

CHAPTER 4

MATERIAL AND METHODS

4.1 Water sampling and its physico-chemical analysis

Surface water from in and around the reservoir were sampled using hand held clean and sterilized polypropylene wide mouth reagent bottles of 125ml in triplets. To sample water from the sources, non-isokinetic samplers, commonly employed open mouth swing samplers were used. The bottle was held at a depth of around 0.5m, with the mouth facing slightly downwards. Afterwards, the bottle was turned upright, acidified immediately for pH less than 2 with nitric acid, replaced with the bottle cap and stored with necessary precautions at a temperature less than 4⁰C in the laboratory to abstain it from any kind of alterations. A total of 120 water samples were collected both during pre & post monsoon reasons i.e. during the month of February and October respectively after the preliminary field survey to interpret the overall impact of industrial and mining activity on the adjoining reservoir. Right away after the sampling of water pH, electrical conductivity, total dissolved solids, dissolved oxygen and temperature were analyzed in the field using well calibrated Hanna Hi 9828 multi parameter apparatus. Thereafter the sample bottles were labelled and brought to the departmental laboratory for analysis of other anions, cations and total suspended solids. For the analysis of ions such as potassium, sodium, calcium, lithium, magnesium, chloride, fluoride, sulphate, nitrate, phosphate - ion chromatography was used while for remaining heavy metals such as iron, copper, zinc, lead, chromium, cadmium, nickel - induced coupled plasma technique was deployed. However, total suspended solids of the water samples was computed using the conventional laboratory analysis method. In this method, TSS was evaluated by filtering 100 ml of water sample through Whatmann filter paper no. 41 using formula whereas total hardness was estimated by complexometric titration method by titrating with a standard solution of ethylene diamine tetra acetic acid (EDTA) – a complexing agent used

in the determination of total hardness present in the water samples. These standard methods for determination of physico-chemical parameters were used in consultation, with the guidelines provided by “standard procedures for treatment of water and waste water” as laid down by (American Public Health Association (APHA) in its 22nd edition while permissible limit was taken in to consideration as per the guideline given by Bureau of Indian Standards (BIS) and World Health Organization (WHO). Figure 4.1 shows sampling location map of study area computed after obtaining satellite data from USGS earth explorer website.

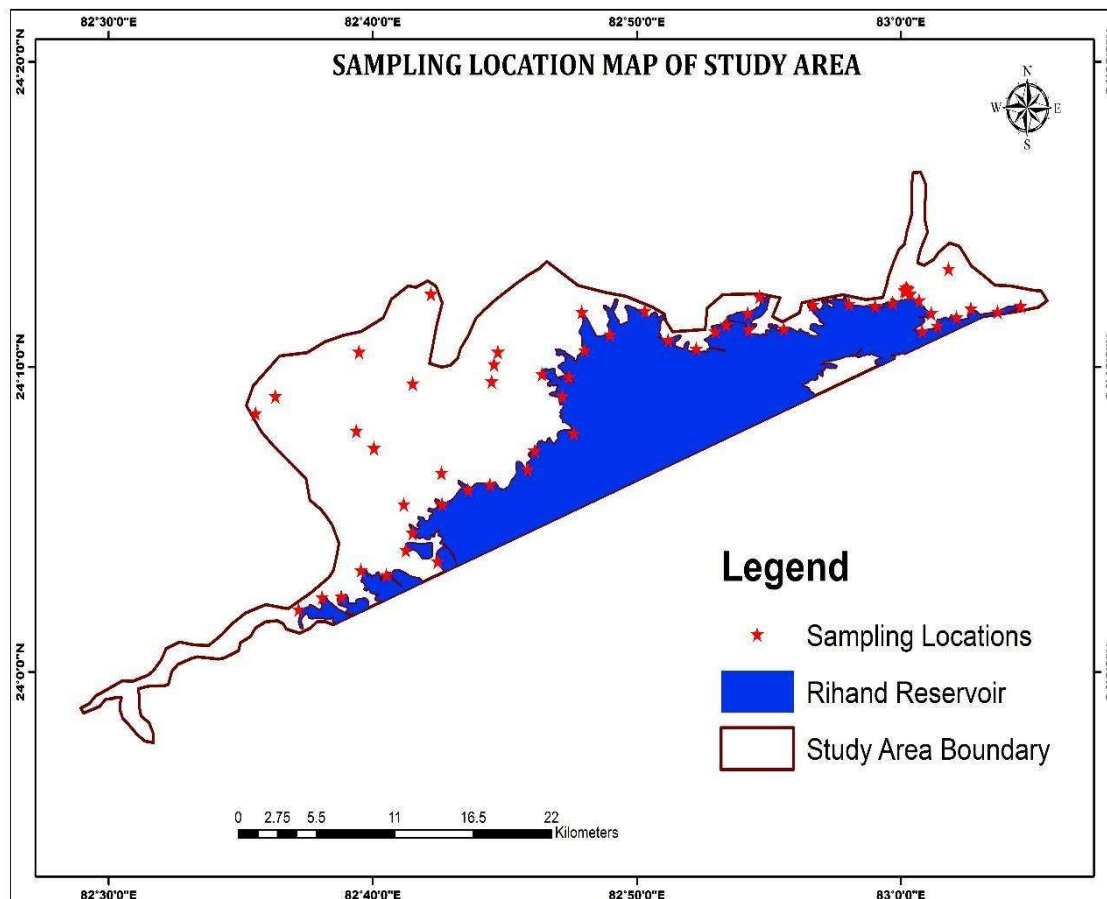


Fig 4.1 Sampling location map of study area



Fig 4.2 Laboratory analysis of water samples through multi-parameters analyzer



Fig 4.3 Laboratory analysis of water samples through Ion Chromatography

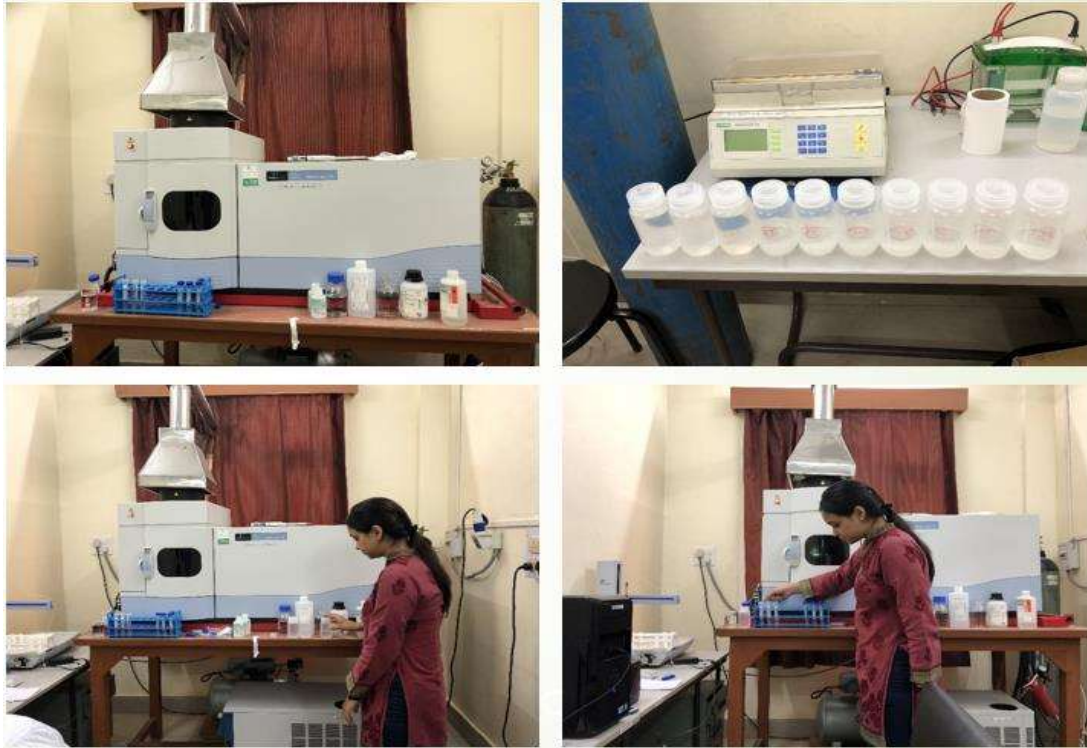


Fig 4.4 Analysis of water samples through Induced Coupled Plasma



Fig 4.5 Laboratory analysis of water samples for Total Suspended Solids

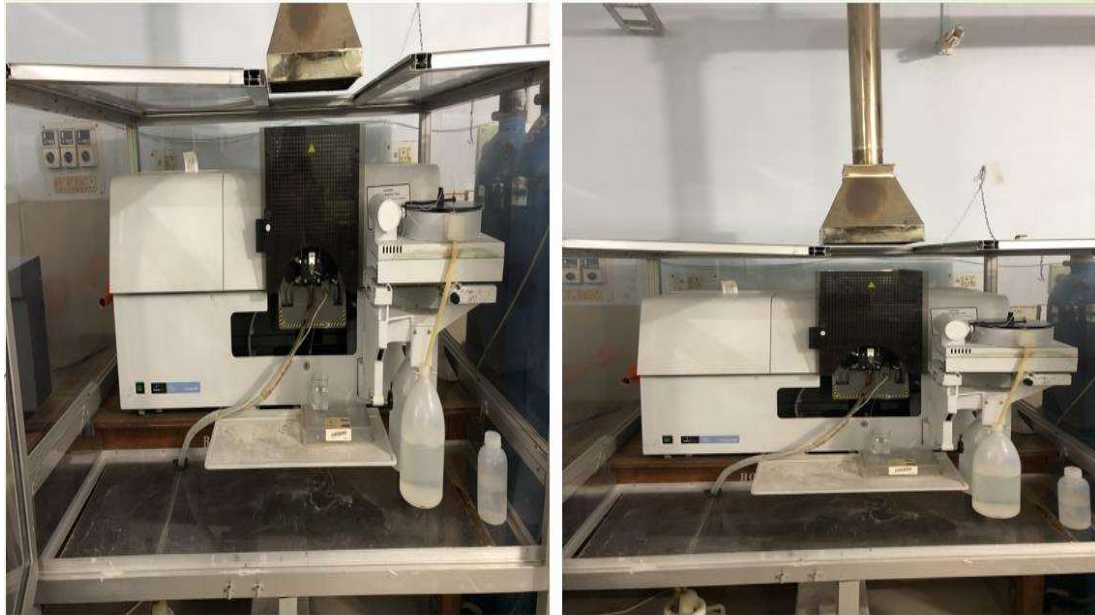


Fig 4.6 Analysis of water samples through Atomic Absorption Spectroscopy



Fig 4.7 Water sample locations in coal mines of NCL



Fig 4.8 Water sample locations in Govind Ballabh Pant reservoir

4.2 Method for calculation of Water Quality Index

The surface water samples (i.e. reservoir water samples and industrial water samples) quality was calculated by using Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) formula developed by the British Columbia Ministry of Environment, Lands and Parks and modified by Alberta Environment. In this method, the final Index value ranges between 0-100 in which the higher value indicate an increase in quality of water while the lower value of Water Quality Index (WQI) represents depletion in quality of water according to the CCME WQI ranked :-

Excellent: (CCME WQI Value 95-100) – water quality is protected with a virtual absence of threat or impairment; conditions very close to natural or pristine levels.

Good: (CCME WQI Value 80-94) – water quality is protected with only a minor degree of threat or impairment; conditions rarely depart from natural or desirable levels.

Fair: (CCME WQI Value 65-79) – water quality is usually protected but occasionally threatened or impaired; conditions sometimes depart from natural or desirable levels.

Marginal: (CCME WQI Value 45-64) – water quality is frequently threatened or impaired; conditions often depart from natural or desirable levels.

Poor: (CCME WQI Value 0-44) – water quality is almost always threatened or impaired; conditions usually depart from natural or desirable levels.

After the body of water, the period of time, and the variables and objectives have been defined, each of the three factors that make up the index must be calculated. The calculation of F_1 and F_2 is relatively straightforward; F_3 requires some additional steps.

F_1 (Scope) represents the percentage of variables that do not meet their objectives at least once during the time period under consideration (“failed variables”), relative to the total number of variables measured:

$$F_1 = \frac{(\text{Number of Failed Variables})}{(\text{Total number of Variables})} \times 100 \quad (1)$$

F_2 (Frequency) represents the percentage of individual tests that do not meet objectives (“failed tests”):

$$F_2 = \frac{(\text{Number of Failed tests})}{(\text{Total number of tests})} \times 100 \quad (2)$$

F_3 (Amplitude) represents the amount by which failed test values do not meet their objectives. F_3 is calculated in three steps.

(i) The number of times by which an individual concentration is greater than (or less than, when the objective is a minimum) the objective is termed an “excursion” and is expressed as follows. When the test value must not exceed the objective:

$$\text{excursion}_i = \frac{(\text{Failed Test Value}_i)}{(\text{Objective}_j)} - 1 \quad (3a)$$

For the cases in which the test value must not fall below the objective:

$$\text{excursion}_i = \frac{(\text{Objective}_j)}{(\text{Failed Test Value}_i)} - 1 \quad (3b)$$

ii) The collective amount by which individual tests are out of compliance is calculated by summing the excursions of individual tests from their objectives and dividing by the total number of tests (both those meeting objectives and those not meeting objectives). This variable, referred to as the normalized sum of excursions, or nse, is calculated as:

$$nse = \frac{\sum excursion}{\# of tests} \quad (4)$$

iii) F_3 is then calculated by an asymptotic function that scales the normalized sum of the excursions from objectives (nse) to yield a range between 0 and 100.

$$F_3 = \frac{nse}{(0.01nse+0.01)} \quad (5)$$

Once the factors have been obtained, the index itself can be calculated by summing the three factors as if they were vectors. The sum of the squares of each factor is therefore equal to the square of the index. This approach treats the index as a three-dimensional space defined by each factor along one axis. With this model, the index changes in direct proportion to changes in all three factors.

The CCME Water Quality Index (CCME WQI):

$$CCMEWQI = 100 - \left(\frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.732} \right) \quad (6)$$

The divisor 1.732 normalises the resultant values to a range between 0 and 100, where 0 represents the “worst” water quality and 100 represents the “best” water quality.

4.3 Mapping of Water Quality/Spatial Distribution Maps

Since, the sampling of reservoir & industrial water was done between Feb 2016 to Dec. 2018, thus covering 3 pre monsoon seasons and 3 post monsoon seasons. The average or mean of each of these seasons has been into consideration for important parameters of the water quality along with water quality index. For mapping of these water quality parameters Arc GIS 10.1 software was used as Geographic Information system can aid in

reinforcing an effective solution to an extensive range of problems associated with water quality assessment. The interpolation of spatial distribution map utilizes mainly two methods i.e. Inverse Distance Weighting (IDW) & Kriging method. The IDW makes predictions using a linear weighted combination based on the inverse of the distance between the points (Shahid et al 2016). The IDW is preferred over Kriging because it assumes that the nearer a sample point is to a cell whose value is to be estimated, the more closely the cell's value will resemble the sample points value. This makes it computationally fast & also makes it to accommodate barriers which shows linear discounting in the surface (Shahid et al 2017). However, in Kriging method, assumes that data is acquired from a stationary stochastic process & the technique requires that this is to be normally distributed (Audu et al 2015). This method is not preferred since the pixel value used per computation of spatial distribution map does not pass through the studied points & the pixel value can go beyond the studied points of the samples. The major limitation of this method is the non-linearity & non stationarity in the data.

Mathematically, IDW is expressed by the equation:-

$$\hat{Z}(S_0) = \sum_{i=1}^n \lambda_i(S_0) \cdot Z(S_i) \quad (7)$$

Where, $\hat{Z}(S_0)$ is the values at unknown locations, $\lambda_i(S_0)$ is the weighting value and $Z(S_i)$ is the values at known locations.

$$\lambda_i(S_0) = \frac{\frac{1}{\beta d(S_0, S_i)}}{\sum_{i=1}^n \frac{1}{\beta d(S_0, S_i)}} ; \beta > 1 \quad (8)$$

Where λ_i is the weight for neighbor i (the sum of weights must be unity to ensure an unbiased interpolator), $d(S_0, S_i)$ is the distance from the new point to a new point to a known sample point, β is the coefficient used to adjust the weights and n is the total number of points in the neighborhood analysis. In this study, IDW was performed using

60 sampling locations for geostatistical analysis in ArcGIS 10.1 software during pre and post monsoon seasons. Here, the seasonal classification was done for important parameters of water quality along with WQI where seasonal mean value during pre and post monsoons were computed separately and used as weights in mapping the study area in terms of parameters and WQI through IDW for the seasons separately. The WQI calculated through CCME method encompassing various physicochemical parameters was performed and a resulting WQI map along with maps for other physicochemical parameters were drawn explaining the outcome of the study for both the seasons during the investigation.

4.4 Basic statistical Analysis

4.4.1 Correlation Analysis

As there were not much variations among the quality of the water samples from the reservoirs though few of the sampling stations of industrial water showed absurd results. To analyze the relationship between various physico-chemical parameters of the water samples taken into account, correlation coefficient (r) was determined among the samples for both the seasons i.e. pre and post monsoon by using the mathematical formula & Microsoft Excel 2010 software.

$$r = \frac{n\sum(xy) - \sum(x)(y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (9)$$

Where x & y are the variables chosen for the analysis & n is the number of data. The correlation coefficient is a measure of linear relationship between the variables. In this, the value ranges between - 1 to +1, where +1, represents a high positive correlation between the variables whereas -1 shows a high inverse relationship between the variables however 0 represents no relationship amongst the variables.

4.4.2 Box Plot Diagrams

The box plot is a diagrammatic representation of results of various physico-chemical parameters taken into consideration under this study. Each box in the diagram is shown by the data in terms of maximum value, minimum value, 1st & 3rd quartile along with median value for the given dataset. This statistical method is useful in depicting the comparative analysis amongst the value of concentration for various physico- chemical parameters. The box plot diagram for the current study has been plotted by using origin 8.1 software for both pre & post monsoon seasons.

4.4.3 Hierarchical Cluster Analysis (HCA)

HCA can be defined as the technique of unsupervised pattern recognition where classes are divided into clusters or groups on the basis of dissimilarities amongst the different classes while classifying them on the basis of similarities within the class. Dendrograms provides a visual summary along with important information of the clustering showing the propinquity amongst the groups (Shrestha 2007). In this study, HCA was determined on the basis of Ward's method using Square Euclidean distance in which the distance between two stations are calculated on the basis of square root of the sum of the squares of the differences between the corresponding values.

$$D_{ij} = \sum_{k=1}^p [(Z_{i,k} - Z_{j,k})^2]^{0.5} \quad (10)$$

Where p = no. of variables, $Z_{i,k}$ & $Z_{j,k}$ are standardized value for variables k at site i & j respectively. The results of HCA are useful in interpreting the data along with indicating the pattern followed by the data used in the study. Cluster analysis not only helps in identifying the similarities or dissimilarities amongst the samples but it also helps in confirming the accuracy rate of the discriminant function along with recognizing the most remarkable physico chemical variable accountable for spatial & temporal variations

in the water quality of the reservoir as well as the water from industrial outlet. This technique is useful in spotting the site corresponding to most polluted areas & setting up management techniques in those areas (Hajigholizadeh et al 2016). The data used in this study for the computation of Dendrogram has been treated through SPSS 22.0 version software.

4.4.4 Piper Diagram

This diagram was developed by piper in (1944) to describe the chemical characteristics of water (Back, 1961). The piper plot classification is based on six chemical components out of which three are cations (Ca, Mg & Na+K) and three are anions (Cl^- , SO_4^{2-} & $\text{CO}_3^{2-} + \text{HCO}_3^-$). These characterization of species allows for easy visual chemical comparison of the chemistry of water samples along with recognition of water samples with interchangeable chemical attributes, assessment of mixing among waters having different characteristics & possible determination of the chemical evolution in water samples (USGS report).

The diagrammatic presentation of the piper plot is represented by a sum of cations & anions accounting to 100% & on the each ternary diagram values of cations & anions are plotted accompanying the central diamond shaped diagram. It helps in identifying the water samples on the basis of dominant presence of cation or anion. In this study, GW_Chart software has been used for computation of piper plots based on the presence of cations and anions in the water samples collected from reservoirs in the water samples collected from reservoirs or industrial outlets.

4.5 Land Use/ Land Cover (LU/LC)

For LU/LC analysis of the study area, remotely sensed satellite data of Landsat 5 MSS and Landsat 8 OLI were recorded. The Landsat 5 satellite consists of MSS (Multispectral sensor) & Thematic Mapper (TM) sensor. The MSS sensor consist of four spectral bands (Bands 4,5,6 & 7) at 60 m resolution with six detectors for each spectral band in addition to ground sampling interval (pixel size) of 57 X 79 m while the TM sensor consists of 6 spectral bands of 30 m resolution and a thermal band (Band 1,2,3,4,5 & 7) with a ground sampling interval of 30m reflective & 120 m thermal. The satellite is a run synchronous near polar orbit satellite at an altitude of 705 km with a 16 days repeat cycle circling the earth at every 99 minutes at a resolution of 30 m and 120 m respectively with a swath width of 185km. While the Landsat 8 satellite is the most recent satellite of the Landsat series consisting of OLI and TIRS sensors. The OLI seasons consists of 9 bands for spectral responses while the TIRS sensors collect the information from the earth's surface at 2 discrete wavelengths within the thermal infrared position of the electromagnetic spectrum. The resolutions of visible & thermal infrared regions are 30m & 100m respectively while that of the panchromatic region is 15m. The swath of this satellite is about 185km unlike that of the Landsat 5 satellite for a resolution of 30m with a repetitive cycle of 16 days. The study utilizes the band inputs of 3, 4 & 5 bands for LU/LC classification in path-142 & row-43 for the evaluation. For the year 1988, 1996 & 2005 the multi temporal satellite imagery were acquired from Landsat 5 MSS for the month of Feb while that for the year 2016 & 2018 satellite imagery from Landsat 8 OLI were taken in the consideration again for the month of February from USGS (United States Geological survey) earth explorer website (www.earthexplorer.usgs.gov). The month of February was chosen as during this period of the year the atmosphere is usually free from the rain bearing clouds there by making it clear. These acquisition of imagery were done

during the same month of the respective year due to the reason to avoid any kind of seasonal variations for the time span of 30 years (1988-2018) to review the spatial data of land use in different years. For this study, Landsat Images were taken into consideration due to their easy availability with medium and high spatial resolution. The Area of interest (AOI), or region of interest (ROI) was delineated from the acquired satellite imagery using the function of subset. In this study for the processing, enhancement and classification of (Region of Interest) ROI or (Area of Interest) AOI ERDAS Imagine 2014 and ArcGIS 10.1 software has been employed for the production of land use land cover maps of 5 years for a time span of 30 years. However, the Toposheets No. 631/11 63L/12, 63L/15, 63L/16, 64I/09, 64I/13, 64P/03 and 64P/04 at the scale 1:50000 were used the spatial and temporal changes in the study area during the period of study. These Toposheets were later used for geo rectification. It provides a supplementary source for mapping LU/LC with primary data. Here, in this study the satellite data has been classified into six LU/LC classes i.e. Habitation, Hilly area/forest, mining Area, water bodies, Fallow /Barren & crop land. For the assessment of LU/LC of the area under study, following steps were followed.

4.5.1 Preprocessing

For classifying a raw data the preliminary step is the correction of data which are mainly used to improve the radiometric geometric & atmospheric error present in the data. In radiometric correction the data is corrected from the season irregularities which is thereby used to accurately represent the reflected or emitted radiation measured by the sensor. However, in geometric correction the image is transformed in such a way that it has the scale & projection properties of a map by assigning ground control points to the image, for registering the map to a precise map of geo referencing the satellite image while in atmospheric correction the scattering caused by atmospheric gases and particles is

corrected by darkest pixel (DP) method according to which the visible spectrum is highly affected while near infrared is free from any kind of atmospheric scattering. Since, water has low reflectance values in the visible bands & negligible in the near infrared band as per the DP method the pixel having lowest DN value is assumed to be zero in each band which may correspond to the darkest object in that band.

4.5.2 Image Enhancement

In order to obtain a hazel free image, this step of image processing is adopted where the using appearance of the objects in the image is improved by linear grey level stretching & spatial filling for enhancing the edges in which a level threshold value is chosen so that pixel values below this are mapped to zero while the pixel values above this are mapped to 255. Rest all the values of pixels between 0-255 are interpolated linearly so that the haziness of the image could be minimized and a contrast between the different objects present in the image could be refined.

4.5.3 Image Classification

The main objective of this method is to automatically classify all pixels in the image to LU/LC classes (Shalaly & Tateshi 2007). For classifying this a specifically designed algorithm is used which mainly relies on the purpose & the kind of the work. Typically, pixel based methods are used for classification in Remote sensing by classifying different pixel values & grouping them into class based on spectral properties. The main techniques used in this are super wised and unsupervised classification. In supervised classification, a prior knowledge about the classes are known acne to which the software classify them into different classes whereas in unsupervised classification, no prior knowledge about the classes is there and the computer group all the pixels according to the spectral relationship between them, thereafter Each group will be assigned a land cover type by

the analyst. In this study, supervised classification has been used for classifying the image into 6 different classes based on the prior knowledge about the study area. These classes are as follows:- Habitation, Hilly Area/Forest, Mining Area, Water bodies, Fallow/Barren and cropland for the years 1988, 1996, 2005, 2016 & 2018. After this, the major area changes for all the classes in different years were observed according to the land use land cover classification

4.6 Post Classification Change Detection

Post classification change detection map (PCC detection map) was prepared for the time period of 1988-2018 i.e. a time span of 30 yrs. This detection technique is concerned in recognizing the rate, nature & location of changes which are primarily based on the independent classification as well as conditions and properties of satellite images during the acquisition (Almutairi & Warner 2010). On the contrary, PCC detection technique represents the quantity & dispersal of each of the changed land cover classes equally within the different multi temporal images (Ojaghi et al 2017). However, this method neglects the temporal correlation between the corresponding pixels in two different time periods (Wu et al 2017). The PCC techniques used in this study firstly considers the LU/LC images classification for 5 different time periods after which the procedure containing overlaying of images using ArcGIS software was embraced for dismissing the spatial changes in LU/LC during 1988-2018. This technique was later on used to describe changes in the study area on the basis of “from-to” change in the various classes of the 6 classes of 5 maps for 5 different years.

4.7 Volume Estimation of water present in the reservoir

To estimate the capacity of the Govind Ballabh Pant Sagar reservoir, Satellite Imagery of Landsat Series was obtained from the United States Geological Survey (USGS) website

for the year 1988, 2013, 2016 and 2018. These imageries were then preprocessed, enhanced and utilized to construct a Digital Elevation Model (DEM) from data surveyed along with the generation of Triangulated Irregular Network (TIN) surface from the data surveyed. DEM data was deployed to produce the elevation of the study area while TIN was implicated to establish a TIN surface for surveyed data set using direct linear interpolation. These data sets are primarily used in the estimation of the volume of the reservoir since they provide a proficient reference source for the interpolation of the surface from the surveyed data set. The annual rainfall data of the study area from 1998 to 2018 was also acquired from the CMPDI NCL. According to the rainfall data having values approximately the same for all the four years were used in the estimation of the volume of the reservoir. After this, the volume of the reservoir was calculated by subtracting the elevation values computed within the bounded surface of the reservoir from the full supply level. However, under the TIN model, it was calculated by summation of the volumes of the components of the triangulated prisms. The DEM and TIN were derived from the ArcGIS 10.1 software. Then, the data obtained from the TIN is converted into the raster format of the image in order to create random points for the extraction of its values from these points of the raster data for the generation of elevation map and estimation of reservoir volume. The capacity of the reservoir for all the four years during a time span of 20 years was calculated separately. The difference obtained between the values of these years were then used for the calculation of sediment deposition and decrease in volume of the reservoir for a time span of 20 years using radial focus in ArcGIS 10.1.

Hence, this chapter described about the detailed methodology adopted in the research work done. The next Chapter presents the results obtained as well as discusses it in a detail manner.