Built Environment and Health in the Age of Big Data

Anne Vernez Moudon, Dr es Sc
Urban Form Lab
University of Washington, Seattle, WA, USA
http://depts.washington.edu/ufl/

This material is extracted from a keynote made at the European Colloquium on Theoretical and Quantitative Geography on 10 September 2017, in York, UK; and a presentation made at the Leeds institute of Data Analytics, 21 September 2017, University of Leeds, UK

Presentation

1. City and transport planning can/should support better health, reduce the incidence of non-communicable diseases, specifically obesity, and help achieve energy balance
2. Research on behavior and built environment (BE)
3. Causal path to obesity: BE influences through walkability and walking
4. Exposure to BE: Theory and measures
5. Spatial modeling of BE and walking/transit use behavior: An example

The Lancet Series September 2016

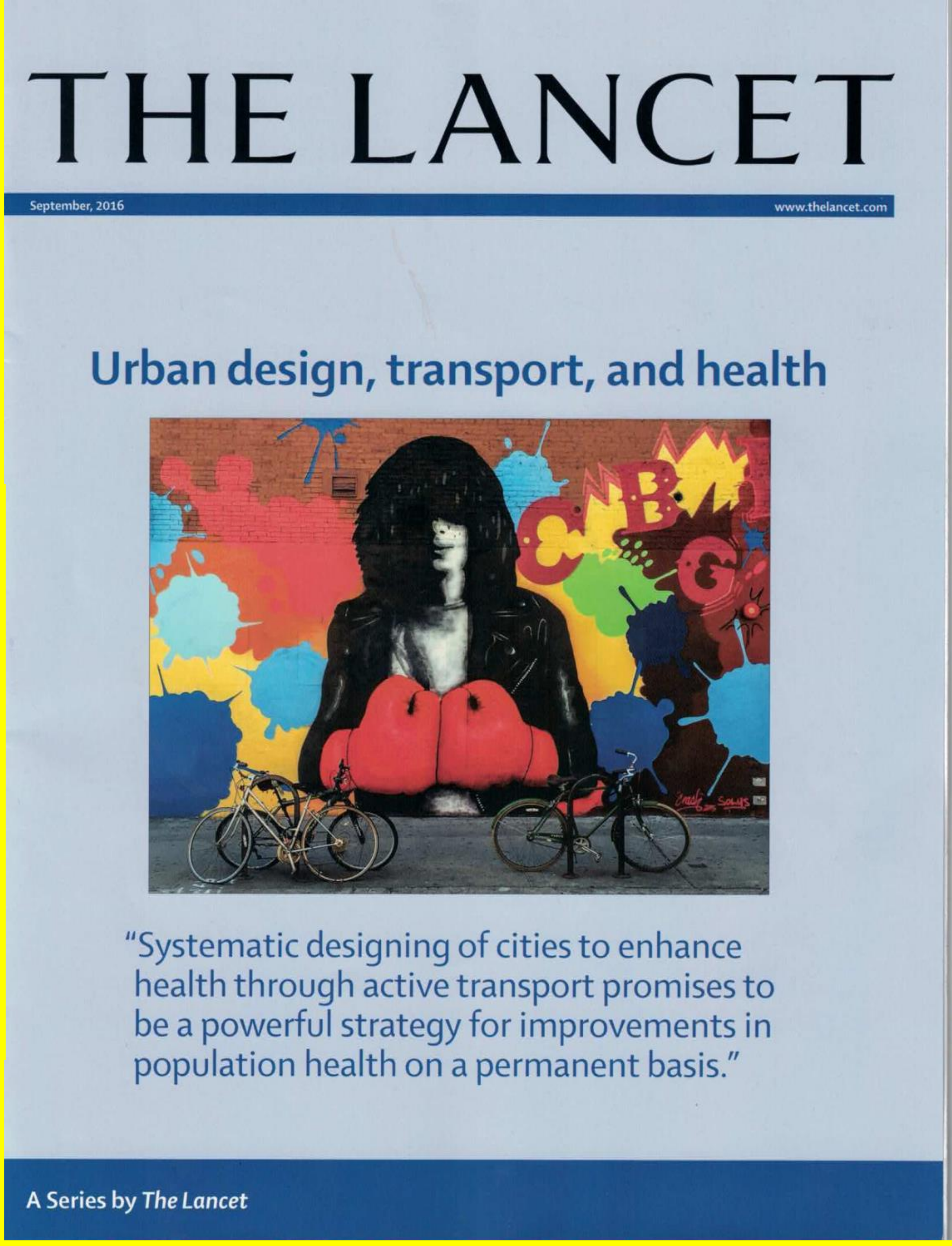
1. HOW CITY DESIGN AND TRANSPORT PLANNING CAN IMPROVE HEALTH

City planning can help reduce air pollution, road trauma, non-communicable diseases (cardiovascular disease and diabetes), and physical inactivity.

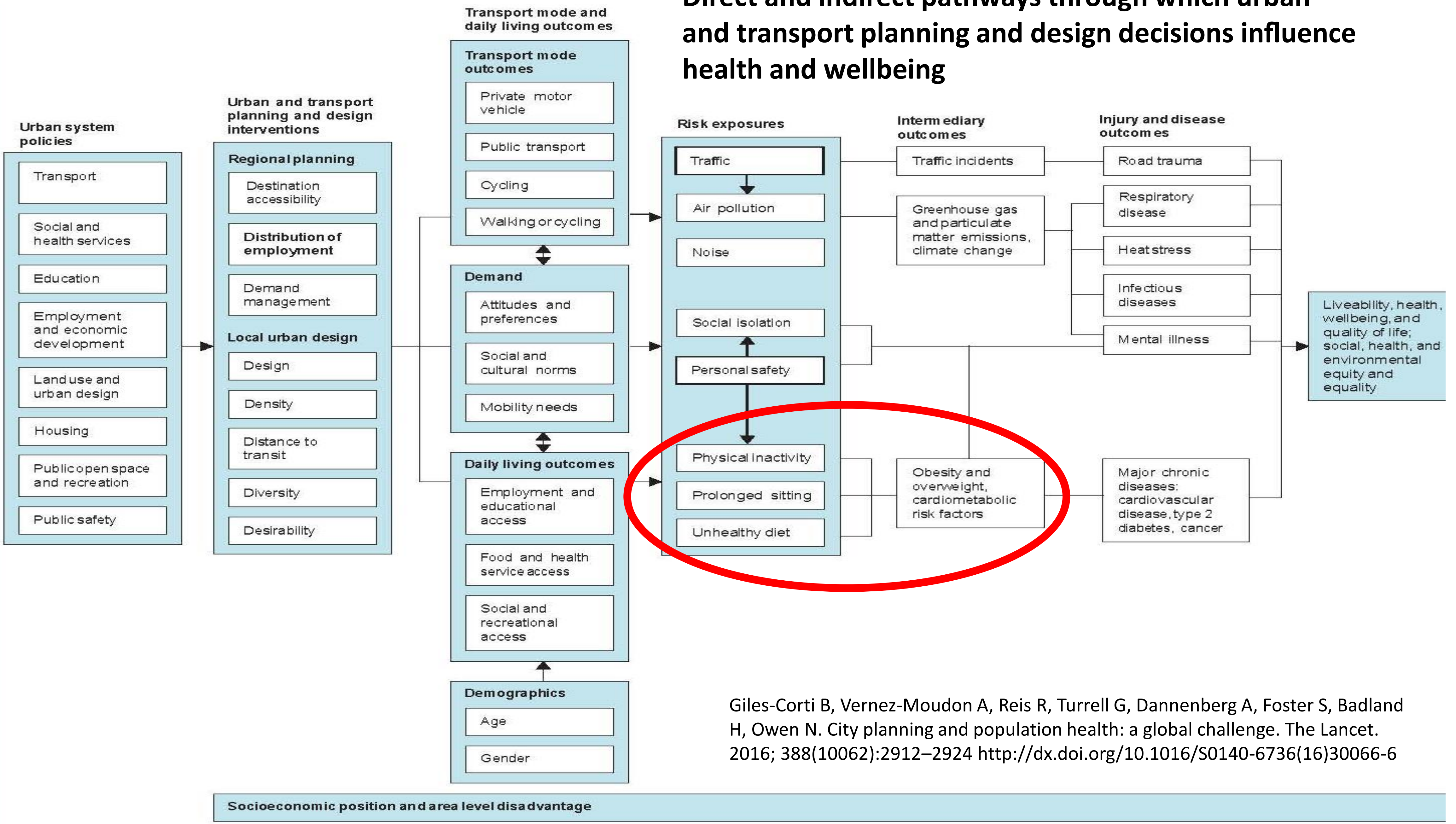
For better health, cities need to

- incentivize a shift from private car use to cycling and walking
- adopt a compact city model where distances to shops and facilities, including public transport, are shorter and within walking distance.

http://www.thelancet.com/series/urban-design



Direct and indirect pathways through which urban and transport planning and design decisions influence health and wellbeing



Giles-Corti B, Vernez-Moudon A, Reis R, Turrell G, Dannenberg A, Foster S, Badland H, Owen N. City planning and population health: a global challenge. The Lancet. 2016; 388(10062):2912–2924 http://dx.doi.org/10.1016/S0140-6736(16)30066-6

2. Behavior – built environment

Data and research

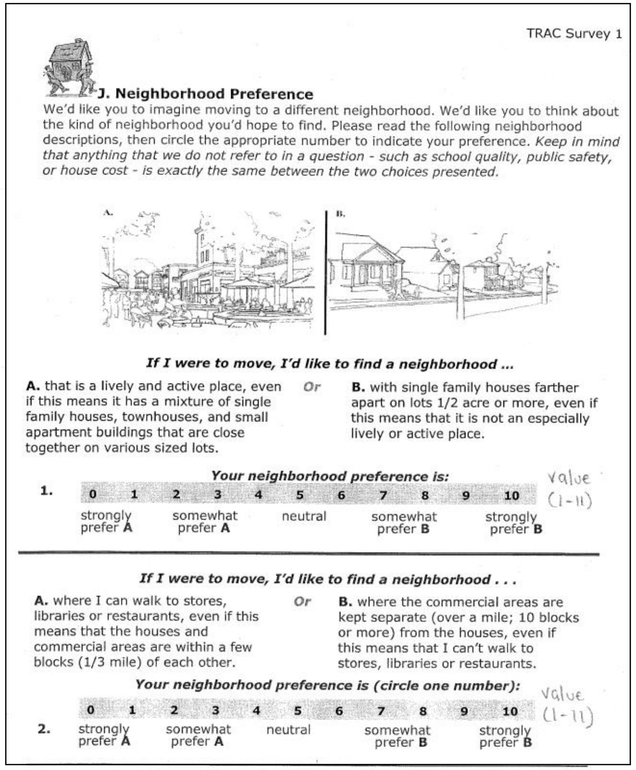
Behavior: Primary data on mobility and activity patterns

Four instruments for 7-day assessment; administered 3 times in 5 years

TRAC and Action projects,
PI BE Saelens

GPS

Accelerometer



TRAC Survey 1

J. Neighborhood Preference

We'd like you to imagine moving to a different neighborhood. We'd like you to think about the kind of neighborhood you'd hope to find. Please read the following neighborhood descriptions, then circle the appropriate number to indicate your preference. *Keep in mind that anything that we do not refer to in a question - such as school quality, public safety, or house cost - is exactly the same between the two choices presented.*

A. B.

If I were to move, I'd like to find a neighborhood ...

A. that is a lively and active place, even if this means it has a mixture of single family houses, townhouses, and small apartment buildings that are close together on various sized lots. *Or* **B.** with single family houses farther apart on lots 1/2 acre or more, even if this means that it is not an especially lively or active place.

Your neighborhood preference is: value (1-11)

	0	1	2	3	4	5	6	7	8	9	10
1.	strongly prefer A		somewhat prefer A			neutral		somewhat prefer B			strongly prefer B

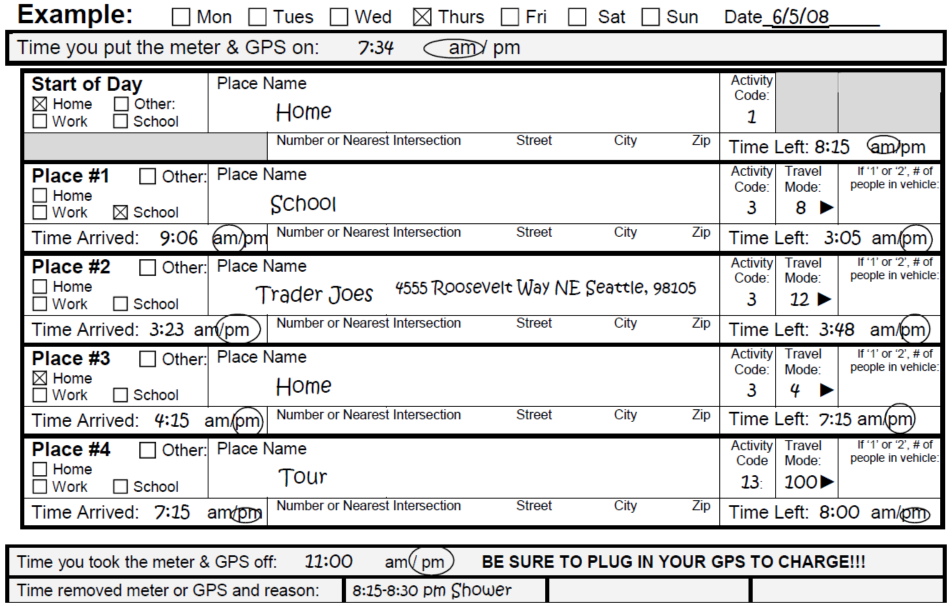
If I were to move, I'd like to find a neighborhood . . .

A. where I can walk to stores, libraries or restaurants, even if this means that the houses and commercial areas are within a few blocks (1/3 mile) of each other. *Or* **B.** where the commercial areas are kept separate (over a mile; 10 blocks or more) from the houses, even if this means that I can't walk to stores, libraries or restaurants.

Your neighborhood preference is (circle one number): value (1-11)

	0	1	2	3	4	5	6	7	8	9	10
2.	strongly prefer A		somewhat prefer A			neutral		somewhat prefer B			strongly prefer B

Survey



Example: ☐ Mon ☐ Tues ☐ Wed ☒ Thurs ☐ Fri ☐ Sat ☐ Sun Date 6/5/08

Time you put the meter & GPS on: 7:34 (am) / pm

Place	Place Name / Number or Nearest Intersection, Street, City, Zip	Activity Code	Travel Mode	If '1' or '2', # of people in vehicle	Time Left
Start of Day ☒ Home ☐ Work ☐ Other: ☐ School	Home	1			8:15 (am)/pm
Place #1 ☐ Home ☐ Work ☐ Other: ☒ School; Time Arrived: 9:06 (am)/pm	School	3	8 ▶		3:05 am/(pm)
Place #2 ☐ Home ☐ Work ☐ Other: ☐ School; Time Arrived: 3:23 am/(pm)	Trader Joes 4555 Roosevelt Way NE Seattle, 98105	3	12 ▶		3:48 am/(pm)
Place #3 ☒ Home ☐ Work ☐ Other: ☐ School; Time Arrived: 4:15 am/(pm)	Home	3	4 ▶		7:15 am/(pm)
Place #4 ☐ Home ☐ Work ☐ Other: ☐ School; Time Arrived: 7:15 am/(pm)	Tour	13	100 ▶		8:00 am/(pm)

Time you took the meter & GPS off: 11:00 am/(pm) **BE SURE TO PLUG IN YOUR GPS TO CHARGE!!!**

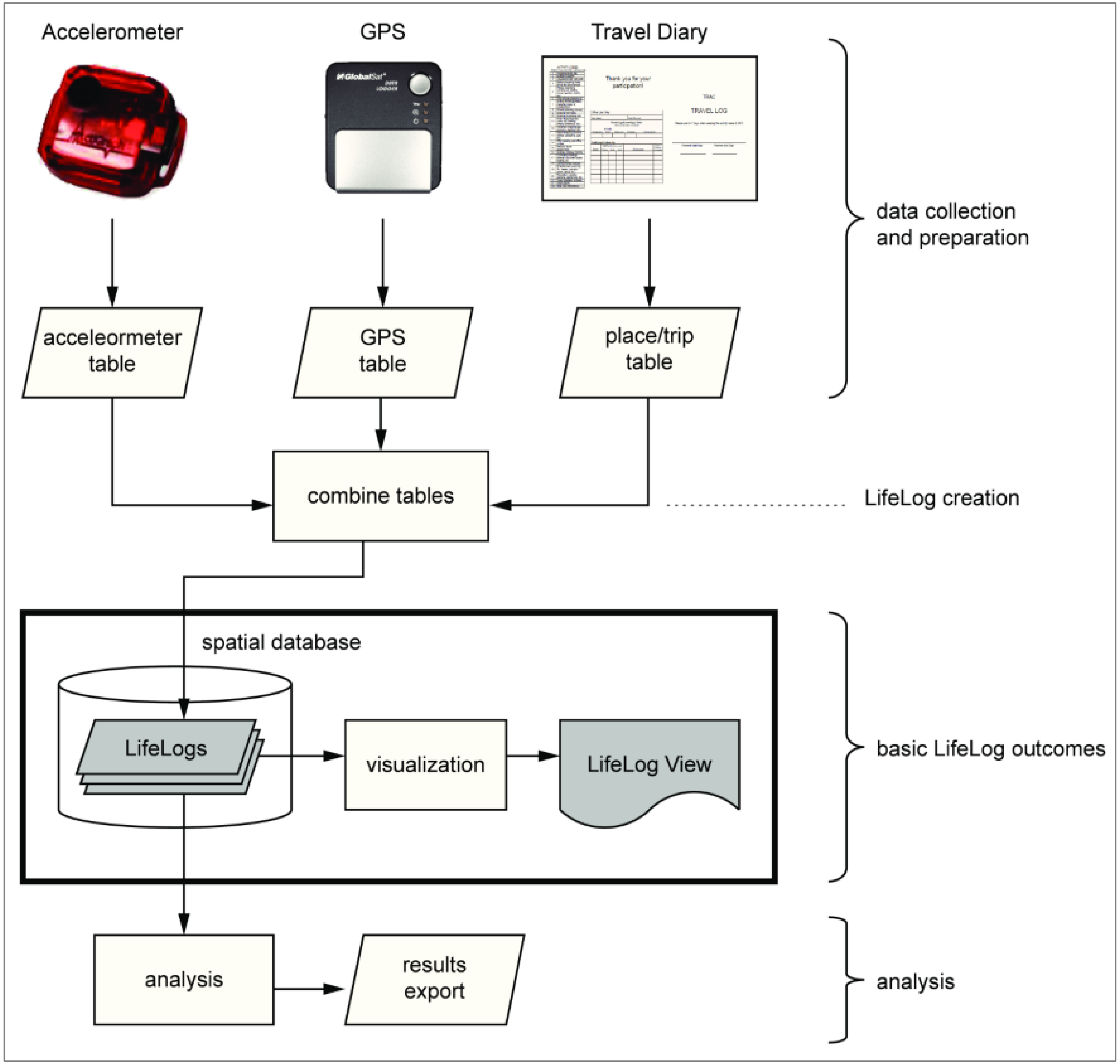
Time removed meter or GPS and reason: 8:15-8:30 pm Shower

Travel Log

Behavior data integration

- Collation of multiple data streams into a single "LifeLog" table containing all source data
- Data joined by common time stamps across tables

Hurvitz PM, Moudon AV, Kang B, Saelens BE, Duncan GE. Emerging technologies for assessing physical activity behaviors in space and time. Front Public Health. 2014 2, 2. PMID: 24479113

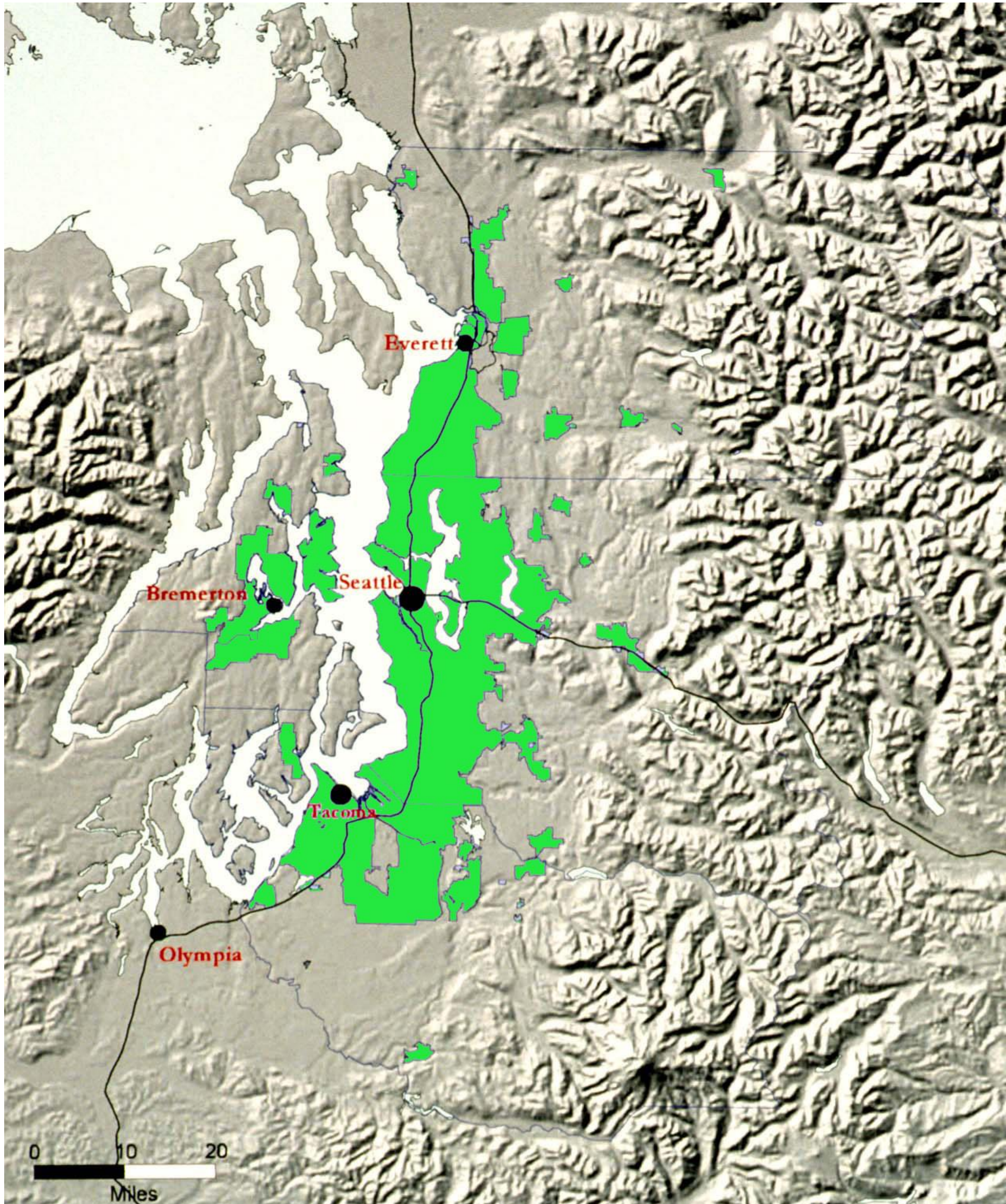


LIFELOG contains "Big Data" on activity and mobility patterns

- Number of records per subject per week
  - GPS: ~33,000 x, y points
  - Accelerometry: ~51,000 counts
  - Travel diary: ~43 places and trips

Built environment data:
Spatial Extent and data unit

Puget Sound Seattle Region:
1,000 mi^2 (2500 km^2)
1,200,000 tax parcels



Built environment data

by domain and variable of interest

Domains	Variables*	Number of discrete observations in King Co
Neighborhood environment	residential units (houses, apartments, condos, mobile homes)	489k parcels
	employment/jobs	21k parcels
	residential wealth (property values)	489k parcels
	vacant lands	51k parcels
Routine destinations	food facilities	1,500 food stores 6,500 restaurants
	physical activity and fitness facilities	880 parcels
	retail services	5652 parcels
	schools and educational facilities	737 parcels
	offices	4393 parcels
	medical offices	769 parcels
	public services (libraries, etc)	
	open space and parks	1541 parks
	facilities in parks	103 types of facilities per park

* measures typically include counts, densities, and distances between features of interest. Both airline and network measures are calculated; UFL = data already has been collected by the UW Urban Form Lab

Built environment data

by domain and variable of interest

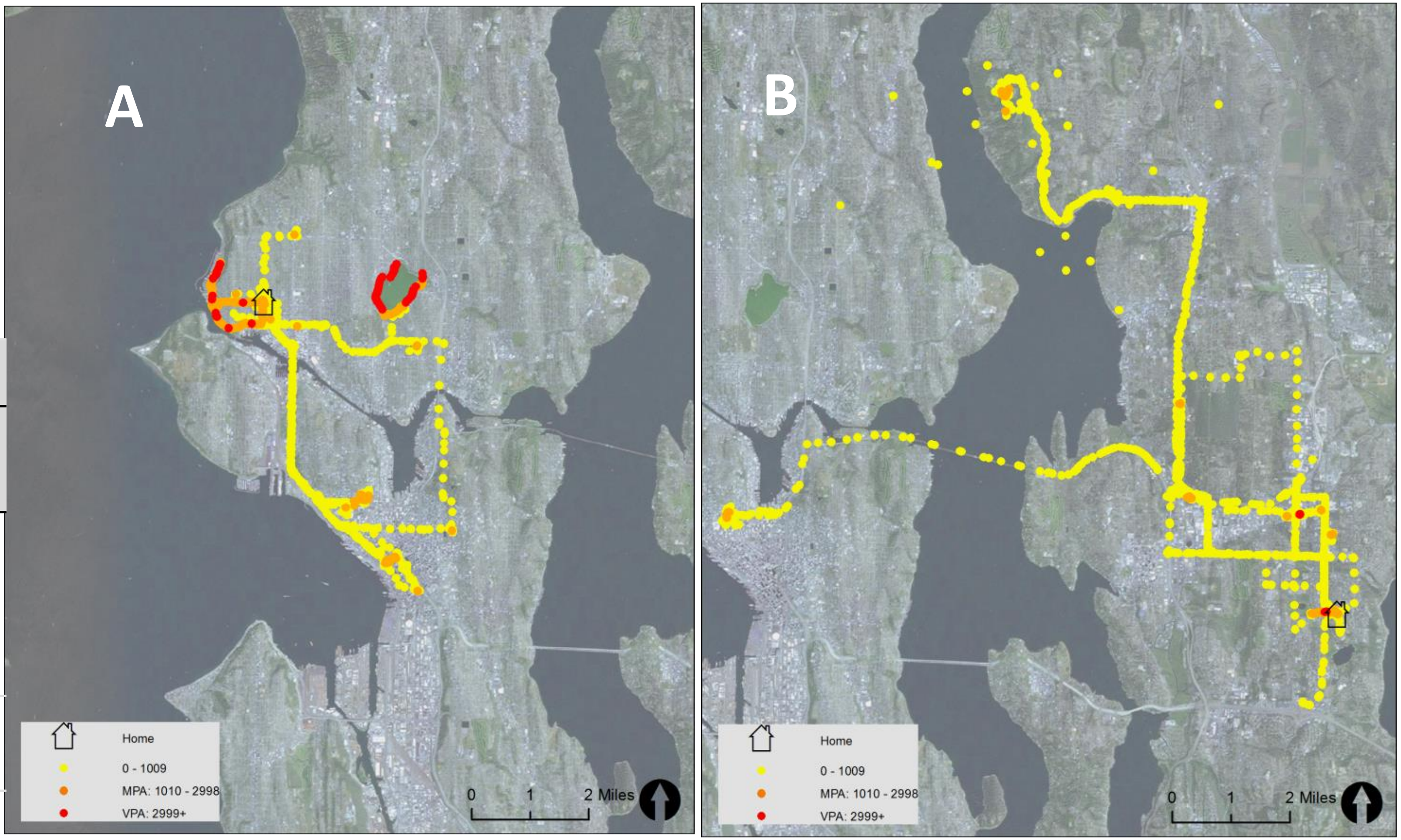
Domains	Variables*	Number of discrete observations in King Co
Transportation infrastructure	street (freeways and expressways, arterials, collector and local streets)	14k linear miles
	intersection density	64k intersections
	trails	829 linear miles
	sidewalks	1708 linear miles
	traffic signals	2000 signals
	parking	2.3 million stalls
	passenger rail stations	17 stations
	bus stops	8635 stops
Traffic conditions	vehicular volumes	86 million daily vehicle-miles
	bus ridership	364k daily trips
	pedestrian/bike collisions	1150 annual collisions

* measures typically include counts, densities, and distances between features of interest. Both airline and network measures are calculated; UFL = data already has been collected by the UW Urban Form Lab

Behavior – built environment data

Mobility and activity patterns

Minutes per day	A		B	
	Accelero-meter	GPS	Accelero-meter	GPS
Total observed	1234	262	1260	741
Total MPA	38	14	24	16
Total VPA	5	4	2	2
Total driving (travel diary)	27		70	



From Troiano et al.(2008). Physical activity in the United States measured by accelerometer. Med & Sci in Sports & Exercise. MPA : accelerometer count >= 1010 per 30-sec epoch; VPA: accelerometer count >= 2999 per 30-sec epoch

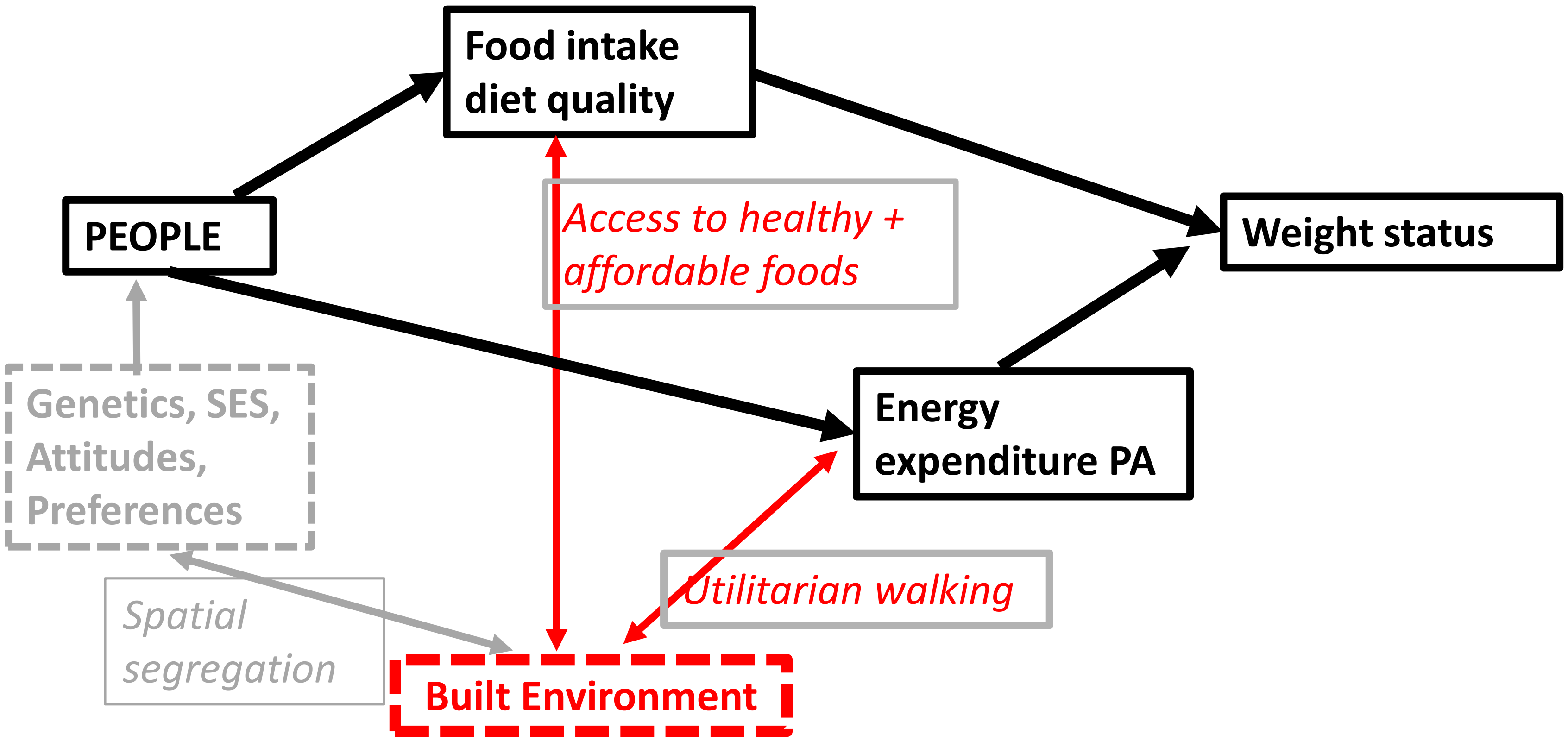
Energy intake

Energy expenditure

3. Built environment influence on behavior

Causal paths to energy balance

BE and causal paths to energy balance



Horn EE, Turkheimer E, Strachan E, Duncan GE. Behavioral and Environmental Modification of the Genetic Influence on Body Mass Index: A Twin Study. Behav Genet. 2015. doi: 10.1007/s10519-015-9718-6. PubMed PMID: 25894925.

Duncan GE, Cash SW, Horn EE, Turkheimer E. Quasi-causal associations of physical activity and neighborhood walkability with body mass index: A twin study. Prev Med. 2015;70C:90-5. doi: 10.1016/j.ypmed.2014.11.024. PubMed PMID: 25482422.

MVPA, walking, walkability, and BMI

Results from twin studies

6376 same-sex adult twin pairs within pair analyses

- Walking and MVPA associated with BMI in phenotypic analyses; associations attenuated but significant in biometric analyses (*Ps* < 0.05).
- Walkability associated with walking (but not with MVPA or with BMI) in both phenotypic and biometric analyses (*Ps* < 0.05), with no attenuation accounting for shared genetic and environmental background.
- *Higher neighborhood walkability is (quasi) causally associated with increased neighborhood walking levels, and, in turn, higher neighborhood walking levels are (quasi) causally associated with reduced BMI*.

Duncan GE, Cash SW, Horn EE, Turkheimer E. Quasi-causal associations of physical activity and neighborhood walkability with body mass index: A twin study. Prev Med. 2015;70C:90-5. doi: 10.1016/j.ypmed.2014.11.024. PubMed PMID: 25482422.

MVPA, walkability and BMI

Results from twin studies

5079 same-sex adult twin pairs

- High levels of MVPA suppressed genetic risk for high BMI, controlling for underlying genetic etiology shared between PA and BMI.
- Neighborhood walkability also had moderating effects on genetic variance in BMI; however, these effects were mediated by MVPA.
- *Interventions focusing on PA, including those that improve aspects of the BE that in turn promote more PA, could help reduce obesity*.

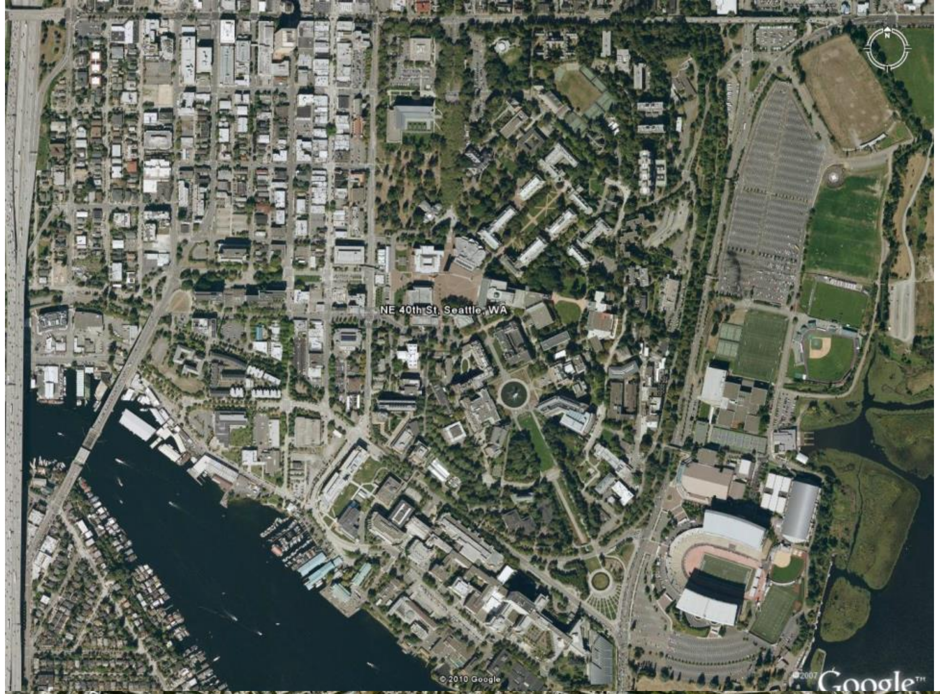
Horn EE, Turkheimer E, Strachan E, Duncan GE. Behavioral and Environmental Modification of the Genetic Influence on Body Mass Index: A Twin Study. Behav Genet. 2015. doi: 10.1007/s10519-015-9718-6. PubMed PMID: 25894925.

Walking - **physical activity**

↕

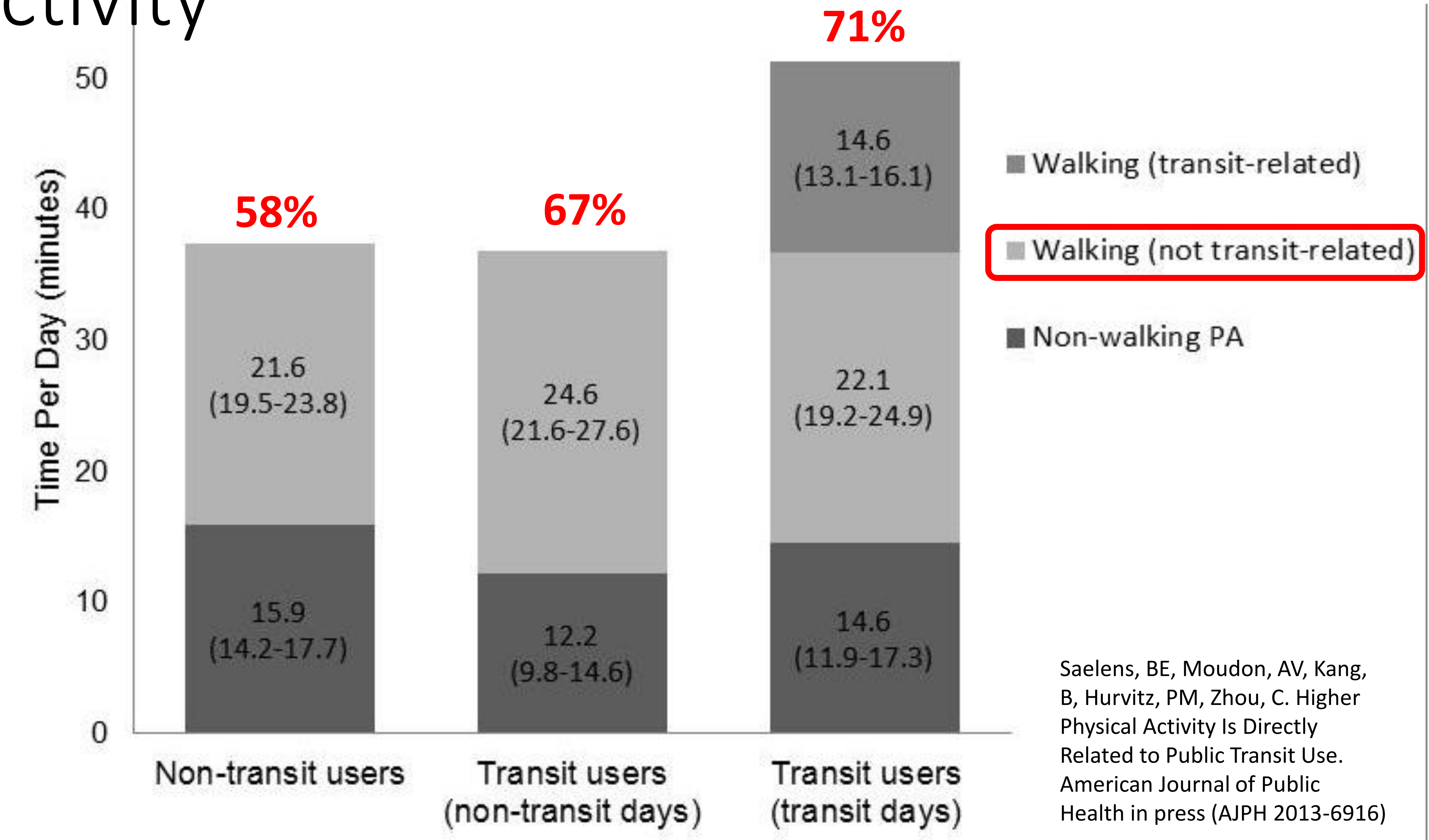
Transit use - **built environment**

- Walking = population-wide preferred mode of physical activity or exercise
- Transit use leads to more walking
- Transit infrastructure systems coexist with higher development densities, higher mix of uses (allowing routine active trips between origins and destinations)



Effect of transit use on walking share of physical activity

- N = 693 participants
- 4432 person-day
- PA = >1000 cpm



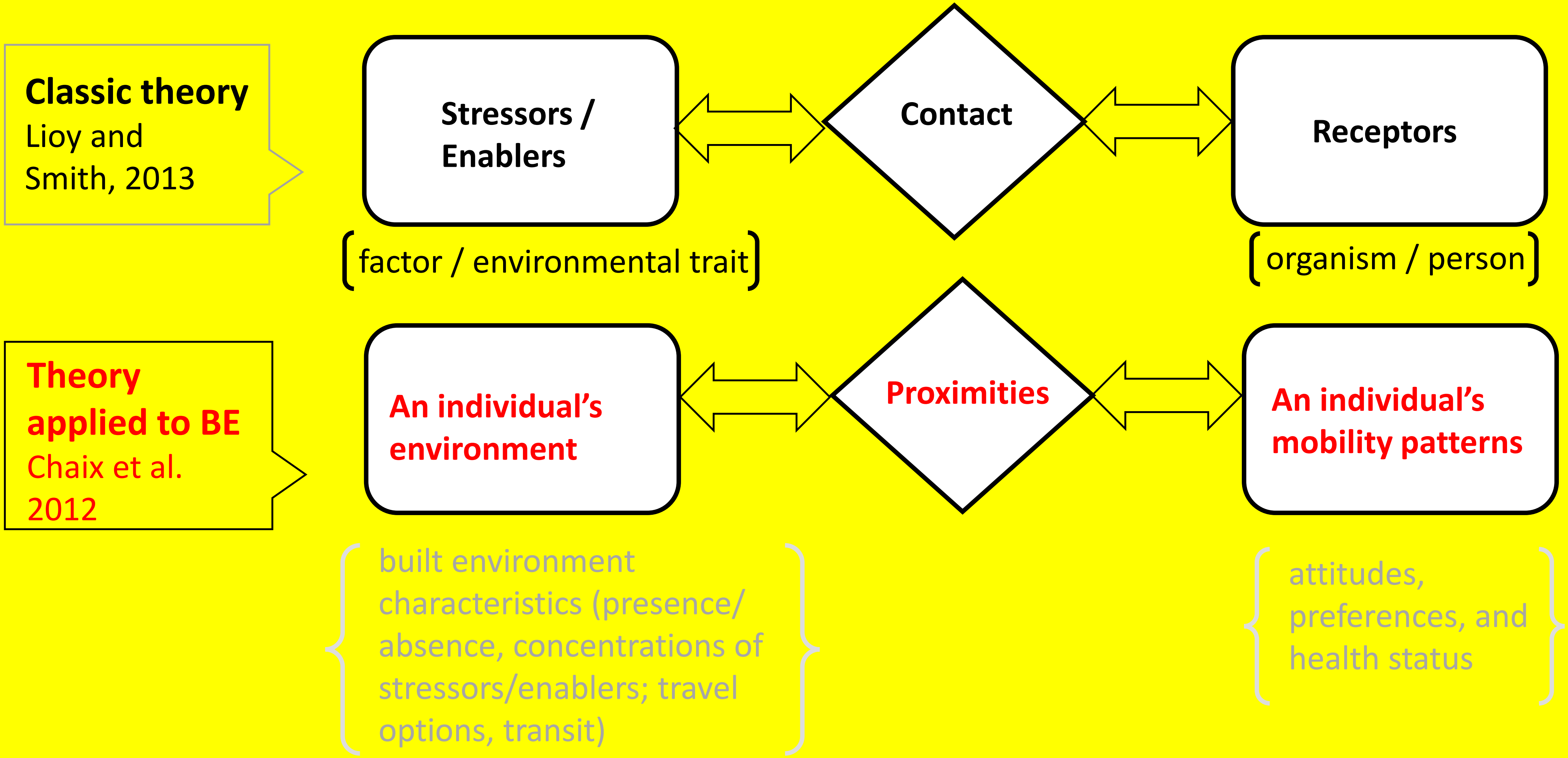
Saelens, BE, Moudon, AV, Kang, B, Hurvitz, PM, Zhou, C. Higher Physical Activity Is Directly Related to Public Transit Use. American Journal of Public Health in press (AJPH 2013-6916)

Untangling the effects of BE on health Methodological challenges

- Participants cannot be randomly assigned to built environments. Causality can be tested through:
  - Longitudinal studies
  - Natural experiments
  - Twin studies
- Unknown direction of causality in cross-sectional studies of behavior - BE
- Self-selection of place of residence and places visited
  - spatial segregation by socioeconomic status
  - selective daily mobility bias
- Poor measures of health behaviors (esp. food intake) and BE
- Poor theory of exposure to BE

4. Theory from exposure science

Lioy and Smith, 2013 + Chaix et al. 2012



Environment affecting energy expenditure

Environment affecting energy intake

BE stressors/enablers

Attributes of BE stressors/enablers

- **ENERGY EXPENDITURE**

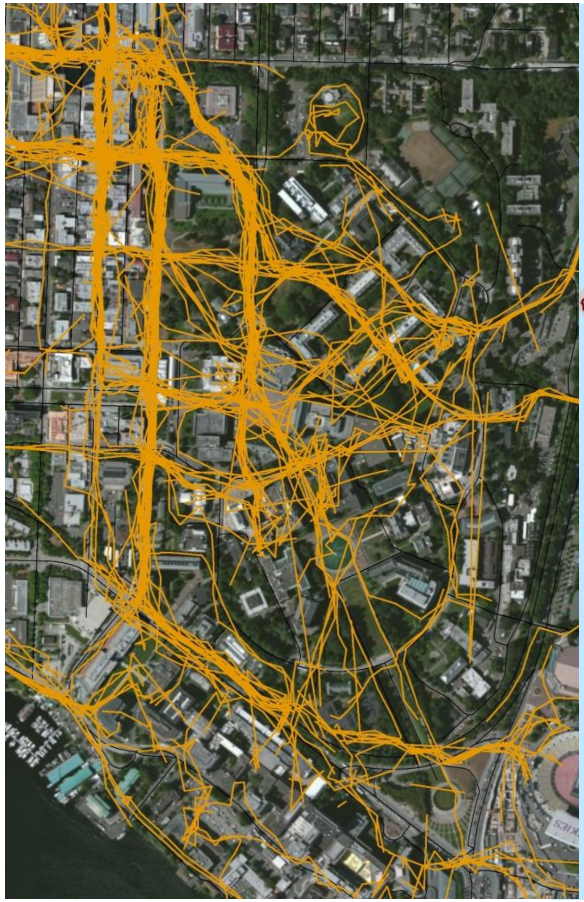
- **Walking (utilitarian and recreation) environment**
  - Development density
  - Nearby destinations for daily living
  - Short street-blocks
  - Sidewalks/ crosswalks
  - Public transport
  - Parks and recreation areas (waterways)
  - Trails

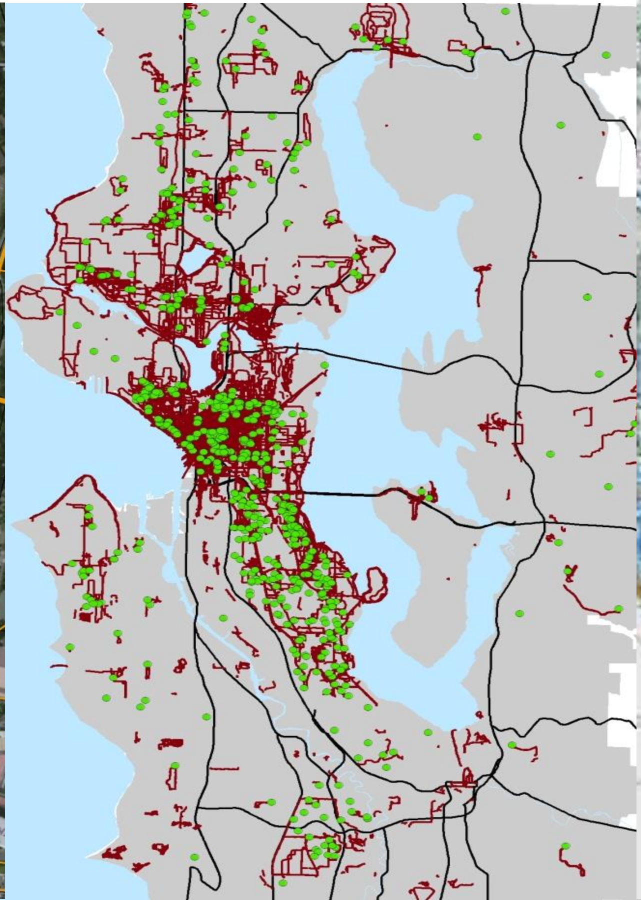
- **ENERGY INTAKE**

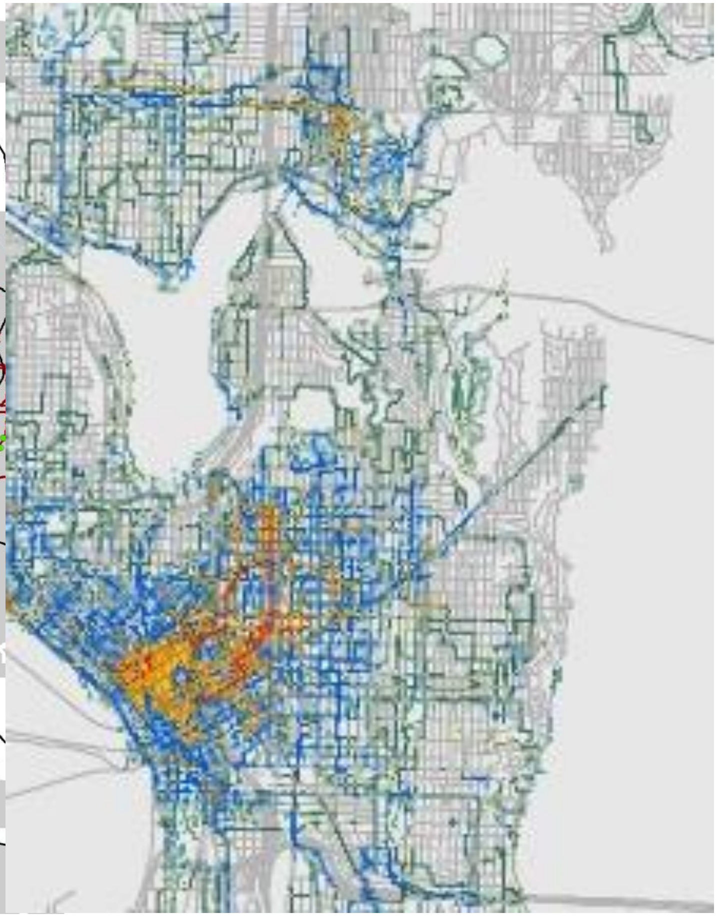
- **Food environment**
  - Healthy and unhealthy food venues
    - Calories versus nutrients
    - Affordability

Geographic context of BE stressors/ enablers

- The spatial distribution/concentration of food establishments is based on market place assumptions of the potential **influence of exposure on use**
    - Numerous convenience stores and FFRs (low-cost, impulse buying),
    - Fewer supermarkets (more medium costs SM than low or high cost SM) (purpose/considered buying)
    - Few trails, sports fields, short street-blocks, etc.
- GLOBAL TREND since 1960s, most BEs contain MORE stressors and FEWER enablers of PA and healthy diet







Receptors

Mobility patterns, neighborhood and places visited

Geographic context of BE receptors

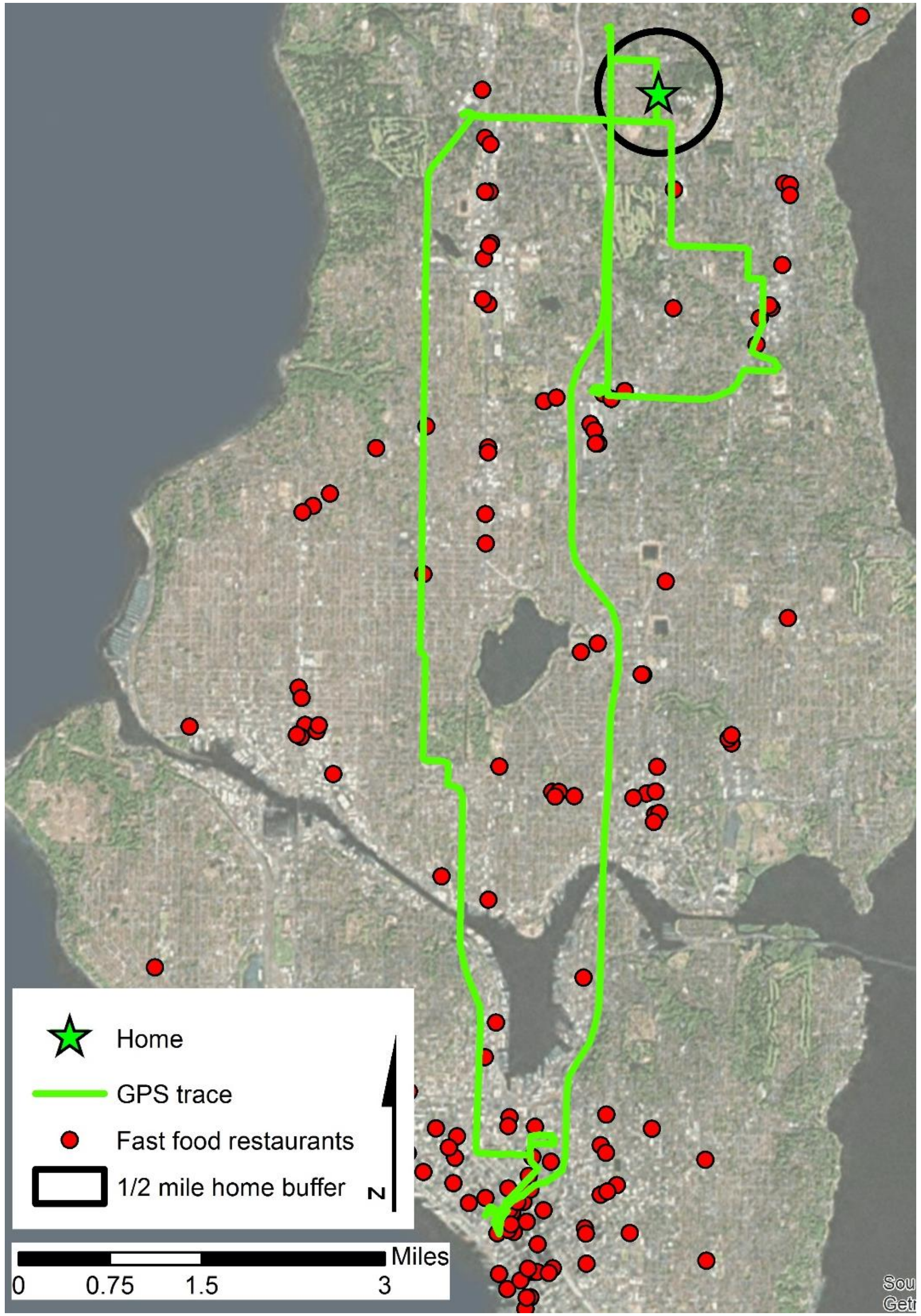
- **Static/dwell areas and places**
  - Residential or work neighborhood
  - Other non-residential and/or non-work
- **Dynamic/move locations**
  - Places visited
  - Travel routes (from travel diaries or GPS traces)

- Uncertain geographical context
  - James et al. in Effects of buffer size and shape on associations between the built environment and energy balance. Health Place. 2014 May;27:162-70. doi:10.1016/j.healthplace.2014.02.003. Epub 2014 Mar 7.
- Spatial polygamy
  - Matthews S, Yang TC. Spatial Polygamy and Contextual Exposures (SPACEs): Promoting Activity Space Approaches in Research on Place and Health. Am Behav Sci. 2013 Aug 1;57(8):1057-1081.
- Moving away from place-based to people-based measures of exposure
  - Kwan MP. From place-based to people-based exposure measures. Soc Sci Med. 2009 Nov;69(9):1311-3. doi: 10.1016/j.socscimed.2009.07.013. Epub 2009 Aug 7
  - Kestens et al. Using experienced activity spaces to measure foodscape exposure. Health & Place 16 (2010) 1094–1103

Example:
Exposure to fast food restaurants

Scully et al. A time-based objective measure of exposure to the food environment. Under review

- **Stressor:**
  - Fast food restaurant counts, densities, percentages, etc., *within proximity* of participant
- **Receptor:**
  - Participant's dwells and moves



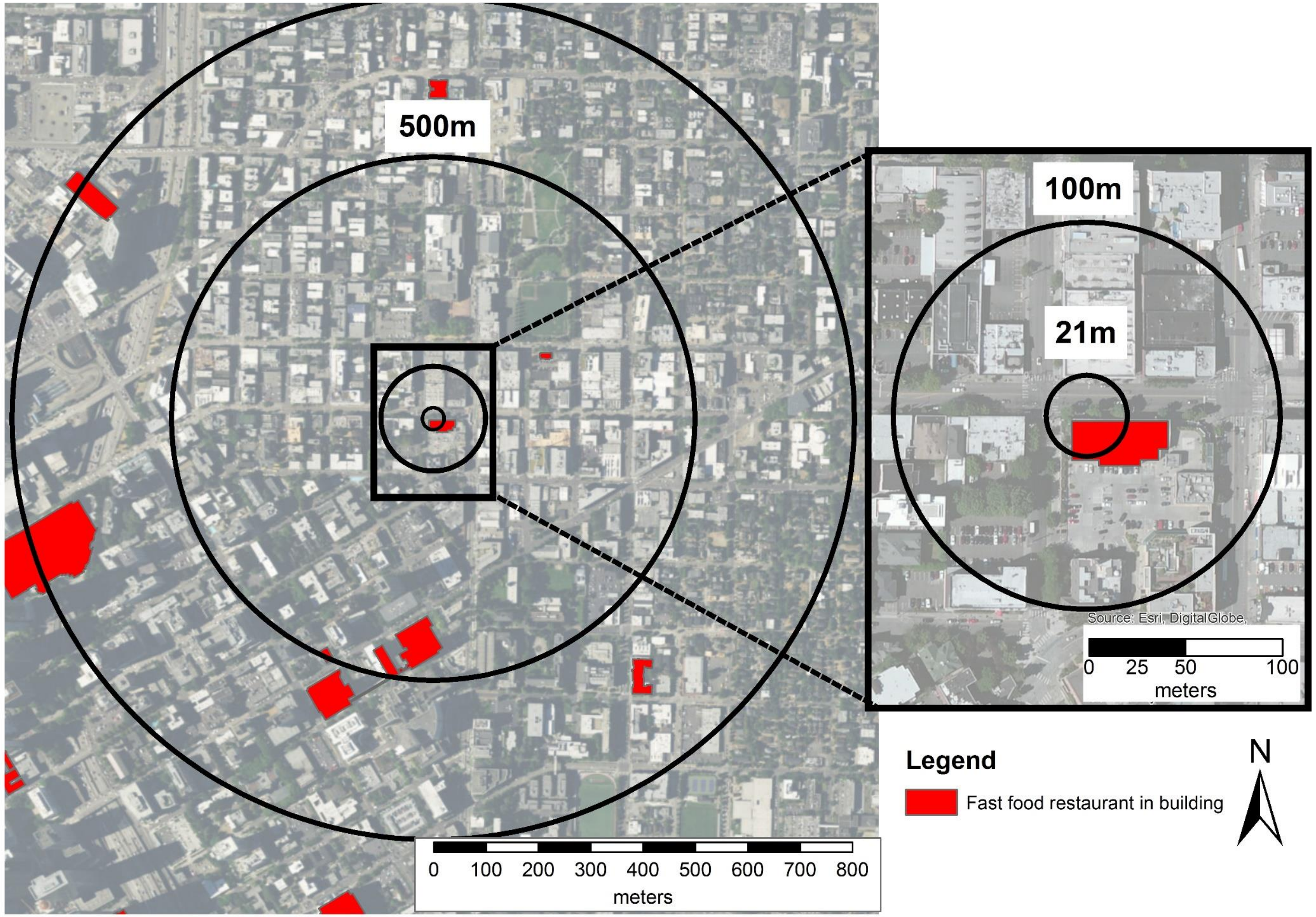
Other references:

- Self-reported activity locations: VERITAS, Chaix et al. 2012
- Work commute path: Burgoine et al. 2013, & Burgoine et al. 2014
- GPS: Zenk et al. 2011, Christian 2012

At what proximity may contact occur?

Scully JY et al.

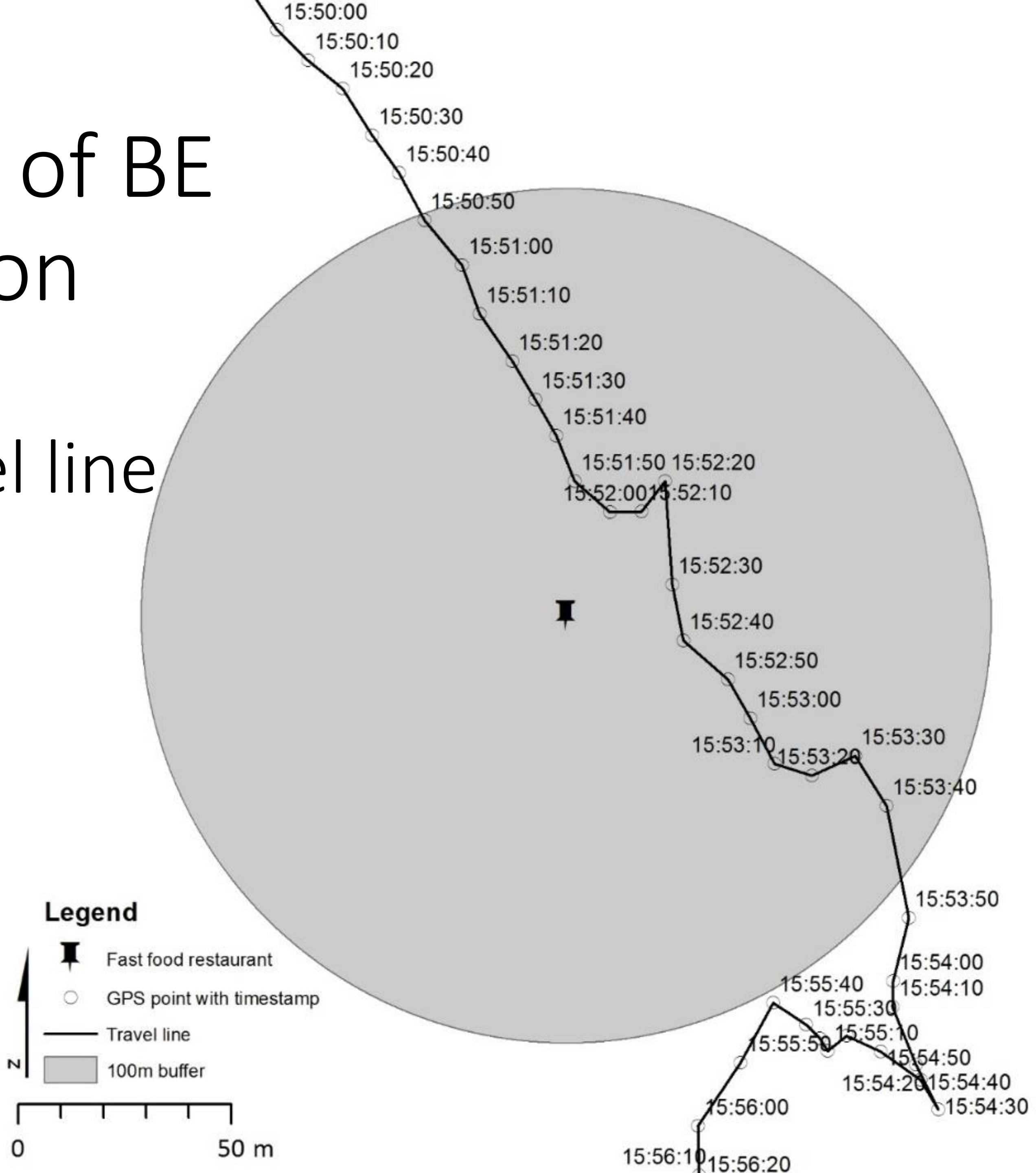
- **21 m**
  - Width of a street
  - Farthest distance at which a human face is recognize
- **100 m**
  - Length of a city block
  - Pedestrian travel
- **500 m**
  - What is accessible by car
- **Half-mile**
  - A 10-minute walk



A novel measure of BE exposure: duration

Scully J Y et al. in progress

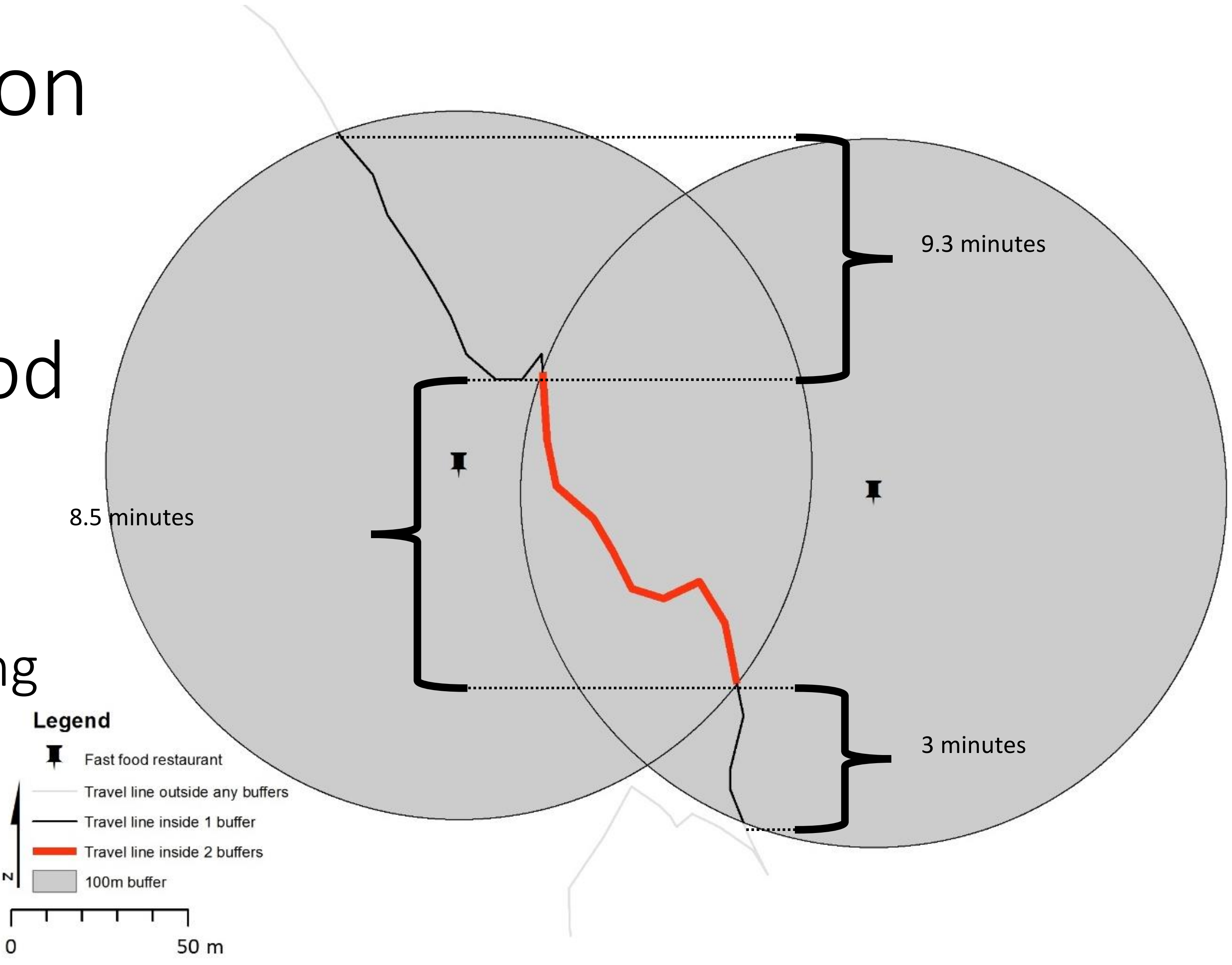
Measuring GPS travel line through FFR buffer



Exposure duration weighted by the number of proximate fast food restaurants

Scully JY et al.

GPS travel line intersecting two or more overlapping buffers = time inside overlapping buffers * number of buffers

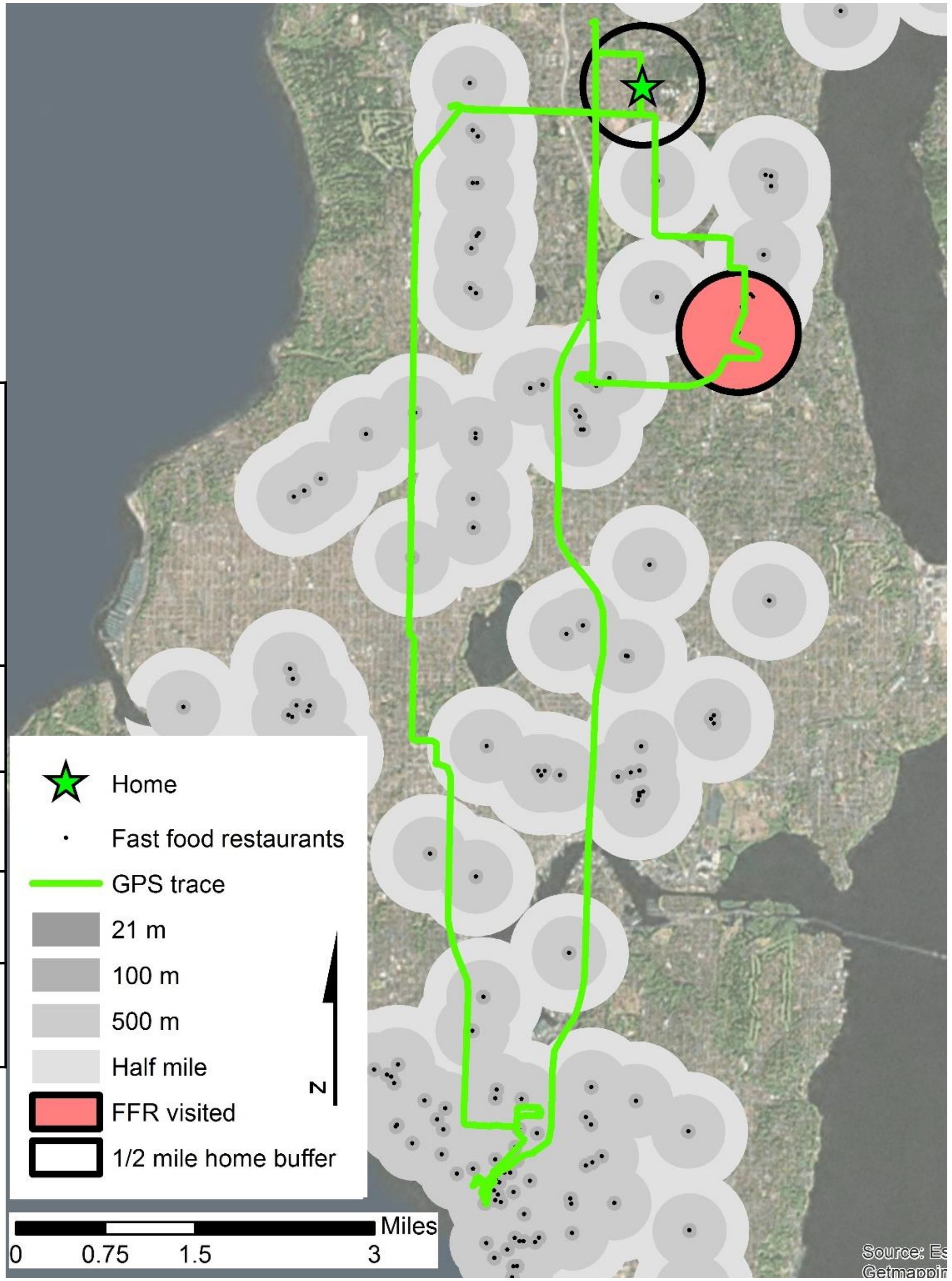


Duration of exposure = 9.3 + 8.5 + 3
Weighted duration of exposure = 9.3*1 + 8.5*2 + 3*1

Results

Scully JY et al.

Proximity	**Fast food restaurant** Mean count/day (SD)	**Exposure duration** Mean minutes/day (SD)
21 m	1.5 (1.1)	1.0 (1.8)
100 m	8.1 (4.5)	17.0 (16.6)
500 m	24.34 (13.2)	84.8 (56.7)
Half mile	34.1 (18.9)	117.7 (69.2)



Results: Odds of visiting > one FFR

Scully J Y et al. In progress

	21 meters			100 meters			500 meters			Half mile		
	Odds	95% CI	p-value	Odds	95% CI	p-value	Odds	95% CI	p-value	Odds	95% CI	p-value
A: By duration of exposure*												
(Intercept)	0.53	0.22-1.27	0.147	0.47	0.19-1.2	0.109	0.7	0.29-1.69	0.426	0.55	0.23-1.36	0.191
Tertile 1	Ref			Ref			Ref			Ref		
Tertile 2	**2.06**	**1.17-3.65**	**0.011**	1.24	0.7-2.18	0.456	1.06	0.61-1.83	0.844	**1.93**	**1.1-3.39**	**0.021**
Tertile 3	**2.8**	**1.58-4.96**	**0.000**	**2.89**	1.65-5.07	**0.000**	**1.72**	**1-2.94**	**0.046**	**2.16**	1.22-3.83	**0.008**
B: By count of FFRs*												
(Intercept)	0.78	0.34-1.83	0.570	0.74	0.31-1.8	0.507	0.77	0.33-1.81	0.541	0.86	0.37-2	0.720
Tertile 1	Ref			Ref			Ref			Ref		
Tertile 2	1.26	0.73-2.18	0.408	1.16	0.66-2.04	0.601	1.32	0.76-2.3	0.323	1.06	0.6-1.86	0.849
Tertile 3	1.41	0.8-2.47	0.229	1.68	0.96-2.93	0.066	1.38	0.76-2.51	0.289	1.49	0.83-2.68	0.175
C: By weighted duration of exposure*												
(Intercept)	0.56	0.23-1.33	0.182	0.49	0.2-1.21	0.117	0.72	0.3-1.75	0.462	0.8	0.34-1.92	0.616
Tertile 1	Ref			Ref			Ref			Ref		
Tertile 2	1.62	0.92-2.85	0.092	1.4	0.79-2.47	0.248	1.15	0.67-1.99	0.606	1.25	0.72-2.17	0.423
Tertile 3	**2.69**	**1.53-4.73**	**0.001**	**3.07**	**1.76-5.36**	**0.000**	1.47	0.86-2.52	0.158	1.15	0.67-1.99	0.600

*** Adjusting for age, gender, race, education, income, number of cars in household, household size, commute distance, and residential density.**

5. SmartMaps

Example: Spatial modeling and mapping of BE and behavior at the micro-level

TELUMI example
Transportation-Efficient Land Use Mapping Index

A tool that identifies locations with demand for alternative travel modes (walking, biking, transit) using land use measures

Objective: to help local jurisdictions make decisions on where to target infrastructure and land use investments to support alternative travel modes

Moudon et al. Transportation-Efficient Land Use Mapping Index (TELUMI), a Tool to Assess Multimodal Transportation Options in Metropolitan Regions. International Journal of Sustainable Transportation. 5:111–133, 2011
DOI: 10.1080/15568311003624262

TELUMI as a multi-functional tool

Interdisciplinary and interpersonal

to visually display land-use conditions associated with different modes of travel,

to perform advanced quantitative analyses of land-use attributes.

to effectively help bridge common communication gaps between lay and professional audiences.

Interactive —use concrete, readily identifiable individual land-use variables

to facilitate scenario building

to target intervention strategies or investment decisions, such as augmenting residential or employment density or building sidewalks, by evaluating their effectiveness in improving transportation efficiency.

Multi-scaled —

to analyze land use at the regional level (macro scale),

with land-use characteristics captured via fine-grained data (micro scale)

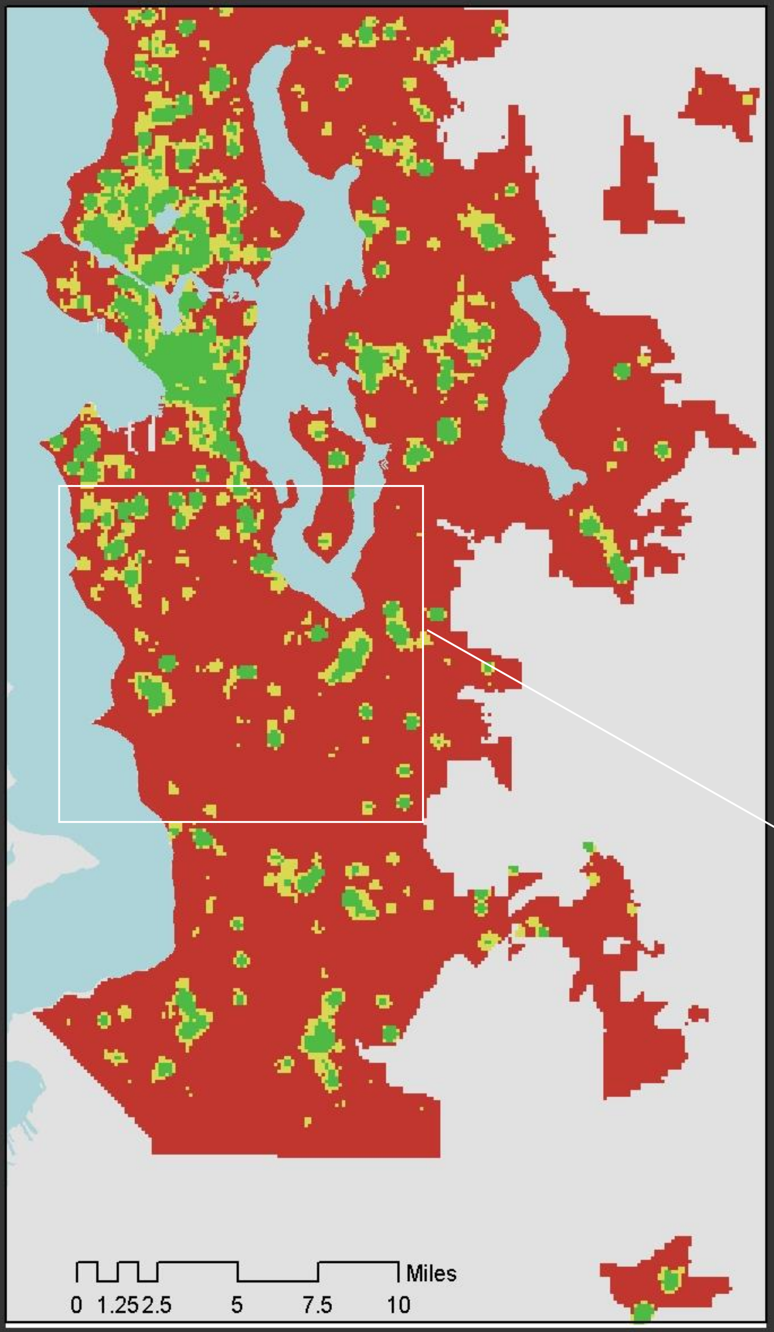
TELUMI

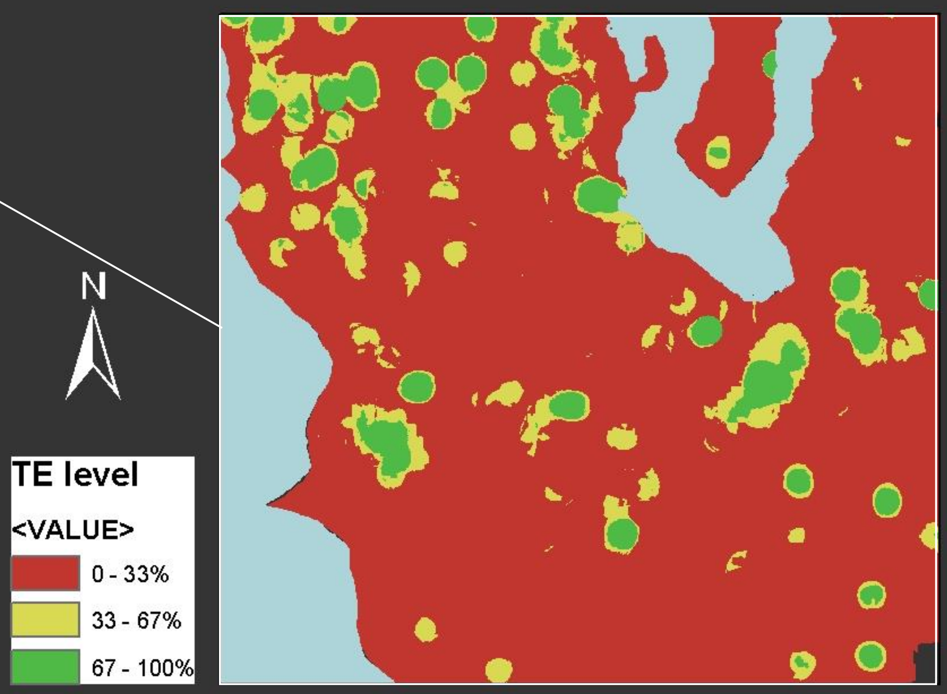
Transportation-**E**fficient **L**and **U**se **M**apping **I**ndex

Final composite layer

King County-Seattle transportation efficiency areas

(green)	HIGH EFFICIENCY
(yellow)	LATENT EFFICIENCY
(red)	LOW EFFICIENCY

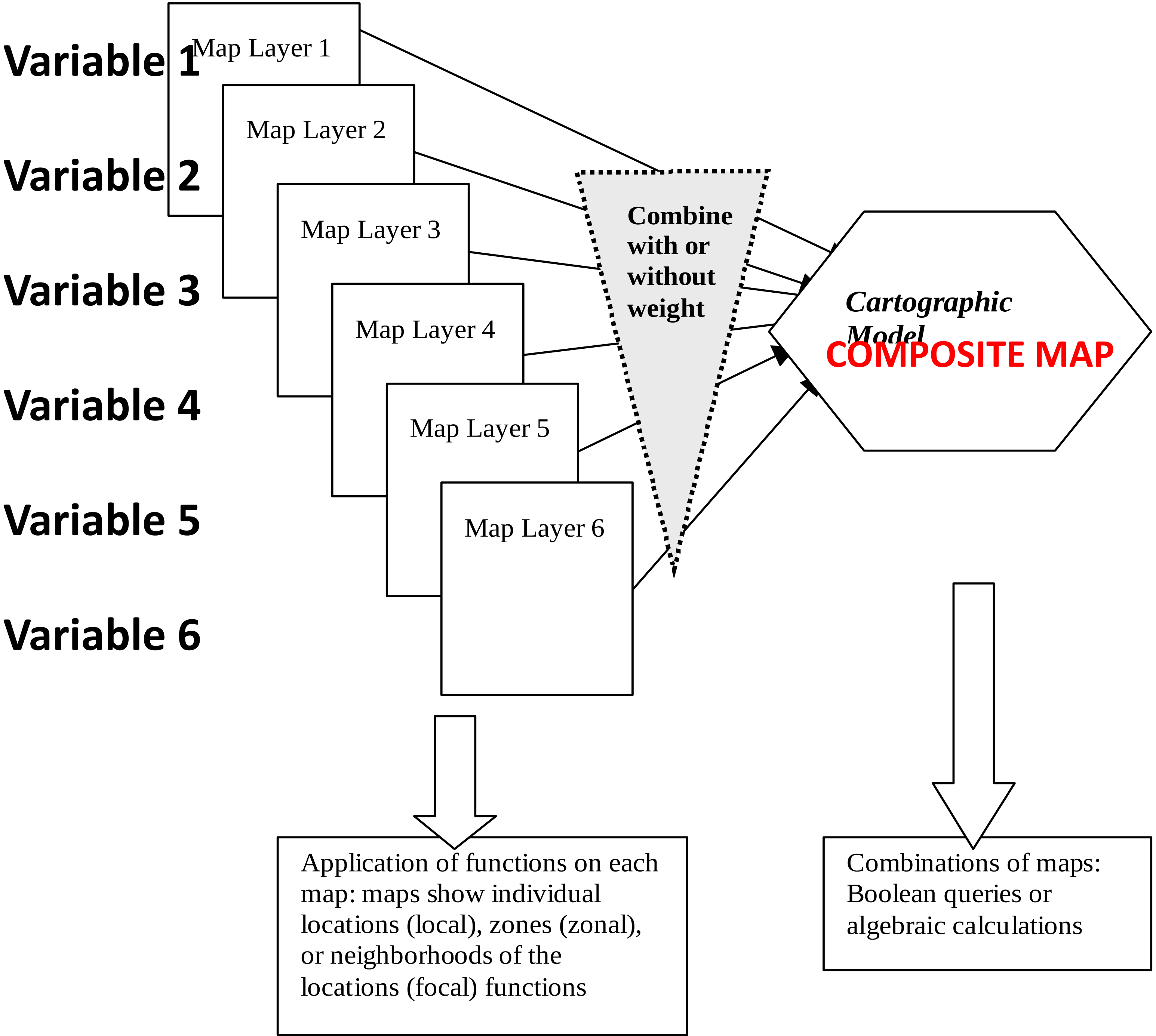




TELUMI
Cartographic Model
Simplified Diagram

Tomlin CD 1970

One land use attribute/variable per map



TELUMI

Six land use domains

Nine variables

	DOMAIN	SPECIFIED VARIABLES/MEASURES
I	Density	• Residential Density [net] • Employment Density [net]
II	Mix of uses	• Proximity to groups of destinations (NC= Neighborhood Center)
III	Network Connectivity	• Average street-block size
IV	Parking supply and management	• % at-grade parking lots in commercial parcels
V	Pedestrian environment	• Topography • Traffic volume (School / Shopping Trips)
IV	Affordable housing	• % of mean assessed residential land and improvement value

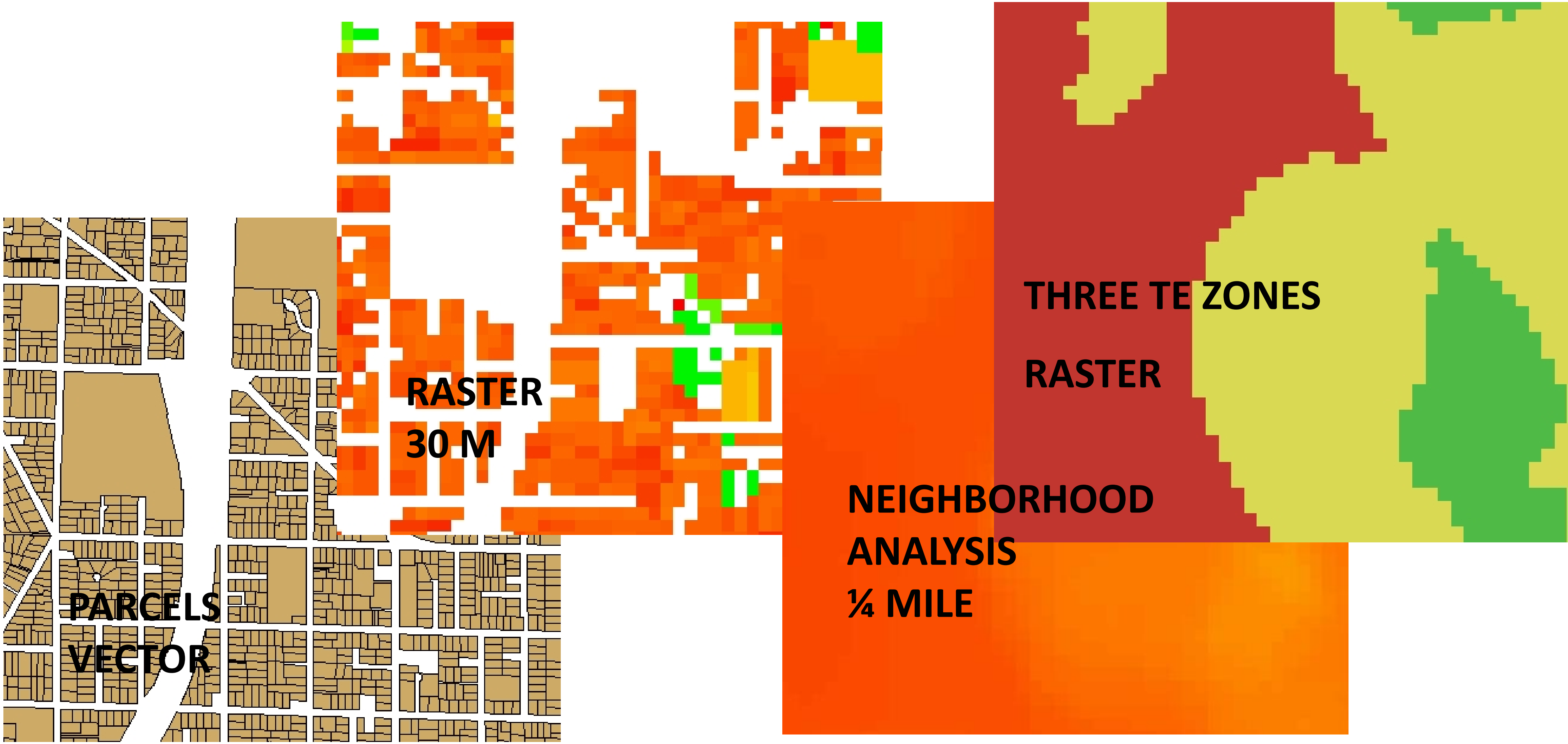
TELUMI

Three levels of Transportation Efficiency (TE)

Transportation Systems		Cartographic Model		
Transportation Options	**Investment Outcomes**	**Zone/Threshold Name**	**Zone Characteristics**	**Example of Threshold Measure**
Low number and types of options	Likely to be ineffective	**Low TE**	Zones with high number of SOV and low number of transit trips	>90+ % of trips in SOV
Medium number and types of options	Likely to be highly effective	**Latent TE**	Zones with medium number of transit or para-transit trips	>75 % of trips in SOV
High number of types and options	Likely to be effective	**High TE**	Zones with high number of transit, para-transit, and non-motorized trips, and low SOV number of trips	<75 % of trips in SOV

TELUMI

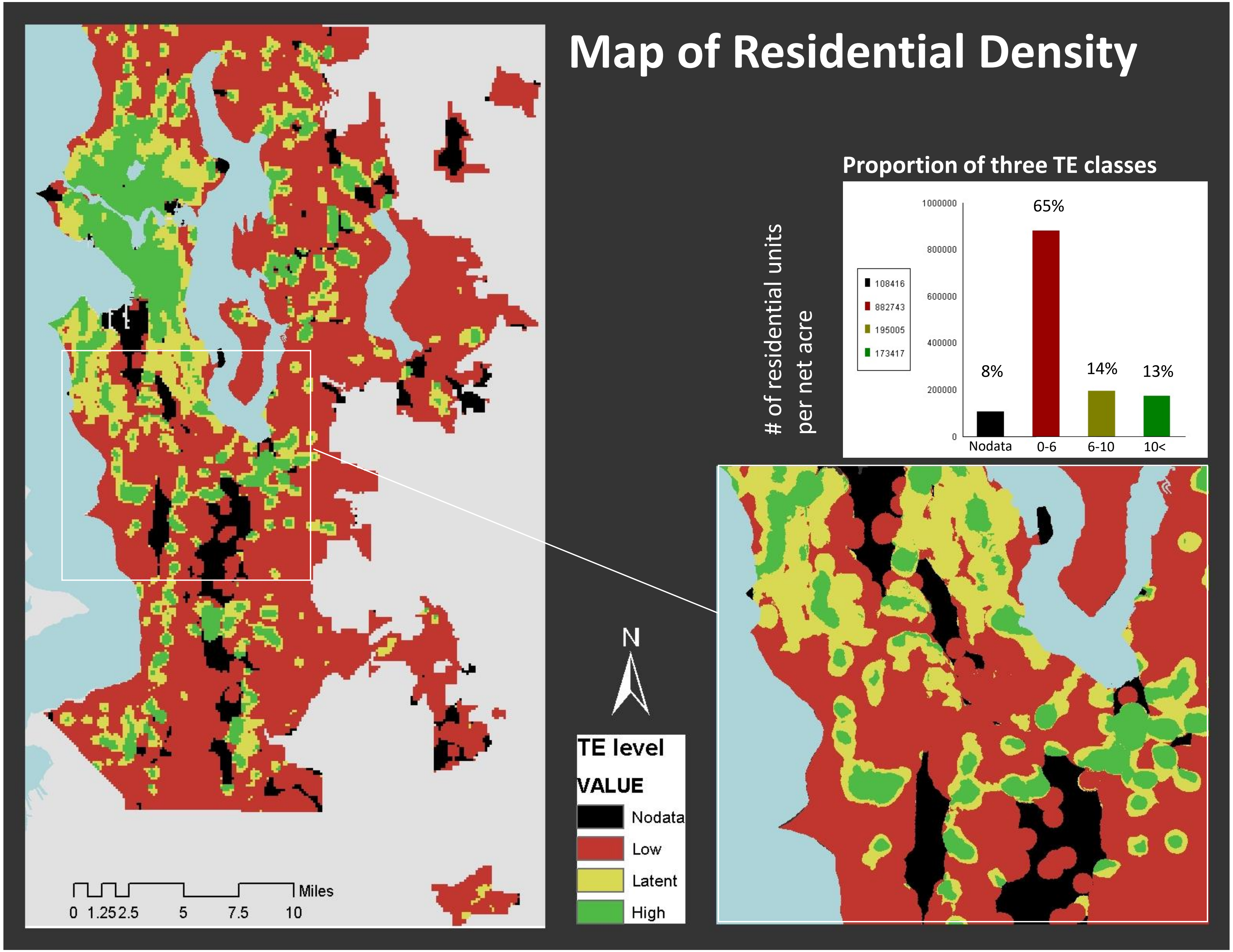
Land use data transformation



TELUMI

Variable 1

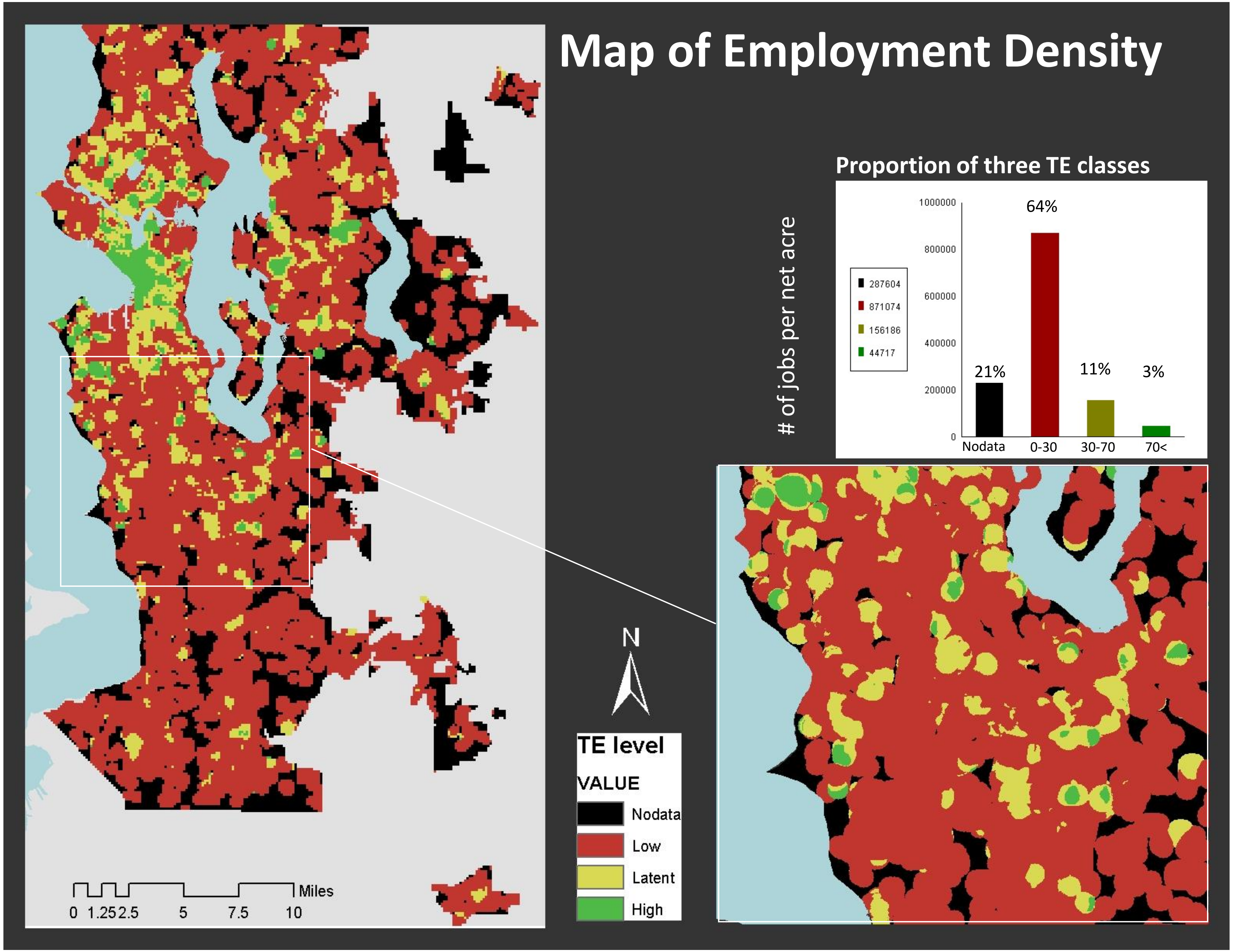
Residential density



TELUMI

Variable 2

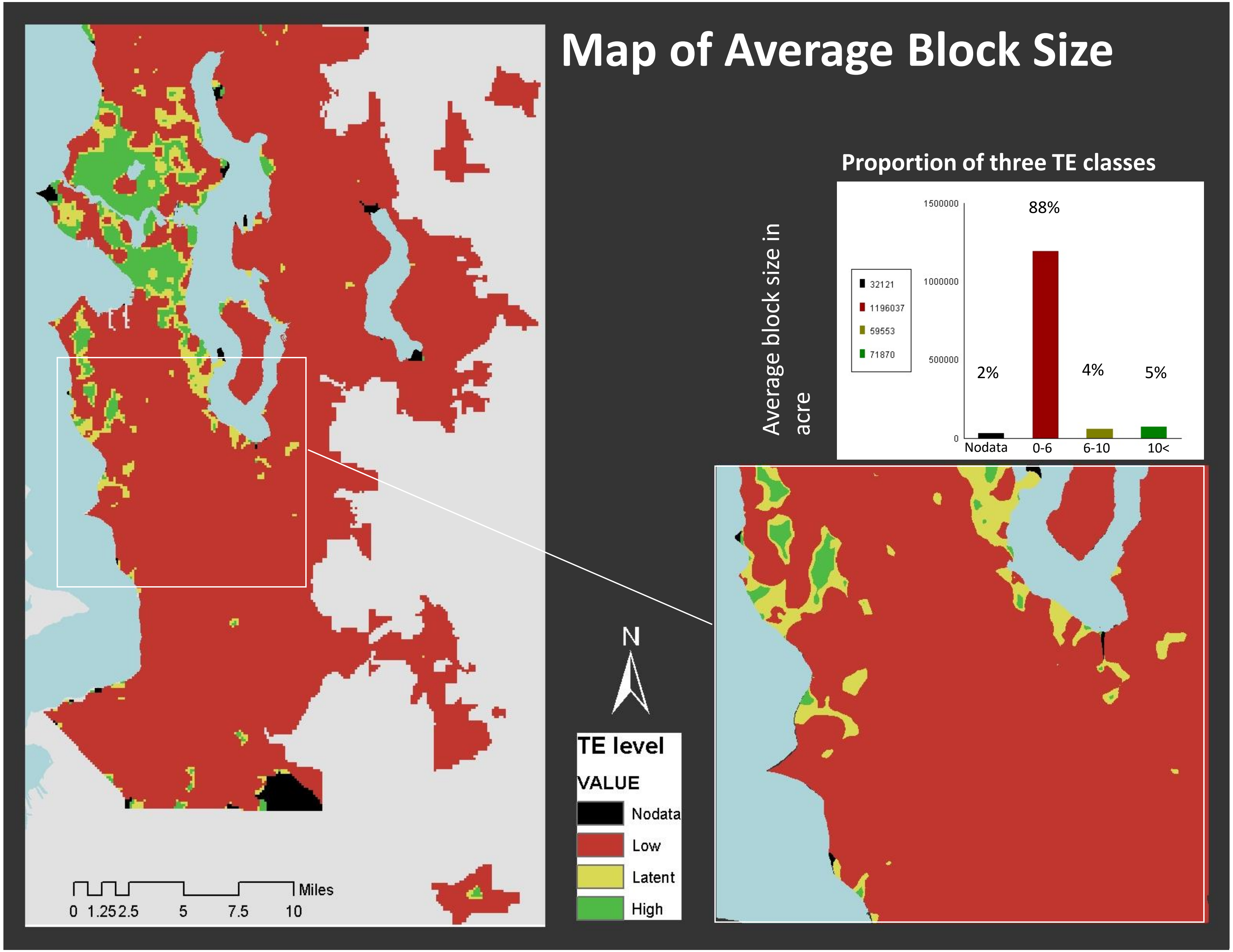
Employment density



TELUMI

Variable 4

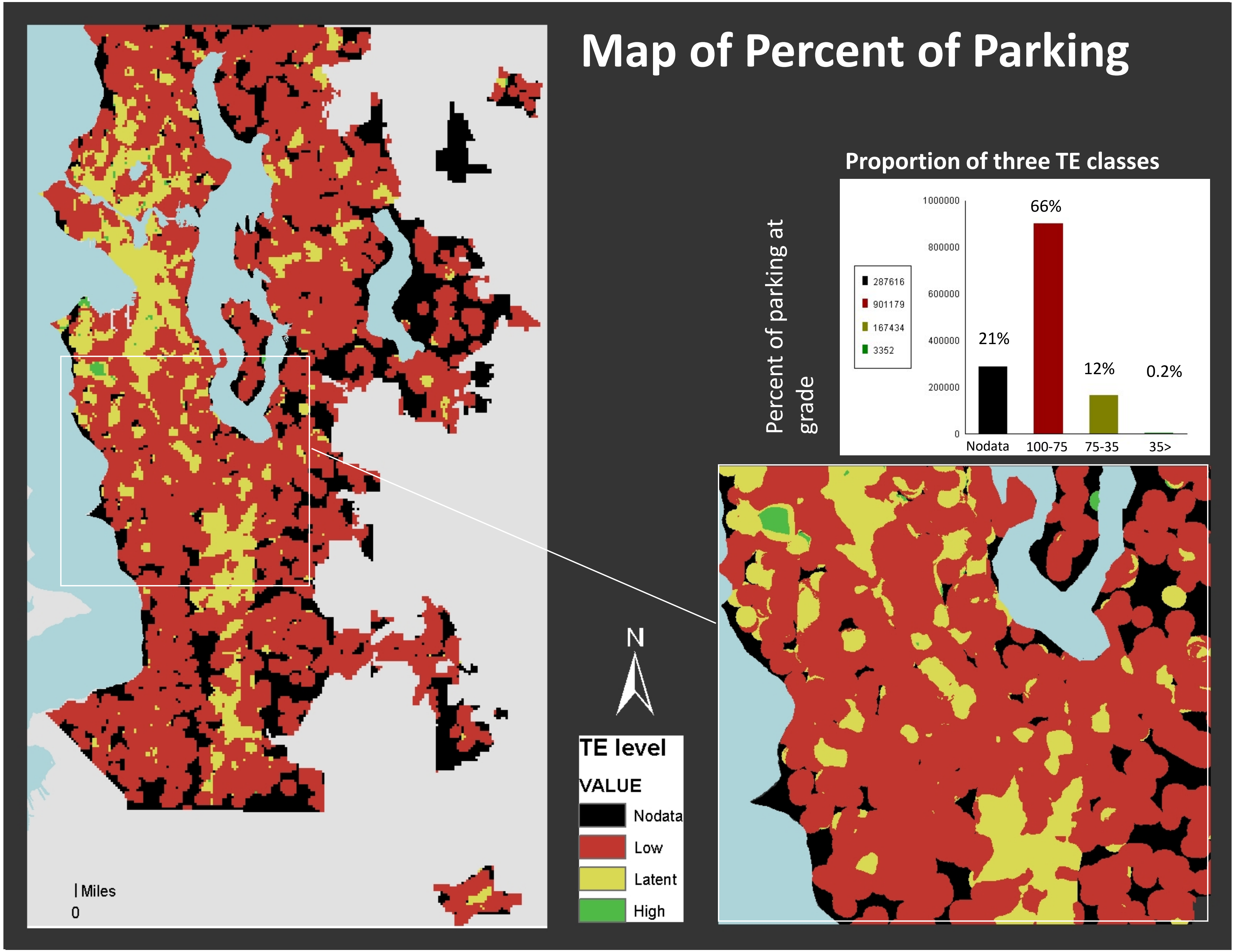
Street-block size



TELUMI

Variable 5

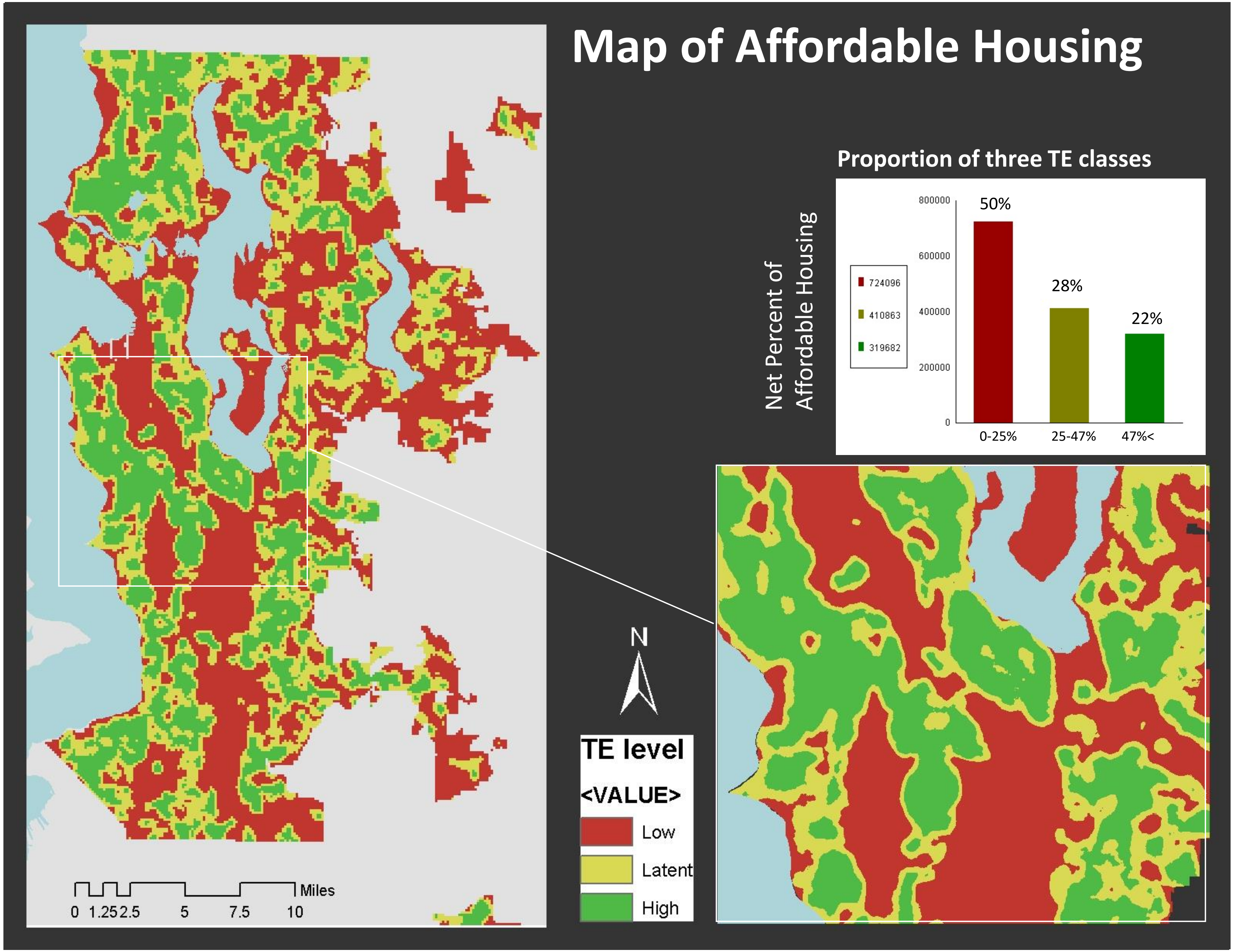
Parking



TELUMI

Variable 9

Affordable housing



TELUMI

Composite Layer

Binary logit model to generate weights for each land use variable

- *Dependent variable:* Dichotomized ridership data <37 versus >37 riders per bus stop per day
  - Threshold of 37 riders per stop (37x4=148 per intersection) divided the sample population of bus users into those in the top 30 percent of higher bus usage, and all the others.
  - Data distribution: 63 percent (3,356 out of 5,363) of the bus stops and 91 percent of boardings and alightings (430,684 out of 473,169) within the Seattle city limits.
- *Independent variables:* 9 TELUMI measures averaged in a quarter-mile radius buffer, centered on bus stop locations

TELUMI

Composite Layer

Logit Model Results

Nagelkerke R-square: 0.344

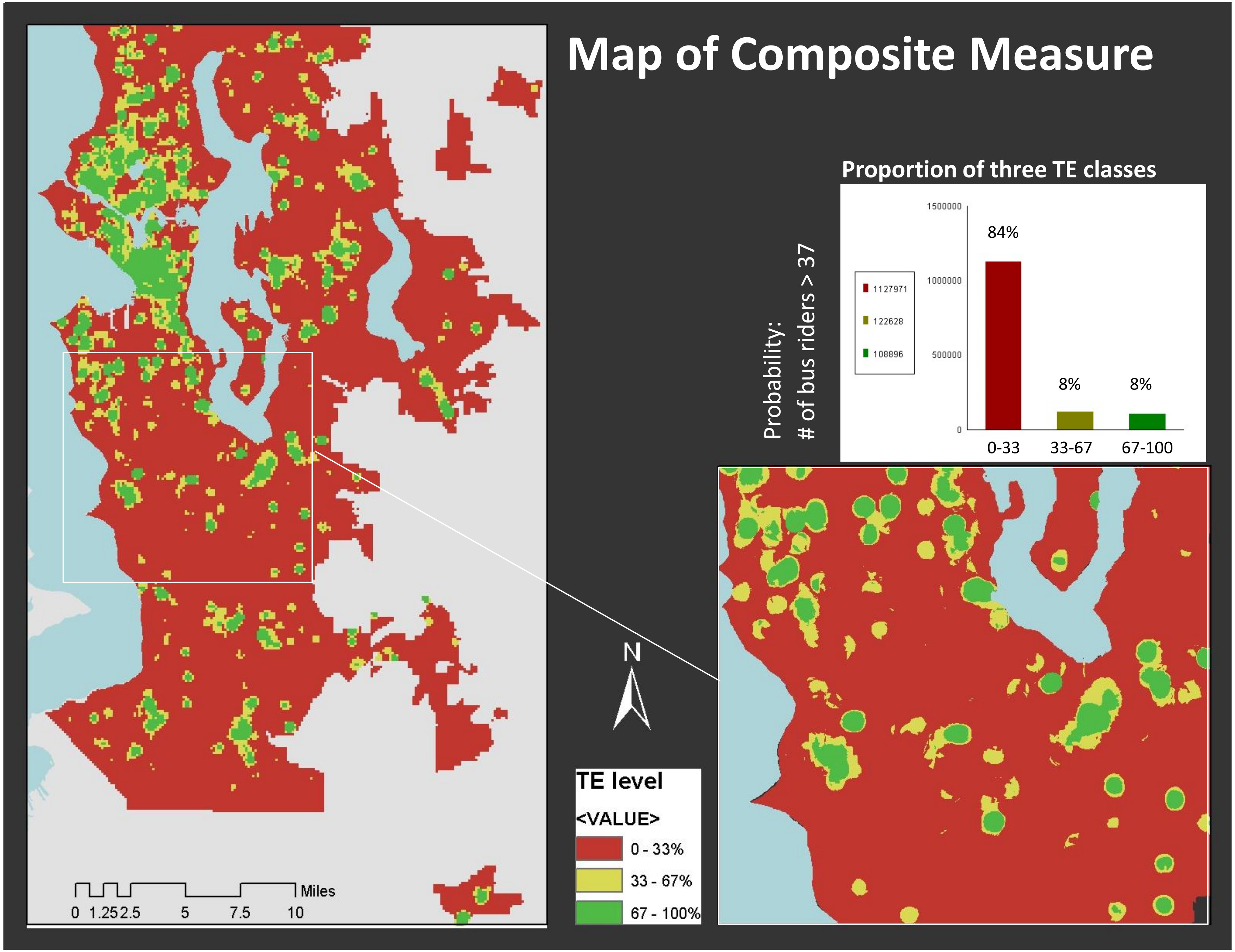
Variable Name	B*	S.E.	Sign.**	Exp(B)
res_den	0.662	0.053	0	1.939
p_parking	0.506	0.076	0	1.659
nc2	0.471	0.08	0	1.602
emp_den	0.416	0.056	0	1.517
slope	0.324	0.07	0	1.383
blk_size	0.311	0.046	0	1.365
sch_traff	0.002	0	0	1.002
ret_traff	0	0	0	1
Constant	-5.181	0.179	0	0.006

*B values are the weights applied to each variable to calculate the composite layer

**Significant at 0.99 level

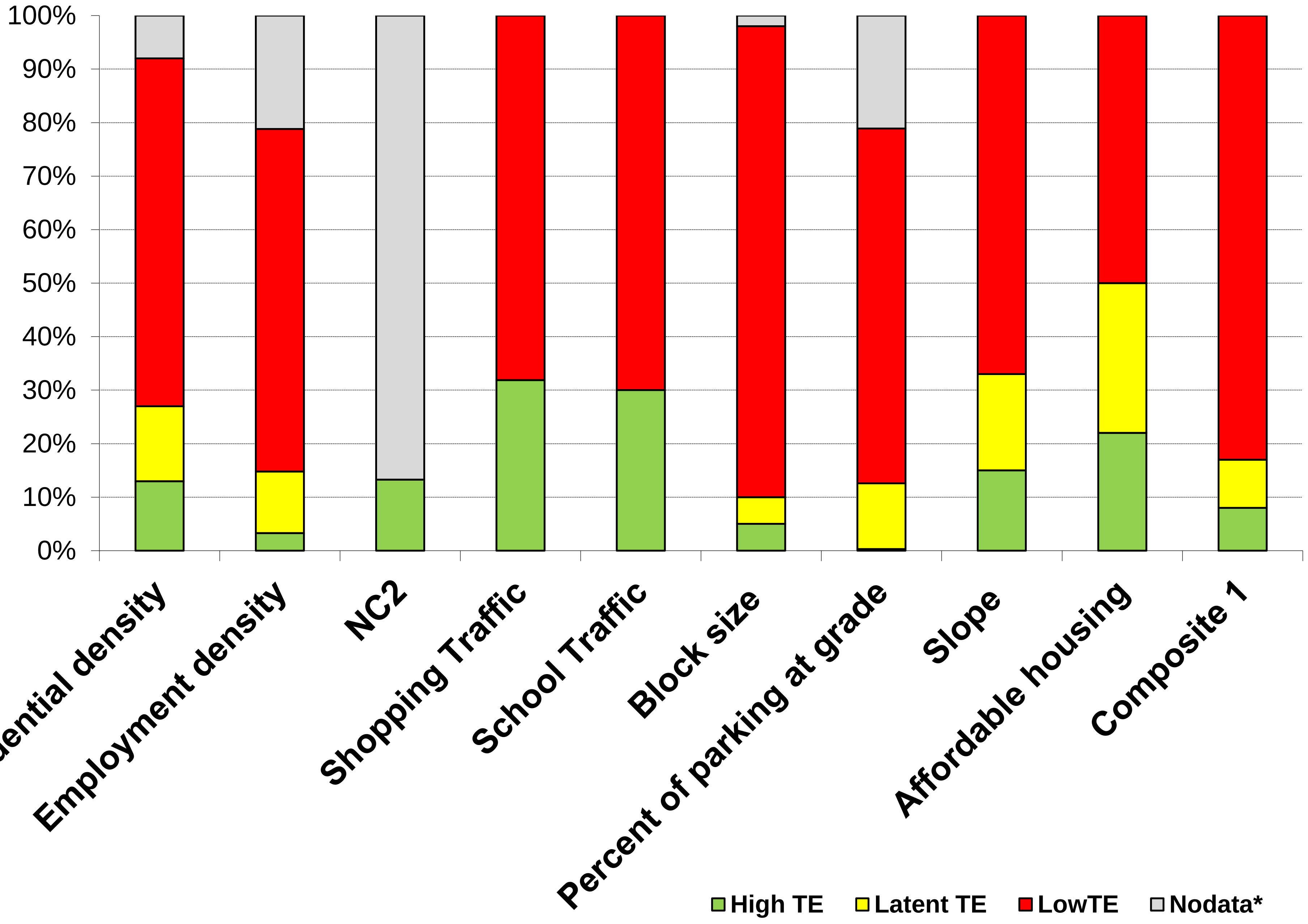
TELUMI

Final composite layer



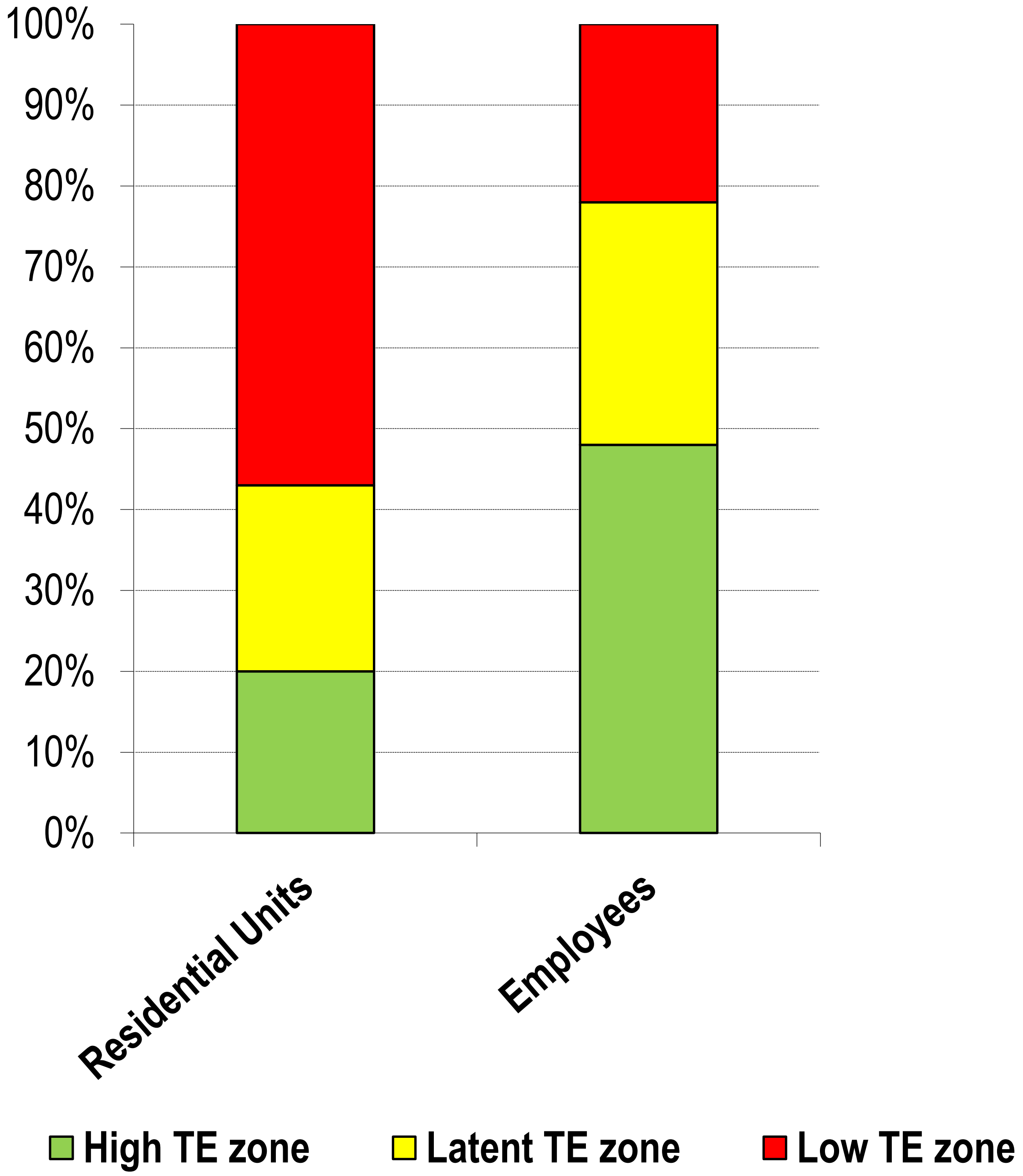
TELUMI

Areas in three TE zones
King County



TELUMI

Distribution of residential units and employment in the three TE Zones



Conclusions

- City planning can improve walkability and support active living
- Transport planning can incentivize transit use and active travel
- City and transport planning can rely on spatial modeling of BE and walking/transit use behavior
- Behavior-built environment research can benefit from advances in sensor technology and exposure theory and measures
    - **Pluses of time-based exposure measures**
    - Travel mode independent
    - Cumulative exposure
    - Removal of places of intended (selected) exposure: Home, work, and places used
    - **Remaining questions**
    - Path selection bias
    - Measuring influence of environmental knowledge on mobility patterns

Thank you