North Carolina Data Submission
October 2011

Data Collection Methodology

The e-NC Authority

10/1/2011

Table of Contents

Executive Summary

The e-NC Authority

The e-NC Authority, created by the N.C. General Assembly under Session Law 2003-425, is dedicated to growing local-level wealth and creating jobs and educational opportunity through increased broadband deployment. Mandated as the primary Internet policy and planning body for the state, e-NC works with citizens, broadband service providers, local and state government and partners across the state. Its responsibilities include:

- Serving as the Broadband Authority for the state, with a focus on rural and urban-distressed areas;
- Conducting research to help guide the state in economic development decision-making;
- Mapping of broadband infrastructure in North Carolina per the requirements of the National Telecommunications & Information Administration (NTIA);
- Providing Technical Assistance to communities and organizations;
- Responding to citizen inquiries;
- Facilitating local-level programs on technology-based economic development (i.e. the e-NC Business & Technology Telecenters); and
- Serving as a grant-making and monitoring organization.

e-NC finds and advocates for solutions to ensure that all North Carolina citizens and businesses increase broadband adoption and usage and have equal access to affordable, high-speed broadband. e-NC also promotes the benefits of broadband investments around commerce, education, healthcare, agriculture and government services to demonstrate greater economic opportunities. e-NC serves as a resource and manager for various statewide broadband initiatives and accomplishes its work through public-private partnerships, targeted research and direct outreach and education. Currently, the e-NC Authority is implementing a five-year project under the NTIA of the U.S. Department of Commerce.

North Carolina's SBDD Grant

The e-NC Authority (through its fiscal agent, the Rural Economic Development Center), is the recipient of the NTIA's State Broadband Data and Development Grant for North Carolina. The SBDD grant program enables North Carolina to collect comprehensive and accurate state-level broadband availability data and to display a state-level broadband map (http://e-ncbroadband.org/), with these efforts aimed at aiding in the development and maintenance of the national broadband map. The e-NC Authority is currently using provider data for its map, but is also evaluating other data collection methodologies including Web crawling techniques and collecting broadband consumer data at the local level. In addition, e-NC uses radio wave propagation prediction modeling (using GIS) to reflect wireless coverage in North Carolina. We recently turned in a report on the evaluation of various data collection techniques to the NTIA SBDD offices . Initial broadband planning funds for the project were used to conduct the 2010 Citizen Survey on broadband usage in North Carolina and the 2010 e-Strategy Survey of businesses, organizations, and households looking at broadband usage and benefits among industry sectors. In addition to the data collection, validation and display work; and the initial broadband planning surveys,

the SBDD funding allows e-NC to undertake the following additional programs: state broadband capacity building, a technical assistance program, a Lifeline Online pilot to improve computer ownership and Internet usage (LITE-UP), and funding to partner with the NC Center for Geographic Information and Analysis on address file improvements for the state, with all these efforts continuing through October 2014.

Spring 2011 Broadband Data Collection and Mapping Process

Data Collection

The official data request letter was sent to all 104 identified providers of broadband service on June 28, 2011, via e-mail and hardcopy mailed letter. Attachments were included explaining the SBDD mapping project effort, the e-NC Authority's role in the endeavor, and all requested parameters for information. Providers were asked to reply to the request on or before July 29, 2011.

Excel and geodatabase templates were shared with providers, along with PDF format instructions summarizing all NTIA requirements and information relevant to each type of provider (fixed wireless, mobile wireless, and wireline). Technical assistance was provided to any organization who requested it.

A secure server hosted by MCNC is configured with an open source, browser-based direct file upload system called eGroupware. Providers were sent a log-in name and password for this upload system once they contacted either Samantha Jackson or Stephanie Jane Edwards to communicate that their data was ready for submission. A confirmation e-mail went to Stephanie Jane once data had been uploaded.

Reminder e-mails were sent to unresponsive providers with usernames and passwords for data upload. An official reminder e-mail was sent out the first week of August to providers of broadband service that were unresponsive to the data request. Phone calls were placed at various times in September to organizations that had not yet responded to the data request or reminders. These phone calls and some background research allowed for e-NC to determine the companies that have gone out of business and those that refused to submit data. One provider which was thought to be out of business was found to still be operational in one McDowell County (AND Wireless and Security, Inc.), and another provider (Cherokee Broadband Enterprises) was identified for the first time. The number of known broadband service providers operating in North Carolina is now at 104.

Integration of Provider Data into NTIA Statewide Geodatabase

For ease of data integration, a front-end Excel format template was offered to all providers, containing notes defining required fields, explanations of which data is required in which formats by which types of providers, and hyperlinks connecting fields to additional tables listing the corresponding NTIA-specified values and codes (for speed tiers, technology types, connection point facility types and capacities, county codes, end user types). A brief description of how census block FIPS codes work was also taken from an internet source and distributed as needed to providers who had questions about how to report this information.

BB Service by Census Block

As requested by the NTIA mapping and planning team, all census block data is included with 2010 census block geometry. Technical assistance was often needed by providers to correctly report served areas by either the 15-digit FIPS codes or in some way by which e-NC staff could derive the appropriate FIPS codes.

BB Service Road Segment

The reporting and mapping of data by street segment presented significant challenges to accurate interpretation of where broadband availability is and is not. This is mainly attributed to the difficulty of standardization among the many data structures by which providers report street segments. Quality of data has improved since some providers have switched to submitting data in shapefile format, and others have been able to start including a Tigerline ID (TLID) field for reference in mapping tabular information. Use of this unique identifier has reduced ambiguity in some tabular datasets and improved data quality upon mapping.

BB Service Address

A few address-level datasets were submitted to e-NC with the latitude/longitude coordinates already included, but most needed to be geocoded. This was done using the NC Master Address file as the primary reference file, significantly increasing the accuracy of matching records. Secondary sources for address records that did not find a match this way included street segment interpolation, ESRI data utilizing the 4-digit ZIP extension, and manual placement/digitizing based on a combination of reference data and online browser maps. Upon completion of geocoding for each provider submitting address data, the address point features were overlain with a 2010 census block layer to add the census block FIPS code attribute, then all address feature points were loaded into the geodatabase feature class. The geocoded shapefiles for each provider are kept with geocode match score and match reference type for every matched address, so the thoroughness of this data type could be tracked and/or improved with more time.

BB Service Wireless

Approximately seven small, fixed wireless providers have been able to share technical information about their transmitting towers, antennae, and frequencies, so that e-NC can produce for them a service coverage shapefile using the contracted services of the University of NC at Greensboro Center for Geographical Information Science (http://cgis.uncg.edu) . An Excel template was developed with all the relevant information that can be filled in by providers with technical assistance in some cases, and the propagation model is field-calibrated to reflect actual ground conditions.

BB Service Overview

Records for overview containing subscriber-weighted nominal speeds of a given provider were generally joined to a template layer of county features, using the option to keep matching records only. Then these matching features and their new attributes were exported as a new shapefile before being loaded into the collective overview feature class. For providers with multiple technology types serving a given county in at least one instance, this information was single-field geocoded using the 5-digit county FIPS code, and then geocoded point features were spatially joined to the county polygon using “within” criteria.

Some detail formatting performed as needed:

- Add state FIPS code and any needed leading zeros onto county code for the new State+County FIPS code. Most providers list just the county code because this was the original NOFA request.
- Change state abbreviation values from “37” to “NC”.

- Change weighted speeds to appropriate units (kbps) and remove unit text.
- Translate to county from weighted speeds reported by RSA/MSA.

BB Service - Critical Anchor Institutions

Only anchor Institutions that could be geolocated were included. Only 17 CAIs were identified that could not be geocoded to a point feature. CAIs were collected by contacting administrative offices of some CAI category types and receiving databases of information, as well as collecting from individual CAI locations for other types using survey emails and follow up phone calls as necessary. There are 5,857 CAI's identified, located, and included in the geodatabase to date.

Census Block data (tabular)

- Fields standardized and transferred into Excel template
- Geocoded to centroids of census blocks using 2010 Census Block layer in WGS1984 projection as reference file for "Address Locator".
- Spatial join of geocoded census block data points to polygon features

Census Block Geometry Conversion (2000 to 2010)

The following providers either are represented based on data submitted in previous data pulls, or submitted new data in census block geometry from 2000, therefore requiring geometry conversion:

- Carolina Mountain Cable
- Cherokee Cablevision
- CoMPAS Cable
- Country Cablevision
- Inteliport
- Lexcom
- North State
- Piedmont TMC
- Skyenet
- Star TMC
- Suddenlink
- Surry TMC
- Tele-Media Corporation
- Windstream Concord Telephone
- Windstream North Carolina

The following steps were carried out in ArcGIS for each provider:

- Select all records in Spring 2011 provider dataset
- Relate FIPS 2000 field in provider data to FIPS 2000 field in statewide crosswalk table
- Relate FIPS 2010 field in statewide crosswalk table to GEOID field in 2010 census block layer.
- Export related/selected 2010 CB records as new layer, and related/selected crosswalk records as a provider-specific dBase table.
- Attribute join on exported Crosswalk subset with Spring 2011 provider data layer based on year 2000 CB number, keeping matching records only. Number of records should stay the same. Export all features as new joined crosswalk DBF.

- Attribute join on new 2010 CB layer with the joined crosswalk DBF (which should now have the relevant provider data) based on 2010 FIPS field, keeping matching records only. Number of records should stay the same.
- Export 2ns join results into new finalized 2010 CB layer with broadband data.
- After adding the new EndUserCat field, this 2010 CB service area layer is loaded into the geodatabase.

Street Data

Some datasets were submitted to e-NC by providers already in shapefile format, and others were reported in various tabular formats (text, Excel, CSV, etc.). Of the tabular datasets, some included a Tigerline ID ("TLID") field along with some or all other fields such as city, state, zip, and census block FIPS.

- For datasets submitted tabular with TLID:
    - Max and Min address ranges were calculated from the FromRight, ToRight, FromLeft, ToLeft format used by most standard street segment reference files and incoming datasets
    - All data formatted into back-end Excel format, including converted speeds if reported at some other granularity.
    - Table geocoded to Tigerline 2010 street segment file using single-field and "TLID" values, with zero offset.
    - Geocoded point features converted to street segment geometry via spatial join using "contains" criteria, keeping matched records only.
- For datasets submitted tabular without TLID:
    - Max and Min address ranges were calculated from the FromRight, ToRight, FromLeft, ToLeft format used by most standard street segment reference files and incoming datasets
    - All data formatted into back-end Excel format, including converted speeds if reported at some other granularity.
    - Table geocoded to Tigerline 2010 street segment file using false midpoint address and either ZIP5 or census block FIPS (whichever available) as address locator zone.
    - Geocoded point features converted to street segment geometry via spatial join using "contains" criteria, keeping matched records only.
- For datasets submitted as shapefiles: VB If/Then statements used to calculate "Max" and "Min" address range attributes required by the NTIA/FCC, converted from the FromRight, ToRight, FromLeft, ToLeft format used by most standard street segment reference files and incoming datasets:
    - **To calculate "Min":**

```
Dim fromRight
Dim toRight
Dim fromLeft
Dim toLeft

fromRight = [FROMRIGHT]
toRight = [TORIGHT]
```

```
fromLeft = [FROMLEFT]
toLeft = [TOLEFT]


Dim minright
If fromRight = 0 And toRight = 0 Then
   minright = 0
ElseIf fromRight = 0 Then
  minright = toRight
ElseIf toRight = 0 Then
  minright = fromRight
Else
  If fromRight < toRight Then
    minright = fromRight
  Else
    minright = toRight
  End If
End If

Dim minleft
If fromLeft = 0 And toLeft = 0 Then
   minleft = 0
ElseIf fromLeft = 0 Then
  minleft = toLeft
ElseIf toLeft = 0 Then
  minleft = fromLeft
Else
  If fromLeft < toLeft Then
    minleft = fromLeft
  Else
    minleft = toLeft
  End If
End If
```

- **To calculate "Max":**

```
Dim fromRight
Dim toRight
Dim fromLeft
Dim toLeft


fromRight = [FROMRIGHT]
toRight = [TORIGHT]
fromLeft = [FROMLEFT]
toLeft = [TOLEFT]


Dim maxright
If fromRight > toRight Then
```

```
  maxright = fromRight
Else
  maxright = toRight
End If

Dim maxleft
If fromLeft > toLeft Then
  maxleft = fromLeft
Else
  maxleft = toLeft
End If

Dim max
If maxleft > maxright Then
  max = CStr(maxleft)
Else
  max = CStr(maxright)
End If
```

Creating last mile and middle mile features

- Formatted numeric fields in Excel as text since the short integer format in the data model for these fields will not accept values from the Excel import's default general format.
- ArcToolbox > Data Management Tools > Layers and Table Views > Create XY Event Layer
- Zoom to Layer, verifying that all points are located inside NC boundaries

Provider-specific notes, functions and corrections performed by e-NC as needed

Access/On Multimedia Inc.

- This is a middle mile only provider
- Provider confirmed no changes since last round so fall data was used

AT&T F11

- Converted subscriber weighted nom speed data from CBSA to county
- Converted max advertised speed data from CBSA to county
- Translated max advertised speeds from KBPS to NTIA codes
- Applied converted speeds to appropriate availability records by county based on FIPS codes, by pasting the CBlock FIPS codes into speed columns and using Find/Replace functions in Excel (ex Find fields with 37001* and Replace with 7). For data by street and CB.
- Copied max advertised speeds into typical speed columns (for which data was not supplied by AT&T)
- Calculated conversion of Left and Right To/From addresses for street segment data to NTIA's required Max/Min values (using "min" and "max" formulas in Excel)
- Checked data by CB for duplicates, 14,399 found and removed.
- Used 2010 TLID field in attribute join to map street segment data.

AT&T Mobility F11

- Merged shapefile features into a single multipart polygon to remove arbitrary grid boundaries.

- Validation: Ran “Eliminate Polygon Part” tool to remove any parts or donut holes less than 0.125 square miles in area.
- Added attributes supplied in Excel spreadsheet.

ATMC

- Added Address field populated with a concatenation formula of component address information.
- Added EndUserCat field and populated with code 5
- Overlay of address points w/CB layer to get FIPS code field
- Created new fields and used Calculate Geometry function in ArcMap to generate Lat and Long attributes

ATMC Wireless

- Clipped shapefile to state boundary
- Eliminated polygon parts less than 0.125 square miles
- Added spectrum attribute

CenturyLink F11

- Reprojected CB and street shapefiles and changed format of some fields for loading
- Ran Delete Identical tool on streets shapefile to remove duplicates based on Tigerline Ref numbers.
- Used If/Then scripts to calculate min and max address fields from left and right max/min ranges in ArcMap field calculator
- Created new fields of compatible type for TransTech and Provider_Type fields

Charter F11

- Projected and formatted attribute fields.
- Streets submitted and mapped in 2010 Tigerline, which has no address range information.

Comporium F11

- Removed duplicate records in Excel during formatting
- Ran address sorter, transferred previously geocoded features into new GDB (with the newer tech and speed attributes) and geocoded new data.

Comcast F11

- Mapped previous CB’s and streets, then new CB’s and Streets submitted this round
- Calculated min/max address ranges for street segment data in Excel
- Geocoded hypothetical midpoint of tabular street segments by address range and spatially joined to street segment features.
- Used Overview data from Fall 2010
- Low quality on streets from previous data pulls (only a 61% match to tiger streets w/CB zone)

Country Cablevision and Carolina Mountain Cable F11

- Converted CB shapefile from Spring 2011 into 2010 geometry
- Duplicated max advertised speeds into typical speed fields via Field Calculator
- Added Provider Type field and populated with code 1

Electronics Service Co of Hamlet

- Customized propagation model for unique antenna setup high up in trees
- Clipped output to state boundary

Electronic Solutions

- Converted coordinates, added negative sign to longitude
- Produced shapefile from data supplied in Tab D. Converted raw speeds to NTIA codes.
- Put weighted speeds into correct units.

Epproach

- Copied Census blocks from Fall 2010 geodatabase
- Merged census block polygons
- Loaded into geodatabase and populated Unlicensed for spectrum field.

Frontier F11

- Started with Spring 2010 Verizon data with legal agreement from both Verizon and Frontier.
- Applied Max Advertised speeds from MSA to CB and Street Segment level based on FIPS codes and relevant counties.
- Missing speed data: duplicated Max speeds for Typical which were not submitted. Speeds were not reported for all CB's and streets reported, and for these the lowest (except for 1 CB) values from Max speed data, NTIA code 5 for down and 3 for up, were applied.
- Middle Mile: assumed "Owned" for Ownership field to substitute for missing information, as instructed by federal program office.
- Created XY Event layer for new last mile points submitted. Learned from follow up with provider that data from previous round still applies, and that 15,000 foot service circle applies to new last mile points for availability.
- Selected 2010 road segments that intersect eitherthe new 15,000 ft radius buffers, the previously reported census blocks, or the previously reported streets. Exported as single streets shapefile for fall.
- Manually deleted stray street segments which mapped far outside of counties reported as served.
- Selected and exported census blocks <2mi^2 which contained a served street segment.
- Erased street segments from new layer which overlap 2010 census blocks <2mi^2.
- Manually attributed provider/tech/speed data (uniform across service area)
- Loaded last mile points from current and previous data pulls. Middle mile points from previous data pull.

Greenlight (City of Wilson) F11

- Re-projected shapefiles into WGS84.
- Added FRN2 field with leading zeroes, Lat, Long, EndUserCat (populated with code 5), and Provider type field (populated with code 1) to address attributes, and re-concatenated "Address" field.
- Removed duplicate addresses using Delete Identical tool in ArcToolbox, checking in Address, TransTech, MaxAdDown and MaxAdUp fields.
- Populated missing Typical speed fields with Maximum Advertised fields.
- Added/populated FRN w/leading zeroes, lat and long fields for middle and last mile

- Attribute join to county template feature class for Overview

Hughes Network Systems F11

- Added leading zeros and concatenated tabular census state, county, tract, and block FIPS component fields.
- Geocode and spatial join on year 2000 census block list. Merge on all census block features into a single polygon. Manual attribution from email and website contents.

Interstar

- Mapped subscriber addresses supplied by the provider, then used the point locations to derive a Minimum Bounding Polygon (Convex Hull) representing available wireless coverage.
- One-to-one spatial join associating provider attributes and speeds (max for served area) with minimum bounding polygon.
- Added spectrum field and populated with code 6.

Level 3

- 11 duplicate address records removed, 209 unique records remaining.

Mediacast F11

- Max Advertised speed values duplicated to populate typical speed fields.

Mediacom F11

- Corrected fields in MidMile (provider name typo, ownership, positive longitude value)
- Removed duplicate address records from data in Excel
- Concatenated street address and other full address components
- Used end user field in supplied data translating “RES” as EndUserCat code 1, and “COM” as EndUserCat code 3 (though size of the business cannot be determined).
- Manual cleanup of some address field values
- Duplicated max advertised speed values for typical speed fields
- Run script to sort out/update previously geocoded addresses and prepare new addresses for processing.

MI-Connection F11

- Deleted 8,071 duplicate records (using address, transtech, and all speed fields)
- Populated unmatched/ungeocoded addresses with placeholder values (-9999)
- Re-concatenated Address field for cleaner, consistent contents

Morris F11

- Use of same address list as Spring 2011, confirmed that these include both current and potential broadband customer locations.
- Learned that speeds for fiber records had been reported by Mbps and converted these to NTIA codes to match other records.
- Spatial join with 2010 census blocks for FIPS field.

North State

- Emailed about missing FIPS digit and inserted (leading zero for tracts) upon their response.

- Speeds were reported as Typical Up/Down only. Substituted these values into Max Ad Up/Down as well.
- Duplicate CB records were given to us for each service tier. Merged into CB shapefile after geocoding by:
    - Splitting into separate shapefiles by tech type (10, 30, and 50)
    - One-to-one spatial join field merge rule taking the maximum value from duplicates' speed fields.
- Middle Mile, Last Mile: Added negative sign to longitude values
- Last Mile point with longitude -70.97528 fell out of state boundaries and was changed to -79.97528 based on locations of all the other last mile locations.

Sprint Nextel F11

- Validation: Ran "Eliminate Polygon Part" tool to remove any parts or donut holes less than 0.125 square miles in area.

Starvision F11

- Parsed tabular address-level data in Excel.
- XY event layer created with address-level data that included coordinates.
- Corrected typo on one point that appeared outside of service area, moved point and applied new census block FIPS code.

Surry TMC and Piedmont TMC

- Contacted for clarification and formatted mislabeled "street" information into address tab
- Removed 7 duplicates from address data in Excel

Skybest and Skyline F11

- Downloaded and created missing .prj file for shapefile exports from provider, based on follow up determining an NAD 83 North Carolina FIPS 3200 ft projection.
- Converted polylines to polygon for each DSL and fiber-to-the-home technology layers.
- Reprojected into WGS 1984
- Selected Tigerline 2010 streets by location inside newly created polygons
- Attributed for TransTech then merged into street layers by provider
- Created fields and attributed manually from contents of provider-supplied Word documents.
- Used VB script in Field Calculator to derive max/min address range information

Sky Catcher

- Wireless Propagation study.
- Created XY Event Layer to map Middle Mile information, deleted duplicate records. Remaining records loaded into geodatabase.

Suddenlink F11

- Deleted 203 census block records and 332 address records that had blank technology/speed info, after confirming with the provider these records are not relevant to broadband.
- Removed 410 duplicates in the address field after concatenating without apartment numbers.
- Mapped census blocks then converted to 2010 census block geometry/attributes.
- Manually deleted 2 census blocks that mapped well outside the reported service area.

TDS Telecom

- 713 out of 741 matches for addresses submitted found in old data. Remaining records geocoded.

Tele-media

- Provider type of 1 assumed and populated.
- Checked for duplicates CB's in Excel, none found

Time Warner Cable

- CB and Streets:
  - Padded FRN w/two zeroes
  - Reprojected into WGS 1984
  - Added Provider Type field and coded as a "1"
  - Input Max Advertised speeds as Typical Speeds as well, since they were not provided.
- Streets: created "AddyMax" and "AddyMin" fields and used If/Then statement to calculate values from LFrom, LTo, RFrom, and RTo fields

T-Mobile F11

- Reprojected shapefiles into WGS 1984.
- Added field to categorize by technology type/T-mobile service tier (3G, 4G).
- Attributed manually from information sent in a text file from T-Mobile.
- Executed spatial Union between coverage of higher speed and the broader 3G coverage, then extracted (Data Export selected features) resulting 3G only features to distinguish max speeds here versus where higher speeds are also available. Merged into single shapefile
- Eliminate Polygon part tool to remove features <0.125 square mile.

Tri-County

- Concatenated address information into single Address field in BackEnd template spreadsheet.
- Duplicates removed by technology type (17 dsl, 3 wireless)
- Lat/longs from provider with address data, so mapped using Create XY Event Layer in ArcToolbox
- Sorted, selected, and exported by TransTech types 70 and 10, then one-to-one overlay of each shapefile with CB layer. Maximum merge rule used for speed information.
- For Tech Type 10: Selected and exported resulting aggregated CB data for CB's <2 mi. These were loaded into the geodatabase with associated broadband data.
- For Tech Type 70: created copy of resulting CB's <2 shapefile and merged all features into one multi-part polygon. This was loaded into the wireless feature class and manually assigned "Unlicensed" spectrum value.
- Address feature layer was clipped using polygons created from merged CB's OVER 2 miles, and those in the clip result were loaded into the geodatabase with associated broadband data.

Verizon Wireless

- Compared submitted shapefile with previously submitted shapefile, differences confirmed.

Windstream (Windstream North Carolina, Windstream Concord Telephone, and Lexcom)

- Sorted 2 Access tables by "DSL" field and deleted all records without a "Y"
- Sorted 2 Access tables by census block size field, dividing up data by CB and streets

- Copy pasted all relevant fields into Excel Template column by column, including number listed indicating company name and MSA/RSA name pasted into Max Advertised Download Speed field.
- Used Find/Replace to populate appropriate Provider, DBA Names and FRN's (sent in emails upon request) and Up/Down Max Advertised Speed info based on contents of cells w/direct relationship to this information.
- Recalculated left/right, to/from street segment address ranges to max and min
- Created false address using the integer midpoint of max and min concatenated with street name provided, then geocoded these "addresses" using Tigerline 2009 overlain with CB 2000 to use as Zone
- Split Windstream NC and Windstream CT geocode results up into two tables, then one-to-one keep common spatial join w/Tigerline 2009 features using "contains" criteria.

Yadtel F11

- Corrected middle mile capacity code after checking with the provider.

Post-processing Functions for Final Integration

Census Block

After Census Block data was loaded into the transfer geodatabase feature class, FIPS code fields were calculated using commands in the Field Calculator and contents of the FullFIPSID field. The following calculation formulas were used:

STATE FIPS = Left ([FULLFIPSID],2)
COUNTYFIPS = Mid([FULLFIPSID],3,3)
TRACT = Mid([FULLFIPSID],6,6)
BLOCKID = Right ([FULLFIPSID],4)

- 1019 duplicate records (with same value for Provider Name, DBA Name, FRN, TransTech, and FullFIPS ID) were removed using the ArcToolbox Delete Duplicates tool. Identified by provider using the ArcToolbox frequency tool, the following number of duplicates were removed:
    - CenturyLink (1010)
    - Covad Communications (1)
    - Ellerbe TMC (3)
    - Time Warner Cable (1)
    - PTC Communications (1)
    - TriCounty Telecom (3)

Address Data

- Exported features into a shapefile, conducted one-to-one, keep all spatial join with CB 2010 using "Is_Within" criteria to produce the associated 15-digit FIPS Code. These features were then reloaded into a clean version of the Address feature class.
- Parsing of 317 records that had null value for building number or street name or city.
- Reverse selection within state boundary used to then export (for record-keeping) and deletion of addresses outside North Carolina.

- Excluded 23 records (from @Communications) that did not meet broadband speed threshold values.
- Sorted out, selected, and field calculated missing End User Category values from "ZZ" default value to "5" for Other/Cannot Determine.
- Calculated geometry for missing Lat/long, for unmatched addresses changed to -9999.

Wireless

- Duplication of multipart coverage polygons to reflect multiple spectrum ranges used, per NTIA/FCC instruction.

Overview

- Field Calculated "Geographic Unit Type" field to CO, and "StateAbbr" field to NC.
- Field Calculated missing Maximum Advertised Up and Down speed fields to "ZZ" "default" values.
- Deleted records of information for wireless technology types.
- Verified that all FRN's were either 9999 or 10 digits with leading zeroes.

Last Mile

- Field Calculated "Ownership" field to -9999, as we do not collect this field. Calculated "StateAbbr" field to NC. Then went back and calculated all "Ownership" field values to "0" for owned since the data model script does not accept the default values we were instructed to use.

Middle Mile

- Spatial join with census block layer to derive the 15-digit FIPS code, then reload features into middle mile feature class including the new values for populating the "FullFIPSID" field.
- Replaced Null Elevation values with -9999 "default" value using Field Calculator.
- Populated State Abbreviation column with "NC".

CAI

- Parsed address information for address fields
- Deleted "DMV Tag Office" in "Charlotte, NC" due to absence of street address information. Was geocoded incorrectly.
- Deleted 526 records for which survey respondents report that they do subscribe to broadband but did not give speed information accepted by the NTIA's script.

Verification Implemented Prior to Fall Data Submission

Data verification methods implemented by e-NC in time for submission at the federal level followed generally along the lines of quality control. Methods most often used are outlined below. Time constraints on existing staff did not allow for the execution of some less basic verification approaches that are in the planning/setup stages, but more substantial verification involving multiple data sources are in development.

Standardizing

The files from datasets received from each provider, except for those few submitted in shapefile format, were manually transferred to a back end Excel-format template with field headers, to create a single-file, standardized field structure for each provider's data that could be used for quick reference and map feature creation. This step also helped staff to ensure that all required components were either present or requested in follow up to the provider, and that the components were reported in the correct format.

Lat/long coordinates

Some information was submitted to e-NC with lat/long coordinates included for the location of point features. This location information was checked during the mapping process, and values were corrected if the provider had made mistakes such as reversing the latitude with the longitude, or forgetting to include the negative sign for the longitude value. In addition, e-NC followed up with providers on point features that showed up in the map outside the state and/or outside the provider's reasonably expected service area. Point features that mapped outside the state after follow up with providers, including those that mapped to zero degrees latitude and longitude due to an unknown location, were deleted from the geodatabase for submission at the federal level. For fixed wireless data generated by propagation model from antenna specs, the latitude/longitude coordinates of the antenna locations reported by the provider to e-NC were verified by e-NC's university GIS research contractor using high-resolution orthoimagery.

Multiple FRNs

In several instances, providers reported multiple FRN's that increased in numerical increments of one for each record of data, and this was found to be a simple error when the providers were trying to paste their organization information down the rows applying to a list of broadband data records. This was checked for and corrected after confirming that the lowest/first reported FRN was the correct one.

Correct technology type codes

Knowledge from our technical staff and online research was sometimes used to supplement data that e-NC had relevant to a provider that was unresponsive or otherwise did not supply this specific piece of the information. For example, a provider may have gaps in their transmission technology field and these were filled in when technical staff could confirm that the provider operates with only a single technology type. Or the staff may know which technology type is used by a provider who simply left this field blank on all records.

Subscriber-weighted nominal speeds

Weighted nominal speed values were checked, and staff followed up with the provider if all values were the same for multiple counties, as this could result from either a single speed tier for a given transmission technology across counties, or in some cases providers were not following the formula provided and had manually entered the same value regardless of differences in subscriber numbers. When these cases were discovered, technical assistance was offered and a new subscriber-weighted nominal speed dataset created to reflect variation between counties.

Wireless model fieldwork

For fixed wireless provider data that was generated as coverage area output from models based on technology and environmental factors, the data was verified by "ground-truthing" with measurements of signal strengths at sample locations within a provider's service area, observation of the influential ground conditions in each location, and comparison to the expected signal strengths at the same

locations in the model. Some calibration of the model was then performed so that the resulting polygons could more accurately reflect what would be found in real life.

Check Geometry

After compiling all datasets into the geodatabase feature classes, the check geometry process in Arc Toolbox's Data Management section was used on each feature class to identify and repair any geometry errors in the features.

Comparisons with Citizen-Sourced Data

The e-NC Authority has recently constructed a mapped layer of input from citizens who report having no access to broadband at their location, from any broadband provider. The compiled layer is collected from local citizen advocates, citizen input on e-NC's website feedback form, and locally conducted surveys. Comparison of provider-sourced data with this source of information has allowed for targeted follow up with providers in order to promote access to broadband for these citizens, as well as to begin refinement of our statewide broadband data. Further data collection from citizen input and comparative analysis approaches will be described in spring 2012.