

# Geostatistical computing for spatial landscape modelling using open source technology

Hariom Singh  
Indian Institute of Technology, Roorkee  
Roorkee, India  
e-mail: [hariom.cs1@gmail.com](mailto:hariom.cs1@gmail.com)

R.D. Garg  
Indian Institute of Technology, Roorkee  
Roorkee, India  
e-mail: [rdgarg@gmail.com](mailto:rdgarg@gmail.com)

A.K. Jha  
Indian Institute of Remote Sensing, ISRO  
Dehradun, India, e-mail: [akjha@iirs.gov.in](mailto:akjha@iirs.gov.in)

Harish Karnatak  
Indian Institute of Remote Sensing, ISRO, Dehradun  
Dehradun, India  
e-mail: [harish@iirs.gov.in](mailto:harish@iirs.gov.in)

Arijit Roy  
Indian Institute of Remote Sensing, ISRO, Dehradun  
Dehradun, India  
e-mail: [arijitroy@iirs.gov.in](mailto:arijitroy@iirs.gov.in)

**Abstract-** Spatial landscape modelling for identification of Biological Richness (BR) areas in Indian landscape was developed by Department of Space (DOS) and Department of Biotechnology (DBT) Govt. of India under national level project on Biodiversity Characterization at landscape level using RS&GIS. The spatial landscape model developed during the study is in desktop GIS environment using proprietary software solutions. The model has many limitations such as licensing issues, version compatibility, single user access, static pre-defined parameters etc. There is no possibility to include user defined GIS layer for further enhancement of the model. To overcome these issue an attempt has been made to develop modified version of spatial landscape model using R programming. The functionalities are further enhanced to describe the geostatistical computing of landscape thematic data such as vegetation type map, digital elevation model, social, physical etc. The computing was initiated by detail description of landscape analysis parameters such as fragmentation, patchiness, porosity, juxtaposition etc. The R project for geostatistical computing is the key open source environment for spatial data exploratory and advanced computational statistics. The geostatistical computing presented in this study is based on R open source software comprising raster, sp, rgdal etc packages. The different algorithms for landscape analysis parameter have been elaborated by using R programming and its graphical user interface Rstudio. The package is under development and there is feasibility to run through a Web Services.

**Keywords-** Geostatistics; landscape analysis parameter; R programming; Web Services;

## I. INTRODUCTION

National biodiversity characterization at landscape level using remote sensing and GIS, a project jointly sponsored by DBT and DOS, was implemented to rapid assessment and mapping of biological richness areas in India [1]. The contribution of the project is to the scientific understanding, landscape characterization and interpreting landscape analysis patterns of Indian natural vegetated ecosystems, their mapping through disturbance index and biological richness map. Geostatistical computing is commonly used to map environmental variable for biodiversity management and mitigation [2]. Spatial information about the predictable values can be easily generated and geostatistics takes a model of the spatial pattern recognition into account. However, geoinformatics also encouraged the use of geostatistics to support decision support problem such as landscape ecology management. By taking the consideration of geostatistical computing and landscape ecological principles for biodiversity characterization, spatial landscape analysis package (SPLAM) was developed at the Indian Institute of Remote Sensing (IIRS) using C++ programming environment [3]. The developed package comprises implemented algorithm for landscape analysis parameter (Fragmentation, Patchiness, Porosity, Interspersion Juxtaposition etc.) for the generation of Disturbance Index (DI) map and Biological Richness (BR) map. The developed model has certain limitations: desktop GIS environment using proprietary software solutions, licensing issues (vendor dependency), version

compatibility with propriety software such as ArcGIS, single user access, static pre-defined parameters etc.

To overcome these issue an attempt has been made to develop modified version of spatial landscape model (SPLAM) using R programming. The ecological parameters computing are under development and were initiated by detail description of landscape analysis parameters such as fragmentation, patchiness, porosity, juxtaposition, interspersion and proximity analysis. An attempt has been made to develop the new modified version of SPLAM, which is going to develop in the statistical environment [4]. Here, the required package such as raster, rgdal etc have been used for spatial landscape modelling. The developed model has the extensive possibilities for visualizing the result in graphical format and export the output in user defined format. The package is under development to run as computational back-end of a Web Services and integration with this package make it easy with other changing services. In this paper we demonstrate the geostatistical algorithm within an example from landscape ecology such as Fragmentation, Patchiness, Porosity, Interspersion, Juxtaposition, Ecosystem Uniqueness, Biodiversity value, Species Richness, Terrain Complexity. The cross-validation of the algorithms can be done easily by using SPLAM output and the suggestion of biodiversity expert. There are several choices to be made by the users of this developed model. In addition, the computation time will depend upon number of observation and size of the geospatial data. The intelligent efforts have been taken to optimize the computation time using R memory management parameter and other optimization technique.

## II. PROBLEM DOMAIN

An implementation of the specific model using open source technology for biodiversity conservation and prioritization without vendor independent component is hardly available. It has been observed that a system can automate the process of thematic map production, but there are possibilities that a computer could prepare thematic maps that are miss-interpreted by humans. Definitely, it required cross-validation and further research for accuracy assessment and performance optimization.

## III. LITERATURE REVIEW

Automap, a geostatistical module written in R programming language, was developed to describe automatically a spatial correlation model. It is a step ahead that is considered to be the most obstacles for automating the spatial interpolation method with

geostatistical algorithms [5]. For the development of eHabitat, computation of the Mahalanobis distances has been developed by using R programming and used for environmental modelling [6]. Geostatistical interpolation package: rtop, when observations have non-point spatial support has been developed by using R programming environment. The interpolation package developed with indicator spatial support by using an example from river network [7].

The application of geostatistical computing have been demonstrated in the biodiversity characterization at landscape level by using geospatial modelling technique have been demonstrated for North-East India as one of the most diverse region with respect to flora and fauna. Satellite remote sensing data have been used for the characterization of the landscape parameters. The modelling technique has been applied using SPLAM for the generation of DI map and BR map [8]. In addition, the fragmentation indexes map of India which was developed by using SPLAM and others tool and technologies. The forest areas have been categorized as high, low, moderate. It measured the tremendous pressure of population and anthropogenic on the forest area across the India. The studies have been represented in the graphical format (Histogram) so that it provide critical inputs to prioritization and conservation of India forest and associated biodiversity. The thematic map such as DI, BR, Vegetation type and Fragmentation map developed using SPLAM and associated technologies have been published in web browser environment (Biodiversity Information System) and it provide user friendly environment to download the data at client side [9].

## IV. RESULT AND DISCUSSION

Geostatistical computing for spatial landscape modelling using open source technology has been demonstrated by implementing landscape analysis parameter such as (1) Fragmentation (2) Patchiness (3) Porosity (4) Interspersion (5) Terrain Complexity. Here, all the landscape analysis parameter has been implemented by using R programming environment. For the algorithm development, Rstudio has been used for geostatistical computing and graphics. It is a free and open source integrated development environment (IDE). The above said landscape parameter is required for the generation of DI map BR assessment.

### *Landscape Modelling*

#### *A. Fragmentation*

The vegetation type index map was generated by using cloud free IRS 1C, 1D and P6 LISS III satellite data (spatial resolution: 23.5 m). The vegetation type map is

carried out for landscape analysis such as fragmentation [9]. The numbers of patches of non-forest and forest type per unit area were computed by using grid cell of  $n \times n$  (where  $n$  represent the kernel size, 500 m). The process was computed by moving window through the entire layer. The fragmentation was computed by using the mathematical equation:

$$\text{Frag} = \int (\text{Nf}, \text{Nnf})$$

Where Frag is the fragmentation, Nf is the forest patches, Nnf is the non-forest patches.

### Algorithm

Landscape analysis parameter (Fragmentation) has been computed by using R programming environment.

Step 1: Add required geostatistical package (raster, rgdal) by using syntax *library (raster)*

Step 2: Add vegetation type map of Uttarakhand

Step 3: Re-compute pixel value such as 0: Non-forest, 1: Forest.

Step 4: Run moving window (500 m) through the layer and calculate mean.

Step 5: Multiply by weight

Step 6: Use invert function

Step 7: Export in GeoTiff format

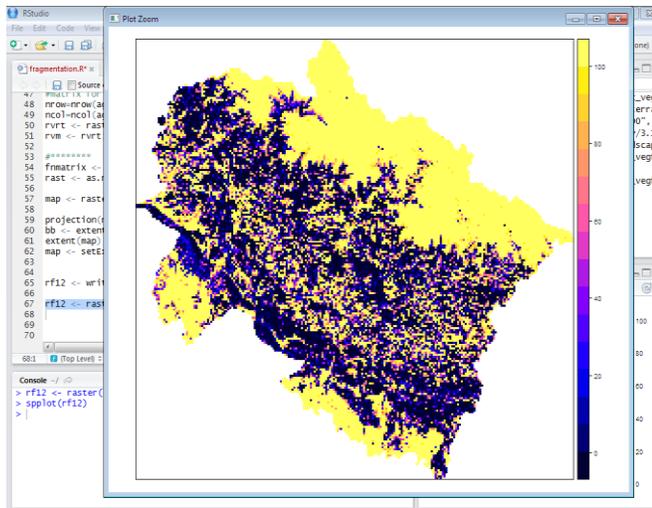


Figure 1: Fragmentation type map

“Fig 1”, the value 0 is representing non-forest and value 1-100 is representing increasing density of forest and further it can be categorized by using text legend. Here, the data source is vegetation type map of Uttarakhand. Rstudio, GUI is divided into many windows as shown in the above said figure. Every window has their utilities

such as workspace, output window etc. The graphical output is shown by using Pan Zoom utility of Rstudio.

### B. Patchiness

Patchiness is the computation of the density of all type of patches

### Algorithm

Step 1: Add required geostatistical package (raster, rgdal) by using syntax *library (raster)*

Step 2: Add vegetation type map Uttarakhand

Step 3: Re-compute pixel value such as 0: Non-forest, 1: Forest.

Step 4: Combine patches of connected cells. Each clump gets a unique ID

Step 5: Run moving window (500 m) through the layer and calculate count of unique patches.

Step 6: Export in GeoTiff format

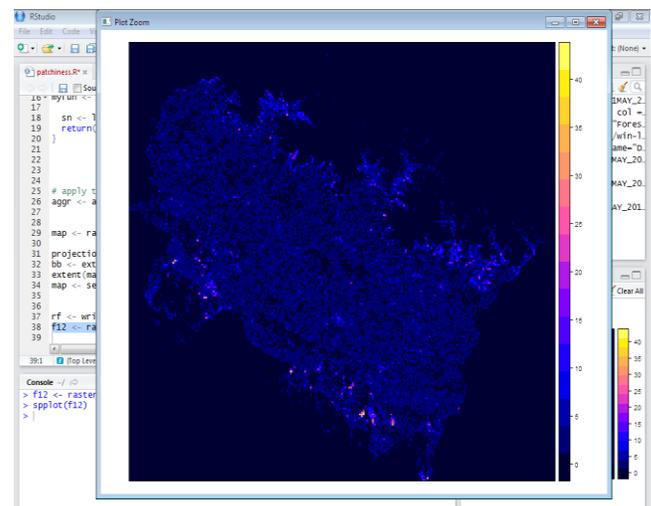


Figure 2: Patchiness map

“Fig 2”, the legend number (0-40) is representing the density of all type of patches. Here, 0 is representing minimum density and 40 are representing maximum density of patches in the map. The graphical output is shown by using Pan Zoom utility of Rstudio. R programming provides the environment for excellent graphical capability and export the output map in user defined format.

### C. Porosity

Porosity is the computation of the count of density of patches within particular type. Minimum porosity values indicate less interaction among landscape patches [8].

### Algorithm

- Step 1: Add required geostatistical package (raster, rgdal) by using syntax *library (raster)*
- Step 2: Add vegetation type map of Uttarakhand
- Step 3: Re-compute pixel value such as 1: Non-forest, 0: Forest.
- Step 4: Combine patches of connected cells. Each clump gets a unique ID
- Step 5: Run moving window (500 m) through the layer and calculate count of unique patches.
- Step 6: Export in GeoTiff format

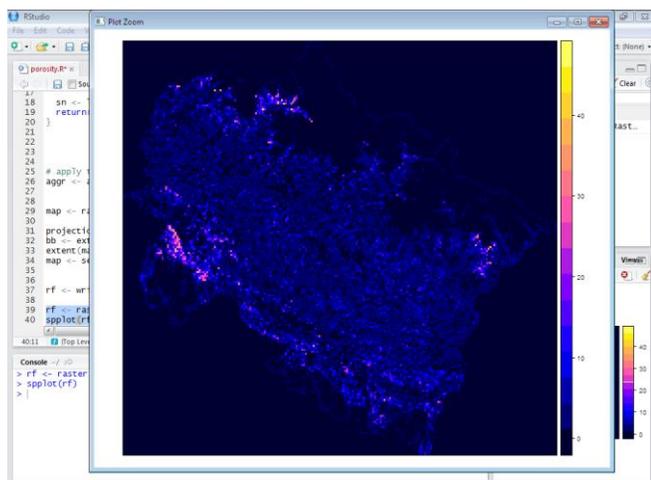


Figure 3: Porosity map

“Fig 3”, the legend number (0-40) is representing the computation of the count of density of patches within particular type. Here, 0 is representing minimum density and 40 are representing maximum density of patches in the map. The graphical output is shown by using Pan Zoom utility of Rstudio.

### D. Interspersion

Interspersion is the computation of spatial intermixing of landuse/habitats. It is computed as the count of surrounding grid cells that differ from the central cell value [8]. It represents homogeneity of the landscape and hence characterizes landscape diversity.

### Algorithm

- Step 1: Add required geostatistical package (raster, rgdal) by using syntax *library (raster)*
- Step 2: Add vegetation type map Uttarakhand
- Step 5: Run moving window (500 m) through the layer and computed as the count of surrounding grid cells that differ from the central cell value.
- Step 6: Export in GeoTiff format

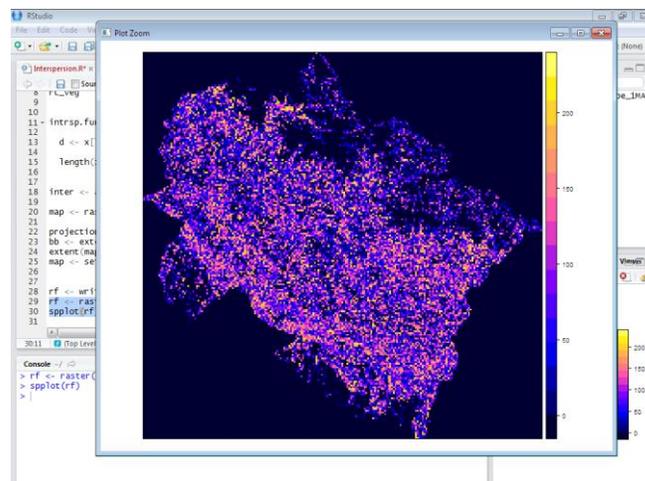


Figure 4: Interspersion map

“Fig 4”, the legend number (0-220) is representing the computation of count of surrounding grid cells that differ from the central cell value. Here, 0 is representing minimum interspersion and 200 are representing maximum interspersion in the map. The graphical output is shown by using Pan Zoom utility of Rstudio.

### E. Terrain Complexity

Terrain complexity has been calculated by using digital elevation model (DEM). It represents the variation of the elevation by driving the terrain complexity of the image. Here, the terrain complexity is represented by terrain steepness of the elevation data [10].

### Algorithm

- Step 1: Add required geostatistical package (raster, rgdal) by using syntax *library (raster)*
- Step 2: Add elevation data (GTOPO) India
- Step 5: Run moving window (500 m) through the layer and computed the log difference from the center value to the corresponding value within the mask.
- Step 6: Calculate the mean of the value within the mask.
- Step 6: Export in GeoTiff format

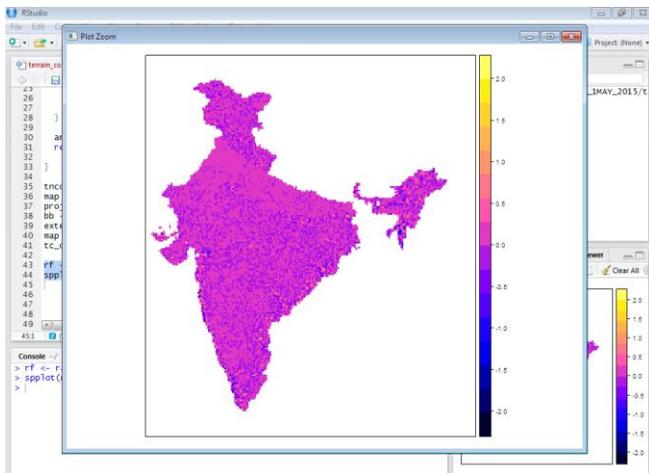


Figure 5: Terrain Complexity Map

"Fig 5", the legend number (0-2) is representing the computation of computed the log difference from the center value to the corresponding value within the mask. Here, 0 is representing less steepness and 2 are representing more steepness in the map. The graphical output is shown by using Pan Zoom utility of Rstudio.

## V. CONCLUSION

An initiative has been taken to overcome the limitation of previously developed SPLAM. The complete model for spatial landscape modelling is under development. The model is an open source and developed within R programming environment. In the present research, some landscape analysis parameter has been discussed and implemented such as Fragmentation, Patchiness, Porosity, Interspersion and Terrain complexity. The discussed landscape analysis parameter is required for the generation of DI map and BR assessment. It will be significant spatial landscape modelling tool for biodiversity conservation and prioritizations. The package is developed to run through a Web Services environment for further research in web browser environment.

**ACKNOWLEDGMENT:** The first author would like to extend his sincere thanks to Institute Computer Center, IIT Roorkee for their help in offering the resources in running the program.

## REFERENCES

- [1] P. S. Roy, H. Karnatak, S.P.S. Kushwaha A. Roy and S. Saran, "India plant diversity database at landscape level on geospatial platform: prospects and utility in today's changing climate", *Current Science*, 2012, vol. 102, pp. 1136-1142.
- [2] J. Chiles and P. Delfiner, "Geostatistics: Modeling Spatial Uncertainty". John Wiley & Sons, New York, 1999, pp. 102-135
- [3] P.S. Roy, S. Singh, C.B.S. Dutt, C. Jeganathan, R.N. Jadav, S.A. Ravan, et al. "Biodiversity characterisation at landscape level using satellite remote sensing and Geographic Information System. DOS-DBT User's Manual, Indian Institute of Remote Sensing (NRSA), Dept. of Space, Govt. of India, 4 Kalidas Road, Dehra Dun", 1999, pp. 14-24.
- [4] R core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2013, (<http://www.R-project.org/>).
- [5] D. Jesus, J. Dubois and P. Hiemstra, "Web-based geostatistics using WPS". In: *Proceedings of the 6th Geographic Information Days*, E. Pebesma, M. Bishr and T. Bartoschek. (Eds), 2008, pp. 199-218. IFGI Prints Band 32, 16-17 April. Muenster Germany.
- [6] G. Dubois, M. Schulz, J. Skoien, L. Bastin and Peedell, S, "eHabitat, a multi-purpose Web Processing Services for ecological modeling. *Environmental Modelling & Software*", 2013, vol. 41, pp.123-133.
- [7] J. O. Skoien, G. Bloschl, Laaha, E. Pebesma, J. Parajka and A. Viglione, "rtop: An R package for interpolation of data with a variable spatial support, with an example from river networks". *Computers & Geosciences*, 2014, vol. 67, pp.180-190.
- [8] P.S. Roy and S. Tomar, "Biodiversity characterization at landscape level using geospatial modelling technique". *Biological Conservation*, 2000, vol. 95, pp. 95-109.
- [9] P.S. Roy, M.S.R. Murthy, A. Roy, S.P.S Kushwaha, S. Singh, C.S. Jha, et al. "Forest fragmentation in India". *Current Science*, 2013 vol.105, pp. 774-780.
- [10] Q. Zhou, X. Liu and Y. Sun. "Terrain complexity and uncertainties in grid-based digital terrain analysis". *International Journal of Geographical Information Science*, 2006, vol.20, pp. 1137-1147

## AUTHORS PROFILE

Mr. Hariom Singh: Working as Ph.D. Research Scholar at Geomatics Engg Division, IIT Roorkee His area of specializations is geostatistical computing, web GIS, spatial DBMS and open source GIS development and implementation.

Dr. Harish Karnatak: Working as a Scientist SF at Indian Institute of Remote Sensing, ISRO Dehradun. He is PhD in computer science and having more than 12 years of experience in Geoinformatics and software development for GIS. His area of specializations is enterprise and distributed GIS, web GIS, SDSS, spatial DBMS and open source GIS development and implementation.

Dr. R.D. Garg: Working as Associate Professor at IIT Roorkee. He is Ph.D. in geomatics engineering and have more than 20 year of experience in the field of Geomatics Engineering, Land Surveying, Remote Sensing, GIS, GPS, Digital Image Processing, SAR Interferometry, GPR.

Dr. Arijit Roy: Working as a Scientist SF at Indian Institute of Remote Sensing, ISRO Dehradun. He is PhD in Ecology and having more than 15 years of experience in Ecological modeling, GIS. His area of specializations is Terrestrial Ecology, Ecological modeling, GIS.

Mr. A.K. Jha: Working as a Scientist SD at Indian Institute of Remote Sensing, ISRO Dehradun. He is M.Tech in Remote Sensing. His area of specializations is Image Processing.