

Round 2 Methodology

Primary Data Collection

In Round 2, as in Round 1, the CPUC sent out a Data Request very widely to potential broadband providers in the state and encouraged broadband service providers to submit broadband availability data to the CPUC. For Round 2 We expressed our preference for providers to use a file geodatabase format when possible. Tabular data was also accepted and template files, record formats, and data submittal instructions were posted on the CPUC Broadband Mapping Website at <http://www.cpuc.ca.gov/PUC/Telco/Information+for+providing+service/BB+Mapping.htm> to assist providers, along with the text of our Data Request, the NOFA, and other relevant documents.

Having had to send our Data Request prior to NTIA's determination to require 2000 Census geography, we requested that 2009 Census geography be used. 2009 results were subsequently converted to 2000 geography prior to our submission to the NTIA.

Following the Data Request, a set of maps showing the individual provider's Round 1 data were created and sent to each provider for verification and feedback. This process allowed the providers to visually examine the output of their previously submitted data, give the CPUC feedback on any errors that were exposed, and prepare improvements for Round 2.

Community Anchor Institutions

Our CAI data was updated to reflect new California Teleconnect Fund (CTF) eligible entries and, geocoded to point locations by the CPUC and loaded into the file geodatabase. The CTF provides 50% discounts on telecommunications bills for qualifying schools, libraries, government-owned and operated hospitals and health clinics, and other community based organizations, thus providing a good initial list of CAIs. Broadband connection technology and speeds were included in our CAI results where known, either through information received from the Institutions themselves (as in the case of libraries), from those service providers who responded to our request for such information.

Provider Participation

The total number of providers in the Round 2 data collection was 73. Of those, 52 submitted fresh data sets while 21 either did not respond or indicated they had no changes to report since Round 1. These providers are responsible for a high percentage of the total broadband connections in California reported to the FCC on form 477, and thus constitute a very complete set of data.

CPUC Initial Data Verification

Submitted files were reviewed by an analyst and tracked using a data inventory spreadsheet. Each file was reviewed against the matching record format to see if mandatory fields were filled in, and each field was checked for the appropriate range of values. Where possible, analysts loaded the submitted data into the corresponding geodatabase table to make certain that appropriate field headers were used and that each field contained the correct data type. When data was found to be missing or incorrect, the provider was contacted and the issue was documented in the spreadsheet. Some providers submitted nearly perfect data sets while others gave incomplete, unexpected, or incorrect data. New information, correspondence with the providers, and fixes made by the CPUC were also documented in the spreadsheet.

Chico GIC Geoprocessing

After the initial CPUC review, data was transferred to the Geographical Information Center at CSU Chico for geocoding, geomatching, propagation of wireless service by antenna, and validation of geographic data. In those cases where the CPUC had received street address level data from broadband providers, such addresses were assigned a point location, (geocoded) and then geomatched to census blocks and street segments.

Providers who offer wireless service but could not submit a shapefile or geographic representation of their service area gave tabular antenna information. Wireless antenna parameters were used to model a service area and shapefiles were created for each provider. The wireless propagation model is based on the Longley-Rice, Irregular Terrain propagation model. Individual unit specifications are used to measure performance based on frequency, transmit power, receiver sensitivity, antenna gain, and height. Signal coverage patterns are produced for each individual unit taking into account terrain and vegetation features that may hinder signal dispersion.

Chico GIC Data Validation Methods

The following validation methods were conducted for Round 2 data submission:

FCC Form 477

FCC Form 477 collects information about broadband connections to end user locations, wired and wireless local telephone services, and interconnected Voice over Internet Protocol (VoIP) services, in individual states, at the Census Tract level. A shape file was created for each provider reflecting the presumed availability of broadband service at each census tract where the provider reported customers to their fixed broadband service. These layers were used to cross reference ISP data submissions to the CPUC.

FCC Cable Plant Flyover

FCC Form 320 contains a snapshot of the interference potential to aircraft communications and navigation over a cable system's distribution plant. Flyovers are restricted to a geographic boundary in which a cable plant is capable of deploying service. Boundaries were obtained and digitized from maps

made available through the FCC, to create GIS layers specific to each provider, to cross reference ISP data submissions to the CPUC.

TeleAtlas Wire Center

The Wire Center Premium product is a comprehensive database for mapping and analyzing wire center service areas. It forms the backbone of the Tele Atlas® Telecommunication Products line. This product lists every Local Exchange Carrier (LEC) landline wire center in the United States. The term “wire center” refers to the location where the telephone company terminates the local lines; this is usually the same location as a central office, although a wire center might house one or more central offices. Buffers were created at 12,000 feet and 18,000 feet from provided Wire Center point datasets to cross reference ISP data submissions to the CPUC.

TeleAtlas Wire Center Region

The Wire Center Premium product is a comprehensive database for mapping and analyzing wire center service areas. It forms the backbone of the Tele Atlas® Telecommunication Products line. This product lists every Local Exchange Carrier (LEC) landline wire center in the United States. The wire center boundary is a representation of the area served by all of the switching equipment housed at that physical location. Wire Center Region polygon GIS layers were provided and used for cross referencing ISP data submissions to the CPUC.

Chico GIC Data Validation Processes

Each individual provider’s submitted and/or created data was validated independently to all applicable validation methods. The following fields were added to each individual provider’s data tables, where appropriate; FCC_477 (FCC Form 477), FLYOVER (FCC Cable Plant Boundary), TA_WC_REG (TeleAtlas Wire Center Region), WC_VAL_12K (TeleAtlas Wire Center 12,000 foot buffer), WC_VAL_18K (TeleAtlas Wire Center 18,000 foot buffer), VAL12k_18k (TeleAtlas Wire Center 12,000 to 18,000 foot buffer ring), DEGRAD_FT (TeleAtlas Wire Center distance), to record validation results and to allow symbology of discrepancies based on validation methods for further interaction with each provider to refine their data submissions. The final step was a summary statistics report of all validation results for all submitted providers. Summary statistics include validity counts and percentages for all validation methods, specific to provider and technology.

Wireline Census Block Validation

A spatial selection was performed on Census Block data, either submitted by provider, or created from submitted address records through a geocoding/spatial selection process, to derive only those blocks which intersect polygons in a given validation layer. Counts are recorded as number of unique blocks which share geographic area with any given validation layer, compared to the total number of unique blocks submitted by, or created for, a given provider. Percentages are recorded as percentage of the total number of unique blocks which share geographic area with any given validation layer, compared to the total number of unique blocks submitted by, or created for, a given provider.

Wireline Street Segment Validation

A spatial selection was performed on Street Segment data, either submitted by provider, or created from submitted address records through a geocoding/spatial selection process, to derive only those segments which intersect polygons in a given validation layer. Counts are recorded as number of unique segments which share geographic area with any given validation layer, compared to the total number of unique segments submitted by, or created for, a given provider. Percentages are recorded as percentage of the total number of unique blocks which share geographic area with any given validation layer, compared to the total number of unique segments submitted by, or created for, a given provider.

Wireless Validation

A spatial selection was performed on Wireless Availability data, either submitted by provider, or created from antenna location and specification information, to select only those polygons which intersect a given validation layer. Results are recorded as a percentage of the total geographic area of wireless coverage sharing geographic area with any given validation layer, compared to the total coverage area submitted by, or created for, a given provider.

CPUC Final Data Verification

The resulting geodatabase was returned from Chico to the CPUC for final review and verification. Data sets were checked again by CPUC analysts and reviewed for unexpected changes resulting from the geocoding /geomatching process. We visually reviewed the data using ArcGIS to verify the validated data from Chico. Another feedback loop was initiated when a set of Round 2 maps were created and sent to providers for review. Final data verification edits were made when the CPUC received corrections or new information from providers.

Deliverable Data

The final dataset was delivered to the NTIA/FCC in a file geodatabase with the following feature classes:

BB_ConnectionPoint_LastMile – not required per Clarification to the NOFA.

BB_ConnectionPoint_MiddleMile – Point between the local “last mile” network and the middle mile network which goes on to connect to the internet backbone. This is a confidential dataset.

BB_Service_Address – not included per the CPUC NDA.

BB_Service_CAInstitutions – Community Anchor Institutions: from CTF applicant list.

BB_Service_CensusBlock – Broadband availability polygons for areas less than 2 square miles

BB_Service_Overview – Service overview by County including Subscriber Weighted Nominal Speed

BB_Service_RoadSegment – Broadband availability line segments for areas 2 square miles and greater

BB_Service_Wireless – Wireless service area polygons.