transmitted using mobile phones.(Fohringer et al., 2015)

Table 3 Crowdsourcing in Disaster management cycle

	Crowd Micro tasker	**Crowd reporter**	**Crowd Social computer**	**Crowd sensor**
Mitigation	✓	✓	✓	
Preparedness	✓			✓
Response		✓	✓	✓
recovery		✓	✓	✓

Source: (Juhász et al., 2016)

Technology wise there are two different approaches which are identified from literature.

- Communication oriented: Interaction between disaster management systems and public in affected areas is considered as communication-oriented approach.
- Data oriented: unstructured sources such as Twitter, (Facebook, Instagram, etc) social media platforms processing and mining to generate emergency alerts is considered as data-oriented approach.

3.2 Research parameters Data

Research parameters were identified through critical evaluation of the literature study. Criteria for selection of parameters include:

1.Flood Prediction Parameters

- LULC
- Geology
- Drainage Density
- Geomorphology
- Soil
- Topography
- Flood Occurrence

2.Social media-based Inundation Model

- Twitter data sets
- Crowdsourcing Data (Chennai Rains.Org)

3.3 Data collection and Visualization Tools

The data collection is of both primary and secondary sources. The data mostly depend on the secondary sources available online. The data is classified into two levels as per approaches. A detailed checklist is attached to this report in Annexure

3.3.1 Twitter API's

Twitter API is a group of parameters which can be accessed with the set of URLs. The URL's gives access the tweet with a #. These access token keys provide public twitter data with geographic location where the user enabled the Location data. Developer account is been created to get access to the token keys.

Consumer key ="FZo5PuU9n08AkDALuKVO7Kcin"
Consumersecret"lpZvfHSuC0ttmii31Qsb60E59Y0BqXmOhcOyBQVHnBYKUCYoaH"
Access Token ="1166310141790048262-HMZaS8FllchVOf3znjwXPSof7zdTNC"
AccessTokenSecret"i2voP37V3eiIuEd4ZdLoi7LRQvoINXsgycf9VN6F2q1hr"

Figure 11 Twitter Consumer API Keys

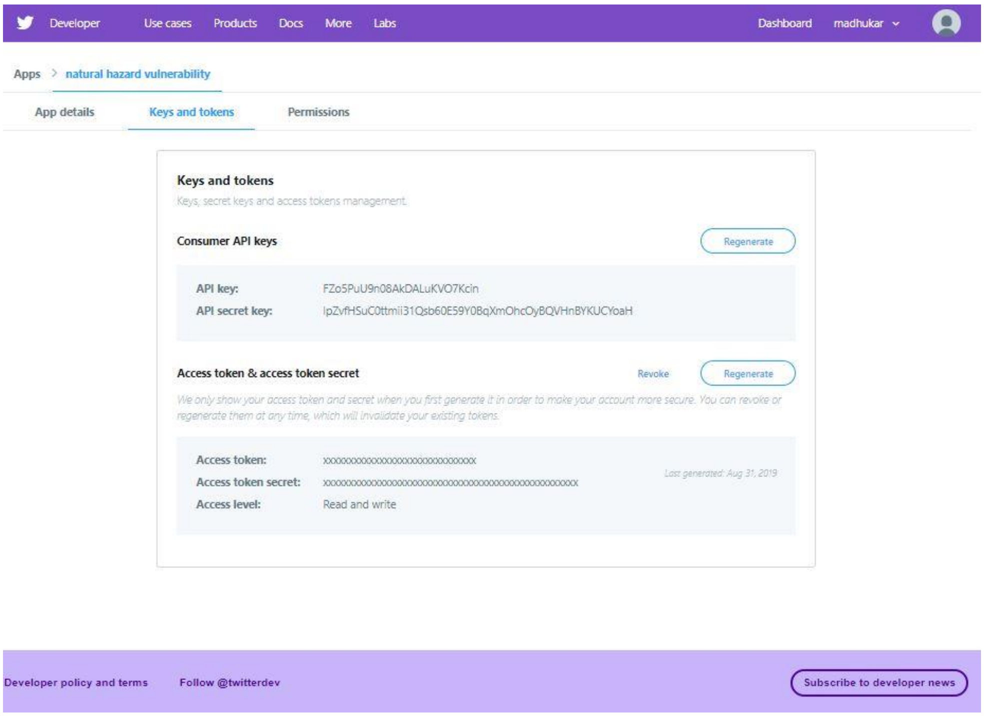
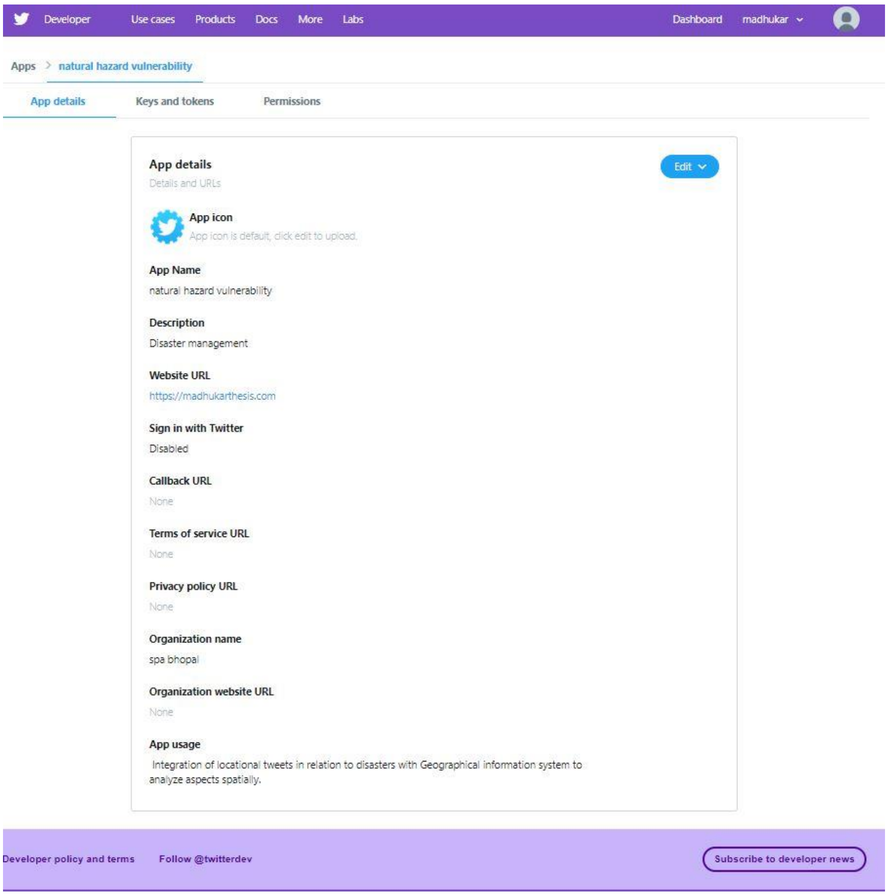


Figure 12 Twitter Developer Account Details



3.3.2 Arc GIS API

ArcGIS API is a modern and powerful tool with python enabled library to perform analysis on web. In this study Arc GIS online is used for data visualization and spatial data management tasks using python scripts. Arc GIS online is a web platform which constructs data structures for visualization and analytics. It enables to get access python modules for information models. Mapping and visualization module represent map images, vector data using Jupyter notebook.

Figure 13 ArcGIS Developer Account

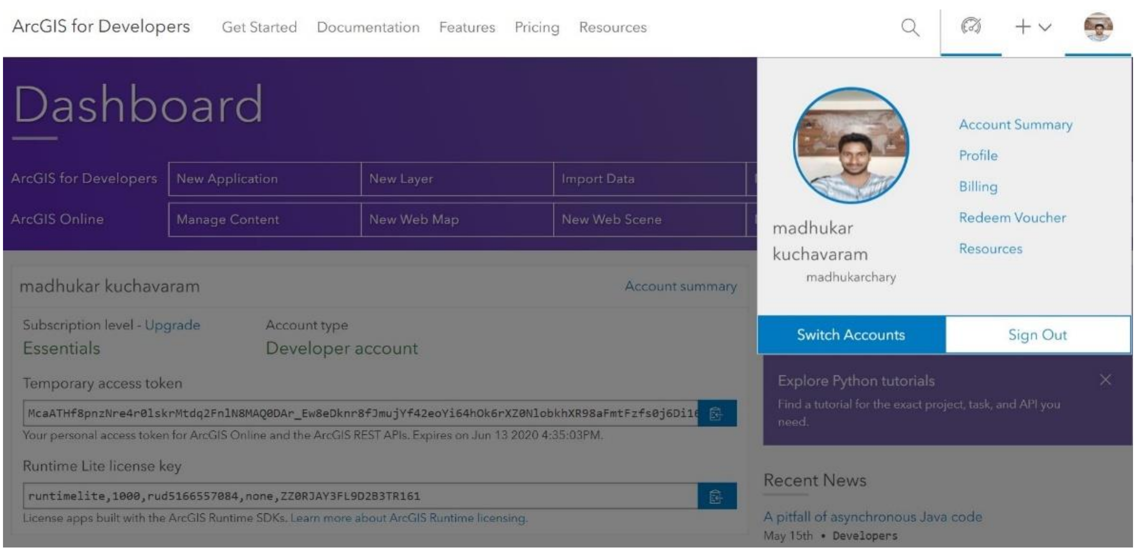
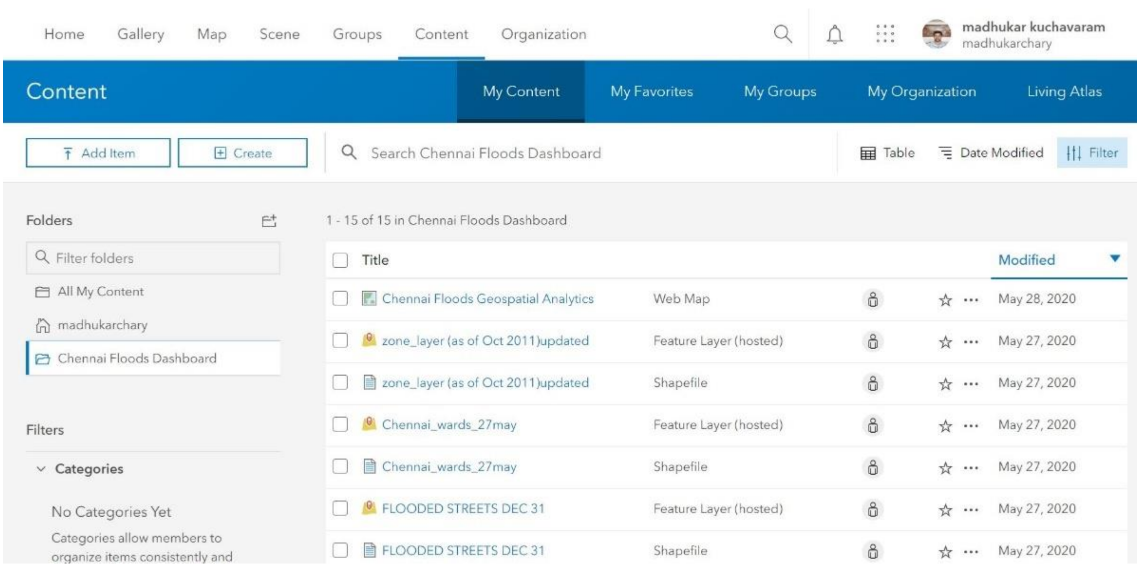


Figure 14 ArcGIS Online Content



3.3.3 Jupyter notebook

Jupyter notebook is an Integrated development Environment web application open source used for data Statistical modelling, visualization, numerical simulation, machine learning, Data cleaning and many more. Jupyter is an integrated development environment (IDE) enables to consolidate source codes. Type Jupyter notebook in commend prompt to launch it on localhost.

Figure 15 Jupyter Notebook CMD http localhost

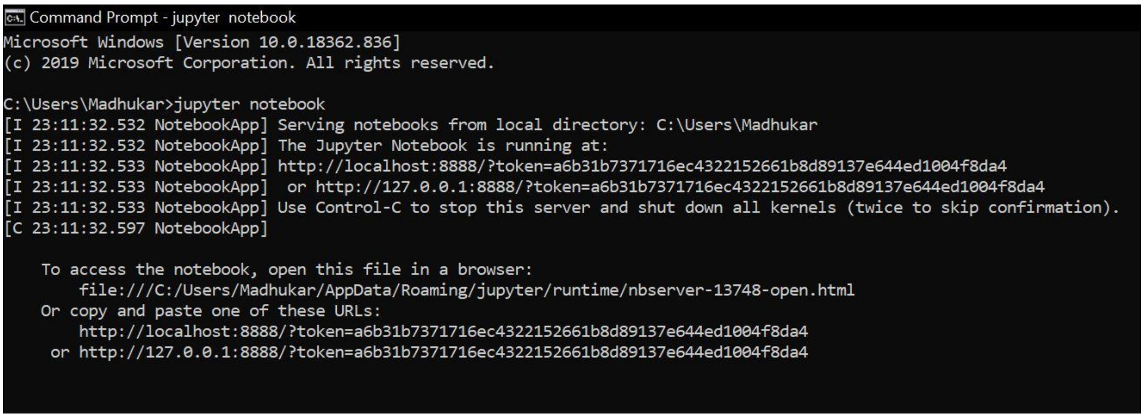
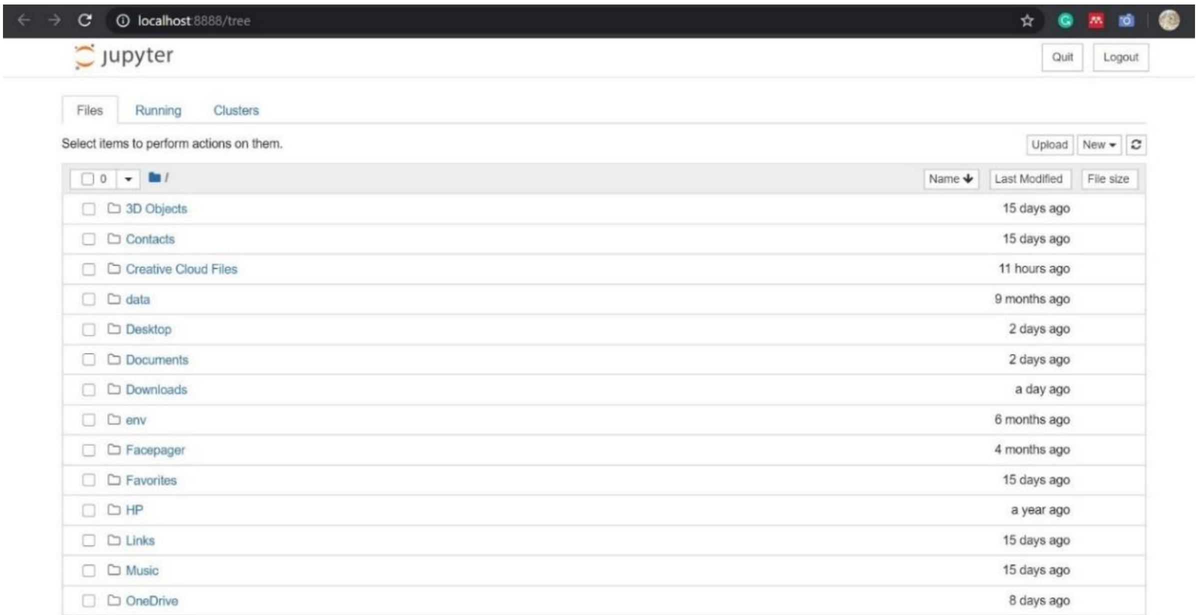


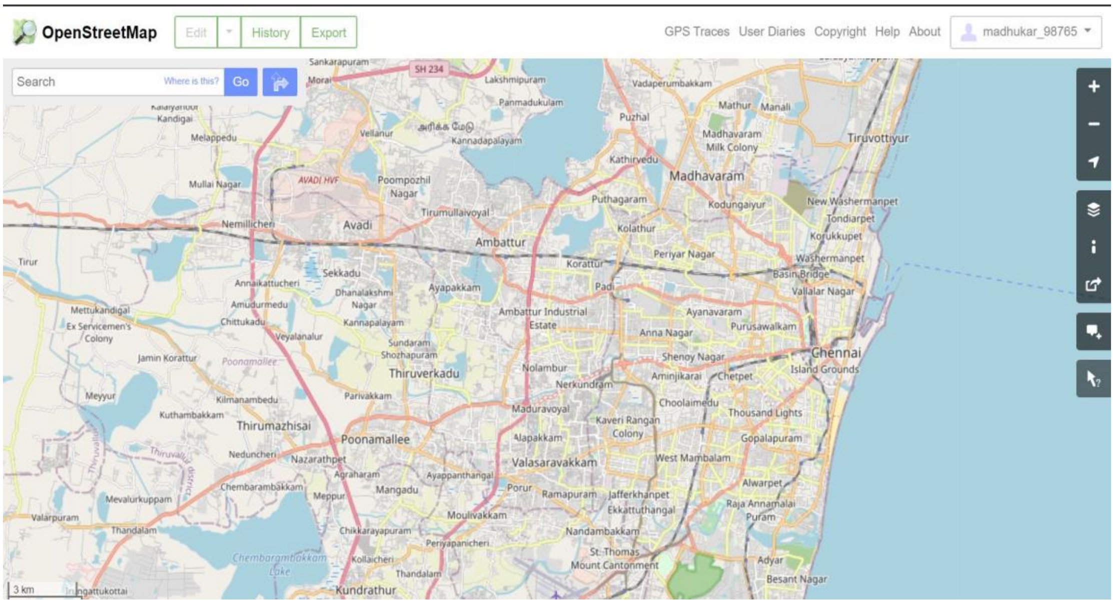
Figure 16 Jupyter Notebook Localhost



3.3.4 Open street maps.org

The open street maps data is collected based on crowdsourcing and is available on open database platform. It is a collaborative platform to edit, create maps and can be downloaded directly from openstreetmap.org.

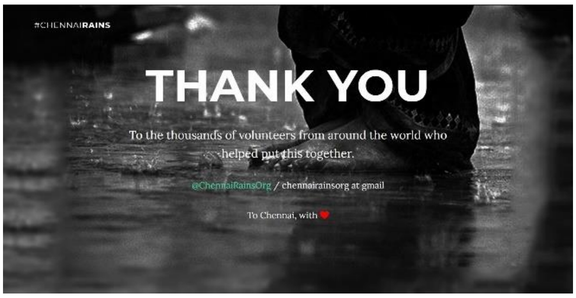
Figure 17 OSM Open portal



3.3.5 Chennairains.org

Chennai rains.org is an open web portal which was used to disseminate the real time data at the time of floods in 2015. Data was collected from various social media crowdsourcing platforms such as Facebook, twitter etc. activated a safety check feature. The portal began with simple spreadsheet to collect data. Overall idea was simple to collected data about reporting inundated areas and knowing people who can offer shelter.(Shi et al., 2019)

Figure 18 Chennairains.org Open Portal



Chapter 4 Literature Review

This chapter Introduces the concepts of Big data, spatial data infrastructure (SDI) and Contribution of bigdata analytics in sustainable development goals for development and humanitarian action. The chapter concludes with understanding the importance of crowdsourcing and social media in emergency management. The case studies which used social media as an emergency management tool in disaster management cycle.

……………………………………………………………………………………………..

4.1 What is Big Data?

Big Data is produced at exponential rates and has several misconceptions attached to it. The term Big Data is used when the size of the data is in exabytes or petabytes or terabytes or gigabytes or larger, but that's is not its definition entirely. Depending on the context a small amount of data can also be referred as Big Data. For instance, 100megabytes file cannot be attached to a mail because the attachment of this size is not supported by the email system. So,100 megabytes are 'Big Data' for email system. Similarly, altering image files around 10 terabytes urgently wouldn't be possible using conventional way. Therefore, the image file of 10 terabyte is referred as Big Data.(*(2) What is Big Data_ (2019) - YouTube*, n.d.)

Figure 19 Big Data

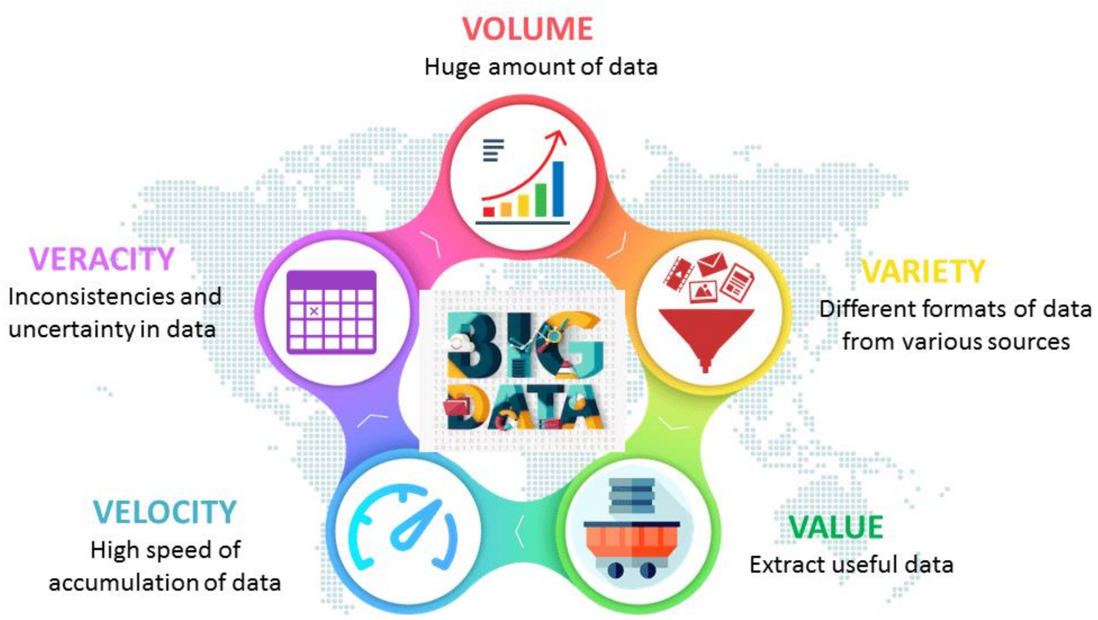
Source: (CloudMoyo, 2019)

4.1.1 5v's of Big data

Velocity, Volume and Variety were previously the three different dimensions of higher degree in which data sets were considered as Big Data. But recently, value and veracity have also been added.(Panneerselvam et al., 2015)

- **Velocity**: The speed of data generation
- **Volume**: the amount of data generated
- **Variety**: the diversity or different types of data
- **Value**: the worthiness
- **Veracity**: the quality, accuracy, or trustworthiness of data

Figure 20 Five V's of Big Data



Source: (GSAM, 2016)

4.1.2 Sources of Big data

The primary sources that generate large amounts of Big Data are- Social data, machine data and transactional data. The data is further classified as structured (refined) and unstructured (unrefined) data. (Minelli et al., 2013)

primary sources of Big data are:

Social media data: Popular social platforms like twitter, Facebook, Google plus, LinkedIn and YouTube are receiving large amounts of data daily. Reports confirm that Facebook receives around 100 terabytes of data, Twitter gets 400 million tweets, Google plus and LinkedIn receives tons of terabytes of data and YouTube uploads 48 hours of new videos every minute. Storing and processing such tremendous amounts of data is insane with an ever increase in the user count. Therefore, data needs to processed with speed as it holds a chunk of important information.(Marr, 2014)

Machine data: Soon the usage of the internet in smart meters, satellites, games, road cameras having sensors will increase substantially and these will deliver the V dimensions of data such as variety, velocity, value and volume.

Transactional data: The data generated daily in the form of transactions such as invoices, payment orders, delivery receipts, storage records (both online and offline) etc is meaningless or in unstructured form. Therefore, it must be sorted so that it makes sense and can be put to some use. Majority of the organizations work towards transforming this transactional data into a useful data.

Figure 21 Data Open Sources

Source:(CloudMoyo, 2019)

4.2 Data Analytics

An efficient way to analyse and inspect large data is Big data analytics and several organisations have adopted it years ago. Previously, collected data would be analysed and the unrefined information stored for future decisions. Presently, data is easily identified and decisions made promptly owing to Big Data's speed and efficiency. (Minelli et al., 2013)

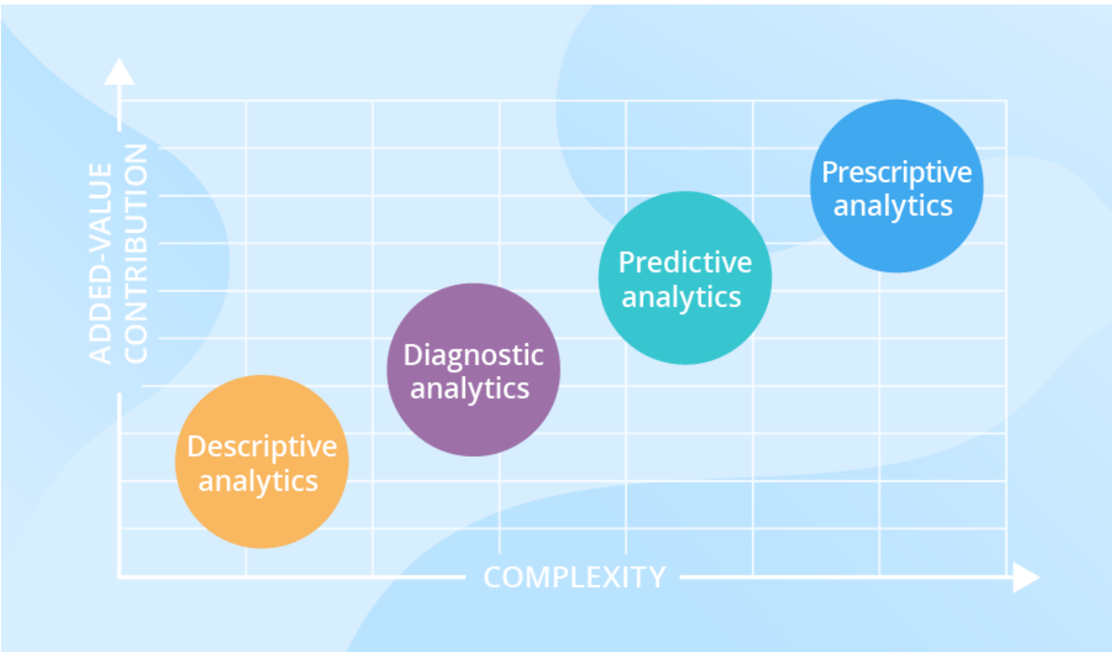
Figure 22 Data Analytics

Source: (Panneerselvam et al., 2015)

4.2.1 Types of Big Data Analytics

There are 4 different types of analytics. With the simplest one being first and list being in increasing order of sophistication. Value increases with the complexity (CloudMoyo, 2019)

Figure 23 Types of Big Data Analytics



Source: (J. G. Lee & Kang, 2015)

- **Descriptive analysis:**

Descriptive analysis arranges primary raw data into many data sources to produce valuable insights from past without reasoning. therefore, this is not suggested by data consultants to highly data driven companies but they prefer to use it along with other data analytics.

- **Diagnostic analytics:**

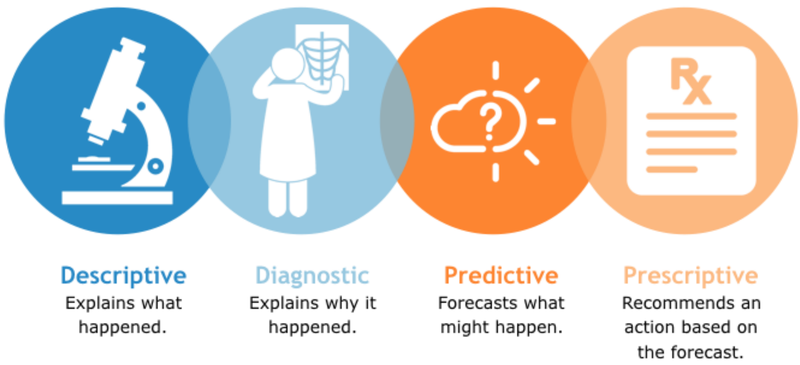
It deals with a particular problem in detail and measures past data trends in relation with data in real time. It resonates “why something happened”. The data should be in a sorted manner otherwise the collection of data may be time consuming and it would depend on one person for every issue.

- **Predictive analytics:**

Predictive analytics predicts what will happen likely. To do that, it uses the data collected from descriptive and diagnostic analytics and detects the clumps and exceptions, making it a valuable tool for projection. Predictive analytics is a proactive approach and is advanced analytics type with several advantages based on machine or deep learning like sophisticated analytics.

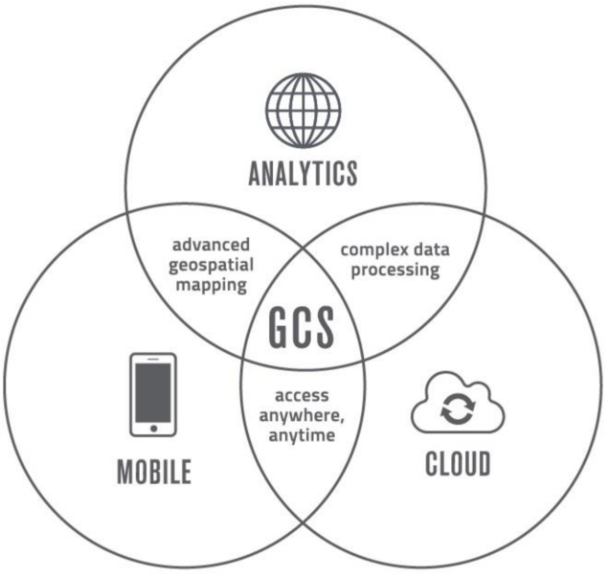
- **Prescriptive analysis:**

It helps to get rid of the problems that would pop up in future and also deals with on present trends. The advanced sophisticated technologies such as machine learning, business rules and algorithms are used by the prescriptive analytics. Additionally, this needs external information apart from previously existing data because it is algorithm-based.



4.3 Geospatial Big data

Figure 24 geospatial data integrating mobile, cloud and Analytics

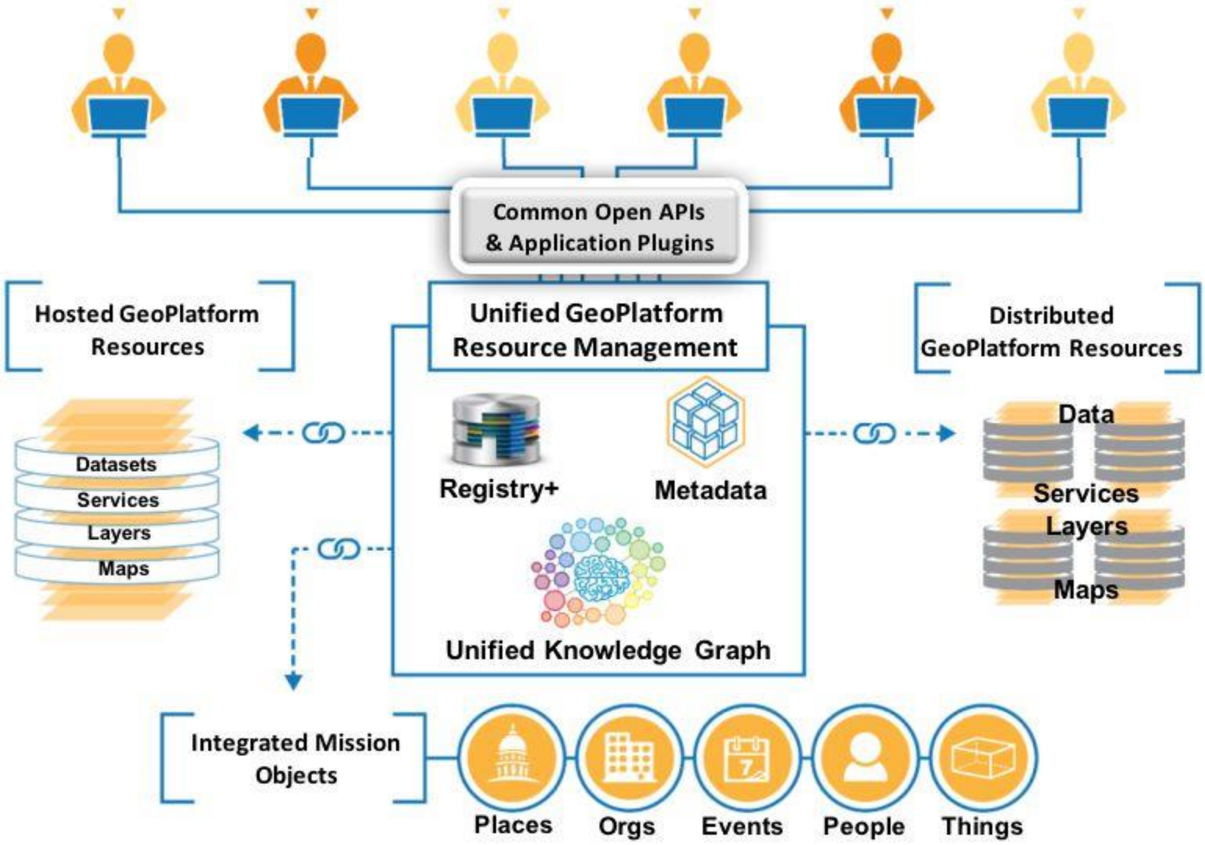


Now a day's big data analytics for locational data is getting considerable importance for enabling users to analyze large amounts of spatial data. The locational data typically refers to geospatial data where the computing capacity excides. A study by McKinsey institute at global level says the generation of personal geospatial data was 1pb in 2009 and grown at 20% per year without including RFID sensor data. According to united nations global geospatial information management shows 2.5 exabytes of data is generating every single day and most of it is spatial data. Google generates 25PB Spatio-temporal data every data. The trend of generating Spatio-temporal data is escalating more because of mobile usage. In India internet traffic from desktop computers is shifting to mobile devices.

4.4 Spatial data infrastructure

US National research council coined the term spatial data infrastructure in 1993 to design a framework of institutional arrangements, policies and technologies to integrate the exchange of spatial data and information associated with geospatial data across the community where information is shared. The framework which enables the integrated sharing of data at broad levels at global, regional and national level. This infrastructure framework provides the uses to produce the data which includes posts automated, exchanging spatial data at institutional levels, evaluating and discovering. The data infrastructure that enable the framework of users, metadata, tools and geographic data that are connected interactively in order to utilize geospatial data in flexible and efficient way. Another definition is "the technology, policies, standards, human resources, and related activities necessary to acquire, process, distribute, use, maintain, and preserve spatial data"(Koswatte et al., 2015)

Figure 25 US National Spatial Data Infrastructure — GeoPlatform Architecture

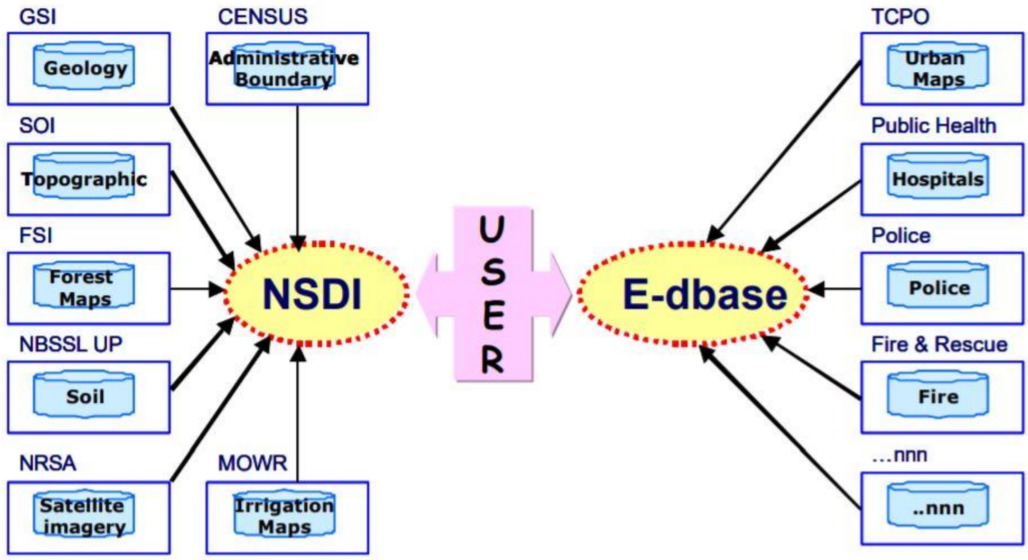


Source: (Report et al., 2018)

4.4.1 National Spatial data infrastructure

National-Spatial-Data-Infrastructure-India (NSDI India), has been implemented as a network of geospatial data nodes across the central and state governments in order to improve access to geospatial data by all stakeholders. A set of geospatial data and process standard specifications from Open Geospatial Consortium (OGC) and International Standardization Organization (ISO) have been taken to enable interoperable sharing of and access to data amongst organizational data nodes and end-users. Standards-based Web Map Services (WMS) are accessible from data providing agencies for use in the development of GIS applications. Attempts are made to reengineer and maintain available data sets for provision of Web Feature Services (WFS) for improving access to processable feature data sets with stakeholders. Collaborative preparation and maintenance of an authoritative, seamless and nationally consistent set of high-resolution National Foundation Spatial Data (NFSD) has been a top priority for NSDI to tackle problems in integration of data sets from multiple data nodes. Several national- and state-level policies have been in place to provide the required governance framework.(Sarda et al., 2018)

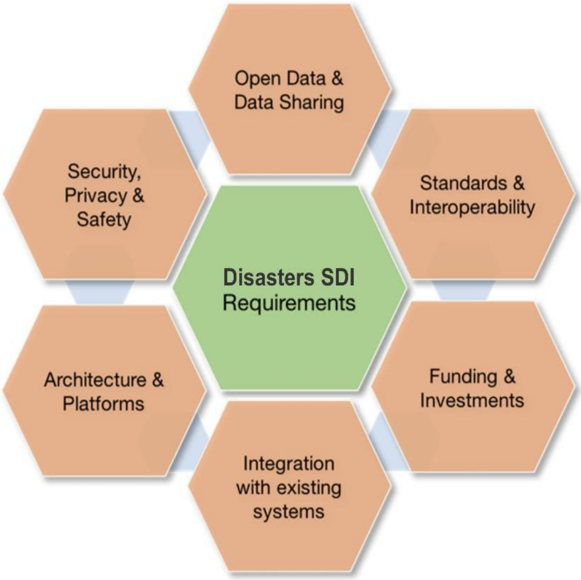
Figure 26 GeoPlatform Architecture – NSDI



Source: (*National Spatial Data Infrastructure _ Department Of Science & Technology*, n.d.)

4.4.2 Spatial data infrastructure in disaster management

Figure 27 Disasters SDI Requirements



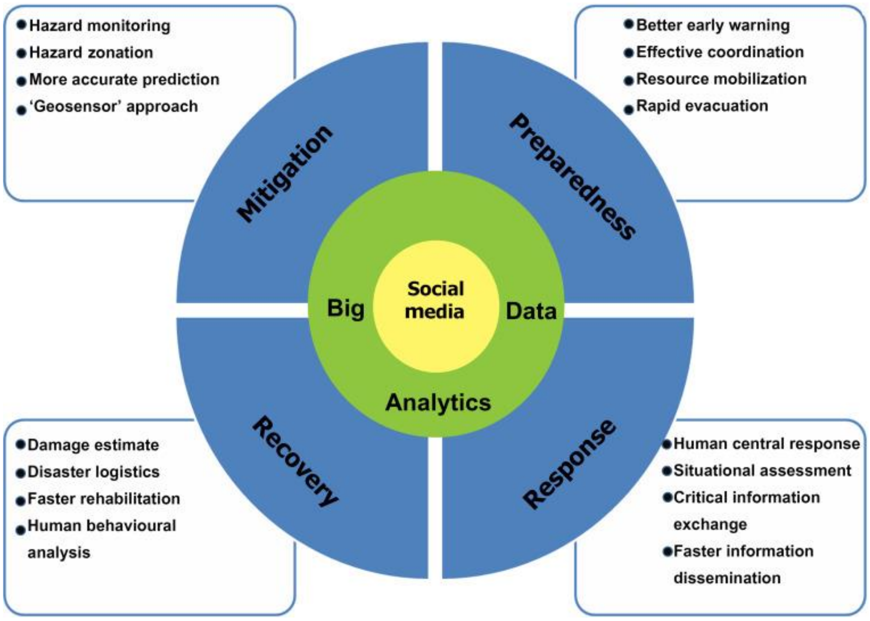
Source :(Juhász et al., 2016)

SDIs have supported in disaster management in the recent past by enabling geospatial data integration between organisations to enable the access and utilise existing geospatial data(Snoeren et al., 2007). The conventional SDIs are crucially institutional focused and carry 'a significant organisational inertia an rapidly growing complex legislative framework that is challenging to change'. However, SDIs are gradually evolving to user centric and process-based models from data centric models.(Juhász et al., 2016) .

4.5 Geospatial Big Data in Disaster Management

Large data sources of social media can be used to enhance communication when hit by a disaster, before being approached by it or after it leaves. It all began with the Haiti earthquake in 2010 and the Tohuku earthquake and tsunami in 2011. The Japan and United States collaboratively started research in these areas, focusing mainly on, tsunami and earthquakes. Social media can also be used in pre disaster backchannel communications, leading to a rapid increase in social capacity of information dissemination and. generation. Studies on social media data in emergency management have come into picture only now-a-days since real-time platforms like Twitter started. According to the study of Lindsay in 2010, social media platforms was ranked as the 4th most popular source for dissemination and retrieving information on predicament. The effective geospatial Big Data aggregation , visualization and integration from Social media platforms will assist disaster management authorities to upgrade the situational awareness and in better decision-making.(Report et al., 2018)

Figure 28 Big Data in Disaster Management



Source: (GSAM, 2016)

- ## Mitigation stage

The decision-making process operations which are supported by several hazard maps at mitigation stage in emergency management are generated by cyber GIS database. Social media platforms which generate the scenario of disaster in real time are used as Geoserver network systems to monitor and detect extreme events. Intelligent transport systems and sensor networks are utilized now a days for early warning systems and for data dissemination, which is integrated with big data platforms. Geospatial Data from social media becomes more accurate with rapid increase in target regions and people. Hence data facilitation prediction of approaching disaster and mitigation. High resolution satellite data and technologies like UAV's can be integrated with crowdsourcing data to obtain accurate hazard maps for risk modeling of natural vulnerability assessment.

- ## Preparedness stage

Haiti earthquake in 2010 information was first disseminated through social media platforms. Critical information of preventable hazards which can be avoided are broadcasted to large groups via SMS. Stakeholders in disaster management authorities responding in real time can be served with dynamic data. During Mahasen cyclone 2013 in Bangladesh used early warning systems under environmental extremes and population project via mobile data. The keywords from tweets about disaster crises in collaborative efforts under UN global program where citizens as sensors was considered as large-scale initiative.

- **Response stage**

Disaster response has many lives at stake. In the past few years, social media has been a vital source communication tool for disseminating emergency response. Twitter can be considered as good example for emergency communication tool used for warning messages for millions of people rapidly.

- **Recovery stage**

Communication processes should be flexible in Disaster recovery systems and it is crucial to know how active participants providing recovery through online communities. Social media to assess the damage by processing the images of structures which are damaged which are captured by mobile phones. During disaster recovery phase, human behaviour can be estimated by geospatial data which is captured from social media platforms and these data can be used for identification of victims trapped in isolation. In recovery process data analytics can provide with vital information about logistics and volunteer coordination in difficult times.

4.5.1 Benefits of Geospatial Big Data Storage

Geospatial Big data is the parameter that enables efficient analytics of large quantities of information obtained through the modern practices. It is the collection of engineering and scientific tools and methods that help in making the best of huge amounts of available data that not only answers accessibility issues, but also issues related to storage, analysis, distribution, and effective visualization of data analytics. Accurately, communication also enables monitoring and understanding the entire public body and open sourced available communication like content and messages on social media where public may be conveying their situation to their loved ones, exchanging messages or seeking for help. However, geospatial data enables a detailed analytics of all information that gives valuable communication of a general validity for the public; information about a disease outbreak for an instance.(GSAM, 2016)

- **Efficient Allocation of Resources:** Geospatial Big data generated from remote sensing platforms and geo-informatics identify the gaps and suggest where to allocate resources for risk mitigation that includes recovery monitoring, focusing early alerts and assessing resilience.
- **Identification of Vulnerable:** It can pin point the most vulnerable section of the community and can be used to pursue 'risk informed' development.

4.6 Big data Analytics challenges in Disaster management

Large volumes of data travelling at high speed on social media is a matter of concern for Big data apart from quality and services including the ever-evolving privacy, security, technologies and considerations. Therefore, a lot of research is still required to be done in the field. Many Geospatial Big Data sources are paid and restricted to public, while Facebook has an open Application Programming Interface (API) to access its data, access to twitter's data stream can be restrictively expensive, some big data sources are view-only. Overestimation, false, noisy data are some other challenges in geospatial Big Data analytics for emergency management.

4.7 Case Studies in Geospatial Big Data and Emergency in Disasters

4.7.1 Case 1: Chennai Floods in India, 2015

In November 2015, south India experienced heavy rains by the northeast monsoon causing extreme pluvial flooding of coastal cities –Pondicherry and Chennai. The NDRF (National Disaster Response Force headquarters) has got hands on over of crucial data on social media via the Internet and was generating responses to those pleading for help. owing to the flood crisis SOS calls were made online. NDRF team got real time updates on personal portal about responses through social media and phone calls. The team was asked to collect each and every message generate in social media to monitor the disaster. Strong fore team of NDRF received responses on official twitter Handle along with emails, phone calls and messages. Chennai flooded in November and December and the tweets increased substantially during this time, with about 70-75% tweets being about the Chennai flood. To have an understanding of type of information generated in Twitter pre and post disaster, tweets regarding Chennai flood were collected from 10 Nov 2015 to 25 March 2016 with the keywords related to page #Chennairains. 16 attributes were formulated to collect tweeter data for Chennai floods and the attributes are User ID, Text i.e. the content or information that the user tweets regarding a particular topic, Language, date created, Favourite Count i.e. the count of likes that the tweet has obtained, Retweet i.e. whether the text is retweeted or not, Retweet Count i.e. the number of times the text is retweeted, Username, User Screenname, date at which Twitter user created his account, User Tweets i.e. the number of tweets that the user has tweeted, User Favourites Count i.e. the count of tweets that the user liked the most, User Followers Count i.e. the number of people who are following the user, User Following Count i.e. the count of the number of people that the user follows, UserListedCount i.e. the count of number of people who have added that particular user to his or her favourites list, User Lang i.e. the language in which the text is tweeted and belongs. The tweets with 16 attributes were given as input to the weka tool.(Nair et al., 2017)

4.7.2 Case 2: Tohoku Earthquake and Tsunami, 2011

On March 11, 2011 at Tohoku on the east coast of Japan, an earthquake of magnitude 9.0 was reported by USGS. The Hashtag #j _ j _help me was used on twitter for emergency responders to identify the people who were stuck. The first letter j stands for Japan and the second one is for Jishin, which translates to “earthquake” in Japanese.

During the 2011 East Japan Earthquake and Tsunami, Twitter and Facebook served as a lifeline for directly affected individuals, a means of information sharing, and a way for people inside and outside Japan to volunteer and to provide information-based support to affected individuals. Social media became a platform to perform crucial relief functions like displaced-persons locating, support for disabled individuals, fund-raising, moral support systems, damage information provision and safety identification.(Akeuchi, 2012)

4.7.3 Case 3: Typhoon Morakot, 2009

In 2009 morakot Typhoon hit china, Philippines and japan. Among these Taiwan was worst affected.(C. S. Lee et al., 2011) In Taiwan PPT was the popular online platform where social media data was disseminated among volunteers in post disaster scenario. Morakot unofficial online platform was created under association of digital culture Taiwan at disaster center and asked public to report damage and help in affected areas in social media platform. Disaster response authorities in Taiwan set up a web portal based on twitter responses and put them in sunk with other canopy web sites integrating the disaster management departments in Taiwan.(Tseng et al., 2011)

Chapter 5 Case study area

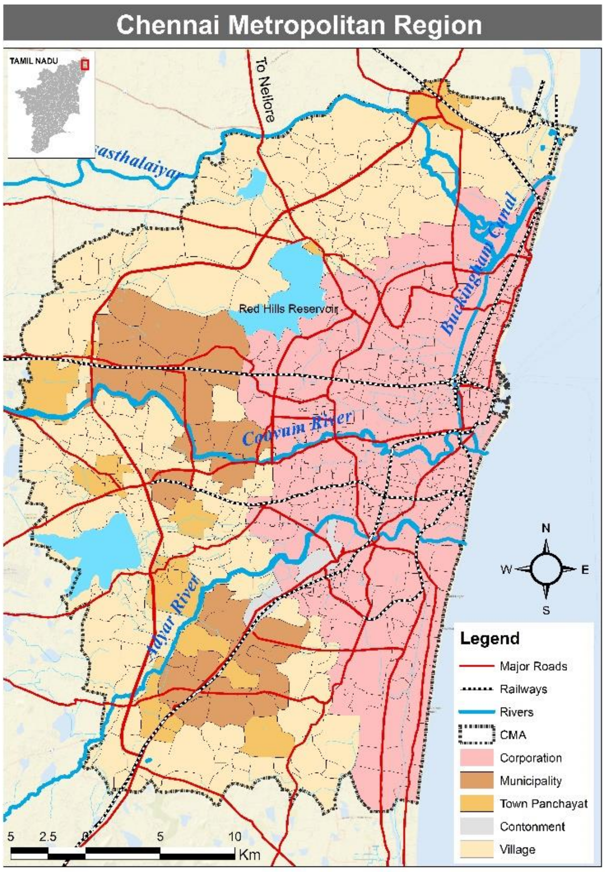
This chapter includes the introduction to the case study area details. Physical features, demographic profile, connectivity. Hydrology of the city is described in context of pluvial flooding and its impact in the past. Reviewing the city disaster management plan at zone level and tabulating the post disaster scenario according to municipal corporation.

………………………………………………………………………………………………

5.1 City Profile

Chennai formerly known as the Madras is the capital of Tamil Nadu. It is also known as "Gateway of South India". Tamil Nadu is one of the most urbanized States in the Country and Chennai city is the 4th most populous Metropolitan in India, followed by Mumbai, Kolkata and Delhi. The City is bounded by the Eastern Longitudes of 80°12' 10" and 80°18' 20" and the Northern Latitudes of 12°59' 10" and 13°08' 50" and It is called the port city owing to a long coastal line of about 43 kms from Kathivakkam in the north to Uthandi in the south, along the Bay of Bengal. The jurisdiction of the city has been expanded in the year 2011 by increasing the area from 174 sq.km to 426 Sq.km covering 3 revenue Districts - Chennai, partly Kanchipuram and Thiruvallur. Chennai city is one amongst the four cities in India that connects the country with the rest of the world through undersea fibre- optic cables, the other three being Mumbai, Kochi, and Tuticorin. (City & City, n.d.)

Figure 29 CMDA Region Map



Source : (CMDA, 2013)

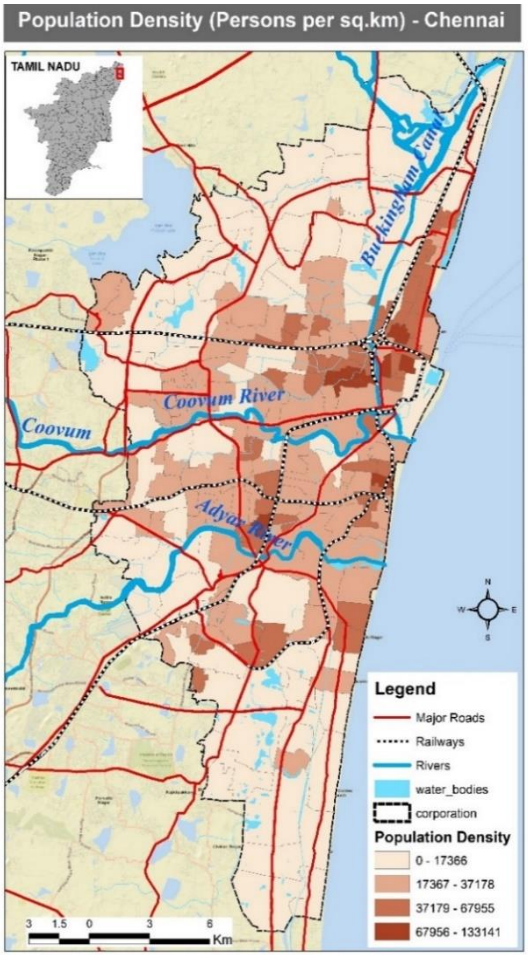
Table 4 CMA Population Projections

Sl.No	Description	Actual	Projection					Gross Density (Per/ Hectare)
		2001	2006	2011	2016	2021	2026	
1	Chennai_City	43.4	46.8	49.5	52.9	55.4	58.6	333
2	Municipalities	15.8	18.2	21.5	25.6	30.2	35.9	149
3	Town_Panchayats	3.8	4.7	5.9	7.4	9.4	12.2	78
4	Village_Panchayats	7.3	8.7	10.5	12.6	15.9	19.8	32
	CMA [Total]	70.1	78.6	88.7	99.6	111.9	125.8	105

Source: (*Chennai_Master_Plan 2026 - Map, Summary & Free Download!*, n.d.)

5.2 Chennai Demographics

Figure 30 GCC Population Density Map



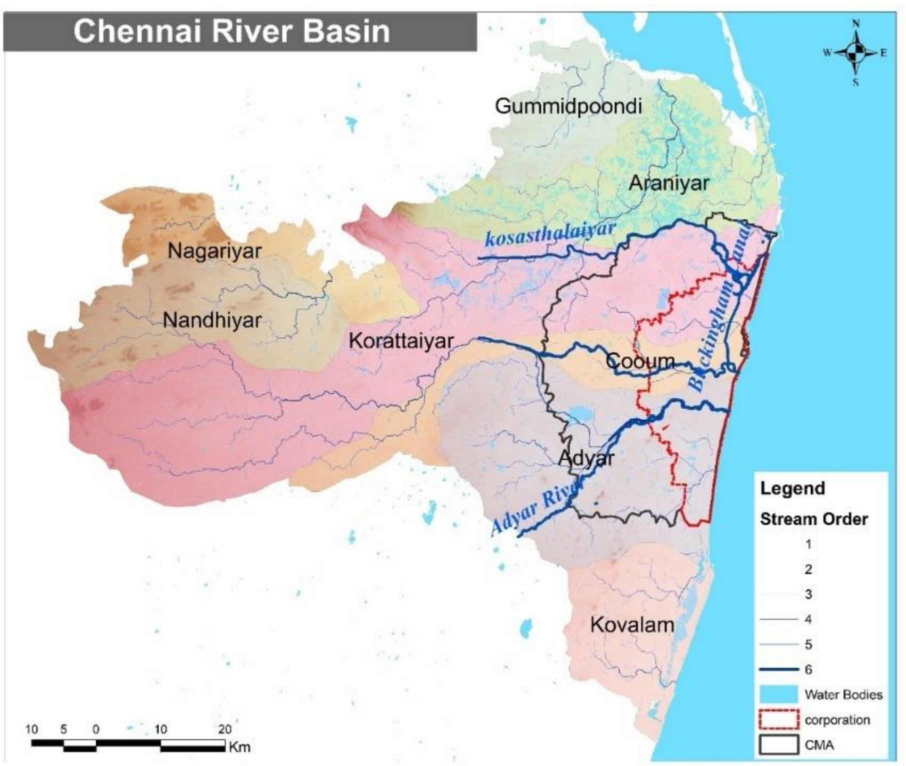
Source : (CMDA, 2013)

The average population of Chennai city is about 66.72 lakhs as per, 2011 census with a sex ratio of 989, much higher than the national average of 929. TNSCB estimates that 1.08 Lakhs EWS families reside in undeveloped slums of which 72,827 families are living in objectionable slums. TNSCB also estimated that about 37,000 families are living in unobjectionable areas and 9,687 families reside in Adyar river banks and 14,237 families in Cooum river banks. The birth and death rates in Chennai are 400 and 160 respectively. The average literacy rate being 81.27%, much higher than the national average of 72.99%.(Office of the Ministry of Home Affairs Government of India, 2011)

5.3 Chennai river basin

The Chennai Basin is situated at the north east corner of Tamilnadu. Andhra Pradesh lies on it north; Palar River basin lies on west and south the Bay of Bengal on the east were the main boundaries of this basin. The total area of the basin is 7282 km². Out of that 5542 km² lies in Tamilnadu and the rest is in Andhra Pradesh. Araniyar (covers 763 sq.km), Kosathalayar (covers 3.240 sq.km), Cooum (682 sq.km) and Adayar (857 sq.km) are the rivers of this basin. This basin group covers partly 26 blocks of Tiruthani, Thiruvallore, Saidapet, Tambaram, Ponneri, Sriperumbudur, Arakonam, and Walajapet taluks of Thiruvallore, Kanjeepuram, Chennai and Vellore districts. The major part of the basin area comes under Thiruvallore and Chennai districts (fully covered) and only a few areas covers Vellore and Kanjeepuram districts. Of the four rivers, the Adayar River carries the floodwater and drainage of Chennai and its environs. It does not have any direct irrigation and carries only the flood discharge during the northeast monsoon period for few days. Within the city limits Cooum River carries drainage and sewerage.(Basin et al., n.d.)

Figure 31 Chennai River Basin Map



Source: (Basin et al., n.d.)

5.4 Land use land cover CMA

Land use land cover Spatio temporal data for Chennai metropolitan area was extracted from Bhuvan using web management server connecting to Desktop Arc GIS. LULC 2005 and LULC 2012 raster is generated and calculated the change of land utilization in sq.km. There is rapid increase in built up both in urban and rural. Water bodies, crop lands, waste lands and agricultural plantation consistently decreased. Because of the changes in landcover built-up vulnerability of pluvial flood is High in Chennai city limits.(Bhuvan _ Thematic Data dissemination _ Free GIS Data _ OGC Services _ Clip and Ship.pdf, 2014)

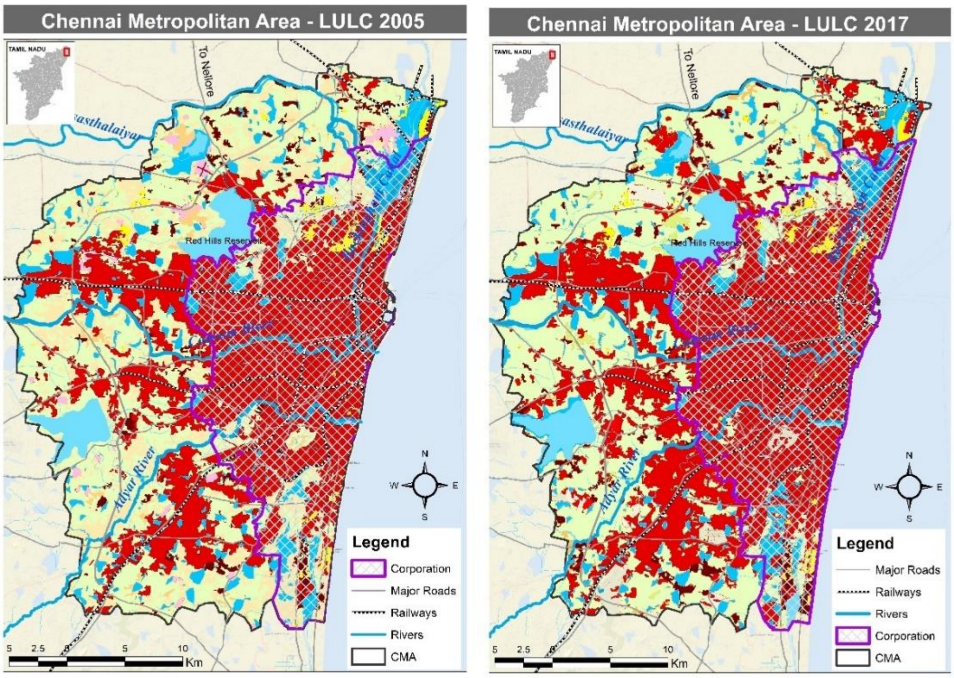


Figure 32 Bhuvan LULC Classification

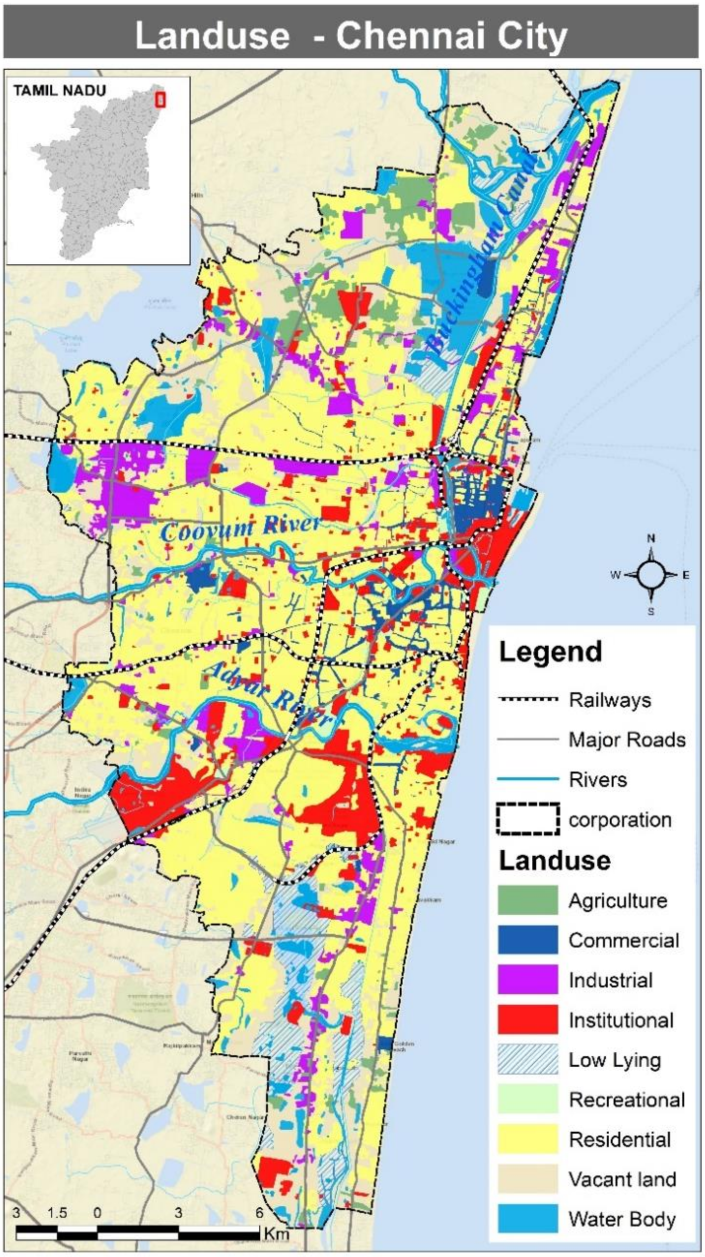
LULC - Classification	2005 (sq.km)	2017 (sq.km)	Change
Agriculture Crop Land	368	324	-43.82
Agriculture Fallow	18	48	30.53
Agriculture Plantation	17	13	-4.82
Barren / sandy Area	47	63	16.00
Waste Lands	102	24	-78.65
Grass/ Grassing	256	170	-85.60
Rural Built Up	14	48	34.37
Urban Built Up	256	411	155.50
Water Bodies	110	86	-23.79
	1188	1188	

5.5 Physical Features

5.5.1 Landuse

The Draft Master Plan 2026 for Chennai Metropolitan Area (1189 sq. km area Including Chennai city area 426 sq.km)., advocates clustering of developed areas for a compact development. It promotes corridor development, with accommodation of future expansion within or around the current developments in conjunction with the development of transport corridors. (CMDA, 2008)

Figure 33 Greater Chennai city Proposed Landuse 2026



Source : (CMDA, 2013)

Table 5 Landuse Classification

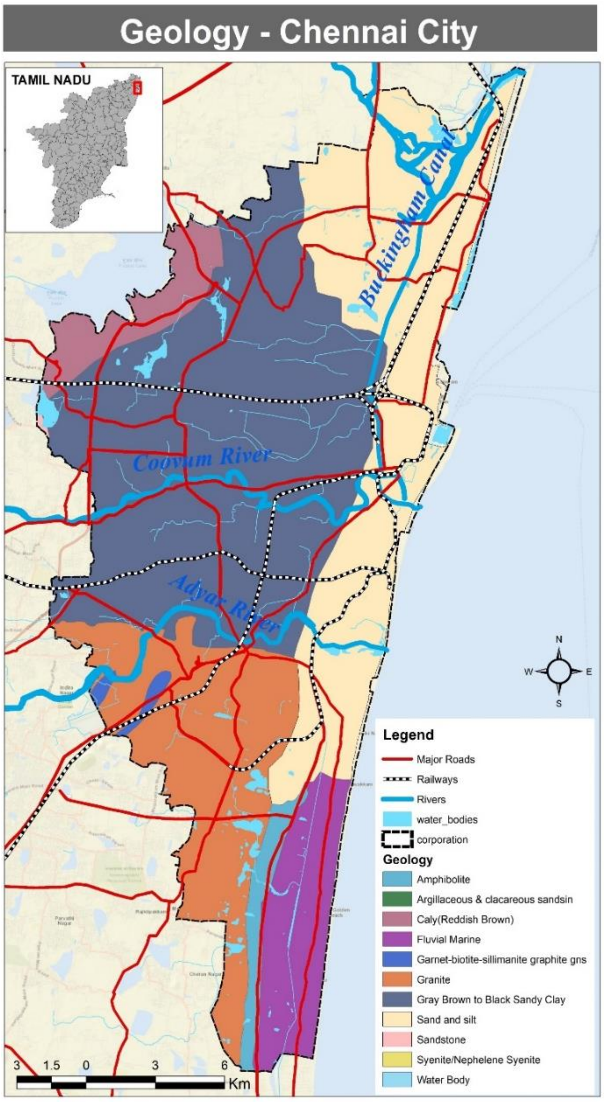
Landuse	Area (in sq.km)
Agriculture	14.02
Commercial	12.93
Industrial	30.19
Institutional	42.51
Low Lying	20.42
Recreational	2.63
Residential	210.08
Vacant land	64.49
Water Body	34.42
	431.69

Landuse and Pluvial Flooding

The flooding which is caused by heavy rains is considered as Pluvial flooding. It is associated with run off and percolation depending on landuse changes. In chennai city issues identifid related to landuse and flooding are : urban sprawl and enchroment of natural drains, degredation of canals, increase in impervious surfaces, narrowed and heavily silted minor drainage system and water bodies enchroment by squatted settlements.(Gupta & Nair, 2010)

5.5.2 Geology

Figure 34 Geology Map Greater Chennai

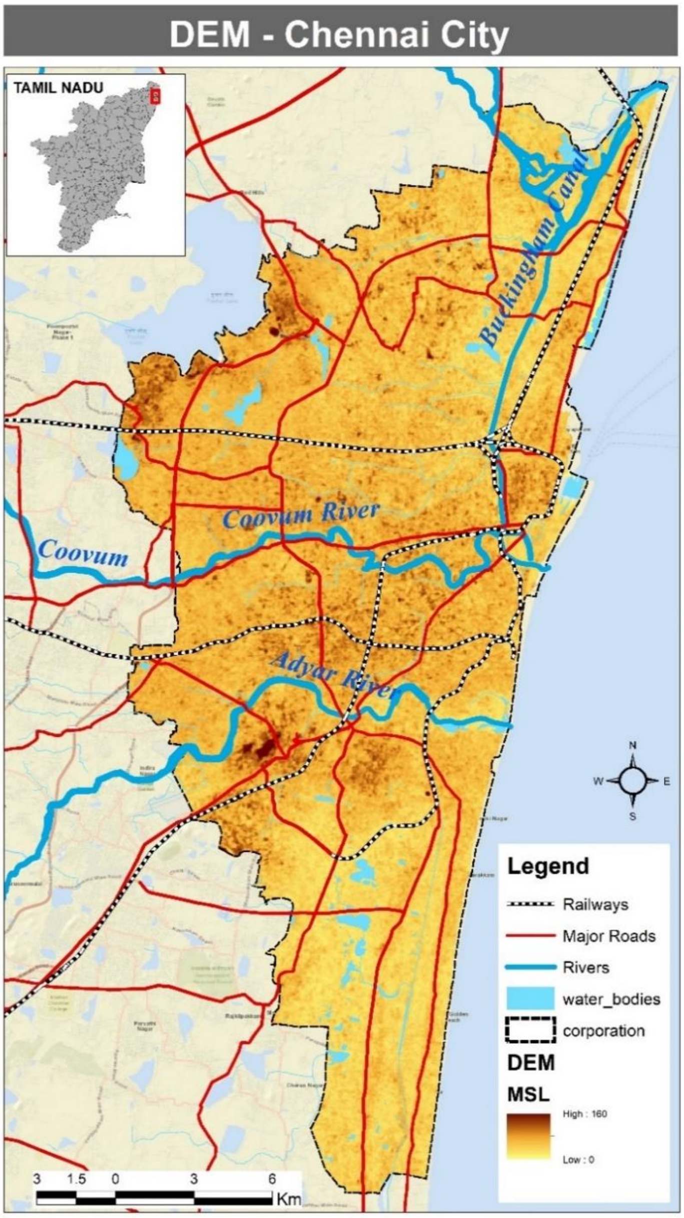


Source :Author, (CMDA, 2013)

The Geology of Chennai city mostly comprises of clay, sandstone and sediment rocks. Based on geology the city is classified into three regions as sandy area (along the river banks), clayey area and hard-rock area. The main part of Chennai city such as West-Mambalam, T-Nagar, AnnaNagar, Perambur and Virugambakkam are clayey. Hard rock areas are Adambakkam and a part in Saidapet. The sandy areas are Tiruvanmiyur, Adyar, Kottivakkam, Santhome, George Town, Tondiarpet and the rest of coastal Chennai.

The groundwater table in most of the area in Chennai is about 4 to 5m depths below the ground. The rainwater run-off in Chennai city percolates very quickly near the coastal area but in hard rock and clayey areas rainwater percolates slowly and held by the soil for a longer period. Geological Formations Includes Alluvium, Gondwana and Tertiary sediments and Archaean crystalline rocks. Charnokite are found along railway track and Adyar river bed.(Pradesh, n.d.)

Figure 35 Topography Map Greater Chennai



Source :Author, (CMDA, 2013)

5.5.3 Topography

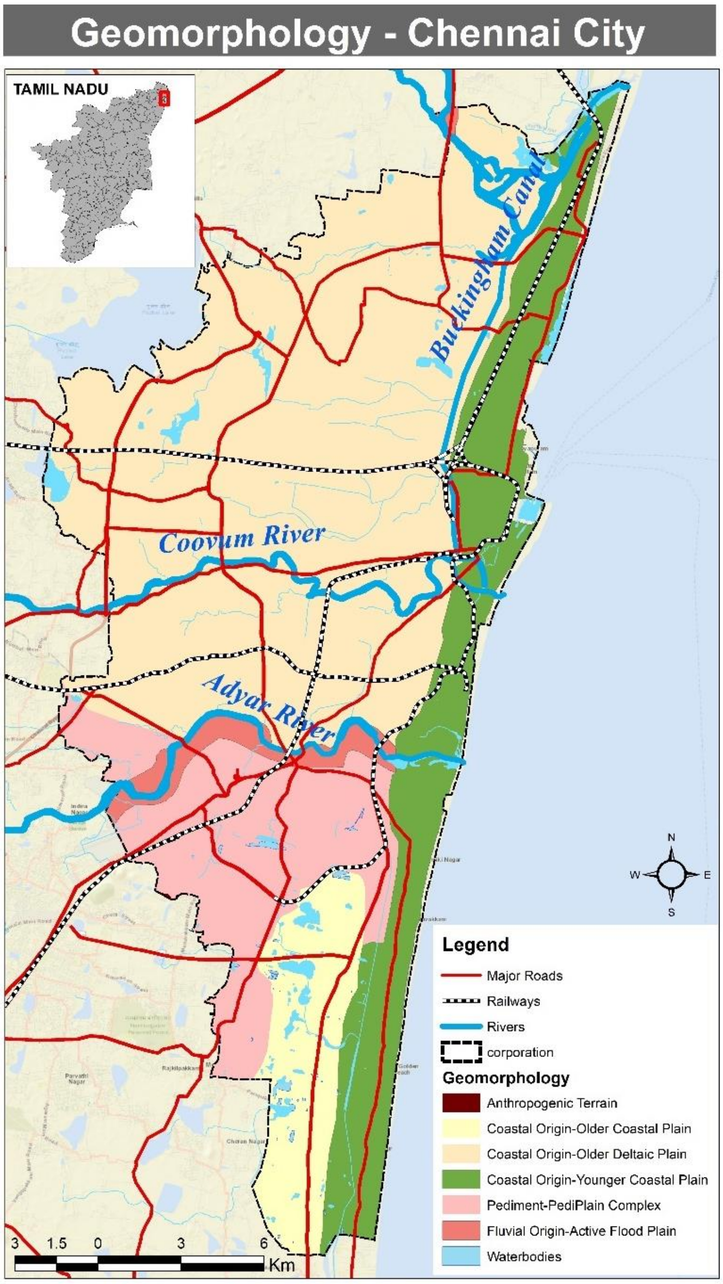
Chennai is a plain terrain, low-lying area with the land surface as flat as a pancake. The contours range from 2 m to 10 m above MSL with a few isolated hillocks in the north near red hills. It rises marginally with distance from the sea-shore but the average altitude of the city is not more than 7 meters above MSL and the average slope varies at 0.7 m per km, while some localities are just at sea level and even in minus level. Terrain slope varies from 1:5,000 to 1:10,000. The bad land topography is in pediment and lies in auroville plateau at a height of 54 meters in the west of kalappetti and in cuddalore town away from Chennai.(Chennai , 2001)

Terrain and flooding: Hydrological properties and terrain characteristics of the region are considered as major factor for floods. Flat terrain and impervious surface throughout the city caused the water to stagnation in the streets.

5.5.4 Geomorphology

Chennai's geomorphology is one of the special types, which covers both hard rock and coastal sedimentary components. The eastern part adjoining the beach and shores covers coastal geomorphic units. The inland topographical units are described as the piedmont geomorphology.(Nadu, n.d.)

Figure 36 Geomorphology Map Greater Chennai



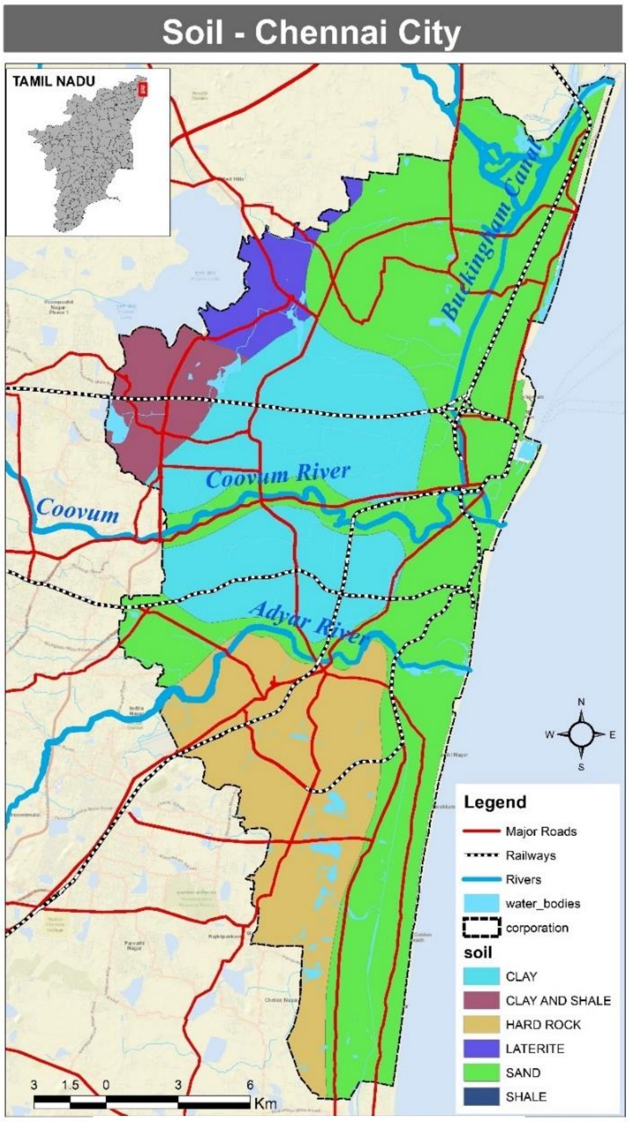
- Beaches
- Beach Ridges
- Beach terraces
- Buried Pediments
- Wash Plains
- Salt Pans
- Swamps
- Swale
- Deltaic Plains
- Deep
- Pediment and Shallow Pediment
- Buried Course & Channels
- Tertiary Uplands
- Flood Plains
- Fluvial flood plain
- Water Bodies

Source :Author, (CMDA, 2013)

5.5.5 Soil

Different types of soils control the hydrological part. Soils are classified on the basis of their color, texture, fertility and chemical combinations which includes salts, minerals and the solution effect over them. From the agricultural and groundwater point of view, the type of soil in the study area is described based on the thematic maps collected.(Series, 2008)

Figure 37 Soil map Greater Chennai



Source :Author, (CMDA, 2013)

Entisols: These alluvial soils comprise sand and sandy materials found on the beaches and at the confluence of rivers and along the side of rivers & channels. Being permeable makes these soils good storehouses of groundwater and unfit for cultivation.

Inceptisols: This major soil group consists of the red sandy to brownish clayey soil fragments spread all along the westward side of the East Coast Road. They are suitable for agriculture and hold moderate groundwater reserves.

Vertisols: The Vertisols are clayey in nature with high specific water retention capacity but poor agricultural assistance. These are found as groundmass in extreme northern portion around Gummidipoondi, Ponneri, Minjur, Madhavaram, and Manali and in the western portion of the East Coast Road around Thirupomr.

5.5.6 Green Cover and Wetlands

Green space distribution plays an imperative role in urban planning since they contribute significantly in enhancing ecological quality of metropolitan areas. Indian National Forest Policy aims at maintaining 33% of countries geographical area under forest and tree cover. Reserved Forest in Chennai city covers 2.71 Sq.km around Guindy National Park.(Meera Gandhi et al., 2015)

Table 6 Land Utilization in Chennai City

Name	Area in Sq. km	Percentage
Buildings	241.5	56.7
Parks	9.2	2.2
Tanks	35.72	8.9
Temples	0.84	0.2
Trees	72.81	17.1
Others	65.93	15.5
Total	426	100

Source:(*Chennai City Disaster Management Plan 2018*, n.d.)

The Pallikaranai marsh land serves as a buffer for rain water draining from Velachery, Pallikaranai, Perungudi, Narayanapuram lake, Kilakattalai, Velachery and from a string of upstream lakes. GCC maintains 525 parks which covers an area of 1.36 Sq.km, 763 Open Space Reservation (OSR) land which covers an area of 16.05 lakhs Sq.km.

5.5.7 Weather and Climate

The geographical location determines the weather and climate, being close to the sea and thermal equator and weather in Chennai is relatively consistent with less variation in the seasonal temperature. The weather in Chennai is mostly hot and humid. The city of Chennai has a tropical climate with 3 major seasons- summer, monsoon and winter. April to June are the hottest months in Chennai and temperature ranges from 38-42 °C.(*Chennai City Disaster Management Plan 2018*, n.d.)

Table 7 Climate Data - Chennai City

Climate data for Chennai													
Month	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Year**
Record high °C (°F)	33 (91)	37 (99)	39 (102)	43 (109)	45 (113)	43 (109)	41 (106)	40 (104)	39 (102)	39 (102)	34 (93)	33 (91)	45 (113)
Average high °C (°F)	29 (84)	31 (88)	33 (91)	35 (95)	38 (100)	38 (100)	36 (97)	35 (95)	34 (93)	32 (90)	29 (84)	29 (84)	33.3 (91.8)
Average low °C (°F)	19 (66)	20 (68)	22 (72)	26 (79)	28 (82)	27 (81)	26 (79)	26 (79)	25 (77)	24 (75)	22 (72)	21 (70)	23.8 (75)
Record low °C (°F)	14 (57)	15 (59)	17 (63)	20 (68)	21 (70)	21 (70)	22 (72)	21 (70)	21 (70)	17 (63)	15 (59)	14 (57)	14 (57)
Average precipitation mm (inches)	16.2 (0.638)	3.7 (0.146)	3.0 (0.118)	13.6 (0.535)	48.9 (1.925)	53.7 (2.114)	97.8 (3.85)	149.7 (5.894)	109.1 (4.295)	282.7 (11.13)	350.3 (13.791)	138.2 (5.441)	1,266.9 (49.878)
Source: Indian Meteorological Department													

Source: IMD

Chennai experiences South West Monsoons from June to Sep and North East Monsoon from Oct to Dec every year. South West Monsoon sets in over the extreme south western tip of the peninsula by the end of May. onset of Monsoon is characterised by a sudden spurt of rainfall activities. It progresses inland in stages and covers the India by the middle of July. The average rainfall of South West Monsoon is 439.10 mm. North East Monsoon essentially contribute the rainfall for the city. The city gets average rainfall of 760 mm during North East Monsoon. The average annual rainfall of the city is about 1,300 mm.

Winters, the most pleasant season, in Chennai are short i.e. November to February. During this time, Chennai maintains an average max and min temperature of 25°C and 21°C respectively. There are sometimes moderate rainfalls adding to the comfort.

The lowest temperature recorded in Chennai is 15.8 °C (60.4 °F) While the highest temperature is 45 °C (113 °F). The highest yearly rainfall in the city is recorded to be 2,570 mm (101 in) in 2005. The formation of Cyclones in the Bay of Bengal has great influence in the city. The winds that prevail in Chennai in between April and October are the south-westerly wind and the north-easterly winds prevail for the rest of the year.

In Chennai IMD has two rain recording stations and 4 automated warning systems, the details are

1. Regional Meteorological Department, Nungambakkam
2. Airport Meteorological Department, Mennambakkam

5.6 Hydrology of City

5.6.1 Water Bodies and Water Ways

There are 5 major water ways in the city maintained by the Public Works Department (PWD) and 31 minor canals maintained by Greater Chennai Corporation. There are 7,360 storm water drains in the city for a length of 1,894 km which are drained to major and minor drainage systems. The storm water drainage catchment of Chennai city is presented in the map. The drainage systems in the city is divided into two types.(*DRAINAGE SYSTEM IN CHENNAI METROPOLITAN*, n.d.)

i) Macro Drainage and

ii) Minor Drainage System.

Figure 38 Hydrology Map Greater Chennai

Source :Author, (CMDA, 2013)

5.6.2 Major Drainage System

Adyar and cooum are the two rivers flowing through the chennai city. Kosathalaiyar is the third river flows through the north side fringes of the city before getting into the Bay of Bengal, near Ennore. The Buckingham Canal, runs along to the coast, linking the Cooum river. OtteriNullah stream flows along north Chennai meets Buckingham canal near Basin Bridge. Several water bodies of varying size are located on the western fringes of the city. The drainage system with water bodies is given in detail in the following map.

Kosasthalaiyar River originates near Kaveripakkam near pallipet Tiruvallur district and flows in eastward direction. It bifurcates into Cooum River and Kosasthalaiyar at Keshavaramanicut. The main branch of kosasthalaiyar River then flows northwards and enters into Poondi Reservoir. Nagar River originating in Chitoor district, is a northern tributary and joins the Kosasthalaiyar River at Poondi reservoir. On the downstream side of Poondi reservoir, Kosasthalaiyar River flows through the Thiruvallur district and Chennai, and finally confluence with sea at Ennore.

Coocum River bifurcates from the main Kosasthalaiyar River at the Keshavaramanicut and flows eastwards through Kanchipuram district into Chennai and finally joins the sea near Napier Bridge. Surplus from about 75 tanks in the catchment reaches are linked with Cooum River.

Adyar River originates from two tank groups namely Pillapakkam and Kavanur in Kanchipuram District and flows through the Chennai before it joins the sea at Adyar Mouth. Surplus from about 450 tanks in the catchment, including a major tank at Chembarambakkam, reaches the Adyar River.

The Buckingham canal has been constructed more than 200 years ago, as a navigation channel. It originates close to Kakinada in Andhra Pradesh and runs along the east coast for a total length of 418 km. Its entry point into Chennai is near Athipattu village and the exit point is near Semmencheri village. While the three rivers run west to east, Buckingham canal runs north to south and connects all these three rivers at different points.

The waterways maintained by Public Works Department are detailed as follows.

Table 8 List of Waterways within Chennai city

Sl. No	List of Waterways	length in Km
1	Adyar River	42.5
2	Buckingham Canal	47.9
3	Cooum River	65
4	OtteriNullah	8.7
5	Virugambakkam & Arumbakkam Canal	6.36
	Total	**170.46**

Source: (*Chennai City Disaster Management Plan 2018*, n.d.)

Besides the above three major river system, there are 31 minor canals runing in the city linked with major drainage systems. While OtteriNullah, Kodungaiyur drain and Captain Cotton Canal drain into the Buckingham canal, Veerangalodai and Velacheri drain discharge into Pallikaranai marsh. Virugambakkam drain discharges into the Cooum River.

5.6.3 Minor Drainage Network

There are 31 minor natural drainages in the city connected with the macro drainages (16 in old city and 15 in the expanded areas) which are maintained by the Greater Chennai Corporation. In addition to this, a network of storm drains to the length of about 1,894 km discharges storm water into the sea through major / minor drains.

Table 9 Minor Canals in GCC

SL. No.	List of Water ways	Zone
1	Thamaraikulam Canal	I
2	Captain Cotton Canal	IV
3	Kodungaiyur Canal	IV
4	Link Canal	IV
5	Old Napalayan Canal	II
6	Thulasi Nagar Canal	II
7	MalaniPudur TNHB Canal	II
8	PeriyaEatchankuzhi Canal	II
9	Kadapakkam Lake Surplus Canal	II
10	Madhavaram - Manali Lake Canal	II
11	Vyasarpadi Canal	IV
12	Ekangipuram Canal	VI
13	Jawahar Canal	IV
14	T.V.S. Canal	VI
15	Nungambakkam Canal	IX
16	Nolambur Canal	VII & XI
17	AmbatturSidco Canal	VII
18	PadiKuppam Canal	VII
19	Trustpuram Canal	X
20	Mambalam Canal	IX & XIII
21	Nandanam Canal	IX

22	Chellammal College and Guindy Industrial Estate Canal	XIII
23	Reddykuppam Canal	X
24	M.G.R. Canal	X
25	Jafferkhanpet Canal	X
26	Rajbhavan Canal	XIII
27	Pallikaranai Canals	XIV
28	Nandambakkam Canal	XI & XII
29	Adambakkam Canal	XII
30	Large Lake Link Canal	XV
31	Secretariat Colony Canal	XV

Source: (*Chennai City Disaster Management Plan 2018*, n.d.)

5.7 Occurrence of Natural Calamities in Past

Chennai has been hit with various natural calamities in the past three decades, a few of which are listed as follows-

A heavy flood in 1976 leading to submergence of TNHB quarters in Kotturapuram bank of Adayar river, adding to the plight were the high tides and storm in the sea.

Again in 1996, the city was flooded by Adayar, Cooum and Kosasthalaiyar rivers. The surplus water from Poondi reservoir collapsed Karanodai Bridge at Redhills. Similarly, surplus water from Chembarambakkam Tank flooded Adayar beyond its capacity and eventually lead to extreme floods.(Kanthimathinathan, n.d.)

In 1998, Thanikachalam Nagar, a residential colony in Madhavaram was flooded by Kodungaiyur canal.

An earthquake in the Indian Ocean occurred on 26 December 2004 with the epicentre off the, Indonesia west coast of Sumatra. The shock had the moment magnitude of 9.1-9.3 and a maximum Mercalli intensity. The undersea megathrust earthquake was caused when the Indian Plate was subducted by the Burma Plate and triggered with devastating tsunami wave along the Indian Ocean. Inundated communities with waves up to 35 metres (100ft) high affected 25 kuppams in 4 coastal revenue villages affected 73,000 people, 30,000 people evacuated, 206 human loss, 9 injured and damaged 17,000 houses and other properties in Chennai city.

During 2005, Chennai experienced a sudden heavy downpour in a day. Cooum, Adayar, Otteri Nullah, B – Canal, Virugambakkam and Arumbakkam Canal flooded and people residing in the nearby areas lost their homes.

On 18th December 2005, a school situated in MGR Nagar, K.K. Nagar was chosen for distributing the relief material by the Government to the people. It was affected by severe flood, 42 people died and 37 injured due to stampede.

In the same year the city was also exposed to a Cyclone named as Fanoos with a wind speed of 85 km/hr east of Chennai on 9th December.

In 2008 again Cyclone Nisha with wind speed of 83 – 102 km/hr hit the city on 26th November and left it in ruins.

The following are some of the cyclones that hit Chennai in different periods, Cyclone Jal in 2010 with the wind speed of 100 – 110 km/hr, Cyclone Thane in 2011 with the wind speed of 140 – 165 km/hr, Cyclone Nilam in 2012 with the wind speed of 85 – 100 km/hr.

5.8 Flood

The North East Monsoon which set on 28th October 2015, initially did not shown any disastrous downpour, however things changed from the second week of November as heavy rain occurred which was recorded as the wettest month of the last 100 years with a rainfall of 1113.80.

This heavy rainfall coupled with high-speed wind resulted in 911 trees falling in the city. Way back in 1918 Chennai received 1088.40 mm in a month which remains the maximum till date. The maximum rainfall recorded as 319.60 mm on 02.12.2015 is the highest record on a single day. The day wise details of rainfall are depicted in the following table.

Table 10 Day wise details of Rainfall

Date	22-10-1969	27-10-2005	09-11-2015	13-01-2015	16-01-2015	02-12-2015	12-12-2016
Rainfall in mm	279.7	272	166.8	147	256	319.6	119.1

Source:(*Chennai City Disaster Management Plan 2018*, n.d.)

Figure 39 Flood extent - Landsat 8

Source :Author, (CMDA, 2013)

In Kanchipuram District 908 out of 912 tanks overflowed due to which excess water fed into Chembarambakkam lake through 170 drains causing flood in Adyar, Coovam rivers and link canals. Similarly, in Kovalam basin 65 tanks reached way past their maximum level and overflowed causing inundation of Sholinganallur and Perungudi areas. In Tiruvallur District surplus water released from Puzhal – Redhills, Poondi reservoir caused flood in Kosasthalaiar basin. Some of the worst affected areas are Mudichur, West Tambaram, Manapakkam, Saidapet, Jafferkhanpet, Kotturpuram etc. Many inundated areas were with more than 2 m height of water. The carrying capacity in Coovum and Buckingham canal exceeded its designed capacity. Canals including 31 minor canals in Chennai City has flooded the nearby areas. As a result, about 859 places in Chennai were inundated by heavy rainfall and Greater Chennai Corporation with Other Departments cleared water stagnation and 911 fallen trees.

City has flooded the nearby areas. As a result, about 859 places in Chennai were inundated by heavy rainfall and Greater Chennai Corporation with Other Departments cleared water stagnation and 911 fallen trees.

The satellite pictures showing nature and intensity of inundation during flood 2015 taken on 3rd December.

Table 11 Chennai city Flood Vulnerable Zones (3 Dec 2015)

Sl. No	Zone	Very High Vulnerability Above 5 Feet	High Vulnerability 3 to 5 Feet	Medium Vulnerability 2 to 3 Feet	Low Vulnerability less than 2 Feet	Total
1	I	1	7	0	6	14
2	II	0	0	0	3	3
3	III	2	6	1	1	10
4	IV	0	7	0	7	14
5	V	0	1	0	12	13
6	VI	0	0	0	36	36
7	VII	0	0	0	6	6
8	VIII	1	1	0	7	9
9	IX	5	20	0	40	65
10	X	0	9	0	0	9
11	XI	4	13	0	9	26
12	XII	0	0	0	8	8
13	XIII	24	17	0	26	67
14	XIV	0	0	0	16	16
15	XV	0	3	0	7	10
TOTAL		**37**	**84**	**1**	**184**	**306**

Source:(*Chennai City Disaster Management Plan 2018*, n.d.)

Chapter 6 Data analysis

This Chapter emphasizes the data collected, evaluation of the parameters in line with Spatial Data Analytics. In connection with this, from Frequency Ratio Model and the Parameters, a Weighted overlay analysis is done which results in the calculation of Flood susceptibility model. Correlation analysis is done in between the FR model, Post Disaster Scenario and Crowdsourced data. The resultant variables are further analyzed with multivariate regression which results in the significance of the SM.

....

6.1 Occurrence Ratio Model Analysis

The parameters of the research are evaluated by understanding the relationship with Flood. Statistical analysis is performed based on flood occurrence in each parameter Individually. Percentage of influencing is evaluated based on the occurrence by frequency in each field of particular parameter. In the present study, based on the literature, six causative factors, Geology, Geomorphology, soil, Rainfall, Land use Land cover and drainage density were considered as major Flood influencing factors and were further analysed for the purpose of the Flood susceptibility zonation mapping. A spatial database was constructed for different Flood causative factors to apply the frequency ratio models.

6.1.1 Prioritisation of Parameters

Parameters which effect the flood are identified and based on occurrence of the flood statical model is used to calculate the percentage of influence.

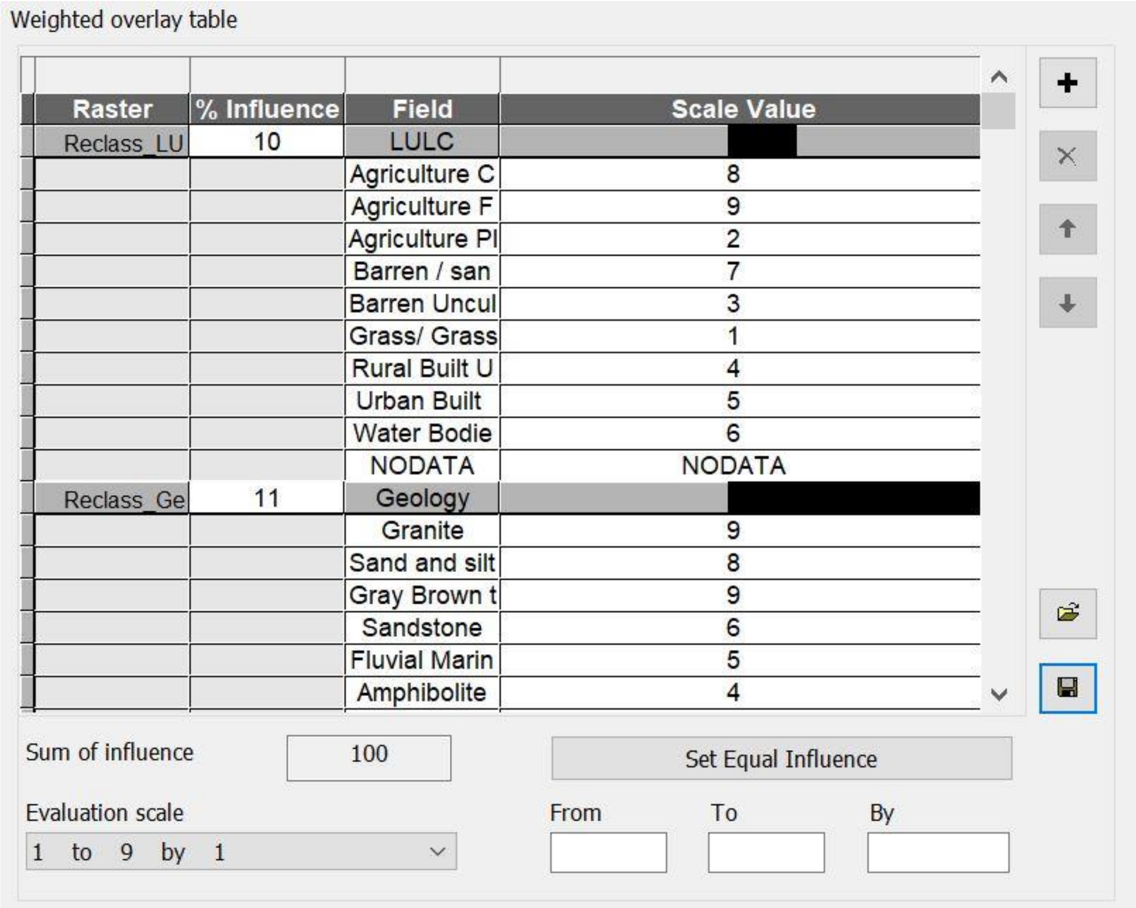
Table 12 Frequency Model Parameters Calculation

		0.21	0.23	0.20	0.21	1.00	0.16	
	Parameter	LULC	Geology	Geomorphology	Soil	Drainage	Slope	
0.21	**LULC**	1.00	1.10	0.95	1.00	4.83	0.76	
0.23	**Geology**	0.91	1.00	0.86	0.91	4.38	0.69	
0.20	**Geomorphology**	1.05	1.16	1.00	1.05	5.08	0.80	
0.21	**Soil**	1.00	1.10	0.95	1.00	4.84	0.76	
1.00	**Drainage**	0.21	0.23	0.20	0.21	1.00	0.16	
0.16	**Slope**	1.31	1.45	1.25	1.31	6.34	1.00	
		5.48	6.04	5.21	5.46	26.47	4.18	52.84

6.1.2 Parametric Analysis of Flooding Model in Arc gis

The values calculated are converted into raster datasets and the scale values of each parameter is fixed based on the occurrence of the flood. Influencing percentage of each parameter from occurrence ratio model used in weighted overlay tool in Arc gis desktop with total influencing sum method. Random outputs are generated to validate.

Figure 40 parameters weighted overlay tool


Weighted overlay table

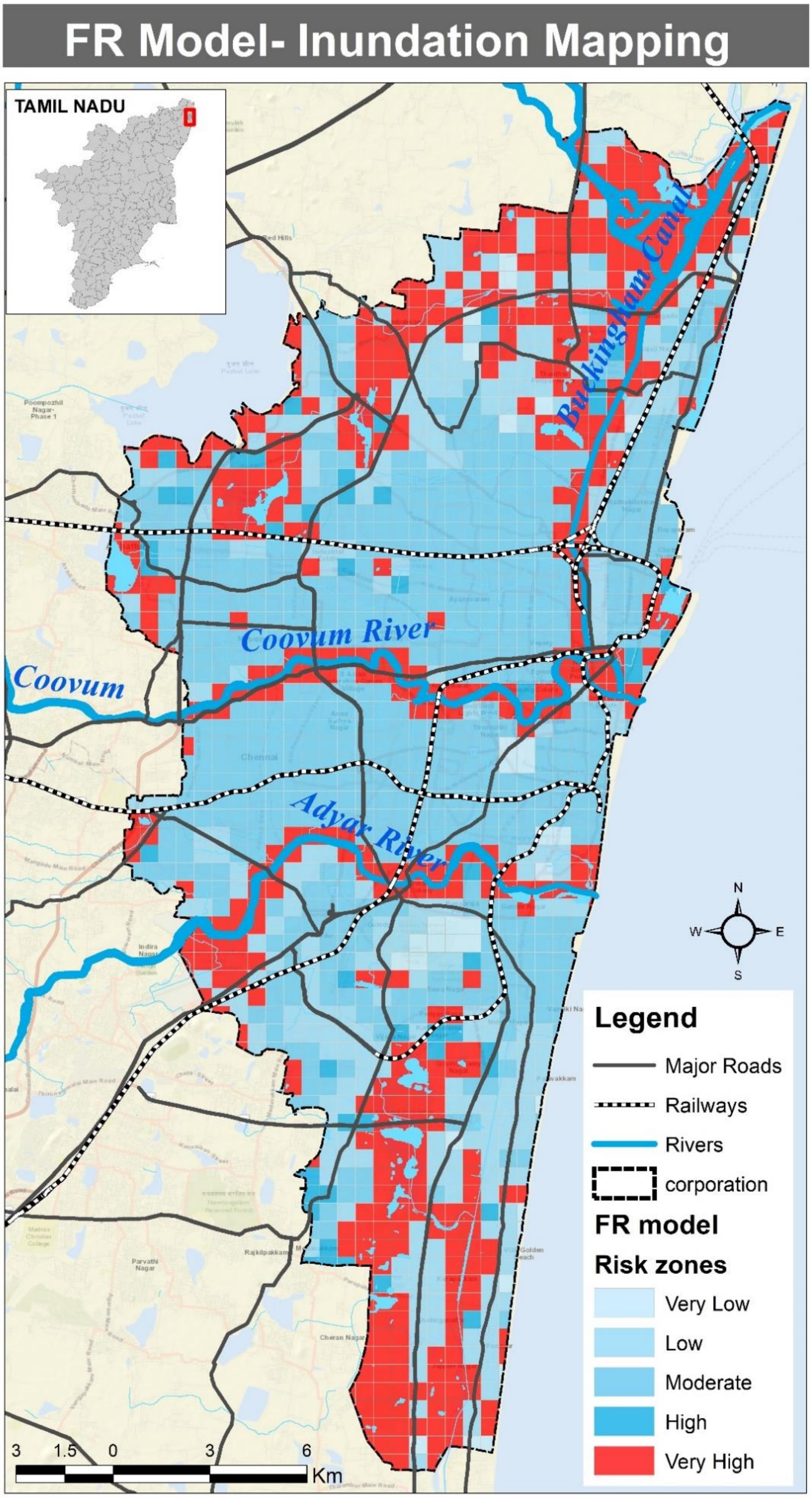
Raster	% Influence	Field	Scale Value
Reclass_LU	10	LULC	
		Agriculture C	8
		Agriculture F	9
		Agriculture Pl	2
		Barren / san	7
		Barren Uncul	3
		Grass/ Grass	1
		Rural Built U	4
		Urban Built	5
		Water Bodie	6
		NODATA	NODATA
Reclass_Ge	11	Geology	
		Granite	9
		Sand and silt	8
		Gray Brown t	9
		Sandstone	6
		Fluvial Marin	5
		Amphibolite	4

Sum of influence 100
Set Equal Influence
Evaluation scale 1 to 9 by 1
From To By

Figure 41 web-based Rainfall Data Integration in Jupyter

```
In [6]: search_rainfall = gis.content.search("Chennai_precipitation",
                                            item_type="Feature Layer",
                                            outside_org=True)
        if len(search_rainfall) >= 1:
            rainfall = search_rainfall[0]
        else:
            # if the "Chennai_precipitation" web layer does not exist
            print("Web Layer does not exist. Re-publishing...")
            # import any pandas data frame, with an address field, as a layer in our GIS
            import pandas as pds
            df = pds.read_csv('data/Chennai_precipitation.csv')
            # Create an arcgis.features.FeatureCollection object by importing the pandas dataframe with an address field
            rainfall = gis.content.import_data(df, {"Address" : "LOCATION"})
```

Figure 42 Prediction model Inundation map



6.2 Jupyter notebook IDE (Integrated development Environment)

6.2.1 Importing required libraries

Libraries in python refers to module at core level which supports the code to execute. Tweepy, NumPy and arc gis libraries are constantly used in overall code.

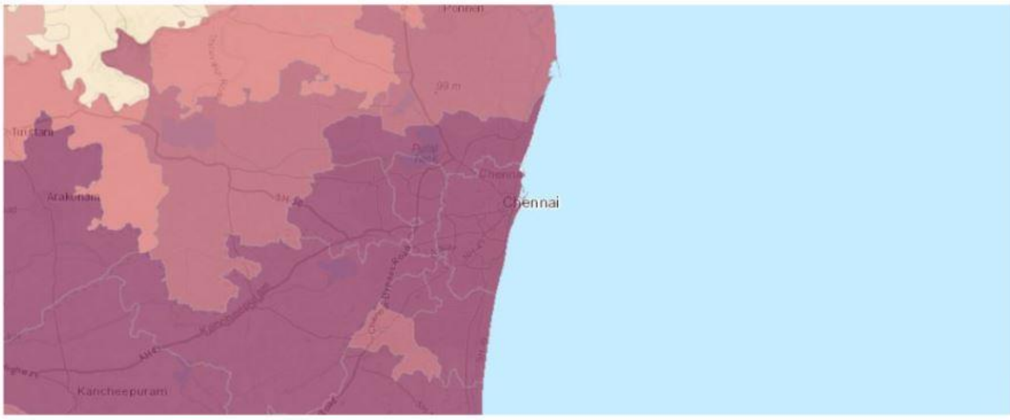
```
In [1]: import datetime

        %matplotlib inline
        import matplotlib.pyplot as pd
        from IPython.display import display, YouTubeVideo

        import arcgis
        from arcgis.gis import GIS
        from arcgis.features.analyze_patterns import interpolate_points
        from arcgis.geocoding import geocode
        from arcgis.features.find_locations import trace_downstream
        from arcgis.features.use_proximity import create_buffers

        gis = GIS(profile = "Madhukarchary")
```

6.2.2 Visualizing spatial data of Chennai city

Arc gis literary is imported in jupyter notebook to visualize Geojson format of Chennai city and population density is analysed at ward level.

```
In [67]: map = gis.map("Chennai")
         map
```

Out[67]:




```
In [4]: chennaipop = gis.content.search("Chennai_Population",
                                        item_type="Feature Layer",
                                        outside_org=True)[0]
        chennaipop
```

Out[4]:

Chennai_Population

Feature Layer Collection by arcgis_python
Last Modified: February 23, 2019
0 comments, 9 views

```
In [58]: map.add_layer(chennaipop, {"renderer":"ClassedColorRenderer",
                                    "field_name": "TOTPOP_CY",
                                    "opacity":0.7})
```

6.2.3 Interpolation of Rainfall stations Data

Rainfall stations locations were collected from IMD web site and interpolated the data in the attributes using arc gis content page. Output of the code generates rainfall map from IMD data.

In [6]:
```
search_rainfall = gis.content.search("Chennai_precipitation",
                                     item_type="Feature Layer",
                                     outside_org=True)
if len(search_rainfall) >= 1:
    rainfall = search_rainfall[0]
else:
    # if the "Chennai_precipitation" web layer does not exist
    print("Web Layer does not exist. Re-publishing...")
    # import any pandas data frame, with an address field, as a layer in our GIS
    import pandas as pds
    df = pds.read_csv('data/Chennai_precipitation.csv')
    # Create an arcgis.features.FeatureCollection object by importing the pandas dataframe with an address field
    rainfall = gis.content.import_data(df, {"Address" : "LOCATION"})
```

```
Web Layer does not exist. Re-publishing...
```

6.2.4 Visualization of weather stations rainfall data

Weather stations data for rainfall is visualized on open street maps in jupyter notebook to create near real time maps.

In [8]:
```
map2.add_layer(rainfall, {"renderer":"ClassedSizeRenderer",
                          "field_name":"RAINFALL" })
```

Here we used the **smart mapping** capability of the GIS to automatically render the data with proportional symbols.

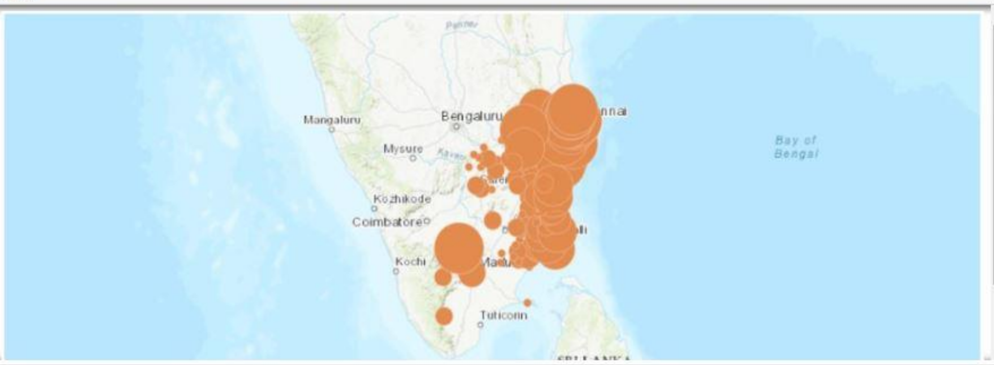
Spatial Analysis

Rainfall is a continuous phenonmenon that affects the whole region, not just the locations of the weather stations. Based on the observed rainfall at the monitoring stations and their locations, we can interpolate and deduce the approximate rainfall across the whole region. We use the **Interpolate Points** tool from the GIS's spatial analysis service for this.

The Interpolate Points tool uses empirical Bayesian kriging to perform the interpolation.

In [68]:
```
map2 = gis.map("Tamil Nadu, India")
map2
```

Out[68]:



In [8]:
```
map2.add_layer(rainfall, {"renderer":"ClassedSizeRenderer",
                          "field_name":"RAINFALL" })
```

Here we used the **smart mapping** capability of the GIS to automatically render the data with proportional symbols.

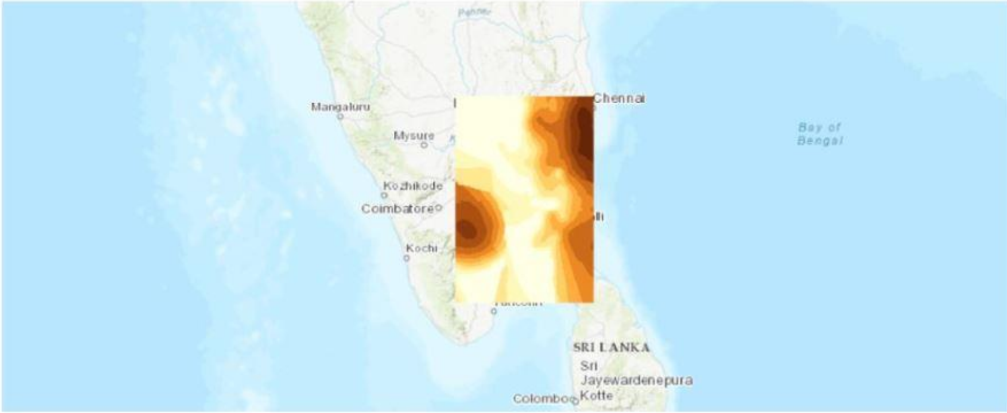
Spatial Analysis

Rainfall is a continuous phenonmenon that affects the whole region, not just the locations of the weather stations. Based on the observed rainfall at the monitoring stations and their locations, we can interpolate and deduce the approximate rainfall across the whole region. We use the **Interpolate Points** tool from the GIS's spatial analysis service for this.

The Interpolate Points tool uses empirical Bayesian kriging to perform the interpolation.

```
In [9]: interpolated_rf = interpolate_points(rainfall, field='RAINFALL')
```

```
In [69]: intmap = gis.map("Tamil Nadu")
         intmap
Out[69]:
```



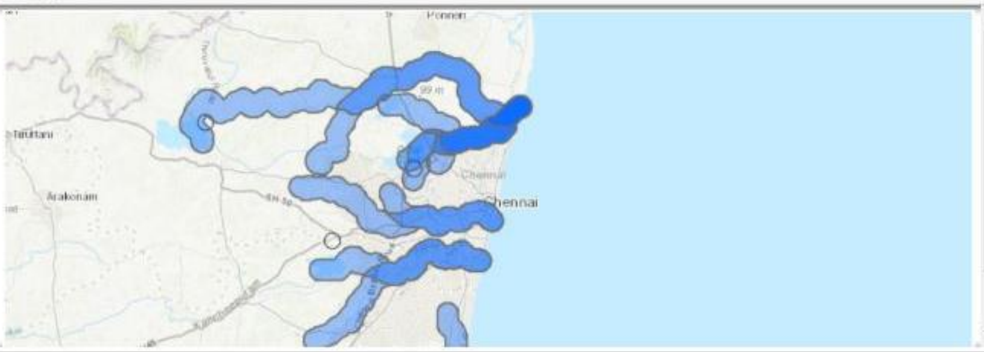

```
In [11]: intmap.add_layer(interpolated_rf['result_layer'])
```

We see that rainfall was most severe in and around Chennai as well some parts of central Tamil Nadu.

6.2.5 Geocoding Water Bodies and major lakes

Water bodies and major lakes are identified from the open street maps and a buffer is created where the reservoirs had huge outflow during the monsoon. To plot the locations geocoding tool from Arc gis desktop is used.

```
In [70]: lakemap = gis.map("Chennai")
         lakemap.height='450px'
         lakemap
Out[70]:
```



Let's have look at the major lakes and water reservoirs that were filled to the brim in Chennai due the rains. We plot the locations of some of the reservoirs that had a large outflow during the rains:

To plot the locations, we use geocoding tools from the `tools` module. Your GIS can have more than 1 geocoding service, for simplicity, the sample below chooses the first available geocoder to perform an address search

```
In [14]: lakemap.draw(geocode("Chembarambakkam, Tamil Nadu")[0],
                      {"title": "Chembarambakkam", "content": "Water reservoir"})
         lakemap.draw(geocode("Puzhal Lake, Tamil Nadu")[0],
                      {"title": "Puzhal", "content": "Water reservoir"})
         lakemap.draw(geocode("Kannampettai, Tamil Nadu")[0],
                      {"title": "Poondi Lake ", "content": "Water reservoir"})
```

```
In [19]: lakemap.add_layer(floodprone_buffer)
```

```
In [18]: floodprone_buffer = create_buffers(downstream, [ 1 ], units='Miles')
```

```
In [15]: search_results = gis.content.search("Chennai_lakes",
                                             item_type="Feature Layer",
                                             outside_org=True)
         search_results

Out[15]: [<Item title:"Chennai Lakes OverFlow Path" type:Feature Layer Collection owner:Balaganesh99>,
          <Item title:"Possible Flood Zones" type:Feature Layer Collection owner:Balaganesh99>,
          <Item title:"Chennai_lakes" type:Feature Layer Collection owner:yjiang_geosaurus>]
```

```
In [17]: downstream = trace_downstream(chennai_lakes)
         downstream.query()

Out[17]: <FeatureSet> 11 features
```

```
In [16]: chennai_lakes = search_results[2]
         chennai_lakes

Out[16]:
```

Chennai_lakes

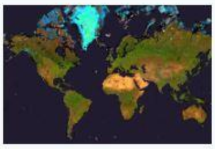
Feature Layer Collection by yjiang_geosaurus
Last Modified: February 22, 2019
0 comments, 5 views

6.2.6 Automation of satellite data sources to detect flood

Automation process for the remote sensing data in jupyter notebook is performed using math algebra tools from arc gis. To analyze the extent of flood to identify the vulnerable zones Landsat 8 is used with the band combinations (5,4,3) Normalized Difference Water Index is calculated.

In [20]:
```
def exact_search(my_gis, title, owner_value, item_type_value, max_items_value=20):
    final_match = None
    search_result = my_gis.content.search(query= title + ' AND owner:' + owner_value,
                                          item_type=item_type_value, max_items=max_items_value, outside_org=True)

    if "Imagery Layer" in item_type_value:
        item_type_value = item_type_value.replace("Imagery Layer", "Image Service")
    elif "Layer" in item_type_value:
        item_type_value = item_type_value.replace("Layer", "Service")

    for result in search_result:
        if result.title == title:
            final_match = result
            break
    return final_match

ls_water = exact_search(gis, 'Landsat GLS Multispectral', 'esri', 'Imagery Layer')
ls_water
```

Out[20]:



Landsat GLS Multispectral
Landsat GLS 30 and 60m Multispectral 8 band images with visual renderings and indices.

Imagery Layer by esri
Last Modified: May 03, 2018
0 comments, 7,363 views

In [ ]:
```
ls_water_lyr = ls_water.layers[0]
```

In [22]:
```
from arcgis.geocoding import geocode
area = geocode("Tamil Nadu, India", out_sr=ls_water_lyr.properties.extent.spatialReference)[0]
ls_water_lyr.extent = area['extent']
```

In the cell below, we will use a band combination [5,4,3] (a.k.a. mid-IR (Band 5), near-IR (Band 4) and red (Band 3)) of Landsat to provide definition of land-water boundaries and highlights subtle details not readily apparent in the visible bands alone. The reason that we use more infrared bands is to locate inland lakes and streams with greater precision. Generally, the wetter the soil, the darker it appears, because of the infrared absorption capabilities of water.

In [23]:
```
# data source option
from arcgis.raster.functions import stretch, extract_band
target_img_layer = stretch(extract_band(ls_water_lyr, [5,4,3]),
                           stretch_type="percentclip", gamma=[1,1,1], dra=True)
```

Use the cell below to filter imageries based on the temporal conditions, and export the filtered results as local images, then show comparatively with other time range. You can either use the where clause e.g. `where="(Year = " + str(start_year) + ")",` or use the temporal filter as shown below.

In [24]:
```
import pandas as pd
from arcgis import geometry
import datetime as dt

def filter_images(my_map, start_year, end_year):
    selected = target_img_layer.filter_by(where="(Category = 1) AND (CloudCover <=0.2)",
                                          time=[dt.datetime(start_year, 1, 1), dt.datetime(end_year, 1, 1)],
                                          geometry=arcgis.geometry.filters.intersects(ls_water_lyr.extent))
    my_map.add_layer(selected)

    fs = selected.query(out_fields="AcquisitionDate, GroupName, Month, DayOfYear, WRS_Row, WRS_Path")
    tdf = fs.sdf
    return tdf
```

In [1]:
```
satmap1 = gis.map("Pallikaranai, Tamil Nadu, India", 13)
df = filter_images(satmap1, 1991, 1992)
df.head()
```

Out[1]:

	AcquisitionDate	DayOfYear	GroupName	Month	OBJECTID	SHAPE	Shape_Area	Shape_Length	WRS_Path	WRS_Row
0	1991-08-25	237	p142r051_5x19910825	8	33340	{"rings": [[[9024389.361699998, 1418416.124200...	3.157366e+10	711646.676375	142	51
1	1991-01-29	29	p142r052_5x19910129	1	33341	{"rings": [[[8988705.9855, 1253944.3066000007]...	3.199999e+10	716418.700482	142	52
2	1991-08-25	237	p142r052_5x19910825	8	33342	{"rings": [[[8988407.6618, 1253748.178100001],...	3.118276e+10	707233.012677	142	52
3	1991-01-29	29	p142r053_5x19910129	1	33343	{"rings": [[[8953190.746100001, 1089696.012499...	3.170226e+10	713084.857110	142	53
4	1991-04-10	100	p143r051_5x19910410	4	33378	{"rings": [[[8854469.6882, 1418030.1556000002]...	3.231922e+10	719956.694300	143	51

Then search for satellite imageries intersecting with the area of interest at 2009.

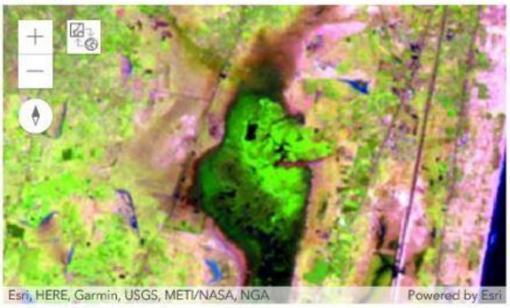
In [2]:
```
satmap2 = gis.map("Pallikaranai, Tamil Nadu, India", 13)
df = filter_images(satmap2, 2009, 2010)
df.head()
```

Out[2]:

	AcquisitionDate	DayOfYear	GroupName	Month	OBJECTID	SHAPE	Shape_Area	Shape_Length	WRS_Path	WRS_Row
0	2009-02-15	46	L5142051_05120090215	2	2711	{"rings": [[[9060430.627500001, 1532858.459600...	3.410318e+10	739655.852473	142	51
1	2009-02-15	46	L5142052_05220090215	2	2712	{"rings": [[[9024333.732299998, 1367472.252799...	3.374737e+10	735793.784840	142	52
2	2009-03-03	62	L5142053_05320090303	3	2713	{"rings": [[[8988128.4335, 1202891.985199999],...	3.349556e+10	733068.511329	142	53
3	2009-07-25	206	L5142054_05420090725	7	2714	{"rings": [[[8950658.1897, 1042055.4193999991]...	3.330632e+10	731037.483585	142	54
4	2009-01-21	21	L5143051_05120090121	1	2747	{"rings": [[[8889603.408100002, 1533573.856199...	3.409762e+10	739594.538531	143	51

In [3]:
```
from ipywidgets import *

satmap1.layout=Layout(flex='1 1', padding='10px', height='300px')
satmap2.layout=Layout(flex='1 1', padding='10px', height='300px')

box = HBox([satmap1, satmap2])
box
```

Out[3]:



6.2.7 Geocoding Flood relief camps

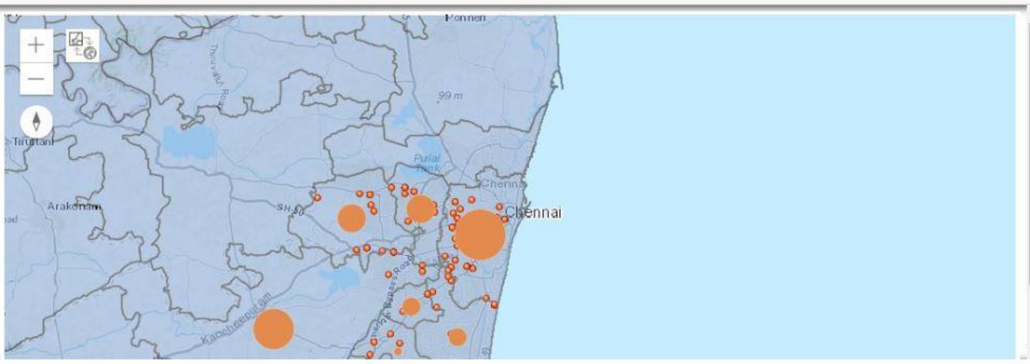
Flood Relief Camps

To provide emergency assistance, the Tamil Nadu government has set up several flood relief camps in the flood affected areas. They provide food, shelter and the basic necessities to thousands of people displaced by the floods. The locations of the flood relief camps was obtained from http://cleanchennai.com/floodrelief/2015/12/09/relief-centers-as-on-8-dec-2015/ and published to the GIS as a layer, that is visualized below:

In [28]:
```
relief_centers = gis.content.search("Chennai Relief Centers")[0]
```

In [76]:
```
reliefmap = gis.map("Chennai")
reliefmap
```

Out[76]:



In [73]:
```
reliefmap.add_layer(chennaipop, {"opacity":0.5})
```

In [74]:
```
reliefmap.add_layer(relief_centers)
```

In [32]:
```
relief_data = relief_centers.layers[0].query().sdf
relief_data.head()
```
Out[32]:

	Contact_No	Division__	FID	F_Locations	No_of_Cent	No_of_fami	No_of_pers	SHAPE	Sl_No_	SymbolID	Zone_____
0	Balamurali, 9445190311	10	1	Poonthotam School, Chennai	10	65	200	{"x": 8919695.334199999, "y": 1464332.82629999...	1		I
1	Jayakumar, 9445190302	2	2	St.Joseph church community Hall, Chennai	0	200	600	{"x": 8936283.704100002, "y": 1469202.8202, "s...	2		
2	Jayakumar, 9445190302	2	3	Nehru Nagar chennai Middle school, Chennai	0	75	250	{"x": 8916764.954599999, "y": 1450941.69069999...	3		
3	Shanmugam, 9445190301	7	4	Kalaimagal School, Chennai	0	15	50	{"x": 8924034.069200002, "y": 1462457.79919999...	4		
4	Rameshkumar, 9445190304	4	5	Ramanathapuram School, Chennai	0	100	300	{"x": 8919695.334199999, "y": 1464332.82629999...	5		

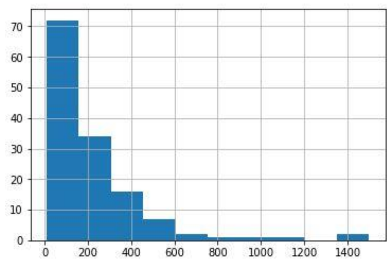
In [33]:
```
relief_data['No_of_pers'].sum()
```
Out[33]: 31478

In [34]:
```
relief_data['No_of_pers'].describe()
```
Out[34]:
```
count     136.000000
mean      231.455882
std       250.334202
min        10.000000
25%        60.000000
50%       150.000000
75%       300.000000
max      1500.000000
Name: No_of_pers, dtype: float64
```

In [39]:
```
aggr_lyr = res['aggregated_layer']
```

In [75]:
```
reliefmap.add_layer(aggr_lyr, { "renderer": "ClassedSizeRenderer",
                               "field_name":"SUM_No_of_pers"})
```

In [35]:
```
relief_data['No_of_pers'].hist()
```
Out[35]: <matplotlib.axes._subplots.AxesSubplot at 0x29b5361a748>



In [37]:
```
chennai_pop_featurelayer = chennaipop.layers[0]
```

In [38]:
```
res = arcgis.features.summarize_data.aggregate_points(  relief_centers,
                                                        chennai_pop_featurelayer,
                                                        False,
                                                        ["No_of_pers Sum"])
```

In [41]:
```
df = aggr_lyr.query().sdf
df.head()
```
Out[41]:

	AVGHHSZ_CY	Age_15_30	Age_30_45	Age_45_60	AnalysisArea	ENRICH_FID	ENRICH_F_1	FEMALES_CY	FID	HasData	...	Wtr_Pump	Wtr_River	Wtr_S
0	3.90	104365	86739	53977	13.622797	33	33	330353	1	1	...	2959	420	
1	3.91	89395	74300	46237	24.148443	25	25	195680	2	1	...	2024	107	
2	3.92	159311	140362	86549	32.373704	30	30	463066	3	1	...	4144	58	
3	3.92	245989	204995	127679	49.687276	26	26	287041	4	1	...	11674	309	
4	3.93	177140	155732	95889	63.265503	29	29	337026	5	1	...	3061	248	

5 rows × 40 columns

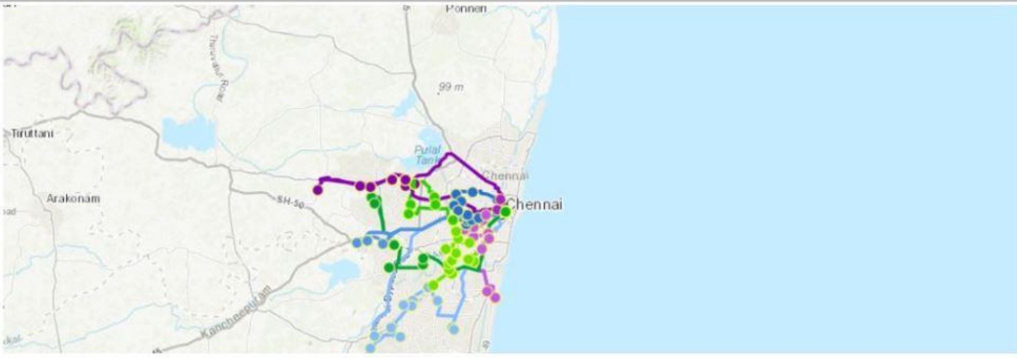
```
In [44]: df = aggr_lyr.query().sdf

         df2 = df[['NAME', 'SUM_No_of_pers']]
         df2.set_index('NAME', inplace=True)
         df2
```

Out[44]:

NAME	SUM_No_of_pers
Alandur	1256
Tambaram	950
Ambattur	1886
Shollingalanur	1632
Poonamallee	2785
Chennai	19765
Chengalpattu	49
Sriperumbudur	3155

```
In [77]: routemap = gis.map("Chennai")
         routemap
```

Out[77]:



6.3 Data Mining Twitter Page – Chennairains.org

Twitter credentials (access token keys) are executed in jupyter notebook to mine data from twitter pages where public responses were captured. The captured data was converted to locational data Geojson format and visualized in the notebook.

```
In [2]: from textblob import TextBlob
        import sys, tweepy
        import matplotlib.pyplot as plt

In [3]: def percentage(part, whole):
            return 100 * float(part)/float(whole)

In [4]: consumerkey ="FZo5PuU9n08AkDALuKVO7Kcin"
        consumersecret ="lpZvfHSuC0ttmii31Qsb60E59Y0BqXmOhcOyBQVHnBYKUCYoaH"
        accessToken ="1166310141790048262-HMZaS8FllchVOf3znjwXPSof7zdTNC"
        accessTokenSecret ="i2voP37V3eiIuEd4ZdLoi7LRQvoINXsgycf9VN6F2q1hr"

In [5]: auth = tweepy.OAuthHandler(consumerkey, consumersecret)
        auth.set_access_token(accessToken, accessTokenSecret)
        api = tweepy.API(auth)

In [6]: searchTerm = input("Enter keyword/hashtag to search about: ")
        noOfsearchTerms =int(input("Enter how many tweets to analyze: "))
        tweets = tweepy.Cursor(api.search, q=searchTerm, ).items(noOfsearchTe

        Enter keyword/hashtag to search about: apple
        Enter how many tweets to analyze: 10

In [7]: #chennaifloodshelp= 0
        negative = 0
        netural = 0
        polarity = 0

In [8]: for tweet in tweets:
            print(tweet.text)
            analysis = TextBlob (tweet.text)

In [2]: import pandas as pd
        chennai_floods=pd.read_csv("G:\jupyter notebook\chennai_dec_1-2.csv")

In [3]: chennai_floods
```

6.4 Data validation

Although social media data which is collected from twitter pages about street level inundation is structured it requires a validation with conventional inundation mapping techniques. To validate the social media data sets satellite imageries and municipal official records are used. Firstly Landsat 8 imagery is processed and Normalized Difference Water Index is calculated. To validated it with twitter inundation map 30meters x 30meters pixels are created with in Chennai city boundary and reason for creation 30sq.m pixel is because the open sourced available imagery resolution is 30meters.after getting both the maps at same resolution post disaster captured by municipal authorities are recorded in the same pixels. Municipal official records include on ground verification after the flooding and high-resolution imageries. City disaster management team made ground truthing survey in 2016 to collected the data regarding the inundation at ward level and streets inundation level. If the data which is unstructured or semi structured needs machine learning models and applications which can convert those data sets into structured and validated in near real time. All these data sets are correlated with each other in order to validate the inundation results. correlation matrix shows, social media is considered as the dependent variable i.e., to check the significance of the Crowdsourcing. To structure the crowdsourcing data in simple way the pages and responses in the social media platforms need to be structured while collecting the data.

Table 13 Correlation Results of Inundation maps

	SM	Post disaster	FR model	Landsat
SM	**1**			
Post Disaster	**-0.11**	**1**		
FR model	**-0.14**	**0.66**	**1**	
Landsat	**0.61**	**0.38**	**0.41**	**1**

Table 14 Pixel based crowdsourced Inundation map

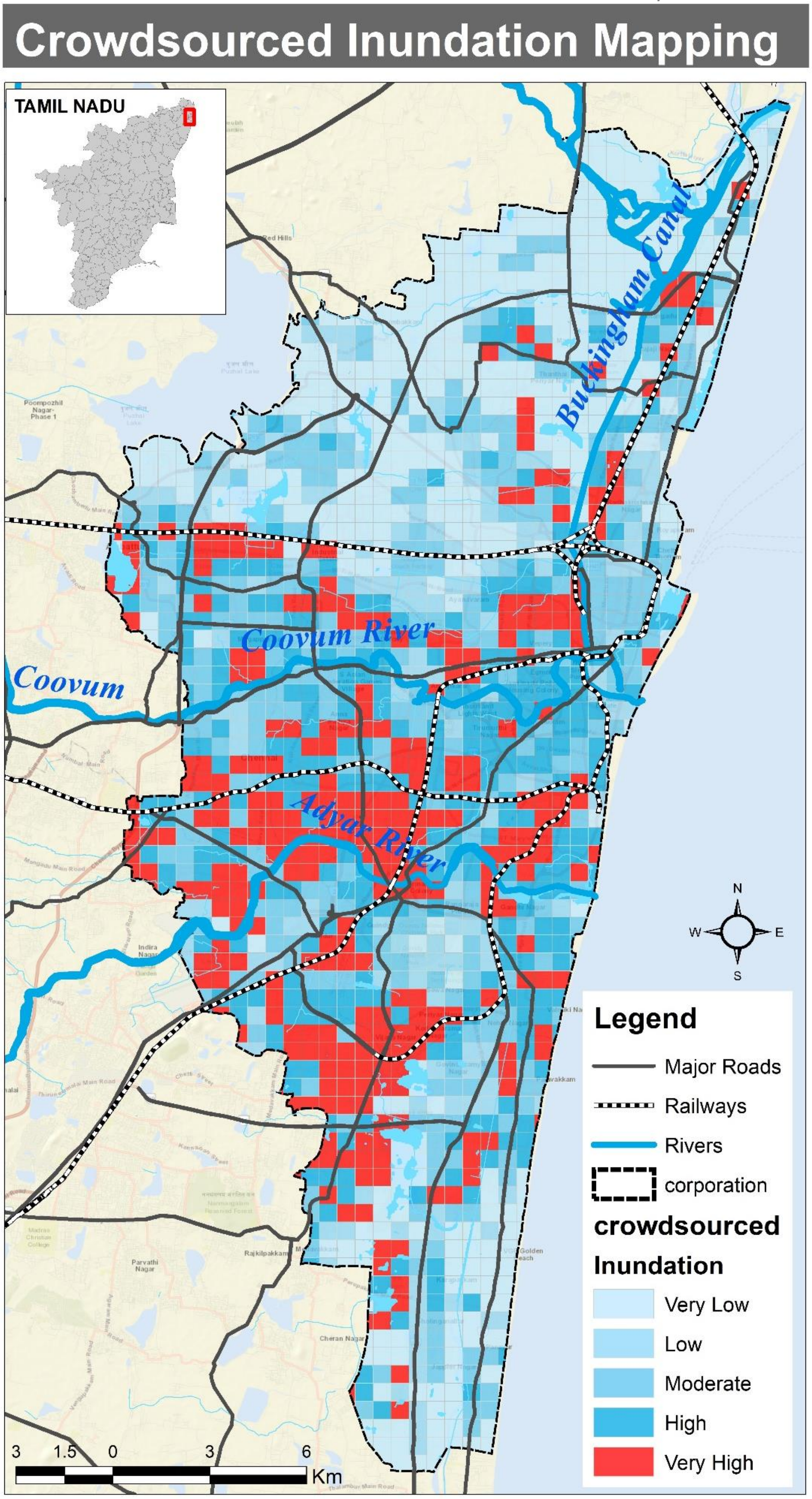


Table 15 pixel-based Landsat Inundation map

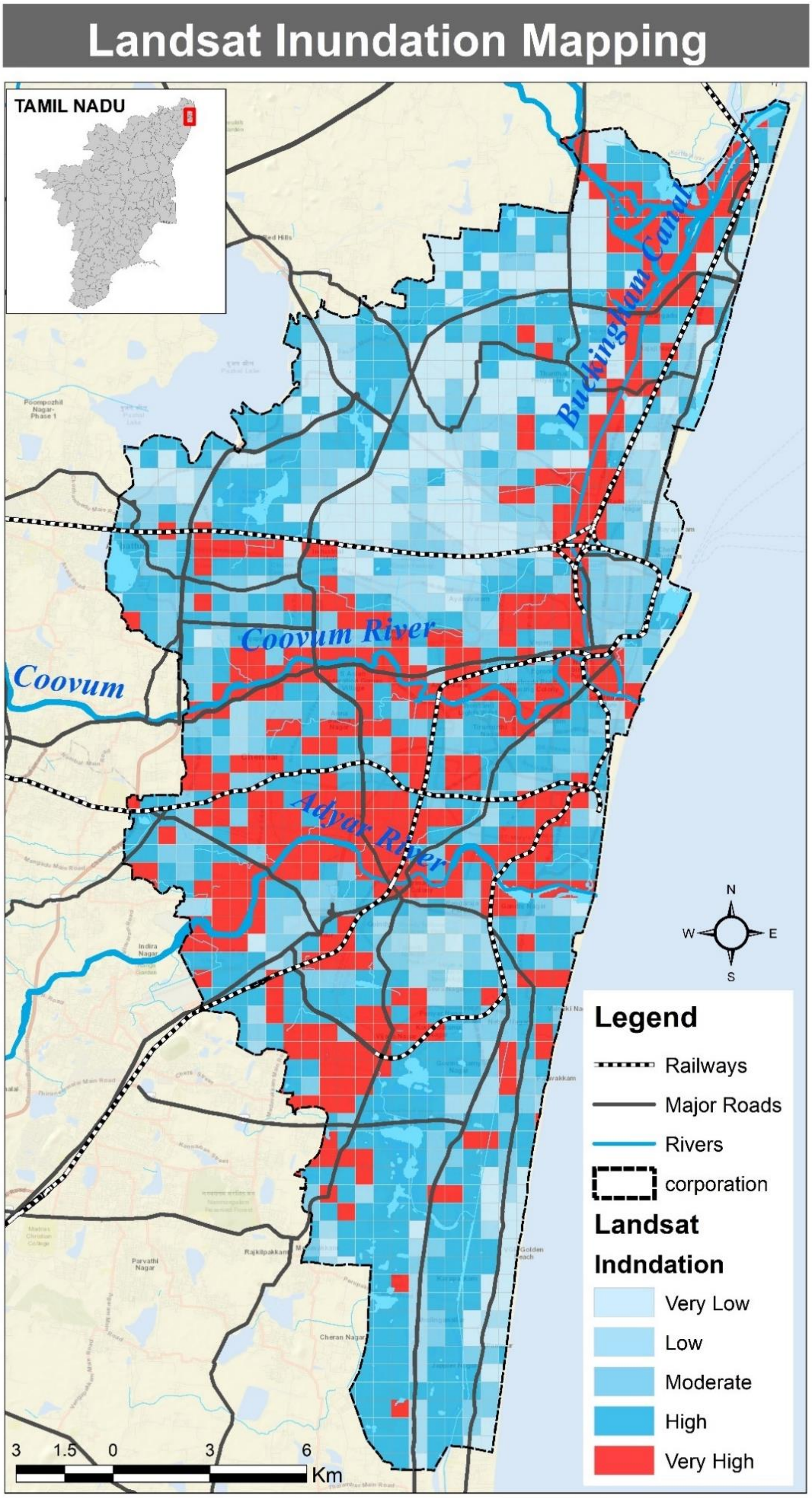
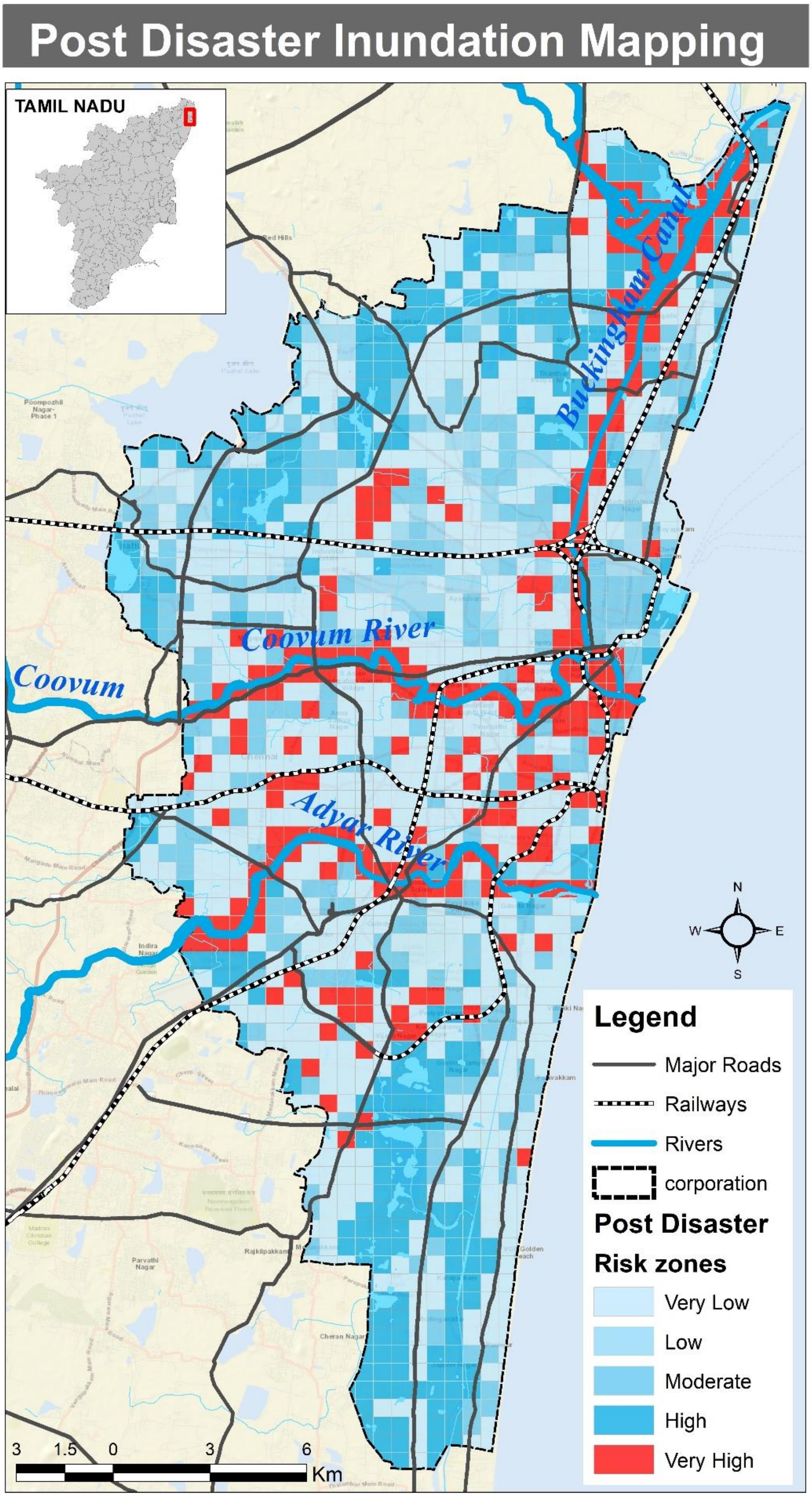
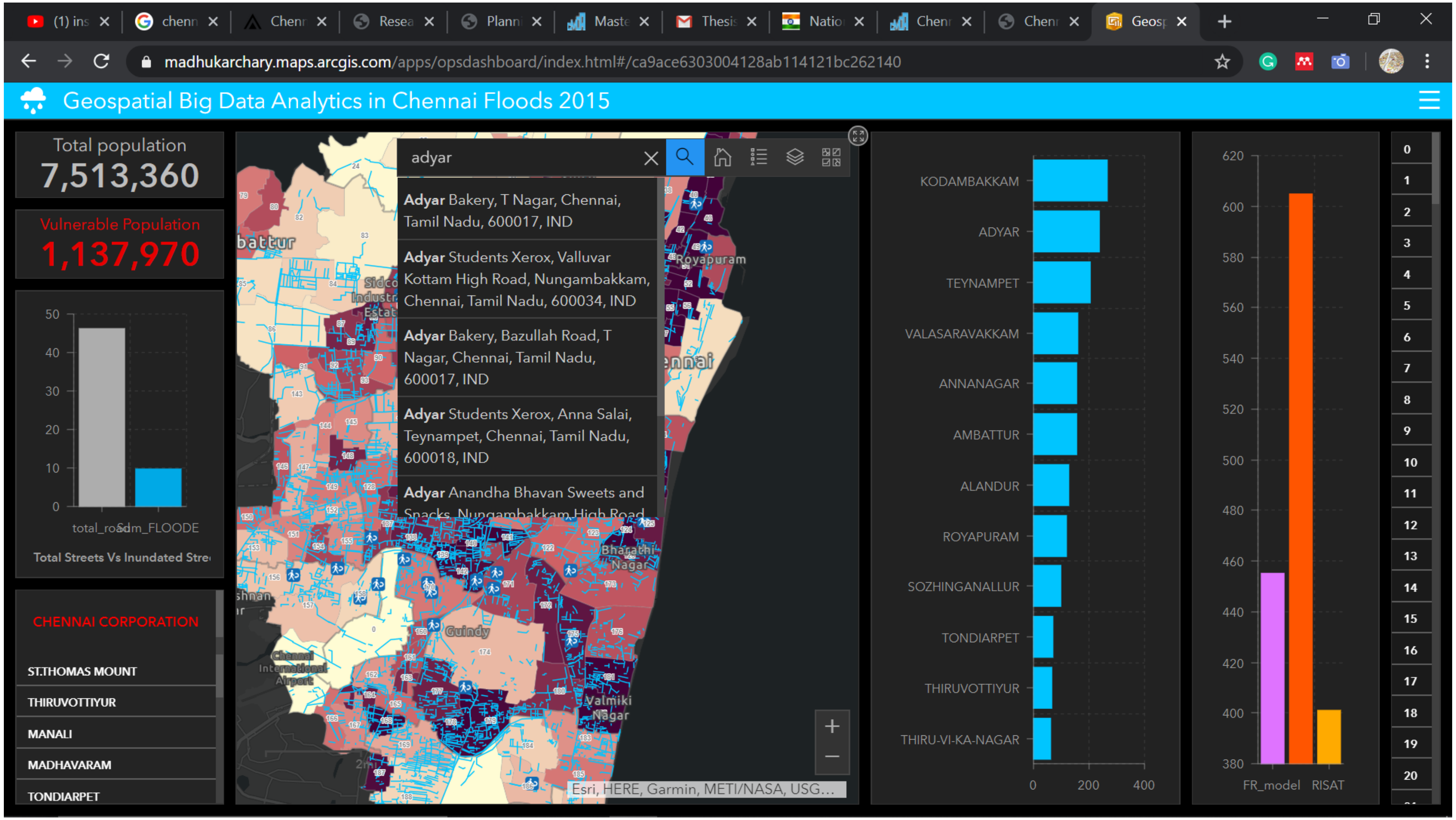


Table 16 Pixel based municipal official records Inundation map



Chapter 7 Geospatial dashboard

This Chapter emphasizes the data visualization in a web management server where ward level data is displayed with Analytics.


madhukarchary.maps.arcgis.com/apps/opsdashboard/index.html#/ca9ace6303004128ab114121bc262140
Geospatial Big Data Analytics in Chennai Floods 2015
Total population
7,513,360
Vulnerable Population
1,137,970
Total Streets Vs Inundated Stre
CHENNAI CORPORATION
ST.THOMAS MOUNT
THIRUVOTTIYUR
MANALI
MADHAVARAM
TONDIARPET
adyar
Adyar Bakery, T Nagar, Chennai, Tamil Nadu, 600017, IND
Adyar Students Xerox, Valluvar Kottam High Road, Nungambakkam, Chennai, Tamil Nadu, 600034, IND
Adyar Bakery, Bazullah Road, T Nagar, Chennai, Tamil Nadu, 600017, IND
Adyar Students Xerox, Anna Salai, Teynampet, Chennai, Tamil Nadu, 600018, IND
Adyar Anandha Bhavan Sweets and
Esri, HERE, Garmin, METI/NASA, USG...
KODAMBAKKAM
ADYAR
TEYNAMPET
VALASARAVAKKAM
ANNANAGAR
AMBATTUR
ALANDUR
ROYAPURAM
SOZHINGANALLUR
TONDIARPET
THIRUVOTTIYUR
THIRU-VI-KA-NAGAR
FR_model
RISAT

Chapter 8 Conclusion

This study analyzed the significance of crowdsourced near real time semi-structured locational data generated on social media platforms viz. twitter. The research also addressed advantages of open sources data dissemination platforms to collect and provide data to researchers and civilians to understand the situations in emergency. The correlation analysis between social media locational data sets from twitter and conventional data generated from satellite Imageries showed positive in the study and further it is validated with spatial data using statistical models. Consideration of text data can expose the emotions of public i.e. situational awareness at time of disasters but it needs proper validation and filtering before utilizing it in real time for decision making. Timing of posts need to be considered to validate in near real time. In this particular research twitter application programming interface was used to mine the data and additional platforms will definitely add accuracy to the output. The social media data sets generated from twitter are limited to create inundation map so contribution from other social media platforms will improve the accuracy of mapping. Through Unstructured data validation was not included in the overall study. Integration of information technology with environment planning very much needed in today's world because problems created by nature are dynamic but to monitor and manage those disasters data and methods are static. Real time assessment and situational awareness is needed very quick for emergency management. The further research includes the consideration of unstructured data validation and Handling crowdsourcing data sharing incentive models in near real time using advanced techniques of artificial intelligence, natural language processing techniques and machine learning.

Chapter 9 References

(2) What is Big Data_ (2019) - YouTube. (n.d.).

Akeuchi, Y. T. (2012). *Utilization of Social Media in the East Japan Earthquake and Tsunami and its Effectiveness Brett D . M . P EARY * Rajib S HAW **. *34*(1), 3–18.

Basin, C., Falling, A., The, W., & Area, B. (n.d.). *Chennai basin*.

Bhuvan _ Thematic Data dissemination _ Free GIS Data _ OGC Services _ Clip and Ship.pdf (p. 4326). (2014).

Big Data for Sustainable Development | United Nations. (n.d.). https://www.un.org/en/sections/issues-depth/big-data-sustainable-development/index.html

Chennai City Disaster Management Plan 2018. (n.d.).

Chennai, G., & Nadu, T. (2001). *CHAPTER 3*. 30–81.

Chennai Master Plan 2026 - Map, Summary & Free Download! (n.d.).

City, S., & City, T. (n.d.). *CHAPTER - IV PROFILE OF THE STUDY AREA - SRIPERUMBUDUR IN KANCHIPURAM DISTRICT Introduction*. 130–152.

CloudMoyo. (2019). *Sources of Big Data: Where Does It Come From? | CloudMoyo*. https://www.cloudmoyo.com/blog/data-architecture/what-is-big-data-and-where-it-comes-from/

CMDA. (2008). Second Master Plan For Chennai Metropolitan Area , 2026. *Strategies*, *I*(190), i–vii.

CMDA. (2013). *Chennai Metropolitan Development Authority, Government of Tamilnadu, India*. http://www.cmdachennai.gov.in

DRAINAGE SYSTEM IN CHENNAI METROPOLITAN. (n.d.).

Fohringer, J., Dransch, D., Kreibich, H., & Schröter, K. (2015). Social media as an information source for rapid flood inundation mapping. *Natural Hazards and Earth System Sciences Discussions*, *3*(7), 4231–4264. https://doi.org/10.5194/nhessd-3-4231-2015

GSAM. (2016). The Role of Big Data in Investing. *GSAM Perspectives*, *September*. https://doi.org/10.4018/978-1-4666-9840-6.ch069

Gupta, A. K., & Nair, S. S. (2010). Flood risk and context of land-uses: Chennai city case. *Journal of Geography and Regional Planning*, *3*(12), 365–372. http://www.academicjournals.org/JGRP

Juhász, L., Podolcsák, Á., & Doleschall, J. (2016). Open Source Web GIS Solutions in Disaster Management – with Special Emphasis on Inland Excess Water Modeling. *Journal of Environmental Geography*, *9*(1–2), 15–21. https://doi.org/10.1515/jengeo-2016-0003

Kanthimathinathan, T. (n.d.). *Session – 3 River and Drainage System in Cma*. 1–45.

Koswatte, S., McDougall, K., & Liu, X. (2015). SDI and crowdsourced spatial information management automation for disaster management. *Survey Review*, *47*(344), 307–315. https://doi.org/10.1179/1752270615Y.0000000008

Lee, C. S., Wu, C. C., Wang, T. C. C., & Elsberry, R. L. (2011). Advances in understanding the "perfect monsoon-influenced Typhoon": Summary from international conference on typhoon morakot (2009). *Asia-Pacific Journal of Atmospheric Sciences*, *47*(3), 213–222. https://doi.org/10.1007/s13143-011-0010-2

Lee, J. G., & Kang, M. (2015). Geospatial Big Data: Challenges and Opportunities. *Big Data Research*, *2*(2), 74–81. https://doi.org/10.1016/j.bdr.2015.01.003

Lokanathan, S., Perera-Gomez, T., & Zuhyle, S. (2017). *Mapping Big Data Solutions for Sustainable Development Goals*. *March*, 1–81.

Management, D. (2017). *ANNUAL*.

Marr, B. (2014). *How Facebook Is Using Big Data: The Good, The Bad and The Ugly | Bernard Marr | LinkedIn*. https://www.linkedin.com/pulse/20140716060957-64875646-facebook-and-big-data-no-big-brother

Meera Gandhi, G., Thummala, N., & Christy, A. (2015). Urban green cover assessment and site analysis in Chennai, Tamil nadu - a remote sensing and gis approach. *ARPN Journal of Engineering and Applied Sciences*, *10*(5), 2239–2243.

Minelli, M., Chambers, M., & Dhiraj, A. (2013). What Is Big Data and Why Is It Important? In *Big Data, Big Analytics* (pp. 1–18). https://doi.org/10.1002/9781118562260.ch1

Nadu, T. (n.d.). *दक्षक्षण न ि ू ी तटीय क्षेत्र , चेन्नई South Eastern Coastal Region, Chennai*.

Nair, M. R., Ramya, G. R., & Sivakumar, P. B. (2017). Usage and analysis of Twitter during 2015 Chennai flood towards disaster management. *Procedia Computer Science*, *115*, 350–358. https://doi.org/10.1016/j.procs.2017.09.089

NASSCOM. (2011). *Integrated ICT and Geospatial Technologies Integrated ICT and Geospatial Technologies*.

National Spatial Data Infrastructure _ Department Of Science & Technology. (n.d.).

Office of the Registrar General Ministry of Home Affairs Government of India. (2011). Census 2011: Provisional Population Totals at a Glance. Tamil Nadu. In *Census* (pp. 1–63). http://censusindia.gov.in/2011-prov-results/data_files/tamilnadu/2-FIGURES AT A GLANCE-tn.pdf%5Cnhttp://censusindia.gov.in/2011-prov-results/data_files/tamilnadu/3.Tamil Nadu_PPT_2011-BOOK FINAL.pdf

Panneerselvam, J., Liu, L., & Hill, R. (2015). An Introduction to Big Data. In *Application of Big Data for National Security* (pp. 3–13). https://doi.org/10.1016/b978-0-12-801967-2.00001-x

Pradesh, A. (n.d.). *Chapter 4 study area*. 4.

Reliance on Social Media in Today's Society - Digital Marketing Blog. (n.d.). https://www.insegment.com/blog/reliance-on-social-media-in-todays-society/

Report, E., Thomas, R., & Report, E. (2018). *Open Geospatial Consortium OGC Development of Disaster Spatial Data Infrastructures for Disaster Resilience*.

Rottach, P. (2004). *Background and Components of Disaster Risk Reduction*.

Sahana, M., & Patel, P. P. (2019). A comparison of frequency ratio and fuzzy logic models for flood susceptibility assessment of the lower Kosi River Basin in India. *Environmental Earth Sciences*, *78*(10), 1–27. https://doi.org/10.1007/s12665-019-8285-1

Sarda, N. L., Acharya, P. S., & Sen, S. (2018). Geospatial infrastructure, applications and technologies: India case studies. In *Geospatial Infrastructure, Applications and Technologies: India Case Studies*. Springer Singapore. https://doi.org/10.1007/978-981-13-2330-0

Scott, G., & Rajabifard, A. (2017). Sustainable development and geospatial information: a strategic framework for integrating a global policy agenda into national geospatial capabilities. *Geo-Spatial Information Science*, *20*(2), 59–76. https://doi.org/10.1080/10095020.2017.1325594

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (Muki). (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, *31*(1), 139–167. https://doi.org/10.1080/13658816.2016.1189556

Series, T. R. (2008). For official use Technical Report Series DISTRICT GROUNDWATER BROCHURE NAGAPATTINAM DISTRICT By Government of India Ministry of Water Resources Central Ground Water Board South Eastern Coastal Region. *Ground Water*, *November*.

Shi, Y., Sayama, T., Takara, K., & Ohtake, K. (2019). Detecting flood inundation information through Twitter: The 2015 Kinu River flood disaster in Japan. *Journal of Natural Disaster Science*, *40*(1), 1–13. https://doi.org/10.2328/jnds.40.1

Snoeren, G., Zlatanova, S., Crompvoets, J., & Scholten, H. (2007). Spatial Data Infrastructure for emergency management: the view of the users. *Proceedings of the 3rd International Symposium on Gi4DM*, *June 2014*, 22–25. http://www.researchgate.net/publication/40106306_Spatial_Data_Infrastructure_for_emergency_management_the_view_of_the_users/file/9c96051b77f02f2c06.pdf

Tseng, S. F., Chen, W. C., & Chi, C. L. (2011). Online social media in a disaster event: Network and public participation. *Communications in Computer and Information Science*, *167 CCIS*(PART 2), 256–264. https://doi.org/10.1007/978-3-642-22027-2_22

UN. (2015). *Goal 15*. Sustainable Development Goals. http://www.un.org/sustainabledevelopment/biodiversity/

WHO. (2012). *WHO | Natural events*. Who. https://www.who.int/environmental_health_emergencies/natural_events/en/

ANNEXURE I

CALCULATION OF LULC AND GEOLOGY OCCURRENCE RATIO

FACTOR	CLASSES	occurrence	occurrence%	class area	class area %	ratio	RF	MAX	MIN	MIN TOTAL	PR
LULC	Agriculture Crop Land	1861	22.43	307.18	22.08	1.02	0.10	0.1	0.02	0.71	0.21
	Agriculture Fallow	2026	24.42	223.87	16.09	1.52	0.15				
	Agriculture Plantation	83	1.00	16.59	1.19	0.84	0.08				
	Barren / sandy Area	1732	20.88	185.93	13.36	1.56	0.15				
	Barren Uncultivable	139	1.68	16.00	1.15	1.46	0.14				
	Grass/ Grassing	26	0.31	21.47	1.54	0.20	0.02				
	Rural Built Up	564	6.80	55.91	4.02	1.69	0.17				
	Urban Built Up	700	8.44	441.86	31.75	0.27	0.03				
	Water Bodies	1165	14.04	122.71	8.82	1.59	0.16				
		8296	100.00	1391.52		10.15					

FACTOR	CLASSES	occurrence	Occurrence%	class area	class area %	ratio	RF	MAX	MIN	MAX TOTAL	PR
GEOLOGY	Amphibolite	87	1.84	8.50	0.74	2.48	0.22	0.22	0.06	0.71	0.23
	Argillaceous & clacareous sandsin	24	0.51	6.73	0.59	0.86	0.08				
	Clay(Reddish Brown)	419	8.85	88.00	7.68	1.15	0.10				
	Fluvial Marine	149	3.15	29.68	2.59	1.21	0.11				
	Garnet-biotite-sillimanite	43	0.91	15.10	1.32	0.69	0.06				
	Granite	1239	26.16	274.81	24.00	1.09	0.10				
	Gray Black Sandy Clay	1253	26.46	370.10	32.32	0.82	0.07				
	Sand and Silt	1083	22.87	223.84	19.55	1.17	0.11				
	Sandstone	418	8.83	121.78	10.63	0.83	0.08				
	Syenite/Nephelene Syenite	21	0.44	6.71	0.59	0.76	0.07				
		4736		1145.25		11.06					

ANNEXURE II

CALCULATION OF SOIL, DRAINAGE DENSITY AND SLOPE OCCURRENCE RATIO

FACTOR	CLASSES	occurrence	Occurrence%	class area	class area %	ratio	RF	MAX RF	MIN RF		TOTAL	PR
SOIL	CLAY	62	1.27	38.18	3.35	0.38	0.08	0.23	0.08	0.15	0.71	0.21
	SHALE	79	1.62	41.90	3.67	0.44	0.09					
	HARD ROCK	1329	27.31	288.24	25.25	1.08	0.23					
	SAND STONE	1043	21.43	233.68	20.47	1.05	0.22					
	SAND	2156	44.30	480.28	42.08	1.05	0.22					
	SHALE	198	4.07	59.04	5.17	0.79	0.16					
		4867		1141.32		4.79						

FACTOR	CLASSES	occurrence	% of occurrence	class area	class area %	ratio	RF	MAX RF	MIN RF		MAX TOTAL	PR
Drainage Density	1.00	188	3.54	440.40	37.12	0.10	0.01	0.71	0.01	0.71	0.71	1.00
	2.00	704	13.24	338.82	28.56	0.46	0.03					
	3.00	1216	22.87	233.99	19.72	1.16	0.08					
	4.00	1592	29.94	139.38	11.75	2.55	0.17					
	5.00	1617	30.41	33.76	2.85	10.69	0.71					
		5317		1186.35		14.95						

FACTOR	CLASSES	occurrence	% of occurrence	class area	class area %	ratio	RF	MAX RF	MIN RF		MAX TOTAL	PR
Slope	1.00	3592.00	35.28	485.85	40.95	0.86	0.15	0.26	0.15	0.11	0.71	0.16
	2.00	3611.00	35.47	475.39	40.07	0.89	0.15					
	3.00	2169.00	21.31	228.28	19.24	1.11	0.19					
	4.00	722.00	7.09	55.39	4.67	1.52	0.26					
	5.00	86.00	0.84	6.64	0.56	1.51	0.26					

		10180.00				5.88						

ANNEXURE III

CALCULATION OF GEOMORPHOLOGY OCCURRENCE RATIO

FACTOR	CLASSES	occurrence	% of occurrence	class area	class area %	ratio	RF	MAX RF	MIN RF		MAX TOTAL	PR
GEOMORPHOLOGY	Anthropogenic Origin-Anthropogenic Terrain	199	3.98	45.63	3.84	1.04	0.11	0.23	0.09	0.14	0.71	0.20
	Coastal Origin-Older Coastal Plain	323	6.46	38.73	3.26	1.98	0.20					
	Coastal Origin-Older Deltaic Plain	2127	42.56	524.28	44.15	0.96	0.10					
	Coastal Origin-Younger Coastal Plain	356	7.12	88.62	7.46	0.95	0.10					
	Denudational Origin-Pediment-Pedi Plain Complex	1600	32.01	440.22	37.07	0.86	0.09					
	Fluvial Origin-Active Flood Plain	285	5.70	38.43	3.24	1.76	0.18					
	Waterbodies	108	2.16	11.51	0.97	2.23	0.23					
		4998		1187.42		9.79						

ANNEXURE IV

URL's

https://madhukarchary.maps.arcgis.com/home/content.html?view=table&sortOrder=desc&sortField=modified&folder=madhukarchary#content

https://developer.twitter.com/en/dashboard

http://localhost:8888/?token=c85bdcc138d1a63f21934dfe0127fdfeac73fa902b31a55c

or http://127.0.0.1:8888/?token=c85bdcc138d1a63f21934dfe0127fdfeac73fa902b31a55c

API	Temporary access token = JpA8539YnsQYFReX-rNZbNu89ksZMoIL5iJ8u_KrCqw1viNdqujVdoMUjVEMJivjq1Tr_OvVS_HNo2MTeC8beR9CiF6FncBukVpg1E7d2UnIyuplwH7g0843dlwSOjjwYMWLZkUcdV2XzQI4cxNvM10OArcc01lqX1N0KnaOxC7ZmlzZAzKIICQWG7OaIbLijE2aOIL9UfZfDILKXGjWIXIPHw7wxqdjlTOCoTo90S8.
	Runtime Lite license key = runtimelite,1000, rud5166557084,none,ZZ0RJAY3FL9D2B3TR161
	https://madhukarchary.maps.arcgis.com/home/webmap/viewer.html