CALIFORNIA HYDROGEN INFRASTRUCTURE TOOL (CHIT) WEBINAR:
SIMULATED TRAFFIC VOLUME DATASET

March 3, 2017

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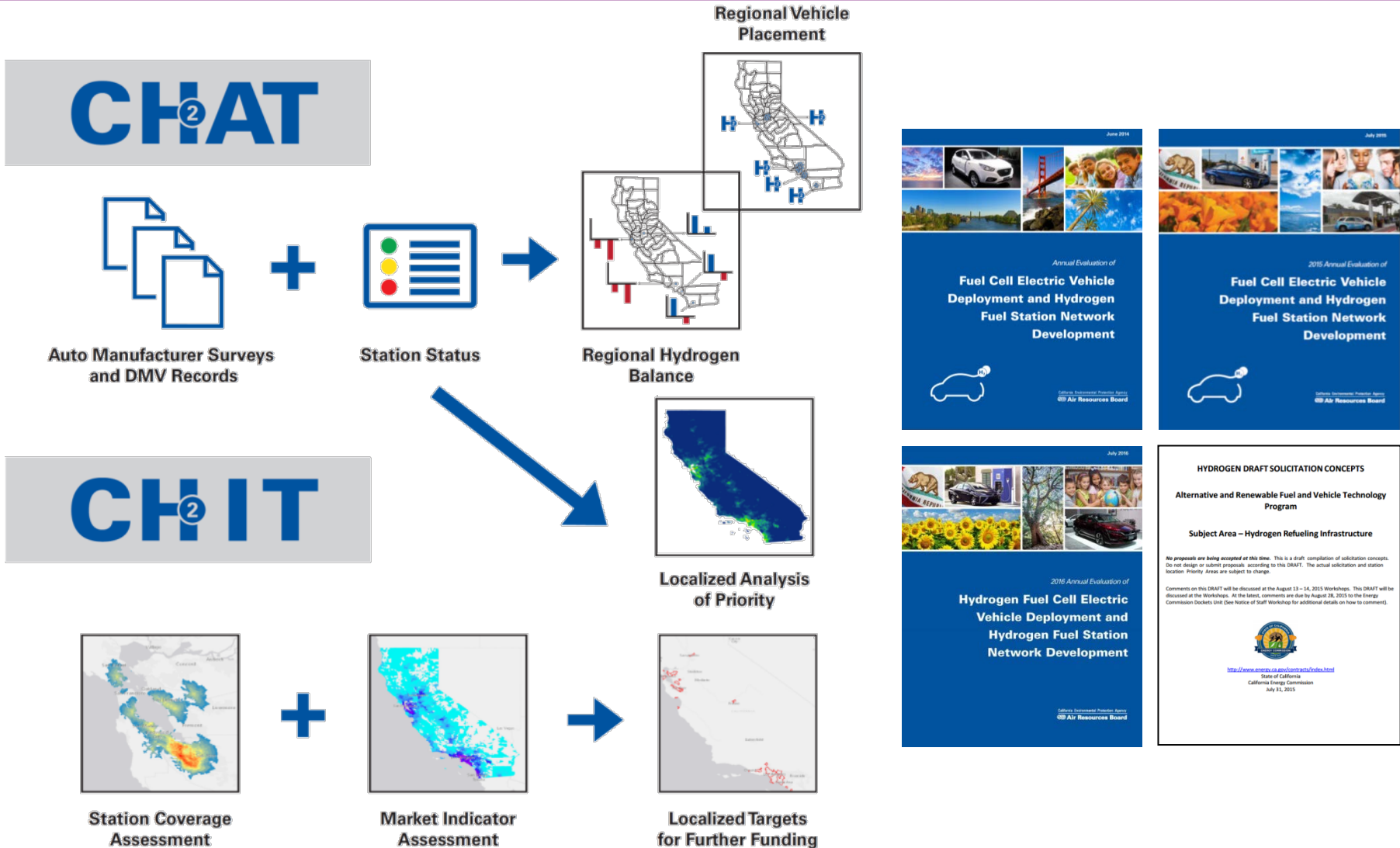andrew.martinez@arb.ca.gov

Discussion Outline

Purpose: To review ARB's proposed implementation of traffic volume data in CHIT. Stakeholders previously suggested that hydrogen fueling station location priorities should include consideration of vehicle traffic. Due to limitations in available observational data, ARB developed a simulated traffic volume data set.

- Review of CHIT motivation, guiding principles, and formulation
- Review of needs for traffic volume data set
- ARB's simulated traffic volume data set
  - Source data
  - Methodology
  - Exploration of simulated traffic volumes
  - Validation with limited available observational data
- Proposed method to combine with established CHIT evaluations
- Proposed method to enable CHIT connector station suggestions
- Consideration of additional market-based weighting factor
- Open discussion and feedback

CHIT REVIEW

Introduction

CHIT/CHAT Tools and AB 8

CHAT

Auto Manufacturer Surveys and DMV Records + Station Status → Regional Hydrogen Balance

Regional Vehicle Placement

Localized Analysis of Priority

CHIT

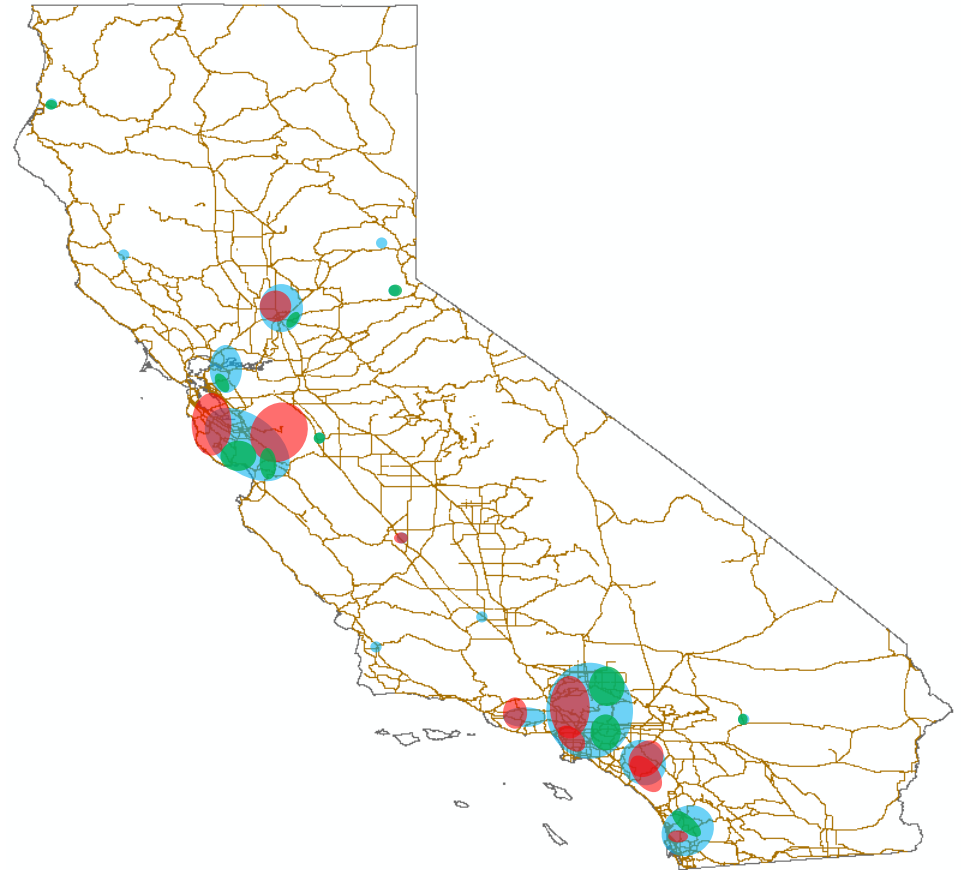Station Coverage Assessment + Market Indicator Assessment → Localized Targets for Further Funding

Introduction

Big Picture Goal

Plan infrastructure placement appropriately for upcoming FCEV releases

1) Identify Market

2) Evaluate current infrastructure
   Existing and potential station coverage

3) Prioritize uncovered market from year-to-year
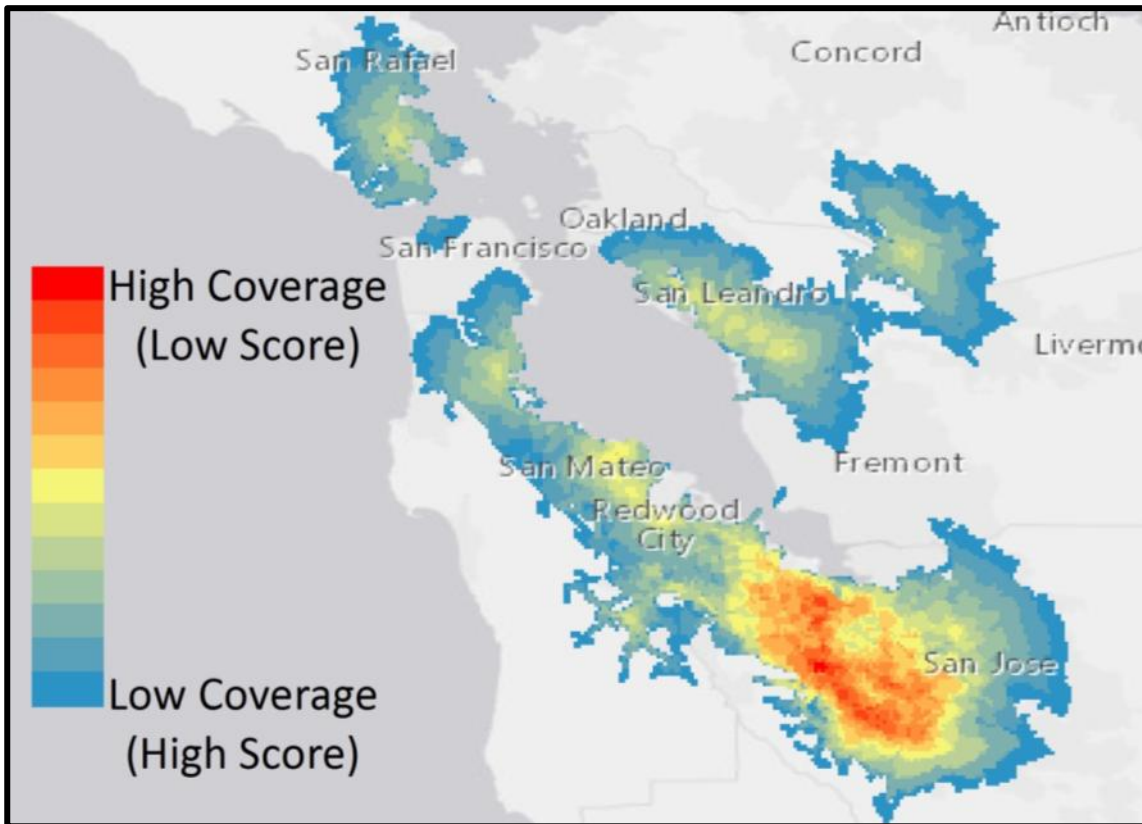
Introduction

Central Theme: Coverage

- Conceptual representation of convenient access to fueling stations

- 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
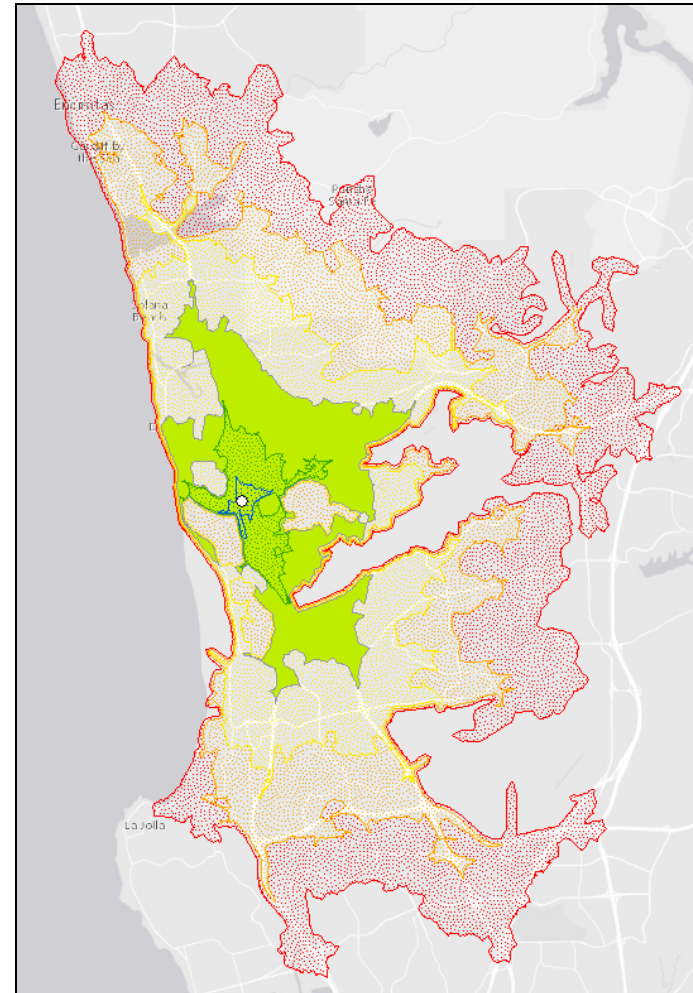
Coverage Algorithms

Goals for Analyzing Existing Coverage

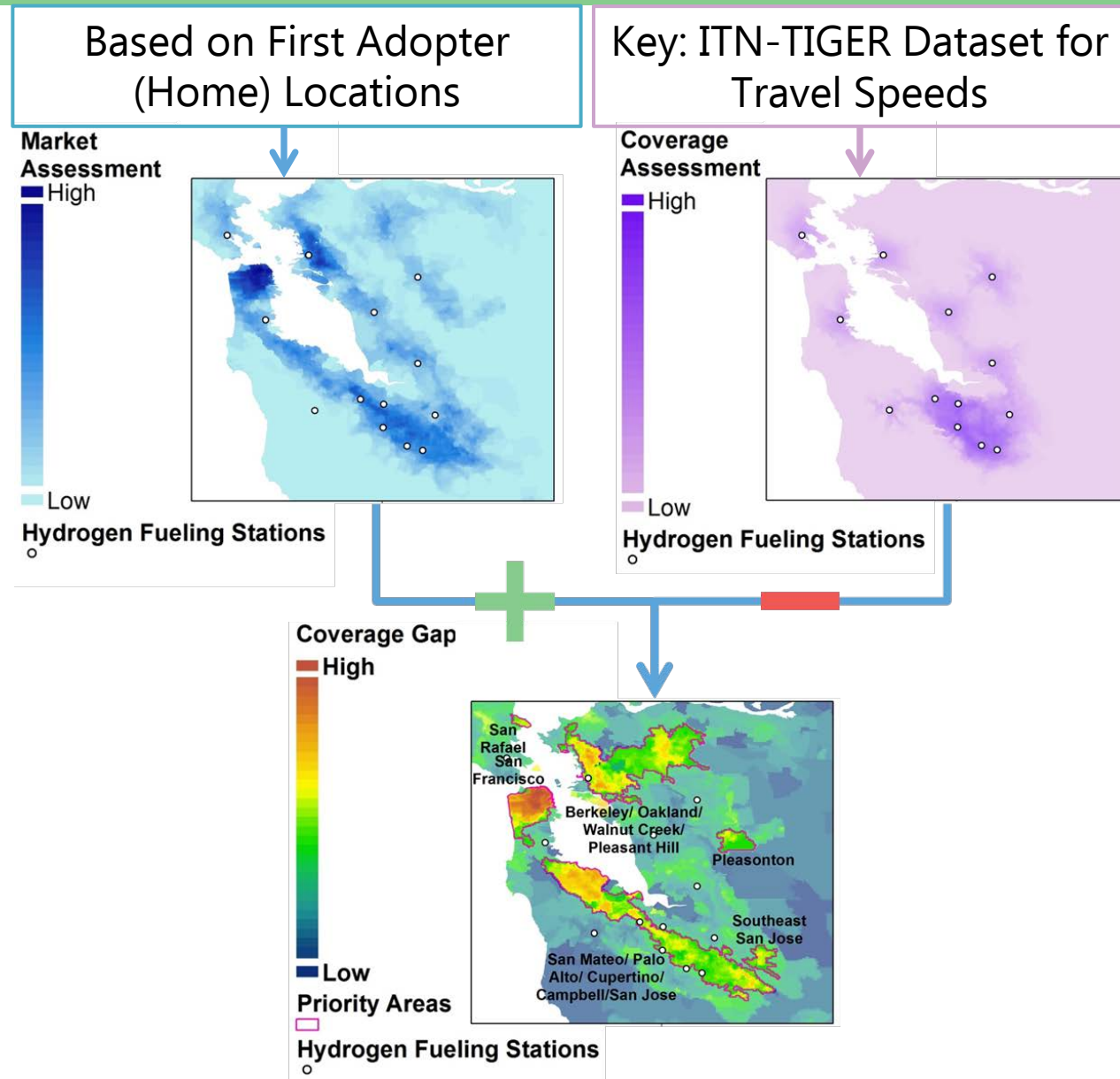*Areas without coverage have no color and score highest

- Provides an estimate of coverage that is more informative than a binary yes/no, allowing for estimation of degrees of coverage

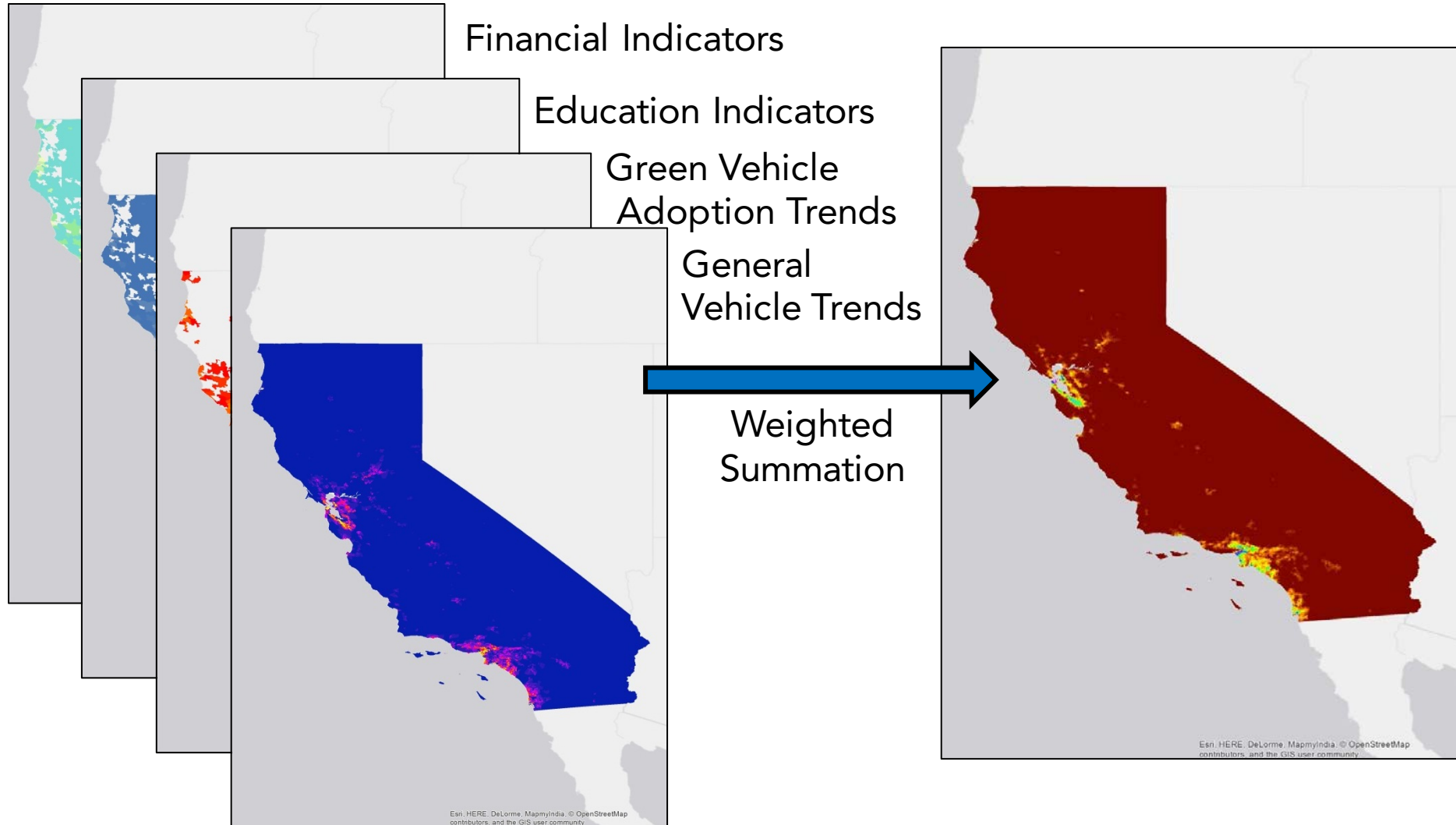- Estimates combined coverage provided by multiple stations that may be reachable within various drive times

CHIT Tool Review

- CHIT is a <u>planning</u> tool intended to provide general direction indicating areas of needed infrastructure

- CHIT evaluates <u>relative</u> need for hydrogen infrastructure based on a gap analysis between a projected market and current infrastructure

Based on First Adopter (Home) Locations

Key: ITN-TIGER Dataset for Travel Speeds

Market Assessment
High
Low
Hydrogen Fueling Stations

Coverage Assessment
High
Low
Hydrogen Fueling Stations

Coverage Gap
High
Low
Priority Areas
Hydrogen Fueling Stations

San Rafael San Francisco
Berkeley/ Oakland/ Walnut Creek/ Pleasant Hill
Pleasonton
Southeast San Jose
San Mateo/ Palo Alto/ Cupertino/ Campbell/San Jose

Review

Analyzing the Early Adopter Market

Financial Indicators

Education Indicators

Green Vehicle Adoption Trends

General Vehicle Trends

Weighted Summation

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
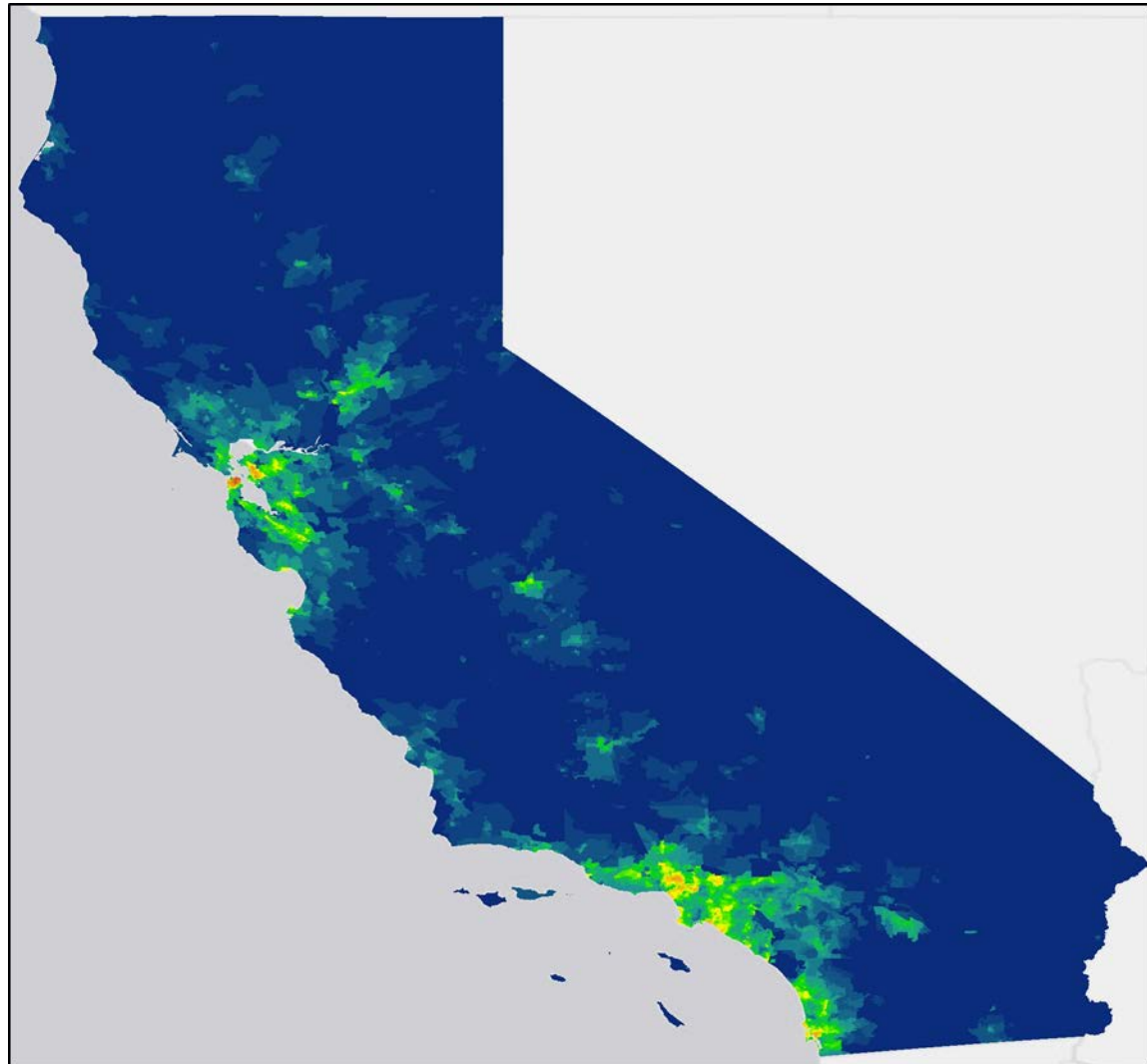
Review

- CHIT is (mostly) a set of custom tools built in ArcGIS ModelBuilder
- Some iterative
- Some nested
- Lots of Spatial Analysis
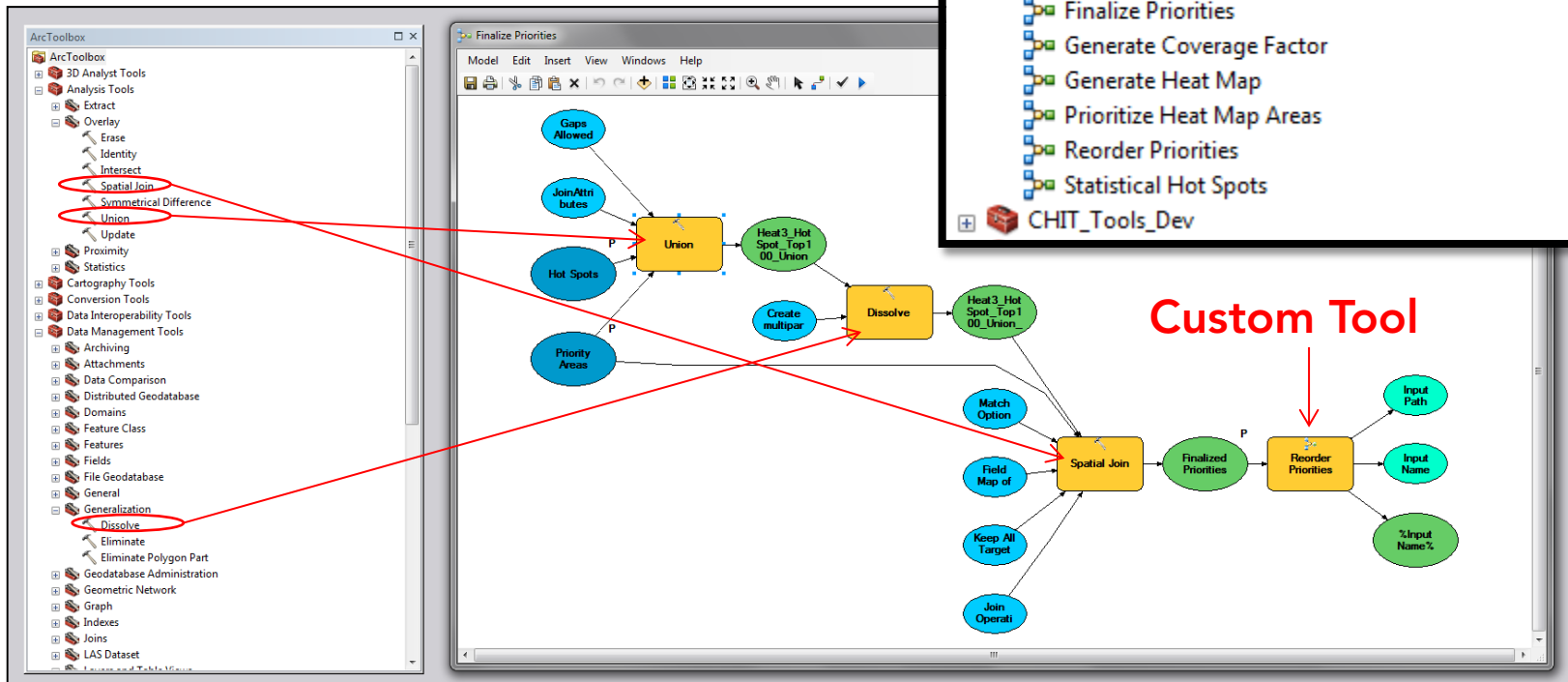
Guiding Principles

CHIT is envisioned as a tool that could be used year-after-year for public planning and reporting purposes, while providing a consistent assessment method across the entire state. Related fundamental principles guide development of CHIT:

- Principle #1: CHIT is a relative assessment
- Principle #2: CHIT is a statewide assessment
- Principle #3: CHIT assesses only the first adopter market
- Principle #4: The FCEV market can be estimated by the relative distribution of multiple demographic indicators
- Principle #5: Accurate assessment of coverage depends on detailed roadway data
- Principle #6: Coverage matches the market when it provides convenient fueling access near FCEV drivers' homes
- Principle #7: CHIT must be a tool that can be shared with the public

TRAFFIC DATA NEEDS

Today's Discussion

Big Picture Goal

Plan infrastructure placement appropriately for upcoming FCEV releases

How to define the market?

CHIT 2015: Identify the areas where likely first adopters live

Stakeholder Feedback: The market is mobile; where FCEV adopters drive matters too

Today's Discussion: Do we have sufficient data for the drive? How could we integrate this?
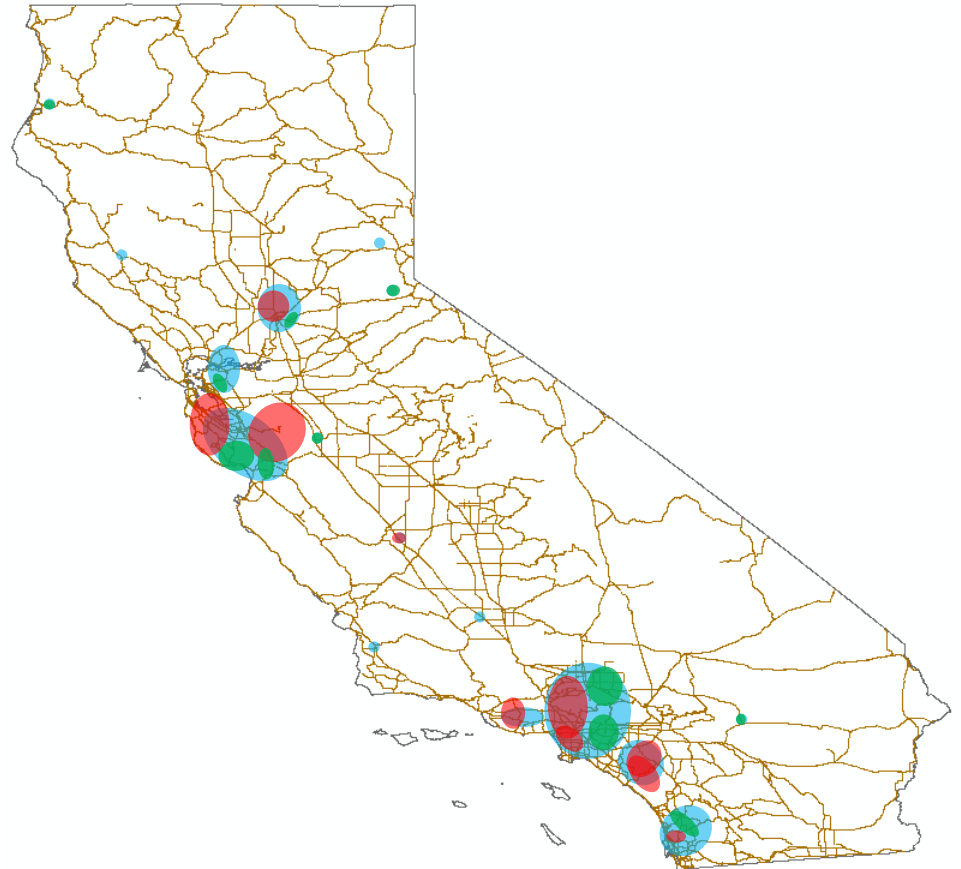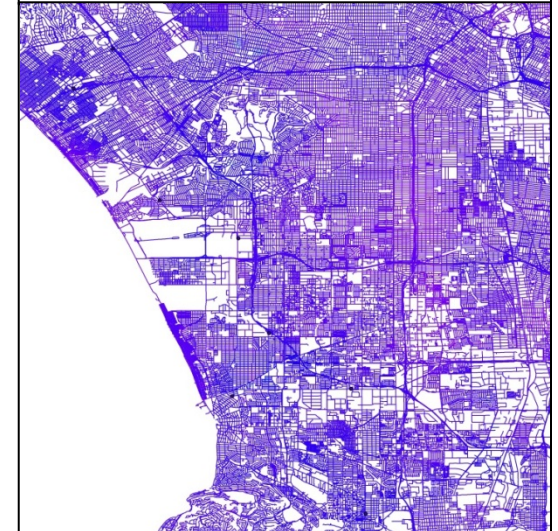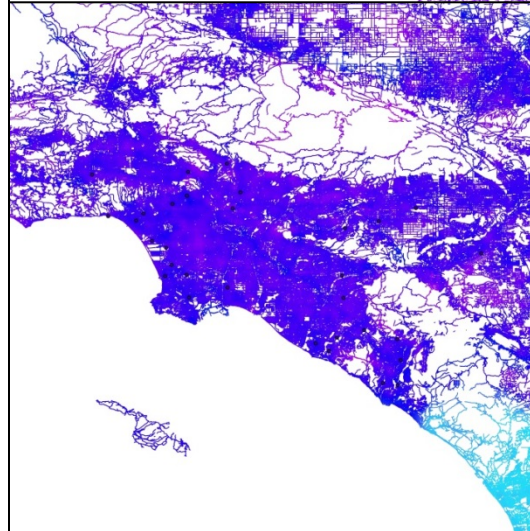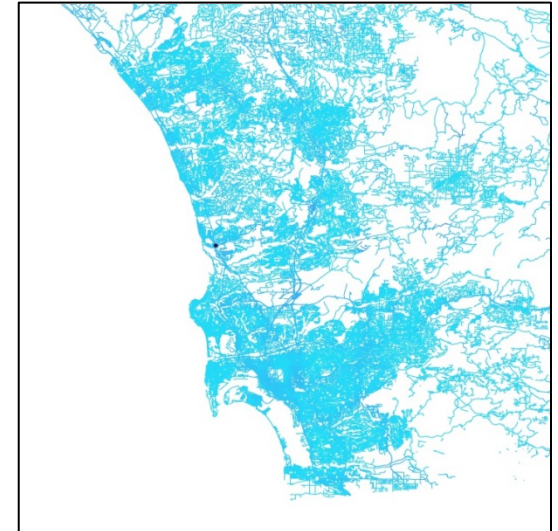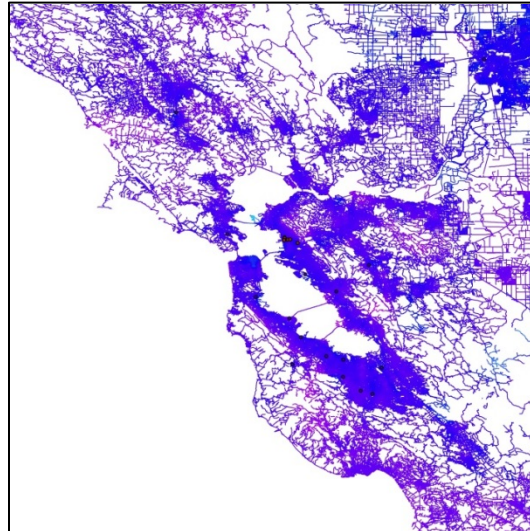
Traffic Volume

- Principle #6: Coverage matches the market when it provides convenient fueling access near FCEV drivers' homes

- Potential Addendum to Principle #6: Coverage is also defined by proximity to commonly-traveled routes

- Principle #4: Identification of the FCEV market can be estimated by consideration of the relative distribution of multiple demographic indicators

- Potential Clarification to Principle #4: Identification of the FCEV **owner** market can be estimated by consideration of the relative distribution of multiple demographic indicators; **the total fueling market potential can be estimated by consideration of both the owner market and commonly-traveled routes**

Prior Look at Traffic Volume

- Investigated utilizing volume data available in ITN

- Followed method similar to speed to estimate PM peak traffic

- ITN-TIGER data quality good for speed, not for volume in all areas

- Explicit and direct consideration requires origin-destination data sets

Prior Look at Traffic Volume

- Investigated other data sources
  - California Household Travel Survey
  - CalTrans Annual Average Daily Traffic
  - CHAPIS
- Principle #5 often not fulfilled: need data on neighborhood-level streets for accurate assessment of convenience and coverage
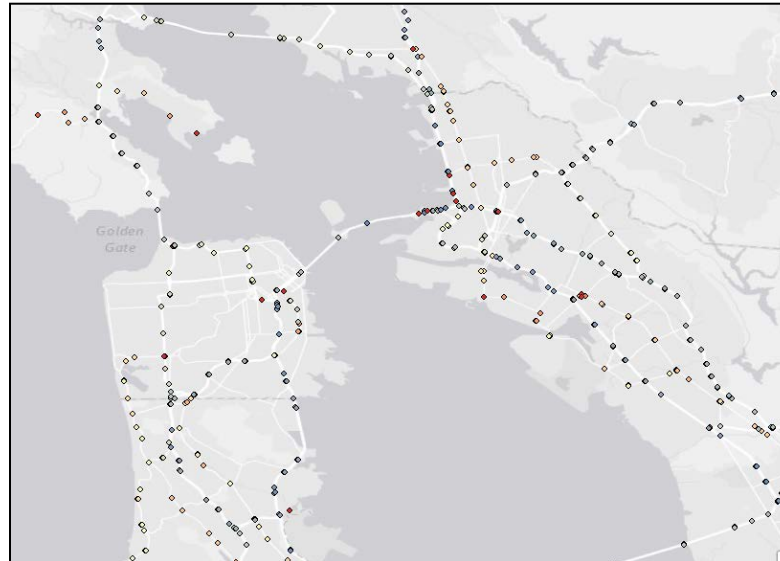
ARB'S TRAFFIC SIMULATION METHOD

Simulation Concept

- Given a sufficiently-detailed street geometry and travel speed data (provided by ITN-TIGER)…

- Availability of Origin-Destination data for commutes may enable route simulation…

- However, simulated data requires elevated scrutiny compared to observational data

- Major lead: Stakeholder suggestion at October 2015 webinar to look into LODES (LEHD Origin-Destination Employment Statistics) data set

Simulation Data Sources

ITN-TIGER (PM peak speeds shown)
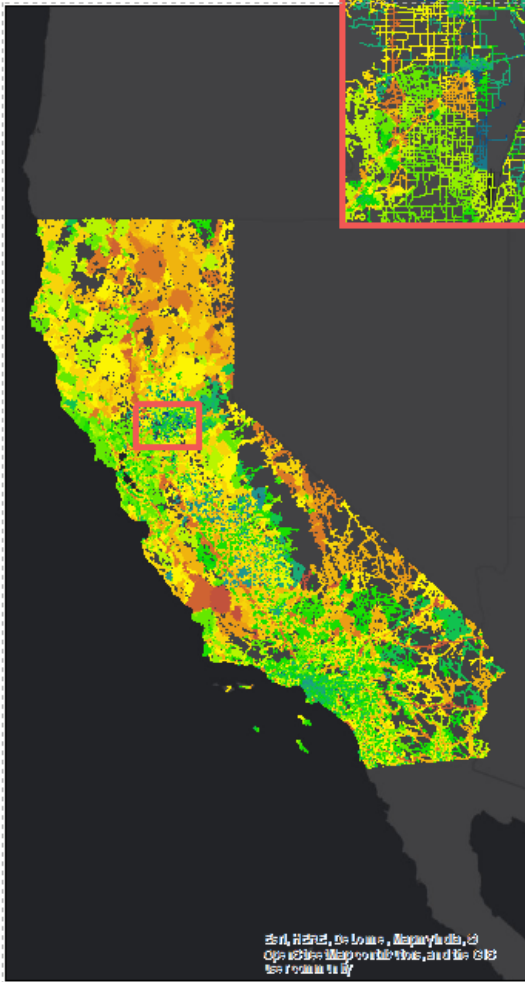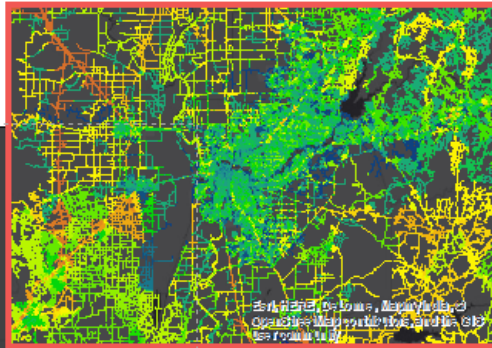
LODES (LEHD Origin-Destination Employment Statistics

- Pairs of home and work census blocks for all respondents
- Includes number of respondents with the same O-D pair
- Census block resolution not exact street address, but still very good
- ARB used the 2014 vintage

lodesca_nonull

Rowid	OBJECTID	W_GEOCODE	H_GEOCODE	S000	CREATEDATE	XW	YW	XH	YH	W_GEOCODES	H_GEOCODES
1	1	60014001001007	60014001001024	1	20150727	-122.235698	37.872464	-122.237696	37.861152	60014001001007	60014001001024
2	2	60014001001007	60014001001058	1	20150727	-122.235698	37.872464	-122.22611	37.852014	60014001001007	60014001001058
3	3	60014001001007	60014004001002	1	20150727	-122.235698	37.872464	-122.255348	37.852128	60014001001007	60014004001002
4	4	60014001001007	60014011002003	1	20150727	-122.235698	37.872464	-122.262336	37.830131	60014001001007	60014011002003
5	5	60014001001007	60014040002011	1	20150727	-122.235698	37.872464	-122.255377	37.824804	60014001001007	60014040002011
6	6	60014001001007	60014075001011	2	20150727	-122.235698	37.872464	-122.193887	37.770364	60014001001007	60014075001011
7	7	60014001001007	60014083002016	1	20150727	-122.235698	37.872464	-122.163857	37.771966	60014001001007	60014083002016
8	8	60014001001007	60014220001049	1	20150727	-122.235698	37.872464	-122.291631	37.858066	60014001001007	60014220001049
9	9	60014001001007	60014238003021	1	20150727	-122.235698	37.872464	-122.250086	37.853585	60014001001007	60014238003021
10	10	60014001001007	60014239021007	1	20150727	-122.235698	37.872464	-122.255373	37.854669	60014001001007	60014239021007
11	11	60014001001007	60014352002006	1	20150727	-122.235698	37.872464	-122.047884	37.681514	60014001001007	60014352002006
12	12	60014001001007	60014513002015	1	20150727	-122.235698	37.872464	-121.796301	37.683946	60014001001007	60014513002015
13	13	60014001001007	60090002202082	1	20150727	-122.235698	37.872464	-120.868282	38.126471	60014001001007	60090002202082
14	14	60014001001007	60133190003008	1	20150727	-122.235698	37.872464	-122.124557	38.002816	60014001001007	60133190003008
15	15	60014001001007	60133211013006	1	20150727	-122.235698	37.872464	-122.110607	37.993219	60014001001007	60133211013006
16	16	60014001001007	60133250003013	1	20150727	-122.235698	37.872464	-122.076419	37.930332	60014001001007	60133250003013
17	17	60014001001007	60133373005010	1	20150727	-122.235698	37.872464	-121.997642	37.947581	60014001001007	60133373005010
18	18	60014001001007	60133400014003	1	20150727	-122.235698	37.872464	-122.071278	37.918425	60014001001007	60133400014003
19	19	60014001001007	60133522011000	1	20150727	-122.235698	37.872464	-122.133137	37.858401	60014001001007	60133522011000
20	20	60014001001007	60133522012004	1	20150727	-122.235698	37.872464	-122.137776	37.845146	60014001001007	60133522012004
21	21	60014001001007	60133591023015	1	20150727	-122.235698	37.872464	-122.311633	38.007406	60014001001007	60133591023015
22	22	60014001001007	60133591033018	1	20150727	-122.235698	37.872464	-122.298059	37.998304	60014001001007	60133591033018
23	23	60014001001007	60133601013027	1	20150727	-122.235698	37.872464	-122.28442	37.985865	60014001001007	60133601013027
24	24	60014001001007	60133630001002	1	20150727	-122.235698	37.872464	-122.313539	37.975192	60014001001007	60133630001002
25	25	60014001001007	60770003001001	1	20150727	-122.235698	37.872464	-121.297406	37.958187	60014001001007	60770003001001
26	26	60014001001007	60855070011031	1	20150727	-122.235698	37.872464	-121.990924	37.227011	60014001001007	60855070011031
27	27	60014001001007	60855077033020	1	20150727	-122.235698	37.872464	-122.06804	37.307666	60014001001007	60855077033020
28	28	60014001001007	60952511001028	1	20150727	-122.235698	37.872464	-122.233616	38.095916	60014001001007	60952511001028
29	29	60014001001007	60971530034003	1	20150727	-122.235698	37.872464	-122.748063	38.432243	60014001001007	60971530034003

Block Resolution Samples

Urban Setting

Downtown and Midtown Sacramento

Block centers relatively good for downtown areas with large buildings that span city blocks

Less precise in dense residential, but wouldn't change results significantly

Rural Setting

East of Coalinga

Especially sparse division of area into census blocks in rural parts of the state
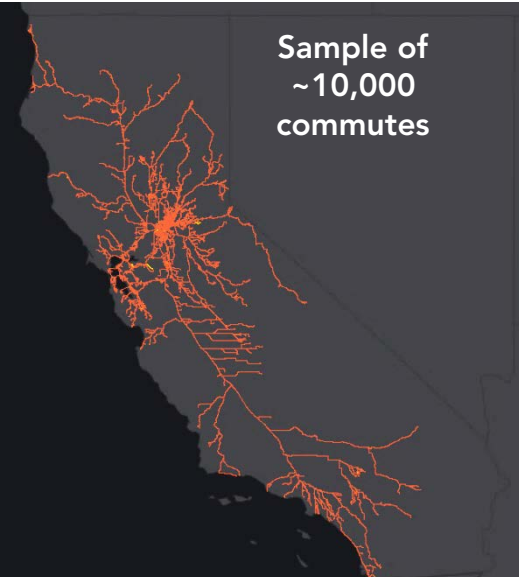
Less precise modeling of door-to-door travel

Method

- Step 1: Simulate all entries in the LODES data file
  - Geocode census block centroids and extract x/y to be able to load origins and destinations in Network Analyst
  - Use Network Analyst to simulate PM commute routes of all entries
    - Inherently assumes travel time optimization
  - Large processing effort: ~7.6 million records for ~8.36 million commutes
  - Executed via stand-alone arcpy across 10 processors on 2 computers
  - Batch processed 10,000 routes on each processor at a time: ~6 months
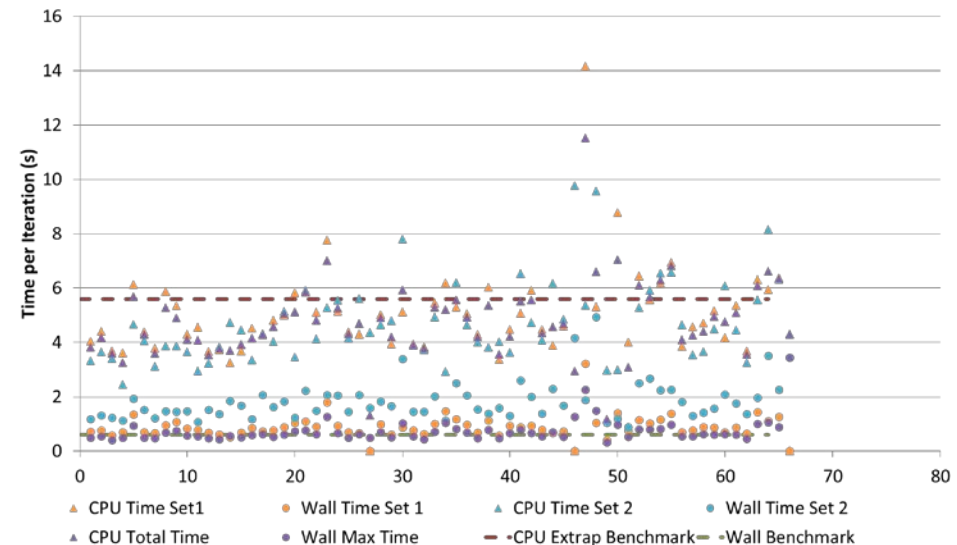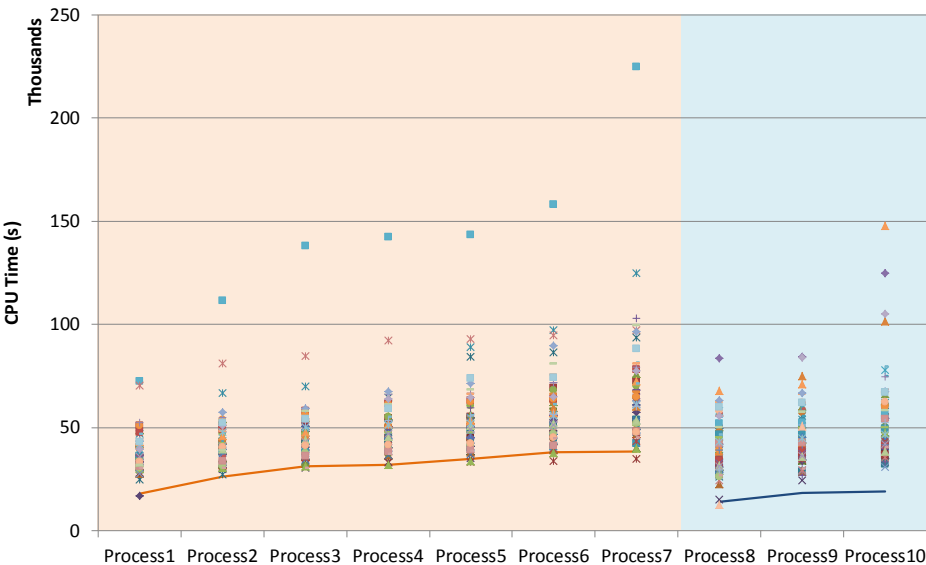
Method
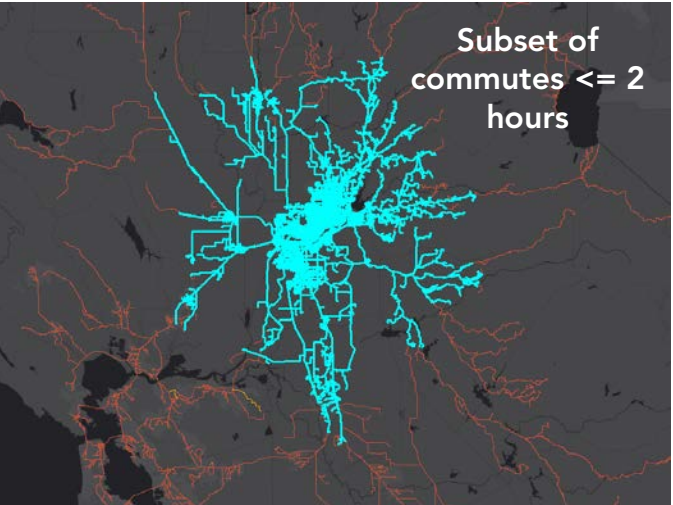
- Step 2: Combine data from resulting ~760 files
  - Down-select within each file to routes with 2 hours or less drive time
  - Spatially join with analysis grid
  - Keep running total of number of commutes through each cell
  - Smaller computing effort: python script executed in ArcMap
  - Batch processed over ~3 weeks

Sample of ~10,000 commutes

Bay-Sacramento Zoom

Subset of commutes <= 2 hours

Results

After down-selecting to commutes <=2 hr, simulations include ~7.14 Million commuters
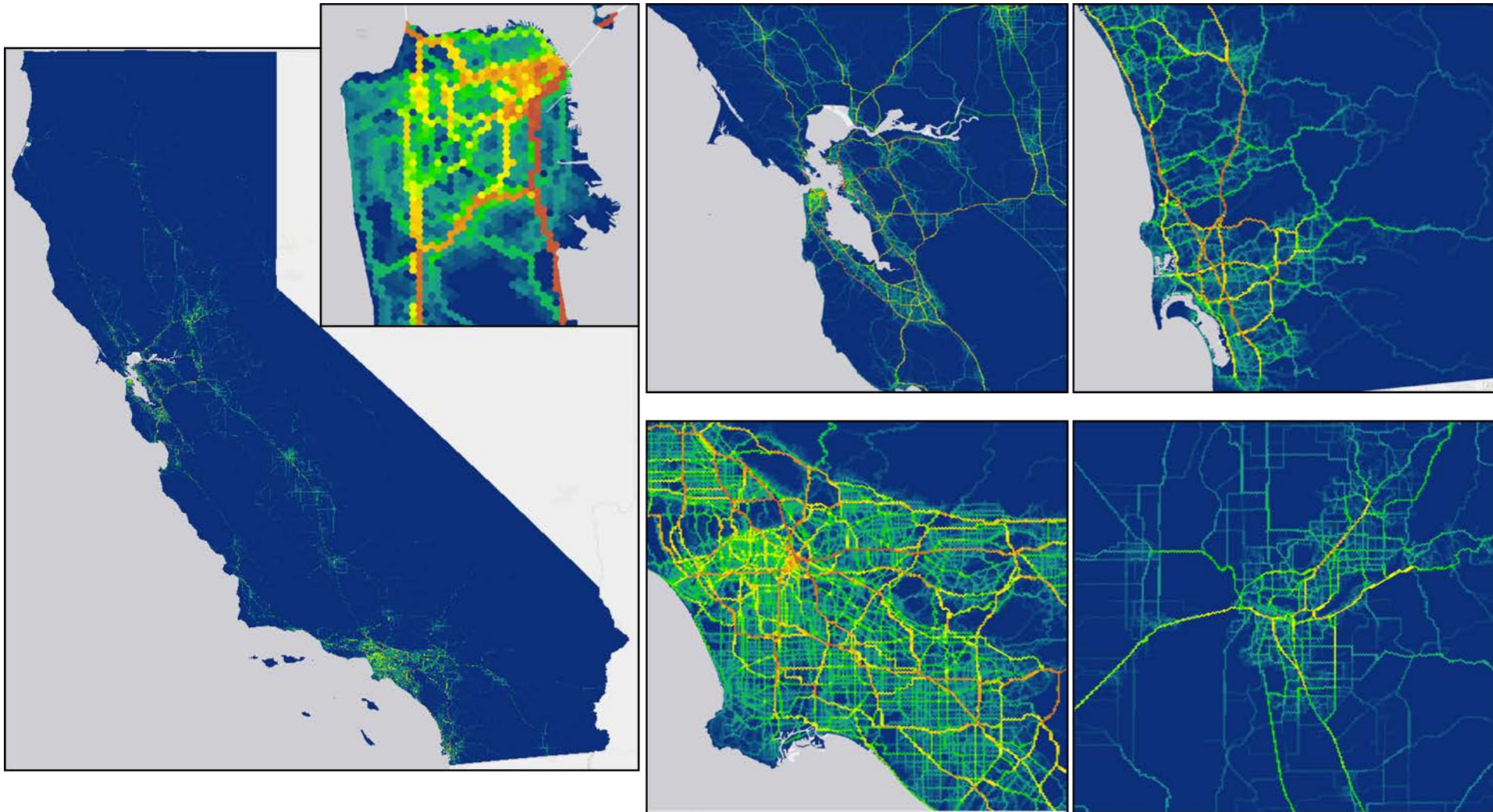
Crosscheck with CalTrans

- CalTrans maintains traffic census GIS data

- Data are sparse for the state:
  - Roughly 7,000 data capture points across state
  - Concentrated on major highways and freeways
  - Little data on lower-classification roads

- Can use to at least explore validity of simulation where data are available (on highways)

- Following slide shows three data points:
  - Smallest circles are CHIT simulation
  - Medium circles are CalTrans in "Back" direction
  - Large circles are CalTrans in "Ahead" direction
  - CalTrans data for peak hour in 2014 (data do not indicate PM or AM peak)
  - All data sets normalized to internal maximum (0-1 range to assess relative intensity)

Crosscheck with CalTrans
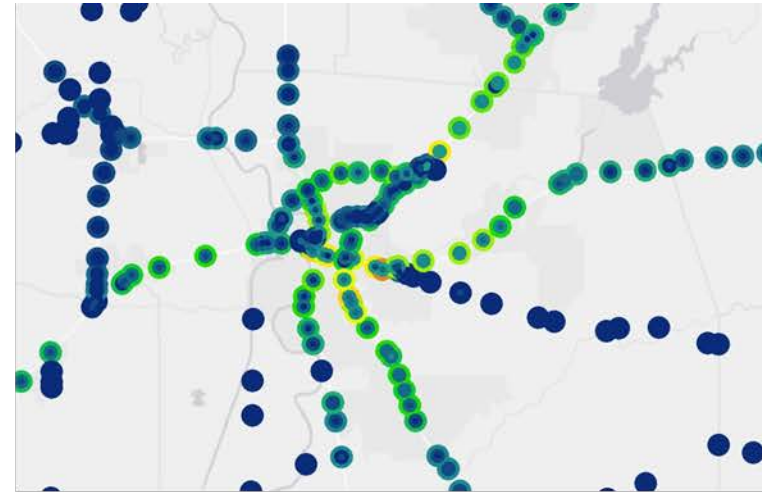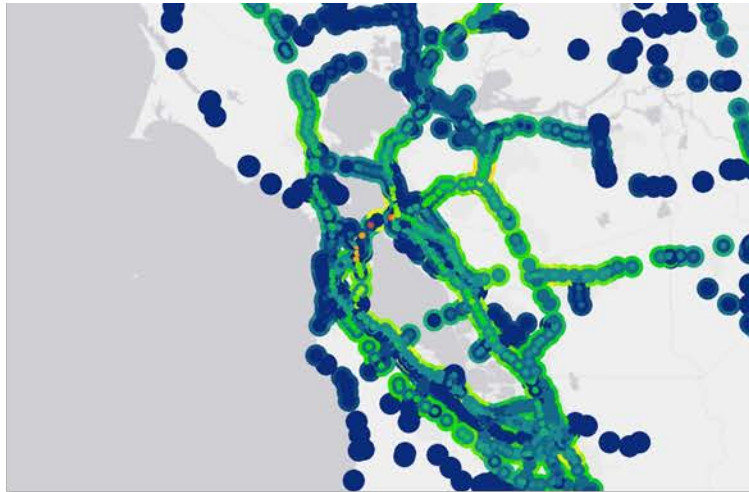
Crosscheck with CalTrans

Minimum Error:
- 0.000000 - 0.100000
- 0.100001 - 0.200000
- 0.200001 - 0.300000
- 0.300001 - 0.400000
- 0.400001 - 0.500000
- 0.500001 - 0.600000
- 0.600001 - 0.700000
- 0.700001 - 0.800000
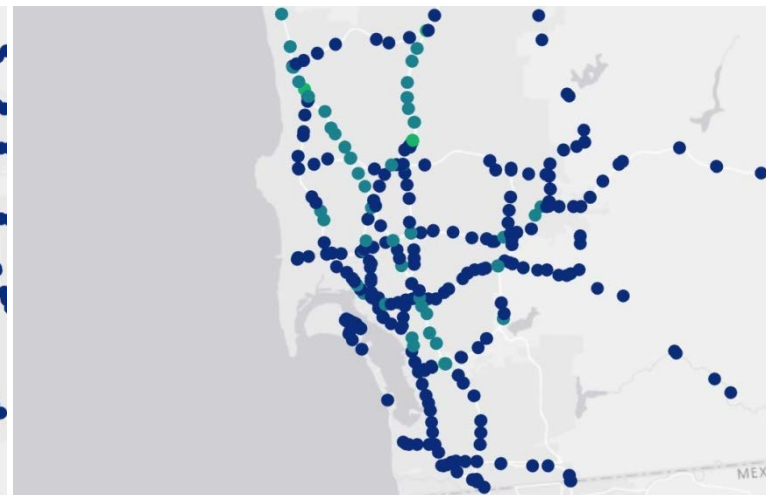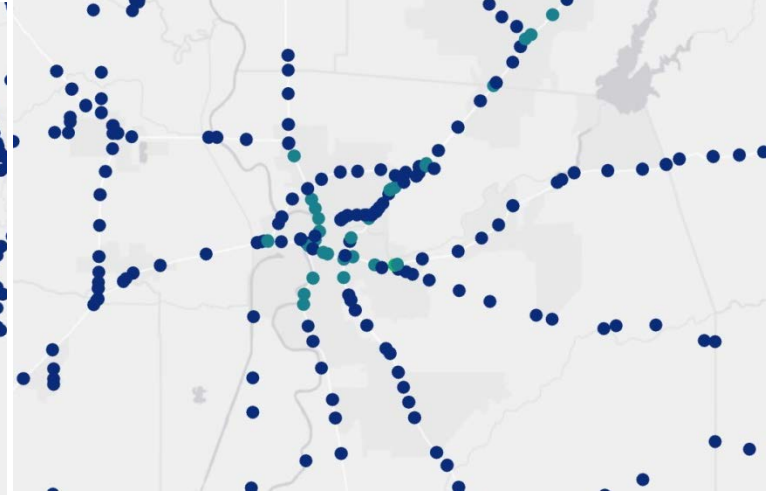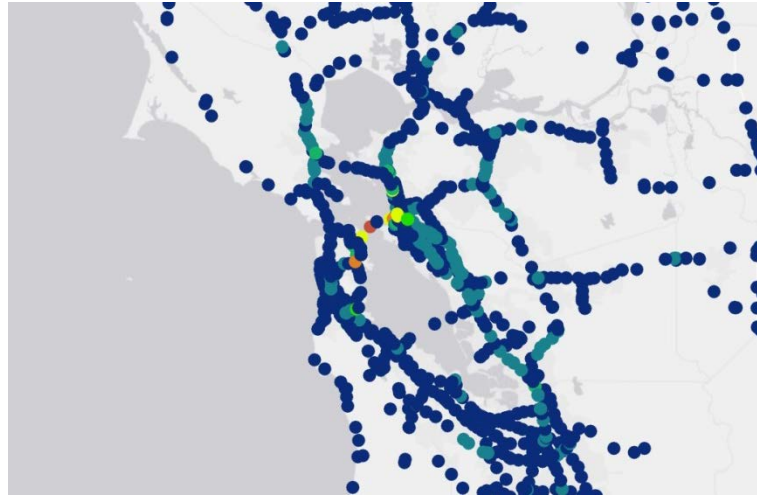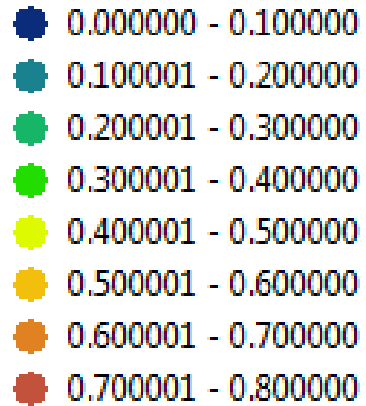
Crosscheck with CalTrans

Locations with error >20 points
Limited to ~100 locations
(>30 points error, ~40 locations)
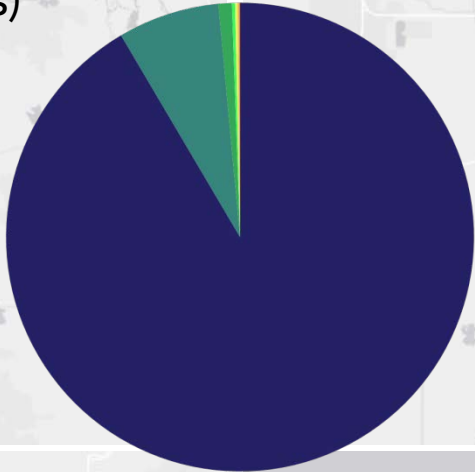
Minimum Error:
- 0.000000 - 0.100000
- 0.100001 - 0.200000
- 0.200001 - 0.300000
- 0.300001 - 0.400000
- 0.400001 - 0.500000
- 0.500001 - 0.600000
- 0.600001 - 0.700000
- 0.700001 - 0.800000
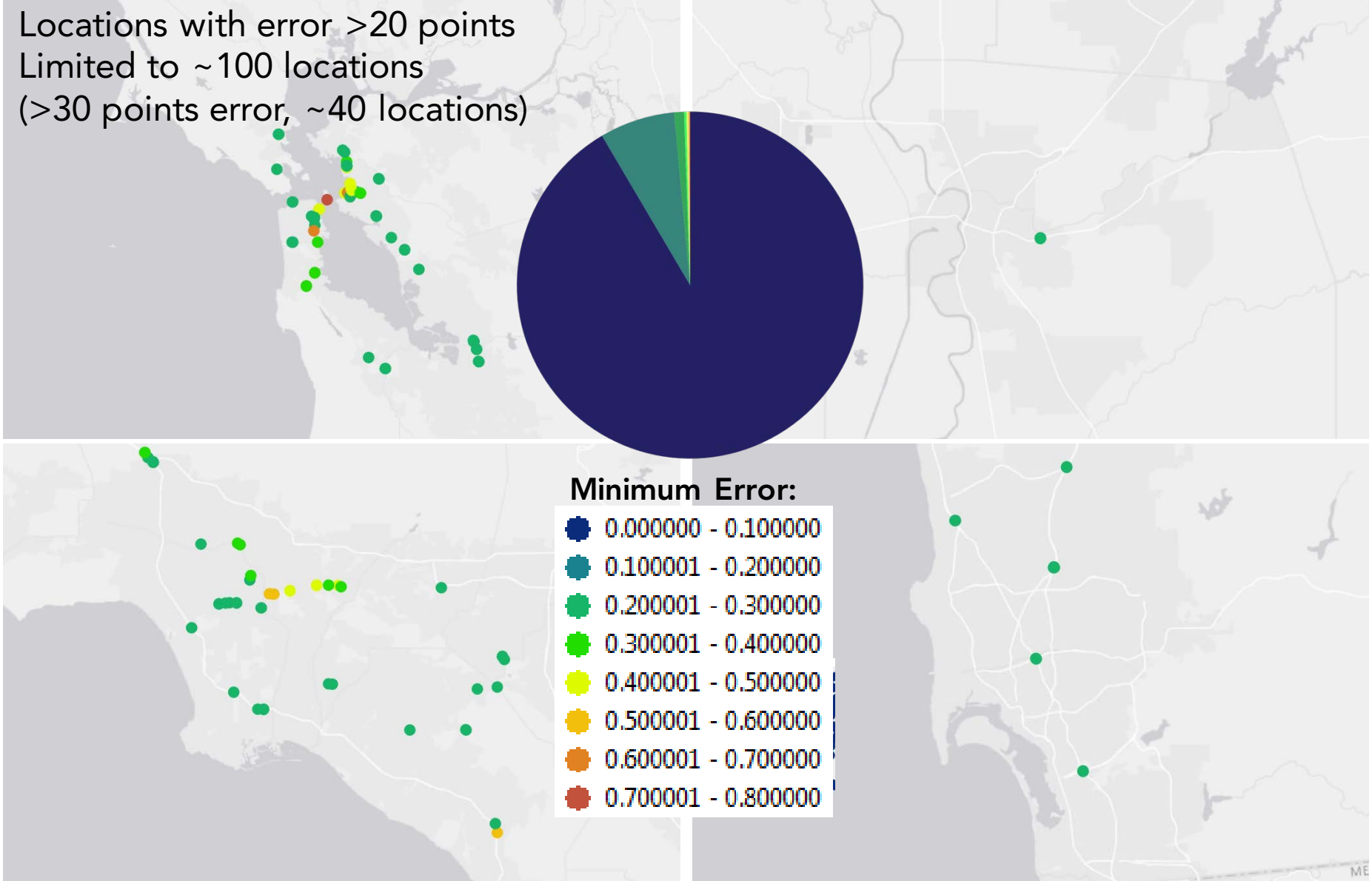
Crosscheck with CalTrans

Equal Interval Scale Classification
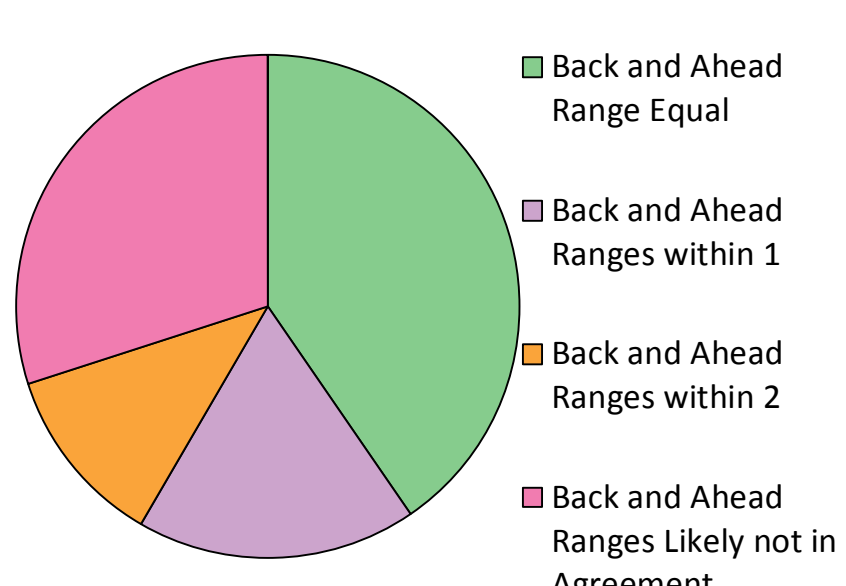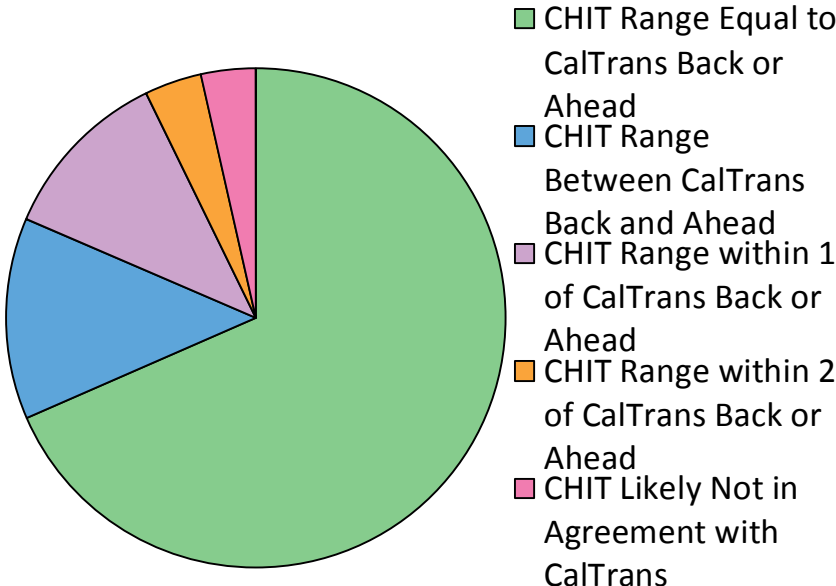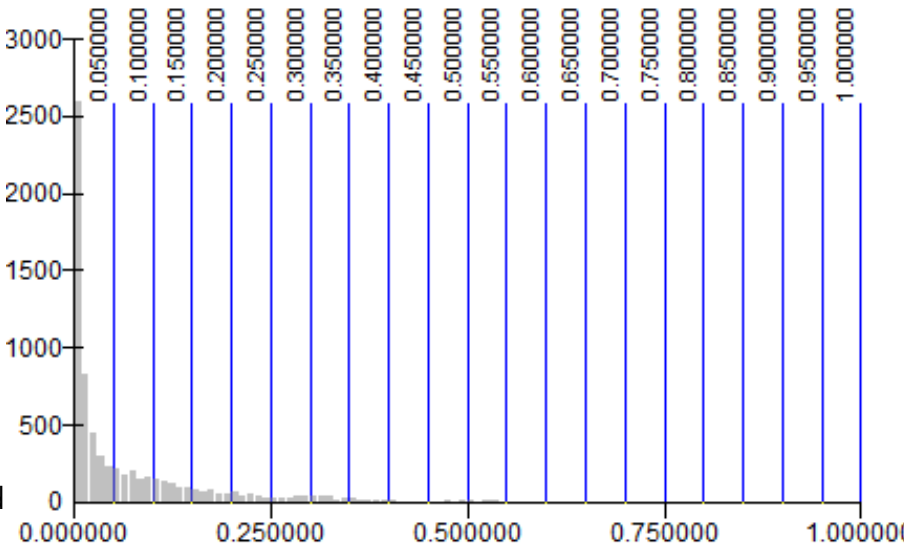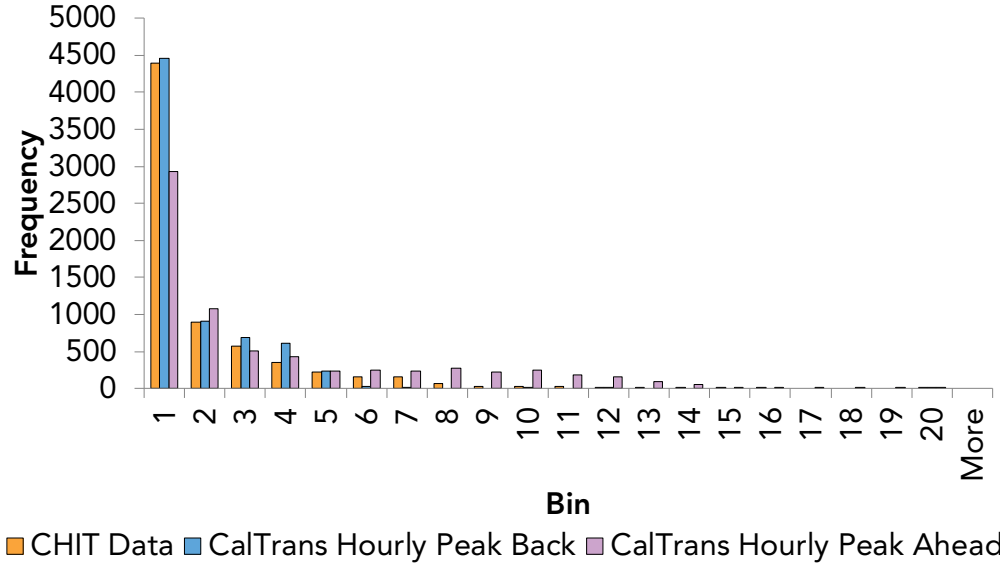
Crosscheck with CalTrans
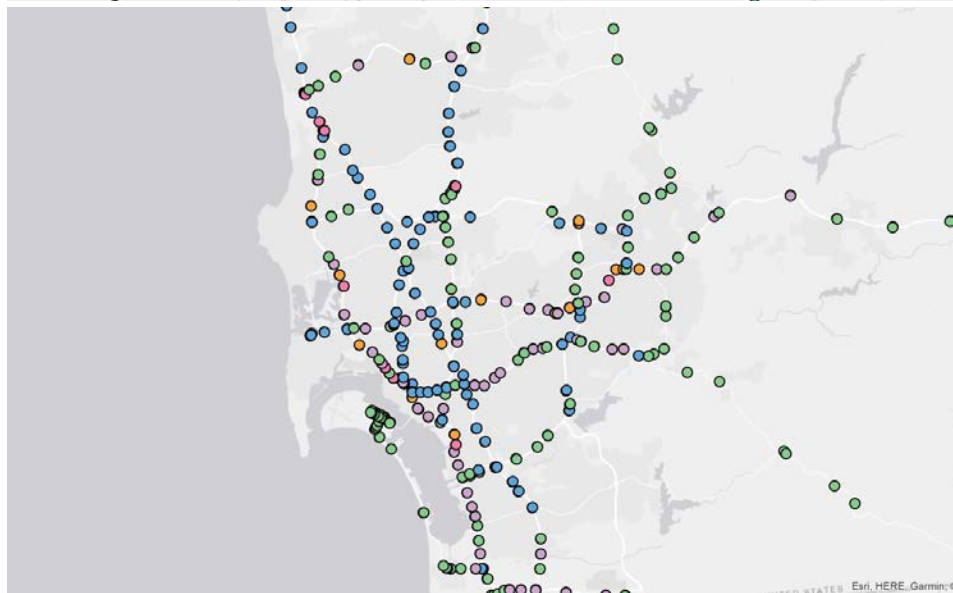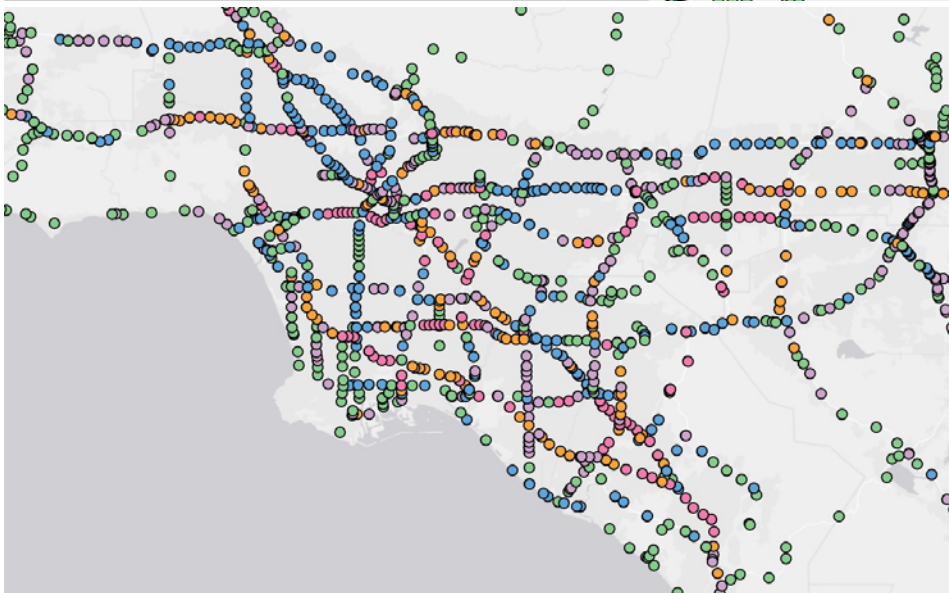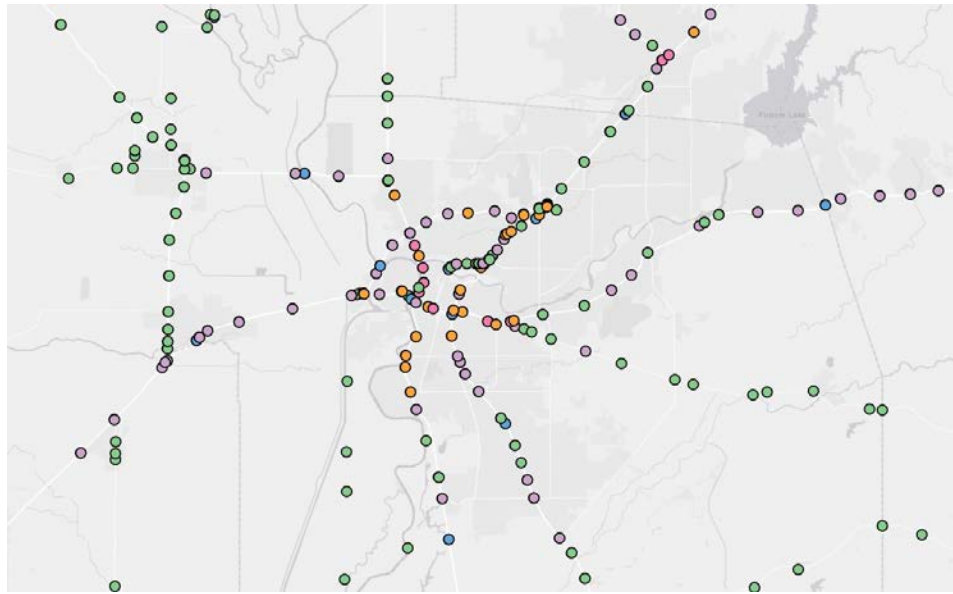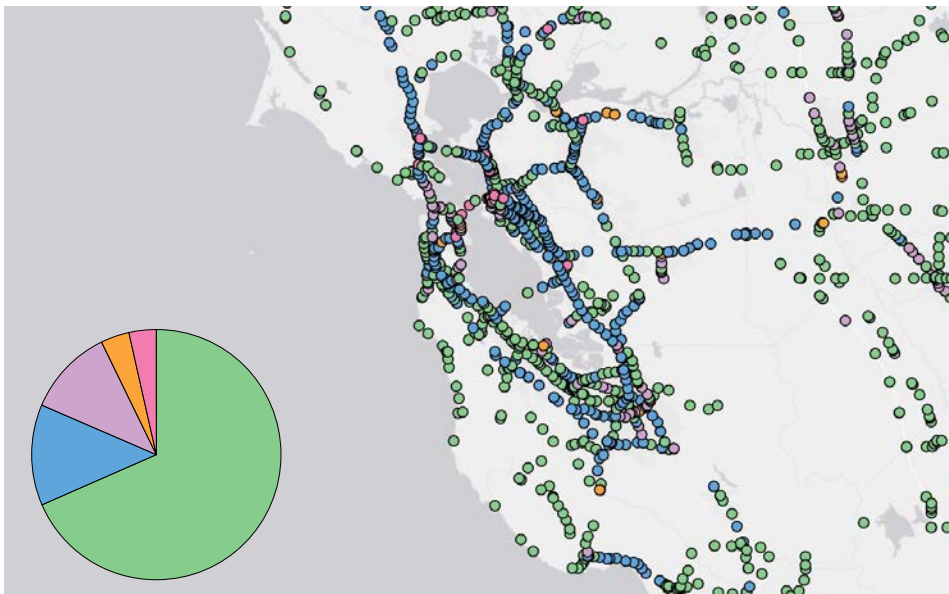
Crosscheck with CalTrans
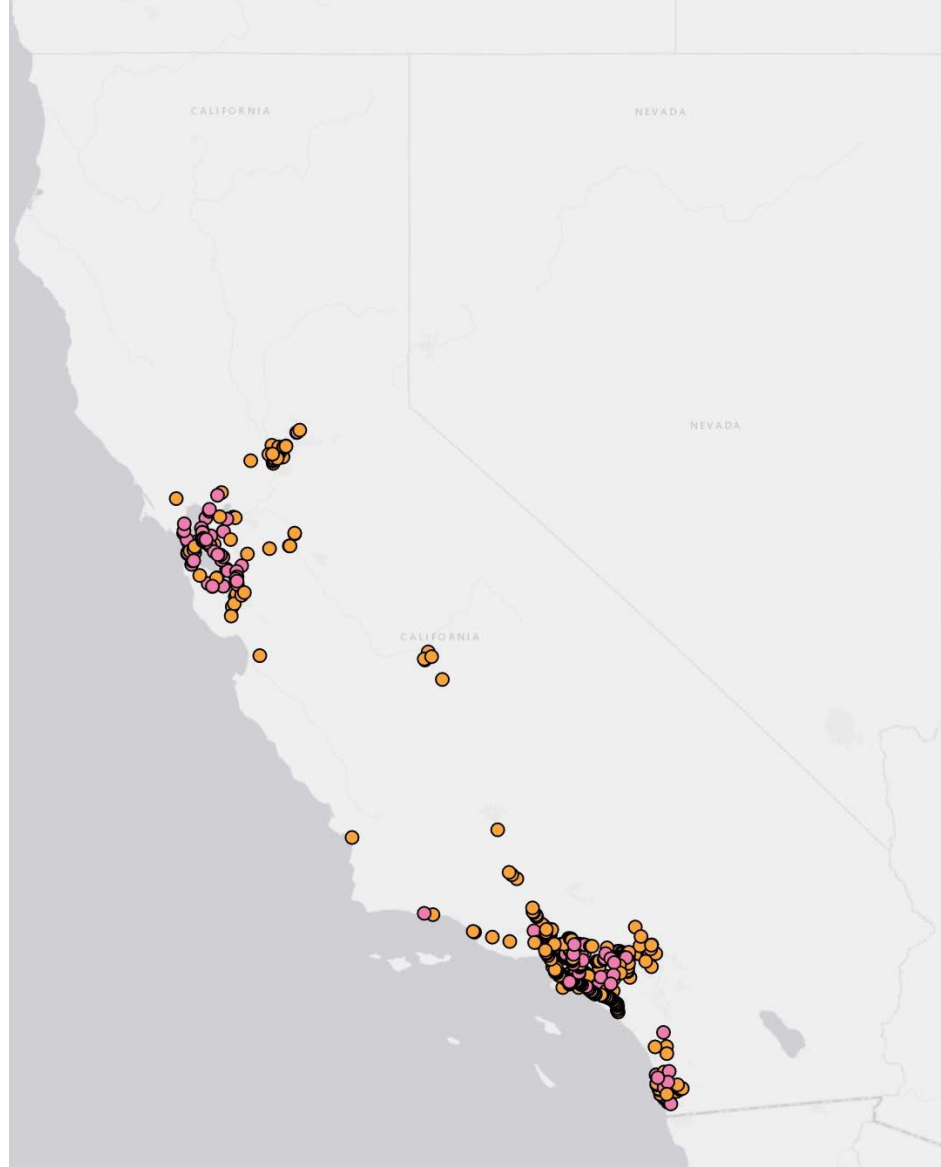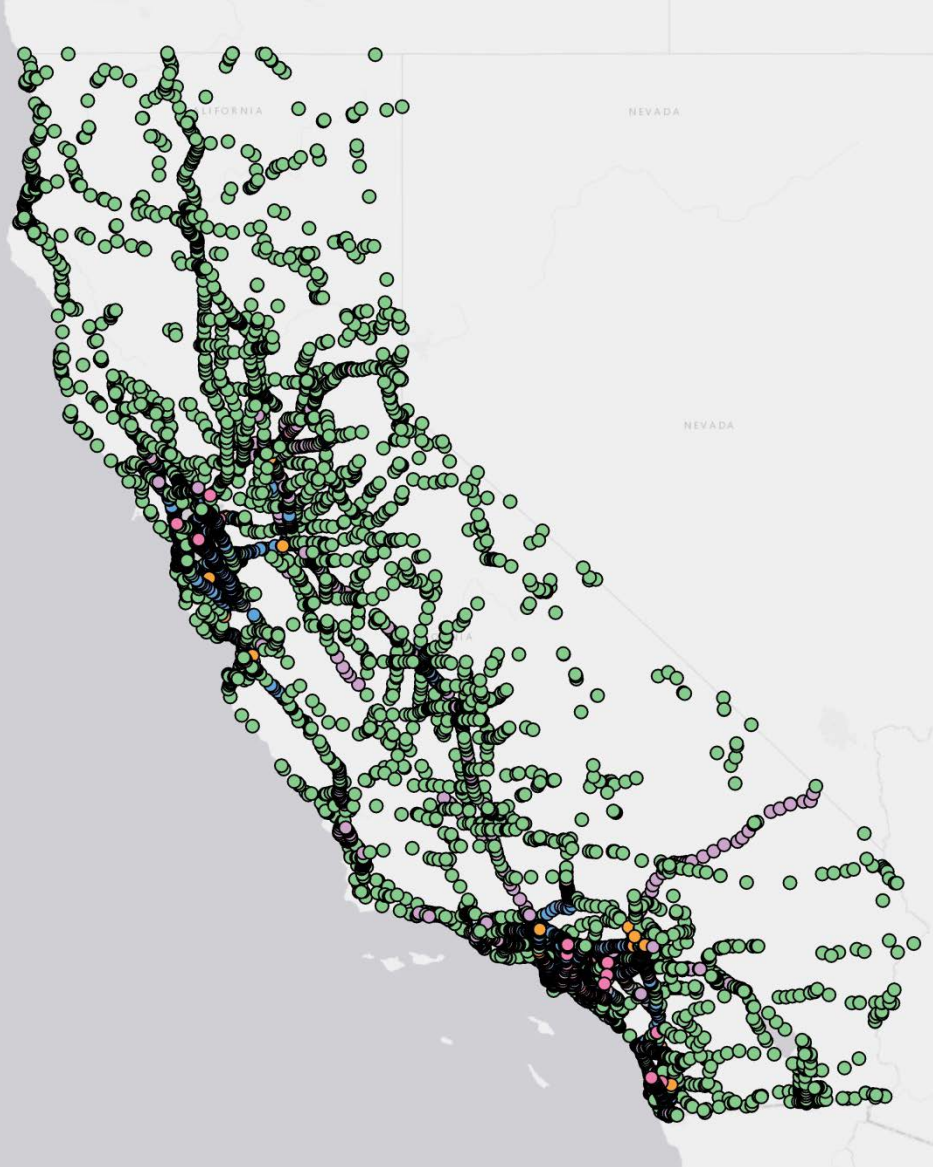
Recap of Simulated Data Set

- What the simulated traffic data **DOES** provide:
  - Detailed routing information, based on speed data at fine resolution
  - The <u>optimized</u> travel route for the simulated commuters, given the input speed data
  - A reasonably-sized sample of the full labor force
    - US Bureau of Labor Statistics estimates ~17.6 million in California's employed work force in late 2014
  - An estimate of what the relative commuter traffic intensity <u>might be</u> in different locations throughout the state
  - A unified estimate that is similarly developed for all locations in the state
  - An estimate of traffic volume that is self-consistent with traffic speeds already implemented in CHIT
  - Cross-checking with limited highway observational data shows fairly convincing agreement with CHIT
  - Differences appear to be regionally-specific

Recap of Simulated Data Set

- What the simulated traffic data **DOES NOT** provide:
  - Observed, actual routes of commuters
  - The full commuter travel pattern
  - Travel patterns of non-commuter vehicles
    - Occupational, Medium-Duty, Heavy-Duty, and other vehicles not included
    - The traffic patterns of these vehicles can affect Light-Duty commuter traffic
    - The intricacies and interactions are not accounted for
  - An assessment of commuter travel via public transportation or other modes
  - Any indication of the travel routes specifically utilized by first adopters
  - Exact door-to-door routes
  - Cannot assess validity of CHIT simulation on lower-classification roads

USING TRAFFIC AND FIRST ADOPTER MARKETS IN CHIT

Revised CHIT Coverage Gap Equation

- CHIT 2015:

  Heat = Coverage *
      Market

  Coverage= 0.5 * Existing +
      0.5 * Potential

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

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

- CHIT 2017 w/ Traffic:

  Heat= Coverage *
      (x * Market +
      y * Traffic Volume)

  Coverage= 0.5 * Existing +
      0.5 * Potential

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

  Financial= 0.34 * Income +
      0.33 * MSRP +
      0.33 * Luxury
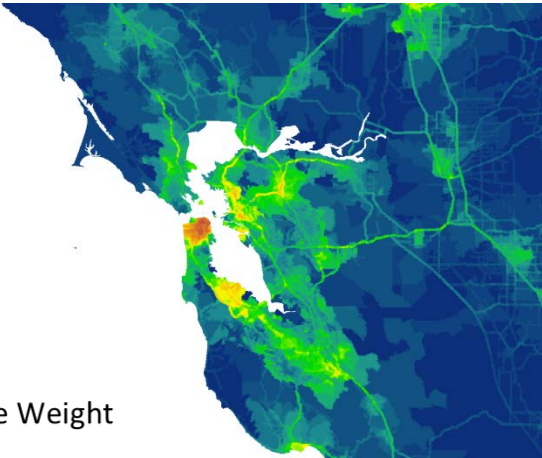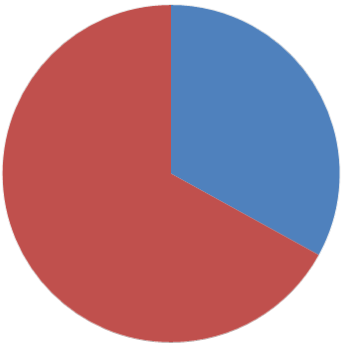
Effect in CHIT
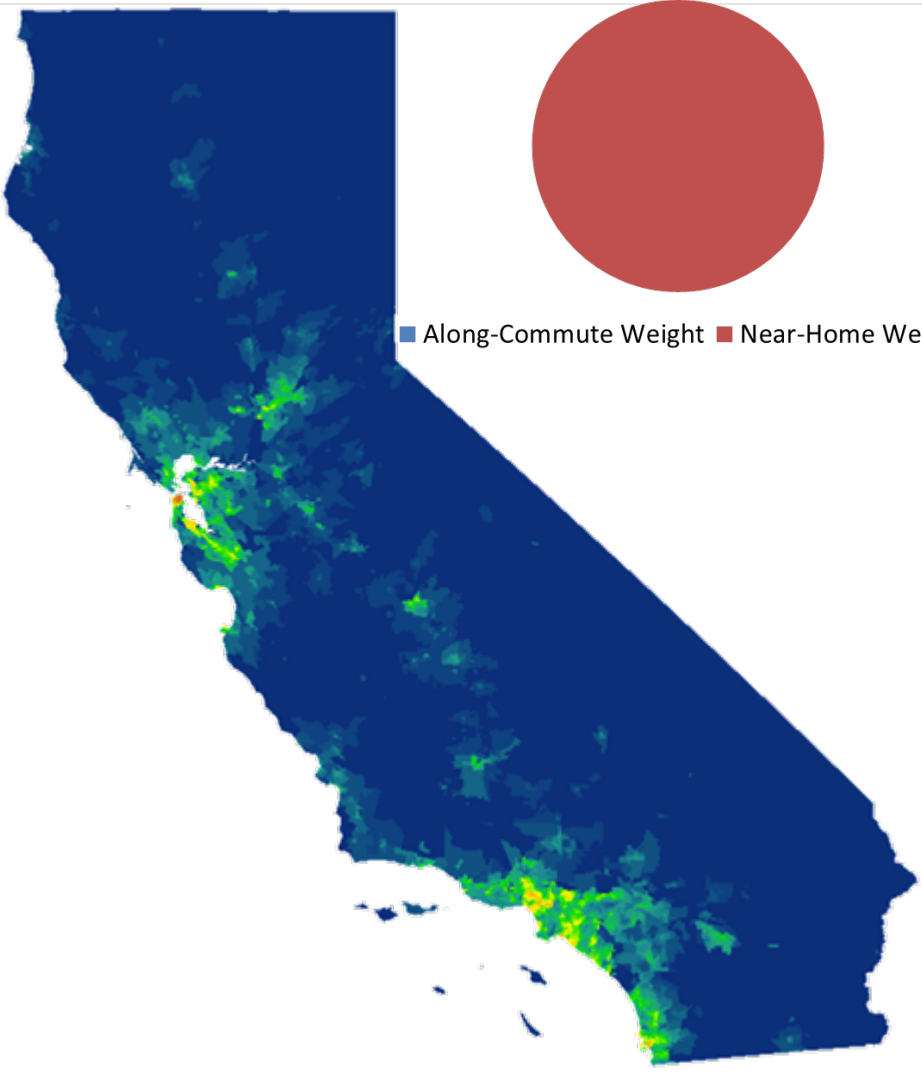
Along-Commute Weight ■ Near-Home Weight
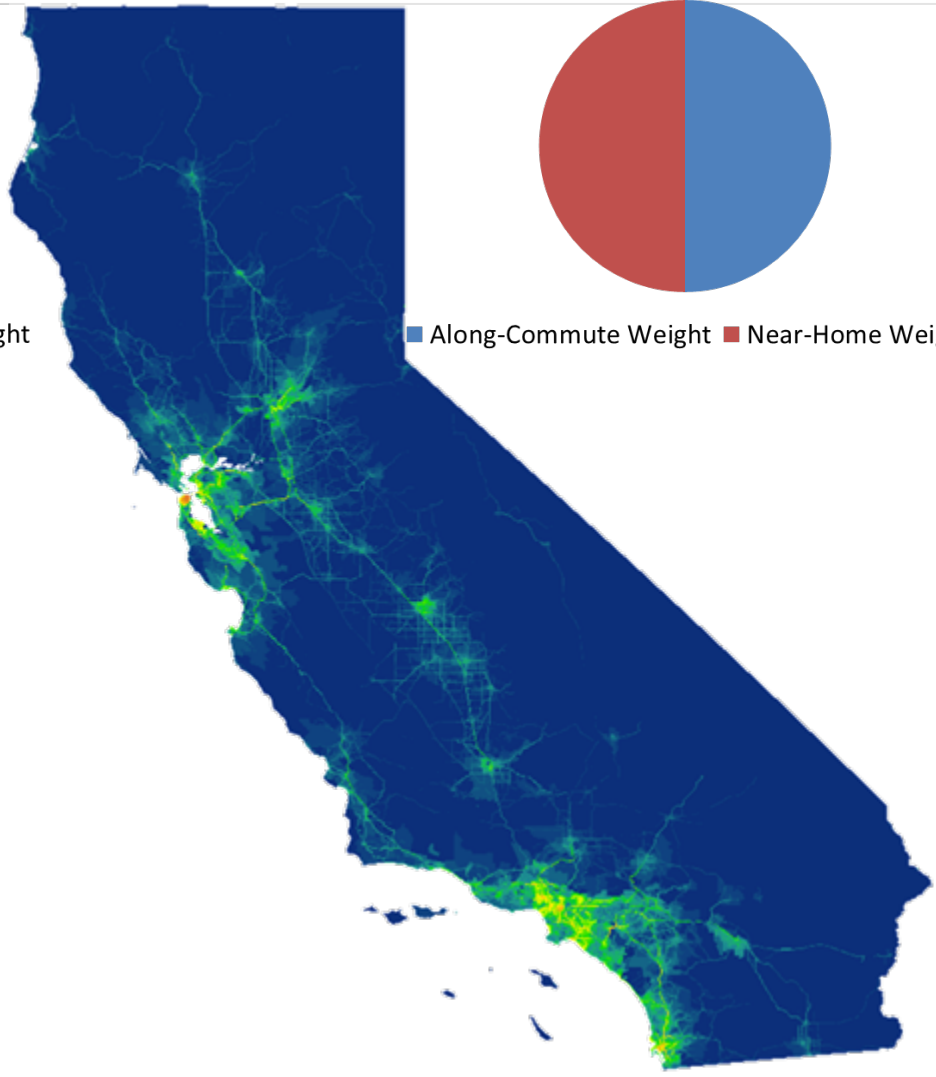
Along-Commute Weight ■ Near-Home Weight
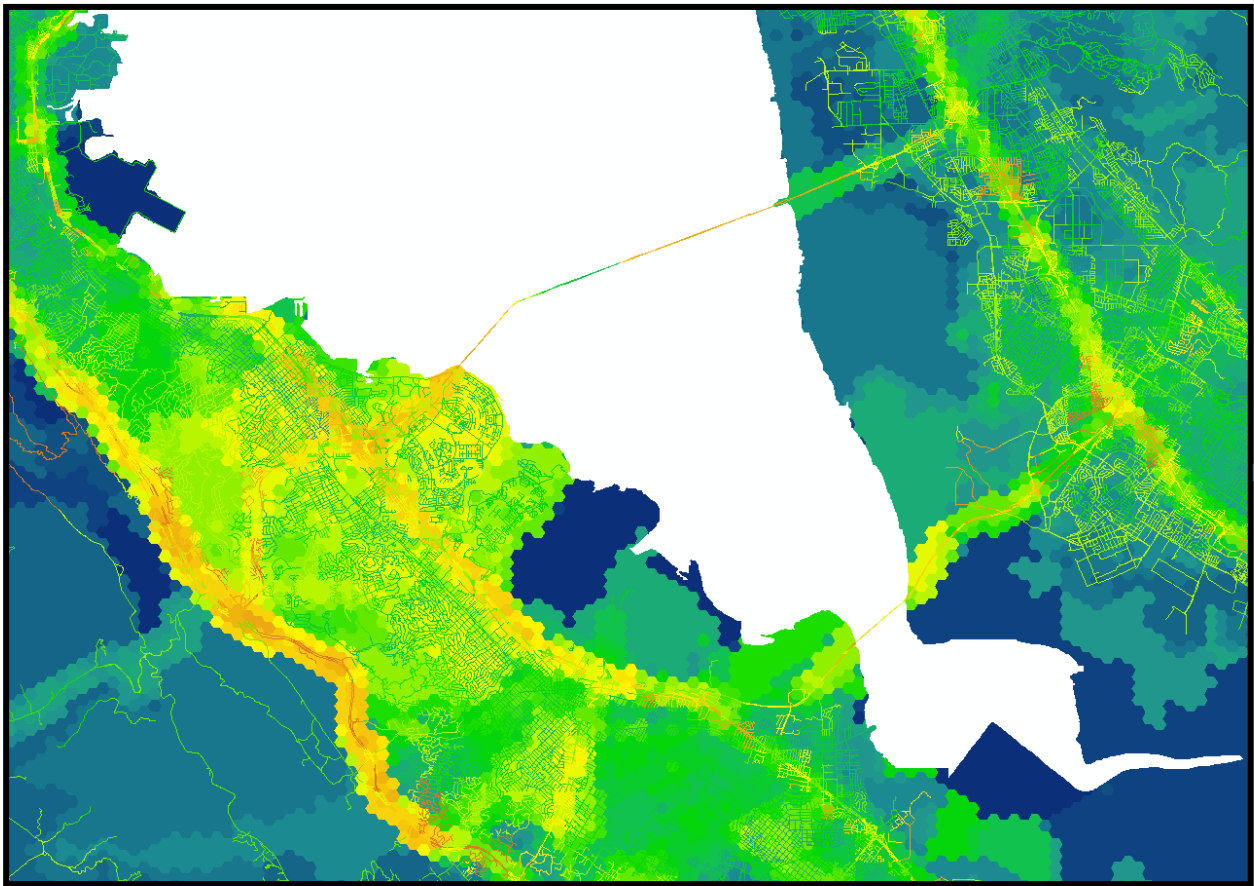
Effect in CHIT

Along-Commute Weight   Near-Home Weight

Effect in CHIT

Along-Commute Weight   Near-Home Weight
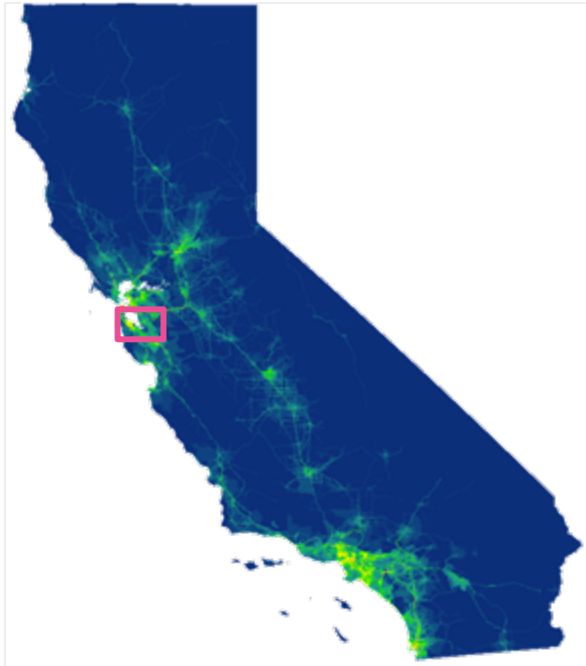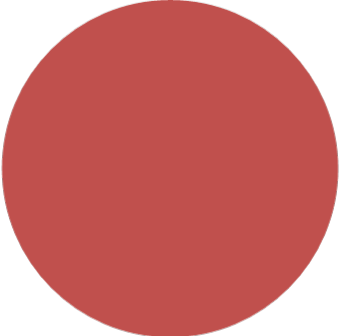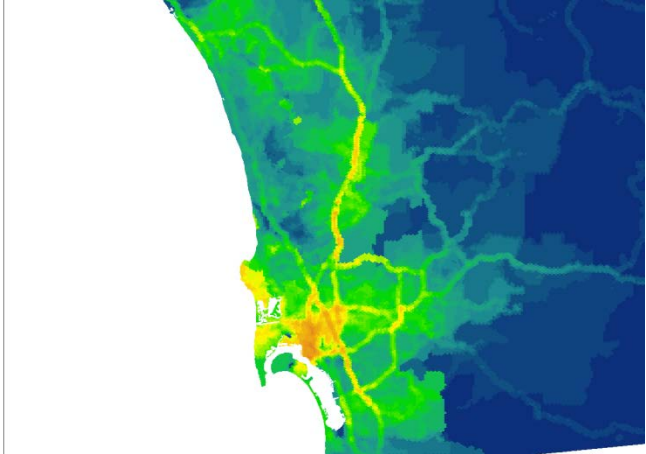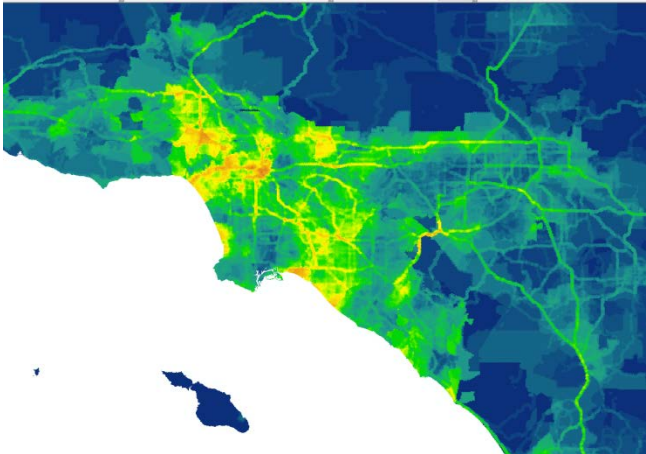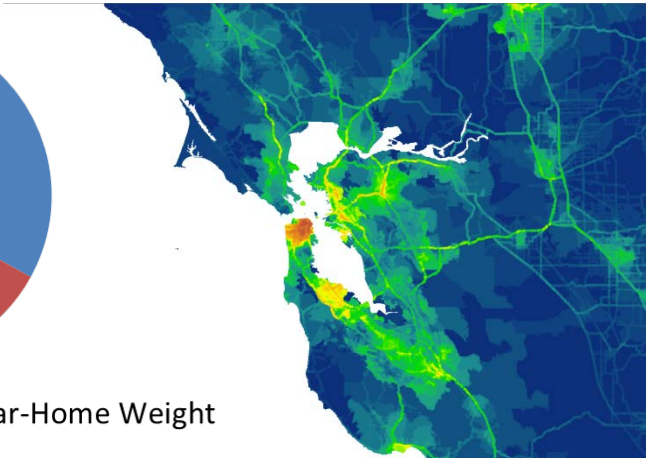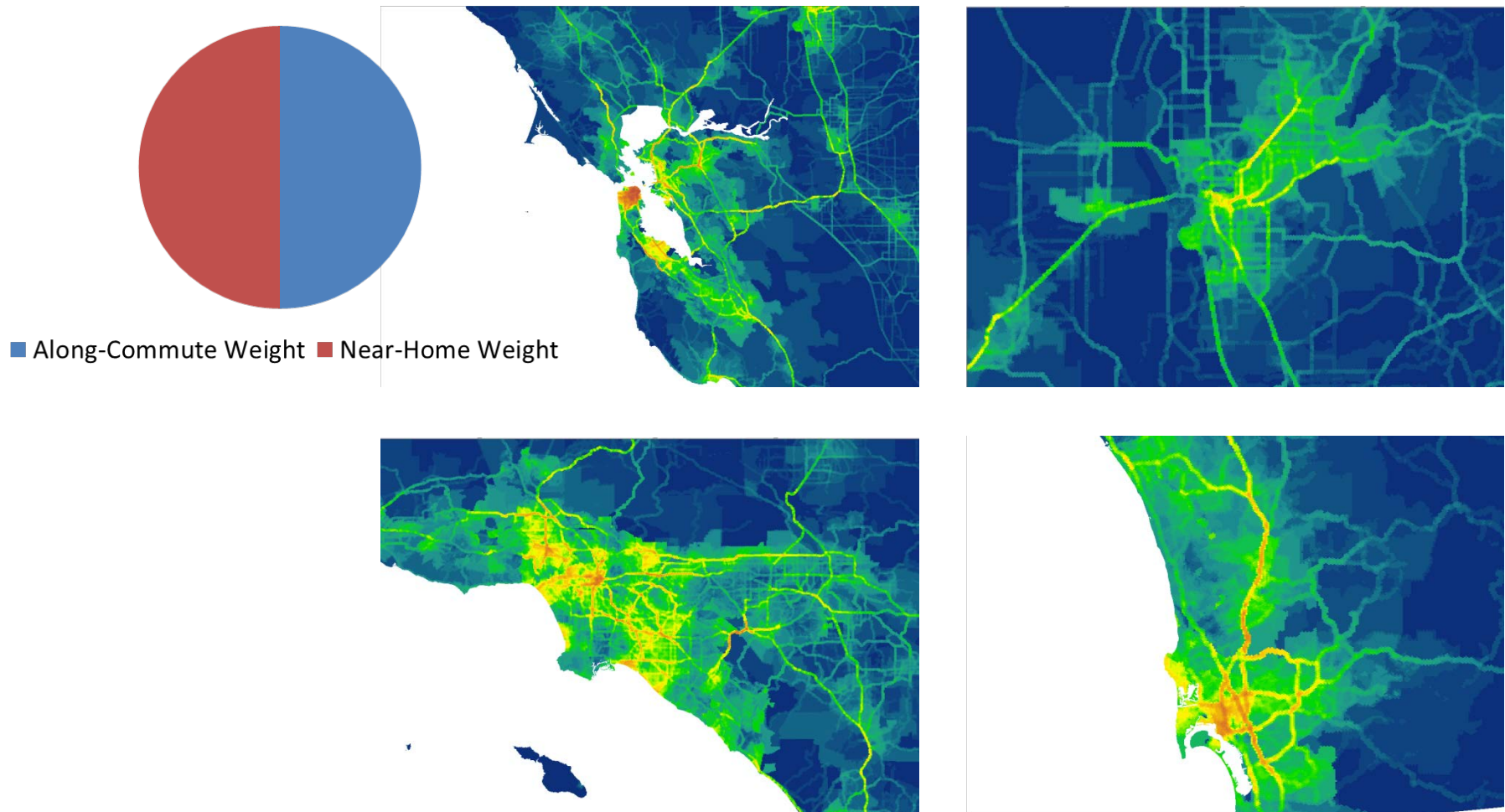
Effect in CHIT

■ Along-Commute Weight ■ Near-Home Weight
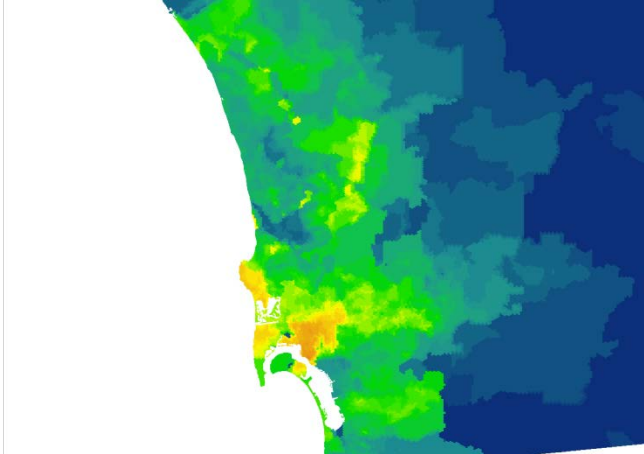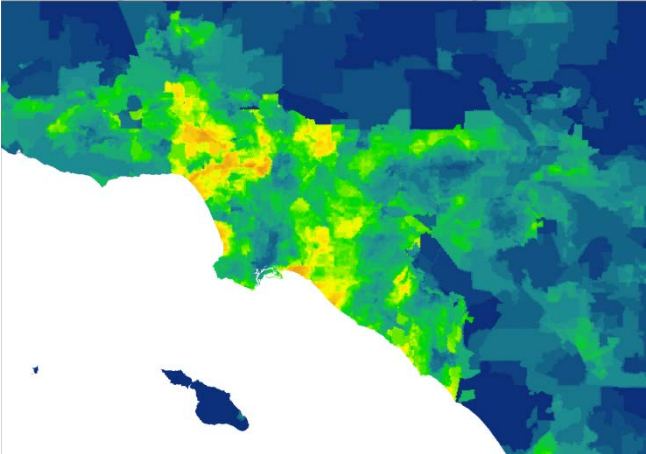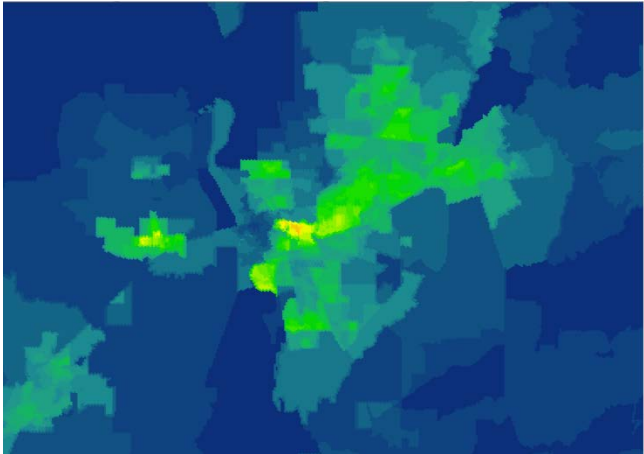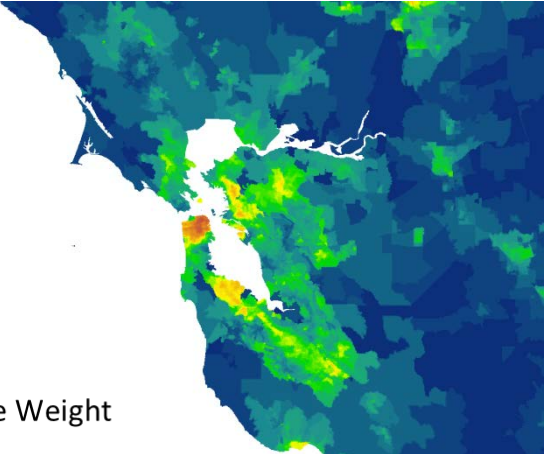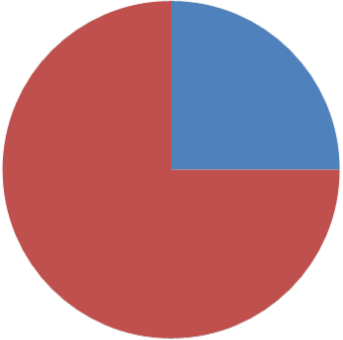
Effect in CHIT

■ Along-Commute Weight ■ Near-Home Weight
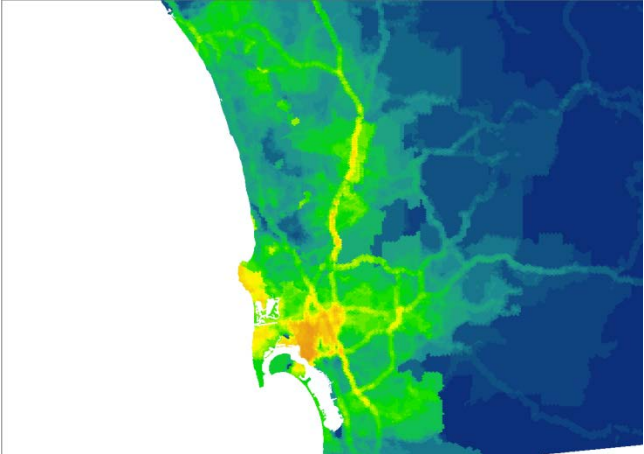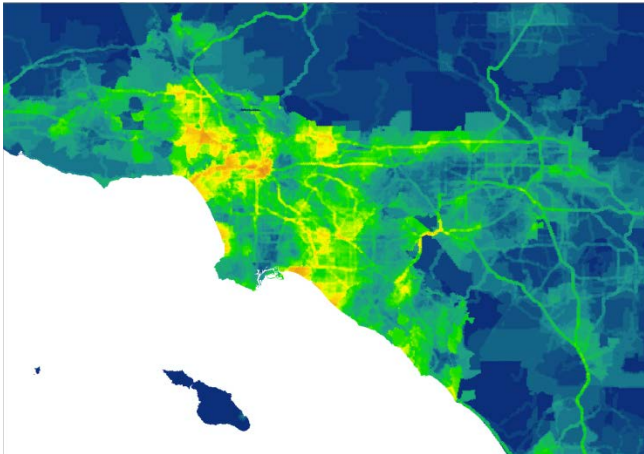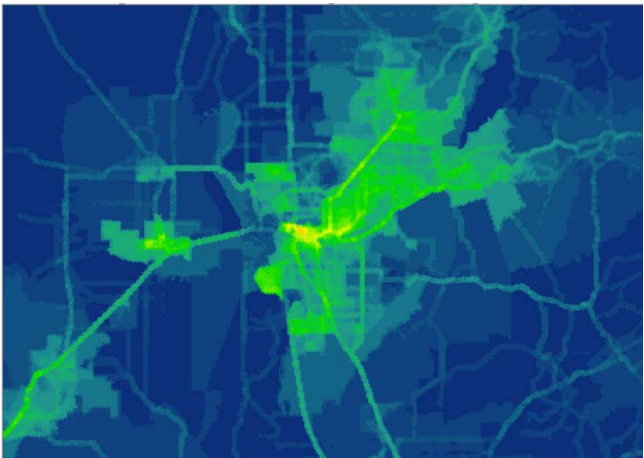
Effect in CHIT

Along-Commute Weight Near-Home Weight

Remaining Considerations

- With a new Coverage Gap (Heat) equation defined to include traffic, analysis of the map and identification of priority areas will likely need to change

- Traffic volume "hot spots" tend to be narrow, small-area stretches unlike hot spots based only on first adopter home locations
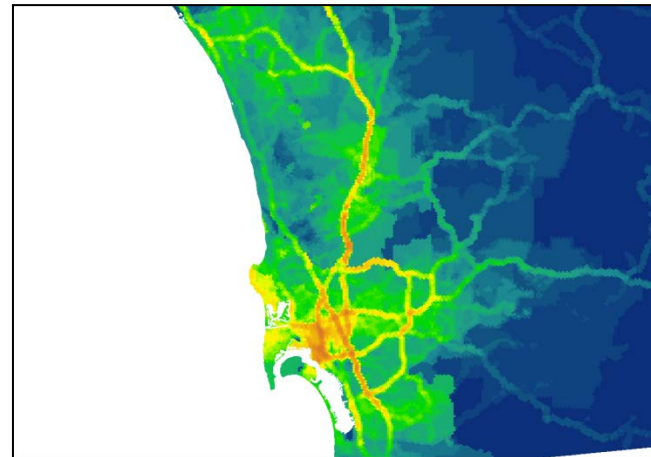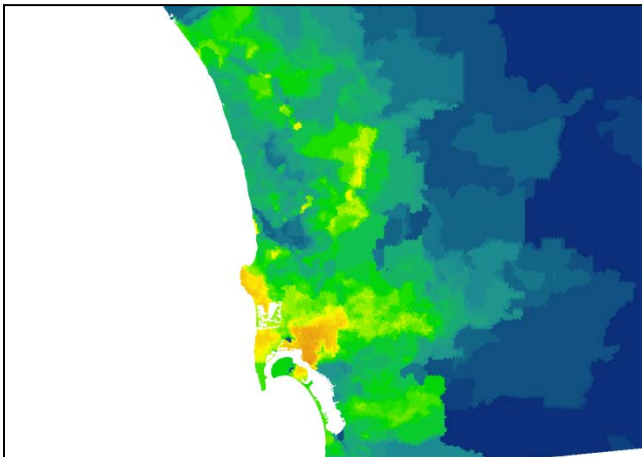
Remaining Considerations

- Narrow and small regions are more difficult to recommend because they provide less flexibility in detailed site selection
- Disconnected home-based hotspots may be joined by high-travel routes
- New map analysis tools may need to be developed
  - May do separate analysis and describe Priority Areas and Corridors separately
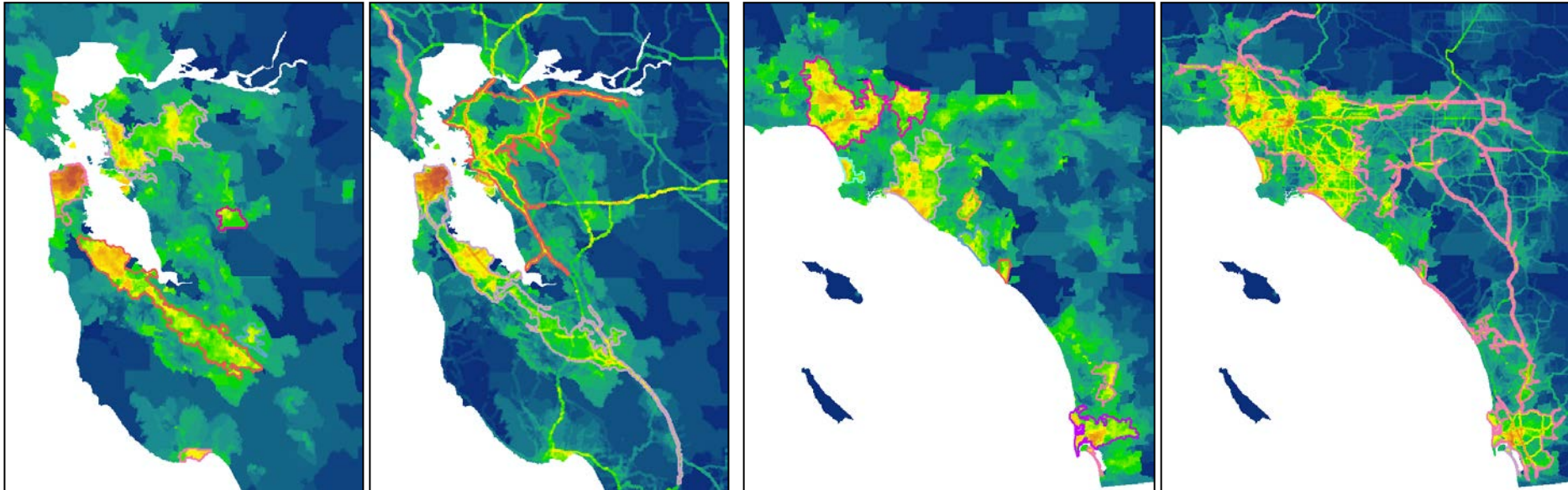
# USING TRAFFIC TO DETERMINE CONNECTOR STATIONS

Coverage Algorithms

Existing Coverage Factor Estimate

0.25mi x 0.25mi

$$Score = \frac{1}{6}$$

Score=5

Score=6

$$Score = \frac{1}{6+5}$$

1-min

3-min
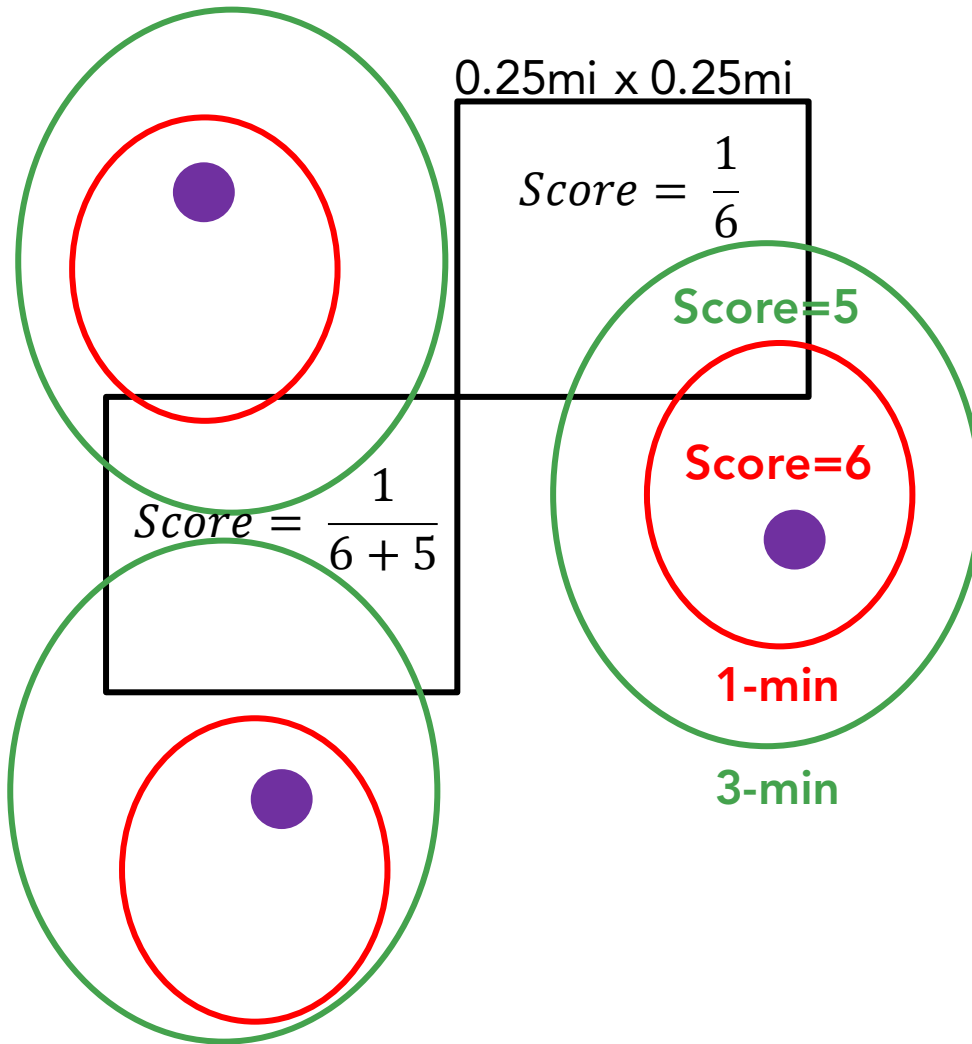
- Account for multiple overlapping service areas at a given "point"

- Assign shortest overlapping service area from each overlapping station to cell

- Shorter drive times assigned higher score

- Score for cell is inverse of sum of overlapping values

- CHIT core market: 1, 3, 6, 9, 12, 15 minutes

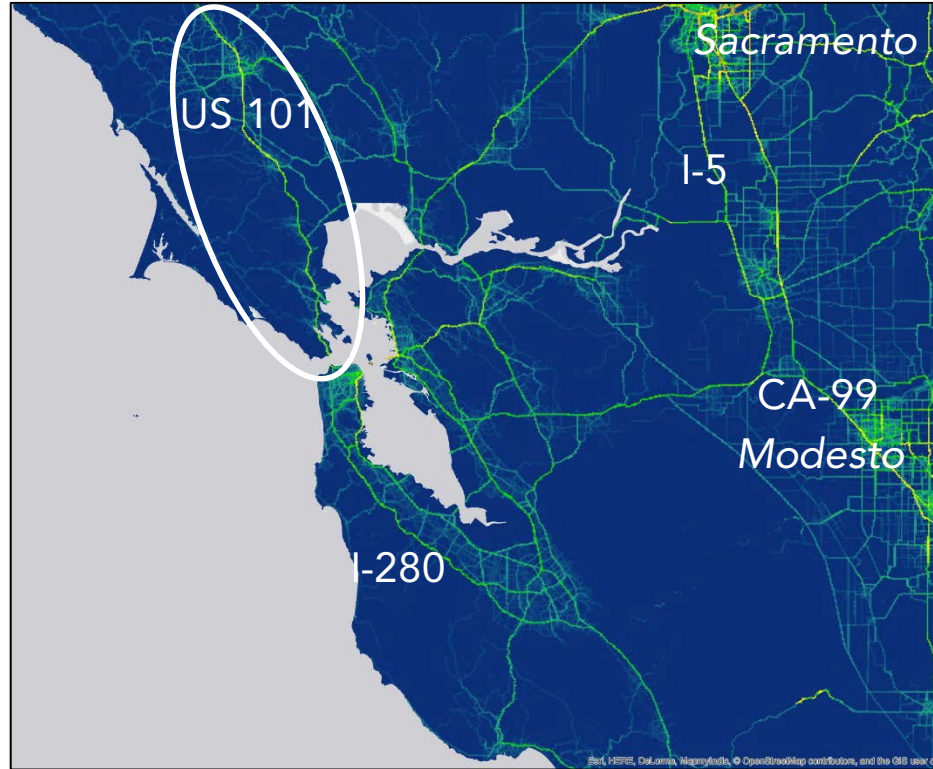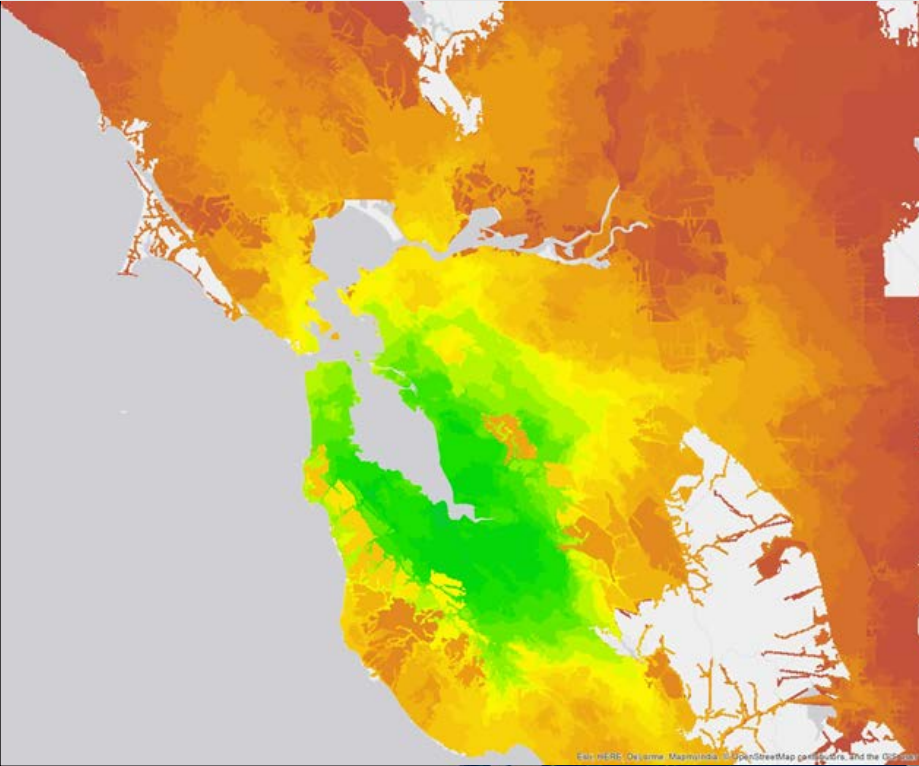- Proposed CHIT Connector: 30, 60, 90, 120 minutes

Use for Connector Stations
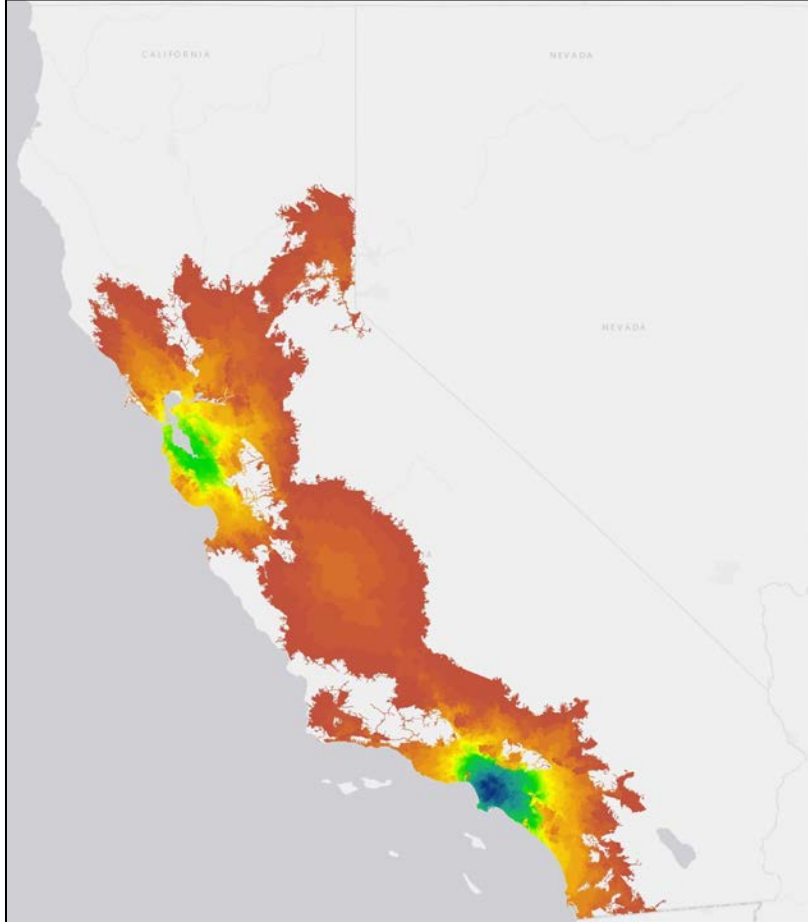
Statewide Overview
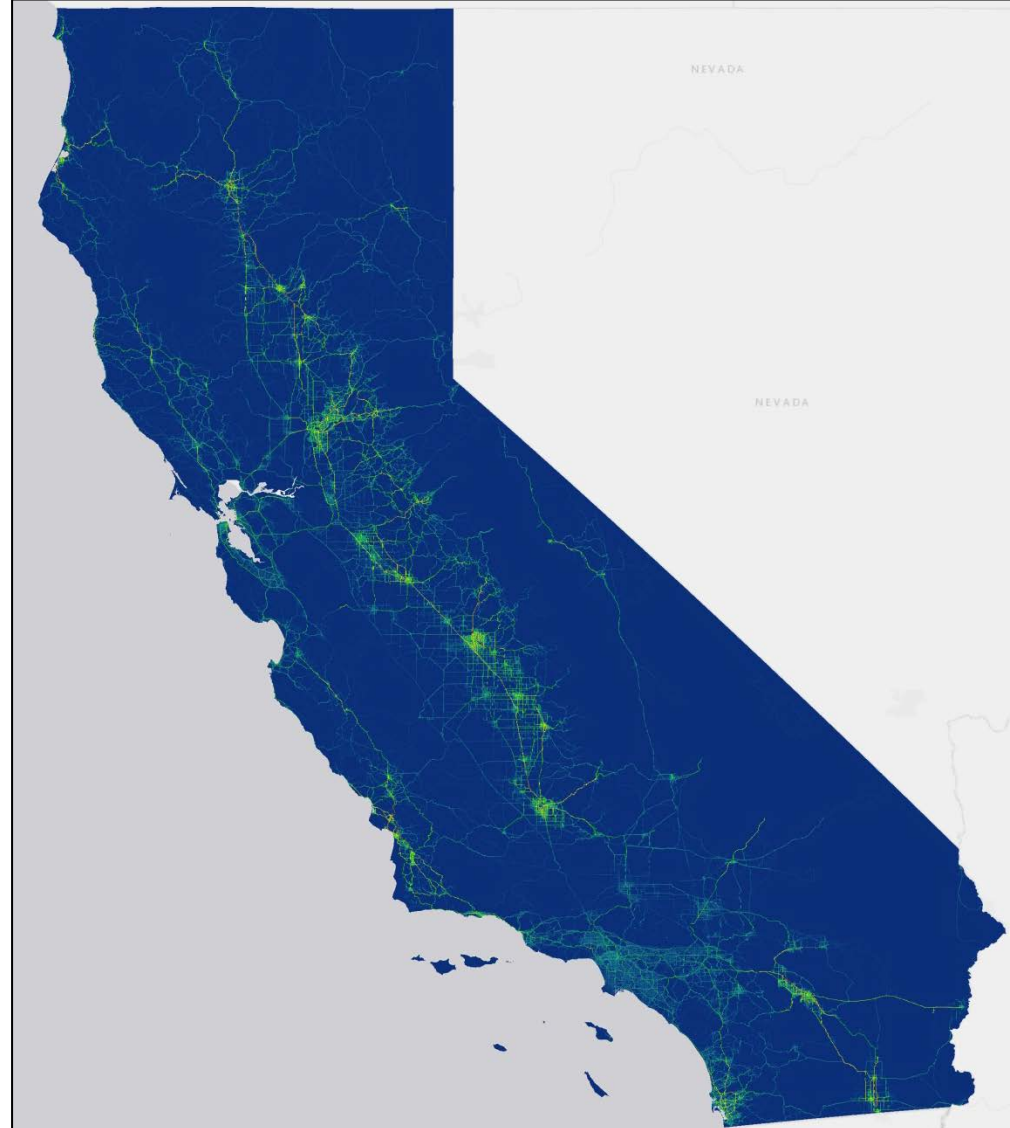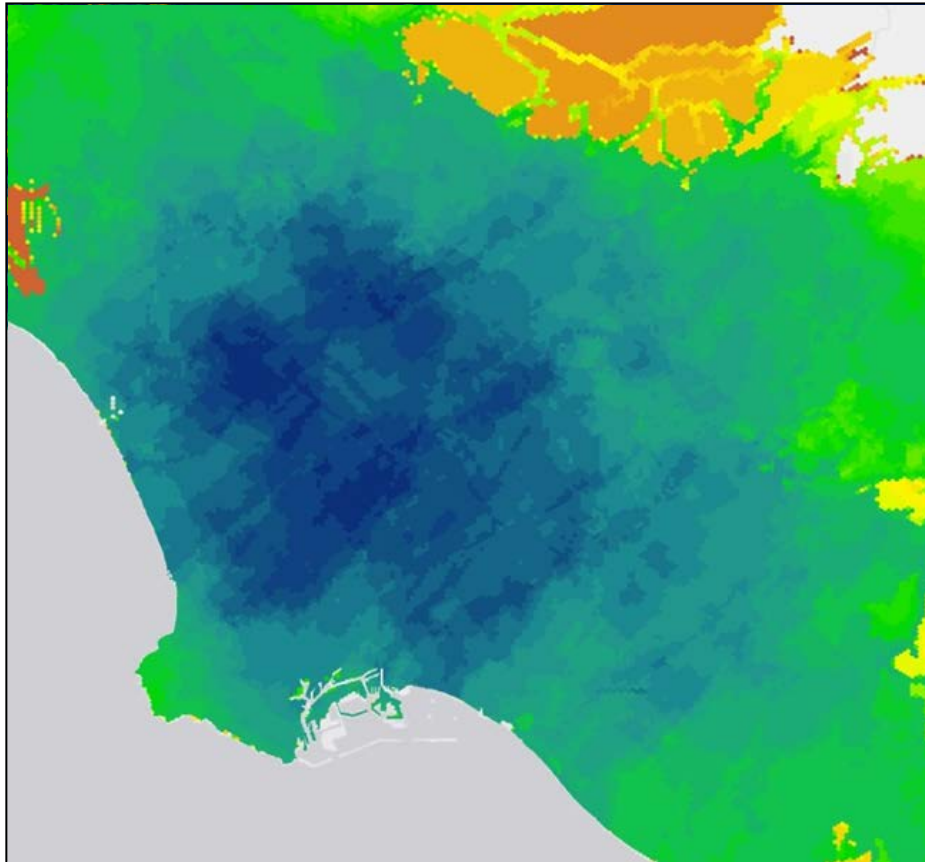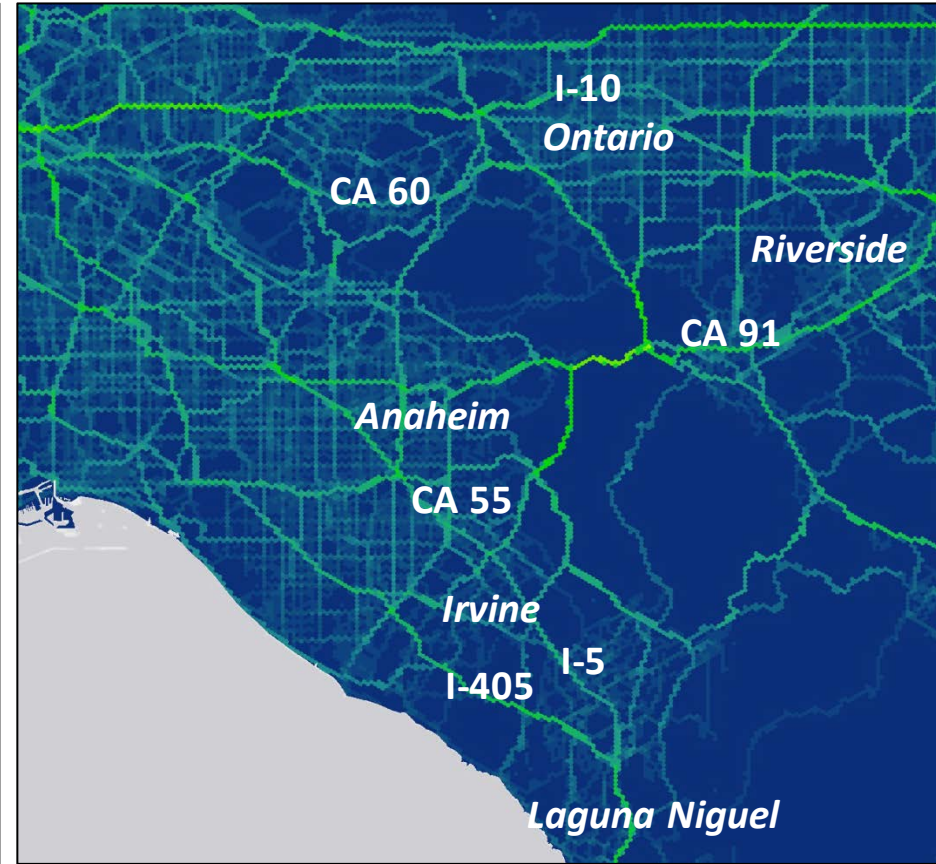
Use for Connector Stations

Bay Area Detail View

Use for Connector Stations

Los Angeles- Area Detail View
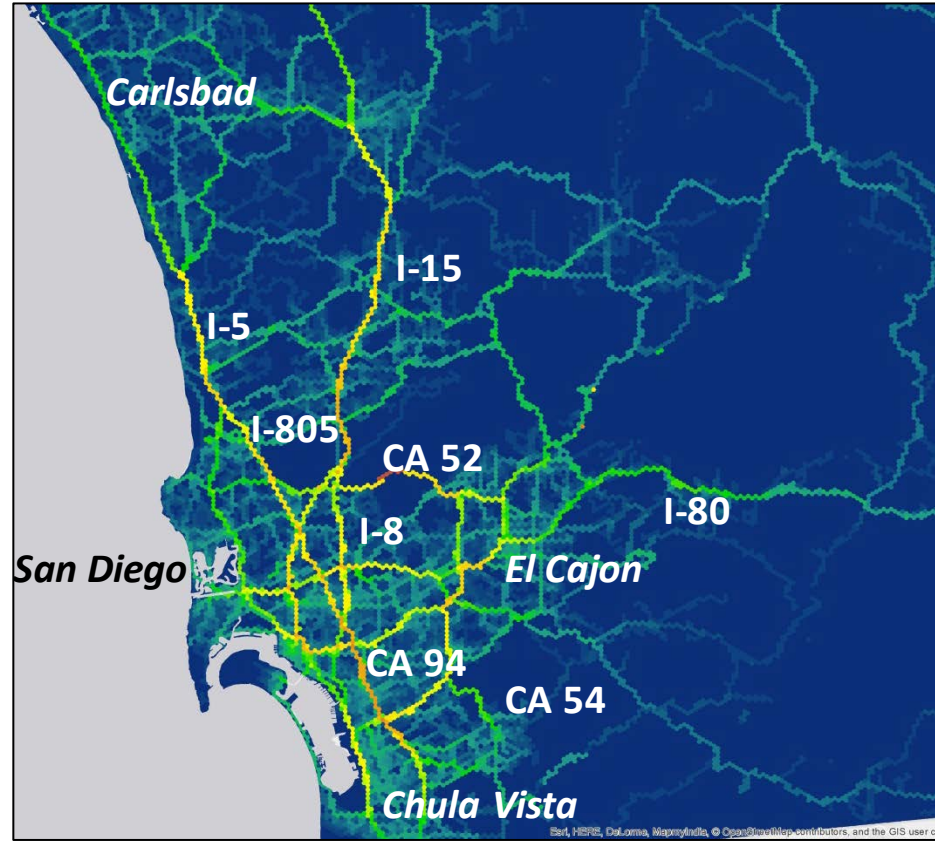
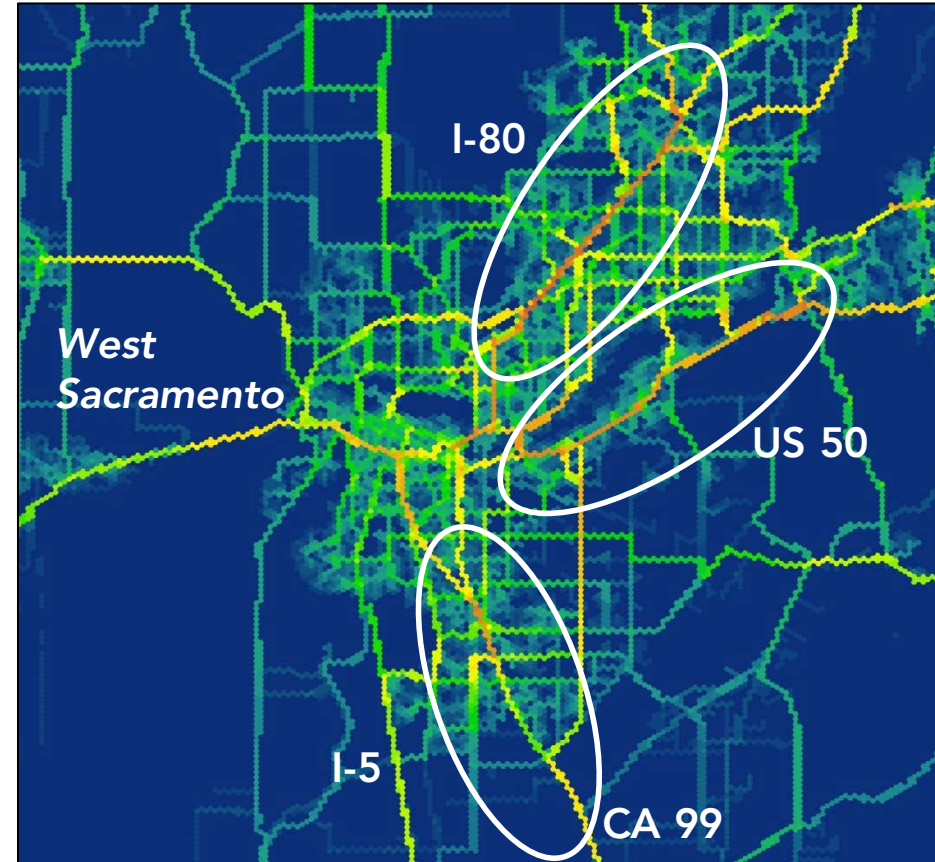# Use for Connector Stations
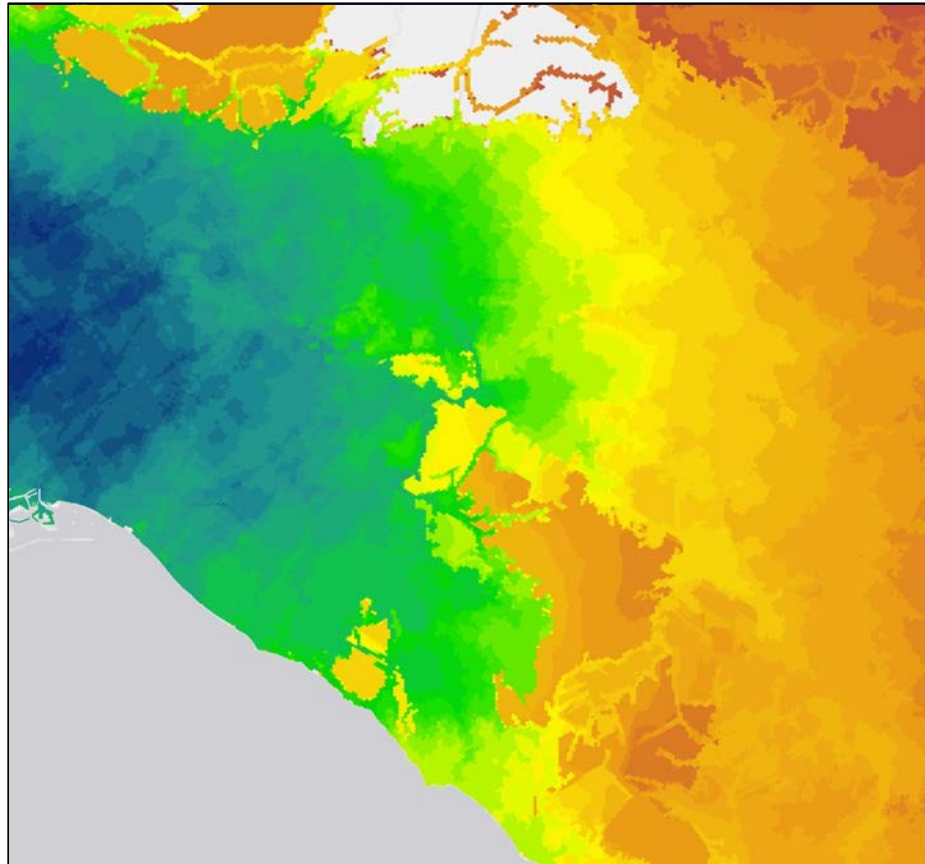
Coastal Orange County Detail View
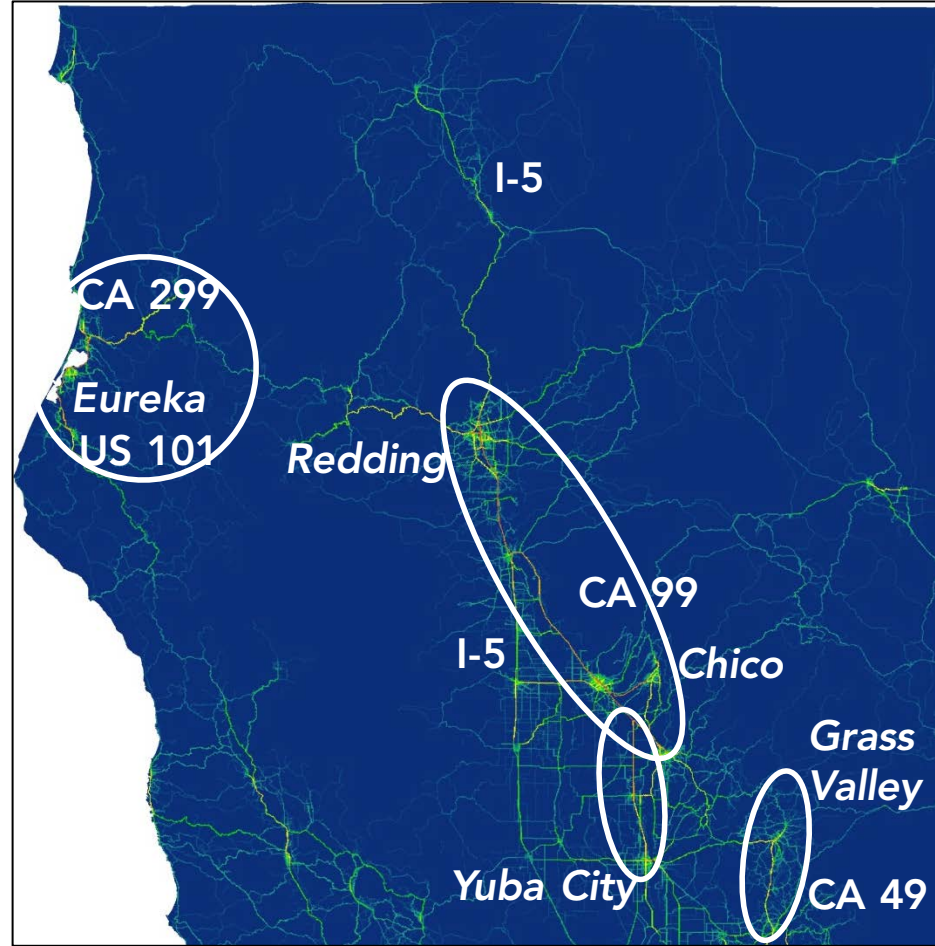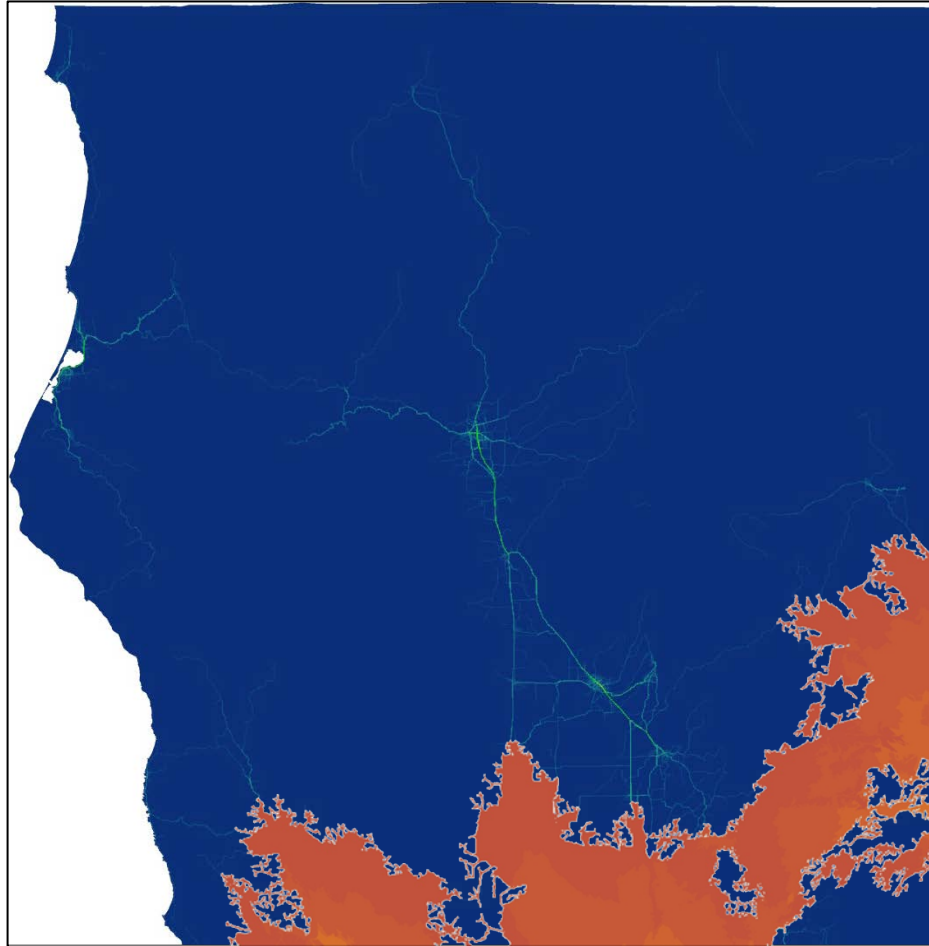
Use for Connector Stations

San Diego Area Detail View
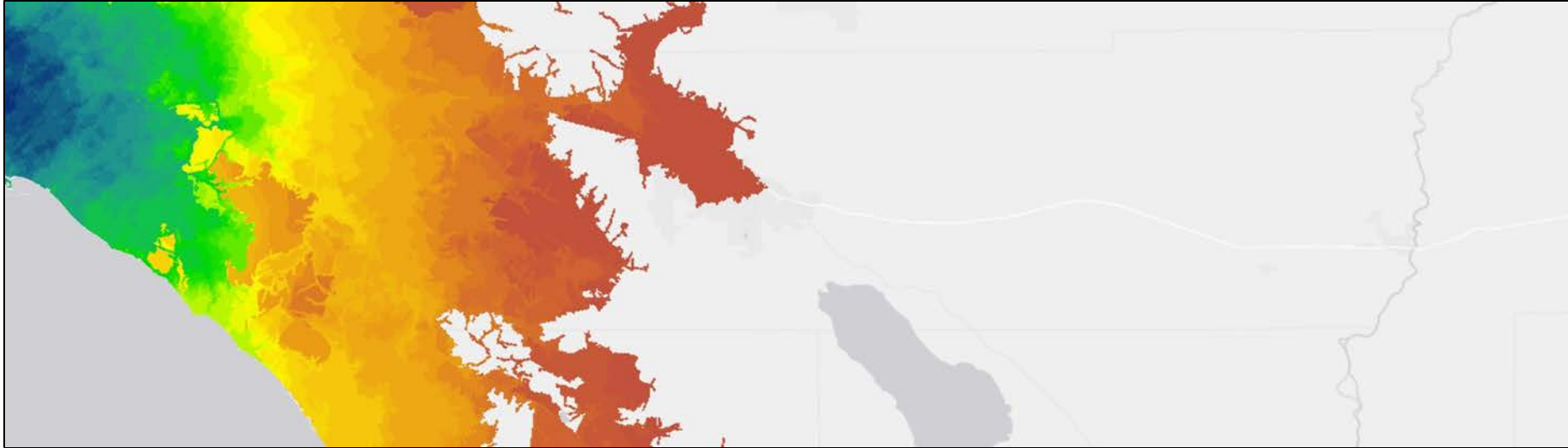
Use for Connector Stations

Sacramento Detail View

Use for Connector Stations

North State Detail View

Use for Connector Stations

Desert Cities Detail View
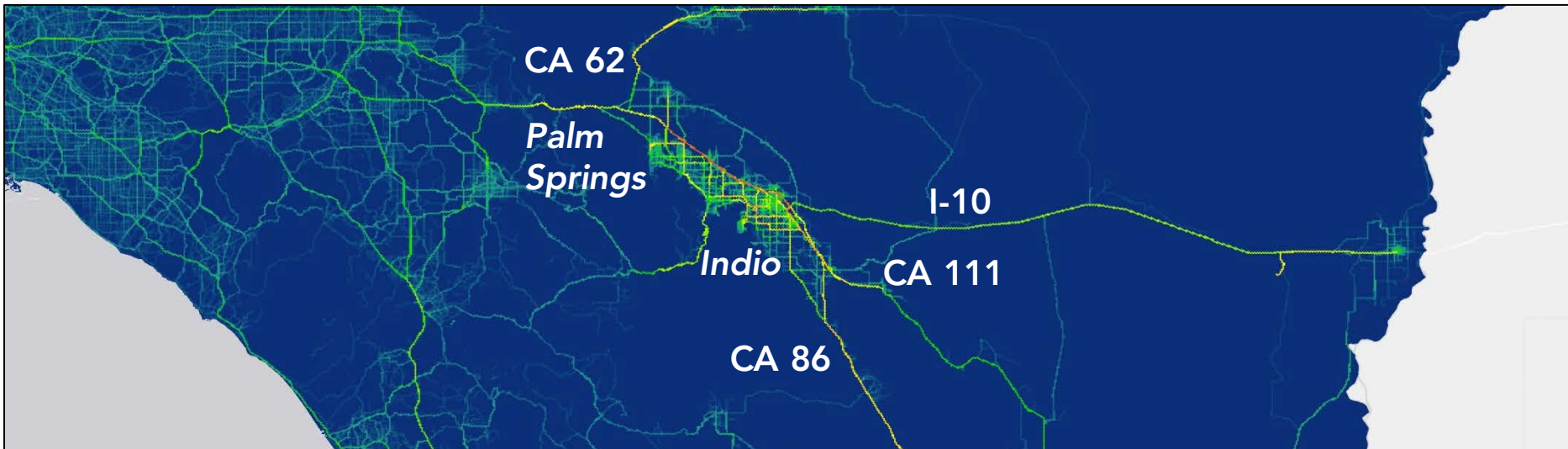
Use for Connector Stations
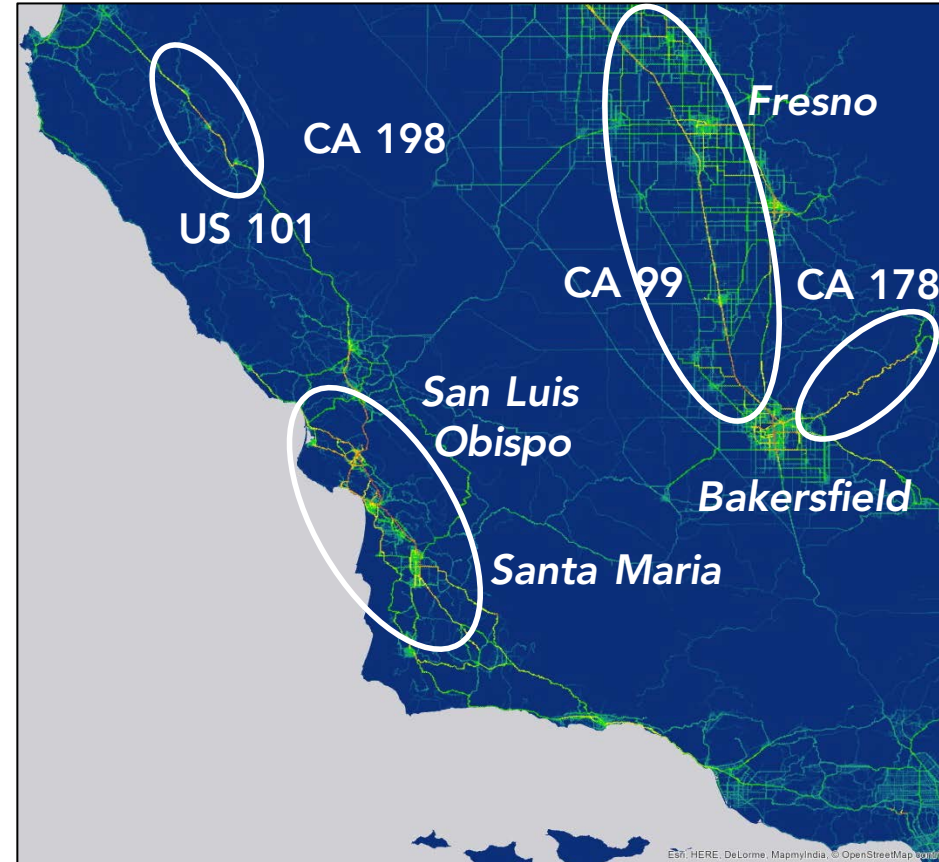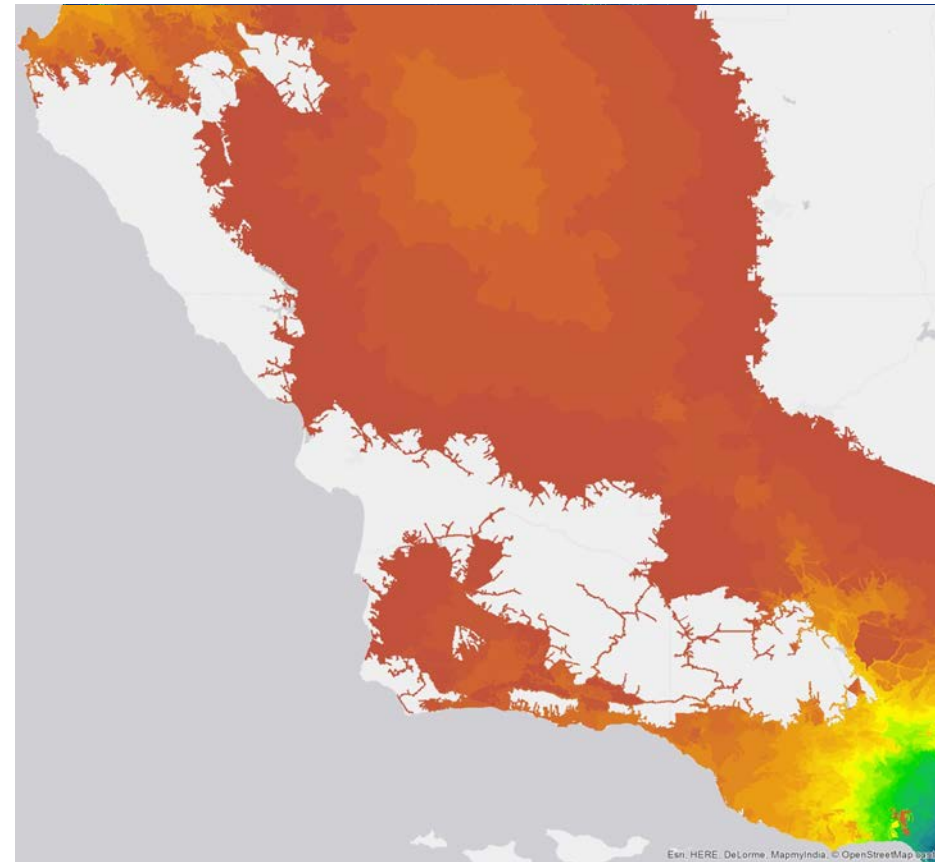
Central Valley and Coast Detail View

Connector Station Implications

- From the preceding figures, the following may be determined:
  - In the Bay Area, current core market stations cover traffic patterns well; an opportunity may exist along US-101 north of the peninsula

  - LA and OC – area long distance connector travel is similarly well-covered by the core market stations

  - San Diego – area connector travel is not well met by the current network. Opportunities exist along I-15 and I-5 leading into the area from the north, and along I-805, I-8, and State Routes 163, 52, 54, and 94 between downtown and El Cajon

  - Major highways near Sacramento are also good candidates, including east-west routes of I-80 and US-50, and the north-south routes of I-5 and CA-99

Connector Station Implications

- From the preceding figures, the following may be determined:

  - In the northern areas of the state, I-5 and CA-99 are strong candidates, especially between Yuba City and Redding. CA-49 between Sacramento and Grass Valley is also a prime candidate, as are US 101 and CA-299 around Eureka.

  - There is strong need between the desert cities around Palm Springs, along I-10 further to the east, Routes 111, 86, and 62, and portions of I-15

  - In the Central Valley, north-south connectors along CA-99 are the greatest need, though east-west routes extending from CA-99 appear viable as well

  - Along the Central coast, opportunities are strong along US-101 and State Route 227 and others, especially near coastal cities of Santa Maria and San Luis Obispo, and on US-101 north of the intersection with State Route 198

POTENTIAL TO FURTHER INTEGRATE MARKET AND TRAFFIC

Market Factor on Travel Routes

- As thus far presented, the consideration of travel routes is generic. It is not in any way explicitly known to be the travel of the first adopters.

- All origins and destination are known in the data set. This could enable weighting of each route by the strength of the first adopter market at the home end of each route.

- This still would not mean the first adopters are directly simulated (since we do not know that all destinations are actually the destinations of first adopters' commutes rather than any neighborhood commute).

Market Factor on Travel Routes

- However, considered as a probability weighting factor, this method could help increase the relative weight put on routes that might be more related to first adopters.

- ARB has not initiated this effort yet. Although the data are available, the development and calculation time are likely to be significant, as well.

- ARB would like to hear feedback from stakeholders on whether this is a necessary step before initiating.

THE WILDCARD

Alternative Option: UBER

- On January 8, 2017, UBER announced Movement

- Movement is a platform that allows planners to leverage data collected by UBER's drivers

- The platform is not yet released and details have not been given of the specific data fields that will be available

> This data is anonymized and aggregated into the same types of geographic zones that transportation planners use to evaluate which parts of cities need expanded infrastructure, like Census Tracts and Traffic Analysis Zones (TAZs). In the weeks ahead, we'll be inviting planning agencies and researchers to access our data and explore zone-to-zone travel times, and will soon make the website freely available to the public.

UBER

- Costs have not been announced

- ARB has signed up to receive notification of availability

OPEN DISCUSSION

Main goal: What potential uses for simulated traffic data should be implemented for CHIT 2017+ analyses?

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