

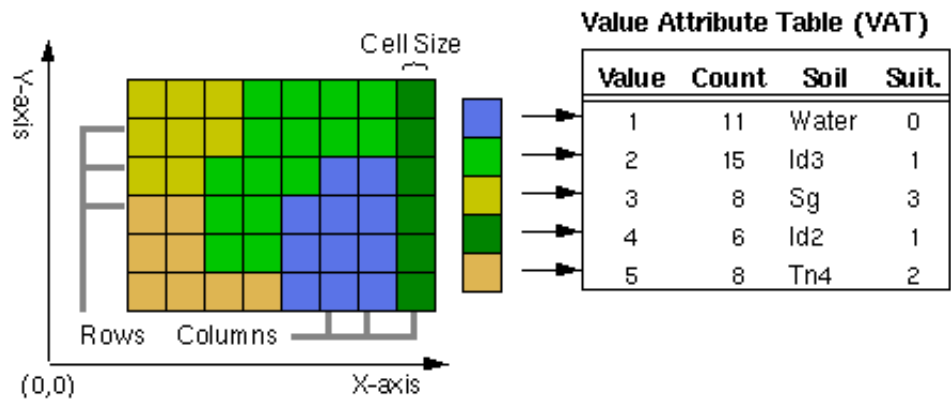
Esri Grid format

Resource Center » Professional Library » Data Management » Geographic data types » Rasters and images » Supported raster data

A grid is a raster data storage format native to Esri. There are two types of grids: integer and floating point. Use integer grids to represent discrete data and floating-point grids to represent continuous data.

Learn more about discrete and continuous data

Attributes for an integer grid are stored in a value attribute table (VAT). A VAT has one record for each unique value in the grid. The record stores the unique value (VALUE is an integer that represents a particular class or grouping of cells) and the number of cells (COUNT) in the grid represented by that value. For example, if 50 cells have a value of 1 representing a forest, then the VAT would show a VALUE = 1 and COUNT = 50 for each of the 50 cells.



Learn more about raster dataset attribute tables

Floating-point grids do not have a VAT because the cells in the grid can assume any value within a given range of values. The cells in this type of grid do not fall neatly into discrete categories. The cell value itself is the attribute that describes the location. For example, in a grid that represents elevation data in meters above sea level, a cell with a value of 10.1662 indicates that the location is about 10 meters above sea level.

The range of data values that can be stored as grid values are as follows:

- Floating-point grids can store values from -3.4×10^{38} to 3.4×10^{38} .
- Integer grids can store values from -2147483648 to 2147483647 (-2^{31} to $2^{31}-1$).

For integer grids, this information applies only to the VALUE item. An integer grid may have other INFO items added to its VAT whose range of values depends on the item definition.

The coordinate system of a grid is the same as that of other geographic data. The rows and columns are parallel to the x- and y-axes of the coordinate system. Since each cell within a grid has the same dimension as other cells, the location and area covered by any cell is easily determined by its row and column. The coordinate system of a grid is thus defined by the cell size, the number of rows and columns, and the x,y coordinate of the upper left corner. Grids also carry additional information, such as the coordinate system associated with the grid.

Caution: As with most formats, a grid should not be named with spaces or any other special characters in its name. A multiple-band grid cannot have more than 9 characters in its file name, and a single-band raster dataset cannot have more than 13 characters.

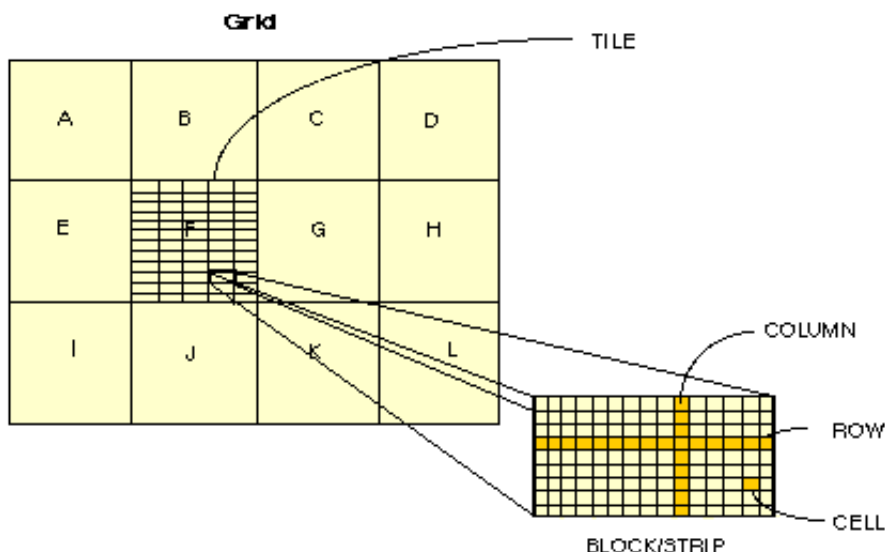
Grid data structure

Grids are implemented using a tiled raster data structure in which the basic unit of data storage is a rectangular block of cells. Blocks are stored on disk in compressed form in a variable-length file structure referred to as a tile. Each block is stored as one variable-length record.

The size of the tile for a grid is based on the number of rows and columns in the grid at the time of creation. The upper limit on the size of a tile is set by the application and is very large (currently set at 4,000,000 x 4,000,000 cells). As a result, most grids used for GIS applications are automatically stored in a single tile. The spatial data for a grid is automatically split across multiple tiles if the size of the grid at the time of creation is larger than the upper limit for the size of a tile.

The blocked storage organization for grids supports both sequential and random spatial access to large raster datasets. The blocking structure imposes no restrictions on the joint analysis of grids. Tiles and blocks from different grids also need not coincide in map space for joint analysis. The tile and block structure of a grid is completely hidden from the user, who always creates and manipulates a grid as though it were a seamless raster of uniformly square cells.

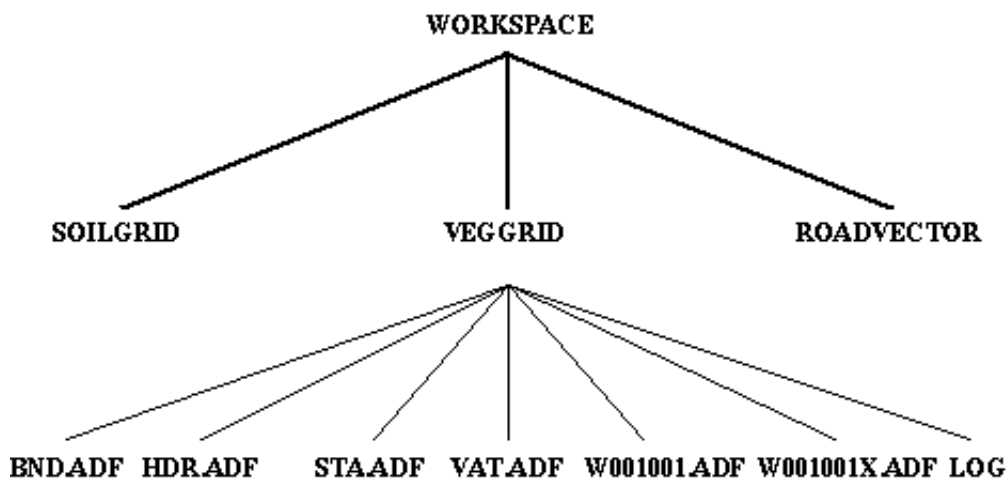
Grids use a run-length raster compression scheme that is adaptive at the block level. Each block is tested to determine the depth (bits per cell) to be used for the block and to determine which storage technique (cell by cell or run length coded) is more efficient. The block is stored in the format that requires less disk space. The adaptive compression scheme is the optimal choice because of its ability to efficiently represent both homogeneous categorical data and heterogeneous continuous data while supporting joint analysis using both types of data. Single layer per-cell operations, such as data reclassification, operate directly on runs of data without decompression. Multilayer per-cell operations on compressed input layers intersect runs of data from the different layers and operate on the intersected runs. Single-layer per-neighborhood operations and multilayer per-cell operations that mix compressed and uncompressed data expand runs into cells and perform traditional cell-by-cell processing transparently.



The tile-block structure of a grid is also transparent to any application programs that access the spatial data in a grid. Programs that manipulate grids access the spatial data by setting a rectangular window defined in map coordinates.

Grid data storage

A grid is stored in an ArcInfo workspace. The grid, like a coverage, is stored as a separate directory with associated tables and files that contain specific information about the grid. In an integer grid directory (originally created by ArcInfo Workstation), the following tables and files are found: the BND table, which stores the boundary of the grid; the HDR file, which stores specific information describing the grid, for example, cell resolution and blocking factor; the STA table, which contains statistics for the grid; the VAT table, which stores the attribute data associated with the zones of the grid; the log file (LOG), which monitors the activity that has occurred on the grid; and the tile file w001001.adf (q0x1y1), which stores the cell data and the accompanying index file w001001x.adf (q0x1y1x) that indexes the blocks in the tile and the LOG. (Some of these, such as the log file, may not exist if created using ArcGIS operators.)



If a grid is altered, the values and information contained in the files and tables are updated immediately. The information contained in an INFO table is accessible to the user and provides information about the grid.

BND table

A grid BND contains the boundary of the grid. The boundary is a rectangle that encompasses the cells of a grid; it is stored in map coordinates. All grid BNDs are stored in double precision.

The minimum coordinates in the BND table are for the lower left corner of the lower left cell in the grid. The maximum coordinates are for the upper right corner of the upper right cell in the grid.

HDR file

The HDR is a binary file. Information stored in the file includes the cell size, type of grid (integer or floating point), compression technique, blocking factor, and tile information.

STA table

The STA table is an INFO table that contains statistical data about a grid. The minimum, maximum, mean, and standard deviation for the grid are stored as floating-point values in the STA table. You should not attempt to alter these values directly.

Because NoData represents an unknown value, NoData is not used in calculating the statistics in the STA table.

When a bilevel grid (containing only 0 and 1 values) is created, the STA table contains the value 0 for the mean and -1 for standard deviation. The standard deviation value -1 indicates that statistics have not been calculated for a grid.

A standard deviation value of -2 indicates that the grid contains only NoData cells.

VAT table

The VAT is an INFO table that stores attributes associated with the zones of a grid. Only integer grids have a VAT associated with them. Every VAT has at least two items, VALUE and COUNT. The VALUE item contains integer values that are used to distinguish the characteristics of one location from the other locations in a grid. All cells that are assigned the same value contain the same characteristics and, therefore, belong to the same zone. COUNT is the number of cells in a zone.

New items can be added to the VAT. The VALUE and COUNT items should not be changed, and the VAT must be kept sorted on the VALUE item. Never add new items before VALUE or COUNT.

Cells containing NoData are not represented in the VAT.

Below is an example of a VAT:

Record	VALUE	COUNT
1	0	628872
2	1	265043
3	2	151150
4	3	3185652
5	4	79983
6	5	4782
7	6	74334
8	7	8877
9	8	1817
10	9	491
11	10	858
12	11	8770
13	12	28789
14	13	72539
15	14	3686
16	15	3932
17	16	13227
18	17	1890
19	18	1305
20	19	427286
21	20	6695

Tile files

The w001001.adf (q0x1y1) and w001001x.adf (q0x1y1x) files store the data and the index for the first, or base tile, in a grid. The upper limit on the size of a tile is very large, and most grids are stored using a single tile. If additional tiles are used, they are automatically numbered based on their spatial relationship to the first tile. Tiles are implemented as variable-length binary files. With versions prior to ARC/INFO 7.x, these files were named q0x1y1 and q0x1y1x and still work with the current software.

LOG file

The LOG file is an ASCII file that contains information about the creation of and alterations to a grid. The LOG monitors the actions performed on the grid, but it does not contain every action performed with the grid. Since all Grid functions result in a new grid, only Grid commands, such

as RENAME and COPY, can alter an existing grid and be entered into the LOG file. The LOG file can be accessed, like all ASCII files, through system commands or any text editor.

Storage limitations

The name of a grid is limited as follows:

- It cannot be stored using spaces.
- It cannot start with a number.
- It cannot be longer than 13 characters (a multiband grid is allowed up to 9 characters).

There is a limit to the number of files that can be stored in an INFO directory for both coverages and grids. This total is approximately 10,000. Therefore, this limits the number of grids you can store in a workspace. For example, the following lists the theoretical maximum number of grid datasets that can be stored in a single workspace directory:

- Fewer than 5,000 floating point grids, or
- Fewer than 3,333 integer grids, with VATs (fewer than 5,000 if no VATs), or
- Fewer than 10,000 grid stacks

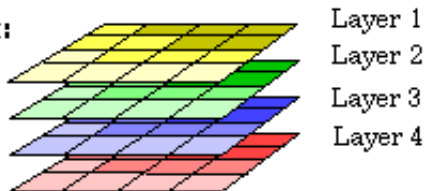
The preceding numbers are the theoretical maximums. If you have a process that will create interim grids (and therefore files in the INFO directory) these numbers will be less. Additionally, if you are storing a mix of files, such as grids and coverages, you will store fewer.

These numbers relate to the number of files in the grid folder that store information in the INFO directory. The limit is 10,000 (9,999), but it's not the total number of files in an INFO directory, it's the number of files pointing to the files in the INFO directory. For each grid, there are two files in the grid's folder pointing to files in the INFO folder: the BND (boundary) files, and STA table (statistics) files ($9999/2 \approx 5000$). When a grid has a VAT, this also points to files in the INFO directory, so the number that can be stored is reduced again ($9999/3 \approx 3333$). A grid stack only has a single file which points to the INFO directory ($9999/1 \approx 9999$).

Stack

A stack consists of an ordered set of spatially overlapping grids (layers) treated as a single entity for multivariate analysis. Cluster analysis, classification, and principal component analysis all work on the layers in a stack.

A Stack:

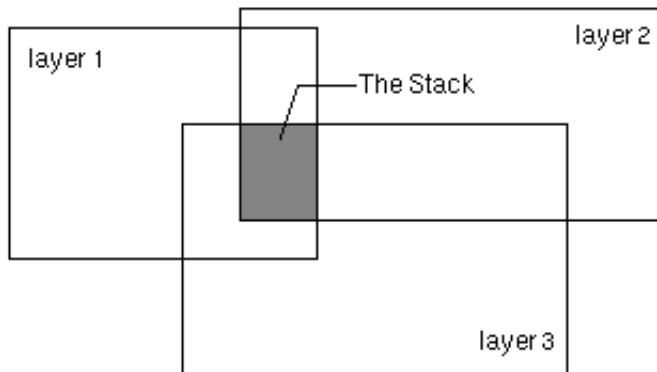


A stack has the following characteristics:

- A set of layers with each layer corresponding to a grid
- A map extent, or BND
- A cell size
- A data type
- A projection

Each layer specified in a stack has an index number indicating its order in the stack. The grids that make up a stack must be in the same workspace.

The boundaries of the input layers can overlap exactly, partially, or not at all, but only the area where layers overlap comprises the stack. The stack's BND is where the boundaries of its layers intersect. The computations of a multivariate analysis function occur on the overlapping area. If there is no common area between the input layers, the stack is empty and no computations occur.



The cell size of a stack defaults to the coarsest layer in the stack.

You can combine any number of data types (real or integer) of the input grids in a stack. However, before applying a multivariate technique, you should be aware of what the values represent, whether categorical or continuous data, and the range or relative range of the values. In certain analyses, the input data type of the stack determines the data type of the output.

Projection information associated with the input grids is stored with the stack. Since a stack is treated as a single entity, all grids in a stack must be in the same projection. The projection information is used to ensure that each grid of the stack occupies the same geographic area.

Storing a grid stack

A stack is stored in a directory structure similar to a grid or coverage. There are two files in the stack directory: an external INFO STK table and an ASCII PRJ file. The actual grids that comprise the stack are not stored in the stack. They are ordinary grids in your workspace. That means any grid can be used in more than one stack. The STK table stores the names of the grids that comprise the stack and their corresponding index values:

```
GRID: LIST JER135.STK
Record  INDEX  GRID
  1      1     jer1
  2      2     jer3
  3      3     jer5
```

The INDEX item gives the position of a grid in the stack, while the GRID item lists the grid names that comprise the stack. The spatial data of the input grids is not duplicated in the stack. As a result, the stack always reflects the latest version of the input grids. The STK file is as accessible as any other INFO file. You can add items for descriptive purposes, such as an item for storing the date that the data was collected, but don't use INFO to alter the values in the INDEX item or names in the GRID item. All manipulations to these items should only be performed using a variety of the stack management commands available in Grid.

The PRJ file, when present, stores the projection information of the stack:

```
Projection STATEPLANE
Zone      4701
Datum     NAD27
Zunits    NO
Units     FEET
Spheroid  CLARKE1866
```

```
Xshift    0.0000000000
Yshift    0.0000000000
Parameters
```

If the projection is unknown for all input grids in the stack, no PRJ file is created.

The name of a grid stack cannot be stored using spaces, cannot start with a number, and cannot be longer than 9 characters.

NoData in a grid

Every cell in a grid has a value assigned to it; however, cells without actual values can be assigned NoData on the grid representing that theme. NoData and 0 (zero) are not the same; 0 is a valid value. For this reason, NoData cells cannot be used in calculating the statistics in a grid's STA table.

Learn more about NoData

Copyright © 1995-2013 Esri. All rights reserved.

8/21/2013

URL: <http://help.arcgis.com/en/arcgisdesktop/10.0/help/009t/009t0000000w000000.htm>