

UNIT 4

Spatial Data Modeling and Analysis

Geographical Information System (CACS454) — BCA IV/VIII

The real analytical power of a GIS lies not in simply storing geographic data, but in modeling spatial relationships and applying analytical operations that convert raw data into decision-making information. This unit builds on the data structures introduced earlier by examining how spatial relationships are modeled, and by surveying the core families of analysis performed on vector and raster data.

1. Spatial Data Modeling

Spatial data modeling is the process of formally representing geographic features and, critically, the spatial relationships between them, in a way that a GIS can query and analyze. While a general data model (Unit 3) defines how features are stored, spatial data modeling focuses specifically on capturing how those features relate to one another in space.

Types of Spatial Relationships

Relationship Type	Description	Example
Topological	Relationships that remain true regardless of scale or projection, such as connectivity, adjacency and containment.	A road segment connects to another at a junction node.
Metric (Geometric)	Relationships based on measurable distance and direction between features.	A school is 500 m from the nearest health post.
Order / Directional	Relationships describing relative position such as above, below, north of.	District B lies north of District A.

Topology in Spatial Modeling

Topology explicitly encodes the relationships of adjacency (what touches what), connectivity (what is linked to what) and containment (what is inside what) between features. Topological modeling allows a GIS to answer questions such as “which parcels share a boundary with this one?” or “which pipe segments are connected to this valve?” — questions that pure geometry alone cannot answer.

- Adjacency — features that share a common boundary (e.g. two neighbouring land parcels).

- Connectivity — features linked end-to-end, essential for network analysis (e.g. road or river networks).
- Containment — one feature lying entirely within another (e.g. a well located inside a ward boundary).

Cartographic (Analytical) Modeling

Beyond representing individual relationships, GIS analysts build cartographic models — step-by-step flowcharts of data layers and operations — to plan complex, multi-layer analyses before they are executed. A cartographic model typically identifies the input layers, the sequence of operations (such as buffer, reclassify or overlay) applied to them, and the resulting output layer, providing a transparent and repeatable analysis workflow.

Note: A well-built cartographic model — for example, one combining slope, land-use and distance-to-road layers for a site-suitability study — can be re-run automatically whenever any input layer is updated.

2. Vector-Based Analysis

Vector-based analysis operates on the precise point, line and polygon geometry of vector features, and is generally best suited to questions involving discrete, well-bounded features such as parcels, roads or administrative units.

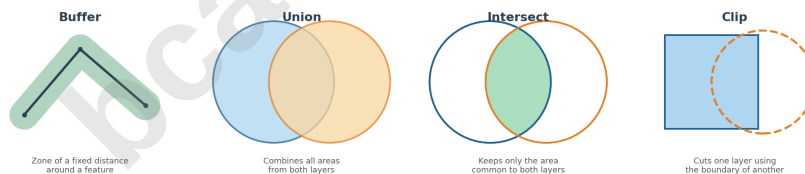


Figure 4.1 — Common vector analysis operations: buffer, union, intersect and clip.

Operation	What It Does
Buffer	Creates a new polygon representing a zone of a specified distance around a point, line or polygon feature.
Overlay — Union	Combines two polygon layers, preserving all areas and attributes from both inputs.
Overlay — Intersect	Retains only the area common to both input layers, combining their attributes.
Clip	Extracts the portion of one layer that falls within the boundary of another, without merging attributes.
Dissolve	Merges adjacent polygons that share a common attribute value into a single feature.
Proximity / Nearest	Identifies the nearest feature(s) in one layer relative to features in another.

Operation	What It Does
Network Analysis	Uses connectivity between line features (e.g. roads) to solve routing, shortest-path and service-area problems.
Spatial Query	Selects features based on a spatial condition, such as “within”, “touches” or “intersects” another layer.

Example: Siting a New Health Post

A planner buffers all existing health posts by a 5 km service radius, then overlays this buffer layer with a settlements layer using an intersect operation to identify which settlements already fall within range. A subsequent spatial query selects the remaining settlements outside every buffer, revealing the underserved areas where a new health post is most needed.

3. Raster-Based Analysis

Raster-based analysis operates cell-by-cell across one or more grids, and is generally best suited to continuous phenomena such as elevation, rainfall or land-cover, and to analyses that combine many layers simultaneously across an entire study area. The set of techniques for combining and transforming raster layers using mathematical and logical operations is known as map algebra.

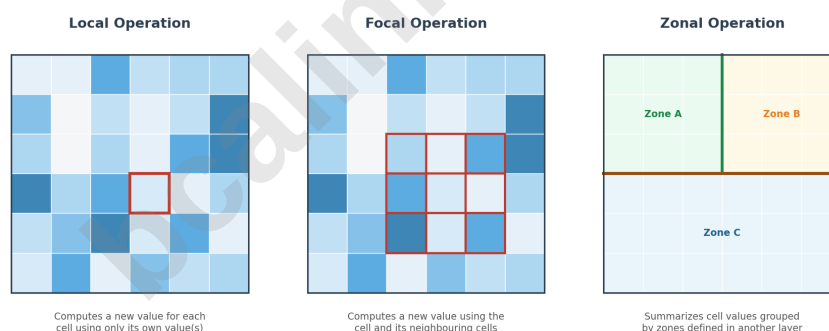


Figure 4.2 — The three fundamental classes of map algebra operation.

Operation Class	What It Does
Local	Computes an output value for each cell based only on the input value(s) at that same cell location (e.g. reclassifying elevation into height classes).
Focal (Neighbourhood)	Computes an output value for each cell based on the values of its surrounding neighbourhood (e.g. smoothing, slope, aspect calculation).
Zonal	Computes summary statistics (mean, sum, majority) for cells grouped by zones defined in a separate zone layer (e.g. average rainfall per district).
Global	Computes an output value for each cell based on all cells in the entire raster (e.g. Euclidean distance from the nearest river across the whole study area).

Common Raster Analysis Techniques

- **Reclassification** — grouping or recoding cell values into new categories (e.g. converting continuous slope values into “flat / moderate / steep” classes).
- **Surface Analysis** — deriving slope, aspect and hillshade from an elevation raster (DEM) to characterize terrain.
- **Weighted Overlay** — combining several reclassified raster layers using weighted map algebra to produce a composite suitability or risk score.
- **Distance / Cost Surface Analysis** — calculating straight-line or cost-weighted distance from a source feature across every cell in the study area.

Note: Vector analysis preserves precise boundaries and is ideal for discrete legal or administrative features, while raster analysis excels at combining many continuous surfaces quickly across a whole area — the two approaches are often used together in the same GIS project.

Key Takeaways

- ✓ Spatial data modeling represents topological, metric and directional relationships between features, going beyond simple geometry.
- ✓ Topology captures adjacency, connectivity and containment, enabling relationship-based spatial queries.
- ✓ Cartographic modeling plans multi-step, multi-layer analyses as a transparent, repeatable workflow.
- ✓ Vector-based analysis (buffer, overlay, clip, dissolve, network analysis) works with precise feature geometry, best for discrete features.
- ✓ Raster-based analysis (map algebra — local, focal, zonal and global operations) works cell-by-cell, best for continuous surfaces.
- ✓ Real-world GIS projects typically combine both vector and raster analysis to reach a final decision.

— *End of Unit 4* —