Georgia Broadband Mapping Project

Product Release White Paper

Contact Name Manager: Rich Calhoun
Contact Phone Number: 404-463-5906
Contact E-mail: richard.calhoun@gta.ga.gov

Submitted By: Rich Calhoun
Contact E-mail: richard.calhoun@gta.ga.gov

Product Specification: Spring 2011 NTIA Data Model
Product/Process: NTIA – April 14th, 2011 Data Deliverable
Dataset Submission QC: NTIA – SBDD_CheckSubmission.py

Document Control

Version	Primary Author(s)	Description	Date Completed
1.0	Kristin Rousseau	Original Draft Document	03/29/11
2.0	Kristin Rousseau	Final Document	04/14/11

Contents

1 Overview

The following describes the Data Gathering, Data Integration, Data Validation and Verification and Quality Control processes utilized to create the Broadband Mapping Project's April 1st, 2011 data submission.

To support various levels of technical and program knowledge, this white paper supplies both a high level summary and a detailed process review.

2 High Level Review

2.1 Data Gathering - Providers

Broadband Service Area, Middle Mile Aggregation Points and Broadband Service Overview
The collection of Broadband Service areas, Middle Mile Aggregation points and Broadband Service Overview information is handled through the following Provider Outreach Process:

- Build and Maintain an Inventory of Broadband Providers through research and State inputs.
- Update Provider Material that describes the data requirements and logistics for data transfer.
- Update NDA for use in project, where applicable
- Maintain multiple protocols for the provider to submit data, including SFTP technology when desired.
- Conduct one-on-one informational discussions with each provider to communicate the following:
    - Requirements of this project
    - Broadband data required to support the product data model
    - Submission protocols available
    - Capability to validate how the supplied data is aggregated

- Download/receive Provider Data
- Establish a repeatable process with Provider. Maintain Provider communication, transaction and data handling records throughout the project (dates contacted, data received, etc.)

2.2 Data Gathering - Community Anchor Institution (CAI)

The collection of CAI information is handled through the following CAI Collection Process:

- Collect and maintain inventory of CAIs through Data Mining, research, and State inputs.
- Maintain web-based CAI portal for institutions to add or confirm attribution, location and enter broadband-specific information.
- Upload web-based data to Core Database for standardization.
- Perform internal cleansing, such as removing duplicate records, identifying gaps in broadband attribution and verifying category.
- Geocode CAI locations.
- Translate Core Database data to deliverable ready format.
- Continue engagement with non-responsive institutions.

2.3 Data Integration Process

The data integration and processing mechanisms currently utilized allow for multiple types of inputs and results in a standardized output that meets the NTIA deliverable requirements. This process is flexible to support data model changes and project requested enhancements.

- Receive inputs from Providers via submission protocols, upload into Sourcing Database and catalog with provider information.
- Review Provider supplied data for completeness and for potential discrepancies that require resolution prior to processing and flag as necessary.
- Categorize input into data type category (addresses, block lists, paper maps, etc.).
- Standardize input based on data type within Staging Database.
- Create Compact Polygons (CP)—(internal methodology for generating area based feature for coverage in Staging Database).
- Apply broadband attribution to CP, Apply metadata to CP
- Perform quality analysis of the CP against the source supplied to identify any completeness or accuracy issues.
- Request additional information from the provider if elements of coverage are missing or contain discrepancies. This is a second manual quality check to ensure data is complete. Following completion of CP creation, process steps within Data Validation & Verification occur
    - Process coverage area to build the required NTIA data model layers

- Process CAI data input into internal standardized format, as mentioned above under CAI Create Product Deliverable based on NTIA and State-level requirements.
- Following the creation of the product, process steps within Data Validation & Verification occur

2.4 Data Validation & Verification

To ensure the data collected and processed is accurate and comprehensive, a holistic approach has been developed to further validate and verify the data. Following the initial mapping of providers' coverage area and serviceability claims, the project team uses the following methods:

- **Third-Party Data Verification:** Visually and programmatically compare the coverage against third-party data.
  Pitney Bowes and American Roamer data are used in cases where a coverage area is questionable. All anomalies identified during this analysis are reviewed with the providers.

- **Broadband Provider Validation – Provider Portal Application:** Providers were trained on and requested to use a secure interactive web application to review their current coverage area(s) and supporting broadband attribution and validate their data or submit change requests to update their data.

  All provider change requests go through the Data Integration Process and a review with the provider to complete validation.

- **Confidence Values:** All Verification, Validation, and manual quality reviews are tracked by provider and then by technology type, which is then stored and maintained within a "Validation" table. A confidence value is assigned based on the collected information to highlight provider coverage areas that require further investigation and enhancements.

2.5 Quality Control

Following collection, processing and analysis of the provider and CAI data, the product is checked manually and algorithmically against the NTIA data model. Some of the items included within these checks are as follows:

- Format Correctness
- Table & Field Structure
- Valid Values
  - Including default values, where applicable
- Geographic Extent and Topology Errors

Prior to data submission, another quality control script supplied by NTIA is run. This script, SBDD_CheckSubmission.py, creates an output in text form that is required to be submitted along with the final deliverable. All errors must come up clean, unless otherwise specified from NTIA.

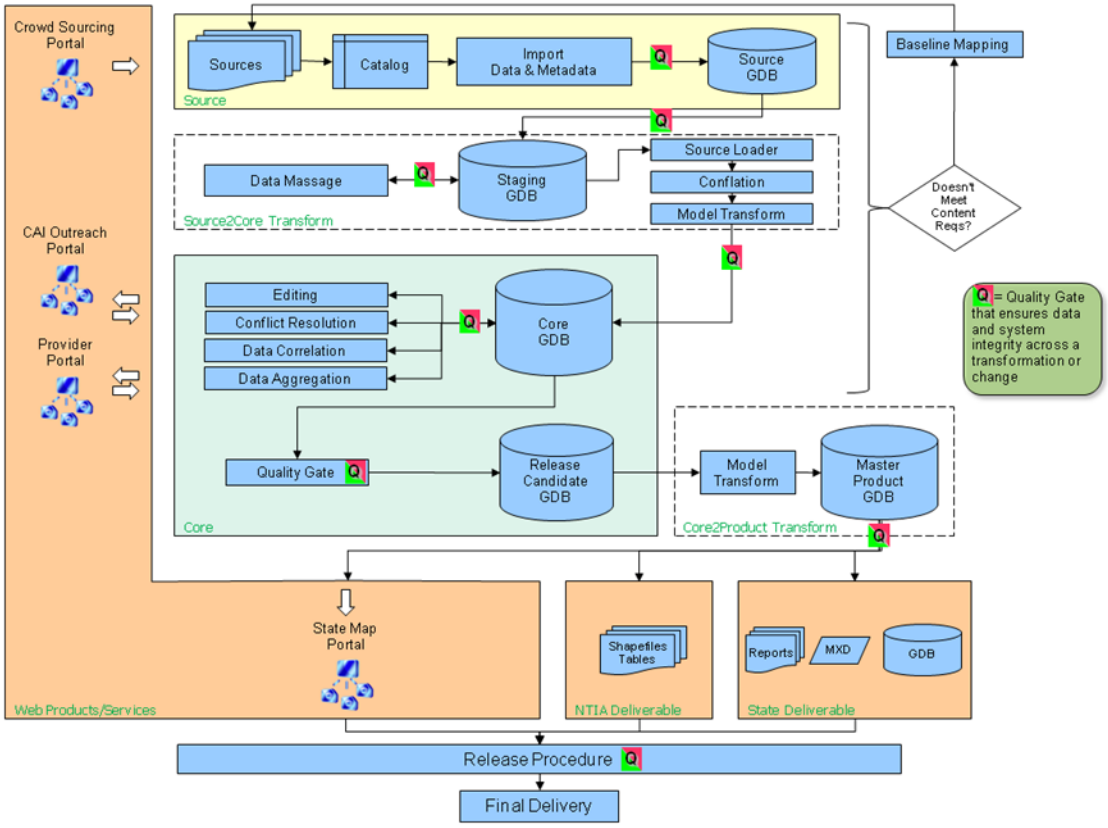
Exceptions to the script as noted by NTIA on the SBDD Workspace on 03/25/11 at the following link:
https://sbdd-granteeworkspace.pbworks.com/w/page/38218329/CheckSubmissionExceptions

- Longitude values for States outside the lower 48 (any table)
- CAI results for Transtech, MaxAdUp, MaxAdDown if BBService is ‘No’ or ‘Unknown’
- Overview MaxAdDown, MaxAdUp if 100% of record level data has MaxAdDown or MaxAdUp populated

3 Detailed Process Review

Below is a detailed review of the data collection, integration and quality control points along the broadband data gathering and mapping process.

Diagram of overall process:



3.1 Provider Outreach

For the April 2011 data submission, an e-mail notification was sent to all providers with supporting deliverable dates. The Provider Portal web application was released and training webinars held so providers could use this application to submit changes to and/or validate their current coverage area(s).

Data was also collected from the providers via e-mail and SFTP, depending on their comfort level to submit data in time for the April 1st deadline.

In support the data collection effort, providers that did not timely respond to the outreach were contacted by phone.

3.2 Outreach Materials

The original provider packet sent via email to the providers included the following documents and files:

1) Letter from the State inviting them to participate in the program
2) Copy of the non-disclosure agreement (NDA)
3) Copy of the Mapping NOFA from the NTIA
4) Copy of the NOFA Clarification from the NTIA
5) Broadband service address example file in CSV format
6) Word document describing service address example file
7) Broadband service block example file in CSV format
8) Word document describing service block example file
9) Broadband service street example file in CSV format
10) Word document describing service street example file
11) Broadband subscriber example file in CSV format
12) Word document describing subscriber example file
13) Broadband wireless coverage area sample shapefile
14) Word document describing wireless coverage area sample shapefile
15) Instructions for downloading, installing, and using the WinSCP secure FTP application

3.3 Outreach Process

The provider outreach process is comprised of the following general steps:

1) Send the provider package and introduction letter to the main point of contact for the provider
2) Follow up with email and call to verify that the main point of contact is correct.
3) If necessary, discuss the NDA further and resolve any redlines.
4) Once the correct primary contact is established, set up a call, if necessary, to learn more about the provider's offerings and direct them to the appropriate outreach materials.
5) If providers are unable to be contacted (non-responsive) or indicate that they are not interested in participating (non-cooperative) mark them as such on the provider tracking sheet. These providers will be escalated to the state for further action.
6) As the providers are collecting the required data, provide instruction on downloading, installing, and using the WinSCP secure FTP application, if required.
7) Arrange with the providers to transfer the data in whatever way they are comfortable. Some providers will find regular email acceptable. Others will want to use the secure FTP application.
8) After data is received and reviewed, it may be necessary to contact a provider for clarification or to address incomplete data sets. In the interest of building and maintaining relationships, care is given not to push the provider but to work with it to obtain accurate data in the best possible format.

3.4 Data Collection

3.4.1 Data Transfer Procedures

There are three primary ways data is collected from providers. These are:

1) Secure FTP using the WinSCP application
2) Regular email
3) Mail

3.4.2 Initial Data Review and Quality Assurance

The initial data review and quality assurance process consists of the following general steps:

1) Access the data from the secure FTP site or email
    a. If emailed, place copy of original data set in the appropriate provider folder on the secure FTP site
2) Place copy of raw data on local computer in a working directory.
3) Review data and determine course of action based on type of data received.
4) Ensure data is complete and contact provider to address any gaps.

Note: The goal is to get as many providers as possible to provide subscriber address data in the correct format. Obviously, this will not be possible with all providers so we will continue to have to process various types of provider-supplied data.

3.5 Data Ingestion

3.5.1 Data Ingestion Overview

The following outlines the process steps taken based on the type of input supplied by the data provider:
Point Data

- Subscriber location
- DSLAM location
- Central Office location
- Broadcast Tower location

Linear Information

- TIGER street segments

Polygonal Information

- Census Blocks
- Coverage Area

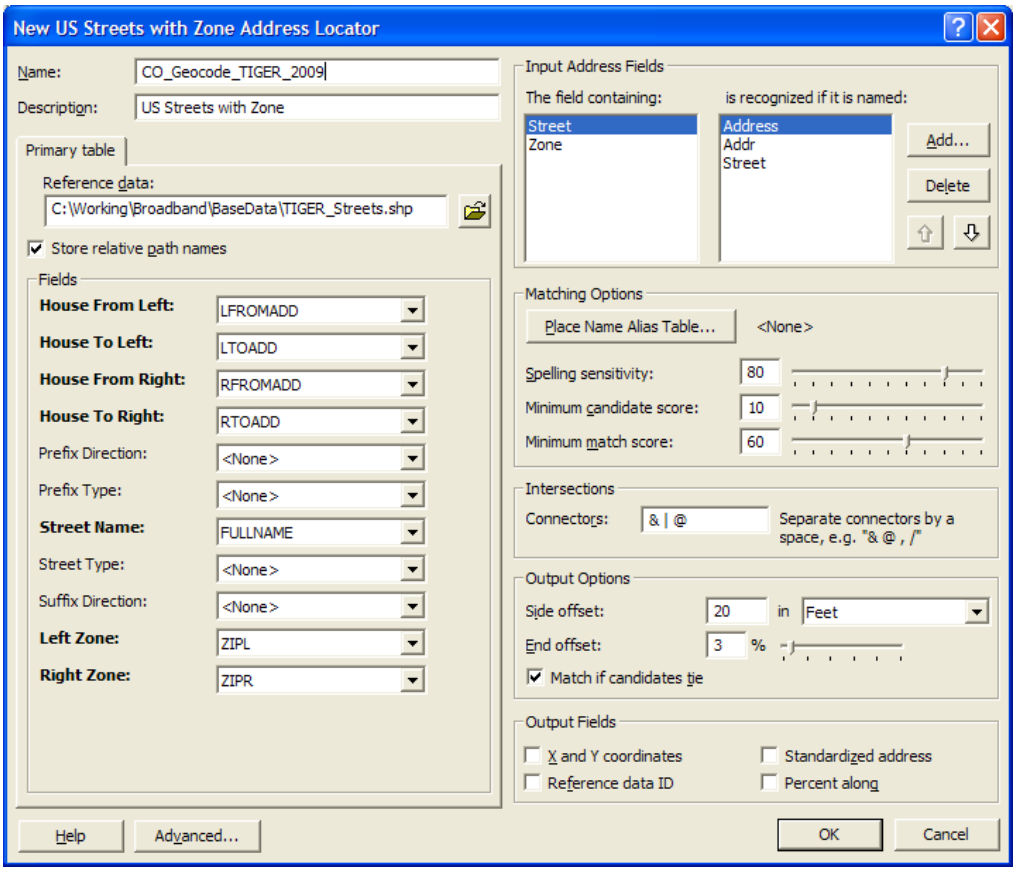
Overall, the process is geared toward taking the provider data supplied and creating polygon shapes to append to the bb_cov feature class. The bb_cov feature class is the interim data set that is then processed using the **makeDeliverable.py** Python scrip to create the MapConnect data layers that will be delivered to the state and, ultimately, to the NTIA. Following are the detailed instructions used in this process.

3.5.2 Point Data

3.5.2.1 Subscriber Location – Address Data

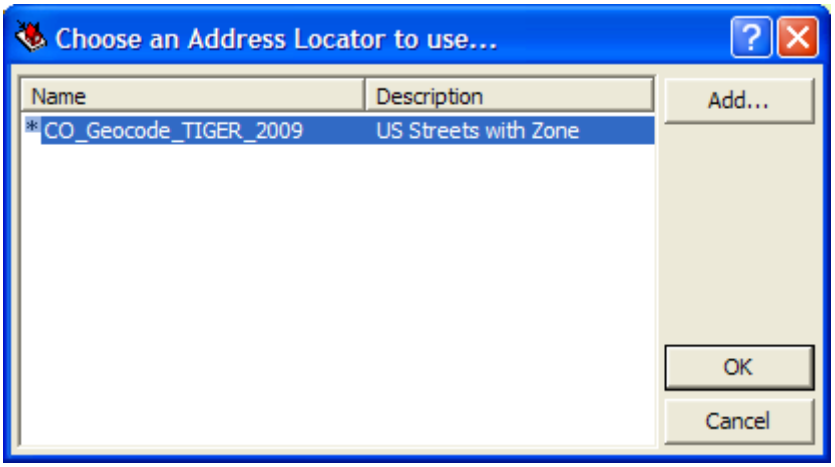
In the event that the data provider supplies subscriber address data the following actions occur:

1) First, convert the address data to a clean Excel spreadsheet in an appropriate address data format.
   a) Usually, this has the following columns: street address (number, pre-directional, pre-modifier, street name, street type, post-directional, and post-modifier concatenated together), city, state, ZIP.
2) Configure the ArcGIS geocoding tool to use the TIGER 2009 streets dataset
   a) In ArcCatalog, create a new Address Locater by right-clicking in the white space of the appropriate directory and selecting New>Address Locater from the dropdown menu.
   b) Select “US Streets with Zone” and press OK.
      i) Note: It is likely that multiple Address Locators will have to set up to handle the variety of provider address data received.
   c) Navigate to the TIGER Streets 2009 file and press OK.
   d) Fill in the dialog box as seen below:

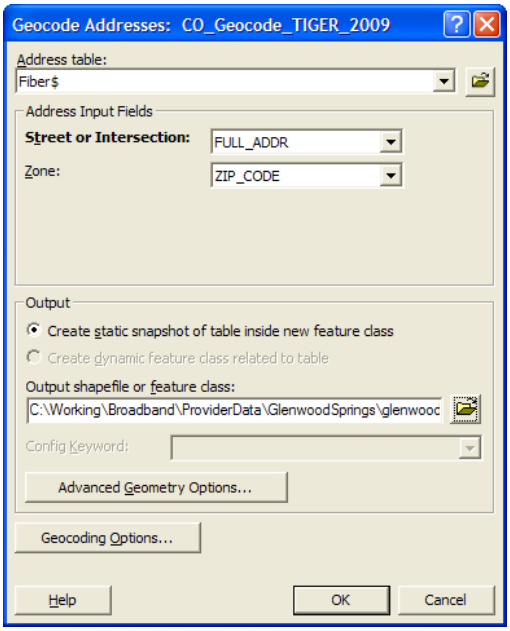


e) Click OK.

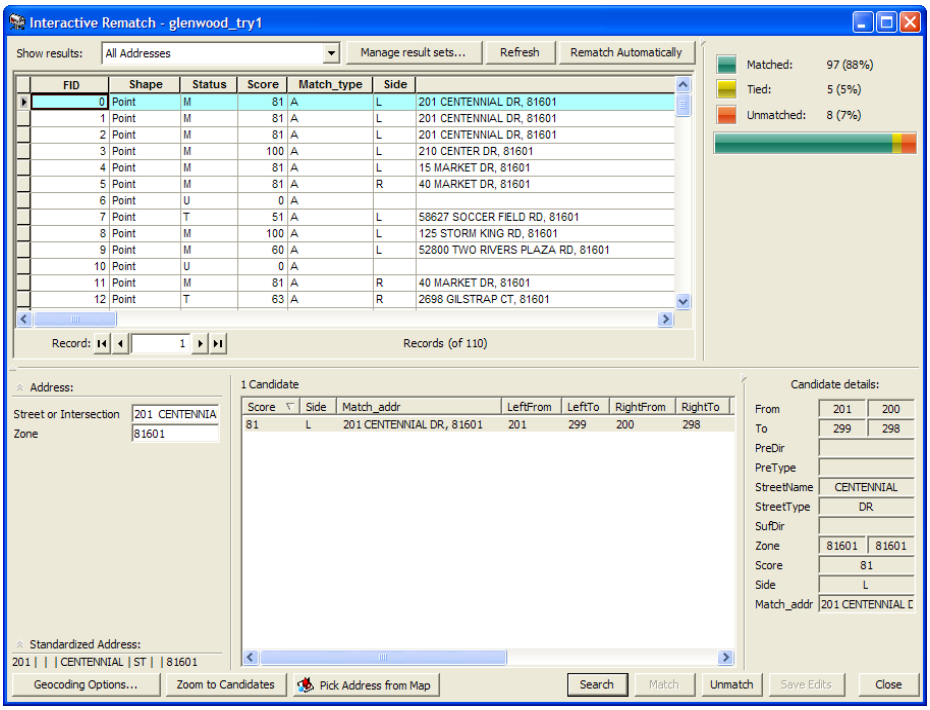
3) Open up ArcMap, and add the Excel spreadsheet with the address information.
4) Right-click on the Excel spreadsheet and select Geocode Addresses from the dropdown menu.
5) Select the appropriate address locator by clicking Add…. then OK.



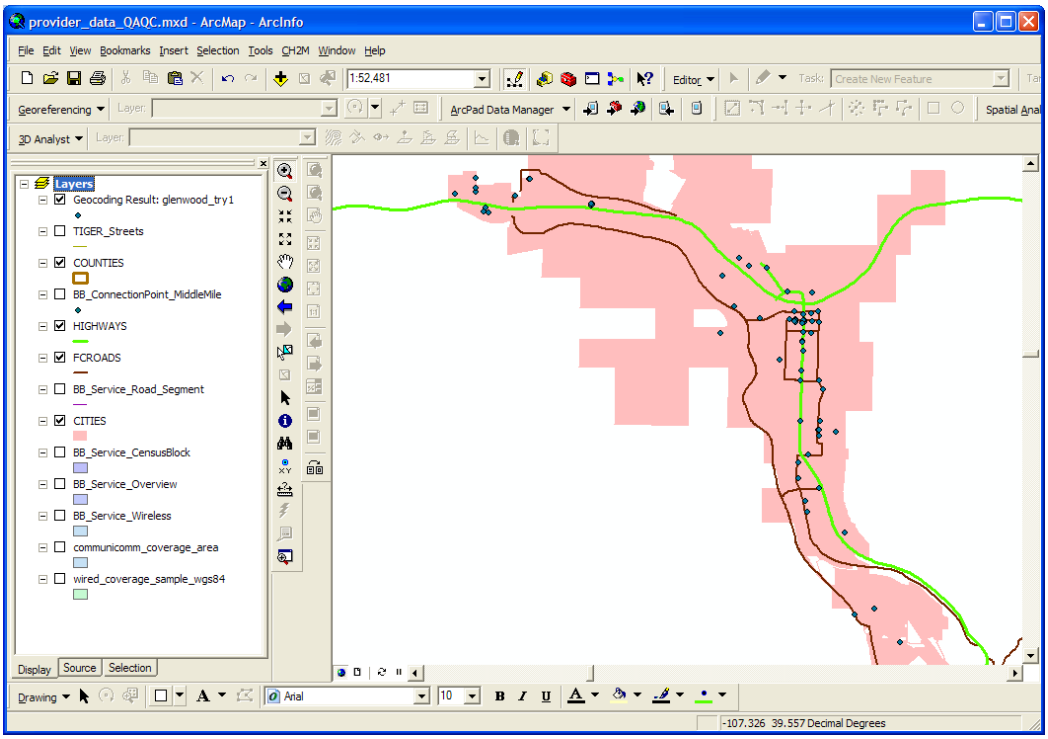
6) Fill out the Geocode Addresses dialog box as shown below:



7) Geocode the list in batch mode using the geocode service set up in Step 2 above, accepting all the default parameters.
8) Review results.



9) Adjust geocoding parameters accordingly and repeat batch to resolve issues.
10) Manually geocode unmatched addresses until target hit rate achieved, generally 90%.
11) Visually inspect the data as seen below:



12) Follow the steps detailed in Subscriber Location – GIS Data below

3.5.2.2 Subscriber Location – XY Data

If the provider supplies a list of subscriber data with accompanying XY data such as latitude and longitude, the steps are as follows:

1) Refine the format in Excel so that the data can easily be opened using ArcMap.

    a. Remove all font color, highlighting, cell colors and borders, clean up column headers and make sure there are no merged cells.

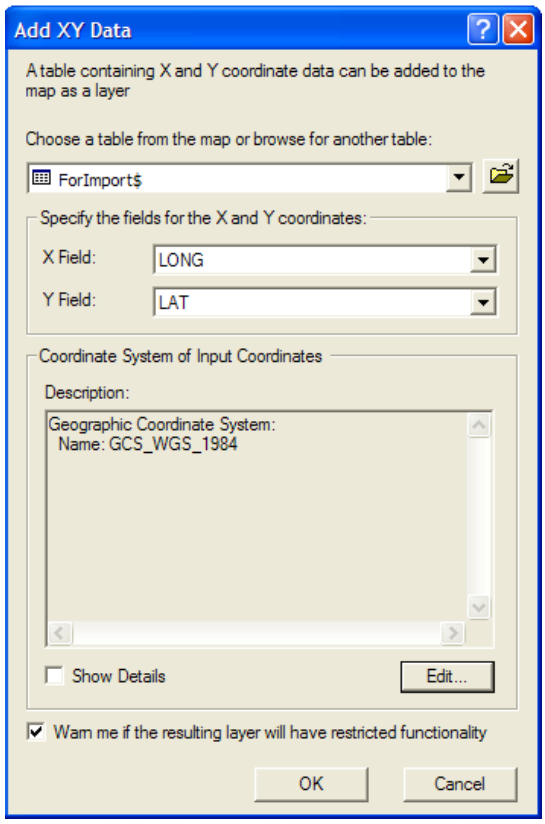
    b. Make sure that XY locations are in decimal degrees.

        i. To convert from degrees, minutes, seconds (39º 26’ 45.67”) to decimal degrees us the following formula: DD + (MM/60) + (SS.SSS/3600).

        ii. Note: if XY locations from some other coordinate system are provided, you can use those in the process below but you must know what the coordinate system is.

2) Open up the Excel worksheet in ArcMap.

3) From the menu bar, select Tools>Add XY Data…



4) Supply the appropriate fields for the X and Y coordinates, choose the appropriate coordinate system and press OK.

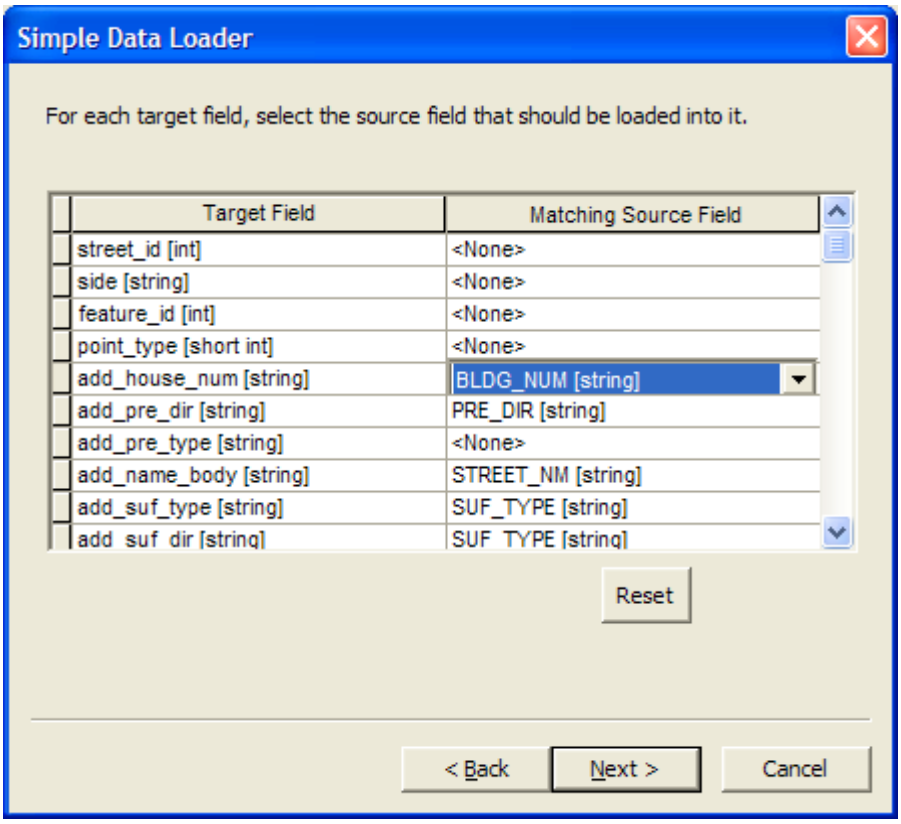
5) Results are an event layer, not a true spatial layer. Export the data by right-clicking the event layer and selecting Data>Export Data… from the dropdown menu.

6) Follow the steps detailed in Subscriber Location – GIS Data below.

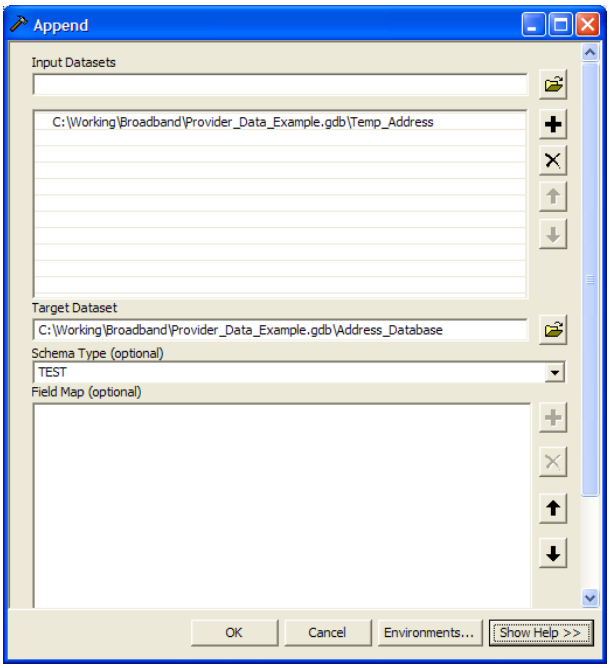
3.5.2.3 Subscriber Location – GIS Data

If the provider supplies subscriber location in GIS format, the only process step is to load that data into the appropriate data schema and it will be ready for processing.

1) First, load the data into the Point Address database schema (please see Appendix D for an example of the Point Address database schema.) using an empty feature class in that schema.
2) In ArcCatalog, right-click on the empty feature class and select Load from the dropdown menu.
3) Navigate to the provider address GIS data set and then map the attribute fields accordingly, as seen in general below:

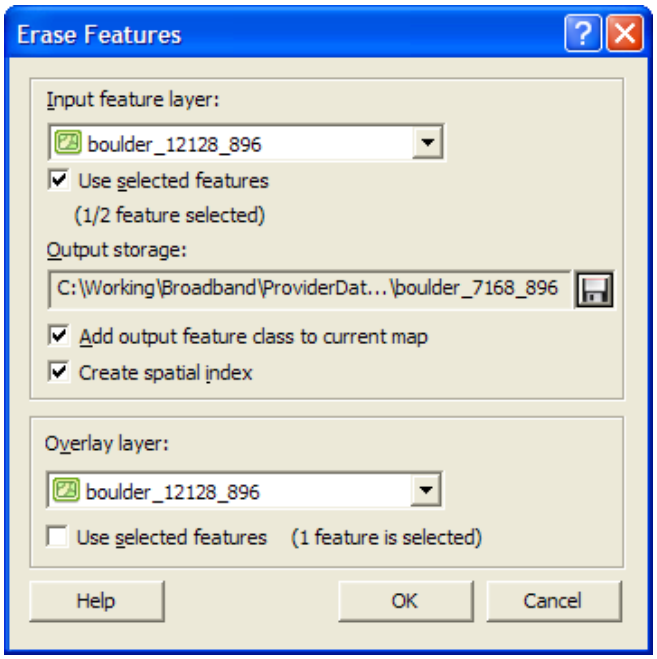


4) Once you have successfully loaded the provider address data into the temporary database with the correct schema, you will now append that data to the overall Point Address database.
5) In ArcToolbox, use the Append command (*Data Management Tools>General> Append)* to add the features into the overall Point Address database, as seen in general below:



6) Since the data is already in the Point Address database schema, there is no n need to alter the Field Map in the Append tool.
7) After appending, calculate metadata reflecting geometry source and representation values.
8) Break provider-specific points into separate county feature classes and perform the following steps per county feature class:
    a. <u>Within ArcGIS</u>
        i. Summarize download and upload speeds [first,last] to determine all speeds available for county.
            1. This will save as a DBF table. Keep track of location for future reference.
        ii. Buffer county address point featureclass to 150'.
            1. During buffer command, dissolve on "ad_down"; "ad_up"; "provider"; "dba"; "frn"; "tt"; 'all metadata fields'; "stctyfips". Save as…. county_fastestdown_fastestup.
            2. (Example using Qwest data: boulder_40128_20128, where boulder=county; 40128=ad_down; 20128=ad_up)
            3. Note: these attribute fields are specific to the Point Address database.
        iii. Select the features that represent the lowest speeds
    b. <u>Using XtoolsPro (*http://www.xtoolspro.com/*)</u>
        i. In the XTools Pro toolbar, select XTools Pro>Layer Operations>Erase Features
        ii. Use the same feature class for Input and Overlay
        iii. Check Use selected features on the Input feature, as seen below.

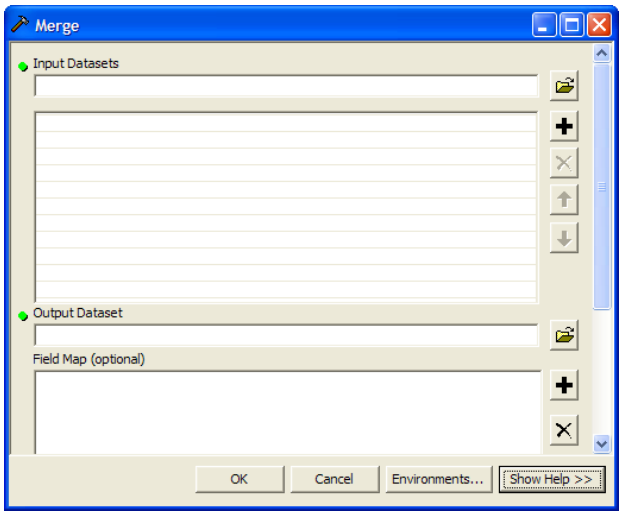
iv. Repeat and erase slowest speeds one speed at a time. Saving each new feature class as the next slowest speed, using the same naming convention as above. A general example is seen below:



c. Within ArcGIS

i. Edit/delete speeds from the attribution table of each feature class, so each remaining feature class has only one speed value.

ii. Merge individual speed feature classes together using the Merge command in ArcToolbox (Data Management Tools>General>Merge). The dialog box is seen below:



iii. Merge individual county feature classes together using the Merge command in ArcToolbox (Data Management Tools>General>Merge).

iv. Since the county files are all in the same schema, do NOT alter the Field Map portion of the command interface.

v. When all the county files are merged together into one dataset, use the Append command in ArcToolbox *(*Data Management Tools>General>Append) to add the features to the bb_cov interim data set. Use the Field Map portion of the Append tool to map the appropriate field values to their corresponding fields in the bb_cov feature class.

3.5.2.4 DSLAM or Central Office Location – Address Data

In the event that the provider supplies DSLAM (digital subscriber line access multiplexer) or Central Office address data please follow the steps below:

1) Follow the process for geocoding points in Subscriber Location – Address Data, above.
2) Follow the steps detailed in DSLAM or Central Office Location – GIS Data below.

3.5.2.5 DSLAM or Central Office Location – XY Data

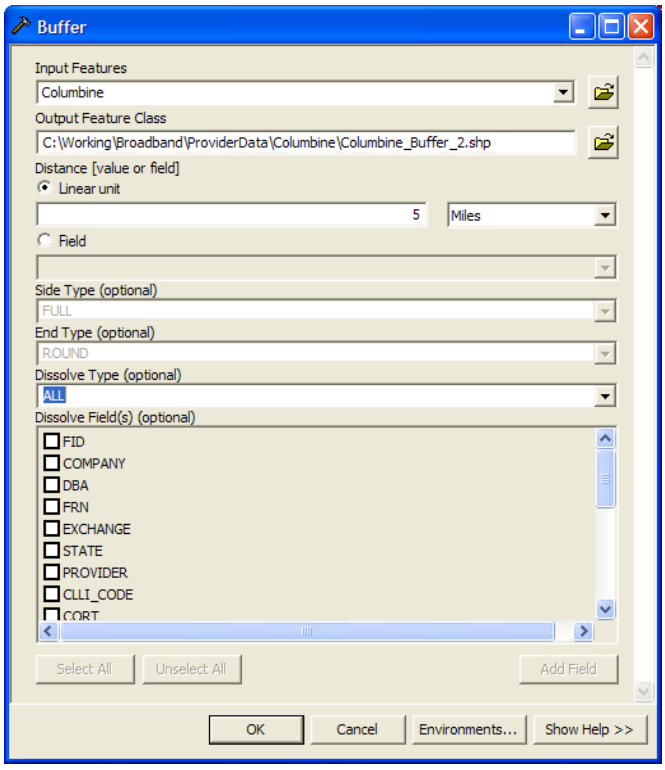
In the event that the provider supplies DSLAM (digital subscriber line access multiplexer) or Central Office XY data please follow the steps below:

1) Follow the process for creating points from XY data in Subscriber Location – XY Data, above.
2) Follow the steps detailed in DSLAM or Central Office Location – GIS Data below.

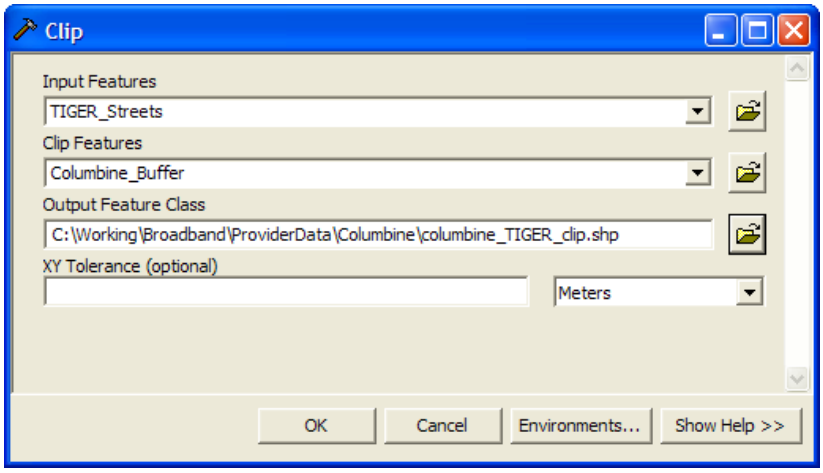
3.5.2.6 DSLAM or Central Office Location – GIS Data

In the event that the provider supplies DSLAM (digital subscriber line access multiplexer) or Central Office GIS data please follow the steps below:

1) Buffer the DSLAM/Central Office points feature class
   a) Add the point feature class to ArcMap
   b) Open up ArcToolbox and go to Analysis Tools>Proximity>Buffer
   c) Set the buffer distance to 5 miles
   d) Set the dissolve type to ALL
   e) Name the output feature class
   f) Typical Buffer tool is seen below:

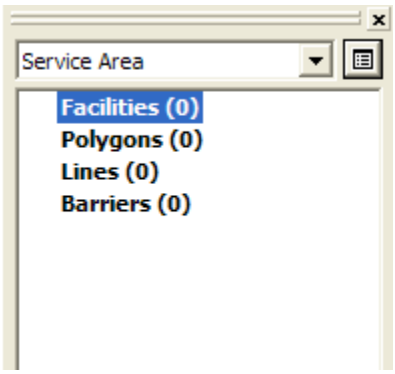


   g) Press OK
2) Use the resulting buffer feature class to clip the TIGER street layer (as described earlier):
   a) Add TIGER street layer to ArcMap
   b) Open up ArcToolbox and go to Analysis Tools>Extract>Clip
   c) Complete the dialog box as seen below:

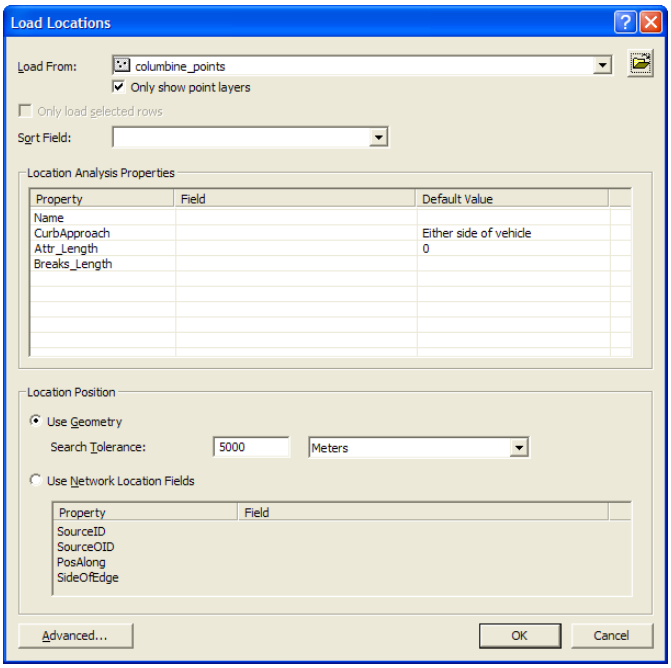


d) Press OK.

3) Using ArcCatalog and within the file geodatabase:
   a) Right Click and create a new Feature Dataset
      i) For the Feature Dataset settings:
         (1) Name the feature dataset accordingly
         (2) Select horizontal coordinate system by importing the coordinate system associated with the clipped TIGER street layer by selecting Import and navigating to the location of that feature class
         (3) No vertical coordinate system needed
         (4) Leave all x,y,z,m values at default.
         (5) Press Finish
4) Import previously created street feature class into new Feature Dataset
5) Right-click Feature Dataset and create new Network Dataset – accept all default setting for the Network Dataset
   a) Note: the Network Analyst extension must be turned on
6) In ArcMap Turn on the Network Analyst Toolbar by going to View>Toolbars>Network Analyst
7) Add the Network Dataset created in Step 5 to ArcMap
8) Using Network Analyst Toolbar drop down – create "New Service Area"
9) Open up the Network Analyst Window by selecting the button.



10) Right click Facilities layer, select Load Locations, and navigate to the DSLAM/Central Office facilities feature class.



11) Press OK.

12) Click the Service Area Properties button 

13) For the following tabs change the following properties:

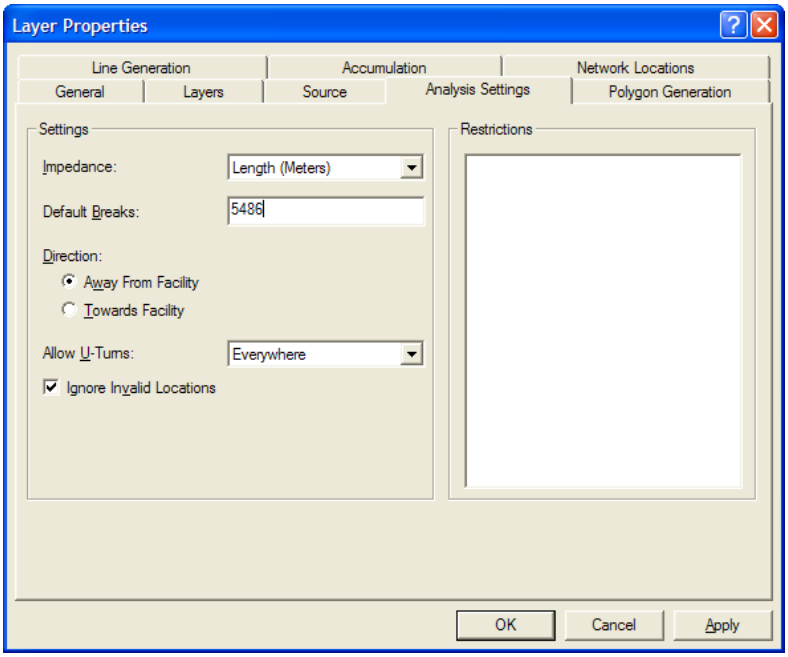
a) "Polygon Generation" tab

i) Select "Merge by break value"

ii) Also disable the Trim Polygons option

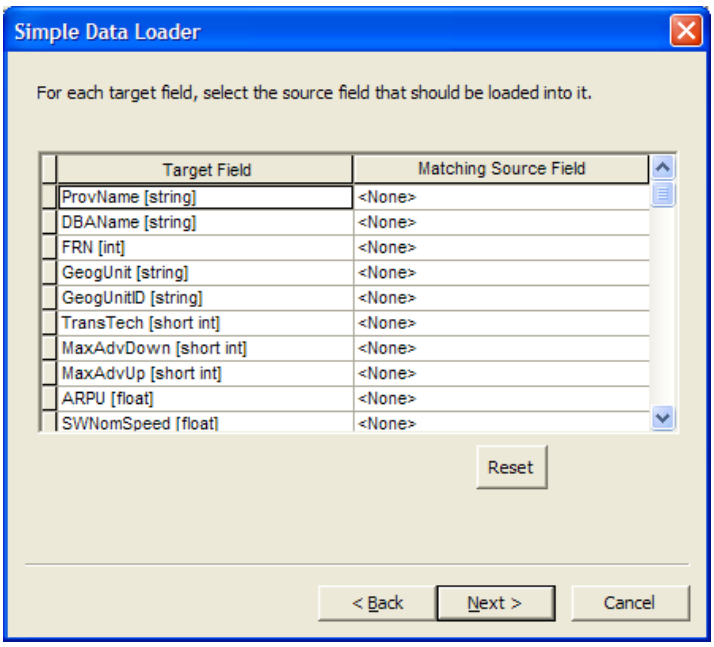
b) "Analysis Settings" tab – using and converting the specified DSLAM buffer distance from feet to meters – input buffer distance value in meters into the "Default Breaks" location

i) Generally, 18,000 feet (5486 meters) from DSLAM or Central Office location is used as the buffer distance



c) Click OK.

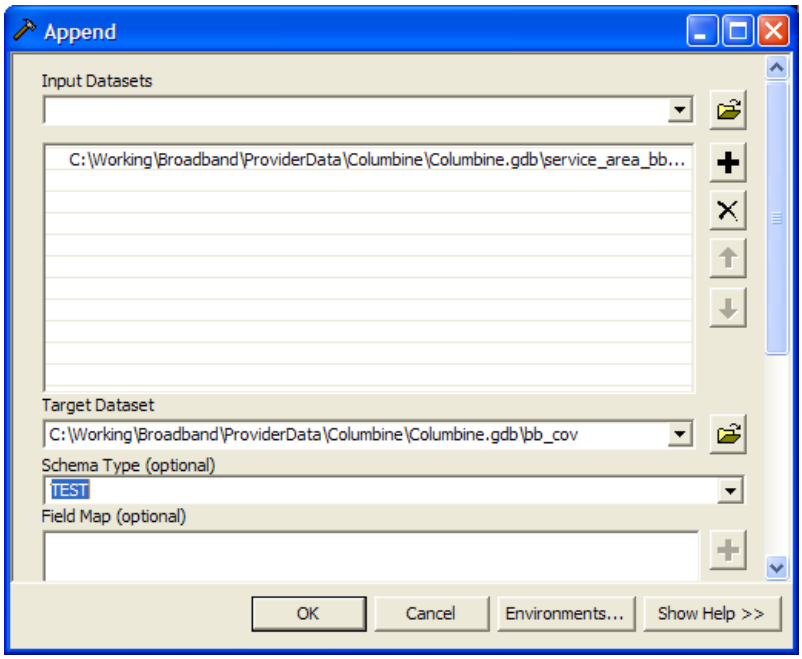
14) On the Network Analyst Toolbar click the "Solve" button  to create service area polygons.
15) Right-click on the created service are polygon in the layer list, and select Data>Export Data from the dropdown list.
16) Export to a feature class in the file geodatabase you created earlier
17) In ArcCatalog, create an empty feature class with the schema of the bb_cov feature class and load the feature class created in Step 16 into it.
    a) Right-click on the empty feature class, select Load>Load data from the dropdown menu and navigate to the location of the service area feature class
    b) Press the Add button, hit Next
    c) Accept the defaults and hit Next
    d) Do NOT attempt to map any fields, as seen below:



e) Press Next, then Next again, then Finish.

18) In ArcToolBox, go to Data Management Tools>General>Append

19) Append the formerly empty feature class to bb_cov, completing the dialog box as seen below:



20) Leave the Schema Type as TEST

21) Press OK.
22) In ArcMap, open up bb_cov for editing and manually input associated attribution.

3.5.2.7 Broadcast Tower Location – Address Data

In the event that the provider supplies wireless broadcast tower location address data please follow the steps below:

1) Follow the process for geocoding points in Subscriber Location – Address Data, above.
2) Follow the steps detailed in Broadcast Tower Location – GIS Data below.

3.5.2.8 Broadcast Tower Location – XY Data

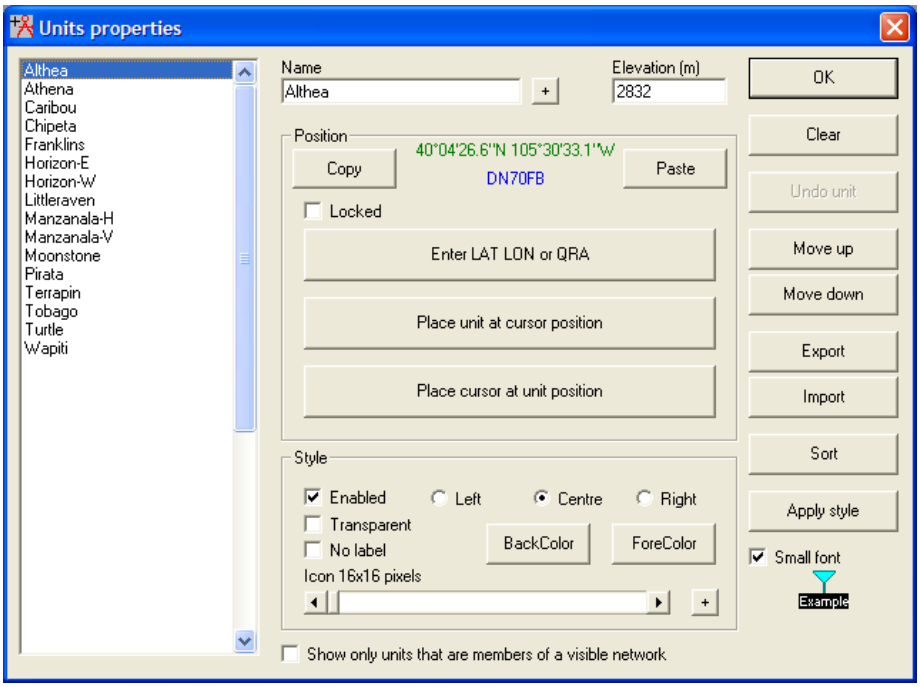
In the event that the provider supplies wireless broadcast tower location XY data please follow the steps below:

1) Follow the process for creating points from XY data in Subscriber Location – XY Data, above.
2) Follow the steps detailed in Broadcast Tower Location – GIS Data below.

3.5.2.8.1 Broadcast Tower Location – GIS Data

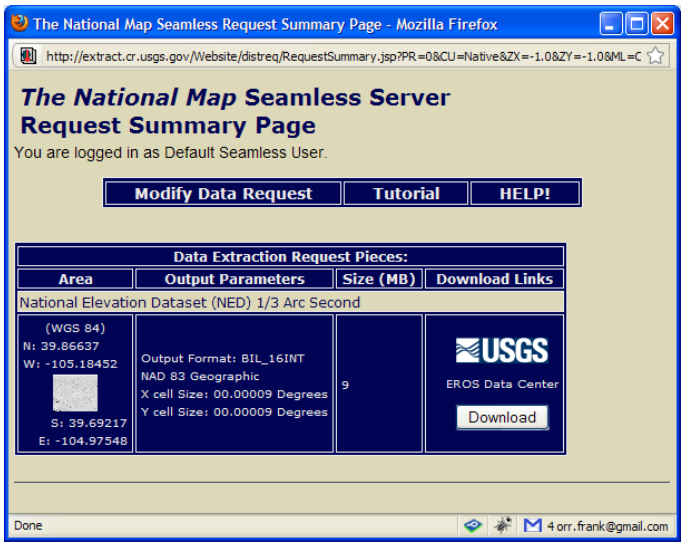
In the event that the provider supplies wireless broadcast tower location GIS data please follow the steps below:

1) Download the required software (Radio Mobile) from the website: http://www.cplus.org/rmw/english1.html
2) Install the software according to the standard directions, found here: http://www.cplus.org/rmw/download/download.php?S=1
3) Open up the application
4) Load the broadcast tower location and elevation information by selecting File>Unit properties. The following dialog box appears:



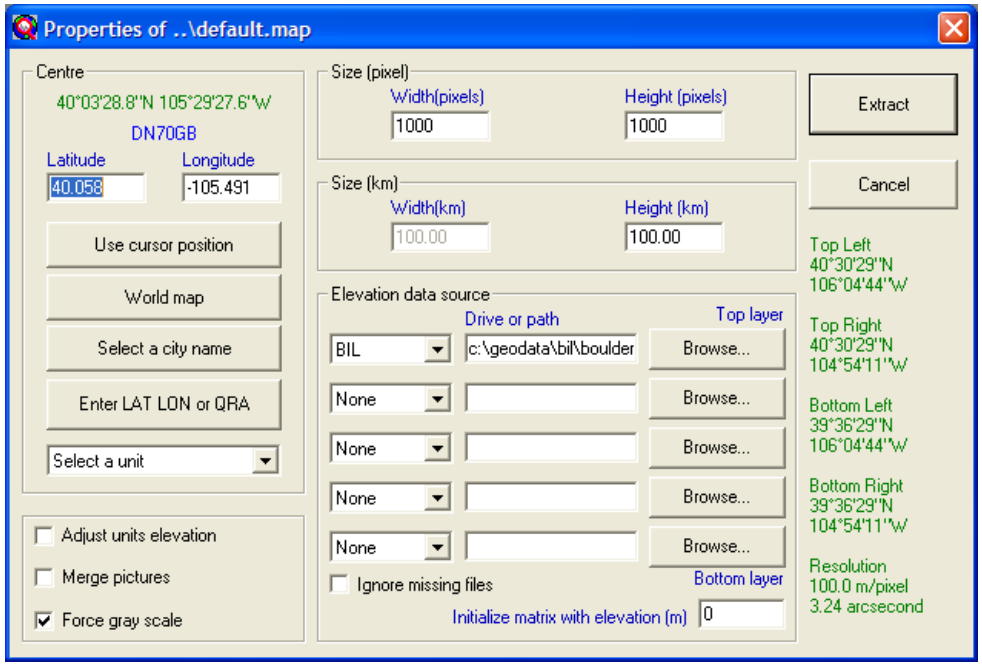
5) Add in the information for all the towers supplied by the WISP data provider, including the elevation. If provider does not supply elevation, this information can be obtained from Google Earth.
    a. I f available, use the Import button to import a Google Earth KML of the tower locations.
6) Go to the National Map Seamless Server (http://seamless.usgs.gov/) and download elevation data sufficient to contain the tower locations.
    a. At least the 1/3” NED data is needed. Select this by clicking the Download button in the upper right of the web site and checking the box nect to 1/3 “ NED.
    b. Zoom to the area of interest and use the Download tools:

to define the area to download.
    c. Click the Modify Data Request button to request the data in BIL_16INT format, not ESRI GRID, as seen below:



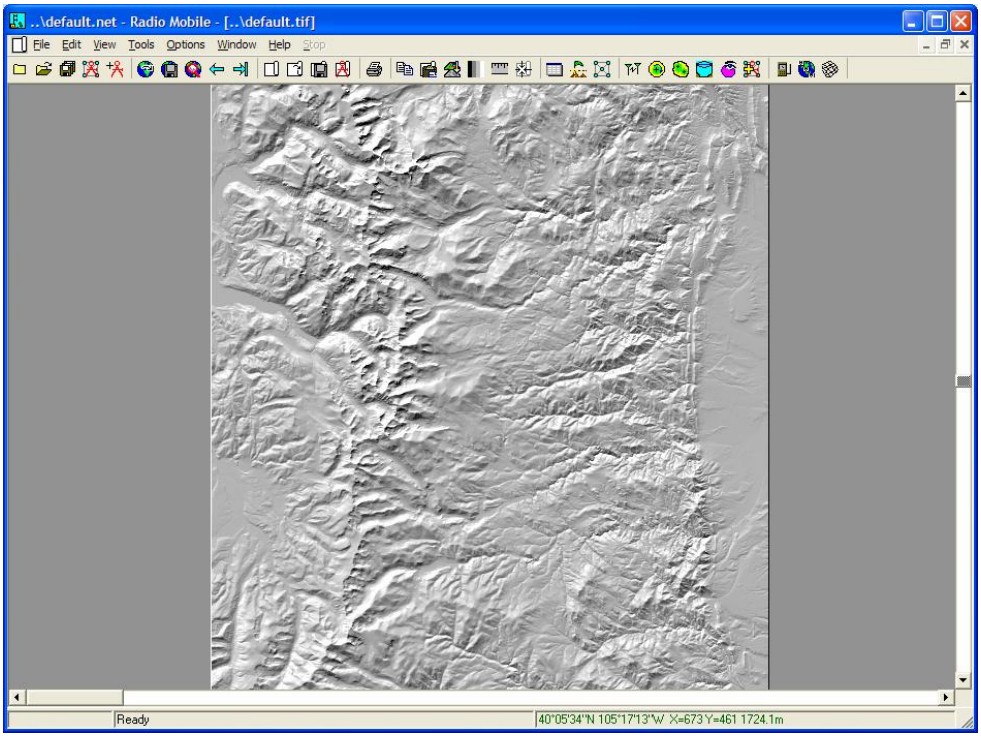
d. Download the data and unzip it.

7) Select File>Map Properties to define the map
8) Enter in a latitude and longitude in the center of the tower locations
9) Set the size (in pixels) and the size (in kilometers) of the map
10) Set the directory path leading to the BIL elevation data just downloaded
11) The dialog box is seen below:

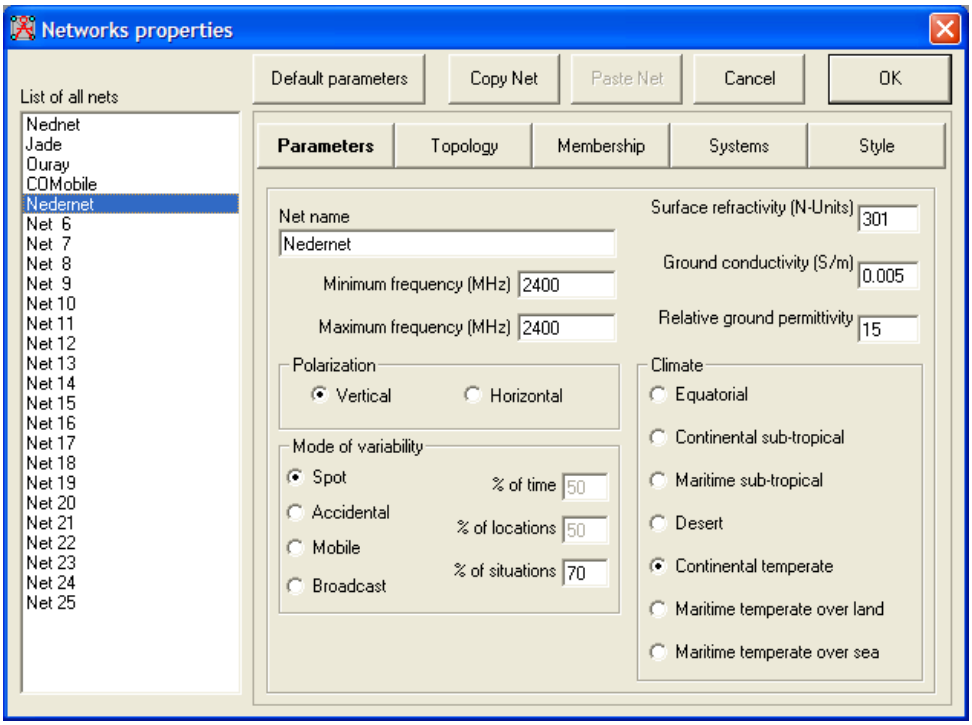


12) Hit Extract.

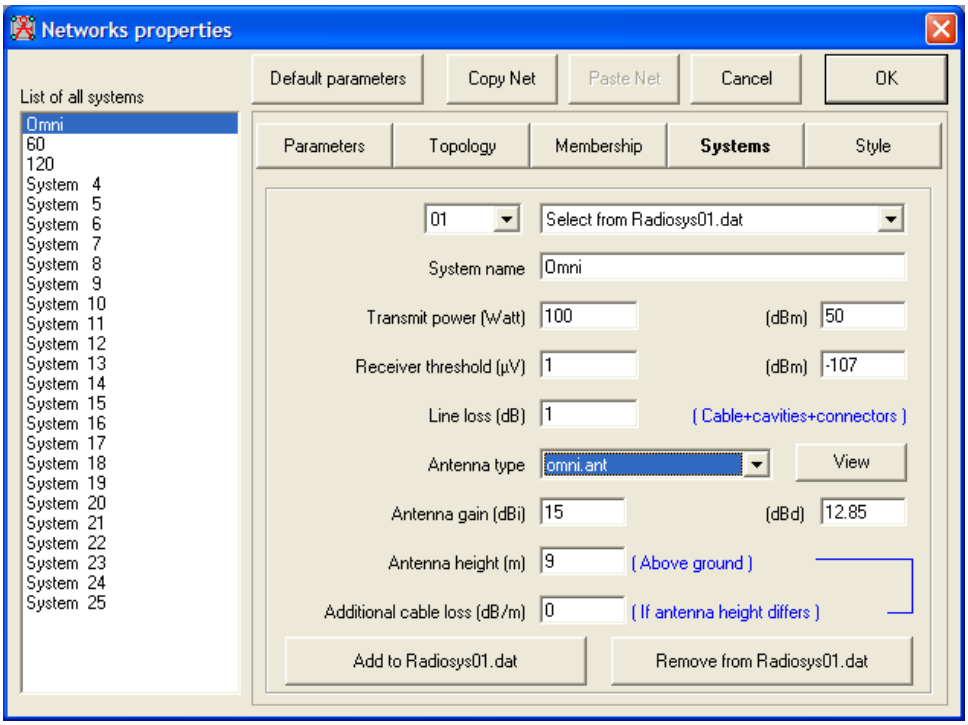
13) The elevation data is render as a hill shade, as seen below:



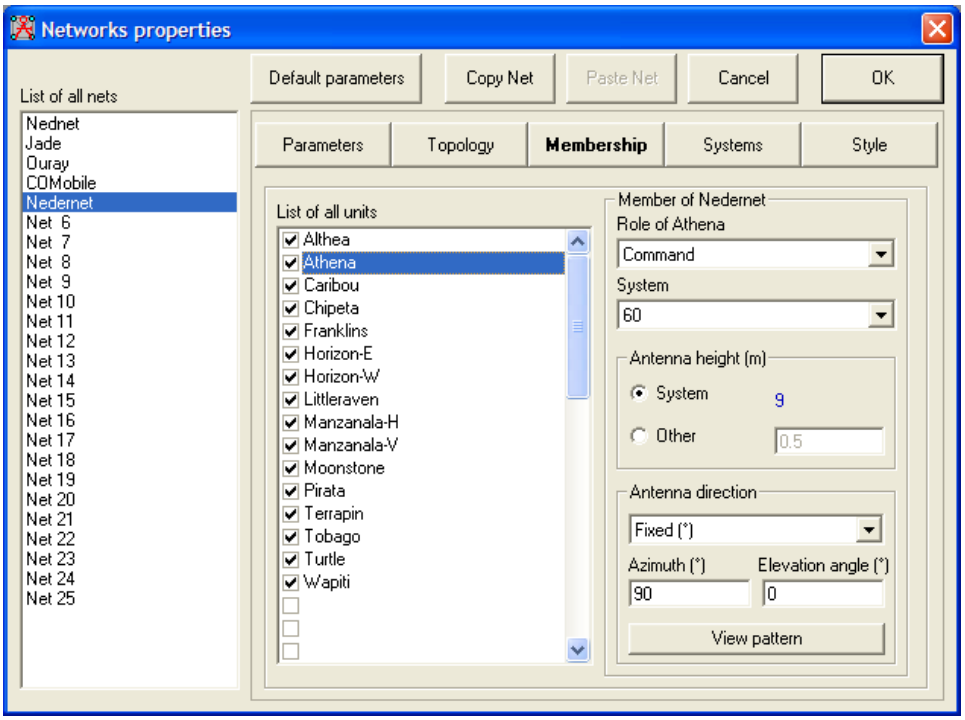
14) Select File>Network properties from the main menu
15) Create a new network and enter in the frequency range under the Parameters tab, as seen below:



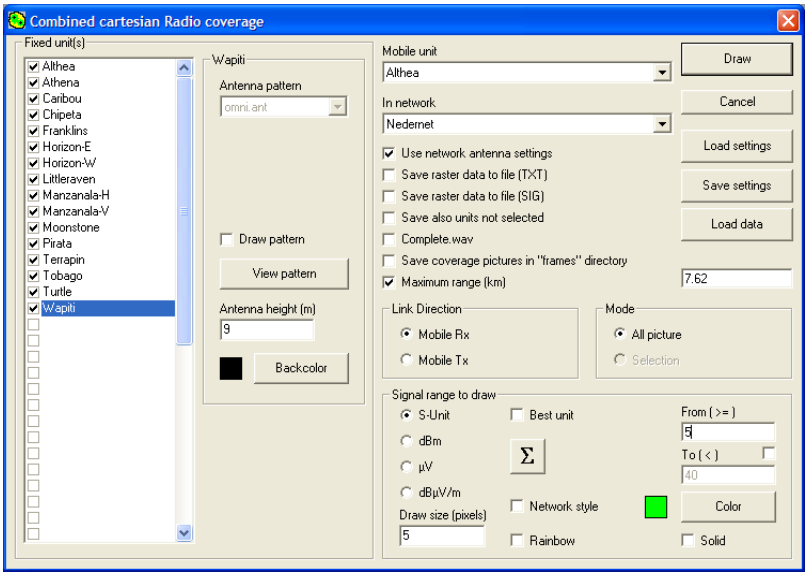
16) Leave all the other values as they appear, and select the Systems tab
17) Create enough systems to cover all the varieties of equipment in the provider network. This will include the antenna type, height, and line loss, as seen below:



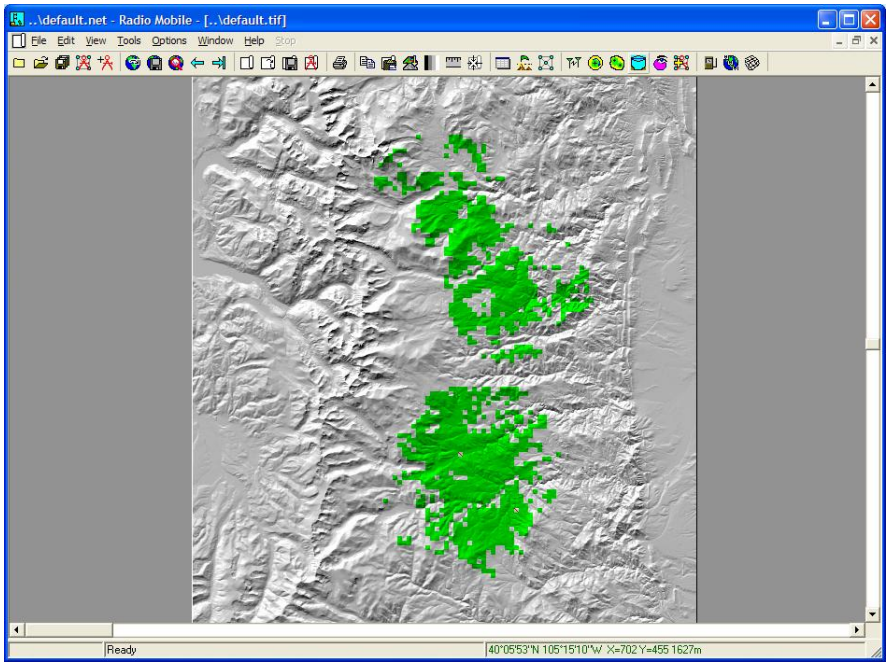
18) Now click on the Membership tab, and assign the individual towers to their respective systems, providing the azimuth for non-omnidirectional antennas, as seen below:



19) Press OK.
20) Select Tools>Radio Coverage>Combined Cartesian from the main menu
21) Complete the dialog box as seen below, providing the Maximum Range from the highest tower beam radius supplied by the provider.
22) Set the Pixel Size at 5 (experiment depending on the area covered to get the right level of granularity) as seen below:



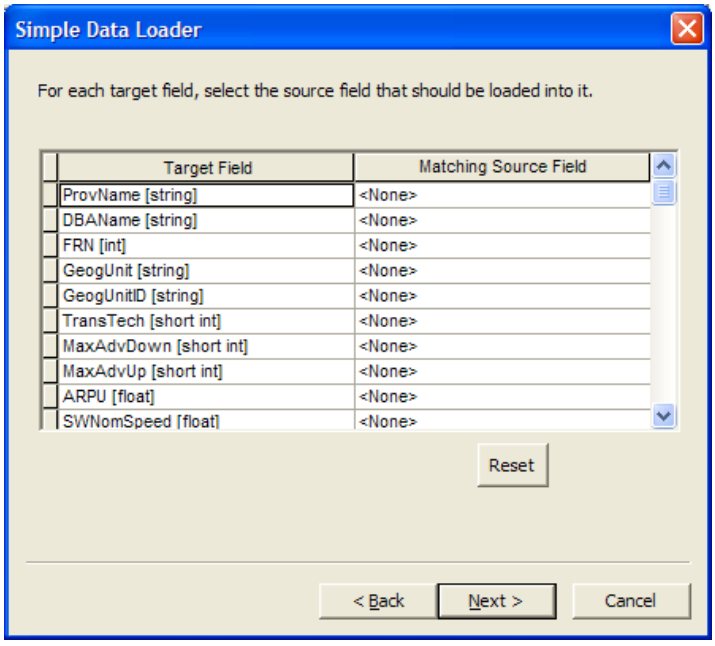
23) Set the signal range to draw to S-Unit and type 5 in the From (>=) box.
24) Press Draw.



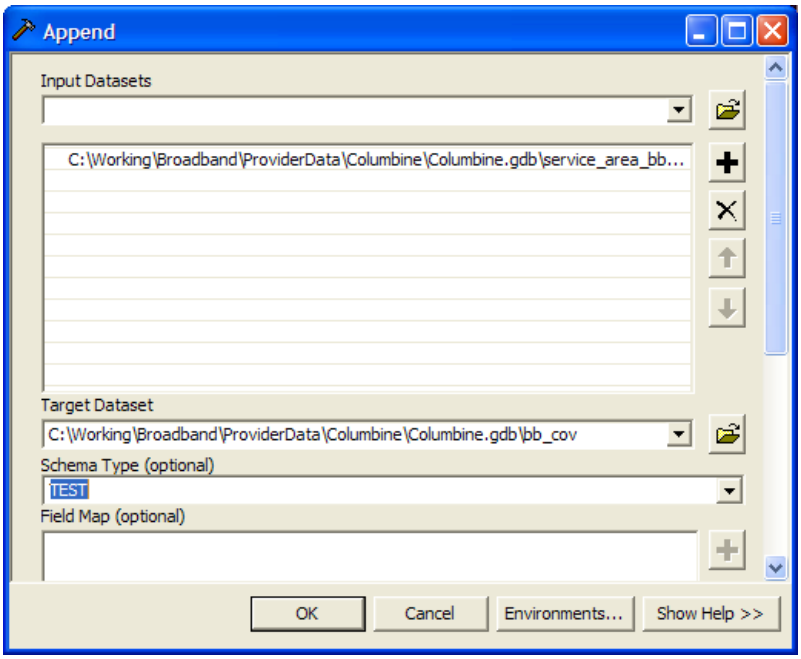
25) Save the resulting image as a TIF by selecting File>Save Picture as.
26) Open ArcMap and load the BIL elevation data you used in Radio Mobile.
27) Load the TIF image you created and georeference it using the corners of the BIL data.
    a. The corners of the data can be seen in the TIF image.

28) Follow the georeferencing directions from the Coverage Area – PDF/JPG/Other Image Format section below.
29) Use the Georeferencing Toolbar to Update the Georeferencing for the TIF data set.
30) In ArcToolbox, select Data Transformations>From Raster>Raster to Polygon and input the georeferenced TIF you just created as seen below:

31) Open the resulting polygon feature class up for editing using the Editing toolbar in ArcMap and clean up as necessary.
32) In ArcCatalog, create an empty feature class with the schema of the bb_cov feature class and load the feature class created above into it.
    a. Right-click on the empty feature class, select Load>Load data from the dropdown menu and navigate to the location of the service area feature class
    b. Press the Add button, hit Next
    c. Accept the defaults and hit Next
    d. Do NOT attempt to map any fields, as seen below:



    e. Press Next, then Next again, then Finish.
33) In ArcToolBox, go to Data Management Tools>General>Append
34) Append the formerly empty feature class to bb_cov, completing the dialog box as seen below:



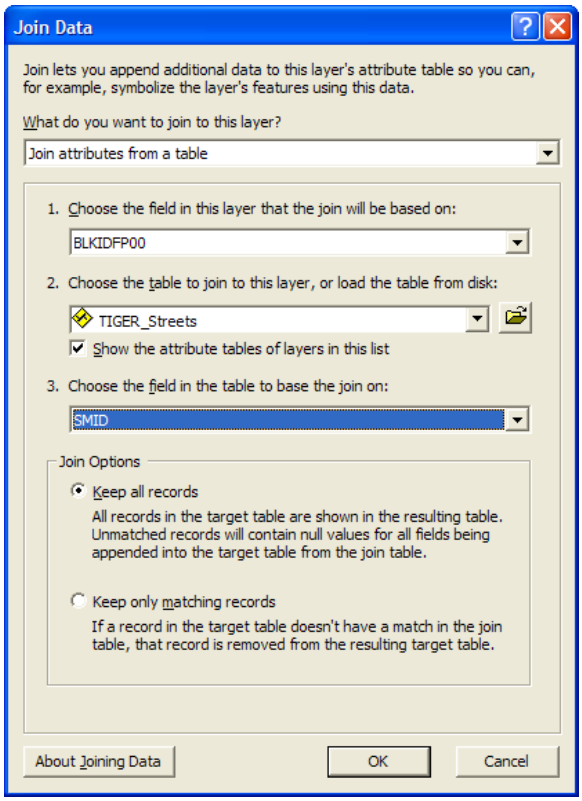
35) Leave the Schema Type as TEST
36) Press OK.
37) In ArcMap, open up bb_cov for editing and manually input associated attribution.

3.5.3 Linear Data

3.5.3.1 TIGER Street Segments – List, Spreadsheet, or GIS Data

In the event that the provider supplies TIGER street segments in list or spreadsheet format please follow the steps below:

1) Join TIGER road segments to 2000 census blocks feature class using one of two methods based on how the data is provided:
   a) If the TIGER data is provided with a Census Block ID, then join the segments to the Census Block geometry based on that ID
      i) Load both data sets into ArcMap
      ii) In the layer list, right-click on the 2000 census block feature class and select Joins and Relates>Join
      iii) In the dialog box, select the TIGER road segments data and the proper attribute fields for joining, as seen below:



iv) Press OK

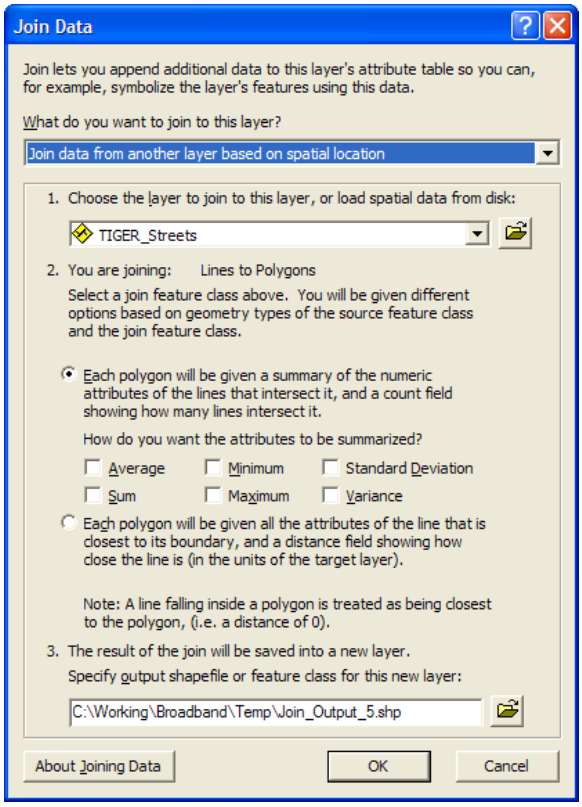
b) If the data provided is a list containing TLIDs, then join to the TIGER line data using the TLID, and use a spatial join to associate the TIGER segment with the coterminous block based on the block ID

i) Load both data sets into ArcMap

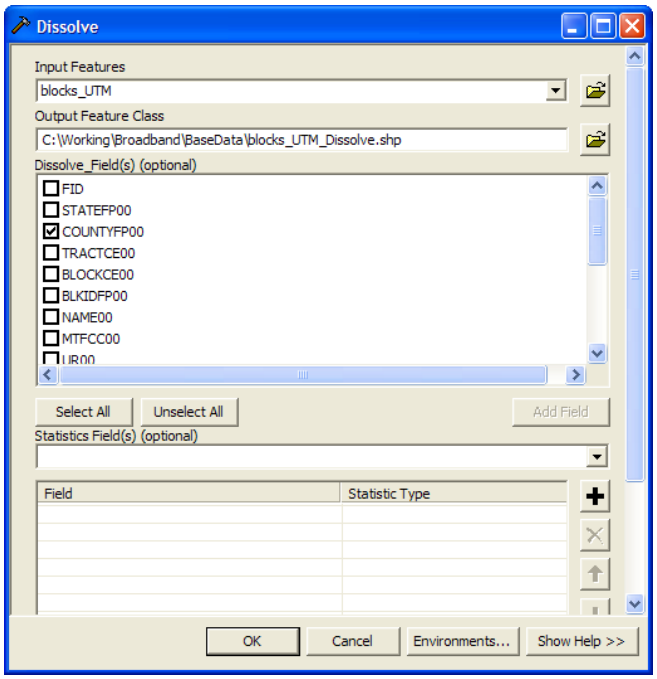
ii) In the layer list, right-click on the 2000 census block feature class and select Joins and Relates>Join

iii) Select "Join data from another layer based on spatial location" from the dropdown menu

iv) Complete the dialog box as seen below and press OK.



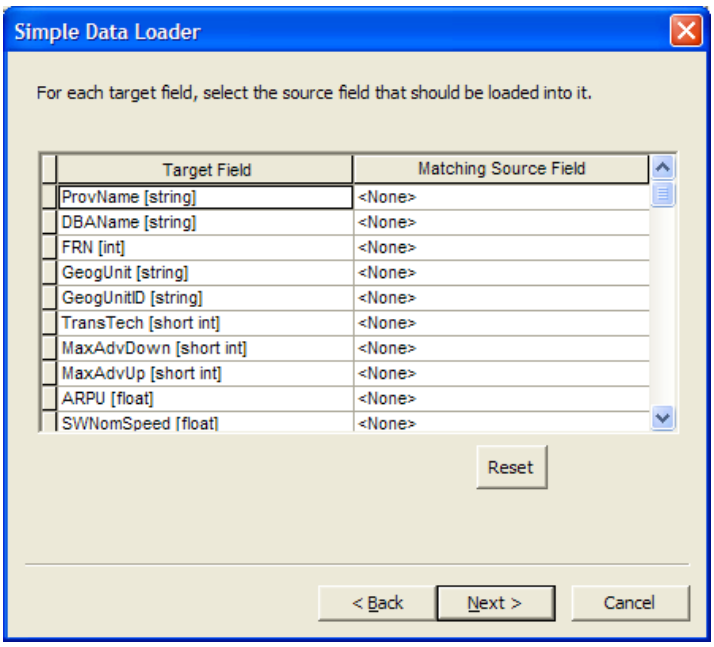
2) Export joined records into a temporary feature class.
3) If joined Census Block geometry is confined to one specific area then dissolve blocks into one record. If joined Census Block geometry is distributed throughout a particular state then dissolve sub-selections of census blocks for each county.
   a) Use the County FIPS code to dissolve by county.
   b) In ArcToolbox, select Data Management Tools>Generalization>Dissolve
   c) Complete the Dissolve dialog box as seen below:



d) Press OK.

4) For each dissolved region, open up the feature class for editing using the Editing tool in ArcMap and remove unnecessary slivers and other small holes. For general guidance on editing features in ArcMap, see http://webhelp.esri.com/arcgisdesktop/9.3/pdf/Editing_Tutorial.pdf

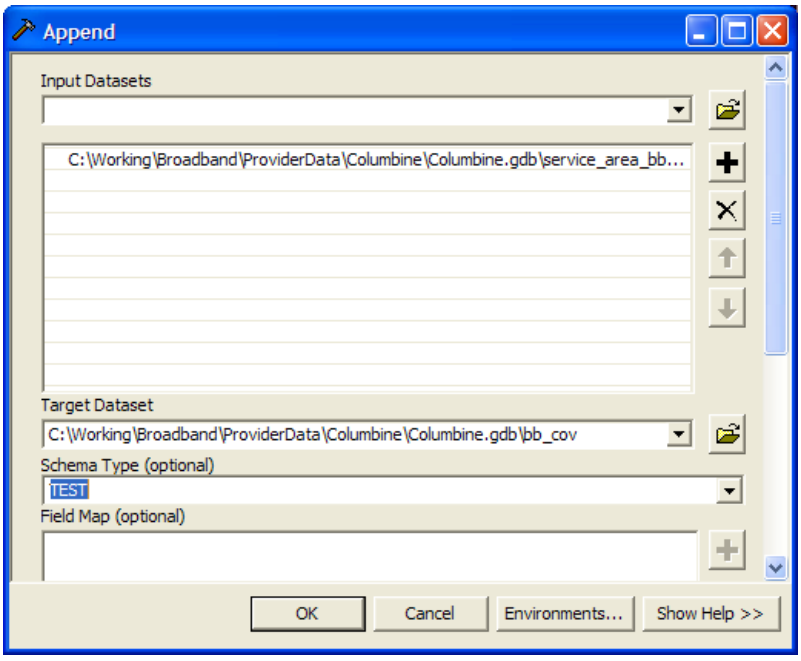
5) In ArcCatalog, create an empty feature class with the schema of the bb_cov feature class and load the feature class created above into it.
   a) Right-click on the empty feature class, select Load>Load data from the dropdown menu and navigate to the location of the service area feature class
   b) Press the Add button, hit Next
   c) Accept the defaults and hit Next
   d) Do NOT attempt to map any fields, as seen below:



a) Press Next, then Next again, then Finish.

6) In ArcToolBox, go to Data Management Tools>General>Append

7) Append the formerly empty feature class to bb_cov, completing the dialog box as seen below:



8) Leave the Schema Type as TEST

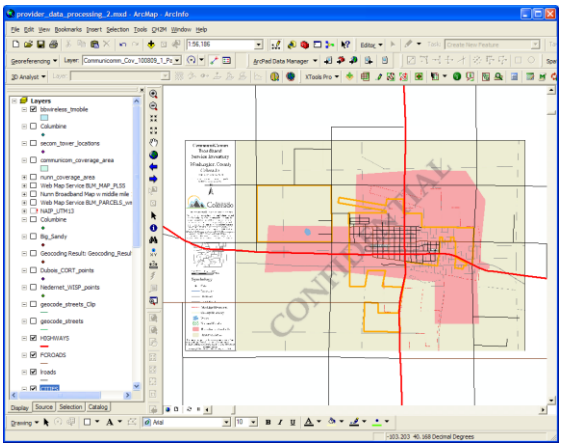
9) Press OK.
10) In ArcMap, open up bb_cov for editing and manually input associated attribution if necessary.

3.5.4 Polygonal Data

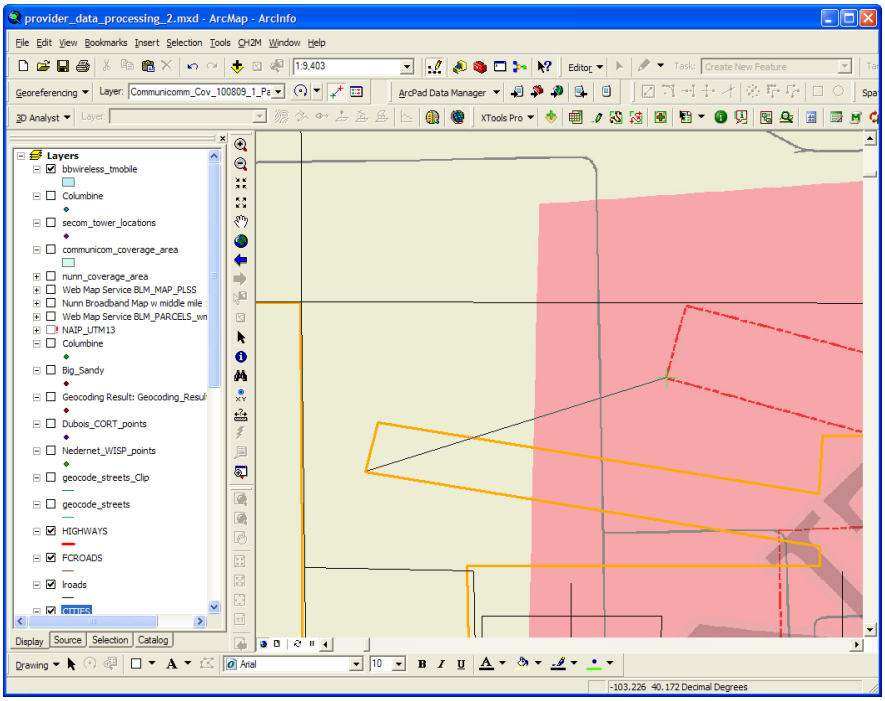
3.5.4.1 Coverage Area – PDF/JPG/Other Image Format

In the event that the provider supplies coverage area data in some image format such as PDF or JPG format please follow the steps below:

1) If in PDF format, open in Adobe Acrobat and Save As… JPG format.
2) Open up the JPG image in ArcMap.
3) Add the required basemap vector data for georeferencing.
   a) This will generally be either the CDOT data or TIGER data
4) Change the coordinate system of the data frame to the desired end coordinate system
5) Zoom to the general location of the JPG map image
   a) This is the location based on the vector data, not the JPG image itself. For example, if you know that the JPG image represents an area around the town of Limon, zoom to the town of Limon in your vector data.
6) Open up the Georeferencing toolbar by selecting View>Toolbars>Georeferencing from the main menu bar.
7) Using the Georeferencing toolbar, select Fit to Display, results seen below:



8) Use the Control Point button  to add control points to the map
9) Use common points in the base data set and the JPG image
   a) For example, find major street intersections, county/city boundaries, etc.
   b) Try to distribute the points more or less in the four corners on the image for the best transformation
10) Click on the location on the image first, then click on the corresponding location on the vector data base map, as in the image below:



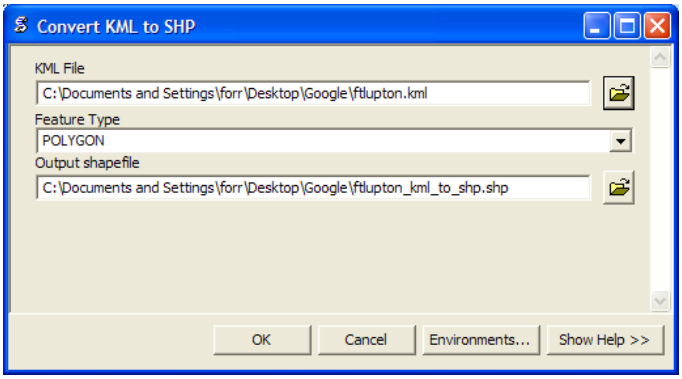
11) After placing each control point, the image transformation will update automatically.
12) Repeat until satisfied with the transformation.
    a) Note: The transformation may take up to four points, although sometimes only two are necessary.
13) When satisfied with the transformation, select Update Georeferencing from the Georeferencing toolbar dropdown.
    a) This will create a "world" file (.jgw in the case of JPGs) in the same directory as the image file.
14) In ArcCatalog, create a new polygon shapefile with the appropriate data schema for a provider coverage area, which can be found in Appendix D.
15) Add the shapefile to ArcMap.
16) Using the Editor Toolbar, select Start Editing. Set the Task: to "Create New Feature."
17) Use the Sketch Tool  to digitize a new coverage polygon using the coverage area outline from the georeferenced JPG and add the required attributes manually.
18) Repeat the above steps for all subscriber speed coverage areas provided.
19) Follow the steps detailed in Coverage Area – GIS Data below.

3.5.4.2 Coverage Area – KML/KMZ

In the event that the provider supplies coverage area data in Google Earth KML or KMZ format please follow the steps below:

1) Use a KML to SHP converter to translate file into an ESRI format
2) http://arcscripts.esri.com/details.asp?dbid=15603

3) Download the script and follow the provided instructions for installing it in ArcToolbox.
4) Double-click on the script in ArcToolbox and navigate to the location of the KML file, as seen below:



5) Add the new shapefile to ArcMap. Repeat for all KML files provided.
6) Follow the steps detailed in Coverage Area – GIS Data below.

3.5.4.3 Coverage Area – CAD Data

In the event that the provider supplies coverage area data in GIS format please follow the steps below:

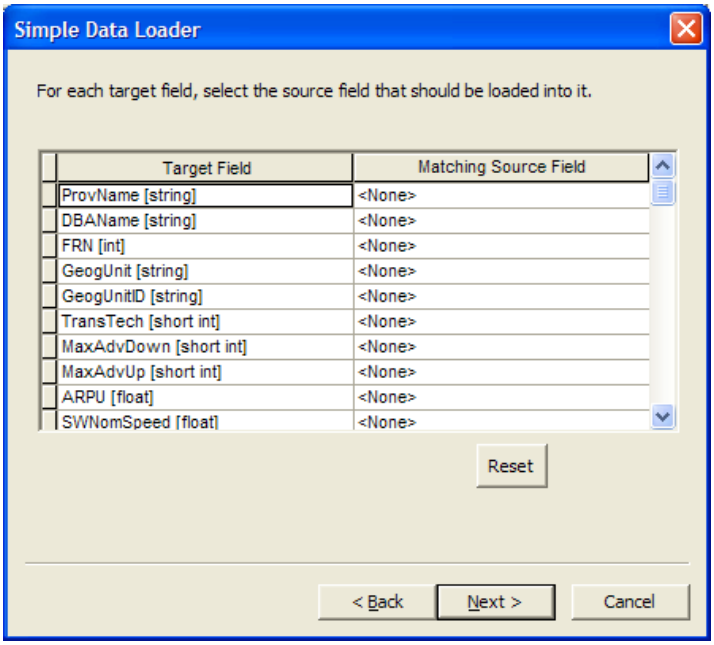
1) Transform the CAD dataset into an ESRI format
2) http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Transforming_CAD_datasets
3) It may be necessary to contact the provider first to determine the coordinate system of the CAD data.
4) If the CAD data is not in a standard coordinate system, it may be necessary to use ArcMap to georeference the CAD data to a known coordinate system first.
   a) To do so, follow the instructions provided above in "Coverage Area – PDF/JPG/Other Image Format."
5) In ArcCatalog, create a new polygon shapefile with the appropriate data schema for a provider coverage area, which can be found in Appendix D.
6) Add the shapefile to ArcMap.
7) Using the Editor Toolbar, select Start Editing. Set the Task: to "Create New Feature."
8) Use the Sketch Tool  to digitize a new coverage polygon using the coverage area outline from the georeferenced CAD file and add the required attributes manually.
9) Follow the steps detailed in Coverage Area – GIS Data below.

3.5.4.4 Coverage Area – GIS Data

In the event that the provider supplies coverage area data in GIS format please follow the steps below:

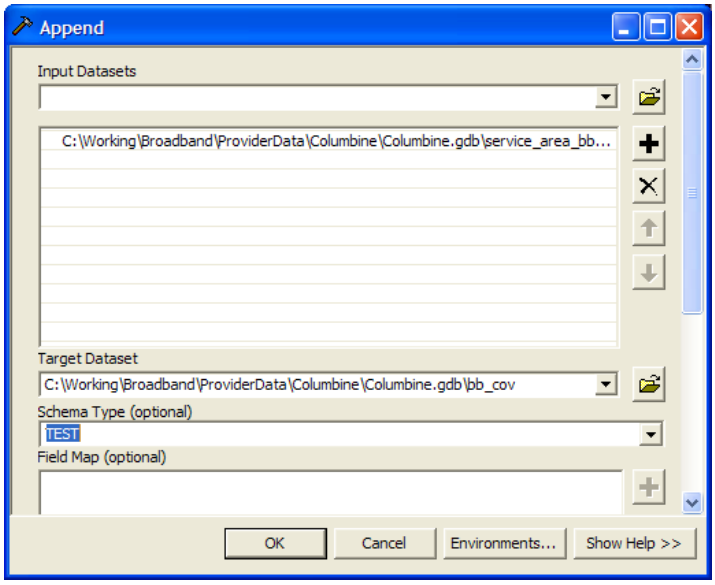
1) In ArcCatalog, create an empty feature class with the schema of the bb_cov feature class and load the GIS feature class either created above or supplied by the provider into it.

a) Right-click on the empty feature class, select Load>Load Data from the dropdown menu and navigate to the location of the service area feature class
b) Press the Add button, hit Next
c) Accept the defaults and hit Next
d) Do NOT attempt to map any fields, as seen below:



e) Press Next, then Next again, then Finish.

2) In ArcToolBox, go to Data Management Tools>General>Append
3) Append the formerly empty feature class to bb_cov, completing the dialog box as seen below:



4) Leave the Schema Type as TEST
5) Press OK.
6) In ArcMap, open up bb_cov for editing and manually input associated attribution, if necessary.

3.5.4.5 Compact Polygon From Subscriber Points

- Geo-code address list using latest state “Composite Locator”
- Verify that your geo-coded file has only one TT (Technology Type). If not export individual geo-coded layers for each Technology Type.
- For each TT check for differences in speed values or speed tiers and create separate layers for each speed value/tier.
- Clean your geo-coding results - remove any points that geo-code to accuracy levels below ZIP+4 (ZIP centroids, carrier route centroids, etc). Also, verify that outliers with acceptable accuracy levels are legitimate, i.e. fall in correct City and Zip.
- Perform spatial join between county polygons (using stcnyfips field) and the cleaned geo-coded subscriber points, in order to carry the county name and stcnty fips.
- Summarize the number of subscribers by county and use the subscriber counts by county to populate the Rate Tier table.

- Un-join the county data from the geo-code subscribers list.

- Create Compact Polygon using cleaned geo-coded layer or sub-selection of using – XtoolsPro – ConvexHull-DetailedHull option. A sub-selection of geo-coded points will be used in areas where more than one polygon will need to be created for one provider's service area.

- Evaluate output Hull carefully – looking for areas that should not be covered by hull polygon.

    - If it is determined that an area or areas should not be represented in coverage area, manually reshape hull polygon until coverage area is adequate.

    - When not obvious and as a general rule, manually resolve compact polygon when the distance between the subscriber points used to define the outer boundary of the compact polygon exceeds 5 miles . When reshaping the hull polygon, snap to the outermost geo-coded points. See figure 2 and 3 for an example.

FIGURE 2- Compact Hull: Manual Resolution Required

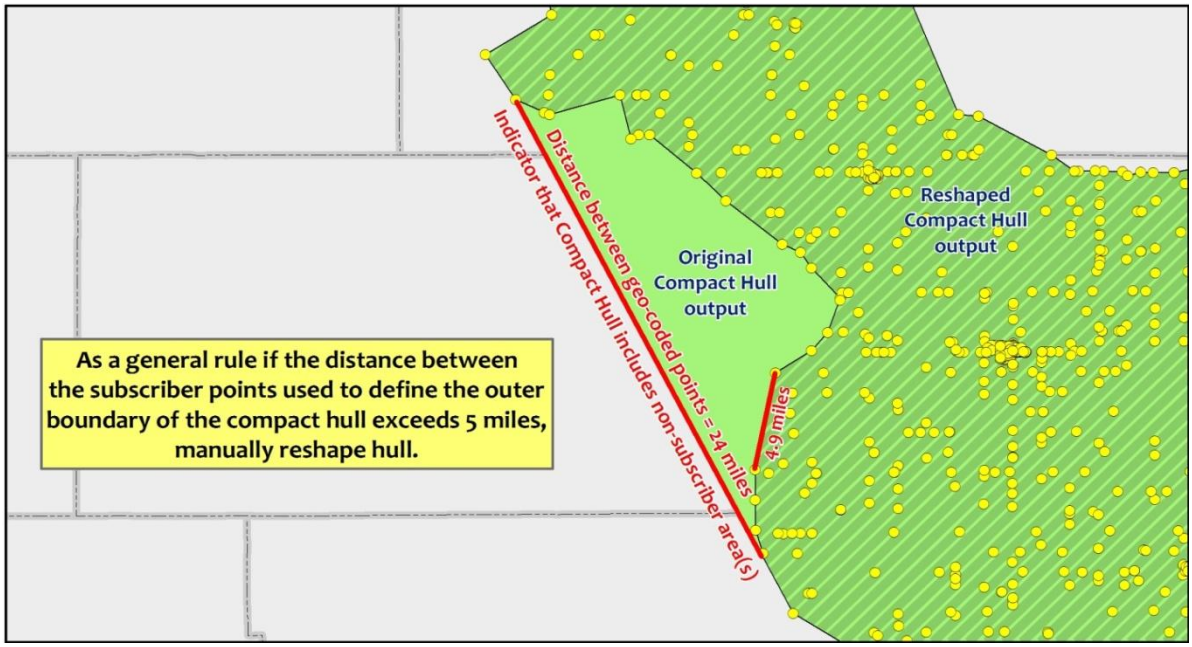


FIGURE 3a- Compact Hull: Manual Resolution Required

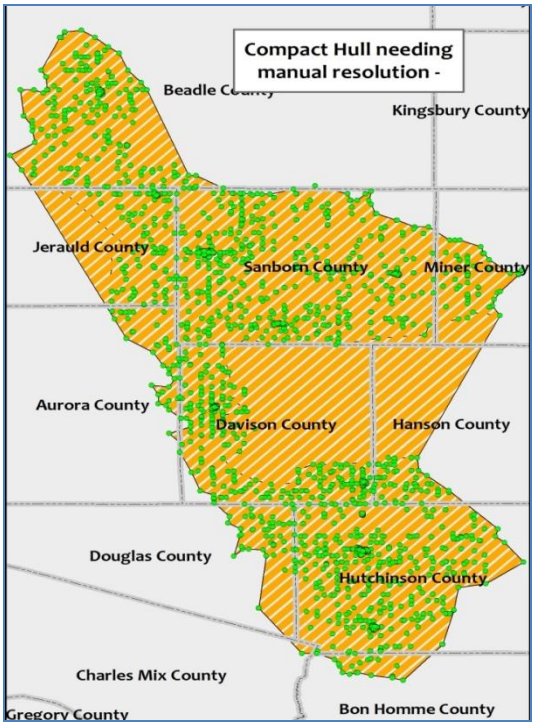
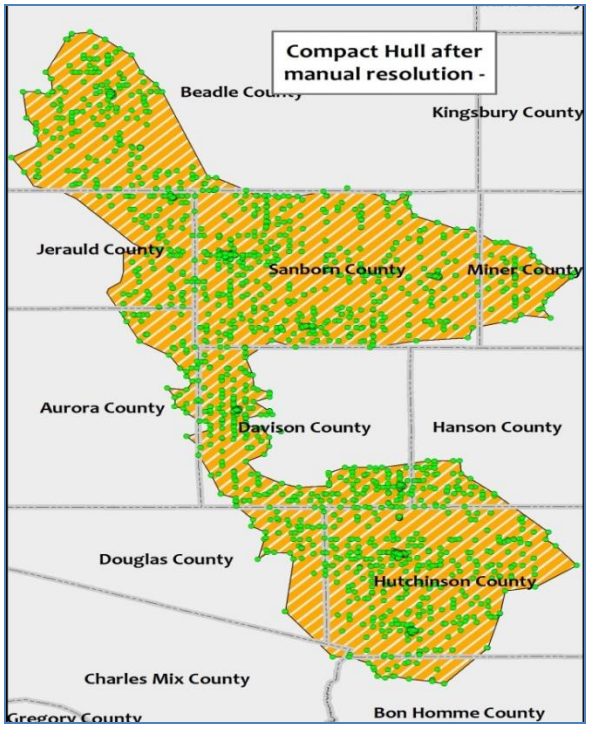


FIGURE 3b- Compact Hull: After Manual Resolution



- To attribute the compact polygon - Perform a "Spatial Join" where your Target Feature Class is the compact polygon and the Join Feature Class is your geo-coded point layer. Export compact hull with joined attributes and name file appropriately.

- Append attributed compact polygon to BroadBand TT template Feature Class and if required manually input any provider attribution that may not have carried over in the append process.

- Intersect compact polygon with county boundaries to create unique records by county and use the state-county-fips field to populate "stcty_fips" field. Also use the county name field to populate the "BBCov_Name" field.

    - Exceptions is where a provider's coverage is distributed throughout more than one area of any given county where the "BBcov_Name" should be populated using an appropriate city or other logical name based on geographical location.

- Export/Load into appropriate BB TT model Dataset.

3.5.4.6 Census Blocks – List or Spreadsheet

In the event that the provider supplies census block data in a list or spreadsheet, please follow the steps below:

1) Ensure block polygons supplied by the provider are 2000 currency
2) If other currency, convert to 2000 currency before proceeding
    a. To do this, remove the trailing letter (a, b, etc.) from the block ID
    b. You will now have two blocks that equate to one block in the 2000 block geometry
    c. Delete duplicate block IDs, retaining the higher service tier in each case
3) Prepare the block list in clean Excel format, removing all Excel-only formatting, merged cells, colors, borders, etc.
4) Import the spreadsheet into ArcMap.
5) Right-click on the 2000 census block feature class in the layer list in ArcMap and select Joins and Relates>Join from the drop down menu. Join the census block list to the 2000 census blocks feature class using the block ID and export joined records in a new feature class. The Join dialog box and process can be seen above in the TIGER Street Segments – List, Spreadsheet, or GIS Data section.
6) Follow the steps in Census Blocks – GIS Data below.

3.5.4.7 Census Blocks – GIS Data

In the event that the provider supplies census block GIS data please follow the steps below:

1) Ensure that the blocks supplied by the provider are in the required data schema and are complete as far as require attribution.
    a. If not, manually enter the required attribution or contact the provider to fill gaps.
2) If census block geometry is distributed throughout more than one county then select Data Management Tools>Generalization>Dissolve in ArcToolbox and dissolve based on County/Provider/TT/Speed Tier so that unique records are created for each unique combination.
    a. The dissolve dialog box can be seen above in the TIGER Street Segments – List, Spreadsheet, or GIS Data section.

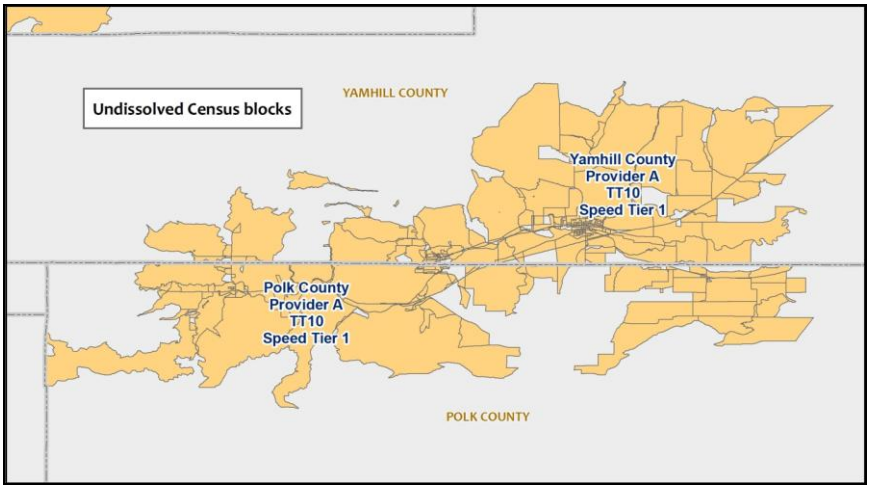


Figure 1: Undissolved census block polygons

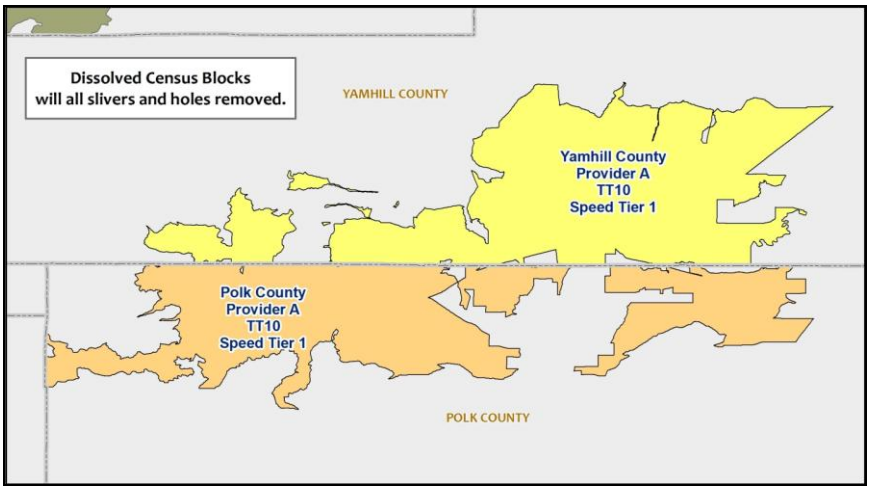
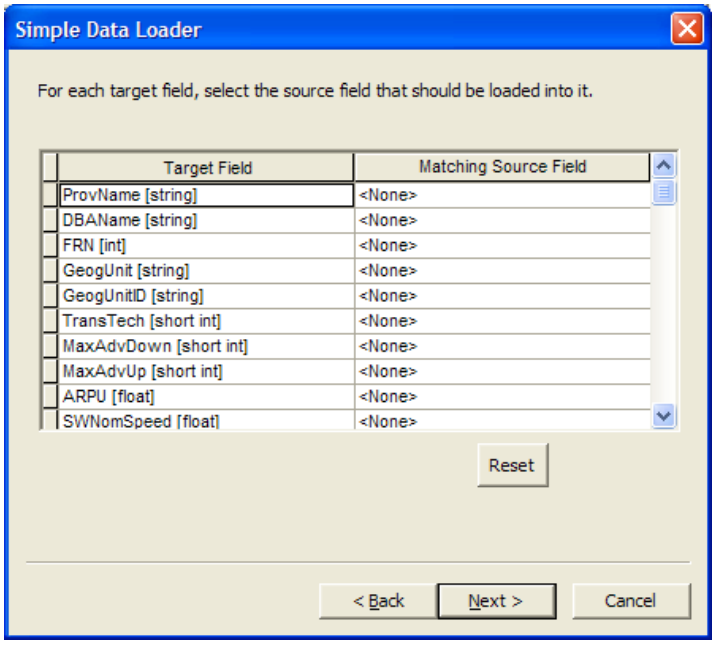


Figure 2: Census block polygons dissolved by county

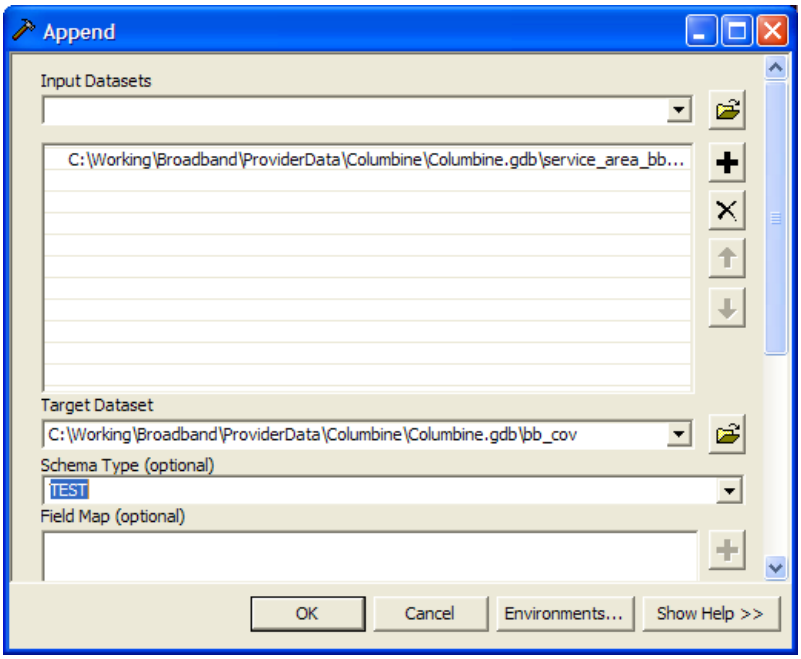
2) For each dissolved region use the Editing toolbar in ArcMap to remove unnecessary slivers and other small holes.
3) In ArcToolbox, select Data Management Tools>General>Merge and merge the processed polygons together into single layer.
4) The merged census blocks will need to have the subscriber's "frn" field added and populated.
5) In ArcCatalog, create an empty feature class with the schema of the bb_cov feature class and load the GIS feature class either created above or supplied by the provider into it.
    a. Right-click on the empty feature class, select Load>Load Data from the dropdown menu and navigate to the location of the service area feature class
    b. Press the Add button, hit Next
    c. Accept the defaults and hit Next
    d. Do NOT attempt to map any fields, as seen below:



e. Press Next, then Next again, then Finish.

6) In ArcToolBox, go to Data Management Tools>General>Append

7) Append the formerly empty feature class to bb_cov, completing the dialog box as seen below:



8) Leave the Schema Type as TEST

9) Press OK.
10) In ArcMap, open up bb_cov for editing and manually input associated attribution, if necessary.

3.6 Metadata Transactions

Following any updates or changes completed within the file geodatabase (fGDB) stored on the GIS-Analysts staging environment, the GIS-Analyst runs transactions to compare that fGDB with the one stored on the Core server to ensure metadata on all changes are recorded.

Below outlines the steps taken to run transactions on the updated Core database:

1. Open a command line window and run generateTransactions.py
    a. Usage: generateTransactions.py [Core fGDB] [Staging Environment fGDB]

    b. Example of command line:

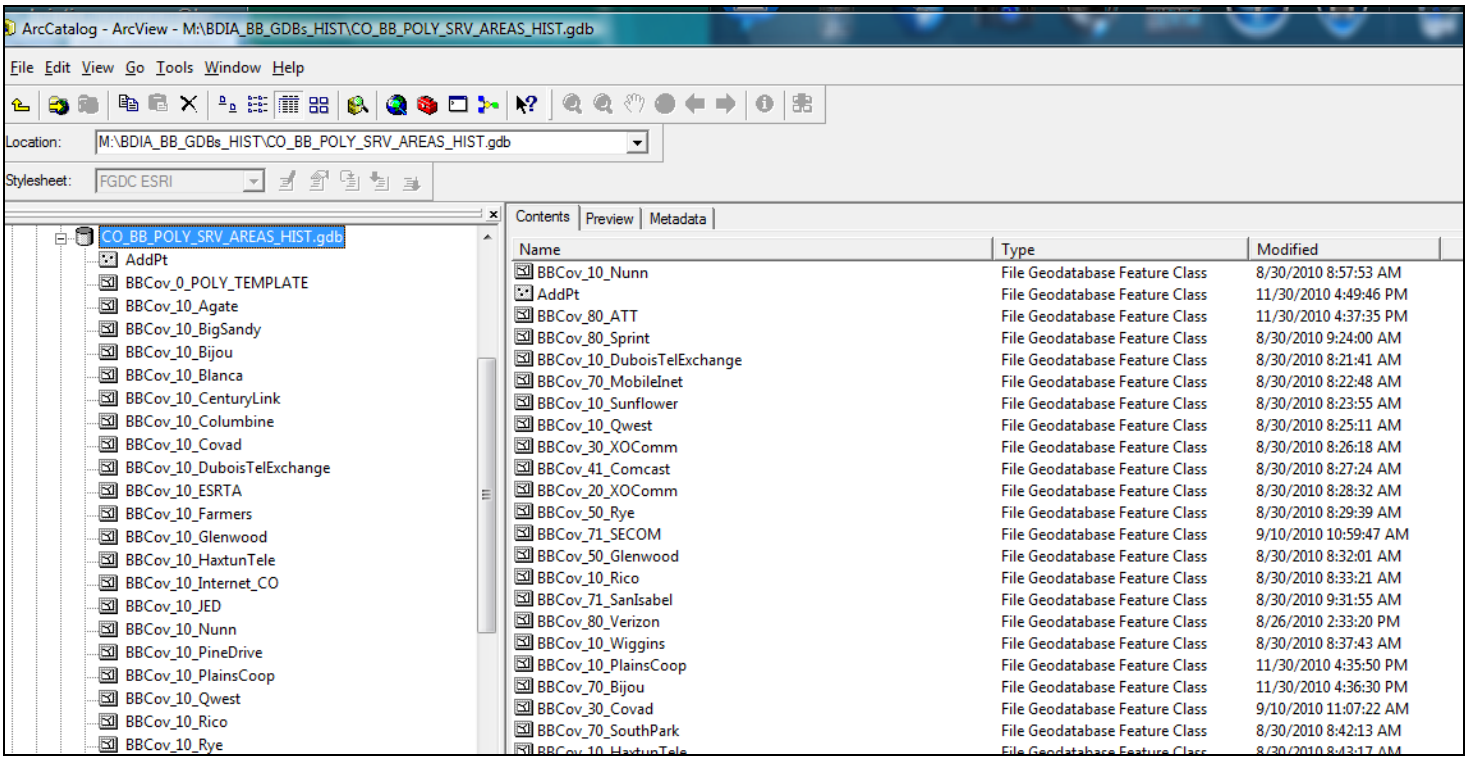
    <path>generateTransactions.py <path>ST_BB_POLY_SRV_AREAS.gdb <path>ST_BB_POLY_SRV_AREAS.gdb

2. Below is an example of the output screen that will be displayed:

```
-----------  Collecting Transactions  ----------------

Calculating rec_id field for BBCov_0_BB_POLY_TEMPLATE
value can not be 0 or less
Trouble creating the progress meter

Calculating rec_id field for BBCov_10_CenturyLink
%   10   20   30   40   50   60   70   80   90   100
----|----|----|----|----|----|----|----|----|----|     Goal = 8
```

```
Merging change: X:\BDIA_BB_GDBs\MS_BB_POLY_SRV_AREAS.gdb\AddPt
Calculating Transaction fields for AddPt
%   10   20   30   40   50   60   70   80   90   100
----|----|----|----|----|----|----|----|----|----|     Goal = 1
**************************************************
X:\BDIA_BB_GDBs\MS_BB_POLY_SRV_AREAS.gdb\AddPt...changes is complete.

Your transaction FeatureClasses are in:
\\michigan\AllAccess\BDIA_BB_GDBs_HIST\MS_BB_POLY_SRV_AREAS_HIST.gdb
------------------------------------------------------------------------------
------------
elapsed time = 2994.4 seconds
```

3. After process has completed, results can be found in the ST_BB_POLY_SRV_AREAS_HIST.gdb
    a. The transactions scripts records changes at a feature level.
    b. Below is a screen shot supporting the directory structure of the historical fGDB.



    c. Attribution associated with each added/removed/changed features is tracked, including the following additional columns appended to the end of each:
        i. Commit_by
            1. Records the GIS-Analyst that committed the changes to the historical fGDB.

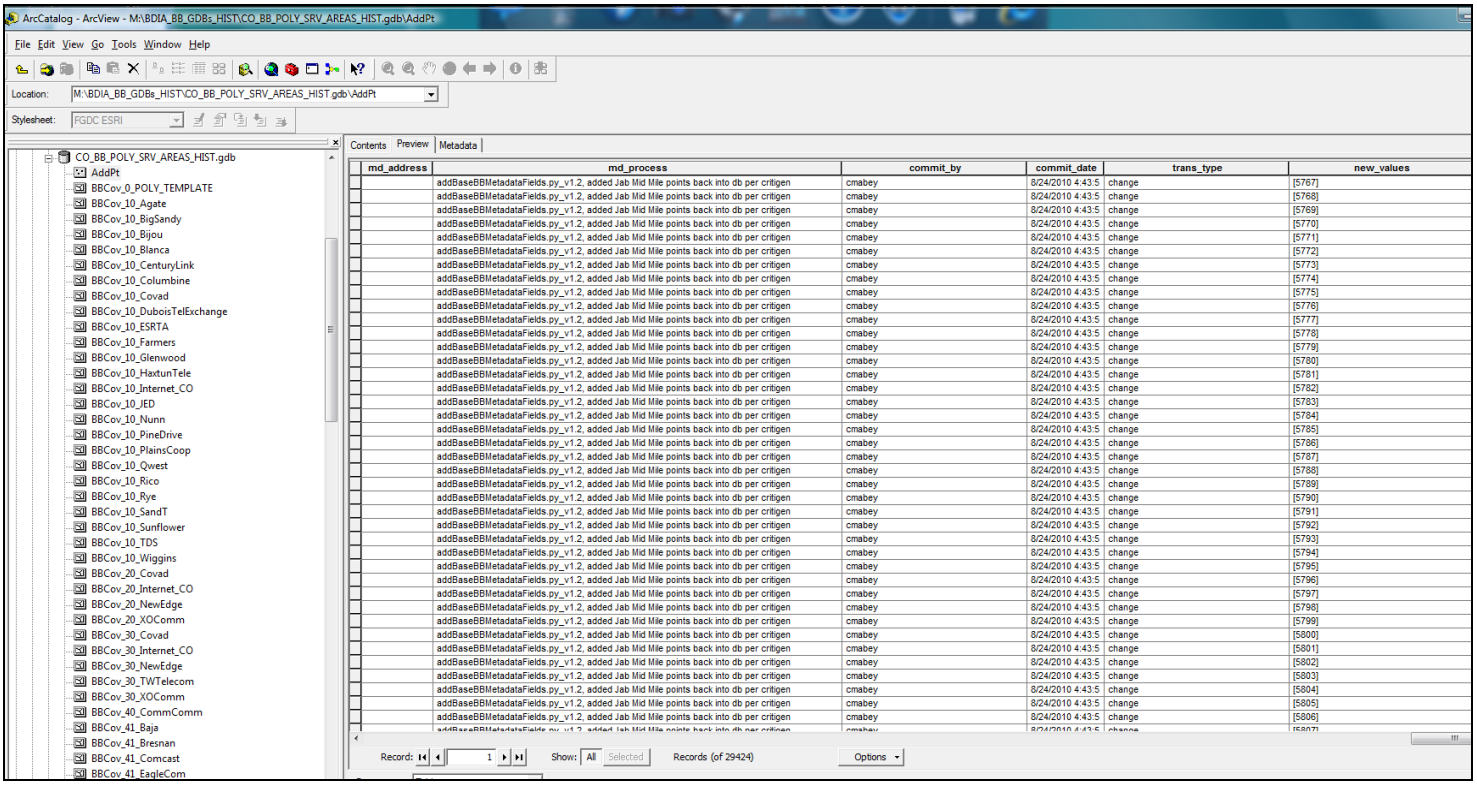
        ii. Commit_date
            1. Records the date and time stamp that the changes were committed.

        iii. Trans_type
            1. This field reflects the type of change recorded.
            2. Categorized by:
                a. Adds/Change/Deletes

        iv. New_values

1. Records the new values when a change was completed on a feature. Example: Name or speed change

d. MD_Process is also transferred from the edited fGDB to the historical fGDB, which states the actions completed by the GIS-Analyst.



3.7 Data Processing

3.7.1 Data Processing Overview

The following items outline the actions required to process the service provider data further to meet the NTIA requirements.

- Weighted Nominal Speed
- Middle Mile
- Broadband Coverage Template

3.7.2 Weighted Nominal Speed

The weighted nominal speed is populated one of the following two ways:

3.7.2.1 Subscriber Data Supplied by Provider

Where we are supplied with subscriber speed information by the data provider, we use the following formula from the NOFA:

$$\frac{\text{(speed tier-1 in kbps} \times \text{no. of tier-1 subscribers) + (speed tier-2 in kbps} \times \text{no. of tier-2 subscribers) + (etc.)}}{\text{Total average monthly subscribers}}$$

Data is initially broken up in the following order:

1) Stcty_fips
2) Transmission technology type
3) Subscriber tiers

3.7.2.2 Value Supplied by Provider

Some providers will supply their weighted nominal speed. In these cases, the data supplied will be populated instead of using the NOFA formula.

When these values have been obtained or calculated, they are used to update the service overview layer. This can be done manually or by creating a table with the provider's FRN and average weighted speed and joining it to the service overview table in ArcMap. To Join, right-click on the layer you would like to join to and select Joins and Relates>Join… from the dropdown menu. Then navigate to the table you want to join and select the join fields from the drop down lists. Then open up the source table (the table in ArcMap) and right-click on the header of the Average Weighted Speed field and select Calculate Field from the drop down menu. Use the value of the average weighted speed from the joined table.

3.7.3 Middle Mile

Middle mile information is generally provided in spreadsheet or text file format. The process is to take what is supplied by the provider and translate it into the required data schema.

1) If the data is supplied with address information, follow the process outlined above in Subscriber Location – Address Data.

2) If the data is supplied with associated XY coordinates, follow the process outlined above in Subscriber Location – XY Data.
3) Once the data is in GIS format, use the Append (Data Management Tools>General>Append) command in ArcToolbox to append the data to the overall middle mile dataset.
4) Set the schema type to NO_TEST and use the Field Map to map the attribute fields from the source to the target dataset.

3.7.4 Broadband Coverage Template

Below is the description of the fields within the BB_Cov layer, which is the interim data set that is used to create the final product deliverable.

Name	Alias	Description
objectid	OBJECTID	Internal Object ID
shape	SHAPE	Internal Shape storage
prov_id	PROVIDER_ID	Unique numeric identifier for each provider
prov_name	PROVIDER_NAME	Unique name for each provider
dba_name	DOING_BUSINESS_AS	An alternative "Doing-Business-As" name for the provider
frn	FCC_REGISTRATION_NUMBER	Provider FCC Registration Number
bbcov_name	BBCOV_NAME	BroadMap Broadband Coverage name
trans_code	TRANSMISSION_CODE	Unique code for the transmission technology type described by this layer
trans_name	TRANSMISSION_NAME	Name for the transmissions technology type
trans_desc	TRANSMISSION_DESC	Description for the transmissions technology type
spect_code	SPECTRUM_CODE	Unique code for the spectrum [WIRELESS ONLY]
spect_name	SPECTRUM_NAME	Name for the spectrum [WIRELESS ONLY]
spect_desc	SPECTRUM_DESC	Description for the spectrum [WIRELESS ONLY]
mad_dwn_t	MAX_AD_DOWN_TIER	Maximum advertised downstream speed available within given area (speed tier)
mad_up_t	MAX_AD_UP_TIER	Maximum advertised upstream speed available within given area (speed tier)
typ_dwn_t	TYPICAL_DOWN_TIER	Typical downstream speed available within given area (speed tier)
typ_up_t	TYPICAL_UP_TIER	Typical upstream speed available within given area (speed tier)
mad_dwn_k	MAX_AD_DOWN_KBPS	Maximum advertised downstream speed available within given area (kbps)
mad_up_k	MAX_AD_UP_KBPS	Maximum advertised upstream speed available within given area (kbps)

Name	Alias	Description
typ_dwn_k	TYPICAL_DOWN_KBPS	Typical downstream speed available within given area (kbps)
typ_up_k	TYPICAL_UP_KBPS	Typical upstream speed available within given area (kbps)
subs	SUBSCRIBERS	Total average monthly subscribers for this provider for this technology for this coverage polygon
md_geom	MD_GEOMETRY	Metadata: Comma separated list of source id's from which the polygon extent was produced
md_exists	MD_EXISTS	Metadata: Comma separated list of source id's used in understanding and editing the provider data for this polygon
md_who	MD_WHO	Metadata: Name of the editor who last edited this feature at the time in md_when
md_when	MD_WHEN	Metadata: Date/timethat this feature was last edited
md_process	MD_PROCESS	Metadata: Comma separated list of processed used to create and/or modify this layer
stcty_fips	STATE_COUNTY_FIPS	State/County FIPS code
rec_id	RECORD_ID	Compound Key formed from STCTY_FIPS+"|"+Provider_ID+"|"+Trans_Code+"|"+BBCov_Name
st_area	ST_AREA(SHAPE)	Area in square decimal degrees
st_length	ST_LENGTH(SHAPE)	Length in decimal degrees
Provider_Type	Type of Provider	Has Subtype (1:Broadband provider as described in the NOFA,2:Reseller,3:Unknown), default value = 1 (New 04/11 Model)

3.7.5 Verification and Validation

3.7.5.1 Provider Validation – Provider Portal/PDF Map Review

Following the collection and aggregation of provider data, the data is then validated by the provider to ensure the data aggregated is an accurate representation of their coverage area and supporting broadband information. This is completed through the Provider Portal web application, which is a secure interactive map displaying their coverage areas and allows the user to validate, submit feedback or request changes. If changes are requested, then the features on the portal are then updated and an automatic request is sent to the provider to complete the validation effort.

For some providers that did not use the Provider Portal, a PDF was sent displaying their coverage map and validation was then completed via e-mail notification.

3.7.5.2 Provider Verification – 3rd Party Source Review

Once the provider has validated their coverage areas, a 3rd party source comparison and analysis is then performed. Where anomalies or discrepancies are identified, a 'SCAN' point is dropped and descriptive comments applied so they can later be reviewed with the provider.

During the provider review, the map is displayed along with the 'SCAN' points and potential refinement is completed based on input from the Provider.

3rd Party Sources Utilized

3rd Party Source Name	Source Type	Verification Type
InfoUSA	Consumer and Business Listings	Community Anchor Institutions Can also be used for demographic information supporting the State websites
Pitney Bowes (PBBI)	Exchange Info Plus (Central Office Locations)	Exchange datasets are used to verify the following Transmission Technologies (TT): Asymmetric xDSL (10), Symmetric xDSL (20), Other Copper Wireline (30), and Optical Carrier/Fiber to the End User (50).
Media Prints	Cable Boundaries	Used to verify the following TT: Cable Modem—DOCSIS 3.0 (40) and Cable Modem—Other (41)
American Roamer	Wireless Coverage Patterns (EVDO, GPRS, WISP, HSPA)	Used to verify the following TT: Terrestrial Fixed Wireless—Unlicensed (70), Terrestrial Fixed Wireless—Licensed (71) and Terrestrial Mobile Wireless (80)
ComSearch	Wireless Spectrum Holdings and Tower Data	Used to verify the following TT: Terrestrial Fixed Wireless—Unlicensed (70), Terrestrial Fixed Wireless—Licensed (71) and Terrestrial Mobile Wireless (80)

3.7.5.3 Assigning Confidence Values

All efforts from the above-mentioned validation and verification activities, plus internal peer quality reviews are combined and tracked in a Validation table. Based on the results of this analysis, a confidence value is assigned for each provider and then each technology.

The confidence values are as follows:

0 = Coverage area has not been reviewed
10 = Extremely Low. Single Source QC.
20 = Very Low. Needs Additional Validation\Verification
30 = Low. Even with Validation\Verification, Coverage is still suspect.
40 = Acceptable, confirm with State prior to shipment.
50 = Meets requirements to be included in shipment.
60 = Moderate. Meets NTIA/State's standards, representative of Technology Type (TT)
70 = High. Accurate representation of coverage based upon TT.
80 = Very High. Multiple validation\verification with most 3rd party sources
90 = Extremely High. Multiple validation\verification sources
100 = Perfect. Multiple validation\verification sources, with complete alignment with sources and ground truth verification activities

This Validation table is then maintained as updates or changes occur for each provider, down to technology type, with the overall goal to improve the confidence values and overall map representation.

Example of the Validation table:

Contents | Preview | Metadata

OBJECTID *	BBCOV	CONFIDENCE_CODE	PROVIDER_ID	PEER_QC	PROVIDER_QC	THIRD_PARTY_VERIFICATION	THIRD_PARTY_ID
1	BBCov_10_Axxis	40	771	11/4/2010	9/27/2010	11/4/2010	3070
2	BBCov_10_BeaverTelCo	80	850	10/18/2010	3/9/2011	6/7/2010	2010
3	BBCov_10_CanbyTelcom	80	706	10/18/2010	9/21/2010	6/7/2010	2010
4	BBCov_10_CascadeUtil	70	3005	11/4/2010		11/4/2010	3070
5	BBCov_10_CenturyLink	70	710	11/4/2010	9/23/2010	11/4/2010	3070
6	BBCov_10_ColtonTel	80	713	11/4/2010	9/16/2010	11/4/2010	3070
7	BBCov_10_Covad	60	717	11/4/2010	9/23/2010	11/4/2010	3070
8	BBCov_10_DataVision	30	767	11/4/2010		11/4/2010	3070
9	BBCov_10_EasternOregonTelcom	60	899	11/4/2010	9/20/2010	11/4/2010	3070
10	BBCov_10_Frontier	70	764	11/4/2010	9/16/2010	11/4/2010	3070
11	BBCov_10_Gervais	90	767	10/18/2010	9/22/2010	6/7/2010	2010
12	BBCov_10_Helix	70	726	11/4/2010	9/22/2010	11/4/2010	3070
13	BBCov_10_Integra	30	790	10/18/2010	9/27/2010	6/7/2010	2010
14	BBCov_10_McMinnville	60	732	11/5/2010	9/27/2010	11/5/2010	3070
15	BBCov_10_Molalla	50	734	10/18/2010	9/8/2010	6/7/2010	2010
16	BBCov_10_MonitorCOOP	70	1100	10/18/2010	9/17/2010	6/7/2010	2010
17	BBCov_10_Monroe_Telephone	80	736	10/18/2010	9/20/2010	6/7/2010	2010
18	BBCov_10_MtAngel	90	707	10/18/2010	3/9/2011	6/7/2010	2010
19	BBCov_10_Nehalem	80	795	10/18/2010	9/28/2010	6/7/2010	2010
20	BBCov_10_NorthStateTel	40	738	3/15/2011	3/15/2011	11/5/2010	3070
21	BBCov_10_OregonTeleCo	20	739	11/5/2010	9/14/2010	11/5/2010	3070
22	BBCov_10_People	80	1012	11/5/2010	9/17/2010	11/5/2010	3070
23	BBCov_10_PineTelephone	70	757	10/15/2010	3/17/2011	6/9/2010	2010
24	BBCov_10_Pioneer	70	740	11/5/2010	9/20/2010	11/5/2010	3070
25	BBCov_10_Qwest	80	1102	11/8/2010	5/7/2010	11/8/2010	3070
26	BBCov_10_Rionet	50	807	11/8/2010	9/27/2010	11/8/2010	3070
27	BBCov_10_Roome	90	746	10/18/2010	9/10/2010	6/7/2010	2010
28	BBCov_10_Sandy	60	873	11/8/2010	9/17/2010	11/8/2010	3070
29	BBCov_10_Scio	90	800	10/15/2010	3/17/2011	6/9/2010	2010
30	BBCov_10_SCS	60	1030	11/8/2010	9/17/2010	11/8/2010	3070
31	BBCov_10_SCTC	70	803	10/18/2010	9/17/2010	11/10/2010	3070
32	BBCov_10_StPaulTel	80	750	3/15/2011	3/15/2011	6/7/2010	2010
33	BBCov_10_TDS	40	752	10/18/2010		6/7/2010	2010
34	BBCov_10_TransCascade	40	709	11/8/2010	9/21/2010	11/8/2010	3070
35	BBCov_20_CanbyTelcom	80	706	10/18/2010	9/21/2010	6/7/2010	2010
36	BBCov_20_ClearCreek	80	712	10/18/2010	9/17/2010	6/7/2010	2010
37	BBCov_20_Covad	60	717	11/4/2010	9/23/2010	11/4/2010	3070
38	BBCov_20_Integra	30	790	10/18/2010	9/27/2010	6/7/2010	2010
39	BBCov_20_NewEdge	20	796	11/8/2010		11/8/2010	3070
40	BBCov_20_QuantumComm	60	1021	11/8/2010	9/23/2010	11/8/2010	3070
41	BBCov_20_Rionet	50	807	11/8/2010	9/27/2010	11/8/2010	3070
42	BBCov_30_CanbyTelcom	80	706	10/18/2010	9/21/2010	6/7/2010	2010
43	BBCov_30_Covad	60	717	11/4/2010	9/23/2010	11/4/2010	3070
44	BBCov_30_Integra	30	790	10/18/2010	9/27/2010	6/7/2010	2010
45	BBCov_30_Lightspeed	20	793	11/8/2010		11/8/2010	3070

Record: 1 | Show: All | Selected | Records (of 122) | Options

Preview: Table

3.7.6 Community Anchor Institution (CAI) Data

3.7.6.1 Data Collection

The CAI data was initially collected from the State to create the baseline inventory. All location information and broadband coverage data supplied was also ingested into the data deliverable.
Additional collection of CAI information was done via data mining and/or webscraping to build out the inventory further. For example: Collection of additional CAIs, address and broadband data.

The state-agency-provided CAI inventory was comprehensive but the challenge is collecting broadband related data; service provider(s), technology and speed data for each CAI. Availability of the CAI portal has not significantly increase submission of this data. Additional promotion to CAIs to utilize the CAI portal will be needed to increase this data for subsequent deliverables.

3.7.6.2 Institution Data

Institution data is obtained from a variety of sources and almost always provided in Excel spreadsheet format. The general process for incorporating this data is below:

1) If the data is provided in Excel or some similar format:

    a. Clean and standardize the Excel spreadsheet, removing any cell formats, merged cells, etc.

b. Standardize the address format as defined in the staging CAI database

c. If the spreadsheet includes X and Y values, such as latitude and longitude, use the Add XY Data tool in ArcMap to create a spatial data layer.

d. If there are only addresses, then follow the geocoding steps outlined above to create spatial data points for each of the institutions.

    i. Institutions that do not geocode based on the TIGER 2009 data set will have to be manually located using Google Maps, Google Earth, or some other information source.

2) If the CAI source data is in GIS format, add the Latitude and Longitude fields and use the Calculate Geometry tool to populate them, using the WGS 84 coordinate system.

3) Using ArcCatalog, load the new data into the staging CAI database.

4) This database is ready for the makeDeliverable.py script to process the information into the final state and NTIA deliverables.

3.7.6.3 Community Anchor Institution (CAI) Portal Updates

A web application has been released to allow for further data collection and validation of anchor institution location information, broadband coverage, and speed test data.

Information collected from the CAI Portal is then ingested into the overall inventory and will later be compared against the provider coverage areas mapped for any potential discrepancies.

3.8 Product Extract

3.8.1 Python Scripts

The following sections make use of Python scripts. In general, to use a Python scrip, you must have Python installed on your computer. To download the latest version of Python, go to http://www.python.org/download/ and download the latest stable version. As of August 2010, this was version 2.7. Once this is installed, the general way to run a script is to type the following at a command prompt: C:\Python27\python.exe C:\<location of script>. Many of the scripts provided have environment variables that must be set before they can be run.

The python code for BroadMap's product extract has been incorporated into a Hudson CI System, which is detailed in the Process Operation and Monitoring section of this document. This was a process improvement activity so all processes can be monitored, controlled and contain historical tracking on each process.

3.8.2 Product Extract Process

Note: specific Python scripts are called out in red font in the sections below.

The MapConnect product extract process, makeDeliverable.py, uses the BB_Cov and BROADMAP_POINTS interim data sets to create the following layers according to the current specifications:

- BB_Service_Road_Segment
  - This layer contains all broadband services associated with specific street segments for census 2000 blocks larger in area than two square miles
- BB_ServiceCensusBlock
  - Contains all broadband services associated with census blocks of no greater than two square miles.
- BB_Service_Wireless
  - This layer contains all wireless services not associated with specific addresses.
- BB_ServiceOverview
  - This layer contains subscriber-weighted nominal speed for each provider's service area at a county level and is meant to act as a summarized view.
- BB_ConnectionPoint_MiddleMile
  - This layer contains middle-mile and backbone interconnection points
- BB_Service_CAInstitutions
  - Broadband Service at Community Anchor Institutions (CAI)
    - Community Anchor Institutions consist of schools, libraries, medical and healthcare providers, public safety entities, community colleges and other institutions of higher education, and other community support organizations and entities.

Due to a NTIA model change for the October 2010 data deliverable, an addition to this code was created to support both models in the case a comparison is later desired or a request is made to revert back to the original model. This script name is bdia2ntia.py and creates the following layers in addition to the layers mentioned above, rolled up to **NATL_Broadband_Map.**

- BB_ConnectionPoint_LastMile
  - This layer contains last mile infrastructure points, which is only populated if data cannot be provided at a more granular level.
- **BB_Service_Address**
  - **Represents broadband availability for service address points.** Address Point availability refers to those individual addresses at which each facilities-based provider of broadband service can provide broadband services of minimal characteristics within 7 - 10 business days.
- State_Boundary
  - State boundary supporting topological validation of point feature classes.
- **NATL_Broadband_Topology**
  - Supports basic topology quality checking. Example: No CAI's or Middle Mile points outside of the state boundary

The following process flow provides a view of how the Core fGDB is extrapolated to the NTIA final deliverable via the makeDeliverable.py script. Following that, the bdia2ntia.py script is run, which limits what's placed in the final layers based on the NTIA modeling standards.

The product scripts and supporting extract were originally created separately per request, in case data model comparisons were to be completed.

3.8.3 Product Statistics

Following the completion of a product extract, the product statistics script (BDIA_ReleaseNotesStats.py) extracts the following information supporting that product deliverable.

- Provider Statistics
  - Collects all provider information, listing by Provider Name
  - Provides output of FRN
  - Counts the number of features supported within the following layers:
    - Census Block
    - Street Segment
    - Max Upstream
    - Wireless Services
    - Infrastructure Points
  - These updates were made to support the Data Package required to accompany every NTIA product deliverable.
- Community Anchor Institution (CAI) Statistics
  - Breaks CAI down to the 8 categories
    - 1: School: K through 12
    - 2: Library
    - 3: Medical/Healthcare
    - 4: Public Safety

- 5: University/College
- 6: Other Government
- 7: Other Community non-government
- None: Unknown Category
  - In cases where this occurs, further investigation is completed prior to product shipment to ensure all CAI's are categorized accurately

- Reports out the following counts
  - Total CAIs within that category
  - Total CAIs that contain partial BB coverage
    - Contains any of the following information for given CAI:
      - BB Subscriber, Transmission Technology, Speed Down Speed Up
  - Total CAIs that contain full BB coverage
    - Contains <u>all</u> of the above-mentioned BB information for given CAI.

The output of this script is two CSV files: AnchorInstitutions.csv and Providers.csv. These files can then be inspected to ensure that there are the expected number of CAIs and providers for every release.

3.9 Quality Assurance

Quality assurance is supported manually and algorithmically on the interim data, BB_Cov file geodatabase, and on the final product. For scheduled product releases, a test product extract and subsequent manual and algorithmic QC run is completed along with a release review. The product specifications, project status reports, previous product release notes are used as references throughout this review.

The following parameters are tested using the methodology listed below each:

- Product Deliverable Format
  - Correct names and format of data deliverables
    - BDIA_QC_SUITES (please see below for details)
  - Correct Projections/Datum
    - Manual interaction with product
  - Metadata Present and Correct
    - Manual interaction with product
- Table Structure
  - All required tables included
    - BDIA_QC_SUITES
  - Extraneous tables identified
    - BDIA_QC_SUITES
- Field Structure
  - All fields included
    - BDIA_QC_SUITES
  - Extraneous fields identified
    - BDIA_QC_SUITES

- Correct field names, types and widths
  - BDIA_QC_SUITES
- Field Domains
  - Values in all tables are constrained to the specified values specified
    - This action is accomplished via BDIA_QC_SUITES and manual review of the product
    - This tends to identify project completeness issues as fields with a null value are identified.
- Geometric Representation
  - Identify if all layers have the correct geometric representation
    - Manual review of the BB_ServiceOverview layer
    - Dependent on NTIA and client requirements
- Geographic Extent
  - Product includes the necessary Geography associated with Product?
    - Manual Review - ArcGIS
  - Is there extraneous geography included in Product?
    - Manual Review - ArcGIS
- Completeness
  - Products contain the expected amount of data?
    - Manual review of product stats relative to weekly State reports and defined expectations.
- Accuracy
  - Product meets the stated accuracy requirements for the deliverable?
    - Sampling procedure to manually review source material to resulting product
    - Provider Validation
    - Verification using 3rd Party Data
    - Verification against reality, where applicable
- Data Regression
  - Any unexplainable data loss or change?
    - This action is accomplished by comparing results within product statistics script (BDIA_ReleaseNotesStats.py) from previous releases, as well as manual review of the product
- Confidentiality
  - Any unauthorized confidential information included in the delivery?
    - Review of NDAs and delivery expectations
- Prior Issues Resolved
  - Have expected internal issues been resolved?
    - Manual review of data against previous product release notes
  - Have agreed upon customer issues been resolved?
    - Manual review of data against previous product release notes, status report and client feedback
- Delivery Medium
  - Has the product medium been verified?
    - Manual review
  - All files present
    - Manual review of SFTP site to ensure all files are copied correctly, including file/directory size

- Correct location
  - Manual review – confirmation of SFTP link, username and password

3.9.1 QC Suite

The BDIA_QC_SUITES consists of four main types of scripts supporting the overall QC process. These scripts are all run in concert and are called from the test_runner script and the test_BDIAProductGDB script.

3.9.1.1 Configuration

These scripts establish the configuration for the test_BDIAProductGDB script which is the core of the QC Suite.

- update_test_config
- active_config
- config_PROCESS01_automated
- config_PROCESS01_manual
- set_active_config

3.9.1.2 Libraries

These scripts provide additional functionality that is called from with the test_BDIAProductGDB script.

- bb_unittest_fixture
- bbcov_structure
- BC_XmlWriter
- file_folder
- search_and_replace
- unittst_fixture
- validate_BB_DB
- validate_BB_GDB
- xmlrunner_gui

3.9.1.3 QC Suite

This is the core script for performing automated QA/QC on the interim and final data deliverables.

- test_BDIAProductGDB

3.9.1.4 Other

These scripts perform other functions detailed below:

- test_runner – this is the main script that runs all the other QC scripts and imports all the necessary scripts and libraries
- which_build – this determines the current build and passes information to the configuration scripts

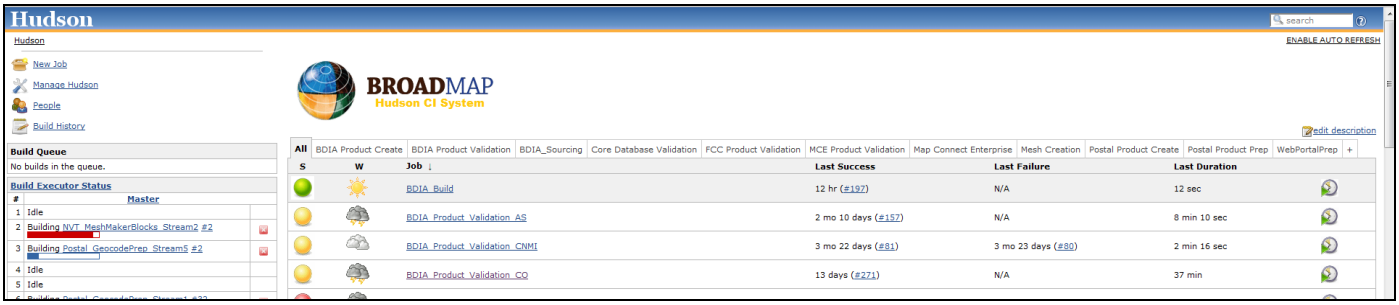
3.10 Process Operation and Monitoring

Product Extract, makeDeliverable.py and bdia2ntia.py, is run within BroadMap using a platform called Hudson that has been enhanced to support BDIA product extraction, process monitoring, as well as product validation. The same platform can be planned for implementation for the State, if desired.

Below are examples of the product create, product validation, product statistics and monitoring processes which are managed within the BroadMap Hudson CI-System. All of the above-mentioned python scripts, with the exception of metadata transactions script, are run via this system.

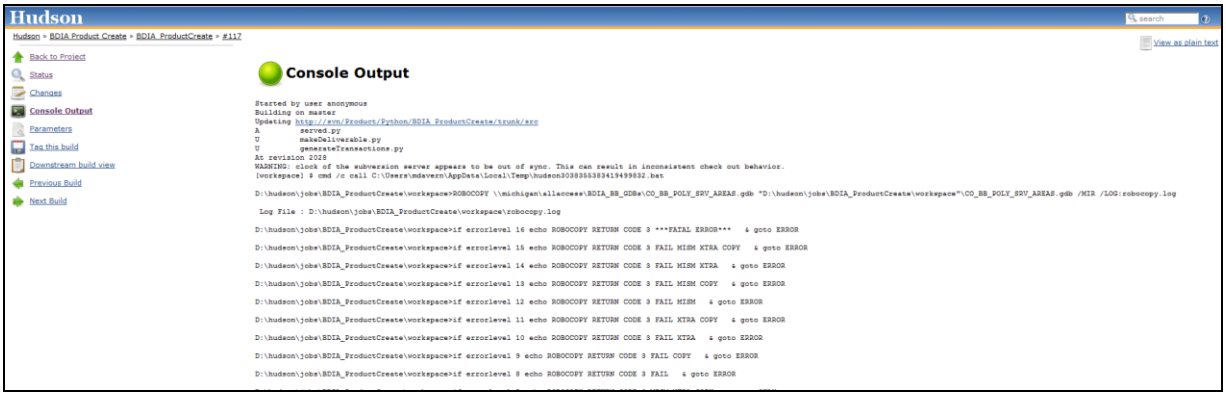
3.10.1 BDIA Product Create

Below is an example of the main page where the type of product build can be selected.

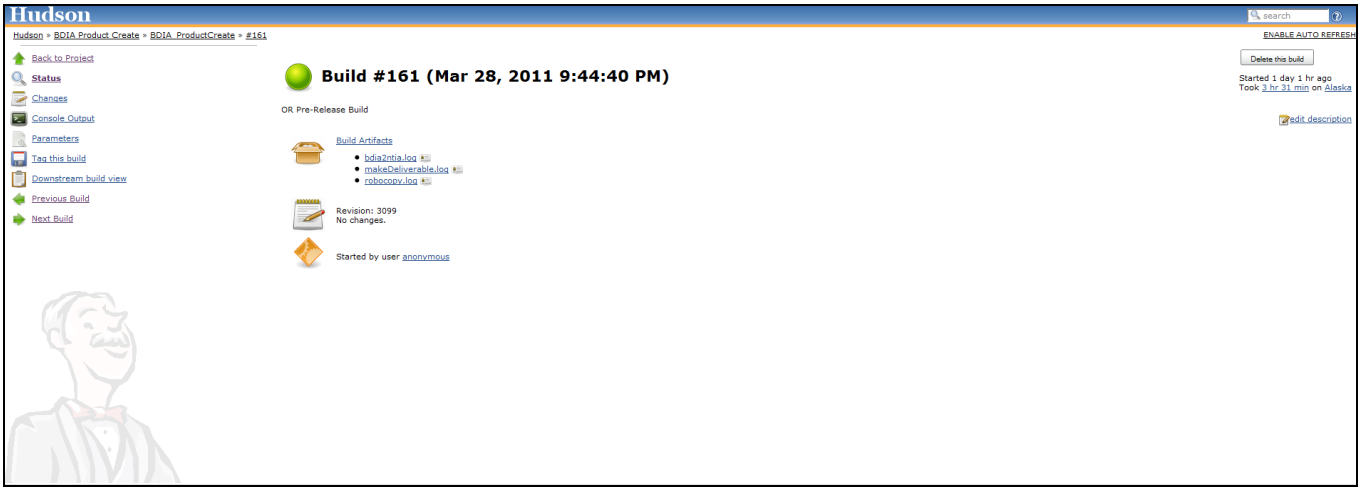


Selecting based on the type of process that will be initiated.

The Console Output can be reviewed to see the progress of product create. Following the completion of each product creation process, an e-mail notification is automatically sent to the team.



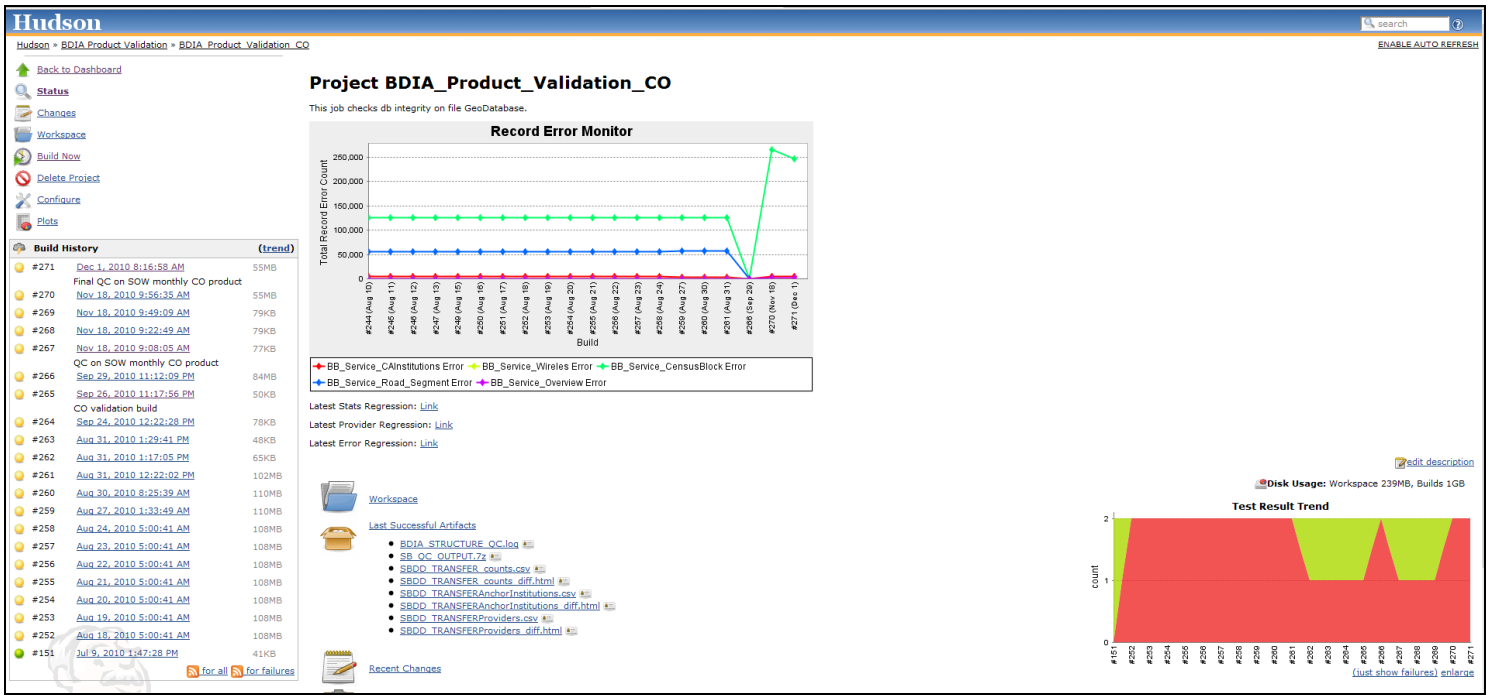
All processes run via the BroadMap Hudson CI-System are stored for historical reporting. Each process can be reviewed, including the Console Output and Build Artifacts from that run.



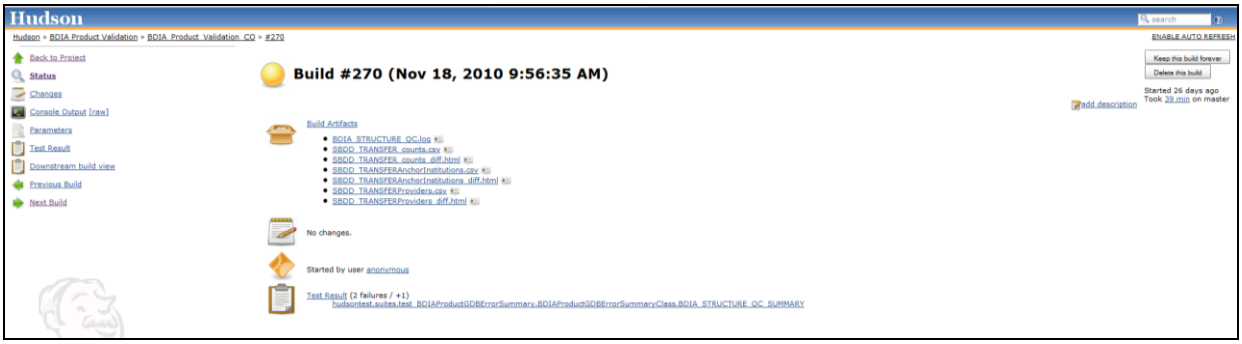
3.10.2 Product Validation and Statistics

Once the product creation process is complete, Product Validation and Statistics are then initiated. These support the BDIA_ReleaseNotesStats.py script and the BDIA_QC_SUITES scripts detailed above.
All statistics and reports are stored for historical review with the capability to place violation criticality on each quality control check allowing the identification of errors due to project status/completeness verses project correctness. Example: Typical Speeds populated.

Below is an example of the report provided based on various control points running over a specified time period:

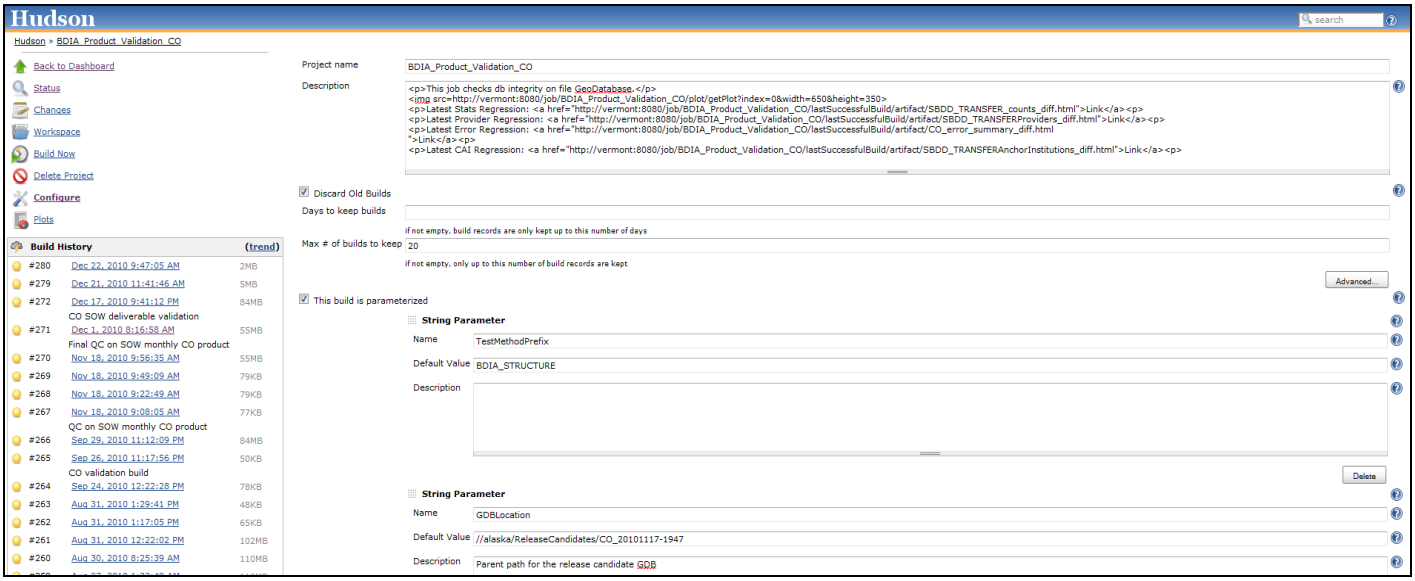


Similar to the Product Create process, all results from the process are maintained:



Results are then reviewed manually to ensure no errors reported are critical or in violation of the NTIA data model or project completion statements. Any errors of concern are communicated ahead of product delivery and included within the product release notes.

Further detail on the Hudson CI System environment can be found by navigating to the following link:
http://wiki.hudson-ci.org/display/HUDSON/Meet+Hudson



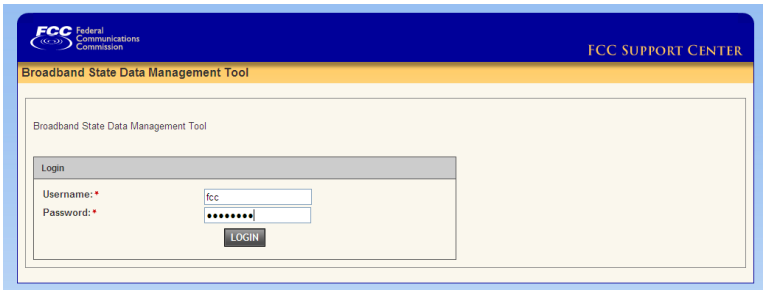
3.11 Product Extract Data Delivery

Product delivery for MapConnect Broadband is handled two ways, depending on client requirements:

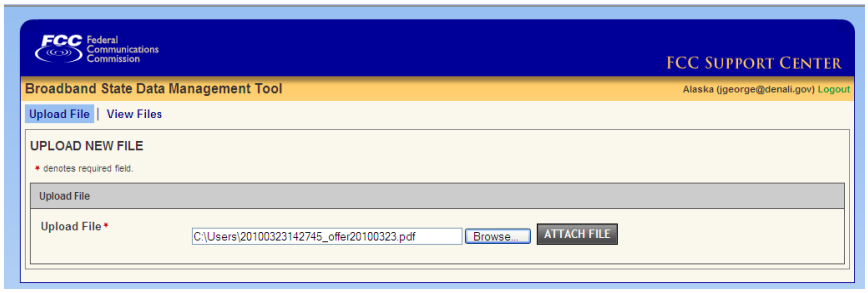
1) State Submittal
   a) Data is submitted via SFTP site
   b) Product Release Notes and QC Test Report accompanies the delivery

2) NTIA Submittal

   a) Directions for using the NTIA State Broadband Data file submission tool
   b) Go to the following WWW web site: https://esupport.fcc.gov/statedata
   c) Enter your username and password as provided to you from the NTIA program administrator.



   d) Click in Upload a file field
   e) Browse to local file for submission using the 'Browse' button. Select file then select ATTACH FILE.



f) Logout / Receipt using the Logout button in the Top Right of the screen

g) A receipt of submission is emailed to username e-mail address