**The Forest Inventory and Analysis: BIGMAP User Documentation in ESRI products**

**October 16th, 2023**

**Contents**

**Preface**

**Abstract**

**Keywords**

**Authors**

**Background**

In 1928, Congress enacted the McSweeny-McNary Forest Research Act which states that, *“… a comprehensive survey of the present and prospective requirements for timber and other forest products of the United States…*” Thus, began the Forest Inventory and Analysis (FIA) program, which collects Forest health parameters such as condition, volume, growth, and use of trees on the Nations Forest lands. FIA publishes yearly summary reports of the lands they survey, which benefits forest management and policies across the United States.

These reports are constructed with the same information that is publicly available and can be found on the Forest Inventory and Analysis Database (FIADB). Most inventories can be located using this database, some historical information might not be available due to the changes in inventory sampling. As a widespread reach for national consistency, all annual inventories use a common plot design and common data collection procedures. These summary reports, written by Forest Service Personal can be found on the geospatial showcase hub.

The Big Data Mapping and Analytics Platform (BIGMAP) is FIA’s could-based national scale modeling, mapping, and analysis environment for US Forests. The foundation of BIGMAP was established with the use of moderate spatial resolution imagery from MODIS to produce Continuous United States (CONUS) raster data sets on numerous forest attributes. BIGMAP intends to provide users with simple ways to estimate forest attributes (and associated uncertainties), eliminating the need to understand how FIA stores and processes information. BIGMAP allows access to data services that are derived from the integration of FIA data and ancillary predictors are based upon peer-reviewed modeling approaches, these models can be constructed to simultaneously protect plot security and provide reliable estimates of various forest attributes across a range of spatial scales. Lastly, individual forest attributes can be combined easily with other content to provide integrative assessments (developed by the Agency as well as its partners) relevant to resource managers, policymakers, and the interested publics to create associated maps. This report should be used to navigate the geospatial showcase to locate BIGMAP layers and utilize them for further analysis.

**Acknowledgements**

-Partnership with ESRI

**Chapter 1: Introduction**

* 1. **Purpose of This Guide**

This guide is user guide to the Big Data Mapping and Analytics Platform (BIGMAP) developed through a partnership between the USFS and ESRI. This guide is the first version of documentation of BIGMAP layers and will be revised and updated as more layers are being developed. Layers developed through BIGMAP are intended to be used by those outside of FIA who are interested in FIA for independent analyses. It is important for users of this data to understand how the layers were derived from the collection of the data through FIA field plots and the satellite imagery that were based on. This guide is meant to be a source of information to aid users through their personal processing of BIGMAP Aboveground Biomass layers.

* 1. **Early Stages of BIGMAP**

Early stages of BigMAP

* + CONUS ([Characterization and visualization of the accuracy of FIA's CONUS-wide tree species datasets | US Forest Service Research and Development (usda.gov)](https://www.fs.usda.gov/research/treesearch/46519)

**Chapter 2: Raster format -Layer Catalog**

Raster images are comprised of pixels each one filled with information that when combined with other pixels create an image. The information that is collected from each pixel is information regarding the color and where it falls on the visible color spectrum, which can be manipulated to and interpretated to match real-time environmental conditions. This is the framework for remote sensing analytics and combined with data collected in the field many things can be harnessed.

BIGMAP uses CONUS Landsat images from 2014-2018 and 212,978 FIA plots using harmonic regression to associate each pixel with the most similar FIA plots. This multivariate association approach is known as “k nearest neighbors,” using XXXX. Models were created for 36 CONUS ecoregions and then mosaicked together using spatial indexing from EMAP hexagons as the sampling design for FIA utilizes a standardized hexagonal grid.

Parallel processing outputs raster imagery which are packaged into a single mosaic dataset, containing 327 species at a 30m pixel resolution for the Aboveground Biomass layers in acres ton, which are then stored on an image service.

Image services exist for public access of the data. To use an image service, first connect to the GIS server and navigate to your source data. Source data for BIGMAP can be found on the [FIA Geospatial Showcase](https://fia-usfs.hub.arcgis.com/) page under FIA BIGMAP General Layer Catalog.

**Chapter 3: Mosaic Dataset**

Mosaic datasets are used to manage collections of raster images stored in the geodatabase. The translation of raster imagery to mosaic dataset allows you to store, manage, view, and query your raster imagery quickly and efficiently. Their advanced raster querying capabilities and processing functions are a result of the referencing of the raster imagery rather than manipulating the source imagery.

A mosaic dataset is created and stored in the geodatabase through the geoprocessing toolbox. At first it is created empty where then the user can add an unlimited amount of raster imagery to the dataset. The images added to a mosaic dataset do not need to be continuous and can be pieced together so long as the coordinate system is the same.

It is best to store the image service layers from the BIGMAP catalog in mosaic datasets because of the ease at which the information can be cataloged and processed. The design at which mosaic datasets can handle varying resolutions all spectrally, spatially, temporarily, and radiometrically. As the mosaic datasets only reference the raster imagery while preforming functions the source pixels are never altered or converted, and the metadata remains intact.

Just like normal spatial layers mosaic datasets have attribute tables that are related to spatial data layers. The attributes associated with mosaic datasets include:

* **ObjectID**
* **Raster**
* **Name**
* **MinPS and MaxPS**
* **LowPS and HighPS**
* **Category**
* **Tag**
* **GroupName**
* **ProductName**
* **CenterX and Century**
* **ZOrder**

The three layers that are created through a mosaic dataset include: boundary, footprint, and image layer. The boundary layer displays the full extent of all the raster imagery added to the mosaic dataset. The footprint layer reflects the extent for each individual raster image added to the mosaic dataset. Each footprint box corresponds to a unique row in the attribute table that can be used to query the data. The image layer is what controls the rendering of the data which can be manipulated through the stretch, RGB band combination, resampling, mosaic method, and other properties.

The seamline layer is another layer that can be added that acts similarly to the footprint layer in mosaicking images together. The functionality of this layer ensures that there is minimal overlap between all the raster images once added to the mosaic dataset. The use of seamlines is recommended when you do not need to view individual mosaic dataset images separately and would like to keep one feature selected throughout the mosaic dataset.

**Chapter 4: Accessing BIGMAP layers**

Graphical user interface, text, application, email

Description automatically generatedBIGMAP species layers are stored on an image service that can be located by through the [USFS geospatial showcase.](https://fia-usfs.hub.arcgis.com/) There are two ways you can process the data: 1) via ArcGIS online 2) ArcGis pro. Access for the already processed and created mosaic dataset can be found at: [FIA\_BIGMAP\_2018\_Species\_Aboveground\_Biomass - Overview (arcgis.com)](https://usfs.maps.arcgis.com/home/item.html?id=1e75650439c84a27ab9ff1f1c169f10b)

**Chapter 5: Viewing and manipulating BIGMAP data in AGOL**

**Section 5.1 Accessing**

**Map viewer classic**

1. Click on the drop down arrow in “Open in Map Viewer”, and open in the species ABG layer in, “Open in Map Viewer Classic”
2. Once open zoom into extent of the United States
3. In the layers pane, under the species layer right click in the eclipses, … for a drop down menu to appear
4. In the drop down menu click on, “ Image Display”
5. In the Image display window, change the “renderer display” to your species of interest
6. Hit apply and your species of interest will be displayed
7. If you are interested in changing the Color Ramp, change:
   1. Symbology type: Classify
   2. Field: Value
   3. Method: Natural Breaks
   4. Color Ramp: [Color of choice]
8. Hit apply and your species of interest and color will be displayed

**5.2 Creating an area of interest**

**Chapter 6: Viewing and manipulating BIGMAP data in ArcGIS Pro**

**6.1 Accessing**

1. In scrolling to the bottom of the page, copy the image service url
2. In ArcGIS Pro, open a new map window
3. Upload the species data using the URL from the image service: Navigate to Map ribbon >> Layer >> Add Data >> Data From Path >> [image service url](https://di-usfsdata.img.arcgis.com/arcgis/rest/services/FIA_BIGMAP_2018_Species_Aboveground_Biomass/ImageServer)
4. To filter layer by species:
   1. Processing Template
      1. In the Data Ribbon>> choose processing template>> species of choice>> click on species of interest
   2. Definition queries
5. To change symbology, right click on the layer in the contents page, scroll down to symbology
6. In the symbology pane change the “Color Scheme” to the desired color ramp

**6.2 Creating an area of interest**

**Chapter 7: Use Cases – Choose 2 or 3**

**Chapter 9: Raster functions**

**Input: - FIA\_BIGMAP\_2018\_Species\_Aboveground\_Biomass**

**-**

|  |  |  |  |
| --- | --- | --- | --- |
| **Function** | **Additional inputs** | **Description** | **Output** |
| **Zonal Statistics** |  |  |  |
|  |  |  |  |

1. Calculate Cell pixel statistics

Zonal statistics as a table

1. Tabulate area
2. Different graphs

**Chapter 10: Validation data**

Packaged alongside the BIGMAP layers is an additional layer of validation data, It’s the result of a comparative accuracy assessment between the modeled BIGMAP estimates and the estimates one would get from the FIA plot data alone. There are three validation layers, each calculated at a different spatial scale and depicting the level of agreement between the two datasets at that scale. Assessment scales are defined by hexes or aggregations of hexes, ranging from Xkm2 (one hex) to Xkm2 (49 hexes). These layers are possible because of the extensive FIA plot dataset with its sampling intensity of one plot for every 6,000 acres (24.29 km2) of forested land.

To be included in the validation assessment, hexagons, and hex aggregations are required to have enough FIA plots to ensure that the summed scaling factor (typically 6000 acres x the number of forested plots) accounts for at least 10% of the assessment area. To be included in the validation set, there needs to be three plots with a specific observation recorded bringing the expansion factor above 64.8 km2. For example: If one Dogwood tree is found on 2 plots in a 3 hexagon range, then you would need to scale this to account for 10% of the total inventory. Another requirement is that hex aggregations be of a similar area to maintain consistency within a single scale. As a result, along the coast and borders one can see where hex aggregations are missing from the map because they did not contain enough land area to qualify. Specifically, for level 2 aggregations, there needs to be a minimum of 6 of the 7 component hexagons present with observations on them and level 3 required a minimum of 37 of the 49 hexagons with observations.

Validation results are presented in three visual ways at each scale:

1. As a map depicting how the BIGMAP estimates compare with not just the FIA estimate, but in relationship to our confidence in the FIA estimate. For example, in a highly variable area, or in an area where we have fewer FIA plots to capture that variability, we know that we have less confidence in that FIA estimate, and this is expressed as a larger “confidence interval” around that estimate. The map depicts where the BIGMAP estimate for a species is within the 90% confidence interval of the FIA plot estimate, or if its higher or lower than the FIA plot estimate and how much higher or lower relative to the size of the confidence interval. The map allows one to see if areas of over- and under-estimation are scattered or clumped in a particular region. It also allows one to examine more closely a particular region of interest.
   1. What to look for – legend description [Visual of the three different map levels]
2. As a scatterplot of BIGMAP estimates vs. FIA plot-based estimates. It is from this scatterplot that the Agreement Coefficients and other metrics are derived. In each of these graphs there are two lines in addition to the points. One is the 1:1 line, to which the relationship is compared. And the second line is the regression line that describes the overall relationship expressed by the points, though in this case a Reduced Major Axis (RMA) regression line is used because of the way it assumes that both estimates being compared could have error. The metrics derived from this graph describe the agreement associated with the entire area over which a species is observed.
3. This same information presented in the map is also depicted in graph form, sorted by the magnitude of the plot-based estimate. This graphing allows one to see whether area of over- and under-estimation are scattered or clustered at low or high biomass values for that species. Chart, line chart

   Description automatically generated

Diagram

Description automatically generatedAs mentioned in #1 above, along with the graphs, several metrics are calculated to quantify and facilitate comparison between the different species and spatial scales of validation. The metrics calculated are: Agreement Coefficient (AC), Systematic Agreement Coefficient (ACsys), Unsystematic Agreement Coefficient (ACunsys), RMA Relationship (y=mx+b), and R2 values.The Systematic agreement coefficient, ACsys calculates the difference between the 1:1 line and the RMA regression line, which describes the bias present. The Unsystematic agreement Coefficient calculates the scatter around the RMA line. A high scatter of points suggests a lot of variation in the species estimates that was difficult to predict from the information used in the modeling. The overall Agreement Coefficient is a combination of both the systematic and unsystematic agreement. The RMA relationship is strongest when there is an equal distribution of scatter around the GMFR line. The R2 is included for comparison because it is a familiar value to many. The R2 value assumes that the FIA estimate is truth and does not distinguish between systematic and unsystematic agreement.

**Chapter 11: Common errors/troubleshooting**

1. [Mosaic Dataset Analyzer: Error 70100—ArcMap | Documentation (arcgis.com)](https://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/mosaic-dataset-analyzer-error-70100.htm)
2. **Error 001491** Cannot process above the size limits of the image service: 'FIA\_BIGMAP\_2018\_Species\_Aboveground\_Biomass'. The allowed maximum number of rows and columns is 4100 and 15000 respectively. Please adjust the output extent and/or cellsize to fit within the limits.

**Chapter 12: Python functionality**

**Chapter 13: Tool development**

1. Dashboard Mike made.
2. Proposed tools and future development

**Literature Cited**

**Appendix**