





QUICK REFERENCE GUIDE on Geographic Information System

Regional Remote Sensing Centre-North National Remote Sensing Centre, ISRO

Quick Reference Guide

On

Geographic Information System

Regional Remote Sensing Centre-North National Remote Sensing Centre Indian Space Research Organisation

भारत सरकार अन्तरिक्ष विभाग राष्ट्रीय सुदूर संवेदन केन्द्र बालानगर, हैदराबाद-500 037, तेलंगाना, भारत देलिफोन : +91 40 23884001 / 04



Government of India Department of Space **National Remote Sensing Centre** Balanagar, Hyderabad-500 037, Telangana, India Telephone: +91 40 23884001/04

डॉ. प्रकाश चौहान / Dr. Prakash Chauhan उत्कृष्ट वैज्ञानिक & निदेशक Outstanding Scientist & Director



Foreword

National Remote Sensing Centre (NRSC) is one of the primary centres of Indian Space Research Organisation (ISRO), Department of Space (DOS) for developing remote sensing applications, establishing ground stations for receiving satellite data and generating high-quality satellite data and aerial data products. Regional Remote Sensing Centres (RRSC) are part of NRSC supporting various remote sensing tasks specific to their regions at the national level.

Geographic Information Systems (GIS) have revolutionized the way we understand and interact with our world. The ability to visualize, analyse, and interpret spatial data has transformed numerous fields, from urban planning and environmental management to disaster response and public health. The growing reliance on GIS underscores its importance in contemporary problem-solving and decision-making processes.

This Quick Reference Guide on GIS, meticulously prepared by RRSC-North, serves as a comprehensive guide for students, researchers, and professionals. It covers a wide range of topics, from the basics of GIS and spatial data collection to advanced analytical techniques and real-world applications. The reference guide is designed to offer clear explanations, practical examples, and conceptual illustrations to facilitate effective learning and application of GIS principles.

I am confident that this reference guide will serve as a valuable resource, equipping readers with the knowledge and skills necessary to leverage GIS technology in their respective fields. It is our hope that this guide will inspire and enable you to harness the power of GIS for innovative solutions and informed decision-making.

(Prakash Chauhan)

September 17, 2024

भारतीय अन्तरिक्ष अनुसंधान संगठन Indian Space Research Organisation

भारत सरकार अंतरिक्ष विभाग भारतीय अंतरिक्ष अनुसंधान संगठन **राष्ट्रीय सुदूर संवेदन केंद्र** प्लॉट सं. ७, प्लानिंग एरिया सेंटर, जे.बी. टीटो मार्ग सादिक नगर, नई दिल्ली 110 049

डॉ. एस. के. श्रीवास्तव

मख्य महाप्रबंधक (क्षेत्रीय केंद्र)

Dr. S. K. Srivastav

Chief General Manager (Regional Centres)



Government of India Department of Space Indian Space Research Organisation **National Remote Sensing Centre** Plot No. 7, Planning Area Centre, J.B. Tito Marg Sadiq Nagar, New Delhi 110 049

दूरभाष/ Telephone: 011-26400754 ईमेल/ Email: cgm@nrsc.gov.in sksrivastav@nrsc.gov.in

From The Chief General Manager's Desk

National Remote Sensing Centre (NRSC) has five Regional Remote Sensing Centres (RRSCs) spread across the country. These centres are involved in addressing the local and regional issues using space and geospatial technology and also actively participate in capacity building as well as outreach activities in their respective regions. RRSC-North, located at New Delhi, caters to the needs of users in the Northern states of India viz. Delhi, Himachal Pradesh, Jammu & Kashmir, Uttar Pradesh and Uttarakhand. RRSC-North also organises training programmes and workshops on topics related to remote sensing, geographic information systems, digital image processing and their applications for government officials, academia and students.

I am extremely happy that RRSC-North has prepared a "Quick Reference Guide on Geographic Information Systems (GIS)." Basic concepts of GIS are covered in a crisp and lucid manner in this compilation. I am sure that this reference guide will be helpful for the beginners to get familiar with the concepts of GIS in a short time.

रूस, के. स्त्रीयाम्बर

S. K. Srivastav Chief General Manager (Regional Centres) NRSC

भारत सरकार अंतरिक्ष विभाग भारतीय अंतरिक्ष अनुसंधान संगठन



Government of India Department of Space Indian Space Research Organisation National Remote Sensing Centre Regional Remote Sensing Centre- North Plot No. 7, Planning Area Centre, J.B. Tito Marg Sadig Nagar, New Delhi 110049

दरभाष/ Telephone: 011-26400746 ईमेल/ Email: sameer_s@nrsc.gov.in amrc n@nrsc.aov.in

राष्ट्रीय सुदूर संवेदन केंद्र क्षेत्रीय सदर संवेदन केंद्र – उत्तर प्लॉट सं.७, प्लानिंग एरिया सेंटर, जे.बी. टीटो मार्ग सादिक नगर, नई दिल्ली 110 049

डॉ. समीर सरन उप महाप्रबंधक, क्षेत्रीय सुदूर संवेदन केंद्र - उत्तर Dr. Sameer Saran

Deputy General Manager, Regional Remote Sensing Centre- North

About the Book

Reference guide on Geographic Information Systems (GIS) provides a thorough understanding of key concepts and practical applications of GIS. It covers an overview of GIS, data models & structures, basic components, functionalities, and significance across various fields. It emphasizes the importance of coordinate systems and map projections for accurately representing the Earth's surface. The book also delves into spatial analysis techniques, focusing on vector and raster analysis for extracting meaningful information from spatial data.

Further, the reference guide explores network analysis principles and techniques, particularly in transportation and utility systems. It details the elements and design principles of effective map composition. The inclusion of Free and Open-Source Software for Geospatial (FOSS4G) applications highlights various tools and techniques. Insights into the Global Navigation Satellite System (GNSS), including GPS and other satellite-based navigation systems, are also provided.

The study material will be helpful for understanding the GIS concepts in an easier and lucid manner for the benefit of the beginners in the field of GIS.

Sameer Saran Deputy General Manager, RRSC-North, NRSC

Index

1.	Introduction to GIS1	
	Dr. Vinod Kumar Sharma	
2.	GIS Data Models	
	Khushboo Mirza	
3.	Coordinate Systems and Map Projections20	0
	Anurag Mishra	
4.	Spatial Analysis: Vector Analysis & Raste	r
	Analysis	l
	Khushboo Mirza	
5.	Network Analysis 41	Ĺ
	Abhinav Kumar Shukla	
6.	Map Composition	
	Anurag Mishra	
7.	FOSS4G - Open-Source Tools & Technique	S
	Dr. Vinod Kumar Sharma	
8.	Global Navigation Satellite System (GNSS)
0.	74	'
	Javant Singhal	
0	AI Applications in Remote Sensing and CI	2
9.	Al Applications in Remote Sensing and Gi	,
	Allulag Milsilla	

Chapter – 1

Introduction to GIS

1.1 Introduction

Geographic Information Systems (GIS) are systems specifically designed for the collection, storage, manipulation, analysis, management, and presentation of spatial or geographic data. GIS can be defined as a computer-based tool that encompasses four key functions for handling georeferenced data: inputting data, managing and retrieving data, manipulating and analysing data, and outputting data. This comprehensive functionality allows users to explore spatial patterns and relationships, thereby enhancing decisionmaking processes in various fields and areas.

In the 1960s, the need for spatial data management and analysis began to emerge. The first recognized GIS was developed by Roger Tomlinson in Canada in 1963, known as the Canada Geographic Information System (CGIS). This system was designed to analyse land use and assist in resource management and planning. Throughout the 1970s and 1980s, advancements in computer technology and the development of more sophisticated software, such as the introduction of the Environmental Systems Research Institute (ESRI) by Jack Dangermond, propelled the growth of GIS. The advent of personal computers in the 1980s and the subsequent proliferation of the internet in the 1990s further expanded the accessibility and functionality of GIS. Today, GIS is an indispensable tool in various fields, including urban planning, environmental management, and public health, driven by continuous technological innovations and an ever-growing repository of spatial data. Recent advancements in GIS, National Spatial Data Infrastructure (NSDI) emphasis on information transparency and sharing, with the recognition that spatial information is a national

resource and citizens, society, private enterprise and government have a right to access it, appropriately.

Geographic Information Systems (GIS) and remote sensing play crucial roles in a variety of sectors across India. The Indian Space Research Organisation (ISRO) leverages GIS and its Indian Remote Sensing (IRS) satellites to monitor and manage natural resources effectively. In agriculture, these technologies are used to estimate crop acreage and yields, which are vital for ensuring food security (https://bhuvan-app1.nrsc.gov.in/agriculture/agri.php). GIS also enhances disaster management by providing real-time data and predictive models to assess and mitigate the effects of natural disasters like floods and earthquakes(https://bhuvanapp1.nrsc.gov.in/bhuvandisaster/). Urban planning benefits from GIS through its applications in designing smart cities, gram panchayat management, controlling urban sprawl, and improving infrastructure development(https://bhuvanpanchayat.nrsc.gov.in/). Moreover, environmental monitoring programs use GIS to track deforestation, soil erosion, and biodiversity, thereby promoting sustainable development initiatives nationwide (https://bhuvanapp1.nrsc.gov.in/moef/).

1.2 Components of GIS

Geographic Information Systems (GIS) are composed of several essential components that collaborate to capture, store, analyse, and present spatial data. These key components include hardware, software, data, people, and methods.

Hardware: This encompasses the physical devices necessary for GIS operations, such as computers, GPS units, and servers, which supply the required computing power and storage.

Software: GIS software includes the tools and applications used to process and analyse spatial data. Notable examples are ArcGIS,

QGIS, and GRASS GIS, which provide functionalities such as data input, spatial analysis, and map creation.

Data: Serving as the core of GIS, data comprises spatial data (geographic locations) and attribute data (descriptive information). Data can be sourced from satellite imagery, aerial photography, and field surveys.

People: Skilled personnel are vital for the effective operation of GIS systems. This category includes GIS specialists, analysts, and managers who interpret data, conduct analyses, and make informed decisions based on GIS outputs.

Methods: These involve the procedures and techniques used for data collection, analysis, and presentation, ensuring the accuracy and consistency of GIS operations, including best practices for data management and spatial analysis.

Each of these components is integral to the successful implementation and operation of GIS, enabling it to provide valuable insights across diverse applications.

GIS Software Producer	Main Product(s)
Esri (Environmental Systems	ArcGIS (ArcMap, ArcGIS Pro,
Research Institute)	ArcGIS Online)
QGIS	QGIS (Quantum GIS)
Autodesk	AutoCAD Map 3D, Autodesk
	InfraWorks
Bentley Systems	Bentley Map, MicroStation
Hexagon Geospatial	ERDAS IMAGINE, GeoMedia
Pitney Bowes Software	MapInfo Pro
GRASS GIS	GRASS GIS (Geographic
	Resources Analysis Support
	System)

Table 1.1: List of GIS software's and their main products

Intergraph Corporation (now part of Hexagon)	GeoMedia, ERDAS Imagine
Mapbox	Mapbox GL, Mapbox Studio
Boundless	Boundless Suite (including
	OpenGeo Suite, Boundless Server)
Cadcorp	Cadcorp SIS (Spatial Information
	System)
SuperMap Software Co., Ltd.	SuperMap GIS
Maptitude	Maptitude GIS
Manifold System	Manifold System
Global Mapper	Global Mapper

1.3 Geospatial data

Spatial data refers to information that describes the geographic location and characteristics of features on the Earth's surface (Fig. 1.1). Attribute data enriches spatial information with additional details, such as population density or land use classifications. Integrating spatial and attribute data empowers analysts to perform thorough spatial analyses, aiding decision-making in diverse fields like urban planning, environmental management, and public health.



Figure 1.1: Type of Geographical data.

Spatial data can be categorized into **discrete features** and **continuous features. Discrete features** represent specific,

identifiable objects with clear boundaries, such as buildings, roads, and trees. **Continuous features**, on the other hand, represent phenomena that vary continuously across space, such as elevation, temperature, and precipitation.

Geospatial database models are structures used to organize and store spatial data efficiently. Common models include the vectorbased feature dataset model and the raster-based grid model (Fig 1.2). The vector-based feature dataset model represents geographic features as points, lines, or polygons with attributes, ideal for complex shapes and precise spatial relationships. Conversely, the raster-based grid model organizes spatial data into cells, suitable for continuous phenomena and tasks like terrain modelling and remote sensing analysis. The raster model prioritizes simplicity and computational efficiency, whereas the vector model emphasizes accuracy.



Figure 1.2: Vector & Raster data model.

Topology refers to the spatial relationships and connectivity between geographic features. It helps ensure data integrity and

facilitates spatial analysis by defining rules for how features can be spatially related. Topology in GIS is defined as the spatial relationships between adjacent or neighbouring features in geographic plane. Mathematically, the topology assumes that geographic features occur on a two-dimensional plane. Through planar enforcement, spatial features can be represented through nodes (0-dimensional cells); edges, sometimes called line (onedimensional cells); or polygons/area (two-dimensional cells). Because features can exist only on a plane, lines that cross are broken into separate lines that terminate at nodes representing intersections rather than simple vertices. Different software offers different tools for maintaining and querying these spatial relationships. For example: disjoint, meet, equal, inside, covered by, contains, covers, overlap etc.

Coverages and **shapefiles** are common file formats for storing vector-based spatial data. Coverages are a legacy format developed by Esri, while shapefiles are widely used for their simplicity and compatibility across GIS software. A coverage stores a set of thematically associated data considered to be a unit. It usually represents a single layer, such as soils, streams, roads, or land use. In a coverage, features are stored as both primary features (points, arcs, polygons) and secondary features (tics, links, annotation).

A **Triangulated Irregular Network (TIN)** is a surface representation model used to depict continuous features, such as terrain elevation, using a network of irregularly spaced triangles. It is efficient in terms of data storage. The irregularity of TIN allows for lesser points to be used to represent smooth terrains. In this sense, TINs are more efficient than the raster format, where all cells are allocated a value, even if it is the same as the value of neighbouring cells. Dynamic segmentation models allow for the linear referencing of spatial data along linear features, such as roads or rivers, by dynamically segmenting them into smaller units based on specific attributes. **Attribute data** provide additional information about spatial features, such as population density, land use classification, or building height. Joining spatial and attribute data allows analysts to combine and analyse both types of data simultaneously, enabling more comprehensive spatial analysis and decision-making processes. Combining spatial and attribute data involves linking geographic information with detailed descriptive data to enhance dataset analysis and visualization. Spatial data, which encompasses coordinates and shapes representing various locations and areas, is integrated with attribute data that describes these locations. This combination facilitates comprehensive GIS analysis, enabling tasks such as mapping, spatial queries, and geostatistical analysis. By merging these datasets, users can uncover deeper insights into spatial patterns, relationships, and trends, thereby supporting more informed decision-making in fields.

1.4 GIS operations

GIS activities can be grouped into following; however, sequence can vary based on application:

- 1. Data capture and input
- 2. Data management
- 3. Data integration and display
- 4. Data exploration
- 5. Spatial analysis
- 6. GIS modelling

GIS operations involve systematic processes for managing, analysing, and visualizing spatial data. These operations begin with data capture from sources like GPS, remote sensing satellites, and surveys. Data management then ensures the quality and organization of both spatial and attribute data. Data integration combines different datasets for comprehensive analysis, while spatial analysis

techniques such as overlay, buffering, and querying reveal patterns and relationships. Geocoding translates addresses into geographic coordinates, and network analysis optimizes routes and connectivity. Remote sensing processes satellite and aerial imagery, and 3D visualization creates detailed models for examination. Spatial statistics uncover trends and correlations, and mapping and visualization produce informative maps and graphics. GIS models simulate the real-world processes and provide insights of geographic data patterns and relationships between them. Digital Elevation Models (DEM) represents terrain surface in 3D and created using terrain elevation data. Hydrological models like SWAT (Soil and Water Assessment Tool and HEC-HMS (Hydrologic Modelling System), models simulate the movement and distribution of water on the earth's surface. Network Models, analyse the connectivity and flow within networks. Spatial statistical models, apply statistical techniques to spatial data to identify patterns and relationships like Kriging and Hot spot analysis.

1.5 Conclusion

Geographic Information Systems (GIS) is a powerful technology that merges spatial and non-spatial data, offering deep insights into the geographic context of various phenomena. Utilizing GIS allows users to gather, manage, analyse, and visualize geographic information, uncovering patterns, relationships, and trends not easily seen through traditional data analysis. GIS is essential in fields like urban planning, environmental management, transportation, and public health, providing vital support for decision-making and strategic planning. The core functions of GIS—data acquisition, data management, spatial analysis, and visualization—collaborate to enable precise and detailed spatial analyses. As GIS technology continues to advance, its range of applications is set to grow, reinforcing its status as an indispensable tool in scientific research and practical applications across various domains.

Chapter 2

GIS Data Models

2.1 Introduction

Geographic Information Systems (GIS) are essential tools for analysing and managing spatial data. Central to the operation of GIS are data models, which provide the structure for storing, organizing, and analysing geographic information (Fig. 2.1). These models define how spatial data is represented within the system, making an understanding of their types crucial for effective GIS analysis and application. This chapter explores the two main types of data models used in GIS: raster and vector data models.



Figure 2.1: Geographic features represented by layers.

2.2 Spatial Data Models

Spatial data models describe the representation of geographic features within a GIS. These models can be broadly classified into

vector and raster models (Fig. 2.2), each suited to different types of spatial data and analyses.



Figure 2.2: Raster and vector representation of real world.

2.2.1 Vector Data Model

The vector data model uses geometric objects such as points, lines, and polygons to represent spatial features. Features are real-world objects such as roads, property boundaries, electrical substation sites and so on. A feature has a geometry (which determines if it is a point, polyline or polygon) and attributes (which describe the feature) (Fig. 2.3).



Figure 2.3: Representation of vector feature.

Every point or vertex associated with vector data contains the x,y location of the point. In vector data layers, the feature layer is linked to an attribute table (Fig. 2.4). Every individual feature corresponds to one record (row) in the attribute table.

Points

Points are the simplest form of vector data, representing discrete locations defined by a pair of coordinates (x, y). Examples include wells or bus stops.

Lines

Lines (or polylines) are sequences of points connected by straight segments. Lines are recorded as a series of ordered x,y coordinates; They are used to represent linear features such as roads, rivers, and pipelines.

Polygons

Polygons are enclosed areas formed by connecting multiple lines. Polygons are recorded as a series of x,y coordinates defining line segments that enclose a polygon. They represent features such as lakes, parks, and land parcels.



Figure 2.4: Vector and raster data model in geographic case.

Advantages of vector data-

- **Precision**: Vector data provides high precision in representing geographic features.
- **Topological Relationships**: Vector data can easily represent topological relationships, which is essential for network analysis and understanding spatial relationships.
- **Storage Efficiency:** Generally, vector data can be more storage-efficient for certain data types compared to raster data.

Disadvantages of Vector Data

- **Complexity:** Managing and processing vector data can be complex, especially with large datasets.
- **Computation Intensive:** Some spatial operations, like overlay analysis, can be computationally demanding.

2.2.2 Raster Data Model

The raster data model represents geographic features as a grid of cells or pixels, where each cell contains a value representing a specific attribute, such as elevation or land cover (Fig. 2.5). The value can be in the form of an integer, floating points or alphanumeric character. A point can be represented by a single pixel in the raster model. A line is a chain of spatially connected cells with the same value. Similarly, a water body in the raster data is represented as a set of contiguous pixels having the same value, representing a homogeneous area.



Figure 2.5: Raster data model.

Cells and Grids

In raster models, the geographic area is divided into a regular grid of cells. Each cell holds a value that represents a particular attribute (e.g., elevation, land cover type).

Continuous and Discrete Data

- **Continuous Data:** Represent phenomena like elevation or temperature, which change gradually over space.
- **Discrete Data:** Represent phenomena with distinct boundaries, such as different land use types (Fig. 2.6).



Figure 2.6: Type of raster data.

Advantages of raster data-

- **Simplicity:** The raster model is straightforward and easy to understand.
- **Analytical Efficiency:** Well-suited for mathematical modelling and spatial analysis, especially for continuous data such as satellite imagery and environmental data.

Compatibility: Integrates well with remote sensing data, useful in applications like climate modelling and terrain analysis.

Disadvantages of raster data-

- **Resolution Dependency:** The quality of raster data depends on its resolution, with higher resolutions requiring more storage.
- **Storage Requirements**: Large raster datasets can be storage-intensive.
- **Less Precision**: Compared to vector data, raster data can be less precise in representing boundaries and linear features.

2.3 Attribute Data Models

Attribute data models, store descriptive information about geographic features. It consists of the characteristics of spatial features that are independent of all geometric considerations. This information is typically organized in tables, with each row representing a feature and each column representing an attribute (Fig.2.7).



Figure 2.7: Attribute information for spatial data.

Attribute Tables

Attribute tables link non-spatial data to spatial features, enabling detailed descriptions and analyses. Each record (row) in the table corresponds to a geographic feature, while each field (column) holds a specific attribute.

Field Types

- Numeric Fields: Store numbers, such as population or area.
- **Text Fields**: Store descriptive text, such as names or types.
- Date Fields: Store dates and times.
- **Boolean Fields**: Store true/false values.

Relational Databases

Relational databases are commonly used to manage non-spatial data. They organize data into tables and define relationships between them using keys. This structure supports complex queries and efficient data management.

Advantages of Non-Spatial Data

- Detailed Information: Provides comprehensive descriptions of spatial features.
- Flexibility: Easily updated and queried.
- Integration: Can be integrated with spatial data for complex analyses and comprehensive reports.

Disadvantages of Non-Spatial Data

- Maintenance: Managing large datasets requires regular updates and maintenance.
- Complex Queries: Advanced queries can require significant computational resources and expertise.

2.4 Database Models

A database model serves as the theoretical framework of a database, dictating how data can be stored, organized, and manipulated within a database system. It establishes the structure provided by a specific database system. Different methods of organizing databases are known as database models. The primary database models include relational, network, hierarchical, and object-oriented database models (Fig. 2.8):

- In the **hierarchical model**, data are organized into a treelike structure with parent-child one-to-many relationships.
- In the **network model**, data are structured with records connected through pointers, classified into record types.
- In the **relational model**, data are stored in tables, with records organized into rows and columns.
- In the **object-oriented model**, data are represented as objects, each with unique attributes and operations, and classified into object types or classes.



Figure 2.8: Database models; (a) relational, (b) network, (c) hierarchical and (d) object-oriented database models.

2.5 Integrating Spatial and Non-Spatial Data

Effective GIS analysis often requires integrating spatial and nonspatial data. This integration allows for more detailed and informed decision-making.

Joining Data

Joining data involves linking attribute tables to spatial features using a common identifier. This process enables the combined analysis of spatial locations and their attributes.

Spatial Analysis

Spatial analysis involves a range of techniques that help in interpreting and understanding geographic data. These techniques enable the assessment of patterns, relationships, and trends within spatial data. Common types of spatial analysis include:

- **Overlay Analysis:** Overlay analysis is a fundamental technique in spatial analysis that involves superimposing multiple layers of spatial data to identify relationships and patterns. This method allows the integration of different datasets, such as land use, vegetation cover, and population density, to generate new insights.
- **Proximity Analysis**: Proximity analysis assesses the distance between spatial features and determines the spatial relationship based on distance. It is crucial for identifying the influence of one feature on another.
 - **Buffer Analysis**: A common proximity analysis technique is buffer analysis, where buffer zones (areas within a specified distance) are created around a feature to analyse their impact or influence.

- **Nearest Neighbour Analysis**: This technique evaluates the closest distances between features.
- Network Analysis: Network analysis examines the relationships and flows within networks, such as transportation, utilities, and communication systems. It helps in optimizing routes, improving connectivity, and managing resources effectively.
 - **Route Optimization:** Network analysis can find the shortest or fastest route between two points.
 - Service Area Analysis: This identifies the area covered within a certain distance from a service location.

2.6 Conclusion

Understanding spatial and non-spatial data models is fundamental to leveraging the full potential of GIS. Spatial data models provide the framework for representing geographic features, while non-spatial data models enrich these features with descriptive attributes. The integration of these models enables powerful spatial analysis, supporting a wide range of applications from urban planning to environmental management. As GIS technology continues to evolve, the effective use of both spatial and non-spatial data models will remain crucial for addressing complex geographic challenges.

Chapter 3

Coordinate Systems and Map Projections

This chapter covers the basic concepts of coordinate systems and map projections that are central to a GIS system. These concepts are used for accurate representation of the complex, curved surface of the Earth on flat maps or digital screen.

3.1 Coordinate Systems

A coordinate system is a framework that allows for the precise location of geographic features on the Earth. It provides a standardized method to define positions through coordinates, typically using a set of numbers. There are two primary types of coordinate systems used in GIS, namely, geographic coordinate systems (GCS) and projected coordinate systems (PCS).

3.1.1 Geographic Coordinate Systems

Geographic coordinate systems are used for determining the location of a feature on three-dimensional spherical earth surface. It uses latitude and longitude to specify the location. Latitude measures the angle north or south of the Equator, while longitude measures the angle east or west of the Prime Meridian (passing through Greenwich England). This system is ideal for global or large-scale maps where the curvature of the Earth needs to be taken into account. The figure below shows the different latitude parallels with respect to equator and longitudes with respect to prime meridian.



Figure 3.1: The latitude lines with respect to equator on the left and longitude lines with respect to prime meridian on earth surface on the right (https://www.britannica.com/science/latitude).

The latitude is determined by the angle between a location on earth's surface and equatorial plane (Fig. 3.1). Similarly, longitude is a measure of angle between prime meridian plane and north-south plane crossing the feature location. For instance, the approximate latitudinal extent of India is from 6.7° N to 37.1° N and longitudinal extent is from 68.1° E to 97.5° E.

The determination of latitude and longitude of an earth feature for its visualization on a geographic information system involves the mathematical modelling of the curved earth surface.

3.1.1.1 Ellipsoid

An ellipsoid is a mathematically defined surface that approximates the shape of the Earth. It is an elongated sphere, or spheroid, with two principal radii: the equatorial radius (semi-major axis) and the polar radius (semi-minor axis). The equatorial radius is slightly larger than the polar radius, making the ellipsoid an oblate shape, which tries to capture the elongated shape of the Earth around the equator. This shape is used because it provides a simpler, more uniform model for calculations compared to the Earth's actual shape. Commonly used reference ellipsoids include the WGS84 and GRS80, which are defined by their semi-major and semi-minor axes, as well as their flattening factor. The figure 3.2 shows a comparison of a sphere and an ellipsoid.



Figure 3.2: Sphere and ellipsoid (https://www.caliper.com/).

3.1.1.2 Geoid

A geoid is a model of the earth's shape that represents the mean sea level across the planet's oceans, extended through the continents. It is a more complex and irregular surface compared to the ellipsoid because it accounts for variations in earth's gravitational field caused by factors such as mountain ranges, ocean trenches, and density differences in the Earth's interior. The geoid undulates due to these gravitational anomalies, providing a more accurate reference for measuring elevations and understanding the earth's gravitational field.



Figure 3.3: The Earth's geoid (https://www.esa.int/).

The geoid (the surface of equal gravitational potential of a hypothetical ocean at rest) serves as the classical reference for all topographical features (Fig. 3.3). The above image shows the Earth's geoid. The areas having stronger gravity field are shown by redyellow color and those having weaker gravity are depicted by blue color.

3.1.1.3 Datum

A datum is a reference framework that defines the position of the ellipsoid relative to the center of the Earth and provides a basis for geographic coordinates. There are two main types of datums: horizontal datums, which specify latitude and longitude coordinates on the Earth's surface, and vertical datums, which measure elevations relative to the geoid. Horizontal datums, such as WGS84 (used in GPS) and NAD83, are crucial for mapping and navigation, ensuring consistent and accurate spatial data across different regions and applications. Vertical datums, such as the North American Vertical Datum of 1988 (NAVD88), are used for measuring and comparing elevations.

3.2 Projected Coordinate System

A projected coordinate system is a method of mapping the Earth's three-dimensional surface onto a two-dimensional plane. This system allows for the representation of geographic features in a flat, rectangular coordinate system, which facilitates easy calculation of distances, areas, and angles. A projected coordinate system uses linear units in place of angular units as are used in geographic coordinate system. Therefore, a projected coordinate system is composed of a geographic coordinate system along with a map projection (Fig. 3.4). The map projection provides a mathematical transformation function that converts geographic coordinates to planar coordinates. The figure below shows the geographic coordinates of a location converted to planar coordinates.



Figure 3.4: The coordinates of a location in GCS (left) and PCS (right) (https://naarm.org.in/).

3.2.1 Map projections

The map projections facilitate mathematical transformation of location information of features on a curved surface to a 2-D plane. The term 'projection' comes from the idea of placing a light source within a transparent globe and projecting shadows of the meridians, and parallels onto a sheet of paper placed tangent to the globe. Each

map projection has certain strengths and weaknesses in terms of the accuracy of shape, area, distance and direction. It is not possible for any projection to retain more than one of these characteristics over a large part of the Earth. It is important to identify that because of the curvature of the Earth, all map projections distort distance and directional relationships. The figure 3.5 shows the projection of point $P'(\phi, \lambda)$ on a 3D sphere to P(x, y) on 2D plane.



Figure 3.5: *Map projection from reference surface to map plane* (*https://kartoweb.itc.nl/geometrics/Map_projections/mappro.html*).

There exist many different types of map projections, each type is intended for a different application. These map projections are broadly classified as following types:

- 1. Projection surface (cylindrical, conical or azimuthal),
- 2. Point of secancy (tangent or secant),
- 3. Aspect (normal, transverse or oblique), and
- 4. Distortion property (equivalent, equidistant or conformal).
3.2.1.3 Map projections based on projection surface

Based on the projection surface, map projections are classified as, (i). Cylindrical, (ii). Conical, and (iii). Azimuthal. The projection system is depicted in the figure 3.6, where each type of projection surface is wrapped around earth in such a way that projection surface is a tangent at points of contact with reference surface.



(https://kartoweb.itc.nl/geometrics/Map_projections/mappro.html)

3.2.1.4 Map projections based on point of secancy

Another class of projections is obtained if the surfaces are chosen to be secant to (intersect with) the horizontal reference surface. In this case, the reference surface is intersected along one closed line (plane) or two closed lines (cone and cylinder). Secant map surfaces are used to reduce or average scale errors because the line(s) of intersection are not distorted on the map (Fig. 3.7).



Figure 3.7: Map projections based on point of secancy (https://kartoweb.itc.nl/geometrics/Map_projections/mappro.html)

3.2.1.5 Map projections based on aspect

This class of projection is defined based on the projection plane's orientation with respect to the globe. There are three possible aspects: normal, transverse, and oblique, depending on how the projection plane intersects Earth's axis. The projection plane is normal, parallel, and at an angle (non-parallel and non-normal) to Earth's axis in normal, transverse, and oblique projections, respectively (Fig. 3.8).



(https://kartoweb.itc.nl/geometrics/Map_projections/mappro.html).

3.2.1.6 Map projections based on distortion property

The distortion properties of a map are typically classified according to what is not distorted on the map. For instance, in a conformal map projection the angles between lines in the map are identical to the angles between the original lines on the curved reference surface. This means that angles and shapes are accurately represented on the map. The figure below shows a Mercator projection which preserves the conformal property by accurately representing local angles and shapes. However, it exhibits large area distortions. Greenland, having area 1/8th of that of South American continent, appears larger than the South America.

Similarly, in an equal-area (equivalent) map projection, the areas on the map are identical to the areas on the curved reference surface (taking into account the map scale), ensuring that areas are represented correctly on the map (Fig. 3.9). No map projection can be both conformal and equal-area. A projection can only be equidistant (true to scale) at certain places or directions.



Figure 3.9: Mercator projection exhibiting area distortions.

3.3 Commonly used map projections

3.3.1 Universal Transverse Mercator

The Universal Transverse Mercator (UTM) projection is a widely used map projection system that divides the world into a series of six-degree longitudinal zones. Each zone uses a transverse Mercator projection cantered on a meridian. This projection is conformal, meaning it preserves local angles and shapes, making it ideal for detailed and accurate mapping. This projection is used extensively for detailed topographic maps, especially for large-scale mapping.

3.3.2 Lambert Conformal Conic

The Lambert Conformal Conic (LCC) projection is a conic map projection that is widely used for aeronautical charts, regional mapping, and weather maps. This projection is conformal, which means it preserves local angles and shapes, making it highly suitable for detailed and accurate representations of regions with larger eastwest extents.

3.3.3 Albers Equal Area Projection

The Albers Equal-Area projection is a conic map projection that is designed to minimize distortion of area, making it ideal for thematic and statistical maps where accurate area representation is crucial. The Albers projection maps the Earth's surface onto a cone that intersects the globe along two standard parallels. These parallels are chosen based on the specific region being mapped and represent the lines of latitude where the cone touches the globe. This projection preserves area, ensuring that the size of features on the map is proportional to their size on the Earth. This is particularly important for thematic maps that require accurate representation of regions sizes, such as population density or land use maps.

3.4 Conclusion

Coordinate systems and map projections form essential cornerstones in geographic information systems (GIS), integral to accurately representing and analysing spatial data. The selection of a coordinate system dictates how geographic positions are defined on the Earth's surface, whether through spherical coordinates like latitude and longitude or projected coordinates on a flat map. Map projections, in turn, convert the Earth's curved surface into a two-dimensional map, balancing considerations of area, shape, distance, and direction depending on the application's requirements. Proficiency in these concepts is critical for GIS professionals, ensuring precise analyses in fields ranging from urban planning and environmental management to navigation and scientific research. The process of choosing the appropriate coordinate system and map projection involves evaluating factors such as scale, distortion characteristics, and the specific intended use of spatial data, underscoring GIS's interdisciplinary relevance across various domains today.

Chapter 4

Spatial Analysis: Vector Analysis & Raster Analysis

Spatial analysis is a crucial component of Geographic Information Systems (GIS), enabling the examination of spatial data to discern patterns, relationships, and trends. In geographic information systems (GIS), vector and raster analyses are fundamental methods for manipulating spatial data. Both techniques offer unique strengths and are applied based on the nature of the data and the requirements of the analysis. In this chapter, we will delve into two primary forms of spatial analysis: vector analysis and raster analysis. Each of these methodologies has unique characteristics, applications, and advantages that make them suitable for different types of spatial data and analysis tasks.

4.1 Vector Analysis

Vector data models represent geographic features as points, lines, and polygons. Vector data is ideal for representing discrete features with precise locations and boundaries, such as buildings, roads, or administrative boundaries. Each feature is defined by its vertices and can store various attributes. Vector analysis focuses on the relationships and interactions between these features.

Basic operations in vector data analysis are-

Buffering creates zones around vector features to analyse spatial relationships and proximity. This operation is fundamental in many GIS applications, such as environmental impact assessments or urban planning (Fig. 4.1).

• **Point Buffers:** Useful for identifying areas within a certain distance of a specific location (e.g., determining the area within 500 meters of a school).

- Line Buffers: Applied to linear features like roads or rivers to define influence zones (e.g., creating a buffer around a highway to assess noise pollution).
- **Polygon Buffers**: Surround entire areas with a buffer zone, useful in environmental studies (e.g., establishing a buffer around a protected area).



Figure 4.1: Buffer around Point, Line, Polygon features.

Overlay analysis is a powerful technique for combining multiple vector layers to investigate spatial relationships. This is crucial for integrating various datasets and drawing insights from complex spatial data. The primary overlay operations are (Fig.4.2):

- Intersection: Identifies common areas where two layers overlap. The resulting layer contains features where both input layers intersect, preserving attribute data from both. For example, finding areas where commercial zoning overlaps with high flood risk zones.
- Union: Merges all features from both input layers into a single layer, combining their geometries and attributes.
- **Difference**: Subtracts the features of one layer from another, identifying areas present in the first layer but not in the second.



Figure 4.2: A visualization of the vector overlay operations.

Spatial querying involves selecting and analyzing features based on their spatial relationships or attributes. It enables users to extract specific subsets of data for further analysis or visualization.

- Attribute Query: Selects features based on attribute values (e.g., all parks larger than 10 hectares).
- **Location Query**: Selects features based on spatial criteria (e.g., all hospitals within 2 km of a major road).

Advanced techniques in vector data analysis are explained below-

Network analysis examines connectivity, flow, and accessibility within spatial networks, such as transportation systems, utilities, or communication networks. It is essential for optimizing routes, evaluating service areas, and understanding network dynamics.

- Shortest Path Analysis: Determines the most efficient route between two points, considering distance, travel time, or other costs (Fig.4.3).
- Service Area Analysis: Identifies areas reachable within a certain distance or time from a facility, useful for emergency response planning.
- Network Flow Analysis: Analyses the flow of goods, services, or people through a network, crucial for logistics and supply chain management. E.g. calculating the shortest path between two locations in a road network.



Figure 4.3: Shortest path from source node 0 to target node 10 using network analysis.

Topological modeling examines the spatial relationships between features, such as adjacency, connectivity, and containment. Topology ensures data integrity and supports complex spatial analysis by maintaining consistent relationships between vector features (Fig. 4.4).

- Adjacency: Identifies features that share a common boundary (e.g., neighboring land parcels).
- **Connectivity**: Analyses how features are connected, crucial for network analysis (e.g., road intersections).
- **Containment**: Determines if one feature is contained within another (e.g., determining which land parcels fall within a specific zoning area).



Figure 4.4: Topological Relationships showing adjacency, connectivity, and containment.

Spatial interpolation estimates values at unsampled locations based on known values from surrounding points. It is used to create continuous surfaces or predict spatial patterns from discrete observations.

• **Inverse Distance Weighting** (IDW): Estimates values based on the inverse of the distance to known points, giving more weight to closer points. Inverse Distance weighting models work on the premise that observations further away should have their contributions diminished according to how far away they are. The simplest model involves calculating the weighted mean for all of the observations

The basic IDW interpolation formula is as below

$$x^* = \frac{w1x1 + w2x2 + w3x3 + \dots + wnxn}{w1 + w2 + w3 + \dots + wn}$$

Where x^* is unknown value at a location to be determined, w is the weight, and x is known point value. The weight is inverse distance of a point to each known point value that is used in the calculation. Simply the weight can be calculated as below

$$w1 = \frac{1}{d_{ix^*}^p}$$

p is a positive real number, called the power parameter.

A value at position x will be determined from sampling points 1, 2, and 3, with the distances to x point are d_{1x} , d_{2x} and d_{3x} . Using the equations, each respective weight will be calculated and then the value at position x will be determined.

• **Kriging:** Kriging is an advanced geostatistical technique for spatial interpolation that leverages both the distance and the degree of variation between known data points to estimate values at unknown locations. Kriging provides not just predictions but also measures of the prediction uncertainty, making it highly valuable for various applications in geology, environmental science, and resource management.

Ordinary Kriging is the most commonly used type. It predicts the value Z(x0) at an unknown location x0 based on weighted sums of the known values Z(xi).

$$Z(x_0) = \sum_{i=1}^N \lambda_i Z(x_i)$$

where:

• $Z(x_0)$ = estimated value at location x_0

- $Z(x_i)$ = estimated value at location x_i
- λ_i = weight assigned to the known value at x_i
- N= number of known sample points

4.2 Raster Analysis

Raster data represents spatial information through a matrix of cells or pixels, where each cell contains a value representing a specific attribute, such as elevation, temperature, or land cover. This format is particularly suited for modeling continuous phenomena and analyzing spatial patterns over large areas.

Basic operations in raster data analysis-

Map Algebra involves performing mathematical operations on raster datasets to create new outputs. It is a flexible tool for spatial analysis and modeling, allowing for a range of operations from simple arithmetic to complex spatial analyses (Fig. 4.5).

- Local Operations: Apply a function to each cell individually based on one or more input rasters. For example, calculating the sum of two elevation rasters.
- **Focal Operations**: Apply a function to a cell based on its neighborhood. This can be used to calculate the average elevation within a specified radius around each cell.
- **Zonal Operations**: Apply a function to all cells within each zone of a raster, such as calculating the mean temperature within different climate zones.



Figure 4.5: Map algebra operations.

Reclassification involves changing the values of raster cells based on a set of rules or criteria. It simplifies complex datasets, making them easier to analyze and interpret (Fig. 4.6).

- **Binary Reclassification**: Converts continuous or categorical data into a binary format (e.g., reclassifying land cover into forest and non-forest).
- **Categorical Reclassification**: Groups data into categories (e.g., converting elevation ranges into altitude classes).
- **Continuous Reclassification**: Adjusts continuous values into new ranges or classes (e.g., reclassifying temperatures into hot, warm, and cold).



Figure 4.6: Reclassification by range of values.

Suitability analysis evaluates multiple criteria to determine the best locations for a specific purpose, often using weighted overlay techniques.

Advanced Techniques in raster data analysis-

Surface analysis derives terrain attributes from raster data, typically digital elevation models (DEMs). It helps in understanding topography, landform characteristics, and environmental processes (Fig. 4.7).

- **Slope**: Calculates the steepness of the terrain, useful for erosion studies and infrastructure planning.
- **Aspect**: Determines the direction of slope faces, important for solar radiation and vegetation analysis.
- **Hillshade**: Creates a shaded relief map that simulates how terrain appears under specific lighting conditions, enhancing topographic visualization.



Figure 4.7: Slope, aspect and the hillshade derived from the DEM. a) Slope, b) Aspect, c) Hillshade.

Raster to vector conversion transforms raster data into vector format, enabling integration with vector datasets and facilitating certain types of spatial analyses (Fig. 4.8).

- **Contour Extraction**: Converts elevation rasters into contour lines, useful for topographic maps.
- **Polygonization**: Transforms classified raster data into vector polygons, such as converting land cover classes into vector polygons for land use mapping.
- Stream Network Extraction: Converts flow accumulation grids into vector stream networks for hydrological analysis.



Figure 4.8: Elevation plot from contour map.

4.3 Conclusion

Both vector and raster analyses are integral to spatial analysis, each offering unique capabilities and suited to different types of data and applications. By mastering spatial analysis techniques, GIS professionals can effectively analyze spatial data, uncover patterns, and inform decision-making processes in various applications. Understanding their differences and strengths allows for the effective application of GIS techniques in various fields, from urban planning to disaster management.

Chapter 5

Network Analysis

5.1 Network Analysis in GIS

Network analysis in GIS examines relationships within spatial networks, consisting of interconnected nodes (points) and edges (lines) that represent real-world systems like roads and utility lines. It aims to solve spatial problems by analysing these connections.

This analysis is crucial for optimizing routes, enhancing accessibility, managing infrastructure, and making informed decisions across various sectors. For example, it aids urban planners in designing efficient transportation systems, helps emergency responders find the quickest routes, and allows logistics companies to optimize delivery paths. Analysing spatial networks improves operational efficiency, reduces costs, and enhances service delivery.

Network analysis has diverse applications, including:

- 1. **Transportation Planning**: Optimizing public transit routes, traffic flow, and planning infrastructure.
- 2. **Emergency Response**: Identifying the fastest routes for emergency vehicles and planning evacuations.
- 3. **Logistics**: Optimizing delivery routes and managing fleet operations.
- 4. **Utilities**: Managing networks of pipelines, electrical grids, and communication lines.
- 5. **Urban Planning**: Analysing accessibility to amenities and planning pedestrian and bicycle paths.
- 6. **Environmental Management**: Studying water flow in river networks and managing natural resources.
- 7. **Healthcare**: Optimizing locations of healthcare facilities and analysing patient access.

Key data structures for network analysis in GIS include:

- Vector Data Model: Represents networks with points (nodes) and lines (edges), with attributes like length and travel time.
- **Raster Data Model:** Uses a grid of cells to represent networks, often for environmental networks.
- **Topological Data Model:** Focuses on relationships and connectivity between network elements, maintaining spatial relationships.

These structures are essential for building and analyzing networks in GIS, enabling accurate analysis and optimization of spatial networks.

5.2 Fundamentals of Network Theory

Graph theory forms the mathematical basis of network analysis, focusing on graphs composed of nodes (or vertices) and edges (or links). A graph is a collection of nodes connected by edges, representing relationships between objects.

Key concepts in graph theory include:

- **Graph**: A structure made up of nodes connected by edges.
- Vertex (Node): The basic unit in a graph, representing locations in the network (e.g., intersections, cities).
- Edge (Link): The connection between nodes, representing relationships or pathways (e.g., roads, communication lines).
- **Degree**: The number of edges linked to a node; in directed graphs, this includes in-degree (incoming edges) and out-degree (outgoing edges).
- **Path**: A sequence of edges that connects a series of nodes.
- **Cycle**: A path that begins and ends at the same node without repeating any edges.

Types of Networks: Directed and Undirected

Networks can be classified based on whether their edges have a direction:

Undirected Networks:

- Edges do not have a direction; the relationship between nodes is bidirectional.
- Typically used to represent mutual relationships, like friendships in social networks.

Directed Networks (Digraphs):

- Edges have a direction, indicating a one-way relationship.
- Used for systems where direction matters, such as road networks with one-way streets or citation networks in academic research.

Nodes, Edges, and Attributes

In network theory, nodes and edges can have various attributes that provide additional information:

- Nodes (Vertices):
 - Attributes: Characteristics such as name, location, type, or other relevant information.
 - Example: In a transportation network, nodes could represent intersections, with attributes like coordinates and traffic signal timings.

• Edges (Links):

- Attributes: Properties such as length, travel time, capacity, or cost.
- Example: In a road network, edges could represent streets, with attributes like distance, speed limit, and number of lanes.

Understanding these fundamental components and their attributes is essential for analyzing and interpreting networks effectively. By applying graph theory, user can explore various network properties, optimize routes, and solve complex spatial problems in GIS.

5.3 Data Preparation for Network Analysis

Effective network analysis in GIS begins with gathering high-quality spatial data from reliable sources. Government agencies offer comprehensive datasets, such as those for road networks, utility lines, and public transit systems. Commercial providers supply proprietary datasets with more detailed or specialized information. Open data platforms like OpenStreetMap provide accessible geospatial data for free. Remote sensing techniques, including satellite imagery and aerial photography, are also valuable sources, while field surveys provide precise and up-to-date data through direct collection.

After data collection, cleaning and preprocessing are crucial to resolve any errors or inconsistencies. This includes correcting errors by identifying and fixing issues like missing values, duplicates, and inaccuracies. It is important to standardize all data to the same coordinate system and projection for consistency. Data integration involves combining datasets from different sources to ensure spatial and attribute alignment. Finally, building topology ensures that the network's connectivity is accurately represented, verifying that roads and intersections are properly depicted. Proper data preparation is essential for conducting reliable and accurate network analysis in GIS.

5.4 Network Data Models in GIS

In GIS, network data models are essential for representing spatial networks effectively. The **vector data model** utilizes points, lines, and polygons to depict nodes, linear features (such as roads and utility lines), and areas, respectively. Each feature can have

associated attributes like road length or capacity, making it ideal for detailed representations in transportation, utility, and communication networks.

Conversely, the **raster data model** employs a grid structure where each cell holds specific values, suitable for environmental networks like hydrological modelling. It excels in handling continuous data over large areas but may lack precision for intricate linear features compared to the vector model.

Topological data models emphasize spatial relationships and connectivity between network elements, ensuring accurate representation through defined nodes, edges, and rules governing their interaction. These models are pivotal for tasks such as route optimization and network tracing in fields like transportation planning and utility management, where precise connectivity is crucial. By maintaining proper connectivity and relationships, these models enable robust analysis and decision-making in GIS applications.

5.5 Types of Network Analysis

5.5.1 Shortest Path Analysis

Shortest path analysis calculates the most efficient route between two points within a network, minimizing distance or travel time based on predefined criteria. For example, in urban planning, determining the shortest route from residential areas to hospitals can significantly enhance emergency response times.

5.5.2 Service area analysis

Service area analysis identifies accessible areas from a specific location within a defined travel time or distance. It aids in delineating coverage areas for facilities like fire stations or distribution centers. For instance, analyzing a 15-minute service area around a retail store can evaluate its market reach and accessibility.

5.5.3 Closest Facility Analysis

Closest facility analysis determines the nearest facility from specified locations, crucial for tasks such as locating the nearest fire station to residential areas or identifying the closest warehouse for efficient logistics operations.

5.5.4 Origin-Destination (OD) Cost Matrix

The OD cost matrix computes travel costs or distances between multiple origin-destination pairs within a network. It offers insights into network connectivity and is useful for analyzing traffic flow patterns and commuter behaviours. For instance, analyzing an OD cost matrix can reveal commuting patterns between residential areas and job centers.

5.5.5 Route Optimization

Route optimization identifies the optimal sequence of stops or waypoints to minimize travel time, distance, or cost. It is essential for logistics companies optimizing delivery routes or public transit systems refining bus schedules. For example, optimizing delivery routes can reduce operational costs and improve service efficiency.

These forms of network analysis play a pivotal role in GIS applications across diverse sectors, providing critical insights that optimize resource allocation, enhance service delivery, and improve overall operational efficiency.

5.6 Algorithms and Techniques for Network Analysis

5.6.1 Dijkstra's Algorithm

Dijkstra's algorithm stands as a fundamental method for determining the shortest path between nodes within a graph characterized by nonnegative edge weights. It operates by iteratively selecting the node with the smallest known distance and updating distances to neighbouring nodes accordingly. This approach finds extensive use in applications necessitating shortest path computations, including GPS navigation systems and network routing protocols.

Let's consider a simplified example where we have a small road network in a GIS:

- Nodes: A, B, C, D
- Edges with weights (distances): A-B (5), A-C (10), B-C (3), B-D (9), C-D (2)

Suppose we want to find the shortest path from node A to node D using Dijkstra's algorithm:

1. **Initialization**: Start at node A with a distance of 0 and set all other nodes' distances to infinity. Place node A in the priority queue.

2. Iteration Steps:

- Visit node A, then explore its neighbors B and C.
- Calculate tentative distances:
 - From A to B: 5 (current distance from A)
 + 5 (distance from A to B) = 10
 - From A to C: 5 (current distance from A)
 + 10 (distance from A to C) = 15
- Update distances: Set B's distance to 10 and C's distance to 15. Place B and C in the priority queue.
- 3. **Next Node**: Select node B (shortest distance in the queue). Explore its neighbor C.
 - Calculate tentative distance from A to C through B: 10 (current distance to B) + 3 (distance from B to C) = 13
 - \circ Since 13 is less than 15, update C's distance to 13.

Continue: Continue selecting the next node with the shortest distance until all nodes are visited. In this example, after processing all nodes, the shortest path from A to D would be A -> B -> D with a total distance of 19.

In GIS, Dijkstra's algorithm is extensively used for tasks such as:

- Finding the shortest route for emergency vehicles or service vehicles in transportation planning.
- Determining optimal paths for logistics and supply chain management.
- Analyzing accessibility to amenities or services in urban planning.

5.6.2 A* Algorithm

The A* algorithm integrates components from Dijkstra's approach with heuristic strategies to effectively ascertain the shortest path in graph structures. Employing a heuristic function that estimates the cost from the current node to the goal, A* directs its search toward the most promising paths. This algorithm proves particularly advantageous in scenarios where accurate heuristic estimates are available, such as in video game pathfinding or logistical route planning.

Consider a scenario with the following network details:

- Nodes: A, B, C, D
- Edges with weights: A-B (5), A-C (10), B-C (3), B-D (9), C-D (2)
- **Heuristic (h-value)**: Euclidean distance from each nodeto-node D (destination).

Let's assume Node A serves as the starting point and Node D as the destination:

- **Initialization**: Begin at Node A with an initial heuristic estimate (h-value) towards Node D.
- **Priority Queue**: Start the algorithm from Node A, calculating f-values based on g-values (actual costs) and heuristic estimates.
- Algorithm Execution: Expand nodes in order of priority until reaching Node D.
- **Path Reconstruction**: Upon reaching Node D, reconstruct the shortest path using stored parent pointers.

In this example, the A* algorithm efficiently determines the shortest path from Node A to Node D by utilizing both actual costs (g-values) and heuristic estimates (h-values). This approach ensures optimized route planning in GIS applications, such as navigation systems and logistics optimization, by guiding the search towards the most promising paths based on heuristic predictions.

5.6.3 Floyd-Warshall Algorithm

The Floyd-Warshall algorithm serves to compute the shortest paths between all pairs of nodes within a weighted graph. By systematically considering all potential intermediate nodes, this algorithm determines optimal paths. It finds applicability in scenarios involving dense graphs or instances requiring comprehensive all-pairs shortest path analyses, such as network connectivity assessments and traffic management.

In GIS, the Floyd-Warshall algorithm is applied in various scenarios, including:

- Calculating shortest paths in transportation networks.
- Analyzing connectivity and accessibility between locations.
- Assessing network robustness and reliability in infrastructure planning.

5.6.4 Heuristics and Metaheuristics

Heuristics and metaheuristics represent problem-solving methodologies utilized to approximate solutions in complex optimization challenges where precise algorithms may be impracticable. Heuristics rely on intuitive strategies or rules of thumb to expediently identify solutions, whereas metaheuristics employ advanced strategies like simulated annealing or genetic algorithms to effectively explore and exploit search spaces. These techniques find application across diverse network analysis tasks, encompassing vehicle routing problems, facility location decisions, and the optimization of network designs.

These algorithms and methodologies serve as integral components of network analysis in GIS, facilitating efficient pathfinding, comprehensive connectivity evaluations, and the optimization of spatial networks across a broad spectrum of applications.

5.7 GIS Software and Tools for Network Analysis

Network Analyst Extension in ArcGIS

ArcGIS features a specialized toolset known as Network Analyst, tailored for advanced spatial analysis based on network structures. This extension offers several key functionalities:

- 1. **Routing Analysis**: Users can determine optimal routes between locations, considering criteria like shortest path, fastest route, or avoidance of specific obstacles.
- 2. Service Area Analysis: This tool enables the delineation of service areas around locations based on travel time or distance, crucial for assessing accessibility and service coverage.
- Location-Allocation Analysis: Facilitates strategic placement of facilities to meet demand, accounting for factors such as demand distribution and service capacities.

4. **Network Dataset Management**: Provides capabilities for creating and managing network datasets, defining connectivity rules, attributing network elements, and ensuring data accuracy.

Plugins and Extensions in QGIS

QGIS extends its core functionality through various plugins and extensions, enhancing capabilities particularly relevant to network analysis:

- 1. **QGIS Network Analysis Library (QNEAT3)**: This plugin augments QGIS with tools for tasks such as shortest path analysis, service area computations, and OD matrix generation.
- 2. **PgRouting**: Integrates with QGIS to offer advanced routing and network analysis using PostgreSQL and PostGIS databases.
- 3. **GRASS GIS Integration**: QGIS seamlessly integrates with GRASS GIS, providing additional tools for spatial analysis and modelling, including tasks related to network analysis.
- 4. **Processing Toolbox**: QGIS's Processing Toolbox includes algorithms dedicated to network analysis, such as shortest path calculations and other specialized functions, customizable to specific project needs.

These tools empower GIS professionals to conduct comprehensive network analysis across various domains, from urban planning and transportation logistics to environmental management and infrastructure development.

5.8 Conclusion

Network analysis in GIS is vital for understanding spatial relationships and optimizing resource use across sectors. Using

graph theory and algorithms, GIS professionals' model complex networks of nodes and edges. Vector, raster, and topological data models enable detailed representations of real-world networks like transportation and utilities. Effective data preparation, including collection, cleaning, and dataset creation, ensures accurate analyses. Various analyses like shortest path calculations and route optimization meet specific spatial planning needs. Algorithms such as Dijkstra's, A*, and Floyd-Warshall enhance capabilities for pathfinding and connectivity. With ArcGIS and QGIS software and their specialized tools, professionals conduct precise spatial analyses, improving efficiency and resource management.

Chapter 6

Map Composition

Map composition is a process of preparing maps with incorporation of thematic layers and their attributes, showing all map elements. The elements are shown using conventional signs and symbols for easy interpretation of the maps. This chapter explores the essential components of map composition, guiding principles, and best practices to ensure that maps are both functional and aesthetically pleasing.

6.1 Basic Map Elements

There are some basic elements of map design that aid in map interpretation. These map elements are title, scale, reference grid, legend, projection, north arrow etc. The figure 6.1 shows some of the important map elements.



Figure 6.1: Common map elements

(https://docs.qgis.org/3.34/en/docs/gentle_gis_introduction/map_producti on.html)

6.1.1 Map Title

A title of the map introduces the user to the map content and its specific details, such as the study area, year of study, or variable under consideration. A map title should be concise, informative, and reflective of the map's content.

6.1.2 North Arrow

A north arrow shows the map's orientation, helping users understand directionality. It is especially important in navigation maps and should be easily visible without being intrusive.

6.1.3 Scale

Scale is the ratio between the distance between two points on a map and the actual distance on the ground. For instance, 1 cm of a map drawn at 1:50000 scale represents 50,000 cm (0.5 km) distance on the ground. A ratio of distance on map to actual distance on the ground is also known as representative fraction. A map with high a representative fraction is called as large-scale map. Conversely, a map with a very small representative fraction is called a small-scale map.

A scale can be represented in various ways, including a scale bar, verbal statement, or numeric scale. Choosing the right scale is essential for conveying the appropriate level of detail. The use of different types of scale is given below.

Statement Scale: 1 cm = 2 km

(Interpretation: 1 cm of map is equal to 2 km on the ground)

Numeric Scale: 1:200,000

(Interpretation: 1 unit on the map is equal to 200,000 units on the ground)

Graphical Scale:

0 7.5 15 km

(Interpretation: Each division of graphical scale represents 7.5 km distance on the ground)

The scale of a map determines the information it carries about the features. For instance, a large-scale map shows a small area of earth in great detail and small-scale map shows a larger area in less detail (Figure 6.2). For example, the figure below shows three maps at 1:100,000 (small scale), 1:50,000 (medium scale), and 1:25,000 (large scale). The small-scale map shows a larger area compared to medium and large-scale maps, but level of detail about blue feature is more in the large-scale map.



1:100,000 1:50,000 1:25,000 Figure 6.2: A representation of feature details at three different scales.

6.2 Projection

A map projection determines how three-dimensional information is translated to a two-dimensional plane. Including information about the projection used is vital for understanding the spatial relationships and potential distortions present (Figure 6.3).



Figure 6.3: World in different projections (Mollweide Equal Area projection on the left and Plate Carree Equidistant Cylindrical projection on the right).

6.3 Reference Grid

The inclusion of a reference grid (or graticule) on a map helps in identifying the spatial location of different features. A grid of latitude and longitude or alpha-numeric codes is used as a reference grid that subdivides the map into different regions for easy interpretation (Figure 6.4).



Figure 6.4: Reference grids dividing the entire image area into 9 subregions (Source: https://desktop.arcgis.com/en/arcmap/latest/map/pagelayouts/what-are-grids-and-graticules-.htm).

6.4 Map Legend

A map is a simplified representation of the real-world features. A map uses symbols to represent real-world objects. The map legend explains the symbols, colors, and patterns used on the map. It is essential for interpreting the data accurately (Figure 6.5). Legends should be simple, uncluttered, and positioned where they can be easily found without obstructing the map's main features.



Figure 6.5: Crop map of Narayanpur command area with legend shown at bottom of image (Source: https://bhuvan.nrsc.gov.in/nhp/webgis-irrigation/map).

Different symbols are used to represent natural and man-made features, including point, linear, and polygon shape features. These are also known as primary visual variables. The size of these shapes and color intensity are often used to depict auxiliary properties of the variables, which are called secondary visual variables.

6.5 Conclusion

Map composition involves preparing maps with thematic layers and attributes, using conventional signs and symbols for easy interpretation. Essential components and best practices ensure maps are functional and aesthetically pleasing. Basic map elements include title, north arrow, scale, projection, reference grid, and legend. Each element aids in map interpretation, orientation, and understanding of spatial relationships. Scale represents the ratio of map distance to ground distance and can be displayed in various formats. Map projections convert 3D information to 2D which is crucial for accurate spatial representation. Map legend explains the symbols, colors, and patterns on the map, essential for accurate data interpretation. Legends should be simple and unobtrusive, aiding in the easy understanding of map features.

Chapter 7

FOSS4G - Open-Source Tools & Techniques

7.1 Introduction

Geospatial technology has advanced significantly, driven by the necessity for sophisticated tools to manage complex spatial data and analysis. A key development in this field is the Free and Open-Source Software for Geospatial (FOSS4G) movement, which has played a crucial role in making powerful geospatial tools accessible to all and fostered a collaborative community. This chapter delves into FOSS4G, examining its core principles, the variety of tools it offers, and the innovative techniques it supports.

FOSS4G (https://foss4g.org/) includes a wide range of freely available and open-source geospatial software, which users can access, modify, and distribute without financial or licensing restrictions. This open model not only lowers costs but also enhances transparency and fosters innovation. Users can examine the source code, tailor functionalities to their needs, and contribute to the software's development. The collaborative spirit of FOSS4G communities accelerates the evolution of geospatial tools, ensuring they stay cutting-edge and meet the diverse requirements of users worldwide.

This chapter provides an overview of essential FOSS4G tools such as QGIS, GRASS GIS, GeoServer, PostGIS, GDAL, and Leaflet, detailing their features, capabilities, and applications. Each tool addresses different aspects of geospatial analysis, from data integration and management to spatial modelling and web mapping. By understanding these tools, users can conduct comprehensive spatial analyses, create interactive maps, and manage geospatial data effectively and efficiently. The chapter also explores various techniques made possible by FOSS4G tools, including data integration, spatial analysis, web mapping, and automation. These techniques enable users to tackle complex spatial challenges, streamline workflows, and extract meaningful insights from geospatial data. Through case studies and practical examples, the chapter demonstrates how FOSS4G tools and techniques are utilized in real-world scenarios, highlighting their impact in fields like urban planning, environmental management, transportation, and public health.

With the growing demand for geospatial solutions, the importance of FOSS4G continues to increase. By adopting open-source geospatial tools and techniques, users can fully exploit spatial data, contribute to the advancement of geospatial technology, and drive innovation in their respective domains. This chapter aims to provide readers with the knowledge and skills needed to effectively use FOSS4G, fostering a deeper understanding of its benefits and applications in the geospatial landscape.

7.2 Key components & benefits of FOSS4G

FOSS4G features several key components that have led to its widespread adoption and effectiveness. One of its major advantages is its accessibility and cost-effectiveness; FOSS4G tools are free, making them accessible to individuals, organizations, and governments regardless of their budget. Additionally, the vibrant global community offers extensive documentation, forums, and development resources, making these tools easy to use and troubleshoot. Key benefits of FOSS4G are:

1. **Transparency and customizability** are central to FOSS4G's appeal. The availability of source code allows users to view and modify the software, tailoring it to specific needs and enhancing transparency. Organizations

can also develop custom plugins or extensions, improving functionality without dependence on proprietary vendors.

- 2. Interoperability and standards compliance are critical elements, with many FOSS4G tools adhering to international standards set by organizations like the Open Geospatial Consortium (OGC). This compliance ensures compatibility with other software and data formats, allowing seamless integration with other systems for efficient data exchange and workflow integration.
- 3. FOSS4G provides a **diverse toolset** that caters to various aspects of geospatial analysis. Tools like QGIS for desktop GIS, GRASS GIS for advanced spatial modeling, GeoServer for web mapping, PostGIS for spatial databases, GDAL for data processing, and Leaflet for interactive web maps cover all stages of geospatial data handling, from acquisition and management to analysis and visualization.
- 4. The benefits of FOSS4G are extensive. One of the primary advantages is **cost savings**, as FOSS4G eliminates the need for expensive licensing fees, significantly reducing the cost of acquiring and maintaining geospatial software. Its open-source nature also fosters collaborative development, distributing costs and efforts across a wider community.
- 5. FOSS4G promotes innovation and collaboration through its global community of developers and users, who contribute to continuous improvement and ensure that the tools remain at the forefront of geospatial technology. The open-source model allows for rapid iteration and the inclusion of new features based on user feedback and emerging needs.
- 6. **Flexibility and control** are additional benefits. Users can modify and extend the software to meet their specific requirements, offering greater control over the functionality and performance of their geospatial tools.
This independence from a single vendor reduces the risk of obsolescence and ensures long-term access and usability.

- 7. **Community support** is crucial, with an active user and developer base providing a wealth of knowledge, resources, and assistance, simplifying the learning and troubleshooting processes. Open forums, user groups, and conferences facilitate the exchange of ideas and best practices, enhancing overall expertise within the community.
- 8. Enhanced data sharing and integration are made possible through compliance with international standards, ensuring that FOSS4G tools can easily integrate with other systems and promote data sharing and collaborative projects. The widespread use of FOSS4G tools across various industries and sectors encourages the development of interoperable solutions and datasets.
- 9. **Sustainability and longevity** are supported by the opensource model, which fosters sustainable software development practices. The community can continue to develop and maintain the software independently of commercial interests. The availability of source code ensures that FOSS4G tools can be maintained and updated over the long term, even if the original developers move on.

7.3 Prominent Open-Source Geospatial Tools

7.3.1 QGIS (Quantum GIS)

QGIS is a user-friendly, open-source desktop GIS application designed for viewing, editing, and analyzing geospatial data (Figure. 7.1). It boasts advanced analysis capabilities, plugin support, and compatibility with various data formats. QGIS is widely used in urban planning, environmental management, resource mapping, and other fields. QGIS is designed with an intuitive, user-friendly interface, making it accessible to both beginners and advanced users. It is compatible with multiple operating systems, including Windows, macOS, Linux, and Android. QGIS supports a wide range of raster and vector data formats, such as shapefiles, GeoTIFF, PostGIS layers, and more. The software provides powerful tools for spatial analysis, geoprocessing, and advanced cartography, enabling users to perform buffer analysis, overlay analysis, network analysis, and additional tasks. QGIS features a robust plugin architecture, allowing users to extend its functionality through the QGIS Plugin Repository, which hosts a variety of plugins for diverse geospatial tasks. It offers comprehensive tools for managing and manipulating geospatial data, including attribute table management, data import/export, and database integration. Advanced cartographic tools are included for creating high-quality maps, with options to design and print maps featuring customized layouts, symbols, and labels.



Figure 7.1: QGIS homepage (https://www.qgis.org/en/site/).

The software includes georeferencing tools essential for integrating scanned maps and aerial imagery into GIS projects. QGIS provides extensive visualization options, including thematic mapping, 3D visualization, and temporal data animation. As open-source software released under the GNU General Public License, QGIS is free to use, modify, and distribute.

A strong, active community of users and developers offers extensive support through forums, mailing lists, and online resources. QGIS can also be integrated with other geospatial software and tools, such as GRASS GIS, GDAL, and SAGA GIS, enhancing its functionality. These features make QGIS a versatile and powerful tool suitable for a wide range of geospatial applications, from simple mapping to complex spatial analysis and geospatial data management.

7.3.2 GRASS GIS (Geographic Resources Analysis Support System)

GRASS GIS, acronym for Geographic Resources Analysis Support System, stands out as a robust open-source GIS platform celebrated for its extensive analytical capabilities spanning raster, vector, and geospatial data (Figure. 7.2). Equipped with a comprehensive suite of tools encompassing spatial modeling, geostatistics, and image processing, it emerges as a pivotal resource for in-depth geospatial analysis. Offering support for advanced functionalities like 3D visualization, topological vector analysis, and temporal data processing, GRASS GIS proves indispensable for sectors ranging from environmental management and urban planning to scientific research. Its modular architecture empowers users to execute intricate operations seamlessly via an intuitive graphical user interface or command line, thereby enabling a heightened level of customization and automation. With its capacity to handle substantial datasets and execute complex geospatial computations, GRASS GIS emerges as a potent asset for addressing a diverse array of spatial challenges and endeavours.



Figure 7.2: GRASS GIS (https://grass.osgeo.org/).

7.3.3 GeoServer

GeoServer is a robust open-source server application designed to share geospatial data and services seamlessly across the web (Figure 7.3). It provides a diverse range of functionalities and services that support efficient data dissemination, spatial analysis, and web mapping.

Functionality: Main functionalities of GeoServer are:

- 1. **OGC Standards Compliance**: GeoServer conforms to Open Geospatial Consortium (OGC) standards like Web Map Service (WMS), Web Feature Service (WFS), and Web Coverage Service (WCS), ensuring compatibility with other geospatial systems.
- 2. **Data Publishing**: It simplifies the publication of geospatial data in various formats, including raster and vector data such as Shapefile, GeoTIFF, and PostGIS, facilitating smooth integration with GIS software.
- 3. **Data Styling and Symbology**: Users can customize the appearance of geospatial data layers within GeoServer, allowing the creation of visually appealing maps and

cartographic outputs by adjusting colors, labels, and symbology.

- 4. **Raster and Vector Data Processing**: GeoServer provides tools for processing raster and vector data, enabling realtime data transformation, reprojection, and spatial analysis to ensure geospatial data can be served in desired formats and projections.
- 5. Security and Access Control: Robust security features are available in GeoServer, allowing administrators to manage access to data and services based on user roles and permissions, ensuring data integrity and confidentiality in multi-user environments.



Figure 7.3: GeoServer homepage (https://geoserver.org/).

Services: Important services provided by GeoServer are-

1. **Web Mapping**: As a foundation for web mapping applications, it empowers users to visualize geospatial data through interactive maps embedded in web pages, facilitating the creation of dynamic, responsive maps accessible from any web-enabled device.

- 2. **Data Dissemination**: GeoServer streamlines the dissemination of geospatial data to a broad audience by offering standardized web services for data access, enabling users to access and download geospatial data in various formats for offline analysis or integration into GIS workflows.
- 3. **Spatial Data Infrastructure (SDI) Development**: Serving as a cornerstone for Spatial Data Infrastructures (SDIs), GeoServer supports data sharing and interoperability, enabling organizations to publish and share geospatial data internally and externally, fostering collaboration and informed decision-making.
- 4. **Geospatial Analysis**: With support for OGC standards and robust data processing capabilities, GeoServer facilitates geospatial analysis workflows by providing access to geospatial data layers and services, allowing users to perform tasks like buffering, overlay analysis, and spatial querying directly through web-based interfaces.

7.3.4 PostGIS

PostGIS is a powerful extension for the PostgreSQL database, designed to enable the storage, management, and analysis of spatial data. By enhancing PostgreSQL's capabilities to include geographic objects, PostGIS allows for the direct storage of spatial features such as points, lines, and polygons within the database. This is supported by a range of spatial data types, including GEOMETRY and GEOGRAPHY, accommodating different coordinate systems and spatial reference models. A key feature of PostGIS (Figure 7.4) is its ability to perform complex spatial queries using standard SQL, facilitating operations such as finding intersections, calculating distances, and querying spatial relationships. Furthermore, the use of R-tree and GiST (Generalized Search Tree) indexing ensures efficient spatial querying, delivering rapid performance even with large datasets.

In addition to basic spatial data storage and querying, PostGIS provides a comprehensive suite of geometric functions that enable users to conduct detailed spatial analyses, such as buffering, union, and convex hull calculations. It also supports raster data management, essential for handling satellite imagery and digital elevation models (DEMs). Adherence to Open Geospatial Consortium (OGC) standards ensures seamless integration with other geospatial tools and applications, including OGIS, GeoServer, and MapServer. This interoperability, combined with PostgreSQL's scalability and robustness, makes PostGIS a formidable tool for managing large-scale spatial databases. The extension also includes topological functions for maintaining and analyzing spatial relationships, support for various spatial reference systems, and parallel processing capabilities, all of which enhance its performance and usability for complex geospatial data management and analysis tasks.



Figure 7.4: PostGIS homepage (https://postgis.net/).

PostGIS is an extension to the PostgreSQL database that adds support for geographic objects. It features spatial queries, spatial indexing, and geospatial data storage and management. PostGIS is used in spatial databases, geographic information systems, and web mapping applications.

7.3.5 PostgreSQL

PostgreSOL, often referred to as Postgres, is a sophisticated opensource relational database management system (RDBMS). Known for its robustness, high performance, and adherence to SQL standards, PostgreSOL (https://www.postgresql.org/) is utilized across a wide range of applications, from web services to data warehousing. It guarantees reliable transactions through ACID compliance, ensuring Atomicity, Consistency, Isolation, and Durability. PostgreSQL supports a diverse array of data types, including JSON, XML, hstore, and arrays, and offers powerful fulltext search capabilities for advanced querying. The system allows users to create custom functions, operators, and data types, and employs Multi-Version Concurrency Control (MVCC) to manage multiple transactions simultaneously without conflict. It also features streaming replication and robust backup solutions for high availability and data integrity, and with extensions like PostGIS, it can handle and query spatial data.



Figure 7.5: PGADMIN of PostgreSQL.

A major advantage of PostgreSQL is its open-source nature, making it free to use, modify, and distribute, thus providing cost savings and flexibility. It boasts a proven track record of stability and data integrity, efficient query optimization, and indexing for high performance, and SQL standards compliance for compatibility and ease of use. PostgreSOL's extensibility is evident with its wide range of plugins and extensions (Figure 7.5), supported by a large, active community offering extensive documentation and support. It operates on various systems, including Windows, macOS, and Linux. However, PostgreSQL can be complex to configure and optimize for peak performance, requiring advanced database administration skills, and it has a steeper learning curve compared to some other RDBMS, especially for new users. High-performance applications may demand substantial hardware resources, and while community support is robust, commercial support options are more limited compared to proprietary databases like Oracle or Microsoft SQL Server. Additionally, some third-party tools and applications might offer limited support or integration relative to more widely adopted commercial databases. In summary, PostgreSQL is a powerful and versatile RDBMS with extensive functionality and numerous benefits, particularly in terms of cost, performance, and extensibility, though it requires significant resources and expertise for effective management in high-performance environments.

7.3.6 GDAL

The Geospatial Data Abstraction Library (GDAL) is an open-source library designed for managing geospatial data formats, offering a broad range of functionalities and advantages for users working with geographic information systems (GIS). GDAL supports numerous raster and vector data formats, including GeoTIFF, JPEG, PNG, shapefiles, KML, and PostGIS, among others. This extensive format support allows for seamless conversion between different geospatial data formats, promoting interoperability and data sharing across various GIS applications. GDAL (https://gdal.org/index.html) provides robust tools for geospatial data processing, including image reprojection, resampling (like nearest neighbour resampling), mosaicking, subsetting, and layer translation. It also supports advanced spatial operations such as raster algebra, contour generation, and vector data rasterization. The library includes a suite of command-line tools for data inspection, transformation, and processing, which are highly useful for batch processing and automation. Additionally, GDAL offers APIs for various programming languages, including C++, Python, and Java, allowing developers to integrate geospatial data handling capabilities into custom applications. As an open-source library, GDAL is freely available, reducing costs associated with proprietary geospatial data handling solutions. Its extensive format support and data translation capabilities ensure compatibility with almost any geospatial data source, promoting interoperability across different systems and applications. GDAL benefits from a large, active user and developer community, providing extensive documentation, forums, and usercontributed code examples. Users can tailor GDAL to meet specific needs through its programmable APIs, making it adaptable for a wide range of geospatial tasks. Designed for efficient data processing, GDAL can handle large datasets and complex spatial operations, making it suitable for both desktop and server environments. It integrates well with other open-source GIS software such as QGIS, GRASS GIS, and MapServer, enhancing its utility within a broader geospatial data ecosystem. In summary, GDAL is a powerful, versatile tool essential for geospatial data management and processing, offering extensive functionalities and significant advantages, particularly in terms of format support, interoperability, and cost-effectiveness.

7.3.7 Openlayers

OpenLayers is an open-source JavaScript library widely used for displaying maps and building interactive web mapping applications. This versatile tool offers numerous advantages, making it a popular choice among developers and organizations. Being open source and freely available under a permissive license, OpenLayers is costeffective. It supports a wide range of map sources, including OpenStreetMap, Google Maps, Bing Maps, and custom tile servers, which provides flexibility in map presentation. Its rich feature set includes extensive tools for creating interactive maps, such as layer management, vector drawing, and spatial analysis. The library is highly customizable and extensible due to its modular architecture and comprehensive API, enabling developers to tailor it to specific requirements. Additionally, OpenLayers (Figure 7.6) works seamlessly across major web browsers, ensuring a consistent user experience. The active community of developers continuously contributes to its improvement, and there is extensive documentation and tutorials available to support users.



Figure 7.6: Openlayers home page (https://openlayers.org/).

However, OpenLayers also has its limitations. Rendering large datasets or complex maps can lead to performance issues, particularly in resource-constrained environments. The extensive functionality and flexibility of OpenLayers come with a steep learning curve, making it challenging for beginners to master quickly. While it is usable on mobile devices, it may not perform as well or be as responsive as some mobile-focused mapping libraries. Managing dependencies and ensuring compatibility with other frameworks can sometimes be challenging, as with many JavaScript libraries. Additionally, OpenLayers has limited native support for 3D visualization compared to some other mapping libraries like CesiumJS, which may be a drawback for applications requiring 3D maps. In summary, OpenLayers is a powerful tool for building interactive web maps, offering significant advantages in terms of being open-source, highly customizable, and compatible with various map sources. Nonetheless, it also presents challenges such as performance concerns, a steep learning curve, and less optimization for mobile and 3D applications.

7.4 Conclusion

The chapter on "FOSS4G - Open-Source Tools & Techniques" provides an in-depth examination of Free and Open-Source Software for Geospatial (FOSS4G), elucidating its core principles, functionalities, and advantages. It elucidates how FOSS4G plays a pivotal role in widening access to potent geospatial tools, fostering collaboration, and stimulating innovation within the geospatial community. Central to the chapter are the key components of FOSS4G, including QGIS, GRASS GIS, GeoServer, PostGIS, GDAL, and Leaflet, each highlighted for their unique features, capabilities, and applications across various sectors like urban planning and environmental management. Moreover, the chapter explores the diverse techniques enabled by FOSS4G, encompassing data integration, spatial analysis, web mapping, and automation, effectively demonstrating their practical utility through real-world case studies and examples. Ultimately, the chapter underscores the indispensable contribution of FOSS4G in propelling advancements in geospatial technology, empowering users, and facilitating solutions to intricate spatial challenges spanning multiple domains.

Chapter 8

Global Navigation Satellite System (GNSS)

8.1 Introduction

Global Navigation Satellite Systems (GNSS) are a collection of satellite constellations that provide global coverage for positioning, navigation, and timing (PNT) services. These systems transmit signals from space, which GNSS receivers use to determine their exact location (latitude, longitude, and altitude) anywhere on Earth, at any time, and in any weather conditions. GNSS encompasses several satellite systems operated by different countries or entities, each contributing to the global PNT infrastructure.

GNSS technology has a multitude of uses that span across various industries and everyday life. In transportation, GNSS is essential for navigation, providing real-time directions and traffic updates for drivers, as well as guiding ships and aircraft with precision. In emergency response and disaster management, GNSS enables efficient search and rescue operations, helps coordinate relief efforts, and tracks the movement of emergency vehicles. The agriculture sector utilizes GNSS for precision farming. GNSS plays a critical role in geospatial and surveying activities, offering accurate location data for mapping, construction, and land surveying.

8.2 Principle behind GNSS

Global Navigation Satellite Systems (GNSS) work on the principle of trilateration (Figure 8.1), which involves determining a position based on the distances from multiple reference points. In the context of GNSS, these reference points are satellites in space.



Figure 8.1: Shows trilateration from GNSS satellites.

GNSS relies on a constellation of satellites orbiting the Earth (Figure 8.1). Each satellite continuously transmits signals that include its location and the precise time the signal was sent. A GNSS receiver on the ground (or in the air or at sea) picks up these signals. By calculating the time, it takes for the signals to travel from the satellite to the receiver, the distance from each satellite to the receiver can be determined. This is known as the time-of-flight or time-delay measurement. The distance D from a satellite to the receiver is calculated using the formula:

 $D = c \times t$

where c is the speed of light (approximately 299,792 kilometres per second) and t is the time delay.

To determine its exact position, a receiver needs signals from at least four satellites. Three satellites are required for calculating the receiver's position in three-dimensional space (latitude, longitude, and altitude). A fourth satellite is needed to correct the receiver's clock error. GNSS receivers do not have atomic clocks like the satellites; thus, they need this additional measurement to precisely synchronize their internal clocks with the satellite clocks.

8.3 GNSS Bands

GNSS satellites transmit signals on specific frequency bands that are divided into multiple channels. Each band serves a unique purpose and offers different advantages. Each band plays a crucial role in ensuring the accuracy, reliability, and efficiency of GNSS services. The frequency bands used by GNSS systems include L1, L2, L5, and other higher frequencies. These bands are carefully selected to balance factors such as signal strength, atmospheric interference, and the need for interoperability between different GNSS systems. Most commonly used GNSS bands are as follows:

1. L1 Band (1575.42 MHz)-

Usage: The L1 band is the primary frequency used by most GNSS receivers. It carries the standard positioning service (SPS) signals, which are freely available to civilian users.

Advantages: Signals in the L1 band are less affected by ionospheric delays compared to lower frequency bands. This band is widely supported by all major GNSS systems, including GPS, GLONASS, Galileo, and BeiDou, ensuring broad compatibility and availability.

Applications: Commonly used in consumer-grade devices such as smartphones, car navigation systems, and handheld GPS units.

2. L2 Band (1227.60 MHz)-

Usage: The L2 band transmits the precise positioning service (PPS) signals, primarily intended for military and authorized users. However, civilian signals such as L2C (a modernized civilian signal) are also available.

Advantages: L2 signals, when combined with L1 signals, allow for more accurate correction of ionospheric errors. This dual-frequency capability improves positioning accuracy.

Applications: Used in high-precision applications such as surveying, geodesy, and scientific research.

3. L5 Band (1176.45 MHz)-

Usage: The L5 band is designated for safety-of-life applications and is used by modernized GNSS systems to provide enhanced accuracy and reliability.

Advantages: Signals in the L5 band have a higher power level and a wider bandwidth, which improves signal robustness and accuracy. They are also less prone to multipath errors and interference.

Applications: Critical for aviation, maritime navigation, and other safety-of-life services where precision and reliability are paramount.

4. Other Bands (E1, E6, B1, B2, etc.)-

Usage: Different GNSS systems use additional frequency bands to provide specific services and enhance overall system performance. For example, Galileo uses E1, E5, and E6 bands, while BeiDou employs B1, B2, and B3 bands.

Advantages: These additional bands support advanced services, improve interoperability between systems, and offer redundancy, ensuring continuous availability of positioning services.

Applications: Used in a variety of specialized applications, including scientific research, high-precision surveying, and augmented GNSS services.

8.4 Sources of error in GNSS

Global Navigation Satellite Systems are highly reliable and accurate, but various sources of error can affect the precision of the positioning data they provide. Understanding these sources of error is crucial for improving GNSS performance and developing techniques to mitigate them. The primary sources of error in GNSS are as follows:

1. Satellite Clock Errors

Each GNSS satellite is equipped with highly accurate atomic clocks. However, even these clocks can experience slight deviations from the true time, leading to errors in the signals transmitted by the satellites. These errors can affect the accuracy of the calculated position, as the receiver uses the time of signal transmission to determine distance from the satellite.

2. Ephemeris Errors

Ephemeris errors, also known as orbital errors, occur due to inaccuracies in the satellite's reported position. GNSS satellites orbit the Earth, and their exact positions are calculated and transmitted in the navigation message. Any errors in these calculations can lead to incorrect positioning data.

3. Ionospheric Delays

The ionosphere is a layer of the Earth's atmosphere that is ionized by solar radiation. GNSS signals passing through the ionosphere are delayed due to the varying density of charged particles. This delay is frequency-dependent, meaning signals on different frequencies experience different delays. Dual-frequency receivers can measure and correct for this error by comparing the delays of signals at different frequencies.

4. Tropospheric Delays

The troposphere is the lowest layer of the Earth's atmosphere, and it contains water vapor and other particulates. GNSS signals slow down when passing through the troposphere, leading to delays. Unlike ionospheric delays, tropospheric delays are not frequencydependent, making them harder to correct. Models and real-time data are used to estimate and mitigate these errors.

5. Multipath Effects

Multipath effects occur when GNSS signals reflect off surfaces such as buildings, water bodies, and the ground before reaching the receiver. These reflected signals can interfere with the direct signals, causing errors in the calculated position. This is particularly problematic in urban environments with many reflective surfaces.

6. Receiver Noise

GNSS receivers themselves can introduce errors due to internal noise and inaccuracies in signal processing. Thermal noise, electronic interference, and imperfections in the receiver's hardware can all contribute to positioning errors. High-quality receivers with advanced signal processing algorithms can reduce the impact of receiver noise.

7. User Equivalent Range Error (UERE)

UERE encompasses several factors, including satellite clock errors, ephemeris errors, and atmospheric delays, that contribute to the overall error in the distance measurement between the satellite and the receiver. It provides a comprehensive measure of the uncertainties affecting the range measurements used to compute the receiver's position.

8.5 Mitigation Techniques

Several techniques and technologies are employed to mitigate these errors and improve GNSS accuracy like:

- 1. **Differential GNSS (DGNSS):** This technique uses a network of ground-based reference stations to provide corrections for GNSS errors. These stations calculate the errors in the GNSS signals and broadcast correction data to nearby GNSS receivers.
- 2. **Satellite-Based Augmentation Systems (SBAS)**: Systems such as WAAS (Wide Area Augmentation System) in the

U.S., EGNOS (European Geostationary Navigation Overlay Service) in Europe, and MSAS (Multi-functional Satellite Augmentation System) in Japan provide additional correction signals from geostationary satellites, enhancing GNSS accuracy and integrity.

- 3. **Real-Time Kinematic** (**RTK**): RTK is a high-precision technique that uses carrier phase measurements and corrections from a nearby base station to achieve centimeter-level accuracy. It is widely used in surveying, agriculture, and other applications requiring precise positioning.
- 4. Advanced Receiver Technologies: Modern GNSS receivers use advanced algorithms and signal processing techniques to mitigate errors. These include multipath mitigation, enhanced ionospheric and tropospheric models, and improved clock error corrections.

8.6 Major GNSS Systems

There are four primary GNSS systems operational today:

- GPS (Global Positioning System): It is operated by the United States Department of Defense and GPS is the most widely recognized and used GNSS. It became fully operational in 1995 and consists of a constellation of at least 24 satellites in medium Earth orbit (MEO). It uses L1, L2 and L5 bands for operations. GPS has global coverage and provides accurate positioning anywhere on Earth. GPS offers a Standard Positioning Service (SPS) for civilian use and a Precise Positioning Service (PPS) for military and authorized users. GPS has major applications in navigation, surveying, agriculture, aviation, and more.
- GLONASS (Global Navigation Satellite System): It is operated by Russian Federation Government. GLONASS is Russia's GNSS, which became fully operational in 1996.

The system comprises a constellation of 24 satellites in MEO. It uses L1, L2 and L3 bands for operations. It also has global coverage comparable to GPS. It has better performance in high latitude regions particularly in northern latitudes due to its orbital design.

- 3. Galileo: It is operated by European Union through the European Space Agency (ESA). Galileo is Europe's GNSS, which aims to provide an independent high-precision positioning system. It became operational in 2016. It uses E1, E5a, E5b and E6 bands for operations. Galileo offers higher precision positioning compared to GPS and GLONASS. Provides free services for civilian use and encrypted services for governmental applications. It is designed to be interoperable with other GNSS systems.
- 4. BeiDou (BDS BeiDou Navigation Satellite System): It is operated by China National Space Administration (CNSA). BeiDou, also known as BDS, is China's GNSS. The system started regional service in 2000 and achieved global coverage in 2020. It uses B1, B2 and B3 bands for operations. It provides both regional (Asia-Pacific) and global services. It has a unique feature allowing two-way communication for short messages.

8.7 Regional GNSS Systems

Other than major GNSS systems, regionals GNSS systems that cover a specific region on Earth are alos available. Major ones are as follows:

- 1. **QZSS (Quasi-Zenith Satellite System):** It is operated by Japan. It has enhanced GNSS coverage and accuracy in the Asia-Oceania region. It complements GPS and provides additional services tailored to regional needs.
- NavIC (Navigation with Indian Constellation): Indian Regional Navigation Satellite System (IRNSS) or NavIC

(operational name) is a regional GNSS system operated by the Indian Space Research Organisation. It provides precise positioning services to users in India and the surrounding region (about 1,500 km around it). Plans are to further extent this range up to 3,000 Km. IRNSS constellation consists of 8 satellites in space, along with two satellites on ground as stand-by. IRNSS has "standard positioning service" for civilian use and "restricted service" with encryption for defence purposes and both the services use L5 (1176.45 MHz) and S band (2492.028 MHz) for operations. The constellation is operational in space in since 2018.

8.8 Conclusion

Global Navigation Satellite Systems are indispensable in the modern world due to their ability to provide accurate and reliable positioning, navigation, and timing services. GNSS have vast number of applications ranging from transportation, agriculture, emergency response, telecommunications, and scientific research to autonomous systems.

Chapter 9

AI Applications in Remote Sensing and GIS

Artificial Intelligence (AI) has revolutionized numerous industries, and its impact on remote sensing and Geographic Information Systems (GIS) is profound. By integrating AI with these technologies, a new potential for data analysis, pattern recognition, predictive modeling, and decision-making has been unlocked. This chapter explores how AI enhances remote sensing and GIS applications, focusing on various techniques, case studies, and future trends.

9.1 AI in Remote Sensing

Remote sensing involves acquiring information about the Earth's surface without direct contact, typically through satellites or aerial imagery. AI algorithms enhance remote sensing by improving image analysis, feature extraction, and image classification.

9.1.1 Image Classification

One of the primary applications of AI in remote sensing is image classification. The machine learning algorithms such as random forest, support vector machine, XGBoost, and neural networks are capable enough to perform Level-1 land use and land cover tasks. Similarly, Deep Learning based methods such as Convolutional Neural Networks (CNNs) excel in recognizing patterns and categorizing images. These models can classify land cover types (forests, urban areas, water bodies) with high accuracy by analyzing spectral signatures in satellite imagery. For instance, Figure 9.1 shows the Level-I classification of Delhi city using the Random Forest algorithm on Sentinel-2 data. The red color represents built-up area, blue represents water, dark green shows naturally vegetated areas, yellow represents vacant or barren land, bright green

represents agriculture, and light green represents grassland or wastelands.



Figure 9.1: Level-1 Classification of Delhi city using Random Forest Algorithm.

9.1.2 Change Detection

ML and DL methods also facilitates change detection in remote sensing. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, including architectures like GRU (Gated Recurrent Unit) and Bi-LSTM (Bidirectional LSTM), are particularly effective in identifying temporal changes in sequential data. By comparing time-series images, these models can detect changes in land use, vegetation growth, urban expansion, and environmental degradation.

 Object Detection: AI-driven object detection models such as YOLO (You Only Look Once) (Fig. 9.2), Faster R-CNN, SSD (Single Shot MultiBox Detector), and RetinaNet are utilized to identify and locate specific objects within remote sensing imagery. Applications include detecting ships in maritime surveillance, vehicles in traffic monitoring, and buildings in urban planning.



Figure 9. 2. *Building detection in Cartosat-3 data in parts of Delhi using Yolov8 algorithm.*

9.1.3 Hyperspectral and Multispectral Analysis

Remote sensing often involves hyperspectral and multispectral imaging, capturing data across various wavelengths. AI algorithms, particularly deep learning models like 3D-CNNs (Three-Dimensional Convolutional Neural Networks) and hybrid CNN-RNN architectures, can analyze these complex datasets to identify materials and assess their properties. This capability is vital for applications like mineral exploration, agriculture monitoring, and environmental assessment.

9.2 AI in GIS

Geographic Information Systems (GIS) involve the storage, analysis, and visualization of spatial data. AI enhances GIS by enabling advanced spatial data analysis, predictive modeling, and decision support systems.

- 1. Spatial Data Analysis: AI techniques such as K-means clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), and Random Forest classification help analyze spatial data more effectively. For example, clustering algorithms can group similar geographic features, aiding in regional planning and resource allocation. Classification algorithms can categorize land parcels based on usage, improving land management practices.
- 2. Predictive Modeling: AI-driven predictive modeling in GIS helps forecast future spatial patterns and trends. Machine learning models such as Gradient Boosting Machines (GBMs), XGBoost, and LightGBM can predict urban growth, environmental changes, and disaster risks. For instance, predictive models can simulate flood risks based on historical data and climate projections, assisting in disaster preparedness and mitigation.
- 3. **Route Optimization**: AI algorithms optimize routing and logistics in GIS applications. Techniques like Genetic Algorithms (GAs), Ant Colony Optimization (ACO), and Reinforcement Learning models such as Q-Learning and Deep Q-Networks (DQN) find the most efficient paths for transportation and delivery services. These models consider various factors such as traffic conditions, road networks, and delivery constraints to minimize travel time and costs.
- 4. **Smart Cities and IoT Integration**: In the context of smart cities, AI and GIS integration is crucial. AI processes data

from Internet of Things (IoT) devices, such as sensors and cameras, using models like Long Short-Term Memory (LSTM) networks and Autoencoders to monitor urban environments in real-time. Applications include traffic management, air quality monitoring, and infrastructure maintenance. GIS platforms visualize this data, providing insights for urban planners and policymakers.

9.3 Case Studies

- 1. **Precision Agriculture**: AI in remote sensing and GIS plays a pivotal role in precision agriculture. By analyzing satellite imagery and sensor data with models such as Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs), AI can assess crop health, soil moisture, and pest infestations. GIS integrates these insights with spatial data, enabling farmers to make informed decisions about irrigation, fertilization, and pest control. This approach enhances crop yields and resource efficiency.
- 2. **Disaster Management:** AI and GIS are instrumental in disaster management. During natural disasters like hurricanes and earthquakes, AI models like UNet and SegNet analyze satellite imagery to assess damage and identify affected areas. GIS platforms visualize this data, guiding emergency response teams in resource allocation and rescue operations. Predictive models like Random Forests and Gradient Boosting Machines (GBMs) also forecast disaster impacts, aiding in preparedness and risk reduction.
- Environmental Monitoring: Environmental monitoring benefits significantly from AI and GIS integration. AI algorithms such as Random Forests, k-Nearest Neighbors (k-NN), and Gradient Boosting Machines (GBMs) analyze remote sensing data to monitor deforestation, water quality, and pollution levels. GIS platforms map these changes,

providing a spatial context for environmental policies and conservation efforts. For instance, AI-driven models can detect illegal logging activities, enabling timely interventions.

4. Urban Planning: Urban planning leverages AI and GIS for sustainable development. AI models like Support Vector Machines (SVMs), Decision Trees, and Neural Networks analyze population growth, land use patterns, and infrastructure needs. GIS visualizes these data layers, supporting urban planners in designing efficient transportation networks, green spaces, and residential areas. This integrated approach promotes balanced urban growth and resource management.

9.4 Conclusion

AI has significantly transformed remote sensing and GIS, offering advanced capabilities for data analysis, predictive modeling, and decision support. The integration of AI with these technologies enables more accurate, efficient, and insightful analyses of spatial data, addressing various challenges in agriculture, disaster management, environmental monitoring, and urban planning. As AI algorithms and big data technologies continue to evolve, the future holds immense potential for further advancements and applications in remote sensing and GIS, contributing to sustainable development.

Glossary

Ι

A Analysis p1, Aspect p25, Attribute p4, Attribute Table p15, Automation p71, Azimuth p26.

Band p76, Buffer p18.

С

B

Class p26, Contour p39, Coordinate System p20, Coverage p6, Cylindrical Projection p56.

D

Data Source p71, Data Model p9, Data Type p12, Database p17, Datum p23, DEM p8.

E

Ellipsoid p21.

G

Geocodingp8,GeographicCoordinateSystem(GCS)p20,Geoprocessingp63,Georeferencedp1, GPSp80.

Identify p8, Image p23, Interpolation p36, Intersect p26.

J

Join p18

K

Key attributes p17, Kriging p 36

L

Latitude p20, Layer p9, Legend p57, Longitude p20

Μ

Map elements p53, Map Projection p20.

Ν

Nearest-neighbor resampling p71, NSDI (National Spatial Data Infrastructure) p1.

0	Т
OGC (OpenGIS Consortium) p61.	Topology p5.
Р	U
Prime meridian p20, Projected	Union p32, UTM (Universal
coordinate system p24, Projection	Transverse Mercator) p29.
p24.	
R	V
Reference system p68.	Vector p5, Vector model p5.
S	W
Scale bar p54, Shortest path p34,	WGS 84 p22.
Slope p39, Spatial analysis p3,	
Spheroid p21.	
	7

Ζ

Zone p29.

इसरो ंडल्व

Deputy General Manager

Regional Remote Sensing Centre-North National Remote Sensing Centre Indian Space Research Organisation New Delhi