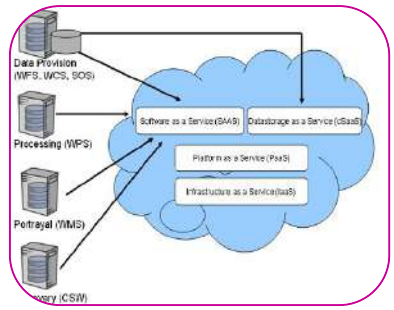
REVIEW OF RESEARCH
IMPACT FACTOR : 5.2331(UIF) UGC APPROVED JOURNAL NO. 48514 ISSN: 2249-894X
VOLUME - 7 | ISSUE - 10 | JULY - 2018

GEOCLOUD 4 HEALTH: CLOUD SDI MODEL FOR MAPPING AND SHARING OF MALARIA DISEASE INFORMATION INFRASTRUCTURE MANAGEMENT IN MAHARASHTRA, INDIA

Sapna. A. Sasane[1], A.S. Jadhav[2], Rabindra K. Barik[3] and V Raghavaswamy[4]
[1]Department of Geography, University of Pune, India.
[2]Department of Geography, Mumbai University, Maharashtra, India.
[3]School of Computer Applications, KIIT Deemed to be University, Bhubaneswar, India.
[4] Visiting Faculty , CSIT, JNT University, Hyderabad , Telangana, India.

ABSTRACT

Spatial Data Infrastructure (SDI) is an important framework for mapping and sharing of spatial big data across the web. Integration of SDI with cloud computing lead to emergence of Cloud SDI model as a tool for processing, transmission and analysis of spatial health data. Cloud computing is a paradigm where computing are employed on web to increase the throughput and reduce latency at the edge of the clients. In this study, a Cloud SDI model named GeoCloud4Health has been proposed and evaluated for mining analytics from spatial health data. It is built on a prototype using QGIS open source software. SDI can offer a real time and dynamic way to represent information through maps and related data. Here, a case study on malaria vector borne disease positive cases and that of number of deaths from spatial district level maps of Maharashtra state in India, from 2001 to 2014. The proposed model, besides visualization it executed overlay analysis and delivery of value added spatial information and outputs. The results of the present research work have shown the efficacy of using open source tools in the proposed framework which is modular, distributive and flexible for enhanced analysis of spatial health data in cloud computing environment.



KEY WORDS : *SDI, spatial health data, Malaria disease, Overlay analysis, QGIS, Cloud computing.*

1. INTRODUCTION

Open Sources GIS (OSGIS) can provide the users an alternative way for modeling the real world problems to produce robust spatial solutions in comparison to costly proprietary GIS software. In software market place, the source codes are shared for the needs of the organizations such that the users can access and solve various problems by using the codes. The proprietary software are costly for small organizations to operate their business model in cost effective manner (Barr, 2005). Camara (2004) emphasized that Open Source Software (OSS) projects can offer continuous improvement, rich functionality from contributing developers in these mapping industries (Barik et al., 2009).

Spatial data are essential for the development of a region and there is amplified credit for the benefits being returned to communities by investing in the development and implementation of spatial information systems. This has resulted in the development of the Spatial Data Infrastructure (SDI) concept at the local to global level. SDI is an inventiveness projected to create an environment in which all stakeholders can cooperate with each other and work together with technology to better achieve their objectives at different administrative and political levels. SDI has become noteworthy in formative the way in which spatial data are used in organization and nation, different regions and the world (GSDI Cookbook, 2008).

For health sector, disease data sharing are significant for the collaborative preparation and recovery stages of disease control. Disease phenomena are strongly associated with spatial and temporal factors. Web-based GIS provides a real-time and dynamic way to represent disease information on maps (Gao et al., 2008). However, data heterogeneity integration, interoperability, and cartographical representation are still major challenges in the health geographic fields (Vanmeulebrouk et al., 2008). These challenges cause barriers in extensively sharing and mapping of health data and restrain the efficacy in understanding and responding to disease outbreaks. To overcome these challenges in health sector, sharing and mapping of spatio-temporal disease information in an interoperable framework.

2. IMPORTANCE OF SDI MODEL

The spatial data communities participate in different levels of decision-making, from strategic to tactical to operational, in the management and utilization of spatial data. The development of SDI supports these decision-making functions at different administrative and political levels. The influence of SDI through decision-making frameworks varies in accordance with the jurisdiction, the differing levels of decision-making functions as well as the support of the spatial community. SDI provides an environment within which organizations interact with technologies to foster activities for producing and managing geographic data. It is the standard and technology to process and maintain the spatial data in different fields (Rawat, 2003).

The data and different categories of people are one category while the second category is the technological components like policy, access network and standards. The SDI model concept is depicted in Figure 1 (Rajabifard et al., 2002; Mansourian et al., 2005; Puri et al.,2007; Rawat 2003)

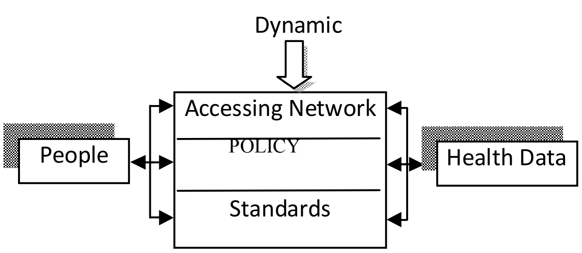


FIGURE 1: THE DYNAMIC NATURE OF SDI MODEL

2.1 Need of SDI model for Malaria Information Infrastructure Mapping

The increasing number of Malaria, Dengue, Chikungunia cases across urban centers of the country has once again re-generated active interest to study the incidence of urban Malaria among the researchers (IPCC 2007). The scope of present study is to throw light on the nature of incidence of Malaria over the last one and a half decade in Maharashtra. With the integration of spatial, web and mobile technology, it has the ability for various functionalities in terms of spatial data sharing and mapping over web. It can offer real time data representation through maps. So, there is an urgent need to establish a robust SDI model where each and every stakeholder can use/access spatial data (Barik et al. 2014). Spatial Web Service is one of the key technologies require for development and implementation of SDI (He et al. 2015). Design and implementation of SDI is used in Cloud Computing technology for sharing/accessing the information about Malaria Information Infrastructure mapping and sharing in Maharashtra, India. It enables the end user or data analyst to quickly look into the problem and gets the information according to their need. Thus, the next section describes the details of related works which has been done with the Cloud SDI model.

2.2 Cloud SDI model

Cloud computing provides enormous store and computable assets for implementation of geo-analysis prototypes. Here, the original model provides a transition from PC to computable servers. Cloud computing and other web processing architectures has delivered an open environment in web to share different assets (Buyya et al. 2008; Chen et al. 2012; Huang et al. 2013; Barik et al. 2014; Barik et al. 2011).

Likewise, Cloud SDI model deploys a unique-instance and permitting various types of client to contribute assets without disrupting each other (Vaccari et al. 2009; Xiaolin 2005; Ramachandra and Kumar 2004; Morris 2006; Leidig and Richard 2015; Yang et al. 2010; Wu et al. 2010; Yue et al. 2013). Figure 2 illustrates the system architecture for Cloud SDI model.

The client tier layer consists of thick clients, thin clients and mobile clients for capabilities of visualizing the spatial information (Evangelidis et al. 2014). The Application Tier comprises the main spatial services executed by servers. It enables intermediate among the different clients and providers. In top of the application tier, dedicated server for application has been operated for different services i.e. WMS, WCS, WFS, CSW and WPS (Bastian et al. 2010; Yang et al. 2011).

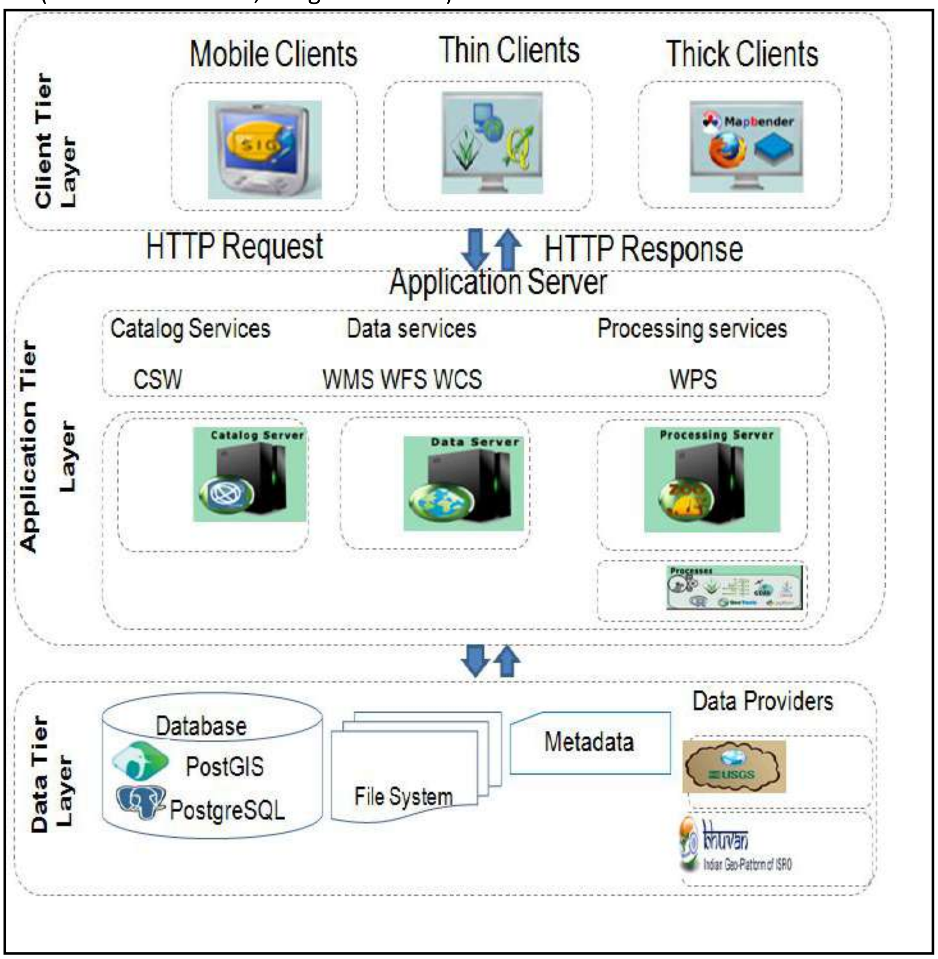


Figure 2: System architecture for Cloud SDI model (Evangelidis et al. 2014)

3. OBJECTIVE

In this paper, the spatial data managed at the verge expending cloud computing platform. It highlighted the following contributions to Cloud SDI model i.e. *GeoCloud4Health* in Malaria Information Infrastructure mapping and sharing of information in Maharashtra, India:

- It has proposed Cloud SDI model i.e. *GeoCloud4Health* which meant to mend output and decrease latent for analysis & storing of spatial data associated with the malaria information infrastructure mapping
- It has executed malaria information infrastructure (Case Study) as positive case and death trends due to malaria in the district of Maharashtra, India
- It has designed the spatial database creation of malaria data with the help of QGIS; and also generated spatial web services for mobile and desktop applications with QGIS Cloud platform
- It has also performed various data and overlay analysis for processing and visualisation of spatial data in *GeoCloud4Health* environment

4. STUDY AREA

Maharashtra is a state in western area of India. It is one of the second most populated state after Uttar Pradesh and third largest state by area in India. Maharashtra is surrounded by Gujarat & the Union territory of Dadra and Nagar Haveli to northwest, Arabian Sea to west, Karnataka to the south, Madhya Pradesh to the north and northeast, Chhattisgarh to the east, Goa to the southwest and Telangana to the southeast. The state covers 9.84% of the total geographical area of India and an area of 307,731 km2 (118,816 sq mi). There are nearly 378 urban centers and 41000 villages in Maharashtra. Maharashtra has one of the highest levels of urbanization of all Indian states (Internet-2 2017). Figure 3 shows the study area location map of Maharashtra, India.

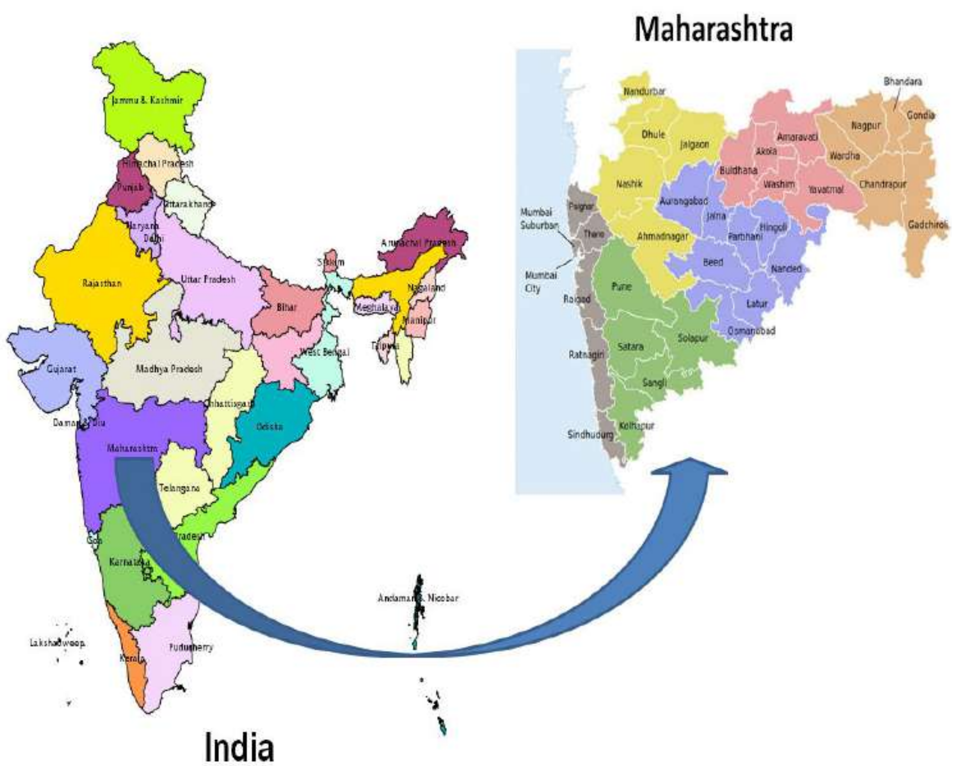


Figure 3: Study area location map of India with State Maharashtra

(Internet-1 2017: https://www.ncbi.nlm.nih.gov; Internet-2 2017: http://www.maharashtra.nic.in)

5. THE SYSTEM ARCHITECTURE OF *GEOCLOUD 4 HEALTH* MODEL

For development of *GeoCloud4Health* model, the main focus has been on the use of a practical approach to explore and extend the concept of SDI for health sector.

GeoCloud4Health model provides an effective and efficient means for sharing spatial health data and non-spatial data and for publication of maps on the web using GIS in a secure way. Figure 4 shows the flow of information for the *GeoCloud4Health* model framework developed in the present work.

The architecture of *GeoCloud4Health* is classified into three important layers. Initial phase of the layer is data tier layer. In this layer, spatial database is developed with the help of QGIS software. There are different data providers i.e. Open Street, Bing and Google map are used with the developed database for various analysis. In the layer of application, catalog services for searching of desired services have been maintained by catalog service. For invocation of WFS & WMS services are maintained thorough the QGIS Cloud Provider. This provider is the responsible for the management of application tier layer. The client tier is categories into three parts as thin, thick and mobile client. Figure 5 is described the system architecture of *GeoCloud4Health* model.

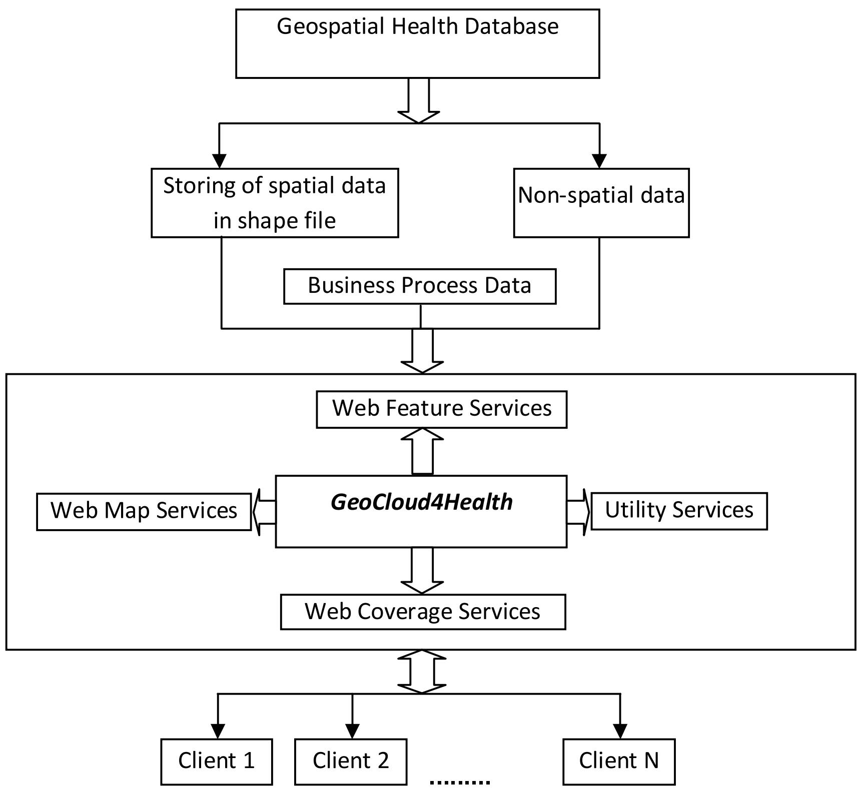


Figure 4:Flow Chart showing *GeoCloud4Health*Framework

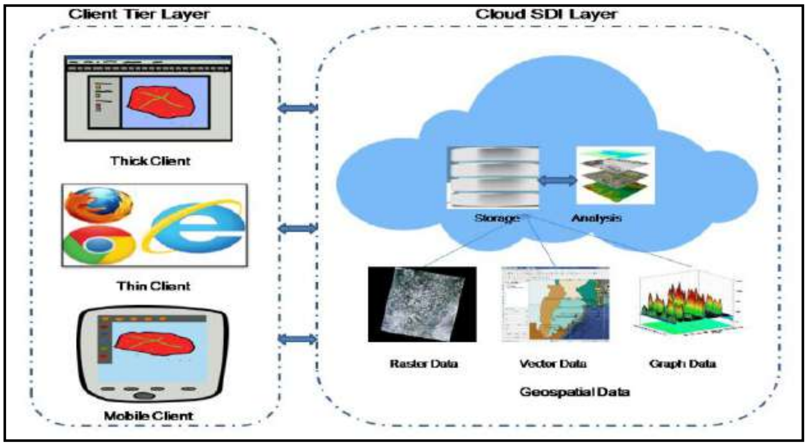


Figure 5: System architecture of *GeoCloud4Health*

For the creation of integrated spatial health database, the present model uses QGIS software and its QGIS Cloud plugin have been integrated for imparting the spatial web capabilities.

5. PROTOTYPE SDI DEVELOPMENT FOR *GEOCLOUD4HEALTH*

The prototype development is based on Jacobson's method of Object Oriented Software Engineering (OOSE) for incorporating strong user focus and the time critical nature (Mall, 2004; Davies, 2003). The software development process adopts a sequence of steps including requirement specification, analysis and design, implementation and testing, complete module and model observation. The process is usually cyclic or incremental in nature and each implementation refines the analysis and design stages through evaluation and testing of a completed module. This successive version development allows to take into consideration more informed view of SDI requirements. Figure 6 shows the complete process model for development of *GeoCloud4Health*.

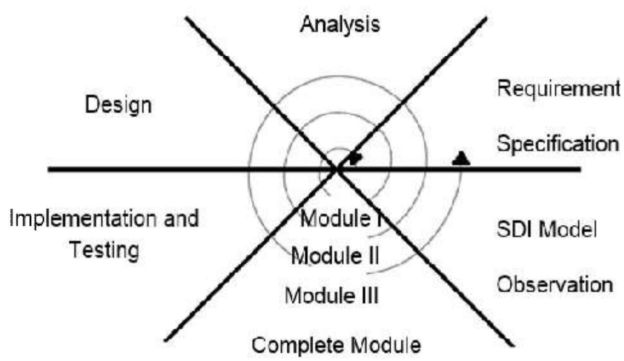


Figure 6: Win-win Spiral model for *GeoCloud4Health*

The use case model is the main aspects in this phase which basically specifies a sequence of actions need to perform (Davies, 2003). The construction of use case model involves how a user would interact with the system. The use case model is associated with administrative user, general user and developer. All these three types of users will have the different role on this system. From the use case model, it can visualize the more efficiency of each and every user's role. Figure 7 shows the use case model for *GeoCloud4Health*

The preliminary investigations determine the feasibility of system requested. There are three stages in the feasibility study portion of the preliminary investigation, namely, technical feasibility, operational feasibility and economic feasibility. Technical feasibility is concerned with specifying equipment and software that will successfully satisfy the user requirements and technical needs of the system. The operational feasibility deals with various operations on the software that should be easy and interactive. The economic feasibility deals with the tools required for developing and running the system, which should be easily available and cheap. In the present study, the selected tools are OSS which are free of cost and are downloaded from the web.

The design phase involves the creation of spatial health database. Spatial data have been created in the form of shape files. These shape files are used for retrieval, maintenance and deployment on the web. The thematic layers created include maps of India with State boundaries, Maharashtra state with block boundaries. It has also proposed the Cloud Spatial Data Infrastructure model for malaria information infrastructure mapping. In Cloud SDI model, spatial database has been developed for the health data of both the positive cases and that of number of deaths due to Malaria in Maharashtra, India from 2001 to 2014 Further, in the design phase, the flow of processes in the system is captured in the form of state diagram which need to be properly designed. In context of the present work, data flow diagram describes the workflow behavior of *GeoCloud4Health* and is shown in figure 8.

The flow of process in the system is captured in the form of state diagram In the present research, state diagram is described with the different state of *GeoCloud4Health* and is elaborated in figure 9. The model consists of three main important modules. Module I is used for registration of new users, Module II is for Maharashtra State Malaria data mapping and sharing where as module III is for utility services for uploading and

Figure 7: Use case model for *GeoCloud4Health*

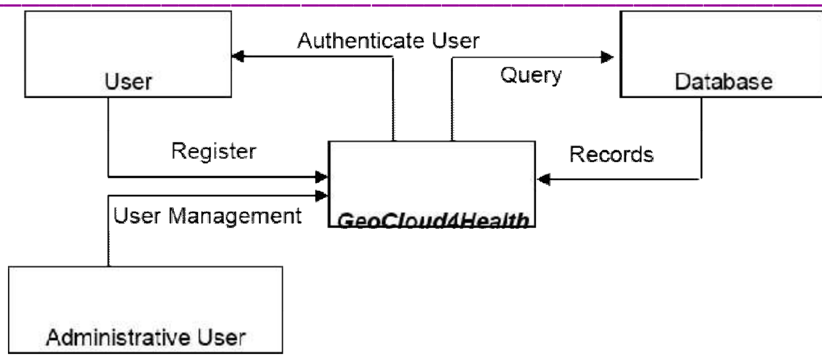


Figure 8: Data flow diagram (DFD) for *GeoCloud4Health*

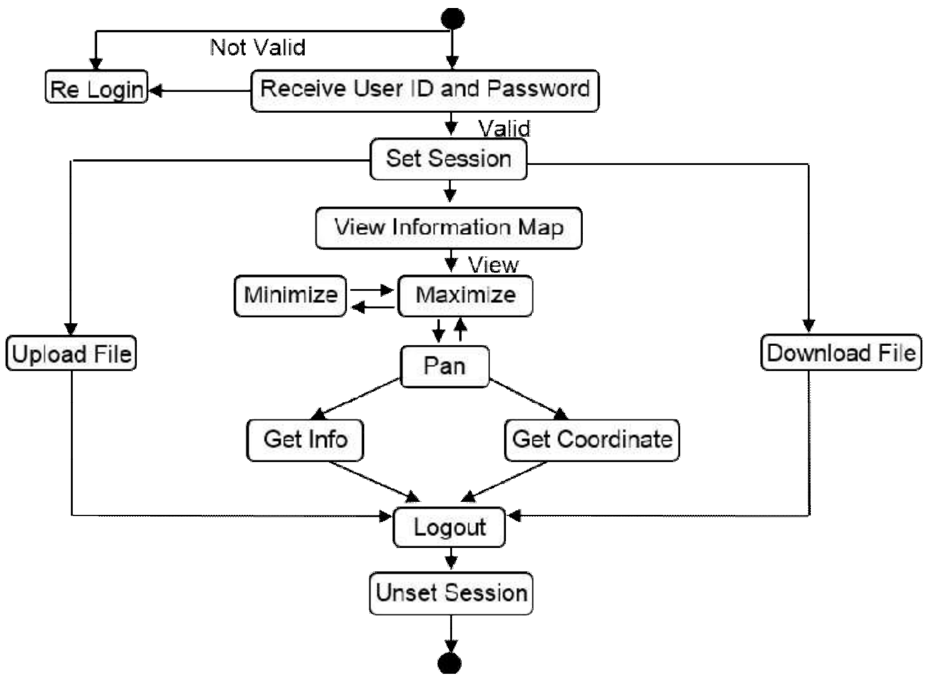


FIGURE 9: STATE DIAGRAM OF *GEOCLOUD4HEALTH*

6. *GEOCLOUD 4 HEALTH* MODEL OBSERVATION

Here data analysis particularly overlay analysis is done. It can superimpose various spatial data in a common platform for better analysis by spatial data analyst with raster and vector spatial data. After schema definition, 2 thematic layers have been created. First layer has been created which indicates the whole India state boundary. In this layer, WGS-84 with EPSG: 4326 coordinate reference system has been chosen. The next thematic layer has been created for malaria information infrastructure mapping in Maharashtra, India. It is observed that overlay analysis of various raster and vector data of particular area has been performed. Initially, the developed spatial databases have been opened with QGIS and performed some join operations. In QGIS, QGIS Cloud plugin is installed. This plugin has the capability of storing vector and raster data set in cloud database. After storing in cloud database, it automatically provides link both in thin and mobile clients for proper visualization and use of both raster and vector data. Figure 10 has shown the overlay analysis on thin client environment in *GeoCloud4Health* model (Internet-3 2017; Internet-42017).

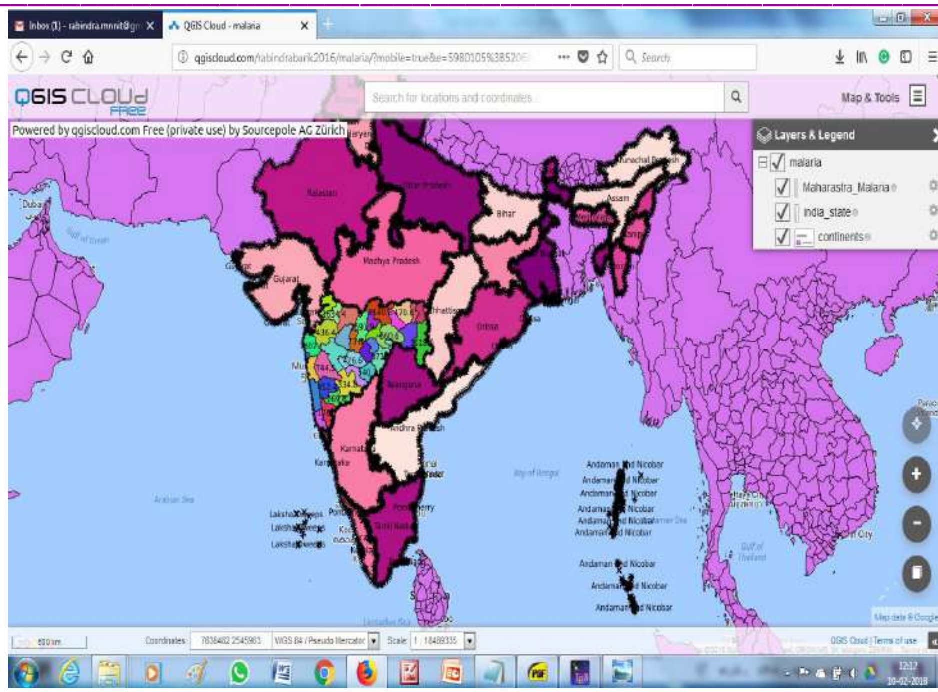


Figure 10: Integration of Malaria Information Infrastructure Mapping in Maharashtra, India on thin client environment in *GeoCloud4Health* model

8. SPATIAL DATA ANALYSIS

Death due to Malaria from 2001-2014 reported very high in Mumbai which is range from 98 – 980 persons because of high rainfall high humidity and uncontrollable growth of urbanization more slums and unhygienic conditions swamy areas all are favorable factor for the incidence as well as death due to malaria and more population on small island Mumbai which has shown in the Table 1 and Figure 10. After this Thane and Nagpur districts shows high incidence and death due to malaria which range from 55-98. After this Gadchiroli, Chandrapur, Bhandara, Gondia these north eastern districts shows the death in the range of 28-55 because of heavy rainfall and high humidity. But Pune comes under rain shadow zone still it shows more incidence and death due to malaria it range from 28-55 and the causative factor is different that is uncontrollable rapid increase of urbanization which comes along with more slums, more unhygienic conditions and dumping ground swamy areas etc. All these conditions are favorable for mosquito breeding ground zone and the diseases. After this Raigad, Ratnagiri, Wardha and Ahmednagar show the death in between the range of 16-28. Rest of all other districts shows the death in between 1-16 surprisingly Sindhudurg is coastal district still comes under this category and shows the incidence and death is very less. It is gratifying thing that Washim district reported not a single death due to malaria which is illustrated in Figure 10. Incidence of malaria from 2001- 2005 the data shows highest incidence was in Mumbai and Bhandara districts followed by Thane, Raigad and Wardha districts.

Incidence of positive cases of malaria from 2006- 2010 the data shows highest incidence was in Mumbai , then Thane, Raigad, Gadchiroli and Bhandara districts. In Figure 10 shows that Incidence of malaria from 2011- 2014 the data shows highest incidence was in Mumbai, Thane, Gadchiroli followed by Raigad, Chandrapurand Ahmednagar. In this period Bhandara district shows very less incidence may be because of some mitigative measures and awareness. Finally the trend map shows the highest incidence is in the Mumbai, Thane and Gadchiroli districts followed by Raigad, Chandrapur and Ahmednagar districts A negative trend shows in the Bhandara districts because from 2001 to 2010 continually it shows highest incidence of positive cases of malaria but in the period of 2011- 2014 it shows negligible incidence because of awareness and other mitigative factors which is really gratifying thing. Figure 11 has also shown the positive cases of malaria in maps from 2001-2014.

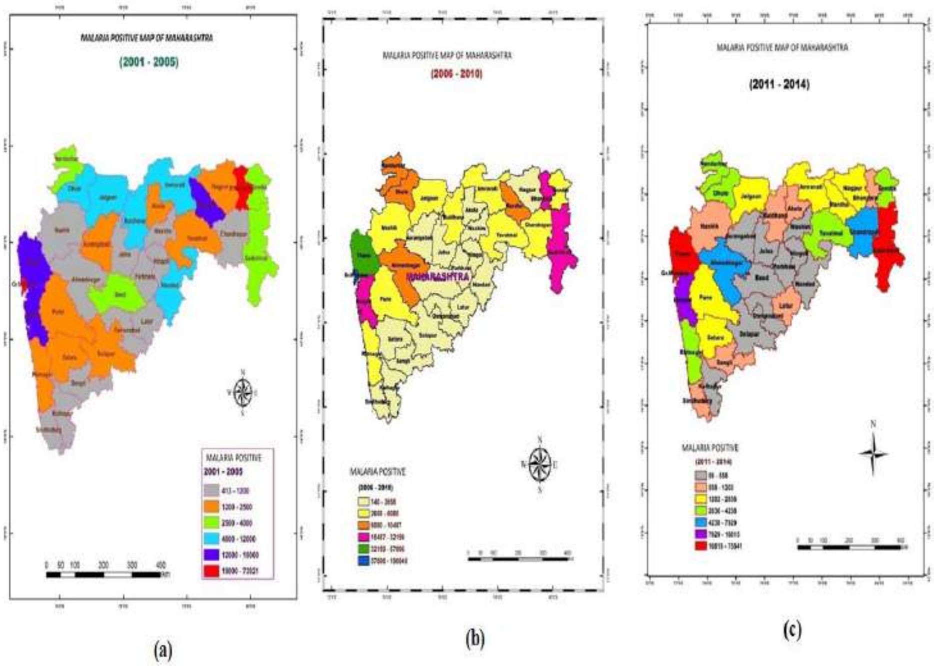


Figure 11: Maps shows the positive cases of malaria patients from (a) 2001-2005 (b) 2006-2010 (c) 2011-2014. (Maharashtra district boundaries: natural earth.com)

8. CONCLUSION

In this paper, it is focused on creating, sharing and accessing the spatial health information on the cloud computing environment in developed *GeoCloud4Health* model. The developed model adopts flexible and distributive structure, and provides an efficient mechanism for the delivery and generation of value-added spatial information by extending the concept of spatial web services in the field of health sector.

The experience gained in using open source GIS software suggests that various tools like QGIS and its QGIS plugin for implementation of spatial web services and creation of spatial datasets. In our future studies, we will add different variety of intelligent services and practicability aspects on SDI model development and implementation of successful SDI Framework at international level in education, watershed management and energy management sector. It will also plan to execute to assist the cloud computing environment by implementing fog computing paradigm at the edge of the client.

ACKNOWLEDGMENT

The authors express thanks to the various departments in the respective universities for extending their support and facilities to carry the research work. Authors are grateful to the National vector borne disease control program (NVBDCP) New Delhi for their support in data collection. The authors are also grateful to the reviewers to give their valuable comments for improvement of the manuscript.

REFERENCES

Barik, R.K. and A.B. Samaddar (2014).Service Oriented Architecture Based SDI model for Mineral Resources Management in India.Universal Journal of Geoscience, *2*(1), 1-6.

Barik, Rabindra K., Arun. B Samaddar, Shefalika G. Samaddar (2011). Service Oriented Architecture based SDI model for Geographical Indication Web Services, International Journal **of** Computer Applications, 25(4), 42-49.

Barik, R.K., A. B. Samaddar,and R.D. Gupta (2009). Investigations into the Efficacy of Open Source GIS Software, International Conference, Map World Forum, on Spatial Technology for Sustainable Planet Earth,Hyderabad, India.

Barr, M. (2005). Open Source versus Proprietary Software, A Discussion, at http://www.matthewbarr.co.uk/opensource.htm, [accessed 27 April, 2017].

Bastian Schaffer, Bastian Baranski, Theodor Foerster (2010). Towards Spatial Data Infrastructures in the Clouds. Spatial Thinking Lecture Notes in Geoinformation and Cartography, 0 399-418.

Buyya, R., C.S. Yeo and S. Venugopal (2008). Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. In High Performance Computing and Communications, 2008.HPCC'08. 10th IEEE International Conference on, 5-13

Camara, G. (2004). Developing Open-Source GIS: What are the Challenges? INPE – Brazil, at http://www.dpi.inpe.br.

Chen, Z., N. Chen, C. Yang and L. Di (2012).Cloud computing enabled web processing service for earth observation data processing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5(6), 1637-1649.

Davies, Jessica (2003). Expanding the spatial data infrastructure model to support spatial wireless applications, PhD Thesis, Department of Geomatics, The University of Melbourne, Australia.

Evangelidis, K., N. Konstantinos, M. Stathis, P. Constantine (2014). Spatial services in the Cloud, Computers & Geosciences, 63, 116-122.

Gao, S., M.Darka, A. Francois, Yi.Xiaolun and David J Coleman(2008). Online GIS services for mapping and sharing disease information, International Journal for Health Geographics.[hosted at: http://www.ijhealthgeographics.com/content/7/1/8]

GSDI Cookbook (2008). Developing spatial data infrastructure: The SDI Cookbook, Ver 2.0., at: http://www.gsdidocs.org/GSDIWiki/ index.php/ Main_Page. [accessed 28 January, 2017]

He, L., P. Yue, L. Di, M. Zhang and L. Hu (2015). Adding spatial data provenance into SDI—a service-oriented approach. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, *8*(2), 926-936.

Huang, Q., C. Yang, K. Liu, L. Kai, C. Xu, J. Li, Z. Gui,M. Sun, and Z. Li (2013). Evaluating open-source cloud computing solutions for geosciences. Computers & Geosciences, 59, 41-52.

Internet-1: https://www.ncbi.nlm.nih.gov › NCBI › Literature › PubMed Central (PMC) (Access on 23rd August, 2017)

Internet-2: http://www.maharashtra.nic.in/images/mapMaharashtraB.jpg (Access on 23rd August, 2017)

Internet-3:

http://qgiscloud.com/rabindrabarik2016/malaria?mobile=false (Access on 23rd January, 2017)

Internet-4:

http://qgiscloud.com/rabindrabarik2016/malaria?mobile=true (Access on 23rd January, 2017)

IPCC (2007). Climate Change: Impacts, Adaptation and Vulnerability. Working Group II Contribution to the IPCC.

Leidig, M., T. Richard (2015). Free software: A review, in the context of disaster management. International Journal of Applied Earth Observation and Geoinformation, 42, 49-56.

Mall, Rajiv (2004).Fundamentals of Software Engineering, Rev.2nd Edition, Prentice-Hall of India Pvt.Ltd, India.

Mansourian, A., A. Rajabifard, MJ.ValadanZoej,and

I. Williamson(2005). Using SDI and web based system to facilitate disaster management, International Journal of Computers and Geosciences, Vol.32, pp. 303-315.

Morris, Steven P. (2006). Spatial Web services and geoarchiving: New opportunities and challenges in geographic information service. Library trends, 55(2), 285-303.

Puri, S. K., S. Sahay, and Y. Georgiadou(2007). A Metaphor-Based Sociotechnical Perspective on Spatial Data Infrastructure Implementations: Some Lessons from India, Research and Theory Advancing Spatial Data Infrastructure Concepts, ESRI Press, pp.161-17.

Rajabifard, A., M.E.F. Feeneyand I.P. Williamson(2002). Future Directions for SDI Development, International Journal of Applied Earth Observation and Geoinformation, ITC, The Netherlands, Vol.4, No.1, pp. 11-22.

Ramachandra, T.V., U. Kumar(2004). Geographic Resources Decision Support System for Land Use, Land, Cover Dynamics Analysis.Proceedings of the FOSS/GRASS User Conference, Bangkok Thailand.

Rawat, S. (2003). Interoperable Geo-Spatial Data model in the Context of the Indian NSDI, Thesis (Master), ITC, The Netherlands.

Vaccari, L., P. Shvaiko, M. Marchese(2009). A geo-service semantic integration in Spatial Data Infrastructure. International Journal of Spatial Data Infrastructures Research, 4, 24-51

Vanmeulebrouk, B., U. Rivett, A. Ricketts and M. Loudon (2008). Open source GIS for HIV/AIDS management, International Journal for Health Geographics, [hosted at: http://www.ij-healthgeographics.com/ content/ 7/ 1/ 53]

Wu, Bian, Xincai Wu, Jian Huang. (2010). Spatial data services within Cloud computing environment.IEEE International Conference on Audio Language and Image Processing (ICALIP), 1577-1584.

Xiaolin, Lu. (2005). An Investigation on Service Oriented Architecture for constructing Distributed WebGIS Application.IEEE International Conference on Services Computing (SCC'05), 1, 191-197.

Yang, Chaowei, Michael Goodchild, Qunying Huang, Doug Nebert, Robert Raskin, Yan Xu, Myra Bambacus, and Daniel Fay (2011). Spatial cloud computing: how can the spatial sciences use and help shape cloud computing? International Journal of Digital Earth, 4(4), 305-329.

Yang, Chaowei, Robert Raskin, Michael, Goodchild, Mark Gahegan. (2010). Spatialcyberinfrastructure: past, present and future. Computers, Environment and Urban Systems, 34(4), 264-277.

Yue, P., Zhou, H., Gong, J. and Hu, L. (2013).Geoprocessing in Cloud Computing platforms–a comparative analysis. International Journal of Digital Earth, 6(4), 404-425.

Sapna. A. Sasane
Department of Geography, University of Pune, India.