

# GIS-based Computational Tools & Techniques for Multidimensional Data Analysis & Visualization

Sanjay Srivas<sup>1</sup> and Dr. P. G. Khot<sup>2</sup>

<sup>1</sup>Research Scholar, Dept. of Electronics & Information Technology, RTM Nagpur University Nagpur, Maharashtra, India.

<sup>2</sup>Professor Dept. of Statistics, RTM Nagpur University Nagpur, Maharashtra, India.

## Abstract

The increasing use of Geographic Information System (GIS) in various domains and data pertaining to it, there is huge demand of analytical tools which can provide best and efficient results for Decision Making System (DSS). This paper explores the tool and techniques available for multidimensional database analysis & visualization. There are various tools available in GIS software's which provides the best results for Geo-spatial data. These tools are used for various overlays, union for combination of attributes of data, extraction and statistics generation. In this paper, we mainly focused on Quantum GIS (QGIS) open source tool which can efficiently support many different visualization functions and spatial data analysis models. Furthermore, user can effectively use features (like, Edit Spatial Information, Analysis, and exporting Graphical Maps.) available in this platform. This tool is available at <https://www.qgis.org/en/site/>

**Keywords:** GIS Analysis, QGIS, Visualization, Geospatial, GIS Tools, QGIS

## INTRODUCTION

With the exponential growth of the usage of web map services, the analysis of geo data has become more popular. The use of three dimensional (3D) visualization techniques for both Remote Sensing and GIS [1] integration has been widespread for quite some time now. In the present day scenario use of the GIS plays very important and significant role in decision making process across various domain such as Real State, Water Management, Crime Mapping and Disease Analysis etc. GIS is one of the most efficient tools for preparation of Decision Support Systems (DSS). Analysis of large amount of spatial data is time and resource consuming. For efficient Spatial Data Analysis (ESDA), there are various tools present which can give better analysis results in less time consuming manner. The Spatial Data Management toolbox provides a rich and varied collection of tools that are used to develop, manage, and maintain feature classes, datasets, layers, and raster data structures.

## QGIS

Quantum GIS (QGIS) is a cross-platform Geographic Information System (GIS). It offers support for vector and raster formats, including spatially enabled tables in PostgreSQL using PostGIS, common GIS vector formats such as Shapefiles, and geo-referenced rasters (TIFF, PNG, and

GEOTIFF). Several add-ons (Plugins) are available to dynamically add new features in QGIS. Plugins are small programs that add functionalities to QGIS. Plugins are controlled through the plugins manager which reads the QGIS repository online and keeps track of all online available repositories. Viewing of GRASS layers (vector and raster) is provided by a complement using additional plugin. GRASS vector layers can be easily edited in QGIS tool.

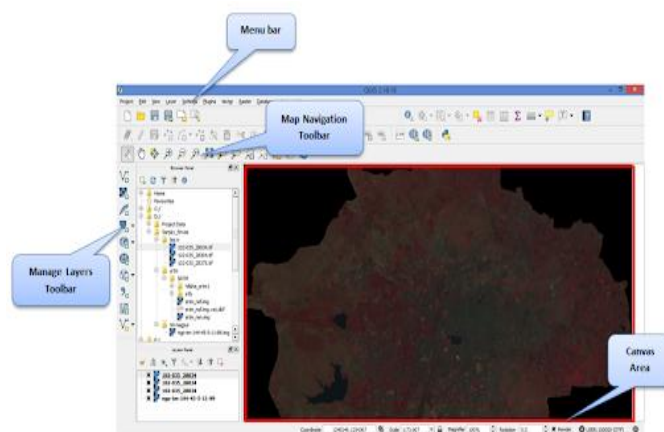


Figure 1: Main page layout of QGIS.

The rest of this paper is organized as follows: Section 2 presents GIS Spatial Data Analysis; Sections 3 Types of GIS Data Analysis; Section 4 describes the Visualization Outline; Section 5 presents Tools & Techniques for GIS Data Analysis & Visualization; and finally Section 6 conclude the paper.

## GIS SPATIAL DATA ANALYSIS

With a GIS software and related database, we can display, query, and organize data geographically and solve problems by uncovering and analyzing trends and patterns. The spatial analysis functions in GIS [2] include distance mapping, proximity mapping, weighted-distance mapping, density function, surface functions and local statistical functions etc. (ESRI, 1996). Spatial queries are achieved by creating a series of queries based on the spatial relationship of map functions. The spatial operators in the queries define the spatial relations existing between the characteristics of the map. Most spatial operators can be combined to respond to complex spatial queries [3].

## TYPE OF GIS DATA ANALYSIS

### A. Spatial overlay

Overlay operations are part of most spatial analysis processes and typically form the core of GIS projects. These operations combine different maps and, therefore, provide new information that was not present in the individual maps. In overlay operations, new spatial elements are created based on multiple input maps. Overlay operations [4] are performed only on raster maps in the integrated soil and water information system (ILWIS). The structure of the raster data is particularly suitable for such operations, since all the maps used in the analysis have the same geographical reference. ILWIS has a powerful tool for combining maps, called Map Calculation. Many maps can be combined simultaneously using arithmetic, relational or conditional operators and many different functions.

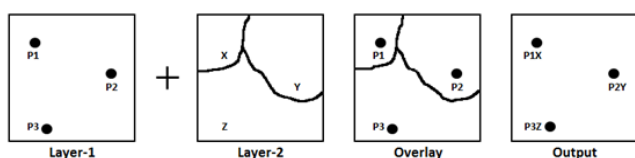


Figure 2: Example of Image overlay

### B. Boundary Analysis

The main purpose of the analysis of limits (Boundary) appears in GIS because "a problem created during spatial analysis, caused by arbitrary or discrete boundaries that are imposed on spatial data that represent non-discrete spatial phenomena or without limits. Boundary analysis [5], which is often referred to as districting, helps define regions according to certain criteria. This procedure is used to define area of specific demographic characteristic. Since districting is normally an iterative process involving the development of numerous scenarios based on various combinations of desired criteria, the computing power of the GIS proves to be a real timesaver. Rather than struggling with paper maps and adding machines, it is able to interactively define proposed boundaries and have related population totals automatically computed in a matter of minutes. The redistricting of the study area is based on the majority ethnic group. The process involves the assignment of a certain number to each group. Prior to using GIS, developing of clusters is done using the statistical techniques.

### C. Buffer Analysis

Buffer analysis [6] is used for identifying areas surrounding geographic features. The process involves generating a buffer around existing geographic features and then identifying or selecting features based on whether they fall inside or outside the boundary of the buffer.

This process is used to identify neighborhood for a specific ethnic group. The purpose of this; is to cluster households of the same ethnic group. The process involves representing each household on the map as a point. Each point is coded with

information that pertained to the type of ethnicity, as well as the other demographic characteristic.

### D. Information Retrieval

Information retrieval is one of the common features in GIS which is used to point at any location, object and any specific area. GIS data is stored in the form of pixels and same can be retrieve by using the information retrieval process from the Database System which holds the information about the GIS images.

In order to understand this process, this always has one question word in-built in it i.e. "Where". If user is trying to locate one Point, location and area, some input has to be given to the system and based on input it will fetch the desired results. Retrieval is the ability of the GIS based database system to get back on required data that were previously stored.

### E. Search by Attribute

Most GIS systems include a relatively simple relational database manager as part of the package or simply rely on the output capabilities of a database system. All DBMS include functions for displaying basic data. Searches by attribute are controlled by the capabilities of the database manager. Find is the basic attribute search. Find is intended to get a single record. Find can be browse or by searches. Examples include show attributes, show records, generate a report, find, recode, select, renumber, sort, compute allows the creation of new attributes based on calculated values, restrict, join, replace; all are examples of data reorganization. Attribute queries aren't deeply toward for geological search as it has slight provision to identify location, hence it is just as modest support in geographical searching needs [7].

### F. Search by Geographies

In a GIS map database, records are stored as features. Spatial retrieval of GIS information is the generating maps, which enable user to visually search for information and highlights the result. For example, to generate a report; the spatial equivalent would to produce a finished map; the spatial equivalent of a find is locate. The spatial equivalents of the DBMS queries determine the location of the feature set or the construction of new GIS layer. Typically, geographic searches include browsing the map and picking features, Spatial searching, Spatial sorting to identify features that result from attribute sorting, Recoding features spatially, that is changing the scope of their attribute, is equivalent to spatial merge, Spatial select is to extract specific features. Combinations of spatial and attribute queries can build some complex and powerful GIS operations, such as weighting e.g. dominant ethnic group in an area. Entire suites of geographic searches are searches and tests by relations of points, lines, and areas. Typical GIS searches are point in polygon, line in polygon, and point distance to line.

## G. Geoprocessing

Geoprocessing [8] is most common feature of any GIS software. The fundamental purpose of geoprocessing is to provide tools and a framework for the analysis and management of geographical data. Under the geoprocessing, modeling and analysis capabilities are present in QGIS software. By using geoprocessing, various operations can be performed like polygon overlays, complex regression analysis and image classification etc.

## VISUALIZATION OUTLINE

The outline of visualization is based on four major components which are commonly used to carry out any GIS data analysis and Visualization:



### A. Data Modelling

The fundamental aim of data models is to help the improvement of data frameworks by giving the definition and organization of information. Spatial information is referenced by its position facilitates (for ex. Longitude & Latitude) in the geographic information system (GIS) for geospatial analysis. The common applications of spatial data usage are Geometric computation and geo-spatial relationships, Topological matching & pattern analysis and Multidimensional data representation. Data modeling [9] can be used in 2D as well as 3D data analysis. Raster-based stacked surfaces represent the surface of each hidden geological unit which is used to hold data on lateral variations. Vector-based stack-unit maps depict the vertical succession of geologic units to a specified depth [10].

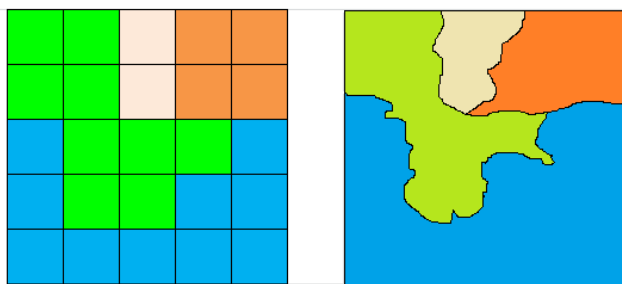


Figure 2: Representation of Raster & Vector data.

### B. Data management

Data management is the process of organizing and maintaining GIS data to meet the specific requirements. Typically GIS data is maintained at centralized server for the particular geographic location, it can be stored in distributed servers using cloud as a service platform. The emphasis on data management comes into picture because effective data management is required to manage GIS operations effectively. Data management contains the tools like Archiving, Data Comparison, Distributed GeoDatabase, GeoDatabase Administration etc.

### C. Data Estimation

Data estimation is carried out using remote sensing and the geographic information system (GIS). It has long been used to estimate the population of large geographical areas. Mainly, these methods are broadly classified into two categories:

- I. Area Interpolation Method
- II. Statistical Modelling Method.

### D. Data Visualization

Visualizing large amounts of information interactively is one of the most interesting and useful GIS features. High-powered computers can alter any element of the screen "on the fly", changing not only the appearance of the graphic image, but also its interpretation. Displaying data using current computer technology and interactive GIS has several advantages compared to traditional paper maps.

- GIS is completely interactive. By adding new data fields, removing them, changing the color scheme or the shape of the map, adding text, moving symbols and a range of different capabilities, the user gets a high level of flexibility and power.
- The GIS screens are expandable and pannable. Moving on the screen gives the user new perspectives, more (or less) details and a new perception.
- Users can take advantage of computationally intensive features such as "draping" a perspective view on a surface (such as a digital elevation model) or creating a three-dimensional print on a 2D screen (the computer screen) using complex shading algorithms and rendering.

## TOOLS & TECHNIQUES FOR GIS DATA ANALYSIS & VISUALIZATION

**GIS as a service (GISAAS):** GIS based services are broadly classified into 3 major categories:

**Table 1:** List of GISAAS services available for use

<b>PaaS – Platform as a Service: GIS data Analysis &amp; Processing Services</b>	
<b>S. No.</b>	<b>Name of the Platform based Services available</b>
1	ArcGIS Online
2	Google Maps Javascript API version
3	Microsoft Bing Geocode Dataflow API
4	Here Maps Javascript API version
5	US Census Geocoder
<b>SaaS – Software as a Service: Software available as a Service on the Internet</b>	
<b>S. No.</b>	<b>Name of the Software based Services available</b>
1	ArcGIS Online – ESRI's cloud based version of ArcGIS
2	QGIS – is a cross-platform Geographic Information System (GIS)
3	Mapbox – is provider of custom online maps for websites
4	CartoDB – It is online mapping platform that offers an open source, cloud based SaaS model
<b>DaaS – Data as a Service: Data or Content Services</b>	
<b>S. No.</b>	<b>Name of the Data based Services available</b>
1	OpenStreetMap
2	ArcGIS Online
3	Microsoft Bing Maps
4	Apple Maps
5	Google Maps

Various tools are available for data visualization, analysis to define the competitive approach in this field. Some of the most commonly used tools are as follows:

**ArcGIS:**

ArcGIS [11] is a desktop software mainly used for working with maps and Geographic Information System. It is used to create and use maps, compile geographic data, analyze mapped information, share and discover geographic information, use maps and geographic information in various applications, and manage geographic information in a database. ArcGIS for Desktop includes several built-in applications such as ArcToolbox, ArcMap, ArcCatalo, ArcScene, ArcGlobe, and ArcGIS Pro. ArcToolbox includes an assortment of tools necessary for analyzing, converting and managing GIS data. GIS is often used as an analysis tool. ArcToolbox contains a set of preformatted analysis tools to enable a user for pinpoint locations fitting into specific criteria.

**MATLAB:** Matlab (matrix laboratory) [12] is licensed software that mainly used in high-performance numeric calculations and visualization. It supports developing applications with graphical user interface (GUI) features. This also has various advance features to deal with the graphical image processing and analysis. Matlab also offer valuable collection of functions that handles import of images and video

signals directly from computer to the Matlab environment which can further used for analysis purpose.

**Google Earth:**

Created by Keyhole Inc. and originally named Earth Viewer, Google Earth [13] maps the earth by superimposing satellite images and aerial photos on a virtual globe. GE allows the user to interactively browse the globe in 3D view and to zoom from space into street level views. Data is streamed from Google's Server upon request to the client computer. Google Earth has the capability to combine the multi-resolution raster image representation of the planet with any kind of georeferenced data.

**UrbanSim:**

UrbanSim [14] is an open source urban simulation system created by Paul Waddell, University of California, Berkeley. It has various collaborators features to support metropolitan land use, transportation, and environmental planning etc. Recent studies show that UrbanSim has become the most used land use model system by US planning agencies.

**Google Charts:**

Google Charts is the most common charting solution for most of the web. It is very flexible and has an excellent set of development tools. It is a particularly useful tool for specialized visualizations, such as geocharts and indicators, and also includes embedded animation and user interaction controls.

**Polymaps:**

Polymaps is a JavaScript library for creating dynamic and interactive maps in modern web browsers. Polymaps provides fast visualization of multiple zoom datasets on maps and supports a variety of visual presentations for tiled vector data. In addition to the usual OpenStreetMap, CloudMade, Bing and other web-based map providers; polymaps provides image-based web maps for effective analysis. Because Polymaps can load data at a full range of Polymaps also uses SVG (Scalable Vector Graphics) to display information.

**SpaceCurve:**

Space Curve can discover underlying patterns in multidimensional geodata. Geodata is different data than normal data as mobile devices create new data really fast and not in a way traditional databases are used to. The platform provides real-time information on complex contextual and ad hoc relationships. This is possible through processing and continuous analysis of geospatial, temporal, social graphic and other large data sets as per the real time data available. This pattern-based approach allows you to scale widely and efficiently within large, distributed, multidimensional data.

**SensorUp:**

SensorUp provides the Cloud hosting and SDKs, based on the Open Geospatial Consortium SensorThings API standard, allowing developers to build applications that connect with Internet of Things sensor observations and tasking. SensorUp's OGC standard-based platform includes sensor data, analytics, and mapping SDK.

**Table 2:** List of GIS Analysis & Visualization Tools

S. No.	Product Name	Product Feature	Details
1.	Google Earth	3D representation of Earth based on satellite image.	Open access
2.	ArcGIS	Overlay, Map Algebra, Surface Approximation, Network Analysis	Licensed
3.	QGIS	Analysis and edit Spatial information, Composing	Open Source

S. No.	Product Name	Product Feature	Details
		and exporting graphical maps.	
4.	MATLAB	MATLAB for Geo Image Optimization	Licensed
5.	TerraFly	Visualization of spatial and remote sensed imagery.	Web-enabled system
6.	UrbanSim	Dynamic traffic assignment, Emissions and land cover change	Open Source
7.	Google Charts	It creates graphical charts from user-supplied information.	Web Service
8.	Polymaps	For making dynamic, interactive maps in modern web browsers	Web-based
9.	SpaceCurve	Patterns analysis in multidimensional Geographical data.	Web-based
10.	SensorUp	Patterns in multidimensional Geo data	Web-based

**CONCLUSION**

This paper has demonstrated various methods & techniques for multidimensional data analysis & visualization. For geographical data analysis, tools should have enough features to explore upto maximum possibility and also should have enough plug-in to explore further. Geo analyst & researcher can use these tools to improve the efficiency in data analysis. High quality data analysis and the level of information can be obtained after using the scientific data analysis tools. In this paper, we have explained QGIS, one of the widely used open source tool for the geographical data analysis and also explored the various data analysis & visualization tools (Open source/Licensed) are discussed in brief.

**ACKNOWLEDGEMENT**

I would like to sincerely thank to Mrs. Tutu Sengupta, Associate Scientist, Maharashtra Remote Sensing Application Centre for providing support and guidance during this studies. Her constant guidance and support during this study kept me motivated all the time.

## REFERENCES

- [1] Goodchild, M.F., Geographical information science, International Journal of Geographical Information Systems, 1992 Vol. 6(1): 31-45.
- [2] Anselin, L., Syabri, I., & Kho, Y. GeoDa: An introduction to spatial data analysis. Geographical analysis, 2006 Vol. 38(1), 5-22.
- [3] Lee, K., Ganti, R. K., Srivatsa, M., & Liu, L. Efficient spatial query processing for big data. Paper presented at the Proceedings of the 22<sup>nd</sup> ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2014.
- [4] Dirk Tiede, A new geospatial overlay method for the analysis and visualization of spatial change patterns using object-oriented data modeling concepts, 2014, 227-234.
- [5] G. M. Jacques S. Maruca M.-J. Fortin, From fields to objects: A review of geographic boundary analysis, Springer 2002, Vol.2, 221-241.
- [6] Lan Mu, A shape-based buffering method, 2008, 01-15.
- [7] Harvey J. Miller, Elizabeth A. Wentz, Representation and Spatial Analysis in Geographic Information Systems, 2003 93(3), 2003, pp. 574–594
- [8] Pons NAD, Pejon OJ, Zuquette LV. Use of geoprocessing in the study of land degradation in urban environments: the case of the city of São Carlos, state of São Paulo, Brazil. Environmental Geology, 2007 727-739
- [9] Wang, H. A Large-scale Dynamic Vector and Raster Data Visualization Geographic Information System Based on Parallel Map Tiling, 2011 .
- [10] Choi, Y. Choi, J. Um, M. Park, Finding least-cost paths across a continuous raster surface with discrete vector networks Cartography and Geographic Information Science, 2014, 75-85
- [11] Lourdes A. Funtanilla, GIS pattern recognition and rejection analysis using MATLAB. In: Proceedings of ESRI, 2004
- [12] Lourdes A. Funtanilla, GIS pattern recognition and rejection analysis using MATLAB. In: Proceedings of ESRI, 2004
- [13] Google Inc., Google Earth, available online at: , 2007.
- [14] P. Waddell, “UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning”, *Journal of the American Planning Association*, Vol. 68(3):297-314, 2002.