Chapter 4: Maps, Data Entry, Editing, and Output

Joaquin Cavieres

Geoinformatics, Bayreuth University

1. Building a GIS database
2. Digitizing: coordinate capture
3. Coordinate transformation
4. Output: Hardcopy maps, digital data, and metadata

Building a GIS database

Introduction

Spatial data entry and editing are frequent activities for many GIS users. A large number of coordinates is needed to represent features in a GIS, and each coordinate value must be entered into the GIS database. But,

Often it is slow, even with automated techniques, and spatial data entry and editing take significant time for most organizations.

Introduction

Spatial data entry and editing are frequent activities for many GIS users. A large number of coordinates is needed to represent features in a GIS, and each coordinate value must be entered into the GIS database. But,

Often it is slow, even with automated techniques, and spatial data entry and editing take significant time for most organizations.

The spatial data sources can be categorized as:

1. Hardcopy: are any drawn, written, or printed documents, including hand-drawn maps, manually measured survey data, legal records, and coordinate lists with associated tabular data (Fig. 1)
2. Digital: is the process of collecting digital coordinates. Digitizing is a common data entry method today, although primarily from satellite and aerial images. In this type of category are the *digital maps* (Fig. 2)

Figure 1: This early map of northern Europe shows approximate shapes and relative location (Bolstad (2016)).

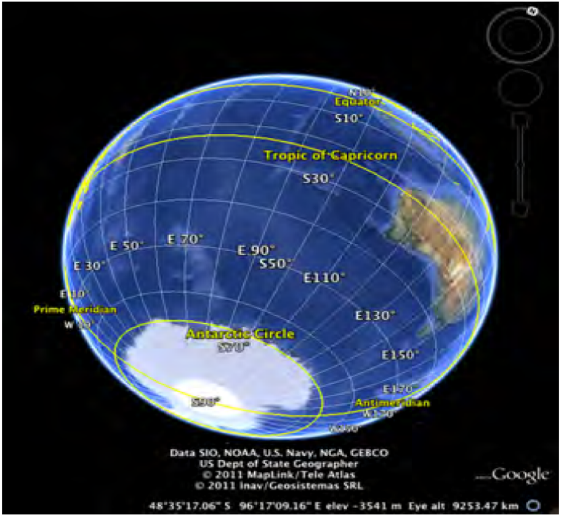


Figure 2: An example of commonly produced digital maps (courtesy Google) (Bolstad (2016)).

Hardcopy or digital maps contain several components:

- *Data area*: occupies the largest part of the map (it contains most of the depicted spatial data).
- *Neatline*: included to provide a frame around all map elements.
- *Insets*: contain additional map elements.
- *Scalebars*: legends, titles, and other graphic elements such as a north arrow are often included.
- *Map scale*: commonly defined as the ratio of the distance on the map to corresponding distance on the ground.

The Fig.3 shows graphically these components.

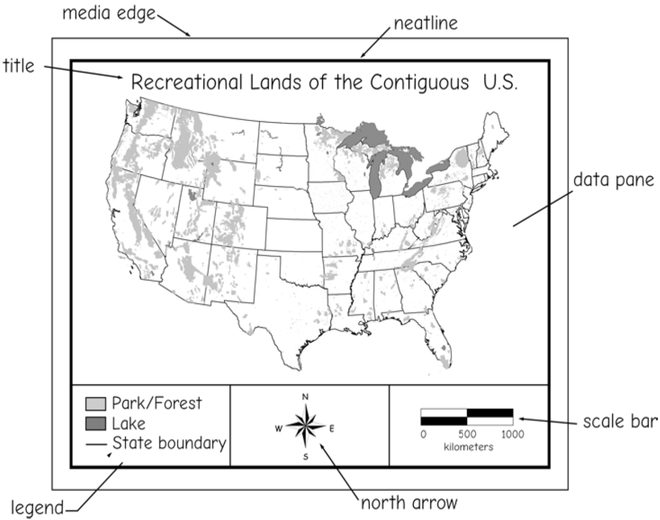


Figure 3: An example of a map and its components (Bolstad (2016)).

Also, maps often depict coordinate lines. If the lines are constant for latitude and longitude, a set of coordinate lines is called a *graticule*. These lines can appear curved, depending on the map scale the map coordinate system, and the location of the area on Earths (Fig.4).

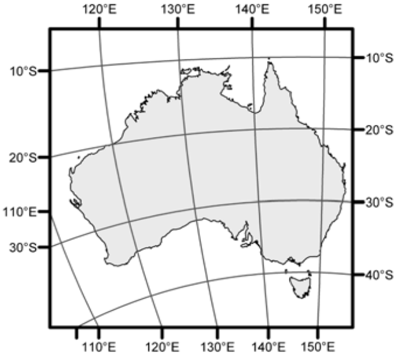


Figure 4: A *graticule* of constant latitude and longitude (Bolstad (2016)).

Maps can also depict a *grid* consisting of lines of constant coordinates. Grid lines are typically drawn in both the x and y directions, and appear straight on most maps (Fig.5).

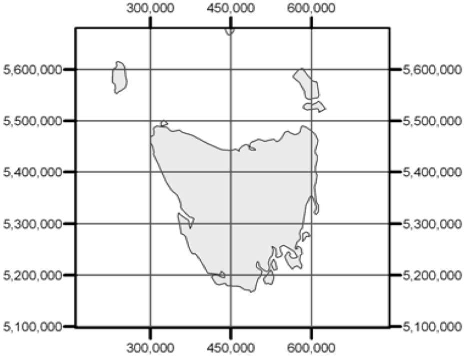


Figure 5: A *grid* of constant x and y coordinates. (Bolstad (2016)).

Graticules and grids are useful because they provide a reference against which location can be quickly estimated.

Archives of hard copy maps are a rich source of geographic data. However, as geographic science migrates to digital formats, these maps often sit in drawers unused. By scanning these maps and managing them using mosaic datasets, they can again be used, analyzed, and shared.

Graticules and grids are useful because they provide a reference against which location can be quickly estimated.

Archives of hard copy maps are a rich source of geographic data. However, as geographic science migrates to digital formats, these maps often sit in drawers unused. By scanning these maps and managing them using mosaic datasets, they can again be used, analyzed, and shared.

In summary,

- Hardcopy data are an important because most geographic information produced had been plotted on *cartometric* quality maps.
- Digital spatial data are those provided in a computer-compatible format. These include complete raster and vector data layers, text files, lists of coordinates, and digital images.

The objective of this chapter is to introduce spatial data entry via digitizing and coordinate surveying.

Map types

The maps depends of characteristics of them, for example:

- *Feature maps* are the one more simple map, because they map points, lines, or areas and provide nominal information (upper left).
- *Choropleth maps* depict quantitative information for areas (upper rigth)
- *Dotdensity maps* are another map form commonly used to show quantitative data (bottom left)
- *Isopleth maps* (aka. contour maps), display lines of equal value (bottom rigth)

See next figure (Fig.6:)

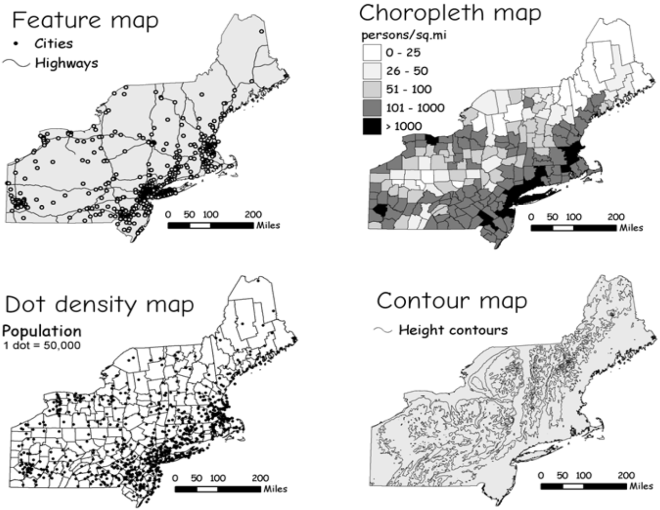


Figure 6: Common hardcopy map types depicting the northeastern United States (Bolstad (2016)).

Isopleth maps are commonly used to depict elevation. In this case the lower elevations typically pass under the higher elevations, and the isopleth is labeled with the tallest height (Fig. 7)

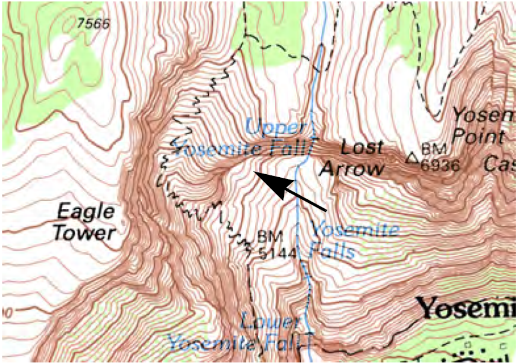


Figure 7: Lines on isopleth maps typically do not cross. However, as shown at the arrow in this image, lines can coincide when there is a common value (Bolstad (2016)).

Choropleths and isopleths have two key differences that separate them: choropleths rely on boundaries, while isopleths rely on density.

- Choropleths: Helps to present data on elections, population density, median income, or any other measure that can be tied to specific locations with boundaries.
- Isopleths: Radar maps used in meteorology are one of the most widely viewed examples of an isopleth area map.

So,

- Choropleths maps $\rightarrow$ display different values or percentages of a data set in the geographical location.
- Isopleths maps $\rightarrow$ rely strictly on continuous data

When should we used them?

- Choropleths maps: the data can be shown through percentages, the data can be split into ranges, the data shows just one variable.
- Isopleths maps: the data is not bound by location or population, the data does not need to be shown in proportions to a whole.

Map Ssale

Map scale is often reported as a distance conversion (Fig. 8).

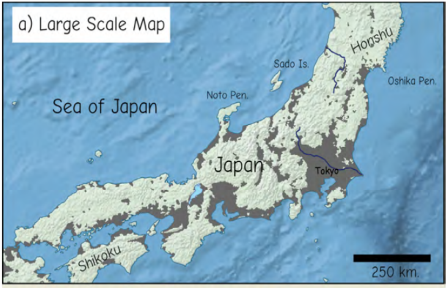


Figure 8: Coverage, relative distance, and detail change from larger-scale (Bolstad (2016)).

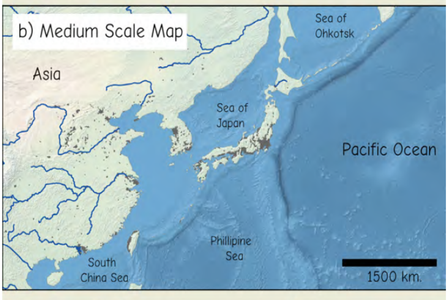


Figure 9: Coverage, relative distance, and detail change from medium-scale (Bolstad (2016)).

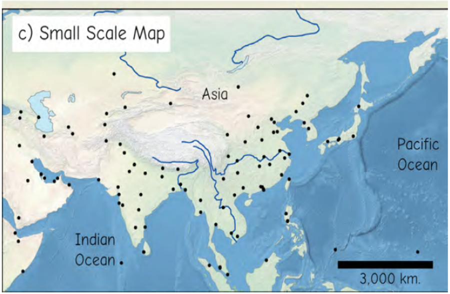


Figure 10: Coverage, relative distance, and detail change from small-scale (Bolstad (2016)).

The notion of large vs small scale is often confused because scale implies a ratio, or fraction.

A larger ratio signifies a large scale map, so a 1:24,000 scale map is considered large scale relative to a 1:100,000 scale map.

The notion of large vs small scale is often confused because scale implies a ratio, or fraction.

A larger ratio signifies a large scale map, so a 1:24,000 scale map is considered large scale relative to a 1:100,000 scale map.

For example:

A 1:100,000-scale map that is 50 cm (20 in) on a side covers more ground than a 1:24,000 scale map that is 50 cm on a side. However, it is the size of the ratio or fraction, and not the area covered, that determines the map scale.

Because maps often report an average scale, and because there are upper limits on the accuracy with which data can be plotted on a map, large-scale maps generally have less geometric error than small-scale maps if the same methods were used to produce them

Map generalization

Maps are abstractions of reality, as are spatial data in a GIS database. This abstraction introduces map generalization, the unavoidable approximation of real features when they are represented on a map.

The mapmaker determines the set of features to place on the map, and selects the methods to collect and represent the shape and location of these features on the map.

Map generalization

Maps are abstractions of reality, as are spatial data in a GIS database. This abstraction introduces map generalization, the unavoidable approximation of real features when they are represented on a map.

The mapmaker determines the set of features to place on the map, and selects the methods to collect and represent the shape and location of these features on the map.

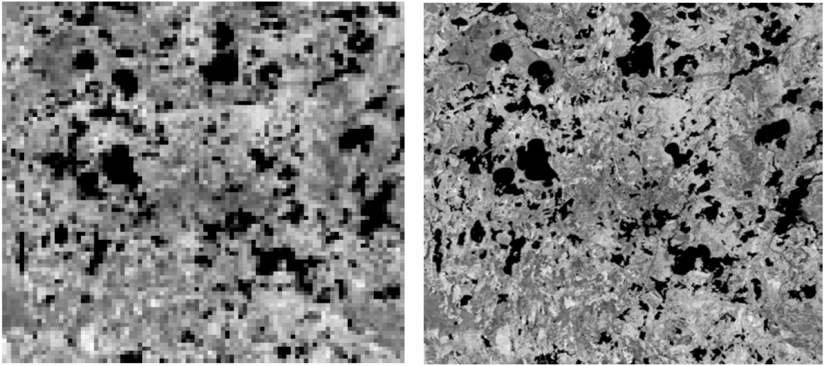


Figure 11: A mapmaker chooses the materials and methods used to produce a map, and so imposes a limit on spatial detail (Bolstad (2016)).

Feature generalization is one common form of generalization. Feature generalization is a modification of features when representing them on a map.

The geographic aspects of features are generalized because there are limits on the time, methods, or materials available when collecting geographic. These limits also apply when compiling or printing a map. These feature generalizations can be classed as:

- *Fused*: multiple features can be grouped to form a larger feature
- *Simplified*: boundary or shape details are lost or "rounded off"
- *Displaced*: features can be offset to prevent overlap or to provide a standard distance between mapping symbols
- *Omitted*: Small features in a group can be excluded from the map
- *Exaggerated*: standard symbol sizes are often chosen, for example, standard road symbol widths, which are much larger when scaled than the true road width.

See the next figure for a better explanation (Fig. 12).

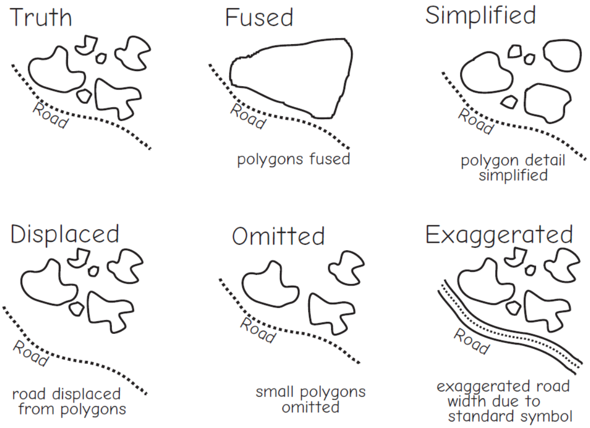


Figure 12: Generalizations common in maps (Bolstad (2016)).

Generalization is present at some level in every map, and should be recognized and evaluated for each map that is used as a source for data in a GIS (Fig. 13, 14 and 15).

Note that the maps are not drawn at true scale to facilitate comparison.

Magnified portion of a 1:24,000-scale map

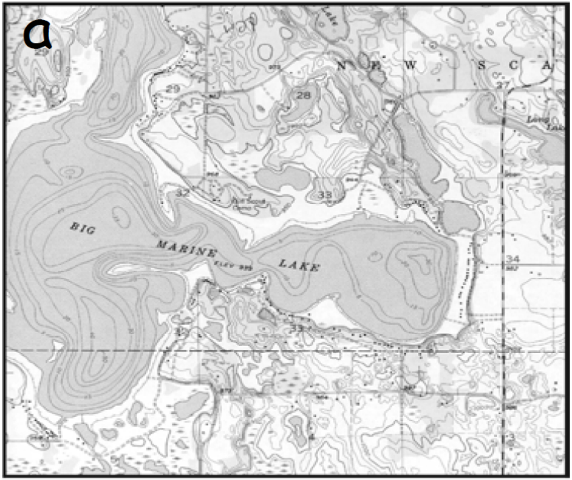


Figure 13: Map generalization from a large-scale (a, 1:24,000) (Bolstad (2016)).

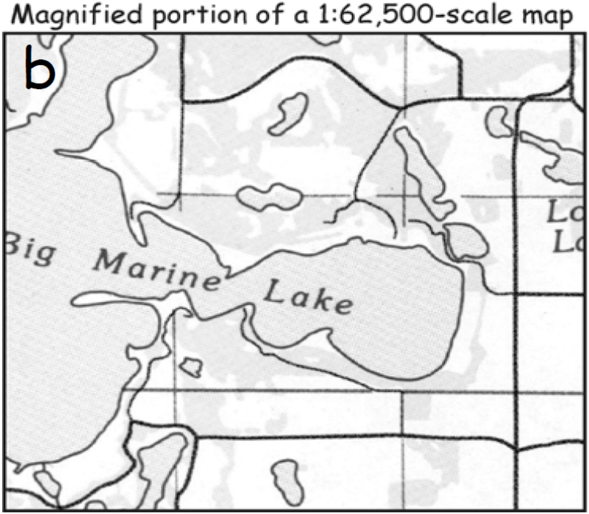


Figure 14: Map generalization from a intermediate-scale (b, 1:62,500) (Bolstad (2016)).

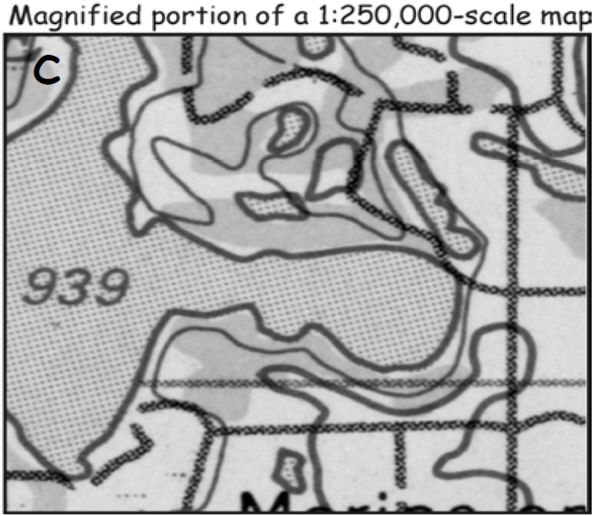


Figure 15: Map generalization from a small-scale (c, 1:250,000) (Bolstad (2016)).

Map boundaries and spatial data

One final characteristic of maps affects their use as a source of spatial data: hardcopy maps have edges, and discontinuities often occur at these edges.

Much digital data have been converted from legacy paper maps, so edge discontinuities have been carried to the present. These errors are disappearing as newer data are collected with digital methods, but will be encountered and should be understood.

Map boundaries and spatial data

One final characteristic of maps affects their use as a source of spatial data: hardcopy maps have edges, and discontinuities often occur at these edges.

Much digital data have been converted from legacy paper maps, so edge discontinuities have been carried to the present. These errors are disappearing as newer data are collected with digital methods, but will be encountered and should be understood.

- Large-scale, high-quality maps generally cover small areas.
- Cartometric maps larger than a meter in any dimension have proven to be impractical for most organizations. Maps above this size are expensive and difficult to print, store, or view. Thus, human ergonomics set a practical limit on the physical size of a map.

So, larger-scale maps generally cover smaller areas.

- A 1:100,000-scale map that is 18 in (47 cm) on a side spans approximately 28 miles (47 km).
- A 1:24,000-scale map that is 18 in on a side represents 9 mi (15 km) on Earths surface.

Problems often arise when adjacent maps are entered into a spatial database because features do not align or have mismatched attributes across map boundaries.

Different interpreters can also cause differences across map boundaries. Large-area mapping projects typically employ several interpreters, each working on different map sheets for a region.

All professional, large-area mapping efforts should have protocols specifying the scale, sources, equipment, methods, classification, keys, and cross-correlation to ensure consistent mapping across map sheet boundaries.

Different interpreters can also cause differences across map boundaries. Large-area mapping projects typically employ several interpreters, each working on different map sheets for a region.

All professional, large-area mapping efforts should have protocols specifying the scale, sources, equipment, methods, classification, keys, and cross-correlation to ensure consistent mapping across map sheet boundaries.

Digitizing: coordinate capture

Digitizing is the process by which coordinates from a map, image, or other sources are converted into a digital format in a GIS.

On the other hand,

Manual digitization is human-guided coordinate capture from a map or image source.

Digitizing is the process by which coordinates from a map, image, or other sources are converted into a digital format in a GIS.

On the other hand,

Manual digitization is human-guided coordinate capture from a map or image source.

Digitizing is the process by which coordinates from a map, image, or other sources are converted into a digital format in a GIS.

On the other hand,

Manual digitization is human-guided coordinate capture from a map or image source.

On-screen digitizing

On-screen digitizing, also known as heads-up digitizing, involves manually digitizing on a computer screen, using a digital image as a backdrop.

Digitizing software allows the operator to trace the points, lines, or polygons that are identified on the image (Fig.16).

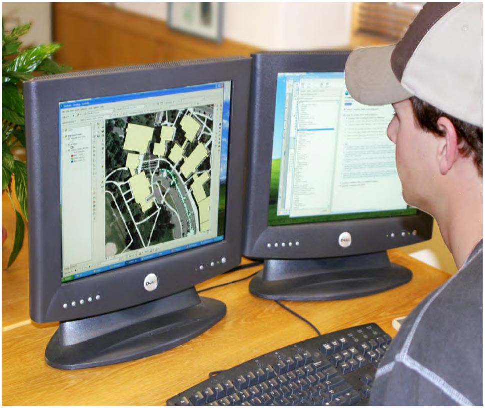


Figure 16: Images or maps are displayed on a computer screen and feature data digitized manually. Buildings, roads, or any other features that can be distinguished on the image can be digitized (Bolstad (2016)).

Digitizing software allows the human operator to specify the type of feature to be recorded, the extent and magnification of the image on screen, the mode of digitizing, and other options to control how data are input.

UNIVERSITÄT BAYREUTH

Hardcopy map digitization

Hardcopy digitizing is human-guided coordinate capture from a paper, plastic, or other hardcopy map.

An operator securely attaches a map to a digitizing surface and traces lines or points with an electrically sensitized puck (Fig.17).

Figure 17: Manual digitizing on a digitizing table (Bolstad (2016)).

While once a major method for capturing spatial data, hardcopy map digitizing is diminishing in importance as most paper documents have been converted to digital forms.

But,

Not all maps are appropriate as a source of information for GIS. The type of map, how it was produced, and the intended purpose must be considered when interpreting the information on maps.

While once a major method for capturing spatial data, hardcopy map digitizing is diminishing in importance as most paper documents have been converted to digital forms.

But,

Not all maps are appropriate as a source of information for GIS. The type of map, how it was produced, and the intended purpose must be considered when interpreting the information on maps.

Characteristics of manual digitizing

Manual digitizing is common because it provides sufficiently accurate data for many, if not most, applications.

- can be at least the accuracy of most maps or images, so the equipment, if properly used, does not add substantial error.
- Also requires low equipment investment, often just the software for image display and coordinate capture.
- Is often best because short training periods are required, data quality can be frequently evaluated, and digitizing equipment is commonly available.

There are a number of characteristics of manual digitization that can negatively affect the positional quality of spatial data.

As noted earlier, map or image scale and resolution impacts the spatial accuracy of digitized data.

The abilities and attitude of the person digitizing affects the geometric quality of manually digitized data. Operators differ in visual acuity, hand steadiness, attention to detail, and ability to concentrate.

There are a number of characteristics of manual digitization that can negatively affect the positional quality of spatial data.

As noted earlier, map or image scale and resolution impacts the spatial accuracy of digitized data.

The abilities and attitude of the person digitizing affects the geometric quality of manually digitized data. Operators differ in visual acuity, hand steadiness, attention to detail, and ability to concentrate.

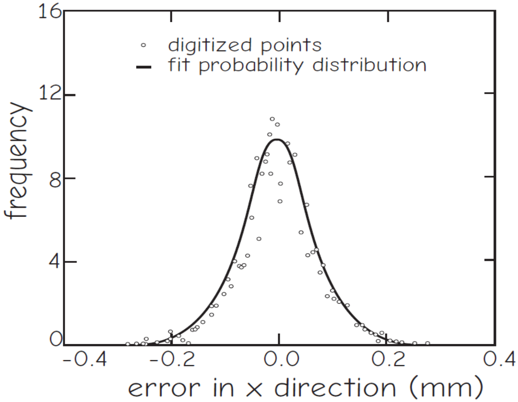


Figure 18: Digitizing error, defined by repeat digitizing. Points repeatedly digitized cluster around the true location and follow a normal probability distribution (from Bolstad et al., 1990).

The digitizing process

Displaying a digital image on screen or placing a map on a digitizing surface, and tracing the location of feature boundaries

↓

Coordinate data are sampled by manually positioning the puck or cursor over each target point and collecting coordinate locations

↓

Features that are viewed as points are represented by digitizing a single location.

↓

Lines are represented by digitizing an ordered set of points, and polygons by digitizing a connected set of lines.

↓

Lines are represented by digitizing an ordered set of points, and polygons by digitizing a connected set of lines.

→

Lines have a starting point, often called a ***starting node***, a set of ***vertices*** defining the line shape, and an ending *node*

↑

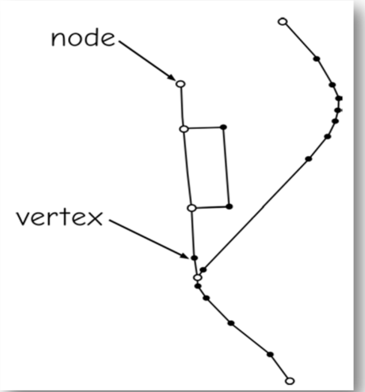


Figure 19: Nodes define the starting and ending points of lines. Vertices define line shape.

Digitizing errors, node and line snapping

Positional errors are inevitable when data are manually digitized. These errors can be small relative to the intended use of the data. However, these relatively small errors can still prevent the generation of correct networks or polygons.

Digitizing errors, node and line snapping

Positional errors are inevitable when data are manually digitized. These errors can be small relative to the intended use of the data. However, these relatively small errors can still prevent the generation of correct networks or polygons.

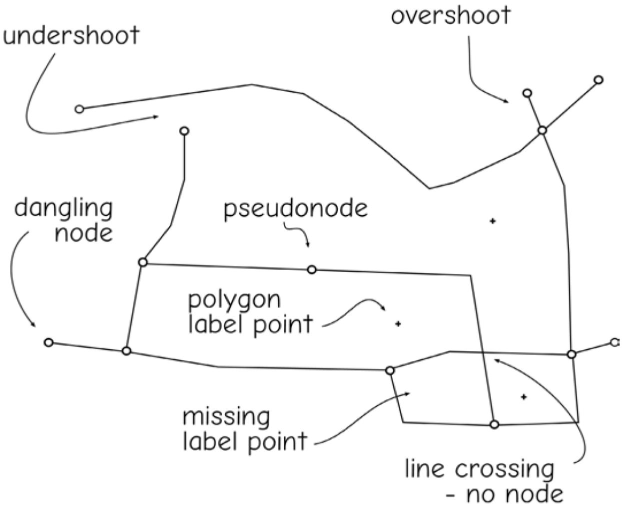


Figure 20: Common digitizing errors (Bolstad et al., 1990)

Undershoots and *overshoots* are common errors that occur when digitizing.

- Undershoots are nodes that do not quite reach the line or another node, and can cause unconnected networks and unclosed polygons.
- Overshoots are lines that cross over existing nodes or lines and typically do not cause problems when defining polygons. However, they can cause difficulties when defining and analyzing line networks.

Node snapping and *line snapping* are used to reduce undershoots and overshoots while digitizing.

Snapping is a process of automatically setting nearby points to have the same coordinates. Snapping relies on a *snap tolerance* or *snap distance*.

This distance can be interpreted as a minimum distance between features. Nodes or vertices closer than this distance are moved to occupy the same location (Fig. 21)

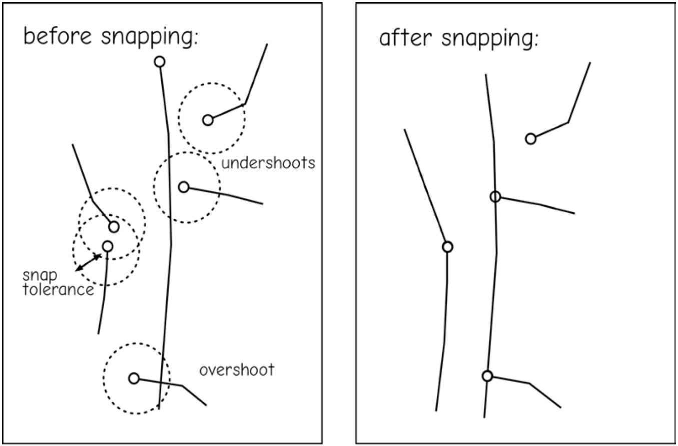


Figure 21: Undershoots, overshoots, and snapping. Snapping can join nodes, or can place a node onto a nearby line segment. Snapping does not occur if the nodes and/or lines are separated by more than the snap tolerance (Bolstad et al., 1990)

- Node snapping prevents a new node from being placed within the snap distance of an already existing node; instead, the new node is joined or snapped to the existing node.
- Line snapping inserts a node at a line crossing and clips the end when a small overshoot is digitized.

Remember:

Nodes are used to define the ending points of a line. By snapping two nodes together, we ensure a connection between digitized lines.

The snap distance must be carefully selected for snapping to be effective. If the snap distance is too short, then snapping has little impact.

Careful selection of the snap distance can reduce digitizing errors and significantly reduce time required for later editing.

Remember:

Nodes are used to define the ending points of a line. By snapping two nodes together, we ensure a connection between digitized lines.

The snap distance must be carefully selected for snapping to be effective. If the snap distance is too short, then snapping has little impact.

Careful selection of the snap distance can reduce digitizing errors and significantly reduce time required for later editing.

Remember:

Nodes are used to define the ending points of a line. By snapping two nodes together, we ensure a connection between digitized lines.

The snap distance must be carefully selected for snapping to be effective. If the snap distance is too short, then snapping has little impact.

Careful selection of the snap distance can reduce digitizing errors and significantly reduce time required for later editing.

Reshaping: line smoothing and thinning

Digitizing software can provide tools to smooth, densify, or thin points while entering data. One common technique uses *spline* functions to smoothly interpolate curves between digitized points and thereby both smooth and densify the set of vertices used to represent a line.

A spline is a set of polynomial functions that join smoothly. Polynomial functions are fit to successive sets of points along the vertices in a line; for example, a function can be fit to points 1 through 5, and a separate polynomial function fit to points 5 through 11 (Fig. 22).

Once the spline functions are calculated they can be used to add vertices.

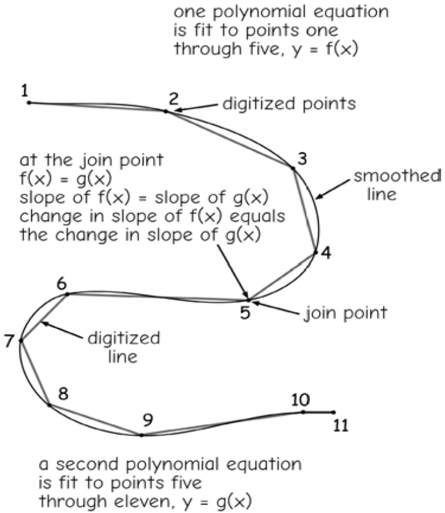


Figure 22: Spline interpolation to smooth digitized lines (Bolstad et al., 1990)

Many point thinning methods use a perpendicular weed distance, measured from a spanning line, to identify redundant points (Fig. 23, top). The Lang method exemplifies this approach. A spanning line connects two nonadjacent vertices in a line. A predetermined number of vertices is spanned initially.

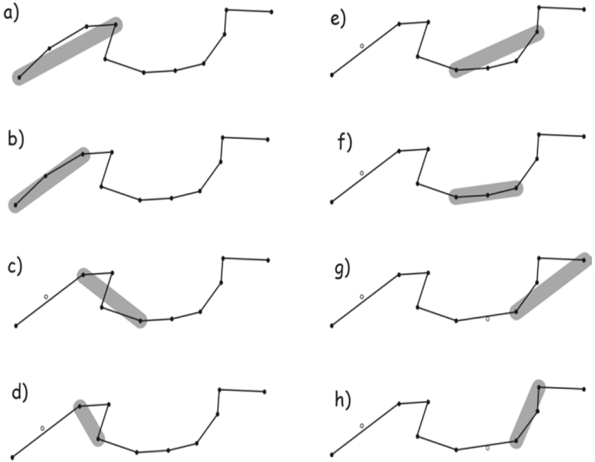


Figure 23: The Lang algorithm is a common line-thinning method. In the Lang method vertices are removed, or thinned, when they are within a weed distance to a spanning line (adapted from Weibel, 1997).

Scan digitizing

Optical scanning is another method for converting hardcopy documents into digital formats. Scanners have elements that emit and sense light.

This device measures both the precise location of the point being sensed and the strength of the light reflected or transmitted from that point. Reflected light intensities are sensed and converted to numbers.

Scan digitization usually requires some form of skeletonizing, or line thinning, particularly if the data are to be converted to a vector data format. Scanned lines are often wider than a single pixel (Fig. 24).

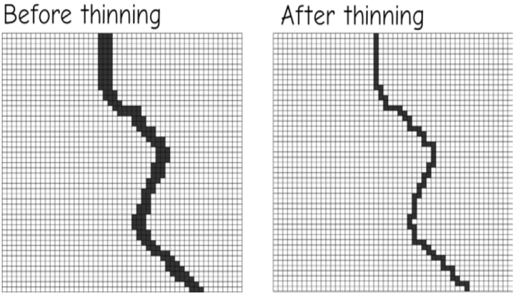


Figure 24: Skeletonizing, a form of line thinning that is often applied after scan digitizing (Bolstad et al., 1990)

Editing geographic data

Spatial data can be edited, or changed, for several reasons. Errors and inconsistencies are inevitably introduced during spatial data entry.

So,

- Identifying errors is the first step in editing. Errors can be identified by printing a map of the digitized data and verifying that each point, line, and area feature is present and correctly located.
- Software helps operators identify potential errors as well. Line features typically begin and end with a node, and nodes can be classified as connecting or dangling.
- Attribute consistency can also be used to identify errors. Operators note areas in which contradictory theme types occur in different data layers.

So,

- Identifying errors is the first step in editing. Errors can be identified by printing a map of the digitized data and verifying that each point, line, and area feature is present and correctly located.
- Software helps operators identify potential errors as well. Line features typically begin and end with a node, and nodes can be classified as connecting or dangling.
- Attribute consistency can also be used to identify errors. Operators note areas in which contradictory theme types occur in different data layers.

So,

- Identifying errors is the first step in editing. Errors can be identified by printing a map of the digitized data and verifying that each point, line, and area feature is present and correctly located.
- Software helps operators identify potential errors as well. Line features typically begin and end with a node, and nodes can be classified as connecting or dangling.
- Attribute consistency can also be used to identify errors. Operators note areas in which contradictory theme types occur in different data layers.

So,

- Identifying errors is the first step in editing. Errors can be identified by printing a map of the digitized data and verifying that each point, line, and area feature is present and correctly located.
- Software helps operators identify potential errors as well. Line features typically begin and end with a node, and nodes can be classified as connecting or dangling.
- Attribute consistency can also be used to identify errors. Operators note areas in which contradictory theme types occur in different data layers.

Many GIS software packages provide a comprehensive set of editing tools. Editing typically includes the ability to select, split, update, and add features (Fig. 25)

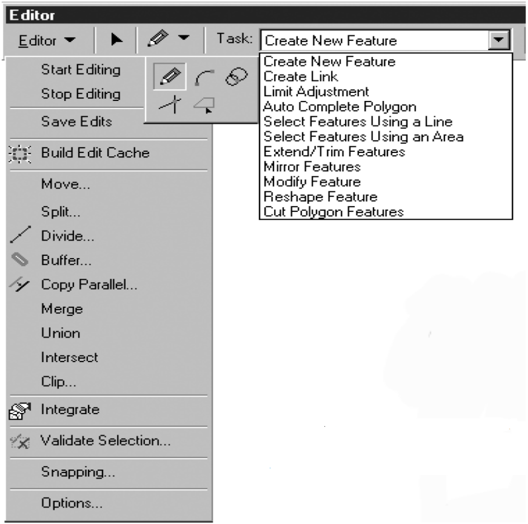


Figure 25: GIS software provides for a flexible and complete set of editing tools (Bolstad et al., 1990)

Features common to several layers

One common problem in digitizing derives from representation of features that occur on different maps or images. These features rarely have identical locations on each map or image, and often occur in different locations when digitized into their respective data layers (Fig. 26).

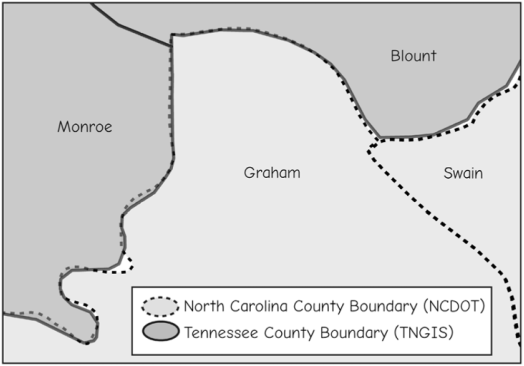


Figure 26: Common features can be spatially inconsistent in different spatial data layers. Note the boundaries are inconsistent in several locations, creating gaps and overlaps along their shared margin. These gaps must be addressed if the data sets are to be combined in an analysis (Bolstad et al., 1990)

There are several ways to remove this common feature inconsistency.

- One involves removing inconsistencies while redrafting the data from conflicting sources onto a new base map. Redrafting is labor intensive and time consuming, but forces a resolution of inconsistent boundary locations.
- A second, often preferable method involves establishing a master boundary that is the highest accuracy composite of the available data sets.

Coordinate transformation

Coordinate transformation:

A coordinate transformation brings spatial data into an Earth based map coordinate system so that each data layer aligns with every other data layer.

A good transformation helps avoid inconsistent spatial relationships such as farm fields on freeways, roads under water, or cities in the middle of swamps, except where these truly exist.

Coordinate transformation is most commonly used to convert newly digitized data from the digitizer/scanner coordinate system to a standard map coordinate system (Fig. 27).

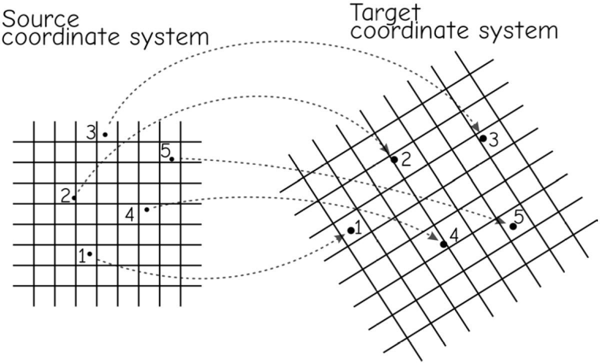


Figure 27: Control points in a coordinate transformation. Control points are used to guide the transformation of a source, input set of coordinates to a target, output set of coordinates (Bolstad et al., 1990)

Figure 28 depicts the application of a coordinate transformation in data development.

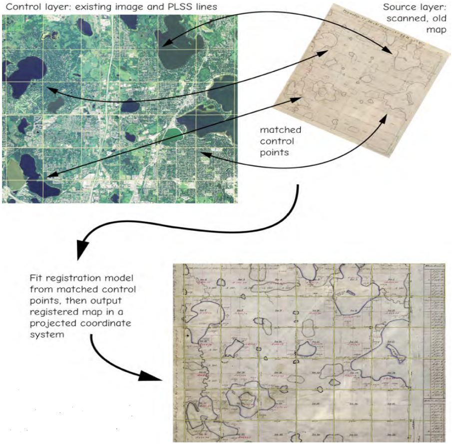
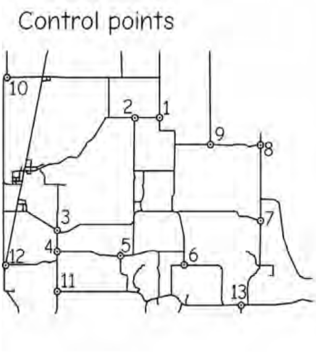


Figure 28: Coordinate transformation involves matching control points from known to target layers, and estimating then applying a transformation model (Bolstad et al., 1990)

Control points

A set of control points is used to transform the digitized data from the digitizer or photo coordinate system to a map-projected coordinate system.

These control points are used to estimate equations that we use for the coordinate transformation (Fig. 29).



ID	Digitizer coordinates X	Digitizer coordinates Y	Projection coordinates (UTM) E	Projection coordinates (UTM) N
1	103.0	-100.1	500,083.4	5,003,683.5
2	0.8	-69.1	504,092.3	5,002,499.5
3	-20.0	-69.0	504,907.5	5,002,499.5
4	-60.0	-47.0	506,493.3	5,001,673.5
5	-102.0	-47.2	508,101.3	5,001,651.7
6	-101.7	10.8	508,090.1	4,999,384.0
7	-86.0	75.8	507,475.9	4,996,849.0
8	-40.0	45.7	505,689.2	4,998,022.0
9	11.0	36.8	503,679.2	4,998,368.0
10	63.0	34.0	501,657.9	4,998,479.5
11	63.0	17.7	501,669.1	4,999,116.0
12	63.0	64.3	501,680.3	4,997,296.0
13	106.0	47.7	500,005.3	4,997,943.5

Figure 29: An example of control point locations from a road data layer, and corresponding digitizer and map projection coordinates (Bolstad et al., 1990)

The affine transformation

The affine coordinate transformation employs linear equations to calculate map coordinates. Map projection coordinates are often referred to as eastings (E) and northings (N), and are related to the x and y digitizer coordinates by the equations:

$$E = T_E + a_1 x + a_2 y \quad (1)$$
$$N = T_N + b_1 x + b_2 y \quad (2)$$

Equations 1 and 2 allow us to move from the arbitrary digitizer coordinate system to the project map coordinate system.

The affine system of equations has six parameters to be estimated, T_E, T_N, a_1, a_2, b_1, and b_2. Each control point provides E, N, x, and y coordinates, and allows us to write two equations.

Example

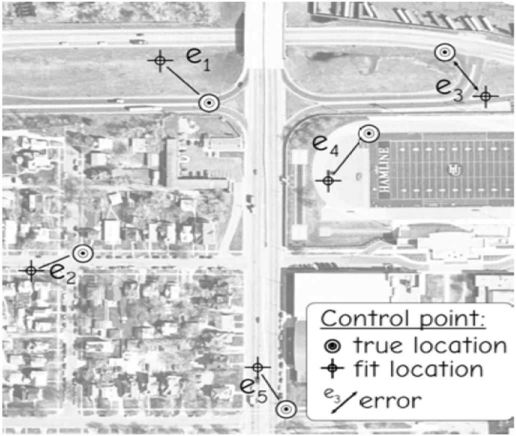


Figure 30: Examples of control points, predicted control locations, and residuals from coordinate transformation (Bolstad et al., 1990)

Other coordinate transformations

The conformal coordinate transformation is similar to the affine, and has the form:

$$
\begin{aligned}
E &= T_E + cd - dy & (3) \\
N &= T_N + dx + cy & (4)
\end{aligned}
$$

The coefficients T_E, T_N, c, and d are estimated from control point data. Like the affine transformation, the conformal transformation is also a first-order polynomial.

Unlike the affine, the conformal transformation requires equal scale changes in the x and y directions.

The x and y coefficients match across equations 3 and 4, thus there is a change in sign for the d coefficient. This results in a system of equations with only four unknown parameters, and so the conformal can be estimated when only two control points are available.

A caution when evaluating transformations

Selecting the best coordinate transformation to apply is a subjective process, guided by multiple goals. We hope to develop an accurate transformation based on a large set of well-distributed control points.

There are no clear rules on the number of points versus distribution of points, but it is typically best to strive for the widest distribution of points.

The transformation equation should be developed with the following observations in mind:

1. Bad control points happen, but we should thoroughly justify the removal of any control point.
2. A lower RMSE does not mean a better transformation. The RMSE is typically lower for a second and other higher-order polynomials than an affine transformation, but this does not mean the higher-order polynomial provides a more accurate transformation.

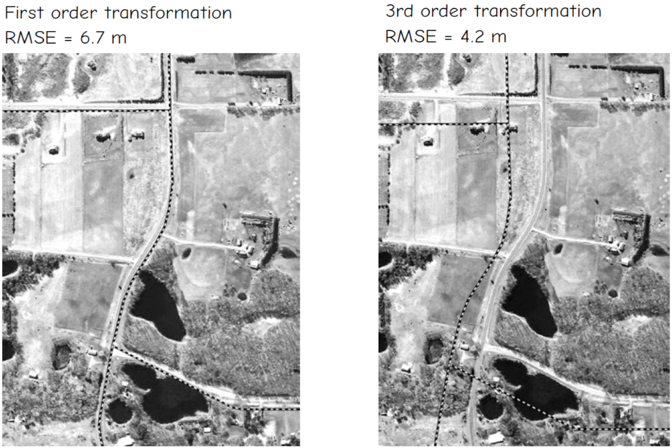


Figure 31: An illustration that RMSE should not be used to compare different order transformations, nor should it be used as the sole criterion for selecting the best transformation (Bolstad et al., 1990)

Control points from existing maps and digital data

Registered digital image data are common sources of ground control points, particularly when natural resources or municipal databases are to be developed for managing large areas. Digital images often provide a richly detailed depiction of surface features (Fig. 32).

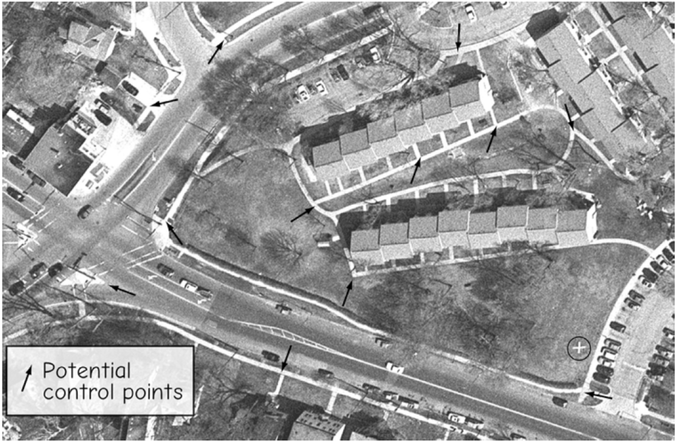


Figure 32: Potential control points, indicated here by arrows, can be extracted from digital images that have been geometrically corrected to provide an accurate, coordinate-projected rendering. (Bolstad et al., 1990)

Raster geometry and resampling

Data often must be resampled when converting between coordinate systems, or changing the cell size of a raster data set (Fig. 33).

Resampling is required when changing cell sizes because the new cell centers will not align exactly with old cell centers.

Common resampling approaches include:

- *Nearest neighbor*: taking the output layer value from the nearest input layer cell center)
- *bilinear interpolation*: distancebased averaging of the four nearest cells
- *cubic convolution*: a weighted average of the sixteen nearest cells,

See next figure for a better understanding (Fig. 33).

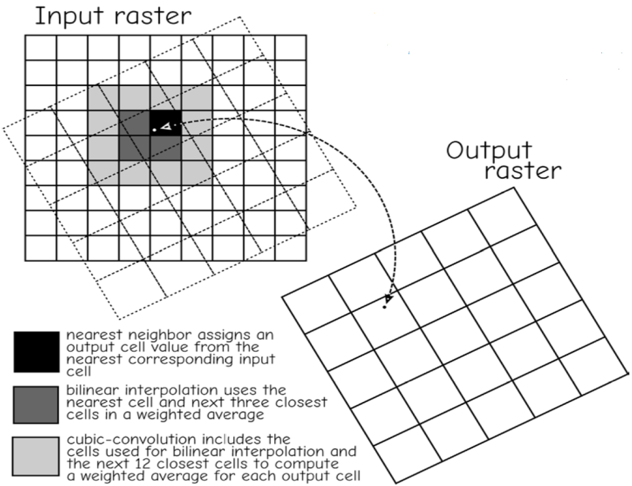


Figure 33: When the orientation or cell size of a raster data set is changed, output cell values are calculated based on the closest (nearest neighbor), four nearest (bilinear interpolation), or sixteen closest (cubic convolution) input cell values. (Bolstad et al., 1990)

Map projection vs. transformation

Map transformations should not be confused with map projections.

- A map transformation typically employs a statistically fit linear equation to convert coordinates from one Cartesian coordinate system to another.
- A map projection, described in Chapter 3, differs from a transformation in that it is an analytical, formula-based conversion between coordinate systems. **No statistical fitting process is used with a map projection**.

Map projection vs. transformation

Map transformations should not be confused with map projections.

- A map transformation typically employs a statistically fit linear equation to convert coordinates from one Cartesian coordinate system to another.
- A map projection, described in Chapter 3, differs from a transformation in that it is an analytical, formula-based conversion between coordinate systems. **No statistical fitting process is used with a map projection**.

Output: Hardcopy maps, digital data, and metadata

We create spatial data to use, share, and archive.

To be widely useful, we must also generate information, or metadata, about the spatial data we have created, and we can have to convert our data to standard forms.

Cartography and map design

A primary purpose of cartography is to communicate spatial information. This requires identification of the

- Intended audience
- Information to communicate
- Area of interest
- Physical and resource limitations

These considerations drive the major cartographic design decisions we make each time we produce a map:

- Scale, size, shape, and other general map properties
- Data to plot
- Symbol shapes, sizes, or patterns
- Labeling, including type font and size
- Legend properties, size, and borders
- The placement of all these elements on a map.

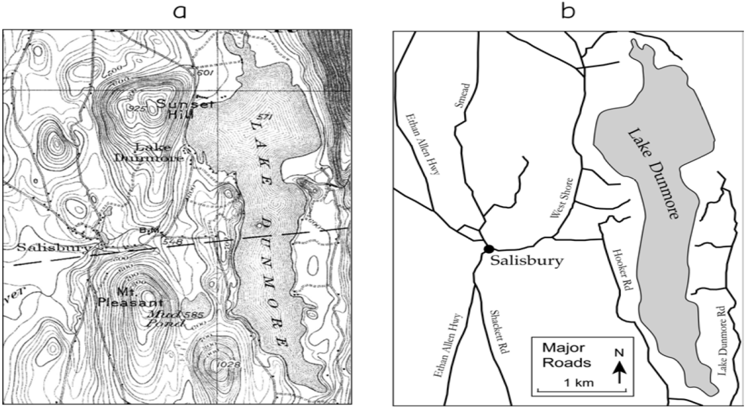


Figure 34: Example of a) a detailed, general-purpose map, here a portion of a United States Geological Survey map, and b) a specialized map focusing a specific set of selected features, here showing roads. (Bolstad et al., 1990)

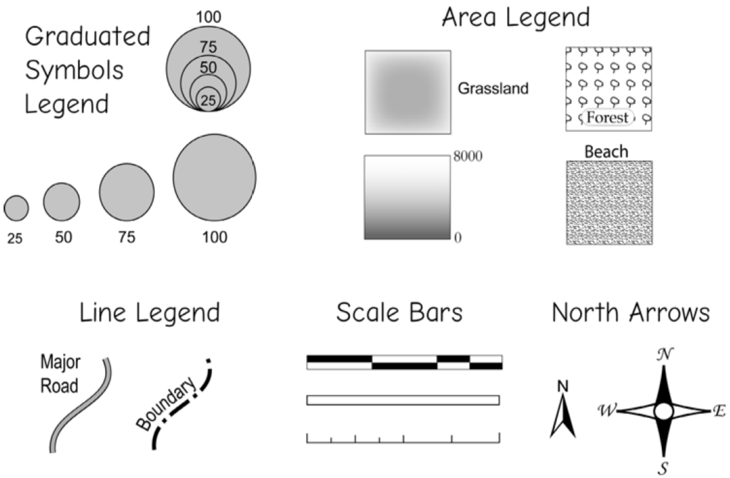


Figure 35: Examples of legend elements and representation of symbols (Bolstad et al., 1990)

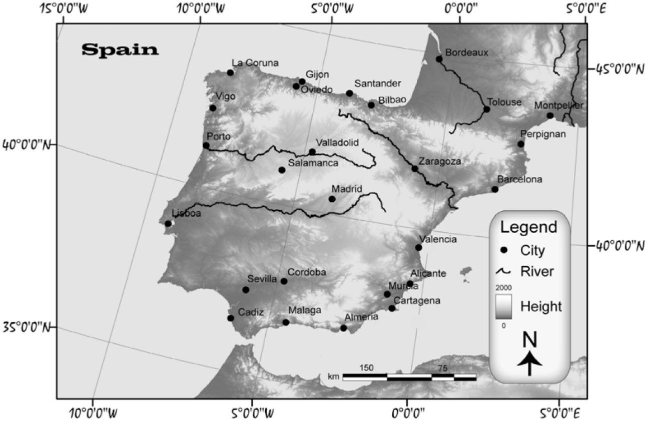


Figure 36: This figure shows a number of mistakes common for the neophyte cartographer, including small labels (cities) and mismatched fonts (graticule labels, title), poor labeling (city labels overlapping, ambiguously placed, and crossing distinctly shaded areas), unlabeled features (oceans and seas), poorly placed scale bar and legend, and unbalanced open space on the left side of the map (Bolstad et al., 1990)

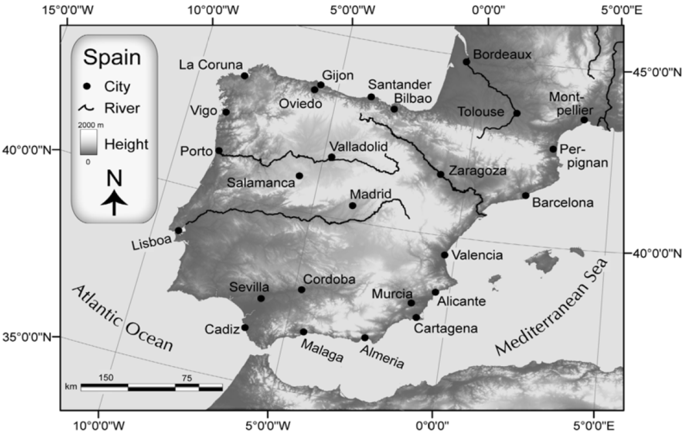


Figure 37: A correct map for the previous figure (Bolstad et al., 1990)

Thank You

References I

Bolstad, P. (2016). *GIS fundamentals: A first text on geographic information systems*. Eider (PressMinnesota).