CALIFORNIA HYDROGEN INFRASTRUCTURE TOOL TECHNICAL FORMULATION WEBINAR

Main Technical Formulation Discussion

October 09, 2015

For questions or comments, contact:
Andrew Martinez
(916) 322-8449
andrew.martinez@arb.ca.gov

Discussion Outline

Purpose: The technical formulation discussion is intended to provide a detailed review of the methods, data sources, and considerations that guided the development of CHIT and allow public vetting of these methods

- Brief Introduction and review of CHIT and AB 8 process
- Guiding principles in CHIT evaluations
- Roadway and travel speed data set
- Coverage algorithms
- Traffic volume assessment
- First adopter FCEV market assessment
- Priority Area identification
- Capacity needs in Market and Priority Areas
- Open discussion of potential future development
- Discussion of potential public distribution

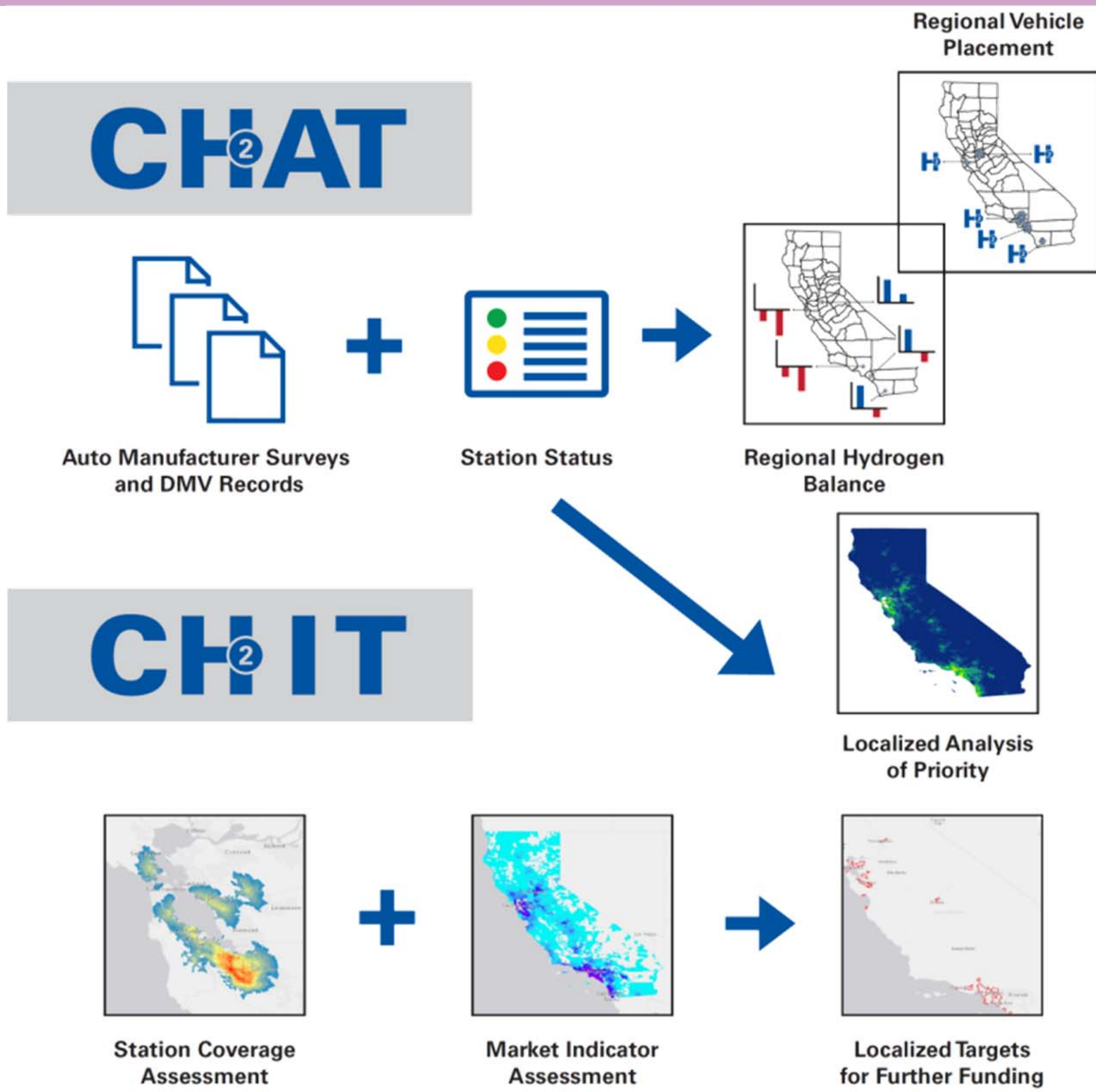
Discussion Outline

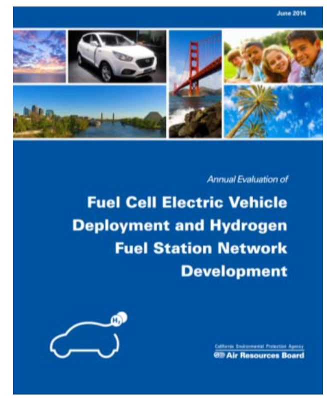
- This discussion will answer questions like:
  - What traffic-related information was utilized in CHIT? What information was not used and why?
  - What is the definition and mechanism of the coverage estimation algorithm?
  - How are financial indicators of the potential FCEV market assessed and compared in CHIT?
  - How does CHIT determine the location and relative rank of a Priority Area?
  - What are the current plans for further development of CHIT?

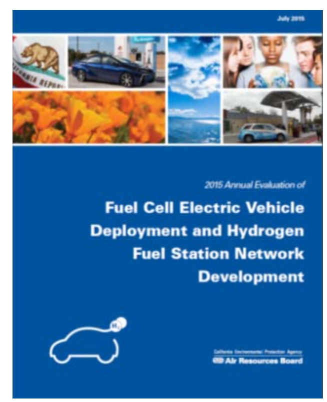
INTRODUCTION AND REVIEW OF CHIT

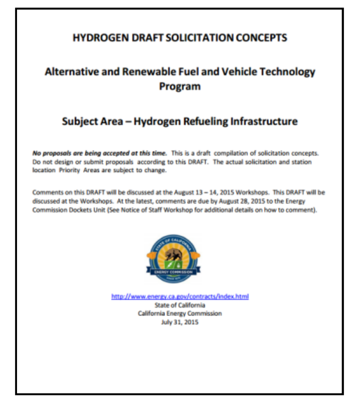
Introduction

CHIT/CHAT Tools and AB 8









HYDROGEN DRAFT SOLICITATION CONCEPTS

Alternative and Renewable Fuel and Vehicle Technology Program

Subject Area – Hydrogen Refueling Infrastructure

No proposals are being accepted at this time. This is a draft compilation of solicitation concepts. Do not design or submit proposals according to this DRAFT. The actual solicitation and station location Priority Areas are subject to change.

Comments on this DRAFT will be discussed at the August 13 – 14, 2015 Workshops. This DRAFT will be discussed at the Workshops. At the latest, comments are due by August 28, 2015 to the Energy Commission Dockets Unit (See Notice of Staff Workshop for additional details on how to comment).

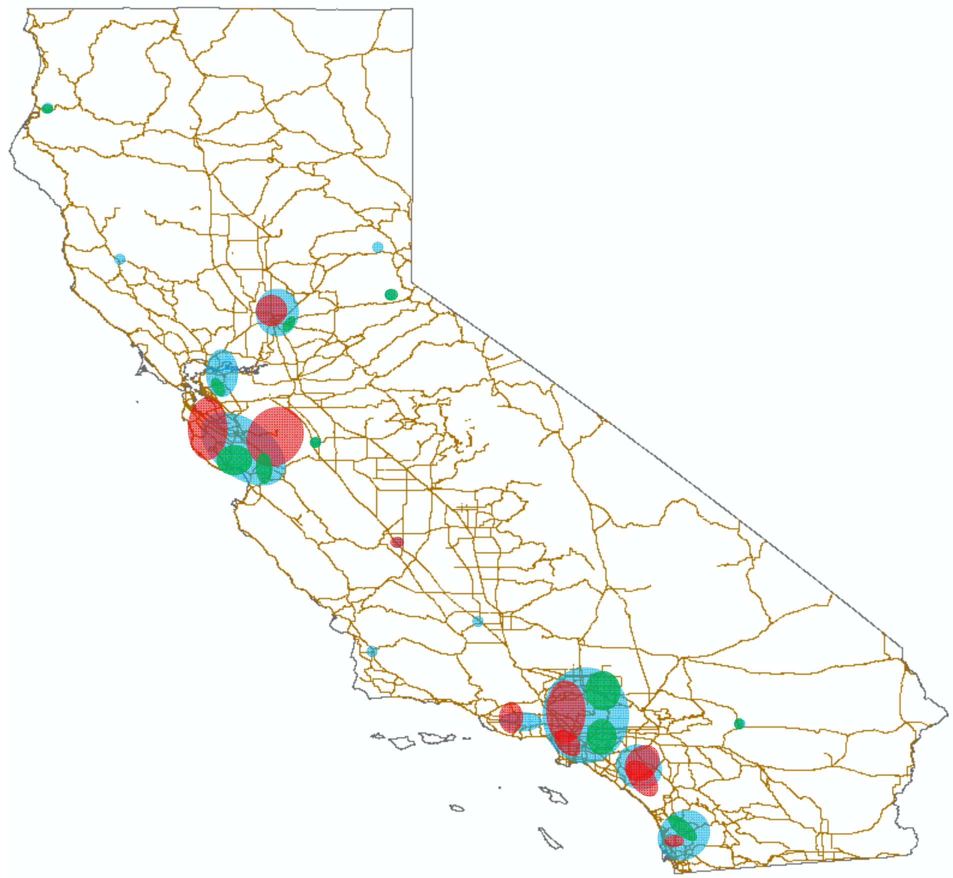
http://www.energy.ca.gov/contracts/index.html
State of California
California Energy Commission
July 31, 2015

Plan infrastructure placement appropriately for upcoming FCEV releases

1) Identify Market
   - Financial indicators
   - Green vehicle indicator
   - Education indicator

2) Evaluate current infrastructure
   - Existing and potential station coverage

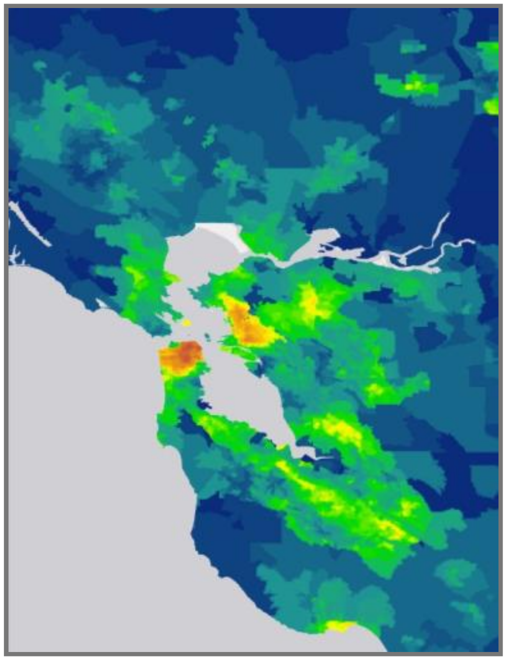
3) Prioritize uncovered market from year-to-year



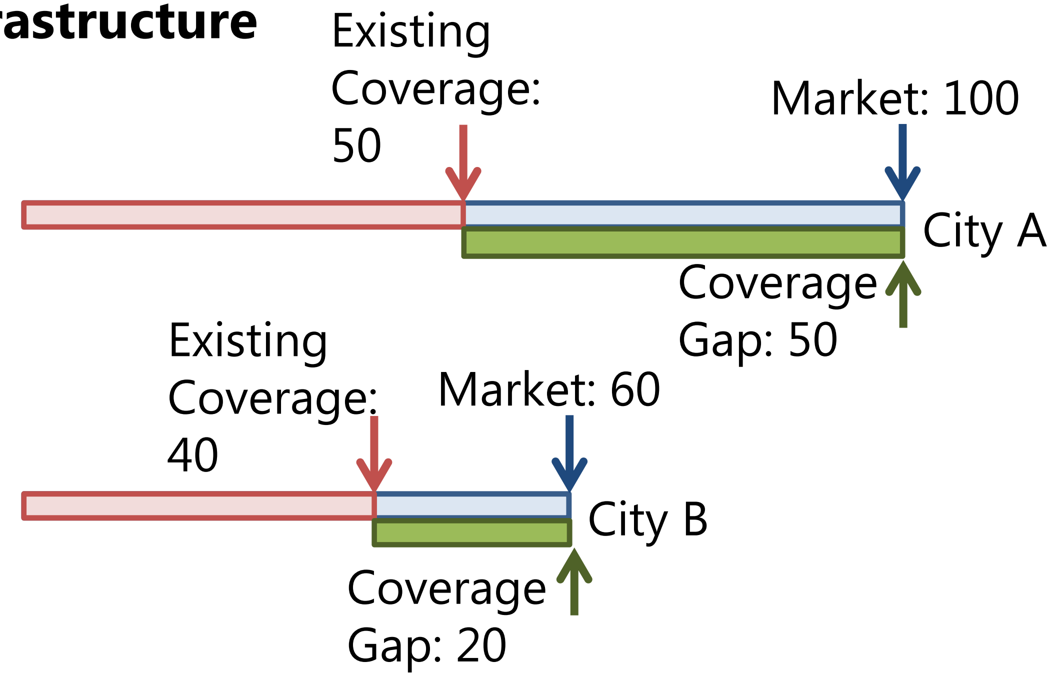
Introduction

CHIT: A Coverage and Market Assessment Tool

- **CHIT is a planning tool intended to provide general direction indicating areas of needed infrastructure**
- **CHIT evaluates relative need for hydrogen infrastructure based on a gap analysis between a projected market and current infrastructure**



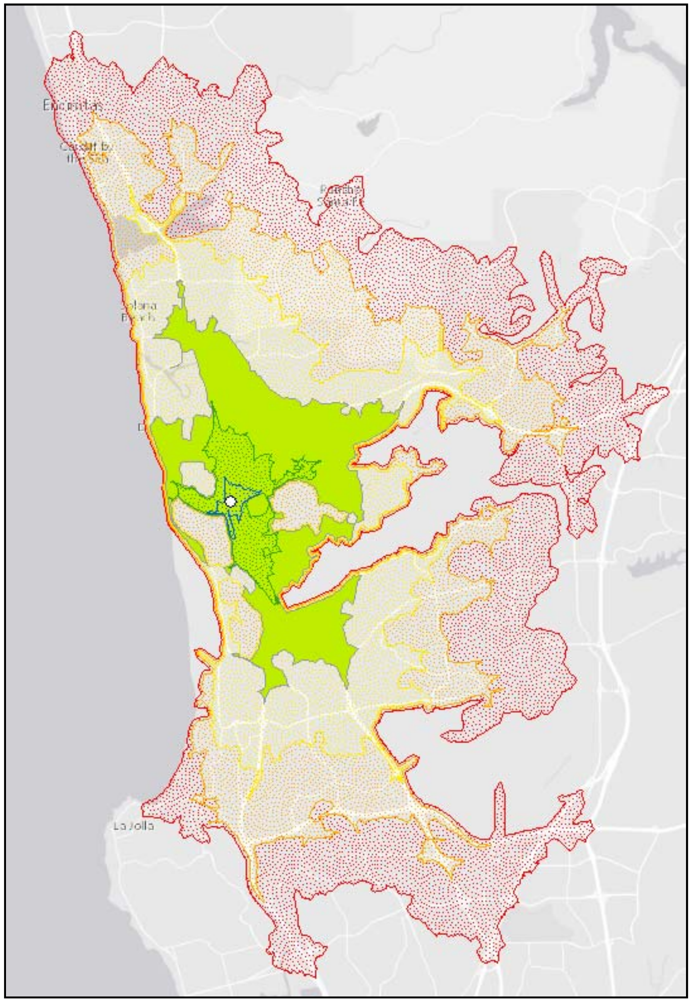
GIS Network Analysis and Station Area Planning



Introduction

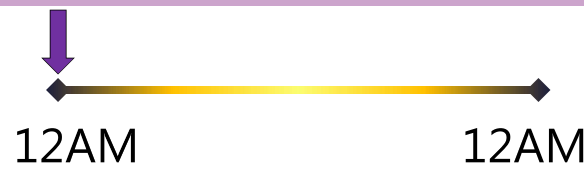
Central Theme: Coverage

- Conceptual representation of convenient access to fueling station
- Often discussed in terms of drive time, e.g. coverage is provided to all neighborhoods within a 6-minute drive of a station.
- Coverage can be conceptualized as binary (yes/no) or as degrees of coverage
- Well-planned coverage increases consumer confidence and adoption of vehicles



Introduction

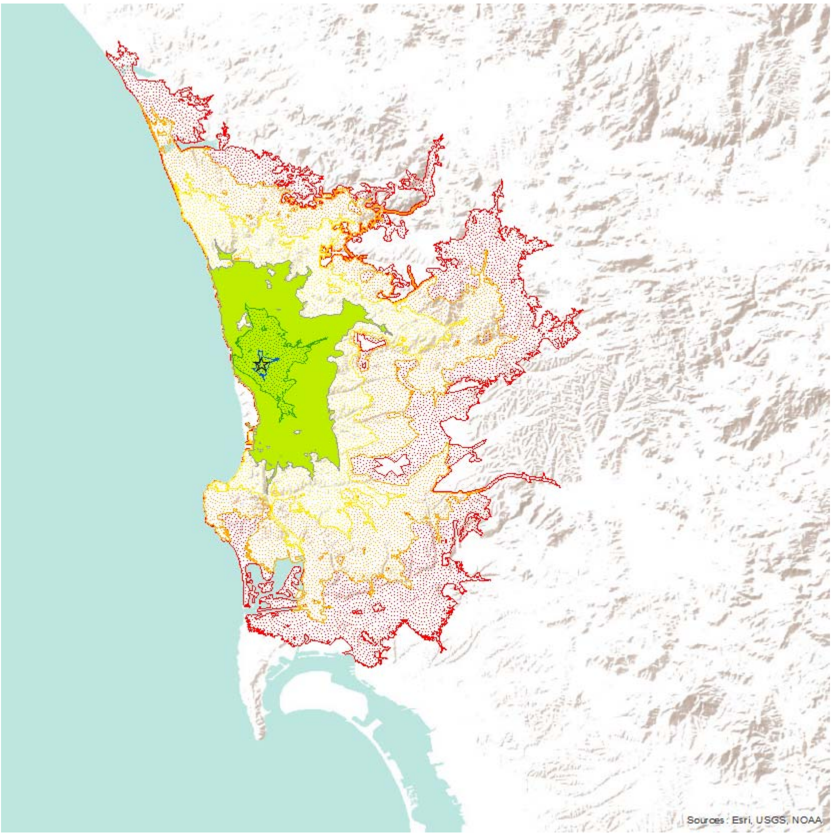
Coverage Depends on Traffic Patterns



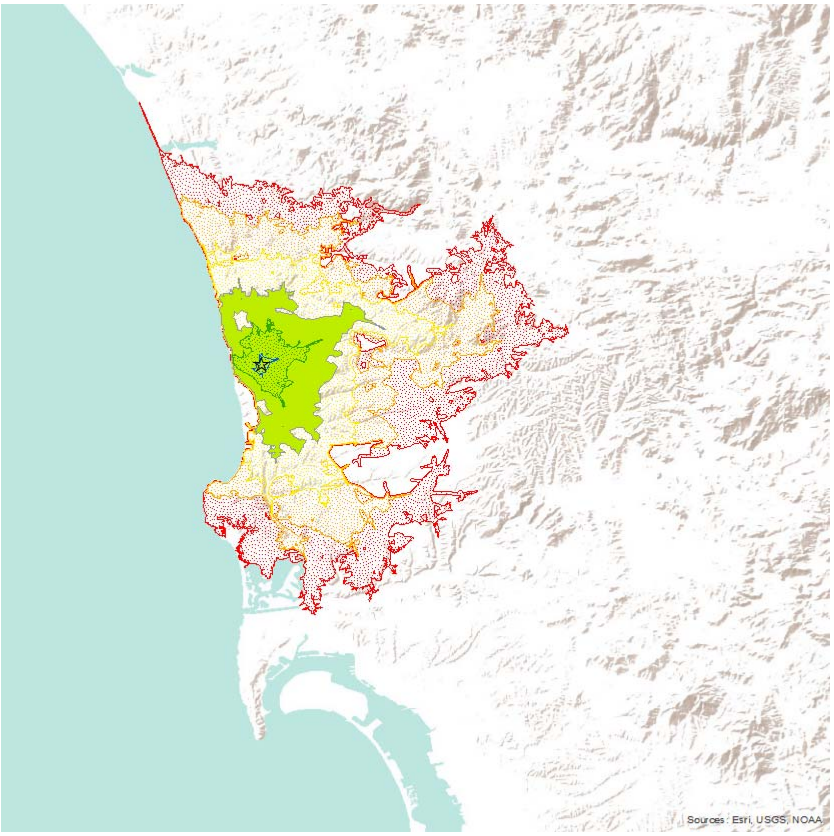
Service Areas for San Diego Station on Weekdays
(Based on Highway travel speed data and posted speed limits on city streets)

Introduction

Coverage Depends on Traffic Patterns



12 AM (No Traffic)

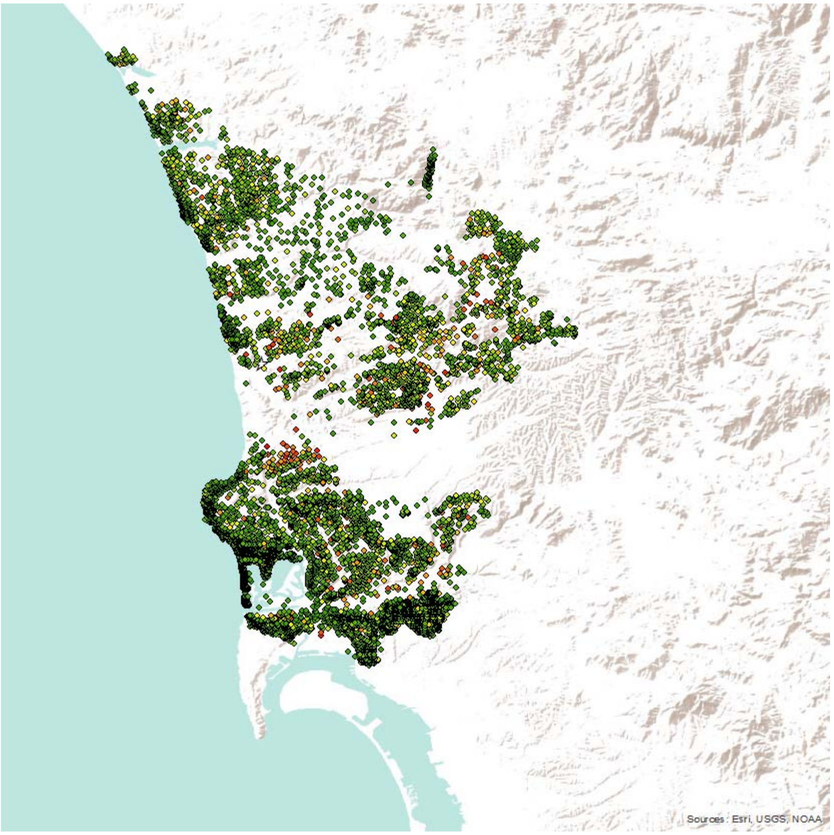


5 PM (High Traffic)

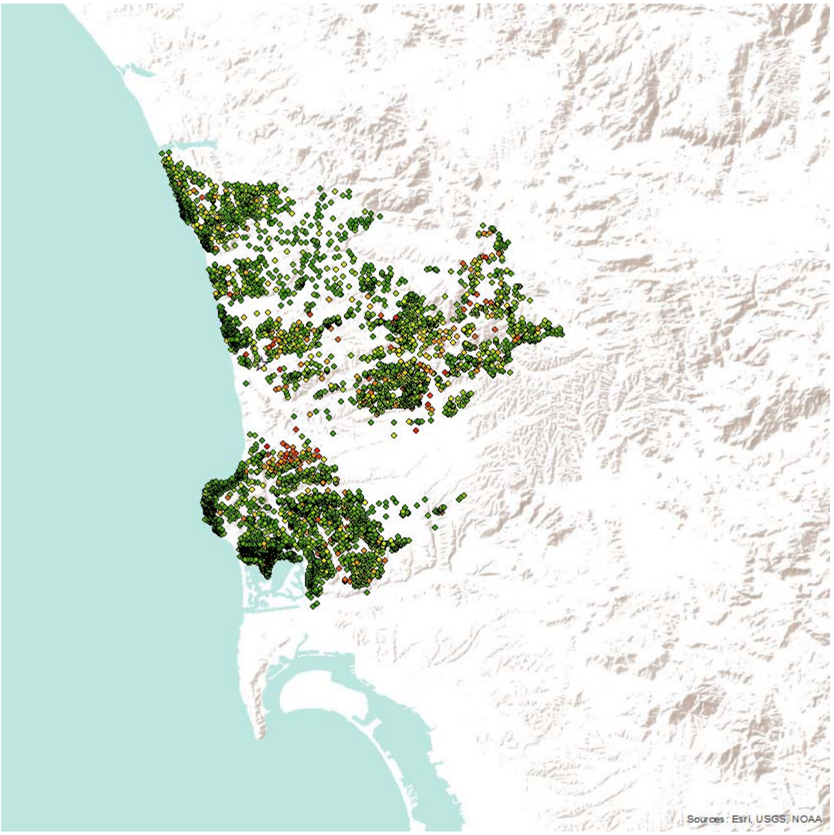
Introduction

Coverage Depends on Traffic Patterns

Low High



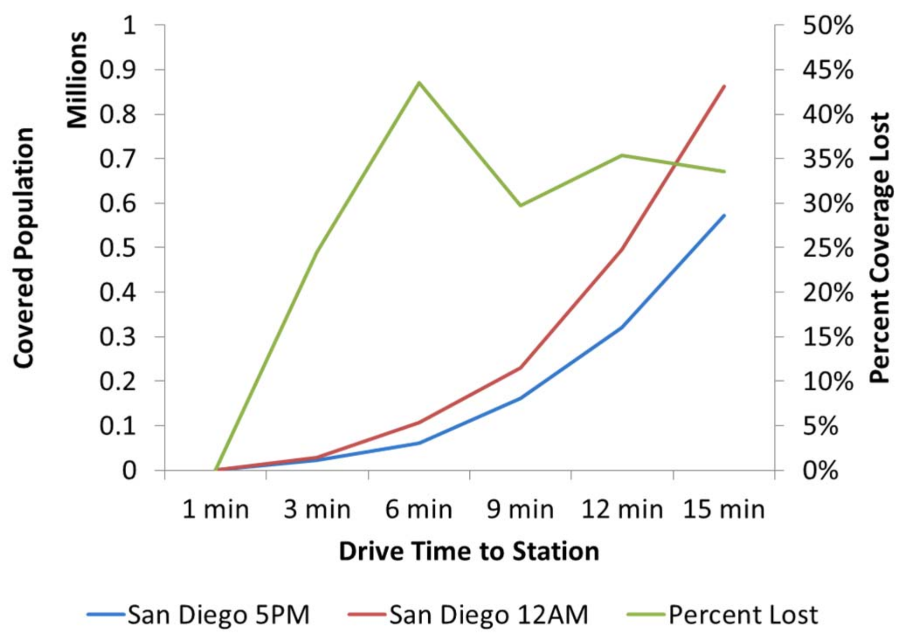
12 AM (No Traffic)



5 PM (High Traffic)

Introduction

Coverage Depends on Traffic Patterns



~1/3-1/2 of Covered Population Lost During Peak Traffic

- There is an observable and quantifiable relationship between coverage and traffic patterns
- The effect of the relationship varies with time of day, directly tied to the timing of peak traffic
- Time-varying analyses are interesting, but a PM peak analysis can allow planning for the worst-case scenario

GUIDING PRINCIPLES OF DEVELOPMENT

Guiding Principles

CHIT is envisioned as a tool that could be used year-after-year for public planning and reporting purposes and could provide a consistent assessment method across the entire state. Because of these motivations, there are a number of fundamental principles that determined the direction of developments in CHIT

- Principle #1: CHIT is a relative assessment
  - When planning for priorities, it is more important to understand how areas compare to one another on a relative scale than an absolute scale, which could be more sensitive to year-to-year changes

- Principle #2: CHIT is a statewide assessment
  - While individual site assessment can require special considerations based on information about the site and its surroundings, a broad statewide planning tool needs to be free of area-specific special considerations in order to provide a consistent basis of relative assessment. Data sources must therefore be as equally viable across the entire state as practicable.

Guiding Principles

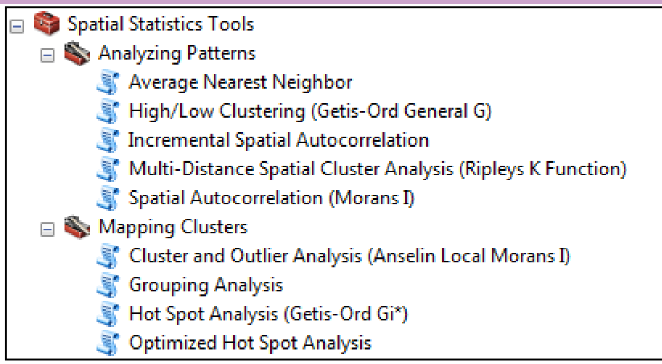
- Principle #3: In its current version, CHIT assesses only the first adopter market
  - The participants in the FCEV market will evolve over time, but the market is currently in its first adopter phase; first adopters are assumed to share certain characteristics
- Principle #4: Identification of the FCEV market can be estimated by consideration of the relative distribution of multiple demographic indicators
  - CHIT does not attempt to predict the absolute number of FCEV drivers in the market, but considers the spatial distribution of estimated attributes of the assumed FCEV drivers to indicate the relative spatial distribution of the market

Guiding Principles

- Principle #5: Accurate assessment of coverage, especially for shorter travel times, depends heavily on finely-detailed data
  - Development of the coverage assessment algorithm in CHIT focused heavily on development of a data set with high resolution and a high degree of completeness for the state's true roadway network
- Principle #6: Coverage matches the market when it provides convenient fueling access near FCEV drivers' homes
  - CHIT adopts the paradigm that convenient fueling is provided by stations near the FCEV drivers' homes. This is closely tied to the early adopter nature of the current market and the need to grow this market through highly visible demonstration of convenience.
- Principle #7: CHIT must be a tool that can be shared with the public
  - Data sources that are used in any analysis of the market and station coverage must not be confidential

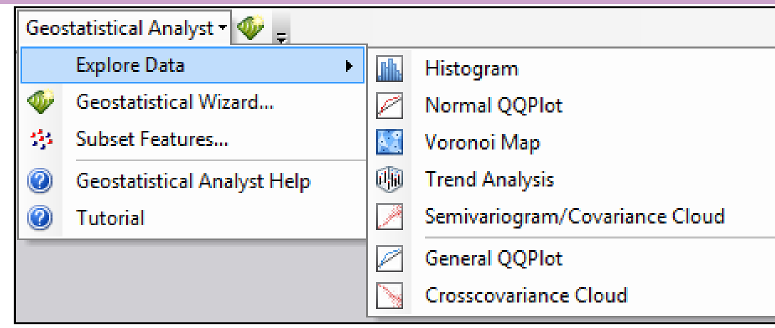
Guiding Principles

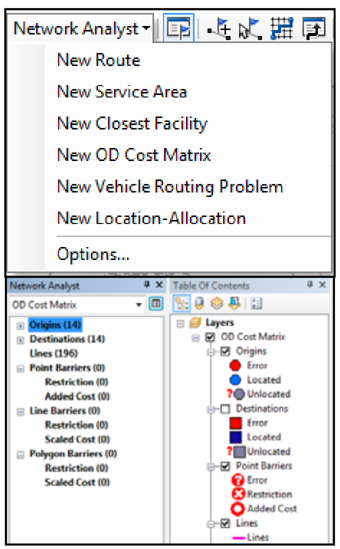
ArcGIS

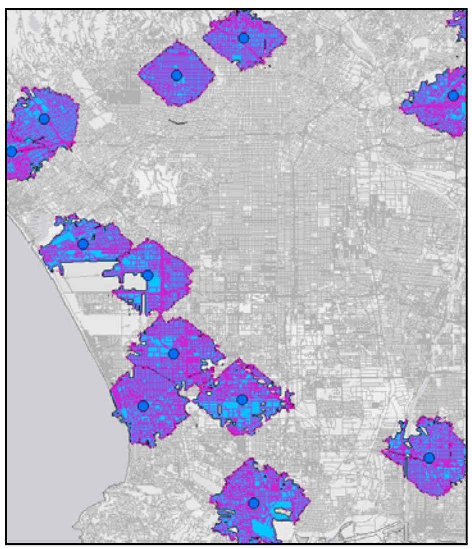


Results analysis

Input data exploration and validation

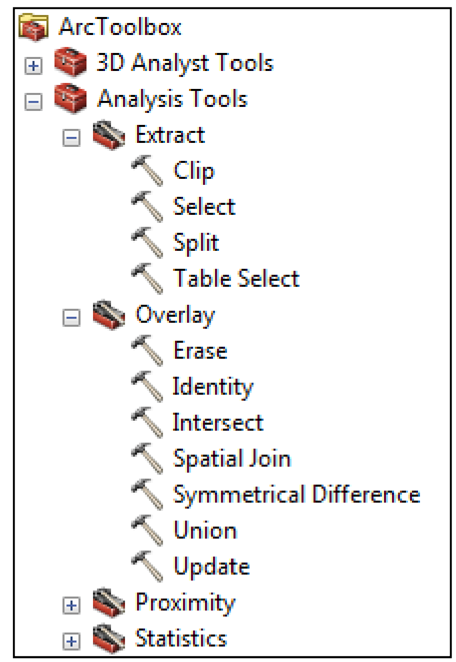






Determination of coverage provided by existing, funded, and potential stations

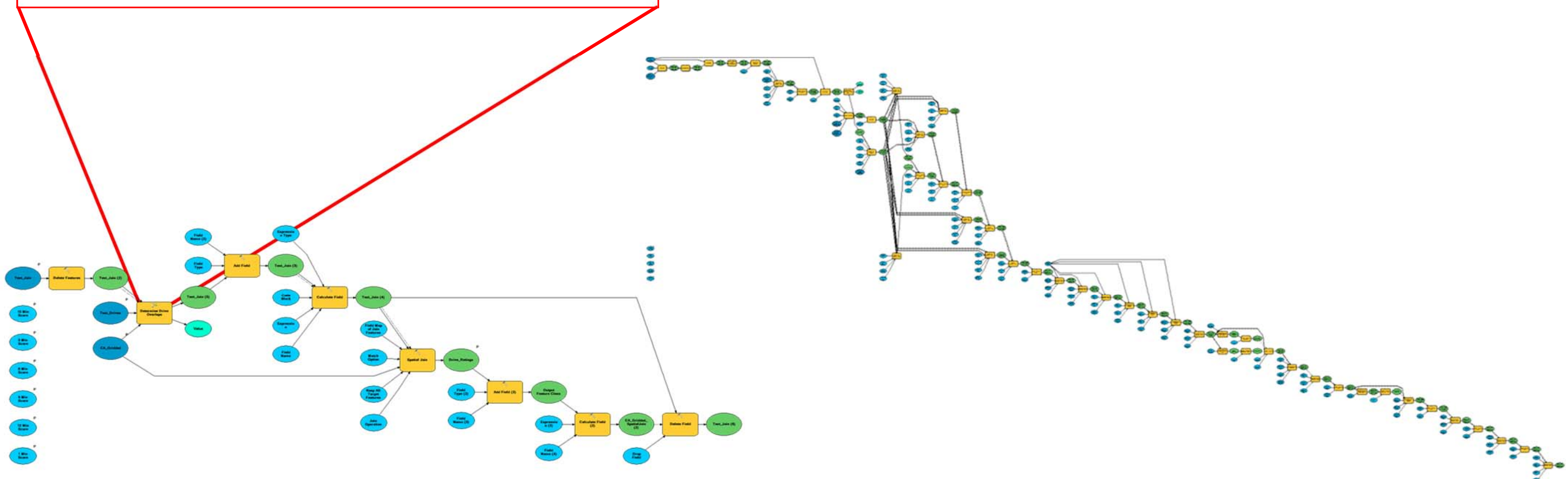
Analysis relies heavily on spatial overlays and correlations



Guiding Principles

ArcGIS

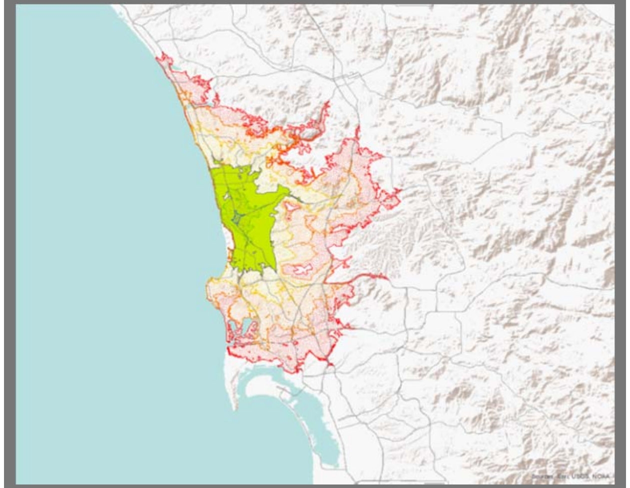
- CHIT_Tools
  - Determine Drive Overlaps
  - Determine Station Counts 6-Factor Match_Norm
  - Determine Station Counts 6-Factor Match_Norm_Alt
  - Filter Priorities by Area
  - Finalize Priorities
  - Generate Coverage Factor
  - Generate Heat Map
  - Prioritize Heat Map Areas
  - Reorder Priorities
  - Statistical Hot Spots

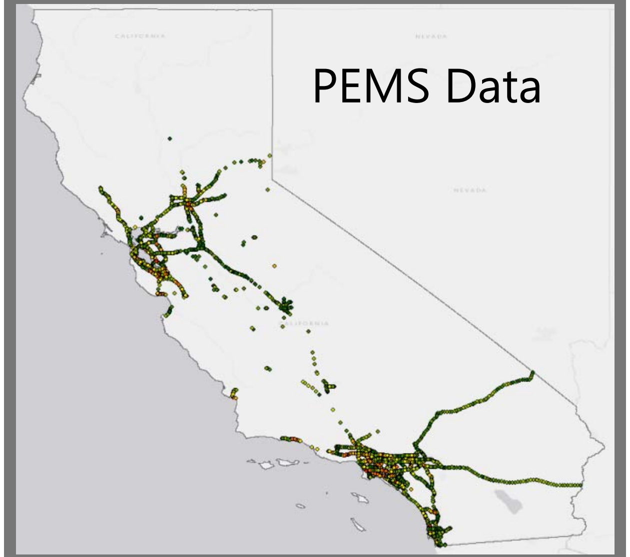


ROADWAY AND TRAVEL SPEED DATA SET

Roadway Data

TIGER from Census and PEMS from Caltrans





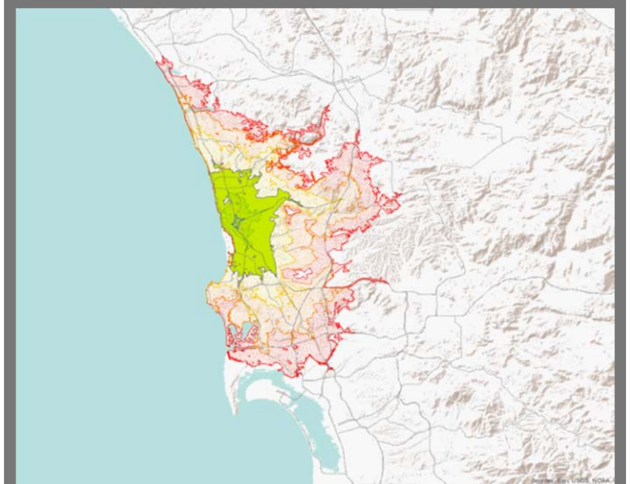
- Service areas calculated with hourly-varying traffic data in Caltrans PErformance Measurement System (PEMS) applied to Census-provided TIGER street geometry

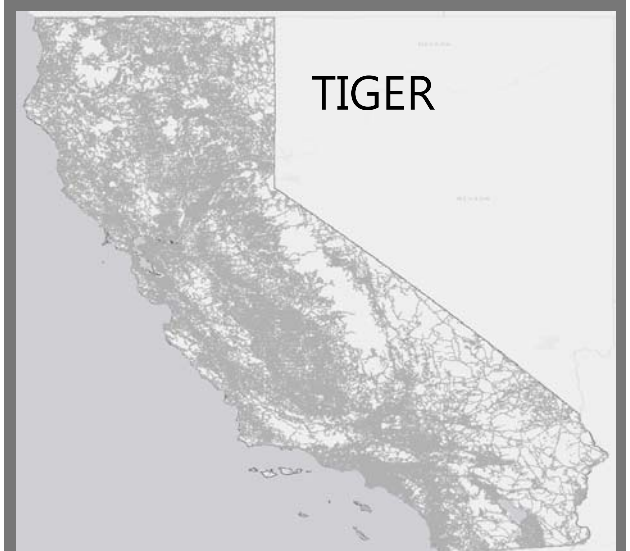
- PEMS provides real-world data of travel speeds at all hours of the day for large portions of the year

- PEMS limited to highways and not a complete coverage of highway mileage throughout the state
- Roads below highway classification require assumption of speed
  - Assumed speed limit based on classification of road
- Service areas in many regions unrealistically large

Roadway Data

TIGER from Census and PEMS from Caltrans





TIGER Network:

- Finely detailed map
- Does not lose resolution in less urban areas
- True-to-life map
- Can be spatially aligned to other map data sources

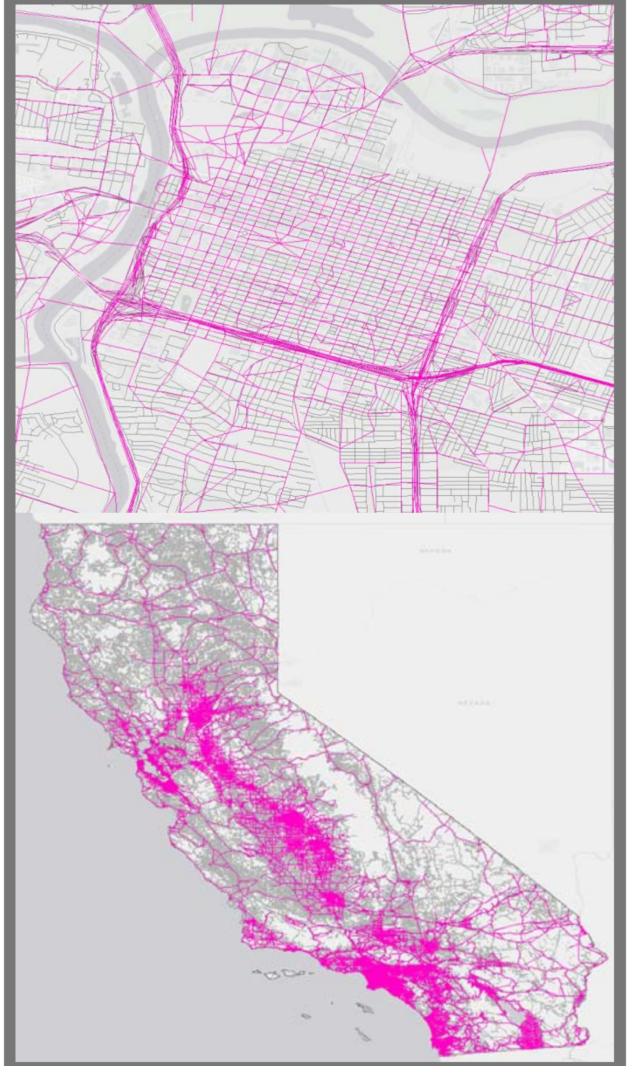
- Contains no performance attributes (volume, speed)
- Road classification system subject to interpretation
- Common in literature to date, make best guess at posted speed limit

SOLUTIONS:

- Find a different road dataset with attributes and use it
- Find a different road dataset with attributes and project it onto TIGER

Roadway Data

TIGER and ITN



Integrated Transportation Network (ITN)

- ITN used internally in conjunction with EMFAC for spatial emissions allocation*
- Stitched-together version of all Municipal Planning Organization traffic models
- Using PM Peak data

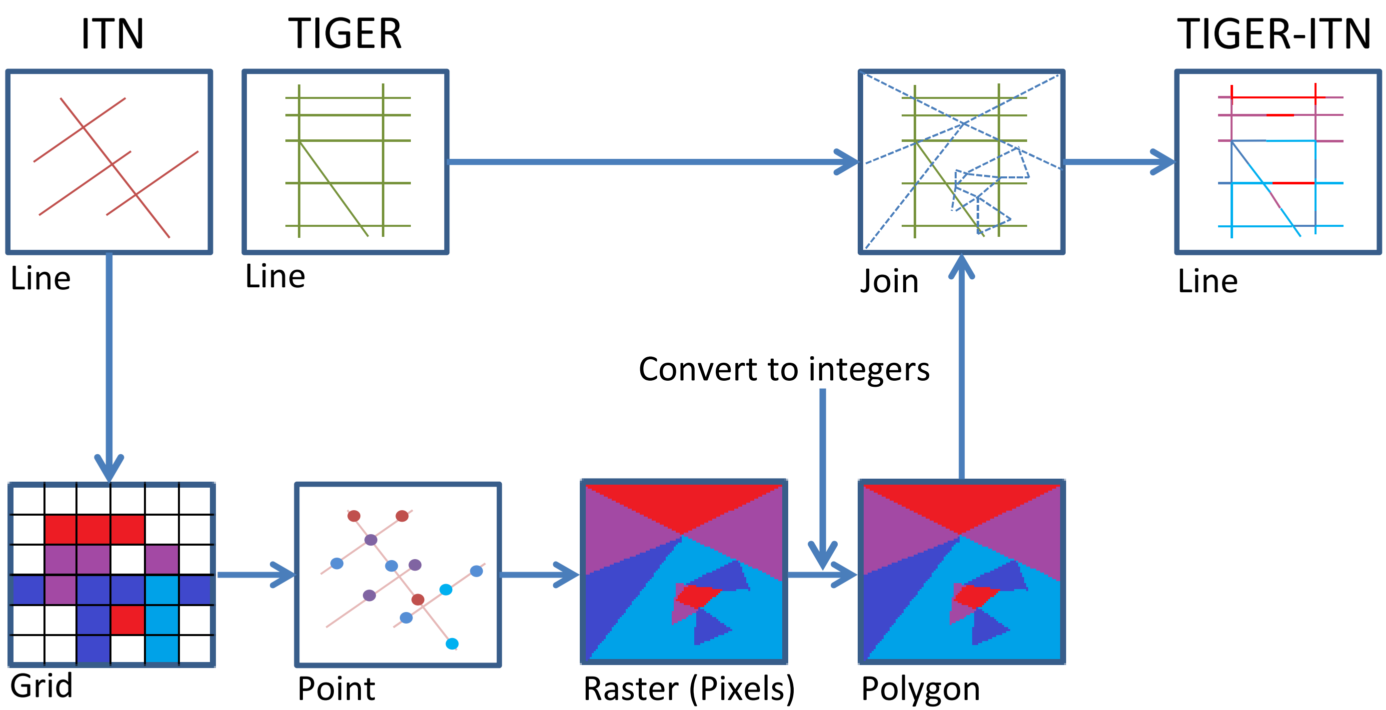
- Not a true-to-life network
- Network is an effectively equivalent model for planning purposes
- Data density follows population density

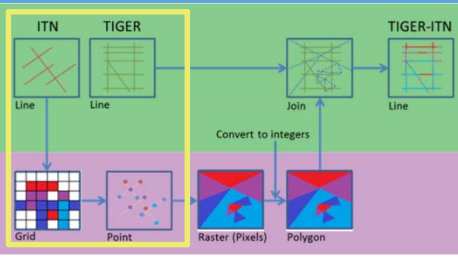
*Thanks to Nesamani Kalandiyur and Harikishan Perugu for guidance and collaboration on implementing ITN

Roadway Data

Interpolation and Extrapolation

- Objectives:
  - Utilize desirable assets of two datasets (data of ITN, geometry of TIGER)
  - Need to fill in data "gaps" of ITN
  - Need to project interpolated ITN attributes onto TIGER geometry
- ArcGIS requires certain geometry types for different steps



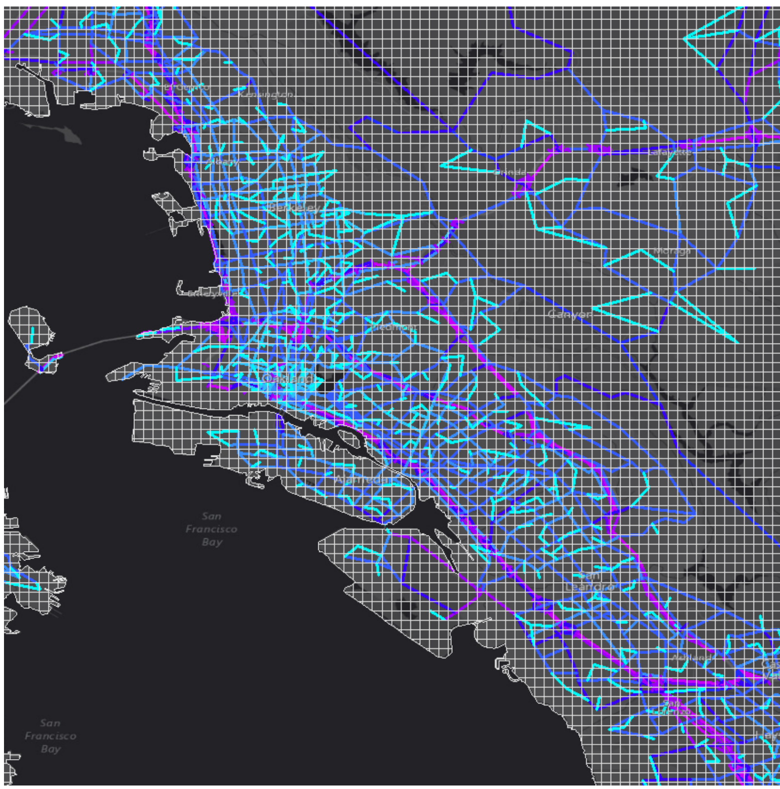


Roadway Data

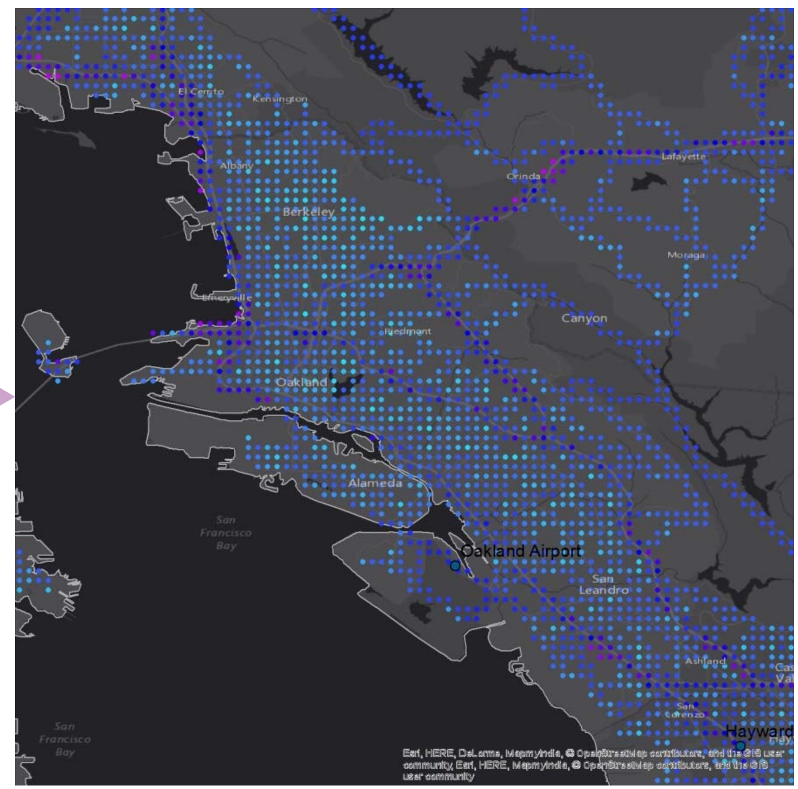
Interpolation and Extrapolation

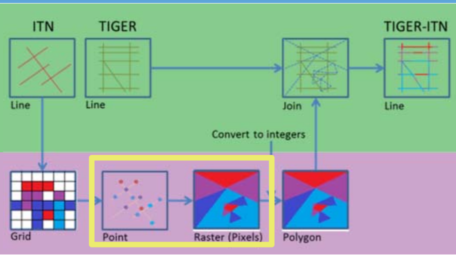
- From ITN Lines to 0.25-mile spaced points, ready for geostatistical evaluation and interpolation

Lines



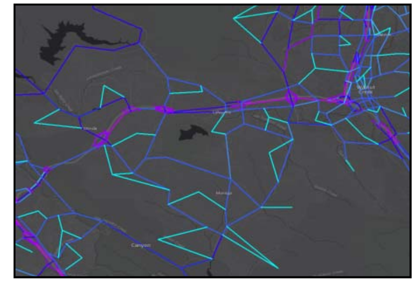
Gridded Points



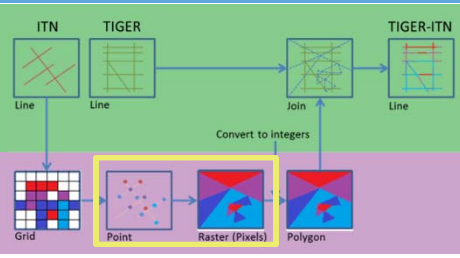


Roadway Data

Interpolation and Extrapolation



- Data clean-up prior to gridding
  - Contains several spikes of zero data
  - Due to "representative model" nature of original ITN
  - Allows for improved solutions in native model, but can skew data in this application
- Interpolation method
  - Roadway data intended to cover entire state; specialized and complex interpolation models require large amounts of information to accurately represent entire state
  - Simpler, general rules investigated for interpolation
  - Inverse Distance Weighted and Kriging (Ordinary Type) investigated
    - IDW directly related to fundamental theorem of geography
    - Kriging simplest available statistical model
  - Allowed ArcGIS to automatically tune parameters

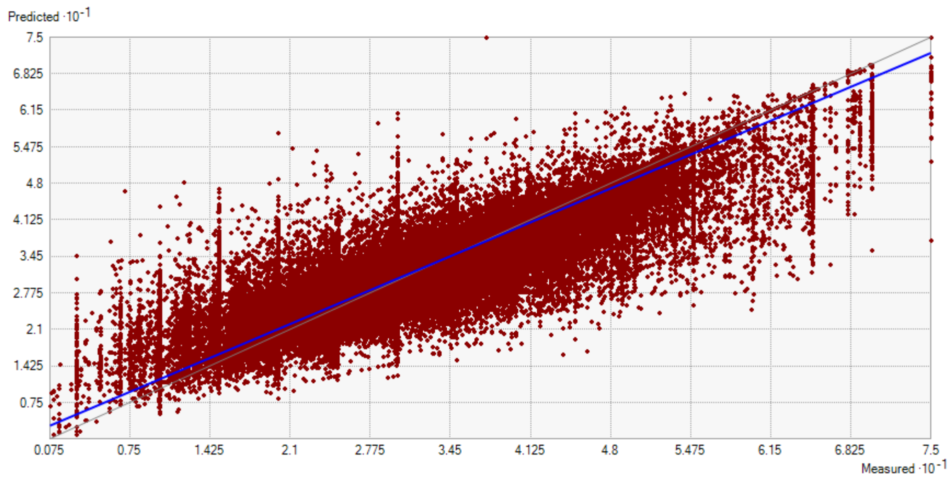


Roadway Data

Interpolation and Extrapolation

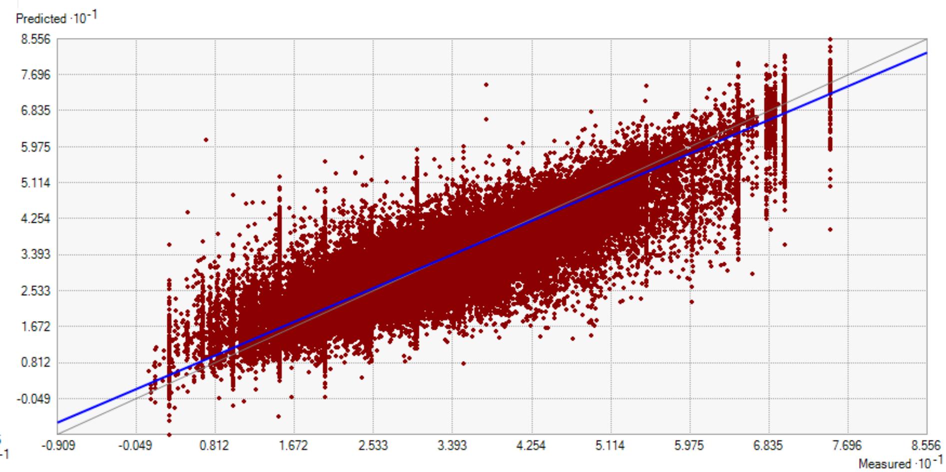
- IDW is exact at input points and will not exceed bounds
- Kriging assumes input may just be a sample of real data; is not necessarily exact and may exceed bounds
- Cross-validation plots and error measures indicate no major improvement by Kriging
- Kriging over/under-predicts maximum and minimum and predicts some negative values

IDW

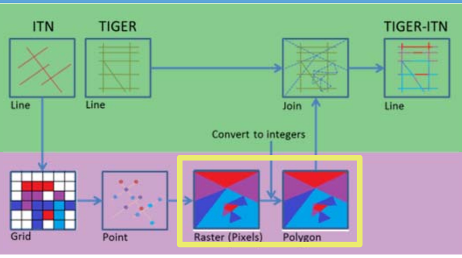


Mean Error: 0.0135
RMS Error: 3.24

Kriging



Mean Error: 0.0138
RMS Error: 3.45

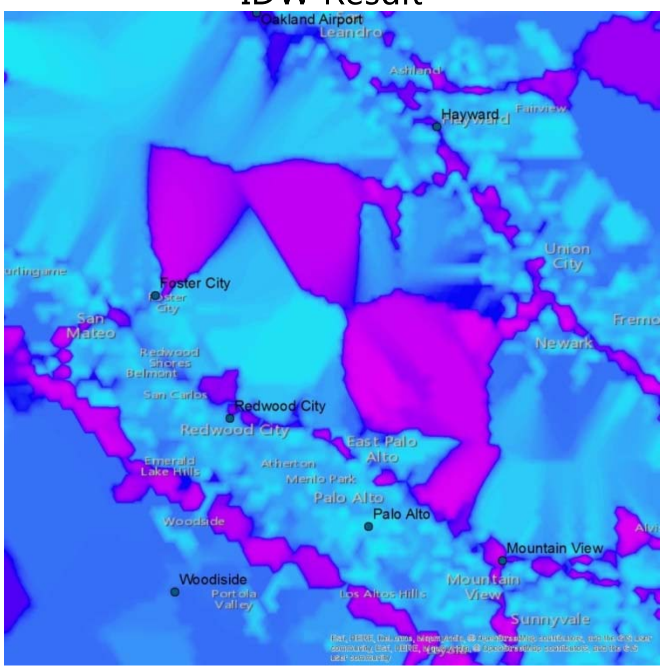


Roadway Data

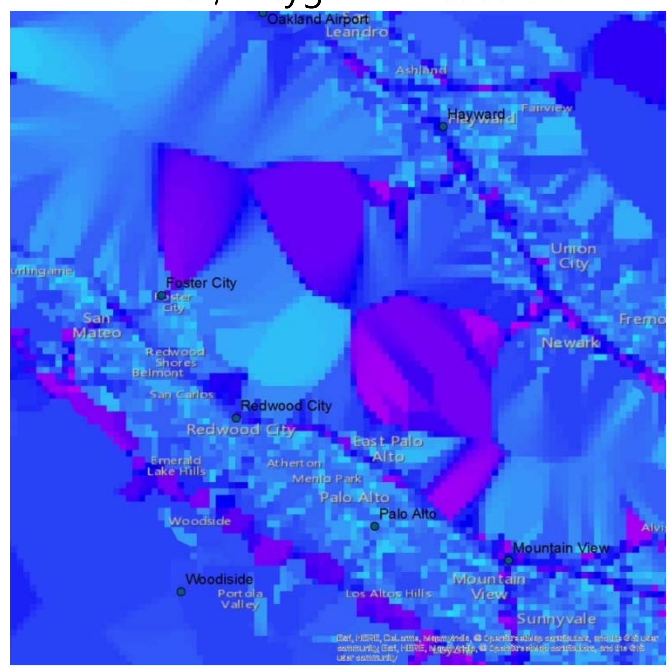
Interpolation and Extrapolation

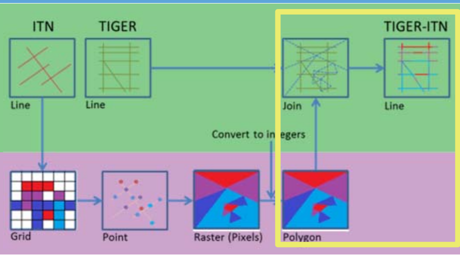
- Export IDW to raster (pixel-level image)
- Convert from raster to polygons to enable cut of TIGER lines by polygon boundaries
- Polygons converted to integer format to reduce number of unique values

IDW Result



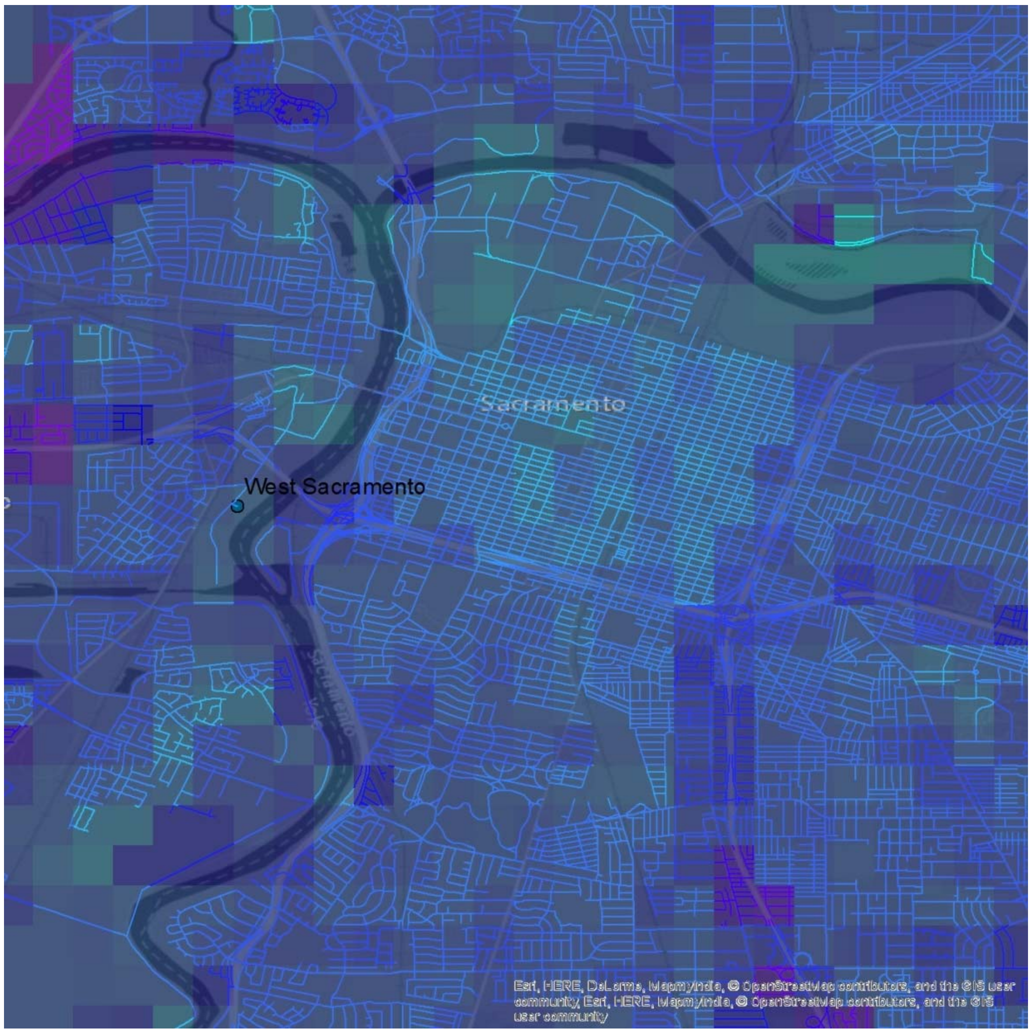
IDW Exported to Raster, Converted to Polygons in Integer Format, Polygons “Dissolved”





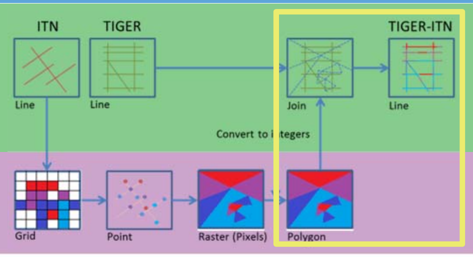
Roadway Data

Interpolation and Extrapolation



- Sacramento area shown
- Computed resolution of interpolation raster from ITN data apparent
- Lines shown are TIGER lines
- Where a TIGER line overlays two or more differently-colored cells, it is split to create new lines, each with the appropriate ITN speed estimate
- Increased number of TIGER lines from ~1.1 to ~1.6 million

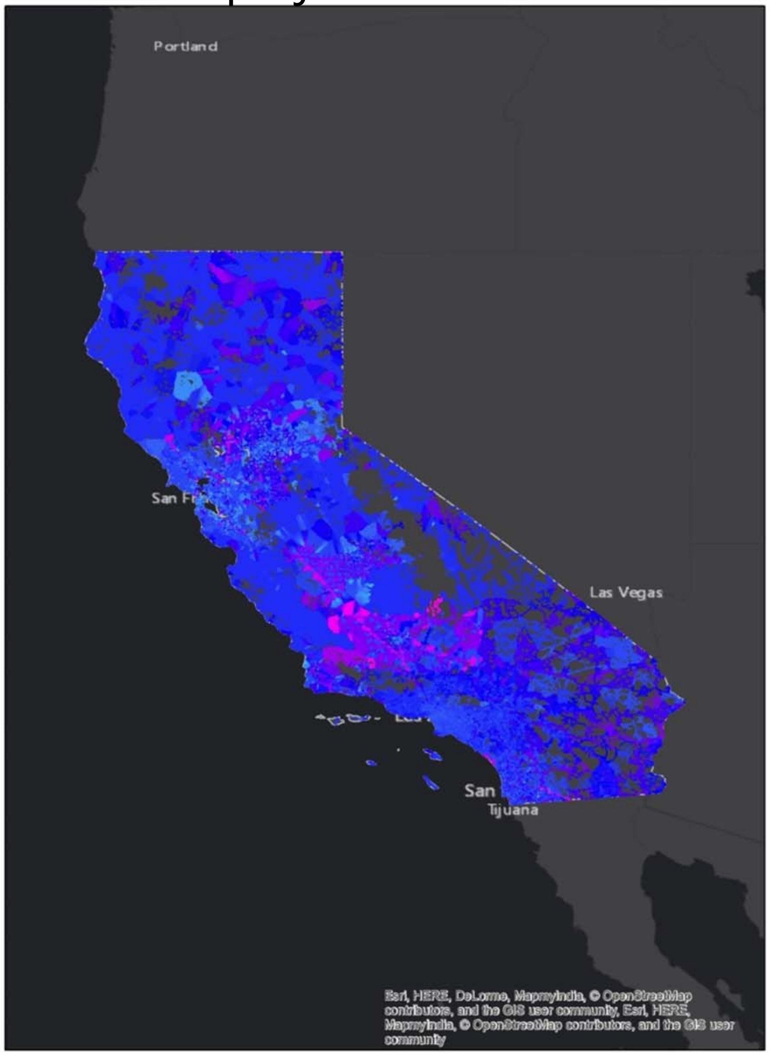
Roadway Data



Interpolation and Extrapolation

ITN Interpolation Map

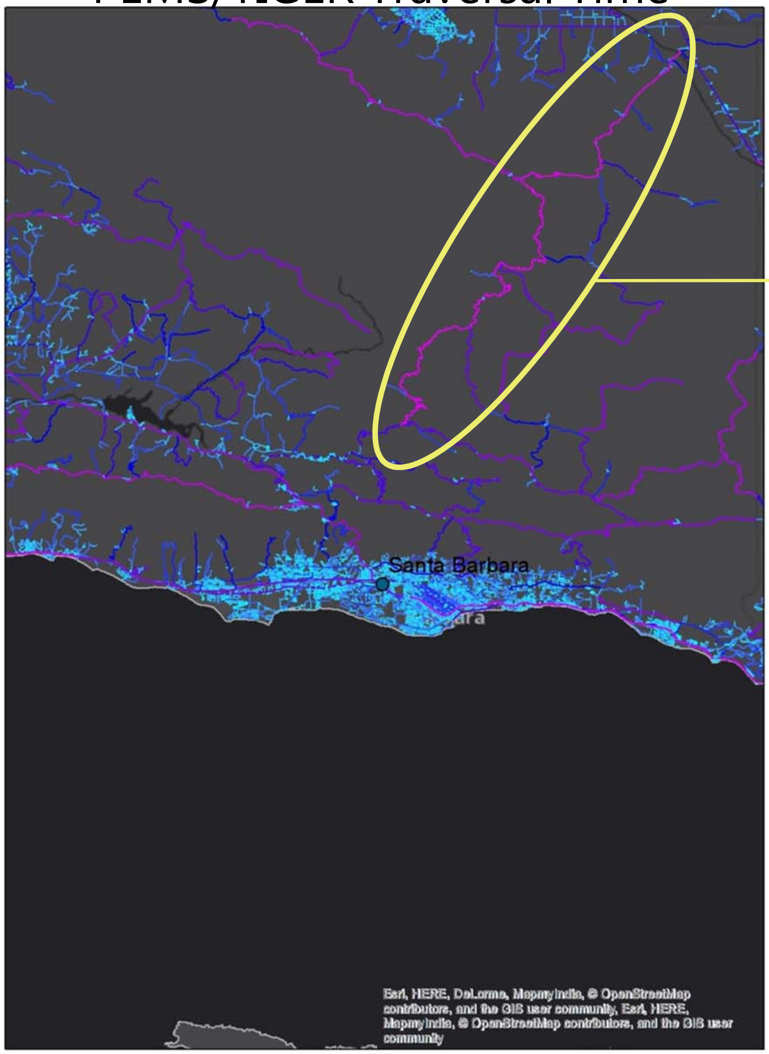
ITN projected on TIGER



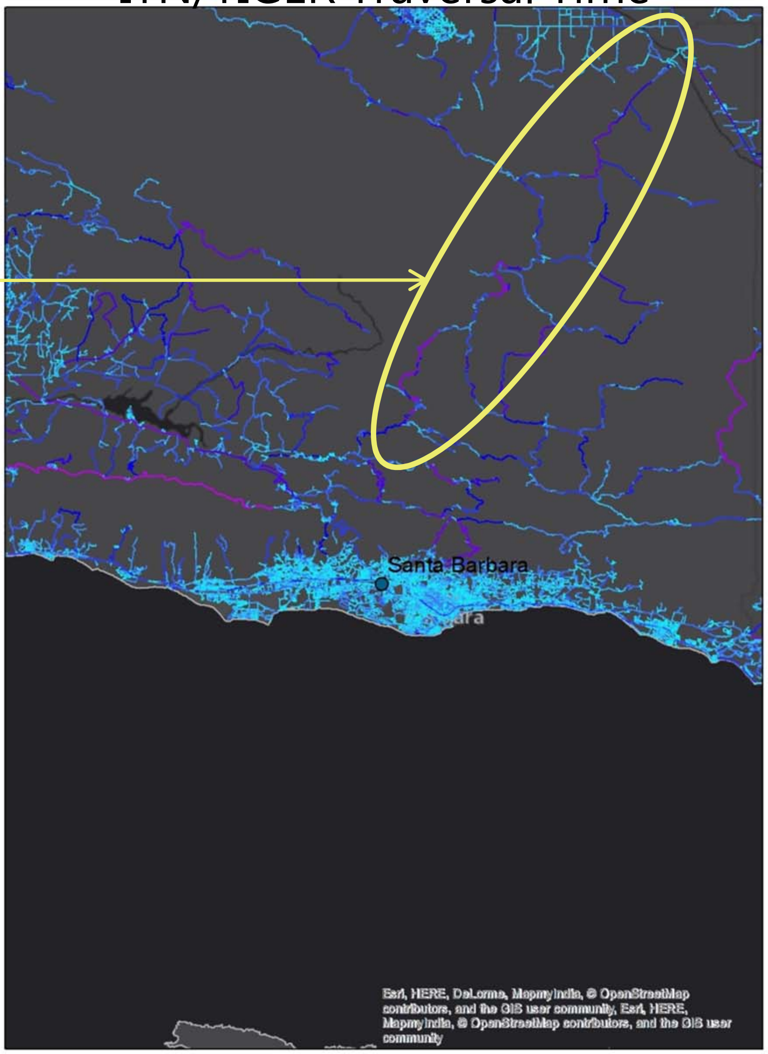
Roadway Data

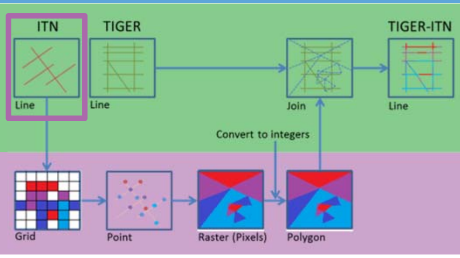
Interpolation and Extrapolation

PEMS/TIGER Traversal Time



ITN/TIGER Traversal Time

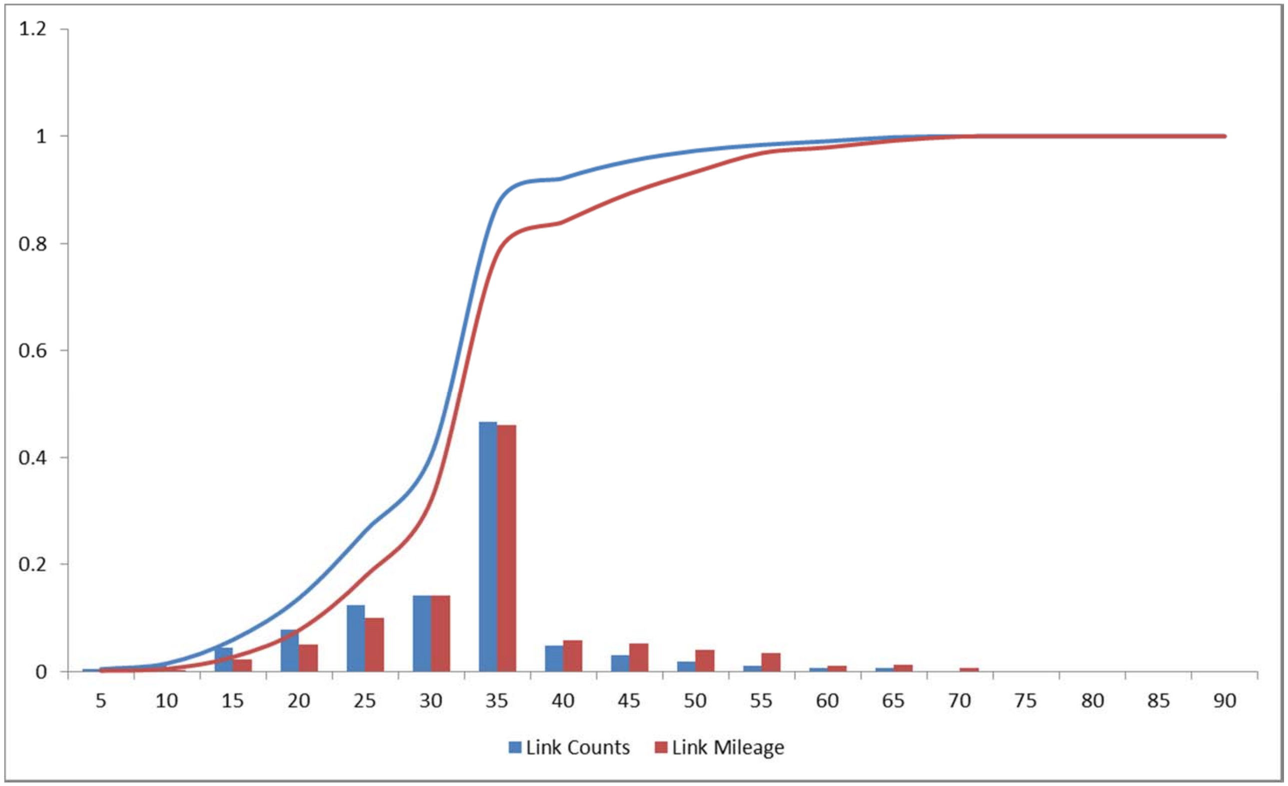


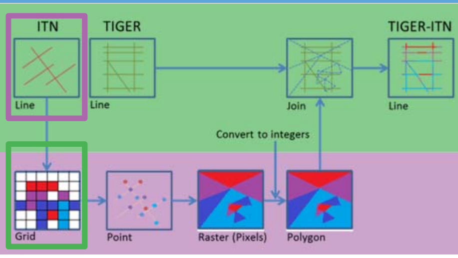


Roadway Data

Validity of ITN/TIGER

- Distributions of speed in ITN master data by count of linkages in each bin and by total mileage of linkages in each bin

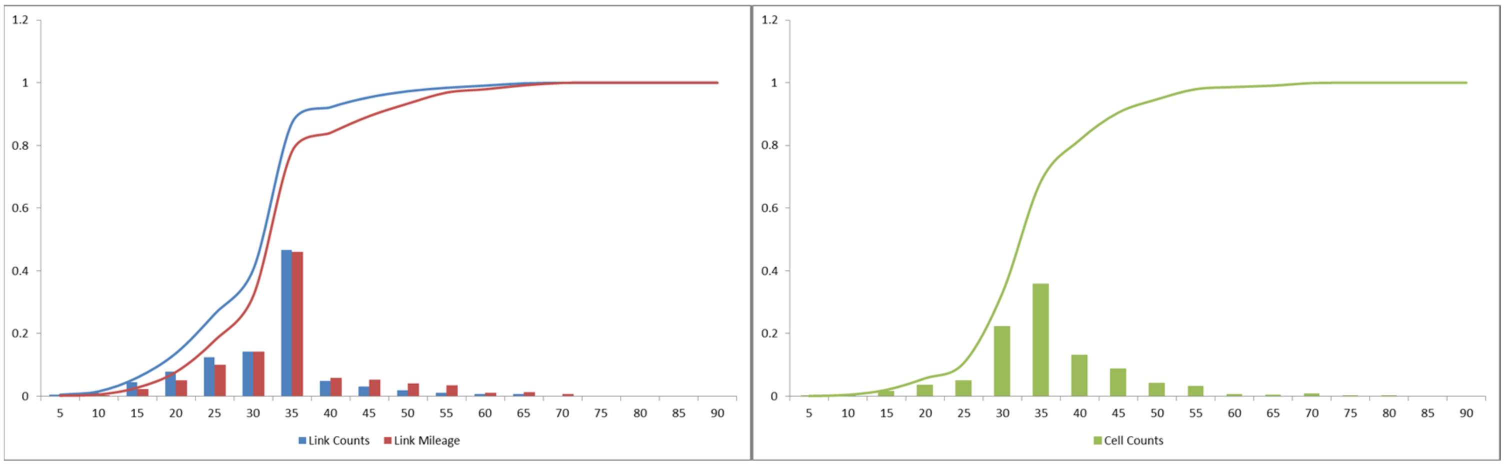




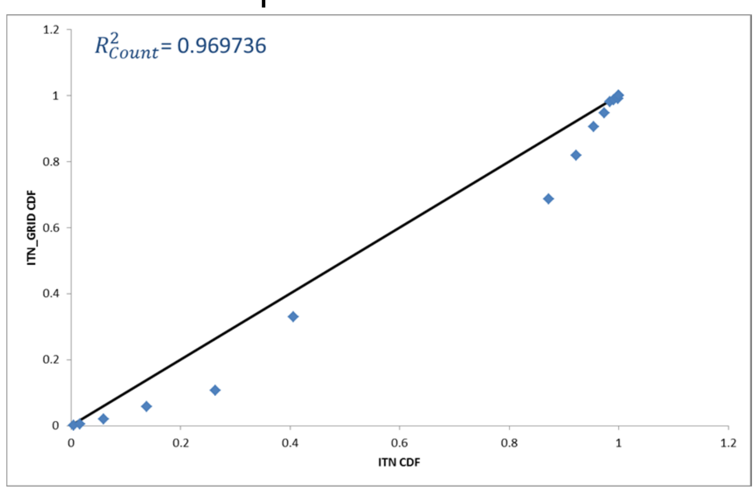
Roadway Data

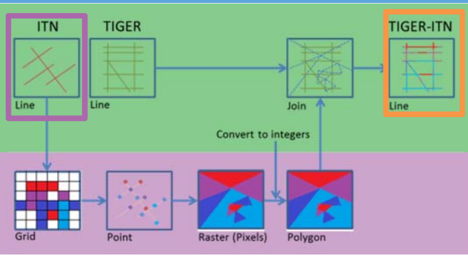
Validity of ITN/TIGER

Distributions of speed in gridded version of ITN by count of cells in each bin



Comparing count cumulative distributions, errors at extremes and some "spread" of values but acceptable overall fit

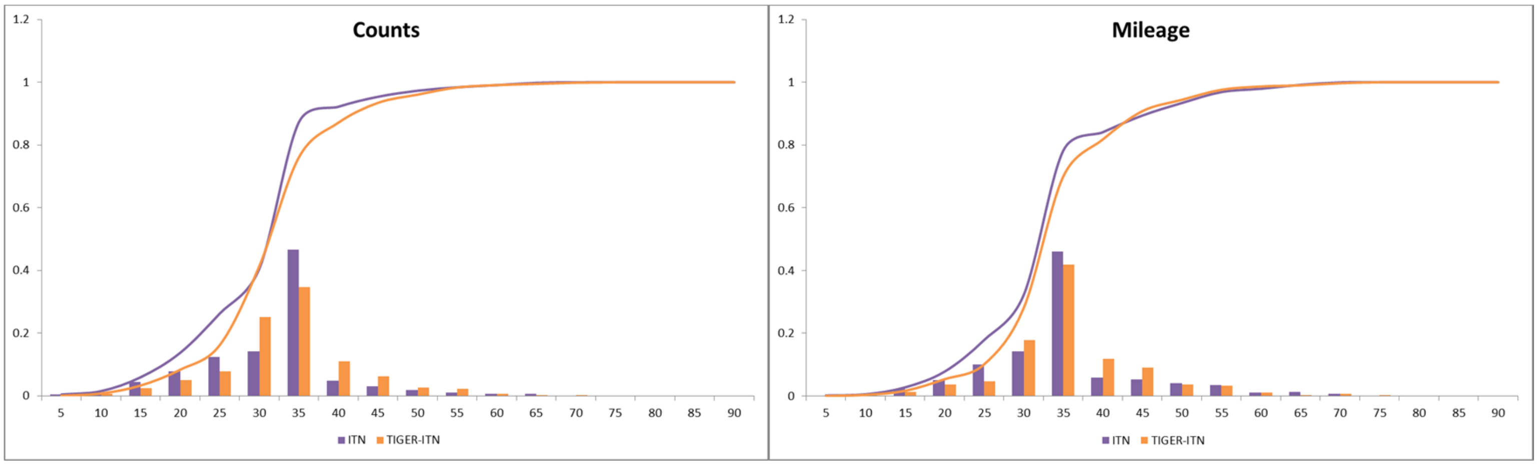




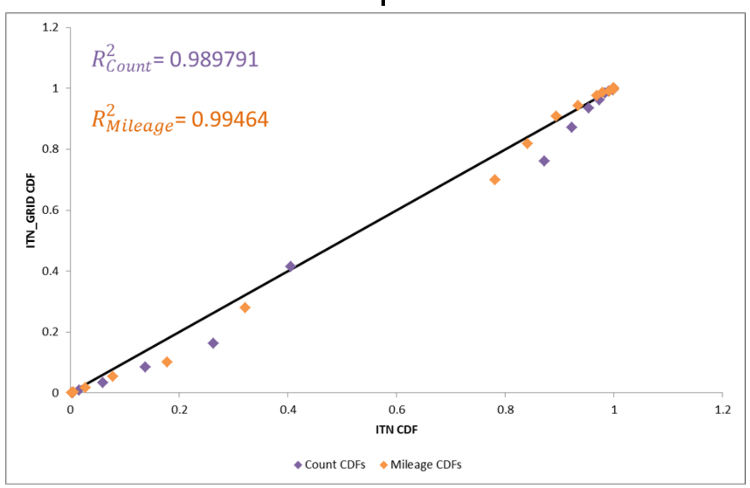
Roadway Data

Validity of ITN/TIGER

Distributions of speed in ITN/TIGER by link count and mileage in each bin



Comparing distributions, achieved better agreement when applied to TIGER lines but still may be room for improvement



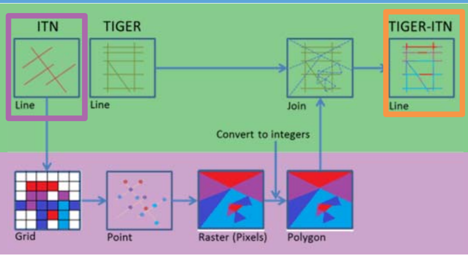
Roadway Data

Validity of ITN/TIGER

- Investigated a number of possible improvements that did not significantly alter match between ITN raw data and interpolated and extrapolated results
  - Length-weighted averaging for ITN link distance within each cell

Cells may contain multiple links from ITN data set of varying lengths. Possible to consider each link's contribution as determined by its length in the cell.

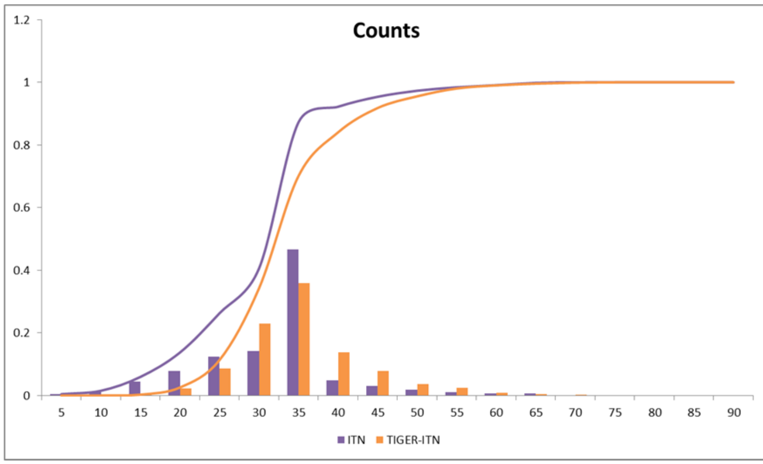
  - Decrease resolution of grid to 1/8-mile square
  - Non-averaging conversion to points
    - Input data for interpolation method follows roadways at defined intervals
    - Uses actual point data at defined intervals

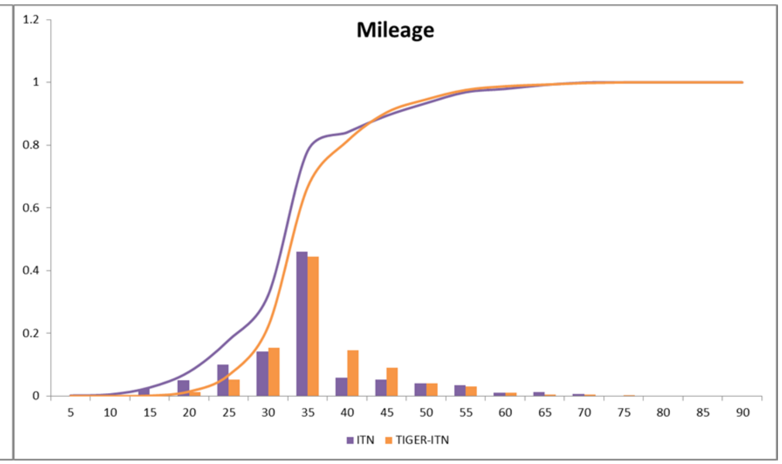


Roadway Data

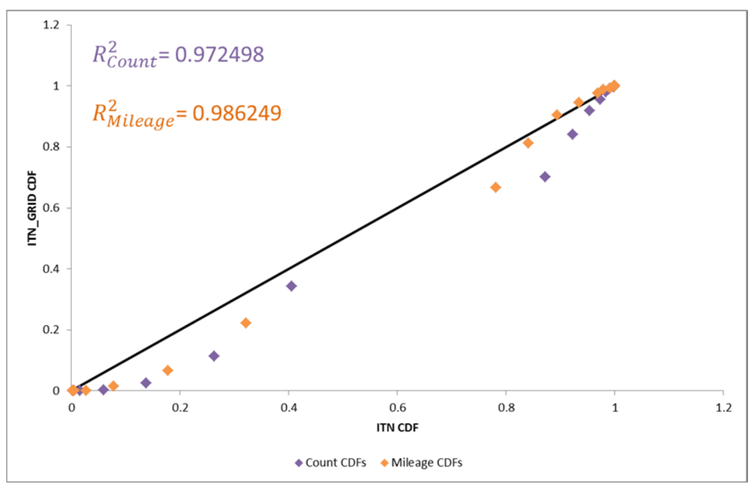
Validity of ITN/TIGER

Distributions when using 1/8th-mile points on lines





Histogram shows better match at peak, but resolution in lost low-speed tail



Roadway Data

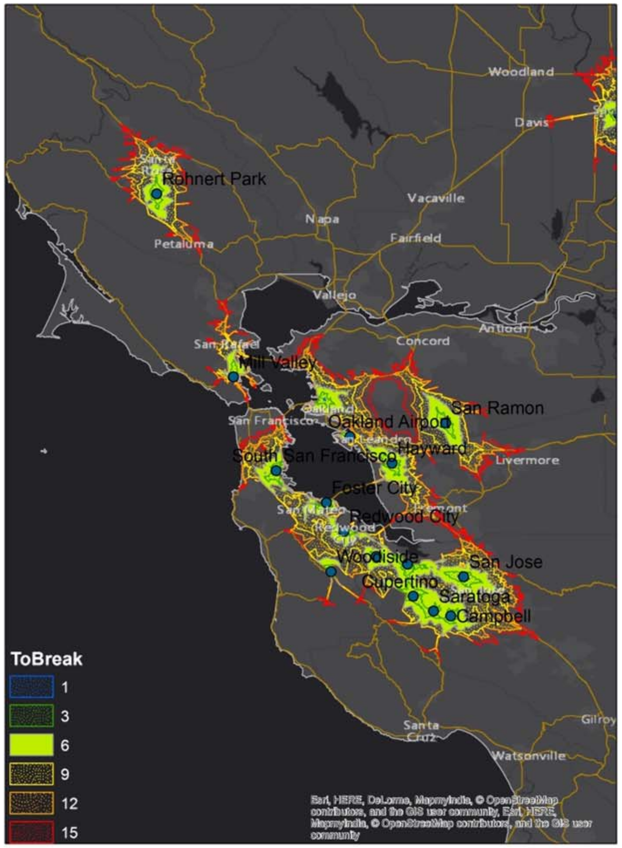
Validity of ITN/TIGER

- Considering all investigated options, settled on IDW interpolation method with non-weighted averaging of ITN data in cells of ¼-mile resolution
  - IDW is a deterministic interpolation method, ensuring that predicted value at locations with an actual data point will match; ensures range of predictions is maintained (no negatives and no overly-high speed projections)
  - Kriging is not deterministic and saw predicted values did not match input data as well
  - Increased resolution and length-weighted averaging did not noticeably improve match to raw data
  - Non-averaging method missed extreme values, especially at low speed end

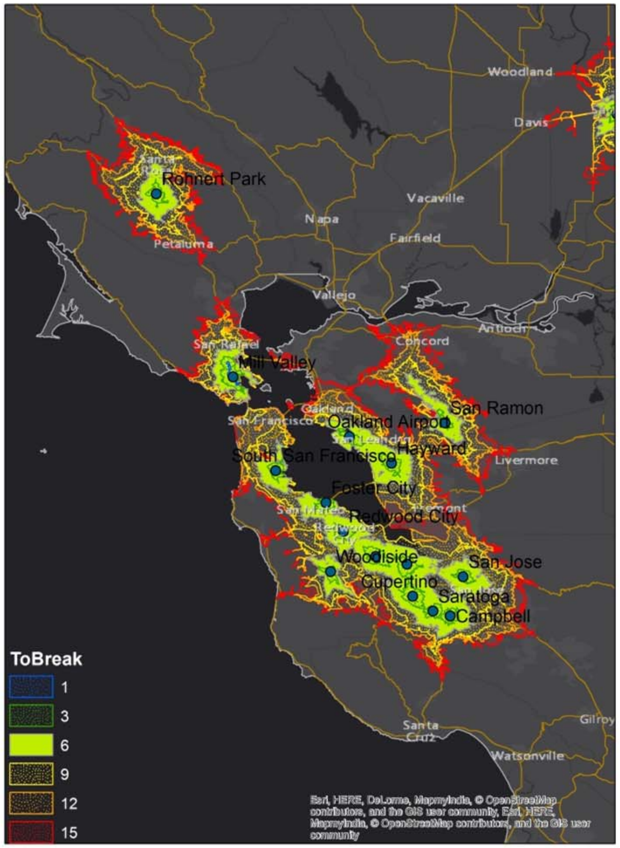
Roadway Data

Comparing Service Areas

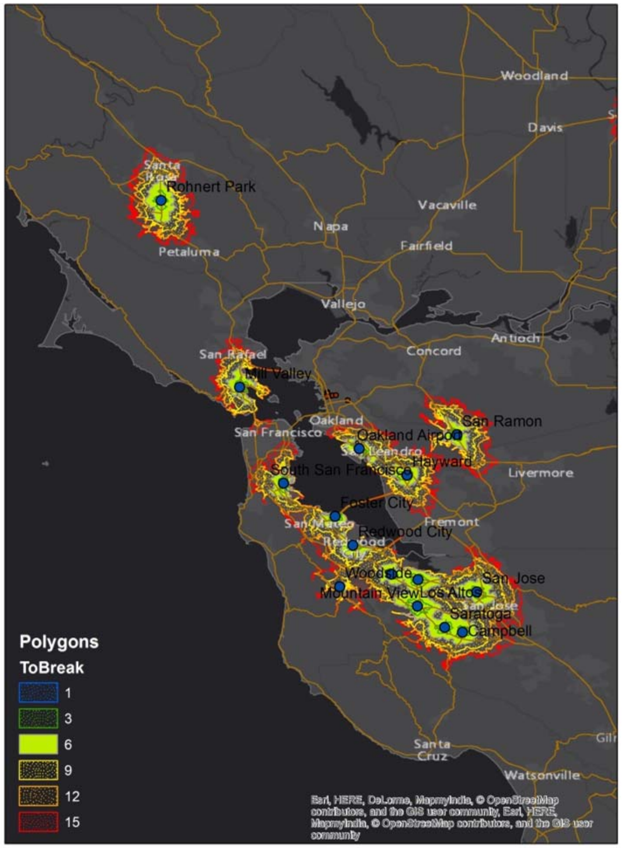
ITN



PEMS/TIGER



ITN/TIGER

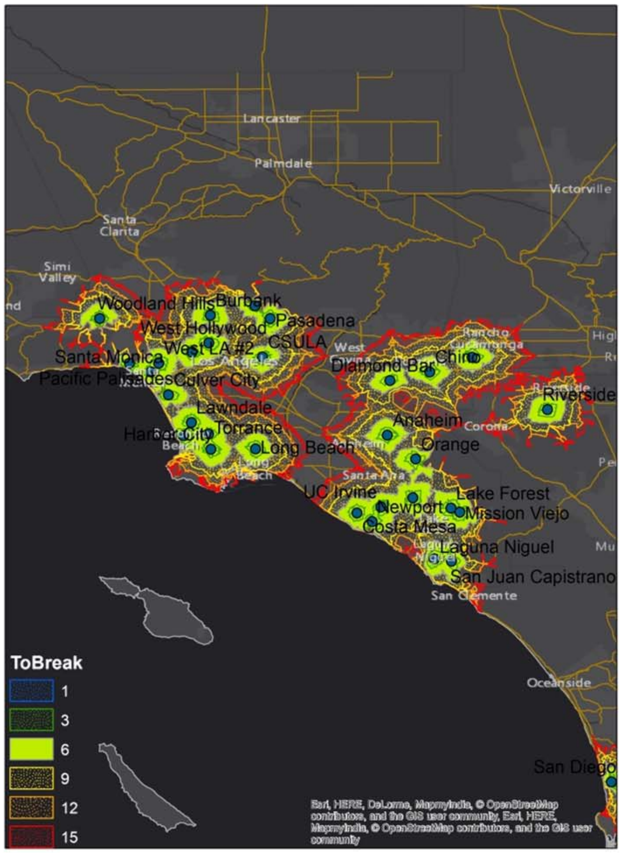


Note: "ITN" case from data in 2012 ITN without projecting onto TIGER, developed with data to 2010
"ITN/TIGER" case from 2015 ITN, developed with data to Dec. 2014

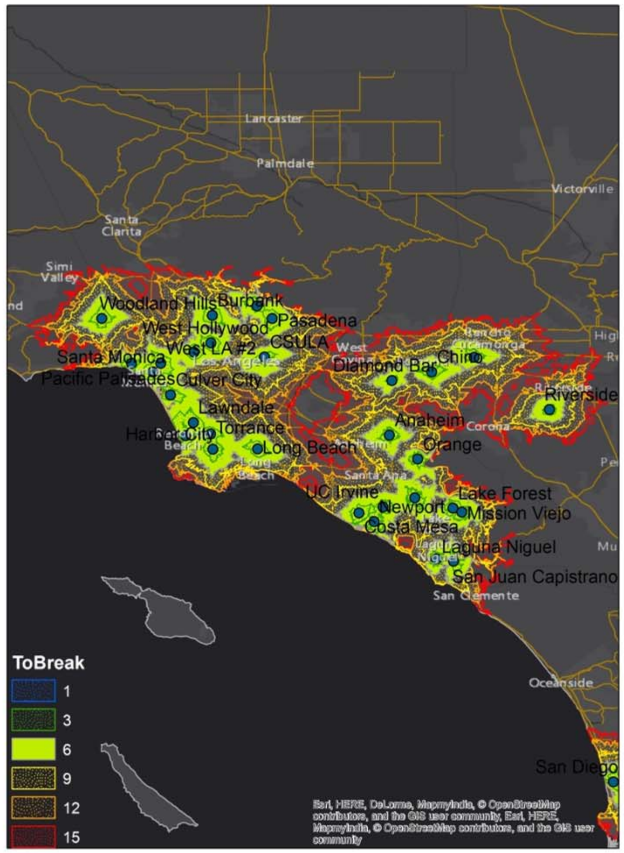
Roadway Data

Comparing Service Areas

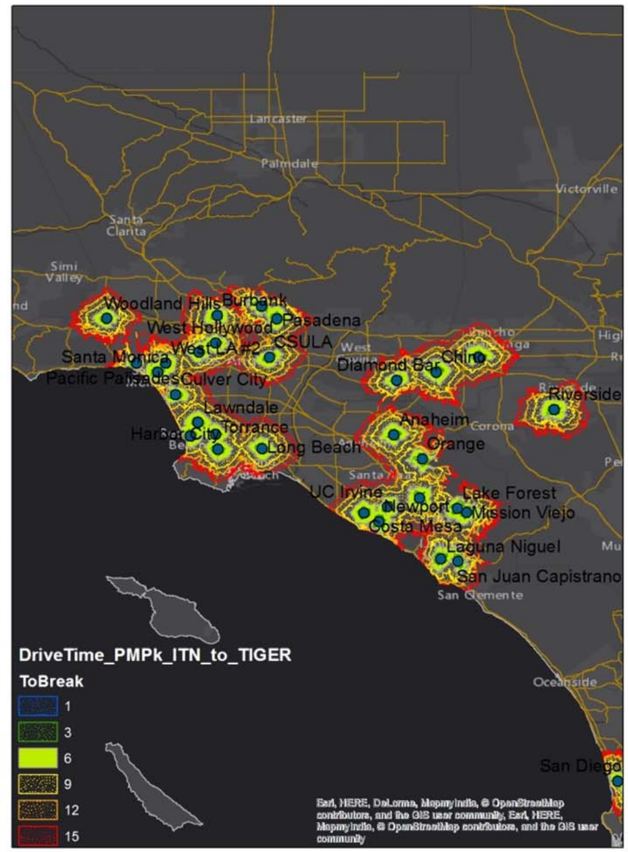
ITN



PEMS/TIGER



ITN/TIGER

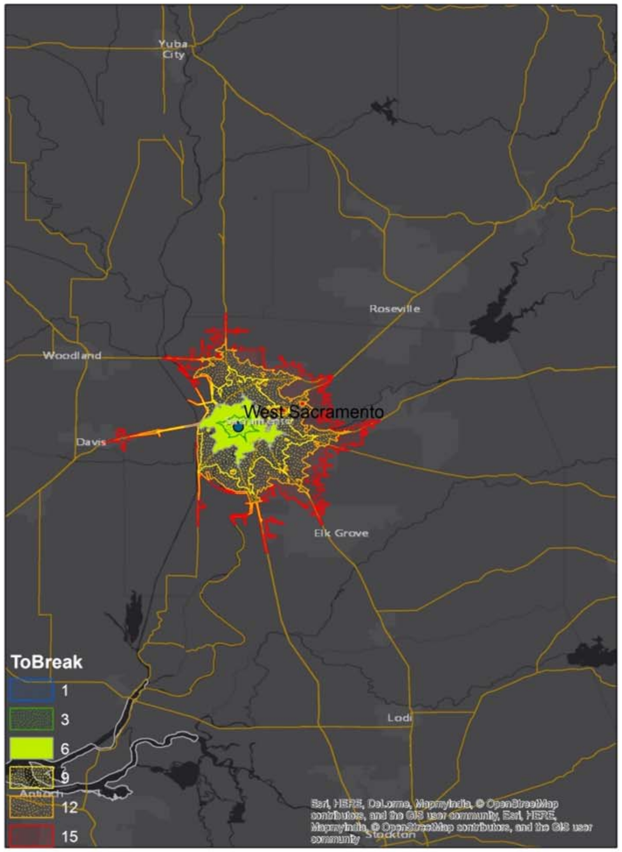


Note: "ITN" case from data in 2012 ITN without projecting onto TIGER, developed with data to 2010
"ITN/TIGER" case from 2015 ITN, developed with data to Dec. 2014

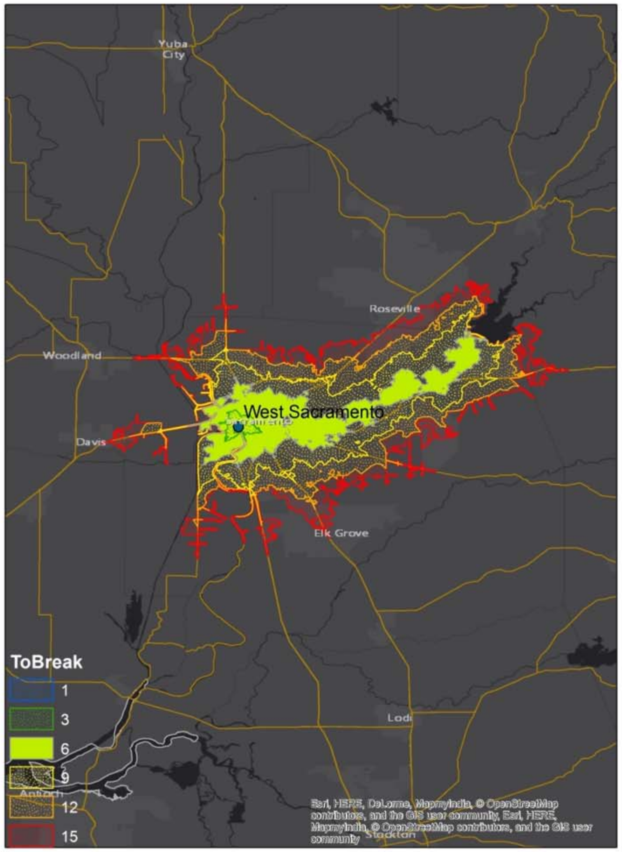
Roadway Data

Comparing Service Areas

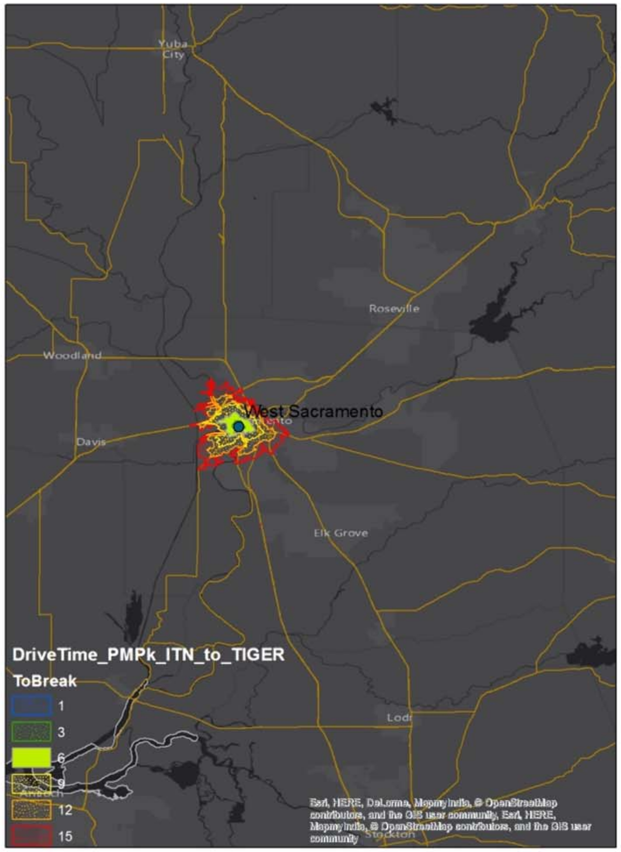
ITN



PEMS/TIGER



ITN/TIGER

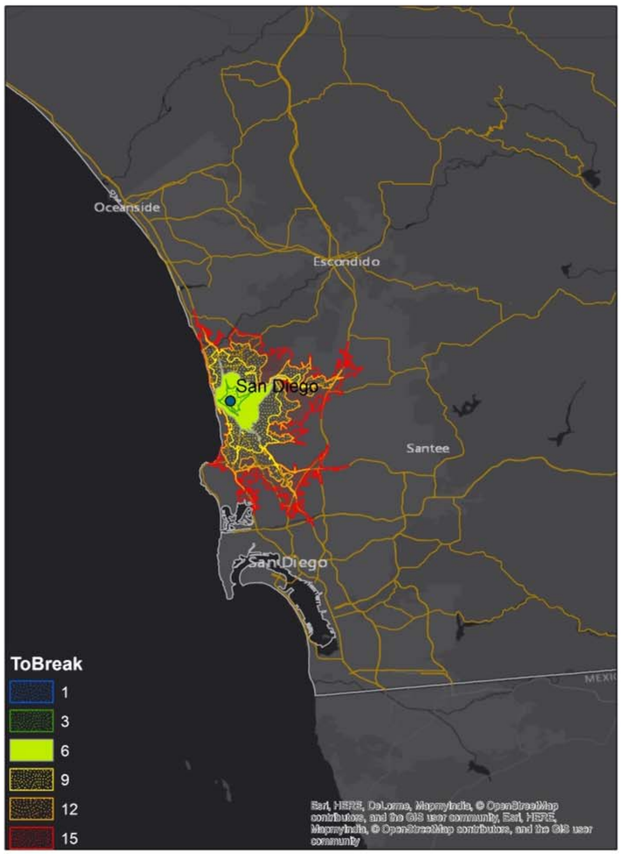


Note: "ITN" case from data in 2012 ITN without projecting onto TIGER, developed with data to 2010
"ITN/TIGER" case from 2015 ITN, developed with data to Dec. 2014

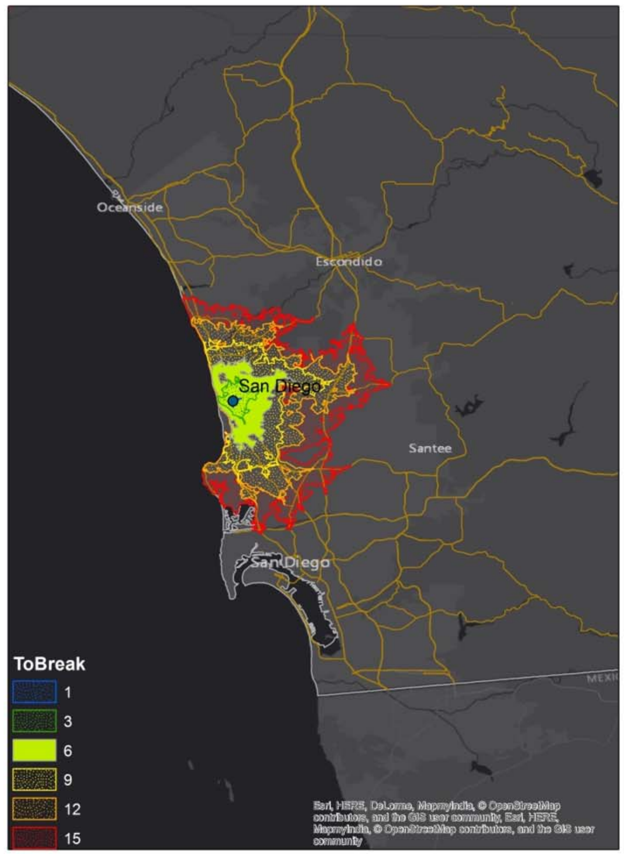
Roadway Data

Comparing Service Areas

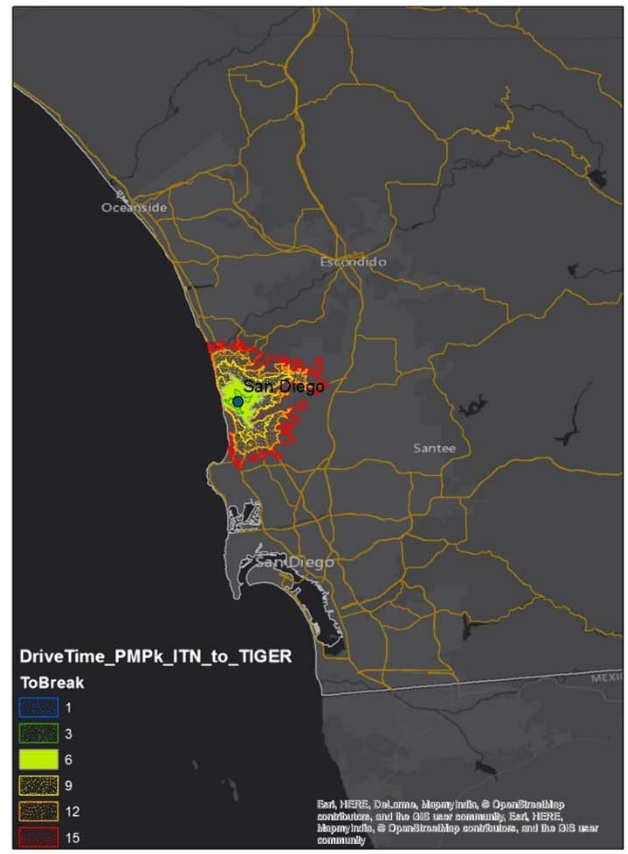
ITN



PEMS/TIGER



ITN/TIGER

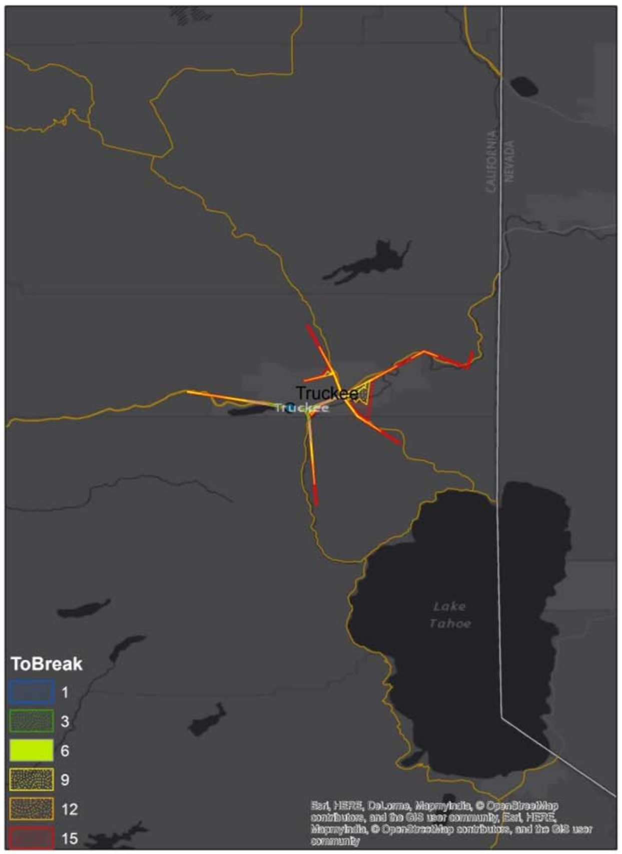


Note: "ITN" case from data in 2012 ITN without projecting onto TIGER, developed with data to 2010
"ITN/TIGER" case from 2015 ITN, developed with data to Dec. 2014

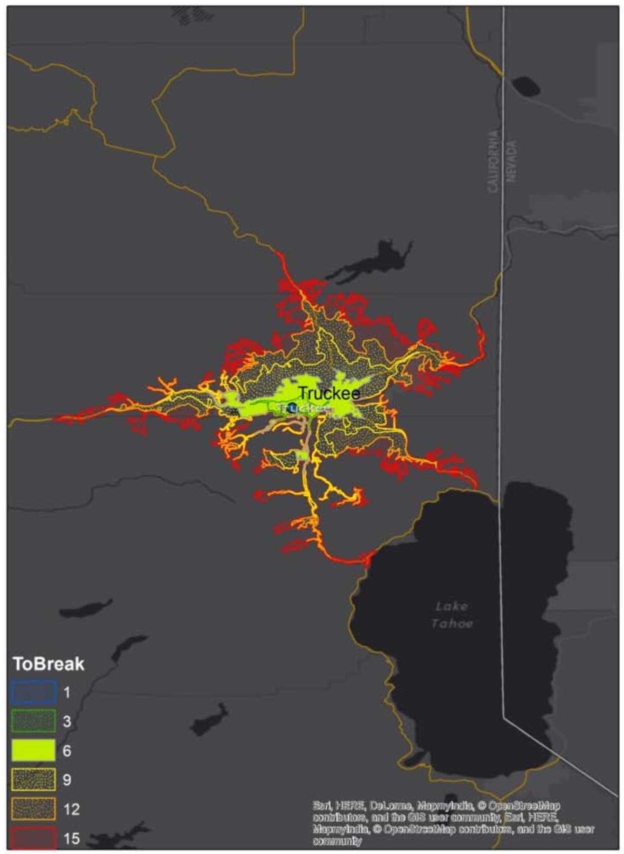
Roadway Data

Comparing Service Areas

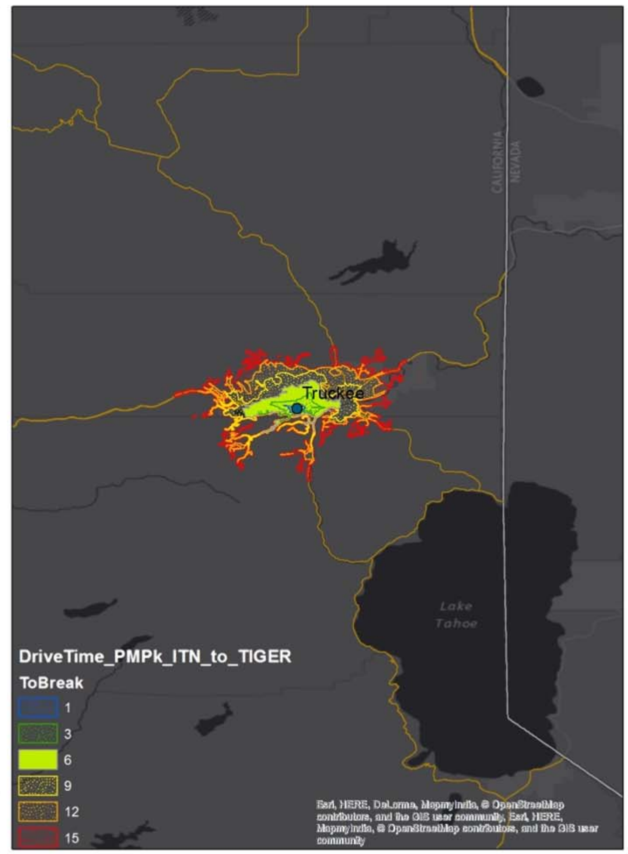
ITN



PEMS/TIGER



ITN/TIGER



Note: "ITN" case from data in 2012 ITN without projecting onto TIGER, developed with data to 2010
"ITN/TIGER" case from 2015 ITN, developed with data to Dec. 2014

Travel Time	Area Ratio (ITN/TIGER:PEMS/TIGER)
15 min	46%
12 min	43%
9 min	40%
6 min	36%
3 min	40%
1 min	59%

- Demonstrated significant difference in Service Area (and therefore coverage) between data set that relies on measured highway speeds only (PEMS) and a data set that relies on measured speeds on all road classifications (ITN)

Roadway Data

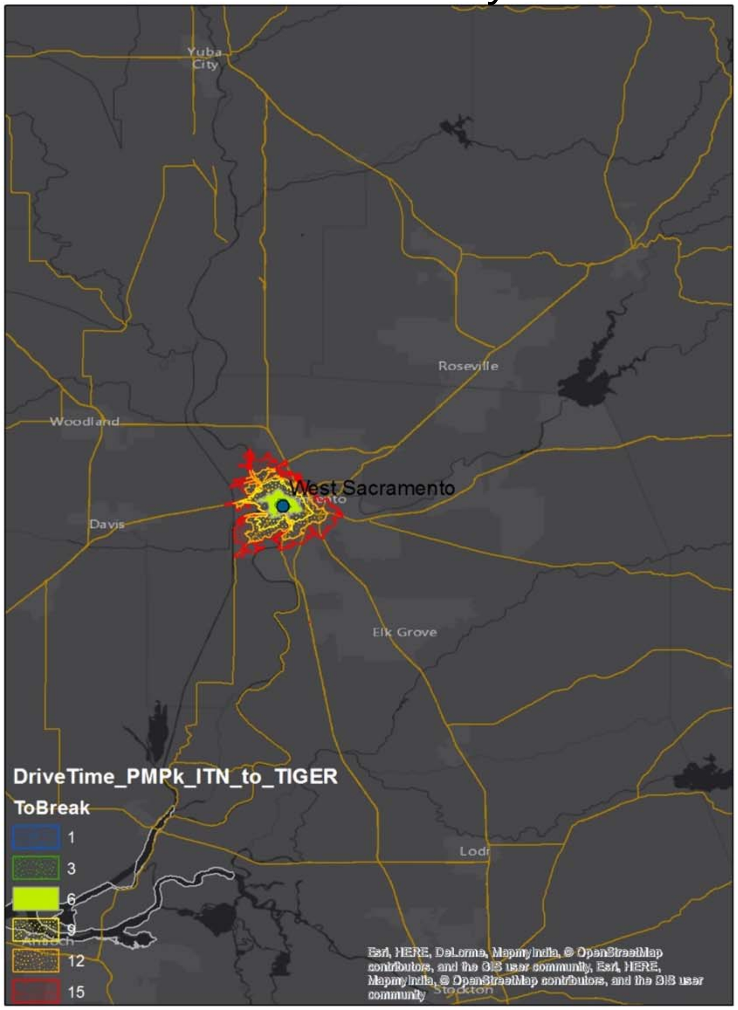
Turn Delays

- All results so far have not used turn delays
- Turn delays indicate the “wait” time to make a left, right, U-, etc… turn
- Can be a global rule applied to all turns of a type, or can make localized rules
- No apparent consensus in literature on “typical” values
  - Recent MS Thesis from UofRedlands: 2,5,3,5 (S,Rev,R,L)
  - Recent MS Thesis from NWMOState: 4,7,3,6
  - Mike Price (active Emergency Responder in GIS Community) Conference Paper: 1,30,2,4
  - All above for focused regional analyses and/or vehicle cases. Global turns will always be a rough assumption for a statewide network like ITN/TIGER
- Following are samples of results with 2,6,2.5,4.5 (median of above sets)

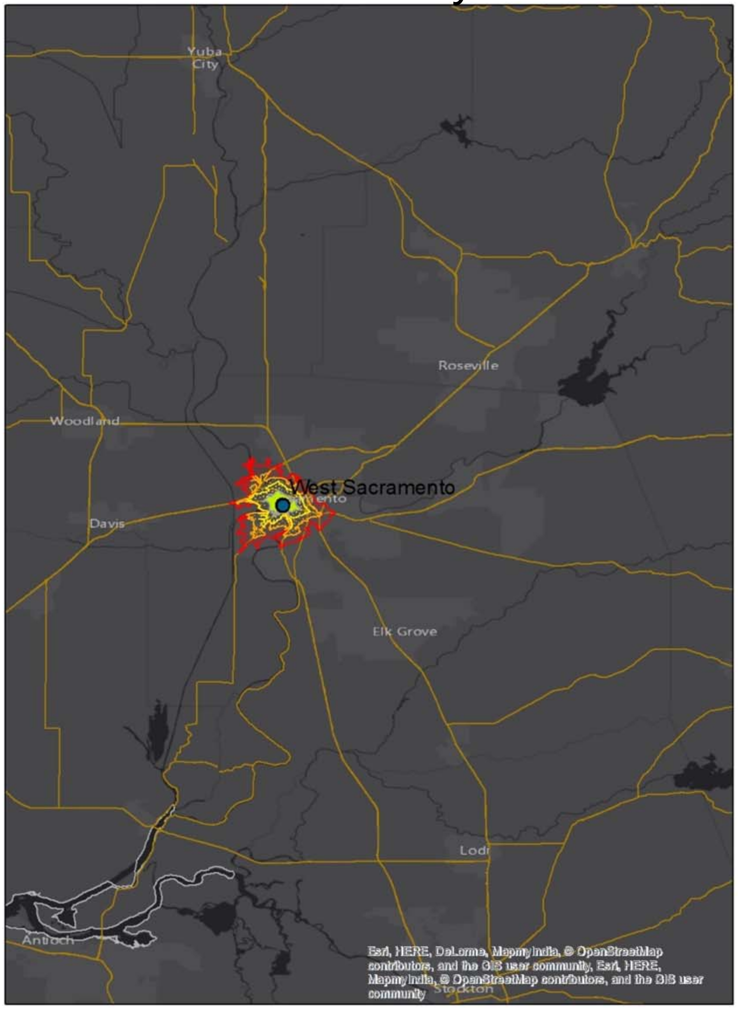
Roadway Data

Turn Delays

No Turn Delays



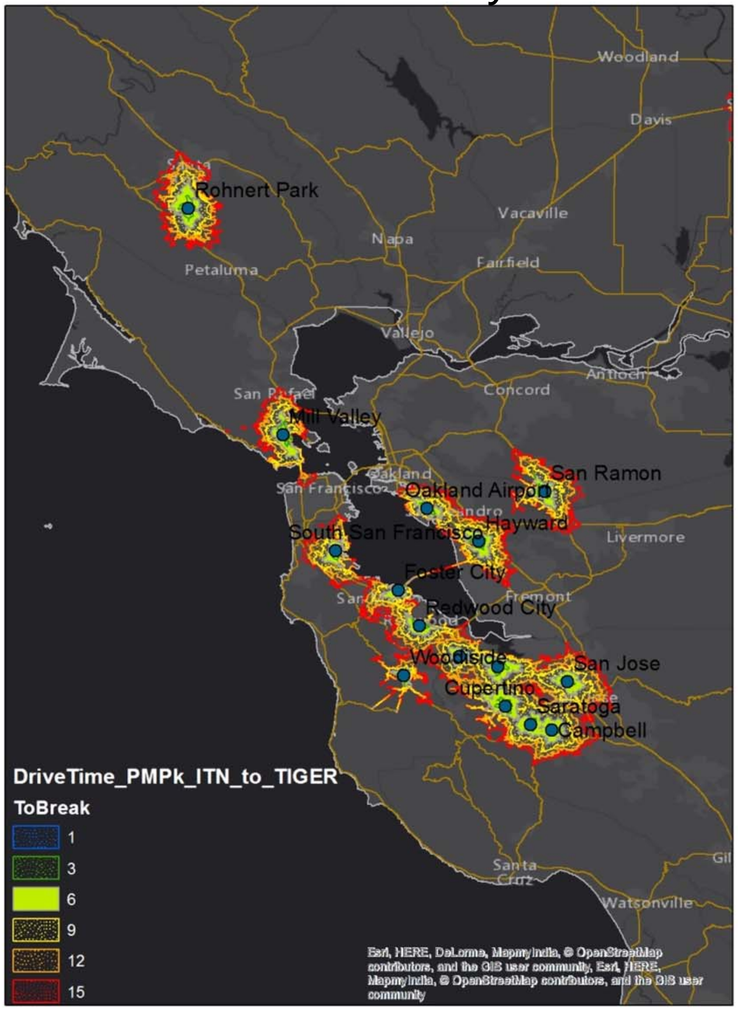
Turn Delays



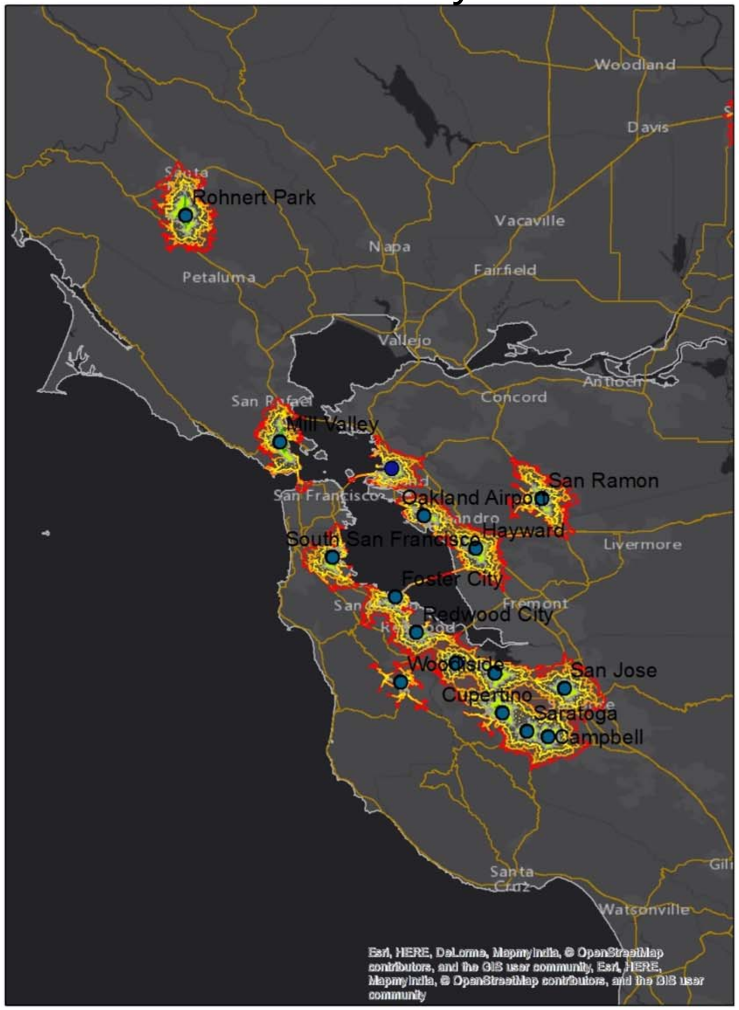
Roadway Data

Turn Delays

No Turn Delays



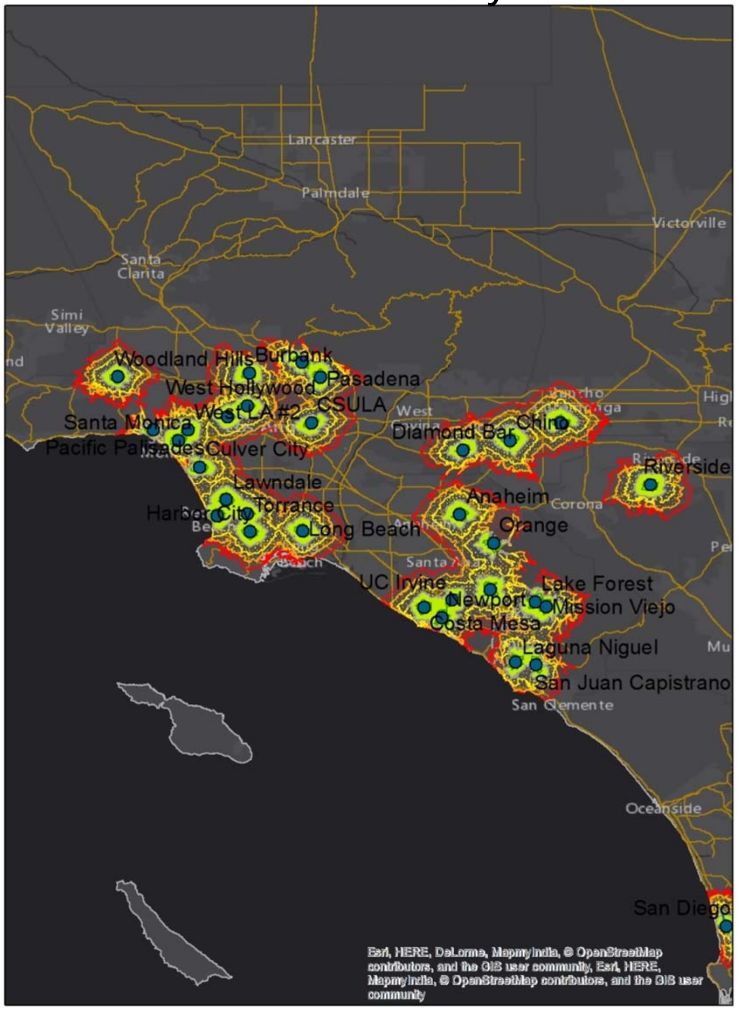
Turn Delays



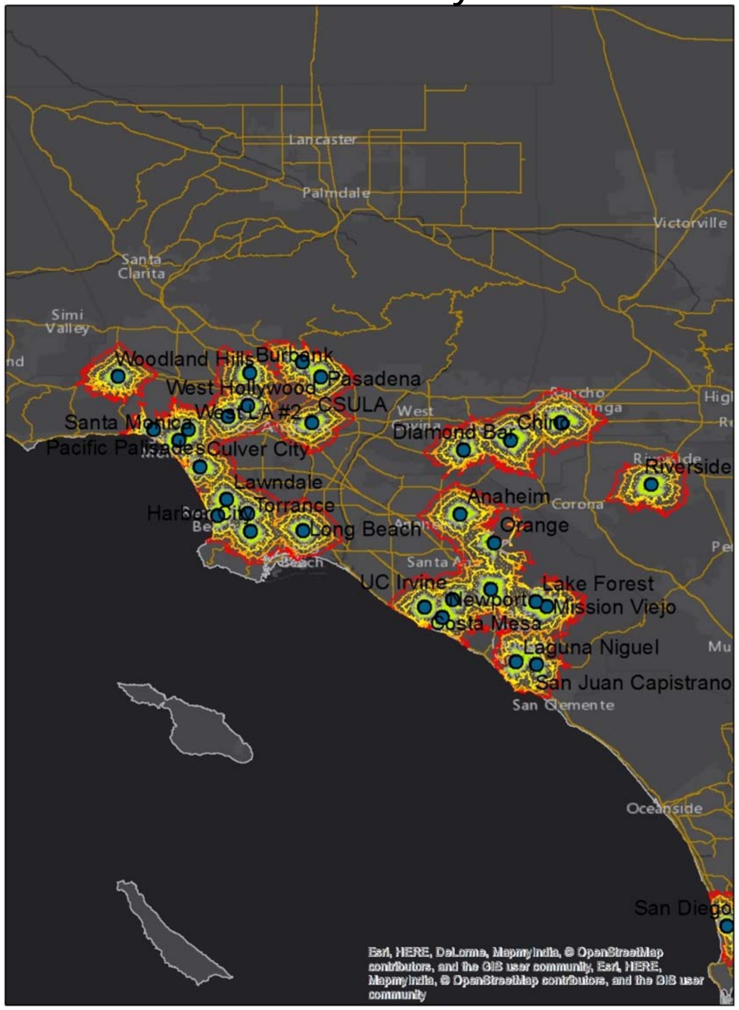
Roadway Data

Turn Delays

No Turn Delays



Turn Delays



Travel Time	Area Ratio (with:without)
15 min	86%
12 min	84%
9 min	80%
6 min	76%
3 min	70%
1 min	59%

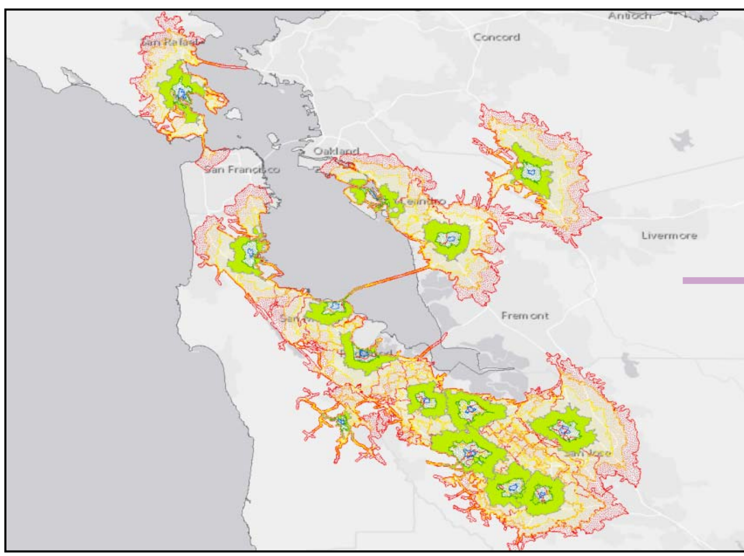
- Difference does exist between Service Areas, with increasing effect for smaller drive times
- However, this was a simple case study based on combining data from unrelated studies on highly specific regions
- Unable to identify a proven generalized rule for global turns, so did not implement

COVERAGE ALGORITHMS

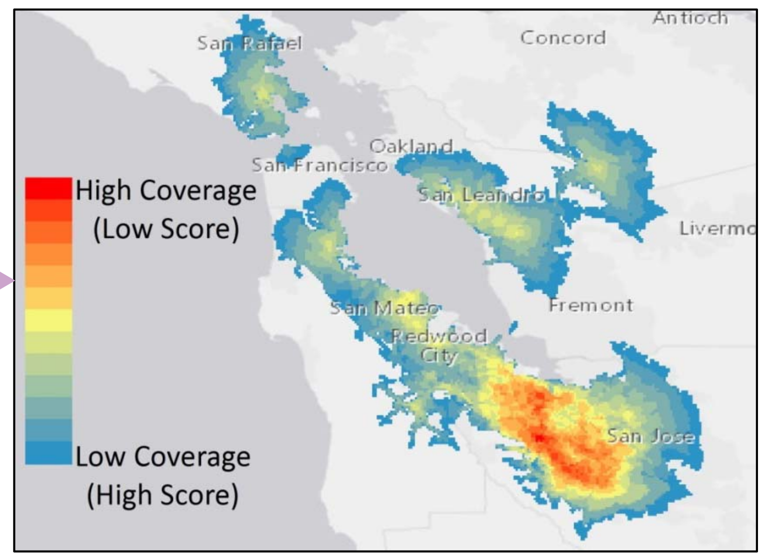
Coverage Algorithms

- Development of ITN/TIGER roadway dataset enables the estimation of Service Areas
- Service Areas form the basis of representation and estimation of coverage provided stations
- Can additionally be utilized to estimate potential coverage

Service Area Calculations

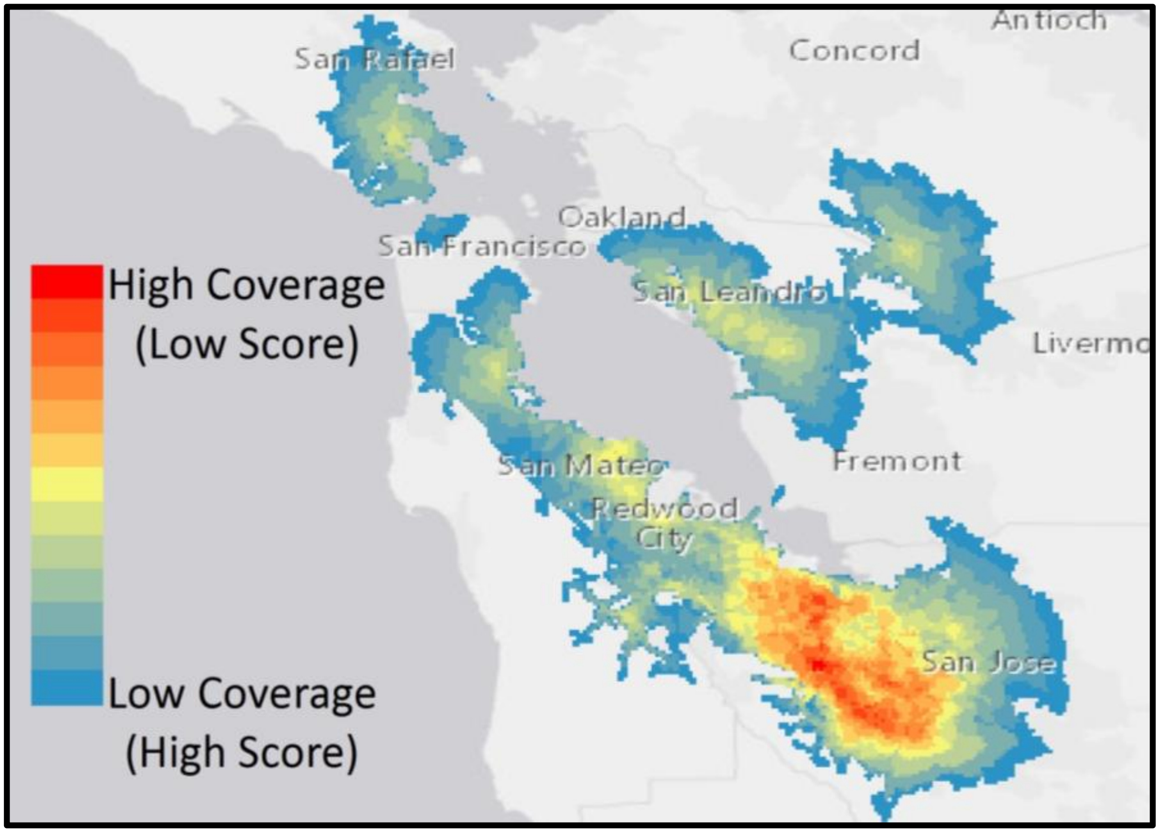


Coverage Estimation



Coverage Algorithms

Goals for Analyzing Existing Coverage

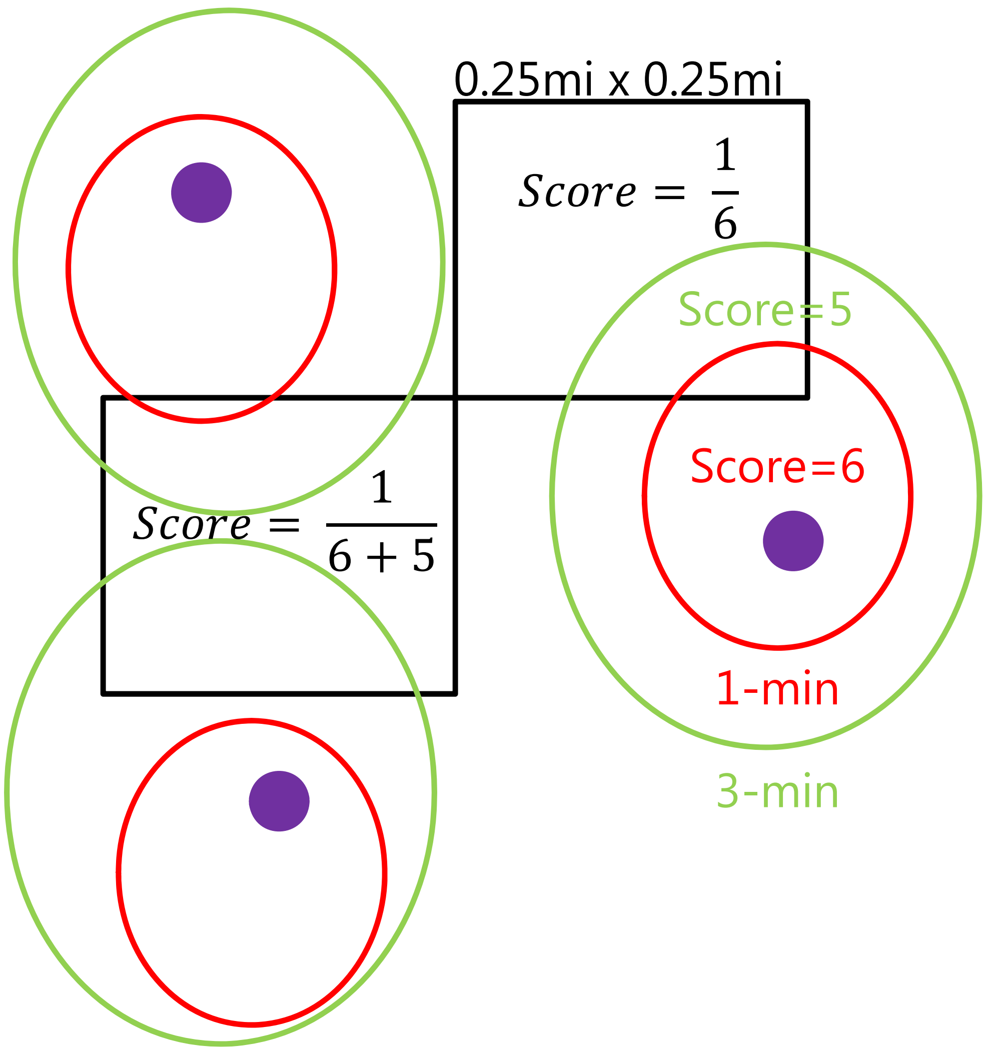


*Areas without coverage have no color and score highest

- Provides an estimate of coverage that is more informative than a binary yes/no, allowing for estimation of degrees of coverage
- Estimates combined coverage provided by multiple stations that may be reachable within various drive times

Coverage Algorithms

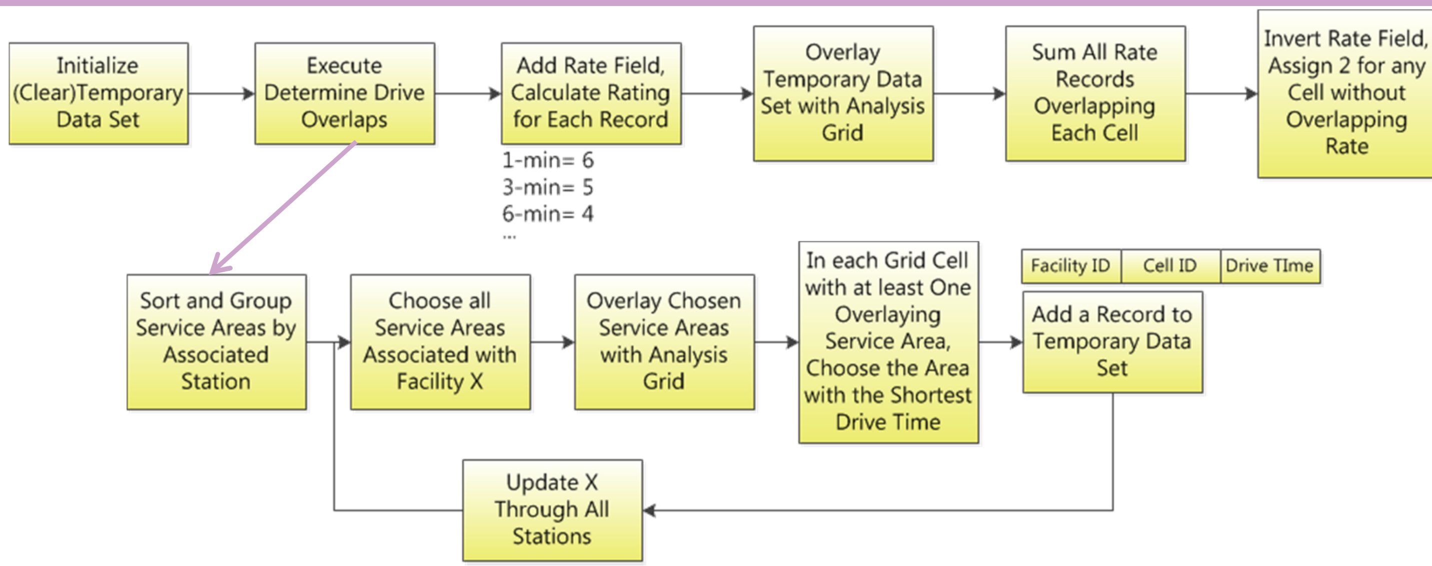
Existing Coverage Factor Estimate

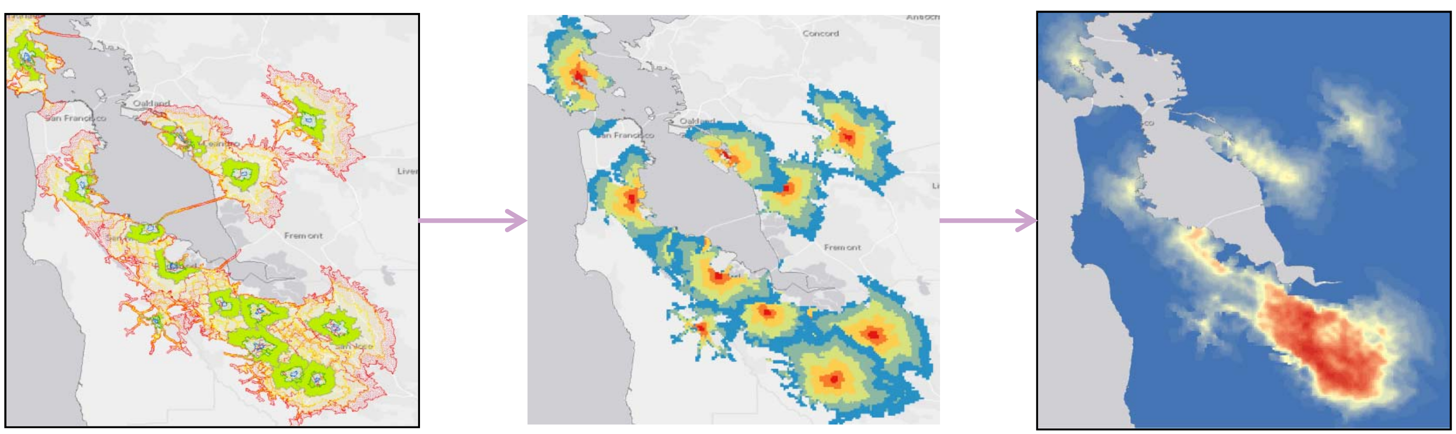


- Need to assess current stations' coverage
- Want to account for multiple overlapping service areas at a given "point"
- Cell resolution of 0.25-miles on each side large enough to include multiple service area coverages from each of multiple stations
- Assign shortest overlapping service area from each overlapping station to cell
- Shorter drive times assigned higher score (1-6)
- Score for cell is inverse of sum of overlapping values

Coverage Algorithms

Implementation in CHIT



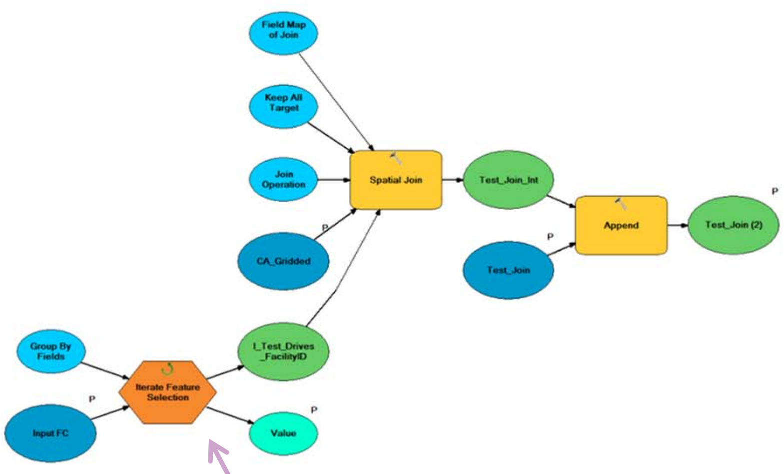


*Note: Full data not shown; many features overlap

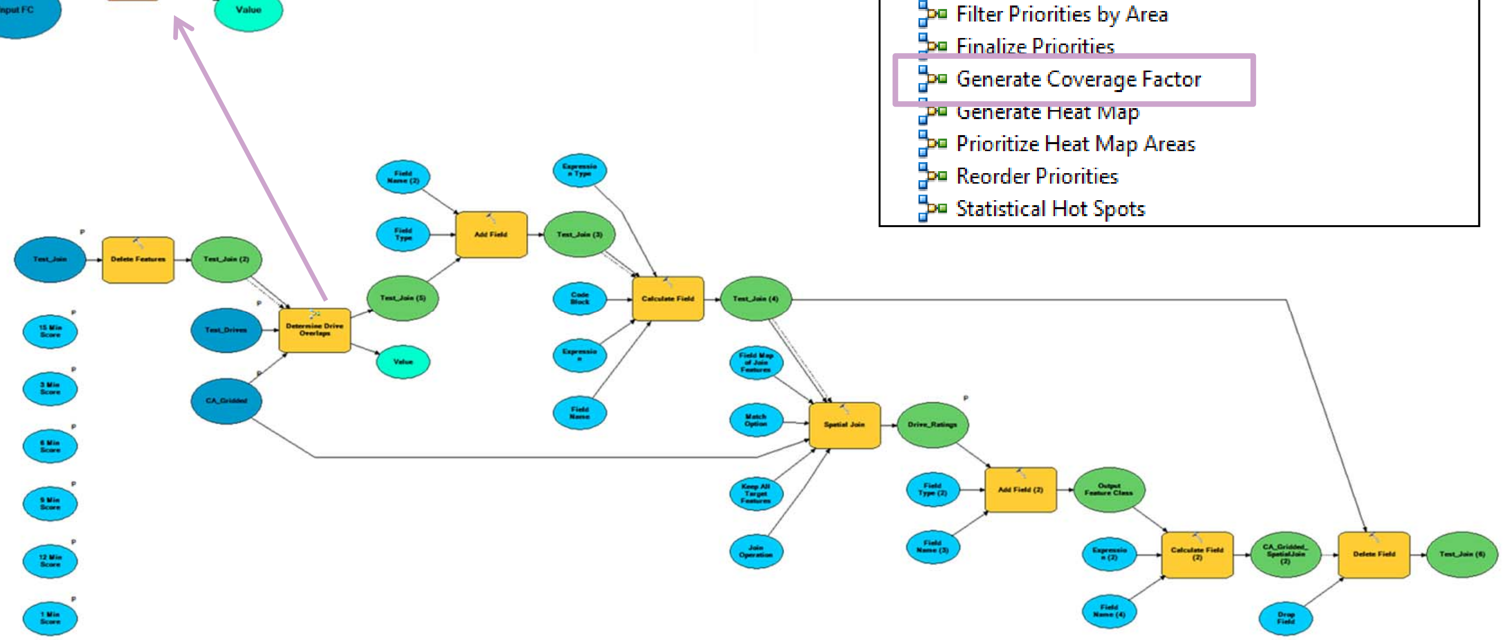
Coverage Algorithms

Implementation in ArcGIS

- More stations and closer stations mean higher coverage (lower coverage gap score)



CHIT_Tools
- Determine Drive Overlaps
- Determine Station Counts 6-Factor Match_Norm
- Determine Station Counts 6-Factor Match_Norm_Alt
- Filter Priorities by Area
- Finalize Priorities
- Generate Coverage Factor
- Generate Heat Map
- Prioritize Heat Map Areas
- Reorder Priorities
- Statistical Hot Spots



Coverage Algorithm

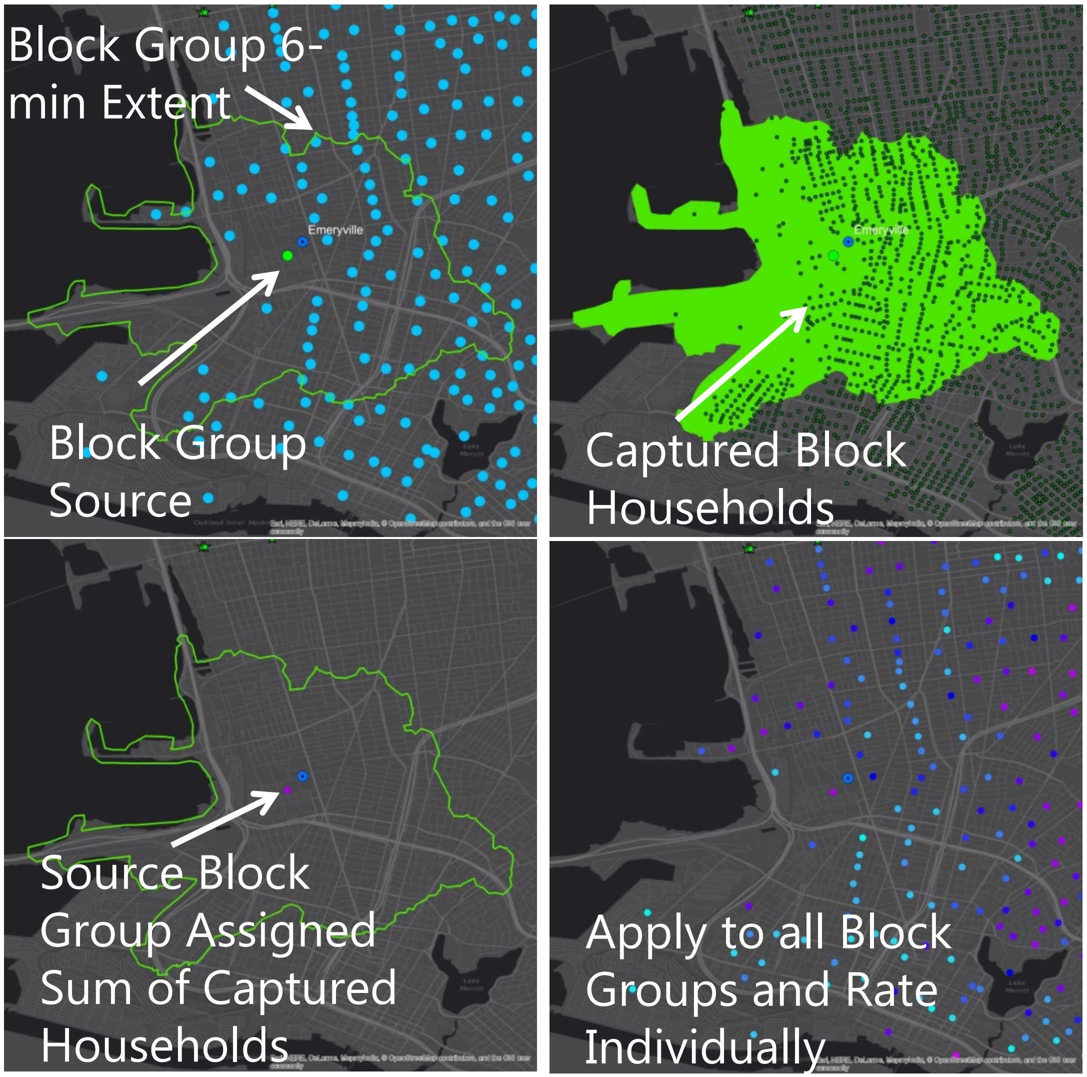
Potential Coverage Factor

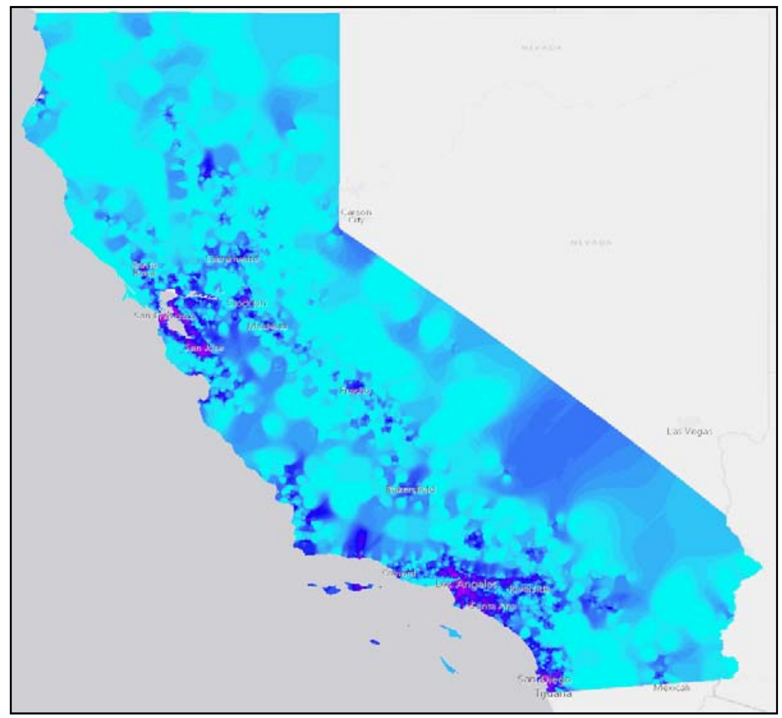
- In addition to Existing Coverage, need to assess the coverage that could be provided by placing a station in a new location
- Require a method that could be applied statewide
- Potential Coverage treated as an input that does not change as often as Existing Coverage
- Formulated a method to answer the question

  "As a potential station site (here modeled as centers of block groups), how many households will have access to me within x minutes (here 6 minutes)?"
- Utilized block group centers as relatively high-resolution set of theoretical potential locations
- Block groups are not uniformly distributed
  - Used interpolation method to fill in between block group centers
  - Investigated interpolation methods similar to ITN/TIGER

Coverage Algorithms

Potential Coverage Factor





TRAFFIC VOLUME ASSESSMENT

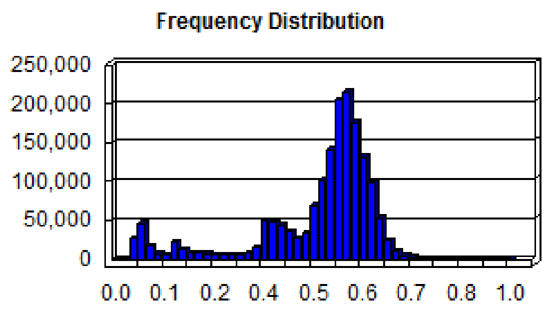
Traffic Volume

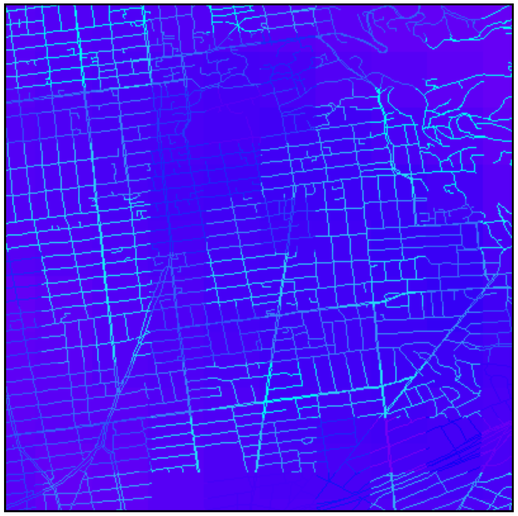
- Principle #6 (convenience is provided by stations within short drive times of vehicle owner homes) requires only on-the-road speeds for Service Areas and estimation of coverage
- An alternative to this principle is that convenience is defined by short deviation from commonly-traveled routes
- Explicit and direct consideration requires origin-destination data sets
- ARB did not identify a dataset sufficiently covering entire state
- As an alternative, investigated utilizing volume data available in ITN
- Followed method similar to speed to define, at all points in the state, the ratio of AM and PM peak travel volumes to the daily average
  - Assumed a road that is a main commuter line would have a large ratio of peak to average volume

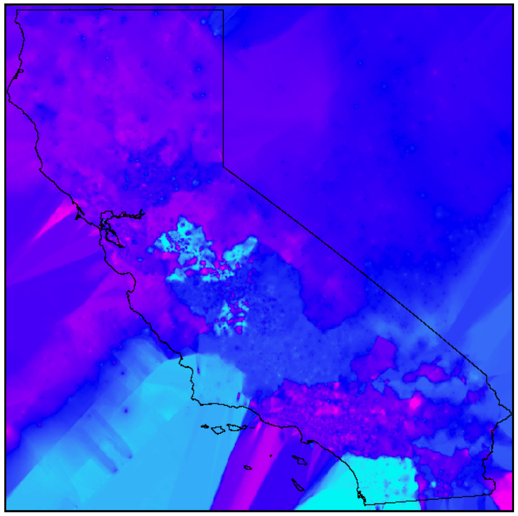
Traffic Volume

Overall Assessment of Volume in ITN/TIGER

- Statewide, volume ratio does display variation that could be used to identify more viable areas
- Some areas show similar variation in smaller scales

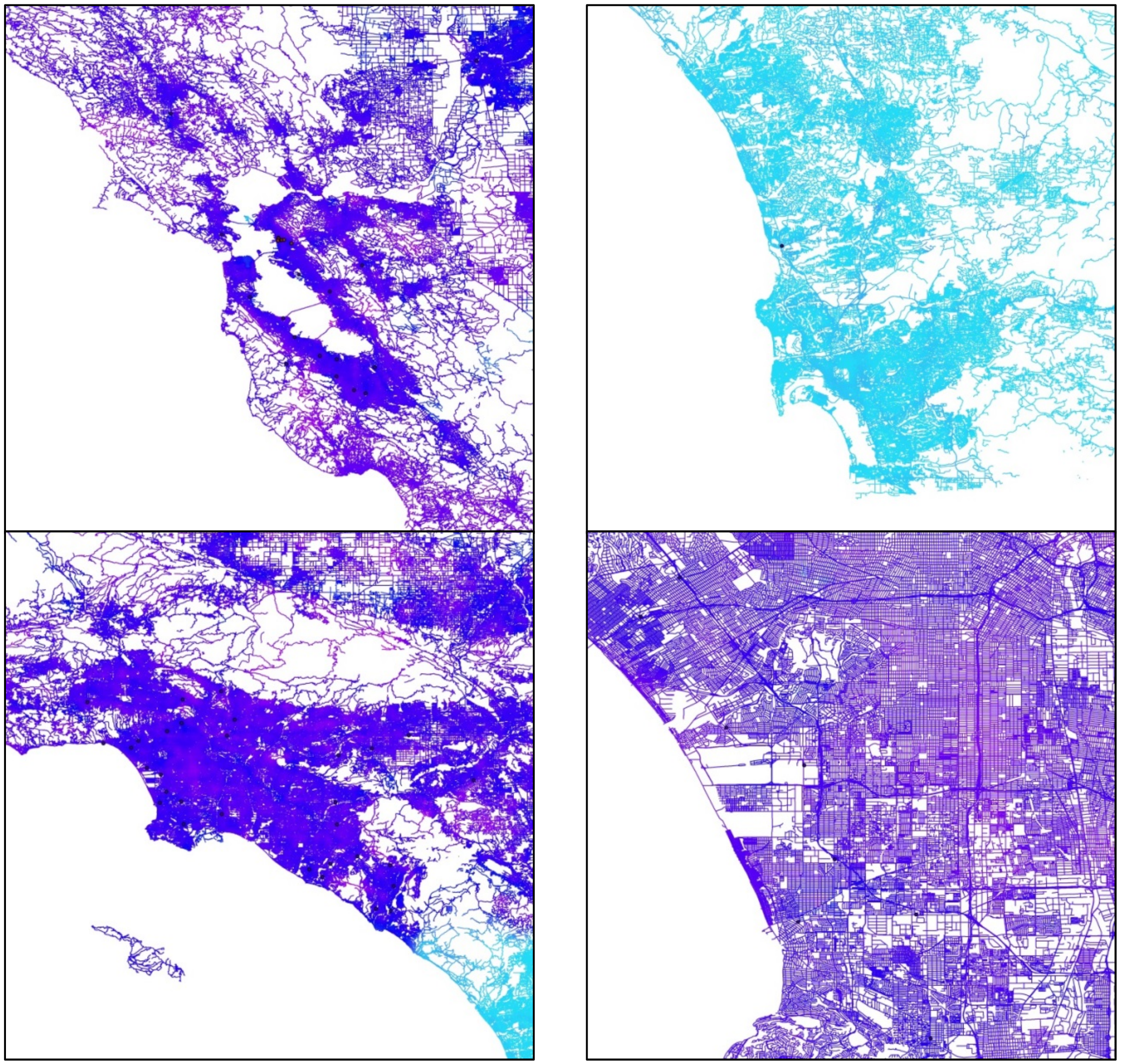






Traffic Volume

Regional Samples of ITN/TIGER

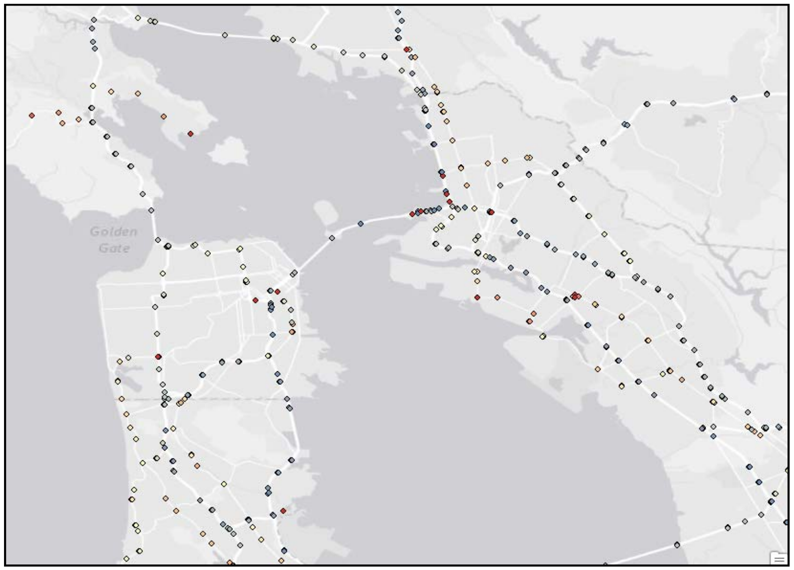


- Many areas do not show much variation in ratio of volume
- Volume ratio not usable across state as appropriate indicator
- Does not indicate traffic volume or Origin-Destination as an indicator are invalid
- Using volume in this manner did not prove viable; if fully-detailed Origin-Destination data become available, could be investigated

Traffic Volume

Other Data Sources

- Have investigated other data sources, but did not find any that included sufficient data density and coverage to be applied equivalently across the entire state
  - California Household Travel Survey
  - CADOT Annual Average Daily Traffic
  - CHAPIS
- Principle #5 often not fulfilled: need data on neighborhood-level streets for accurate assessment of convenience and coverage



FIRST ADOPTER FCEV MARKET ASSESSMENT

Market Assessment

- CHIT utilizes a number of demographic-based indicators for determining areas where FCEV early market adopters are likely to be located

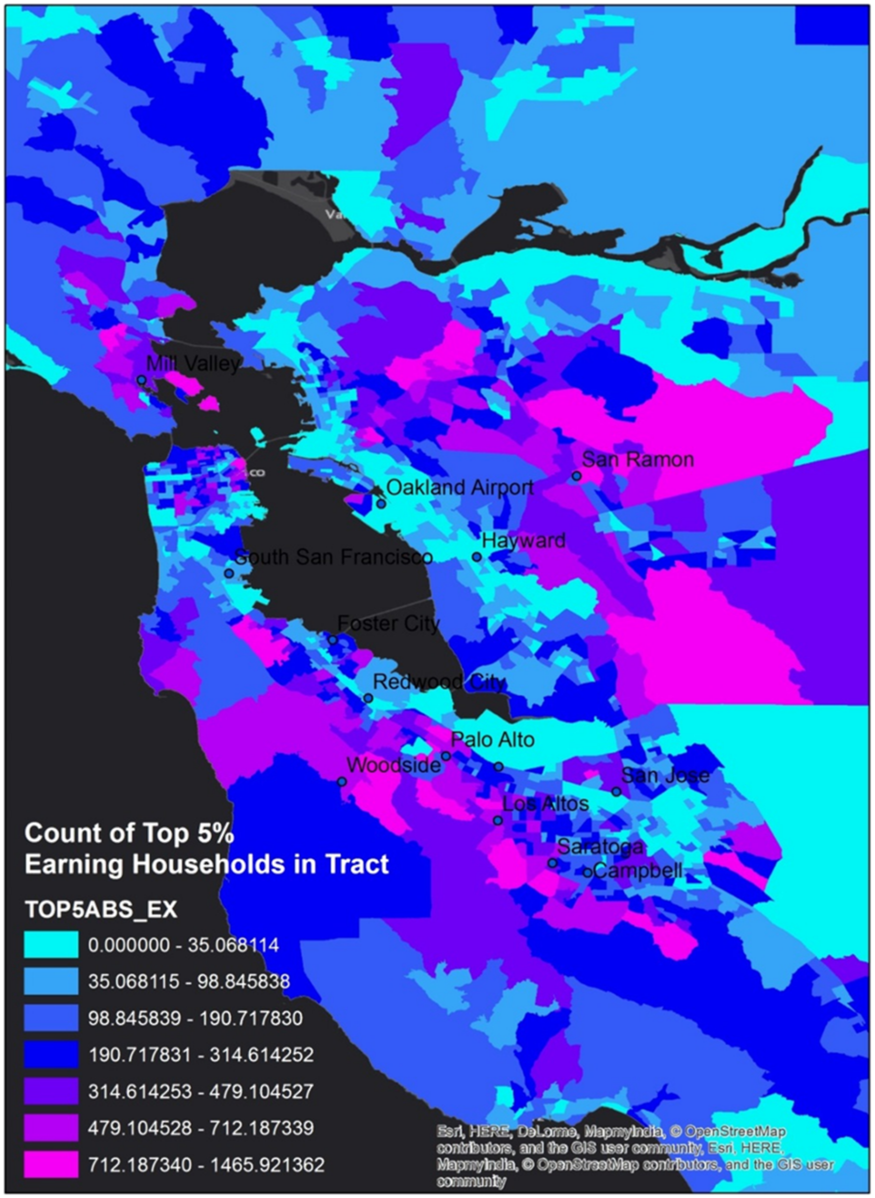
- Three main categories are incorporated:
  - Financial indicators
  - Green vehicle adoption indicator
  - Educational attainment indicator

- Important to keep in mind that overlaying indicators is not equivalent to locating all households that individually meet all attributes

  CHIT does not identify the locations of all households that meet an income threshold, own a certain number of green vehicles all of given makes, and have at least one household member with a certain degree

Market Assessment

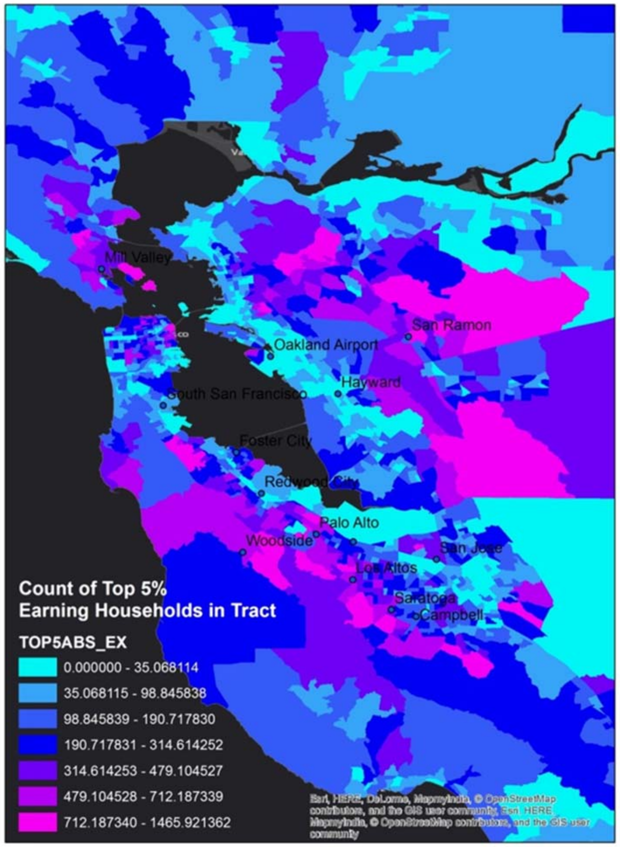
- Incorporated factors appearing in literature and from discussions with stakeholders
- Geographical resolution based on data source
  - Income (Census- Tract): count of households at or above mean for top 20% households
  - Vehicle Data (DMV- ZIP):
    - Luxury: Select brands, previous 5 years
    - MSRP: Hi/Lo Range, previous 5 years
    - Green: PHEV & HEV, up to first 7 years of segment
  - Education (Census- Tract): count of post-graduate degrees



Market Assessment

Income Factor

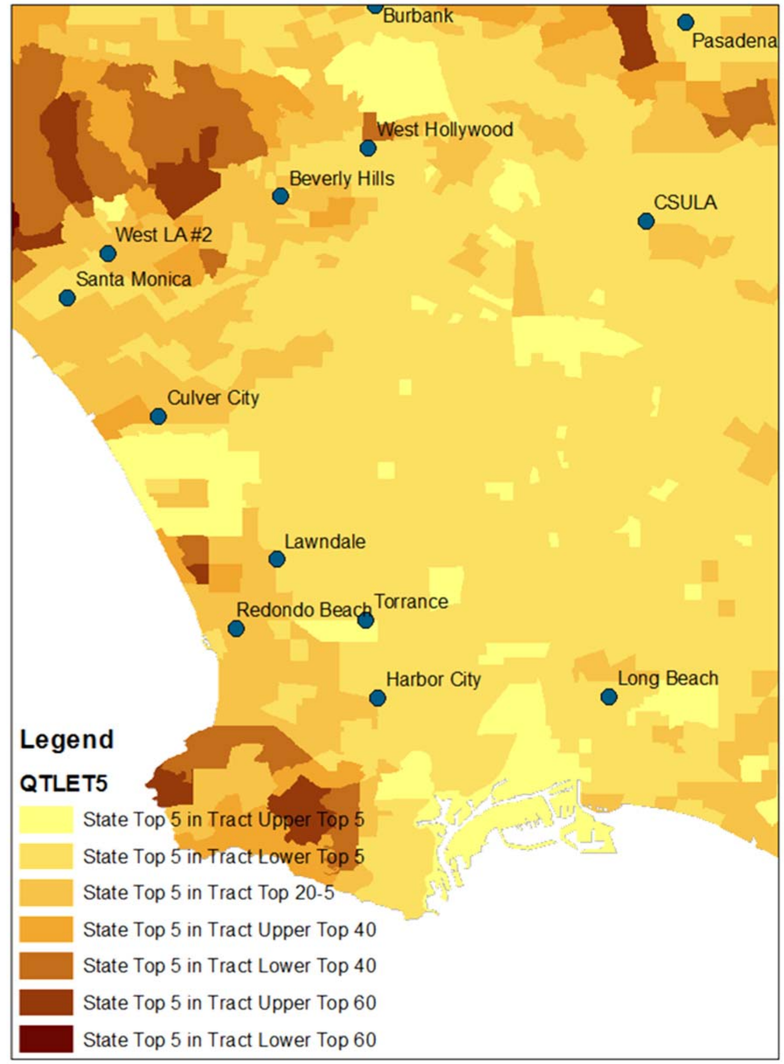
- Anticipate first market vehicles will have a retail cost premium; first adopters likely to be in high-earning households
- Based on statewide income data from ACS 2013 5-year estimates
  - Top 20% of households in 2014 earn $230k on average
  - Alternatively, top 5% earn $230k+; easier to implement as validation on any interpolation or extrapolation
- ACS data present a challenge: tract-level income distribution data saturate at $200k
- Considered a number of methods to approximate where these households are concentrated
  - Rating based on tract-level quantile containing threshold
  - Modeling tract-level income as lognormal and interpolating/ extrapolating count
  - Modeling tract-level income as exponential and interpolating/extrapolating count



Market Assessment

Income Factor

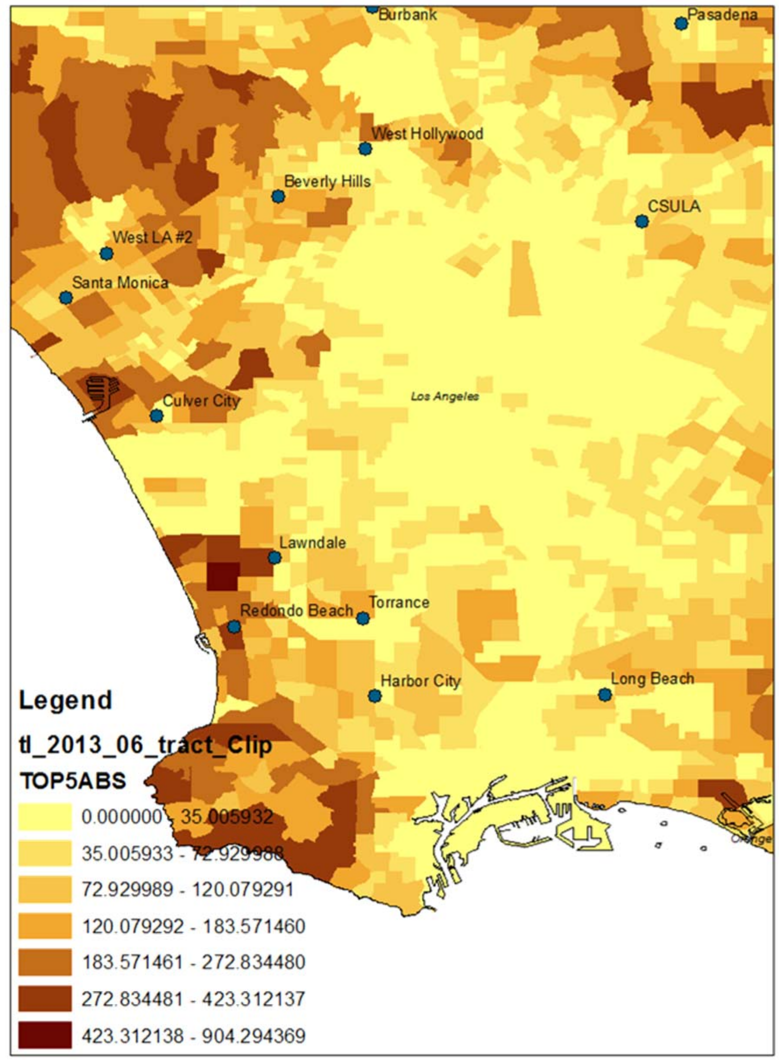
- Assuming distribution on tract level cannot be reconstructed, can compare $230k threshold to quantile cutoffs and means in quantiles on tract level
- Provides a means of directly using ACS data without extrapolation
- Inherently qualitative: requires development of a rating system to implement in a numerical assessment



Market Assessment

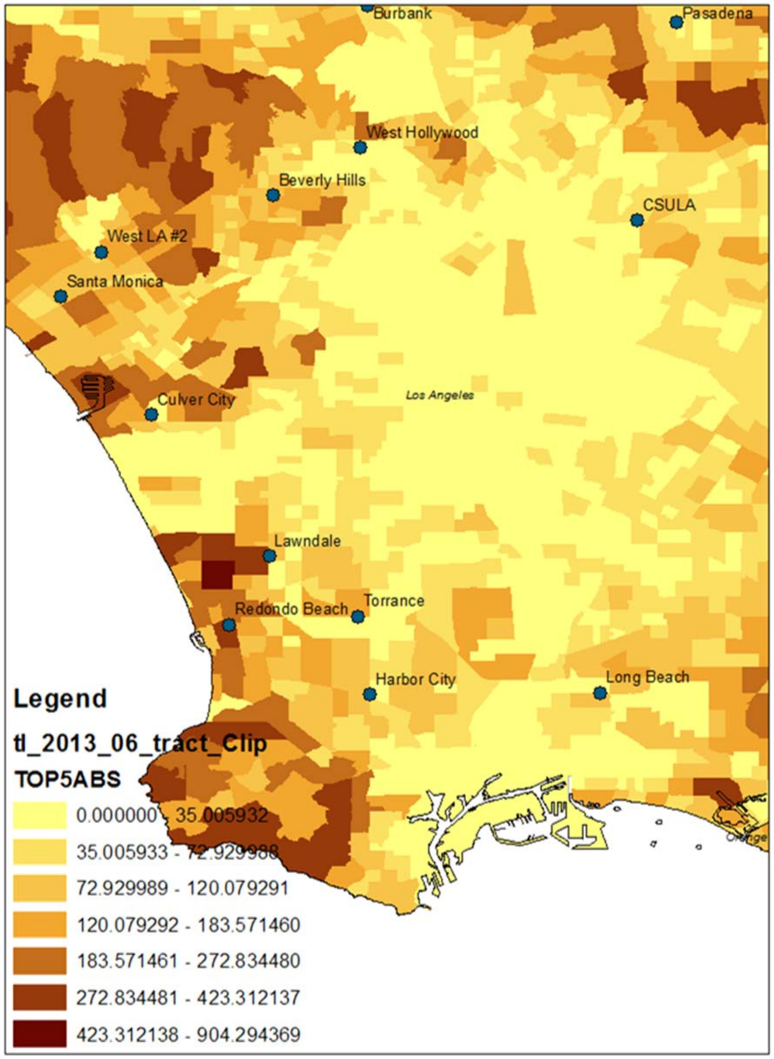
Income Factor

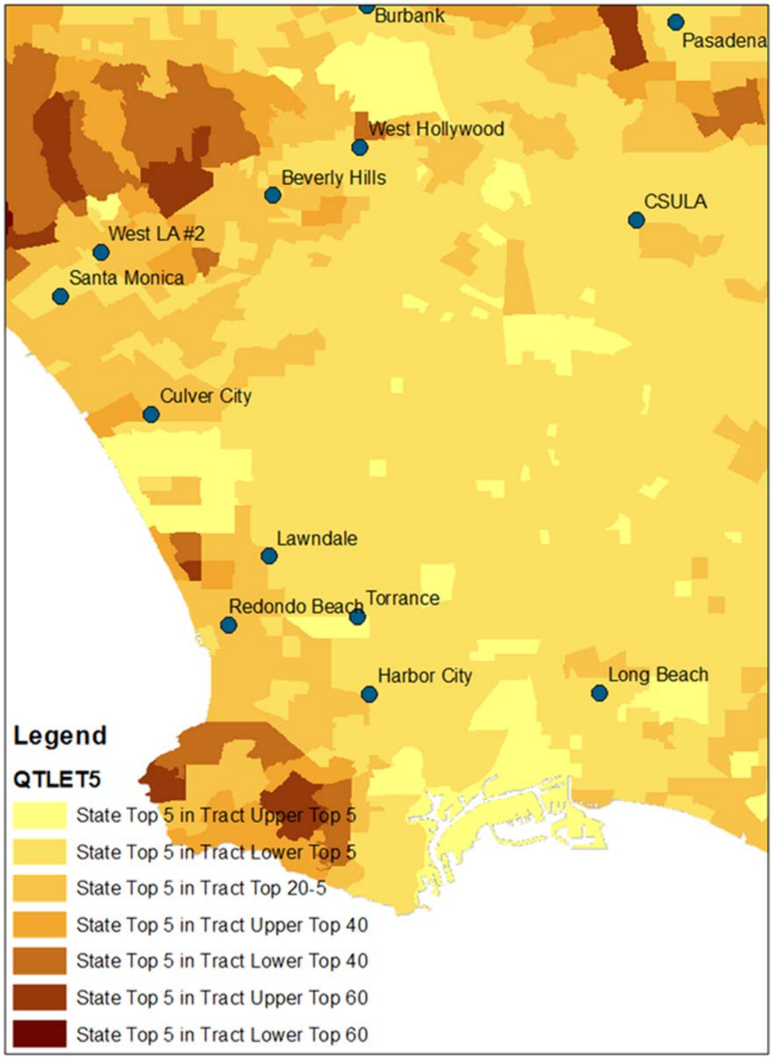
- With summary statistic data available from ACS, it may be possible to reconstruct the shape of the income distributions on tract level
- Statistical inference method based on lognormal (used in many literature studies) assumption, median, and margin of error on tract level
- Estimating number of households (HH) in each tract above the $230k threshold
- Arrived at ~5.7% of HHs after calculation
- ACS documentation is unclear on formulation of reported medians and margins of error



Market Assessment

Income Factor





Market Assessment

Income Factor

- Identified recent works indicating that income distribution at high incomes is exponential, i.e. income distribution is two different models based on cutoff*
- Additionally moved to using empirical survival function (inverse of a cumulative distribution) in place of attempting to fit a model with a median and margin of error
- Survival function based on proportion tables of population in income brackets and exponential distribution assumption
  - Fit top 3 data points for each tract to an exponential survival function to extrapolate proportion above $230k

Subject	Households	
	Estimate	Margin of Error
Total	115,610,216	+/-238,223
Less than $10,000	7.2%	+/-0.1
$10,000 to $14,999	5.4%	+/-0.1
$15,000 to $24,999	10.8%	+/-0.1
$25,000 to $34,999	10.3%	+/-0.1
$35,000 to $49,999	13.6%	+/-0.1
$50,000 to $74,999	17.9%	+/-0.1
$75,000 to $99,999	12.2%	+/-0.1
$100,000 to $149,999	12.9%	+/-0.1
$150,000 to $199,999	4.9%	+/-0.1
$200,000 or more	4.8%	+/-0.1

*Thanks to Jeff Austin for guidance and collaboration on developing the method

Market Assessment

Income Factor

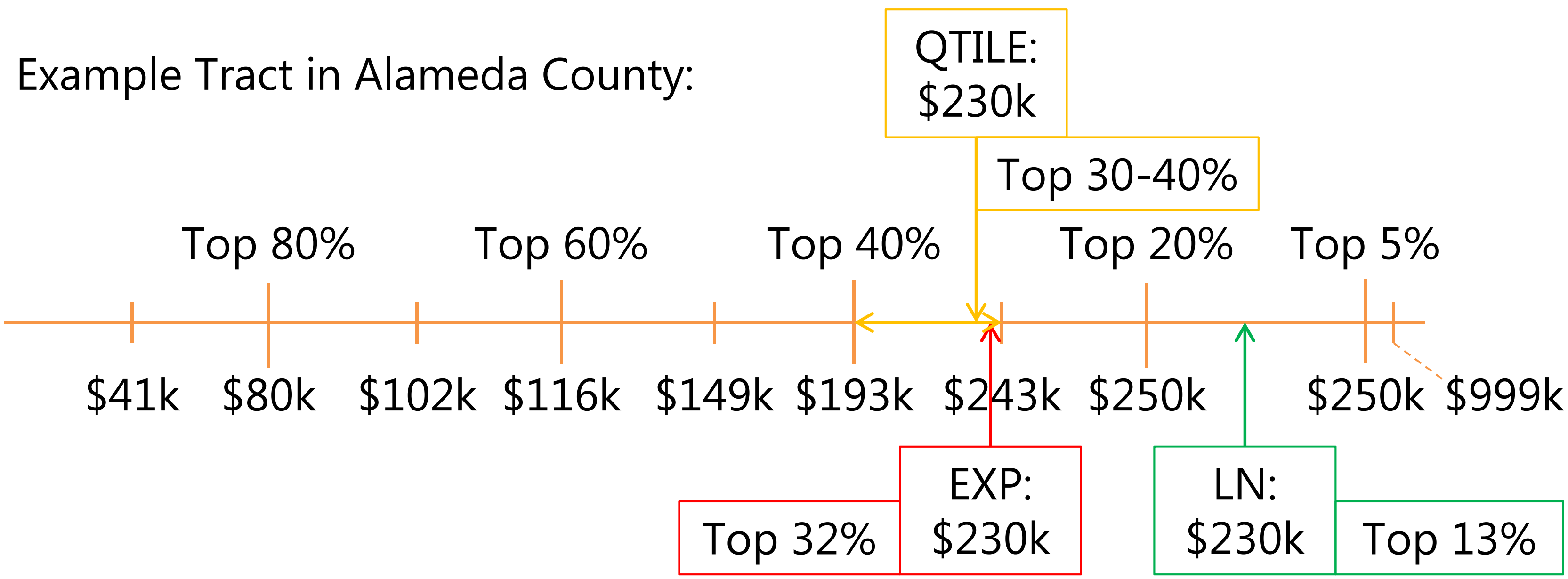
- Exponential modeling method provides 5.1% of households in top 5% when estimates summed over all tracts
- Exponential tract-level estimates range from 0 to 58%
- Lognormal estimates range from 0 to 30%
- Quantile (QTILE) data indicates range of 0 to 60%

Market Assessment

Income Factor

- To evaluate which may be more appropriate, compared to quantile measure
  - From either statistical model method, know a precise estimate for the proportion within the tract of the target population. (i.e. in a given tract, 10.5% of the households are in the top 5% of earners statewide)
  - From tract-level quantile data and means in quantiles, can estimate lower limit of quantile that the target value falls within, (i.e. if the target value is above the cutoff of top 20% of earners in the tract, it can be estimated that at least 20% of that tract's households are in the top 5% of earners statewide)

Example Tract in Alameda County:

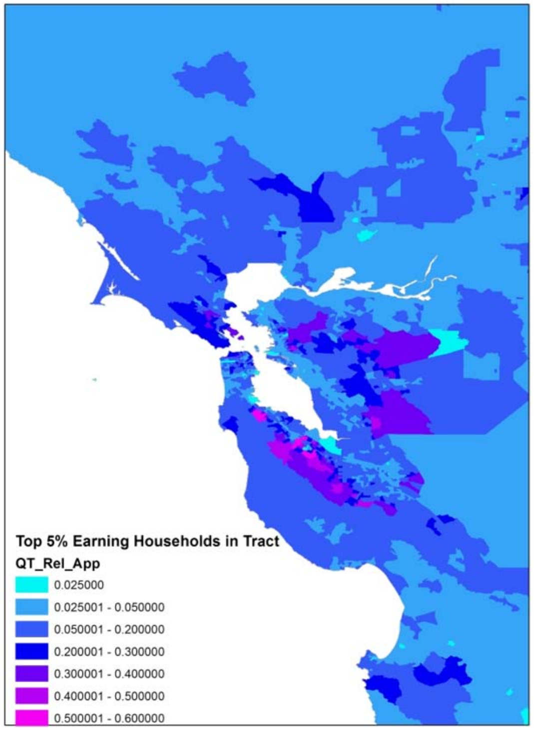


Market Assessment

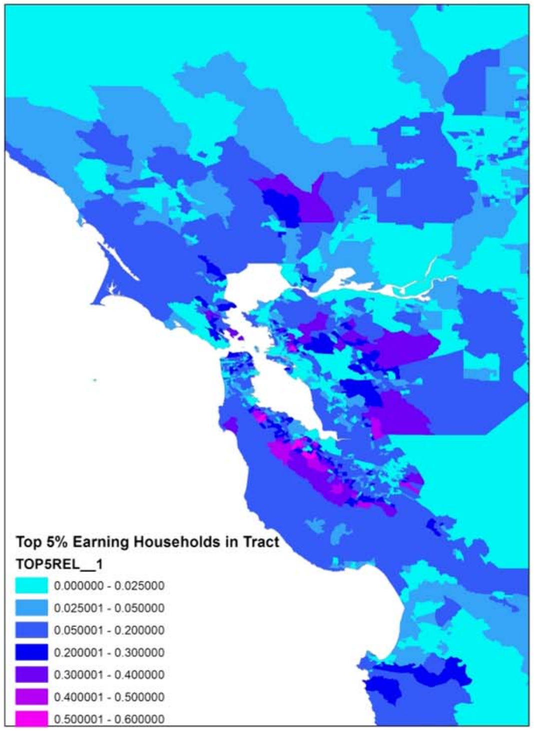
Income Factor

- Comparisons of population proportion in each bin show good match between spatial distribution of QTILE and EXP
- Key point: the higher values are the focus of the model, where EXP predicts more accurately

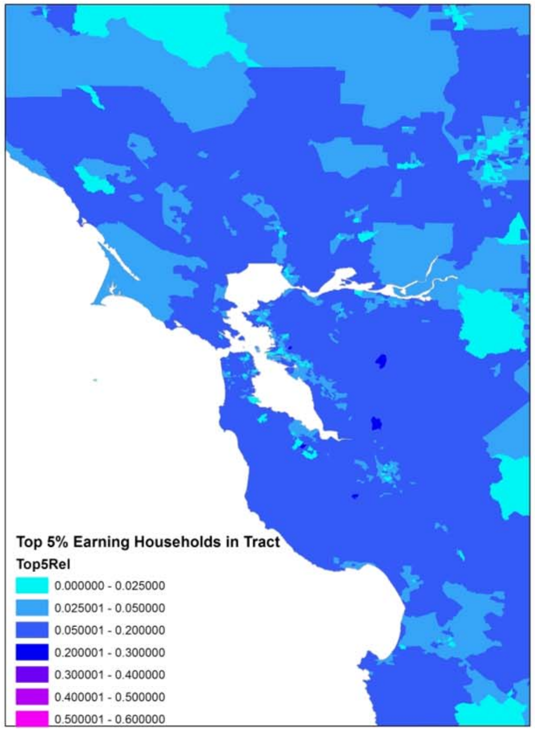
QTILE



EXP



LN



Market Assessment

Vehicle Attributes

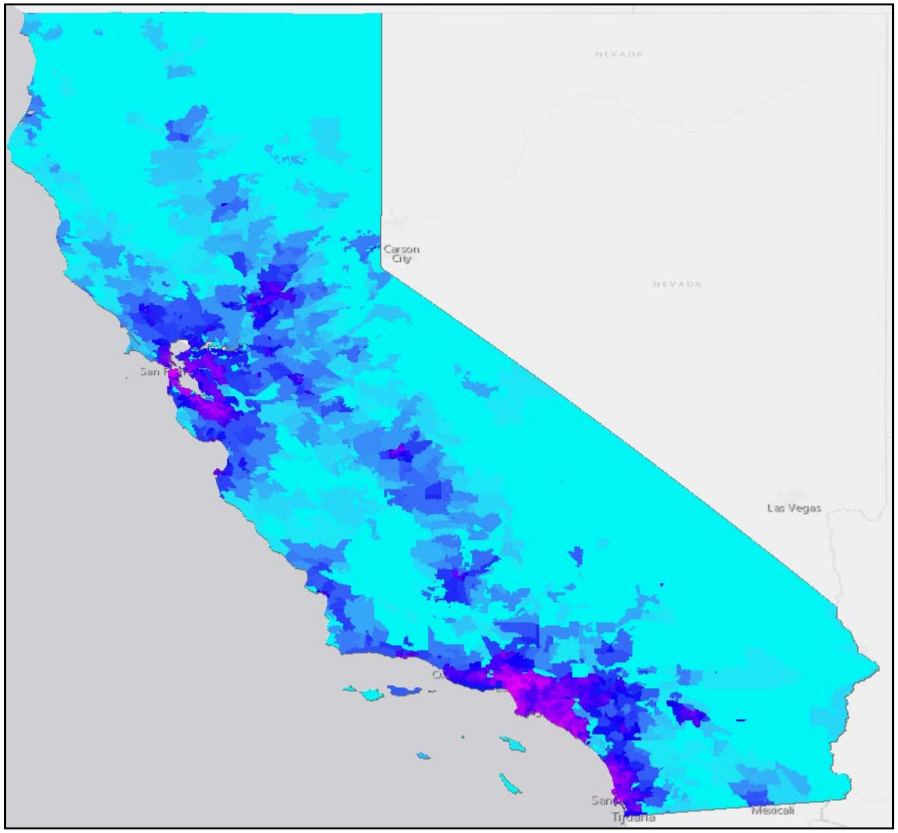
- Multiple vehicle-based attributes investigated as additional indicator factors
- Based on DMV data available at ZIP code resolution
- Luxury Vehicles: Indications that luxury brand buyers may be more likely to be first adopters of vehicle technologies; implemented counts of registered luxury brand vehicles in the last 5 model years
- MSRP Range: DMV data include a range of MSRP values for all models. Based on reported values for pre-commercial FCEV vehicles (excluding FCHV-adv). Counted all vehicle registrations in last 5 years that fall within this estimated range.
- Green Vehicles: First adopters of previous technologies may be more likely to become first adopter of FCEV. Limited to hybrid and plug-in hybrid as these do not require a change in fueling behavior (similar to paradigm for FCEV). Counted vehicle registrations in first 8 years of each technology’s market participation.

Market Assessment

Luxury Vehicle Registrations

- > 1 million total registrations with sufficient data quality

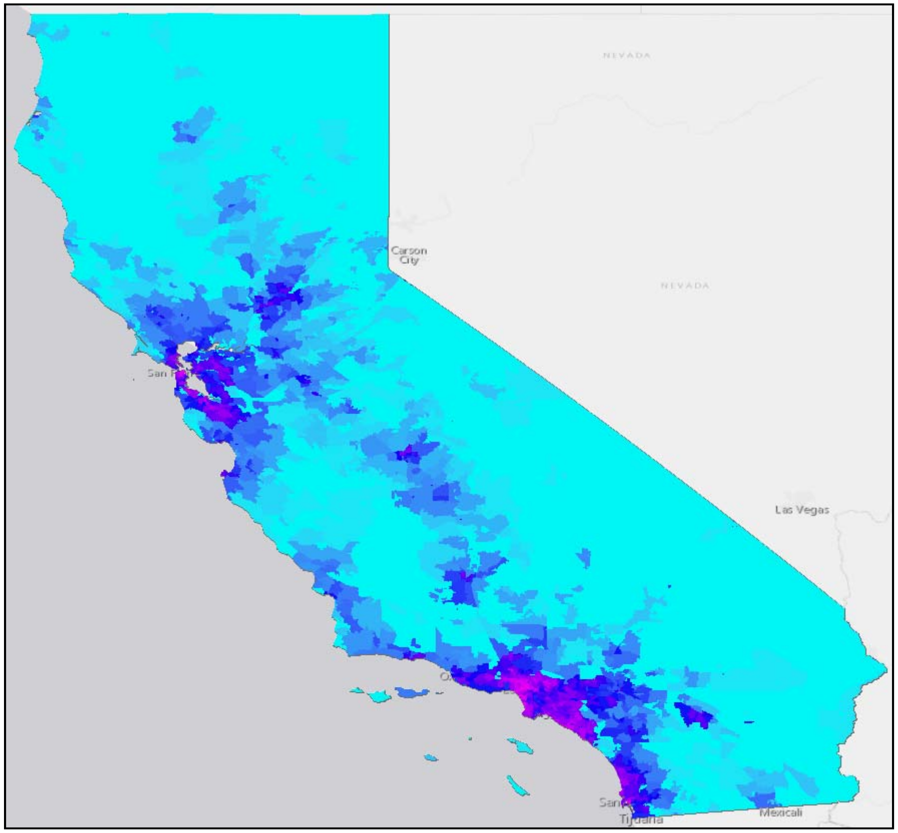
- Based on registrations of
  - Acura
  - Audi
  - BMW
  - Cadillac
  - Infinity
  - Land Rover
  - Lexus
  - Lincoln
  - Mercedes-Benz
  - Porsche
  - Volvo
  - Tesla



Market Assessment

MSRP Range

- Considered Median, Max, and Min of reported Hi/Lo on DMV data historical FCEV data and known MSRP from current and announced vehicles
- Defined Hi/Lo Range as Low MSRP at least $50k and High MSRP at most $75k
- Based on first-owner registrations of vehicles with MSRP matching within the historically-based hi/lo MSRP range for FCEVs



Market Assessment

Covariation of Financial Indicators

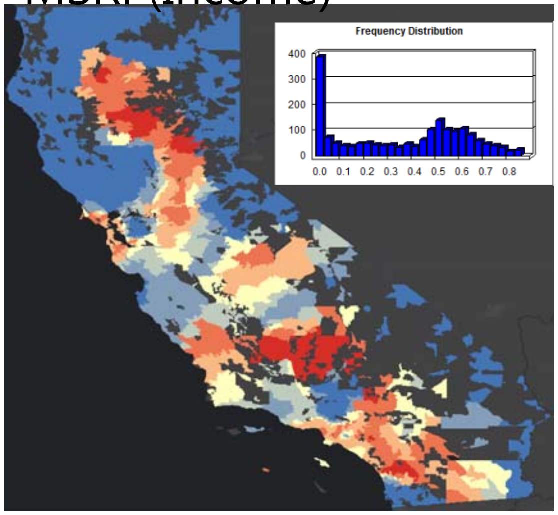
- Qualitatively, appears to be some match between spatial variation in MSRP, income, and luxury registrations
  - Conceptually, all have an element of finances/income
- Explored spatial covariation through a number of models with Geostatistical Analyst
- Spatial patterns of MSRP more related to luxury vehicle registrations than income
- Slight improvement in model when both included
  - Global R^2 for MSRP(income): ~0.25
  - Global R^2 for MSRP(luxury): ~0.56
  - Global R^2 for MSRP(luxury, income): ~0.58
- Both luxury vehicles and income are statistically significant factors in purchase price

Market Assessment

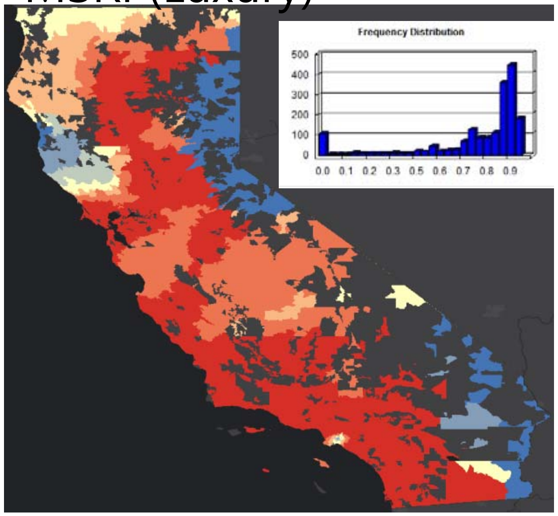
Covariation of Financial Indicators

- Localized R^2 Values
- Blue is low R^2 (low explanatory power)
- Red is high R^2

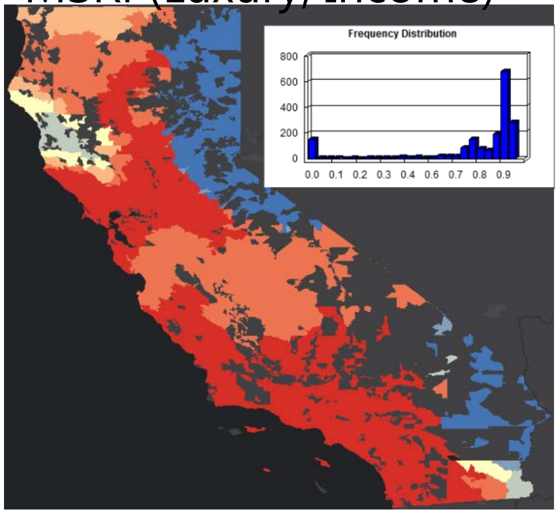
MSRP(Income)



MSRP(Luxury)



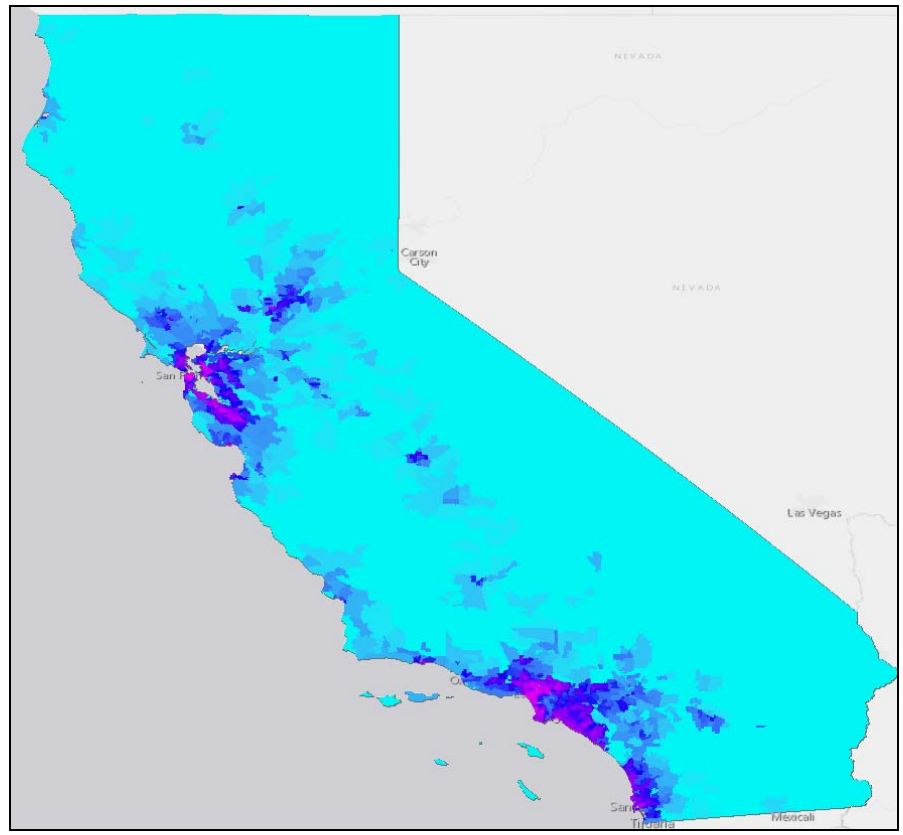
MSRP(Luxury, Income)



Market Assessment

Green Vehicle Adoption

- Assumed eight-year period for first adopter market
- Utilized counts of Hybrid Vehicles in MY 2000-2007 and Plug-In Hybrid in MY 2010-2014
- >300k total registrations with sufficient quality data

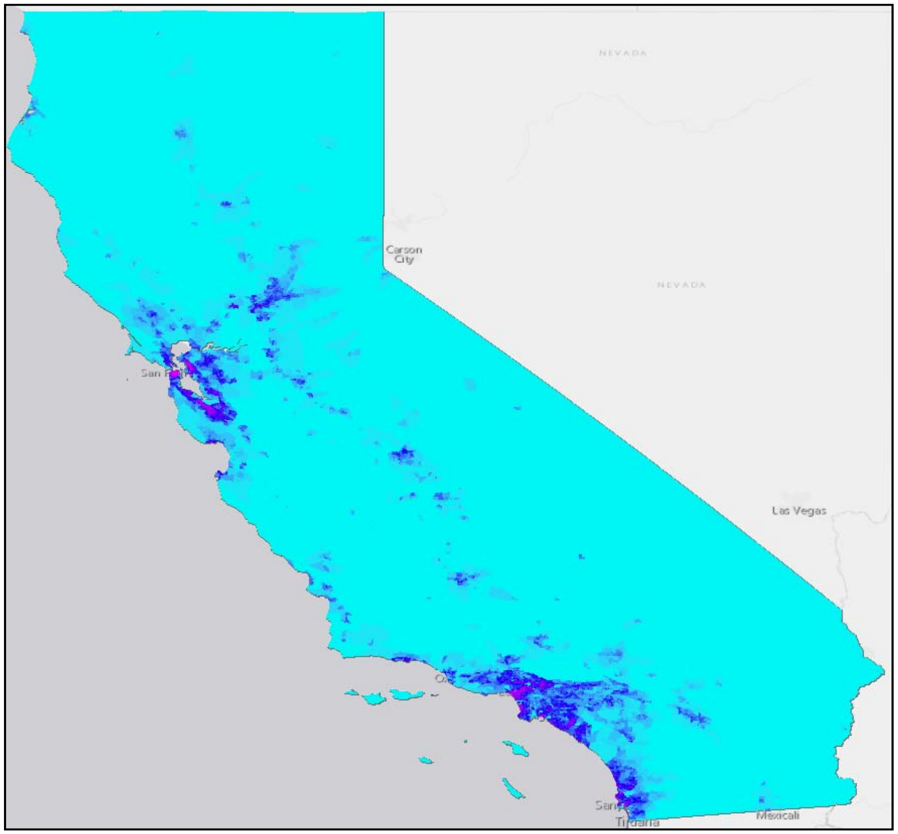


Market Assessment

Education

- Indications that areas with households having increased education levels may be more likely to be early adopters of green/advanced technologies

- Utilized census data counts of number of post-graduate degrees in population 25 and older

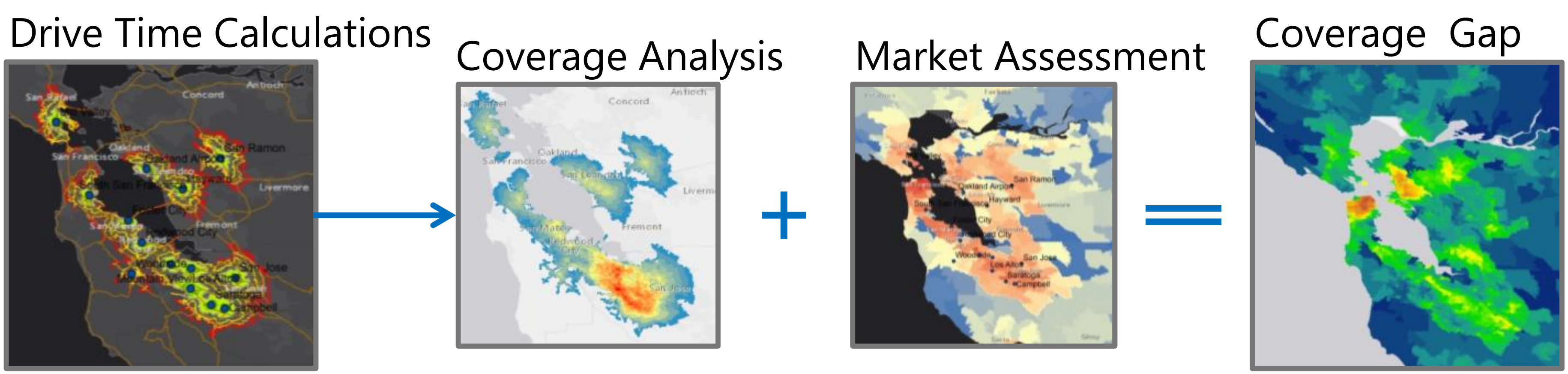


PRIORITY AREA IDENTIFICATION

Priority Areas

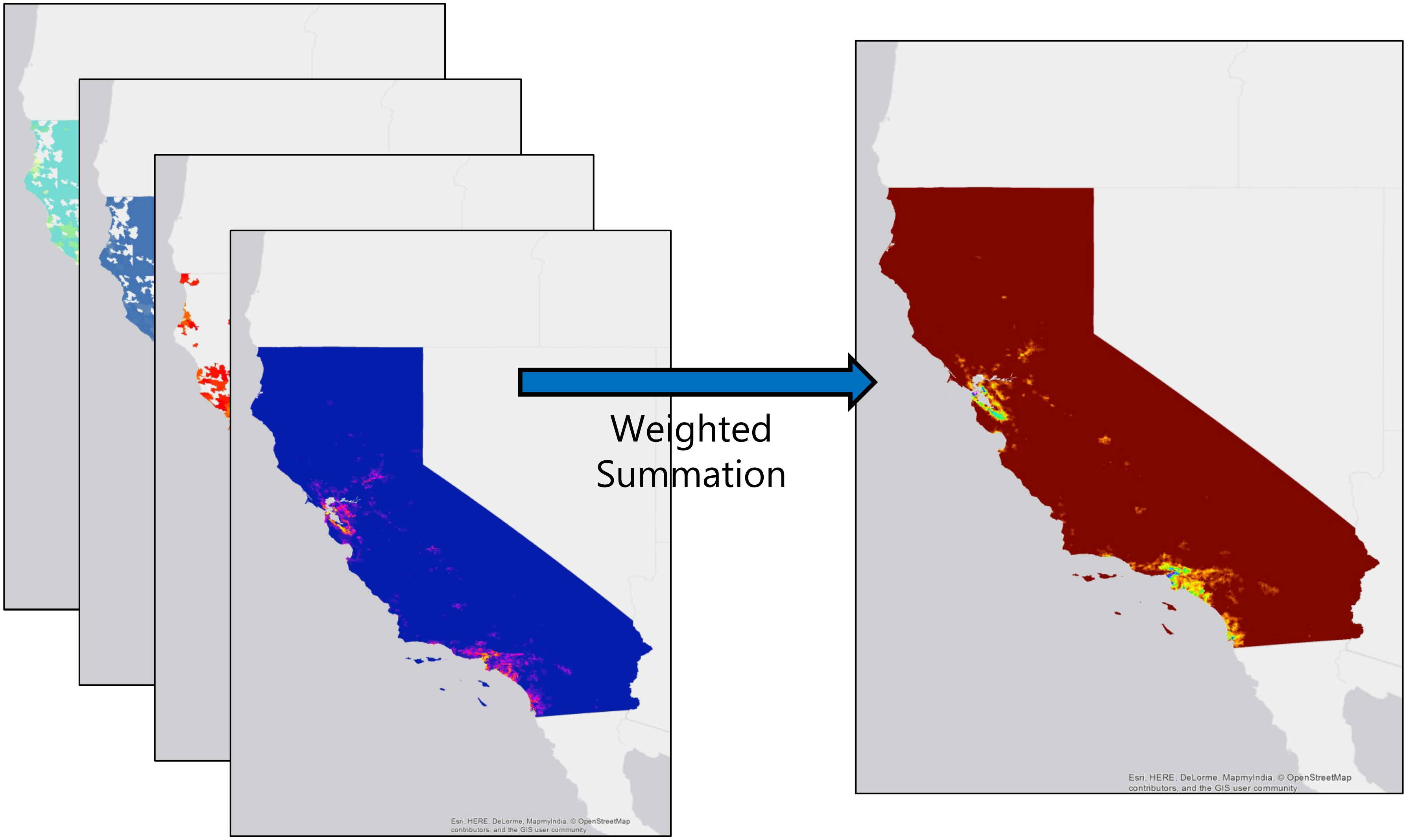
Structure of CHIT Evaluation

- Compares evaluation of market and coverage to determine gaps
- Enables separate identification of market potential and areas of greatest coverage and capacity needs
- Enables annual dynamic evaluation to adjust planning as deployment progresses



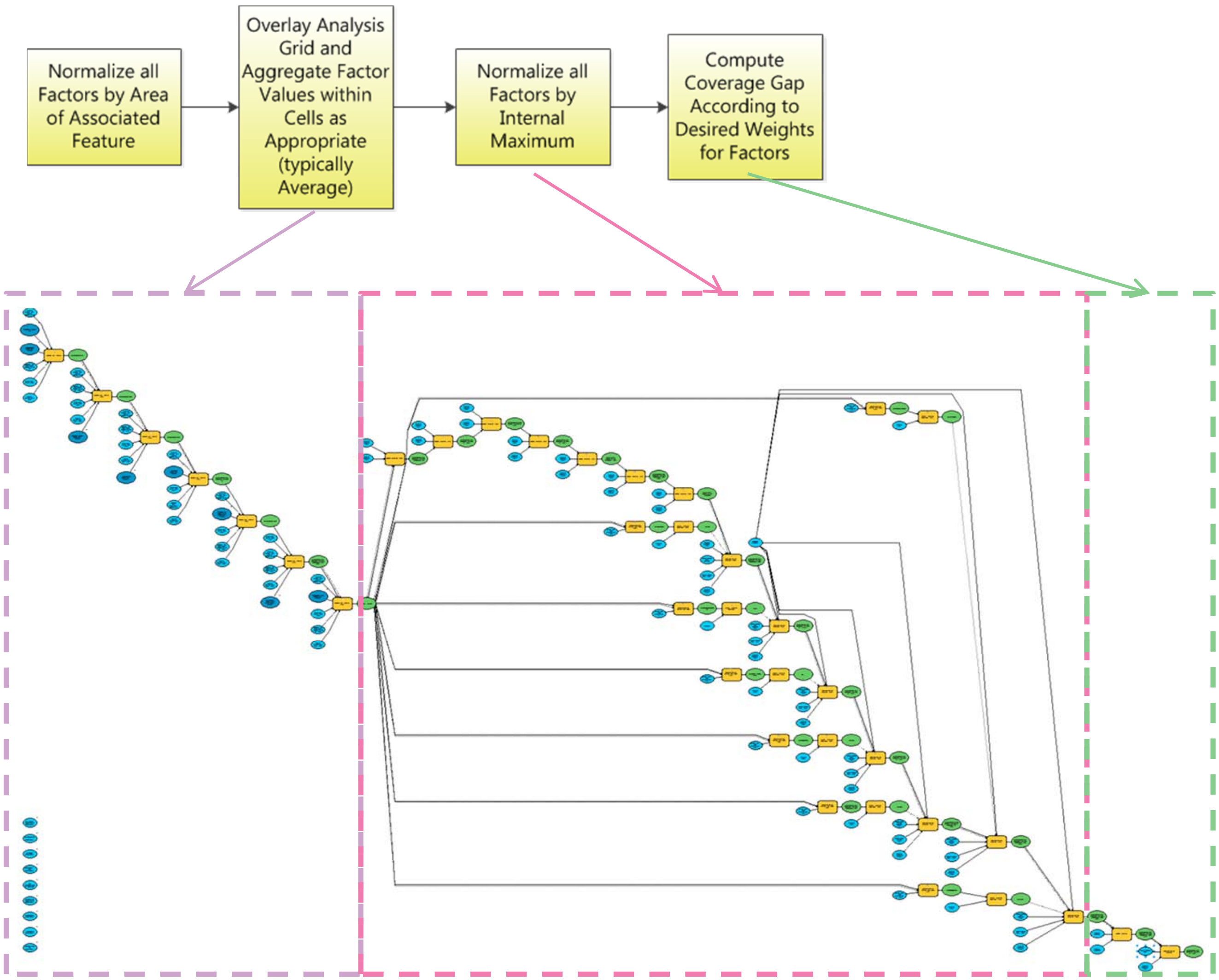
Priority Areas

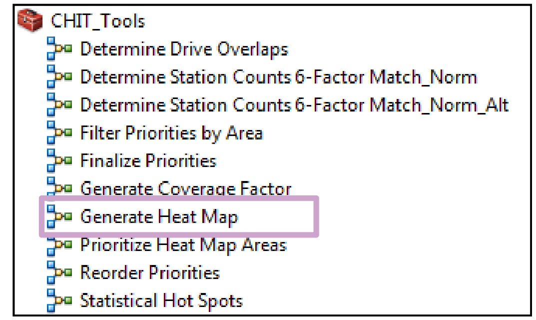
Analyzing the Early Adopter Market



Priority Areas

Steps in Considering Multiple Factors





- Normalizing factors by internal maximum
  - All factors have a rating value of 0 to 1
  - Allows for consideration of varying measures together
  - In accordance with Principle #1, mostly interested in the areas where indicators simultaneously have concentrations of high values

Priority Areas

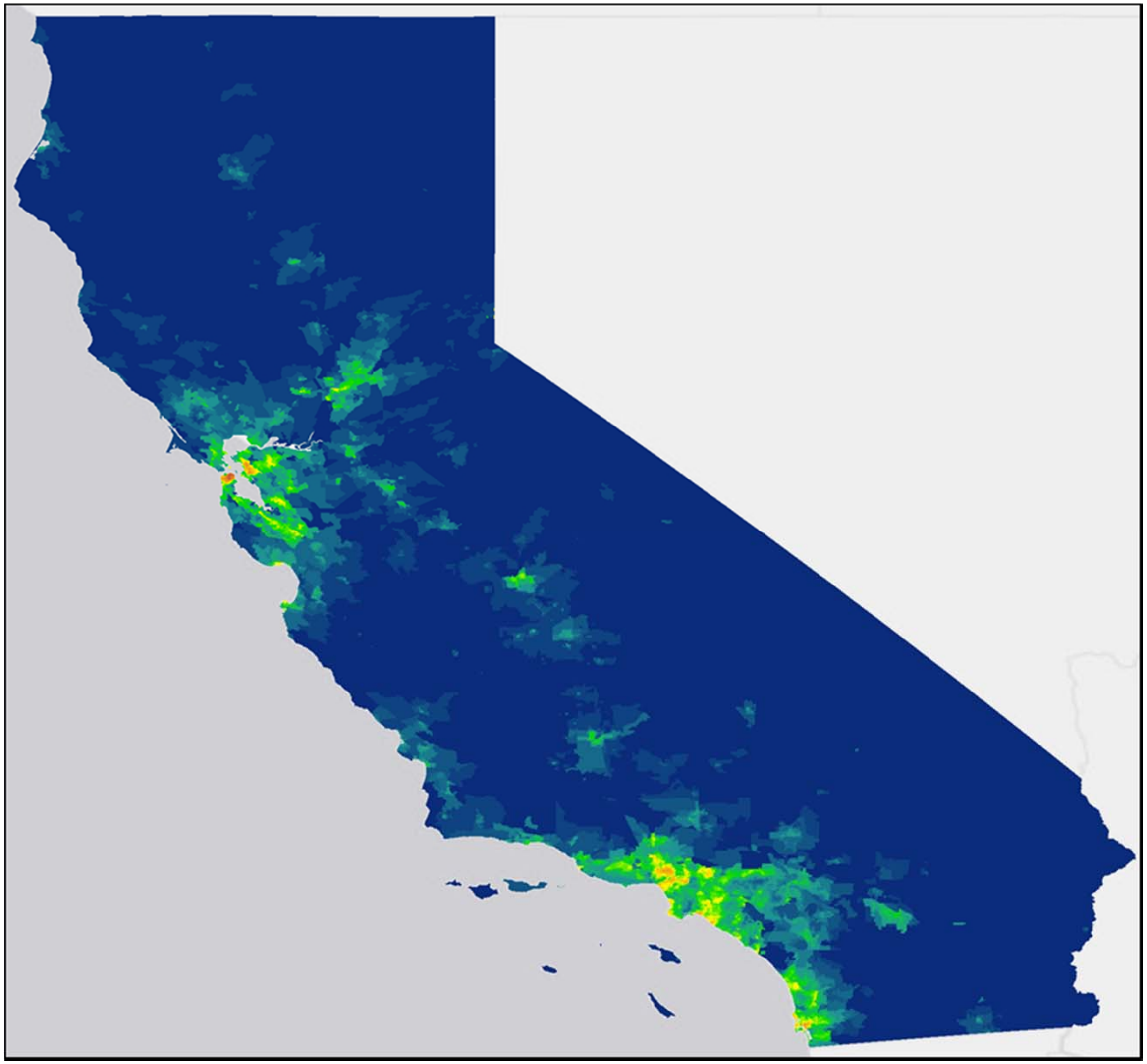
Coverage Gap Map Formulation

Heat = Coverage *
Market

Coverage= 0.5 * Existing +
0.5 * Potential

Market= 0.5 * Financial +
0.3 * P/HEV +
0.2 * Edu

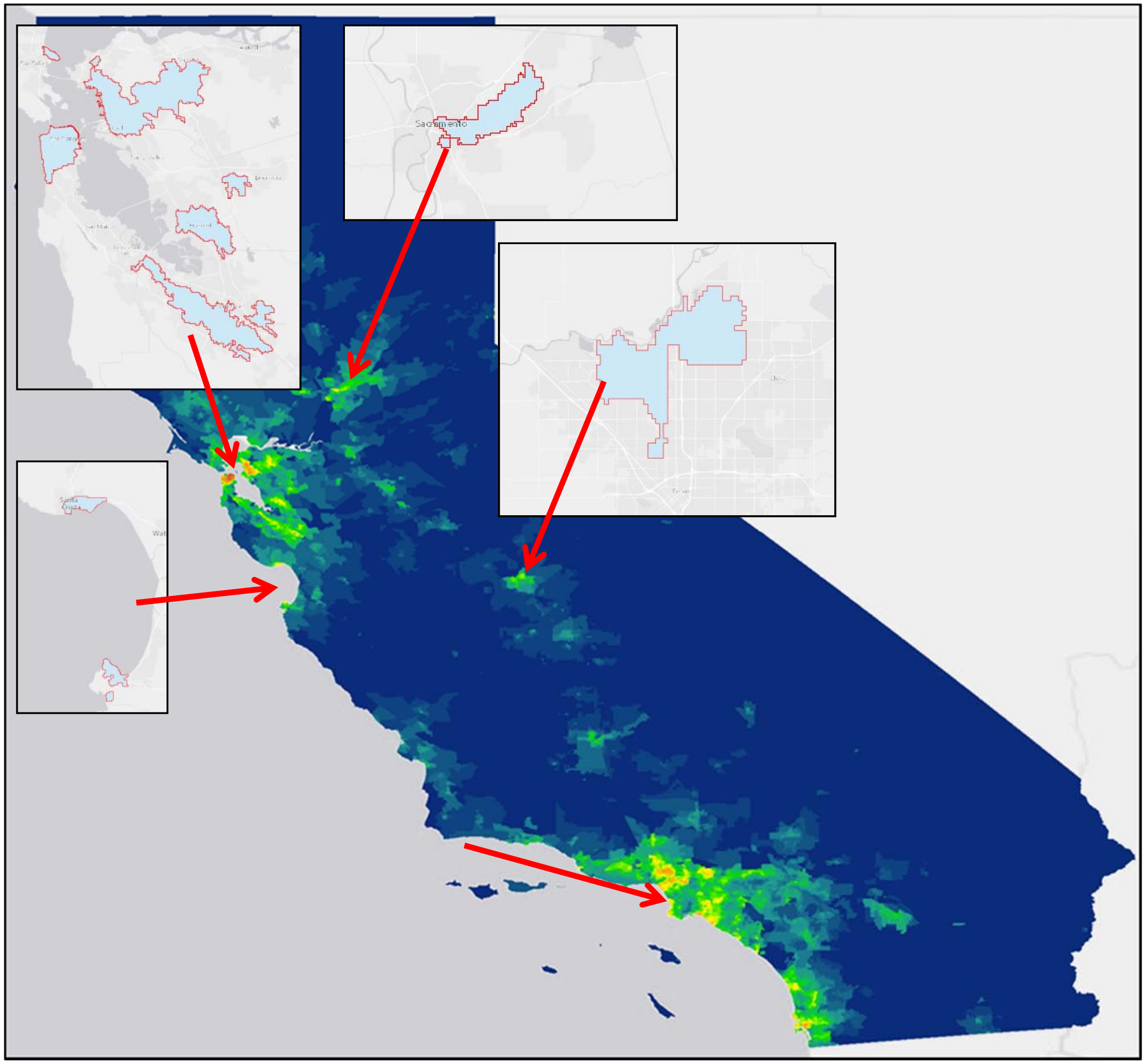
Financial= 0.34 * Income +
0.33 * MSRP +
0.33 * Luxury



Priority Areas

Analyzing Coverage Gap Map

- Gap analysis provides heat map, but must still analyze the map for meaningful interpretation
- Use generalization and statistical methods to identify and rank hot spots on the map
- Requires three steps:
  - Identify global highest scores
  - Identify local Hot Spots
  - Merge Solutions
- Rank by coverage gap score



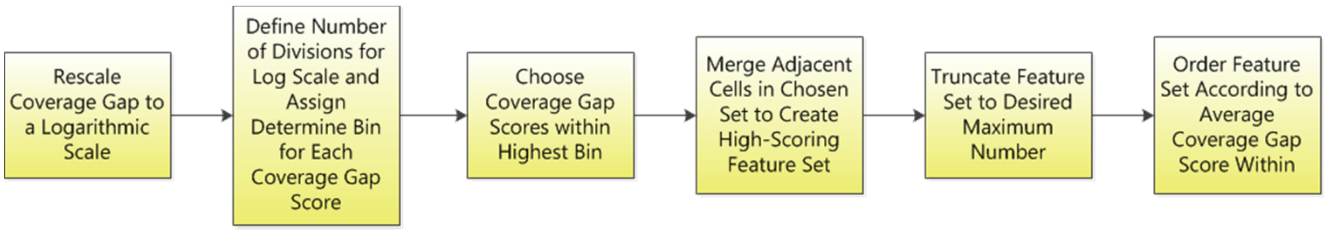
Priority Areas

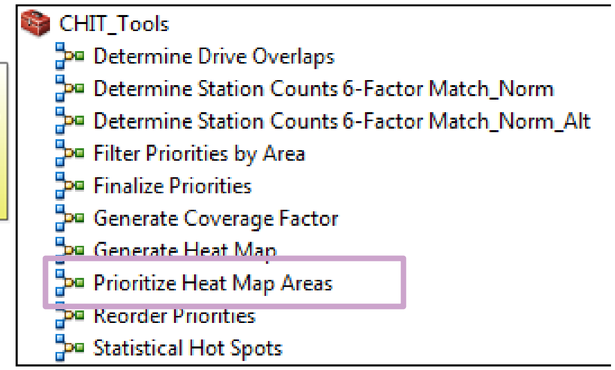
Analyzing Coverage Gap Map

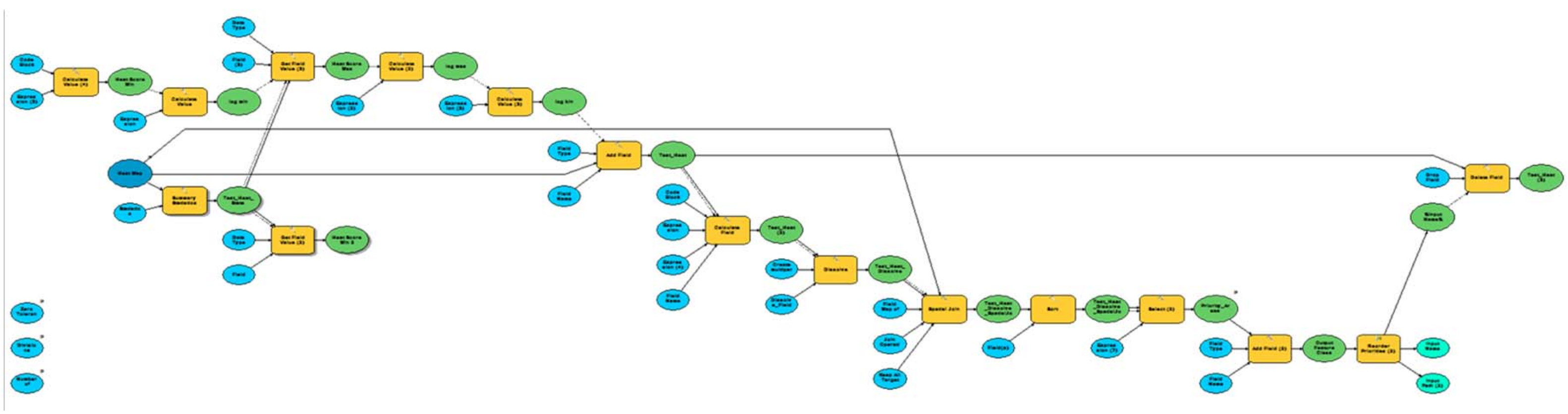
- Exploration of Coverage Gap Map values indicates range of 0 to ~0.54
  - Indication that there is no area where an exactly ideal location exists
- Distribution of coverage gaps heavily weighted to low end
- Few high-scoring areas that are much higher than low and mid-scoring areas
- Interpreted results as logarithmically distributed
- Generalization of results necessary to avoid too narrow of a definition of priority area

Priority Areas

Analyzing Coverage Gap Map



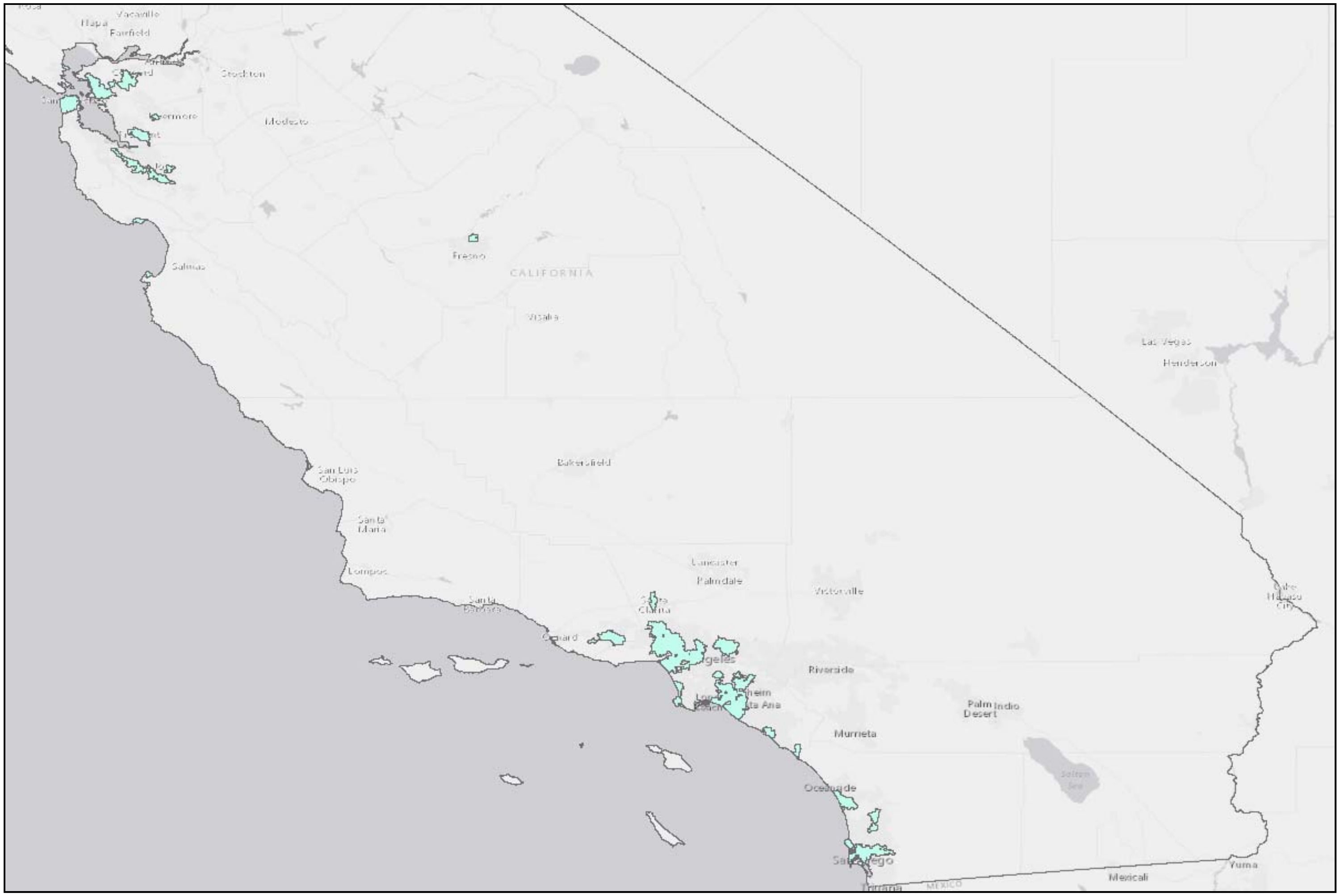




- For AB 8 2015 report, divided scale of coverage gap scores into 5 classes and chose the highest ranking class

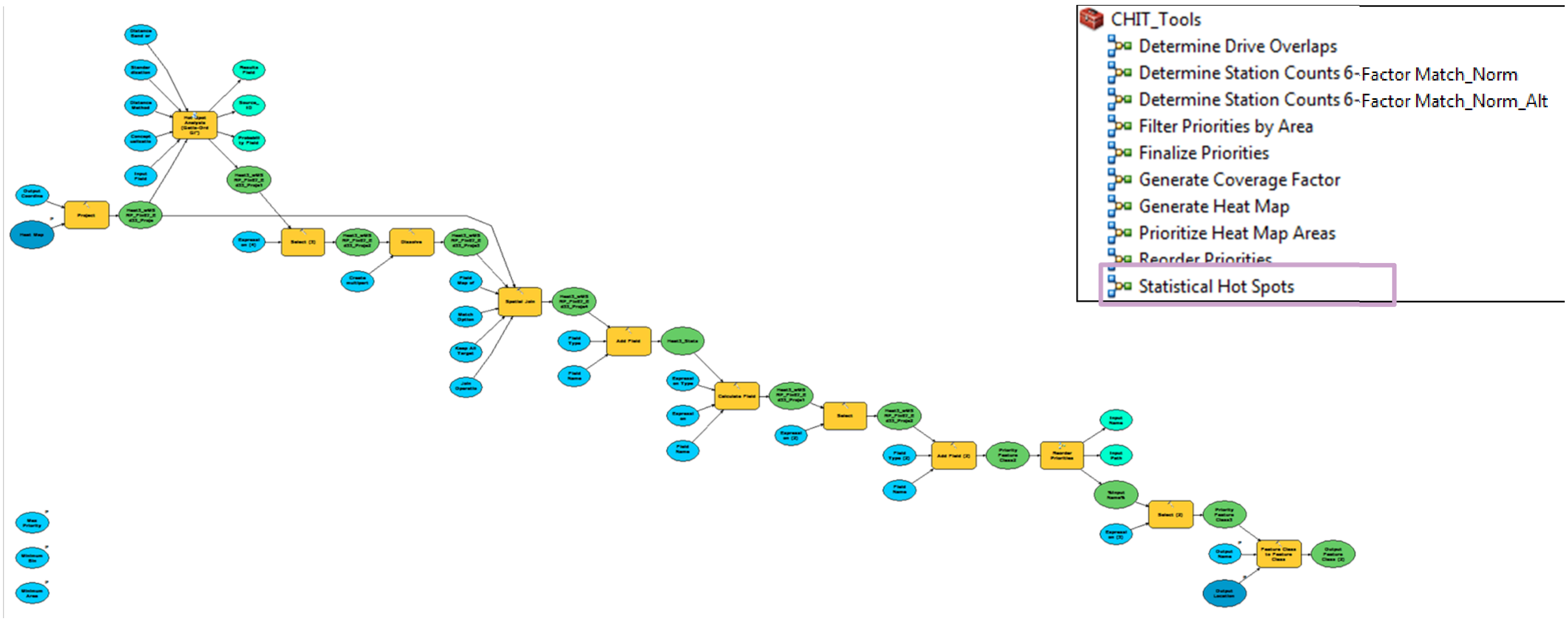
Priority Areas

Analyzing Coverage Gap Map



Priority Areas

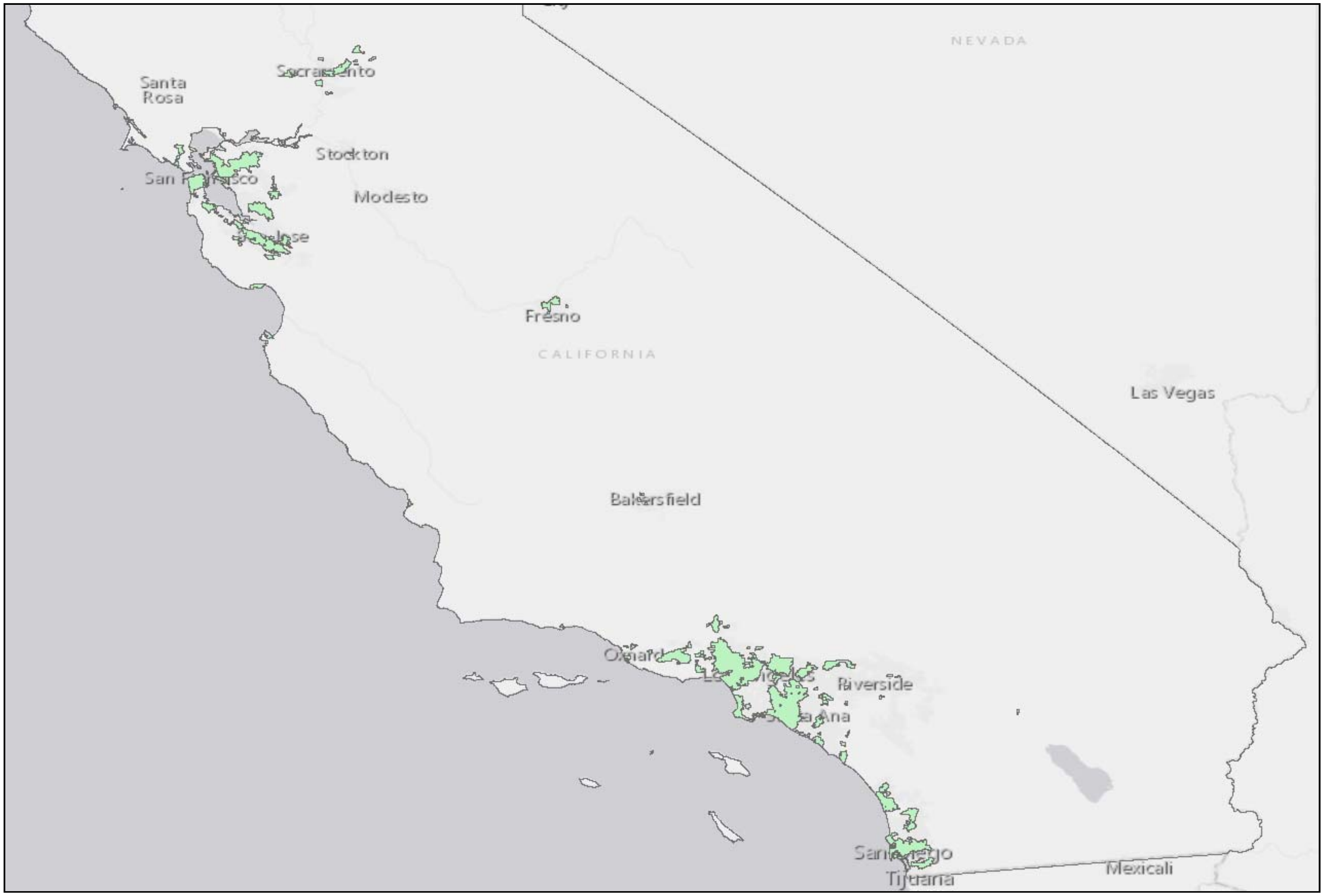
Analyzing Coverage Gap Map



- Simply runs ArcGIS' built-in Hot Spot analysis, providing tuning parameter options and ability to limit total number of output areas

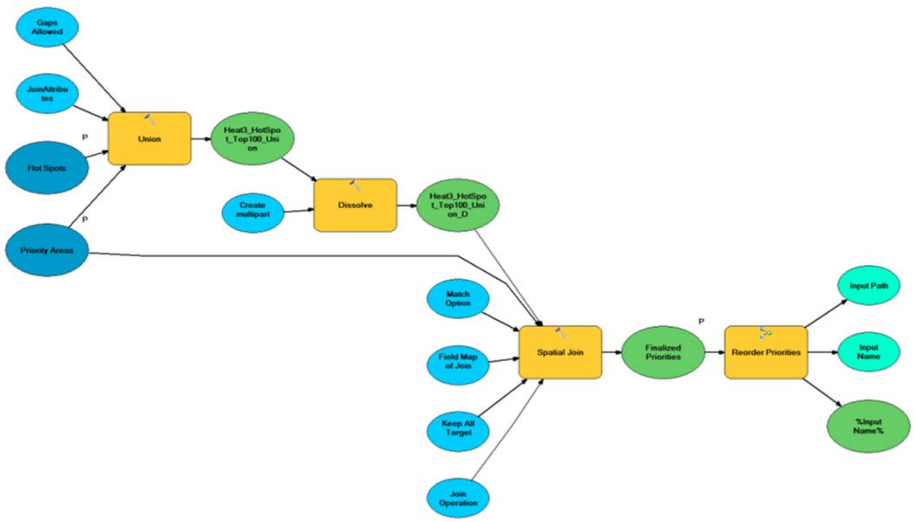
Priority Areas

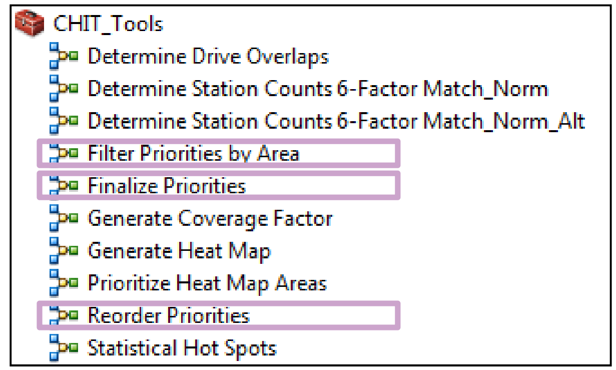
Analyzing Coverage Gap Map



Priority Areas

Analyzing Coverage Gap Map



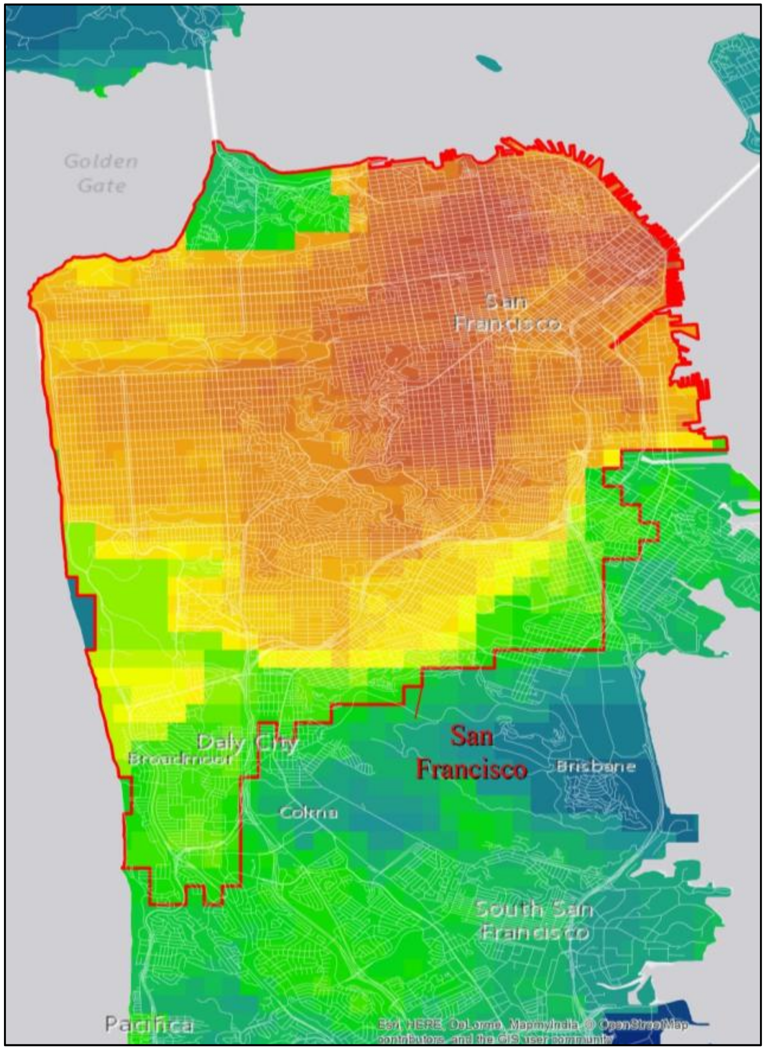


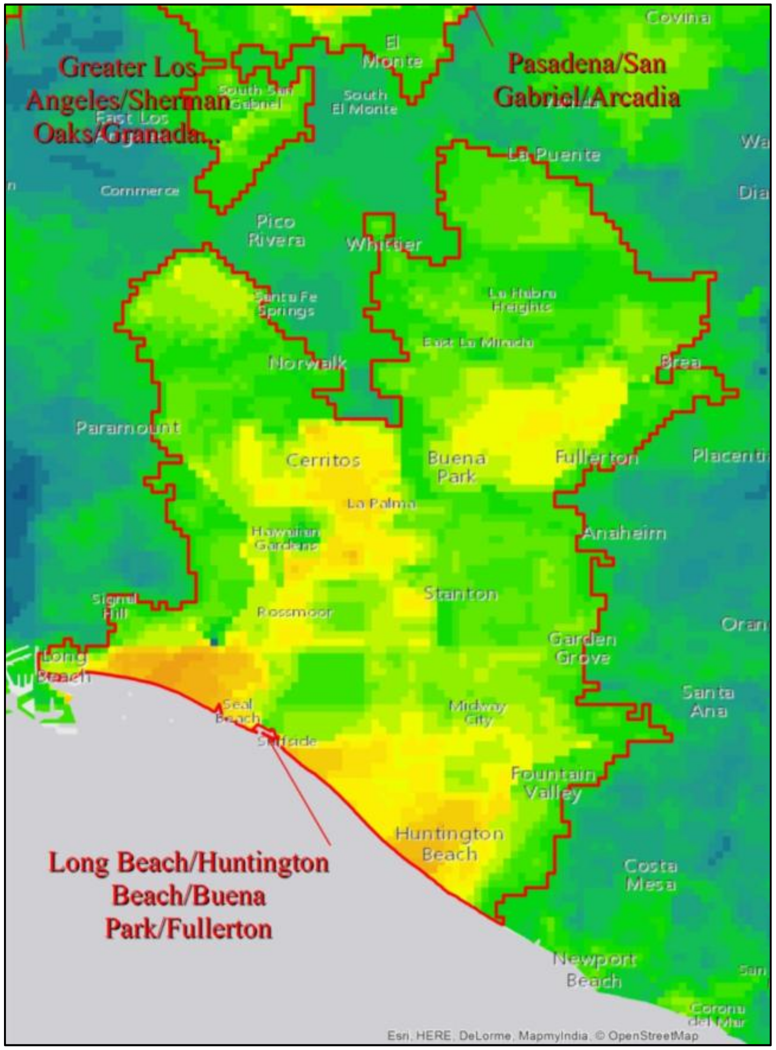
- Union of two feature classes assures final output contains areas that are locally desirable and globally high-ranking

- In 2015 AB 8 report, also filtered areas to at least 1.5 square miles; CHIT provides tools to allow filtering by any user-desired size and then re-ordering resultant set

Priority Areas

Sample Priority Areas



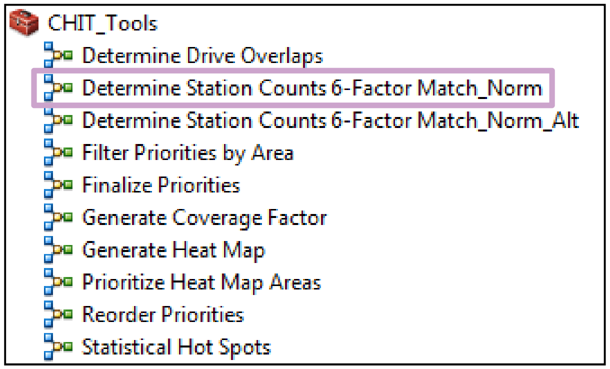


CAPACITY NEEDS IN MARKET AND PRIORITY AREAS

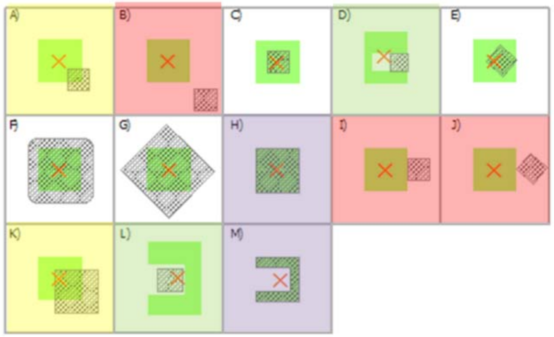
Capacity Needs

Projecting Capacity Needs

- Determine capacity needs according to likely distribution of vehicles in full first adopter market



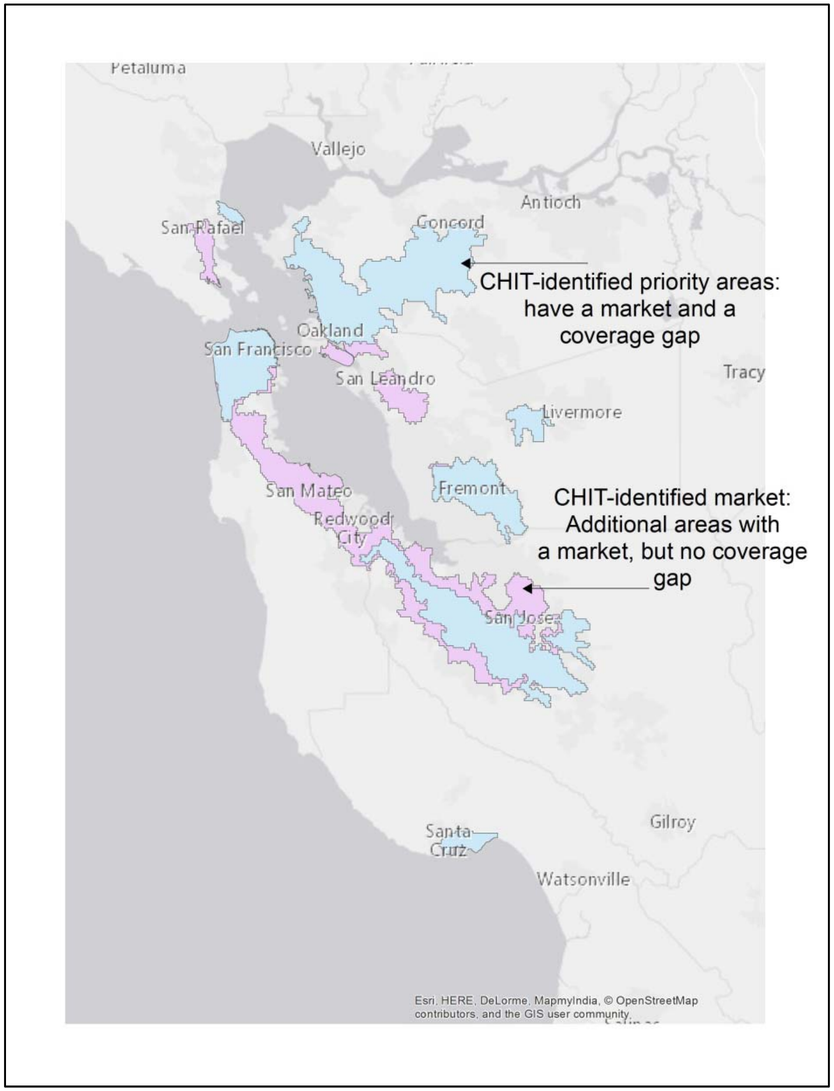
- Account for existing stations' capacity
  - Assumed a station services a market or priority area if its 6-minute Service Area overlaps the market or priority area
  - Because of ArcGIS' rules for how to count features as overlapping, required 6 steps of overlapping to appropriately include stations
    - Fundamental relationships
    - ArcGIS provides options for sets of fundamental relationships
    - Needed to combine theses sets to arrive at desired solution



Capacity Needs

Steps in Projecting Capacity Needs

- Step 1: Identify Full Early Adopter Market
  - Find areas with high market values
  - Follow entire CHIT process, with Coverage Factor weights set to 0
- Step 2: Separate out Priority Areas
  - Find areas with high coverage gap values
- Step 3: Assign Existing Station Capacity
  - Determine capacity provided by current stations
- Step 4: Allocate Vehicles to Determine H_2 Balance
  - Use CHAT to project vehicle population
  - Assign FCEVs to priority and market areas proportional to market score and population (0.6 market + 0.4 population)
  - Vehicle count and existing capacity determine H_2 need or surplus



Capacity Needs

Capacity Needs Output

- For PON Concepts, analyzed 2021 need based on published 34,300 vehicles
- Used CHAT-calculated vehicle projections from 2015 to 2021 to extrapolate need in intervening years
- Based recommendations of number of stations in 2015 AB 8 report on a near-term projection over a few years

Area	Capacity Need (kg/day)
San Francisco	2070
Berkeley/Oakland/Walnut Creek/ Pleasant Hill	1120
San Diego/La Mesa	990
Greater Los Angeles/Sherman Oaks/Granada Hills/Glendale	1700
South San Diego/Coronado	320
Torrance/Palos Verdes/Manhattan Beach/Redondo Beach	320
Pasadena/San Gabriel/Arcadia	540
Long Beach/Huntington Beach/Buena Park/Fullerton	1520
Santa Cruz	330
Encinitas/Carlsbad	400
Fremont	390
Sacramento/Land Park	220
Sacramento/Carmichael	370
Thousand Oaks	330

		Area	Stations	Purpose
First Priority	1	San Francisco	6	Establish Core Market
	2	Berkeley/Oakland/Walnut Creek/ Pleasant Hill	3	Establish Core Market
	3	San Diego/La Mesa	3	Expand Core Market Coverage
	4	Greater Los Angeles/Sherman Oaks/Granada Hills/Glendale	3	Core Market Capacity
	5	South San Diego/Coronado	1	Expand Core Market Coverage
	6	Torrance/Palos Verdes/Manhattan Beach/Redondo Beach	1	Core Market Capacity
	7	Pasadena/San Gabriel/Arcadia	1	Expand Core Market Coverage
	8	Long Beach/Huntington Beach/Buena Park/Fullerton	2	Expand Core Market Coverage
	9	Santa Cruz	1	Future Market
	10	Encinitas/Carlsbad	1	Connector/Future Market
	11	Fremont	1	Future Market
	12	Sacramento/Land Park	1	Expand Core Market Coverage
	13	Sacramento/Carmichael	1	Expand Core Market Coverage
	14	Thousand Oaks	1	Future Market

OPEN DISCUSSION OF POTENTIAL FUTURE DEVELOPMENT

Future Development

ARB Concepts

- Traffic volume
  - Significant interest has been communicated in assessing traffic volume patterns. To date, ARB has not been able to identify a data set that satisfies Principle #2. ARB is interested in suggestions of data sets to explore.
- Gasoline stations
  - Current funding paradigms emphasize co-location on existing gasoline stations. Proximity to gasoline stations could be an influential factor. Throughput of existing gasoline stations could also influence relative importance.
  - Data exist and can be obtained by ARB, but concerns of confidentiality of throughput data may limit implementation.

Future Development

ARB Concepts

- Origin-Destination studies and data
  - Origin-Destination would present a new paradigm to be included in CHIT, potentially supplanting Principle #6
  - Valid arguments have been advanced in the community and literature in support of O-D modeling; however, ARB is not currently aware of data set that is complete enough to cover the full range of trips representative of the state's population. ARB is interested in suggestions of data sets to explore.
- Hydrogen production facilities
  - Inclusion of centralized production facilities as a factor could provide a more holistic view of the full hydrogen supply chain
  - Could move CHIT more towards an optimization than a planning tool
  - Unclear of the importance at the current moment

Future Development

ARB Concepts

- Disadvantaged communities, environmental justice, and other socio-political considerations
  - The State has many goals and programs related to environmental justice
  - CHIT could employ these principles and/or data either as part of its calculations or in post-processing to assess how CHIT calculations may address these goals
  - CalEnviroScreen may be a rich resource for integrating these issues

- Adjusting the Potential Coverage factor
  - Currently uses only a 6-minute drive
  - Could be expanded to consider and weight other drive times, as in the Existing Coverage factor
  - Could also weight captured population by market factors

Future Development

Open Discussion

- ARB welcomes feedback regarding these concepts
  - Appropriateness
  - Relative importance
  - Suggestions for data sets to explore

- ARB additionally welcomes suggestions for concepts that have not yet been identified and suggestions to improve the concepts and data that have been implemented

OPEN DISCUSSION OF POTENTIAL PUBLIC DISTRIBUTION

Public Distribution

ARB's Ideal Plan

- Principle #7 highlighted the desire for ARB to create a tool that could be publicly shared
- ARB provided summary results and its own assessment of the tool's results in the June AB 8 report and intends to continue to do so
- Further sharing of CHIT involves both the fully detailed results and access to the tool itself:
  - Ideally, ARB would be able to host an interactive map application for stakeholders to be able to obtain various output data from the CHIT tool on their own
  - Additionally, ARB would provide a pre-packaged download and instructions for setting up the data and the tool so stakeholders can explore, implement, and modify for their own analysis purposes

Public Distribution

Challenges

- ARB does not currently have the capability to host interactive pages (internal Information Services only supports static pages)
  - A potential solution is to provide maps of small areas of the state, which can be chosen by clicking on an area of a static map image
  - Pre-generated PDFs of the map areas would then be shown on screen and able to be downloaded
  - This strategy is employed by other GIS and map databases in the public sphere
- ArcGIS is not free software
  - Open-source options exist but are not supported by ARB Information Services and redevelopment of CHIT in other programs will take significant time and are not guaranteed to provide exactly the same capability or solutions

Public Distribution

Questions for Open Discussion

- How important is an interactive online mapping application for exploring the output data from CHIT?
- Is the proposed alternative sufficient to meet prospective needs for gathering and exploring CHIT output data?
- What level of interest exists for access to CHIT itself?
- Is cost of ArcGIS a significant barrier?
- Any other considerations and/or concerns?

FURTHER DISCUSSION

For questions or comments, contact:
Andrew Martinez
(916) 322-8449
andrew.martinez@arb.ca.gov