Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2471-2477
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.338

Research Article

A Spectral Structural Approach for Building Extraction from Satellite Imageries

Minakshi Kumar[1], Pradeep Kumar Garg[2], Sushil Kumar Srivastav[1]

[1]Indian Institute of Remote Sensing, Dehradun, Uttarakhand, India
[2]Uttarakhand Technical University, Dehradun, Uttarakhand, India

Publication Date: 22 February 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.338

Abstract Automatic feature extraction from high resolution satellite imagery remains an open research area in the field of remote sensing, computer vision and machine learning. While many algorithms have been proposed for automatic building extraction, none of them solve the problem completely. This paper proposes a system for increasing the degree of automation in extraction of building features from high resolution multispectral satellite images. Image segmentation is a prerequisite for processing of very high spatial resolution imageries. Most image segmentation methods use spectral information of an image alone for generating image objects. A novel image segmentation method for very high spatial resolution multispectral images using combined spectral and structural information is proposed in this paper. The method involves computation of textural parameters from high resolution multispectral imagery and is combined with the spectral bands for extracting spectral-structural characteristics. Hence in addition to the spectral information, the tone, texture and shape information is evaluated for an object-oriented analysis. The support vector machines classification rules are applied on the generated object primitives. The proposed image segmentation method is well applicable to the segmentation of imagery over urban and suburban areas for large scale building extraction.
Keywords *Image segmentation; Multi-resolution segmentation; Support vector machines; Textural analysis*

1. Introduction

Image segmentation involves the partitioning of a given image into a number of homogeneous regions according to a given criterion. Image segmentation is a key and prerequisite step for processing of very high-resolution imagery. The quality of image segmentation directly affects the quality of subsequent object-based image analysis and applications. While the traditional pixel-based approach for remote sensing image classification is based on the statistical analysis of multispectral features of the pixels in an image, object-based image analysis (OBIA) allows the use of a wide range of additional information (Aguilar et al., 2012, Blaschke et al., 2014). The OBIA approach involves two steps: segmentation and classification. After segmentation, a very large number of features can be calculated for the resulting objects. The main advantages of OBIA, compared with pixel-based approaches, is the larger number of available features and the fact that the features convey more information when they are calculated on real objects than when sampled on a square grid (Baltsavias 2004; Barrile et al., 2008; Su et al., 2008) having a very large number of features poses two problems. First, the larger the number of features used in classification, the longer the computing time needed.

Second, using a very large number of explanatory features, especially when some of them are redundant, noisy or information less, might result in a less accurate classification. This is the so-called curse of dimensionality or the Hughes effect, which is an important issue in optimization and machine learning. Previous work done using the above methods mainly compared the accuracy improvement on pixel based classification by object based classification (Kumar M, 2013; Kumar M et al., 2009a). Because of the diverse composition of features in an urban environment, extracting urban features from high resolution imagery is a difficult task. In urban area, previous work focused on extracting on extracting either buildings (Kumar M, 2009 a, b; Ettarid, 2008; Xiaoying, 2005) or roads (Kumar M, 2014) where high resolution image was segmented and classified to obtain requisite buildings.

Most existing image segmentation methods use spectral information of an image alone and the exploration of new and sophisticated image segmentation methods has been a focus of image processing and analysis. A novel image segmentation method for very high spatial resolution multispectral images using combined spectral and structural information is proposed in this paper. The main objective of this study is to evaluate different object primitives which have homogeneous colour, similar texture, and constrained shape which when combine with spectral information results in accurate image segments then using spectral data alone. This paper proposes a system for increasing the degree of automation in extraction of building features from high resolution multispectral satellite images by identification of a suitable classification method for the above object primitives.

1.1. Study Area and Data Used

Chandigarh is a city and a union territory of India that serves as the capital of the Indian states of Punjab and Haryana. Chandigarh is bordered by the state of Punjab to the north, west and south, and to the state of Haryana to the east. Chandigarh was one of the early planned cities in the post-independence India and is internationally known for its architecture and urban design. The master plan of the city was prepared by Swiss-French architect Le Corbusier. It is located near the foothills of the Sivalik range of the Himalayas in northwest India. It covers an area of approximately 114 km^2. It has an average elevation of 321 metres (1053 ft).

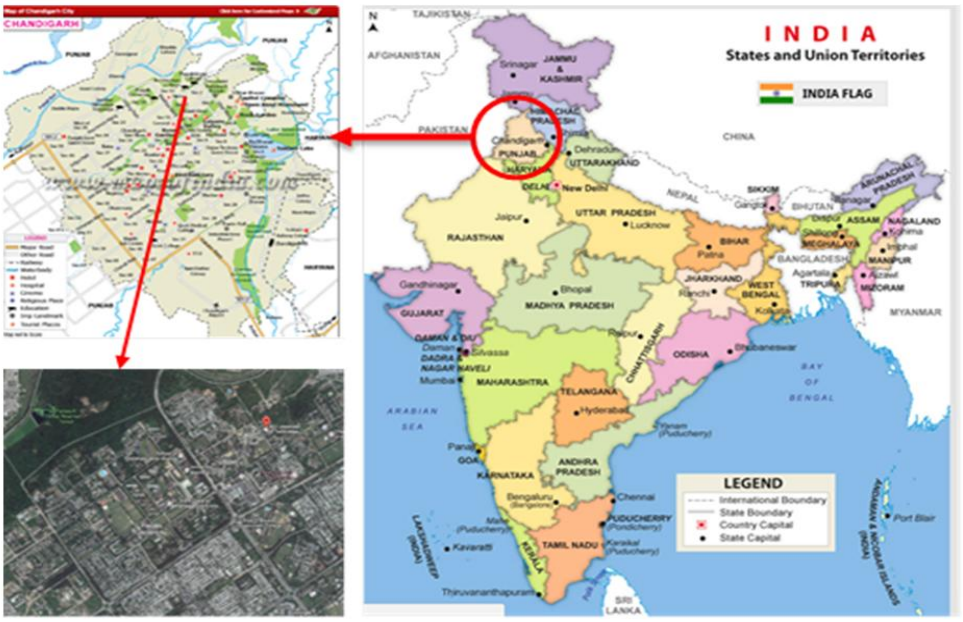


Figure 1: *Location map of study area*

Chandigarh has a humid subtropical climate characterized by a seasonal rhythm: very hot summers, mild winters, unreliable rainfall and great variation in temperature (−1 °C to 46 °C OR 30.2 °F to 114 °F). The average annual rainfall is 1110.7 mm. The city also receives occasional winter rains from the Western Disturbance originating over the Mediterranean Sea. As of 2011 India census, Chandigarh had a population of 1,055,450, [1] making for a density of about 9,252 (7,900 in 2001) persons per square kilometer. There are numerous educational institutions in Chandigarh like Panjab University, Post Graduate Institute of Medical Education and Research (PGIMER), Government Medical College & Hospital, PEC (Punjab Engineering College) University of Technology and many other private and public institutions (https://en.wikipedia.org/wiki/Chandigarh). The test area is a part of Chandigarh covering major institutions. The location map of Study area is presented in Figure 1. The satellite data used is a part of Chandigarh captured by World View 2 Satellite on May 11, 2015. The specifications of satellite data are tabulated in Table 1.

Table 1: *Satellite data specifications*

Satellite Data Used : Worldview 2	
Resolution	Panchromatic: 0.46 meters GSD at nadir*, 0.5 meters GSD at 20° off-nadir
	Multispectral: 1.84 meters GSD at nadir*, 2.0 meters GSD at 20° off-nadir
Wave length Range Panchromatic Band	450 - 800 nm
	8 Bands
Multispectral Bands Wave length Range	Coastal: 400 - 450 nm
	Blue: 450 - 510 nm
	Green: 510 - 580 nm
	Yellow: 585 - 625 nm
	Red: 630 - 690 nm
	Red Edge: 705 - 745 nm
	Near-IR1: 770 - 895 nm
	Near-IR2: 860 - 1040 nm

2. Materials and Methods

The method involves computation of textural parameters from an original high resolution multispectral image and combined with the spectral bands for quantification of spectral-structural characteristics of a pixel. The multi-resolution image segmentation is then applied on this combined spectral-structural image, resulting in the formation of the different level of polygon primitives at different space scale. Hence in addition to the spectral information, the tone, texture and shape information is evaluated for an object-oriented analysis. Object primitives which have homogeneous colour, similar texture, and constrained shape are generated providing us different view of the scene at different resolution. The segmentation guided classification rules were applied on the previously generated object primitives. The methodology flow is presented in Figure 2.

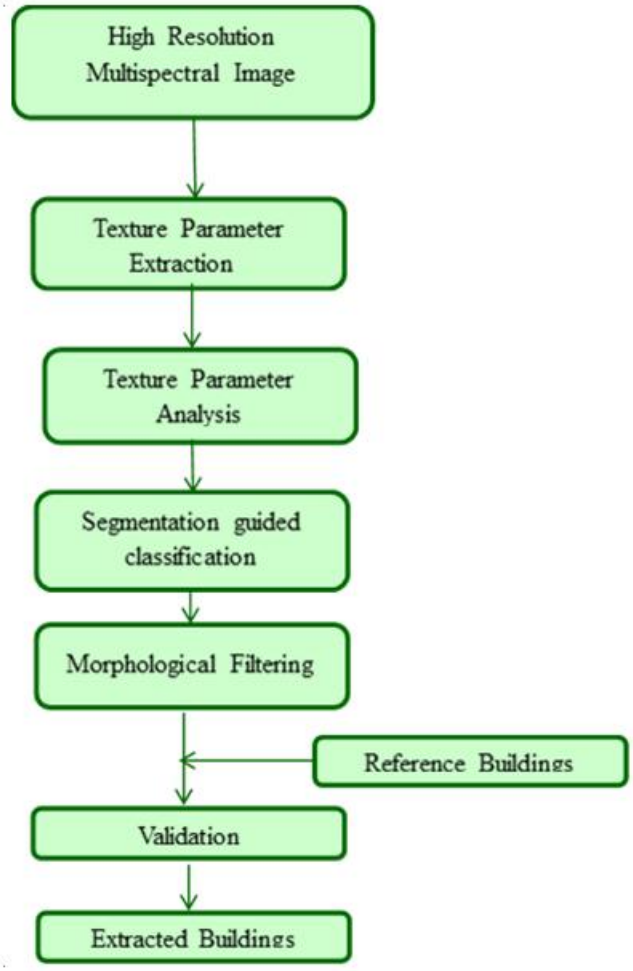


Figure 2: *Methodology flowchart*

3. Results and Discussion

The Grey level co-occurrence matrix (GLCM) was computed and second order textures parameters were computed on the eight bands of World view 2 Imagery. The GLCM based measures can be

grouped into Contrast Group which uses weights related to distance from GLCM diagonal. This group has Contrast, Dissimilarity and Homogeneity textures. The Orderliness group defines the orderliness of a window and uses weighted average of GLCM elements. This group has Angular Second Moment, Energy and Entropy parameters. The third group consists of GLCM mean, GLCM variance and GLCM Correlation is statistics which can be computed. It was observed that Contrast, Dissimilarity behaved similarly as GLCM variance and GLCM Entropy behaved similar to GLCM Entropy. GLCM mean and Correlation are more independent.

In the present case study GLCM data range, dissimilarity aided in boundary detection. GLCM Entropy was useful in discriminating Vegetation areas and the colour composite of GLCM Mean, Variance and Correlation segregated the shadows. The layers were used for segmentation using Multiresolution approach with scale parameter varying between 25- 40 for different shape and sized buildings. A higher scale parameter was used for larger buildings and a smaller one for large urban areas with small houses. The image was then classified using segmentation guided classification. Accuracy

assessment was done for building classification using test samples. Accuracy achieved for buildings extracted was 87% after combination of Textural information as compared to 76% without structural information.

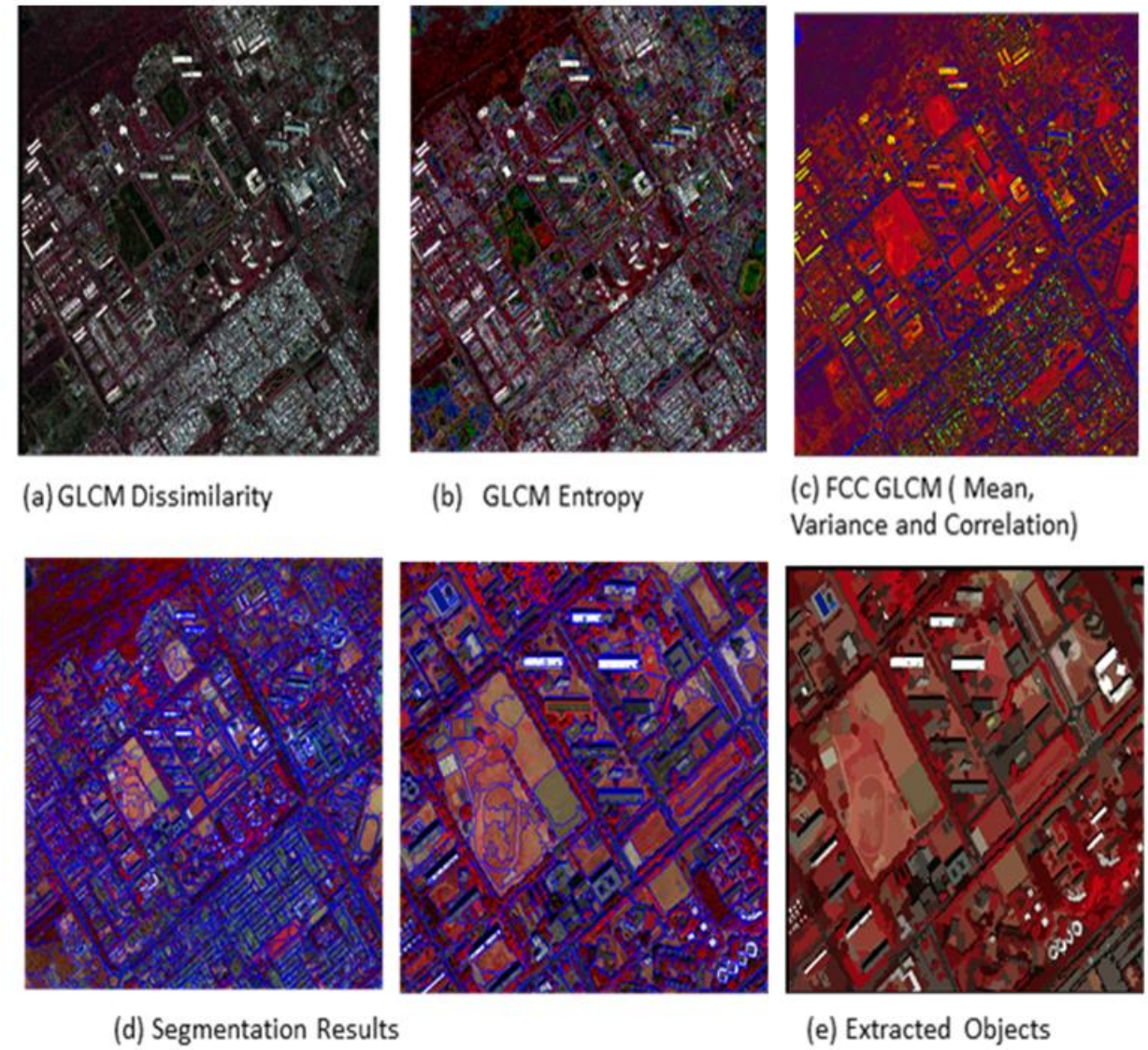


Figure 3: *Results*

4. Conclusion

The accuracy of segmentation results depend upon the use appropriate parameter values in image segmentation. Most studies have proposed trial and error methods of determining segmentation parameters yet there is no standard or widely accepted method. The determination of segmentation parameters is dependent on the analyst's experience. Segmentation evaluation is an essential dispensable processing step. In this study visual inspection and reference based discrepancy method was employed to fully evaluate the effectiveness of the proposed method.

Visual inspection is still one of the most commonly used methods. It considers the geometrical position as well as the membership of one object class to a single region. For quantitative evaluation of segmentation, the discrepancy based method was adopted, where the segmentation result is quantitatively compared with reference data. The reference data were acquired by manually

delineating polygons from different land cover classes. Accuracy achieved for buildings extracted was 87% after combination of Textural information as compared to 76% without structural information. The development of new and sophisticated image segmentation methods is crucially important for object-based analysis. High Resolution images usually have high spatial information; the combination of spectral and spatial information in image segmentation may produce more accurate results. This study proposed a new image segmentation method using spectral and textural information. Experimental results demonstrate that the joint use of spectral and structural information outperformed the use of either information alone. The proposed image segmentation method is well applicable to the segmentation of imagery over urban and suburban areas for large scale building and road extraction.

References

Aguilar, M.A, Vicente, R., Aguilar, F.J., Fernández, A. and Saldaña, M.M. 2012. *Optimizing object-based classification in urban environments using very high resolution geoeye-1 imagery*. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Melbourne, Australia.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, R.Q., van der Meer, F., van der Werff, H., van Coillie, F. and Tiede, D. 2014. Geographic object based image analysis – towards a new paradigm. *J. Photogrammetry & Remote Sens.*, 87, pp.180-191.

Baltsavias, E.P. 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58, pp.129-151.

Barrile, V. and Bilott, G. 2008. An application of remote sensing: object-oriented analysis of satellite data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII, Part B8.

Su, W., Li, J., Chen, Y., Liu, Z., Zhang, J., Low, T., Suppiah, I. And Hashim. S.A.M. 2008. Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. *International Journal of Remote Sensing*, 29, pp.3105-3117.

Kumar, M. and Roy, P.S. 2013. Utilizing the potential of World View-2 for discriminating urban and vegetation features using object based classification techniques. *Journal of Indian Society of Remote Sensing*, 41 (3).

Kumar, M., Tiwari, P.S., Pande, H., and Dadhwal, V.K. 2009a. *A semi-automatic segmentation approach procedure for agriculture field boundary extraction in remotely sensed imagery*. Indian Society of Remote Sensing Symposia, Nagpur, India.

Kumar, M. and Shrivastava, N. 2009b. *Multiresolution segmentation and classification based building extraction using merged CARTOSAT and LISS IV Data*. Indian Society of Remote Sensing Symposia, Nagpur, India.

Ettarid, M., Rouchdi, M. and Labouab, L. 2008. Automatic extraction of buildings from high resolution satellite images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII, Part B8.

Xiaoying, J. 2005. Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information. *EURASIP Journal on Applied Signal Processing*, 14, pp.2196-2206.

Liu, Z.J., Wang, J. and Liu, W.P. 2005. *Building extraction from high resolution imagery based on multi-scale object oriented classification and probabilistic Hough transform*. International Geoscience and Remote Sensing Symposium.

Kumar, M., Singh, R.K., Raju, P.L.N. and Krishnamurthy, Y.V.N. 2014. Road network extraction from high resolution multispectral satellite imagery based on object oriented techniques. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-8, pp.107-110.

Chandigarh. Available from: https://en.wikipedia.org/wiki/Chandigarh, Accessed on 20/6/2017.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2478-2517
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.337

Research Article

Estimation of the Near-Surface Air Temperature during the Day and Nighttime from MODIS in Berlin, Germany

Forough Marzban[1], Tim Conrad[2], Pouria Marzban[3], Sahar Sodoudi[1]

[1]Institut für Meteorologie, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany
[2]Department of Mathematics and Computer Science, Freie Universität, Berlin, Germany
[3]Department of Computer Engineering, Islamic Azad University, Sari, Iran

Publication Date: 22 February 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.337

Abstract Air temperature (T_{air} or T_{2m}) is an important climatological variable for forest biosphere processes and climate change research. Due to the low density and the uneven distribution of weather stations, traditional ground-based observations cannot accurately capture the spatial distribution of T_{air}. In this study, T_{air} in Berlin is estimated during the day and night time over six land cover/land use (LC/LU) types by satellite remote sensing data over a large domain and a relatively long period (7 years). Aqua and Terra MODIS (Moderate Resolution Imaging Spectroradiometer) data and meteorological data for the period from 2007 to 2013 were collected to estimate T_{air}. Twelve environmental variables (land surface temperature (LST), normalized difference vegetation index (NDVI), Julian day, latitude, longitude, Emissivity31, Emissivity32, altitude, albedo, wind speed, wind direction and air pressure) were selected as predictors. Moreover, a comparison between LST from MODIS Terra and Aqua with daytime and night time air temperatures (T_{day}, T_{night}) was done respectively and in addition, the spatial variability of LST and T_{air} relationship by applying a varying window size on the MODIS LST grid was examined. An analysis of the relationship between the observed T_{air} and the spatially averaged remotely sensed LST, indicated that 3 × 3 and 1 × 1 pixel size was the optimal window size for the statistical model estimating T_{air} from MODIS data during the day and night time, respectively. Three supervised learning methods (Adaptive Neuro Fuzzy Inference system (ANFIS), Artificial Neural Network (ANN) and Support vector machine (SVR)) were used to estimate T_{air} during the day and night time, and their performances were validated by cross-validation for each LC/LU. Moreover, tuning the hyper parameters of some models like SVR and ANN were investigated. For tuning the hyper parameters of SVR, Simulated Annealing (SA) was applied (SA-SVR model) and a multiple-layer feed-forward (MLF) neural networks with three layers and different nodes in hidden layers are used with Levenber-Marquardt back-propagation (LM-BP), in order to achieve higher accuracy in the estimation of T_{air}. Results indicated that the ANN model achieved better accuracy (RMSE= 2.16°C, MAE = 1.69°C, R^2 = 0.95) than SA_SVR model (RMSE= 2.50°C, MAE = 1.92°C, R^2 = 0.91) and ANFIS model (RMSE= 2.88°C, MAE= 2.2°C, R^2 = 0.89) over six LC/LU during the day and night time. The Q-Q diagram of SA-SVR, ANFIS and NN show that all three models slightly tended to underestimate and overestimate the extreme and low temperatures for all LC/LU classes during the day and night time. The weak performance in the extreme and low temperatures are a consequence of the small number of data in these temperatures. These satisfactory results indicate that this approach is proper for estimating air temperature and spatial window size is an important factor that should be considered in the estimation of air temperature.

Keywords *ANFIS; ANN; Cross-validation; MODIS; Simulated annealing; SVR*

1. Introduction

The standard meteorological T_{air} is measured in a shelter at 2m height (Brunel, 1989; Jin and Dickinson, 2010). It is an important indicator of terrestrial environmental conditions across the earth (Prihodko and Goward (1997); Peón et al., 2014) and one of the most widely used climatic variables in global change studies. It plays an important role in multiple biological and physical processes among the hydrosphere, atmosphere and biosphere (Stisen et al., 2007; Shamir et al., 2014; Benali et al., 2012). Regarding ecosystem, it influences the distribution of plant species (Cabrera 2002) and affects the dynamics of the soil-plant-water system (Chartzoulakis and Psarras, 2005; Zavala, 2004), being included in evapotranspiration models (Allen et al., 2006; Carlson et al., 1995) as well as hydrological models (Purkey et al., 2007; Yates et al., 2005). At the individual level, temperature affects plant growth and net primary productivity since photosynthetic and respiration rates depend on it. Moreover, T_{air} plays a critical role in vegetation distributions, phenology, and growth (Benavides et al., 2007; Stahl et al., 2006). The maximum temperature also shows significant relationship with the occurrence of wildfire on hot and sunny days (Aldersley et al., 2011; Litschert et al., 2012). Therefore, detailed knowledge of the spatial variability of air temperature is of interest for many research and management.

In addition, T_{air} plays an important role in energy balance and is a key input in various environmental models and applications, such as crop evapotranspiration estimation (De Bruin et al., 2010), distributed hydrology (Gao et al., 2014) and climate change models (Lofgren et al., 2011). Moreover, the importance of temperature in urban area are related to heat stress and human health. Meteorological measurements provide accurate discrete T_{air} information for specific locations but have limited ability to describe its spatial heterogeneity over large areas (Benali et al., 2012; Willmott and Robeson, 1995). The non-uniform spatial distribution of weather station locations within most networks and the complexity of the land surface conditions and patterns make it a challenge to get spatial-continuous T_{air} data.

However, weather stations are usually sparsely distributed in mountainous regions, especially in high-elevation areas, and thus may not optimally represent all environments (Rolland, 2003). Given the large spatial heterogeneity of T_{air} in complex terrain (Holden et al., 2011), it is difficult to accurately characterize the distribution of T_{air} over mountainous areas (Carrega, 1995). Different interpolation methods have been used to generate spatially continuous Tair from point station measurements (Benavides et al., 2007; Dodson and Marks, 1997; Duhan et al., 2013; Kurtzman and Kadmon, 1999; Stahl et al., 2006). However, the performance of interpolation methods is highly dependent on the spatial density and distribution of weather stations (Chan and Paelinckx, 2008; Vogt et al., 1997), which is not considered satisfactory in mountainous areas.

Satellite remote sensing observations from global imaging sensors, such as the Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS), represent a potentially valuable alternative to characterize spatially-detailed Tair patterns across large areas. A split window technique was applied to AVHRR (Pinheiro et al. 2006), MODIS (Wan et al., 2002), and Meteosat (Atitar and Sobrino, 2009) thermal data to estimate Land Surface Temperature (LST). The science-grade quality of the LST data collected by MODIS has proven valuable for monitoring land surface dynamics over large areas (Benali et al., 2012, Mostovoy et al., 2006, Lin et al., 2012).

The earth's surface is heated by solar radiation, while the atmosphere is mainly heated from the ground up through longwave infrared radiation (Frederick et al., 2006). The relationship between Land

Surface Temperature and T_{air} may vary with time and location, because the land surface energy balance is a complex phenomenon that depends on multiple factors (e.g., cloud cover, surface roughness, wind speed and soil moisture), whereas some of them (e.g., wind speed) are usually not available from satellite (Goward et al., 1997; Prince et al., 1998; Stisen et al., 2007).

An accurate estimation of T_{air} and the mapping of its spatial distribution are useful for predicting ecological consequences of climate change. For example, climate warming will lead to higher temperatures and an increase of extreme weather conditions, which are associated with changes in wildfire regime (Westerling et al., 2006; Chen et al., 2011; Manzo-Delgado et al., 2009), forest biomass distribution (Reich et al., 2014) and crop yield (Ruane et al., 2014; Rosenzweig et al., 2014). The demand for accurate spatial T_{air} data over a large scale has continued to rise (Oyler et al., 2015; Beier et al., 2012). However, the spatial distribution of the weather stations in many parts of the world, is often limited which restricts the use of T_{air} measurements over a large spatial domain (Vancutsem et al., 2010). LST, but on the other hand, is measured in a global extent with significant higher spatial coverage (Jin and Dickinson, 2010). The US National Research Council and the Intergovernmental Panel on Climate Change (IPCC) expressed the need for long-term remotely sensed LST data in global warming studies to overcome the limits of conventional surface T_{air} measurements (IPCC, 2007, Jin, 2004). Remote sensing data has great potential to estimate spatial-temporal patterns of T_{air} which can further our knowledge, on both the climate and terrestrial biological processes at regional and global scales (Benaliet al., 2012). Monitoring and understanding the trends of T_{air} and LST are crucial in the study of regional and global climate changes (Yoo et al., 2011). LST can be monitored and modelled from multiple daily satellite observations, such as the MODIS LST. Studies have shown that LST can be used for linear regression estimates of daily minimum and maximum T_{air} on a local scale (Mostovoy et al., 2006; Vancutsem et al., 2010; Zhang et al., 2011a; Yoo et al., 2011; Evrendilek et al., 2012; Benali et al., 2012; Zhu et al., 2013). Cresswell et al. (1999) found an over and underestimation of T_{air} during the day and at night, respectively, from Meteosat LST observations. They attempted to correct these errors and produce a proxy of T_{air} T_{air} by applying a solar zenith angle correction on the Meteosat geostationary observations. They achieved an accuracy of 3°C for over 70% of the Meteosat temperatures. Similarly, Jin and Dickinson (2010) have studied the differences in the diurnal cycles of LST and T_{air} over a single site. Some studies (Florio et al., 2004) have used several statistical approaches that combined a simple AVHRR Spilt-Window Technique (SWT) with ground meteorological station measurements in the prediction of T_{air}. Other studies (Wloczyk et al., 2011) have used the Landsat LST data to derive T_{air}. They have attempted to assign the satellite-derived T_{air} to a certain height above the ground and have investigated the possibility of a simple correction for reference height. They also considered the link between T_{air} spatial pattern and the window-size of the Landsat LST pixels. Xu et al. (2012) used four empirical regression models to estimate the relationship between T_{air} measurements and the MODIS-Aqua LST and found different relationships between the two different LC types in their study. They also assessed the effect of the MODIS LST window-size on the agreement between the two variables and found that spatial averaging over multiple pixels improves the accuracy of T_{air} estimates. Zaksek and Schroedter-Homscheidt (2009) reviewed the types of methods commonly used to estimate T_{air} based on LST, dividing them into three distinct groups:

1) Statistical approaches which are based on regression techniques, can be simple if only based on LST and T_{air} (e.g. Mostovoy et al., 2006; Vogt et al., 1997) or advanced, when more than one independent variable is used such as solar zenith angle (SZA), elevation, altitude, Julian day among others (Lin et al., 2012; Cresswell et al., 1999; Jang et al., 2004). Lin et al. (2012) used stepwise linear regression method to estimate daily maximum air temperature (T_{max}) and daily minimum air temperature (T_{min}) with MAE = 1.9, agreement index = 0.79 and MAE = 1.9 °C, agreement index = 0.92, respectively, over east Africa. Fu et al. (2011) used linear regression between MODIS LST and

T_{max} from stations on the northern Tibetan Plateau. In general, these methods perform well within the spatial and time frame they were developed, but the accuracy might decrease when extended in time and space (Stisen et al., 2007). Statistical methods generally perform well within the spatial and time frame they were derived in, but have limited generalization and require large amounts of data to train the algorithms (Stisen et al., 2007).

2) The second category is index-based such as Temperature-Vegetation index (TVX). It is based on the assumption that for an infinitely thick canopy, the top-of-canopy temperature is the same as within the canopy (Czajkowski et al., 2000; Prihodko and Goward (1997), Nemani and Running et al., 1989; Nieto et al., 2011) and uses the Normalized Difference Vegetation Index (NDVI) as a key input variable. However, the assumption of linear and negative slope between LST and NDVI is not always applicable and is influenced by the seasons, the type of ecosystem and soil moisture variability (Sandholt et al., 2002; Vancutsem et al., 2010). Zhu et al. (2013) used the TVX method to estimate daily T_{max} with RMSE (the root mean square error) =3.709 °C, MAE (the mean absolute error) = 3.03 °C and r (correlation coefficient) = 0.83 in Xiangride River Basin of China. However, Vancutsem et al. (2010) found that TVX method did not adapt to different ecosystems over Africa because non-significant relationship between LST and NDVI in their study. Karnieli et al. (2003) found that the approaches based on this negative NDVI/LST relationship have minimal utility in energy-limited environments (e.g., high latitude and elevations) compared to moisture-limited environments because vegetation-expressed NDVI response is more related to available solar radiation than land surface conditions (e.g. soil moisture).

3) The final approach uses surface energy balance parameterizations based on physically-based models (Sun et al., 2005). The sum of incoming net radiation is considered equal to the sum of the soil heat flux, sensible flux and latent heat flux (Zakšek and Schroedter-Homscheidt, 2009; Meteotest 2010; Sun et al., 2005). However these methods require large amounts of information that are usually not only from remote sensing (e.g., roughness, soil physical properties) (Benali et al., 2012, Mostovoy et al., 2006, Prince et al., 1998).

Most of the previous studies have focused on daily estimations or instantaneous T_{air}. The TVX method has been widely used for T_{air} estimation. Czajkowski et al. (2000) estimated Tavg for a weekly period with associated RMSE between 1.72 and 3.48 °C and R2=0.64. Stisen et al. (2007) and Prihodko and Goward (1997) estimated T_{air} with RMSE higher than 2.5 °C and R^2 between 0.64 and 0.86. Cresswell et al. (1999) used a statistical method to derive instantaneous T_{air} with an associated RMSE below 3 °C for more than 70% of the sampled data. Zaksek and Schroedter-Homscheidt (2009) used a more sophisticated method, which was based on the energy balance to estimate instantaneous T_{air} with an RMSE of 2°C. Vancutsem et al. (2010) used 1 km MODIS data to estimate weekly T_{min} and T_{max}. They reported correlations between LST and T_{min} ranging from 0.01 to 0.96 for several stations and T_{max} was estimated with an R^2=0.92 and RMSE=1.83 °C.

Moreover, in previous studies, several variables were employed to estimate air temperature. For example, the variables used by Benali et al. (2012) included LST, Julian Day, elevation, and the distance to coast. Benali et al. (2012) used both weekly daytime LST data (LST_{day}) and night time LST data (LST_{night}) to estimate the average, maximum and minimum weekly temperature. They found that there was a higher correlation between average weekly temperature and averaged weekly LST_{night}, which indicates the potential of LST_{night} in estimating averaged weekly temperature. The variables used by Kim and Han (2013) included LST, NDVI, altitude, and solar zenith angle. The variables used by Cristóbal, Ninyerola and Pons (2008) included LST, NDVI, and albedo. The variables used by Zakšek and Schroedter-Homscheidt (2009) included LST, NDVI, solar zenith, albedo, solar radiation,

and altitude. After comprehensive consideration of these variables, twelve variables were selected as the predictors for the modelling of air temperature during the day and night time: LST, NDVI, Julian day, latitude, longitude, Emissivity31, Emissivity32, altitude, albedo, wind speed, wind direction and air pressure.

The main objective of this study was to estimate the air temperature during day and night time with high spatial resolution in Berlin from Moderate Resolution Imaging Spectroradiometer (MODIS) data by for different land cover types.

First, this research presents the comparison of state-of-the-art remote sensing-based LST data from MODIS with T_{air} for the six LC/LU. Within this study, we compared the relationship between T_{air} and the Four LST products of MODIS over Berlin in order to analyze the agreement between LST from MODIS Terra and Aqua and T_{air} for the period of 2007 to 2013 based on different land cover classes. Specifically, the spatial scale effects of the relationship between T_{air} and LST were first analysed to determine the best window size to retrieve T_{air} in the study area. The comparison is done by using statistical parameters such as the correlation coefficient, the slope and the intercept with the y-axis of the regression line, mean bias error (MBE), and normalized mean bias which known as bias. The MBE is calculated by the difference between LST and T_{air} divided by the amount of observed time steps. If the MBE is positive, the LST detects warmer temperatures than the measured T_{air}, and vice versa (Hachem et al., 2012). Then Adaptive neuro fuzzy system (ANFIS), Artificial Neural Network (ANN) and support vector machine (SVR) models were developed to estimate T_{air} , and the accuracy of these models were assessed by comparison with the observed air temperature data from weather stations and the cross validation (CV) approach, in order to find the best model with high accuracy during the day and night time. The errors associated with T_{air} estimation based on remote sensing data are often large and strongly limit its applicability (e.g. Czajkowski et al., 2000; Vazquez et al., 1997; Vogt et al., 1997). One of the objectives of this work is to provide T_{air} estimations with an accuracy, which will potentiate the future applications. Moreover, tuning the hyper parameters of some models like SVR and ANN were investigated. In order to select the hyper parameters of SVR, Simulated Annealing (SA) was applied and a multiple-layer feed-forward (MLF) neural networks with three layers and different nodes in hidden layers are used with Levenberg–Marquardt back-propagation (LM-BP) in order to achieve higher accuracy in the estimation of T_{air} during the day and night time over six LC/LU.

2. Materials and Methods

2.1. The Study Area

Berlin is the capital city of Germany. It is located in the northeast of the country, covers an area of 892 km^2. Berlin is located on a mostly flat topography. Regarding land use patterns, Berlin is characterized by a significant amount of green areas and water bodies. Outside the inner city, there is relatively low buildings and population density, with many allotment gardens for private cultivation and recreation. There are a considerable number of urban brownfield sites, despite the trend of population growth in the last decade. Berlin consists of 45% water bodies and urban green spaces (forested and unforested, allotment gardens), almost 20% transport and infrastructural areas (streets and railways), and around 35% built-up areas (e.g. for residential use). Table 1 shows the location and the related land use of the weather stations in Berlin which are used in this study.

2.2. Data Description

Three main datasets for the period of 2007-2013 according to the availability of meteorological station record and MODIS data were used:

1. Ground measurements from 20 meteorological stations in Berlin.
2. Remotely sensed data.
3. Digital elevation model (Berlin Digital Environmental Atlas).

Table 1: *Information about weather stations over Berlin including their LC/LU, latitude, longitude and elevation*

Station	LC/LU	Lat	Long	Elevation(m)
Botanischer-Garten	Green urban area	52.45	13.30	46.88
Fasanen	Industrial, commercial, public, military	52.51	13.33	34.08
Tegel-Forstamt	Forest	52.60	13.27	39.58
Gatow	Industrial, commercial, public, military	52.47	13.13	47.09
Marzahn1	Green urban area	52.54	13.58	50.61
Pichelsdorf	Evergreen needle leaf tree	52.50	13.19	29.66
Wannsee	Evergreen needle leaf tree	52.43	13.18	40.77
Dahlem-FU	Industrial, commercial, public, military	52.45	13.31	67.50
Tegel	Airport	52.56	13.30	35.25
Schonefeld	Airport	52.38	13.53	45
Buch	Industrial, commercial, public, military	52.63	13.50	65.45
Marzahn2	Green urban area	52.54	13.55	63.29
Kaniswall	Agriculture, semi-natural and wet area	52.40	13.73	32.57
Tempelhof	Airport	52.46	13.40	47.74
Eiskeller	Agriculture, semi-natural and wet area	52.58	13.13	31.78
Kreuzberg	Industrial, commercial, public, military	52.49	13.40	34.91
Wannsee-meteo	Evergreen needle leaf tree	52.43	13.18	43.49
Adlershof	Industrial, commercial, public, military	52.42	13.52	35.15
Potsdam	Industrial, commercial, public, military	52.38	13.11	33.79
Insulaner	Green urban area	52.45	13.35	43.75

2.2.1. Meteorological Data

Air temperature observations were obtained from 20 different meteorological ground stations in this study area. The measurements included daily T_{air}, wind speed, wind direction, air pressure and Julian day. The meteorological station records were obtained from the Deutscher Wetterdienst (ftp://ftp-cdc.dwd.de/pub/CDC) and from the Freie university Berlin meteorological station (http://mevis-www.met.fu-berlin.de/devel/mevis). The accuracy of observation data in meteorological stations is as following:

1. 2m air temperature: ± 0.2 K
2. Wind speed: ± 0.3 of measured value
3. Wind direction: ± 5C^{o}
4. Relative humidity: ± 0.3% up to 0.5%
5. Air pressure: ± 0.1 hpa

2.2.2. MODIS Data

The second source of data is the satellite data. MODIS sensors were launched on board the National Aerodynamics and Space Administration (NASA) Observing System (EOS) Terra and Aqua satellites in December 1999 and May 2002, respectively (Zhu et al., 2013). Both sensors are on board sun-synchronous polar orbiting satellites. MODIS Terra data is available during 10:30–12:00 a.m. and p.m. (daytime/night time) local time, while MODIS Aqua sensor collects the imagery during 1:00–3:00 a.m. and p.m. (daytime/night time). In this study, the following products of MODIS were used:

1. MODIS daily land surface temperature at 1km resolution, from Terra (MOD11A1.005)
2. MODIS daily land surface temperature at 1km resolution, from Aqua (MYD11A1.005)
3. MODIS monthly vegetation index at 1km resolution from Terra (MOD13A2.005)
4. MODIS monthly vegetation index at 1km resolution from Aqua (MYD13A2.005) product

Daily MODIS LST and monthly NDVI from Aqua and Terra were extracted at the nearest points to the stations. The LST product from MODIS has been used in previous studies to derive Tair (Benali et al., 2012, Vancutsem et al., 2010, Zhu et al., 2013). All MODIS LST data used in this study were acquired from the U.S. Geological Survey (USGS) website (Piao et al., 2009). We used two MODIS LST products MOD11A1 and MYD11A1 from Terra and Aqua satellites, respectively. The MODIS LST consists of daytime and night time data at a spatial resolution of 1 km. Thus, in total there are four LST datasets: Aqua daytime, Aqua night time, Terra daytime and Terra night time.

2.2.2.1. *Vegetation Index*

Normalized difference vegetation index is the most common remote sensing index used to parameterize vegetation status (Zhu et al., 2013, Stow et al., 2004, Raynolds et al., 2008). The absorption and reflectivity of the vegetation cover are correlated with their structural properties, such as leaf area index (LAI), fractional vegetation cover (FVC), and their physiological condition (Bustos et al.2014, Raynolds et al., 2006). The values of NDVI vary between −1 and 1, where the range between 0.2 and 0.9 is mostly common in continuous vegetation cover (Bustos et al.2014). In this study, the NDVI was extracted from Terra (MOD13A2.005) and Aqua (MYD13A2.005) products with 16-day temporal and 1km resolution as mentioned in Table 2.

Table 2: *Data source and variables*

Variable	Source	Explanation
Land surface temperature	MODIS	Land Surface temperature derived over the 2007–2013 time period using MYD11A1, MOD11A1 product
Julian Date	Meteorological data	The continuous count of days was from 1 January to the last day every year
Emissivity31	MODIS	Emissivity31 derived over the 2007–2013 time period using MYD11A2, MOD11A1 product
Emissivity32	MODIS	Emissivity32 derived over the 2007–2013 time period using MYD11A2, MOD11A1 product
Normalized Difference Vegetation Index	MODIS	vegetation index at 1km resolution from Terra (MOD13A2.005) and Aqua (MYD13A2.005) products with 16day temporal resolution
Albedo	MODIS	Albedo at 1Km resolution from Terra (MCD43B3.005) product with 16 day temporal resolution
Relative humidity	Meteorological Data	The RH was extracted for each station during the year of 2007–2013
Altitude	DEM	The altitude extracted from a 5m resolution digital elevation model (DEM) according to the location of meteorological stations
Latitude	Meteorological data	The geographical location of meteorological stations was extracted from meteorological metadata
Wind direction	Meteorological data	The WD was extracted for each station during the year of 2007–2013
Wind speed	Meteorological data	The WS was extracted for each station during the year of 2007–2013
Air pressure	Meteorological data	The AP was extracted for each station during the year of 2007–2013

2.2.3. Auxiliary Data

In addition to MODIS products (LST, emissivity31, emissivity32), Albedo and NDVI, some auxiliary variables were used, including latitude, altitude, Julian day, air pressure, wind speed, wind direction and relative humidity which is presented in Table 2. These auxiliary variables either have a known impact on T_{air} and LST or influence the relationship between T_{air} and LST. Latitude, LC/LU and altitude were derived from the location of meteorological stations. Altitude was obtained from a 5m resolution digital elevation model (DEM) (downloaded from https://www.eea.europa.eu/data-and-maps/data/urban-atlas). Moreover, Julian day was also considered as proxies for the fraction of solar energy absorption during the day and emission during the night, influencing the diurnal amplitude of T_{air} throughout the year. Julian day is the continuous count of days from 1 January every year. In addition, the LC/LU of each meteorological station was extracted in terms of its position, and reclassified into urban, industrial, forest, airport, needle leaf trees and agriculture based on a 5m resolution map of LC/LU, which was downloaded from https://www.eea.europa.eu/data-and-maps/data/urban-atlas.

In addition, all data (Auxiliary and MODIS data) were combined to create a single dataset for each LC/LU for day and night time. The collinearity of independent variables was detected using variance inflation factor (VIF > 10) and pair wise correlation (r > 0.75) (Zurr et al., 2010, Dormann et al., 2013).

2.3. LST Pre-Processing

A certain number of pre-processing steps were required to convert the original LST product in HDF format to raster layers with a versatile projected coordinate system. Firstly, raster subsets of the LST product were extracted based on the boundary extent of the study area. LST L3 product is gridded in the global Sinusoidal projection, and the grid containing data for the study area is located at column 18 (h18) and line 03 (v03). It is important to eliminate low quality data in the MODIS LST data because remote sensing based T_{air} estimates are strongly influenced by errors (e.g., errors caused by clouds and large sensor viewing angles, uncertainties in surface emissivity (Wan et al., 2004). Validation studies of MODIS LST show that under clear sky conditions the precision is approximately 1 K or less, but higher errors would be observed at large viewing angles and in semiarid regions (Wan et al., 2008). So only the pixels of the targeted land cover types that were flagged in the MODIS quality assurance data as cloud free and of high quality were retained.

2.3.1. Calculating LST of Weather Station Location

LST data under clear sky conditions at weather stations are retrieved by the following steps:

- A total of 5110 MODIS HDF format (MOD11A1 and MYD11A1, h18v03, Collection 5, from 1 January 2007 to 31 December 2013 over Berlin) in HDF (Hierarchical Data Format) format were re-projected to WGS_1984_UTM_ zone_33N using the nearest neighbor resampling method. The corresponding layers (LST_Day_1km, LST_Night_1km, Daytime LST observation time, and Night time LST observation time) were extracted. However, Daytime and Night time LST observation time were used in order to identify the approximate overpass time of MODIS at local time.
- MODIS LST data for the pixels in which the weather stations are located are extracted from MODIS using nearest neighbor algorithm.
- All these LST data (DN value) were converted to Celsius temperature using the following equation:

$$T\ (C^{o}) = 0.02 * DN - 273.15 \ldots (1)$$

Where is the C^{o} is the Celsius temperature and 0.02 is the scale factor of the MODIS LST product.

- Removing low quality data: MODIS LST products are not available for a location (pixel) if clouds are present (Wan 2008). However, there are some pixels that are lightly covered or contaminated by clouds. These pixels are not removed because the contamination is very small and cannot be detected by the cloud-removing mask algorithm (Ackerman et al., 2008, Williamson et al., 2013; Xu and Shen 2013). To avoid this kind of data, only the pixels of the targeted land cover types that were flagged in the MODIS quality assurance data as cloud-free and of high quality were retained.

2.4. The Relationship between Observed T_{air} and the Four LST Products from MODIS

The influence of the time of observation on the estimation of T_{air} has been discussed in several studies, which resulted in different conclusions. Benali et al. (2012) stated that the use of both aqua LST_{day} and LST_{night} could improve the estimation of T_{day} and T_{night} , (T_{day} and T_{night} are not the maximum and minimum temperature of a day and night time)respectively, because the MODIS Aqua overpass time is closer to the time of both T_{day} and T_{night} than Terra's. In contrast, Zhu et al. (2013) showed that both terra LST_{day} and LST_{night} were better than aqua LST_{day} and LST_{night} for T_{air} estimations in Xiangride River basin of China. In another study, Mostovoy et al. (2006) found that the difference between the satellite overpass (Terra and Aqua) had little impact on the estimation accuracy of T_{air}.

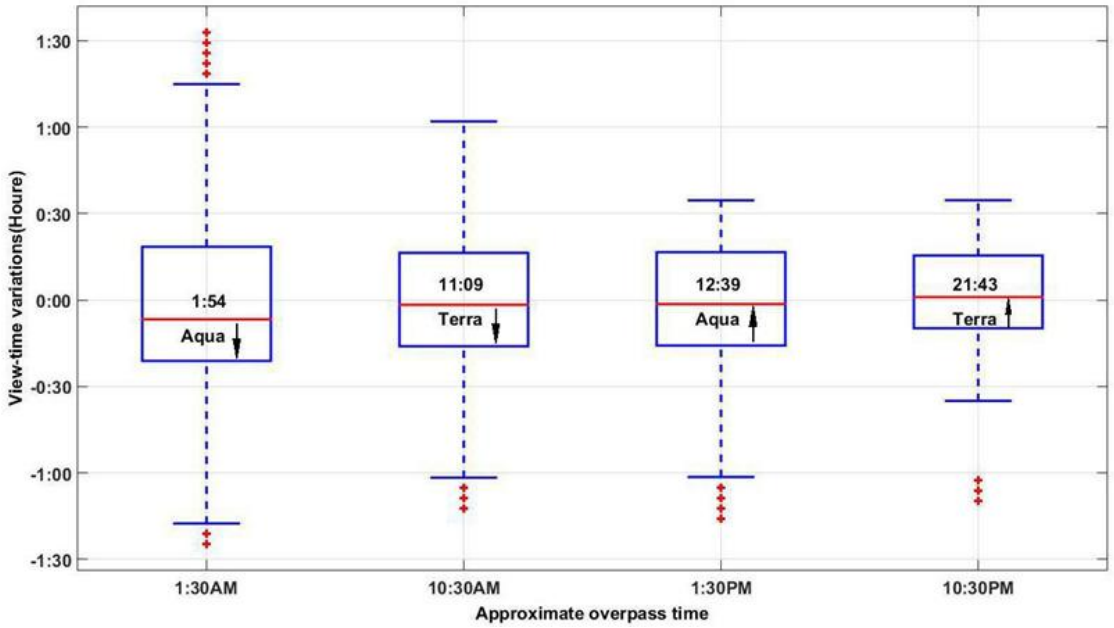


Figure 1: *Average viewing times (local solar) & overpass nodes (shown as labels and arrows), maximum variations from the mean observation times (in hour shown by lower and upper caps of whiskers), median times (middle line), lower (25th) and upper (75th) quartiles of all observation times (lower and upper edges of boxes) of four overpasses of MODIS (onboard Terra and Aqua, two overpasses each) over the study area of 7 years (2007–2013). The mean local solar observation time of each overpass is subtracted from the series (scaled to 0) but is labeled on each box*

2.4.1. Temporal Matching of T_{air} to LST Observations

The temporal frequency of the MODIS LST L3 product is four observations per day (Terra passes over the equator at approximately 10:30 am, 10:30 pm each day, Aqua satellite passes over the equator at approximately 1:30 pm and 1:30 am) in cloud free conditions, which are derived from a composite of several MODIS overpasses with different view angles (Wan 1999; Zhu et al., 2013). Depending on the local longitude (which results in changes in the sensor's viewing angle) and latitude, the local solar observation times at each pixel can vary up to 120 minutes or more over a repeating cycle (16 days) of the MODIS twin sensors (Figure 1). Other than that, overpass times do not follow a regular period during the day and over the sensor's repeat cycle. On the other hand, T_{air} data from the weather stations are provided at an hourly and by the minute frequency in Berlin in standard time (MEZ and UTC). This complicates the matching of the MODIS observation times with T_{air} time-series. To overcome this issue, for those stations that are in minute temporal resolution, we only need to convert from MEZ to UTC, but for other stations that are in hourly temporal resolution, a linear equation was considered for the synchronizing of T_{air} with LST form HDF file. For the creation of a data set for use during the day and night time for each LC/LU separately, we need to consider the overpass time of MODIS over Berlin. Another point is that the data has a significant number of missing points due to clouds in our study area, therefore in order to increase the maximum of usable observations, as did Alcantara et al. (2013) in his research, the terra and aqua data were considered. As shown in Figure 1, the LST day from terra and aqua in descending and ascending orbit were considered as a day time series, respectively and only terra in ascending orbit was considered as night time series because of higher correlation which was observed between LST and T_{air} in this time.

2.4.2. Day/Night analysis

Apart from spatial variations, the observation time can affect the relationship between LST and T_{air} time-series. To identify any variability in LST and T_{air} relationship in a diurnal basis, time-series of both variables were separated, based on the MODIS overpass times to produce four series over a single pixel window from MODIS-Terra and MODIS-Aqua day and night overpasses (four in total) were used in this analysis.

2.4.3. Statistical methods

A simple linear relationship is often assumed between LST and T_{air} in literature (Brunel, 1989; Mostovoy et al., 2006). In view of this, a univariate linear regression analysis with the MODIS LST as the independent (or explanatory) and T_{air} as the dependent (or response) variable was applied to analyze LST and T_{air} relationship. The correlation coefficient, r, is reported as a quantitative measure to evaluate the strength of the agreement between LST and T_{air} time series in different steps of the analysis. Significance levels (*p-values*) are reported in the results to express how unlikely the given *r* values would occur if no relationship between the explanatory and response variables did exist. The smaller the *p-level* is, the more significant the relationship. Moreover, two other statistical measurements such as RMSD and Normalized mean Bias (Bias) were considered as following:

$$Normalized\ mean\ bias = \frac{\sum_{i=1}^{n}(M-O)}{\sum_{i=1}^{n} O} \quad \ldots (2)$$

Where n is the number of data, M is LST value and O is the temperature.

3. Theory and Methodology

In this section, a brief overview of a SVR, Simulated annealing (SA), ANN and ANFIS will be discussed.

3.1. Support Vector Regression

Support vector machine (SVM) is a very promising artificial intelligence method applied extensively for solving the classification problems. Support Vector Regression (SVR) method is derived from the SVM, which is a powerful technique to solve a nonlinear regression problem, but it has received less attention, due to the fact that SVR algorithm is sensitive to users' defined free parameters. The involved hyper parameters of the SVR model consist of penalty parameter, insensitive loss function parameter, and the parameter for kernel function. Inappropriate parameters in SVR can lead to over fitting or under fitting problems.How to properly use the hyper parameters is a major task, which has a significant impact on the optimal generalization performance and the SVR regression accuracy (Schölkopf and Smola 1998). Recently, a number of new algorithms like genetic algorithm, grid search optimizing, cross-validation and particle swarm optimization (PSO) have been proposed for the optimization of the SVR parameters (Sartakhti et al., 2011; Ustün et al., 2005; Wang et al., 2016; Chen and Wang, 2007; Hu et al., 2010; Keerthi, 2002; Ito and Nakano, 2005). In this work, the SA algorithm was applied for tuning the parameters of SVR.

3.1.1. Brief Overview of SVR

In this section, the basic SVR concept is concisely described; for detailed description, please see (Cristianini and Taylor, 2000; Smola and Schölkopf, 2004; Ito and Nakano, 2005; Keerthi, 2002). Suppose a given training data of elements $\{(x_i, y_i), = 1, 2, \ldots N\}$, where x_i denotes the ith element in n-dimensional space; that is $x_i = \{x_1 i , \ldots x_n i\} \in R^n$, and $y_i \in R$ is the output value corresponding to x_i. According to mathematical notation, the SVR algorithm builds the linear regression function as follows:

$$f(x, \omega) = (\omega \cdot \varphi(x) + b), \quad : R^n \rightarrow F, \omega \in F, \qquad \ldots (3)$$

Where w and b are the slope and offset coefficients and x denotes the high-dimensional feature space, which is nonlinearly mapped from the input space x. The previous regression problem is equivalent to minimizing the following convex optimization problem shown in equation (4):

$$Min \ \frac{1}{2}\|\omega\|^2$$

Subject to

$$y_i - (x_i) - b \leq \varepsilon$$
$$\omega\varphi(x_i) + b - y_i \leq \varepsilon. \qquad \ldots (4)$$

In this equation, an implicit assumption is that a function f essentially approximates all pairs (xi, yi) with Ɛ precision, but sometimes this may not be the case. Therefore, by introducing two additional positive slack variables ξi and ξ_i^*, the minimization is reformulated as the following constrained optimization problem shown in equation 5:

$$\min\ R(\omega,\xi,\xi^*) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$$
$$y_i - (\omega,\varphi(x)) - b \le \varepsilon + \xi_i^*$$
$$(\omega,\varphi(x)) + b - y_i \le \varepsilon + \xi_i$$
$$\xi_i,\xi_i^* \ge 0, i = 1,2,\dots,N,\ \varepsilon \ge 0, \quad \dots (5)$$

Where the parameter c is the regulator, which is determined by the user and it influences a tradeoff between an approximation error and the weights vector norm and slack variables that represent the distance from actual values to the corresponding boundary values of tube. According to the strategy outlined by Scholkopf and Smola (1998), by applying Lagrangian theory and the KKT condition, the constrained optimization problem can be further restated as the following equation: was applied in the study, which has the ability to universally approximate any distribution in the feature space. With an appropriate parameter, the radial basis function (RBF) usually provides a better prediction performance, so it is adopted in this study as shown in the following formula:

$$f(x,\omega) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\,K(x_i,x) + b$$
$$s.t.\ \ \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0. \quad \dots (6)$$

Here αi and αi* are the Lagrange multipliers. The term (xi, x) is defined as the kernel function. The nonlinear separable cases could be easily transformed to linear cases by mapping the original variable into a new feature space of high dimension using (xi, x). The RBF was applied in the study, which has the ability to universally approximate any distribution in the feature space. With an appropriate parameter, RBF usually provides a better prediction performance, so it is adopted in this study as shown in equation 7:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad \dots (7)$$

where xi and xj are input vector spaces and σ^2 is the bandwidth of the kernel function. In the above equations, there exist three hyper-parameters to be determined in advance, that is, the penalty parameter C, insensitive parameter Ɛ, and the related kernel function parameters σ^2. They heavily affect the regression accuracy and computation complexity of SVR. The penalty parameter C controls the degree of punishing the samples whose errors go beyond the given value. The insensitive parameter Ɛ controls the width of the Ɛ-insensitive zone used to fit the training data. The value of Ɛ can enhance the generalization capability; with the increase of Ɛ, the number of support vectors will decrease, and the algorithmic computation complexity will also reduce. The bandwidth of σ the kernel function has a great influence on the performance of the learning machine. In this study, one optimization method, that is, simulated annealing (SA), is presented to determine the optimal hyper parameters of the SVR model. According to research of Ustün and Melssen (2005), the general range of C, σ^2, and Ɛ has been given. In the trial operation, we narrowed it to avoid blindness in the optimization process. In this study, the set of hyper parameter (C, σ^2, Ɛ) is initialized in the given range $C \in [0, 1000]$, $\sigma^2 \in [0, 2]$, and $Ɛ \in [0, 0.0001]$, where optimization method (SA) is to seek the global optimal solutions.

3.1.2. Simulated Annealing Optimization Method

Simulated annealing is a local search algorithm capable of escaping from local optima. Its ease of implementation and convergence properties and its use of hill climbing moves to escape local optima have made it a popular technique over the past two decades. Survey articles that provide a good overview of simulated annealing's theoretical development and domains of application include (Eglese, 1990; Fleischer, 1995; Henderson et al., 2003; Koulamas et al., 1994; Romeo et al., 1991; Anily and Federgruen, 1987; Suman and Kumar, 2006; Abramson et al., 1999; Ben-Ameur, 2004; Aarts and Korst, 1989; van Laarhoven and Aarts, 1988; Aarts and Lenstra, 1997). This study proposed an SA-based approach for parameter tuning in the SVR. For convenience, the SVR model with SA is referred to as a SA-SVR method. The idea, is to find the parameters that minimize the generalization error of the algorithm at hand. This error can be estimated on some data which has not been used for learning. To achieve this aim, the three basic decision variables as mentioned before must be tuned in proper manner. We propose here a methodology for automatically tuning multiple parameters for the SVR. The process of SA-SVR algorithm approach is briefly summarized as follows:

Algorithm: Simulated annealing algorithm.

Step 1: Solution space X
Object function F
Neighborhood structure N

Step 2: Current =An initial solution, among all possible state (X)
$S_{optimal}$=Current
T0=INFINITY
T=T0
Iteration=MAX_Iter
Epoch=1
Select temperature reduction function alpha, 0.8≤ alpha ≤0.99

Step 3: Repeat
Next= randomly selected from N (Current)
ΔF=F [Next] – F [Current]
If ΔF> 0:
Current=Next
else
r=rand (0, 1) % Generate a random number r ϵ (0, 1)
if $r < e^{(-\Delta f/T)}$:
Current=Next
Until Epoch <= Iteration
T=alpha*T
If F [Current] < F [$S_{optimal}$] :
$S_{optimal}$ =Current
Until stop condition is met

Step 4: Return $S_{optimal}$ as an approximation to the global minimum solution

The proposed parameter values of SA-SVR approach were set as follows: Iteration = 200, T_0 was set to a sufficiently large number, while the set of hyper parameters (C, σ^2, Ɛ) is initialized in the given range C ∈ [0, 10000], σ^2 ∈ [0, 2], and Ɛ ∈ [0, 0.0001], where optimization method (SA) is to seek the global optimal solutions. The best solution among these possible solutions is then selected as the optimal solution in the SA-SVR.

3.2. An Adaptive Neuro-Fuzzy Inference System

An adaptive neuro-fuzzy inference system or adaptive network-based fuzzy inference system (ANFIS) is a kind of artificial neural network that is based on Takagi–Sugeno fuzzy inference system (Sugeno and Tanaka, 1992; Takagi and Sugeno, 1985). The technique was developed in the early 1990s (Jang and Shing (1991, 1993)). Since it integrates both neural networks and fuzzy logic principles, it has a potential to capture the benefits of both in a single framework. Its inference system corresponds to a set of fuzzy IF–THEN rules that have learning capability to approximate nonlinear functions (Abraham, 2005). Hence, ANFIS is considered to be a universal estimator (Jang, Sun and Mizutani, 1997).

3.3. Neural Network

Artificial neural network models are universal approximations with the ability to generalize through learning non-linear relationships between provided variables of input(s) and output(s) (Hájek and Olej 2012). ANN are organized and interconnected collections of processing units (neurons or nodes), whose operation is analogue to a neural structure (Müller and Fill, 2003). ANN extract its computational power from its solid parallel distribution structure and ability to learn/generalize, allowing the resolution of complex propositions in many known areas (Haykin 2001). ANN execution is inspired on the human brain (Haykin, 2001) and has been used in many applications with success. In agreement with Galvão et al. (1999), by the reason of its nonlinear structure, the ANN can acquire more complex data characteristics, which are not always possible using traditional statistical techniques (Maier et al., 2010; Razavi and Tolson, 2011). ANN is a robust computational technique which is primarily used for pattern recognition, classification, and prediction (Bose and Liang, 1996; Haykin, 1999; Panchal et al., 2011). The use of ANNs in meteorological applications includes the prediction of ozone concentration, sulfur dioxide concentration, tornadoes, storms, solar radiation, carbon dioxide, pollutants, and monsoon rainfall (Gardner and Dorling, 1998), monthly and year precipitation levels (Bodri and Cermak, 2000), tide charts (Steidley et al. 2005), wave heights (Wedge et al. 2005), flash floods (Luk et al., 2000), and air temperature (Jain et al., 2003; Smith et al., 2006; Maqsood et al. 2004), estimation of dew point temperature (Mittal and Zhang, 2003; Shank et al., 2008). Bilgili and Sahin (2010) used ANN for predicting long-term monthly temperature and rainfall in Turkey. Kisi and Shiri (2011) introduced new hybrid wavelet-AI models for precipitation forecasting. Smith et al. (2005) developed an enhanced ANN for air temperature prediction by including information on seasonality and modifying parameters of an existing ANN model.

3.3.1. Determining Hidden Node

Many researchers put their best effort in analyzing the solution to the problem that how many neurons are kept in hidden layers in order to get the best results (Rivals I. and Personnaz L. 2000; F. Fnaiech.et al. 2001; Kortmann-Unbehauen 1988; Onoda 1995; Md. Islam and Murase (2001); Stuti Asthana and Rakesh K Bhujade (2011); Kazuhiro Shinike 2010; Doukim et al. 2010; Yuan et al. 2003; Wu and Hong 2010; Panchal et al. 2011; Hunter et al. 2012; Shuxiang et al. 2008; Ke and Liu (2008)), but unfortunately no one succeeded in finding the optimal formula for calculating the number of neurons that the neural network training time can be reduced and also accuracy in determining the target output can be increased. Usually some rule-of-thumb methods are used for determining the number of neurons in the hidden nodes.

1. The number of hidden layer neurons are 2/3 (or 70% to 90%) of the size of the input layered. If this is insufficient then the number of output layer neurons can be added later on (Boger and Guterman, 1997).

2. The number of hidden layer neurons should be less than twice of the number of neurons in input layer (Berry and Linoff,1997).
3. The size of the hidden layer neurons is between the input layer size and the output layer size (Blum, 1992).

But the above three methods are not considered to be always true because not only the input layer and the output layer decide the size of the hidden layer neurons, but also the complexity of the activation function applied on the neurons, the neural network architecture, the training algorithm, and most important the training samples of the database on which the neural network is designed to execute. In this work, we decided to use the cross validating approach in the 3-layers MLP in the following simulations, in order to select the number of hidden nodes in the second layer. The 3-layer MLP contains an input layer, one hidden layer with nonlinear transfer functions and an output layer with linear transfer functions. The training algorithm is Back Propagation (BP) in order to get the configuration that minimizes the RMSE in the test phase while keeping an eye on over fitting and the train set error.

3.3.2. Assess Predictive Performance of Models

In a real application, cross-validation is a model assessment technique (Allen, 1974; Stone, 1974; Geisser, 1975) used to evaluate a machine learning algorithm's performance in making predictions on new datasets which has not been trained on. This is done by partitioning a dataset and using a subset to train the algorithm and the remaining data for testing. Because cross-validation does not use all of the data to build a model, it is a commonly used method to prevent over fitting during training. Each round of cross-validation involves, randomly partitioning the original dataset into a training set and a testing set. The training set is then used to train a supervised learning algorithm and the testing set is used to evaluate its performance. This process is repeated several times and the average cross-validation error is used as a performance indicator (Hastie et al., 2009; Yang, 2007b). Common CV techniques include, k-fold, Holdout, Leave out, repeated random sub-sampling, Stratify, Substituting. In this work, we apply K-fold CV (with k=4) techniques, in order to test how well our model is able to be trained by some data and then to estimate the data it hasn't seen before and then to select the best model.

3.4. Data Normalization

Before computing, data of both input and output variables were normalized. In this study, data of all variables used were normalized into the range [0, 1] with:

$$X_{norm} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad \ldots(8)$$

where X_{norm} is the normalized value, Xi is the original value, and X_{min} and X_{max} are the minimum and maximum values out of the sample of Xi. This was due to the eliminating influence of different dimensions of data and to the avoidance of overflows of the model during calculations, as a result of very large or small weights towards a maximization of model parsimony with considering computational effort. After the computation, output values were transformed back to the real prediction data.

3.5. Model Calibration and Validation

Cross-validation was used to evaluate the generalizability of a model for estimating the air temperature with the LST data. The observations were randomly divided into two parts. 70% of the observations were used for model calibration, and the rest were used as test dataset for model validation. The accuracy of the estimated air temperature obtained from three estimating models, ANFIS, NN and SA-SVR, have been assessed by a set of statistic measures, including: Root Mean-square Error (RMSE), coefficient of determination R-squared (R2), Mean Bias Error (MBE) and Mean Absolute Error (MAE), respectively. The RMSE (was mainly used in the development process of the model and represents residual errors, which gives a global perspective of the differences between the observed and estimated values (Sousa et al., 2007; Zheng et al., 2013; Willmott et al., 2005). The RMSD is calculated similarly to RMSE. These goodness of fit criteria are expressed as equations (9-12):

$$R = \frac{\sum_{i=1}^{M}(O_i - \bar{O})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{M}(O_i - \bar{O})^2} \times \sqrt{\sum_{i=1}^{M}(S_i - \bar{S})^2}} \quad (9)$$

$$MBE = \frac{1}{M}\sum_{i=1}^{M}(O_i - S_i) \quad (10)$$

$$MAE = \frac{1}{M}\sum_{i=1}^{M}|O_i - S_i| \quad (11)$$

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(O_i - S_i)^2} \quad (12)$$

where, M is the total number of the observation data, O and S are the average of the observed and estimated T_{2m}, and O_i and S_i are the observed and estimated T_{2m} of the i^{th} data, respectively. In addition, graphical goodness-of-fit criteria such as quantile-quantile (Q-Q) diagram, bar plot of RMSE in train and test phases were applied for the comprehensive evaluation of simulation results. Although, the R^2 criteria is a measure of goodness-of-fit of the model and higher values are indicative that the predictive model fits the data in a better way. By definition, R^2 is the proportional measure of variance of one variable that can be predicted from the other variable. Thus, ideally the values of R^2 to approach one is always desirable. However, a high R^2 tells you that the curve came very close to the points, but in reality, it does not always indicate the model quality (Maddala, 2001). In order to have a reliable statistical comparison between the models, both the MAE and RMSE can be used together to ascertain the variation in errors in a given set of estimation. It should be noted that in MAE, all the individual errors have equal weight on the average, making it a linear score, but the RMSE has a quadratic error rule, where the errors are squared before being averaged. As a result, a relatively high weight is given to large errors. This could be useful when large errors are undesirable in a statistical model (Chai and Draxler, 2014; Armstrong, 2002).

4. Theoretical Concepts for Selecting Input Parameters

Spatial and temporal variation in temperature are governed by physical processes. For example, land surface temperature at some 'locations' in space and time ($s0, t_0|s \in S, t \in T$) is a function of incoming solar radiation, cooling factor by wind, land cover, temperature inversion and other effects. The temperature patterns differ between day and nighttime also; during the night temperature patterns are

mainly determined by land cover, air humidity and proximity to water bodies and/or soil moisture (van Leeuwen et al., 2011). In urban and industrial areas, temperature is often locally somewhat higher due to heat emissions from industrial activities or heating (e.g. Cheval and Dumitrescu 2009). Moreover, near-surface air temperature is driven more by land surface temperature than by direct solar radiation (Zakšek and Schroedter-Homscheidt 2009), making LST an important variable for estimating T_{air}.

Other parameters, such as vegetation cover, soil moisture, solar radiation, and albedo also have some influence on air temperature. In previous studies, several variables were employed to estimate air temperature. For example, the variables used by Benali et al. (2012) included LST, Julian Day, elevation, and distance to coast. The variables used by Kim and Han (2013) included LST, NDVI, altitude, and solar zenith angle. The variables used by Cristóbal, Ninyerola, and Pons (2008) included LST, NDVI, and albedo. The variables used by Zakšek and Schroedter-Homscheidt (2009) included LST, NDVI, solar zenith, albedo, solar radiation, and altitude. After comprehensive consideration of these variables, twelve variables were selected as the predictors for modelling air temperature: LST, NDVI, latitude, longitude, altitude, albedo, wind speed, wind direction, emissivity31, emissivity32, relative humidity and Julian day. The reasons for selecting these variables as input to our model for estimating the air temperature are summarized as follows:

1 The latitude, longitude and elevation were selected as an input parameter to model because the incoming solar radiation can be globally derived as a function of this factors. Moreover, latitude, longitudes and elevation are always the underlying effect relative to temperature (Zhao and Cheng, 2005; Samanta et al., 2012; Stahl et al., 2006).

2 Emissivity is important, because all objects at temperatures above absolute zero emit thermal radiation. However, for any particular wavelength and temperature the amount of thermal radiation emitted depends on the emissivity of the object's surface. Emissivity is defined as the ratio of the energy radiated from a material's surface to that radiated from a blackbody (a perfect emitter) at the same temperature and wavelength and under the same viewing conditions. The emissivity of a surface depends not only on the material but also on the nature of the surface. The emissivity also depends on the temperature of the surface as well as wavelength and angle. Knowledge of surface emissivity is important both for accurate non-contact temperature measurement and for heat transfer calculations. Moreover, Surface emissivity is a measure of inherent efficiency of the surface in converting heat energy into radiant energy above the surface (Sobrino et al., 2001). Therefore, land surface emissivity is critical for determining the thermal radiation of the land surface (Caselles et al., 1995). The emissivity of a surface is controlled by some factors such as water content, chemical composition, structure, roughness, and the observation conditions (i.e. wavelength, pixel resolution and observation angle) (Snyder et al., 1998). For these reasons, in our study, due to considering six different LC/LU, the land surface emissivity also considered as an input parameter.

3 LST is the radiative temperature of the land surface (Ghent et al., 2010). It is influenced by albedo, vegetation cover and soil moisture (Land Surface Temperature Copernicus Global Land Service). The “surface” can include snow and ice, bare soil, grass, or the roofs of buildings (Land Surface Temperature: Global Maps, 2016). Near-surface air temperature “is a measurement of the average kinetic energy of the air near the surface of the Earth” (Near Surface Air Temperature - GES DISC-Goddard Earth Sciences Data and Information Services Center, 2016). Usually LST is measured by remote sensing whereas air temperature is measured 1-2 m above the ground. Near-surface air temperature is a consequence of complex effects of the turbulent heat transports produced by nearby heated surfaces (Unger, et al., 2009). The advantage of using MODIS LST is that, they account for small differences in

temperature that are due to different land cover, moisture content which cannot modeled with constant physical parameters such as elevation, latitude, longitudes.

4 The Julian day is proxies for the fraction of solar energy absorption during the day and emission during the night, influencing the diurnal amplitude of T_{air} throughout the year. The Julian day included the information of vegetation cover changes with seasons. Julian day is the continuous counting of days from 1st January every year.

5 The NDVI and Albedo reflect the seasonal variation of land cover.

6 The Relative humidity (RH) is the ratio of the partial pressure of water vapor to the equilibrium vapor pressure of water at a given temperature. Relative humidity depends on temperature and the pressure of the system of interest. It requires less water vapor to attain high relative humidity at low temperatures; more water vapor is required to attain high relative humidity in warm or hot air (Perry and Green, 2007).

7 Moreover, Seasonal variation in some parameters such as relative humidity, wind speed, wind direction and air pressure contribute to explaining seasonal variation air temperature over six LC/LU.

8 The MODIS LST can be used to improve spatial prediction of ground-measured values.

5. Results and Discussion

5.1. MODIS LST versus T_{air} time-series over a single pixel

Before analyzing the effects of MODIS window size, the daily variability of LST and T_{air} relationship was examined by using separate LST series (over 1x1 window) (Diurnal differences). In this section, LST series used in this analysis is a composite time series which includes four daily LST observations (except for cloudy days) from both the MODIS Terra and Aqua day and night overpasses (approximately at 1:30, 10:30, 13:30, 22:30) supplied in the LST L3 product.

Table 3: *Statistical analyses between MODIS LST products and T_{air} observation from automatic meteorological stations. MOD_{day}, MOD_{night}, MYD_{day} and MYD_{night} are representative of MOD11A1 LST_{day}, MOD11A1 LST_{night}, MYD11A1 LST_{day} and MYD11A1 LST_{night} from Terra and Aqua respectively for urban and industrial LCT*

Dataset	Urban				Industrial			
	R^2	RMSD	MBE	Bias	R^2	RMSD	MBE	Bias
$^{MOD}day,^{T}day$	0.88	3.72	0.25	0.01	0.86	3.69	-0.58	-0.03
$^{MOD}nigt,^{T}night$	0.87	3.63	-1.89	-0.21	0.80	4.57	-2.65	-0.28
$^{MYD}day,^{T}day$	0.87	4.22	1.57	0.09	0.86	3.81	0.23	0.01
$^{MYD}nig^{T}night$	0.88	2.94	-1.50	-0.21	0.80	4.21	-2.55	-0.35

Table 4: *Statistical analyses between MODIS LST products and T_{air} observation from automatic meteorological stations. MOD_{day}, MOD_{night} , MYD_{day} and MYD_{night} are representative of MOD11A1 LST_{day}, MOD11A1 LST_{night}, MYD11A1 LST_{day} and MYD11A1 LST_{night} from Terra and Aqua respectively for agriculture and needle leaf trees LCT*

Dataset	Agriculture				Needleleaftrees			
	R^2	RMSD	MBE	Bias	R^2	RMSD	MBE	Bias
$^{MOD}day,^{T}day$	0.91	2.89	-1.42	-0.08	0.85	4.21	0.49	0.02

$^{MOD}night, ^{T}night$	0.85	3.23	-0.82	-0.11	0.87	3.61	-1.93	-0.20
$^{MYD}day, ^{T}day$	0.92	3.05	-0.22	-0.01	0.83	4.41	1.11	0.06
$^{MYD}night, ^{T}night$	0.85	2.83	-0.43	-0.09	0.87	3.15	-1.59	-0.21

The comparison between MODIS LST data and the T_{air} observations shows that LST_{day} and LST_{night} from both Terra and Aqua, with the mean relative bias above and under zero tended to overestimate T_{day} and underestimate T_{night} (Table 3-5) respectively as Cresswell et al. (1999) found the same result.

Table 5: *Statistical analyses between MODIS LST products and T_{air} observation from automatic meteorological stations. MOD_{day}, MOD_{night} , MYD_{day} and MYD_{night} are representative of MOD11A1 LST_{day}, MOD11A1 LST_{night}, MYD11A1 LST_{day} and MYD11A1 LST_{night} from Terra and Aqua respectively for airport and forest LCT*

Dataset	Airport				Forest			
	R^2	RMSD	MBE	Bias	R^2	RMSD	MBE	Bias
$^{MOD}day, ^{T}day$	0.90	3.99	1.65	0.10	0.89	4.43	2.46	0.13
$^{MOD}night, ^{T}night$	0.86	4.32	-2.92	-0.31	0.83	3.69	0.73	0.08
$^{MYD}day, ^{T}day$	0.89	4.80	2.72	0.15	0.88	4.82	3.54	0.17
$^{MYD}night, ^{T}night$	0.86	3.71	-2.41	-0.32	0.88	2.85	1.15	0.18

As shown in the table, a higher relative RMSD and bias values were seen for the Aqua $LST_{daytime}$ than the Terra $LST_{daytime}$ which might be given to the fact that more solar radiation has been received at the time of the Aqua MODIS overpass later in the day. Considering the scatterplots of LST_{night} and T_{night} from Aqua for the industrial LC type, has higher scattering than daytime observations which are more spread around the 1:1 line (Figure 5). This is undeniable evidence of the negative impact of urbanization on a surface urban heat island (UHI) and global warming This indicates the urban heat island with RMSD=4.21°C (Nguyen et al., 2015).

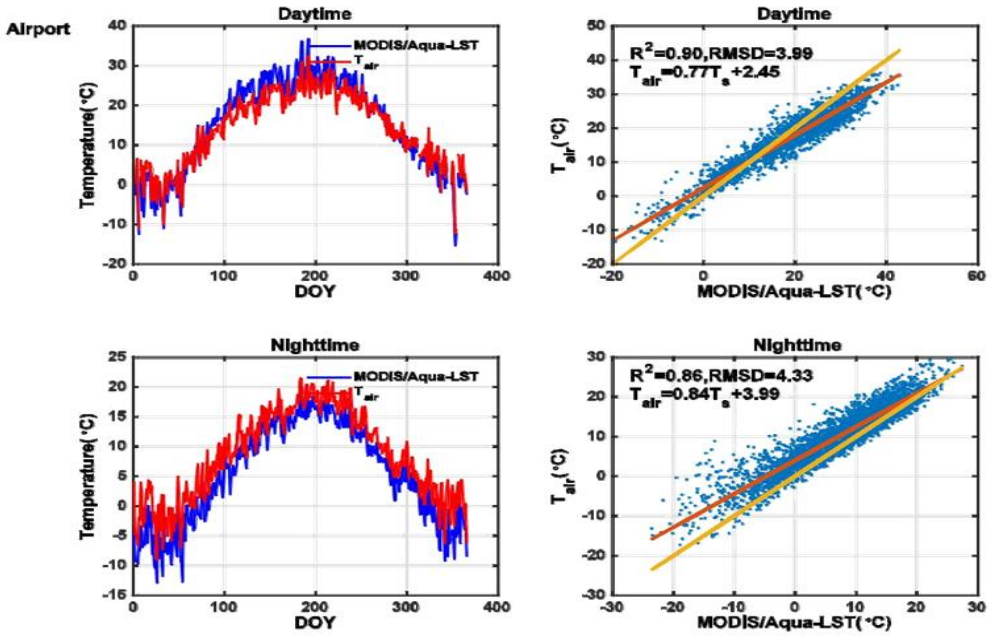
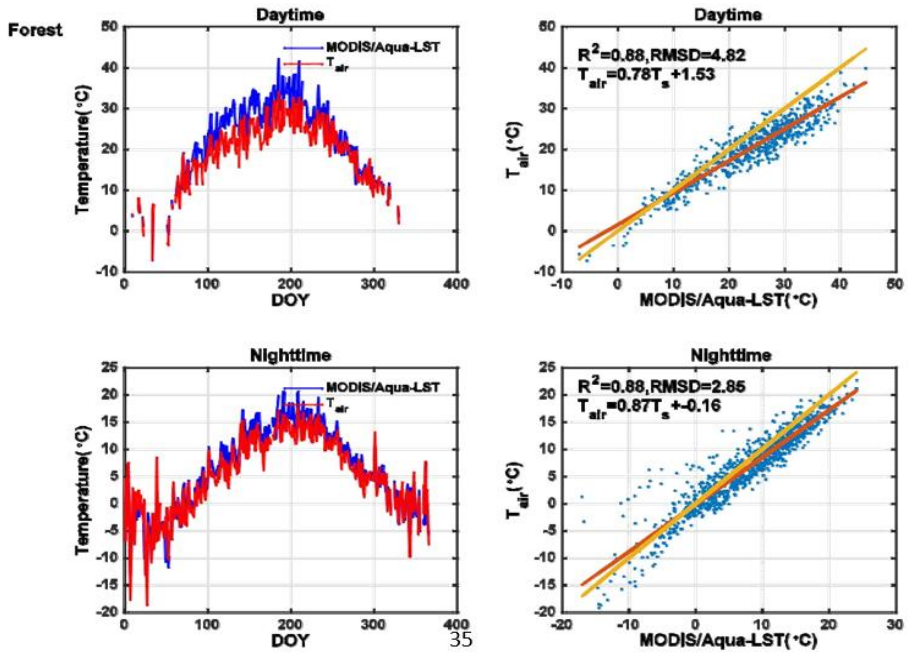


Figure 2: *Correlations between LST and Tair time-series separated based on approximate overpass times of MODIS-Aqua for seven years between 2007 to 2013, where each scatterplot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and nighttime observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for airport LCT with P-value<0.01. DOY means day of year*



35

Figure 3: *Correlations between LST and Tair time-series separated based on approximate overpass times of MODIS-Aqua for seven years between 2007 to 2013, where each scatterplot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and nighttime observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for forest LCT with P-value<0.01. DOY means day of year*

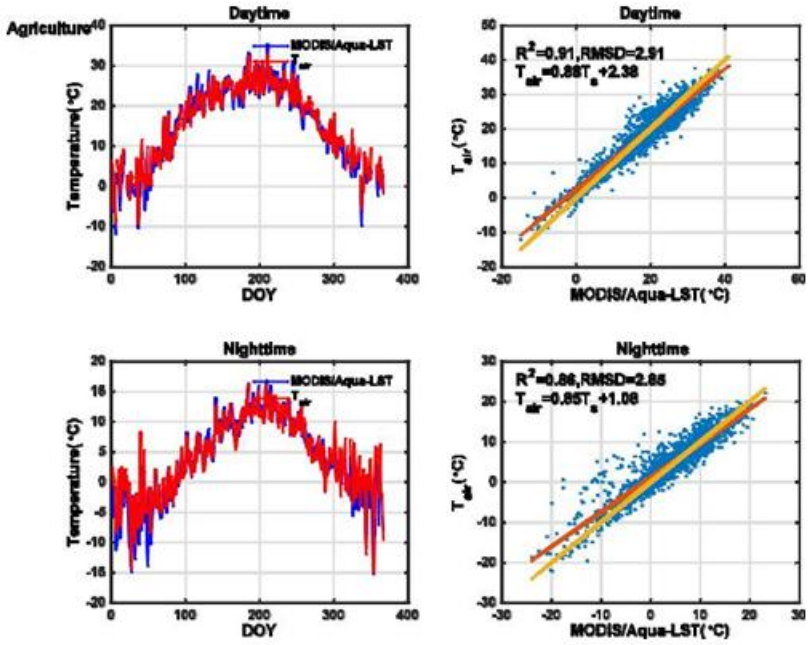


Figure 4: *Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua for seven years between 2007 to 2013, where each scatterplot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and nighttime observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for agriculture LCT with P-value<0.01*

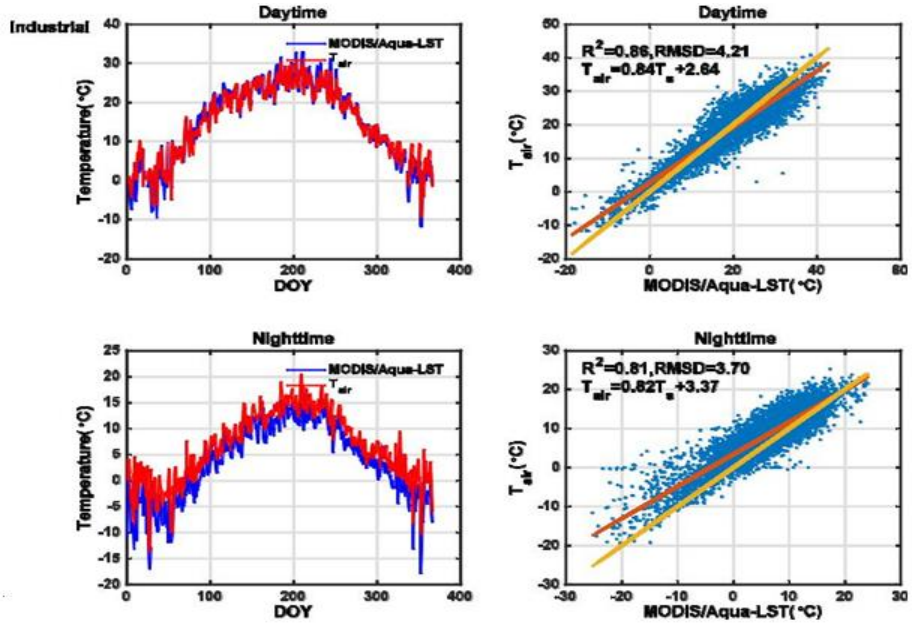


Figure 5: *Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua for seven years between 2007 to 2013, where each scatterplot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and nighttime observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for industrial LCT with P-value <0.01. DOY means day of year*

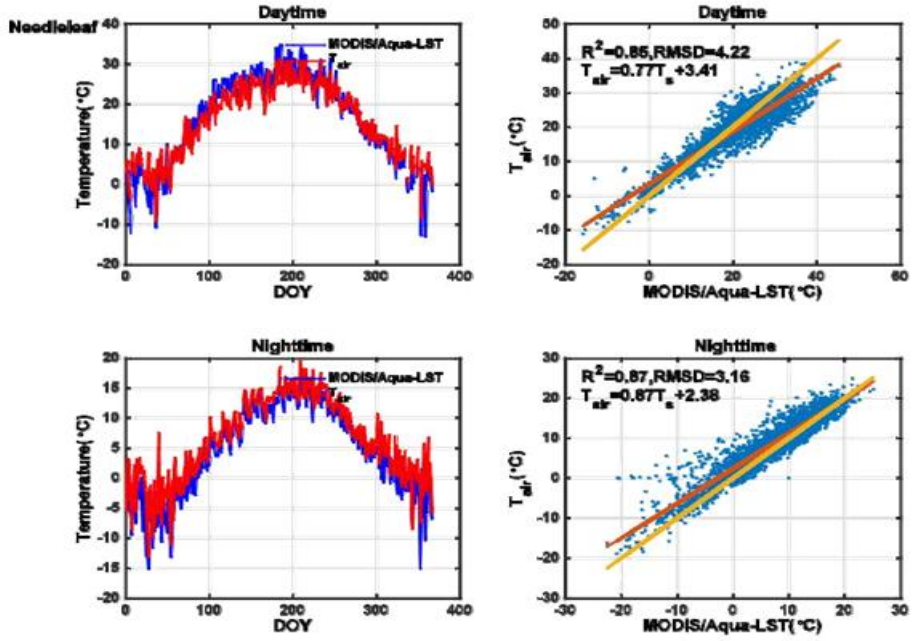


Figure 6: *Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua for seven years between 2007 to 2013, where each scatterplot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and nighttime observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for needle leaf trees LCT with P-value <0.01. DOY means day of year*

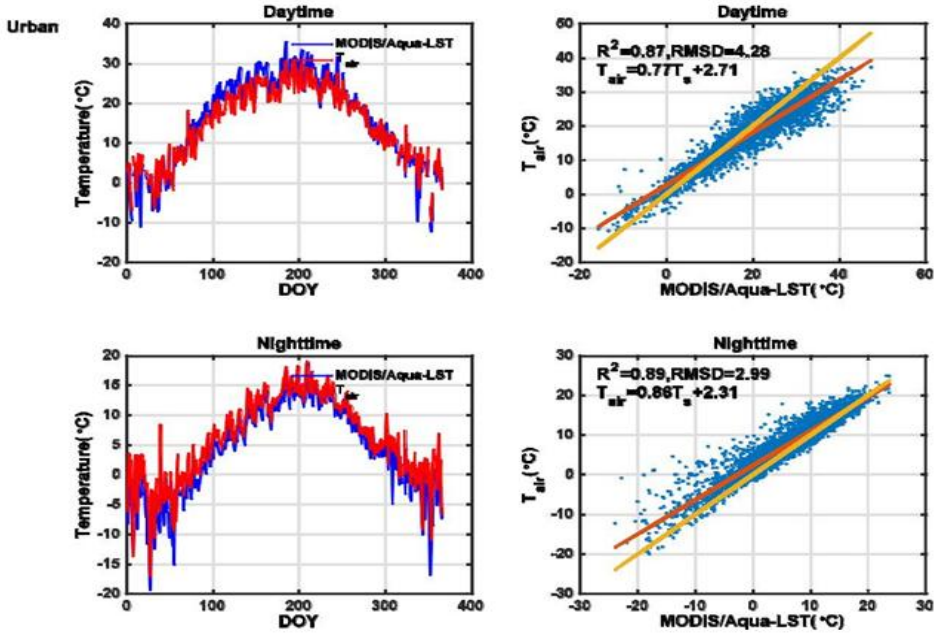


Figure 7: *Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua for seven years between 2007 to 2013, where each scatterplot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and nighttime observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for urban LCT with P-value <0.01. DOY means day of year*

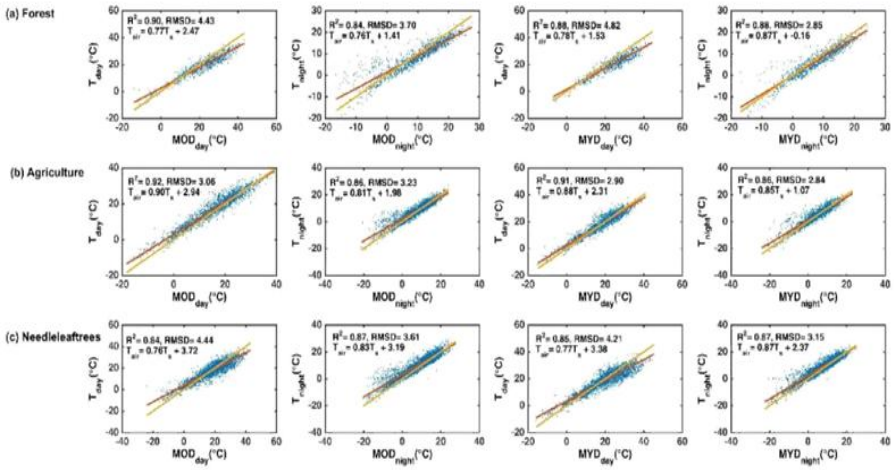


Figure 8: *Scatter plots between observed T_{air} (T_{day} and T_{night}) and LST from Four MODIS products (MOD_{day}, MoD_{night}, MYDday, MYD_{night}) which are represented for three different LC/LU for considered year (2007 to 2013). RMSD is the Root-Mean_Squar Deviation and it's calculated as the same as RMSE. R^2 is adjusted correlation coefficient between T_{air} and T_S. Tsis land surface temperature*

Both Aqua and Terra LST_{night} underestimated the T_{night} as well except for forest. Moreover, according to RMSD from Table 3-5 and MODIS LST from Terra, a higher RMSDs is found for industrial and airport LC types during night time which indicates the UHI phenomena (with RMSD = 4.57°C and 4.32°C respectively). Moreover Table 3-5 show that, correlations between the MODIS LST from Terra data are generally stronger from the daytime series compared with those from the night series, except

for needle leaf trees. The needle leaf tree type showed more complex correlation patterns from day and night observations. The possible reason for this, is that the values of LST recorded by MODIS observation on this particular LC type is not exactly a representative of the skin temperature of the soil, but rather affected by the temperature near the top of the trees (canopy temperature). In addition, LST and T_{air} are correlated to a certain degree, with some drawbacks depending on factors, such as land cover type (Jin et al., 2010; Mildrexler et al., 2011). In general, Figure 2-7 show that the time-series of the MODIS LST over six LC/ LU classes were correlated individually during the day and night time. They are highly correlated with $R^2 > 0.80$. Moreover, Figure 2-7 show that, during the warm months the LST_{day} is higher than T_{day} due to strong radiation, while as expected during the cold months LST_{day} is lower than T_{day} for almost all LC/LU. Moreover, almost for all LC/LU, the LST_{night} is close to T_{night}. As due to long wave, radiation from surface LST and T_{air} at night are closer.

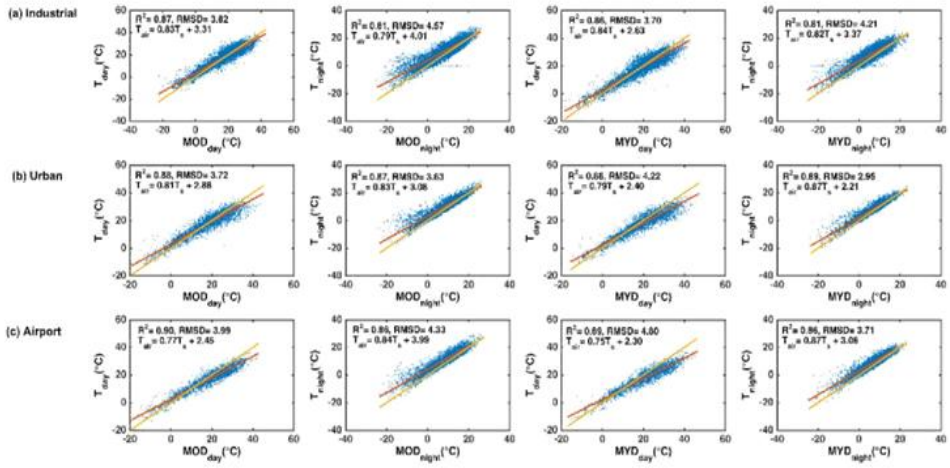


Figure 9: *Scatter plots between observed Tair (T_{day} and T_{night}) and LST from Four MODIS products (MOD_{day}, MoD_{night}, MYD_{day}, MYD_{night}) which are represented for three different LC/LU for considered year (2007 to 2013). RMSD is the Root-Mean_Squar Deviation and it's calculated as the same as RMSE. R2 is adjusted correlation coefficient between T_{air} and TS. Tsis land surface temperature*

Both the Terra and Aqua LST products were compared with the ground-based T_{air} as shown in the Figure 8 and 9, the night time LST datasets (MOD_{night} and MYD_{night}) and the observed T_{air} are more linearly concentrated along the fitting line than the daytime datasets. Strong correlations were observed between the night time LST and T_{night} with minimal bias ($0.81<R^2<0.89$, RMSE< 4.80 and MBE < 2.91 °C). Specifically, the MYD_{night} tends to be more accurate for the estimation of T_{air} with lower intercepts, smaller RMSD and MBE than MOD_{night}. For T_{day}, the MYD_{day} had good agreement than MOD_{day} with lower intercept. This is most likely because the Aqua overpass time (1:30 the time when maximum temperature was recorded). However, LST from Aqua and Terra seems to be best for estimation of T_{day} among the LST products.

To sum up, the relationship between LST and T_{air} may vary with time and location because the land surface energy balance is a complex phenomenon that depends on multiple factors (e.g., cloud cover, surface roughness, wind speed and soil moisture). In addition, the LST and T_{air} are different in principle. The satellite remotely sensed LST is a measure of the surface radiation. LST was calculated from the emissivity's surface, which is sensitive to LC, especially during daytime and another reason is the heat capacity or specific heat of LC. However, the specific heat varies significantly from one LC to another. The variation of the difference between LST_{day} and T_{day} may be due to the different heat capacities or specific heats of LC types (Marzban et al., 2017, Voogt and Oke, 2003). The heat

capacity changes with temperature, which may result in different relations at the different times even over the same LC. Our results showed that the MODIS LST correlates best with T_{air} measurement during the daytime. To some extent, this outcome was contradictory to the other works in the literature (e.g., Zhang et al., 2011a; Benali et al., 2012, Marzban et al., 2017) where they have reported a stronger correlation at night time compared to daytime. It must be noted, though, that Benali et al. (2012) used MODIS-Terra but not MODIS-Aqua observations. Variations in the MODIS, overpasses time in its 16-day repeated cycle which enabled us to reconstruct the diurnal LST profile over a 7 years period. Although, many studies have shown a higher agreement between LST and T_{air} at night (Zhang et al., 2011a; Benali et al., 2012), this is not the case for all hours of the day or night. During some hours of the night the LST - T_{air} relationship is weaker than some hours during the day. These differences could be understood, as not only being the time of observation, but also geographical location affecting the relationship between LST product and T_{air} and therefore, affecting the accuracy estimation of T_{air} based on LST products.

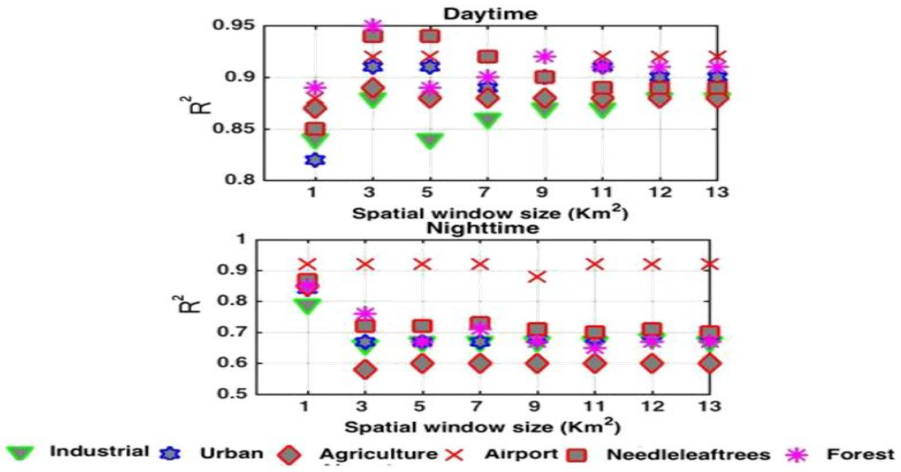


Figure 10: *Variation of the correlation coefficient between T_{day}, T_{night} and LST with the varying spatial window size over six LCT*

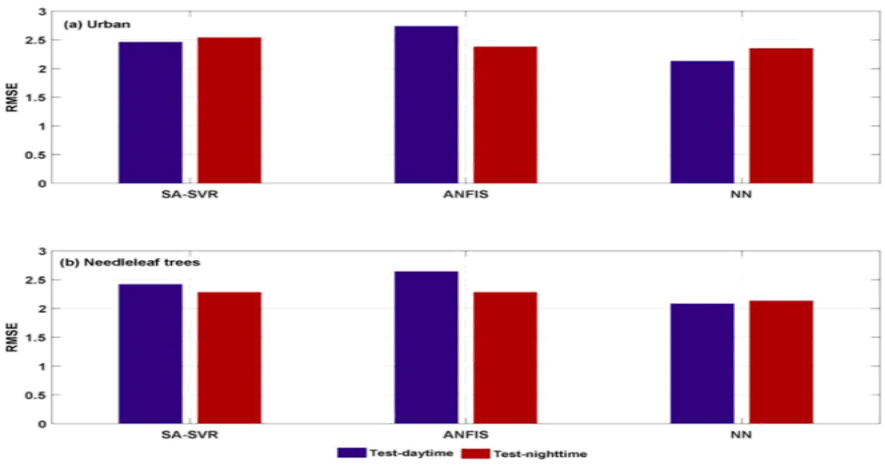


Figure 11: *Bar plot of estimated T2m versus measured temperature during day and nighttime in test phases using SA-SVR, ANFIS and NN for (a) urban and (b) Needle leaf trees LCT*

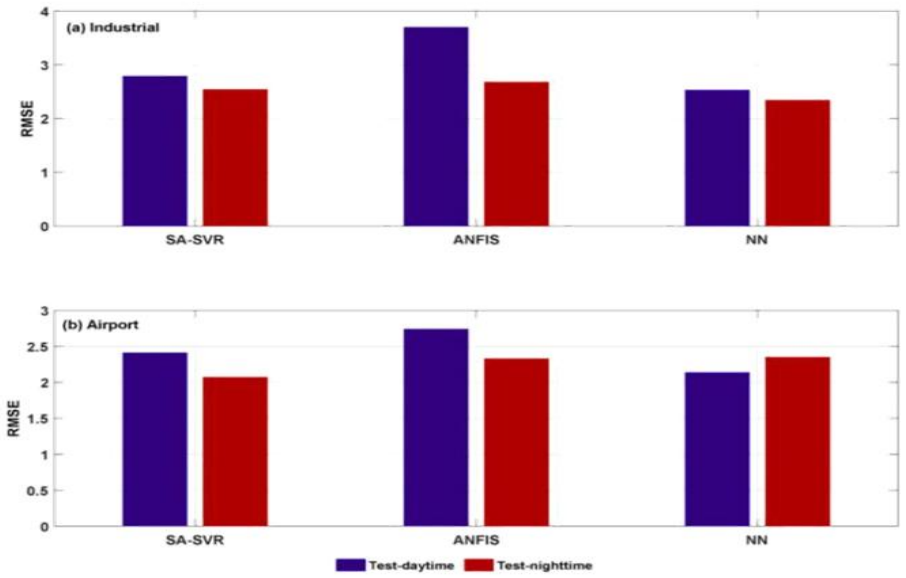


Figure 12: *Bar plot of estimated T2m versus measured temperature during day and nighttime in test phases using SA-SVR,ANFIS and NN for (a) Industrial and (b) Airport LCT*

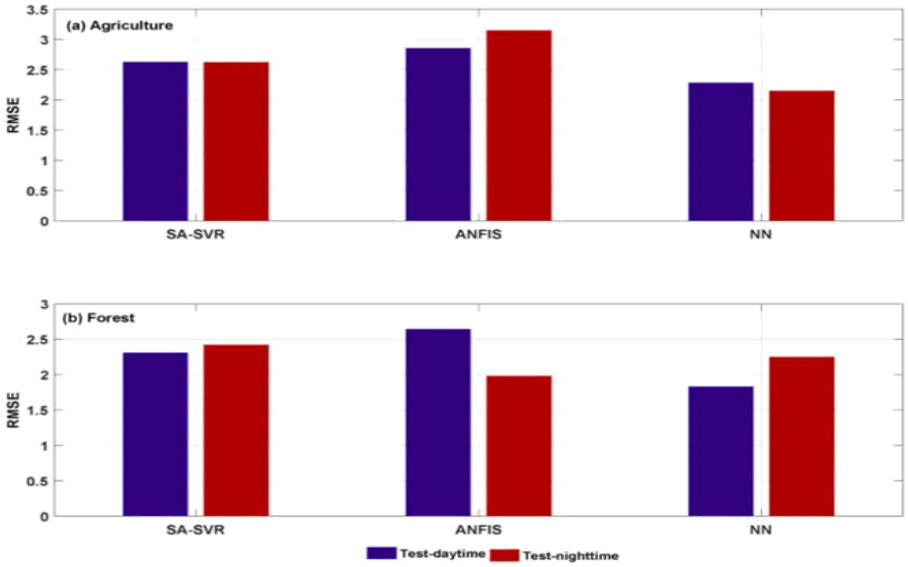


Figure 13: *Bar plot of estimated T2m versus measured temperature during day and nighttime in test phases using SA-SVR, ANFIS and NN for (a) Agriculture and (b) Forest LCT*

5.2. Multiple LST window size

The relationship between the observed $T_{night} - T_{day}$ and LST is not limited to a single pixel, because the temperature of the near-surface air mass in a given area, is influenced by many factors such as energy exchanges with the land surface over a larger area. On the other hand, the T_{air} is impressed by both the local radiation budget and air advection from the surrounding areas, thus, for better understanding of the spatial variability in LST- T_{air} relationship, a spatial window with a varying size is examined to discover the optimal spatial extent over which LST agrees best with the T_{air}

measurements. In order to describe the effects of LST window-size on the LST and T_{air} relationship better, firstly, the time-series of LST from a single pixel (1x1 window size) overlapping each weather station were retrieved from the MODIS LST grid and then the LST of 3x3, 5x5, 7x7, 9x9, 11x11, 13x13 and 15x15 pixels were generated, respectively. Secondly, in order to determine the proper spatial window size for estimating air temperature, correlation coefficient analysis was made for different LC/LU. As shown in figure 8, the correlations were improved very slightly when the window size was increased from 1 to 3 pixels for daytime. The highest correlation values were achieved with 3x3 window for all LC/LU during the daytime and at the 1x1 during the night time. Significance levels of all correlations were found to be at which can be interpreted from p-values (all *p-values* were <0.01). According to these results, the window size was selected for all LC/LU prior to model development for day and night time data set.

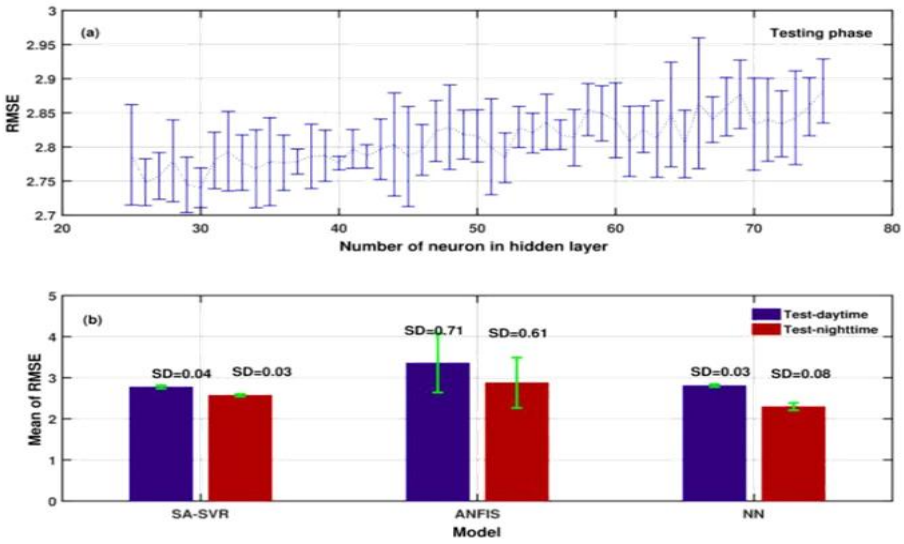


Figure 14: *This two subplots show the effect of K-fold cross validation (with k=4) in three models. In subplot (a) x and y-axes show the average of cross validation error (RMSE) and number of nodes in hidden layer in testing phase respectively. In subplot (b), x and y-axes show the type of model and the average of cross validation error (RMSE) in three models respectively*

5.3. Discussion

Three different methods namely SA_SVR, ANN and ANFIS were employed to estimate T_{air} during the day and night time in Berlin by using the twelve variables as predictors. The performance of the three models was assessed using cross-validation with k=4 fold over different LC/LU during the day and night time. All samples from each LC/LU were used in turn as the validation data set to test the model, while the remaining samples were used as the training data set to fit the model. RMSE, R^2, MBE and MAE were calculated from the measured and estimated T_{air} values to assess model performance. As shown in Table 6-11, ANN model with three layers structure, has higher adjusted R^2 value ranged from 0.93 to 0.97, RMSE ranged from 1.83 °C to 2.53 °C and MAE ranged from 1.53 °C to 1.94 °C in test phases for all LC/LU for estimating T_{day}. The results showed that all models have similar capability in the training phase for estimating T_{night} but the ANN has a higher adjusted R^2 which ranged from 0.89 to 0.93, RMSE and ranged from 2.13 °C to 2.35°C and also MAE ranged from 1.54 °C to 1.84 °C values in the test phase in comparison to ANFIS and SA-SVR. The bar plots of RMSE for the three methods on testing data for each LC/LU are shown in Figure 11-13, respectively. As shown in Table 6-11 and Figure 11-13, the three models SA-SVR, ANFIS and ANN have satisfactory been able to capture the

relationship between the process variables. The bar plots depicted the performance of the NN model on the testing data which was better than those of ANFIS and SA-SVR models for the whole of LC/LU during the day and night time, but again we applied a CV approach to assess the models performance in the test phase for the three mentioned models. As can be seen from Figure 14 (b), the SA-SVR and NN models are more robust and stable than ANFIS model regarding their SD values (ranged from 0.03 to 0.08) during the day and night time and we can say that, these two models are more reliable than the ANFIS model.

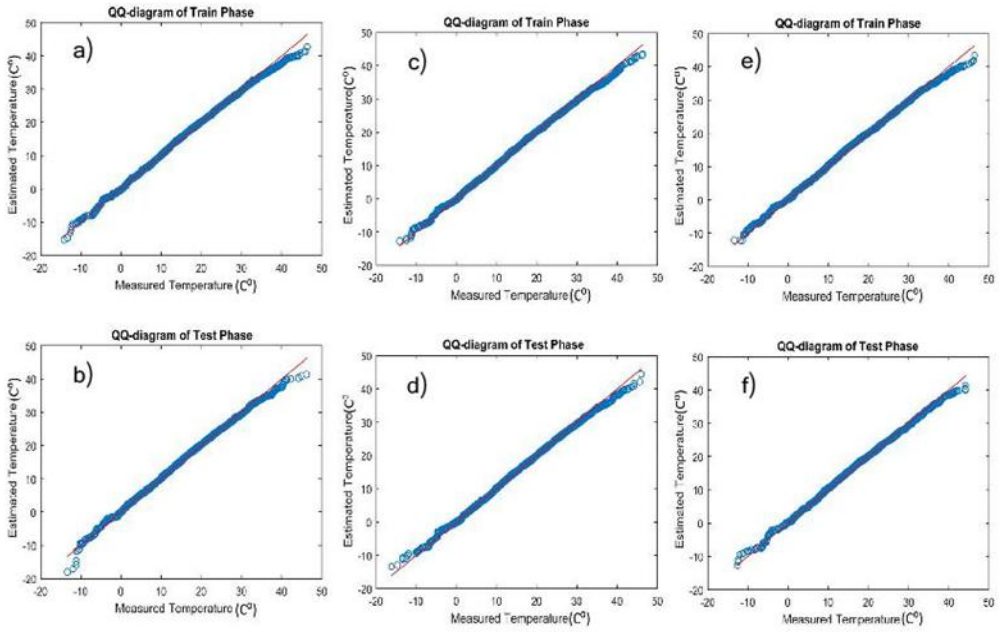


Figure 15: *Q-Q diagram of estimated T2m versus measured temperature during daytime for Industrial LCT using SA-SVR, ANFIS and ANN in testing phase*

Table 6: *Statistic indices between estimated T_{day} values obtained by SA-SVR and measured value from meteorological stationover six LCT in test phase*

LCT	RMSE	MAE	MBE	R2
Agriculture	2.62	2	0.13	0.92
Forest	2.31	1.84	0.06	0.91
Industrial	2.79	2.12	0.09	0.91
Urban	2.46	1.89	0.13	0.92
Airport	2.41	1.87	0.02	0.92
Needleleaf trees	2.42	1.85	0.09	0.93

Moreover, in order to find the optimum number of neurons in hidden layer, various numbers of neurons are used in the MLP and the optimum number of hidden neurons is determined using the CV approach to get the configuration that minimizes the RMSE in the test phase. Figure 14 (a) shows that after a certain number of hidden neurons are added, the model will start over fitting our data and give bad estimates on the test set. This indicates that over fitting starts to occur when the number of neurons is greater than 30, and in this point the model has lowest RMSE, and obviously we can conclude that the optimal number of hidden neurons should be 30, but if we consider the error bar which is the indicator of standard deviation, the less variation was observed at point 40, and then we can say that the model is more stable at this point as compared to point 30 (which is the number of neurons).

Table 7: *Statistic indices between estimated T_{day} values obtained by SA-SVR and measured value from meteorological station over six LCT in test phase*

LCT	RMSE	MAE	MBE	R2
Agriculture	2.62	1.87	0.26	0.88
Forest	2.42	1.67	0.20	0.88
Industrial	2.54	1.84	0.16	0.89
Urban	2.56	1.87	0.16	0.89
Airport	2.07	1.49	0.16	0.92
Needleleaf trees	2.28	1.61	0.18	0.91

Table 8: *Statistic indices between estimated T_{day} values obtained by ANFIS and measured value from meteorological station over six LCT in test phase*

LCT	RMSE	MAE	MBE	R2
Agriculture	2.85	2.21	0.12	0.91
Forest	2.64	2.08	0.69	0.90
Industrial	3.70	2.78	0.17	0.88
Urban	2.74	2.03	0.35	0.90
Airport	2.75	2.08	0.36	0.90
Needleleaf trees	2.64	2.06	0.24	0.90

Table 9: *Statistic indices between estimated T_{night} values obtained by ANFIS and measured value from meteorological station over six LCT in test phase*

LCT	RMSE	MAE	MBE	R2
Agriculture	3.15	2.40	-0.06	0.84
Forest	1.98	1.45	-0.04	0.92
Industrial	2.68	1.91	0.19	0.88
Urban	2.38	1.72	-0.18	0.90
Airport	2.33	1.69	-0.16	0.90
Needleleaf trees	2.28	1.65	0.50	0.92

Table 10: *Statistic indices between estimated T_{day} values obtained by ANN and measured value from meteorological station over six LCT in test phase*

LCT	RMSE	MAE	MBE	R2
Agriculture	2.28	1.82	0.05	0.97
Forest	1.83	1.54	-0.29	0.97
Industrial	2.53	1.94	-0.102	0.93
Urban	2.13	1.62	-0.07	0.97
Airport	2.14	1.69	0.06	0.95
Needleleaf trees	2.08	1.57	0.14	0.95

Moreover, Figure 15 shows Q-Q diagram of SA-SVR (left), ANFIS (middle) and ANN (right) models. Q-Q diagrams are often used to determine whether the model could extract the behavior of the observed data (Chambers et al., 1983). As shown in Figure 15, the models cannot estimate the high temperature for all LC/LU during the day and night time. The weak performance of all models at high temperature are a consequence of a small number of data in these temperatures, and this is also highly related to the study area condition (Berlin) which has a short summer and has only a few number of high temperatures. In these cases, the learning algorithm of the three mentioned models

have the tendency to underestimate the temperature. Therefore, the generalization of these models for the high temperature is reduced.

Table 11: *Statistic indices between estimated T_{night} values obtained by ANN and measured value from meteorological station over six LCT in test phase*

LCT	RMSE	MAE	MBE	R2
Agriculture	2.15	1.59	0.16	0.90
Forest	2.15	1.54	-0.01	0.89
Industrial	2.34	1.73	0.03	0.91
Urban	2.35	1.81	-0.06	0.93
Airport	2.35	1.84	0.10	0.92
Needleleaf trees	2.13	1.55	0.06	0.92

6. Conclusions

In this study, the comparison between the LST and T_{air} observations was done. The comparison shows that LST_{day} and LST_{night} from both Terra and Aqua, with the mean relative bias above and under zero tended to overestimate T_{day} and underestimate T_{night} respectively, and also a higher relative RMSD and bias values were seen for the Aqua $LST_{daytime}$ than the Terra $LST_{daytime}$ which might be given the fact that more solar radiation has been received at the time of the aqua MODIS overpass later in the day. The scatterplots of LST_{night} and T_{night} from Aqua for industrial LC/LU has higher scattering than daytime observations which are more spread around the 1:1 line (Figure 4). This indicates UHI phenomena with RMSD = 4.21°C. Moreover, according to RMSD from Table 3-5 and MODIS LST from Terra, a higher RMSDs is found for industrial and airport LC/LU types during the night time which indicated the UHI phenomena (with RMSD= 4.57°C and 4.32°C respectively). The results show that, the correlations between the MODIS LST from Terra data are generally stronger from the daytime series compared with those from the night time series except for needle leaf trees. The needle leaf tree type showed a more complex correlation pattern from day and night observations. The reason is that the values of LST recorded by MODIS observation on this particular LC type is not exactly a representative of the skin temperature of the soil, but rather affected by the temperature near the top of the trees. In general, the results showed that the time-series of the MODIS LST over six LC/LU classes were correlated individually during the day and night time. They are highly correlated with $r > 0.80$. In addition, the results indicate that, during the warm months the LST_{day} is higher than T_{day} while as expected during cold months LST_{day} is lower than T_{day} for almost all LC/LU. Moreover, for almost all LC/LU, the LST_{night} is close to T_{night}. Overall, the relationship between LST and T_{air} is varied with time and location because the land surface energy balance is a complex phenomenon that depends on multiple factors (e.g., cloud cover, surface roughness, wind speed and soil moisture). In the other words, The LST–T_{air} relationship is mainly controlled by the surface energy balance, but it also depends on factors that are closely linked to energy processes (Prince et al., 1998; Zhang et al., 2015).

Moreover, in this study, the air temperature during the day and night time in the period from 2007 to 2013 was estimated for the Berlin area over six LC/LU, using 1 km Aqua and Terra/MODIS data. The correlation coefficient between observed T_{air} and remotely sensed LST shows an increasing trend, with a spatial window size increasing from 1 km × 1 km to 3 km × 3 km, and subsequently decreasing slightly at window sizes larger than 3 km × 3 km for the daytime, but for the night time this correlation coefficient between observed T_{air} and LST showed a decreasing trend, with spatial window size from 1 km × 1 km to 13 km × 13 km, and subsequently decreasing slightly at window sizes larger than 1 km × 1 km. These window sizes were therefore used to spatially average five satellite-derived environmental

variables, (NDVI, Albedo, Emissivity31, Emissivity32) which were used as predictors of T_{air} in the three models.

In addition, a difficult task with ANN involves choosing the hidden nodes' number. Here, the ANN with one layer was used and the hidden nodes' number was determined using error and trials. For the ANFIS model, Gaussian membership function (MF) and 250 iterations were used. Different number of membership functions were tested and the best of which gave the minimum RMSE and was selected, which was 4 MFs for each variable. For the adjustment, the parameter in SVR model, the simulated annealing was applied. The ANN, ANFIs and SA-SVR models are compared in the test phase based on Table 6-11. The ANN model, among the six LC/LU during the day and night time performed better than the two other models with RMSE which ranged from 1.83°C to 2.53°C and from 2.13°C to 2.35°C during the day and night time respectively. The RMSE of SA-SVR model is ranged from 2.07 to 2.79 °C during the day and night time over six LC/LU and also the highest RMSE was observed in the ANFIS model with a range from 2.64°C to 3.70°C during the day and night time over six LC/LU. These results indicated that the ANN model out performs the SA-SVR and ANFIS models for almost whole LC/LU during the day and night time, but based on Figure 14 and the cross-validation results, the SA-SVR and ANN models out performs the ANFIS model. Moreover, the results showed that there was a high similarity between the training and testing table which demonstrates that the over-fitting has not been occurred in the SA-SVR, ANFIS and ANN. The Q-Q diagram of SA-SVR, ANFIS and ANN shows that all three models slightly tended to underestimate and overestimate the extreme and low temperature for all LC/LU during the day and night time. The weak performance in the extreme and low temperature are a consequence of a small number of data in these temperatures. In these cases, the generalization of these models reduce for estimating the high and low temperature. In addition, despite moderate to high correlations between LST and T_{air}, LST cannot be directly used for estimating air temperature due to the large difference in MBE (Table 3-5), while by applying some additional parameters, in three models (Table 6-11), It can be seen that the MBE was reduced notably, in all LC/LU during day and night time.

Moreover, prediction of long-term monthly air temperature using ANFIS and ANN had been done in the study of Kisi and Shiri (2014). They applied station latitude, longitude and altitude values as input variable to predict the long-term monthly temperature values. They found that the ANN models generally performed better than the ANFIS model in the test period. The ANN models generally performed better than the ANFIS model in the test period and they found that for the ANN model, the maximum and minimum determination coefficient values were between 0.921 and 0.995. The maximum and minimum determination coefficient values were found as 0.99 and 0.876 for the ANFIS model in different stations. Testing results of the ANN and ANFIS models in the study of Kisi and Shiri (2014) show the RMSE values range from 0.1.53 to 4.20◦C and 1.18 ◦C to 9.25◦C for each station, respectively.

Moreover, in the study of Xu. et al. (2014), they applied spatially averaged values of LST, NDVI, modified normalized difference water index (MNDWI), latitude, longitude, distance to ocean, altitude, albedo and solar radiation as predictors of T_{air} in linear regression and random forest models for estimating T_{air} in summer periods from 2003 to 2012. In their study, prior to model development, they also investigated the window size effect on the relationship between LST and T_{air}. The Cross-validation results of their study show that the random forest model (MAE = 2.02°C, R^2 = 0.74) outperforms the linear regression model (MAE = 2.41°C, R^2 = 0.64) and the distribution of residuals from the random forest model slightly overestimates T_{air}, with a mean residual value of 0.09°C.

To sum up, in our study, instead of estimation monthly air temperature and only using the geographical input data, we estimate air temperature during day and night. Moreover, different parameters such as NDVI, Albedo, relative humidity, wind speed, wind direction and Julian day have been take into consideration, which are representative of seasonal changes. The satisfactory results suggested that this modelling approach is appropriate for estimating air temperature in Berlin over six different LC/LU. In addition, the results indicate that MODIS time series of LST can be successfully combined with ground measurements of temperature to produce accurate and more detailed predications of temperature during day and night time. Although the air temperature estimated from satellites tends to be higher than ground-based measurement, the use of satellite remote sensing data can help to overcome the spatial problem of estimating T_{air} particularly in areas with low station density using satellite-based land surface temperature estimation and ground-based relationship between LST and T_{air}. To reduce the biases in satellite-estimated air temperature, it can be effective to use retrieval method based on land surface heat budget (e.g. Kato and Yamaguchi, 2005) in future work.

Acknowledgments

This research was supported by Helmholtz graduate research school GeoSim. We would like thank Dr. Klaus Müller and Thomas Dümmel from the Institute of Meteorology, Freie Universität Berlin, who provided meteorological data in this research. We would also like thank Patricia Margerison for proofreading this manuscript. Our thanks also go to anonymous reviewers for their very thorough and constructive comments.

References

Abraham, A. 2005. *Adaptation of fuzzy inference system using neural learning*. Nedjah, Nadia; de Macedo Mourelle, Luiza, Fuzzy Systems Engineering. Theory and Practice, Studies in Fuzziness and Soft Computing, Springer Verlag, Germany, pp.53-83.

Aarts, E.H.L. and Korst, J. 1989. *Simulated annealing and boltzmann machines: a stochastic approach to combinatorial optimization and neural computing*. Wiley.

Aarts, E.H.L. and Lenstra, J.K. 1997. *Local search in combinatorial optimization*. Chichester, Wiley.

Abramson, D., Krishnamoorthy, M. and Dang, H. 1999. Simulated annealing cooling schedules for the school timetabling problem. *Asia-Pac. J. Oper. Res.*, 16, pp.1-22.

Anily, S. and Federgruen, A. 1987. Simulated annealing methods with general acceptance probabilities. *J. Appl. Probab.*, 24, pp.657-667.

Allen, D.M. 1974. The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, 16, pp.125-127.

Armstrong, J.S. and Collopy, F. 2002. Error measures for generalizing about forecasting methods: empirical comparisons. *Int. J. Forecast*, 8(1), pp.69-80.

Asthana, S. and Bhujade, R.K. 2011. Handwritten Multiscript Pin Code Recognition System having Multiple hidden layers using Back Propagation Neural Network. *International Journal of Electronics Communication and Computer Engineering*.

Ben-Ameur, W. 2004. Computing the initial temperature of simulated annealing. *Comput. Optim. Appl.*, 29, pp.369-385.

Bose, N.K. and Liang, P., 1996. *Neural network fundamentals with graphs, algorithms and applications*. McGRaw-Hill Series in Electrical and Computer Engineering. New York: McGraw-Hill.

Bodri, L. and Cermak, V. 2000. Prediction of extreme precipitation using a neural network: application to summer flood occurrence in Moravia. *Advances in Engineering Software*, 31, pp.311-321.

Boger, Z. and Guterman, H. 1997. *Knowledge extraction from artificial neural network models*. IEEE Systems, Man, and Cybernetics Conference, Orlando, FL, USA.

Berry, M.J.A. and Linoff, G. 1997. *Data Mining Techniques*. NY: John Wiley & Sons.

Blum, A. 1992. *Neural Networks in C++*. NY: Wiley.

Benali, A., Carvalho, A., Nunes, J., Carvalhais, N., and Santos, A. 2012. Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sensing of Environment*, 124, pp.108-121.

Brunel, J. 1989. Estimation of sensible heat flux from measurements of surface radiative temperature and air temperature at two meters: application to determine actual evaporation rate. *Agricultural and Forest Meteorology*, 46(3), pp.179-191.

Beier, C.M., Signell, S.A., Luttman, A. and DeGaetano, A.T. 2012. High-Resolution Climate Change Mapping with Gridded Historical Climate Products. *Landsc. Ecol.*, 27, pp.327-342.

Bilgili, M. and Sahin, B. 2010. Prediction of long-term monthly temperature and rainfall in Turkey. *Energy Sources, Part A*, 32, pp.60-71.

Chai, T. and Draxler, R.R. 2014. Root mean square error (RMSE) or mean absolute error (MAE) – Arguments against avoiding RMSE in the literature. *Geosci Model Dev.*, 7(3), pp.1247-1250.

Czajkowski, K.P., Goward, S.N., Stadler, S. and Walz, A. 2000. Thermal remote sensing of near surface environmental variables: Application over the Oklahoma Mesonet. *The Professional Geographer*, 52(2), pp.345-357.

Cresswell, M.P., Morse, A.P., Thomson, M.C., and Connor, S.J. 1999. Estimating surface air temperatures, from Meteosat land surface temperatures, using an empirical solar zenith angle model. *International Journal of Remote Sensing*, 20(6), pp.1125-1132.

Chen, Y., Randerson, J.T., Morton, D.C., DeFries, R.S., Collatz, Y. and Marlier, M.E. 2011. Forecasting fire season severity in south america using sea surface temperature anomalies. *Science*, 334, pp.787-791.

Cristóbal, J. Ninyerola, M. and Pons, X. 2008. Modeling air temperature through a combination of remote sensing and GIS Data. *Journal of Geophysical Research*, 113, D13106.

Cheval, S. and Dumitrescu, A. 2009. The July urban heat island of Bucharest as derived from MODIS images. *Theor Appl Cli-matol*, 96(1), pp.145-153.

Caselles, V., Coll, C., Valor, E. and Rubio, E. 1995. Mapping land surface emissivity using AVHRR data: Application to La Mancha, Spain. *Remote Sensing Reviews*, 12, pp.311-330.

Chen, K.Y. and Wang, C.H. 2007. Support vector regression with Genetic algorithms in forecasting tourism demand. *Tourism Management*, 28, pp.215-226.

Doukim, C.A., Dargham, J.A. and Chekima, A. 2010. *Finding the number of hidden neurons for an MLP neural network using coarse to fine search technique*. Information Sciences Signal Processing and their Applications (ISSPA), pp.606-609.

Doukim, C.A. and Chekima, A. 2010. *Finding the number of hidden neurons for an MLP neural network using coarse to fine search technique*. International Conference on Information Sciences, Signal Processing and their Applications (ISSPA '10), pp.606-609.

Eglese, R.W. 2014. Simulated annealing: a tool for operational research. *Eur. J. Oper. Res.*, 46, pp.271-281.

Evrendilek, F., Karakaya, N., Gungor, K. and Aslan, G. 2012. Satellite-based and mesoscale regression modeling of monthly air and soil temperatures over complex terrain in Turkey. *Expert Systems with Applications*, 39(2), pp.2059-2066.

Florio, E.N., Lele, S.R., Chi Chang, Y., Sterner, R. and Glass, G E. 2004. Integrating AVHRR satellite data and NOAA ground observations to predict surface air temperature: a statistical approach. *International Journal of Remote Sensing*, 25(15), pp.2979-2994.

Fleischer, M.A. 1995. *Simulated annealing: Past, present, and future*. In: Alexopoulos, C., Kang, K., Lilegdon, W.R. and Goldsman, D., (eds.) Proceedings of the 1995 Winter Simulation Conference, pp. 155-161. IEEE Press, Arlington, Virginia.

Fnaiech, F., Fnaiech, N. and Najim, M. 2001. *A new feed forward neural network hidden layer neuron pruning algorithm*. IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings. (ICASSP '01).

Geisser, S. 1975. The predictive sample reuse method with applications. *Journal of the American Statistical Association*, 70, pp.320-328.

Goward, S.N., Cruickshanks, G.D., and Hope, A.S. 1985. Observed relation between thermal emission and reflected spectral radiance of a complex vegetated landscape. *Remote Sensing of Environment*, 18(2), pp.137-146.

Galvão, C., Valença, M., Vieira, V., Diniz, L., Lacerda, E., Carvalho, A. and Ludermir, T. 1999. *Sistemas inteligentes: Aplicações a recursos hídricos e ciências ambientais*. Porto Alegre, UFRGS/ABRH, 246 pages.

Gardner, M.W. and Dorling, S.R. 1998. Artificial neural networks (the multilayer perceptron) - A review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14-15), pp.2627-2636.

Ghent, D., Kaduk, J. and Balzter, H. 2010. Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. *J. Geophys. Res. Atmos*.

Hájek, P. and Olej, V. 2012. Ozone prediction on the basis of neural networks, support vector regression and methods with uncertainty. *Ecol. Inform.*, 12, pp.31-42.

Henderson, D., Jacobson, S.H. and Johnson, A.W. 2003. *Handbook of Metaheuristics*. Kluwer, Boston, MA.

Haykin, S. 1999. *Neural networks: a comprehensive foundation*. Prentice Hall, Upper Saddle River, NJ.

Haykin, S. 2001. *Redes Neurais: princípios e prática*. Porto Alegre: Editora Bookman, 900 pages.

Hastie, T., Tibshirani, R. and Friedman, J. 2009. *The elements of statistical learning: data mining*. Inference and Prediction, Springer, New York.

Hu, G., Hu, L. and Liu, W. 2010. *Grid resources Prediction with support vector regression and particle swarm optimization*. International Joint Conference on Computational Sciences and Optimization (CSO'10), 1, pp.417-422.

Hong, J.S. 2007. *A literature review of wind forecasting technology in the world*. IEEE Lausanne Power Tech, pp.504-509.

Hunter, D., Hao, Y. and Wilamowski, B.M. 2012. Selection of proper neural network sizes and architecture- A comparative study. *IEEE Trans Indust Inf.*, 8(2), pp.228-240.

IPCC Climate Change. 2007. *The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. In: Solomon, S. and Miller, H.L. (Eds.), United Kingdom: Cambridge University Press, Cambridge.

Ito, K. and Nakano, R. 2005. Optimizing Support Vector regression hyper-parameters based on cross-validation. *International Joint Conference on Neural Networks*, 3, pp.871-876.

Islam, M.M. and Murase, K. 2001. A new algorithm to design compact two-hidden-layer artificial neural networks. *Neural Networks*, 14, , pp.1265-1278.

Jang, Jyh-Shing R. 1991. *Fuzzy modeling using generalized Neural Networks and Kalman Filter Algorithm (PDF)*. National Conference on Artificial Intelligence, Anaheim, CA, USA, pp.762-767.

Jang, J.S.R. 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, 23(3).

Jang, Sun, Mizutani. 1997. *Neuro-fuzzy and soft computing*. Prentice Hall, pp.335-368.

Jang, J.D., Viau, A.A. and Anctil, F. 2004. Neural network estimation of air temperatures from AVHRR data. *International Journal of Remote Sensing*, 25(21), pp.4541-4554.

Jin, M. and Dickinson, R.E. 2010. Land surface skin temperature climatology: benefitting from the strengths of satellite observations. *Environmental Research Letters*, 5(4), p.044004.

Jin, M. 2004. Analysis of land skin temperature using AVHRR observations. *Bulletin of American Meteorological Society*, 85, pp.587-600.

Jain, A., McClendon, R.W., Hoogenboom, G. and Ramyaa, R. 2003. *Prediction of frost for fruit protection using artificial neural networks*. American Society of Agricultural Engineers, St Joseph, MI, ASAE Paper 03-3075.

Korrmann, M. and Unbehauen, H. 1988. Structure detection in the identification of nonlinear systems. *Automatique productive informatique Industrielle (APH)*, 22, pp.5-25.

Keerthi, S. 2002. Efficient tuning of SVM hyper parameters using radius/margin bound and iterative algorithms. *IEEE Transactions on Neural Networks*, 13, pp.1225-1229.

Koulamas, C., Antony, S.R. and Jaen, R. 1994. A survey of simulated annealing applications to operations research problems. *OMEGA-Int. J. Manage. Sci.*, 22, pp.41-56.

Kazuhiro Shin-ike. 2010. *A two phase method for determining the number of neurons in the hidden layer of a 3-layer neural network*. SICE Annual Conference, The Grand Hotel, Taipei, Taiwan.

Ke, J. and Liu, X. 2008. Empirical analysis of optimal hidden neurons in neural network modeling for stock prediction. *Pac Asia Work Comput Intell Indust Appl.*, 2, pp.828-832.

Kisi, O. and Shiri, J. 2011. Precipitation forecasting using wavelet-genetic programming and wavelet-neuro-fuzzy conjunction models. *Water Resource Management*, 25(13), pp.3135-3152.

Luk, K.C., Ball, J.E. and Sharma, A. 2000. A study of optimal model lag and spatial inputs for artificial neural network for rainfall forecasting. *Journal of Hydrology*, 227, pp.56-65.

Land Surface Temperature - Copernicus Global Land Service. 2016. Available online: http://land.copernicus. eu/global/products/lst (accessed on 1 March 2016).

Land Surface Temperature: Global Maps. 2016. Available online:http://earthobservatory.nasa.gov/GlobalMaps/view.php?d1=MOD11C1_M_LSTDA. Accessed on 1 March 2016.

Maier, H.R., Jain, A., Dandy, G.C. and Sudheer, K.P. 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environ. Model. Softw.*, 25, pp.891-909.

Muller, M. and Fill, H. 2003. *Redes Neurais aplicadas na propagação de vazões, in: Simpósio Brasileiro de Recursos Hídricos*. Curitiba, Brazil, unpaginated CD-Rom Proceedings.

Mostovoy, G.V., King, R.L., Reddy, K.R., Kakani, V.G. and Filippova, M.G. 2006. Statistical Estimation of Daily Maximum and Minimum Air Temperatures from MODIS LST Data over the State of Mississippi. *GI Science & Remote Sensing*, 43(1), pp.78-110.

Mittal, G.S. and Zhang, J. 2003. Artificial neural network-based psychrometric predictor. *Biosystems Engineering*, 85(3), pp.283-289.

Maqsood, I., Khan, M.R. and Abraham, A. 2004. An ensemble of neural networks for weather forecasting. *Neural Computing & Applications*, 13, pp.112-122.

Maddala, G.S. 2011. *Introduction to econometrics*. 3rd ed. John Wiley & Sons.

Meteotest. 2010. Meteonorm handbook, Part III: Theory Part 2. Accessed online in 9th February 2011. Availabe from: http://www.meteonorm.com/media/pdf/mn6_software.pdf.

Manzo-Delgado, L., Sánchez-Colón, S. and Álvarez, R. 2009. Assessment of Seasonal Forest Fire Risk Using Noaa-Avhrr: A Case Study in Central Mexico. *Int. J. Remote Sens.*, 30, pp.4991-5013.

Marzban, F., Sodoudi, S. and Preusker, R. 2017. The influence of land cover type on the relationship between NDVI-LST and LST-T_{Air}. *International Journal of Remote Sensing*.

Cristianini, N. and Taylor, L. 2000. *An introduction to support vector machines*. Cambridge University Press, Cambridge, UK.

Nieto, H., Sandholt, I., Aguado, I., Chuvieco, E. and Stisen, S. 2011. Air temperature estimation with MSG-SEVIRI data: Calibration and validation of the TVX algorithm for the Iberian Peninsula. *Remote Sensing of Environment*, 115(1), pp.107-116.

Near Surface Air Temperature - GES DISC-Goddard Earth Sciences Data and Information Services Center. 2016. Available online: http://disc.gsfc.nasa.gov/hydrology/dataholdings/parameters/near_surf_air_temp.shtml. Accessed on 17 January 2016.

Nguyen, O. and Suwandana, E. 2015. Temporal change and its spatial variety on land surface temperature and land use changes in the Red River Delta, Vietnam, using MODIS time series imagery. *Environ Monit Assess*.

Oyler, J.W., Ballantyne, A. and Running, S.W. 2015. Creating a topoclimatic daily air temperature dataset for the conterminous united states using homogenized station data and remotely sensed land skin temperature. Int. J. Climatol, 35, pp.2258-2279.

Onoda, T. 1995. *Neural network information criterion for the optimal number of hidden units*. iIEEE International Conference on Neural Networks, pp.275-280.

Panchal, G., Ganatra, A. and Panchal, D. 2011. Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. *Int J Comput Theory Eng.*, 3(2), pp.332-337.

Prihodko, L. and Goward, S.N. 1997. Estimation of air temperature from remotely sensed surface observations. *Remote Sensing of Environment*, 60(3), pp.335-346.

Prince, S., Goetz, S., Dubayah, R., Czajkowski, K. and Thawley, M. 1998. Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using Advanced Very High-Resolution Radiometer satellite observations: comparison with field observations. *Journal of Hydrology*, 212-213, pp.230-249.

Peón, J., Carmen, R. and Javier, F.C. 2014. Improvements in the estimation of daily minimum air temperature in peninsular spain using MODIS land surface temperature. *Int. J. Remote Sens.*, 35, pp.5148-5166.

Panchal, G. and Panchal, D. 2011. Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. *International Journal of Computer Theory and Engineering*, 3, pp.332-337.

Razavi, S. and Tolson, B.A. 2011. A New Formulation for Feedforward Neural Networks. *IEEE Trans. Neural Netw.*, 22, pp.1588-1598.

Romeo, F. and Sangiovanni-Vincentelli, A. 1991. A theoretical framework for simulated annealing. *Algorithmic*, 6, pp.302-345.

Rivals, I. and Personnaz, L. 2000. *A statistical procedure for determining the optimal number of hidden neurons of a neural model*. Second International Symposium on NeuralComputation (NC'2000), Berlin.

Reich, P.B., Luo, Y. and Oleksyn, J. 2014. Temperature drives global patterns in forest biomass distribution in Leaves, Stems, and Roots. *Proc. Natl. Acad. Sci. USA*, 111, pp.13721-13726.

Ruane, A.C., McDermid, S. and Cecil, L.D. 2014. Carbon-temperature-water change analysis for peanut production under climate change: a prototype for the agmip coordinated climate-crop modeling project (C3mp). *Glob. Chang. Biol.*, 20, pp.394-407.

Rosenzweig, C., Elliott, J. and Khabarov, N. 2014. Assessing Agricultural Risks of Climate Change in the 21st Century in a Global Gridded Crop Model Intercomparison. *Proc. Natl. Acad. Sci. USA*, 111, pp.3268-3273.

Snyder, W.C., Wan, Z., Zhang, Y. and Feng, Y.Z. 1998. Classification based emissivity for land surface temperature measurement from space. *International Journal of Remote Sensing*, 19, pp.2753-2774.

Sobrino, J.A., Raissouni, N. and Li, Z.L. 2001. A comparative study of land surface emissivity retrieval from NOAA data. *Remote Sensing of Environment*, 75, pp.256-266.

Shuxiang, X. and Chen, L. 2008. *A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining*. International Conference on Information Technology and Application, pp.683-686.

Sousa, S., Martins, F., Alvimferraz, M. and Pereira, M. 2007. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environ. Model. Softw.*, 22, pp.97-103.

Sartakhti, J.S. and Mozafari, K. 2011. Hepatitis disease diagnosis using a novel hybridmethod based on support vector machine and simulated annealing (SVM-SA). *Computer Methods and Programs in Biomedicine*, 108, pp.570-579.

Stisen, S., Sandholt, I. and Eklundh, L. 2007. Estimation of diurnal air temperature using MSG SEVIRI data in West Africa. *Remote Sensing of Environment*, 110(2), pp.262-274.

Steidley, C., Sadovski, A., Tissot, P. and Bachnak, R. 2005. *Using an artificial neural network to improve predictions of water level where tide charts fail*. In: Ali, M. (Editors), Innovations in Applied Artificial Intelligence, Springer, Bari, Italy.

Suman, B. and Kumar, P. 2006. A survey of simulated annealing as a tool for single and multi objective optimization. *J. Oper. Res. Soc.*, 57, pp.1143-1160.

Sugeno, M. and Tanaka, K. 1992. Successive identification of a fuzzy model and its applications to prediction of a complex system. *Proc Fuzzy Sets Syst.*, 42, pp.315-334.

Smith, B.A., McClendon, R.W. and Hoogenboom, G. 2006. Improving Air Temperature Prediction with Artificial Neural Networks. *International Journal of Computational Intelligence*, 3(3), pp.179-186.

Stone, M., 1974. Cross-validation choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, 36, pp.111-147.

Sandholt, I., Rasmussen, K. and Andersen, J. 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of Environment*, 79(2-3), pp.213-224.

Smola, A.J. and Schölkopf, B. 2004. A tutorial on support vector regression. *Statistics and Computing*, 14, pp.199-222.

Sun, Y.J., Wang, J.F., Zhang, R.H., Gillies, R.R., Xue, Y. and Bo, Y.C. 2005. Air temperature retrieval from remote sensing data based on thermodynamics. *Theoretical and Applied Climatology*, 80(1), pp.37-48.

Shamir, E. and Georgakakos, K.P. 2014. MODIS land surface temperature as an index of surface air temperature for operational snowpack estimation. *Remote Sens. Environ.*, 152, pp.83-98.

Smith, B.A., McClendon, R.W. and Hoogenboom, G. 2005. An enhanced artificial neural network for air temperature prediction. *World Academy of Science, Engineering and Technology (PWASET)*, 7, pp.7-12.

Shank, D.B., Hoogenboom, G. and McClendon, R.W. 2008. Dewpoint temperature prediction using artificial neural networks. *Journal of Applied Meteorology and Climatology*, 47, pp.1757-1769.

Stahl, K.R.D., Moore, J.A. Floyer, M.G. and Mckendry, I.G. 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agricultural and Forest Meteorology*, 139, pp.224-236.

Schölkopf, B. 1998. Learning with kernels, Ph.D. Thesis, GMD, Birlinghoven, Germany.

Takagi, T. and Sugeno, M. 1985. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans Syst Man Cybern*, 15(1), pp.116-32.

Unger, J., Gál, T. and Tobak, Z. 2009. *Air temperature versus surface temperature in urban environment*. International Conference on Urban Climate, Yokohama, Japan.

Ustün, B., Melssen, W.J. and Oudenhuijzen, M. 2005. Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization. *Analytica Chimica Acta*, 544, pp.292-305.

Van Laarhoven, P.J.M. 1988. *Theoretical and computational aspects of simulated annealing*. Centrum voor Wiskunde en Informatica, Amsterdam, Netherlands.

Vancutsem, C., Ceccato, P., Dinku, T., and Connor, S.J. 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sensing of Environment*, 114(2), pp.449-465.

Van Laarhoven, P.J.M. and Aarts, E.H.L. 1987. *Simulated annealing: Theory and applications*. D. Reidel; Kluwer, Dordrecht, Boston, Norwell, MA.

Van Leeuwen, T., Frank, D., Jin, Y., Smyth, P. and Randerson, J. 2011. Optimal use of land surface temperature data to detect changes in tropical forest cover. *J Geophys Res*., 116, p.G02002.

Vogt, J., Viau, A.A. and Paquet, F. 1997. Mapping regional air temperature fields using satellite derived surface skin temperatures. *International Journal of Climatology*, 17, pp.1559-1579.

Vazquez, D.P., Reyes, F.J.O. and Arboledas, L.A. 1997. A comparative study of algorithms for estimating land surface temperature from AVHRR data. *Remote Sensing of Environment*, 62, pp. 215-222.

Wloczyk, C., Borg, E., Richter, R., and Miegel, K. 2011. Estimation of instantaneous air temperature above vegetation and soil surfaces from Landsat 7 ETM+ data in northern Germany. *International Journal of Remote Sensing*, 32(24), pp.9119-9136.

Wedge, D., Ingram, D., McLean, D., Mingham, C. and Bandar, Z. 2005. *A global-local artificial neural network with application to wave overtopping prediction*. In: Duch, W. and Kacprzyk, J. (Editors), Artificial Neural Networks: Formal Models and Their Applications – ICANN, Springer, Warsaw, Poland.

Willmott, C.J. and Matsuura, K. 2005. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. *Clim. Res.*, 30, pp.79-82.

Wan, Z., Zhang, Y. and Zhang, Q. 2004. Quality assessment and validation of the MODIS global land surface temperature. *Int. J. Remote Sens.*, 25, pp.261-274.

Wan, Z.. 2008. Radiance-based validation of the V5 MODIS land-surface temperature product. *Int. J. Remote Sens.*, 29, pp.5373-5395.

Westerling, A.L., Hidalgo, H.G., Cayan, D.R. and Swetnam, T.W. 2006. Warming and Earlier Spring Increase Western U.S. *Forest Wildfire Activity*, 313, pp.940-943.

Wang, J. and Tan, Z. 2012. An annual load forecasting model based on support vector regression with differential evolution algorithm. *Applied Energy*, 94, pp.65-70.

Xu, Y., Qin, Z. and Shen, Y. 2012. Study on the estimation of near-surface air temperature from MODIS data by statistical methods. *International Journal of Remote Sensing*, 33(24), pp.7629-7643.

Xu, Y. and Shen, Y. 2013. Recunstruction of the land surface temperature time series using harmonic analysis. *Computers & Geosciences*, 61, pp.126-132.

Yoo, J.M., Won, Y.I., Cho, Y.J., Jeong, M.J., Shin, D.B., Lee, S.J., Lee, Y.R., Oh, S.M., and Ban, S.J. 2011. Temperature trends in the skin/surface, midtroposphere and low stratosphere near Korea from satellite and ground measurements. *Asia-Pacific Journal of Atmospheric Sciences*, 47, pp.439-455.

Yang, Y. 2007b. Consistency of cross validation for comparing regression procedures. *The Annals of Statistics*, 35, pp.2450-2473.

Yuan, H.C. and Huai, X.Y. 2003. Amethod for estimating the number of hidden neurons in feed-forward neural networks based on information entropy. *Computers and Electronics in Agriculture*, 40, pp.57-64.

Zhu, W. and Jia, S. 2013. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment*, 130, pp.62-73.

Zhang, W., Huang, Y., Yu, Y. and Sun, W. 2011a. Empirical models for estimating daily maximum, minimum and mean air temperatures with MODIS land surface temperatures. *International Journal of Remote Sensing*, 32(1), pp.1-26.

Zaksek, K. and Schroedter-Homscheidt, M. 2009. Parameterization of air temperature in high temporal and spatial resolution from a combination of the SEVIRI and MODIS instruments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(4), pp.414-421.

Zheng, X., Zhu, J. and Yan, Q. 2013. Monthly air temperatures over Northern China estimated by integrating MODIS Data with Gis Techniques. *J. Appl. Meteorol. Climatol*, 52, pp.1987-2000.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2518-2539
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.340

Research Article

Assessment of Soil Erosion Risk in Mubi South Watershed, Adamawa State, Nigeria

S.R. Thlakma[1], E.O. Iguisi[2], A.C. Odunze[3], D.N. Jeb[4]

[1]Department of Geography, Federal University Kashere, Gombe State, Nigeria
[2]Department of Geography Ahmadu Bello University Zaria, Kaduna State, Nigeria
[3]Department of Soil Science Ahmadu Bello University Zaria, Kaduna State, Nigeria
[4]National Center of Remote Sensing Jos, Plateau State, Nigeria

Publication Date: 22 February 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.340

Abstract The study aims to assessment soil erosion risk in Mubi South watershed with the aid of RUSEL model and Geospatial techniques. RUSLE model parameters such as rainfall, soil map, topography map, cover management and conservation practice factor map were derived. The method employed includes the use of RUSLE model and Geospatial techniques using ArcGIS 10.3 Software, for analysis and presentation of result. It was found that sandy soil are the dominant soil of the watershed which covered about 65%, 18% silt and 17% clay. The landuse landcover has about 29 % of area covered by agricultural activities, 19% were covered by forest, and 25% were not cultivated and covered by bare land. The study area has about 0.58 to -0.07 normalized difference vegetation index (NDVI) with majority of the area within the lower topography of 570 m above sea level. The Soil cover management factor ranges from the higher value of 0.5 to the lower value of 0.01 in the watershed, 15.8 mm to 15.7 mm occurrences of daily rainfall and 492.34 mm rainfall and runoff covered when rainfall per-day is greater than 15 mm rainfall. The results of the study also show that average rate of soil detachment is 1t ha^{-1} yr^{-1}. The average transport capacity of overland flow is 1.5 t ha^{-1} yr^{-1}. Average soil per detachability by raindrop is 69.6 t ha^{-1} yr^{-1} total soil particle detachments is 69.66 t ha^{-1} yr^{-1} and average estimated soil erosion of 3.52 t ha^{-1} yr^{-1}. It is recommended that other soil erosion model to be applied in the study area for further comparative analysis of soil erosion risk.
Keywords *Erosion; Risk; Soil; RUSLE and Geospatial techniques; Watershed area*

1. Introduction

Soil erosion is a natural process of soil material removal and transportation through the action of erosive agents such as water, wind, gravity, and human disturbance (Aksoy et al., 2009; Asdak, 2009; Kefi et al., 2009; Hacısalihoglu et al., 2010). However, if soil erosion is occurring faster than necessary due to human disturbance, it will cause negative impacts on the environment and e economy (Kefi and Yoshino, 2010).

Soil erosion potential risk is determined by all natural phenomena, which could cause erosion damages (Auerswald, 1993). Soil erosion actual risk is the potential risk plus human induced intensification of the potential risk. The actual erosion and soil erosion risk is determined by all natural and human caused phenomena, which lead to soil erosion (Auerswald, 1993).

Soil erosion by water is estimated as the most extensive erosion type and results from excess surface runoff. The scope of water erosion is influenced by type of soil; slope and land cover (Verheijen et al., 2009). Through the removal of surface soil (including organic matter and nutrients) from soil mass, effective soil functioning is affected. About one-third of land used for agriculture at global level has been affected by soil degradation. Most of this damage was caused by water and wind erosion (Braimoh and Vlek, 2008).

In twenty-first century, soil erosion by water has become a worldwide issue because of progressive decrease in the ratio between natural resources and population and to climate change. Soil erosion negatively impacts on ecology and can lead to reduced crop productivity, worsened water quality, lower effective reservoir of water levels, flooding and habitat destruction (Park et al., 2011). In both the past and present day, soil erosion is one major and most widespread environmental threat. Risk assessment of soil erosion caused by water is indispensable to the creation of effective policies and measures on water and soil resource conservation.

In Nigeria, World Bank (1990) estimated that soil erosion affects over 50 million people and account for loss of resources that amount to US 3000 million dollars per year. For decades, soil erosion has been a major environmental problem in Nigeria (Olofin, 1994). Erosion is the most serious natural hazard in Nigeria, affecting several parts of the country. It has killed people, destroyed roads, destroyed homes, schools and farmlands and displaced poor people (Federal Republic of Nigeria, 2007).

To analyse soil erosion and suggest appropriate management plans, Revised Universal Soil Loss Equation (RUSLE) Models was selected for this research. Also, RUSEL was chosen in this research over other model such as; Soil and Water Assessment Tool (SWAT), European Soil Erosion Model (EUROSEM), and Annualized Agricultural Non- Point Source (AnnAGNPS) because these models applied worldwide to soil loss estimation and their convenience in application and compatibility with GIS (Kouli et al., 2009; Pandey et al., 2009; Bonilla et al., 2010). RUSLE models was selected and applied in this research because of their simplicity and flexibility in use as compared to other models and needs less data than most of the other erosion estimation models. Revised Universal Soil Loss Equation Models is easy in integration with GIS and their performance at a watershed/catchment level in Mubi South is not yet known to the best of the researcher's knowledge; hence, need to apply two models in this study in order to estimate soil erosion in Mubi South Local Government Area of Adamawa State, Nigeria with the aid of geospatial technology.

1.1. Statement of Research Problems

Soil erosion and related degradation of land resources are highly significant spatio-temporal phenomena in many countries (Pandey et al., 2009). Soil problems have become a threat to sustainable agricultural production and water quality. In many regions, unchecked soil erosion and associated land degradation have made vast areas economically unproductive. Often, a quantitative assessment is needed to infer extent and magnitude of soil erosion risk so that effective management strategies can be resorted to. The complexities of variables make precise estimation or prediction of soil erosion difficult.

Soil erosion of various types and extent are also found in various parts of Adamawa state but most especially where man's activities have stripped off vegetation that normally holds and protects the soil. In Adamawa state, researches have shown that the different causes of soil erosion sprang from human activities for various purposes such as; intensive cultivation, over grazing, bush burning and

deforestation. These are the principal determinants of variation in types and intensity of soil erosion and Mubi Local Government Area is not exceptional like any other part of Adamawa State (Tekwa, Laflen and Yesuf, 2014).

Previous studies have been conducted within and outside the study area. Tekwa et al. (2014) conducted a research titled “estimation of monthly soil loss from ephemeral gully erosion features in some parts of Mubi North and Mubi South, L.G.A of Adamawa State, Northeastern Nigeria”. They found that soil aggregate particles in the study area were mainly sandy; with silt content range of 18-25 % and clay contents in the range of 19-26 % that did not differ significantly among selected sites. Organic matter content was low. The monthly area of soil loss ranges from 1.5 to 143 m^2, and volume of soil loss was 0.4 to 131 m^3, that were significantly higher in the months of August and September than in the months of June, July and October. Ephemeral gully erosion rates for Muvur and Digil sites were greater than at other sites. The monthly rates of ephemeral gully erosion ranged from 35 to 132 m^3, and 15 to 79 m^3 in terms of surface area and volume of soil loss respectively. The soil loss rates thereafter decreased from 18 m^2 to 5 m^2 and 11 m^2 to 2 m^2 in terms of surface area and volume of soil loss respectively. The researcher recommended that future researches should consider developing empirical soil loss predictor model (s) for Mubi and environs. The researches concentrated on only gully erosion for selected sites and concentrated mainly on their chemical properties. Also, the authors took soil samples only in gully areas and did not employ GIS techniques to estimate and predict spatial distribution of soil erosion risk.

Interest in soil erosion risk was triggered by a growing awareness of off-site impacts of soil erosion. These impacts are predominantly associated with movement of eroded soil, sediment particles and changes in water flows (both through and across the soil). The off-site problems are often more evident, and include loading and sedimentation of water courses and reservoirs, increases in stream turbidity; all of which can disturb aquatic ecosystems and upset the geomorphological functioning of river systems (Owens, 2005).

To reduce such limitations; geostatistic techniques that interpolate data for an entire catchment from appropriately sampled point measurements, are readily available (Tesfahunegn et al., 2011a and b). Mapping through conventional methods demand intensive data collection, which is often difficult to practice in complex terrains (Tesfahunegn et al., 2011b). The Geographic Information System (GIS) techniques can provide easy and time effective tools to map and analyze erosion input data of hydrophysical parameters (Tesfahunegn et al., 2011b).

1.2. Aim and Objectives

The aim of this study is to estimate soil erosion risk in Mubi South watershed area with the aid of RUSLE models and Geospatial techniques.

1.3. The Scope of the Study

On the basis of spatial extent, the research was carried out in Mubi south watershed area, Adamawa State and focused on analysis of soil erosion risk using Revised MMF and RUSEL model with GIS techniques. The watershed area consists of the following fourteen (14) villages and Local Government Headquarters, namely: Sebbore, Gude, Gudere, Wafa, Chaba, Masuwa, Lunguwa, Gyakwar, Wuro Babbowa, Gavayi, Gella 2DH, Gella, Giranburum and Uro Gella which is the Local Government Headquarters of Mubi South. On content, 80 soil samples were collected in July 2016 and coordinates of the sample points were derived using grid system in GIS software environment as shown in Figure 2. The parameters used as input data were collected from different sources such parameters such as

rainfall, soil map, topography map, cover management and conservation practice factor map were derived. The research covered soil erosion data as at July, 2016.

1.4. The Study Area

Location and Description of the Study area

Mubi south Local Government Area is located in Northeast Nigeria between latitudes 10º 4′ 30”N - 10º 15′ 0” N, and Longitudes 13º 20′E 0” - 13º 27′ 0”E of the Greenwich Meridian. The study catchment area covered about 148.43 km^2 (sq km). The study area is bordered by Lamurde from North-East, Gella Local Government Area to the East, Wuro Bobbowa and Girgi in the South-West. The map and location of study area is show on Figure 1.

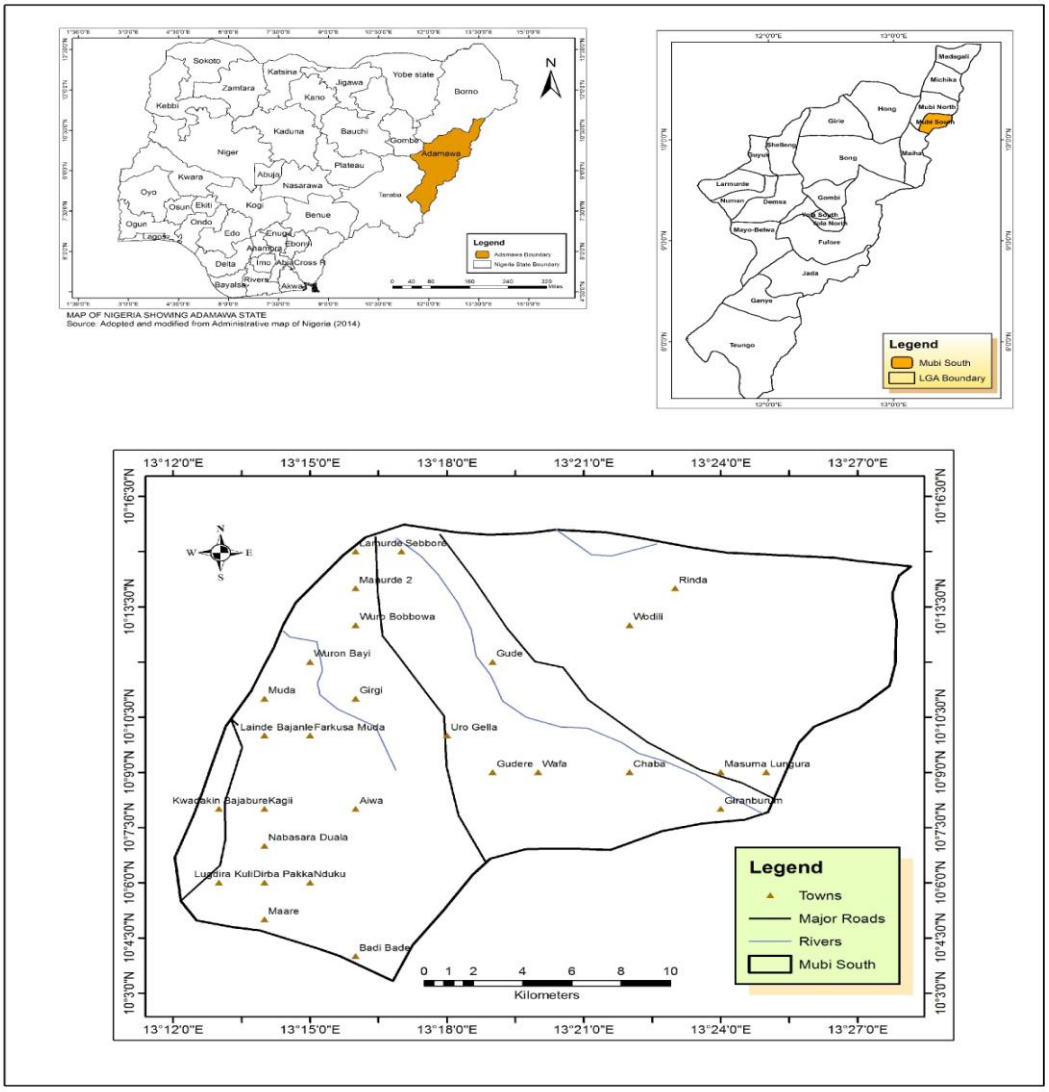


Figure 1*: Map of the study area [Source: Modified from the Administrative Map of Adamawa State and field Survey (2015)]*

The climate of the study area is typical of the West African Savanna climate. Temperature in this climatic region is high because of the radiation income, which is relatively evenly distributed. However, there is usually a seasonal change in the temperature. There is gradual increase in temperature from January to April. There is also a distinct drop in temperature at the onset of rains due to the effect of cloudiness. A slight increase after the cessation of rain (October to November) is common before the onset of harmattan in December the temperature in Yola reach 40^{o}C particularly in April and while minimum temperature can be as low as 18^{o}C in the south to 27.8^{o}C in the northeastern part in December (Adebayo and Tukur, 1999). Rainfall Erosivity ranges between 481m to 192m with about 15.5mm to 15.8mm rainfall per day and 4.5 m to 4.6 m rate of potential evapotranspiration.

The area is characterized by a typical tropical wet (April-October) and dry (November-March) climate with a mean annual rainfall ranging from 700 mm to 1,050 mm (Adebayo, 2004). The vegetation is a typical Sudan savanna with short grasses interspersed with shrubs and few trees (Adebayo, 2004; Tekwa and Usman, 2006).

The study area is usually characterized by orchard-type vegetation due to its limitation in inherent fertility (Nwaka et al., 1999). The major vegetation formations in the State are the Southern Guinea Savannah, Northern Guinea Savannah, and the Sudan Savannah. Within each formation is an interspersion of thickets, tress savannah, Open grass savannah and fringing forest in the river valleys. It is however necessary to note that large scale deforestation resulting from indiscrimination extraction of wood for fuel and expansion of agricultural land areas have left large area within each vegetation type with few indigenous woody plant species. Most areas especially those close to settlements are covered with exotic species such as the neem and eucalyptus trees.

Soils of the study area belong to the order lithosols (Agboola, 1979; Adebayo, 2004; Tekwa and Usman, 2006). Lithosols constitute one of the upper categories of FAO/UNESCO soil classification system (Aduayi et al., 2002). They refer to soils with rock-basements within shallow depths from the soil surface and this implies shallowness and stoniness of the surface soil depths. Arenosols and Regosols: There are relatively young soils or soils with very little or no profile developments, or very homogenous sands, are grouped together. These are found on mountain sites within the 213 and 232 units. On these types of soils, weathering is slight and involves no accumulation of the products of weathering. The B. horizon may not be very clear and reddish in colour, while the original carbon content is most of the time leached out. The study area have soil moisture 0.072 %, bulk density of 1.63 Mgm^{-3}, 2.33 gkg^{-1} soil particle densities, 6.66 gkg^{-1} organic carbon, 0.68 mm of soil porosity and 11.46 gkg^{-1} organic matter.

Geology of the area consists of Precambrian Basement rocks, while parent material of the soil is undifferentiated Basement Complex, represented by migmatite-gneisses, schists, quartzites aplite, medium and coarse-grained granites, pegmatite, diorite, and amphibolites (Adebayo, 2004). The dominant landuses in the study area are; agricultural land, forestry/vegetation, water body, builtup Area and bareland. Moreover, the town has become center of learning with numerous tertiary and secondary institutions established in the metropolis.

The study area has a total projected population of 126,378 people (National Population Census, 2009) in 2015. The growth of Mubi town is traced to agricultural, administrative, and commercial functions it performs.

2. Methodology

2.1. Reconnaissance Survey

Reconnaissance survey was carried out by the researcher to get acquainted with the study area in terms of selections of coordinate location points, choice for major land use classes, ground thruthing and major crop types selected for the study.

2.2. Type and Sources of Data Used

The types and sources of data used for this research are summarized in Table 1.

Table 1: *Types, sources and used of data*

Types of Data	Sources of Data	Uses
Landsat thematic mapper of 2016 with 30m resolution	Download from GLCF web	Input Parameter for the Model as land use type
ASTER Image (DEM) (Advanced Spaceborne Thermal Emission and Reflection Radiometer)	Download from GLCF web	Input Parameter for RUSLE Model
Rainfall data	Geography department ADSU Mubi	Input Parameter for the RUSLE Model
Soil texture (Particle size distribution)	Laboratory determination	Input Parameter for the RMMF Model
Crop types cover	Field survey	Input Parameter for RUSLE Model
Vegetation cover	Landsat imagery of 2015 in ArcGIS 10.3	Vegetation cover conditions for RUSLE
Slope steepness	From ASTER image	Input Parameter for RUSLE Model

Source: Adopted from Iguisi (2003)

2.3. Input Parameters for the RUSLE Model

The Input Parameters used for the RUSLE Model were presented on Table 2.

Table 2: *Inputs parameters for RUSLE*

Factor	Parameter	Definition and remarks
Rainfall	R	Rainfall-runoff erosivity factor (MJ.mm/ ha.hr.year) (Arnoldus, 1980)
Soil	K	Soil texture/ erodibility in ton.ha.hr/ (MJ.mm.ha) (USDA, 1978)
Topographic	LS	Slope length and slope steepness (m) (Moore and Burch, 1986 a, b)
Cover management	C	Ratio of soil loss from land cropped under specific conditions to the corresponding loss from clean-tilled, continuous fallow conditions (Zhou et al., 2009).
Conservation practice	P	Soil conservation operations or other measures that control the erosion. The values of P-factor ranges from 0 to 1 (Renard et al., 1997).

Source: Compiled by the Author (2016)

2.3.1. Deriving Inputs of RUSLE Model

The RUSLE Model equation is a function of five input factors in raster data format: rainfall erosivity, soil erodibility, slope length and steepness, cover management and support practice. These factors vary over space and time and depend on other input variables. Therefore, soil erosion within each pixel was estimated with the RUSLE. The RUSLE method is expressed as:

$$A = R \times K \times LS \times C \times P \quad \ldots (1)$$

Where A is the computed spatial average of soil loss over a period selected for R, usually on yearly basis (t ha^{-1} y^{-1}); R is the rainfall-runoff erosivity factor (MJ mm t ha^{-1} y^{-1}); K is the soil erodability factor (t ha^{-1} $y^{-1}MJ^{-1}$ mm^{-1}); LS is the slope length steepness factor (dimensionless); C is the cover management factor (dimensionless, ranging between 0 and 1.5); and P is the erosion control (conservation support) practices factor (dimensionless, ranging between 0 and 1).

Rainfall erosivity (R): The rainfall factor, an index unit, is a measure of the erosive force of a specific rainfall. This was determined as a function of the volume, intensity and duration of rainfall and can be computed from a single storm, or a series of storms to include cumulative erosivity from any time period. Raindrop/splash erosion is the dominant type of erosion in barren soil surfaces. Rainfall data of 11 years (2004 to 2015) collected from Department of Geography Meteorological unit Adamawa State University Mubi was used for calculating R-factor using the following relationship developed by Wischmeier and Smith (1978) and modified by Arnoldus (1980):

$$R = \sum_{i=1}^{12} 1.735 \times 10^{\left(1.5 \log_{10}\left(\frac{P_i^2}{P}\right) - 0.08188\right)} \quad \ldots (2)$$

Soil Erodibility Factor (K)

Different soil types are naturally resistant and susceptible to more erosion than other soils and is a function of grain size, drainage potential, structural integrity, organic matter content and cohesiveness. Erodibility of soil is its resistance to both detachment and transport. Soil texture map of the study area was used for the preparation of K factor map and soil types were grouped into major textural classes. The corresponding K values for soil types were identified from soil erodibility nomograph (USDA, 1978) by considering particle size, organic matter and permeability class.

Slope Length and Steepness Factor (LS)

Length and steepness of a slope affects total sediment yield from the site and is accounted by the LS-factor in RUSLE model. In addition to steepness and length, other factors; such as compaction, consolidation and disturbance of the soil were also being considered while generating LS-factor. Erosion increases with slope steepness but, in contrast to L-factor representing effects of slope length. The combined LS-factor was computed for the watershed by means of ArcGIS Spatial analyst extension using DEM, as proposed by Moore and Burch (1986a, b). The flow accumulation and slope steepness were computed from the DEM using ArcGIS Spatial analyst.

$$LS = (\text{Flow accumulation} \times \text{Cell size}/22.13)^{0.4} \times (\sin \text{slope}/0.0896)^{1.3} \quad \ldots (3)$$

Where flow accumulation denotes the accumulated upslope contributing area for a given cell, LS=combined slope length and slope steepness factor, cell size=size of grid cell (for this study 30 m) and sin slope =slope degree value in sin.

Cover management factor (C)

The C-factor represents effect of soil- Cover management factor (C). The C-factor represents effect of soil-disturbing activities, plants, crop sequence and productivity level, soil cover and subsurface bio-mass on soil erosion. Due to the variety of land cover patterns with spatial and temporal variations, satellite remote sensing data sets were used for the assessment of C-factor (Karydas, 2009; Tian et al., 2009). The Normalized Difference Vegetation Index (NDVI), an indicator of the vegetation vigor and health was used along with the following formula (eq. 4) to generate C-factor value image for the study area (Zhou et al., 2009; Kouli et al., 2009).

$$C = \exp\left[-\alpha \frac{\mathrm{NDVI}}{(\beta - \mathrm{NDVI})}\right] \quad \ldots (4)$$

Where a and b are unitless parameters that determine shape of the curve relating to NDVI and the C-factor. VanderKnijff et al. (2000) found that this scaling approach gave better results than assuming a linear relationship and values of 2 and 1 were selected for parameters a and b, respectively. This equation was applied for assessing C-factor of areas with similar terrain and climatic conditions (Prasannakumar et al., 2011a; Prasannakumar et al., 2011b).

Conservation practice factor (P)

The support practice factor (P-factor) is the soil-loss ratio with a specific support practice to the corresponding soil loss with up and down slope tillage (Renard et al., 1997). In this study, the P-factor map was derived from the land use/land cover and support factors. The values of P-factor ranges from 0 to 1, in which the highest value is assigned to areas with no conservation practices (deciduous forest); the minimum values correspond to built-up-land and plantation area with strip and contour cropping. The lower the P value, the more effective the conservation practices.

2.4. Image Processing

The Satellite image of the study area was corrected geometrically to remove distortions and subsequently enhanced to improve visual interpretation. This followed by classification into different land use types. Supervised classification was employed because of its high accuracy and the researcher's knowledge of the training areas. Ten coordinates location for each land use class were collected with the aid of GPS during ground thruthing. This was done to aid supervised classification. This is to identify sets of pixels that accurately represent spectral variation present within each information region. The datasets were classified into classes of water body, vegetation, dare land, built-up area and Agriculture. These are adopted from Anderson, Hardy, Roach, and Witner (2001) to suit the study area.

2.5. Techniques of Data Analysis

The stated objectives were achieved through the following:

2.5.1. Estimate Soil Loss in the Catchment Area

This was done using spatial distribution of the rate of soil detachment by rain drop (F) and spatial data layers such as unchanneled and channeled flows (erosion) were input to RUSLE model in ArcGIS 10.3 software environment and predicted annual pixel level soil loss using Equation 1.

$$A = R \times K \times LS \times C \times P \quad \ldots (5)$$

Average annual soil loss of various land use/land cover types was estimated and analyzed to understand causes of erosion in the watershed in context to spatial distribution of erosion factors together with RUSLE Model.

3. Results and Discussion

3.1. Spatial Distribution of Land Morphological Factors

Land morphological factors are those features which are considered to be important determinant factors in soil development and also affect soil erosion of an area. The spatial distribution of land use of Mubi South watershed is show in Figure 2 and Table 3 shows the land covered.

3.2. Land Use, Landcover in the Watershed

Table 3 presents result of land use land cover of the study area. The land uses were categorized based on five Major land uses in the study area, which were: Agricultural land, forestry, water body, built up and bare land.

As shown on Table 3, about 29% of the watershed was covered by agricultural activities, 19% was covered by forest, and 25% was not cultivated and covered by bare land while 17% was covered by built up areas and 10% by water bodies. From the result it was inferred that major land use of the watershed was agricultural activities. However, most to be area was also covered by bare land where no agricultural activities are taking place neither covered by forest.

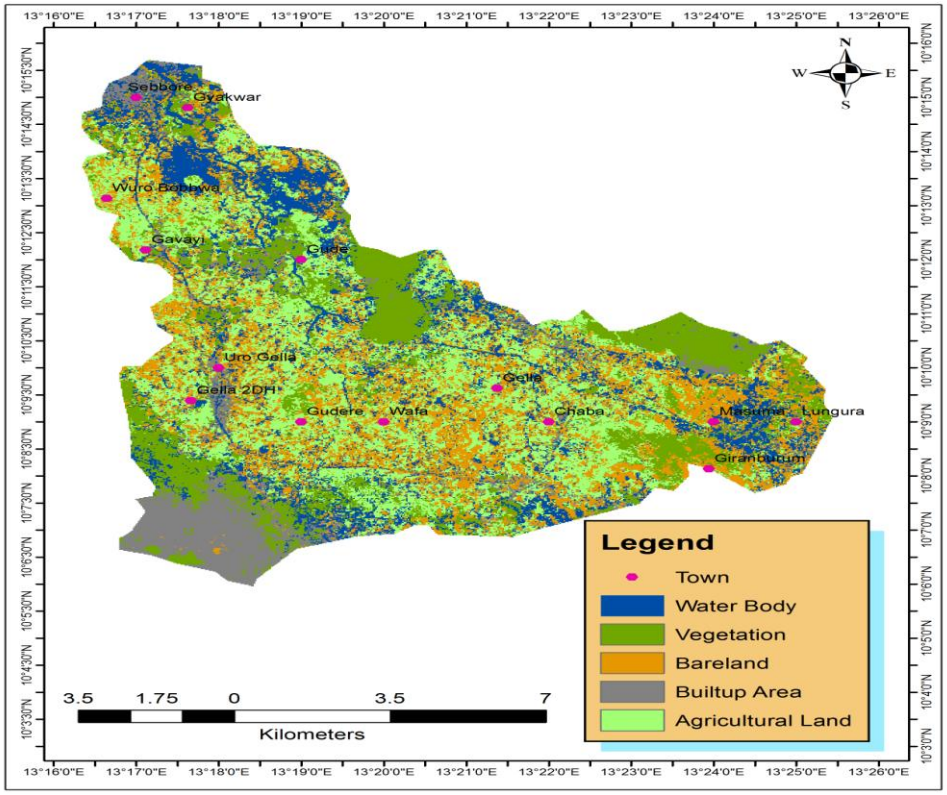
Table 3: *Land use landcover of the study area*

S/No	Land use type	Area Sqkm	Percentage %
1.	Agriculture	42.32	28.6%
2.	Forestry/Vegetation	28.05	19.0%
3.	Water body	14.72	10.0%
4.	Built up Area	25.30	17.1%
5.	Bareland	37.33	25.3%
Total		**147.72**	**100**

Source: Author's Analysis (2016)

3.2.1. Land Uses Land Cover within the Watershed

Figure 2 shown that Mubi South watershed is dominantly covered by Agricultural activities with agricultural land constituting about 28% of the area, followed by bare land with about 25% of the land covered.



Source: Author's Analysis (2016)

Figure 2: *Land used with the watershed*

Figure 2 show that built up area covered about 10% of land use whereas water bodies covered 17% of the land in the study area. Also, forest covered about 19% of the study area. On the basis of the analysis, it can be inferred that Mubi South is a dominant agricultural area with most of the forest cleared and cultivated as on agrarian land (Figure 2). Also, it was observed that there was a higher rainfall in the study area.

3.2.2. The Normalized Difference Vegetation Index (NDVI)

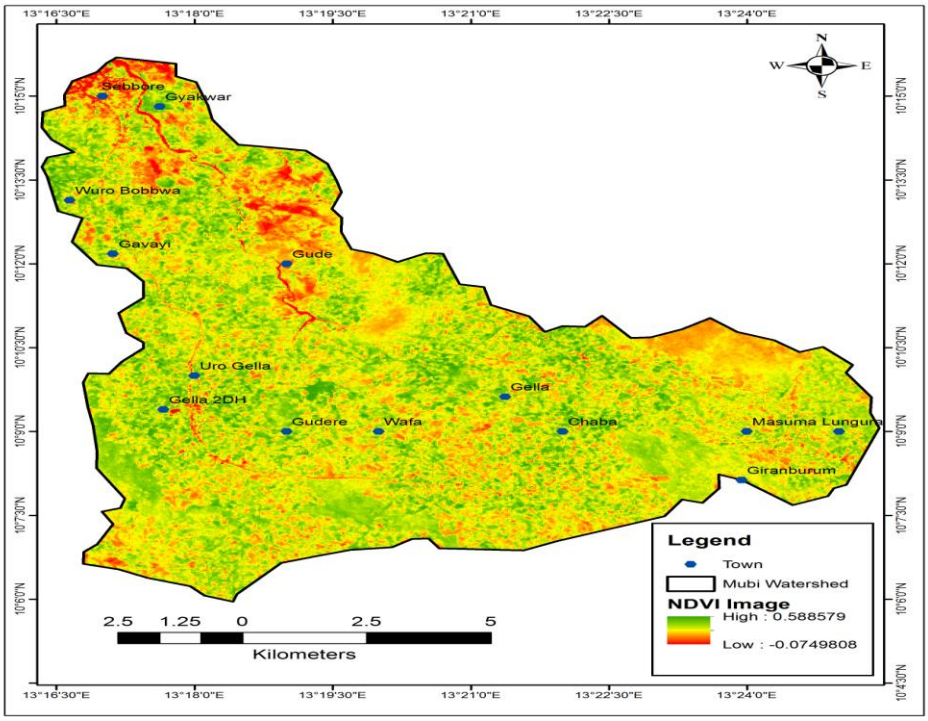
The Normalized Difference Vegetation Index (NDVI) is an indicator of vegetation vigor and health used in determining cover management factor (c), which represents effect of soil disturbing activities, plants, crop sequence and productivity level, soil cover and subsurface biomass on soil erosion risk. Figure 3 shows spatial distribution of NDVI in the study area. It reveals that majority of the watershed area was covered by higher NDVI, which was about 0.58 with a lower value of -0.07. These show that the target area being observed contains live green vegetation at the ground cover. Yellow colour on Figure 4 represents bare soil which reflects moderately in both the red and infrared portion of the electromagnetic spectrum as supported by Holme et al. (1987).

This study noted that the study area has dense vegetation because the Normalized Difference Vegetation Index (NDVI) takes value between -1 and 1, with values 0.5 indicating dense vegetation and value less than 0 indicating no vegetation as shown in Figure 3. Symeonakis and Drake (2004) and Tateishi et al., (2004) studies were compared using satellite imagery to produce maps of vegetation-related variables for soil erosion and found out that the normalized difference vegetation index (NDVI) was the most useful in estimating soil erosion.

3.2.3. Digital Elevation Model

The digital elevation model contains information derived from long-track, 15m ASTER optical data acquired in near infrared bands 3N and 3B. The topographic data used to derive slope and slope aspect are basic to all aspect of land surface. The DEM of Mubi South watershed is shown on Figure 4.

From Figure 4, it was inferred that spatial distribution of DEM in the study area ranges from higher value of about 1261 m lower value of about 570 m respectively. Figure 4 shows that majority of the study area has was within the lower topography of 570 m above sea level and also all settlements of the study area were located within the lower topography with the exception of Giranburum town which is located in higher elevation area. The evidence of higher elevation of about 1261 m is the Mandara Mountain which serves as a border between Nigeria and Cameroon.



Source: Author's Analysis (2016)

Figure 3: *NDVI within the watershed*

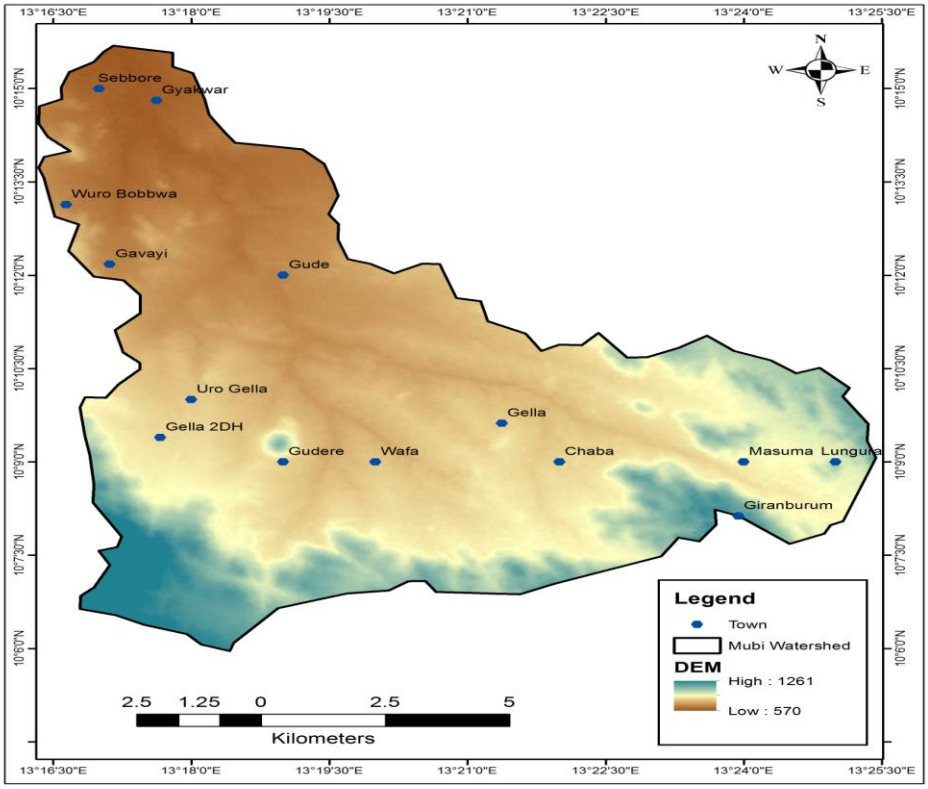
The area covered by high elevation or mountainous area has mixed up vegetation as shown in Figure 2 and this attributed agricultural activity taking place in the area and people from the communities engaged on deforestation in the area for domestic fuel and selling them to solve their financial problems as it were.

3.2.4. Soil of Mubi South

Soil of the Mubi South watershed is shown on Figure 5. The soil map was produced for the soil sample collected from the study area during field work. Laboratory analysis was conducted and the result obtained from the laboratory was used as basis for producing soil map of the study area and compared with FAO (1978) global soil data result. From Figure 5, it was observed that there are three dominant soil classes in the study area which comprised of Luvisols, Regosols and Arenosols. Luvisols constitute about 30% Regosols 35% and Aresonols 35% respectively.

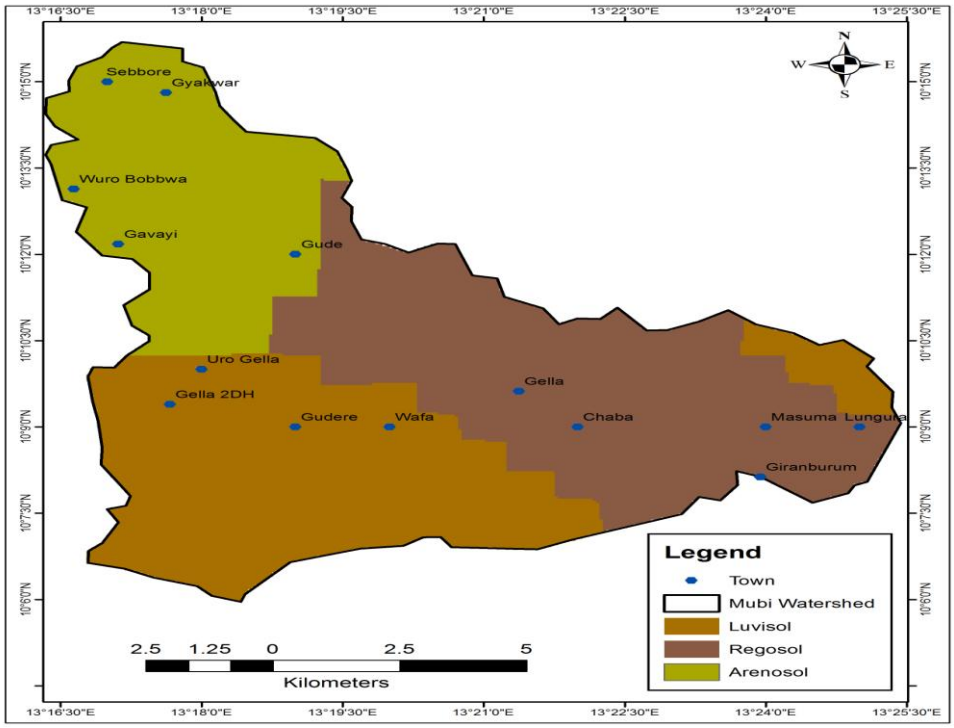
The Luvisol was in pediplain of the watershed, which has about 570 m above sea level as shown in Figure 5, along sabbore, Gyakwa, Wuro Babbwa, Gavayi and some part of Gude town of the study area. Lowland and low vegetation in the study area resulted to the presence of Luvisols in the watershed. Also, Regosol and Arenosol were located between mixed vegetation in the study area and also agricultural areas.

The movement of soil detachability from highland area contributed to the presence of Regosol and Arenoso, because nature of topography, time climate and vegetation types plays a major role on soil formation and Mubi south watershed area is not exception from the factors. Dominant soil groups of the study area are shown in Appendix III as well as their sequence, dominant soil, association and inclusion supported by FAO (2006). The soil group of study area is further divided into soil types.



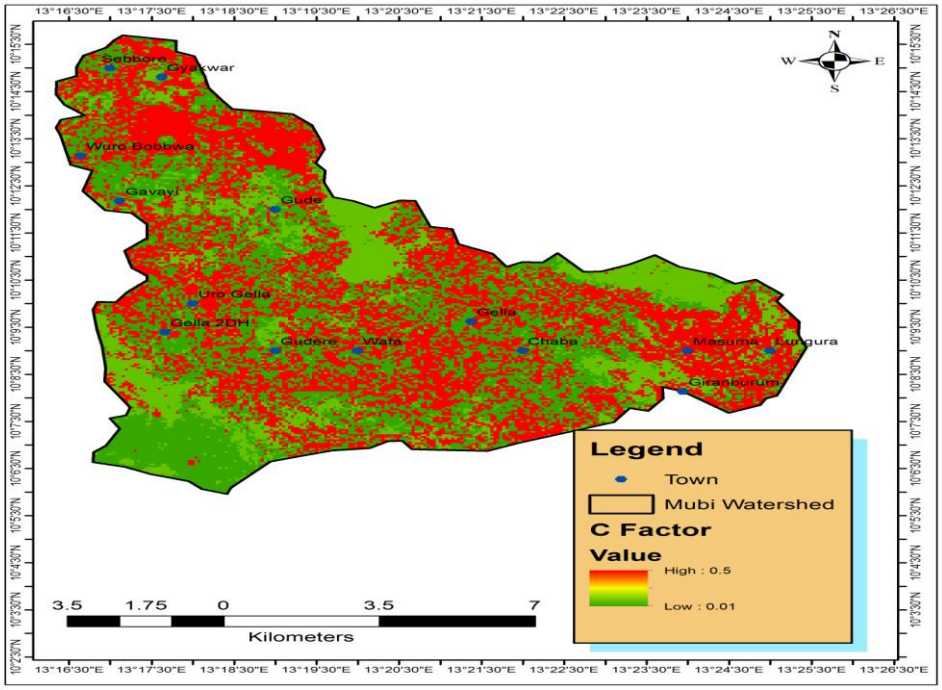
Source: Author's Analysis (2016)

Figure 4: *DEM within the watershed*



Source: Author's Analysis (2016)

Figure 5: *Soil within the watershed*



Source: Author's Analysis (2016)

Figure 6: *Cover Management Factor (C-Factor)*

3.3. Spatial Distribution of Factors used in Assessing Revise Universal Soil Lost Model(RUSLE)

In order to assess spatial distribution of RUSLE, five factors were chosen based on the model which served as input in the equation for assessment of soil loss in the watershed area.

Amongst factors used are: cover management factor (factor), soil erodibility factor (k-factor), slope length and steepness factor (Ls-factor), conservation practice factor (p-factor) and Rainfall erosivity (R-factor). All the factors were derived from land morphological factors; except rainfall factor. That asserts the need for assessment of morphological factors for watershed of the study area for proper analysis and assessment of soil erosion risk in the study area. The factor maps result obtained from analysis is presented in Figures 6 to Figure 10:

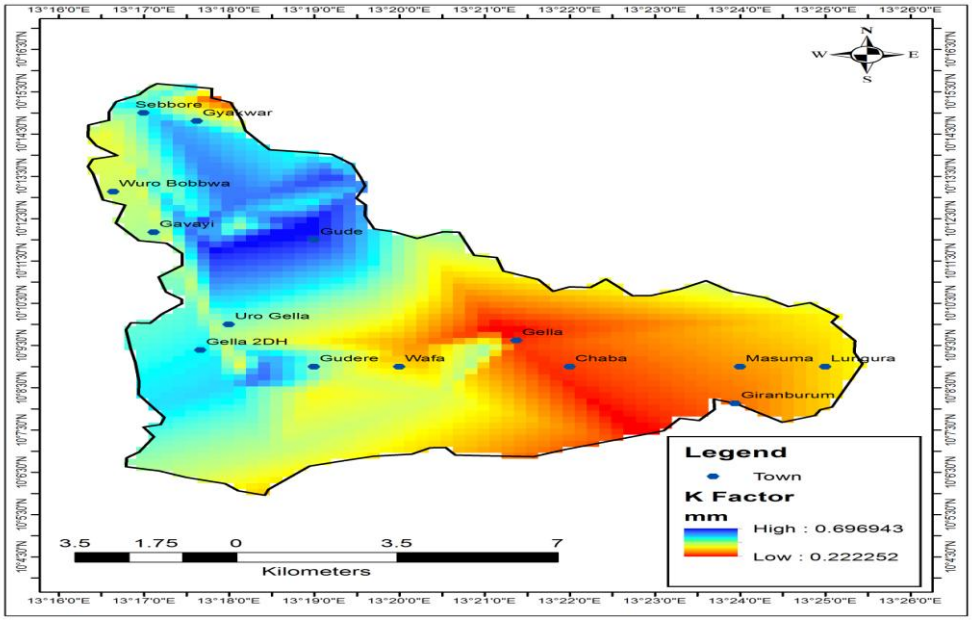
3.3.1. Cover Management Factor

The result of cover management factor (c-factor) is presented in Figure 6 to show spatial distribution of cover management factor that range from higher value of 0.5 to lower value of 0.01 in the watershed.

Result of Figure 6 was due to land use covered within watershed which shows that most of the vegetation and agricultural activities were around the river. There is an intersection of low and high cover management practices in the study area. High cover management factor was observed around mountainous area of the watershed. This finding shows that land use slope and hillshade plays vital role in determining cover management practices in an area.

3.3.2. Soil Erodibility Factor

Result of soil erodibility of the watershed is shown in Figure 7. The soil erodibility factor (k-factor) measures susceptibility of soil particle to detachment and transport by rain or runoff (Goldman et al., 1986; Mitchell and Bubenzer, 1980).



Source: Author's Analysis (2016)

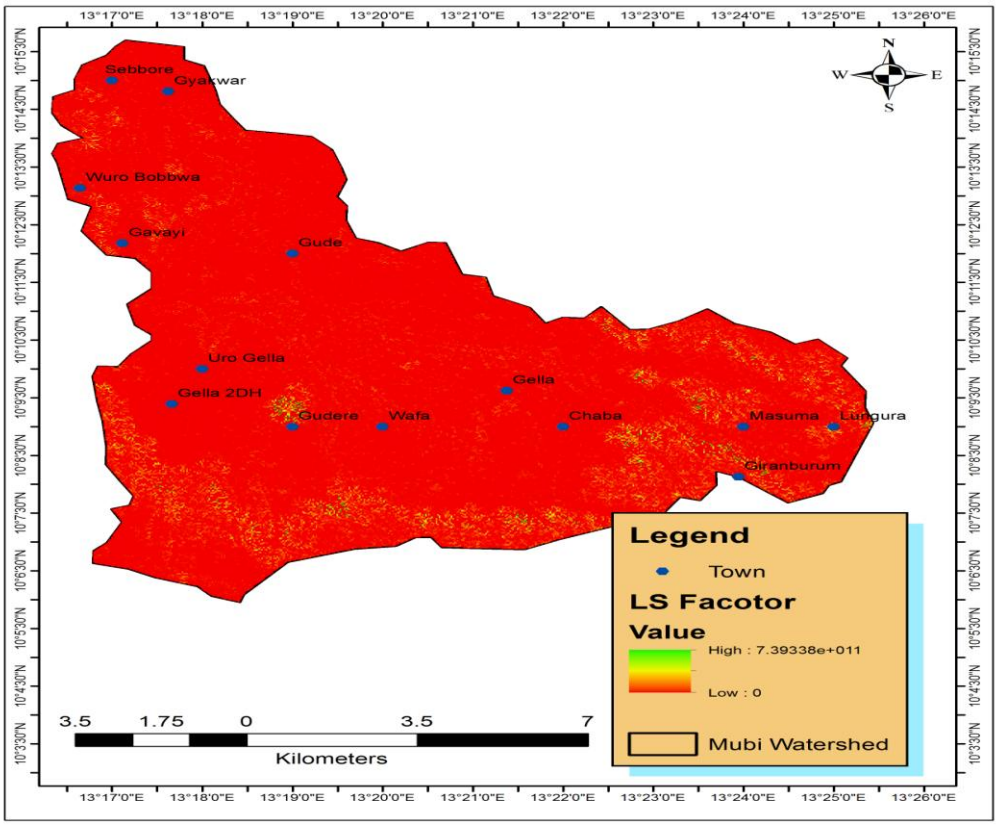
Figure 7: *Soil Erodibility Factor (K-Factor)*

From Figure 7, it was inferred that soil erodibility factor reveals that different soils erode at different rates. For example, comparing this result with soil map on Figure 5, it was observed rate of soil erodability in heigher (0.69) around areas having Arenosol. Also, Arenosol had highest clay content and lower silt content in the study area. Luvisol had moderate (0.4) erodibility factor and characterized with moderately clay and silt content and sandy soil deposit.

The findings of soil erodibility factor range in value from 0.02 to 0.69 and support Goldman et al. (1986) and Mitchell and Bubenzer (1980). This result shows that there was lower permeability in Southeast; moderately in Southwest and higher at the Northern part of the watershed. This was as result of soil type, landuse and hillshed of the area which show vidence of resistance of soil to detachment by rainfall impact and surface flow. Research Toy et al. (2002) shows that soil with larger sand and silt properties are more vulnerable to water erosion due to lack of stability of the soil particle. Soil texture with large particles are resistant to transport because of the greater force required to entrain them and fine particles are resistant to detachment because of their cohesiveness. The least resistant particles are silts and fine sands. Thus, soils with silt content above 40 per cent are highly erodible (Richter and Negendank, 1977).

3.3.3. Slope Length and Steepness (Ls-factor)

Slope length and steepness factor represent effect of slope length on soil erosion. The slope length is the ratio of soil loss from field slope length and soil loss increase more rapidly with slope steepness than it does with slope length. The Ls-factor result of the area is presented on Figure 8.



Source: Author's Analysis (2016)

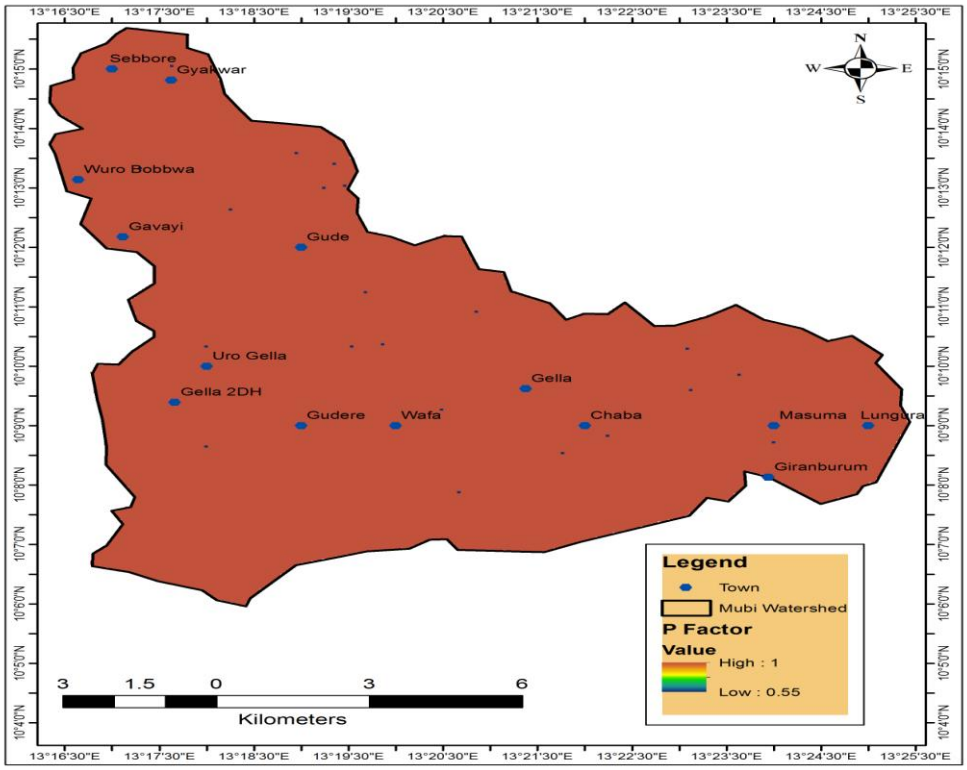
Figure 8: *Slope Length and Steepness (LS-Factor)*

From Figure 8, the result of LS-factor shows even distribution of slope steepness along each watershed. Research by Tay et al. (2002) shows susceptibility of soil to water erosion depends on soil length and is more prevalent in sloping area (Angima et al., 2003). Also, the result of these studies suggests a curvilinear relationship between soil loss and slope steepness, with erosion initially increasing rapidly as slope increases from gentle to moderate, reaching a maximum on slopes of about 7° and then decreasing with further increases in slope. Such a relationship would apply only to erosion by rain splash/sheet and surface runoff. It would not apply to landslides, piping or gully erosion by pipe collapse.

Again, studies of Toy et al. (2002) show that slope length has effect on soil loss for steep slope. Also reported that greater sensitive of slope had effects on soil loss due to differences in rainfall. Areas having about 7.39m length of cell slope length and steepness in the watershed as show on Figure 8, will have greater soil loss as supported by Toy et al. (2002), than those areas having 3m and 0m length of cell slope length respectively.

3.3.4. Conservation Practice Factor (P-factor) Within the Watershed

Conservation practice (p-factor) is the support practice factor. Reflect effect of practices that will reduced amount and rate of water runoff; hence reduce the amount of erosion (Renard et al., 1997).



Source: Author's Analysis (2016)

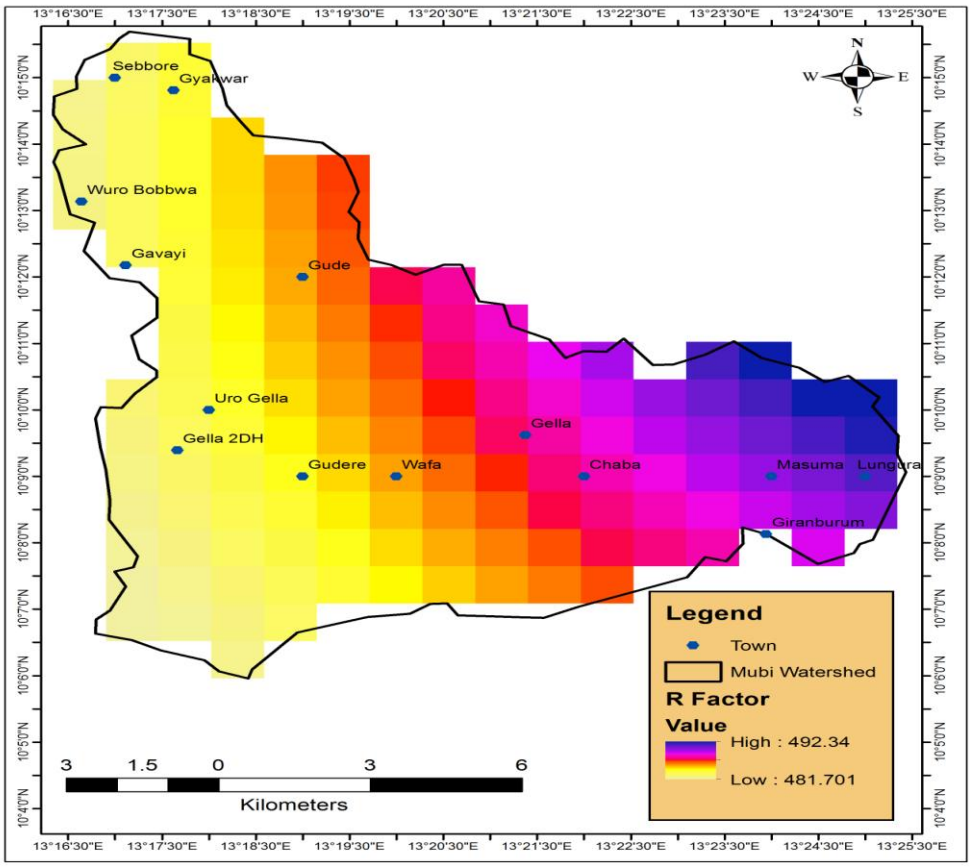
Figure 9: *Conservation Practice Factor (P-factor)*

As shown in Figure 9, areas which shows the spatial distribution value of 1 from the legend show no conservation practice (deciduous forest) while the minimum value of 0.55 corresponds to built-up with strip and contour cropping as supported by Renard et al. (1997). The lower the p-value (0.55) as in the watershed, the more effective is conservation practice in the study area. Also, result of the practice

correspond to nature and land use landcover map of the watershed. The result also proves that practice factors supported in the watershed was contouring and contour strip cropping with the value ranges between (0.350 - 0.600).

3.3.5. Rainfall Erosivity (R-factor)

Rainfall erosivity is a measure of the erosive force of a specific rainfall. Rainfall is the main climatic characteristics that influence soil erosion; given the extraordinary importance of soil detachment process due to drop impact and runoff shear (Wischmeier and Smitt, 1978). Rainfall erosivity map of the watershed is shown on Figure 10.



Source: Author's Analysis (2016)

Figure 10: *R-Factor*

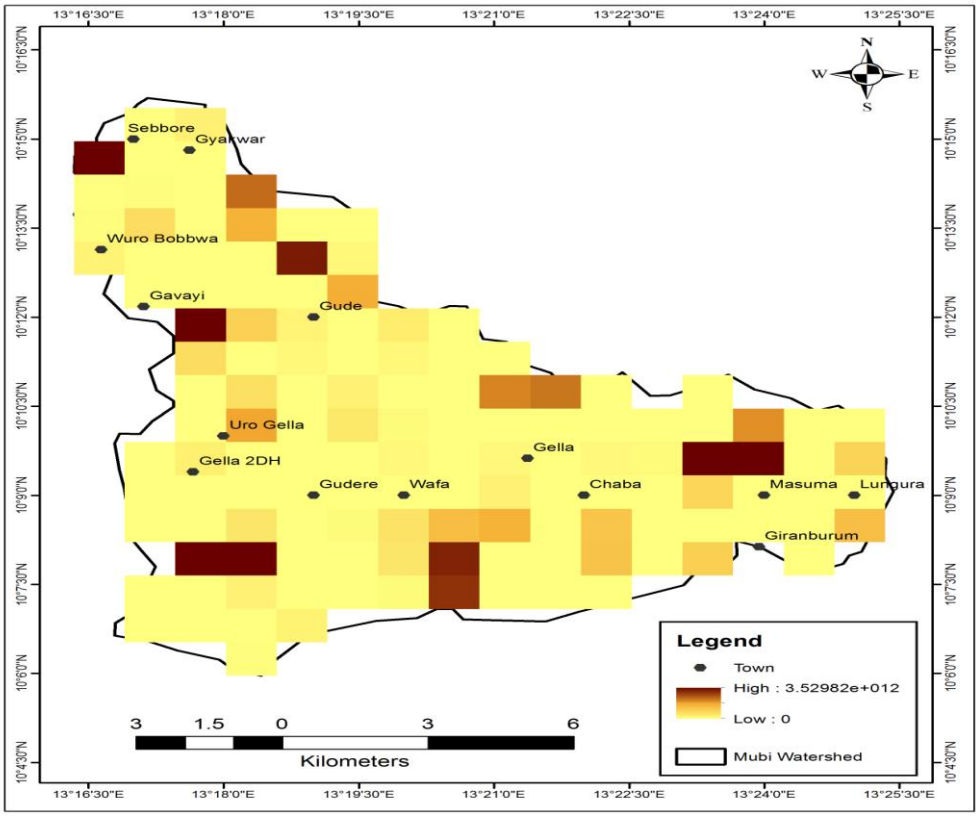
As shown in Figure 10, spatial distribution of rainfall map of the watershed indicates that it ranges between 481m to 192m. It also shows that the rainfall is higher towards northwest of the watershed and the area of higher rainfall were in the mountainous area of the watershed along Masuma and Lumgura which shares boundary with the Republic of Cameroon.

The spatial distribution of five factors map produced were used as input for derivating GIS-based RUSLE and also served as input for some of the RMMF model. These analyses explored the importance of GIS and remote sensing in integrating both spatial and non-spatial data/information; hence the needs for the technique in assessing and analyzing spatial distribution of soil erosion data. Results obtained were used as input in estimating rate of soil erosion in objective four of this study.

3.4. Estimated Spatial Distribution of Soil Loss Risk in the Watershed Area

The result of the estimated spatial distribution of soil loss in the study area is presented in Figure 11.

Figure 11 indicates that soil erosion loss in the watershed area was 3.5 t h^{-1} yr^{-1} as at the time of this research. It was observed that higher soil loss was obtained were crop cover management is high; with about 4.5 rate of potential evapotranspiration and between 0.08 to 0.24 high soil moisture. It was also observed that soil loss was high between 15.8 mm to 15.7 mm high to medium occurrences of daily rainfall.



Source: Author's Analysis (2016)

Figure 11: *Estimated soil erosion using RUSLE*

Also, it was noticed that area with high soil loss was observed within low soil porosity area, except around Sebbore which was located in an area of high porosity with lower slope steepness length mostly having low to moderate silt content (52.8 to 223.9) with 1.4 Mgm^{-3} to 1.6 Mgm^{-3} bulk density.

Contributions to Knowledge

1. The outcome of this research provides blueprint showing that estimated spatial distribution of soil loss in Mubi South indicated high soil erosion loss of 3.5 t h^{-1} yr^{-1} in the watershed.
2. Results of this study serve as a document that can help town planners, road construction planners, engineers and site constructors to design urban land use plan that are less vulnerable to erosion.

4. Conclusion

Studies had shown that recent global land degradation caused by increase in soil erosion risk lead to land degradation and in Mubi South Watershed and Nigeria are not exceptions of degradation. Remote sensing data and GIS successfully enabled rapid, as well as detailed assessment of estimating soil erosion and show spatial distributions of soil erosion related factors and features.

Recommendations

Reduced on-site impacts of soil erosion: Vegetative buffers (trees, understory, and ground cover) combat soil erosion as they protect the soil from erosion processes, allow greater infiltration of water and trap sediment entering from cultivated areas. Appropriately designed windbreaks can also significantly reduce soil loss from fields by reducing wind velocity as recommended by Salah et al. (2008).

Recommendation for Further Research

This present research did not measure the area coverage of individual soil erosion types; hence, the need for further research to measure the magnitude of all forms of soil erosion.

It is also recommended for research like this to be conducted in Mubi North Local Government because it was confirmed during field work that the area is also experiencing soil erosion and there are the needs to know the areas vulnerable to soil erosion risk for proper conservation practices, planning and management.

References

Adebayo, A.A. and Tukur, A.L. 1999. *Adamawa State in Maps*. 1st ed., Yola: Paraclete Publishers, p. 8.

Adebayo, A.A. 2004. *Mubi Region: A geographical synthesis*. 1st ed. Yola-Nigeria: Paraclete Publishers, pp.32-38.

Aduayi, E.A., Chude, V.O., Adebusuyi, B.A. and Olayiwola, S.O. 2002. Fertilizer use and Management Practices for Crops in Nigeria. 3rd ed. Federal Fertilizer Department, Federal Ministry of Agriculture and Rural Development, Abuja, Nigeria, Garko International Limited.

Aksoy, E., Ozsoy, G. and Dirim M.S. 2009. Soil mapping approach in GIS using landsat satellite imagery and DEM data. *African Journal of Agricultural Research*, 4(11), pp.1295-1302.

Anderson, J.R., Hardy, E.E., Roach, J.T. and Witner, R.E. 2001. Landuse and land cover classification system for use with remote sensor data. Geological Survey Professional Paper 964, United State Geological Survey Circular 671. United State Government Printing Office, Washington.

Angima, S., Stott, D. E, O'neill, M., Ong, C. and Weesies, G. 2003. Soil Erosion Prediction using RUSLE for Central Kenyan highland conditions. *Agriculture, Ecosystems & Environment*, 97(1), pp.295-308.

Arnoldus, H.M.J. 1980. An approximation of the rainfall factor in the Universal Soil Loss Equation. In: De Boodt, M. and Gabriels, D. (eds.). *Assessment of Erosion*. UK: Wiley, Chichester, pp.127-132.

Asdak, C. 2009. *Hydrology and Watershed Management*. Yogyakarta: UGM Press.

Auerswald, K. 1993. Bodeneigenschaften und Bodenerosion. Wirkungswege beiunterschiedlichen Betrachtungsmaßstaben. Borntr¨ager, Berlin, p.208.

Bonilla, C.A., Reyes, J.L. and Magri, A. 2010. Water erosion prediction using the Revised Universal Soil Loss Equation (RUSLE) in a GIS framework, central Chile. *Chilean Journal of Agricultural Research*, 70(1), pp.159-169.

Braimoh, A. and Vlek, P. 2008. *Land use and soil resources*. Springer Verlag.

FAO. 1978. Report on the Agro-Ecological Zones Project, World Soil Resources Report. FAO, Rome. Methodology and Results for Africa, 1, p.48.

FAO. 2006. World Reference Base (WRB) for Soil Resources. A framework for international classification and communication. World Soil Resources Report.

Federal Republic of Nigeria. 2007. Population Census 2006 Official Gazette. Abuja. Nigeria.

Goldman, S.J., Jackson, K. and Bursztynsky, T.A. 1986. Erosion and Sediment Control Handbook. New York: McGraw Hill Book Co.

Hacısalihoglu, S., Oktan, E. and Yucesan, Z. 2010. Predicting Soil Erosion in Oriental Spruce (Picea orientalis (L.) Link.) Stands in Eastern Black Sea Region of Turkey. *African Journal of Agricultural Research*, 5(16), pp.2200-2214.

Holme, A. McR., Burside, D.G. and Mitchell, A.A. 1987. The development of a system for monitoring trend in range condition in the Arid Shrublands of Western Australia. *Australia Rangland Journal*, (9), pp.14-20.

Karydas, C.G., Sekuloska, T. and Silleos, G.N. 2009. Quantification and site-specification of the support practice factor when mapping soil erosion risk associated with olive plantations in the Mediterranean island of Crete. *Environmental Monitoring and Assessment*, 149(1-4), pp.18-28.

Kefi, M., Yoshino, K., Zayani, K. and Isoda, H. 2009. Estimation of soil loss by using combination of erosion model and GIS: case of study watersheds in Tunisia. *Journal of Arid Land Studies*, 19(1), pp.287-290.

Kefi, M. and Yoshino, K. 2010. Evaluation of the economic effects of soil erosion risk on agricultural productivity using remote sensing: case of watershed in Tunisia. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Kyoto Japan.

Kouli, M., Soupios, P. and Vallianatos, F. 2009. Soil erosion prediction using the Revised Universal Soil Loss Equation (RUSLE) in a GIS framework, Chania, Northwestern Crete, Greece. Environmental Geology, 57, pp.483-497.

Morgan, R.P.C. *Soil Erosion*. London: John Wiley and Sons Ltd., pp.17-62.

Moore, I.D., Burch, G.J. and Mackenzie, D.H. 1988. Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. *Transactions of the American Society of Agricultural Engineers*, 34, pp.1098-1107.

Moore, I.D. and Burch, G.J. 1986a. Physical basis of the length slope factor in the universal soil loss equation. *Soil Science Society of America*, 50(5), pp.1294-1298.

Nwaka, G.C., Alhassan, A.B. and Kunduri, A.M. 1999. A study of soils derived from Basalt in North Eastern, Nigeria 11. Physico-chemical characteristics and fertility status. *Journal of Arid Agriculture*, 9, pp.89-98.

Olofin, E.A. 1994. The application of SLEMSA in Estimating Soil Erosion and issues of Productivity in the drylands of Nigeria. 37th Annual Conference of the NGA, Ikere, Ekiti.

Pandey, A., Mathur, A., Mishra, S.K. and Mal, B.C., 2009. Soil erosion modeling of a Himalayan watershed using RS and GIS. *Environmental Earth Sciences*, 59(2), pp.399-410.

Park, S., Oh, C., Jeon, S., Jung, H. and Choi, C. 2011. Soil erosion risk in Korean watersheds, assessed using the revised universal soil loss equation. *Journal of Hydrology*, 399 (3-4), pp.263-273.

Prasannakumar, V., Shiny, R., Geetha, N. and Vijith, H. 2011a. Spatial prediction of soil erosion risk by remote sensing, GIS and RUSLE approach: a case study of Siruvani river watershed in Attapady valley, Kerala, India. *Environmental Earth Sciences*, 64, pp.965-972.

Prasannakumar, V., Vijith, H., Geetha, N. and Shiny, R. 2011b. Regional scale erosion assessment of a sub-tropical highland segment in the Western Ghats of Kerala, South India. Water Resources Management, 25, pp.3715-3727.

Renard, K.G., Foster, G.R., Weesies, G.A., McCool, D.K. and Yoder, D.C. 1997. Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE). US Department of Agriculture, Washington, DC, 703, pp.1-251.

Richter, G. and Negendank, J.F.W. 1977. Soil erosion processes and their measurement in the German area of the Moselle river. *Earth Surface Processes*, 2, pp.261-278.

Salah, D., Christopher, W., Gretel, G., Erika, S. and Julienne, R. 2008. Watershed management approaches, policies, and operations: Lessons for Scaling Up The World Bank, Washington, DC, Water Sector Board Discussion Paper Series, Paper No.11 May 2008, p.63.

Symeonakis, E. and Drake, N. 2004. Monitoring desertification and land degradation over sub-Saharan Africa. *International Journal of Remote Sensing*, 25(3), pp.573-592.

Tateishi, R., Shimazaki, Y. and Gunin, P.D. 2004. Spectral and temporal linear mixing model for vegetation classification. *International Journal of Remote Sensing*, 25(20), pp.4203-4218.

Tay, J.T., George, R.F. and Kenneth, G.R. 2002. Soil Erosion. New York: John Wiley and Sons Inc.

Tekwa, I. J. and Usman, B. H. (2006). Estimation of soil loss by gully erosion in Mubi Adamawa State, Nigeria. *Journal of the Environment*, 1: 35-43

Tekwa, I.J., Laflen, J.M. and Yusuf, Z. 2014. Estimation of monthly soil loss from ephemeral gully erosion features in Mubi, Semi-arid Northeastern Nigeria. *Agricultural Science Research Journal*, 4(3), pp.51-58.

Tesfahunegn, G.B., Tamene, L. and Vlek, P.L.G. 2011b. Catchmentscale spatial variability of soil properties and implications on site-specific soil management in northern Ethiopia. *Soil and Tillage Research*, 117, pp.124-139.

Tian, Y.C., Zhou, Y.M., Wu, B.F. and Zhou, W.F. 2009. Risk assessment of water soil erosion in upper basin of miyun reservoir. *Environmental Geology*, 57, pp.937-942.

Toy, T.J., Foster, G.R. and Renard, K.G. 2002. *Soil erosion: processes, prediction, measurement and control.* New York, USA: John Wiley and Sons, p.338.

USDA. 1978. Predicting rainfall erosion losses. A Guide to Conservation Planning, Washington DC.

Verheijen, F., Jones, R., Rickson, R. and Smith, C. 2009. Tolerable versus actual soil erosion rates in Europe. *Earth Science Review*, 94, pp.23-38.

Wischmeier, W.H. and Smith, D.D. (1978). Predicting rainfall erosion losses: a guide to conservation planning. Agriculture Handbook No. 537. US Department of Agriculture Science and Education Administration, Washington, DC, USA, p.163.

World Bank. 1990. Towards the development of an environmental action plan for Nigeria, West Africa Department.

Zhou, P., Luukkanen, O., Tokola, T. and Nieminen, J. 2009. Effect of vegetation cover on soil erosion in a mountainous watershed. CATENA, 75(3), pp.319-325.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2551-2559
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.346

Research Article

Runoff and Sediment Yield Prediction Using Agriculture Non-Point Source (AGNPS) Model in Ata-Gad Watershed, Uttarakhand, India

Deepa Naik[1], Pramod Kumar[2], Aniruddha Deshmukh[2]

[1]Department of Geography, S.P. University of Pune – 411 007, Maharashtra, India
[2]Indian Institute of Remote Sensing, Dehradun – 248 001, Uttrakhand, India

Publication Date: 7 April 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.346

Abstract The present study was undertaken to predict the runoff and sediment loss from Ata-gad watershed, Chamoli district, Uttarakhand, India. The land use/land cover (LULC) map was prepared using IRS-P6 LISS-III data. Digital Elevation Model (DEM) from ASTER and soil information from Soil and Land Use Survey of India (SLUSI) was used for runoff and sediment yield prediction. It was observed that large part of the watershed is forested (71.9%) and agricultural activity is ongoing in lower reaches of the valley (18%). The watershed area is mostly under moderately steep (15-35%) to very steep slope (50-75%). LULC, Soil, DEM and other inputs were fed into Agriculture Non-point Source (AGNPS) model through AGNPS Data Generator (ADGen) interface of image processing software. The AGNPS model helps to visualize the effect of slope, rainfall, LULC, etc. on runoff and sedimentation characteristics of a watershed. It was observed that nearly fifty percent area of the watershed produced 2.54 cm of runoff corresponding to 17.8 cm of rainfall. As large part of the watershed is under forest and consequently 64.24% of its area produced less than 1.42 cumec and only 0.11% of the area showed more than 49.55 cumec of peak runoff. Twenty-one percent area of the watershed is having steep slope (slope>75%) and showed the maximum rate of erosion as 48.67 tons/ha. Erosional characteristics vis-à-vis other properties of the landscape were also analyzed. It was also observed that with the increase in slope, though the soil erosion has increased but the slope factor solely does not affect erosional characteristics.
Keywords *AGNPS; GIS; Remote sensing; Runoff; Soil erosion*

1. Introduction

The hydrological behaviour of a catchment is a complex phenomenon, which is controlled by large number of climatic and physiographic factors that vary in time and space. The models are required not only to predict water yield and subsequently to build design parameters of hydraulic structures, but also for understanding and to evaluate the anthropogenic and disaster-induced effects on the hydrological regime of a river basin.

The Himalayan regions are adversely affected with erosional processes due to high elevation differences, denuding forest cover, varying climatic conditions, agriculture practices and inhabitation, etc. in comparison to other mountainous regions. The hilly region is also inhabited by livestock population which results in overgrazing and that induces soil erosion. The rainfall is a major triggering

factor which enhances the chances of water erosion. By adopting appropriate soil conservation practice and proper water management, this problem could be tackled to a great extent.

The hydrological models help to identify the cause and sink areas of sediment transport, runoff and nutrients that leave its original place. The Agricultural Non-point Source Pollution Model is a event based distributed computer simulation model developed by Agriculture Research Service, United States Dept. of Agriculture with the assistance of National Resource Conservation Service and the Minnesota Pollution Control Agency (Bosch et al., 1998)

Many GIS based AGNPS interfaces are available for the preparation of input parameters and to run the model. Among them, the AGNPS Data Generator (ADGen) and ERDAS Interface and Map Window Interface (MWAGNPS) are most popular and widely used. Grunwald and Norton (1999) compared surface runoff and sediment yield using AGNPS water quality simulation model for 52 rainfall-runoff events, 22 for calibration and 30 for validation for two small watersheds in Bavaria, Germany. Ma and Bartholic (2003) used AGNPS model used in combination with GIS tools to assess the feasibility of water quality effluent trading for phosphorus in Morrow Lake sub-watershed, Kalamazoo, Michigan, USA. Chowdary et al. (2001) have studied AGNPS for quantitative assessment of nonpoint source pollution within the Karso watershed, Damodar river valley, Hazaribagh district of Bihar state, India. Zema et al. (2012) used AnnAGNPS model to assess runoff water amount and quality as well as sediment yield in small to large monitored watersheds in different climatic and geomorphologic conditions. Jianchang and Luoping (2008) used AGNPS model for Wuchuan catchment in Fujian Province, China for ten storms. Haregeweyn et al. (2003) evaluated the AGNPS model on Augucho catchment in western Hararghe region of Ethiopia using observed data of 8–10 years. AnnAGNPS did not simulate base-flow, hence to compare the model predicted runoff to observed runoff; base-flow was separated from the observed runoff using the straight-line method (Sarangi et al., 2007). Najim et al. (2006) tested the suitability of AGNPS pollution model for a mixed forested watershed. The simulated runoff volume reasonably matched with the observed runoff volume, with coefficient of performance of 0.09. Rainis (2004) compared the effects of slope information derived from three sources on sediment yield estimated using AGNPS model.

1.1. Study Area

The present study has been carried out for the Simli watershed which is also known as Ata-Gad watershed and falls within Pindar Catchment in Bageshwar district of Uttarakhand province in India. The geo-bounds of the study are 30º05' N to 30º16' N longitude and from 79º10' E to 79º17' E latitude. Agriculture is the prime occupation of watershed inhabitants. Paddy and Millets are the major *Kharif* crops, which are practiced during rainy season whereas Wheat and Mustard are grown as winter or *Rabi* crops. The river at its initial course flows through sedimentary rock and further to the south meanders through quartz, schist and granite found in abundance in the study area. The drainage pattern is dominantly dendritic. The nature and type of soil found in the study area varies from place to place and along with it changes the vegetation it supports. There is rich humus present in the thin soil layers.

2. Data Used and Methodology

There are four important parameters which are required to run the AGNPS model. These are DEM, hydrological soil group, watershed boundary and land use/land cover (LULC). These maps should have same spatial resolution (30 m in the present study). The hydrological soil group map has been prepared using soil map and later, attributes were added to determine soil erodibility factor (K factor) using nomograph method. LULC map was prepared using supervised classification method and later

contextual refinement was conducted to improve the results. Some attributes have also been attached to LULC map like SCS CN for Antecedent Moisture Condition-II (AMC), Manning's roughness factor (N), cropping factor, and surface condition constant. Figure 2 explains the methodology followed in the present study.

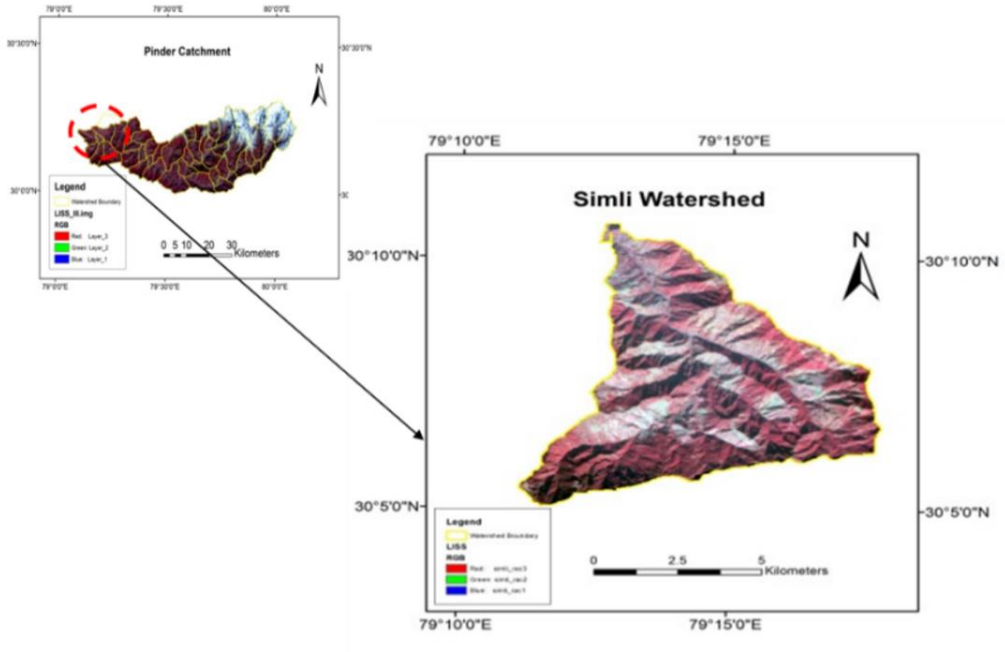


Figure 1: *Location map of Simli watershed within catchment of Pinder river*

In the present study, IRS-P6 LISS-III acquired on 26 December 2007 and ASTER Digital Elevation Model (DEM) has been used. The Survey of India (SOI) topographical map at 1:50,000 scale has been used for preparing the base map for the study area. The existing soil map, rainfall data (event-specific), SCS CN table has been used for running the model. The ADGen (AGNPS data generator) interface has been used for estimating runoff and soil loss. The Arc Hydro tool has been used to derive several data sets that collectively describe the drainage and topographic properties of a watershed. The raster analysis is performed to generate data on flow direction, flow accumulation, stream definition, stream segmentation and watershed delineation. These data are later used to develop a vector representation of catchments and drainage lines. Using these information, the geometric network was constructed. Later, the Arc Hydro data model has been used for micro-watershed delineation.

2.1. Model Description

The model uses Natural Resources Conservation Service (NRCS) Soil Conservation Service (SCS) Curve Number (CN) method to estimate runoff which uses the following equation:

$$Q = (P-0.2S)^2/(P+0.8S) \quad \ldots\ldots (1)$$

Where, Q is runoff depth (mm), P is rainfall (mm), and S is retention parameter (mm) which is defined as:

$$S = (1000/CN) - 10 \quad \ldots\ldots (2)$$

The erosion is estimated using a modified version of Universal Soil Loss Equation (Wischmeier and Smith, 1978)

$$SL = (EI)\ K\ LS\ C\ P \ldots\ldots (3)$$

Where, SL is soil loss, EI is product of storm kinetic energy and maximum 30 minute Intensity, K is soil erodibility factor, LS is topographic factor, C is cover management factor and P is conservation practice factor. The peak flow is estimated using empirical relationship developed for CREAM model by Smith and Williams (1980) as given below:

$$Q_p = 3.79\ (A)^{0.7}\ (CS)^{0.16}\ (R/25.4)^{0.9A\hat{}0.02}\ (L^2/A)^{-0.19} \ldots\ldots (4)$$

Where, Q_p is peak runoff rate (m^3/s), A is watershed area (km^2), CS is channel slope (m/km), R is runoff volume (mm), L is the watershed length (km).

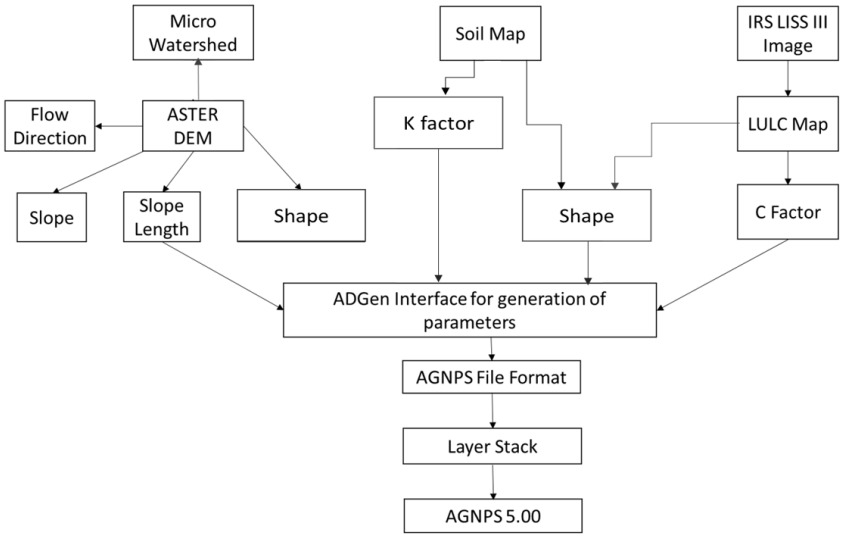


Figure 2: *Methodology flow chart*

The AGNPS model has following limitations: a) all runoff and associated sediment, nutrient, and pesticide loads for a single day are routed to the watershed outlet before the next day simulation begins (regardless of how many days this may actually take), b) there are no mass balance calculations tracking inflow and outflow of water, c) there is no tracking of nutrients and pesticides attached to sediment deposited in stream reaches from one day to the next, and d) point sources are limited to constant loading rates (water and nutrients) for entire simulation period, and there is no allowance for spatially variable rainfall.

Rainfall at antecedent periods of 5-30 or more days prior to the storm are commonly used as indices of watershed wetness. These are only rough approximations, as they do not include the effect of evapotranspiration and infiltration on watershed wetness. The ADGen is a 22 parameter driven interface that facilitates compatible database generation for AGNPS model. Table 1 shows the utility of the parameters and the sources from where these are derived. The AGNPS Model generates a *.nps file after successful execution of the input data. The ADGen interface provides a useful tool to convert the *.nps files to ERDAS Imagine *.img format such that user can visualise the output results.

3. Results and Discussion

The AGNPS model estimates runoff, soil loss and nutrients on a cell-to-cell basis for entire watershed. It generates the runoff map for various hydrologic-soil-land use/land cover complexes and estimates runoff and soil loss corresponding to various rainfall events. As the area is mountainous and receives high amount of rainfall, the AMC of soil is important as it gets saturated easily and immediately produces surface runoff. The major LULC units in the Ata-Gad watershed are classified as forest, agriculture, barren land, scrub, settlement, snow cover and water body. It was observed that large part of the watershed is forested (61.9%) and agricultural activity is ongoing in lower reaches of the valley (18%). The watershed area is mostly under moderately steep (15-35%) to very steep slope (50-75%) (Table 2). During the peak runoff analysis, it was observed that 64.24% of watershed area produced less than 1.42 cumec and only 0.11% of the area showed more than 49.55 cumec of peak runoff (Table 3). Other parts of the watershed showed the runoff in the range of 1.42-49.55 cumec. One of the important parameter that affects runoff and soil erosion is slope. 1196 ha area of the watershed is having moderate to steep slope (slope>75%) and showed the maximum rate of erosion as 48.67 tons/ha. The surface runoff for Ata-Gad watershed has been assessed using AGNPS model. Figure 3a shows the runoff contributing area on cell-to-cell basis. Figure 3b shows the overland runoff from the cell. It is observed that maximum runoff is observed from agricultural region and it involves the transport of agricultural chemical via surface runoff which could be a threat to downstream ecosystems. The runoff from scrub, open forest and dense forest areas are comparatively lesser. The figure 3c shows upstream concentrated flow. As AGNPS gives information on cell-to-cell basis, thereby it shows the upstream concentrated flow that is received by each pixel from upstream. Figure 3d shows the upstream accumulated runoff computed on a cell to cell basis. Figure 3e shows downstream accumulated runoff i.e. runoff emanating from each pixel. Figure 3f shows the downstream concentrated flow. All these figures help to understand hydrological processes, human influence activities related to land use and natural impacts related to climate. Figure 3g depicts upstream and downstream sediment yield generated at various parts of watershed. The agriculture fields are contributing higher soil erosion followed by the scrub and then forest land due to poor vegetation characteristics and over grazing practices.

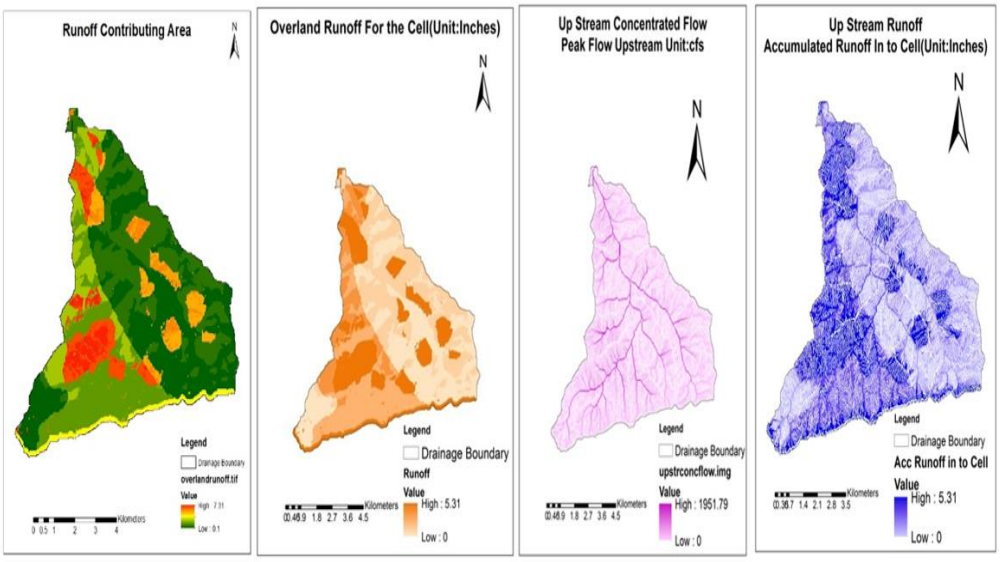


Figure 3: a. *Runoff contributing area, b. Overland runoff from the cell, c. Stream contributing area, d. Upstream accumulated runoff*

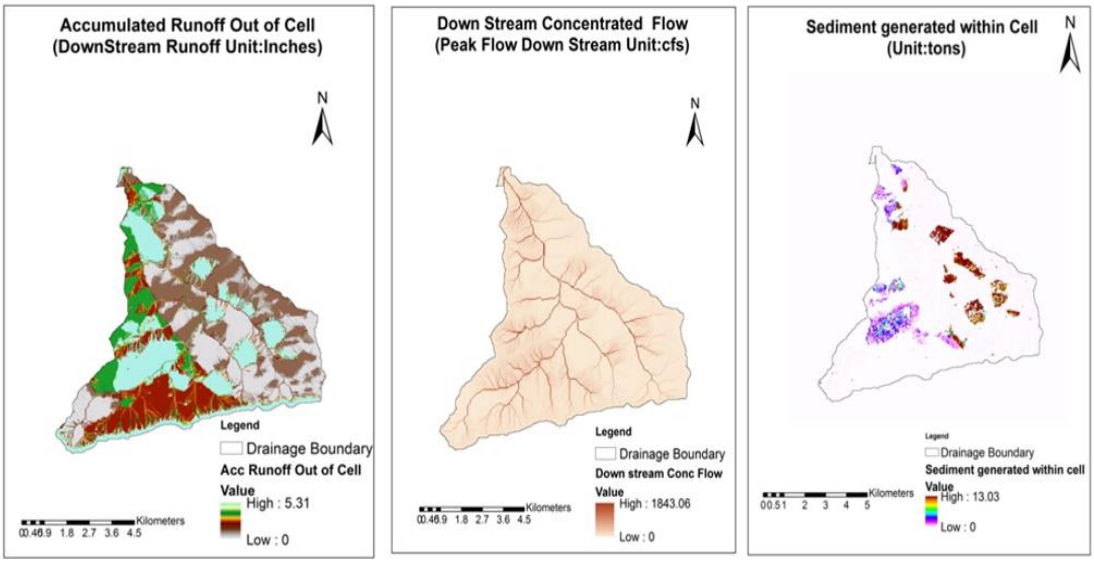


Figure 3: *e. Downstream accumulated runoff, f. Downstream concentrated flow, g. Sediment generated within cell*

The runoff parameter has been studied using SCS CN method incorporated within AGNPS model. It was observed that 50% area of the watershed produced 2.54 cm of runoff corresponding to 17.8 cm of rainfall. Similarly, the peak runoff characteristics of the watershed depicts that 64% of the Ata-Gad watershed produced peak runoff of 1.42 cumec corresponding to 17.78 cm of rainfall (Table 3).

Table 1: *Input parameters used in AGNPS model*

S. No	Parameter	How To Derive	Utility In the Model
1	Cell Number	Divide watershed into small cells of convenient size and number them as 1 for top left cell and proceed by each row, till the end	Identification
2	Receiving Cell	Based on drainage direction	To compute drainage direction for any cell
3	SCS Curve Number	Using standard curve number table based on land use map, hydrologic soil group map and antecedent moisture condition.	To compute retention factor and surface runoff
4	Land Slope	Using DEM	To determine velocity of overland flow
5	Filed Slope Length	Based on field observation of overland flow areas.	To calculate peak discharge
6	Channel Slope	Based on drainage map, topographic data and field observation	To calculate width of channel (for peak discharges)
7	Channel Side Slope	To calculate width of channel (for peak discharge)	Based on field observation and land use
8	Manning Roughness Coefficient	To calculate velocity in channel	Based on land use and field observation
9	Surface Condition Constant	To calculate time of overland flow	Based on topographic contour information
10	Flow Direction	To determine length of channel	Based on topographic contour information

11	Channel Length	Based on drainage	To determine length of channel
12	Soil Erodibility Factor (K)	To identify erodibility of soil	Based on soil type using USLE
13	Surface Condition	To identify soil surface condition	Based on land use
14	Aspect	Indicates drainage direction	Based on DEM
15	Soil Texture	Identification of clay, silt, sand	Based on soil map
16	Fertilization level	Based on field data	To identify fertilizer requirement
17	Fertilizer availability Factor	Based on field data	This estimates percentage availability of fertilizer left in top 1 cm of soil
18	Point source indicator	Based on field data	Indicates existence of a point source input within a cell
19	Gully source level	Based on field data	To estimate Gully Erosion In Cell (Tons)
20	Chemical oxygen level	Based on land use	Based on land use type
21	Impoundment indicator	Based on present terraces in the cell	Indicating presence of an impoundment terrace system within cell
22	Channel indicator	To determine the presence of channel	Based on Drainage

Table 2: *Erosional characteristics of Ata-Gad watershed under various slope categories*

S. No.	Slope category (%)	Area (ha)	Max. soil erosion (Tons/ha)	Mean soil erosion (Tons/ha)
1	0-5%	16.56	17.55	0.74
2	5-10%	48.53	14.97	0.63
3	10-15%	92.70	23.40	0.72
4	15-35%	955.86	29.39	0.99
5	35-50%	1396.57	31.14	1.26
6	50-75%	1980.65	36.43	1.21
7	> 75%	1196.33	48.67	0.99
	Total	5687.19		

Table 3: *Peak runoff characteristics of Ata-gad*

Peak Runoff (cumec)	Area (ha)	% of Area
<1.42	3653.28	64.24
1.42 - 7.08	738.81	12.99
7.08 - 14.16	382.77	6.73
14.16 - 21.24	381.60	6.71
21.24 - 28.32	326.16	5.73
28.32 – 35.40	164.34	2.89
35.40 – 42.48	26.73	0.47
42.48 – 49.56	7.02	0.12
>49.56	6.48	0.11
	5687.19	100.00

4. Conclusion

Soil is an important element essential for the sustenance of biotic systems and therefore, it is utmost necessary to preserve soil and its contents. The productivity of soil is determined by the nutrients and these nutrients and other soil sediments are removed by different agents like water and wind, etc. The AGNPS model was used in the present study to assess the runoff and sediment loss from Ata gad watershed of Pinder River in Uttarakhand hills. It was observed that large part of the watershed is

forested (61.9%) and agricultural activity is ongoing in lower reaches of the valley (18%). The watershed area is mostly under moderately steep (15-35%) to very steep slope (50-75%).

The erosional and runoff characteristics vis-à-vis other properties of the watershed have been analyzed. During the peak runoff analysis, it was observed that 64.24% of watershed area produced less than 1.42 cumec and only 0.11% of the area showed more than 49.55 cumec of peak runoff. Other parts of the watershed showed the runoff in the range of 1.42-49.55 cumec. One of the important parameter that affects runoff and soil erosion is slope. 1196 ha area of the watershed is having moderate to steep slope (slope>75%) and showed the maximum rate of erosion as 48.67 tons/ha. It is seen that slope factor solely does not affect the erosion characteristics. Such kind of studies will be helpful to manage resource utilization practices. The anthropogenic effects have impact on upstream-downstream interaction and water transport processes. The good catchment management practices at upstream can provide better opportunities for downstream communities and a clean and sustainable water supply for irrigation. The poor catchment management practices may not only degrade upstream environmental conditions, but will also limit the opportunities downstream.

References

Bisantino, T., Bingner, R., Chouaib, W., Gentile, F. and Trisorio Liuzzi, G. 2015. Estimation of runoff, peak discharge and sediment load at the event scale in a medium-size mediterranean watershed using the AnnAGNPS model. *Land Degradation & Development*, 26(4), pp.340-355.

Bosch, D., Theurer, F., Bingner, R., Felton, G. and Chaubey, I. 1998. *Evaluation of the AnnAGNPS water quality model*. Annual International Meeting, Orlando, Florida.

Choi, K.S. and Blood, E. 1999. Modeling developed coastal watersheds with the agricultural non-point source model. *Journal of the American Water Resources Association*, 35(2), pp.233-244.

Chowdary, V.M., Kar, S. and Adiga, S. 2004. Modelling of non-point source pollution in a watershed using remote sensing and GIS. *Journal of the Indian Society of Remote Sensing*, 32(1), pp.59-73.

Grunwald, S. and Norton, L.D. 1999. An AGNPS-based runoff and sediment yield model for two small watersheds in Germany. *Transactions of the ASAE*, 42(6), pp.1723-1731.

Haregeweyn, N. and Yohannes, F. 2003. Testing and evaluation of the agricultural non-point source pollution model (AGNPS) on Augucho catchment, western Hararghe, Ethiopia. *Agriculture, Ecosystems & Environment*, 99(1-3), pp.201-212.

Jena, S.K. and Tiwari, K.N. 2016. *Hydrologic modeling of watersheds using remote sensing, GIS and AGNPS*. Modeling Methods and Practices in Soil and Water Engineering. Apple Academic Press, pp.131-165.

Jianchang, L., Zhang, L., Zhang, Y., Huasheng, H.O.N.G. and Hongbing, D.E.N.G. 2008. Validation of an agricultural non-point source (AGNPS) pollution model for a catchment in the Jiulong River watershed, China. *Journal of Environmental Sciences*, 20(5), pp.599-606.

Ma, Y. and Bartholic, J. 2003. GIS based AGNPS assessment model in a small watershed. *Nat. Sci*, 1(1), pp.50-56.

Mishra, S.K., Chaudhary, A., Shrestha, R.K., Pandey, A. and Lal, M. 2014. Experimental verification of the effect of slope and land use on SCS runoff curve number. *Water Resources Management*, 28(11), pp.3407-3416.

Mostaghimi, S., Park, S.W., Cooke, R.A. and Wang, S.Y. 1997. Assessment of management alternatives on a small agricultural watershed. *Water Research*, 31(8), pp.1867-1878.

Najim, M.M.M., Babel, M.S. and Loof, R. 2006. AGNPS model assessment for a mixed forested watershed in Thailand. *Science Asia*, 32, pp.53-61.

Perrone, J. and Madramootoo, C.A. 1997. Use of AGNPS for watershed modeling in Quebec. *Transactions of the ASAE*, 40(5), pp.1349-1354.

Rainis, R. 2004. Estimating sediment yield using Agricultural Non-Point Sources (AGNPS) model: The effects of slope information from different GIS softwares. *Journal of Spatial Hydrology*, 4(2).

Sarangi, A., Cox, C.A. and Madramootoo, C.A. 2007. Evaluation of the AnnAGNPS model for prediction of runoff and sediment yields in St Lucia watersheds. *Biosystems Engineering*, 97(2), pp.241-256.

Smith, R.E. and Williams, J.R. 1980. Simulation of surface water hydrology. In: CREAMS, A field scale model for chemicals, runoff, and erosion from agricultural management systems. *USDA Conservation Resources Report*, 26(1), p.15.

Young, R.A. and Shepherd, R.G. 1995. *AGNPS-Agricultural Nonpoint Source Model*. Workshop on Computer Applications in Water Management, 33.

Yuan, Y., Bingner, R L. and Rebich, R.A. 2001. Evaluation of AnnAGNPS on Mississippi Delta MSEA watersheds. *Transactions of the ASAE*, 44(5), p.1183.

Zema, D.A., Bombino, G., Denisi, P., Licciardello, F. and Zimbone, S.M. 2012. *Prediction of Surface Runoff and Soil Erosion at Watershed Scale: Analysis of the AnnAGNPS Model in Different Environmental Conditions*. Research on Soil Erosion, 2012.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2560-2573
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.347

Research Article

Fraction of Vegetation Cover and Its Application in Vegetation Characterization in the Hazaribagh Wildlife Sanctuary, Jharkhand, India

Saurabh Kumar Gupta, A.C. Pandey

Center for Land Resource Management, Central University of Jharkhand, Jharkhand, India

Publication Date: 7 April 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.347

Abstract The Vegetation Indices (VI) used for estimation of variations in vegetation cover in past decades. The Fraction of vegetation cover was calculated from pre-computed reflectance database based on inversion of PROSAIL model. It was found that the very low and low forest vegetation cover area was dominated with an area of about 104 Km^2 and 69.49 Km^2 respectively. The very high forest vegetation cover area is very low of about 24 Km^2. The high and medium vegetation cover forest shows intermediate dominance of about 51.74 Km^2 and 57 Km^2 in area respectively. The result indicates that NDVI, RVI and PSSRA value closely related to forest vegetation cover. Also, in the regression analysis the same was observed as the high relationship is found between FCOVER AND PSSRA (R^2 = 68%) followed by NDVI (R^2 = 66%) and RVI (R^2= 65%). The low relationship is found in the relationship of FCOVER and MCARI (R^2 = 36.7%) followed by DVI (R^2 =53.7%) and IRECI (R^2 = 59%).

Keywords *Fractional vegetation cover; Forest vegetation cover; Vegetation indices*

1. Introduction

The fraction of vegetation cover (FCOVER) gives the information about the biophysical status of the forest. The FCOVER is an imperative variable for many spatial biophysical and biogeochemical models and used for the measurement of land cover change (Zhang, 2006). According to remote sensing techniques, FCOVER may be defined as the vegetated area which is directly visible from the sensor (Purevdor et al., 1998). Quantitative information on the vegetation cover is required in many studies for observing the global and local changes in forest landscapes.

With the advent of satellite imaging technology, it is much more common to use remote sensing techniques to monitor forest data, in particular, tropical deforestation. Remote sensing based methods for retrieving FCOVER have developed rapidly during the last few decades. The upcoming of Sentinel 2 satellites offer new open doors for a nonstop checking of the land and vegetation with regards to the climate change and global warming (POENARU, 2017).

Satellite sensors Sentinel-2 has the capability to forecast the forest activities and spectral resolution of 10 m at its red and NIR bands make it more reliable. It is ESA's mission to provide high-resolution imagery and enhanced data continuity for Landsat and SPOT (Drusch, 2012). The

several studies have been implemented that show the potential of Sentinel -2A for measuring many biophysical and biochemical parameters like FCOVER (Vander Meer, 2014).

Theoretical analyses and field studies have shown that VIs is near-linearly connected to the fraction of absorbed photosynthetic active radiation (fAPAR), a fraction of vegetation cover (FCOVER), roughness length for turbulent transfer, emissivity and albedo (Glenn, 2008). Remotely Sensed vegetation indices help to analyze vegetation parameters such as leaf area, biomass and physiological activities (Baret and Guyot, 1991; Verrelst et al., 2008). Spectral vegetation indices that are based on red and near-infrared reflections have the highest correlation with leaf area index and vegetation cover (Broge and Leblac, 2000). Spectral vegetation indices (VI) are very helpful in characterizing forest vegetation based on satellite data and spatially and temporally variables in vegetation conditions (Yang, 2017; Pricope, 2015). These Studies suggest the important relation of spectral indices to extract plant signal.

The present study utilizes the inversion of the PROSAIL radiative model which permitted us to recover vegetation biophysical variables (e.g. FCOVER) (Jacquemoud et al., 1995) from multispectral Sentinel 2A imagery. The proposed method was applied to monitor the forest vegetation cover in the Hazaribagh wildlife sanctuary, located in the eastern parts of India. The present study also utilizes scatter plot matrix which analyses the relationship between FCOVER and VIs. The scatter plot gives the linear relationship between the model output and the uncertain component using regression analysis (Srivastava et al., 2012).

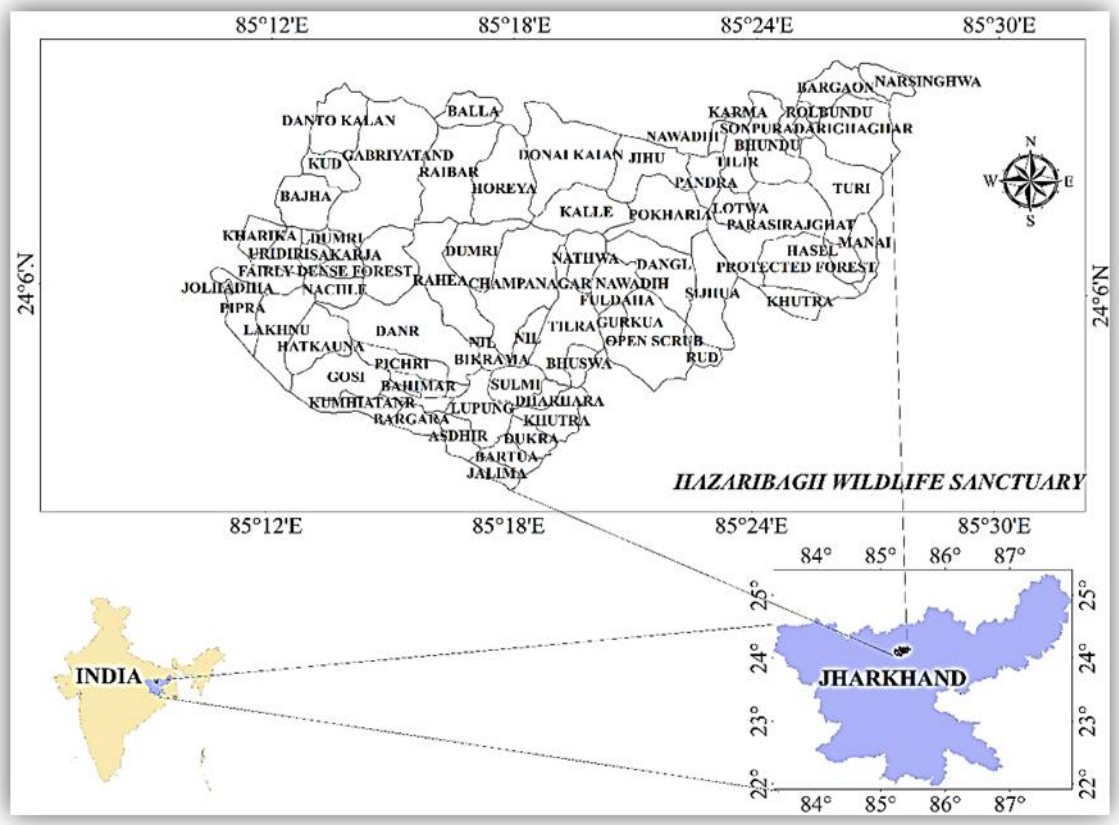


Figure 1: *Study area- Hazaribagh Wildlife Sanctuary*

1.1. Study Area

This study was performed in the Hazaribagh wildlife sanctuary, Jharkhand, India (Figure 1). It lies between 24°45'22" N to 24°08′20″N latitude and 85° 30'13" E to 85°21′58″E longitude and also includes bird Sanctuary and other biodiversity parks. The climate of the region is tropical having hot summers and chilly winters. In summer, greatest temperature rises to up to 41 degrees and low of 19 degrees. In winter, most extreme high and low temperature decreased are 19 degrees and 7 degrees respectively. The sanctuary has Sambhar, Deer, Bison and various mammalian fauna. The Cheetah, Kakar, Nilgai, Sambar and Wild Boar are among the most effective and frequently spotted creatures, especially close to the waterholes at the time of the sunset.

2. Methodology

The Figure 2 illustrates the flow chart of methodology used in the present study and its details is given according to it below. It involves the acquisition of sentinel 2A imagery and their processing for retrieving FCOVER, FCC and vegetation indices. The accuracy was done through the help of aerial map (Google Earth) and observing False Color Composite (FCC). The classification FCOVER and sample points to raster was converted by the use of Arc GIS software and further confusion matrix, overall accuracy estimation and kappa analysis were generated using QGIS software.

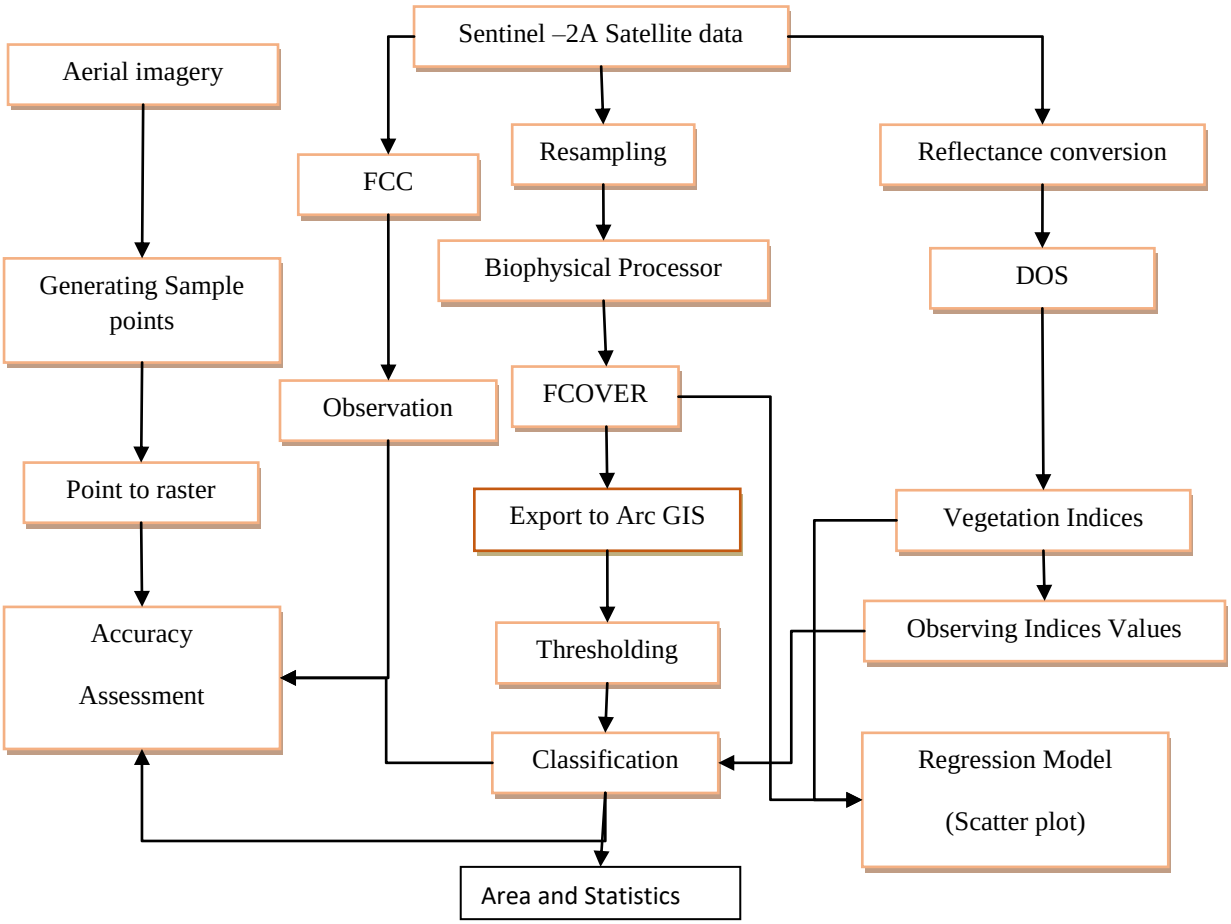


Figure 2: *Flow chart of methodology adopted in the present study*

2.1. Data and Software Used

In this study, Sentinel-2A optical satellite data were acquired from the USGS earth explorer (https://earthexplorer.usgs.gov/) for the period of February-2016 (Table 1). It contains 12 spectral bands with a spatial resolution of 10m, 20m, and 60m. This data is processed with SNAP software developed by ESA, which is excellent software for the processing of Sentinel satellite imagery. Q-GIS software was used for the DN to the reflectance conversion of the imagery and then the processed data was used for computation of various indices. The result obtained was used as an input parameter in SAGA GIS software for sensitive analysis between FCOVER and Vegetation Indices. Finally, all the results were exported using ArcGIS software.

Table 1: *The satellite data used in the present study - Fractional of vegetation cover (FCOVER) estimation*

Sensor	Date	Id	Processing Level	Product type
Sentinel2A	2016-02-13	20160213T101928_A003362_T45QUG	Level-1C	S2MSI1C

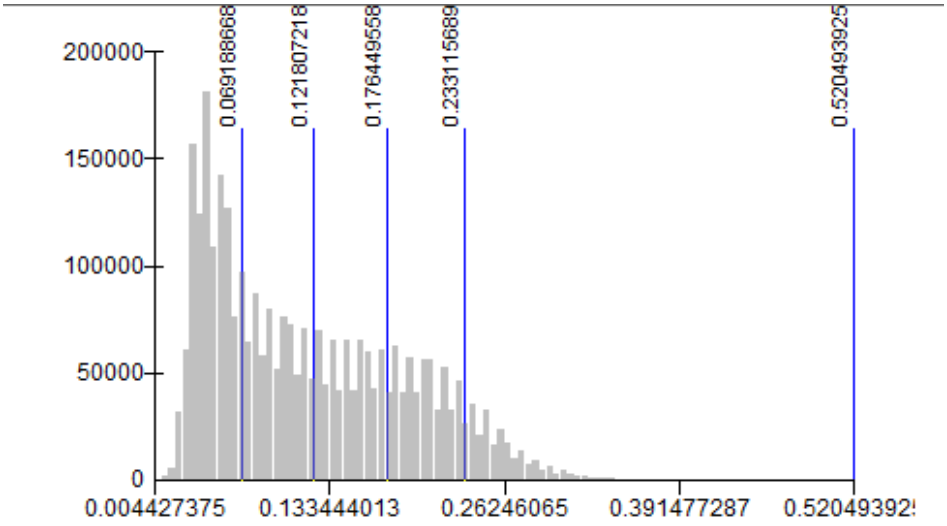


Figure 3: *Classification of FCOVER*

The FCOVER is a canopy inherent variable that depends upon the structure of the canopy (LAI, Leaf angle and clumping) (Bacour, 2006). The FCOVER derived from the analysis of the PROSAIL model inversion technique of artificial neural network approach from multispectral Sentinel 2 data. In this perspective, the portion or fraction of vegetation cover in the forest area of the Hazaribagh wildlife sanctuary was calculated by using inversion of PROSAIL model embedded in the biophysical processor of the SNAP ESA Toolbox. It is already proven to be good at estimating FCOVER. It uses trained neural network algorithm having the input of 11 normalized data (B3, B4, B5, B6, B7, B8a, B11, B12, cos (Viewing_Zenith), cos (Sun_Zenith), cos (relative_azimuth_angle)) of the Sentinel-2A Sensor, one hidden layer with tangent sigmoid transfer function contains 5 neurons and an output layer with a linear transfer function (Vuolo et al., 2016). The training database contains canopy reflectance spectra with the corresponding output variables. The FCOVER obtains from the processing step of the artificial neural network and its inversion techniques, from the encoded coefficient computed during the training phase. The training process depends on the mean square error between the reflectance database and the output FCOVER. The network planning is determined by trials and errors, to get the optimal performances with minimum number of inputs layers and neurons. The

validation of the FCOVER was done through ground truthing during its training period. For the calculation of FCOVER the sentinel 2A bands firstly resample to 60m resolution, the resolution was chosen to ease the exporting task in Tiff format. Resampling is essential for converting each input bands into same resolution, as bands in sentinel 2A come in different resolution of 10m, 20m and 60m.

With the use of FCOVER, the forest vegetation cover area has been classified by thresholding it in ARC GIS (Figure 3). The value of the FCOVER between 0.23 to 0.52 considered as very high vegetation cover area, the value consisting 0.176 to 0.23 is considered as high vegetation area, the value between 0.12 to 0.176 is considered as a medium vegetation cover area, 0.069 to 0.12 as low vegetation cover area value of 0 to 0.069 as very low vegetation cover area and e statistical analysis was carried out in classified images of FCOVER, which gives area and percentage of classified FCOVER.

2.2. Atmospheric Correction

The estimation of radiometric vegetation indices depended on the reflectance properties of bands on Sentinel 2A satellite data. The bands require atmospheric correction before their use. For this reason, the raw satellite imagery of Sentinel 2A was first converted into reflectance. The DN (Digital Number) to reflectance conversion (Calibration) needs a header document given in the downloaded satellite imagery which contains all the necessary inputs like acquisition date, sun height, gain and bias for correcting the imagery (Lu, 2001). Then, the images were corrected by the dark object subtraction (DOS) to get real surface reflectance of the object. The DOS strategy is an image-based technique to counterbalance the haze component caused by additive scattering from remote sensing data (Chavez, 1988). For this purpose, QGIS software was utilized for the atmospheric correction of Sentinel 2 A imagery. After these processing the bands were selected and mathematically operated in raster calculator in QGIS for the calculation of vegetation indices. The mathematical operation was accord to the vegetation index formula (table2) for the calculation of the selected vegetation indices for evaluating its performance in comparison to FCOVER and vegetation cover types.

2.3. Vegetation Indices

The six types of different vegetation indices were computed utilizing the mathematical operation on bands of satellite data to represent the amount and structure of vegetation (Figure 4).The most commonly used bands are RED and NIR because they are sensitive to the soil and vegetation reflectance respectively. The Difference Vegetation Index (DVI) is sensitive to the amount of vegetation and difference soil from vegetation. In DVI zero indicates bare soil, values less than zero indicate water, and those greater than zero indicate vegetation. The range of this index is theoretically very large, but if reflectance images are used, it tends to give values between 0 and 1. The Inverted Red-edge Chlorophyll Index (IRECI) needs four bands to estimate forest chlorophyll content. The red-edge is the region of vegetation spectra (680nm±740nm, red bands) that is sensitive to leaf scattering and chlorophyll absorption. The exact range value of it for vegetation and other land feature was not determined yet. The Ratio Vegetation Index (RVI) is high for vegetation and low for water, ice and soil. It indicates the amount of vegetation and reduces the atmospheric effects and topography. Note that the SR is not bounded; its values can increase far beyond 1. Generally, very high SR values are on the order of 30. Typical ranges are a little more than 1 for bare soil to more than 20 for dense vegetation. The Pigment Specific Simple Ratio (PSSRA) exhibits the linear relationship with the canopy concentration per unit area of chlorophyll and carotenoids. The exact range value of it for vegetation and other land feature was not determined yet. The Normalized Difference Vegetation Index (NDVI) is the most commonly used vegetation indices used in the analysis of vitality and

strength of vegetation. Data from vegetated areas will yield positive values for the NDVI due to high near-infrared and low red or visible reflectance. The NDVI increases in value up to nearly 1 as the amount of green vegetation increases in a pixel (picture element).In contrast, bare soil and rocks generally show similar reflectance in the near-infrared and red or visible, generating positive but lower NDVI values close to 0. The red or visible reflectance of water, clouds, and snow are larger than their near-infrared reflectance, so scenes containing these materials produce negative NDVIs. The typical range of actual values is about 0.1 to 0.2 for bare soils and 0.2 to 0.9 for vegetation. The Modified Chlorophyll Absorption Ratio Index (MCARI) used to observe the chlorophyll variation and ground reflectance. The exact range value of it for vegetation and other land feature was not determined yet.

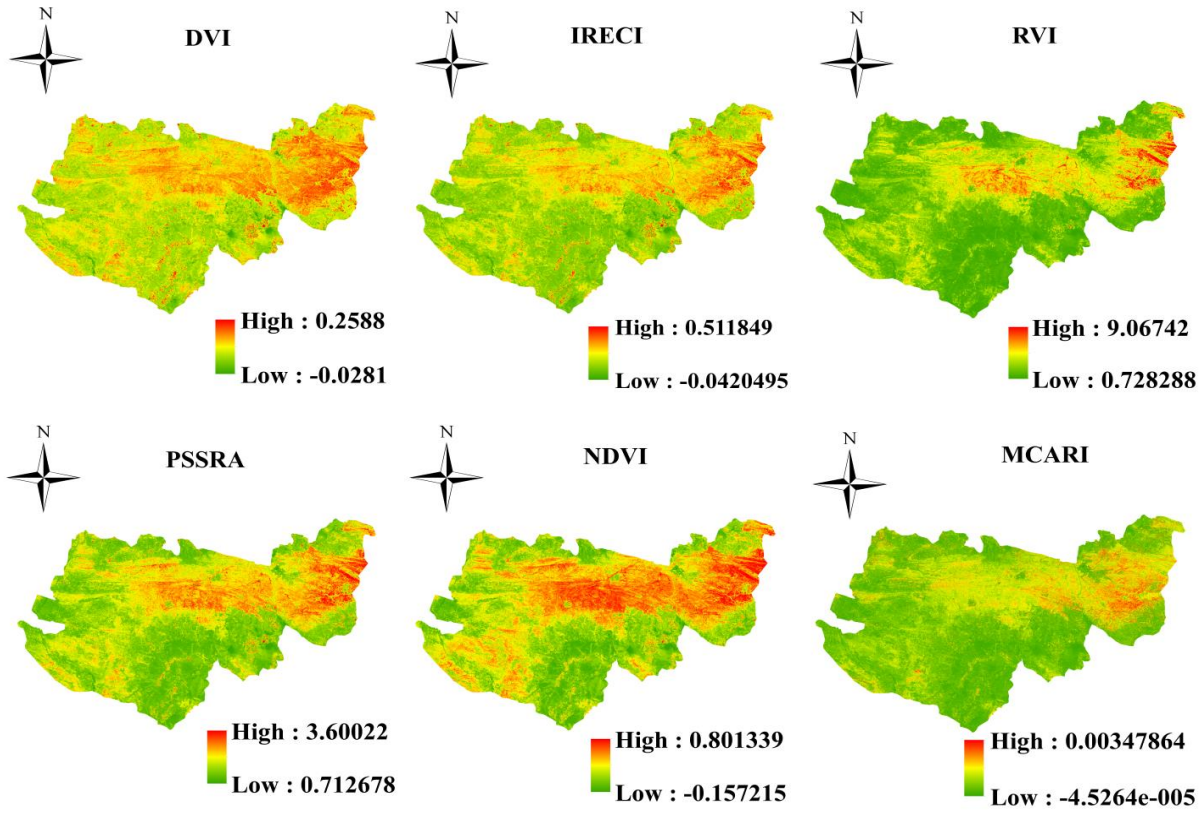


Figure 4: *Spectral vegetation index used in the study area*

2.4. Regression Model

The scatter plot was preferred for the regression analysis between two raster variables. The variables here are rasters of vegetation indices and FCOVER. The potential of estimation of forest vegetation cover by indices is analyzed based on regression equation and R^2, which define mathematically the relation between two variables.

2.5. Accuracy Assessment

The validation of the result has been always a main criterion for the assessment of remote sensing based analysis. In general, validation for the most part refers to evaluating the uncertainties in remote sensing indirectly comes about by means of correlation with ground truth or between examination with

another option data that apparently speaks to the true condition of an object (Wang, 2016). In this paper, the validation was done through observing the false color composite (FCC) of imagery and the taking sample point of vegetation cover type from Google earth (sampling methods).

Table 2: *Spectral vegetation indices and its formula for Sentinel-2 A bands*

Vegetation indices	Formula	References
DVI	(B8-B4)	Richardson and Wiegand (1977)
IRECI	(B7-B4)/B5/B6	Frampton et al. (2013)
RVI	B8/B4	Birth and McVey (1968).
PSSRA	B7/B4	Black burn, 1998
NDVI	(B8−B4)/ (B8 +B4)	Rouse et al. (1974)
MCARI	[(B5-B4)-0.2 * B5-B3] * B5/B4	Daughtry, et al. 2000

The accuracy assessment was done for the validation of the classified vegetation cover areas. These were assessed in two ways. Firstly, the classification was compared with the false color composite (8-4-3 bands) of the same sentinel imagery. The visual and color interpretation was done to observe the vegetation condition in the forest. Secondly, the training sample point was generated from the Google earth imagery. The point was taken by observing the different classes of vegetation cover according to classified FCOVER. Around 250 random sampling was done and then these points were converted to raster format in ARC GIS. These raster was taken as reference classified map for the accuracy assessment with the classified FCOVER**.** Then, the overall accuracy, confusion matrix and kappa coefficient were retrieved using QGIS software.

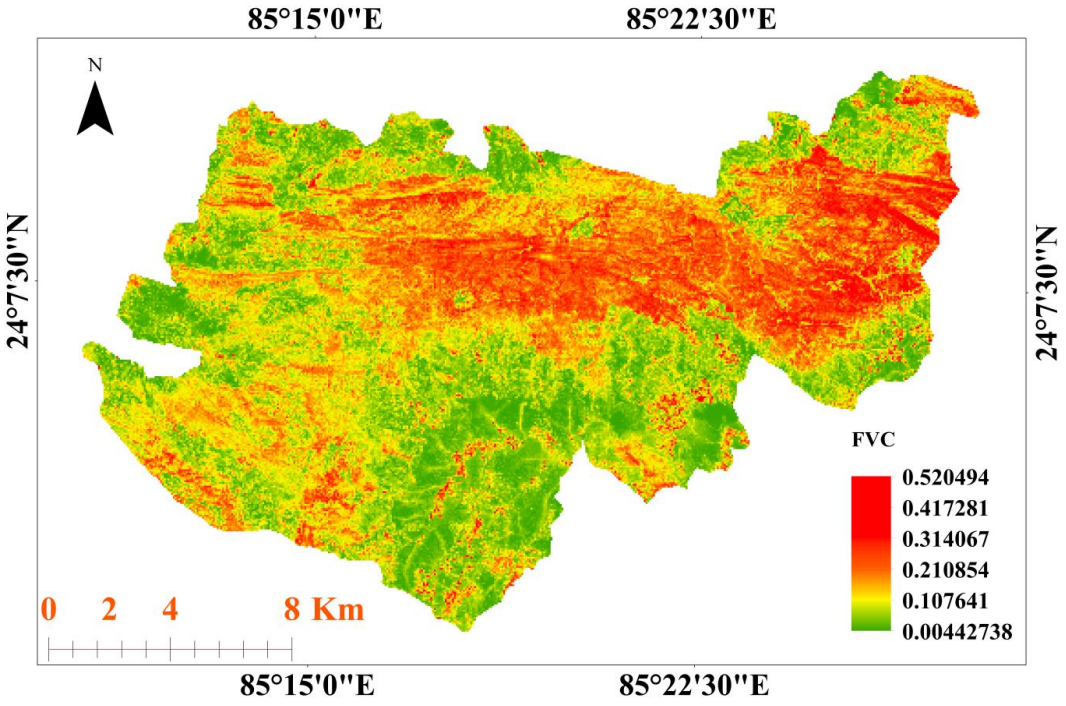


Figure 5: *Fraction of vegetation cover (FCOVER) map of the wildlife sanctuary*

3. Result and Discussion

3.1. Classification and Area Estimation of Vegetation Cover

The fraction of vegetation cover was used to estimates the vegetation cover in the forest (Figure 3). The value of FCOVER indicates the type of forest vegetation cover. Its value ranges from 0 to 1. In Figure 5, FCOVER of the study area is given; its high range is 0.52 whereas the low value is 0.0044. The classification can also be observed in the Figure that represents the classified FCOVER by natural breaks (Jenks) methods of classification in ARC GIS. Figure6indicates the classification of the forest vegetation covers according to FCOVER value. The statistics of an area and the percentage of FCOVER are given (Table 3). It is found that the very low and low forest vegetation cover was dominant, with the area of about 104 Km^2 and 69.49 Km^2 respectively, indicates deforestation and degradation in the part of the area. The very high forest vegetation cover area observed in small patches across the study area which contributing 24 Km^2. The high and medium vegetation covers forest shows intermediate dominance, about 51.73 Km^2 and 57 Km^2 in area respectively. In this study, we identified that forest vegetation monitoring with FCOVER is a best suited methodology and future change detection analysis can be possible. The Sentinel 2A is a new satellite and in the coming year, it can be helpful in change detection or seasonal change monitoring.

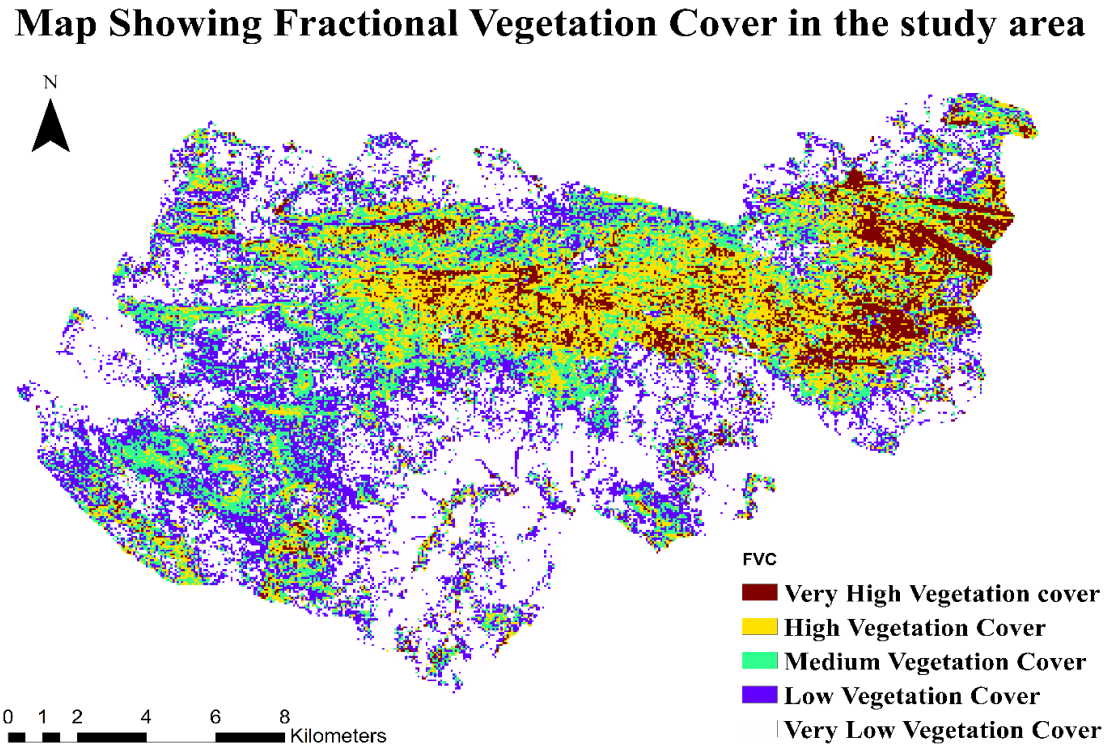


Figure 6: *Forest vegetation classification based on FCOVER*

3.2. Accuracy Assessment

In the Figure 7, the false color composite of the sentinel 2A imagery is given, which indicates the true condition of vegetation and can be taken for validating the obtained result of classified FCOVER. In the FCC, the high vegetation cover area appears in dark red color. Sometimes it is misunderstood by the

shadow, as the forest canopy cover creates shadows depends upon the sun geometry. The more darkness in red found in the very high vegetation cover area and darkness decrease in the high vegetation cover followed by other medium and low vegetation cover area. The low and very low vegetation cover area looks pink and gray white respectively. Thus, the above validation for the obtained result is satisfactory.

In the Table 4, the accuracy assessment is given which gives the details about confusion matrix generated from reference map (sampling method) and classified FCOVER. The overall accuracy was observed 95.65% with kappa coefficient of 0.94.

Table 3: *Area and percentage of forest vegetation classes based on FCOVER*

FCOVER	Forest vegetation cover	Percentage %	Area (Km^2)
0.5	Very High	7.84	24.163200
0.4	High	16.79	51.735600
0.3	Medium	18.8	57.916800
0.2	Low	22.56	69.494400
0.1	Very Low	33.9	104.731200

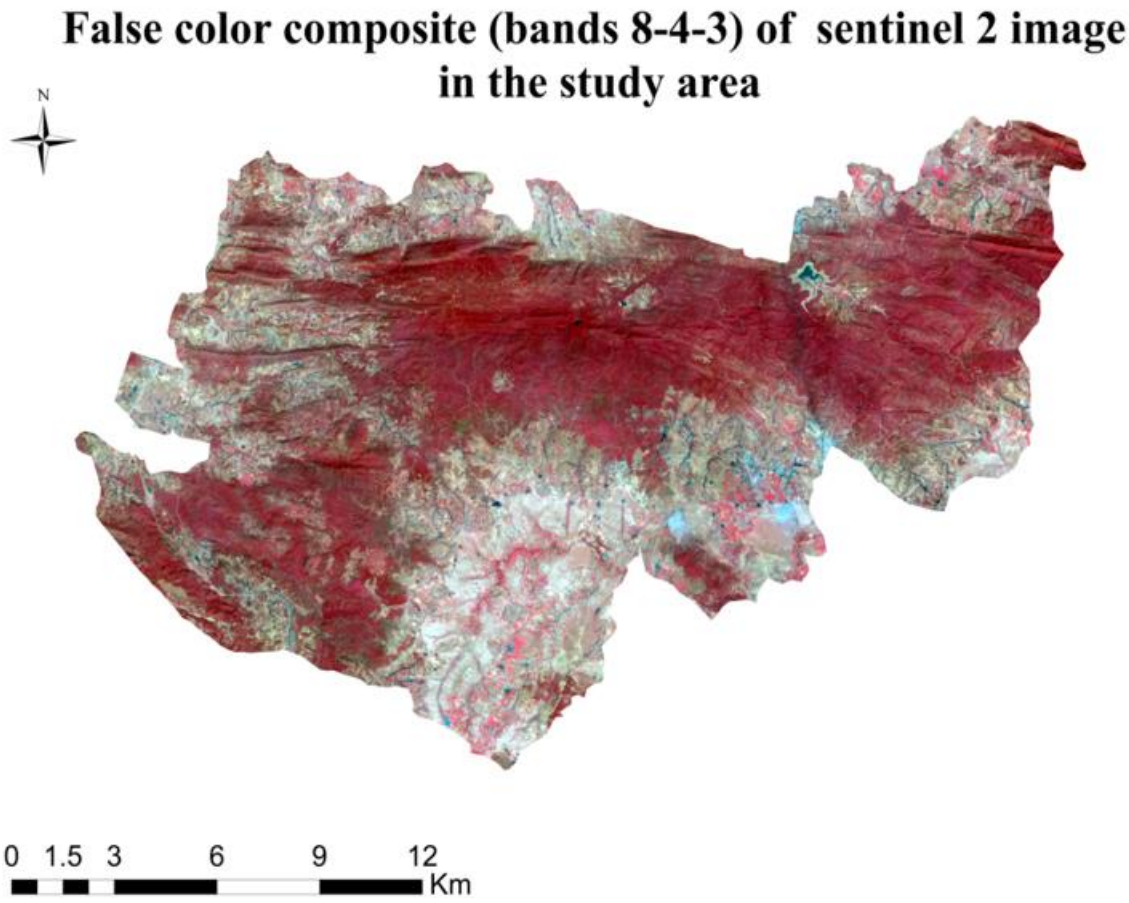


Figure 7: *False color composite of Sentinel imagery*

Table 4: *Accuracy assessment*

> Confusion MATRIX						
Vegetation cover > Reference	**Very Low**	**Low**	**Medium**	**High**	**Very High**	**Total**
Very Low	1047276	0	0	0	0	1047276
Low	0	694908	0	0	0	694908
Medium	0	46908	532260	0	0	579168
High	0	0	54144	430452	32760	517356
Very High	0	0	0	0	241632	241632
Total	1047276	741816	586404	430452	274392	3080376
Overall accuracy [%] = 95.6559848538 Kappa hat classification = 0.943128853533						

3.3. Vegetation Indices Thresholding and Its Performance

The value of NDVI ranges from 0.81 to -0.15. The value of 0.6 to 0.82 indicates the very high vegetation cover area. The value from 0.47 to 0.6 is the high vegetation cover. The value that ranges from 0.33 to 0.47 is the medium vegetation cover area. Lastly, the value from 0.2 to 0.3 indicates a low vegetation cover and very low vegetation cover ranges from 0 to 0.2. The value that is close to zero indicates bare soil and the negative value indicates rocks and water bodies

The value of DVI range is between 0.2588 to -0.021. The value closes to zero, indicate bare soil and the negative value indicate rocks and water bodies. The value above 0.116 is for very high vegetation cover. The value 0.116 to 0.098 indicates a high vegetation cover; the value of 0.098 indicates 0.076 medium vegetation cover and less than this indicates low and very low vegetation cover. There is difficulty in the classification of low and very low vegetation cover and the value of DVI indicating FCOVER based classification, doesn't properly matched in all classes.

The range of RVI obtained is 0.9 to 0.72. The value that is greater than 0.395 indicates a very high vegetation cover. The value ranges from 3 to 3.95 indicates a high vegetation cover. The value ranges from 2 to 3 comprises medium vegetation cover. The value ranges from 1 to 2 consists of low vegetation cover and value less than 1 indicates very low vegetation cover. The RVI is somewhat well related to FCOVER based classification than DVI.

The PSSRA value ranges from 3.6 to 0.71. The value of 2 to 3.6 indicates a very high vegetation cover. The value ranges from 1.76 to 2 indicates a high vegetation cover. The value of the PSSRA from 1.48 to 1.76 indicates a medium vegetation cover and its value from 1.199 to 1.48 indicates low and very low vegetation cover jointly, means doesn't able to categorize them. The value less than 1.199 indicates water and rocks.

The MCARI value ranges from 0.0035 to -0.000057. The value that retrieved is very low to define any forest vegetation cover classes. However, the indication is brought by effort to examine the forest vegetation classes in its smallest value. The value that ranges from 0.0007 to 0.0035 is the high vegetation cover area. The value ranges from 0.00044 to 0.0007 is the high vegetation cover. The value that ranges from 0.00018 to 0.00044 is the medium vegetation cover area. Lastly, the value that ranges from -0.000057 to 0.00018 indicates both low and very low vegetation cover areas jointly, means relationship of MCARI and vegetation cover is uncertain at these values to indicate classes properly.

The value of IRECI ranges from 0.51 to -0.042. The value that ranges from 0.1413 to 0.514 indicates higher vegetation cover area. The value ranges from 0.1077 to 0.14 is the high vegetation cover. The value that ranges from 0.074 to 0.1077 is the medium vegetation cover area. Lastly, the value that ranges from 0.007 to 0.074 indicates both low and very low vegetation cover areas jointly that is from difficulties in classification. The value that is close to zero indicates bare soil and the negative value indicates rocks and water bodies

This analysis shows that they doesn't totally good match with the vegetation cover types and these indices at a certain range show just an indication of vegetation types. However, NDVI, RVI and PSSRA show some potential in the classification that more relates to vegetation classification.

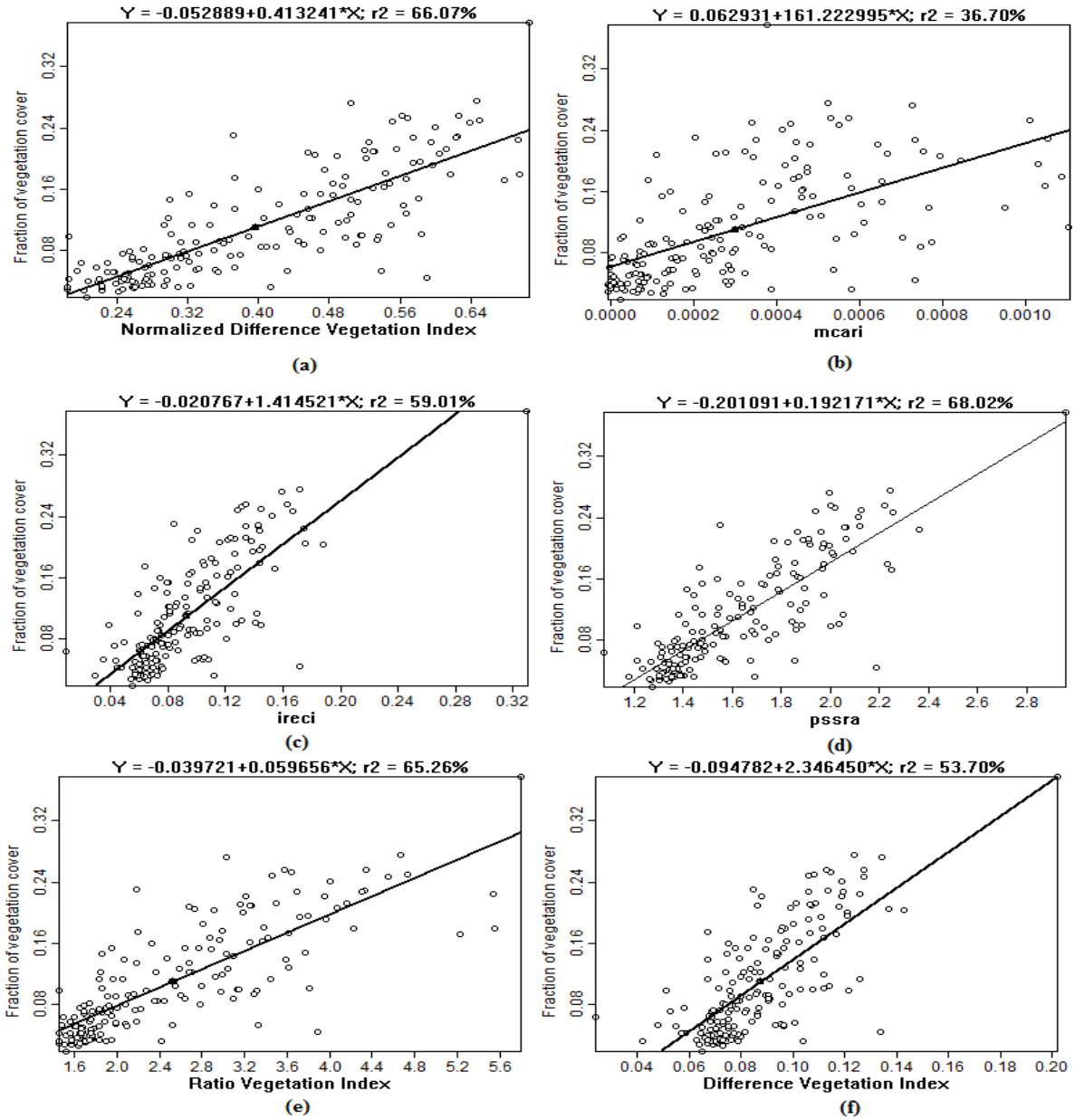


Figure 8: *Scatter plot matrix for regression analysis of FCOVER with VIs; a) FCOVER vs NDVI, (b) FCOVER vs MCARI, (c) FCOVER vs IRECI, (d) FCOVER vs PSSRA, (e) FCOVER vs RVI and FCOVER vs DVI*

3.4. Regression Analysis

It is necessary to analyze the relationship of FCOVER with vegetation indices because the indices can be helpful for the estimation of forest biophysical variables indirectly, therefore helps in forest monitoring. Theoretical analysis and field study shows that the FCOVER is related to Vegetation indices. In Figure 8, the scatter,plot is given which gives the relationship between various vegetation indices with the FCOVER. The regression analysis by using the scatter plot was done to observe the relationship of vegetation indices with FCOVER. The high relationship was found between FCOVER AND PSSRA (R^2 = 68%) followed by NDVI (R^2 = 66%) and RVI(R^2= 65%). The low relationship was found in the relation of FCOVER and MCARI (R^2 = 36.7%) followed by DVI (R^2 =53.7%) and IRECI (R^2 = 59%). The detail of the relationship between FCOVER and VIs is given (table4). The PSSRA, NDVI and RVI shows greater response toward FCOVER. Therefore, they show potential in forest monitoring activities.

Table 5: *Regression analysis of FCOVER with different VIs*

Y	X	Regression equation	R^2
FCOVER	NDVI	-0.05 + 0.41X	66.07%
FCOVER	MCARI	0.06 + 161.2X	36.7%
FCOVER	IRECI	-0.02+ 1.41X	59.01%
FCOVER	PSSRA	-0.2+ 0.19X	68.02%
FCOVER	RVI	-0.04+ 0.06X	65.26%
FCOVER	DVI	-0.09 +2.35X	53.7%

4. Conclusion

The present work uses FCOVER analysis to find the response of vegetation indices for vegetation cover monitoring; it is accurate and reliable because of its utilization and processes are easy to implement. In the study area, the lower level of FCOVER shows the condition of the forest is not good. FCOVER based classification of forest vegetation proved to be useful in defining several classes of forest vegetation cover. It requires less ground truth information for accuracy assessment. However, validation was done by the help of the FCC and Google imagery to observe the correctness of the obtained classification and found satisfactory. The overall accuracy was observed 95.65% with a kappa coefficient of 0.94.

The estimation of FCOVER works on pre-trained neural PROSAIL model which was pre-tested and proved to be efficient in a monitoring fraction of vegetation cover (FCOVER) in the forest.

The response of vegetation indices toward the vegetation cover was examined through the use of comparing the pixel value of classified vegetation cover types to VIs and from regression analysis of FCOVER and VIs. The FCOVER and VIs show positive and linear relationships among them, but NDVI and PSSRA show the high relationship both by observing the thresholds of VIs in vegetation cover and from regression analysis. The result indicates that NDVI, RVI and some PSSRA value ranges closely related to FCOVER based classification of forest vegetation cover. The NDVI, PSSRA and RVI can be used for monitoring forest vegetation and their quantifications.

The result reveals that the regression analysis and adopted thresholding of vegetation indices proves to be useful in the selection of potential indices for forest vegetation monitoring.

References

Baret, F. and Guyot, G. 1991. Potentials and Limits of Vegetation Indices for LAI and APAR Assessment. *Remote Sensing of Environment*, 173, pp.161-173.

Birth, G., and McVey, G. 1968. Measuring the color of growing turf with a reflectance spectrophotometer. *Agronomy Journal*, 60, pp.640-643.

Blackburn, G.A. 1998. Quantifying chlorophylls and carotenoids at leaf and canopy scales: an evaluation of some hyperspectral approaches. *Remote Sens. Environ.*, 66, pp.273-285.

Broge, N.H. and Leblanc, E. 2000. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.*, 76, pp.156-172.

Chavez, P.S. Jr. 1989. Radiometric calibration of Landsat thematic Mapper multispectral images. *Photogrammetric Engineering and Remote Sensing*, 55, pp.1285-1294.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., Brown de Colstoun, E. and Mcmurtey, J. 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.*, 74, pp.229-239.

Drusch, M., Del Bello, U.S., Carlier, O., Colin, V., Fernandez, F., Gascon, P. and Bargellini, P. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, pp.25-36.

Frampton, W.J., Dash, J., Watmough, G. and Milton, E.J. 2013. Evaluating the Capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, pp.83-92.

Glenn, E.P., Huete, A.R., Nagler, P.L. and Nelson, S.G. 2008. Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*.

Lu, D., Mausel, P., Brondizio, E. and Moran, E. 2002. Assessment of atmospheric correction methods for Landsat TM data applicable to Amazon basin LBA research. *International Journal of Remote Sensing*, 13, pp.2651-2671.

Poenaru, V., Badea, A., Negula, I.D. and Moise, C. 2017. Monitoring vegetation phenology in the braila plain using sentinel 2 data. *Land Reclamation, Earth Observation Surveying, Environmental Engineering*, 9, pp.175-180.

Pricope, N.G., Gaughan, A.E., All, J.D., Binford, M.W. and Rutina, L.P. 2015. Spatio-temporal analysis of vegetation dynamics in relation to shifting inundation and fire regimes: disentangling environmental variability from land management decisions in a Southern African Transboundary Watershed. *Land*, 4(3), pp.627-655.

Purevdor, T., Tateishi, R., Ishiyama, T. and Honda, Y. 1998. Relationships between percent vegetation cover and vegetation indices. International Journal of Remote Sensing, 18, pp.3519-3535.

Richardson, A.J. and Wiegand, C.L. 1977. Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, 43, pp.1541-1552.

Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, W.D. 1973. *Monitoring vegetation systems in the Great Plains with ERTS*. In: Third ERTS Symposium. NASA SP-351, pp.309-317.

Srivastava, P.K., Han, D., Gupta, M. and Mukherjee, S. 2012. Integrated framework for monitoring ground waterpollution using a geographical information system and multivariate analysis. *Hydrological Sciences Journal*, 57, p.1453e1472.

Van der Meer, F.D., van der Werff, H.M.A. and van Ruitenbeek, F.J.A. 2014. Potential of ESA's Sentinel-2 for geological applications. *Remote Sensing of Environment*, 148, pp.124-133.

Verrelst, J., Schaepman, M.E., Koetz, B. and Kneubühler, M. 2008. Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. *Remote Sensing of Environment*, 112, p.2341e2353.

Vuolo, F., Atzberger, C., Richter, K., D'Urso, G. and Dash, J. 2010. Retrieval of biophysical vegetation products from RapidEye imagery. ISPRS TC VII Symposium – 100 Years. *ISPRS*, 38, pp.281-286.

Wang, S., Li, X., Ge, Y., Jin, R., Ma, M., Liu, Q., Wen, J. and Liu, S. 2016. Validation of regional-scale remote sensing products in china: from site to network. *Remote Sens.*, 8, p.980.

XIV World Forestry Congress. 2015. FAO. Available from: www.fao.org/about/meetings/world-forestry-congress/en/. Assessed 10.09.2015.

Yang, X. and Guo, X. 2014. Quantifying responses of spectral vegetation indices to dead materials in Mixed Grasslands. *Remote Sens.*, 6(5), pp.4289-4304.

Zhang, X., Yan, G., Li, Q., Li, ZL., Wan, H., and Guo, Z. 2006. Valuating the fraction of vegetation cover based on NDVI spatial scale correction model. *International Journal of Remote Sensing*, 27(24), pp.5359-5372.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2466-2470
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.333

Methodology Article

Urban Area Delineation Using Pattern Recognition Technique

Tutu Sengupta[1], Satish Sharma[2]

[1]Maharashtra Remote Sensing Application Centre,VNIT Campus, South Ambazari Road, Nagpur & Research Scholar, Department of Physics, RTM Nagpur University, Nagpur, Maharshtra, India
[2]Department of Electronics and Computer Science, RTM Nagpur University, Nagpur, Maharashtra, India

Publication Date: 9 January 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.333

Abstract Urban areas are most dynamic in terms of land use changes. The temporal changes can be monitored through satellite data and reflects the development taking place. Pattern recognition techniques are used to monitor the dynamics. In the present paper, Nagpur city urban sprawl has been studied using band ratioing techniques and fuzzy classification. The result shows an increase in the impervious layer which is the bench mark of urbanization and a corresponding decrease in the vegetation cover and the open spaces. This type of output can be used for planning of industries, garden, rejuvenation of water bodies and monitor the development activities.
Keywords *Remote sensing; Geographic information system; Normalised difference built up index; Fuzzy classification*

1. Introduction

Urbanization refers to the change of rural lifestyles into urban ones (Antroo, 2003). Rapid urbanization and urban expansion may have significant impact on conditions of urban ecosystems (Bouhennache et al., 2005; Turner, 2005; Lopez et al., 2005; Alphan, 2005). The urban conglomerates are of immense importance to mankind. All these urban areas act as centres for growth and development. The rural areas are also of immense importance especially those in the vicinity of the growth centers. These rural or semi urban areas are called as peri urban, and they convert to urban areas in due course of time. Remote sensing technology uses satellite imaging to retrieve information about the changes on Earth. These changes when properly recorded and monitored, give us an insight into the growth pattern, and the suggestive reasons for the drivers for such change. Now-a-days, there are plenty of commercial and non-commercial satellites with different specifications to map Earth surface. The improvements in the satellite data availability and the sensors have given the scientist the opportunity to use different spatial, spectral and temporal data.

Remotely sensed data provide synoptic view, spatial resolution in range of 30 cm and more, multi spectral data, revisit etc. This data is an invaluable source of information for monitoring changes on earth surface and management of resources. Satellite imagery is an effective way to reveal the changes in land use patterns. In order to separate urban and non-urban details, plenty of spectral urban indices have been developed by scientists including Normalized Difference Built-Up Index (NDBI) (Zha et al., 2003), Index-based Built-Up Index (IBI) (Xu, 2008) Urban Index (UI) (Kawamura et al., 1996) and Enhanced Built-Up and Bareness Index (EBBI) (As-syakur et al., 2012).

Image processing software like ERDAS Imagine 2013 is a versatile tool which can be used to formulate new algorithms for delineation of urban areas. Along with the image processing software, GIS software like ARC GIS 9.1 can also be used for conversion of classified image to vector map, topology building and statistic generation. Users can develop their own models using modeler icon and try new methods of feature extraction. In the paper, an attempt has been made to delineate urban areas using various pattern recognition methods like band rationing and fuzzy classification (Mather, 1999) etc. Attempts to improve the classification using supervised classification with fuzzy option has been tried. Manifestation of urban areas on the satellite imageries are in the form of speckled cyan colour sprawls on medium to high resolution satellite data, and for very high-resolution satellite data, the features are almost clearly visible when viewed as false colour composite or as true colour. However, for feature extraction heads up digitization is more useful due to enhanced spatial resolutions in high resolution satellite data.

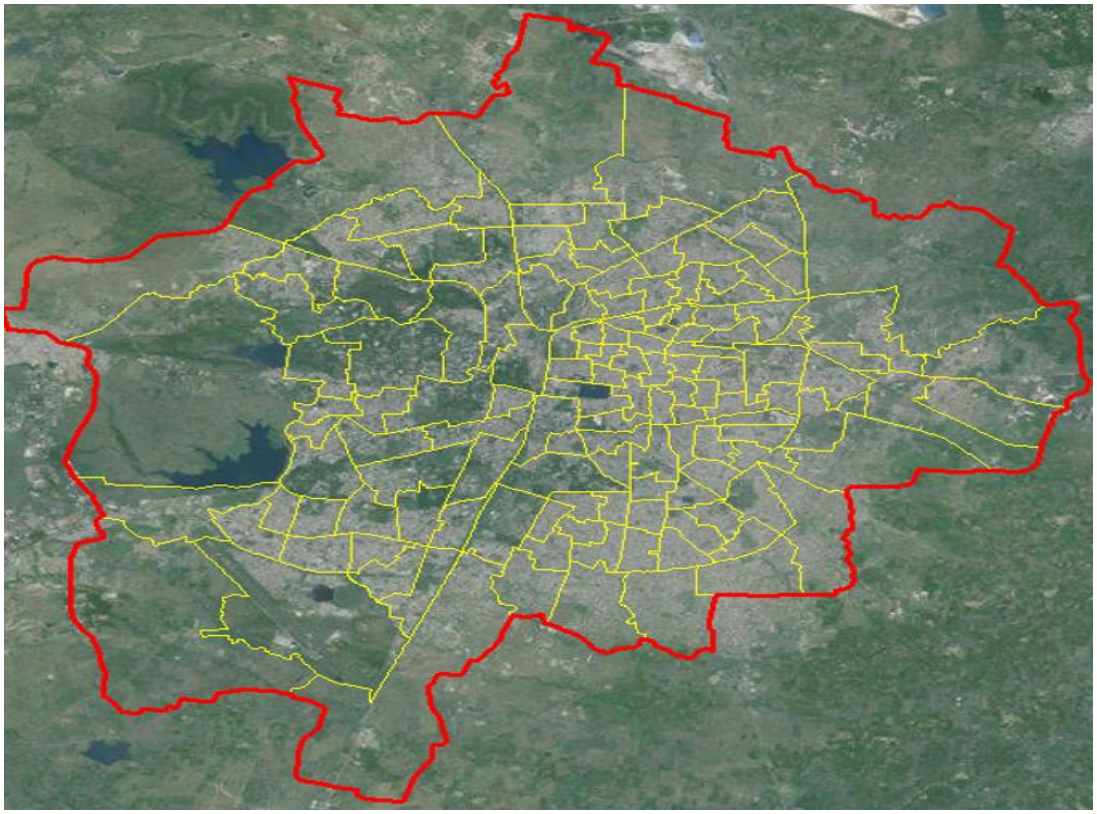


Figure 1: *Nagpur city sprawl made from ward boundary map*

2. Study Area and Satellite Data Used

The study area chosen is Nagpur city urban sprawl. Nagpur city is the winter capital of the state of Maharashtra, India, and the third largest city in the State after Mumbai and Pune. It is the largest city in Central India. Under the scheme of the Govt. of India, to handle the large-scale increase in the urbanization, Nagpur has been chosen under the smart city project and development activities are in progress. It is also one of the most livable cities in India. Thus, an inventory and study to map the changes in the impervious layer is required to ascertain the amount of increase in urbanization. Figure 1 shows the Nagpur city urban sprawl with administrative boundaries called ward boundaries. Satellite Data used is of IRS P6 LISS IV containing the SWIR as well as the NIR. An attempt has been made to

use these two bands for the delineation of urban area. The model has been developed in the "Model Maker" in ERDAS 2013.

3. Methodology

Fuzzy classification in ERDAS IMAGINE 2013 is carried out on LISS IV image. The Normalised Difference Built up Index (NDBI) is used. It is defined as:

$$NDBI = (\lambda_{SWIR} - \lambda_{NIR})/(\lambda_{SWIR} + \lambda_{NIR}) \quad (1)$$

The image so obtained gives a clear demarcation of urban and non-urban areas. The urban and the non-urban areas are separated by masking. The non-urban areas are marked as zero (0) and the urban areas as one (1). A supervised classification is then done with a fuzzy classification option on the urban area. The fuzzy classification consists of a multi-layer classification output file and optional Distance File. Layer 1 contains class values for the best classification, layer 2 with the second best, and so on. The number of class layers to compute is specified in the Best Classes Per Pixel number field. Fuzzy classification works on data that may not fall into exactly one category or another. Fuzzy classification works using a membership function, wherein a pixel's value is determined by whether it is closer to one class than another. A fuzzy classification does not have definite boundaries, and each pixel can belong to several different classes. Using this concept, the urban areas were clearly delineated in the layer 1 as seen in Figure 2. The output clearly shows that the areas of built-up (impervious layer) and water bodies get clearly distinguished.

The classes identified were impervious layer (built up), vegetated areas, open spaces and water bodies. It is clearly seen that the impervious layer consisting of mainly artificial surfaces and structures like pavements, parking lots, buildings etc. that are covered by impenetrable materials like asphalt, concrete, brick, stone and rooftops of houses and buildings are clearly delineated.

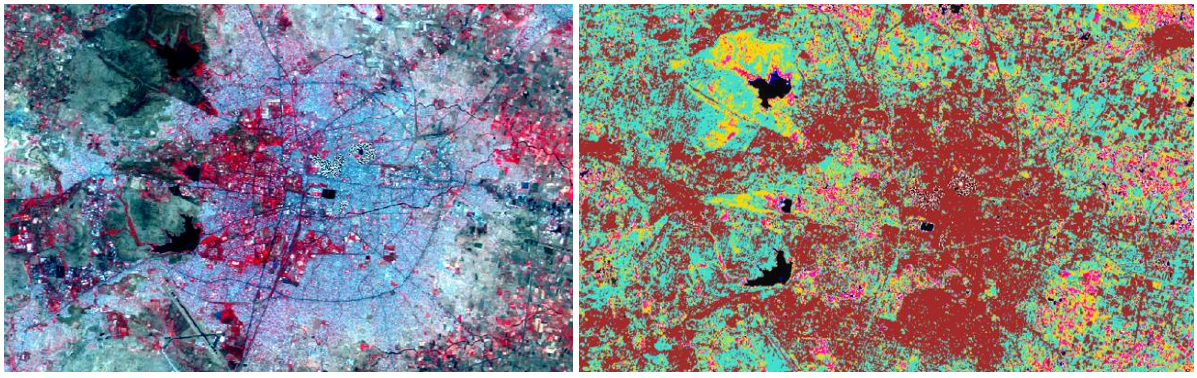


Figure 2: *Nagpur city image and output based on fuzzy classification*

4. Results and Discussion

The temporal changes in the land use / land cover of the three years 2005-06 and 2011-12 and 2015-16 were compared for Nagpur city. The Nagpur city ward boundary map was used to get the outer boundary. The total area of this area of interest (AOI) is 21061.9 Ha. It is found that the area under the built-up class increased significantly over the years from 14458.58 Ha in the year 2005-06 to 15448.83 ha in year 2011-12 and to 15797.98 Ha in 2015-16. The results are shown in the Table 1, Table 2 and Table 3 as temporal changes in Nagpur City (AOI) during 2005-06, 2011-12 and 2015-16.

Table 1: *Urban sprawl in Nagpur city - 2005-06*

LULC Classes	Area in Ha	Area in %
Built Up	14458.58	68.65
Open space	1503.21	7.14
Vegetated	4696.69	22.30
Waterbodies	403.36	1.92
Total	21061.85	100.00

Table 2: *Urban sprawl in Nagpur city - 2011-12*

LULC Classes	Area in Ha	Area in %
Built Up	15448.83	73.35
Open Space	1200.38	5.70
Vegetated	3975.35	18.87
Waterbodies	437.30	2.08
Total	21061.85	100.00

Table 3: *Urban sprawl in Nagpur city - 2015-16*

LULC Classes	Area in Ha	Area in %
Built Up	15797.98	76.77
Open Space	1423.70	5.0
Vegetated	3356.17	15.93
Waterbodies	484.00	2.30
Total	21061.85	100.00

These tables clearly show that there has been an increase in the urban sprawl (impervious layer) which is seen in both the NDBI and output of fuzzy classification. The built up or impervious layer has increased from 68.65 % in 2005-06 to 73.35 % in 2011-12 to 76.77 % in 2015-16. Similarly, in the year 2011-12, as there was very less rainfall, the receding water bodies have contributed to increase in the bare lands. The mixed vegetation part, especially in western Nagpur has dwindled due to the developmental activities. The water availability has also decreased as very few activities are taken up to clean and de-silt the water bodies. The problems of a rapidly developing metropolis are the choked drainages, unplanned growth in urban conglomerates, reduction in the greenness, increase in temperature and pollution. These signs are manifested in Nagpur. Using remote sensing and various pattern recognition techniques, an area of growth in urban classes, presence of bare lands, water bodies can be delineated and used for detailed feature extraction etc.

The present work has shown that the method of supervised classification and fuzzy logic is a very good method of identifying the built up, water bodies, vegetated and open spaces from satellite data.

References

Alphan, H. 2005. Land use change and urbanisation in Adana, Turkey. *Land Degradation and Development*, 14(6), pp.575-586.

Antroo, M. 2003. Landscape change and the urbanisation process in Europe. *Landscape and Urban Planning*, 67, pp.9-26.

As-syakur, A.R., Adnyana, W.S., Arthana, W. and Nuarsa, W. 2012. Enhanced Built-Up and Bareness Index (EBBI) for mapping built-up and bare land in an urban area. *Remote Sensing*, 4, pp.2957-2970.

Bouhennache, R., Bouden, T., Taleb, A.A. and Chaddad, A. 2005. Extraction of urbanland features from T.M. landsat image using land feature index and tasseled cap transformation. *Recent Advances on Electroscience and Computers*, pp.142-147.

Kawamura, M., Jayamana, S. and Tsujiko, Y. 1996. Relation between social and environmental conditions in Colombo, Sri Lanka and the urban index estimated by satellite remote sensing data. *International Arch. Photogramm. Remote Sens. Part B7*, pp.321-326.

Lopez, G., Bocco, M., Mendoza, E. 2005. Predicting land cover and land use change in Urban Fringe, a case in Morelia City, Mexico. *Landscape and Urban Planning*, 55, pp.271-285.

Mather, P. 1999. *Computer Processing of Remotely Sensed Images, An Introduction*, 2nd edition. John Wiley & Sons, Ltd.

Turner, B.L. 2005. Local faces, global flows: the role of land use and land cover in global environmental change. *Land Degardation and Development*, 5, pp. 71-78.

Xu, H. 2008. A new index for delineating built up land features in satellite in satellite imagery. *International Journal of Remote Sensing*, 29, pp.4269-4276.

Zha, Y., Gao, J. and Ni, S. 2003. Use of normalised difference builtup index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24, pp.583-594.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2540-2550
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.342

Case Study

Hydro Geomorphological Characteristics and Delineation of Ground Water Potential Zone - A Case Study of Rushikulya and Bahuda Basin, Ganjam Odisha

P.C. Sahoo[1], P.K. Panda[2], K.C. Sahu[1], D.S. Pattainak[3]

[1]Department of Oceanography, Berhampur University, Berhampur, Odisha, India
[2]Department of Civil Engineering, Centurion University of Technology and Management, Rajaseetapuram, Odisha, India
[3]Department of Earth Science, Utkal University, Bhubaneswar, Odisha, India

Publication Date: 27 February 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.342

Abstract The present study attempts to delineate different groundwater potential zones using remote sensing and geographic information system (GIS) in Rushikulaya Bahuada Basin of Ganjam district, Orissa. Thematic maps of Geology, Hydrology, geomorphology, land use and land cover, drainage density, were prepared using the Landsat Thematic Mapper data. Relationship of each layer to the groundwater regime has been evaluated. The major hydro geomorphic units identified in the area are, weathered denudation hills, residual hills, Pediments, spit, Valley Fills, beach ridges, Alluvial Palin, flood plain. Most part of the study area is occupied by alluvial plains with various thicknesses and the ground water potential is directly related to thickness of alluvial plain. Field observations showed that ground water occurs under unconfined conditions with water table at shallow to deep depth. From the lineament map, the lineament density and lineament intersection maps prepared to understand the impacts on groundwater percolation. Finally, the hydrogeomorphology and Lineament maps are overlaid following the weighted index overlay method, which delineates groundwater potential zones. An integrated remote sensing and Geographic Information System (GIS) based approach has been used for demarcating groundwater potential zones in the study area.
Keywords *Geomorphology; Ground water; Remote sensing*

1. Introduction

Ground water is a vital natural resources for the reliable and economic provision of potable water supply in both urban and rural environment. It plays significant role in human well beings, as well as some the aquatic and terrestrial ecosystem. At present, ground water contributes 34% of the total annual water supply and is an importance to fresh water resources; so it as an urgent need to study the ground water assessment using the recent technologies such as Remote sensing and GIS. In the recent years it has been observed that extensive use of satellite remote sensing has made easier to define the spatial distribution of different ground water potential zones based on geomorphology, and its associated features. Remote sensing techniques are also useful for ground water exploration, especially delineation Hydrogeomorphological units. Complex geology and geomorphic set up, skewed distribution of rainfall result in diverse hydrogeological conditions in an area. Ground water occurrence being subsurface phenomenon, its identification and location is based on indirect analysis of some

directly observable terrain features like geological structures, geomorphic features and their hydrologic characters (Das et al., 1997). GIS and satellite remote sensing have credibility to obtain various features related to land and water. GIS techniques also facilitate integrated and conjunctive analysis of large volumes of multi-disciplinary data sets include spatial and non-spatial, with the same georefrencing scheme. To obtain the ground water potential zone, it is necessary to integrate data from the various terrain characteristics such as topography, lithology, geological structures, and geomorphology, weathering depth, slope characteristics and drainage pattern using geographical information system techniques (Jones, 1986; Sinha et al., 1990; Chi and Lee, 1994; Bahuguna et al., 2003; Kumar et al. 2007). The present study focused delineation of ground water potential zone based on integrated study using all the controlling parameters such as geomorphology, drainage, slope characteristics, geological structure.

1.1. Study Area

The study area, a Coastal tract is situated on the eastern margin of Ganjam District (Figure 1). It lies between latitudes of 19°05' - 19°30' N and longitudes of 84°35' - 85°05' E and cover all area of 350 sq. km. The width of the coastal tract varies from 4 to 6 km and extends for a length of about 45 km. It is bound by Rushikula River on north and Bahuda Estuary on south and Bay of Bengal on the east.

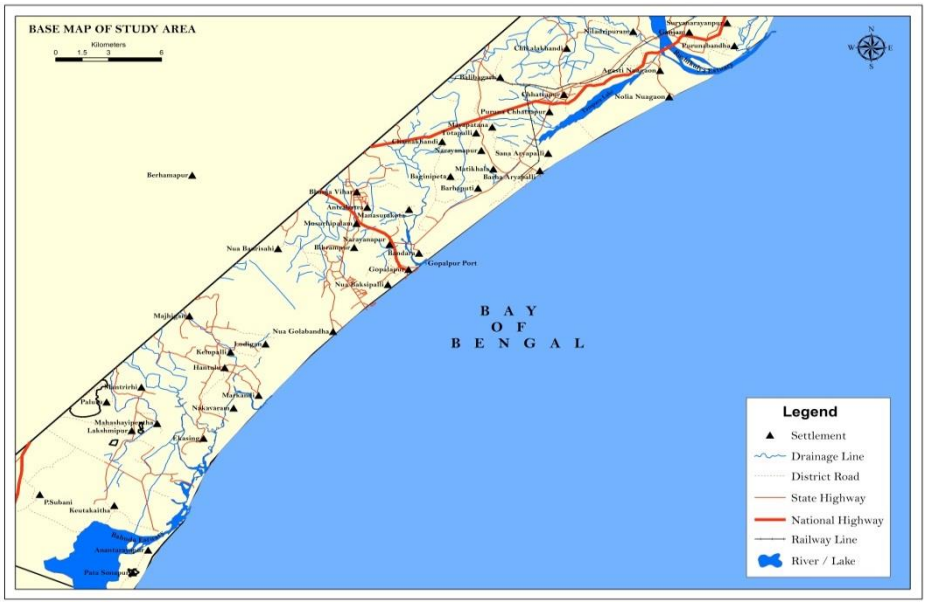


Figure 1: *Base map of study area*

1.2. Geology

Major parts of study area are underlain by hard crystalline rocks of Archaean age. Rushikulya Bahuda River occupies recent to sub recent sediments along the costal track. It has been observed sometimes laterites occur as a capping over khondalites. The Archaean crystalline of the Eastern Ghat Group comprises of Granite and Granite gneiss, Khondalite suite, Charnockite suite, Pegmatites and Quartz vein. Figure 2 shows Geology map the study area.

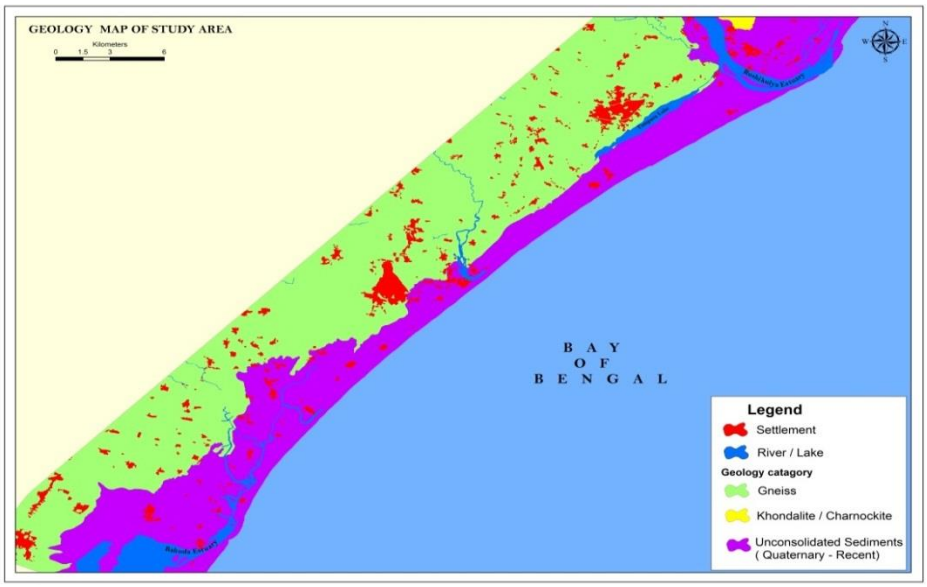


Figure 2: *Geology map of study area*

1.3. Geomorphology

The study area presents a unique geomorphological setting with highly rugged mountains region in the West, North, Central, and northern part and with dense forest covered and gently undulating plains and some isolated hillocks in the eastern part. The altitude in the hilly terrain ranges between 300m to 700m above the mean sea level. The low laying area which includes the flood plain of the river Rushikulya and the intermountain valleys are characterised by undulating topography, scattered hill. The major geomorphic units are classified as Structural hills, Denudational hills, Residual Hills, Inselberg, pediment, buried pediplain, flood plain, coastal plain and sandiness (Tripathy et al., 1996).

1.4. Hydrogeology

With respect to water bearing properties of different geological formations play an important role in the occurrence and movement of ground water. In the present study area the crystalline rock of Archean age which is occupies more than 90% of the total area. The patches of recent to sub recent alluvium along the river courses occupy 374 sq.km areas. As per hydrology is concerned the weathered and fractured zones of the crystalline and porous alluvium and costal deposits contains ground water in the district depending upon the water yielding properties of various formation, it can be grouped into 3 hydrological units such as Consolidated formation, Semi consolidated formation and Unconsolidated formation.

1.5. Structure

The study area falls in the parts of the Eastern Ghats. These rocks have undergone poly phase deformation revealed by structural features like Joints, floods, fault and foliation etc. Foliations are well developed in khondalites and granitic rocks, crude foliations are also common in charnockite. From the structural analysis it has been observed that Landsat imagery revealed five major tectonic events represented by NE-SW, ENE-WSW, N-W, NW-SE, and NNE-SSW tectonic pattern in chronological order (Krishan, 1992). The major NE-SW, NE-SE, N-S and E-W lineaments are closely related to the fold movements.

2. Material and Methods

2.1. Data Collection

The landsat thematic mapper TM data has been used for the present study. Survey of India toposheet (No.74E/3, 74A/15 and 74A/16) at 1:50,000 scales have been used. Hydrology and ground water well data has been collected from Central Ground Water Board, Bhubaneswar and supporting field data collected through VGIS. ARC GIS and ERDAS have been used for interpretation, analysis and mapping of the individual layers.

2.2. Satellite Data Analysis

In order to produce the thematic map a step by step procedure has been followed to do analysis and interpretation from the satellite data. The processing has been done the various digital images processing such as Enhancement, filtering, classification and followed by GIS analysis. (ESRI-ArcView GIS, 1996). Subsequently, Selective field checking was carried out.

2.3. Spatial Database Building

All the primary and secondary data brought in to the GIS data base. Digitization of all maps and collateral data followed by the transformation and conversion from raster to vector, gridding, buffer analysis interpolation and analysis were done in GIS. In this stage the layers has been created such as Geomorphology, Drainage, Drainage density, lithology, surface water body, lineament and lineament density and slope etc.

2.4. Satellite Data Interpretation

The final stage involved combining all thematic layers using the method that are modified from DRASTIC model. This model is used to assess ground water pollution vulnerability by Environmental Protection Agency of the United State of America (Aller et al., 1985). The output has been classified into five groups such as very high, high, moderate, low and very low using the quintile classification method (ESRI-ArcView GIS, 1996). Groundwater potential index (GWPI) in the present study has been determined by the formula as shown below (Nath et al., 2000):

GWPI = (GgwGgr + DgwDgr + DdwDdr + LhwLhr + LtwLtr + LdwLdr + WwwWwr + SswSsr)/ Total Weight.

Where: Gg: geomorphology, Dg: drainage, Dd: drainage density, Lh: lithology, Lt: lineament, Ld: lineament density, Ww: surface water body, Ss: slope. With “w” representing weight of a theme and “r” the rank of a feature in the theme. GWPI is a dimensionless quantity that helps in indexing the probable groundwater potential zones in an area. It is classified as Excellent GP (>7.1), Very good GP (5.8-7.1), Good GP (4.4-5.8), Moderate GP (3.0-4.4) and Poor GP (<3.0).

3. Results and Discussion

3.1. Generation of Thematic Maps

Lineaments have been identified from Landsat TM and also water bodies interpreted. SRTM digital elevation model is also used preparing the slope map along with the Survey of India toposheet. These

mentioned above, there data sets are useful for delegation of the ground water potential zones coupled with limited field study.

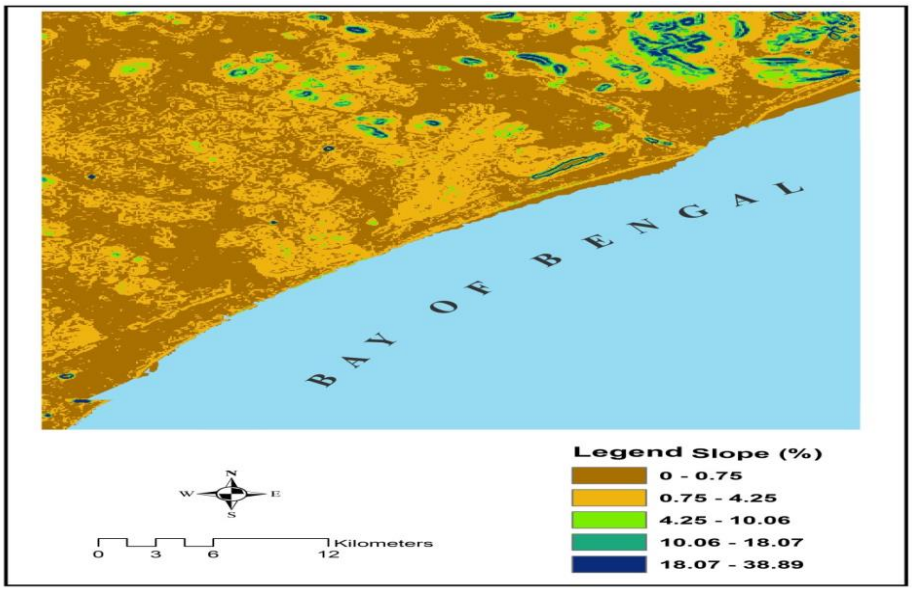


Figure 3: Slope map (%) of the study area (*Biswas Arkoprovo et al., 2012*)

3.2. Lithology Map

Geology or lithology is a very important aspect to identify and predict ground water potential zones. From the visual interpretation skill extraction of geological information from satellite data depends on the identification of different patterns on an image resulting from the spectral arrangement of different tones and textures. Depending on the rock reflectance properties, satellite images are used and they play important role in rock identifications. A lithology map is prepared using the IRS IC and Landsat TM Digital Data and simultaneously ground check verification is done in field (Biswas Arkoprovo et al., 2012) (Figure 4).

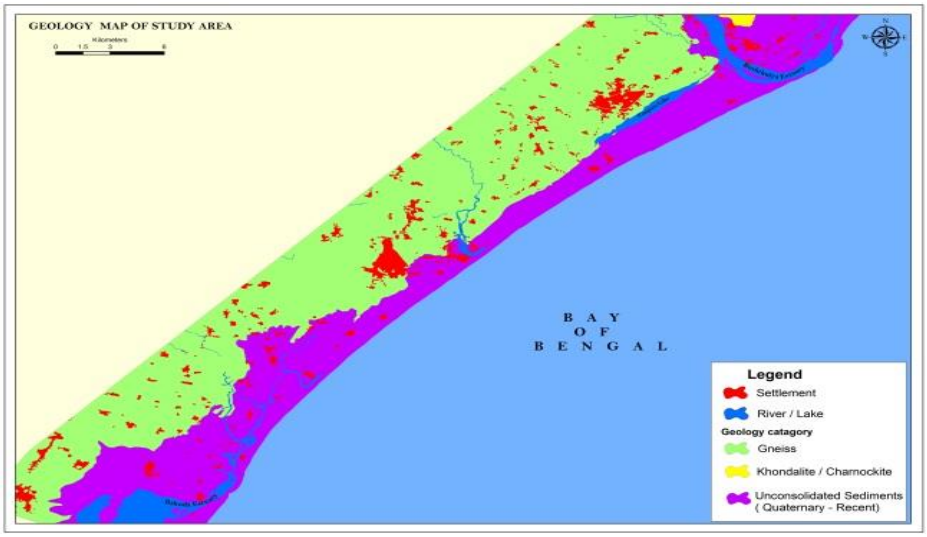


Figure 4: *Geology map of study area*

3.3. Geomorphology Map

Geomorphic features have been identified using visual interpretation from Landsat TM on scale of 50,000 and hydrogeological characteristics have been identified. The different geomorphological elements (Figure 5) such as Flood plain, pediment, costal plain, beach rides and sandiness identified (Biswas Arkoprovo et al., 2012)

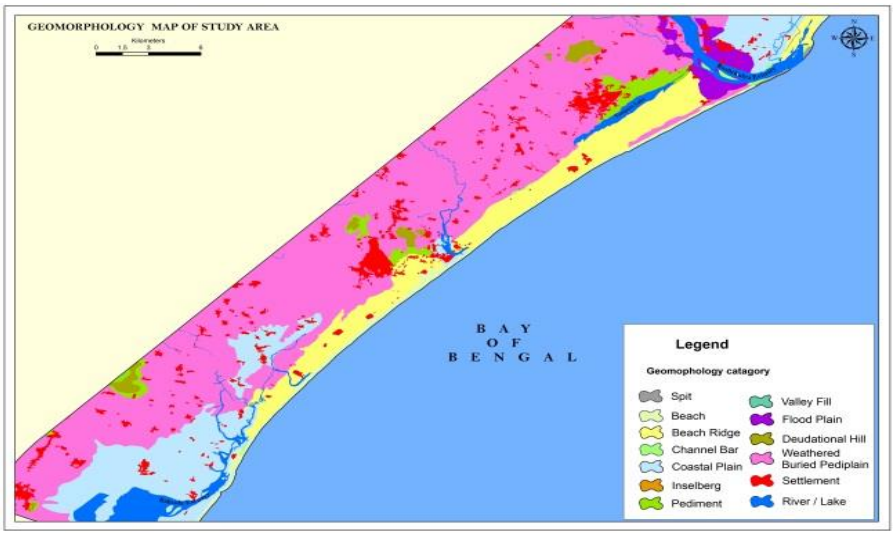


Figure 5: *Geomorphology map of the study area*

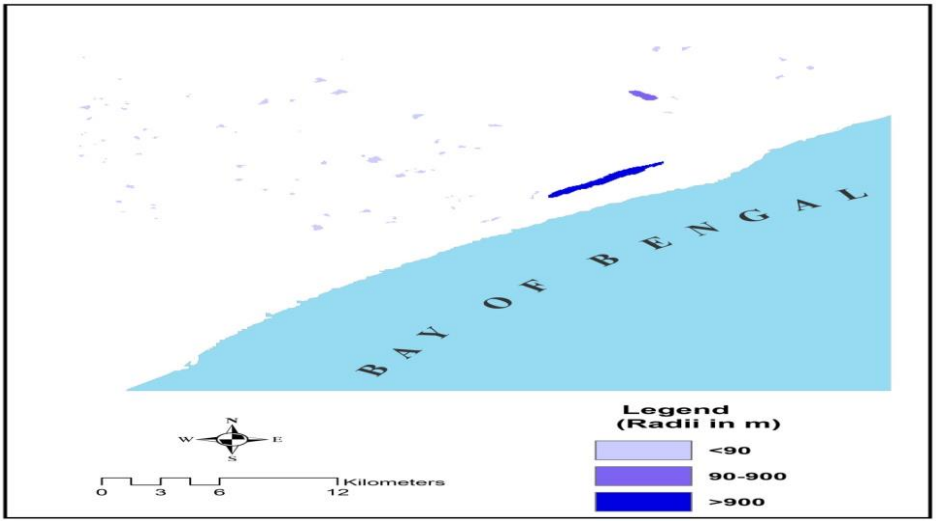


Figure 6: *Surface water body map (Biswas Arkoprovo et al., 2012)*

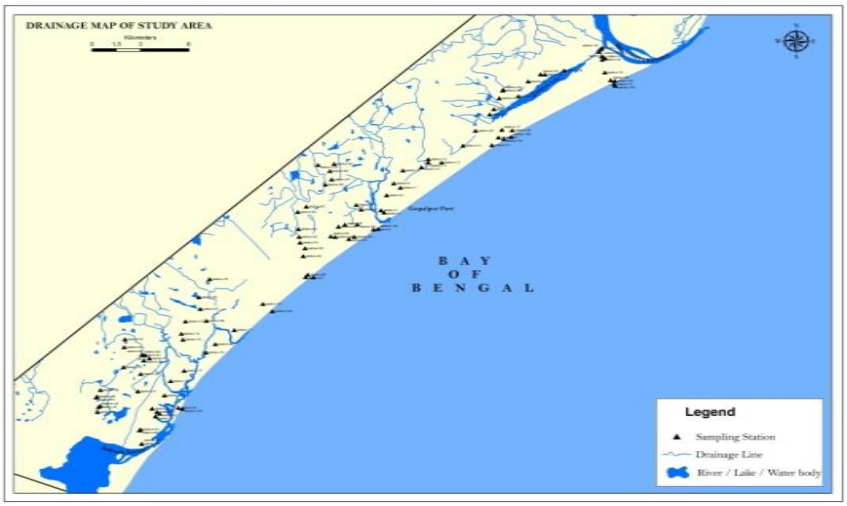


Figure 7: *Drainage map of study area*

3.4. Slope Map

Using the SOI Toposheet and SRTM data of the area, a slope map of the area is prepared. The area, in general is very gentle slope. Slope always plays a crucial role in groundwater potential mapping. However, in the north-eastern part of the study area, there is an increase in slope. Despite of this very gentle slope, a slope map has been prepared according to the following class interval (Biswas Arkoprovo et al., 2012) (Figure 3).

3.5. Surface Water Body Map

Surface water body map is prepared from Landsat TM Satellite data. The supervised maximum likelihood classification is used to delineate the surface water body present in different parts of the area. The study area is covered by many water logged bodies, ponds, lakes and rivers (Figure 6).

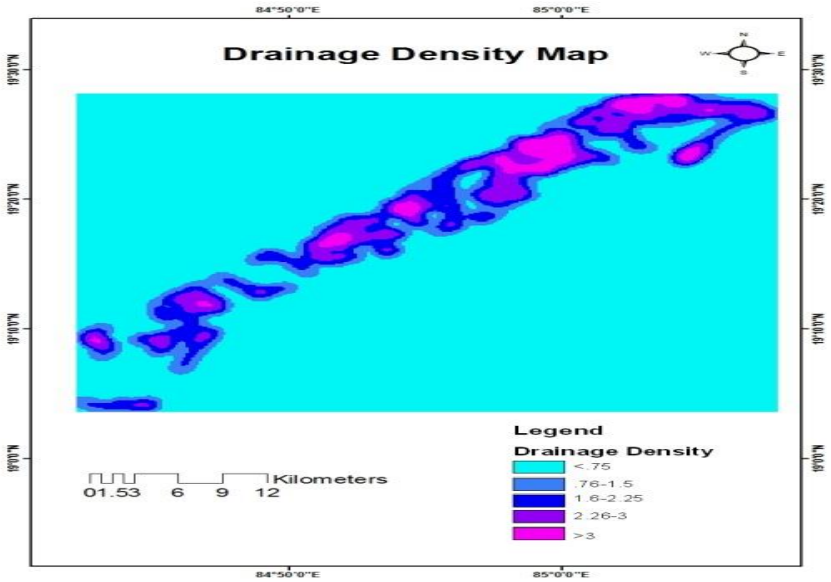


Figure 8: *Drainage density map*

3.6. Drainage and Drainage Density Map

A surface drainage map is prepared from SOI Toposheet at 1:50,000 scale and satellite data. The study area is covered by Rushikulya and Bahuda River. Mostly the drainage pattern is dendritic in nature but locally it exhibits structurally controlled area (Figure 7). From the above Drainage map, a density map is generated. In this density map, the values have been assigned depending on the density of the drainage pattern (Figure 8).

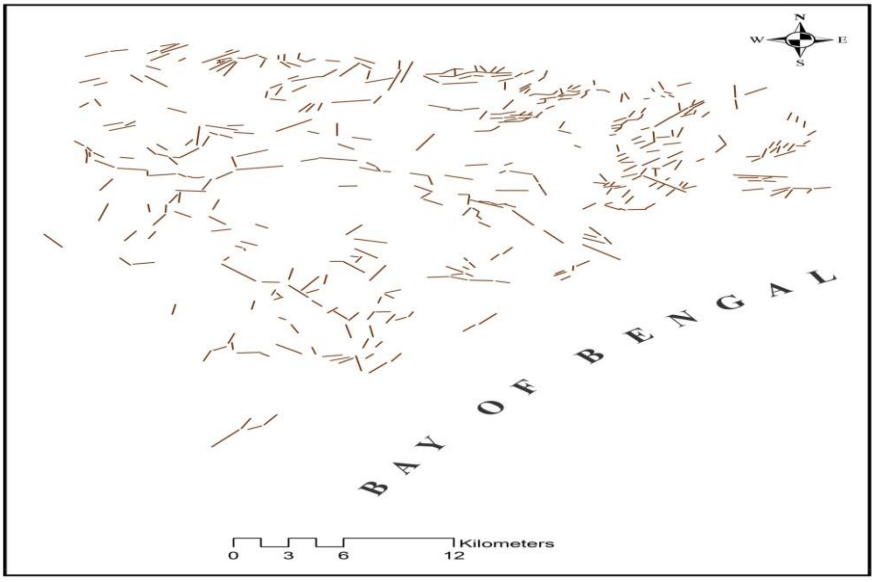


Figure 9: *Lineament map of the study area (Biswas Arkoprovo et al., 2012)*

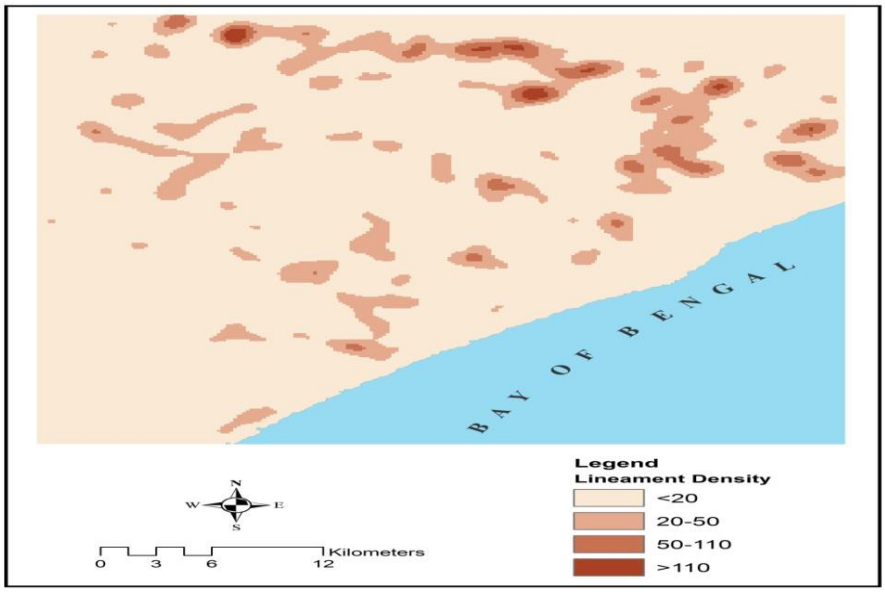


Figure 10: *Lineament density map (Biswas Arkoprovo et al., 2012)*

3.7. Lineament and Lineament Density Map

A surface lineament map is prepared from the SRTM DEM data. The study area exhibits a structurally controlled region. The areas have much high altitude hilly structures which are linearly aligned (Figure 9). From the above Lineament map, a lineament density map is generated. The lineament density map shows a low density in most of the area comparatively to other parts of the study area (Biswas Arkoprovo et al., 2012) (Figure 10).

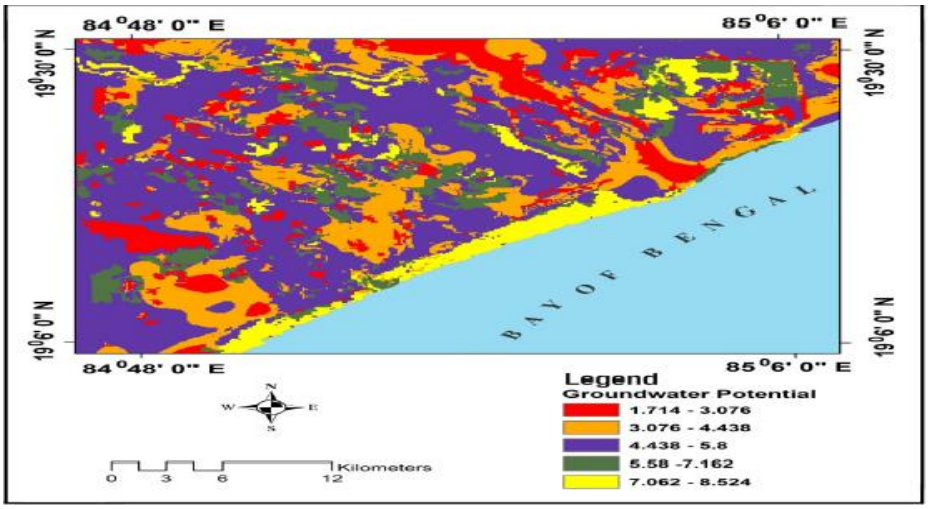


Figure 11: *Ground water potential map (Biswas Arkoprovo et al., 2012)*

3.8. Integration of Thematic Layers and Modelling through GIS

Weighted Index Overlay Model

Depending on the groundwater potentiality, each class of the main eight thematic layers (geomorphology, lithology, slope, drainage density, lineament density and surface water body) are qualitatively placed into one of the following categories viz., i.e. Excellent, ii. Very good, iii. Good, iv. Moderate, v. Poor. Suitable weightage on a scale of 'Six' has been given to each class of a particular thematic layer based on their contribution towards ground water potentiality. The rank of each thematic map is scaled by the weight of that theme. All the thematic maps are then registered with one another through ground control points and integrated step by step using normalized aggregation method in GIS for computing groundwater potential Index of each feature. All the thematic maps have been integrated using GWPI formula in GIS. A final groundwater potential map (Figure 11) is prepared based on the above technique. In the present study, the groundwater potential zones have been categorized into five types' viz., excellent, very good, good, moderate and poor (Biswas Arkoprovo et al., 2012).

4. Conclusion

From the above study it has been observed that that remote sensing and GIS can provide the appropriate platform for convergent analysis of large volume of multi-disciplinary data and decision making for groundwater studies. The Weighted index overlay model has been found very useful in the mapping of groundwater prospective zones. Mainly IRS IC was used for preparing the thematic maps viz., lithology, geomorphology, drainage. Landsat TM was useful in demarcating the water bodies and also in identifying numerous lineaments. SRTM was also used in preparing DEM from which slope

map is prepared. In this study area, five categories of groundwater potential zones have been delineated based on remote sensing and GIS technique. The five categories are: excellent, very good, good, moderate and poor. The groundwater potential map (Figure 11) near the coast line and in some parts of the north-eastern and north-western region shows excellent potential whereas greater part of the area shows good groundwater potential. Poor groundwater potentials are confined mostly in the hilly terrain and in settlement area. The quantitative amount of groundwater potential zones can be carried out if geophysical resistivity data is available.

References

Aller, L., Bennett, T., Lehr, J.H. and Petty R.J. 1985. *DRASTICA standardized system for evaluating ground water pollution potential using hydrogeological settings*. U.S. Environmental Protection Agency report EPA/600/2-85/018.

Antony Ravindran, A. 2012. Azimuthal square array resistivity method and groundwater exploration in Sanganoor, Coimbatore District, Tamilnadu, India. *Research Journal of Recent Sciences*, 1(4), pp.41-45.

Biswas Arkoprovo, Jana Adarsa and Sharma Shashi Prakash. 2012. Delineation of Groundwater Potential Zones using Satellite Remote Sensing and Geographic Information System Techniques: A Case study from Ganjam district, Orissa, India. *Research Journal of Recent Sciences*, 1(9), pp.59-66.

Central Ground Water Board (CGWB). 1995. Report on the groundwater resources and developmental potentials of Ganjam District, Bhubaneswar, Orissa.

ESRI-ArcView GIS. 1996. The Geographic Information System for everyone, Environmental Systems Research Institute, USA.

Frazier, P.S. and Page, K.J. 2000. Water body detection and delineation with landsat TM data. *Photogram. Eng. Rem. Sen.*, 66(10), pp.1461-1467.

Jensen, J.R. 1986. Introductory digital image processing. Third Edition. Englewood Cliffs, NJ: Prentice Halls, p.544.

Khan, M.A. and Moharana, P.C. 2002. Use of remote sensing and geographical information system in the delineation and characterization of ground water prospect zones. *Journal of the Indian Society of Remote Sensing*, 30(3), pp.131-141.

Krishnamurthy, J., Arul Mani, M., Jayaraman, V. and Manivel, M. 1997. *Selection of sites for artificial recharge towards groundwater development of water resource in India*. Asian Conference on Remote Sensing, Kuala Lumpur.

Krishnamurthy, J., Venkatesa, K.N., Jayaraman, V. and Manivel, M. 1996. An approach to demarcate ground water potential zones through remote sensing and a geographical information system. *International Journal of Remote Sensing*, 7(10), pp.1867-1884.

Krishnan, M.S. 1982. *Geology of India and Burma*. Sixth Edition. Delhi: CBS Publishers and Distributors, p.536.

Lillesand, T.M. and Kiefer, RW. 2000. *Remote sensing and image interpretation*. Fifth Edition, Singapore: John Wiley and Sons (Asia) Pte. Ltd, p.820.

Manimaran, D. 2012. Groundwater geochemistry study using GIS in and around Vallanadu Hills, Tamilnadu, India. *Research Journal of Recent Sciences*, 1(6), pp.32-37.

Mohanty, B.K. and Devdas, V. Geological mapping of Quaternary formations in Rushikulya river basin in parts of Ganjam District, Orissa, *Rec. Geol. Surv. India*, 122(3), pp.5-6.

Narayanaswamy, S. 1975. Proposal for charnockite khondalite system in the Archean shield of Peninsular India. *Misc. Geol. Soc. India*., 23(203), pp.1-16.

Nath, S.K., Patra, H.P. and Shahid, S. 2000. *Geophysical prospecting for groundwater*. New Delhi: Oxford and IBH Publishing Co. Pvt. Ltd, pp.145-152.

Obi Reddy, G.P., Mouli, K.C., Srivastav, S.K., Srinivas, C.V. and Maji, A.K. 2000. Evaluation of groundwater potential zones using remote sensing data - a case study of Gaimukh Watershed, Bhandara District, Maharashtra. *Journal of the Indian Society of Remote Sensing*, 28(1), pp.19-32.

Phukan, I., Ravindran, K.V. and Banerjee, D.M. 1999. *Delineation of groundwater prospect zones using remote sensing and geographic information system in the West Garo Hills District, Meghalaya*. "Geoinformatics: Beyond 2000", an International Conference on Geoinformatics for Natural Resource Assessment, Monitoring and Management, Jointly Organized by Indian Institute of Remote Sensing (National Remote Sensing Agency) Dehradun, India and ITC (International Institute for Aerospace Survey and Earth Sciences) The Netherlands, pp.203-211.

Sankar, K. 2002. Evaluation of groundwater potential zones using remote sensing data in Upper Vaigai River Basin, Tamil Nadu, India. *Journal of the Indian Society of Remote Sensing*, 30(3), pp.119-129.

Saraf, A.K. and Choudhury, P.R. 1998. Integrated remote sensing and GIS for groundwater exploration and identification of artificial recharge sites. *International Journal of Remote Sensing*, 19(10), pp.1825-1841.

Srivastav, P. and Bhattacharya, A.K. 2000. Delineation of groundwater potential zones in hard rock terrain of Bargarh District, Orissa using IRS. *Journal of the Indian Society of Remote Sensing*, 28(2-3), pp.129-140.

Todd, D.K. 1980. *Groundwater Hydrology*. Second Edition. New York: John Wiley and Sons Inc., p.556.

Tripathy, J.K., Panigrahy, R.C. and Kumar, K.V. 1996. Geological and Geomorphological studies of a part of Ganjam district, Orrisa by remote sensing techniques. *Journal of the Indian Society of Remote Sensing*, 24(3), pp.169-177.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2643-2655
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.359

Research Article

Investigating Image Fusion Techniques on CHRIS/Proba Space Borne Hyperspectral Data for Material Identification

Veeramallu Satya Sahithi[1], Iyyanki V. Murali Krishna[2]

[1]Jawaharlal Nehru Technological University, Hyderabad, India
[2]Research Centre Imarat, Defence Research and Development Organization (DRDO), Hyderabad, India

Correspondence should be addressed to *sahithi.geo@gmail.com*

Publication Date: 31 May 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.359

Abstract Image fusion can be defined as the process of combining information from two images with unique characteristics to obtain a resultant image which has an assemblage of qualities of both the inputs. Of late, many studies were carried out in fusing multispectral data with panchromatic data, but fusion of hyperspectral data with multispectral data is a topic of interest till date. In the present study, image fusion technique was used to obtain a spatially and spectrally rich hyperspectral (HS) image which can aid in improved material identification. Two conventional pixel basedfusion techniques namely – Gram Schmidt (GS) fusion and High Pass Filtering (HPF) technique were used for fusing the HS CHRIS image with the multispectral (MS) LISS IV image. An overall spectral, spatial/visual and quantitative analysis was carried out to examine the quality offused images. Quantitative analysis of the individual bands was made using SNR and entropy measures. Classification of the fused images helped in identifying five different concrete materials and four different vegetation types within the study area. An overall classification accuracy of 88.33% was obtained using GS fused image, 79.25% with HPF fused image and 73.66% with the CHRIS data.
Keywords *Classification; GS method; HS data fusion; HPF; Quantitative analysis*

1. Introduction

The space borne remote sensing sensors provide data covering different regions of the electro-magnetic spectrum at different spatial and spectral resolutions, where the spatial resolution represents smallest resolvable area (e.g. pixel) and the spectral resolution corresponds to smallest wavelength that can be detected in spectral measurement (Lillesand and Kiefer, 2000; Shashi, 2009). The presently available data from space borne hyperspectral (HS) sensors (like Hyperion, CHRIS, AVIRIS) have a medium spatial resolution (ranging between 17m-30m) with rich spectral information (8-20 nm). Also, the high spatial resolution multispectral (MS) sensors like LISS IV, Quick bird, World View-2, Geo eye etc provide data at high to very high spatial resolution but with a limited spectral resolution (100-150 nm). Both these HS and high resolution MS sensors have their own importance in identification and in delineation and classification of various land use/land cover classes. But, due to the system limitation of the sensors, it is not possible to acquire a HS image with high spatial resolution (Pande et al., 2013). Although airborne HS sensors can give these kind of high spatio spectral images, these have constraints in terms of cost and security. Hence, an alternative to this can be image fusion

technique that can combinethe properties of spectrally rich HS image and spatially rich MS image which yields inincreased interpretation capabilities and improved material identification.

Image fusion can be defined as the process of dealing with data and information from multiple sources to achieve refined/improved information for decision making (Hall, 1992). Many studies were made based on the fusion of MS and panchromatic data that have highlighted the advantages and importance of fusion (Scarp, 2014; Alparone, 2008). However fusion of MS and panchromatic images can improve the qualities of output image in spatial context, but cannot improve the spectral properties of fused output (Luciano, 2008). Hence, in the present work, we have attempted to fuse the spectrally rich HS image with a spatially rich MS image. Fusion of HS images with high resolution MS images may result in an image with high spectral and spatial resolution in which properties of both of the input images could be preserved. Thus, the former can support better delineation of features spectrally while the later can help in identification of features spatially. This may also help in bringing out some inherent information that cannot be identified in the HS and MS datasets when considered alone. The input images used for fusion should satisfy certain conditions in order to obtain good results upon fusion. These include - a)accurate co- registration, b)same temporal characteristics- identical time or season, c) identical illumination conditions, d) common/overlapping spectral range for both input images. There are various image fusion techniques available which can be used for fusing MS and panchromatic datasets likesimple averaging, Principal Component (PC) sharpening, Brovey transform, Gram Schmidt fusion, Color Normalized technique, Ehler's fusion (Abdikan et al., 2014), Wavelet based fusion, High Pass Filtering (HPF) method etc., each technique having its own advantages and limitations. Gomez (2001) made a research on the wavelet based fusion of HS and MS datasets. Not all algorithms used for MS datasets are suitable for fusing HS data with MS data. Only some of the transformation based algorithms like Intensity Hue Saturation (IHS) (Witharana et al., 2014), Wavelet decomposition, Neural Networks (Vivone et al., 2014), Knowledge-based image fusion, Gram Schmidt (GS) technique (Dahiya et al., 2013), Color Normalized Transform (CNT) etc. (Darvishi et al., 2005) can help in fusing HS with MS datasets. Solanki and Patel (2011) stated that, fused image using pixel-based technique exhibits good contrast and works well for input images with similar levels of contrast. Shashi (2009) made a comparison of three different fusion techniques for fusing the Hyperion and Ikonos datasets and concluded in her work that after fusion there is remarkable spatial enhancement of features in the output image when compared to the Hyperion but not much gain was noticed when compared to Ikonos data.

Image classification helps in obtaining an accurate LULC /thematic map when provided with good training samples and appropriate classification parameters. The conventional MS classifiers like Maximum Likelihood and Minimum Mistance to Mean etc are not suitable for classifying the HS images due to its high volume and data redundancy. Various improved classification techniques like spectral angle mapper, artificial neural networks, spectral unmixing, support vector machine etc. were developed for classifying HS images. In the present work, the most widely used Support Vector Machine (SVM) algorithm was used to obtain a classified map from fused images.

The main objective of the present study is to integrate the spatial information from the high spatial resolution MS data and spectral information from HS data for improved material identification. Further analysis includes spatial and spectral analysis of the fused images, quantitative analysis of the fused outputs, and classification of the images for identifying various materials. Two different pixel based fusion methods were used which includeGram Schmidt transform based fusion and HPF (High Pass Filter) fusion technique. The Gram Schmidt technique is an improved version of PCA transform, and the HPF fusion is based on wavelet transform. A brief review of these two techniques is presented below.

Algorithms

HPF Fusion

The HPF based fusion method can be performed with much smaller computation time and data space requirements than when compared to the other image transformation methods. The general steps in algorithm include (Jawaketal., 2013)

(i) Calculating a ratio of MS cell size to high-resolution cell size (taken as R).
(ii) High-pass filtering of high spatial resolution image.
(iii) Resampling multi-spectral image to pixel size of high-pass image.
(iv) Adding HPF image to each multi-spectral band.
(v) Stretching new multi-spectral image to match statistics of the original multi-spectral image.

The size of the high-pass kernel (HPK) is a function of the relative input pixel sizes, R. The HPF image is weighted relative to global standard deviation of multi-spectral bands and is also a function of R. The value of weight applied to HPF image, prior to addition to multi-spectral image, depends on both R and standard deviations (SD) of both HPF image and multi-spectral band it is being merged with. The weight (W) is determined by the formula (Ehlers and Xu, 2013).

$$W = (SD(MS) / SD(HPF) \times M) \text{ -------} \quad (1)$$

$$\text{Pixel (out)} = [\text{Pixel (in)}] + [HPF \times W] \text{ ------} \quad (2)$$

Where,

W = weighting multiplier for HPF image value; SD(MS) = standard deviation (SD) of the MS band to which the HPF image is being added; SD(HPF) = standard deviation (SD) of the HPF image and M = modulating factor to determine the crispness of the output image.

Gram Schmidt Sharpening Technique

The Gram Schmidt algorithm is similar to Principal Component Analysis algorithm. In the first step, a simulated low-resolution pan band is generated from existing MS bands. The second step is to perform a Gram-Schmidt transformation on simulated panchromatic band and multi spectral bands, using the simulated panchromatic band as the first band. In this, starting with the simulated pan band as the first vector, all bands are treated as high dimensional vectors and are made orthogonal using the Gram Schmidt vector orthogonalization. In Gram Schmidt transformation, first the angle between red band and pan band is calculated, and then red band is rotated to make it orthogonal to pan band. A similar process is executed with the green band and pan and green band is rotated so that it is made orthogonal to both the red and pan bands. Thus, the Gram-Schmidt forward transform helps in de-correlating the bands (Maurer, 2013). In the next step, the simulated low resolution pan band is replaced by high-resolution pan band and an up-sampling is performed to spatial resolution of pan band. Finally, an inverse Gram-Schmidt transform is applied on up-sampled bands using the same transform co-efficients. The resultant output from this transform gives a pan sharpened image. (Maurer et al., 2013; Despini et al., 2014).

$$Pan_{Sim} = \sum_{k=1}^{n} w_k \, MS_k$$

Where,

k is the band number; W is the weight factor; M is the Covariance matrix and S is the simulated pan band.

Support Vector Machine Classifier

SVM is a classification technique based on the statistical learning theory. This technique is said to be independent of dimensionality of the feature space. The main idea behind this classification is to separate the classes with a surface that maximizes the margin between them, using boundary pixels (Gualtieri et al., 1999). The surface thus generated is termed as a decision plane/hyper plane (Melgani et al., 2004). The data points that are close to the hyper plane are termed 'support vectors'. The hyper planes always need not be straight lines, as it is not possible in real world classification problems. SVM's were initially applied on binary classification problems where a linear separating hyper plane is used (Pal M. et al., 2004). Later on, the multi class SVMs were developed where the data is projected on to a higher dimensional feature space to construct the decision planes (Pal M. et al., 2005). Multiclass SVMs use certain kernels - either linear or non-linear to solve the classification problems. The performance of the support vector mainly depends on the kernel types used (Waske et al., 2010). SVMs are known for their better performance in classifying complex and noisy data. In practical considerations, the major advantage of support vector machine classifier is that even a single training pixel for a particular class would be enough for the SVM to classify the corresponding matching spectra in the image to that class.

Foody and Mathur (2004) made a comparison between different classification methods viz. discriminant analysis, decision trees, feed forward neural networks and SVM classification and concluded in his results that SVM yielded highest accuracy than the other classifiers. Shi D. and Yang X. (2012) and Candade N. and Dixon B. (2003) showed in their results that SVM classifier outperformed the MLC and other classifiers. Dixon (2004) in his study identified that SVMs show a better accuracy even with a small number of training samples and don't have the problem of over-fitting, local minima and sensitive to dimensionality of the data.

Study Area and Datasets

The study area is a part of Suratgarh airbase station located in Sri Ganganagar district, Rajasthan, India. It falls in the latitude and longitude extent of 29˚24'49"N to 29˚22'24"N and 73˚52'44"E to 73˚55'54"E. The airbase is mostly composed of different concrete materials and two to three kinds of vegetation. A location map of the study area is given in Figure 1.

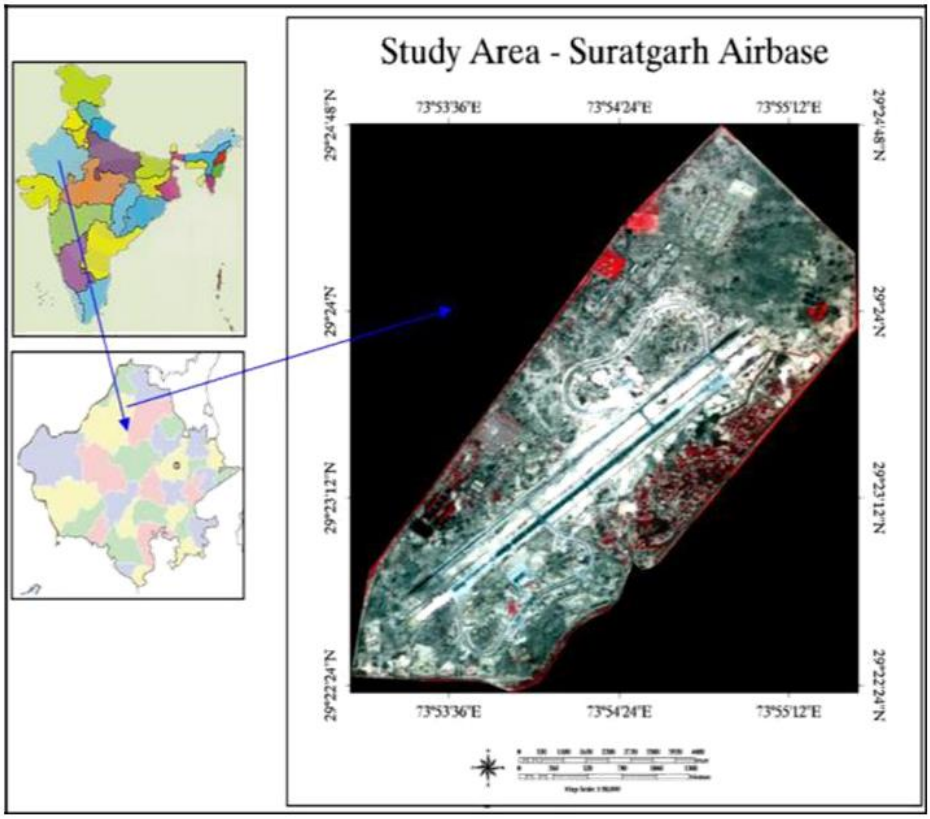


Figure 1: *Location map of the study area*

Datasets Used

Datasets used in the present study include a HS image from ESA's CHRIS (Compact High-Resolution Imaging Spectrometer)/ PROBA sensor and an MS image from Indian Remote sensing Satellite series - LISS IV (Linear Imaging Self Scanner) sensor. The Compact High Resolution Imaging Spectrometer on board Proba mission of European Space Agency was launched on 22nd October 2001 with an orbital period of 1 year. CHRIS/ Proba distinguishes itself from other sensors due to its multi-viewing capability and its programmable spectral channels. The sensor acquires images in five different fly by zenith angles of +55°, +36°, nadir, -36°, -55° with a time period of 2.5 seconds. The negative zenith angles (-36°, -55°) correspond to acquisitions for which the satellite has already flown over the target while the positives represent the geometry before nadir pass. CHRIS monitors the earth in five different modes each one having its own application like vegetation studies, water and oceanic studies, atmosphere, land use/land cover applications etc. In this study, nadir viewing mode 3 image that operates at a spatial resolution of 17m and spectral resolution of 18bandswas used. The specifications of the datasets are presented in the Table 1.

Table 1: *Technical specifications of the datasets used in the study*

Name	Spatial resolution	Spectral resolution	Date of acquisition
CHRIS	17m	18 bands (400-1050nm)	29-02-2008
LISS IV	5.8m	3 bands(500-900nm)	01-03-2008

2. Methodology

Pre-processing of HS data is an important step in hyperspectral data processing. In the pre-processing step, CHRIS data was radiometrically corrected for any bad bands and vertical/ horizontal line drop outs. The weighted averaging technique was used for correcting line drop outs. Later, radiometrically corrected CHRIS image and LISS IV image were atmospherically corrected using FLAASH atmospheric correction module (MODTRAN based model). Various parameters like satellite elevation, sun azimuth angle, aerosol optical thickness, average elevation of the study area and satellite, date and time of acquisitions, lat/long of the image etc. were given as inputs for FLAASH module for obtaining accurate results. The corrected datasets were then fused using HPF and GS image fusion techniques after which resultant fused outputs were classified. A detailed step by step methodology adopted in the study is presented below.

Co-registration

Co-registration is theprocess of geographically aligning one image with respect to other. The accuracy of co-registration plays an important role in quality of fusion. The atmospherically corrected LISS IV image was initially geo-corrected using 22 ground control points (collected using Trimble GPS during the field visits) with an RMSE of 0.23. This image was used for co-registering CHRIS nadir image, with LISS IV as master and CHRIS as slave. An overall RMSE of 0.124 was obtained after co-registration.

Fusion

The co-registered images were fused using pixel based Gram Schmidt transform and High Pass Filter based fusion techniques. In the Gram Schmidt technique, the green band of the LISS IV image was fused with the corresponding bands falling in green region of CHRIS dataset. Similarly, red and IR bands of LISS IV image were fused with the corresponding red and IR bands of CHRIS dataset. Then all the fused layers were stacked to get fused HS output. A similar procedure was followed for HPF method.

Spectral and quantitative analysis

The results obtained after fusion using GS and HPF technique were tested for spectral and spatial property preservance. The spectral profiles of various materials in the study site were compared visually. Later, a quantitative analysis was made to find the quality of the fused image in preserving the signal to noise characteristics of the input images. Mean, standard deviation, SNR and entropy measures were used to observe for any increase/decrease in final outputs.

SNR (Signal to Noise Ratio) and Entropy

SNR is defined as the ratio of signal in image to amount of noise in the image. In this work, the SNR for each band was calculated individually and finally an average SNR of all these bands is considered for analysis. Entropy is an index to evaluate the information quantity contained in an image. An increase in the value of entropy after fusion indicates an improvement in the information content after fusion. Entropy is defined as:-

$$f_9 = -\sum_i \sum_j p(i,j) \log(p(i,j)) \text{---------- (4)}$$

Where, P(i,j) is the probability distribution of grey levels or reflectance values; log is the base 2 algorithm.

Classification

The statistical learning theory based Support Vector Machine classification algorithm was used. The reason for choosing this technique is its potentiality in classifying voluminous HS data using very few training samples (Sahithi et al., 2016). SVMs are known for their low computational time and high accuracy. Classification was performed using an RBF kernel with a gamma value of 0.75 and a penalty parameter of 500. These values were attained using trial and error approach. End members were collected for five different concrete materials and three vegetation species from both the original CHRIS and the fused images individually and used for classification.

3. Results and Discussion

Visual Analysis

Visual interpretation of fused images depicted that spatial resolution of the both GS and HPF fused images had reasonably increased as compared to CHRIS 17m data. The contrast and color of CHRIS image were well preserved in GS fused method and hence GS fused image was visually appealing as compared to HPF fused image. HPF fused image was observed to have sharp edges and staircase appearance at linear features like roads, airbase etc. The high pass filtering process in HPF method induced some unwanted salt and pepper noise in the fused image which degraded the visual quality of the image. These confusions were not seen in the GS fused image as the image was clear with good color and spatial quality preservance.

Quantitative analysis

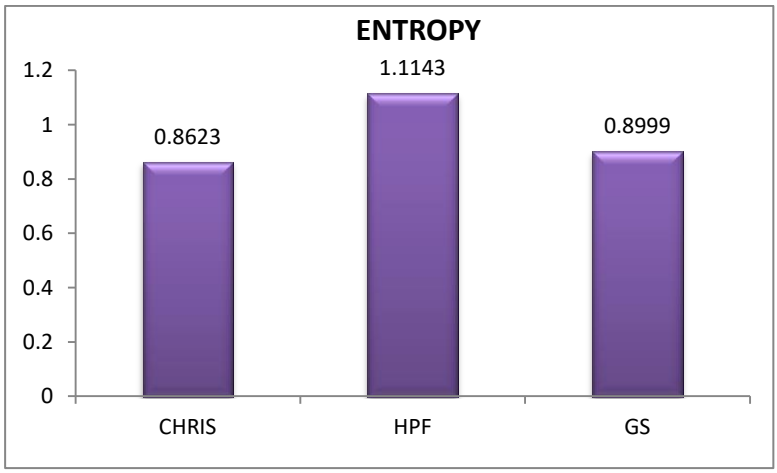
SNR and entropy measures were used for quantitatively analyzing the fused outputs. For calculating the SNR, few samples were collected from the brightest/foreground pixels and few other samples from darkest/background pixels from the image. A ratio of mean to the standard deviation of the foreground and background pixels gave SNR of the image. The SNR and entropy statistics calculated for HPF fused and GS fused images are given in Table 2 & 3. The SNR and entropy values of both the fused images have shown an increase, indicating an improvement in information content of the fused image. The Entropy of HPF fused image has drastically increased by 0.252 as compared to CHRIS image indicating an increase in randomness (due to the salt and pepper noise) in the image, while that of GS fused image has increased by 0.0376. The SNR values also indicated a good increase in HPF method than GS fusion method.

Table 2:*Statistics of HPF and GS fused images*

Band	GS fused image			HPF fused image		
	Mean	Std.	SNR	Mean	Std. dev	SNR
2	3871.29	397.63	9.74	4425.80	282.86	15.65
3	4008.70	385.72	10.39	4741.68	317.32	14.94
4	4240.16	397.35	10.67	5168.41	365.28	14.15
5	4113.15	377.87	10.89	5083.90	367.95	13.82
6	4899.63	453.86	10.80	5397.39	418.34	12.90
7	5010.03	450.44	11.12	5561.09	436.84	12.73
8	5036.44	461.95	10.90	5607.41	448.00	12.52
9	4952.83	410.25	12.07	5602.29	448.59	12.49
10	4904.05	411.77	11.91	5588.20	456.43	12.24
11	4850.29	411.26	11.79	5529.75	453.52	12.19

12	5025.78	400.33	12.55	5813.84	494.13	11.77
13	4843.87	382.00	12.68	5638.64	484.78	11.63
14	5081.64	399.47	12.72	5925.78	517.39	11.45
15	5271.05	414.67	12.71	6202.61	558.14	11.11
Average	4722.06	411.04	11.50	5449.06	432.11	12.83

Table 3: *Entropy of band 4 (551.00 nm) calculated for fused images and CHRIS image*



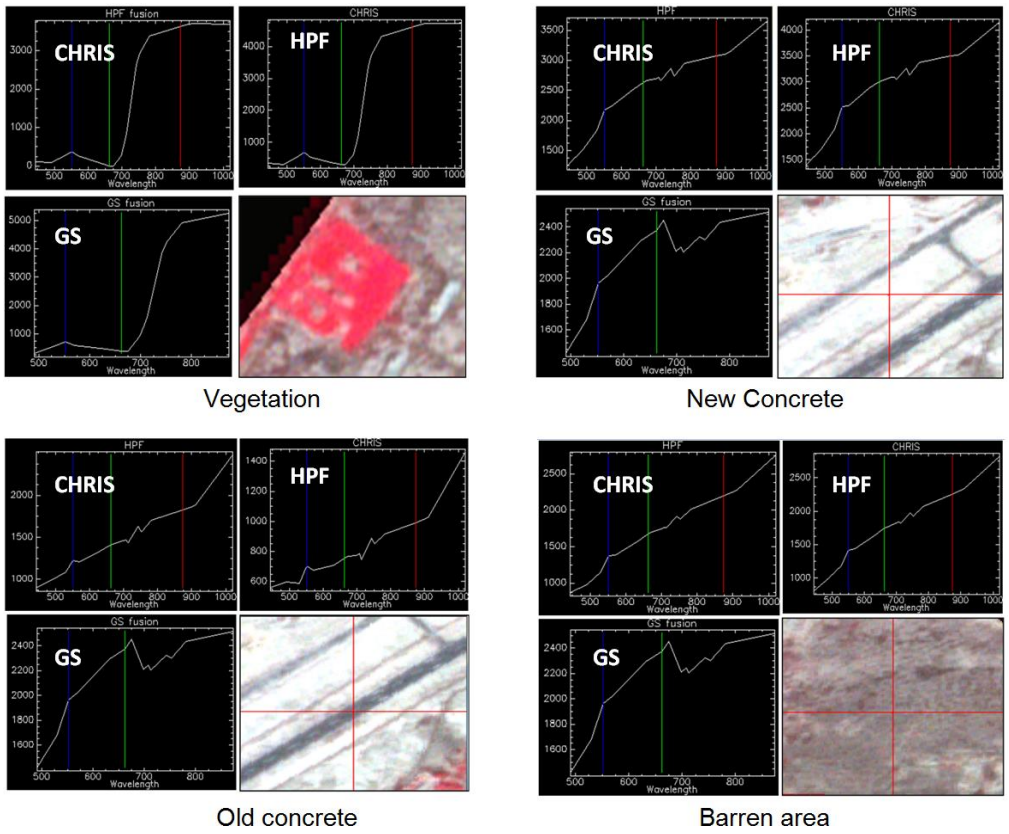


Figure 2:*Spectral profiles analysis of various materials –CHRIS, HPFand GS fusion respectively*

Spectral analysis

Gram Scmidt fusion of CHRIS and LISS IV images resulted in a spatially and spectrally rich fused product with 15 bands in 400 - 1050nm region with a spatial resolution of 5.8m. HPF fused image contained 14 bands with 5.8m spatial resolution. A comparison of spectral profiles of vegetation and urban materials in GS fused, HPF fused and original CHRIS images is presented in Figure 2. Analysis of spectral profiles has depicted that both GS and HPF fusion methods preserrved the characteristics of vegetation features but the spectral properties of urban materials were well preserved in HPF method as compared to GS technique. A sudden dip is observed at 670nm region for urban materials in GS fused image.

Classification Results

The results of SVM classification of GS and HPF fused images are presented in Figure 3. Overall classification accuracy for both the classified maps was assessed using ground GCPs and ground truth information.

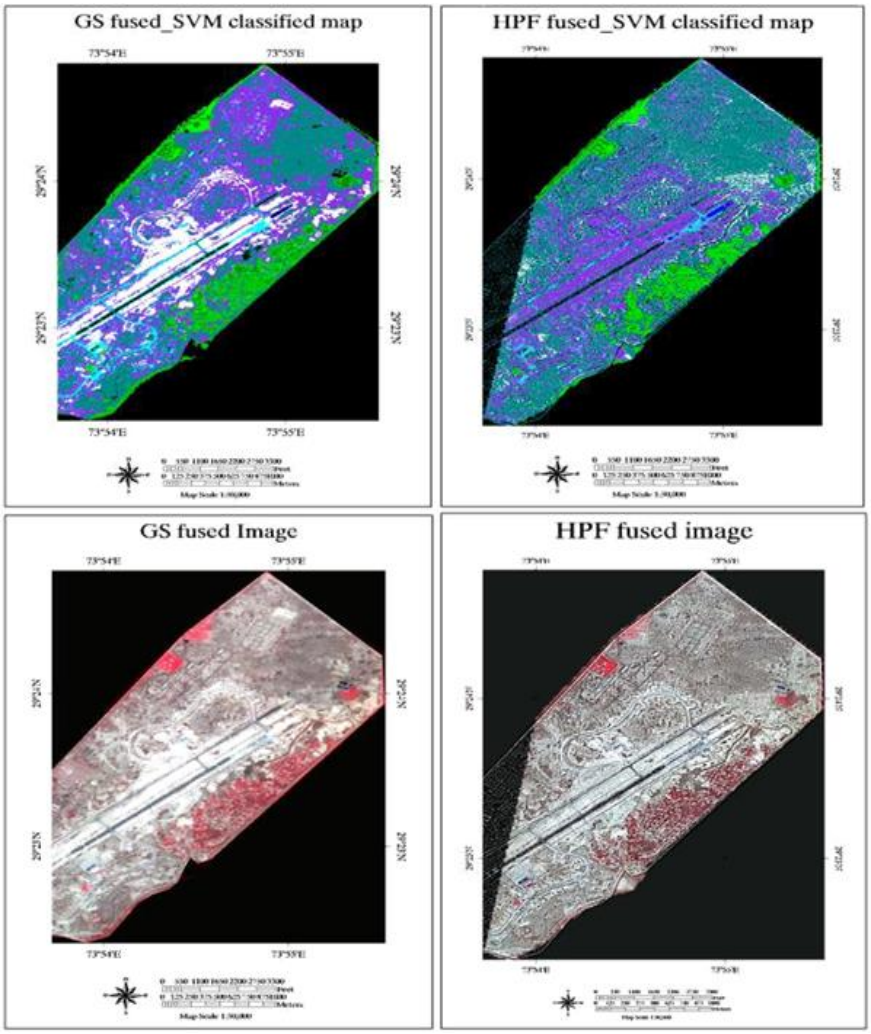


Figure 3: *a) Classified GS fused map b) Classified HPF fused map c) GS Fused Image d) HPF fused image*

It was observed that accuracy of GS fused image was 88.33% while that of HPF fused image is 79.25%. Many classes that were misclassified in CHRIS image due to mixing of pixels were revealed in fused image. Misclassifications observed in the classes of new concrete and old concrete were cleared in GS fused image. The noisy HPF fused image had certain unclassified pixels and barren area class was noticed to be negligible in HPF classified image. The classified map obtained from GS fused image yielded highest information as compared to classifications obtained from original CHRIS image and HPF fused image. Clear distinction between old concrete, new concrete, barren area and asphalt were observed in GS fused image.

Comparison was also made between classified maps obtained from CHRIS, LISS IV and GS fused image. Same number of classes were used for classification in all the three images and certain discriminations which could not be identified in the input images but well evident in the fused images are presented in Figure 5 (left). It can be seen that pathway in airbase station, crop lands etc are well identified after fusion.

In Figure 5 (right), the additional spatial and spectral information from the GS fused image helped in differentiating spectrally confusingclasses like cropland and bushes, which were confusing in the input datasets. Barren area/pathway is clearly delineated in GS fused image due to increased spatial and spectral information. The same path got mixed with old concrete in LISS IV image (due to lack of spectral information) and scrub in CHRIS image (mixing of spectra due to moderate spatial resolution). Also, a water body (circled) seen in CHRIS classified image but not identified in LISS IV classified map is delineated in the fused image. Thus, image fusion can help in enhancing certain inherent information and also preserving the characteristics of input data.

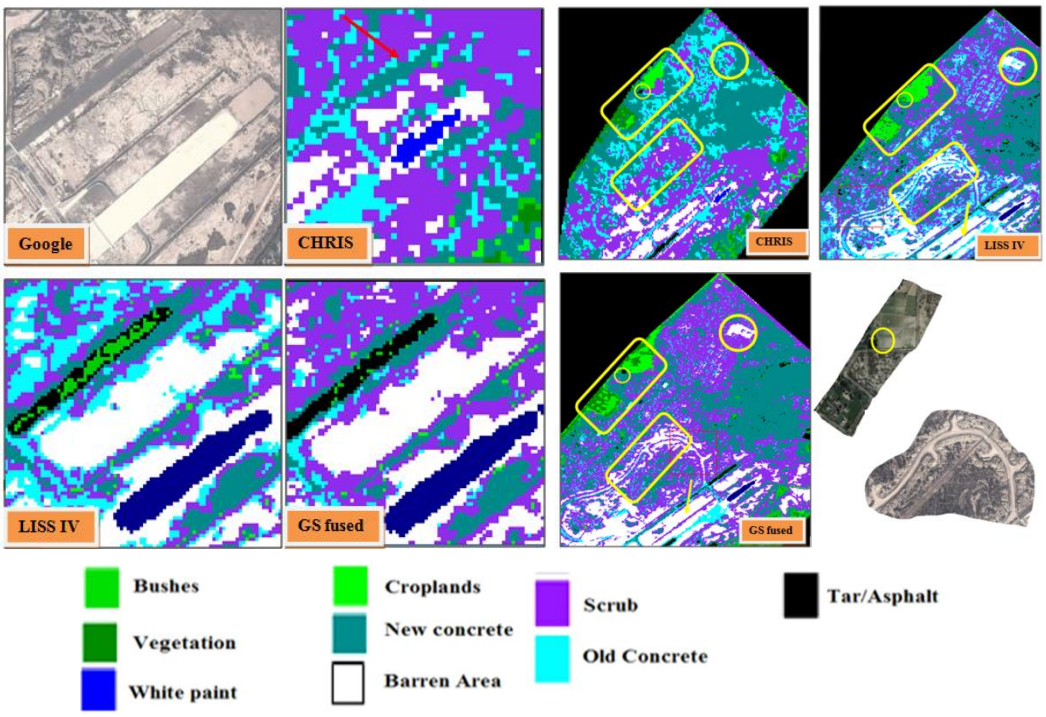


Figure 4. *Detailed observation of regions- comparison of classified maps from CHRIS, LISS IV & GS fused images*

4. Conclusion

In this work, we have made an attempt to fuse hyperspectral image from CHRIS sensor with multispectral image from IRS-LISS IV sensor, to obtain a spatially and spectrally rich hyperspectral image. Two pixel-based image fusion techniques were used – Gram Schmidt and High Pass Filtering. The GS fused image was visually appealing with good color preservance and also gave good classification results. But on the spectral and statistical end, HPF methods had a relatively better performance than the GS fusion methods. The spectral properties of various urban materials were well preserved in HPF method but the high pass filtering led to some unwanted salt and pepper noise in the fused output. The increased SNR and entropy values after fusion indicated that, fusion method not only preserves the information from the input images but also enhances some inherent information that is unavailable in the raw input images alone. Classification results concluded that fusion of HS image with MS image is a reliable image processing method for improving the accuracies of the classified maps and target detections. However, the spectral and temporal characteristics of the input images and accurate co-registration must be taken care to obtain desired results.

Acknowledgement

We would like to thank the Director of Research Centre Imarat (RCI - DRDO), Hyderabad for all the support that he extended during the project work. Also, we would like to thank the Department of Science and Technology (DST) for the financial support extended.

References

Abdikan, S., Fusun, B.S. and Manfred, E. 2014. A comparative data-fusion analysis of multi-sensor satellite images. *International Journal of Digital Earth*, 8, pp.671-687.

Ali, D., Kappas, M. and Erasmi, S. 2005. *Hyper-Spectral/High-Resolution Data fusion: Assessing the Quality of EO1-Hyperion/Spot-Pan & Quickbird-MS Fused Images in Spectral Domain*. Unpublished online document.

Candade, N. and Dixon, B. 2003. *MS classification of Landsat images: A comparison of support vector machine and neural network classifiers*. ASPRS Annual Meeting Proc., Denver. 43(6), pp.1882-1889.

Dahiya, S., Garg, P. and Jat, M.K. A comparative study of various pixel-based image fusion techniques as applied to an urban environment. *International Journal of Image and Data Fusion*, 4, pp.197-213.

Despini, F., Sergio T. and Baraldi, A. 2014. *Methods and metrics for the assessment of Pan-sharpening algorithms*. *International SPIE Proceedings*, *Image and Signal Processing for Remote Sensing XX,* 924403.

Foody, G.M. and Mathur, A. 2004. A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42, pp.1335-1343.

Gualtieri, J. Anthony and Robert, F. 1999. Cromp.Support vector machines for hyperspectral remote sensing classification. *27th AIPR Workshop: Advances in Computer-Assisted Recognition*. International Society for Optics and Photonics, Vol. 3584.

Hall, D.L and Llinas, J. 1997. An introduction to multisensor data fusion.*Proceedings of the IEEE*, 85(1), pp.6-23.

Hsieh and Pei-Jyun, 2015. An automatic kernel parameter selection method for kernel nonparametric weighted feature extraction with the RBF kernel for hyperspectral image classification. *Geoscience and Remote Sensing Symposium (IGARSS),* IEEE International.

Jawak, Shridhar D. and Luis, A.J. 2013. A comprehensive evaluation of PAN-sharpening algorithms coupled with resampling methods for image synthesis of very high resolution remotely sensed satellite data. *Advances in Remote Sensing*, 2, p.332.

Kohram, M.N. 2008. Support vector classification of remote sensing images using improved spectral kernels. *Journal of TeknologiMaklumat*, 20(1). pp. 14-27.

Lillesand, M.T. and Kiefer, W.R. 2000. *Remote sensing and image interpretation*. New York: John Wiley and Sons.

Luciano, A., Aiazzi, B., Baronti, S. and Selva, M. 2008. MS and panchromatic data fusion assessment without reference. *Photogrammetric Engineering & Remote Sensing*, 2, pp.193-200.

Luciano, A., Baronti, S., Garzelli, A. and Nencini, F. 2004. Landsat ETM+ and SAR image fusion based on generalized intensity modulation. *IEEE Transactions on Geoscience and Remote Sensing*, 12, pp.2832-2839.

Maurer, T. 2013. How to pan-sharpen images using the Gram-Schmidt pan-sharpen method-a recipe.International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W1, pp.239-244.

Melgani, Farid, and Lorenzo Bruzzone. 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on geoscience and remote sensing, 42(8)*, pp.1778-1790.

Pal, M., and Mather, P.M. 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26(5), pp.1007-1011.

Pal, Mahesh, and Paul M. Mather, 2004. Assessment of the effectiveness of support vector machines for hyperspectral data. *Future Generation Computer Systems*, 20(7), pp. 1215-1225.

Pande, H. and Tiwari, P.S. 2013. *High-resolution and HS data fusion for classification*. New advances in image fusion. InTech. pp.57-77. Available from: *http://www.intechopen.com/books/new-advances-in-image-fusion/high-resolution-and-HS-data-fusion-for-classification.* Last accessed 12th April, 2018.

Richard, B.G. and Jazaeri, A. 2001. Wavelet-based HS and MS image fusion. In Aerospace/Defense Sensing, Simulation and Controls. *International Society for Optics and Photonics*, pp.36-42.

Sahithi, V.S. and Iyyanki V Murali Krishna. 2016. Performance evaluation of dimensionality reduction techniques on CHRIS hyperspectral data for surface discrimination. *Journal of Geomatics*, 10(1), pp.7-11.

Sarp, G. 2014. Spectral and spatial quality analysis of pan-sharpening algorithms: A case study in Istanbul. *European Journal of Remote Sensing*, 47, pp.19-28.

Shi, D. and Yang, X. 2012. Support vector machine for landscape mapping from remote sensor imagery. Auto Carto.

Solanki C.K., Patel, N.M. .2011. Pixel based and Wavelet based Image fusion Methods with their Comparative Study. National Conference on Recent Trends in Engineering & Technology.

Vapnik, V. 1995. *The Nature of Statistical Learning Theory*. New York: Springer.

Vivone, G., Luciano A. and Jocelyn C. 2014. *A critical comparison of pan sharpening algorithms*. In Geoscience and Remote Sensing Symposium (IGARSS), IEEE International, pp.191-194.

Wald, L. 1998. *Data fusion: A conceptual approach for an efficient exploitation of remote sensing images*. International Conference. Fusion of Earth Data: merging point measurements, raster maps and remotely sensed images, pp.17-24.

Wald, L. 1999. Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 3, pp.1190-1193.

Waske and Björn, 2010. Sensitivity of support vector machines to random feature selection in classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 48, pp. 2880-2889.

Witharana, C., Civco, D.L. and Meyer, T.H. 2014. Evaluation of data fusion and image segmentation in earth observation based rapid mapping workflows. *Journal of Photogrammetry and Remote Sensing*, 87, pp.1-18.

Xu, S. and Ehlers, M. 2017. Hyperspectral image sharpening based on ehlers fusion. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*. Volume XLII-2/W7, pp.941-947.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2633-2642
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.358

Research Article

Water Resources of Barak Valley, India: Spatial Assessment of Lentic and Lotic System Using Remote Sensing and GIS at 5.8 m Resolution

Demsai Reang, Aparajita De, Ashesh Kumar Das

Department of Ecology and Environmental Science, Assam University, Silchar, Assam, India

Publication Date: 31 May 2018

Correspondence should be addressed to *aparajitade.ecology@gmail.com*

DOI: https://doi.org/10.23953/cloud.ijarsg.358

Abstract High resolution remote sensing data are suitable for quantifying or analyzing very small land cover components. In this study, the spatial distribution of water bodies and the perennial rivers of Barak valley were mapped and studied. Mapping was done for the first time using the high resolution IRS-R2 LISS4 FX data. The study adopted a hybrid technique incorporating supervised technique using the maximum likelihood algorithm and Visual Interpretation technique after ground validation. The river Barak along with its tributaries and sub-tributaries were identified and mapped. 16 tributaries and sub-tributaries have been mapped and reported for the first time after ground validation. This is a first detailed study on the spatial assessment of the waterbodies in this region. A total of 549 surface water bodies have been mapped and studied. The result indicated that the total area under water cover during dry season was *ca* 21,186 ha of the total geographical area of the study region. The study revealed that maximum numbers of the surface water bodies were between the sizes of 1 to 50 ha in area. Sone beel (wetland) was found to be the largest surface water body of the valley covering a total area of 1,348 ha. The final map composed represents the surface water of Barak valley and their spatial distribution. Thus, this paper may significantly help decision makers and researchers for sustainable water management, monitoring, controlling and planning development at local and regional scales.

Keywords *Anua; Beels; Floodplains; Haor; LISS4 data; Remote Sensing; River*

1. Introduction

Water body is the accumulation of water on the earth's surface and is an important resources and vital assets for human life and biological communities as a whole. The rivers and other wetlands in our country is habitat to various plants and animals; ranging from the simplest unicellular organisms to those belonging to the highest taxonomic order. The natural wetlands of India can be classified into seven categories. They are: himalayan lakes, flood plain wetlands, saline and temporary wetlands located in dry areas, coastal wetlands, mangrove swamps, coral reefs and marine wetlands (Prasad et al., 2002). It has been noted by other workers that the freshwater wetlands of India are home to ca 20% of the total biological diversity (Deepa and Ramachandra, 1999). The freshwater ecosystem services are particularly valuable to the inhabitants who are poor and entirely depend on nature's services directly for their livelihoods (Bhagawati and Gupta, 2015). Remote sensing technology helps

in better management of the water resources (Huang et al., 2016). The hydrology of the wetlands can be better understood if the surface water bodies can be surveyed and delineated accurately using remote sensing and GIS techniques (Du et al., 2016).

Monitoring water bodies accurately is an important and basic application in remote sensing (Du et al., 2016) and monitoring of surface water generally requires high spatial resolutions (Huang et al., 2016). The Landsat images have 30-m resolution, which is relatively low for use in detailed land study (Iqbal and Khan, 2014). A combination of remote sensors have been used in recognizing and observing surface water since the 1970s, for example, Landsat Multispectral Scanner (MSS) and Thematic Mapper (TM)/Enhanced Thematic Mapper Plus (ETM+) (Rango and Salomonson, 1974; Feyisa et al., 2014; Li et al., 2013), NOAA Advanced Very High Resolution Radiometer (AVHRR) (Barton and Bathols, 1989; Jain et al., 2006) and Moderate Resolution Imaging Spectroradiometer (MODIS) (Chen et al., 2013; Huang et al., 2014).

The Survey of India topographical maps are the earliest true maps of India showing various landuse/landcover classes in detail. This also includes the waterbodies. The first mapping of wetlands in India was carried out in 1992-93 using IRS LISS I/II data (Garg, 2015), the spatial resolution of which is 72/36 m respectively. However this did not include the rivers, streams etc. Further updating was done by the Space Application Centre using IRSP6 AWiFS data of 2004-2005. This had a spatial resolution of 56-70 m. The third attempt was the inventory of wetlands under the National Wetland Inventory and Assessment (NWIA) project using IRS LISS III data (NWA: Assam, 2010). This had a spatial resolution of 23.5 m. The present study is the first attempt to spatially assess the water distribution of both lentic and lotic system of Barak Valley, Assam by use of very high resolution satellite data (IRS-R2 LISS4 FX) of 2011 to 2013 and GIS techniques. The spatial resolution of the data used is 5.8 m. Hence it was possible to map the major river and their various tributaries and sub-tributaries precisely for the first time. The other water bodies including charas (small streams), haors (floodplain lakes), beels (wetlands) and anuas (oxbow lakes) were also delineated and studied.

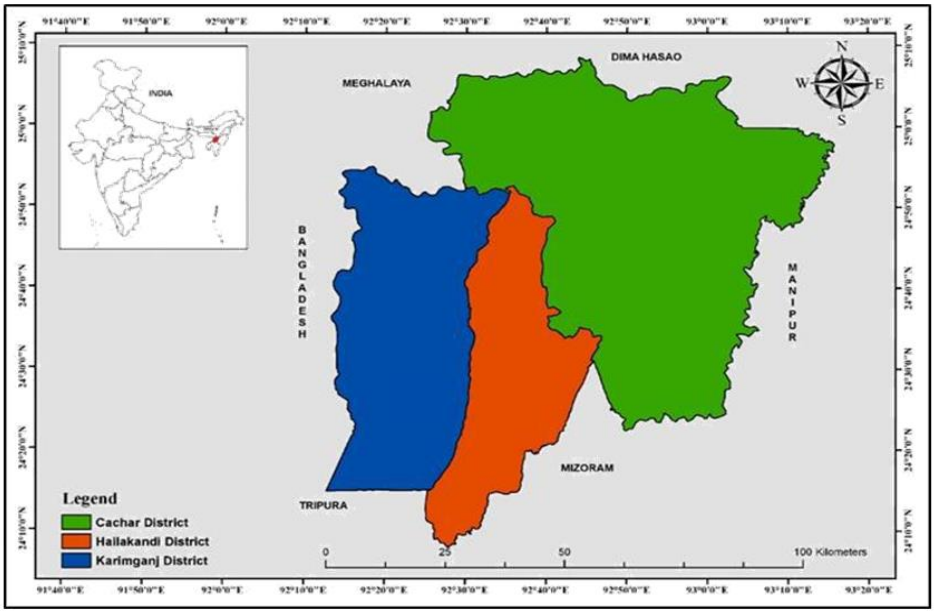


Figure 1: *Geographical location of the study area*

2. Materials and Methods

The study was conducted in Barak valley. Barak valley lies in the southernmost part of Assam, Northeast India and is named after the river Barak. It comprises of three different districts, viz., Cachar, Hailakandi and Karimganj (Figure 1). Geographically it is situated between $24^0$8'-$25^0$8' N latitude and $92^0$15'-$93^0$15' E longitude covering a total area of 6,922 km^2. The valley shares its boundary with Jaintia hills district of Meghalaya, Dima Hasao district of Assam on the North, Mizoram on the South, Manipur towards the east, the state of Tripura and Sylhet district of Bangladesh on the west. The largest district is Cachar covering an area of 3786 km^2, followed by Karimganj district (1809 km^2) and Hailakandi district (1327 km^2). Barak valley has an undulating topography which is characterized by hills, hillocks, wide plains and low-lying waterlogged areas. The valley has very gentle slope of average height of 20 meters above mean sea level. The hills situated on the northern part of the valley have a maximum elevation of ca 1500 meters (Ghosh et al., 1993). The valley is surrounded by forest on three sides. The southern hills are locally known as tillahs. In the satellite imagery they are seen to be projecting inwards intermittently into the valley (Figure 2). They have a maximum elevation of *ca* 200 msl. The valley is located at an altitude of 14-20 msl (Gupta, 2012) and is criss-crossed by a larger number of tributaries and sub-tributaries that forms the life support system for the local inhabitants. As a result the areas between these tributaries and sub-tributaries turn into large floodplain lakes during the monsoon months of June-September.

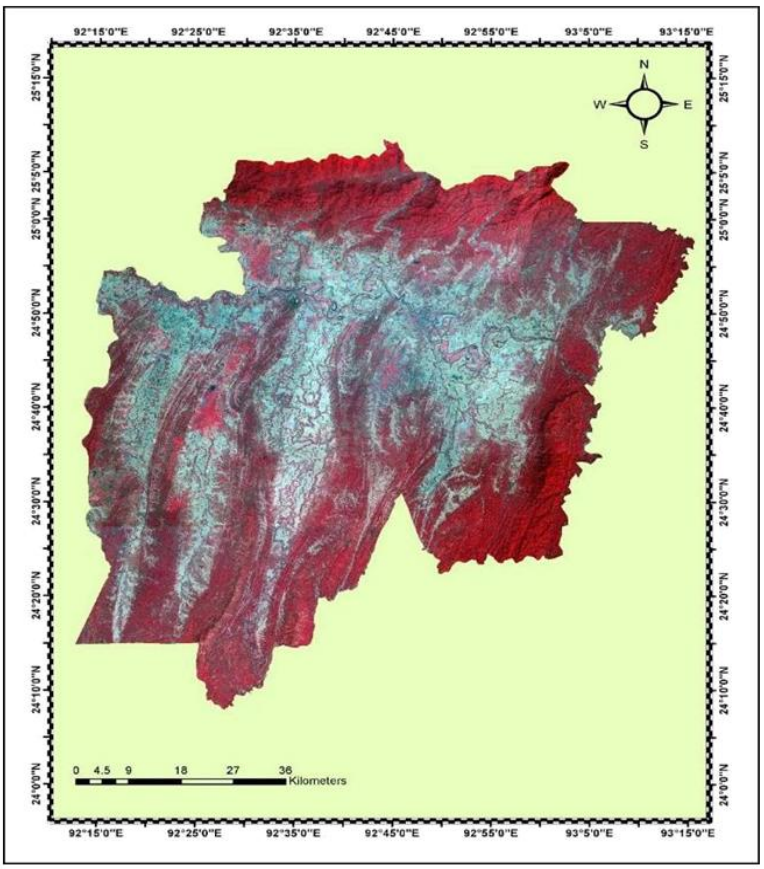


Figure 2: *Area of interest (False Colour Composite of IRS-R2 LISS4 FX Image)*

River Barak enters the Cachar district of Barak Valley near Lakhipur and flows in a westerly direction for *ca* 130 km through the plains of Cachar and Karimganj districts up to Haritikar near Bhanga, where it bifurcates into Surma and Kushiyara that enter Bangladesh. It therefore divides the entire study area into the northern and southern part. It has several tributaries and sub-tributaries branching in both in the northern and southern parts of the study area. The northern tributaries of river Barak are Jiri, Chiri, Madhura, and Jatinga-Doloo. The Dhaleshwari, Ghagra, Sonai, Katakhal and Rukni tributaries are found in the southern part of the valley, as reported previously (www.india-wris.nrsc.gov.in/wrpinfo, 2015). Rivers Baleshwar, Kalaincherra and Gumra are the major tributaries of River Surma. The Longai river in the Karimganj district is a tributary of River Kushiara. The River Singla is found in the southern side of Sone beel, while there are two small tributaries *viz*., Rivers Kakra and Kachua that flow on the northern side of the Sone beel and fall into the Kushiara river.

The region has numerous wetlands (locally known as beels), oxbow lakes (locally known as anuas), floodplain lakes (locally known as haors), and other temporary reservoirs. These floodplain wetlands are locally called haors, while the more depressed parts of the basin that retain water for a longer time are called beels (Gupta, 2012). Chatla, Hilara, Jabda, and Lucca haors of Cachar, Bauwwa and Bakri haors of Hailakandi and Son beel and Rata beel of Karimganj districts are among the most important floodplain wetlands of Barak Valley (Choudhury, 2000; Gupta, 2010). Kar et al. (Kar et al., 2008), reported a total area of 140,590 ha under lentic and lotic system in Assam.

This region experiences a warm humid climate having mean annual rainfall of 2,440–4,100 mm, that are mostly received during the southwest monsoon season (May to September) and a mean monthly maximum and minimum temperatures of 33.9°C and 9.2°C. The soil of the region falls under two textural classes; the plateau and the plain area, and is covered by silty clay and loamy soil while coarse sandy loamy soil is found in the hillocks.

For the study of surface water body, use of correct time period satellite data with a high resolution is important. The size, shape and nature of the water bodies, particularly of the smaller water bodies change with the change in seasons. Since the study area remains wholly or partially inundated by water during the monsoons and post monsoon periods, the dry season is the best time to identify and delineate them effectively. Therefore the satellite data of the dry season was selected for the present study. The selected satellite data sensor was IRS-R2 LISS4 FX. Cloud free data was procured from National Remote Sensing Centre, Hyderabad. IRS-R2 LISS4 FX satellite image has three spectral channels viz. green (0.52-0.59 µm), red (0.62-0.68 µm) and near IR (0.77-0.86 µm) with a temporal resolution of 5 days revisit for any given ground area and 70 km swath width. A total of 5 scenes were required to cover the entire study area and their details are given in Table 1.

Table 1: *Satellite data used for the study*

Sat ID	Sensor	Path	Agency	Row	Scene date	Resolution
IRS-R2	LISS4 FX	112	NRSC	55	25-Feb-11	5.8 m
IRS-R2	LISS4 FX	112	NRSC	54	25-Feb-15	5.8 m
IRS-R2	LISS4 FX	111	NRSC	54	10-Mar-11	5.8 m
IRS-R2	LISS4 FX	111	NRSC	55	10-Mar-13	5.8 m
IRS-R2	LISS4 FX	111	NRSC	54	10-Mar-15	5.8 m

Pre-processing of any satellite image is necessary for extracting and quantifying meaningful information from a remotely sensed data (Iqbal and Khan, 2014; Butt et al., 2015; Naqvi et al., 2014). Image pre-processing was performed using ERDAS Imagine application in order to extract complete information from the study area. IRS-R2 LISS4 FX data consist of 3 bands and so the geo-rectification

of all bands/layers and their quality enhancement was processed using ERDAS Imagine. The projection and datum selected were the Universal Transverse Mercator (UTM) and WGS84 respectively. The geo-rectified bands were then stacked. All the 5 scenes were pre-processed following the same above technique and the band combination is then changed to False Colour Composite image (FCC). The entire stacked satellite data was then mosaicked and a subset was then run to extract the required Area of Interest (AOI)/study area (Figure 2). Post-processing of the image was done using the ArcGIS software. Firstly, for the extraction of river, tributaries and sub tributaries, a supervised classification was performed using the maximum likelihood algorithm. The output clusters was then reclassified by changing the clusters old values and converted to polygon, and then edited using the ArcGIS tools.

Next, a thorough on-screen visual interpretation of the satellite data was carried out on the FCC image for delineating all the other waterbodies that were not classified by the unsupervised classification technique due to their relatively small size and spectral mixing with their neighbouring pixels. This step was done following the basic image interpretation elements like tone, texture, shape, size and location. For the purpose of performing this process accurately, a priori knowledge of the desired class occurrence was used (Foody, 1995). Delineation was done after a rigorous field visit of the different waterbodies and recording their geographic coordinates with the help of a handheld GPS. The secondary information regarding the names of the rivers and their specific locations was collected from the South Assam Forest Circle and other local inhabitants of the valley. After the successful digitization of all the waterbodies, the final delineated waterbodies layer was added to the earlier edited river and their tributaries layer for the final map composition. The final map obtained thus represented the distribution of different water bodies which is a very important component of the larger landscape in the three districts of Barak valley.

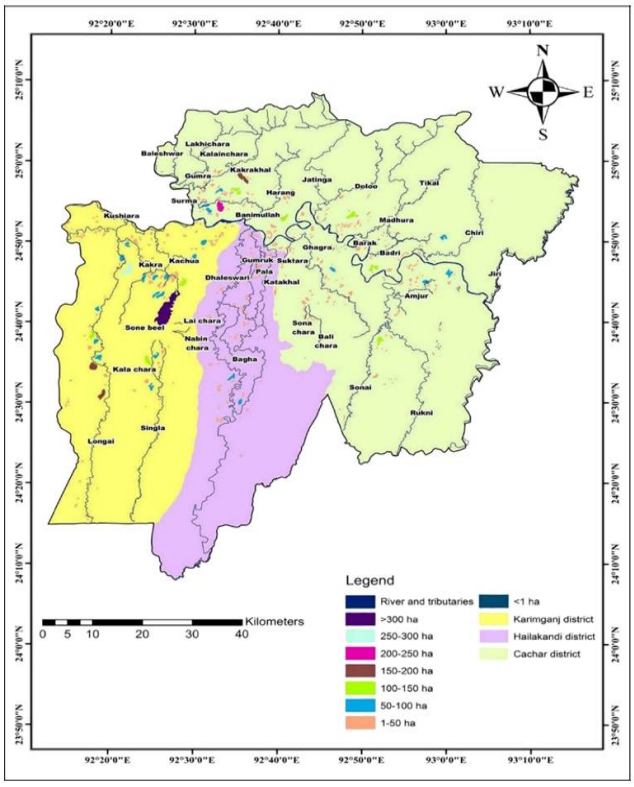


Figure 3: *Water resources of Barak valley, south Assam*

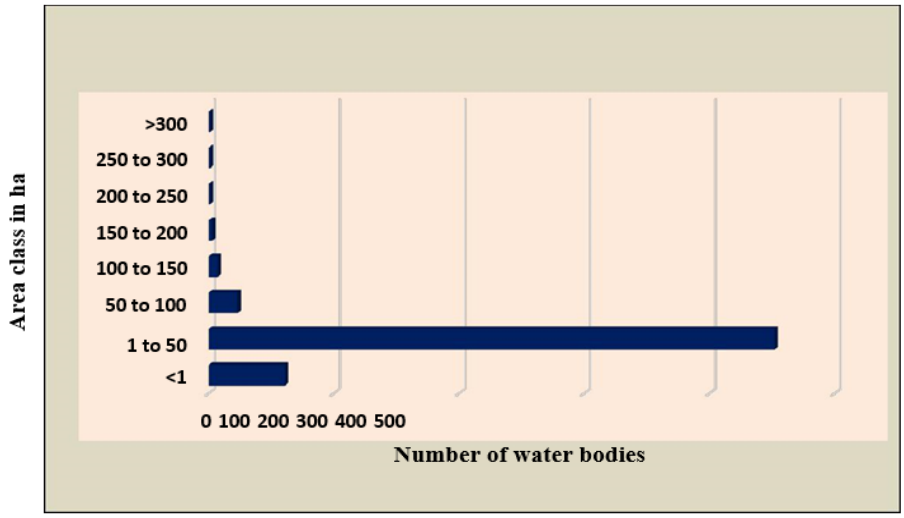


Figure 4: *Number of water bodies under different area class in hectare*

3. Results and Discussion

The final map (Figure 3) shows the spatial distribution of water bodies in Barak valley of Assam. The various rivers, their tributaries and sub-tributaries were identified and mapped in detail for the first time. A total of 549 lentic and lotic water bodies including the beels, haors and anuas were successfully delineated. The total area under the various waterbodies was found to be *ca* 21,186 ha in Barak valley. An increase of about 6,581 ha in the waterbody area was observed compared to an earlier study conducted in this region (Devi et al., 2012). This could be due to the higher resolution of imagery used in the present study as compared to the previous one.

The waterbodies delineated were classified into several categories based on their area (Figure 4). The size of the waterbodies ranged from less than a ha to 1,348 ha. Most of the waterbodies were present in the category of 1 ha to 50 ha. The Sone beel was found to be the largest waterbody comprising an area of 1,348 ha. The areas of most of the waterbodies fluctuate annually depending upon the rainfall. The present study was carried out using satellite imagery procured in the dry season (February-March) in order to map the areal extent of waterbodies accurately. Some of the important wetlands of the region and their areal extent is given in Table 2. The other important beels in the valley were; Salchapra beel, Doloo beel, Petumara beel, Hairuguri beel etc. Puneer haor, Chatla haor, Bakri haor were the major haors, Baskandi anua, Sibnarayanpur anua, Fulbari anua, etc. are some of the major anuas.

Table 2: *Some larger wetlands of Barak valley and their areal extent*

Name	Areal extent (in ha)
Sone beel (Karimganj district)	1348
Rata beel (Karimganj district0	117
Bakri haor (Hailakandi district)	13
Chatla haor (Cachar district)	15.3

The river Barak is the principal river flowing through the valley. There are a large number of tributaries and sub-tributaries forming a network of water channels in the valley. In the present study a total of 16 such tributaries and sub-tributaries have been studied and mapped for the first time. Many of the

tributaries remain dry during the dry season. However for the present study only those tributaries and sub-tributaries were mapped that had a perennial source of water. For instance, water was observed only in parts of River Kachua, which is a small tributary arising from Sone beel and falling into river Kushiara. Similarly several small rivulets were identified both in the satellite imagery as well as in the ground that had very little water. These were locally termed as Charas. Although small in size, they serve as important sources of water for the local inhabitants. Some of the charas that were digitized are Lai chara, Kala chara and Nabin chara in Karimganj district, Sona chara, Lakhi chara, Kalain chara and Bali chara in Cachar district. Some sub-tributaries of Dhaleshwari and Katakhal were also mapped. These include Bagha, Gumruk and Pala. The first is a sub tributary of Dhaleshwari while the second and third are sub tributaries of Katakhal. In the north of River Barak, the small rivers that were mapped and reported for the first time are Badri, Tikal, Banimullah, Lakhichara and Kakrakhal. The Tikal is a sub tributary of river Madhura while Badri is a small tributary of the river Barak (Figure 3). The river Harang was also found to bifurcate into two branches. The left branch drained into river Surma while the right branch, known as Banimullah was a sub-tributary of river Harang that drained into the river Barak. Two small tributaries of river Barak were observed in the southern side of river Barak, *viz.* Amjur in Cachar district and Suktara in Hailakandi district. Kakrakhal, Lakkhichara and Kalainchara are small tributaries of river Surma flowing in the north western part of the study area.

Figure 5: *a) River bank farming (Pala river); b) Stone Crushing in Bakri haor (Hailakandi district)*

Silchar is located in the floodplain area between river Barak and its tributary, river Ghagra, while a smaller portion is similarly enclosed by river Barak and river Madhura. The river Ghagra also forms the Chatla floodplain basin that functions as a natural floodwater storage and drainage area on the south bank of river Barak along with the Barak- Madhura floodplain on the north bank (Gupta, 2012). The present study helped in the identification of several other floodplain areas such as, The Kakra-Kachua-Kushiara floodplain in Karimganj district, The Dhaleswari-Katakhal-Barak floodplain in Hailakandi district and The Barak-Harang floodplain in northern part of Cachar district. In the absence of appropriate legal provisions for their protection and conservation, the wetlands of Barak valley have undergone severe degradation and shrinkage (Gupta, 2012). The floodplains of Barak Valley needs to be conserved as these help to store the excess floodwater.

During the field survey, it was observed that the rivers were being exploited at unsustainable levels through vegetable cultivation in the banks, collection of sand and gravel from the river bed/banks, fishing activities through erection of sluice gates and temporary barricades in the rivers, and proliferation of brick kilns and stone crushers in the floodplain areas. Even the small sub-tributaries like Pala, Gumruk etc. were being used in this manner (Figure 5). History and practice have shown that where wetlands are degrading, poverty generally increases, increasing pressure on remaining wetland resources and leading to further wetland degradation and poverty. The development of tourism has increasingly been regarded as a possible solution to the reduction of poverty in wetland areas (Van der Duim, 2007). A sustainable development of wetland based tourism in Barak valley could help in the alleviation of poverty in the region.

4. Conclusion

River Barak and its tributaries constitute the main drainage system in the Barak valley region of south Assam and is also the major source of useable water along with the floodplain lakes, ox-bows, ponds and other lentic freshwater systems (Gupta, 2003; Das and Gupta, 2008). Hence the detailed mapping of this very important land cover system will help in better management of this resource. The present study thus provides a detailed atlas of all the major waterbodies of Barak valley. As high resolution imagery was used for the present study it was possible to identify and map several smaller tributaries and sub-tributaries of the major rivers. Therefore the present study provided a holistic overview of the waterbodies of Barak valley and would help as it provided vital baseline data for related research activities.

Acknowledgment

The authors wish to acknowledge Department of Science and Technology and its FASTTRACK programme for funding the project (SERB/F/2489/2011-12 Dated 2/3/2012). The authors also acknowledge the Forest department (South Assam Circle). The authors are also thankful to the local people of Barak valley for helping in the ground identification of the streams during the field survey.

References

Barton, I.J. and Bathols, J.M. 1989. Monitoring floods with AVHRR. *Remote sensing of Environment*, 30(1), pp.89-94.

Bhagawati, R.R. and Gupta, S. 2015. Ecosystem health of Lake Tamrangabeel, bongaigaon district, Assam, India with special reference to aquatic insect assemblage. *Current World Environment*, 10(2), p.500.

Butt, A., Shabbir, R., Ahmad, S.S. and Aziz, N. 2015. Land use change mapping and analysis using Remote Sensing and GIS: A case study of Simly watershed, Islamabad, Pakistan. *The Egyptian Journal of Remote Sensing and Space Science*, 18(2), pp.251-259.

Chen, Y., Huang, C., Ticehurst, C., Merrin, L. and Thew, P. 2013. An evaluation of MODIS daily and 8-day composite products for floodplain and wetland inundation mapping. *Wetlands*, 33(5), pp.823-835.

Choudhury, A.U. 2000. *The Birds of Assam*. Gibbon Books and WWF India, N.E.R. Office, Guwahati.

Das, T. and Gupta, A. 2008. An appraisal on the physico-chemical properties of water along a disturbance gradient in River Barak, Assam, North East India. *The Bioscan*, 3(1), pp.109-114.

Deepa, R.S. and Ramachandra, T.V. 1999. *Impact of Urbanization in the Interconnectivity of Wetlands*. The National Symposium on Remote Sensing Applications for Natural Resources: Retrospective and Perspective (XIX-XXI 1999), Indian Society of Remote Sensing, Bangalore.

Devi, N.P., Reddy, C.S., De, A. and Dutta, B.K. 2012. Spatial assessment of land use in Barak valley, Assam using satellite remote sensing data. *National Academy Science Letters*, 35(5), pp.493-495.

Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W. and Li, X. 2016. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sensing*, 8(4), p.354.

Feyisa, G.L., Meilby, H., Fensholt, R. and Proud, S.R. 2014. Automated water extraction index: a new technique for surface water mapping using landsat imagery. *Remote Sensing of Environment*, 140, pp.23-35.

Foody, G.M. 1995. Using prior knowledge in artificial neural network classification with a minimal training set. *International Journal of Remote Sensing*, 16(2), pp.301-312.

Garg, J.K. 2015. Wetland assessment, monitoring and management in India using geospatial techniques. *Journal of Environmental Management*, 148, pp.112-123.

Ghosh, J.K. Lamar, H. and Roel, N. 1993. Forest cover and land use mapping of a region of barak valley of Assam, India Using IRS LISS-II Imagery. *International Archives of Photogrammetry and Remote Sensing*, 29, pp.435-435.

Gupta, A. 2003. *Flood and floodplain management in North East India: an ecological perspective. In Proceedings*. International Conference on Hydrology and Water Resources in Asia Pacific Region, 1, pp.231-236.

Gupta, A. 2010. Forests and wetlands of Barak Valley (Assam): a natural capital perspective. *Northeast Researches*, 1(1), pp.17-29.

Gupta, A. 2012. Water resource management in Barak valley, Assam, North East India: Status, needs and Priorities. *Bionano Frontier*, 5, pp.1-5.

Huang, C., Chen, Y. and Wu, J. 2014. Mapping spatio-temporal flood inundation dynamics at large river basin scale using time-series flow data and MODIS imagery. *International Journal of Applied Earth Observation and Geoinformation*, 26, pp.350-362.

Huang, C., Chen, Y., Zhang, S., Li, L., Shi, K. and Liu, R. 2016. Surface water mapping from Suomi NPP-VIIRS imagery at 30 m resolution via blending with Landsat data. *Remote Sensing*, 8(8), p.631.

Iqbal, M.F. and Khan, I.A. 2014. Spatiotemporal land use land cover change analysis and erosion risk mapping of Azad Jammu and Kashmir, Pakistan. *The Egyptian Journal of Remote Sensing and Space Science*,17(2), pp.209-229.

Jain, S.K., Saraf, A.K., Goswami, A. and Ahmad, T. 2006. Flood inundation mapping using NOAA AVHRR data. Water Resources Management, 20(6), pp.949-959.

Kar, D., Barbhuiya, A.H. and Saha, B. 2008. *Wetland Diversity of Assam: their present status*. In: Sengupta, M. and Dalwani, R. (Eds.) Taal 2007. Proceedings of the 12th World Lake Conference, pp.1844-1857.

Li, W., Du, Z., Ling, F., Zhou, D., Wang, H., Gui, Y., Sun, B. and Zhang, X. 2013. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sensing*, 5(11), pp.5530-5549.

Naqvi, H.R. Siddiqui, L., Devi, L.M. and Siddiqui, M.A. 2014. Landscape transformation analysis employing compound interest formula in the Nun Nadi Watershed, India. *The Egyptian Journal of Remote Sensing and Space Science*, 17(2), pp.149-157.

National Wetland Atlas: Assam, SAC/RESA/AFEG/NWIA/ATLAS/18/2010. 2010. Space Applications Centre (ISRO), Ahmedabad, India, p.174.

Prasad, S.N., Ramachandra, T.V., Ahalya, N., Sengupta, T., Kumar, A., Tiwari, A.K., Vijayan, V.S. and Vijayan, L. 2002. Conservation of wetlands of India-a review. *Tropical Ecology*, 43(1), pp.173-186.

Rango, A. and Salomonson, V.V. 1974. Regional flood mapping from space. *Water Resources Research*, 10(3), pp.473-484.

Van der Duim, V.R. and Henkens R.J. 2007. Wetlands, poverty reduction and sustainable tourism development: opportunities and constraints. Wetlands International; Wageningen, The Netherlands; Available from: library.wur.nl.

Water resources management system. Available from: www.india-wris.nrsc.gov.in/wrpinfo, Date of access 19/02/2018.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2656-2667
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.360

Research Article

Using Geographical Information System for Developing Education Information System: A Case Study of Srinagar City, Jammu & Kashmir, India

Faizan Jalal[1], Junaid Qadir[2]

[1]Department of Earth Sciences, University of Kashmir, Hazratbal, Srinagar – 190006, Jammu & Kashmir, India
[2]Department of Geography & Regional Development, University of Kashmir, Hazratbal, Srinagar – 190006, Jammu & Kashmir, India

Publication Date: 31 May 2018

Correspondence should be addressed to *qadirjunaid10@gmail.com*

DOI: https://doi.org/10.23953/cloud.ijarsg.360

Abstract The purpose of this study was to collect appropriate data of Education facilities around Srinagar city using GPS and to store the data into the Geodatabase built to manage them more efficiently. The Education data pertaining to these Education facilities was classified into different classes such as: Primary, Primary with upper primary, Secondary, Higher Secondary and Schools operating together. A total of 352 Government Schools were mapped out of them 122 were Primary Schools, 138 primary with upper primary Schools, 52 were Secondary Schools and 29 were Higher Secondary Schools. During the study it was found 10 Schools were operating together at same place, so their single coordinates have been taken and few of the Schools couldn't be mapped due to some issues such as non-existence. The present study focuses on using Geographic information system and integrating with the Education information system to generate a baseline data for the planning, development and preservation of these Education facilities.

Keywords *Education information system; Geographic information system; GPS; Geodatabase; Education Facilities*

1. Introduction

Education information system is a GIS based information system that stores all information about Educational facilities – their Name, location, category, No. of class rooms, No. of boys, No. of girls, Total no. of students, No. of teachers (male, female), Qualification of Teachers, Geographic coordinates of education institutes, Availability of: playground, computer lab, library, physical teacher, sports goods, NT (Non-teaching staff) strength, canteen facility etc. The trend towards using Geographic information systems and to support decision making for educational facilities are becoming very significant for planning purposes (Okan, 2012). Geographical information systems (GIS) can be described as a set of integrated software programs designed to store, analyze, and display geographical data-information (Fitzpatrick & Maguire, 2000). According to Makino and Watanabe, (2002) Geographic Information system has become an important asset and plays a vital role in analyzing the school-planning situation. Besides several other significant steps such as

explanation of educational administration, and expansion of compulsory education, enforcing the clear school planning policies and guidelines, and creation of the digital database that can be used in GIS, will have more positive and progressive impact on improving school services and imparting quality education to current and future children .

The current study has tried to map each and every Educational institute located in the Srinagar city in a comprehensive Geodatabase format, where in within a single click of the mouse we can know attributes such as Pupil data (enrolment, age, repetition), Teacher data (Qualification, experience, placement), School inventory data (location, geographic co-ordinates, stories, type, number of classrooms, equipment etc.), Provisioning of materials (textbooks etc.), Performance Data, Monitoring of internal management initiatives (e.g. Special projects) etc. The study holds significance for the Educational department which will find it handy for planning various projects relating to the development & improvement of our Educational facilities in Srinagar City in the future.

2. Study Area

Summer capital of the state of Jammu and Kashmir, Srinagar and is located in the valley of Kashmir at an altitude of 1,730 m above sea level. The city lies on both banks of the Jhelum River, a tributary of the Indus River. Srinagar lies between the coordinates 34^0 01` N to $34^0$ 27`N latitude and 74^0 36 E to 75^0 30 E, of over an area of 105 Sq. km (Figure 1).

Srinagar is surrounded by Budgam district in the west, district Pulwama in South and district Ganderbal in north. The city has the unique physiographic setup with steep hills in the east and north east, low lying paddy fields forming flood plain of Jhelum in the south and west and raised plateau lands in the south. The valley is surrounded by the Himalayas on all sides. Winters are cool, with daytime temperature averaging to 2.5 °C (36.5 °F) and drops below freezing point at night. Srinagar has Mediterranean type of climate with percentage humidity varying from 90% in winter to 78% in summer months. The average annual rainfall is around 720 millimetres (28 in). The highest temperature reliably recorded is 38.3 °C (100.9 °F) and the lowest is −20.0 °C (−4.0 °F). Geologically speaking, the city lies in the middle of great rock depression, with a volcanic hillock of Takat-i-sulaiman formed during the carboniferous and Permian times when glaciers were crawling in some parts of North kashmir (Chadda, 1991).

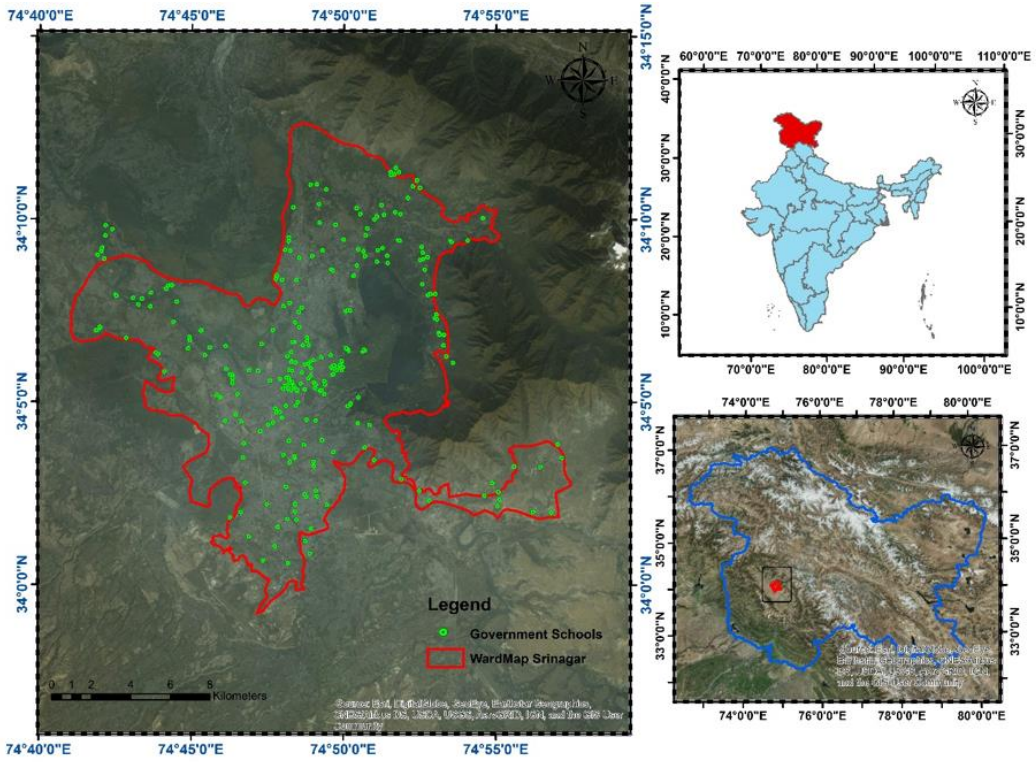


Figure 1: *Ward Map of Srinagar City, showing all Government schools.*

3. Methodology

In this Study, GPS was used to collect the location (geographic co-ordinates) of various Education facilities as per the data provided by Chief Education officer (C.E.O). These GPS points were then overlaid on high resolution GeoEye-1 image. The overall methodology adopted for the current study is shown in (Figure 2).

The Education data pertaining to these Education facilities was classified into different thematic maps or layers in Arc Map 10.1 GIS as given below:

1. Primary Schools.
2. Primary with Upper Primary Schools.
3. Secondary Schools.
4. Higher Secondary Schools.
5. Schools operating Together.

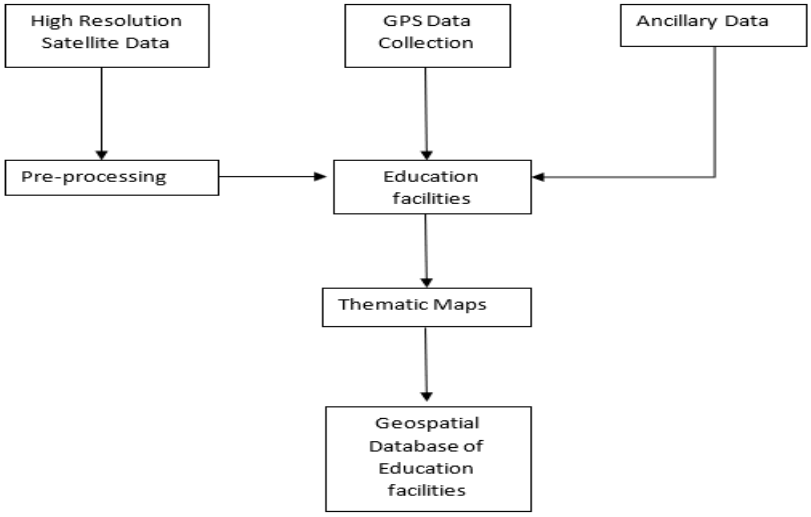


Figure 2: *Flow chart of Methodology.*

3.1 Data Sets

To carry out the study following data sets were used:

1. GeoEye-1.
2. Global Positioning System (GPS) data.
3. Ancillary data.

GeoEye-1 was launched on September 6, 2008 and is capable of acquiring image data at 0.46 meter panchromatic (B&W) and 1.84-meter multispectral resolution (https://www.satimagingcorp.com/satellite-sensors/geoeye-1/). The detailed sensor specification of GeoEye-1 is described in (Figure 3) below.

3.2 GPS Data Collection

In the present study an in-depth mapping of Srinagar city using handheld Juno SB GPS was carried out to generate the geospatial database of Educational facilities.

3.3 Ancillary Data

The ancillary data was generated from the data given at the C.E.O (Chief Education Office). The ancillary data in the form of hard copy format was then digitized to generate the database of Education facilities of Srinagar city in Excel format. It includes name and location of these institutes.

3.4 Pre-Processing

For the present study a high resolution GeoEye-1 of October 2008 was used as the satellite data source. Same data was first geo-referenced by field GCPs in order to make it reliable for the study. Also, the Education institutional data collected by GPS was processed in GPS pathfinder in order to make the data readable in Arc Map 10.1 and remove errors that had occurred during the field survey.

Scheduled Launch Date	2008		
Camera Modes	• Simultaneous panchromatic and multispectral (pan-sharpened) • Panchromatic only • Multispectral only		
Resolution	0.41 m / 1.34 ft* panchromatic (nominal at Nadir) 1.65 m / 5.41 ft* multispectral (nominal at Nadir)		
Metric Accuracy/Geolocation	CE stereo: 2 m / 6.6 ft LE stereo : 3 m / 9.84 ft CE mono: 2.5 m / 8.20 ft These are specified as 90% CE (circular error) for the horizontal and 90% LE (linear error) for the vertical with no ground control		
Swath Widths & Representative Area Sizes	• Nominal swath width - 15.2 km / 9.44 mi at Nadir • Single-point scene - 225 sq km (15x15 km) • Contiguous large area - 15,000 sq km (300x50 km) • Contiguous 1° cell size areas - 10,000 sq km (100x100 km) • Contiguous stereo area - 6,270 sq km (224x28 km) (Area assumes pan mode at highest line rate)		
Imaging Angle	Capable of imaging in any direction		
Revisit Frequency at 684 km Altitude (40° Latitude Target)	Max Pan GSD (m)	Off Nadir Look Angle (deg)	Average Revisit (days)
	0.42	10	8.3
	0.50	28	2.8
	0.59	35	2.1
Daily Monoscopic Area Collection Capacity	Up to 700,000 sq km/day (270,271 sq mi/day) of pan area (equivalent to about the size of Texas) Up to 350,000 sq km/day (135,135 sq mi/day) of pan-sharpened multispectral area (equivalent to about the size of New Mexico)		

Figure 3: *Sensor Specification of GeoEye-1.*

3.5 Data Integration

The Educational facilities database developed in Excel format was converted to .dbf format (Table 1) and integrated with the point locations of Educational institutes through a joining process in Arc Map 10.1.

Table 1: *Type of Data sets collected from various Government schools operating in Srinagar city*

Name of School	Category	Class Rooms	No. of Students	Location	Type	Latitude	Longitude
No. of Teachers	**Computer Lab**	**Library**	**Playground**				

4. Results and Discussions

During this Study, an extensive field Survey was carried out to map the Education facilities of Srinagar City using GPS. From a total of 410 Education institutes, 352 were mapped to make a Geodatabase containing the Spatial & Non-Spatial information of each institute. Total of 352 Sites were mapped to make a Geodatabase containing the Spatial & Non-Spatial information of each institute. The government Schools were divided in 5 different categories (Table 2(a), 2(b)): Primary schools, Primary with upper primary, Secondary, Higher Secondary and schools operating together.

Table 2(a): *(Attribute date); Total of 10 Sites taken from each Category, showing EIS for Primary, Primary with upper primary, Secondary, Higher Secondary and Schools operating together.*

School Name	School Category	Class rooms	Boys	Girls	Total Teachers.	Library	Computer Lab	playground	physical Teacher	Building/stories
PS CHEKI HARGOPAL	Primary	2	18	18	2	N.A	NO	N.A	N.A	1
BPS LALOO	Primary	5	60	60	4	N.A	NO	N.A	N.A	1
GPS PEERBAGH	Primary	2	4	6	6	N.A	NO	N.A	N.A	1
PS GORIPORA	Primary	3	9	8	5	N.A	NO	N.A	N.A	1
BPS BAGAT BARZULLA	Primary	2	6	3	3	N.A	NO	N.A	N.A	1
GPS CHECK NADIRGUND	Primary	1	7	7	2	N.A	NO	N.A	N.A	1
BPS RAWALPORA	Primary	2	31	30	12	N.A	NO	N.A	N.A	1
GPS WATTAL PORA RAWALPORA	Primary	2	18	24	3	N.A	NO	N.A	N.A	1
MPS CHECK RAWALPORA	Primary	3	30	30	3	N.A	NO	N.A	N.A	1
PS PARABAGH RAWALPORA	Primary	1	4	9	3	N.A	NO	N.A	N.A	1
MS ALLOUCHI BAGH	Primary with Upper Primary	5	35	30	6	N.A	NO	Available	Available	1
GMS AMIRA KADAL	Primary with Upper Primary	9	35	165	17	N.A	NO	Available	Available	1
BMS TULSI BAGH	Primary with Upper Primary	8	28	13	16	N.A	NO	Available	Available	1
GMS TULSI BAGH	Primary with Upper Primary	9	24	11	13	N.A	NO	Available	Available	1
BMS KUSHIPORA	Primary with Upper Primary	10	47	0	7	N.A	NO	Available	Available	1
GMS UMERABAD	Primary with Upper Primary	7	16	23	6	N.A	NO	Available	Available	1
GMS KHUSHIPORA	Primary with Upper Primary	4	0	83	5	Available	NO	Available	Available	1
GMS ZAINAKOTE	Primary with Upper Primary	7	102	119	12	Available	NO	Available	Available	1
BMS BEMINA	Primary with Upper Primary	5	40	63	12	Available	NO	Available	Available	1
GMS BOAT COLONY	Primary with Upper Primary	7	59	90	10	Available	NO	Available	Available	1
GHS BATMALOO	Secondary	15	28	65	24	Available	yes	Available	Available	2
GHS SHALTENG	Secondary	8	50	39	17	Available	yes	Available	Available	2
GHS CHATTABAL	Secondary	12	20	18	10	Available	yes	Available	Available	2
BHS CHATTABAL	Secondary	5	15	31	11	Available	yes	Available	Available	2
BHS LAWAYPORA	Secondary	7	19	36	12	Available	yes	Available	Available	2
BHS BARTHANA	Secondary	14	194	159	24	Available	yes	Available	Available	2
HS RAMBAL GARH	Secondary	7	100	139	13	Available	yes	Available	Available	2
GHS BARZULLA	Secondary	7	28	28	15	Available	yes	Available	Available	2
BHS HYDERPORA	Secondary	8	62	47	12	Available	yes	Available	Available	2
GHS CHATTERHAMA	Secondary	8	42	101	17	Available	yes	Available	Available	2
KGBV GASOO	Schools operating Together	8	0	22	6	Available	NO	Available	Available	2
KGBV GORIPORA	Schools operating Together	3	0	38	6	Available	NO	Available	Available	2
MPS OLD PAMPOSH COLONY	Schools operating Together	2	37	35	2	N.A	NO	N.A	N.A	1
GPS PANZINARA	Schools operating Together	3	6	19	4	N.A	NO	N.A	N.A	1
BPS PARIMPORA	Schools operating Together	3	10	12	2	N.A	NO	N.A	N.A	1
GPS PARIMPORA	Schools operating Together	2	10	10	2	N.A	NO	N.A	N.A	1
MPS TALI MOHALLA UMARHAIR	Schools operating Together	1	14	10	2	N.A	NO	N.A	N.A	1
BPS UMERHAIR	Schools operating Together	6	41	0	3	N.A	NO	N.A	N.A	1
PS KHOJA BAGH	Schools operating Together	1	5	9	2	N.A	NO	N.A	N.A	1
PS HANAN ABAD MALOORA	Schools operating Together	1	4	6	2	N.A	NO	N.A	N.A	1
BHSS Hazratbal	Higher Secondary	NA	305	0	11	Available	yes	Available	Available	3(4)
BHSS Zainakote	Higher Secondary	NA	82	80	16	Available	yes	Available	Available	3(4)
BHSS Batmaloo	Higher Secondary	NA	401	0	22	Available	yes	Available	Available	2(3)
GHSS Amira-kadal	Higher Secondary	NA	0	1915	16	Available	yes	Available	Available	2(3)
GHSS Khanyar	Higher Secondary	NA	0	569	19	Available	yes	Available	Available	3(3)
GHSS Kothibagh	Higher Secondary	NA	0	1561	28	Available	yes	Available	Available	2(3)
BHSS Jawahir Nagar	Higher Secondary	NA	1081	0	10	Available	yes	Available	Available	2(3)
S.P.Hr.Sec	Higher Secondary	NA	3470	0	20	Available	yes	Available	Available	4(2)
GHSS Soura	Higher Secondary	NA	0	626	13	Available	yes	Available	Available	2(4)
GHSS Nawakadal	Higher Secondary	NA	0	1543	15	Available	yes	Available	Available	2(3)

Table 2(b): *(Attribute date); Total of 10 Sites taken from each Category, showing EIS for Primary, Primary with upper primary, Secondary, Higher Secondary and Schools operating together*

School Name	School Category	Class rooms	Boys	Girls	Total Teachers	Library	Computer Lab	playground	physical Teacher	Building/stories
MPS BANPORA	Primary	3	5	3	4	N.A	NO	N.A	N.A	1
PS MAHRAJ PORA	Primary	1	6	11	2	N.A	NO	N.A	N.A	1
PS SHEIKH DAWOOD COLONY	Primary	1	10	9	2	N.A	NO	N.A	N.A	1
BPS PARIMPORA	Primary	3	10	12	2	N.A	NO	N.A	N.A	1
GPS PARIMPORA	Primary	2	10	10	2	N.A	NO	N.A	N.A	1
PS GHAT PARIM PORA	Primary	1	18	14	2	N.A	NO	N.A	N.A	1
PS DAR MOHALLA	Primary	1	16	14	2	N.A	NO	N.A	N.A	1
BPS SHALTENG	Primary	3	16	10	2	N.A	NO	N.A	N.A	1
PS ABAN SHAH	Primary	2	12	15	2	N.A	NO	N.A	N.A	1
PS KUMAR MOHALLA KHUSHIPORA	Primary	2	11	15	2	N.A	NO	N.A	N.A	1
BMS MALOORA	Primary with Upper Primary	8	111	48	10	Available	NO	N.A	Available	1
MS KOJABAGH MALOORA	Primary with Upper Primary	3	37	56	3	Available	NO	N.A	Available	1
GMS CHEK MUJGUND	Primary with Upper Primary	3	31	22	4	Available	NO	N.A	Available	1
GMS MUJGUND	Primary with Upper Primary	5	84	126	12	Available	NO	N.A	Available	1
BMS TENGPORA	Primary with Upper Primary	5	26	26	10	Available	NO	N.A	Available	1
GMS BARTHANA	Primary with Upper Primary	4	11	11	5	Available	NO	N.A	Available	1
MS FIRDOUSABAD FRUIT MANDI	Primary with Upper Primary	3	24	24	4	Available	NO	N.A	Available	1
BMS S.K COLONY	Primary with Upper Primary	5	85	93	11	Available	NO	N.A	Available	1
BMS PANZINARA	Primary with Upper Primary	7	25	23	13	Available	NO	N.A	Available	1
GMS BARZULLA	Primary with Upper Primary	4	13	19	14	Available	NO	N.A	Available	1
GHS RAJBAGH	Secondary	9	3	105	24	Available	Yes	Available	Available	2
BHS SONWAR	Secondary	10	63	48	22	Available	Yes	Available	Available	2
BHS PANTHACHOWK	Secondary	13	51	47	23	Available	Yes	Available	Available	2
BHS BADAMI BAGH	Secondary	8	6	3	15	Available	Yes	Available	Available	2
GHS PANDRETHAN	Secondary	7	20	21	18	Available	Yes	Available	Available	2
GHS LASJAN	Secondary	10	0	109	15	Available	Yes	Available	Available	2
BHS NATIPORA	Secondary	10	114	108	22	Available	Yes	Available	Available	2
GHS R.N.MANDIR	Secondary	9	26	32	21	Available	Yes	Available	Available	2
GHS ZAINDAR MOHALLA	Secondary	3	20	12	17	Available	Yes	Available	Available	2
BHS G.G. MOHALLA	Secondary	4	7	13	12	Available	Yes	Available	Available	2
GOVT.HSS GUND HASSI BHAT	Schools operating Together	7	17	23	13	Available	Yes	Available	Available	2
GHS HYDERPORA	Schools operating Together	10	10	3	15	Available	Yes	Available	Available	2
BHS BARZULLA	Schools operating Together	6	14	14	16	Available	Yes	Available	Available	2
BHS TAKENWARI	Schools operating Together	8	66	40	20	Available	Yes	Available	Available	2
BHS NANDPORA	Schools operating Together	3	43	53	15	Available	Yes	Available	Available	2
BHS NEW THEED	Schools operating Together	6	26	0	19	Available	Yes	Available	Available	2
BHSS NOWPORA	Schools operating Together	8	8	7	23	Available	Yes	Available	Available	2
BHS RAINAWARI	Schools operating Together	11	50	0	22	Available	Yes	Available	Available	2
GHSS RAINAWARI	Schools operating Together	6	0	53	34	Available	Yes	Available	Available	2
BHS LASJAN	Schools operating Together	9	14	0	16	Available	Yes	Available	Available	2
BHS GASSI MOHALLA	Higher Secondary	11	123	124	21	Available	yes	Available	Available	3
BHS CHANDIHAR	Higher Secondary	9	56	86	16	Available	yes	Available	Available	2
GHS NARWARA	Higher Secondary	9	0	39	20	Available	yes	Available	Available	2
BHS NARWARA	Higher Secondary	9	35	0	18	Available	yes	Available	Available	2
GHS GOORIPORA	Higher Secondary	8	39	27	13	Available	yes	Available	Available	2
BHS SANGUM	Higher Secondary	10	78	64	19	Available	yes	Available	Available	2
GHS HARWAN	Higher Secondary	7	0	184	19	Available	yes	Available	Available	2
BHS HARWAN	Higher Secondary	12	165	0	20	Available	yes	Available	Available	2
GHS DRUGJAN	Higher Secondary	9	38	150	20	Available	yes	Available	Available	2
BHS BRANE	Higher Secondary	11	38	0	18	Available	yes	Available	Available	2

4.1 Primary Schools

The term primary school is derived from French école primaire, which was first used in 1802.The primary schools are those type of elementary institutes in which children between the age group of 4-10 years of age receive first stage of compulsory education. In India schools having classes up to V (Starting from Play schools, Nursery, LKG, UKG) are categorized as primary school or elementary schools. In the current study a total of 112 Primary Schools (Figure 4) were mapped using GPS and the generated point theme was then stored in the Geodatabase.

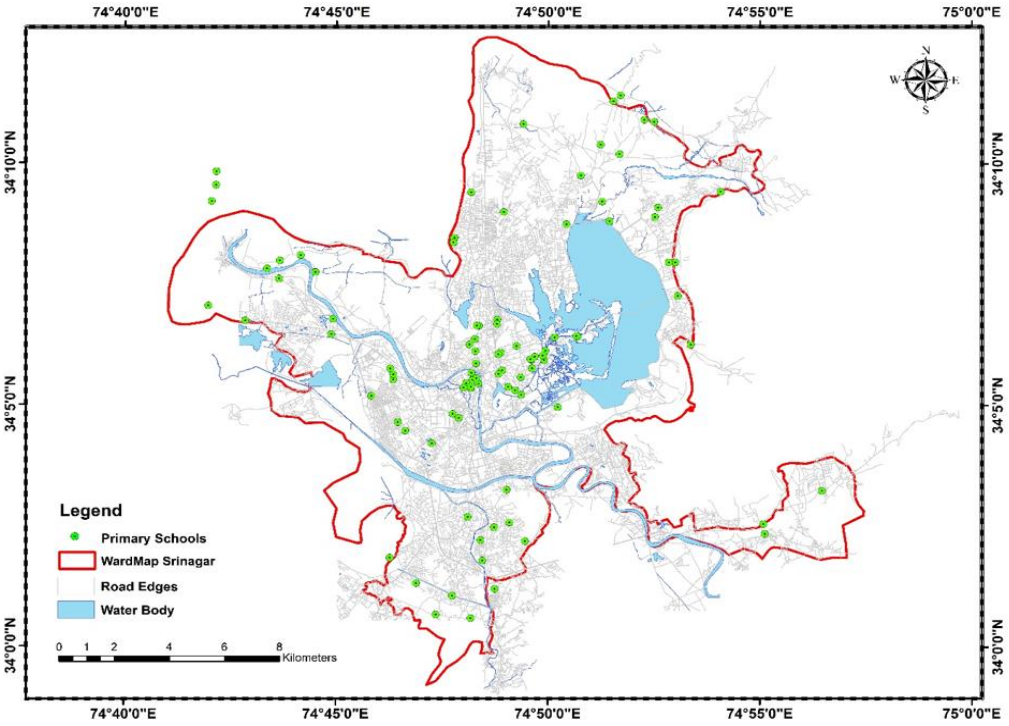


Figure 4: *Total number of Government Primary Schools in Srinagar city (J&K).*

4.2 Primary with Upper Primary School

Primary with upper primary or Middle schools are those type of elementary institutes in which children between the age group of 11 – 14 years of age receive second stage of compulsory education. In India schools having classes from VI – VIII are categorized as primary with upper primary school or middle schools. It is important to mention here Government of India made education free for children for 6 to 14 years of age or up to class VIII under the “Right of Children to Free and Compulsory Education Act 2009” (Ministry of Law and Justice Legislative Department). A total of 138 Government middle Schools (Figure 5) mapped using GPS and the generated point theme was then stored in the Geodatabase.

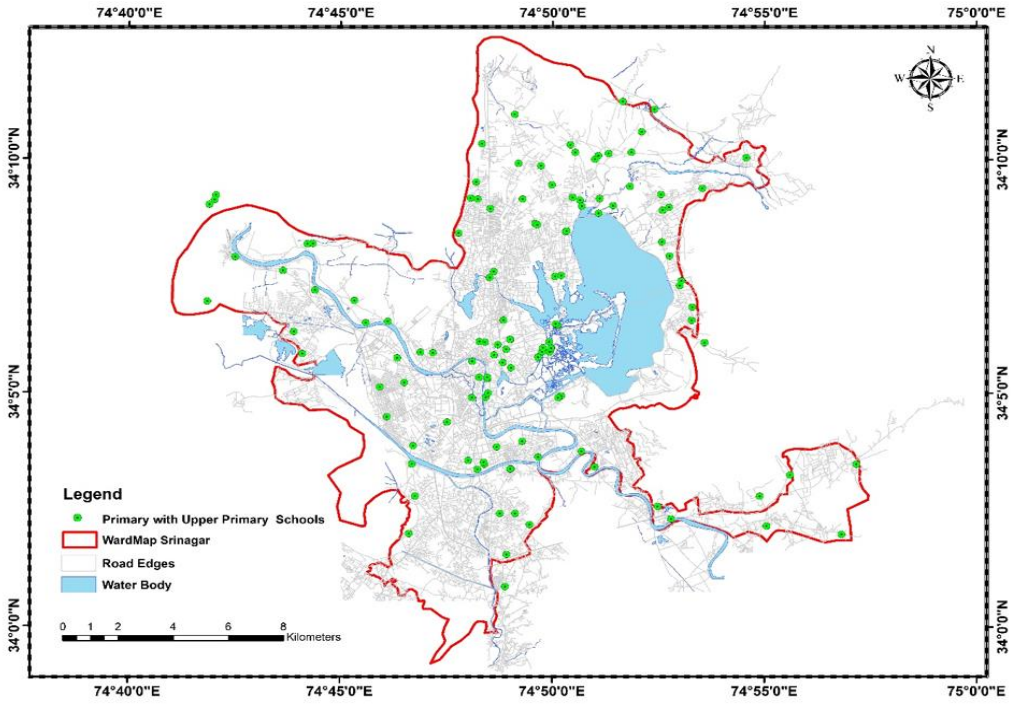


Figure 5: *Total number of Government Primary with upper primary Schools in Srinagar city (J&K).*

4.3 Secondary Schools

Depending on the system, schools for this period, or a part of it, may be called secondary Schools or High schools. In India children between the age group of 15 – 16 years of age are enrolled and cover two year of study of classes IX & X. The exact boundary between primary and secondary education also Varies from country to country and even within them. Here comes the main stage for a student's Career and that is Board exams of class X, till this period every student has same subjects to follow if they are from same Board of School Education (BOSE) or Central Board of Secondary Education (CBSE). A total of 52 Government Secondary Schools (Figure 6) mapped using GPS and the generated point theme was then stored in the Geodatabase.

4.4 Higher Secondary Schools

Higher secondary schools are those type of institutes in which children between the age group of 17 – 18 years of age are enrolled. These two years of study are the compulsory and important stage for which a pass certificate is needed to carry out further studies (College or University). In India schools having classes from XI – XII are categorized as Higher secondary schools. A total of 29 Government Higher Secondary Schools (Figure 7) mapped using GPS and the generated point theme was then stored in the Geodatabase.

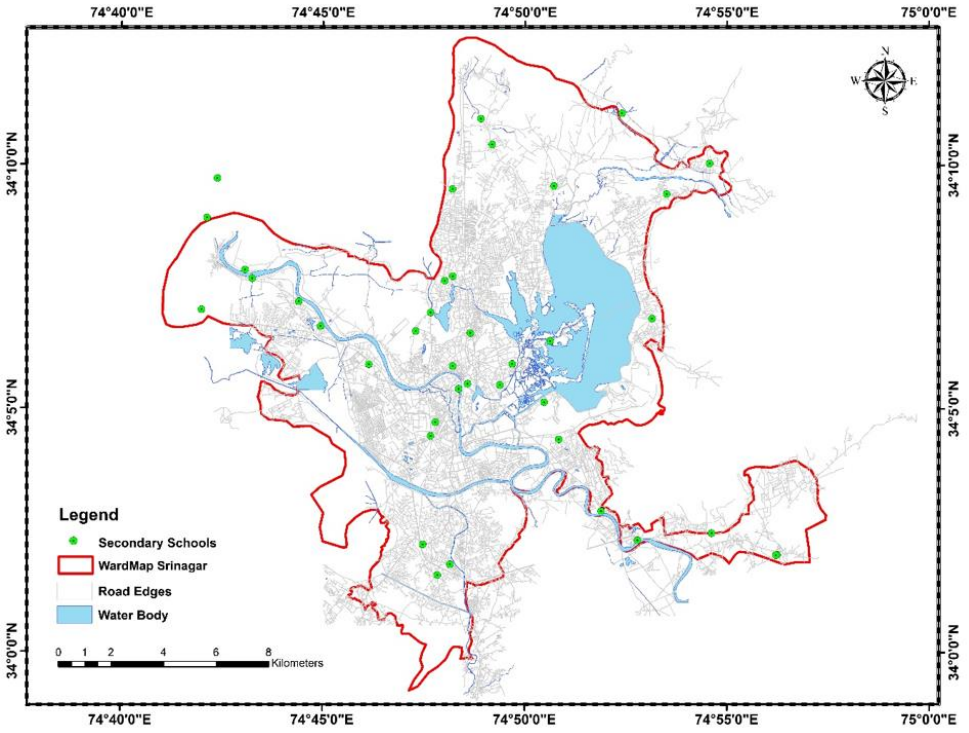


Figure 6: *Total number of Government Secondary Schools in Srinagar city (J&K)*

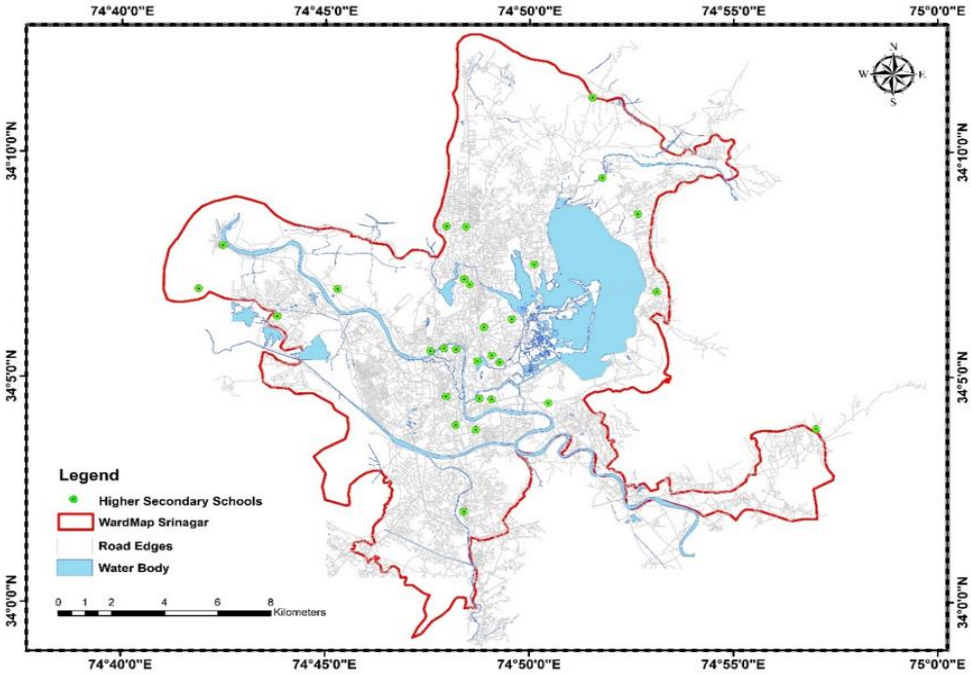


Figure 7: *Total number of Government Higher Secondary Schools in Srinagar city (J&K)*

4.5 Schools Operating Together

Schools of different blocks which are operating together at same place e.g. Ps Kushipora and Bms Kushipora HMT etc. So, these Schools constitute same geographic Location. A total of 10 such schools (Figure 8) were mapped constituting primary & primary or Primary & primary with upper primary schools.

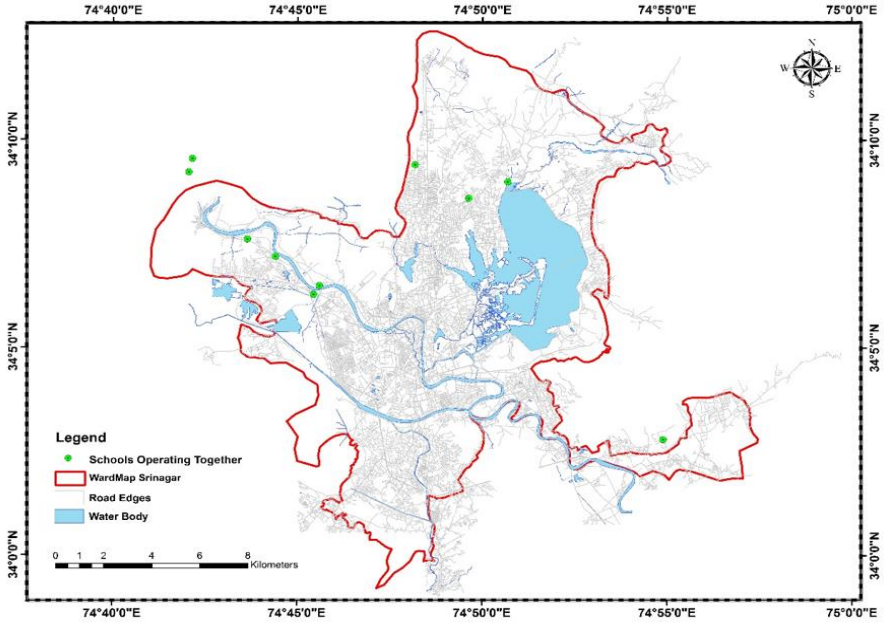


Figure 8: *Total number of Government Schools operating together in Srinagar city (J&K)*

Figure 9: *Snapshots of Different Education facilities in the Srinagar City, J&K taken during field survey. (A, B) Showing Government Primary Schools. (C, D) Government Primary with upper primary Schools. (E, F) Government (G, H) Government Higher secondary schools. (I, J) Government Schools Operating together*

5. Conclusion

During the study it was found that majority of schools lacked basic facilities like Toilets, Playgrounds, libraries, labs, etc. Most of the Schools are located in rented buildings and most of the primary schools operate in one or two rooms. These schools lack basic infrastructure. The classes are crowded and teachers are not able to give proper attention to children's. Most of the Primary with upper primary is also operating from rented accommodations lacking proper facilities. Most of the Secondary and Higher Secondary schools lack facilities like playground, Sports equipment, Lab equipment, Computer facility etc. In addition to this, the quality of education provided is not up to the standard. Many areas have unequal distribution of Schools, with some areas having more than two schools in their vicinity and others don't have even single school. The participation of Government Schools in various academic and extracurricular activities is negligible thus preventing Students from having a healthy exposure.

This study led to the vision for the generation of EIS for Srinagar City using GIS and would be used as a baseline data to improve planning, organisational efficiencies, data collection and analysis, information sharing & transparency in the education sector, especially as these activities may relate to helping meet objectives relate to Education for its development goals.

Authors Contribution

Each author contributed to the research in terms of conception, research design, cross-checking data analysis and co-writing the paper. Both authors read and approved the final manuscript.

Acknowledgement

Authors would like to thank Head of Department Earth Sciences, University of Kashmir for providing requisite facilities and support and C.E.O (Chief Education Office) for valuable suggestions and Esri, Digital Globe, USGS for the accessing High resolution data.

References

Chadda S.K. 1991. Kashmir Ecology and Environment. Mittal Publications, New Delhi, India.

Fitzpatrick, C. and Maguire, D.J. 2000. *GIS in schools: infrastructure, methodology and role*. In Green, D.R. (Ed.), GIS: A sourcebook for schools. New York: Taylor & Francis, pp.61-62.

GeoEye-1 Satellite Sensor (0.46m) Satellite Imaging orp. 2008. Available from: *http://www.satimagingcorp.com/satellite-sensors/geoeye-1/.*

Makino, Y. and Watanabe, S. 2002. The application of GIS to the school mapping in Bangkok. Retrieved from Asian association on Remote Sensing.

Ministry of Law and Justice (Legislative Department) (27 August 2009). The Right of Children to Free and Compulsory Education Act. Archived from the original on 19 September 2009.

Okan, E. 2012. Application of Geographic Information System (GIS) in Education. *Journal of Technical Science and Technologies*, 1(2), pp.53-58.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2668-2674
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.361

Research Article

Geographic Information System for Flood Management in Greater Colombo Area

T.D.C. Pushpakumara, T.P. Vidyadari

Department of Civil Engineering, University of Moratuwa, Sri Lanka

Publication Date: 9 June 2018

Correspondence should be addressed to *thudugalagedon@gmail.com*

DOI: https://doi.org/10.23953/cloud.ijarsg.361

Abstract Nonstructural approach in flood management plays a major role in disaster management practice. This research paper focuses on flood management using GIS software which comes as a non-structural approach. Land use is essential ingredient in determining the flood plain area of the study area. Colombo as the highest populated city in Sri Lanka, there is a higher pressure on land for human settlement and related urban services. Increasing impervious area and reduction in retention areas due to rapid development has resulted in faster runoff. Thus, flooding has been one of the costliest disaster in terms of both property and human damage in this area. This research paper focuses on analyzing the changes in the land use in the area over the decade with the use of Arc GIS software. Hec-Hms software has been used for determining flood plain of the area flood risk map has been developed using the obtained results.
Keywords *Catchment; DEM; Flood management; GIS; HEC-HMS*

1. Introduction

There are two approaches in flood management practice such as structural approaches and non-structural approaches. Structural approach can be defined as flood mitigation projects such as dams, detention ponds etc. Non-structural approach refers to flood forecasting, proper early warnings and conducting awareness programs among the flood affected community. Flood hazard mapping is considered as one of the priority tasks to be accomplished in disaster management practice since it is very effective in mitigating the impact of flooding compared to structural measures.

Flood management using GIS software which comes under non-structural approach has been initiated in this research. Application of GIS for mapping flood prone areas will make it easy to reduce the flood damages and risks involved. Such type of technologies can create accurate and current floodplain maps with improved efficiency, speed and public safety. GIS is ideally used for various floodplain management activities such as base mapping, topographic mapping, and post disaster verification of mapped floodplain extent and depths. GIS is useful in capturing and communicating a vast amount of information about the study area (Briviot et al., 2002).

The study area of the project, greater Colombo covers Colombo city which is the largest commercial city in the country and the area of four surrounding local authorities. Greater Colombo area has a total

extent of 104 km^2 and population of 1.2 million. This area has been vulnerable to flooding as most of the area consists of low lying lands which is less than 6m above the sea level and some areas are even less than 1m above the sea level. More paved areas and less open areas due to rapid development has resulted in faster runoff. Citizens of this area have suffered from frequent floods every year and their economic and social activities have often been interrupted. Therefore, there is a necessity of such kind of flood management system.

In this case, Satellite images were used to obtain the land use patterns in greater Colombo area in order to find the permeable and impermeable areas of the study area. Use of satellite images enables to gather invisible information rather than traditional survey approach. Sub catchments of the study area were obtained by analyzing the created DEM using hydrology tool (Sanyal et al., 2003). HEC-HMS software has been used to identify the flood depth.

1.1. Objectives

Under this research it is required to fulfill few objectives. Main objective of this research is to analyse a selected catchment and evaluate the risk of flood. In order to fulfill the main objective, there are few sub objectives such as to create catchments for the study area using digital elevation model, compare the land use patterns with the help of satellite images, and discuss the effect of land use patterns for flooding.

2. Methodology

Following procedure was followed in this research to reach objectives. As the first step, literature review was done to identify the methods which are available for preparing flood risk maps with the help of Arc GIS software. Required procedure for the research could be determined through the literature review.

5m DEM was created as by using elevation data which is received from the survey department. Then the created DEM was used to obtain sub catchment with the help of Hydrology tool which comes under spatial analyst tools in GIS software.

As an example, one sub catchment was selected to carry out analysing part in order to identify the flood risk. Analyzing the land use patterns for the selected sub catchment were done by digitizing the satellite images of study area(Samarasinghe et al., 2010). Effect of changing land use patterns and percentage of impervious area in flooding has been analyzed by comparing 2000 land use map and the 2010 land use map.

Finally, peak runoff and flood plain for the selected catchment were obtained by using hydrology software (Chen et al., 2009). In this case, HEC-HMS software was used to analyse the catchment using rainfall data. Variations of run-off flow due to changing the permeable area was modelled.

Flood plain was obtained through the HEC-HMS software and represented in the contour map in Arc GIS. Similarly, the other catchments can also be analysed.

2.1. Data

Data Collection Source

Flood affecting areas were analysed by using 5m DEM which should have to be developed since the available 90m DEM was not accurate enough for the purpose. For this case, Elevation data were obtained from the Survey Department and developed the 5m DEM for the Colombo area.

Satellite images which were used to digitized the 2010 land use map. For this case, high resolution satellite images were used in order to identify the land use features. Those high resolution maps were provided through the supervisor by purchasing from Map mart. Google earth was used to verify the digitized map. 2000 land use map was created by converting created AutoCAD file to GIS file. That AutoCAD file was provided by Urban Development Authority, Colombo.

Collected Data

Elevation data required for creating 5m DEM were obtained from the Survey Department. Then, high resolution satellite images should have to be used for the purpose since it should consist of pixels smaller than 1m. Therefore 1: 150,000 scale, eleven satellite images which were taken in year 2010 were used for the purpose. Identified features were verified by using Google earth.

For the comparison of land use variations in the study area, obtained 2000 land use map was used. Land use in year 2000 was obtained through the Auto-CAD map. Addition to that, maximum limit of the impervious area was taken as 65% according to the Colombo municipal reports.

A design of rainfall hyetograph which was required to simulate the rainfall in the hydrology software was taken from the design reports of Metro Colombo Urban Development Project. 100mm/hr Rainfall with a 10 year return period was used for analysing.

Analysis

Analysis of the data was done by using ArcGIS tools and Hec-Hms software. As shown in the Figure 1, 5m DEM was created by giving elevation to 5m squares. A section from the initial DEM was taken to produce sub catchment by considering the easiness of analysis. Right side of the Figure shows the section of the DEM which was used for analysing.

As the first step, hydrology tools which comes under the spatial analysis tools in GIS were used to produce the subcatchments in the created digital elevation model. One of the sub catchments was selected to carry out further analysing steps.

Table 1: *Elevation – Area data of the selected catchment*

Elevation (m)	Area (1000 m^2)
1.62	0.15
2	82.119
3	184.544
4	460.569
5	985.443
6	1875.293
7	2856.767

8	3384.373
9	3551.657
10	3573.369
11	3575.258
12	3575.573

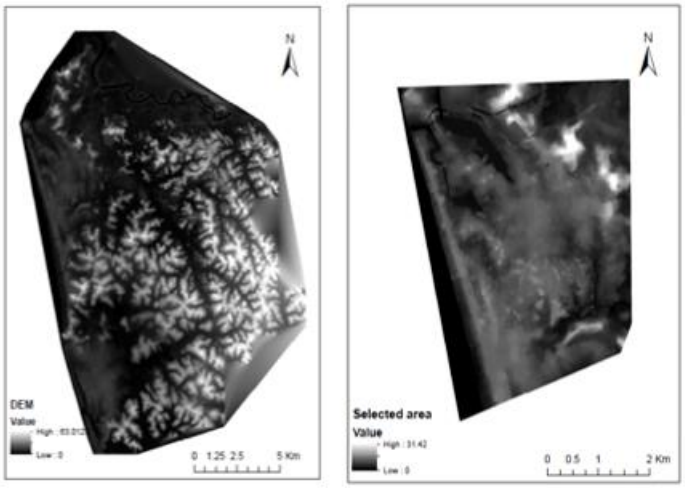


Figure 1: *Created 5m DEM and DEM used for analysing*

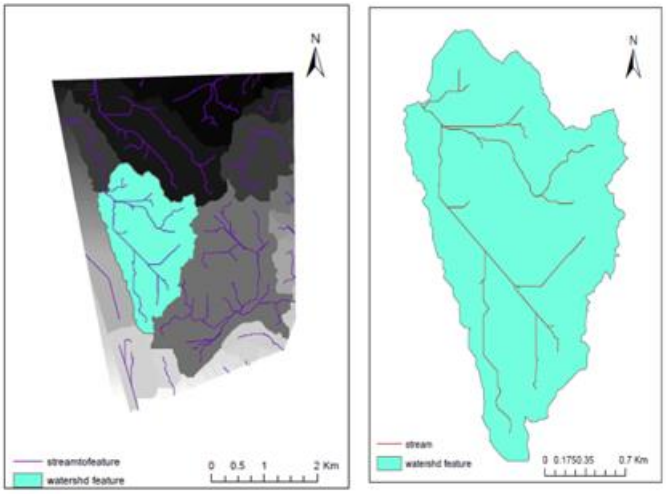


Figure 2: *Selected sub catchment*

Sub catchment which belongs to Bambalapitiya and Kollupitiya area was delineated by using hydrology tools as shown in the Figure 2. After selecting the catchment, land use features were attached by clipping the total features in order to get total impervious area in the selected catchment and compare the land use patterns over a decade.

Characteristics of the selected catchment were used for calculating the runoff of the catchment by using HEC-HMS software. SCS curve method was used for the determination of flood peak flow. Area of the catchment was 3.58 km^2. SCS curve number was 86.5 by considering the conditions of the

selected area. SCS curve number and SCS unit hydrograph were taken as the loss method and transform method respectively.

Flood plain analysis was done in HEC-HMS software by assuming storage method as area elevation method. For that, area elevation data were obtained through the created DEM. Surface volume function in Arc GIS was used for the purpose and obtained the area elevation curve.

Table 2: *Land use variations over a decade*

Land use patterns	2000	2010
	%	%
Building	29.21%	40.44%
Roads	6.71%	8.26%
Water Bodies	3.29%	3.09%
Total impervious area	35.9%	48.7%

Table 3: *Peak discharges for different impervious areas*

Percentage of impervious area (%)	Direct runoff (m^3/s) Peak discharge
35.92	60.9
48.7	62.8
65 (Maximum)	65.1

3. Results

According to the analysis of selected catchment, the land use variation and percentage over a decade can be shown in Table 2.

Percentages of impervious area which were obtained by digitizing were used for calculating the direct runoff of the catchment as shown in Table 3. Following Flood risk map (Figure 3) was obtained from analyzing the selected catchment in HEC-HMS software.

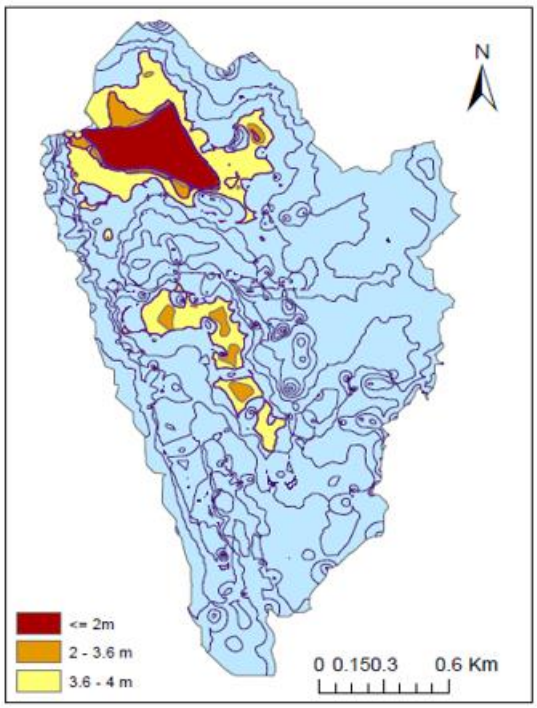


Figure 3: *Flood risk map*

4. Discussion

According to the land use percentages obtained for the selected catchment, there is an increment in the area belongs to buildings and roads in year 2010 compared to year 2000. This can be totally due to urban expansion. When considering water bodies, the area of water bodies has been reduced due to the land fillings. Although there is no marshy lands in selected catchment, there is a significant reduction in marshy lands when considering the entire Colombo area. Satellite images which were taken in year 2000 and 2010 have to be used for digitizing since they were the most available images.

Percentage of impervious area was taken as the final outcome of digitizing the land use patterns. HEC-HMS software was used for determining the surface runoff and the flood plain elevation. Analyse of surface runoff was done by according to the SCS curve method since it has a limited data requirement and high practical applicability. SCS curve number was obtained by considering two elements such that soil type and antecedent soil moisture of the catchment.

In this case, Group C has been selected for the type of soil since the selected catchment is contained of soils high in clay. Type 3 has been selected as the antecedent moisture condition by considering the most critical condition which occurs at the high moisture content.

According to the results obtained from the surface runoff of the selected catchment, surface runoff has been increased when increasing the percentage of impervious area. This implies that flood risk has been increased with increment of land use for buildings and roads. Elevation of the peak flood was obtained by considering the maximum limit of impervious area (65%) since it is the most conservative approach.

There can be deviations between flood elevation obtained from the model and the actual situations. Since few assumptions have been made during the modeling such as SCS curve number, number of outlets, created DEM etc.

After identifying those areas as flood risk areas, we can manage land use to minimize flood damage to people and their properties also. And also drainage system can be modified such that minimize the flood risk.

5. Conclusion

Identification of flood risk areas plays an important role in avoiding or minimizing the hazards. Among the factors affecting floods, land use variation is very important since it can be controlled. Digitizing the satellite images method was used to represent the land use patterns since it can identify number of land uses at a time and difficult to use manual techniques where having a huge variation in land use patterns.

GIS had been used at all these stages to support and speedup the data processing and analysis. GIS has the ability to perform and display different types of professional analysis including to “create, manipulate, analysis and display” all types of geographically or spatially referenced data. HEC-HMS software has the ability to represent the peak flood flow and the flood plain with the minimum number of data.

Analysis is done for only one catchment in greater Colombo area due to the time limitations. But it can also be done for overall research area by analysing each catchment and compiling them.

References

Briviot, P.A., Colombo, R., Maggi, M. and Tomasoni, R. 2002. Integration of remote sensing data and GIS for accurate mapping of flooded areas. *International Journal of Remote Sensing*, 23(3), pp.429-441.

Chen, J., Arleen, A.H. and Lensyl, D.U. 2009. A GIS based model for flood urban flood inundation. *Journal of Hydrology*, 373, pp.184-192.

Samarasinghe, S.M.J.S., Nandalal, H.K., Weliwitiya, D.P., Fowze, J.S.M., Hazarika, M.K. and Samarakoon, L. 2010. Application of remote sensing and GIS for flood risk analysis: a case study at Kalu-Ganga River, Sri Lanka. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, 38(8), pp.110-115.

Sanyal, J. and Lu, X.X. 2003. Application of remote sensing in flood management with special reference to monsoon Asia: a review. *Natural Hazards*, 33, pp.283-301.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2589-2602
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.352

Research Article

Using Geographic Information System (GIS) to Develop Health Information System (HIS) for Srinagar City, Jammu and Kashmir

Qaiser Shafi[1], Junaid Qadir[2*], Faizan Jalal[1]

[1]Department of Earth Sciences, University of Kashmir, Hazratbal, Srinagar – 190006, Jammu & Kashmir, India
[2]Department of Geography & Regional Development, University of Kashmir, Hazratbal, Srinagar – 190006, Jammu & Kashmir, India

Correspondence should be addressed to ***qadirjunaid10@gmail.com.***

Publication Date: 11 May 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.352

Abstract Health Information System (HIS) is a GIS based information system that stores all information about health facilities – their name, location, category, no. of rooms, no. of male doctors, no. of female doctors, total no. of employee, no. of beds, qualification of doctor's, Geographic coordinates of health Centers, availability of medicine, 24 hours service availability, maternity service, canteen facility etc. HIS allows the preservation of the information value of health Centers digitally and offers new exploitation possibilities, like the immediate connection of different kinds of data for analysis, or the digital documentation of these institutes for its improvement. The purpose of this study was to collect appropriate data of Health Centers around Srinagar city and to store the data into the Geodatabase built to manage them more efficiently. The data pertaining to these Health Centers was classified into different categories such as: Public Health Center, Sub Health Center, Allopathic Center, and private Hospitals. At total of 73 Health Centers were mapped out of them 13 were Public Health Center, 29 were Sub Center, 18 were Allopathic Center, 13 were Private Hospitals. The present study focuses on using Geographic information system and integrating with the Health information system to generate a baseline data for the planning, development and maintenance of these health facilities in Srinagar city.

Keywords *Health Information System (HIS), Geographic Information System (GIS), Health Centers*

1. Introduction

With the onset of the epidemiological transition and as the prevalence of infectious diseases decreased through the 20th century, primary health began to put more focus on chronic diseases such as cancer and heart disease. Previous efforts in many developed countries had already led to dramatic reductions in the infant mortality rate using preventative methods. During the 20th century and early in the next, the dramatic increase in average life span is widely credited to primary health achievements, such as vaccination programs and control of many infectious diseases including polio, diphtheria, yellow fever and smallpox; effective health and safety policies such as road traffic safety and occupational safety; improved family planning; tobacco control measures; and programs designed to decrease non-communicable diseases by acting on known risk factors such as a person's background, lifestyle and environment.

One of the major sources of the increase in average life span in the early 20th century was the decline in the "urban penalty" brought on by improvements in sanitation (Denis Postle., 2012). These improvements included chlorination of drinking water, filtration and sewage treatment which led to the decline in deaths caused by infectious waterborne diseases such as cholera and intestinal diseases. Meanwhile, large parts of the developing world remained plagued by largely preventable/treatable infectious diseases and poor maternal and child health outcomes, exacerbated by malnutrition and poverty. Since the 1980s, the growing field of population health has broadened the focus of primary health from individual behaviors and risk factors to population-level issues such as inequality, poverty, and education. Modern primary health is often concerned with addressing determinants of health across a population. There is recognition that our health is affected by many factors including where we live, genetics, our income, our educational status and our social relationships - these are known as "social determinants of health" (Burke, 2013). A social gradient in health runs through society, with those that are poorest generally suffering the worst health. However even those in the middle classes will generally have worse health outcomes than those of a higher social stratum. The new primary health seeks to address these health inequalities by advocating for population-based policies that improve health in an equitable manner.

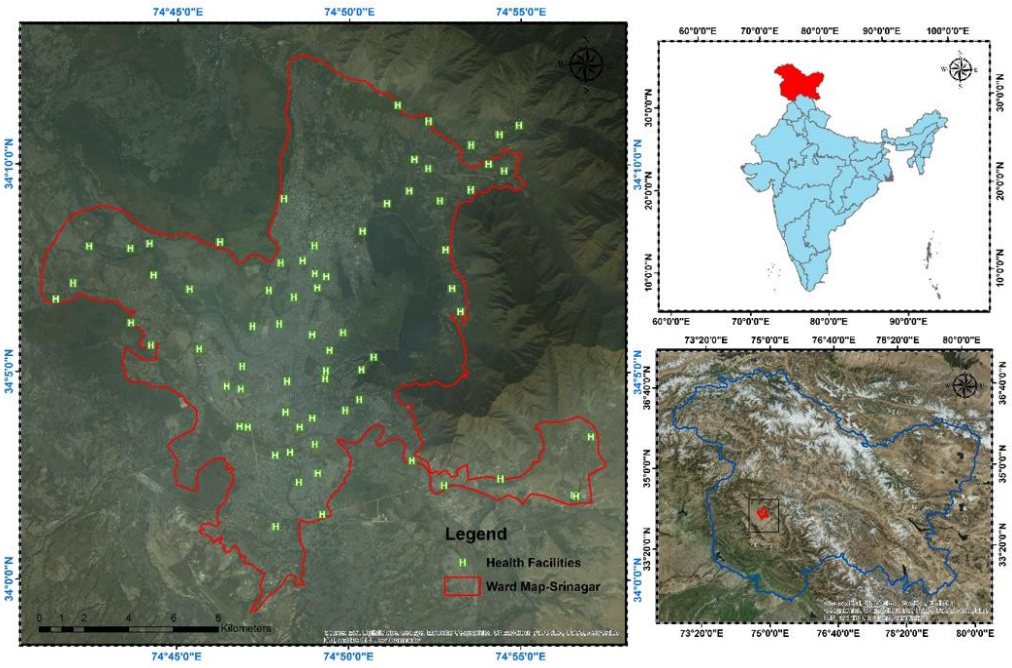


Figure 1: *Ward Map of Srinagar City, showing all Health Facilities*

Public health applications of GIS technology are still in the early stages. Many challenges remain that need to be addressed before the full potential of GIS technology can be realized for public health practice, planning, and research. One of the greatest challenges is to incorporate epidemiologic principles and methods into the analysis to be mapped. GIS technology provides public health practitioners and researchers with several new types of data. Public health practitioners can also use digital imagery from satellites or aerial photos to add details to improve the accuracy of a mapping projects. As new GIS methods are developed, they can be added to the "toolkits" of epidemiology and health services research. Exeter in 1998, also concluded that Compared with tables and charts, maps developed using GIS technology can be an extremely effective tool to help community decision makers visualize and understand a public health problem (Richards, 1999). The current study has tried to map each and every health facility located in the Srinagar city in a comprehensive Geodatabase

format, where in within a single click of the mouse we can know attributes such as location, name of health facility, in charge, no. of doctors, facilities available etc. the study holds significance for our health department which will find it handy for planning various projects relating to the maintenance and development of health facilities.

1.1. Study Area

Srinagar city is located in the valley of Kashmir at an altitude of 1,730 m above sea level. The city lies on both banks of the Jhelum River, a tributary of the Indus River. The city has the unique physiographic setup with steep hills in the east and north east, low lying paddy fields forming flood plain of Jhelum in the south and west and raised plateau lands in the south. The valley is surrounded by the Himalayas on all sides. Srinagar lies between the coordinates 34^0 01` N to $34^0$ 27`N latitude and 74^0 36 E to 75^0 30 E, of over an area of 105 Sq. km (Figure 1). Winters are cool, with daytime temperature averaging to 2.5°C (36.5 °F) and drops below freezing point at night. Srinagar has Mediterranean type of climate with percentage humidity varying from 90% in winter to 78% in summer months. The average annual rainfall is around 720 mm.

2. Methodology

In this Study, GPS was used to collect the location (geographic co-ordinates) of various Health Centers as per the data provided by Chief Medical officer (CMO). These GPS points were then overlaid on high resolution GeoEye-1 image. The overall methodology adopted for the current study is shown in (Figure 2).

The Health data pertaining to these Health facilities were classified into different thematic maps or layers in Arc Map 10.1 as given below:

1. Public Health Center
2. Sub Health Center (Sub-Center)
3. Allopathic Center
4. Private Hospitals

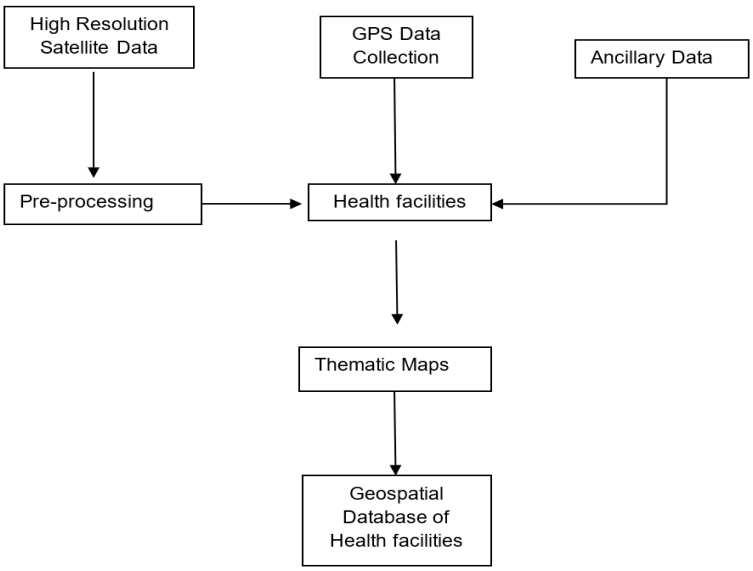


Figure 2: *Flow chart of Methodology*

The above generated thematic maps were stored in the Geodatabase format to generate the final geospatial database of Health Centers in Srinagar City.

2.1. Data Sets

To carry out the study following data sets were used:

1. GeoEye-1
2. Global Positioning System (GPS) data
3. Ancillary data

GeoEye-1 was launched on September 6, 2008. The satellite separated successfully from its Delta II launch vehicle at 12:49 pm 58 minutes and 56 seconds after launch (Justin Ray, 2008). The satellite provides 41 centimetres (16 in) panchromatic and 1.65 meter multispectral imagery in 15.2 km swaths. The spacecraft is intended for a sun-synchronous orbit at an altitude of 681 km (423 mi) and an inclination of 98 degrees, with a 10:30 a.m. equator crossing time. GeoEye-1 can image up to 60 degrees off nadir. It is operated out of Herndon, Virginia and was built in Arizona by General Dynamics Advanced Information Systems (https://en.wikipedia.org/wiki/GeoEye).

2.2. GPS Data Collection

In the present study an in-depth mapping of Srinagar city using handheld Juno SB GPS was carried out to generate the geospatial database of Health Centres.

2.3. Ancillary Data

The ancillary data was generated from the data given at the CMO (Chief Medical Office). The ancillary data in the form of hard copy format was then digitized to generate the database of Health Centers of Srinagar city in Excel format. It includes name and location of these institutes.

2.4. Pre-Processing

For the present study a high resolution GeoEye-1 of October 2008 was used as the satellite data source. Same data was first geo-referenced by field GCPs in order to make it reliable for the study. Also, the Health institutional data collected by GPS was processed in GPS pathfinder in order to make the data readable in Arc Map 10.1 and remove errors that had occurred during the field survey.

Table 1: *Showing the type of data obtained from various Health Centers in Srinagar city, J&K*

Name of Public Health Center	**Medical Amenity**	**Name of the town**	**No. Of doctors**	**No. of beds**	**No. of rooms**	**Latitude**	**Longitude**
Facilities available	No. of nurses and ward boys	Elevator	Blood Bank	Parking	Canteen	Mode of electricity	Infrastructure

2.5. Data Integration

The Health Centres database developed in Excel format was converted to .dbf format (Table 1) and integrated with the point locations of Health institutes through a joining process in Arc Map 10.1.

3. Results and Discussion

3.1. Mapping of Different Health Facilities

During this Study, an extensive field Survey was carried out to map Health Centers of Srinagar City using GPS. At present total of 73 Health Centers were mapped out of them 13 were Public Health Center, 29 were Sub Health Center, 18 were Allopathic Center, 13 were Private Hospitals to make a Geodatabase containing the Spatial & Non-Spatial information of each Site. During the field survey following information was also incorporated to develop a robust Health Information System:

- Health records.
- Non-Professional staff.
- Doctors' data (Number, Qualification, experience, placement).
- Facility inventory data (location, geographic co-ordinates, stories, type, number of rooms, equipment.
- Facilities available.
- Infrastructure.
- Monitoring of internal management initiatives (e.g. special projects).
- No of beds, rooms, ambulances, doctors etc.

3.2. Public Health Centres

These are facilities that serve to protect the health of city residents. They are central institutions for community health and sanitation. Public health Centers conduct specialized programs, such as food sanitation, environment sanitation, emotional health, and tuberculosis/contagious, disease, countermeasures. In addition, in order to provide accessible health services to everyone in the community. These sites were mapped using GPS and the generated point theme was then stored in the Geodatabase. The total of 13 primary health Centers was mapped in Srinagar district. Some of the public health facilities lack the staff, equipment's, facilities, medicine etc. While most of them are in satisfactory condition. The distribution of these health facilities is shown in (Figure 3).

3.3. Sub Health Center

In the public sector, a Sub-health Centre (Sub centre) is the most peripheral and first contact point between the primary health care system and the community (Shinde et al., 2014). As per the population norms, one Sub centre is established for every 5000 population in plain areas and for every 3000 population in hilly/tribal/desert areas (Shinde et al., 2014). A Sub centre provides interface with the community at the grass-root level, providing all the primary health care services (Sharma et al., 2015). As sub centres are the first contact point with the community, the success of any nationwide programme would depend largely on well-functioning sub centres providing services of acceptable standard to the people.

Of particular importance are the packages of services such as immunization, antenatal, natal and postnatal care, prevention of malnutrition and common childhood diseases, family planning services

and counseling. They also provide elementary drugs for minor ailments such as ARI, diarrhea, fever, worm infestation etc. and carryout community needs assessment. Besides the above, the government implements several national health and family welfare programmes which again are delivered through these frontline workers. There are total of 29 Sub Health centers in Srinagar city that are mapped here in (Figure 4).

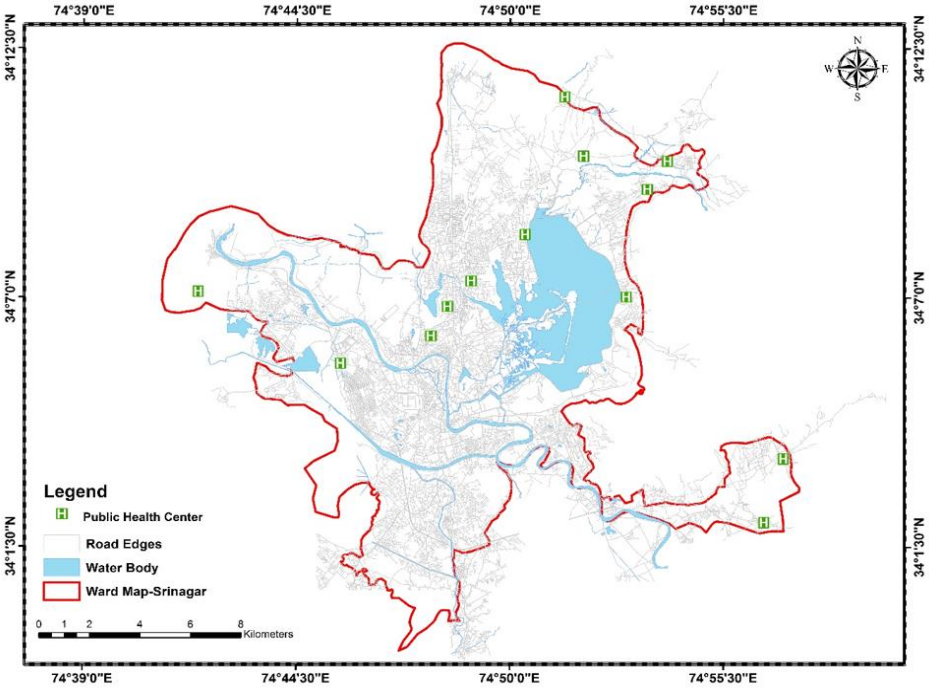


Figure 3: *Total number of Public Health centers in Srinagar city*

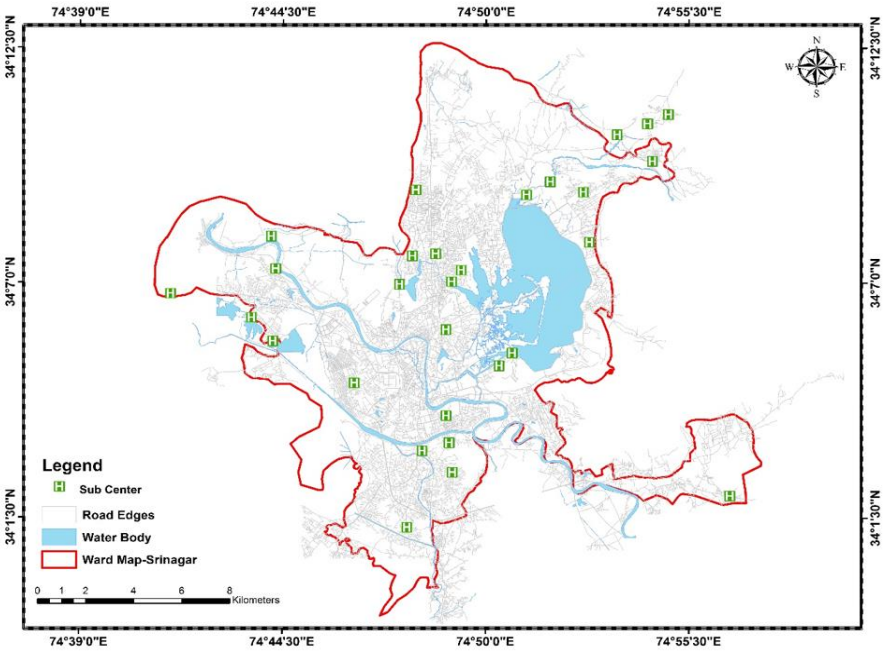


Figure 4: *Total number of Sub Health centers in Srinagar city*

3.4. Allopathic Dispensaries

Allopathic means treating disease by using a substance antagonistic to the condition, such as antibiotics to treat an infection; or replaces something in the body. Facilities that handle allopathic medicines prescribed in association with traditional Western treatment may call themselves pharmacy, while those that deal with alternative and complementary medicine like Chinese herbs may be called dispensaries. Trained staff, at dispensary can process written orders for medications, which may arrive electronically if the system is connected to such systems. They can confirm the medication and the dosage, prepare it, and package it appropriately with directions for use. These sites were mapped using GPS and the generated point theme was then stored in the Geodatabase. The totals of 18 allopathic dispensaries were mapped in Srinagar city. Some of these Allopathic facilities lack the staff, equipment's, facilities, medicine etc. While most of them are in satisfactory condition. The distribution of these health facilities is shown in (Figure 5).

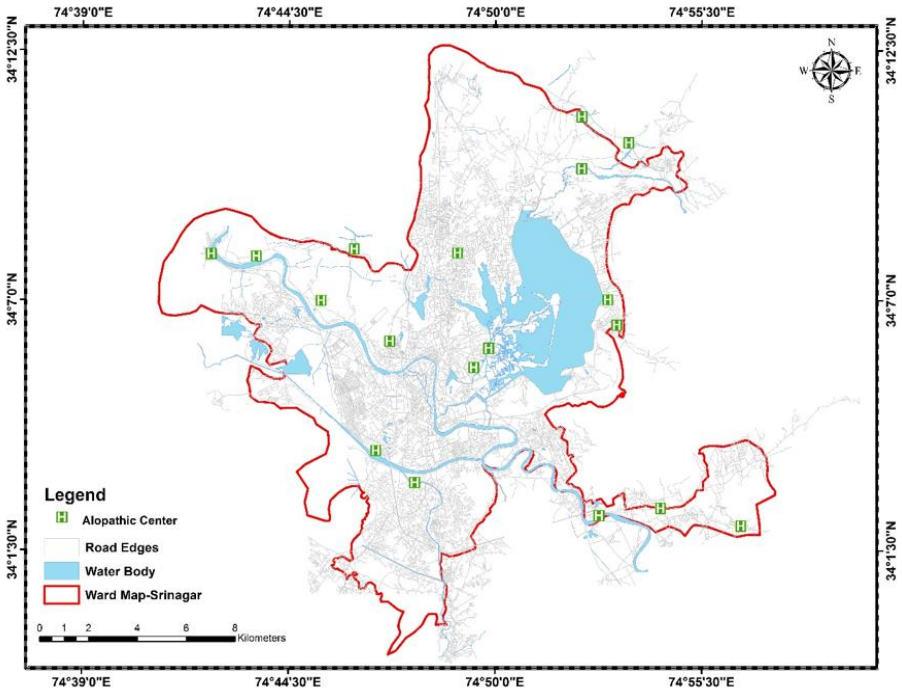


Figure 5: *Total number of Allopathic centers in Srinagar city*

3.5. Private Hospitals

A hospital similar to a group hospital except that it is controlled by a single practitioner or by the practitioner and the associates in his or her office (Farlex Partner Medical Dictionary, 2012). A private hospital are owned by a non-profit organization and privately funded through payment for medical services by patients themselves, by insurers. Private healthcare is when doctors, dentists, and other healthcare providers are paid for through private insurance and (occasionally) out of private bank accounts as this is in contrast with a public system, in which they are paid by the government money (http://www.ehow.co.uk/facts_6782488_definition-private-healthcare.html). The biggest distinguishing feature of private health care is that it is run with the goal of making money (http://www.ehow.co.uk/facts_6782488_definition-private-healthcare.html). A total of 13 Private Hospitals from Srinagar city were mapped and distribution of these Health facilities is shown in (Figure 6).

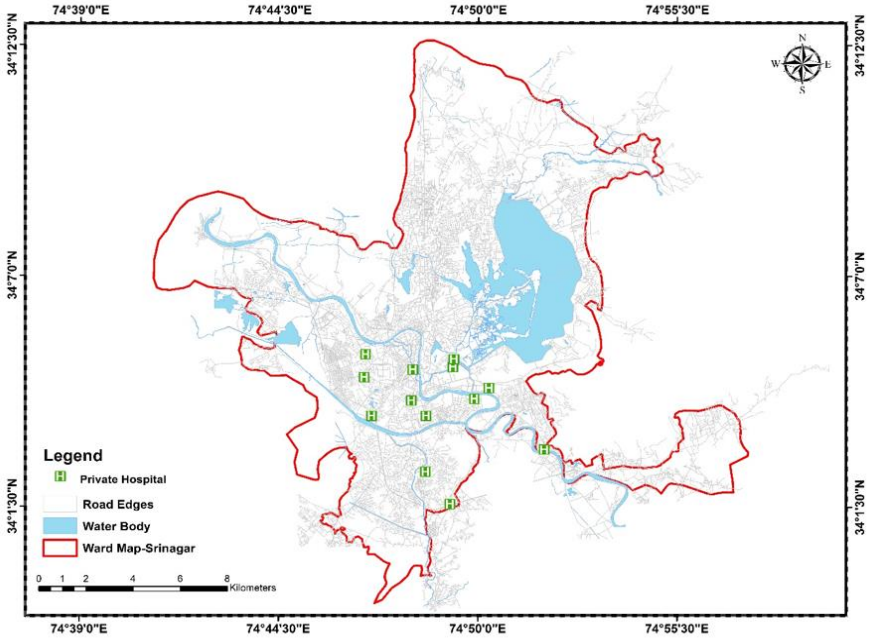


Figure 6: *Total number of Private Hospitals in Srinagar city*

4. Using GIS to develop Health Information Systems (HIS) for Srinagar City,J&K.

Health Information System (HIS) is a GIS based information system that stores all information about health facilities – their name, location, category, no. of rooms, no. of male doctors, no. of female doctors, total no. of employee, no. of beds, qualification of doctor's, Geographic coordinates of health Centers, availability of medicine, 24 hours service availability, maternity service, canteen facility etc. Health information system allows preservation the information value of Health facilities digitally and offers new exploitation possibilities, like the immediate connection of different kinds of data for analysis, or the digital documentation of these facilities for its improvement.

GIS plays a critical role in the decisions on where and when to intervene, improving the quality of care and accessibility of services, finding the most cost-effective delivery modes, and protecting patient confidentiality while satisfying needs of the research community on data accessibility (Hanjagi et al., 2007). Geographic Information System is able to organize all the routes that a health care professional has to follow and it can take into account other parameters, too (Fradelos et al., 2014). Generally, GIS application areas might be applied towards Strategic Planning, Research and Evaluation, emergency preparedness and both response and location of health care services, too (Smith et al., 2007). Geographic Information Systems provides a tremendous convenience for health care providers as regards the organization and the management of these services (Fradelos et al., 2014). The use of GIS and spatial representation of various health issues make professionals arrive at conclusions in a faster and better way in the field of both public health and decision-making (Hanchette et al., 2003). Hence, the organization and coordination of various services would be easier and more efficient. The healthcare provider may direct quickly and efficiently the patient to suitable health care services (Najafabadi, 2009). The purpose of this study was to highlight Geographic Information Systems also acting as a decision-making tool in health care and contribute to the formulation of policies regarding the health sector in the Srinagar city. Geo Database prepared for different Health facilities (Public Health Center, Sub Health Center, Allopathic and Private Hospital) are shown below (Table 2-5).

Table 2: *Attribute Data; Health Information System (HIS) of Public Health Center Hazratbal, Narwara, Nishat & Lal Bazaar areas of Srinagar city.*

S.no	Name of town	Hazratbal	Narwara	Nishat	Lal Bazar
1	Medical amenity	Public Health Center			
2	No. of doctors	7	3	2	2
3	No. of male doc.	4	1	*	*
4	No of female doc.	3	2	2	2
5	No of storeys	2	2	2	1
6	No of beds	4	4	4	6
7	No of ambulances	2	*	*	*
8	No of rooms	12	12	27	8
9	No of Op. theaters	1	*	*	*
10	No of ward boys	2	*	*	*
11	No of Nurses	2	*	*	*
12	No of Fmphw worker	*	2	2	2
13	No of NO'S	2	3	1	1
14	No of pharmacists	2	1	1	*
15	Facility available	Lab,xray,Usg Etc	Lab	Lab,XRay,Usg Etc	HB
16	Maternity service	No	N O	No	No
17	Elevator facility	No	No	No	No
18	24 hours service	Yes	No	No	No
19	Blood bank	No	No	No	No
20	OPD availability	Yes	Yes	Yes	Yes
21	Medicine availability	Yes	Yes	Yes	Yes
22	Canteen facility	No	No	No	No
23	Parking facility	Yes	No	No	No
24	Infrastructure	Satisfactory	Satisfactory	Satisfactory	Satisfactory
25	Mode of electricity	AC	AC/DC	AC	AC

Table 3: *Attribute Data; Health Information System (HIS) of Sub Health Center Jawahar Nagar, Mehjoor Nagar, Kushipora & Rawalpora areas of Srinagar city*

S.no	Name of town	Jawahar Nagar	Mehjoor nagar	Khushipora	Rawalpora
1	Medical amenity	Sub Center			
2	No. of doctors	1	1	*	*
3	No. of male doc.	*	*	*	*
4	No of female doc.	1	1	*	*
5	No of storeys	1	1	1	1
6	No of beds	3	4	*	*
7	No of ambulances	1	1	*	*
8	No of rooms	5	4	3	3
9	No of op. theaters	*	*	*	*
10	No of ward boys	2	*	*	*
11	No of Nurses	1	3	*	*
12	No of Fmphw worker	*	*	2	2
13	No of NO'S	*	*	*	*
14	No of pharmacists	1	1	*	*
15	Facility available	Hb	Hb	*	*
16	Maternity service	No	No	No	No
17	Elevator facility	No	No	No	No
18	24 hours service	No	No	No	No
19	Blood bank	No	No	No	No
20	OPD availability	Yes	Yes	Yes	Yes
21	Medicine availability	Yes	Yes	Yes	Yes
22	Canteen facility	No	No	No	No
23	Parking facility	No	No	No	No
24	Infrastructure	Satisfy	Satisfy	satisfy	Satisfy
25	Mode of electricity	AC	AC	AC	AC

Table 4: *Attribute Data; Health Information System (HIS) of Allopathic Center Chattar Hama, Tengpora, Rainawari & Miskeen Bagh areas of Srinagar city*

S.no	Name of town	Chatter Hama	Tengpora	Rainawari	Miskeen Bagh
1	Medical amenity	Allopathic Center			
2	No. of doctors	1	2	2	1
3	No. of male doc.	*	*	*	*
4	No of female doc.	1	2	2	1
5	No of storeys	1	1	2	1
6	No of beds	1	3	2	5
7	No of ambulances	*	*	*	*
8	No of rooms	4	4	3	6
9	No of op. theaters	*	*	*	*
10	No of ward boys	*	*	*	*
11	No of Nurses	*	*	*	*
12	No of Fmphw worker	1	3	3	1
13	No of NO'S	*	*	3	1
14	No of pharmacists	1	1	1	1
15	Facility available	*	hb	hb	*
16	Maternity service	No	No	No	No
17	Elevator facility	No	No	No	No
18	24 hours service	No	No	No	No
19	Blood bank	No	No	No	No
20	OPD availability	Yes	Yes	Yes	Yes
21	Medicine availability	Yes	Yes	Yes	Yes
22	Canteen facility	No	No	No	No
23	Parking facility	No	No	No	No
24	Infrastructure	Rented	satisfy	satisfy	satisfy
25	Electricity mode	AC	AC	AC	AC

Table 5: *Attribute Data; Health Information System (HIS) of Private Hospitals Gogji Bagh, Hazuri Bagh, Raj Bagh, Bishembar Nagar areas of Srinagar city*

S.no	Name of town	Gogji Bagh	Hazuri Bagh	Rajbagh	Bishember Nagar
1	Medical amenity	Ramzana	Khanam's	Modern	Valley Orthocraft
2	No. of doctors	4	4	8	5
3	No. of male doc.	1	1	0	5
4	No of female doc.	3	3	8	0
5	No of storeys	3	4	5	4
6	No of beds	35	25	50	18
7	No of ambulances	2	1	2	1
8	No of rooms	45	50	65	29
9	No of op. theaters	2	3	3	2
10	No of ward boys	20	5	8	6
11	No of nurses	12	9	40	8
12	No of Fmphw worker	*	*	*	*
13	No of NO'S	*	*	*	*
14	No of pharmacists	3	3	12	8
15	Facility available	Lab,usg, Ecg Xray	Lab,usg, Ecg Xray	Lab,usg, Ecg Xray	Lab, x ray
16	Maternity service	Yes	Yes	Yes	No
17	Elevator facility	Yes	No	Yes	Yes
18	24 hours service	Yes	Yes	Yes	Yes
19	Blood bank	No	No	No	No
20	OPD availability	Yes	Yes	Yes	Yes
21	Medicine availability	Yes	Yes	Yes	Yes
22	Canteen facility	Yes	Yes	Yes	Yes
23	Parking facility	Yes	Yes	Yes	Yes
24	Infrastructure	Yes	Yes	Yes	Yes
25	Mode of electricity	AC/DC	AC/DC	AC/DC	AC/DC

5. Conclusion

The growing rate of diseases, improper health care to women and children is the biggest issue today. Viewing these issues from a historic standpoint is important to understand where and why these issues began. By knowing the historic information, it is easier to discover what has and has not worked, and what issues continually transpire with health in the entire region. During this Study, an extensive field Survey was carried out to map Health Centers of Srinagar City using GIS. At present total of 73 Health Centers were mapped out of them 13 were Public Health Center, 29 were Sub Health Center, 18 were Allopathic Center, 13 were Private Hospitals. During the study it was found the most of the Health institutes (Public Health Center, Sub Center & Allopathic Center's) lack basic infrastructure and even some of them operate in rented accommodations. Fund allocated by the state & central agencies are not properly utilized and this leads to shortage staff, medicines, latest equipment's, and Lab facilities etc. Lack of competent Doctor's & supporting staff, Unbalanced patient - Doctor ratio, Management issue's, Drug laws and corruption in these facilities has taken a toll on the lives of the people at large. Reluctance by the Doctor's to serve in the rural areas has affects the lives of the people at large scale, which in turn has resulted in overburdening of these Health institutes which at times leads to the chaos. On the other side Private Hospitals have been flourishing at greater pace, there has been considerable amount of upgradation and improvement in the facilities provided to people.

Acknowledgement

Authors would like to thank Head of Department Earth Sciences, University of Kashmir for providing requisite facilities and support and CMO (Chief Medical Office) for valuable suggestions and Esri, Digital Globe, USGS for accessing High resolution data.

References

Anand Sharma, Vibhakar Mansotra, Sourabh Shastri. 2015. An Exploratory Analysis of Public Healthcare Data: A Case Study of Jammu & Kashmir State. *Asian Journal of Computer and Information Systems,* 3(5).

Exeter, D.J. An evaluation of cartographic visualization techniques for epidemiology [master's thesis]. Auckland [NZ]: University of Auckland Department of Geography; 1998. Available at: ftp://ftp.geog.auckland.ac.nz/pub/outgoing/dxthesis.

Fradelos, E.C., Papathanasiou, I.V., Mitsi, D., Tsaras, K., Kleisiaris, C.F. and Kourkouta, L. 2014. Health based Geographic Information Systems (GIS) and their applications. *Acta Informatica Medica*, 22(6), pp.402-405.

Geo Eye-1. Available from: https://en.wikipedia.org/wiki/GeoEye-1. Accessed 07 Jan 2017.

Hanjagi, A., Srihari P. and Rayamane, A. 2007. *A public health care information system using GIS and GPS: a case study of Shiggaon*. In: Lai P.C., Mak A.S.H. (eds) GIS for Health and the Environment. Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg.

Hanchette, C.L. 2003. *Geographic Information Systems*. In: O'Carroll, P.W, Yasnoff, Y.A. editors. Public Health Informatics. New York: Springer, pp.431-466.

Justin, R. 2018. *Delta 2-335 Mission Status Center*. Accessed 07 Jan 2018.

Burke, M.C. 2013. *Human trafficking: interdisciplinary perspectives*. Routledge.

Mahadeo B. Shinde, P.M. Durgawale. 2014. Nursing Audit of Health Workers Providing Health Services in Rural Area with Special Emphasis to Community Satisfaction in Satara District. *International Journal of Science and Research. pp.*2319-7064.

Najafabadi, A.T. 2009. Applications of GIS in Health Sciences. *Shiraz E-Medical Journal*, 10(4), pp.221-230.

Postle, Denis. 2012. Therapy Futures: Obstacles and Opportunities, London: Wentworth Learning Resources.

Proprietary hospital. (n.d.) Farlex Partner Medical Dictionary. 2012. Retrieved January 7, 2018. Available from: https://medical-dictionary.thefreedictionary.com/proprietary hospital.

Richards, T.B., Croner, C.M., Rushton, G. and Fowler, L. 1990. Geographic information systems and public health: mapping the future. *Public Health Reports*, 114, pp.359-373.

Smith, M.J., Goodchild, M.F. 2007. Geospatial Analysis – The comprehensive guide to principles, techniques and software tools. Available from: http://www.spatialanalysisonline.com/.

Sam, G. 2017. Definition of Private Hospital. Available from: http://www.ehow.co.uk/facts_6782488_definition-private-healthcare.html. Accessed on 07 Jan 2018.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2719-2727
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.368

Research Article

Site Suitability Analysis for Local Airport Using Geographic Information System

T.D.C. Pushpakumara, A.N.C. Rodrigo

Department of Civil Engineering, University of Moratuwa, Sri Lanka

Correspondence should be addressed to *thudugalagedon@gmail.com*

Publication Date: 3 July 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.368

Abstract Geographic Information System (GIS) are now recognized broadly as a valuable tool for managing, analyzing and displaying large set of data relevant many local and regional planning and analyzing activities. The complex nature of Airport site suitability issues can be sorted it out using GIS tools. This research will discuss some potential of GIS application in site suitability analyzing considering factors that can affect to the Airports for entire Sri Lanka (considering forest area, tanks, airborne restricted area, and streams) and closer study for Polonnaruwa district, as there is a higher tourist attraction for historical and ancient cities. Anuradhapura and Polonnaruwa districts give considerable amount of contribution for that. Therefore Polonnaruwa district will be closely analyzed considering all the factors that will affect the selection of area for Airports. In this study, the factors considered, will lead to generate areas which are not suitable for local airports in Sri Lanka. Hence using Erase tool under GIS tool box overlay function erase all those areas as not suitable areas for local Airports. Rest can be defined as suitable areas for local airports. Also the buffering tool can be used to identify the most suitable areas by considering land use patterns, road and railway networks.
Keywords *Buffer and GIS; Erase; Overlay; Site suitability analyze*

1. Introduction

Geographic Information system has been applied to many disciplines including geography, forestry, urban planning, tourism planning, land use, site suitability analysis (Ahesan and Marson, 2015) etc. Nowadays GIS is recognized widely as a valuable software for managing, analysing, and displaying large volumes of diverse data related to many local and regional planning activities (Pareta, 2013). GIS can be considered as a tool box of techniques and technologies which is widely applicable to analyse site suitability.

Inhere GIS software is used to propose most suitable location for local airports by considering Environment, Economic, Social and Cultural aspect. Consideration of Environment, Economic, social aspects are as follows:

Environment – distance from forest area, animals and its behavior near to location, weather condition, and main environment effect of selected location.

Economic – Land value, mobilization cost, no of users, road network, distance from other airports, tourism. Effect of industrial, agriculture, service sector.

Social and cultural- distance from residential area, effect of living condition and quality, distance from places which have local people and tourist attraction.

Land use analysis is more important for site suitability analysis of local airports. As it is difficult to identify the nature of land, variation of land usage of particular time duration etc, without land use analysis.

Thus by using GIS software it is possible to find out most suitable locations for local airports which have less effect an environment, animals and community and more effective for economy.

Objectives

Main Objective

- The main objective is to find out most suitable locations for local Airports in Sri Lanka.

Other objectives

- To promote Tourists attraction towards Sri Lanka.
- Provide a proper Transport mode to local transportation system.
- To Prevent Economic, Environment and Social issues which can be occurred during construction stage and after completion.
- Develop peoples' living standards by providing new utilities.

2. Methodology

Following steps were followed to reach the above mention objectives:

- Studying of the background of site suitability analysis and Airport selection.
- Getting familiar with the topic with past research and publications on site suitability analysis and remote sensing and GIS techniques.
- Collecting data from Survey Department of Sri Lanka, from Gazette and from internet and google earth.
- Buffer all data with relevant distance according to Airport site selection techniques.
- Erase buffering areas from base map which are identified as unsuitable areas for Local Airports.
- Verification and discussion on analyzed data.
- Generate site suitability Map for Study area.
- Conclusion about final results.

Data Collection

Main source of data collection was the Survey Department of Sri Lanka. Most of the Shape files which are used for analysis were created based on data given by Survey Department. Following Figures 1, 2 illustrate the shape files which used to analyze site suitability for local Airport.

Figure 1: *Map of Konduruwawa area*

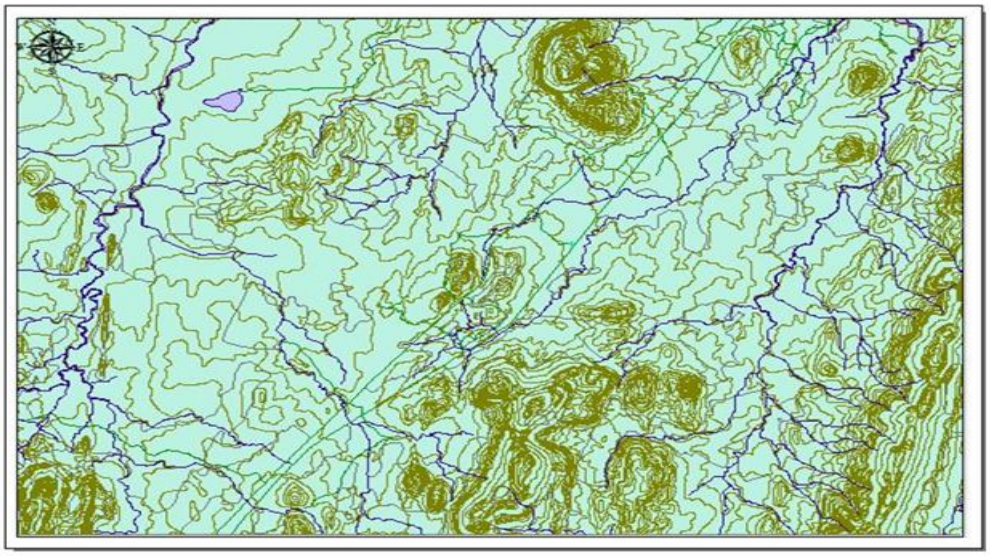


Figure 2: *Forest shape file*

Similarly, tank, streams, road and railway shape file are used.

- The Gazette of the Democratic Socialist Republic of Sri Lanka (EXTRAORDINARY), Air Navigation (Air Defense) Regulation No. 1 of 2007
- Google Earth

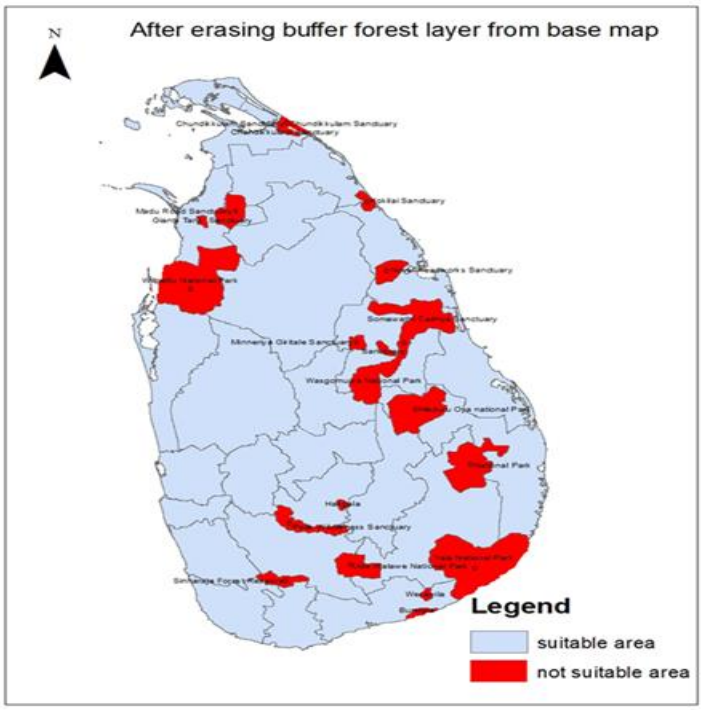


Figure 3: *Base map after erasing buffer forest layer*

3. Data Analyzing and Results

Main tool used for the analysis was ArcMap 10.3 which is one of the premier computer software used for Remote Sensing and GIS studies. Analyzing stage follows the following steps:

Analysis for Entire Sri Lanka

Forest Area in Sri Lanka

Input all Shape files (district, forest, tank, streams and road) which were collected from Survey Department. Then using ArcMap proximity buffer tool buffer forest layer with 1km along the perimeter of the forest boundary and using overlay erase tool erase buffered area from base map as not suitable areas for local Airports.

All Tanks in Sri Lanka

The next step is to consider the tank layer for the analysis. Therefore, first of all using Extract clip tool tank layer should be clip in to the Base map which is after erasing buffer forest layer (Figure 3). Then using buffer tool buffer tank layer with 1km along perimeter of the tank boundary and using overlay erase tool erase buffered tank area from base map after erasing buffer forest layer.

There is another method to do this thing which is buffer tank layer using Base map (Original map) along 1 km from tank boundary and erase this area from the Base map and using overlay Intersect tool get intersection between Base map after erasing buffer tank area and Base map (Figure 4) after erasing the buffer forest area. Both methods are generated same result.

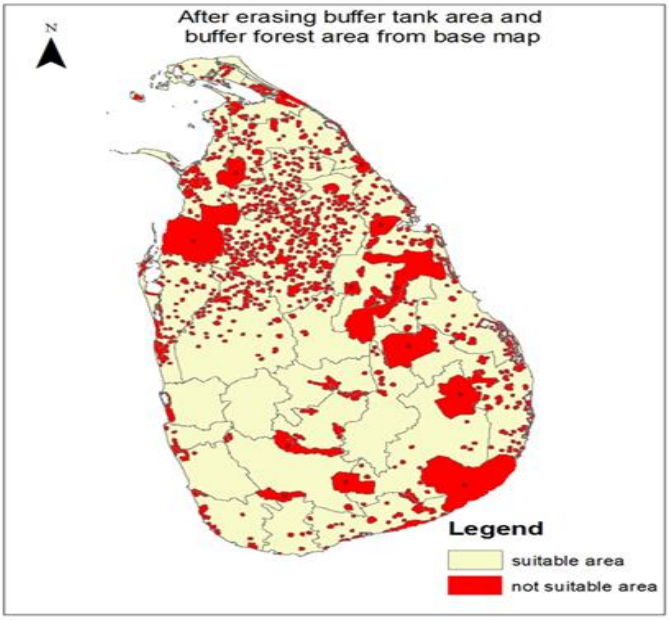


Figure 4: *Map after erasing buffer forest layer and erasing buffer tank base layer*

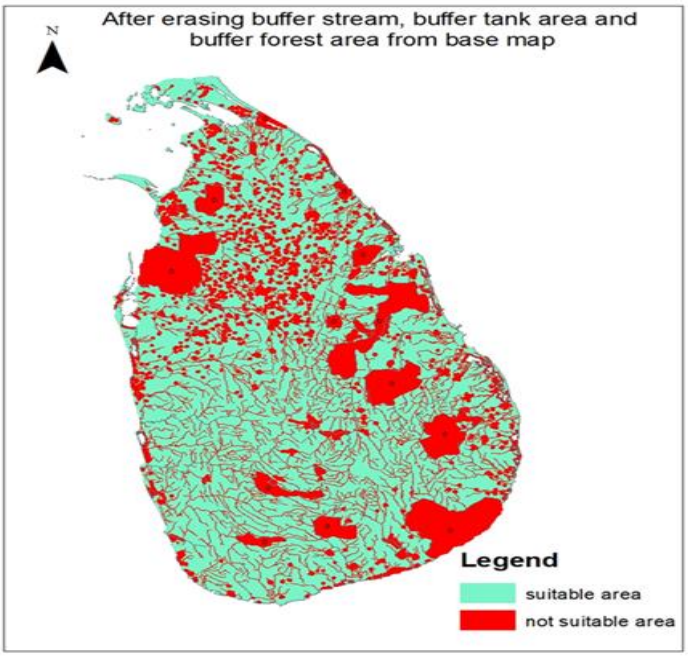


Figure 5: *Base map after erasing buffer stream, buffer forest layer and erasing buffer tank layer*

All streams in Sri Lanka

Same procedure was done which mentioned clause 5.1.2 to analyze stream.

Air Defense Restricted area

This is the final step of the site suitability analysis for entire Sri Lanka. Next steps will be carried out to analyze the closer study area. Because the free data available which relevant to in this study are limited. As the result of that closer study will be carried out in Polonnaruwa District considering Land use pattern later. In this step created shape file for the air defense restricted area (see Figure 5 and Table 1).

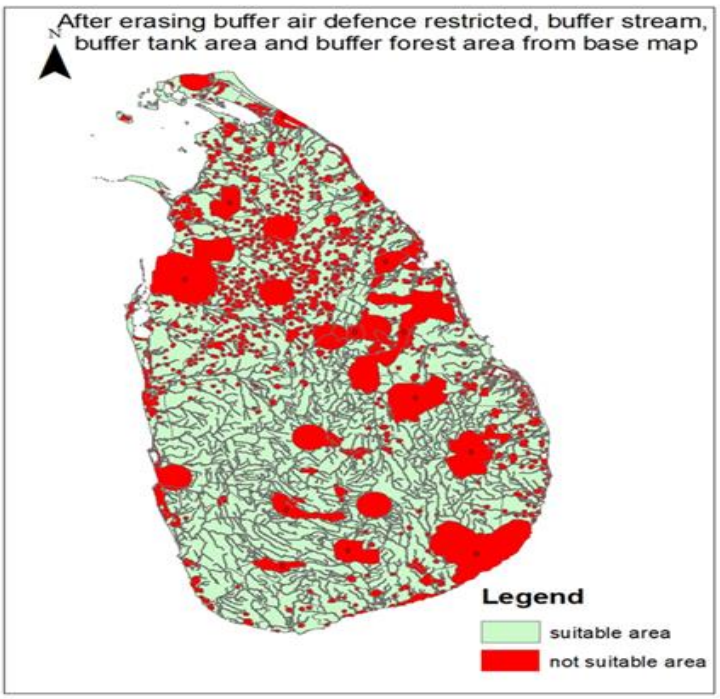


Figure 5: *Base map after erasing buffer air defense restricted area buffer stream buffer forest layer and erasing buffer tank layer*

Table 1: *Suitability of different type of areas for locate airport site*

Type of land area	Suitability to locate Airport site
Build up areas	Very suitable
Sand area	Suitable
Cultivation area, bare and boggy area	Not suitable

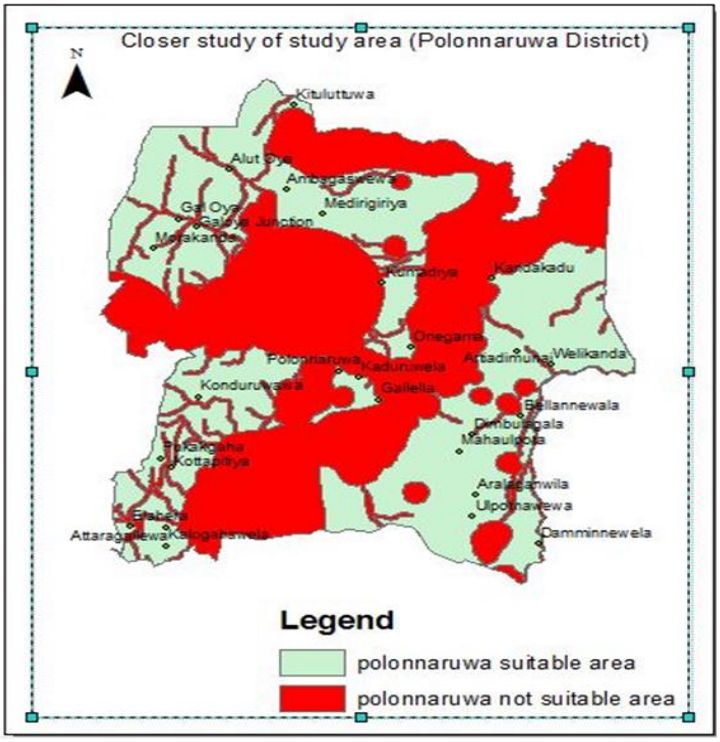


Figure 6: *Site suitability map for Polonnaruwa district*

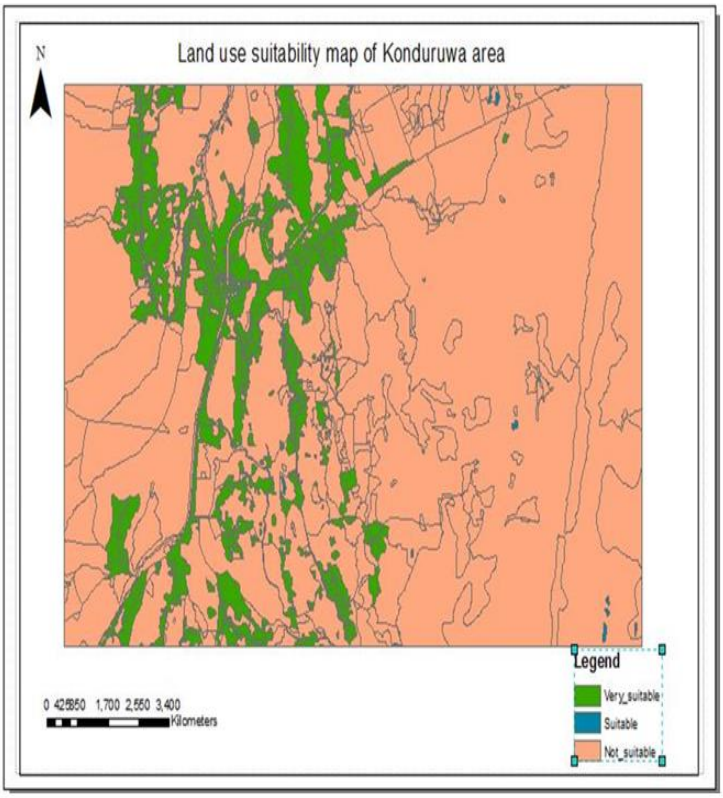


Figure 7: *Land use suitability map for Konduruwawa area*

Analysis of Closer Study Area

After all steps mention in above are carried out, closer study area should be selected based on economic, social and environment facts due to data and resources availability. Therefore scope is concentrated to the Polonnaruwa district. Following map is site suitability map for Polonnaruwa district.

Konduruwawa area was selected among suitable locations of Polonnaruwa district for further analysis. Land use pattern was considered for area to identify the most suitable location for local Airport. Also comparison and analysing will be done for this area to find out final suitable location.

Land Use

Using land use data given by survey department generate shape files. Those shape files defined as cultivation area, bogy area, bare area, rock area, build up area and sand area. Those areas suitability for Airport site was categorized as follows by considering economic, social and environment aspects.

4. Discussion

Basically this study was conducted to find and analyze the suitable locations for local Airports. As mentioned above different factors were considered in order to find suitable locations. All those factors were negative when consider in the suitability of airport site. That's why those areas classified as unsuitable area for local airports. Not only the exact area of forest, tank and streams but also some additional areas also defined as unsuitable areas. As buffering was done for these layer by giving some tolerant distance. It may vary according to the effect of considered factors and regulations of the airports sites or characteristics of itself.

Study area was reduced due to lack of data and resources. Polonnaruwa district was selected by considering tourist attraction and distance from other tourist attraction places. Figure 6 shows the suitable areas for local Airport. This is just approximation. Due to that reason, land use data have to be used to locate exact locations for Airport sites. Among these areas, Konduruwawa, Madirigiriya, Aranaganwila areas can be considered as most suitable areas. Among them Konduruwawa area was selected for further analysis as there weren't sufficient land use data available for Madirigiriya and Aranaganwila areas.

Figure 7 shows land use analysis for 40km^2 of Konduruwawa area. Buildup areas can be identified as most suitable areas for Airport. As they are having already developed infrastructure facilities and services which may increase the tourist attraction. Also these areas have good accessibility and mobility. Those areas are highlighted in green colour in Figure 7. Even though buildup areas are defined as suitable areas, airport sites should be located far from human sensitive buildings such as school, hospital etc. Otherwise it may create social issue, as there will be some disturbance to those facilities due to noise and vibration produced by aircrafts during operational stage.

Cultivation areas, bare and boggy areas in which flora and fauna live are also defined as unsuitable areas. These areas can be considered as environment sensitive areas. Also these areas can be flooded during rainy seasons and it will help to protect other areas from flooding. Environment impact can be eliminated or reduced by omitting these areas as unsuitable areas for Airport sites. As a result of that environment issues which are very sensitive to the community will be not formed during construction and operational stage of Airport.

5. Conclusion

According to the results obtained by analyzing above mentioned data, it can be conclude that Konduruwawa area is more suitable for Airport site. However, only by using Geographic Information System (GIS), it is difficult to propose exact and correct location for Airport site. In that case, a closer social and environment survey has to be done.

In this study, only the site suitability of Konduruwawa area was considered, due to lack of data availability. When considering the available data such as forest, tank, steam, and air defense restricted areas in Polonnaruwa District. Location suitability depend on factors which were considered for analysis and type and purpose of the Airport site. In here basic few factors were considered. There are so many other factors such as Regional plan, Airport use, Ground accessibility, Topography, Obstructions, Visibility, Wind, Noise nuisance, Grading, drainage and soil characteristics, Future development, Availability of utilities from town.

Consideration of all those factors for analysis is big and complicated task. But the analysis with more factors will increase the accuracy of the suitable locations. In order to find the most suitable location for local Airport in Polonnaruwa district, it is need to analyze other suitable location those are showed in Figure 6 closely by considering land uses. After that, the most suitable location for a local Airport in Polonnaruwa district can be identified by comparing those results.

References

Ahesan, M.A. and Marson, T.B. 2015. GIS Application to De termaine Suitable Sites for Automatic Rice Mills in Joypurhat District, Bangladesh. *International Journal of Advanced Remote Sensing and GIS*, 4(1), pp.883-894.

Pareta, K. 2013. Remote Sensing and GIS Based Site Suitability Analysis for Tourism Development. *International Journal of Advance Research in Engineering and Applied Science*, 2(5), pp.43-48.

Data Resource Form, Survey Department of Sri Lanka.

Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2018, Volume 7, Issue 1, pp. 2712-2718
ISSN 2320 – 0243, Crossref: 10.23953/cloud.ijarsg.366

Research Article

Flood Modelling and Analyzing of Attanagalu Oya River Basin Using Geographic Information System

T. D. C. Pushpakumara, T. V. A. Isuru

Department of Civil Engineering, University of Moratuwa, Sri Lanka

Correspondence should be addressed to *thudugalagedon@gmail.com*

Publication Date: 29 June 2018

DOI: https://doi.org/10.23953/cloud.ijarsg.366

Abstract GIS (Geographic Information Systems) are used to store and process terrain data. Using GIS for could be very effective in flood modelling as it becomes very accurate with the use of the Digital Elevation Models. These DEMs could be analyzed with various tools in GIS software and by processing these terrain models, data necessary for the development of flow and flood models can be extracted with a minimum effort in the field. In this research Attanagalu Oya river basin was selected due to its constant exposure to floods and the model for Attanagalu Oya would be processed and the data extracted would be exported in to Hec-Hms and Hec-Ras software to model the flow and create a flood simulation to determine the extent of the flood. By using future rainfall forecasts in the flood simulation developed in this research future floods can be forecasted with great accuracy. As this model could be used to forecast the future floods and their extent it makes this model a warning system for future floods and could be used to mitigate the property loss and the damage incurred.
Keywords *GIS; Geographic Information Systems; Arc map; Hec-Hms; Hec-Ras; DEM; Digital Elevation Model; Attangalu Oya basin; flood modelling*

1. Introduction

Constant exposure to large scale floods has become a major problem in Gampaha district. Many lives households and cultivations and many other more aspects are affected by this problem. With the increase in population and the limited amount of resources, man has done a lot more damage to the nature in search of these resources. Weather man made or not many reasons have led to the increased the frequency of large scale floods in a Gampaha district. The Land Reclamation Commission of Sri Lanka is trying to solve the flood problem in Gampaha district and they have selected the Attanagalu oya basin to do a GIS modelling and analyze the data to find possible solutions for the problem

Total catchment area of the Attanagalu Oya basin is about 736km^2 and it extends till it reaches the sea at Negambo (Wijesekara, 2016). Attanagalu Oya basin was specially selected depending on the availability of data. A Digital Elevation Model (DEM) of the area was used for the studies of the research. This study focuses mainly on the ability use GIS to forecast the future flood events in the Attanagalu Oya basin. The forecast would further include the extent of the flood event corresponding

to a series of rainfall data. Furthermore, this kind of a model could be used as a pre-disaster warning system and in the hand of the relevant authorities it would be very useful in reducing the casualties.

Objectives

- Developing a model to check the past flood events and their extent.
- Forecasting the flood extent of the Attanagalu Oya river basin.

2. Methodology

Development of the model was carried out through a series of steps which included several software (Figure 1). Mainly the following were used:

- Arc map 10.3
- Hec-Hms 4.2.1
- Hec-Ras 5.0.3

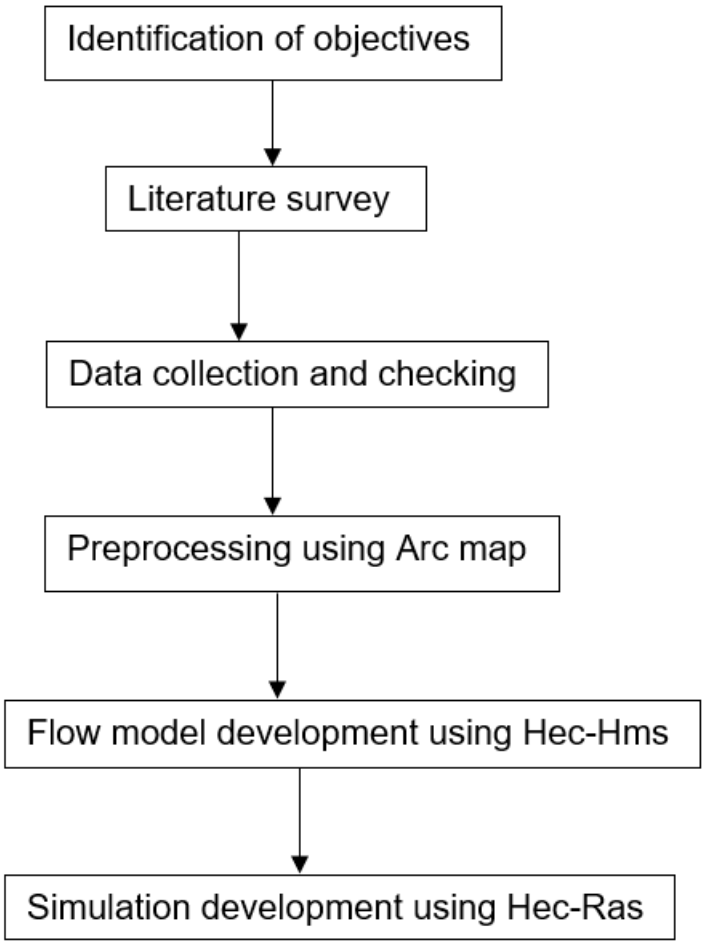


Figure 1: *Methodology of flow chart*

First Arc map was used to extract the necessary geometry data gathered from the survey department of Sri Lanka (Survey department of Sri Lanaka, n.d.) and then these data were exported to Hec-Hms software to create a flow for the Attanagalu Oya. Then this data along with the geometry data from Arc map was imported to Hec-Ras to create the flood simulation for the area. The simulation has the ability to represent the flood depth along the river for any given date with in the simulation time. If correct future rainfall forecasts are available this model would be able to simulate the flood levels in the future.

Preparation of Data

Geometry data of Attanagalu Oya area was needed for the model development. For this aerial photograph of the basin were obtained from drones and these photographs were used to create a DEM (Digital Elevation Model) file of cell size 1mx1m (Figure 2) (Mitchell, 2012). The entire pre-processing part of the Arc map software was based on this model.

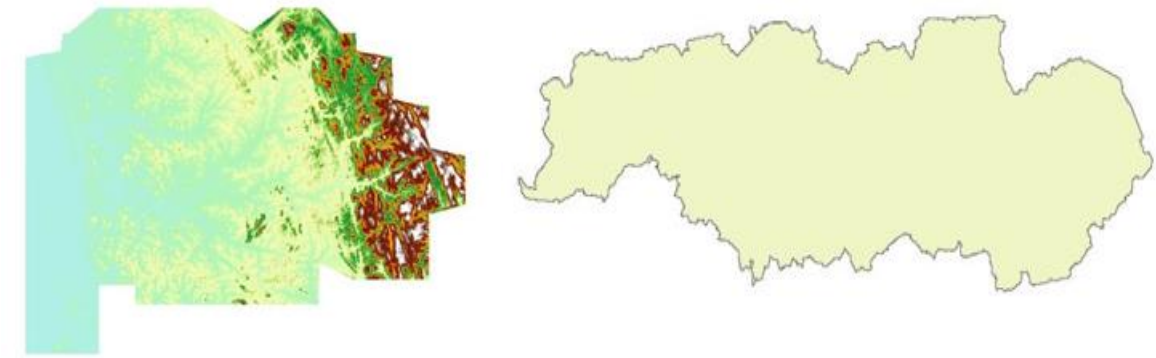


Figure 2: *DEM of the catchment*

Pre-processing using Arc Map

The DEM file of the area was processed in arc map to find the following parameters needed for the next steps of model development in Hec-Hms and Hec-Ras

1. Area of the Catchment (Figure 3).
2. Longest flow path (Figure 4)
3. Slope of the catchment
4. River cross section data (Figure 5)

Arc hydro and spatial reference tools were used to process the terrain data. Data obtained from terrain processing was used in creating the flow model in Hec-Hms.

Figure 3: *Processed catchment*

Table

LongestFlowPathCat

Shape *	OID *	Shape_Length	HydroID	DrainID	LengthDown
Polyline	1	41776.732081	<Null>	<Null>	<Null>
Polyline	2	17338.161892	<Null>	<Null>	<Null>
Polyline	3	26013.481728	<Null>	<Null>	<Null>
Polyline	4	53240.473632	<Null>	<Null>	<Null>

Figure 4: *Attribute table of longest flow path raster*

River cross section data were obtained using the HEC GEO RAS tools (Chndramali, 2012). XS cut lines were drawn on a base map and then the data along the lines were extracted from the DEM. Data along the cut lines were taken as the cross-section data. These cross-section data were used in Hec-Ras flood simulation to model the river.

Figure 5: *XS cut lines on a base map*

Flow Model Development using Hec-Hms

SCS unit hydrograph method was recognized as the ideal transformation method of the Hec-Hms model (Halwathura, 2013).

Following data from Arc map pre-processing was imported to the Hec-Hms flow model.

1. Catchment size - 191.82 km^2
2. Slope - 12%

Along with these data soil types of the Attangalu Oya basin were used to determine the SCS curve number of the basin. “Red Yellow” podzolic was determined as the major soil type appearing in the Attanagalu Oya basin (Wijesekara, 2016) .Based on the characteristics of the basin and the soil type 45 was selected as the curve number for the Attanagalu Oya basin.

Rainfall data for the Attanagalu Oya basin was taken from the department of Irrigation, Sri Lanka (Department of Irrigation Sri Lanka, n.d.). For the model simulation rainfall data during a flood event May 2016 was used (Department of meteorology Sri Lanaka, n.d.).

Hec-Hms produced the outflow for the sub basin relevant to the precipitation data provided (Figure 6). Kirpich formula was used to calculate the time of concentration of the sub basin and from the result the lag time for the sub Basin was derived. Kirpich formula for the time of concentration can be mathematically expressed as:

$$t_c = 0.0078 \left[\frac{L^{0.77}}{S^{0.385}}\right]$$ (Rajapakae, 2013)

Where t_c is the time of concentration, L is the travel length and S is the slope of the catchment.

Kirpich formula for lag time can be expressed mathematically as:

$$t_L = 0.6\, t_c$$ (Rajapakae, 2013)

Where t_L is the lag time and t_c is the time of concentration for the sub basin.

Results obtained from the Hec-Hms was used in the Hec-Ras flood simulation (Figure 7).

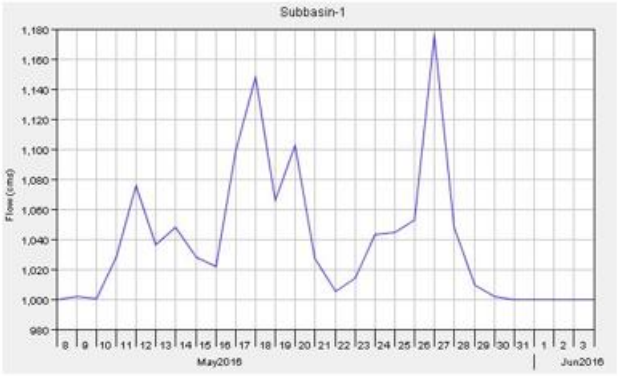


Figure 6: *Out flow of the Hec-Hms model*

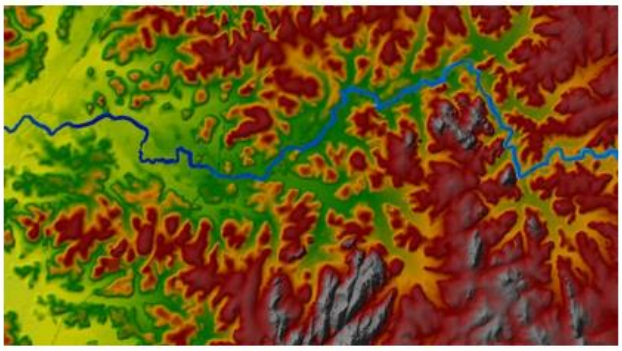


Figure 7: *Hec-ras simulation*

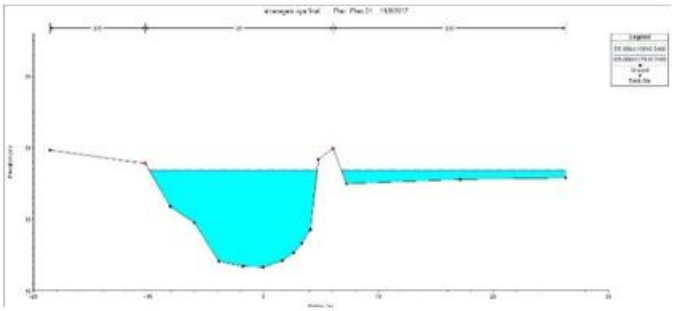


Figure 8: *Cross section in Hec-Ras*

Simulation Development using Hec-Ras

Cross section data from Arc map preprocessing and flow from Hec-Hms flow model was exported to the Hec-Ras model.

From Hec-Ras a simulation based on the rainfall data from May 2016 was created. From this simulation variation of the flood level at any given time could be visualized given that time fall in the range of the simulation run (Figure 8).

3. Results and Conclusion

Models developed in this research were able to simulate the May 2016 flood event accurately. The extent of the flood prone area could be calculated from the Arc GIS software.

A warning system could be modelled based on the results of these models. The model can be used to create a future flood event with a high accuracy and the people living in the area could be warned and evacuated beforehand the flood and the damage could be mitigated.

Accuracy of the model was the main concern of this research. Accuracy depends on mainly on two factors.

- Accuracy of the DEM
- Accuracy of the rainfall data

DEM has to be accurate, without any discrepancies of high level. They could occur due to the photographs used in to create the DEM capturing the tree tops not the land. If any high-level discrepancies were found a field survey has to be done to determine the validity of the relevant points. If the validity of the points is found to be compromised they should be removed manually and this would require field surveying.

With accurate rainfall data, the simulations could be more and more accurate. For the forecasting of future floods rainfall forecasts of high accuracy should be used as the output extent of the flood depends on it.

References

Chndramali, E.A.G. and Wijesekara, N.T.S. 2016. Field Data Collection for Land Information Systems – A Case Study on Quantification of Resource Requirement for a GIS. *Journal of the Institution of Engineers*, 45(4), p.51.

Halwathura, D. and Najim, M.M.M. 2013. Application of the HEC-HMS model for runoff simulation in a tropical catchment. *Environmental Modelling & Software*, 46, pp.155-162.

Wijesekara, N.T.S. and Perera, K.R.J. 2016. Potential on the use of GIS watershed modeling for river basin planning - case study of Attanagalu Oya basin, Sri Lanka. *Journal of the Institution of Engineers*, 45(4), p.13.

Rajapakae, R.L.H.L. and Wijesekara, N.T.S. 2013. Mathematical Modelling of Watershed Wetland Crossings for Flood Mitigation and Groundwater Enhancement - Case of the Attnagalu Oya River Basin. *Journal of the Institution of Engineers*, 36(3), pp.55-67.

Department of Survey, Sri Lanka. Available from: *http://www.survey.gov.lk/.*

Department of Irrigation, Sri Lanka. Available from: *http://www.irrigation.gov.lk.*

Department of Meteorology, Sri Lanaka. Available from: *http://www.meteo.gov.lk/index.php?lang=en.*

Mitchell, A. 2012. Esri Arc user. Available from: *www.esri.com/news/arcuser/1012/gis-modeling-explained-and-explored.html*