Talking about Rasters in ArcGIS

Transcript

Welcome to our ESRI Instructional Series podcast entitled *Talking about Rasters in ArcGIS*. I am Colin Childs from Educational Services at ESRI in Redlands, California. I teach a host of ArcGIS courses that range from geodatabase design classes to those that teach GIS analysis and modeling. Today I will be exploring some broad topics and concepts related to how we deal with rasters in ArcGIS. I hope to share with you some of my knowledge about the raster formats, cell size, statistics, and pyramiding. This discussion is tailored to users of the ArcGIS software who display, store, and analyze raster data and who'd like to check their knowledge and perhaps glean some new insights into dealing with rasters.

So let us begin this podcast with a brief look at the variety of different rasters that you as an ArcGIS user may encounter and deal with. The rasters you work with may take the form of geographic data for example, rasters that represent a specific geographic extent; it might be a DEM containing elevation information or an author rectified photo or satellite imagery. Or, perhaps you raster may take the form of pictures or graphics that are associated with a geographic object. Maybe they make up an attribute of a feature, for example, an attribute of some vector data such as a picture of a pipe joint or a fire hydrant or a telephone pole. Now each of these rasters and the different rasters' types, of course, may have a number of properties that describe them. Some of these properties may include things like the resolution or the cell size; the number of rows and columns, which define the size; the spatial extent; the number of bands, single, multiband, or hyperspectral; perhaps pixel resampling; in other words whether pyramiding is present or not; maybe the pixel type, which refers to the type of values, integer or a floating point that the raster can store; maybe a little information about pixel or bit depth, which defines the range of values that the raster may store; whether a compression is applied or not, whether statistics are present; and whether a special reference may be present for the raster.

Now I'm going to spend the next few minutes describing some of these properties and, in future podcasts, I'm hoping to explain in greater detail more about several of these properties that you are most likely to set and work with and decisions that you might make with respect to these properties. You can check out a raster's properties in ArcCatalog by right-clicking on the raster in the Catalog tree and then choosing the Properties option. And that will give you the option of listing out all these properties. I guess one of the first things you'd be interested in knowing is

what raster formats are supported within the ArcGIS product. Well these formats or the different raster formats range from a host of formats that you have the option of reading and writing to some of the formats that are read-only and, of course, a whole collection of military formats that are essentially read-only as well.

Some of the read/ write formats that you will encounter is firstly the ESRI Grid or Grid stack, which is the native format of raster in ArcGIS. You also have the option, of course, of writing ERDAS IMAGINE (IMG files), TIF, and GeoTIF files. You have the option of also writing ArcSDE personal and file-based geodatabase rasters in the form of raster datasets. You have the option of writing JPEG 2000, JPEG, bitmap, GIF, and PNG files.

Some of the read-only formats that you will encounter are ECW, MrSID, the USGS DEM format, BIL, BIP, and BSQ format, ERDAS 7.5's LAN, GIS, and RAW formats, the ER Mapper ERS format, Intergraph CIT and COT formats, Idrisi’s .RST format, the PCIDSK PIX format, MAT format, XPM format from X-Pixmap; and of course new to us is the option of reading HDF4 formats. Some of the military formats that you may encounter are DTED levels 1 and 2, ADRG, CADRG, CIB, NITF, and DIGEST formats.

Now of course many of these raster formats do support compression and the compression applied to them may be in the form of lossless compression or lossee compression. In a lossless compression of course all the original values are maintained in the resulted compressed dataset, but the compression ratios may vary generally around two to one. In a lossee compression where you the user who is applying this compression, you can choose how much you want to lose. You may apply a compression to a JPEG, which would give you between 5 to 1 and 15 to 1 compression, or you could apply a 20 to 1 or 50 to 1 compression to say JPEG 2000, MrSID or ECW format. The loss, of course, will be greater with the higher compression ratios, and with a lossee compression, remember your original values are all going to be changed by the compression algorithm it deployed and you won't be able to return the exact values you started off with if you ever decide to uncompress the compressed raster.

While working with rasters in ArcGIS, you'll notice us use some terminology; you may have heard the term “raster dataset” or “raster catalog”. In ArcGIS, the term *raster dataset* refers to any supported raster that represents an individual set of pixels consisting of, say, one or more bands covering the same area. A *raster catalog*, however, refers to a collection of raster datasets that are organized as a feature class. So effectively each row of the feature class represents a different raster dataset that may be participating within the raster catalog. Another important aspect of what you may want to look at is the number of bands that a raster may have. Some rasters are single band or they may measure only a single characteristic, for example soil type or vegetation or elevation. Rasters may of course also be multiband where they're representing and recording several different values per cell or pixel. Typical examples of single-band rasters are elevation data, grayscale aerial imagery, scanned maps and documents, and maybe black and white TIFs which are one-bit data. Typical examples of multiband rasters are remotely sensed data, or color aerial photography, or maybe data from Thematic Mapper, or SPOT Image, or Space Imaging.

Another important characteristic for you to be aware of is cell size. The cell size defines or determines the level of detail of features or phenomena that are represented by the raster and often your cell size is a very important criteria because the smaller the cell size the larger the raster might be, but of course the smaller the cell size the more detail you would have. Smaller cell sizes will give you greater storage space and also result in longer processing time.

Now let's take a look at pixel type and pixel depth, which you may recall is a property I mentioned earlier as well. Pixel type refers to the type of values that are stored in the raster, whether they are integer or floating point, and whether they are signed or unsigned. This is an important criteria for us by the way, pixel type, because it determines what kind of analysis we may be able to perform on the rasters if we were going to use them in modeling and analysis. Pixel depth is another interesting criteria and this defines the range of values that a particular raster can store. For example, the raster with a pixel depth of one bit can really only store values in the cells that are zeros or ones, where as a two-bit raster may store zeros and up to three, so your cell values a zero, one, two, or three. In a four-bit raster, or four-bit pixel depth, raster values can range between zero and fifteen. Now here's an interesting one for you. In an eight-bit

raster that is unsigned your cell values may range between zero to 255, but in an eight-bit signed raster the values may range only between negative 128 up to positive 127. Another important criteria you need to think about when working with data or raster data is the concept of NoData and NoData is the option where within the ArcGIS product you can isolate specific cells from being processed by assigning them a NoData value. We'll talk more about these in later podcasts as well, by the way.

Now ArcGIS also maintains statistics for rasters and the statistics are used when applying contrast stretching or classifying the raster when it's being rendered. If it is not able to be maintained in the header of your raster, it will be maintained in what we call an AUX or auxiliary file. The AUX file can also store things like color maps, pointers to pyramid files, coordinate system information, and transformation information if need be. So to improve display performance on rasters, ArcGIS offers you the option to pyramid a raster. And pyramiding is very essential to generating rasters that can be displayed at high speed. Pyramids are typically going to take up around 33 percent more storage for the raster, but the improved display performance will far outweigh the storage needs that you may have with respect to the raster. It's also interesting to note that if you build pyramids on raster datasets that are highly compressed the pyramid files can actually be larger than the raster dataset on disk.

As you've heard ArcGIS supports a host of different raster formats. These differ in what pixel type and pixel depths are used for the cell values. You've also learned a little more about the differences between raster datasets and raster catalogs and we've touched on some of the additional properties such as the number of bands, statistics, NoData, and raster pyramiding. For further resources, please check out our instructor-led training courses at www.esri.com. This discussion has also touched on topics that are covered in the *Working with ArcGIS Spatial Analyst* instructor-led training class and in the Virtual Campus class entitled *Working with Rasters in ArcGIS 9*.

I want to thank you for tuning into this session of our ESRI Instructional Podcast series. Stay tuned for future broadcasts where we will explore many more of these topics discussed in this presentation in greater detail.