Overview

Broadband availability is depicted on the California Interactive Broadband Map at various levels of granularity, depending on the granularity of the data provided to the CPUC. This creates a challenge when depicting broadband availability at a detailed level, such as for a specific address. It is not uncommon for the map to show all of the households in a census block as "served" at the threshold of 6 megabits per second down or greater and 1.5 megabits per second up or greater even though some households in that block remain under- or unserved (false positive). At the same time, there are instances of opposite problem as well where the map shows a census block being 100% unserved when there are indeed some households that are served (false negative).

Given the Legislative and Commission directives to implement the California Advanced Services Fund (CASF) program, the validation procedures currently in place tend to favor false negatives over false positives. In other words, the Interactive Map may understate availability of broadband rather than overstate it. Nevertheless, the California Interactive Broadband Map is not the final arbiter of CASF eligibility. Indeed, the CASF application includes a challenge process, which allows providers to identify portions of a census block that are in fact served.

The following validation methods are conducted on each provider's data submission. The 2015 update of the Interactive Map, which is based on broadband availability data as of December 31, 2014, includes "red zone" and "purple zone" validation layers for each provider. The red zone layer shows areas for which either no validation method exists to verify the existence of a broadband provider's service, or public feedback contradicted the provider's claim of service. The purple zone layer shows areas for which no validation method could verify a broadband provider's advertised downstream or upstream throughput. This does not mean there is no service, or service is definitely not available at the speeds submitted, just that we haven't been able to confirm the presence of service with the data sources available to us.

The table below summarizes the validation method, type of data, and to which type of broadband connection the validation method applies.

Data Source	Data Type	Fixed: Wireline	Fixed: Wireless	Mobile Wireless
FCC Form 477 Subscriber Data	Number of subscribers by upstream and downstream speeds by census tract **used to validate availability and speed at census tract**	YES	YES	NO
TeleAtlas Wire Center	Serving wire center locations of telephone companies **used to**	YES	NO	NO

	validate DSL coverage			
CPUC Mobile Field Test Upstream and Downstream Interpolation	Interpolated coverage based on mean minus 2 standard deviation used to validate availability at census block	NO	NO	YES
CPUC Mobile Field Test Results Point Data	Provider-specific, "In coverage" location results showing "No Effective Service" (point data)	NO	NO	YES
CalSPEED results	Speed test results from LTE-capable devices and "No Effective Service" results from ANY device	NO	NO	YES
Customer address service and speed information	Provider-supplied list of customers showing their address and subscribed speeds – used to validate availability and speed at census tract	YES	YES	NO
Public Survey	Reports of "no service" for a specific provider used to validate availability at census block ("no service" = block becomes unavailable for that provider)	YES	YES	YES
Tower data and/or EDX propagation image	Coverage propagation of fixed wireless provider based on tower, radio, and antenna data submitted by the provider used as a baseline for availability footprint	NO	YES (footprint only)	NO

Details on the Validation Data

- **FCC Form 477 Subscriber Data -** For fixed services, the FCC collects data from each broadband provider twice a year. This includes the number of broadband connections by technology type and speed tier (upstream and downstream) for each census tract where the provider has customers. If a provider indicates it has broadband service in a particular census block but has not reported customers to the FCC for the census tract where that block resides, the Form 477 data cannot validate the actual presence of service. In the case of speed validation, if a provider

has not reported any subscribers in any of the census blocks nested within the applicable census tract, then Form 477 cannot validate the speed for the entire census tract. As with any validation technique, there are inherent errors. For example, if Form 477 data show that a particular provider has customers in a census tract and at the maximum advertised speeds submitted to us, we consider all blocks within that census tract to be validated for speed and/or availability for that provider. Because Form 477 data is only available at the census tract level, this validation tool tends to yield false positives and overstate validation for individual census blocks. In contrast to fixed broadband service data, mobile broadband service data are aggregated at the state level and are not useful for census tract level validation.

- **TeleAtlas Wire Center data** 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 its local lines; this is usually the same location as a central office, although a wire center might house multiple central offices. Buffers were created at 12,000 feet from provided Wire Center point datasets to cross reference ISP data submissions to the CPUC. The wire center boundary is a representation of the area served by all of the switching equipment housed at that physical location. When a provider indicates broadband availability in a particular census block, and that location is within the distance from the wire center to support a given speed, that census block is considered validated. If the census block is beyond 12,000 feet from the central office, the speed cannot be validated. This methodology is used for DSL technologies only.

- **CPUC Mobile Field Test Upstream and Downstream Interpolation** uses data generated by the semi-annual mobile field tests, which cover 1,990 randomly selected points across the state and measure broadband performance for the four major mobile wireless operators: Verizon, AT&T, Sprint, and T-Mobile USA. The mean minus two standard deviation results are interpolated to create a kriging map. This map is used to estimate availability, upstream, and downstream throughputs. We compare the interpolated model against each provider's stated coverage and speed. In cases where the estimate is below the provider's stated coverage, we create a purple zone for the census block(s) that fall under all or part of that area. In cases where the estimate shows no coverage but the provider's map does show coverage, we create a red zone.

- **CPUC Mobile Field Test Point Data** come from our semi-annual field tests. The mean minus two standard deviation point data from the Fall 2014 tests were compared against each operator's advertised availability in the census block where the test was conducted. In census blocks where the test result for a particular operator was zero or "No Effective Service," but the operator advertised coverage there, the coverage for that census block was considered un-validated.

- **CalSPEED Results** are crowd-sourced mobile test results from the CPUC's Android mobile testing application. The CPUC launched CalSPEED on Google Play's app store on April 5, 2013. The point data results through May 2014 were compared against each operator's advertised availability in the census block where the test was conducted. These results included operators beyond the four tested for the bi-annual mobile field testing. In census blocks where the test result for a particular operator was zero or "No Effective Service," but the operator advertised coverage there, the coverage for that census block was considered un-validated.

- **Customer Address Service and Speed Information**
  Where we were unable to validate any areas of a provider's availability (their entire footprint was a red zone), we requested customer address information to use as a validation data source. Census tracts where customers resided were considered validated.
- **Public Survey -** As part of our effort to collect and incorporate information from the public, we created an online as well as downloadable paper survey that member of the public fill out to tell us which providers they have and at what speeds they subscribe. The survey also captures whether they have been denied service or do not have access to specific providers claiming to offer service to their area. There is also a section where they can tell us the results from speed tests. The survey and FAQ are available on the CPUC web site at: http://www.cpuc.ca.gov/General.aspx?id=5868 . Results through May 30, 2015 were used to validate broadband availability. Reports of "no service" override other validation methods.
- **Tower data and/or EDX propagation image -** For fixed wireless providers, we used tower location and system parameter information, where available, to propagate a fixed wireless provider's coverage area using EDX's Signal software, version 11.2. The wireless propagation model is based on the Anderson-2D propagation model. System parameters included frequency, transmit power, receiver sensitivity, antenna gain, and height. EDX produced coverage patterns for each tower/sector combination taking into account terrain and land use/clutter that may hinder signal dispersion. For terrain, we used two data sets, EDX universal .201 and SRTM 3-second .HGT format. For land use/clutter, we used is GCATTN_2011_clutter 30-meter .151 files. A separate propagation shapefile was created for each downstream and upstream speed tier combination, and all shapefiles were later overlaid and dissolved to where only the fastest advertised speed available was visible.

Regarding Mobile Data Interpolation

Through the mobile field testing program, the CPUC has shown that carrier-reported "highest advertised speeds" are not representative of the typical user experience. The FCC similarly rejects the adequacy of the carrier-reported maximum advertised speeds collected by the NTIA under its Broadband Data Initiative, and instead requires carriers to report their lowest advertised speeds on FCC Form 477 Deployment. The FCC has not yet determined whether "lowest advertised speeds" now being collected will be any better at representing that experience.

Assuming a normal distribution of data, adopting a speed standard at two standard deviations below the mean provides that estimate speeds would meet or exceed the speed standard 98% of the time. While the test results do not fall into a normal distribution, and the actual percent probability of availability will vary, we believe that speeds two standard deviations below the mean is more representative of a typical user experience than average speeds.

For the interpolation model used for mobile provider validation, we calculate, for every test location within a provider's footprint, a mean minus two standard deviation value for both upstream and

downstream speeds. The standard deviation is calculated from the 40 test results (20 for upstream, 20 for downstream) collected at each test location for each provider. We take the mean upstream and downstream speeds for each provider from the most recent mobile field test for each location (averaging smartphone and tablet speeds) and subtract two standard deviations for those upstream and downstream speeds from the means. The resulting mean minus two standard deviation values form the basis of the kriging (interpolation) model created for each provider. The image created by the kriging process looks similar to a heat map with color shading denoting high speeds, low speeds, and no service.