Maine SBDD Data Submittal to NTIA
Technical Whitepaper

Version 1.0

This document is written and maintained by the James W. Sewall Company (Sewall). Version numbers are used to identify the contents of the document at different points in time and a revision history for the document is provided in the table below:

Version	Release Date	NTIA Submission	Description
1.0	10/1/10	10/1/10	Initial delivery

Table of Contents

List of Figures

1 Introduction

As an NTIA State Broadband Data and Development (SBDD) grant recipient, the State of Maine is undertaking a statewide project to inventory and map broadband services for inclusion in both national and state broadband maps. The SBDD grantee project team for Maine consists of the ConnectME Authority (ConnectME), the Maine Office of GIS (MEGIS), and the James W. Sewall Company (Sewall). The team is collecting broadband service availability data, including speeds and types of technology, as well as information on Community Anchor Institution (CAI) locations across the entire state. The collected service data undergoes geospatial processing and verification steps before it is loaded into Maine's broadband geodatabase. This geodatabase is used to satisfy NTIA's bi-annual submission requirements as well as support the ConnectME Authority's statewide initiatives and programs.

This whitepaper provides an account of the methodologies and processes used throughout the project and is organized by the projects major components.

2 Provider Outreach and Data Gathering

Mapping broadband coverages across the State begins by identifying potential providers and contacting them to determine service capabilities and level of participation. If a provider offers broadband level Internet service in Maine, the provider will be invited to participate in the project. After executing a non-disclosure agreement (NDA), the provider submits data showing where services are offered, technology of transmission used, and maximum advertised downstream and upstream speeds. The project team has developed a step by step process that has been captured by the high-level workflow shown in ***Figure 1***. Starting with contacting a service provider, the workflow allows a user to determine whether a provider should be included and if so what types of service are offered.

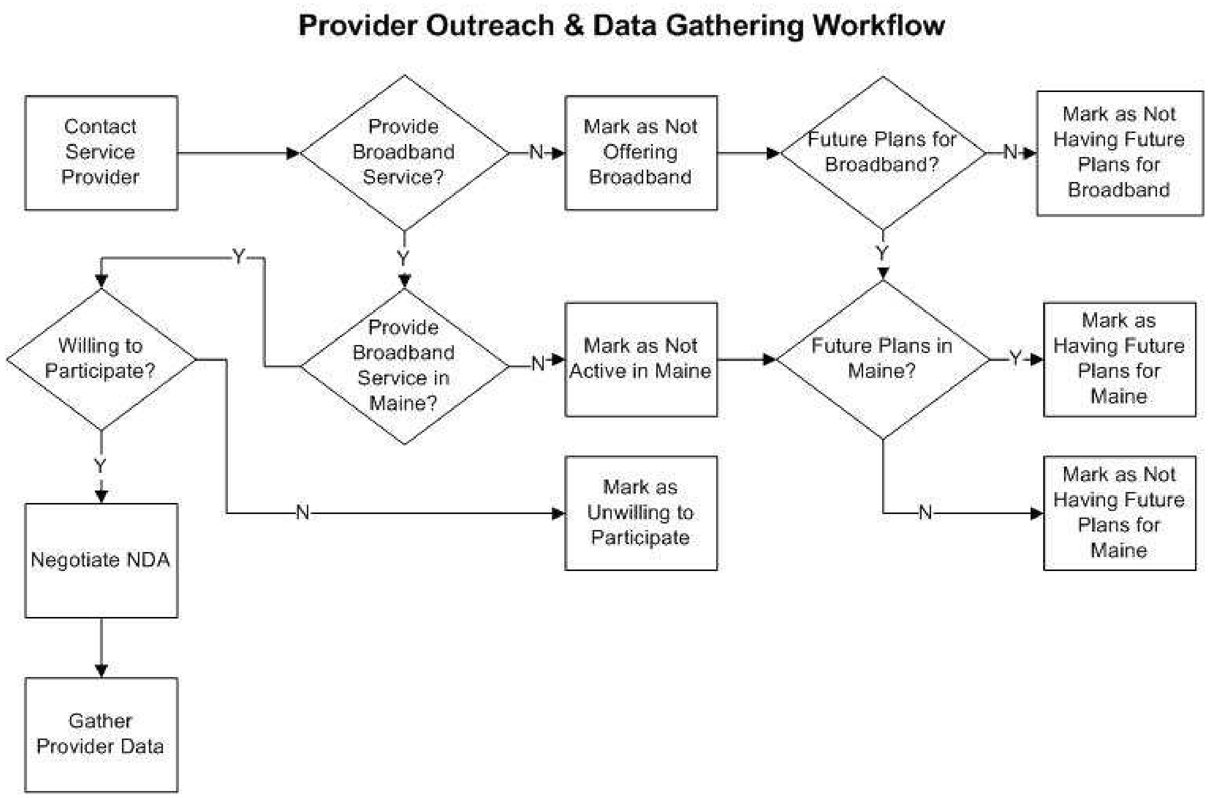


***Figure 1** - Provider Outreach and Data Gathering Workflow*

The task of reaching out to the provider community and gathering service data has five main tasks: Research Service Providers, Execute NDA, Gather Provider Data, Assess Provider Data, and Categorize Data for Production.

2.1 Research Service Providers

The Maine project team has established a service provider contact database, which contains contact information for all of the potential broadband service providers in the state. The initial set of providers was obtained from state and industry lists as well as Internet research. Ongoing management of the list is required because new providers begin offering services that qualify as broadband and changes occur to existing provider companies through mergers or acquisitions.

Sewall initially contacts each provider by phone and introduces the project. One purpose for the initial contact is to identify the individual at the provider company with whom the team should be working. In some instances, especially for larger companies it may take multiple attempts before the appropriate person is reached.

Another purpose is to determine if the company's services meet the requirements for inclusion in the project. If a company offers broadband level service in Maine then the next step is to determine the type(s) of service being offered, whether the service offerings are as an end-user provider or as a middle mile/back haul provider, and whether the company owns facilities or re-sells services using another carrier's network. Data from back haul carriers and resellers are included in the project.

A third purpose behind the initial contact is to confirm that the provider wants to participate in project and is willing to submit data that represents its service offerings and coverages. Provider companies who elect to participate are invited to execute an NDA to protect those data items considered to be confidential or proprietary. If a provider company does not want to participate, Sewall may look for assistance from the ConnectME Authority and the NTIA SBDD project team to encourage participation.

2.2 Execute Non-Disclosure Agreement (NDA)

The process of executing an NDA starts with sending a letter of introduction along with an NDA template and a copy of a ConnectME Protective Order. **Appendix A** contains a sample letter. The NDA template was drafted by the Maine law firm, Rudman & Winchell, based on confidentiality guidelines presented by NTIA and can be found in **Appendix B**. A copy of the ConnectME Protective Order signed on 21 December 2009 at the request of many of the service providers is in **Appendix C**.

Changes to the NDA template are negotiated with individual companies as needed. Once finalized, the NDA is signed by the provider company, Sewall, and the ConnectME Authority before the data gathering process begins.

2.3 Gather Provider Data

More often than not after an NDA has been executed, a different individual at a provider company is identified as the primary contact for data submittals. Once the contact is confirmed, a data submittal information sheet prepared by the project team is sent to the contact. The data submittal sheet identifies the data items desired and has definitions from the SBDD NOFA. The items requested include:

- FRN or provider FCC Registration Number
- Location and extents of service coverage
- Technology of service

- Speeds of service including maximum advertised downstream & upstream speeds
- Tower and transmitter locations and transmission attributes (for fixed wireless service)
- Middle mile and back haul connection points
- Customer service locations (for wired and fixed wireless service)
- Failed service locations (for wired and fixed wireless service)
- Community anchor institutions

After sending the data submittal information Sewall follows up with the provider contact to review the requested data items and discuss potential formats for submitting data. The team is cognizant of the wide range of environments operated by the provider companies and recognizes the need to accommodate submissions in many different formats including tabular (CSV, Excel, DBF), GIS (ESRI shapefile, ESRI geodatabase, MapInfo, Google KML/KMZ, CAD (AutoCAD, Microstation), and hardcopy. The team also understands that many of the smaller providers in Maine are handicapped by a lack of resources in trying to comply with the project's data submission requirements. Some of the issues facing these providers include small staff sizes, lack of mapping technical expertise, and proprietary digital systems. Sewall lends technical assistance and expertise as needed.

A file transfer site is currently used for providers to submit data. The site utilizes HyperText Transmission Protocol, Secure (https), with password protected user accounts and separate folders for each provider. The project team is also continuing its implementation of a web-based application for managing data submittals, which can track file uploads and has built-in change detection for tabular data submitted in *.csv* file format.

2.4 Assess Provider Data

After data has been submitted by a provider, Sewall catalogues it and assesses the data files to see if all of the requested items were provided and what data types were received. Sewall also verifies the locations and spatial definitions for the data items and checks for missing attribute information. Any questions generated are sent to the provider for clarification. It is common for the initial submission to need multiple iterations of data exchanges and feedback before the submission is completed.

Once an initial set of broadband service data is in place, follow-up rounds of data gathering will incorporate modifications to existing service coverages, service types, or service speeds. Later submittals by a provider could consist of an entire set of data records or may only contain updates since the previous submission. Sewall's integration processes are equipped with GIS and database tools to fold newer versions of provider records into the existing baseline. The team anticipates that further development and refinement of these processes and tools will be made as more update submissions are received.

2.5 Categorize Data for Production

When data from a provider has been received and assessed, production processes are needed to integrate the data into the project database. **Section 4** of this paper describes the various workflows to turn the submitted data into the SBDD data transfer model features and attributes.

3 Community Anchor Outreach and Data Gathering

Community Anchor Institutions (CAI), as defined by NTIA NOFA category codes, consist of the following:

Category 1: School – K through 12
Category 2: Library
Category 3: Medical/Healthcare
Category 4: Public Safety
Category 5: University, College, Other post secondary
Category 6: Other community support – government
Category 7: Other community support – non-governmental

The three primary steps with the CAI are data gathering, data processing and attribution.

3.1 Data Gathering

Several data sources were utilized to represent all CAI categories across the state.

State of Maine, Office of Geographic Information Systems (MEGIS)

- ARMORIES
- CEMA (County Emergency Management Agency)
- COLLEGES
- FIRE
- HOSPITAL
- HAS (Hospital Service Areas)
- MEAIR (Airports)
- POLICE
- REDCROSS
- RESCUE
- SCHLIB (School Libraries)

NAVTEQ-NAVSTREETS (Points of Interest)

- NAVTEQ-COMMSVC
- NAVTEQ-EDUINSTS
- NAVTEQ-HOSPITAL
- NAVTEQ-TRANSHUBS

State of Maine – State Facilities

- State Facilities File

Maine Department of Health & Human Services (DHHS) – Maine Care Services

- Hospitals
- Clinics/Rehab/Nursing
- Schools
- Pharmacies
- Home Care
- Counseling/Psychologists

- Shared Living
- Mental Health
- School Departments
- Health related businesses
-

Service Provider Data

- CAI data submitted by provider companies

3.2 Data Processing

The data processing task involved an in-depth cleaning and sorting of all CAI source records. Data is initially sorted as spatial (e.g., GIS layer) and non-spatial (e.g., table) data. The spatial data consisted of points and generally needed minimal formatting before loading into a personal geodatabase. The non-spatial data required some initial format revisions to prepare the data for geocoding to generate spatial geometry. The following descriptions associated with ***Figure 2*** below outline the overall workflow and processes involved.

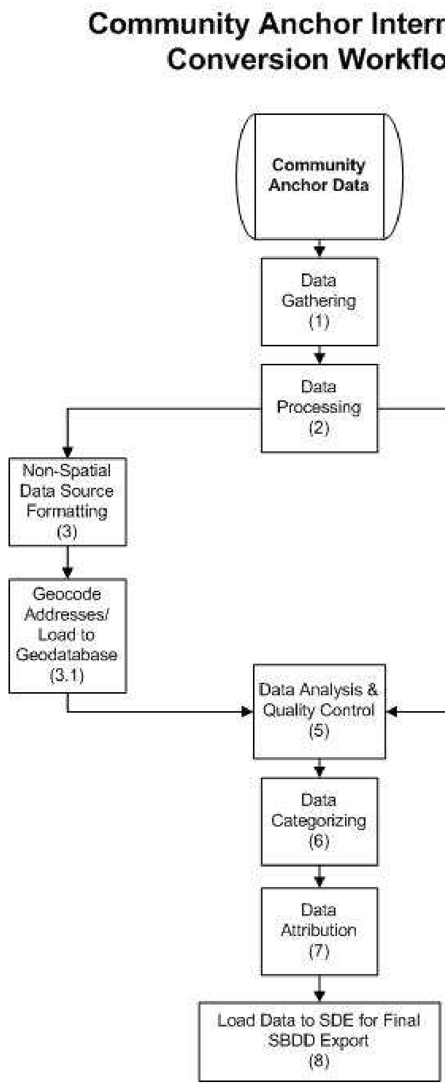


Figure 2 - Community Anchor Internal Workflow

(1) Data Gathering
Data gathering involves acquiring source data involving the seven categories defined by NTIA NOFA. Data may originate from several sources including state, county, town, outreach programs, service providers and more. Records are documented for metadata and given a level of confidence reflecting the data source, spatial accuracy and processing enhancements.

(2) Data Processing
The data processing phase separates the data sources into two types: flat file (non-spatial) and spatial. A flat file refers to data or a table that contains 1 record per line, generally in the format of an *.xls* spreadsheet or *.dbf* table. Without spatial coordinate values to translate to points, this type of data must be geocoded in ArcGIS. Spatial data contains pre-defined coordinate values or is already in a format containing spatial geometry with a defined projection and can be imported directly.

(3) Non-Spatial Data Source Formatting
Non-spatial data files are scrubbed to ensure that all necessary fields are present and are formatted to run through the geocoding process.

(3.1) Geocode Addresses/Load to Geodatabase
Using the geocoding tool in ArcGIS, an address locator file must first be setup. The address locator file maps out the ConnectME street centerline fields and is used as a reference for the non-spatial data during the geocoding process. The non-spatial data is saved as a *.csv* file. Shown below is a typical record formatted to geocode.

Name	Address1	City	State	Zip
Healthworks	10 Bangor	Bangor	ME	04401

In this example, the geocoding process will reference or match this address record to the ConnectME street address locator and place a point at this location in the map layer. All records in the source file are processed at once. Points are generated, based on how matching parameters or set. Points are then loaded into personal geodatabase for final scrubbing and quality acceptance.

(4) Spatial Data Source Formatting
Spatial data sources are received as flat files with spatial coordinate values or reside in a GIS layer as points. Each source type is processed differently.

Flat files with coordinate values:

- Prepare field name formats
- Prepare coordinate values in decimal degrees

Name	Address1	City	State	Latitude	Longitude
Healthworks	10 Bangor St	Bangor	ME	46.1252	-67.8422

- Add X,Y data into ArcGIS, generating the point locations on the fly
- Output to personal geodatabase for final scrubbing and quality acceptance

Point files:
- Export file to shapefile format if necessary
- Project file to state coordinate system (UTM NAD83 Zone19 Meters) for compatibility with other data layers
- Output to personal geodatabase for final scrubbing and quality acceptance

(4.1) Load Spatial Data to Geodatabase
All spatial data types (point files) are loaded into a personal geodatabase for final scrubbing and quality acceptance.

(5) Data Analysis and Quality Control
A final analysis is completed on all points loaded in the personal geodatabase to identify any issues. The table below indicates the primary types of issues, the means to detect them, and the resulting solution.

Issue ⇒	***Identification*** ⇒	***Result***
Duplicate Points	Selection by location/imagery review	Delete incorrect record
Unmatched geocoded records	Google Maps review	Matched record
Inaccurate CAI locations	Imagery review	Modify point location
Unsuitable CAI	-	Delete record

(6) Data Categorizing
Once the CAI records have gone through the data analysis and quality control, the records are given a category value of 1 to 7, as discussed in the introduction.

(7) Data Attribution
CAI attributes are the most difficult to acquire at the data gathering stage and are typically acquired through additional steps, including contacting each CAI. The required attributes are:
- Broadband Service
- Technology of Transmission
- Advertised Downstream and Upstream Speeds
-

The project team is in the process of acquiring the results of two surveys – one conducted by the Maine Fiber Company for the 'Three Ring Binder' project and the other conducted by the Maine's DHHS. Both surveys included broadband service questions for CAI's. In addition, the project team will initiate an outreach program where CAI's with missing broadband service information will be contacted by phone.

(7) Load Data to SDE for Final SBDD Export
CAI data is loaded from the personal geodatabase to the SDE environment for final export to SBDD format.

4 Data Analysis and Conversion

Data is analyzed and converted with different processes, depending on its type and characteristics.

4.1 Fixed Wired Transmission

Fixed wired service provider companies in the state of Maine range from small to large businesses and utilize several distinct types of technology to deploy broadband service. In order to accommodate the varied inputs, Sewall has developed a flexible and comprehensive workflow to incorporate provider information into a state broadband map developed by Sewall in conjunction with the ConnectME Authority.

The ConnectME model depicts broadband service provider coverage at the street segment level. The model uses a street centerline as the spatial component of the coverage, and a related table stores provider specific information for street segments. Sewall developed production tools to accommodate the incorporation of service provider data into this ConnectME model and instill quality control into the process.

The steps in the process for analyzing and converting Fixed Wired Transmission data are outlined in ***Figure 3*** and described below.

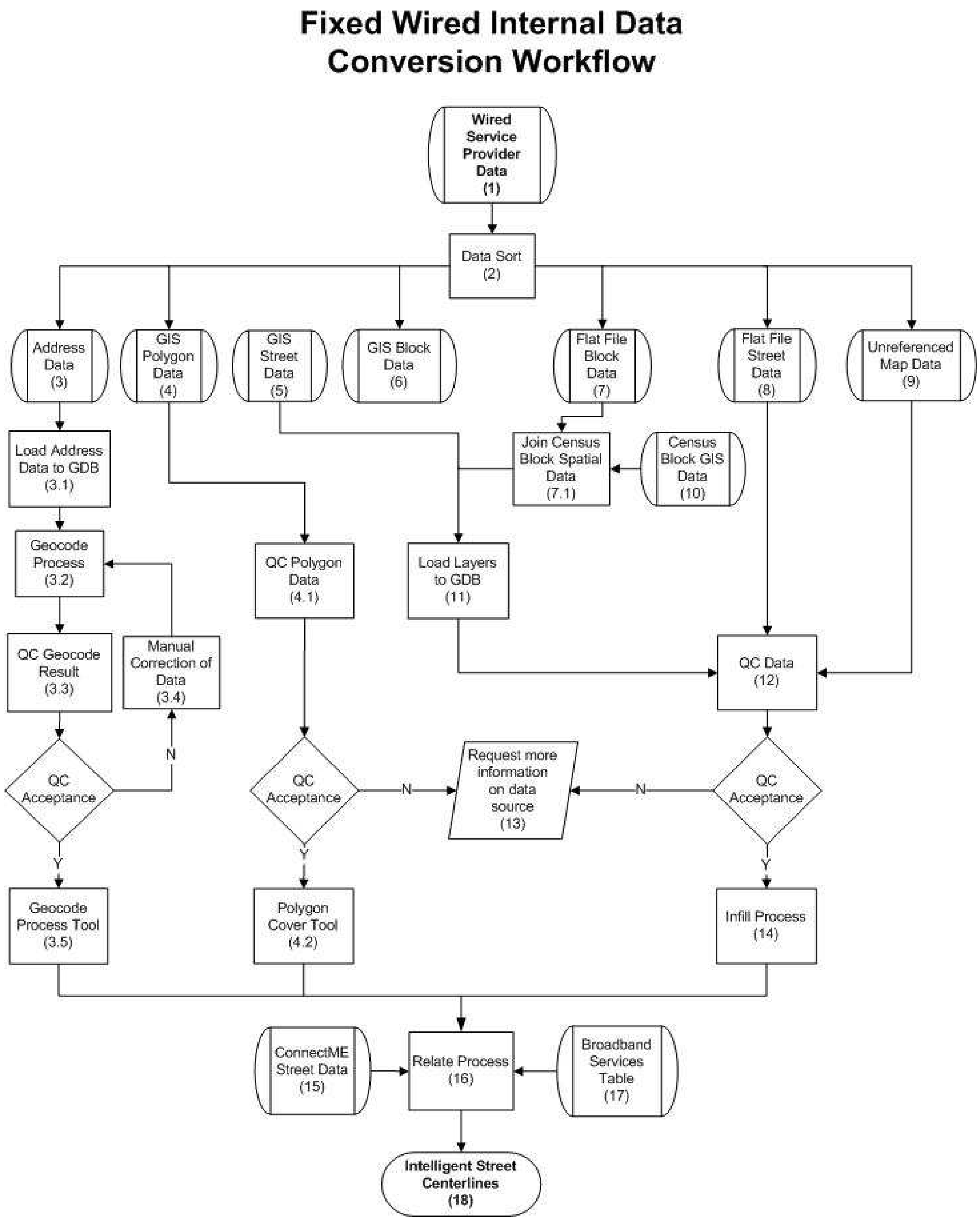


Figure 3 - Data Flow for Fixed Wired Transmission Providers

(1) Wired Service Provider Data
The data bin is the storage location for wired broadband service provider data gathered by Sewall.

(2) Data Sort
The data sort phase immediately follows the data collection process. Analysts sort the wired data by provider and by data characteristics. The wired data can consist of address data, predefined coverage data, flat file coverage data and unreferenced maps. Individual workflows have been developed by Sewall for the various data formats.

(3) Address Data
The address data bin is reserved for service provider data that is at the address level. Examples of address data formats received are spreadsheet and text file format.

(3.1) Load Address Data to Geodatabase
Address data is formatted to meet the ArcGIS geocoder standards and loaded into the geodatabase for processing. The formatting of the address data will include ensuring fields with the full street address and town name are populated in the dataset.

(3.2) Geocode Process
Formatted address data is geocoded using the ConnectME street centerline dataset. The address locator style used in this process is the ArcGIS *US Streets with Zone*. For this process, the city fields of the ConnectME street dataset are utilized in the zone component of the locator.

(3.3) QC Geocode Result
Analysts review the address data geocode result for the following:

- Overall geocode hit rate
- Town geocode hit rates
- Data anomalies

If address data fails any of these checks the data will not pass QC acceptance.

(3.4) Manual Correction of Data
Address data that has not passed the QC acceptance is evaluated for corrections necessary for the data to pass QC acceptance. Corrections to town names and updates to street names are commonly required to match the naming conventions in the ConnectME roads dataset.

(3.5) Geocode Process Tool
Sewall has developed an ArcGIS tool named *Geocode Process Tool* that translates the accepted geocoded address data into tabular address range records related to the accompanying ConnectME street centerlines. This tool is shown in ***Figure 4*** below.

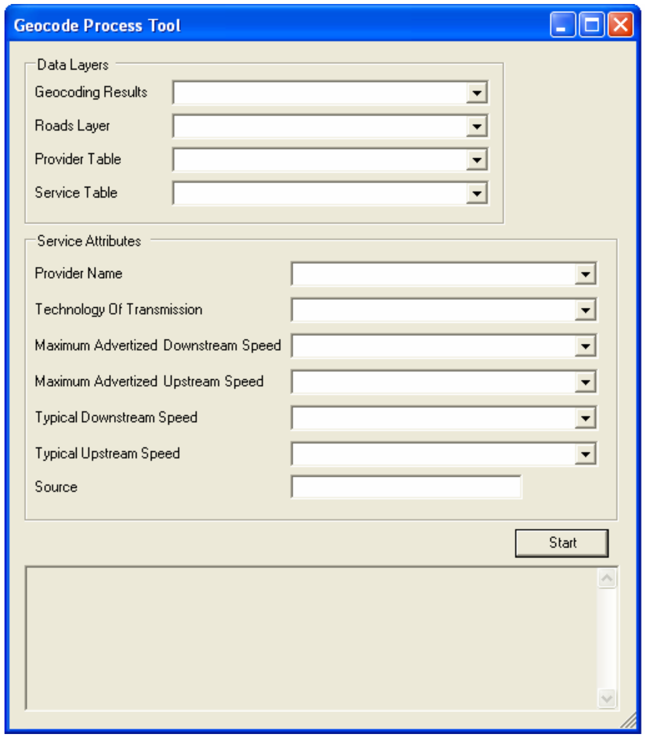


Figure 4 - Geocode Process Tool
Data Layers:(1) Geocoding Results - geocoded layer of address data (2) Roads Layer - ConnectME roads data layer (3) Provider Table - table of provider specific information (4) Service Table - broadband service output table where the service provider street address ranges are stored.
Service Attributes: The first six values are necessary to populate fields in the deliverable. Source is used to designate that the records created are from the Geocode Process Tool.

In ArcMap the user specifies which layers in the map correspond with the data layer inputs for the tool as well as the service provider service attributes that correspond with the geocode address point layer. Once the information is set the user clicks Start and the process begins.

Each geocoded address point within the geocode layer has as an attribute the street segment that the address was geocoded to. Using this street link, the tool can locate all of the geocoded address points assigned to a given street segment and build a modified street range of broadband service for the street segment. The tool then creates a record in the Broadband Service table that contains a link to the street segment in the ConnectME street feature class and populates the record with the derived broadband service street segment range and specified service provider information. This process is repeated for each unique street segment listed in the geocoded address point layer.

(4) GIS Polygon Data
The GIS polygon data bin is for service provider data that represents a coverage area of broadband availability and is delivered in a GIS format.

(4.1) QC Polygon Data

Datasets from the GIS polygon data bin are reviewed by an analyst. The QC routine ensures that the data has spatial integrity and includes the necessary attribution for inclusion to the state broadband project.

(4.2) Polygon Cover Tool

Sewall has developed an ArcGIS tool named *Polygon Cover* that converts service provider coverage area polygons into street segment related tabular records. Each tabular record created by the tool incorporates the service provider broadband specification information as well as modified street ranges representing provider street coverage.

This tool was initially created by Sewall for use on the fixed wireless viewshed datasets but was incorporated into the wired workflow for service providers that provided polygon regions of service coverage.

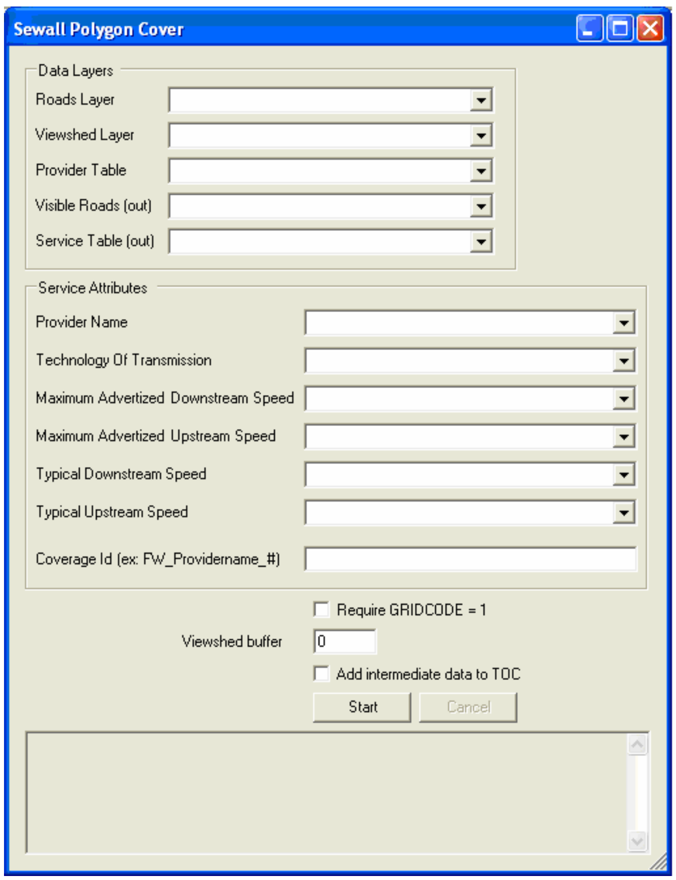


Figure 5 - Polygon Cover Tool

<u>Data Layers:</u> (1) Roads Layer - ConnectME street centerline data layer with address ranges (2) Viewshed Layer - viewshed layer used in delineating visible polygons for clipping road segments. For wired providers this would be the polygon layer that depicts a provider's coverage area. (3) Provider Table - internal processing flag (4) Visible Roads (out) - output feature class that stores the clipped road segment geometry (5) Service Table (out) - output table that the extracted address ranges populate.
<u>Service Attributes:</u> The first seven values are necessary to populate fields in the deliverable.
<u>Require GRIDCODE = 1:</u> Toggle is unchecked when running a wired broadband provider dataset that is represented as a coverage area.

In ArcMap the user specifies which layers in the map correspond with the data layer inputs for the tool as well as setting the service attributes for the service provider polygon layer. While running the Polygon Cover tool for fixed wired service regions analysts ensure the *Require GRIDCODE = 1* toggle is unchecked. Since this tool was initially created for use with a viewshed polygon output, the tool will not run on a non-viewshed layer unless this toggle is unchecked. Once the information is set the user clicks Start and the process begins.

The tool selects street segments from the input Roads layer that intersect the input polygon coverage and exports the street segments to a separate working file. These streets are then clipped to the polygon coverage. Next the tool runs a length ratio process that assigns each street segment a fractional value based on the clipped and original lengths. The tool then populates modified street range attributes based on the length ratio of a segment and the original street range of a segment. These modified street range values represent the broadband service street range of the provider. For each street segment the tool also creates a record in the Broadband Service table that contains a link to the original street segment in the ConnectME street feature class and populates the record with the modified broadband service street segment range and specified service provider information.

(5) GIS Street Data
The GIS street data bin is for wired broadband provider data at the street segment level that is delivered in a GIS format.

(6) GIS Block Data
The GIS block data bin is for provider data that is delivered at the census block level in a GIS format.

(7) Flat File Block Data
Census block service data delivered in a flat file format is stored in the flat file block data bin. Examples of flat file data are spreadsheets, text files and database files.

(7.1) Join Census Block Spatial Data
Flat file block provider coverage information is joined to a spatial census block layer using the full census block id value. Blocks with provider information joined are exported creating a spatial representation of the provider's census block broadband coverage.

(8) Flat File Street Data
The flat file street data bin is where provider data is stored when Sewall receives street level information in a format that cannot be associated spatially. Examples of files types delivered in a flat file format are spreadsheet, database and text file.

(9) Unreferenced Map Data
Provider data that cannot be referenced in ArcGIS are stored in the unreferenced map data bin. Examples of this type include paper maps and *PDF* documents.

(10) Census Block GIS Data
This data is Census 2000 block data in GIS format for the state of Maine that has been downloaded from the US Census website.

(11) Load Layers to GDB
Provider GIS data is loaded into the Sewall SDE geodatabase. A feature class is created for each provider's dataset. Sewall workflow tracking attributes are added to the feature classes.

(12) QC Data
Datasets are sent to a Sewall analyst for QC. The QC routine is to ensure that the data includes the necessary information for inclusion to the state broadband project. Provider data is cross-referenced with information on broadband availability that has been gathered from other sources. The QC of datasets with spatial data includes additional QC routines to ensure spatial integrity.

(13) Request more information on data source
Broadband provider data that does not meet the QC acceptance criteria Sewall initiates a request order to the provider for additional information. This request includes a detailed listing of the deficiencies found in the data as well as inquiries regarding spatial inaccuracies and anomalies discovered in the analysis.

(14) Infill Process
Sewall developed a tool named *Infill* to interact with the ConnectME street segments and populate related tabular records for fixed wired service provider availability. The *Infill Tool* allows a user to configure a specific set of service provider parameters, select ConnectME street segments, and then view and edit the related broadband availability information in the Broadband Services table that corresponds with the configured attributes. This tool is used to input fixed wired broadband availability data that Sewall received as census block, street or unreferenced map data. The majority of fixed wired service provider datasets utilize the *Infill Tool* for processing. A screenshot of the configuration dialog box is shown as Figure 6 below.

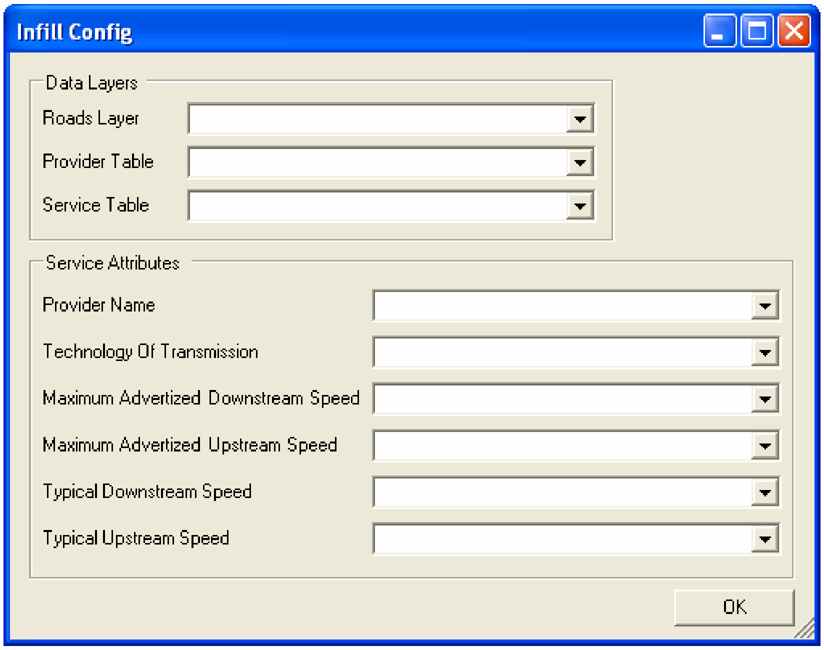


Figure 6 - Infill Tool Configuration
Data Layers: (1) Roads Layer: ConnectME roads data layer (2) Provider Table: Internal processing flag (3) Service Table: Broadband Service output table where the service provider street address ranges are stored.
Service Attributes: These fields are necessary to populate fields in the deliverable.

The first time a user uses the *Infill* tool in an ArcMap session, the *Infill Config* screen appears. The user enters the input data layers and the attributes for the service provider dataset that the tool will utilize during processing.

Once the *Infill Config* screen has been set a user selects one or more ConnectME road segments. Using the unique primary key values of the selected streets and the specified provider name and technology of transmission the tool searches the Broadband Services table for existing matching tabular records. If matches are found from this search, the tool reports the information in the *Infill* window. For selected street segments where no match was found in the Broadband Services table, the tool populates the *Infill* window with street segment road name and street range attributes representing potential broadband service ranges for the provider on the selected streets. These street range attributes can be updated in the *Infill* window based on provider sources. This *Infill* tool window is shown as ***Figure 7***.

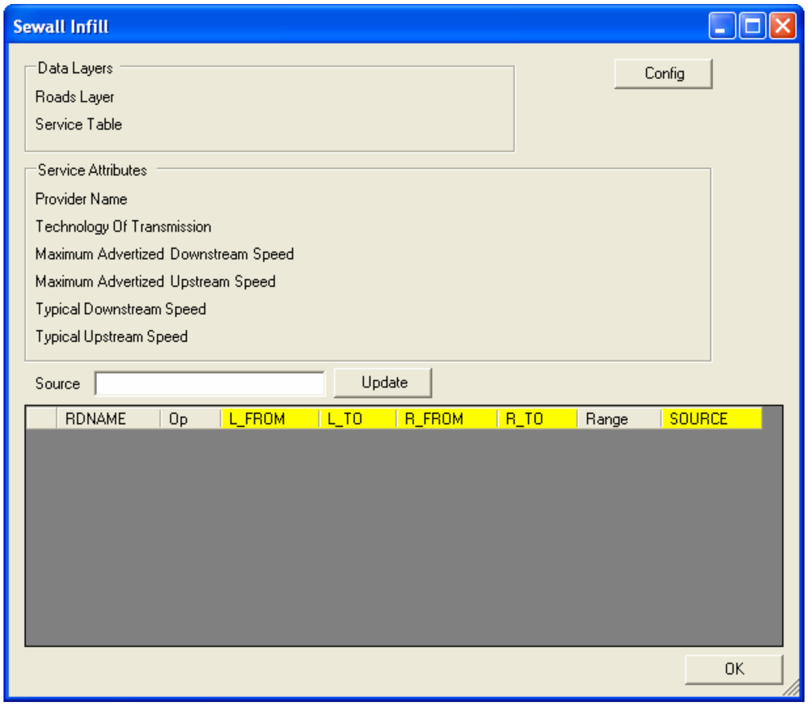


Figure 7 - Infill Tool
Data Layers: (1) Roads Layer: ConnectME roads data layer (2) Service Table: Broadband Service output table where the service provider street address ranges are stored
*Config: Opens the Infill Config window (**Figure 6**)*
Service Attributes: These fields are necessary to populate fields in the deliverable.
Source: Internal flag for source of service availability
Update: Updates selected tabular records SOURCE field to the value entered in the Source field
Tabular Record Attributes: (1) RDNAME: Name of ConnectME road segment (2) Op: Operation being performed {INSERT-new tabular record, UPDATE-update existing tabular record, DELETE-delete tabular record} (3) L_FROM: "Left from" broadband address range of ConnectME road segment (4) L_TO: "Left to" broadband address value of ConnectME road segment (5) R_FROM: "Right from" broadband address value of ConnectME road segment (6) R_TO: "Right to" broadband address value of ConnectME road segment (7) Range: Reports either "full" or "partial" and is a comparison for each tabular record of the broadband provider street range to the accompanying ConnectME street range (8) SOURCE: Internal process flag.

Once the user has reviewed the values pressing OK will perform the operations listed in the *Op* field.

(15) ConnectME Street Data

The ConnectME street data bin contains the street centerline dataset used in the geocode and street relate processes. The Maine Office of GIS E-911 street centerline file was used to create the base street segments and gives the project the most accurate street centerline file for the State of Maine. The NAVTEQ street centerline dataset NAVSTREETS was utilized to infill street segments in areas where gaps were assessed in the MEGIS E-911 file.

(16) Relate Process

Through the use of Sewall developed tools the data gathered for fixed wired broadband service providers gets stored in the Broadband Services table as availability street ranges associated with street centerline segments. Each record in the Broadband Services table is associated by a foreign key/primary key relationship with a street segment in the ConnectME street centerline dataset. This relationship allows for clean and easy access to street level availability of service providers.

(17) Broadband Services Table
The Broadband Services geodatabase table was developed by Sewall to store broadband service provider information and street range coverage. NTIA requirements and formats were utilized when creating the fields to ensure the records stored in the Broadband Service table are compatible with the SBDD data model.

(18) Intelligent Street Centerlines
The output from the fixed wired workflow is a comprehensive intelligent street centerline network comprised of street centerlines and related service availability tabular records.

4.2 Fixed Wireless Transmission

The initial stage of mapping terrestrial fixed wireless service territories depends on the quality of the data received. To process any service footprint of a particular transmitter, the initial resources acquired during the data collection phase of the project are critical.

Terrestrial Fixed Wireless technology is clouded by many variables that determine the overall performance of each transmitter signal. Inaccurate data pertaining to location, height of a transmitter, horizontal and vertical limitations, signal range and many more factors present potential obstacles to producing an accurate representation of any transmitter's service footprint. Some of these factors have not been considered during the mapping process due to lack of data needed for modeling them. For example, while a 10-meter DEM is used to represent the surface terrain, we have not incorporated obstructions on the surface such as trees and other man-made obstacles that could influence a transmitter's propagation model.

The data collection process and subsequent conversion workflow is designed to accommodate a variety of data sources received from the service providers and production tools have been developed to build efficiencies and quality control into the workflow. When received by the service providers, supplemental data is used throughout the conversion workflow to help verify the mapping results. However, a larger scale verification process is described in **Section 5**.

The data conversion process for fixed wireless transmission is represented by ***Figure 8*** and described below.

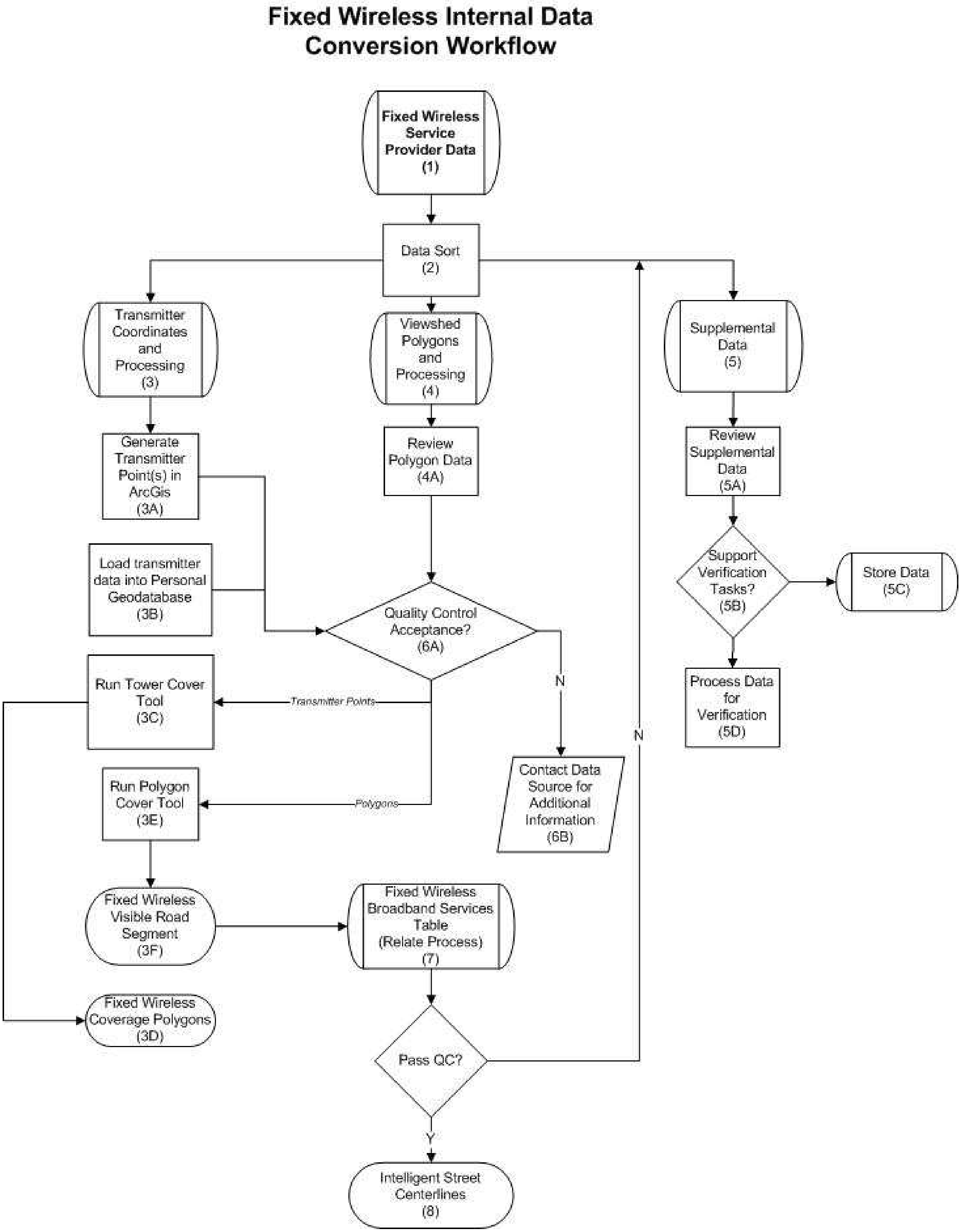


Figure 8 - Fixed Wireless Internal Conversion Workflow

(1) Fixed Wireless Service Provider Data
Service provider data gathered during the data collection phase. Data is cataloged in separate folders by provider and managed according to task and technology of transmission.

(2) Data Sort
The data sort phase of production immediately follows the data collection process. During this task, a thorough review of the service provider data determines the type of data received. Fixed wireless data generally consists of three types: transmitter coordinates and attributes, pre-defined polygons and attributes, and supplemental data. Each type of data follows unique internal processing steps.

(3) Transmitter Coordinates and Processing
Transmitter coordinate data is essentially the raw data necessary to generate a viewshed for each transmitter. In order to be processed, the transmitter source data must have certain required fields such as latitude and longitude, spot (ground elevation), equipment height at the transmitting and receiving ends, horizontal and vertical limitations, and range of transmission. The content of the transmitter data is carefully reviewed for completeness and overall consistency prior to the next step. Once completed, the data is imported into ArcGIS for continued processing and quality control.

(3B) Load Transmitter Data into Personal Geodatabase
Using the newly scrubbed *.csv* file, transmitter points are created in ArcGIS and the transmitter location points are displayed. A final comparison against supplemental data is performed to ensure the transmitter locations are in the correct locations. Supplemental data includes such layers as imagery, political boundaries, and road centerlines.

(3C) Run Tower Cover Tool
This tool was designed and developed by Sewall to batch process 1 or more transmitter point viewsheds. A screenshot of the tool is shown below as ***Figure 9***.

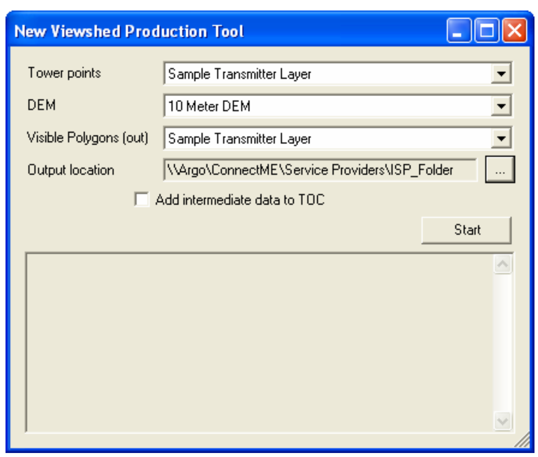


Figure 9 - Tower Cover Tool (Viewshed Production)
<u>Tower Points:</u> The data layer containing records of all transmitters that need a viewshed generated. Originally received from ISP and pre-processed by Sewall for format compatibility.
<u>DEM:</u> 10-meter digital elevation model obtained from MEGIS as the primary surface model for generating the viewshed

Visible Polygons (out): Visible polygons (only) output to an SDE layer
Output location: Location of output to personal geodatabase workspace to be used for additional processing.

(3D) Fixed Wireless Coverage Polygons

The *Tower Cover Tool* generates raster data sets depicting the visible and non-visible surfaces representing each transmitter. As a final output, the tool extracts the visible components of the raster data and outputs to polygon vector layers stored in the SDE environment as supplemental reference data.

(3E) Run Sewall Polygon Cover Tool

This tool was designed and developed by Sewall to facilitate several production steps.

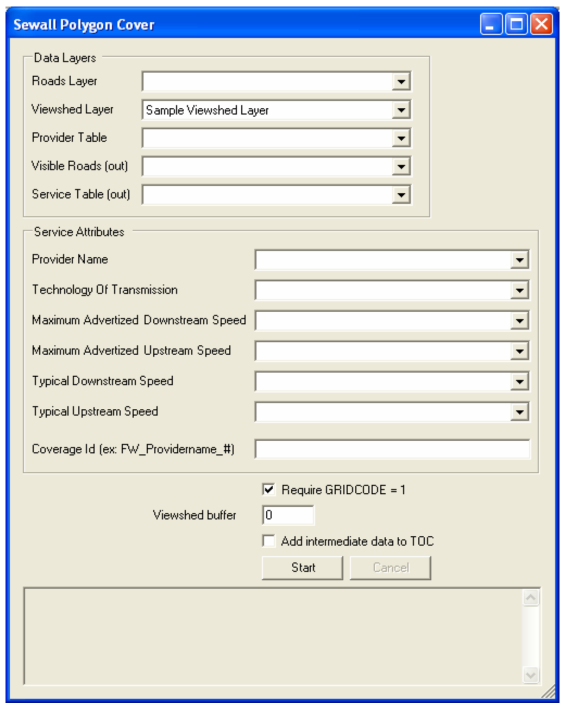


Figure 10 - Polygon Cover Tool
Data Layers: (1) Roads Layer - ConnectME Street data layer with address ranges (2)Viewshed Layer - viewshed layer used toe delineate visible polygons for clipping road segments (3)Provider Table - internal processing flag (4) Visible Roads (out) - output feature class that stores the clipped road segment geometry (5) Service Table (out) - output table that the extracted address ranges populate.
Service Attributes: These fields are populated, if data is available, to meet NTIA NOFA requirements.

(3F) Fixed Wireless Visible Road Segments

The *Polygon Cover Tool* clips road segments that are within visible polygon viewsheds and writes them out to a polyline vector layer stored in the SDE environment as supplemental reference data.

(4) Viewshed Polygons and Processing

Although not as common, another source of data received from the service providers is a polygon dataset that has already been generated to represent visible service territory of transmitters. Service providers or third party vendors will frequently run their own propagation models to be used for broadband mapping. Polygon formats include ESRI shapefiles, MapInfo files, Google *.kml* files, and raster files. Each format requires a thorough review to determine the subsequent processing steps.

(4A) Review Polygon Data
Although each format listed is unique, the data eventually runs through the *Polygon Cover* tool so that the address ranges within the polygons can be clipped out. Each format is carefully inspected for content, spatial characteristics and accuracy. The general workflow for each format is as follows:

- Shapefile: Review content > Edits > Project > QC > Load for processing > Run Sewall Polygon Cover Tool
- MapInfo: Review content > Translate to ESRI shapefile > Edits > Project > QC > Load for processing > Run Sewall Polygon Cover Tool
- Google *.kml*: Review content > Translate to ESRI shapefile > Edits > Project > QC > Load for processing > Run Sewall Polygon Cover Tool
- Raster: Review content > Translate raster to polygon > Edits > Project > QC > Load for processing > Run Sewall Polygon Cover Tool

(5) Supplemental Data
Supplemental data received by service providers is generally used for verification to support internal processing results. It is not used as a data source to generate transmitter locations or viewsheds. Supplementary data includes, but is not limited to, failed service locations, customer service locations, hard copy plots, *PDF* files, and other digital reference files. In most circumstances, the data can be used for cross-referencing.

(5A) Review Supplemental Data
Each format is unique and so are the processing steps that are necessary to prepare the data for use.

- Failed Service Locations: Provides an excellent source for cross-referencing to viewshed polygons (visible and non-visible) but must have complete address in order to geocode location of address.
- Customer Service Locations: Provides an excellent source for cross-referencing to the viewshed polygons (visible and non-visible) but must have a complete address in order to geocode location of address.
- Hard copy plots: May be used for verification purposes if the content of the material is applicable.
- *PDF* files: May be used for verification purposes if the data content is applicable.
- Other data sources: All sources are reviewed for potential use.

(5B) Support Verification Tasks
Supplemental data sources are reviewed to determine if they hold any value to the project workflow. Value added data will be stored and utilized as needed to support internal processing.

(5C) Store Data
Data received from service providers that does not have any given value to the project is organized and stored under the service provider folder.

(5D) Process Data for Verification Tasks
Supplemental data sources are scrubbed for compatibility and processed.

(6) Quality Control Acceptance
Quality control procedures are implemented at each of the three production stages depending on the data (transmitter coordinates, viewshed polygons, or supplemental data). Because the service provider data is received in numerous formats, styles, and content, much of the initial QC is completed during the data collection stage. When data is received from a service provider, an initial review is done to determine what is received and what is outstanding. This cycle of communication with the providers continues until all the necessary data is either received or clearly understood that it will not be received. Throughout the data collection process, Sewall keeps an inventory of receivables.

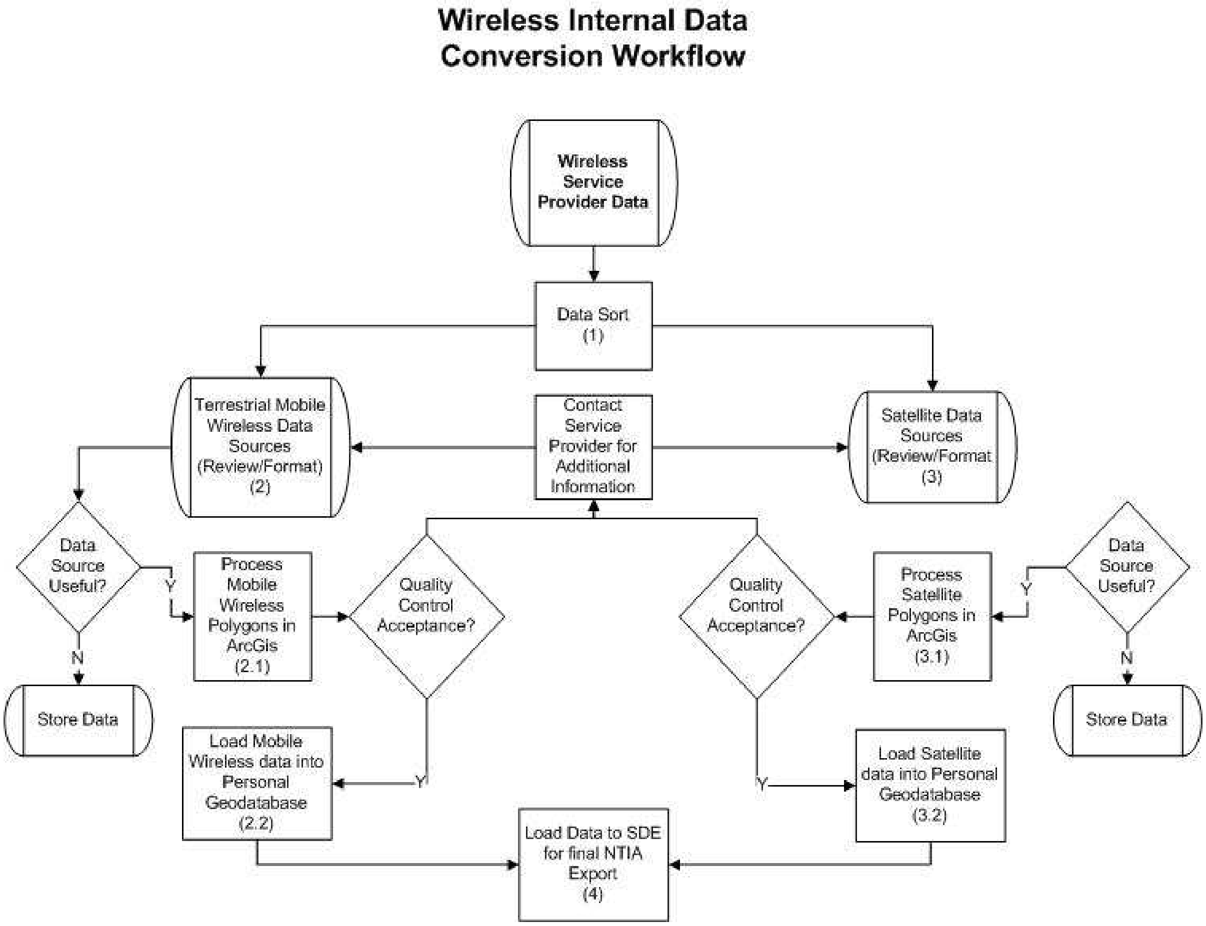
(6A) Contact Data Source for Additional Information
During the data collection phase of the project, questions or clarifications may have been overlooked, or items may present road blocks at some point later during the processing. If an internal quality review does not resolve an issue, the service provider is contacted for additional information or clarification.

(7) Fixed Wireless Broadband Services Table (Relate Process)
The *Polygon Cover Tool* has two outputs, both generated using the visible polygons created by the *Tower Cover Tool*: (1) road segments, and (2) calculated address ranges. While the visible road segments are not part of the NTIA deliverable, they are stored as a reference file named *CONNECTME.FW_VISIBLE_ROAD_SEGMENTS*.

(8) Intelligent Street Centerlines
The output from the fixed wireless workflow is a comprehensive intelligent street centerline network comprised of street centerlines and related service availability tabular records.

4.3 Mobile/Satellite Transmission

Wireless broadband technology consists of all facilities-based providers of wireless broadband service that is not address specific. For the State of Maine, this includes terrestrial mobile wireless and satellite broadband service. Mapping mobile wireless and satellite coverage requires less processing than other technologies that are address-based, such as wired and fixed wireless service. Data consists of polygons generated by the providers or third party vendors, representing areas where broadband service is offered. As shown in the workflow below, the data received from providers is sorted, processed and loaded into a geodatabase. Minimal steps are required to process this data, but established internal workflows are taken to ensure that proper protocols and quality assurance are met. The



primary steps of the internal workflow are shown in ***Figure 11*** and described below.
Figure 11 - Wireless Internal Conversion Data Workflow

(1) Data Sort
Upon receiving data from a mobile or satellite service provider, Sewall initially sorts and stores the data by technology - terrestrial or satellite.

(2) Terrestrial Mobile Wireless Data Sources (Review)
After the data is sorted, an initial data analysis is performed to determine if the data received appears to be intact spatially and is accompanied by the proper attribution required for adherence to the SBDD data model. Follow-up with the service provider continues until all necessary information is acquired.

(2.1) Process Mobile Wireless Polygons in ArcGIS
After determining that the data has value, the polygons are projected into the proper coordinate system to complement the internal workflow. Depending on the source data, additional data processing routines may be necessary before loading the data into the geodatabase.

(2.2) Load Mobile Wireless data into Personal Geodatabase
Although the primary quality control procedures are completed during the verification process, initial acceptance testing to ensure the data is spatially valid is performed by cross-referencing to additional data sources such as aerial imagery or information taken from the service provider website. Discrepancies are documented for use in subsequent verification processes. Once quality checks are complete, the data is loaded into a personal geodatabase

(3) Satellite Data Sources (Review)
When all the spatial and attribute information is received, the satellite data follows the same internal workflow as mobile wireless data (Steps 2, 2.1 and 2.2).

(4) Load Data to SDE for final SBDD Export
Mobile wireless and satellite data is loaded to SDE environment for final export to SBDD format.

4.4 Middle Mile Locations

Middle Mile and Internet Backhaul Connection Points are defined by NTIA as "interconnection points that typically enable relatively fast data rates, are built to handle substantial capacities, and may be service-quality assured." At this stage of the mapping, middle mile data has been the most difficult to obtain from service providers during the data collection process. Service provider networks can include as little as one middle mile location such as a backhaul connection point or as many as dozens, operating as interconnection points within a fixed wireless network reaching out to end users. Furthermore, some service providers may offer middle mile connection points only as a service, such as a splice into a fiber line to support a lateral to a central office or business.

Regardless of the technical framework, all middle mile locations that meet the NTIA definition are captured in a point feature class with additional attribution including the ownership of the facility, serving facility capacity and serving facility type.

The outline of workflow is shown as ***Figure 12***. The description of each step follows.

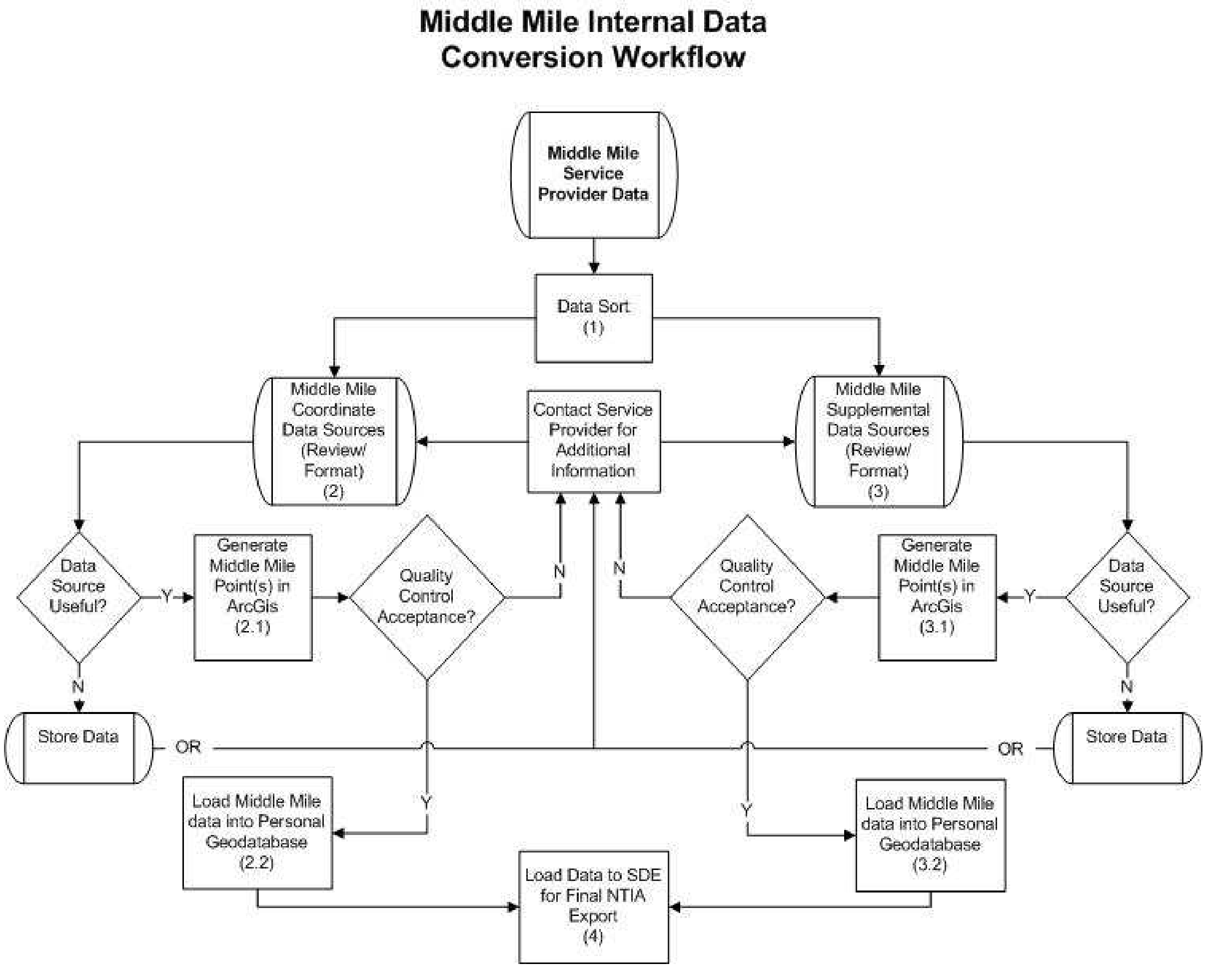


Figure 12 - Middle Mile Internal Data Conversion Workflow

(1) Data Sort

The initial data sort separates the data and distinguishes formats more compatible to the database model, such as middle-mile coordinate values listed in a spreadsheet or ESRI shapefiles. Data received in compatible formats require minimal processing steps. Supplemental data sources generally require additional processing steps. Examples may include the conversion of *.kml* files to ESRI shapefiles or polyline files that require points to be added at splice or lateral connections.

(2) Middle Mile Coordinate Data Sources Review

Sewall reviews the data to ensure that the information is a valid input. If so, the data is reformatted and loaded into in ArcGIS. Sources deemed as invalid are stored, or the service provider is contacted for additional information if necessary.

(2.1) Generate Middle Mile Points in ArcGIS

Points are loaded into ArcGIS. Sewall analysts run acceptance procedures to verify data translation to ArcGIS and spatial accuracy and completeness using supplemental data sources provided such as addresses, imagery or descriptive information about the point locations. In addition to the point geometry, all attribution carried over in the translation is confirmed.

Conflicts or questions are referred back to the service provider for further clarification if necessary.

(2.2) Load Middle Mile Data into Personal Geodatabase
Middle-Mile data is loaded to a personal geodatabase. Additional data received by the service providers or revisions will cycle through the same process and be stored in the personal geodatabase prior to loading to the SDE environment for final export.

(3) Middle-Mile Supplemental Data Sources (Review)
Supplemental data sources may involve additional processing during this step in order to proceed. Some of the more common supplemental data sources include, but are not limited to, the following:

- Google *.kml* files
- *.jpg* images showing middle-mile locations
- AutoCAD point or polyline files
- e-mails with descriptions of locations
- Other miscellaneous information

Once the data has been fully reviewed and normalized, the remaining steps follow the same internal workflow as coordinate data sources (Steps 2.1 and 2.2).

(4) Load Data to SDE for final SBDD Export
Middle mile data is loaded from the personal geodatabase to the SDE environment for final export to SBDD format.

4.5 Service Overview

Broadband service providers that participate in the state broadband mapping project have been asked to provide broadband service territory footprints at the address, street, census block or county level. The service overview dataset contains the information that has been delivered at the county level.

The workflow developed by Sewall integrates the gathered data from broadband service providers into a consistent spatial format that is stored in a geodatabase designed to be compatible with the SBDD deliverable.

The service overview workflow is described below and depicted in ***Figure 13***.

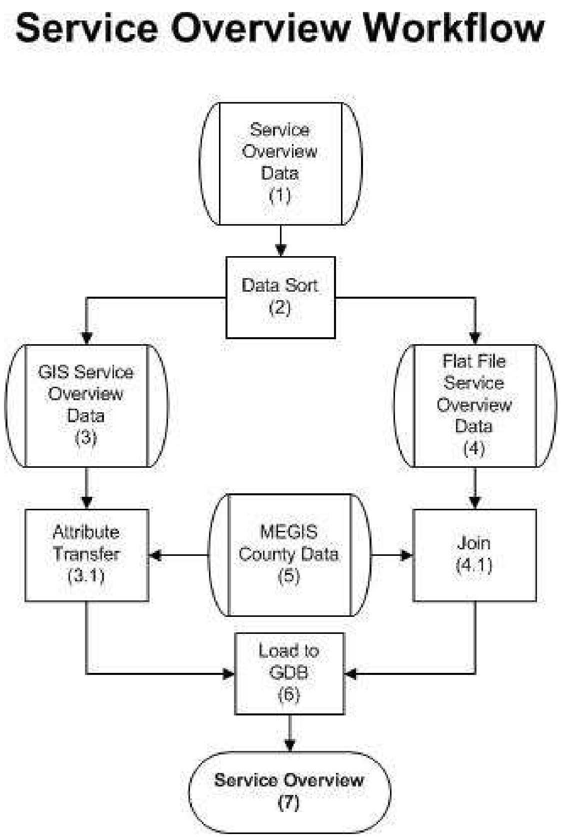


Figure 13 - Service Overview Workflow

(1) Service Overview Data

The Service overview data bin is the storage location for service overview specific broadband service provider data gathered by Sewall. Sewall specifies what information is necessary for this deliverable and what formats are acceptable when contacting each provider during the data gathering phase of the project.

(2) Data Sort

The service overview data is sorted into categories by data type.

(3) GIS Service Overview Data

The GIS data bin is used to store provider data that has been delivered to Sewall with service overview attribution and is in the requested GIS format.

(3.1) Attribute Transfer

Attributes contained in the GIS data are sent through an attribute transfer process that populates county data from the MEGIS County data. This step ensures that there is one consistent spatial dataset utilized as a basemap in the service overview.

(4) Flat File Service Overview Data

The flat file data bin is used to store provider data that has been delivered to Sewall with service overview information in a flat file format.

(4.1) Join
Using county name information provided in the flat files the MEGIS county data is joined to the flat files. The joined dataset is exported and stored in the GIS service Overview data bin.

(5) MEGIS County Data
The shapefile *cnty24p.shp* was downloaded from the MEGIS website (megis.maine.gov) and utilized for county spatial representation of the service overview dataset during the workflow.

(6) Load to Geodatabase
Once the service overview data has been processed, the data is reviewed for content and accuracy and then loaded to the ConnectME production database.

(7) Service Overview
The output of the service overview workflow is a polygon dataset that is compatible with the SBDD data model.

5 Data Verification

The verification process is used to ensure that the data delivered is in fact valid and current. For verification purposes, the broadband data provided is sorted into two groups: (1) mobile coverage, and (2) wired and fixed wireless and satellite coverage.

Mobile coverage consists of data from providers who offer mobile broadband services to consumers through devices such as smartphones or mobile laptop aircards. Common providers of this type of broadband service in Maine are AT&T, Verizon Wireless, and Sprint.

Wired and fixed wireless coverage consists of data from providers who offer broadband service through wired services (such as cable and DSL) and fixed wireless transmitted signals. Common providers of this type of broadband service in Maine are FairPoint Communications, Time Warner Cable, Axiom Technologies, and Aroostook Internet. Satellite coverage consists of data from providers who offer broadband service via satellite. Common providers of this type in Maine are Starband Communicates and WildBlue Communications.

Once the data has been collected, processed and verified, the results are statistically analyzed and plotted atop the original provider data coverages in GIS. Any ‘holes’ or inconsistencies in the data from the service provider are reported to the provider in a feedback loop to ensure all parties involved are aware of the potential issues with the broadband service in an area.

5.1 Mobile Coverage Verification

Mobile coverage data is received by Sewall from the service providers in the form of GIS polygon files. After these files have been reviewed and properly projected (see **Section 4.3** for details), they can be analyzed in the verification process. The mobile coverage file is compared against the State of Maine boundary file in a GIS application in order to assess the size and location of the coverage area with respect to the State.

The methodology developed by the ConnectME Authority to verify mobile coverage in Maine is to select a series of points throughout a provider’s coverage and have field crews run tests at these predetermined locations. A minimum of 37 points per coverage area are needed in order for the statistical analysis on the field data to be valid.

To select the points for field verification, a 28-square-mile grid was created in GIS and layered with the provider’s coverage area, the E911 road layer and the state boundaries. One point was placed per grid block within the provider’s coverage network. Each point was placed on a road, usually at road intersections for ease of access by the field crew. Once all the points were placed, the points were divided into groups for distribution to field crew personnel.

The points were assigned attributes of point ID, latitude and longitude. The attribute table was then exported to an Excel file for further editing. The columns *field connect*, *upload* *speed*, *download speed* and *notes* were added to the spreadsheet. The *field connect* column holds values to describe whether the field crew was able to log on to the provider's network., speeds collected from the state website at that location are stored in the *upload speed* and *download speed* columns. The spreadsheet was loaded onto the field laptops for data entry.

Crews utilized Microsoft Streets & Trips to assist in navigating to each of the field points across the state. The software, which was loaded on each of the field laptops, has a GPS component that could track and direct field crews. The spreadsheet used for data entry was also loaded into the software so the points could be plotted based on given coordinates. The field crews could properly identify each of the points based on the *Point Nam*e attribute.

The program turned each of the points into a "stop." The start and ending points of the trip were also added, allowing the software to calculate an optimized route to reduce driving time and mileage. After optimization, the software also provided driving directions, which were saved and loaded onto the field laptops.

Mobile broadband aircards from each of the mobile service providers were purchased outright directly from the providers. This eliminated the need for a service contract so that the aircards can be deactivated after the verification process without a contract cancellation fee. Service providers activated the mobile aircards with a month-to-month data package of 5GB.

Aircards from each of the providers were then loaded onto the field crew laptops. The software from the aircards was installed, aircard functionality was checked, and any updates were installed prior to crews leaving the office.

Each time verification tasks are performed, the points are visited by a field crews who are equipped with a field laptop enabled with the mobile broadband aircard of the corresponding service provider and proper navigation information. The field crews drive to each of the points, log onto the service provider's network and navigate via Internet Explorer to an internet speed test website created by the James W. Sewall Company specifically for the ConnectME Broadband Mapping Project.

For each test point, the point number, service provider and date are entered into the internet speed test website (e.g., Test_745_verizon_20100521) and a test is executed. Results are recorded both in the speed test database (automatically) and in the spreadsheet. Once all of the points are completed, crews return to the office and spreadsheets are combined. Data columns are filled in with corresponding broadband upload and download speeds for sites with connectivity.

Data points are then plotted on maps to view where broadband coverage is full strength or where it is lacking. If there are large 'holes' in the coverage areas, the points are revisited to ensure that readings were accurate and not subject to user or equipment error.

The merged field spreadsheets are then handed off to a statistician for the statistical analysis of the data.

5.1.1 Statistical Analysis for Mobile Coverage Verification

Large data sets are often expressed best in terms of summary statistics. It is often easier to look at commonly defined statistics (stats) to get a quick overview of what the data describes, than to look at all the raw data.

In analyzing this data, we chose statistics using the following criteria:

- Commonly used and understood
- Fit the data (data type) in question
- Had practical application to the reader in understanding what the data was describing

We believe that the statistics presented can be beneficial in several ways:

- Description/Summary: they consolidate many data observations into a few summary stats that can be quickly compared
- Quantification: they describe which portion of the data falls within or outside of the limits of acceptable criteria
- Reliability/Prediction: in some cases, they attest to the reliability of the data collection

The following statistics were used:

- Number of samples (n): number of data points in the sample
- Average (xbar): arithmetic mean or the mean value of a set of integers, terms, or quantities, expressed as their sum divided by their number.
- Standard Deviation (sd): used as a measure of the dispersion or variation in a distribution, equal to the square root of the arithmetic mean of the squares of the deviations from the arithmetic mean.
- Percentages (%): a proportion or share in relation to a whole; a part; a fraction or ratio with 100 understood as the denominator (e.g., 0.98 equals a percentage of 98).
- Hypothesis testing: statistical process used when trying to determine if it is reasonable to conclude that the entire population possesses a certain characteristic by the analysis of a sample.

Explanation of choices made:

- Quantitative statistics were only applied on sample data that fell within the published service area of the provider in question. This was possible because the area was "bounded" by the geographic area described in the "service area." Outside the service area there is no bound (limit), so these same statistics would not be reliable as used with our methodology.
- Assumed a normal distribution because this is the most common and typical distribution type for this type of data, and we had no evidence to counter this assumption.

- Chose sample statistics because we were not dealing with the whole population (almost unlimited sample points possible).
- Chose hypothesis testing because we wanted to have the most valid predictor of the population parameters given the variability of our sample data.
- Chose student's T-distribution when sample size was equal to or less than 30 (n=<30) and Z-test when populations were above 30 (n>30).
- Used one-tailed tests because we were interested in the area above the curve from a single lower parameter (criteria of minimum speed).

Procedures:

- Field procedure described in previous section

Statistics used in ConnectME Analysis of Service Coverage:

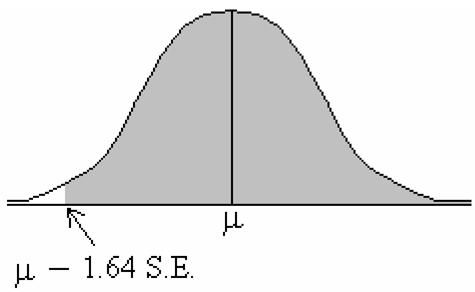
Data was sorted to yield only those sample points that fell within the published service area of the provider in question.

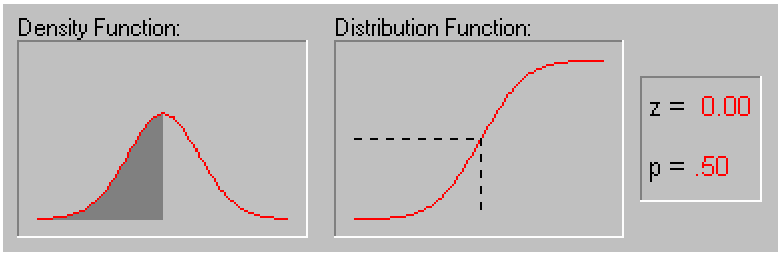
Then the following information was calculated:

- n = number of total sample points
- Degrees of Freedom (df) = n-1
- Selection of t-distribution (df < 30) or standard normal curve (df >= 30)
- Percent of points where connection was established
- Percent of points where both tested upload and download speeds were equal to or greater than (=>) broadband speeds (200 and 768 kb/sec respectively).
- Percent of points where either the upload or download speed was equal to or greater than (=>) broadband speed, but not both.
- Percent of points where neither the tested upload or download speeds was equal to or greater than (=>) broadband speeds.

Using all data points within the designated service provider coverage that registered an upload speed during the test, the following were calculated:

- Average # of points where a connection was made that had an upload speed equal or greater than broadband minimums.
- Average upload speed (xbar/upload)
- Standard deviation of the sample (SD/upload)
- Statistical prediction of percent of points that would meet minimum 3G upload speed in subsequent samplings (using one-tailed t-test or z-score, depending on df) – see schematic below





Using all data points within the designated service provider coverage that registered a download speed during the test, the following were calculated:

- Average # of points where a connection was made that had a download speed equal or greater than broadband minimums.
- Average download speed (xbar/download)
- Standard deviation of the sample (SD/download)
- Statistical prediction of percent of points that would meet minimum 3G upload speed in subsequent samplings (using one-tailed t-test or z-score, depending on df) – see schematic above.

5.2 Wired, Fixed Wireless, and Satellite Verification

In order to verify the existence of wired and fixed wireless coverage in an area, direct access to the provider's service is needed. Logistically this would not be an easy task because transmission receivers, accounts and other equipment would have been required for each of the providers. Instead, the project team has opted to gather information through other means including:

- Customer service locations and failed service locations
- Surveys
- Speed test results
- FCC Consumer Broadband Test (CBT) data
- FCC Form 477 aggregate data

5.2.1 Customer service locations and failed service locations

Service locations are where a provider is currently offering service and can usually be provided as an export from the company's billing system. Failed service locations track where a potential customer asked to obtain service, but the service could not be established. For DSL service, for example, this failure could be the result of the distance from the central office or a device in the loop make-up that prevents the DSL technology from passing through. For fixed wireless service this failure could be due to vertical obstructions that prevent the signal from reaching the location.

While service locations and failed service locations are helpful in trying to establish the extents of a provider's coverage, not all providers are willing or able to provide this data. If the provider does submit location data, the project team geocodes these locations and overlays the data on the coverage created from the provider's submission. Both service locations and failed service locations help identify anomalies and areas of interest in a provider's service coverage that require follow-up and further research.

5.2.2 Surveys

The project team is surveying residents and businesses in Maine with a questionnaire about their current internet connections. There are approximately 473,000 residences and 59,000 businesses in the State of Maine. The ConnectME Authority has opted to survey yet undetermined percentage of the residences and businesses. A rate of survey return of between 3% and 5% is expected between the hardcopy and electronic versions of the survey.

The current draft of the survey is comprised of 10 questions and takes 1.5 minutes to complete. There is also an online version of the survey which people can access by navigating to a link indicated on the delivered hardcopy of the questionnaire. The electronic version, once completed, directs the person to the ConnectME internet speed test website, which reports the upload and download speeds of the user's internet connection. The speeds are recorded in a database that tracks entered physical address and speed test results for future analysis.

The survey also identifies the consumer by the physical address, which can be geocoded against a street centerline file in GIS to create a point file. The data associated with each address (e.g., transmission type and provider) is analyzed by layering the consumer information with the coverage data provided by the service provider. Sewall can analyze the layers to verify if each service provider does cover the areas represented by the data it submitted. In addition, if an area shown to have no service by a provider appears in the consumer survey, the provider in question can be contacted to confirm and provide updated coverage information.

5.2.3 Speed test results

For the SBDD project, the ConnectME Authority has implemented an online speed test tool. The website was developed by Ookla Net Metrics and was brought online on 13 January 2010. Through 30 September 2010, 10,512 tests had been recorded. The speed test stores downstream and upstream speeds as well as the user's address and the user's ISP. The results from the speed test tool are geocoded and the information is used to help verify service coverages and service speeds for wired, fixed wireless, and satellite providers.

5.2.4 FCC Consumer Broadband Test (CBT) data

The Consumer Broadband Test data provided by the FCC consists of three datasets: Speed Test records, Mobile Broadband Speed Test records, and Broadband Dead Zone Report records. The project team plans to incorporate the FCC speed test records along with those records captured by the ConnectME speed test tool. However, the name of the service provider is not included with data, so a method for mapping the IP address in these records to the appropriate provider must be developed.

The dead zone reports are used to identify locations reported to be without coverage. The addresses from these records are geocoded and then are cross-referenced with service provider coverages in the areas.

5.2.5 FCC Form 477 aggregate data

The FCC has provided SBDD grantees and their teams access to the FCC Form 477 aggregate data. This data contains information on service providers in Maine at an aggregate or granularity higher than the SBDD data, but is useful for checking the list of providers and their locations at Census Track level.

The project team has recently developed a tool that compares the records in the Form 477 aggregate data to the provider data in the SBDD project database. The tool lists out by Census Track each provider that includes the tract in the Form 477 filing. Each provider that has service data that falls within the tract is considered a match. Using this data, the team has been able to find potential providers that were not previously included in the study, as well using the tract locations as a cross-reference to where each provider has service. The team has plans to further enhance the tool to provide a set of results centric to each provider.

6 Validation

Once initial broadband service territories are mapped, the Sewall project team will generate maps for each provider company representing the current status of data received by Sewall at the time of the mapping. This will give each service provider the opportunity to validate their broadband service footprint and provide feedback to the Sewall project team.

Figure 14 below represents a fixed wired validation map where a provider company's broadband service (DSL) foot print is symbolized in red. Depending on the size of a service footprint and map density, additional information, such as road names, may be represented.

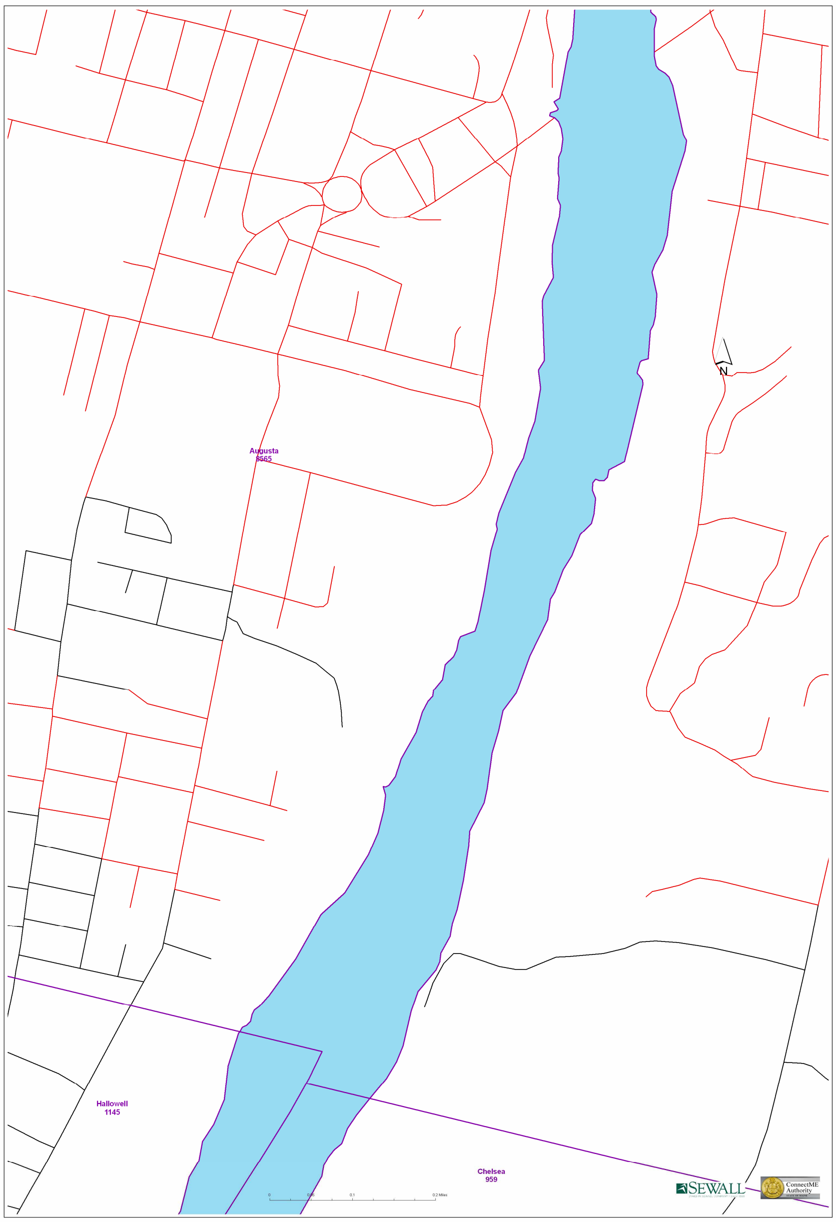


Figure 14 - Fixed Wired Validation Map

7 Data Delivery

Service provider data that has been processed to the Sewall production model needs to be transferred to the SBDD data model for delivery. In order to accomplish this Sewall has developed a process by which the Sewall production datasets are exported to the current SBDD data model structure.

The Sewall production model was designed with the NTIA delivery model in mind and, in as many cases as possible, the production model utilizes the NTIA delivery defined attribute definitions and domain values. Through the use of this design philosophy, Sewall has mitigated the pitfalls for exporting to the SBDD data model.

To facilitate the transfer of data stored in the Sewall production model to the SBDD model for delivery Sewall has developed an ArcCatalog tool named *State Broadband Data Export*. This tool reads a source geodatabase set of features and writes to a destination geodatabase set of features. A screenshot of the tool dialog box is shown in ***Figure 15***.

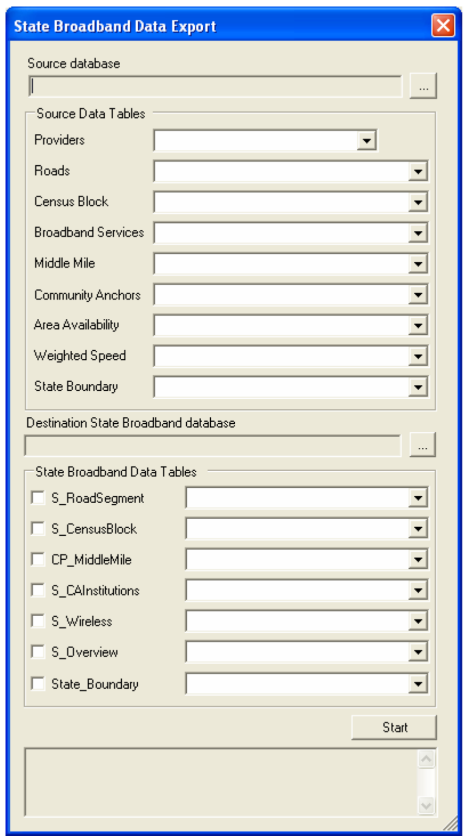


Figure 15 - State Broadband Data Export Tool
<u>Source database:</u> Sewall production geodatabase location.
<u>Source Data Tables:</u> (1) Providers - Geodatabase table with list of provider specific information (2) Roads - ConnectME street centerline feature class (3) Census Block - Census 2000 block geodatabase feature class (4) Broadband Services - Geodatabase table containing broadband provider characteristics and street ranges linked to ConnectME street centerline segments (5) Middle Mile - Geodatabase point feature class containing broadband service provider middle mile locations (6) Community Anchors - Geodatabase point feature class containing community anchor institution locations (7) Area Availability - Geodatabase polygon feature class

containing mobile wireless and satellite broadband provider coverage (8)Weighted Speed - Geodatabase polygon feature class service overview data (9) State Boundary - Geodatabase polygon feature class portraying the Maine state boundary.
Destination State Broadband database: *SBDD geodatabase location.*
State Broadband Data Tables: *These are the required SBDD deliverables.*

On launching the ArcCatalog tool, the user selects the source and destination geodatabases for the transfer process. The source geodatabase is the Sewall internal production model, and the destination geodatabase is the empty SBDD model. Next the user matches the items listed in the Source Data Tables section to the production model features. Once complete, the user checks which deliverables the tool will export in the State Broadband Data Tables section. Clicking Start will begin the export process.

The road segment and census block exports are performed simultaneously in the *State Broadband Data Export Tool* with road segments being reported in census blocks greater than 2 square miles and census blocks being reported in areas up to 2 square miles. The tool reads the service provider data stored in the Sewall production geodatabase and performs an analysis through which the deliverables are extracted. The analysis process by which the tool extracts the road segments and census block data is outlined in the whitepaper entitled "Misalignment between 2000 Census Blocks & Maine E911 Streets: Technical Whitepaper," dated 1 September 2010. This paper is included in **Appendix D**.

Middle mile and community anchor institution data are stored as point features in the Sewall production model and are extracted utilizing a standard export routine. The datasets are reprojected from the production UTM projection to the SBDD WGS84 projection and LAT/LON attributes are populated. Once complete, the points are loaded into the destination feature classes of the SBDD geodatabase.

Wireless, service overview and state boundary data are stored as polygon features in the Sewall production model and a standard export routine extracts these to the SBDD features. The datasets are reprojected from the production UTM projection to the SBDD WGS84 projection as features are loaded.

During the export process features with front-end business rule violations get reported. The report is than reviewed by a Sewall analyst, and necessary corrections are made to the base datasets. This reporting mechanism ensures the data delivered in the SBDD geodatabase is as complete and accurate as the provided data sources allow.

Once the SBDD transfer file geodatabase has been created and its content validated, the geodatabase files are included in the data submittal zip file along with the other submittal files including 'datapackage.xls,' schema modifications report, data verification summaries, and this technical whitepaper.

Appendix A - Sample Letter to Service Providers

[date]

Sewall
P.O. Box 433
136 Center St.
Old Town, ME 04468
207-827-4456

[address]
[address]
[address]
[address]

Dear Mr. [name]:

The National Telecommunications and Information Administration (NTIA) of the U. S. Department of Commerce has been charged by Congress under the American Recovery and Reinvestment Act of 2009 and the Broadband Data Improvement Act (BDIA) to develop and maintain a comprehensive, interactive, and searchable nationwide inventory map of existing broadband service capability and availability in the United States that depicts the geographic extent to which broadband service is deployed and available from a commercial or public provider throughout each state (the Program).

The ConnectME Authority (the Authority) is responsible for developing and maintaining these data for the State of Maine and for serving as the conduit for this information to the NTIA. The Authority has contracted with James W. Sewall Company of Old Town, Maine, to undertake the initial mapping and to consult with the Authority on how best to update and maintain these data going forward.

We are writing to insure that you are familiar with this Program and to invite your collaboration in teaming with us in this important, statewide initiative. (See the URL's provided at the end of this letter for further information.) Indeed, your organization's collaboration is essential to the Program's success, and we thank you in advance for your participation.

To comply with the Program, the NTIA requires each state to provide structured data that includes:

- the availability of broadband service at the address level;
- advertised and "expected actual" speeds of broadband service;
- the technology used to deliver broadband service;
- location and capability of critical broadband related infrastructure (this data will not be publicly displayed on the national broadband map);
- the spectrum used by wireless broadband service providers.

We expect that the publicly searchable national broadband map and database will contain:

- geographic areas in which broadband service is available;
- the technologies used to provide broadband service in such areas;
- the speed at which broadband service is available in such areas;
- broadband service availability at public schools, libraries, hospitals, colleges, and all public buildings used by the state or municipalities.

- other economic or demographic data that may enable Federal efforts to provide usable and searchable data on a variety of issues pertinent to the public interest.

We recognize that some of the data we will ask you to provide is proprietary. Consequently, we include a Protective Order authorized by the ConnectME Authority and an accompanying non-disclosure agreement (NDA) for your review and execution. Please note, however, that the NTIA requires that this NDA may not restrict the Authority from providing all data collected to the NTIA or restrict the NTIA's use of such data as contemplated under this Program, including sharing such data with the FCC or other federal agencies. Furthermore, the NTIA prohibits the Authority or Sewall from agreeing to a more restrictive definition of Confidential Information than that adopted by the NTIA. Currently, as required under the BDIA, the NTIA identifies Confidential Information as any information, including trade secrets, or commercial or financial information, submitted under the Program that:

- identifies the location, type and technical specification of infrastructure owned, leased or used by a specific broadband service provider; or
- explicitly identifies a broadband service provider in relation to its specific service area or at a specific service location.

Confidential Information will not be made publicly available pursuant to the limits set forth in the BDIA except as required by applicable law or judicial or administrative action or proceeding, including Freedom of Information Act requirements. From the BDIA (§ 106(h)): "Notwithstanding any provision of Federal or State law to the contrary, an eligible entity shall treat any matter that is a trade secret, commercial or financial information, or privileged or confidential, as a record not subject to public disclosure except as otherwise mutually agreed to by the broadband service provider and the eligible entity." Sewall was chosen to lead this task in part because of its long history of handling confidential information for a variety of industries. Finally, should your organization apply for a Broadband Technology Opportunities Program (BTOP) grant to support the deployment of broadband infrastructure in unserved and underserved areas, enhance broadband capacity at public computer centers or to encourage sustainable adoption of broadband service, the NTIA requires that you participate in this mapping Program.

The NTIA has set a very aggressive Program schedule, with many deliverables due by November 2009 and all initial deliverables due in March 2010. Consequently, a representative from the Sewall team will be contacting you soon to discuss any questions you may have and to facilitate completion of the NDA and your participation. If we should be in communication with others in your organization concerning either the NDA or the data transfers, please inform the Sewall representative as soon as possible. Thank you again and we look forward to working with you.

Sincerely,

Phillip W. Lindley, Executive Director
ConnectME Authority

James H. Page, CEO
James W. Sewall Company

URLs for:
www.maine.gov/connectme
www.sewall.com
www.ntia.doc.gov/press/2009/BTOP_mappingtotals_090909.html

Appendix B - ConnectME Authority Protective Order

STATE OF MAINE — December 21, 2009

CONNECTME AUTHORITY — PROTECTIVE ORDER
(Proprietary Business Information)

Pursuant to 35-A M.R.S.A. § 9207(1) and Rule Chapter 101, § 4, the ConnectME Authority (Authority) may designate information as confidential to protect the legitimate competitive or proprietary interests of communications service providers and mobile communications service providers. The Authority may designate information as confidential only to the minimum extent necessary to protect such legitimate competitive or proprietary interests. Information designated as confidential is not a public record under 1 M.R.S.A. § 402(3).

The Authority is currently conducting a Broadband Mapping and Inventory Project with the services of a private contractor, James Sewall Company (Sewall). Sewall is required to obtain data from service providers (Provider) by the Authority and the National Telecommunications and Information Administration (NTIA) pursuant to the Broadband Data Improvement Act (BDIA) and the NTIA Notice of Funds Availability (NOFA). The NTIA requires that the Authority agree to comply with confidentiality requirements in section 106(h)(2) of the BDIA.

It is anticipated that providers submitting data to Sewall or the Authority may have a need to provide information considered to be confidential, in that the information provided may involve commercially sensitive and/or proprietary information regarding information that identifies (i) the location, type, and technical specifications of infrastructure owned, leased, or used by providers or (ii) explicitly identifies providers in relation to their specific service area or at a specific service location (collectively, the "Confidential Information"). The Authority has determined that such Confidential Information is generally not disclosed publicly, and that the public disclosure of such Confidential Information without restriction would cause competitive harm to the applicant or provider.

Accordingly, the following terms shall apply unless and until modified by the Authority or a court of competent jurisdiction:

1. Data submitted to Sewall or the Authority falling within the above definition of Confidential Information, as well as any data submitted to Sewall or the Authority pursuant to the Non-Disclosure Agreement set forth in Attachment A , (collectively, "Designated Confidential Information") shall be deemed to be competitively sensitive and/or proprietary in nature and such Designated Confidential Information shall be and remain exempt from public disclosure pursuant to the terms of this Protective Order and the articles referenced therein.

2. All Designated Confidential Information shall be and remain exempt from public disclosure pursuant to the terms of this Protective Order, unless removed from the coverage of this Protective Order as provided below or otherwise by a court of competent jurisdiction. No persons provided access to any Designated Confidential Information by reason of this Protective Order shall use such information for any purpose other than the purposes designated by the Authority. Every person provided access to Designated Confidential Information shall use his or her best efforts to keep the Designated Confidential Information secure and shall not publicly disclose it or accord public access to it to any person not authorized by the terms of this Protective Order.

3. Any person or the Authority may challenge the designation of any document or other information as Designated Confidential Information. The Authority will provide reasonable prior notice to the applicant or provider and an opportunity for hearing prior to ruling on any such challenge. In considering any such challenge, the usual burdens of proof and production shall apply and no additional presumption shall be given as a result of the prior acceptance by the Authority of material as Designated Confidential Information. In the event the Authority should rule over the objections of the person providing the Designated Confidential Information that any information should no longer

be subject to the terms of this Protective Order, such information shall not be publicly disclosed until the later of five (5) business days after the Authority so orders or, if the person files within such five day period an appeal or request for stay of such order, the date upon which such appeal or request for stay is decided; provided, however, that said periods may be extended in accordance with any stay ordered by the Authority or a reviewing court. Upon the entry of a final unappealed decision by the Authority or a reviewing court granting public disclosure, the terms of this Protective Order shall cease to bind any person with respect to the information that the order granting disclosure shall have expressly and clearly removed from the coverage of this Protective Order.

4. Any person provided access to Designated Confidential Information shall review and be bound by the terms of this Protective Order. Prior to obtaining access to any Designated Confidential Information, such person shall sign an acknowledgment of his or her obligation to abide by the terms of this Protective Order in the Non-Disclosure Agreement (NDA) attached hereto as Attachment A.

5. Unless modified by the Authority or a court of competent jurisdiction, access to Designated Confidential Information shall be limited to Authority Staff, Sewall, any independent consultants or experts retained by the Authority, the National Telecommunications and Information Administration, and those designated persons, who have signed the NDA.

6. No copies of Designated Confidential Information shall be circulated to persons other than those authorized under paragraph 5 of this Protective Order. Persons authorized under paragraph 5 hereof also may take such notes as may be necessary. Such notes shall be treated as Designated Confidential Information.

7. The restrictions upon, and obligations accruing to, persons who become subject to the terms of this Protective Order shall not apply to any Designated Confidential Information submitted in accordance with this Protective Order if the Authority rules, after reasonable notice to the applicant or provider and an opportunity for hearing, that such Designated Confidential Information was publicly known at the time it was furnished or has since become publicly known.

8. Where reference to Designated Confidential Information is required in any Authority document, such reference shall be by citation of title or attachment number only or by some other non-confidential description to the extent possible.

9. Designated Confidential Information furnished to the Authority pursuant to this Protective Order shall remain in the possession of the Authority, under seal, and subject to the terms of this Protective Order, until the Authority or a court of competent jurisdiction shall otherwise order.

10. The terms of this Protective Order may be modified on motion of any person or on the Authority's own motion upon reasonable prior notice to the applicant or provider and an opportunity for hearing.

BY ORDER OF THE CONNECTME AUTHORITY

Phillip Lindley, Executive Director

ATTACHMENT A [Non-Disclosure Agreement]

Appendix C - Template for Non-Disclosure Agreement

NON-DISCLOSURE AGREEMENT

THIS AGREEMENT is made this ________ day of _________________, 20___, by and between ____, a ________________________________ having a principal place of business at ____ ("PROVIDER") and ConnectME Authority, a body corporate and politic and a public instrumentality of the State of Maine established pursuant to 35-A M.R.S.A. § 9203 (the "AUTHORITY") and James W. Sewall Company, a corporation organized under the laws of the State of Maine and having a principal place of business at 136 Center Street, Old Town, Maine 04419 ("SEWALL") (AUTHORITY and SEWALL individually or collectively referred to as "RECIPIENTS") (PROVIDER AND RECIPIENTS collectively referred to as the "Parties").

Recitals

WHEREAS, the National Telecommunications and Information Administration (the "NTIA") of the United States Department of Commerce has been charged by Congress under the America Recovery and Reinvestment Act of 2009 (the "ARRA") and the Broadband Data Improvement Act (the "BDIA") to develop and maintain a comprehensive, interactive, and searchable nationwide inventory map of existing broadband service capability and availability in the United States that depicts the geographic extent to which broadband service is deployed and available from a commercial or public provider throughout each state (the "Data"); and

WHEREAS, the AUTHORITY is responsible for developing and maintaining the Data for the State of Maine and for serving as a conduit for the Data to the NTIA; and

WHEREAS, SEWALL is contracted by the AUTHORITY to undertake the initial mapping and to consult with the AUTHORITY on how best to update and maintain the Data going forward; and

WHEREAS, the PROVIDER has trade secrets and commercial or financial information relating to the location, type, and technical specifications of infrastructure owned, leased, or used by PROVIDER, which is included in the Data (the "PROVIDER Information"); and

WHEREAS, the PROVIDER has agreed to provide PROVIDER Information to SEWALL and/or the AUTHORITY pursuant to the requirements of the ARRA and the BDIA for use by the NTIA.

NOW THEREFORE, for and in consideration of the mutual promises and covenants contained herein, and for other good and valuable consideration, the receipt and sufficiency of which is hereby acknowledged, the Parties agree as follows:

1. As requested in writing by PROVIDER, RECIPIENTS agree to hold in absolute and strict confidence and shall not disclose or reveal in any manner or form to any entity other than the NTIA any PROVIDER Information identified as confidential that identifies (i) the location, type, and technical specifications of infrastructure owned, leased, or used by PROVIDER or (ii) explicitly identifies PROVIDER in relation to its specific service area or at a specific service location (collectively, the "Confidential Information"), whether such disclosure was made orally, in writing, or in any other form, without prior written permission from PROVIDER.

Notwithstanding the foregoing, Confidential Information shall not include the following:

(a) information that now is or hereinafter becomes publicly known or available otherwise than through unauthorized disclosure by RECIPIENTS;
(b) information that was in RECIPIENTS' possession at the time of disclosure and was not acquired, directly or indirectly, from PROVIDER;
(c) information that RECIPIENTS received in good faith from a third party who is not under a similar restriction of confidentiality and having a right to disclose the Confidential Information; or

(d) information that is required to be disclosed pursuant to applicable law or judicial or administrative action or proceeding, including the Freedom of Information Act requirements.

2. RECIPIENTS agree not to use for any purpose the Confidential Information except as provided for under the ARRA and the BDIA, without prior written permission from PROVIDER.

3. This Agreement shall be governed by the laws of the State of Maine and applicable federal law, except for the State of Maine's conflict-of-laws provisions, as applicable. The Parties to this Agreement each specifically consent to jurisdiction in Maine in connection with any dispute between the Parties arising out of this Agreement or pertaining to the subject matter hereof, with venue being in a court of competent jurisdiction located in Penobscot or Kennebec County, Maine, United States of America.

4. This Agreement shall inure to the benefit of and be binding on the Parties and their respective successors and assigns.

5. This Agreement constitutes the complete and exclusive agreement of the Parties hereto with respect to the matters set forth herein. The terms of this Agreement may not be modified or amended except by an instrument in writing signed by each of the Parties hereto.

6. This Agreement shall be construed without regard to any presumption or other rule requiring construction against the drafting Party.

7. This Agreement may be executed in counterparts and each Party hereto may execute each such counterpart, each of which when executed and delivered shall be deemed to be an original and both of which counterparts taken together shall constitute but one and the same instrument. This Agreement shall become binding when all counterparts taken together shall have been executed and delivered by all Parties. Execution and delivery of this Agreement may be made by facsimile transmission, and each Party agrees that the delivery of the Agreement by facsimile shall have the same force and effect as delivery of original signatures and that each Party may use such facsimile signatures as evidence of the execution and delivery of the Agreement by all Parties to the same extent that an original signature could be used.

IN WITNESS WHEREOF, the Parties have executed this Agreement the day and year first above written.

WITNESSED BY:	PROVIDER
______________________________	By: ______________________________
	Title: ______________________________
	ConnectME Authority
______________________________	By: ______________________________
	Title: ______________________________
	James W. Sewall Company
______________________________	By: ______________________________
	Title: ______________________________

Appendix D - White Paper: *Maine-SBDD Census Block-Street Segment Misalignment*

Misalignment between 2000 Census Blocks & Maine E911 Streets

Technical Whitepaper

1 September 2010

Introduction

Importing broadband service provider data into the State Broadband Data Development (SBDD) Map Data Transfer Model at the census block versus street segment level has created challenges for the grantees. For the State of Maine one of the challenges involves the spatial misalignment between the 2000 Census Block polygon geometries and Maine's street centerline dataset.

In order to better understand the challenge that Maine is encountering it is necessary to review how the State is collecting and maintaining broadband service provider data.

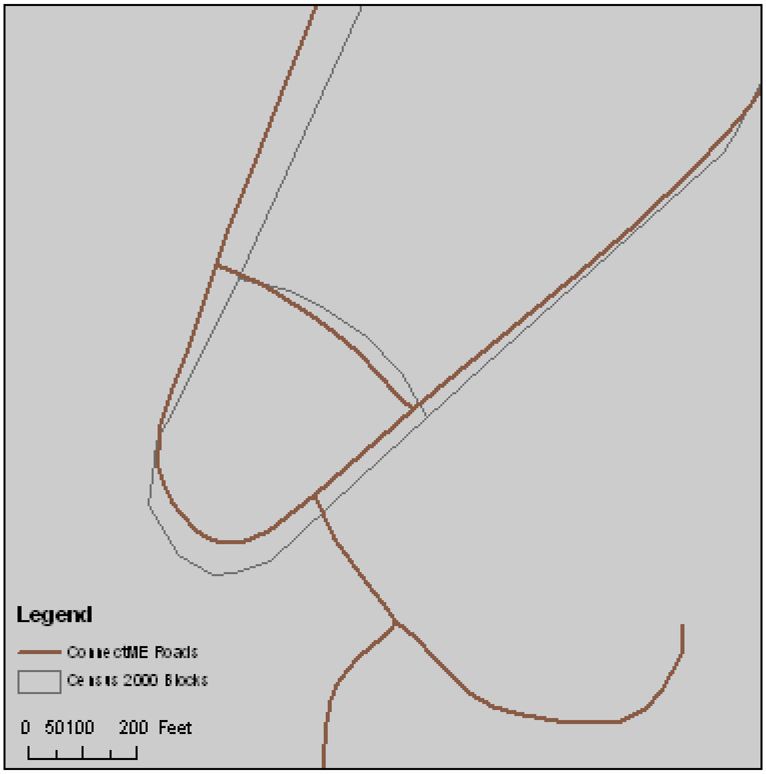
As a result of Maine's geographic population distribution, mapping broadband service at a census block level does not satisfy the State's requirements for statewide broadband tracking and development. Instead of utilizing the hybrid census block-street centerline model outlined in the SBDD NOFA, the State is collecting service provider coverages at a street level for wired and fixed wired technologies. The State has developed a relational model to best represent the one-to-many relationship between a street segment and its broadband service provider coverages.

The street segment data that the State is utilizing is based primarily on the State's E911 street centerline GIS layer with additional street coverage added from a 3rd party dataset for those towns not yet participating in the E911 project. For information on the broadband service providers, a database table was developed based on the required attribution descriptions outlined in the NOFA.

Now that the existing data structure has been presented the challenge of importing this data into the transfer model can be discussed along with the State's proposed solution for addressing misalignment and helping minimize its impact on the broadband data processing.

The Challenge

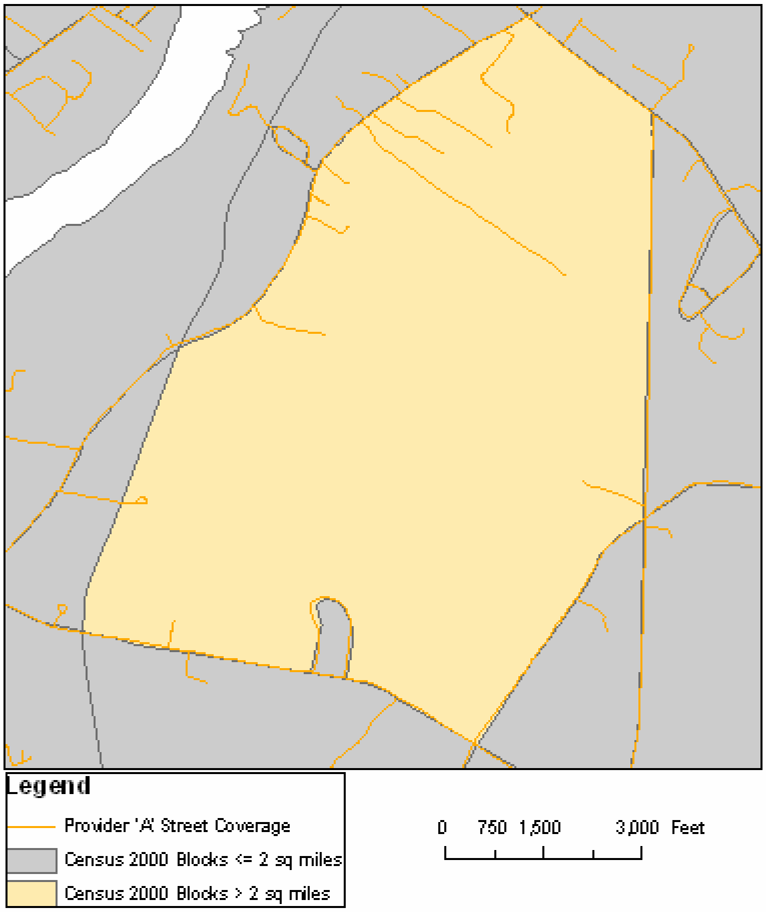
Year 2000 Census Block geometry is spatially misaligned with the Maine's street centerlines.



As shown in the above screen capture the typical misalignment between these two datasets is between 50 and 100 feet.

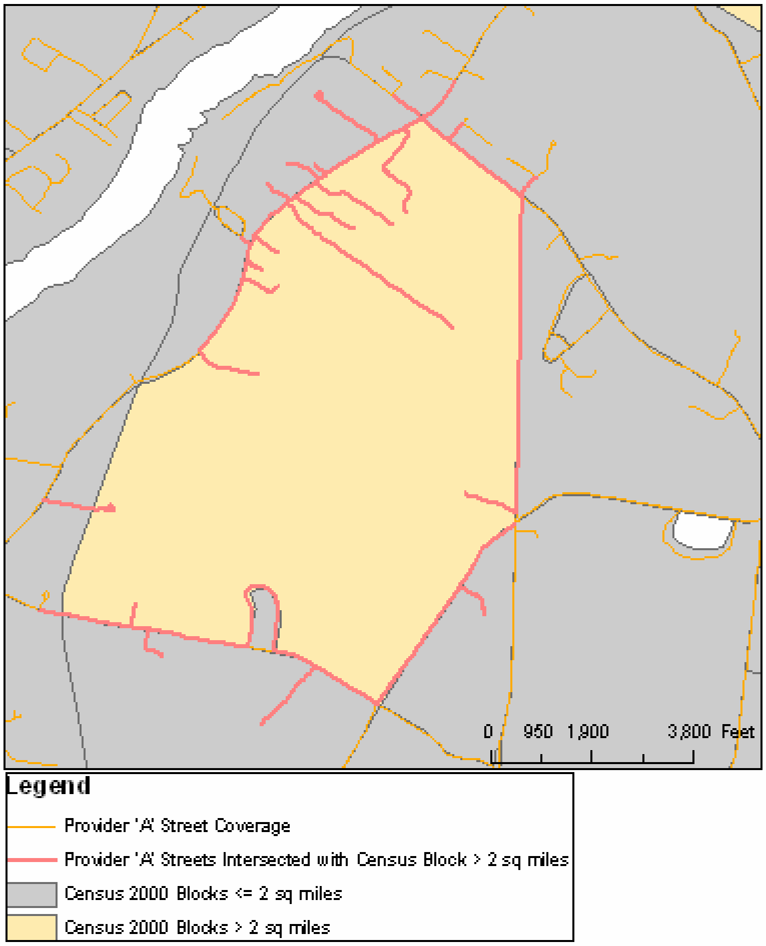
Since Maine is storing all broadband service providers' information as records associated with street centerlines this misalignment causes considerable challenges when trying to accurately export this information into the new SBDD data transfer model. The misalignment is great enough that utilizing basic intersect methodology is not enough to provide NTIA with a highly accurate representation of broadband coverage in Maine.

Example: Basic Intersect



The above screen capture shows an example of a 2000 Census Block that is greater than 2 square miles and Provider ‘A’ street coverage data that is to be reported.

Performing an intersect between the greater than 2 square mile census block and the street network for Provider 'A' results in the highlighted streets being reported.

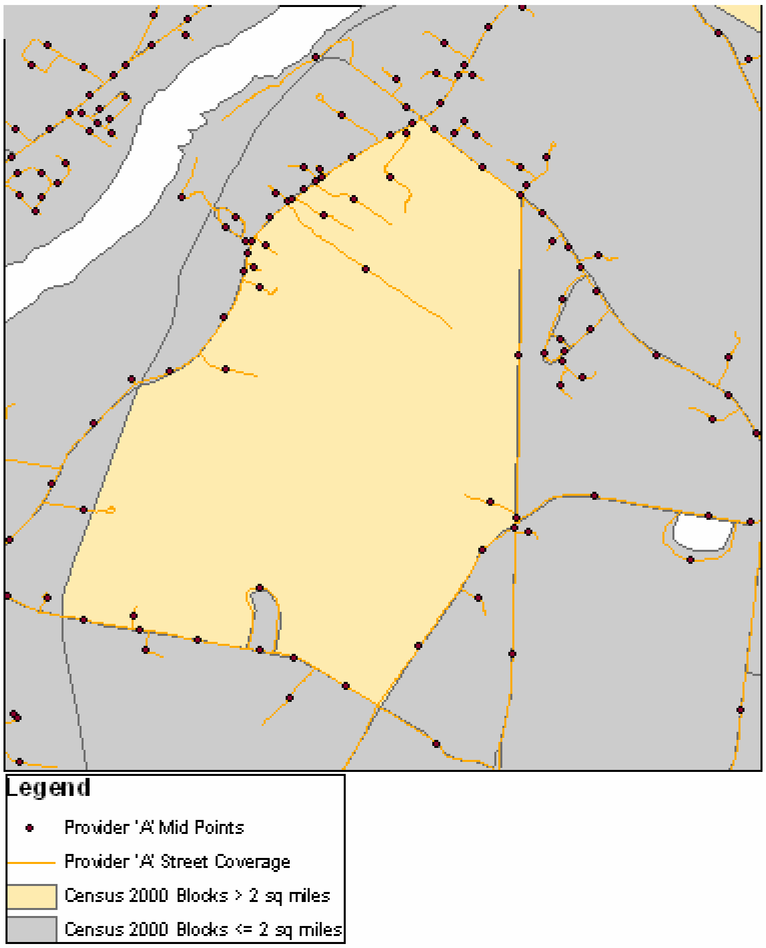


It is clear from the screen capture that several extra streets where selected and a few streets were missed by using the intersection method.

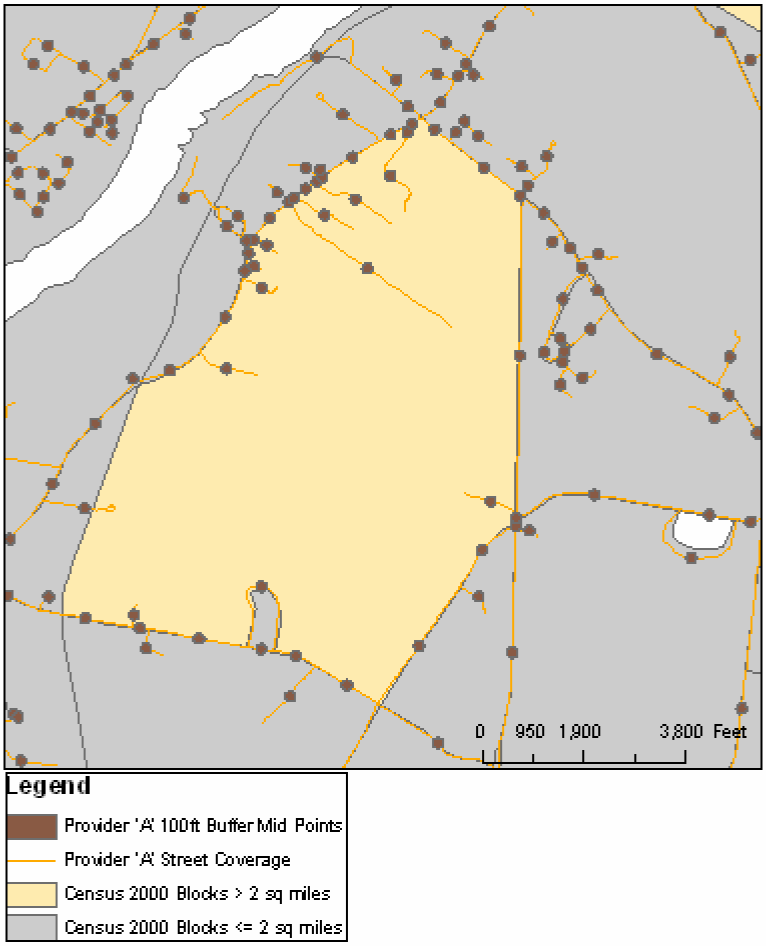
Proposed Technical Solution

The solution to this challenge is a multi-step process that needs to be run on each street segment with intelligent analysis employed to minimize errant representation of broadband service in census blocks greater that 2 square miles.

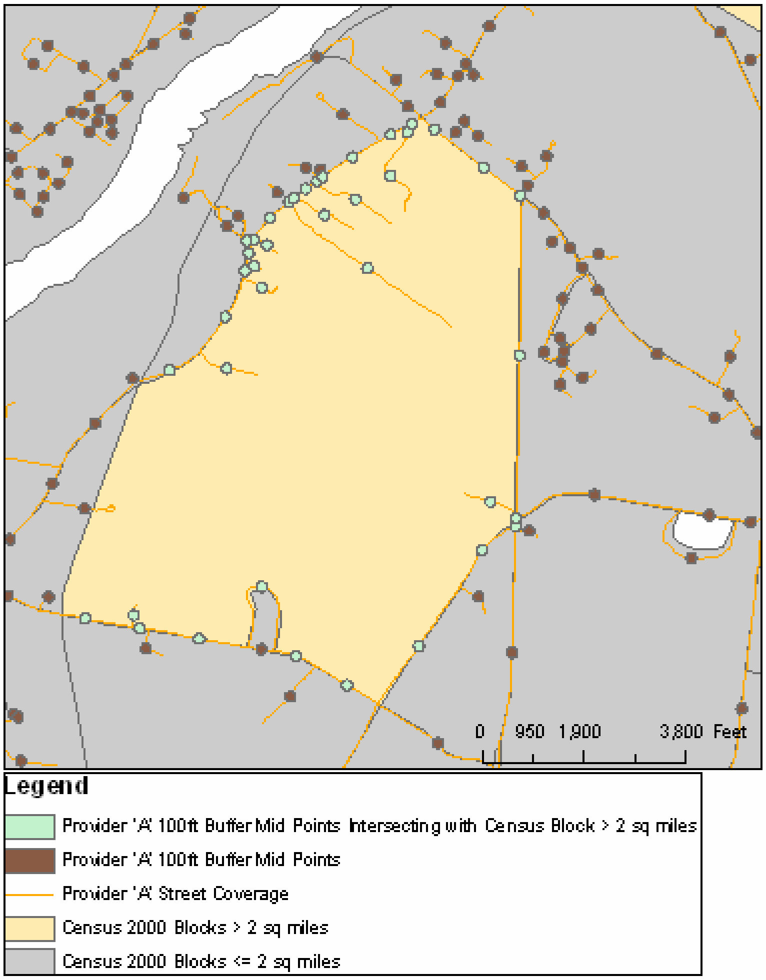
The first step is to create mid points of the street centerlines for Provider ‘A’.



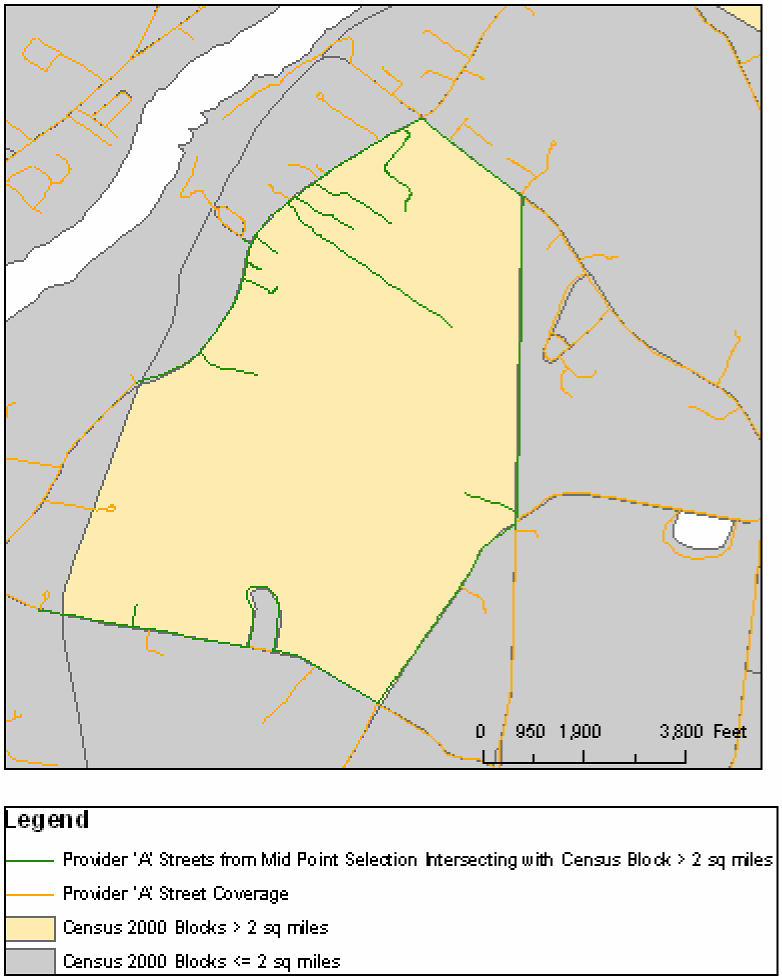
The next step is to create a buffer around the mid points using a distance to compensate for the misalignment in the census blocks. The distance found to have the best return for this process was determined to be 100 feet.



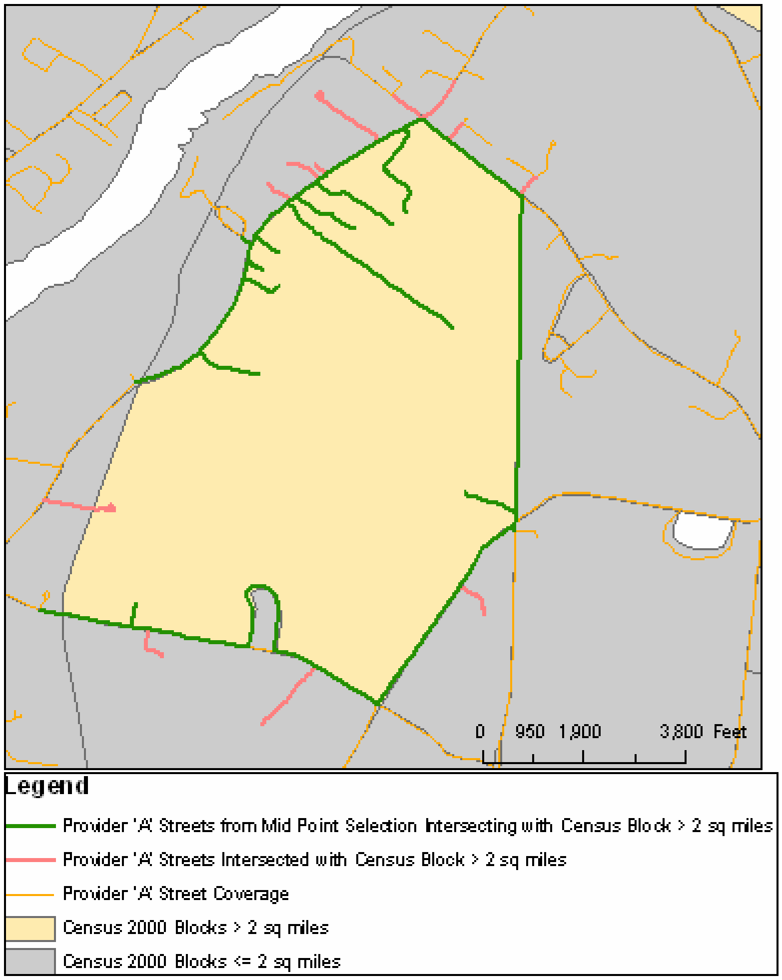
Selecting the buffered mid points that intersect the greater than 2 square miles census block returns the following results:



The selected buffered mid points relate back to the following street selection:

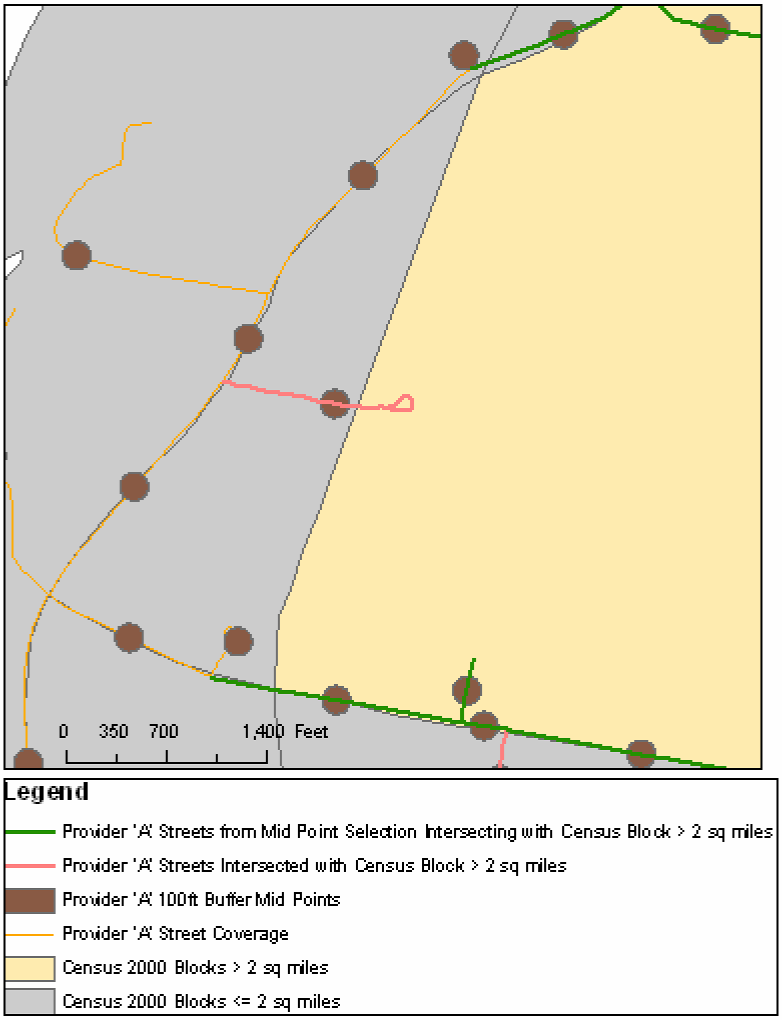


Compare this selection to the original intersection process selection:



The result of the mid point buffering process is a much better representation of streets contained within the greater than 2 square miles census block. A large number of the erroneous streets initially marked as included in the census block have been dropped providing a much improved report.

Taking a look at the left hand side of the map there is a street that intersects the census block but is not reported in the mid point buffering process. A closer look reveals why.

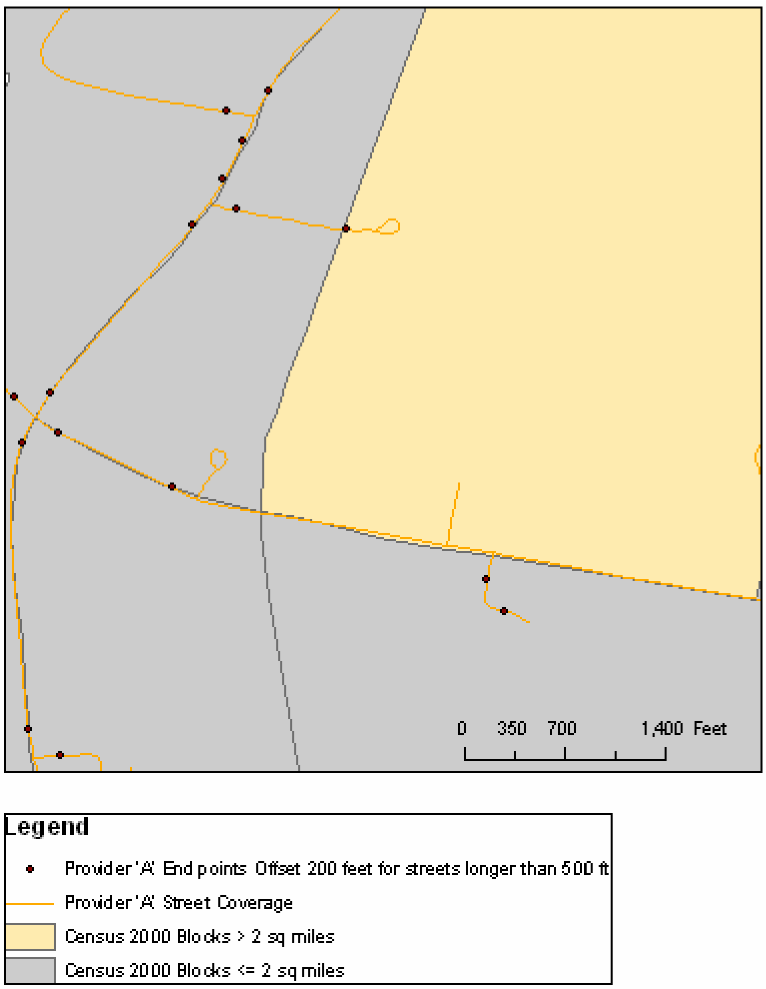


The street in question is relatively long in length and has a midpoint that is located outside of the greater than 2 square miles census block resulting in it not being reported.

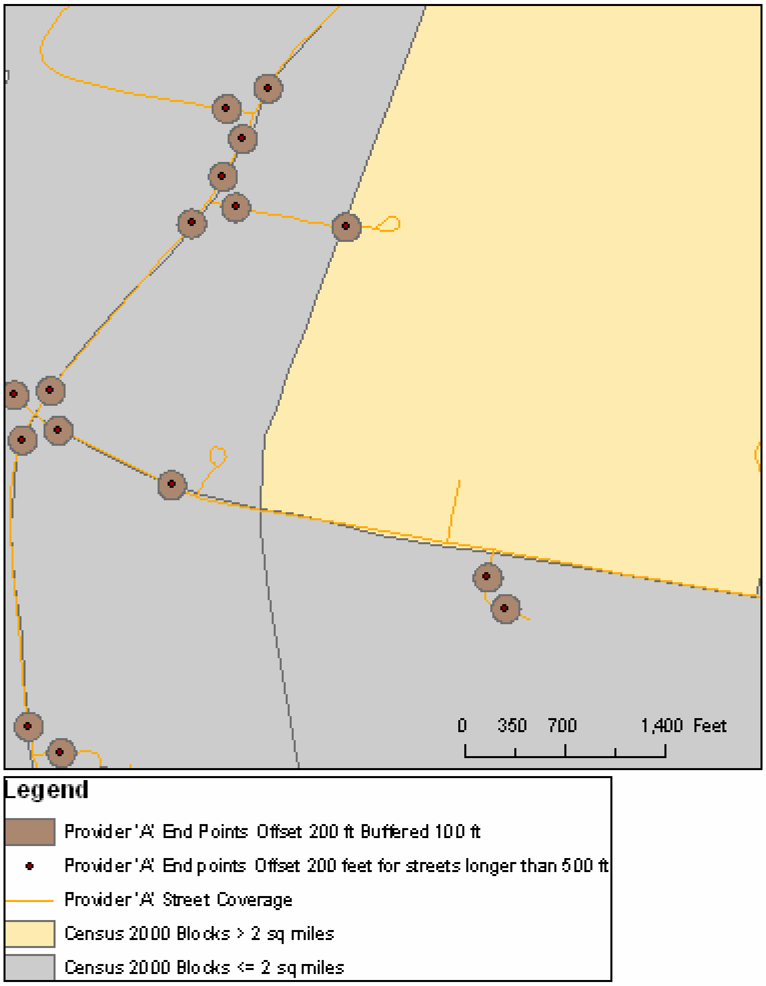
Building onto what has been performed already an additional automation check can locate and incorporate these long streets into the dataset.

The Proposed Solution: Additional Intelligence

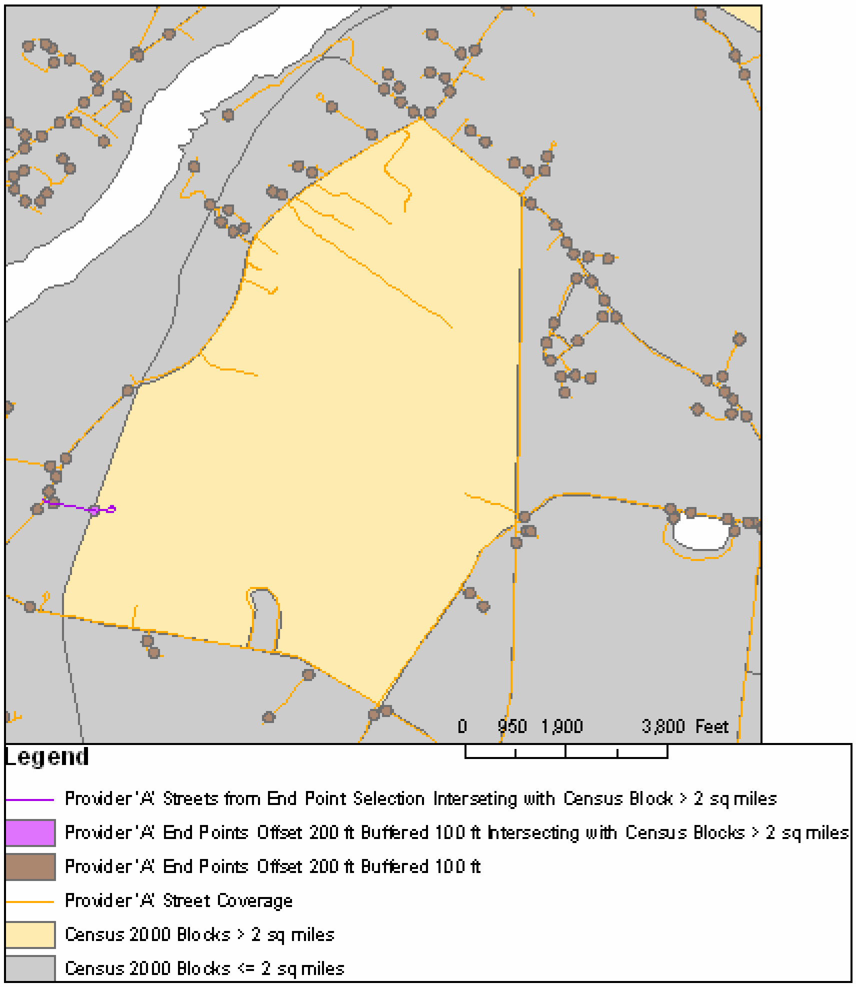
The first step in this additional iteration is to select streets that have not been flagged as being contained within a census block greater than 2 square miles and are longer than 500 feet. Then create points that are offset 200 feet from each end of the selected streets.



Next these 200 feet offset points are buffered 100 feet:

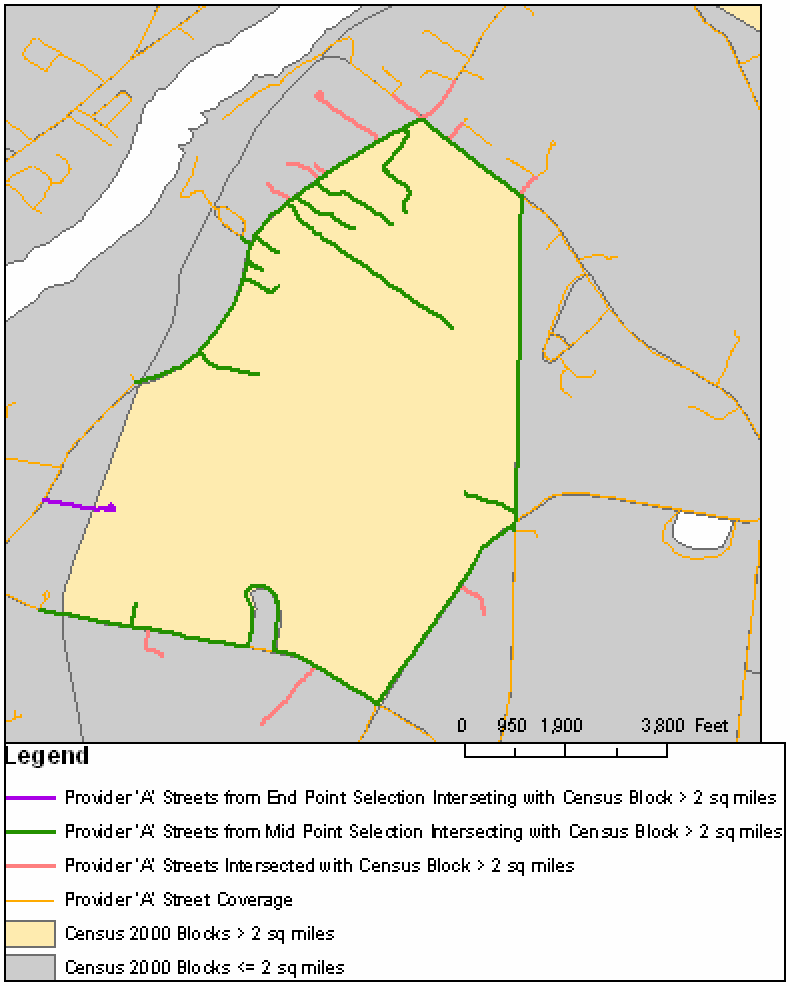


Then by selecting the buffers that intersect the greater than 2 square miles census block and selecting the associated streets, the process results in the following:



The Results

The screen capture below shows the streets reported using the two step process in comparison to the basic intersect method of reporting street segments.



The following table shows the results of the processes for Provider 'A' for this particular census block:

Method	Missed Streets	Extra Streets	%Error
Basic Intersect Process	2	11	35.14
MID Point Process	1	2	8.11
MID and END Point Process	0	2	5.41

The proposed solution gives a much better representation of the data set and minimizes the errors induced by using a basic intersection process.

Summary

The SBDD data submission requirements involving census blocks and street segments have created a challenge for the grantees to accurately represent broadband service provider information. In particular the State of Maine has a significant offset between the 2000 Census Block geometries and the corresponding street centerlines that the State is utilizing to map broadband availability data. A basic spatial intersect method has proven to be highly inaccurate in identifying street centerline data in census blocks greater than 2 square miles.

Through it analysis, the State has found that using a two step process using mid-point and offset end point buffering provides improved results for street centerlines in the greater than 2 square mile census blocks. The State expects this methodology to improve the accuracy of street segment determination by approximately 50% for these regions. Unless instructed otherwise by the NTIA project team, the State intends to utilize this two step process to develop the SBDD deliverables for street centerlines in census blocks greater than 2 square miles.

In addition, it is the hope of Maine's project team that the spatial misalignment between census block geometries and the E911 street centerlines will be remedied by the 2010 Census data as a result of the work done under the Boundary and Annexation Survey (BAS). Maine is one of the states that participated in the 2010 BAS with the Census Bureau to better align the census boundary data with the State's E911 street data.