Integrated Models and Data Needs for Large-Scale Simulations:

Accessibility, Time-Space Prisms, Choice Sets, and Related Issues

Konstadinos (Kostas) Goulias

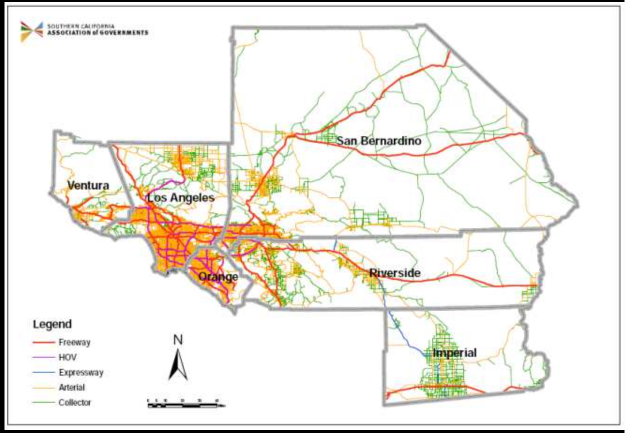
UNIV. OF CALIFORNIA SANTA BARBARA GEOGRAPHY

Outline

- “Typical” emerging schema
- “Typical” emerging data collection
- Behavior in Space-Time
- Modeling the environment & the agent
- Integration needs
- Computational issues

Southern California Region

- **Simulator of Activities, Greenhouse gas Emissions, Networks, and Travel (SimAGENT) in Southern California**
- Six Counties (population, percent change 2000-2010)
  - Imperial (174,528, D%=22.6)
  - Ventura (823,318 , D%=9.3)
  - San Bernardino (2,035,210, D%=19.1)
  - Riverside (2,189,641, D%=41.7%)
  - Orange (3,010,232, D%=5.8%)
  - Los Angeles (9,818,605, D%=3.1)



Typical Schema SimAGENT

Baseline Year (t=1)

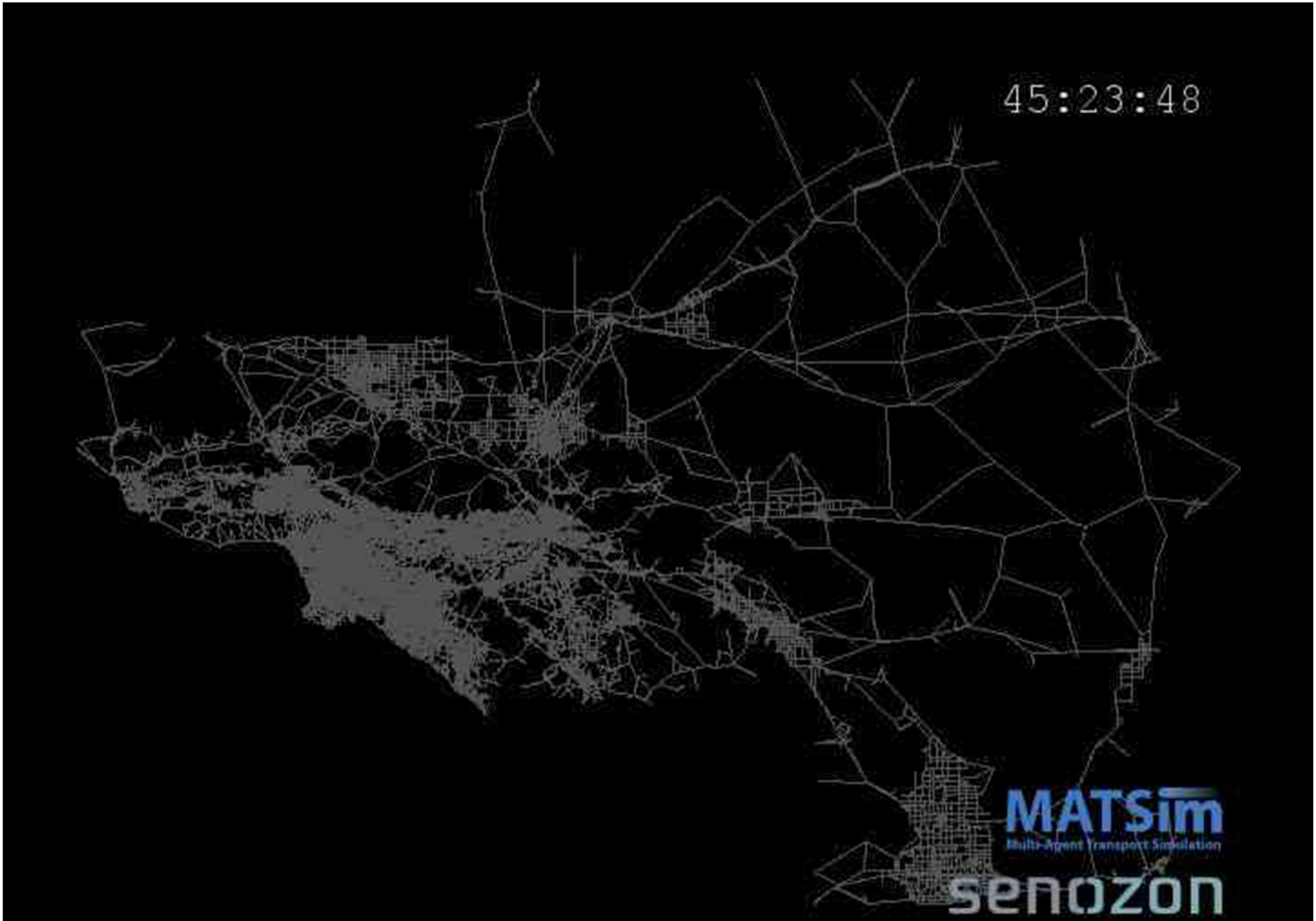
- Synthetic Population
- Accessibility by Time-of-Day
- Long Term Choices
- Car Ownership and Type
- Activity and Travel Scheduling
- Routes & Assignment
- Energy Consumption & Emissions

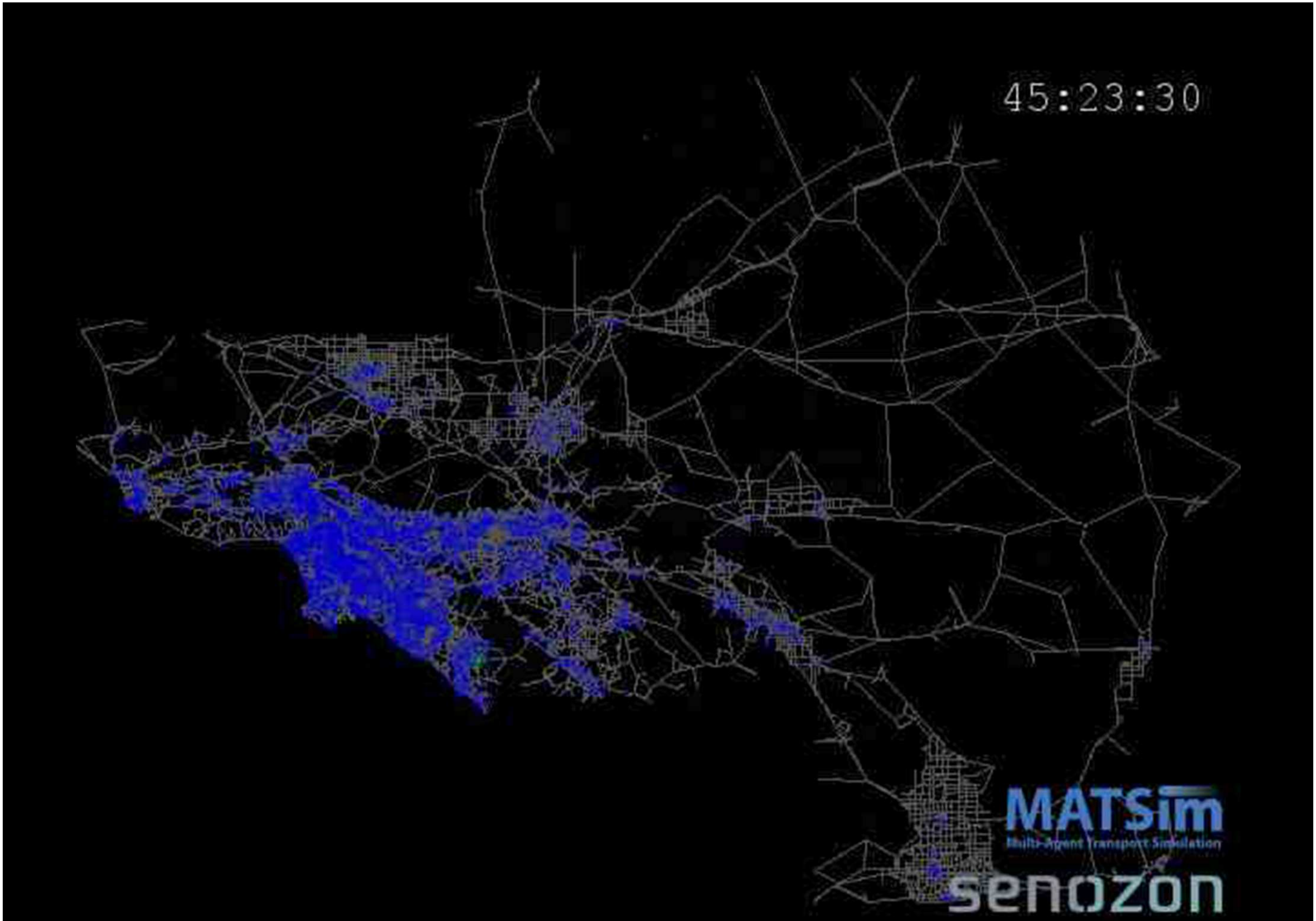
Agent and Environment Evolution

- Population Evolution
- Urban Landscape Evolution
- Infrastructural Changes
- Scenario Databases
- Information Fusion
- Accessibility Computation

One Year Later (t=t+1)

- Synthetic Population
- Accessibility by Time-of-Day
- Long Term Choices
- Car Ownership and Type
- Activity and Travel Scheduling
- Routes & Assignment
- Energy Consumption & Emissions


45:23:48
MATSim
Multi-Agent Transport Simulation
senozon


45:23:30
MATSim
Multi-Agent Transport Simulation
senozon

Data Needs

- Core Behavior and Household Characteristics
  - Other aspects – policy dependent (cars and costs, long term choices and lifestyle, attitudes)
- Other agents (firms, institutions, plans, and so forth)
- Landscape/Environment/Context
  - Activity locations
  - Homes/Jobs/Schools
  - Availability over time

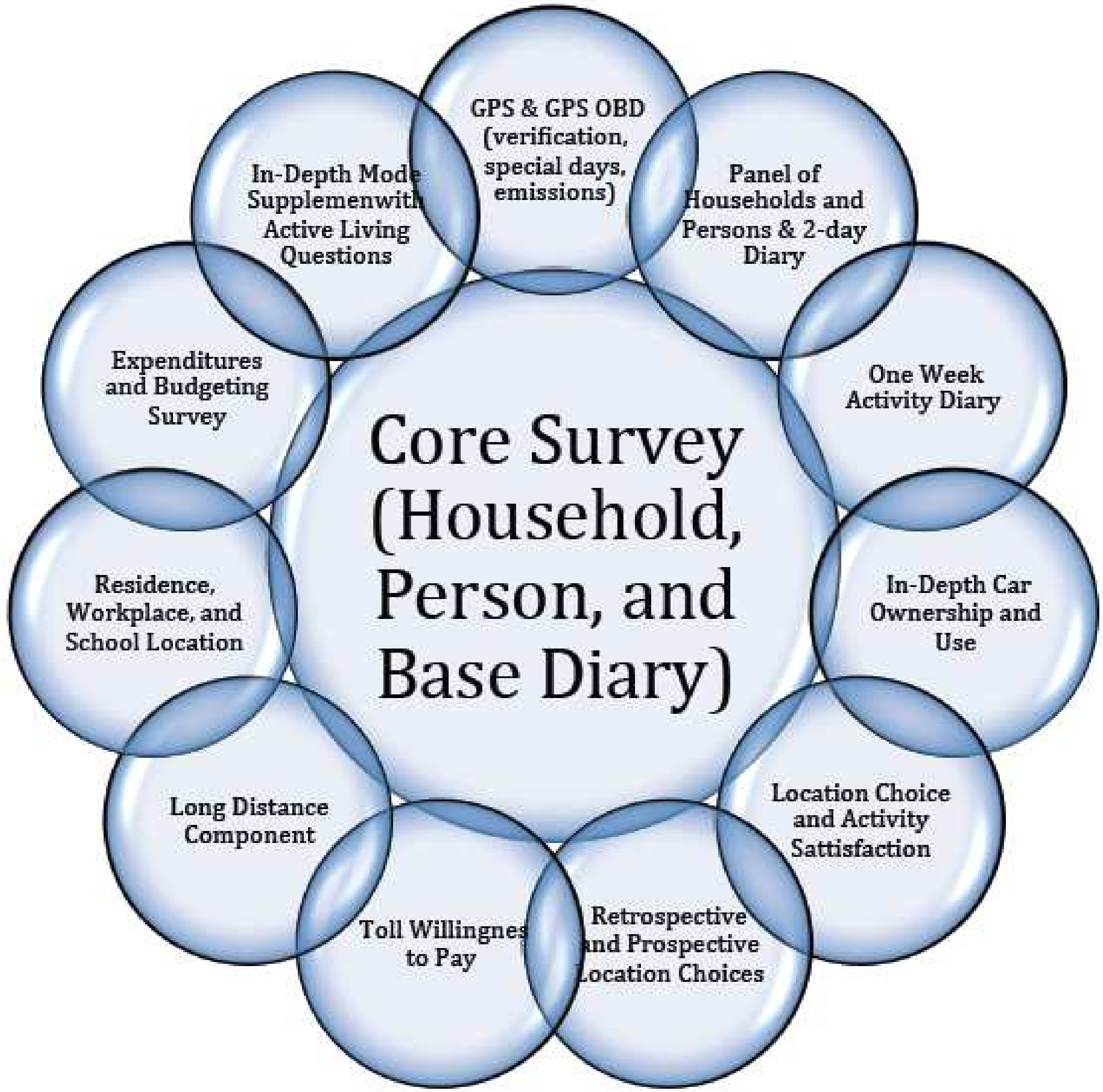


Figure 4 The Data Collection Overall Scheme

The Many Ds in California Land Use & Transp. Policy (change in each indicator)

- Density = (Population + Employment)/square mile
- Diversity =
- Design = a*street network density + b*sidewalk completeness+ c*route directness
- Destinations = from an origin zone i Sum over j of (attractions*impedance)
- Distance to transit

Local = depends on application; **street network density** = length of street in miles/area of neighborhood in square miles, **sidewalk completeness** = total sidewalk centerline distance/total street centerline distance, **route directness** = average airline distance to center/average roadway distance to center

INDEX 4D uses a=0.0195, b=1.18, and c=3.63

Goals

- Develop the base background data and information to develop the Ds
- Create indices at a fine spatial level (display activity “deserts”)
- Add the temporal (time of day) dimension to the indices
- Create a method that can be used in forecasting/scenario building
- Focus on Destinations intended as **opportunities** and not as a gravity model or other functions that require behavioral assumptions

Objectives

- Map an entire region (even a state) to show access to activity opportunities
- Use 15 different types of "industries" (retail, education, health, etc)
- Account for different speeds on network by time of day
- Account for different opening-closing hours by county and industry type
- Use largely available data with any needed rectifications using secondary sources of information
- Develop a multi-scale process (TAZ/tract, blockgroups, blocks, land parcels)
- Compute these for different years (2000-2003, 2008, 2010, 2025, 2035) –future years are for land use scenarios
- Test these indicators in behavioral model specifications
- Use these indicators to study spatial equity for sociodemographic groups

Ingredients

- A geocoded database that enumerates all opportunities classified by the type of activity – first order of approximation can be number of business establishments classified based on NAICS-SIC
- Employment by type of industry to use as verification but also as a direct indicator of attractiveness
- Other data to fuse, merge, verify, validate built environment estimates

- A detailed network with all models – start with highways and add other travel modes
- Estimates of travel speeds on each link

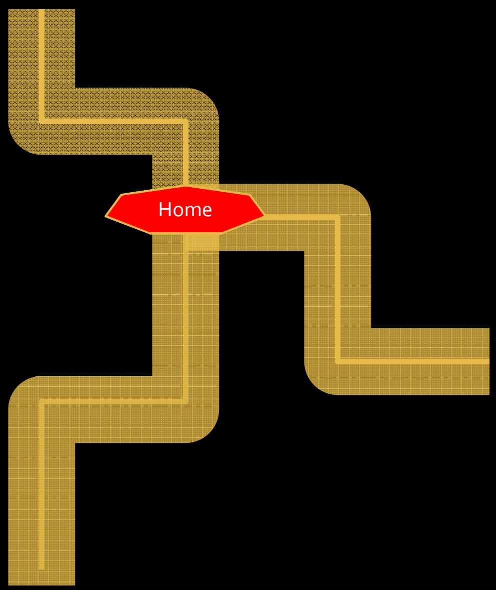
- An algorithm to build the indicators

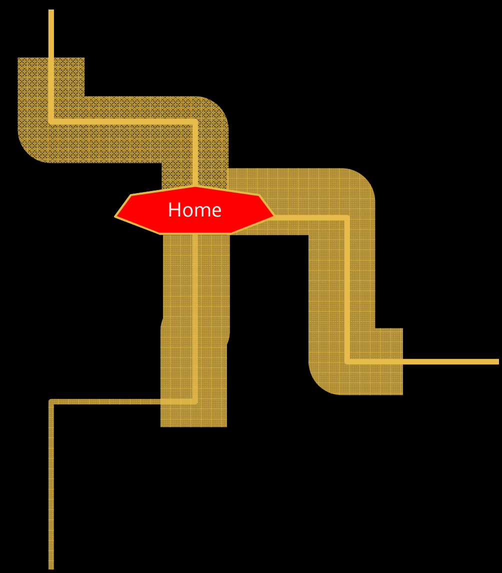
Data Used
(in Los Angeles+)

- CTPP vintage 2000 (considered as a baseline)
- Dunn & Bradstreet post processed to provide summaries by block group (via a joint project with LANL & UCOP)
- US Census Block, block groups, TAZs, and Tracts
- SCAG network (highway and transit)
- Teleatlas California 2000
- InfoUSA (later vintage) just for comparison
- SCAG Travel Survey Post-Census Commute Arrival to Work and Departure from Work

- All this is usually available for MPOs

Conceptually we Build Buffers





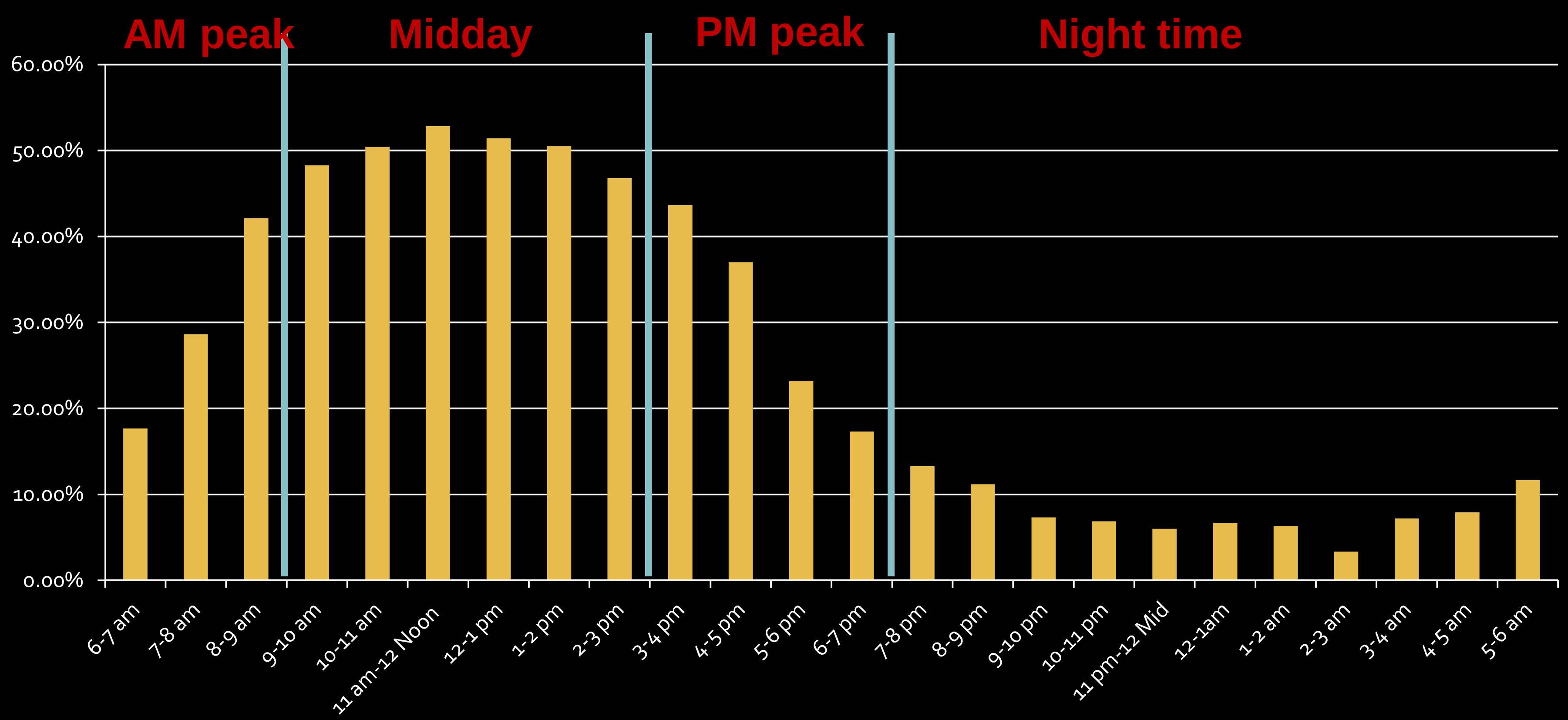
Off Peak Network-based Access to an opportunity area (potential path area)

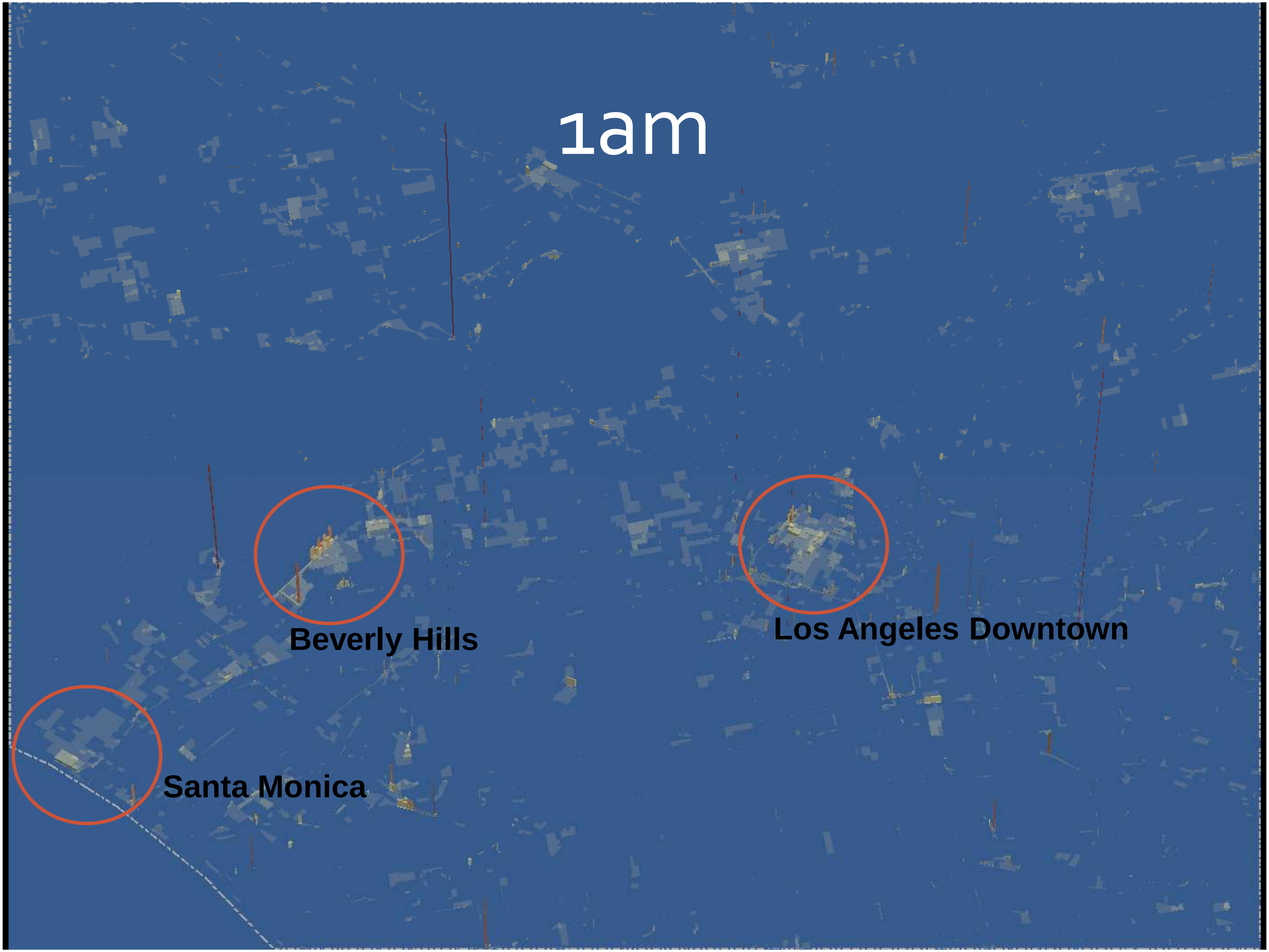
Peak Period Network-based Access to an opportunity area

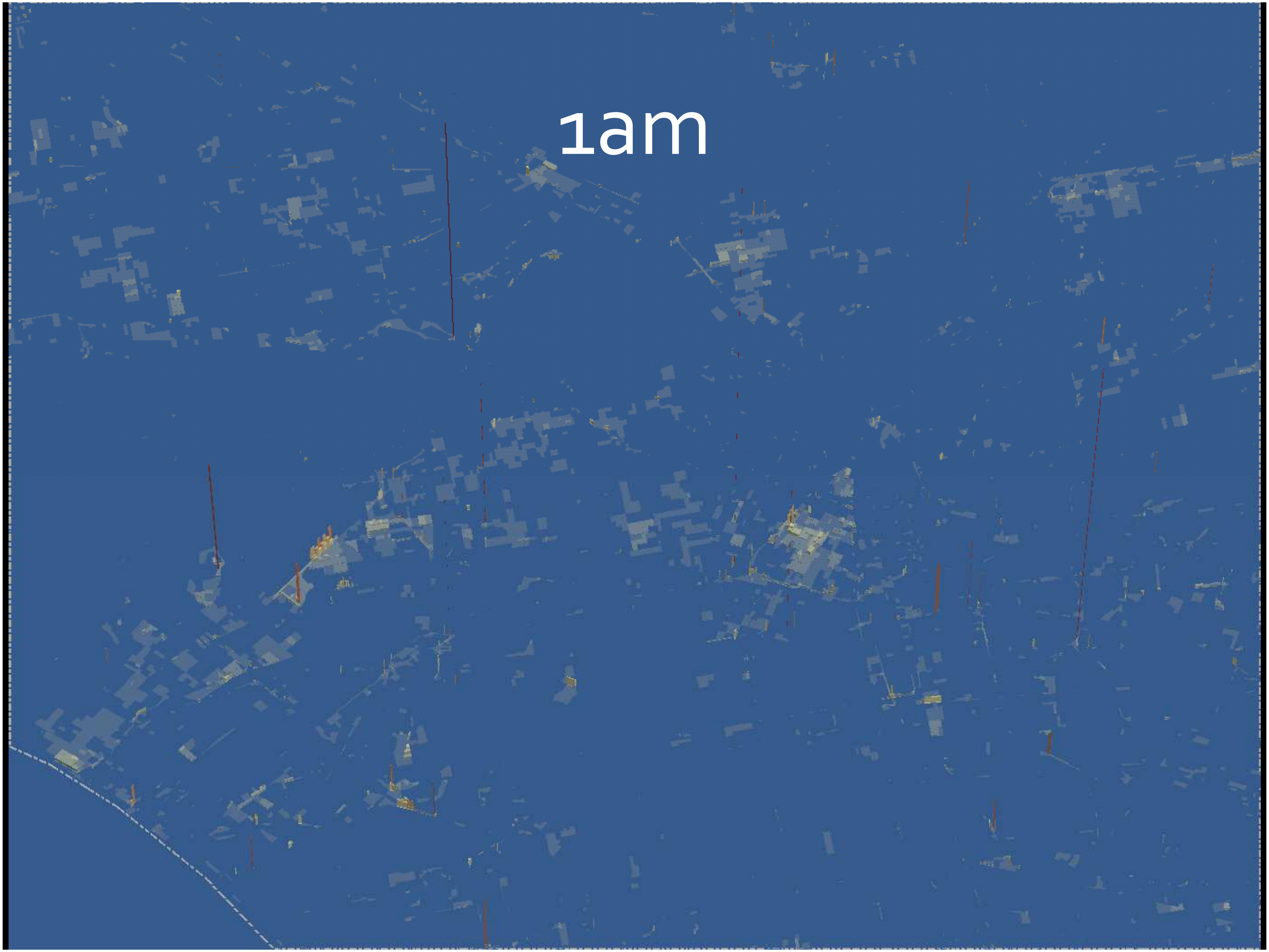
Tracking density of opportunity by time of day – Just based on the information about arrival and departure from work of retail employees

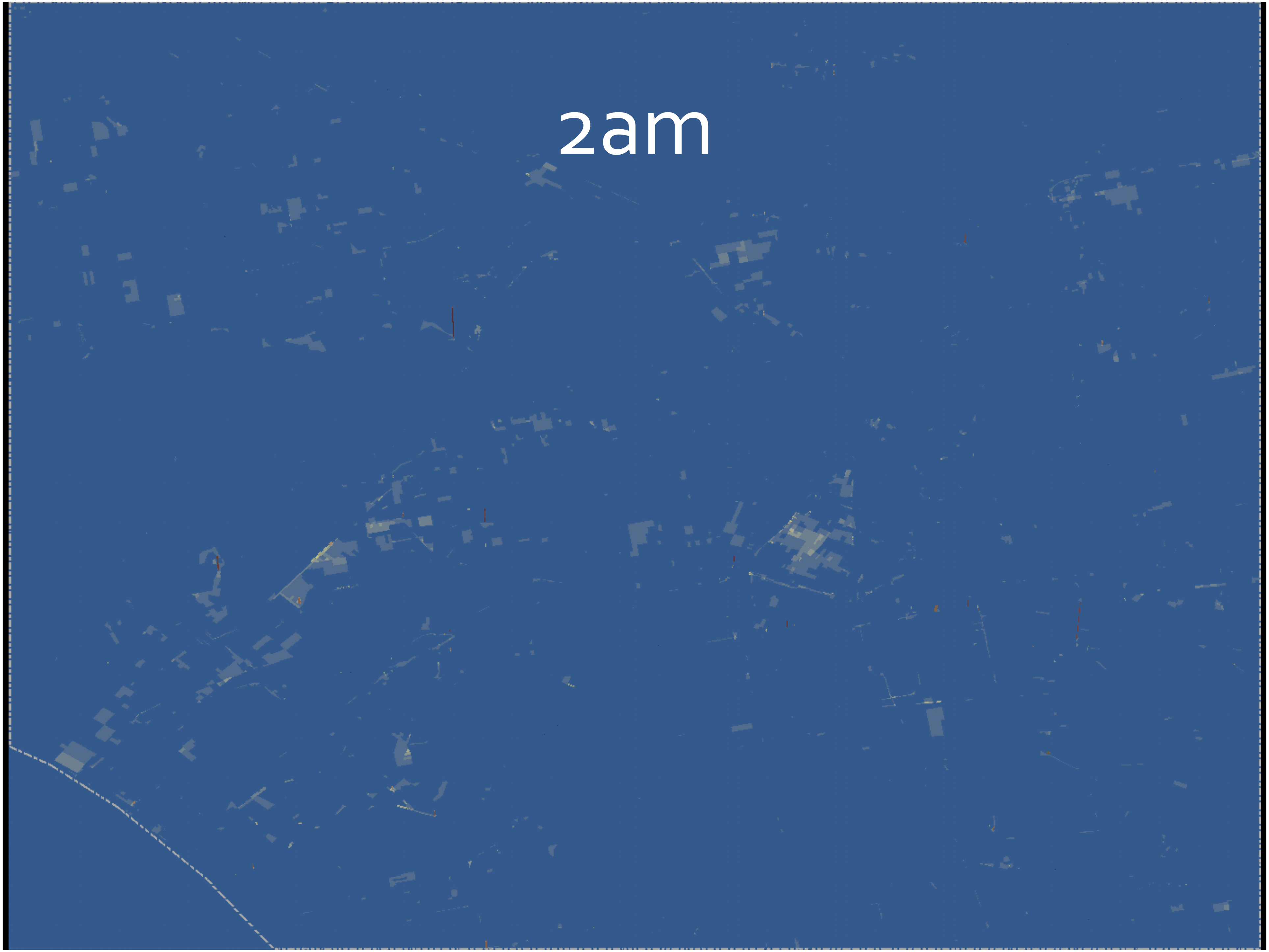
Time of day profile of available opportunity

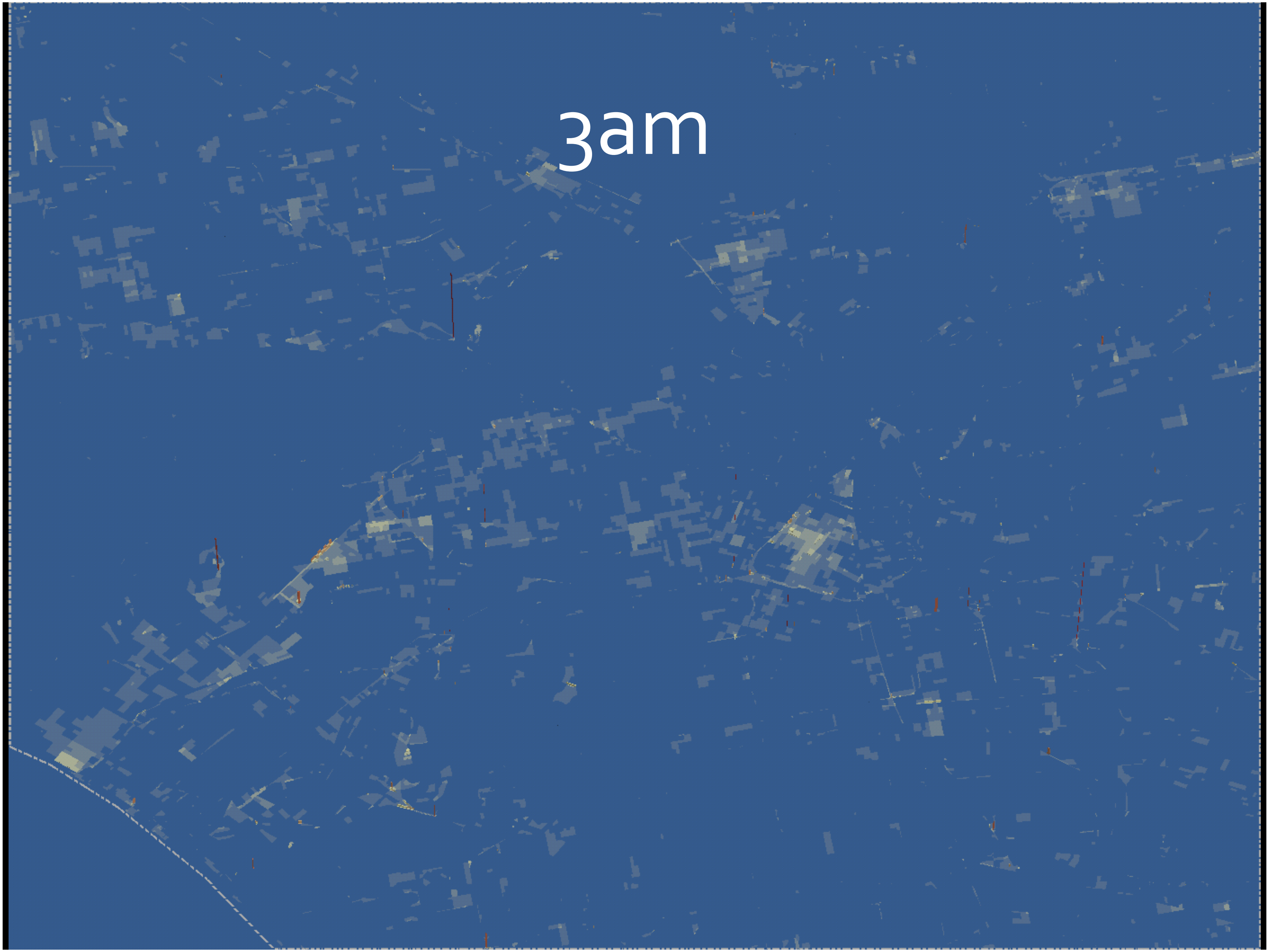
- Arrival and departure time of workers in travel survey (for each county and each industry type)
- Retail workers at work in LA county

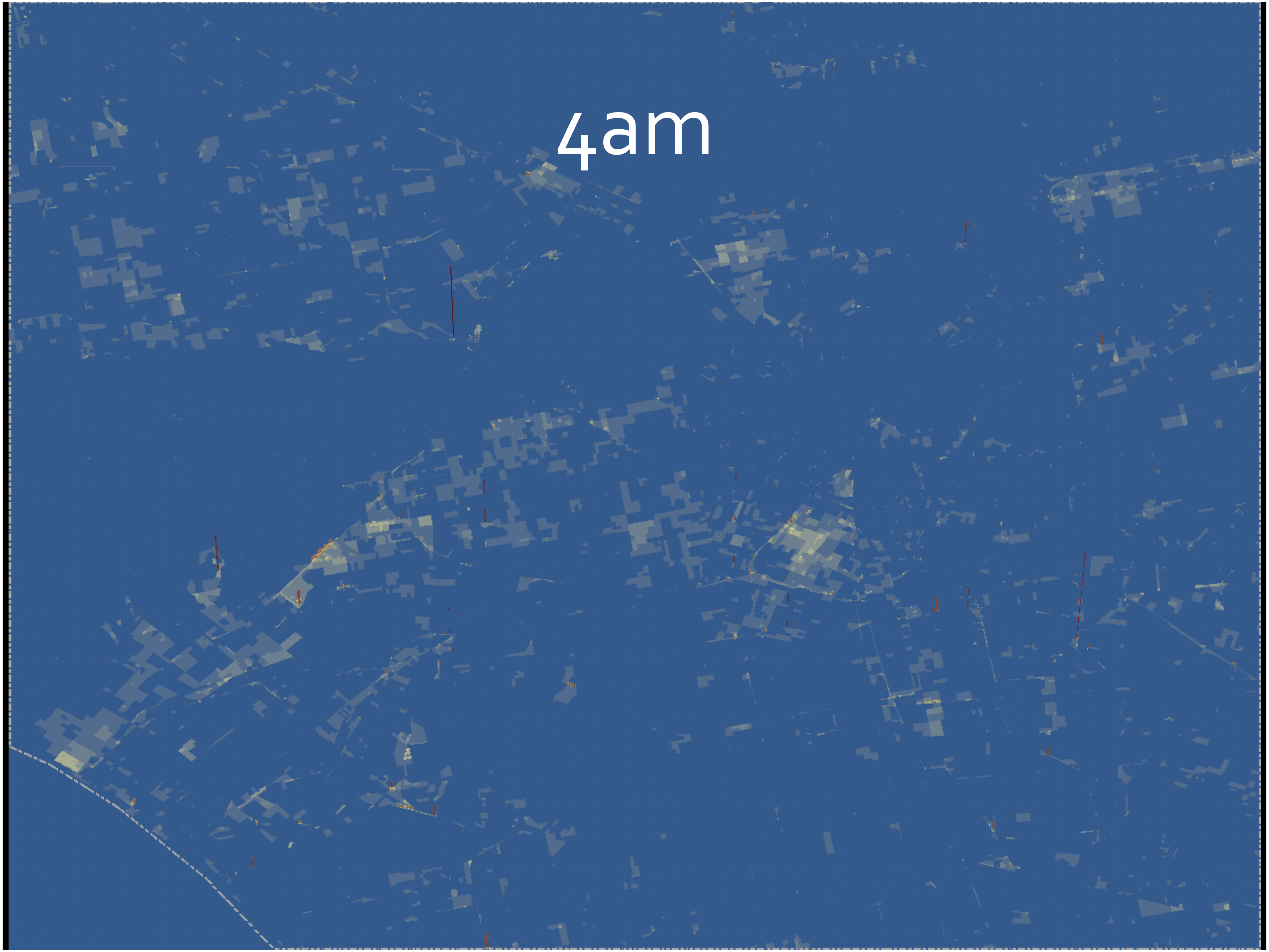


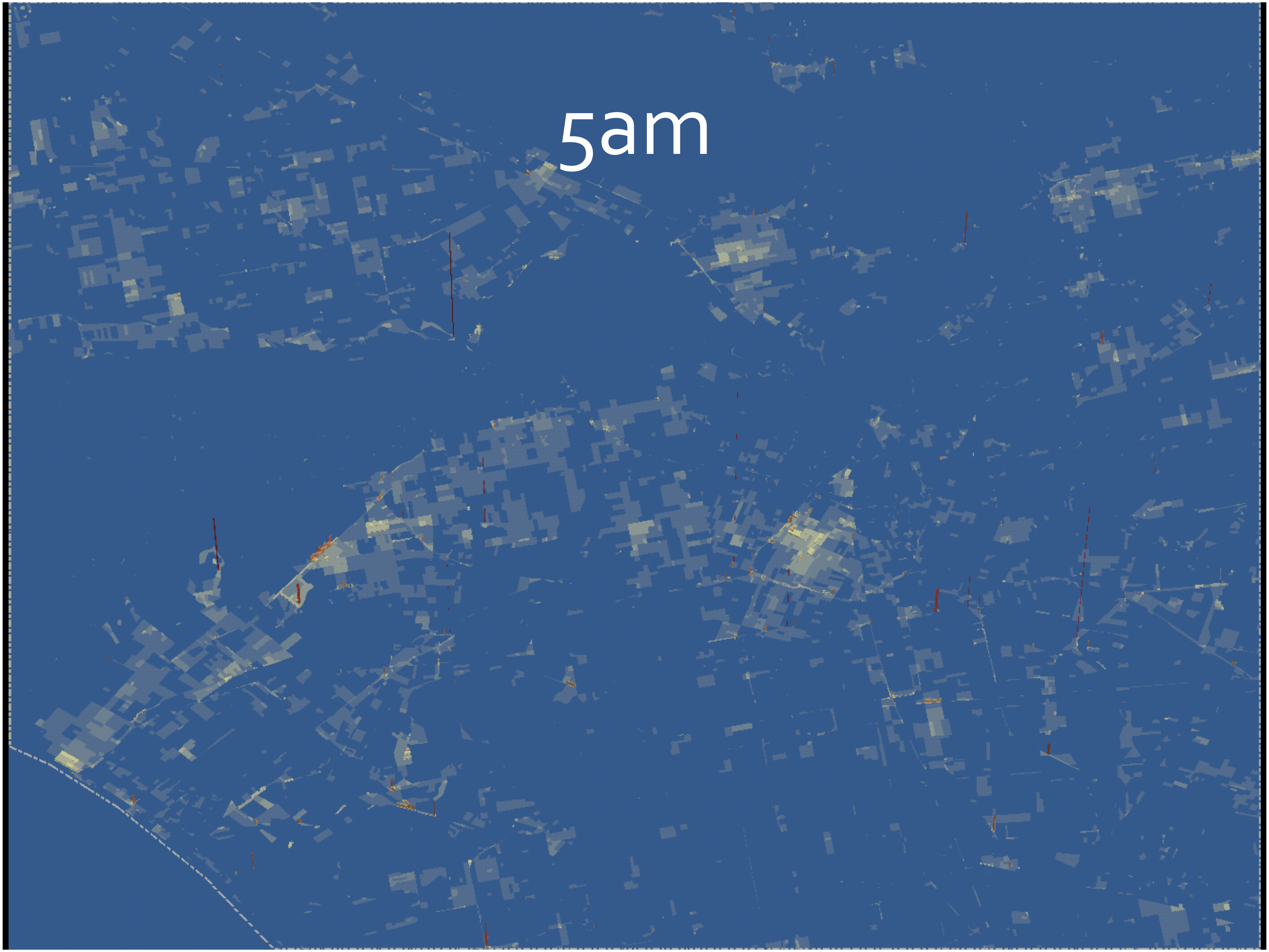

1am
Beverly Hills
Los Angeles Downtown
Santa Monica

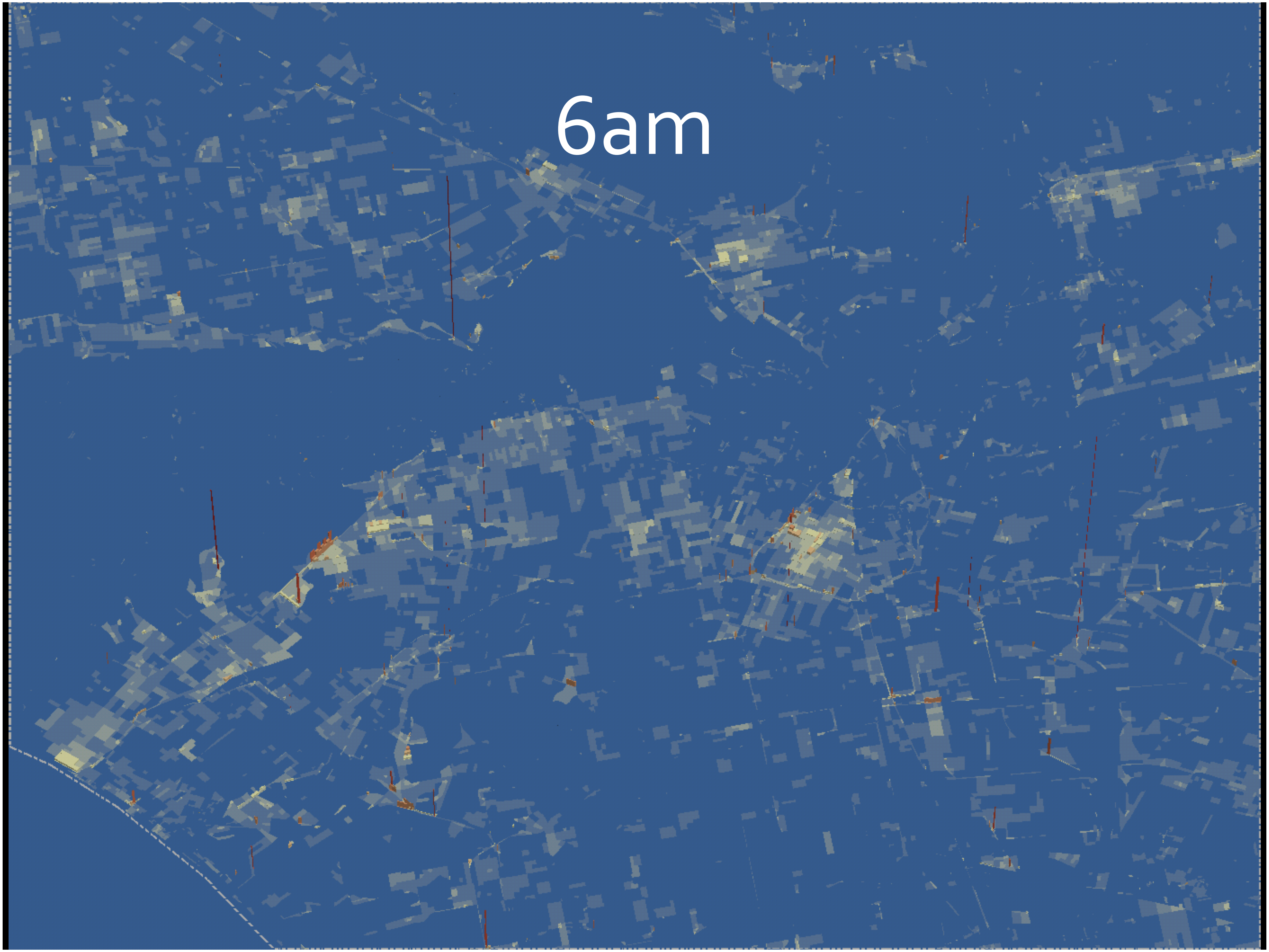

1am

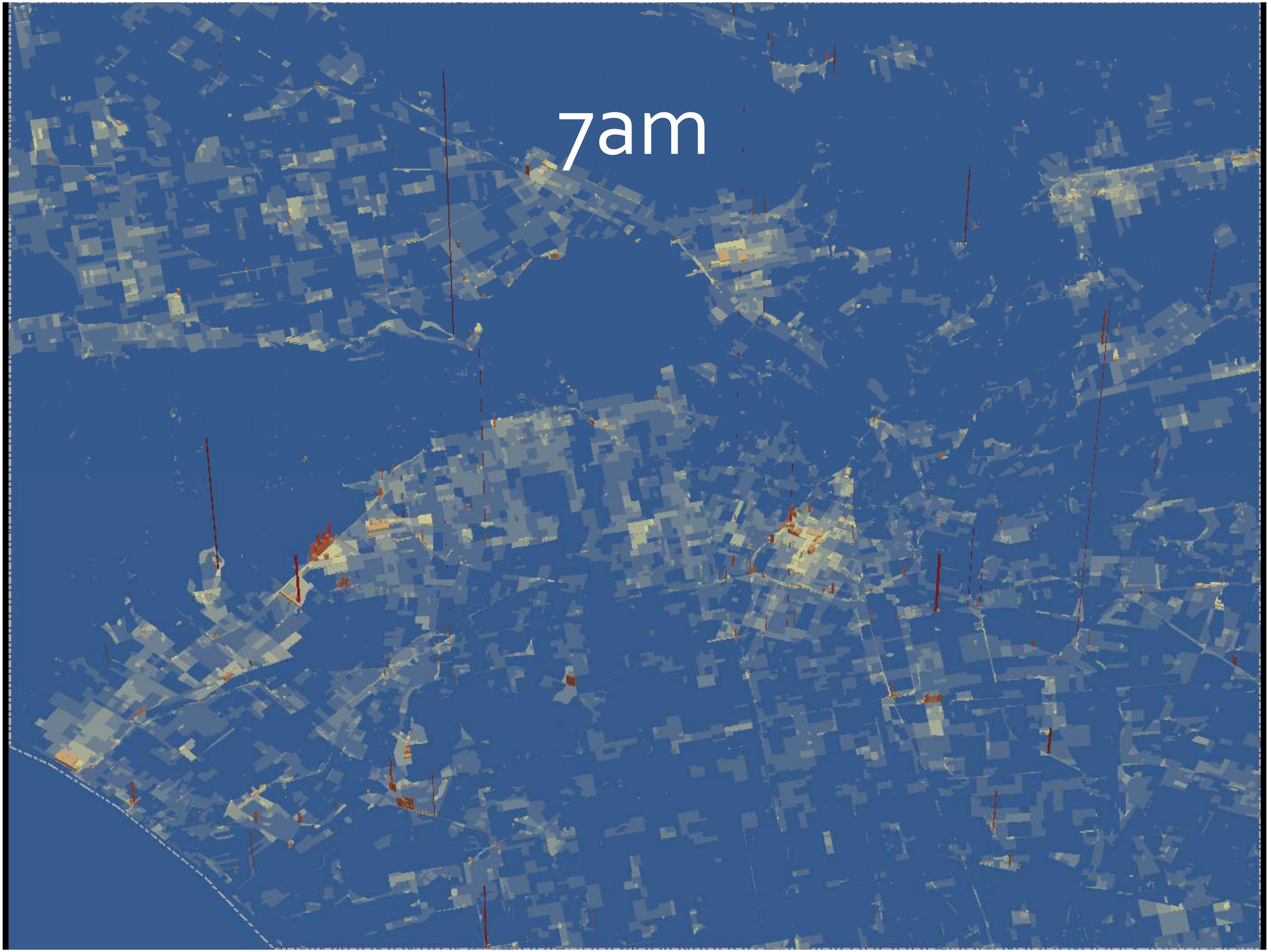

2am

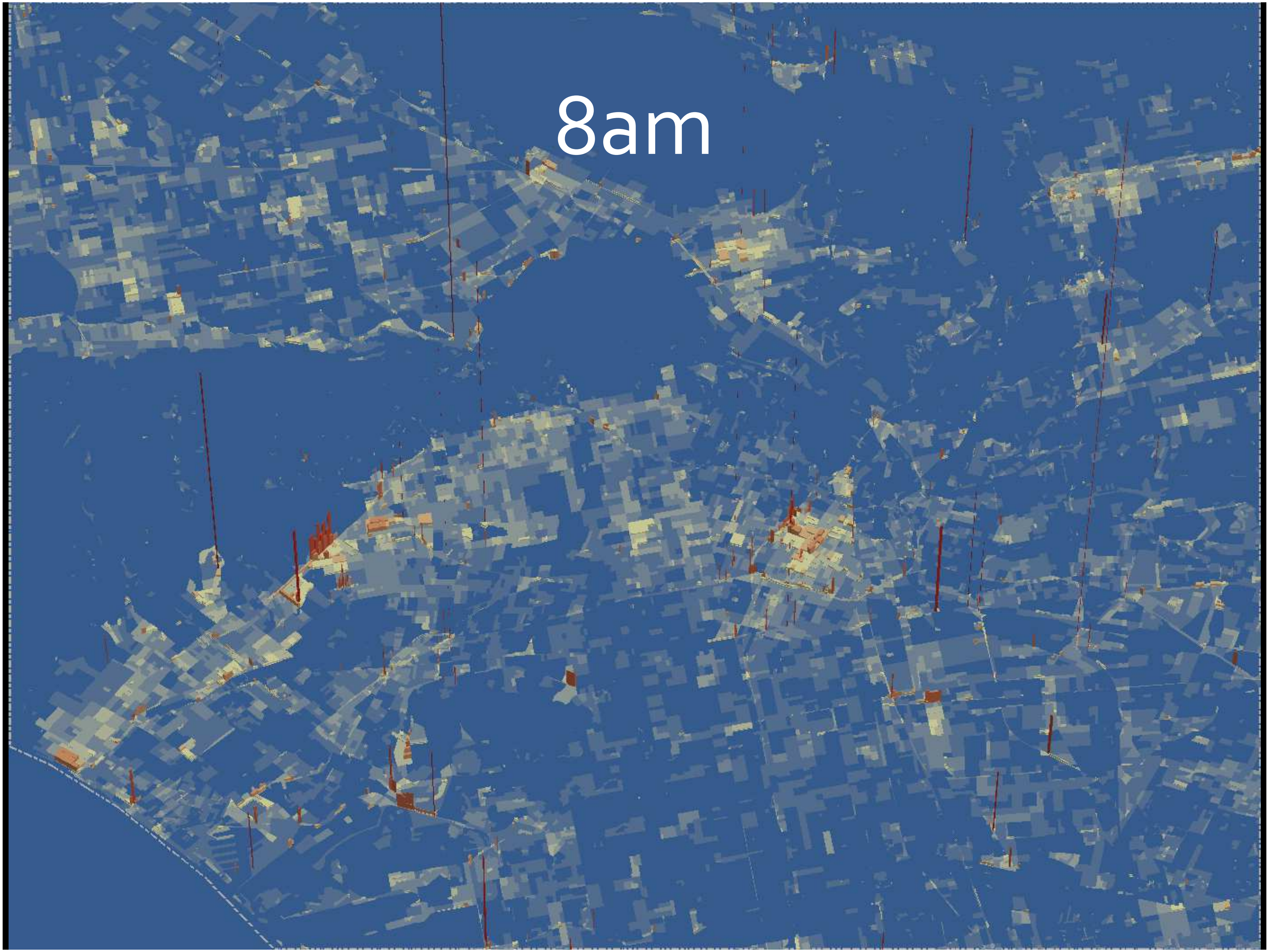

3am

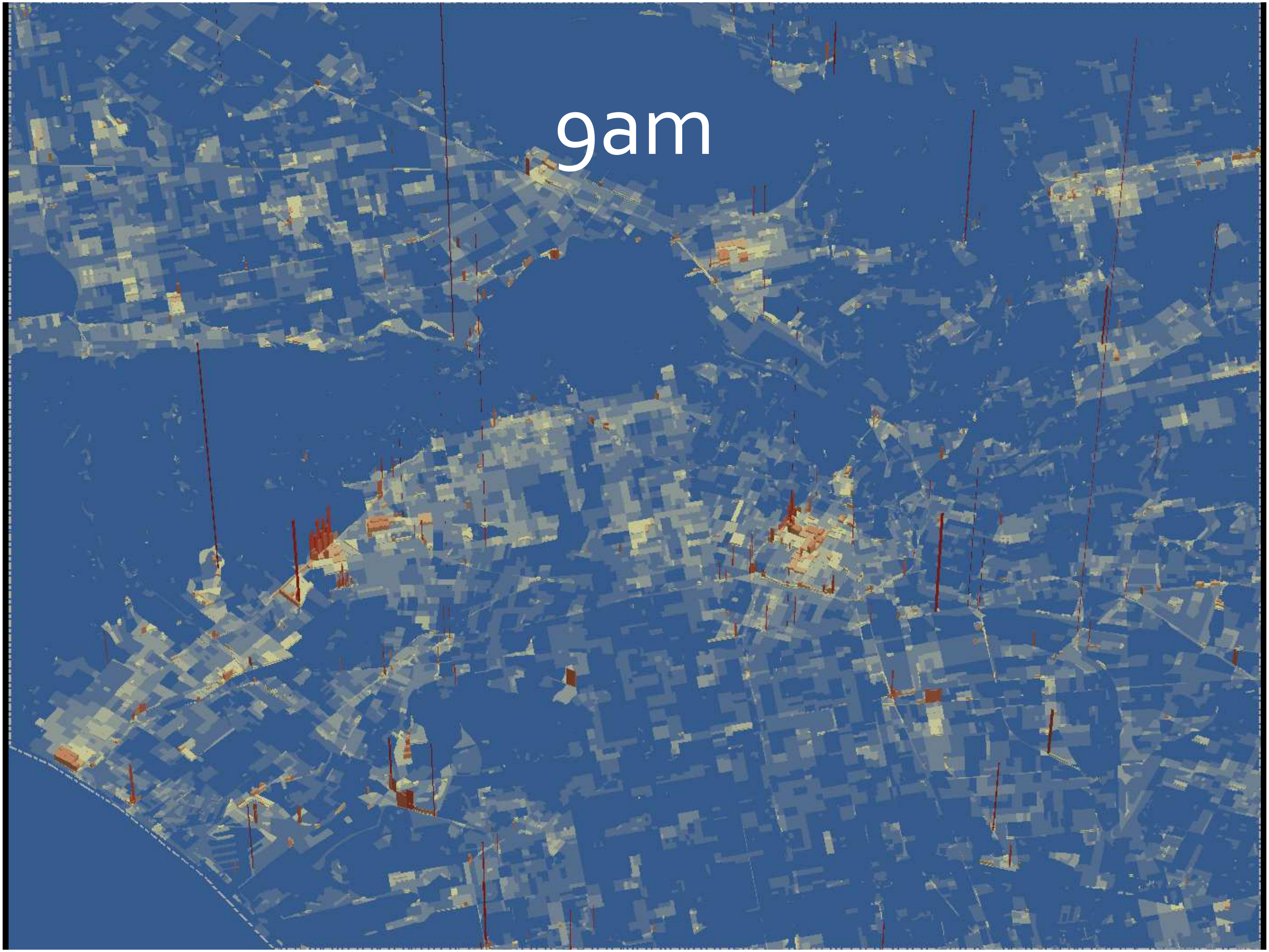

4am

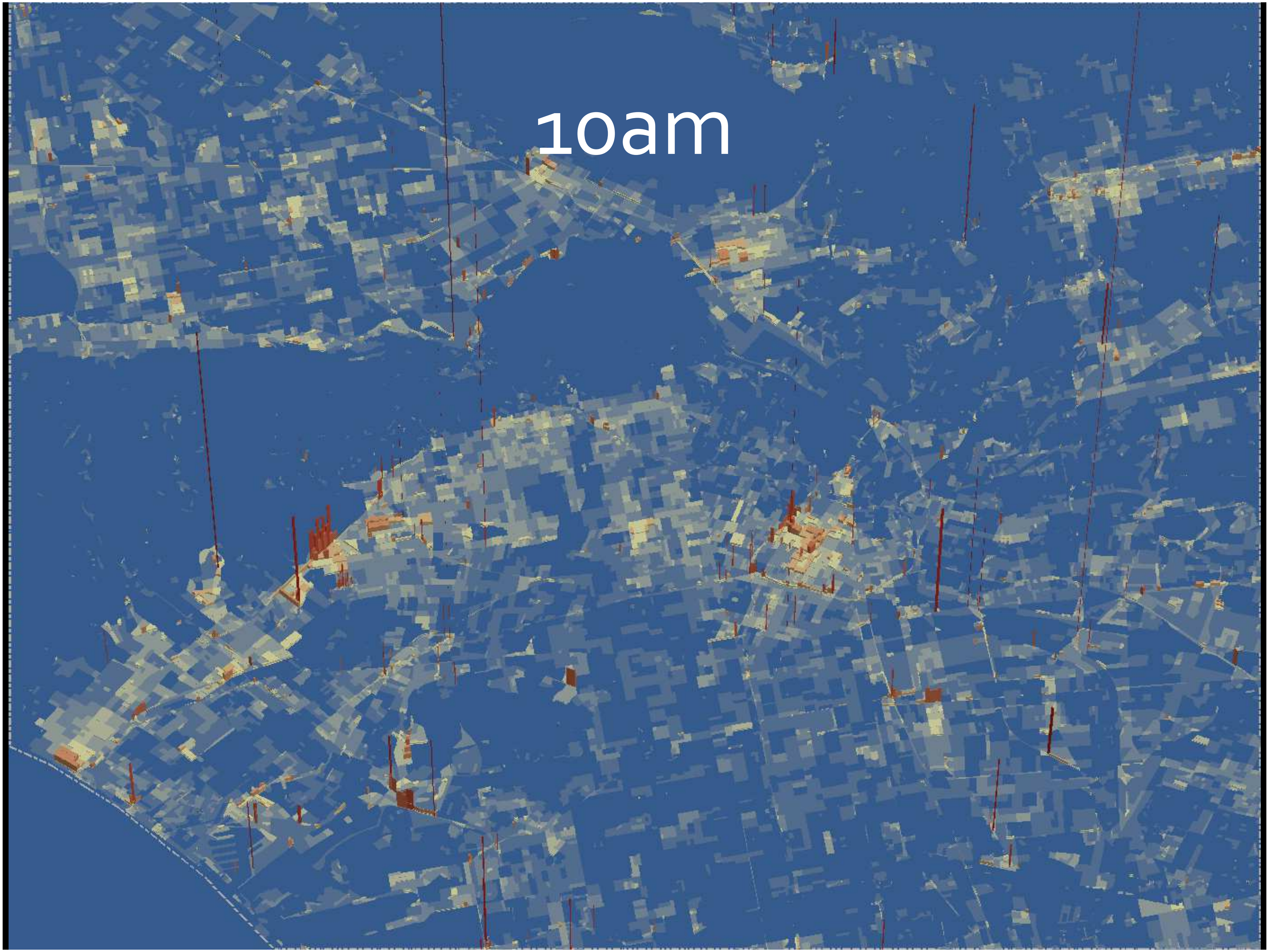

5am

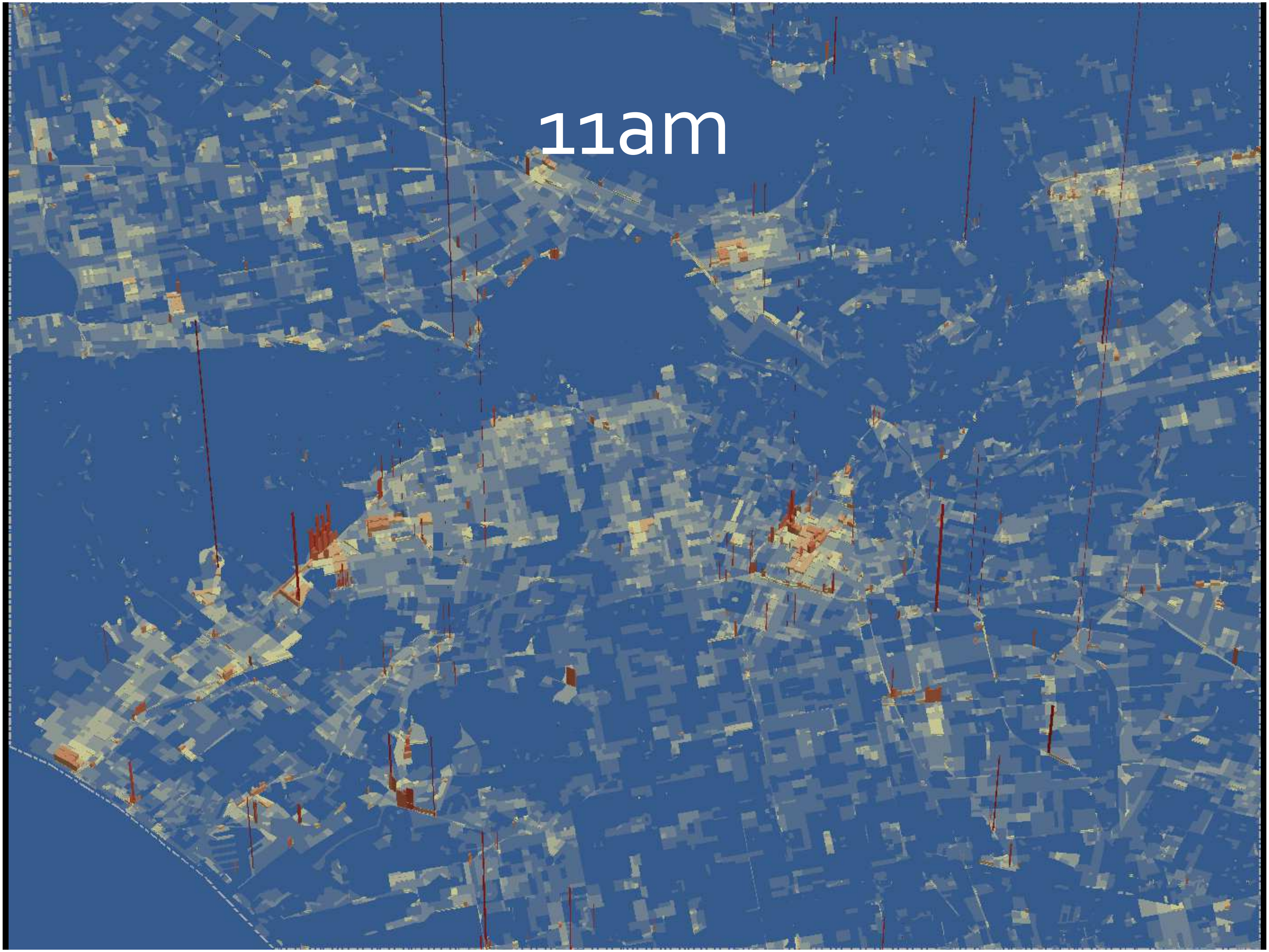

6am

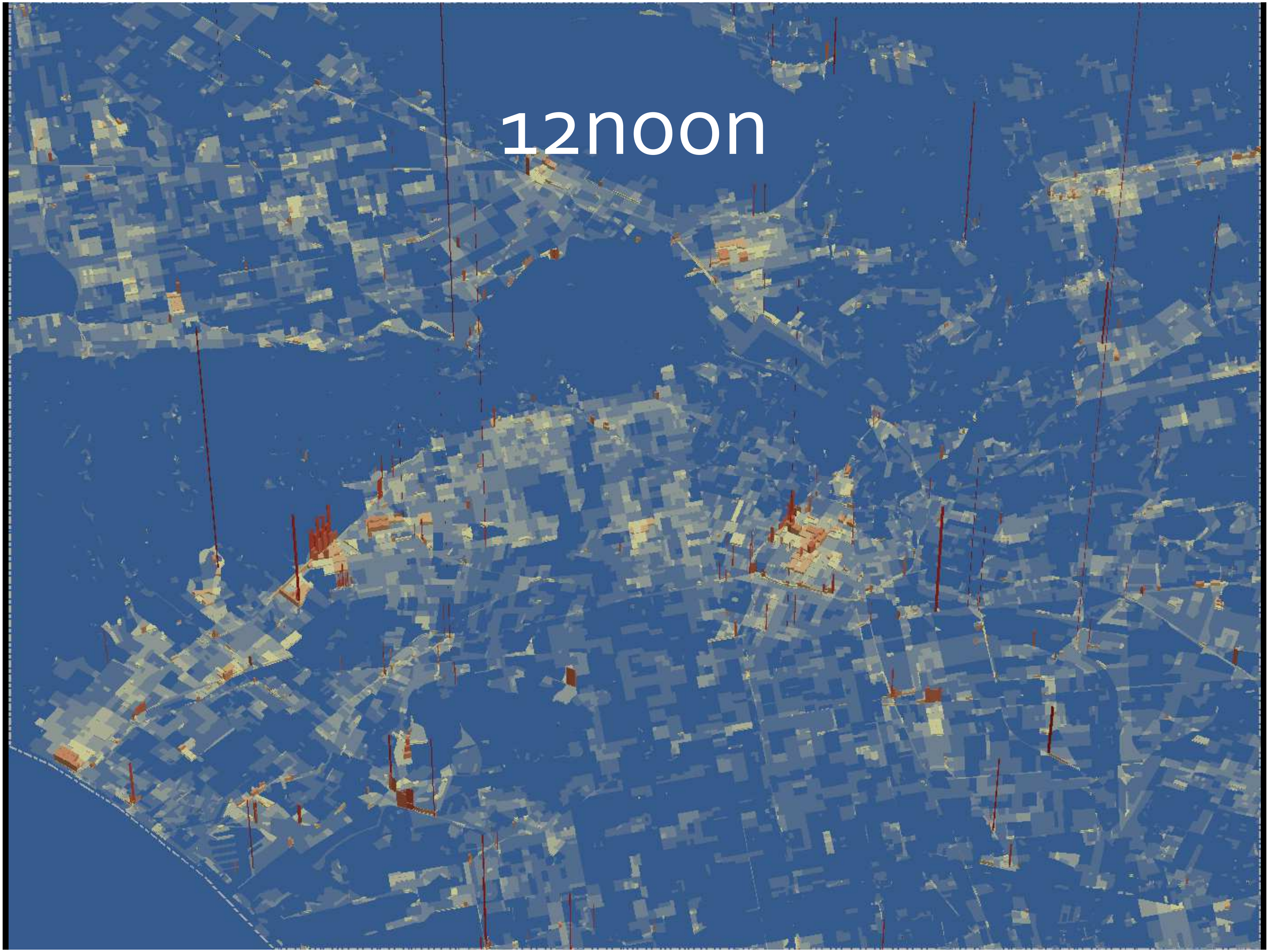

7am

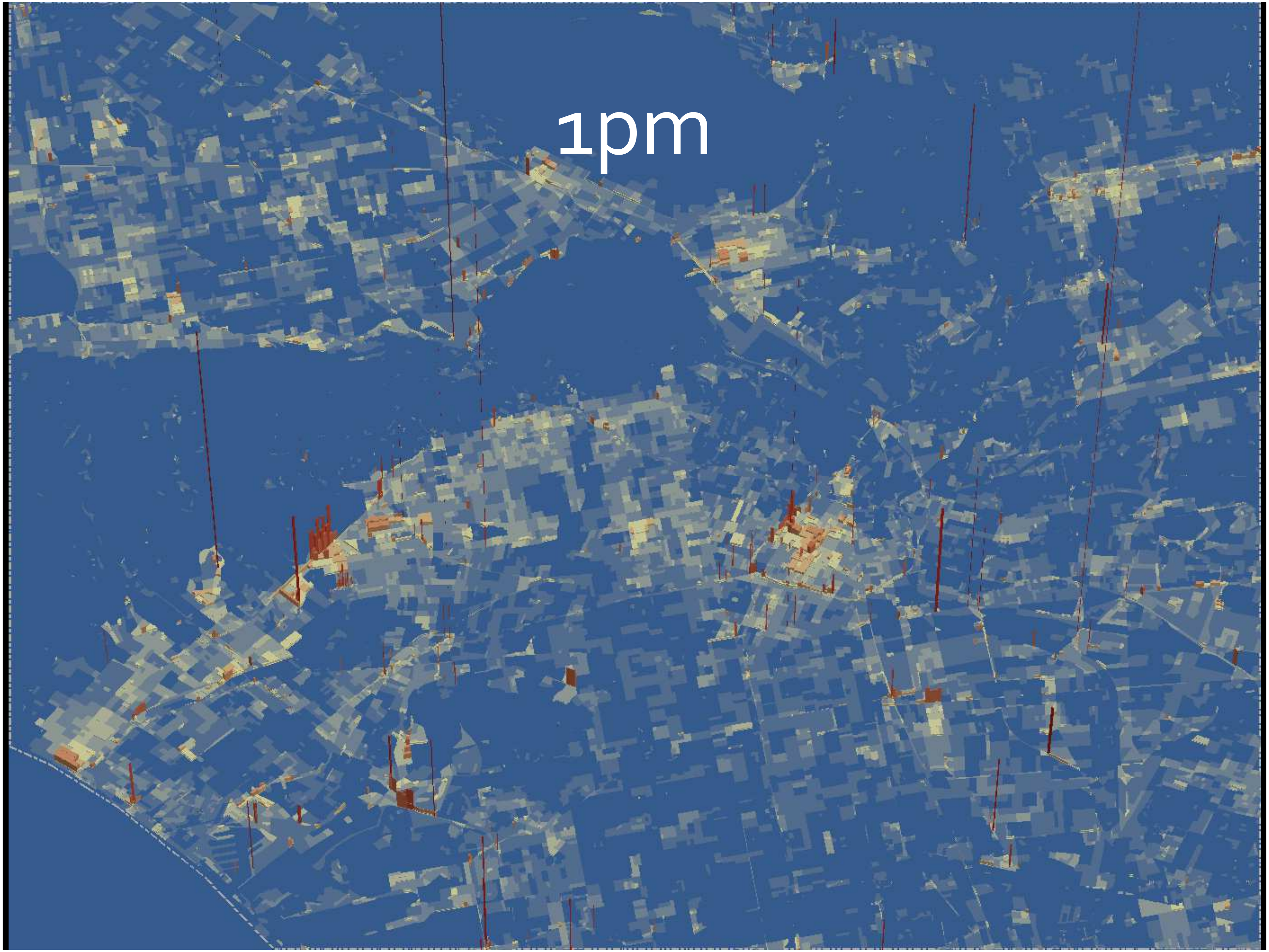

8am

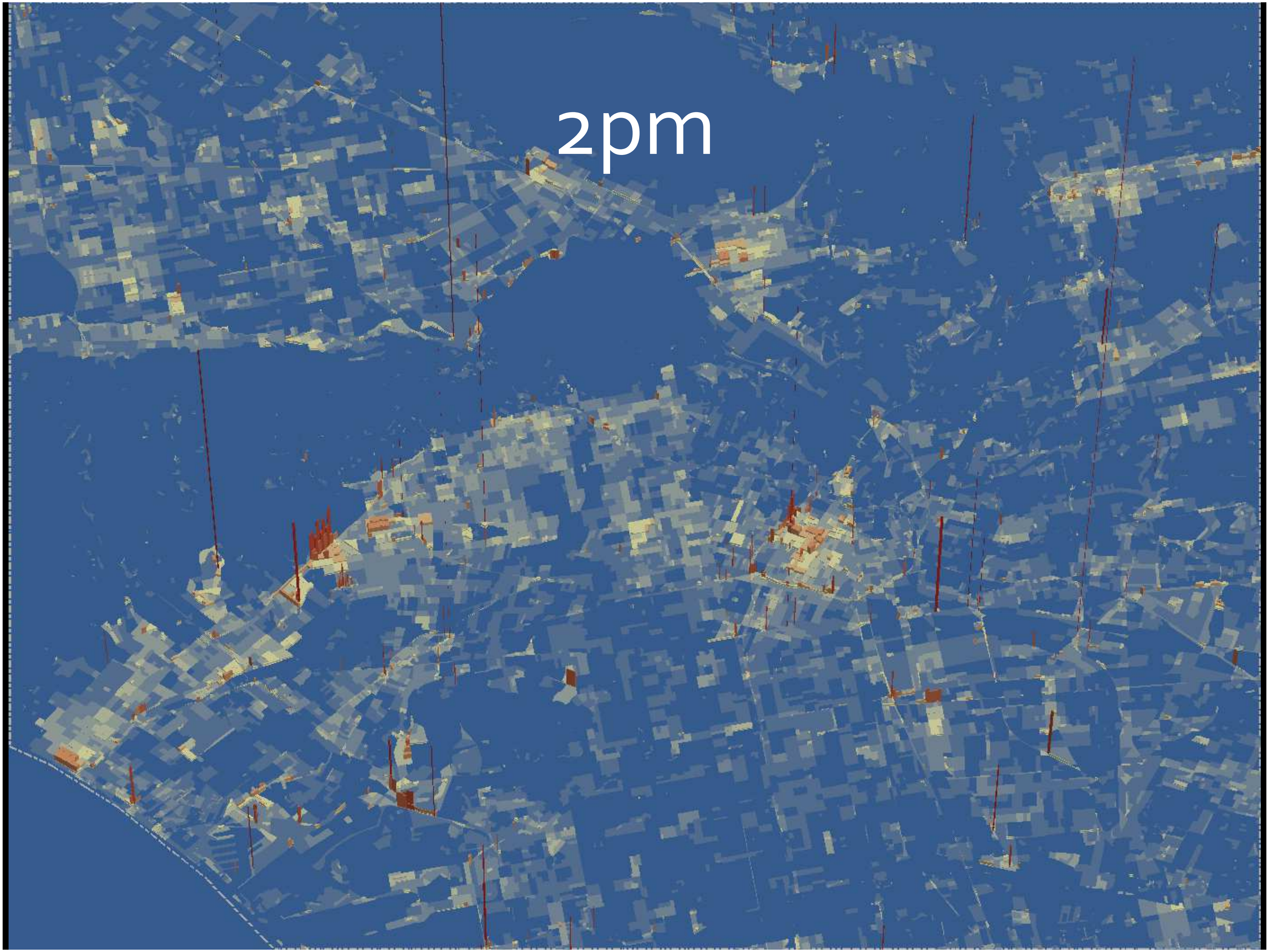

9am

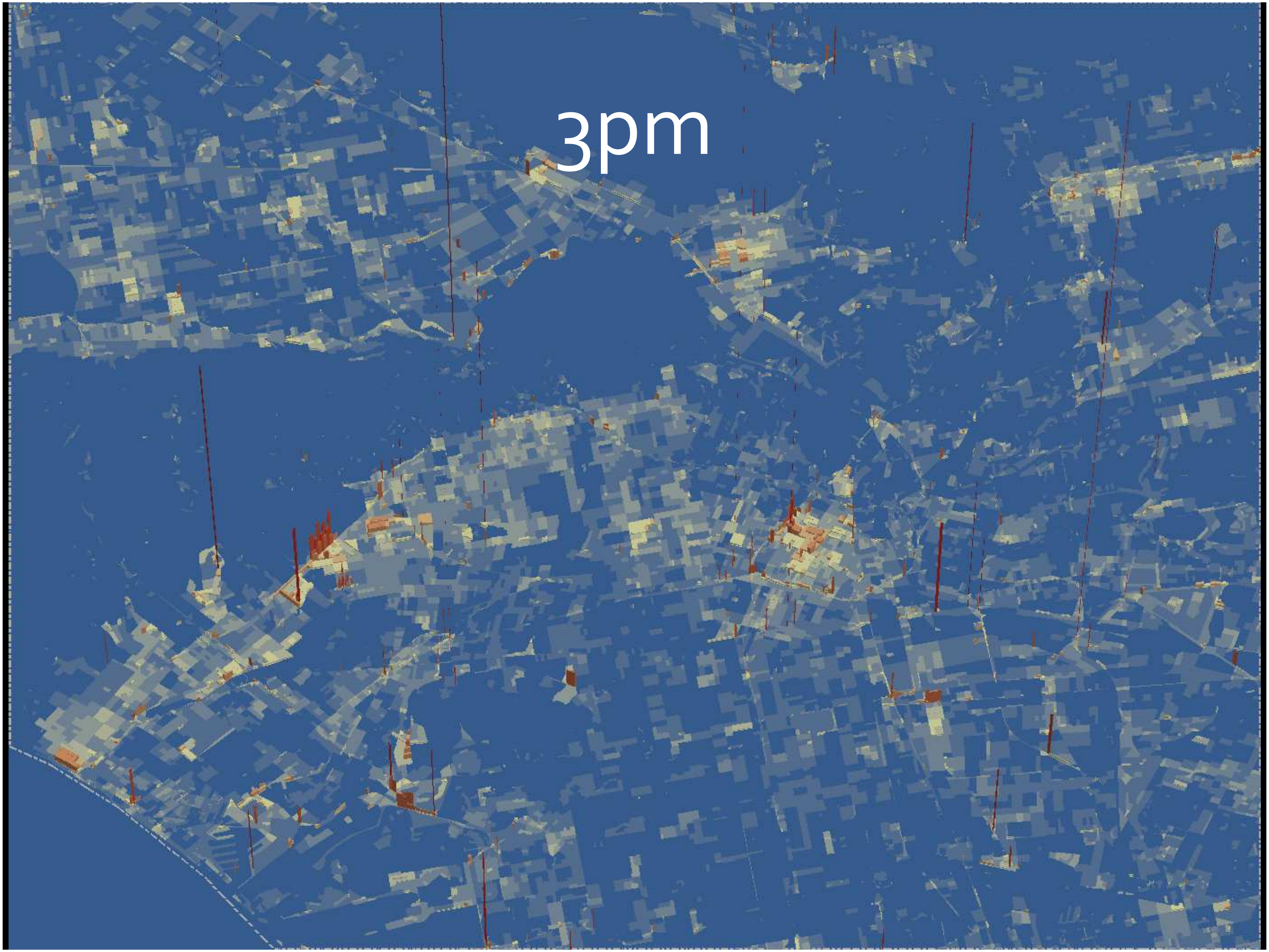

10am

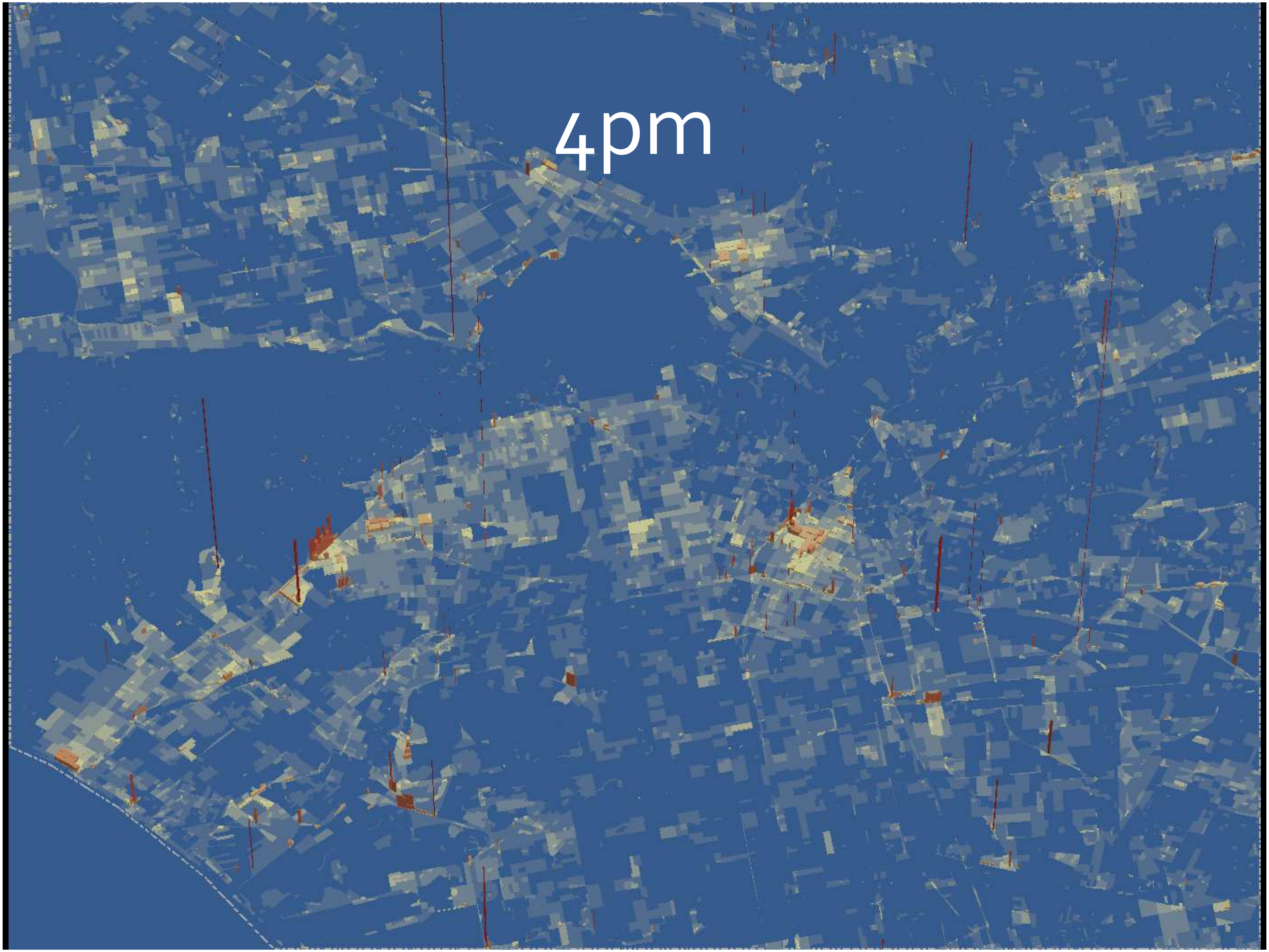

11am

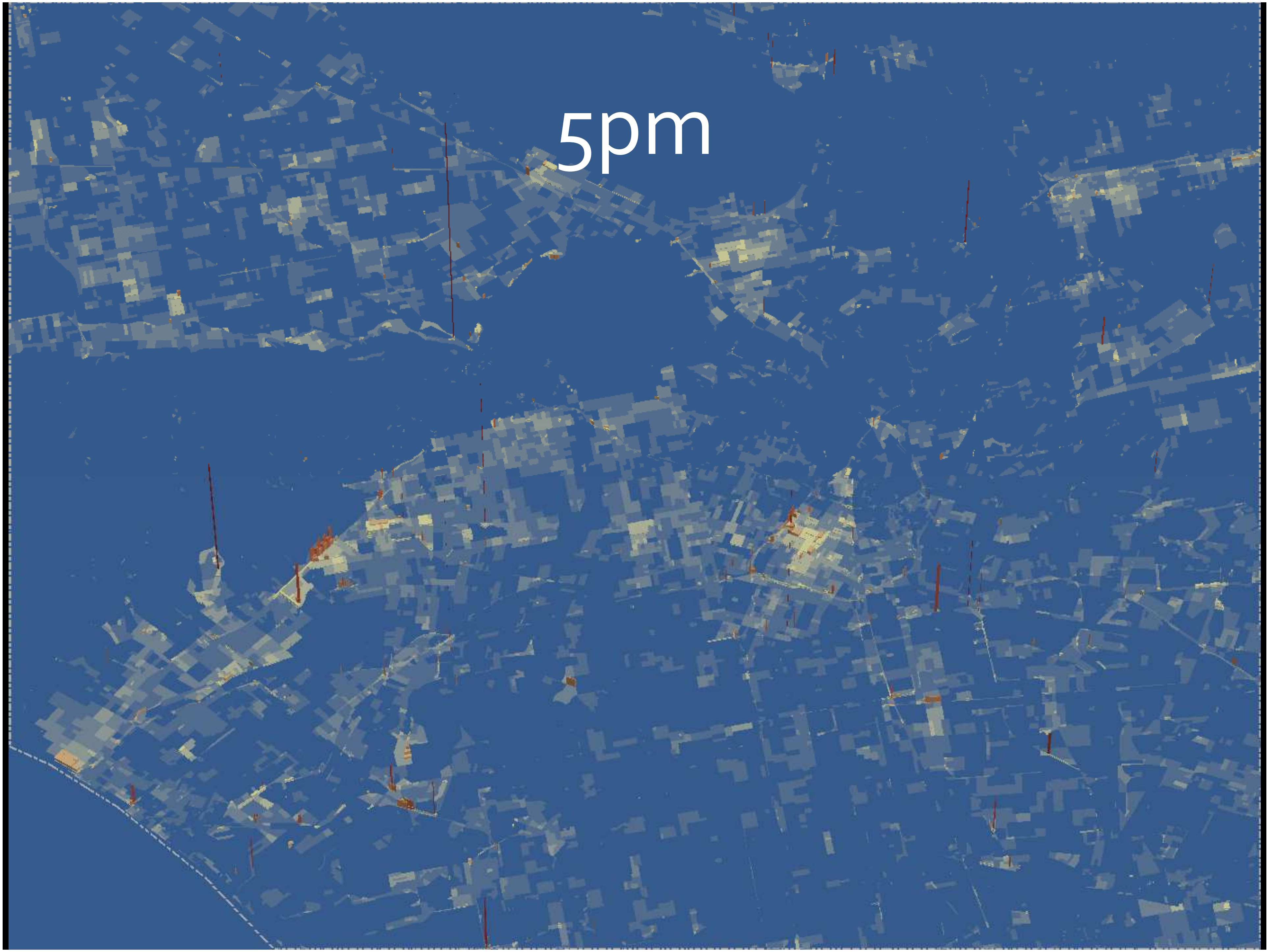

12noon

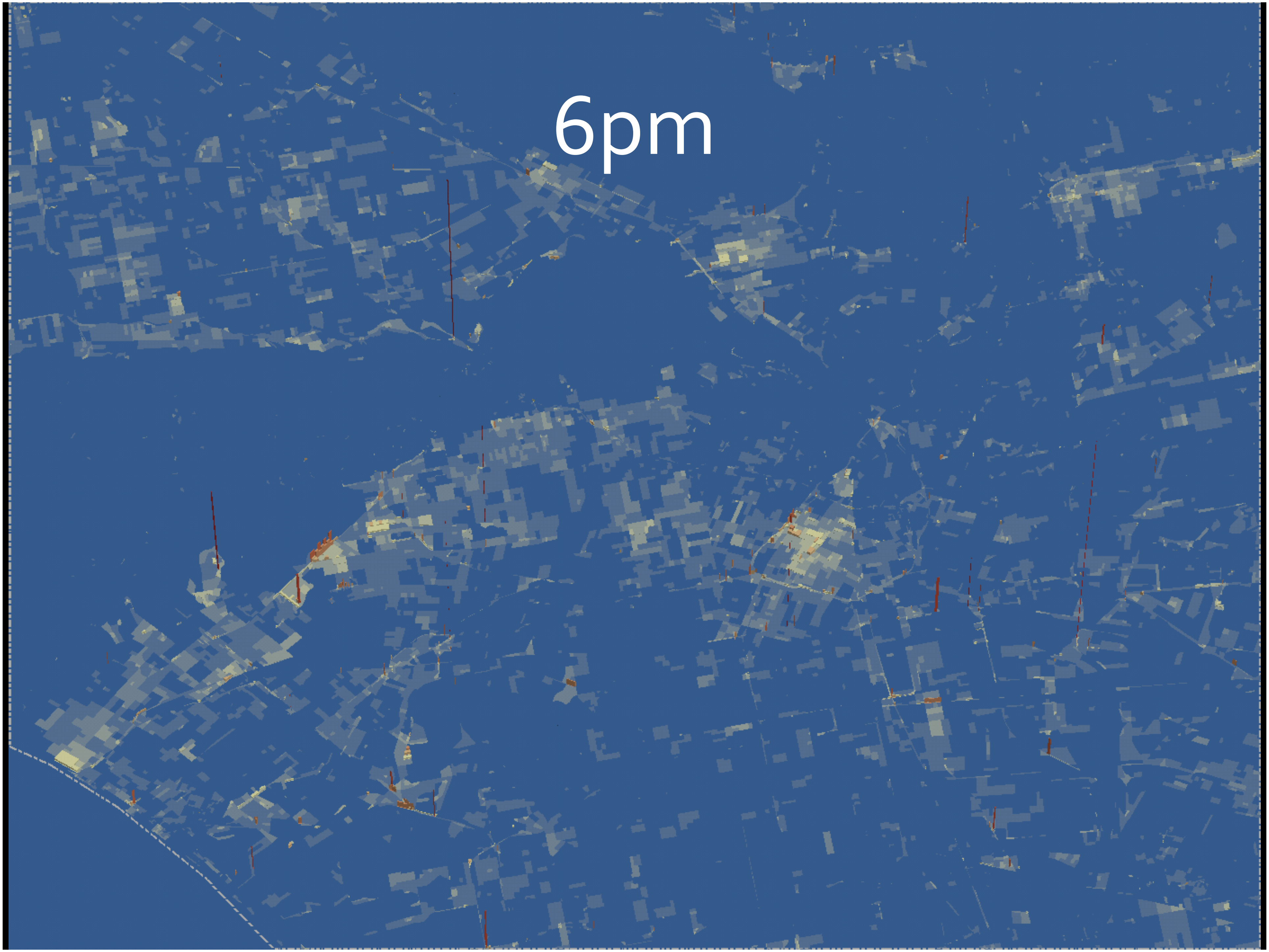

1pm

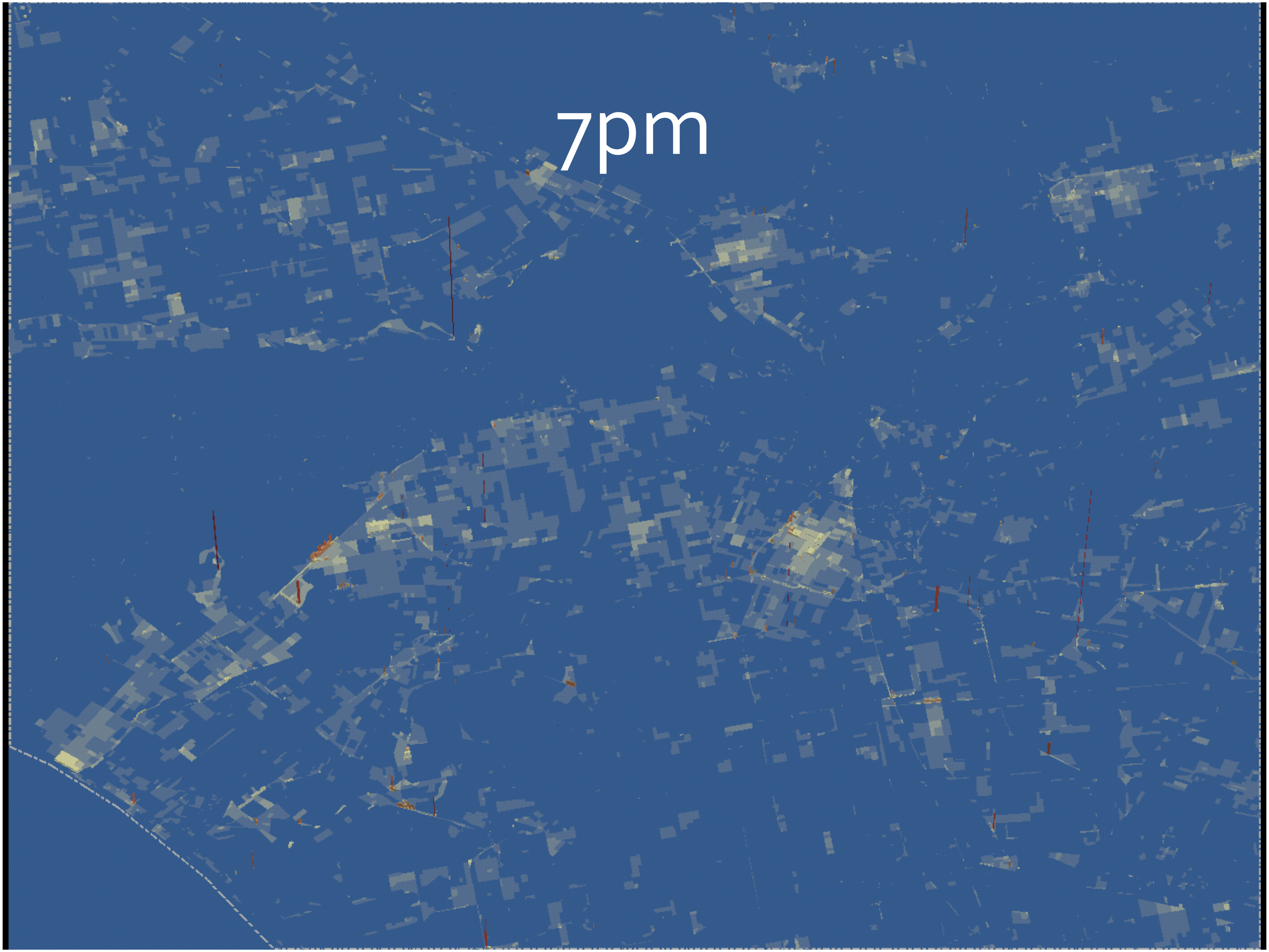

2pm

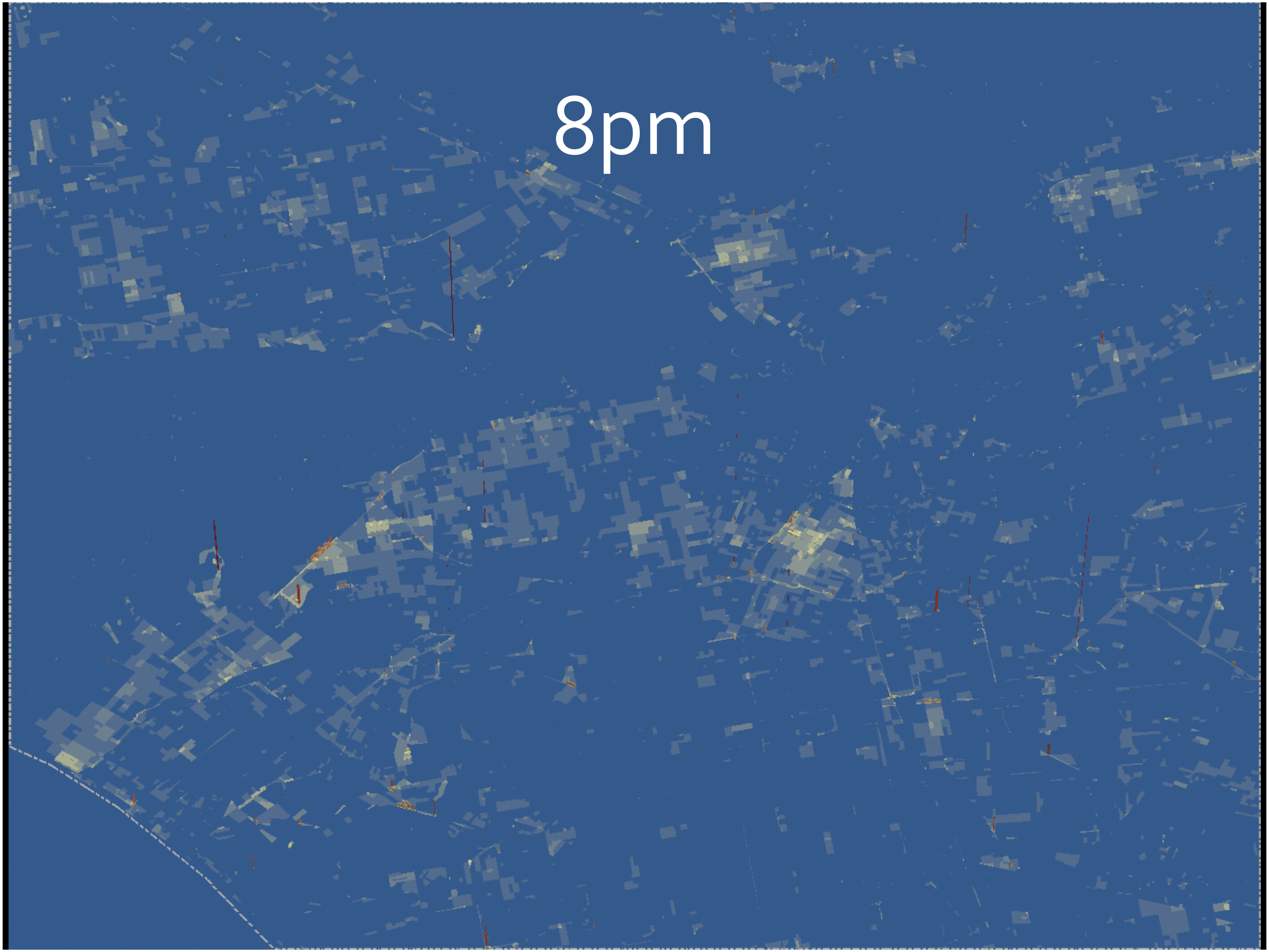

3pm

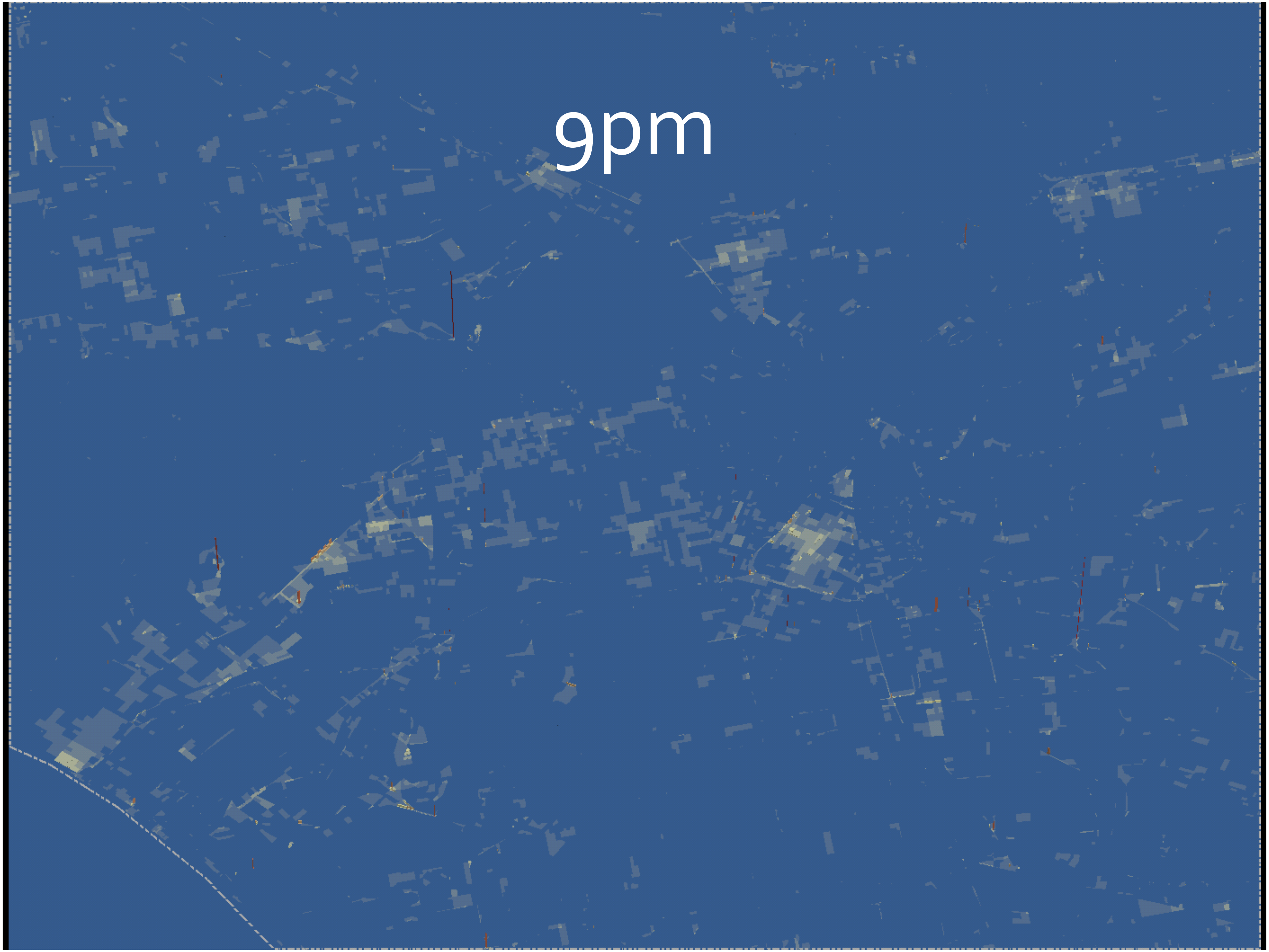

4pm

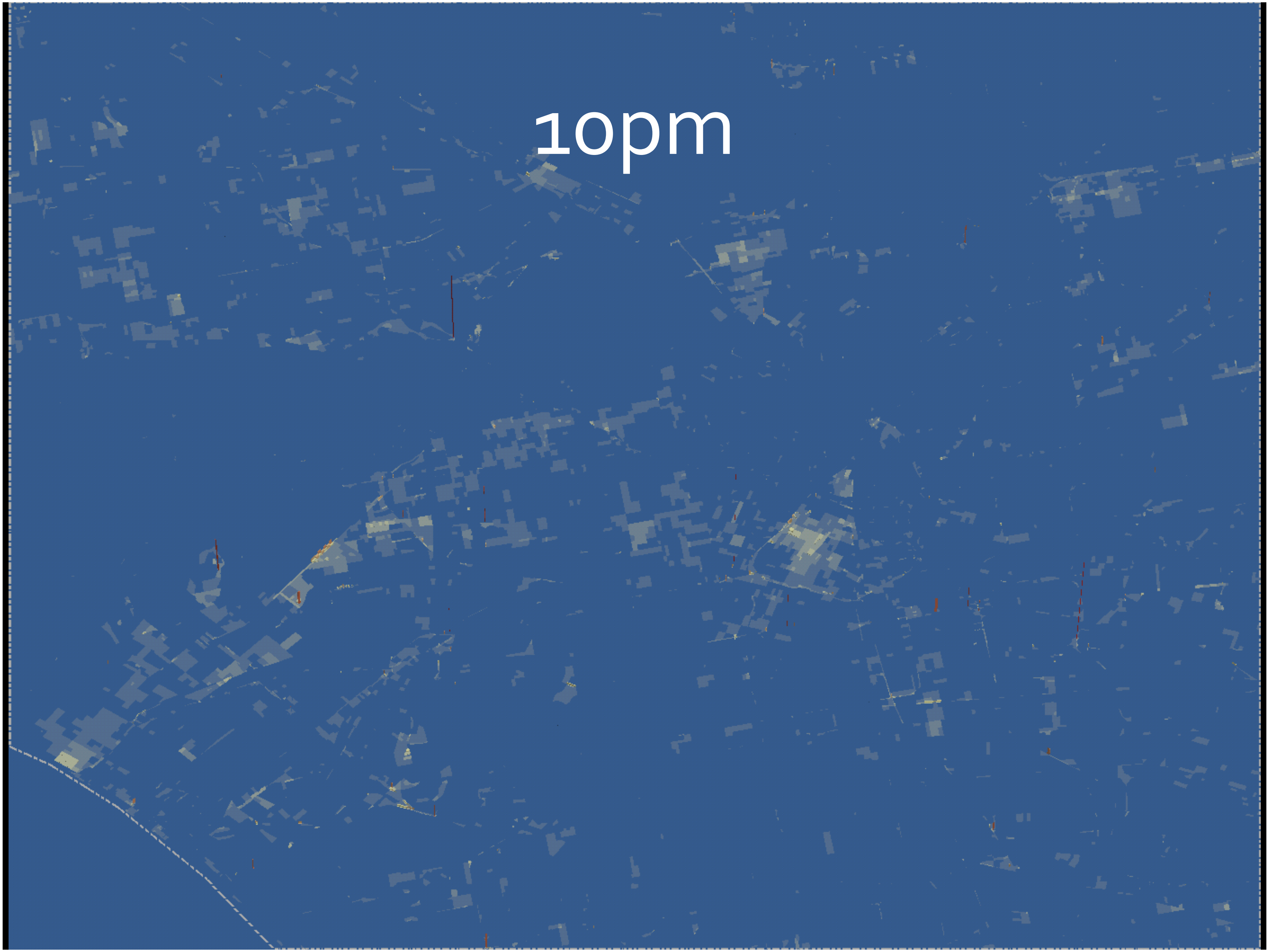

5pm

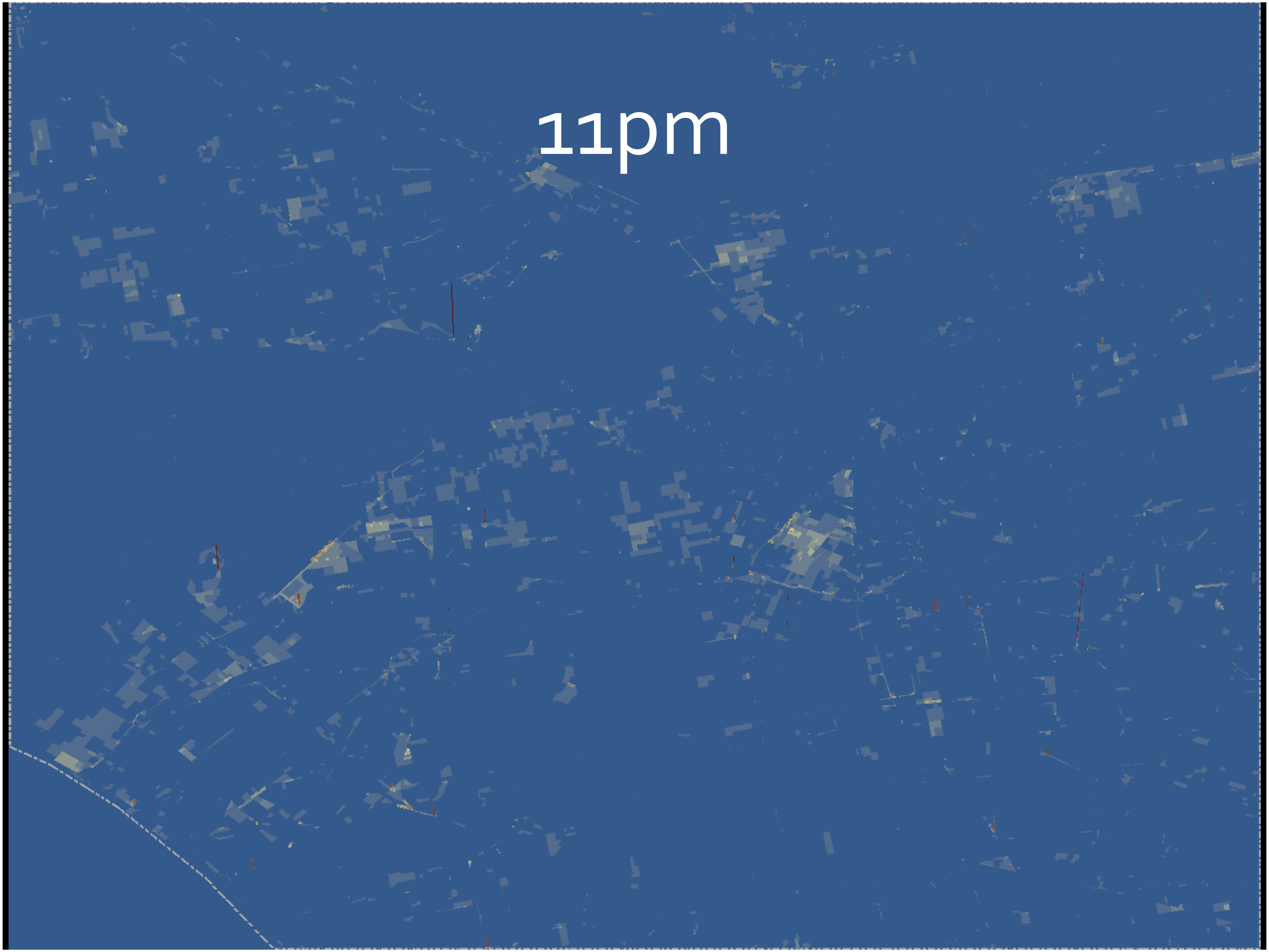

6pm


7pm

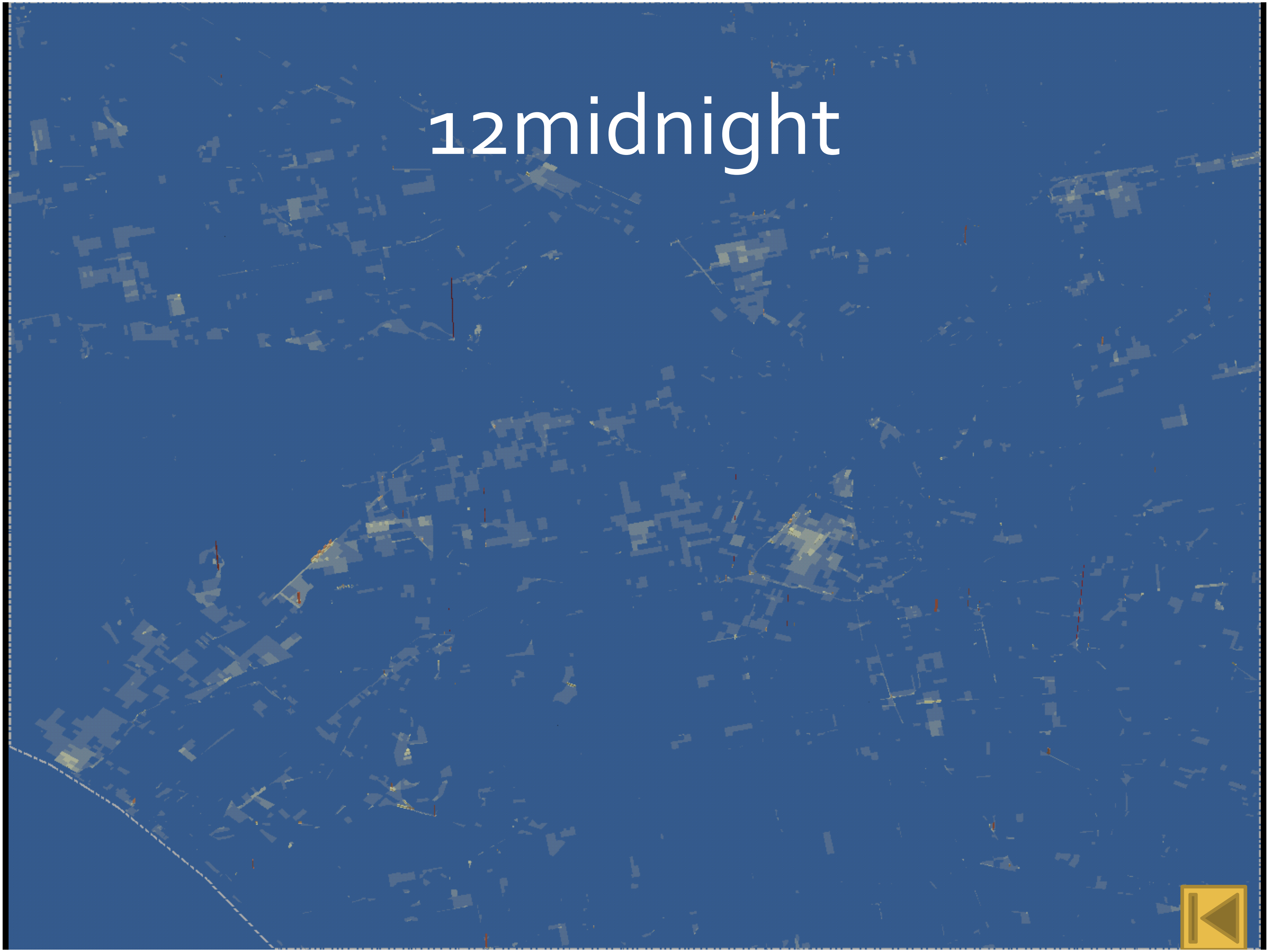

8pm


9pm


10pm

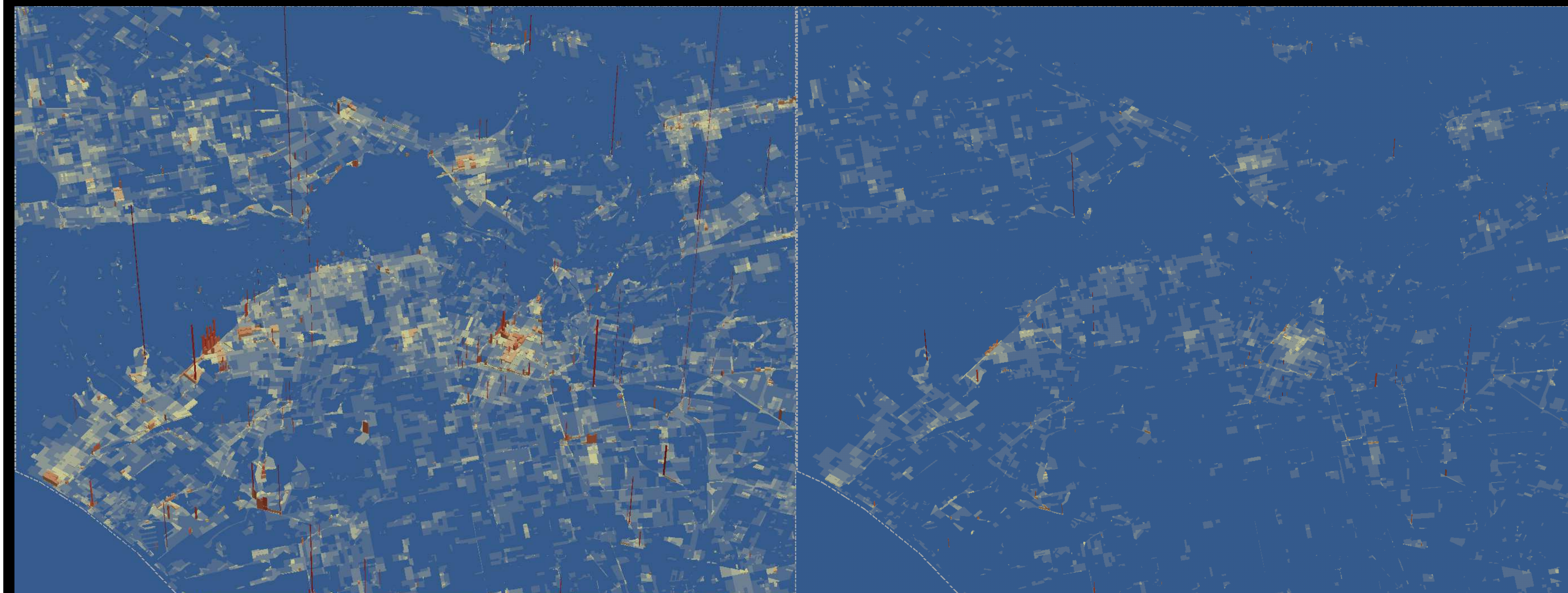

11pm

12midnight

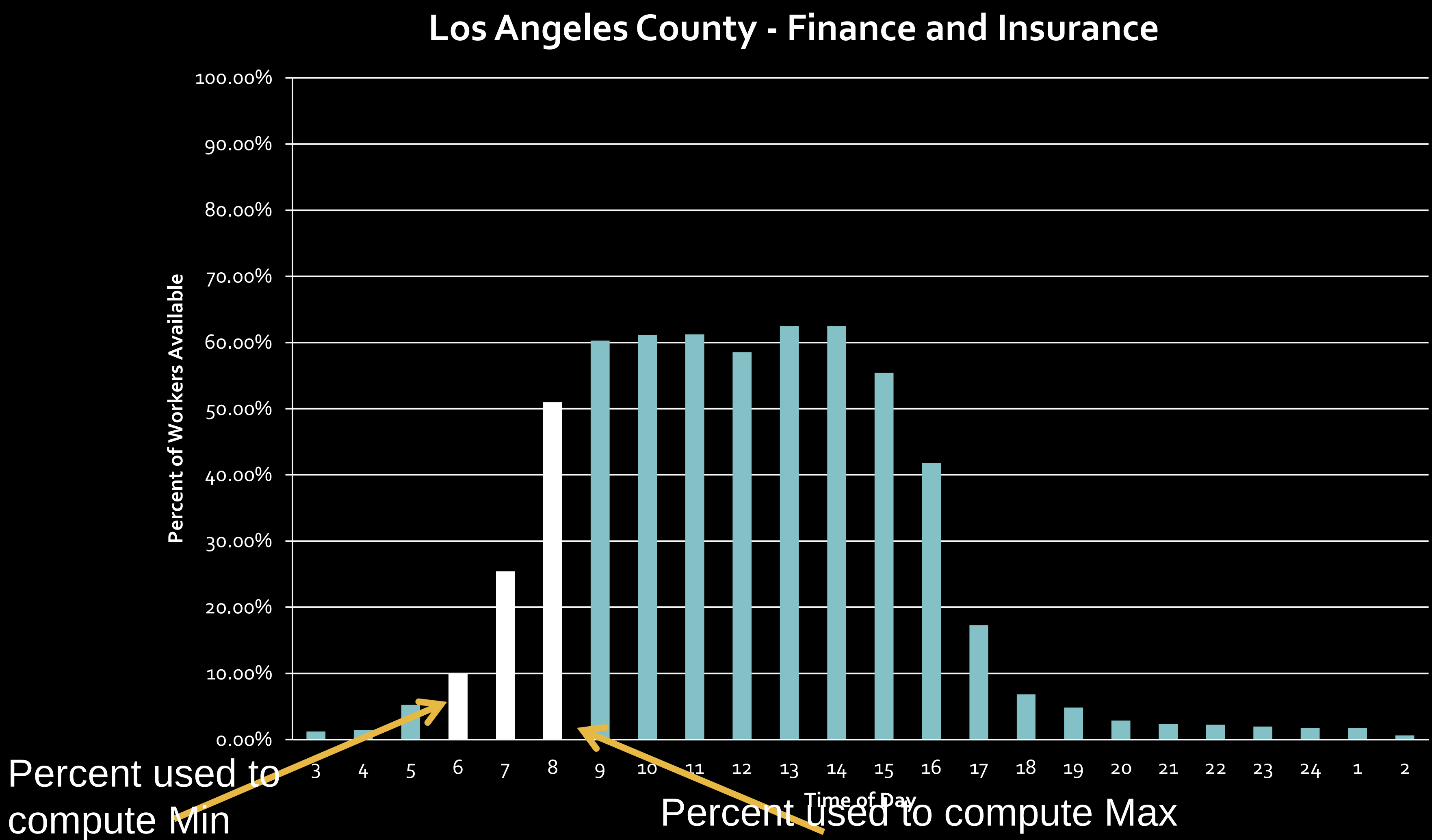


11AM

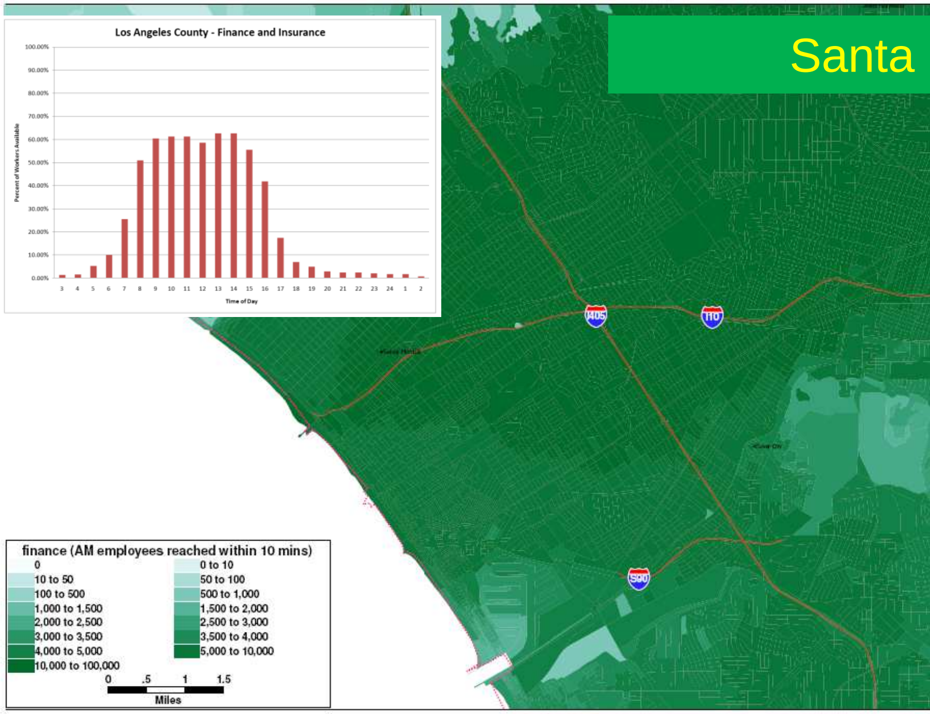
8PM

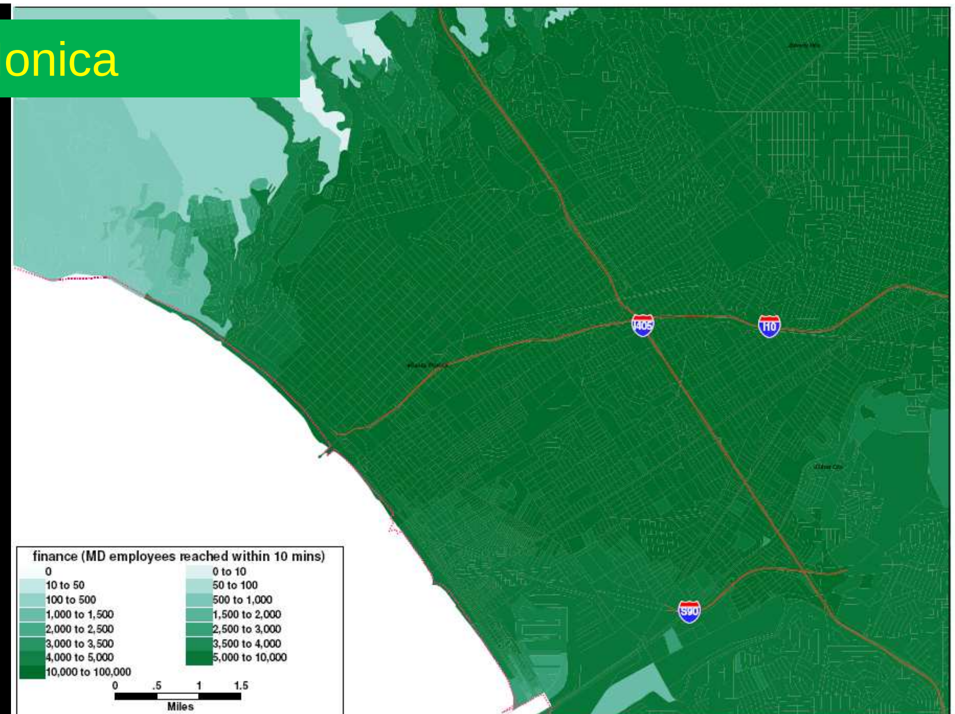


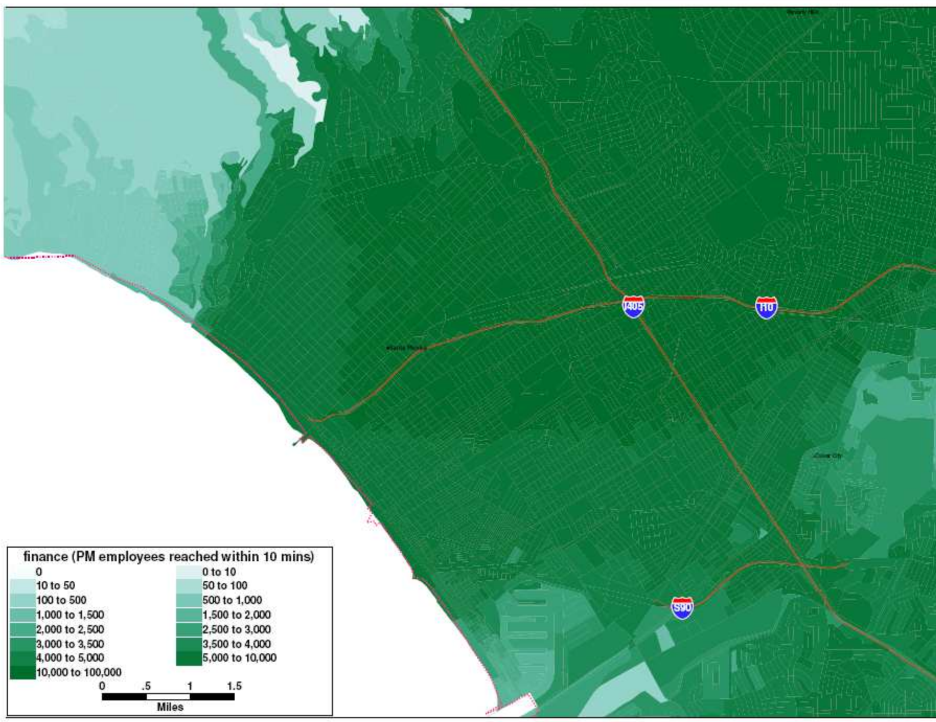
Min and Max in AM Peak

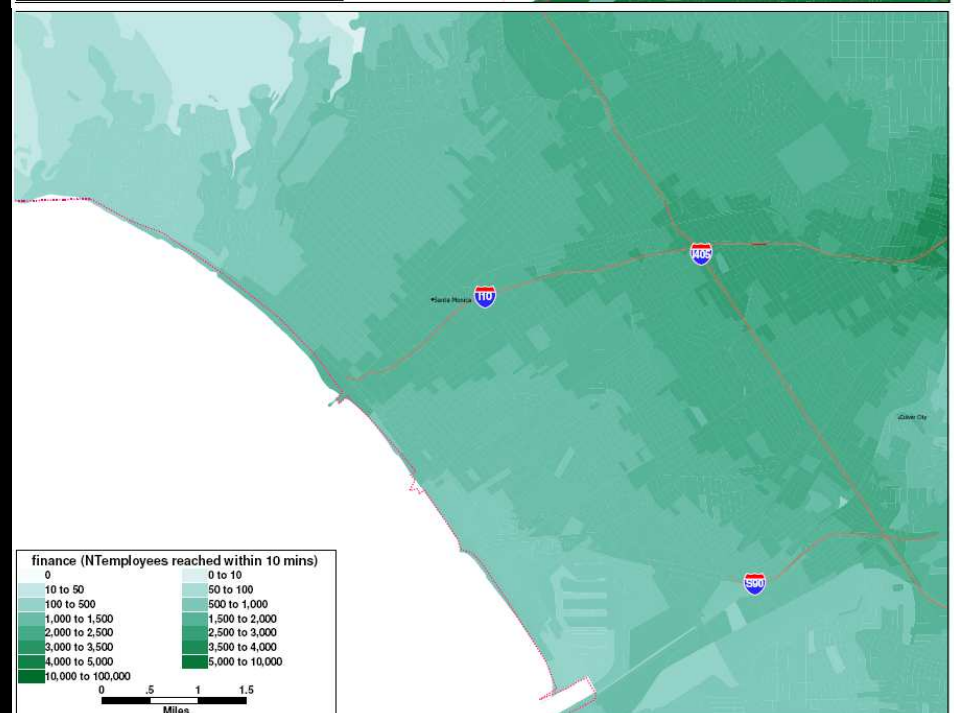


Santa Monica

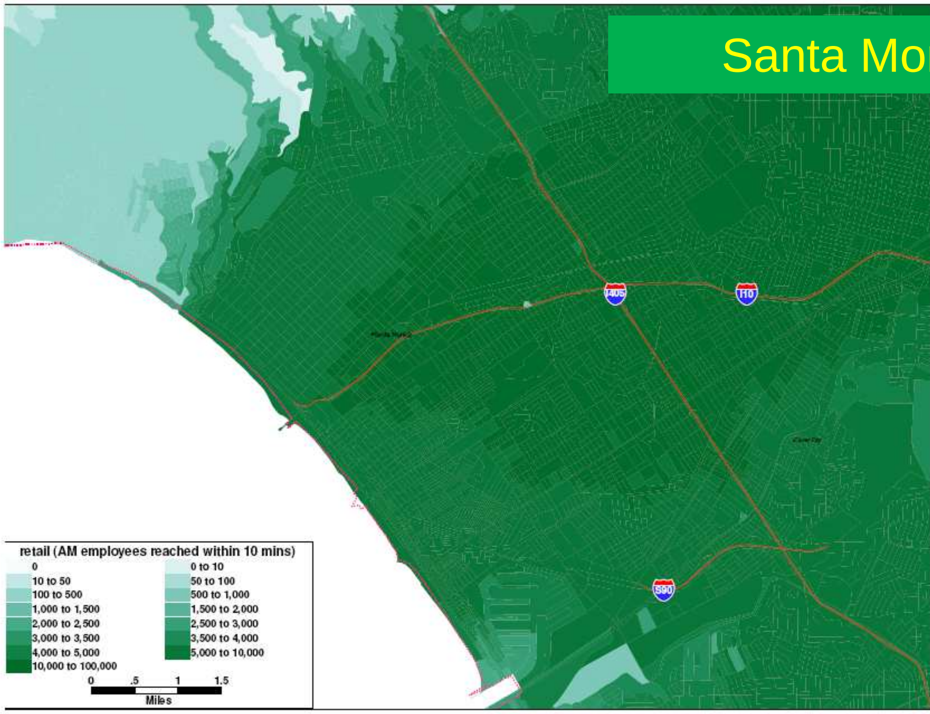


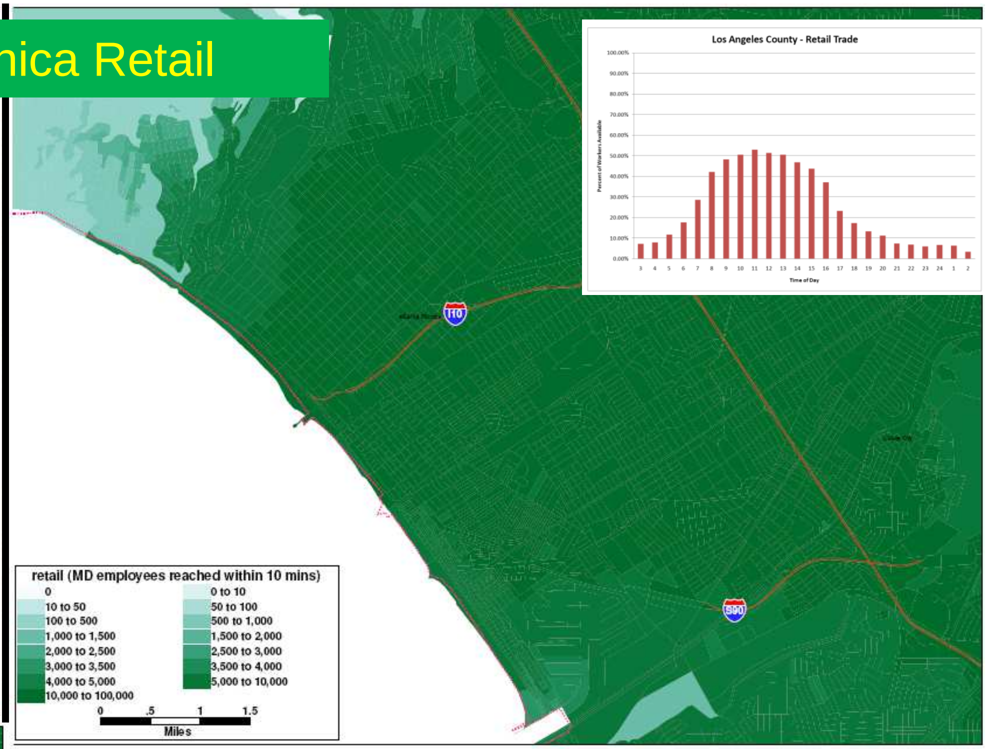


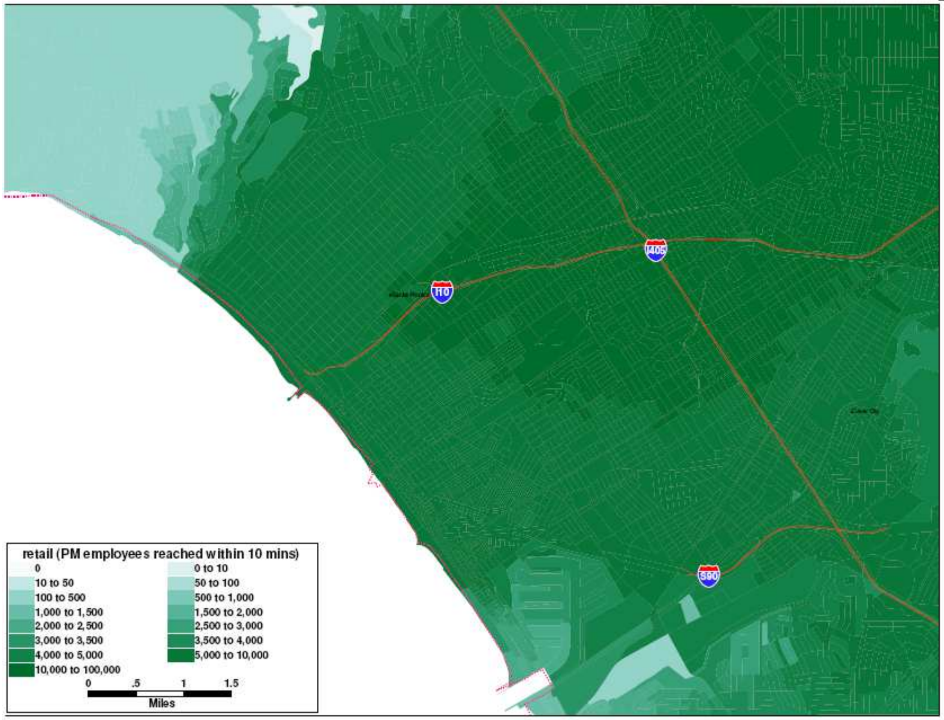


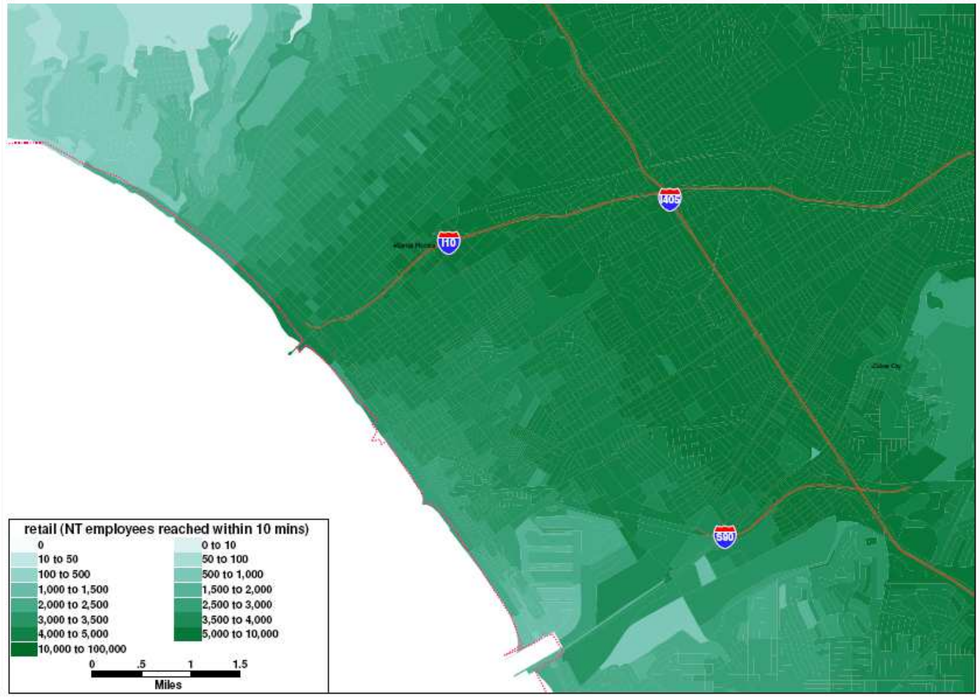


Santa Monica Retail

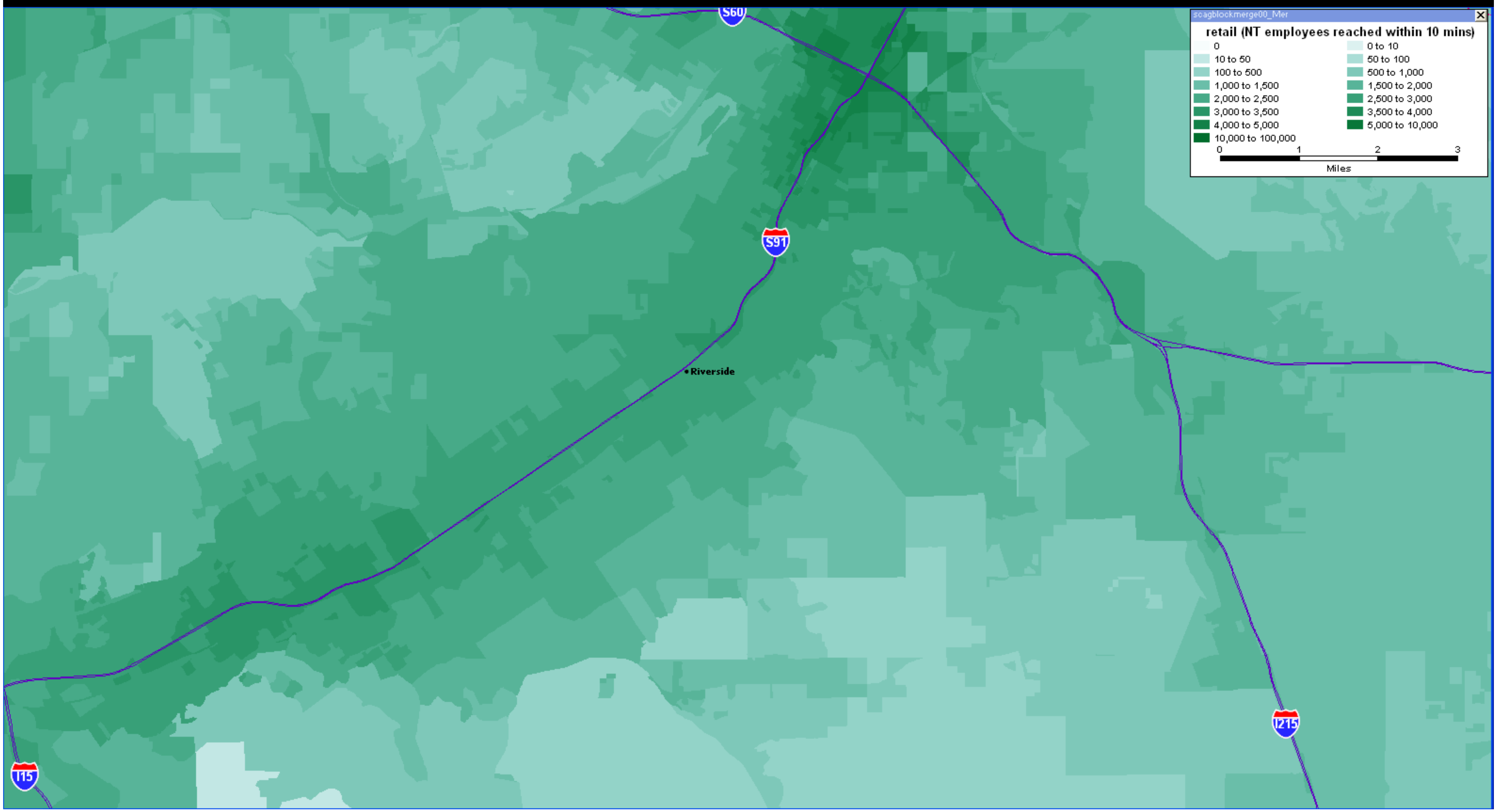




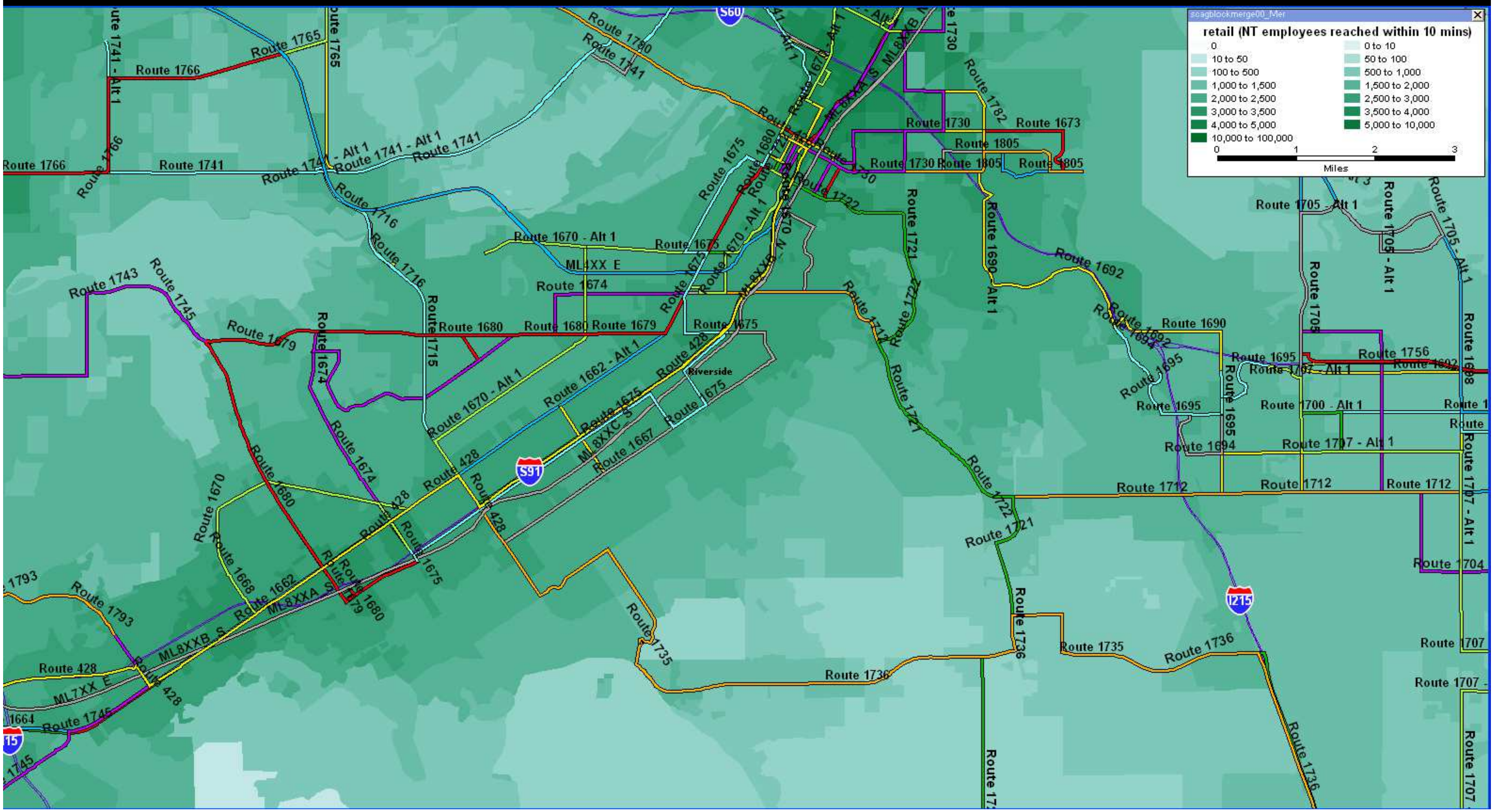




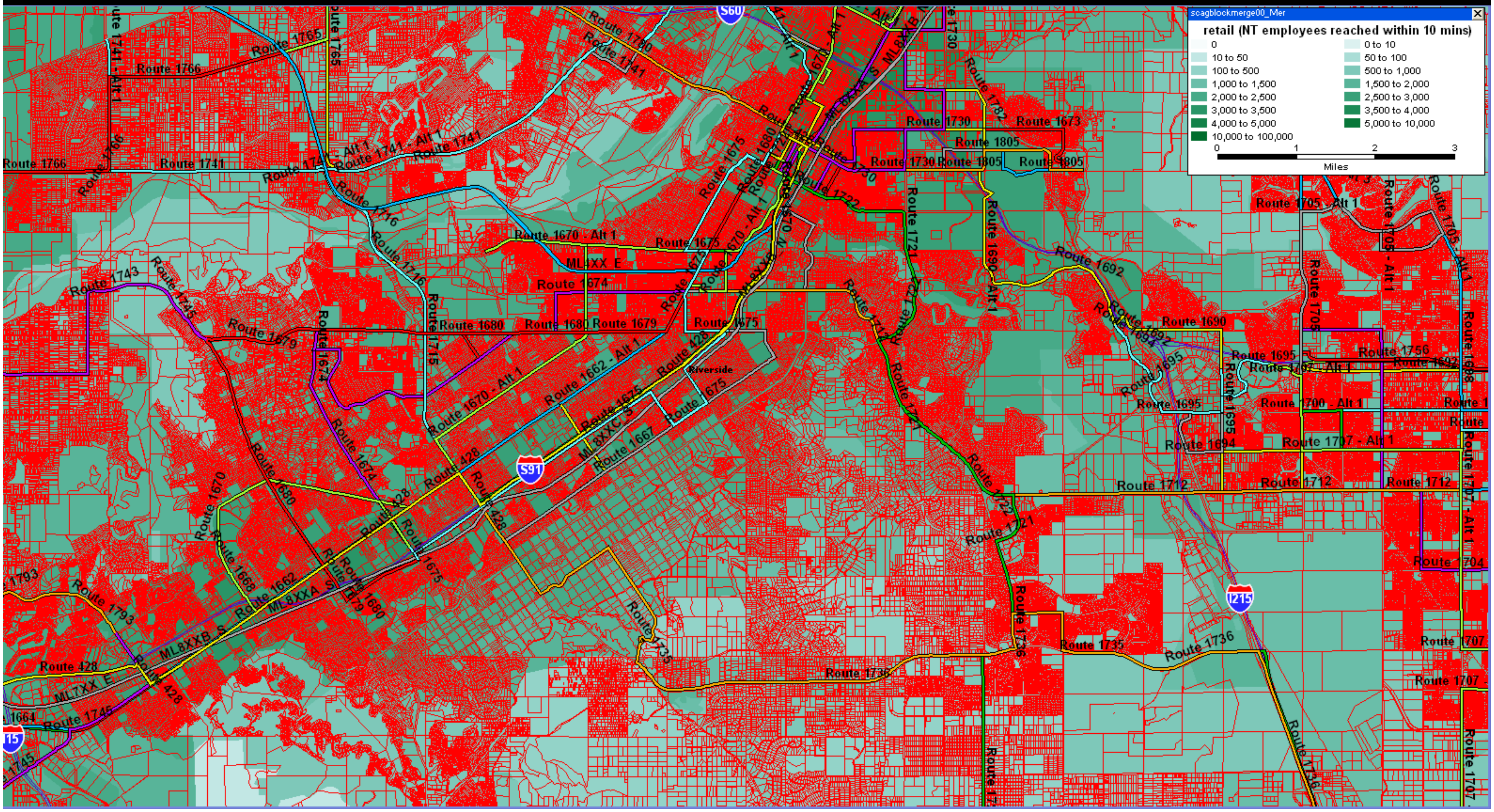
Riverside opportunity map
Max in 7PM to 6AM (Night Time) period of retail within 10 minutes for each block



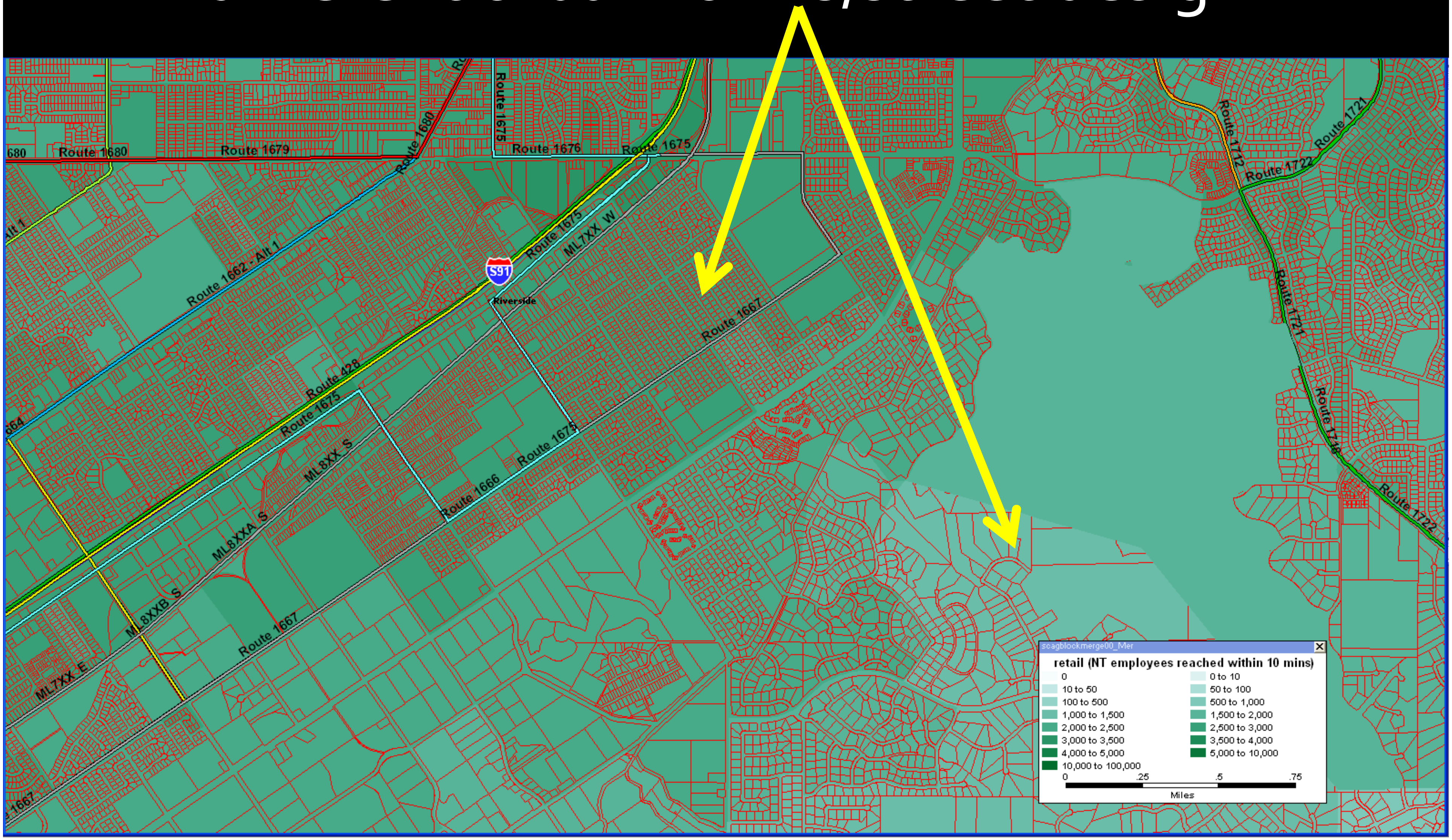
Superimposing Transit Routes



And the Land Parcels!



Compare retail accessibility between two different urban forms/street design



Accessibility and Policy

- Land-use strategies aiming at changing business location -> impact the opportunities available and the accessibility of zones for activity purposes.

- In modeling we need to represent the opportunities available as a measure of attraction

- But we may not have all this information available -> create a framework, test its feasibility, and then derive indicators at different levels – depending on data availability

- Enhance modeling framework by computing different accessibility measures and accommodating these in the re-estimation

Integration at Multiple Levels

- Land Use-Activity Opportunities
  - Spatial shifts/change in economic activity -→ spatial shifts in activity opportunities
  - Spatial shifts/change in employment → spatial shifts in work, residence, and school location choices
  - Spatial shifts in economic opportunity (jobs) → immigration/out-migration
  - Demographic evolution → location/relocation patterns, labor market changes, activity location preferences change and so on and on….
  - WE NEED A TIME SEQUENCED SPATIO-SOCIAL EVOLUTIONARY **SYSTEM!**
  - Verify all this with “clean” longitudinal data --> we are working with some noisy data

Feasibility of Using Time-Space Prism to Represent Available Opportunities and Choice Sets for Destination Choice Models

Based on: Yoon SY, K. Deutsch, and K.G. Goulias (2011) Feasibility of using time-space prism to represent available opportunities and choice sets for destination choice models in the context of dynamic urban environments. Paper 12-2707 presented at the *91st Annual Meeting of the Transportation Research Board, Washington, D.C.,* January 22-26, 2012.

In the previous slides we reviewed accessibility indicators based on:

- Spatial units = 203,621 blocks
- Highway network and travel time
  - almost every local road of the network with an estimate of travel time between blocks
- Temporal units
  - Estimated speeds based on 4 time periods (AMpeak, PMpeak, Midday, and Night time) for travel time
  - 24 time periods for opportunity availability (based on opening and closing time)
  - Activity schedules by minute from an ABM (not reviewed here but in other seminars)

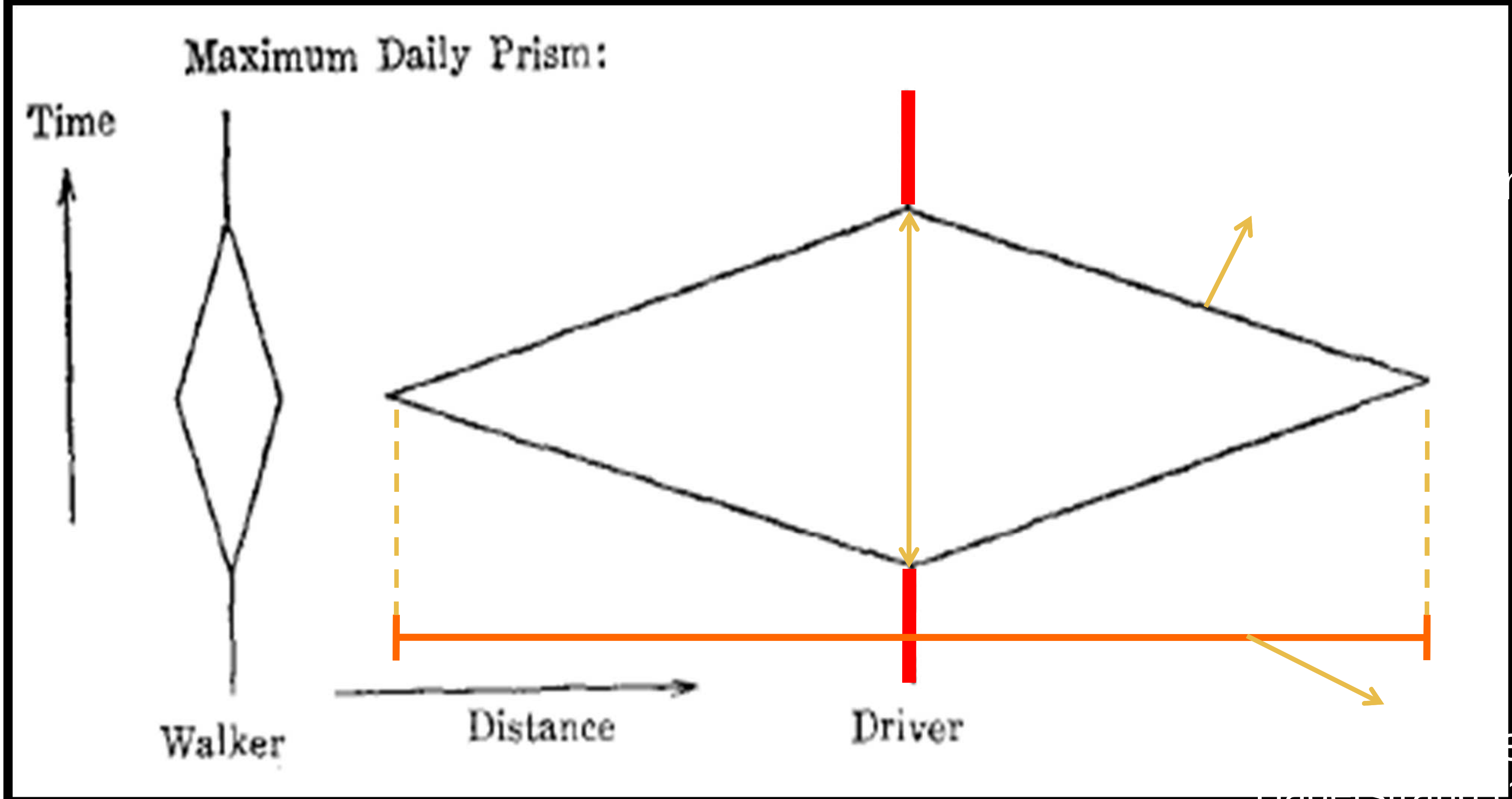
Components considered

- Available opportunities by time of day
  - 24 time periods and 15 industry types
  - a) Agriculture, forestry, fishing and hunting and mining; b) Construction; c) Manufacturing; d) Wholesale trade; e) Retail trade; f) Transportation and warehousing and utilities; g) Information; h) Finance, insurance, real estate and rental and leasing; i) Professional, scientific, management, administrative, and waste management services; j) Educational; k) Health; l) Arts, entertainment, recreation, accommodation and food services; m) Other services (except public administration); o) Public administration; p) Armed forces
- Travel time by time of day
  - 4 time periods
  - AM peak (6am-9am), PM peak (3pm-7pm), off-peak (9am-3pm), and night time (7pm-6am)
- Individual differences in activity scheduling

Data Used

- **Household Travel Survey for SCAG** (collected from spring 2001 to spring 2002) containing 16,939 households and 40,376 individuals.
- **Detailed roadway network including time dependent travel speed/travel time**
- **Employment data of 15 industry types** that combine the US Census Transportation Planning Package (CTPP) of year 2001 and Dun & Bradstreet (D&B)
- **Geographic information of the census blocks and the block groups** in SCAG region (available from the Census website, http://www2.census.gov/cgi-bin/shapefiles2009/state-files?state=06) (for more details see Chen et al, 2011)
- **InfoUSA** that is a database enumerating all the business establishments in the region
- **Land parcel database** to locate and enumerate major shopping centers

Time-space prism

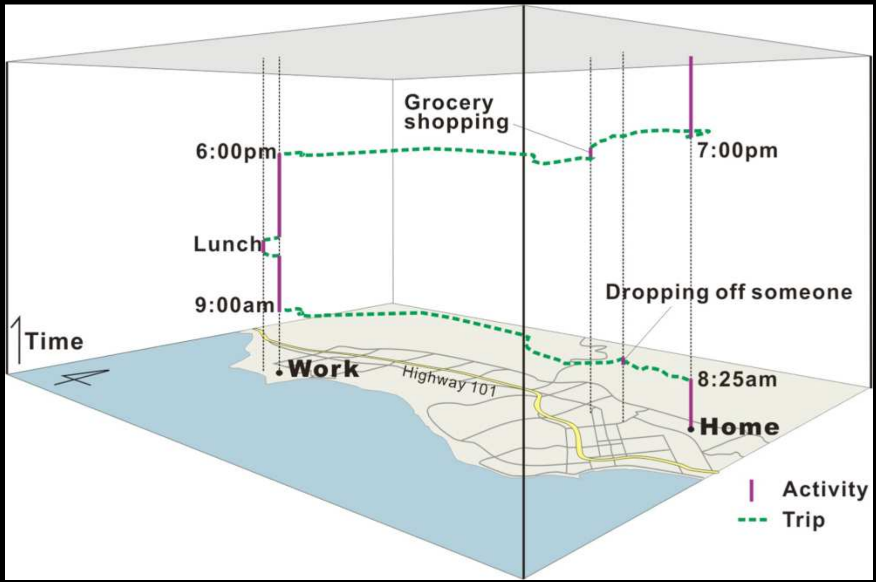


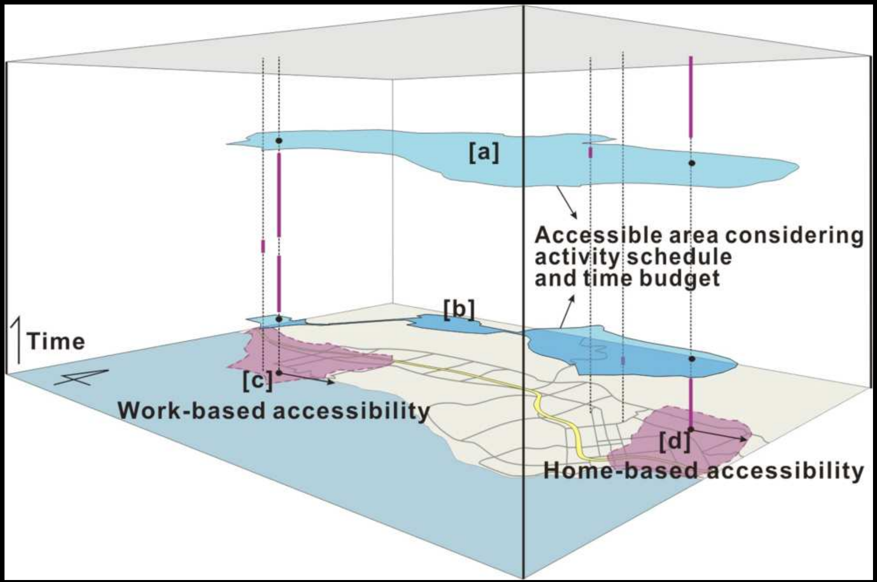
m

ea

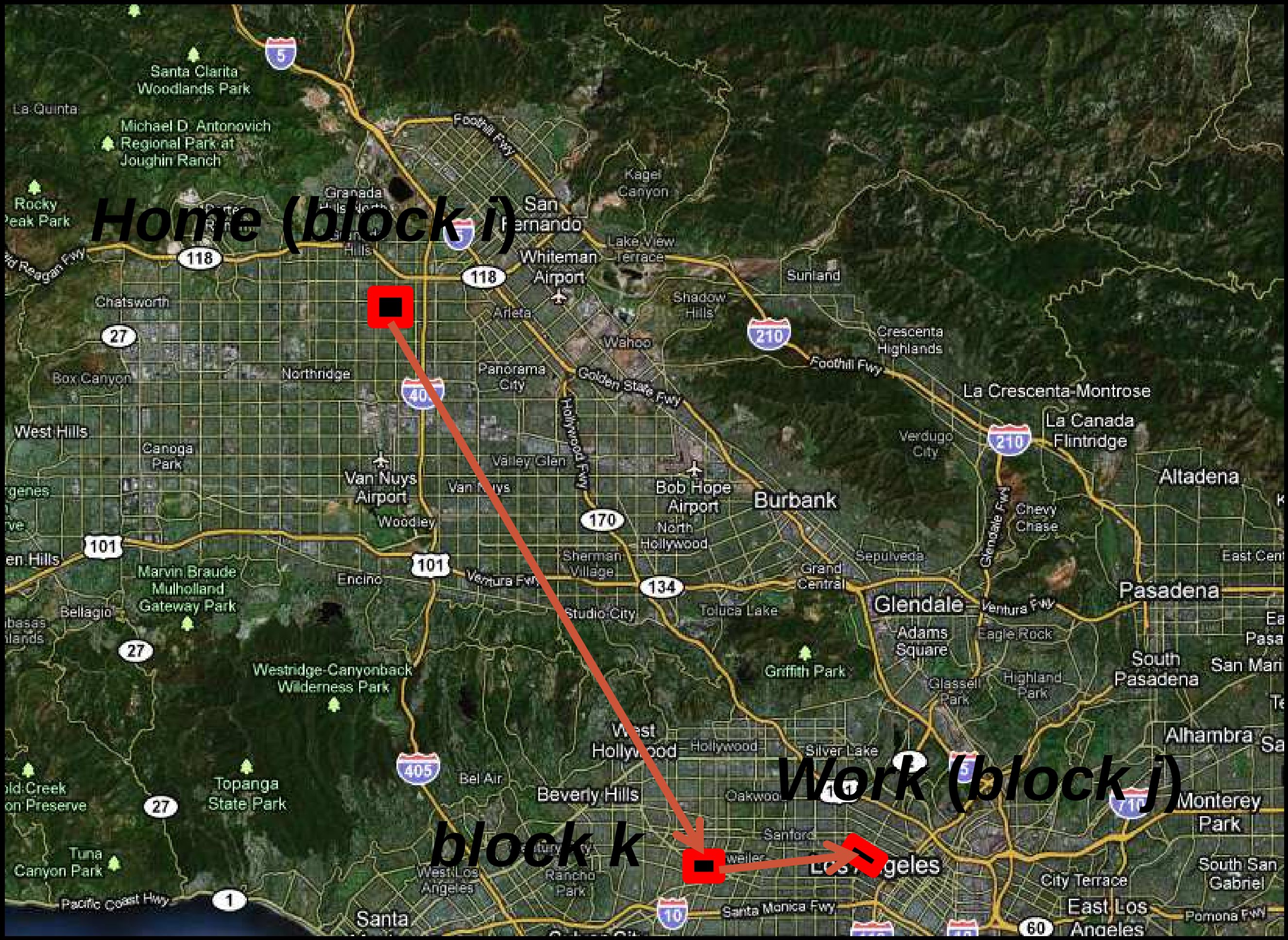
Hagerstrand (1970)

Activity scheduling





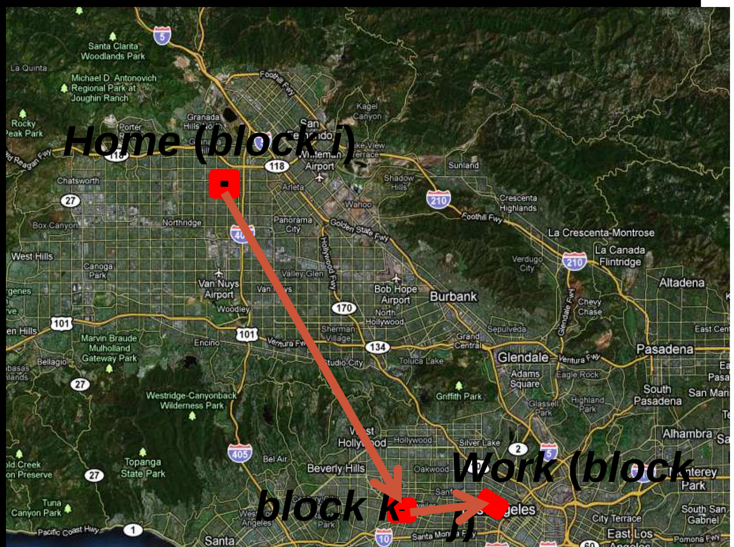
Computation of Potential Path Area

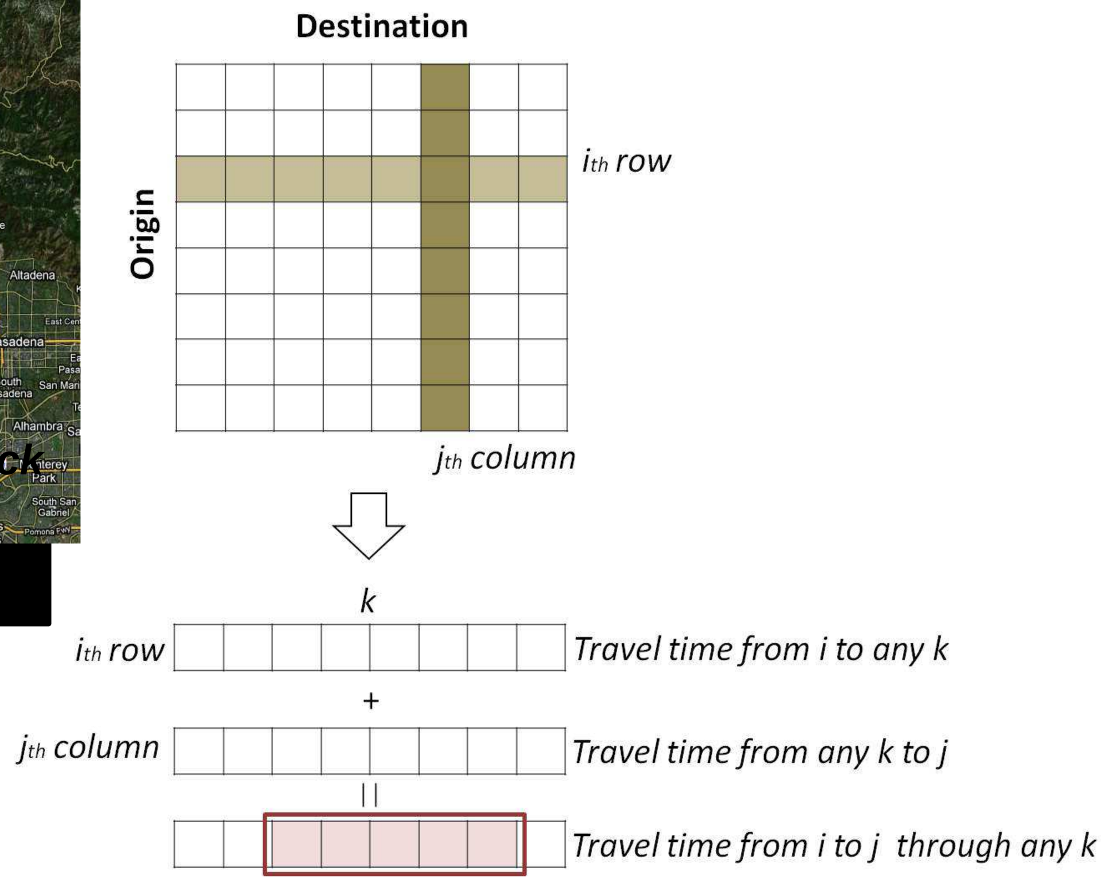


Travel time matrix

Origin	1	2	3	4	5	6	7	8	9	10	11	12
1	0.0	1.3	3.7	10.9	13.0	20.1	19.8	19.8	22.4	22.6	19.6	21.1
2	1.3	0.0	2.4	9.6	11.6	21.4	21.1	21.1	23.7	23.9	20.9	22.4
3	3.7	2.4	0.0	7.3	9.3	23.8	23.5	23.5	26.1	26.3	23.2	24.8
4	10.9	9.6	7.3	0.0	2.0	31.0	30.7	30.8	33.4	33.5	30.5	32.0
5	13.0	11.6	9.3	2.0	0.0	33.0	32.7	32.8	35.4	35.5	32.5	34.0
6	19.1	20.4	22.7	30.0	32.0	0.0	8.7	8.8	11.4	11.5	8.5	10.0
7	22.4	23.7	26.1	33.3	35.4	3.3	0.0	3.1	2.6	2.8	11.7	4.3
8	19.6	20.9	23.3	30.5	32.6	0.5	9.3	0.0	11.9	12.1	9.0	10.6
9	19.8	21.1	23.4	30.7	32.7	0.7	0.5	0.4	0.0	0.2	9.1	1.7
10	19.6	20.9	23.3	30.5	32.5	0.5	0.3	0.3	2.9	0.0	8.9	1.5
11	19.8	21.1	23.5	30.8	32.8	0.8	0.4	0.5	3.1	3.3	0.0	1.7
12	21.0	22.3	24.7	32.0	34.0	2.0	1.7	1.6	4.4	4.6	10.4	0.0

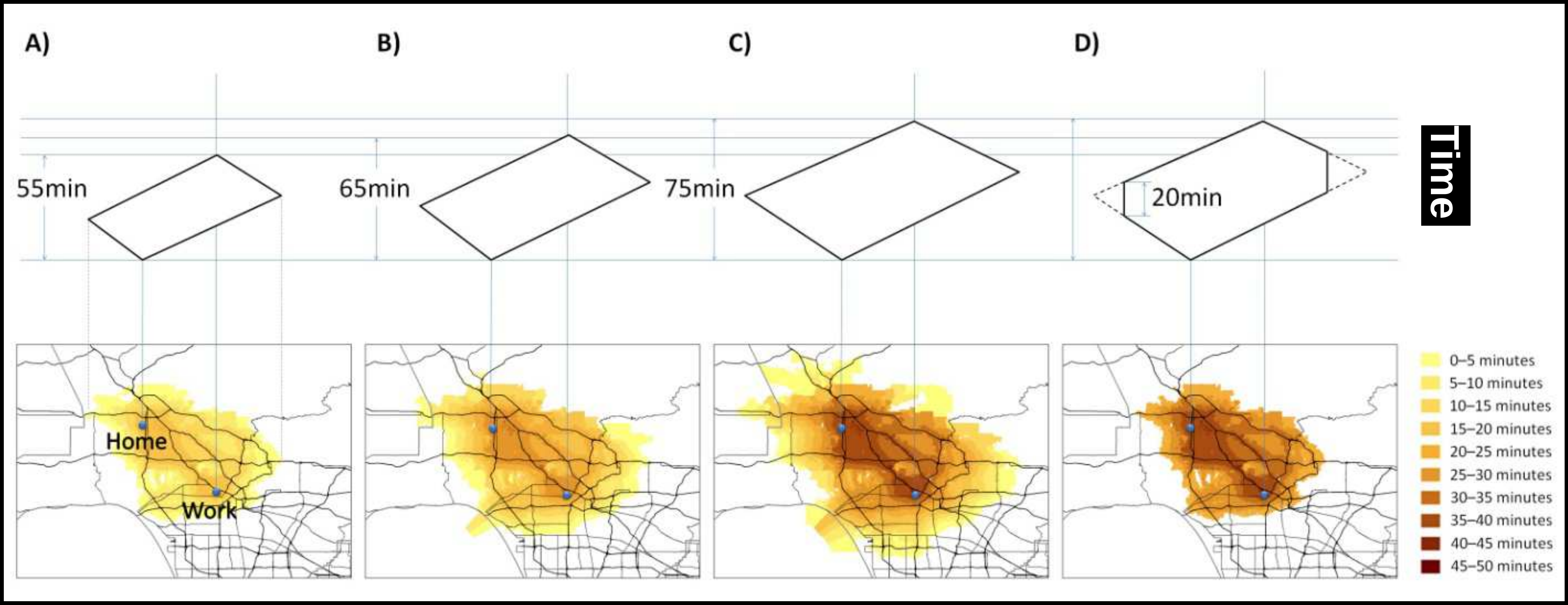
Computation of Potential Path Area



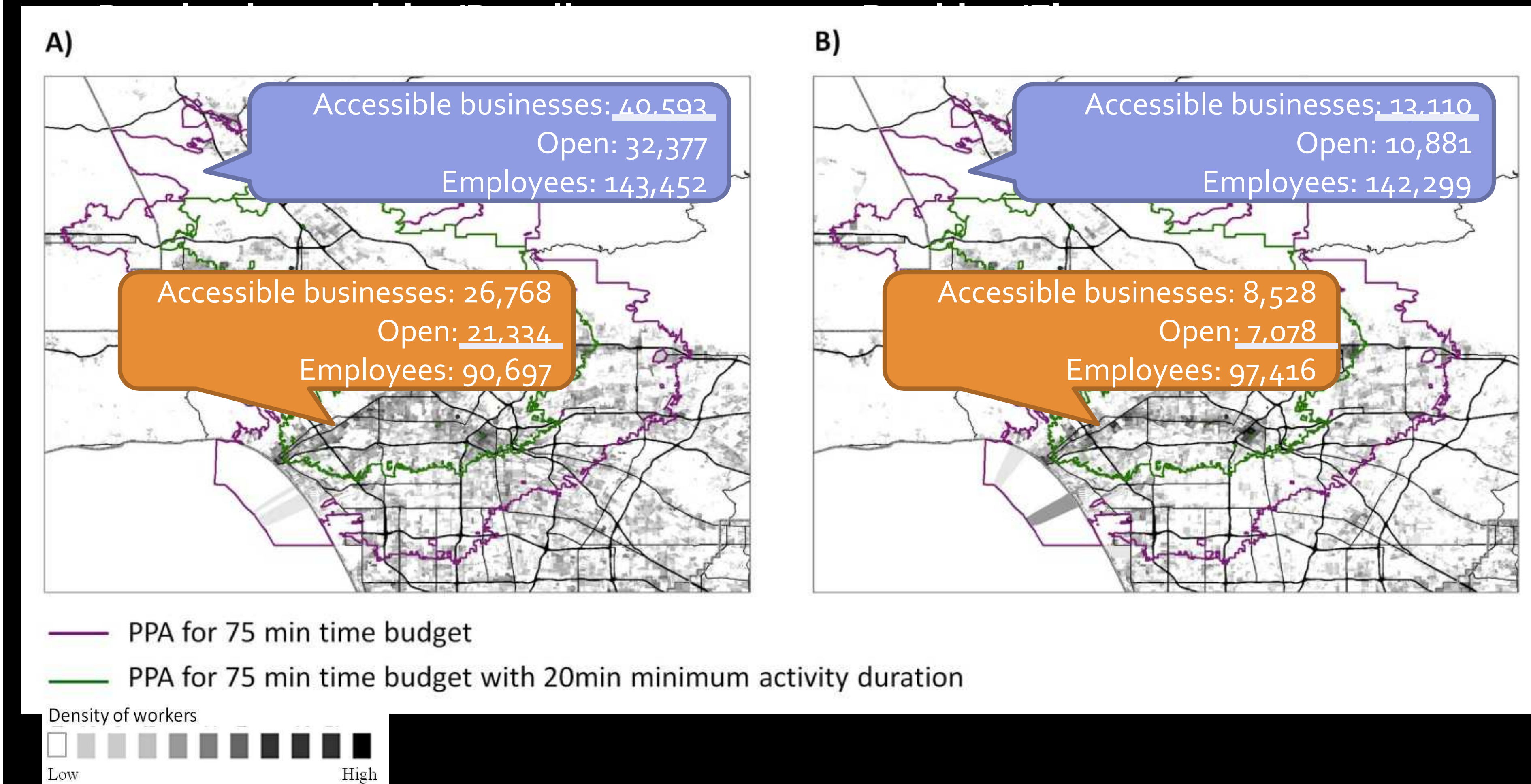


Blocks that are accessible within the time budget: PPA

Potential path areas for different time windows (AMpeak)

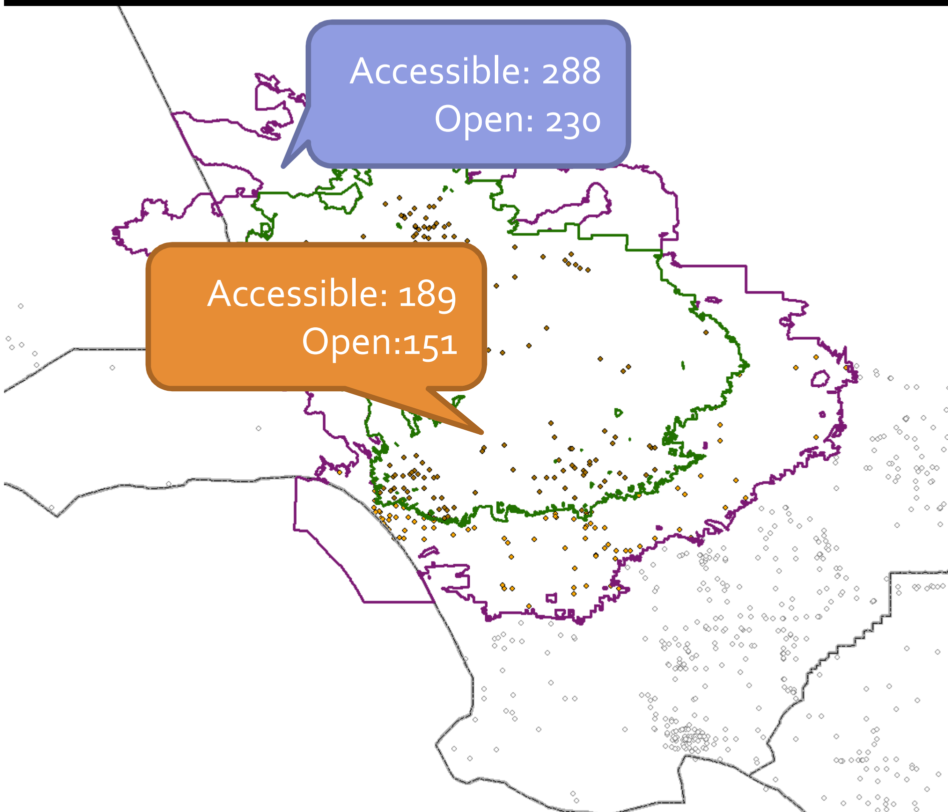


Available opportunities within PPA for different activity types (8am to 9am)

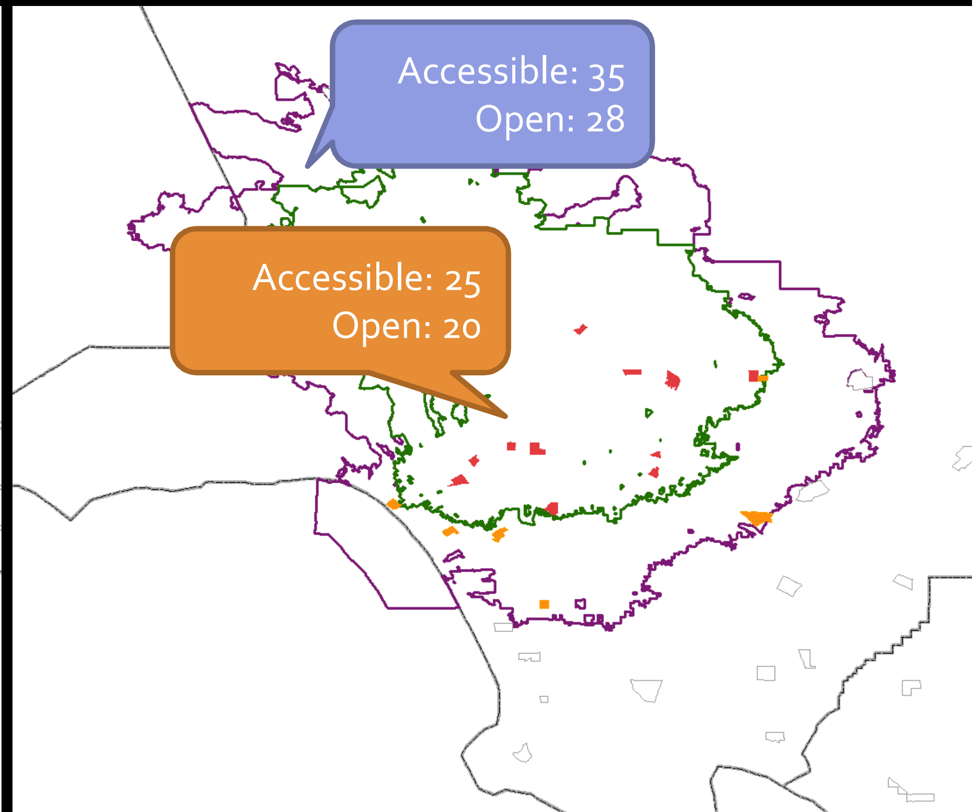


Different purchasing activity types

Grocery stores



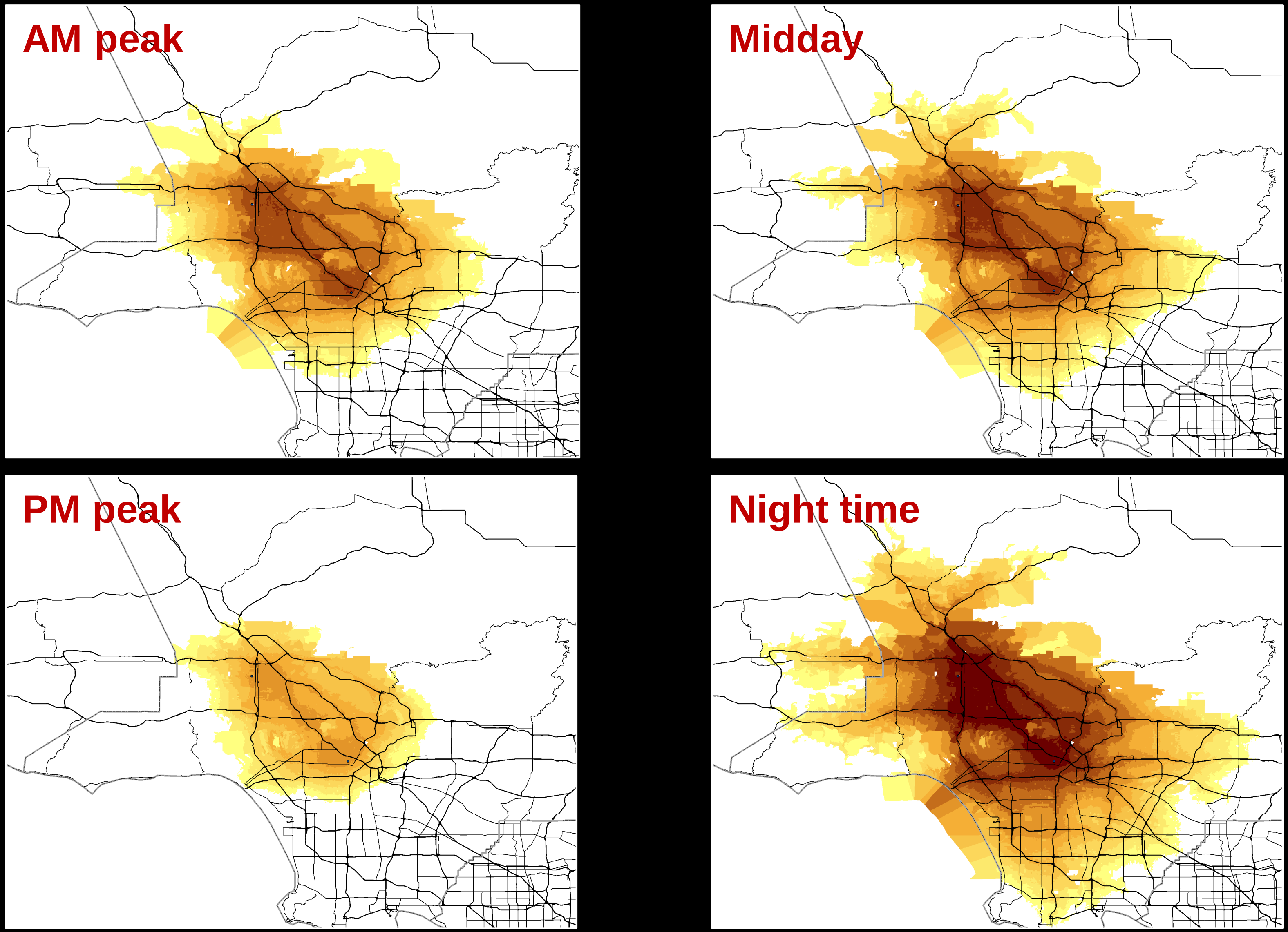
Regional shopping centers



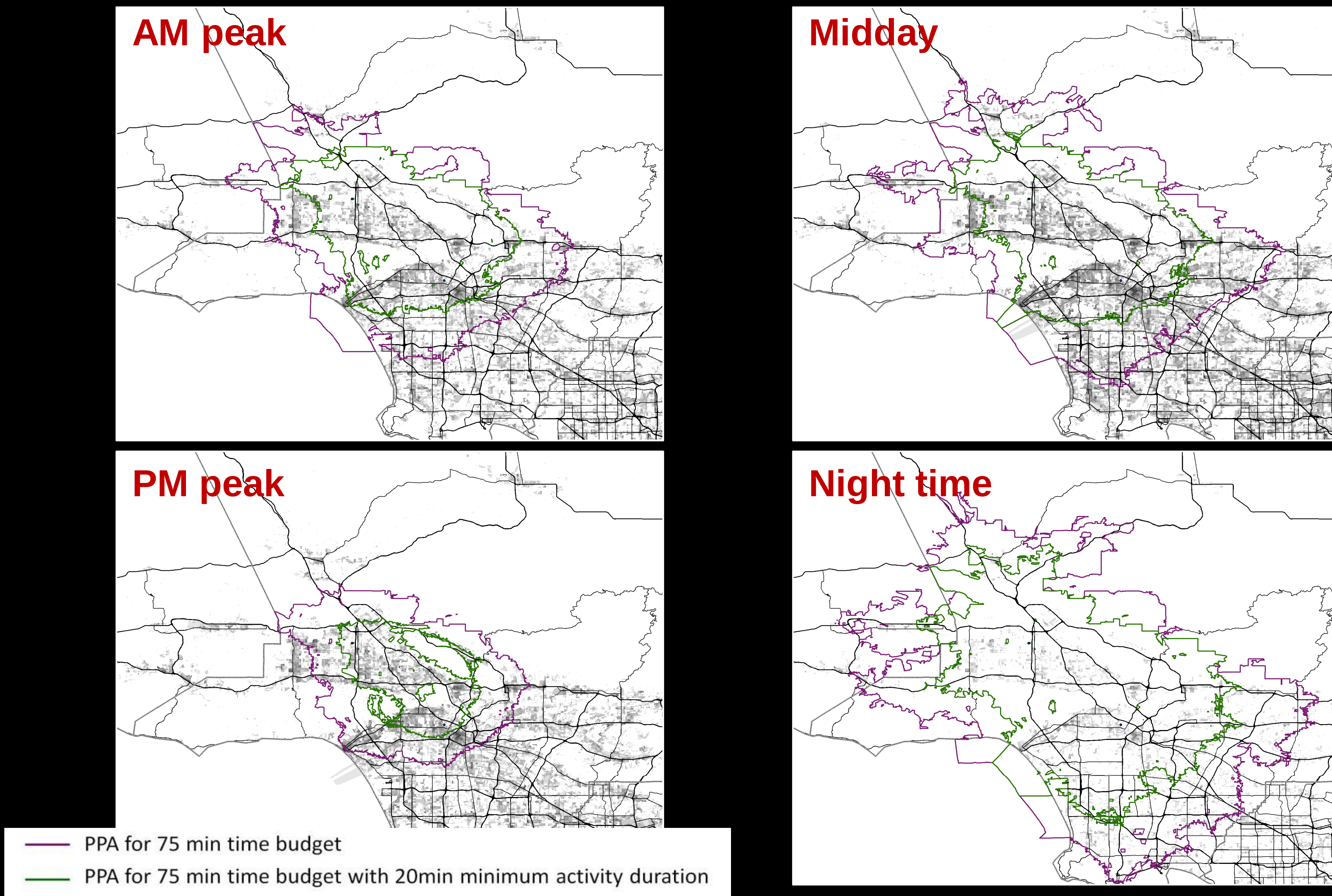
PPA for 75 min time budget

PPA for 75 min time budget with 20min minimum activity duration

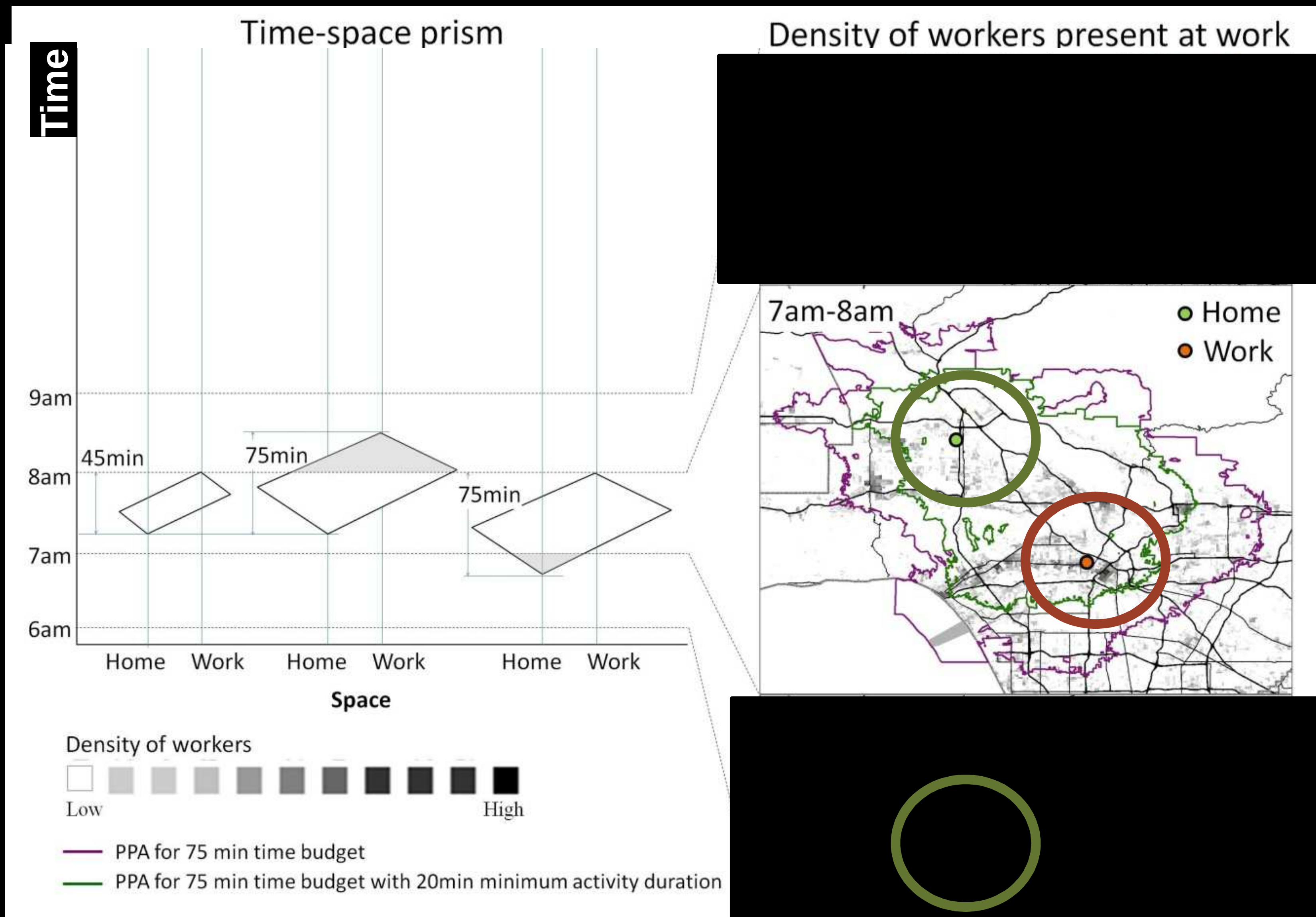
PPA by time of day

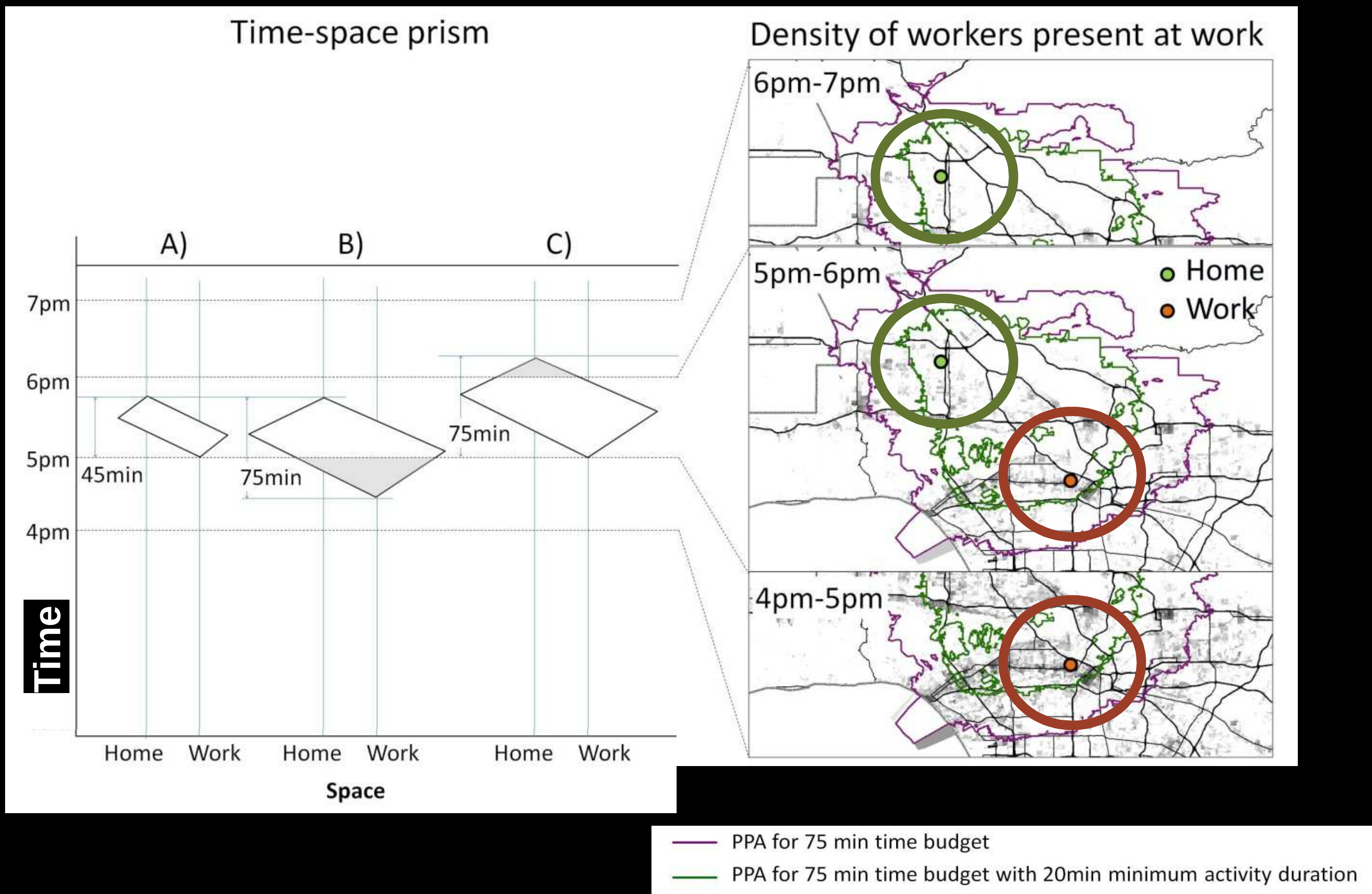


Available opportunities by time of day



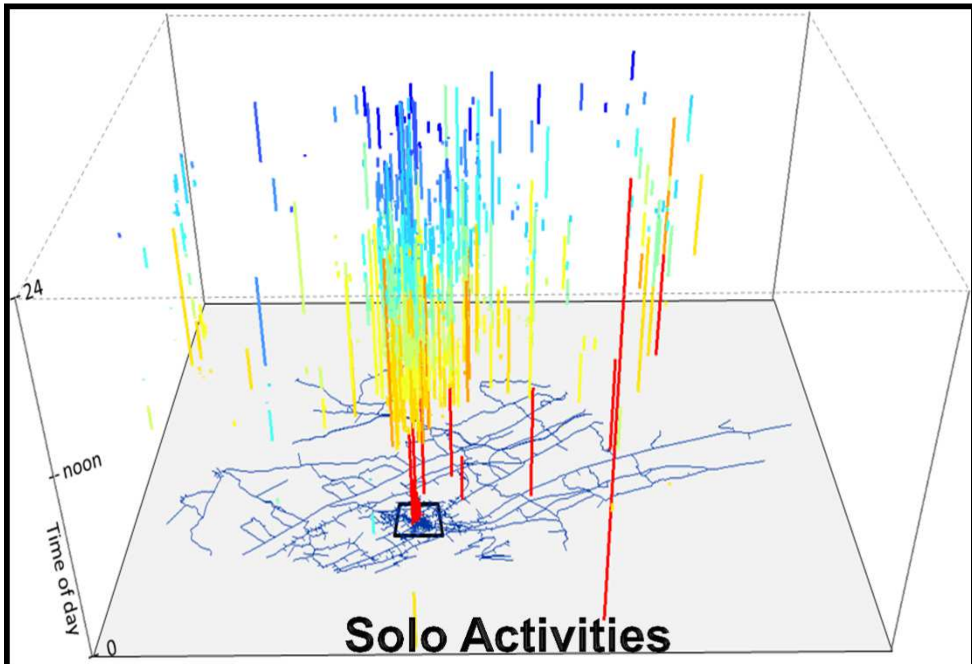
If have flexibility in arrival and departure time. Which value of accessibility matters to behavior (home, work, route or all)?

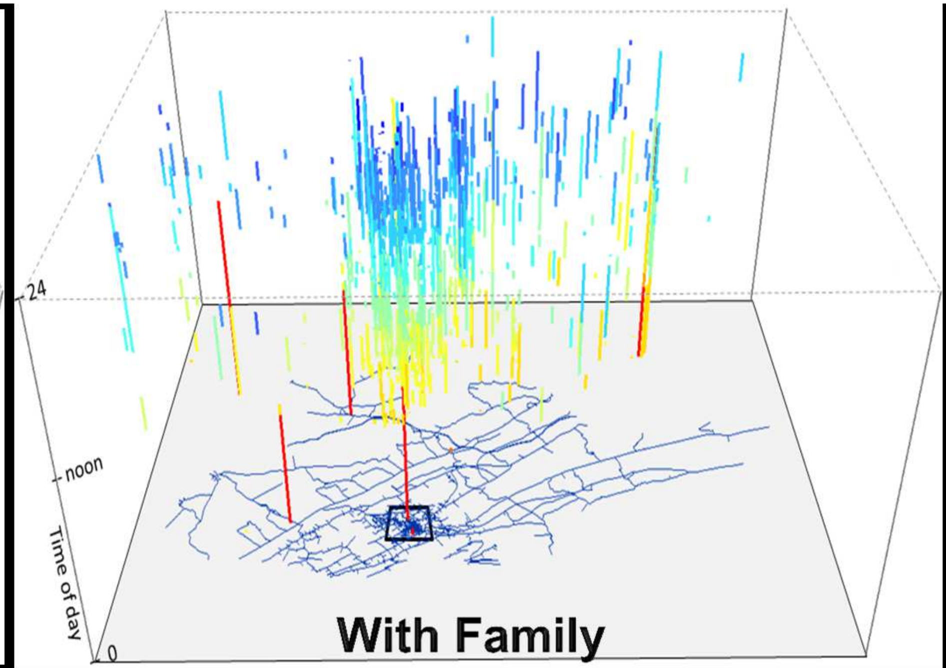




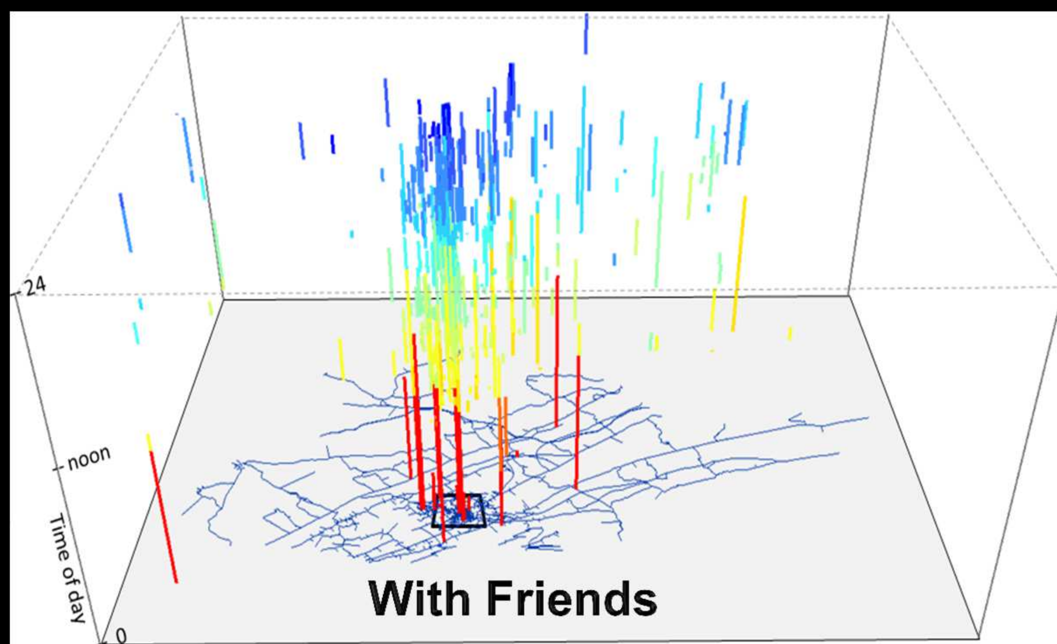
Time-space prism
Density of workers present at work
A)
B)
C)
7pm
6pm
5pm
4pm
45min
75min
75min
Time
Home
Work
Home
Work
Home
Work
Space
6pm-7pm
5pm-6pm
4pm-5pm
Home
Work
PPA for 75 min time budget
PPA for 75 min time budget with 20min minimum activity duration

Nature of Place Changes with Time of Day and With Whom we Interact


24
noon
Time of day
0
Solo Activities


24
noon
Time of day
0
With Family


Beginning of activity
22:00-23:59
12:00-13:59
00:00-1:59


24
noon
Time of day
0
With Friends

Some of the findings of this section

- Spatial distribution of alternatives has impacts on behavior
- Temporal dimensions of behavior also have impacts on
  - the spatial extent of choice sets and
  - the nature of alternatives that are distributed over space
- Flexibility (when, where, and how) makes some alternatives more attractive and some other less attractive
- These have to be accounted for in simulation models to achieve behavioral realism and to enhance model's ability to predict the impact of policies related to time (i.e., staggered arrival at work, flexible work hours)
- Further development
  - Intra-household interaction and life cycle stages (Yoon and Goulias, 2010) and
  - Subjective perception of space (mental maps) and attitudes

Integration Needs

- We need estimates of travel time on all links of the entire network – traffic assignment should give us all this information by some time interval and we do not know the right interval to use.
- Travel times need to be multi-modal (walk + bus) of any combination analyzed for policy – not a trivial task but feasible (see appendix)
- Availability of opportunities in this example is point-based (individual establishment) → model of open-closed hours of establishments?
- Knowledge (spatial cognition) and preferences/attitudes (sense of place) are not included and we know are important in delimiting spatial regions of considered opportunities → integrate with cognitive filters?

Computation

- Initial versions take days to run on 40 core servers
- Second iteration of newer versions take hours to run
- BUT: slow improvement in handling input-output from different components-still need human intervention to check databases and verify interfaces work properly
- ONE core application designed to integrate modules is needed now!

Acknowledgement

GeoTrans lab.

University of California Transportation Center

United States Department of Transportation Eisenhower Fellowship program

University of California Office of the President UC Lab Fees Program

University of California Office of the President Multicampus Research Program Initiative Sustainable Transportation

Southern California Association of Government

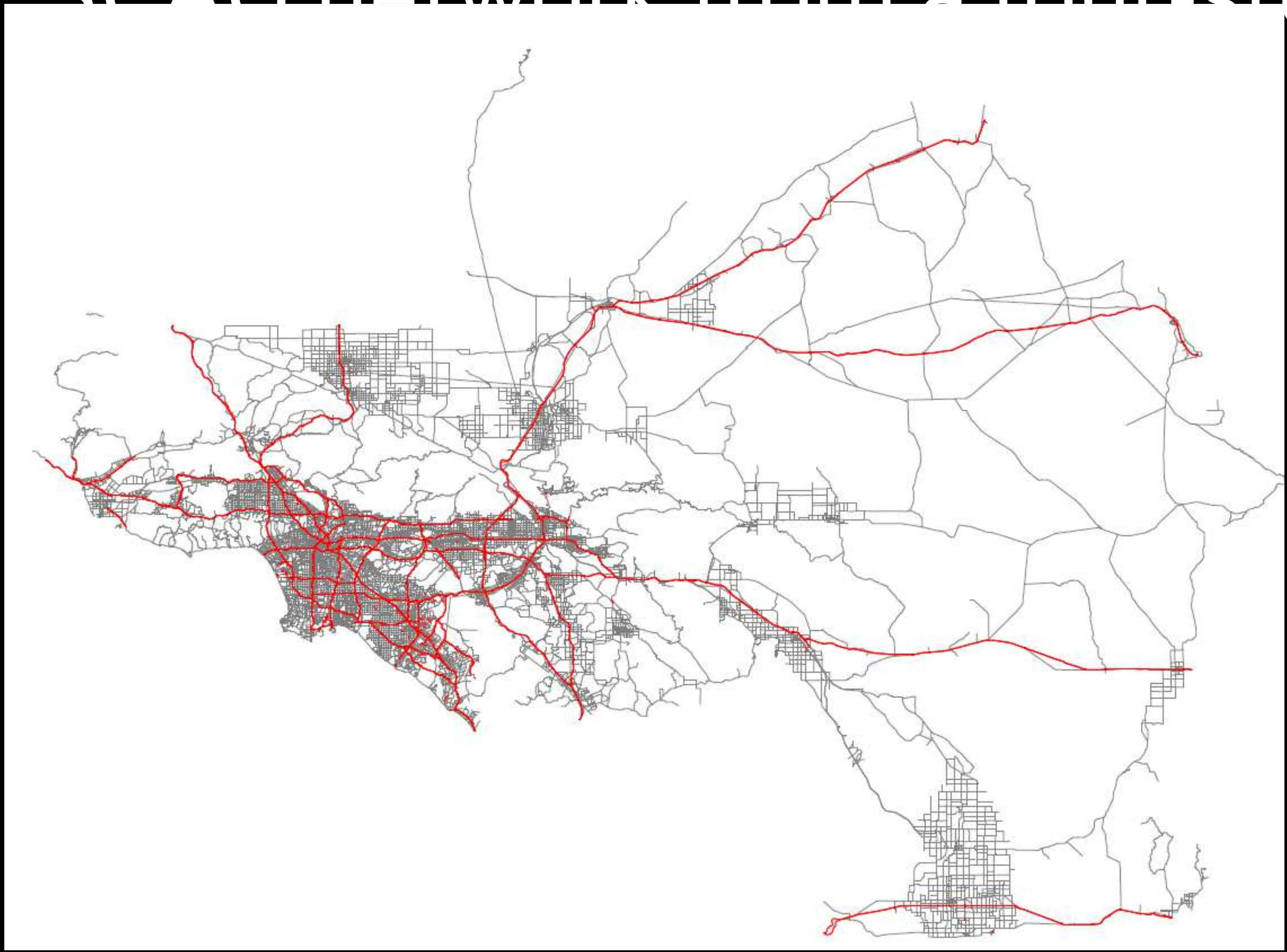
Thank You to The Contributors:

Dr. Yali Chen, Dr. Seo Youn Yoon, Dr. Ting Lei,
Kathleen Deutsch, and Srinath Ravulaparthy

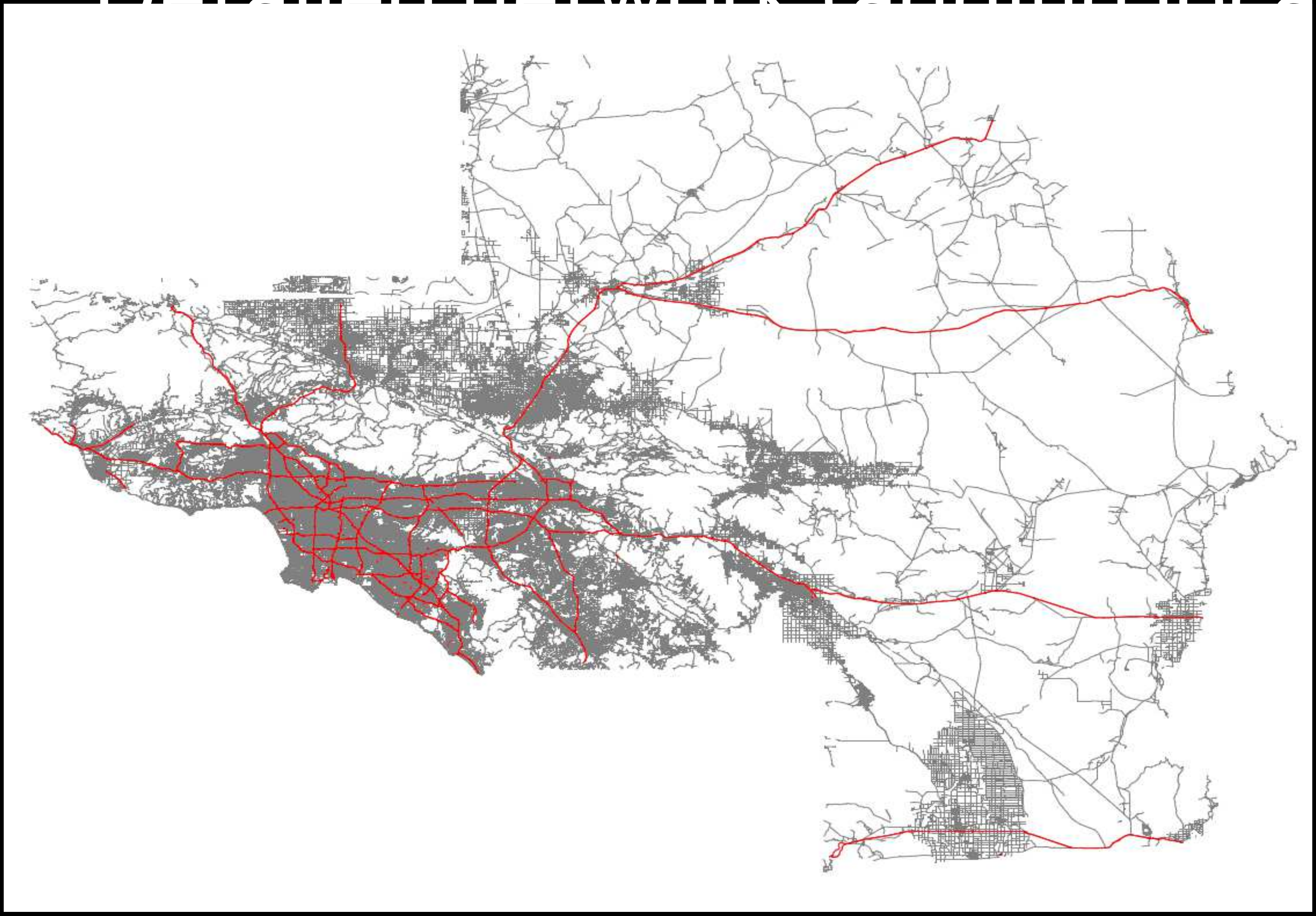
Base map of opportunities

FROM BLOCK GROUPS TO BLOCKS & ADDING NETWORK DETAIL

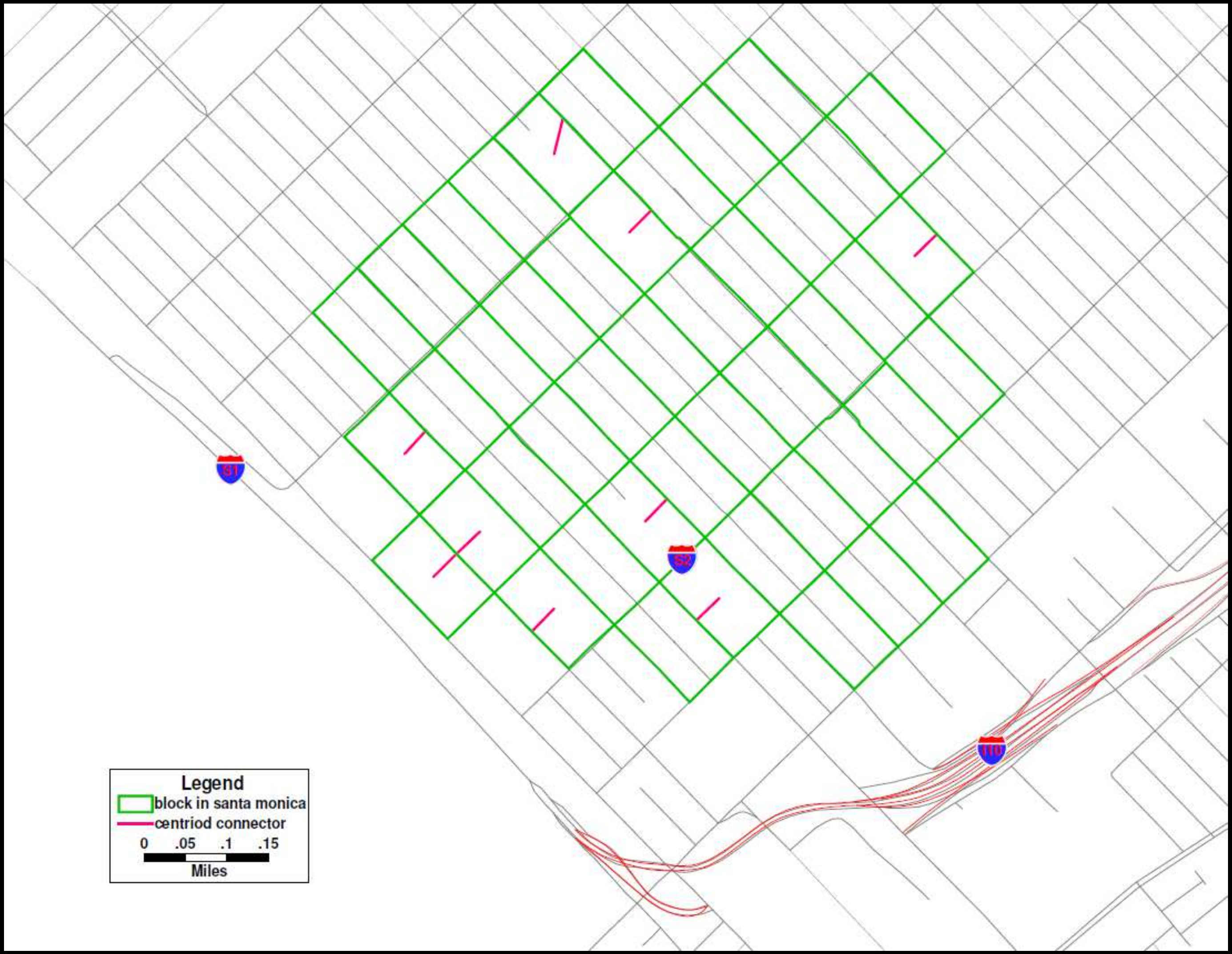
SCAG network from a four-step



Detailed network (adding local



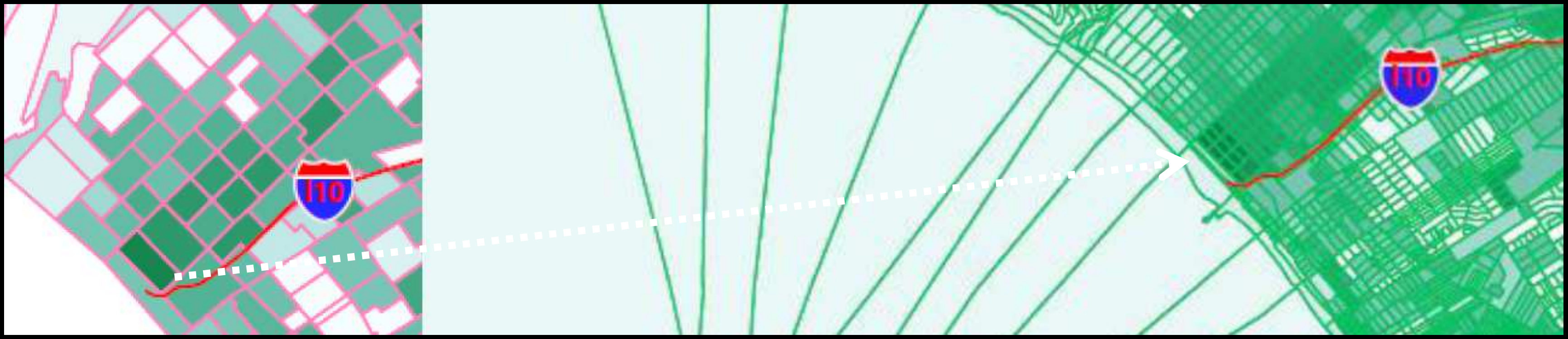
Block Level Centroid Connectors



Allocation from Block groups to Blocks

- Function of land area, population, amount of: freeways, arterials, collectors, ramps (negative and positive influence depends on industry type).

- Tested many different regression models and decided on Poisson regression (count data regression models)
- Kept the total equal to the block group for each block group.



Block Groups

Blocks

Using these regression models (one for each industry type), the block group employment for each industry type is distributed into each block within a block group as shown below:

a. Calculate the estimated block employment ($\hat{E}_i$) for block i by applying the block characteristics and the estimated regression model
b. Calculate the percentage of employment (P_i) for each block i within each block group j.

$$P_i = \frac{\hat{E}_i}{\sum_{i \in B_j} \hat{E}_i}, \text{ where } B_j \text{ is the set of all blocks within block group } j.$$

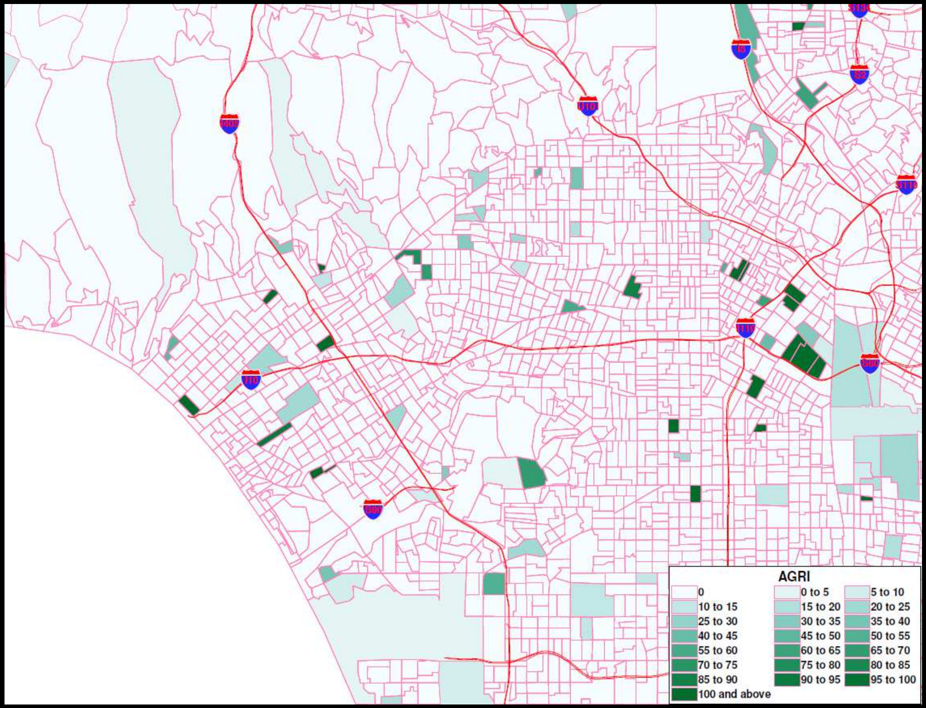
c. Calculate the block employment (E_i) by multiplying the percentage with the block group employment

$$E_i = P_i \times (\text{employment in block group } j \text{ that block } i \text{ belongs to}).$$

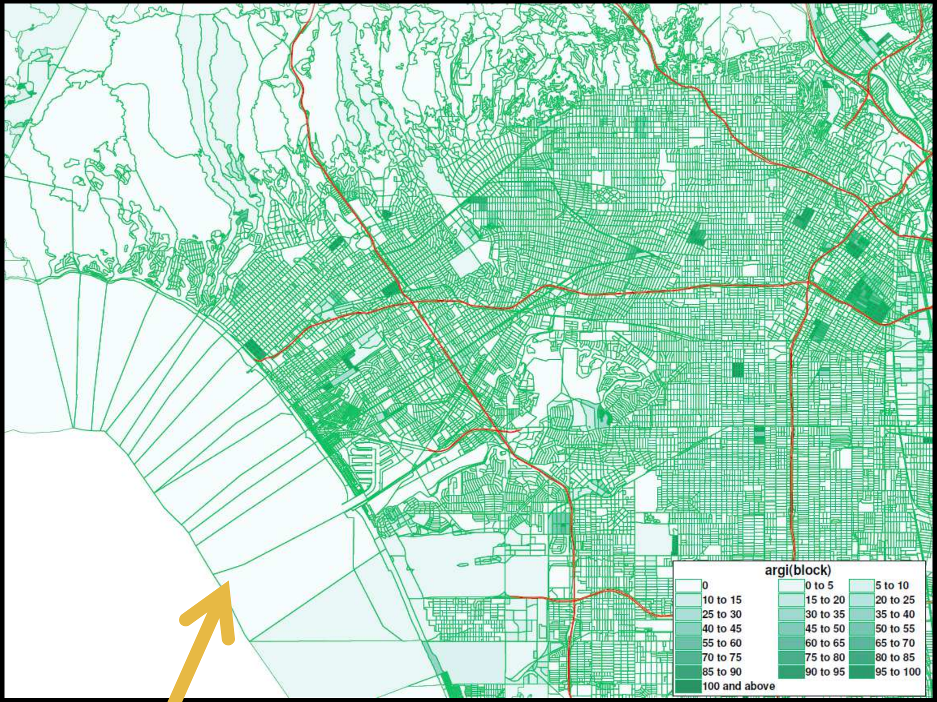
The same process is repeated for j = 10,631 block groups and 15 industry types.

Agriculture Density

Block group level(observed)



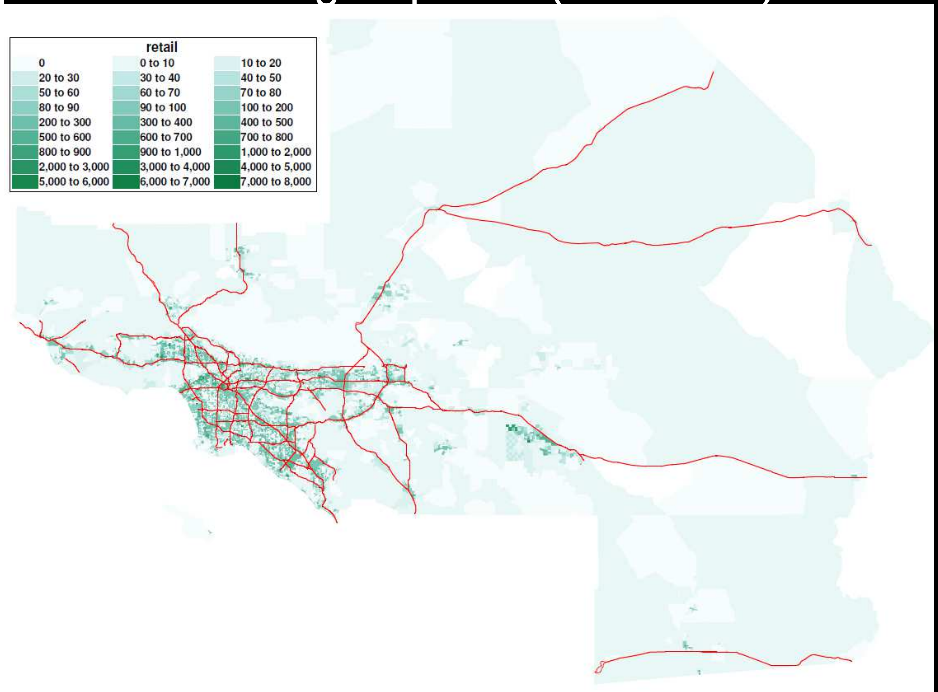
Block level(predicted from model)



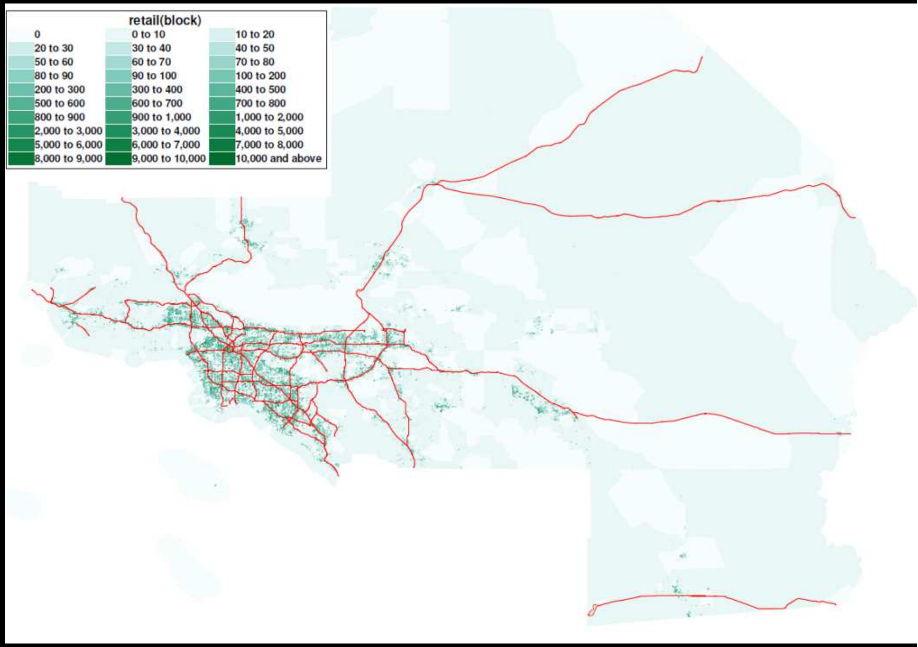
Blocks in the Ocean – special treatment

Retail Density

Block group level(observed)

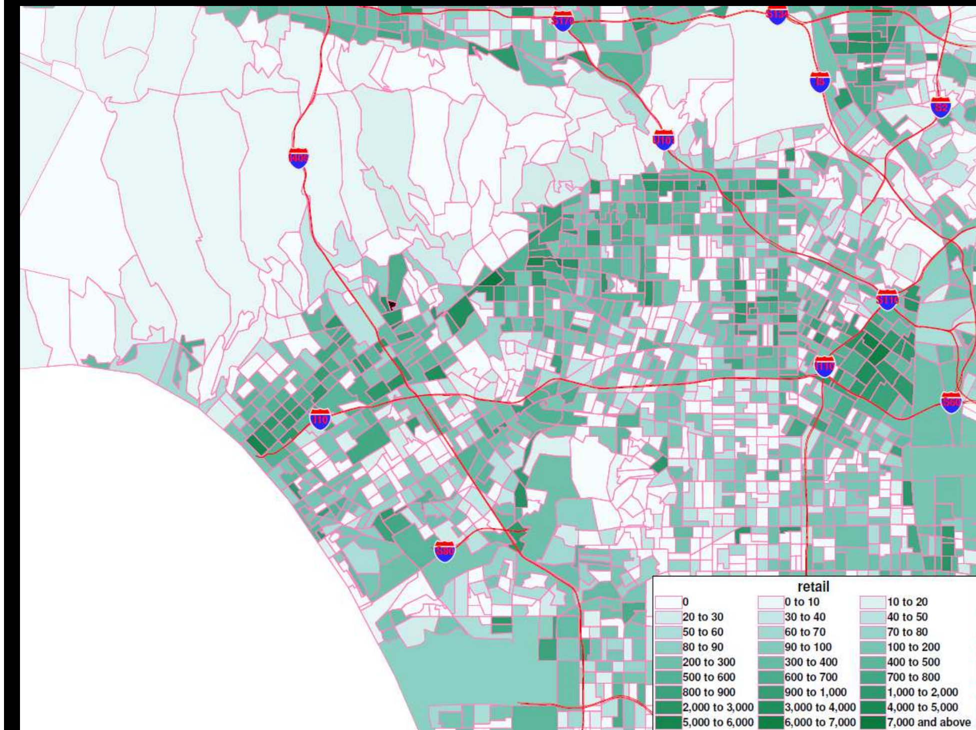


Block level(predicted allocation model)

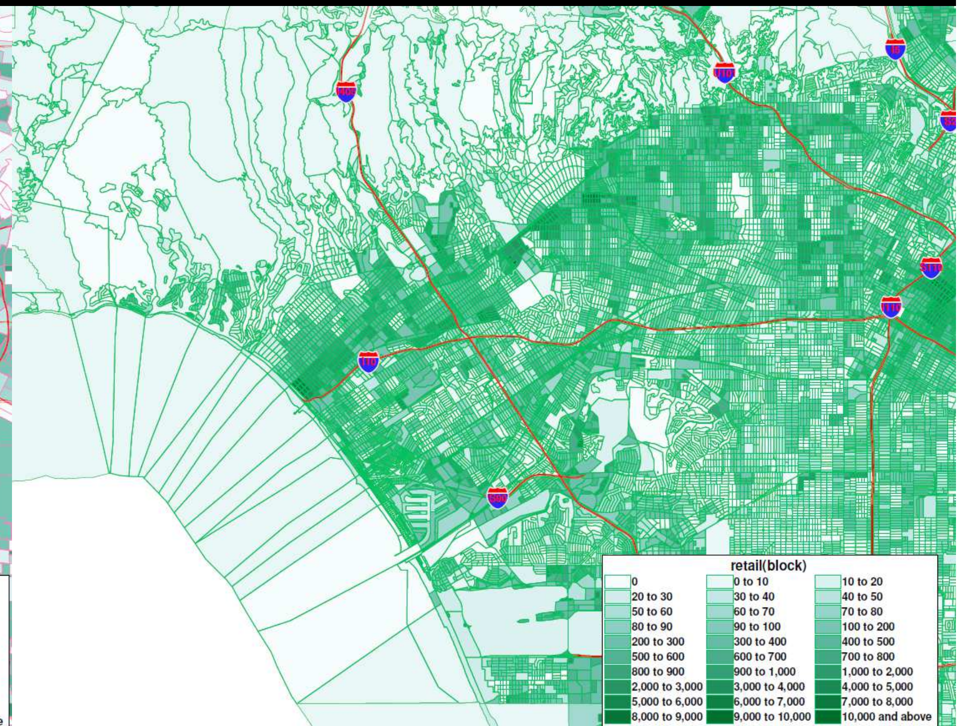


Retail Density

Block group level(observed)

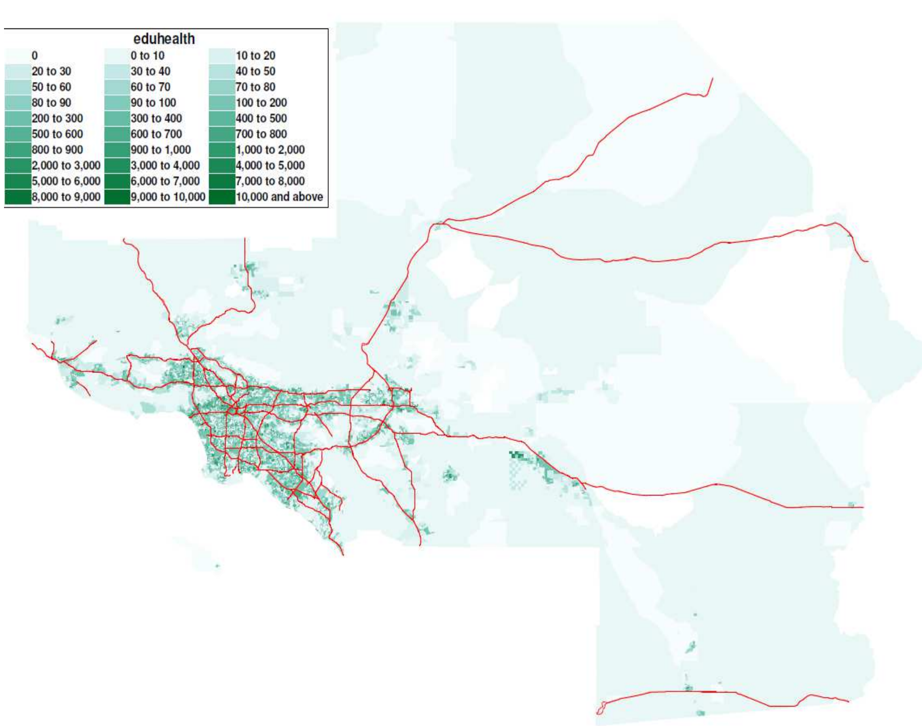


Block level(predicted from model)

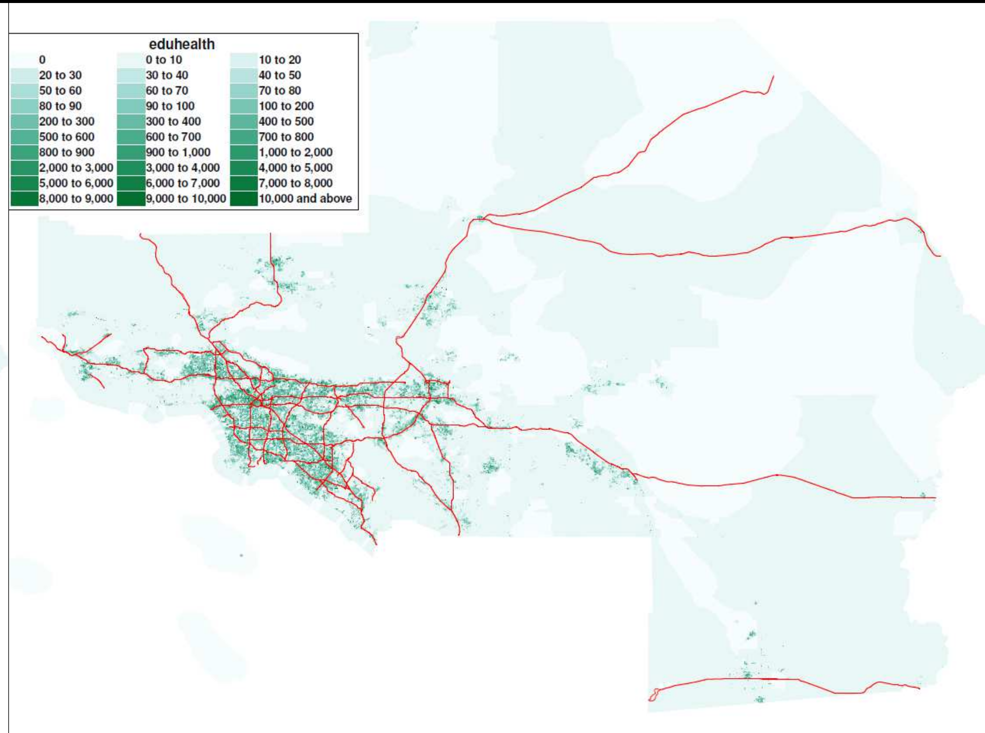


Education & Health Density

Block group level(observed)



Block level(predicted from model)

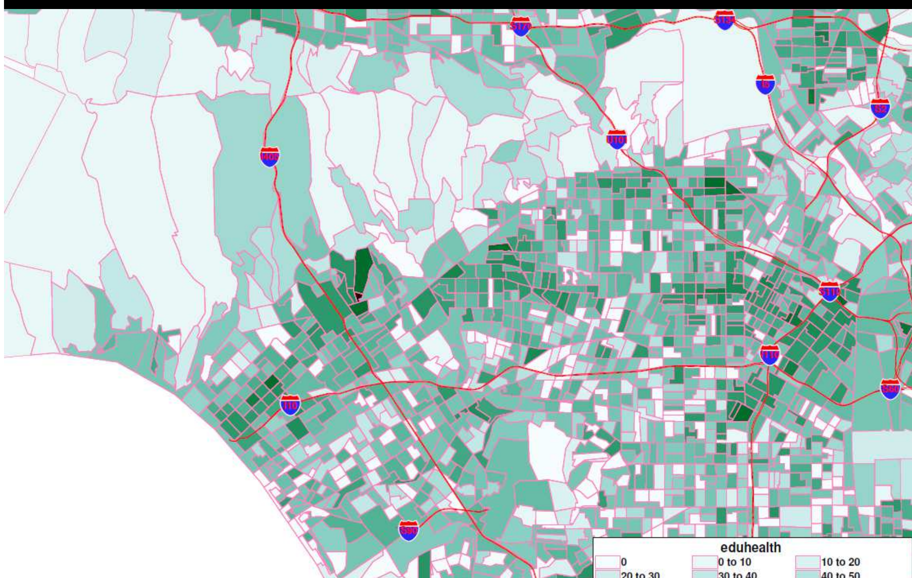


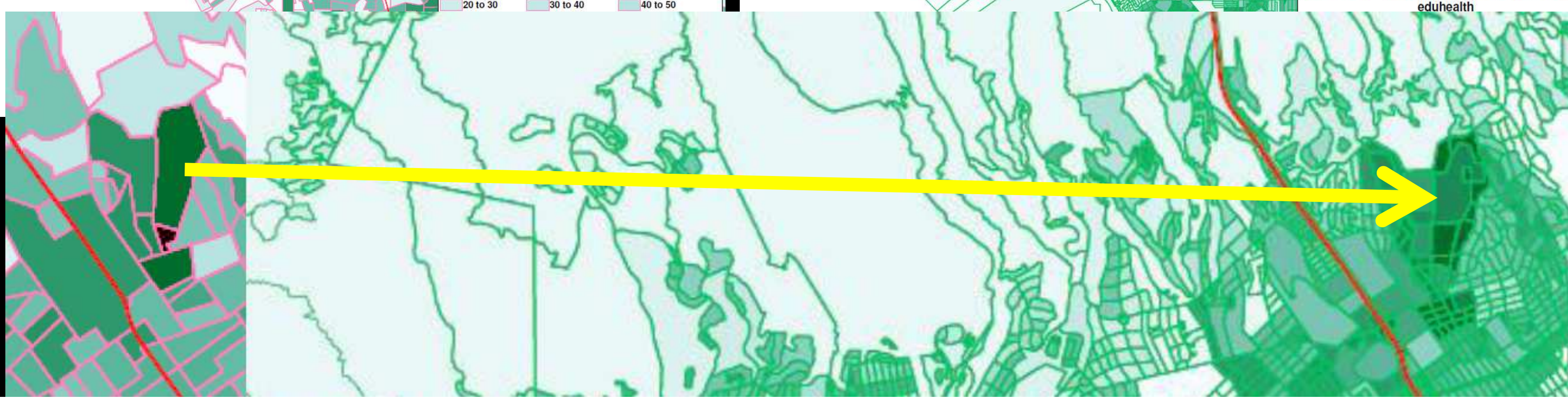
We split this in two separate sets of indicators

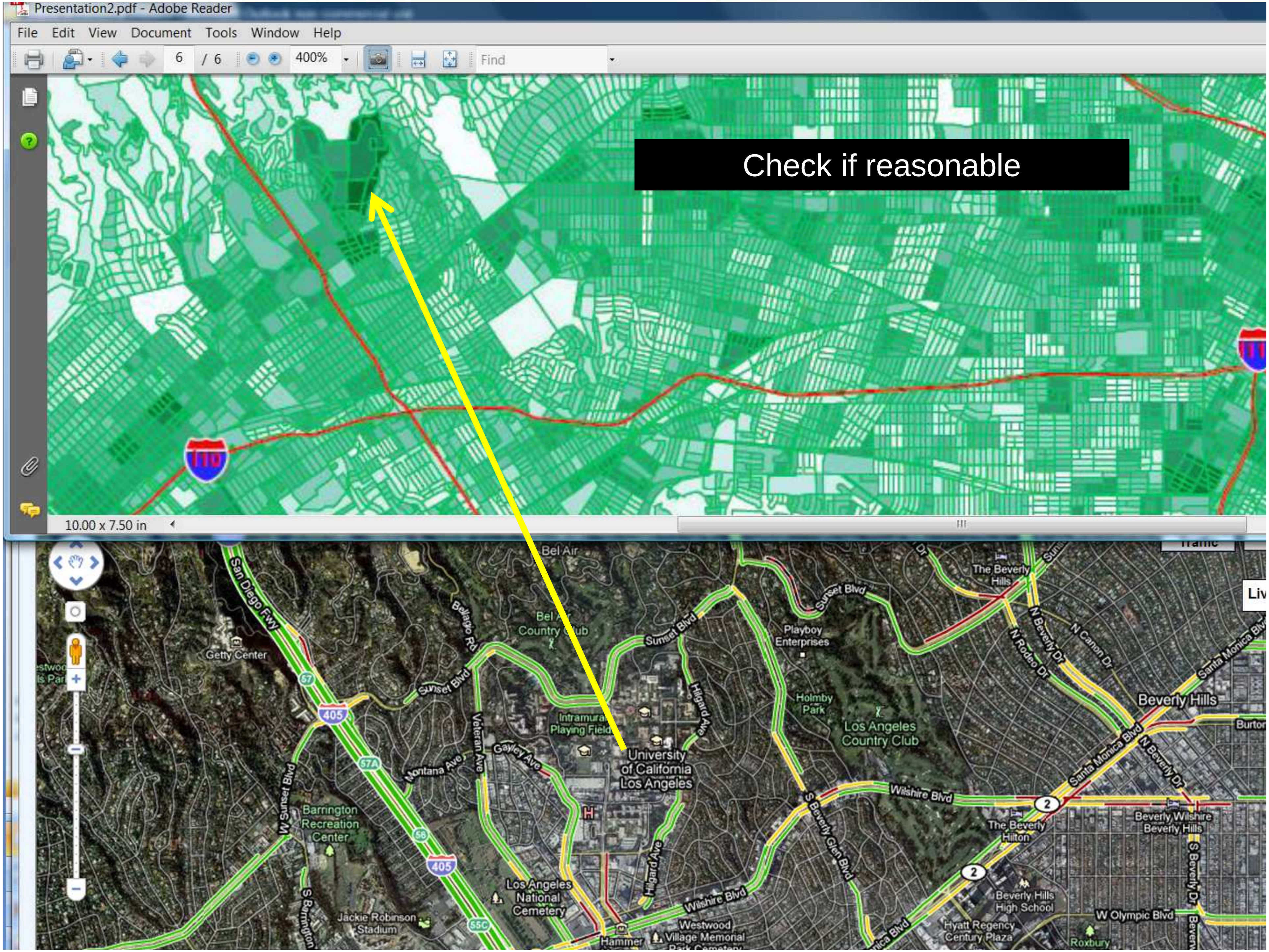
Education & Health Density

Block group level(observed)

Block level(predicted from model)






Presentation2.pdf - Adobe Reader
File Edit View Document Tools Window Help
6 / 6
400%
Find
Check if reasonable
10.00 x 7.50 in
Bel Air
Bel Air Country Club
Getty Center
San Diego Fwy
Bellagio Rd
Sunset Blvd
The Beverly Hills
Playboy Enterprises
N Beverly Dr
N Rodeo Dr
N Canon Dr
Santa Monica Blvd
Beverly Hills
Burton
Hilgard Ave
Holmby Park
Los Angeles Country Club
Intramural Playing Field
Veteran Ave
Gayley Ave
Montana Ave
University of California Los Angeles
Wilshire Blvd
W Sunset Blvd
Barrington Recreation Center
S Beverly Glen Blvd
The Beverly Hilton
Beverly Wilshire Beverly Hills
Los Angeles National Cemetery
Jackie Robinson Stadium
Westwood Village Memorial Park Cemetery
Beverly Hills High School
Hyatt Regency Century Plaza
W Olympic Blvd
S Beverly Dr
Roxbury
Hammer
S Barrington

Issues Found and Resolved

- Block groups with employees but no businesses
- Block groups with businesses but no employees
- Disagreements between D&B number of persons by industry and CTPP
- Differences in NAICS-SIC classifications
- CTPP reports education and health as one industry (split in two)

- You can find additional details at: Chen, Y., S. Ravulaparthy, K. Deutsch, P. Dalal, S.Y. Yoon, T. Lei, K.G. Goulias, R.M. Pendyala, C.R. Bhat, and H-H. Hu (2011) Development of Opportunity-based Accessibility Indicators. *Transportation Research Record: Journal of the Transportation Research Board, No. 2255,* Transportation Research Board of the National Academies, Washington D.C., 2011, *pp.58-68.*

AVAILABILITY OF OPPORTUNITIES BY TIME OF DAY

Time of Day Components

- Travel time on the network -> use travel speed from calibrated four step model that has four wide time periods (AM Peak - 6AM to 9AM, Midday - 9AM to 3PM, PM Peak - 3PM to 7PM, Night - 7PM to 6AM)
- Opening-closing times for businesses -> use arrival and departure time of workers in post Census survey (16k hhs)
- To account for the "wide" periods of the network times we use MIN and MAX for the number of employees by industry that can be reached within a temporal buffer of 10, 20, and 50 minutes
- The buffers 10 vs 50 resemble local vs regional accessibility
- Shortest path computed using aprox. 200,000 by 200,000 origin destination matrix using TRANSCAD.

Employee Classification

1. Wholesale Trade
2. Retail Trade
3. Administrative and Support Services
4. Educational Services
5. Healthcare and Social Assistance
6. Arts, Entertainment and Recreation
7. Accommodations and Food Services
8. Public Administration
9. Agriculture/Forestry/Fishing
10. Construction
11. Manufacturing – Durable and Nondurable goods
12. Transportation and Warehousing
13. Information
14. Finance or Insurance
15. Professional, Scientific and Technical Services

Using a survey of workers for each respondent q that works, we have the arrival time at work (say t_{qa}) and departure time from work (say t_{qd}).

Define a dummy variable δ_{qgh} that takes a value of one if individual q works in industry type g and county h, and zero otherwise.

Consider a specific hour of the day k, and let T_s be the start time of this hour k, and let T_e be the end time of this hour k.

Define another dummy variable α_{qk} that takes a value 1 if $t_{qs} \geq T_s$ and $t_{qe} \leq T_e$, and zero otherwise.

Intuitively, α_{qk} equals one if the employee q is at the work place at the hour k and zero otherwise.

From the travel survey, we also have a weight W_q for person q, which expands the individual so that the expanded sample across all surveyed individuals is representative of the population.

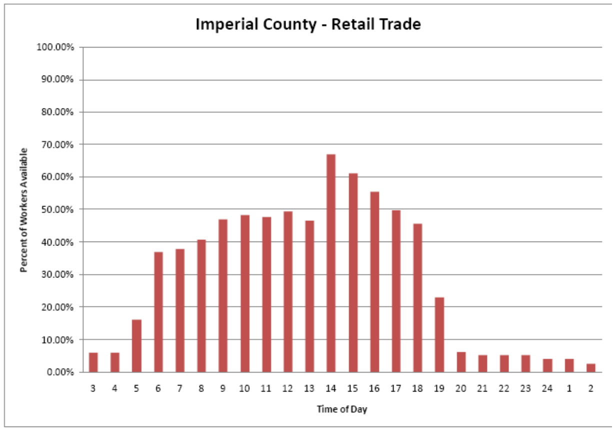
An estimate of the number of reachable employees in industry type g at county h at time k (A_{ghk}) is as follows:

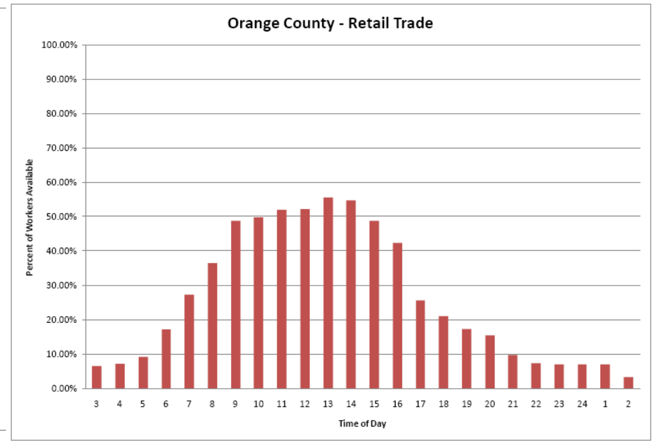
$$A_{ghk} = \sum_{q} \delta_{qgh} \alpha_{qk} W_q$$

The total number of reachable workers in industry *g* at county *h* at time *k* can be expressed as a function of the percentage of reachable workers in industry *g* and county *h*:

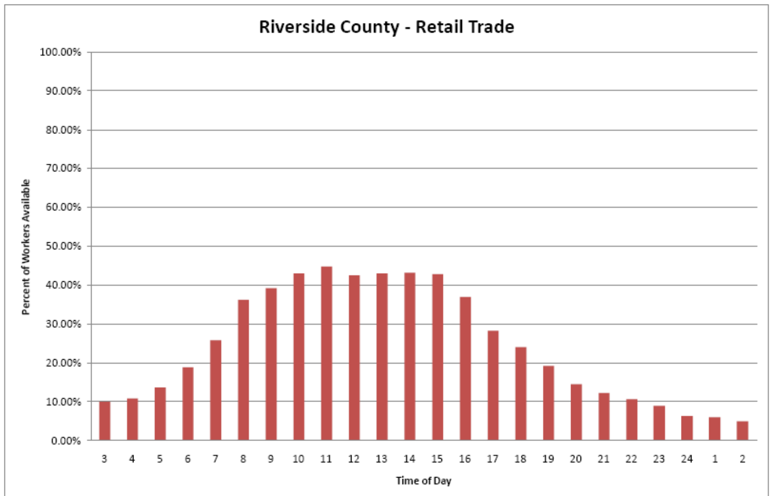
$$P_{ghk} = \frac{A_{qhk}}{\sum_{q} \delta_{qgh} W_{q}}.$$

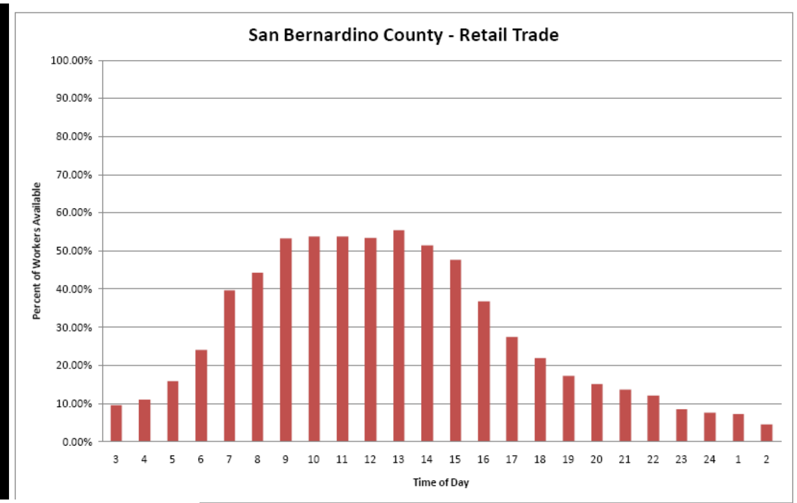
The next slide shows examples of the reachable percent of workers in Los Angeles County by time of day.


Imperial County - Retail Trade
Percent of Workers Available
Time of Day

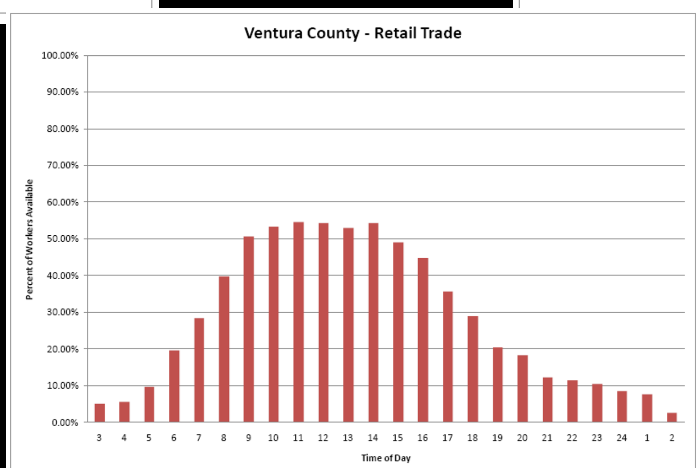

Orange County - Retail Trade
Percent of Workers Available
Time of Day


Los Angeles County - Retail Trade
Percent of Workers Available
Time of Day


Riverside County - Retail Trade
Percent of Workers Available
Time of Day


San Bernardino County - Retail Trade
Percent of Workers Available
Time of Day

Pghk

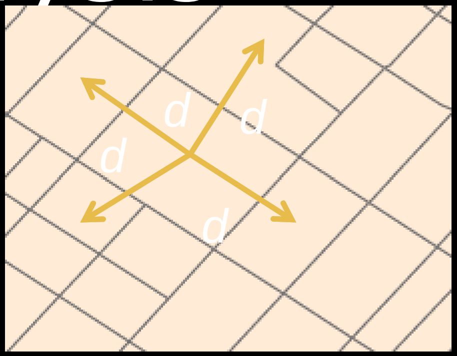

Ventura County - Retail Trade
Percent of Workers Available
Time of Day

Identification of Activity Opportunity Clusters

Based on: Ravulaparthy S., P. Dalal, Y. Chen, and K. G. Goulias (2011) An Exploratory Analysis of Spatial Hierarchical Clustering Using Opportunity-based Accessibility and Reported Land Values and its Relationship with Residents Characteristics in Los Angeles County. Paper 12-1297 presented at the *91st Annual Meeting of the Transportation Research Board, Washington, D.C.,* January 22-26, 2012.

Step 2: G* Analysis

- Spatial clusters analysis
  - Spatial dependence
  - Spatial heterogeneity
- Getis-Ord *G*(d)* statistic is a local measure of spatial dependence
  - Measures spatial association between any given point in space to all other neighboring points
  - Defined by a distance of *d*
  - Neighborhood blocks sharing a boundary
- Getis-Ord G* statistic calculates the z-score
  - Statistical significance of clustering
- z-score significance
  - Higher or lower values indicates high intensity of clustering
  - Near zero indicates no clustering in study area

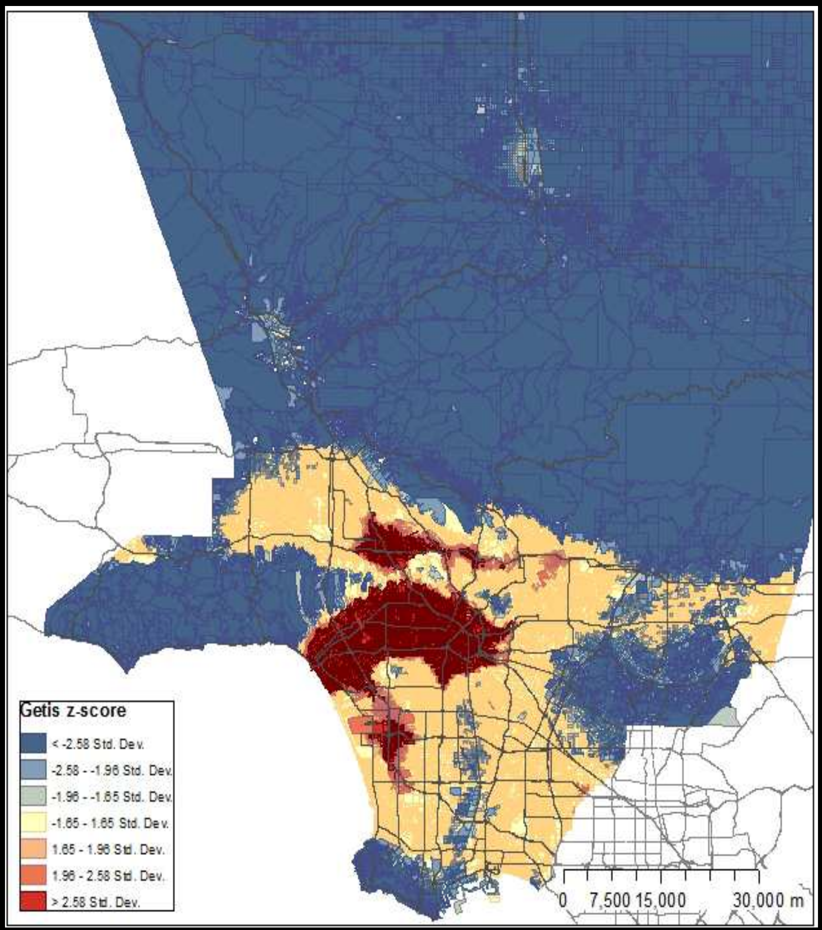


Jacquez (2008) , Getis and Ord (1992) and www.esri.com

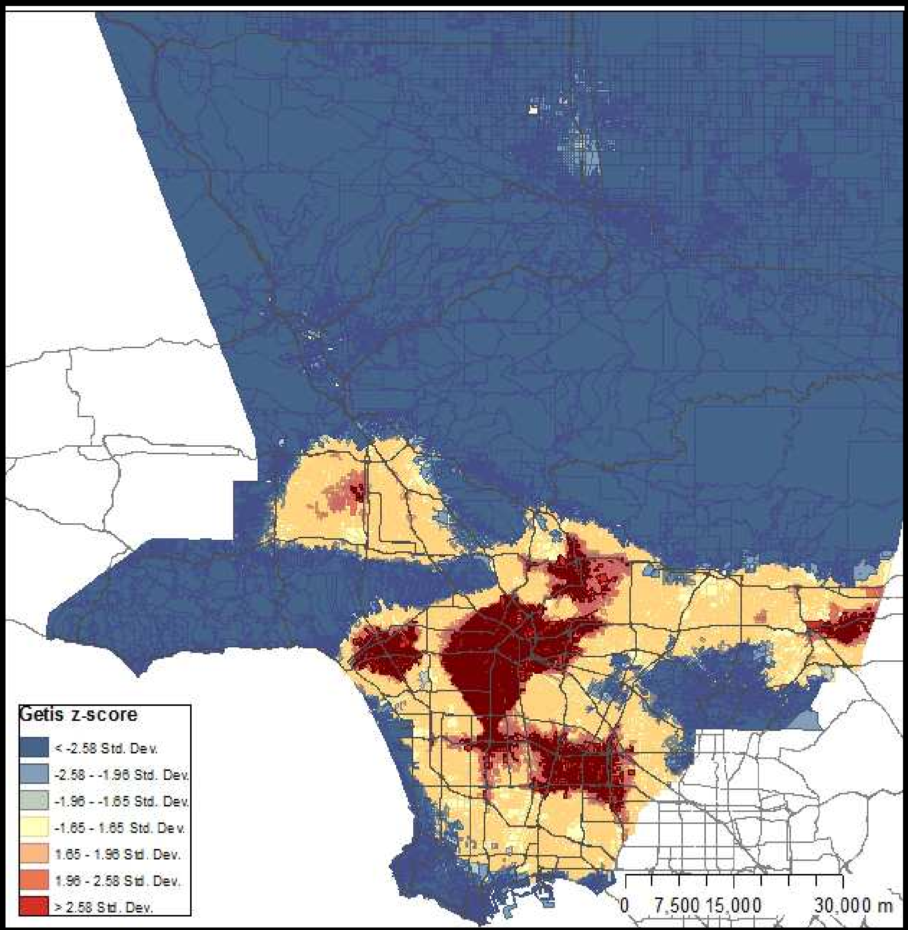
Step 2: G* Analysis

– Data input for Getis-Ord G* statistic
  - AM peak (6AM to 9AM) block level accessibility measures within 10 minutes for ***finance, retail trade, arts/entertainment, healthcare*** **&** ***education***
  - Average land values (price per unit area)
– Output
  - Computed z-scores for the six spatial indicators at block level
  - High opportunities and high land values are positive z-scores
  - Low opportunities and lower land values are negative z-scores
– G* clusters
  - Red indicates high value and clusters
  - Blue indicates low value and clusters
  - Yellow indicates lack of clustering

G* Clusters

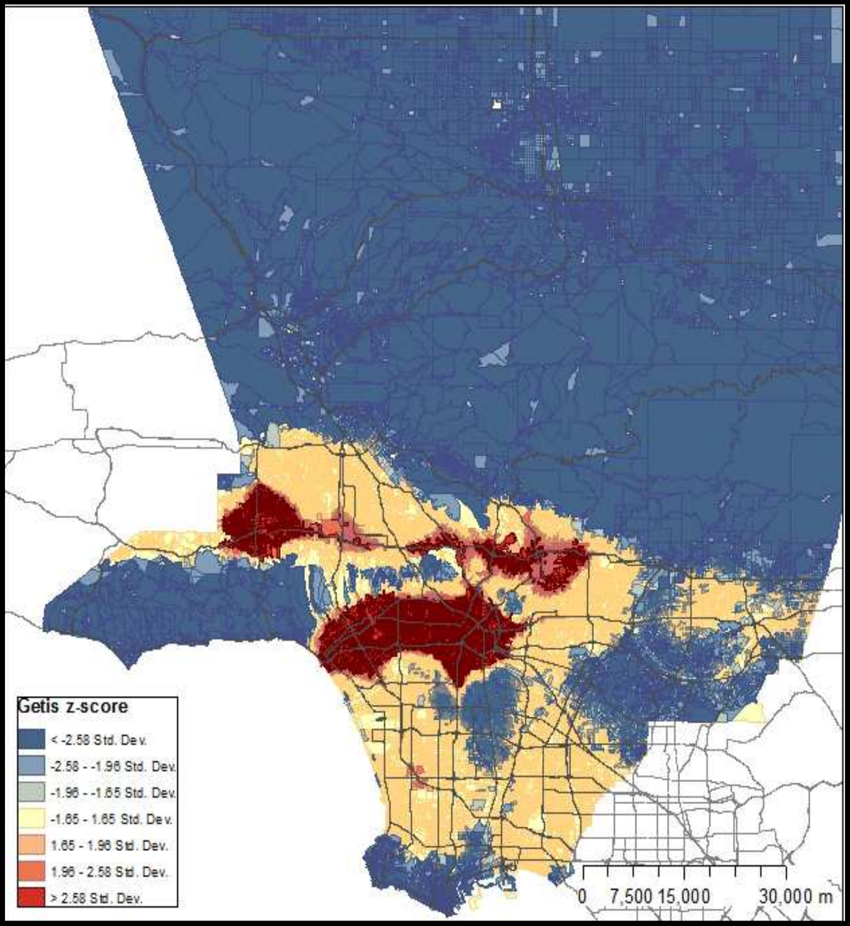


Arts/Entertainment

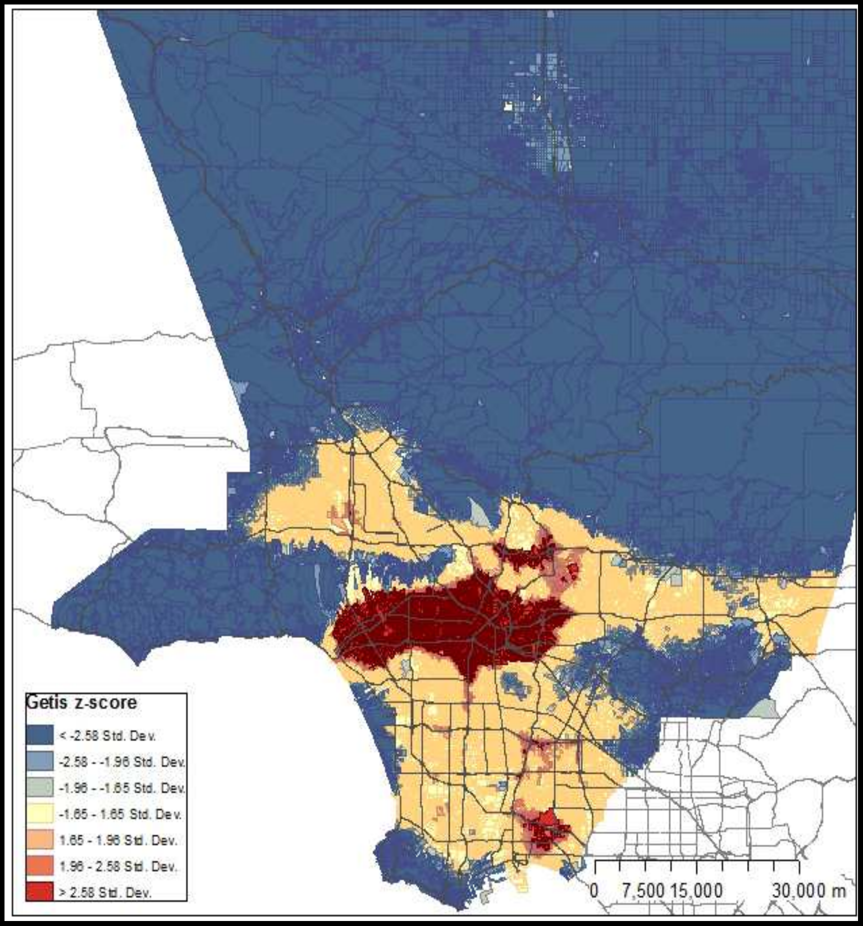


Education

G* Clusters

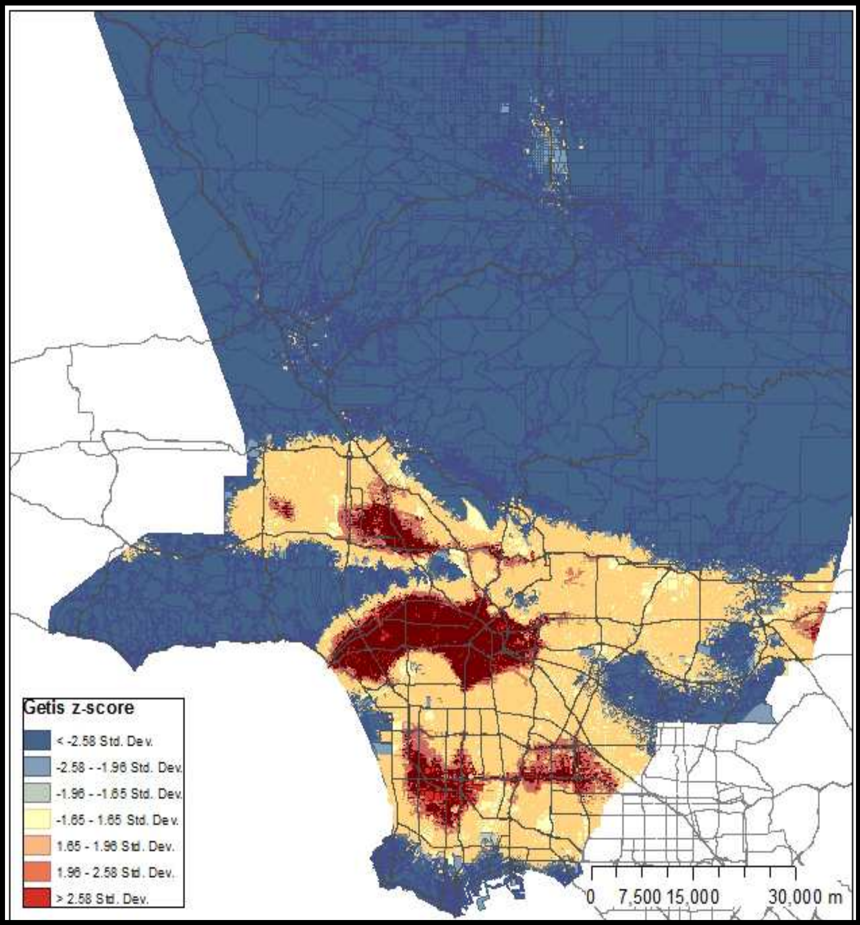


Finance

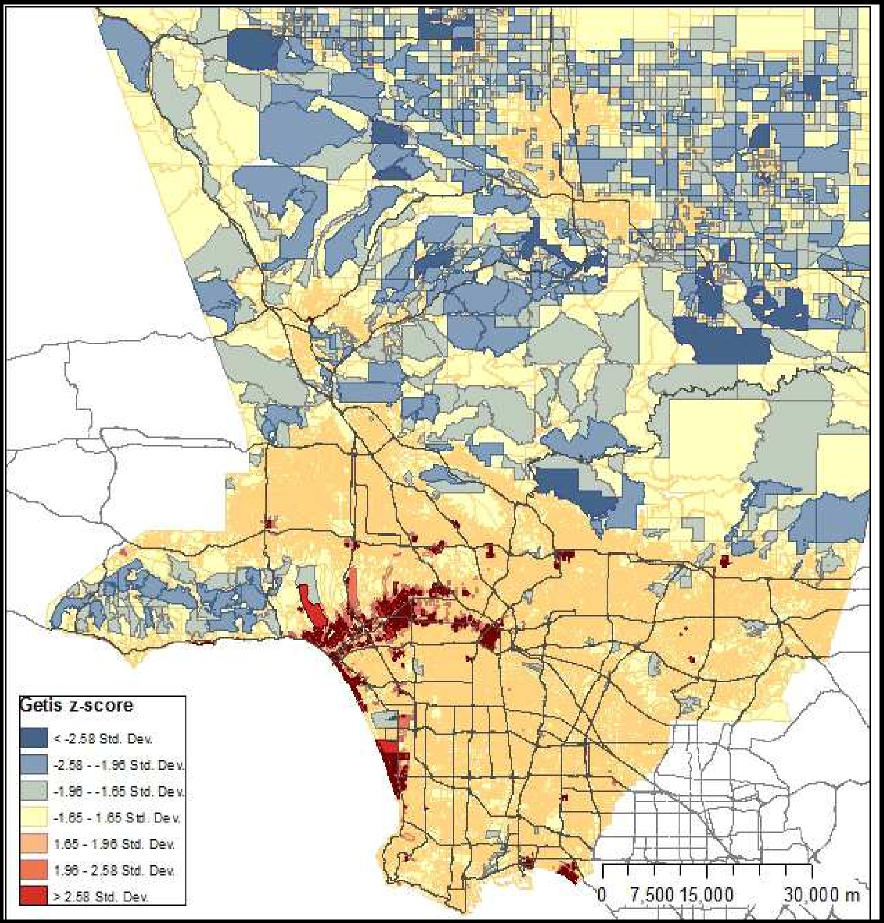


Healthcare

G* Clusters



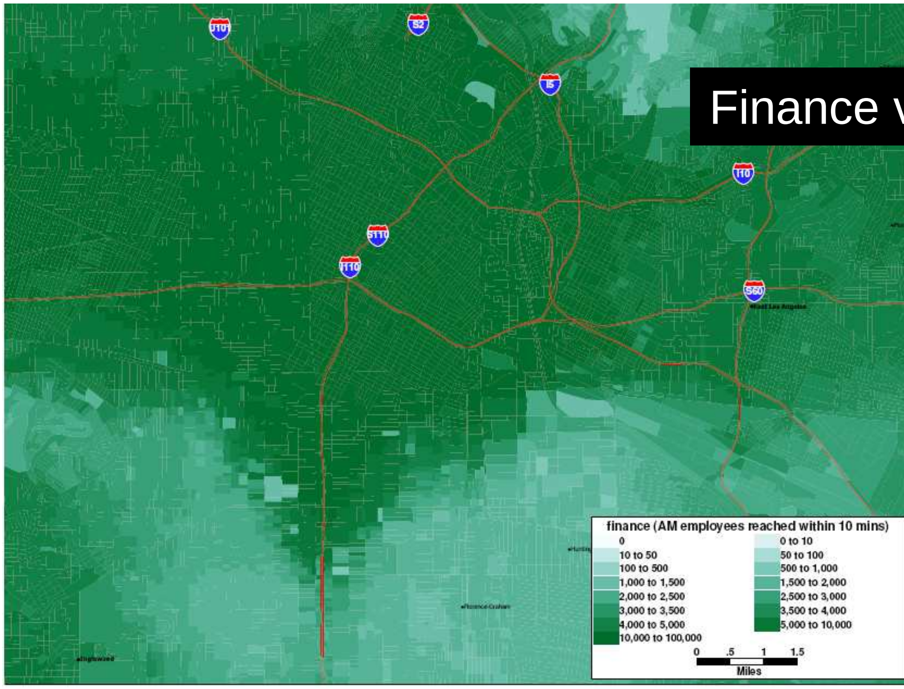
Retail



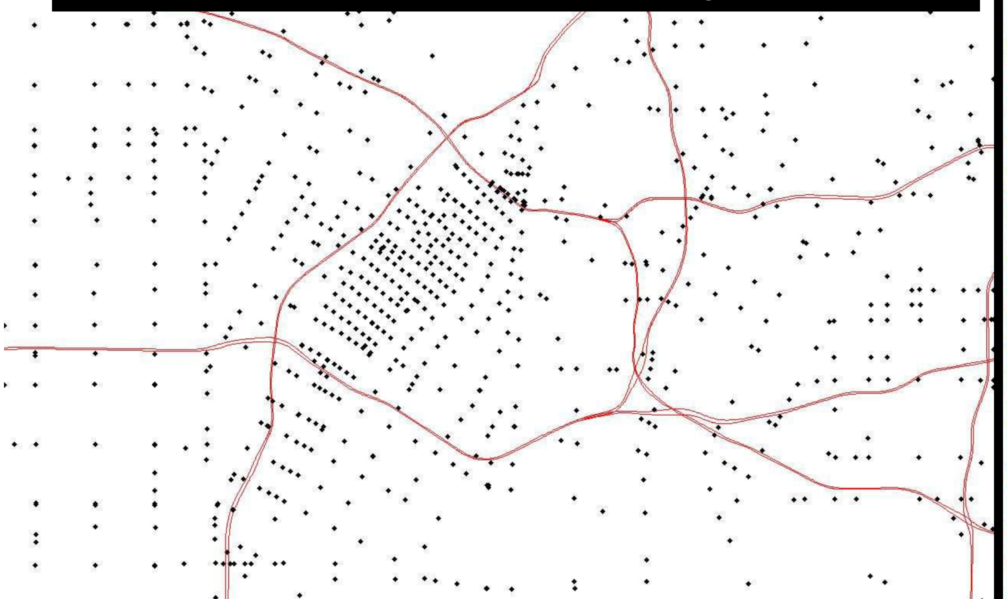
Land Value

Based on: Lei T., K. G. Goulias, and Y. Chen (2011) Opportunity-Based Dynamic Transit Accessibility in Southern California: Measurement, Findings, and Comparison with Automobile Accessibility. Paper 12-3813 presented at the January 2012 *91st Annual Meeting of the Transportation Research Board, Washington, D.C.,* January 22-26, 2012

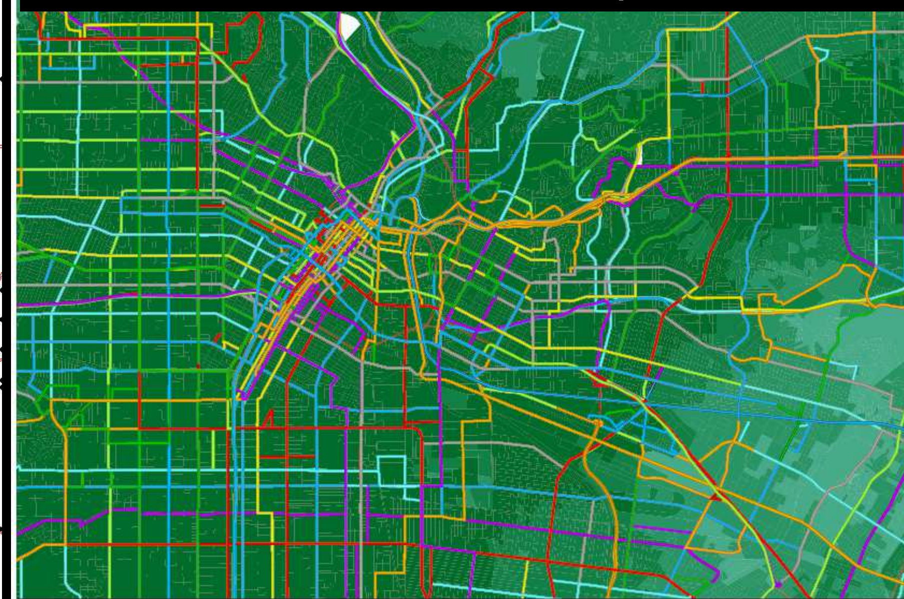
Finance vs Manufacturing

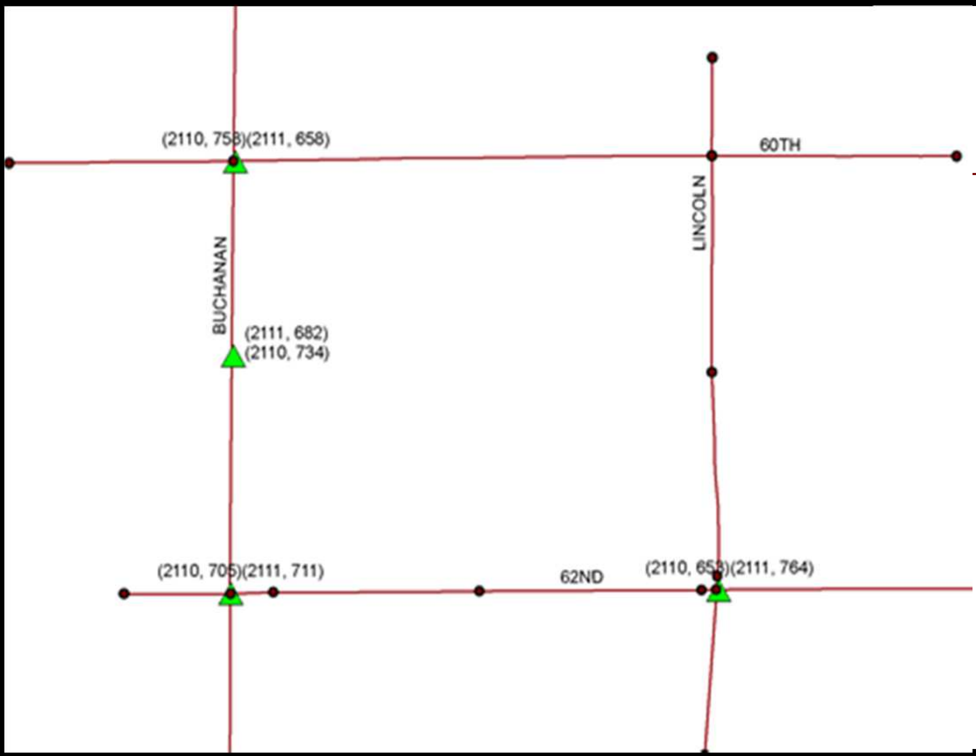


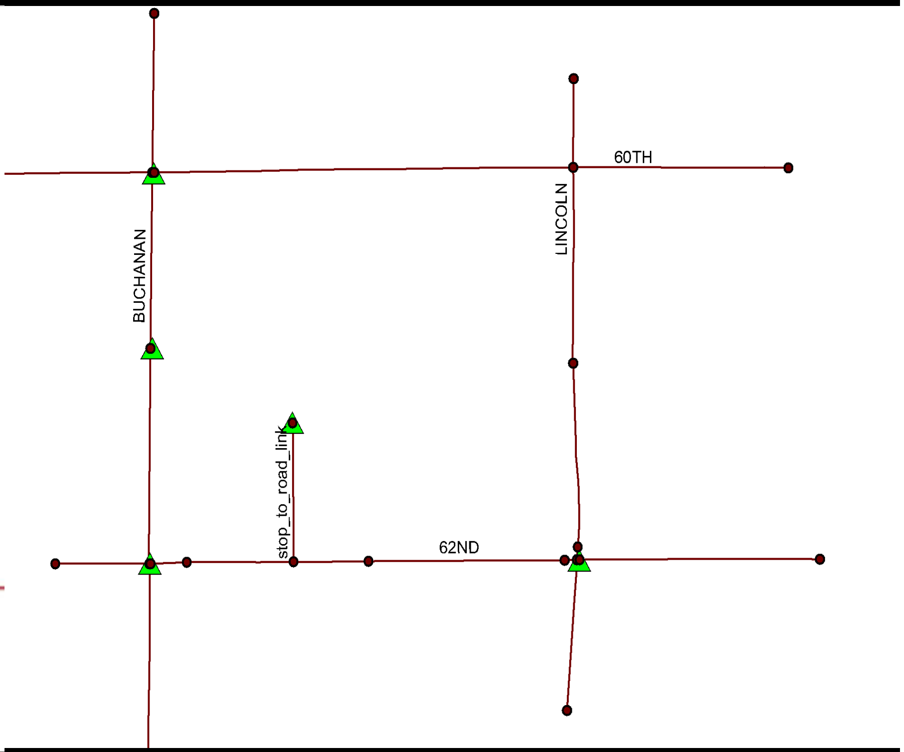
Access Points to Public Transportation

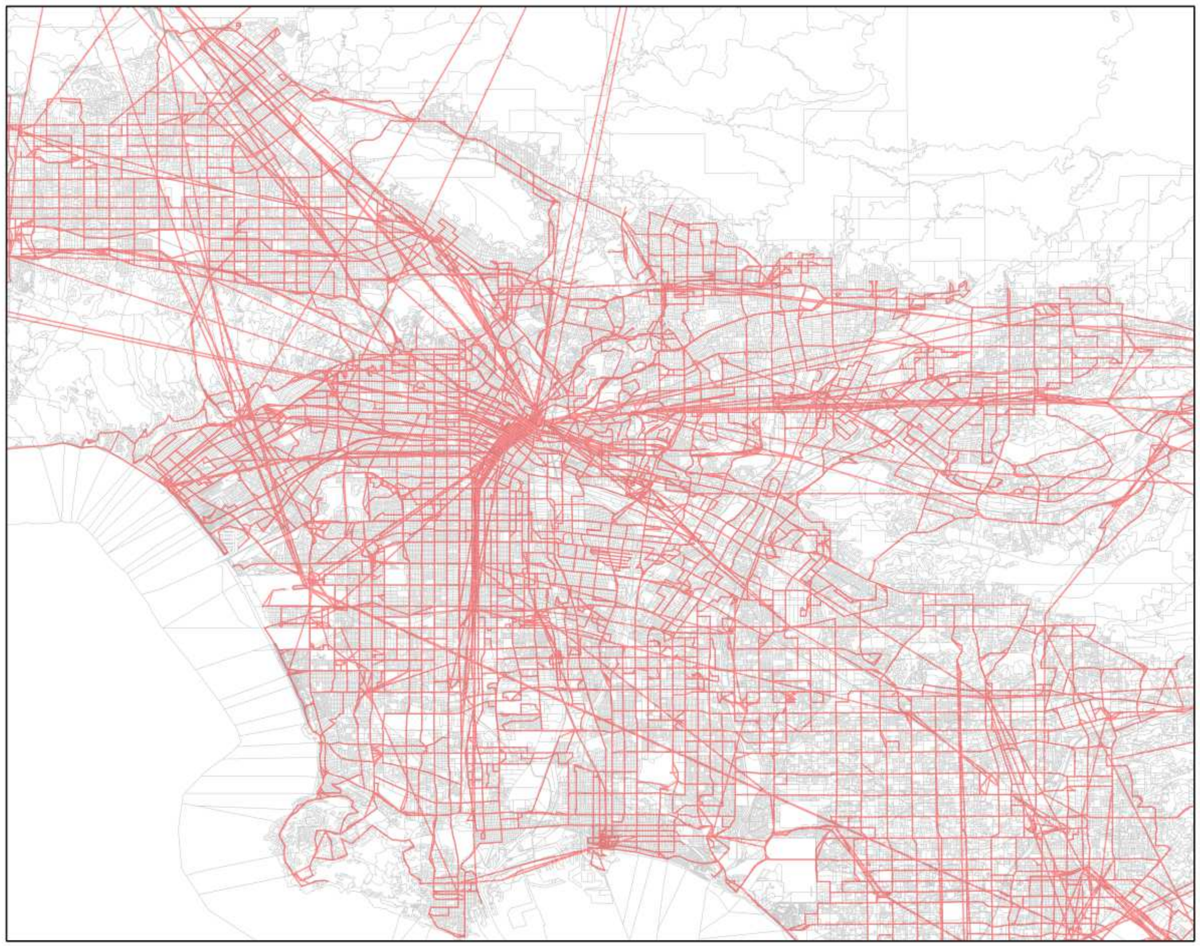


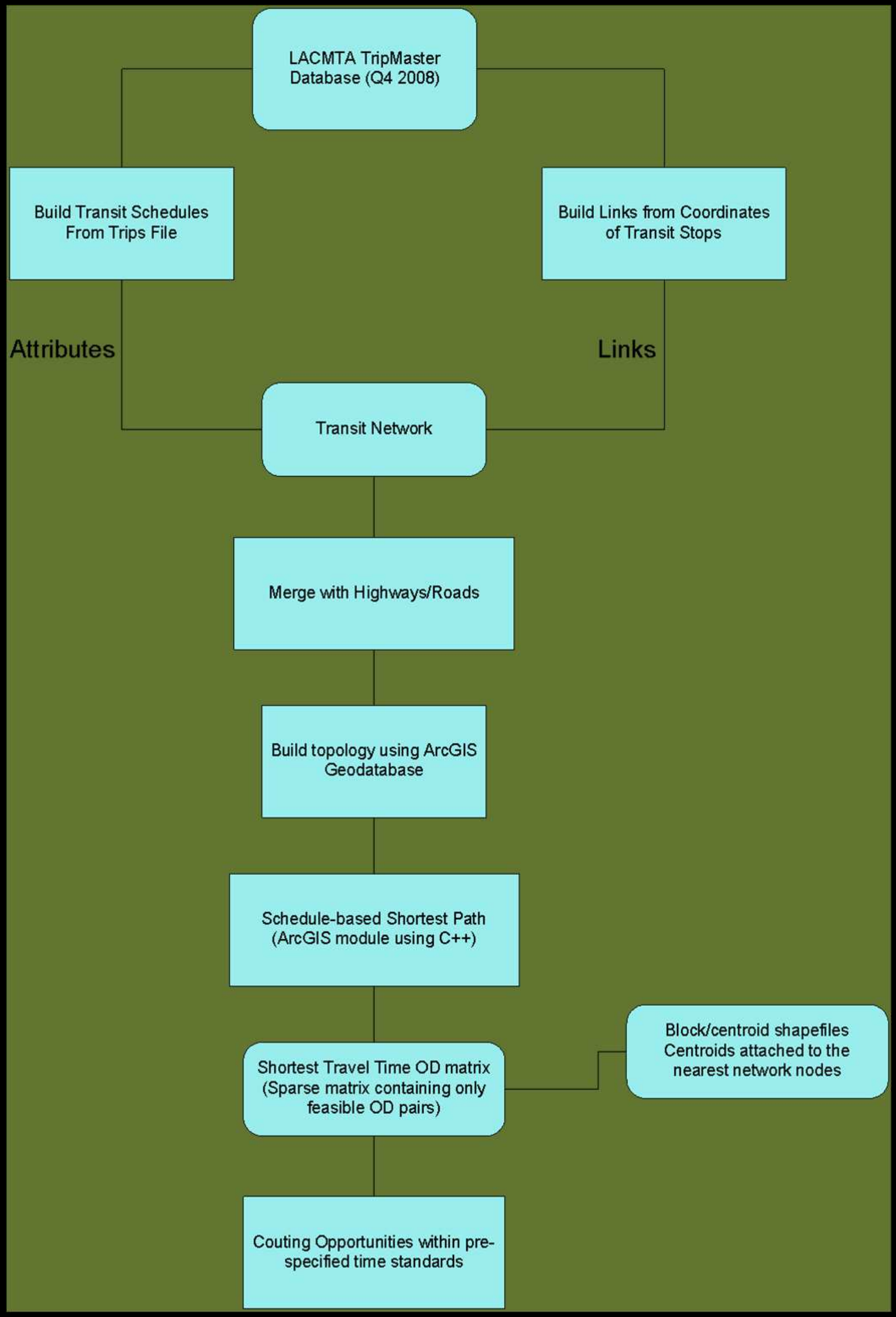
Routes of Public Transportation

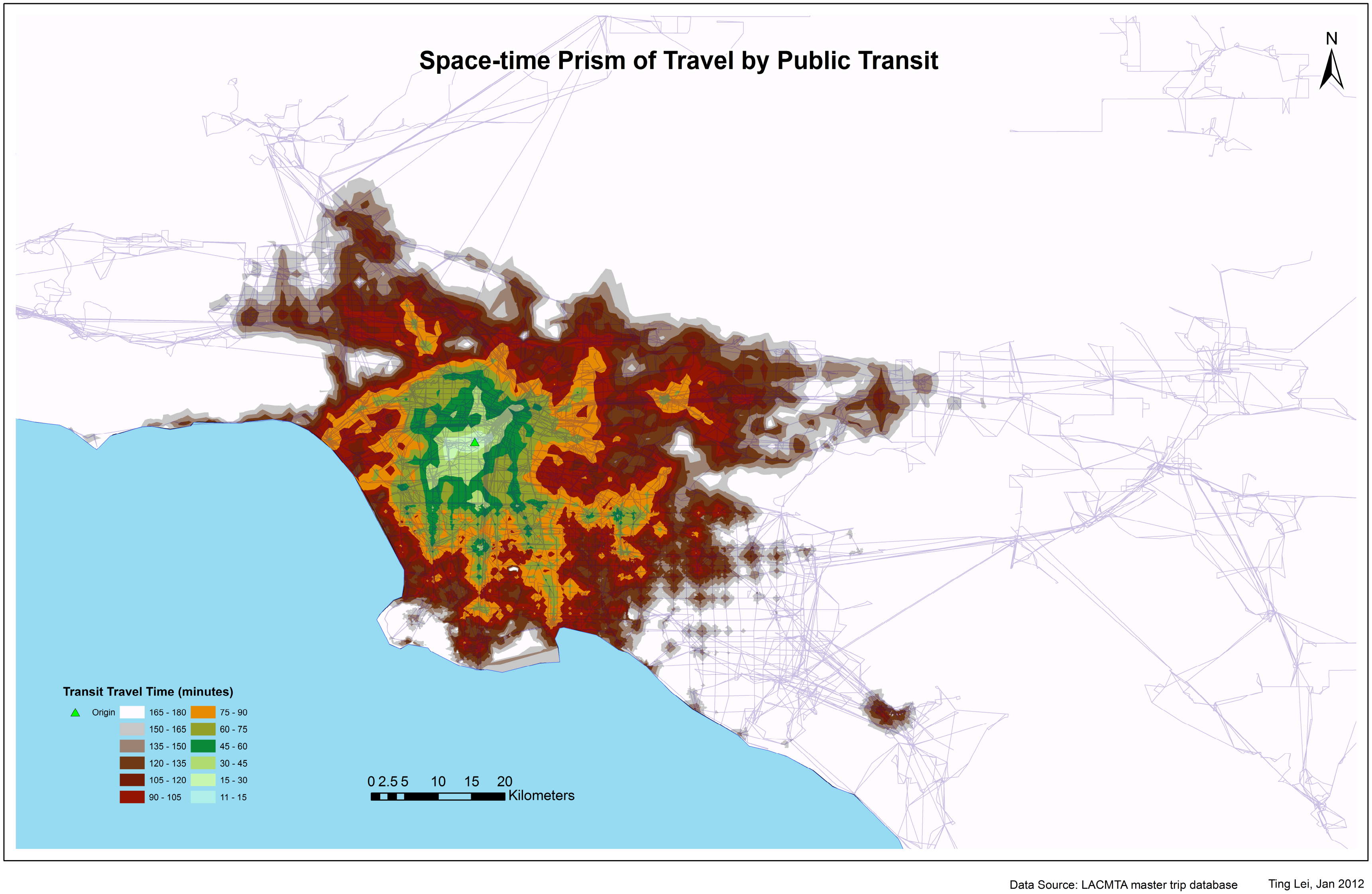



(2110, 758)(2111, 658)
60TH
LINCOLN
BUCHANAN
(2111, 682)
(2110, 734)
(2110, 705)(2111, 711)
62ND
(2110, 658)(2111, 764)


60TH
LINCOLN
BUCHANAN
stop_to_road_link
62ND




LACMTA TripMaster Database (Q4 2008)
Build Transit Schedules From Trips File
Build Links from Coordinates of Transit Stops
Attributes
Links
Transit Network
Merge with Highways/Roads
Build topology using ArcGIS Geodatabase
Schedule-based Shortest Path (ArcGIS module using C++)
Shortest Travel Time OD matrix (Sparse matrix containing only feasible OD pairs)
Block/centroid shapefiles Centroids attached to the nearest network nodes
Couting Opportunities within pre-specified time standards


Space-time Prism of Travel by Public Transit
N
Transit Travel Time (minutes)
Origin
165 - 180
150 - 165
135 - 150
120 - 135
105 - 120
90 - 105
75 - 90
60 - 75
45 - 60
30 - 45
15 - 30
11 - 15
0 2.5 5 10 15 20
Kilometers
Data Source: LACMTA master trip database
Ting Lei, Jan 2012




Maximum number of reachable retail employees for a 20-minute buffer by transit by time of day in Central Los Angeles.




Maximum number of reachable education employees for a 20-minute buffer by transit by time of day in Central Los Angeles.


Maximum number of reachable retail employees for a 20-minute buffer by time of day in Central Los Angeles by Private Car

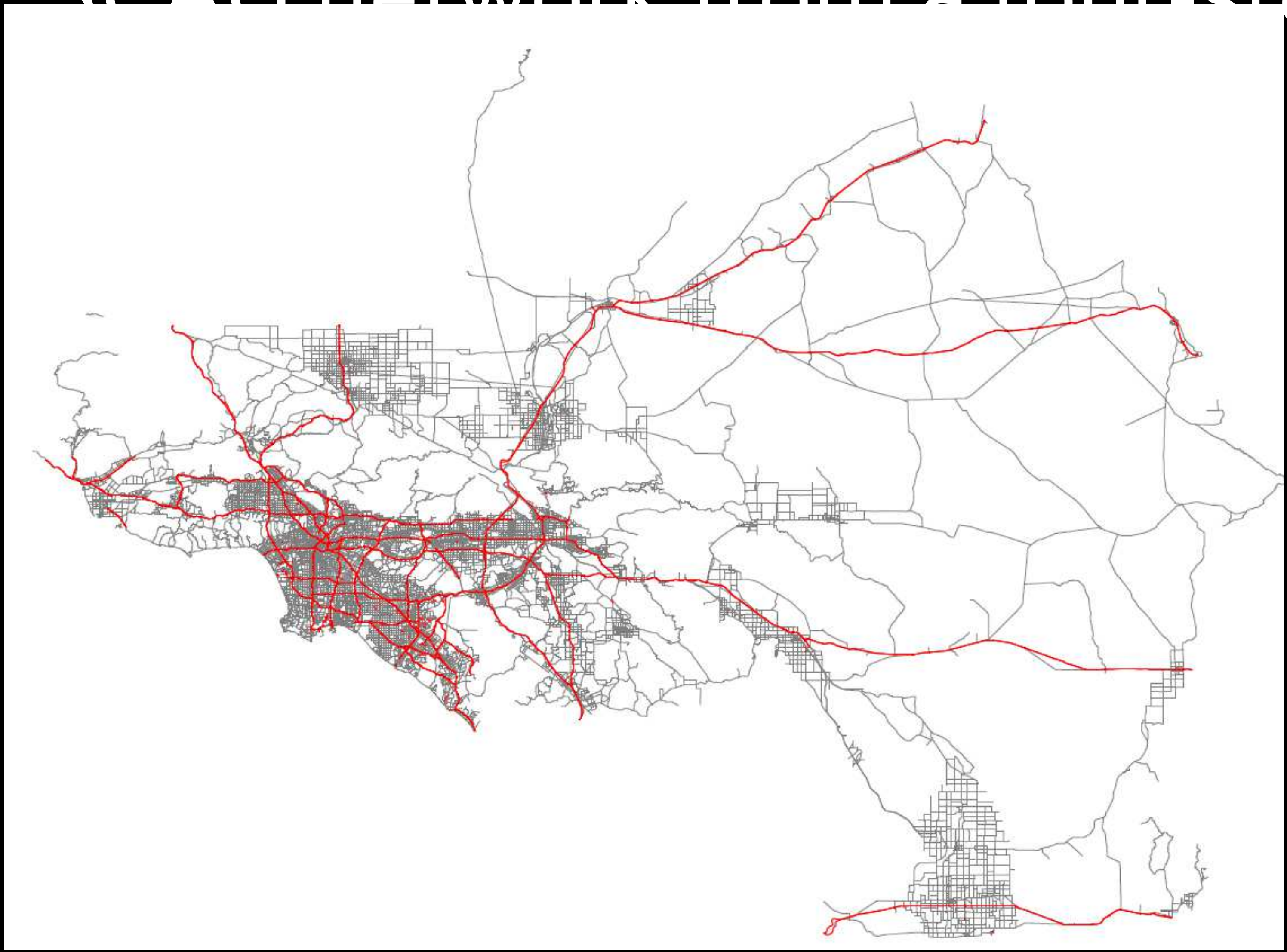

Maximum number of reachable education employees for a 20-minute buffer by time of day in Central Los Angeles by Private Car

- Moving to finer resolutions (more TAZs or even points) increases ability to represent built environment but creates the need to maintain databases over time
- There are some interesting issues about spatial allocation – for now we used count data regression models
- Assign synthetic households to blocks but also assign parcels/housing units to synthetic households
- Start the path toward a demographic microsimulator for at least 25 years and identify methods for fine spatial resolution
- Interface with land use model can offer added information but the space resolution will always be an issue to consider
- Interface with the more recent “sustainability tool(s)” that use GIS maps of parcels open a new gateway to possibilities

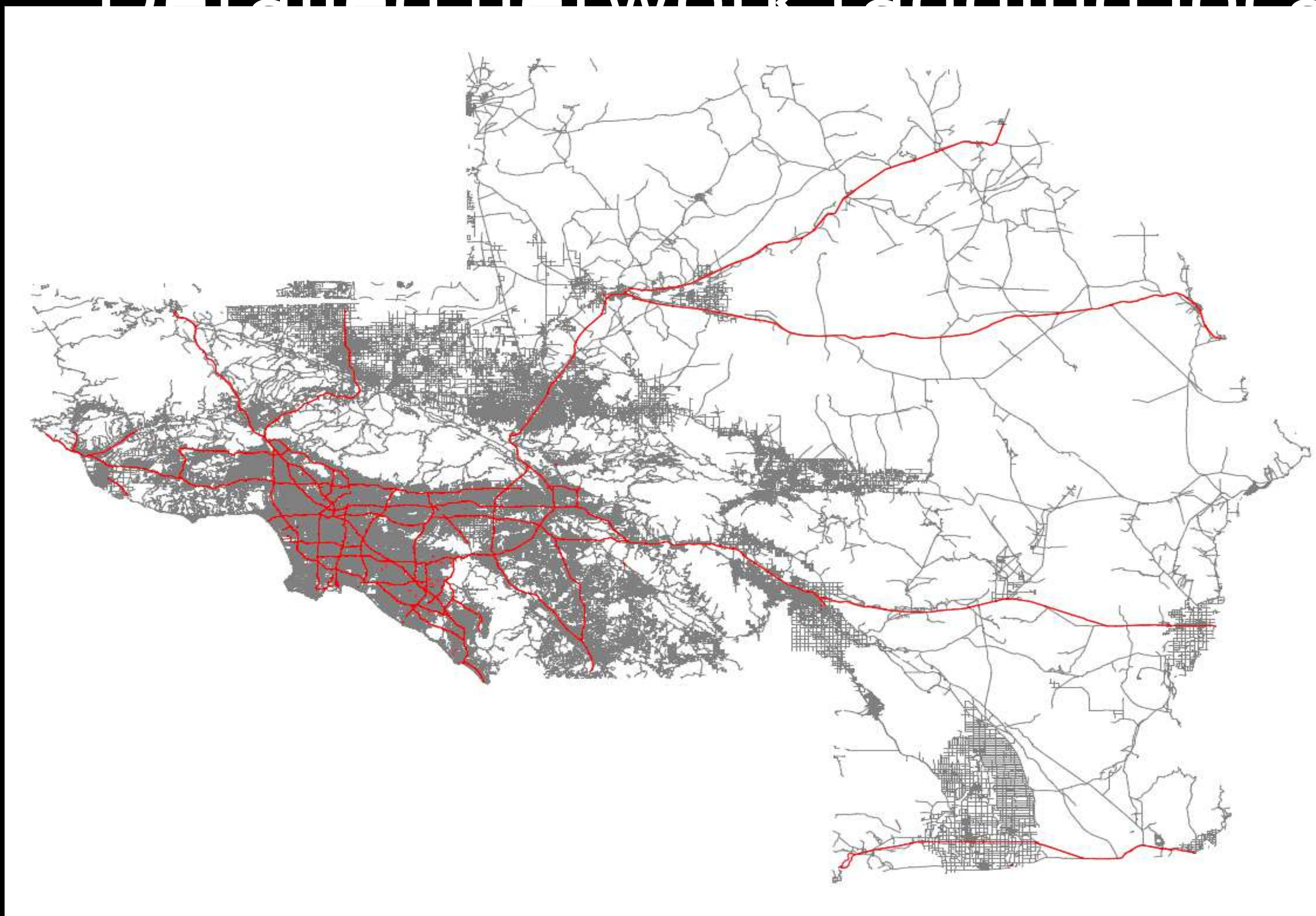
Base map of opportunities

FROM BLOCK GROUPS TO BLOCKS & ADDING NETWORK DETAIL

SCAG network from a four-step



Detailed network (adding local

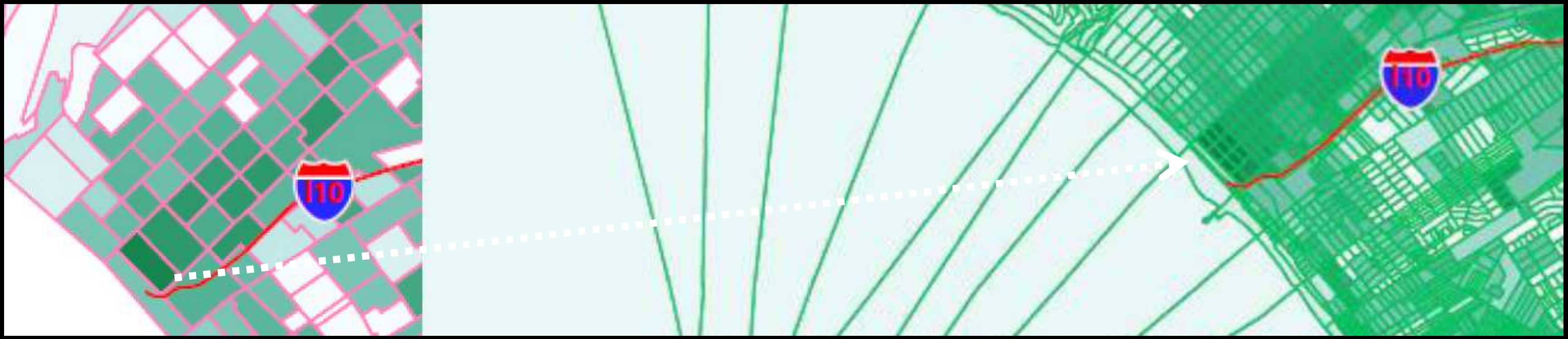


Block Level Centroid Connectors

Allocation from Block groups to Blocks

- Function of land area, population, amount of: freeways, arterials, collectors, ramps (negative and positive influence depends on industry type).

- Tested many different regression models and decided on Poisson regression (count data regression models)
- Kept the total equal to the block group for each block group.



Block Groups Blocks

Using these regression models (one for each industry type), the block group employment for each industry type is distributed into each block within a block group as shown below:

a. Calculate the estimated block employment ($\hat{E}_i$) for block i by applying the block characteristics and the estimated regression model

b. Calculate the percentage of employment (P_i) for each block i within each block group j.

$$P_i = \frac{\hat{E}_i}{\sum_{i \in B_j} \hat{E}_i}$$, where B_j is the set of all blocks within block group j.

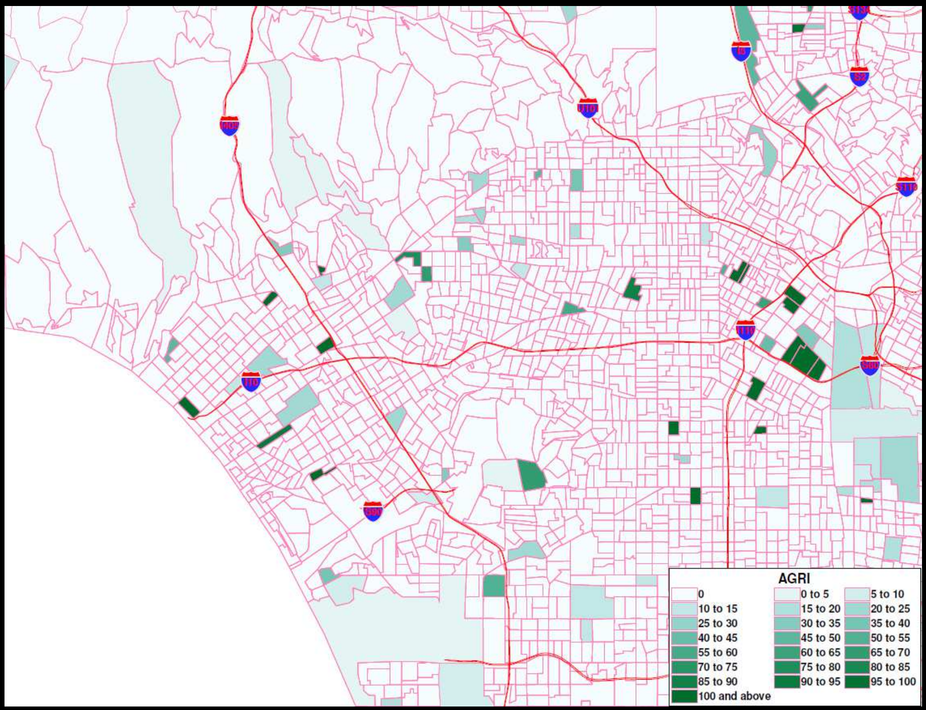
c. Calculate the block employment (E_i) by multiplying the percentage with the block group employment

$$E_i = P_i \times (\text{employment in block group } j \text{ that block } i \text{ belongs to})$$.

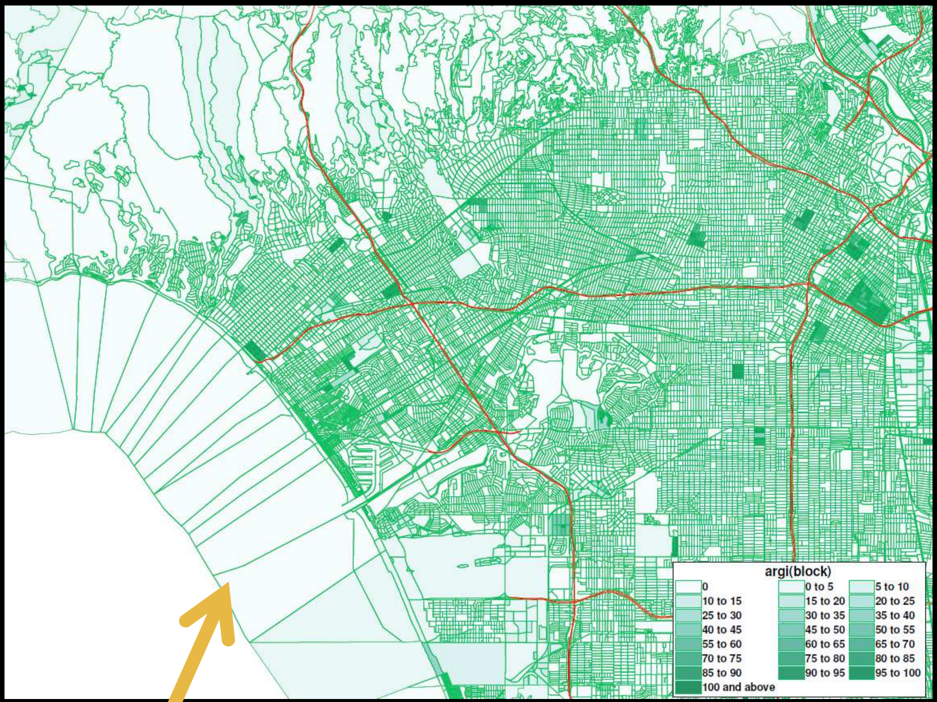
The same process is repeated for j = 10,631 block groups and 15 industry types.

Agriculture Density

Block group level(observed)



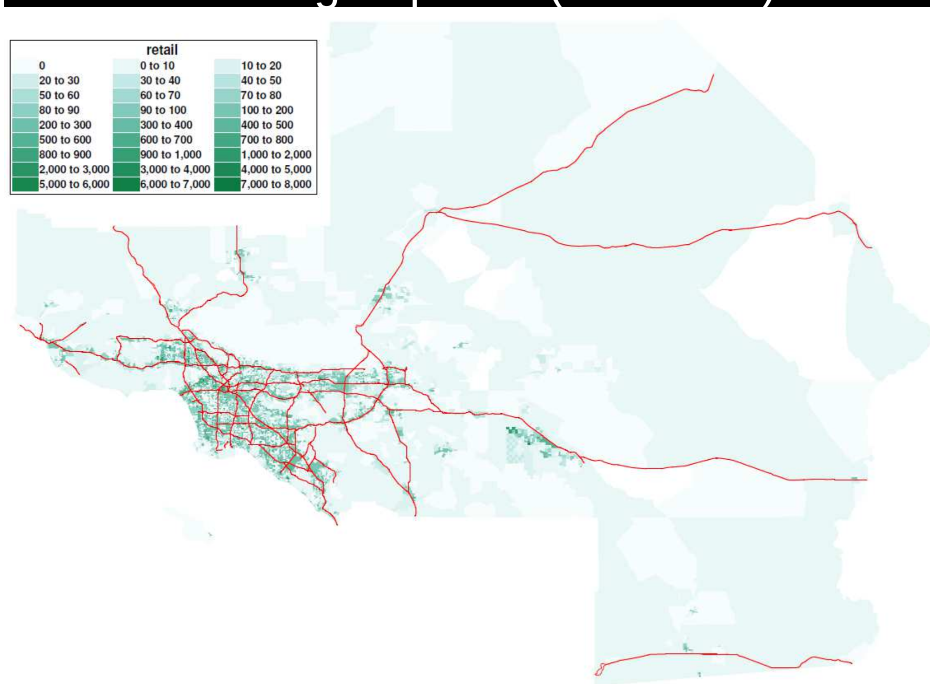
Block level(predicted from model)



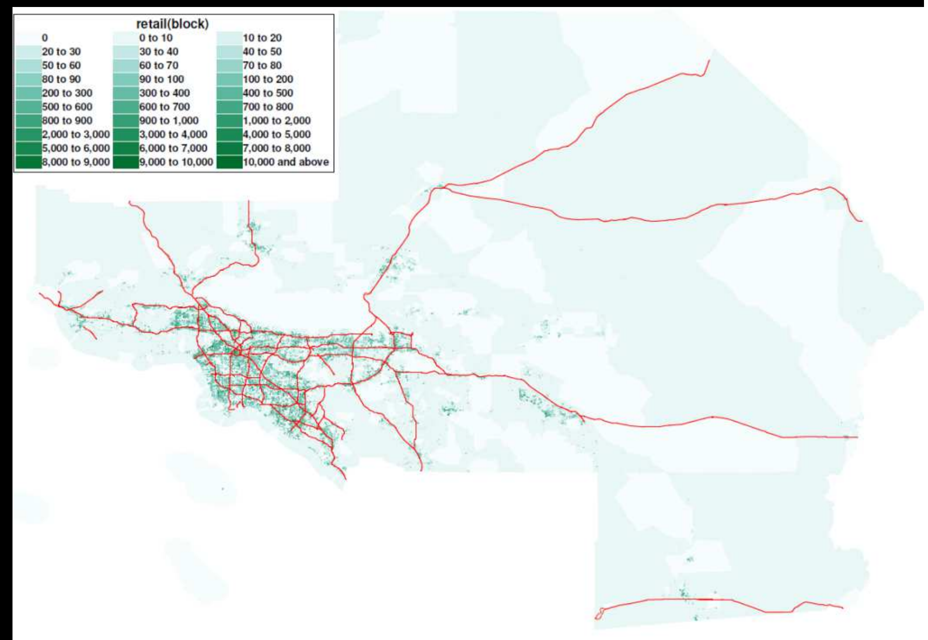
Blocks in the Ocean – special treatment

Retail Density

Block group level(observed)

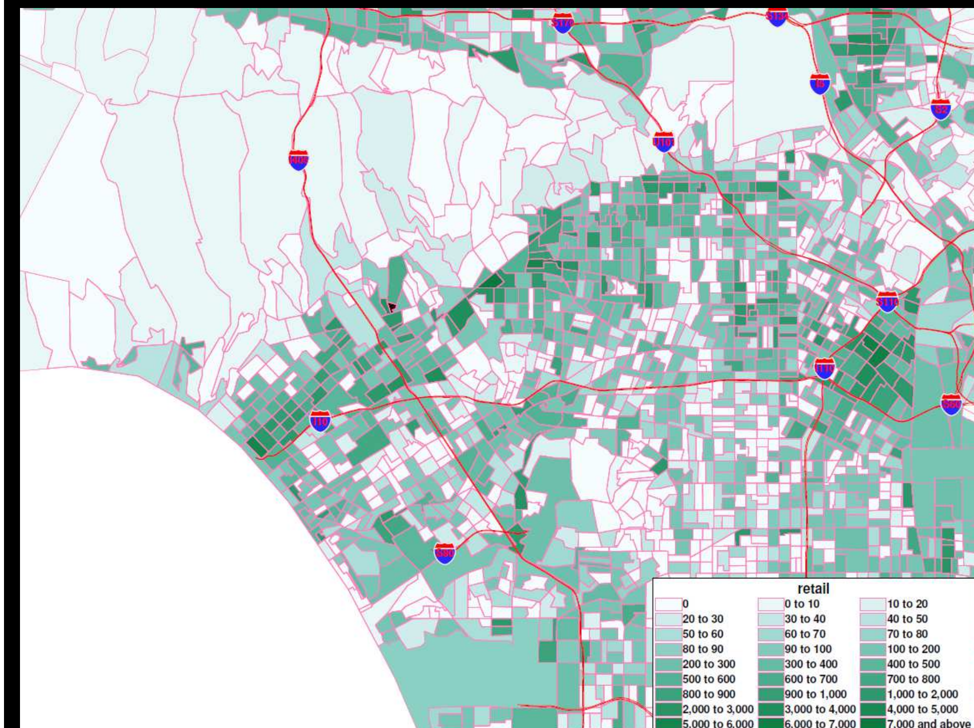


Block level(predicted allocation model)

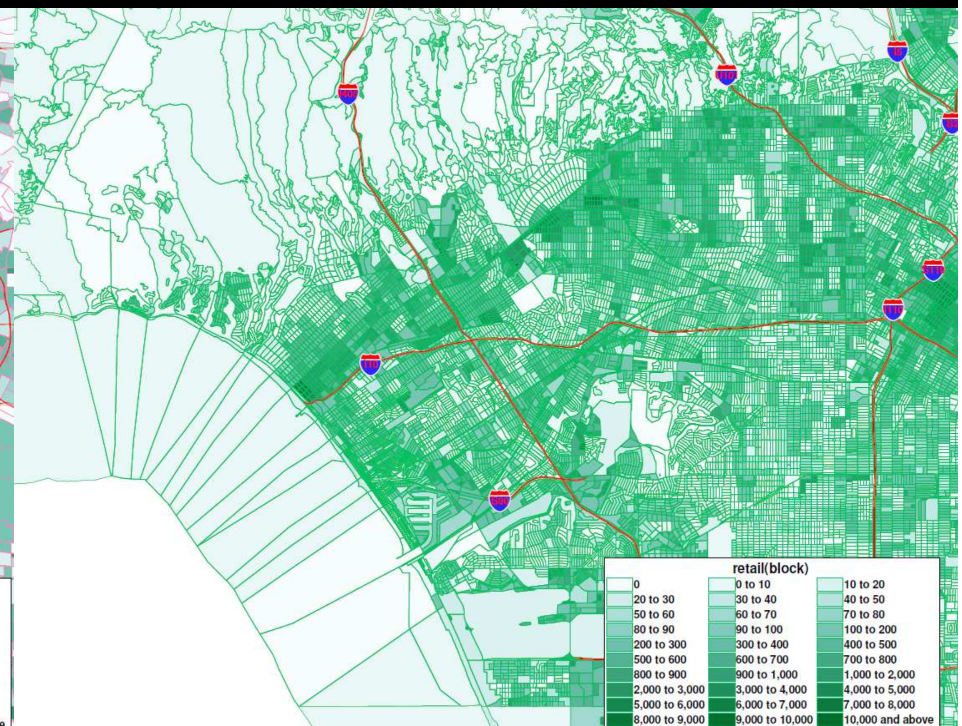


Retail Density

Block group level(observed)



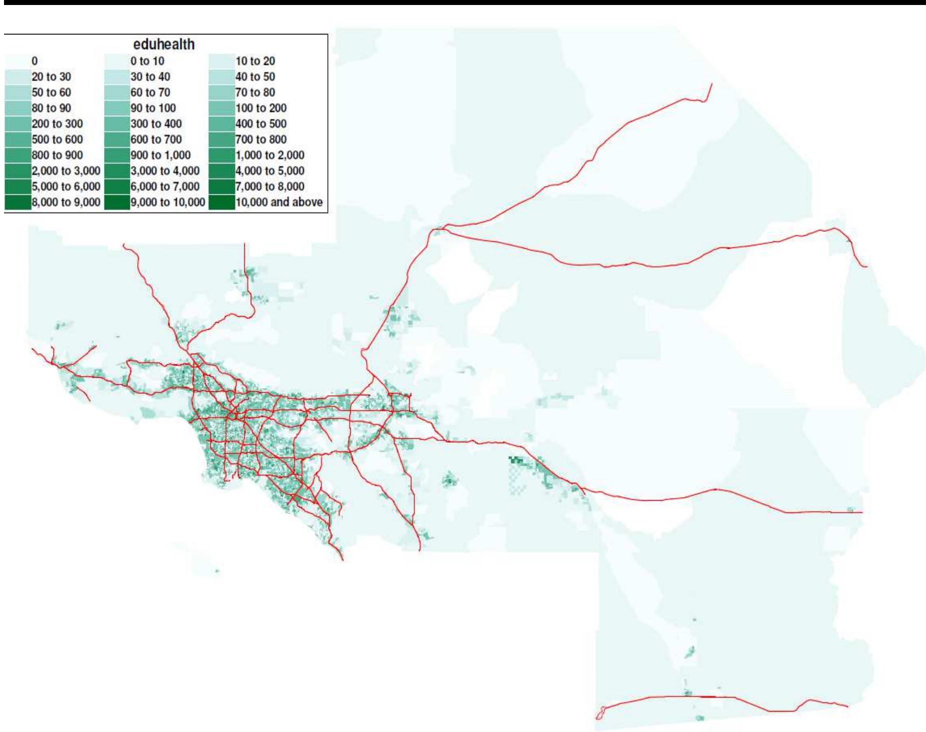
Block level(predicted from model)

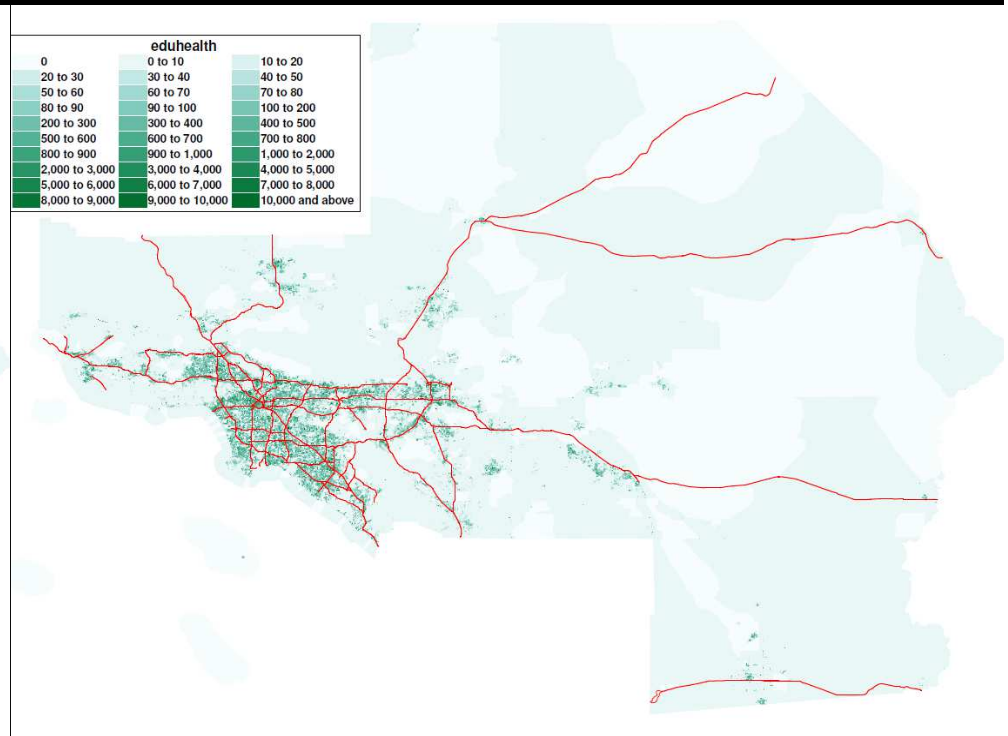


Education & Health Density

Block group level(observed)

Block level(predicted from model)



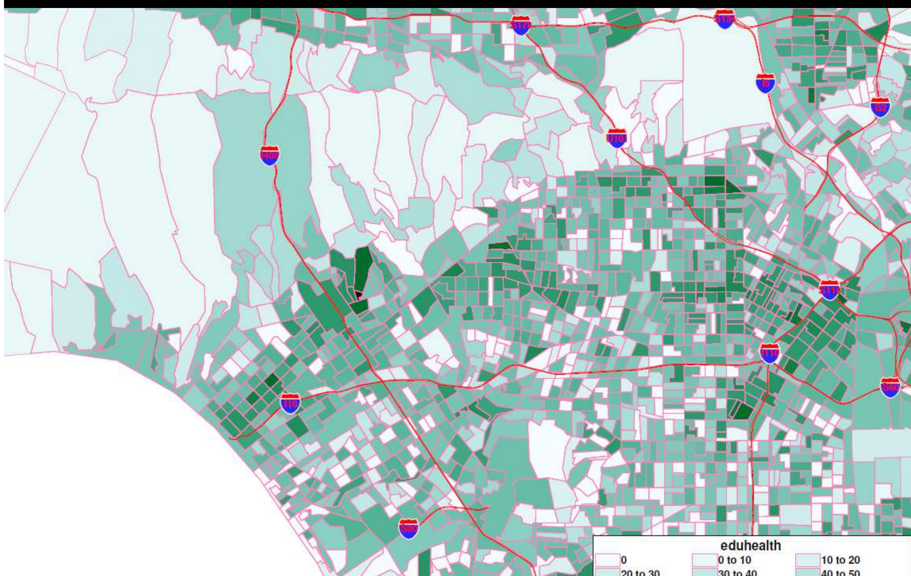


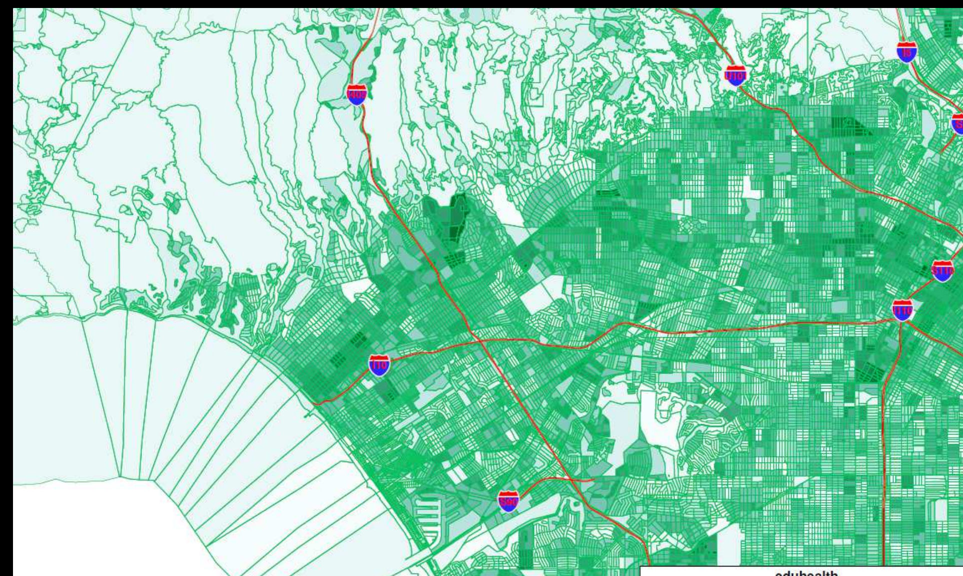
We split this in two separate sets of indicators

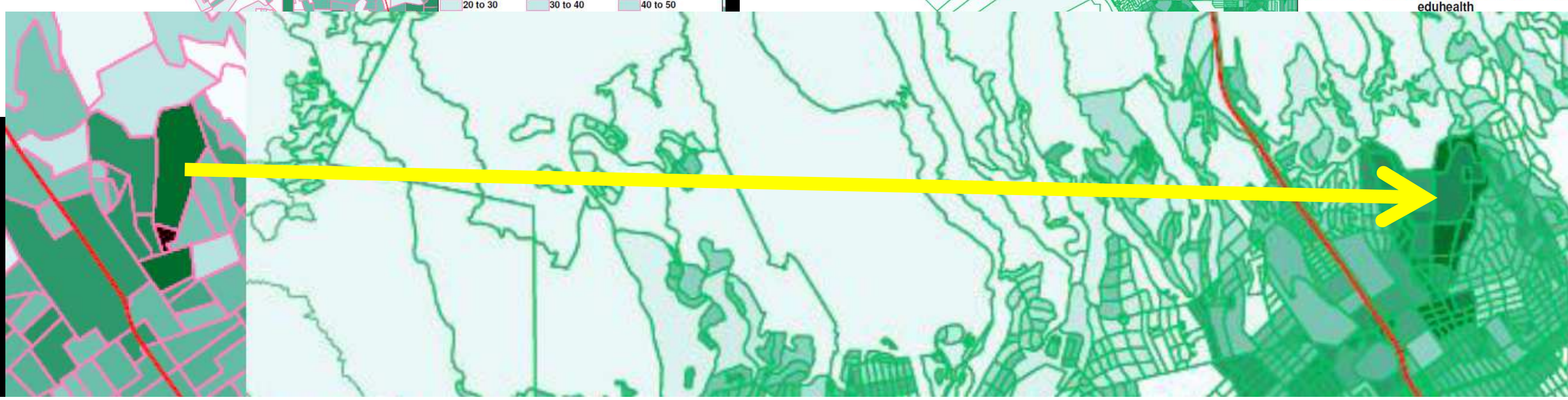
Education & Health Density

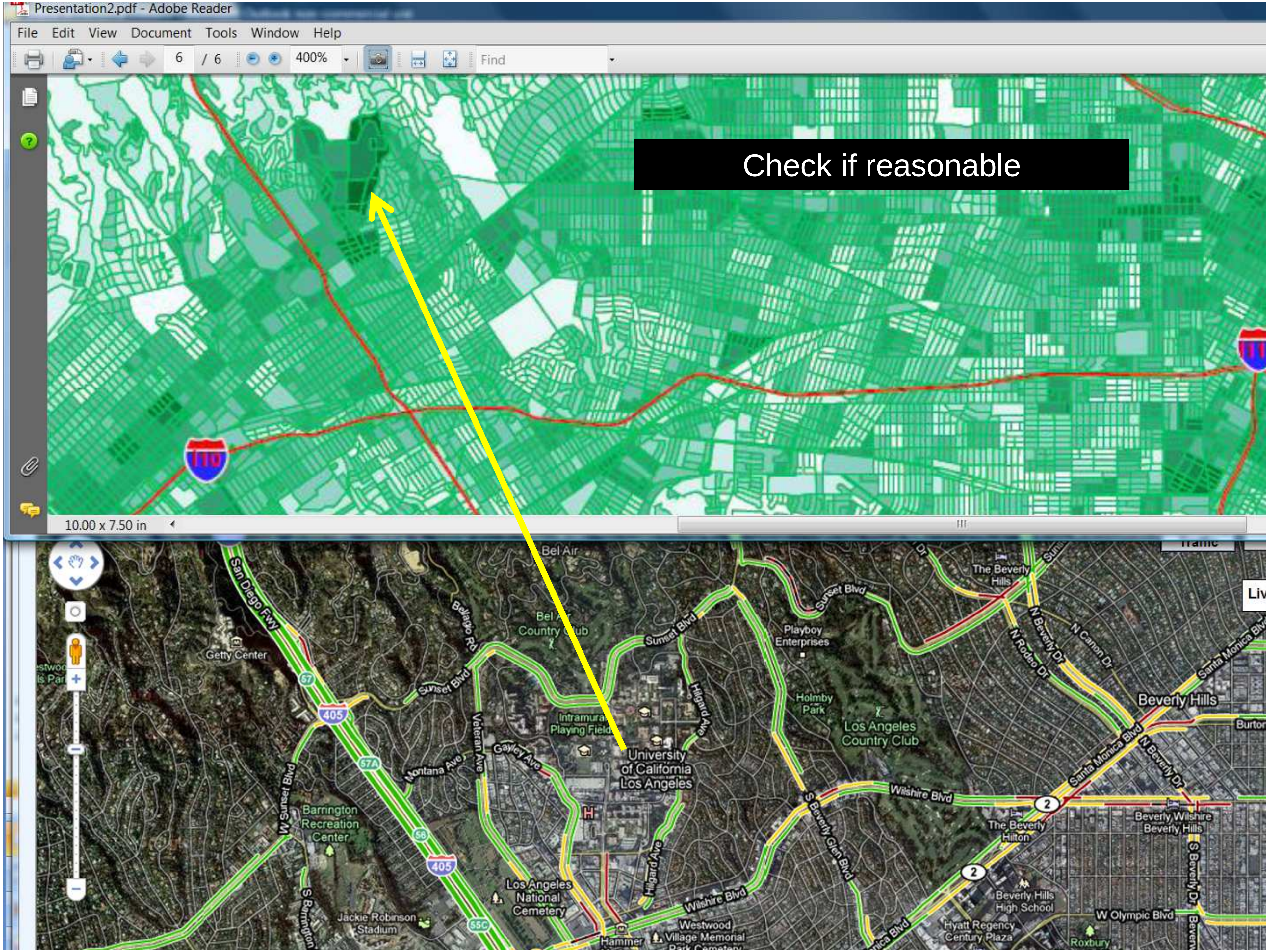
Block group level(observed)

Block level(predicted from model)








Check if reasonable
Presentation2.pdf - Adobe Reader
File Edit View Document Tools Window Help
6 / 6
400%
Find
10.00 x 7.50 in
Bel Air
Bel Air Country Club
Getty Center
San Diego Fwy
Sunset Blvd
Playboy Enterprises
Holmby Park
Los Angeles Country Club
The Beverly Hills
N Beverly Dr
N Rodeo Dr
N Canon Dr
Santa Monica Blvd
Beverly Hills
Burton
Hilgard Ave
Intramural Playing Field
Gayley Ave
Veteran Ave
Montana Ave
University of California Los Angeles
Wilshire Blvd
S Beverly Glen Blvd
The Beverly Hilton
Beverly Wilshire Beverly Hills
S Beverly Dr
Barrington Recreation Center
W Sunset Blvd
Los Angeles National Cemetery
Jackie Robinson Stadium
S Barrington
Westwood Village Memorial Park Cemetery
Hammer
Beverly Hills High School
Hyatt Regency Century Plaza
W Olympic Blvd
Roxbury

Issues Found and Resolved

- Block groups with employees but no businesses
- Block groups with businesses but no employees
- Disagreements between D&B number of persons by industry and CTPP
- Differences in NAICS-SIC classifications
- CTPP reports education and health as one industry (split in two)

- You can find additional details at: Chen, Y., S. Ravulaparthy, K. Deutsch, P. Dalal, S.Y. Yoon, T. Lei, K.G. Goulias, R.M. Pendyala, C.R. Bhat, and H-H. Hu (2011) Development of Opportunity-based Accessibility Indicators. *Transportation Research Record: Journal of the Transportation Research Board, No. 2255,* Transportation Research Board of the National Academies, Washington D.C., 2011, *pp.58-68.*

AVAILABILITY OF OPPORTUNITIES BY TIME OF DAY

Time of Day Components

- Travel time on the network -> use travel speed from calibrated four step model that has four wide time periods (AM Peak - 6AM to 9AM, Midday - 9AM to 3PM, PM Peak - 3PM to 7PM, Night - 7PM to 6AM)
- Opening-closing times for businesses -> use arrival and departure time of workers in post Census survey (16k hhs)
- To account for the “wide” periods of the network times we use MIN and MAX for the number of employees by industry that can be reached within a temporal buffer of 10, 20, and 50 minutes
- The buffers 10 vs 50 resemble local vs regional accessibility
- Shortest path computed using aprox. 200,000 by 200,000 origin destination matrix using TRANSCAD.

Employee Classification

1. Wholesale Trade
2. Retail Trade
3. Administrative and Support Services
4. Educational Services
5. Healthcare and Social Assistance
6. Arts, Entertainment and Recreation
7. Accommodations and Food Services
8. Public Administration
9. Agriculture/Forestry/Fishing
10. Construction
11. Manufacturing – Durable and Nondurable goods
12. Transportation and Warehousing
13. Information
14. Finance or Insurance
15. Professional, Scientific and Technical Services

Using a survey of workers for each respondent q that works, we have the arrival time at work (say t_{qa}) and departure time from work (say t_{qd}).

Define a dummy variable δ_{qgh} that takes a value of one if individual q works in industry type g and county h, and zero otherwise.

Consider a specific hour of the day k, and let T_s be the start time of this hour k, and let T_e be the end time of this hour k.

Define another dummy variable α_{qk} that takes a value 1 if $t_{qs} \geq T_s$ and $t_{qe} \leq T_e$, and zero otherwise.

Intuitively, α_{qk} equals one if the employee q is at the work place at the hour k and zero otherwise.

From the travel survey, we also have a weight W_q for person q, which expands the individual so that the expanded sample across all surveyed individuals is representative of the population.

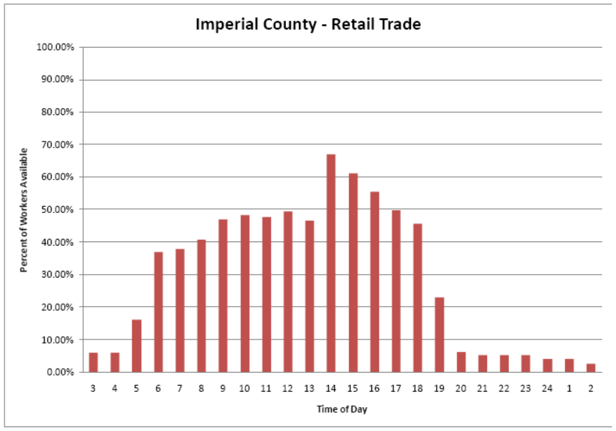
An estimate of the number of reachable employees in industry type g at county h at time k (A_{ghk}) is as follows:

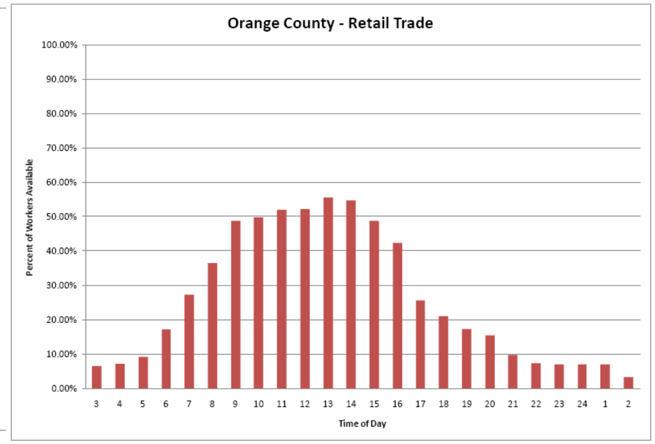
$$A_{ghk} = \sum_{q} \delta_{qgh} \alpha_{qk} W_q$$

The total number of reachable workers in industry g at county h at time k can be expressed as a function of the percentage of reachable workers in industry g and county h:

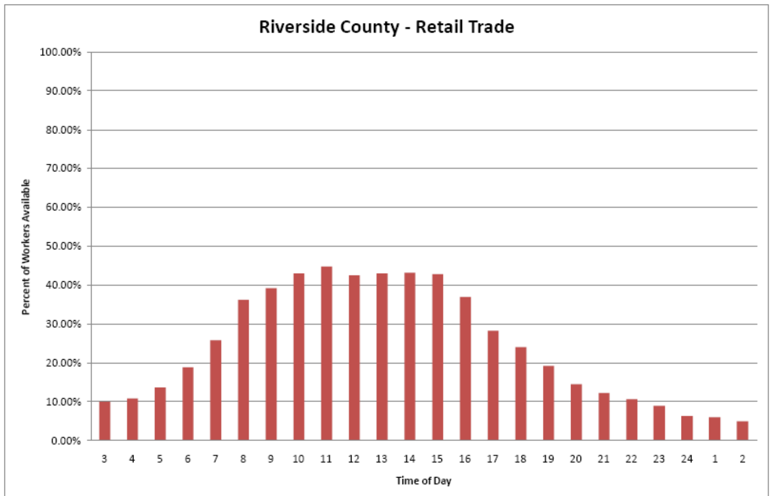
$$P_{ghk} = \frac{A_{qhk}}{\sum_{q} \delta_{qgh} W_q}.$$

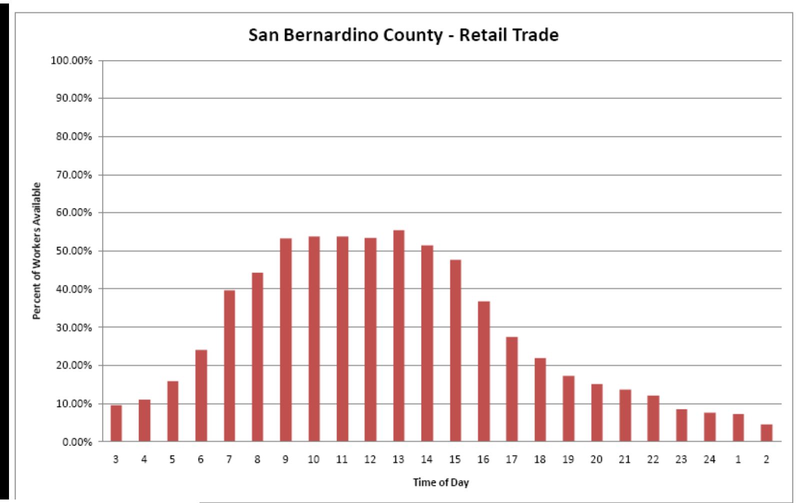
The next slide shows examples of the reachable percent of workers in Los Angeles County by time of day.


Imperial County - Retail Trade
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
Percent of Workers Available
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 1 2
Time of Day

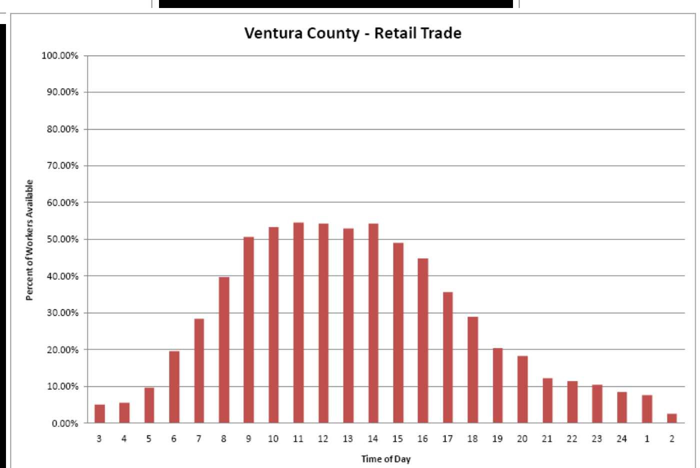

Orange County - Retail Trade
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
Percent of Workers Available
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 1 2
Time of Day


Los Angeles County - Retail Trade
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
Percent of Workers Available
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 1 2
Time of Day


Riverside County - Retail Trade
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
Percent of Workers Available
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 1 2
Time of Day


San Bernardino County - Retail Trade
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
Percent of Workers Available
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 1 2
Time of Day

Pghk

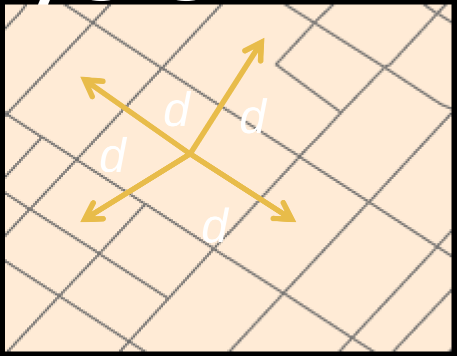

Ventura County - Retail Trade
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
Percent of Workers Available
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 1 2
Time of Day

Identification of Activity Opportunity Clusters

Based on: Ravulaparthy S., P. Dalal, Y. Chen, and K. G. Goulias (2011) An Exploratory Analysis of Spatial Hierarchical Clustering Using Opportunity-based Accessibility and Reported Land Values and its Relationship with Residents Characteristics in Los Angeles County. Paper 12-1297 presented at the *91st Annual Meeting of the Transportation Research Board, Washington, D.C.,* January 22-26, 2012.

Step 2: G* Analysis

- Spatial clusters analysis
  - Spatial dependence
  - Spatial heterogeneity
- Getis-Ord *G*(d)* statistic is a local measure of spatial dependence
  - Measures spatial association between any given point in space to all other neighboring points
  - Defined by a distance of *d*
  - Neighborhood blocks sharing a boundary
- Getis-Ord G* statistic calculates the z-score
  - Statistical significance of clustering
- z-score significance
  - Higher or lower values indicates high intensity of clustering
  - Near zero indicates no clustering in study area

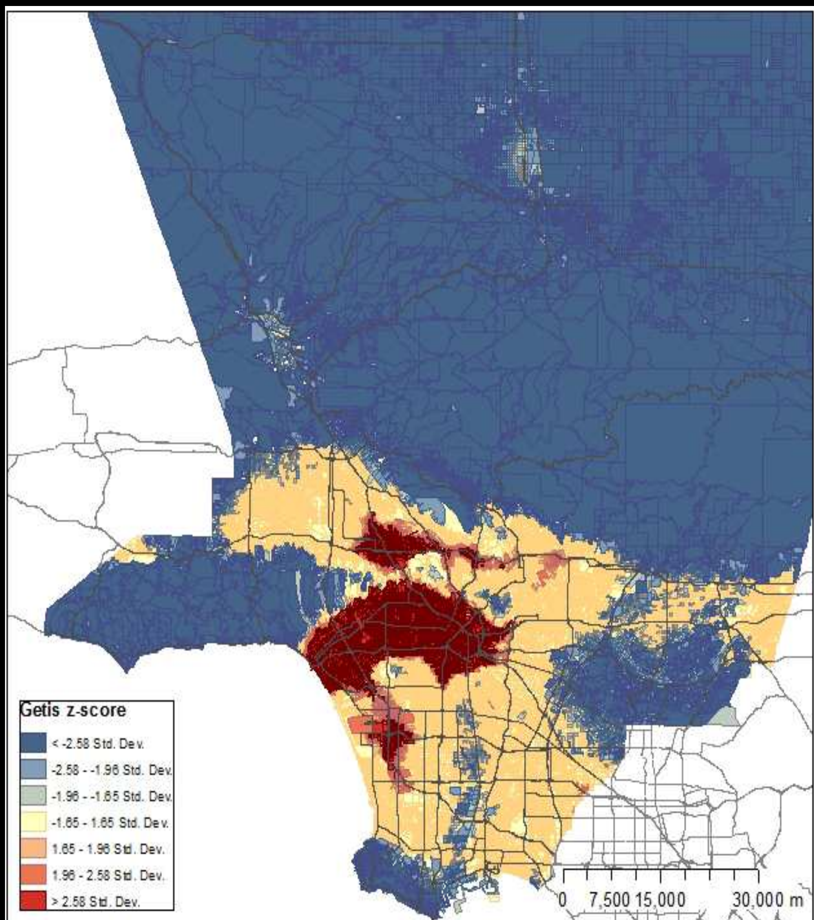


Jacquez (2008) , Getis and Ord (1992) and www.esri.com

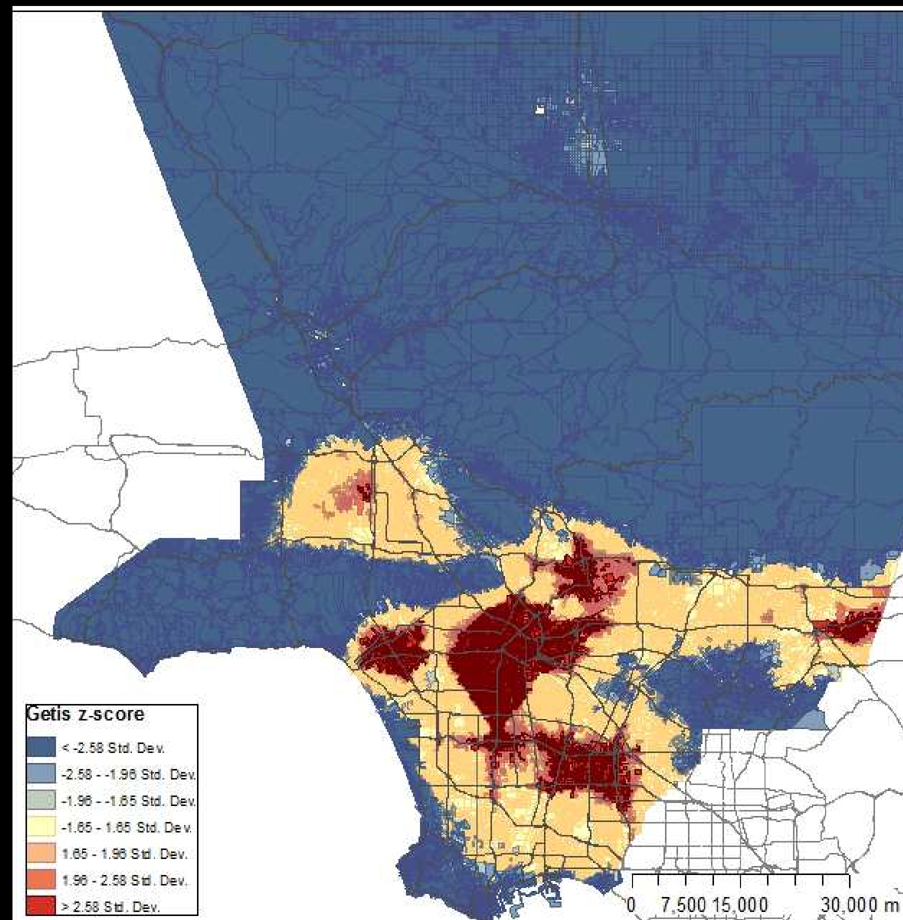
Step 2: G* Analysis

- Data input for Getis-Ord G* statistic
  - AM peak (6AM to 9AM) block level accessibility measures within 10 minutes for ***finance, retail trade, arts/entertainment, healthcare & education***
  - Average land values (price per unit area)
- Output
  - Computed z-scores for the six spatial indicators at block level
  - High opportunities and high land values are positive z-scores
  - Low opportunities and lower land values are negative z-scores
- G* clusters
  - Red indicates high value and clusters
  - Blue indicates low value and clusters
  - Yellow indicates lack of clustering

G* Clusters

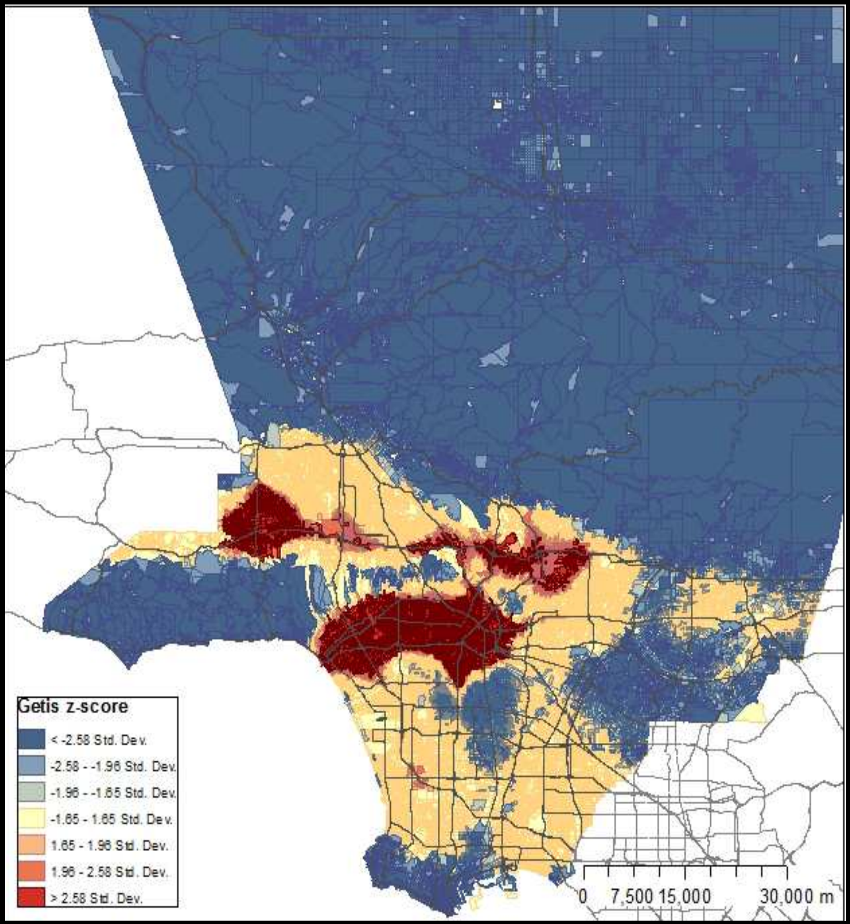


Arts/Entertainment

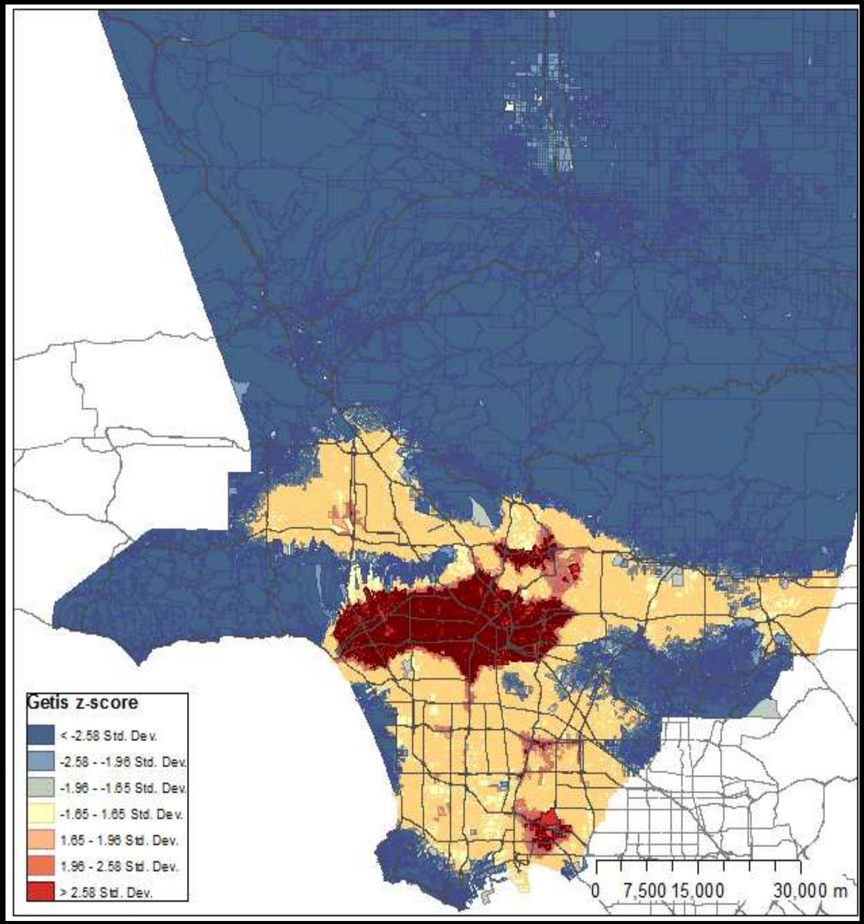


Education

G* Clusters

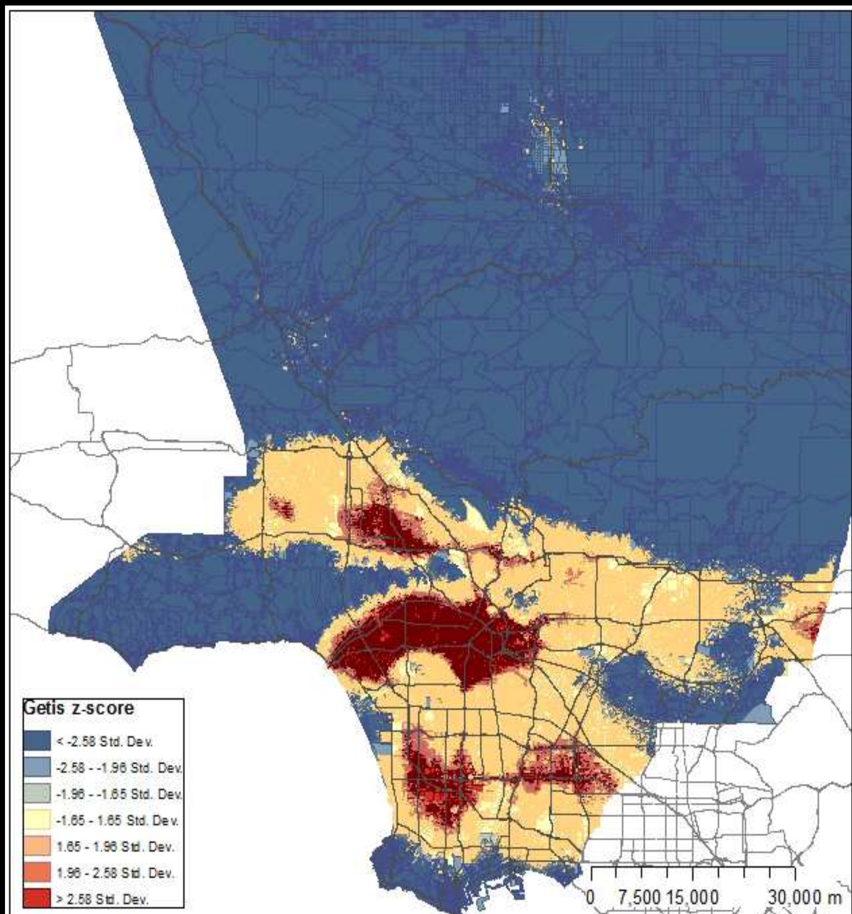


Finance

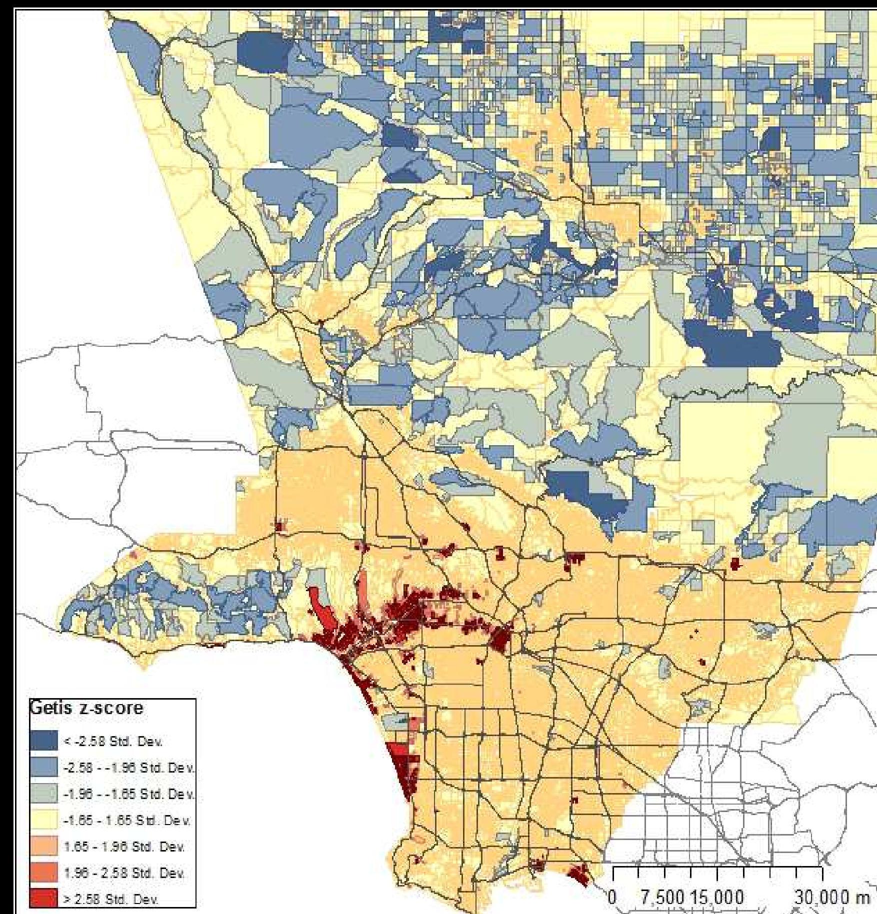


Healthcare

G* Clusters



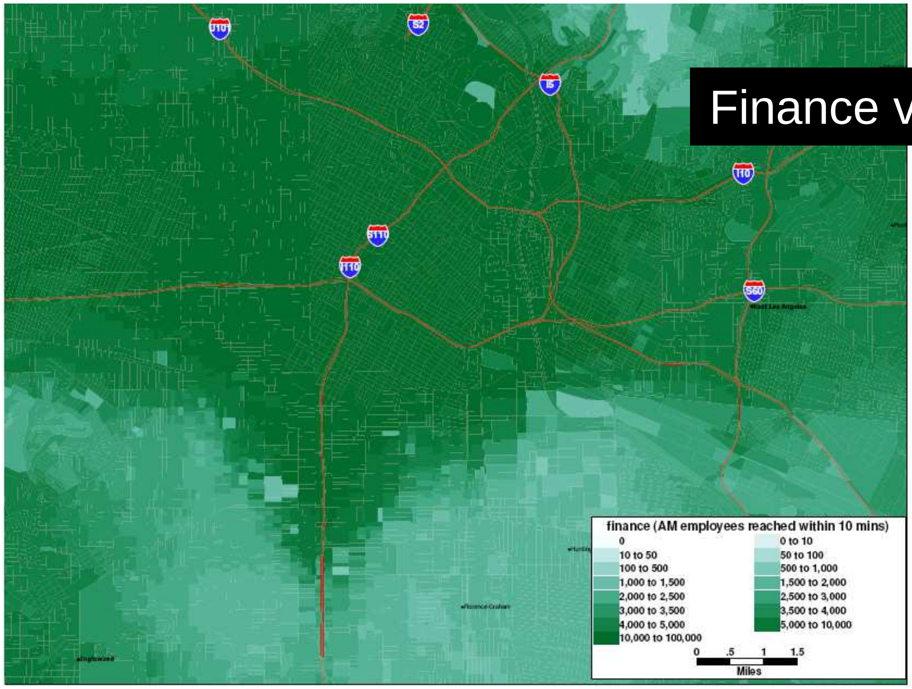
Retail



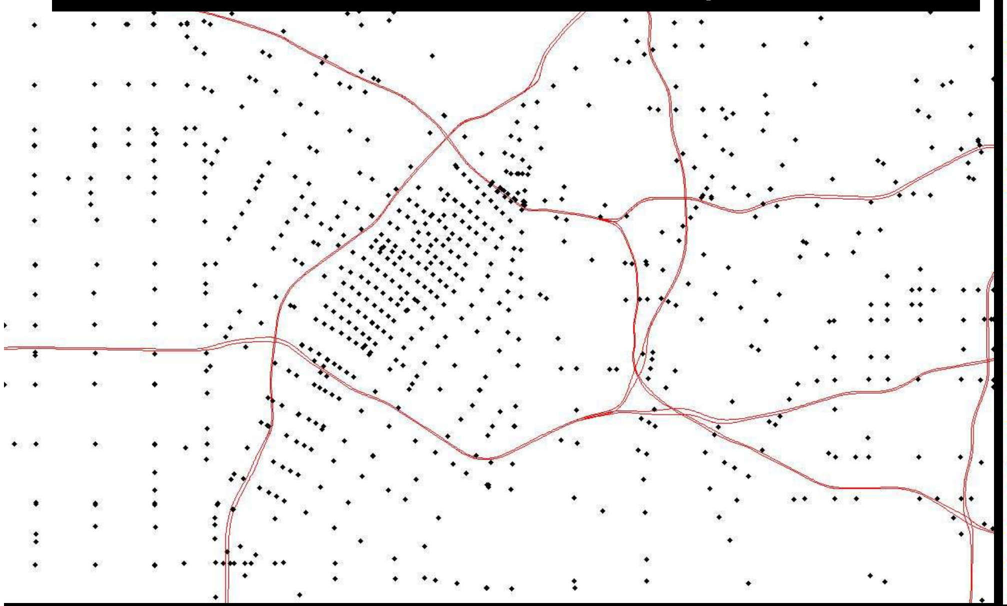
Land Value

Based on: Lei T., K. G. Goulias, and Y. Chen (2011) Opportunity-Based Dynamic Transit Accessibility in Southern California: Measurement, Findings, and Comparison with Automobile Accessibility. Paper 12-3813 presented at the January 2012 *91st Annual Meeting of the Transportation Research Board, Washington, D.C.,* January 22-26, 2012

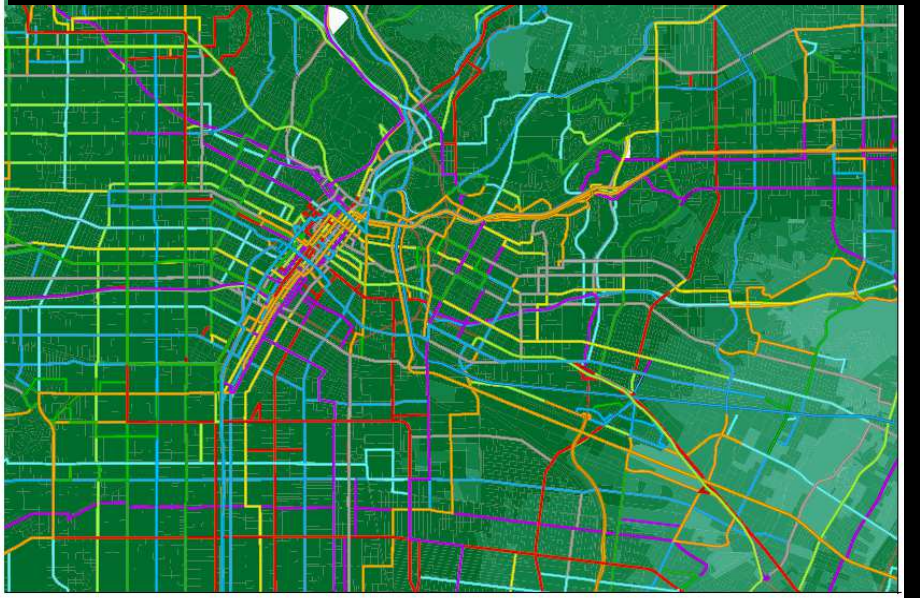
Finance vs Manufacturing

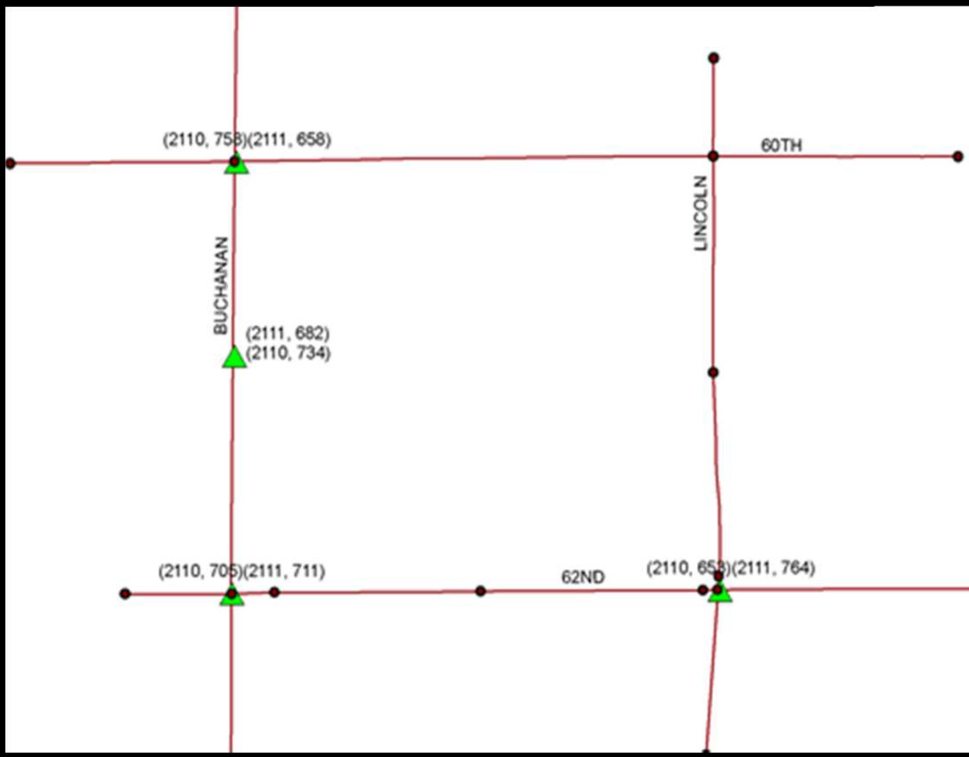


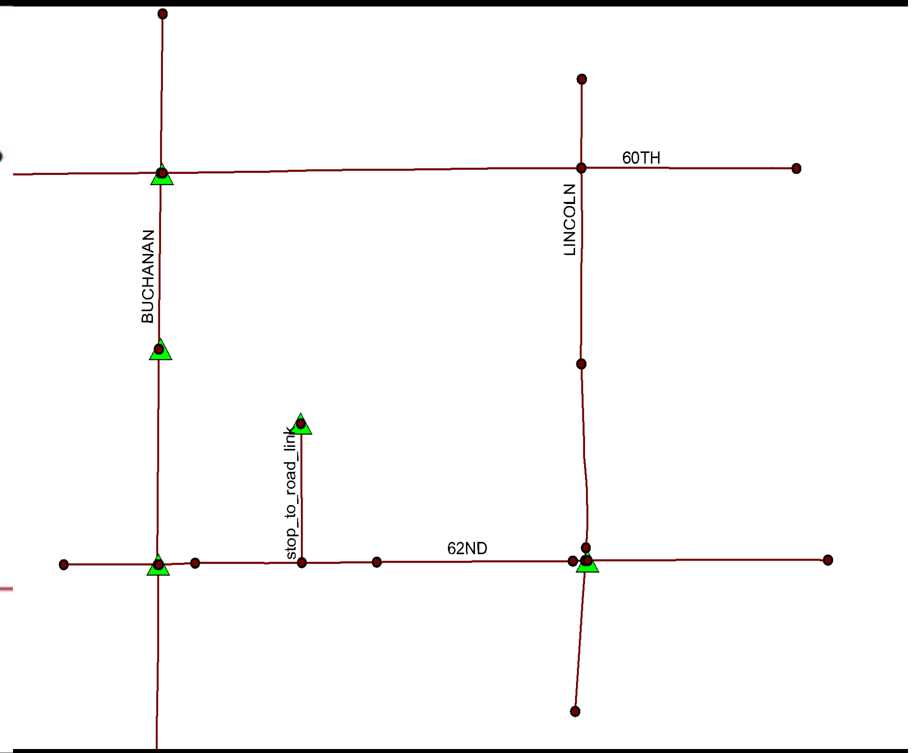
Access Points to Public Transportation

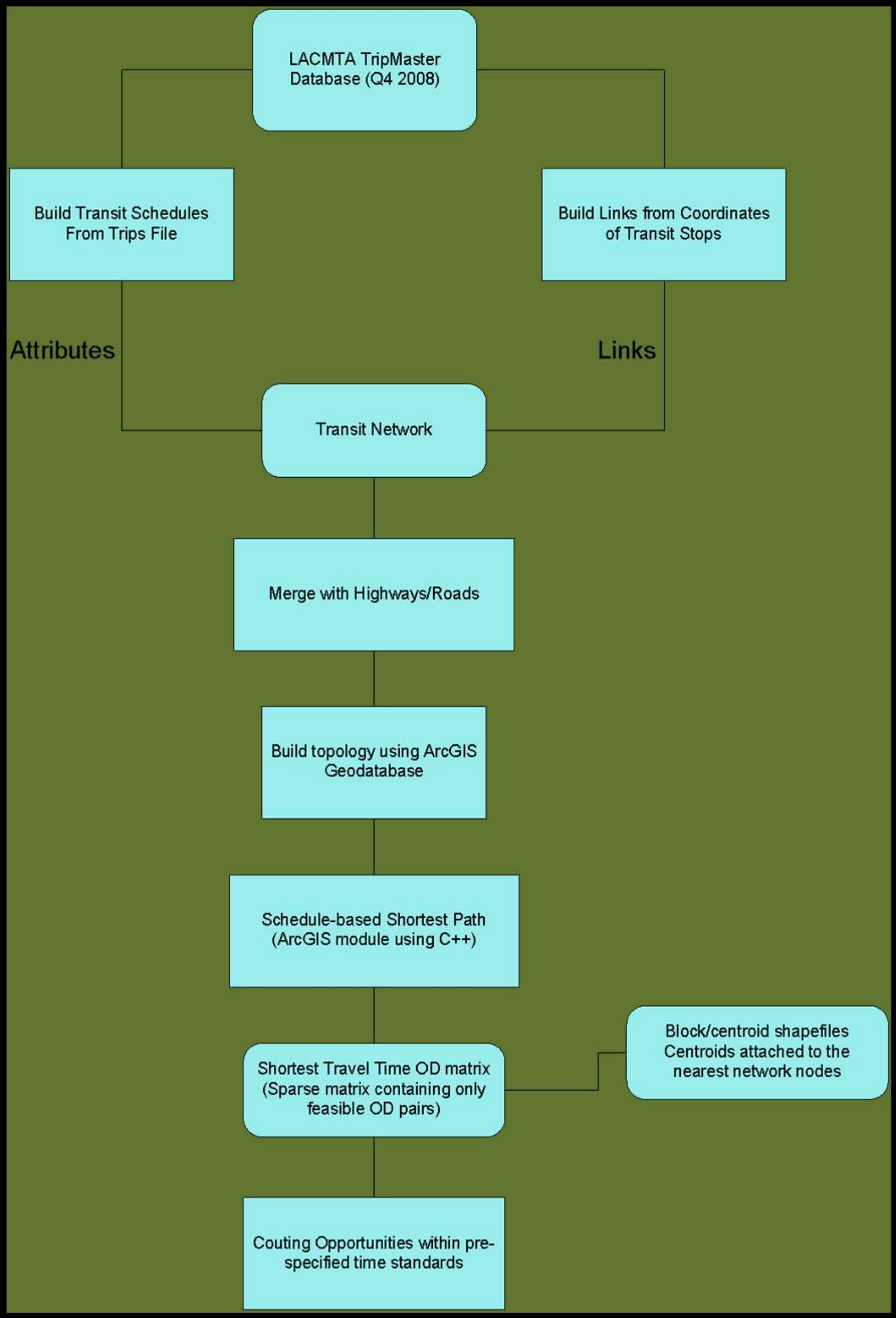


Routes of Public Transportation

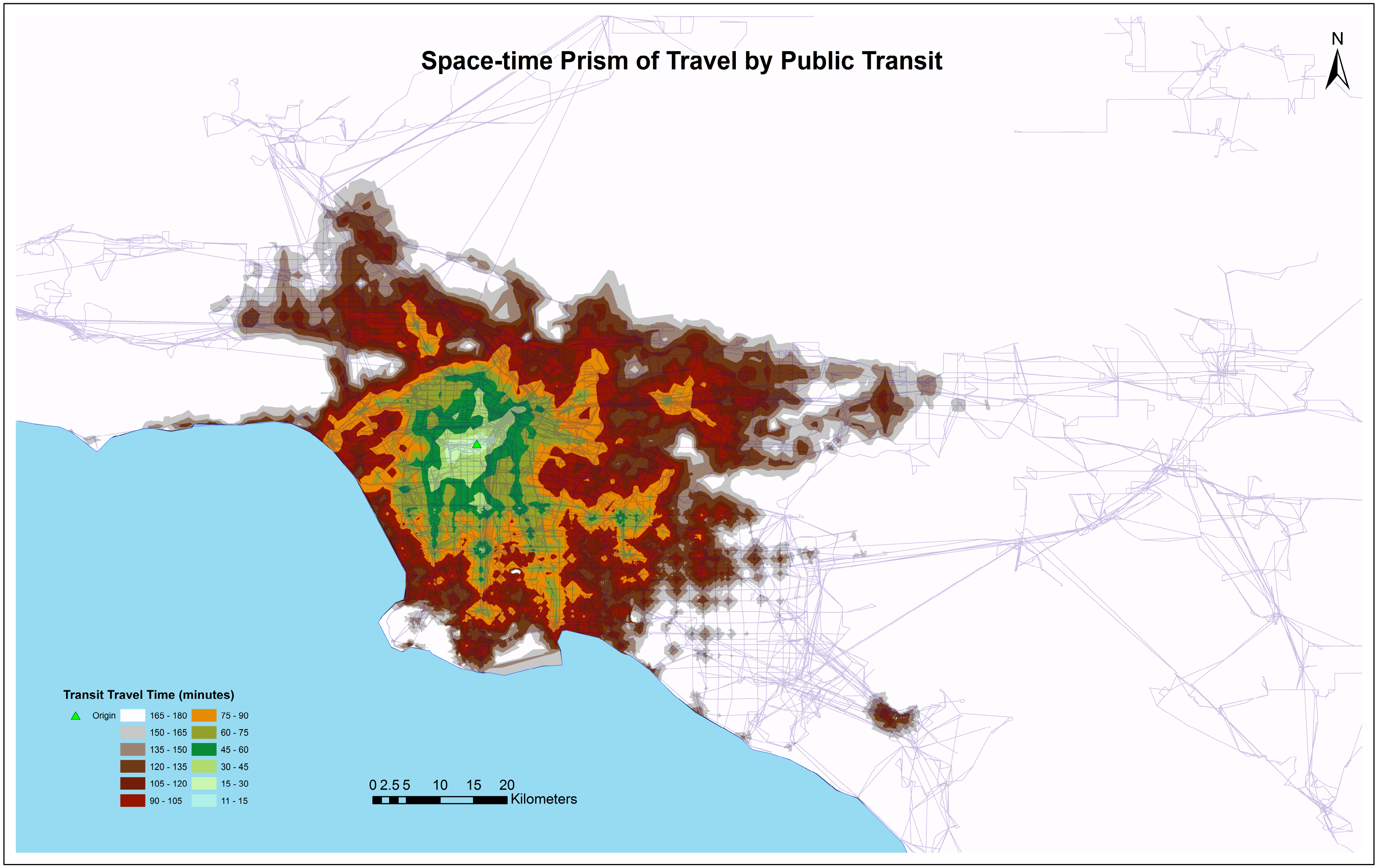



(2110, 758)(2111, 658)
60TH
LINCOLN
BUCHANAN
(2111, 682)
(2110, 734)
(2110, 705)(2111, 711)
62ND
(2110, 653)(2111, 764)


60TH
LINCOLN
BUCHANAN
stop_to_road_link
62ND

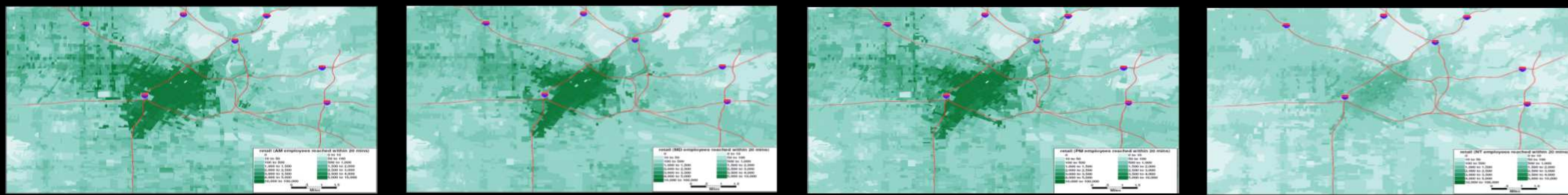


LACMTA TripMaster Database (Q4 2008)
Build Transit Schedules From Trips File
Build Links from Coordinates of Transit Stops
Attributes
Links
Transit Network
Merge with Highways/Roads
Build topology using ArcGIS Geodatabase
Schedule-based Shortest Path (ArcGIS module using C++)
Shortest Travel Time OD matrix (Sparse matrix containing only feasible OD pairs)
Block/centroid shapefiles Centroids attached to the nearest network nodes
Couting Opportunities within pre-specified time standards

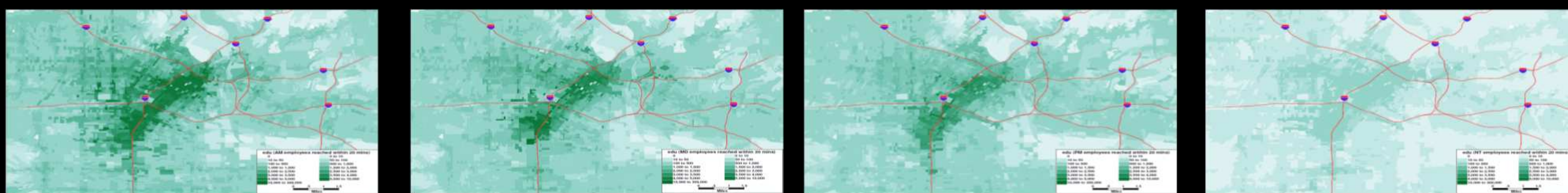


Data Source: LACMTA master trip database

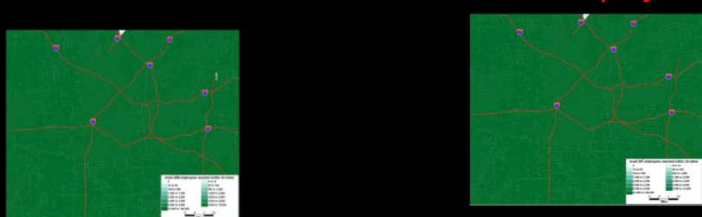
Ting Lei, Jan 2012



Maximum number of reachable retail employees for a 20-minute buffer by transit by time of day in Central Los Angeles.



Maximum number of reachable education employees for a 20-minute buffer by transit by time of day in Central Los Angeles.



Maximum number of reachable retail employees for a 20-minute buffer by time of day in Central Los Angeles by Private Car

Maximum number of reachable education employees for a 20-minute buffer by time of day in Central Los Angeles by Private Car

- Moving to finer resolutions (more TAZs or even points) increases ability to represent built environment but creates the need to maintain databases over time
- There are some interesting issues about spatial allocation – for now we used count data regression models
- Assign synthetic households to blocks but also assign parcels/housing units to synthetic households
- Start the path toward a demographic microsimulator for at least 25 years and identify methods for fine spatial resolution
- Interface with land use model can offer added information but the space resolution will always be an issue to consider
- Interface with the more recent “sustainability tool(s)” that use GIS maps of parcels open a new gateway to possibilities

Thank You

&

Questions?