PARALLEL COMPUTING BASED WEB GEOPROCESSING SERVICES FOR DERIVING FOREST FRAGMENTATION INDEX IN THE HIMALAYA

Hariom Singh[1], R.D. Garg[2], Harish C. Karnatak[3] and Arijit Roy[4]

[1]Geomatics Engineering Group, Civil Engg. Department, IIT Roorkee, Uttarakhand, India
Email: hariom.cs1@gmail.com

[2]Geomatics Engineering Group, Civil Engg. Department, IIT Roorkee, Uttarakhand, India
Email: garg_fce@iitr.ac.in

[3]Indian Institute of Remote Sensing, ISRO, Dehradun, Uttarakhand, India
Email: harish@iirs.gov.in

[4]Indian Institute of Remote Sensing, ISRO, Dehradun, Uttarakhand, India
Email: arijitroy@iirs.gov.in

KEY WORDS: web processing service, forest fragmentation, parallel processing, web coverage service, distributed GIS

ABSTRACT: Mapping and assessment of the extent of forest fragmentation index is one of the key requirements for undertaking any eco-restoration work. Commercial-Off-The-Shelf (COTS) based desktop softwares have certain limitations such as limited computation resources at user end, complexity in software deployment, platform depended solutions, software/algorithm updation issues etc. To overcome these limitations, a web processing service (WPS) based fragmentation index algorithm is developed and implemented during this study using free and open source geospatial solutions. The proposed algorithm is derived by using the multi-layer moving window based approach on geospatial data. The algorithm is based on parallel computing environment for better performance and efficiency when dealing with large geospatial data in multi-user environment. In order to demonstrate web geoprocessing based fragmentation algorithm, the Western Himalayan regions are taken as region of interest (ROI) for the online geoprocessing. As an input, land use land cover (LULC) map of the study area for different time periods is published as web coverage service (WCS) by using geospatial data server. Furthermore, the WCS of LULC maps are considered as input dataset for the derived algorithm. On the basis of ROI, the algorithm automatically computes the extent of fragmentation index. As the capability of the algorithm, the derived output is automatically categorized as intact, low, moderate, high and very high fragmented regions across the study area in distributed GIS environment. From the derived results, it is observed that large area is in very high category in spite of tremendous population pressure, indicating effective protection. At the end, parallel computing based web geoprocessing services approach and an interactive and responsive graphical user interface (GUI) provide the web environment for the computation of forest fragmentation index which can be utilized in the conservation of forest and associated biodiversity.

INTRODUCTION

Due to distributed and heterogeneous nature of geospatial data archives, the web based geoprocessing models are becoming one of the major requirements of the geospatial community. The available geoprocessing models are mainly Commercial-Off-The-Shelf (COTS) based desktop softwares. These softwares have certain limitations i.e. platform dependent solutions, complexity in software deployment, software/algorithm updation issues etc. and also require specific skills for installation and running of these type of geospatial software. At present, most of the geoportals are providing simple spatial query and geovisualization functionality to the users. During last decades, the researchers have initiated various open source and platform independent environment for online geoprocessing and analysis (Karnatak et al. 2012). However, the online geoprocessing is very limited due to various technological limitations at server and web client's end.

Open Geospatial Consortium (OGC) based Web Processing Service (WPS) standard provide a service interface for publishing and sharing geospatial processes to web clients (Hinz et al. 2013). The inputs of WPS can be web-accessible URLs or local files through other OGC standards such as Web Feature Service (WFS) and Web Coverage Service (WCS). The output of the geoprocessing services can be stored as web-accessible URLs or local files. In order to implement web based geoprocessing, the most popular and open source based OGC WPS specification implementations are PyWPS, 52° North WPS4R, ZOO WPS platform, Deegree WPS framework that enable deployment of geospatial processes on the web in a standardized manner. These software specifications are based on Service Oriented Architecture (SOA) implementation in which the published services can be accessed over a network and can be reused in an interoperable geoprocessing application (Giuliani et al. 2012; Yue et al. 2016). These frameworks have their pros and cons and depend on different programming languages. The 52° North WPS4R specification allows R programming script to deploy and integrate into a standardized web service environment.

In the last two decades there has been a large-scale degradation of contiguous forest land into smaller patches of forest. These isolated patches of forests are separated by agriculture, roads, utility corridors, or other human development activities. Due to immense population growth, the country like India which is tenth in the world in terms of extent of forest cover faces a serious problem of forest fragmentation. The geoprocessing based modelling and analysis provides the effect of forest fragmentation on landscape patterns and ecological process. There are various studies addressing the status of forest fragmentation and mapping with respect to vegetation types in the Indian landscape using customized software program (Roy et al. 2013; Chakraborty et al. 2017). However, the forest fragmentation algorithm is customized in COTS based desktop software. The algorithm requires optimization to process large and multi-temporal spatial datasets by using multi-layer moving window based approach in parallel computing environment with the use of state-of-the-art open source geospatial solutions.

The main purpose of this manuscript is to provide web based geoprocessing to deploy optimized forest fragmentation algorithms as an annotated script using distributed GIS based developed system. The implementation of the algorithm is executed by multi-layer moving window based approach for raster geocomputation under parallel computing environment. It is platform independent solution and based on open system architecture. The web based geoprocessing approach in this study takes essential inputs as annotated script using R programming language and OGC data service specifications such as WFS and WCS. Here, the algorithm of forest fragmentation has been modified and implemented on updated and suitable 52° North WPS4R computing environment. The WPS4R is the R back-end of the 52° North WPS. It is developed by using OGC WPS standard and provides the flexible environment for sharing of algorithms and geospatial data in web browser environment. The interoperable web process environment is used to manage deployed algorithms as services and make ready to available geoprocesses for WPS client. A case study of environmental application is developed and demonstrated to provide web based geoprocessing algorithms such as forest fragmentation using OGC WPS standard. The algorithm is available for geoprocessing services in an interoperable web application. As an input, land use land cover (LULC) map of India for different time periods (1985, 1995, and 2005) is published as WCS in geospatial data server repository of the application. The LULC datasets are used for the demonstrative case study; however, the geoprocessing environment can be used for any other geospatial datasets. The derived results of the parallel algorithms have correctly characterized forest fragmentation in the landscape of Western Himalayan state of India as Region of Interest (ROI). The derived output of the algorithm is categorized as intact, low, moderate, high and very high in equal interval classification in distributed GIS environment. The GUI of the application provides the basic functionality to the end user for the online computation of forest fragmentation index.

RELATED WORK

The desktop and COTS software based spatial landscape modelling acts as supporting tool for geospatial data processing (Roy and Behera, 2005; Karnatak et al. 2007). The desktop based model was developed by using software component provided by ESRI Inc. These proprietary software modules require commercial licences and upgradation. The main problem of this type of software is platform dependency. The installation and running of these systems also require lot of technical skills. The web based geoprocessing models using free and open source geospatial solution have been developed to address the key issue that COTS based desktop application faces in geospatial modelling (Geller and Turner , 2007; Dubois et al. 2013). Initially, the concept of "Model Web" has been proposed by Geller and Turner (2007), which was based on interoperable framework using web services. By using OGC based WPS protocol, the various web based geoprocessing and analysis applications have been developed for case studies like ecological modelling, climate change modelling and ecological forecasting (Clerici et al. 2013; Dubois et al. 2013). The focus of these studies is very specific in web based geospatial modelling and analysis.

Forest fragmentation is also computed by the desktop based spatial pattern analysis and landscape modelling software like FRAGSAT. However, it depends on COTS based platform and requires lot of technical skill and effort to derive output and its geovisualization. It processes the geospatial dataset in sequential rather than parallel environment and requires lot of time to process large datasets. With rapidly increasing volume of geospatial data in heterogeneous and distributed nature, it has become challenging to process and analyse data in an efficient manner using COTS desktop software APIs. With the use of this type of software APIs, various case studies on forest fragmentation have shown the characterization of fragmentation trends of the forest around the world to solve challenges in a real world context (Ren et al. 2017; Chakraborty et al. 2017).

APPROACH

Forest fragmentation algorithm determines the edge index on the landscape with the input parameter as vegetation type map/ LULC map. The algorithm of forest fragmentation was represented by the n×n convoluted moving window on the LULC/ vegetation map with the condition of identifying the number of forest patches within the moving window through the entire layer. It is considered as the number of forest patches identified in a landscape (Roy and Tomar, 2000; Roy et al. 2013). An output raster layer with patch number is determined and an associated normalized LUT of the layer is computed within the normalized range. The mathematical equation of the modified forest fragmentation algorithm is described as follows.

$$L_k(x, y) = \begin{cases} NA & if\ f(x,y) = 0 \text{ where } f(x,y) \text{ input map} \\ 1 & if\ f(x,y) = fcls \\ 0 & if\ f(x,y) \neq fcls \end{cases} \quad (1)$$

$$F_k(x,y) = \frac{1}{m\times n}\sum_{s=-i}^{i}\sum_{t=-j}^{j} \left(w_k(s,t)\ L(x+s, y+t)\right)\ \forall x,y \in L_k \quad (2)$$

where *f(x, y)* is the input LULC map; *fcls* is the user defined forest class values. The values of x and y are varied so that each pixel in *w(s, t)* considers every pixel in *L(x, y)*. Here, *w(s, t)* is the moving window and *L(x, y)* is the derived input map. The size of moving window $m\times n$ is described with $m=2i+1$ and $n=2j+1$, where i and j are equal and positive integers. The value of k depends on the number of input layers.

Furthermore, the computational complexity of the forest fragmentation algorithm is analysed based on its time and space complexity in parallel computing environment. In order to carry out parallel processing, the input raster dataset is sub-divided into number of smaller tiles and each tile is assigned to an independent computing processor (multi-core computing). The multi-processor results are then merged into single output with the objective of executing an algorithm in less time. Here, the approach describes the development of multi-layer based forest fragmentation algorithm, suitable for the computation of large spatial datasets using parallel processing. The algorithm is scalable to handle large spatial data cube in web based geoprocessing environment. In general, the speedup *(S)* of an algorithm is defined as the ratio of the serial runtime of the sequential algorithm (T_s) for executing a problem to the time taken by the parallel algorithm (T_p) to execute the same problem on p processors: $S= T_s/T_p$. The number of processors (p) is an important factor to analyse the speedup of a parallel algorithm. The systemic flowchart of the modified forest fragmentation algorithm is represented in Fig. 1. In the initial step, the algorithm considers different time periods LULC input map as WCS. The ROI is defined by providing polygon geometry for sub-area setting of the input datasets. The resampling operation is performed on the input datasets for similar resolution by setting rules for interpolating values across the new pixel sizes. In the next step, it is required to set the forest and non-forest class values. Thereafter, parallel computing based operations are executed by setting multi-layer moving window and then merge operation is executed to combine the multiple outputs into single output and then followed by various other mathematical operations. Thereafter, normalization and writing the derived output are carried out in the same dimension and coordinate system as input.

Web based computation is a widely used geoscientific method for distributed geoprocessing. OGC based WPS standard is the pivotal component for providing geoprocessing environment. In order to carry out web based geoprocessing and analysis, annotated script of R programming language syntax is deployed under the framework of 52º North WPS4R environments for invoking geospatial processing service of forest fragmentation algorithm. The developed script of the algorithm can be deployed by system administrator for an interoperable web application. The system is developed using open system architecture and is platform independent solution for web based geoprocessing. The implementation of this application includes both the script developer and web developer for providing a geoprocessing tool to the biodiversity users.

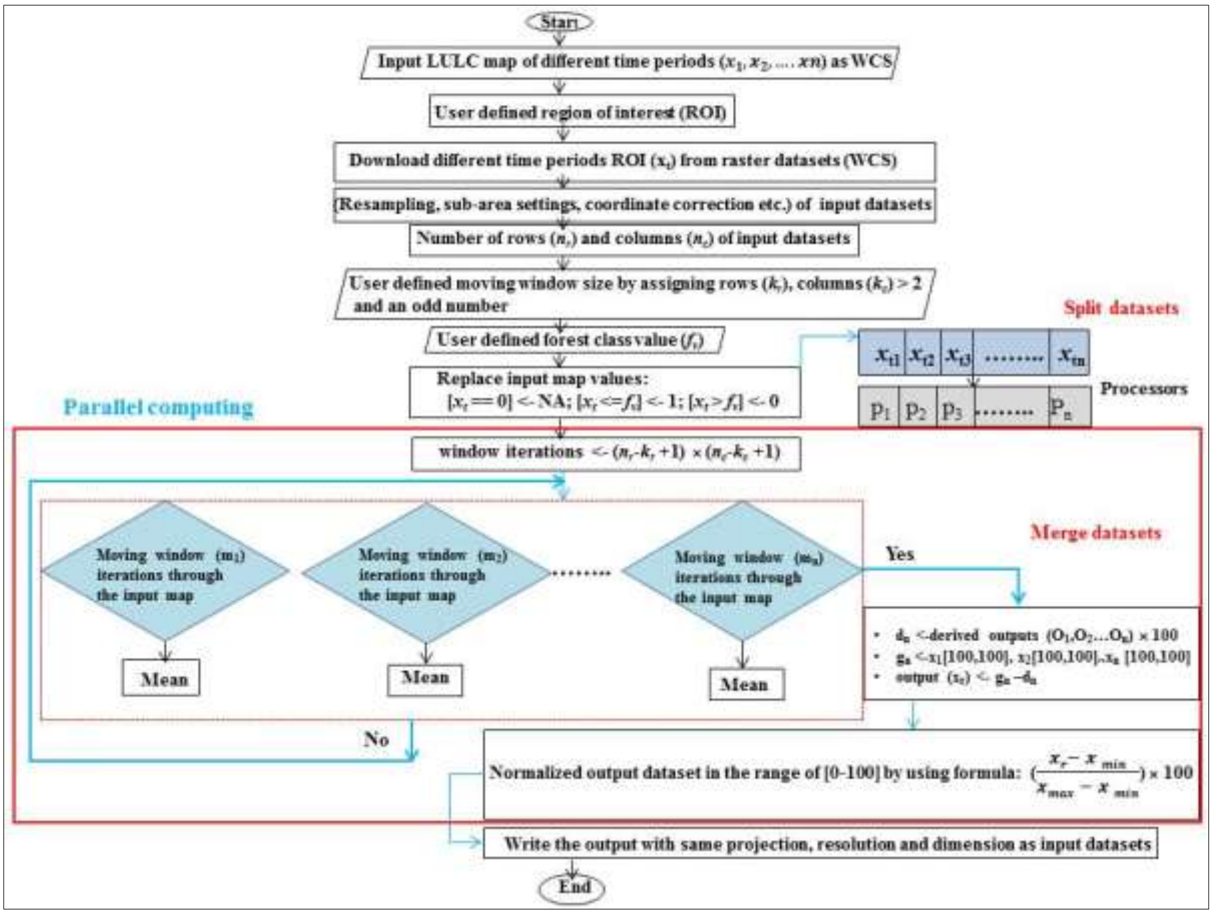

Figure 1: Flowchart for parallel computing and multi-layer moving window based algorithm of forest fragmentation.

The client/server based system architecture of the geoprocessing model is shown in Fig. 2. The architecture of the geoprocessing algorithm is divided into three main phases. The first phase is *"Inputs"* which include LULC datasets of India (1985, 1995 and 2005) as WCS. The raster datasets are stored as service repository in geospatial data server of the application. In the runtime, the defined ROI as polygon geometry is uploaded by the application users. The LULC maps as services and ROI are essential part of the annotated R script for geospatial data inputs. The system administrator has the responsibility to deploy R script in interoperable and standardized web processes under the algorithm repository of geoprocessing engine. The second phase of the developed algorithm is named as "*Geoprocessing Engine*" which is an important component of the system. The geoprocessing engine comprises algorithms repository, geospatial data accessing using web services, geoprocesses, and returns the derived output. The 52º North WPS4R provides a geoprocessing solution which is based on annotated R-script of the algorithm. The developed framework is flexible to upload new algorithms for any application area which require geospatial modelling. The key advantages of geoprocessing engine are interoperability of network-enabled geospatial data processing, extension of distributed computing approach, reusability, flexibility, scalability, and security. Here, the client/server based geoprocessing is feasible by using TCP/IP communication protocol, which is managed by Rserve package of the R. It provides the network based architecture for executing local R-script as an algorithm or mathematical equation on the web. The package acts as a socket server (TCP/IP) which allows binary requests and response to be sent to R. Every client/server based connection has a separate working directory and workspace. GUI implementations are available for popular languages such as PHP and Java, allowing any web application to use facilities of R software functionality. Thereafter, it manages and executes the geoprocess for selecting and downloading the defined extent of the dataset for algorithm computation.

In order to carry out web based geoprocessing and analysis, geospatial data server (GeoServer, MapServer etc.) is an essential component for publishing and geoprocessing data online. Here, GeoServer is used to publish LULC datasets as web services and MapServer is used to manage geovisualization and automatic classification.

The process of web based software development includes open source and freeware geospatial data server mainly used for interoperability and sharing geospatial data by using OGC web services standards (WMS, WFS, WCS etc.).

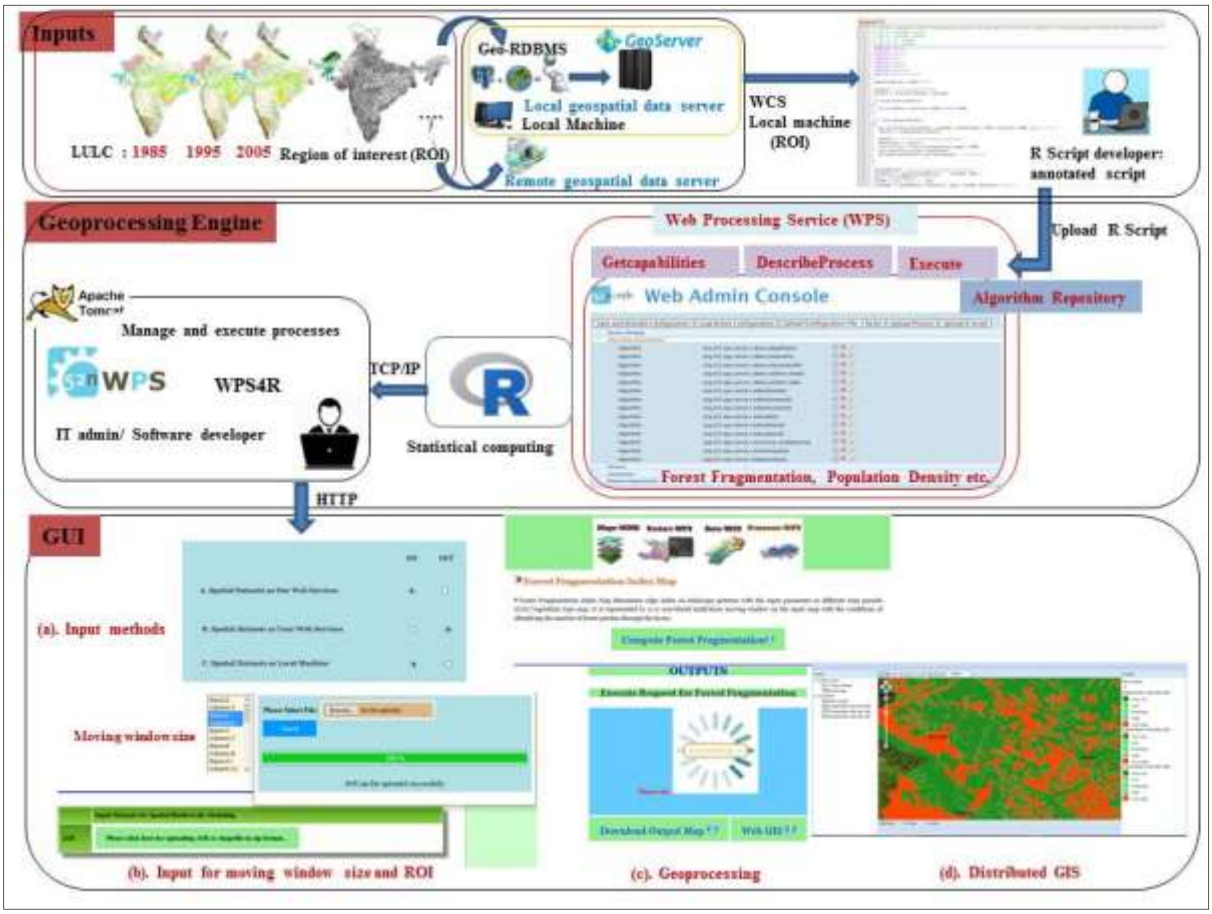

Figure 2: System architecture for web based geoprocessing services for forest fragmentation

The final phase of the algorithm is named as "*GUI*" which allows the end users to execute the algorithms in distributed and interoperable environment. In order to execute the forest fragmentation as web based geoprocessing algorithm, the end user needs to select input method, moving window size and AOI based on Indian landscape, geoprocessing, downloading derived output and geovisualization in distributed GIS environment, which is described in GUI section of Fig. 2(a), (b), (c) and (d). The distributed GIS environment is implemented by raster data classification (e.g. equal interval) for automatic geovisualization and mapping. In order to achieve this functionality, state-of-the-art open source GIS solutions such as HTML5, PHP, OpenLayers, ExtJS, GeoExt etc. are used for simple GUI.

EXPERIMENS AND RESULTS

This section defines the performance and scalability of the forest fragmentation algorithm in web based geoprocessing environment by using WPS. The online computation is carried out for forest fragmentation of Western Himalayan state in India, although the algorithm can be run for any part of the world with the availability of datasets. The geoprocessing framework is generic and flexible to execute the algorithms and can be deployed for several geospatial modelling applications. In order to compute the algorithm, the experiments and results have been demonstrated in the following sections. In the initial scenario, the LULC maps of Indian landscape are published as WCS in geospatial data server while the user can provide the dataset from remote data server using WCS. In order to process the datasets of certain extent/area, the user needs to upload ROI as vector data (geometry polygon). The sub-area setting, resampling operation etc. are performed to process the spatial data cube of Western Himalayan states as user defined ROI. In the second scenario, the geoprocessing engine of the system generates the outcome of forest fragmentation by analysing the effectiveness of the algorithm as geospatial tool to characterize and quantify land cover changes.

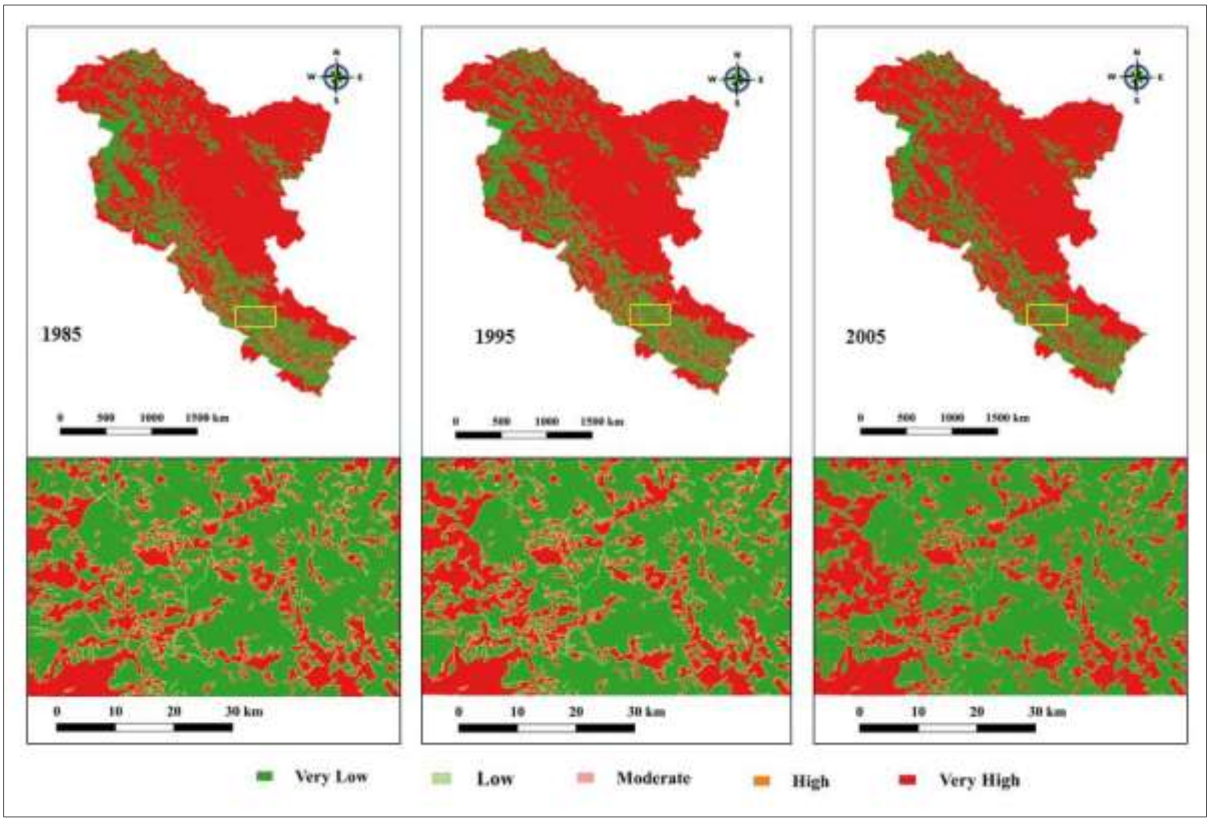

Figure 3: Forest fragmentation maps (1985, 1995, and 2005) are representing temporal fragmentation of Western Himalayan state of India.

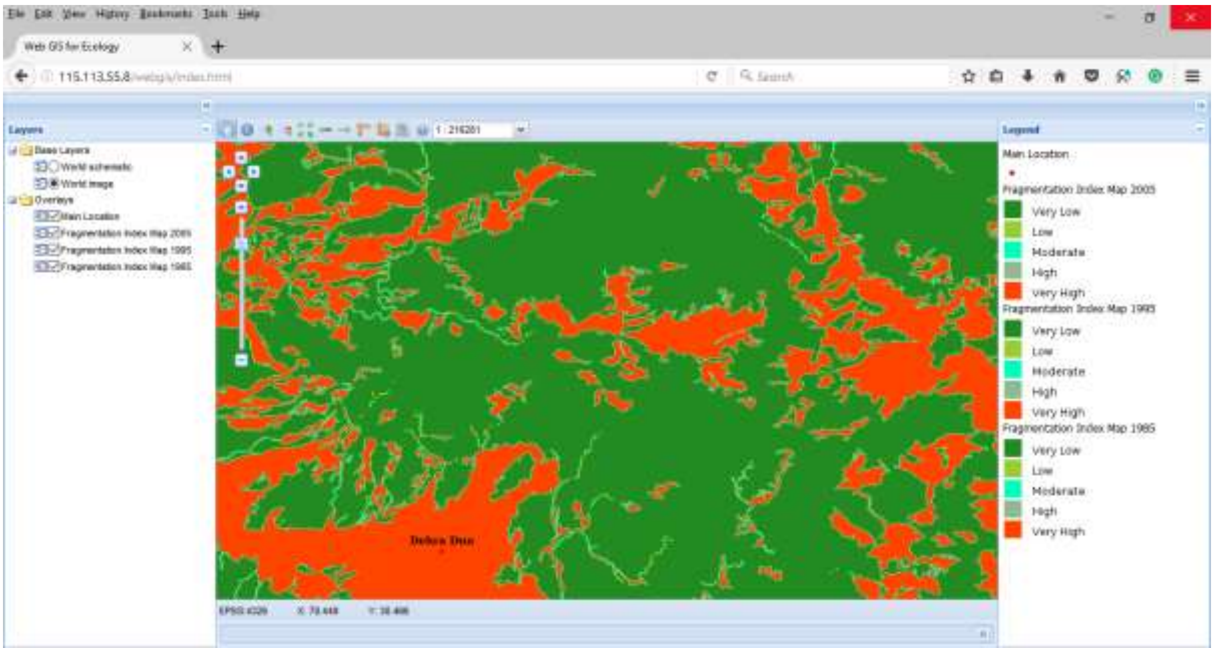

Figure 4: Forest fragmentation algorithm results in distributed GIS environment.

In order to reduce edge effects in the three time-period LULC maps (1985,1995, 2005), the algorithm has been computed by using multi-layer based moving window approach (e.g., 5x5), with user defined size, and running pixel by pixel computation through multiple layers. Here, the three landscape metrics have been processed with multi-moving-window based approach by assigning value at that pixel in the derived outputs. The outputs are the new maps with a derived pixel value defining the local landscape metrics that quantify the spatial characteristic of forest and non-forest patches. The derived algorithm is quantifying the pattern of forest and non-forest categories of patches. The computed results of forest fragmentation algorithm are shown in Fig. 3. The obtained results will help to identify the effects of land cover changes and fragmentation on ecosystem services and biodiversity in the region. In order to categorize the derived results, the output maps are classified using natural breaks approach as very low, low, moderate, high and very high. The very low category represents intact forest landscape while very high category represents highly fragmented region. In the third scenario, there is provision to download the derived outputs or perform mapping and geovisualization in distributed GIS environment. Here, the automated equal interval classifications as well as geovisualization operations are performed to dynamically visualize the forest fragmentation (1985, 1995, and 2005) result of the Western Himalaya region. The distributed GIS results of the forest fragmentation algorithm are shown in Fig. 4. Furthermore, the performance evaluation of this algorithm is analysed in sequential as well as parallel computing environment. The speedup, S= T_1/T_4, where T_1 (56054.76 seconds) is the time taken by one sequential processor and T_4 (34351.25 seconds) is the time taken by 4 parallel processors.

DISCUSSION

As the extensive degradation of forest land over the last decades, forest fragmentation is recognized as a serious threat to biodiversity across the globe. The process of forest fragmentation has been identified as the important factor to the loss and extinction of species diversity in Western Himalayan region of Indian landscape. The geospatial modelling is playing a very crucial role in generating information on LULC changes at a landscape level. The periodic assessment of forest fragmentation is a valuable attempt to quantify landscape pattern to understand the change characteristics of forest ecosystem. LULC dynamics is a fundamental parameter to quantify isolation of habitats and anthropogenic forest fragmentation. Change detection techniques based on multi-temporal satellite dataset have demonstrated their potential as an effective tool to understand forest fragmentation dynamics. There are several research studies in India as well as the globe, which have documented forest fragmentation using COTS platform in desktop based environment. The algorithms as well as the desktop based COTS software have certain limitations such as platform dependency, sequential computation, single moving window, closed architecture etc. In order to overcome these limitations, the forest fragmentation algorithm is modified by including the state-of-the-art functionality such as multi-moving-window approach, interoperability using GIS web service (WFS, WCS etc.), exploiting the power of parallel and distributed computing, fast and reliable access to near real-time computations, reusability, flexibility, scalability, security etc.

The modified forest fragmentation approach at landscape level is an optimized algorithm, which identifies the forest/non-forest patches from LULC/vegetation type map of large volume and distributed datasets. The need of hour is an online geoprocessing model for sharing and computing local and distributed datasets through GIS web services. In order to address the online geoprocessing algorithm, the WPS based forest fragmentation algorithm has been developed using state-of-the-art open source technologies such as 52º North WPS4R framework, OpenLayers, GeoServer etc. The developed geoprocessing framework is generic in nature and can be used for other algorithms. Furthermore, the pre-and post-processing GIS operation such as crop/mask, resampling, AOI settings, classification etc. are executed by the geoprocessing engine as it is the most time consuming process in desktop based geospatial modelling approach. The experimental run of the forest fragmentation algorithm has shown the online geoprocessing and analysis for the Western Himalayan region of Indian landscape, which acceptably characterize the forest/non-forest patches in the study area. The validation of the derived results is based on desktop based customized model i.e. SPLAM (Roy et al. 2012). The forest fragmentation map of this desktop based model can be accessed and downloaded by using biodiversity information system (URL-http://bis.iirs.gov.in/). Furthermore, the modified approach is reusable and scalable to process spatial data cube of LULC map simultaneously by using multi-moving-window approach rather than processing in fixed grid size. In order to process large volume spatial data cube of LULC/vegetation type map, the modified algorithm has included parallel processing to overcome the limitation of traditional geospatial modelling by using R as open source statistical computing software. Apart from the parallel processing, the algorithm is scale independent where it can be used for high resolution as well as coarse resolution spatial data cube.

CONCLUSION

The present study indicates the WPS based online forest fragmentation in parallel processing and multi-moving-window environment to handle large spatial data cube of LULC map. The geoprocessing engine for the algorithm is dynamic in nature to deploy more algorithms and handle distributed data archives through GIS web services. In order to overcome the limitation of COTS software as well as forest fragmentation algorithm, the modified algorithm includes the implementation of reusability, flexibility, interoperability, parallel computing, multi-layer moving window, GIS operation etc. The implementation of the algorithm is based on the open source geospatial technologies and open system architecture. The geoprocessing engine of the algorithm is based on SOA architecture to deploy annotated and optimized algorithm to any real-world application. In this study, forest fragmentation algorithm provides the valuable information on intact and very high fragmented region of Western Himalayas of Indian landscape during the time periods (1985, 1995 and 2005). In order to carry out web based geoprocessing, an interactive and response environment is provided for inputs, geoprocessing and outputs for the computation. In order to present experimental results, the Western Himalaya region of Indian landscape is defined as study area to implement the outcome of the modified forest fragmentation algorithm. The derived output of the algorithm can be downloaded or it can be geovisualized in distributed GIS environment. The outputs of the algorithm are validated by using already developed desktop based modelling approach. From the experimental run of the present study, it has been concluded that the optimized algorithm has shown significant improvement in the performance to handle large spatial data cube of LULC map and the derived forest fragmentation outcomes.

ACKNOWLEDGEMENT

The first author would like to acknowledge MHRD, Govt. of India for providing the financial assistantship during the research and also want to extend his sincere thanks to IIRS/ISRO, Dehradun for their help in offering the resources in running the program. Thanks to Dr. A. Senthil Kumar, Director IIRS/ISRO and Dr. S.K. Srivastav, GH RSGG IIRS/ISRO for their encouragement and technical support for this research study. Authors express their special thanks to Dr. P.S. Roy, Professor emeritus University of Hyderabad and former Director IIRS/ISRO for sharing his overall suggestions and guidance to develop this web application.

REFERENCES

Chakraborty, A., Ghosh, A., Sachdeva, K., Joshi, P.K., 2017. Characterizing fragmentation trends of the Himalayan forests in the Kumaon region of Uttarakhand, India. Ecological Informatics, 38, pp. 95-109.

Clerici, M., Combal, B., Pekel, J.F., Dubois, G., et al., 2013. The eStation, an Earth Observation processing service in support to ecological monitoring. Ecological Informatics, 18, pp. 162-170.

Dubois, G., Schulz, M., Skøien, J., Bastin, L., et al., 2013. eHabitat, a multi-purpose Web Processing Service for ecological modeling. Environmental Modelling & Software, 41, pp. 123-133.

Geller, G.N., Turner, W., 2007. The model Web: A concept for ecological forecasting. International Geoscience and Remote Sensing Symposium, IGARSS 2007, IEEE International, Barcelona, Spain, 23-28 July 2007.

Giuliani, G., Nativi, S., Lehmann, A., Ray, N., 2012. WPS mediation: An approach to process geospatial data on different computing backends. Computers & Geosciences, 47, pp. 20-33.

Hinz, M., Proß, B., Nüst, D., Pebesma, E., 2013. Spatial Statistics on the Geospatial Web. Agil-Leuven, pp. 14-17.

Karnatak, H.C., Saran, S., Bhatia, K., Roy, P.S., 2007. Multicriteria spatial decision analysis in web GIS environment. Geoinformatica, 11, pp. 407-429.

Karnatak, H.C., Shukla, R., Sharma, V.K., Murthy, Y.V.S., et al., 2012. Spatial mashup technology and real time data integration in geo-web application using open source GIS-a case study for disaster management. Geocarto Internationl, 27, pp. 499-514.

Ren, X., Lv, Y., Li, M., 2017. Evaluating differences in forest fragmentation and restoration between western natural forests and south eastern plantation forests in the United States. Journal of Environmental Management, 188, pp. 268-277.

Roy, P.S., Behera, M.D., 2005. Assessment of biological richness in different altitudinal zones in the Eastern Himalayas, Arunachal Pradesh, India. Current Science, 88, pp. 250-257.

Roy, P.S., Murthy, M.S.R., Kushwaha, S.P.S., Singh, S., et al., 2013. Forest fragmentation in India. Current Science, 105, pp. 774-780.

Roy, P.S., Roy, A., Karnatak, H., 2012. Contemporary tools for identification, assessment and monitoring biodiversity. Tropical Ecology, 53, pp. 261-272.

Roy, P.S., Tomar, S., 2000. Biodiversity characterization at landscape level using geospatial modelling technique. Biological Conservation, 95, pp. 95-109.

Yue, P., Guo, X., Zhang, M., Jiang, L., et al., 2016. Linked Data and SDI: The case on Web geoprocessing workflows. ISPRS Journal of Photogrammetry and Remote Sensing, 114, pp. 245-257.