North Carolina Data Submission
April 2011

Data Collection Methodology

The e-NC Authority

4/15/2011

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. 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 112 identified providers of broadband service on January 26, 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 parameters requested information. Providers were asked to reply to the request on or before February 28, 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 (mobile wireless, fixed 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 has 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 in mid-March to providers of broadband service that were unresponsive to the data request. Phone calls were placed during the weeks of March 14th and 21st 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. The number of known broadband service providers operating in North Carolina is now at 102.

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 2000 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.

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 2000 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. 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 4,224 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 2000 Census Block layer in WGS1984 projection as reference file for "Address Locator".
- Spatial join of geocoded census block data points to polygon features

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 2009 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 2009 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: 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 dFROMRIGHT As Double
Dim dTORIGHT As Double
Dim dFROMLEFT As Double
Dim dTOLEFT As Double

dFROMRIGHT = val([RFROMADD])
dTORIGHT = val([RTOADD])
dFROMLEFT = val([LFROMADD])
dTOLEFT = val([LTOADD])


Dim minright As Double
If dFROMRIGHT = 0 And dTORIGHT = 0 Then
   minright = 0
ElseIf dFROMRIGHT = 0 Then
  minright = dTORIGHT
ElseIf dTORIGHT = 0 Then
  minright = dFROMRIGHT
Else
  If dFROMRIGHT < dTORIGHT Then
    minright = dFROMRIGHT
  Else
    minright = dTORIGHT
  End If
End If

Dim minleft As Double
If dFROMLEFT = 0 And dTOLEFT = 0 Then
   minleft = 0
ElseIf dFROMLEFT = 0 Then
  minleft = dTOLEFT
ElseIf dTOLEFT = 0 Then
  minleft = dFROMLEFT
Else
  If dFROMLEFT < dTOLEFT Then
    minleft = dFROMLEFT
  Else
    minleft = dTOLEFT
  End If
End If

Dim min As String
If minleft = 0 And minright = 0 Then
   min = "0"
ElseIf minleft = 0 Then
  min = Str(minright)
ElseIf minright = 0 Then
   min = Str(minleft)
```

```
Else
   If minleft < minright Then
     min = Str(minleft)
   Else
     min = Str(minright)
   End If
End If
```

- **To calculate "Max":**

```
Dim dFROMRIGHT As Double
Dim dTORIGHT As Double
Dim dFROMLEFT As Double
Dim dTOLEFT As Double

dFROMRIGHT = val([RFROMADD])
dTORIGHT = val([RTOADD])
dFROMLEFT = val([LFROMADD])
dTOLEFT = val([LTOADD])


Dim maxright as string
If dFROMRIGHT > dTORIGHT then
  maxright = dFROMRIGHT
Else
  maxright = dTORIGHT
End if

Dim maxleft as string
If dFROMLEFT > dTOLEFT then
  maxleft = dFROMLEFT
Else
  maxleft = dTOLEFT
End if

Dim max as string
If maxleft > maxright then
  max = maxleft
Else
  max = 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

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

- 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)
- Parsed Street names from street types using Find/Replace functions in Excel
- 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, none found.

ATMC

- Missing End User Category, Typical Speeds data on address data
- 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
- Substituted/duplicated max advertised speed values in typical speed fields for wireline and wireless address-level data.

CenturyLink

- Reprojected CB and street shapefiles and changed format of some fields for loading
- Excluded 593 CB's with speed codes outside the NTIA domain for ADSL (1, 2, and 9 for max down)
- 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

Comcast

- Deleted records from fall as indicated in CB and Streets information
- Mapped new CB's and Streets
- Used Overview data from Fall 2010
- Low quality on streets (only a 61% match to tiger streets w/CB zone)

Country Cablevision and Carolina Mountain Cable

- Copied CB shapefile from Fall2010
- Duplicated max advertised speeds into typical speed fields via Field Calculator
- Added Provider Type field and populated with code 1

Covad/DIECA

- Geocoded streets to Tigerline 2009 using false address created from integer midpoint of max and min, and 5 digit ZIP code as zone. 77% match (1640 out of 2138). Zero offset from centerline for matched "addresses".
- Spatial join with Tigerline 2009 (one to one, keep common, contains). 1504 found matches.

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

- Used 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.
- Verification: checked tabular CB data for duplicates, none found.
- Streets: were provided in text files without city, zip, or Tigerline ID, but did have CB FIPS.
    - Overlay to associate statewide Tigerline streets to corresponding CB FIPS codes
    - Creation of address locator for geocoding tabular street data as street segment midpoint features.
    - Modification of Address locator properties to remove default 20 ft. offset from street centerline reference features.
    - Tabular street data geocoded to points using a false address created by the midpoint integer of the min max address range and census block FIPS
    - Select by Location on Tigerline layer for features that intersect geocoded points, exported selection as new shapefile
    - Spatial join w/new selected streets shapefile and geocoded broadband data points, one-to-one keeping matches only where segment lines CONTAIN geocoded midpoints w/broadband data

Greenlight (City of Wilson)

- Re-projected shapefiles into WGS84.
- Added FRN2 field with leading zeroes, Lat, Long, and Provider type field (populated with code 1) to address attributes, and re-concatenated "Address" field.
- Added/populated FRN w/leading zeroes, lat and long fields for middle and last mile
- Attribute join to county template feature class for Overview

Level 3

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

Mediacom

- Corrected fields in MidMile (provider name typo, ownership, positive longitude value)
- Used address data from Fall 2010, which contained many duplicates
    - Exported fall addresses into new shapefile
    - Exported .dbf of shapefile into new Excel file
    - Removed 11,992 duplicates in Excel
        - Data > Remove Duplicates (went from 108,043 records to 96,051)
    - Created an Address locator using Fall Mediacom's non-duplicate address points only
    - Re-geocoded new Excel table containing unique records.
    - Added short integer TransTech and Provider_Type fields and loaded into spring geodatabase
    - Excluded 286 address records that were PO Boxes only.

Morris

- Considerable cleanup and re-parsing to the provider sourced address-level data
- Found 3 address records with speed values of 14, 15, and 50 in all 4 speed fields, changed these to "ZZ". Then went back and deleted these per NTIA instructions. Will follow up with provider to learn what these values should be.

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.
- Excluded 2 CB's that only had 14 digits in FIPS code field and could not be geocoded to the correct area.

Randolph Telephone and Randolph TMC

- Data was reported all by street segment in a text file, with no CB number, TLID, or Zip code field.
- Created false address from integer midpoint of address ranges concatenated with street name
- Built Address Locator using NC Streetmap reference file with city name as Zone, removed default offset of output.
- Geocoded street data false addresses, then spatially joined to line features
- One-to-One spatial join of geocoded street data POINTS to 2000 CB, using merge rule of Maximum for max advertised and typical speed fields. All tech types were 10. Sorted resulting polygons by area and exported just the CB polygons < 2 mi. These were loaded into the geodatabase with associated broadband attributes

- Clipped geocoded and joined street data POLYLINES by polygons created from merged CB's OVER 2 square miles. These were loaded into the geodatabase with associated broadband attributes.

Sprint Nextel

- Fixed topological errors
- Merged polygons with like spectrum, transtech, and speeds.

Star TMC

- Excluded 106 census blocks (out of 1641) whose numbers did not find an exact match.

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

- Duplicated Maximum advertised speeds into Typical speed fields which were not provided.
- Ran skyline through address sorter
- Geocoded address data

Sky Catcher

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

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

- Reprojected shapefiles into WGS 1984.
- Repair Geometry to fix some incomplete polygons
- Executed spatial Union between coverage of HSPA Plus (higher speed) and the broader 3G coverage, then extracted (Data Export selected features) resulting 3G only features to distinquish max speeds here versus where HSPA Plus is also available.
- Simplify Polygon tool to delete excess vertices and simplified features <0.125 square mile.
- Loaded final features into geodatabase feature class and then attributed manually from information sent in a text file from T-Mobile.

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.

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)

- 1033 duplicate records (with same value for Provider Name, DBA Name, FRN, TransTech, and FullFIPS ID) were removed using a python script created for this purpose.

Address Data

- Exported all features into a shapefile, conducted one-to-one, keep all spatial join with CB 2000 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.
- Reverse selection within state boundary used to then export (for record-keeping) and deletion of addresses outside North Carolina.
- Sorted out, selected, and field calculated missing End User Category values to "ZZ" default value
- 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.
- To remove "donut holes" in coverage shapefiles less than 0.125 square mile in area as instructed by the NTIA/FCC:
  - Separated feature classes into unique attribute records. For each provider's feature class…
    - Created a polygon feature class with one large polygon covering all of NC, called BACKGROUND
    - Performed a Union between BACKGROUND and the wireless feature class so that gap areas were then polygons. (1)
    - Dissolved with multi-part feature unchecked to explode the multipart features, but preserved the attribution. (2)
    - Selected areas that were both part of the original polygon and are less than 0.125 sq mi (3484800 sq ft) in area. Dissolved the selection to remove donut holes; did not create multipart features. (3)
    - Selected areas that are greater than 0.125 sq mi (3484800 sq ft) in area. Exported selection to remove small islands. (4)
    - Dissolved again, this time creating multipart features.
  - Joined each feature class to the original datasets to re-acquire the attribute information.
  - Used append to merge the provider files back into one feature class per provider.

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 Spring 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.

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.