

Vol. 13 • No. 1 • April 2019

ISSN: 0976 - 1330

Journal of GEOMATICS



INDIAN SOCIETY OF GEOMATICS

Journal of Geomatics

(A publication of the Indian Society of Geomatics)

Editorial Board

Chief Editor: Dr. A.S. Rajawat

(Address for Correspondence: Group Director, Geosciences, Hydrology, Cryosphere Sciences & Applications Group, Space Applications Centre, ISRO, Ahmedabad - 380 015, India)

Phone: +91-79-26914018 (O), +91-79-29795665 (R), Email: asrajawat@sac.isro.gov.in, editorjogisg@gmail.com

Associate Editor:

R. P. Singh Ahmedabad, Email: rpsingh@sac.isro.gov.in

Assistant Editors:

R. Ratheesh Ahmedabad, Email: ratheeshr@sac.isro.gov.in

S.V.V. Arun Kumar Ahmedabad, Email: arunkumar@sac.isro.gov.in

Ritesh Agrawal Ahmedabad, Email: ritesh_agrawal@sac.isro.gov.in

Members:

A.R. Dasgupta Ahmedabad, Email: arup@ieee.org

P.K. Garg Dehradun, Email: gargpfce@iitr.ernet.in

P.K. Verma Bhopal, Email: drpkverma@rediffmail.com

Ashok Kaushal Pune, Email: akaushal1960@yahoo.co.in

T.T. Medhavy Australia, Email: medhavy.thankappan@ga.gov.in

I.V. Murali Krishna Hyderabad, Email: ivm@ieee.org

S.M. Ramasamy Tiruchirapalli, Email: grucc@ruraluniv.ac.in

P.S. Roy Hyderabad, Email: psroy1952@yahoo.in

Milap Chand Sharma New Delhi, Email: milap@mail.jnu.ac.in

Tara Sharma Canada, Email: sharmatara@yahoo.com

P. Venkatachalam Mumbai, Email: pvenk@csre.iitb.ac.in

Claudio Zucca Morocco Email: c.zucca@cgiar.org

Advisory Board

Paul J. Curran University of London, **UK**

V. Jayaraman Bengaluru, **India**

R. Krishnan Thiruvananthapuram, **India**

P. Nag Varanasi, **India**

M.P. Narayanan New Delhi, **India**

R.R. Navalgund Bengaluru, **India**

Y.S. Rajan ISRO H.Q., Bengaluru, **India**

Josef Strobl University of Salzburg, **Austria**

**Indian Society of Geomatics
Executive Council 2017 – 2020**

President	Tapan Misra , Space Applications Centre, ISRO, Ahmedabad - 380015
Vice-President	Y.V.N. Krishna Murthy , Indian Institute of Space Science and Technology, Thiruvananthapuram – 695547 Raj Kumar , Space Applications Centre, ISRO, Ahmedabad - 380015
Secretary	Shashikant A. Sharma , Space Applications Centre, ISRO, Ahmedabad - 380015
Joint Secretary	K.P.R. Menon , Kerala Remote Sensing and Environment Centre, Thiruvananthapuram - 695102
Treasurer	P. Jayaprasad , Space Applications Centre, ISRO, Ahmedabad - 380015
Members	P.L.N. Raju , North Eastern Space Applications Centre, DOS, Umiam - 793103 K.S. Jayappa , Mangalore University, Mangalore - 575001 A.K. Singh , Indus University, Ahmedabad - 382115 R.J. Bhanderi , Space Applications Centre, Ahmedabad - 380015 K.L.N. Sastry (Retd.) , Space Applications Centre, Ahmedabad – 380015
Ex-Officio (Immediate Past President)	A.S. Kiran Kumar , Indian Space Research Organisation, Bengaluru - 560231

Secretary: (address for correspondence)

6202, SAC Bopal Campus, Ahmedabad – 380058, India
Email: secretary@isgindia.org; sasharma@sac.isro.gov.in

Journal of Geomatics

(A Publication of the Indian Society of Geomatics)

Vol. 13, No. 1

Research articles

April, 2019

1	A Spatial data mining approach applied in urban planning Mohammed Midoun and Hafida Belbachir	1
2	Predicting soil organic carbon concentration using digital soil mapping techniques in eastern Mau forest - Nakuru County, Kenya Ian Ocholla Ariko and Patroba Achola Odera	8
3	Modelling uncertainties in differential global positioning system dataset I. Yakubu and I. Dadzie	16
4	Performance enhancement of standard fuzzy majority voting-based fusion of probabilistic classifiers Mahmoud Salah	24
5	Morphological investigation on degrading minor irrigation tanks: a case study in Hunsur taluk of Mysore district- Karnataka, India Pradeep Raja K.P. and Suresh Ramaswamyreddy	34
6	Online geospatial transaction system using open source libraries Sajeewan G., Danish Shaikh, Mugdha Magare and Ashish Kuvelkar	47
7	Spatial pattern of urban growth using remote sensing and landscape metrics Akintunde John Akinrinola	53
8	Study of discontinuity adaptive MRF models with kernel-based noise classifier Ishuita SenGupta, Anil Kumar and Rakesh Kumar Dwivedi	61
9	Empowering the Electoral Planning and Monitoring System (EPMS) using GeoWeb tools and services – a case study on the Meghalaya legislative assembly election, 2018 Dibyajyoti Chutia, Nilay Nishant, Siddhartha Bhuyan, P. Subhash Singh, Avinash Chouhan, Ritu Anil, Victor Saikhom, Manoj Lokare and P.L.N. Raju	74
10	Landslide hazard evaluation and zonation in Dilbe Town and its surrounding areas, North-western Central Ethiopia – A GIS based grid overlay statistical approach Sendu Demessie and Tarun Kumar Raghuvanshi	79

Special Section: Selected papers of National Symposium on Advancements in Geospatial Technology for Societal Benefits, Dec 03-07, 2018, Ahmedabad (Part-1)

	Preface R.P. Singh, Ratheesh R. and Surisetty V.V. Arun Kumar	93
1	Remote sensing based time-series analysis for monitoring urban sprawl: A case study of Chandigarh capital region Varinder Saini and Reet Kamal Tiwari	94
2	Evaluation of the consistency of DMSP-OLS and SNPP-VIIRS Night-time light datasets Reshma Jeswani, Anurag Kulshrestha, Prasun Kumar Gupta and S.K. Srivastav	98
3	Techniques developed for large area Mars image mosaic using ISRO's Mars Color Camera (MCC) data Indranil Misra, S. Manthira Moorthi and Debajyoti Dhar	106

4	Stress mapping of Sundarban mangroves with Sentinel-2 images using Discriminant Normalised Vegetation Index (DNVI) and Fuzzy classification technique	111
	Sudip Manna and Barun Raychaudhuri	
5	Subsurface utility mapping using multi frequency ground penetrating radar: A case study of road collapse	118
	Tathagata Chakraborty, Dharmendra Kumar Pandey, Jayaprasad P., Deepak Putrevu and Arundhati Misra	
6	Development of universal geospatial data collection application and visualisation platform	129
	Arpit Agarwal	
7	Interannual variation of atmospheric carbon dioxide concentration derived from Orbiting Carbon Observatory-2 data compared with its local and large-scale variations obtained from NASA Giovanni database	134
	Barun Raychaudhuri and Tithi Sen Chaudhuri	
8	Offshore wind farm site suitability and assessment along Tamil Nadu coast using RISAT-1 SAR and synergetic scatterometer data	138
	Surisetty V. V. Arun Kumar, Avinash N. Parde, Jagdish and Raj Kumar	
9	Seasonal variability of rip current probability along a wave-dominated coast using high resolution satellites and wave data	149
	T. Sridevi, Surisetty V. V. Arun Kumar and Raj Kumar	
10	Change detection and trend analysis for Oceansat-2 Ocean Color Monitor (OCM-2) time series data	156
	Ravi Kamal Choudhary, Tushar Shukla, Sampa Roy and Debajyoti Dhar	
11	Spatial enhancement of SWIR band from Resourcesat-2A by preserving spectral details for accurate mapping of water bodies	163
	Indranil Misra, Yatharath Bhateja, Neha Gaur and Vivek Sharma	
12	Extraction of water body, cloud shadow and cloud detection using object-based classification	168
	Sudhanshu Raghubanshi, Ritesh Agrawal, A.K. Sharma and A.S. Rajawat	
13	An approach for generation of multi temporal co-registered optical remote sensing images from Resourcesat-2/2A sensors	174
	Indranil Misra, Vivek Sharma, S. Manthira Moorthi and Debajyoti Dhar	
	Symposium Report	180
	C.P. Singh	
	Indian Society of Geomatics: Awards	v
	Indian Society of Geomatics: Fellows	x
	Instruction for Authors	xi
	Journal of Geomatics: Advertisement Rates	xiii
	Indian Society of Geomatics: ISG Membership Form	xiv
	Indian Society of Geomatics: Membership Fees	xv

A Spatial data mining approach applied in urban planning

Mohammed Midoun^{1*} and Hafida Belbachir²

¹Dept. Informatique Université des Sciences et de la Technologie Mohamed Boudiaf Oran USTO Oran, Algeria

²Dept. Informatique Université des Sciences et de la Technologie Mohamed Boudiaf Oran USTO

*Email: midoun.mohamed@univ-usto.dz

(Received: Apr 09, 2018; in final form: Apr 24, 2019)

Abstract: The nature of spatial information generates a set of problems of incompatibility with the principles of data mining. Spatial data mining is an extension of data mining that considers the interactions in space. It involves various techniques and methods in various areas of research. It takes into account the specificities of spatial information such as spatial relationships that can be topological, metric or directional. These relationships are implicit and difficult to represent. A Bayesian network is a graphical model that encodes causal probabilistic relationships among variables of interest, which has a powerful ability for representing and reasoning and provides an effective way to spatial data mining. Moreover, spatial data cubes allow storage and exploration of spatial data. They support spatial, non-spatial and mixed dimensions. A spatial dimension may contain vector and raster data. The spatial hierarchies can represent topological relationships between spatial objects. In this article we propose to use Bayesian networks for knowledge discovery in spatial data cubes. The goal of our approach is first to consider spatial relationships in the data mining process, and secondly to benefit from the strength of the data warehouses to apply spatial data mining on different aggregation levels according to the topological relations between spatial data.

Keywords: Spatial data mining, Bayesian networks, Spatial warehouse Spatial Analysis

1. Introduction

Spatial data mining is an extension of data mining that considers the interactions in space. It takes into account the specificities of spatial information, such as spatial relationships and spatial dependence. Many studies have been done, where association rules, clustering, classification methods or Bayesian networks are used. A Bayesian network is a graphical model that encodes causal probabilistic relationships among variables of interest, which has a powerful ability for representing and reasoning and provides an effective way for spatial data mining.

On the other hand, spatial data cubes are cubes that contain dimensions or facts which are spatially referenced and can be represented on maps (Bédard et al, 2001) They allow the storage and exploration of spatial data. They support spatial, non-spatial and mixed dimensions. A spatial dimension can contain vector data or raster data. The nature of spatial information generates a set of problems of incompatibility with the principles of data mining. First, the spatial data is linked, while the methods of datamining consider that the data are independent. On the other hand, the spatial relationships are implied and are seldom stored in databases.

The spatial relations are multiple, they may be topological (adjacency, intersection...) or metric (distance) and the analysis can be mono or multi thematic. This makes it difficult to choose the correct spatial relationship.

To represent the spatial relationships in relational databases, we can use the spatial joint index or contiguity matrix. Another approach is to model the spatial information into spatial data cubes. A spatial data cube is an ideal environment for data mining, it allows analysis and spatial queries on several levels of spatial aggregation. Several works on data mining on spatial data cubes were

made. However, few studies have applied Bayesian networks on spatial data cubes. This, due to the complexity of spatial data sets.

Our major contribution is to propose a platform for the application of Bayesian networks on spatial data cubes for data mining purposes. To represent spatial relationships, we use a spatial hierarchy of vector layers that will respect the topological relationships between spatial objects. The spatial aggregation will be used to calculate the measures and then apply data mining on different levels of the spatial hierarchy.

The main interest of our contribution is to use Bayesian networks to apply spatial data mining on different levels of aggregation of spatial hierarchy. This by considering the spatial relationships. We will use spatial analysis to confirm the validation of our approach and view the results on a map.

In the next section we give an overview of some existing works pertaining to spatial data mining and spatial data cubes. Then we define our approach and we propose a framework of spatial data mining based on Bayesian networks. The results and evaluation of our approach will be discussed in experiments section. Finally, we end this paper with some conclusions.

2. Relevant literature

Spatial data mining is the application of data mining techniques to spatial data. It can be defined as the discovery of interesting, implicit and previously unknown knowledge from large spatial data bases (Bédard et al, 2001). The main objective of the spatial data mining is to discover relationship and characteristics that may exist implicitly in spatial databases. It has been used in various fields like remote sensing, medical imagery and visual data mining. Spatial Data Mining extends relational data mining with

respect to special features of spatial data, like mutual influence of neighboring objects by certain factors (topology, distance, direction). Extracting interesting and useful patterns from spatial datasets is more difficult than extracting the corresponding patterns from traditional numeric and categorical data due to the complexity of spatial data types, spatial relationships, and spatial autocorrelation.

Many works have been proposed for spatial data mining, they relate to the various tasks of data mining, such as classification (Ester et al, 1997; Warrender and Augustejn, 1999), association rules (Kamber et al, 1997), or clustering (Han et al, 2001).

Moreover, the application of Bayesian Networks for spatial data mining and knowledge discovery was introduced by (Han et al, 2001). Bayesian networks provide a coherent framework of representation and reasoning for spatial problems. The process of spatial data mining based on Bayesian networks includes two parts, one is structure learning, and the other is learning the parameters of the network. Many studies have focused on the learning of structure (Lam and Bacchus, 1994; Huang et al, 2004), and many others on the study of algorithms and methods of learning parameters (Oniško et al, 2001; Feelders and Van der Gaag, 2006).

As for a spatial data mining method, Bayesian networks can be used for spatial knowledge representation, spatial classification, spatial clustering, and spatial prediction (Huang and Yuan, 2007). Several studies have been conducted: (Porwal et al, 2006) used Bayesian network classifiers for mineral potential mapping (Liebig, et al, 2009), developed an algorithm that can be applied to large trajectory collections (Walker et al, 2005), proposed Spatial Bayesian learning algorithms for geographic information retrieval, and (Li et al, 2012) proposed a bayesian method for assessing vulnerability to natural disasters to catastrophic risk.

Data warehouses are databases of information dedicated to the analysis and decision making (Kimball et al, 1996). A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process (Inmon, 1996). Spatial data warehouse is data warehouse where some dimension members or some facts are spatially referenced and can be represented on a map. Spatial data warehouses contain geographic data, for example, satellite images, and aerial in addition to non-spatial data.

A number of studies have been conducted for spatial data mining in spatial data cubes. They relate in particular to the use of association rules, classification methods, and exploitation of raster databases (Image) (Han et al, 1998; Bédard et al, 2001).

The main difficulty in spatial data mining is the recognition of spatial relationships in databases. These spatial relationships are implicit and difficult to be represented. Several solutions have been proposed to solve this problem. The spatial relationships between objects in a spatial

framework are often modeled by a contiguity matrix. A contiguity matrix can be representing a neighborhood relationship defined using the Euclidean distance or contiguity. Another solution proposed by Valduries (Valduries 1987) is to add a joint index to speed up the joints as part of a relational database. The extension to spatial data has been proposed by Zeitouni et al. in (Zeitouni et al, 2001). This extension consists of adding a third attribute that represents the spatial relationship between two objects. More models of spatial relationships using hypergraphs are available in the literature.

Malinowski and Zimányi, (2005) propose to model topological relationships through spatial hierarchies of spatial data cube. They define the different types of spatial hierarchies. In addition, they classify topological relationships between hierarchical levels according to the procedures required for ensuring correct measure aggregation. A spatial data cube can include numerical measures and spatial measures and pointers to spatial objects at different levels of aggregation. Aggregation of spatial objects is not easy; it requires the use of a spatial hierarchy.

Few studies have applied Bayesian networks on spatial data cubes. Our contribution is to provide a methodology for the application of Bayesian networks on spatial data cubes. With the aggregation of spatial and non-spatial measures, our work allows to take in account the spatial relationships, including topological relationships between different objects, and perform a knowledge discovery in various aggregation levels of a spatial hierarchy.

3. Proposed approach

Figure 1 shows the approach we propose for the application of Bayesian networks on spatially referenced data. We use GIS data that we store in a database. These data are then integrated in a cube of spatial data, where the measurements are aggregated according to the different levels of the spatial hierarchy which corresponds to the topological relations of the spatial objects. Then we apply the Bayesian networks for data mining purposes to predict the progress of constructions in a housing program. The results obtained are then validated by comparing them with the results observed in the field

3.1 Dataset

In a Geographic Information System (GIS), there are vector data (geometric: point, line, polygon) and raster data (pixels). Our approach is based on GIS vector data. We get the data from multiple heterogeneous sources of spatial data (vector) and non-spatial data (attribute).

Then, we apply the pre-processing steps, such as, converting the vector data or adding spatial projection. We perform these pre-processing steps in a GIS environment, and then we build our spatial database.

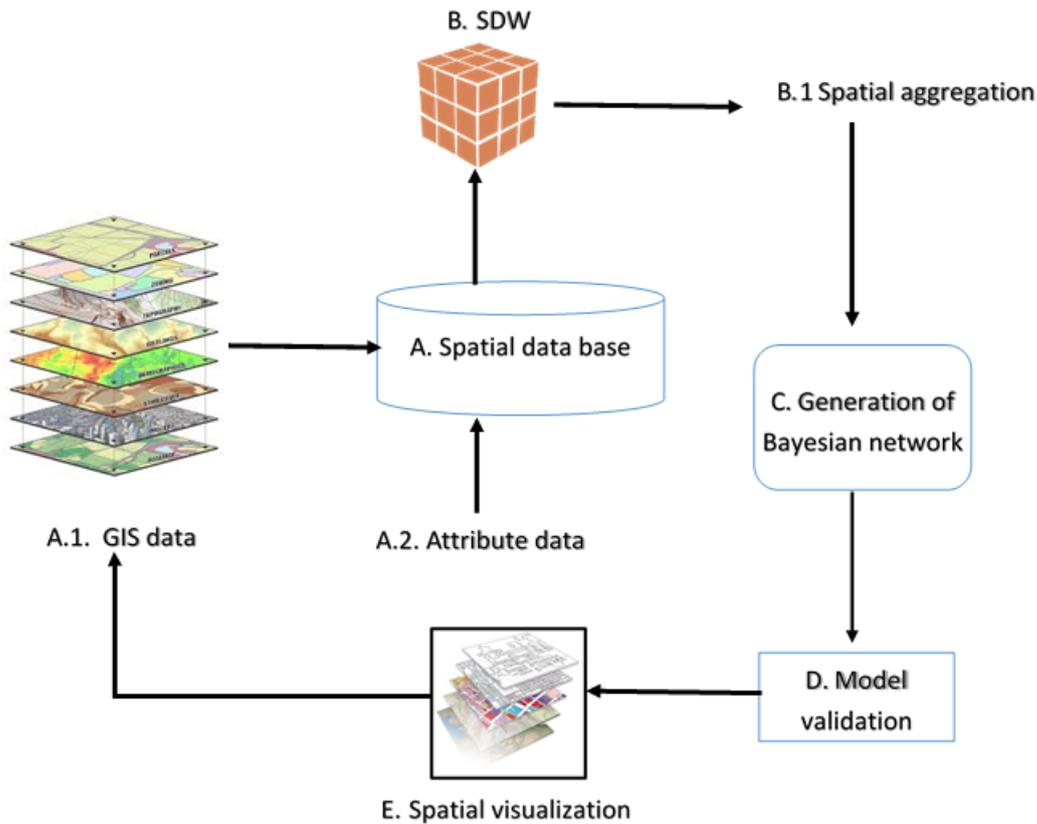


Figure 1: Proposed spatial data mining process

In the spatial data mining environment, we use not only GIS to manage and visualize spatial data, but also as a means of calculating spatial measures using spatial analysis techniques.

3.2 Spatial data cube

In spatial datawarehouse, spatial information can be integrated as dimensions or measures. Spatial data cubes are cubes for which members of dimensions or facts (via spatial measures) are spatially referenced and can be represented on maps (Bédard et al, 2001).

There are two types of spatial data cubes, vector cubes and raster cubes. They contain at least one dimension where some or all members are geometric.

In a data cube, data is organized in dimensions which describe in a natural way most of the attributes associated with the data of interest (Inmon 1992). The dimensions are in turn organized into hierarchies, with data aggregated at each level. As for the dimension hierarchies, topological relationships have hierarchical structures. these relations correspond to the hierarchical semantic relationships between spatial objects.

Therefore, our approach is based on the use of these topological relationships to add levels to the spatial hierarchy of our spatial data cube. The measures of the fact table will be aggregated and calculated according to each level of aggregation of the spatial dimension. They will represent the variables on which we apply Bayesian networks

3.3 Bayesian networks

Bayesian networks are graphical models for defining probabilistic relationships between variables. An advantage of Bayesian networks is that they capture knowledge in a form people can understand intuitively, and which allows a clear visualization of the relationships involved.

Bayesian networks use a directed acyclic graph (DAG) to represent assertions of conditional relationships. The nodes in the graph represent the variables and the directed arcs define the conditional relationships.

The advantages of directed graphic models over undirected models are the notion of causality. Causality indicates that if an arc is directed from A to B in the network, then A causes B. Bayes' theorem is used to calculate causal inference about the variables. Bayes' theorem states:

$$P(\mathbf{B} | \mathbf{A}) = \frac{P(\mathbf{B} | \mathbf{A}) \cdot P(\mathbf{A})}{P(\mathbf{B})}$$

The construction of Bayesian networks is a difficult task, and the number of possible structures and parameters can be huge in this kind of structure.

Learning a Bayesian network from data involves two tasks: Estimating the probabilities for the conditional probability tables (learning parameters) and deriving the structure of the network.

The process of building the Bayesian network consists of three steps: variables definition, structure learning, and parameter estimation.

- Variable definition. Defines the relevant variables and the relationship between them.
- Structure learning. Determine the directions of all edges based on prior knowledge and the given data set. Structure learning of Bayesian networks is the key step to perform reasoning and predicting.
- Parameter estimation. It refers to define the conditional probabilities of the relationships. This step defines the conditional probabilities associated with each node.

As for a spatial data mining method, Bayesian networks can be used for spatial knowledge representation, spatial classification, spatial clustering, and spatial prediction.

Bayesian networks involves different search algorithms for constructing the network topology. The heuristic algorithms include K2, DAG, Hill Climbing (HC), and TAN (Tree Augmented Naive (Bayes)).

In our approach, we use Bayesian networks on measures of spatial data cube. The measures are calculated using aggregated levels of spatial hierarchy. Then, they will be discretized, and several Bayesian networks can be built from these measures.

3.4 Evaluation and validation

Once the Bayesian networks are built, each network must be evaluated. For this purpose, we compare the results obtained with real results observed in the field. The accuracy of the evaluation and the calculation of the Kappa index will allow us to evaluate the results obtained by our approach based on Bayesian networks.

3.5 Spatial analysis

Spatial analysis is a set of methods and tools which enable to understand, evaluate and interpret the spatial distribution of phenomena in order to discover and / or highlight the general rules of organization of space (Pumain and Saint-Julien, 1997). Spatial analysis can be applied to the interrogation of thematic, geometric and topological components of the spatial information contained in the GIS.

Once the Bayesian model is generated and validated, we integrate the parametric data of the Bayesian model into thematic GIS layers. This will not only allow spatial analysis to visualize the results on a map and compare them with the results obtained in the field, but will also validate the analysis of the generated Bayesian network.

4. Experiments

This section describes the experiment conducted to evaluate the proposed approach. We apply Bayesian networks for Urban Planning in order to predict the progress of housing construction programs in Algeria. We use real data stored in a spatial data cube.

4.1 Experimentation environment

We applied our approaches under Windows environment, with SQL Server 2012 as database management system and ArcGIS 10.3 desktop as GIS.

4.2 Dataset

For the purposes of our experiments, we used a database comprised of vector GIS data prepared in ArcGIS environment, and non-spatial data from progress reports on housing construction programs in Algeria. Spatial data concerns the graphic representation of zones, islets, parcels and buildings. The non-spatial data relate to the progress rates of the constructions collected in the field as well as the information on each of the space objects.

These data were processed and integrated in spatial data cube modeled in snowflake schema as shown in figure 2. This cube has four dimensions: Report, Date, Phase, and Buildings, and one spatial hierarchy with four levels: Building / Parcels/ Islets / Area . This hierarchy represents the topological relationship between spatial objects classes.

There are 21 100 measures in the fact table of spatial data cube. They provide information on the progress of construction of buildings.

The measures are aggregated and calculated according to the spatial hierarchy, so as to apply Bayesian networks on different levels of aggregation of spatial data cube.

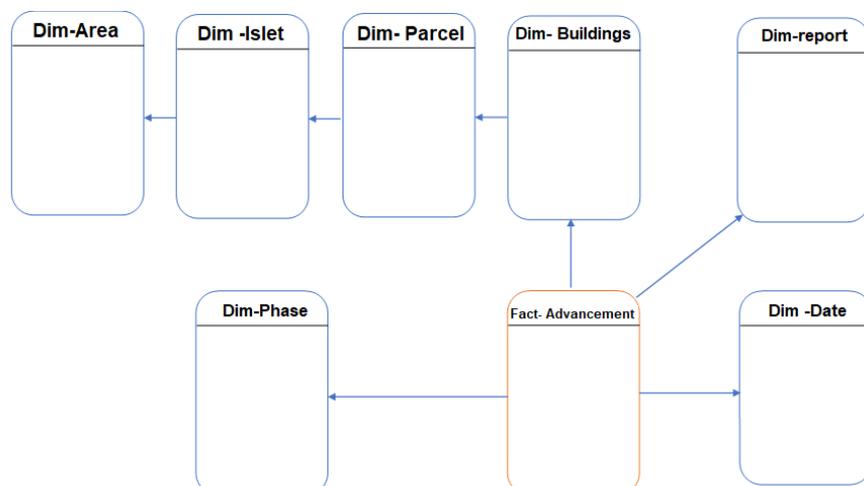


Figure 2: Spatial data warehouse

4.3 Bayesian networks

In our implementation we use the K2 algorithm to calculate the Bayesian network structure. K2 algorithm is the most famous score-based algorithm in Bayesian network. It recovers the underlying distribution in the form of DAG efficiently. We begin by defining variables. The variables are described as following: Earthwork, Boundary marking, Concrete dosage, Verification of verticality, Verification of stability, Floor Coating, Partitioning, Coating, Window installation, Waterproofing, Painting.

The values of its variables are represented in the form of measures in the spatial data cube. We perform the discretization and aggregation of these measures according to the levels of the spatial hierarchy.

After defining the domain variables and data preparation, we can obtain the structure of the Bayesian network and then we should compute the conditional probabilities of the relationship.

Figure 3 shows the structure of the Bayesian network applied for monitoring the building construction process. In this structure, we have 11 measurements, that are spread over six phases of construction (Infrastructure, Superstructure, Masonry, Coating, Water roofing and Painting). The progress of the construction phases is conditioned by the rate which represents each measure. The final rate of progress is conditioned by the rate of progress of the construction phases.

We apply Bayesian networks for each level of aggregation of spatial hierarchy (Buildings, parcels, islets). The Parametric results of the Bayesian network distribute the buildings, parcels and islets in six classes we have defined to represent the different stages of progress of the construction process.

4.4 Evaluation and validation

We selected 255 cases for testing the validity of the model. Table I shows a confusion matrix. It shows the results of the experiment. We compare the results obtained with real results observed in the field. The accuracy of the evaluation

is 84.7% and the Kappa index is 0.848. The experimental results validate the proposed approach for spatial data mining. The parametric results obtained by the Bayesian networks represent the estimated rates of construction progress for the buildings, then aggregated for the parcels and islets according to the spatial hierarchy of the data cube. These estimated results are compared to the observed field results that represent the measures stored in the data cube.

Table 1: Confusion matrix generated by Bayesian network

Evaluated results	Observed results						
	I	II	III	IV	V	VI	Total
I	17	0	2	1	2	1	23
II	2	20	3	1	0	1	27
III	0	1	36	3	0	2	42
IV	1	0	4	57	3	1	66
V	1	0	2	1	30	3	37
VI	0	2	0	1	1	56	60
Total	21	23	47	64	36	64	255
Evaluation accuracy /%	80,9	86,9	76,5	89,0	83,3	87,5	

4.5 Spatial analysis

After validating the model, we integrate the parametric data of the Bayesian model in the GIS thematic layers: Buildings, parcels and islets. This allows us to perform spatial analysis, visualize the results on a map and compare them with the results observed in the field. The spatial analysis carried out were applied to the analysis of the progress of construction of buildings, plots and islets.

Figures 4 and 5 show the spatial distribution of buildings and islets. The construction process has six phases. In these figures, we visually compare the predictive results obtained with those observed in the field by the agents. This is done at several levels of aggregations that correspond to the topological relationships of spatial objects. Spatial analysis have enabled us to spatially visualize results and confirm the validation of our approach.

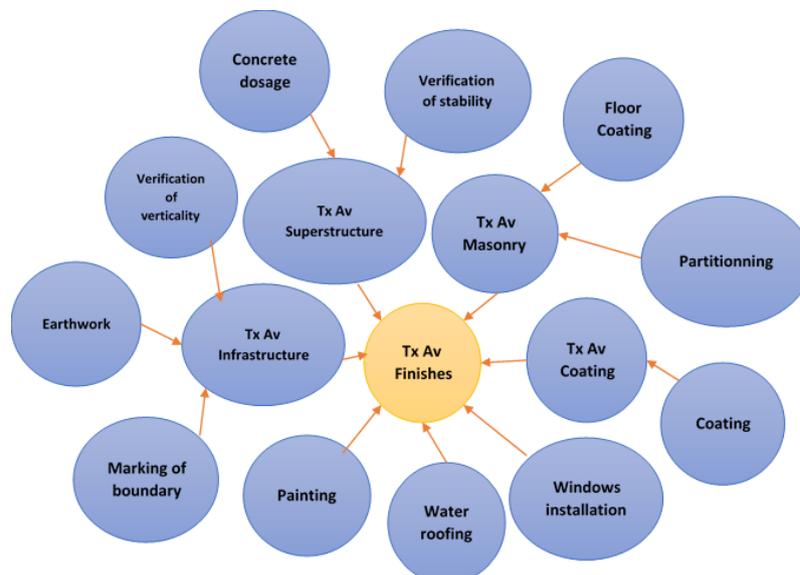


Figure 3: Bayesian network applied for urban planning

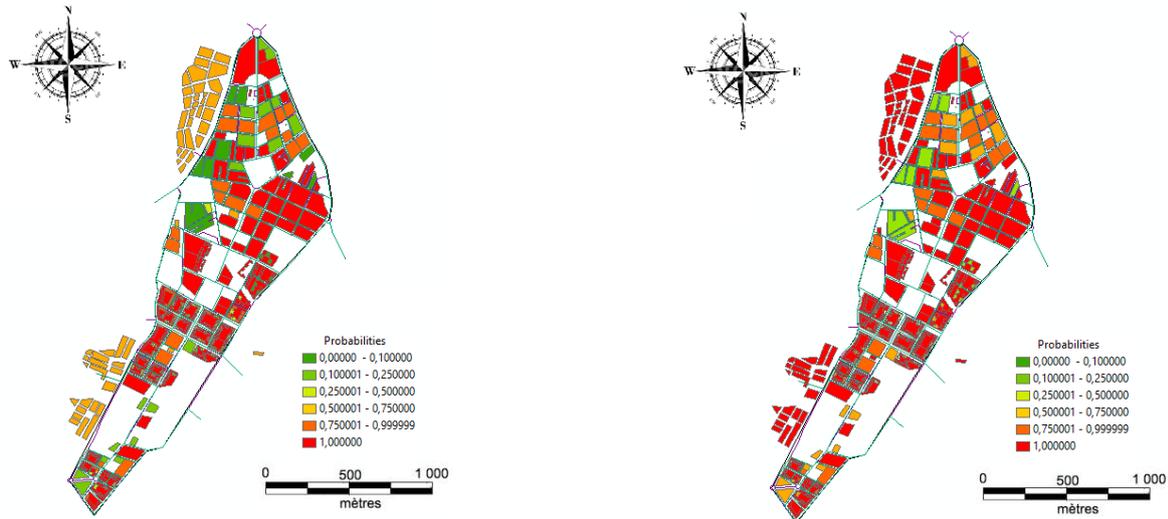


Figure 4: Spatial analysis of the distribution of buildings into classes according to the rate of progress of construction and comparison of results with observed results

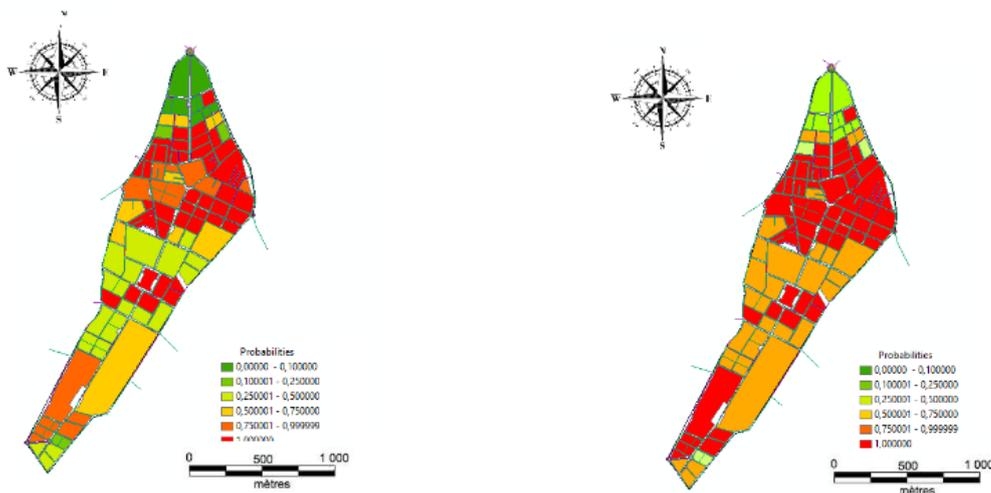


Figure 5: Spatial analysis of the distribution of islets into classes according to the rate of progress of construction and comparison of results with observed results

4.6 Discussion

As shown in Table I, the estimation accuracy was 84.7% and the Kappa index was 0.84 which is considered as a good result for prediction. On the other hand, using spatial analysis, the comparison of the results obtained with the results observed in the field validates our approach by a visual analysis. Data mining gives detailed results whereas spatial analysis gives a general description of the results. Spatial analysis is used to confirm visually the results obtained by data mining, but cannot be enough to give reliable results on its own.

We can conclude that the experimental results thus validate the feasibility of the proposed approach for knowledge discovery in spatial data. Moreover, the application of data mining on a spatial data cube allows a knowledge discovery about the different levels of aggregation of spatial hierarchy. Our approach allows not only to predict the construction progress of each building, but also the overall assessment of the construction process on the different islets and parcels of the study area. Another advantage of

our method is to use GIS to visualize, validate and locate the results on a map. We can therefore say that our approach is a good way for spatial data mining in spatial data cubes.

5. Conclusion

Spatial data mining is an extension of data mining that takes into account the spatial relationships. Spatial relationships are difficult to be represented in databases. Few studies have used Bayesian networks for knowledge discovery in spatial data cubes. In this article, we first explain the concepts related to data mining and spatial data cubes. Then, we propose a framework for data mining in spatial data cubes using Bayesian networks. Furthermore, we showed a case study and used the experimental data to validate the applicability of Bayesian networks for spatial data mining. Consequently, we consider our approach as a good way to explore the spatial data.

The first interest of our approach is that it takes into consideration the spatial relationships including topological relationships. In addition, it allows knowledge discovery

about the different levels of aggregation of spatial hierarchy. Another advantage of our method is to use spatial analysis and GIS to evaluate, visualize and locate the results on a map.

In conclusion, our study presents multiple perspectives, such as the development of a decision support tool that combines spatial analysis and Bayesian networks, or the development of new algorithms for Bayesian networks taking into account the spatial relationships in the process of knowledge discovery.

References

Bédard, Y., T. Merrett and J. Han (2001). Fundamentals of spatial data warehousing for geographic knowledge discovery. *Geographic data mining and knowledge discovery*, 2, pp. 53-73.

Ester, M., H.-P. Kriegel and J. Sander (1997). Spatial data mining: A database approach. *International Symposium on Spatial Databases*, Springer.

Feelders, A. and L. C. Van der Gaag (2006). Learning Bayesian network parameters under order constraints. *International Journal of Approximate Reasoning*, 42(1), 37-53.

Han, J., M. Kamber and A. Tung (2001). Spatial clustering methods in data mining: A survey. H. Miller and J. Hun, eds. *Geographic Data Mining and Knowledge Discovery*, Taylor & Francis, London, UK.

Han, J., N. Stefanovic and K. Koperski (1998). Selective materialization: An efficient method for spatial data cube construction. *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer.

Huang, J., H. Pan and Y. Wan (2004). An algorithm for cooperative learning of Bayesian network structure from data. *International Conference on Computer Supported Cooperative Work in Design*, Springer.

Huang, J. and Y. Yuan (2007). Construction and application of Bayesian network model for spatial data mining. *2007 IEEE International Conference on Control and Automation*, IEEE.

Inmon, W. (1992). *Building the data warehouse*, QED Technical Pub. Group.

Inmon, W. (1996). *The data warehouse environment*. In *The Building the Data Warehouse*, Wiley and Sons, New York.

Kamber, M., J. Han and J. Chiang (1997). Metarule-Guided Mining of Multi-Dimensional Association Rules Using Data Cubes. *KDD*.

Kimball, R., M. Ross, W. Thornthwaite and B. Becker (1996). *The Data Warehouse Toolkit*, John Wiley&Sons. Inc., New York.

Lam, W. and F. Bacchus (1994). Learning Bayesian belief networks: An approach based on the MDL principle. *Computational intelligence*, 10(3), 269-293.

Li, L., J. Wang, H. Leung and S. Zhao (2012). A Bayesian method to mine spatial data sets to evaluate the vulnerability of human beings to catastrophic risk. *Risk Anal*, 32(6), 1072-1092.

Liebig, T., C. Körner and M. May (2009). Fast Visual Trajectory Analysis Using Spatial Bayesian Networks. *ICDM Workshops*.

Malinowski, E. and E. Zimányi (2005). Spatial hierarchies and topological relationships in the spatial MultiDimER model. *British National Conference on Databases*, Springer.

Oniśko, A., M. J. Druzdzel and H. Wasyluk (2001). Learning Bayesian network parameters from small data sets: Application of Noisy-OR gates. *International Journal of Approximate Reasoning*, 27(2), 165-182.

Porwal, A., E. J. M. Carranza and M. Hale (2006). Bayesian network classifiers for mineral potential mapping. *Computers & Geosciences*, 32(1), 1-16.

Pumain, D. and T. Saint-Julien (1997). *L'Analyse spatiale. 1. Localisations dans l'espace*. Paris: Armand Colin, coll. Cursus.

Valduriez, P. (1987). "Join indices." *ACM Transactions on Database Systems (TODS)*, 12(2), 218-246.

Walker, A. R., B. Pham and M. Moody (2005). Spatial bayesian learning algorithms for geographic information retrieval. *Proceedings of the 13th annual ACM international workshop on Geographic information systems*, ACM.

Warrender, C. E. and M. F. Augusteijn (1999). Fusion of image classifications using Bayesian techniques with Markov random fields. *International Journal of Remote Sensing*, 20(10), 1987-2002.

Zeitouni, K. and L. S. Géographiques (2000). Fouille de données spatiales. *Revue internationale de géomatique*, (4/99).

Zeitouni, K., L. Yeh and M.-A. Aufaure (2001). Join indices as a tool for spatial data mining. *Temporal, Spatial, and Spatio-Temporal Data Mining*, Springer, pp.105-116.

Predicting soil organic carbon concentration using digital soil mapping techniques in eastern Mau forest - Nakuru County, Kenya

Ian Ocholla Ariko¹ and Patroba Achola Odera^{2*}

¹Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

²University of Cape Town, Private Bag X3, Rondebosch, South Africa

*Email: patroba.odera@uct.ac.za

(Received: Mar 13, 2018; in final form: May 01, 2019)

Abstract: Soil organic carbon (SOC) concentration is one of the most important indicators of soil fertility and soil quality. Accurate information about the spatial variation of SOC concentration is critical to sustainable soil utilization and management. Eastern Mau forest reserve is one of the major water towers in Eastern Africa, providing essential ecosystem services and it is critical in storage of organic carbon. The objective of this study is to predict the spatial distribution of SOC concentration using digital soil mapping (DSM) techniques. Two prediction techniques; multi-linear regression and multi-linear regression kriging have been used. Seventy-five percent (75%) of soil sample data have been used for model(s) calibration while 25% for model(s) validation. Various variables including terrain attributes, climate data, land use and soil properties data are used in the prediction. Results indicate that multi-linear regression model has a lower R² of 37.4% compared to multi-linear regression kriging, with R² of 42.3%. The results further show that the western parts of eastern Mau (largely forest land) has the highest concentration of SOC, while lowest SOC concentration is observed on the eastern section (largely crop land). The study demonstrates that multi-linear regression kriging performs better than multi-linear regression in capturing the spatial distribution of SOC across the study area.

Keywords: soil organic carbon, digital soil mapping, multi-linear regression, multi-linear regression kriging, Eastern Mau forest

1. Introduction

Forest reserves are most affected by human population for settlement and agricultural activities. The SOC pool, estimated at 1,550 Pg to 1-meter depth is about twice the atmospheric pool or 2.8 times the biotic pool (Batjes, 1996). Soil organic matter contains 58% carbon on average (Chan, 2008) and is essential in regulating climate, water supplies and biodiversity which are vital to human well-being (Kumar et al., 2016).

Countries which are signatories to the United Nations Framework Convention on Climate Change (UNFCCC) under the Kyoto protocol are expected to monitor changes in the soil organic carbon (SOC storage) within their countries (Razakamanarivo et al., 2011). Consequently, better and accurate techniques for estimation of SOC concentration are necessary to inform policy making on actions that enhance organic carbon storage and suggest measures to counter areas with depreciating SOC storage. The recent development of geospatial technologies has allowed for a spatial quantitative prediction approach involving modeling of continuous soil properties (based on factors of soil formations) besides the assessment of accuracy and uncertainties of the predictions. This approach referred as digital soil mapping (Mora-Vallejo et al., 2008) is much better compared to the conventional mapping techniques which only generate qualitative maps whose accuracy cannot be assessed. The conventional method(s) use polygons which do not consider spatial variability within an area because a whole polygon is normally given a constant value.

Soil organic matter (SOM) makes up just 2-10% of soils mass but has a critical role in physical, chemical and biological function of agricultural soils. SOM is formed by

decay of organic material that enters the soil system. Soil organic carbon (SOC) is the major constituent of SOM. SOC is normally expressed as a percentage carbon by weight, that is, g C per 100 g of soil (Chan, 2008).

In Kenya, deforestation has led to a decline in most of the forest reserves, leaving only five major water towers, which are closely monitored for ecosystem sustainability. This paper identifies possible significant covariates to be used in the prediction model(s) and generates predicted SOC concentration map using multi-linear regression and multi-linear regression kriging techniques in eastern Mau forest.

The study area (eastern Mau forest reserve) is in Nakuru County, Kenya. It lies between latitudes 0° 15' S and 0° 40' S and longitudes 35° 40' E and 36° 10' E as shown in figure 1. It has an area of approximately 650 km², with an elevation ranging from 2,210 to 3,070 m. The climate is cool and humid with an average of 93.5 mm of annual precipitation and a mean annual temperature ranging from 9.8 to 17.5° C. The Njoro and Naishi rivers drain from the eastern slopes into Lake Nakuru, while River Nessuiet flows northwards into Lake Bogoria. The major land uses are forest, agriculture and grassland. Despite the rampant deforestation and degradation experienced since mid-1990s, because of illegal logging, charcoal burning and encroachment of approximately 61,023 ha for human settlement (UNEP, 2009), eastern Mau forest remains the largest Afromontane forest in Eastern Africa.

2. Theoretical background of digital soil mapping

Digital soil mapping (DSM) technique is defined by (Lagacherie and McBratney, 2007) as “the creation and

population of spatial information systems by numerical models inferring the spatial and temporal variation of soil types and soils properties from soil observation and knowledge from related environmental variables". DSM is like conventional soil mapping except that the functional relationship between the soils attributes and model factors are formulated using statistical models rather than conceptual models (Thompson et al., 2012). These statistical models are fitted using geo-referenced soil data. We discuss some models in the following sections.

2.1 State factors or CLORPT model

For the latter half of 20th century, scientific rationale for soil mapping has been the state factors or CLORPT model, which stands for (CL=climate, O=organisms, R=relief, P=parent material, and T=time). The logic of CLORPT model was based on the equation of Jenny (Jenny, 1992) and formulated from the recognition of the factors of soil formation. The state factor equation can be expressed as,

$$S = f(cl, o, r, p, t) \quad (1)$$

where, S represents the soil, considered to be a function of (cl) climate, organism (o) or vegetation, relief (r), parent material (p) acting through time (t).

The Jenny equation illustrates that by correlating soil attributes with observable difference in one or more of the state factors, a function (f) or model can be developed that explains the relationship between the two, which can be used to predict soil attributes at new locations when the state factors are known. However, the state factors do not constitute factors that institute pedogenic processes.

2.2 SCORPAN model

In the last decade, McBratney et al., 2003 generalized and formulated a new equation with the objective of modeling the variables responsible for the processes of soil formation, through an empiric quantitative description of the relationships among other spatially geo-referenced factors which are used as spatial prediction functions. It is an improvement of Jenny equation, the scorpan model is expressed as,

$$S = f(s, c, o, r, p, a, n) \quad (2)$$

where, (s) is the soil attributes, (a) represent the age or time factor and (n) the space or spatial position. The other symbols have their usual meanings as given in equation 1.

This model differs from the clorpt as it is intended for quantitative spatial prediction rather than explanation, this distinction justifies the inclusion of soil and space as factors because soil attributes can be predicted from other soil attributes and spatial information (Thompson et al., 2012).

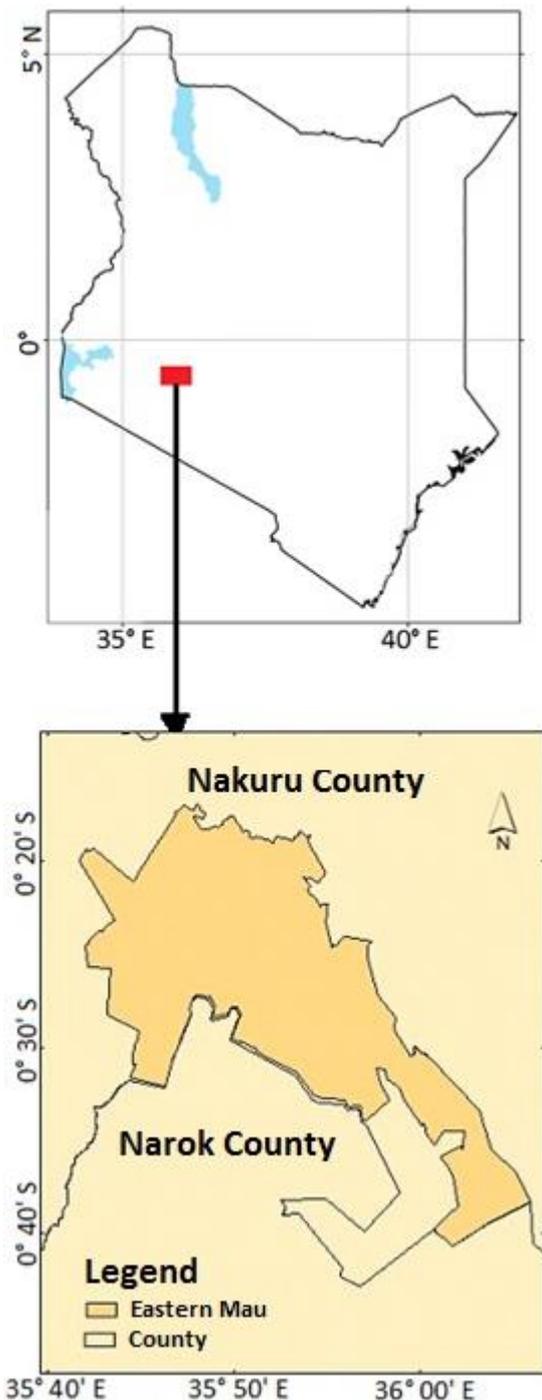


Figure 1: Eastern Mau forest reserve.

2.3 Geostatistics

Geostatistics offers a way of describing the spatial continuity of natural phenomena and provides adaptations of classical regression techniques (Hengl, 2009). Geostatistics is used to predict values of a sample variable over the whole area of interest as it is used in combination with various geospatial Information Systems (GIS) layers. Geostatistics differs from the conventional statistics in that in the later the samples taken from a statistical community are independent from each other and the presence of one sample does not show any information about the next sample. While in geostatistics the spatial structure or correlation among variables in a region are investigated (Abadi et al., 2015), with the use of semi-variograms to quantify spatial autocorrelation.

We discuss only one type of interpolation technique here, Kriging. Kriging is one of the geostatistical methods, based on the theory of regionalized variables and variogram model. It is considered as the best linear unbiased predictor (*BLUP*) that satisfies a certain optimality criterion. It is named after a South African mine engineer D.G. Krige who used the technique in the mining industry in the early 1950's as a means of improving ore reserve estimation (Krige, 1951). Kriging is a suitable method in the presence of spatial dependence as it is beneficial to model the deterministic component of soil spatial variation as a function of the environmental covariates and the stochastic component (Thompson et al., 2012).

We now describe ordinary kriging and regression kriging. Ordinary kriging is often regarded more appropriate interpolation technique, as it adapts to local fluctuation of the mean by limiting the domain of stationarity of the mean to the local neighborhood (Mulder et al., 2011). Ordinary kriging is used to improve the prediction by interpolating the environmental variables. On the other hand, Regression kriging is a hybrid method that combines either a simple or multi-linear regression (MLR) model with ordinary kriging of the regression residuals.

Multi linear regression is commonly used in up-scaling approach to model the linear relationship between independent variable and secondary variables (predictors). However, MLR generates a process which is stationary and assumes the residuals are identical and independently distributed. Errors associated with MLR are large since the approach does not consider the varying relationship between the environmental variable and the SOC across space (Ge et al., 2007).

Regression kriging (RK) involves the use of environmental variables as it consists of three components, it can be expressed as (Zhang et al., 2012),

$$Z_{(S)} = Z_{(S)}^* + E'_{(S)} + E'' \quad (3)$$

where, $Z_{(S)}$ is the RK prediction formed by summing the regression prediction from the covariates, $Z_{(S)}^*$ is the deterministic component (ordinary kriging of the residuals), $E'_{(S)}$ is the stochastic component and E'' is the pure noise.

In this study, multi linear regression and multi-linear regression kriging techniques have been used to predict SOC concentration based on other similar studies (Mora-Vallejo et al., 2008; Sumfleth and Duttmann, 2008). Peng et al., 2013 indicate that the use of multi linear regression model is simple and direct.

3. Materials and methods

3.1 Soil data

Two hundred and twenty (220) soil samples of soil organic carbon, clay, sand and silt (up to 30 cm depth) were

provided by the Kenya soil survey. The soil samples were georeferenced within the study area, the soil samples were then randomly divided into two, that is, 75% calibration set, and 25% validation set. Figure 2 shows the distribution of SOC data points in the study area.

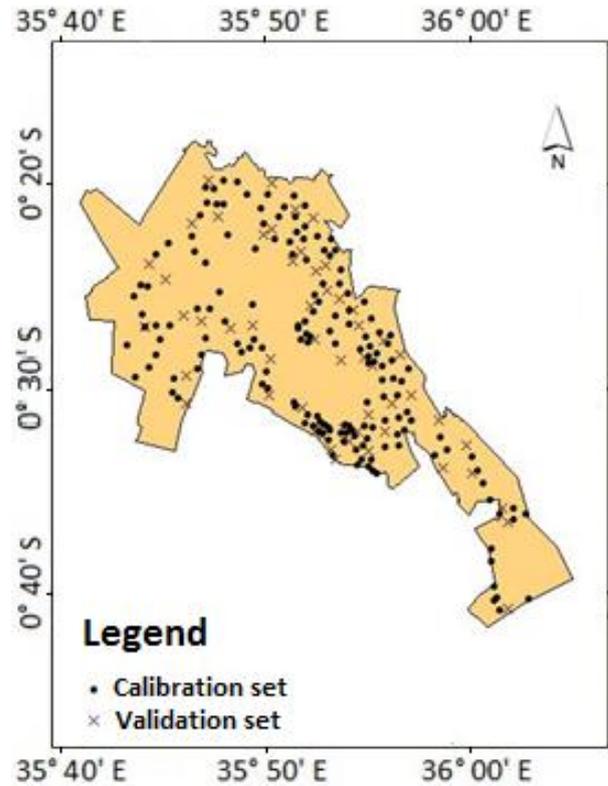


Figure 2: Distribution of SOC data points.

3.2 Auxiliary data

Various sources of data retrieved and analyzed to capture the spatial variation of the soil forming factors in the study area included climate, normalized difference vegetation index (NDVI) and digital elevation model (DEM). The mean annual temperature and precipitation on a 30 arc-second raster data were obtained from www.worldclim.org. The datasets were resampled to 30 m resolution (admittedly with some errors). Multi-spectral remote sensed satellite imagery data (Landsat 5) was obtained from <http://earthexplorer.usgs.gov/>. Landsat 5 was selected because soil sample were collected in 2011. Subsurface reflectance band values were used in calculating NDVI values using the following expression,

$$NDVI = \frac{NIR - R}{NIR + R} \quad (4)$$

where, R and NIR are the Red and Near Infrared bands respectively (Rouse et al., 1993). The land cover of the study area was also generated from the relevant bands. A 30 m Shuttle Radar Topographic Mission (SRTM) data was obtained from www.jpl.nasa.gov, it was used to derive different topographic attributes as aligned with relief as part of soil forming factors. The parameterized derivatives included elevation, slope, aspect, plan curvature, profile curvature and flow accumulation. Topographical Wetness Index (TWI) which is a secondary topographic derivative was also used. The equation used for TWI is given as (Wilson and Gallant, 2000),

$$TWI = \frac{\ln(\text{flow accumulation})}{\tan(\text{slope} \div 57.29577951)} \quad (5)$$

3.3 Selection of explanatory variables for predictive model

The SOC calibration dataset was used to extract points from the co-variables through map overlay. The extracted co-variable-values were then linearly regressed against the response variable (SOC). Regression analysis is essential in characterizing the relationship between the predictors and response variable, also it is used in estimating their correlation. Stepwise regression analysis using Akaike Information Criterion (AIC) was applied on the initial model to obtain a reduced model, which accomplishes a desired level of prediction with as few predictor variables as possible (Akaike, 1973). The best model has the smallest AIC value. The coefficients of the reduced model were fitted into a multi-linear regression equation to map the spatial distribution of SOC concentration, using the following expression (Kumar and Lal, 2011),

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_kX_k + e \quad (6)$$

where, *Y* is the SOC concentration, *B*₀ is the *Y* intercept (a constant term) *B*_{*k*} are model coefficients, *X*_{*k*} are independent variables, and *e* is an error of disturbance.

3.4 Spatial structure of the model

The residuals obtained from the reduced regressed model were assessed and fitted into a semi variogram. The nugget to sill ratio was then used to characterize the importance of the random component and provide quantitative measures of spatial dependence. The residuals were interpolated using kriging to incorporate the spatial correlation of the errors of the multi-linear regression model.

The final prediction map was obtained by spatial overlay of the multi-linear regression model surface with the kriged interpolated residual surface in a regression kriging approach (equation 3). Validation test using mean error (ME), root mean square error (RMSE) and coefficient of determination (*R*²) were performed to both models to evaluate the prediction accuracy.

4. Results and discussion

4.1 Modelling results

Results of the multi linear regression analysis of various covariates indicated that the covariables could account for 56% of SOC concentration variability within the study area. The results from the stepwise regression analysis are shown in table 1.

Model number four (4) had the lowest AIC values (Table 1). The coefficients of the model parameters were determined and the reduced multi linear regression model obtained as,

$$SOC = -7.2832 + (-0.0895468 \times Silt) + (-0.771998 \times band4) + (68.5149 \times band1) +$$

$$(9.2718 \times NDVI) + (0.0019903 \times \text{flow accumulation}) + (0.003212 \times \text{elevation}) \quad (7)$$

where SOC and NDVI have the same meanings as earlier defined.

Model parameters such as silt concentration, NDVI, flow accumulation and elevation are given in figures 3, 4, 5 and 6 respectively. We also present rainfall, sand soil concentration and major land use/cover in figures 7, 8 and 9 respectively.

Table 1: Model parameters with their AIC values.

Model	Covariables	AIC values
1	TWI, Slope, Curvature, plan curvature, rainfall, sand, NDVI, band 1,4,5, Elevation, Silt, flow accumulation	39.34
2	Plan curvature, rain, flow accumulation, NDVI, Silt, band 1,4,5, Elevation	36.50
3	Rain, Elevation, band 1,4,5, NDVI, Silt, Sand, flow accumulation	35.15
4	Rain, Sand, Band 1,4, flow accumulation, NDVI, Silt, Elevation	34.88

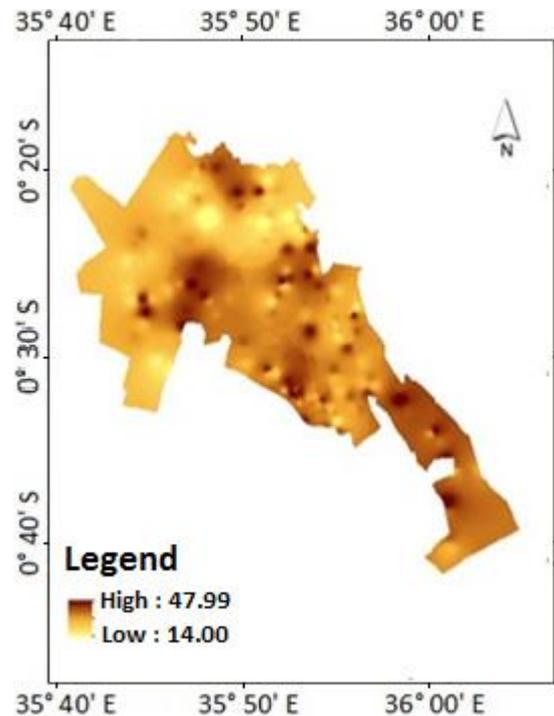


Figure 3: Silt concentration distribution (units in %).

Using equation 7, the reduced multi linear regressed SOC surface was generated as shown in figure 10. The residuals were then assessed if they fulfilled the requirements for spatial dependence i.e. they should not be biased and should have a uniform distribution. This assessment was done through fitting the residuals into a histogram and normal quarter quantile plot, the residuals certified the requirements. The residuals were then fitted into an

experimental semi-variogram with a spherical function, to measure the spatial dependence using the nugget effect (N) to sill (S) ratio.

the data. It also meant that the distance between sample points were far apart lowering the spatial dependence of the residuals.

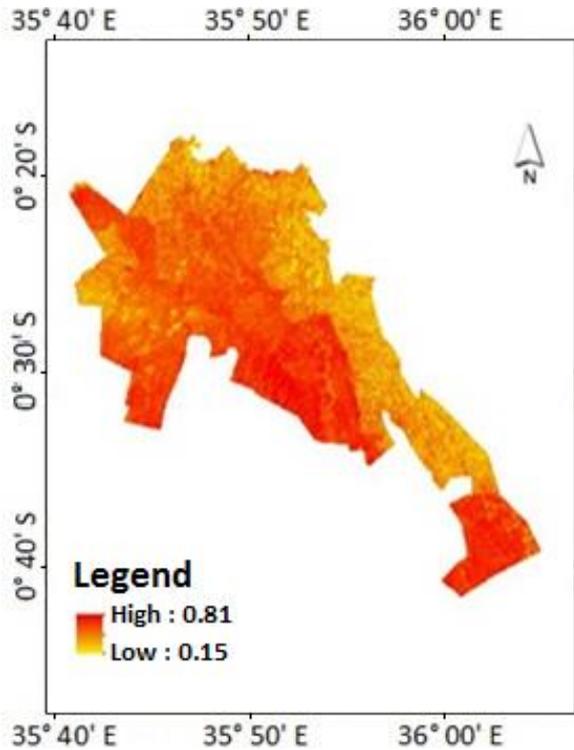


Figure 4: Normalized difference vegetation index (NDVI).

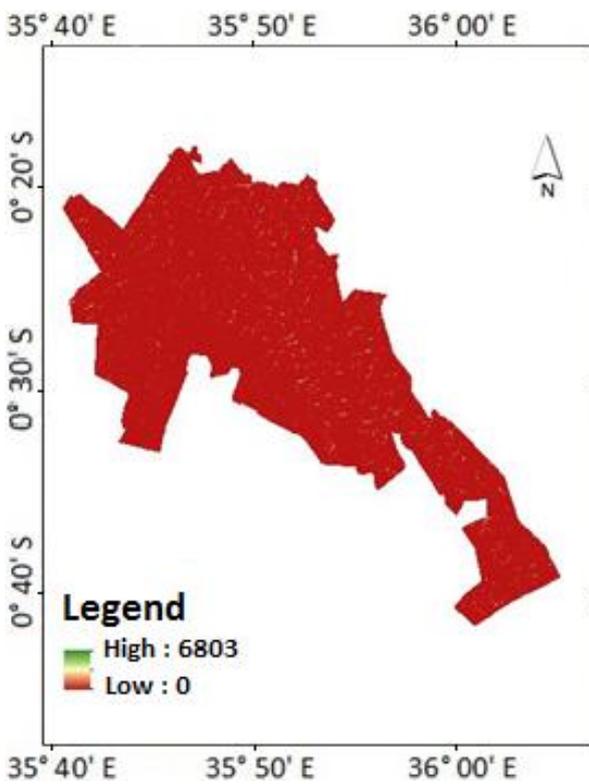


Figure 5: Flow accumulation.

The N:S ratio was 56.34% which indicated a moderate spatial dependence of the residuals. The high nugget value as shown in table 2 indicated the measurement error within

Table 2: Semi-variogram parameters of the residuals.

	Nugget	Sill	Nugget/sill
SOC residuals	0.54	0.9657	0.56

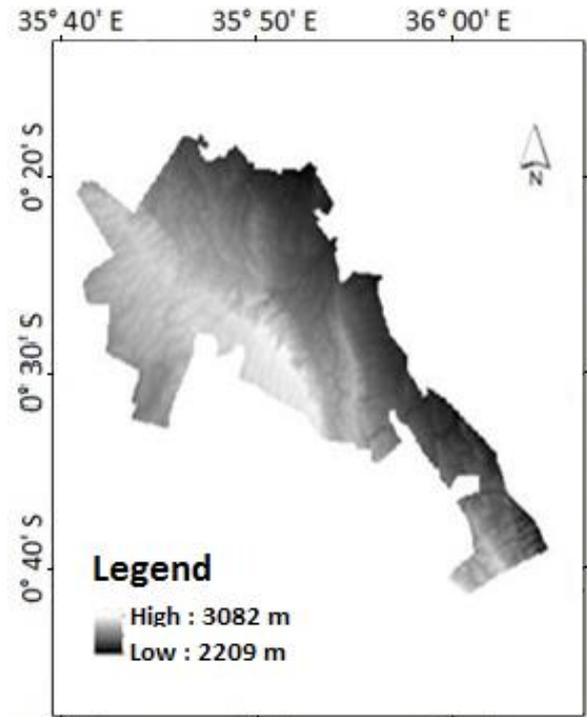


Figure 6: Elevation (units are in m).

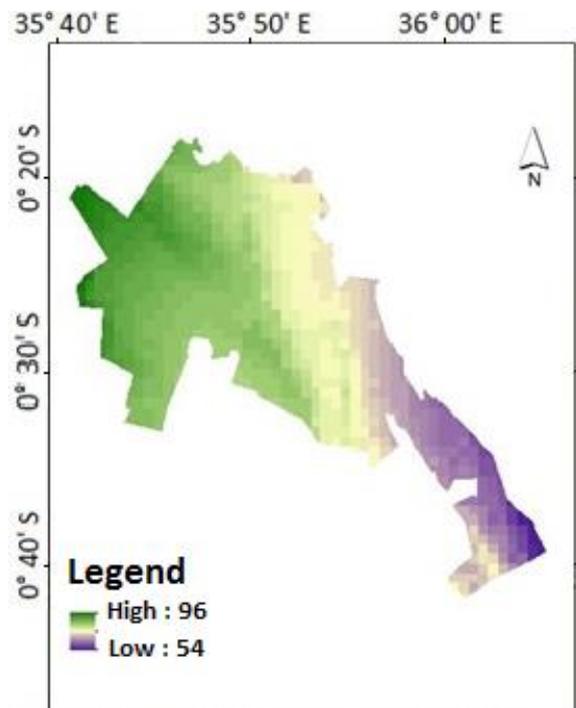


Figure 7: Mean annual rainfall (units are in mm).

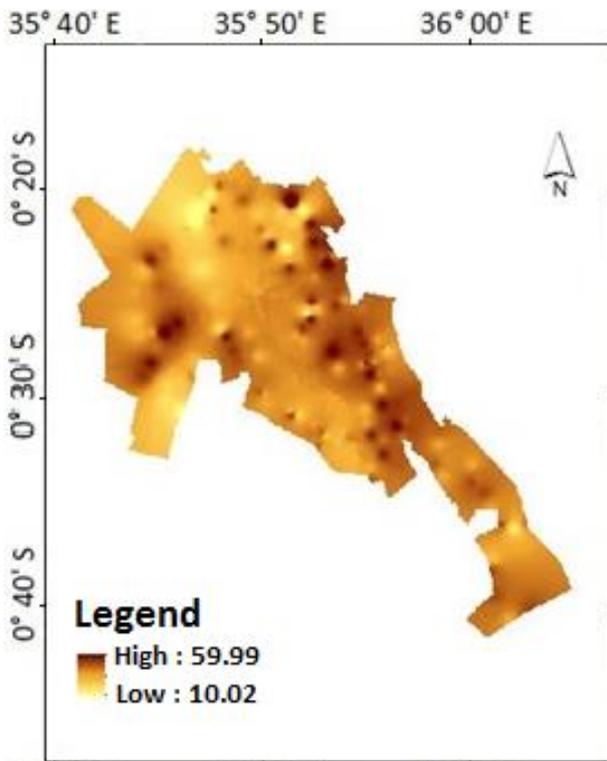


Figure 8: Sand soil concentration (units are in %).

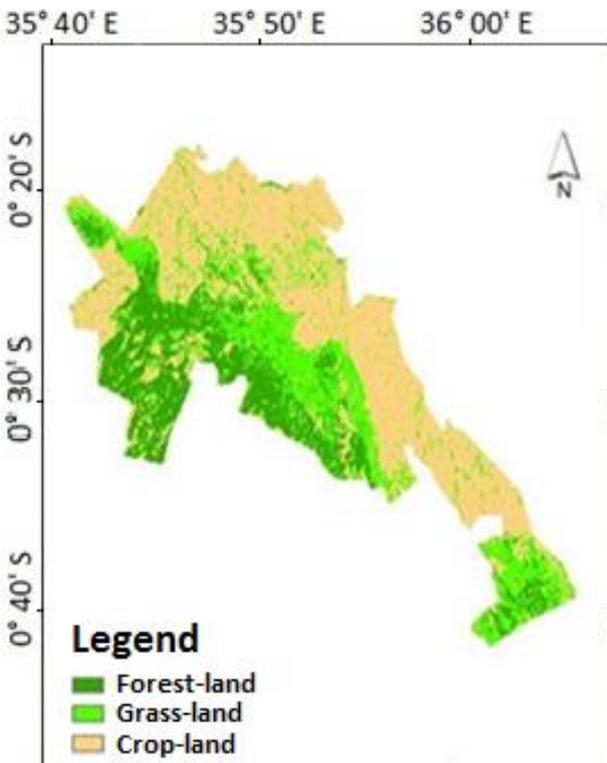


Figure 9: Major land use/cover.

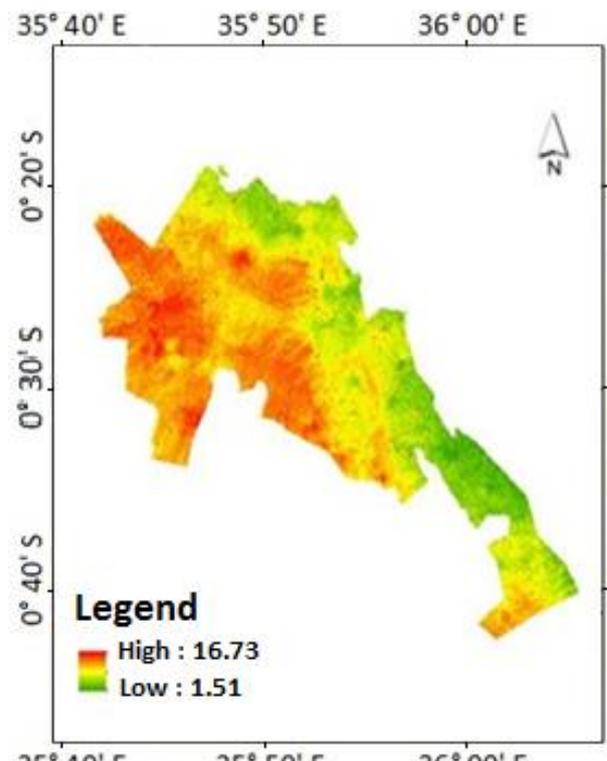


Figure 10: Reduced regressed surface of SOC concentration (units are in %).

The residuals were then interpolated using ordinary kriging as shown in figure 11. The reduced multi-linear regressed model surface and the kriged surface maps were spatially overlaid to obtain the final regression kriged surface of SOC concentration (Figure 12).

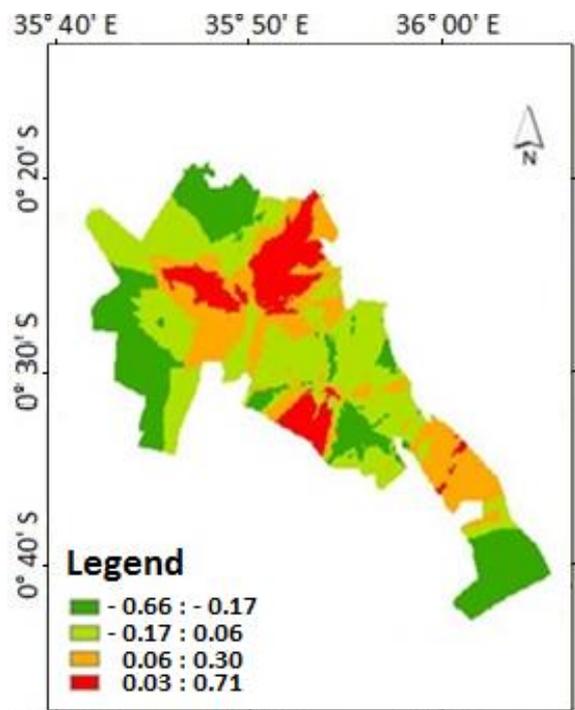


Figure 11: Interpolated residual surface.

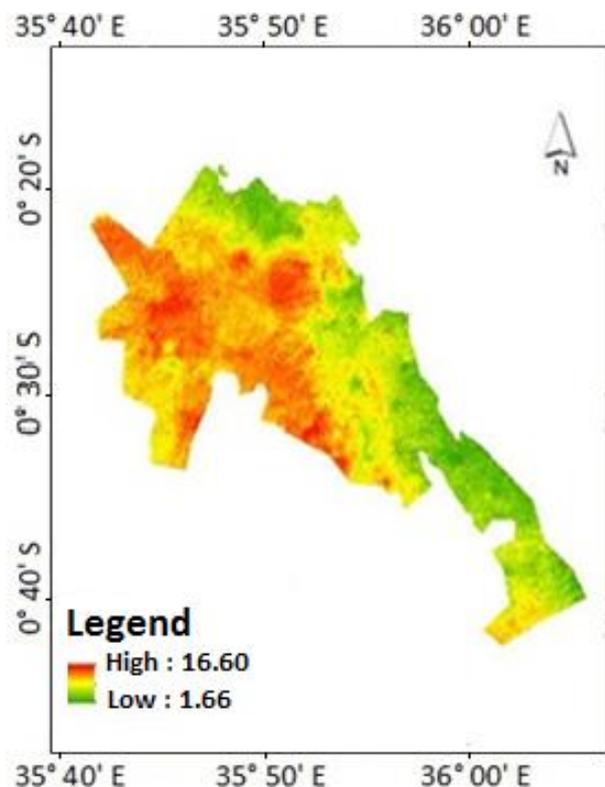


Figure 12: Regression kriged surface of SOC concentration (units are in %).

4.2 Model assessment

Fifty-five independent SOC data points were used for validation of the two models. The results of the validation are given in table 3. The R^2 of the regression and regression kriging models are 37.38 and 42.25%, respectively. The better performance by regression kriging (RK) is because it considers uncertainty due to regression and measurement in the form of kriged variables compared to the regression model.

The low R^2 (42.25%) of the overall RK model can be attributed to the moderate spatial structure of the residuals and errors due to the interpolation of datasets. In table 3, RK model had a lower ME (0.2794) compared to the regression model (0.3304). The RK model also had a better RMSE value (1.3970) compared to the regression surface RMSE value (1.4590). From the validation, the regression kriging model proved to be a better prediction technique in the study area.

Table 3: Statistics of models' assessment.

	Regression	Regression kriging
Minimum difference	0.0392	0.0690
Maximum difference	5.5494	4.9121
Mean difference	0.3304	0.2794
RMSE	1.4590	1.3970
R^2	0.3738	0.4225

4.3 Discussion

The spatial distribution of regression-kriging showed that SOC concentration decreased from the western section to the south-eastern section in the forest reserve, as shown in

figure 12. The spatial distribution of SOC concentration in the study area could be predicted using terrain attributes, soil, texture, climate and land use.

Band 4 and band 1 (Landsat data) were significant variables in estimation of SOC concentration, band 4 captures the near infrared reflectance which distinguishes vegetation varieties and conditions while band 1 provides information capable of differentiating soil and rock surfaces from vegetation. From the two bands, areas which have vegetation cover tend to have higher concentration of SOC compared to bare soil and rock surfaces.

A comparison of the final SOC prediction map (Figure 12) with the land use/cover map (Figure 9), reveals that SOC concentration is relatively higher in the forest areas and lower in the cropland areas. These results show that land use/cover type has a significant impact on the spatial SOC concentration patterns. This can be attributed to the fact that forest land fix plentiful SOC because of the flourishing soil plant roots and thicker forest litter layers, which are easily absorbed and beneficial to SOC. In cropland, there is continuous tillage exposing the nutrient for decomposition leading to decrease in the SOC concentration.

High elevation regions also recorded high SOC concentration, this is because the high-altitude areas had low temperatures which lower nutrient decomposition and these areas have low human interference, compared to areas of lower elevation which have increased anthropogenic activities which expose soil organic carbon for decomposition.

5. Conclusions

This study compared two prediction techniques (multi-linear regression and multi-linear regression kriging) in predicting soil organic carbon concentration in the eastern Mau forest reserve. The two models were calibrated and verified using independent validation datasets. Results show that multi-linear regression model has a lower R^2 of 37.4% compared to multi-linear regression kriging (42.3%). The western part of eastern Mau, largely forest land, has the highest concentration of SOC, while lowest SOC concentration is observed on the eastern section, largely crop land. Given the characteristics of the study area, the number of observations used and data distribution, the prediction of SOC by multi-linear regression kriging is satisfactory. These results can be improved by accurate selection and representation of soil formation factors and related spatial residuals.

Acknowledgements

We would like to thank the Kenya Soil Survey Department for providing valuable datasets. We are grateful to Dr. Kennedy Were for his support in terms of data and technical expertise in carrying out this study. Finally, we appreciate the valuable comments by two anonymous reviewers, that have been used to improve the quality of the manuscript.

References

- Abadi, K.M., R. Jafarinia and J. Varvani (2015). Analyzing kriging and cokriging methods by using ArcGIS software in preparing SP map of Farahan plain soil. *Biological Forum-An International Journal*, 7(1), 836–846.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Petrov, B.N. and F. Csaki (eds.), *2nd International Symposium on Information Theory*, Akademiai Kiado, Budapest, pp. 267–281.
- Batjes, N.H. (1996). Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*, 47, 151–163.
- Chan, Y. (2008). Increasing soil organic carbon of agricultural land. In: Prime facts, NSW Department of Primary Industries, available online at: https://www.dpi.nsw.gov.au/data/assets/pdf_file/0003/210756/Increasing-soil-organic-carbon.pdf.
- Ge, Y., J.A. Thompson, C.L. Morgan and S.W. Searcy (2007). VNIR diffuse reflectance spectroscopy for agricultural soil property determination based on regression-kriging. *Transactions of the American Society of Agricultural and Biological Engineers*, 50, 1081–1092.
- Hengl, T. (2009). A practical guide to geostatistical mapping. Office for Official Publications of the European Communities, Luxembourg.
- Jenny, H. (1992). *Factors of soil formation: A system of quantitative pedology*. Dover, New York.
- Krige, D.G. (1951). A Statistical approach to some basic mine valuation problems on Witwatersrand. *Journal of chemical, Metallurgical and Mining Society of South Africa*, 52(6), 119–139.
- Kumar, P., P.C. Pandey, B.K. Singh, S. Katiyar, V.P. Mandal, M Rani, V. Tomar and S. Patairiya (2016). Estimation of accumulated soil organic carbon stock in tropical forest using geospatial strategy. *Egyptian Journal of Remote Sensing and space Sciences*, 19(1), 109–123.
- Kumar, S. and R. Lal (2011). Mapping the organic carbon stocks of surface soils using local spatial interpolator. *Journal of environmental monitoring*, 13, 3128–3135.
- Lagacherie, P and A.B. McBratney (2007). Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In Lagacherie, P., A.B. McBratney and M. Voltz, (eds.), *Digital soil mapping: An Introductory Perspective*, *Developments in Soil Science*, 31, pp. 389–399.
- McBratney, A.B., M.L. Mendonca-Santos and B. Minasny (2003). On digital soil mapping. *Geoderma*, 117 (1-2), 3–52.
- Mora-Vallejo, A., L. Claessens, J. Stoorvogel and G.B.M. Heuvelink (2008). Small scale digital soil mapping in southeastern Kenya. *Catena*, 76(1), 44–53.
- Mulder, V.L., S. De Bruin, M.E Schaeppman and T.R. Mayr (2011). The use of remote sensing in soil and terrain mapping-A Review. *Geoderma*, 162, 1–19.
- Peng, G., W. Bing, G. Guangpo and Z. Guangcan (2013). Spatial distribution of soil organic carbon and total nitrogen based on GIS and geostatistics in a small watershed in a hilly area of northern China. *PLoS ONE*, 8(12), 1–9.
- Razakamanarivo, R.H., C. Grinand, M.A. Razafindrakoto, M. Bernoux and A. Albrecht (2011). Mapping organic carbon stocks in eucalyptus plantations of the central highlands of Madagascar: A multiple regression approach. *Geoderma*, 162, 335–346.
- Rouse, J.W., R.H. Haas, J.A. Schell and D.W. Deering (1973). Monitoring vegetation systems in the great plains with ERTS. In proceedings of the third Earth Resources Technology Satellite symposium, Goddard Space Flight Center, National Aeronautics and space Administration, SP-351, Science and Technical Information Office, NASA, Washington, D.C., pp. 309–317.
- Sumfleth, K. and R. Duttman (2008). Prediction of soil property distribution in paddy soil landscape using terrain data and satellite information as indicators. *Ecological Indicators*, 8, 485–501.
- Thompson, J.A., S. Roecker, S. Grunwald and P.R. Owens (2012). *Digital Soil Mapping: Interactions with and applications for hydrogeology*. Academic Press, Boston, pp. 665–709.
- UNEP. (2009). *Kenya: Atlas of our changing environment*. Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP).
- Wilson, J.P. and J.C. Gallant (2000). Secondary topographic attributes. In Wilson, J.P. and J.C. Gallant (eds.), *Terrain Analysis: Principles and Applications*, John Wiley and Sons, Inc., New York, pp. 87–132.
- Zhang, S., Y. Huang, C. Shen, H. Ye and Y. Du (2012). Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. *Geoderma*, 171, 35–43.

Modelling uncertainties in differential global positioning system dataset

I. Yakubu^{1*} and I. Dadzie²

¹Department of Geomatic Engineering, University of Mines and Technology, Tarkwa, Ghana

²Department of Geomatic Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

*Email: yissaka@umat.edu.gh

(Received: Jan 18, 2019; in final form: May 07, 2019)

Abstract: The quest for precision and accuracy in Differential Global Positioning System (DGPS) data requires modelling uncertainties in the data collected and eliminating outliers which hinder the precision and accuracy. In this paper, dual frequency DGPS receivers' dataset collected repeatedly over a control station at thirty (30) minutes interval were mathematically modelled for uncertainties using five soft computing and classical methods namely: Backpropagation Artificial Neural Networks (BPANN), Generalised Regression Neural Networks (GRNN), Multivariate Adaptive Regression Splines (MARS), Radial Basis Functions Neural Network (RBFNN), and Total Least Square. The results revealed that all the models produced were satisfactory. The Mean Horizontal Error (MHE), Root Mean Square Error (RMSE), and Standard Deviation (SD) performance criteria indices were applied. GRNN outperformed BPANN, MARS, RBFNN, and TLS in modelling DGPS data uncertainties. In terms of their mean horizontal displacement and standard deviation, GRNN achieved 4.5314E-11 m and 1.3200E-13 m compared to TLS, BPANN, MARS and RBFNN which achieved: 7.3901E-06 and 8.7500E-14; 2.8311E-06 and 2.2300E-08; 0.0088 m and 3.3158E-05 m; and 1.2016E-04 m and 1.2195E-06 m respectively. It can be concluded that all the models used can be applied in detecting and eliminating uncertainties in DGPS data. There is therefore, the need to apply these methods in modelling uncertainties in DGPS applications in sensitive areas such as deformation monitoring of high rise buildings, bridges and dam embarkment.

Keywords: Data Uncertainties, Soft Computing, Regression Techniques, DGPS

1. Introduction

In recent years, the use of Global Positioning System (GPS) for positioning has gained more grounds than the traditional surveying techniques which required the use of a theodolite or total station. Schuessler and Axhausen (2009), conducted a research and concluded that Differential Global Positioning Systems (DGPS) has provided surveyors with more accurate results than the traditional techniques. The use of the DGPS help pinpoints satellite locations of interest. Most of the GPS data has to be processed before becoming meaningful to the user. Hence, scientists and engineers find the term post-processed GPS data appropriate to describe GPS data which is not processed directly on the field, rather obtained after some processing in the office.

Nonetheless, the use of the GPS receivers come with some restrictive conditions including hazy weather conditions, shade, an insufficient number of required satellites for positioning. These conditions cause a delay in obtaining position information of the observation point thereby causing long observational periods and introduce errors into the observed data.

These errors, classified as either parametric or non-parametric, need to be removed or minimized with the use of statistical tools such as the box and whisker plot (box plot) and the modified z-score which could be used to identify outliers in a given dataset (Ben-Gal, 2005). Outliers are blunders which are committed on the field and are non-adjustable. They have to be eliminated from datasets since they are either the very small or big values identified during data analysis (Carlson and Goodman, 2014). On the other hand, systematic errors which are usually fixed and have a mean of zero (0) follow a pattern, and hence are predictable. Random errors are considered

the remaining errors after all other errors have been removed (Filzmoser, 2004).

Least square regression models invented by Gauss (1823) have been applied for solving majority of problems in geoscientific field, notably adjustment of DGPS survey networks (Yakubu *et al.*, 2018; Peprah and Mensah, 2017; Ansah, 2016; Annan *et al.*, 2016; Okwuashi and Eyoh, 2012), determination of GPS coordinate transformation parameters (Ziggah *et al.*, 2013), and datum transformation parameters (Ziggah *et al.*, 2016) with the purpose of identifying and eliminating the uncertainties in the datasets. However, (Acar *et al.*, 2006) noted that some outliers in some datasets still remain after adjustments mainly due to inefficiencies of the mathematical models to eliminate the outliers. This study, therefore, seeks to investigate more sophisticated and advanced models for data pruning, denoising and eliminating outliers or uncertainties in DGPS datasets.

The invention of soft computing techniques has revolutionized data pruning and adjustment mainly because they have the capabilities to denoise datasets (Kutoglu, 2006) and give better accuracy in points estimation (Akyilmaz *et al.*, 2009) and have been applied to solve numerous scientific problems including improving classification (Sharma *et al.*, 2015; Kumar *et al.*, 2014), adjusting DGPS networks (Yakubu *et al.*, 2018), modeling stream networks (Achour *et al.*, 2012), earthquake modelling (Pinho *et al.*, 2008), landslide modelling (Zahra, 2010), coordinate transformation (Ziggah *et al.*, 2016), GPS height conversion (Fu and Liu, 2014; Liu *et al.*, 2011), geodetic deformation modelling (Bao *et al.*, 2011; Du *et al.*, 2014; Maxime *et al.*, 2005), earth orientation parameter determination (Liao *et al.*, 2012), estimating energy demand (Alreja, *et al.*, 2015), slope stability analysis (Samui, 2013; Lall *et al.*, 1996).

In this study, soft computing techniques such as Backpropagation Artificial Neural Networks (BPANN), Generalised Regression Neural Networks (GRNN), Radial Basis Functions Artificial Neural Networks (RBFNN), and Multivariate Adaptive Regression Splines (MARS) were used in modelling the uncertainties in DGPS datasets. The performances of these soft computing techniques were compared with classical techniques such as Total Least Squares (TLS) which has the capability to adjust errors in both the observation and design matrices (Acar *et al.*, 2006), and recommendations made.

2. Resources and methods used

Tarkwa is the study area (Figure 1) which located in the Western Region of the Republic of Ghana (Asklunel and Eldvall, 2005) with geographic coordinates between longitude 2° 10' 00" W - 1° 45' 00" W and latitude 4° 30' 00" N - 5° 25' 00" N with an average topographic height of about 78 m above Mean Sea Level (MSL). Geographically, the topography is generally undulating with steep slopes parallel to each other and to the strike of the rocks in the north-south direction (Kortatsi, 2004). The type of coordinate system used in the study area is Ghana projected grid derived from the Transverse Mercator with 1° W Central Meridian and the WGS84 (UTM Zone 30N)

(Yakubu *et al.*, 2018; Peprah *et al.*, 2017). Tarkwa is a mining town, with three major mining companies, namely: Goldfields Ghana Limited, Tarkwa Mine; AngloGold Ashanti, Iduaprim Mine and Ghana Manganese Company Limited, Nsuta.

DGPS survey was carried out on a known control in Tarkwa, the study area. The DGPS field observations were conducted at 30-minute intervals over the same control to obtain one hundred and four (104) redundant coordinates data for the control. These observations were made to check the consistency of the output data for the same position. The base station was a Continuous Operating Reference Station (CORS) located at the University of Mines and Technology, Tarkwa.

One of the contributing factors affecting the estimation of accuracy is related to the quality of datasets used (Devi and Karthikeyan, 2015; Dreiseit and Ohno-Machado, 2002; Ismail *et al.*, 2012), several precautions including time of observation, avoiding overhead cables, multipath errors (Yakubu and Kumi-Boateng, 2011) were considered to ensure the reliability of the observed datasets. Table 1 shows sample DGPS dataset used in modelling the uncertainties

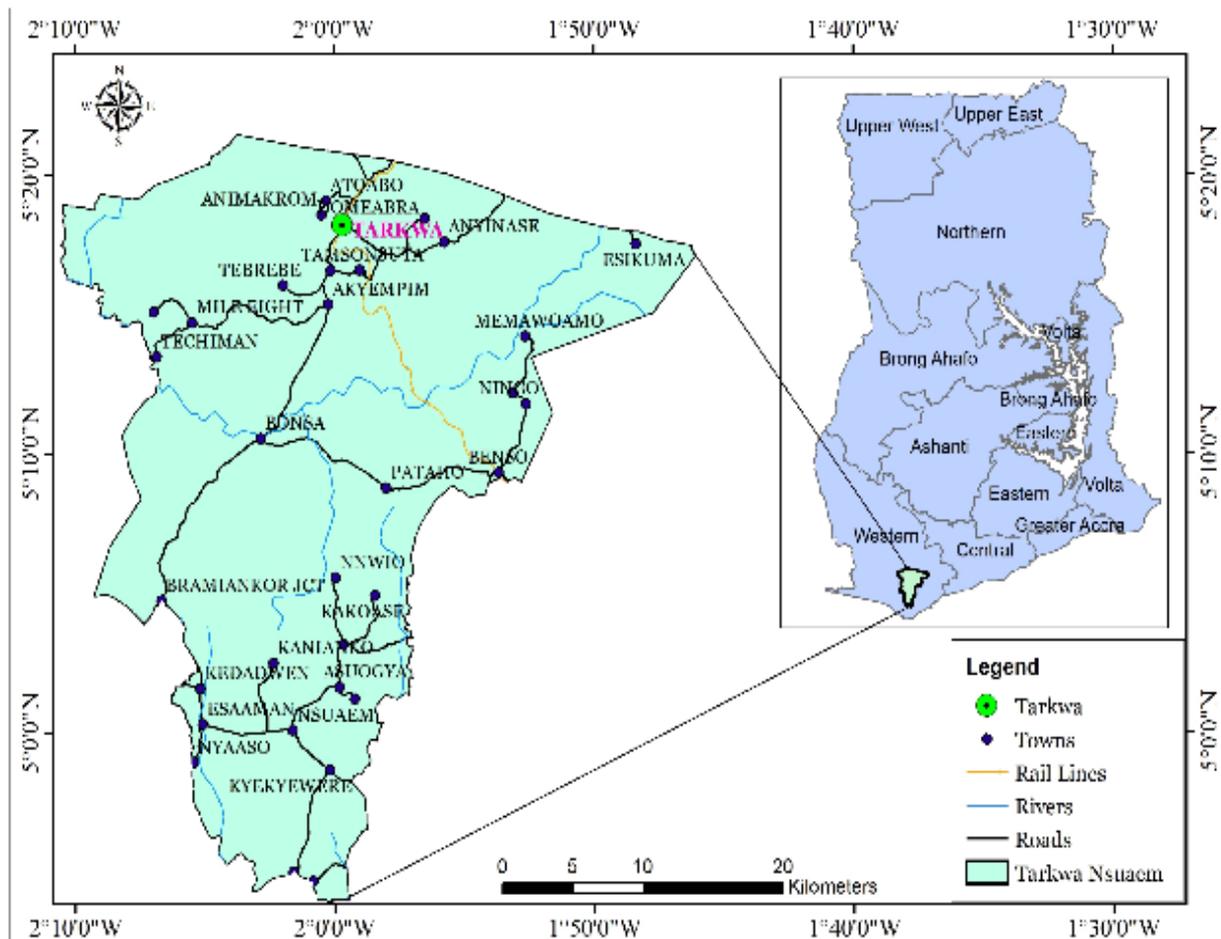


Figure 1: Study area

Table 1: Sample of data used for the study (Units in meters)

SN	Easting (E)	Northing (N)	Elevation (Z)	SN	Easting (E)	Northings (N)	Elevation (Z)
1	163364.5833	69577.1700	76.1957	11	163364.7337	69577.1290	76.3378
2	163364.5338	69577.1707	76.2743	12	163364.5816	69577.1695	76.1958
3	163364.2055	69577.2139	76.2182	13	163364.5829	69577.1701	76.1997
4	163364.3048	69577.3004	75.6254	14	163364.5169	69577.1430	76.1977
5	163364.4388	69577.2636	75.1524	15	163364.5776	69577.2190	75.7588
6	163364.5803	69577.1708	76.2003	16	163364.5784	69577.1706	76.1886
7	163365.1671	69576.8470	76.4418	17	163364.5882	69577.2844	76.6442
8	163364.3452	69576.9806	76.1168	18	163364.5821	69577.1674	76.1937
9	163364.5871	69577.1773	76.2827	19	163364.5820	69577.1680	76.1889
10	163364.5821	69577.1703	76.2027	20	163364.7380	69577.1282	75.8962

2.1 Methods

2.1.1 Outlier Detection

All DGPS data were filtered and cleaned for errors arising from dilution of precision, inadequate number of satellites during observations *etc.*, as advised by Schuessler and Axhausen (2009). Techniques like the Grubb’s Test, Tietjen-Moore Test, Generalised Extreme Studentized Deviate (ESD) Test (Anon., 2013) were applied to remove outliers. Probability plots of the eastings, northings and elevations were conducted to show their central tendencies as shown in figures 2, 3 and 4.

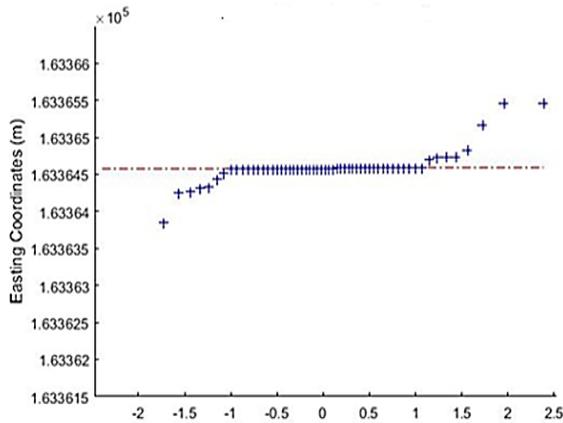


Figure 2: Probability plot of easting coordinates

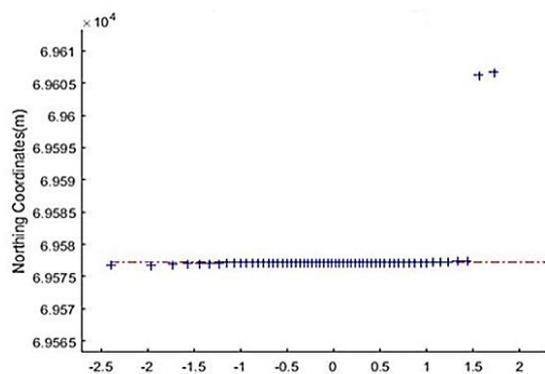


Figure 3: Probability plot of northings coordinates

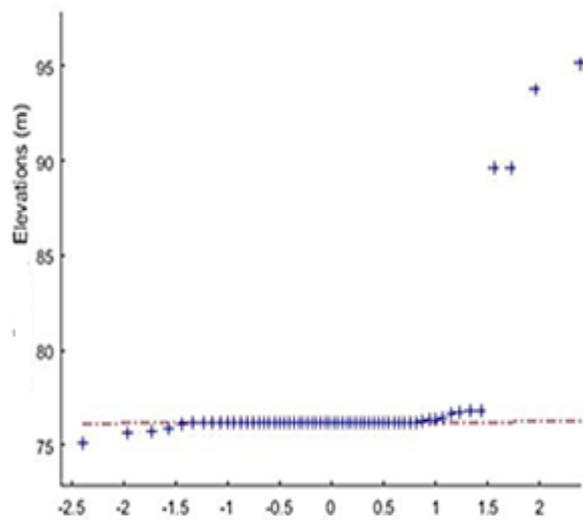


Figure 4: Probability plot of elevations

Error Fitting

A Gaussian distribution fit was applied to the dataset to monitor the distribution of the errors. Figures 5, 6 and 7 show the plots of the error distribution.

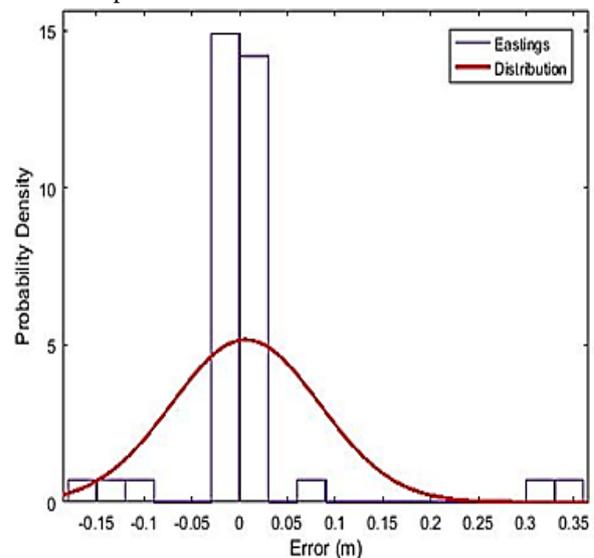


Figure 5: Error fitting for eastings

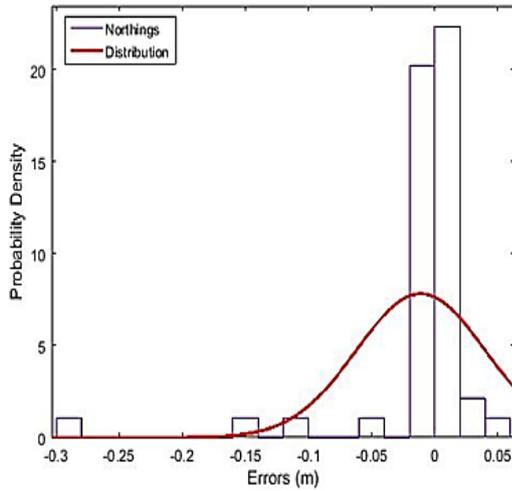


Figure 6: Error fitting for northings

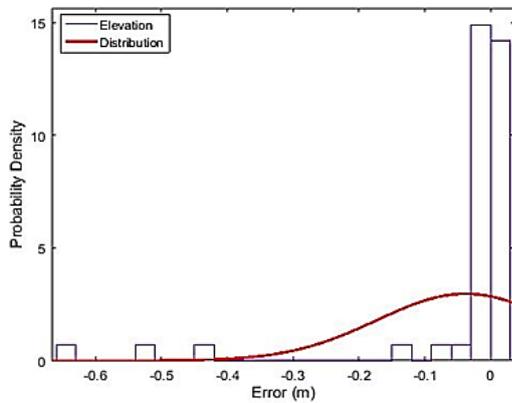


Figure 7: Error fitting for elevations

The fits were realized to be skewed with the summary statistics shown in table 2.

Table 2: Summary statistics for error distribution curve

Coordinate	Parameter		
	Skewness		Standard Deviation
X	-2.7795	2.7459	0.0764
Y	4.2465	-4.1948	0.0505
Z	3.5389	-3.4934	0.1330

2.1.2 Backpropagation Artificial Neural Network (BPANN)

The BPANN training algorithm involves three stages: the input layer, hidden layer and output layer. In this study, the input variables were the Eastings, Northings, and Elevations denoted as $(E_{i,j}, N_{i,j}, Z_{i,j})$ and the output variables were the Eastings and Northings denoted as $(E_{\text{output}}, N_{\text{output}})$. In the BPANN model formulation, the dataset is normalized to ensure constant variation in the model. The selected input and output variables were normalized between the intervals $[-1, 1]$ as presented by Equation (1) (Mueller and Hemond, 2013).

$$y_i = y_{\min} + \frac{(y_{\max} - y_{\min}) \times (x_i - x_{\min})}{(x_{\max} - x_{\min})} \quad (1)$$

where y_i denotes the normalized data; x_i is the measured GPS dataset values; x_{\min} and x_{\max} are the minimum and maximum values of the measured GPS dataset with y_{\max} and y_{\min} values set at 1 and -1, respectively.

Bayesian Regularization learning algorithm was used for the training to find the optimum weight combination. The datasets were divided into training (70 %) and testing (29 %). At the point where the error starts to increasing the training is stopped (Chakraborty and Goswami, 2017). The tansig and purelin activation functions were used for the hidden and output layer respectively in the network training. BPANN is an iterative training procedure, therefore the network was trained varying the number of hidden neurons ranging from 1 to 30 until the optimal model was achieved.

2.1.3 Radial Basis Function Neural Networks (RBFNN)

RBFNN model is an unsupervised learning algorithm and consist of three layers namely; an input layer, a hidden layer and an output layer (Ziggah *et al.*, 2016). The input layer is made up of sensory units that connect the network to its environment. In the second layer, the only hidden layer in the network applied a nonlinear transformation from the input space to the hidden space. The output layer is linear, supplying the response of the network to the activation pattern applied to the output layer. In this study, the input variables were the Eastings, Northings, and Elevations denoted as $(E_{i,j}, N_{i,j}, Z_{i,j})$ and the output variables were the Eastings, Northings and Elevations denoted as $(E_{\text{output}}, N_{\text{output}}, Z_{\text{output}})$. The dataset used for the formulation of the model were divided as training data which consist of 60 % of the total dataset and testing data which consists of 29 %. RBFNN is an exact interpolator (Erdogan, 2009), hence a linear function is used in the input neurons and the connection between the input and hidden layers are not weighted (Kaloop *et al.*, 2017). In this presented study, the Gaussian function is applied, and the output neuron is a summation of the weighted hidden output layer given by Equation (2) (Erdogan, 2009) as

$$y(x) = \sum_{j=1}^n \kappa_j \chi_j(x) \quad (2)$$

where n is the number of hidden neurons, $x \in R^M$ is the input, κ_j are the output layer weights of the radial basis function network, $\chi_j(x)$ is Gaussian radial basis function given by Equation (3) as (Srichandan, 2012; Idri *et al.*, 2010):

$$\chi_j(x) = e^{\left(\frac{-\|x_i - c_j\|^2}{\sigma_j^2} \right)} \quad (3)$$

where $c_j \in R^M$ and σ are the centre and width of j th hidden neurons respectively, $\| \cdot \|$ denotes the Euclidean distance.

2.1.4 Generalized Regression Neural Networks (GRNN)

GRNN is a different kind of Radial Basis Function Neural Network (RBFNN) which is based on Kernel regression networks (Hannan *et al.*, 2010) with one pass learning algorithm and highly parallel structure (Dudek, 2011). It comprises of four layers namely; input layer, pattern layer (radial basis layer), summation layer, and an output layer. It was developed by Specht (1991). In this study, GRNN is being adopted and applied on DGPS dataset to model uncertainties. The input variables were the Eastings, Northings, and Elevations denoted as (E_{ij}, N_{ij}, Z_{ij}) and the output variables were the Eastings and Northings denoted as (E_{output}, N_{output}) . The number of input units in the first layer depends on the total number of the observational parameters. The first layer is linked to the pattern layer and in this layer, each neuron is being presented by a training pattern and its output. The pattern layer is connected to the summation layer. The summation layer consists of two different types of summation namely, single division unit and summation unit (Hannan *et al.*, 2010). The summation with output layer combined perform a normalization of output datasets. In training of the network, radial basis and linear activation functions are used in hidden and output layers. Each pattern layer unit is connected to two neurons in the summation layer. One neuron unit computes the sum of the weighted response of the pattern, and the other neuron unit computes unweighted outputs of pattern neurons. The output layer divides the output of each neuron unit by each other yielding the predicted output variables (Equation (4)):

$$y_i = \frac{\sum_{i=1}^n y_i \cdot \exp - G(x, x_1)}{\sum_{i=1}^n \exp - G(x, x_1)} \quad (4)$$

where y_i is the weighted connection between the i th neuron in the pattern layer and the summation neuron, n is the number of training patterns, G is the Gaussian function given by Equation (5) as

$$G(x, x_1) = \sum_{(k=1)}^m \frac{(x_1 - x_{1k})^2}{\sigma} \quad (5)$$

where m is the number of elements of an input vector, x_l and x_{lk} are the j th elements of x and x_i respectively. During the network training, the spread parameter was varied between 0 and 1 until the output with minimal residuals in terms of statistical analysis was achieved. This same procedure was also done when training the RBFNN.

2.1.5 Multivariate Adaptive Regression Splines (MARS)

The MARS is nonparametric regression technique which works by dividing the variables into regions, producing each region as a least squares equation (Friedman, 1991;

Leathwick *et al.*, 2005). Unlike Ordinary Least Squares, MARS assumes no functional relationship between the target and the predictor variables. The MARS model employed in this study is adopted from the work of Samui and Kurup (2012). The estimation of MARS model is developed in two steps. In the first step (the forward algorithm), MARS is estimated with an excessive number of knots in order to get a better estimate of the predictor variable (Samui and Kim, 2012). In the second step, the knots that contribute significantly to the overall estimation are retained whiles eliminating the less significant once. To ensure the goodness of fit, the Generalized Cross-Validation (GCV) is use to remove the redundant basis functions (Samui and Kothari, 2012; Craven and Wahba, 1979). In this present study, the Salford Predictive Model (SPM) software was adopted to train the MARS model (Yakubu *et al.*, 2018).

2.1.6 Total Least Square (TLS)

TLS is an iterative least squares estimation technique of determining the structure and estimating unknown parameters of a given model (Golub and Van Loan, 1980). In TLS, the orthogonal function estimates the model's parameters one at a time. Also, the percentage reduction with respect to the average of the squared residuals as well as the relative contribution of each term is estimated. Therefore, TLS takes into account the observational errors on both target and the predictor variables which in literature, gives more accurate results (Golub and Van Loan; 1980; Yanmin *et al.*, 2011). The implementation of TLS technique in this study is adopted from Yakubu *et al.* (2018). The adequacy of the estimated model will be tested to ascertain its overall fitness.

2.1.7 Models performance assessment

It is very essential to assess and validate the performance of each estimated model so as to establish whether the model is optimal. This study employs three of the frequently used Performance Indicators (PIs) such as the Root Mean Square Error (RMSE), Mean Horizontal Error (MHE), and Standard Deviation (SD) (Yakubu *et al.*, 2018). The selected PIs will be used in selecting the optimal model to represent each technique considered in this study. After the selection of the optimal models, the best of the techniques for modelling the uncertainties will be selected based on the technique (s) accounting for the least errors.

3. Results and discussion

Statistical methods and analysis were applied to prune the dataset in order to detect and eliminate outliers. The statistical analysis of the two-dimensional shift of the observed data are tabulated in table 3 and table 5. From the tables, it can be observed that the points gradually deviate from their true position. This can be due to the shift in the earth crust, delayed in the transmission of the propagated signal, or the type of the instrument used. This can cause serious problems in higher engineering works which requires a high degree of precision and accuracy.

Table 3: Model results for soft computing techniques (units in meters)

BPANN (Eastings)			
PCI	ME	RMSE	SD
Training	-8.8485E-08	3.0680E-06	8.0400E-09
Testing	-3.6570E-06	1.7798E-05	9.5000E-08
BPANN (Northings)			
PCI	ME	RMSE	SD
Training	3.2645E-07	2.5749E-06	5.1000E-09
Testing	5.2468E-08	8.5937E-07	9.1200E-09
RBFNN (Eastings)			
PCI	ME	RMSE	SD
Training	1.1974E-10	3.0754E-06	9.3500E-09
Testing	-3.9056E-04	1.9850E-03	1.3456E-05
RBFNN (Northings)			
PCI	ME	RMSE	SD
Training	1.0394E-12	3.0449E-07	3.2200E-10
Testing	1.4257E-05	5.4049E-05	3.6000E-07
GRNN (Eastings)			
PCI	ME	RMSE	SD
Training	2.0789E-12	4.9562E-11	8.1300E-13
Testing	8.0286E-12	4.5858E-11	3.4100E-12
GRNN (Northings)			
PCI	ME	RMSE	SD
Training	6.2365E-13	1.9755E-11	9.0400E-15
Testing	-5.0179E-12	1.9486E-11	1.2200E-12
MARS (Eastings)			
PCI	ME	RMSE	SD
Training	0.0012	0.0047	5.7164E-05
Testing	-0.0132	0.0742	0.0003
MARS (Northings)			
PCI	ME	RMSE	SD
Training	-0.0005	0.0024	1.9578E-05
Testing	0.0009	0.0023	7.4954E-05

The optimal equation for modelling the uncertainties in both the Eastings and Northings is given by Equation 6 and Equation 7 respectively. The basis functions used is tabulated in table 4.

$$E_i = 163364 + 1 \times BF11.98271e^{(-005)} \times BF2 + 2.0567e^{(-005)} \times BF6 \quad (6)$$

$$N_i = 69576.8 + 1 \times BF1 + 2.92989e^{(-005)} \times BF2 - 3.6429e^{(-005)} \times BF4 + 7.69176e^{(-006)} \times BF12 \quad (7)$$

Table 4: Basis functions used by the MARS model

Eastings	Northings
BF1 = max (0, E-163364);	BF1 = max (0, N-69576.8);
BF2 = max (0, E -163365);	BF2 = max (0, N-69577.2);
BF6 = max (0, E-163365);	BF4 = max (0, N-69577.2);
	BF12 = max (0, N-69577.2);

Table 5: Results for all models (units in meters)

Model	MHE	RMSE	SD
TLS	7.3901E-06	7.3901E-06	8.7500E-14
GRNN	4.5314E-11	5.2345E-11	1.3200E-13
RBFNN	1.2016E-04	1.0750E-03	1.2195E-06
BPANN	2.8311E-06	1.0215E-05	2.2300E-08
MARS	0.0088	0.0404	3.3158E-05

4. Conclusions and recommendations

Outliers that remain uncertainties in datasets can affect the accuracy of evaluation procedures and estimated parameters if not eliminated. Several researchers have come up with many mathematical models such as classical least squares which have been used for decades. In recent times, with the advancement of science and technology, soft computing techniques have revolutionized the difficulties and deficiencies with the use of classical methods due to its capabilities in denoising datasets to yield a better estimate than the classical methods. This present study assessed the performance of soft computing techniques in modelling the uncertainties of DGPS dataset of two control stations whose coordinates are known to a certain degree of accuracy. The soft computing methods adopted were the BPANN, GRNN, MARS, and RBFNN. The study also compares the performance of these soft computing methods to classical methods such as the TLS due to the wide recommendation by researchers about its efficiency in modelling dataset to give a better estimate. The statistical analysis of the study reveals that all models gave satisfactory result in modelling the DGPS dataset. GRNN outperformed BPANN, RBFNN, MARS, and TLS in modelling the uncertainties in the dataset to give a better estimate. It can, therefore, be stated that all the methods can be used in modelling uncertainties in DGPS datasets.

But with respect to this research the GRNN model has demonstrated superiority over the other models. Hence, natural and engineered structures being monitored for deformation should take into account these methods for modelling the uncertainties in the dataset collected. This can also be automated and integrated in any deformation monitoring process.

This study does not only have a localised significance but will also open more scientific discourse into the applications of soft computing techniques in solving some of the problems in surveying and related disciplines.

References

- Acar, M., M. T. Ozuledemir, O. Akyilmaz, R. N. Celik and T. Ayan (2006). Deformation analysis with Total Least Squares, *Natural Hazards and Earth System Sciences*, 6, 663-669.
- Achour, H., N. J. V. D. Rebai, Driessche and S. Bouaziz, (2012). Modelling uncertainty of stream networks Derived from elevation data using two free softwares: R and SAGA, *Journal of Geographic Information System*, 4, 153-160.
- Akyilmaz, O., M. T. Ozludemir, T. Ayan and R. N. Celik, (2009). Soft computing methods for geoidal height transformation, *Earth Planets Space*, 61, 825-833.
- Alreja, J., S. Parab, S. Mathur and P. Samui, (2015). Estimating hysteretic energy demand in steel moment resisting frames using multivariate adaptive regression spline and least square support vector machine, *AinsShans Engineering Journal*, doi:<http://dx.doi.org/10.1016/j.as.ej.2014.12.006>, 1-7,
- Annan, R. F., Y. Y. Ziggah, J. Ayer and C. A. Odutola, (2016). Accuracy Assessment of heights obtained from Total station and level instrument using Total Least Squares and Ordinary Least Squares Methods, *Journal of Geomatics and Planning*, 3(2), 87-92.
- Anon. (2013). Nist/Sematech e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>.
- Ansah, E., (2016). DGPS Networks adjustments using least squares collocation, (Unpublished) BSc Project Work, University of Mines and Technology, Tarkwa, Ghana, pp. 1-57.
- Asklunel, R. and B. Eldvall, (2005). Contamination of water resources in Tarkwamining area of Ghana, (Unpublished) MSc Thesis, Department of Engineering Geology, Royal Institute of Technology, pp. 1-72.
- Bao, H., D. Zhao, Z. Fu, J. Zhu and Z. Gao, (2011). Application of genetic algorithm improved BP neural network in automated deformation monitoring, Seventh International Conference on Natural Computation, Shanghai-China, Institute of Electrical and Electronics Engineers, doi: 10.1109/ICNC.2011.6022149, pp. 1-5.
- Ben-Gal, I. (2005). Outlier Detection, *Data mining and knowledge discovery handbook*, Springer, pp. 131 – 146.
- Carlson, B. and D. Goodman, (2014). Network least Square adjustment demystified, Carlson Incorporated, Accessed: October 24, 2016, 32 pp.
- Craven, P. and G. Wahba, (1979). Smoothing noisy data with spline function: estimating the correct degree of smoothing by the method of generalized cross-validation, *Numerische Mathematik*, 31, pp. 317-403.
- Devi, K. M. and R. Karthikeyan, (2015). A survey on outlier detection for uncertain meteorology data, *International Journal of Engineering Sciences & Research*, 4(12), 203-208.
- Dreiseit, S., and L. Ohno-Machado, (2002). Logistic regression and artificial neural network classification models: A Methodology Review, *Journal of Biomedical Science*, 35(5-6), 352-359.
- Du, S., J. Zhang, Z. Deng and J. Li, (2014). A new approach of geological disasters forecasting using meteorological factors based on genetic algorithm optimized BP neural network, *Elektronika IR Elektro Technika*, 20(4), 57-62.
- Dudek, G. (2011). Generalized regression neural network for forecasting time series with multiple seasonal cycles, *Springer-Verlag Berlin Heidelberg*, 1, pp. 1-8.
- Erdogan, S. (2009). A comparison of interpolation methods for producing digital elevation models at the field scale, *Earth Surface Processes and Landforms*, 34, 366-376.
- Filzmoser, P. (2004). A multivariate outlier detection method, *Proceedings of the Seventh International Conference on Computer Data Analysis and Modeling*, S. Aivazian, P. Filzmoser & Y. Kharin (eds.). Belarusian State University, Minsk, 1, pp. 18 – 22.
- Friedman, J. H. (1991). Multivariate adaptive regression splines, *Annals Statistics*, 19, 1-67.
- Fu, B., and X. Liu, (2014). Application of artificial neural network in GPS height transformation, *Applied Mechanics and Materials*, 501(504), 2162-2165.
- Gauss, C. F. (1823). *Theoria Combinationis observationum erroribus minimis obnoxiae*, Werke, 4, Gottingen, Germany, 1-5.
- Golub, G. H. and C. F. Van Loan, (1980). An analysis of the Total Least Squares problem, *SIAM Journal on Numerical Analysis*, 17(6), 883-893.
- Hannan, S. A., R. R. Manza and R. J. Ramteke, (2010). Generalized regression neural network and radial basis function for heart disease diagnosis, *International Journal of Computer Applications*, 7(13), 7-13.
- Idri, A., A. Zakrani and A. Zahi, (2010). Design of radial basis function neural networks for software effort estimation, *International Journal of Computer Science*, 7(4), 11-17.
- Ismail, S., A. Shabri and R. Samsudin, (2012). A hybrid model of self-organizing maps and least square Support Vector Machine for River flow forecasting, *Hydrological Earth System Science*, 16, 4417-4433.
- Kalooop, M. R., Rabah, M., Hu, J. W. and A. Zaki, (2017). Using advanced soft computing techniques for regional shoreline geoid model estimation and evaluation, *Marine*

- Georesources & Geotechnology, doi: 10.1080/1064119x.2017.1370622, 1-11.
- Kortatsi, B. K. (2004). Hydrochemistry of groundwater in the mining area of Tarkwa-Prestea, Ghana, (Unpublished) PhD Thesis, University of Ghana, Legon-Accra, Ghana, pp. 1-45.
- Kumar, S., A. Shukla and N. Tiwan, (2014). Bayesian analysis of a stationary AR (1) Model and Outlier, *Electronic Journal of Applied Statistical Analysis*, 7, Issue 1, 81-93.
- Kutoglu, H. S. (2006). Artificial neural networks versus surface polynomials for determination of local geoid, 1st International Gravity Symposium, Istanbul, Turkey, pp. 1-6.
- Lall, U., T. Sangoyomi and H. D. I. Abarbanel, (1996). Nonlinear dynamics of the Great Salt Lake: nonparametric short-term forecasting, *Water Resources Research*, 32, 975-985.
- Leathwick, J. R., D. Rowe, J. Richardson, J. Elith and T. Hastie, (2005). Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish, *Freshwater Biology*, 50, 2034-2051.
- Liao, D. C., Q. J. Wang, Y. H. Zhou, X. H. Liao and C. L. Huang, (2012). Long-term prediction of the earth orientation parameters by the artificial neural network technique, *Journal of Geodynamics*, 62, 87-92.
- Liu, S., and J., Li and S. Wang, (2011). A hybrid GPS height conversion approach considering of neural network and topographic correction, *International Conference on Computer Science and Network Technology*, China, Institute of Electrical and Electronics Engineers, pp. 1-5, doi: 10.1109/ICCSNT.2011.6182386.
- Maxime, C., V. Thierry and S. Romuald, (2005). Modelling uncertainties in pillar stability analysis, *Post Mining 2005*, November 16-17, Nancy, France, pp. 1-12.
- Peprah, M. S. and I. O. Mensah, (2017). Performance evaluation of the Ordinary Least Square (OLS) and Total Least Square (TLS) in adjusting field data: an empirical study on a DGPS Data, *South African Journal of Geomatics*, 6(1), 73-89.
- Peprah, M. S., Ziggah, Y. Y., and Yakubu, I. (2017). Performance evaluation of the earth gravitational model (EGM2008) – A Case Study, *South African Journal of Geomatics*, 6(1), 47-72.
- Pinho, R., H. Crowley and J. J. Bommer, (2008). Open-source software and treatment of epistemic uncertainties in earthquake loss modelling, *The 14th World Conference on Earthquake Engineering*, October 12-17, 2008, Beijing, China, pp. 1-8.
- Samui, P. (2013). Multivariate Adaptive Regression Spline (MARS) for prediction of elastic modulus of jointed rock mass, *Geotechnical and Geological Engineering*, 31, 249-253.
- Samui, P. and D. P. Kothari, (2012). A multivariate adaptive regression spline approach for prediction of maximum shear modulus (Gmax) and minimum damping ratio (ξ_{min}), *Engineering Journal*, 16(5), 1-10.
- Samui, P. and D. Kim, (2012). Modelling of reservoir-induced earthquakes: A multivariate adaptive regression spline, *Journal of Geophysics and Engineering*, 9, 494-497.
- Samui, P. and P. Kurup, (2012). Multivariate Adaptive Regression Spline (MARS) and Least Squares Support Vector Machine (LSSVM) for OCR prediction", *Soft Computing*, 16, 1347-1351.
- Schuessler, N. and K. Axhausen, (2009). Processing raw data from global positioning systems without additional information, transportation research record, *Journal of the Transportation Research Board*, 2105, 28 – 36.
- Sharma, P. K., H. Haleem and T. Ahmad, (2015). Improving classification by Outlier Detection and Removal, *Advances in Intelligent Systems and Computing*, 2(338), 621-624.
- Specht, D. (1991). A general regression neural network, *Institute of Electrical and Electronics Engineers Transactions on Neural Networks*, 2(6), 568-576.
- Srichandan, S. (2012). A new approach of software effort estimation using radialbasis function neural networks, *ISSN (Print)*, 1(1), pp. 2319-2526.
- Yakubu, I. and B. Kumi-Boateng, (2011). Control position fix using single frequency global positioning system receiver techniques – A case study, *Environmental and Earth Sciences Research Journal*, 3(1), 32-37.
- Yakubu, I., Y. Y. Ziggah and M. S. Peprah, (2018). Adjustment of DGPS Data using artificial intelligence and classical least square techniques, *Journal of Geomatics*, 12(1), pp. 13-20.
- Yanmin, J., T. Xiaohua and L. Lingyun, (2011). Total least squares with application in geospatial data Processing, *Proceedings of Nineteenth International Conference on Geoinformatics*, Shanghai, China, pp. 1-3.
- Zahra, T. (2010). Quantifying uncertainties in landslide runout modelling, (Unpublished) MSc Thesis, Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands, pp. 1-94.
- Ziggah, Y. Y., H. Youjian and C. A. Odutola, (2013). Determination of GPS coordinate transformation parameters of geodetic data between reference datums, A case study of Ghana geodetic reference network, *International Journal of Engineering Sciences and Research Technology*, 2(4), 2277-9655.
- Ziggah, Y. Y., H. Youjian, A. Tierra, A. A Konate, and Z. Hui, (2016). Performance evaluation of artificial neural networks for planimetric coordinate transformation-A case study, *Ghana, Arabian Journal of Geosciences*, Vol. 9, pp. 698-714, doi: 10.1007/s12517-016-2729-7.

Performance enhancement of standard fuzzy majority voting-based fusion of probabilistic classifiers

Mahmoud Salah

Department of Surveying Engineering, University of Benha, 108 Shoubra Street, Cairo, Egypt

Email: engmod2000@yahoo.com

(Received: Dec 16, 2018; in final form: May 10, 2019)

Abstract: Combining classifiers is essential for feature extraction and mapping applications. This paper proposes an approach to improve the performance of one of the most frequently used Multiple Classifier Systems (MCSs), namely the Fuzzy Majority Voting (FMV). First, a set of texture attributes has been generated from a 0.82mpan-sharpened IKONOS image covers the test area. The generated attributes along with the original image have been applied as input for three-member classifiers: Artificial Neural Networks (ANN); Support Vector Machines (SVM); and Classification Trees (CT). Before combination, a weighting criterion has been determined, based on the performances of each member classifier, and assigned to the output of that classifier. After that, The FMV has been applied for combining the weighted results from the three-member classifiers to extract buildings (B), roads (R) and vegetation (G). The proposed method has been tested and compared with the three-member classifiers as well as the standard FMV. The results have been analyzed considering four different aspects: (1) overall accuracy; (2) class accuracy; (3) sensitivity to training sample size; and (4) computational complexities. The proposed method resulted in an overall classification accuracy of about 95.60%, which is 3.88, 6, 8.51 and 1.24% better than ANN, SVM, CT, and standard FMV respectively. On the other hand, most of the class-accuracies are much better and less variable than those obtained by any member classifier as well as the standard FMV. While the proposed method is stable and always outperforms individual classifiers even in the cases of small size training samples, its computational cost is still comparable with that of standard FMV.

Keywords: MCSs, FMV, High Resolution (HR) satellite imagery, feature extraction.

1. Introduction

Feature extraction from HR satellite imagery is an important task in remote sensing (RS) and geographic information system (GIS) applications. However, Classification of remotely sensed imagery is still a challenge because of the low illumination and low spatial resolution of satellite imagery, noise, ambiguity and uncertainty in the decision-making process. There is no single classifier that can be optimal for all classification problems since single classifiers sometimes lead to inefficient classification accuracy. Combining classifiers can improve the classification accuracy by integrating the outputs of multiple individual classifiers through some linear or non-linear MCSs (Moustakidis et al., 2012). Over the last decade, MCSs can be considered as one of the most important advancements in the field of pattern recognition (International workshop on multiple classifier systems, 2007).

The appropriate combination of individual classifiers usually results in better performance in terms of classification accuracy and/or CPU time. MCSs can outperform any individual classifier in cases of small training dataset, local optima problem and a huge amount of data (Woz'niak et al., 2014). As well, it can improve the performance of weak classifiers and stabilize the decision of random ones such as ANN and CT.

On the other hand, the diversity of member classifiers of a given MCS can perfectly handle the problem of noisy data (Ponti-Jr. and Papa, 2011). It is worth mentioning that classifiers with correlated results may provide lower accuracy than the worst classifier. In the case of correlated classifiers, the MCSs error will be almost the

same as the average error of the member classifiers. On the other hand, it will be n times smaller than the average error of the members in the case of statistically independent classifiers (Tumer and Ghosh, 1996). In this regard, the successfulness of a MCS is based on the degree of diversity between individual classifiers. Diversity can be achieved by using: different input features, different training samples, or different classifiers. A detailed review of the most common diversity measures is given by Ranawana and Palade (2006). Recently, MCSs have been introduced to remote sensing applications in: satellite image classification (Maulik and Chakraborty, 2010); land cover mapping (Han et al., 2012); and change detection (Du et al., 2013).

In general, MCSs can be grouped into three categories: parallel; serial; and hierarchical (Lv et al., 2000). In the parallel form, the same data are used as input for each individual classifier independently, and the final decision is made by combining their independent results. In the serial form, individual classifiers are applied in sequence. This form starts with a primary classifier, the classifier with the cheapest computational cost, followed by the secondary classifiers, the ones with higher computational cost (Fumera et al., 2004). The hierarchical method combines both parallel and serial techniques in order to obtain optimal combination results (Ranawana and Palade, 2006). The majority of publications are focused on the parallel combination approach since it is simple to implement, easy to analyze and can take advantages of all member classifiers (Woz'niak et al., 2014).

Segrera and Moreno (2005) categorized the methods for building MCSs into two groups: ensemble, and hybrid. Ensemble MCSs combine classifiers with the same

learning models while modifying the input training data set for each classifier. Ensemble-based combination typically has lower generalization errors than any of its individual models (Ahn et al., 2007). Bagging and boosting techniques have been considered in most of the previous studies as a typical example of ensemble-based approaches. Both bagging and boosting introduce diversity by considering different training samples and only one base classifier. Dietterich (2000) concluded that with low noise data, boosting outperforms bagging technique. On the other hand, bagging outperforms boosting in the case of high noise data (Khoshgoftaar et al., 2011). The main problem with Bagging and boosting is the large number of classifiers in the final MCS. As well, measuring diversity is still an open research area (Cavalcanti et al., 2016). Hybrid MCSs, on the other hand, combine classifiers with different learning algorithms. In this regard, many combination rules have been considered in the literature: the intersection of decision regions (Haralick, 1976); voting methods (Mazurov et al., 1987); prediction by top choice combinations (Wemecke, 1992); Dempster-Shafer theory (Xu et al., 1992); and ranking methods (Ho et al., 1994).

Dai and Liu (2010) proposed a MCS to combine the results from six base classifiers: maximum likelihood (ML); SVM; ANN; spectral angle mapper (SAM); minimum distance (MD); and CT. A voting strategy was applied for the combination. The results confirmed that the MCS performed much better than any base classifier. Ghimire et al. (2012) have compared three combination techniques based on decision trees (DT): bagging; AdaBoost; and random forest (RF). For the three cases, the MCS has outperformed the DT-base classifier. Kumar and Meher (2013) proposed an effective MCS based on multiple rules of granular neural networks (GNN) with improved performance classification accuracy. Khosravi and Beigi (2014) applied bagging and boosting to construct a MCS based on SVM to classify a hyperspectral dataset. The results showed a perfect performance of the MCS for classifying high dimensional data. Chen et al. (2017) constructed a MCS to combine SVM, DT and ANN using the AdaBoost technique. The results showed that the MCS has effectively improved the classification accuracy as compared with individual classifiers.

In order to select the most suitable combiner for a given problem, some guidelines have been given in the literature: majority voting (Kimura and Shridhar, 1991) for combining class labels; averaging techniques (Hashem and Schmeiser, 1995) for combining posterior probabilities; fuzzy logic (Zadeh, 1965), Dempster-Shafer theory of evidence (DS) (Rogova, 1994) or ANN (El-Melegy and Ahmed, 2007) for combining fuzzy membership values. Detailed reviews of MCSs are given by Wozniak et al. (2014). The most recent techniques are usually presented in the International Workshop on MCSs and in Machine Learning and Pattern Recognition Journals.

It is worth mentioning that the fuzzy set theory is more suitable for pattern recognition in the case of remotely

sensed data where classes are normally ill-defined and overlapped (Kuncheva, 2000). On the other hand, it is not depending on the input data which is the main drawback of most of the existing MCSs. Many attempts have been made for RS image analysis and classification using fuzzy sets (Chen, 2000; Tso and Mather, 2001; Ghosh et al., 2008). Salah et al. (2010) applied the standard FMV to combine non-weighted ANN, CT and SVM classifiers using aerial images and LiDAR data. The results demonstrate that the standard FMV has improved the classification accuracy as compared with the best single classifier. FMV is the most commonly used fuzzy sets technique since it is easy to apply and able to manage imperfect data (Ponti-Jr., 2011).

To the best of the author knowledge, the effect of weighting classifiers and training sample size on the accuracy and robustness of FMV-based fusion have not been considered in the literature. The objective of this paper is to define clear guidelines to explain under which conditions the FMV are able to improve the performance of individual classifiers. To meet the objective, ANN, SVM and CT have been adopted as base classifiers. The three classifiers have different modelling and learning criteria which lead to different errors and then complementary information. In this regard, the parallel combination technique has been applied. The three-member classifiers have been applied to classify the test area using IDRISI Taiga software (Clark labs, 2012). The proposed fusion method in this research has been implemented through a set of codes generated by the author in Matlab environment. For the rest of the paper, the term WFMV will be used to refer to the FMV-based fusion of the weighted classifiers, while the term SFMV will be used to refer to the FMV-based fusion of the standard or non-weighted classifiers.

2. Study area and data sources

2.1 Pan-sharpened satellite image

An IKONOS image covers the test area was collected on April 17, 2010, and supplied in a digital TIFF format. The IKONOS image has been created using a pan-sharpening process that combines the 0.82m panchromatic band with the 3.2m multispectral bands to create 0.82m colour image. Table 1 summarizes IKONOS satellite specifications. The test area is a dense urban area with medium size residential buildings, a large network of main and minor roads, as well as open vegetation areas as shown in figure 1.

Table 1: IKONOS satellite specifications

Imaging Mode	Panchromatic	Multispectral
Pixel Size	0.82 meter	3.2 meter
Spectral Range	450-900 nm	450-520 nm (blue) 520-600 nm (green) 625-695 nm (red) 760-900 nm (NIR)
Dynamic Range	11 bit/pixel	11 bit/pixel



Figure 1: The one-meter pan-sharpened IKONOS image covers the study area

2.2 Reference data

In order to evaluate the performance of the proposed combination method, B, R and G were digitized in the image and applied as reference data as shown in figure 2, Class "G" mainly corresponds to grass and trees. All recognizable objects were digitized independently of their size. Joined buildings that were obviously separated were digitized as separate buildings; otherwise, they were digitized as one building.



Figure 2: Reference data used for evaluating the performance of the proposed method. Red: B, green: G, black: R

2.3 Feature attributes

In order to describe classes effectively, a wide variety of spectral attributes have been generated and only the most useful ones, as shown in figure 3, have been statistically selected based on Yang (2007). The selected attributes and the original multispectral image have been used simultaneously as input data for the classification process. The objective is to solve for two common problems associated with HR digital imagery which are: 1) shadows caused by buildings and trees; 2) and spectral variability within the same land use/cover class (Lu and Weng, 2007). On the other hand, it provides useful information for improved land use/cover classification (Hirose et al., 2004). For more details about the formulas used for calculating attributes, one can refer to Russ (2002).

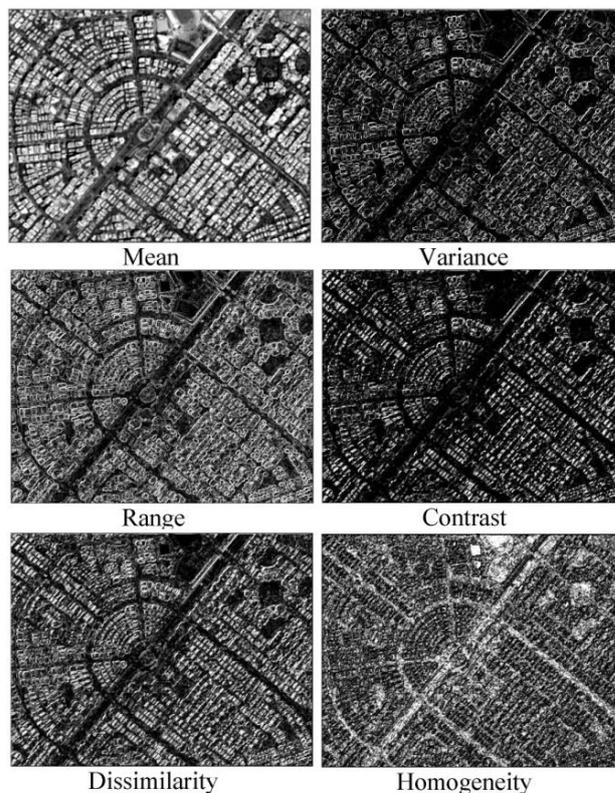


Figure 3: The set of attributes that have been applied as the input for individual classifiers

2.4 Training Datasets

Training datasets assemble a set of statistics to describe the spectral pattern for each land use/cover class in the image. A minimum of $(n+1)$ pixels is required for a signature with n is the number of bands (Lillesand and Kiefer, 2004). The training data used are sets of manually digitized samples from the image for each land use/cover class. Polygons of approximately the same areas were digitized for B, R and G classes. As recommended by Kuncheva (2004), the same training samples have been applied to train all the member classifiers. Digital numbers (DNs) in a range between 0 and 255, corresponding to reflectance values, have been applied to generate the training samples. The selected signatures are compared in a graph representing DN values for each signature from the red band as shown in figure 4. The clear separation for most DN values indicates that the selected signatures represent a completely distinct set of pixels, which is essential for good classification results.

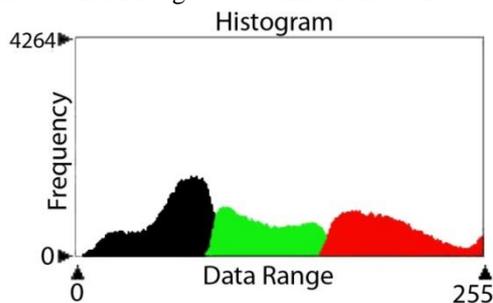


Figure 4: Minimum and maximum DN values for signatures, from the red band: black for R; green for G; and red for B

3. Methodology

The proposed FMV-based fusion of ANN, SVM and CT classifiers has been implemented in several steps as shown in figure 5. The proposed MCS has four phases. First, training data are fed into the three-member classifiers to obtain individual decisions. After that, the obtained probabilities were weighted according to the relative importance of each classifier. The 10-fold cross-validation technique was then applied to tune the shape and position of the fuzzy membership function. At the end, FMV is applied to combine the weighted probabilities from the three-member classifiers and form the final decision.

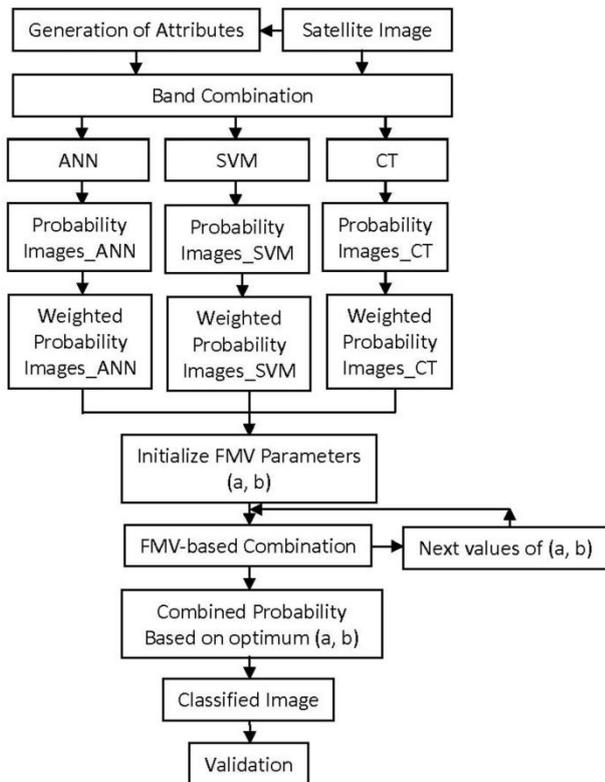


Figure 5: The proposed FMV-based fusion workflow

3.1 Base Classifiers

In order to improve the performance of the MCS, individual classifiers should have different mathematical concepts and offer complementary information. As well, in the case of two base classifiers, a limited improvement in classification accuracy can be obtained by the MCS (Chen et al., 2017). In this regard, three different algorithms have been applied as member classifiers. These classifiers represent different learning criteria and include SVM as a machine learning classifier; ANN as an artificial neural networks classifier; and CT as a statistical classifier. The output of each classifier is a degree of membership of every pixel for each class.

3.1.1 ANN

ANN is a self-learning algorithm that can compensate for uncertainty in information and can perfectly handle the problem of high spectrum confusion in remotely sensed data. This can be done by setting the number of nodes in the hidden layers (Coppin et al., 2010). The ANN is

trained based on randomly chosen initial weights (Hu, 2000). The most common and widely used feed-forward back propagation neural algorithm, multi-layer perceptron (MLP), has been applied. The network consists of three layers: input; hidden; and output. The number of input nodes of the MLP is the number of the input features; the number of output nodes is the number of classes; and the number of hidden nodes is between $2N$ to $3N$ where N is the number of classes (Ghosh and Uma Shankar, 2010). In this regard, the MLP was a seven hidden layers MLP with nine input neurons, one for each input variable, and three output neurons, one for each class. Except for the input nodes, a weight is calculated for each node as the sum of the output at the nodes to which it is connected in the preceding layer. In order to derive the final output and fed it to the nodes in the next layer, the weighted sum is passed through a transfer function as follow:

$$net_v = \sum_u W_{uv} O_u + bias_v \text{ and } O_v = S(net_v) \quad (1)$$

Where:

w_{uv} : the weight for the connection between nodes u and v

$bias_v$: the bias for node v

O_u : the output at node u

S : the sigmoid activation function. This function can perfectly handle nonlinear problems (Cybenko, 1989).

For weights updating, MLP uses a back-propagation learning algorithm to reduce the sum of square error between the obtained and desired output in a descending manner as follows (Haykin, 1998):

$$\Delta W_{uv}(n+1) = \alpha \Delta W_{uv}(n) + \eta \delta_v O_u \quad (2)$$

Where n , α , η and δ are the iteration number, momentum parameter, learning rate and node error respectively. In this regard, each input pattern is assigned to the class that corresponds to the highest node value obtained at the output of the MLP. In order to improve the performance of the MLP with reasonable processing time, a set of parameter values suggested by Kavzoglu and Mather (2003) have been applied as shown in table 2.

Table 2: The basic architecture to start the MLP classifier

Parameter	Value
n	10,000
α	0.5
η	0.05
δ	0.0001

3.1.2 CT

CT was introduced by Breiman et al. (1984). It is a non-parametric technique that uses an iterative procedure known as binary recursive partitioning. In this regard, a heterogeneous sample of training data with multiple classes is hierarchically and progressively subdivided into more homogeneous classes based on a binary splitting rule to form the tree. The tree is then used to classify the whole datasets. Classification trees have proved to be strong, simple to implement, ideal for noise minimization, highly automatic and perfect for complex

data such as multi-source and/or multi-scale data. In the classification process, only the most useful attributes are selected and used (Chen et al., 2017).

Three models can be used with CT as splitting criteria which include: Entropy, Gain Ratio, and Gini. The Entropy algorithm (after Shannon, 1949) has proved to be preferable for classification problems from HR digital imagery (Salah et al., 2011) and has been applied as the splitting criterion in this study. The method decreases the entropy until a terminal node that has zero entropy (contains pixels from one class) is reached. In order to identify class x_i of a training dataset in node N , the entropy can be described as in equation 3 where $P(x_i)$ is the probability of class x_i .

$$Entropy(N) = - \sum_{i=1}^l P(x_i) \log_2 P(x_i) \quad (3)$$

A 10-fold cross-validation process has been applied for pruning the trees. This technique has proved to be highly accurate and requires no independent dataset for assessing the performance of the model.

3.1.3 SVM

Recently, SVM has become a common tool to classify linear and nonlinear problems. It is based on statistical learning theory and has excellent learning performance especially when applied to remotely sensed data. First, the input feature space is transformed into a high dimensional one, and then an optimal hyperplane is fitted into data to separate 0 and 1 classes by maximizing the margin between them. The closest data points to the hyperplane are referred to as support vectors (Vapnik, 1995).

Since the One-Against-One (1A1) technique normally results in a huge number of binary SVM as well as intensive computations, the One-Against-All (1AA) technique has been applied to solve for the 0/1 classification problem. The radial basis function (RBF) kernel can nonlinearly map more complex data into a higher dimensional space with reasonable processing times and has proved to be effective for remote sensing applications. In this regard, it has been applied for fitting the hyperplane into data (Van der Linden et al., 2009). The general mathematical representation of the RBF kernel is shown in equation 4. The gamma term, γ , is a user-controlled parameter and its correct definition can significantly improve the performance of the SVM classification.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (4)$$

The 10-fold cross-validation technique has been applied to determine the optimal γ value, 0.03. This technique has proved to be effective to prevent over-fitting problems and usually results in better performance (Hsu et al., 2009). The sequential minimal optimization (SMO) algorithm has been applied for training the SVM through breaking the large size optimization problems into a series of smallest ones. This can speed up the computations and minimize memory requirements (Platt, 1999).

3.2 WFMV-Based fusion

FMV has proved to be a powerful technique to handle the uncertainties and imprecision in remotely sensed data by defining a fuzzy membership function. A membership function is a relation that shows how a certain point in the input space will be mapped as a membership value in the output space. Before incorporating the probabilities, pp_i , from ANN, SVM, and CT into the FMV, the probabilities are weighted by the accuracies estimated for the corresponding classifiers as an importance weight. This weighting process minimizes errors for the FMV in the case of independent outputs. Let the classification accuracies obtained from the accuracy assessment process are: α_{c1} ; α_{c2} ; and α_{c3} for ANN, SVM, and CT respectively. The weighted probabilities, pp_i , for a classifier c_i may be given as follows:

$$w_i = \log \left(\frac{\alpha_{ci}}{1-\alpha_{ci}} \right) \quad (5)$$

$$pp_i = w_i * pp_i \quad (6)$$

The idea behind the FMV is to give some semantics, meanings, to the weights. Based on this semantics, the weights can be determined directly (Zadeh, 1983). First, the membership function of relative quantifiers are defined as in equation 7 (Herrera and Verdegay, 1996). With parameters $a, b \in [0, 1]$ and pp_i is the weighted class membership of pixel i .

$$Q_{P_i} = \begin{cases} 0 & \text{if } pp_i < a \\ \frac{pp_i - a}{b - a} & \text{if } a \leq pp_i \leq b \\ 1 & \text{if } pp_i > b \end{cases} \quad (7)$$

The value 1 of Q_{P_i} indicates that the quantifier is completely satisfied. On the other hand, the quantifier is not fulfilled at all if $Q_{P_i} = 0$ and any intermediate value

Q_{P_i} indicates an intermediate fulfilment degree. The optimal selection of fuzzy parameters (a, b) has a direct impact on the FMV performance since they control the shape and position of the membership function. Unfortunately, there is no precise mathematical method to define these parameters (Saheb et al., 2013). A grid-search on a and b using a 10-fold cross-validation was used for this purpose. Basically, pairs of (a, b) were tested and the one with the best cross-validation accuracy was selected. In this regard, a grid with an interval of 0.1 for both a and b has been applied. Then, the weights based on the linguistic quantifier can be determined as in equation 8, with i is the order of a given classifier after ranking Q_{P_i} in a descending order and N is the total number of classifiers (Yager 1998):

$$w_{P_i} = Q_{P_i} \left(\frac{i}{N} \right) - Q_{P_i} \left(\frac{i-1}{N} \right), \text{ for } i = 1, \dots, N \quad (8)$$

The final combined probability can be determined as in equation 9, with k is the number of classes.

$$P_{WFMV} = \underset{k}{\operatorname{argmax}} \left[\sum_{i=1}^N w_{p_i} p_{p_i} \right] \quad (9)$$

3.3 Accuracy assessment

In order to evaluate the performance of the proposed method, the results have been compared with the reference data. The overall classification accuracy *OA* has been determined as in equation 10 with *NCP* is the total number of correctly classified pixels and *NRP* is the total number of reference pixels. Since the overall classification accuracy is just a global measure for the performance of the combination process, the users and producers accuracies (*UA* and *PA*) have been used. Unlike overall classification accuracy, *UA* and *PA* clearly indicate how the proposed methods improve or deteriorate the results for individual classes as shown in equations 11 and 12. *CP* is the correct class predictions, *TP* is the total predictions and *TCP* is the total class pixels.

$$OA = \frac{NCP}{NRP} \quad (10)$$

$$UA = \frac{CP}{TP} \quad (11)$$

$$PA = \frac{CP}{TCP} \quad (12)$$

4. Results and discussion

Initially, classification of the satellite image has been performed and the parameters for each classifier have been estimated using the labeled training samples. Once the optimum parameters were selected, each classifier has been applied to classify the whole image. The obtained results are nine probability images, three for each classifier, representing the membership of each pixel to each class. The probabilities were then modified by assigning weights derived from the classification accuracy of the corresponding classifier. The membership values are true probabilities in the range of 0 to 1 as shown in figure 6.

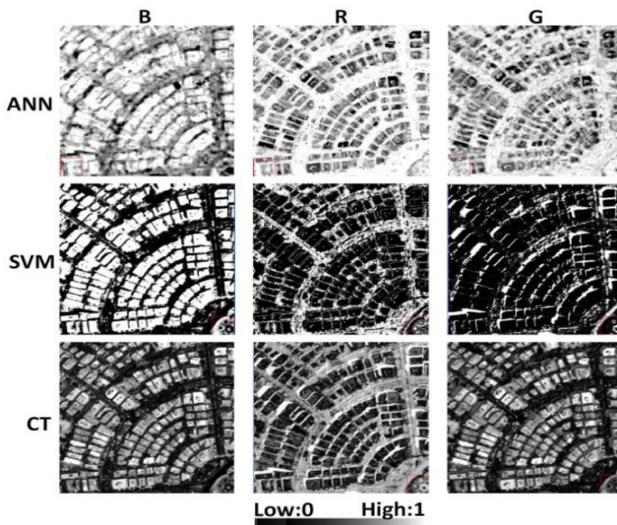


Figure 6: A typical example showing the weighted membership images of individual classifiers

Before applying the WFMV algorithm to combine these probabilities, a grid-search on *a* and *b* using a 10-fold cross-validation and grid interval of 0.1 for both *a* and *b* has been applied. As a result, a relative quantifier with parameters (0.1, 0.5) has performed the best for the membership function Q_{P_i} in equation 7 as graphically depicted in figure 7. Once *a* and *b* have been determined, they were applied with the nine weighted probability images to perform the WFMV-based combination. Figure 8 is a typical example of the WFMV output which is three probability images representing the membership values of every pixel for each class.

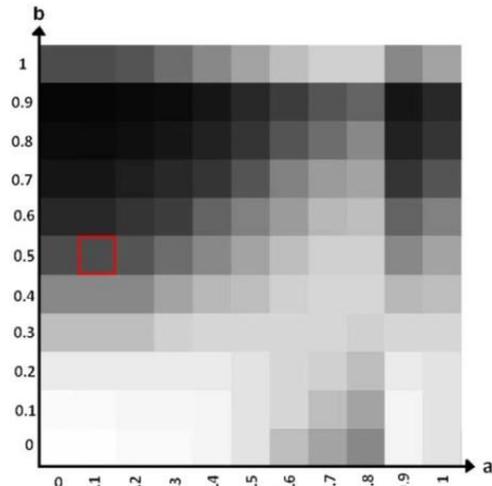


Figure 7: The grid search results using the input data

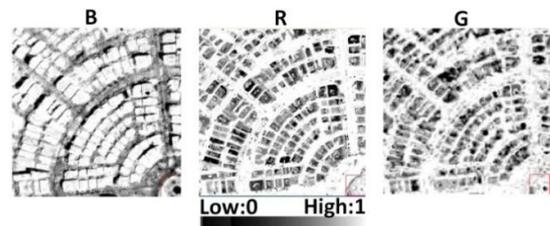


Figure 8: A typical example showing the membership images of the WFMV system

For each pixel, the membership values for all classes were compared and the class with the highest value was assigned to that pixel to create a WFMV-based classification image. Figure 9 is a typical example illustrates the original image, the classification results obtained for individual classifiers, SFMV and WFMV. By focusing on the buildings inside the white squares, one can find that the WFMV has detected complete buildings much better than SVM and CT, and as good as ANN and SFMV. On the other hand, the white circled regions indicate that WFMV has detected separate buildings more accurate than any individual classifier, as well as SFMV. However, many vegetation was classified as roads by the three-member classifiers and hence by SFMV and WFMV as graphically depicted in the white rectangles. An expected reason for that can be the high degree of similarity between the spectral reflectance of roads and vegetation in the used pan-sharpened IKONOS image. One possible solution is to use a Normalized Difference Vegetation Index (NDVI) which may increase the classification accuracy due to its ability to detect vegetation accurately.

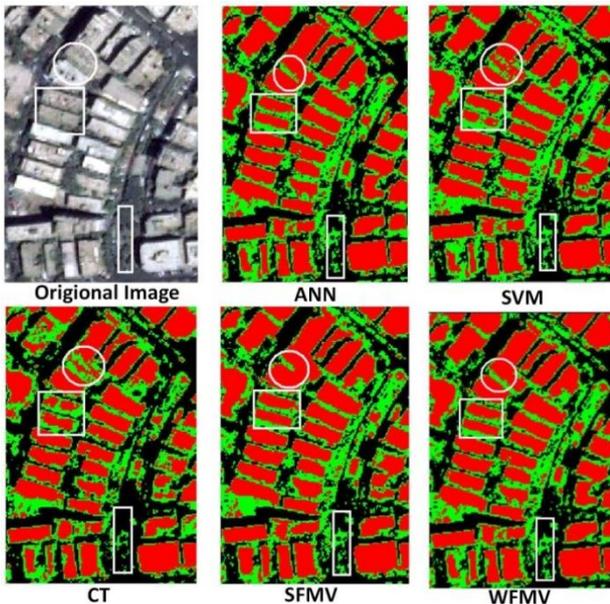


Figure 9: Results of individual classifiers, SFMV and WFMV

4.1 Overall accuracy

In order to evaluate the performance of the WFMV, it has been compared with single classifiers and with SFMV. Table 3 summarizes the obtained accuracies for the three classifiers and their combinations by SFMV and WFMV. The WFMV has statistically performed better than any individual classifier as well as SFMV. For individual classifiers, ANN performed the best with 91.72% overall accuracy, followed by SVM with 89.60% overall accuracy and CT with an overall accuracy of 87.09%. The WFMV resulted in overall classification accuracy of about 95.60%, which is 3.88, 6, 8.51 and 1.24% better than ANN, SVM, CT and WFMV respectively.

Table 3: Overall accuracy of different classification methods

classifier	ANN	SVM	CT	SFMV	WFMV
Overall Accuracy	91.72	89.60	87.09	94.36	95.60

4.2 Class accuracy

In terms of class accuracy, the three-member classifiers resulted in different class accuracies for the same test area as shown in table 4. No single classifier has performed the best for all classes. A typical example is that CT resulted in lower UA for vegetation, 58.75%. On the other hand, it outperformed the ANN and SVM in classifying buildings with UA of about 99.99%. These results confirm that classifiers with different algorithms are complementary and result in different classification accuracies for different classes.

Table 4: Classification accuracies obtained for different classifiers

classifiers	B		R		G		Average	SD
	UA	PA	UA	PA	UA	PA		
ANN	99.13	91.09	98.02	91.12	70.19	94.52	90.68	10.59
SVM	91.34	70.53	81.94	99.03	94.55	96.06	88.91	10.75
CT	99.99	85.28	96.15	85.85	58.75	94.38	86.73	14.89
SFMV	99.23	94.02	98.79	89.33	73.37	96.72	91.91	9.79
WFMV	98.15	95.38	87.99	96.75	99.41	94.62	95.38	4.02

Assessments of class accuracies confirmed that the WFMV-based fusion performed the best in most cases as shown in table 4. Most of the class-accuracies are improved by the WFMV. Whereas the application of ANN, SVM, CT and SFMV resulted in average class accuracies (average of UA and PA) of 90.68, 88.91, 86.73 and 91.91% respectively, the application of WFMV fusion displayed a significant improvement and resulted in average class accuracy of 95.38%. Another advantage of WFMV-based fusion is that the obtained errors are less variable. Whereas the application of ANN, SVM, CT and SFMV resulted in standard deviations (SD) of 10.59, 10.75, 14.89 and 9.79 respectively; the WFMV-based fusion resulted in SD of 4.02. Thus it conforms to the requirement of Anderson et al. (1976) that the class accuracies of different classes should be about equal.

4.3 Sensitivity to training sample size

In order to obtain a robust decision about the performance of the WFMV system, five different training samples (100, 200, 300, 400 and 500 pixels) evenly distributed through the test area were selected and tested. As can be observed from figure 10, WFMV always improves the performance of individual classifiers and outperforms the SFMV even in the cases of small size training samples. This behaviour can be clearly observed in the cases of training samples of size less than 300 pixels. On the other hand, WFMV is the most stable classifier followed by SFMV, ANN, SVM and CT respectively. Decreasing the sample size from 500 to 100 pixels has decreased the obtained classification accuracies by 12.83, 14.38, 22.7, 30.23 and 35.42% for WFMV, SFMV, ANN, SVM and CT respectively.

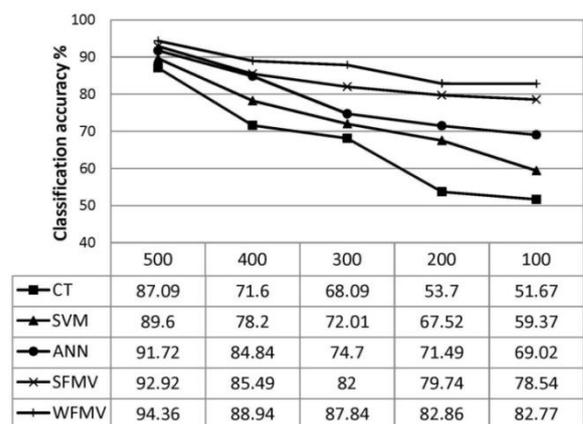


Figure 10: Performance evaluation of individual classifiers along with their combination using SFMV and WFMV on different data samples

4.4 Computational complexities

Another problem to be investigated is the computational cost associated with each model. In this regard, it is worth mentioning that the computer system used is of: Genuine Intel (R) CPU T2130, 1.86GHz and 783MHz, and 896 MB of RAM along with a test area of approximately one Km². In order to evaluate the performance of the proposed model, it has been compared with the three-member classifiers and SFMV. Table 5 shows the elapsed time in seconds during the combination process. In terms of individual classifiers, CT is the cheapest classifier with almost 12 second/km² processing time, followed by SVM and ANN with almost 15 and 17 second/km² respectively. The time required by the SFMV technique is almost 24 seconds/km² which is almost two times more than that required by the cheapest classifier. WFMV, on the other hand, is the most complex classifier with 27 second/km². However, it is still comparable with SFMV.

Table 5: Computational complexity comparison of WFMV-based fusion with individual classifiers and SFMV

classifier	CPU time (second/km ²)
ANN	17.082077
SVM	15.333318
CT	12.165660
SFMV	24.294089
WFMV	27.922584

5. Conclusions

In this paper, a MCS using ANN, SVM and CT has been proposed to produce land use/cover maps from 0.82m pan-sharpened IKONOS satellite imagery. ANN was the most accurate individual classifier (91.72%), followed by SVM (89.60%) and CT (87.09%). A modified weighted majority voting scheme, WFMV, has been applied for combining the results obtained for individual classifiers. The architecture of the WFMV has been obtained through two steps: 1) weighting the decisions of the member classifiers; and 2) setting the parameters of the WFMV model using a 10-fold cross-validation. The proposed model has been tested and evaluated considering four different aspects: 1) overall accuracy; 2) class accuracy; 3) sensitivity to training sample size; and 4) computational complexity. The results showed an improvement of about 3.88% in the classification accuracy over the best individual classifier. The application of WFMV resulted in an average class accuracy of 95.38% which is 4.7 and 3.5% better than the best individual classifier and SFMV respectively. Another advantage of WFMV fusion is that the obtained errors are less variable and class accuracies of the different classes are almost equal. In terms of computational complexities, WFMV is still comparable with SFMV. Although ANN can perform as well as WFMV when large training samples are available, WFMV performs much better than ANN in the case of small training samples. The proposed WFMV approach can be practically extended to integrate any number of member classifiers. As future work, Deep Learning Based (DLA) approach should be incorporated into the

classification process to further improve the accuracy of the obtained results.

Acknowledgements

The author would like to acknowledge Giza Utility Data Centre, Giza Governorate, Egypt for the provision of the 0.82mpan-sharpened IKONOS image.

References

- Ahn, H., M.J. Fazzari, N. Lim, J.J. Chen and R.L.Kodell (2007). Classification by ensembles from random partitions of high dimensional data. *Computational Statistics and Data Analysis* 51(12), 6166–6179.
- Anderson, J.R., E.E. Hardy, J. T. Roach and R. E. Witmer (1976). A land use and land cover classification system for use with remote sensor data. *Geological Survey Professional Paper No. 964*, U.S. Government Printing Office, Washington DC, 28.
- Breiman, L., J. Friedman, R. Olshenand and C. J. Stone (eds) (1984). *Classification and regression trees*. Chapman & Hall, New York, 358 p.
- Cavalcanti, G., L. Oliveira, T. Moura and G. Carvalho (2016). Combining diversity measures for ensemble pruning. *Pattern Recognition Letters*, 74(2016), 38–45.
- Chen, C.F. (2000) Fuzzy training data for fuzzy supervised classification of remotely sensed images, In: *Proceedings of the 20th Asian Conference on Remote Sensing (ACRS 1999)*, 22-25 November 1999, Hong Kong Convention and Exhibition Centre, Venue, Hong Kong, China, pp. 460–465.
- Chen, Y., P. Peng Dou and X. Xiaojun Yang (2017). Improving Land Use/Cover classification with a multiple classifier system using Ada Boost integration technique. *Remote Sensing*, 9(10), 1055; doi: 10.3390/rs9101055.
- Clark labs (2012). Guide to IDRISI software [online]. Clark University. Available from: <http://www.clarklabs.org/applications/upload/land-change-modeler-IDRISI-focuspaper.pdf> [Accessed 23 October 2018].
- Coppin, P., I. Jonckheere, K. Nackaerts, B. Muys and E. Lambin (2010). Digital change detection methods in ecosystem monitoring: A review, *International Journal of Remote Sensing* 2010(25), 1565–1596.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, 2(4), 303–314.
- Dai, L. and C. Liu (2010). Multiple classifier combination for land cover classification of remote sensing image. In: *Proceedings of the 2010 2nd International Conference on Information Science and Engineering (ICISE)*, 4–6 December 2010 Hangzhou, China, pp. 3835–3839.
- Dietterich, T. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization, *Machine Learning*, 40, 139–157.

- Du, P., S. Liu, J. Xia and Y. Zhao (2013). Information fusion techniques for change detection from multi-temporal remote sensing images, *Information Fusion*, 14(1), 19–27.
- El-Melegy, M.T. and S.M. Ahmed (2007). Neural networks in multiple classifier systems for remote-sensing image classification. In: Nachtegaele, M., Vin der Weken, D., Kerre, E.E. and Philips, W., ed. *Soft Computing in Image Processing: Recent Advances*, Volume 210 of *Studies in Fuzziness and Soft Computing*, pp. 65–96, Springer, Germany.
- Fumera G., I. Pillai and F. Roli (2004). A two-stage classifier with reject option for text categorisation. In: Fred A., Caelli T.M., Duin R.P.W., Campilho A.C., de Ridder D., ed. *Structural, Syntactic, and Statistical Pattern Recognition (SSPR/SPR 2004)*. Lecture Notes in Computer Science, 3138. Springer, Berlin, Heidelberg, pp. 771–779.
- Ghimire, B., J. Rogan, V.F. Rodriguez-Galiano, P. Panday and N. Neeti (2012). An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience and Remote Sensing*, 49(5), 623–643.
- Ghosh, A. and B. Uma Shankar (2010). Neuro-fuzzy-combiner: an effective multiple classifier System. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 2(2), pp. 107–129.
- Ghosh, A., S. K. Meher and B.U. Shankar (2008). A novel fuzzy classifier based on product aggregation operator, *Pattern Recognition*, 41(3), 961–971.
- Han, M., X. Zhu and W. Yao (2012). Remote sensing image classification based on neural network ensemble algorithm, *Neurocomputing*, 78(1), 133–138.
- Haralick, R.M. (1976). The table look-up rule, *Communications in Statistics-Theory and Methods A5*, pp. 1163–1191.
- Hashem, S. and B. Schmeiser (1995). Improving model accuracy using optimal linear combinations of trained neural networks, *IEEE Transactions on Neural Networks*, 6(3), 792–794.
- Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation*, 2nd ed., Prentice Hall.
- Herrera, F. and J. L. Verdegay (1996). A Linguistic decision process in group decision making. *Group Decision Negotiation*, 5, pp. 165–176.
- Hirose, Y., M. Mori, Y. Akamatsu and Y. Li (2004). Vegetation cover mapping using hybrid analysis of IKONOS data. *ISPRS Archives*, XXXV (B7), 12–23 July 2004 Istanbul, Turkey.
- Ho, T.K., J.J. Hull and S.N. Srihari (1994). Decision combination in multiple classifier systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1), 66–75.
- Hsu, C.W., C.C. Chang and C.J. Lin (2009). A practical guide to support vector classification [online]. Department of Computer Science, National Taiwan University, Available from: <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf> [Accessed 23 October 2018].
- Hu, Y. (2000). *Handbook of neural network signal processing*, 1st ed., CRC Press, Inc., Boca Raton, FL, USA, 2000.
- Kavzoglu, T. and P.M. Mather (2003). The use of back propagating artificial neural networks in land cover classification. *International Journal of Remote Sensing*, 24(3), 4907–4938.
- Khoshgoftaar, T., J. Van Hulse and A. Napolitano (2011). Comparing boosting and bagging techniques with noisy and imbalanced data. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* (41), 552–568.
- Khosravi, I. and M. Mohammad-Beigi (2014). Multiple classifier systems for hyperspectral remote sensing data classification. *Journal of the Indian Society of Remote Sensing*, 42(2), 423–428.
- Kimura, F. and M. Shridhar (1991). Handwritten numerical recognition based on multiple algorithms, *Pattern Recognition*, 24(10), 969–983.
- Kumar, D.A. and S.K. Meher (2013). Multiple classifiers systems with granular neural networks. In: *Proceedings of the 2013 IEEE International Conference on Signal Processing, Computing and Control (ISPCCC)*, 26–28 September 2013 Solan, India, pp. 1–5.
- Kuncheva, L.I. (2000). *Fuzzy Classifier Design*, Springer-Verlag.
- Kuncheva, L.I. (2004). *Combining Pattern Classifiers: Methods and Algorithms*, Wiley-Interscience.
- Lillesand, T. and R. Kiefer (2004). *Remote sensing and image interpretation*. Fourth Edition, John Wiley & Sons, Inc., New York.
- Lu, D. and Q. Weng (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823–870.
- Lv, Y., P.F. Shi and Y.M. Zhao (2000). Voting principle for combination of multiple classifiers. *Journal of Shanghai Jiao Tong University*, 34(5), 680–684.
- Maulik, U. and D. Chakraborty (2010). A robust multiple classifier system for pixel classification of remote sensing images, *Fundamentalia Informaticae*, 101, 286–304.
- Mazurov, V.D., A.I. Krivonogov and V.S. Kazantsev (1987). Solving of optimization and identification problems by the committee methods. *Pattern Recognition*, 20, 371–378.
- Moustakidis, S., G. Mallinis, N. Koutsias and J.B. Theocharis (2012). SVM-based fuzzy decision trees for classification of high spatial resolution remote sensing images. *IEEE Transaction on Geoscience and Remote Sensing*, 50, 149–169.
- Platt, J. (1999). Fast training of support vector machines using sequential minimal optimization. In: B. Schölkopf,

- C. J. C. Burges, and A. J. Smola, ed. *Advances in Kernel Methods - Support Vector Learning*, Cambridge, MA.: MIT Press, pp. 185-208.
- Ponti-Jr, M. (2011). Combining Classifiers: from the creation of ensembles to the decision fusion. In: *Proceedings of the 2011 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutorials*, 28-30 August 2011, Alagoas, Brazil, pp. 1-10.
- Ponti-Jr. M. and J. P. Papa (2011). Improving accuracy and speed of optimum-path forest classifier using combination of disjoint training subsets. In: *Proceedings of the 10th International Workshop on Multiple Classifier Systems (MCS 2011) LNCS 6713*, 15-17 June 2011, Naples, Italy, pp. 237-248.
- Ranawana, R. and V. Palade (2006). Multi-classifier systems: review and a roadmap for developers. *International journal of hybrid intelligent systems*, 3(1), 35-61.
- Rogova, G. (1994). Combining the results of several neural network classifiers. *Neural Networks*, 7(5), 777-781.
- Russ, J.C. (2002). *The image processing handbook*, fourth edition. Boca Raton, FL: CRC Press.
- Saheb, B., k. Subba rao and S. Phani kumar (2013). A survey on voting algorithms used in safety critical systems. *International Journal of Engineering and Computer Science* ISSN: 2319-7242, 2(7), 2272-2275.
- Salah, M., J. Trinder and A. Shaker (2011). Performance evaluation of classification trees for building detection from aerial images and Lidar data: A Comparison of Classification Trees Models. *International Journal of Remote Sensing*, 32(20), 5757-5783.
- Salah, M., J. Trinder, A. Shaker, M. Hamed and A. Elsagheer (2010). Integrating multiple classifiers with fuzzy majority voting for improved land cover classification. In: *Papadoditis N., Pierrot-Deseilligny M., Mallet C., Tournaire O., ed. IAPRS, XXXVIII (3A)*, 1-3 September 2010, Saint-Mandé, France, pp. 7-12.
- Segrera, S. and M. Moreno (2005). Multi classifiers: applications, methods and architectures. In: *Proceedings of the International Workshop on Practical Applications of Agents and Multiagents Systems, IWPAAMS05*, pp. 263-271.
- Shannon, C. E. (editor) (1949). Reprinted 1998, *The mathematical theory of communication*, (Urbana, IL: University of Illinois Press), 324.
- Tso, B. and P.M. Mather (2001). *Classification methods for remotely sensed data*, Taylor and Francis, London.
- Tumer, K. and J. Ghosh (1996). Error correlation and error reduction in ensemble classifiers, *Connection Science*, 8(3), 385-204.
- Van der Linden, S., A. Rabe, A. Okujeni and P. Hostert (2009). *Image SVM classification, application manual: imageSVM version 2.0*.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer-Verlag, New York.
- Wemecke, K.D. (1992). A coupling procedure for the discrimination of mixed data, *Biometrics*, 48, 497-506.
- Wozniak, M., M. Graña and E. Corchado (2014). A survey of multiple classifier systems as hybrid systems. *Information Fusion*, 16 (2014), 3-17.
- Xu, L., A. Krzyzak and C.Y. Suen (1992). Methods of combining multiple classifiers and their applications to handwriting recognition, *IEEE Transactions on Systems, Man, and Cybernetics*, 22, 418-435.
- Yager, R.R. (1998). On ordered weighted averaging aggregation operators in multicriteria decision making, *IEEE Transactions on Systems, Man, and Cybernetics*, 18, 183-190.
- Yang, Z. (2007). An interval based attribute ranking technique. Unpublished report, ITT Visual Information Solutions.
- Zadeh, L.A. (1965). Fuzzy sets. *Information Control*, 8, 338-353.
- Zadeh, L.A. (1983). A computational approach to fuzzy quantifiers in natural languages, *Computers and Mathematics with Applications*, 9, 149-184.

Morphological investigation on degrading minor irrigation tanks: a case study in Hunsur taluk of Mysore district- Karnataka, India

Pradeep Raja K.P.* and Suresh Ramaswamyreddy

Department of Civil Engineering, BMS College of Engineering, Basavanagudi, Bengaluru, Karnataka, India

*Email: kppradeepraja@yahoo.co.in

(Received: July 12, 2018; in final form: May 15, 2019)

Abstract: The objective of this morphological investigation is to ascertain and study the main reasons for the degradation of Minor Irrigation (MI) tanks and focusses its analysis and discussions on the Bilikere and Halebidu tanks and their combined catchment (BKHB) - a part of Hunsur taluk of Mysore district, Karnataka, India which were found completely deteriorated and degraded after 2004. In this regard it is relevant to mention here that analysis of the daily rainfall data from 1975 to 2014 reveals that the average annual rainfall has remained normal whereas the mean daily intensity has decreased and the number of rainy days has increased. Drainage morphometry in relation to hydrology of BKHB catchment is very useful in understanding the reasons behind degradation of these MI tanks. The study has made use of CARTOSAT data in generating Digital Elevation Model (DEM) of BKHB catchment covering 44.67 km² which is a part of Lakshman Theertha river basin which is a sub-basin of Cauvery river in the semi-arid region of Mysore district. A comprehensive study of linear aspects reveals that the basin is dominated by lower order streams, the mean bifurcation ratio (R_b) of the BKHB catchment is 4.17; that of 3rd order Micro Watersheds (MWS) is 3.9 shows that the drainage pattern is not influenced by geological structures and length of overland flow (L_g) indicates flat terrain with less surface run-off and more infiltration. Similarly, the study of areal aspects like form factor, Gravelius index, shape factor, circularity ratio and elongation ratio indicates that the 3rd order MWS are elongated. Other aspects like drainage density, texture etc., shows that the catchment is highly permeable and dominated by 1st order streams. Relief aspects indicates low degree slope of the land form and resistant bed rocks in the terrain. The dendritic drainage pattern shows uniform bedrock terrain with insignificant faults and joints. A sum total of all the indices lead to the fact that the BKHB catchment is more permeable with high infiltration and less runoff working as a contributory factor towards degradation of lakes.

Keywords: Morphometry, CARTOSAT, Bilikere and Halebidu, MI Tanks, BKHB Catchment.

1. Introduction

Drainage basins are the primary units for any hydrological study. Precipitation, floods and droughts are the primary atmospheric activities which alters the landforms. The drainage basin acts as a funnel by collecting all the water within the area covered by the basin and channels it to a single point. The fluvial activities leave imprints on the landforms or drainage basins. The number, the size and the shapes of the stream channels are the evidences of these imprints and are exhibited in the surficial topography. The formation of the channels depends on the type of soil, geology and vegetation. Characterization and evaluation of the basin hydrology is an important factor in the study of surface runoff, ground water potential and management of environment. In this investigation, Bilikere and Halebidu (BKHB) catchment is delineated using SOI topomaps. GIS analysis is carried out using CARTOSAT DEM and Hydrology tool in Arc-GIS software. The combined catchment is further divided into 2 Sub-watersheds (SBW) namely Bilikere and Halebidu taking into consideration their proximity, interdependency being a cascade system and comprising of 13 micro watersheds (MWS). Prioritization of 3rd order MWS is considered because of the domination of lower order streams.

The importance of surface drainage networks and its development has been given attention from past several years using conventional methods (Horton, 1945;

Strahler, 1952; Strahler, 1956; Strahler, 1965). ASTER data integrated with GIS was used for the characterization of 3rd order watersheds where the lower order streams dominated the basin (Rama, 2014). GIS based methodology for the assessment of drainage morphometric parameters combined with Remote Sensing (RS) data is more suitable than the conventional methods (Pareta and Pareta, 2012). Morphometric analysis carried out in Najanagud taluk of Mysore district revealed that the elongated shape of the basin is due to the guiding effect of thrusting and faulting (Subhan and Rao, 2011). Landsat imageries and topographical maps were used to illustrate the hydrological behavior of a subtropical Andean basin in Argentina (Mesa, 2006). RS and GIS has been demonstrated to be an effective tool in the process of delineating drainage and morphometric study (Lone et.al, 2012). LISS-III +PAN merged image and SOI topomaps were used to study and analyze the sub-watersheds in the district of Tumkur, Karnataka, India and demonstrated that RS techniques are capable tool in morphometric analysis (Srinivasa et al., 2004). Bio-physical and hydrological features like population, rainfall etc., and land use/land cover of various watersheds can be comprehended by RS and GIS techniques (Ramachandra et al., 2014). The inter-relationships between landforms and land resources can be better realized for planning and management at a micro watershed level (Boobalan et.al., 2014). Disaster due to floods and other natural calamities could be properly managed using advanced satellite images and

technologies (Altaf et.al., 2013). A correlation between drainage density and stream frequency was calculated for different watersheds and found to be positive (Malik et.al., 2011). The work reported in this paper is intended to find out the hydrological behavior of the BKHB catchment using the morphometric parameters and consequently establish the reason behind the degradation of the water bodies in the study area.

2. Study area

Hunsur Taluk of Mysore district has seven Minor Irrigation (MI) tanks with overall Culturable Command Area (CCA) of 740 Hectares. Among seven lakes two major lakes; Bilikere lake and Halebidu lake have dried up completely since 2004. These lakes are situated between 12°21'10.163"N & 12°17'37.353"N latitude and 76°25'46.296"E & 76°31'4.167"E longitude as shown in figure 1 & 2. Both the tanks are rain fed waterbodies at an altitude of 695 m and 680 m Mean Sea Level (MSL) respectively located west of Mysore, Karnataka, India in a suburban area. The water was used for agriculture, horticulture, fish culture and domestic purposes. The combined catchment BKHB has an area of 44.67 km² with a highest elevation of 774 m and lowest elevation of 680 m above MSL. Geologically, the area comprises of granites, gneisses and charnockite rock stratum. The catchment is primarily dominated by agricultural land and major part of the land is cropland, sparse vegetation and poor soil cover. The soil types are red sandy soil, red loamy soil and deep black soil of varying thickness upto 6 m (Basavarajappa et al., 2014). Variation in rainfall leads to recurring drought and over usage of ground water which characterizes the study area.

3. Details of tanks

There are about 27 water bodies including ponds and lakes; ranging from 0.7 Hectare to 62 Hectare in the BKHB catchment. This catchment is a part of Lakshman Theertha River which is a sub-basin of Cauvery river. The Bilikere lake has a catchment area of 22.87 km², live capacity of 21 MCft (Million Cubic feet) with water spread area of 36.8 Hectares and a total physiographical area of 62.12 Hectares. The Halebidu lake has a catchment area of 21.80 km², live capacity of 18 MCft

with a water spread area of 30 Hectares and a total physiographical area of 58.33 Hectares.

4. Data used and methodology

Open series Survey of India topographical maps D43W7 & D43W11 (2005 edition) on 1:50,000 scale are used as the base maps. Morphometric analysis is carried using CARTOSAT-DEM; Version 1.1, R1, developed by ISRO, India, derived from the CARTOSAT-1, stereo payload launched in 2005, vertical accuracy is 8 m at 90% confidence and horizontal resolution of one arc second (approximately 30m). Landsat-5 TM and ETM images (Table 1) with spatial resolution of 30 m and 8-bit radiometric resolution downloaded from USGS were utilized to find the temporal changes in the waterbodies. The morphometric parameters were analyzed in three aspects; linear, areal and relief. Various parameters are derived using GIS tools and other empirical formulae (Ven, 1964) as presented in table 2.

5. Results and discussion

5.1 Linear aspects

Stream network is generated using Cartosat DEM for the BKHB catchment. Stream links and junctions characterize linear aspects of the catchment. Stream definition is calculated for a grid cell size of 50 pixels covering 4.67 Hectares of land (0.1% of catchment area) because of the small size of BKHB catchment; so that small streams (30 m length) can be identified. The linear aspects include stream order, stream length, mean stream length, stream length ratio, bifurcation ratio, length of the basin, length of overland flow and rho coefficient as listed in the tables 3 & 4.

5.1.1 Stream order (S_u)

According to Strahler (1952, 1957 and 1964) and Horton (1945), Stream order is a measure of the relative size of streams. In the study area 5th order stream is the highest order. There are 357 streams identified in the catchment out of which 283 - 1st order (79.3%), 57 - 2nd order (16%), 13 - 3rd order (3.6%), 3 - 4th order (0.8%) and 1 - 5th order (0.3%). It is also observed that overall 35% of 1st order streams directly contribute to the higher order streams.

Table 1: Details of satellite data

Sl.No	Name of the Satellite	Sensor	Date of acquisition	Path	Row
1	Landsat -5	MSS	05/12/1994	144	052
2	Landsat -5	MSS	06/11/1996	144	052
3	Landsat -5	MSS	11/09/1999	144	052
4	Landsat -5	MSS	16/12/2004	144	052
5	Landsat -5	MSS	31/12/2006	144	052
6	IRS-P6	LISS-III	03/11/2009	099	065

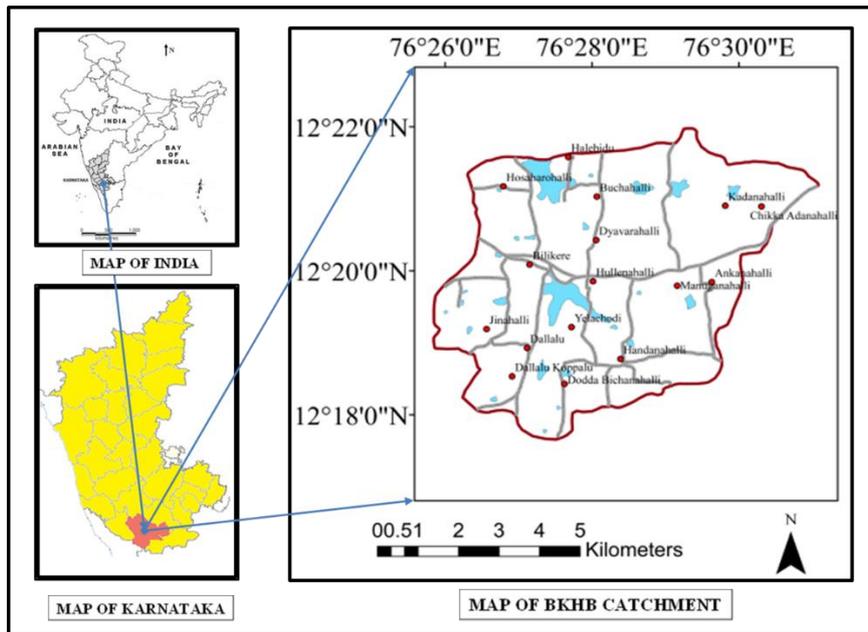


Figure 1: Location map of the study area

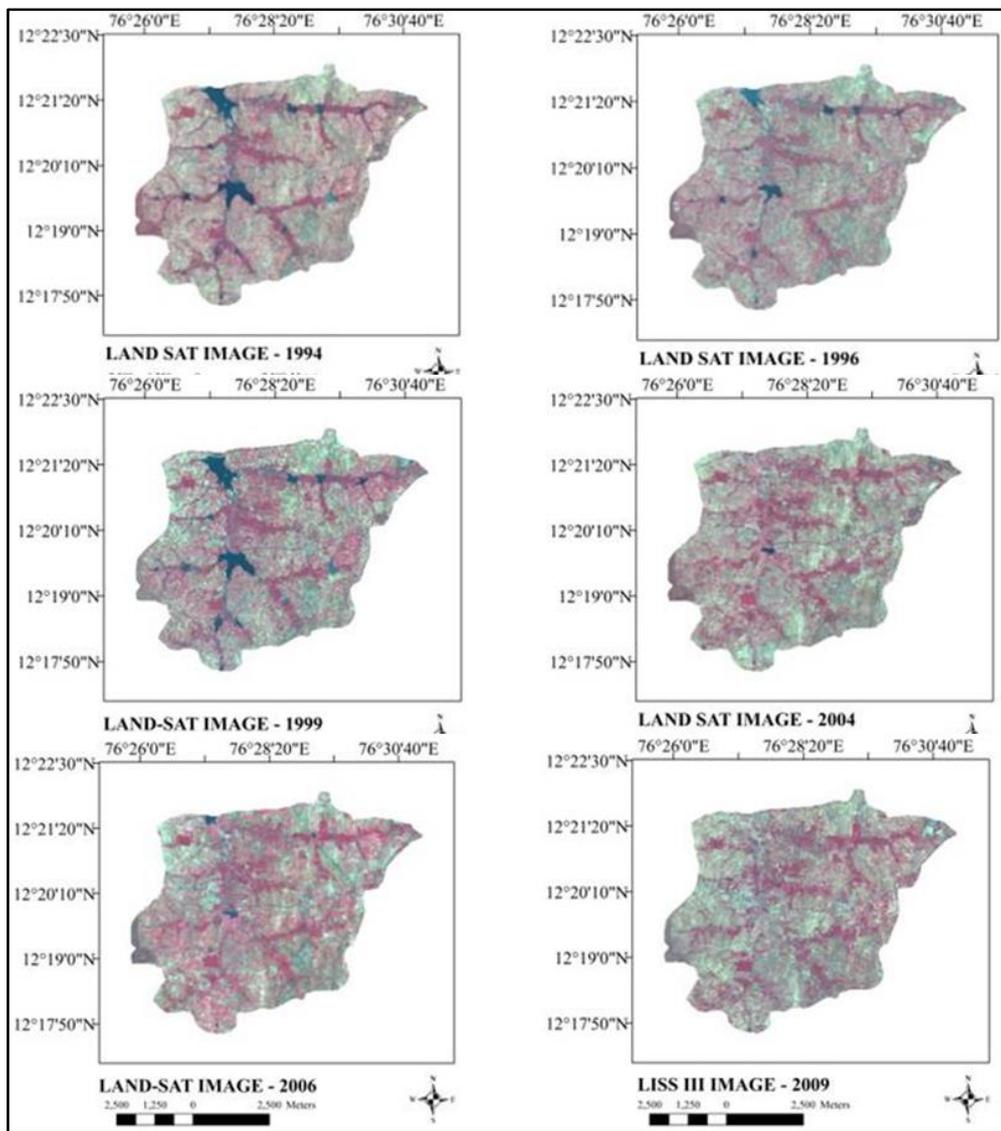


Figure 2: Temporal satellite images of the study area

Table 2 : Determination of morphological parameters

Sl. No	Particulars	Formulae	References
LINEAR PARAMETERS			
1	Stream Order (S_u)	Hierarchical ranking system	Strahler (1964)
2	Stream Length (L_u)	Law of stream lengths	Horton (1945)
3	Stream Number (N_u)	Law of stream Number	Horton (1945)
4	Mean Stream Length (L_{um})	$L_{um} = L_u/N_u$	Strahler (1964)
5	Stream Length ratio (L_{ur})	$L_{ur} = L_u/L_{u-1}$	Horton (1945)
6	Length of overland flow (L_g)	$L_g = 1/2D_d$	Horton (1932)
7	Bifurcation Ratio (R_b)	$R_b = N_u/N_{u+1}$	Strahler (1964)
8	RHO Coefficient (R_{ho})	$R_{ho} = L_{ur}/R_b$	Mesa (2006)
AREAL PARAMETERS			
9	Area (A)	Area calculated from GIS Tools	GIS
10	Perimeter (P)	Calculated using GIS Tools	GIS
11	Basin Length (L_b)	Calculated using GIS Tools	GIS
12	Form Factor (F_r)	$F_r = A/L_b^2$	Horton (1945)
13	Gravelius Index (G_i)	$G_i = 0.284P/\sqrt{A}$	Gravelius (1914)
14	Shape Factor (S_f)	$S_f = L_b^2/A$	Horton (1932)
15	Circularity Ratio (R_c)	$R_c = 12.57A/P^2$	Miller (1957)
16	Elongation Ratio (R_e)	$R_e = D/L_b = 1.128\sqrt{A}/L_b$	Schumm (1956)
17	Drainage Density (D_d)	$D_d = \sum L_u/A$	Horton (1945)
18	Drainage Texture (D_t)	$D_t = \sum N_u/P$	Smith (1950)
19	Texture Ratio (T_r)	$T_r = \sum N_1/P$	Schumm (1965)
20	Stream Frequency (F_s)	$F_s = \sum N_u/A$	Horton (1945)
21	Infiltration Number (I_f)	$I_f = F_s D_d$	Faniran (1968)
22	Constant of channel maintenance (C_m)	$C_m = 1/D_d$	Schumm(1956)
23	Lemniscate's ratio (K)	$K = 3.14L_b^2/4A$	Chorley (1957)
RELIEF PARAMETERS			
24	Watershed Relief (R)	$R = H - h$	Strahler (1952)
25	Relief ratio (R_f)	$R_f = R/L_b$	Schumm(1956)
26	Relative relief ratio (R_r)	$R_r = R/P$	Schumm(1956)
27	Slope gradient (S_g)	$S_g = R/L_b^2$	Gravelius (1914)
28	Ruggedness number (R_n)	$R_n = HD_d$	Strahler (1964)

5.1.2 Stream number (N_u)

The number of streams, of different orders in a given drainage basin tends closely to approximate an inverse geometric series in which first term is unity and the ratio is the bifurcation ratio (Horton 1945), (Figure 3).

5.1.3 Stream length (L_u)

Stream length is one of the most important hydrological features of the basin as it reveals the surface runoff behaviors. The length of the stream is a clue of the gradient of the catchment and of the degree of the basin. In general, streams are smaller in number, greater in length; in more permeable strata whereas larger in number, smaller in length in a steep well drained basin. The number of streams of various orders in a sub-watershed is counted and their lengths from mouth to drainage divide are measured. The stream length ' L_u ' has been computed based on the law proposed by Horton

(1945). The stream length of MWS & SWS is calculated using GIS tools and it is observed that the sum of stream length is minimum (1.23 km) in BK-ManuganaHalli MWS and maximum (14.84 km) in HB-ChikkaKadanahalli MWS.

5.1.4 Mean stream length (L_{um})

Mean Stream length is a dimensional property revealing the characteristic size of components of a drainage network and its contributing watershed surfaces (Strahler A N, 1964). It is obtained by dividing the total length of stream of an order by total number of segments in the order; mean stream length found to vary from 0.12 km to 0.34 km for I order; 0.09 km-1.02 km for II order; 0.26 km-3.16 km for III order streams. The results indicate no structural disturbance in the formation of streams in all the MWS of 3rd order.

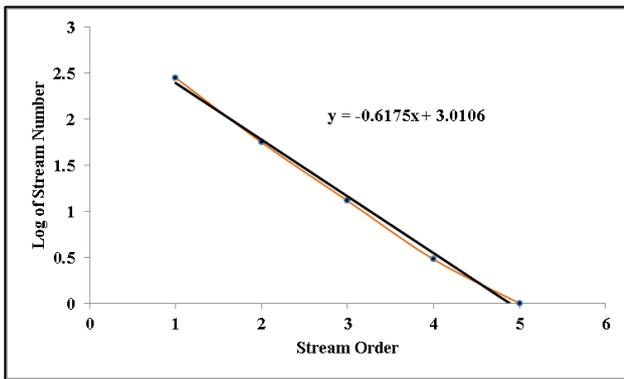


Figure 3: Relationship of Stream Numbers and Stream Order

5.1.5 Stream length ratio (L_{ur})

Horton (1945) states the ratio of the mean (L_{um}) of segments of order (S_u) to mean length of segments of the next lower order (L_{um-1}), which tends to be constant throughout the successive orders of a basin. His law of stream lengths refers that the mean stream lengths of stream segments of each of the successive orders of a watershed tend to approximate a direct geometric sequence in which the first term is the average length of segments of the first order. Changes of stream length ratio from one order to another order indicating their late youth stage of geomorphic development (Singh and Singh, 1997). The variation in the values of ' L_{ur} ' for different order streams within a basin indicates the permeability of the surface contributing to the stream network of the basin. It is apparent from the values of the BKHB catchment that the ' L_{ur} ' for the III order stream is higher (2.72) than the ' L_{ur} ' for the streams of the other orders. It is followed by the II order stream (2.24). This

shows the absorbent nature of the region through which the 2nd and 3rd order streams flow.

A graphical plot (Figure 4) between the order of the stream (x-axis) and log of mean stream lengths (y-axis) illustrates a direct relationship upto the 5th order of BKHB catchment. The equation for the trend line is given by

$$\log Y = -0.0861X + 0.5208 \quad 1$$

Where Y, is the mean stream length and X, is the order of the stream. The regression coefficient – 'R' squared value of 0.57 shows the statistical significance of linear regression fit and confirms the law of stream length ratio proposed by Horton.

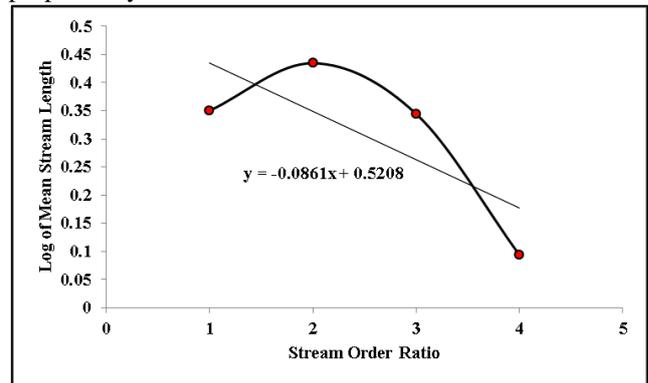


Figure 4: Relationship of stream length and stream order ratio

Table 3: Linear aspects of Sub-Watersheds

SBWS	Stream Number- N_u						Bifurcation Ratio- R_b				
	I	II	III	IV	V	Total	R_{b1}	R_{b2}	R_{b3}	R_{b4}	Mean
Bilikere	151	31	8	2	0	192	4.87	3.88	4.00		4.25
Halebidu	126	24	6	2	1	159	5.25	4.00	3.00	2.00	3.56
Combined-BKHB	283	57	13	3	1	357	4.96	4.38	4.33	3.00	4.17
Average	186.67	37.33	9.00	2.33	1.00	236.00	5.03	4.09	3.78	2.50	3.99

SBWS	Stream Length- L_u						Mean Stream Length- L_{um}				
	I	II	III	IV	V	Total	I	II	III	IV	V
Bilikere	37.67	17.10	9.47	7.54	0.00	71.78	0.25	0.55	1.18	3.77	
Halebidu	32.40	15.13	10.15	5.42	1.20	64.30	0.26	0.63	1.69	2.71	1.20
Combined-BKHB	71.40	32.20	19.98	10.18	4.22	137.98	0.25	0.56	1.54	3.39	4.22
Average	47.16	21.48	13.20	7.71	1.81	91.35	0.25	0.58	1.47	3.29	2.71

SBWS	Stream Length Ratio- L_{ur}					RHO-Coefficient-RHO					L_g
	L_{u2}/L_{u1}	L_{u3}/L_{u2}	L_{u4}/L_{u3}	L_{u5}/L_{u4}	Mean	R_{ho1}	R_{ho2}	R_{ho3}	R_{ho4}	Mean	
Bilikere	2.21	2.15	3.18	-	2.51	0.45	0.55	0.80		0.60	0.16
Halebidu	2.45	2.68	1.60	0.44	1.79	0.47	0.67	0.53	0.22	0.47	0.16
Combined-BKHB	2.24	2.72	2.21	1.24	2.10	0.45	0.62	0.51	0.41	0.50	0.16
Average	2.30	2.52	2.33	0.84	2.14	0.46	0.62	0.61	0.32	0.52	0.16

Table 4: Linear aspects of BKHB 3rd order MWS

MWS-No	MWS-Name	Stream Number- N_u				Bifurcation Ratio- R_b			
		I	II	III	Total	R_{b1}	R_{b2}	Mean	
1	HB_Chikka Kadanahalli	28.00	6.00	1.00	35.00	4.67	6.00	5.33	
2	HB_Hosa Harohalli	9.00	3.00	1.00	13.00	3.00	3.00	3.00	
3	HB_Kadanahalli	12.00	2.00	1.00	15.00	6.00	2.00	4.00	
4	HB_BiliKere	15.00	2.00	1.00	18.00	7.50	2.00	4.75	
5	HB_DyavaraHalli	17.00	3.00	1.00	21.00	5.67	3.00	4.33	
6	BK_AnkanaHalli	9.00	3.00	1.00	13.00	3.00	3.00	3.00	
7	BK_HullenaHalli	6.00	2.00	1.00	9.00	3.00	2.00	2.50	
8	BK_ManuganaHalli	5.00	2.00	1.00	8.00	2.50	2.00	2.25	
9	BK_JinaHalli	30.00	6.00	1.00	37.00	5.00	6.00	5.50	
10	BK_Dallalu	4.00	2.00	1.00	7.00	2.00	2.00	2.00	
11	BK_Dallalu Kappalu	28.00	4.00	1.00	33.00	7.00	4.00	5.50	
12	BK_Dodda BichanaHalli	15.00	4.00	1.00	20.00	3.75	4.00	3.88	
13	BK_HandanaHalli	22.00	5.00	1.00	28.00	4.40	5.00	4.70	
	Average	15.38	3.38	1.00	19.77	4.42	3.38	3.90	
MWS-No	MWS-Name	Stream Length- L_u				Mean stream Length- L_{um}			
		I	II	III	Total	I	II	III	Total
1	HB_Chikka Kadanahalli	7.76	4.79	2.29	14.84	0.28	0.80	2.29	3.37
2	HB_Hosa Harohalli	2.69	1.39	1.74	5.82	0.30	0.46	1.74	2.50
3	HB_Kadanahalli	2.82	1.23	1.42	5.47	0.24	0.62	1.42	2.27
4	HB_BiliKere	2.31	1.07	1.79	5.17	0.15	0.54	1.79	2.48
5	HB_DyavaraHalli	3.84	1.06	2.55	7.45	0.23	0.35	2.55	3.13
6	BK_AnkanaHalli	3.07	1.14	0.93	5.14	0.34	0.38	0.93	1.65
7	BK_HullenaHalli	1.61	1.14	0.26	3.01	0.27	0.57	0.26	1.10
8	BK_ManuganaHalli	0.58	0.28	0.37	1.23	0.12	0.14	0.37	0.63
9	BK_JinaHalli	7.60	3.63	3.16	14.39	0.25	0.61	3.16	4.02
10	BK_Dallalu	1.16	0.73	0.50	2.39	0.29	0.37	0.50	1.16
11	BK_Dallalu Kappalu	6.90	4.08	1.00	11.98	0.25	1.02	1.00	2.27
12	BK_Dodda BichanaHalli	3.53	0.35	1.32	5.20	0.24	0.09	1.32	1.64
13	BK_HandanaHalli	6.53	2.60	2.66	11.79	0.30	0.52	2.66	3.48
	Average	3.88	1.81	1.54	7.22	0.25	0.50	1.54	2.28
MWS-No	MWS-Name	Stream Length Ratio- L_{ur}			RHO-Coefficient- RHO			L_g	
		L_{u2}/L_{u1}	L_{u3}/L_{u2}	Mean	R_{ho1}	R_{ho2}	Mean		
1	HB_Chikka Kadanahalli	0.62	0.48	0.55	0.13	0.08	0.11	0.17	
2	HB_Hosa Harohalli	0.52	1.25	0.88	0.17	0.42	0.29	0.17	
3	HB_Kadanahalli	0.44	1.15	0.80	0.07	0.58	0.32	0.18	
4	HB_BiliKere	0.46	1.67	1.07	0.06	0.84	0.45	0.18	
5	HB_DyavaraHalli	0.28	2.41	1.34	0.05	0.80	0.43	0.18	
6	BK_AnkanaHalli	0.37	0.82	0.59	0.12	0.27	0.20	0.18	
7	BK_HullenaHalli	0.71	0.23	0.47	0.24	0.11	0.18	0.15	
8	BK_ManuganaHalli	0.48	1.32	0.90	0.19	0.66	0.43	0.15	
9	BK_JinaHalli	0.48	0.87	0.67	0.10	0.15	0.12	0.16	
10	BK_Dallalu	0.63	0.68	0.66	0.31	0.34	0.33	0.17	
11	BK_Dallalu Kappalu	0.59	0.25	0.42	0.08	0.06	0.07	0.17	
12	BK_Dodda BichanaHalli	0.10	3.77	1.94	0.03	0.94	0.48	0.16	
13	BK_HandanaHalli	0.40	1.02	0.71	0.09	0.20	0.15	0.16	
	Average	0.47	1.22	0.85	0.13	0.42	0.27	0.17	

5.1.6 Bifurcation ratio (R_b)

It is a dimensionless number denoting the ratio between the number of streams of one order ' N_u ' and those of the next higher order ' N_{u+1} ' in a drainage network. According to Horton (1945), bifurcation ratio indicates the relief and dissipation. Strahler (1957) demonstrates that bifurcation ratio shows a small range of variation for different regions except where the powerful geological control dominates. It is observed that the bifurcation ratio characteristically ranges between 3.0 and 5.0, for the basin in which geology is reasonably homogeneous and with no structural disturbances. The lower values of ' R_b ' indicate less structural disturbances. ' R_b ' is a measure of proneness to flooding. Higher the bifurcation ratio greater the probability of flooding. The observed mean value of ' R_b ' (3.38) of 3rd order MWS is less than 5 signifying that there is no structural disturbance on the drainage network. Out of thirteen, 3rd order MWS, only two 3rd order watersheds have a value more than 5 indicating structural control over the development of drainage network in these MWS. ' R_b ' value for the 2nd order streams vary from 2.0 to 7.5, four MWS among thirteen show the value of ' R_b ' greater than 5 indicating that the drainage network is influenced by the structural disturbances.

Figure 3 shows a graphical presentation between stream order as abscissa and log of stream number as ordinate. The best fitting regression equation for the linear relationship is given by

$$\log Y = -0.6175X + 3.0106 \quad 2$$

Where ' Y ' is the number of streams and ' X ' is the order of the stream. The regression coefficient – ' R ' squared value of 0.99 shows the statistical significance of linear regression fit and confirms the law of stream order proposed by Horton.

Overall the mean value of ' R_b ' for 2nd order stream is 4.42, for 3rd order it is 3.38, for BKHB catchment it is 4.17 suggesting that there is no structural disturbances in the formation of drainage network and there are about 4.2 times as many numbers of streams of any given order to that of the next higher order.

5.1.7 Length of overland flow (L_g)

(Horton R E, 1932) describes overland flow as the tendency of water to flow horizontally across soil surfaces when rainfall exceeds the capacity of infiltration. Length of overland flow is the length of the run of the surface water on the land surface before it is assigned into definite channels. Horton has taken L_g as the length equal to half the reciprocal of the drainage density. Higher the values of ' L_g ' lower the permeability and lower the value higher the permeability. The observed value of 3rd order streams of all MWS vary between 0.15-0.18 km/km² with a mean value of 0.17 km/km², BKHB catchment has a value of 0.16 km/km² indicating that the catchment is having a low slope, smaller flow paths, Less surface runoff and more infiltration.

5.1.8 Length of the basin (L_b)

Schumm defines the basin length as the longest dimension of the basin parallel to the principal drainage line (S. Schumm, 1956). The length of the basin for MWS varies from 0.8 km to 3.36 km, with a mean length of 2.28 km. The length of the BKHB Catchment found to be 7.2 km.

5.1.9 RHO coefficient

It defines the relationship between the drainage density and the development of the earth's features in the basin. It evaluates the storage capacity of the drainage network (Horton 1945). Higher values of RHO exhibit higher water storage during floods and essentially weaken the erosion effect during elevated discharge. The average R_{HO} coefficient of BKHB catchment is 0.5; that of MWS is 0.27 indicating less storage capacity of the channel network. **Areal aspects**

The physical characteristics of a catchment rely upon the size, shape, and gradient, drainage density of the watershed; size and length of the contributing streams. Areal aspects of a catchment of a particular order is defined as the total area projected upon a horizontal plane, contributing overland flow to the channel segment of that particular order including all branches of lower order. The size and shape of the catchment has an important relation to the drainage discharge characteristics. For instance, a circular catchment with a low bifurcation ratio can have a peak discharge compared to an elongated catchment with high bifurcation ratio may have a fluctuated flood discharge. Runoff, sediment processes and rate of discharge also depend heavily on the shape of the catchment. Parameters like form factor, circularity ratio, drainage density, compaction coefficient, elongation ratio etc., define the characteristics of a catchment. Table 5 shows the areal parameters for the MWS and BKHB catchment.

5.2.1 Form factor (F_f)

According to Horton (1932), Form Factor is the ratio of basin area to square of the basin length. It is a dimensionless number. The value of form factor would always be greater than 0.754 for a perfectly circular watershed. Smaller the value of form factor, more elongated will be the watershed. The mean value of the all 3rd order MWS found to be 0.44 indicating, elongated MWS.

5.2.2 Gravelius index (G_i)

Gravelius Index also known as compactness co-efficient (Gravelius, 1914), of a watershed is the ratio of perimeter of a watershed to circumference of circular area, which equals the area of the watershed. A circular catchment yields the shortest time of concentration before peak flow occurs in the basin. $G_i=1$ indicates that the catchment behaves like a circular catchment. $G_i>1$ shows that the basin deviates from circular to elongated and hence the time of concentration also increases. ' G_i ' of 3rd order MWS found to vary from 1.23-1.79; with a mean of 1.44, which indicates most of the MWS are elongated.

5.2.3 Shape factor (S_f)

It is the ratio of square of the basin length to the basin area (Horton R E, 1932) It is used to measure the degree of similarity of catchment shapes. The value of $S_f=1$ for a perfect square catchment, If $S_f>1$ then the catchment is elongated and If $S_f<1$ then the catchment is a circular one. The shape factor for all MWS varies from 1.43 to 3.65, indicating elongated shape of the MWS, ' S_f ' for BKHB catchment is 1.16 which suggests that the total catchment is slightly elongated.

5.2.4 Circularity ratio (R_c)

Circularity ratio is defined as the ratio of the basin area to the area of a circle having the same perimeter as the perimeter of the basin (Miller, 1957). It signifies the dissection stages of the study area with low, medium and high values, which represent youth, mature and old stages of the cycle of the tributary watershed of the region. According to Miller, If the value ranges between 0.4-0.5, It implies that the basin area is elongated, highly permeable and having a homogeneous lithology. The average ' R_c ' for MWS is 0.49 and that of BKHB is 0.38, which means the basin area is elongated, more infiltration, low discharge and the subsoil is highly absorbent.

5.2.5 Elongation ratio (R_e)

Schumm (1956) defines it as the ratio of the diameter of the circle having the same area of the basin to the maximum length of the basin. Mean ' R_e ' value of MWS in the study area is 0.74 indicating elongated shape with low relief. ' R_e ' for BKHB is 1.05; reveals that the basin area is typically low relief area.

5.2.6 Drainage density (D_d)

It is defined as the ratio of the total length of all the streams to the total area of the drainage basin and is a measure of catchment characteristics like infiltration, runoff and land use. Higher values of drainage densities indicate more runoff and lower value indicates more infiltration or vegetation. The mean value of ' D_d ' for the MWS is found to be 3 which indicate that the basin is medium textured. It also shows that the basin is highly permeable with low relief.

5.2.7 Drainage texture (D_t)

Horton (1945) defined drainage texture as the total number of stream segments of all order per perimeter of the basin. Drainage texture depends on the underlying lithology, infiltration capacity and relief aspect of the terrain. (Smith, 1950) has classified drainage texture into 5 different textures i.e., very coarse (<2), coarse (2 to 4), moderate (4 to 6), fine (6 to 8) and very fine (>8). The drainage texture for the MWS in the study area lies from 1.52 to 3.58 with a mean of 2.52 indicating very coarse to coarse texture; BKHB catchment has a value of 9.3 signifying that the 1st order streams dominate the basin.

5.2.8 Texture ratio (T_r)

According to (Schumm, 1965), it is expressed as the ratio between the first order streams and perimeter of the basin

and it depends on the underlying rocks, infiltration capacity and relief aspects of the terrain. It ranges between 0.89-2.86 for MWS with an average of 1.91, for BKHB it is 7.34; which reveals that the basin is controlled by first order streams.

5.2.9 Stream frequency (F_s)

According to Horton (1945) Stream frequency is referred to as number of streams per unit area of the catchment. It shows the relation of the lithology with the catchment. For 3rd order MWS it ranges from 6.63 to 21.05 per km²; with a mean value of 9.35 /km² suggesting that the catchment is moderately drained.

5.2.10 Infiltration number (I_f)

It is defined the infiltration number as the product of drainage density and stream frequency, Lower value of ' I_f ', higher the rate of infiltration and higher the value of ' I_f ' lower the infiltration rate (Faniran, 1968). The mean value of the ' I_f ' is found to be 28.39 for MWS; with a minimum of 29.69 and maximum of 68.14. The entire catchment has a value of 24.69 suggesting a moderate infiltration rate in the study area.

5.2.11 Constant of channel maintenance (C_m)

It has been characterized as the inverse of the drainage density by Schumm (1956). This constant provides an approximation of the extent of catchment required to maintain a unit length of the channel. The mean value of ' C_m ' for all MWS is 0.334 km²/km which indicates that about 0.334 km² of area is required to support one kilometer of the channel.

5.2.12 Lemniscate's ratio (K)

Lemniscate or pear shape, which defines the shape of the basin; it is more consistent with empirical reality than an ideal circular shape of a basin (Chorley, 1967). Chorley suggested that if the K value<0.6, then the shape of the basin is circular; if it is between 0.6-0.9 then it is oval; if K>0.9 elongated; Accordingly, the values of ' K ' for MWS varies between 1.12 to 2.87 with a mean of 1.90 indicating that the 3rd order basins are elongated. The ' K ' for the BKHB catchment is 0.91 which implies it is less elongated.

5.2.13 Drainage pattern

In morphological analysis, the flow pattern formed by the streams is called as drainage pattern. The pattern is controlled by geology of the area like dominating hard and soft rocks, slope of the terrain and topography of the catchment or land. In the present study, the pattern is identified as dendritic (Figures 5 & 6) which occurs in horizontal sedimentary or in intrusive igneous rocks with homogeneity of rock mass. This pattern is the most common form of drainage system. There are many twigs of streams which are then joined into the tributaries of the main stream or lakes. Dendritic pattern develops in a terrain which has uniform bedrock and where faulting and jointing are insignificant.

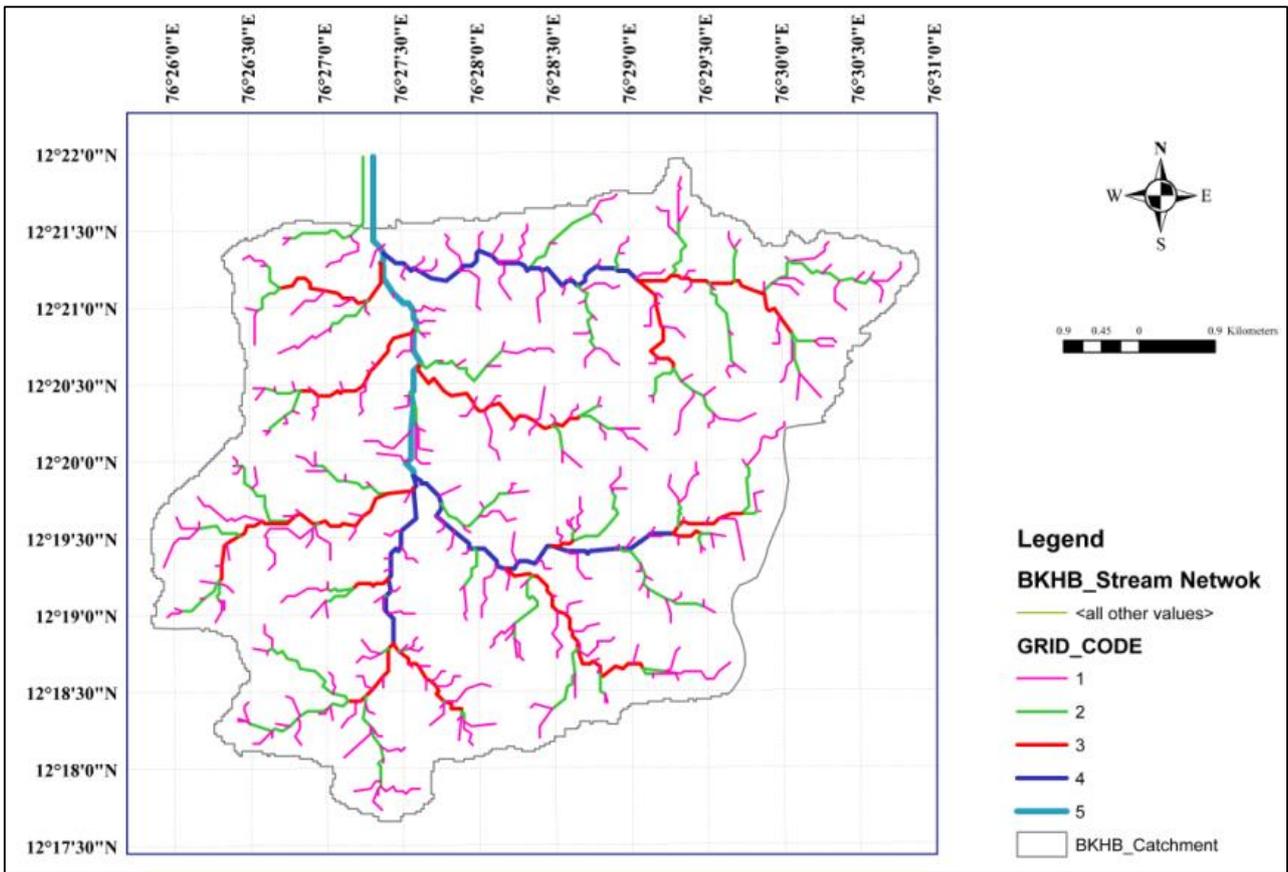


Figure 5: Drainage map of BKHB Catchment.

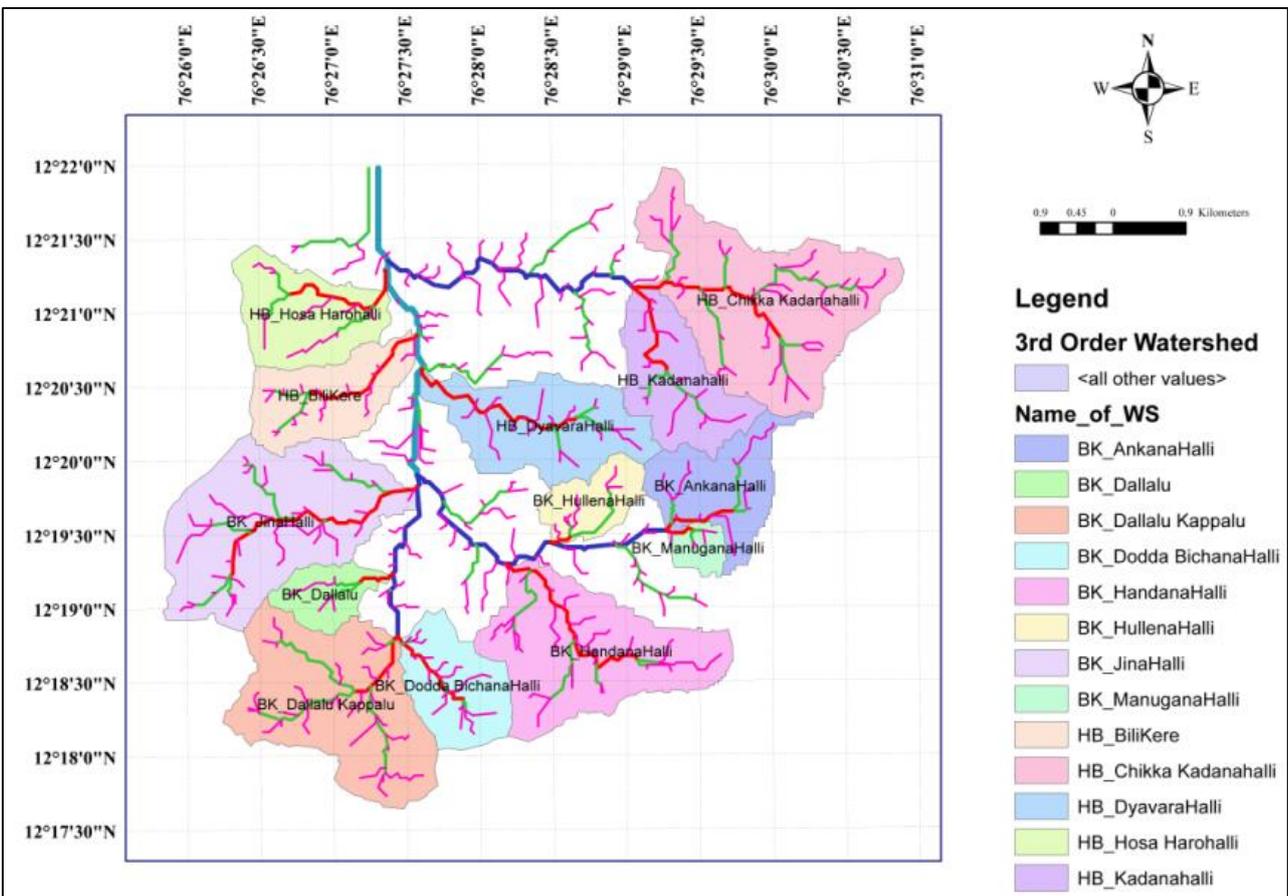


Figure 6: 3rd Order Micro watershed with Streams

Table 5: Areal parameters of BKHB- MWS & SBW

MWS-No	Name of MWS	A	P	L _b	F _f	G _i	S _f	R _c	R _e
1	HB_Chikka Kadanahalli	4.94	12.45	3.36	0.44	1.58	2.29	0.40	0.75
2	HB_Hosa Harohalli	1.96	6.85	1.90	0.54	1.38	1.84	0.52	0.83
3	HB_Kadanahalli	2.02	7.48	2.31	0.38	1.49	2.64	0.45	0.69
4	HB_BiliKere	1.83	6.67	2.33	0.34	1.39	2.97	0.52	0.65
5	HB_DyavaraHalli	2.63	9.10	3.10	0.27	1.58	3.65	0.40	0.59
6	BK_AnkanaHalli	1.82	8.57	1.95	0.48	1.79	2.09	0.31	0.78
7	BK_HullenaHalli	0.91	4.63	1.50	0.40	1.37	2.47	0.53	0.72
8	BK_ManuganaHalli	0.38	2.73	0.80	0.59	1.25	1.68	0.64	0.87
9	BK_JinaHalli	4.58	10.48	3.50	0.37	1.38	2.67	0.52	0.69
10	BK_Dallalu	0.80	4.51	1.60	0.31	1.42	3.20	0.49	0.63
11	BK_Dallalu Kappalu	4.03	10.03	2.40	0.70	1.41	1.43	0.50	0.94
12	BK_Dodda BichanaHalli	1.65	5.59	1.85	0.48	1.23	2.07	0.66	0.78
13	BK_HandanaHalli	3.83	10.28	3.03	0.42	1.48	2.40	0.46	0.73
	Mean	2.41	7.64	2.28	0.44	1.44	2.42	0.49	0.74
SWS	Name of SWS	Data of the sub-watershed							
BK	BK_Bilikere	22.87	32.10	4.30	1.21	1.91	0.83	0.27	1.24
HB	HB_Halebidu	21.80	29.16	6.50	0.49	1.80	2.03	0.31	0.79
BKHB	Combined Catchment	44.67	38.56	7.20	0.86	1.63	1.16	0.38	1.05
MWS-No	Name of MWS	Dd	Dt	T _r	F _s	I _f	C _m	K	
1	HB_Chikka Kadanahalli	3.00	2.81	2.25	7.09	21.28	0.333	1.79	
2	HB_Hosa Harohalli	2.97	1.90	1.31	6.63	19.69	0.337	1.45	
3	HB_Kadanahalli	2.71	2.01	1.60	7.43	20.11	0.369	2.07	
4	HB_BiliKere	2.83	2.70	2.25	9.84	27.79	0.354	2.33	
5	HB_DyavaraHalli	2.83	2.31	1.87	7.98	22.62	0.353	2.87	
6	BK_AnkanaHalli	2.82	1.52	1.05	7.14	20.17	0.354	1.64	
7	BK_HullenaHalli	3.31	1.95	1.30	9.89	32.71	0.302	1.94	
8	BK_ManuganaHalli	3.24	2.93	1.83	21.05	68.14	0.309	1.32	
9	BK_JinaHalli	3.14	3.53	2.86	8.08	25.38	0.318	2.10	
10	BK_Dallalu	2.99	1.55	0.89	8.75	26.14	0.335	2.51	
11	BK_Dallalu Kappalu	2.97	3.29	2.79	8.19	24.34	0.336	1.12	
12	BK_Dodda BichanaHalli	3.15	3.58	2.68	12.12	38.20	0.317	1.63	
13	BK_HandanaHalli	3.08	2.72	2.14	7.31	22.50	0.325	1.88	
	Mean	3.00	2.52	1.91	9.35	28.39	0.334	1.90	
SWS	Name of SWS	Data of the sub-watershed							
BK	BK_Bilikere	3.20	5.98	4.70	8.57	27.47	0.312	0.65	
HB	HB_Halebidu	3.08	5.45	4.32	7.62	23.50	0.324	1.59	
BKHB	Combined Catchment	3.09	9.26	7.34	7.99	24.69	0.324	0.91	

5.2 Relief aspects

It is the signature of the direction of flow. It helps in determining the degree of erosion in the catchment. It consists of watershed relief, relief ratio, relative relief, slope, slope gradient and ruggedness number. Figure 7 shows the DEM of 3rd order MWS and figure 8 shows the slope map of the BKHB catchment. The results are shown in table 6.

5.3.1 Watershed relief (R)

It is the difference in elevation between the highest point on the ridge line of the catchment to the mouth of the watershed. In the combined BKHB catchment; the highest elevation point is 774m and lowest point is 680m above MSL, it is extracted from Cartosat DEM.

5.3.2 Relief ratio (R_r)

Relief ratio is defined as the ratio of total relief to the basin length (Schumm 1956). Schumm correlated the relation between the hydrological characteristics of the

basin to the relief ratio. He observed that areas with low to moderate relief and slope are characterized by moderate value of relief ratios. Low value of relief ratios are mainly due to the resistant bed rocks of the basin and low degree of slope. The value of 'R_r' ranges from 0.018 to 0.049 with an average value of 0.025; 0.013 for BKHB indicating the terrain is of low relief for all MWS

5.3.3 Relative relief (R_r)

Relative relief is the ratio of total relief to the perimeter of the basin in percentage (Melton, 1958). R_{hp} for 3rd order MWS ranges between 0.47%-1.43%; with an average value of 0.76%, for BKHB catchment it is found to be 0.24%.

5.3.4 Slope

Slope analysis is a valuable criterion in geomorphic studies. The slope aspects are controlled by the terrain features and lithological elements like underlying bed rocks of different resistance. For the management of the

watershed and waterbodies; it is essential to analyze the gradient of the terrain.

5.3.5 Slope gradient (S_g)

Slope gradient is one of the factors which influence the drainage density.

5.3.6 Ruggedness Number (R_n)

Strahler’s ruggedness number is defined as the product of the basin relief and the stream density and usually combines slope with its length. Accordingly it is calculated for all the MWS and BKHB catchment.

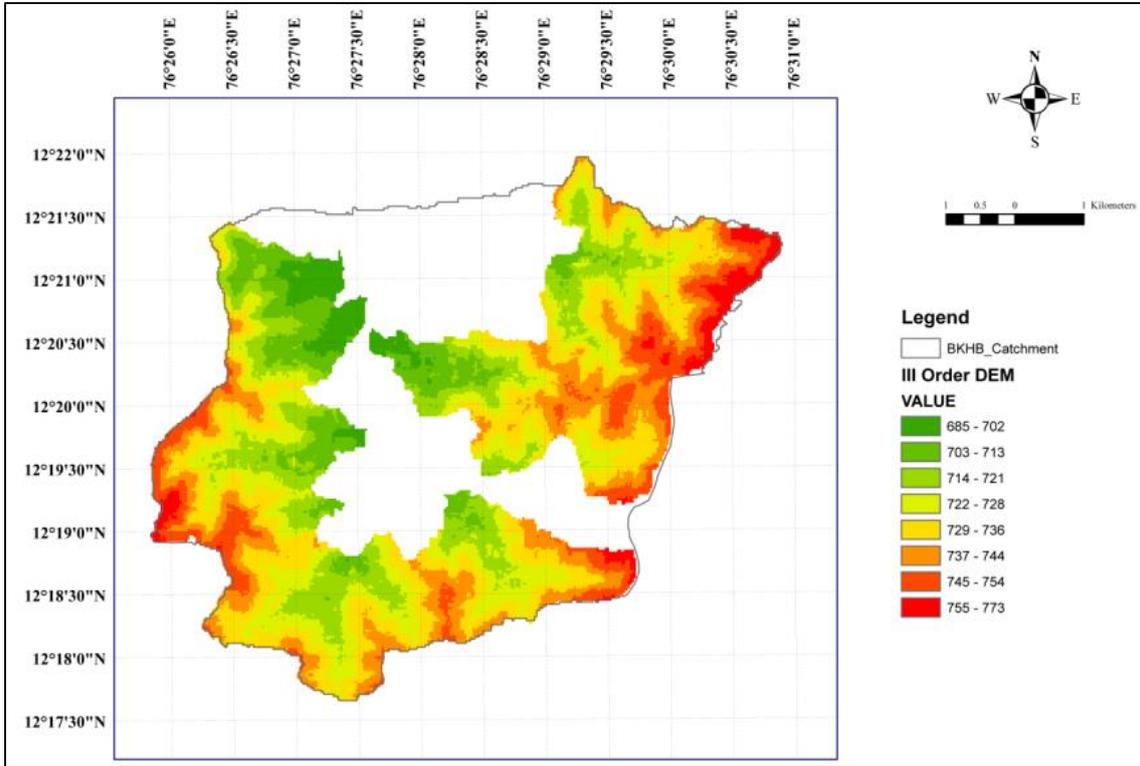


Figure 7: DEM of 3rd Order MWS of BKHB Catchment

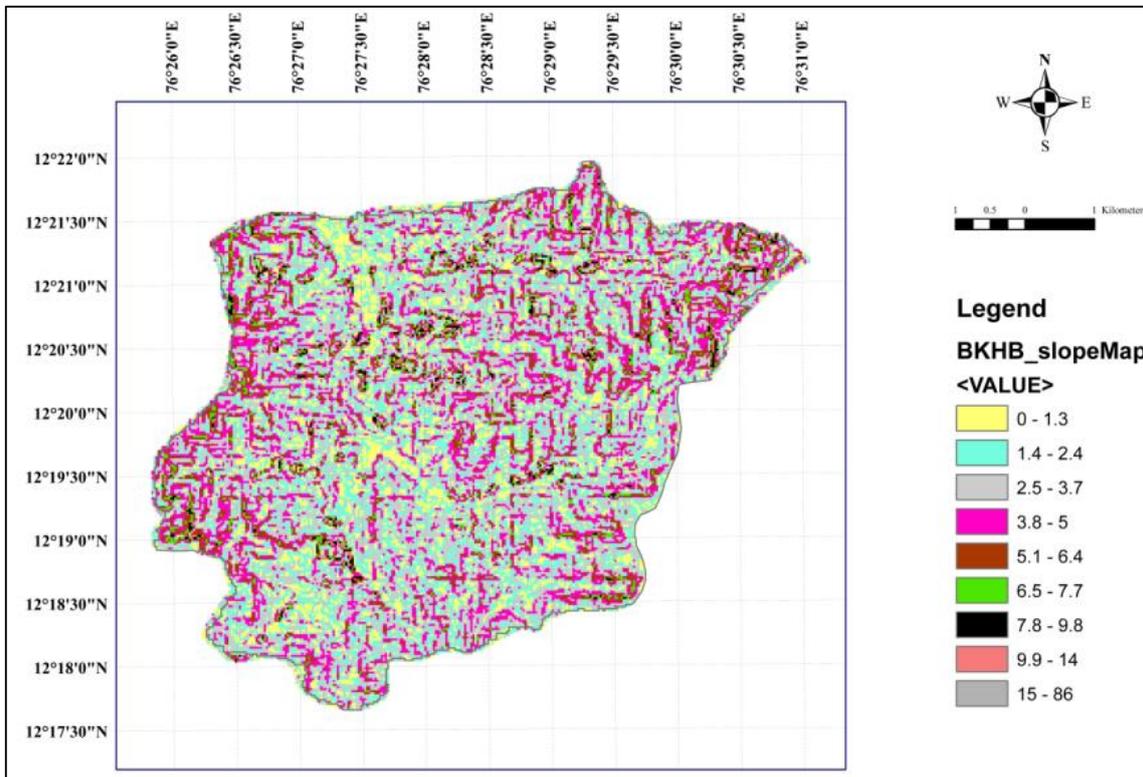


Figure 8: Slope map of BKHB Catchment

Table 6: Relief parameters of BKHB MWS & SBW

MWS-No	Name of MWS	H(m)	H(m)	R(m)	R _f	R _r (%)	S _g	R _n
1	HB_Chikka Kadanahalli	773	708	65	0.019	0.52	5.76	0.195
2	HB_Hosa Harohalli	742	685	57	0.030	0.83	15.79	0.169
3	HB_Kadanahalli	757	708	49	0.021	0.66	9.18	0.133
4	HB_BiliKere	749	689	60	0.026	0.90	11.05	0.170
5	HB_DyavaraHalli	748	691	57	0.018	0.63	5.93	0.161
6	BK_AnkanaHalli	767	707	60	0.031	0.70	15.78	0.169
7	BK_HullenaHalli	746	712	34	0.023	0.73	15.11	0.112
8	BK_ManuganaHalli	760	721	39	0.049	1.43	60.94	0.126
9	BK_JinaHalli	766	698	68	0.019	0.65	5.55	0.214
10	BK_Dallalu	754	704	50	0.031	1.11	19.53	0.149
11	BK_Dallalu Kappalu	756	709	47	0.020	0.47	8.16	0.140
12	BK_Dodda BichanaHalli	748	709	39	0.021	0.70	11.40	0.123
13	BK_HandanaHalli	761	707	54	0.018	0.53	5.88	0.166
	Mean				0.025	0.76	14.62	0.156
SWS	Name of SWS	Data of the sub-watershed						
BK	BK_Bilikere	767	695	72	0.017	0.22	3.89	0.231
HB	HB_Halebidu	774	680	94	0.014	0.32	2.22	0.290
BKHB	Combined Catchment	774	680	94	0.013	0.24	1.81	0.290

6 Conclusions

From the morphometric analysis of linear, areal and relief aspects; few major geomorphological conclusions are that the Stream order analysis shows that 79.3% of the catchment is dominated by first order streams. Further linear aspects like mean stream length, stream length ratio, bifurcation ratio, length of overland flow, length of the basin and Rho coefficient, as a subset of the major set of linear aspect reveals that no structural disturbance in the formation of streams of 3rd order, the catchment is absorbent in nature, low slope, smaller flow paths, less runoff and more infiltration with less storage capacity of channels. Areal parameters like form Factor, Gravelius index, Shape factor, Circularity index, Elongation ratio, Stream frequency, Drainage density and texture show that all MWS are elongated in shape with high permeability. MWS with Elongated shape and high permeability enhances time of concentration. Constant of channel maintenance result suggest that a minimum of 33 Hectares of catchment is required to maintain a channel of 1 km in the study area. The relief parameters reveal that the catchment is having a very low slope, less runoff and more infiltration. These conclusions may be used to supplement the rainfall-runoff analysis as a contributing factor for the degradation of lakes.

References

- Altaf, F., G. Meraj and S. A. Romshoo (2013). Morphometric analysis to infer hydrological behavior of Lidderwatershed, Western Himalaya, India. *Geography Journal*, 2013, 14.
- Basavarajappa, H. T., T. Parviz., M. C. Manjunatha and A. Balasubramanian (2014). Integration of geology, drainage and lineament on suitable landfill sites selection and environmental appraisal around Mysore city, Karnataka, India through remote sensing and. *Journal of Geomatics*, 8(1), 119–124.
- Boobalan, C., B. Gurugnanam and M. Mahesh (2014). Drainage morphometry and its influence on landform in a crystalline terrain, Sarabanga Sub-basin, Cauvery River, South India – a remote sensing and GIS approach. *International Journal of Advanced Geosciences and Technical Research*, 4, 765–781.
- Chorley, R. J., Donald, E. G., Malm and H. A. Pogorzelski (1957), A new standard for estimating drainage basin shape, *American Journal of Science*, 255, 138-141
- Faniran, A. (1968). The index of drainage intensity - A Provisional New Drainage Factor. *Australian Journal of Science*, 31, 328–330.
- Gravelius, H. (1914). *Grundrifi der gesamten Gewissserkunde Band I: Flufikunde Compendium of Hydrology*, 1, Goschen, Berlin.
- Horton, R. E. (1932). Drainage basin characteristics. *American Geophysical Union*, 13, 350–361.
- Horton, R. E. (1945). Erosional development of streams and their Drainage Basins. *Bulletin of the Geological Society of America*, 56, 275–370.
- Lone, M., D. Nagaraju., G. Mahadeva swamy., Lakshamma and S. Siddalingamurthy (2012). Morphometric Analysis of HeggadadevanaKote(H D Kote) Taluk, Mysore District, Karnataka, India, 5(2), 288–293.
- Malik, M. I., M. S. Bhat and N. A. Kuchay (2011). Watershed based drainage morphometric analysis of Lidder catchment in Kashmir valley using geographical information system. *Recent Research in Science and Technology*, 3(4), 118–126.

- Melton, M. A. (1958). Geometric properties of mature drainage systems and their representation in an E4 phase space. *The Journal of Geology*, 66(1), 35–54.
- Mesa, L. M. (2006). Morphometric analysis of a subtropical Andean basin (Tucuman, Argentina). *Environmental Geology*, 50, 1235–1242.
- Miller, V. C. (1957). A quantitative geomorphic study of drainage basin characteristics in the Clinch Mountain area, Virginia and Tennessee. *The Journal of Geology*, 65(1), 112–113.
- Pareta, K. and U. Pareta (2012). Quantitative geomorphological analysis of a watershed of Ravi River basin, H.P. India. *International Journal of Remote Sensing and GIS*, 1(1), 41–56.
- Rama, V. A. (2014). Drainage basin analysis for characterization of 3rd order watersheds using Geographic Information System (GIS) and ASTER data. *Journal of Geomatics*, 8(2), 200–210.
- Ramachandra, T. V., N. Nagar, S. Vinay and B. H. Aithal (2014). Modelling hydrologic regime of Lakshmanatirtha watershed, Cauvery river. 2014 IEEE Global Humanitarian Technology Conference - South Asia Satellite, GHTC-SAS 2014, 64–71.
- Schumm, S. (1956). Evolution of drainage systems and slopes in badland at Perth Amboy, New Jersey. *Geological Society of America Bulletin*, 67(5), 597–646.
- Schumm, S. A. (1965). *Geomorphic research: Applications to erosion control in New Zealand: Soil and Water Conservation and Rivers Control Council*, 1, 21–24.
- Smith, K. G. (1950). Standards for grading texture of Erosional Topography. *American Journal of Science*, 248, 655–668.
- Singh, S. and M. C. Singh (1997). Morphometric analysis of Kanhar river basin. *National Geographical Journal of India*.
- Srinivasa V. S., S. Govindaiah and H. H. Gowda (2004). Morphometric analysis of sub-watersheds in the Pavagada area of Tumkur district, South India using remote sensing and GIS techniques. *Journal of the Indian Society of Remote Sensing*, 32(4), 351–362.
- Strahler, A. N. (1952). Hypsometric Analysis of erosional topography. *Bulletin of the Geological Society of America*, 63, 1117–1142.
- Strahler, A. N. (1956). Quantitative slope analysis. *Bulletin of the Geological Society of America*, 67, 571–596.
- Strahler, A. N. (1964). Quantitative geomorphology of drainage basin and channel network. In *Hand Book of Applied Hydrology*, (39–76).
- Strahler, A. N. (1965). Quantitative geomorphology of drainage basin and channel Networks. *Hand book of applied Hydrology*.
- Subhan, L. M and K. Rao (2011). Morphometric analysis of Nanjangud taluk, Mysore District, Karnataka, India, using GIS Techniques. *International Journal of Geomatics and Geosciences*, 1(4), 721–734.
- Ven, T.C (1964). *Handbook of applied hydrology*. USA: McGraw-Hill.

Online geospatial transaction system using open source libraries

Sajeevan G.*, Danish Shaikh, Mugdha Magare and Ashish Kuvelkar
Centre for Development of Advanced Computing, C-DAC Innovation Park, Pune 411008, India
*Email: sajeevan@cdac.in

(Received: Mar 01, 2019; in final form: June 09, 2019)

Abstract: Geospatial Information System (GIS) plays a very significant role in our daily life. In India, GIS establishments have massive databases of geospatial data. However, for appropriate planning and development activities, the digital vector data requires improvement in spatial accuracy. Inaccuracies in geospatial data exist due to various reasons such as anomalies in digital conversion processes. Under *Pradhan Mantri Gram Sadak Yojana* (PMGSY) National GIS, a reliable national level geospatial database is created. Nevertheless, achieving the spatial accuracy for so many spatial layers in short duration is a herculean task. A web-based geospatial data editing application is, therefore, designed and developed using open source software, which could initially be implemented for PMGSY. The application can access base maps from external sources like the Indian Space Research Organisation's (ISRO) Bhuvan and OpenStreetMap; the existing spatial features can be modified and new features be added. With the right data and software, geospatial data services can be used by authorised users for other projects. In future, rather than hunting for reliable digital spatial data, the officials can focus on analysis, planning, modeling, etc. The architectural framework and the application can be extended to other national GIS applications or to a national geospatial framework.

Keywords: geospatial transaction, Web GIS, WFS, open source, geospatial framework, PMGSY

1. Introduction

Geographic Information System (GIS) is enriched with spatial data that help in analyzing the associated features on spatial relations, geometry and visualization as dynamic maps. Reasonably accurate geospatial data is essential for planning, development and timely decision-making activities by the government or by anyone dealing with spatial data. Inaccurate geospatial data severely affect planning and development activities and may even result in disputes. One of the major reasons for GIS project failure is the unavailability of reasonably accurate data (Sajeevan, 2017a). The applications of spatially accurate geospatial data are innumerable; be it finding nearest election polling station, identifying precise locations of military targets, number of households and their spatial distribution in a village, locating a ship and its trajectory, coverage of forests along a road, optimum distribution of road network for handling smooth vehicle movement, emergency response, etc. The geospatial data gaps need to be filled at the earliest to cope up with the fast pace of nation building.

On analyzing completed GIS projects, it was observed that the spatial features such as district boundary, road and settlements are displaced on the digital map (Sajeevan, 2017b). It is crucial that a feature is placed on the map at the same location where it is spatially located on the ground. Looking at the trend, it may be observed that a good part of a GIS project time is spent on creating digital base data rather than getting into meaningful spatial analysis. While certain projects struggle to get reasonably accurate digital spatial data, many private agencies own and operate with relatively higher accuracy geospatial data, such as the data for vehicle navigation systems (Sajeevan, 2017b).

The major reason for spatial inaccuracy in digital data is due to the incorrect procedures adopted for data conversion process. Overlaying spatial data from different sources clearly reveals any positional shifts, which would have crept in due to inappropriate data conversion process. Figure 1 shows visible shift between satellite data and vector layer overlays on ISRO's Bhuvan (Bhuvan, 2017). Regarding Bhuvan, our experiments show that the satellite data is placed with pretty good accuracy compared with locations from Global Positioning System (GPS) and Google maps data. The inaccuracy in vector layers is assumed to be due to digital conversion issues of data that are compiled under different projects by different agencies. Similar inaccuracies are observed on other Web GIS applications as well. The spatial mismatch of data from different sources is a common trouble faced by the GIS industry (Sajeevan, 2005). These are some of the constraints in geospatial data sharing. To address the national issues described earlier, there is a crucial need of a GIS application for spatial data correction. Online Geospatial Transaction System is, therefore, designed and developed using open source libraries. It is proposed to be implemented in PMGSY.

In the past, pilot implementation of GIS was carried out for a few states, wherein GIS enabled Road Information Management and Monitoring System (GRIMMS) was developed for PMGSY (Sajeevan et.al, 2013, 2012, 2007, 2006ab; Sajeevan 2012). However, the spatial accuracy of the GIS data was poor and the spatial shift was similar to those shown in figure 8 or worse than that. Although such spatial data may be used in isolation for limited GIS analysis; they will not be of meaningful use for geospatial data interoperability and related spatial analysis.

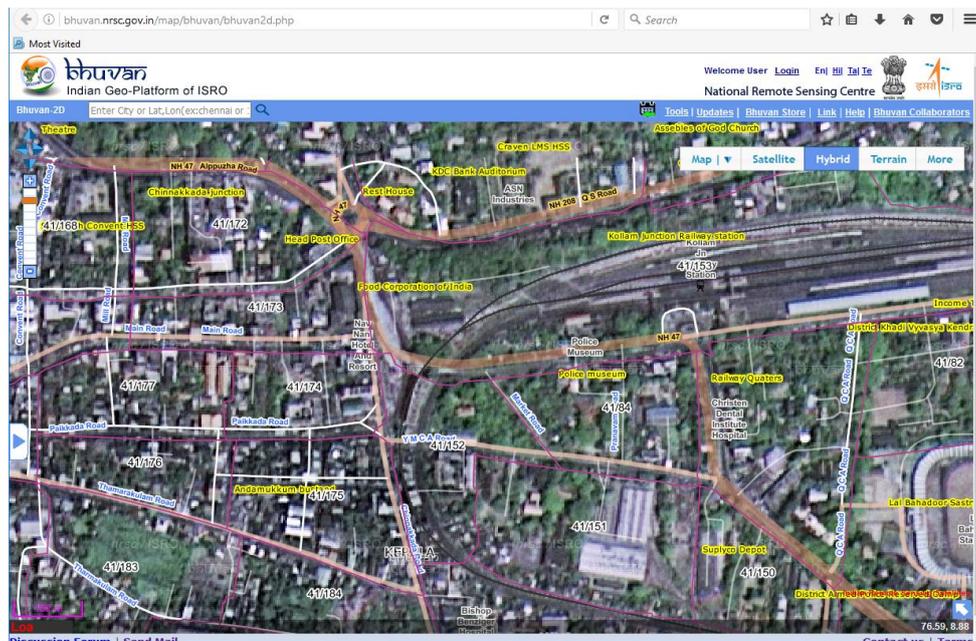


Figure 1: Spatial shift of vector layers on Bhuvan (accessed and screen captured on 22.02.2017)

2. About PMGSY National GIS

Currently, the project titled 'National implementation of Web based Geographic Information System for *Pradhan Mantri Gram Sadak Yojana*' is being executed for National Rural Infrastructure Development Agency (NRIDA), Ministry of Rural Development, Government of India in collaboration with Centre for Development of Advanced Computing (C-DAC) and all the states. The project is commonly referred as PMGSY National GIS.

Conforming to the PMGSY National GIS guidelines (PMGSY, 2016), all the states created digital geospatial data. The geospatial data creation process deviated from the traditional way of digitization to ensure position accuracy of features. The base data and the enforcing boundaries were discarded. Along with that, un-projected latitude-longitude coordinate reading on WGS84 datum was enforced. Spatial accuracy of any location on a map highly depends on how the appropriate projection parameters are used (Sajeewan, 2005). If the datum is changed, any selected location can have different geodetic coordinates (Sajeewan, 2008). Geospatial features could also be directly captured using GPS receivers rather than losing spatial accuracy through map projection conversions. By following the GIS guideline, the traditional map mosaicing task was also avoided. It was a herculean task to create national level geospatial data and correct with respect to satellite data services of Bhuvan. This enables users to access the precise feature locations on field with reference to GPS receivers either on mobile or other ways. The States departments dealing with PMGSY now have the reliable spatial data that can be used for further planning and development.

The GIS application developed under PMGSY National GIS is named Geospatial Rural Road Information System (GRRIS, <http://pmgsy-grris.nic.in>). GRRIS handles only the spatial data component. All other information about the features are accessed live from Online Management,

Monitoring and Accounting System (OMMAS, <http://omms.nic.in>) and then reports/maps are generated dynamically. OMMAS and GIS data updating and management were carried out independently, which ensured the success and sustainability of the GIS application. The Web GIS application enables effective planning, management and monitoring of PMGSY using geospatial technologies by the concerned officials. As the application is open to public, online dissemination of geospatial information to general public ensure transparency of processes and increased accountability. The Web GIS application was developed over a commercial GIS engine that lacked spatial data editing function. This constraint also led to the development of online geospatial data editing application.

3. Objective

Geospatial data has the location information of the feature geometry represented in geographic coordinates. The geospatial features are represented in map as point, line and polygon. The geospatial data types, structures and transactions are complex. Spatial features can be retrieved across the Web using Open Geospatial Consortium (OGC) compliant Web Feature Service (WFS) and the features can be created, deleted or updated using transactional Web Feature Service (WFS-T) (OGC, 2017). With the following objectives, online geospatial data editing application was developed using open source software libraries.

- Using a web browser, user should be able access, add, delete or update spatial features
- A generic architecture
- Design and development of a scalable web-based GIS application
- Deployment of the application
- Geospatial data service

The objective includes designing and developing a generic web based geospatial data editing application that can satisfy the specific GIS requirements of PMGSY.

4. Methodology

The methodology used for developing the Online Geospatial Transaction System is explained subsequently. The application is developed under Microsoft .NET environment utilising the following open source software:

- Database server and spatial extension
- GIS server
- GIS SDK

4.1 Geospatial DBMS

Geospatial data may be saved in intrinsic GIS data formats or in database servers with spatial extension. Examples of open source database servers are PostGIS (spatial extensions for PostgreSQL), MySQL Spatial (spatial extensions for MySQL), SpatialLite (spatial extensions for SQLite database). Examples of commercial database servers are Oracle Spatial (spatial extensions for Oracle database) and SQL Server Spatial (spatial extensions for Microsoft SQL Server).

PMGSY National GIS required spatial in shape file format (.shp). The spatial data of Odisha state is used for the initial analysis and testing of the online geospatial data editing application. The GIS layers are also kept in PostGIS. PostGIS provides spatial objects for the PostgreSQL database, allowing storage and query of information about location and mapping (PostGIS, 2017). The aim of storing the layers in both the formats is for analyzing the response time for geospatial data retrieval and transaction from those formats.

4.2 GIS Server

There exist a number of GIS servers, which provide maps on request. GIS servers have the capability to read data from GIS data files, as well as from database servers. While the GIS servers may generate and provide user requested map in standard image formats, the geospatial features are transferred using GML that is an extension of XML. Examples of open source GIS server are GeoServer and MapServer; and commercial servers ArcGIS Server and ERDAS Apollo. Designed for interoperability, GeoServer publishes data from any major spatial data source using open standards (GeoServer, 2017). Powerful and having rich functions, we have used Geoserver as GIS server.

4.3 Open Source GIS Library / SDK

Apart from generic development environment, additional libraries/SDK are required to perform GIS functions. The client site libraries are generally JavaScript, but, the Server-side libraries are of many types. The well-known commercial GIS servers like ArcGIS Server and MapXtreme have .Net and Java SDKs. OpenLayers, Leaflet and Cesium are examples of open source libraries. We have used the well-known OpenLayers library, which

is having rich support for GIS functions. OpenLayers helps to integrate dynamic map in any web page and can display map tiles, vector data and markers loaded from any source. OpenLayers has been developed to further the use of geographic information of all kinds (OpenLayers, 2017). Web clients require OpenLayers JavaScript library, which can send OGC complaint request to the map server.

4.4 Attribute Database

The attribute information about each feature in geospatial database could be accessed from diverse sources. It is advisable to maintain the spatial data separately and then link the spatial data to various databases online as per the requirements of the software application (Sajeevan, 2012). For PMGSY, all that attribute information is available in Microsoft SQL Server database of OMMAS. The OMMAS database is updated continuously by all the states. These data need to be accessed live and then make appropriate joins to the geospatial data. To achieve this, Foreign Data Wrappers in PostgreSQL can be used.

4.5 Multilingual Support

In the Indian context, the multilingual support of user interface is an important utility. GIS applications report many issues while supporting Indian languages. The language of names on the map may be set through GeoServer, however, the existing Indian language data may need to be converted to a uniform encoding. The multilingual display is configured at GeoServer level so that any software that is able to read OGC complaint services can make use of the multilingual map. Regarding OMMAS, names of the states, districts and blocks are currently available in Hindi language. Figure 2 shows the block map along with Hindi names as Web service.

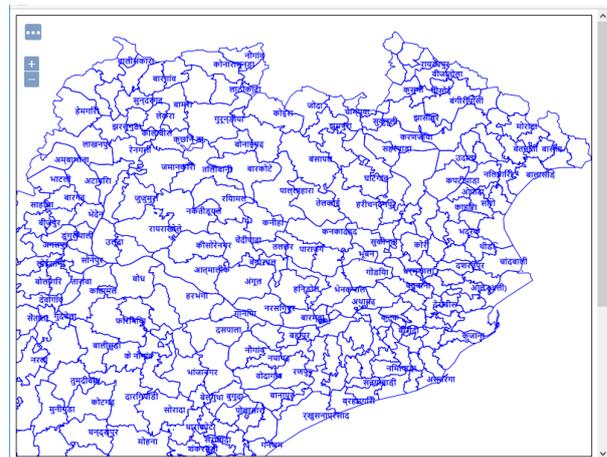


Figure 2: Block map service with name in Hindi

4.6 System Architecture

The architecture of the application is shown in figure 3. The clients should support JavaScript and need to download the required utility from the server. The External Map Servers can be Bhuvan or Google that can provide map services as WFS, WMS, etc. The client can also access these services directly from the external servers or can access through the local GIS server / Web Server. Internet Information Services (IIS) is used as Web server.

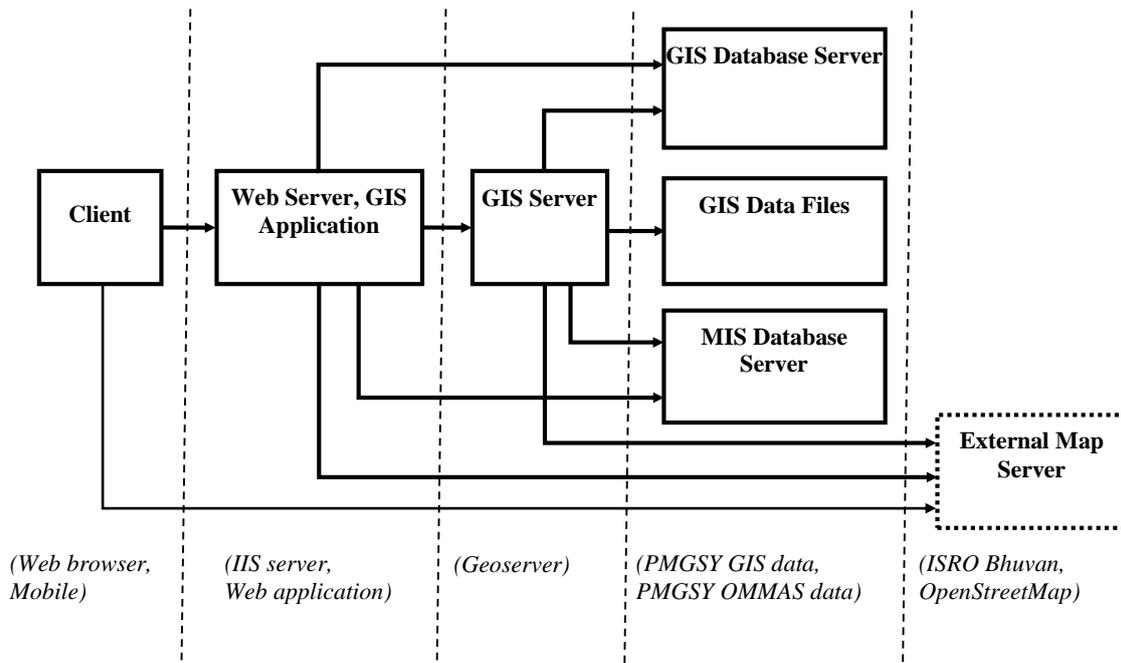


Figure 3: The system architecture

5. The Web GIS Application

With flexible architecture, design and development, a pilot of online geospatial editing application made for PMGSY. It is developed in Microsoft .NET environment. Experiments have been conducted on Apache JMeter to study the map access time by GeoServer from PostGIS and shape file data. The architecture and design of the application on open source software could become the backbone for national geospatial data programs. On successful login (role based) user would be directed to the page along with authorized functional features and spatial extent (Figure. 4).

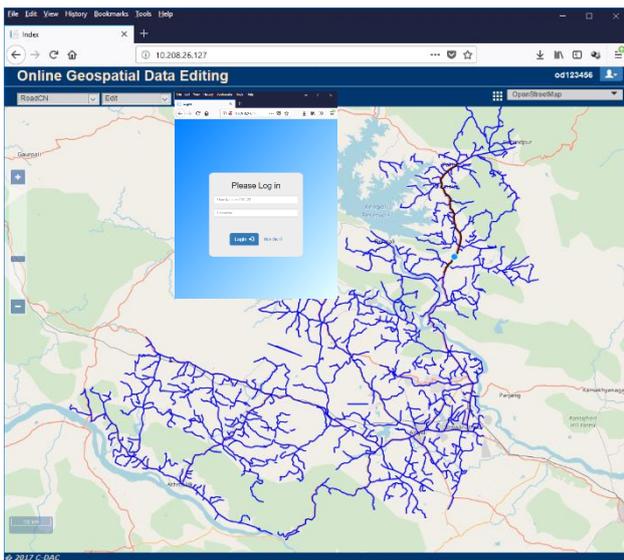


Figure 4: Edit layers on successful login

The user has the options of selecting a layer to edit, and if required, selecting a base map. User can add, edit and delete features of the selected vector layer. The user login

helps assigning/restricting user activity to a selected region; for example, user updating can be limited to a block or a district. The user can select any base map that can be from map services of internal GIS layers or from external sources like Bhuvan, OpenStreetMap and Google map. Figure 5 shows base map as WMS from Bhuvan. The WMS map is rendering of satellite data and the resolution depends on the zoom level.

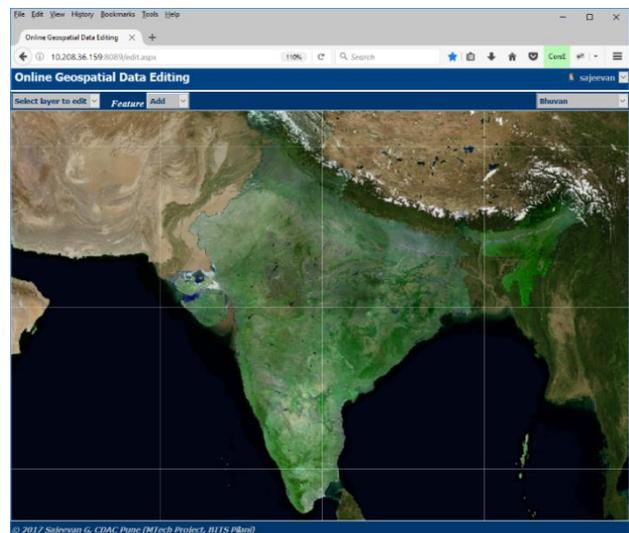


Figure 5: WMS from ISRO's Bhuvan

On selecting a layer to edit, the geometry of the features to be added would change automatically as per the layer geometry. For example, if block layer is selected for editing, feature geometry will be changed to polygon automatically. These capabilities ensure reducing the cognitive load on users and increasing the usability of the application. Figure 6 shows adding a new feature in block layer (for demonstration). The modifications made on the layer can be saved automatically or manually on the server.

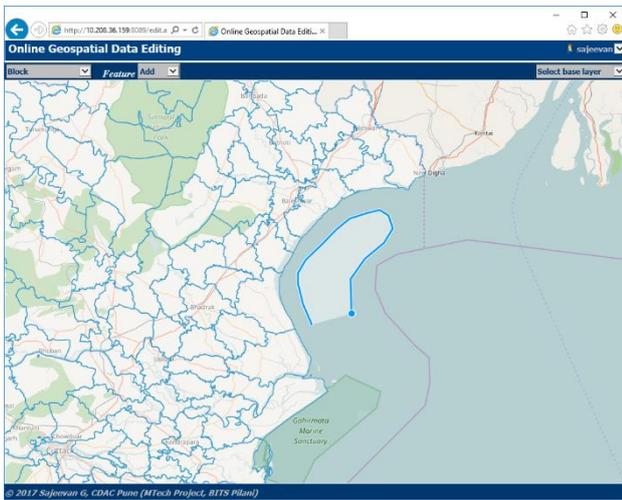


Figure 6: Adding a new feature in block layer

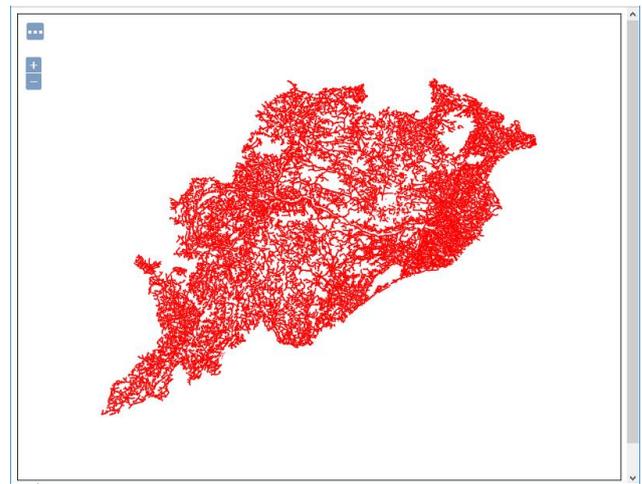


Figure 7: Road data service from shape file

6. Geospatial Data Service

The OGC standards web services started as part of the project, like WMS and WFS, can be used by any other application with or without authentication. The map services could also be accessed from any web browser as shown in figure 7.

Figure 8 shows WFS layer (BITS-OR:road_drrp_i) and WMS layer (OR layers) accessed by an open source desktop GIS software (QGIS). It shows spatial shift of old GIS layers (used in GRIMMS) on Bhuvan satellite data service. Those vector layers are corrected under PMGSY National GIS.

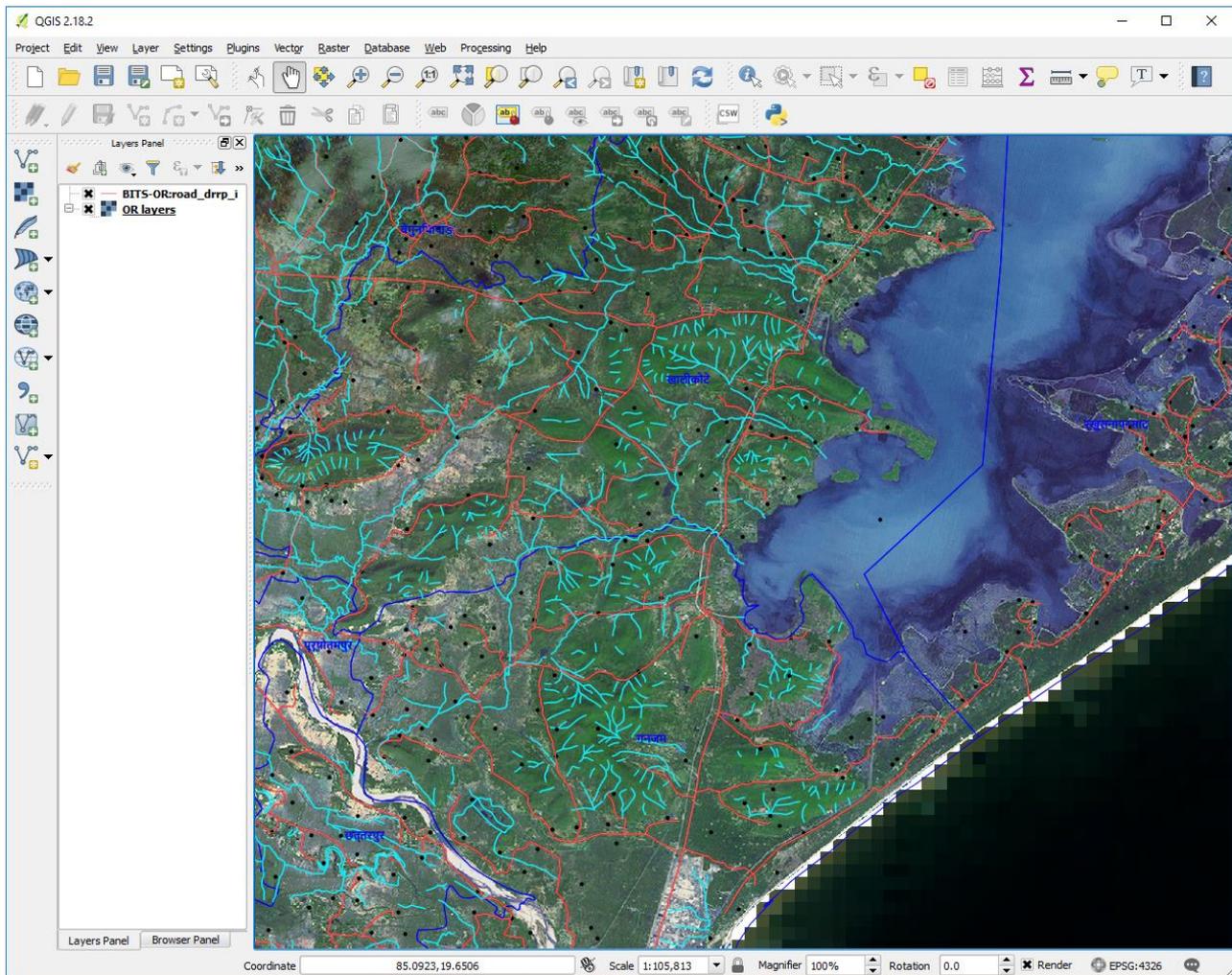


Figure 8: Geospatial data service

7. Conclusion

The massive volume of geospatial data of the country captured in digital form essentially requires spatial correction. The online geospatial transaction application would help immensely in achieving the spatial accuracy that is essential for planning, development and decision making. The application is developed using Free and Open Source Software (FOSS), thereby its benefits and aligning to the Government of India policy towards adaption of FOSS. The architecture, design and application may be used for national level GIS implementations for PMGSY geospatial data creation, planning, analysis, management and monitoring requirements. The application may also be extended to other GIS implementations. Moving the application to cloud environment can be beneficial for scalability, reliability, management, etc. Mobile applications can be plugged into the system for adding spatial data, data improvement, etc. The map services may be made available to other national geospatial projects. Considering the realistic digital spatial data requirements, crowd-sourcing can help incorporating various geospatial layers on fast pace.

The government agencies involved in planning and development should have access to accurate geospatial digital base data. It should be a collective responsibility of the government and the domain experts to bring about a change; a change for making accurate digital spatial data easily available for nation building. Rather than focusing on the elementary data conversion domain, the GIS industry should diversify on location-based services, GIS data processing, complex geospatial data capture/updating, etc. This can propel requirements for advanced spatial data and software (Sajeevan, 2017).

Various GIS technological advancements have been applied to the PMGSY National GIS so that the output could contribute immensely for the national GIS. The holistic approach initiated together with the central and state governments and private agencies delivered successful outputs. In future, the officials could focus on analysis, planning, modeling, etc, rather than hunting for reliable digital spatial data. The proposed online geospatial transaction system having geospatial data service could change the way GIS projects are handled and the GIS initiative is expected to open an ocean of opportunities for the government as well as the GIS industry. This will help decision-makers to arrive at reliable judgments easily.

Acknowledgements

The authors take this opportunity to express indebtedness to C-DAC, under the Ministry of Electronics and

Information Technology for supporting the project execution and NRIDA, Ministry of Rural Development for supporting with the data. The authors wish to express their gratitude to the reviewers for their observations that helped in improving the manuscript.

References

- Bhuvan (2017), ISRO. <http://bhuvan.nrsc.gov.in>
- GeoServer (2017). <http://geoserver.org>
- OGC (2017). <http://www.opengeospatial.org>
- OpenLayers (2017). <https://openlayers.org>
- PMGSY (2016). PMGSY National GIS Guidelines. http://pmsgsy.nic.in/pmsgsygis_v5.pdf
- PostGIS (2017). <http://www.postgis.net>
- Sajeevan, G. (2005). Correct usage of projection parameters for spatial accuracy, Map India Conference, 2005
- Sajeevan, G. (2008). Latitude and longitude – A misunderstanding, Current Science 94(5), pp. 568.
- Sajeevan, G. (2012). Dynamic composite key to link spatial database to external database, Journal of Geomatics 6(1), 5-6.
- Sajeevan, G. (2017a). Online Geospatial Data Editing, MTech dissertation, BITS Pilani.
- Sajeevan, G. (2017b). Nation building through GIS in India, GIM International. <https://www.gim-international.com/content/blog/nation-building-through-gis>, 03 May, 2017.
- Sajeevan, G., J. Mhatre, P. Ranade, U. Dutta and S. Londhe (2013). Rural road management gets on fast track with GIS, India Geospatial Digest.
- Sajeevan, G., J. Mhatre, P. Ranade, U. Dutta and S. Londhe (2012). GRIMMS Web - Connecting India, Geospatial Today 11(9), 42-43.
- Sajeevan, G., S. Ailawar and S. Chhillar (2007). Integrating GIS and MIS, GIM International 21(2), 50-51.
- Sajeevan, G., S. Baliga, A. Kaushal, S. Srivastava, C. P. Johnson, P. Kumar, Y. Singh, U. Dutta, J. Mhatre, B. Banerjee, V. K. M. Raja and A. Mishra (2006a). Integrated use of Standalone and Web GIS Software - A Case Study, International Conference on b-GIS, 7-8 December 2006, Technopark, Trivandrum.
- Sajeevan, G, V. K. M Raja, J. Mhatre and B. Banerjee (2006b). Development through GRIMMS Web, Geospatial Today 4(11), 30 – 32.

Spatial pattern of urban growth using remote sensing and landscape metrics

Akintunde John Akinrinola*

Department of Geography and Planning, Faculty of Environmental Sciences, University of Jos, Nigeria

*Email: akintundej@unijos.edu.ng

(Received: Sep 14, 2018; in final form: Jun 11, 2019)

Abstract: The growth of cities, occurring in uneven and unplanned patterns, affects land use and land cover and it changes the spatial distribution of urban residents. This study seeks to critically understand the pattern of urban growth in Ibadan metropolis, Nigeria by applying heuristic techniques. This study utilized Landsat 5 TM, 7 ETM+ and 8 OLI-TIRS satellite imageries of 1986, 2000 and 2013 respectively of Ibadan. These were integrated into a GIS environment using post-classification change detection approach and employing selected Landscape Metrics (TA, NP, MPS, TE, ED, AWMPFD) to analyze the pattern of urban growth in the area. The built up area in Ibadan has grown from 13302ha in 1986 to 45868ha in 2013 at an average growth rate of 2 and 12% per annum during 1986-2000 and 2000-2013 study periods respectively. Landscape metrics analysis reveal fragmented process of development along the fringes of Ibadan throughout the study periods with substantial increase of urban patches occurring during the second period of urbanization (2000-2013). The core of the city underwent compact growth by infilling of open spaces and through edge expansion over time. Remote sensing and landscape metrics proved valuable for the description of processes in the study.

Keywords: Urban growth, Urban Pattern, Landscape metrics, Urban patches, Fragmentation

1. Introduction

Urban growth is widely viewed as an essential driver of environmental and social problems. It causes the loss of informal open space and the fragmentation of wildlife habitats. Appropriate and exact evaluations of future urban development situations and related ecological effects are crucial for urban planning, approach choice, and natural resource administration. Since modern transformation toward the end of eighteenth century, world urban population has expanded exponentially with rapid speed. In 1830, urban population was about 1 billion and it expanded to 7 billion in 2011. In addition, urban population expanded more rapidly compared with rural population. Urban population increased from 14 percent in 1900 to 50.5 percent in 2010. By 2030, more than 60 percent of population are expected to be urban population (Wu et al., 2010).

Describing and understanding the dynamic patterns of urban growth is basic, given that urbanization continues to be one of the major global environmental changes in the nearest future. As a result of urbanization, the physical processes of urban land use changes are under study and investigation (Seto and Fragkias, 2005). Urban growth process significantly affects the land use patterns, influencing utilitarian parts of the landscape (Frohn and Hao, 2006; Akintunde et al., 2016). Spatial pattern of development on urban zones are changing impressively. Urban regions are expanding in rural zones and urban sprawl is taking place. In view of consistent change in structure and urban growth pattern, these zones have been in constant focal point of scientists (Seto and Fragkias, 2005).

While geographers and economic experts are creating geometric models that depict and clarify the morphology of urban communities for over a century (Herold et al., 2005), numerous components of urban spatial

configuration have proven elusive. Availability of temporal remotely detected information procured through space-borne sensors helps distinguishing the urban landscape progression in connection to urban growth (Chen et al., 2000; Epstein et al. 2002; Lo and Yang 2002; Ji et al., 2001; Yeh and Li 2001; Sudhira et al., 2003; Ramachandra et al., 2012). This guides in describing the spatio-temporal patterns of urban growth process and development (Zerah, 2008). More current research that merges satellite/GIS data with landscape metrics is equipped for examining land cover fragmentation, diversity and richness, and compactness within and across urban areas. Computation of metrics and displaying based on multi-temporal spatial data gives a premise for predicting urbanization processes.

This information supports policy making for an effective urban planning with natural resources conservation. Further temporal dynamics information with spatial metrics gives insights to the urbanization pattern (i.e., property, complexity and size of the urban zone), which enables the sustainable regional development (Hill et al., 2004; DeFries, 2008; Bhatta, 2009).

2. Description of the study area

Ibadan is located in the South Western part of Oyo State, Nigeria (Figure 1). Ibadan metropolitan area is located at latitudes 7°14'15''N to 7°36'34''N and longitudes 3°42'00''E to 4°06'56''E. It is located about 145 km north-east of Lagos, the federal capital of Nigeria with a population of 2,550,593 according to 2006 census results, including 11 local government areas. As the dominant urban center in Oyo State, its administrative and commercial functions transcend beyond the city boundaries. Ibadan metropolis covers a total land area of 3,123km² with 586 persons per km² as the overall population density of which the main city covers a land area 463.33 km².

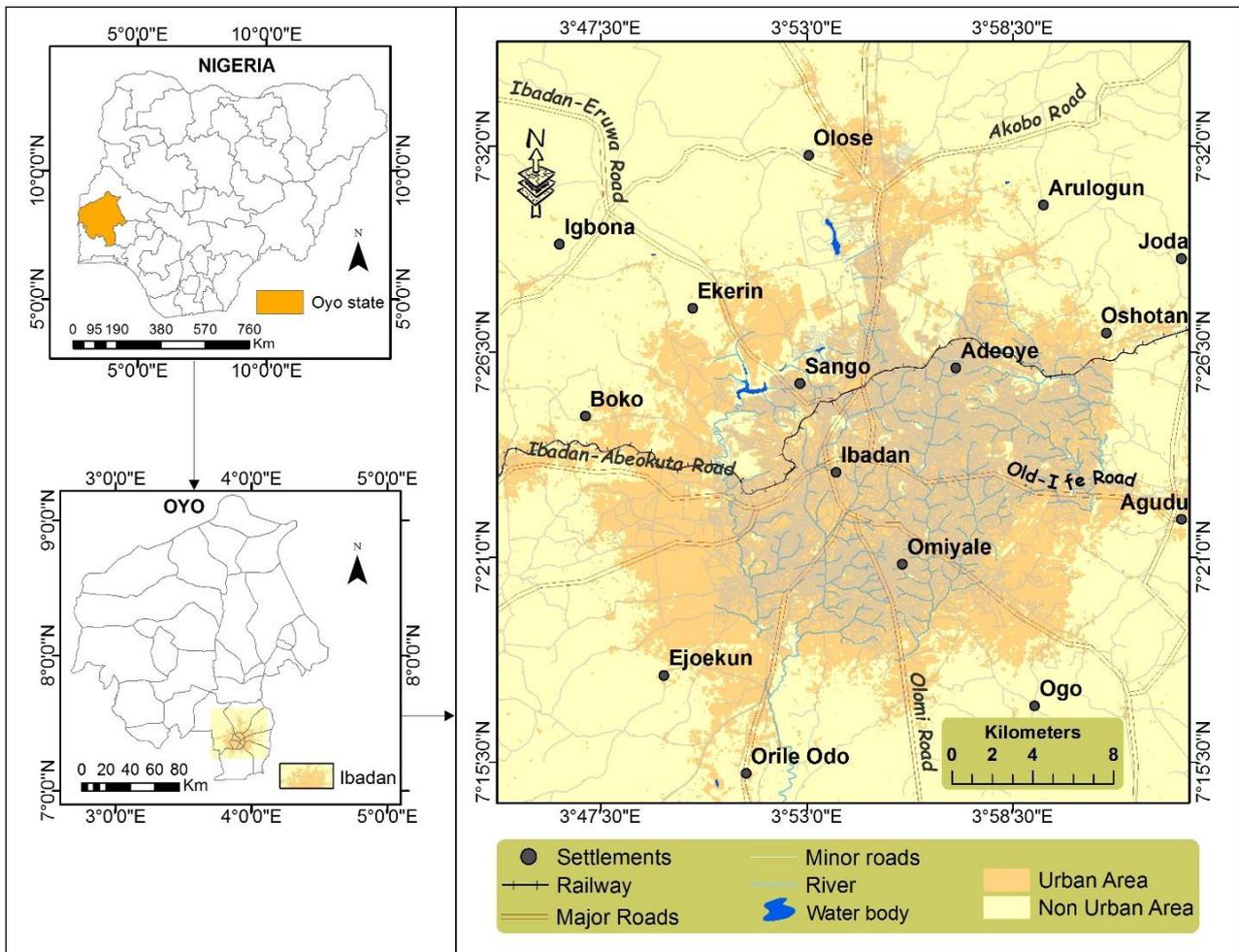


Figure 1: The Study Area.

Ibadan city is encountering unguided and uncontrolled expansion in all directions bringing about large scale urban growth and urban land use changes. Therefore, it is meaningful to extend landscape metrics application to fast developing urban communities. The current institutional structure of urban and regional planning framework (Master planning) cannot adequately address the issue of urban growth and its physical characteristics in Nigeria and Ibadan Metropolis in particular (Oyesiku, 2004; Alabi, 2008).

Therefore, this study will explore the spatio-temporal patterns of Ibadan's urban growth process and measure the hidden spatial configuration of the urban landscape. To achieve this task remote sensing and spatial metrics tools are employed. The combined use of these tools is believed to lead to new levels of understanding the urban development process which can assist city planners and policy makers to make informed decisions (Herold et al., 2005).

3. Methods and data analysis

3.1 Remote sensing image classification

The urban growth trends and patterns of Ibadan for a period of 27 years are analyzed using three multi-temporal medium resolution Landsat imageries (Table 1). All images are of the same spatial resolution, 30m.

Table 1. Data used and characteristics.

Data	Source	Year	Purpose
Landsat 5 TM	USGS	1986	Land cover and Land use analysis
Landsat 7 ETM+	USGS	2000	Land cover and Land use analysis
Landsat 8 OLI	USGS	2013	Land cover and Land use analysis
Topography Map [Scale (1:50000)]	OSGOF	1963	Generate boundary and Base layer maps
High resolution image	Google Earth	2014	Visual interpretation

Images used are acquired geometrically corrected and geo-referenced. Supervised maximum likelihood classification algorithm was applied in ENVI 5.1 software environment to run image classification due to its popularity and wide acceptance in classifying remote sensing images. Accordingly, the images were classified into different land cover classes which finally ended up generating three different year land cover maps of the study area. Pixels with maximum likelihood are categorized into the matching class as shown in figure 2.

The land cover maps are composed of two major land cover classes namely; built up and non-built up. The built up consist of commercial, residential, road and impervious features, residential, industrial and commercial units, road and railway networks, parking lots, sport and leisure facilities, etc. while the non-built up includes cropland (agriculture land), parks, grasslands, forests, green spaces, bare soil and others.

3.2 Accuracy assessment

In remote sensing land cover mapping, classification accuracy is the most important aspect to assess the reliability of the final output maps. In this study, Accuracy Assessment is done through comparison of Kappa coefficients (Congalton *et al.*, 1983). For this purpose, a confusion matrix was calculated. Accuracy assessment and Kappa coefficient are common measurements used in various publications to demonstrate the effectiveness of the classifications (Congalton, 1991; Lillesand and Kiefer, 2005). The main purpose of assessment is to assure classification quality and user confidence on the product (Foody, 2002). In the present study, accuracy was assessed after several classes were merged and classified to come out with three classes of interest.

3.3 Change detection

The method used in this analysis is the post classification comparison technique in which GIS overlay of the independently produced classified images in ArcGIS 10.5 (Alphan *et al.*, 2009). The subsequent land cover maps are then visually compared and change areas are simply those areas which are not classified the same at different times. This method is the most straightforward and intuitive change detection method. Following this method, maps are produced to show the built up class between each subsequent years, i.e. 1986-2000 and 2000-2013 (Yang and Lo, 2002). In combination with class area landscape metrics, these make it possible to quantify the spatial extent and rate of urban growth over time.

3.4 Measuring urban growth pattern using landscape metrics

Spatio-temporal patterns of Ibadan's growth were analyzed using landscape metrics for the time period 1986-2013. Landscape metrics are powerful tools to quantitatively describe and compare multi-date thematic maps. Metrics are computed only for the built up class in the study. The outputs for the selected metrics presented in tables are generated for the whole study area and calculated in Patch Analyst v5.0 and Fragstats 4. These metrics were picked based on their intuitiveness, ease interpretation and their ability to describe the composition and configuration of urban landscape pattern. Nevertheless, the analysis is conducted including areas outside Ibadan, but only focusing on built up land cover class. Landscape metrics describe four dimensions: relative size, absolute size, spatial distribution of patches and complexity of urban form. Landscape metrics employed in this study are given below (Mcgarigal and Marks, 1995).

1) *Total (Class) Area (TA)*: Total area measures how much of the landscape comprises of a particular type of patch.

$$TA = \sum_{j=1}^n a_{ij} \left(\frac{1}{10000} \right) \quad (1)$$

a_{ij} = area (m²) of patch ij.

2) *Number of Patches (NP)*: NP is a measure of isolated urban areas in the landscape. During epochs of rapid urban nuclei development, NP is expected to increase but may experience decrease if urban areas expand and merge into continuous urban fabric (Seto and Fragkias, 2005).

$$NP = N \quad (2)$$

N = total number of patches in the landscape.

3) *Mean patch size*: MPS measures the number of urban patches per the size of each urban area which increases or decreases over time. (Seto and Fragkias, 2005).

$$MPS = \frac{A}{N_{patch}} (10000) \quad (3)$$

A = area (m²) of all patches of the corresponding patch.
N = total number of patches in the landscape.

4) *Total Edge*: TE sums up the lengths (in meters) of all edge segments that contain the similar patch type.

$$TE = E \quad (4)$$

E = total length (m) of edge in landscape.

5) *Edge Density (ED)*: ED is computed by dividing the total length of the urban boundary to the total landscape area. ED has direct relationship with NP.

$$ED = \frac{E}{A} (10000) \quad (5)$$

E = total length (m) of edge in landscape.
A = total landscape area (m²).

6) *Area weighted mean patch fractal dimension (AWMPFD)*: AWMPFD metric describes the degree to which the shape of an urban area is irregular or complex. Values range between 1 and 2 with values closer to 1 indicating areas with relatively simple shapes such as squares or circles. Values that approach 2 represent irregular and complex shapes.

$$AWMPFD = \sum_{i=1}^m \sum_{j=1}^n \left(\frac{2 \ln(2.5 a_{ij})}{\ln a_{ij}} \right) \left(\frac{a_{ij}}{A} \right) \quad (6)$$

a_{ij} = area (m²) of patch ij.
A = total landscape area (m²)

4. Results and discussion

4.1 Accuracy assessment

The overall accuracy of classified images was found to be greater than 85%. This is considered to be a good result for analysis performed using remote sensing images (Herold *et al.*, 2005). Tables 2 and 3 below present the accuracy of classified images used.

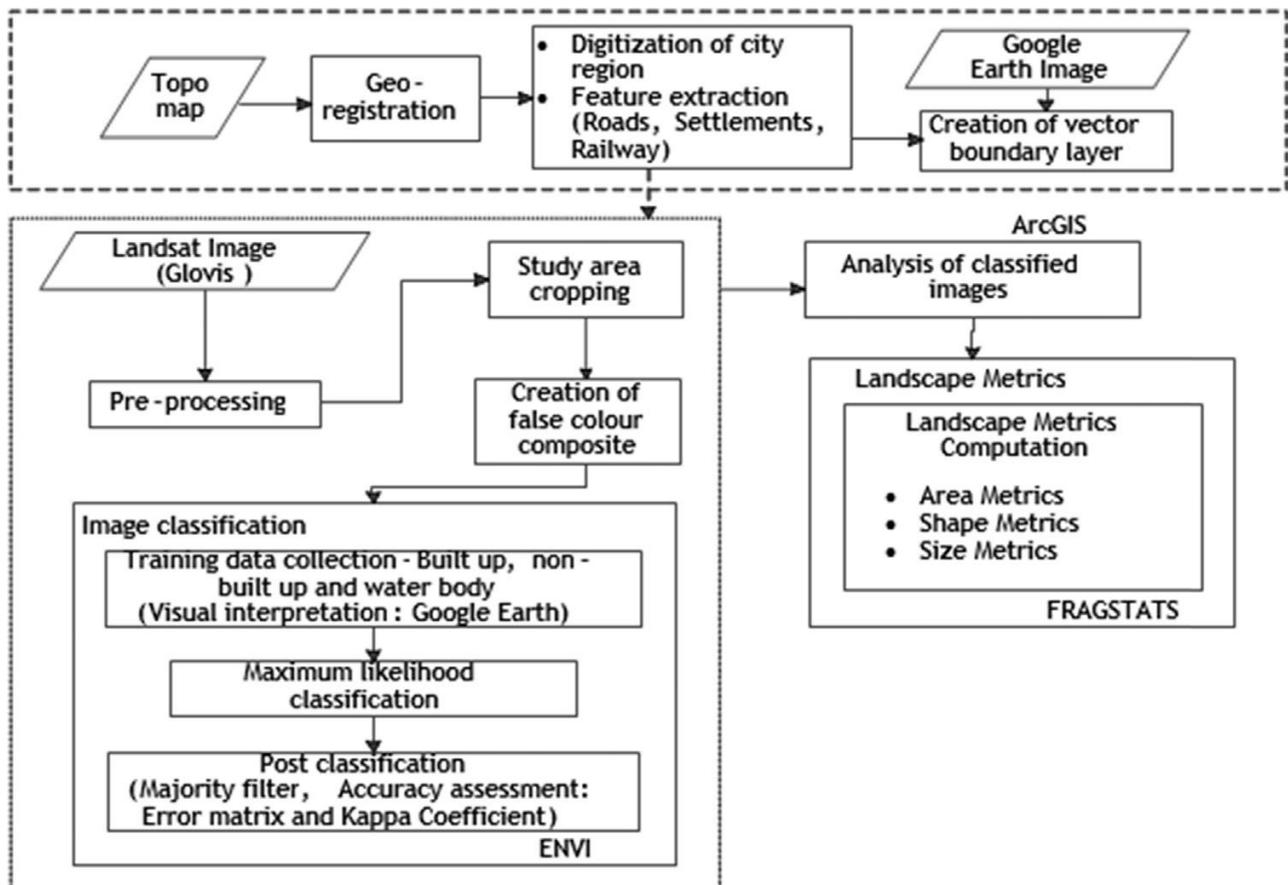


Figure 2: General Methodology workflow

Table 2. Error matrix of classified images (1986-2013)

1986				
Class	Builtup	Non Builtup	Waterbody	Row Total
Builtup	2772	0	0	2772
Non Builtup	0	6186	1	6187
Waterbody	0	0	87	87
Column Total	2772	6186	88	9046
2000				
Class	Builtup	Non Builtup	Waterbody	Row Total
Builtup	2771	0	0	2771
Non Builtup	1	5985	11	5997
Waterbody	0	201	77	278
Column Total	2772	6186	88	9046
2013				
Class	Builtup	Non Builtup	Waterbody	Row Total
Builtup	13904	20	1	13925
Non Builtup	0	9445	0	9445
Waterbody	0	0	386	386
Column Total	13904	9465	387	23756

4.2 Image classification

The classification of the multi-temporal satellite images into built up, non-built up and water body for the three different time periods of 1986, 2000, and 2013 (Figure 3a, b and c) shows a highly simplified and abstracted depiction of the study area.

The maps show a clear increase in the pattern of urban expansion extending from the city core to the adjoining non-built up areas along the major transportation corridors. Figure 3d shows the spatial and temporal pattern of urban growth in the study area experiencing rapid expansion along the fringes of the built up areas. Post classification composition of images classified revealed the pattern of urban growth of the city in different directions, specially, the open spaces experiencing infilling amid already built up regions and the dynamics of expansion of the urban regions in the study area. However, it is imperative to assist the findings with statistical evidences as it is useful to describe the spatial extent and different urban growth patterns that have been occurring in the study area. This will help understand how the city is changing over time and to compare the various growth patterns taking place quantitatively in different time epochs.

4.3 Spatio-temporal analysis of urban growth pattern using landscape metrics

The highest rate of urban growth is observed during the second period of urbanization (2000-2013) with an increase in the built up area of more than six times (160%) within 13 years (Table 4).

Table 3. Accuracy assessment of classified images (1986-2013)

Land Use Class	Reference Total	Classified Total	Number Correct	Number Wrong	Producer's Accuracy (%)	User's Accuracy (%)
1986						
Builtup	2772	2772	2772	0	100	99.85
Non Builtup	6187	6186	6186	0	99.79	100
Waterbody	87	88	87	1	99.74	100
Total	9046	9046	9045	1		
2000						
Builtup	2771	2772	2771	1	99.96	100
Non Builtup	5997	6186	5985	201	96.75	99.8
Waterbody	278	88	77	11	87.5	27.7
Total	9046	9046	8833	213		
2013						
Builtup	13925	13904	13904	0	100	98.85
Non Builtup	9445	9465	9445	20	99.79	100
Waterbody	386	387	386	1	99.74	100
Total	23756	23756	23735	21		

1986: Overall accuracy = 99.988, Kappa coefficient = 0.999 **2000:** Overall accuracy = 97.645, Kappa coefficient = 0.948 **2013:** Overall accuracy = 99.911, Kappa coefficient = 0.998

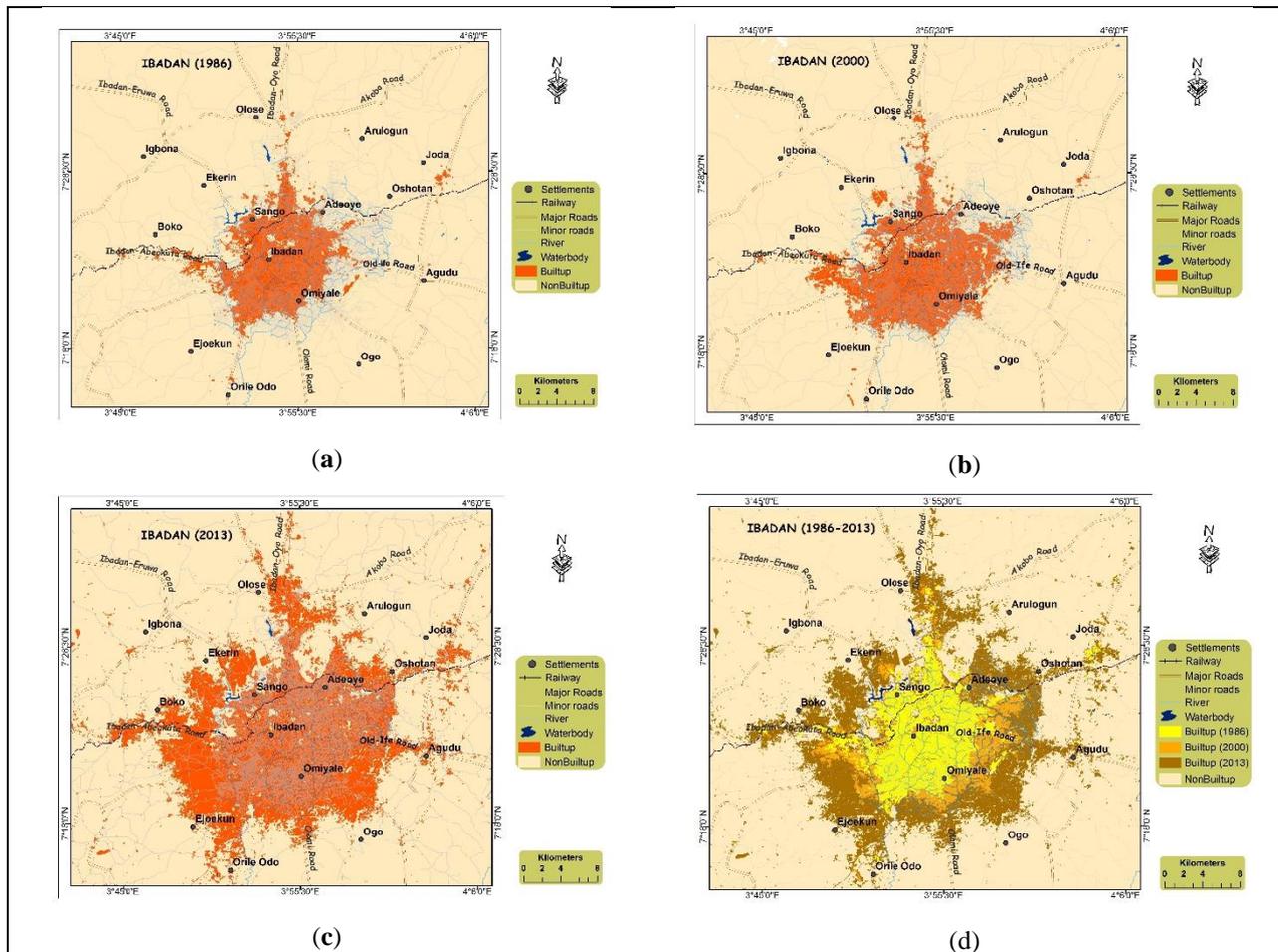


Figure 3: (a) Image classification, 1986; (b) Image classification, 2000; (c) Image classification, 2013; (d) Spatio-temporal growth map of the study area (1986-2013).

Table 4. Analysis of built up area expansion based on total area (TA) metrics.

Study period	Change (ha)	Change (%)	Time span	Growth rate/year	Average
1986-2000	4325	33	14	2	7
2000-2013	28241	160	13	12	

This is followed by 33% during the first period of urbanization (1986-2000). This indicates that more rapid urban growth took place in the study area during the period of 2000-2013 compared to the first period. As the statistics obtained from the area metrics computation confirms, the built up area increased at an average annual growth rate of 2 and 12% during the periods 1986-2000 and 2000-2013 respectively (Figure 4).

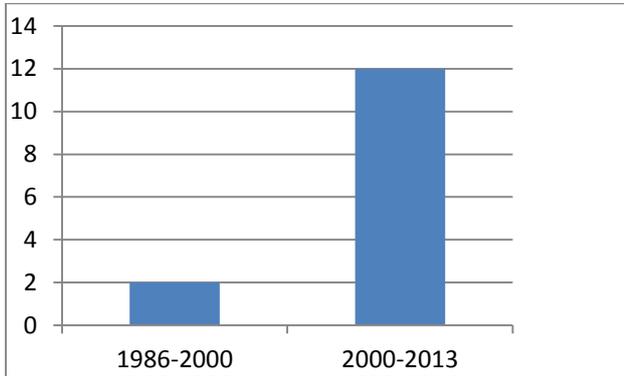


Figure 4: Built up area growth rate (%) per annum per study period

The results presented in Table 5 show that the total built-up area (TA) has grown from 13302ha in 1986 to 17627ha in 2000 and to 45868ha in 2013.

Table 5. Landscape metrics at the entire landscape.

Year	LUC	TA	NP	MPS	TE	ED	AWMPFD
1986	Builtup	13302	473	28	758400	4	1.38
2000	Builtup	17627	523	34	1122300	6	1.41
2013	Builtup	45868	2212	21	3078780	16	1.43

In terms of absolute change in (ha) of land cover the second period 2000 to 2013 (Figure 5) remains the highest witnessing the conversion of 28241ha of non-built up land to urban land. The first period of urbanization (1986-2000) experienced 4325ha of land changed to built up area. Totally 32566ha of non-built up land has been converted to built up land over the period 1986 to 2013.

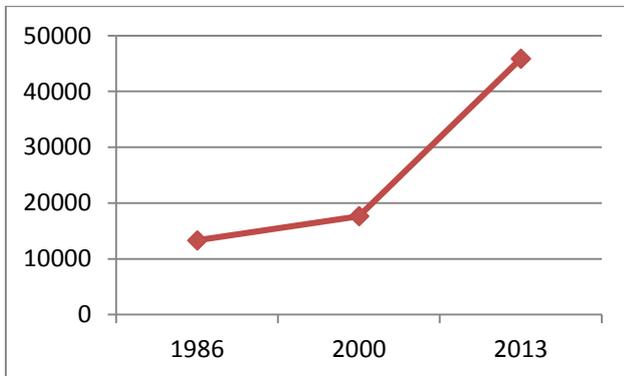


Figure 5: Total Built up area (TA)

The continuous rise of number of patches (NP) has led to and is revealed by the rapid urban growth process in the study area landscape throughout the study periods. In 1986 the NP in the region was 473 and gradually increased to 523 in 2000, and rapidly increased to 2212 in 2013 (Figure 6). This could be an indication of fragmented and

heterogeneous process of urban growth taking place in the study area. During the 2000 to 2013 period, there was a significant change observed in NP. However, the peak occurred in 2013 indicating the continuing development of scattered and fragmented urban patches in the study area. This situation can be attributed to a development of small and irregular built up patches around the periphery of the city and in peri-urban regions. This could happen as the city expands outward in the form of scattered development, the gap between the peri-urban regions and the urban core will decrease by increasing the attractiveness of the peri-urban area for development.

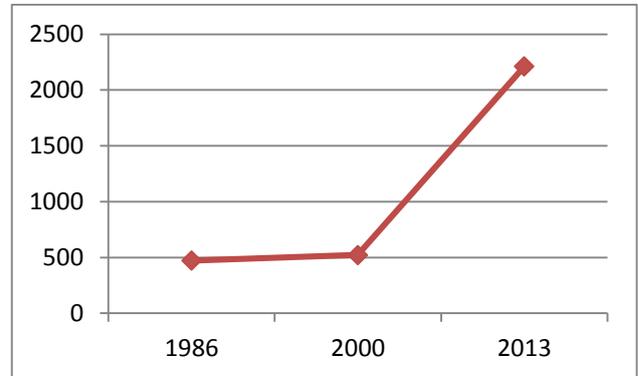


Figure 6: Number of Patches (NP)

Mean patch size (MPS) presents the relationship between urban patches land area and their number. The reduction in MPS shows that new patches have been developed. The increase in MPS shows the extension of existing urban patches. Figure 7 shows that the value of this metric increased during 1986-2000. Since 1986, it has decreasing trend. In 1986, the MPS was 28. With the joining of new patches, MPS extended to 34 in 2000. With the expansion of new patches, the MPS witnessed considerable reduction, 21 in 2013. The fluctuation is associated to the growth of central core and annexation of patches surrounding the central core until 2000. Since 2000, the urban areas experienced more developments of the new patches.

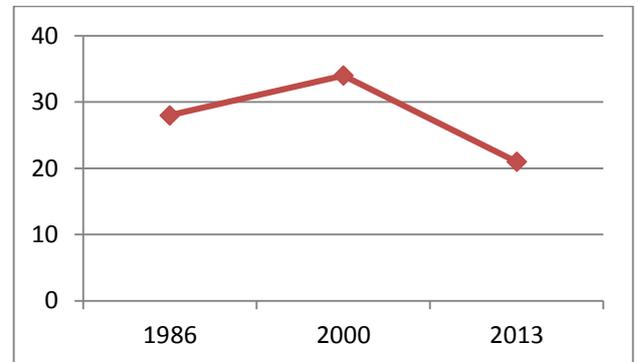


Figure 7: Mean Patch Size (MPS)

Total Edge (TE) considers true edges values greater than or equal to zero. Larger continuous patches indicate edges with larger values. Figure 8 indicates that during 1986 and 2000 the edges were smaller and hence there were discontinuous patches as the landscape was fragmented. In 2000 and 2013, larger edges indicated that the urban edges are ubiquitous and continuous.

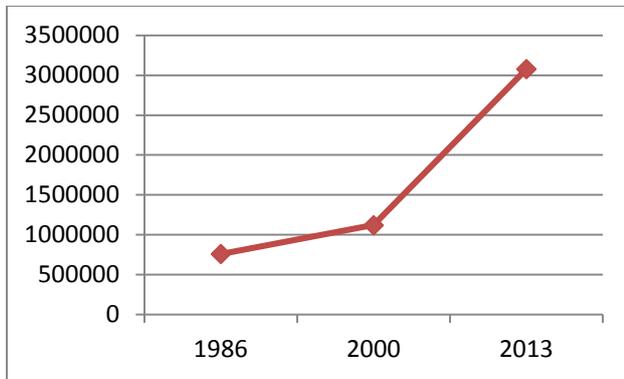


Figure 8: Total Edge (TE)

The result of this study revealed that the edge density (ED) increased from 4 in 1986 to 6 in 2000 and to 16 in 2013 (Figure 9). This shows that there has been significant urban growth with the emergence of various fragmented urban patches observed in the study landscape.

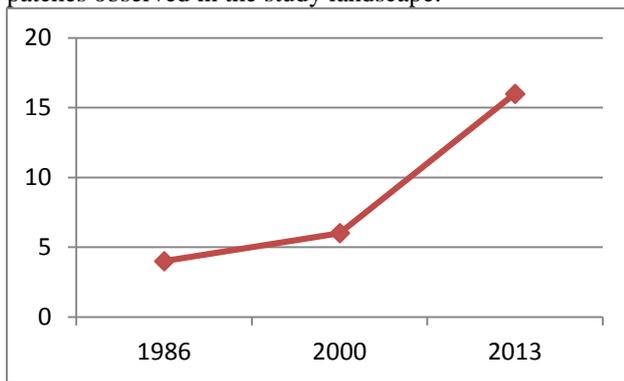


Figure 9: Edge Density (ED)

The area weighted mean patch fractal dimension (AWMPFD) with a consistent increasing trend observed in figure 10 shows the complexity and growing irregularity of urban patches due to fragmentation. This can be associated with the partial integration of existing individual patches and probably the formation of fewer new patches during both periods.

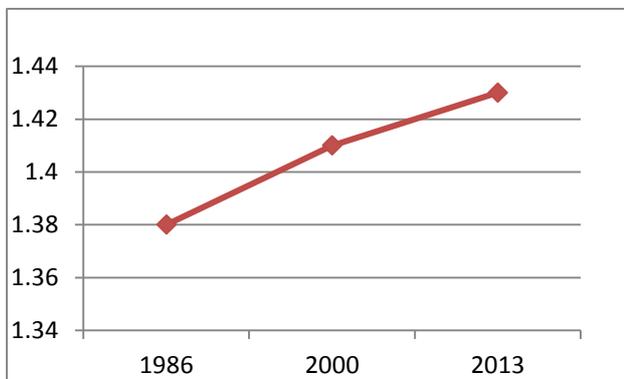


Figure 10: Area Weighted Mean Patch Fractal Dimension (AWMPFD)

5. Conclusions and recommendations

The study has made it possible to successfully capture the changing subtleties of the urban growth pattern at metropolitan (landscape) urban level. Ibadan metropolis experienced fragmented urban growth process,

particularly, at the fringe areas with substantial built up increase while, the core of the city underwent relatively compact growth by infilling open spaces and through edge expansion over time. The built up area in the metropolis has grown from 13302ha in 1986 to 45868ha in 2013 at an average growth rate of 2 and 12% per annum during 1986-2000 and 2000-2013 study periods respectively. In total, 32566ha of non-built up has been converted to urban area.

Analyzing the spatial extent and rate of urban growth as well as identifying the growth directions alone does not give sufficient insight in to the patterns of urban growth processes, which are important to having a better understanding of the urban pattern. To bridge this gap, landscape metrics are used. Six metrics namely: total area (TA), number of patches (NP), mean patch size (MPS), total edge (TE), edge density (ED) and area weighted mean patch fractal dimension (AWMPFD) were utilized to evaluate the patterns of urban growth and processes experienced in Ibadan and its environs at landscape level. Based on the number of patches (NP), the built up area experienced fragmented growth process all through study periods with the second period of urbanization, 2000 to 2013 witnessing substantial increase of built up area (TA). The fluctuation in metric value of mean patch size (MPS) is linked to the enlargement of the central core and annexation of patches surrounding the central core until 2000. Since 2000, there is a clear increase in developments of the new urban patches. The area weighted mean patch fractal dimension (AWMPFD) showed increasing trend. This illustrates that the entire built up area will keep on getting more complex and thus, fragmented over time mainly at the fringe areas. Unorganized development that could be due to poor planning scheme could have played an inevitable part in the fragmented process of development of Ibadan metropolis.

Since the value of information extracted from landscape metrics is dependent on the quality of image classification, future studies could attempt to improve on the classification accuracy of the satellite images utilized in this study or perhaps use images from the same sensor, for instance all images from Landsat ETM+. This could help in solving the issues involved with image consistency.

References

Akintunde, J. A., E. A. Adzandeh and O. O. Fabiyi (2016). Spatio-temporal pattern of urban growth in Jos Metropolis, Nigeria. *Remote Sensing Applications: Society and Environment* 4, 44–54

Alabi, M. (2008). *Urban Sprawl, Pattern and Measurement in Lokoja, Nigeria* Owei, Ede, ObinnaAkarolo, (2008): *Land Market Distortions in Nigerian Cities and Urban Sprawl*, 44th ISOCARP Congress Land market distortions in Nigerian cities and urban sprawl: the case of Abuja and Port Harcourt.

Alphan, H., H. Doygun and Y. Unlukaplan (2009). Post classification comparison of land cover using multitemporallandsat and ASTER imagery: the case of

- Kahramanmaras, Turkey. Environmental monitoring and assessment, 151(1-4), 327-336.
- Bhatta, B. (2009). Analysis of urban growth pattern using remote sensing and GIS: a case study of Kolkata, India, *International Journal of Remote Sensing* 30, 4733-4746.
- Chen, S., S. Zeng and C. Xie (2000). Remote Sensing and GIS for urban growth analysis in China, *Photogrammetric Engineering and Remote Sensing* 66, 593-598.
- Congalton, R., R. Oderwald and R. Mead (1983). Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques, *Photogrammetric Engineering and Remote Sensing* 49, 1671-1678.
- Congalton, R. (1991). A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment* 37, 35-46.
- Defries, R. (2008). Terrestrial vegetation in the coupled human–earth system: contributions of remote sensing, *Annual Review of Environment and Resources* 33, 369-390.
- Epstein, J., K. Payne and E. Kramer (2002). Techniques for mapping suburban sprawl, *Photogrammetric Engineering and Remote Sensing* 63, 913-918.
- Foody, G. (2002). Status of land cover classification accuracy assessment. *Remote sensing of environment*, 80(1), 185-201.
- Frohn, R. and Y. Hao (2006). Landscape metric performance in analyzing two decades of deforestation in the Amazon Basin of Rondonia, Brazil, *Remote Sensing of Environment* 100, 237-251.
- Herold, M., H. Couclelis and K. Clarke (2005). The role of spatial metrics in the analysis and modeling of urban land use change. *Computers, Environment and Urban Systems*, 29(4), 369-399.
- Hill, J., P. Hostert and A. Roder (2004). Long-term observation of Mediterranean ecosystems with satellite remote sensing, In: Mazzoleni S., di Pasquale G., Mulligan M., di Martino P., Rego F. (Eds.), *Recent Dynamics of the Mediterranean Vegetation and Landscape*, John Wiley and Sons Ltd., Chichester, 33- 43.
- Ji, C., P. Lin, X. Li, Q. Liu, D. Sun, and S. Wang (2001). Monitoring urban expansion with remote sensing in China, *International Journal of Remote Sensing* 22, 1441-1455.
- Lillesand, T. and R. Kiefer (2005). *Remote sensing and image interpretation*, John Wiley and Sons, New York.
- Lo, C. and X. Yang (2002). Drivers of land-use/land-cover changes and dynamic modelling for the Atlanta, Georgia Metropolitan Area, *Photogrammetric Engineering and Remote Sensing* 68, 1062-1073.
- McGarigal, K. and B. Marks (1995). FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. USDA for. Serv. Gen. tech. rep. PNW-351.
- Oyesiku, O. (2004). *Town and Country Planning Law and Administration in Nigeria*. Readings in Urban and Regional Macmillan Nigeria Publishers Limited. Pp. 481-520.
- Ramachandra, T., A. Bharath and D. Durgappa (2012). Insights to urban dynamics through landscape spatial pattern analysis, *International Journal of Applied Earth Observation and Geoinformation* 18, 329-343.
- Seto, K. and M. Fragkias (2005). Quantifying spatio-temporal patterns of urban land-use change in four cities of China with time series landscape metrics. *Landscape Ecology*, 20 (7), 871-888.
- Sudhira, H., T. Ramachandra and K. Jagadish (2003). Urban sprawl pattern recognition and modelling using GIS, Paper presented at Map India, January 28-31, New Delhi, <http://www.gisdevelopment.net/application/urban/sprawl/mi03142.htm>
- Wu, J., G. Jenerette, A. Buyantuyev and C. Redman (2010). Quantifying spatiotemporal patterns of urbanization: The case of the two fastest growing metropolitan regions in the United States. *Ecological Complexity*.
- Yang, X. and C. Lo (2002). Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. *International Journal of Remote Sensing*, 23(9), 1775-1798.
- Yeh, A. and X. Li (2001). Measurement and monitoring of urban sprawl in a Rapidly growing region using Entropy, *Photogrammetric Engineering and Remote Sensing* 67, 83-90.
- Zerah, M. (2008). Splintering urbanism in Mumbai: Contrasting trends in a multi-layered society, *Geoforum* 39, 1922-1932.

Study of discontinuity adaptive MRF models with kernel-based noise classifier

Ishuita SenGupta^{1*}, Anil Kumar² and Rakesh Kumar Dwivedi¹

¹College of Computing Sciences and Information Technology, TeerthankarMahaveer University, Moradabad, India

²Indian Institute of Remote Sensing, Dehradun, India

*Email: ishuitasengupta8@gmail.com;

(Received: Jan 12, 2019; in final form: June 14, 2019)

Abstract: The paper aims to determine the effect of adding contextual models and kernel functions with fuzzy based noise classifier with remote sensing data. The non-linearity between class boundaries can be handled through the kernel functions and contextual models eliminates the probability of isolated pixels. Nine different Kernel functions have been combined with conventional Noise Clustering without Entropy classification method (KNC), to classify data obtained from Landsat-8 and Formosat-2 satellites. For contextual support Markov Random Field (MRF) models were introduced with KNC. Standard regularization model (smoothness prior) and four Discontinuity Adaptive (DA) models (edge preserving priors) have been studied with KNC and abbreviated as KNC-S-MRF, KNC-DA1-MRF, KNC-DA2-MRF, KNC-DA3-MRF, KNC-DA4-MRF, respectively. An increase in overall accuracy has been observed when a comparative analysis has been done with the established Noise classifier.

Keywords: Kernel functions, Kernel based Noise Clustering without Entropy (KNC), Markov Random Field (MRF) models, Regularization Model, Discontinuity Model.

1. Introduction

Expansion of Remote Sensing applications have directed to availability of colossal quantity of data. Challenges to sustain the quality of such data also have increased, requiring more robust framework for processing and analysis of these data. Traditional classification techniques designate each pixel to a single land cover class resulting in a hard (or 'crisp') partitioning (Zhang and Foody, 2001). Due to coarse spatial resolution, more than one land-cover type may exist within a pixel, such a pixel is termed as mixed pixel (Foody, 1996) and ignorance of it resulted in a reduction in classification accuracy. Incorporation of mixed pixel has been facilitated in all stages of a classification process to produce accurate and meaningful land cover classifications from remote sensing images (Ibrahim et al., 2005).

The extensive use of fuzzy logic (Zadeh, 1978) for classification leads into soft classifiers. Among the most prominent fuzzy classifiers, Fuzzy *c*-mean (FCM) had been successfully used for estimation and mapping of sub-pixel level land cover composition (Foody, 2000, Fisher and Pathirana, 1990), although it failed to handle noise. Possibilistic *c*-Means (PCM) was developed to overcome the drawback of FCM as PCM was able to surpass the effect of hyperline constant found in FCM (Chawla, 2010). The challenging problem of noise removal was considered from different perspectives (Jolion and Rosenfeld, 1989, Krishnapuram and Freg, 1992) and among them Noise clustering was found to give the best performance (Dave, 1991; Dave 1993). Lately, it was proven that the Noise clustering algorithm is a generalization, where PCM and FCM are its special cases (Dave and Sen, 1997).

Studies related to spatial contextual information in the classification process illustrates improvement in the classifiers robustness against noise when compared to purely spectral based classification algorithm.

(Krishnapuram and Keller, 1996; Foody, 2000). Inclusion of contextual information while classifying helps in removing isolated pixel problem. MRF based contextual methods were used for classification and fusion of multi-source data and it was proven that the classification accuracy has improved and is more reliable over other contextual methods (Solberg et al., 1996; Binaghi et al, 1997). A Robust Fuzzy *c*-Means (RFCM) algorithm was developed by adding contextual information to the objective function of FCM using MRF, while performing image segmentation of Magnetic Resonance Images of brain (Pham, 2001). An Adaptive Bayesian Contextual classifier, which combines the advantages of Adaptive classifier and Bayesian Contextual classifier demonstrated, using MRF modeling of joint probabilities of classes of each pixel and its neighborhood could improve the classification accuracy by mitigating the effect of Speckle error (Jackson and Landgrebe, 2002). Providing contextual support to Noise classifier was proposed earlier with the aim to overcome sensitivity of noise and outliers on the classification result using S-MRF or DA-MRF models (Harikumar, 2014). Integration of contextual information onto support vector machines classifier using MRF model was achieved by reformulating the prior energy function in terms of suitable SVM-like kernel expansion (Moser and Serpico, 2010).

The kernel methods map the input data to a higher dimensional space where the data turn out to be linearly separable (Awan and Sap, 2005). Studies related to kernels and assessment of fusion with fuzzy based classifier has been done earlier also. An unsupervised Kernel Noise clustering algorithm was also proposed (Choti wattana, 2009) based on distances of kernel method (Gaussian and higher order polynomial) and was found to be relatively more resistant against noise. PCM (Possibilistic *c*-Mean) has been modified with KPCM by replacing Euclidean norm with Gaussian Kernel, resulting to increase in robustness to noise (Ganesan and Rajini, 2010). To deal with the drawback of fuzzy clustering

KFCM was introduced (Ravindraiah and Tejaswini, 2013). Local kernel like KMOD and inverse multiquadratic kernel as well as the global kernels were studied and incorporated to enhance the capability of FCM (Bhatt and Mishra, 2013). (Rhee et. al., 2012) proposed a kernel based possibilistic clustering technique, in which Fuzzy Kernel c-Means (FKCM) algorithm for initialization of PCM was used and PCM was modified using kernel induced metric replacing Euclidean distance measure and showed better results than FCM, PCM, and FKCM. Incorporation of eight kernels with Fuzzy *c*-Means classification to handle the nonlinearity among classes has shown improved accuracy (Byju, 2015). Entailing Kernels with fuzzy based classifier have shown effective results than the conventional ones.

The objective of present paper is to develop a novel method that combines the positives of spectral classification with the contextual spatial information. Supervised Noise Clustering has been opted as the base classifier, and adding nine different kernel functions as the distance functions with it lead to derive a kernel based classifier, termed as, KNC (Sengupta *et.al*, 2019). For contextual support Markov Random Field (MRF) models have been incorporated with KNC. Standard regularization model (smoothness prior) and four

Discontinuity Adaptive (DA) models (edge preserving priors) have been studied with KNC and abbreviated as KNC-S-MRF, KNC-DA1-MRF, KNC-DA2-MRF, KNC-DA3-MRF, KNC-DA4-MRF, respectively. Image to image accuracy assessment has been formulated jointly with computation of overall accuracy using Fuzzy error matrix (FERM) of every kernel specified.

2. Study area and dataset used

The datasets used have been acquired from Landsat-8 and Formosat-2 satellites. Landsat8 provides moderate-resolution imagery, from 15 meters to 100 meters, of Earth's land surface and operates in the visible, near-infrared, short wave infrared and thermal infrared spectrums. Formosat2 captures panchromatic and multispectral data simultaneously with 2meters and 8meters resolution respectively. The sensors' spectral wavebands specifications are enlisted in table 1a and b. The site for the study work is situated in Haridwar district in the state of Uttarakhand, India. Area extends from 29°52'49" N to 29°54'2" N and 78°9'43" E to 78°11'25" E. The site is identified with five land cover classes i.e. Water, Wheat, Forest, Riverine Sand, Fallow Land.

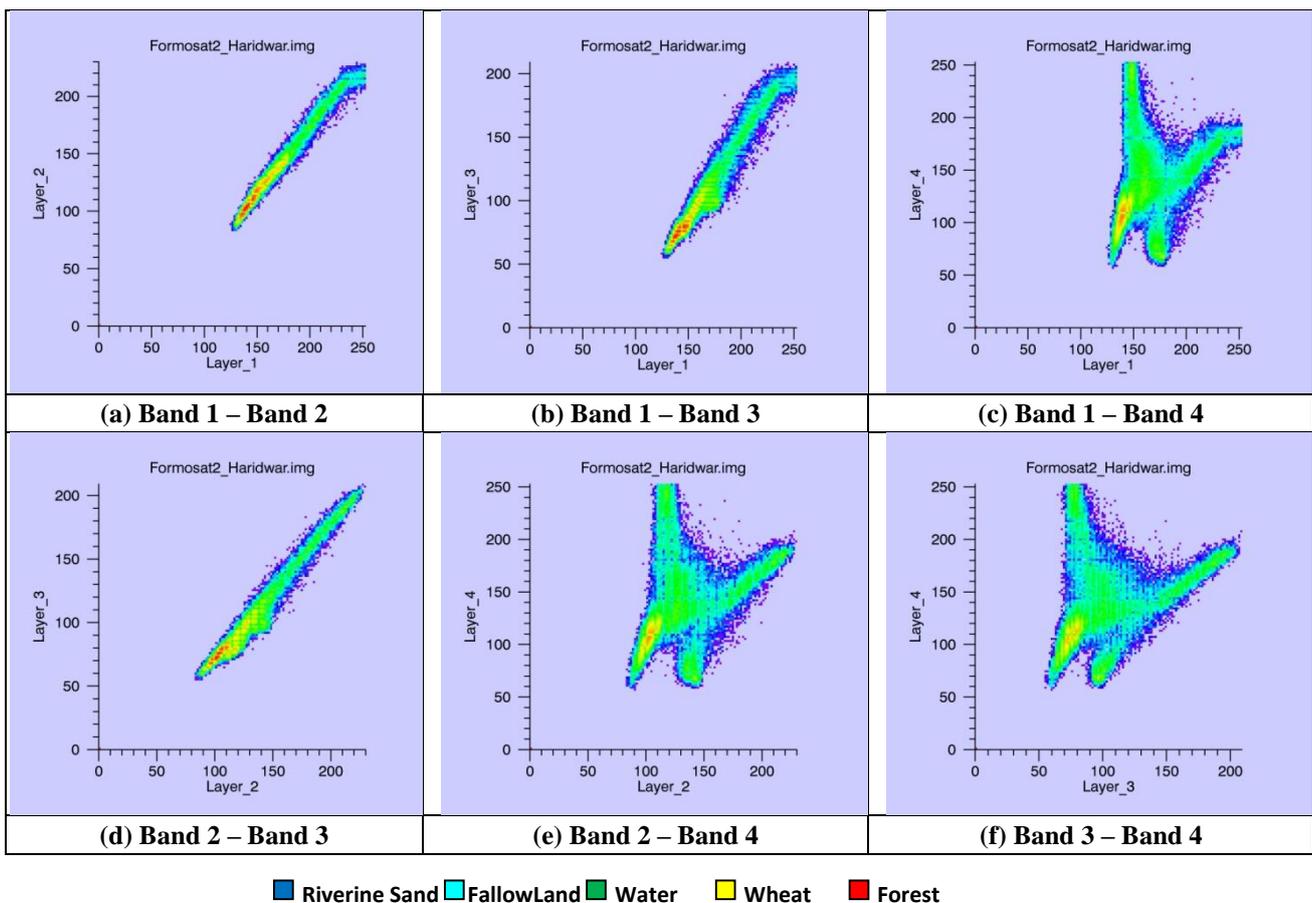


Figure 1: Non-linearity in different classes as 2D scatter plots for Formosat2 for all classes identified. (Generated using ENVI 5.0) (Sengupta *et.al*, 2019)

Table 1: (a) Data Details of Landsat8

Spectral Band	Wavelength (μm)	Resolution (m)
Band 1 - Coastal aerosol	0.43 - 0.45	30
Band 2 - Blue	0.45 - 0.51	30
Band 3 - Green	0.53 - 0.59	30
Band 4 - Red	0.64 - 0.67	30
Band 5 - Near Infrared (NIR)	0.85 - 0.88	30
Band 6 - Short Wavelength Infrared 1	1.57 - 1.65	30
Band 7 - Short Wavelength Infrared 2	2.11 - 2.29	30
Band 8 - Panchromatic	0.50 - 0.68	15
Band 9 - Cirrus	1.36 - 1.38	30

Table 1: (b) Data Details of Formosat2

Spectral Band	Wavelength (μm)	Resolution (m)
Band 1 - Blue	0.45 - 0.52	8
Band 2 - Green	0.52 - 0.60	8
Band 3 - Red	0.63 - 0.69	8
Band 4 - Near Infrared (NIR)	0.76 - 0.90	8
Band 5 - Panchromatic	0.45 - 0.90	2

The 2D scatter plot in figure 1 shows presence of non-linear data in the specified dataset. Samples taken from the site cannot easily identify individual classes indicating the presence of non-linearity or cannot be separated linearly. While using 4th band as NIR (Near Infrared), there is drastic change of reflectance energy, either increasing for vegetation case or decreasing in water case that is why while using 4th band scatter plot is non-linear.

3. Hybrid classification – Kernel based noise classification with MRF Models

Incorporating contextual features with fuzzy based classifiers have quantified the classification. Markov Random Fields (MRF) used for modelling spatial contextual information and integrated into the objective function of the noise classifier and have shown positive impact in the classification accuracy (Harikumar, 2014). The novelty of the present work is to incorporate kernel methods with supervised Noise Clustering, and to integrate contextual MRF models with it.

3.1 Kernel methods used

The aim of kernel method is to identify a linearly separating hyperplane that separates the classes in higher dimensional feature space (Hofmann *et al.*, 2008). The feature map (ϕ), given in Eq. (3.1), is the mapping function that non-linearly maps the data to a higher dimensional feature space and the kernel function (K),

mentioned in Eq. (3.2), implicitly computes the dot product between two vectors \mathbf{x} and \mathbf{x}_i in higher dimensional feature space without explicitly transforming \mathbf{x} and \mathbf{x}_i to that higher dimensional feature space.

$$\Phi : R^p \rightarrow R^q, \text{ where } p < q \quad (3.1)$$

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x, & x_i \end{matrix}\right) = \phi(x)\phi(x_i) \quad (3.2)$$

A total of nine kernels functions have been considered in this study categorized as: four local kernels, three global kernels, spectral kernel, hypertangent kernel.

3.1.1 Local kernels

They are based on evaluation of the quadratic distance between training samples and the mean vector of the class. Only feature vectors that are close or in proximity of each other have an influence on the kernel value (Kumar, 2007). In this research, the value of the input vector was normalized between [0, 1] and thus acceptable result can be produced at "σ" equals 1. The different local kernels were defined as follows:

Radial Basis Function (RBF)

The RBF kernel is defined by exponential function as shown in equation (3.3). Here, \mathbf{x}_i is the feature vector in the data and \mathbf{v}_j is the mean vector of class j . σ determines the width of the kernel; a and b are the constants. By replacing a and b by 1 the Gaussian kernel can be obtained. In this study the value of a and b were taken to be 2 and 3 respectively (Kandpal, 2016).

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x_i, & v_j \end{matrix}\right) = e^{-\left(\frac{\|\mathbf{x}_i^a - \mathbf{v}_j^b\|^2}{2\sigma^2}\right)}, \text{ where } \sigma, a, b > 0 \quad (3.3)$$

KMOD- (Kernel with Moderate Decreasing)

KMOD is the distance based kernel function (Ayat, *et al.*, 2001) as shown in equation (3.4). It shows better result in classifying closely related datasets (highly correlated) and has shown better accuracy than Radial Basis Function (RBF) and polynomial kernel.

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x_i, & v_j \end{matrix}\right) = e^{-\left(\frac{\gamma}{\sigma^2 + \|\mathbf{x}_i - \mathbf{v}_j\|^2}\right)^{-1}}, \text{ where } \sigma, \gamma > 0 \quad (3.4)$$

The parameter γ and σ controls the decreasing speed of the kernel function and the width of the kernel respectively. In this study the value of γ was taken to be one.

Gaussian

The Gaussian kernel is a special case of radial basis function kernel (Byju, 2015), shown in equation (3.5). Here, \mathbf{x}_i is the feature vector in the image and \mathbf{v}_j is the mean vector of the class.

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x_i & v_j \end{matrix}\right) = e^{\left(\frac{\|x_i - v_j\|^2}{2\sigma^2}\right)}, \text{ where } \sigma > 0 \quad (3.5)$$

Inverse Multi-Quadratic (IMQ)

The inverse multi-quadratic kernel is defined as in equation (3.6) (Byju 2015; Kandpal, 2016). Here the value of c was taken to be one.

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x_i & v_j \end{matrix}\right) = \frac{1}{\sqrt{\|x_i - v_j\|^2 + c}}, \text{ where } c > 0 \quad (3.6)$$

3.1.2 Global kernels

In global kernels, the samples that are far away from each other have an influence on the kernel value. All the kernels which are based on the dot-product are global (Kumar, 2007). The different global kernels are as follows:

Linear kernel

Linear kernel is one of the simplest kernel functions. It is defined as the inner product of the input feature vectors, as shown in equation (3.7).

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x_i & v_j \end{matrix}\right) = x_i \cdot v_j \quad (3.7)$$

Polynomial

The polynomial kernel is a positive definite kernel i.e. each element of the kernel matrix (a kernel matrix is a $n \times n$ matrix of feature vector) is positive, shown in equation (3.8). P defines the degree of the polynomial function and c is the constant (Kandpal, 2016). In this work value of P has been taken from 1 to 4. The value of c has been taken to be zero.

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x_i & v_j \end{matrix}\right) = (x_i \cdot v_j + c)^P, \text{ where } c \geq 0 \quad (3.8)$$

Sigmoid

Sigmoid kernel is a hyperbolic tangent function, as shown in equation (3.9). The parameter α work as scaling parameter for the kernel function and defines width of the kernel. The best possible value for α and c were when $\alpha > 0$ and $c < 0$ (Byju, 2015).

$$K\left(\begin{matrix} \rightarrow & \rightarrow \\ x_i & v_j \end{matrix}\right) = \tanh(\alpha \cdot x_i \cdot v_j + c) \quad (3.9)$$

3.1.3 Spectral kernel

The spectral kernel takes into consideration the spectral signature concept (Kandpal, 2016), as shown in equation (3.10). These kernels are based on the use of spectral angle (α) to measures the distance between the feature vector x and the mean vector of the class v_i . It is expressed as follows:

$$\alpha(x, v_i) = \arccos\left(\frac{(x \cdot v_i)}{\|x\| \|v_i\|}\right) \quad (3.10)$$

3.1.4 Hyper tangent kernel

The hyper tangent kernel is a hyperbolic tangent function, as shown in equation (3.11). The adjustable parameter σ defines the width or the scale of the kernel. Here x and v_i are the feature vectors in the data. It has been seen that the hyper tangent kernel outperforms other kernels when applied to a large data set (Kandpal, 2016).

$$K(x, v_i) = 1 - \tanh\left(-\frac{\|x - v_i\|^2}{\sigma^2}\right) \quad (3.11)$$

3.2 Markov Random Field (MRF) models

Contextual information refers to the relationship of an entity with its neighbourhood and in context of an image pixel; it refers to the information obtained from the neighbourhood pixels. Proper use of context can improve the classification accuracy (Jackson and D. A. Landgrebe, 2002; Solberg *et al.*, 1996; Tso and Mather, 2009; Magnussen *et al.*, 2004). Markov Random Field (MRF) is a useful tool for modelling the contextual information and widely used to image segmentation and restoration problem (Besag, 1974; Li, 2009).

Study over MRF Models have been accomplished and propagated stating the relevance of neighbourhood pixel with local interaction (Harikumar, 2014). A prior in an image context, refers to the information about the image data available beforehand. Analytical regularizers are used for representing the prior energy. The general form of the regularizer is given in equation 3.12.

$$U(f) = \sum_{n=1}^N U_n(f) = \sum_{n=1}^N \lambda_n \int_n^b g(f^{(n)}(x)) dx \quad (3.12)$$

Here $U(f)$ is the prior energy represented using the n th order regularizers, $g(f^{(n)}(x))$ is the Potential function that in turn is dependent on the irregularity in $f^{(n-1)}(x)$, N is the highest order considered and λ_n is the weighting factor and is always greater than or equal to 0. Over smoothening of the boundaries can lead to blurred image boundary, therefore, to control smoothening the Adaptive Potential Function (APF) placed within the regularizers and hence four different APFs have been used and hence four DA models. Table 2 demonstrates the mathematical models of the MRF Models to study in the present work.

Table 2: Mathematical Design of the priors (Li 1995; Li, 2009; Harikumar,2014).

Contextual Model	Mathematical Design of Priors
Smoothness Prior - S	$g(f^{(n)}(x)) = g(\eta) = \eta^2$
Discontinuity Adaptive Prior (Type 1)- DA1	$g_{1\gamma}(\eta) = (-\gamma e)^{-\frac{\eta^2}{\gamma}}$
Discontinuity Adaptive Prior (Type 2) - DA2	$g_{2\gamma}(\eta) = \frac{-\gamma}{1 + \frac{\eta^2}{\gamma}}$
Discontinuity Adaptive Prior (Type 3)- DA3	$g_{3\gamma}(\eta) = \gamma \ln\left(1 + \frac{\eta^2}{\gamma}\right)$
Discontinuity Adaptive Prior (Type 4)- DA4	$g_{4\gamma}(\eta) = \gamma \eta - \gamma^2 \ln\left(1 + \frac{\eta^2}{\gamma}\right)$

3.3 Kernel Based Noise Clustering without entropy classification (KNC)

The KNC classifier has been derived by using kernel methods with Noise clustering without entropy classifier (NC). The objective function of the NC in fuzzy mode (Dave, 1991; Hathaway *et.al*, 1996; Harikumar, 2014) expressed as shown in equation (3.13):

$$J_{NC}(U, V) = \sum_{i=1}^N \sum_{j=1}^C (\mu_{ij})^m D(\vec{x}_i, \vec{v}_j) + \sum_{i=1}^N (\mu_{i,c+1})^m \delta \quad (3.13)$$

Where $U = N \times C + 1$ matrix, $V = (v_1 \dots v_C)$, C is the number of classes, N is the total number of pixels in the image, m is the fuzzification factor and is normally positive (Sengar *et.al*, 2012), μ_{ij} represent the membership value of i^{th} pixel in the j^{th} class, $\mu_{i,c+1}$ represents the membership values of the noise class, v_j is the mean value (cluster center) of the j^{th} class, x_i is the vector value of the i^{th} pixel, D is the Euclidean distance between \vec{x}_i and \vec{v}_j and δ is a positive constant called the Noise distance.

Replacing the distance function D with (3.14), KNC objective function derives, as stated in equation (3.15) (Sengupta *et.al*. 2019).

$$D\left(\vec{x}_k, \vec{v}_i\right) = \|\varphi(x_k) - \varphi(v_i)\| = K(x_k, v_i) \quad (3.14)$$

$$J_{KNC}(U, V) = \sum_{i=1}^N \sum_{j=1}^C \mu_{ij}^m K(x_i, v_j) + \sum_{i=1}^N \mu_{i,c+1}^m \delta \quad (3.15)$$

Furthermore, KNC has been modelled with the smoothness prior, and discontinuity adaptive (DA). Thus, the hybrid classifier mentioned from equation (3.16) to (3.20) will be referred as KNC S-MRF, KNC DA1-MRF, KNC DA2-MRF, KNC DA3-MRF and KNC DA4-MRF

classifiers respectively. $U\left(\frac{u_{ij}}{d}\right)$, denotes the posterior probability, β is the weight factor associated with a pixel's neighbors and N_j represents the neighborhood window around pixel i

$$U\left(\frac{u_{ij}}{d}\right) = (1-\lambda) \left[\sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m K\left(\vec{x}_i, \vec{v}_j\right) + \sum_{i=1}^N (u_{i,c+1})^m \delta \right] + \lambda \left[\sum_{i=1}^N \sum_{j=1}^C \sum_{j' \in N_j} \beta (u_{ij} - u_{ij'})^2 \right] \quad (3.16)$$

$$U\left(\frac{u_{ij}}{d}\right) = (1-\lambda) \left[\sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m K\left(\vec{x}_i, \vec{v}_j\right) + \sum_{i=1}^N (u_{i,c+1})^m \delta \right] + \lambda \left[\sum_{i=1}^N \sum_{j=1}^C \sum_{j' \in N_j} (-\gamma e)^{-\frac{\eta^2}{\gamma}} \right] \quad (3.17)$$

$$U\left(\frac{u_{ij}}{d}\right) = (1-\lambda) \left[\sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m K\left(\vec{x}_i, \vec{v}_j\right) + \sum_{i=1}^N (u_{i,c+1})^m \delta \right] + \lambda \left[\sum_{i=1}^N \sum_{j=1}^C \sum_{j' \in N_j} \frac{-\gamma}{1 + \frac{\eta^2}{\gamma}} \right] \quad (3.18)$$

$$U\left(\frac{u_{ij}}{d}\right) = (1-\lambda) \left[\sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m K\left(\vec{x}_i, \vec{v}_j\right) + \sum_{i=1}^N (u_{i,c+1})^m \delta \right] + \lambda \left[\sum_{i=1}^N \sum_{j=1}^C \sum_{j' \in N_j} \left(\gamma \ln\left(1 + \frac{\eta^2}{\gamma}\right) \right) \right] \quad (3.19)$$

$$U\left(\frac{u_{ij}}{d}\right) = (1-\lambda) \left[\sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m K\left(\vec{x}_i, \vec{v}_j\right) + \sum_{i=1}^N (u_{i,c+1})^m \delta \right] + \lambda \left[\sum_{i=1}^N \sum_{j=1}^C \sum_{j' \in N_j} \left(\gamma|\eta| - \gamma^2 \ln\left(1 + \frac{\eta^2}{\gamma}\right) \right) \right] \quad (3.20)$$

3.4 Mean-variance method

Verifying the edge preservation is significant; a standard method to analyze separately the distributions of grey levels of the two regions on either side of the edge, where the difference between the averages values within the two regions indicates the steepness of the edge (Wen and Xia, 1999).

The membership value of a unit in a fraction image is high if the pixel exists at a location of a known class (Class A) and for the unknown class it is low (Non-Class A), elaborated in figure 2. Consequently, the mean of the membership value will be high and the variance will be low in case of a homogeneous area for a known class location in a fraction image, leading to edge preservation. This concept has been used here to verify the edge preservation, to optimize contextual parameters.

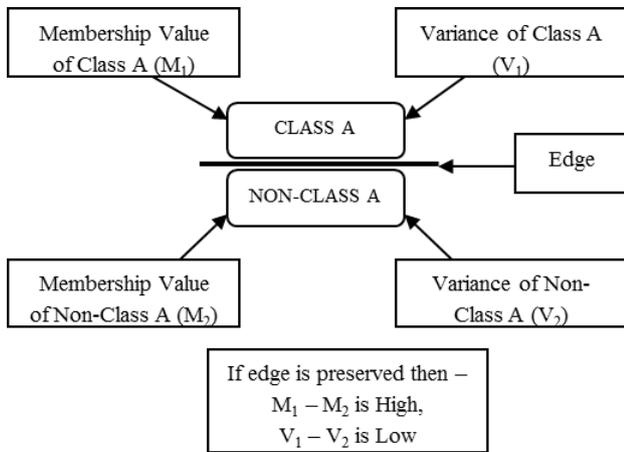


Figure 2: Method to verify edge preservation

4. Accuracy assessment techniques

To assess the discussed soft classifier, simulated image technique has been opted as well as FERM (Fuzzy error matrix) has been opted for computing the accuracy of KNC-S-MRF, KNC-DA-MRF models and NC S-MRF, NC-DA-MRF models. Accuracy assessment of sub-pixel classified output has been done with Java based tool (Kumar et.al, 2006).

4.1 Simulated image technique

The simulated image technique has been introduced to evaluate the fuzzy based classifier behaviour. The concept is to assign membership values to feature vectors from mean vector of the classes on the basis of distance measure. It is generated on the sample data for each class with desired number of bands. The technique facilitates to compare the classifier output with known input over defined location and also makes easy to identify the behaviour of classifier with the mixed pixels. Distribution is such that, the membership values of pure pixel in the classified output of a class must be maximized (close to 1). The mixed pixels were simulated with two variations, one with composition of 50:50 in between two different classes and other with composition of 30:30:40, the target membership value of 0.50, 0.40 and 0.30 is expected from the pixel with 50%, 40% and 30% belongingness for a class respectively (Figure 3).

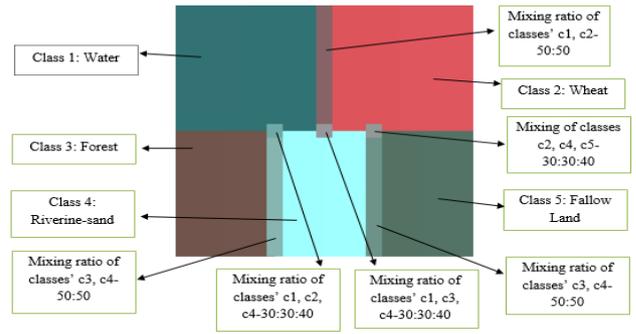


Figure 3: Simulated Image of Formosat 2 (Class Distribution)

4.2 FERM (Fuzzy Error Matrix)

It is a square array of positive fractional value varying between [0, 1]. The column R_N usually represent the sample elements assigned to the reference class n while the rows indicate the sample elements assigned to the classified class m (Binaghi et al., 1999). The element in fuzzy error matrix (M) at row m and column n for a feature vector x is computed as shown in equation (4.1).

$$M(m, n) = \sum_{x \in X} \min(\mu_{C_m}(x), \mu_{R_n}(x)) \quad (4.1)$$

In equation (4.1), x is the overall sampled data set. μ_{C_m} and μ_{R_n} are the membership values for the classified and referenced data. The "min" operator is the traditional fuzzy set operator, it returns the minimum membership value between the classified and referenced data set for a class.

5. Results and analysis

5.1 Parameter estimation

The objective function of all KNC S-MRF and KNC DA-MRF classifiers involves certain parameters, which need to be initialized before the optimization of the membership values, thus, implementation of this hybrid classifier has been done in Java. Base classifier estimation has been done by a series of classification upon simulated image, using different kernels for every combination of m ranging between [1.1, 5.0] and the resolution parameter, δ , taken in the range of 10 to 10^6 . Hybrid parameters have been estimated through simulated annealing (Bertsimas and Tsitsiklis, 1993) and mean variance method, initial T_0 has been set to 3 where optimized final temperature has been taken to be 0.90; λ has been defined in the range between 0 and 1, range of β to be 1 to 100, and that of γ in between 0 and 1.

5.1.1 Base classifier parameter estimation

A series of kernel-based classification has been applied upon simulated image with every combination of defined m and δ . For brief demonstration, figure 4, displays the kernel wise membership values of wheat class, with varying δ , where it stops increasing at $\delta=10^4$, here, n is representing the degree of resolution parameter and is related as $\delta=10^n$. Remaining classes have also shown similar behaviour. Similarly, membership values

have been computed for both pure and mixed pixel variations to optimize m . Table 3 demonstrates kernel wise membership values across varying m from 1.5 to 5.0. The tabular representation consists of three variations of pure pixel and mixed pixels as elaborated in section 4.1. Variation in optimized value of m have been seen from kernel to kernel therefore a specific range has been

defined as an optimal range of m that is found to be [2.7, 5.0], leading to stability of membership values. Polynomial Kernels with degree 1 to 4 have shown poor performance leading to minimal membership values close to 0. Membership values computed, as shown in figure 4 and table 3 supports integer based values, hence, small fractional values rounded off to zero.

Table 3: Kernel wise membership value representation of wheat class with varying fuzzification factor (m)
(a) Pure Pixel Composition (b) Mixed Pixel Composition (50:50) (c) Mixed Pixel Composition(30:30:40)

(a)

Kern els → m↓	Linea r	Hype rtang ent	Gaus sian	Sigm oid	KMO D	IMQ	Radi al	Spect ral	P1	P2	P3	P4
1.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
2.5	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.10	0.00	0.00
2.7	0.98	0.98	0.98	0.99	0.98	0.98	0.98	0.89	0.98	0.12	0.00	0.00
3	0.94	0.93	0.93	0.93	0.92	0.93	0.93	0.93	0.94	0.16	0.00	0.00
3.5	0.86	0.86	0.86	0.86	0.85	0.91	0.86	0.86	0.86	0.20	0.00	0.00
4	0.78	0.79	0.79	0.79	0.77	0.78	0.78	0.75	0.78	0.24	0.00	0.00
4.5	0.71	0.72	0.72	0.72	0.70	0.71	0.71	0.69	0.71	0.25	0.01	0.00
5	0.64	0.65	0.66	0.66	0.64	0.65	0.65	0.64	0.64	0.26	0.02	0.00

*P1 – Polynomial (Degree 1), P2 – Polynomial (Degree 2),
P3 – Polynomial (Degree 3), P4 – Polynomial (Degree 4)

(b)

Kern els → m↓	Linea r	Hype rtang ent	Gaus sian	Sigm oid	KMO D	IMQ	Radia l	Spect ral	P1	P2	P3	P4
1.5	0.23	0.36	0.36	0.04	0.21	0.35	0.35	0.05	0.23	0.00	0.00	0.00
2	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.00	0.00	0.00
2.5	0.20	0.25	0.26	0.15	0.25	0.21	0.25	0.16	0.20	0.00	0.00	0.00
2.7	0.19	0.21	0.22	0.15	0.22	0.20	0.20	0.16	0.19	0.00	0.00	0.00
3	0.19	0.24	0.24	0.17	0.22	0.23	0.23	0.18	0.19	0.00	0.00	0.00
3.5	0.19	0.24	0.24	0.17	0.22	0.23	0.23	0.18	0.19	0.00	0.00	0.00
4	0.18	0.22	0.23	0.18	0.22	0.22	0.22	0.18	0.18	0.01	0.00	0.00
4.5	0.18	0.22	0.22	0.18	0.21	0.22	0.22	0.18	0.18	0.02	0.00	0.00
5	0.16	0.22	0.23	0.18	0.22	0.22	0.22	0.18	0.16	0.04	0.00	0.00

*P1 – Polynomial (Degree 1), P2 – Polynomial (Degree 2),
P3 – Polynomial (Degree 3), P4 – Polynomial (Degree 4)

(c)

Kern els → m↓	Linea r	Hype rtang ent	Gaus sian	Sigm oid	KMO D	IMQ	Radia l	Spect ral	P1	P2	P3	P4
1.5	0.01	0.02	0.02	0.02	0.04	0.03	0.03	0.01	0.01	0.00	0.00	0.00
2	0.05	0.07	0.07	0.07	0.07	0.07	0.07	0.05	0.05	0.00	0.00	0.00
2.5	0.11	0.13	0.14	0.14	0.15	0.14	0.14	0.11	0.11	0.00	0.00	0.00
2.7	0.11	0.15	0.14	0.16	0.15	0.15	0.15	0.14	0.11	0.00	0.00	0.00
3	0.13	0.15	0.16	0.15	0.16	0.16	0.16	0.13	0.13	0.00	0.00	0.00
3.5	0.14	0.16	0.16	0.16	0.17	0.16	0.16	0.15	0.14	0.01	0.00	0.00
4	0.14	0.17	0.17	0.16	0.17	0.17	0.17	0.15	0.14	0.02	0.00	0.00
4.5	0.15	0.17	0.18	0.17	0.18	0.18	0.18	0.16	0.15	0.03	0.00	0.00
5	0.15	0.18	0.18	0.17	0.18	0.18	0.18	0.16	0.15	0.04	0.00	0.00

*P1 – Polynomial (Degree 1), P2 – Polynomial (Degree 2),
P3 – Polynomial (Degree 3), P4 – Polynomial (Degree 4)

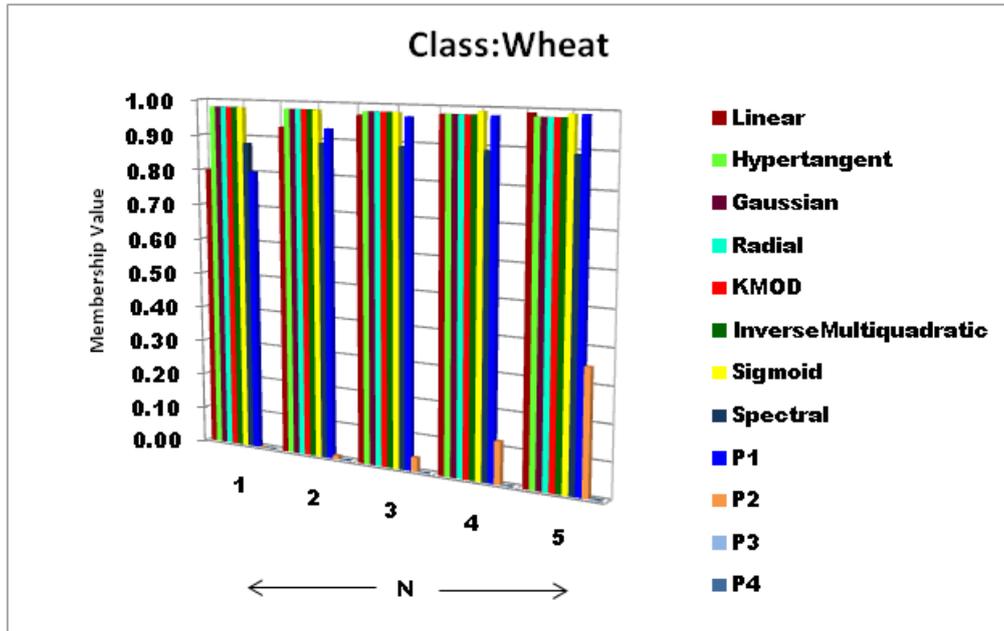


Figure 4: Kernel wise membership value representation of wheat class in accord with Noise Distance ($\delta = 10^N$). *P1 – Polynomial (Degree 1), P2 – Polynomial (Degree 2), P3 – Polynomial (Degree 3), P4 – Polynomial (Degree 4)

5.1.2 Contextual parameter estimation

Contextual parameters include, the weight factor which controls the spatial and spectral component (λ), neighborhood weight in the case of in case of S-MRF models (β) and constant involved in the case DA model (γ). Estimation has been done upon the fractional images of KNC-S-MRF and KNC DA-MRF

classification for Landsat8. Table 4 displays the optimal range of hybrid parameters in case of KNC-S-MRF found to be λ lying between 0.7to 0.9, $\beta=7$ to 20, similarly in KNC DA1-MRF $\lambda=0.1$ with $\gamma=0.1$, KNC DA2-MRF $\lambda=0.7$ with $\gamma=0.7$ to 0.9, KNC DA3-MRF $\lambda=0.7, 0.8$ with $\gamma=0.6$ to 0.9 and KNC DA4-MRF $\lambda=0.9$ with $\gamma=0.8, 0.9$.

Table 4: a) KNC estimation over edge verification for Landsat8 data (Class Water) (Part – 1)

Classifier → Contextual Models ↓	Hybrid Parameters		Linear		Hypertangent		Gaussian		Radial		KMOD		Multiquadratic	
	λ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	-1	0	174	103	176	135	169	149	156	200	167	149
DA1-MRF	0.1	0.1	56	467	179	39	181	44	175	48	162	70	173	47
DA2-MRF	0.8	0.8	0	0	188	-25	187	-41	183	-49	169	21	185	-53
DA3-MRF	0.7	0.8	0	0	172	108	177	125	166	134	153	208	167	157
DA4-MRF	0.8	0.9	0	0	170	23	178	-89	170	-122	90	7784	165	55

(Part - 2)

Classifier → Contextual Models ↓	Hybrid Parameters		Sigmoid		Spectral		Polynomial Degree=1		Polynomial Degree=2		Polynomial Degree=3		Polynomial Degree=4		NC (Euclidean)	
	β/γ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	148	575 2	132	387	-1	0	0	0	0	0	0	0	-1	0
DA1-MRF	0.1	0.1	184	15	133	142	56	467	0	0	0	0	0	0	56	467
DA2-MRF	0.8	0.8	182	39	134	138	0	0	0	0	0	0	0	0	0	0
DA3-MRF	0.7	0.8	177	124	134	44	0	0	0	0	0	0	0	0	0	0
DA4-MRF	0.8	0.9	60	848 3	92	538 2	0	0	0	0	0	0	0	0	0	0

*MD – Mean Difference, VD – Variance Difference

Table 4: b) KNC estimation over edge verification for Landsat8 data (Class Wheat)
(Part – 1)

Classifier → Contextual Models ↓	Hybrid Parameters		Linear		Hypertangent		Gaussian		Radial		KMOD		Multiquadratic	
	λ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	0	0	155	103	158	140	151	148	137	159	150	150
DA1-MRF	0.1	0.1	25	113	164	-16	167	0	161	-4	147	2	159	3
DA2-MRF	0.8	0.8	-1	0	154	2	153	171	149	112	85	6134	145	194
DA3-MRF	0.7	0.8	-1	0	158	3	158	89	158	106	145	117	160	56
DA4-MRF	0.8	0.9	-1	0	148	77	140	-150	141	181	134	302	140	186

(Part – 2)

Classifier → Contextual Models ↓	Hybrid Parameters		Sigmoid		Spectral		Polynomi al Degree=1		Polynomi al Degree=2		Polynomi al Degree=3		Polynomi al Degree=4		NC (Euclidean)	
	λ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	165	161	52	3298	0	0	0	0	0	0	0	0	0	0
DA1-MRF	0.1	0.1	178	57	114	126	25	113	0	0	0	0	0	0	25	113
DA2-MRF	0.8	0.8	135	4814	117	165	0	0	0	0	0	0	0	0	-1	0
DA3-MRF	0.7	0.8	174	45	106	197	0	0	0	0	0	0	0	0	-1	0
DA4-MRF	0.8	0.9	163	179	106	195	0	0	0	0	0	0	0	0	-1	0

*MD – Mean Difference, VD – Variance Difference

Table 4:c) KNC estimation over edge verification for Landsat8 data (Class Forest)
(Part – 1)

Classifier → Contextual Models ↓	Hybrid Parameters		Linear		Hypertangent		Gaussian		Radial		KMOD		Multiquadratic	
	λ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	0	0	178	54	182	61	176	76	166	98	176	73
DA1-MRF	0.1	0.1	48	305	178	54	181	61	176	76	166	98	175	72
DA2-MRF	0.8	0.8	-1	0	178	54	181	61	176	76	166	98	175	72
DA3-MRF	0.7	0.8	0	0	182	-4	186	-18	180	8	171	58	174	36
DA4-MRF	0.8	0.9	0	0	190	-106	193	-16	182	73	167	268	185	67

(Part – 2)

Classifier → Contextual Models ↓	Hybrid Parameter s		Sigmoid		Spectral		Polynomi al Degree=1		Polynomi al Degree=2		Polynomi al Degree=3		Polynomi al Degree=4		NC (Euclidean)	
	λ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	181	64	92	98	53	0	0	0	0	0	0	0	0	0
DA1-MRF	0.1	0.1	179	74	92	98	48	305	0	0	0	0	0	0	48	305
DA2-MRF	0.8	0.8	179	74	92	98	-1	0	0	0	0	0	0	0	-1	0
DA3-MRF	0.7	0.8	194	-88	104	37	0	0	0	0	0	0	0	0	0	0
DA4-MRF	0.8	0.9	195	320	129	161	0	0	0	0	0	0	0	0	0	0

*MD – Mean Difference, VD – Variance Difference

Table 4: d) KNC estimation over edge verification for Landsat8 data (Class Riverine)**(Part - 1)**

Classifier → Contextual Models ↓	Hybrid Parameters		Linear		Hypertangent		Gaussian		Radial		KMOD		Multiquadratic	
	λ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	-1	0	183	119	195	95	181	141	163	207	181	138
DA1-MRF	0.1	0.1	32	117	189	117	200	114	187	149	168	206	188	148
DA2-MRF	0.8	0.8	0	0	183	122	195	121	182	154	163	200	184	228
DA3-MRF	0.7	0.8	0	0	175	48	186	56	173	97	154	165	174	80
DA4-MRF	0.8	0.9	0	0	144	5863	208	-83	111	10433	126	5108	79	10864

(Part - 2)

Classifier → Contextual Models ↓	Hybrid Parameters		Sigmoid		Spectral		Polynomi al Degree=1		Polynomi al Degree=2		Polynomi al Degree=3		Polynomi al Degree=4		NC (Euclidean)	
	λ	β/γ	MD	VD	MD	V D	MD	V D	MD	V D	MD	V D	MD	V D	MD	V D
S-MRF	0.8	20	20 1	- 451	118	-73	53	0	0	0	0	0	0	0	-1	0
DA1-MRF	0.1	0.1	19 9	- 492	118	-80	32	11 7	0	0	0	0	0	0	32	11 7
DA2-MRF	0.8	0.8	21 8	75	115	-167	0	0	0	0	0	0	0	0	0	0
DA3-MRF	0.7	0.8	20 6	- 180	126	18 2	0	0	0	0	0	0	0	0	0	0
DA4-MRF	0.8	0.9	21 8	- 184	130	11 8	0	0	0	0	0	0	0	0	0	0

*MD – Mean Difference, VD – Variance Difference

Table 4: e) KNC estimation over Edge Verification for Landsat8 data (Class Fallow)**(Part - 1)**

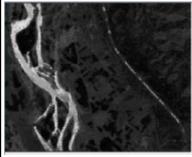
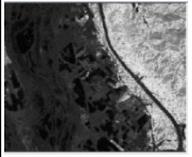
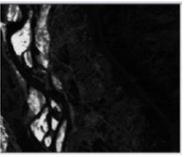
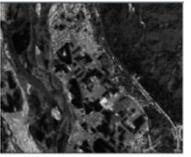
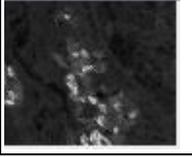
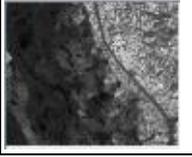
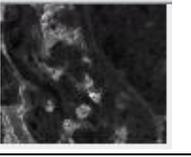
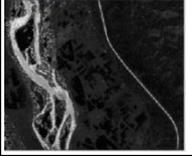
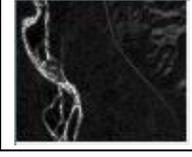
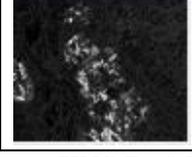
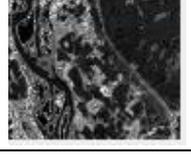
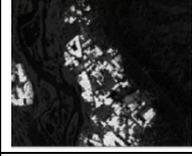
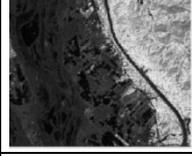
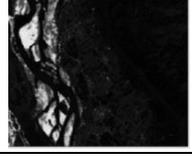
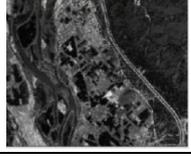
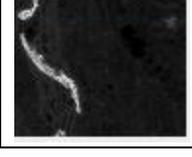
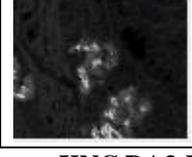
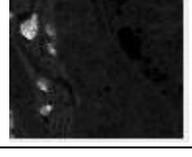
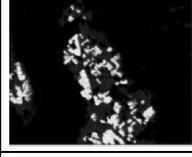
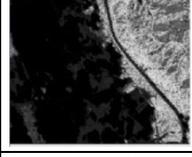
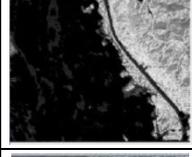
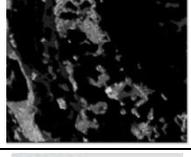
Classifier → Contextual Models ↓	Hybrid Prameters		Linear		Hypertangent		Gaussian		Radial		KMOD		Multiquadratic	
	λ	β/γ	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	0	0	132	4315	156	235	152	236	137	276	119	4543
DA1-MRF	0.1	0.1	20	55	162	115	162	138	156	161	142	172	154	141
DA2-MRF	0.8	0.8	0	0	170	75	165	102	167	126	67	6783	164	70
DA3-MRF	0.8	0.8	0	0	168	283	168	367	163	463	150	646	161	450
DA4-MRF	0.8	0.9	0	0	97	10454	129	11975	62	11023	-11	100	56	11030

(Part - 2)

Classifier → Contextua l Models ↓	Hybrid Prameter s		Sigmoid		Spectral		Polynomia l Degree=1		Polynomia l Degree=2		Polynomia l Degree=3		Polynomia l Degree=4		NC (Euclidean)	
	λ	β/γ	M D	VD	M D	VD	MD	VD	MD	VD	MD	VD	MD	VD	MD	VD
S-MRF	0.8	20	158	24	19	637	0	0	0	0	0	0	0	0	0	0
DA1-MRF	0.1	0.1	168	49	63	111	20	55	0	0	0	0	0	0	20	55
DA2-MRF	0.8	0.8	170	-281	46	101 9	0	0	0	0	0	0	0	0	0	0
DA3-MRF	0.8	0.8	167	94	64	247 4	0	0	0	0	0	0	0	0	0	0
DA4-MRF	0.8	0.9	172	14 5	3	849	0	0	0	0	0	0	0	0	0	0

*MD – Mean Difference, VD – Variance Difference

Table 5: Fractional images obtained from KNC S-MRF and KNC DA-MRF classifiers on Formosat 2 and Landsat 8 against the optimal parameters

Classifier	Water	Wheat	Forest	Riverine Sand	Fallow
KNC S-MRF					
Gaussian (Formosat2)					
Gaussian(Landsat8)					
KNC DA1-MRF					
Sigmoid (Formosat2)					
Sigmoid (Landsat8)					
KNC DA2-MRF					
Hypertangent (Formosat2)					
Hypertangent (Landsat8)					
KNC DA3-MRF					
Sigmoid (Formosat2)					
Sigmoid(Landsat8)					
KNC DA4-MRF					
Hypertangent (Formosat2)					
Hypertangent(Landsat8)					

5.2 Accuracy assessment

Studied contextual classification has been applied upon Landsat8 and Formosat2 image and fractional images generated with the optimized base, as shown in Table 5. FERM based accuracy assessment has been done using the output of Landsat8 and output of Formosat2 as reference map for KNC S-MRF and for KNC DA1-MRF, KNC D2-MRF, KNC DA3-MRF, KNC DA4-MRF models using all nine kernels. Table 6 shows the overall accuracy of best performing kernels along with the overall accuracy of conventional NC based MRF models. Analyzing the performances of various kernels with these contextual models, KNC DA1-MRF is found to be more promising from the perspective of classification accuracy. Overall accuracy of kernel based contextual models is better than the conventional Euclidean distance based contextual models.

Table 6: Accuracy assessment results for MRF based NC and MRF based KNC using single kernel

CLASSIFIER	CONTEXTUAL MODELS				
	S-MRF	DA1-MRF	DA2-MRF	DA3-MRF	DA4-MRF
NC (Euclidean)	4.00%	6.64%	0.79%	0.91%	0.90%
Gaussian Kernel	76.63%	81.09%	70.38%	59.42%	52.59%
Sigmoid	73.51%	75.38%	66.30%	51.04%	37.97%
Hypertangent	75.95%	78.75%	70.09%	60.91%	50.85%

6. Conclusions

The study focused to realize handling of non-linearity between class boundaries by integrating spatial features with kernel based noise classifier. This model takes account of involving MRF models into supervised kernel based Noise classifier (KNC). KNC S-MRF (smoothness prior) and four different Discontinuity Adaptive KNC DA1-MRF, KNC D2-MRF, KNC DA3-MRF, and KNC DA4-MRF models have been introduced and experimented to characterize both the contextual as well as spectral information. From the experiments performed we found that the KNC DA1-MRF model has performed better than the remaining models. Gaussian Kernel followed by Hypertangent Kernel has shown better output amongst the nine kernels. The study concludes that contextual MRF models when associated with kernel based have shown increase in accuracy as compared to Noise Classifier with Euclidean distance.

References

Ayat, N. E., M. Cheriet, L. Remaki and C. Y. Suen (2001). KMOD-A new support vector machine kernel with moderate decreasing. In Document Analysis and Recognition, (2001). Proceedings, Sixth International Conference, pp.1215–1219.

Awan, A. M., and M. N. M. Sap (2005). Clustering spatial data using a kernel-based algorithm. In Proceedings of the Annual Research Seminar, 306–310

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society. Series B (Methodological), 192-236.

Bhatt, S. R. and P. K. Mishra (2013). Study of local Kernel with fuzzy C mean algorithm. International Journal of Advanced Research in Computer & Software Engineering, 3(12), 636–639.

Binaghi, E., P. Madella, M. Grazia Montesano and A. Rampini (1997). IEEE Transactions on Geoscience and Remote Sensing, 35, 326-340.

Binaghi, E., P. A. Brivio, P. Chessi and A. Rampini (1999). A fuzzy set-based accuracy assessment of soft classification. Pattern Recognition Letters, 20(9), 935–948.

Byju, A. P. (2015). Non-Linear separation of classes using a Kernel based Fuzzy c -Means (KFCM) Approach. MSc. Thesis, ITC, University of Twente, The Netherlands.

Chawla, S. (2010). Possibilistic-c-Means-Spatial contextual information based sub-pixel classification approach for multi-spectral data. MSc. Thesis, University of Twente, Faculty of Geo-Information and Earth Observation (ITC), Enschede.

Choti wattana, W. (2009). Noise Clustering algorithm based on Kernel method. Advance Computing Conference (2009). IACC 2009. IEEE, pp. 56-60

Dave, R. N. (1991). Characterization and detection of noise in clustering. Pattern Recognition Letters, 12, 657-664.

Dave, R. N. (1993). Robust fuzzy clustering algorithms. Second IEEE International Conference on Fuzzy Systems 1993, pp. 1281-1286.

Dave, R. and S. Sen (1997). Noise clustering algorithm revisited. Fuzzy Information Processing Society, NAFIPS'97. Annual Meeting of the North American, IEEE, pp.199-204.

Fisher, P. F and S. Pathirana (1990). The evaluation of fuzzy membership of land cover classes in the suburban zone. Remote Sensing of Environment, 34, 121-132.

Foody, G. M. (1996). Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data, International Journal of Remote Sensing, 17(7), 1317-1340.

Foody, G. M. (2000). Estimation of sub-pixel land cover composition in the presence of untrained classes. Computers & Geosciences, 26(4), 469-478.

Ganesan, P and V. Rajini (2010). A method to segment color images based on modified Fuzzy-Possibilistic-C-Means clustering algorithm. Recent Advances in Space Technology Services and Climate Change 2010 (RSTS

- and CC-2010) , 157–163, 2010.[http://doi.org/ 10.1109/RSTSCC.2010.5712837](http://doi.org/10.1109/RSTSCC.2010.5712837).
- Harikumar, A. (2014). The effects of discontinuity adaptive MRF models on the Noise classifier. MSc. Thesis, ITC, University of Twente, The Netherlands.
- Hathaway, R. J., J. C. Bezdek and W. Pedrycz (1996). A parametric model for fusing heterogeneous fuzzy data, *IEEE Transactions on Fuzzy Systems*, 4, 1277-1282.
- Hofmann, T., B. Scholkopf and A. J. Smola (2008). Kernel methods in machine learning. *The Annals of Statistics*, 36(3), 1171–1220.
- Ibrahim, M. A., M. K. Arora, and S. K. Ghosh (2005). Estimating and Accommodating Uncertainty through the Soft Classification of Remote Sensing Data. *International Journal of Remote Sensing*, 26(14), 2995–3007.
- Jackson Q. and D. A. Landgrebe (2002). Adaptive Bayesian contextual classification based on Markov random fields, *IEEE Transactions on Geoscience and Remote Sensing*, 40(11), 2454 – 2463.
- Jolion, J. M. and A. Rosenfeld (1989). Cluster detection in background noise. *Pattern Recognition*, 22, 603-607.
- Kandpal, N. (2016). Non-Linear Separation of classes using a Kernel based Possibilistic c -Means. MSc. Thesis, ITC, University of Twente, The Netherlands.
- Krishnapuram, R. and C. P. Feg (1992). Fitting an unknown number of lines and planes to image data through compatible cluster merging. *Pattern Recognition*, 25, 385-400.
- Krishnapuram, R. and J. M. Keller (1996). The possibilistic c-means algorithm: insights and recommendations. *Fuzzy Systems, IEEE Transactions*, 4, 385-393.
- Kumar, A., S. Ghosh and V. Dadhwal (2006). Sub-pixel land cover mapping: SMIC system. *ISPRS International Symposium “Geospatial Databases for Sustainable Development”*, Goa, India, September 27-30, 2006.
- Kumar, A. (2007). Investigation in sub-pixel classification approaches for land use and land cover mapping. Ph.D. Thesis. IIT Roorkee, India.
- Li, S. Z. (1995). On discontinuity-adaptive smoothness priors in computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(6), 576-586.
- Li, S. Z. (2009). Markov random field modeling in image analysis, *Advances in Computer Vision and Pattern Recognition*, Springer.
- Magnussen, S., P. Boudewyn and M. Wulder (2004). Contextual classification of Landsat TM images to forest inventory cover types. *International Journal of Remote Sensing*, 25 (12), 2421–2440.
- Moser, G. and B. S. Serpico (2010). Contextual remote-sensing image classification by support vector machines and Markov random fields. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3728-3731, 2010
- Pham, D (2001). Spatial models for fuzzy clustering. *Computer vision and image understanding. Computer Vision and Image Understanding*, 84(2), 285-297.
- Ravindraiah, R., and K. Tejaswini (2013). A survey of image segmentation algorithms based on Fuzzy clustering. *International Journal of Computer Science and Mobile Computing*, 2(7), 200–206.
- Rhee, F., K. Choi and B. Choi (2012). Kernel approach to possibilistic C-means clustering. *International Journal of Intelligent Systems*, 24(3), 272–292. <http://doi.org/10.1002/int.20336>.
- SenGupta, I., A. Kumar and R. K. Dwivedi (2019). Performance Evaluation of Kernel-Based supervised noise clustering approach, *Journal of Indian Society of Remote Sensing.*, 47(2), 317–330.<https://doi.org/10.1007/s12524-019-00938-2>
- Solberg, A. H., T. Taxt and A.K. Jain (1996). A Markov random field model for classification of multisource satellite imagery. *Geoscience and Remote Sensing, IEEE Transactions*, 34 (1), 100-113.
- Tso, B. and P. M. Mather (2009). *Classification of remotely sensed data*. Second edition. Boca Raton, CRC, Press.
- Wen, W. and A. Xia (1999). Verifying edges for visual inspection purposes, *Pattern recognition letters*, 20(3), 315–328.
- Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 100, 9–34.
- Zhang, J. and G. M. Foody (2001). A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing*, 19(14), 2721-2738.

Empowering the Electoral Planning and Monitoring System (EPMS) using GeoWeb tools and services – a case study on the Meghalaya legislative assembly election, 2018

Dibyajyoti Chutia*, Nilay Nishant, Siddhartha Bhuyan, P. Subhash Singh, Avinash Chouhan, Ritu Anil, Victor Saikhom, Manoj Lokare and P.L.N. Raju

North Eastern Space Applications Centre, Department of Space, Government of India,
Umiam, Shillong, Meghalaya, India- 793103

*Email: d.chutia@nesac.gov.in

(Received: Sep 14, 2018; in final form: June 16, 2019)

Abstract: This paper showcases the effective utilization of Information and Communication Technology (ICT) and Geospatial Technology based tools and services to empower the electoral planning and monitoring activities in the State of Meghalaya. The aim of the Electoral Planning and Monitoring System (EPMS) is to provide actionable inputs during pre-election preparation and also during the day polls. Live Geo-visualization of polling updates and turnout at the control room on the day of election is the vital component of system. It also generates the live heat map based on the live feed of polling updates from the various polling stations. EPMS is one of the unique initiatives of North Eastern Space Applications Centre (NESAC), Umiam, Meghalaya executed for the Office of the Chief Electoral Officer (CEO) of Meghalaya. The application was effectively utilized for the first time in the country during the last Meghalaya Legislative Assembly Election, 2018.

Keywords: ICT, geospatial technologies, GeoWeb Tools and Services, Electoral System

1. Introduction

India is the largest democracy of the world with the second largest population. In such a democratic system, time bound free and fair elections are compulsory phenomena. Elections decide the composition of the government, the membership of the two houses of the parliament, the state and the union territory legislative assemblies, the Presidency and the vice-presidency (ECI portal). Elections for the Lok Sabha and every State Legislative Assembly are required to take place every five years, unless called earlier by the Election Commission of India (ECI) due to extra-ordinary circumstances. Conducting elections for either the Lok Sabha or Vidhan Sabha in a country as populous as India with tremendous diversity in social, cultural and topographical levels is a hectic exercise. It demands involvement of large human resources with government machinery and technical endeavours. Currently, advanced technological tools are being employed in the election planning, monitoring and result dissemination. The ECI has been exploring the modern Information and Communication Technology (ICT) tools to improve the fidelity of the electoral rolls (Chutia 2011). A number of web applications and including Mobile Apps are already developed for ECI for planning electoral activities. For example, Returning Officer (RO) Net of ECI was tuned by the Chief Electoral Officer (CEO), Punjab for assisting Returning Officers in the arc of their duties right from the nomination process till the counting of votes. Some other instance is the ECI 360 degree mobile app used in the general elections 2016 as a prototype to share necessary information (e.g. Advancement of the nomination process), facilitate interactions among the electoral officials, including political parties and media during the election. Suraksha is another important App of ECI to provide details security related aspects during the election. The ECI Apps Suite is another step of ECI during 2018 towards empowering the electoral process in a democratic manner. In addition, few studies in the applications of geospatial technologies have been reported

towards planning and management of election in India. A list of some applications that have been created to aid the election planning and monitoring process are presented in table 1. The Election Commission portal is made available via Bhuvan geoportal of Indian Space Research Organization (ISRO) to provide critical information on the polling stations in terms of 17 important parameters as function of pre-poll requirements. It eases the tool for vehicle tracking and incident reporting. Users can upload the information on basic minimum facility. Heat map algorithm (Susewind 2014) was proposed for demarcation of constituency boundaries using polling booth datasets published by ECI. Election Management System (Chutia 2011) developed for the state of Meghalaya as per the suggestion of CEO, Meghalaya is a spatial decision support system (SDSS) on electoral process. It provides the optimal locations of polling stations in the spatial domain as per the guidelines of ECI.

Nevertheless, none of the applications were synergized together with ICT and geospatial technology for real time planning and monitoring of the electoral scheme. Hence, considering the updated database of EMS (Chutia 2011) as a foundation, we conceived a novel EPMS as per the suggestions of CEO, Meghalaya to plan the last Meghalaya Legislative Assembly Election, 2018 more effectively.

2. Methodology

Proposed Electoral Planning and Monitoring System (EPMS) has three major modules; 1) e-Atlas Manager, 2) Live Dashboard and 3) Updation module. The e-Atlas module of EPMS is the core of the application with different capabilities and functionalities and internetworked with the other modules. Availability of geospatial information and the services of the modules are synoptically illustrated in figure 1. The capabilities and functionalities of each module are as listed in table-2. The landing page of e-Atlas Manager of EPMS hosted at election.nedrp.gov.in (NEDRP portal) is shown in figure

2. It provides an interactive and responsive user interface (UI) for visualization of geospatial layers, on the fly-statistics with proximity and multi-dimensional querying capability. Live Dashboard and Updation modules are hyperlinked from the e-Atlas Manager. Live Dashboard is one of the powerful and unique modules of the entire application to offer live visualization of polling updates and turn out at the time of the Election Day.

It also provides the heat map representing the occurrences polls based on the live feed of polling updates from the various polling stations. Updating module is another important tool for updating the details of polling stations or the addition of any new polling station in spatial domain without interacting with any GIS software or Application Programming Interface (API). The result of a multidimensional query is depicted in figure 3 as an example

Table 1: List of few ICT and geospatial based applications developed for electoral planning

Previous applications	Domain	Service features	Goals
RO Net	ICT	Web, MobileApp & SMS	DSS on electoral planning
ECI 360	ICT	MobileApp	Overall knowledge base on electoral process
Suraksha App	ICT	MobileApp	DSS on security related issues
EMS	ICT & Geospatial Tech.	WebGIS	SDSS on electoral planning & optimization of polling station locations
Bhuvan	ICT & Geospatial Tech.	WebGIS and MobileApp	SDSS on Important pre-poll requirements
Susewind R. (2014)	Geospatial Tech.	GIS	Demarcation of boundary for constituencies
ECI Apps Suite 2018	ICT	MobileApp	Single window platform of all ECI web and MobileApps

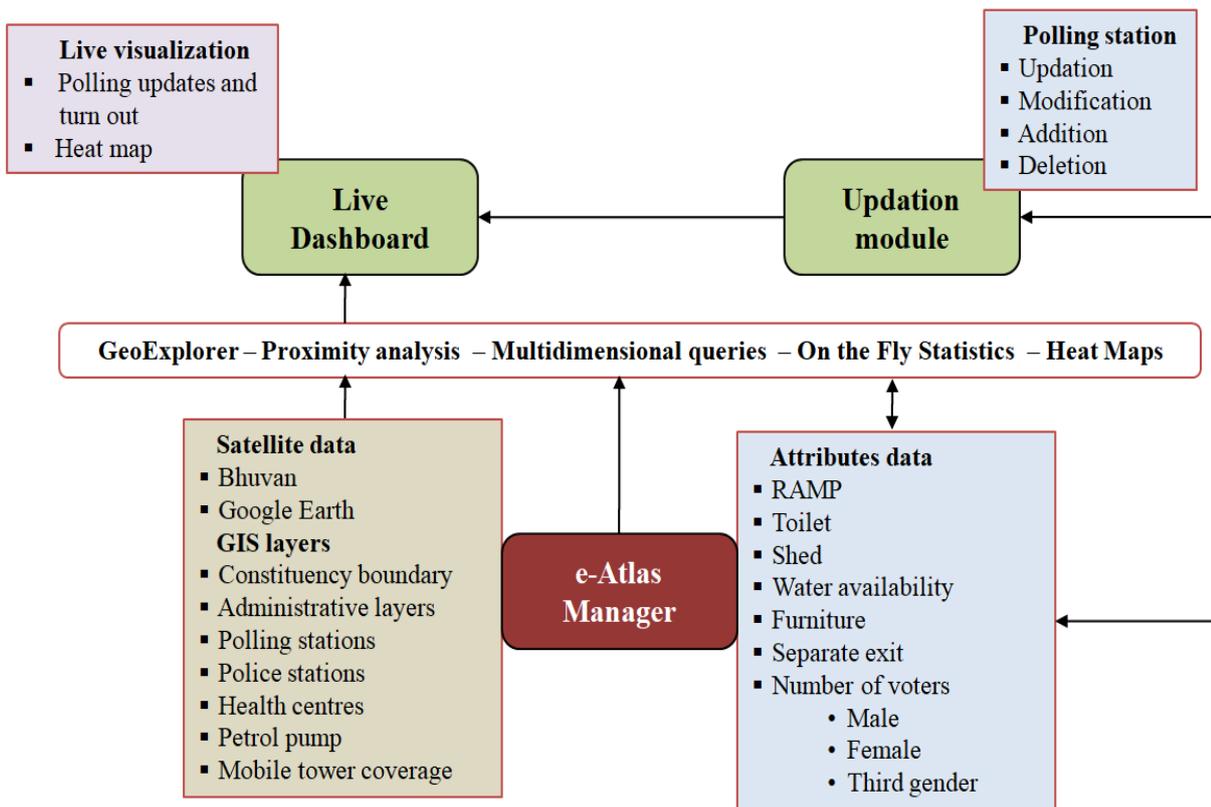


Figure 1: Availability of geospatial information and the services of respective modules are illustrated synoptically.

Table 2: Capabilities and functionalities of different modules of the developed EPMS

Modules	Capabilities and functionalities
e-Atlas Manager	<ul style="list-style-type: none"> • GeoExplorer: roads, settlements, police and polling stations, health centres, mobile networks, Assembly and Parliamentary Constituencies boundary and satellite base map (Bhuvan/Google) • Proximity to: roads, police and polling stations, assured minimum facility (AMF), health centres, State and International boundary • Multi-dimensional query: between proximity, AMF and voter counts • Heat map based on spatial frequency of polling stations
Live Dashboard	<ul style="list-style-type: none"> • Live visualization of polling updates and turn out • Heat map generation based on live polling updates • Categorized view of polling statistics
Updation module	<ul style="list-style-type: none"> • Updation of existing polling stations details • Addition of new polling stations • Live updation of polls for Live Dashboard

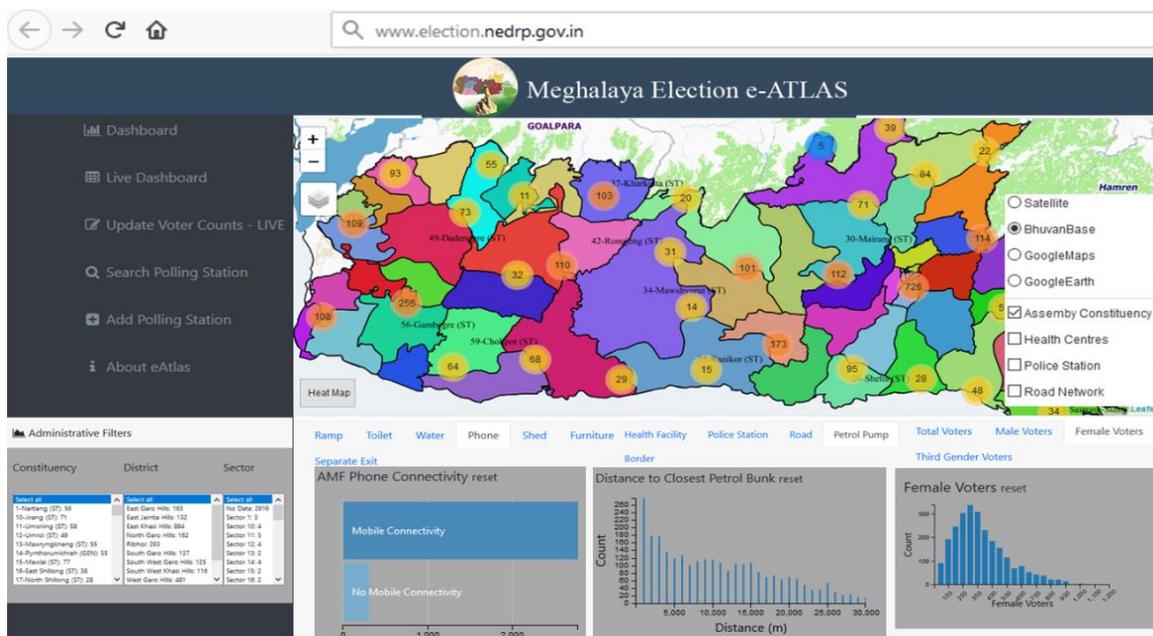


Figure 2: The landing page of EPMS hosted at election.nedrp.gov.in

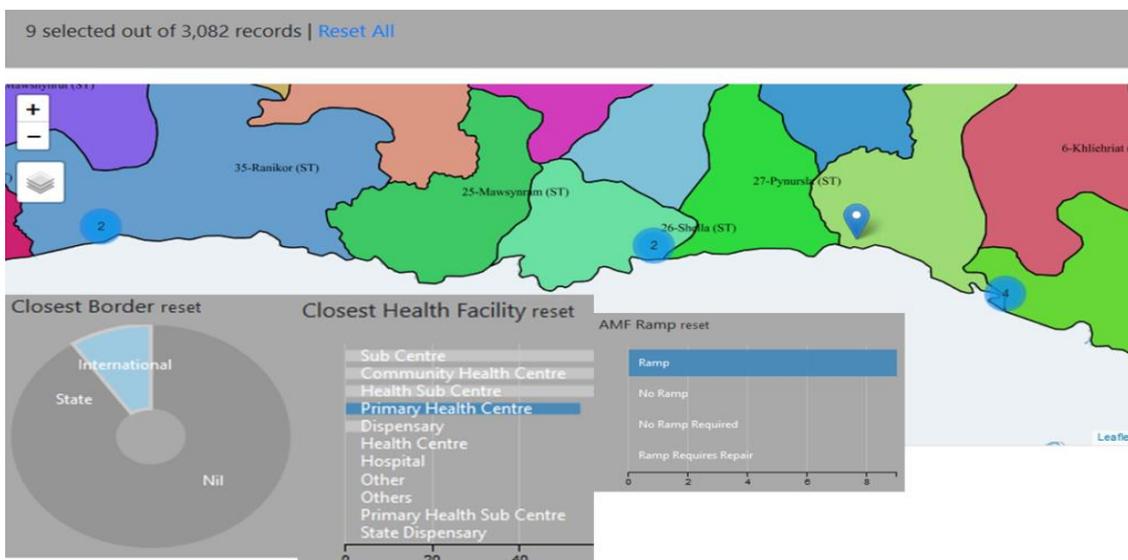


Figure 3: An example of multidimensional query fired to find out the list polling stations nearer to the International boundary having primary health centre and assured minimum facility (AMF)

2.1 Architecture of EPMS

The application comprises of a 3-tier architecture; 1) Database, 2) GIS server and the 3) Web/client mapping tier. The database used for the application is PostgreSQL 9.6. PostgreSQL was chosen as it offers a rich library of spatial querying methods when used with the PostGIS extension. PostgreSQL is enriched with flexible permissions system, with the ability to bundle out specific privileges to specific roles, and provide users with the powers of one or more of those roles. In addition, it uses multiple different systems to authenticate users. The geometry of the vector layers has been stashed away in the database in the Well Known Text (WKT) format. The access to the data is improved by the presence of indexes which were created keeping two factors in mind: numerical queries and spatial questions. The Generic Index Structure (GiST, Douglas 1979) was preferred for spatial queries as it was more effective in spatial indexing as compared by to B-Tree (Hellerstein et al. 1995) Geospatial data is served to the client end via Open Geospatial Consortium (OGC) services. These services are generated with the help of GIS servers and in this case Geoserver is used to serve the information as web mapping services (WMS) as well as web feature services (WFS). The architectural design follows the service oriented architecture (SOA). The services are being provided to the other modules by application interface, through a communication protocol over a network. Geoserver acts as a platform to host these services as well as to attach additional data onto the layers such as styling and layer projection. GeoServer allows access to be determined on a per-layer basis. Providing access to layers is routed via roles. Layers and roles are linked in a file located in the security directory. It contains the rules that control access to workspaces and layers. The client side application is a browser based application. Services generated by Geoserver need to be hosted in a WWW browser – this is accomplished using JavaScript based leaflet APIs. Leaflet acts as a platform for visualizing the geospatial data that are hosted via Geoserver. Map data are passed to the user browser as JavaScript Object Notation (JSON) objects.

The application offers the functionality to query between live polling mode as well as poll station information management mode. User can also query based on the parameters which are attached to the spatial layer as attributes (the query is carried out on the client side with the help of the JavaScript based API, Cross Filter which is a platform for carrying out multi-attribute queries utilizing clients' resources, allowing server to be extensible for data hosting) as well as perform a search upon multiple numeric and non-numeric parameters like distance from nearest petrol bank and police station or nearest hospital. The search results can be used to navigate to the appropriate polling station entry. Further queries can be fired dynamically based on its major attributes. Access to the application needs to pass through the central authentication system (CAS). The architectural flowchart is illustrated in figure 4.

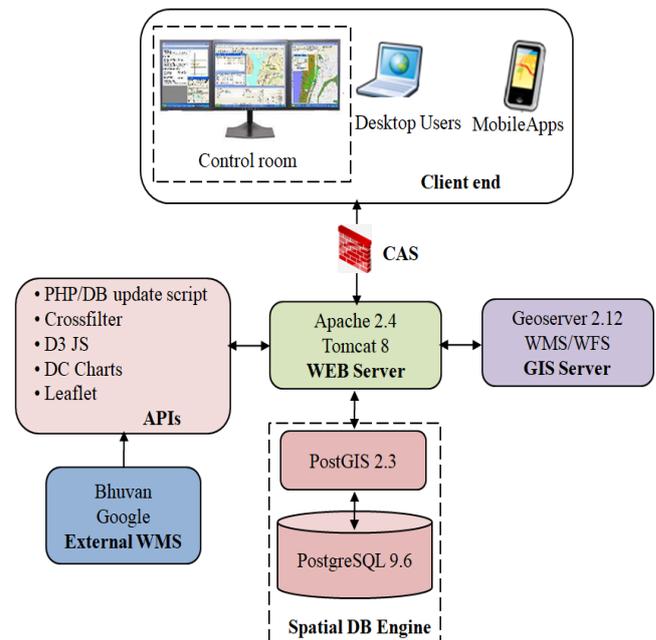


Figure 4: Architectural flowchart of the EPMS

3. Discussion

EPMS was developed in a phase manner to meet the requirements of CEO, Meghalaya. New tools and functionalities have been incorporated into each phase of development as per the suggestions. The application was made functional for testing and updating of polling details much before the last Meghalaya Legislative Assembly Election, 2018 held on February 27, 2018. CEO, Meghalaya also launched the election e-Atlas on September 27, 2017 during a State level function. The greatest challenge was to deal with the live feeding of polls of more than 3082 polling stations of Meghalaya during the day of election. Two control rooms were set-up; one at the Office of Deputy Commissioner (DC), RiBhoi district and another one at North Eastern Space Applications Centre (NESAC), Umiam, Meghalaya. The key aim was to capture the live feeds of the polls and to display into the display panel of the DC Office (Figure 5) via GeoExplorer.

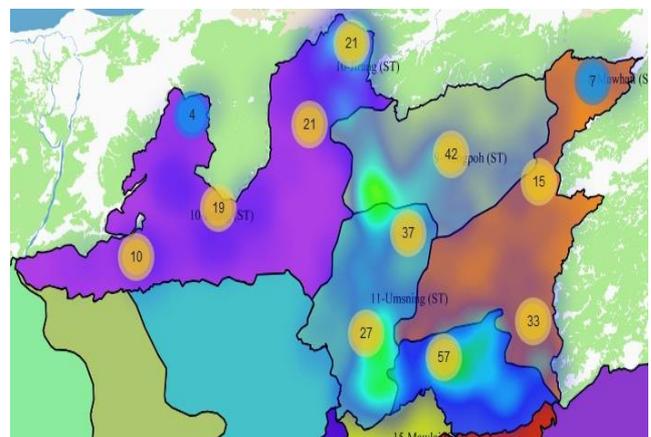


Figure 5: Live visualization of the density of polls demonstrated via a heat map in RiBhoi district

Subsequently, the portal was effectively utilized during the 43-Williamnagar (ST) by-election held on April, 2018. Before that the application was further amended with the enhanced functionalities as per the suggestion of DC, Williamnagar. A team of NESAC staff was deployed at the district Hq. to support the live feeding of polls. Live visualization of polling updates and turnout in 43-Williamnagar (ST) Assembly Constituency is depicted in the figure 6.

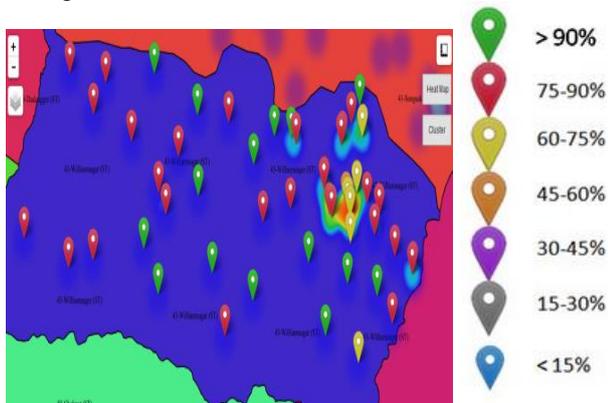


Figure 6: Geo-visualization of live polling turn out or polling percentage in 43-Williamnagar (ST) Assembly Constituency

The percentages given in the right side of the figure 6 with different colours reports the percentage of polls of different polling stations at any moment of time. At the end of the polling, the turnout report may be generated for inspection.

4. Conclusion

The study described here presents the utmost usage of ICT and geospatial technology and services for empowering the electoral process in the State of Meghalaya. It ensures

the increased accountability, transparency and efficiency of the electoral system. This is one of the unique initiatives of CEO, Meghalaya towards strengthening the democratic process and this may be replicated to the country level operational activity. Recently, the application is also made available at apps.nesdr.gov.in/election.

References

Chutia, D. (2011). Application of Geospatial Technologies for development of an efficient Election Management System in Meghalaya. *Journal of Geomatics*, 5(2), 115.

Douglas, C. (1979). The Ubiquitous B-Tree. *Computing Surveys* 11 (2), 123–137

Election Commission of India (ECI), The Function (ElectoralSystem), http://eci.nic.in/eci_main1/the_function.aspx.

Election Commission portal, Bhuvan, <http://bhuvan.nrsc.gov.in/election.html>.

Joseph M. Hellerstein, Jeffrey F. Naughton and AviPfeffer(1995). Generalized Search Trees for Database Systems. Proc. 21st Int'l Conf. on Very Large Data Bases, Zürich, pp. 562–573.

North Eastern District Resources Plan (NEDRP) portal hosted at www.nedrp.gov.in.

Open Geospatial Consortium (OGC): <http://www.opengeospatial.org>.

Susewind, R. (2014). GIS shapefiles for India's parliamentary and assembly constituencies including polling booth localities. Data set/Database, doi: 10.4119/unibi/2674065. DOI: 10.4119/unibi/2674065.

Landslide hazard evaluation and zonation in Dilbe Town and its surrounding areas, North-western Central Ethiopia – A GIS based grid overlay statistical approach

Sendu Demessie¹ and Tarun Kumar Raghuvanshi^{2*}

¹Department of Geology, Woldia University, PO Box 400, Woldia, Ethiopia

²School of Earth Sciences, College of Natural Sciences, PO Box 1176, Addis Ababa University, Addis Ababa, Ethiopia

*Email: tkraghuvanshi@gmail.com

(Received: Apr 04, 2019; in final form: June 18, 2019)

Abstract: The present study was carried out in Dilbe town and its surrounding areas, North-western Central Ethiopian, located about 576 km from Addis Ababa city. The main objective of the present study was to carry out landslide evaluation and Zonation of the study area for which an integrated GIS based Grid overlay statistical approach was followed. For landslide hazard zonation (LHZ) six causative factors; slope-material, slope, aspect, elevation, land use/land cover and groundwater surface traces were considered. These causative factors were derived from topographical sheets, secondary maps, digital elevation model and through field investigation. The quantitative relationship between these causative factors and the past landslides in the area was made through overlay analysis and the Landslide susceptibility index (LSI) was computed. Further, geo-processing was done by overlaying a polygon grid (10x10 m) over each factor map in GIS environment. This forms the basis to evaluate LHZ map of the study area. The results showed that 11.55% of the study area falls into 'very high hazard', 19.84% area falls into 'high hazard', 14.36% area fall into 'medium hazard', 38.25% area falls into 'low hazard' and remaining 16% of the area falls into 'very low hazard'. Further, validation results showed that 74% of the past landslides fall within 'very high hazard' and 'high hazard' zones of the prepared LHZ map. The prepared LHZ map has reasonably validated with the past landslide data. Thus, various landslide hazard zones delineated can safely be applied for future developmental planning in the present study area.

Keywords: Landslides, Landslide hazard evaluation, Landslide susceptibility index, Landslide hazard zonation

1. Introduction

Landslides are well known devastating natural hazard in mountainous terrain. The landslides result into wide spread damage to the infrastructure, land degradation and significant loss to the human life and injury throughout the world (Varnes, 1996; Parise and Jibson, 2000; Dai et al., 2002; Sarkar and Kanungo, 2004; Crozier and Glade, 2005; Kanungo et al., 2006; Pan et al., 2008; Marraqu and Jakka, 2014; Raghuvanshi et al., 2014a; Girma et al., 2015; Raghuvanshi et al., 2015). Several internal and external factors, in combination, result into landslide. The primary internal factors are related to geo-morphology, geology and hydrology (Anbalagan, 1992; Ayalew et al., 2004; Wang and Niu, 2009; Hamza and Raghuvanshi, 2017) whereas main external factors are rainfall (Collison et al., 2000; Dai and Lee, 2001; Dahal et al., 2006; Raghuvanshi et al., 2014a), seismicity (Parise and Jibson, 2000; Keefer, 2000; Bommer and Rodri'guez, 2002; Raghuvanshi et al., 2014a) and manmade activities (Gorsevski et al., 2006; Raghuvanshi et al., 2014a). These factors are responsible for slope failure either by reducing the shear strength of the slope material or these factors may increase the shear forces (Sarkar and Kanungo, 2004).

Landslide hazard evaluation and zonation involves delineation of future hazardous zones based on analysis of governing parameters (Arnous, 2011) and to classify land into different zones of potential or actual landslide hazard (Varnes, 1984; Anbalagan, 1992; Raghuvanshi et al., 2014a). In order to carryout landslide hazard evaluation, prediction and zonation, several techniques are proposed by different researchers which can broadly be classified into expert evaluation (Anbalagan, 1992; Raghuvanshi et al., 2014b), statistical methods (Carrara et al., 1992;

WestenVan et al., 1997; Dai and Lee, 2001) and the deterministic approach (Fall et al., 2006; Raghuvanshi et al., 2015). Each of these techniques has their own merit and demerits over others (Leroi, 1997; Guzzetti et al., 1999; Casagli et al., 2004; Kanungo et al., 2006; Fall et al., 2006; Raghuvanshi et al., 2014a; 2014b). The selection of a technique for landslide hazard zonation primarily depends on factors, namely; area to be covered for hazard zonation, scale at which zonation has to be done, geological and geo-morphological factors to be considered, method by which parameter data has to be acquired and skill set or capability of an evaluator (Carrara, 1992; Ermias et al., 2017).

In Ethiopia, landslide is a common geo-environmental hazard in the highlands (Ayalew and Yamagishi, 2004; Ayenew and Barberie, 2005; Abebe et al., 2010; Raghuvanshi, 2014a; 2014b). The present study area, Dilbe town and its surrounding area, which is located in the north western central Ethiopian highland, is a rugged mountainous terrain. The area is well known for its devastating landslides which has caused considerable damage to the roads, houses and the agricultural land in the area. Thus, there was a necessity to evaluate and carry out landslide hazard zonation in the area, so that concerned authorities and local people may be made aware about the hazardous zones in the area. Such information may be helpful for the safe land use planning and development in the area. Besides, this may also help to mitigate the problems related to landslides in the study area. The main objective of the present study was to evaluate the factors influencing the landslide hazard and to produce the landslide Hazard Zonation (LHZ) map of the study area.

2. The Study area

The present study area is located in the North-western central Ethiopia, North Wello Zone in Gubalafto District (Figure 1). The total study area is about 50 km². The study area is bounded by UTM co-ordinates 1318000m to 1326500m N and 543200m to 552800mE. The study area is about 576 km from Addis Ababa, the capital city of Ethiopia, on way to Woldiya Town. The elevation in general ranges from 2243m to 3591m a.s.l. The study area has rugged topography and is characterized by sharp valleys, hills and ridges. The climate of the area is characterized as temperate (Gemetchu, 1997). The long term average mean annual rainfall in the study area is 661.5 mm (1995 to 2016) and the maximum monthly average precipitation recorded was 401.7 mm in the month of August in 2010. The maximum and minimum temperature in the area is 24.3°C and 1.2°C, respectively. The drainage in the study area is characterized as dendritic to sub parallel and all major and minor streams drain into Gembora River.

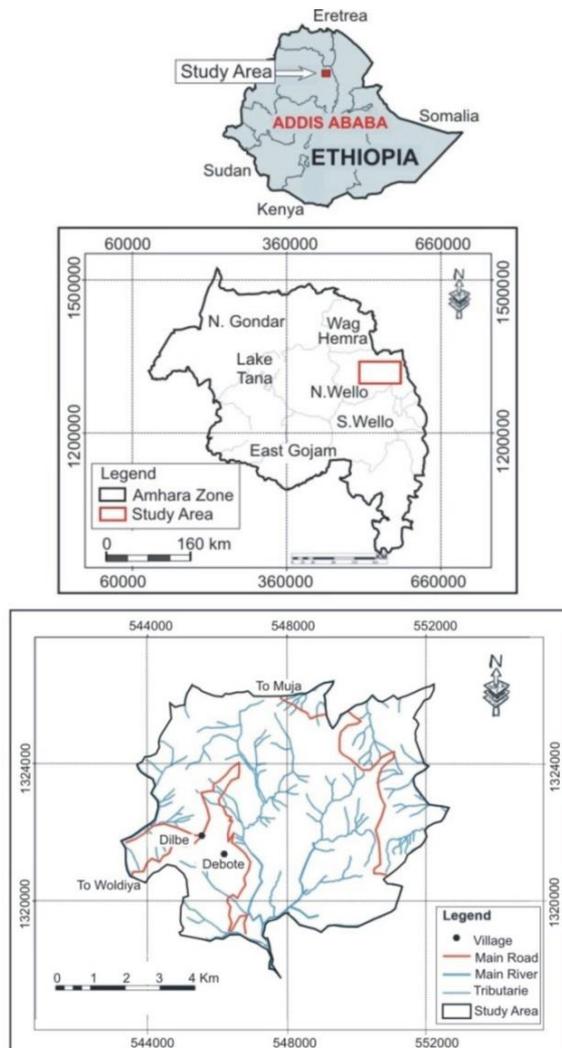


Figure 1: Location map of the study area

3. Geology

The present study area is located in the South-western block of the Afar Depression, where the Main Ethiopian Rift (MER) gradually funnels towards the Afar Depression. Ashangi basalts are the earliest fissural volcanic rocks

exposed in the area (Tefera et al., 1996) which belongs to Eocene-Oligocene period. The Ashangi basalts are followed by Dessie basalt and Tarmaber Megezez Formations, belonging to Oligo-Miocene period (Demissie et al., 2010) (Figure 2). Ashangi basalt Formation is exposed mainly over the steep slopes, low lying flat plains, stream beds and on the gentle slopes. These are overlain unconformably on the Dessie basalt formation. Ashangi basalts are highly weathered, jointed and fractured and are oriented in different directions (Demissie et al., 2010). Aphanitic basalts with columnar jointing are commonly exposed in the study area which belongs to Ashangi basalt Formation. These rocks are black, dark gray and greenish gray in color with aphanitic to coarse grained in texture. Dessie basalt Formation is mainly exposed in the western plateau area and comprises aphanitic, porphyritic, massive and vesicular basalts (Demissie et al., 2010). The aphanitic basalt belonging to Dessie basalt Formation is dominated by fine microcrystalline matrix with fine plagioclase microlites and porphyritic basalts (Demissie et al., 2010). The Tarmaber Megezez Formation is exposed on the gentle slopes, along the stream beds and along the road sections.

This Formation forms a conformable contact with the Dessie basalt Formation. The rocks belonging to this Formation are characterized by dark gray to black coarse grained vesicular and columnar jointed aphanitic basalts (Demissie et al., 2010).

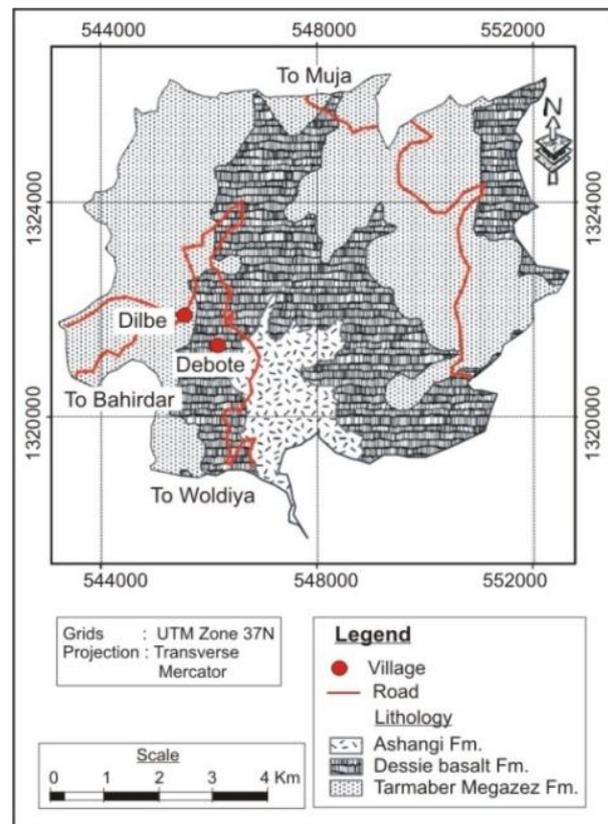


Figure 2: Geology of the study area

4. Methodology

In order to prepare the landslide hazard zonation (LHZ) map of the study area, a GIS based grid overlay statistical approach was followed. This technique was followed as it

is based on the principle that “the past and the present are the keys to the future”, the future landslides, most likely will occur under similar conditions which has prevailed during the past or the present times (Van Den Eeckhaut et al., 2009). Based on the statistical analysis for the interrelationship of causative factors and the past activities, quantitative estimates can be made for those areas where similar conditions prevailed. Each of the causative factor maps can be overlaid on past landslide map to know the relative contribution for each factor and subclass in inducing landslides in the area. Thus, from this data, respective weights can be developed to be applied to each causative factor subclass and finally a landslide hazard can be deduced for the given area (Dai and Lee, 2001).

For the present study six causative factors; (i) slope material, (ii) slope, (iii) aspect, (iv) elevation, (v) land use/land cover and (vi) groundwater surface traces were considered for the evaluation of LHZ (Ayalew et al., 2004; Girma et al., 2015; Raghuvanshi et al., 2015). Based on the field observations, it was realized that these factors were the significant factors that have resulted into past landslides in the area. It was further assumed that “the past is the key for future” (Varnes, 1984; Carrara et al., 1991). This means that the conditions that were responsible for the past landslides, if reoccur in other areas, again landslides can occur (Dai and Lee, 2001; Raghuvanshi et al., 2015). Thus, in order to understand the relationship of these causative factors with the past landslides, statistical analysis was made between each of the causative factors and the past landslides. For this purpose, individual factor maps were overlaid on the past landslide map in GIS environment and quantitative prediction was made through density analysis between the past landslides and each of the causative factor sub-classes (Dai and Lee, 2001; Lee et al., 2004; Girma et al., 2015; Raghuvanshi et al., 2015; Chimidi et al., 2017; Hamza and Raghuvanshi, 2017). With the help of this statistical density analysis, Landslide Susceptibility Index (LSI) was computed for each parameter sub-class (Raghuvanshi et al., 2015). Further, for grid overlay analysis a polygon grid (10 x 10 m) was overlaid on each parameter theme and geo processing was done to know the presence of each parameter sub-class within each grid cell. Later, respective LSI values for each factor sub-class within each grid cell were assigned. Thus, the sum total of LSI values for six parameters within each grid cell provided Total Landslide Susceptibility Index value (TLSI). These TLSI values were further utilized to define the landslide hazard zonation in the study area.

The present study was carried out in three phases; Pre-field work-desk study, Field investigation and Post field study. During the pre-field-desk study required data from secondary sources was collected. This includes collection of data on meteorology, topographical maps and satellite data. Besides, previous reports, data and maps available for the study area were reviewed. Further, slope facet map, geological and soil cover maps were prepared, later to be utilized during the field work. In addition, digital elevation model (DEM) was obtained at a resolution of 30m from the ASTER data set and factor maps for elevation, slope and aspect were prepared. Also, land use land cover map

of the present study area was prepared from the Sentinel-2A data through supervised classification.

The major activities carried out during the field investigation includes inventory mapping for past landslides and verification of the factor maps that were prepared during the pre-field-desk study. The past landslide inventory mapping was carried out by identifying the landslides for its type, dimension, material involved, morphology of failed slope, failure mechanism and the possible triggering factors. The data on landslide inventory was collected through GPS readings along the periphery of the landslide, visual observations on various aspects and through personal interviews with the local residents. Prior to the field work various past landslides in the area were tentatively identified through the Google Earth image and were marked as point data on the facet map. Later during the field work they were further verified and necessary inventory data/ information, as mentioned above were collected.

Besides, factor maps on lithology and soil cover were verified in the field and necessary modifications were made. In order to collect the primary data in the field and to verify factor maps prepared during the pre-field-desk study, slope facet map was used. The facet map was prepared during the pre-field-desk study by delineating major or minor hill ridges and streams on the topographic map by using Arc GIS software (Anbalagan, 1992). Slope facet is basically required to demarcate the area into more or less identical land unit which have nearly similar slope direction and inclination. Therefore, facet provides a means to recognize the land area within which observations can be made. Since, the facets are bounded by streams and ridges therefore they can easily be recognized in the field with the help of topographical map. Thus, the primary purpose of facet is to demarcate the study area into various units which can easily be recognized in the field and within which observations can easily be made for various causative factors.

A total of 54 slope facets were delineated in the study area (Figure 3). In addition, springs present in the study area were located and marked with the help of GPS reading on the elevation map.

The post field study includes processing of data collected during the field investigation stage and to carry out further analysis to prepare landslide hazard zonation map of the study area. The past landslide inventory data, collected as GPS point data along the periphery of the landslides during the field investigation, was converted to polygon data by digitization over the Google-Earth image.

All factor maps and landslide inventory map were brought to the GIS environment for further processing. Besides, individual factor maps were overlaid on past landslide map in GIS environment and quantitative predictions were made through density analysis between the past landslides and each of the causative factor class. Thus, the LSI values for each parameter sub class were computed. Further, overlay analysis was made by overlying grid on each parameter map. Finally, with the help of grid overlay

analysis results and the respective LSI values for each factor sub-class, landslide hazard zonation in the study area was worked out. Figure 4 shows the general methodology followed during the present study.

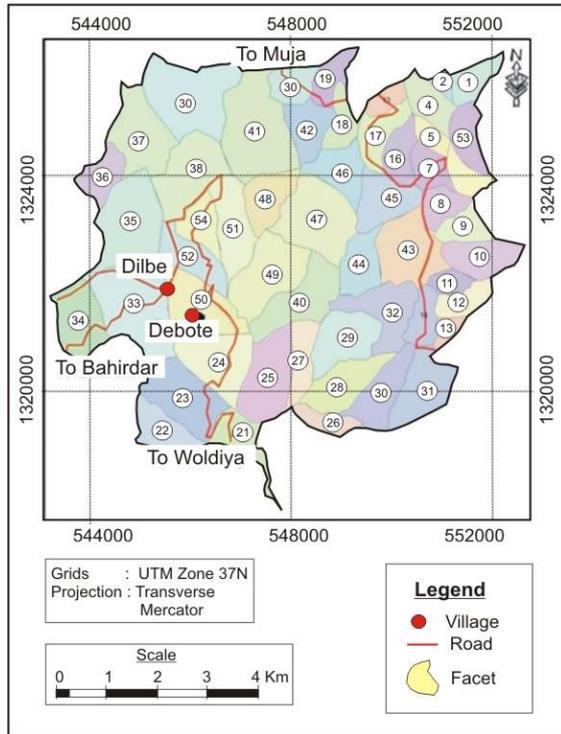


Figure 3: Facet map used for the field data collection

5. Data collection, processing and analysis

For the present study the required data for landslide evaluation and zonation was obtained from the secondary and primary sources. The secondary data that was utilized for the present study includes; topographical maps, geological map, soil cover map, satellite data, digital elevation model (DEM) and the meteorological data (Table 1). The primary data that was used in the present study was mainly obtained through field investigation. This includes verification of factor maps prepared during the pre-field-desk study from secondary data sources and the collection of inventory data and information for the past landslides in the study area.

5.1 Landslide inventory

The landslide inventory data provides useful information in understanding the relative influence of various causative and triggering factors on the slope instability and possible mechanism that has resulted into slope failures. Through landslide inventory data it is possible to understand the conditions that were responsible for the landslides in the area and with this understanding it is possible to know the probable areas where future landslides can occur (Dai et al., 2002; Lan et al., 2004; Raghuvanshi et al., 2015; Chimidi et al., 2017).

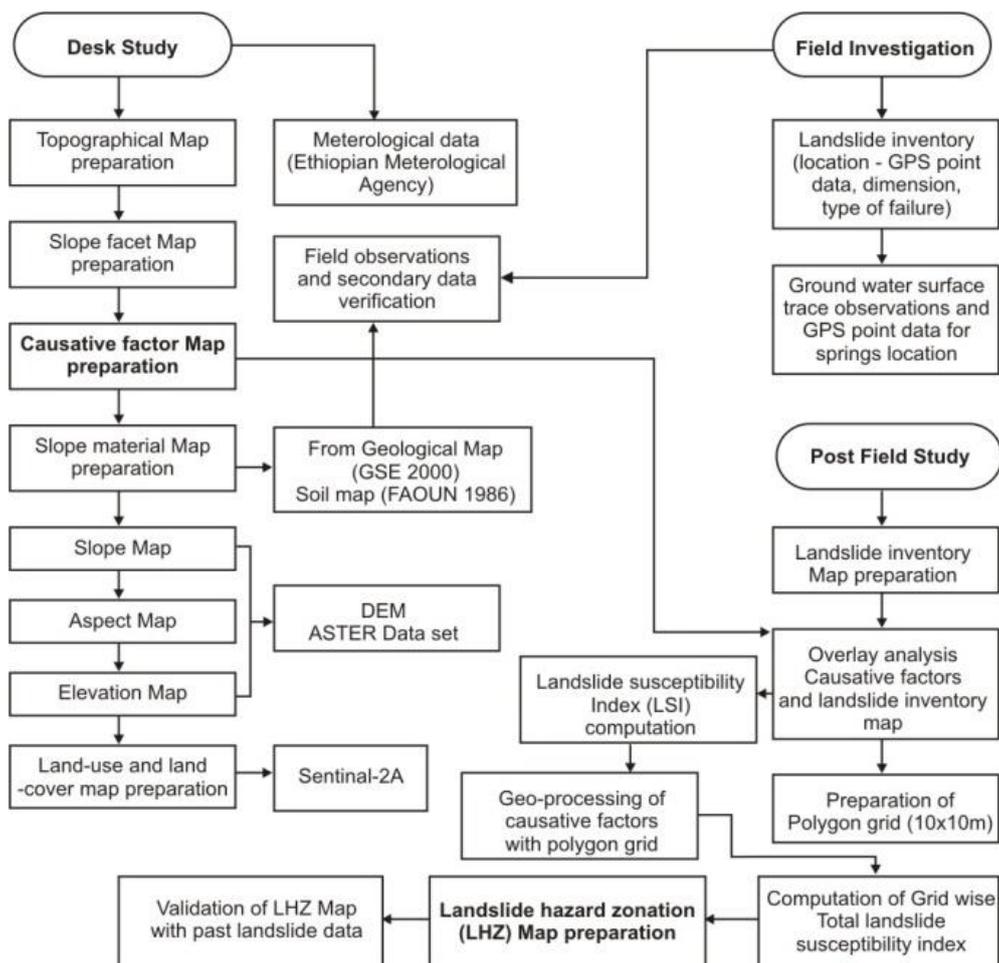


Figure 4: General Methodology followed

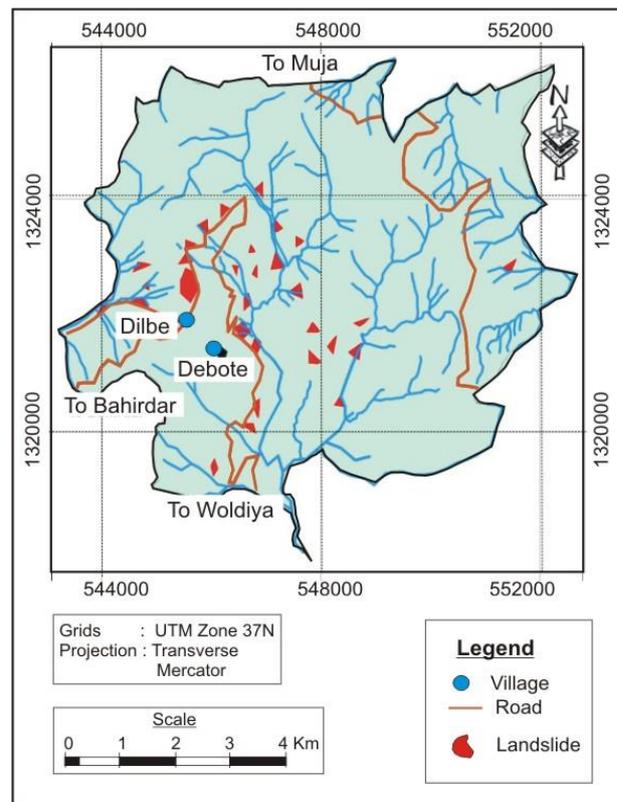
Table 1: Causative factors respective data source

Causative factors	Data Source
Landslide inventory	Field observations – GPS point data along landslide periphery and by using Google Earth image.
Slope material	Slope material map was prepared from the Geological map of Dessie area prepared by Geological Survey of Ethiopia with a scale of 1:250000 (GSE, 2010), soil map of Ethiopia prepared by Food and Agriculture Organization of United Nations (FAOUN, 1986) and through field investigation and mapping.
Slope Aspect Elevation	DEM data with a resolution of 30 m ASTER elevation data set
Land use land cover	Sentinale-2A and field observation and mapping
Groundwater surface trace	Field observations - GPS point data collected at springs location during field work and delineation of hydrological homogeneous zones over topographical maps.

For the present study landslide inventory data was collected through the field investigation. In order to identify the past landslides in the study area traverse mapping was done and the location of all past landslides was marked over the facet map. Besides, GPS point data was also collected and recorded along the periphery of each landslide. In addition, data on type of failure, failure mechanism, dimension, material involved and failed slope morphometry was also collected. In total 30 past landslides were identified in the study area (Figure 5).

These landslides have mainly failed by following four different modes; rotational, translational, fall and complex mode of failures. The translational and complex modes of failures were observed mainly in residual and alluvial soils, respectively. Further, most of the rotational mode of failure was observed in colluvial and alluvial deposits. On the other hand, rock fall in the area were observed in the disintegrated rock mass, mainly along Woldiya - Bahir Dar road section. Besides, the local administration offices were approached to collect systematic records on past landslides, particularly for landslide occurrence date, time, duration, failure mechanism etc., however no such data is being maintained by the concerned offices.

Therefore, to have such information, local residents were approached and informally interviewed through pre-designed questionnaires. According to the local respondents, most of the past landslides occurred during mid of July to September in past years. However, respondent failed to provide information on exact date and time for the past landslides in the area.

**Figure 5: Landslide inventory map**

Further, the meteorological data showed relatively high rainfall during this period. Thus, it can safely be concluded that main triggering factor for the past landslides in the present study area was heavy rainfall.

5.2 Causative factors evaluation

The stability of a slope is mainly governed by the causative intrinsic parameters (Raghuvanshi et al., 2014a; Wang and Niu, 2009; Ayalew et al., 2004; Anbalagan, 1992). The causative factors considered for the present study are; slope-material, slope, aspect, elevation, land use/land cover and groundwater surface traces. These causative factors were selected based on the field observations and their possible relative contribution in inducing instability to the slopes in the present study area. For the purpose of landslide hazard zonation (LHZ) in the present study, attempt was made to evaluate causative factors with respect to their quantitative relationship with the past landslides in the area.

5.2.1 Slope material

Slope material map was prepared by combining soil map and lithological map of the study area. The soil map was extracted from the soil map of Ethiopia, prepared by Food and Agriculture organization (FAOUN, 1986) whereas; lithological map was extracted from the map prepared by the Geological survey of Ethiopia (GSE, 2010). The slope material map was further verified and modified with the field survey. In the present study area, three types of soils are present, these are; residual, alluvial and colluvial. The rocks exposed in the study area are classified as disintegrated and blocky rock mass (Figure 6a). In order to evaluate contribution of slope material on past landslides, overlay analysis was performed. The overlay analysis revealed that about 38.5% of past landslides has occurred

in disintegrated rock mass, 27% in alluvial soils, 26% in colluvial soil and remaining 8.5% in residual soils (Figure 7).

5.2.2 Slope

The slope map of the study area was extracted from the DEM at 30m resolution from ASTER data set (Figure 6b). In general, slope inclination in the present study area varies from 0 to 70°. For the present study slope inclination was distributed into five classes; 0-5°, 5-10°, 10-25°, 25-38° and slopes >38°. These slope classes were made based on the expert decision and the general topography of the area.

The overlay analysis of the past landslides with slope inclination map revealed that 7% of the past landslides falls in 0-5°, 15% falls in 5-10°, 27% falls in 10-25°, 20% falls in 25-38° and 31% falls in the slope class >38° (Figure 7). About 51% of the past landslides fall in slopes having inclination greater than 25°. Also, about 42% of the past landslides were observed in the gentle slope sections (slope inclination 5 – 25°). It was observed that slopes having inclination in between 5 – 25° are mostly occupied by unconsolidated material which is considered to be highly susceptible for the slope instability (Raghuvanshi et al., 2014a: 2014b).

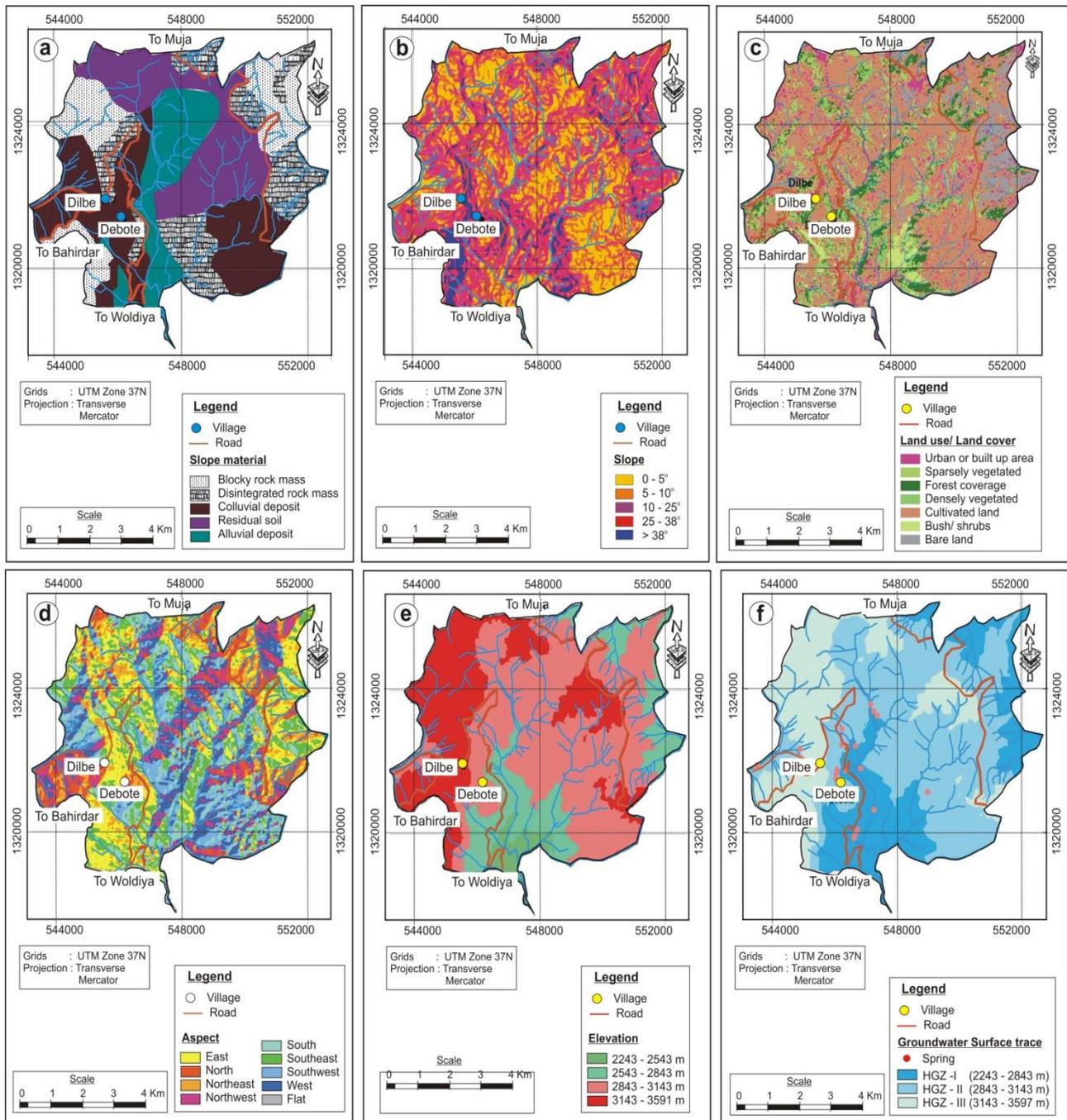


Figure 6: Causative factors map - (a) slope material, (b) slope, (c) land use and land cover, (d) aspect, (e) elevation and (f) ground water surface trace

5.2.3 Aspect

For the present study the aspect map was extracted from the DEM at 30m resolution from ASTER data set. For the present study aspect has been classified as; (i) Flat (1), (ii) North (0– 22.5°), (iii) North-east (22.5–67.5°), (iv) East (67.5–112.5°), (v) South-east (112.5–157.5°), (vi) South (157.5–202.5°), (vii) South-west (202.5–247.5°), (viii) West (247.5–292.5°), (ix) North-west (292.5–337.5°), and (x) North (337.5–365°) (Figure 6d). The overlay analysis revealed that 20%, 18% and 16% of the past landslides occurred in the slopes that are inclined towards South, Southeast and East directions, respectively. Further, 11.2%, 9.2% and 7.7% of landslides occurred on slopes that are oriented towards Southwest, North and Northeast directions, respectively (Figure 7). On the other hand, slopes oriented towards West and Northwest directions have 7.68% and 6.62% of landslides, respectively

5.2.4 Elevation

The elevation of the study area was extracted from the DEM at 30m resolution from ASTER data set. The elevation of the study area was classified into four classes; 3143-3597m, 2843-3143m, 2543-2843m and 2243-2543m (Figure 6e). The overlay analysis showed that about 35% of the past landslides occurred in elevation class of 2243-2543m, 31.5% occurred in 2543-2843m elevation class, 26% occurred in elevation class 2843-3143m and the remaining 7.5% landslides occurred in elevation class 3143-3597m (Figure 7). From these results it can be noticed that about 66.5% past landslides occurred in two elevation classes, 2243-2543m and 2543-2843m.

5.2.5 Land use and Land cover

For the present study land use and land cover map was prepared from the Sentinel-2A data through supervised classification by using ERDAS Imagine. The land use and land cover of the present study area was classified into seven classes; urban or built up area, bare land, bush or shrubs land, sparsely vegetated land, densely vegetated area, forest coverage and cultivated land (Figure 6c). The overlay analysis showed that 26.3% of past landslides fall in the sparsely vegetated land, 24% of landslide falls in Bush or shrubs land, 20% of landslide falls in cultivated land, 8.4% falls in forest coverage, 7.7% falls in densely vegetated area, 7% falls in built up area and remaining 6.6% landslides falls in bare land (Figure 7). The results show that about 70% of the past landslides are concentrated in three classes; sparsely vegetated land, bush or shrubs land and cultivated land. As observed, these land use and land cover classes are mainly occupied by disintegrated rock mass, colluvial and alluvial soils. These unconsolidated materials are susceptible for slope instability (Raghuvanshi et al., 2014b).

5.2.6 Ground water condition

Groundwater is the most important factor that is responsible in inducing instability to the slopes (Raghuvanshi et al., 2014a; Girma et al., 2015; Chimidi et al., 2017). Assessment of groundwater condition throughout the area is practically not feasible for landslide hazard evaluation studies. However, groundwater condition can be assessed by indirect means of surface

manifestations, such as presence of springs over the slopes (Anbalagan, 1992; Chimidi et al., 2017).

Presence of springs on the slope face is an indication of proximity to groundwater and it also suggests general saturation of the slope material. It is also believed that spring locations have direct correlation with the landslides occurrences (Girma et al. 2015; Raghuvanshi et al., 2015; Chimidi et al., 2017). In the present study 22 springs were identified in the field during the inventory mapping. Thus, with the help of spring density and the respective elevation range, hydrological homogeneous zones were delineated in the study area (Raghuvanshi et al., 2015).

The hydrological homogeneous zones that were delineated in the study area are; HGZ-I (2243-2843m), HGZ-II (2843-3143m) and HGZ-III (3143-3597m) (Figure 6f). The overlay analysis with the past landslide data clearly indicates that 46.8% landslides fall in HGZ-I (2243-.2843m) zone, 33.2% falls in HGZ-II (2843-3143m) zone and the remaining 20% landslides fall in HGZ-III (3143-3597m) zone (Figure 7).

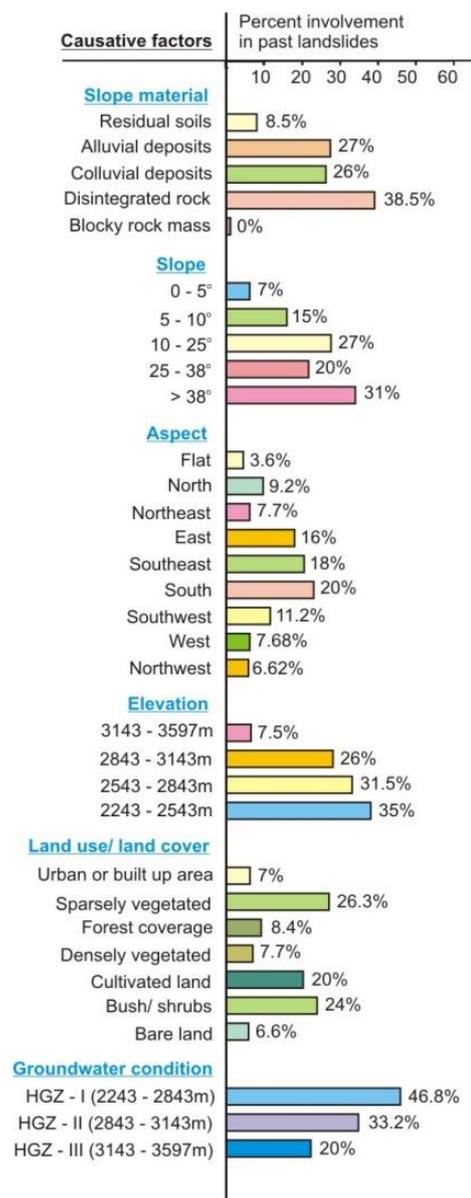


Figure 7: Causative factors influence on past landslide

5.3 Landslide Susceptibility Index (LSI)

Through landslide inventory data it is possible to understand the relative contribution of various causative factors on landslides in the area. The general assumption in this regard is that “the past and the present is the key for the future”. It means that the conditions that were responsible to initiate landslides in the past if reoccur in some other area; again landslides can occur (Dai et al., 2002; Lan et al., 2004; Chimidi et al., 2017). Thus, to understand the quantitative relationship between the past landslides and the factor sub-class of each causative factor Landslide Susceptibility Index (LSI) was computed. The LSI was originally proposed by Sarkar et al., (1995) and later it was modified by Raghuvanshi et al., (2015) and is expressed by Eq.1

$$LSI = Hazardindex * \frac{LSV}{100} \dots (1)$$

Where; ‘LSI’ is the Landslide Susceptibility Index, ‘LSV’ is the Landslide susceptibility value and ‘Hazard index’ is the ratio between “total pixel counts of a sub-class within a Landslide” to the “total pixel count of that sub-class in the area of study”.

In order to calculate LSI for each sub-class of the causative factors, raster calculator tool in ArcGIS was used and Hazard index for each sub class of causative factors was determined. By overlaying the past landslide map over each factor map, total pixel count of a sub-class within area covered by Landslide and the total pixel count of that sub-class in the area of study were determined. Through the raster calculator in ArcGIS, it was found that the total number of pixels for the entire study area is 54257 whereas the number of pixels covered by the landslide is equal to 1020. Table 2 shows the total pixel counts of a sub-class of each causative factor within a Landslide and the total pixel count of that sub-class in the area of study. Further, based on the comparative significance of each causative factor in inducing landslides in the area, LSV were assigned to each causative factor. These LSVs were assigned on a scale of 100 with proportionate distribution to respective causative factors. Landslide is a complex process and it is resulted from contributions of various causative factors. Practically, it is not possible to evaluate contributions of each causative factor in quantitative terms however, through expert evaluation an effort was made to assign LSV to respective causative factors based on the observations made on the past landslide activities in the study area and through the evaluation of the terrain condition (Raghuvanshi et al., 2015). LSV values thus, assigned to the respective causative factors are presented in Table 2. Through past landslide inventory data it was realized that slope material, groundwater surface trace and land use and land cover factors are relatively most prominent factors and have almost contributed significantly for landslide occurrence in the study area. Thus, a LSV value of 20 was assigned to each of these causative factors. Further, elevation and slope factors were found to be less significant therefore these factors were assigned with a LSV value of 15 each. Aspect was given a LSV value of 10, as it was not found to be relatively significant.

5.4: Grid overlay analysis for landslide hazard evaluation

In order to perform geo-processing by overlay analysis the entire study area was divided into 10 x 10m regular square polygon cells. For this a grid was prepared in AutoCAD map where the study area boundary was imported and a grid with 10 x 10m was created. The total study area (50 sq km) was covered by 5226 polygon grid cells. Later, this grid file was exported as a shape file. In order to assign unique IDs to each grid cell ‘*.dbf’ component of the shape file was edited in MS Excel program and unique IDs to each grid cells (1 to 5226) were assigned. Later, this grid file was utilized for overlay analysis with each causative factor maps in GIS environment. Further, the grid file was overlaid on each individual causative factor themes and geo-processing was done.

The primary purpose of this geo-processing was to know the presence of various factor sub-classes in each grid cell. Thus, the overlay analysis resulted into six files, each containing grid cells with various sub-classes of respective causative factors. Later, all six geo-processed causative factor themes were merged together by using ‘merge theme’ option in ArcMap. Finally, a single composite grid file showing intersection with various sub-classes of all six causative factors were obtained. This composite geo-processed grid file also showed attribute data for specific factor sub-classes in 6 columns for each grid cells. Later, ‘*.dbf’ component of composite grid shape file was edited in MS Excel and corresponding LSI values were assigned to each factor sub-class by using find and replace command in MS Excel. Further, for each grid cell LSI values for all 6 causative factors were summed up to get a Total Landslide Susceptibility value (TLSI). Finally, these TLSI values, obtained for each grid cell, formed the basis for the Landslide Hazard Zonation (LHZ) of the study area.

5.5 Landslide Hazard Zonation (LHZ)

For the present study area the minimum TLSI value was found to be 0.58 and the maximum value was 1.34. Further, the TLSI values were classified into various hazard classes as; very low hazard (VLH), low hazard (LH), moderate hazard (MH), high hazard (HH) and very high hazard (VHH). The initial distribution of the TLSI values for various hazard classes was based on logical judgment. Later, attempts were made on trial basis by considering different distributions of TLSI values for various hazard classes and the corresponding LHZ maps were prepared. For every such attempts overlay analysis was made to validate the LHZ map with the past landslide inventory map. Thus, the TLSI values for various hazard classes that gave the best validation with the past landslide data was finally considered. The TLSI values distributed for various landslide hazard classes used to prepare the final LHZ map for the present study is presented in Table 3. Further, the LHZ map prepared during the present study is presented as figure 8. A perusal of figure 8 clearly shows that 11.55% (5.78km²) of the study area falls into ‘very high hazard’ (VHH), 19.84% (9.92km²) area falls into ‘high hazard’ (HH), 14.36% (7.18km²) area falls into ‘medium hazard’ (MH), 38.25% (19.13km²) area falls into ‘low hazard’ (LH) and remaining 16% (8km²) of area falls into ‘very low hazard’ (VLH).

Table 2: Landslide causative factor classes with their respective LSV and LSI

Causative factors and corresponding factor class	Pixel count for Landslide did not occur		Pixel count for Landslide occurred		Hazard index (b/a)	Landslide Susceptibility Index LSI	Landslide Susceptibility Value LSV
	Count	Ratio (a) %	Count	Ratio (b) %			
(a) Slope materials							
Residual soils	12043	22.20	58	5.69	0.26	0.052	20
Alluvial deposits	8141	15.00	261	25.59	1.71	0.342	
Colluvial deposits	15053	27.74	272	26.67	0.96	0.192	
Disintegrated rock mass	10705	19.73	429	42.06	2.13	0.426	
Blocky rock mass	8315	15.33	0	0.00	0.00	0.000	
Total	54257	100.00	1020	100.00			
(b) Slope							
0-5	12076	22.26	175	17.16	0.77	0.116	15
5-10	16498	30.41	270	26.47	0.87	0.131	
10-25	13452	24.79	299	29.31	1.18	0.177	
25-38	8914	16.43	189	18.53	1.13	0.170	
>38	3317	6.11	87	8.53	1.40	0.210	
Total	54257	100.00	1020	100.00			
(c) Aspect							
Flat (0°)	375	0.69	16	1.57	2.28	0.227	10
N (0°-22.5°)	2496	4.60	18	1.76	0.38	0.038	
NE (22.5°-67.5°)	6147	11.33	107	10.49	0.93	0.093	
E (67.5°-112.5°)	9374	17.28	220	21.57	1.25	0.125	
SE (112.5°-157.5°)	8122	14.97	192	18.82	1.26	0.126	
S (157.5°-202.5°)	6701	12.35	143	14.02	1.14	0.114	
SW (202.5°-247.5°)	6728	12.40	117	11.47	0.93	0.093	
W (247.5°-292.5°)	6393	11.78	104	10.20	0.87	0.087	
NW (292.5°-337.5°)	5678	10.47	83	8.14	0.78	0.078	
N (337.5°-360°)	2243	4.13	20	1.96	0.47	0.047	
Total	54257	100.00	1020	100.00			
(d) Elevation							
2243-2543	6319	11.64	185	18.14	1.56	0.234	15
2543-2843	14191	26.16	361	35.39	1.35	0.203	
2843-3143	24401	44.97	276	27.06	0.60	0.090	
3143-3597	9346	17.23	198	19.41	1.13	0.170	
Total	54257	100.00	1020	100.00			
(e) Land-use and Land-cover							
Urban/ built up area	3593	6.63	60	5.88	0.89	0.178	20
Forest coverage	4580	8.44	80	7.84	0.93	0.186	
Densely vegetated	3371	6.21	60	5.88	0.95	0.190	
Cultivated land	31290	57.67	566	55.5	0.96	0.194	
Bare Land	4368	8.05	87	8.53	1.06	0.212	
Sparsely Vegetated	5757	10.61	137	13.43	1.27	0.254	
Bush land	1298	2.39	30	2.94	1.23	0.246	
Total	54257	100.00	1020	100.00			
(f) Ground water surface traces							
HGZ-I (2243-2843)	12600	23.22	300	29.41	1.27	0.254	20
HGZ-II (2843-3143)	28321	52.20	520	50.98	0.98	0.196	
HGZ-III (3143-3597)	13336	24.58	200	19.61	0.80	0.160	
Total	54257	100.00	1020	100.00			

6. Results and discussion

6.1 Causative factors relation with landslides

The landslide inventory carried out during the present study showed the presence of 30 past landslides in the study area. These landslides are mainly present in the central, western and the southern parts of the study area (Figure 5). The landslides, as observed in the study area, have failed by following fall, transitional, rotational and complex mode of failures. Out of total observed landslides 76% has failed by rotational mode of failure, 17% failed as fall and remaining 7% as translational and complex mode of failures. Most of the rotational mode of failure was observed in colluvial and alluvial deposits whereas transitional and complex modes of failures were observed mainly in the residual and alluvial soils, respectively. Besides, rock fall in the area were mainly observed in disintegrated rock mass. As revealed by the local respondents all these landslides in the present study area have occurred during the rainy season (mid of July-September). This indicates that the main triggering factor for past landslides in the study area is heavy rainfall.

Table 3; Landslide hazard zonation based on Total Landslide Susceptibility Index (TLSI) value

No	Value for zonation	Zone Designation	Class
1	0.58-0.78	VLH	Very low hazard
2	0.79-0.99	LH	Low hazard
3	1.00-1.11	MH	Moderately hazard
4	1.12-1.33	HH	High hazard
5	>1.33	VHH	Very high hazard

Further, past landslide data revealed that about 38.5% landslides occurred in disintegrated rock mass, 27% in alluvial soils, 26% in colluvial soil and the remaining 8.5% in residual soils. No landslides were recorded in the blocky rock mass. These figures clearly show that 61.5% of landslides have occurred in slopes that are covered by alluvial, colluvial or residual soils and 38.5% of landslides occurred in disintegrated rock mass. The analysis further revealed that probability of landslides is high in disintegrated rock mass and the alluvial deposits as the Hazard index values are 2.13 and 1.71, respectively (Table 2). It may be noted that, Hazard index value of greater than '1' indicates more probability of the landslide occurrence (Girma et al. 2015; Chimidi et al. 2017). Perusal of Table 2 shows that slope material classes; residual soils and blocky rock mass have relatively less probability for landslide occurrence as the hazard index values for these classes is less than '1'. Also, colluvial material shows hazard index value equal to 0.96 which is nearly close to 1; thus indicating some probability for landslide occurrence. High concentration of landslides in the slopes covered by the unconsolidated materials; disintegrated rock mass, alluvial deposits and colluvial deposits is related to the low shear strength of the material. Also, such material when saturated may become more prone for instability (Anbalagan 1992; Raghuvanshi et al. 2014a; 2014b).

The overlay analysis of past landslides with slope inclination map further revealed that 31% of landslides fall

in slope class $> 38^\circ$, 27% fall in slope class 10-25° and 20% in slope class 25-38°. Also, it was found that the hazard index values for slope classes; $>38^\circ$, 10-25° and 25-38° are 1.4, 1.18 and 1.13, respectively (Table 2). Since all these hazard index values are greater than '1' therefore it shows that the slope classes $>38^\circ$, 10-25° and 25-38° are more susceptible for landslides. The high concentration of landslides in steeper slope sections are possibly related to the fact that as the slope becomes steeper the shear stress and tangential component of weight within the slope increases, thus tendency of slope instability increases (Ahmed 2009; Raghuvanshi et al. 2015). Also, about 42% of the past landslides were observed in the gentle slope sections (slope inclination 5 – 25°). It was observed that slopes having inclination in between 5 – 25° in the study area are mostly occupied by unconsolidated material which is considered to be highly susceptible for the slope instability (Raghuvanshi et al. 2014a; 2014b).

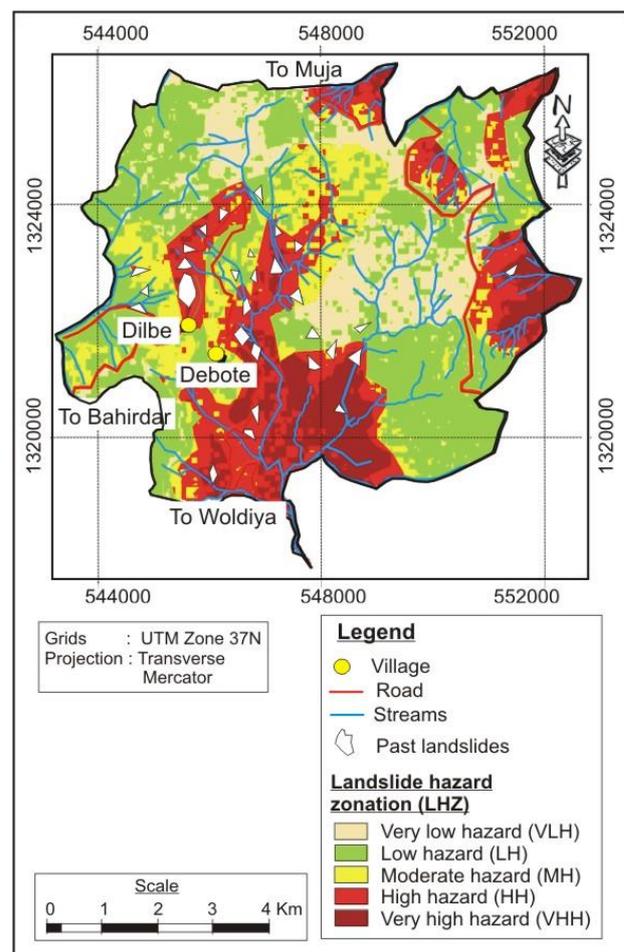


Figure 8: Landslide Hazard Zonation (LHZ) map

The overlay analysis of past landslides with aspect map showed that about 54% of the past landslides occurred on the slopes that are oriented towards South, Southeast and East directions. Also, it was found that hazard index values for aspect classes South, Southeast and East are 1.14, 1.26 and 1.25, respectively. Since all these hazard index values are greater than '1', therefore slope sections which are inclined towards South, Southeast and East directions show more probability for landslide occurrence. Further, when past landslides are compared to the distribution of springs in the area, it was found that most of the springs

are present on the slopes that are inclined towards South or East directions. Thus, the high concentration of the past landslides in the area may possibly be related to the general groundwater flow direction, presence of unconsolidated material over the slopes and oversaturated slope material. All these conditions are responsible for slope instability which might have possibly resulted into landslides (Arora 1997; Hoek and Bray 1981; Raghuvanshi et al. 2014a).

The overlay analysis between past landslides and the elevation map of the study area showed that about 66.5% past landslides occurred in two elevation classes, 2243-2543m and 2543-2843m. Also, the hazard index values computed for the elevation classes 2243-2543m and 2543-2843m are 1.56 and 1.35, respectively (Table 2). These hazard index values clearly show that elevation classes 2243-2543m and 2543-2843m are susceptible for slope instability as these hazard index values are greater than '1'. Further, it was observed that these elevation classes in the study area are dominated by high concentration of springs, disintegrated rock mass, colluvial and alluvial soils. Besides, slopes on these elevation classes are being utilized for cultivation purpose. Cultivation practice may possibly trigger slope instability by frequent unplanned irrigation that saturates unconsolidated material (Raghuvanshi et al. 2014a). Further, disintegrated rock mass, colluvial and alluvial soils with presence of springs make these slopes more susceptible for instability (Raghuvanshi et al. 2015).

The past landslide data also showed that 26.3% of past landslides fall in the sparsely vegetated land, 24% of landslide falls in Bush land and 6.6% of landslides fall in bare land. Also, Hazard index values for sparsely vegetated land, bush land and bare land are 1.27, 1.23 and 1.06, respectively (Table 2). Since, hazard index values for sparsely vegetated land, bush land and bare land are greater than '1' therefore these land use and land cover classes have more probability for landslide occurrence. Further, it was also observed that these land use and land cover classes are mainly occupied by disintegrated rock mass, colluvial and alluvial soils. These unconsolidated materials are susceptible for slope instability (Raghuvanshi et al. 2014b).

The past landslides data further showed that about 80% of the landslides fall within hydrological homogeneous zone classes HGZ-I and HGZ-II. Also, Hazard index value for HGZ-I class is 1.27 which shows high probability for landslide occurrence (Table 2). Similarly, Hazard index value for HGZ-II class is 0.98 which is nearly close to '1' thus it also shows relative probability for landslide occurrence. The high concentration of landslides in HGZ-I and HGZ-II zones show direct relation of landslides with the springs in the area. Also, it may be seen that HGZ-I and HGZ-II zones are dominated by disintegrated rock mass, colluvial and alluvial soils. These materials have weak shear strength and become more susceptible to instability when they are relatively saturated (Raghuvanshi et al. 2015). Thus, all these conditions make HGZ-I and HGZ-II zones more susceptible for slope instability. For this reason, only about 80% of the landslides fall within

hydrological homogeneous zones HGZ-I and HGZ-II in the study area.

6.2 Landslide hazard zonation (LHZ) - distribution

Landslide hazard evaluation shows that 11.55% (5.78km²) of the study area falls into 'very high hazard' (VHH), 19.84% (9.92km²) area falls into 'high hazard' (HH), 14.36% (7.18km²) area falls into 'medium hazard' (MH), 38.25% (19.13km²) area falls into 'low hazard' (LH) and remaining 16% (8km²) of area falls into 'very low hazard' (VLH). Further, perusal of Fig. 8 shows that VHH and HH zones are mainly concentrated in the southern, northern and eastern parts of the study area. The area delineated as VHH and HH zones have good concentration of springs. In general, it has been observed that spring locations generally have direct correlation with the landslides occurrences (Girma et al. 2015; Raghuvanshi et al. 2015; Chimidi et al. 2017). Also, majority of the areas in VHH and HH zones have slopes that are inclined at 10 to 38° and generally have unconsolidated deposits and disintegrated rock mass. As stated earlier, the unconsolidated materials and disintegrated rock mass have relatively low shear strength and such material when saturated may become more prone for instability (Anbalagan 1992; Raghuvanshi et al. 2014a; 2014b). Further, the MH zones are distributed towards Northern, Central and Western regions of the study area (Figure 8). On the other hand; LH zones are scattered in the study area. The VLH zones are mainly concentrated in the Northern and the Eastern parts of the study area. Also, VLH and LH zones fall mainly in slopes that are inclined at slope angles less than 5° and are mainly composed of blocky rock mass.

6.3 LHZ map validation

In order to check the validity of the LHZ map, prepared during the present study, an overlay analysis was made with the past landslides activity in the area (Figure 8). It is believed that if the past landslides fall either within high hazard (HH) or very high hazard (VHH) zones the prepared LHZ map is validated and it has possibly delineated all hazard zones appropriately in the study area. The overlay analysis results showed that 74% (22) of the past landslides fall within VHH and HH zones of the prepared LHZ map. Further, 17% (5) of the past landslides fall into moderate hazard (MH) zone that also have reasonable probability of landslide occurrence. Only 9% (3) of the past landslides fall within low hazard (LH) and very low hazard (VLH) of the LHZ map. Thus, it may be concluded from these results that the prepared LHZ has reasonably validated with the past landslide data. About 9% of the past landslides that fall within LH or VLH zones do not validate with the prepared LHZ map. This variation in the validity of the LHZ map may be due to the limitation of the methodology followed in the present study. The present study was conducted on medium scale and many factors that are responsible for instability of slopes cannot be considered at this scale (Ayele et al. 2014). These factors are discontinuity characteristics and the relationship of discontinuities with the slope, water pressures within the slope, shear strength of the material along the potential discontinuity surfaces etc. (Girma et al. 2015; Chimidi et al. 2017; Hamza and Raghuvanshi 2017).

7. Conclusion

Landslide hazard evaluation in the present study shows that 11.55% (5.78km²) of the study area falls into 'very high hazard' (VHH), 19.84% (9.92km²) area falls into 'high hazard' (HH), 14.36% (7.18km²) area falls into 'medium hazard' (MH), 38.25% (19.13km²) area falls into 'low hazard' (LH) and remaining 16% (8km²) of area falls into 'very low hazard' (VLH). Further, validation results showed that 74% (22) of the past landslides fall within VHH and HH zones of the prepared LHZ map. Further, 17% (5) of the past landslides fall into moderate hazard (MH) zone that also have reasonable probability of landslide occurrence. Only 9% (3) of the past landslides fall within low hazard (LH) and very low hazard (VLH) of the LHZ map. The prepared LHZ map has reasonably validated with the past landslide data. Finally, in general it may be concluded that about 31% area falls into VHH and HH zones and about 14% of the area falls into MH zone. Thus, about 45% of the area is prone for the landslide hazard. The major factors responsible for landslide hazard in the area are the susceptible slope material comprising mainly; disintegrated rock mass, alluvial, colluvial and residual soils. Also, slopes that are inclined at moderate to steep slope angles and are oriented towards South, Southeast and East directions are more susceptible for instability. Besides, slopes which fall in between elevations 2243 to 2843m have also shown potential instability. The landslides in the study area have been triggered mainly during the rainy season. Since the landslides in the area have been causing considerable damage to the roads, houses and the agricultural land therefore there is a need to implement mitigation measures, particularly in hazardous zones delineated through the present study.

Acknowledgements

We thankfully acknowledge all kinds of support extended to us by the head and the staff of the School of Earth Sciences, College of Natural and Computational Sciences, Addis Ababa University. The meticulous review and comments by the esteemed anonymous reviewers is also duly acknowledged.

References

Abebe, B., F. Dramis, G. Fubelli, M. Umer and A. Asrat (2010). Landslides in the Ethiopian highlands and the rift margins, *Journal of African Earth Sciences*, 56, 131–138.

Ahmed, S. (2009). Slope stability analysis using GIS and numerical modeling techniques (Unpublished MSc thesis), Vrije Universiteit, Brussel.

Anbalagan, R. (1992). Landslide hazard evaluation and zonation mapping in mountainous terrain, *Engineering Geology*, 32, 269–277.

Arora, K.R. (1997). Soil mechanics and foundation engineering, Standard Publishers Distributors, Delhi, India, pp. 475.

Arnous, M.O. (2011). Integrated remote sensing and GIS techniques for landslide hazard zonation: a case study WadiWatieer area, South Sinai, *Egyptian Journal of Coastal Conservation* 15, 477–497

Ayalew, L., H. Yamagishi and N. Ugawa (2004). Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan, *Landslides*, 1, 73–81.

Ayalew, L. and H. Yamagishi (2004). Slope failures in the Blue Nile basin, as seen from landscape evolution perspective, *Geomorphology*, 61, 1–22.

Ayele, S., T.K. Raghuvanshi and P.M. Kala (2014). Application of remote sensing and GIS for landslide disaster management—a case from Abay Gorge, Gohatsion–Dejen section, Ethiopia. *Landscape ecology and water management, proceedings of International Geographical Union (IGU) Rohtak Conference, Advances in Geographical and Environmental Sciences* (M. Singh, R.B. Singh and M.I. Hassan: Editors), Springer, Japan, Vol. 2, pp. 15–32

Ayeneu, T. and G. Barbieri (2005). Inventory of landslides and susceptibility mapping in the Dessie area, Northern Ethiopia, *Engineering Geology*, 77, 1–15

Bommer, J. J. and C. E. Rodri'guez (2002). Earthquake-induced landslides in Central America. *Engineering Geology*, 63, 189–220

Carrara, A., M. Cardinali and F. Guzzetti (1992). Uncertainty in assessing landslide hazard and risk, *ITC Journal*, 2, 172–183

Carrara, A., M. Cardinali, R. Detti, F. Guzzetti, V. Pasqui and P. Reichenbach (1991). GIS techniques and statistical models in evaluating landslide hazard. *Earth Surface Process Landform*, 6, 427–445

Casagli, N., F. Catani, C. Puglisi, G. Delmonaco, L. Ermini and C. Margottini (2004). An inventory-based approach to landslide susceptibility assessment and its application to the Virginio River Basin. *Italy Environ, Journal of Environmental and Engineering Geosciences*, 10 (3), 203–216

Chimidi, G., T. K., Raghuvanshi and K. V. Suryabagavan (2017). Landslide hazard evaluation and zonation in and around Gimbi town, western Ethiopia – a GIS-based statistical approach, *Journal of Applied Geomatics*, 9 (4), 219–236.

Collison A., S. Wade, J. Griffiths and M. Dehn (2000). Modelling the impact of predicted climate change on landslide frequency and magnitude in SE England, *Engineering Geology*, 55, 205–218

Crozier, M.J. and T. Glade (2005). Landslide hazard and risk: issues, concepts, and approach. Glade, *Landslide*

- Hazard and Risk, (T. Anderson and M. Crozie: Editors) Wiley, Chichester, pp. 1–40
- Dahal, R.K., S.Hasegawa, T.Masuda and M.Yamanaka (2006). Roadside slope failures in Nepal during torrential rainfall and their mitigation. *Disaster Mitigation of Debris Flows, Slope Failures and Landslides*, pp 503–514
- Dai, F. C., C. F. Lee and Y. Y. Ngai (2002). Landslide risk assessment and management: an overview. *Engineering Geology*, 64, 65–87
- Dai, F. C. and C. F. Lee (2001). Terrain-based mapping of landslide susceptibility using a geographical information system: a case study, *Canadian Geotechnical Journal*, 38, 911–923
- Demissie, T., G. Yohannes, A. Mammo, Y. Tesfaye, Y. Teshome, G. Burusa, M. Edris and M. Wenduante (2010). *Geology geochemistry and gravity survey of the Dessie area*, unpublished report, pp. 68.
- Ermias, B., T. K. Raghuvanshi and B. Abebe (2017). Landslide Hazard Zonation (LHZ) around Alemketema Town, North Showa Zone, Central Ethiopia - A GIS based expert evaluation approach *International Journal of Earth Sciences and Engineering*, 10 (1), 33 -44
- Fall, M., R. Azzam and C. Noubactep (2006). A multi-method approach to study the stability of natural slopes and landslide susceptibility mapping, *Engineering Geology*, 82, 241–263
- FAOUN (Food and Agriculture Organization of the United Nations) (1986). *Ethiopian highlands reclamation study*, Food and Agriculture Organization of the United Nations, Rome, pp. 334
- Gemechu, D. (1977). *Aspect of climate and water balance in Ethiopia*. Addis Ababa University press, Addis Ababa, pp. 79
- Girma, F., T.K. Raghuvanshi, T. Ayenew and T. Hailemariam (2015). Landslide hazard zonation in AddaBerga District, Central Ethiopia—a GIS based statistical approach, *Journal of Geomatics*, 9, 25–38
- Gorsevski, P.V., P. Jankowski and P.E. Gessler (2006). A heuristic approach for mapping landslide hazard by integrating fuzzy logic with analytic hierarchy process. *Control Cybernet*, 35 (1), 121–146
- GSE (Geological Survey of Ethiopian), (2010). *Geosciences Report on Dessie area*, Unpublished Report, Geological Survey of Ethiopian, Addis Ababa, Ethiopia, pp. 68
- Guzzetti, F., A. Carrara, M. Cardinali and P. Reichenbach (1999). Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, central Italy. *Geomorphology*, 31(1–4), 181–216
- Hamza, T. and T. K. Raghuvanshi (2017). GIS based landslide hazard evaluation and zonation—a case from Jeldu District, Central Ethiopia. *Journal of King Saud University–Science*, 29, 151–165
- Hoek, E and J. W. Bray (1981). *Rock slope engineering (revised third edition)*. Institute of Mining and Metallurgy, London, pp. 358
- Kanungo, D. P., M. K. Arora, S. Sarkar and R.P. Gupta (2006). A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas, *Engineering Geology*, 85,347–366
- Keefer, D. V. (2000). Statistical analysis of an earthquake-induced landslide distribution—the 1989 Loma Prieta, California event. *Engineering Geology*, 58, 231–249
- Lan, H. X., C. H. Zhou, L. J. Wang, H. Y. Zhang and R. H. Li (2004). Landslide hazard spatial analysis and prediction using GIS in the Xiaojiang watershed, Yunnan. China, *Engineering Geology*, 76, 109–128
- Lee, S., J.H. Rhyu, J. S. Won and H. J. Park (2004). Determination and application of the weights for landslide susceptibility mapping using an artificial neural network at south. *Korean Engineering Geology*, 71, 289–302
- Leroi, E. (1997). *Landslide risk mapping: problems, limitation and developments. Landslide risk assessment*, (F. Cruden : Editor), Balkema, Rotterdam, pp. 239–250
- Marrapu, B. M and R. S. Jakka (2014). Landslide Hazard Zonation Methods: A Critical Review, *International Journal of Civil Engineering Research*, 5, 215-220
- Pan, X., H. Nakamura, T. Nozaki and X. Huang (2008). A GIS-based landslide hazard assessment by multivariate analysis Landslides. *Journal of Japan Landslide Society*, 45 (3), 187–195.
- Parise, M. and R.W. Jibson (2000). A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January, 1994 Northridge, California earthquake, *Engineering Geology*, 58, 251–270
- Raghuvanshi, T. K., L. Negassa and P.M. Kala (2015). GIS based grid overlay method versus modeling approach—a comparative study for landslide hazard zonation (LHZ) in Meta Robi District of west Shewa Zone in Ethiopia. *Egyptian Journal of Remote sensing and Space Science*, 18(2), 235–250
- Raghuvanshi, T. K., J. Ibrahim and D. Ayalew (2014a). Slope stability susceptibility evaluation parameter (SSEP) rating scheme—an approach for landslide hazard zonation. *Journal of African Earth Sciences*, 99, 595–612
- Raghuvanshi, T.K., P.M. Kala and M. Singh (2014b). Landslide disaster management and reduction- An approach through remote sensing and GIS. *Landscape*

- ecology and water management, proceedings of International Geographical Union (IGU) Rohtak Conference, Advances in Geographical and Environmental Sciences (M. Singh, R.B. Singh and M.I. Hassan: Editors), Springer, Japan, Vol. 2, Springer Japan, pp. 33-40
- Sarkar, S. and D.P. Kanungo (2004). An integrated approach for landslide susceptibility mapping using remote sensing and GIS, *Photogrammetric Engineering & Remote Sensing*, 70 (5), 617–625
- Sarkar, S., D.P. Kanungo and G.S. Mehrotra (1995). Landslide hazard zonation: a case study in Garhwal Himalaya, India, *Mountain Research and Development*, 15 (4), 301–309.
- Tefera, M., T. Chernet and W. Haro (1996). Explanation of the geological map of Ethiopia, *Ethiopian Institute of Geological Surveys*, Addis Ababa, 3: 79
- Van Den Eckhaut M., P. Reichenbach, F. Guzzetti, M. Rossi and J. Poesen (2009). Combined landslide inventory and susceptibility assessment based on different mapping units: An example from the Flemish Ardennes, Belgium. *Journal of Natural Hazards Earth System Science*, 9, 507–521.
- Varnes, D.J. (1996). *Landslide Types and Processes, Landslides: investigation and mitigation*, transportation Research Board Special Report 247, (A.K. Turner and R.L Schuster: Editors), National Academy Press, National Research Council, Washington, D.C., pp. 36-75
- Varnes, D.J. (1984). *Landslide hazard zonation: a review of principles and practice*. UNESCO, Paris, pp. 1–63
- Wang, X. and R. Niu (2009). Spatial forecast of landslides in three gorges based on spatial data mining, *Sensors*, 9, 2035–2061
- WestenVan, C. J., N. Rengers and M. T. J. Terlien (1997). Prediction of the occurrence of slope instability phenomena through GIS-based hazard zonation. *Geologische Rundschau*, 86, 4004–441

Preface

Special Section: Selected papers of National Symposium on Advancements in Geospatial Technology for Societal Benefits, Dec 03-07, 2018, Ahmedabad (Part-1)

Remotely sensed applications have grown and matured into operational uses in planning and management of natural resources in our country. It is also providing valuable information on climate change in terms of occurrence of extreme events, glacier retreat, increase in air pollution, sea level rise, depletion of ground water and addressing various dimensions of Earth Sciences. As the field of remote sensing continues its rapid expansion, many interesting papers were presented in the “National Symposium on Advancements in Geospatial Technology for Societal Benefits”, jointly organised by the Indian Society of Geomatics & Indian Society of Remote Sensing during Dec 03-07, 2018, hosted at Space Applications Centre, ISRO, Ahmedabad.

Papers were presented on various themes such as Geosciences, Agriculture, Navigation and Mobile Apps, Planetary Sciences, Environment, Geospatial Technology, Soil/Land Degradation, Machine Learning, Forestry, Coastal Applications, Hyperspectral Applications, Urban Applications, Marine Sciences, Microwave Applications, Advancements in Data Processing, Cryosphere and Water Resources. Some of the papers of the Symposium were submitted for publication in the Journal of Geomatics. These papers are being published as Special Section in two parts. Part-I in the current issue includes thirteen papers related with urban sprawl, night-time light, generation of Mars image mosaic, mangroves, ground penetrating radar, development of visualization platform, atmospheric carbon dioxide assessment, offshore wind farm site suitability, rip currents, change detection and trend analysis, spatial enhancement of SWIR band, object-based classification and generation of multi temporal co-registered optical remote sensing images. Few more research papers shall be published as Part-II of the Special Section in the forthcoming October issue.

We thank all the authors who have contributed in this Special Section. We are grateful to all the reviewers who have taken out their precious time to provide valuable suggestions and comments to help us in selecting and improving the respective manuscripts.

R.P. Singh
R. Ratheesh
Surisetty V.V. Arun Kumar

-Editors

Remote sensing based time-series analysis for monitoring urban sprawl: A case study of Chandigarh capital region

Varinder Saini* and Reet Kamal Tiwari

Department of Civil Engineering, Indian Institute of Technology Ropar, Nangal Road, Rupnagar, Punjab-140001, India

*Email: varinder.saini@iitrpr.ac.in

(Received: Jan 02, 2019; in final form: May 09, 2019)

Abstract: Urban growth is a global phenomenon, which is happening at unprecedented rates. According to the United Nations, the current world population of 7.6 billion is expected to reach 8.6 billion in 2030 out of which 5 billion will be urban population. Much of this urbanization will unfold in Africa and Asia, bringing huge social, economic and environmental transformations. The main reasons for urban growth are an unorganized expansion, increased immigration, and population explosion. In this context, changes in land cover are considered one of the central components for managing natural resources and monitoring environmental impacts. In the present study, a supervised classification applied to a time-series data of Landsat collected over time (1990, 2001, and 2015) gave an insight into the urban sprawl patterns in the Chandigarh capital region. The results show an increasing trend of urban sprawl in the outskirts of the city, especially in the northwestern and southern directions as a result of the development of an international airport, new sectors, and approach roads. Post-classification change detection points to the fact that maximum changes were seen in the class vegetation as it has rapidly changed to built-up/urban areas. Based on the results of such studies, the city and town planners may be able to take timely and appropriate steps to minimize the environmental implications associated with such urbanization.

Keywords: Urban sprawl, land cover classification, environmental impact, Chandigarh capital region.

1. Introduction

Urbanization is a global phenomenon, which is happening at unprecedented rates. According to the United Nations, the current world population of 7.6 billion is expected to reach 8.6 billion in 2030 out of which 5 billion will be urban population (UNDESA 2017). Urban sprawl, also called sprawl or suburban sprawl, is the rapid expansion of urban areas outside the geographic extent of existing cities and towns. It is a natural process of urbanization but is haphazard and unplanned expansion and therefore, its boundary cannot be determined. Sprawl is characterized by low-density, mono-functional or single-use communities where use of private automobile is the main source of transportation. The process of sprawl begin with initiation of migration of people from other areas to the urban areas. With time, migration leads to increase in population and in the long run the sprawl will occur due to natural growth. The rates of urban sprawl are not accelerating in industrialized nations as 75% of their population has already been urbanized (UNDESA, 2004). However, by 2050, 2.5 billion people will be added to world's urban population, with almost 90% of this growth unfolding in Africa and Asia, bringing huge social, economic and environmental transformations (UNDESA, 2018).

One very glaring consequence of sprawl in developing countries is the change of agricultural land to urban built-up areas. Infact, many cities are growing at their fringes transforming the villages and agricultural lands to industrial, commercial or low density residential clusters (Huang et al., 2009). A number of Indian cities have been studied in the context of urban sprawl (Bhat et al., 2017; Jain and Sharma, 2018; Anees et al., 2018). Thus, urban sprawl has become a major issue facing many cities and it is critical to study and monitor it in a spatial context. In the

case study described here, urban sprawl is seen as an increase in built-up area transformed from rural and agricultural lands. One of the methods to detect changes in an area is by comparing multi-temporal classification derived from satellite images. Remote sensing technology coupled with geographic information system (GIS) tools by virtue of its several advantages has cemented its place in many applications and thus may also prove useful in characterize the dynamics of land cover change in the present study as well. In the present paper, firstly we instigate the context of the study, then we describe the study area and the methodology adopted, and later discuss the results of our study.

2. Materials and Methods

2.1 Description of study area

Chandigarh, also called the City Beautiful, is the first planned city of India which is located in North India at 76° 47' 14" E and 30° 44' 14" N with an average elevation of 335 m above mean sea level. It was inaugurated in 1952 by the first Prime minister of independent India Jawaharlal Nehru and was designed by French architect- Le Corbusier. Chandigarh was declared a union territory on 01 November 1966 and serves as the administrative capital of the states of Punjab and Haryana and lies 160 miles north of the national capital New Delhi. The total area of the city is 114 m². It is bordered on three sides by Punjab (north, west, and south) and on the eastern side by Haryana. It is a part of the Chandigarh capital region (CCR) which includes Chandigarh and the neighbouring cities of Panchkula (in Haryana), Sahibzada Ajit Singh (SAS) Nagar, Zirakpur, Kharar, and Mullanpur (all in Punjab). For the present study, this region (total area approx. 675 sq km) has been taken into consideration (Figure 1).

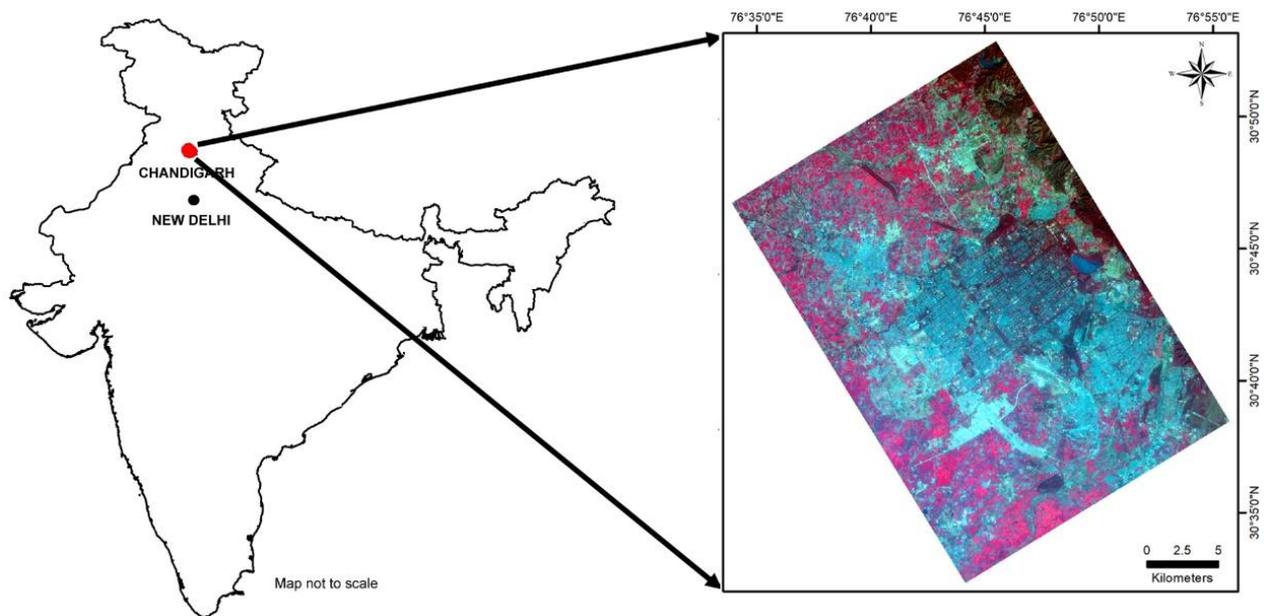


Figure 1: Geographic location and Landsat OLI image (FCC: R = NIR, G = Red, B = Green, Dec. 2015) of the study area

2.2 Data used

Cloud-free multi-temporal datasets of the study area, acquired from TM, ETM+ and OLI sensors on-board Landsat satellites, and covering a span of 25 years from 1990 to 2015, were acquired from the USGS Landsat archives as L1 data products. The optical data is available at 30 m spatial resolution and is available free of cost for anybody to use. Ancillary data such as Survey of India topographical maps (Nos. H43K9, H43K10, H43K13, H43K14); at a scale of 1:50000 were used to georeference the satellite data.

2.3 Land cover classification

In order to obtain information about land cover and urban sprawl areas, a time-series data of Landsat pertaining to three years - 1990 (TM), 2001 (ETM+) and 2015 (OLI) have been processed and classified. The data belonged to same season so as to avoid any uncertainty due to inter-annual variability. Although several image classification techniques have been proposed in literature (Mountrakis et al., 2011; Ranga et al., 2011; Wenzhi et al., 2012) but the most prominent classification techniques is pixel based image classification which is largely based on the spectral information of the pixels.

Digital image processing software ERDAS Imagine (v. 2016) has been used to process, analyse, and integrate the spatial data. Firstly, the multi-temporal data was geometrically registered and later atmospheric correction was carried out using the ATCOR module of ERDAS Imagine. The corrected images were then classified into five classes namely Built-up, Sparse to Moderate vegetation, Dense vegetation, Barren Land and Water. A congruous number of training pixels were selected to train the classifier. Then, Maximum Likelihood Classifier algorithm was applied to each image and three different land cover maps were produced. Post-classification comparison was carried out by following the procedures defined in literature (Jensen et al., 1987; Dimiyati et al., 1996; Ward et al., 2000). This method offers the advantage

to allow the creation and the update of GIS databases, as class/categories are given, and quantitative values of each class can be determined (Fischera et al., 2012). Since our main focus is on the extent of Built-up/Urban area

2.4 Accuracy assessment

The accuracy of the classified objects was estimated using the error matrix approach (Congalton, 1991). For an unbiased assessment, the equalized random sampling method (Jensen et al., 1987) was used to select samples for each class. Three types of accuracies were calculated for each image namely- user's (UA), producer's (PA) and overall accuracy (OA) along with Kappa index of Agreement (KIA).

3. Results and discussion

In order to understand, characterize, and monitor the urban expansion process, the availability of time-series data is essential. The results of this study are shown in figure 2, which indicate that the built-up/Urban area has considerably modified the land cover of the study area, with significant land conversions. Table 1 shows the area of different land cover classes. The urbanization happened in all directions but maximum urbanization took place in the south (Zirakpur) and north-west (Kharar) directions outside the city, as a result of the development of an international airport, new sectors and approach roads on the vegetated areas. This exponential growth of built-up or urban areas has probably happened because of the increased educational and employment opportunities and better living standards in the CCR. Table 2 provides the results of the accuracies attained by means of a confusion matrix. The accuracies of all the three years is above the minimum accuracy threshold of 85% (Anderson et al., 1976). Out of the three years, the highest overall accuracy of 92.7% with a KIA of 0.912 was for the year 2015. A number of classes were more accurately identified than others such as class 'Sparse to Moderate Vegetation' (UA 93.7% and PA 95.8%) and 'Water' (UA 93.7% and PA 97.8%). The high accuracy of 2015 OLI data could be

attributed to its high radiometric resolution. Also, it is because of the fact of availability of more detailed and higher resolution reference maps (such as Google Earth imagery) in recent times for carrying out accuracy assessment.

Over the entire study period, agriculture land was the predominant land cover type, although it declined over

time by being changed to built-up/urban or barren land. Since we were mainly concerned with the extent of urban sprawl, therefore, for post-classification change detection, we compared the increase in built-up area versus Non built-up area for the three years (Figure 3). It could be observed that the built-up/urban area showed significant increase (~79%) over the 25-year period.

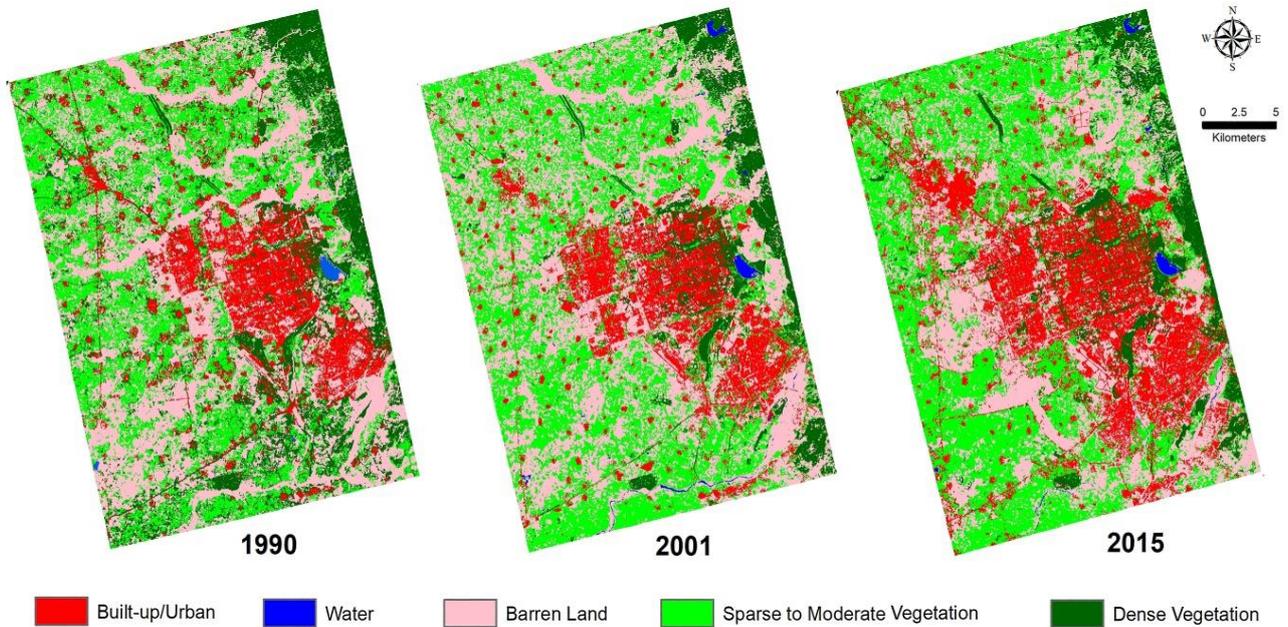


Figure 2: Land cover maps for the three years. Note the expansion in Built-up area (red) over the 25 year period

Table 1: Area of different land-cover classes

Land Cover Class	1990	2001	2015
Barren Land	85.7	94.9	153.3
Built-up	129.2	66.5	67.5
Sparse to Moderate Vegetation	192.1	233.2	227.7
Dense vegetation	266.9	277.5	224.4
Water	1.9	3.5	2.7

Table 2: Summary of the classification accuracies

Land cover class	1990 (TM)		2001 (ETM+)		2015 (OLI)	
	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)
Built-up/Urban	87.5	100	87.5	82.3	87.5	93.3
Water	93.7	100	87.5	100	93.7	97.8
Barren Land	93.8	87.5	93.8	78.9	92.4	84.3
Sparse to Moderate Veg.	91.5	78.9	85.7	93.3	94.6	88.2
Dense Vegetation	88.3	87.5	92.4	98.7	93.7	95.8
Overall Accuracy (%)	91.2		90.0		92.7	
Kappa Statistic	0.891		0.875		0.912	

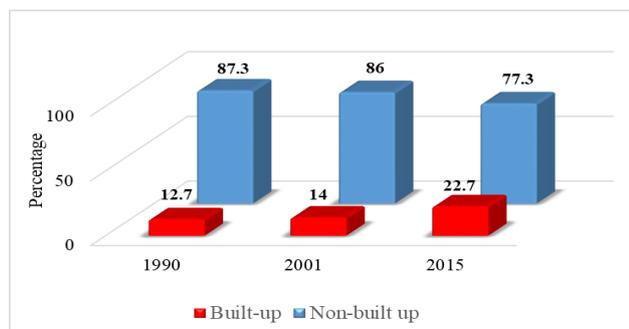


Figure 3: Percent cover of study area under built-up and non built-up

Urban sprawl is a complex diffusion process that is spreading dramatically and affecting rural landscape differently in space and at different scales. The continuing land cover changes in the CCR have different implications. Such kind of drastic change in land cover can negatively alter the potential use of an area and may ultimately lead to loss of productivity. This could, in turn, affect the local people by reducing the means of livelihood of those who depend on agriculture. Other implication associated with the increasing urbanization is increasing pressure on the available environmental resources and infrastructure of the city. Furthermore, such studies can provide information as indicators of the direction of change in the study area over the given period. The city and town planners may thus be able to take timely and appropriate steps to minimize the environmental implications associated with such urbanization.

Acknowledgements

Varinder Saini is grateful to the Science and Engineering Research Board (SERB), New Delhi, for providing financial assistance in the form of National Post-Doctoral Fellowship (NPDF) under grant number PDF/2017/002647. The authors thank the USGS (<http://glovis.usgs.gov/>) for providing the Landsat data used in this work free of cost through the internet. The authors also thank the anonymous reviewers for giving their valuable comments and improving the quality of the article.

References

Anderson, J. R., E. E. Hardy, J. T. Roach and R. E. Witmer (1976). A land use and land cover classification system for use with remote sensor data, US Geological Survey Professional Paper No. 964, USGS: Reston, WV, USA, pp. 28.

Anees, M.M., S. Sajjad and P.K. Joshi (2018). Characterizing urban area dynamics in historic city of Kurukshetra, India, using remote sensing and spatial metric tools, *Geocarto International*, DOI: 10.1080/10106049.2018.1499819.

Bhat, P.A., M. Shafiq, A.A. Mir and P. Ahmed (2017). Urban sprawl and its impact on landuse/land cover dynamics of Dehradun city, India, *International Journal of Sustainable Built Environment*, 6, 513-521.

Congalton, R. (1991). A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37, 35–46.

Dimiyati, M., K. Mizuno, S. Kobayashi and T. Kitamura (1996). An analysis of land use/ cover change using the combination of MSS Landsat and land use map-a case study in Yogyakarta, Indonesia, *International Journal of Remote Sensing*, 17, 931–944.

Fischera, C.R., G. Modica and M. Pollino (2012). Land Cover classification and change-detection analysis using multi-temporal remote sensed imagery and landscape metrics, *European Journal of Remote Sensing*, 45, 1-18.

Huang, B., L. Zhang and B. Wu (2009). Spatio-temporal analysis of rural-urban land conversion, *International Journal of Geographical Information Science*, 23, 379–398.

Jain, G. V. and Sharma, S. A. (2018). Spatio-temporal analysis of urban growth in selected small, medium and large Indian cities, *Geocarto International*, DOI: 10.1080/10106049.2018.1450450

Jensen, J.R., E.W. Ramsat, H.E. Mackey, E.J. Christensen and R.P. Sharitz (1987). Inland wetland change detection using aircraft MSS data, *Photogrammetric Engineering and Remote Sensing*, 53, 521-529.

Mountrakis, G., J. Im and C. Ogole (2011). Support vector machines in remote sensing: A review, *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259.

Ranga, R. V., B. Eddie, V. Chandola, B. Bhaduri, C. Anil and G. Jordan (2011). Machine learning approaches for high-resolution urban land cover classification: a comparative study. 2nd International Conference on Computing for Geospatial Research & Applications (ACM), Article 11, pp. 1-10.

UNDESA (United Nations Department of Economic and Social Affairs), (2018). World Urbanization Prospects: The 2018 Revision. (New York: United Nations). Available <https://esa.un.org/unpd/wup/>

UNDESA (United Nations Department of Economic and Social Affairs), (2017). World population prospects: The 2017 revision (New York: United Nations). Available <https://www.un.org/development/desa/publications/world-population-prospects-the-2017-revision.html>

UNDESA (United Nations Department of Economic and Social Affairs), (2004). World Urbanization Prospects: The 2003 Revision (New York: United Nations). Available <http://www.un.org/esa/population/publications/wup2003/WUP2003Report.pdf>

Ward, D., S.R. Phinn and A.T. Murray (2000). Monitoring growth in rapidly urbanizing areas using remotely sensed data, *Professional Geographer*, 52, 371-386.

Wenzhi, L., R. Bellens, A. Pizurica, W. Philips and P. Youguo (2012). Classification of hyperspectral data over urban areas using directional morphological profiles and semi-supervised feature extraction. *IEEE J-STARS*, 5(4), 1177–1190.

Evaluation of the consistency of DMSP-OLS and SNPP-VIIRS night-time light datasets

Reshma Jeswani^{*1,2}, Anurag Kulshrestha², Prasun Kumar Gupta¹ and S.K. Srivastav¹

¹Indian Institute of Remote Sensing (IIRS), Kalidas Road, 248001, Dehradun, India

²Faculty of Geoinformation Science and Earth Observation (ITC), University of Twente, Netherlands

*Email: reshmajeswani888555@gmail.com

(Received: Jan 02, 2019; in final form: May 13, 2019)

Abstract: Night-time light remote sensing is used to monitor human impact on earth's surface. DMSP-OLS was the famous source of night-time light data until the release of a newer version of night-time light remote sensing satellite: SNPP-VIIRS, which has several improvements over DMSP-OLS. However, the dissemination of DMSP-OLS datasets was ceased for public use after 2013. The SNPP-VIIRS satellite products are available from April 2012. Due to this unavailability of the datasets prior to 2012 for VIIRS and after 2013 for DMSP, raises the major importance of homogeneous long term datasets in understanding the impact of the past and to have a comparative and change analysis prior and later to 2013. Due to the need of a long term homogeneous datasets, the evaluation of the consistency of the DMSP-OLS and SNPP-VIIRS has been foreseen. For this, both the datasets are harmonized temporally, spatially and radiometrically. The study presents the inter-calibration between DMSP and VIIRS using pseudo invariant features (PIFs). For extracting suitable PIFs, combinations of Getis Ord statistics (G_i^*) and coefficient of variation (CV) are used on 3×3 and 5×5 pixel window sizes. Various thresholds are applied to G_i^* and CV and the optimal combination is selected to identify suitable PIFs. Logarithmic and power functions are used for DMSP and VIIRS to find the optimal fit. The calibrated image is validated against socio-economic indicators: Gross State Domestic Product (GSDP) and Electricity consumption (EC). The model is found to be a good fit at national scale (India) but doesn't fit very well on regional scales. Significant improvements can be made to ensure better consistency by the application of different techniques of harmonization, selection of PIFs and application of the inter-calibration model.

Key words: DMSP-OLS, SNPP-VIIRS, Inter-calibration, Simulation, socio-economic indicators

1. Introduction

Night-time light imagery is a unique source which provides a distinct view of Earth's surface and human activities. It has been observed by various studies, that night lights have a significant importance in monitoring the earth surface activities and atmospheric phenomena. Night-time visible imaging was initiated by the Defence Meteorological Satellite Program (DMSP) Operation Linescan System (OLS) in 1960s. This mission was the only source of night time visible images until the launch of Suomi National Polar-orbiting Partnership's visible Infrared Imaging Radiometer Suite (SNPP-VIIRS) in October 2011 which continued the acquisition of night time visible images with some quality enhancements over DMSP-OLS.

The DMSP satellite of the U.S Department of Defence provides a global coverage every 24 hours. The on board OLS sensor acquired images in visible, near infrared and thermal infrared region of the electromagnetic spectrum. It was the longest running time series of night time lights data (Huang, et al., 2014). DMSP-OLS had an oscillating scan radiometer with two spectral bands: Visible Near infrared (VNIR) for Night-time Light (NTL) and thermal Infrared. The night-time overpass is between 20:30 and 21:30 UTC. The data is reported in Digital Number (DN) values on a six-bit scale that ranges from 0 (no light) to 63 (maximum light). A total of 9 satellites from F10 to F18 have collected OLS data. The data is mainly of three types: stable lights, radiance calibrated and average digital number (Doll, 2008).

SNPP-VIIRS provides imagery during the day as well as at night. It has a greater sensitivity in comparison to

DMSP-OLS and can reveal urban details. VIIRS has better spatial resolution than OLS (742 m), wider radiometric detection range, vast reduction in the pixel footprint, wider dynamic range, finer quantization, in-flight calibration and considerable better geometric quality. It also eliminates the critical problems of saturation, blooming and the lack of on-board calibration that DMSP has (Elvidge et al., 2017). Additionally, Day-Night Band (DNB) pixels don't expand and distort towards the edge of the scan as in DMSP-OLS data. (Seaman, 2013).

1.1 Need for relative normalization (Inter-Calibration)

VIIRS DNB has many key improvements over DMSP-OLS, but both have a common fundamental purpose of detection of clouds using moonlight as an illumination source. Also, the stable lights datasets obtained from both sensors have proven to be beneficial in natural and social sciences (Elvidge et al., 2017). Due to the lack of DMSP datasets in the last 6 years, it is not possible to use these data for long-term analysis until the present time, which would provide a great insight for understanding the various application perspectives in a long run. Pandey et al. (2017) reviewed various methods for inter-calibration of DMSP-OLS and highlighted future possibility of inter-annual correction using VIIRS.

1.2 In multi-temporal socio-economic context

Various studies have shown the possibility of correlations between NTL and several known demographic and economic variables. Since the NTL data is globally uniform and continuously measured, it constitutes an important data source where reliable statistics are otherwise lacking (Henderson et al., 2012). Night time lights images provides potential possibilities to quantify

the socio-economic variables with accuracy and spatial information. NTL data can provide essential foundation on estimation of GDP, population, electric power consumption, greenhouse gas emissions, poverty index and other parameters on the basis of spatio-temporal continuity, independence and objectiveness. Previous studies of Ghosh et al. (2010) and Mellander et al. (2015) have indicated the importance of NTL as a proxy for measuring economic growth of a nation, especially when the statistics show important shortcomings.

The main aim of the research is to evaluate the consistency of annual DMSP-OLS and SNPP-VIIRS datasets and to develop a statistical model for the inter-calibration using Pseudo Invariant features (PIFs), and also to validate the model at different scale level against socio-economic parameters.

2. Study area and datasets

The study area selected for this research was entire geographical span of India (6.45° N to 37.6° N and 68.7° E to 97.25° E). India is bounded by Arabian Sea on the southwest, Indian Ocean on the south and Bay of Bengal on the southeast. It shares the land boundary with China, Bhutan, Myanmar, Bangladesh, Pakistan and Nepal. India is world's second most populous country with the number of residents reported roughly as 1.3 billion (Ministry of Home Affairs, 2011). India is a diverse country with variations in the development at various regional and state levels. It covers areas of illumination regions which depicts the heterogeneity of luminosity spread over the whole country. This diversity and variations make it suitable to study the growth from the NTL images and to encounter differences in radiances at national and regional scales.

The datasets of DMSP-OLS and SNPP-VIIRS are taken from NOAA. The datasets are available in GeoTIFF format. The DMSP-OLS data is available annually under Version 4 stable light products from 1992 to 2013. The annual product of the year 2013 was used. The SNPP-VIIRS-DNB datasets are available monthly from April 2012 to December 2017. The monthly composites from January to December for the year 2013 were used. The DMSP-OLS images are available as global tiles, having stable 6-bit radiometric quantization. This stable light product is resampled at a resolution of 1 km and is provided by the NOAA in WGS84 coordinate system. From the global tile, the region of India is extracted for the research.

Due to unavailability of the annual composite of SNPP-VIIRS-DNB for the year 2013, there is a need for an annual composite to temporally harmonize VIIRS with DMSP-OLS. For this, the annual composite for VIIRS 2013 is created using monthly images of India from January 2013 to December 2013 excluding the month of May. This exclusion is done to avoid improper outputs due to the presence of data artefacts in May 2013 (Cao, 2013).

3. Methodology

Flow chart representing the methodology adopted for this research is shown in figure 1.

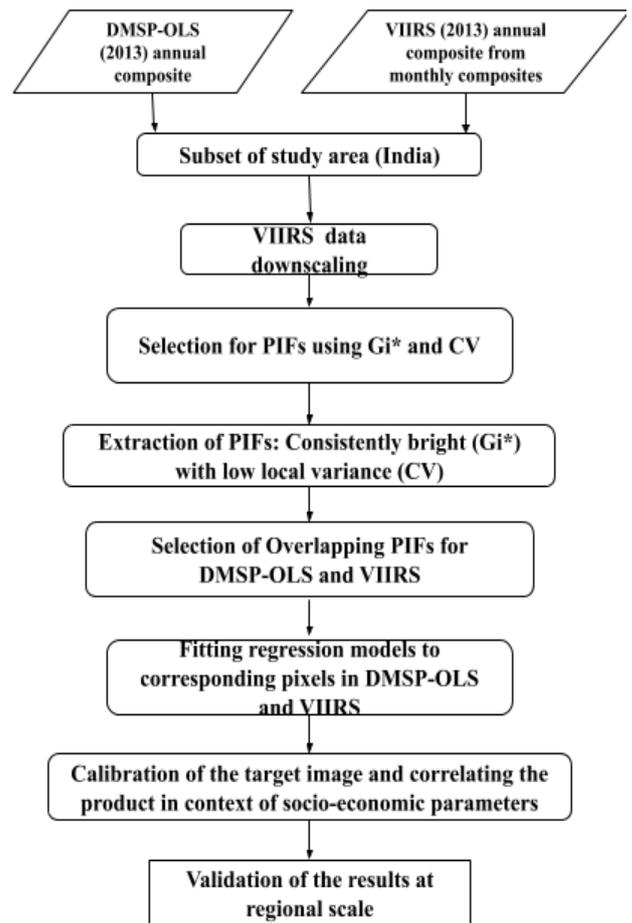


Figure 1: Flow chart representing the methodology adopted for this research

3.1 Masking for background removal

3.1.1 DMSP: Saturation correction

One most important issue in DMSP dataset is the saturation of the pixel's values in the bright cores of the urban areas. This leads to loss of information of the correct light intensities and loss of urban variations. This leads to incorrect analysis where the saturated regions are considered. This saturation takes place due to the limited dynamic range and 6-bit quantization of DMSP-OLS. This issue needs to be resolved for proper analysis of the variation in the luminosity. For this reason, values=63 are masked out.

3.1.2 VIIRS: Point based sampling

The unwanted data in the VIIRS dataset is due to the presence of negative radiances and the radiances due to the presence to the airglow effect in the uninhabited regions where the probability of illumination is zero or very low (Seaman et al., 2014).

Point-based Analysis is performed by the selection of uninhabited regions like rivers, forests and dark regions using Google map layer (Land use map). These regions are

considered to have a lower or zero probability of illumination according to the land use cover. The suitable value in these regions combined can be deducted from the overall values as a threshold. This is done by selection of the mean value and masking of all the values below it.

This technique is approached in two ways: By calculating the range and statistics and by manual inspection. Figure 2 shows the various uninhabited regions near the capital city of India, Delhi. These areas are selected for calculating the threshold.

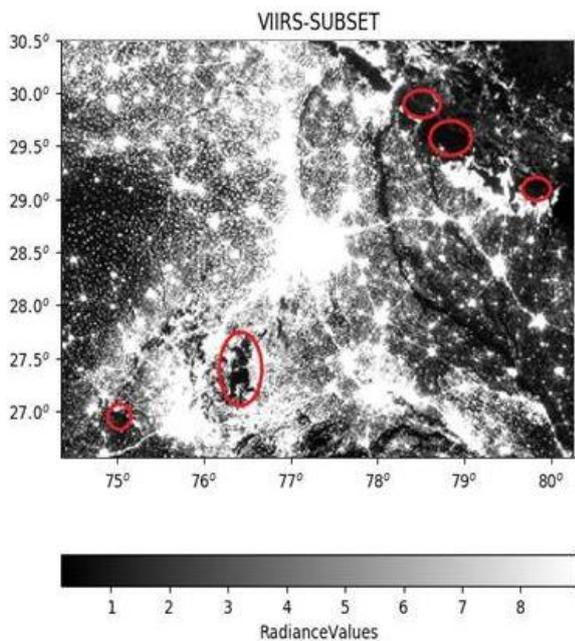


Figure 2: Representation of the subset area selected and uninhabited regions identified

3.2 Extraction of Pseudo Invariant Features (PIFs)

Extraction of Spatially homogeneous clusters and regions with low spatial variability gives PIFs. For identifying PIF, G_i^* statistics, and coefficient of variation are combined (Mukherjee, 2017). This selects relatively bright clusters with low local variability. In this case, the G_i^* is estimated and 90% and 95% significance levels are combined with the CV at five threshold levels of 5%, 10%, 15% and 20% in 3×3 windows and five threshold levels of 3%, 5%, 10%, 15% and 20% in 5×5 windows. The estimations of G_i^* and CV is carried out for DMSP and VIIRS individually. The next sections explain the G_i^* and CV in selection of PIFs.

3.2.1 Getis ord statistics (G_i^*)

G_i^* is a local indicator of spatial association (LISA). It is a measure of identifying significant clusters spatially of hotspots or cold spots. It is famous for hotspot analysis, where hotspots depict high-value clusters ($G_i^* > 0$) and cold spot depict low-value clusters ($G_i^* < 0$). G_i^* computation results in z-scores which indicate a place of a particular value relative to mean and standard deviation in a dataset. The value of z equal to zero ($z=0$) means the value of the statistic is equivalent to mean, value less than zero ($z < 0$) means the value is less than mean and the value greater than zero ($z > 0$) means value greater than the mean.

It considers the value of the points and their neighbouring pixels with a fixed distance d , where all the values within the distance d are considered for computation. It is calculated by comparing local averages to global averages by the analysis of spatial weights. The clusters are formed due to the concentration of weighted points defined in a local area.

G_i^* statistic is a local version of G_i statistic, the difference is that the G_i exclude the value of a particular point and only computes its neighbouring pixels with a certain distance d . G_i^* provides statistical significance to the hotspot analysis. The values of $G_i^* \geq 1.645$, $G_i^* \geq 1.960$, $G_i^* \geq 2.576$ and $G_i^* \geq 3.291$ in the computation denotes 90%, 95%, 99% and 99.9% of significance level respectively are calculated for DMSP and VIIRS.

3.2.2 Coefficient of variation

CV is a standard measure of variability, defined as:

$$cv = \frac{S}{\bar{x}}$$

Where,

S is the measure of standard deviation and of the radiances measurements, \bar{x} is the mean of radiance measurements in a predefined window.

CV is calculated for DMSP-OLS 2013 and VIIRS composite aggregated images on 3×3 and 5×5 windows. The calculations are done on the original images of the datasets of DMSP and VIIRS (composited and aggregated) clipped for India. Various thresholds of low variances are considered which are further used in attaining adequate PIFs. The thresholds of 3%, 5%, 10%, 15% and 20% are calculated.

3.3 Calculation of Net PIFs

After the individual PIF calculations of both the datasets, the net PIFs are calculated. The Net PIFs gives the overlapping PIF pixels free from the background noise of DMSP and VIIRS datasets. From this, the mask was already prepared in the previous steps. The mask generated was multiplied by the Total PIFs to calculate the Net PIFs after the background removal. These Net PIFs were used for extracting the values from original images of DMSP and VIIRS from inter-calibration.

The following steps are followed in the calculation of the PIFs:

- DMSP PIFs = G_i^* AND CV
- VIIRS PIFs = G_i^* AND CV
- Total PIFs = DMSP PIFs AND VIIRS PIFs
- Net PIFs = Total PIFs \times Total mask

Net PIFs generated using $G_i^* > 1.645$ and $CV < 20\%$ for India is shown in figure 3. Number of pixels calculated for DMSP PIFs, VIIRS PIFs, Total PIFs and Net PIFs are shown in table 1.

3.4 Inter-calibration model

This research adopts inter-calibration of VIIRS data with respect to DMSP. The Inter-calibration model is based on

the temporal overlap of DMSP and VIIRS images. For this, the common Net pseudo invariant features (Net PIFs) were considered and the corresponding values from DMSP and VIIRS are extracted. Logarithmic and Power model are applied for inter-calibration and the best suited one is picked for simulation.

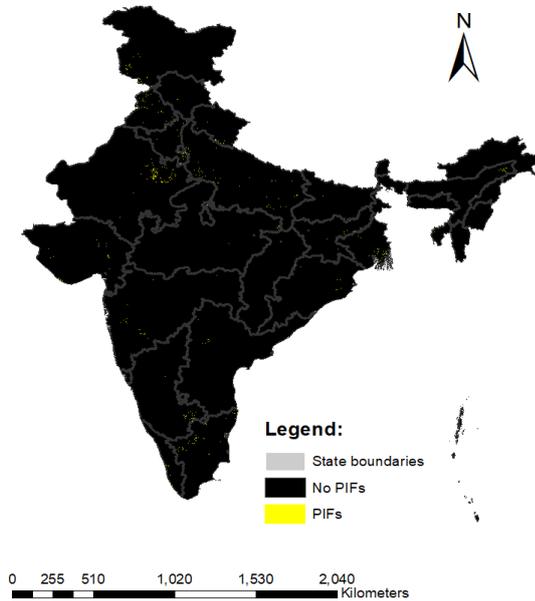


Figure 1: Net PIFs generated using $G_i^* > 1.645$ and $CV < 20\%$ for India

Using logarithmic function

The relationship between DMSP and VIIRS fits the logarithmic model:

$$y = a * \log \log (x) + b$$

Where, x denotes VIIRS value, y denotes the DMSP/OLS value, a and b are coefficients.

Inter-calibration equations:

$$y (dmsp) = a * \log \log (viirs) + b$$

Where y signifies the calculation of calibrated (simulated VIIRS) by altering the equation with coefficients.

Using power function

The power function is found to improve the comparability and to describe the non-linear relationship between the two dataset (Li et al., 2017).

$$y = ax^b$$

Where, x denotes VIIRS value, y denotes the DMSP/OLS value, a and b are coefficients.

Inter-calibration equations:

$$y (dmsp) = a(viirs)^b$$

Where y signifies the calculation of calibrated (simulated VIIRS) by altering the equation with coefficients.

3.5 Calculation of Sum of Lights (SOL) Index

The sum of light (SOL) index is the sum of all pixel values of the night-time light image defined by a particular area. SOL index checks the efficiency and consistency of the

inter-calibration process. It is calculated for different states of India for DMSP and simulated/calibrated VIIRS. This is done to evaluate the differences and errors in the inter-calibration. For this, as the shape file of administrative boundaries of India was considered. The SOL is calculated using Zonal statistics which calculated the sum of values state-wise. SOL for the simulated image is calculated after the equalization to the value range of DMSP.

Table 1: Number of pixels calculated for DMSP PIFs, VIIRS PIFs, Total PIFs and Net PIFs

VIIRS PIFs		
$G_i^* > 1.645$	$CV < 3\%$	No PIFs
(3X3 window)	$CV < 5\%$	34
	$CV < 10\%$	1061
	$CV < 15\%$	5935
	$CV < 20\%$	16597
$G_i^* > 1.645$	$CV < 3\%$	No PIFs
(5X5 window)	$CV < 5\%$	No PIFs
	$CV < 10\%$	24
	$CV < 20\%$	951
DMSP PIFs		
$G_i^* > 1.645$	$CV < 3\%$	96264
(3X3 window)	$CV < 5\%$	160656
	$CV < 10\%$	815609
	$CV < 15\%$	1151764
	$CV < 20\%$	1327312
$G_i^* > 1.645$	$CV < 3\%$	13749
(5X5 window)	$CV < 5\%$	30815
	$CV < 10\%$	340570
	$CV < 15\%$	686836
Total PIFs		
DMSP AND VIIRS	$CV < 5\%$	22
(3X3 window)	$CV < 10\%$	786
	$CV < 15\%$	4747
	$CV < 20\%$	14131
DMSP AND VIIRS	$CV < 10\%$	20
(5X5 window)	$CV < 15\%$	814
Net PIFs		
Total PIFs X Total Mask	$CV < 10\%$	330
(3X3 window)	$CV < 15\%$	2942
	$CV < 20\%$	10626

3.6 Correlation with socio economic indicators

Two famous indicators are considered for testing the inter-calibrated model at national and regional scales which are: Gross State Domestic Product (GSDP) and Electricity Consumption (EC) for the year 2013.

GSDP in India accounts for the sum of the values added by all the industries for the states and union territories. State-wise Gross Domestic Product for India at the current prices on 2011-2012 series was considered for the financial year of 2013-2014. The GSDP is measured as

Indian Rupees (in crore i.e., 10^7). State-wise total energy consumption by ultimate consumers in India was considered for the final year of 2013-2014. The Electricity consumption state-wise has a unit of Gigawatt hours (GWh).

4. Results and discussion

4.1 Extraction of PIFs

PIFs are spatially homogeneous and of low local spatial variability (Refer to section 3.6). The two regions of spatial homogeneous clusters and with low local variances are integrated (using Boolean AND) between outputs of various thresholds applied. Figure 3 shows the Net PIFs calculated using the value of $G_i^* > 1.645$ and $CV < 20\%$ which gives 10,626 pixels of concern. These are the major cities of India having more illumination as compared to the other regions. Table 1 lists the calculations made for G_i^* and CV individually for DMSP, VIIRS and the ANDing result for the Net calculation.

4.2 Relationship between DMSP-OLS and SNPP-VIIRS

4.2.1 Logarithmic relationship

- At 15% CV threshold: the acquired points from the Net PIFs (2942) are extracted from DMSP and VIIRS and the data is applied to the Model: $a \cdot \log(x) + b$. The Model fits with the value of $R^2 = 0.805$.
- The Coefficients of the model fit are found to be $a = 36.957$ and $b = 12.185$.
- At 20% CV threshold: the acquired points from the Net PIFs (10626) are extracted from DMSP and VIIRS and the data is applied to the Model: $a \cdot \log(x) + b$. The Model fits with the value of $R^2 = 0.775$. The coefficients of the model fit are found to be $a = 38.669$ and $b = 12.071$. Figure 4 illustrates the fitting of the points for Log Model.

4.2.2 Power relationship

- At 15% CV threshold: the acquired points from the Net PIFs (2942) are extracted from DMSP and VIIRS and the data is applied to the model: $a \cdot x^b$. The Model fits with the value of $R^2 = 0.734$. The
- Coefficients of the model fit are found to be $a = 14.758$ and $b = 0.448$.
- At 20% CV threshold: the acquired points from the Net PIFs (10626) are extracted from DMSP and VIIRS and the data is applied to the model: $a \cdot x^b$.

The model fits with the value of $R^2 = 0.702$. The coefficients of the model fit are found to be $a = 14.904$ and $b = 0.471$. Figure 5 illustrates the fitting of the points for Power Model.

4.3 Simulation of VIIRS dataset

The simulation or calibration of VIIRS datasets (predicted DMSP) is performed by selecting the optimal model. This is done based on the value of R-square. The suitable model developed with the adequate number of PIFs is found to be the logarithmic fit at 20% threshold of CV. The related coefficients are extracted and original VIIRS image is calibrated using:

$$\text{Predicted DMSP} = 38.67 * \log(\text{VIIRS}) + 12.07$$

The simulated image is then equalized to DMSP by clipping the greater values above 63 and removing the negative values.

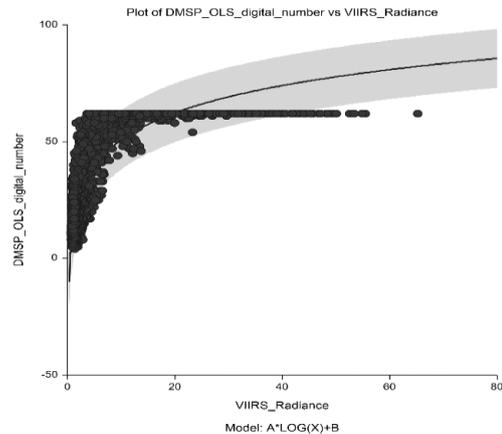


Figure 2: Logarithmic model fit line and the distribution of the points of extracted PIFs from DMSP and VIIRS at 20% CV threshold

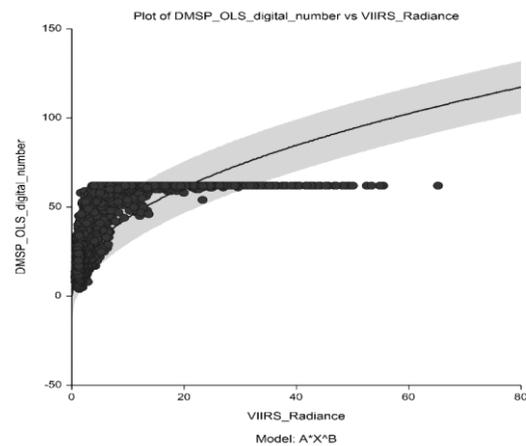


Figure 3: Power model fit line and the distribution of the points of extracted PIFs from DMSP and VIIRS at 20% CV threshold

Figure 6 show the normalized and equalized image of VIIRS in the range of DMSP dataset. This simulation is further used to correlate with socio economic indicators for the year 2013.

The simultaneous comparison between DMSP and VIIRS has been made to know about the accuracy of the simulation. The attempt in this research is been made for a single year 2013, but the idea applies to many years and can be used in a series analysis. This can be further researched and used for different datasets in future.

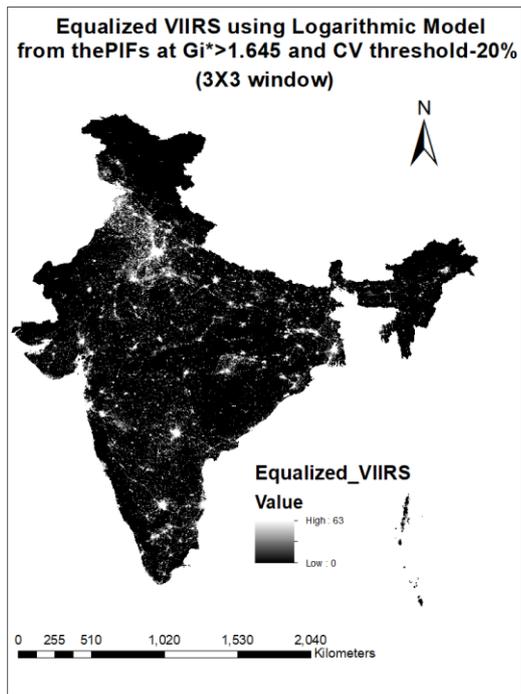


Figure 4: Equalized simulated VIIRS (Predicted DMSP) image from the Log model

4.4 Validation with socio-economic Indicators

4.4.1 National level

Gross state domestic product

DMSP

The equation of the straight line relating SOL and state-wise gross domestic product for the year 2013 of 29 Indian states has been evaluated. The value of R-Squared, the proportion of the variation in SOL that can be accounted for by variation in State wise GSDP for 2013, is found to be 0.716. The correlation between SOL and State wise GSDP 2013 is found to be 0.846.

Simulated VIIRS (predicted DMSP)

The equation of the straight line relating SOL and state-wise gross domestic product for the year 2013 of 29 Indian states has been evaluated. The value of R-Squared, the proportion of the variation in SOL that can be accounted for by variation in State wise GSDP for 2013, is found to be 0.606. The correlation between SOL and State wise GSDP 2013 is found to be 0.778.

Electricity Consumption

DMSP

The equation of the straight line relating SOL and State-wise electricity consumption for the year 2013 of 30 Indian states has been evaluated. The value of R-Squared, the proportion of the variation in SOL that can be accounted for by variation in State wise electricity consumption for 2013, is found to be 0.867. The correlation between SOL and State wise electricity consumption is found to be 0.931.

Simulated VIIRS (Predicted DMSP)

The equation of the straight line between simulated SOL and State-wise electricity consumption for the year 2013

of 30 Indian states has been evaluated. The value of R-Squared, the proportion of the variation in SOL that can be accounted for by variation in State wise electricity consumption for 2013, is found to be 0.690. The correlation between simulated SOL and State wise electricity consumption 2013 is found to be 0.831.

4.4.2 Regional level

Five regions or zones considered are: North, North East, East, West and south. The SOL index is calculated zone wise and validated against GSDP and EC for DMSP and VIIRS individually.

Gross state domestic product

Gross state domestic product is correlated with SOL index of DMSP and SNPP-VIIRS. The coefficient of determination comes out to be 0.905 for DMSP but comes out to be significantly low for simulated VIIRS, 0.529. The relationship seems stronger with the original DMSP than that of simulated VIIRS (Predicted DMSP) equalized to the level of DMSP.

Electricity consumption

Electricity consumption by ultimate consumers is correlated with SOL index of DMSP and SNPP-VIIRS. The coefficient of determination comes out to be 0.936 for DMSP but comes out to be significantly low for simulated VIIRS, 0.539. The relationship seems stronger with the original DMSP than that of simulated VIIRS (Predicted DMSP) equalized to the level of DMSP.

Figure 7 and 8, shows the variability in the relationship of DMSP and VIIRS with GSDP and EC of the year 2013 zone wise. Also, from the state wise variations in the values of sum of light (SOL) for DMSP-OLS and SNPP-VIIRS datasets. It is observed that the model underestimates the values of simulated VIIRS almost in all states except the states of Jammu & Kashmir, Jharkhand, and Meghalaya.

4.5 Discussion

The study innovates at calculating the common invariant regions for the former and the newer source of NTL data and developing an inter-calibration model. The research focused on harmonization and inter-calibration of DMSP and VIIRS for a consistent dataset the study makes use of the resources available for public use and finding out the ways in which the objectives are achieved satisfactorily. The inter-calibration process helps to radiometrically harmonize both the datasets. The process is accomplished by the use pseudo invariant features as the target sites for modelling the relationship between the two datasets. For the calculation of the PIFs, methods of G_i^* and CV are applied. The PIFs are used for intercalibration and simulation of VIIRS in harmony with DMSP. The logarithmic function was fitted to VIIRS with function coefficients and inter-calibrated/simulated VIIRS (Predicted DMSP) image was drawn. The simulated image range was observed from -165.95 to 138.247. This range was made saturated from values greater than 63 and negative values were removed. The range was then equalized to DMSP by converting the radiances to the

integer. This step has radiometrically harmonized VIIRS with the level of DMSP.

covering northern, north eastern, eastern, western and southern regions of India.

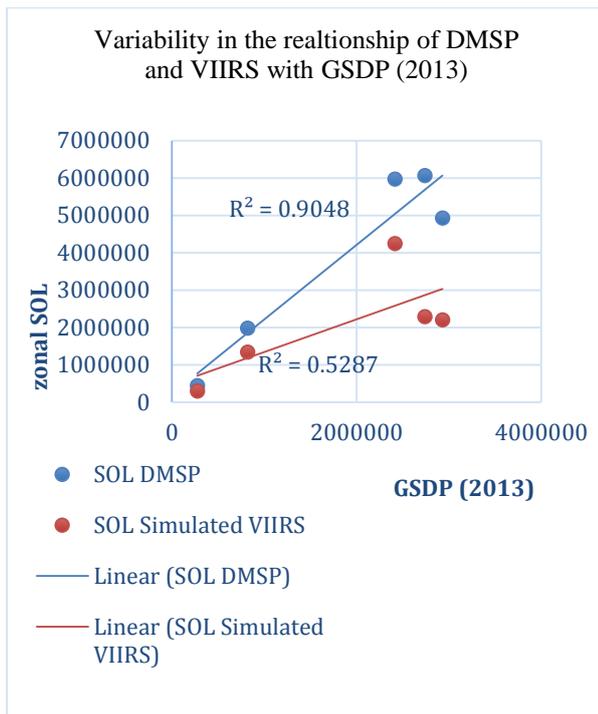


Figure 5: Relationship of zonal SOL against GSDP 2013 (Rs. In crore)

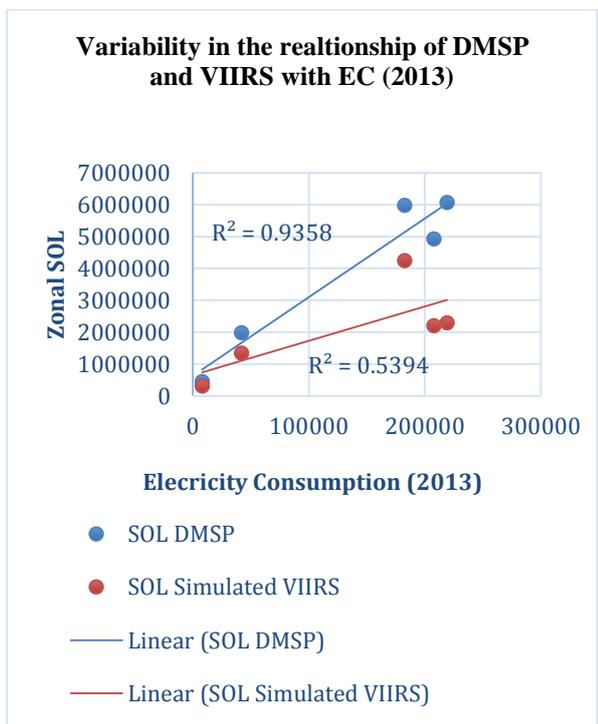


Figure 6: Relationship of zonal SOL against EC 2013 (GWh)

After simulation, SOL was calculated for simulated VIIRS (Predicted DMSP) and original DMSP which are assessed against Gross state domestic product (GSDP) and Electricity consumption (EC) for the year 2013-2014. The indicators were collected state-wise and the correlation was found at the national level and on regional levels

4.6 Limitations of the approach

The inter-calibrated model generated using the methodology adopted cannot be directly applied to the consequent years after 2013 for making a consistent time series of DMSP. This is because the DMSP data are not radiometrically correct. The DMSP data needs pre-processing for the correction of saturation problem and blooming effect. However, piece wise fitting of regression line can produce a better fit of the inter-calibration model. The approach takes monthly images and then the annual composite is constructed using averaging techniques. This doesn't include the sensor's parameters and the consideration of lightning due to other sources. This affects the performance of the inter-calibrated model.

5. Conclusions

For the importance and necessity for consistent night-time light datasets, harmonization and inter-calibration of the two datasets are performed. For harmonization, the foremost requirement to ensure the equality on the basis of temporal, spatial, spectral and radiometric harmonization. Here, the VIIRS dataset is altered to make it consistent to DMSP. This requirement is fulfilled by monthly compositing for temporal harmonization spatial aggregation for spatial harmonization and inter-calibration for radiometric harmonization. The two datasets are already spectrally harmonized, collecting the night-time lights in one specific band.

For inter-calibration, pseudo invariant features were used as target sites denoting the regions where the light does not drastically vary and are spatially homogeneous and stable. The PIFs calculated are made free from the saturated pixels in DMSP and background noise in VIIRS are used for simulation successfully.

The validation was performed on National level covering the states of India and at regional levels covering the zones and their corresponding states. This research finds out that the inter-calibrated Model performs well on the national scale but doesn't perform well at regional scales. Improvements can be made to ensure better consistency by applying different harmonization techniques, selection of PIFs and inter-calibration model selection.

For instance, ancillary products such as Normalised Urban Areas Composite Index (NUACI) (Liu et al., 2015) data can be used to remove non-urban areas which have DN values in DMSP/NTL data. This can help in improving the inter-calibration process.

Moreover, the blooming effect can be attempted to remove using optimal thresholding technique as described by Liu et al. (2015). Inter-calibration within all DMSP/NTL datasets is also necessary since it involves various satellites with different radiometry. This process may automatically bring 2013 dataset also in uniformity.

References

- Cao, C. (2013). NOAA Technical Report NESDIS 142A Visible Infrared Imaging Radiometer Suite (VIIRS) Sensor Data Record (SDR) User's Guide, (September).
- Doll, C. N. H. (2008). CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications, pp. 1–41.
- Elvidge, C. D., K. Baugh, M. Zhizhin and F.C. Hsu (2017). VIIRS night-time lights, *International Journal of Remote Sensing*, 38(21), 5860–5879. <https://doi.org/10.1080/01431161.2017.1342050>
- Ghosh, T., R.L. Powell, C.D. Elvidge, K.E. Baugh, P.C. Sutton and S. Anderson, S. (2010). Shedding light on the global distribution of economic activity, *The Open Geography Journal*, 3, 148–161. <https://doi.org/10.2174/1874923201003010147>
- Henderson, B. J. V., A. Storeygard and D.N. Weil (2012). Measuring economic growth from outer space, *American Economic Review* 2012, 102(2), 994–1028, <http://dx.doi=10.1257/aer.102.2.994>
- Huang, Q., X. Yang, B. Gao, Y. Yang, Y. Zhao (2014). Application of DMSP/OLS Night time Light Images: A Meta-Analysis and a Systematic Literature Review, 6844–6866. <https://doi.org/10.3390/rs6086844>
- Liu, X., G. Hu and B. Ai (2015). A Normalized Urban Areas Composite Index (NUACI) based on combination of DMSP-OLS and MODIS for mapping impervious surface area, *Remote Sensing*, 7(12), 17168–17189. <https://doi.org/10.3390/rs71215863>
- Li, X., D. Li, H. Xu and C. Wu (2017). Inter calibration between DMSP/OLS and VIIRS night-time light images to evaluate city light dynamics of Syria's major human settlement during Syrian Civil War, *International Journal of Remote Sensing*, 38(21), 5934–5951. <https://doi.org/10.1080/01431161.2017.1331476>
- Mellander, C., J. Lobo, K. Stolarick and Z. Matheson (2015). Night-time light data: A good proxy measure for economic activity? *PLoS ONE*, 10(10), 1–18. <https://doi.org/10.1371/journal.pone.0139779>
- Ministry of Home Affairs. (2011). Census of India Website: Office of the Registrar General & Census Commissioner, India. Retrieved May 9, 2019, from <http://censusindia.gov.in/>
- Mukherjee, S. (2017). Quality analysis of inter-calibration of DMSP- OLS night-time images. University of Twente Faculty of Geo Information and Earth Observation (ITC).
- Pandey, B., Q. Zhang, K.C. Seto (2017). Comparative evaluation of relative calibration methods for DMSP/OLS night time lights, *Remote Sensing of Environment*, 195, 67–78. <https://doi.org/10.1016/j.rse.2017.04.011>
- Seaman, C. (2013). Beginner's Guide to VIIRS Imagery Data VIIRS Intro. Curtis Seaman CIRA/Colorado State University, 10/29/2013.
- Seaman, C., D. Hillger, T. Kopp, R. Williams, S. Miller and D. Lindsey (2014). NOAA Technical Report Visible Infrared Imaging Radiometer Suite (VIIRS) Imagery Environmental Data Record (EDR) User ' s Guide NOAA Technical Report Visible Infrared Imaging Radiometer Suite (VIIRS) Imagery Environmental Data Record (EDR) User ' s, (April).

Techniques developed for large area Mars image mosaic using ISRO's Mars Color Camera (MCC) data

Indranil Misra*, S. Manthira Moorthi and Debajyoti Dhar

Optical Data Processing Division, SIPG, Space Applications Centre (ISRO), Jodhpur Tekra, Ahmedabad – 380015, India

*Email: indranil@sac.isro.gov.in

(Received: Jan 02, 2019; in final form: May 13, 2019)

Abstract: Multispectral remote sensing planetary datasets covering same and adjacent locations at different acquisition times allow for image mosaicking to study local, regional and large-scale processes of planetary bodies. ISRO's Mars Orbiter Mission launched on November 5th 2013 increases our curiosity to examine planet Mars using scientific instruments on-board MOM. Mars Color Camera (MCC) is among the five science payloads which operates in visible range (0.4 μm to 0.7 μm) to image the surface feature of Mars with varying resolutions and scales from highly elliptical orbit. The systematic processing procedure for mosaicking MCC images primarily requires radiometrically corrected MCC data with areographic co-ordinates tagged in planetary data system (PDS) standard as input and involves techniques of image processing domain that includes geometric and photometric corrections, image registration, blending and normalization. These techniques are described in this paper to generate mosaic product of Mars terrain using different MCC datasets at variety of scales to illustrate compositional diversity, morphological changes to view the features of Mars and geological issues using different perspectives. In addition, the techniques are also used to generate planet level seamless mosaic, North Pole mosaic using available MCC datasets to monitor dynamic behavior of dust devils/storms and cloud conditions over larger area of Mars.

Key words: Image Mosaic, Mars Color Camera, Photometric Correction, Image Registration, PDS

1. Introduction

Mars Color Camera (MCC) images obtained from Mars Orbiter Mission (MOM) are gaining scientific popularity since Mars Orbiter Mission (MOM) insertion into an elliptical orbit around Mars on 24th Sep, 2014 (Moorthi et al., 2015; Arya et al., 2015). MOM is still active beyond four years in orbit, the coverages from MCC are sufficient to begin constructing mosaic products on a regional and global scale. MCC is a RGB Bayer camera which operates in visible range (0.4 to 0.7 μm). The detector array of MCC has 2048x2048 elements on a pixel pitch of 5.5 μm . MCC imaging is influenced by the special characteristics of MOM's orbit, which is highly elliptical with a periapsis of 261 km and an apoapsis of 78000 km. As a consequence, the ground resolution of MCC images varies from 15m to 4 km that can map a specific crater with more feature details to the full disc images covering large area of Mars in coarser resolution (MCC Team. 2013).

Multiple MCC frames taken at different time can stitch together at areoreferenced space (equivalent to georeferencing of Earth observation data) to generate any mosaic product. The construction and use of MCC mosaic datasets can increase the present understanding of Mars in many aspects that include geo-morphological studies, change detection, dust devil/clouds at different locations, crater counting, active processes, climate and geological history. The systematic data processing procedure to create large scale mosaics is illustrated in the subsequent sections of this paper and present some unique MCC mosaics that can cover the longest canyon system, a large shield volcano and multiple craters together.

2. Data processing steps

MCC data processing system is an automation intensive software system that processes instrument data for edited

and calibrated products, derives meta data about mission events, spacecraft operations, instrument operations, processing parameters, orbit and housekeeping details from ancillary data to generate data products following PDS standards (Moorthi et al., 2015). The MOM data Processing system is ingested with decompressed, formatted, time tagged MCC raw data along with orbit and attitude information.

2.1 Radiometric processing

The first step involves radiometric correction that includes photo response non-uniformity correction (PRNU) using pre-launch calibration data sets and then demosaic all colors at every pixel locations from Bayer filter mosaic. The aim of a demosaicing algorithm is to reconstruct a full color image from the spatially under sampled color channels output from the color filter array (CFA) (Roy et al., 2014).

2.2 Geometric processing

The next step is to compute areographic coordinates of Mars for every pixel using the orbit and attitude information provided along with the data. Areographic coordinates should be precise enough to locate DEM height, demands an extra procedure enough to adjust the geometric accuracy of the mapping using a Mars image mosaic of known accuracy. Mars Digital Image Model (MDIM) 2.1 is a widely used Mars reference image mosaic for asserting location accuracies by Mars science communities (Som et al., 2008). The MDIM 2.1 is a 231 meters/pixel global Viking/Mariner mosaic of Mars having improved geodetic accuracy from its earlier version. Geometrically, MDIM 2.1 is an orthoimage dataset, draped on the Mars Orbiter Laser Altimeter (MOLA)-derived radius model (Archinal et al., 2004). MCC images are co-registered with respect to MDIM 2.1 making it geometrically confirming to MDIM and MOLA data sets. The image registration automatically detect and

match features, estimate transformation between MCC and MDIM 2.1 to correct the input MCC images. Registered MCC images ensure geometric features continuity between adjacent scenes which is pre-requisite for large area mosaic (Misra et al., 2012).

2.3 SPICE computations

The MCC data processing also uses SPICE kernels to compute photometric angles, solar angles and sensor view angles at every MCC pixel location. The SPICE kernels are structured parameter files that describe, among other things, the ephemeris and attitude of the spacecraft, the spatial orientation of the instrument, the ephemeris and physical parameters of the Earth, Mars and other natural bodies in the Solar System, the relationship between spacecraft time and time as measured on Earth, and the locations and orientations of ground stations. SPICE kernel files are intended for use with the SPICE software library, which contains functions that perform computations for space mission-related and astronomical applications (Action et al., 1998).

2.4 Topographic correction

MCC images are also corrected for topographic effects to normalize the radiance measures before considering the data for science analysis (Smith et al., 1999). It is proposed here to use non Lambertian Minnaert semi empirical approach for correcting MCC image photometry further used for deriving results for Mars surface science. The methodology outlined here uses terrain parameters such as slope and aspect values derived from MOLA DEM to compute local illumination angle. Topographically corrected images were evaluated for the improvement in its radiometry quantitatively and it is found to reduce topographic shading and improve the overall image quality (Misra et al., 2015).

3. Mosaic data selection and processing

Image mosaicking is initiated with the selection of high quality and well calibrated data. During the MOM mission, images were both systematically and priority targeted, potentially covering the same area on the planet many times over. Prior to mosaicking, visual examination of

every image is often included as an additional step. Data that are often excluded include images containing elevated line-to line or white noise and excessive repeated dropouts, which makes blending of several images difficult, and images collected during periods of high atmospheric dust condition where the overall contrast of the image is reduced. Datasets are selected based on highest visual quality and other lower quality data is removed, reducing the probability of blending poorly registered data. The sequence of steps performed for mosaic processing is described in the below sub-sections.

3.1 Contrast stretching

Mosaic processing is ingested with MCC reduced data records (RDR) available in PDS standard, which is radiometrically, geometrically and topographically corrected data derived from MCC experimental data record (EDR). In order to display the data for viewing, linear contrast stretch must be applied to scale the intensity of the data and maximize the dynamic range. Linear contrast stretching for normalization attempts to improve an image by stretching the range of intensity values it contains to make full use of possible values. Contrast stretching is restricted to a linear mapping of input to output values. Linear contrast stretch is used in this work for stretching that maximizes the spectral variation while retaining most of the morphological context. The Figure-1 shows original MCC image with no stretch and MCC image with linear contrast stretch for better understanding. In the blending process, MCC datasets are intelligently interwoven by incorporating information from overlapping pixels using two dimensional linear ramp. In the first step, overlap region is extracted between adjacent MCC images using polygon intersection in areographic space. According to the overlap region statistics between the images, normalized weighting factors are computed which is applied to the subsequent image to generate seamless MCC mosaic. The blending algorithm does not blend all MCC images at a time rather a pair is blended into an output and the final mosaic generated by taking the average intensities of overlapping images (Edwards et al., 2011).

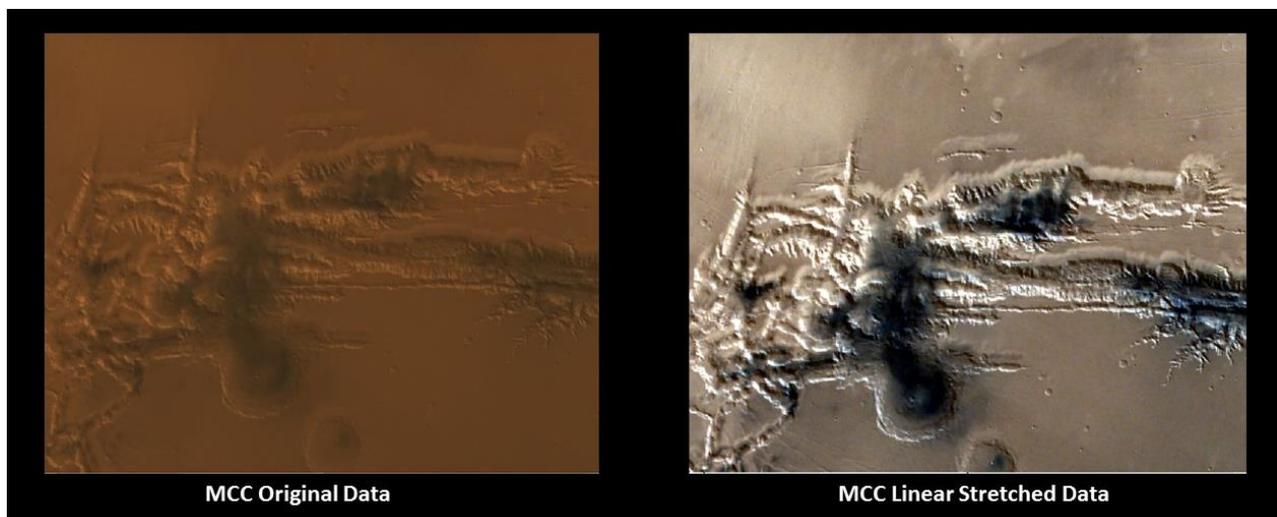


Figure 1: MCC contrast stretch for data viewing (Region: part of Valles Marineris)

3.2 Image to image blending

To summarize, topographic correction is applied before linear stretch and then multi-dimensional blend is applied to be inserted into the final mosaic area.

4. MCC mosaic processing workflow

The processing workflow is divided into three major processing stages. In the first stage of processing MCC datasets are radiometrically corrected. In geometric processing stage, MCC data is tagged with areographic coordinates and photometric angles are derived from SPICE kernels. Automatic image registration with MDIM 2.1 standard reference is executed to improve the absolute location accuracy of the MCC product which is a prerequisite for image mosaicking. Topographically corrected MCC data sets can better resolved different features than its uncorrected counterpart. The Figure-2 (Mangala Valles) clearly show a flow pattern with channel bar in the MCC image indicating release of vast quantities of water in this area by catastrophic floods. Topographic correction better demarked the features for scientific

analysis. The images shown in figure 1 and figure 2 are linearly stretched for better visual analysis.

In the final stage of mosaic procedure images are picked automatically for mosaic construction, contrast stretched and blended as depicted in figure 3.

5. Results

The MCC mosaic product at regional scale and global scale aid better understanding of Mars. At regional scale, Valles Marineris Grand Canyon system mosaic is generated. The Valles Marineris is a large tectonic crack present on the Martian crust running up-to a length of around 4000km, 200 km wide and 7 km deep. Figure 4 shows MCC multi-scene images (from Nov, 2014 acquisitions) mosaic portraying complete Valles Marineris feature produced by applying geometric and radiometric corrections. As proposed, different canyons in the Valles Marineris are seen in MCC image such as Ophir Chasma, Hydras Chasma and Melas Chasma

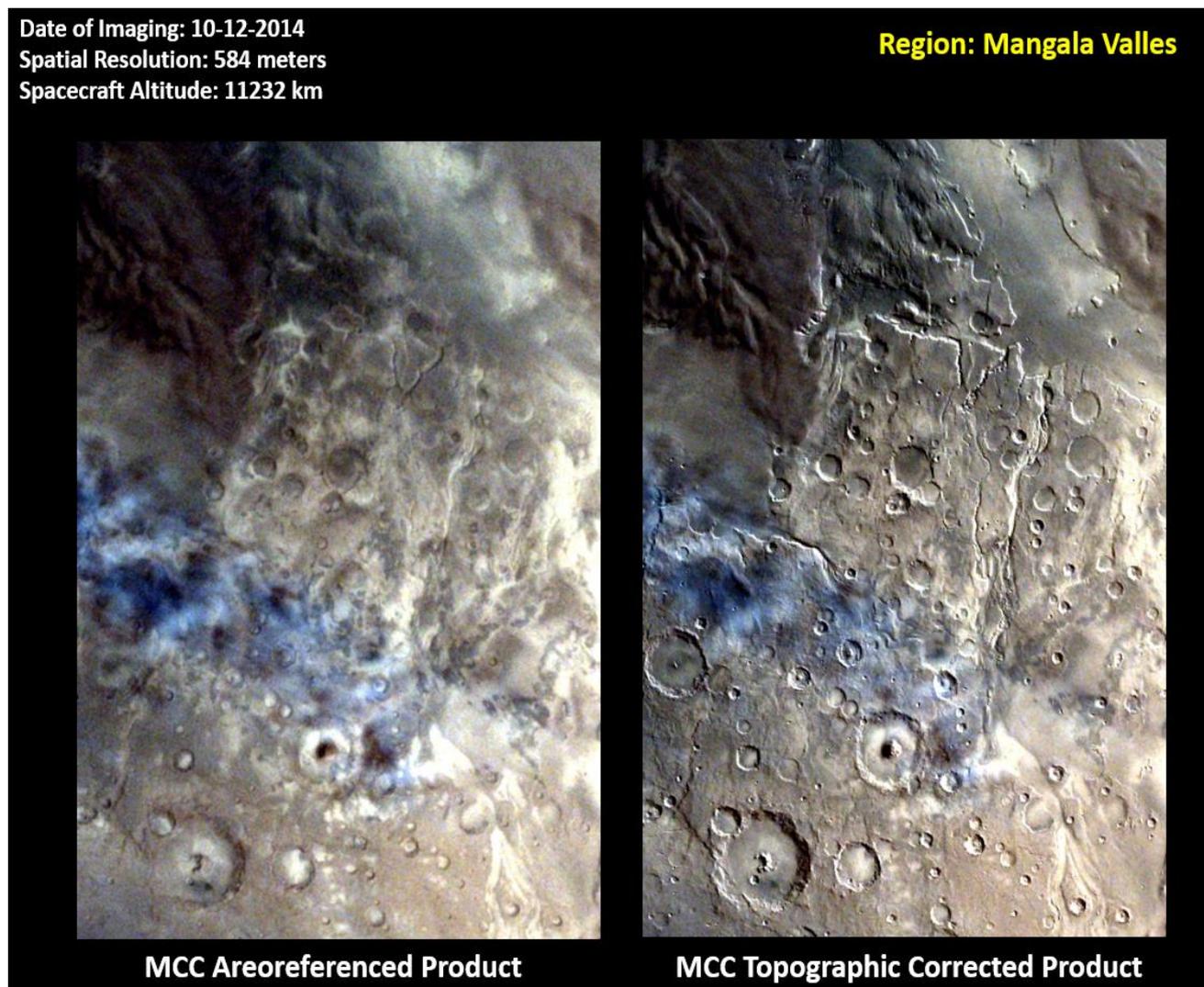


Figure 2: Comparison of MCC original Areoreferenced image (Left) and MCC topographic corrected image (Right) (Mars Region: Mangala Valles)

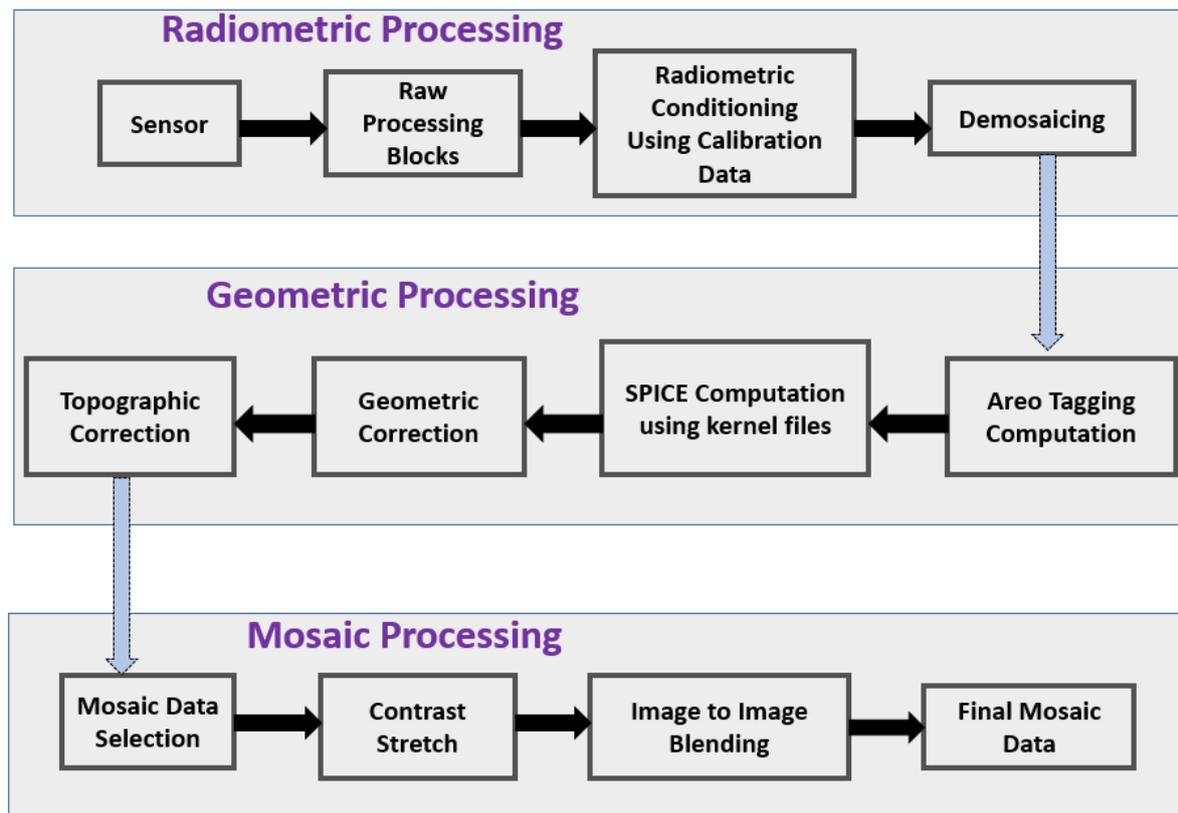


Figure 3: MCC mosaic processing workflow

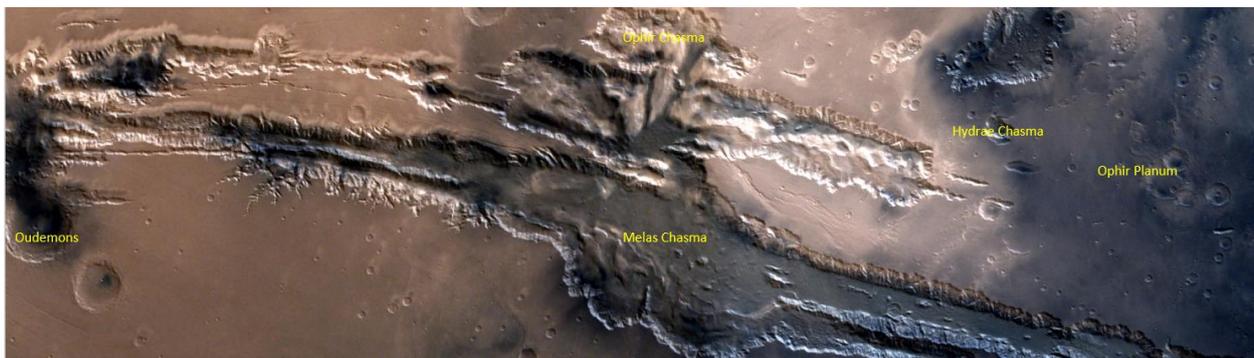


Figure 4: Valles Marineris mosaic using MCC datasets

At global scale, MCC full disc images are used which are actually obtained in perspective view. These images have been rectified using geometric correction steps including a map projection step. Each one of the full disc images covers partial portion of the Mars disc but not complete. While mosaicking the images, relative geometric differences were removed by additional image registration procedure. A seamless Mars full disc canvas was prepared shown in figure 5, adjusting the colour differences between images. Eight full disc images obtained during Dec2015 and Jan 2016 MCC were used. The Pixel Resolution was uniformly scaled to 4 km pixels.

Figure 5 also shows MCC 3D view by warping global mosaic over Mars globe. The global mosaic in figure 5 still left with unbalanced colors. Further improvements can be bought by fine-tuning the present procedure.

Conclusion

MCC mosaic processing system can generate visually appealing mosaics using multiple MCC images acquired at different times covering large area of Mars. The systematic processing of MCC datasets to generate different views of Mars can be used for geological change detection studies. The construction of the mosaics using the techniques described in this paper, provide the ability to view the surface of Mars and geologic problems through many different perspectives. The current processing system only generate mosaic product for visual interpretation and geo-morphological analysis. The future work may use advance techniques for colour processing of Mars images.

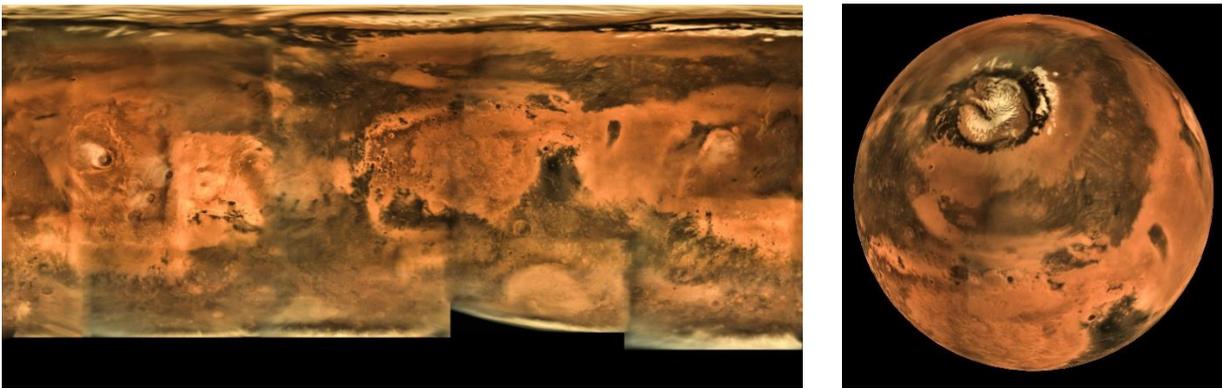


Figure 5: MCC global mosaic using full disc images and MCC 3D view generated using mosaic

Acknowledgements

The authors acknowledge Shri D K Das, Director, Space Applications Centre (SAC), ISRO for his encouragement in our research activities. The authors also thank other team members of optical data processing for their support during the course of this work.

References

- Action, C, N. Bachman, E. Lee, S. Boris and W. Edward (1998). SPICE: A real example of data system re-use to reduce the costs of ground data systems development and mission operations, <https://descanso.jpl.nasa.gov/RCSGSO/paper/A0068Paper.pdf>.
- Archinal, B.A, E.M. Lee, R.L. Kirk, T.C. Duxbury, R.M. Sucharski, D.A. Cook and J.M. Barrett (2004). A new Mars Digital Image Model (MDIM 2.1) Control Network, International Archives of Photogrammetry and Remote Sensing.
- Arya, A.S., R.P. Rajasekhar, Rimjhim B Singh, Koyel Sur, Prakash Chauhan, S.S. Sarkar, S. Manthira Moorthi, A. R. Srinivas, Vishnu Patel, Sampa Roy, Indranil Misra, Rajdeep Kaur Gambhir, Kamlesh K Patel, D.R.M. Samudraiah and A.S. Kiran Kumar (2015). Mars Color Camera onboard Mars Orbiter Mission: Initial observations & results, 46th Lunar and Planetary Science Conference.
- Edwards, C.S., K. J. Nowicki, P. R. Christensen, J. Hill, N. Gorelick and K. Murray (2011). Mosaicking of global planetary image datasets: 1. Techniques and data processing for Thermal Emission Imaging System (THEMIS) multi-spectral data, *Journal of Geophysical Research*, 116, E10008, doi:10.1029/2010JE003755.
- MCC Team (2013). Pre-shipment Review document on Mars Colour Camera (MCC), SAC/MOM.
- Misra, I., S.M. Moorthi and D. Dhar (2015). Minnaert Topographic Correction of Mars Color Camera images from Mars Orbiter Mission, IEEE Conference ICIT, Bhubaneswar, India.
- Misra, I., S.M. Moorthi, D. Dhar and R. Ramakrishnan (2012). An automatic satellite image registration technique based on Harris corner detection and Random Sample Consensus (RANSAC) outlier rejection model; IEEE Recent Advances in information technology, ISM Dhanbad, India.
- Moorthi, S.M., A.S. Arya, Kurien Mathew, R.P. Singh, Prakash Chauhan, S.S. Sarkar, Sampa Roy, Arvind K Singh, Meenakshi Sarkar, Ajay Kumar Prashar, K. Suresh, Indranil Misra, Rajdeep Kaur Gambhir, Arpita Gajaria, Ashutosh Gupta, Tanul Gupta, D. Dhar, T.P. Srinivasan, K. Padia, R. Ramakrishnan, B. Gopalakrishna, S. Chowdhury and A.S. Kiran Kumar (2015). Mars Orbiter Mission: Science data products and archive pipeline, 46th Lunar and Planetary Science Conference.
- Roy, S., D. Dhar, S.M. Moorthi and S.S. Sarkar (2014). Comparative analysis of demosaicking techniques for Mars Colour Camera Data, IEEE International Conference on Contemporary Computing and Informatics.
- Smith E David, Maria T. Zuber, Sean C. Solomon, Roger J. Phillips, James W. Head, James B. Garvin, W. Bruce Banerdt, Duane O. Muhleman, Gordon H. Pettengill, Gregory A. Neumann, Frank G. Lemoine, James B. Abshire, Oded Aharonson, C. David Brown, Steven A. Hauck, Anton B. Ivanov, Patrick J. McGovern, H. Jay Zwally and Thomas C. Duxbury (1999). The Global Topography of Mars and implications for surface evolution, *Research Article*, www.sciencemag.org, Science, 284.
- Som, M.S., M.G. Harvey and R.M. David (2008). The Mars Orbiter Laser Altimeter dataset: limitations and improvements, *The International Journal of Mars Science and Exploration*, Mars, 4, 14-26.

Stress mapping of Sundarban mangroves with Sentinel-2 images using Discriminant Normalised Vegetation Index (DNVI) and Fuzzy classification technique

Sudip Manna* and Barun Raychaudhuri

Department of Physics, Presidency University, Kolkata-700073, West Bengal, India

*Email: sudipmarine@gmail.com

(Received: Jan 02, 2019; in final form: May 14, 2019)

Abstract: The present work reports stress mapping of Sundarban mangroves implementing fuzzy classification technique to Sentinel-2 data. A recently developed health index for mangroves, namely Discriminant Normalised Vegetation Index (DNVI) was used as a tool for extracting the signatures of stressed and healthy vegetation from Sentinel-2 image along with field survey data. Fuzzy classification of stress and health conditions allowed the pixels to acquire partial membership of different classes. The partitioning of the classes was resolved by convolving the fuzzy classes based on DNVI and normalised difference vegetation index (NDVI). A saturating tendency of NDVI was noted when compared to DNVI and consequently the later was assigned the first layer for determining fuzzy convolution weightage. The precise result in the form of micro-level stress map indicates that the stress is a probable function of local geomorphology, topography and physiography. This method appropriately represents the fuzzy pattern of natural forest cover rather than that obtained with fixed algorithm based hard classification methods. The approach also highlights the need of mapping the stress of different assemblages discretely instead of a single health index. Similar index value for different mangroves may not represent similar health conditions for all of them because of their different physiology.

Key words: Stress mapping, DNVI, Mangroves, Sentinel-2, Fuzzy classification.

1. Introduction

Paradigm changes in the global climate result in increasing global temperature, changing precipitation pattern, sea level rise, prolonged droughts, heat waves and continued intensification of storm events (Hansen et al., 1988; Knutson et al., 2010; Allen et al., 2010; Trenberth, 2011; Hoffman et al., 2017; Nerem et al., 2018; Mal et al., 2018). Mitigation of such a state and its further abatement requires assessment of the basic elements of the ecosystems. Forests, in addition to water bodies, atmospheric dynamics and human activities function as such elements and are the most important terrestrial ecosystems globally affected by anthropogenic activities (Práválie, 2018).

Mangrove forests are categorically evergreen sturdy vegetation thriving in intertidal regions of tropics and sub-tropics. They support coastal communities by means of invaluable ecosystem services (Himes-Cornell, 2018). Almost 40% of world's population is living within 150 km of coastlines (Cohen et al., 1997) and the consequent pressure has incurred a massive loss to the mangrove area (Polidoro et al., 2010). Most of the residents of tropical and subtropical coasts greatly depend on mangrove ecosystems for their livelihood and sustenance, either directly or indirectly. Apart from that, these forests provide a protective buffer to the coasts against natural calamities like cyclones and storm surges. Presently these forests are spatially dwindling with '*cryptic ecological degradation*' (Dahdouh-Guebas et al., 2005). Monitoring the mangroves has always been a challenge owing to their complex structure, muddy substrate and tough accessibility. Moreover, these vast ecosystems are sensitive to disturbances and may take over a decade to restore (Smith et al., 1994). In order to analyse such ecosystems and to address their present-day status, researchers use remote sensing as a synoptic tool.

Remote sensing of mangroves requires substantial inputs due to differential resolutions of the satellite data. Literature survey illustrates the use of vegetation indices developed at various times for the interpretation of biophysical parameters using optical remote sensing (Kuenzer et al., 2011) in mangroves. Almost all of the established vegetation indices use spectral response at red and near infrared wavelengths (Bannari et al., 1995; Broge and Leblanc, 2001; Adam et al., 2010) as they are relatable to the leaf pigments and canopy structure. The knowledge of the stress on mangroves induced by geomorphology and other associated parameters is essential for understanding the forest dynamics. To date there has been little work on mangrove stress mapping using remote sensing.

Chellamani et al. (2014) used NDVI for assessment of health status of mangroves in India from SPOT-VGT sensor and categorised mangroves into poor, moderate, health and very healthy. However, it was not stated as to how NDVI can directly portray the health, except for referring to previous studies where NDVI was reported (Tucker, 1979) to be sensitive to green leaf biomass. The health of mangroves cannot be depicted solely based on their canopy density or chlorophyll concentration because different mangrove species have their respective compositional construct and morphology. Domination of a single species covering top-canopies and association with other species standing as under-canopies results in the admixture of canopy reflectance and makes the mapping of the absolute composition challenging. However, active remote sensing is capable of interpreting subsurface organisation of flora (Kuenzer et al., 2011). In the case of analysis of spaceborne optical image, the combination of different spectra in a moderate resolution image is perplexed with background responses raising the need of fine resolution data. The phenology of the mangroves indicates their behaviour like evergreen forests as they continuously shed the senescent leaves to be replaced by the young ones.

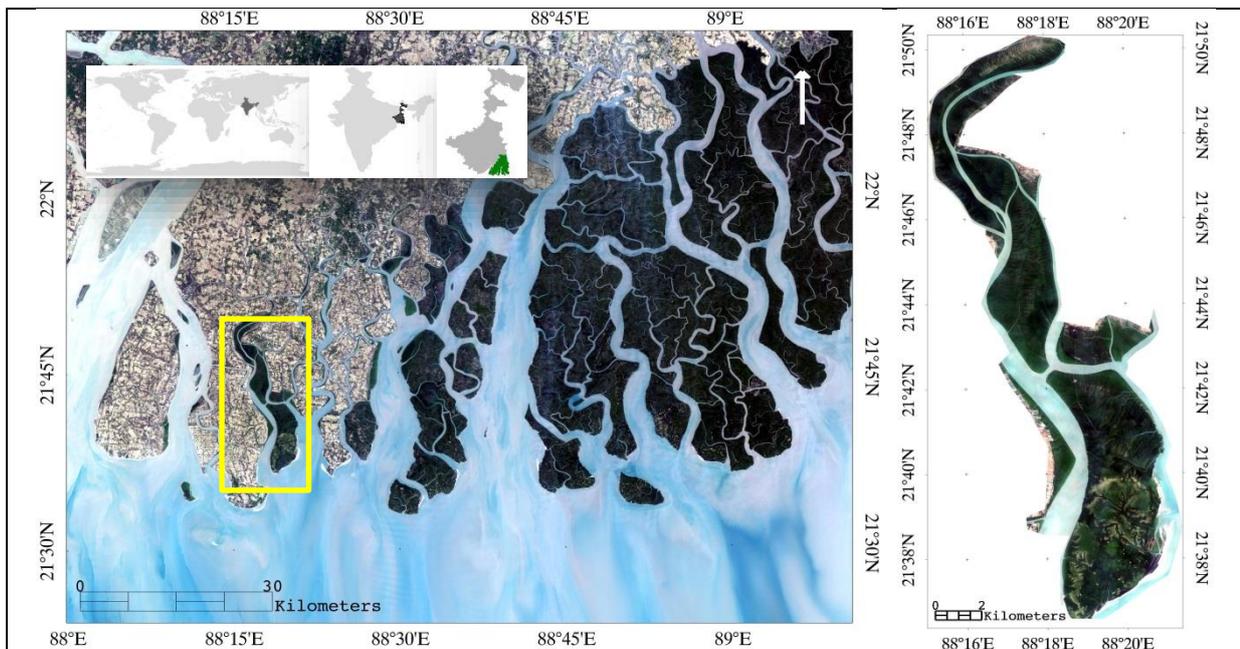


Figure 1: Map showing Indian part of Sundarban (left) and Lothian island with adjoining islands of the study area. Inset maps not to scale.

Stressed mangroves shed leaves resulting into gradual reduction of canopy cover. Various environmental stressors could also lead to reduction of stem density and overall basal area. The dwarfing is another effect leading to morphological changes including branching pattern, wood density and canopy height. Apart from that, mangroves located much inland and receiving infrequent tidal inundation also suffer from stresses (Saenger 2002). Therefore, health mapping of mangroves must account for various parameters like leaf density, chlorophyll concentration, water content, cellulose-lignin content, and nutrients; and most importantly the species or cohort specific assessment owing to difference in dissimilar and congeneric species. All these parameters are comprehensively not relatable to indices solely based on red and near infrared reflectance, hence require the integration of SWIR responses. The stress indicating index used for mangrove assessment, termed as DNVI was developed recently with Sentinel-2 data (Manna and Raychaudhuri 2018). Under circumstances like hypersaline substratum and inadequate influx of tidal and fresh water combined with sediment deposition, the mangroves stands gradually transform into saline blanks with population of dwarf and stunted individuals. Such regions also illustrate some rank higher than absolute zero on NDVI scale. DNVI on the other hand being derived from SWIR's is sensitive to changes in the mangroves structural units (Kuenzer et al., 2011, Zhang et al., 2014) can portray the condition much competently.

Mangrove species assemblages have specific affinity and spatial preferences from the perspective of distance from water and elevation from mean sea level. These stands propagate spatially with diverse species composition and inconstant ecological parameters which induces uneven growth among same species individuals and vice versa. Different species having dissimilar health condition may appear similar entities when observed by remote sensing and therefore, gives rise to the confusion or fuzziness in

their mapping. Here we present an approach for delineating species assemblage specific stress mapping based on DNVI developed from SWIR bands using fuzzy classification method. The objectives of the research were to (i) map the spatial distribution of mangroves based on their stress or health conditions and (ii) to delineate the mangrove species assemblage specific stress condition using fuzzy classification employing ground observation, DNVI and NDVI.

2. Methods

2.1 Study area

Sundarban is the largest deltaic contiguous mangrove patch in the world. The Indian part of the forest is subdivided into a core zone of 1700 km², manipulation zone of 2400 km², restoration zone of 230 km², and a development zone of 5300 km² (Nandy and Kuswaha, 2011). As reported by IUCN, it is a habitat to a wide range of flora; 334 plant species belonging to 245 genera and 75 families, 165 algae and 13 orchid species. The meso-macrotidal estuary gets inundated and exposed twice a day by diurnal tides having amplitude of 2.5–7 m. The soil texture is chiefly clayey-loam whereas certain parts also have sandy-loam and silty soils. The present study was conducted in a wildlife sanctuary (Figure 1) in the Indian part of Sundarban and although situated very close to human settlements, it is having almost no influence from anthropogenic activities.

2.2 Satellite image processing

Sentinel-2 cloud free image archived by European Space Agency was used for the study and downloaded from the Sentinel data hub (<https://scihub.copernicus.eu/>). The reflectance image is comprised of 13 bands having different spatial resolutions (Table 1). All the bands were resampled using ESA-Sentinel Application Platform (SNAP) freeware. We used 10 m and 20 m resolution

bands for the classification purpose and the 60 m bands were used to derive the DNVI.

Table 1: Spatial and spectral resolution of Sentinel S2A MSI

Bands	Spatial resolution (m)	Central wavelength (nm)	Bandwidth (nm)
B1	60	443.9	27
B2	10	496.6	98
B3	10	560.0	45
B4	10	664.5	38
B5	20	703.9	19
B6	20	740.2	18
B7	20	782.5	28
B8	10	835.1	145
B8a	20	864.8	33
B9	60	945.0	26
B10	60	1373.5	75
B11	20	1613.7	143
B12	20	2202.4	242

It has been already depicted by Zheng et al. (2017) that for Sentinel 2 data, downscaling pixel size of a coarse resolution band performs better than upscaling with respect to classification accuracy or delineation capacity. Additionally, the upscaling could deprive the entire dataset from spatial information contained in the fine spatial resolution bands. DNVI developed from 60 m bands (resampled to 10 m for compatibility with NDVI) is used only as a thematic parameter for guiding the fuzzy convolution. The raster data was processed, analysed and interpreted with the help of ENVI software.

2.3 In-situ sampling

Field inventory for the health condition of different species and their assemblages were conducted during March 2016. Eighty ground locations with details of species, physiological conditions like health, slenderness coefficients (Vovides et al., 2014) of trees and substratum conditions were recorded using GPS 72 (Garmin Ltd.). The soil salinity, slenderness coefficient (ratio of total height to girth), canopy density- leaf area index were recorded for the assessment of the health of mangroves. Leaf architecture and condition (thickened and or short leaves, crumbled or flat lamina), tree structure (straight trunked or gnarled) and distance of trees from the intertidal zone which mutually indicate the health condition of mangroves were also recorded. Apart from that, as the weaker trees are more susceptible to insect infestations, the locations of such canopies were also recorded for the purpose. Random sampling was done in order to ensure the complete coverage of the study area and its floral compositions. In addition, the pure and mixed species assemblages were located precisely for their mapping.

2.4 Species mapping

Based on the field inventory and using Support vector machine (SVM) algorithm in ENVI (Exelisvis Inc., USA), the Sundarban mangrove area under consideration was delineated into different mangrove species, assemblage types and non-mangrove landcover including mudflats. Tree class abbreviations are elucidated in table 2. SVM is a non-parametric supervised learning model that exercise

user defined signatures and has been used for precise mangrove mapping recently (Heumann 2011; Manna and Raychaudhuri, 2018).

Table 2: Abbreviations used for mapping classes

	Abbreviation	Species details
1	AA_m	<i>Avicennia alba</i> matured assemblages
2	AA_y	<i>Avicennia alba</i> young assemblages
3	AM	<i>Avicennia marina</i>
4	AM_sc	<i>Avicennia marina</i> scrub assemblages
5	AO	<i>Avicennia officinalis</i>
6	AR	<i>Agialites rotundifolia</i>
7	EA	<i>Excoecaria agallocha</i>
8	Mixed	Mixed tall assemblages
9	Mixed_sc	Mixed scrub assemblages
10	PP	<i>Phoenix paludosa</i>

2.5 Fuzzy classification

Sundarban mangroves having heterogeneous species composition provide an opportunity to map this fuzzy variation with respect to individual species using fuzzy classification technique where there is no precise threshold between two similar yet different targets. Such classification is a type of *soft classification* with certain degree of uncertainty of the classified image. The targets are classified with multiple membership values considering the probability of them belonging to any class, which is the actual scenario of a diversely populated forest. The ambiguity is put in order by fuzzy clustering where the information from neighbouring pixels helps determining the actual parent category.

The classification is based on fuzzy logic comprising the following steps. The reflectance image containing thirteen bands, each of 10 m spatial resolution was used to collect the signatures specific to a species or assemblage. The signature collection was executed by selecting thirty areas of interest (AOI) of different features and of different pixel counts based on ground survey, DNVI and species map. Based on the collected signatures, fuzzy classification was carried out considering two best classes per pixel. Using this classified image as the input along with the distance file of the two bands of DNVI and NDVI, fuzzy convolution was executed, which created a single classification layer of the total weighted inverse distance of all the classes in a window of pixels thereby creating a context based classification. The general expression for the total weighted distance of window for class k is given by

$$T(k) = \sum_{i=0}^s \sum_{j=0}^s \sum_{l=0}^s \frac{w_{ij}}{D_{ijl}(k)}$$

where i and j are row and column index of window respectively, of the window of size s (3×3 used here) and class value k , w_{ij} being the weight table for the window. $D(k)$ represents the distance file value for each window element for class k . The block diagram shown in figure 2 represents the methodology in a schematic flowchart.

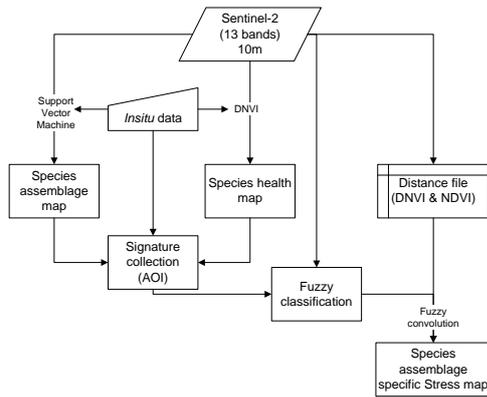


Figure 2: Schematic representation of the methodology used for species assemblage specific stress of mangroves

3. Results and discussion

The fuzzy convolution technique assigns the centre pixel of the running window within the class of the largest total inverse distance summed over the entire fuzzy classified bands. Classes with small distance values remain unchanged whereas those with large values may change to a neighbouring value, if sufficient number of neighbouring pixels with class value exists. In the present case (Figure 3a), 3×3 window size was used (Figure 3b). A larger window size (7×7) might lead to over-generalization (Figure 3c).

The health condition of mangroves is an apparently variable parameter, as their stand comprises different, congeneric species. Conventionally, the most popular vegetation index NDVI is used as indicator to several biophysical parameters including fractional vegetation cover estimates, leaf area index, vigour and even biomass (Curran et al., 1992, Jiang et al., 2006, Manna et al., 2014). However, several studies have depicted the saturating tendency of NDVI in predicting the biophysical properties of trees, especially of canopies which are mostly a voluminous and multi-strata entity. In mapping the overall health of a mangrove forest using remote sensing, an index like DNVI (Manna and Raychaudhuri, 2018) is capable of spatially portraying the stressed and healthy assemblages of mangroves. A comparative analysis (Figure 4) of NDVI with health index DNVI has depicted a linear relation with a saturating tendency at higher NDVI values.

It is concluded from the saturating tendency of NDVI (Figure 4) that the discrimination of mangroves health condition by DNVI is more efficient than that using NDVI as a proxy. The selective difference in their capabilities might be due to the spectral ranges of reflectance used to derive the indices. NDVI is derived using *red* and *near infrared* bands indicative of leaf pigments and cell structure, whereas DNVI is developed using *shortwave infrared* bands responsive to structural properties associated with stress conditions, such as water content, leaf biochemicals, protein, lignin and cellulose (Kuenzer et al., 2011, Zhang et al., 2014). A general stress map generated using DNVI indicates the distribution of saline blanks and stressed vegetation precisely (Figure 5a). The different regions in figure 5a are indicated by similar

colours but all the mangroves under the same category do not belong to same species (Figure 5a, b).



Figure 3: A part of the study area depicting a) the true colour composite from Sentinel-2 data, b) fuzzy convolution result using 3×3 window and, c) fuzzy convolution result using 7×7 window

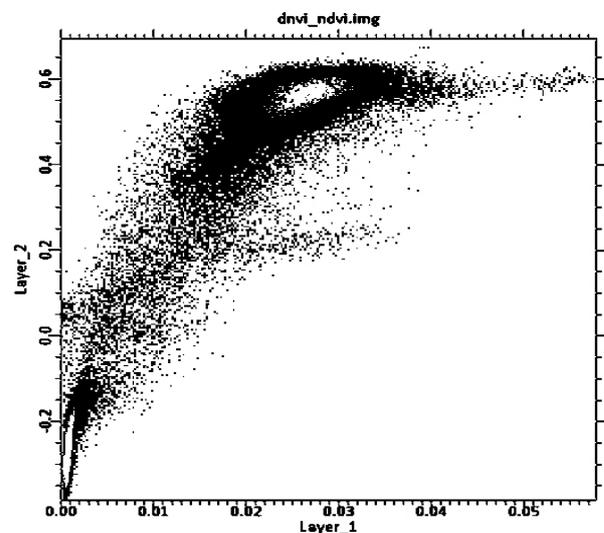


Figure 4: Scatter plot of DNVI (x axis) versus NDVI (y axis) depicting the saturating tendency of NDVI

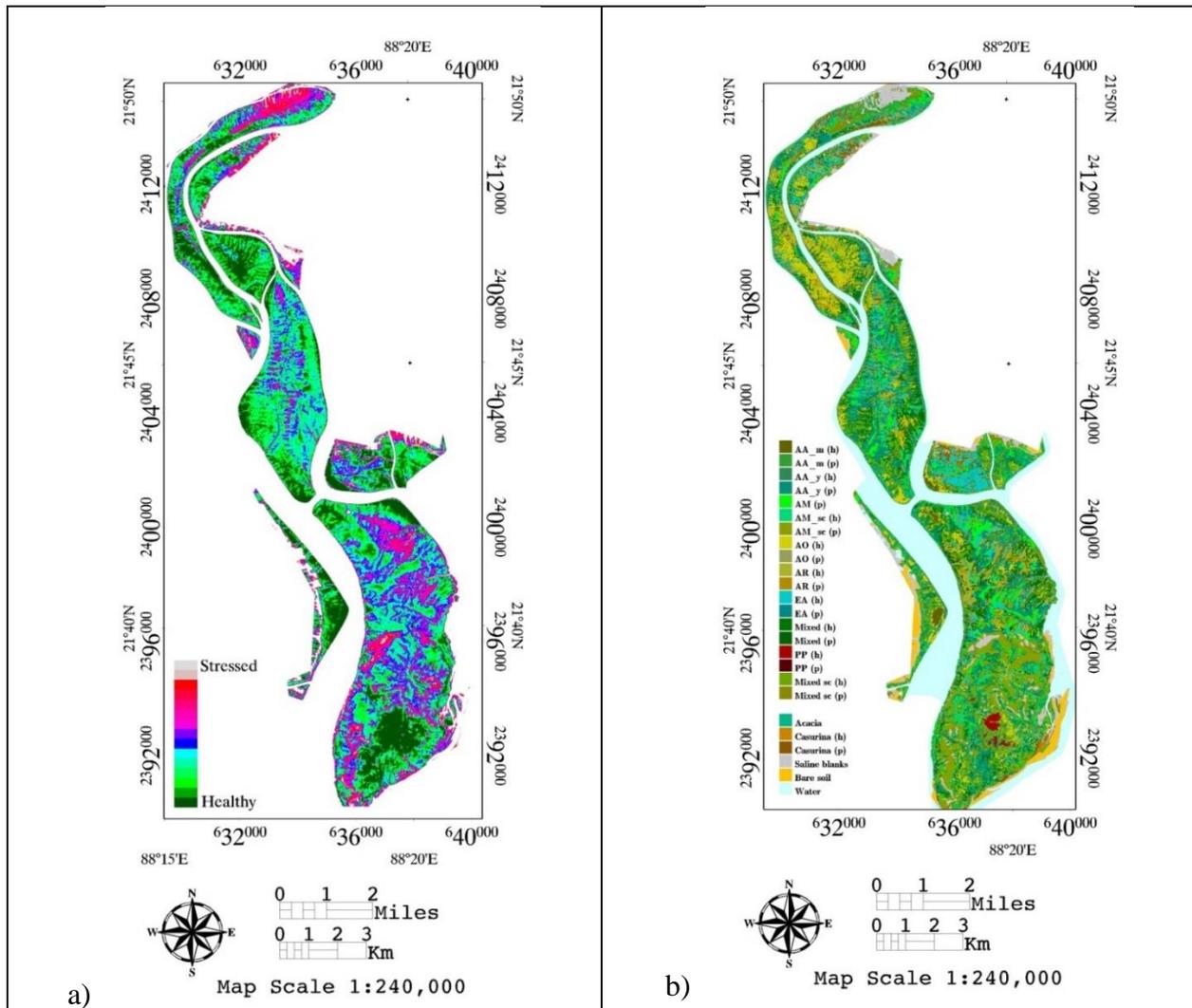


Figure 5: Maps of Lothian wildlife sanctuary, Sundarban showing a) DNVI based generalised stress map and b) Species assemblage specific stress map* derived using fuzzy classification. (* the legend details are provided in Table 2; h and p in parentheses indicate health and poor conditions respectively).

A similar index value for two different species does not indicate similar physiological conditions for both the candidates. Such variation in health condition might be due to their respective physiology, growing capability in particular eco-region and differential association with other species. The geographical variation also plays a crucial role in the growth and health conditions of mangroves. For instance, a species namely *Ceriops decandra* is reported to robustly grow up to a height of ≈ 5 m in the Bangladesh part of Sundarban (Hossain et al. 2012), whereas this species is found to grow mostly as shrub, bushy and stunted in the Indian counterpart. A possible reason is the increasing salinity stress in Indian dominion due to gradual obstructions in paleo-channels bringing fresh water supply to the delta (Gopal and Chauhan, 2006). Moreover, in order to conserve and manage a forest stand with several species, the micro-level stress assessment is much more essential than representing a synoptic health status.

4. Conclusion

DNVI, a health index developed from Sentinel-2 high resolution free data and validated with airborne

hyperspectral satellite data was utilized for mapping the stress of mangroves in Sundarban. The representative species assemblage specific stress map was generated using the fuzzy principle where the ground data, established species map, and DNVI and NDVI were used as parameters for training the parametric classifier. While NDVI has always been used as an indicator to various biophysical parameters of vegetation, this study revealed that in mapping the health it saturates disabling the fine distinction among assemblages of good health.

The implication of fuzzy classification and use of DNVI and NDVI for fuzzy convolution revealed the fuzzy nature of mangroves in its spatial distribution. The vertically multi-stratum canopied forest was mapped better using fuzzy classification than the fixed algorithm based hard classification methods. Given the observations and findings from the study, the approach highlighted the need of mapping the stress of different assemblages discretely that could not be portrayed by single stress image. The necessity of this approach is also supported by the fact that same index value for different mangroves does not indicate an equal health or stress. Furthermore, precise mapping of mangroves stress could be realized by plausible fusion of

high resolution non-commercial data like Sentinel-2 with compatible microwave data for more realistic assessment of mangrove forests having complex morphology and community structure. Precise micro-mapping in the cases like the above requires a superior classification technique rather than a hard classification method, especially in the case of variable forest cover where the confusion probability is higher in multispecies natural system like Sundarban. Therefore, the stress mapping at species and assemblage level was done using fuzzy classification that resulted in fine scale health-stress map of the mangroves in study area.

Acknowledgements

The authors would like to thank the financial support of DST NPDF grant (PDF/2017/002858) from Department of Science and Technology (DST), Science and Engineering Research Board (SERB), Government of India, and the infrastructural support of Presidency University, 86/1 College Street, Kolkata 700073. The authors extend their thanks to the European Space Agency for availing data through 'The Copernicus Open Access Hub' and to the West Bengal Forest Department for permitting and assisting in the field inventories.

References

Adam, E., O. Mutanga and D. Rugege (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review, *Wetlands Ecology and Management*, 18(3), 281-296.

Allen, C. D., A. K. Macalady, H. Chenchouni, D. Bachelet, N. McDowell, M. Vennetier, T. Kitzberger, A. Rigling, D. D. Breshears, E. T. Hogg and P. Gonzalez (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests, *Forest ecology and management*, 259(4), 660-684.

Bannari, A., D. Morin, F. Bonn and A. R. Huete (1995). A review of vegetation indices, *Remote sensing reviews*, 13(1-2), 95-120.

Broge, N.H. and E. Leblanc (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density, *Remote sensing of Environment*, 76(2), 156-172.

Chellamani, P., C.P. Singh and S. Panigrahy (2014). Assessment of the health status of Indian mangrove ecosystems using multi temporal remote sensing data, *Tropical Ecology*, 55(2), 245-253.

Cohen, J.E., C. Small, A. Mellinger, J. Gallup, and J. Sachs (1997). Estimates of coastal populations, *Science*, 278(5341), 1209-1213.

Curran, P.J., J.L. Dungan, H.L. Gholz (1992). Seasonal LAI in slash pine estimated with landsat TM, *Remote Sensing of Environment*, 39, 3-13.

Dahdouh-Guebas, F., L.P. Jayatissa, D. Di Nitto, J.O. Bosire, D.L. Seen, and N. Koedam (2005). How effective were mangroves as a defence against the recent tsunami?, *Current biology*, 15(12), R443-R447.

Gopal, B. and M. Chauhan (2006). Biodiversity and its conservation in the Sundarban Mangrove Ecosystem. *Aquatic Sciences*, 68(3), 338-354.

Hansen, J., I. Fung, A. Lacis, D. Rind, S. Lebedeff, R. Ruedy, G. Russell and P. Stone (1988). Global climate changes as forecast by Goddard Institute for Space Studies three-dimensional model, *Journal of Geophysical Research: Atmospheres*, 93(D8), 9341-9364.

Himes-Cornell, A., L. Pendleton, and P. Atiyah (2018). Valuing ecosystem services from blue forests: A systematic review of the valuation of salt marshes, sea grass beds and mangrove forests, *Ecosystem Services*, 30, 36-48.

Hoffman, J.S., P.U. Clark, A.C. Parnell and F. He (2017). Regional and global sea-surface temperatures during the last interglaciation, *Science*, 355(6322), 276-279.

Hossain, M., M.R.H. Siddique, A. Bose, S.H. Limon, M.R.K. Chowdhury and S. Saha (2012). Allometry, above-ground biomass and nutrient distribution in *Ceriops decandra* (Griffith) Ding Hou dominated forest types of the Sundarbans mangrove forest, Bangladesh, *Wetlands Ecology and Management*, 20(6), 539-548.

Jiang, Z., A.R. Huete, J. Chen, Y. Chen, J. Li, G. Yan and X. Zhang (2006). Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction, *Remote sensing of environment*, 101(3), 366-378.

Knutson, T.R., J.L. McBride, J. Chan, K. Emanuel, G. Holland, C. Landsea, I. Held, J.P. Kossin, A.K. Srivastava and M. Sugi (2010). Tropical cyclones and climate change, *Nature geoscience*, 3(3), 157.

Kuenzer, C., A. Bluemel, S. Gebhardt, T.V. Quoc and S. Dech, (2011). Remote sensing of mangrove ecosystems: A review, *Remote Sensing*, 3(5), 878-928.

Mal, S., R.B. Singh, C. Huggel and A. Grover (2018). Introducing linkages between climate change, extreme events, and disaster risk reduction. In *Climate Change, Extreme Events and Disaster Risk Reduction* (pp. 1-14). Springer, Cham.

Manna, S., S. Nandy, A. Chanda, A. Akhand, S. Hazra and V.K. Dadhwal (2014). Estimating aboveground biomass in *Avicennia marina* plantation in Indian Sundarbans using high-resolution satellite data, *Journal of Applied Remote Sensing*, 8(1), 083638.

Manna, S. and B. Raychaudhuri, B., (2018). Mapping distribution of Sundarban mangroves using Sentinel-2 data and new spectral metric for detecting their health condition, *Geocarto International*, 1-19.

- Nandy, S., S. Kushwaha, (2011). Study on the utility of IRS 1D LISS-III data and the classification techniques for mapping of Sunderban mangroves, *J. Coast. Conserv.*, 15, 123–137.
- Nerem, R.S., B.D. Beckley, J.T. Fasullo, B.D. Hamlington, D. Masters and G.T. Mitchum (2018). Climate-change–driven accelerated sea-level rise detected in the altimeter era, In Proceedings of the National Academy of Sciences, p.201717312.
- Polidoro, B.A., K.E. Carpenter, L. Collins, N.C. Duke, A.M. Ellison, J.C. Ellison, E.J. Farnsworth, E.S. Fernando, K. Kathiresan, N.E. Koedam and S.R. Livingstone (2010). The loss of species: mangrove extinction risk and geographic areas of global concern, *PloS one*, 5(4), 10095.
- Práválie, R. (2018). Major perturbations in the Earth's forest ecosystems. Possible implications for global warming, *Earth-Science Reviews*. 185, 544-571.
- Saenger, P. (2002). *Mangrove ecology, silviculture and conservation*. Springer Science & Business Media.
- Smith, T.J., M.B. Robblee, H.R. Wanless and T.W. Doyle (1994). Mangroves, hurricanes, and lightning strikes, *BioScience*, 44(4), 256-262.
- Tucker, C.J., (1979). Red and photographic infrared linear combinations for monitoring vegetation, *Remote sensing of Environment*, 8(2), 127-150.
- Trenberth, K.E., (2011). Changes in precipitation with climate change. *Climate Research*, 47(1-2), 123-138.
- Vovides, A.G., J. Vogt, A. Koller, U. Berger, U. Grueters, R. Peters, A.L. Lara-Domínguez and J. López-Portillo (2014). Morphological plasticity in mangrove trees: salinity-related changes in the allometry of *Avicennia germinans*, *Trees*, 28(5),1413-1425.
- Zhang, C., J. Kovacs, Y. Liu, F. Flores-Verdugo, and F. Flores-de-Santiago, (2014). Separating mangrove species and conditions using laboratory hyperspectral data: A case study of a degraded mangrove forest of the Mexican Pacific, *Remote Sensing*, 6(12), 11673-11688.
- Zheng, H., P. Du, J. Chen, J. Xia, E. Li, Z. Xu, X. Li, and N. Yokoya (2017). Performance evaluation of downscaling Sentinel-2 imagery for land use and land cover classification by spectral-spatial features, *Remote Sensing*, 9(12),1274.

Subsurface utility mapping using multi frequency ground penetrating radar: A case study of road collapse

Tathagata Chakraborty*, Dharmendra Kumar Pandey, Jayaprasad P., Deepak Putrevu and Arundhati Misra
Space Applications Centre, Ahmedabad – 380015, India
*Email: tathagata@sac.isro.gov.in

(Received: Dec 20, 2019; in final form: May 17, 2019)

Abstract: In Urban areas, non-invasive or non-destructive detection and mapping of various subsurface utilities is a major challenge. However, this is extremely important to detect any defects or damage in the pipelines or cables which can cause hurdles in various aspects. Traditional methods such as digging soil and roads without prior information about exact locations are destructive, time-consuming and labour intensive. Moreover, this activity only provides single point source information. In contrary, Ground Penetrating Radar (GPR) is an extremely useful geo-physical technique to image subsurface in non-destructive and faster way for extracting information about buried utilities such as electric and telephone cables, water and sewage pipes and other infrastructure in dense urban areas. In order to explore multi-frequency GPR capabilities, GPR profiles/signatures have been collected and analysed for various surface and subsurface utilities, like concrete road, peat road, underground pipes made of different materials, manholes and various cables have been brought out in this paper. Moreover, the present article also emphasizes on a case study carried out in Ahmedabad City, Gujarat, India to examine disturbances in the soil and road layers associated with the road collapse due to damages in the underground pipes. The results conclude that GPR technique is highly efficient in identifying most of the underground utilities made of different materials and also the deformation features in the road and soil layers.

Keywords: Ground Penetrating Radar, Subsurface Utilities, Road Collapse

1. Introduction

High-resolution subsurface mapping of the ground in rapid, economical and non-destructive way has always been a necessity in the field of subsurface exploration. However, until recently there were no effective methods that can meet the above requirements. Traditional methods such as digging soil and roads are destructive and time-consuming. In contrary, Ground-penetrating radar (GPR) is a non-destructive geo-physical technique which can image the subsurface with higher resolution in faster and cost effective manner. GPR detects electrical discontinuities in the shallow subsurface (typically < 50 m) by generation, transmission, propagation, reflection and reception of discrete pulses in the megahertz frequency range (Neal, 2004). It is used as near-surface remote sensing tool to detect buried objects and to characterize the subsurface structure and properties in a wide variety of applications such as mineral and groundwater exploration, geotechnical and archaeological investigations, as well as rock mechanics and mine development requirements, subsurface utility detection, road condition analysis etc. (Davis et al., 1989, Mellet, 1995; Annan, 2002; Lambot et al., 2004). Hence, to demonstrate the capabilities of GPR, a detailed study has been carried out over multiple sites in Ahmedabad, Gujarat for deducing signatures of various surface and subsurface utilities, like concrete road, asphalt road, underground pipes made of different materials, manholes and other utilities. The results obtained from the present study are showcased in the present paper.

2. GPR principle

Ground Penetrating Radar (GPR) is a geophysical sensor that uses high-frequency (e.g. 20 to 1,500 MHz) electromagnetic pulses to image the subsurface in non-destructive way. A GPR transmitter emits microwave

pulses into the ground. When the energy encounters a buried object or interfaces between materials having different permittivity, it reflects or scatters back to the surface. A receiving antenna records the variations in the return signal (Figure 1). Reflecting interfaces may be soil horizons, the groundwater surface, soil/rock interfaces, man-made objects, or any other interface possessing a contrast in dielectric properties. However, dielectric properties of materials are not measured directly. The method is most useful for detecting anomalies and changes in the geometry of subsurface interfaces.

Two physical parameters of materials are important in subsurface wave propagation at GPR frequencies. One property is conductivity (σ), measured in mS/m (1/1000 Ω m). The other physical property is dielectric constant (ϵ), which is dimensionless. Good amount of energy of an EM field is consumed during interaction with water molecules or other polarizable materials. Thus, waves propagating through such material go slower and are subject to more attenuation (US EPA, 2016).

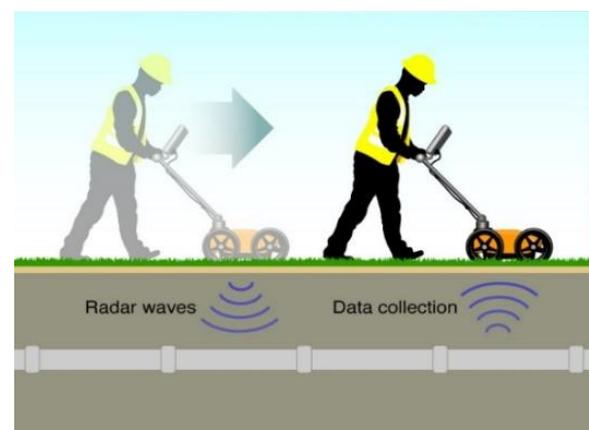


Figure 1: Schematic diagram representing GPR survey (<http://saarit.in/gpr.php>)

Velocity (V) of Radar wave in the medium:

$$V = (3 \times 10^8) / \sqrt{\epsilon}$$

Attenuation (A) of Radar wave in the medium:

$$A = 1.69\sigma / \sqrt{\epsilon}$$

As the GPR signals are in the form of radar pulses, so, as the antenna moves across a point object the range of the object from the antenna changes. When, the antenna is just above the object the range is minimum. Hence, in the time-distance plot, radar image of a point object takes the form of a hyperbola (Figure 2). Amplitude of the reflected radar signal also depends on the material type. Reflection from the metallic objects are more than non-metallic objects. Therefore, hyperbolas appear from metallic objects are more prominent than non-metallic pipes (Amran et al., 2017, Narayana et al., 2018).

3. Experiment setup and study area

GPR can be operated in different modes such as common offset mode, common midpoint mode and WARR (Wide

Area Refraction and Reflection) mode. In the present study, GPR is operated in the common-offset mode, where the transmitter and receiver are maintained at a fixed distance and moved along a profile line (Figure 3). A commercial GSSI based GPR with 100 MHz and 400 MHz antennas are used for subsurface mapping (Figure 4).

According to the prior information, the GPR experiments are conducted at SAC New Bopal Campus, SAC Main Campus, Bopal Cross Road (Near Vakil Saheb Bridge), Ranna Park and Jamalpur area of Ahmedabad, Gujarat. Based on the observations at these sites, the effective results are grouped into various themes and documented here for better understanding. The various object themes, location of the studied objects and the GPR antenna used to study are detailed in table 1.

Raw data from GPR profile was position corrected and processed with different filters for enhancing the subsurface reflections as part of post-processing using RADAN software. It was observed that after post-processing, all previously known underground objects are easily identifiable and their location (apparent) can also be retrieved from GPR 2D profiles.

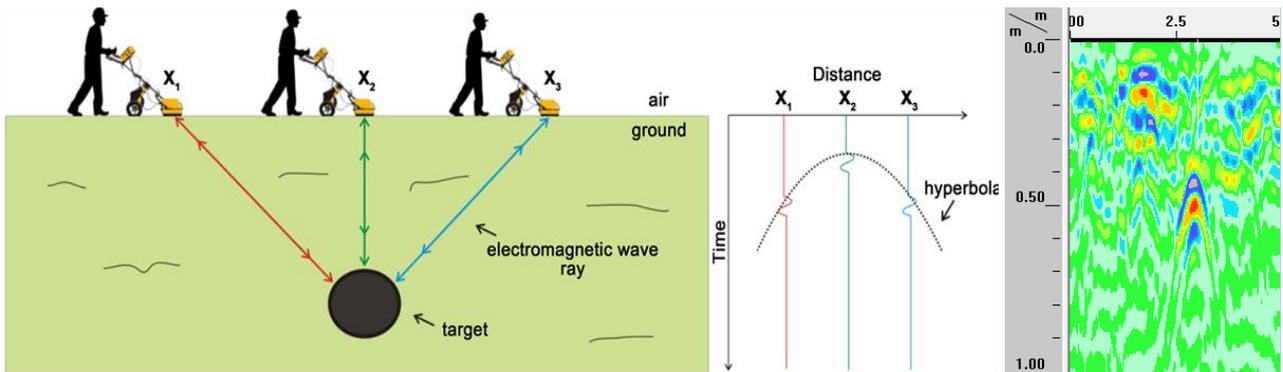


Figure 2: Inverted hyperbola formation from point object in GPR profiles (Poluha et al., 2017)

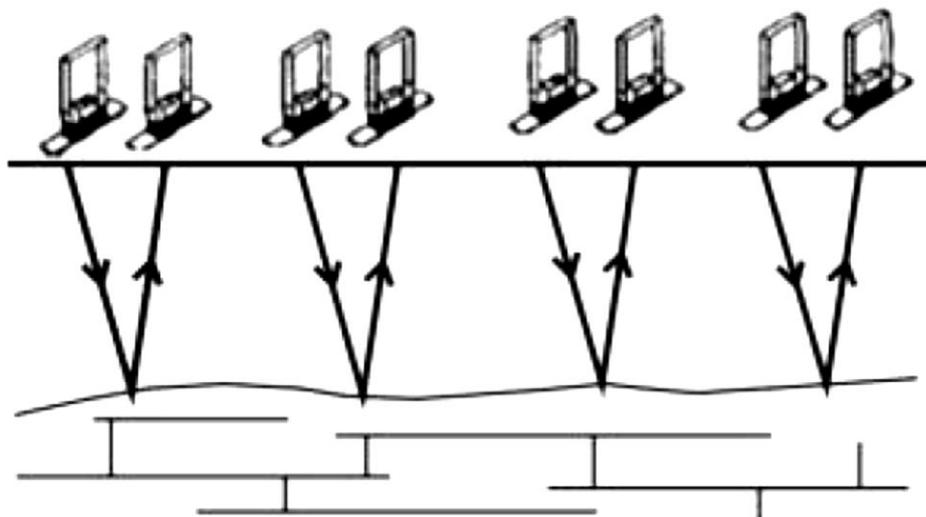


Figure 3: Common offset mode of GPR operation



Figure 4: GSSI 400 MHz GPR antenna (red) with control panel and carrier three-wheel cart. Left side HydraGo (blue) soil moisture instrument for measuring dielectric constant of soil.

Table 1: Various GPR experiments conducted in the present study

Sl. No.	Theme	Experiment Location	GPR Antenna used
1	Concrete Road	New Bopal SAC Campus	400 MHz
2	Asphalt/ Tar Road	All the sites	400 MHz
3	Manholes	New Bopal SAC Campus, Bopal Cross Road, Ranna Park and Jamalpur	400 MHz
4	Different Pipes and Utilities	All the sites	100 MHz, 400 MHz
5	Road and Subsurface soil profile	All the sites	400 MHz (Range 100 ns and 50 ns)

Sl. No.	Theme	Experiment Location	GPR Antenna used
6	Road collapsed and cave-in situation	Ranna Park	400 MHz
7	Cave-in repaired sites	Bopal Cross Road	400 MHz
8	Subsurface water leakage	SAC Main Campus	400 MHz

4. Results and interpretations

4.1 Concrete road

Concrete structures reinforced by iron mesh or rebar are prominently identifiable using GPR study. Iron is a metallic substance, hence, presence of iron in the concrete give rise to high dielectric constant variation which leads to strong reflection of radar signals. Therefore, in GPR profiles the rebar or iron mesh appears very prominently with high reflected signal power.

In New Bopal SAC Campus, a GPR profile is collected using 400 MHz antenna. The rebar structure looks like densely populated small hyperbolas (Figure 5c, 5d) in a line due to strong reflection of the radar signals from iron mesh.

4.2 Asphalt / Tar road

Dielectric constant variation between asphalt layer and soil is low hence asphalt layer mostly remain transparent to radar signals. However, due to direct coupling, straight horizontal bands appear on top of the GPR profile. Direct coupling is a combination of the transmit pulse in air and surface reflection from the top of the material. So, the direct coupling carries information about the structure. Hence, if the asphalt layer is continuous, the top layer in the GPR profile becomes straight. Any disturbances in the road layer causes discrepancy in the top layers in GPR profile.

In all the investigated sites, GPR profiles have been collected mostly on the asphalt road. In figure 6, continuous asphalt road layers are visible as the top layer in the GPR profile. Below the asphalt road layer, parallel soil layers prepared for construction of road are also visible in the GPR profile (Figure 6).

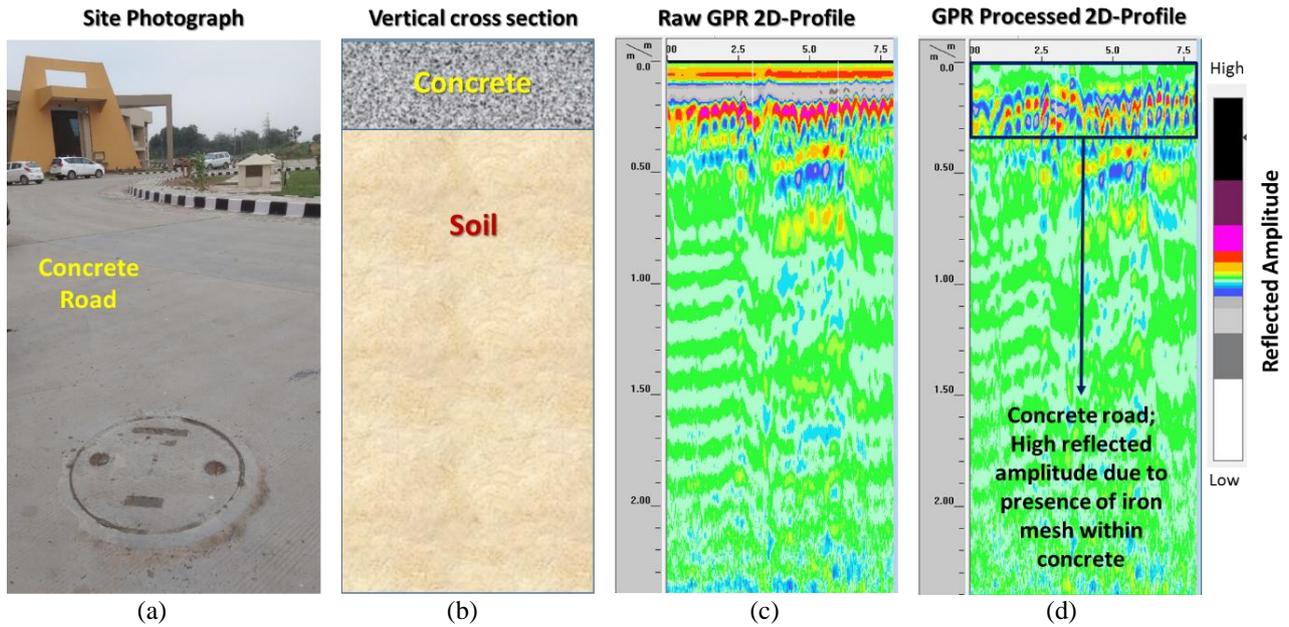


Figure 5: (a) Field Photograph where the GPR profile has been collected; (b) Cross-section of the profile based on prior information along which GPR signals are collected; (c) Position corrected GPR profile conducted on concrete road. Pink small densely populated hyperbolas are indicator of rebar in concrete road; (d) Background removed GPR profiles to highlight the rebar structure.

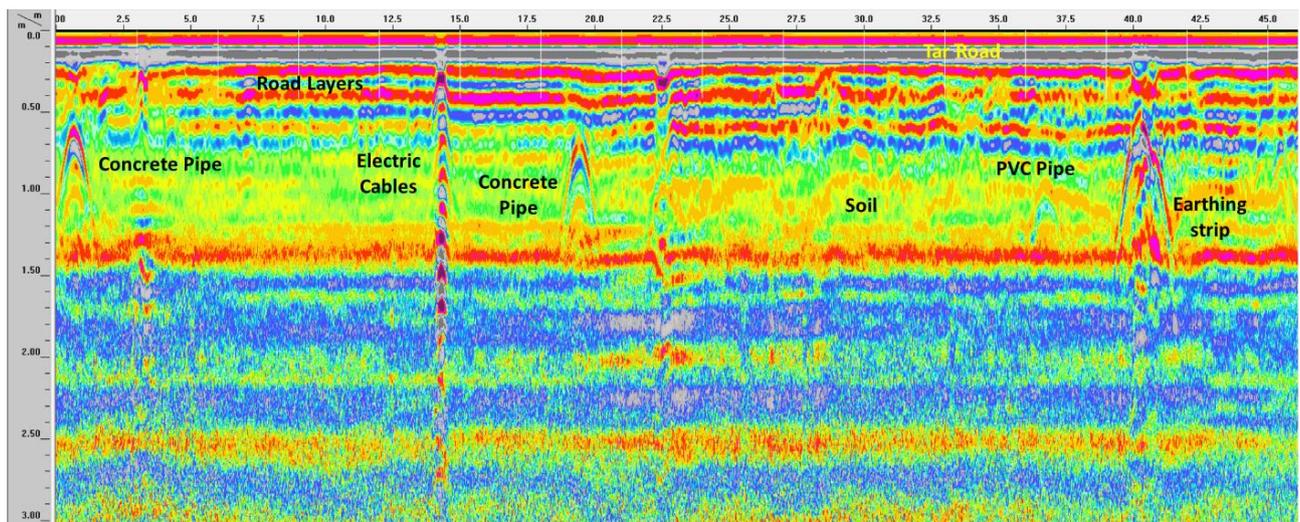


Figure 6: Continuous Asphalt layer as the top layer of the GPR profile.

4.3 Manholes

Manholes are vertical shafts filled with air, hence, act as air gaps. As air and surrounding soil has high dielectric contrast, hence, radar reflection from the manhole is strong (GSSI, 2018). Therefore, manholes appear as very prominent vertical structures in GPR profile (Figure 7a). Minute observation of the responses obtained from manhole reveals that the response is like ‘M’ shape and vertical stacks of ‘M’s give rise to the complete manhole structure. The ‘M’ shape is due to merging of two hyperbolas appearing from the walls of the manhole present either side of the air gap (Figure 7b).

For understanding responses from manhole structures, GPR profiles have been collected in New Bopal SAC campus, Ranna Park, Bopal Cross Road and Jamalpur area using 400 MHz antenna. The manhole structures are prominently decipherable in all the GPR profiles. However, in 100 MHz antenna due to long wavelength, manholes become transparent in the GPR profiles and not determinable. Thus, wavelength of the radar also plays very important role to decipher target object. If the object size is comparable with the wavelength it will appear in the GPR profile, else it will become invisible in the GPR profile. Figure 8 exhibit the series of manhole structures obtained in New Bopal SAC Campus.

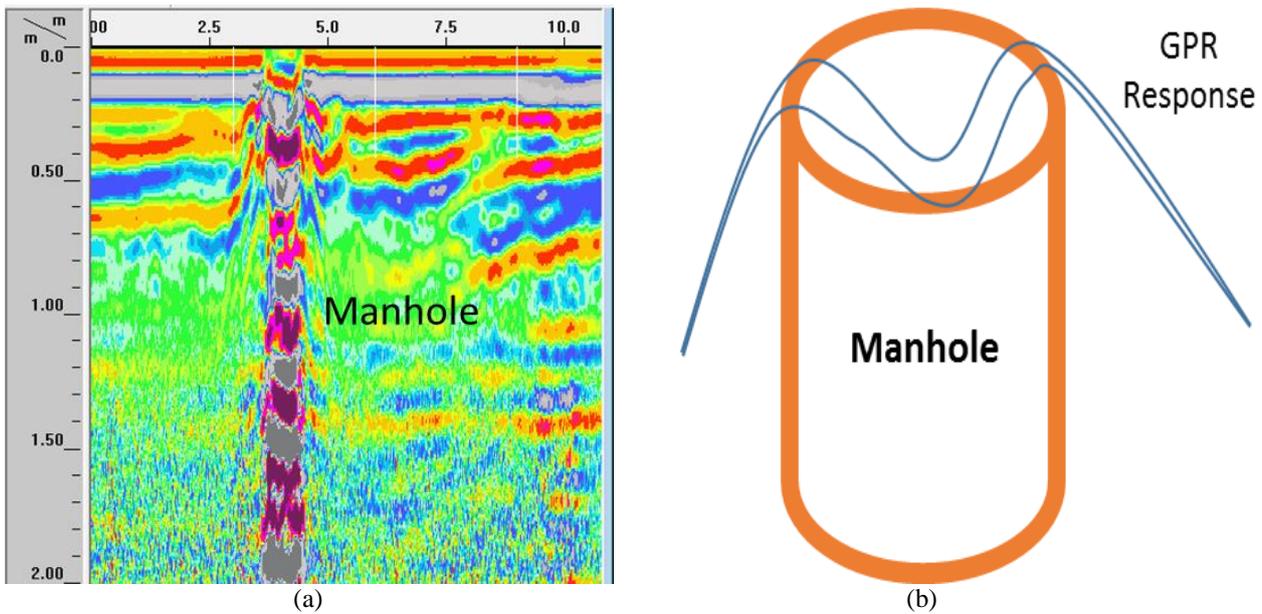


Figure 7: (a) GPR profile of a vertical manhole structure; (b) Reflected radar response from manhole structure

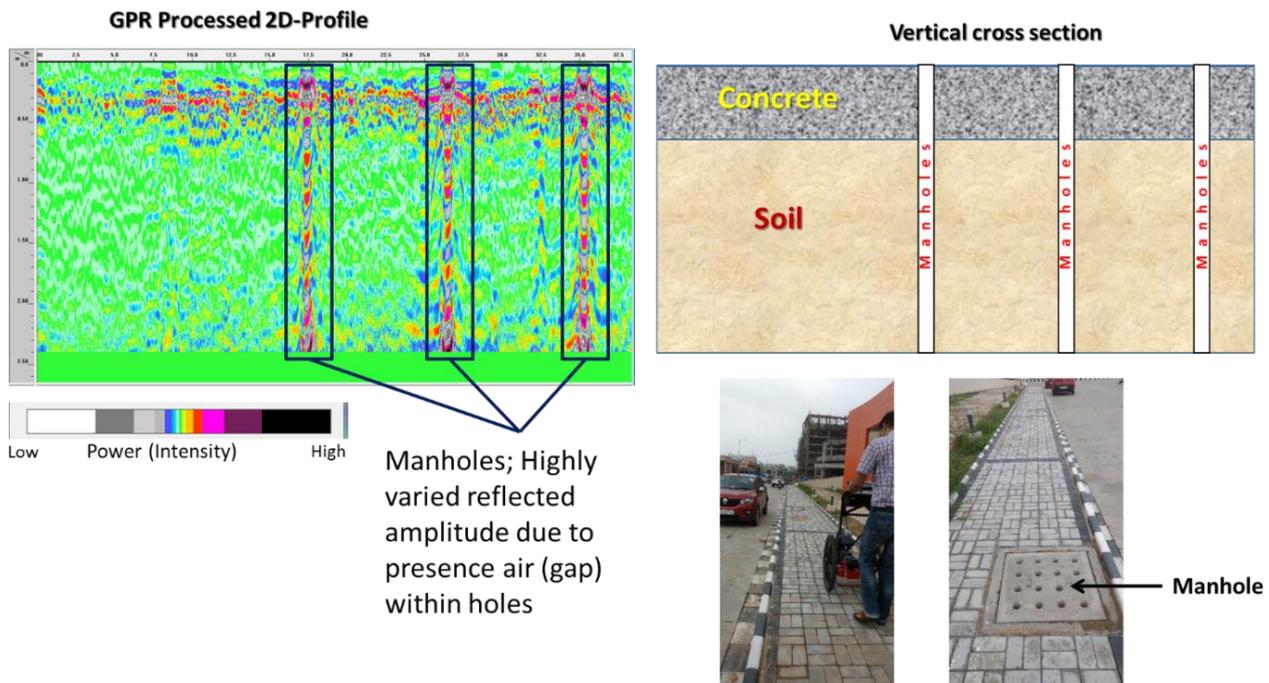


Figure 8: Series of manholes obtained from GPR profile in New Bopal SAC Campus

4.4 Different types of pipes and utilities

Underground utilities such as metallic and non-metallic pipes, electric cables and strips are successfully detectable using GPR survey. Underground utilities mostly act as a point object and their response mostly appear as hyperbolas. Metallic utilities have strong dielectric contrast with soil, so, strong radar reflection takes place leading to prominent appearance in the GPR profile. However, non-metallic utilities are less prominent. PVC pipes are mostly transparent, however, presence of air or water in the pipe helps to reflect radar signal from the pipe (GSSI, 2018). Hence, PVC pipes appear as weak hyperbola in the GPR profile. Figure 6 reveals various underground utilities like concrete pipes, electric cable, PVC pipe and Earthing strip obtained from GPR profile

collected in SAC Main Campus using 400 MHz antenna. Here electric cable and earthing strip are metallic object, so, they have strong radar reflection. Concrete pipes are also appearing prominent in GPR profile. However, radar reflections from PVC pipes are feeble.

Along with the material property, GPR profiles also help to delineate the size of the underground utilities in relative manner. Figure 9 represents a GPR profile which is having subsurface pipes of various sizes. According to the size of the pipes, the size of the hyperbola also varies. The length of the crest part of the hyperbola is comparable with the perimeter of the semicircle of the pipe looked along in the GPR cross-section.

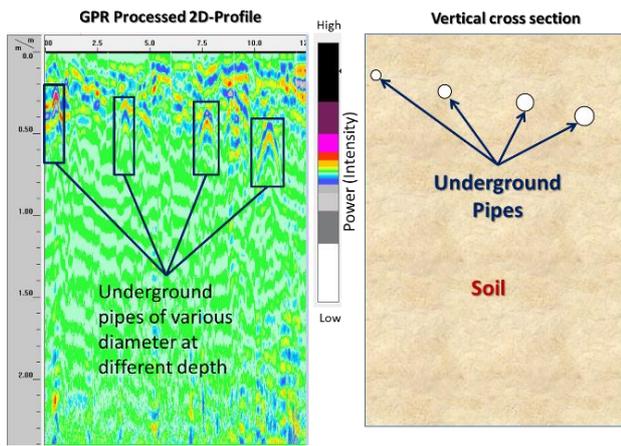


Figure 9: GPR profile showing subsurface pipes of different size present at different depth

If more than one pipes are present together, hyperbolas appeared from individual pipes in the GPR profiles will overlap with each other. Number of hyperbolas help to determine the number of pipes occurring together. However, if the distance between two pipes are less than the wavelength of the GPR signal, then GPR treat the two pipes as single object and only single hyperbola will appear in that case. As metallic objects act as strong reflector, GPR is not able to look below the metallic object. Hence, any pipe or object present below the metallic object is not decipherable in GPR profile. Figure 10, shows the occurrence of multiple pipes together in GPR profiles collected in SAC Main Campus. The depth of the object appeared in the GPR profiles are in apparent terms. The vertical depth accuracy of the object depends on the dielectric constant of the medium. More precise

input of dielectric constant give rise to increased vertical accuracy. In SAC Main Campus, the dielectric constant of the soil measured with dielectric probe and using the value the pipes showed in figure 10 appeared in 40-45 cm below surface. Ground measurement reveals that exact depth of the pipes are 46 cm below ground level.

400 MHz antenna having capacity to penetrate shallow subsurface upto 3m below ground. However, for locating pipes present in greater depth, 100 MHz GPR antenna is useful which can penetrate upto 8-10m below surface. However, the object size should be comparable with the radar wavelength in 100 MHz, else, the object will not be visible in the GPR profile. As requested by AMC, a study is conducted at Jamalpur, Ahmedabad using 100 MHz GPR antenna to locate deep-seated underground pipeline. Figure 11 demonstrates the result of locating large diameter concrete pipes present at the depth of 4m below surface.

4.5 Road and subsurface soil profile

GPR technology proves its potential in imaging subsurface soil profiles which is useful for geological understanding. While surveying in all the sites, the soil morphology in the GPR profiles can be prominently delineated. The uniform soil body appears to be uniform in GPR profile due to absence of any contrast in dielectric properties. However, if any variation in the soil formation prevails then due to variability, the interfaces between the soil layers can be decipherable easily. Moreover, any deformation in the soil layer can also be understood from GPR profile.

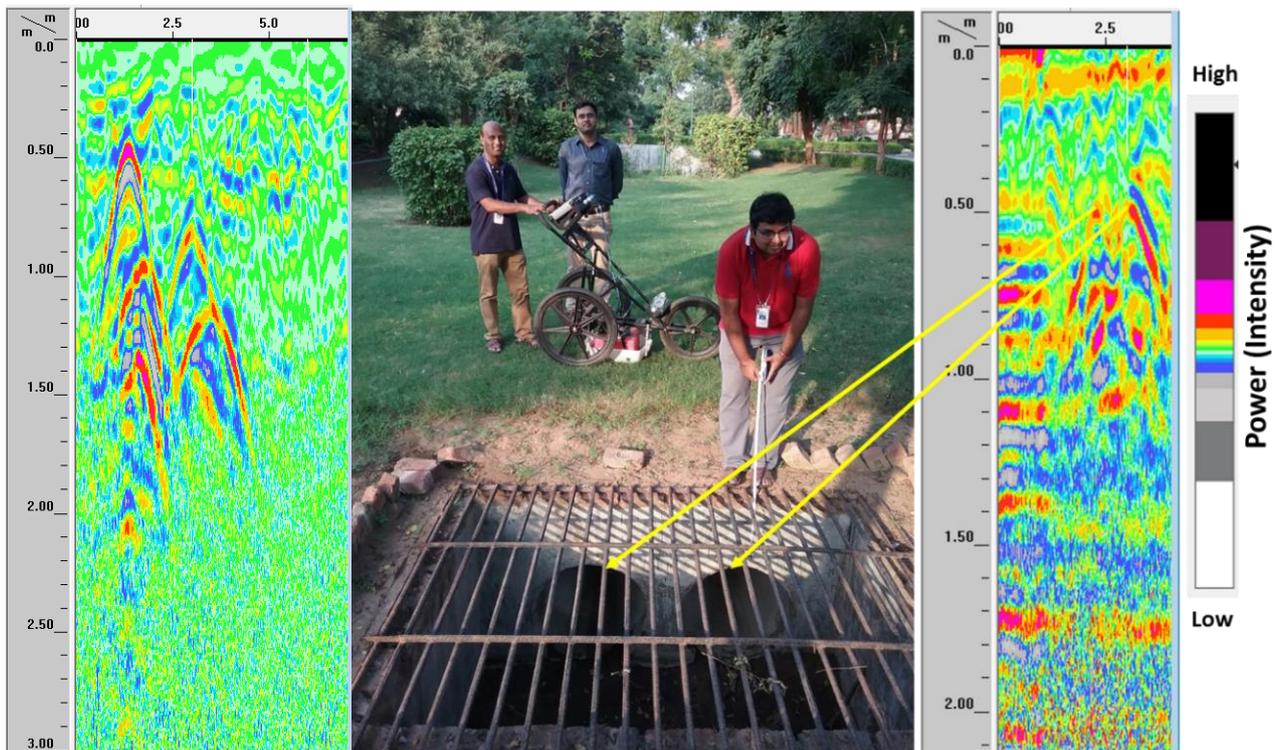


Figure 10: GPR profile showing occurrence of multiple pipes together

Figure 12 reveals the potential of GPR to identify uniform and layered soil profile present alongside. The survey is conducted at Ranna Park, Ahmedabad, with the help of officials of Ahmedabad Municipal Corporation (AMC), using 400 MHz GPR antenna with range value 50 ns and 100 ns. In 50 ns the penetration depth of the radar is upto 3m, however, in 100 ns GPR can image subsurface upto 6m. So, the profiles show that a uniform soil layer is

existing upto greater depth and within that some part of the soil was excavated for installation of pipeline and the excavated pit was filled by various soil layers. Thus, while GPR survey it is important to conduct the survey in various range to understand the subsurface profile in overall details. Figure 13 also demonstrates the capacity of GPR to image both uniform and layered soil profile distinctly.

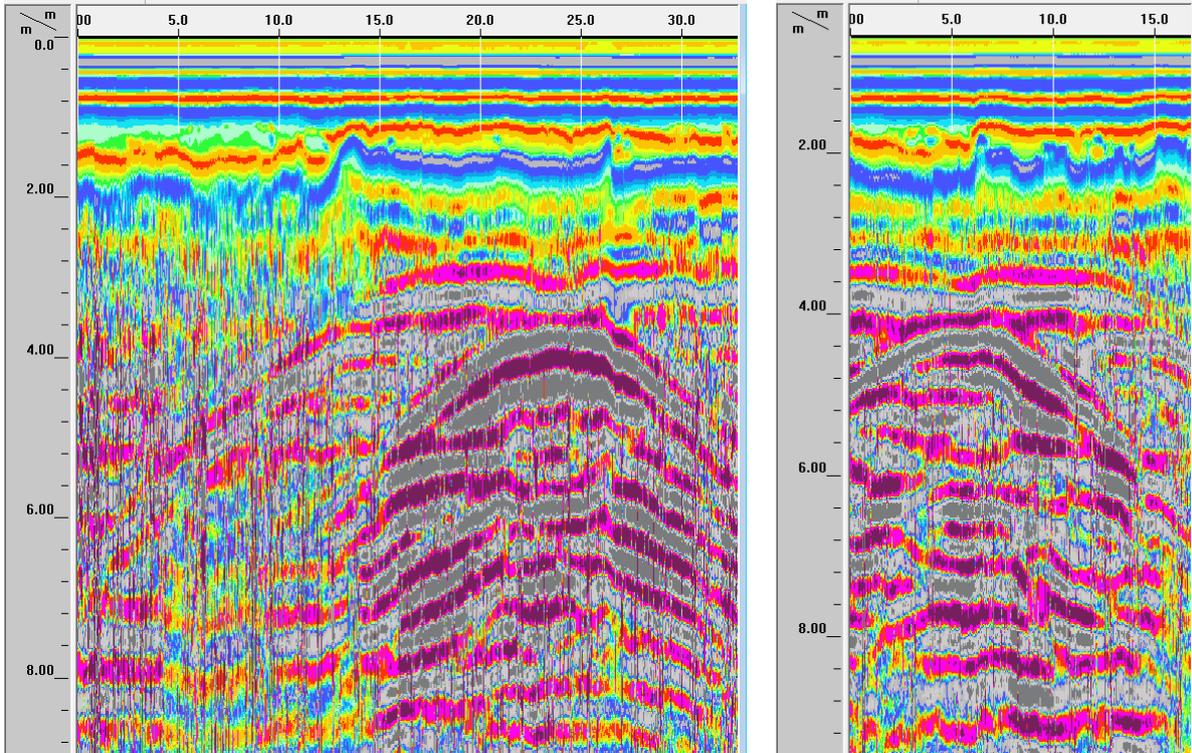


Figure 11: GPR profile showing deep-seated subsurface pipe using 100 MHz GPR antenna

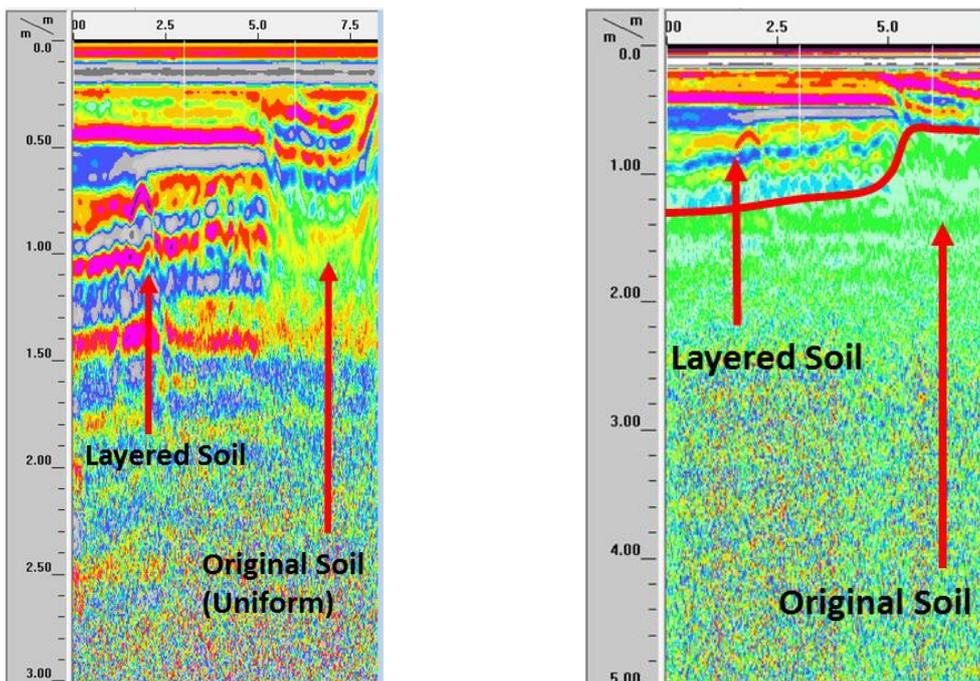


Figure 12: GPR profile showing uniform and layered subsurface soil profile using 400 MHz and in range value 50 ns (left) and 100 ns (right)

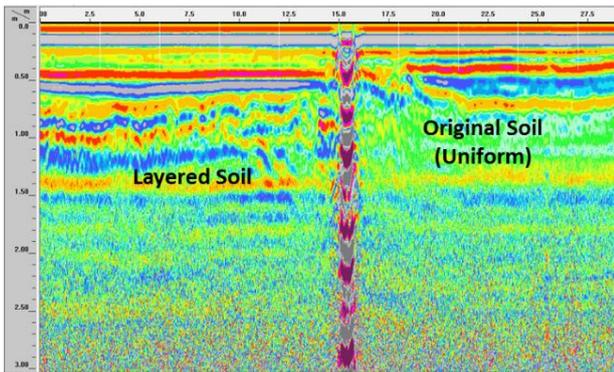


Figure 13: GPR profile showing uniform and layered subsurface soil profile.

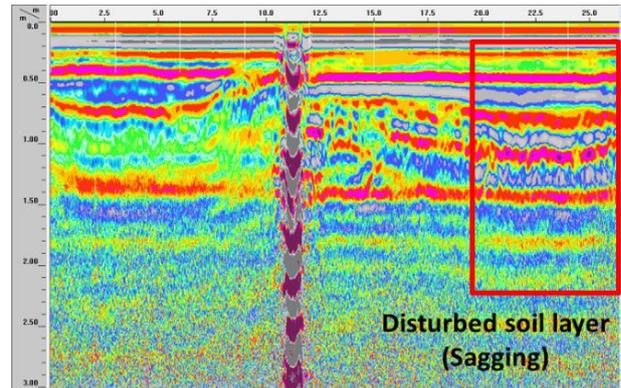


Figure 14: GPR profile showing deformation in the soil profile (sagging layers)

As GPR is able to exhibit the various soil layers, so, any deformation in the soil layers are easily distinguishable in the GPR profiles. Any bending, void formation and sagging of soil layers are understandable from GPR profiles. Figure 14 showing the GPR profiles collected in close vicinity of the cave-in area in Ranna Park, Ahmedabad using 400 MHz GPR antenna. The soil layers are sagged or deformed due to effect of cave-in. In another example from Jamalpur area, Ahmedabad, disturbed and undisturbed soil layers can be easily discriminated from GPR profiles collected using 400 MHz antenna (Figure 15). Thus GPR is useful in studying soil profiles and deformations occurring within it.

4.6 Road collapse and cave-in situation

In a case study, GPR survey has been conducted around cave-in area in Ranna Park, Ahmedabad with the help of Ahmedabad Municipal Corporation (AMC) officials and staffs. According to AMC officials, around 6m below the surface the concrete sewage pipe is placed in the area. The sewage pipe is having manholes which link the pipes with surface.

Due to presence of sewage material in the pipe, methane gas releases which accumulates in the roof of the pipe. The gas corrodes the concrete roof and reduces roof stability. When the roof of the pipe become unstable to resist overburden pressure, it fails and all the overburden soil then flows within the pipe to create a cave-in structure. This cave-in can extend upto the surface leading to road collapse (Figure 16).

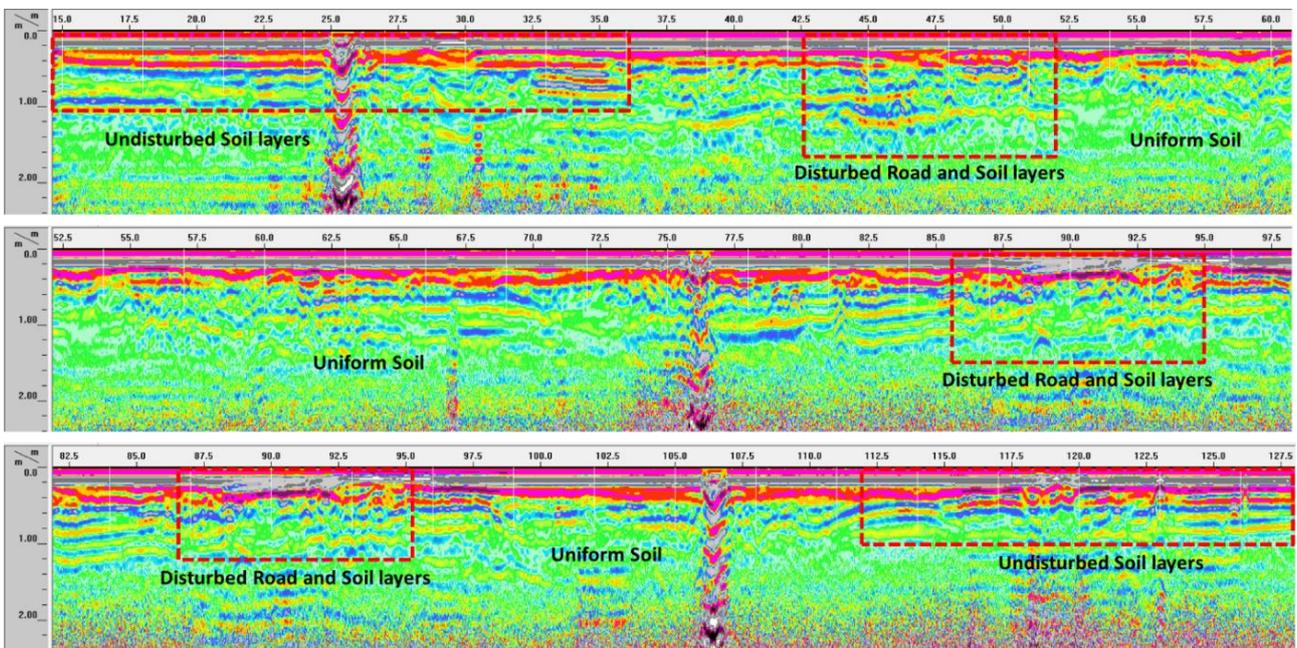


Figure 15: GPR profile showing undisturbed and disturbed soil profiles

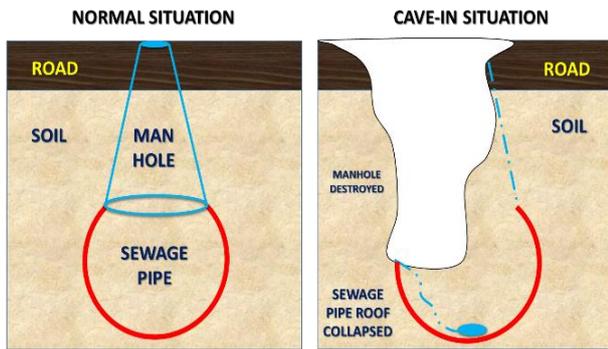


Figure 16: Occurrence of cave-in due to roof collapse of sewage pipe

In Ranna Park also similar road collapse associated with cave-in took place. A manhole connected to sewage pipe was present in that location, which got destroyed due to cave-in. Cave-in area is mainly a cavity structure filled with air. So, due to high dielectric constant variation in between soil and air, cave-in areas appear prominently in

the GPR profiles (GSSI, 2018). In this area, road collapse has occurred in small extent, however, the subsurface cavity generated below the road has larger dimension. GPR profiles are collected above the road below which the underground cavity still persist. Figure 17a shows such GPR profile taken on subsurface cavity in Ranna Park, Ahmedabad. Figures 17a and 17b show position corrected GPR profile to locate the subsurface cavity and background removed GPR profile to delineate the extent of the cavity and the field photograph.

4.7 Cave-in repaired sites

The repaired cave-in sites appear to be different than the original layers prevailing in the area. A GPR survey has been conducted in a cave-in repaired site in Bopal Cross Road, Ahmedabad. The profiles show that the road and soil layers in the repaired site are random and not in continuity with the prevailing original layers. Hence, clearly the cave-in repaired sites can be decipherable using GPR study (Figure 18).

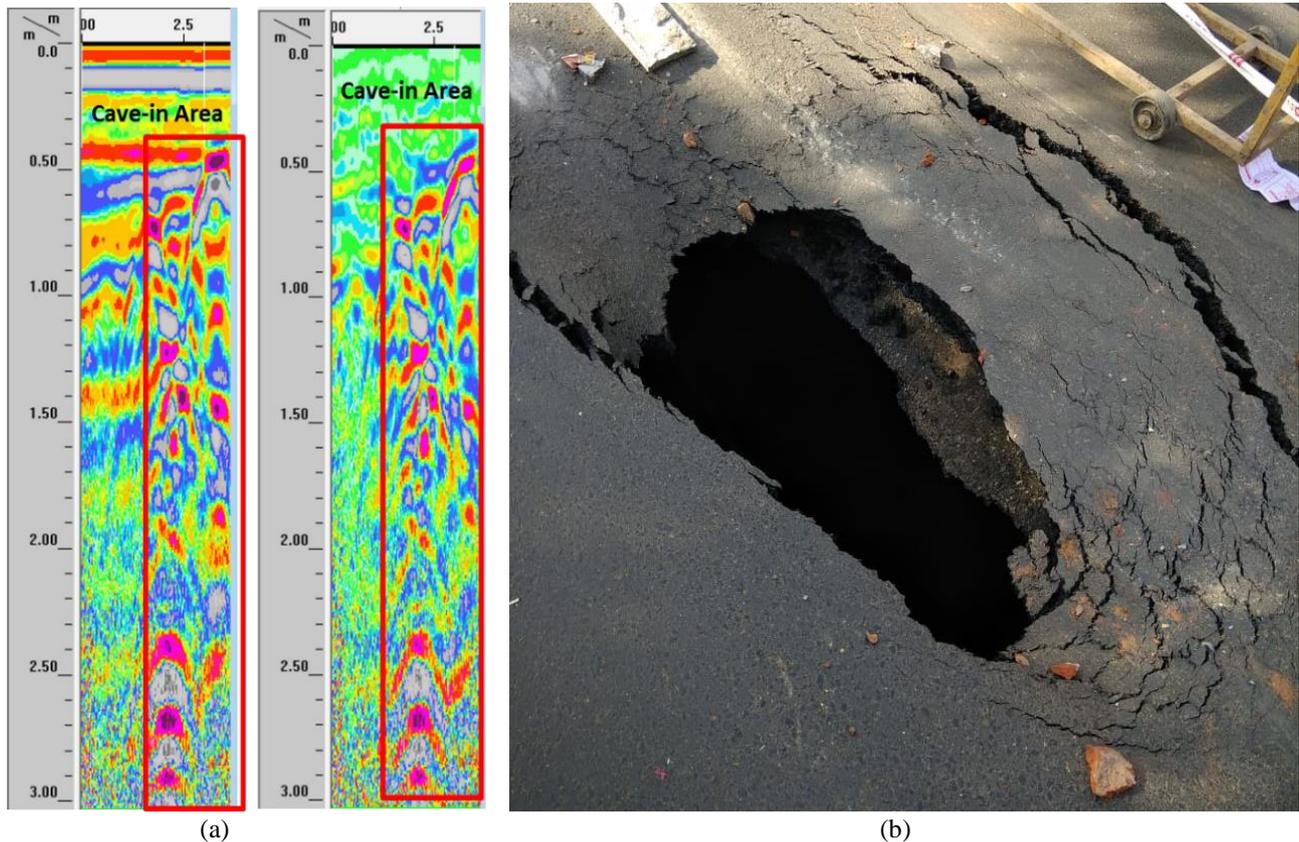


Figure 17: (a) GPR profile just above the cave-in area, where left profile is position corrected and right profile is background removed to highlight only the cave-in area; (b) Field photograph of the road collapse associated with cave-in in Ranna Park, Ahmedabad

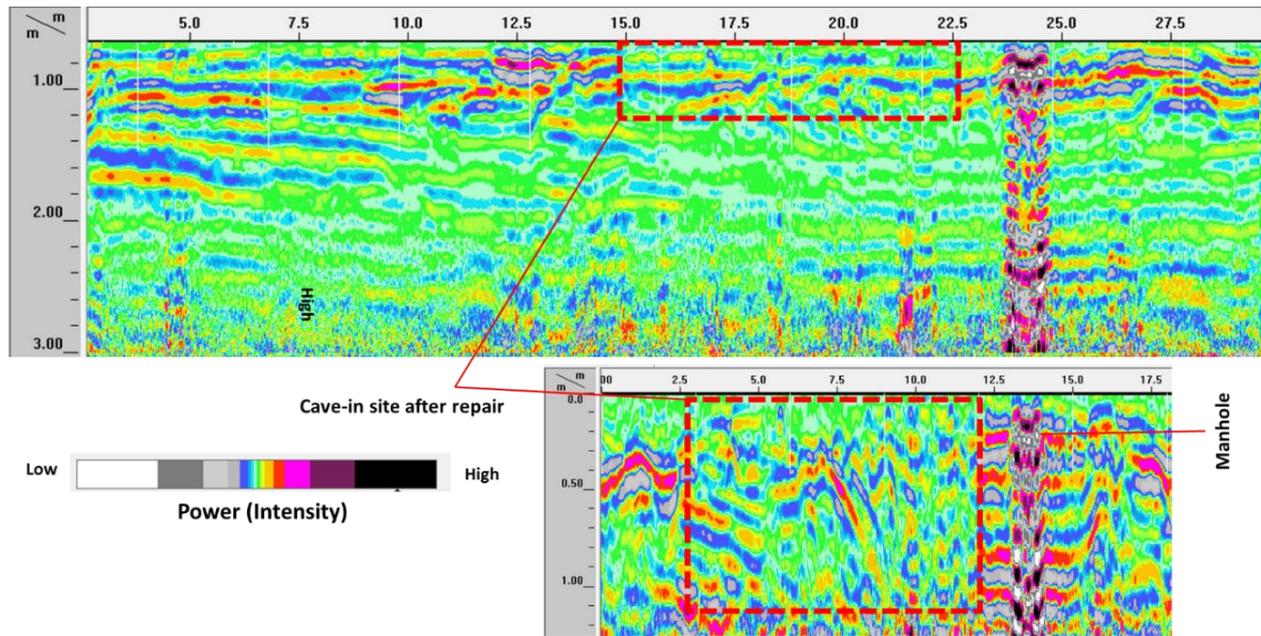


Figure 18: GPR of cave-in repaired site obtained using 400 MHz GPR antenna near Bopal Cross Road, Ahmedabad

4.8 Subsurface water leakage

Another case study was conducted to understand the potential of GPR to detect subsurface water leakage in SAC Main campus, using 400 MHz GPR antenna. As water and soil having contrasting dielectric properties, so, radar reflection for the wet area in background of dry soil is high and appears prominently (GSSI, 2018). The zone having high radar reflection also delineates the extent of the wet soil.

Figure 19a reveals that the wet soil area shows high radar intensity compared to the surrounding dry soil. This phenomenon is more understandable in B Scan GPR profile (Figure 19b) where, individual A scans in the wet soil area showing strong positive radar reflectance

compared to nearby dry soil, which helps to determine the extent of the water leakage.

5. Conclusion

The present study successfully explored the utilisation potential of GPR in detecting various underground pipes and utilities. The results show that GPR has greater proficiency of detecting concrete (rebar), manholes, metallic and concrete pipes compared to PVC pipe. Furthermore, GPR also has the ability to provide understanding about type of material and size of the pipe. In case of multiple occurrence of pipes, GPR is observed to locate the pipes depending on the distance between the pipes and the GPR wavelength.

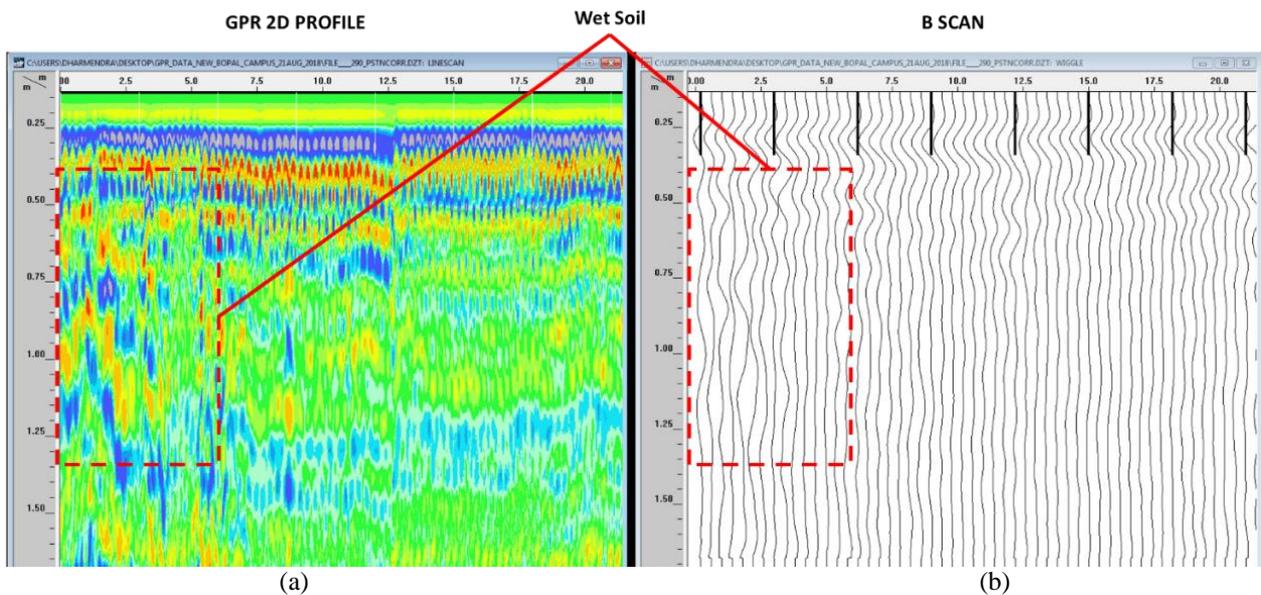


Figure 19: (a) GPR profile of the wet soil in background with dry soil; (b) GPR B scan of wet soil where individual A scan shows strong positive radar reflection compared to nearby dry soil

Further, the field survey also demonstrates the effectiveness of GPR technology in various subsurface road and soil layers. The survey also unveils its capability to detect subsurface deformations related to cave-in. GPR survey with lower frequency (100 MHz) antenna also effectively locates deep-seated utilities. Finally, GPR delineates efficiently presence of subsurface water.

Therefore, a significant amount of data collection has given an idea of the competency and efficiency of GPR, which led to understanding of extended application of GPR technology on mapping and determining various types of underground objects. Most importantly, this study allows the GPR frequency 100 MHz and 400 MHz to be used with confidence as a tool to detect underground pipes and utilities.

Acknowledgements

We would like to express our sincere gratitude and thanks to Shri D K Das, Director, and Dr. Raj Kumar, Deputy Director EPSA, for providing a platform for working in this area. Our deepest and heartfelt thanks also goes to EPSA team for their support whenever required. We would like to thank CMG team for their assistance and providing prior information about the study areas within SAC Campuses. We would also like to thank officials and staffs of Ahmedabad Municipal Corporation (AMC) for providing their continuous support and information during GPR survey in various locations of Ahmedabad, Gujarat. Finally, we like to thank GSSI, India team for extending technological assistance to us for solving any trouble encountered while GPR survey.

References

Amran, T.S.T., M.P. Ismail, M.R. Ahmad, M.S.M. Amin, S. Sani, N.A. Masenwat, M.A. Ismail and S.H.A. Hamid (2017). Detection of underground water distribution piping system and leakages using ground penetrating radar (GPR), AIP Conference Proceedings 1799, 030004, doi: 10.1063/1.4972914

Annan, A. (2002). GPR—History, trends, and future developments, *Subsurf. Sens. Technol. Appl.*, 3, 253–270.

Davis, J. L. and A. P. Annan (1989). Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy, *Geophysical prospecting*, 37, 531-551

GSSI, Geophysical Survey Systems, Inc. (2018). *Utility Locating handbook*, MN72-615 Rev B

Lambot, S., E.C. Slob, I. Bosch, B. Stockbroeckx and M. Vanclooster (2004). Modeling of Ground-Penetrating Radar for Accurate Characterization of Subsurface Electric Properties, *IEEE Transactions on Geoscience and Remote Sensing*, 42(11), 2555-2568

Mellet, J. (1995). Ground penetrating radar applications in engineering, environmental management, and geology, *Journal of Applied Geophysics*, 33, 157–166

Narayana, R.B., D.K. Pandey, D. Putrevu and A. Misra (2018). Impact of sub-surface roughness on GPR performance using numerical modelling and simulations, SAC Internal Report, SAC/EPSA/AMHTDG/MTDD/TR-04/2018

Neal, A. (2004). Ground-penetrating radar and its use in sedimentology: principles, problems and progress, *Earth-Science Reviews*, 66, 261–330.

Poluha, B., J.L. Porsani, E.R. Almeida, V.R.N. Santos and S.J. Allen (2017). Depth estimates of buried utility systems using the GPR method: Studies at the IAG/USP Geophysics Test Site, *International Journal of Geosciences*, 8, 726-742

US EPA (2016). *Environmental Geophysics-Ground-Penetrating Radar*, Available: https://archive.epa.gov/esd/archive-geophysics/web/html/ground-penetrating_radar.html

Development of universal geospatial data collection application and visualisation platform

Arpit Agarwal

Space Applications Centre, ISRO, Ahmedabad, India

Email: arpitagarwal@sac.isro.gov.in

(Received: Dec 20, 2018; in final form: May 31, 2019)

Abstract: Collection and efficient dissemination of ground truth data is a challenge faced by many research groups. Collecting data from multiple sources, transforming them to a common format, filtering the data based on parameters of interest is cumbersome. Presence of multimedia information such as images further adds to the complexity of collecting and storing the data efficiently. To address these issues, a centralised data collection tool is developed, which could collect data along with geo-spatial information. The collection tool is an Android application where a schema for data collection can be defined. Collections forms can be auto generated from the schema on the app. This data can be pushed to centralised data store. A visualisation portal is provided to view the collected data on Maps. The data can be downloaded in commonly used geospatial file formats like GeoJSON and kml.

Keywords: Data collection, Mobile App, Data Management, Web-GIS

1. Introduction

Efficient collection of ground truth has always remained a challenge for the researchers, particularly when collection and distribution is required across multiple functional groups. The available ground truth data is difficult to search/retrieve, filter, operate on and transform. There are some common traits in the ground truth data used for remote sensing applications viz. geo-spatial coordinates and time stamp. Other than that, every domain has some domain specific information to be collected. For example, ground truth of crop data may contain crop name, adjacent crops, disease information etc., while ground truth for wetland data may contain the name of wetland, its type (lake, pond, river etc.) and other related information.

An android application is developed which allows creation of multiple projects and could automatically collect the geospatial and time information. Each project can define its own fields using a schema. Based on the schema defined for a project, a data collection form is generated on the fly with geo-spatial and temporal information filled in. The information could be saved directly on servers. In cases where internet connectivity is not available, data can be collected on the device with option to sync it to the servers later.

For collection of multimedia data, multimedia asset management services are developed which allow upload/download of images to/from the servers. A special data type is defined in schema so as to enable upload of image for the specified field.

To make this data available to the users, a web interface is developed which allows the data to be viewed on maps using Web-GIS techniques. The interface allows search based on project or date of collection. It also allows a free text search over the collected data. The data collected through the app is available for download in most common and widely used formats such as GeoJSON and shape files.

2. State of the art

Various content management systems (Margaret, 2016) are available which allow creation, storage, search and visualization of data across multiple collaborating users. A content management system typically has two parts viz. a content creation/modification platform and a content delivery application. Drupal (Tomme, 2017), WordPress (Paulik, 2013) etc., are some of the commonly used content management systems.

Geospatial content management systems allow data to be stored along with spatial information viz. latitude and longitude. Google maps and map-server are examples, which are not only geospatial content management systems, but also provide a building block for other such systems.

Open Data Kit (ODK) (Ghosh and Dasgupta, 2015) provides tools to build forms (ODK Build), collect Data (ODK Collect) and a server (ODK Aggregate) which could save data to a database and provide generic visualization on the data. Mobile Data Conversion Kit can be used to export the data to GIS software like ESRI, QGIS or Google Earth.

3. System design

The system should be capable to store data with varying attributes along with geospatial information, provide efficient search over the collected data, support multiple users and possibly, multiple languages. The system should be able to handle multimedia data like images. Moreover, the collection should be possible in areas where internet connectivity is not available with the device.

Considering the challenges, a high level system design of the universal data collection infrastructure is shown in figure 1. The major components are explained below.

3.1 Web Service

The Web Service provides HTTP REST APIs. It is developed using the Java Spring Boot (projects.spring.io) framework. It uses an embedded tomcat server to listen to the requests, and does not need a separate tomcat server to be deployed.

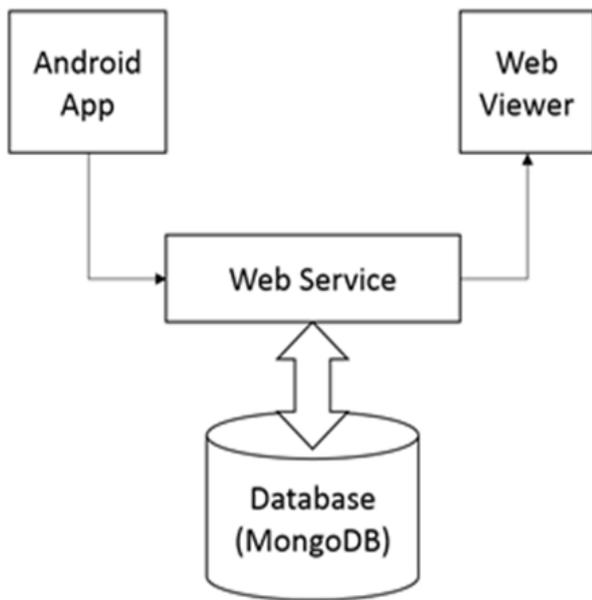


Figure 1: Architecture of universal data collection infrastructure

3.2 Database

MongoDB (Daylay, 2014) is used as database, which is a no-sql document store. As the data to be stored has variable fields, a document store seems apt for the use case. MongoDB supports UNICODE, enabling support for multiple languages. Moreover, it also provides text indices for efficient free text search. MongoDB has been proven to store millions of documents and outperform some other databases available (Bhat, 2015). The latencies generally remain less than one second.

3.3 Android app

The android app enables the user to define projects, schemas and add data. It uses the REST APIs provided by the Web Server.

3.4 Web viewer

Web Viewer enables the user to visualize the data collected using the Android App on maps using Open Layers 3. It uses the Web Server to query the requested data.

4. Multimedia asset management

Storing and transferring images is a challenge in offline stores. Images were typically stored on hard drives and the users store the path with other data elements, typically lying in a csv, excel or a DB. This way, visualisation of data along with associated image(s) becomes difficult. This also applies to data transfer. While transferring any data, images needed to be copied and provided separately.

The multimedia asset management services are developed to overcome this challenge. An image can be encoded as a string and uploaded to the server as a byte stream. The server reconstructs the image from the binary stream and stores the file on a network storage. The file path is stored in a database along with a unique identifier which is returned to the user. The user can use this identifier to download the image back using a browser whenever

needed. While downloading the image, the server searches the image path in the database based on the provided identifier. Once the path is found, the image is sent in the response as byte stream, which can be rendered in the browser.

5. User interfaces

There are two main user interfaces in the universal geospatial data collection viz. Data Collection App and Data Visualization tool.

5.1 Data collection app

The data collection app is an android application. which mainly includes three pages:

(i) Project listing page

This page lists all the available projects in the system with option to edit the project or add data to the project. The projects can be saved offline for later use. In case Internet is not available, offline list can be displayed. The project listing page is shown in figure 2. Clicking on edit will take the user to schema definition page. The page also provides an option to create a new project.

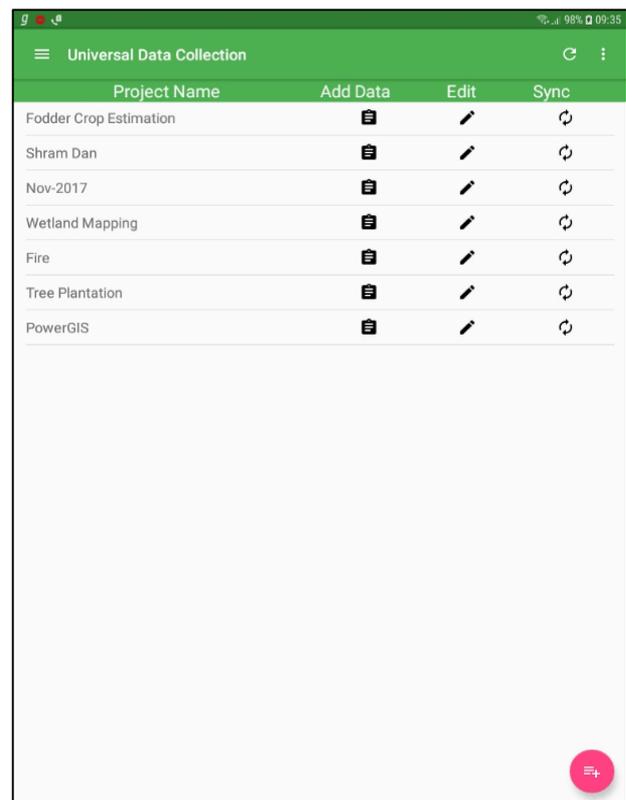


Figure 2: Project listing page

(ii) Project/Schema definition page

This page allows a user to define fields that are required to be collected as part of the project. Figure 3 shows a schema definition page for a sample project. Figure 4 shows the interface to define a field in the schema and the supported data types. A defined schema can be saved to be used later offline. This page also allows a user to edit or delete a field.



Figure 3: Schema definition page

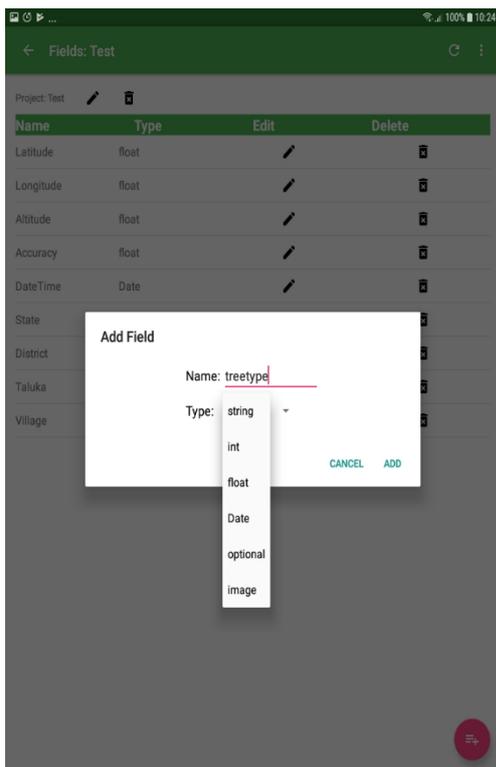


Figure 4: Field definition window

(iii) Data collection page

Based on the defined schema, this page auto-generates a form where the values corresponding to the expected fields can be filled and saved. Figure 5 shows an example of the form generated for project schema defined in figure 3. Geospatial information fields and timestamp are included as default and non-modifiable fields to every project. The geospatial fields include Latitude, Longitude, Altitude and

Accuracy. The geospatial information and the current time stamp are auto populated in the respective fields. The geospatial information is read from the GPS sensor. The geospatial information can also be read from NavIC receivers making it possible to integrate the system with devices carrying NavIC receivers. In addition to the coordinates, the state, taluka and village details are also added as default and filled by the system to facilitate people to identify the place correctly. Users can modify the system suggested state, taluka and village, in case the suggestion is not correct.

The data can be directly sent to the servers if Internet is available, otherwise, it allows the data to be saved on the device with an option to sync the same later.

5.2 Data visualization tool

This is an independent web application which allows the user to visualize the project wise data on maps using Open Layers 3. It uses the same service layer as used by the android app, to retrieve the information from the database. Several Overlay layers, including administrative boundaries and satellite images, have been provided for reference. The satellite reference layers provided are RISAT, LISS 3, Bhuvan High Resolution Maps etc. Google Map is also available as a reference layer. Any layer, available as Web Map Service can be overlaid for value addition.

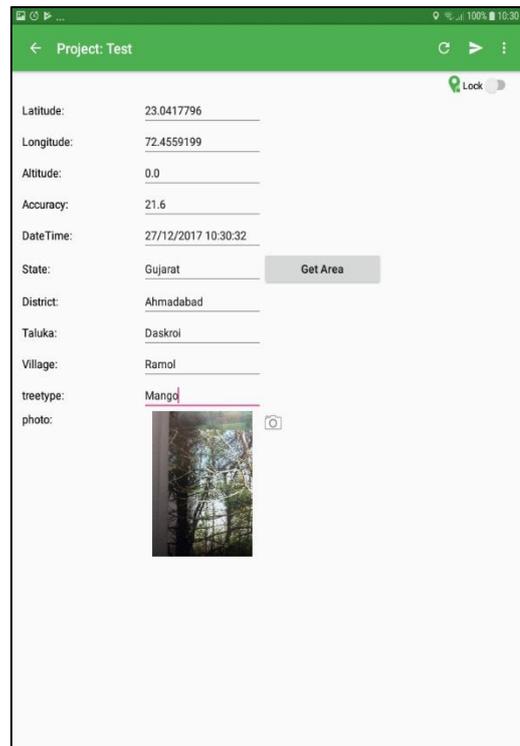


Figure 5: Data creation page

The data can be searched based on date of collection as well as any text string. For example, if a user searches for *mango*, all the data points which contain *mango* as a sub-string are returned. The data is plotted on the map based on geospatial coordinates. Clicking on a point shows the other information contained at that point.

Figure 6 shows the selection page for data visualization and figure 7 shows the popup which displays the information contained at a given sample point.

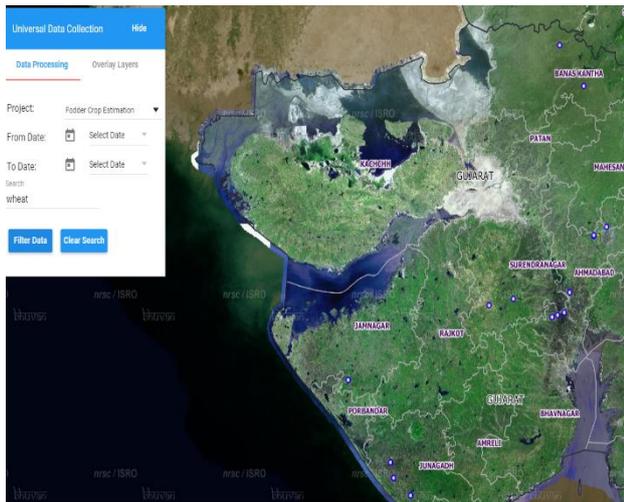


Figure 6: Content delivery application – data filtering

6. Data validation

All the data entered for a selected project is validated against the schema definition of the project. The data type of every field should match with the one defined in the schema. For example, if a user has defined a field, say quantity as integer, then filling “abc” for this field is considered invalid. This validation takes place both at the android app and service backend. A user should have a contributor role for the project he wants to collect and upload data for. A user can be assigned roles by the project administrator. By default, a user who creates a project, becomes its administrator, Data validation at service backend ensures that even if someone bypasses the app, and tries to call the REST end points through scripts, database is not polluted with unintended or malicious entries.

7. Deployment

The Web Server is deployed using an embedded tomcat web server. There are no CPU intensive operations and the web server can be run on commodity hardware. In case of high load, typically millions of calls per minute, RAM and networks are the first things to saturate. As MongoDB supports replication and sharing, the system can be scaled horizontally by adding more nodes.

MongoDB is set up as a service with configured paths for data and logs. Images are stored on a network storage. The images are organised based on a nested folder hierarchy based on year, month and date.

The Web Viewer is available for access at: http://vedas.sac.gov.in/vstatic_1/UDC/. The data collection android app is available at VEDAS (Visualisation of Earth Data and Archival Systems) website.

https://www.vedas.sac.gov.in/vedas/downloads/Android_Apk/Universal_Data_Collection.apk.

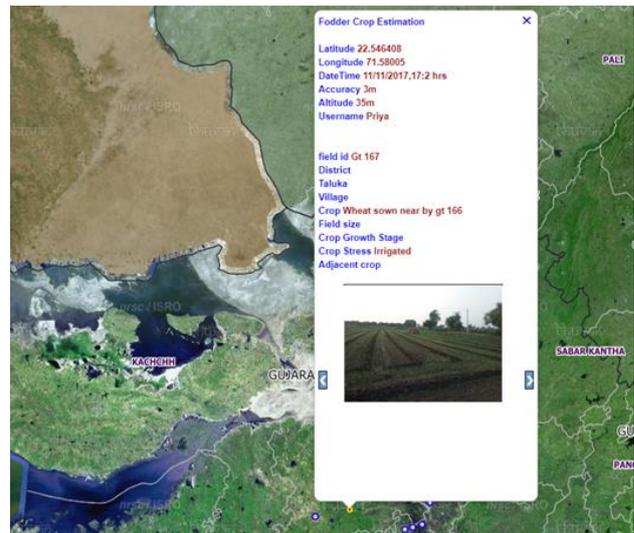


Figure 7: Data selection in content delivery application

8. Discussion

The app is being used by teams working in the field of agriculture and wetland mapping. It is also being used to track tree plantations by staff members at Space Applications Centre.

In future, plan is to increase the adoption by onboarding more teams and users. Work is in progress to provide authentication and authorization for data access and upload so that projects not open for public viewing can also leverage the platform.

9. Conclusion

A user friendly app has been developed for field data collection and sharing of data. An intuitive web interface is provided for data visualization. Researchers from various themes can use this platform to collaborate data among the community. The app is available in the downloads section of VEDAS website <https://www.vedas.sac.gov.in/vedas>.

Acknowledgements

The author sincerely thanks Shri Shashikant A Sharma, Group Head, VEDAS Research Group and Dr. Markand P Oza, Head, Content Generation and Dissemination Division at Space Applications Centre, Ahmedabad for their support, guidance and vision throughout the development process.

References

Bhat, S. (2015). High Performance Benchmarking: MongoDB and NoSQL Systems. Available: <https://www.mongodb.com/blog/post/high-performance-benchmarking-mongodb-and-nosql-systems>. [Accessed: 01-01-2018].

Ghosh, M. and S. Dasgupta (2015). How to use Open Data Kit (ODK): A brief tutorial, Global Change Programme-Jadavpur University Working Paper # GCP/JU/15/02.

Margaret, R. (2016). Content Management System (CMS). <http://searchcontentmanagement.techtarget.com/definition/content-management-system-CMS>. [Accessed: 01-01-2018].

Paulik, C. (2013). A Complete guide to WordPress content management. <https://managewp.com/wordpress-content-management-complete-guide>. [Accessed: 01-01-2018]

Projects.spring.io. Spring Boot. Available: <https://projects.spring.io/spring-boot/>. [Accessed: 01-01-2018].

Tomme, K.V. (2017). Concept: Drupal as a Content Management System. https://www.drupal.org/docs/user_guide/en/understanding-drupal.html. [Accessed: 01-01-2018].

Interannual variation of atmospheric carbon dioxide concentration derived from Orbiting Carbon Observatory-2 (OCO-2) data compared with its local and large-scale variations obtained from NASA Giovanni database

Barun Raychaudhuri* and Tithi Sen Chaudhuri
Department of Physics, Presidency University, Kolkata, India
*Email: barun.physics@presiuniv.ac.in

(Received: Dec 20, 2018; in final form: Jun 01, 2019)

Abstract: The temporal change of global atmospheric carbon dioxide (CO₂) exhibits a positive trend superimposed by an annual cycle, which is ascribed to the combined effect of ecosystem productivity and anthropogenic activities. This work intends to present an analytical treatment of the phenomena in terms of periodicity, local data and their averaging. The time variation of CO₂ concentration derived from NASA-JPL Orbiting Carbon Observatory-2 (OCO-2) database was analyzed with Panoply open source software for two consecutive years 2016 and 2017. The temporal change of the monthly average zonal mean CO₂ flux was obtained from NASA Giovanni earth science database for contrast areas, such as dense forest, desert and crowded city for the period of 2010-16. The results yielded wide differences in the nature of change of both the increasing trend and the seasonal alteration. The Fourier transform indicated the existence of two periodicities superimposed including one sub-annual variation. Possible explanations were given for such general trends in the global change of CO₂. The work was more concentrated on the Indian subcontinent. Prominent difference was found in the positive trend of temporal change at the east coast and the west coast regions over last ten years. The local disparity of CO₂ flux over and around small regions of Northern India exhibited irregular periodic increase and decrease, which when averaged, led to more regular periodicity.

Keywords: Carbon dioxide, OCO-2, NASA Giovanni

1. Introduction

Since the beginning of the industrial age, the concentration of atmospheric carbon dioxide (CO₂) has been increasing at alarming rate. Especially during the last two decades, anthropogenic activities like fossil fuel burning and deforestation have enhanced the CO₂ concentration more rapidly (Etheridge et al., 1996; Leung et al., 2014). Global satellite-borne monitoring systems, such as *Scanning Imaging Absorption Spectrometer for Atmospheric Cartography* (SCIAMACHY) (Bovensmann et al., 1999), *Greenhouse gases Observing Satellite* (GOSAT) (Hamazaki et al., 2004) and *Orbiting Carbon Observatory-2* (OCO-2) (Crisp et al., 2017) and site-specific ground-based systems, such as *Total Carbon Column Observing Network* (TCCON) (Wunch et al., 2011) have been developed for spatial and temporal sampling of atmospheric CO₂ concentration as part of the worldwide mitigation mission for CO₂ emission. Comparative findings of the above sensing systems are reported recently (Wunch et al., 2017; Liang et al., 2017).

The general finding in the temporal change of global atmospheric CO₂ is a positive trend superimposed by an annual cycle, which is ascribed to the combined effect of ecosystem productivity and anthropogenic activities. (Keeling et al., 1976; Basu et al., 2014; Jiang et al., 2016). However, there are wide fluctuations and uncertainties in the rate of CO₂ increase due to various climatic effects, such as natural sink variability (Fu et al., 2017), non-uniform warming (Li et al., 2018), anthropogenic contributions (Xueref-Remy et al., 2018) and wind direction (Roman-Cascon et al., 2019). Accurate monitoring of the above features is necessary for the prediction of future atmospheric CO₂ condition. This work intends to present an analytical treatment of the temporal change in CO₂ in terms of periodicity, local data and their

averaging. In particular, the local variations of CO₂ flux over and around small regions of Northern India are studied.

2. Materials and methods

This work utilizes sample data on the spatial and temporal changes of atmospheric CO₂ concentration downloaded from the following two resources. The daily information on CO₂ concentration (ppm) for the years 2016 and 2017 over an area of 22.5°–23.0° N and 86.0°–89.0° E were obtained by using the data produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology and obtained from the OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center. The monthly averages of CO₂ concentration (ppm) for the period of 2010-16 were procured from NASA Giovanni v4.28 online environments for geophysical parameters.

The Giovanni data were procured for contrast geographic features, such as dense forest, barren desert and crowded urban regions, as mentioned in table 1. More attention was laid on selective places of India, especially those of West Bengal containing Kolkata and around, where ground spectroscopy and airborne imaging were carried out. The OCO-2 data were downloaded with python program and were analyzed with Panoply (v4.9.2) open-source Java application. OCO-2 estimates the column-averaged concentration of atmospheric CO₂ from absorption spectroscopy of reflected solar radiation detected around 0.765 μm (molecular oxygen A band) and two carbon dioxide bands centered around 1.61 μm and 2.06 μm. The CO₂ concentration values derived from the 1.61 μm band are used here for analysis. Sorting from the OCO-2 daily database, the concentration values over an area of 22.5°–23.0° N and 86.0°–89.0° E for the Julian days of years 2016 and 2017 were noted for tracking the annual variation. The

spatial variation of CO₂ concentration was derived from AVIRIS-NG image in March 2016 procured for Howrah and Kolkata, regions within and adjacent to the above. The methodology of estimating CO₂ from AVIRIS-NG image is given below in brief.

Table 1: Regions of CO₂ assessment from NASA-Giovanni database

Location	Geographic Feature
27.3°N – 28.83°N 72.7°E – 74.3 °E	Barren Desert
26.2°N – 27.69 °N 70.4°E – 72.5 °E	Barren Desert
26.01°N – 27.5 °N 76.9°E – 78.8 °E	Crowded Urban area
26.26°N – 27.4 °N 80.9°E – 82.7 °E	Crowded Urban area
23.5°N – 24.52 °N 92.17°E – 93.27 °E	Dense Forest
25.4°N – 26.4 °N 92.6°E – 94.5 °E	Dense Forest
27.38°N – 28.52 °N 93.4°E – 95.38 °E	Dense Forest

The reflectance values of different features varied widely but the wavelength dependence for a certain feature did not change much over the CO₂ absorption bands. So the image pixels of differing reflectance for each narrow band were separated by unsupervised classification and the relative differences of the surface reflectance for the pixels of different classes were reduced by multiplying individual pixel values with suitable constants. The pixel values were normalized for path radiance by subtracting the radiance value of pure waterbody pixels having negligible reflectance at these wavelengths. The above algorithm was repeated for each non-absorbing and absorbing AVIRIS-NG band corresponding to CO₂ and O₂ determined from the radiance spectra and the CO₂ concentration was assessed by the usual differential optical absorption spectroscopic (DOAS) technique. Assuming oxygen as one-fifth of the air with a standard molecular density, the CO₂ columnar density was estimated.

3. Results and discussion

Figure 1 displays the variation of atmospheric CO₂ concentration for the period of January, 2016 to December, 2017 retrieved by OCO-2 using the 1.61 μm band for random points, as available, within the region of 22.5°–23.0° N and 86.0°–89.0° E. The data points fitted with Gaussian formula indicate a clear annual cyclic variation in the concentration with the maximum in local summer season. These are the average concentration values over the whole atmospheric column. The results may be compared with the spatial variation of CO₂ concentration shown in figure 2 derived from AVIRIS-NG image (March 2016). The ground level concentration, however, varied over a much wider range, e.g. 386–504 ppm (Raychaudhuri and Chaurasia 2017).

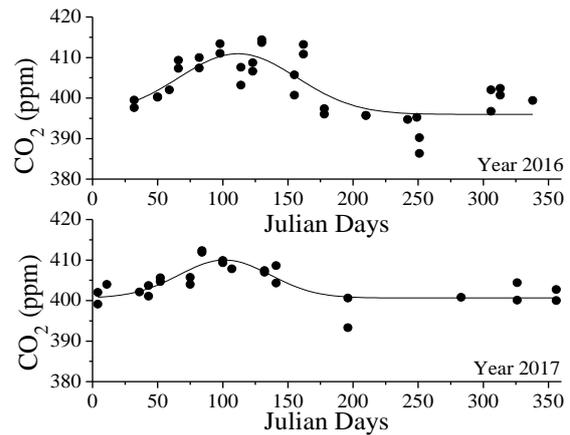


Figure 1: Variation of atmospheric CO₂ concentration within the region of 22.5°–23.0° N and 86.0°–89.0° E for the Julian days of years 2016 and 2017, as retrieved from 1.61 μm band of OCO-2.

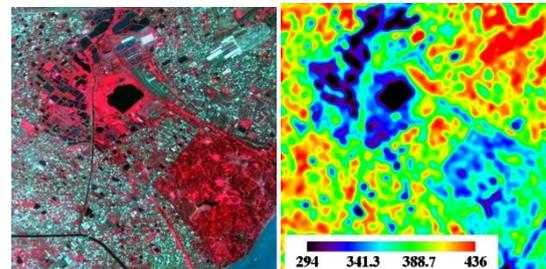


Figure 2: A sample of Spatial variation of CO₂ (ppm) (right panel) for different surface features in RGB image (left panel) derived from AVIRIS-NG image.

The above feature of annual cyclic variation of CO₂ concentration is more illustrated with the long-time variation obtained from Giovanni data. Figure 3 shows the variation of CO₂ concentration during 2010–2016 for individual area of desert (26.27°–27.63°N, 70.34°–72.34°E), urban area (26.01°–27.49°N, 76.95°–78.88°E) and forest (27.38°–28.53°N, 93.45°–95.38°E). Figure 4 presents the variation of CO₂ concentration during 2010–2016 for the east coast (8.68°–8.94°N, 77.76°–78.05°E) and the west coast (8.98°–9.82°N, 76.53°–77.32°E) regions of India.

Discussion on the above results makes it apparent that the CO₂ concentration has a general trend of steady increase with time superimposed by an annual cycle irrespective of geographic features. The annual periodicity becomes more prominent (Figure 3d), when the data are averaged over seven diversified regions of table 1. This explains why the global CO₂ change depicts a well-defined periodicity (Keeling et al., 1976). All the regions related to figure 1 and figure 3 are of tropical climate and exhibit maximum CO₂ flux in summer season. Also there is difference in the temporal steady increase. Figure 4 deals with the CO₂ concentration over the two coastal regions of the same Indian peninsula. Both the regions undergo steady CO₂ increase superimposed by its annual seasonal fluctuation. However, because of difference in geographic and climatic conditions, the data fitted with straight line result in prominent difference in the slopes indicating different rates of CO₂ increase for these two regions.

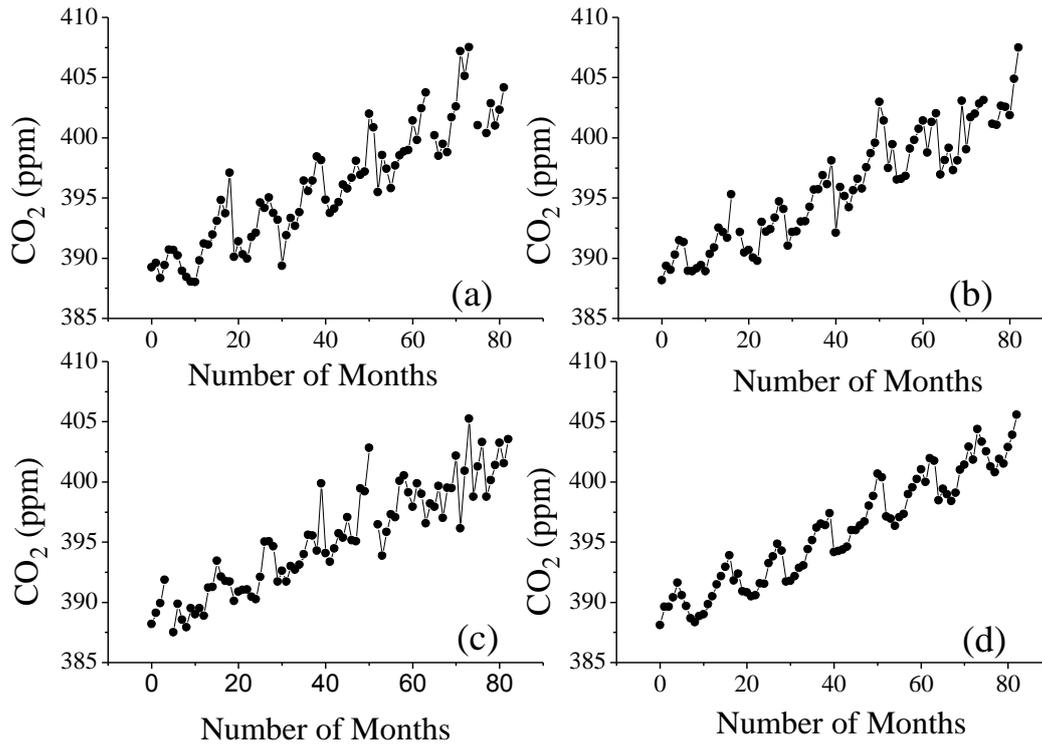


Figure 3: Annual variation of atmospheric CO₂ concentration during 2010-2016 for individual area of (a) desert, (b) urban area, (c) forest and (d) average of seven such regions given in table 1.

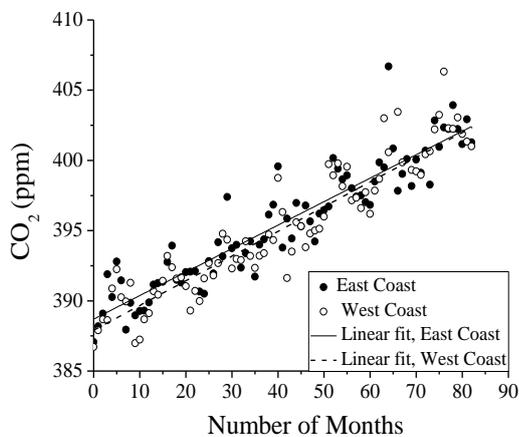


Figure 4: CO₂ concentration during 2010-2016 for the east and west coast regions of India

For the sake of subtler investigation, the linear increment part of the average CO₂ in figure 3d was eliminated and Fourier transform was carried out for the seasonal alteration, as shown in figure 5. It is noted that in addition to the prominent yearly periodicity caused by the regular annual climatic change, there exists a secondary semi-annual peak and also other weaker peaks. Thus the resultant temporal change in CO₂ flux is the combined effect of several factors, such as local climatic change, CO₂ uptake by vegetation and man- made changes.

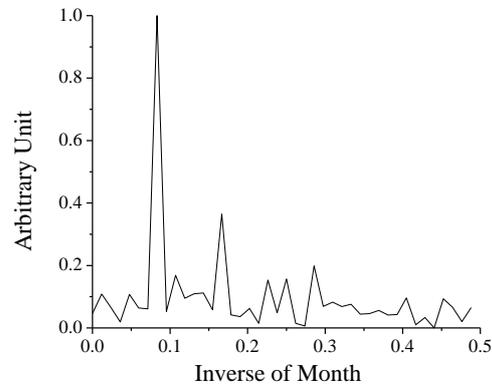


Figure 5: Fourier transform of the CO₂ seasonal periodic component obtained from figure 3d

4. Conclusions

The variation of atmospheric CO₂ concentration (ppm) during 2016-17 at around Kolkata was tracked using OCO-2 data archive and the extent of concentration was compared with that obtained from AVIRIS-NG image. An annual cyclic change in the CO₂ concentration was noted. In order to find the trend of such change during the last one decade, the monthly average values from Giovanni data archive were extracted for several places of contrast geographic features in India. A steady increase in CO₂ concentration was noted for all the places but the periodicity was not well defined. Local variation in the rate of increase was also noted. However, it is found that on

averaging the local values over several such places, the annual periodicity became more prominent as obtained in the global trend. Several periodicities were found to be present in the annual change of CO₂ flux from Fourier transform of the temporal change.

Acknowledgements

The authors thankfully acknowledge the financial support of Space Applications Centre, Indian Space Research Organization (vide AVIRIS-NG AO Project Ref. EPSA/4.20/2017 dt. 19.12.2017) and the infrastructural facilities extended by Presidency University. The authors thankfully acknowledge the use of OCO-2 and Giovanni data mentioned in the text.

References

- Basu, S., M. Krol, A. Butz, C. Clerbaux, Y. Sawa, T. Machida, H. Matsueda, C. Frankenberg, O. P. Hasekamp and I. Aben (2014). The seasonal variation of the CO₂ flux over Tropical Asia estimated from GOSAT, CONTRAIL, and IASI, *Geophysical Research Letters* 41, 1809-1815.
- Bovensmann, H., J.P. Burrows, M. Buchwitz, J. Frerick, S. Noel and V.V. Rozanov (1999). SCIAMACHY: Mission objectives and measurement modes, *Journal of the Atmospheric Sciences* 56, 127-150.
- Crisp, D et al., (2017). The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, *Atmospheric Measurement Techniques* 10, 59-81.
- Etheridge, D.M., L.P. Steele, R.L. Langenfelds, R.J. Francey, J.-M. Barnola and V.I. Morgan (1996). Natural and anthropogenic changes in atmospheric CO₂ over the last 1000 years from air in Antarctic ice and firn., *Journal of Geophysical Research* 101(D2), 4115-4128.
- Fu, Z., J. Dong, Y. Zhou, P. C. Stoy and S. Niu (2017). Long term trend and interannual variability of land carbon uptake – the attribution and processes, *Environmental Research Letters* 12, p. 014018.
- Hamazaki, T., Y. Kaneko and A. Kuze (2004). Carbon dioxide monitoring from the GOSAT satellite. *Proceedings of XXth ISPRS Conference* 3-5, Istanbul, Turkey, July 12-13, 2004.
- Jiang, X., D. Crisp, E.T. Olsen, S.S. Kulawik, C.E. Miller, T. S. Pagano, M. Liang and Y. L. Yung (2016). CO₂ annual and semiannual cycles from multiple satellite retrievals and models, *Earth and Space Science*, 3(2), 78-87.
- Keeling, C., R.B. Bacastow, A.E. Bainbridge, C.A. Ekdahl, Jr, P.R. Guenther and L.S. Waterman (1976). Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, *Tellus*, 28(6), 538-551.
- Leung, D.Y.C., G. Caramanna and M.M. Maroto-Valer (2014). An overview of current status of carbon dioxide capture and storage technologies, *Renewable and Sustainable Energy Reviews*, 39, 426-443.
- Li, Z et al., (2018). Non-uniform seasonal warming regulates vegetation greening and atmospheric CO₂ amplification over northern lands, *Environmental Research Letters*, 13, 124008.
- Liang, A., W. Gong, G. Han and C. Xiang (2017). Comparison of Satellite-Observed XCO₂ from GOSAT, OCO-2, and Ground-Based TCCON, *Remote Sensing*, 9, 1033-1059.
- Raychaudhuri, B. and S. Chaurasia (2017). Ground-based hyperspectral measurement at visible-infrared wavelengths for validation of the atmospheric carbon dioxide absorption wavebands in the context of AVIRIS-NG flight over Kolkata-Howrah, *Journal of Geomatics* 11(2), 167-173.
- Roman-Cascon, C., C. Yague, J. A. Arrillaga, M. Lothon, E. R. Pardyjak, F. Lohou, R. M. Inclin, M. Sastre, G. Maqueda, S. Derrien, Y. Meyerfeld, C. Hang, P. Campargue-Rodriguez and I. Turki (2019). Comparing mountain breezes and their impacts on CO₂ mixing ratios at three contrasting areas, *Atmospheric Research*, 221, 111-126.
- Wunch, D., G. C. Toon, J.-F. L. Blavier, R. A. Washenfelder, J. Notholt, B. J. Connor, D. W. T. Griffith, V. Sherlock and P. O. Wennberg (2011). The Total Carbon Column Observing Network, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1943), 2087-2112.
- Wunch, D et al., (2017). Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO₂ Measurements with TCCON, *Atmospheric Measurement Techniques* 10, 2209-2238.
- Xueref-Remy, I., E. Dieudonné, C. Vuillemin, M. Lopez, C. Lac, M. Schmidt, M. Delmotte, F. Chevallier, F. Ravetta, O. Perrussel and P. Ciais (2018). Diurnal, synoptic and seasonal variability of atmospheric CO₂ in the Paris megacity area, *Atmospheric Chemistry and Physics*, 18, 3335-3362.

Offshore wind farm site suitability and assessment along Tamil Nadu coast using RISAT-1 SAR and synergetic scatterometer data

Surisetty V. V. Arun Kumar^{1,*}, Avinash N. Parde², Jagdish^{1,3} and Raj Kumar¹

¹Space Applications Centre (ISRO), Ahmedabad, Gujarat, India

²Savitribai Phule Pune University, Pune, Maharashtra, India

³Department of Mathematics, Gujarat University, Ahmedabad, Gujarat, India

*Email: arunkumar@sac.isro.gov.in

(Received: Jan 02, 2019; in final form: Jun 06, 2019)

Abstract: The possibility of expanding the wind power production to the offshore regions is viable especially in the regions where winds tend to be higher and consistent. Out of all the coastal states of India, Tamil Nadu coastal region is having higher and persistent winds throughout the year. This study focusses on the utilisation of satellite remote sensing in identifying most suitable sites favourable for installing offshore wind farms. Four years (2012-16) of Radar Imaging Satellite (RISAT)-1 Synthetic Aperture Radar (SAR) and long-term synergetic scatterometer (Quick Scatterometer - QuikSCAT, Oceansat Scatterometer - OSCAT, Advanced Scatterometer -ASCAT-A and ASCAT-B) data for 16 years have been processed to generate mean wind speed maps at 80 m hub height. A Geospatial Information System (GIS) based methodology was developed to characterize the offshore wind power potential along the study region within the Exclusive Economic Zone (EEZ) by using ten site suitability parameters – mean wind speed, bathymetry, distance from the ports, electrical transmission lines, cyclone risk, seismic risk, avian exclusion, visual exclusion, ecosystem exclusion, potential shipping lane exclusion. After assigning suitable relative weightages to each parameter, the offshore wind power potential sites of Tamil Nadu (TN) coast have been evaluated, using heat maps with ranking score between 1 and 10, where 10 denotes the most suitable site. Based on the analysis, it has been observed that the offshore region between Kolachel and Tuticorn, TN has good potential for wind power generation. The overall electrical power generation in the most suitable sites have been estimated using Vestas 3MW and Gamesa 5MW offshore turbine curves. Similar methodology can be adopted for identifying potential sites of wind farms for the entire Indian coast.

Key words: offshore winds, site suitability, RISAT, scatterometer, India

1. Introduction

With the exhaustion of fossil fuels and increasing awareness on environmental issues, harnessing energy from available resources has become a prime focus in the developing countries (Yan et al., 2010). Among the renewable resources, wind power is a very large energy source, with proven commercial technology and very low or negligible CO₂ emissions (Archer and Jacobson, 2005). In comparison to land wind, offshore winds are strong and steady because of the absence of physical barriers (Xin, 2010). Currently, United Kingdom is the leading nation having installed offshore wind power capacity of 6836 MW, followed by Germany (5355 MW), China (2788 MW), Denmark (1271 MW), and Netherlands (1118 MW), totalling over 18,814 MW (Global Wind Energy Council, 2014). India has no offshore wind farms as on today.

Now the Government of the India is focusing to harness the offshore wind resources within the Indian Exclusive Economic Zone (EEZ) with the “National Offshore Wind Energy Policy 2015” by encouraging the development of offshore wind farms. The First Offshore Wind Project of India (FOWPI) with the capacity of 1 GW is initiated in the coastal waters off Gujarat. With the utilization of onshore and the upcoming more offshore wind farms, the government of India has set a target to achieve 60 GW power generation capacities by 2022 and reduction in carbon emissions by 33-35% of its Gross Domestic Product (GDP) by 2030. Hence, there is an immense requirement of assessing the suitable shelf areas along the coastal region. Indian coastal ecosystem comprising of mudflats, sandy beaches, estuaries, creeks, mangroves

coral reef, marshes, lagoon, sea grass beds and sandy and rocky beaches extended up to 42,000 sq. km (approx. 12 nm), so this region should be avoided for any developments (Kumar et al., 2006; Venkataraman and Wafar 2005). Therefore, finding suitable sites beyond of this region within the Exclusive Economic Zone (EEZ) remains a scientific and technical challenge for the engineers and investors.

It has been observed from the work done by Arun Kumar et al. (2016) that Tamil Nadu coastal region has higher and persistent wind speed all throughout the year. Hence, in this study we have attempted to utilise satellite remote sensing data to identify favorable sites for installing offshore wind farms using Geospatial Techniques (GIS).

2. Data and Methods

2.1 Satellite data

2.1.1 Scatterometer data

Scatterometers are the missions with specific goal of observing the equivalent-neutral surface ocean-vector winds across the world. Orbit-wise scatterometer wind products at 12.5 km spatial resolution from Quick Scatterometer - QuikSCAT (L2B version 3), Oceansat Scatterometer - OSCAT (L2B version 2) for the whole mission period 2000-2009 and 2010-2014 respectively and Advanced Scatterometer - ASCAT (L2B version 2 coastal winds) for 2012-2016 have been processed to generate long-term synoptic mean in the Indian region (Lat.0-25°N; Long. 65-95°E). The mission targeted wind vector precision and accuracy are, respectively, 3.0 m s⁻¹ for ASCAT, 0.5 m s⁻¹ for QuikSCAT (Vogelzang et al., 2011)

and OSCAT. A combined wind speed product generated from the four scatterometers have been extracted for the Tamil Nadu coastal region and used in this work. The method of combining the scatterometer data is described in Arun Kumar et al. (2016).

2.1.2 SAR data

Around 3000 Radar Imaging Satellite (RISAT)-1 images in HH polarization of Coarse Resolution ScanSAR (CRS) and Medium Resolution ScanSAR (MRS) modes with 36×36 m and 18×18 m spatial resolution respectively have been obtained from National Remote Sensing Centre (NRSC) over Indian region for the period 2012-2016. The SAR images are processed from L2 to 10 m wind speed using C-band Geophysical Model Function (GMF) CMOD-5N (Hersbach, 2010). The data has been averaged to around $1 \text{ km} \times 1 \text{ km}$ to reduce the inherent speckle noise, a random phenomenon in radar, and eliminate effects of tilt and modulation error from longer waves, before wind inversion. SAR-based ocean surface wind retrieval is described in Dagestad et al. (2013) and references therein. Space Applications Centre (SAC) developed a reprocessing chain to massively process RISAT-1 images. For HH data, the polarization ratio with both incidence angle and wind direction relative to the radar look angle is used. The wind directions were obtained from European Centre for Medium-range Weather Forecasts (ECMWF) atmospheric model at spatial grid of 0.125° . The data closest to the time of RISAT-1 pass are interpolated in space before wind inversion. Validation of RISAT-1 SAR derived wind speed is described in Jagdish et al. (2018). For wind power assessment, this volume of the data may be insufficient, but there were no studies so far to assess the validity of SAR data for wind resource assessment in India. An average wind speed dataset derived from the RISAT-1 SAR has been cropped to the study area for further analysis.

2.1.3 Wind computation methodology

Scatterometer and SAR derived wind data are available at 10 m above m.s.l. In order to estimate the speed at the hub height over the water, log-law applied by assuming the neutral stability of the atmosphere and an oceanic surface roughness $z_0 = 0.2 \text{ mm}$ (Manwell et al., 2010). The data were extrapolated to the required height ($h=80 \text{ m}$ in this case) using this logarithmic approach. The log-law states that a wind speed at a given hub height.

$$v_h = v_0 \left(\frac{\log(h/z_0)}{\log(10/z_0)} \right) \quad (1)$$

where, v_0 is the wind speed (ms^{-1}) at the 10 m above the sea level and v_h wind speed (ms^{-1}) at a hub height respectively. In this paper, two offshore wind turbines from different manufacturers: Vestas V112 3MW and Gamesa G128 5MW have been used to derive the electrical power production. The power curves of the two turbines are shown in figure 1.

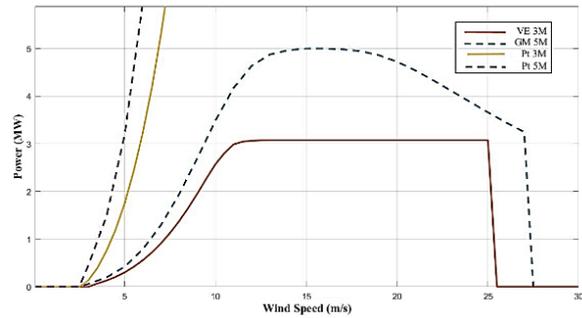


Figure 1: The Power characteristics curve for Vestas V112 offshore 3 MW (VE 3M) and Gamesa G128 5MW (GM 5M) Turbines

The individual effective footprint (i.e. Array spacing) on the seabed was calculated using the equation (Sheridan et al., 2012).

$$\text{Array spacing} = D^2 \times v_{df} \times v_{cf} \quad (2)$$

where, D is the rotor diameter (m); v_{df} , v_{cf} are downwind and crosswind spacing factor 10 and 5 respectively. Thus, the corresponding array spacing for the VM 3M and GM 5M turbines are 0.63 and 0.82 km^2 respectively. The number of turbines commissioned within the area was determined using

$$\text{Number of turbines} = \frac{\text{Total area available}}{\text{Array spacing}} \quad (3)$$

The nameplate wind power capacity (i.e. total installed capacity) was computed by multiplying the rated power of single turbine with the number of installed turbines. However, this is not achievable for any turbine due to several losses such as availability, schedule maintenance, unplanned shutdown, wake effect, turbine efficiency. Hence, the practical power estimates were also computed using the manufacturer's power curve. In decision-making contexts, it is often necessary to compare renewable energy resources such as offshore wind to traditional generation sources (i.e. coal, nuclear, natural gas). In such a comparison, it is more useful to use the 'average output' than the installed nameplate capacity (Sheridan et al., 2012).

2.2 Site suitability criteria

Ten preliminary site suitability criteria viz., mean wind speed (potential), bathymetry, ports, electric transmission lines, cyclone risk, seismic risk, avian exclusion, visual exclusion, ecosystem exclusion, potential shipping lane exclusion have been calculated using data from different sources. Arc GIS (version 10.4) software has been used to create an "Offshore Wind Farm Suitability model". Different site suitability criteria were combined by assigning an appropriate weightage and computed the overall suitability map following (FOWIND, 2015). The detailed modelling steps are shown in the self-explanatory flowchart shown in figure 2. The final suitability scores in the study area were determined by reclassifying the scores derived from the weighted overlay.

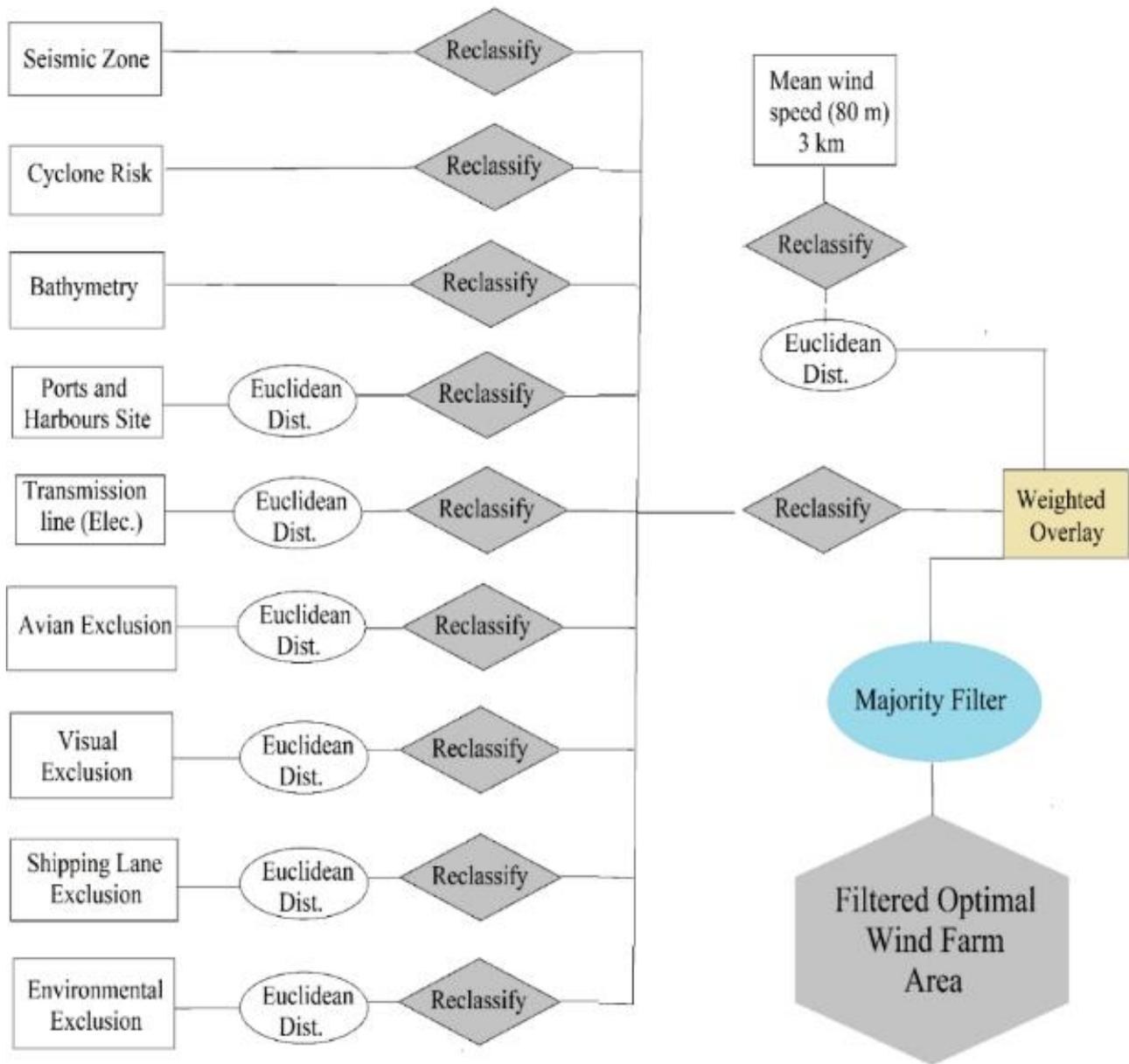


Figure 2: Flowchart for modelling suitability of offshore wind farm development

3. Results and discussion

The data from various resources have been processed in ArcMap and created rasters of 1 × 1 km spatial grid to maintain uniformity. Entire analysis is confined to the Tamil Nadu coastal region between the longitudes 76°E – 81°E and 7°N – 14°N within the EEZ. A suitable score from 1-10 have been assigned with the 10 for the most suitable criteria (Table 1). Following are the details of each parameter used in the study.

a) Potential Wind Speed at hub height (80 m)

The average mean wind speed distribution at 80 m hub height derived from RISAT-1 SAR (Figure 3a) and

synergetic combined scatterometer data (Figure 3b) within the study area has been used separately for the wind farm site suitability analysis. It has been observed that the wind speed is high in the offshore waters between Tuticorn and Kolachel in the south Tamil Nadu. The spatial variation in the wind is high in the SAR data as compared to the scatterometer data. It is because few SAR images available for temporal averaging.

Whereas in the case of scatterometer data, huge volume of each pass data was processed, which makes the synergetic scatterometer data reliable. The benefit of the SAR data is the high resolution (1 km grid) as compared to the scatterometer data (12.5 km).

Table 1: Weighting of spatial influences

Parameter	Criteria	Description	Relative Weightage (%)	Ranking Score	
				Class	Score
Potential Wind Speed at hub height	> 6 m/s	Best practices for economic sustainability	40	<5 m/s – 6.0 m/s	1-3
				6.0 – 7.0 m/s	4-5
				7.0 - 8.0m/s	6-7
				8.0 – 9.5 m/s	8-10
Bathymetry	> -5 m and < -500 m	Best practices for economic sustainability and technology	30	> -200 m	1-3
				-200 - -100 m	4-5
				-100 - -50 m	6-7
				-50 - - 5 m	8-10
Port and Harbours consideration	Nearest	Best practices for economic	8	> 300 km	1-3
				200 – 300 km	4-5
				100 – 200 km	6-7
				0 – 100 km	8-10
Cyclone Risks	Within EEZ	Best practices to minimize impacts	1	0.0205 – 0.0122	1-3
				0.0122-0.0068	4-5
				0.0068-0.0028	6-7
				0.0028-0	8-10
Electric Transmission Line	Nearest	Best practices for economic	10	> 800 km	1-3
				300 – 800 km	4-5
				100 – 300 km	6-7
				0-100 km	8-10
Seismic Risk	Within EEZ	Best practices to minimize impacts	1	65 – 530	1-3
				17 - 65	4-5
				3 – 17	6-7
				0 – 3	8-10
Avian Exclusions	< 2 km	Best practices to minimize impacts	1	0 – 5 km	1-3
				5 – 10 km	4-5
				10 – 100 km	6-7
				> 100 km	8-10
Visual Exclusions	< 10 km	Best practices to minimize impacts	2	0 – 50 km	1-3
				50 – 100 km	4-5
				100 – 200 km	6-7
				> 200 km	8-10
Potential Shipping lanes Exclusions	< 1 km	Best practices to minimize impacts	2	0 – 5 km	1-3
				5 – 10 km	4-5
				10 – 50 km	6-7
				> 50 km	8-10
Ecosystem Site Exclusion	< 10 km	Best practices to minimize impacts	5	0 – 50 km	1-3
				50 – 100 km	4-5
				100 – 200 km	6-7
				> 200 km	8-10
Total			100		

b) Bathymetry

It is very important to determine water depth at any potential offshore wind development site for appropriate foundation technology. On this basis, we can choose the foundation technology and estimate installation cost. In this study, bathymetric data for the Indian EEZ (ocean area up to 200 nm from the baseline) was obtained from the General Bathymetric Chart of the Oceans (GEBCO) (Becker et al., 2009) with a resolution of one arc-minute. Entire bathymetric data was processed and classified into four-water depth ranges see in figure 4. Later the bathymetry data was reclassified from 1-10 considering

the depths between -5 and -500 m only. The score with 1 being most unsuitable and with 10 being most suitable for the wind farm sites. It has been observed that the study region is having steep continental shelf within the EEZ region. Generally, wide and shallow continental shelf is preferred for the wind farms. In addition, the EEZ is very narrow of ~ 20 nautical miles with the presence of international waters sharing with Sri Lanka region. Very shallow waters are existing surrounding Gulf of Mannar up to Nagapattinam coast. A gradual change in the depth is observed southwards from Kolachel.

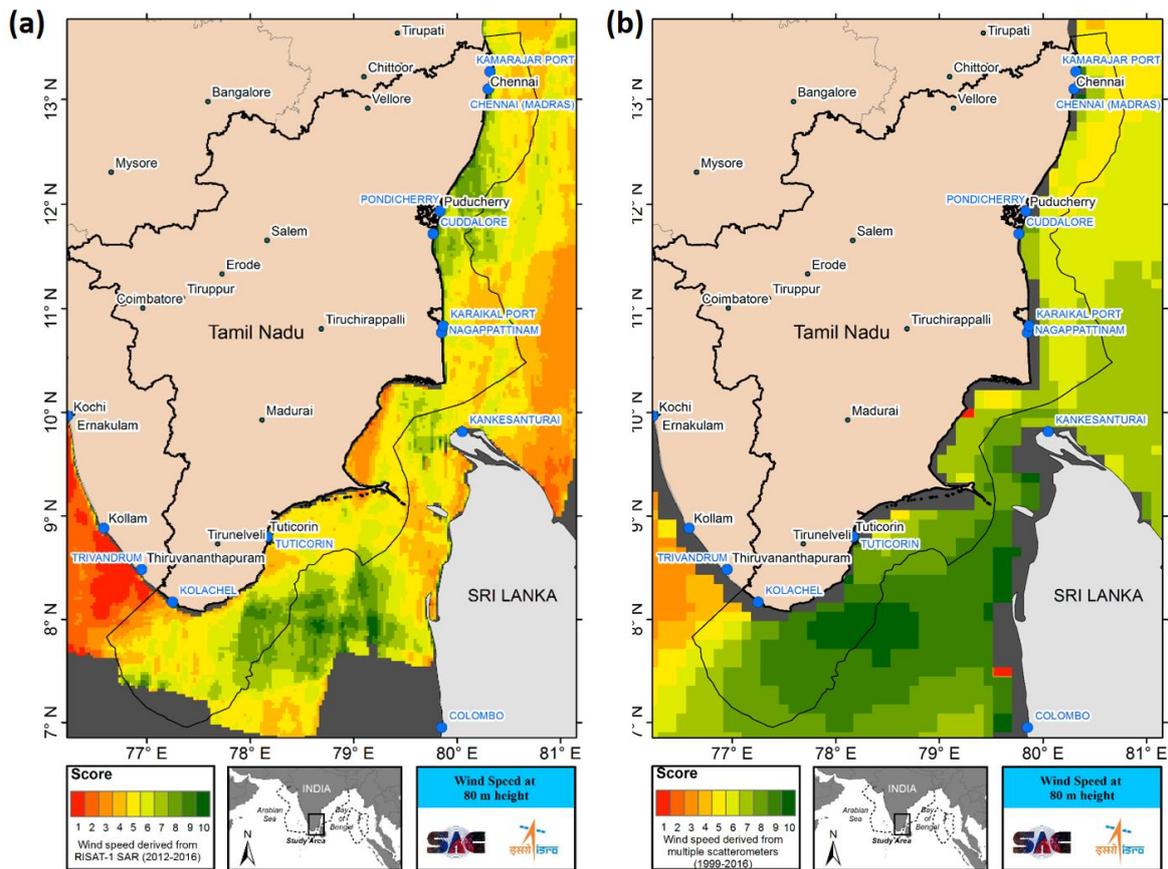


Figure 3: Offshore Wind Speed Heat Map of Tamil Nadu derived from (a) RISAT-1 SAR and (b) multiple scatterometers.

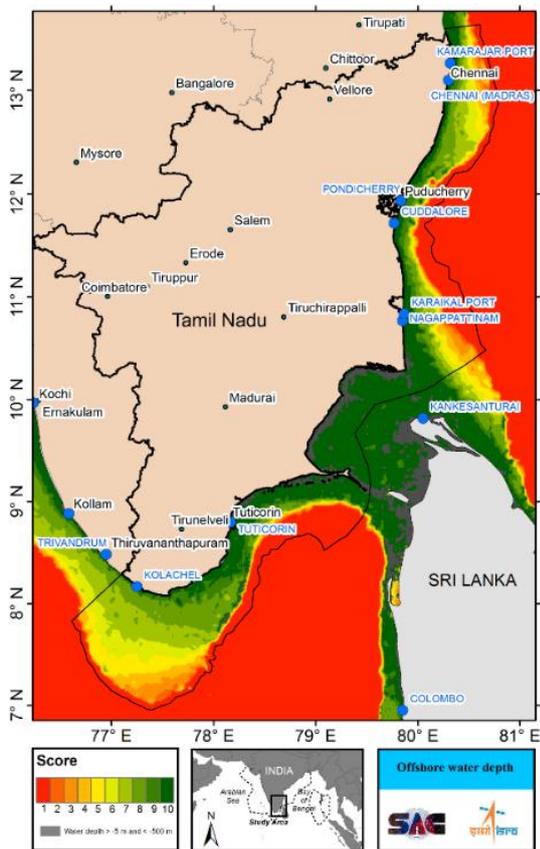


Figure 4: Offshore Water Depth Heat Map

c) Port and Harbours consideration

The port is one of the most important components in offshore wind farm construction. Ports handle manufacturing, storage and transportation of wind farm components. In general, all manufacturing facilities would be located on the coast, within the port closest to the offshore wind farm. Overall, this parameter having a goal to reduce the logistic costs. In this study, major port data (in the form of shape file) has been obtained from the World Port Index database. As the distance from the port increases, the region would become unfavourable for the wind farms (Figure 5). Although there are major and minor ports existing along the Tamil Nadu coast, the minor ports were not considered in this study due to unavailability of the data.

d) Cyclone Risks

Consideration of tropical cyclones in the design of offshore wind turbine farm is very important. In order to understand the impact of cyclones, the cyclone track data for the period of 1900 – 2016 has been obtained from International Best Track Archive for Climate Stewardship IBTrACS (source: <https://www.ncdc.noaa.gov/ibtracs>). A density map of cyclones over North Indian Ocean has been generated (Figure 6). There were few number of cyclones crossed the region. Especially, the southern Tamil Nadu coast is rarely affected by any cyclone during the past 116 years' period. Hence, that region is suitable in terms of cyclonic activity.

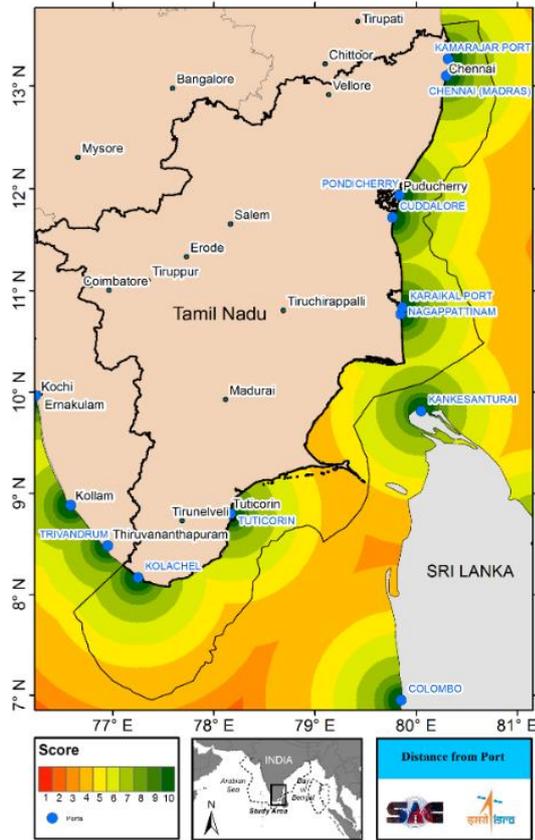


Figure 5: Offshore ports Heat Map

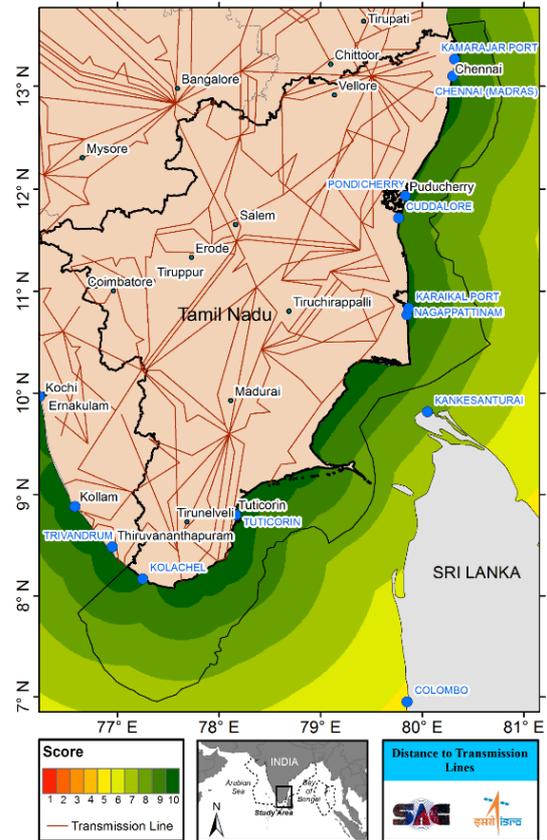


Figure 7: Offshore electric transmission line heat map

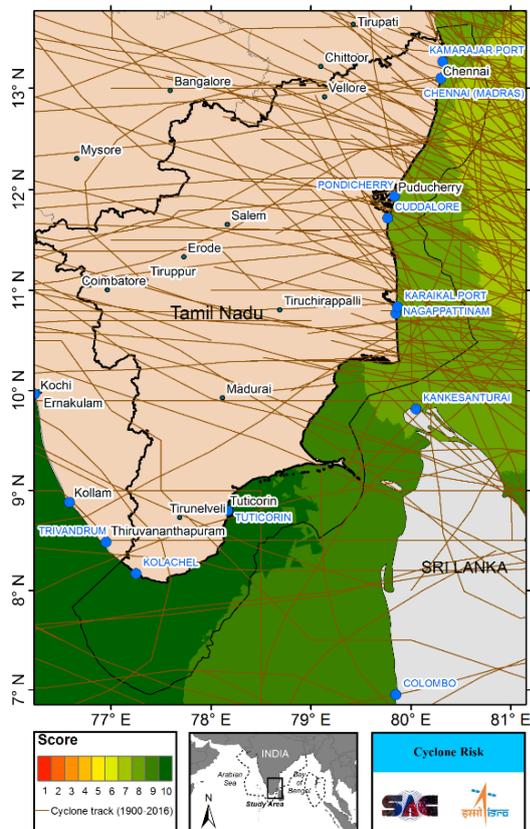


Figure 6: Offshore cyclone risk map

e) *Electric Transmission Line:*

In order to transmit power from the offshore wind farm to the onshore grid system, a dedicated electrical transmission network is required. The wind farm situated far away from the onshore electric transmission lines/grid is less economically viable. Hence, generally wind farm installations are preferred as near as possible to the electrical grids. The data on electrical transmission lines obtained from the Central Electric Authority (CEA) has been processed to estimate the feasibility. It has been observed that the regions closer to the coast are favourable with good score (Figure 7). Submarine electrical transmission lines are preferred for the offshore wind farms. However, lack of data availability restricted the present analysis to the onshore electric transmission lines.

f) *Seismic Zone Exclusion:*

Earthquake is a design concern in seismic areas such as East Asia and Western United states. Global seismic data containing earthquake epi centres have been downloaded (source: <http://earthquake.usgs.gov>) and prepared a detailed seismic risk map for the study region. However, to date there were no past records of seismic events in the coastal waters of Tamil Nadu. Hence, the entire coastal region is suitable with respect to the seismic hazard is concerned (Figure 8).

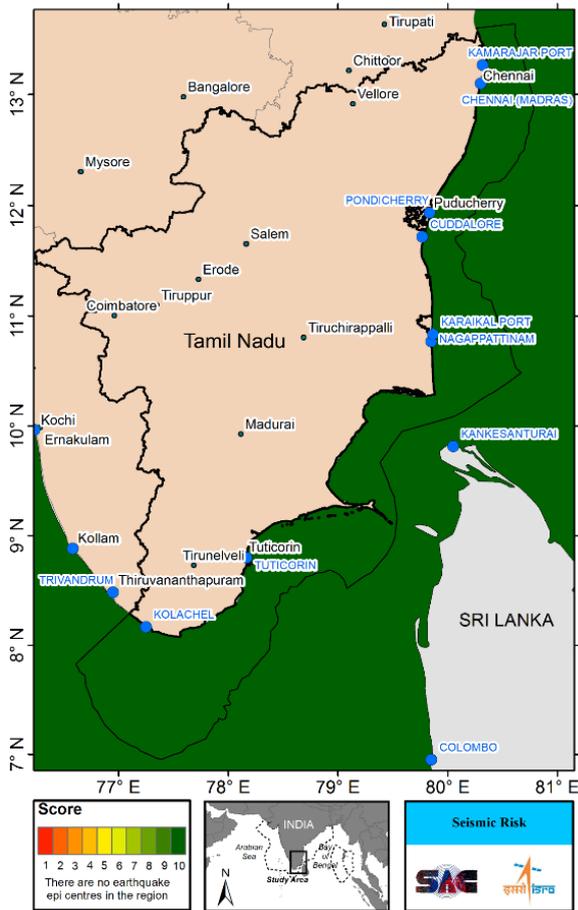


Figure 8: Offshore seismic risk Heat Map

g) *Avian Exclusions:*

It is very important to study the effects of the offshore wind farms (before construction and during operation) on the Environment. The Potential impacts considering avian mortality is to be included in the study for best practices. In this study, we have considered potential avian impacts within 1 nm from the coast (Figure 9) by excluding the zone parallel to the coastline to avoid interference with the migratory birds (Sheridan et al., 2012). Euclidean distance was computed from the coastline and excluded the region within 1 nm from the analysis (score = 1).

h) *Visual Exclusions:*

Visual impacts could potentially decrease tourism revenue if people choose to stop visiting a beach when turbines are visible. By considering impact on the coastal residents and tourist sites, offshore wind farms are to be implemented over the several nautical miles away from the shore. Some studies found that even coastal residents, the group with the highest valuation of uninterrupted views, reported very little additional willingness-to-pay for moving turbines beyond 9 miles (8 nautical miles) offshore (Sheridan et al., 2012). In this paper, the visual exclusion of 8 nm away from the coast has been considered (Figure 10).

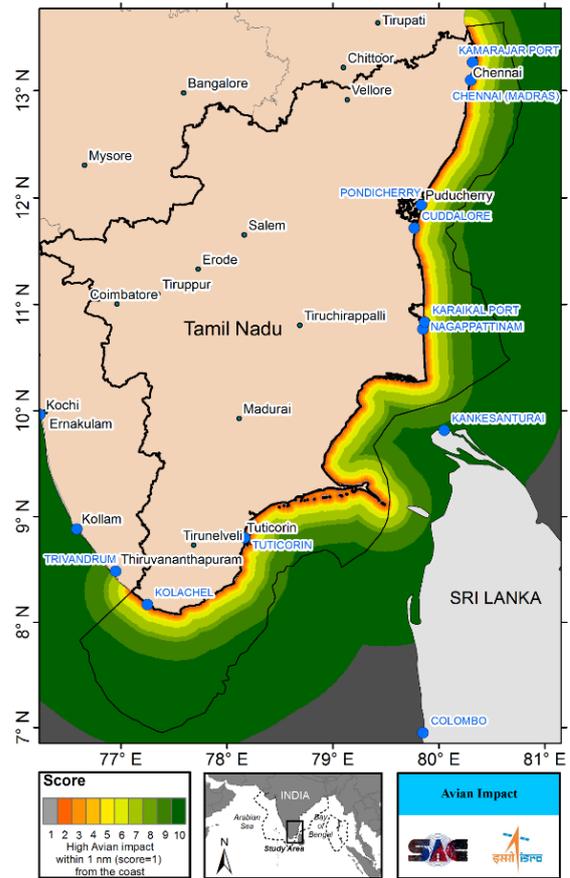


Figure 9: Offshore avian exclusion Heat Map

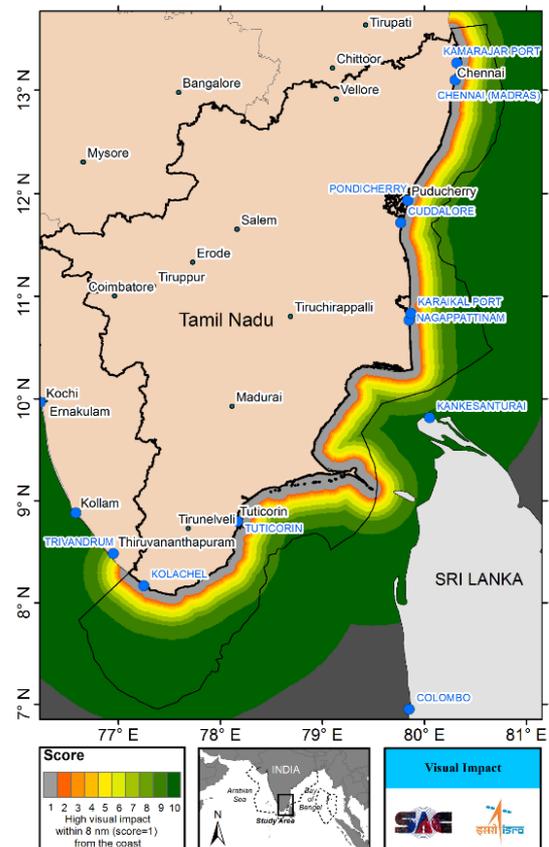


Figure 10: Offshore visual impact Heat Map

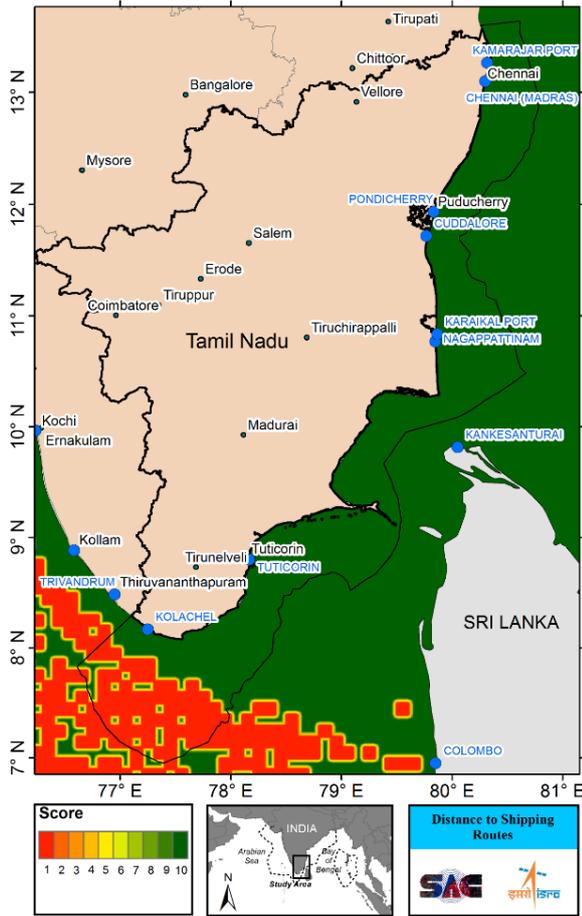


Figure 11: Offshore shipping lane exclusion Heat Map

i) Potential Shipping lanes Exclusions:

In order to understand the potential shipping traffic in the ocean, the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) dataset has been considered. ICOADS is a data set of global marine surface conditions and locations collected and reported by a fleet of about 3700 ships known as the Voluntary Observing Ships (VOS) which shown in the $0.1^0 \times 0.1^0$ grids size. In this paper, a setback of approximately 1 km away from the shipping lanes has been considered in order to avoid collision with the ships plying around that area. It has been observed that major ship traffic is existing along the southern region in relatively deeper waters (Figure 11). The planned *Sethusamudram shipping canal project* may also affect the wind farm development along Rameswaram region, which needs to be further studied and considered.

j) Ecosystem Site Exclusion:

Tamil Nadu coastline contains an abundant variety of ecosystems for example beach, mangrove, coral reef, estuary, island, coastal wetland, etc. Erecting offshore wind turbines will affect their habitats. Therefore, LULC (Land Use Land Cover) data (source: Coastal Zone Information System CZIS, SAC) has been used for preparing a general coastal sensitive ecosystem maps (corals and mangroves) (Figure 12). As the wind turbine foundations are hard structures, it is suggested to exclude these areas by considering marine protection reasons.

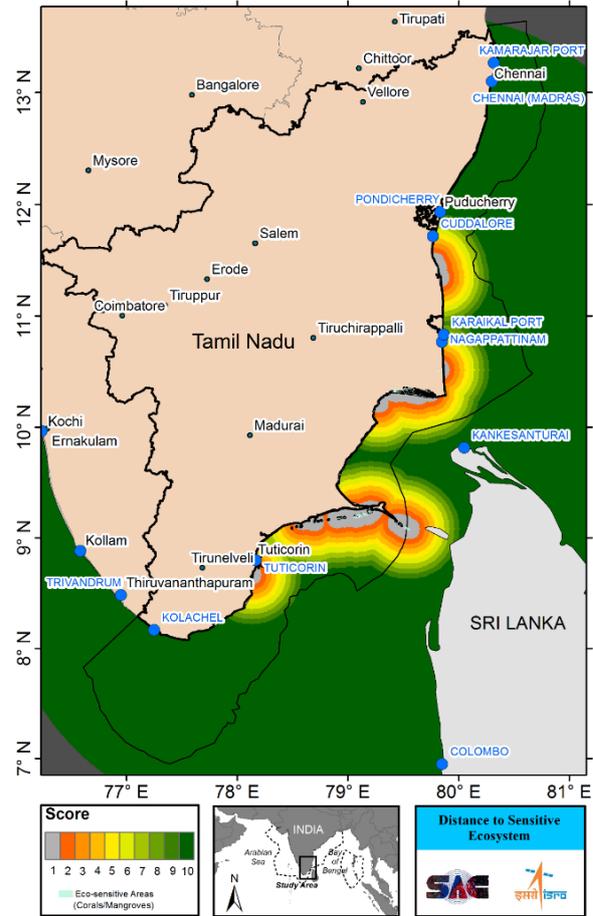


Figure 12: Offshore ecosystem exclusion Heat Map

In order to aid the selection of zones a scoring mechanism has been derived which takes into account the key technical and consenting factors considered.

Based on the site suitability criteria and wind potential, the most feasible areas suitable for wind farm installation in the study is obtained. The available shelf area, number of installable turbines, name plate capacity and the total expected power generation from the VE 3M and GM 5M turbines are given in table 2 and table 3 respectively. Due to variations in the wind speed inputs, the statistics and optimal sites are found different for SAR and scatterometer based winds (Figure 13). It has been observed that the study region is having more shelf area between 0 and 35 m. Around 1500 – 2000 VE 3M turbines can be installed in < 35 m depth, whereas around 500 – 1990 GM 5M turbines can only be installed in the same shelf region due to its vast turbine area. However, more power can be generated by 5 MW turbines as compared to the 3 MW turbines. On an average, around 1300 kW of power can be generated using the former turbines and with the latter turbines, around 1900 kW is possible. An overall output of 1300 MW or 1900 MW is possible from both the turbines in the optimal region, which converts in to an average annual production of 11966 GWh/year (VE 3M) or 16907 GWh/year (GM 5M) for 0 – 35 m depth range obtained from scatterometer data. The wind production at other two depths are shown in the tables.

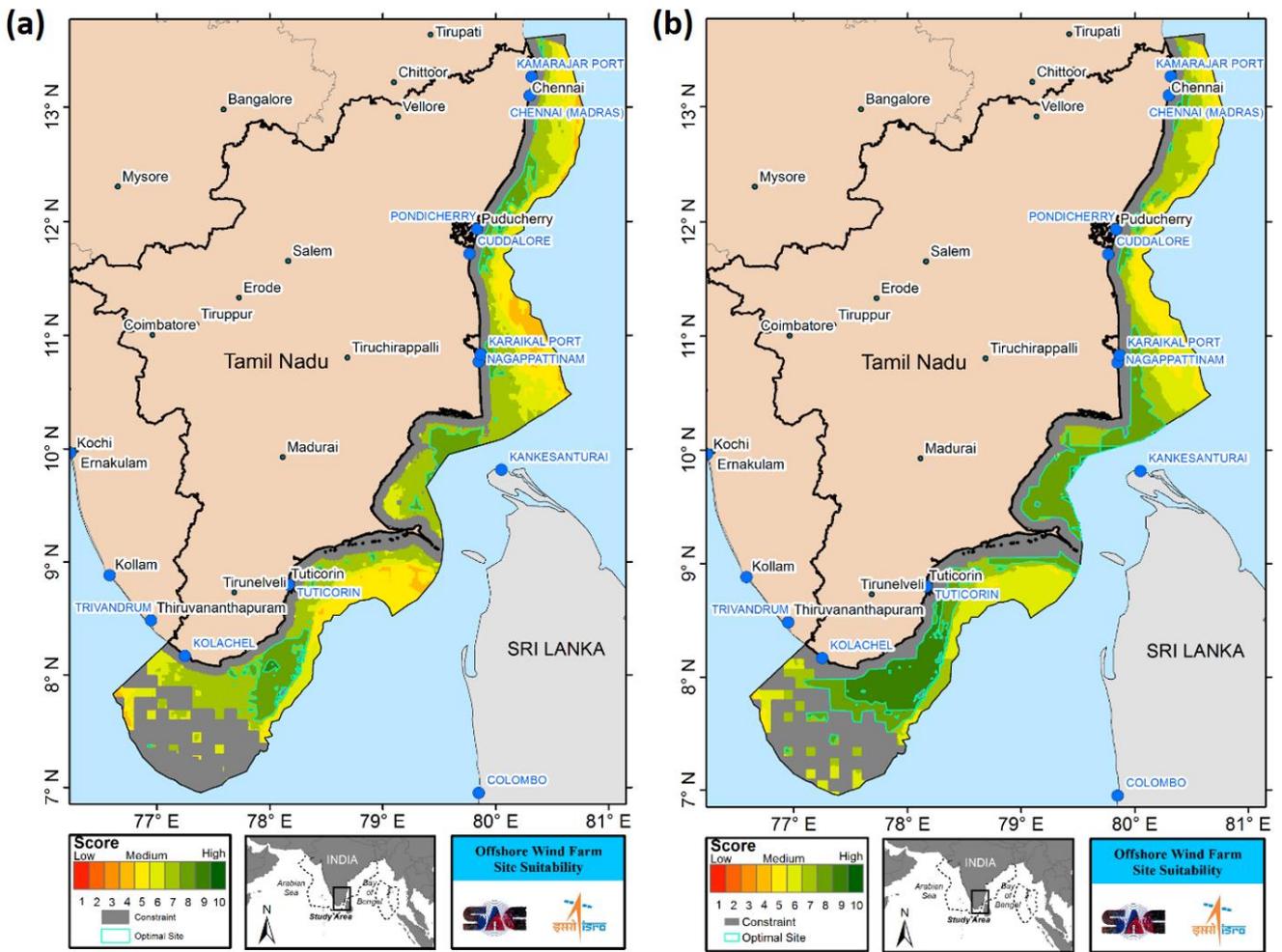


Figure 13: Offshore heat map with potential development zones of Tamil Nadu derived from (a) RISAT-1 SAR data and (b) synergetic scatterometer data

Table 2: Statistics of wind power potential obtained from VE 3M turbine derived from SAR and scatterometer

Parameters	Satellite	Monopile	Jacket	Advanced Jacket	Total
		0-35 m	35-50 m	50-100 m	
Available Shelf Area (km ²)	SAR	425	173	55	653
	SCAT	981	286	202	1469
Number of Turbines	SAR	675	275	87	1037
	SCAT	1557	454	321	2332
Name Plate capacity (MW)	SAR	2025	825	261	3111
	SCAT	4671	1362	962	6995
Range (kW)	SAR	771-1893	847-1753	1017-1725	-
	SCAT	1058-1817	1152-1817	1137-1808	
Mean (kW)	SAR	1187	1364	1416	3965
	SCAT	1393	1588	1533	4514
Total output (MW)	SAR	505	236	78	819
	SCAT	1366	454	309	2129
Total Average Power generation (GWh/year) (EPG)	SAR	4424	2067	683	7174
	SCAT	11966	3977	2707	18650

Table 3: Statistics of wind power potential obtained from GM 5M turbine derived from SAR and scatterometer

Parameters	Satellite	Monopile	Jacket	Advanced Jacket	Total
		0-35 m	35-50 m	50-100 m	
Available Shelf Area (km ²)	SAR	425	173	55	653
	SCAT	981	286	202	1469
Number of Turbines	SAR	518	211	67	796
	SCAT	1196	349	246	1791
Name Plate capacity (MW)	SAR	2590	1055	335	3980
	SCAT	5982	1744	1232	8957
Range (kW)	SAR	1185-2707	1306-2509	1557-2483	-
	SCAT	1495-2569	1629-2569	1608-2556	-
Mean (kW)	SAR	1733	1979	2043	5755
	SCAT	1967	2245	2166	6378
Total output (MW)	SAR	737	342	112	1191
	SCAT	1930	642	437	3009
Total Average Power generation (GWh/year) (EPG)	SAR	6456	2996	981	10433
	SCAT	16907	5624	3828	26359

4. Conclusions

The study provides detailed analysis of potential offshore wind power in the Tamil Nadu EEZ using RISAT-1 SAR and synergetic scatterometer data. Geospatial suitability model has been developed using ten important criteria for the suitable site identification. The southern parts of Tamil Nadu have been identified as one of the potential regions for installing offshore wind farms. In particular, the regions surrounding Tuticorn and Kolachel exhibited the highest suitability. As India is rapidly growing in the economy and population leading to a continuous increase in electricity demand, the offshore wind farms could suffice an excellent alternative energy source because of its negligible greenhouse gas emissions, utilization of offshore surface, cost competitiveness, price stability, and energy security. The methodology can be adopted for identifying potential sites of wind energy for the entire Indian coast. The analysis of production cost and the actual availability of a selected area can be further studied for detailed site suitability.

Acknowledgments

The authors are thankful to the Director, Space Applications Centre (ISRO) for his constant support and encouragement. The first author is thankful to Dr A.S. Rajawat, Group Director (GHCAG), Shri. Arun Kumar Sharma, Head (GSD) and Dr. Rashmi Sharma, Head (OSD) for their suggestions and encouragement. The authors are grateful to all the data providers. This work is a part of SAMUDRA TDP R&D project at SAC and the second author has carried out some part of this study during his SMART training at SAC. They are equally thankful to Dr. Sathiyamoorthy, Head (MRTD) and Dr. Satya Prakash Ojha for the support.

References

- Archer, C. L. and M. Z. Jacobson (2005). Evaluation of global wind power, *Journal of Geophysical Research: Atmospheres*, 110, D12110.
- Becker, J. J., D. T. Sandwell, W. H. Smith, J. Braud, B. Binder, J. Depner, D. Fabre, J. Factor, S. Ingalls, S. Kim and others (2009). Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30PLUS, *Marine Geodesy*, 32, 355-371.
- Dagestad, K.-F., J. Horstmann, A. Mouche, W. Perrie, H. Shen, B. Zhang, X. Li, F. Monaldo, W. Pichel, S. Lehner, and others (2013). Wind retrieval from synthetic aperture radar-an overview. 4th SAR Oceanography Workshop (SEASAR 2012): *Advances in SAR Oceanography*.
- FOWIND (2015). Pre-feasibility study for offshore wind farm development in Tamil Nadu. Tech. rep., Global Wind Energy Council (GWEC).
- Global Wind Energy Council, G. W. (2014). Global Wind Report: Annual market update 2015. URL <http://gwec.net/global-figures/graphs/>. [Accessed October 15, 2018].
- Hersbach, H. (2010). Comparison of C-band scatterometer CMOD5. N equivalent neutral winds with ECMWF, *Journal of Atmospheric and Oceanic Technology*, 27, 721-736.
- Jagdish, Arun Kumar S.V.V., A. Chakraborty and Raj Kumar (2018). Validation of wind speed retrieval from RISAT-1 SAR images of the North Indian Ocean, *Remote Sensing Letters*, 9, 421-428.

- Kumar, V. S., K.C. Pathak, P. Pednekar, N.S.N. Raju, and R. Gowthaman (2006). Coastal processes along the Indian coastline, *Current science*, 91(4), 530-536.
- Manwell, J. F., J.G. McGowan and A.L. Rogers (2010). *Wind energy explained: theory, design and application*. John Wiley and Sons.
- Sheridan, B., S.D. Baker, N.S. Pearre, J. Firestone and W. Kempton (2012). Calculating the offshore wind power resource: Robust assessment methods applied to the US Atlantic Coast. *Renewable Energy*, 43, 224-233.
- Arun Kumar, S.V.V., Jagdish and Raj Kumar (2016). Evaluation of offshore wind energy resources for power generation based on scatterometer and SAR data along the Indian coast. *Remote Sensing of the Oceans and Inland Waters: Techniques, Applications, and Challenges*, 9878, pp. 98780P.
- Venkataraman, K., and M.V. Wafar (2005). *Coastal and marine biodiversity of India*. NISCAIR, CSIR.
- Vogelzang, J., A. Stoffelen, A. Verhoef and J. Figa-Saldaña (2011). On the quality of high-resolution scatterometer winds, *Journal of Geophysical Research: Oceans*, 116.
- Xin, H.-L. (2010). Aspect on the development of offshore wind energy in China. *Periodical of Ocean University of China*, 40, 147-152.
- Yan, Q., Y.-C Chen, A.-J. Wang, G.-S., W.-J. Wang Yu and Q.-S. Chen (2010). Development obstacles of new energies in China and countermeasures: A review on global current situation. *Diqiu Xuebao (Acta Geoscientica Sinica)*, 31, 759-767.

Seasonal variability of rip current probability along a wave-dominated coast using high resolution satellites and wave data

T. Sridevi*, Surisetty V. V. Arun Kumar and Raj Kumar

Earth, Ocean, Atmosphere, Planetary Sciences and Applications Area (EPSA) Space Applications Centre (ISRO),
Ahmedabad, Gujarat, India

*Email: sirimeteor@gmail.com

(Received: Jan 04, 2019; in final form: Jun 06, 2019)

Abstract: Rip currents at RK Beach, Visakhapatnam in the northern coastal Andhra Pradesh are causing havoc to the beach users. Recent studies revealed that rip currents are responsible for most drownings happened along this beach. This creates an interest in the public and raises a research challenge on what factors favorable for such dangerous currents. It has been well established that waves and beach morphology play an important role in the formation mechanism of the rip currents. However, due to limited studies, their variability with the season is less understood. In this paper, we have utilized high-resolution satellite imageries and wave data from the nearby buoy to establish a relationship between beach stage and wave parameters. RK beach is a wave-dominated coast and receives persistent swells during south-west and post monsoon seasons. The strong swells during monsoon season and extreme events induce changes in the beach morphology and thereby changes the beach stage from dissipative type to reflective type followed by an intermediate stage. High-resolution optical satellite imageries have been processed using ArcMap and cropped to the study region. Beach types have been identified in few satellite imageries to understand the beach response to the prevailing waves for the period 2013-2017. The study revealed that the rip current probability is higher during intermediate beach stage, and is responsible for major drowning cases along the study area.

Key words: rip currents, beach stage, RK Beach, satellite data, buoy data, GIS

1. Introduction

Many beaches worldwide are characterised by the presence of narrow and seaward flowing rip currents that extend from the shoreline, through the surfzone and sometimes even few kilometres away. Rip currents are fundamentally generated due to the action of breaking waves owing to alongshore gradients in wave-induced radiation stresses and pressure (Bowen, 1969). The offshore flow velocities of rip currents range between 0.3 - 2 m/s (MacMahan et al., 2004). Rip currents can occur along any coastline due to the interaction between waves, currents, water level and near shore bathymetry and these currents flow faster near the water surface (Haas and Svendsen, 2002). Rip currents can quickly carry the surfzone bathers to the offshore with all swimming abilities (Drozdowski et al., 2015).

Rip currents are one of the most common causes for drownings at the beach in south India. On average, 39 people per year are drowning due to rip currents in the Indian beaches (Arun Kumar and Prasad, 2014). Most of the drowning cases are under reported due to lack of knowledge and popularity in the country. It was estimated that around 320 beach users drowned due to rip currents in the Visakhapatnam beaches during the 2000-2010 (Arun Kumar and Prasad, 2014). Out of all beaches, Rama Krishna (R.K.) beach is declared as the most dangerous for having more than 159 drowning deaths. In Visakhapatnam, more drowning cases due to rip currents were reported during August and October months. The causative factors that drive these dangerous currents are less understood due to manifold complexity in the processes. Waves, tides and nearshore bathymetry are the three important factors seem to be responsible for the majority of the cases. In general, rip current velocities vary with the tide with greater velocities found at low tide and

lower velocities found at high tide in response to changes in wave energy dissipation (Aagaard et al., 1997; Brander, 1999; Brander and Short, 2000). It is believed that heavy surf as a consequence of storms will lead to erosional rips (Short, 1985), which develop with a quasi-regular spacing as a consequence of flow and morphological instabilities (Turner et al., 2007). Consequently, rip current velocities increase in response to increasing wave heights and decreasing tidal elevations (MacMahan et al., 2005). In addition to the erosion rips, accretion and topographically controlled rip current systems exist in nature. Topographical rip currents are controlled by the local geology and shelf bathymetry that controls the refraction patterns and wave focussing.

In this study, we are attempting to understand which dominant beach type is responsible for drownings at RK Beach using in-situ wave data and satellite remote sensing.

2. Study Area

The port city of Visakhapatnam (17° 41' N and 83° 17' E) is located almost midway between Chennai and Kolkata on the east coast of India. The valley city of Visakhapatnam is surrounded on three sides by rocky hills and on the other side, the shore of Bay of Bengal. The climate in this region is mainly controlled by the Indian monsoons. The annual average rainfall is about 975 mm. The wind blows from east and southeast during January – March, changes to south and south-west during October – December. The waves approach the coast from the southeast during monsoon months (March to September) and change to east through southeast during other months (Kumar et al., 2006). It is a wave-dominated coast with wave heights ranges from 0.5 – 2.0 m; higher during monsoon and cyclonic conditions and lower during other seasons.



Figure 1: Location map and study area (red circle denotes the position of wave rider buoy)

The wave-induced longshore currents along this coast are generally towards northeast from March to September and towards southwest from November to February with magnitudes, in general, vary between 0.02 to 0.3 m/s (Kumar et al., 2006). The estimated annual net littoral drift along this coast is about 0.5 - 0.7 million cubic meters towards north-easterly direction (Chandramohan and Liang, 1985; Kumar et al., 2001; Panigrahi et al., 2010). The tides in this region are semi-diurnal with a mean spring tidal range of 1.43 m and the neap tidal range of

0.54 m (Kumar et al., 2001). Depressions and Deep depressions occur frequently during SW monsoon season with maximum occurrence during August. Severe cyclones occur in the region during pre-monsoon months (April – May) and post-monsoon months (October – November).

3. Data and methods

3.1 Satellite data

Rip currents are very small-scale features and therefore their features or signals cannot be observed in coarser resolution satellite (LISS-III, LANDSAT etc.) imageries. Few High-resolution satellite imageries have been ordered from Pleiades, Digital globe, Airbus for the period 2013-17 and some of them were used in this paper. In addition, few imageries have been extracted from the Google Earth software. The details of the satellite data used in this paper are listed in Table 1.

3.2 Wave data

Wave data from the nearby wave rider buoy moored at 25 m depth (Figure 1) have been obtained for the study period from Indian National Centre for Ocean Information Services (INCOIS). The wave parameters like significant wave height (H_s), zero-upcross wave period (T_z), wave direction (θ) data are available at 30 minute interval and were converted to daily data for simplicity.

Table 1: Statistics of wave and computed morphological parameters for the reported drowning events as compared with the satellite imageries (GE = Google Earth, PL = Pleiades)

Drowning event	Satellite image (source)	Data availability (days)	H_s (m)	θ (deg.)	T_m (s)	β	W (m)	ϵ	Beach type
22-09-2013			1.81	160.30	7.27	0.06	71.89	18.45	LBT
01-01-2014	25-12-2013 (GE)	-7	0.71	168.56	5.94	0.11	19.01	4.39	LTT
14-03-2014	14-03-2014 (GE)	0	0.62	165.90	6.56	0.10	18.68	3.64	LTT
	17-10-2014 (GE)								LBT
22-03-2015			0.62	171.60	5.80	0.11	15.72	3.55	LTT
	14-05-2015 (GE)		1.37		8.03	0.07		12.24	RBB
	29-05-2015 (PL)		1.07		5.11	0.10		24.10	TBR
11-10-2015		+1	0.99	157.50	6.12	0.09	29.51	7.51	TBR
	12-10-2015 (GE)		0.90	158.02	6.12	0.09	26.52	6.51	LTT
	17-01-2016 (GE)		0.34		6.19	0.13	9.06	1.52	RFL
08-05-2016			1.21	154.69	6.94	0.08	43.33	10.16	RBB
10-05-2016			0.92	136.41	6.85	0.09	30.30	6.44	LTT
	22-05-2016 (GE)	+12	2.32		7.29	0.06	95.24	26.74	DIS
28-05-2016			1.63	177.19	6.44	0.07	52.16	14.75	RBB
09-06-2016			2.10	163.13	6.04	0.07	67.21	22.99	DIS
22-06-2016			0.95	154.70	6.06	0.09	27.85	7.06	LTT
29-06-2016			2.10	170.20	7.02	0.06	79.50	22.39	DIS
14-09-2016			1.27	170.20	6.45	0.08	40.56	10.50	RBB
25-09-2016			1.33	143.44	5.66	0.09	37.07	11.58	RBB
	12-10-2016 (GE)	-2	1.12	167.58	7.03	0.08	39.40	8.76	TBR
14-10-2016			1.13	156.09	8.96	0.07	54.90	9.09	TBR
	18-02-2017 (PL)		0.64		5.85	0.11	16.87	3.84	LTT

3.3 Nearshore morphology and surf parameters computation

The change in morphological state is driven by a change in the surf scaling parameter (ϵ) described by (Guza and Inman, 1975)

$$\epsilon = \frac{a_b \omega^2}{g \tan^2 \beta} \tag{1}$$

where, a_b is the breaker amplitude ($H_b/2$), ω is the radian frequency of the incident waves ($2\pi/T$, T = wave period), g is acceleration due to gravity and $\tan\beta$ is the gradient of the beach/surf zone. The beach slope (β) has been estimated using the formula by (Sunamura, 1984)

$$\tan \beta = 0.12g^{0.25} H_b^{-0.5} T^{0.5} d_{50}^{0.25} \tag{2}$$

here, $H_b = H_o \times K_s \times K_r$ is the breaker height, K_s and K_r are the shoaling and refraction coefficients. The value for d_{50} has been considered as 0.6 mm (Arun Kumar and Prasad, 2014) for RK Beach.

As wave energy decreases following a storm, the beach changes from dissipative ($\epsilon > 20$) to reflective ($\epsilon < 2.5$) by passing through each of four intermediate states: longshore bar-trough, rhythmic bar-beach, transverse bar and rip and low tide terrace ($2.5 > \epsilon < 20$) (Table 2). Each of these intermediate states is characterized (to varying degrees) by the presence of accretion rips. As the shore face evolves, there is an increase in the rip current velocity as the cross-sectional rip channel area decreases and the ability of the

bar system to block return flows increases (Brander, 1999), until the innermost bar completely welds to the beach face. These rips do not have the characteristic form and may be difficult to identify from the perspective of a beach user. Further, they persist through both storm and non-storm conditions.

The surf zone width is an important parameter that decides the strength of the rip current. It has been computed using $W = H_b / (0.43 * \tan \beta)$ (3) It has been observed that larger the surf zone width, stronger is the rip current.

Table 2: The surf scaling parameter, beach type, rip current probability and the strength as defined by Wright and Short (Wright & Short, 1984)

ϵ	Beach type	Rip current formation probability	Rip current strength
< 2.5	Reflective	No	Nil
2.5 – 7.5	Ridge runnel (RR) or low tide terrace (LTT)	Yes	Moderate
7.5 – 10	Transverse bar and rip (TBR)	Yes	Very Strong
10 – 15	Rhythmic bar and beach (RBB)	Yes	Strong
15 – 20	Longshore bar trough (LBT)	Yes	Weak
>20	Dissipative	No	Nil

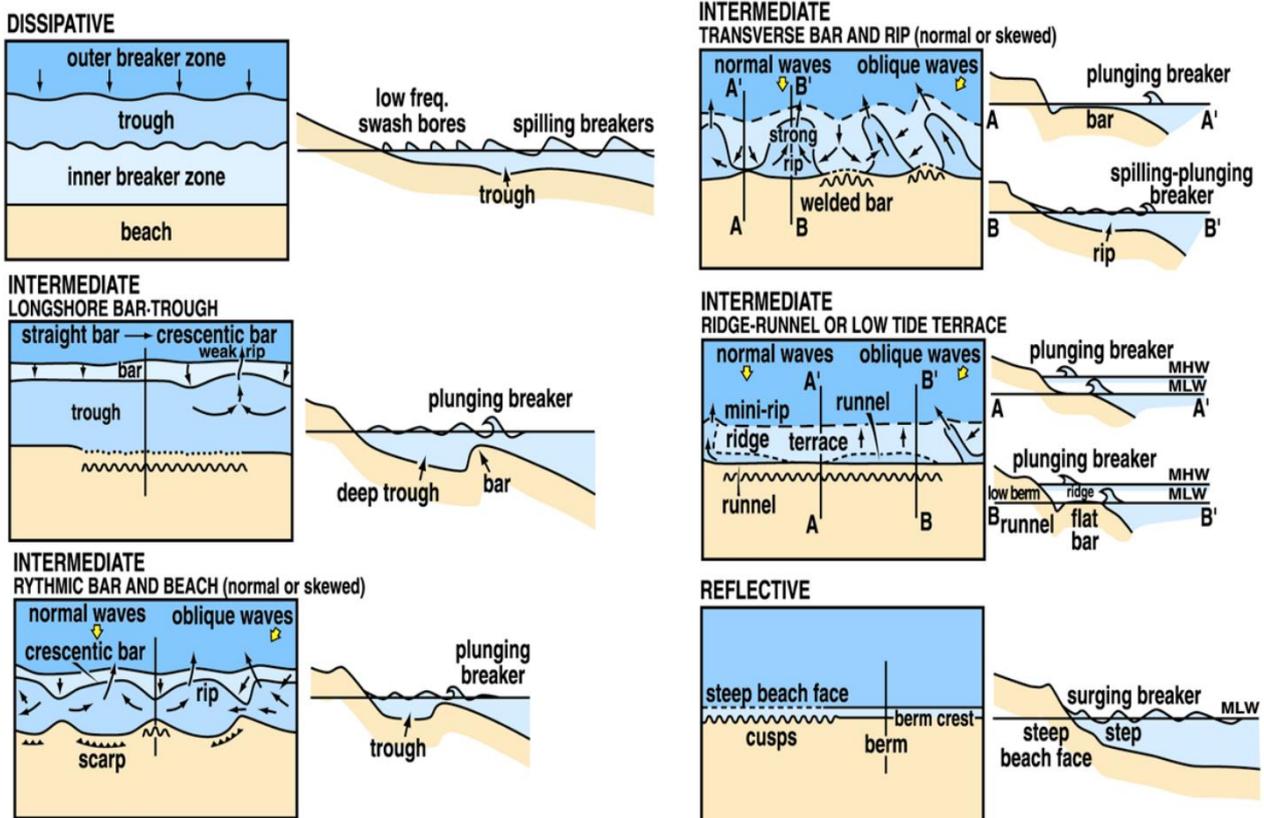


Figure 2: Configuration of the six major beach types. Modified from Wright and Short (Wright & Short, 1984).

4. Results and discussion

Nearshore bars migrate offshore in response to, and onshore in recovery from the storms. Migration offshore is caused by breaking waves affecting the bar crest which creates an offshore-directed current (Ruessink et al., 1998; Plant et al., 2001). In this section, we have explained the application of satellite remote sensing to observe the transition of a beach from dissipative (high wave energy) to the reflective type (calm sea) and discussed the morphology and migration of nearshore bars affect circulation and formation of rips.

4.1 Beach type analysis from satellite imageries

In general, nearshore bars migrate onshore during calm conditions and offshore during high-energy conditions. Rip currents develop as the bar migrates landward. If a reset event occurs, the bar will migrate offshore and rip channel formations will be lost until the bar migrates landward again, restarting the sequence of stages in the bar cycle. This cycle has been examined from a series of high resolution satellite imageries to infer the rip current hazard for different beach stages. In the absence of on-site observations and measurements, the satellite images can be used to classify various beach states as mentioned below, by following the classification descriptions suggested by (Wright and Short, 1984).

Accordingly, the beach is classified into three stages Dissipative, Intermediate and Reflective. The Intermediate stage is further classified into four stages Longshore bar-trough (LBT), Rhythmic bar and beach (RBB), Transverse bar and rip (TBR) and Low-Tide Terrace (LTT) or Ridge Runnel (RR) (Figure 2).

4.1.1 Dissipative

This type of beach occurs generally during stormy condition and monsoons, where the surf zone is full of breaking waves (Figure 3a). High sea-state attract fewer people to venture in to the surf during this time. No rips are possible during this stage. However, due to high energy waves, there are chances of drowning due to wave attack.

4.1.2 Intermediate

Longshore bar-trough (LBT)

A longshore bar-trough (LBT) morphology is characterized by a straight or crescentic nearshore bar with a trough on its landward edge and a shoreline that may be straight or have cusps. If there is any discontinuity in the bar, a weak rip is expected. In figure 3b, a significant bar along with the beach cusps can be clearly seen. After a few days of the storm, this bar starts coming closer to the shore.

Rhythmic bar and beach (RBB)

The rhythmic bar and beach (RBB) morphology presents a crescentic bar with horns migrating more landward and aligning with mega cusp horns alongshore (Figure 3c). Rip currents may form at the bay sections of the crescentic bar. However, their signatures may not be clear from the beach view.

Transverse bar and rip (TBR)

The transverse bar and rip state (TBR) develops during an accretionary sequence when the horns of the crescentic bar welds with the beach face. Rip currents are created when water flows seaward through the topographic channels created from the partial welding of the nearshore bar (Wright and Short, 1984; Sonu, 1972) producing the strongest rip current circulation. The cusped shoreline of the RBB beach state is softened by the partial welding of the bars and creates a more undulating shoreline profile (Figure 3d). This beach type will have the strongest rip current and pose danger to the swimmers often occurs in pre and post monsoon.



Figure 3a: Dissipative beach type classification with satellite image on 22-05-2016



Figure 3b: Longshore bar-trough beach type classification with satellite image on 17-10-2014



Figure 3c: Rhythmic bar and beach type classification with satellite image on 14-05-2015



Figure 3d: Transverse bar and rip beach type classification with satellite image on 29-05-2015



Figure 3e: Ridge-runnel (RR) or low-tide terrace (LTT) beach type classification with satellite image on 18-02-2017

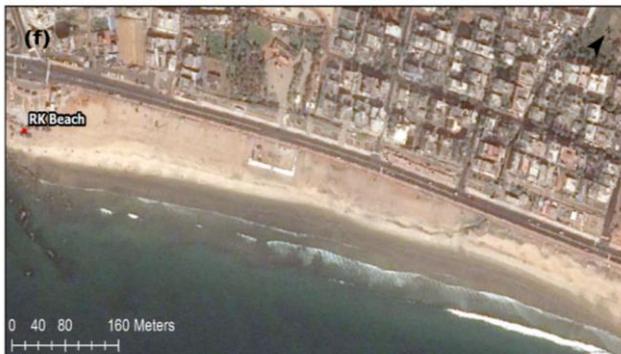


Figure 3f: Reflective beach type classification with satellite image on 17-01-2016

Ridge-runnel (RR) or low-tide terrace (LTT)

The low tide terrace or ridge and runnel beach state (LTT) is formed as the bar welds almost entirely onto the beach face creating a terrace at low tide. Small rip currents may be present during this stage but are more dominant during the TBR and RBB beach states (Figure 3e). This type of beach generally occurs during January and February.

4.1.3 Reflective

If calm conditions continue, the bar will continue to attach to the shore and rip channels will disappear (Ruessink and Kroon, 1994; Houser and Greenwood, 2005; Aagaard et al., 2004). The beach has a steep face and does not pose danger to the general public. Figure 3f shows a typical reflective type of beach in the study region. This type of beach often occurs during January or February.

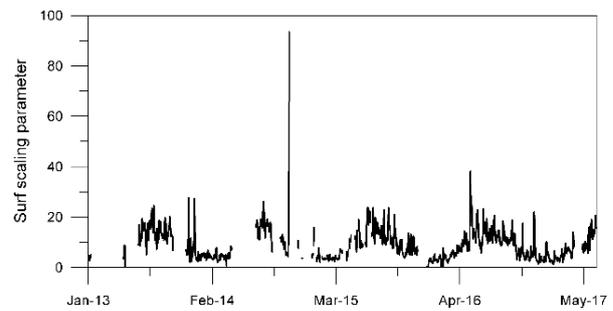


Figure 4: Variation in the Surf scaling parameter

4.2 Rip current hazard variability

From the previous section, it has been observed that the rip current probability and its strength can be easily interpreted from the satellite imageries based on its classification stage. If the wave data is available, the beach stage can be computed using the equation (1) and therefore rip current can be predicted with or without any other supporting data. The surf similarity parameter is a proxy for the rip current formation. It is observed that the values are high during the monsoon season (June to September) and low during January – March (Figure 4). From Table 2, it is clear that stronger rip currents form when the ϵ ranges between 7.5 and 15. However, the rip formation probability is seen for $\epsilon = 2.5$ to 20.

The wave height, wave period, the computed surf zone width and the beach slope are shown in figure 5. A significant seasonal cycle can be observed in all the parameters (excluding wave period, wave direction – not shown). The variation of beach slope is inversely related to wave height. It is evident that the higher sea state erodes the beach face and thereby decreases its slope.

The dependency of wave height (or breaker wave height), surf zone width and the beach slope have been examined in figure 6. It is observed that the surf scaling parameter is high for higher wave heights, larger surf zone widths but smaller beach slopes.

4.3 Rip current drowning analysis

It has been observed that the rip current formation and its strength are strongly related to the surf scaling parameter, which in turn depends on the surf zone width, beach slope and wave height. In order to prove this, the actual drowning cases reported at RK Beach during 2013-2017 have been examined in this section. During this period, around 14 cases have been reported due to rip currents. However, there are no actual observations at the time of the event. Using the method described in the above section, surf zone and nearshore morphological parameters have been calculated for the drowning events. As anticipated, none of the events reported during dissipative and reflective stages. All the events happened during the intermediate stage. Out of all, most of the rip current related drownings reported in the LTT beach stage. It is expected that more people venture into the sea following the storm due to the calm sea state. However, due to alongshore irregularities in the bar, rip current pose danger and lead to drowning.

The beach stage obtained from the wave data for the rip current cases have been compared with the satellite imageries. Due to lack of coincident high-resolution satellite data at the time of the event, the scene within 15 days of the event has been considered for this comparison. Surprisingly, the computations from the wave data are perfectly representing the beach stage on the satellite imagery during the event. On 14-03-2014, LTT type of beach was present and the same type has been obtained from the wave data. In another case, a rip current drowning was reported on 14-10-2016, where the beach stage was TBR as obtained from the wave data. There is no coincident satellite image to verify on that day. However, the same TBR stage was obtained from the satellite image two days prior (12-10-2016). It is often understood that TBR stage will remain at least 2-7 days unless a reset event occurs.

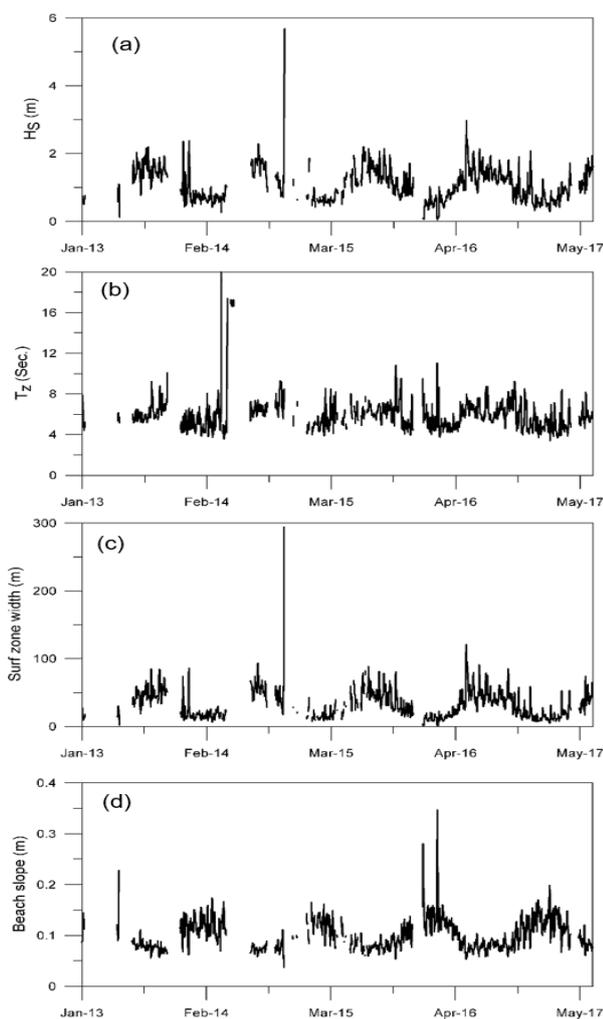


Figure 5: Time series of wave height, wave period, computed surf zone width and beach slope.

Whereas in the case of 10-05-2016 event, the beach stage was LTT, the nearest available satellite image was on 22-05-2016. But by that time the sea state was entirely changed to the dissipative domain. Hence, cannot be compared. From this analysis, it is very clear that the rip current events can be well predicted just by observing the beach stage classification. However, the interpretation is very subjective and requires skills and experience. Also,

high resolution satellite imageries are expensive and sometimes covered with clouds. Whereas, the same job can be easily done using wave data if obtained from a nearby buoy or from an accurate model forecast. The early detection of a change in beach stage can be very much helpful in saving valuable lives in the beaches.

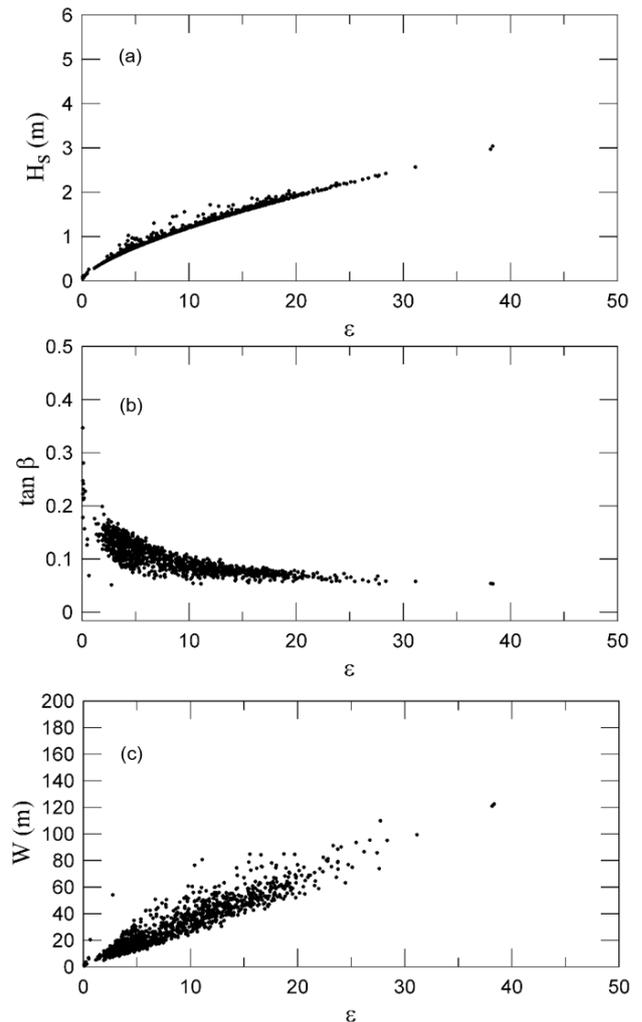


Figure 6: Relationship between H_s , beach slope, and surf zone width with surf scaling parameter

5. Conclusion

High resolution optical satellite imageries have been used to analyse different beach stages along RK Beach, Visakhapatnam where a record drowning reported. The satellite data is supported by the in-situ wave data from a nearby buoy in the shallow waters. Surf scaling parameter - a proxy to the rip current formation was computed from the wave data. It has been observed that the surf scaling parameter has strong seasonal variability and related well with the surf zone width, beach slope and wave height. Based on the drowning data for the period 2013 – 2017, we have observed that most of the drownings happened at low-tide terrace beach type (LTT) and occur during pre and post monsoon seasons. The method proposed in this paper can be used to detect the beach stage and thereby the rip current probability and its strength. The method will be used to examine other beaches along the Indian coast in future.

Acknowledgements

The authors are thankful to the Director, Space Applications Centre (ISRO) for his constant support and encouragement. The first and second authors are thankful to Dr A.S. Rajawat, Group Director (GHCAG), Shri. Arun Kumar Sharma, Head (GSD) and Dr Rashmi Sharma, Head (OSD) for their suggestions and encouragement. They are equally thankful to INCOIS for providing wave rider buoy data and NRSC for providing satellite imageries from Digital Globe, Pleiades, AirBus etc. We acknowledge the support of Google Earth for providing satellite imageries. The first author is thankful to DST-SERB for providing fellowship.

References

- Aagaard, T., R. Davidson-Arnott, B. Greenwood and J. Nielsen (2004). Sediment supply from shoreface to dunes: linking sediment transport measurements and long-term morphological evolution, *Geomorphology*, 60, 205-224.
- Aagaard, T., B. Greenwood and J. Nielsen (1997). Mean currents and sediment transport in a rip channel, *Marine Geology*, 140, 25-45.
- Bowen, A. J. (1969). Rip currents: 1. Theoretical investigations, *Journal of Geophysical Research*, 74, 5467-5478.
- Brander, R. W. (1999). Field observations on the morphodynamic evolution of a low-energy rip current system, *Marine geology*, 157, 199-217.
- Brander, R. W. and A. D. Short (2000). Morphodynamics of a large-scale rip current system at Muriwai Beach, New Zealand, *Marine Geology*, 165, 27-39.
- Chandramohan, J. and L.-K Liang (1985). Bernoulli, multinomial and Markov chain thinning of some point processes and some results about the superposition of dependent renewal processes, *Journal of applied probability*, 22, 828-835.
- Drozdowski, D., A. Roberts, D. Dominey-Howes, and R. Brander (2015). The experiences of weak and non-swimmers caught in rip currents at Australian beaches, *Australian Geographer*, 46, 15-32.
- Guza, R. T. and D. L. Inman (1975). Edge waves and beach cusps, *Journal of Geophysical Research*, 80, 2997-3012.
- Haas, K. A. and I. A. Svendsen (2002). Laboratory measurements of the vertical structure of rip currents, *Journal of Geophysical Research: Oceans*, 107.
- Houser, C. and B. Greenwood (2005). Profile response of a lacustrine multiple barred nearshore to a sequence of storm events, *Geomorphology*, 69, 118-137.
- Arun Kumar, S. V. V. and K. V. S. R. Prasad (2014). Rip current-related fatalities in India: a new predictive risk scale for forecasting rip currents, *Natural hazards*, 70, 313-335.
- Kumar, V. S., K. A. Kumar and N. S. N. Raju (2001). Nearshore processes along Tikkavanipalem beach, Visakhapatnam, India, *Journal of Coastal Research*, 271-279.
- Kumar, V. S., K. C. Pathak, P. Pednekar and N. S. N. Raju (2006). Coastal processes along the Indian coastline, *Current science*, 91(4), 530-536.
- MacMahan, J. H., A. J. H. M. Reniers, E. B. Thornton and T. P. Stanton (2004). Infragravity rip current pulsations, *Journal of Geophysical Research: Oceans*, 109.
- MacMahan, J. H., E. B. Thornton, T. P. Stanton and A. J. H. M. Reniers (2005). RIPEX: Observations of a rip current system, *Marine Geology*, 218, 113-134.
- Panigrahi, J. K., V. S. Kumar and J. K. Tripathy (2010). Littoral drift by alongshore flow at Visakhapatnam-East Coast of India, *Journal of Hydro-environment Research*, 4, 317-327.
- Plant, N. G., M. H. Freilich and R. A. Holman (2001). Role of morphologic feedback in surf zone sandbar response, *Journal of Geophysical Research: Oceans*, 106, 973-989.
- Ruessink, B. G., M. G. Kleinbans and P. G. L. Van den Beukel (1998). Observations of swash under highly dissipative conditions, *Journal of Geophysical Research: Oceans*, 103, 3111-3118.
- Ruessink, B. G. and A. Kroon (1994). The behaviour of a multiple bar system in the nearshore zone of Terschelling, the Netherlands: 1965 -1993, *Marine Geology*, 121, 187-197.
- Short, A.D. (1985). Rip-current type, spacing and persistence, Narrabeen Beach, Australia, *Marine geology*, 65, 47-71.
- Sonu, C.J. (1972). Field observation of nearshore circulation and meandering currents, *Journal of Geophysical Research*, 77, 3232-3247.
- Sunamura, T. (1984). Quantitative predictions of beach-face slopes, *Geological Society of America Bulletin*, 95, 242-245.
- Turner, I. L., D. Whyte, B. G. Ruessink and R. Ranasinghe (2007). Observations of rip spacing, persistence and mobility at a long, straight coastline, *Marine Geology*, 236, 209-221.
- Wright, L. D. and A.D. Short (1984). Morphodynamic variability of surf zones and beaches: a synthesis, *Marine geology*, 56, 93-118.

Change detection and trend analysis for Oceansat-2 Ocean Color Monitor (OCM-2) time series data

Ravi Kamal Choudhary*, Tushar Shukla, Sampa Roy and Debajyoti Dhar
Space Applications Centre, ISRO, Ahmedabad, India
*Email: choudharyravi@sac.isro.gov.in

(Received: Jan 05, 2019; in final form: Jun 06, 2019)

Abstract: Satellite ocean-colour observations are now widely recognized as an important component of international remote sensing programs. Ocean Color Monitor (OCM-2), one of the payloads aboard Oceansat-2 (2009), is designed to obtain quantitative information of ocean-colour variables e.g. chlorophyll-a concentration, etc. The two important tools for analyzing multi-temporal data obtained from earth observation satellite are Change Detection and Time series Trend Analysis. The remote sensing data has become a heart of change detection and trend analysis techniques because of its high temporal frequency and wider selection of spatial and spectral resolution. The general objectives of change detection in remote sensing include recognizing the geographical location and type of changes, quantifying the changes, and assessing the accuracy of change detection results. Change detection is useful in many applications such as land use changes, habitat fragmentation, rate of deforestation, coastal change, urban sprawl, and other cumulative changes. In time series analysis, the goal is to estimate the future value using the behaviors in the past data. Trend Analysis predicts the values over cloudy and missing data, thus helping in generating a uniform time series. This paper discusses remote sensing techniques namely Differencing and Principal Component Analysis (PCA) for Change Detection and ARIMA based method for Trend Analysis on multi-temporal OCM-2 data.

Keywords: Change Detection, Trend Analysis, OCM-2, Oceansat-2, Vegetation fraction, Chlorophyll-a, ARIMA, Image Differencing, PCA, Time series.

1. Introduction

OCEANSAT-2 spacecraft of Indian Space Research Organization (ISRO) is the second satellite in ocean series, which carried three main instruments namely i) Ku band pencil beam Scatterometer, ii) Ocean Colour Monitor (OCM) and iii) Radio Occultation Sounder of Atmosphere (ROSA) instrument of Italian Space Agency (ASI). The main objectives of OceanSat-2 are to study surface winds and ocean surface strata, observation of chlorophyll concentrations, monitoring of phytoplankton bloom, study of atmospheric aerosols and suspended sediments in the water. The OCEANSAT-2 OCM is mainly designed to provide continuity to the OCEANSAT-1 OCM instrument and OCM collects data in 8 spectral bands operating in the Visible-Near IR spectral range and the imaging principle of OCM is based on push-broom technique. The OCM application lies in ocean application like identifying potential fishrie zone (PFZ). The configuration of the OCM-2 payload is identical to the OceanSat-1 OCM except that the spectral band is modified for band 6 and band 7. For band 6, the central wavelength is shifted from 670 nm to 620 nm to improve the reflectance from suspended sediments; for band 7, the cenetral wavelength is shifted from 760 nm to 740 nm to avoid oxygen absorption. Ocean Color Monitor (OCM-2) instrument is designed to obtain quantitative information of ocean-colour variables e.g. chlorophyll-a concentration, vertical diffuse attenuation of the light, (K_d) and total suspended matter (TSM) concentration in coastal waters, apart from ocean-colour information OCM data is also useful for studying the aerosol transport and terrestrial bio-sphere.

Change detection is the process of identifying differences in the state of an object or phenomenon by observing it at different times. It involves the ability to quantify temporal

effects using multi temporal data sets (Singh, 1989). It involves the analysis of two registered multispectral remote sensing images acquired in the same geophysical area at different times to measure how the attributes of a particular area have changed between two or more time periods. Two remote sensing techniques, including Image Differencing and Principal Component Analysis (PCA) were used to detect the changes in multi-temporal OCM-2 data, such as vegetation fraction (VF), land surface water (LSW), chlorophyll-a concentration (CLO) and aerosol optical depth (AOD).

A Time Series (TS) is a collection of observations made sequentially through time. The TS data used to provide visual information to the unpredictable nature of the data. It is an ordered sequence of observations of a variable or captured object at equally distributed time interval. TS is anything which is observed sequentially over the time at regular interval like hourly, daily, weekly, monthly, quarterly etc. TS data is important when you are predicting something which is changing over the time using past data. To carry out the change detection techniques, trend analysis and ARIMA forecast, Oceansat-2 (OCM-2) satellite data were used.

The following basic things have been addressed in this paper

1. Change Detection methodology
2. Change Detection techniques- Image Differencing and Principal Component Analysis (PCA)
3. Trend Analysis and Auto-Regressive Integrated Moving Average (ARIMA) forecast model

2. Change detection

The remote sensing is very important part of the change detection techniques. The time and accuracy of change detection on the earth's surface can provide a better understanding of the relationship and interaction between human and natural phenomena. The change detection problem discussed in this paper is as follows: we are given two images of the same scene taken at several different times. The goal is to identify the set of pixels that are "significantly different" between these two images, these pixels comprise the change mask. The change mask may result from a combination of underlying factors, including appearance or disappearance of objects, motion of objects relative to the background, or shape changes of objects.

There are many different methods for doing change detection (Jianya et al., 2008); this paper analyses the Image Differencing and PCA change detection techniques using OCM-2 satellite images. Image differencing is pixel-based method while principal component analysis is transformed based method of change detection. Classification based methods such as post classification methods and two-date image clustering specify the occurrence of changes in image pixels and label them. Various methods have been proposed to detect changes captured by satellite images. The algebraic methods like image differencing is relatively simple and easy to understand. The disadvantage of these methods are in choosing an appropriate threshold. However, this method does not extract details of the changes completely. Transformation methods focus on reducing the data between bands, and value different information in derived components.

A basic change detection algorithm takes the image sequence as input and generates a binary image 'B' containing [0, 1] called a change mask that identifies changed regions in the last image according to the following generic rule (Radke et al., 2005):

$$B(x) = \begin{cases} 1, & \text{if there is a significant change at} \\ & \text{pixel } x \text{ between consecutive images} \\ 0, & \text{otherwise} \end{cases}$$

Before implementing change detection analysis, the following conditions must be satisfied (Lu et al., 2004):

1. Precise registration of multi-temporal images;
2. Precise radiometric and atmospheric correction or normalization between multi-temporal images;
3. Region/ Area of Interest: same geographic location, free of clouds in the area of analysis;
4. Remote sensing system consideration- spatial, spectral, radiometric and temporal: whenever possible, select images acquired from the same type of sensors, with the same spectral and spatial resolutions, and the same seasonal timeframe in order to minimize unwanted variances.

2.1 Change detection methodology

The methodology used for analysis of different change detection methods using satellite images is divided into five major parts as shown in figure 1. The method involves processing of Oceanasat-2 (OCM-2) images that are detecting the change in study area. The image processing

techniques using first preprocessed the images then image processing is used for different change detection algorithms. The block diagram of proposed system as shown in figure 1.

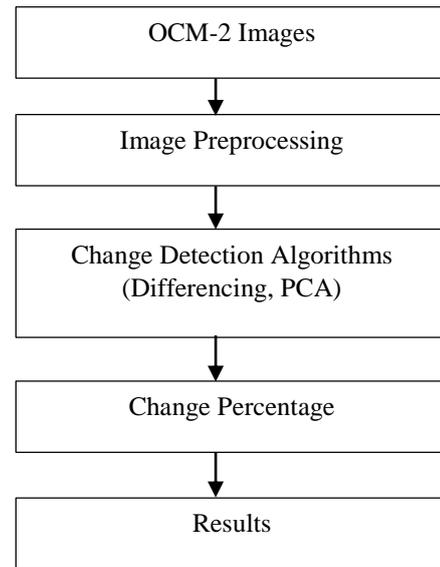


Figure 1: Block diagram of Change Detection System

2.2 Change detection techniques

After the pre-processing stage was completed, two different change detection algorithms were applied to both the images. These were Image Differencing and Principal Component Analysis. The first is pixel based method while the second is transformation based method of change detection.

2.2.1 Image differencing

Image differencing is very intelligible method of the change detection techniques. It is applied to a wide variety of the images and geographical data. It is generally conducted on the basis of gray level images. Image differencing is used widely because of its simplicity to implement and interpret. It involves absolute subtraction of the second-date image from a first-date image, pixel by pixel. The changed and unchanged area is determined by selecting the appropriate threshold value of gray level subtraction image using Otsu's method (Otsu, 1979).

Threshold value (T) is determined from difference image, which decides the quality of change detection. Choosing this suitable threshold value can be maximum separated for the areas of real change. A critical aspect of image differencing method is deciding where to place the threshold boundaries between change and no-change pixels displayed in the histogram. Otsu's automatic thresholding algorithm is used to find out the threshold. Change mask is then prepared by assigning '1' to pixels which are significantly changed in second-date image as compare to first-date image and '0' otherwise representing change and no-change regions respectively. Mathematically,

$$\begin{aligned} \text{Difference image, } D(x) &= \text{Image1} - \text{Image2} \\ \text{Change mask, } B(x) &= \begin{cases} 1 & \text{if } |D(x)| > T \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

2.2.2 Principal Component Analysis (PCA)

The principal components transformation is a linear transformation that defines a new, orthogonal co-ordinate system such that that data can be represented without correlation. PCA is a technique to emphasize variation and bring out strong patterns in a data set. It converts a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components (Gong, 1993; Fung and LeDrew, 1987). It is a widely used method for dimensionality reduction. The three main steps towards implementing PCA algorithm are:

1. Difference image generation and Eigen vector space (EVS)
2. Building the feature vector space (FVS)
3. Clustering of the feature vector space and change map

2.2.3 Difference image generation and Eigen vector space (EVS)

The difference image has the absolute valued differences of the intensity values of the corresponding pixels of the 2-grayscale images. The computed difference image would hence be such that the values of the pixels associated with land changes will have values significantly different from those of the pixels associated with unchanged pixels.

$$\begin{aligned} \text{difference image } (i, j) \\ = | \text{Image1 } (i, j) - \text{Image2 } (i, j) | \end{aligned}$$

To construct the EVS, we take non-overlapping blocks of size $N \times N$ from the difference image and flatten them into row vectors to make vector set from the difference image. PCA takes this vector set and determines its co-variance matrix after performing mean normalization on it. The Eigen vectors and Eigen values of the co-variance matrix are computed (giving us the EVS) and then the Eigen vectors are sorted in the descending order of Eigen values.

2.2.4 Building the Feature Vector Space (FVS)

Taking $N \times N$ blocks from the difference image, flattening them, and lastly projecting them onto the EVS, forms feature vector space. This time the blocks are overlapping. A vector space (VS) is first made by constructing one vector for each pixel of the difference image such a way that one $N \times N$ block is actually a pixel's $N \times N$ neighborhood. FVS is constructed by projecting the vector set on to the EVS, simply means to perform the following matrix multiplication:

$$\begin{aligned} \text{Feature Vector Set (FVS)} \\ = \text{Vector Set (VS)} \cdot \text{Eigen Vector Space (EVS)} \end{aligned}$$

2.2.5 Clustering of the feature vector space and change map:

The feature vectors for the pixels carry information whether the pixels have characteristics of a changed pixel or an unchanged one. Having constructed the feature vector space, we now need to cluster it so that the pixels can be grouped into two disjoint classes, changed and unchanged class. To do that we used K-means algorithm. Thus, each pixel will get assigned to a cluster in such a way that the distance between the cluster's mean vector and the

pixel's feature vector is the least. Each pixel gets a label from 1 to K, which denotes the cluster number that they belong to. It can be postulated that the cluster that has the highest mean is the cluster that belongs to the changed class. The reason behind the cluster which has the highest value of mean belongs to changed class is that the values of the difference image pixels in a region where some changes have occurred are higher than the values of pixels in the regions where there is no change.

Then we build a change map – a binary image to show the output of change detection. We are showing the background blackcurrant, i.e., intensity value of those pixels is zero and the changes in yellow, i.e., intensity value of those pixels is 1. Thus

$$\begin{aligned} \text{Change map } (i, j) \\ = \begin{cases} 1, & \text{if } (i, j) \in \text{highest mean cluster} \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

3. Trend analysis

A time series is said to be continuous when observations are made continuously through time and is said to be discrete when observations are taken only at specified times, usually equally spaced. The special feature of TS analysis is the fact that successive observations are usually not independent and that the analysis must take into account the time order of the observations. The first step in the analysis is usually to plot the observations against time to give what is called a time plot, and then to obtain simple descriptive measures of the main properties of the series. A graph will not only show up trend and seasonal variation, but will also reveal any wild observations or outliers that do not appear to be consistent with the rest of the data. Other features to look for in a time plot include sudden or gradual changes in the properties of the series. If there is some sort of discontinuity in the series, then different models may need to be fitted to different parts of the series. The prediction of the future values of the series is another important task of time series analysis.

This paper includes plotting the data and looking for trends, seasonal fluctuations and forecasting of time series data based on model called autoregressive integrated moving average (ARIMA). Traditional methods of time series analysis are mainly concerned with decomposing the variation in a series into components representing trend, seasonal variation and other cyclic changes. Any remaining variation is attributed to 'irregular fluctuations'. This approach is not always the best but is particularly valuable when the variations dominated by trend seasonality. Vegetation fraction over a specific area for year 2012 is chosen for trend analysis. Median is chosen to remove effect of extreme values in time series and to capture spatial variation coming from experimental area. The missing value in time series is estimated by applying cubic spline fitting to the known data. Trend and Seasonal variations are removed from cubic spline fitted time series to make it stationary before applying ARIMA forecasting model.

The components, by which time series is composed of, are called component of time series data. The four basic component of time series data are: Trend, Seasonal variations, Cyclical variations and Irregular (random) variations.

3.1 Time series forecast models

Most of the Time series (Das, 1994) models work on the assumption that the time series is stationary. TS is said to be stationary if its statistical properties such as mean, variance remain constant over time. Intuitively, we can say that if TS has a particular behavior over time, there is a very high probability that it will follow the same in the future. Almost none of the TS are stationary. There are two major reasons behind the non-stationarity of TS:

1. Trend: varying mean over time.
2. Seasonality: variations at specific time frames.

Therefore, to make TS stationary, we need to remove trend and seasonal variations from TS. There are different ways to estimate and eliminate trend from a TS, and some of commonly used are:

1. Aggregation: taking average for a time period like monthly/weekly average.
2. Smoothing: taking rolling averages.
3. Polynomial fitting: fit a regression model.

Seasonal variations are eliminated from a time series by doing seasonal adjustment to get seasonally adjusted TS.

3.1.1 Auto-Regressive Integrated Moving Avergae (ARIMA)

ARIMA (Box and Jenkins, 1970; Mondal et al., 2014) is a statistical technique that uses time series data to predict future. The ARIMA model is a combination of autoregressive (AR), integration (I) - referring to the reverse process of differencing to produce the forecast, and moving average (MA) operations (Farhath et al., 2016) . In ARIMA (p, d, q) model; p, d and q are integers greater than or equal to zero and refer to the order of the autoregressive components, the number of differencing operators, and the highest order of the moving average term of the model respectively.

ARIMA modelling will take care of trend, seasonal and cyclic variations of a data set while making forecasts. ARIMA requires stationary TS to forecast. In ARIMA, ‘d’ parameter controls the forecast of TS based on stationarity. In ARIMA model a non-stationary time series is made stationary by applying finite differencing of the data points. The ARIMA forecasting for a stationary time series is nothing but a linear (like a linear regression) equation. The predictors depend on the parameters (p, d, q) of the ARIMA model:

1. Number of AR (Auto-Regressive) terms (p): AR terms are just lags of dependent variable.
2. Number of MA (Moving Average) terms (q): MA terms are lagged forecast errors in prediction equation.
3. Number of Differences (d): These are the number of non-seasonal differences.

To determine the values of ‘p’ and ‘q’, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are used.

1. Autocorrelation Function (ACF): It is a measure of the correlation between the TS with a lagged version of itself.
2. Partial Autocorrelation Function (PACF): This measures the correlation between the TS with error term.

Autoregressive models

ARIMA methodology attempts to describe the movements in a stationary time series as a function of "autoregressive and moving average" parameters. An AR model with k parameters may be written as

$$X(t) = A(1) * X(t - 1) + A(2) * X(t - 2) + \dots + A(k) * X(t - k) + E(t)$$

Where,

X(t) = time series under investigation

A(k) = the autoregressive parameter of order k

X(t-k) = the time series lagged k period

E(t) = the error term of the model

The p, d, and q parameters are retrieved from experimental area. The best suited ARIMA model p, d and q values are carried out for experimental data based on minimum MAPE value. These values for various parameters are shown in table 1.

Table 1: Best suited ARIMA (p, d, q) parameters retrieved from experimental data

S. No.	Parameter	ARIMA Model (p,d,q)
1.	Vegetation fraction	(5,1,0)
2.	Land Surface Water	(3,1,0)
3.	Chlorophyll-a concentration	(2,1,0)
4.	Aerosol Optical Depth	(2,1,0)

4. Results

The Change Detection (CD) techniques (Image Differencing, Principal Component Analysis (PCA)) are tested on multiple OCM-2 datasets. In time series forecasting experiments, ARIMA model is fitted with the best-suited model parameters (p, d and q) for experimental data. We tested the original time series with the ARIMA forecasted values and observed that the forecasted values are nearly matching with the original time series. The Mean Absolute Percentage Error (MAPE) was used to measure the prediction accuracy of the forecasting model. Change Detection results for vegetation fraction are shown in figure 2, where change map is binary map containing one for changed pixels and zero for no changed pixels and change map magnitude shows how much value of pixels are changed between these two dates.

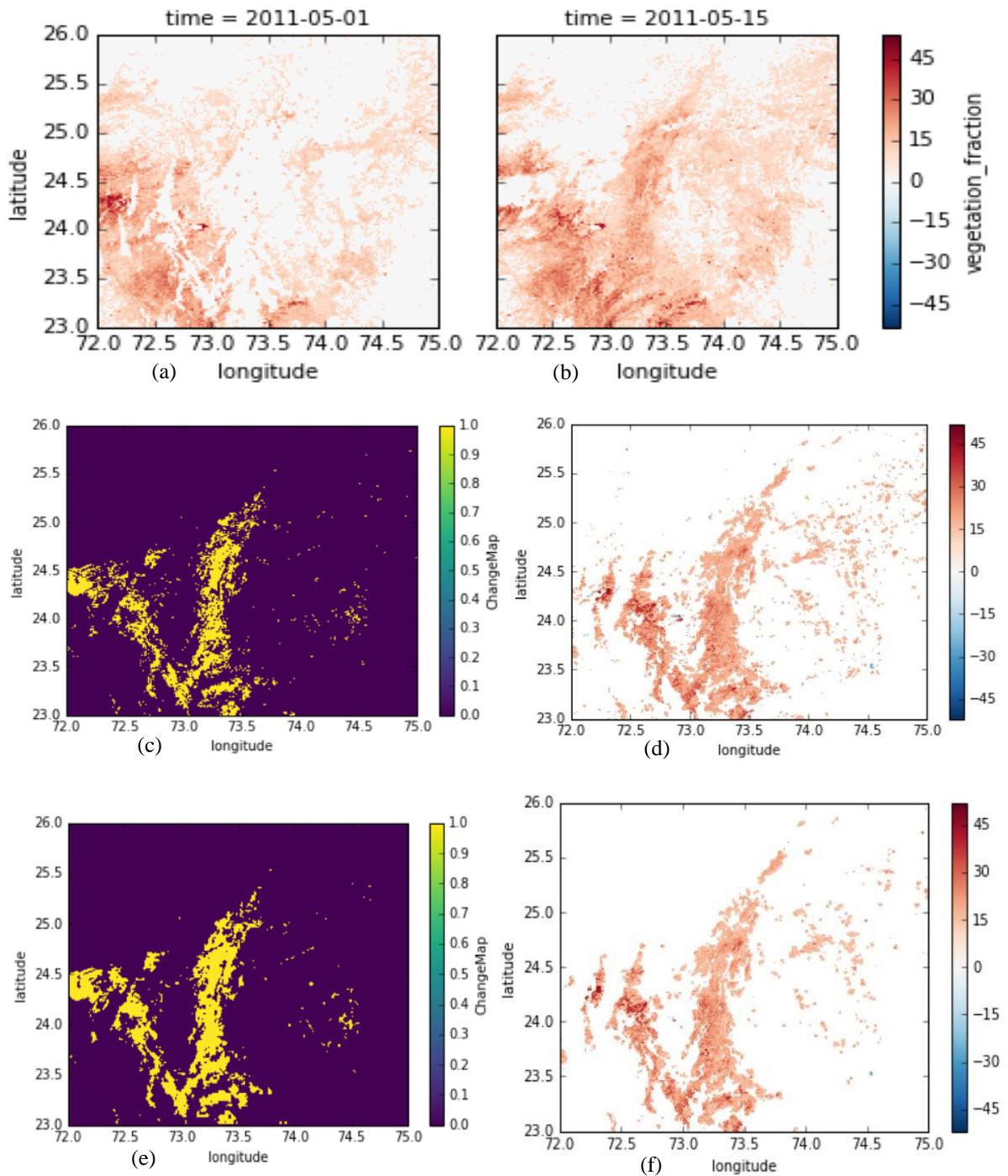


Figure 2: Vegetation fraction of date (a) 01-05-2011, (b) 15-05-2011, (c) Image differencing change Map, (d) Image differencing change map magnitude, (e) PCA change map, (f) PCA change map magnitude

Trend Analysis results are shown in figure (3a - 3d), where plot (a) shows the median time series of a parameter for the chosen dates, (b) shows the cubic spline fitting using non-zero values of time series, (c) shows Train data, Test data; which are 66% and 34% of seasonally adjusted TS respectively and ARIMA forecast for Test data, (d) shows Original TS (Cubic spline fitted TS), Smoothed TS (after

removing trend from spline fitted TS), Seasonally Adjusted TS (after removing seasonal variation from smoothed TS) and ARIMA forecast TS for Test data. As it can be seen from figure 3a that OCM-2 data contains lot of missing values, therefore to estimate the missing values the cubic spline fitting is done to non zero data.

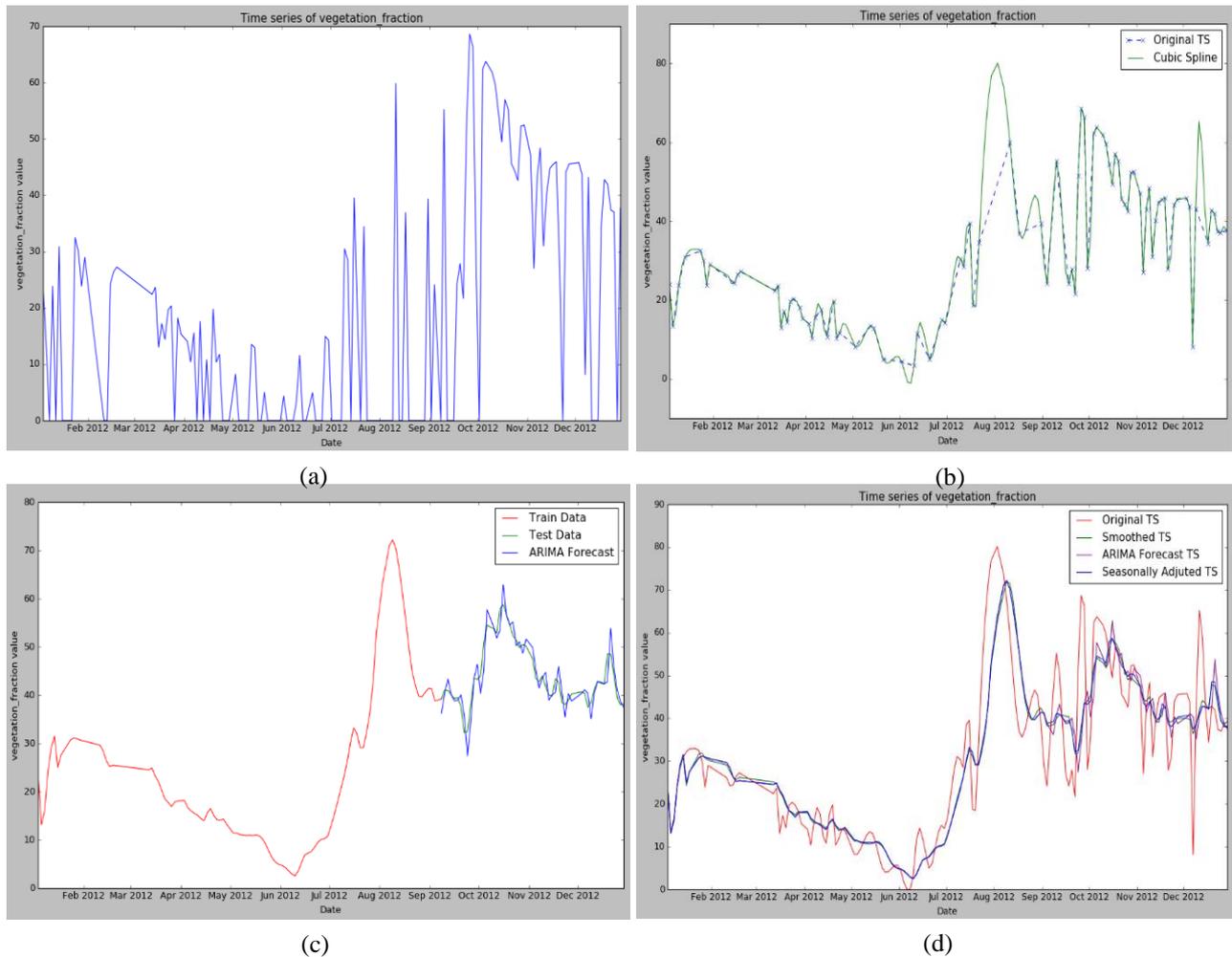


Figure 3: (a) Median Time Series (TS) of vegetation fraction from 01JAN2012 to 31DEC2012, (b) Original non-zero and Cubic Spline fitted TS, (c) Train, Test data and ARIMA forecast for Test, (d) Original TS, Smoothed TS, Seasonally Adjusted TS and ARIMA forecast TS for Test data

MAPE for various parameters are shown in table 2.

Table 2: MAPE for various parameters

S.No.	Parameter	MAPE
1.	Vegetation fraction	4.52
2.	Land Surface Water	0.77
3.	Chlorophyll-a concentration	3.99
4.	Aerosol Optical Depth	2.49

5. Conclusion

Two remote sensing techniques namely Differencing and Principal Component Analysis (PCA) for Change Detection and ARIMA based method for Trend Analysis on multi-temporal OCM-2 time series data is discussed in detail. The change detection outcomes are change map, change map magnitude and percentage of area changed between two images. The percentages of changed area for Differencing and PCA are 7.4 and 8.8 respectively for experimental data.

This paper also forecast various parameters using statistical method, ARIMA, where best suited (p,d,q)

values are estimated from minimum MAPE values. The MAPE values of 4.52, 0.77, 3.99 and 2.49 are observed for Vegetation Fraction, Land Surface Water, Chlorophyll-a concentration and Aerosol Optical Depth respectively for experimental data using best suited (p,d,q) values.

References

- Box, G.E.P., and G.M. Jenkins (1970). Time Series Analysis: Forecasting and Control, CA: Holden-Day, San Francisco.
- Das, S. (1994). Time Series Analysis. Princeton University Press, Princeton, NJ.
- Farhath, Z.A., B. Arputhamary, L. Arockiam (2016). A survey on ARIMA forecasting using time series model, International Journal of Computer Science and Mobile Computing, 5(8), 104-109.
- Fung, T. and E. LeDrew (1987). Application of principal components analysis to change detection, Photogrammetric Engineering and Remote Sensing 53(12), 1649-1658.
- Gong, P. (1993). Change detection using principal component analysis and fuzzy set theory, Canadian Journal of Remote Sensing, 19(1), 22-29.

- Jianya, G., S. Haigang, M. Guorui and Z. Qiming (2008). A review of multi-temporal remote sensing data change detection algorithms, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B7), 757-762.
- Lu, D., P. Mausel, E. Brondizio and E. Moran (2004). Change detection techniques, *International journal of remote sensing*, 25(12), 2365–2401.
- Mondal, P., L. Shit and S. Goswami (2014). Study of effectiveness of time series modelling (ARIMA) in forecasting stock prices, *International Journal of Computer Science, Engineering and Applications*, 4(2), 13.
- Otsu, N. (1979). A threshold selection method from gray-level histograms, *IEEE transactions on systems, man, and cybernetics*, 9(1), 62-66.
- Radke, R.J., S. Andra, O. Al-Kofahi and B. Roysam (2005). Image change detection algorithms: A systematic survey, *IEEE Transactions on Image Processing*, 14(3), 294-307.
- Singh, A. (1989). Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing*, 10(6), 989–1003.

Spatial enhancement of SWIR band from Resourcesat-2A by preserving spectral details for accurate mapping of water bodies

Indranil Misra*, Yatharath Bhateja, Neha Gaur and Vivek Sharma
Space Applications Centre (ISRO), Jodhpur Tekra, Ahmedabad – 380015, India
*E-Mail: indranil@sac.isro.gov.in

(Received: Dec 21, 2018; in final form: Jun 07, 2019)

Abstract: Inland water bodies monitoring using multi spectral images is an important application of remote sensing. Resourcesat-2A (RS-2A) launched on December 7th, 2016 provides data at multi-tier imaging scheme through three instruments viz. AWiFS, LISS-3 and LISS-4 to assess the crop pattern, urban sprawl monitoring, snow and glacier studies, change detection, map land surface water bodies and delineating their spatial distribution to understand the hydrological processes. Short Wave Infrared (SWIR) band from RS-2A LISS-3 sensor is available at best observed spatial resolution of 24 meters and wavelength range of 1.55 μm -1.70 μm plays a vital role for spectral water index determination especially for modified normalized difference water index (MNDWI) which is calculated from green and SWIR bands. The water index accuracy can be enhanced further by improving the spatial resolution of SWIR band from RS-2A platform. In this paper, SWIR band spatial enhancement techniques are described in detail and is divided mainly in two major processing stages. In the first stage, SWIR band spatial resolution is improved directly by combination of laplacian operator to highlight the finer details with smooth gradient to enhance the prominent edges and downscale the image to 12 meters with lanczos based resampling kernel without affecting much the dynamic range of the gray level of SWIR band. In the second stage, spatial enhanced SWIR band of LISS-3 is used for overlap region extraction with same time acquisition data of LISS-4 having spatial resolution of 5 meters and perform weighted average using brovey based image fusion technique to generate SWIR band of RS-2A at spatial resolution of 5 meters by preserving the spectral information. The processing workflow developed can increase SWIR band spatial resolution from 24 meters to 5 meters and improve land surface water bodies mapping performance.

Keywords: SWIR, MNDWI, Resourcesat-2A, Spatial Enhancement, Image Fusion

1. Introduction

Resourcesat-2A provides multi spectral images at varying resolution using three cameras: AWiFS at 56 meters, LISS-3 at 24 meters and LISS-4 FMX at 5 meters spatial resolution. This multi-tier imaging dataset is potentially important for regional and global water bodies mapping over Indian terrain due to systematic and frequent revisit capabilities. At present, Indian remote sensing data can help us to do routine monitoring for land surface water bodies, which is substantially different from in situ measurements. One of the reliable, user friendly and computationally effective method is Modified Normalized Difference Water Index (MDNWI) proposed by (Xu, 2006) that uses green and short wave infrared (SWIR) bands to enhance water information and can extract water bodies with greater accuracy.

Green band is available by Resourcesat-2A LISS-4FMX camera at a finer spatial resolution of 5 meters. But SWIR band is only available at best possible spatial resolution of 24 meters through LISS-3 Camera. This limitation forces us to improve the spatial resolution of SWIR band at ground for accurate water bodies mapping. In this paper, spatial enhancement techniques are discussed in details to bring the spatial resolution of SWIR from 24 meters to 5 meters. Both spatial interpolation and image fusion methods are applied to increase the spatial resolution of SWIR band. Spatially enhanced SWIR can detect smaller size open water bodies with more explicit and accurate boundaries (Yan et al., 2016).

2. Spatial interpolation for image enhancement

The spatial interpolation is applied to coarser spatial resolution data directly and does not use any additional datasets. The image spatial resolution is improved by applying spatial filters with interpolation kernels in combination to improve the overall spatial information of the datasets. The major steps in spatial interpolation are as follows:

2.1 Image resampling

Resampling is a process that involves the extraction and interpolation of gray levels from pixel locations in the original distorted image and their relocation to the approximate matrix coordinate location in the rectified (corrected) image (Parkar et al., 1983). Image values in non-integer coordinates are computed by lanczos based sinc interpolation kernel.

Lanczos is used as low-pass filter over the remote sensing data to smoothly interpolate the values between its samples. Lanczos kernel is actually a windowed sinc function typically used to resize the digital image and considered to be the “best compromise” among several spatial filters. In our case, RS-2A SWIR image is resampled to 12 meters using lanczos resampling kernel.

2.2 Spatial filtering

SWIR image at 12 meters’ spatial resolution is improved spatially using image derivative by applying laplacian operator to highlight the finer feature details in SWIR band (Xie et al., 2009). But it also introduces additional noise in the image that is suppressed by applying smooth gradient mask filter that enhance the prominent edges without

affecting much the dynamic range of the gray level of SWIR band.

3. Image fusion with LISS-4 FMX

Remote Sensing Image fusion generates single hybrid image from a collection of input satellite images and helps us to extract maximum information from the remotely sensed datasets to achieve optimal spatial and spectral resolution. The image fusion is therefore useful for integrating a high spectral resolution image with high spatial resolution image such as spatially improved LISS-3 data at 12 meters spatial resolution with LISS-4 data at 5 meters spatial resolution to produce high quality and accurate data at 5 meters which contains the characteristic of both the multispectral information (object identification) and the spatial detail (object localization and texture). Image fusion in automatic mode requires sequential execution of different phases for generating quality data products. Digital Image Processing techniques are used to generate fused image in different stages. The main steps/phases involved in image fusion are:

3.1 Geometric transformation

In fusion process, the multi sensor data taken as input have different spatial resolution and generally stored in different projection system. So it is needed to bring the datasets into same projection system and all images are need to resampled to same pixel size using standard resampling technique described in section 2.1. This geometric transformation makes the images ready for fusion to achieve better spatial and spectral characteristics.

3.2 Image registration

Image Registration is the process of aligning different images of the same scene acquired at different times, different viewing angles, and/or different sensors (Misra et al., 2012a). It plays an important role in remote sensing and applied in wide variety of tasks such as image fusion. The image fusion results in science quality data product only when multi sensor data are corrected or modeled for relative geometric error. The feature based image registration techniques are better approach for registering spatially enhanced LISS-3 data with LISS-4 data.

3.3 Image fusion

Fusion described by (Lucien, 2003) as a “formal frame work in which are expressed means and tools for the alliance of data originating from different sources. The integration of low and high resolution optical images from Indian Remote Sensing satellites facilitates better visual and automatic image interpretation that aims at retrieving more information of greater quality. The fusion technique used for merging multi sensor data is color normalized fusion which is a variant of Brovey Transform and has the

capability of generating fused product with optimal spectral and spatial resolution.

4. Image fusion methodology

The fusion technique used here for merging multi resolution optical data sets is color normalized fusion which is a variant of Brovey Transform and has the capability of generating fused product with optimal spectral and spatial resolution. Details of this technique are explained further in the following sections.

4.1 Brovey Transform (BT)

The BT uses ratios to sharpen the SWIR image. It was created to produce high-resolution image by merging. Many researchers used the BT to fuse a low-resolution multi spectral band with a high-resolution image. The basic procedure of the BT multiplies each multispectral band by the high-resolution data and then divides each product by mean of the low-resolution multispectral band.

4.2 Color Normalized Transformation (CN)

Color Normalized is an extension of the BT. CN transform also referred to as an energy subdivision transform. The CN transform separates the spectral space into hue and brightness components. The transform low-resolution LISS-3 band with high-resolution LISS-4 FMX data, and these resulting values are each normalized by dividing the mean of the low-resolution multispectral band (Misra et al., 2012b). In the fusion process, a study is carried out to determine most suitable high resolution LISS-4 band by measuring the correlation co-efficient between them and SWIR band of LISS-3. It is noticed that 5m LISS-4 FMX Near Infrared (NIR) band have the greatest correlation coefficient with LISS-3 SWIR and there by chosen as high resolution band for image fusion.

5. Spatial enhancement processing workflow

The processing workflow for improving the spatial resolution of RS2A comprises of two major processing stages. In the first stage, spatial interpolation based image enhancement is carried out that uses the image itself and improve the feature details in the image and generate SWIR band at 12 meters spatial resolution. In the second processing stage, spatially enhanced LISS-3 data is fused with same date acquisition of LISS-4 data on-board Resourcesat-2A platform to generate product at 5 meters spatial resolution. The processing workflow comprises of geometric transformation and multi sensor image registration to make LISS-3 data eligible for image fusion with corresponding high-resolution LISS-4 data. Figure 1 shows the processing workflow developed for SWIR band spatial enhancement.

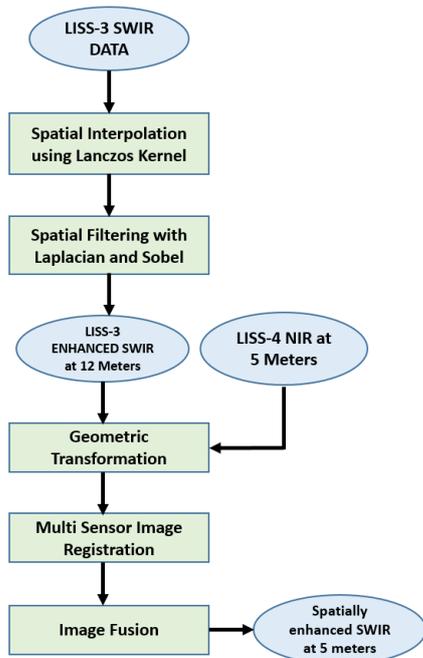


Figure 1: Spatial enhancement processing workflow

6. Results

RS-2A LISS-3 SWIR data is processed using the developed processing workflow. The datasets details is shown in table 1. LISS-3 and LISS-4 same date acquisition data are used to improve the spatial resolution of LISS-3 SWIR band

LISS-3 SWIR data is improved in first iteration using spatial filtering with interpolation technique and downscale the image from 24 meters to 12 meters. Spatially enhanced SWIR at 12 meters along with LISS-4 NIR data at 5 meters is merged to generate SWIR band at 5 meters. Figures 2 & 3 show the SWIR band original, enhanced and finally fused output together to better understand the improvement in each processing stage. In the spatial processing it has been taken care that spectral characteristics of LISS-3 SWIR band should be preserve while improving the spatial resolution.

Table 1: Spatial enhancement dataset details

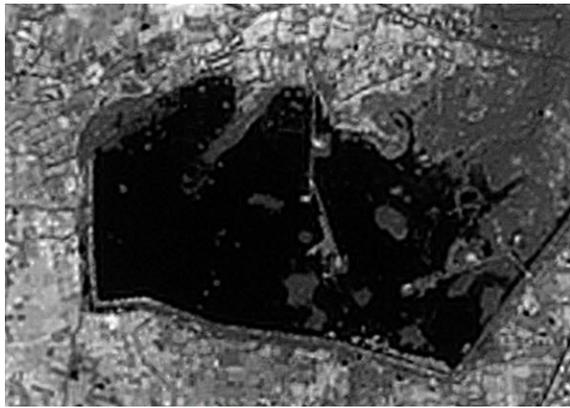
Satellite	Sensor	Location	Path/Row	Date of Pass	Map Projection	Spatial Resolution (in meters)
RS2A	LISS-3	Ahmedabad,	93/56	26 Nov 2017	UTM	24 meters
RS2A	LISS-4	Gujarat, India		26 Nov 2017	UTM	5 meters



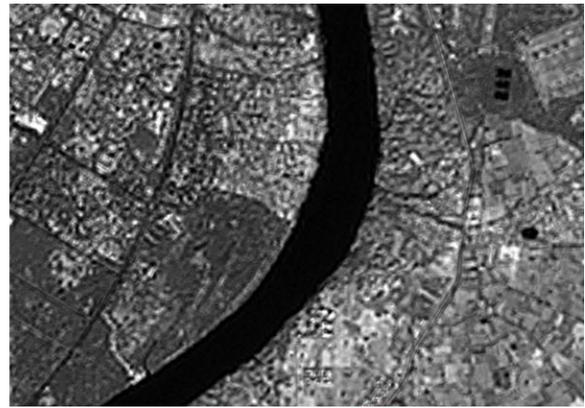
Figure 2: Spatial improvement in LISS-3 SWIR data over Ahmedabad region



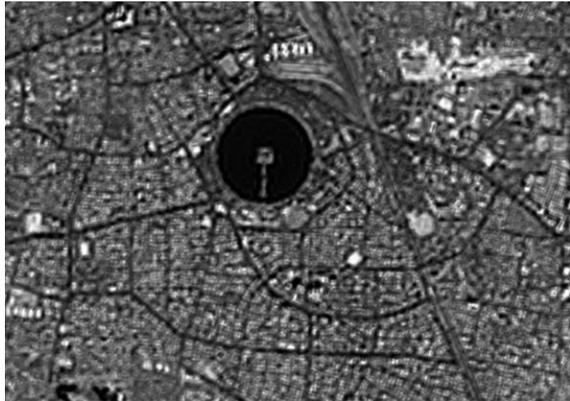
Figure 3: Spatial improvement in LISS-3 SWIR data over Gandhinagar region



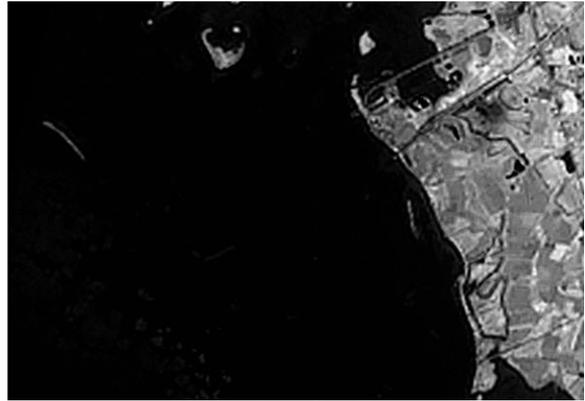
(a) Thol Lake



(b) Sabarmati River



(c) Kankaria Lake



(d) Nalsarovar

Figure 4: Water bodies of Ahmedabad and its surrounding

SWIR constructed at 5.0 meters need to be compared with original LISS-3 SWIR which is taken as reference for spatial enhancement technique. The original LISS-3 SWIR is at 24.0 meters which need to be cubic interpolated at 5.0 meters for comparison with fused SWIR. Table 2 shows the different image quality metrics (Hore and Ziou, 2010) computed with LISS-4 NIR Band as reference. The fused SWIR at 5.0 meters shows high Signal-to-Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM) and low Root Mean Square Error (RMSE), Mean Absolute Error (MAE) as expected than its original interpolated LISS-3 SWIR. It shows the overall image quality is improved in fused SWIR than its original counterpart.

Quantitative analysis is performed on different water bodies that includes river and lakes. Figure 4 shows some prominent water bodies located in and around Ahmedabad. For assessing the spectral characteristics of fused SWIR, overlap water body region (group of pixels) is extracted from both original SWIR and fused SWIR. The original LISS-3 SWIR at 24.0 meters is radiometrically calibrated and meets the absolute radiometry accuracy specifications. Table 3 shows the average DN count differences between original and fused overlap regions taken from different water bodies as shown in figure 4. The average radiometric accuracy achieved is around 90% with respect to original LISS-3 SWIR which is good enough to map wetland boundaries at better map scale.

Table 2: Image quality metrics comparison between fused and original SWIR

S.No	Test Images	SNR	PSNR	SSIM	RMSE	MAE
1.	Fused LISS-3 SWIR	9.9907	26.1059	0.9981	44.5106	28.9225
2.	Original Interpolated LISS-3 SWIR	8.2663	24.3815	0.9925	54.2853	34.2580

Table 3: Radiometry accuracy assessment table

S.No	Water Body Location	Original SWIR (Avg. DN Count)	Fused SWIR (Avg. DN Count)	Radiometric Accuracy (%)
1.	Thol Lake	10.1	8.8	87.1
2.	Sabarmati River	11.7	10.3	88.0
3.	Kankaria Lake	15.3	14.7	96.7
4.	Nalsarovar	5.4	4.6	85.1

7. Conclusion

Spatial enhanced LISS-3 SWIR band generated can be used for accurate water body mapping by computing spectral water index MNDWI at higher spatial resolution. The processing workflow is tested with multiple cloud free LISS-3 datasets with simultaneous LISS-4 acquisition and looks to provide better accuracy in most of the cases. But also need to see the accuracy of the product for other water bodies such as coastal region and river delta which is not covered in this paper. The future work includes super resolution using single image with deep learning approaches and plug the modules in the existing processing workflow for high quality SWIR data generation.

Acknowledgements

The authors acknowledge Shri D K Das, Director, Space Applications Centre (SAC), ISRO for his encouragement. The authors also thank other team members of optical data processing for their support during the course of work.

References

- Hore, A., and D. Ziou (2010). Image Quality Metrics: Psnr vs ssim. 20th International Conference on Pattern Recognition, Istanbul, Turkey, August 23-26, 2010.
- Lucien, W. (2003). Data Fusion: Definitions and Architectures. Les Presses de l'Ecole des Mines, Paris.
- Misra, I., R.K Gambhir, S.M. Moorthi, D. Dhar and R. Ramakrishnan (2012a). An efficient algorithm for automatic fusion of RISAT-1 SAR data and Resourcesat-2 Optical images, IEEE Intelligent Human Computer Interaction (IHCI), IIT Kharagpur, December 27-29 2012.
- Misra, I., S.M. Moorthi, D. Dhar and R. Ramakrishnan (2012b). An Automatic satellite image registration technique based on Harris corner detection and Random sample Consensus (RANSAC) outlier rejection model, IEEE Recent Advances in information technology, ISM Dhanbad, India, March 15-17 2012.
- Parkar, A., J. Robert, V. Kenyon and D.E. Troxel (1983). Comparison of interpolating methods for image resampling, IEEE transaction on Medical Imaging, 2(1), 31-39.
- Xie, X., S. Zaifeng, G. Wei and Y. Suying (2009). An adaptive image enhancement technique based on image characteristic, 2nd International Congress on Image and Signal Processing, Tianjin, China, October 17-19 2009.
- Xu, H. (2006). Modification of Normalized Difference Water Index (NDWI) to enhance open water features in remotely sensed imagery, International Journal of Remote Sensing, 27(14), 3025-3033.
- Yan, D., Y. Zhang, F. Ling, Q. Wang, W. Li and Xiaodong (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, 354, 1-19.

Extraction of water body, cloud shadow and cloud detection using object-based classification

Sudhanshu Raghubanshi*, Ritesh Agrawal, A.K. Sharma and A.S. Rajawat
Space Applications Centre, Ahmedabad, 380015, India
*Email: sraghubanshi@gmail.com

(Received: Jan 26, 2019; in final form: June 12, 2019)

Abstract: Shadow plays a crucial role in Satellite Image Interpretation. Shadow may sometimes obscure ground features. In Optical Imagery, the Spectral Signature (DN value) of Water Body and Cloud Shadow are similar. Features which lie within shadow, reflect less energy and are difficult to identify. This results in difficulties in digital classification, which depends completely on reflectance value (DN value). The object based classification approach utilizes spectral value as well as Shape, Texture and Context information. Such additional attributes are helpful for the detection of the shadow and water body separately. In this study, LISS-3 data of part of Rajkot district, Gujarat was used for detection of cloud shadow, cloud and water bodies. The three parameters such as average DN, reliability and threshold values were used for the shadow and cloud detection. Reliability is a criterion for providing priority to desired class in case of class mixing. Clouds were separated by keeping value of minimum threshold as 185. The water body is differentiated from the shadow by providing reliability of 0.4 as compared to providing reliability of 0.3 in case of the shadow. The results clearly show that in optical satellite images, cloud shadows can be separated from the water body and the cloud can be detected using Object-Based techniques.

Keywords: Reliability, Shadow, Classification, Cloud

1. Introduction

Accuracy of the classified image depends upon the classification technique. Digital classification of satellite data, which uses spectral signature (DN value) of the individual pixel (Willhauck, 2000), works well when the classes are spectrally separable. These can be carried out using supervised classification, unsupervised classification or combination of these (Enderle and Weih, 2005). In case of spectral mixing, these techniques fail and importance of the neighboring pixels (Spatial and Contextual information) may also be neglected.

In contrast, object-based image analysis incorporates not only spectral value, but also shape, texture and context information (Flanders et al., 2003) for classification. Image is divided into groups of homogeneous pixels, which are called objects-created from the segmentation process. On applying proper rules, the objects are classified to ground cover features. Basic entity is group of pixels rather than the single pixel. Object-based image analysis of satellite data has been utilized for decades (Ryherd and Woodcock, 1996; Flanders et al., 2003), but in recent years it has been utilized in different areas such as vegetation monitoring (Yu et al., 2006), forest cover analysis (Heyman et al., 2003), water body extraction (He et al., 2016).

In satellite image interpretation, shadows play very crucial role. It aids in interpretation as well as creates difficulty. Features can be identified as their association with shadow, such as, shadow and water body appear similar, but identified separately due to shadow association with the cloud and hills. Features which lie within shadow, reflects

less energy and are difficult to identify. In optical imagery, the spectral signature (DN value) of water body, cloud shadow and hill shadow are similar. Individual identification of these are difficult using digital classification methods which solely depends upon spectral value. Object-based approach incorporates not only spectral value, but also shape, texture and context information. Using object-based approach, classes which are spectrally similar can be separated out.

In present study, separation of cloud shadow from the water body and cloud detection using Object-Based approach has been attempted. The data used was LISS-3 (Optical) data of part of Rajkot district, Gujarat. This analysis makes use of DN, reliability and threshold value using open source Software - Inter IMAGE for the desired objective. Using reliability value, priority will be given to the class where there is mixing between two classes. The results show that cloud shadow can be separated from the water body and cloud can be detected using Object-Based technique.

2. Study Area and data used

The Study Area was part of Rajkot district located in Gujarat State associated with ground cover features water body, cloud, cloud shadow, settlement and vegetation. The data used was Indian Remote Sensing-P6 (IRS-P6) Satellite, LISS-3 acquired on 27-09-2016 with the resolution of 23.5 m. Wetland boundary of year 2016 (Wetland inventory of India, 2016) was utilized as reference data for the accuracy assessment. Both the data sets are shown in figure 1.

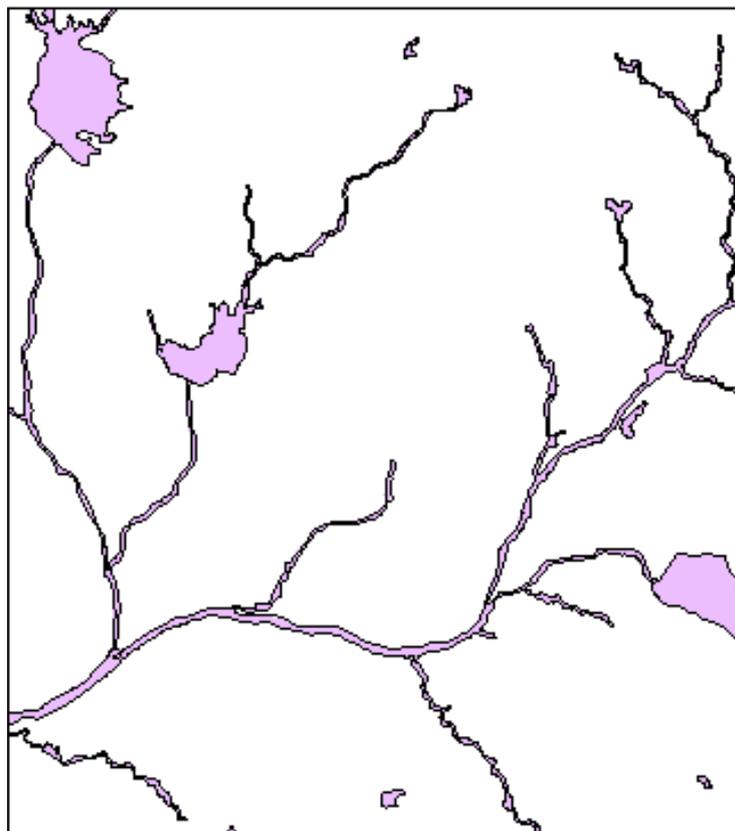
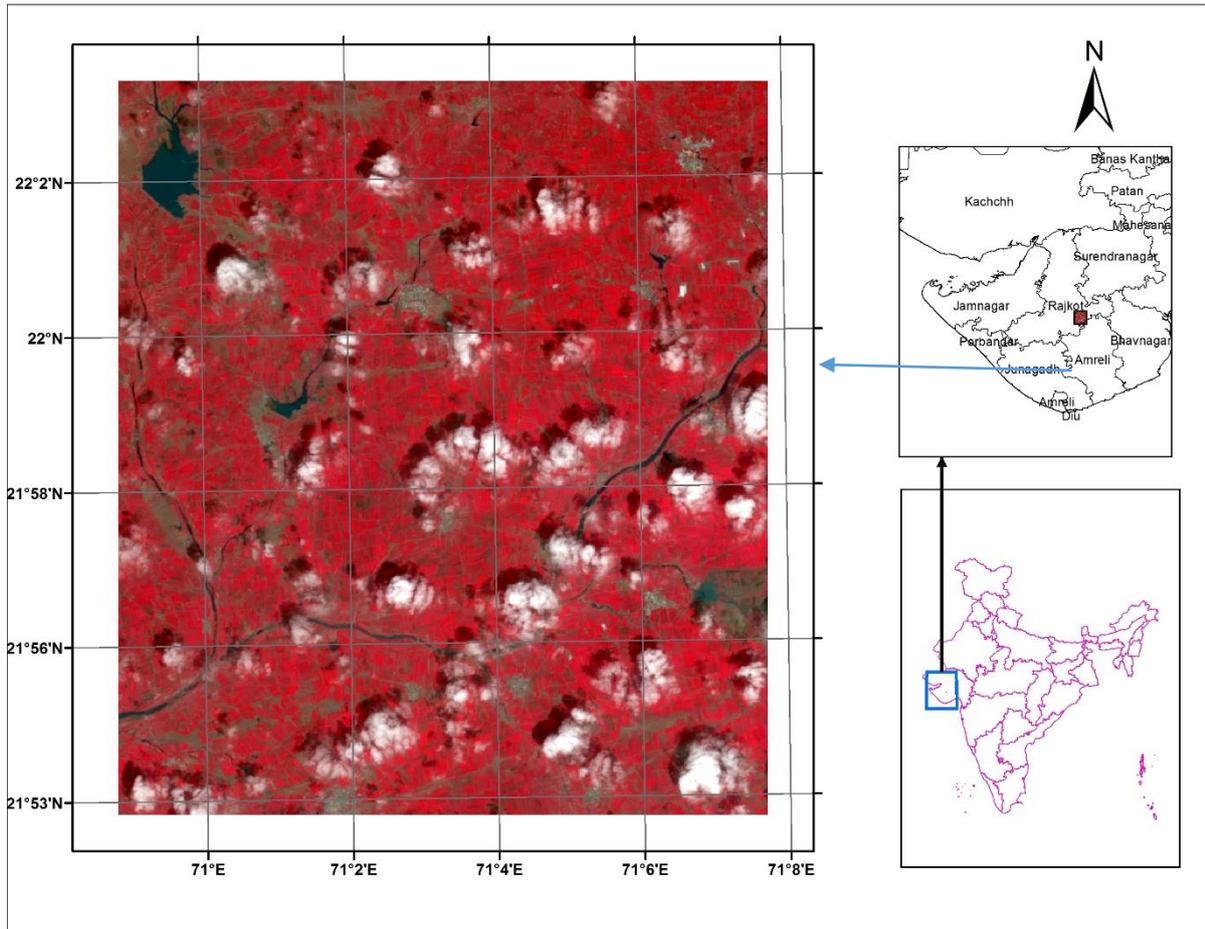


Figure 1: LISS-3 data and wetland boundary of part of Rajkot district, Gujarat.

3. Objective

Objective of this study was detection and separation of water body-cloud Shadow and identification of Cloud using Object-Based classification.

4. Methodology

The objective of study was achieved using object-based classification technique which utilized spectral (DN value) as well as shape, texture and contextual information. For the classification, Inter IMAGE- an open source software was utilized. The image was classified into 4 classes viz. water, cloud shadow, non-shadow and cloud. Objects are created from the multiresolution segmentation approach (Baatz and Schape 2010) using scale parameter 30, compactness weight 0.5 and color weight 0.8. Heterogeneity and closeness of pixels between the objects are governed by scale and compactness respectively. Further, the objects are classified into various classes by applying different rules. Rules which are ratios and averages of DN values in different spectral bands (Green, Red, NIR and SWIR) were incorporated for the achievement of objective. These rules were combined with two additional parameters, reliability and threshold value. Reliability gives the higher weightage of class in the case of class mixing. Priority will be given with higher reliability value to the class where geographic overlay exists between two classes. Different threshold value had been checked for the classification of cloud and shadow. Initially, image was divided into water and non-water with the rule WBI (Water Body Index) associated with water class. WBI is the ratio of reflectance (DN) value associated with green and near infra-red wavelength bands (Green-NIR)/(Green + NIR) (McFeeters, 1996). Positive value of WBI indicates water bodies. Non-water class is divided into cloud and non-cloud with the rule average DN values of all bands associated with non-cloud class and threshold value 185. Finally, non-cloud class was divided into non-shadow and shadow with the same rule as non-cloud but decreased threshold value (100) associated with shadow class. Threshold value 185 and 100 gave the best possible result for the cloud and shadow identification. Rules which

have been incorporated for the classification are summarized in the following table 1.

Table 1: Classification rules

Class	Rules
Water Body	$(\text{Green-NIR})/(\text{Green} + \text{NIR})$
Cloud and Shadow	$(\text{Green} + \text{Red} + \text{NIR} + \text{SWIR})/4$

Extracted classes water body (Blue), cloud shadow (Black) and cloud (White) are shown in figure 2.

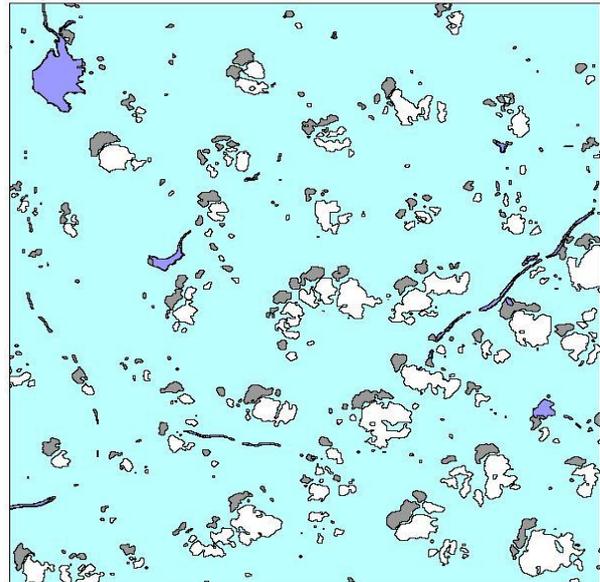


Figure 2: Classified water body (Blue), cloud shadow (Black) and cloud (White)

When the reliability of shadow class (0.3) is higher than the water class (0.2), there was mixing of water body and cloud shadow. Increased reliability of water class (0.4) as compared to shadow class (0.3) gives the separation of water body and cloud shadow (Figure 3).

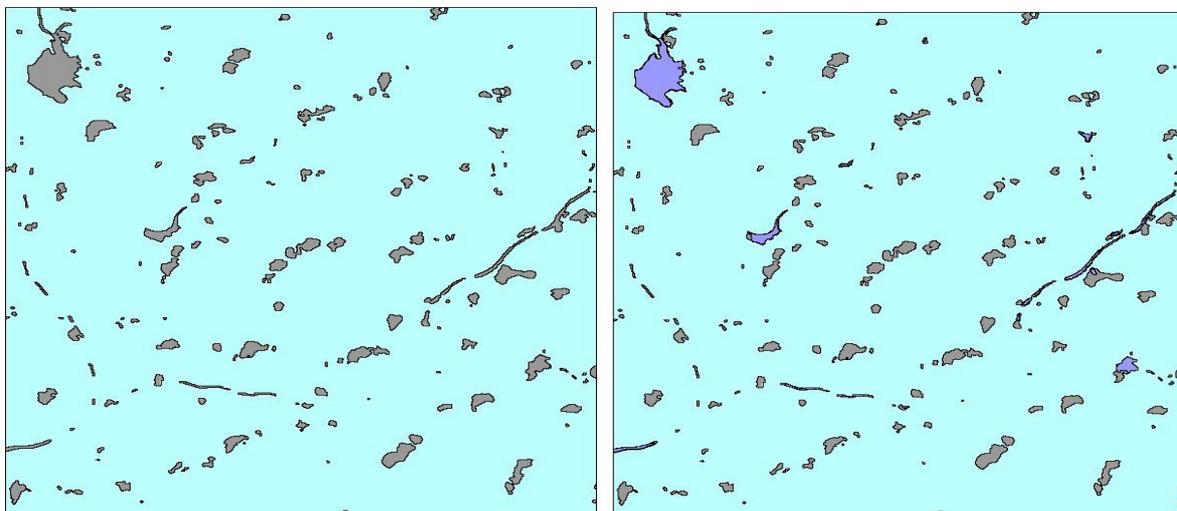


Figure 3: Mixed water body and cloud shadow and their separation

The sequence in which classes was extracted is shown in following flow-chart (Figure 4).

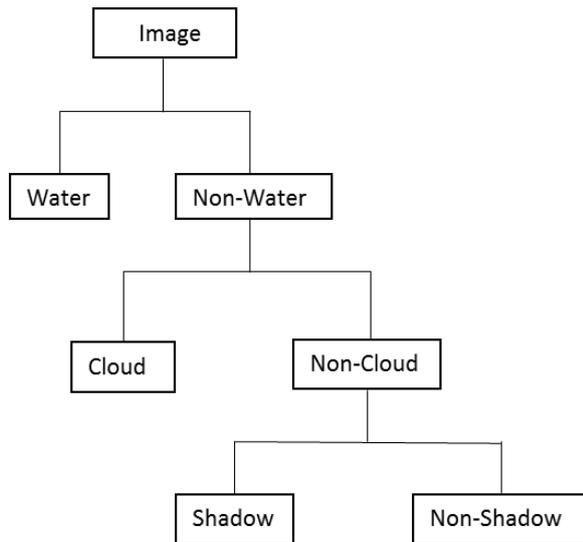


Figure 4: Flow-Chart: Extracted classes

For comparison of object and pixel based classification, pixel-based supervised classification was performed using ERDAS Imagine software. Supervised classification is shown in figure 5.

5. Accuracy Assessment

Accuracy assessment was performed visually. Classified water bodies, cloud shadows and clouds polygons were visually counted and compared to classes in data set taken as reference data for cloud shadow and cloud. For water

body, wetland boundary of year 2016 was used as a reference data (Figure 1). Error matrix for the object-based and supervised classification are shown in tables 2 & 3 respectively.

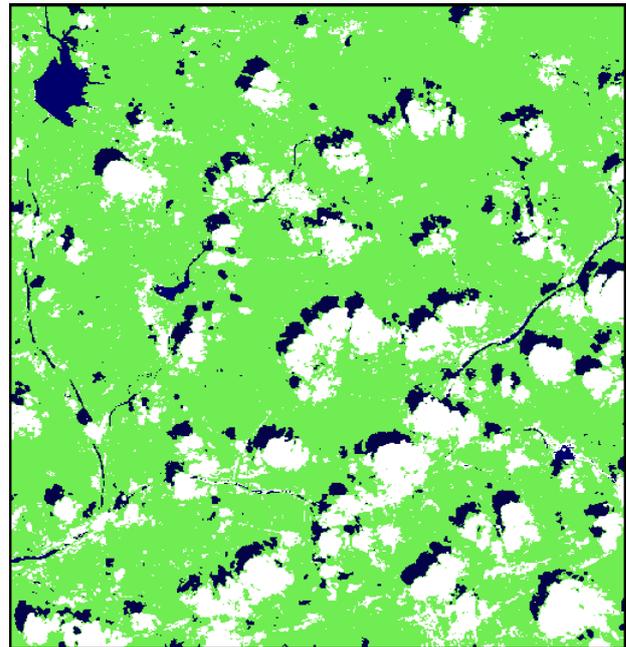


Figure 5: Supervised classification-water body (Blue), cloud shadow (Black) and cloud (White)

Producers accuracy, user’s accuracy, overall accuracy and kappa coefficient of object-based and supervised classification are shown in tables 4 and 5 respectively.

Table 2: Error matrix of object-based classification

	Water Body	Cloud Shadow	Cloud	Others	Row Total
Water Body	24	4	0	0	28
Cloud Shadow	12	60	0	0	72
Cloud	0	0	64	0	64
Others	0	0	0	0	0
Column Total	36	64	64	0	164

Table 3: Error matrix of supervised classification

	Water Body	Cloud Shadow	Cloud	Others	Row Total
Water Body	10	14	0	0	24
Cloud Shadow	26	50	0	0	76
Cloud	0	0	64	16	80
Others	0	0	0	0	0
Column Total	36	64	64	16	180

Table 4: Result of object-based classification

	Producers Accuracy	Users Accuracy	Overall Accuracy	Kappa Coefficient
Water Body	66.67	85.71	90.24	0.8473
Cloud Shadow	93.75	83.33		
Cloud	100	100		

Table 5: Result of supervised classification

	Producers Accuracy	Users Accuracy	Overall Accuracy	Kappa Coefficient
Water Body	27.78	41.67	68.89	0.5323
Cloud Shadow	78.13	65.79		
Cloud	100	80		

6. Results and Discussions

Error matrix of object-based and pixel-based classification are shown in tables 2 and 3 respectively. From table 2, it had been observed that 12 polygons of water body were misclassified in cloud shadow and 4 polygons of cloud shadow was included in water body. However, from table 2, it had been observed that 26 polygons of water body were misclassified in cloud shadow and 14 polygons of cloud shadow was included in water body. 16 polygons of other classes were also included in cloud class.

From the object-based classification, separation of water body and cloud shadow and detection of cloud was achieved (Figure 2). Water can be extracted using water body index and shadow can be identified using average DN value of all bands. Reliability and threshold value plays a crucial role in this study. Higher reliability value of water as compared to shadow makes possible of the separation of these two (Figure 3) while higher threshold value of cloud as compared to shadow makes possible the identification of cloud (Figure 2). Cloud was classified with 100% accuracy and overall accuracy and kappa coefficient was quite good (Table 3). From pixel-based (Supervised) classification, water body and cloud shadow are mixed (Figure 5) because of same spectral signature

associated of these two (Figure 6). Cloud was extracted but other (settlement) classes are also included in cloud class (Figure 5). Cloud was classified with 80% accuracy and overall accuracy and kappa coefficient was less (Table 5) as compared to object-based classification (Table 4).

It was observed that stream channel and water body extension was not fully detected (Figure 2).

7. Conclusion

In this study, separation of water body and cloud shadow and detection of cloud was achieved using object-based classification. Cloud shadow was separated from the water body and cloud was detected using three different parameters- DN value, reliability and threshold value with significant accuracy.

From visually assessed accuracy, it was found that classification for stream channel and extensions of water body were not fully achieved. Attempt is required to overcome this problem. This study was made over a specific area. Attempts are required to test this classification technique over a different area to see how the reliability and threshold value will affect the results.

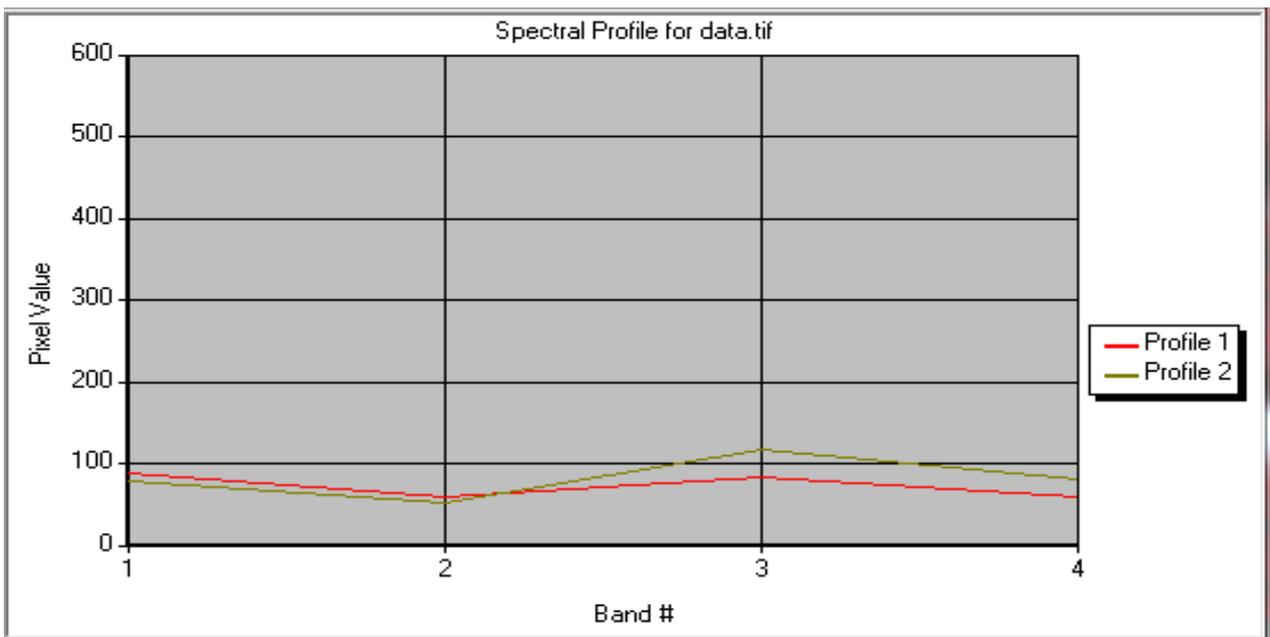


Figure 6: Spectral profile of water body and cloud shadow (Profile 1-water, Profile 2- Shadow)

Acknowledgements

Authors express their sincere gratitude to Sri D K Das, Director, Space Applications Centre (SAC) and Dr. Raj Kumar, Deputy Director, EPSA, SAC for their keen interest, guidance and encouragement. The authors wish to sincerely thank the reviewers for their critical comments and suggestions for improving the quality of the paper.

References

- Baatz, M., and A. Scape (2010). Multiresolution segmentation: An optimization approach for high quality multiscale image segmentation. AGIT-Symposium Salzburg 2000, pp. 12–23.
- Enderle, D. and R.C Weih (2005). Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification, *Journal of the Arkansas Academy of Science*, 59, 65-73.
- Flanders, D., M. Hall-Beyer and J. Pereverzoff (2003). Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. *Canadian Journal of Remote Sensing*, 29(4), 441–452. doi:10.5589/m03-006
- He, Y.R., X.X. Zhang and L.Z. Hua (2016). Object-based distinction between building shadow and water in high-resolution imagery using fuzzy-rule classification and artificial bee colony optimization, *Journal of Sensors*, 1–10. doi:10.1155/2016/2385039
- Heyman, O., G.G. Gaston, A.J. Kimerling and J.T. Campbell (2003). A per-segment approach to improving aspen mapping from high-resolution remote sensing imagery, *Journal of Forestry*, 101(4), 29–33.
- Inter Image operation guide (beta version 0.09).
- McFeeters, S.K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features, *International Journal of Remote Sensing*, 17(7), 1425–1432. doi:10.1080/01431169608948714
- National Wetland Inventory Assessment (2016-17), Space Applications Centre, ISRO, Ahmedabad.
- Ryherd, S and C. Woodcock (1996). Combining spectral and texture data in the segmentation of remotely sensed images, *Photogrammetric Engineering and Remote Sensing*, 62(2), 181–194.
- Willhauck, G. (2000) Comparison of object oriented classification techniques and standard image analysis for the use of change detection between SPOT multispectral satellite images and aerial photos, In: *International Archives of Photogrammetry and Remote Sensing*, XXXIII, Supplement B3.
- Yu, Q., P. Gong, N. Clinton, G. Biging, M. Kelly and D. Schirokauer (2006). Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery, *Photogrammetric Engineering & Remote Sensing*, 72(7), 799–811. doi:10.14358/PERS.72.7

An approach for generation of multi temporal co-registered optical remote sensing images from Resourcesat-2/2A sensors

Indranil Misra*, Vivek Sharma, S. Manthira Moorthi and Debajyoti Dhar

Space Applications Centre, ISRO, Ahmedabad 380015, India

*E-Mail: indranil@sac.isro.gov.in

(Received: Jan 02, 2019; in final form: Jun 17, 2019)

Abstract: Multi temporal remote sensing data availability is crucial for remote sensing community to study the planet earth. The multi temporal data can be used for scientific studies only after sub-pixel level image registration of all multi date acquisitions. Multi temporal co-registered remote sensing data is essential for many real time applications such as crop forecasting, forestry, inland water mapping, change detection and time series analysis. Multi temporal co-registered data is the first pre-processing step for generation of Analysis Ready Data (ARD) cube that allows immediate analysis with minimal user efforts. Image Registration at sub-pixel level requires optimal geometric transformation of all datasets such that data stack is geometrically aligned over each other. The major step in image registration is feature detection to generate a collection of tie points with outliers removed, match the feature points and finally estimate the transformation parameters. In this paper, we present an approach for multi temporal image registration that employs Scale Invariant Feature Transform (SIFT) technique along with a segmented affine based transformation model for different image segments to correct the data at geographic coordinate space to achieve sub-pixel level geometric accuracy. The key purpose of this work is to generate co-registered data stack for image analysis. Resourcesat-2/2A (RS-2/RS-2A) LISS-3 data from Indian Remote Sensing Satellite (IRS) is used for multi temporal co-registration task. The RS-2/2A LISS-3 data have spatial resolution of 24 meters and combination of both RS-2 and RS-2A LISS-3 gives better temporal repetivity to cover the same region in less number of days. The technique developed is tested with LISS-3 data of same region acquired during time interval from 2012 to 2018.

Keywords: LISS-3, SIFT, ARD, Affine Transform, Image Registration, Multi Temporal

1. Introduction

Remote Sensing data acquisitions at different timeline encourages researchers to study the feature changes that can be used for diverse space applications such as agricultural monitoring, disaster management and has the capability to solve complex problems related to earth surface studies (Haiganga and Zhou, 2008). The multi temporal data analysis requires robust preprocessing of data stack to generate sub-pixel level image registration accuracy. This open new door for development of next generation methods of image registration for multi spectral remote sensing data. The critical steps in image registration is to detect stable tie points by removing outliers and estimation of transformation model that establishes a mapping between the images. The image registration task is the most important requirement of remote sensing world where large amount of multi temporal satellite images need to be co-registered for every kind of scientific studies. The data stack generated should be evaluated for both relative and absolute geometric accuracies. The relative geometric accuracy helps us to know the image-to-image registration performance of the stack whereas absolute geometric accuracy of the stack can be evaluated with reference images or ground control points (GCP) to measure the location of an object in data with respect to its true location on the earth.

1.1 Feature detection and description using SIFT

The image registration process starts with efficient feature detection mechanism. Detecting the appropriate feature point is the crux of image registration problem (Aksakal, 2013; Zitova and Flusser, 2003). In literature, many kind of interest point detectors are reported to detect stable

feature points (Schmid et al., 2000). But Scale Invariant Feature Transform (SIFT) is one of the robust feature detection technique that gives better performance and it is invariant to scale, rotation, illumination conditions and image noise (David, 1999). The advantage of SIFT over other feature detector is that it is not only detect the feature points but it also provides the description of the feature point in comparison to the other feature detectors like Harris Corner. SIFT is also capable to detect feature points at sub-pixel level which is needed for accurate transformation parameter estimation.

The SIFT algorithm starts with construction of scale space extrema (David, 1999) using Gaussian Kernel (G). The eq. 1 shows the relation between Smooth Image (L) and Input Image (I) at point (x, y)

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

$$\text{Where, } G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

It has been found that Laplacian of Gaussian (LOG) is able to provide stable features and give excellent notion of scales but it is computationally costly. So the Difference of Gaussian (DOG) pyramid is created which is considered to be the close approximation of Laplacian of Gaussian (LOG) (David, 1999).

The next task is to scan each DOG image to locate the minimum and maximum around all the neighboring points including the scale (Zheng et al., 2008). Figure 1 shows the DOG pyramid generated from scale space and extrema located from the DOGs.

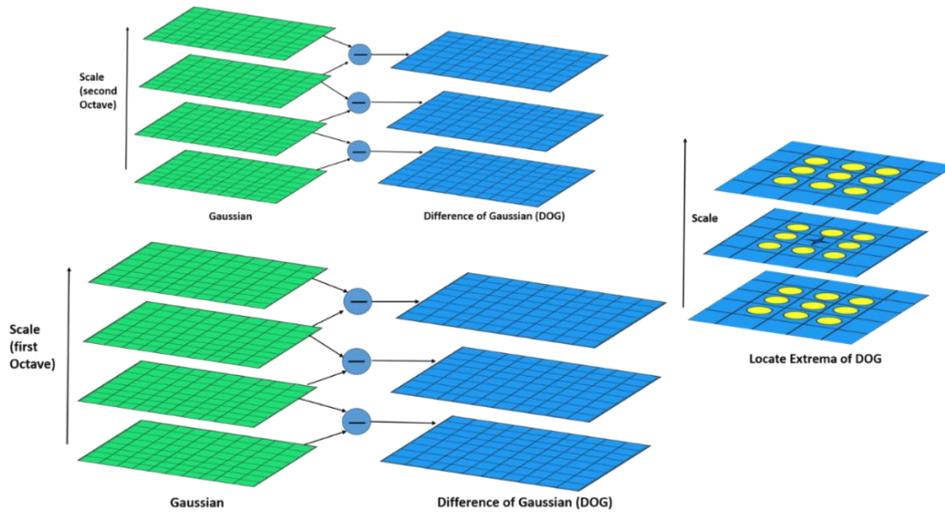


Figure 1: DOG pyramid and extrema localization

The potential stable feature point at sub-pixel level localization and initial outliers can be removed by using Taylor Series approximation

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D^T}{\partial x^2} x \quad (3)$$

and differentiate set to zero

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (4)$$

to get location in terms of (x, y, σ) .

The outliers that are still left can be removed by doing contrast limit filtering and edge response elimination. The key points orientation is computed using gradient of each blurred image. The orientation assignment achieves rotation invariance. It mainly computes central derivatives, gradient magnitudes (m) and direction (θ) of smooth image (L) at the scale of keypoint (x, y) . The weighted direction histogram in a neighborhood of a keypoint in form of bin is created and finally select the peak as direction of the keypoint.

$$m(x, y) = \sqrt{\frac{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}{2}} \quad (5)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (6)$$

The descriptor is built by sampling the point around keypoints. Rotate the gradients and coordinates by previously computed orientation. Then separate the region into sub-region. The histogram is created for each sub-region with specified bins. The descriptor is stored as element vector for each keypoints. The feature points are detected using SIFT independently for both input image and reference image. The matching finds the nearest neighbor i.e. a keypoint with minimum Euclidean Distance (Rabin et al., 2008).

2. Segmented affine transformation estimation model

The input remote sensing image is radiometrically corrected and each pixel is tagged with latitude/longitude (lat/lon) information. In this approach, the matched

featured points are divided into multiple segments and affine transformation model is computed for each segment. The affine transform (Zhou, 2016) for a segment is shown as:

$$aff(k) = \{a_{0k}, a_{1k}, a_{2k}, a_{3k}, a_{4k}, a_{5k}\} \quad (7)$$

Where k = segment number and a_{ik} = affine transform parameters

The tagged input (lat, lon) of image point (x, y) transformed to corrected geo-point (lat_{cor}, lon_{cor}) depending upon the segment number k of the point (x, y) .

$$\begin{bmatrix} lat_{cor} \\ lon_{cor} \end{bmatrix} = \begin{bmatrix} a_{0k} & a_{1k} \\ a_{2k} & a_{3k} \end{bmatrix} \begin{bmatrix} lat \\ lon \end{bmatrix} + \begin{bmatrix} a_{4k} \\ a_{5k} \end{bmatrix} \quad (8)$$

Figure 2 shows an optical remote sensing data divided into four segments in pixel wise direction. Each segment contains matched feature points which is used for affine model estimation for that segment. Segment size is determined on the basis of input image size and pixel size. The four different affine models estimated for the segmented image regions and geographic grid is modified using its corresponding affine transformation model shown in eq. 8 to generate a single final modified geometric grid that maintains the continuity of transitions in the image segments. The output grid generated with segmented affine model approach reduces internal distortion in the final geo-corrected data and provide sub-pixel level registration accuracy with respect to the reference image.

3. Processing workflow developed

The approach described in above sections need to be realized in a data processing workflow for generation of the multi temporal co-registered products. The input to the workflow is multi spectral radiometrically corrected remote sensing data with tagged geographic coordinate information. Using the latitude and longitude coordinates, corresponding reference tile need to be fetched from reference database as shown in figure 3.

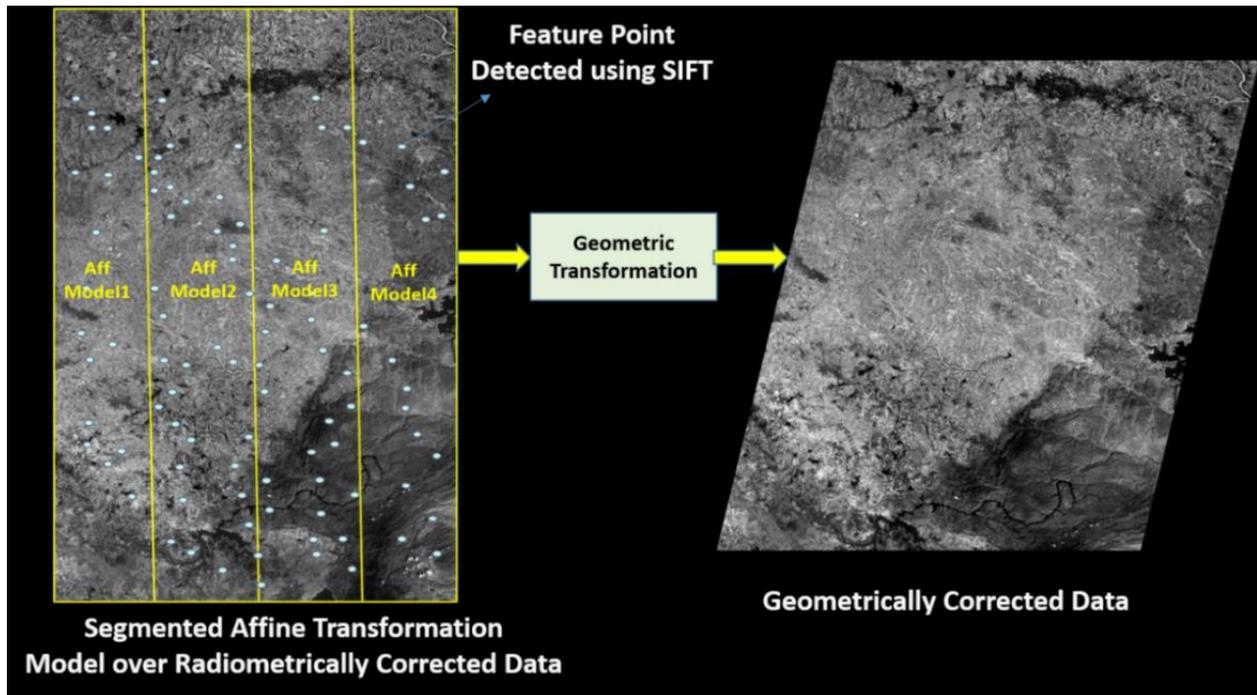


Figure 2: Segmented affine transformation model over remote sensing data

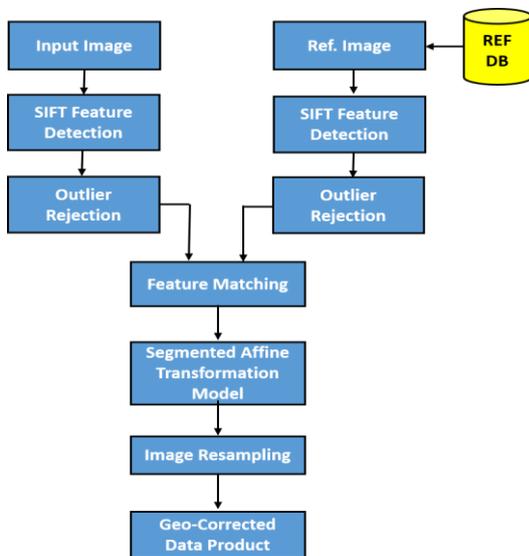


Figure 3: Data processing workflow

The reference tiles are stored as standard geo-tiff ortho-rectified images in UTM map projection. The absolute location accuracy of reference images is less than 12 meters and it remain same for both RS-2/RS-2A. SIFT features are detected in both the images and matched to get the set of key points required for estimation of affine transformation. The segmented affine transformation models are generated as per the approach in section-III. The input image is geometrically transformed using image resampling technique at final stage of data processing to generate geometrically corrected data product. Since resampling kernel is used for input to output transformation. Figure 3 shows the processing workflow developed for multi temporal image registration. The execution time of the algorithm is less than a minute for a scene of 6000 lines*6000pixels image size

4. Data used and results achieved

The processing workflow developed is tested with recent acquisitions of LISS-3 data on board Resourcesat-2A. Resourcesat-2A satellite was launched by ISRO in Dec. 2016. The LISS-3 sensor is primarily used for earth resource monitoring that includes agriculture, forest and other natural resource observation and planning. The details of sensor, data used for testing and its geo-meta information is shown in table 1. Figure 4 shows the false color composite (FCC) image of the study area acquisition by Resourcesat-2A LISS-3 (L3). The multi date acquisitions of the same study area (Figure 4) is considered for image registration exercise.

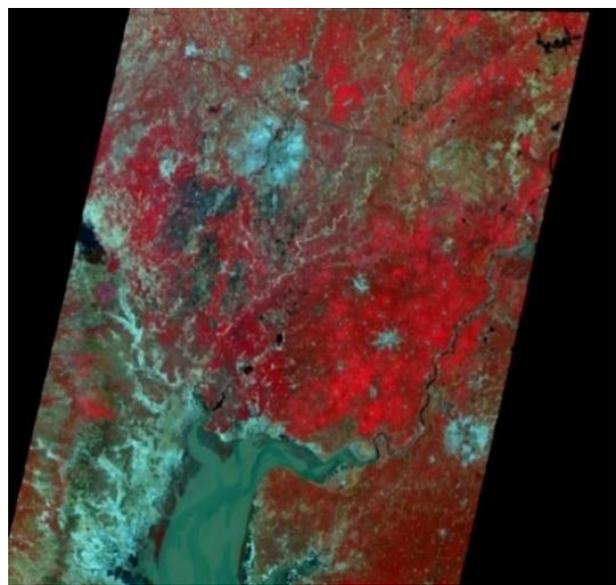


Figure 4: RS-2A LISS-3 FCC image (Path/Row: 93/56)

Table 1: LISS-3 sensor and data details

Spatial Resolution (Meters)	24.0
Swath (Kilometers)	141
Spectral Bands (Microns)	0.52-0.59 0.62-0.68 0.77-0.86 1.55-1.70
Quantisation (Bits)	10
Date Rate (MBPS)	105
Revisit Time (in Days)	24
Path/Row Scene (Referencing Scheme) Used for testing	93/56
Region (Cities) Covered	Ahmedabad, Gandhinagar, India
Multi Date Cloud Free Data Acquisitions	1. 25 th Dec 2016 2. 18 th Jan 2017 3. 31 st March 2017 4. 18 th May 2017
Geographic Corner Coordinates (lat/lon) of Scene	UL: 23d24'16.04"N, 71d46'20.12"E UR: 23d25'54.53"N, 73d37'49.12"E LR: 21d53'27.37"N, 73d38'43.88"E LL: 21d51'56.03"N, 71d48'28.95"E

The RS-2A LISS-3 multi date datasets of the scene information tabulated in table 1 are processed using the workflow developed. The performance of registration is examined both visually and quantitatively for detailed

analysis. The visual inspection is carried out by swiping the registered product over reference tile. All the dates are seen at same region with reference tile and the geometric fidelity shows all the images are registered over reference image. Figure 5 shows the image swipes of all the four dates listed in table 1 with reference at sub-pixel level zoom view.

The image swipes help us to visualize the registration performance at particular regions in the image, which is not enough to quote the final registration accuracy. The quantitative parameters are needed to conclude the overall image registration performance. The evaluation of registered product can be done by computing root mean square error (RMSE). The RMSE is calculated by mapping the relation of stable matched feature points obtained from registered data and reference data [9]. The RMSE computed is an indicator to show that relative geometric accuracy between the images. The SIFT detector is used here to fetch the feature points. The matched correspondence is tabulated in table 2 and it shows the RMSE is within a pixel for all the datasets with respect to reference.

Table 2: Registration Performance Table

Date of Acquisition	No. of Control Points	RMSE (in pixels)
25 th Dec 2016	300	0.44
18 th Jan 2017	454	0.38
31 st Mar 2017	345	0.42
18 th May 2017	286	0.46

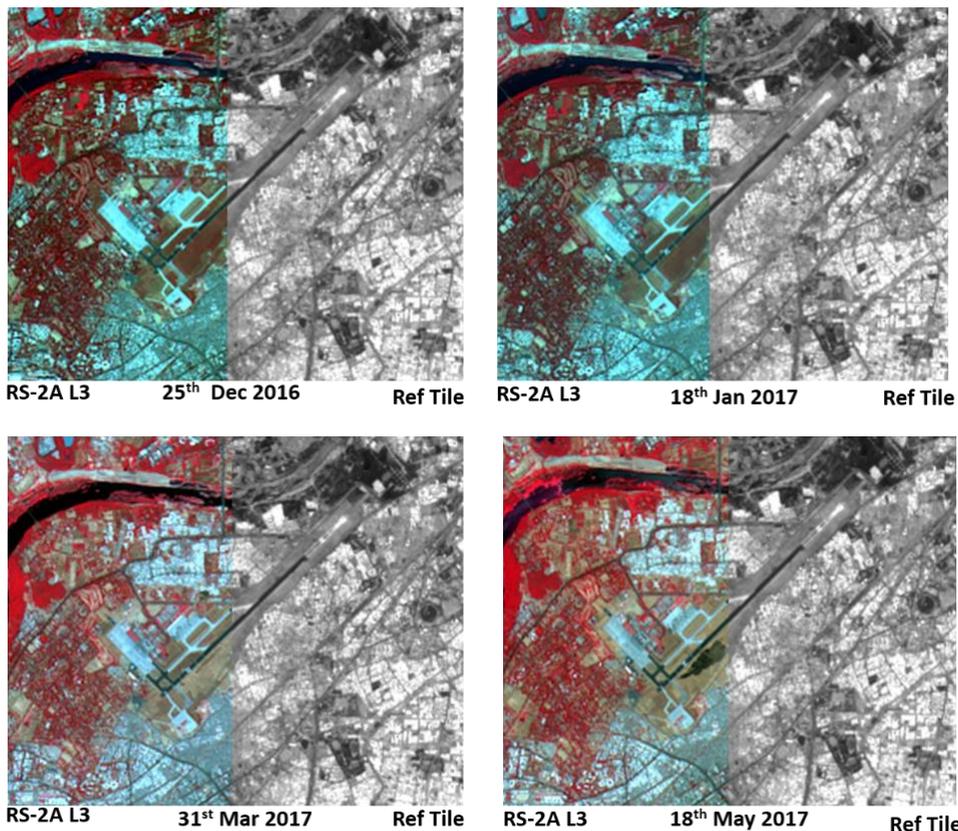


Figure 5: Multi date image swipes (Overlay) against Ref Tile

In addition, multi temporal products in combination of Resourcesat 2 and 2A from the year 2012 to 2018 generated using the processing workflow covering mixed terrain also includes partially cloudy data. The number of multi temporal products considered for analysis is 24 across the span of 6 years acquisition of Resourcesat mission. It is observed that all the data products are well registered within a 0.5 pixel across the images with Band-3 of 23-May-12 as reference shown in figure 6. Also absolute location accuracy of all multi temporal data products is within a pixel. The performance of the registration is checked at multiple patches dominated by urban area, hilly terrain, agriculture land and vegetation cover as shown in figure 7. The RMSE in pixel is computed for the patches with respect to the reference and shown in table 3. All 24 LISS-3 multi temporal registered data products stack meets the specifications of less than 0.25 pixels band-to-band registration (BBR) between the multi spectral channels and also internal distortion is at sub-pixel level which is needed for any kind of change detection studies and analysis. The approach developed is tested in multi temporal coastal scenes dominated by ocean and it has been found that relative registration error is at sub-pixel level registration accuracy

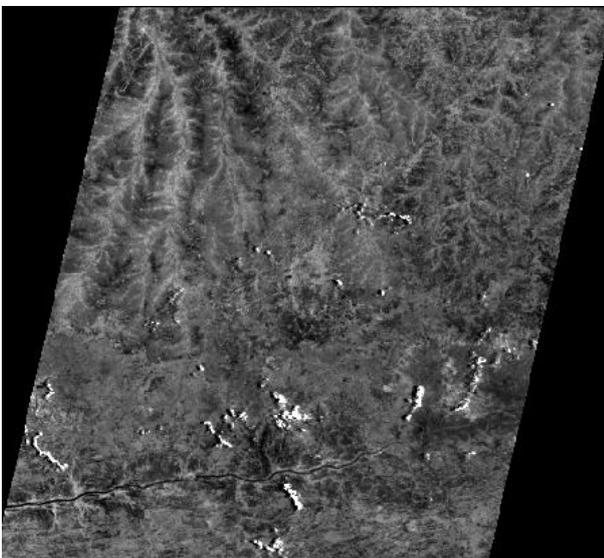


Figure 6: RS2 LISS-3 BAND-3 23 May 2012 (Path/Row: 93/56)

Table 3: Registration Accuracy at different patches

Regions	No. of Control Points	RMSE (in pixels)
Urban	60	0.47
Hilly	23	0.53
Agriculture	18	0.42
Vegetation	34	0.45

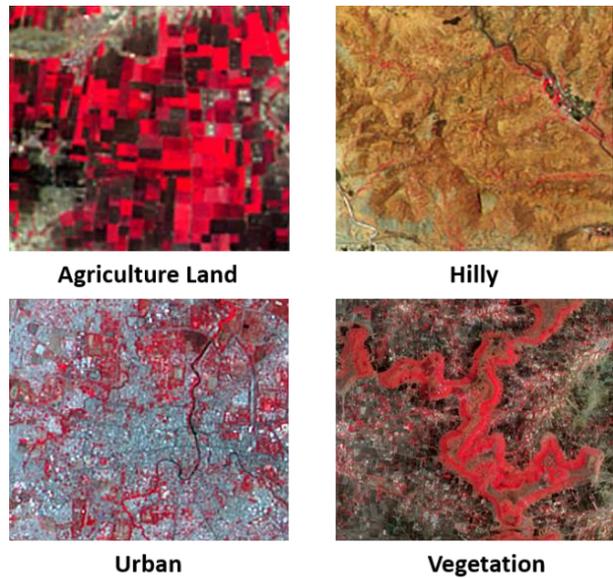


Figure 7: Multiple patches extracted from image to check registration accuracy

5. Conclusion and future work

The image registration for multi temporal data sets is achieved at sub-pixel level accuracy, which is evaluated both visually and quantitatively. SIFT with segmented affine transformation model is the approach developed for co-registration of multi date remote sensing data acquisitions. The co-registered data becomes analysis ready data product which can be used for any kind of earth monitoring studies. The workflow need to be tested with many other multi date acquisitions covering different terrains. The next task is to compare SIFT based approach developed with other feature detection techniques. At present the processing workflow handles same sensor multi temporal data. The future work would include multi-sensor data registration from different satellite platforms and different sensors.

Acknowledgements

The authors acknowledge Shri D K Das, Director, Space Applications Centre (SAC), ISRO for his encouragement. The authors also thank other team members of optical data processing for their support during the course of work.

References

- Aksakal, S.K. (2013). Geometric accuracy investigations of SEVIRI High Resolution Visible (HRV) Level 1.5 Imagery, *Remote Sensing*, (5), 2475-2491; doi: 10.3390/rs5052475.
- David, G. L. (1999). Object recognition from local scale-invariant features, *International Conference on Computer Vision*, 1150-1157.
- Haiganga, M.G and Q. Zhou (2008). A review of multi-temporal remote sensing data change detection algorithms, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XXXVII. Part B7. Beijing, pp. 757-762.

Rabin, J., J. Delon and Y. Gousseau (2008). A statistical approach to the matching of local features, *IEEE ICPR* (40, 41).

Schmid, C., R. Mohr and C. Bauckhage (2000). Evaluation of interest point detectors, *International Journal of Computer Vision*, 37(2), 151–172.

Zheng, Z., C. Zhiguo and X. Yang (2008). Multi-spectral remote image registration based on SIFT, *Electronics Letters*, 44, 107-108.

Zhou, J. (2016). A novel image registration algorithm using SIFT feature descriptors, *International Conference on Smart City and Systems Engineering*, November 25-26, 2016.

Zitova, B. and J. Flusser (2003). Image registration methods: A survey, *Journal of image and Vision*, Elsevier, 21, 977-1000.

Symposium Report

National Symposium on Advancements in Geospatial Technology for Societal Benefits

December 05-07, 2018

Jointly organised by ISG & ISRS

Hosted by Space Applications Centre, ISRO, Ahmedabad



The National Symposium on Advancements in Geospatial Technology for Societal Benefits (ISG-ISRS-NS-2018) and Annual Conventions of Indian Society of Geomatics (ISG) and Indian Society of Remote Sensing (ISRS) with Silver Jubilee Celebrations of the Indian Society of Geomatics was organized at Space Applications Centre (SAC), ISRO, Ahmedabad during December 05-07, 2018. The Symposium was jointly organised by Indian Society of Geomatics, Ahmedabad and Indian Society of Remote Sensing, Dehradun and hosted by Indian Society of Geomatics, Ahmedabad Chapter (ISG-AC), Indian Society of Remote Sensing, Ahmedabad Chapter (ISRS-AC) and Space Applications Centre, ISRO, Ahmedabad.

Pre-Symposium tutorials on five themes were organised during December 03-04, 2018 at SAC, Bopal Campus, Ahmedabad and CEPT University, Ahmedabad. Inaugural session was conducted at SAC, Bopal Campus. Dr Raj Kumar, DD, EPSA, SAC, Ahmedabad presided the function. Theme 1: Applications of IRNSS & GPS in Navigation & Mapping was coordinated by Dr. Ashish Shukla, SAC which had around 13 participants, Theme 2: Planetary Science and Data Processing was coordinated by Dr. A. S. Arya, SAC which had around 12 participants, Theme 3: SAR data Processing and Applications was coordinated by Dr. Arundhati Misra, SAC which had around 11 participants, Theme 4: Big Data Analytics and Machine Learning was coordinated by Shri Shashikant Sharma, SAC which had around 14 participants, and Theme 5: Sustainable Smart Cities: Geospatial Solutions, was coordinated by Prof. Anjana Vyas, CEPT University, which had around 20 participants. All the tutorials were well received. The valedictory session of all the tutorials was conducted at SAC, Bopal Campus on December 4, 2018.

The Symposium was inaugurated on December 05, 2018 at the Yash Pal Auditorium of SAC in the august presence of the Chief Guest, Shri P.P. Kale, Former Director, VSSC & SAC and the Guest of Honour, Lt. General Girish Kumar, VSM, Surveyor General, SOI. The inaugural ceremony started with invocation, lightening of Lamp and Welcome remarks by Shri Nilesh Desai, Associate Director, SAC. Shri D K Das, Director, SAC could not come but he expressed his delight in being the host for ISG-ISRS-NS-2018, complimented the organizing team and wished a grand success to the Symposium. The Presidential address was given by Shri Tapan Misra, President, ISG & ISRS, Senior Advisor, ISRO. Lt. General Girish Kumar, VSM, Surveyor General, SOI also addressed the august gathering as Guest of Honour.

After the inaugural ceremony, many prestigious awards of ISG and ISRS were given by the Chief Guest, Guest of Honour and President ISG & ISRS. Various ISG Awards given were: National Geomatics Award for Excellence, 2016 to Prof. Arup Ranjan Dasgupta, Deputy Director (Retd.), SAC, Ahmedabad; National Geomatics Award-Technology, 2018 to Dr. Harish Chandra Karnatak, Scientist, IIRS, Dehradun; President's Appreciation Medal for Contribution to the ISG, 2018 to Shri N.S. Mehta and Shri K.P. Bharucha, Scientists (Retd.), SAC, Ahmedabad; Prof. Kakani Nageswara Rao Endowment Young Achiever Award, 2018 to Dr. S.V.V. Arun Kumar Surisetty, Scientist, SAC, Ahmedabad; ISG Best Chapter Award, 2018 to ISG-Ludhiana Chapter. Various ISRS Awards were: Bhaskara Award, 2017 to Shri Sanatanu Chowdhury, Director, NRSC, Hyderabad and Dr. Ajai, Group Director (Retd.), SAC, Ahmedabad; Satish Dhawan Award, 2017 to Dr. Raj Kumar, Outstanding Scientist and Deputy Director, SAC, Ahmedabad and Dr. S.N. Omkar, IISc, Bengaluru; National Geospatial Award for Excellence-2017 to Prof (Er.) J.K. Ghosh, IIT-Roorkee; P.R. Pisharoty Memorial Award, 2017 to Shri Satadru Bhattacharya, Scientist, SAC, Ahmedabad and Dr. Hitendra Padalia, Scientist, IIRS, Dehradun; Indian National Geospatial Award, 2018 to Dr. John Mathew, Scientist, ISRO HQ, Bengaluru, and Mr.

Shridhar D. Jawak, Former Project Scientist, NCPOR, Goa and currently RS Officer, SIOS, Svalbard; President's Appreciation Medal, 2018 to Dr. R.N. Sahoo, Scientist, IARI, New Delhi; ISRS Best Chapter Award, 2017 to ISRS-Visakhapatnam Chapter. Dr. S.K. Basu, Scientist (Retd.), SAC, Ahmedabad was conferred ISRS-Fellowship, 2018. Best Papers awards in the Journal of Indian Society of Remote Sensing were received by Dr. Arundhati Misra, Group Director, SAC, Ahmedabad and Dr. Manisha Gupta, Scientist, SAC, Ahmedabad.



Chief Guest Shri P.P. Kale, Former Director, VSSC & SAC lighting up the Lamp during the Inaugural Function



Guest of Honour Lt. General Girish Kumar, VSM, Surveyor General, SOI addressing the delegates

Thereafter, address by Chief Guest Shri P.P. Kale, Former Director, VSSC & SAC followed by ISG-Millennium Lecture on "Sustainable Development Goals: Role of Geomatics" was delivered by Shri Kartikeya V. Sarabhai, Director, Centre for Environment Education (CEE), Ahmedabad.

The Symposium Souvenir cum Abstract Volume, recent issues of Journal of Geomatics and ISG-Newsletter were released. Release of Book "पृथ्वी के प्रेक्षण से संबंधित कैमरों के निर्माण में उत्कृष्टता की दिशा में भारत की यात्रा" authored by Prof. George Joseph was also done and a brief about the book was given by him. Vote of thanks was proposed by Dr. C.P. Singh, Organising Secretary.



Shri Kartikeya V. Sarabhai, Director, CEE, Ahmedabad delivering ISG-Millennium Lecture



Dignitaries releasing the Souvenir

The Symposium attracted a huge gathering of over 350 delegates from various scientific institutes, academia, industries and students from different parts of the country. Members of the industry sponsored and showed keen interest in utilizing the forum to their benefit. Three Plenary Sessions with a total of five invited talks by distinguished speakers viz., Dr. P.G. Diwakar, Director, EDPO, ISRO-HQ, Bengaluru, Shri Nilesh Desai, Associate Director, SAC, Ahmedabad, Dr. Sanatnu Chowdhury, Director, NRSC, Hyderabad, Dr. V.N. Tiwari, Director, NGRI, Hyderabad, Prof. P. Janardhan, Dean, PRL, Ahmedabad were organised. Plenary Sessions also included six Industry presentations.

The Technical Sessions were organized into Parallel Sessions comprising of Eighteen (Oral) and Three (Poster) Sessions under the Themes viz., Web GIS, Geosciences, Agriculture, Navigation and Mobile Apps, Planetary Sciences, Environment, Geospatial Technology, Soil/Land Degradation, Machine Learning, Forestry, Coastal Applications, Hyperspectral Applications Urban Applications, Marine Sciences, Microwave Applications, Advancements in Data Processing, Cryosphere and Water Resources. Several young researchers and students from all parts of the country also participated enthusiastically. Overall, there were 227 contributory papers out of which 156 papers were presented in the Symposium.



Group photo of the Symposium

The 23rd annual general body meeting (AGM) of ISG took place on December 5, 2019 at Yashpal Auditorium of SAC, Ahmedabad and various matters of society were discussed. The AGM was followed by Silver Jubilee Celebration event of the society. The founder members, all current and past office bearers of ISG were felicitated during this event. A silver jubilee celebration memoir book was also released during this occasion. Aavishkar Group of Ahmedabad performed classical Gujarati folk dances as part of the cultural evening during Silver Jubilee celebration of ISG.

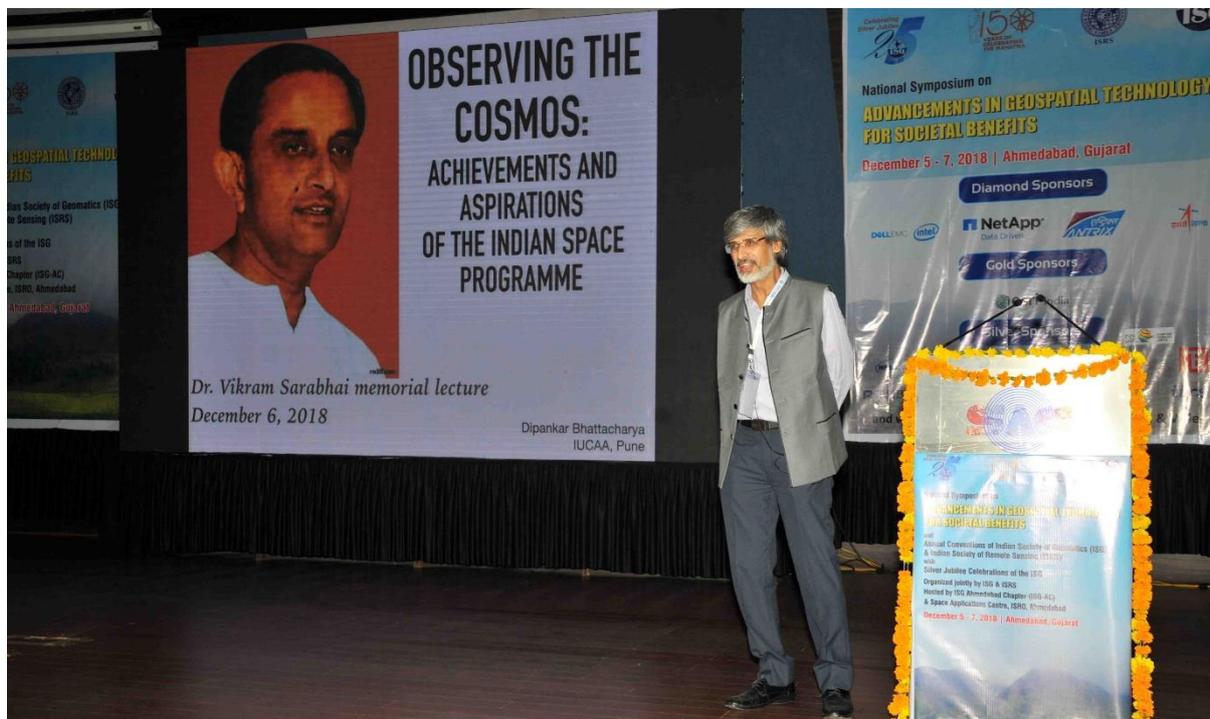


Release of Silver Jubilee Celebration Memoir of ISG



Aavishkar Group performance during Cultural evening as part of Silver Jubilee Celebration of ISG

Dr. Vikram Sarabhai Memorial Lecture of ISRS was delivered by Dr. Dipankar Bhattacharya, Distinguished Professor and Dean, Core Academic Programmes, Inter-University Centre for Astronomy and Astrophysics (IUCAA), Pune on the topic, "Observing the Cosmos: Achievements and Aspirations of the Indian Space Programme". The 45th AGM of ISRS was organised on December 6, 2019 at Yashpal Auditorium of SAC, Ahmedabad and society matters were discussed.



Dr. Vikram Sarabhai Memorial Lecture of ISRS delivered by Dr. Dipankar Bhattacharya, Distinguished Professor and Dean, Core Academic Programmes, Inter-University Centre for Astronomy and Astrophysics (IUCAA), Pune

The Concluding Session of the Symposium was Chaired by Shri Tapan Misra, President, ISG & ISRS and Senior Advisor, ISRO. A group of experts participated in the panel discussions. Symposium recommendations were discussed and finalised. Recommendations are given in subsequent section.



Panel Discussion and Concluding Session

Dr. A.S. Rajawat, Chairman, Technical Programme Committee announced awards for best paper presentation in students and general categories. In the General category, Shri Indranil Misra (First); Dr. Rahul Nigam (Second); Dr. Amit Kumar Dubey and Dr. Surisetty Arun Kumar (Third) and Shri Rishitosh K. Sinha (Best Poster) were awarded. In the student category, Ms. Garima Sodha (First); Shri Harsh Agrawal (Second); Shri Dhwanilnath Gharekhan (Third); Ms. Sadaff Altaf Bakshi, Ms. Simone Darji, Ms. Anasua Chakraborty (Consolation) were awarded in oral presentation, whereas Ms. Maya Raghunath Suryawanshi (First); Shri Tariq Abdullah and Ms. Alpna Revdandekar (Second), Ms. Shreekumari Patel (Third) were awarded in the poster presentations. A huge appreciation was conveyed from all the participants to the Symposium Organising committee for making the event a grand success.



Group photo of the Symposium Organising Committee

Symposium Recommendations

Agriculture

Noting that:

Increasing requirements of geospatial applications for crop insurance including mix of crop types, abiotic-biotic stress detection & quantification, solutions to agro-industries, soil health, crop water requirement,

Realizing that:

Capability of developing micro satellites, advanced sensors related to imaging spectroscopy and Synthetic Aperture Radars,

Recommending that:

Constellation of very high resolution satellites in multispectral, hyperspectral, thermal, microwave and UAV-based imaging are essential to solve the complexities of agriculture and related sectors for Block-to-village-farm level assessment and planning.

Forest and Environment

Noting that:

Forest carbon sequestration and climate change impact especially in hilly region, forest species mapping, forest risk zoning, need of seamless India-specific air quality index, development of alert system for dust events and impact assessment on Ecosystem health,

Realizing that:

Availability of long-term observational records from multiple satellite sensors as well as data availability of recent advanced global climate missions and existing *in-situ* measurement network,

Recommending that:

Environmental Satellites to monitor Essential Climate Variable (ECVs), Ecosystem health and Air Quality are needed to fill up the gaps in global climate missions.

Geosciences

Noting that:

Emphasis is on identifying hydrothermal alteration zones, improving mineral exploration, understanding geo-hazards,

Realizing that:

The potential of space-borne hyperspectral sensors, SAR interferometry, integration of in-situ measurements and advanced modelling for the above requirements,

Recommending that:

Advanced sensors and techniques in geological and structural mapping and modeling, improved 3-D terrain modeling, strengthening of insitu data collection, early warning systems for geo-hazards need to be developed.

Water*Noting that:*

There is strong requirement of monitoring river / inland water level, quality, surface and ground water storage, water harvesting, recharge and quality, flood warning and vulnerability,

Realizing that:

Assimilation of satellite data from multiple sensors into distributed hydrological model has the potential to address the above requirements,

Recommending that:

Constellation of satellites (H-Train) is required with multisensors capability (microwave - Altimeters, radiometers, Multispectral, hyperspectral and thermal imaging and wide-swath high repeat SAR) along with a dense network of in-situ sensors connected through IoT.

Cryosphere*Noting that:*

There is need to understand Himalayan and Polar Cryosphere in the context of climate change and associated role in hydrology, avalanche forecast and glacier lake outburst modelling,

Realizing that:

There is immense potential of earth observation satellites in monitoring and understanding cryosphere processes,

Recommending that:

High resolution multi-temporal DEM for measuring changes in elevation of glacier surfaces, high resolution, wider swath and multi-temporal multispectral sensors for glacier dynamics, hyperspectral sensors for snow-pack characteristics and multi-frequency SAR for snow/glacier dynamics and depth estimation are required.

Development of modelling framework by assimilating satellite derived cryosphere elements and in-situ measurements (ground network and field expeditions) for understanding snow, glacier and sea-ice dynamics and related disaster applications.

Urban*Noting that:*

There are increasing requirements of Urban sprawl and growth forecasting, heat island monitoring (UHI), built up material composition, roof-top solar energy under ‘Smart City’ programmes,

Realizing that:

Increasing population growth, rural-urban transformation, increasing structures and lesser green cover, are major concerns of sustainable development goals (SDGs) and related action plans,

Recommending that:

LiDAR based observations, UAV based imaging, hyperspectral sensing, thermal imaging are required for 3D city modelling, heat island expansion monitoring, aiding green energy measures. Data collection through crowd sourcing and geospatial models for sustainable urban growth are required.

Marine, Ocean dynamics & Coastal*Noting that:*

There is a strong requirement of off-shore wind farm site suitability, beach erosion forecast, river mouth sediment plume dynamics, rip current probability, shore-line changes, detection of new marine production and community structure, medicinally beneficial coral macro-algae, pigment composition and fish productivity,

Realizing that:

Capability exists to retrieve coastal and marine bio-geo-chemical parameters, assimilate in related models, simulate as well predict various processes,

Recommending that:

There is a need to have high temporal resolution sensing with appropriate spectral channels from geo-stationary platform for improved detection, estimation of bio-geo-chemical parameters, supported by appropriate network of calibration and validation sites. Process based models is required to understand coupling of various parameters of bio-geo-chemical cycle.

Planetary Sciences and Astronomy*Noting that:*

There is a need to have in-depth studies on crater morphology, new minerals, water presence, flow characteristics, planetary atmosphere characterization leading to understanding of various forms, structures, processes and evolution of planets,

Realizing that:

There is large potential of planetary research using data from various Indian and Global Missions,

Recommending that:

More capacity building is required to understand and analyse data from various Planetary Missions (including data from Chandrayaan-2 and future Venus and Mars Missions) and thrust is required to develop new techniques of data interpretation and bring out new scientific findings.

Advanced Data Processing*Noting that:*

There is a global requirement to provide Analysis Ready Data (ARD) from Indian sensors such as surface reflectances, surface temperature, Sigma naught,

Realizing that:

Majority of global EO data sets are provided in harmonised data cubes and has demonstrated various applications,

Recommending that:

Data cubes, harmonized data from Indian satellites with common georeferenced, operational products should be delivered to users through improved data processing tools with less Turn-around-Time (TAT).

Web GIS, Navigation, Mobile Apps*Noting that:*

There is an increasing trend to get online access of geospatial data, information, forecasts, advisories,

Realizing that:

Different AI tools, IoTs, Big data analytics, Crowd sourcing, satellite navigation, cloud computation has bigger role to play,

Recommending that:

Web-based geoinformatics and apps connected to these should be encouraged for disseminating free, real-time ARD and value-added information/services at variety of temporal and spatial scales. Synergy must be established among IoT, Big data analytics, machine learning tools, crowd sourcing, cloud computation, satellite-based navigation and Web-GIS through industry interface.

INDIAN SOCIETY OF GEOMATICS: AWARDS

National Geomatics Award for Excellence

This award has been instituted to recognize outstanding and conspicuously important contribution in promoting geomatics technology and applications at the country level. The contributions should have made major impact on the use of this technology for national development.

Areas of contribution considered for the award are:

1. Geographical Information System
2. Global Positioning System
3. Photogrammetry
4. Digital Cartography
5. Applications of Geomatics

The award shall consist of Rs. 50,000/- in cash, a medal and citation.

Eligibility

Any citizen of India, engaged in activities related to geomatics technology and its applications is eligible for this award. The prize is awarded on the basis of work primarily done in India.

The age limit for awardees is 45 years or above as on June 30 of the year of award.

Selection

A duly constituted Award Committee will evaluate all nominations received. The committee shall consist of eminent experts in the field of geo-spatial technology, to be identified by the Executive Council, ISG. The committee shall forward selected name/s to ISG – EC for approval and announcement. Apart from those persons, whose nominations have been received, the Committee may consider any person or persons who, in their opinion, have made outstanding contributions to development of geo-spatial technology and applications.

The award can be withheld in any year if, in the opinion of the committee, no candidate is found suitable in that particular year.

Presentation of the Award

The award shall be presented during the Annual Convention of ISG. Local Hospitality shall be taken care by ISG & Air fare (low cost) may be reimbursed if awardees request for it.

How to make Nomination

The nominations can be proposed by Head of a major research institute/ centre; Vice-Chancellor of a university; Secretary of Government Scientific Departments; President of a National Academy, President, Indian Society of Geomatics / Indian Society of Remote Sensing / Indian National Cartographic Association / ISG fellow or two life members of the society with more than 10 year old membership.

A candidate once nominated would be considered for a total period of two years. Nomination should be sent in the prescribed format to Secretary, ISG.

It is advised to refer ISG website www.isgindia.org for information related to last date for receiving nominations.

Format for nomination of Geomatics Award for Excellence

1. Name of the Nominee
2. Postal Address
3. Academic Background (Bachelor degree onwards)
4. Field of Specialisation
5. Important positions held (in chronological order)
6. Professional Experience including foreign assignments.
7. Important Awards / Honours
8. Important Publications/Patents: (A set of ten most important publications to be enclosed with this form)
9. Contributions of Nominee based on which the nomination is sent (in 1000 words, also provide a statement in 50 words which may be used for citation.):
10. Other Relevant Information:

Proposer:

Signature
Name
Address
Phone/ Fax
E-mail
Life Membership No. (in case of ISG Member):

Place & Date

Endorsed by (in case nomination is by 2 ISG Life members)

Signature
Name
Address
Phone/ Fax
E-mail
Life Membership No. (in case of ISG Member):

Place & Date

(The proposer should give a brief citation of the nominee's work)

National Geomatics Award

National Geomatics Award to be given each year: a) for original and significant contribution in Geomatics technology, b) for innovative applications in the field of Geomatics. Each award comprises a medal, a citation and a sum of Rs 25,000/- The guidelines for these awards are available on ISG website.

ISG Chapter Award for Best Performance

The best chapter award will be given to an active chapter of Indian Society of Geomatics, which has made significant contribution to further the mandate and goal of the society. The award consists of a citation and medal

President's Appreciation Medal for Contribution to the ISG

This award will be given to a member of the society, who has made noteworthy contribution to the growth of the ISG (its main body or any chapter). The Award consists of a Medal and a Citation.

Prof. Kakani Nageswara Rao Endowment Young Achiever Award

Indian Society of Geomatics instituted a new award from year 2013 named "Prof. Kakani Nageswara Rao Endowment Young Achiever Award", to encourage young researchers/scientists/academicians pursuing research in the field of geospatial technology/applications. The award carries a cash prize of Rs. 10,000/- along with a citation.

NATIONAL GEOMATICS AWARD

Indian Society of Geomatics has instituted two National Geomatics Awards to be given each year for (a) Original and significant contribution in Geomatics technology, (b) Innovative application(s) in the field of Geomatics. Each award comprises a medal, a citation and a sum of Rs. 25,000/-.

The guidelines for the award are as under

Areas of contribution considered for the award (both technology and applications)

1. Geographical Information System
2. Global Positioning System
3. Photogrammetry
4. Digital Cartography
5. Remote Sensing

Eligibility

Any citizen of India engaged in scientific work in any of the above-mentioned areas of research is eligible for the award.

The awards are to be given for the work largely carried out in India.

- First award will be given for original contribution in the field of Geomatics technology supported by publications in a refereed journal of repute.
- Second award will be given for carrying out innovative application(s). Supported by publications in peer reviewed Journals of repute.
- The contribution for the first award should have been accepted by peers through citation of the work.
- Work based on the applications of existing technologies will not be considered for the first award.
- The work should have made impact on the overall development of Geomatics.

How to Send Nomination

Nominations should be sent in the prescribed format, completed in all aspects to the Secretary, Indian Society of Geomatics, Space Applications Centre Campus, Ahmedabad 380 015. It is advised to refer ISG website www.isgindia.org for information related to last date for receiving nominations.

Selection Process

An expert committee, consisting of at least three members, constituted by the Executive Council of the Indian Society of Geomatics, will scrutinize the nominations and recommend the awardees' names to the Executive Council. The Council will decide on the award based on the recommendations.

FORMAT FOR AWARD NOMINATION

1. Name of the Candidate:
2. Present Position:
3. Positions held earlier (chronological order):
4. Academic qualifications (Bachelor's degree onwards):
5. Names of at least three Indian Scientists/Technologist in the area as possible referees *:
6. Brief write up on the work (500 words) for which award is claimed:
7. Publication(s) on the above work (reprint(s) to be enclosed):
8. List of other publications of the candidate:
9. Citation of the work for which award is claimed:
10. Impact of the work (for which award is claimed) on the development in the field of Geomatics (500 words):
11. Whether the work has already won any award? If so, give details:

The Applications in the above format (five copies) should be submitted (by Registered Post or Speed Post) to

The Secretary, Indian Society of Geomatics,
Space Applications Centre Campus,
Ahmedabad-380015

It is advised to refer ISG website www.isgindia.org for information related to last date for receiving nominations.

*ISG is, however, not bound to accept these names and can refer the nomination to other experts/peers

INDIAN SOCIETY OF GEOMATICS: FELLOWS

Shri Pramod P. Kale, Pune
 Dr George Joseph, Ahmedabad
 Dr A.K.S. Gopalan, Hyderabad
 Dr Prithvish Nag, Varanasi
 Dr Baldev Sahai, Ahmedabad
 Shri A.R. Dasgupta, Ahmedabad
 Dr R.R. Navalgund, Bengaluru
 Shri Rajesh Mathur, New Delhi
 Dr Ajai, Ahmedabad
 Prof P. Venkatachalam, Mumbai
 Dr Shailesh Nayak
 Prof I.V. Murli Krishna
 Prof SM Ramasamy, Tiruchirapalli
 Dr Ashok Kaushal, Pune
 Shri A.S. Kiran Kumar, Bengaluru
 Prof. P.K. Verma, Bhopal
 Maj. Gen. Siva Kumar, Hyderabad

INDIAN SOCIETY OF GEOMATICS: PATRON MEMBERS

- P-1 Director, Space Applications Centre (ISRO), Jodhpur Tekra Satellite Road, Ahmedabad - 380 015
 P-2 Settlement Commissioner, The Settlement Commissioner & Director of Land Records-Gujarat, Block No. 13, Floor 2, Old Sachivalay, Sector-10, Gandhinagar – 382 010
 P-3 Commissioner, Mumbai Metropolitan Region Development Authority, Bandra-Kurla Complex, Bandra East, Mumbai - 400 051
 P-4 Commissioner, land Records & Settlements Office, MP, Gwalior - 474 007
 P-5 Director General, Centre for Development of Advanced Computing (C-DAC), Pune University Campus, Ganesh Khind, Pune - 411 007
 P-6 Chairman, Indian Space Research Organization (ISRO), ISRO H.Q., Antariksha Bhavan, New BEL Road, Bengaluru 560 231
 P-7 Director General, Forest Survey of India, Kaulagarh Road, P.O. I.P.E., Dehra Dun – 248 195
 P-8 Commissioner, Vadodara Municipal Corporation, M.S. University, Vadodara - 390 002
 P-9 Director, Centre for Environmental Planning and Technology (CEPT), Navarangpura, Ahmedabad - 380 009
 P-10 Managing Director, ESRI INDIA, NIIT GIS Ltd., 8, Balaji Estate, Sudarshan Munjal Marg, Kalkaji, New Delhi - 110 019
 P-11 Director, Gujarat Water Supply and Sewerage Board (GWSSB), Jalseva Bhavan, Sector – 10A, Gandhinagar - 382 010
 P-12 Director, National Atlas & Thematic Mapping Organization (NATMO), Salt Lake, Kolkata - 700 064
 P-13 Director of Operations, GIS Services, Genesys International Corporation Ltd., 73-A, SDF-III, SEEPZ, Andheri (E), Mumbai - 400 096
 P-14 Managing Director, Speck Systems Limited, B-49, Electronics Complex, Kushiaguda, Hyderabad - 500 062
 P-15 Director, Institute of Remote Sensing (IRS), Anna University, Sardar Patel Road, Chennai - 600 025
 P-16 Managing Director, Tri-Geo Image Systems Ltd., 813 Nagarjuna Hills, PunjaGutta, Hyderabad - 500 082
 P-17 Managing Director, Scanpoint Graphics Ltd., B/h Town Hall, Ashram Road, Ahmedabad - 380 006
 P-18 Secretary General, Institute for Sustainable Development Research Studies (ISDRS), 7, Manav Ashram Colony, Goplapura Mod, Tonk Road, Jaipur - 302 018
 P-19 Commandant, Defense institute for GeoSpatial Information & Training (DIGIT), Nr. Army HQs Camp, Rao Tula Ram Marg, Cantt., New Delhi - 110 010
 P-20 Vice President, New Rolta India Ltd., Rolta Bhavan, 22nd Street, MIDC-Marol, Andheri East, Mumbai - 400 093
 P-21 Director, National Remote Sensing Centre (NRSC), Deptt. of Space, Govt. of India, Balanagar, Hyderabad - 500 037
 P-22 Managing Director, ERDAS India Ltd., Plot No. 7, Type-I, IE Kukatpalli, Hyderabad - 500 072
 P-23 Senior Manager, Larsen & Toubro Limited, Library and Documentation Centre ECC Constr. Gp., P.B. No. 979, Mount Poonamallee Road, Manapakkam, Chennai - 600 089.
 P-24 Director, North Eastern Space Applications Centre (NE-SAC), Department of Space, Umiam, Meghalaya 793 103
 P-25 Programme Coordinator, GSDG, Centre for Development of Advanced Computing (C-DAC), Pune University Campus, Pune – 411 007
 P-26 Chief Executive, Jishnu Ocean Technologies, PL-6A, Bldg. No. 6/15, Sector – 1, Khanda Colony, New Panvel (W), Navi Mumbai – 410 206
 P-27 Director General, A.P. State Remote Sensing Applications Centre (APSRAC), 8th Floor, “B” Block, Swarnajayanthi Complex, Ameerpet, Hyderabad- 500 038
 P-28 Director, Advanced Data Processing Res. Institute (ADRIN), 203, Akbar Road, Tarbund, Manovikas Nagar P.O., Secunderabad –500 009
 P-29 Managing Director, LEICA Geosystems Geospatial Imaging Pvt. (I) Ltd., 3, Enkay Square, 448a Udyog Vihar, Phase-5, Gurgaon- 122 016
 P-30 Director, Defense Terrain Research Limited (DTRL), Ministry of Defense, Govt. of India, Defense Research & Development Organisation, Metacafe House, New Delhi – 110 054
 P-31 Chairman, OGC India Forum, E/701, Gokul Residency, Thakur Village, Kandivali (E), Mumbai – 400 101
 P-32 Managing Director, ML Infomap Pvt. Ltd., 124-A, Katwaria Sarai, New Delhi – 110 016
 P-33 Director, Rolta India Limited, Rolta Tower, “A”, Rolta Technology Park, MIDC, Andheri (E), Mumbai – 400 093
 P-34 Director, State Remote Sensing Applications Centre, Aizawl – 796 012, Mizoram

Instructions for Authors

The journal covers all aspects of Geomatics – geodata acquisition, pre-processing, processing, analysis and publishing. Broadly this implies inclusion of areas like GIS, GPS, Photogrammetry, Cartography, Remote Sensing, Surveying, Spatial Data Infrastructure and Technology including hardware, software, algorithm, model and applications. It endeavors to provide an international forum for rapid publication of developments in the field – both in technology and applications.

A manuscript for publication must be based on original research work done by the author(s). It should not have been published in part or full in any type of publication nor should it be under consideration for publication in any periodical. Unsolicited review papers will not be published.

The Editorial Board or the Indian Society of Geomatics is not responsible for the opinions expressed by the authors.

Language

The language of the Journal will be English (Indian). However, manuscripts in English (US) and English (British) are also acceptable from authors from countries located outside India.

Manuscript Format

Each paper should have a title, name(s) of author(s), and affiliation of each of the authors with complete mailing address, e-mail address, an abstract, four to six keywords, and the text. The text should include introduction/background, research method, results, discussion, followed by acknowledgements and references. The main text should be divided in sections. Section headings should be concise and numbered in sequence, using a decimal system for subsections. Figures, images and their captions should be inserted at appropriate points of the text. Figures, images and tables should fit in two column format of the journal. If absolutely necessary, figures, images and tables can spread across the two columns. Figures and images, however, should not exceed half a page in height. A title should be provided for each Table, Image and Figure. All figures and images should be in 600 dpi resolution and sized as per column/margin width. Authors must ensure that diagrams/figures should not lose easy readability upon reduction to column size. The SI (metric) units and international quantities should be used throughout the paper. In case measurements are given in any other system, equivalent measurements in SI (metric) units should be indicated in brackets.

Use MS Word with English (UK/US) or English (Indian) dictionary. The page size should be A4 paper, with 2 cm margin on all sides. Title, authors and affiliation should be centred. Abstract should be justified across margins. The manuscript text should be in two columns of 8.2 cm each with a gutter of 6mm between them. Use only Times New Roman fonts. Title should be 12 points bold. Authors and affiliation should be 9 points. All other text including headings should be 10 points. Heading numbering scheme should be decimal e.g. 1, 1.1, 1.2.3, etc. Headings should be in bold.

Normally length of a published paper should be about 6-10 pages in A4 size including figures. Use of illustrations in colour should be restricted and resorted to only where it is absolutely necessary and not for enhancing the look of the paper. If the number of colour illustrations exceeds five, authors' institution may be asked to reimburse the extra cost involved, which at current rates is about Rs. 2500 per coloured figure/diagram/plate/illustration.

Submission of Manuscript

Submissions should be in electronic form via email. The manuscript may be sent by email to drajai1953@gmail.com. In exceptional cases hard copy submission in camera ready form may be allowed with the prior permission of the Chief Editor. Submission in any other form will be returned to the author. To speed up the review process, authors are advised to provide a list of three probable reviewers with their institutional address and e-mail IDs.

Guidelines for Citing References

Names of all cited publications should be given in full. No abbreviations should be used. Following procedure is to be adopted.

Journal Publications

Bahuguna, I.M. and A.V. Kulkarni (2005). Application of digital elevation model and orthoimages derived from IRS-1C Pan stereo data in monitoring variations in glacial dimensions, *Journal of the Indian Society of Remote Sensing*, 33(1), 107- 112. (to be referred to in the text as Bahuguna and Kulkarni (2005) or if more than two sets of authors are to be referred to, as (Bahuguna and Kulkarni, 2005; Jain et al., 1994)) When more than two authors are to be referred to, use Jain et al. (1994). However, in References, all authors are to be mentioned.

Publication in a Book

Misra, V.N. (1984). *Climate, a factor in the rise and fall of the Indus Civilization – Evidence from Rajasthan and Beyond in Frontiers of the Indus Civilization* (B.B. Lal and S.P. Gupta: Chief Editors) Books and Books, New Delhi, pp. 461-489

Papers Published in Seminar/ Symposium Proceedings

Jain, A., A.R. Shirish, M. Das, K. Das, M.C. Porwal, and P.S. Roy (1994). Remote Sensing and Geographic Information System – An approach for the assessment of biotic interference in the forest ecosystem. *Proceedings. 15th Asian Conference on Remote Sensing*, Bangalore, November 17-23, 1994, pp. 65-72.

Books

Possehl, Gregory L. (1999). *Indus Age: The beginnings*. Oxford and IBH Publishing Corporation, New Delhi.

Journal of Geomatics**Reviewing**

Each paper will be reviewed by three peers. Papers forwarded by members of the Editorial or Advisory Boards along with their comments would get processed faster and may be reviewed by two referees only.

Sample format for Authors is available in downloadable form at ISG website: www.isgindia.org/JOG/Sample_format.doc

Copyright

The copyright of the paper selected for publication will rest with the Indian Society of Geomatics. Corresponding author shall be required to sign a copyright assignment form, on behalf of all authors, once the paper is selected for publication. Authors are, however, at liberty to use this material elsewhere after obtaining permission from the Indian Society of Geomatics.

If the authors have used any copyright material in their

Vol 13, No. 1, April 2019

manuscript, it is understood that they have obtained permission from the owner of the copyright material and they should convey the same along with the manuscript to the Chief Editor.

Certificate of Original Work

The authors will also provide a certificate that the paper is an original work, not published or being considered for publication elsewhere.

In the event the certificate turns out to be false, the Journal shall ban the author(s) from publishing in the Journal for a period of five years and inform the same to all other related publications.

Reprints

Authors will be allowed to download the (PDF) of their paper from ISG Website www.isgindia.org, No hard copy reprints will be provided.

Journal of Geomatics		
Advertisement Rates		
	1 Issue	4 Issues
Back Cover Page in colour	Rs. 25,000	Rs. 80,000
Inside Cover Page in colour	Rs. 20,000	Rs. 64,000
Full Page inside in colour	Rs. 15,000	Rs. 48,000
Full Page inside in B/W	Rs. 10,000	Rs. 32,000

Advertisement Details

Mechanical Details
Double Spread/Center Spread (42 x 29.7) cm
Full page bleed (21 x 29.7) cm
Full page non-bleed (19 x 27.7) cm

Art Requirements

Negatives: Art must be right reading, emulsion, down. Film must be supplied in one piece per color, each identified by color. Camera-ready art is accepted for black & White adds; however, film is preferred. Electronic Files are also accepted.

Electronic File Requirements: All material must be received before ad close dates.

Software: Adobe illustrator 9.0 (saved as EPS). Adobe Photoshop CS (saved as EPS or TIFF). Please convert higher versions down. If you can only supply an IBM format, the file must be in viewable EPS or TIFF format with fonts embedded as that format.

Colour Ads: Colour separations must be provided, right reading, emulsion down. Please note that files using RGB or Pantone colours (PMS) must be converted to CMYK before we receive files.



ISG

To,
The Secretary, Indian Society of Geomatics
6202, Space Applications Centre (ISRO)
AHMEDABAD – 380 015. INDIA

Sir,

I want to become a Member/ Life Member/ Sustaining Member/ Patron Member/ Foreign Member/ Student Member of the Indian Society of Geomatics, Ahmedabad for the year _____. Membership fee of Rs. _____/- is being sent to you by Cash/ DD/ Cheque. (In case of DD/ Cheque No. _____ dated _____ drawn on Bank

_____. I agree to abide by the Constitution of the Society.

Date:

Place:

Signature

• Name: Mr/Ms/Mrs/Dr _____

• Address: _____

_____ PIN: _____

Phone: _____ Fax: _____ Email: _____

_____ • Date of Birth _____

• Qualifications _____

• Specialisation: _____

• Designation: _____ Organisation. _____

• Membership in other Societies: _____

• Mailing Address: _____

_____ PIN: _____

Proposed by:

(Member's Name and No)

Signature of Proposer

For Office Use: A/P/L Member No.		Receipt No.		Date:	
----------------------------------	--	-------------	--	-------	--

Indian Society of Geomatics (ISG), Room No. 6202 Space Applications Centre (ISRO), Ahmedabad-380015, Gujarat. Url: www.isgindia.org Phone: +91-79 26916202 Email: secretary@isgindia.org or sasharma@sac.isro.gov.in Fax +91-79-26916287

MEMBERSHIP FEES

Sr. No.	Membership	Life/Patron Membership fees		Annual Subscription
	Category	₹ Indian	US \$ Foreign	₹ Indian
1.	Annual Member	10	---	300
2.	Life Member			
	a) Admitted below 45 years of age	2500	250	
	b) Admitted after 45 years of age	2000	200	
3.	Sustaining Member	---	---	2000
4.	Patron Member	50000	3000	---
5.	Student Member	10	---	100

MEMBERSHIP GUIDELINES

- Subscription for Life Membership is also accepted in two equal instalments payable within duration of three months, if so desired by the applicant. In such a case, please specify that payment will be in instalments and also the probable date for the second instalment (within three months of the first instalment).
- A Member of the Society should countersign application of membership as proposer.
- Subscription in DD or Cheque should be made out in the name of '**Indian Society of Geomatics**' and payable at Ahmedabad.
- Direct deposit in ISG A/Cs must include bank fee RS. 25/- for cash payment.
- Financial year of the Society is from April 1 to March 31.
- For further details, contact Secretary, Indian Society of Geomatics at the address given above.
- ISG has chapters already established at the following places. Ahmedabad, Ajmer, Bhagalpur, Bhopal, Chennai, Dehradun, Delhi, Hissar, Hyderabad, Jaipur, Ludhiana, Mangalore, Mumbai, Mysore, Pune, Shillong, Trichi, Srinagar, Vadodara, Vallabh Vidya Nagar, Visakhapatnam and Trivandrum. Applicants for membership have the option to contact Secretary/Chairman of the local chapter for enrolment. Details can be found at the website of the Society: www.isgindia.org.
- Journal of the Society will be sent to Life Members by softcopy only.

Indian Society of Geomatics (ISG), Room No. 6202 Space Applications Centre (ISRO), Ahmedabad-380015, Gujarat. Url: www.isgindia.org Phone: +91-79 26916202
 Email: secretary@isgindia.org or sasharma@sac.isro.gov.in Fax +91-79-26916287

Geomatics Revealed

IGiS

Integrated GIS & IP Software

VERSION 2.0



MAKE IN INDIA



National Awards on Technology
By The Former President of India,
Dr. A. P. J. Abdul Kalam



Launch of IGiS Version 2.0
By Padam Shri AS Kiran Kumar, Chairman, ISRO and
Shri Tapan Mishra, DIRECTOR, SAC, ISRO.

What's new in IGiS

IGiS Version 2.0 is full of enhancements which you'll appreciate every day. New advanced GIS/IP and SAR modules are vital now a days. New COM Based Architecture makes you even more productive. The more you do with IGiS Version 2.0, the more you'll wonder how you ever did without it.

Enhancements in IGiS Version 2.0

- Advanced GIS / Image Processing
- Microwave SAR Analysis
- Meteorological Analysis
- COM Based Scalable Architecture
- New Ribbon Bar GUI
- Python Customization
- OGC Standards



Product Development Partner



Government of India | Department of Space
Indian Space Research Organisation - (ISRO)



Scanpoint Geomatics Ltd.

www.scanpointgeomatics.com

Scanpoint Geomatics Ltd.

Corporate Office : 12, Abhishree Corporate Park, Iskcon - Ambli Road, Ahmedabad - 380 058. Gujrat (India)
[P] +91 2717 297096-98 [F] +91 2717 297039 [E] info@scanpointgeomatics.com [W] www.scanpointgeomatics.com

INDIAN SPACE RESEARCH ORGANISATION
GOVERNMENT OF INDIA

A Smart Destination For Geospatial Solutions

National Remote Sensing Centre
Hyderabad, India
www.nrsc.gov.in
www.bhuvan.nrsc.gov.in
data@nrsc.gov.in

nrsc



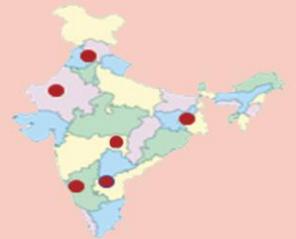
Only Organization in the Country
to Acquire & Supply
Satellite Data to Users



Aerial Acquisition for Specific
User Demands &
Disaster Management Support



Open Data & Value Added
Products Dissemination
Through Bhuvan



Region Specific Solutions



Capacity Building in
Remote Sensing Applications