[DM-02-007] The Raster Data Model

Abstract

The raster data model is a widely used method of storing geographic data. The model most commonly takes the form of a grid-like structure that holds values at regularly spaced intervals over the extent of the raster. Rasters are especially well suited for storing continuous data such as temperature and elevation values, but can hold discrete and categorical data such as land use as well. The resolution of a raster is given in linear units (e.g., meters) or angular units (e.g., one arc second) and defines the extent along one side of the grid cell. High (or fine) resolution rasters have comparatively closer spacing and more grid cells than low (or coarse) resolution rasters, and require relatively more memory to store. Active research in the domain is oriented toward improving compression schemes and implementation for alternative cell shapes (such as hexagons), and better supporting multi-resolution raster storage and analysis functions.

Keywords: compression, continuous data, discrete data, mixed pixel, pixel, resolution, tessellation data models

Author & citation

Pingel, T. (2018). The Raster Data Model. The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2018 Edition), John P. Wilson (Ed.).
DOI: 10.22224/gistbok/2018.3.11

This Topic is also available in the following editions: DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., and Wentz, E. (2006). The Raster Model. The Geographic Information Science & Technology Body of Knowledge. Washington, DC: Association of American Geographers.

Explanation

1. Definitions
2. Description
3. Examples
4. Storage and File Formats
5. Advantages and Disadvantages of the Raster Data Model

1. **Definitions**

2.5D: A system of recording values on a raster in which each grid cell has one and only one z-value.

continuous data: Field-like data in which values are present at any point within the spatial extent, such as elevation or temperature.

digital elevation model (DEM): A data model used to process, store, analyze, and display elevation data.

digital surface model (DSM): A type of DEM that represents a maximum value within a grid cell, thereby recording the tops of buildings, trees, and other objects.

digital terrain model (DTM): A type of DEM that aims to represent an idealized land surface where surface objects (buildings, trees, etc.) have been digitally removed.

discrete data: Object-like data, in which the spatial extent or boundaries of the features are definable.

extent: The area or distance in real space over which some geographic entity exists. In cartography and GIS, the extent of a representation is the size of the real space being represented.

file format: The specification for how data is stored a computer file. Important distinctions include those between binary and plaintext approaches, and between proprietary and free and open formats.

mixed pixel: A condition whereby more than one category of object is present within a single grid cell.

pixel: A portmanteau of “picture element”, the smallest unit of a raster. Sometimes referred to as a cell or grid point.

resolution: The degree of detail to which a phenomenon is detected or represented. Data are stored and rendered at some degree of representation resolution. In raster sensor arrays, resolution is defined by the dimensions of the individual sensors in terms of ground units (i.e., the width of one pixel in meters on the Earth).

2. **Description**

The raster data model, along with the vector data model, is one of the earliest and most widely used data models within geographic information systems (Tomlin, 1990; Goodchild, 1992, Maguire, 1992). It is typically used to record, analyze and visualize data with a continuous nature such as elevation, temperature, or reflected or emitted electromagnetic radiation. The term raster originated from the German word for screen, implying a series of orthogonally oriented parallel lines. Its origin as a description for images comes from the drawing performed by electron beams on cathode ray tube (CRT) screens in the early days of analog television, and the metaphor was subsequently extended to digital images as well. Digital rasters most often take the form of a regularly spaced, grid-like pattern of rows and columns, with each element referred to as a cell, pixel, or grid point. The raster is sometimes referred to as an image, array, surface, matrix, or lattice (Wise, 2000). Cells of the raster are most often square, but may be rectangular (with differing resolutions in x and y directions) or other shapes that can be tessellated such as triangles and hexagons (Figure 1; Peuquet, 1984).

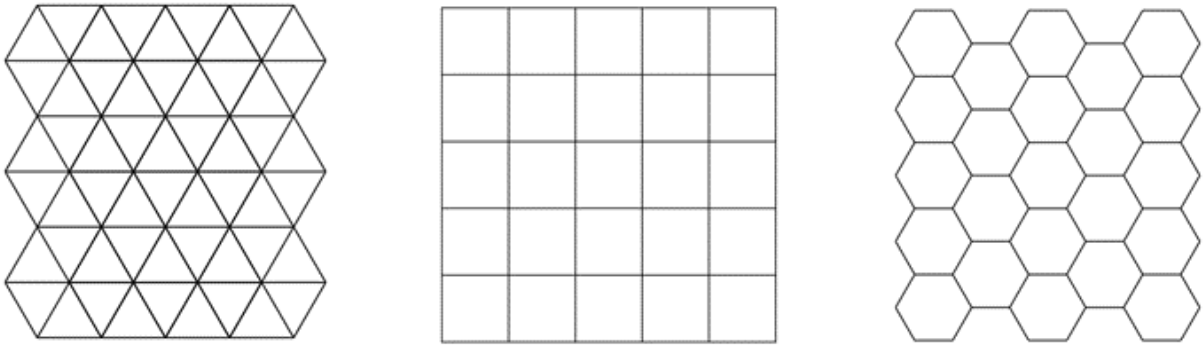


Figure 1. Tessellations of triangles, squares, and hexagons that can be used as the basis for cell shape in a raster model.

The size or extent of each cell indicates the resolution of the raster, and is given in linear units of distance (e.g., number of feet, meters, kilometers along one side of the cell) or in degrees or fractions of degrees of latitude and longitude (e.g., one arc second, or one-third arc second). The resolution of the raster is one component that dictates the memory storage requirements for it, with finer resolutions requiring more space in memory. The number of cells in a raster increases quadratically as resolution increases; a doubling of the resolution of a raster reduces the linear distance of a cell by one-half (e.g., moving from a 2 m cell to a 1 m cell), but increases the number of cells by a factor of four (doubling the number of cells in two directions). Figure 2 illustrates the effect of increased resolution on a digital elevation model (DEM).

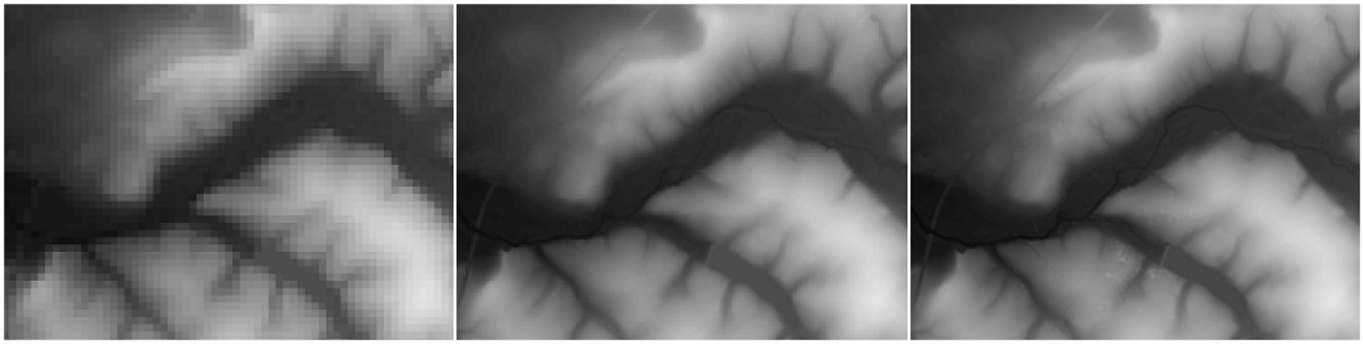


Figure 2. A digital elevation model (DEM) at 30 m (left), 10 m (center), and 3.3 m resolution (right).

In the most common case, when pixels represent a 2D area, pixels may be thought of as a "bin", where the pixel's value is a summary statistic (e.g., mean, median, standard deviation) of all values of the field within the bounds of the pixel. In contrast, pixels could represent the value at the grid's exact center; this is sometimes referred to as a grid (Briese, 2010) or lattice type raster (Wise, 2000).

Rasters are most commonly used to represent continuous data, as they allow for more efficient storage of values than an equivalent vector or point-based lattice system at the point densities generally required. This is because coordinates are stored implicitly as a

position in a data table rather than explicitly (as coordinates). However, rasters are often used to represent categorial (e.g., land use) or discrete data as well. In these cases, an area corresponding to the cell of the raster may be mixed with more than one category present in the pixel. For instance, a grid cell with a resolution of 100 meters may have both residential and industrial uses within it. There are several common strategies for dealing with the mixed pixel problem, including (a) majority-wins or winner-take-all, (b) using a separate category to specifically indicate a mixed pixel, (c) using the value nearest to the center of the cell, or (d) assigning a threshold percentage for a given category (e.g., if at least 25% of the area within the pixel is water, it will be recorded as water). Figure 3 illustrates the mixed pixel problem by overlaying relatively coarse (30 m) resolution land cover designations over a relatively fine resolution (1 m) digital surface model (DSM) showing buildings, roads, trees, and other fine-scale features from which the land cover designations are derived.

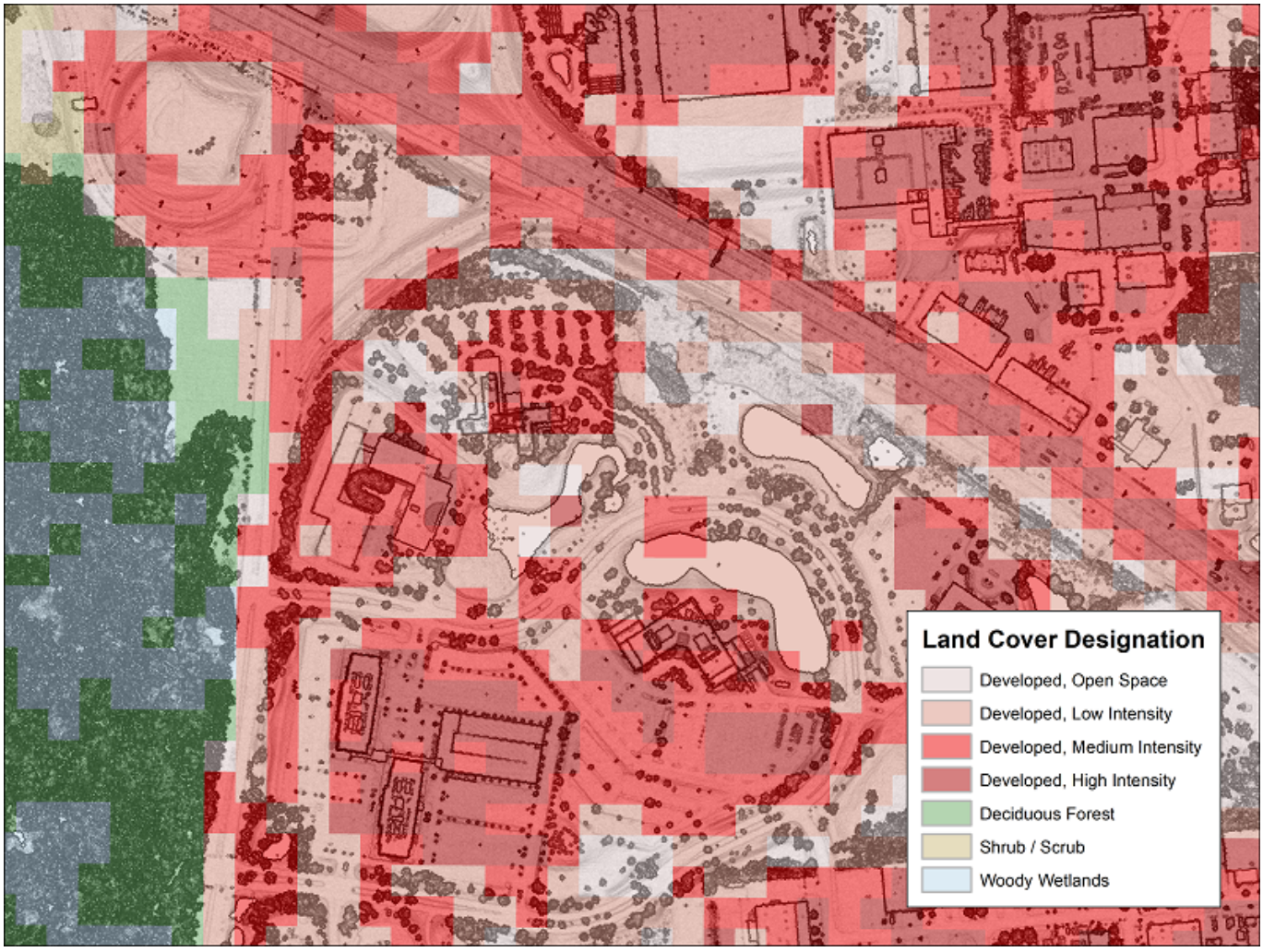


Figure 3. A National Land Cover Database (30 m) raster overlaid on a slope shaded, lidar-derived digital surface model (1 m), highlighting the mixed-pixel problem.

Rasters are most commonly distributed and used as either single, 2.5D surfaces with only one data value per cell (e.g., elevation) or as images with multiple bands. An example of the latter type is orthoimagery, with red, green, blue, near infrared, and potentially many other layers embedded in the same raster. In such cases, representation of the raster often involves linking each band to a red, green, or blue (RGB) visualization channel. Rasters may also extend the grid structure multidimensionally to form cubes (or voxels) or hyperspatial

equivalents, which can represent a volume of space, time, attribute space, or any combination of these. Index values may also be used in place of measurements to link to external attribute information held in an external database management system (DBMS) via lookup tables. Tomlin's (1990) work on the MAP model is an example of this.

3. **Examples**

The raster data model is widely used to encode GIS data. Examples include:

- Digital Elevation Models (DEMs) such as the ETOPO1 Global Relief Model.
- Elevation-derived visualizations and products (Kennelly, 2017).
- Remotely sensed data, including aerial and radar imagery.
- Meteorological variables such as temperature and rainfall, interpolated from point sources.
- Categorical rasters, such as the National Land Cover Database.
- The Gridded Population of the World dataset, which grids values originally recorded via administrative units.
- Digital scans of historical maps.
- Exported cartographic products, such as Digital Raster Graphic topographic maps.
- Cellular automata models such as SLEUTH (Chaudhuri and Clarke, 2013).

4. **Storage and File Formats**

Rasters record one or more values at each grid point. Grid values may be very simply recorded row-by-row staring at the upper left corner in the same manner that text is written in English. Alternative ordering systems may be used to improve efficiency, including row prime, Morton, or Peano-Hilbert methods. Memory demands for rasters are influenced by the type of data recorded (e.g., Boolean, integer, float, or string), the resolution and spatial extent of the raster, and any compression applied to the image. Compression can be used to exploit repetition or redundancy in the data to reduce overall storage demands, and many schemes are available accomplish this, including chain codes, block codes, run length codes and quadtrees.

Rasters may be stored in any number of formats or containers. They are typically stored in a binary format for the sake of efficiency, but plaintext formats are not uncommon, with ESRI's ArcInfo ASCII Grid perhaps the most widely used of this type. Binary raster formats are too numerous to exhaustively list, and the reader may wish to consult the Geospatial Data Abstraction Library (GDAL) project for a list of its 154 readable formats. Many formats originally designed for photographic images have been used as raster containers, including JPG, JPEG 2000, PNG, and TIFF; these have the advantage of easily moving raster data into and out of non-GIS data processing systems. Similarly, many raster formats have been designed specifically for geodata, including ESRI's ArcGRID, Erdas's Imagine format, and MrSID. NetCDF was designed by Unidata as a container for many types of array-based scientific data, and is commonly used for raster geodata in the atmospheric sciences where holding multiple layers of time-series coverage data in a single file is desirable.

Nearly all containers feature a mechanism for compression of data to reduce file sizes, but

vary in the degree to which they allow for lossless (vs. lossy) compression. In some cases, some reduction of data fidelity may not greatly diminish the value of the dataset; for example, digital orthophotos have been widely distributed in the JPEG container, even though some – often imperceptible – data loss will occur in the process of encoding the data for storage. In other cases (e.g., digital elevation models) perfect data fidelity is deemed more critical, and so these are often distributed in formats that feature lossless compression (e.g., GeoTIFF). Formats often differ in the types of data that they can hold. For instance, PNGs can hold only integer-based data with a maximum of four bands, while TIFFs can hold float data and a larger number of bands. The recent BigTIFF format extends the capabilities of the TIFF format, allowing for much larger (greater than 4 gigabyte) file sizes. Image-based containers have also proven popular because they interoperate more easily with non-GIS software. Proprietary storage formats (e.g., MrSID) often feature higher compression rates, but concerns with documentation, longevity, and interoperability have prompted the development of a number of efficient open / non-proprietary alternatives (e.g., JPG 2000).

Georeferencing information for raster images is embedded either directly in the header of the container file, or via ancillary files that are distributed with the raster. A world file is a common example of the latter approach, where a separate plaintext file describes the position in geographic space of the center of the upper-left pixel, as well as the x and y resolution of the raster. Compression of rasters is achieved via one of several approaches, including run-length (Holroyd and Bell, 1992), chain (Freeman, 1974; Žalik et al., 2015), block, quadtree (Finkel and Bentley, 1974; Mark and Lauzon, 1984; Samet, 1984, Martin, 1992), and wavelet methods.

5. **Advantages and Disadvantages of the Raster Data Model**

The raster data model is often contrasted with the vector data model. Both are highly useful, and the choice of which model works better is entirely task dependent. Raster data models excel in cases where the underlying data itself is continuous in nature, as there are significant gains in efficiency of storage and indexing owing to the regular spatial pattern of the grid. The regularized pattern similarly speeds arithmetic calculation times between multiple raster layers, and reduces the required time for operations such as spatial interpolation of missing values, or for resampling. Because raster layers can be interpreted as binary or grayscale images, many operations originally designed for computer vision can be easily applied to problems of classification or machine learning.

Rasters have several disadvantages as well. Many containers support only a single level of resolution, although the construction of ancillary image pyramids can reduce the impact of this on visualization. Many newer raster storage formats, including MrSID and JPEG 2000 feature multi-resolution data storage natively which can make level-of-detail analysis processing possible in software that supports it. Another limitation of raster formats is that reprojection and/or resampling typically results in varying degrees of data degradation. For this reason, it is advisable that rasters are projected and/or resampled from their original source rather than serially. Rasters with coarse resolutions relative to the objects or attributes they represent can cause a “blocky” appearance to the data. Increasing resolution can not fully solve this problem for both practical and theoretical reasons, as increased resolution puts heavier demands on memory, storage, and processing power

(Fisher, 1997). Further, the approach will nearly always yield smaller, but still-mixed pixels along the object edge leading to the same “blocky” appearance if the user zooms in far enough.

Raster cells are typically considered topologically connected to their neighbors to the right, left, top and bottom, but may also be connected to neighbors on the diagonal (or further away) as well. Topological connections may be quite important when rasters are used in modeling, such as least-cost path determination. In this context, the conversion of vector features (such as lines representing a road network) may introduce errors in a raster-based analysis if either extra grid cells are improperly included, or grid cells are assigned improper values during the conversion.

References

Briese, C. (2010). Extraction of Digital Terrain Models. In G. Vosselman and H-G. Maas (Eds.), Airborne and Terrestrial Laser Scanning (pp. 135-167). Dunbeath, Scotland: Whittles Publishing.

Chaudhuri, G., & Clarke, K. (2013). The SLEUTH land use change model: A review. Environmental Resources Research, 1(1), 88-105.

Finkel, R. A., & Bentley, J. L. (1974). Quad trees a data structure for retrieval on composite keys. Acta Informatica, 4(1), 1-9.

Fisher, P. (1997). The pixel: A snare and a delusion. International Journal of Remote Sensing, 18(3), 679-685.

Freeman, H. (1974). Computer processing of line-drawing images. Computing Surveys, 6, 54-97.

Goodchild, M. F. (1992). Geographical Data Modeling. Computers & Geosciences, 18(4), 401-408.

Holroyd, F. & Bell, S. B. M. (1992). Raster GIS: Models of Raster Encoding. Computers & Geosciences, 18(4), 419-426.

Kennelly, P. J. (2017). Terrain Representation. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2017 Edition), John P. Wilson (ed.).

Maguire, D. J. (1992). The raster GIS design model – A profile of ERDAS. Computers & Geosciences, 18(4), 463-470.

Mark, D. M., & Lauzon, J. P. (1984). Linear quadtrees for Geographic Information Systems. In Proceedings of IGU Symposium on Spatial Data Handling, 20-24 August, Zurich, pp. 412-431.

Martin, J. J. (1992). Organization of geographical data with quadtrees and least squares approximation. In Proceedings of the Symposium on Pattern Recognition and Image Processing (PRIP), Las Vegas, Nevada. IEEE Computer Society, pp. 458-465.

Peucker T. K., & Chrisman, N. (1975). Cartographic Data Structures. The American Cartographer, 2(1), 55-69.

Peuquet, D. J. (1984). A Conceptual Framework and Comparison of Spatial Data Models. Cartographica 21(4), 66-113.

Samet, H. (1984). The Quadtree and Related Hierarchical Data Structures. ACM Computing Surveys (CSUR), 16(2): 187-260.

Tomlin, C. D. (1990). Geographic Information Systems and Cartographic Modeling,1st Edition. Englewood Cliffs, New Jersey: Prentice-Hall (now Pearson Education).

Wise, S. (2000). GIS data modelling-lessons from the analysis of DTMs. International Journal of Geographical Information Science, 14(4), 313-318.

Zalik, B., Mongus, D., & Lukač, L. N. (2015). A universal chain code compression method. Journal of Visual Communication and Image Representation, 29, 8-15.