Analysis and Visualisation of Complex Networks

Human Mobility: data, theory and models

Mattia Mazzoli - UniTo

Overview

- Mobility types
- Mobility data
- First human mobility studies
- Overview of mobility models
- Applications:
  - Urban spaces
  - Epidemics

Mobility types

Mobility defined by purpose and distance:

Can you guess what categories we use to define mobility?

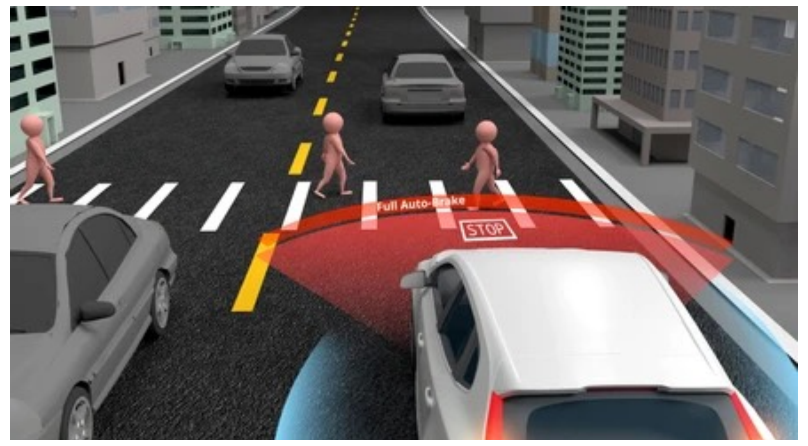
Mobility types

Mobility defined by purpose and distance:

Can you guess what categories we use to define mobility?

Types of mobility:

- Short range: pedestrian (indoor, outdoor, sidewalks)
- Mid-range: commuting (home to work / school), leisure (night time, weekends)
- Long-range: air travel, tourism, migration



Mobility types

Mobility study defined by scale:

Can you guess what spatial scales are used to study mobility?

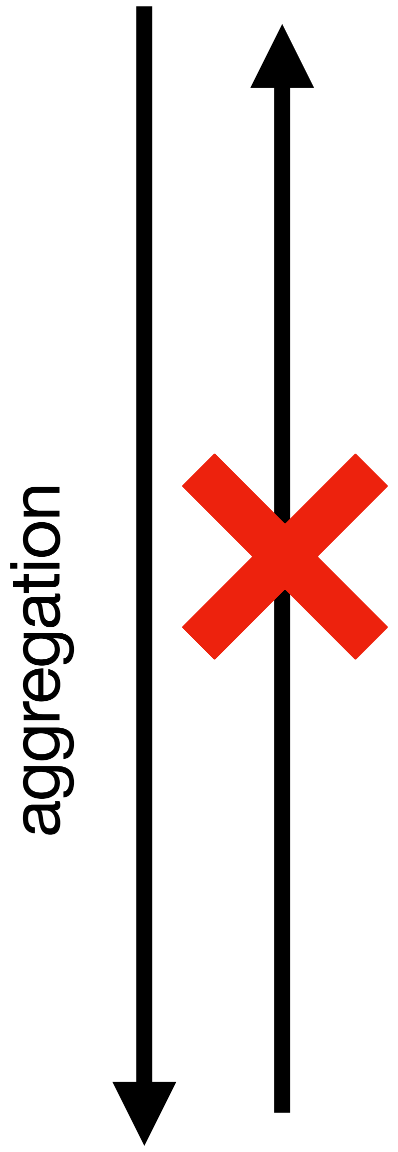
Mobility types

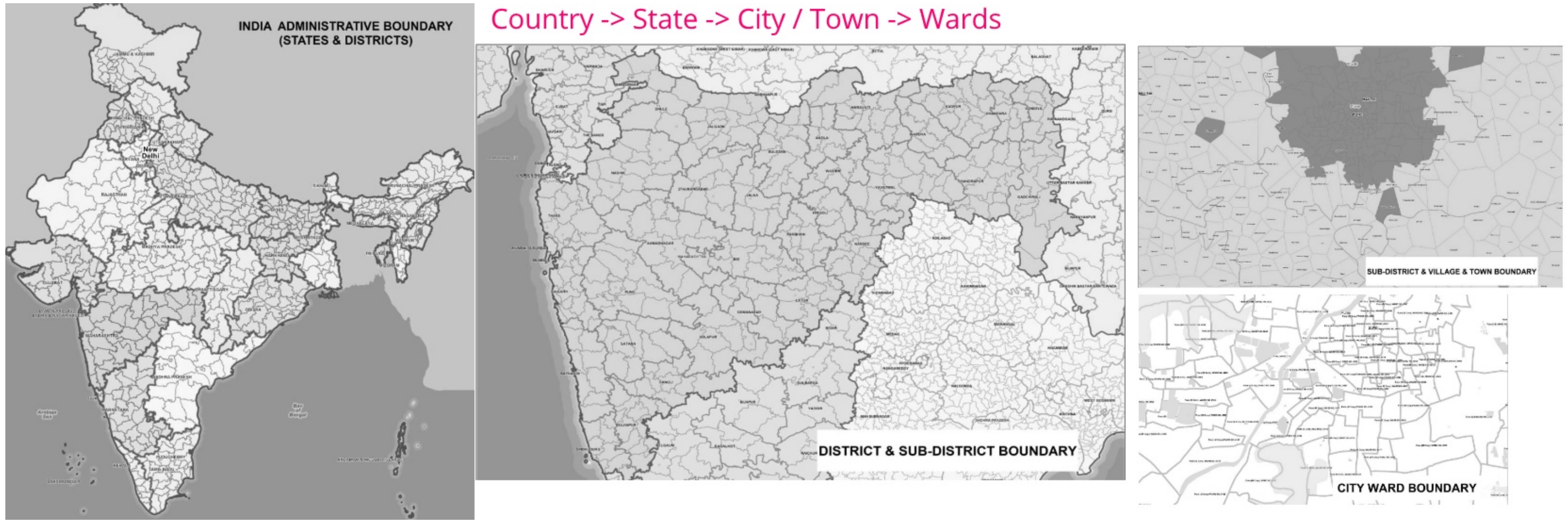
Mobility study defined by scale:

Can you guess what spatial scales are used to study mobility?

Scales of mobility:

- Punctual (lat, lon)
- Spatial grid, ~1x1 sq km.
- Census areas
- Municipalities
- Province
- Regions
- Countries





https://community.geodesignhub.com/

Mobility types

Mobility study defined by scale:

Can you guess what spatial scales are used to study mobility?

Scales of mobility:

- Punctual (lat, lon)
- Spatial grid, ~1x1 sq km.
- Census areas
- Municipalities
- Province
- Regions
- Countries

- Short range: pedestrian (indoor, outdoor, sidewalks), cars routes

Mobility types

Mobility study defined by scale:

Can you guess what spatial scales are used to study mobility?

Scales of mobility:

- Punctual (lat, lon)
- Spatial grid, ~1x1 sq km.
- Census areas
- Municipalities
- Province
- Regions
- Countries

- Mid-range: commuting (home to work / school), leisure (night, weekends)

Mobility types

Mobility study defined by scale:

Can you guess what spatial scales are used to study mobility?

Scales of mobility:

- Punctual (lat, lon)
- Spatial grid, ~1x1 sq km.
- Census areas
- Municipalities
- Province
- Regions
- Countries

- Long-range: air travel, tourism, migration (internal, cross-country)

Why does it matter?

Implications of understanding human mobility:

Can you guess what fields human mobility brings important contributions to?

Why does it matter?

Implications of understanding human mobility:

Can you guess what fields human mobility brings important contributions to?

Fields affected by human mobility:

- Epidemiology (communicable and non-communicable diseases, health accessibility)
- Urban planning (sustainable mobility, smart cities, land-use mixing)
- Transportation and infrastructure engineering (travel demand, traffic regulation, logistic and goods)
- Environment and ecology (pollution, car emissions)
- More…

Why does it matter?

Some research questions in human mobility:

- What makes people move? [Determinants of mobility, collective models]
- How much do people move, new places or old places? [Individual mobility models]
- Where movements will mostly occur? When? How? [Attraction points]
- How many people usually move from there and where do they go? [Transportation planning]
- Traffic viability design [Traffic regulation]
- How much pollution is generated by traffic [Environmental policies]
- Where and how do migrants move? [Humanitarian response]
- How does traveling affects epidemics? [Epidemiology]

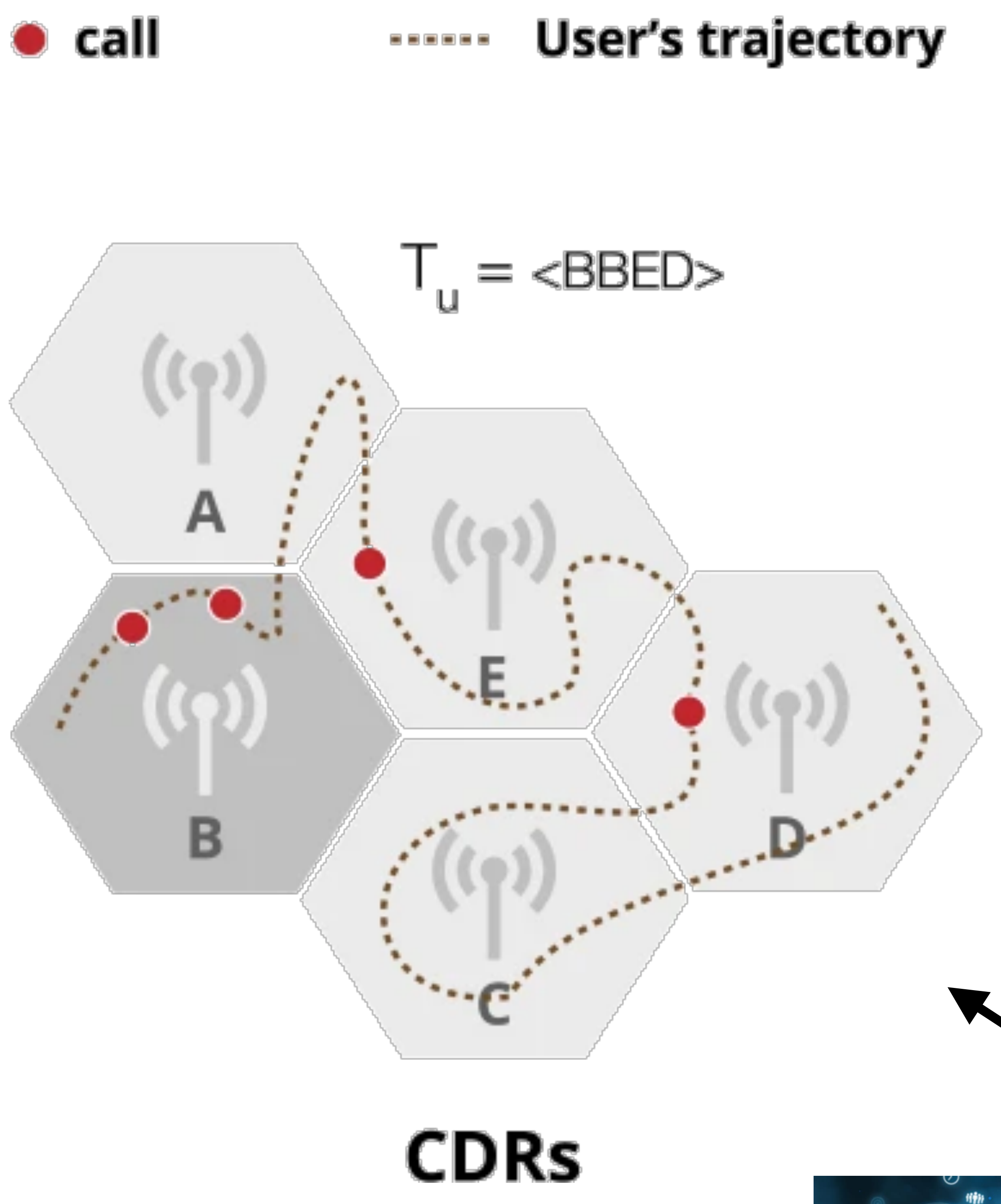
Mobility data

Types of data:

- Mobile phone (CDR, XDR)
- Census (commuting)
- International travel (IATA, Meta Travel patterns)
- GPS traces (Cuebiq, SafeGraph, Meta Co-location, Google location history, Meta Movement range)
- Activity based records (Google mobility reports)
- Surveys

Mobile phone data (CDR & XDR)

… and then, the massive adoption of mobile phones occurred ~ 2000s



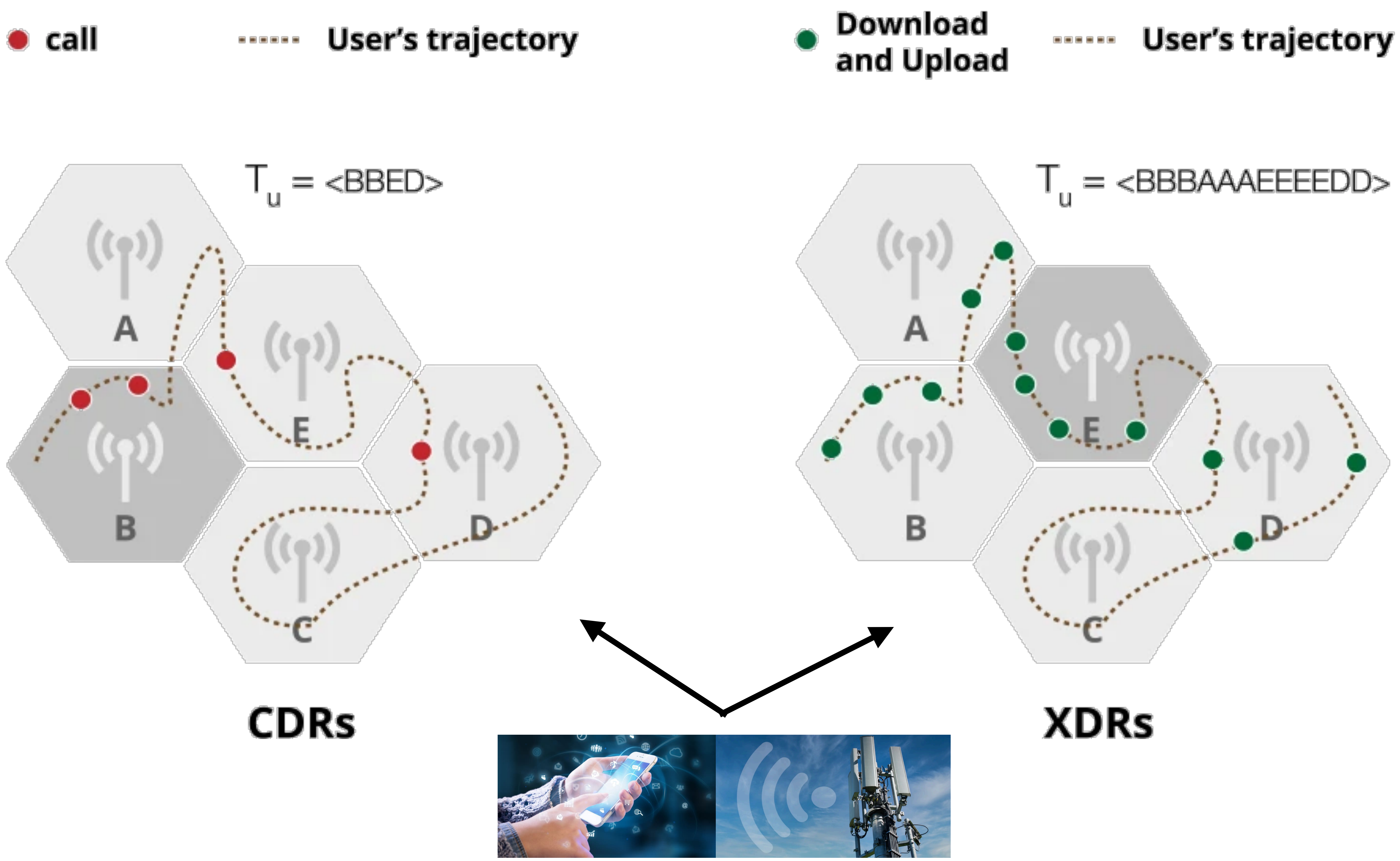
CDR (Call Detail Records)

- Billing purposes
- Covers most of the population
- Pinged when user calls or sends SMS
- Sparse data
- Low spatial and temporal resolution
- Used since years 2000s

Pappalardo, L., Ferres, L., Sacasa, M. *et al.* Evaluation of home detection algorithms on mobile phone data using individual-level ground truth. *EPJ Data Sci.* **10**, 29 (2021).

Mobile phone data (CDR & XDR)

… and then, the massive adoption of mobile phones occurred ~ 2000s



CDR (Call Detail Records)

- Billing purposes
- Covers most of the population
- Pinged when user calls or sends SMS
- Sparse data
- Low spatial and temporal resolution
- Used since years 2000s

XDR (eXtended Detail Records)

- Billing purposes
- Covers most of the population
- Pinged also by app behaviour
- Dense data
- High temporal resolution
- Recently deployed

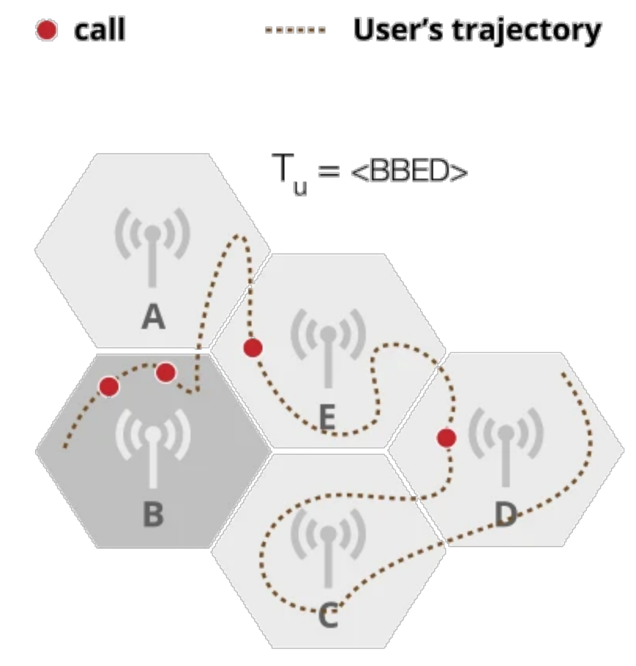
Pappalardo, L., Ferres, L., Sacasa, M. *et al.* Evaluation of home detection algorithms on mobile phone data using individual-level ground truth. *EPJ Data Sci.* **10**, 29 (2021).

Data structure (CDR & XDR)

CDR

00:00:18,MT,38,B92F1982701DF65DF2FD2CAA,,E14E3DCF280350EBB2C69E6B,41575,38

Timestamp: 00:00:18
Call / SMS: MT
Call Length: 38
Origin UUID: B92F1982701DF65DF2FD2CAA
Dest UUID: E14E3DCF280350EBB2C69E6B
Dest Cell Tower ID: 41575

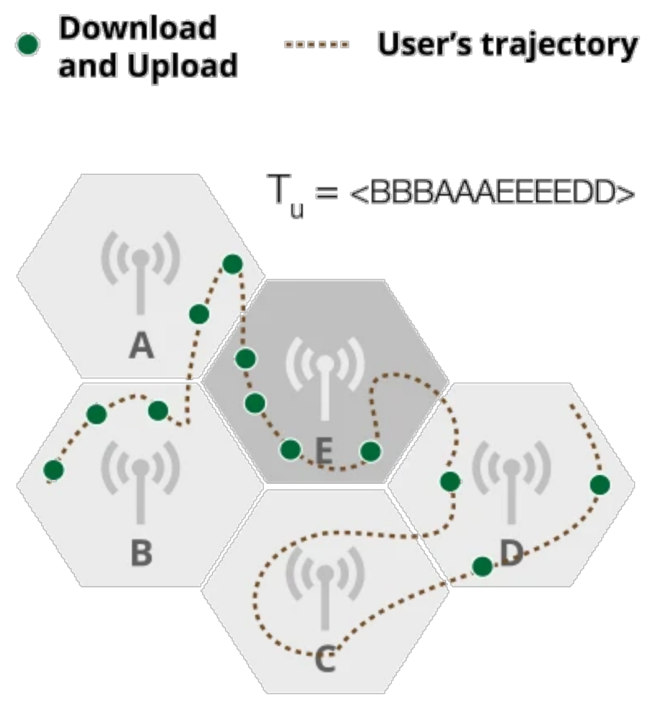


CDRs

XDR

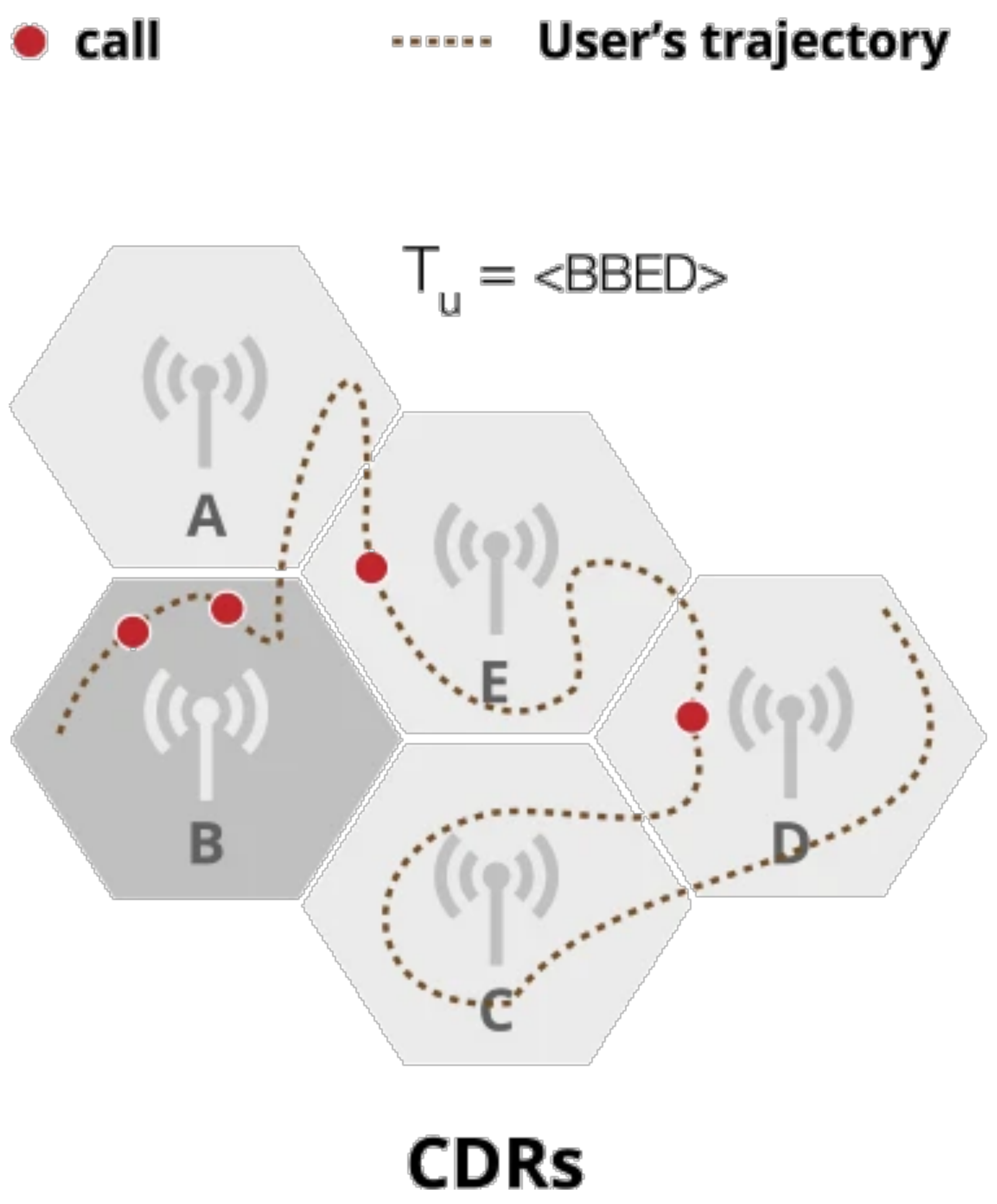
	phone_id	cell_id	technology_cell_cd	timestamps	trafico
▶	0000e2806116ad9dab7be18125ce4e8e365c38c4a84b4ef713008ed8fba8e395	1078306	4	2022-03-07 16:05:17	4
	0000ee66426b94c621683c5e11ab356c2136ba94134fbf2f1d5a0e0bc05bdee0	159815199	4	2022-03-07 23:47:42	52
	0002a5c62fb5f16d3a0100bfea38a17e98fd918212c558ac7d64ffff2a3047d2	44031	3	2022-03-07 03:27:19	1
	0002b61943bcb27ce0aca2aa244d588254e237494ee2f055005723dcc5c60712	59685	3	2022-03-07 23:19:38	3
	0002f2463d2be1263754374cf78f47c80781c187a7ee7e0153f5cc9ddc8c2660	15509	3	2022-03-07 10:58:16	28

Encrypted device ID — Cell tower ID — Time — Data (KB)

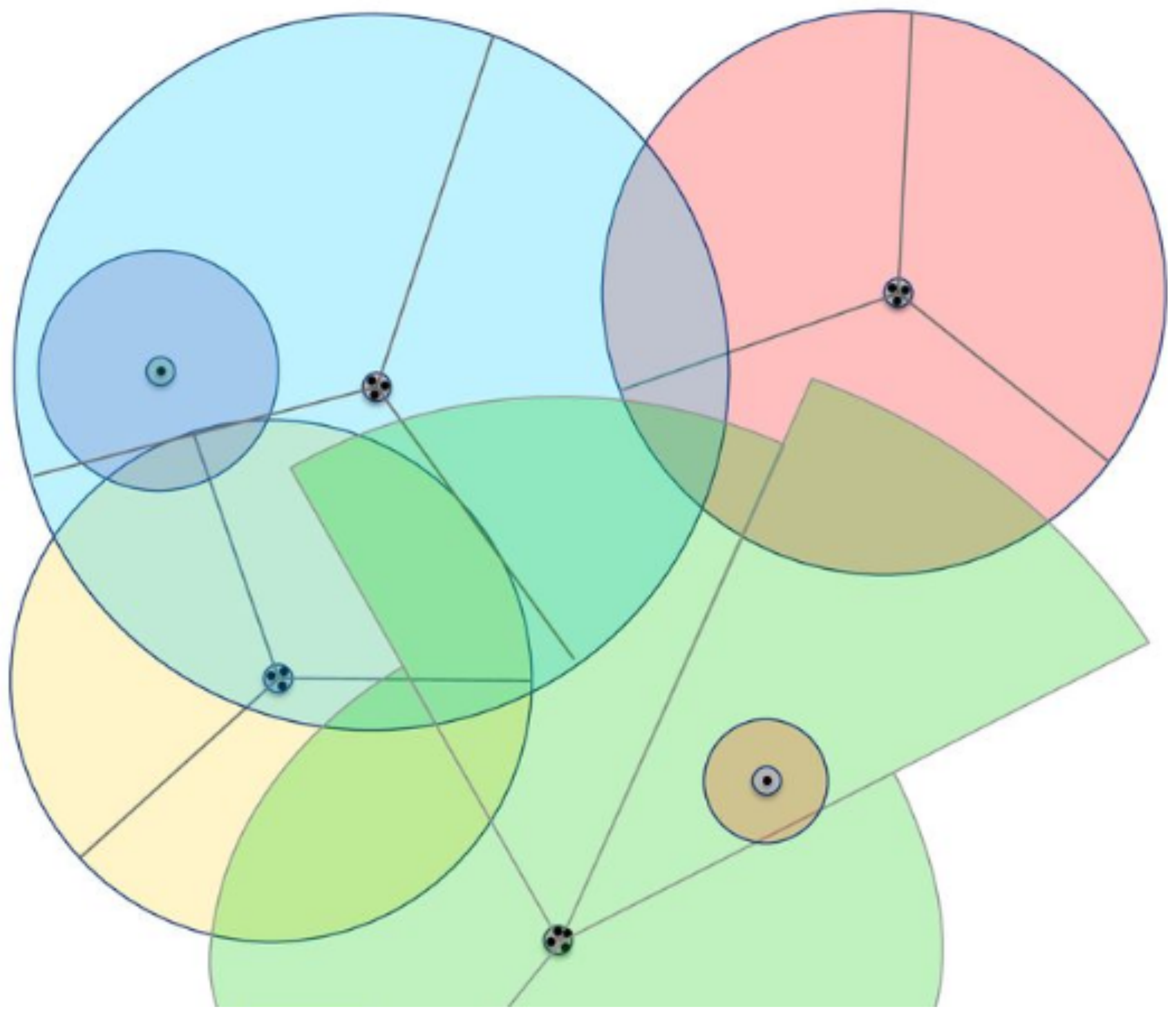


XDRs

Mobile phone data (CDR & XDR)

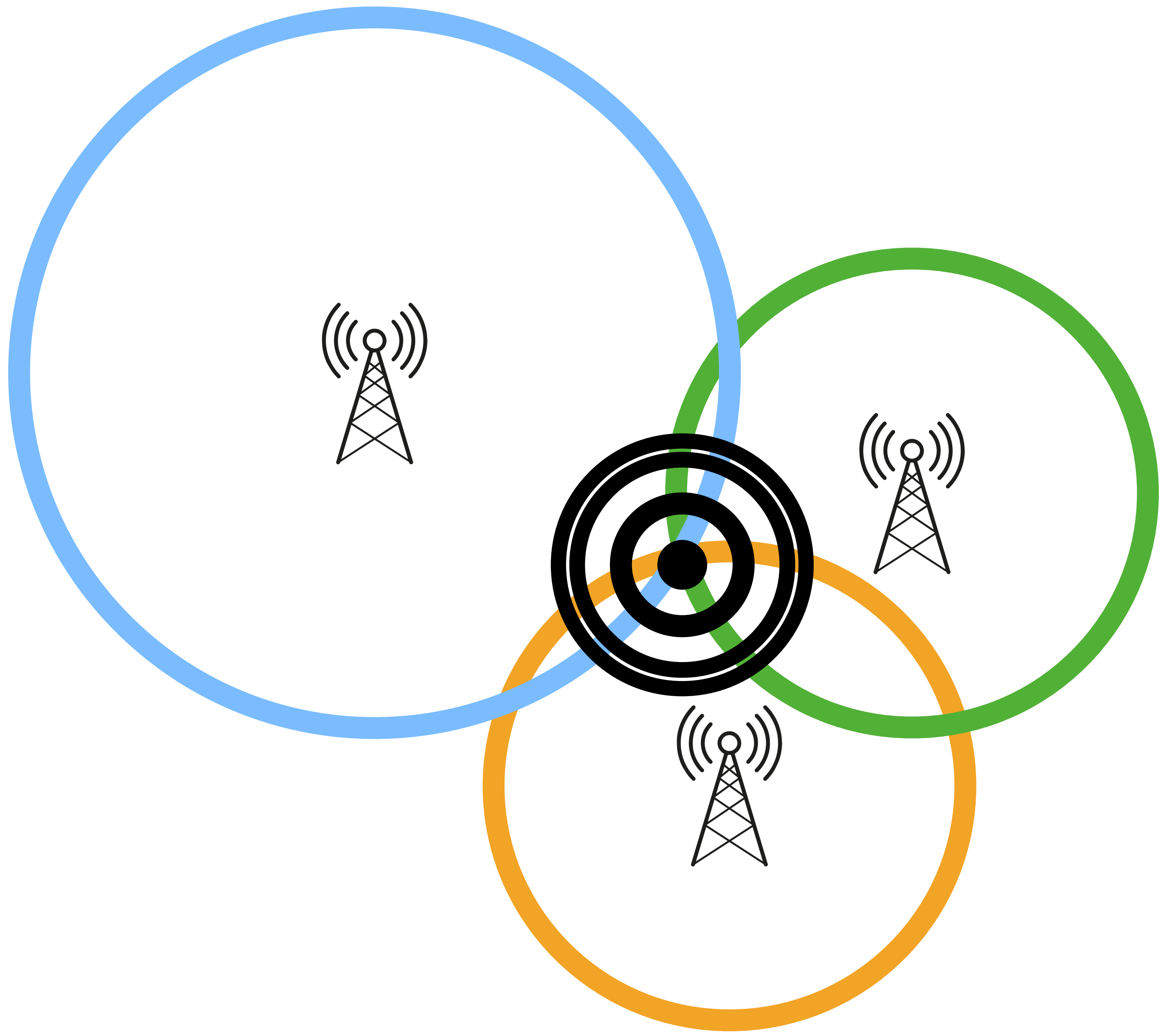


Actual tower cells coverage overlap



Device connects to the tower with the best signal, not always the closest

Towers triangulation



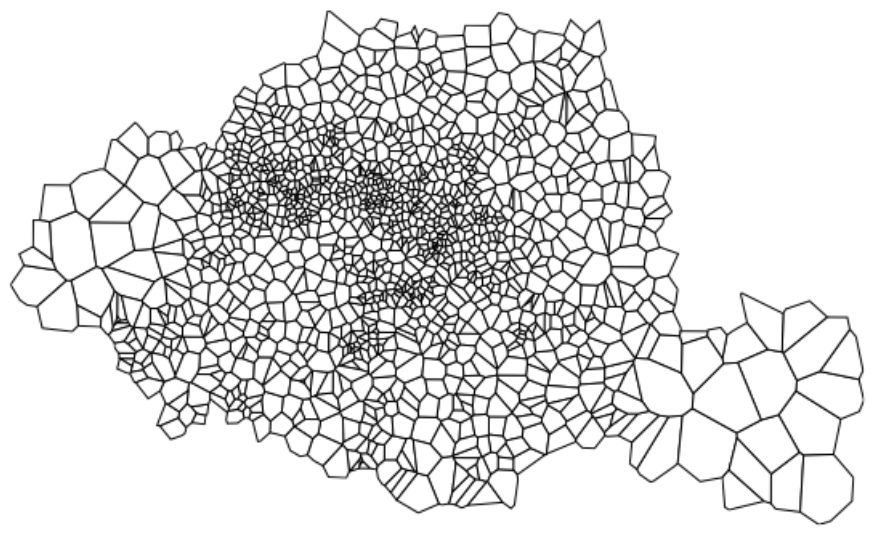
Not performed routinely
For forensic purpose

Pappalardo, L., Ferres, L., Sacasa, M. *et al.* Evaluation of home detection algorithms on mobile phone data using individual-level ground truth. *EPJ Data Sci.* **10**, 29 (2021).
Mobile Phone technology - Maria Sanders https://slideplayer.com/slide/14069191/

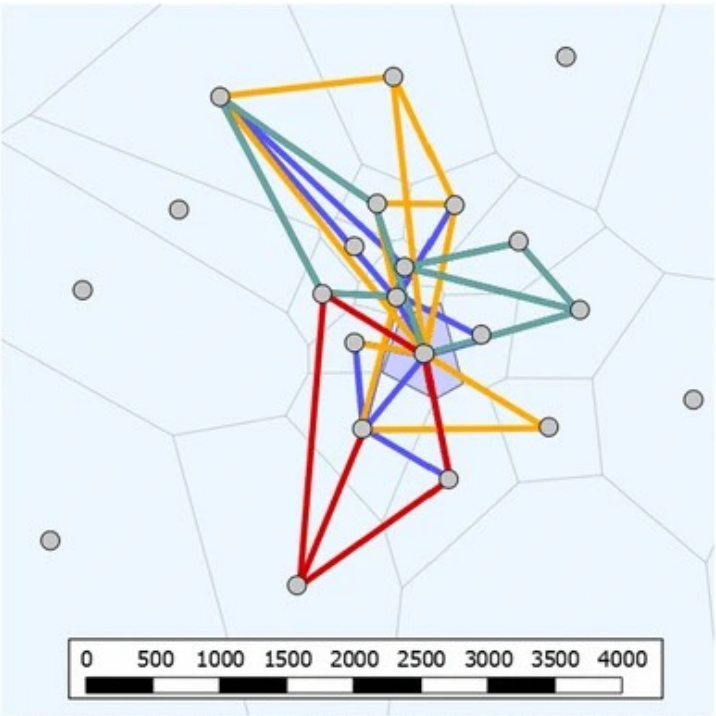
Mobile phone data (CDR & XDR)

Assumption: devices connect to the closest tower (not always true) ~ Voronoi tessellation
CDR and XDR **provide the position of the tower**, not the position of the device!

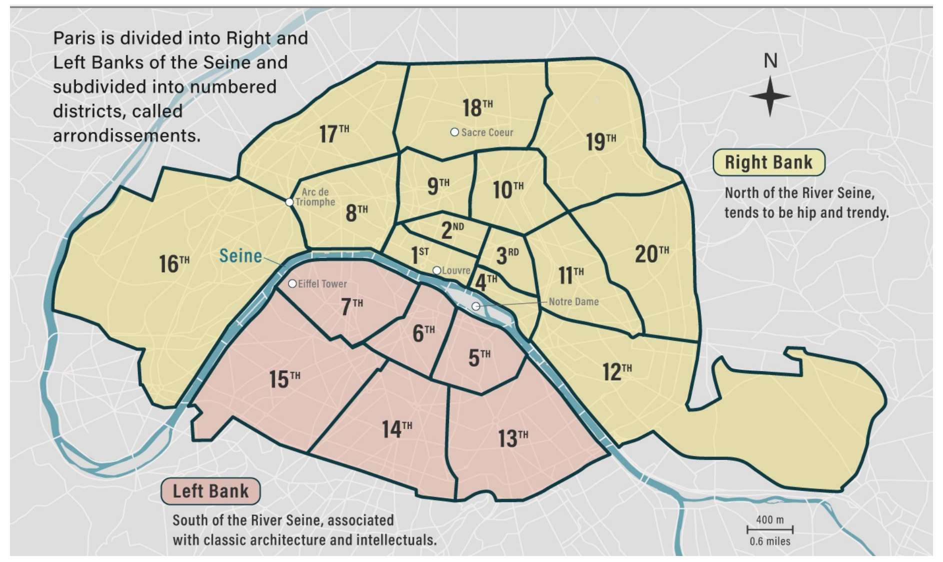
Paris Voronoi cells

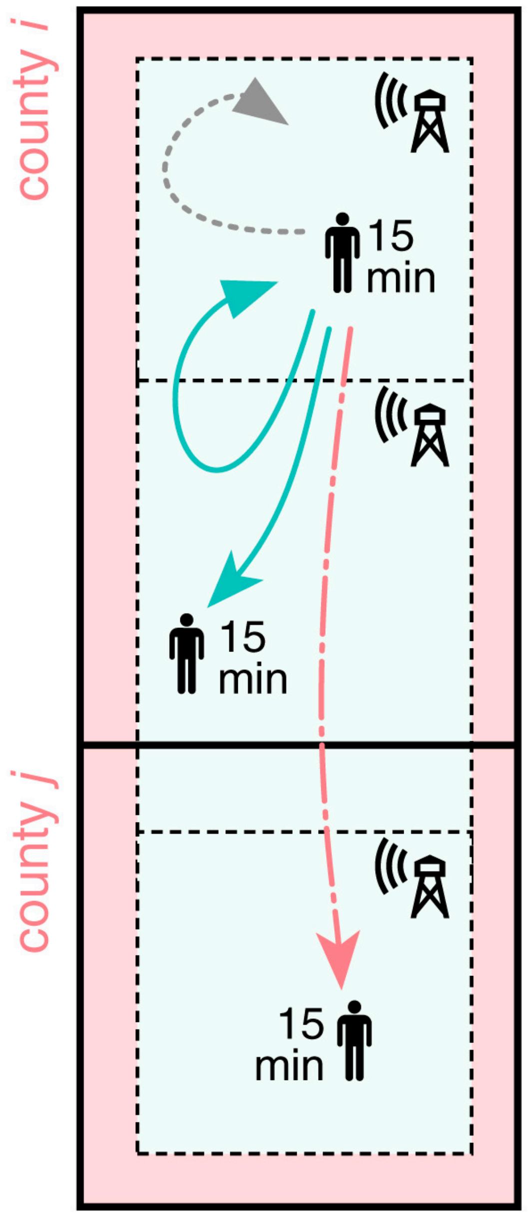


Voronoi cells trips



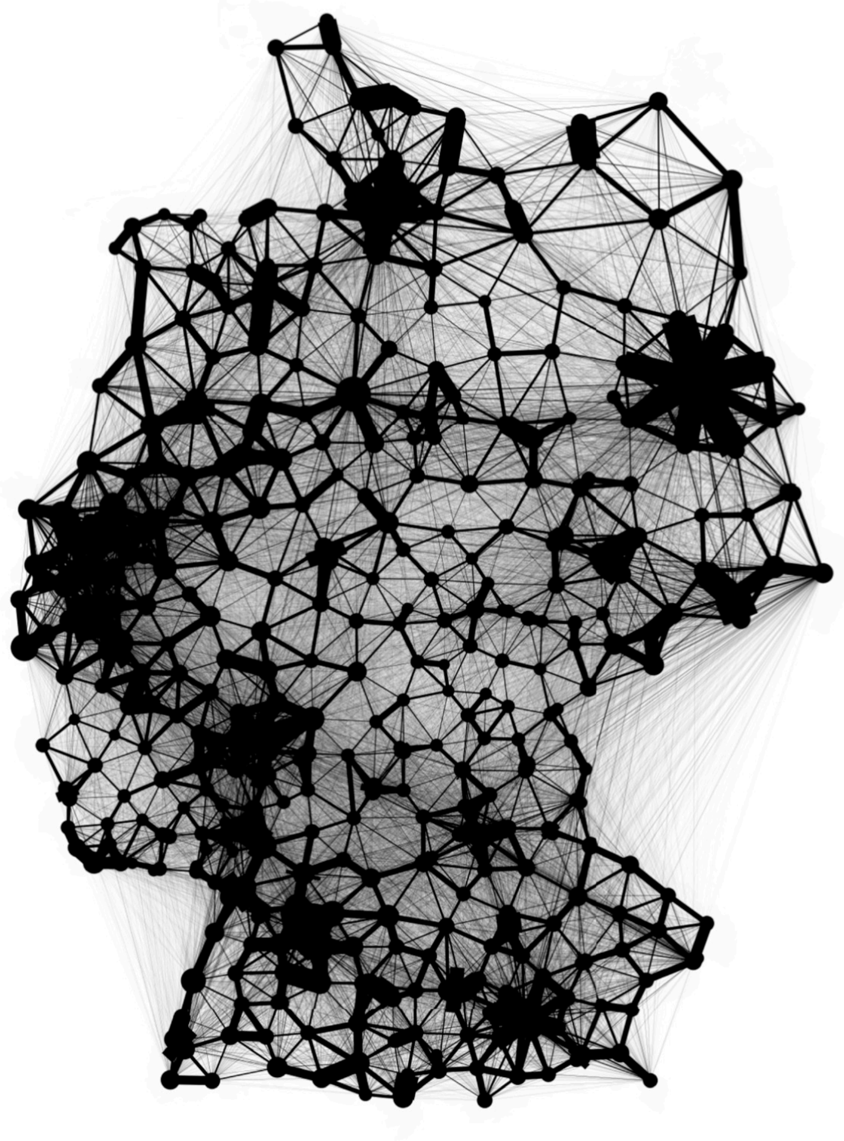
Paris inner admin. boundaries





German provinces

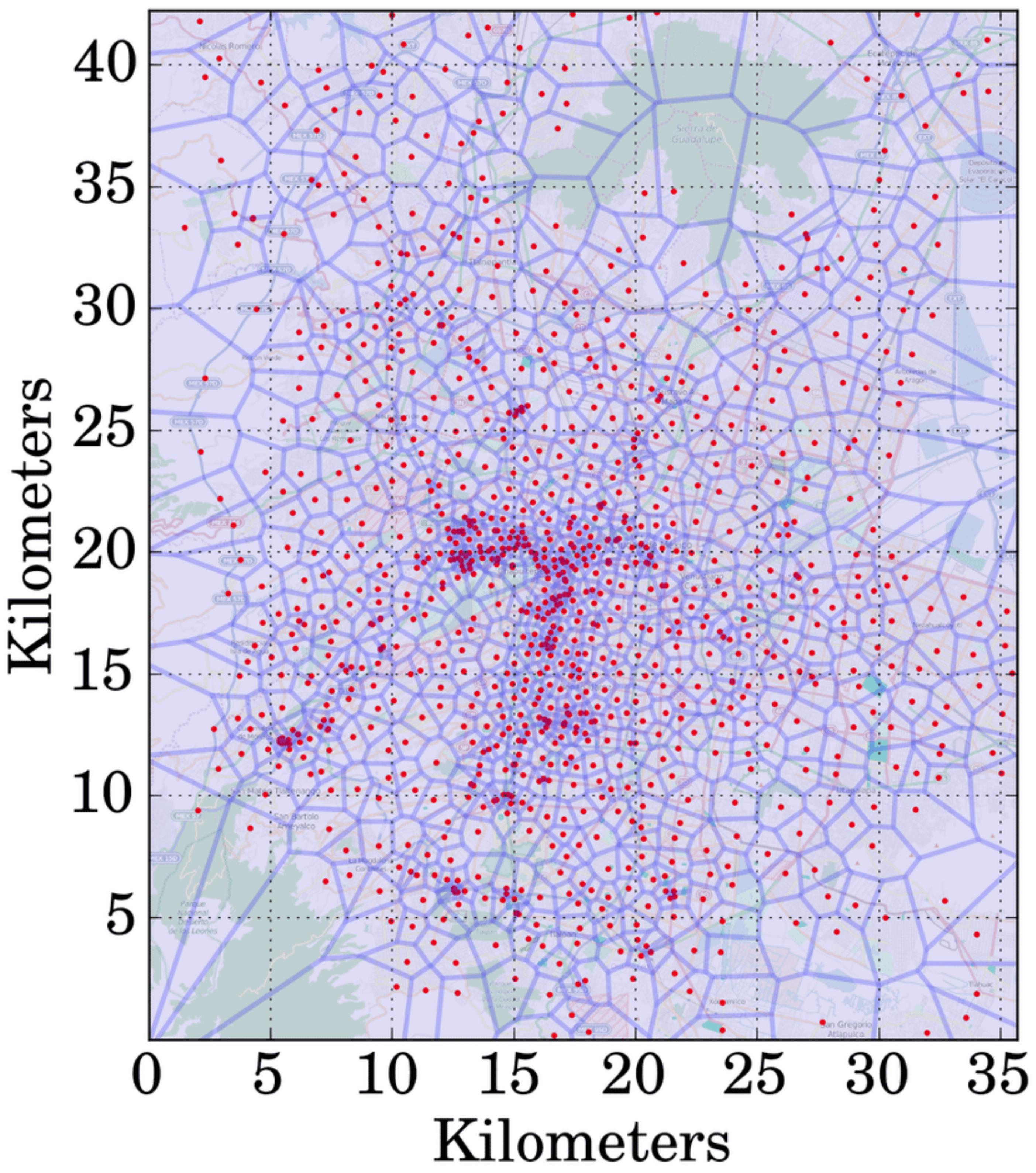
Provinces trips



Louf, Rémi. "Wandering in cities: A statistical physics approach to urban theory." *PhD Thesis* (2015).
Pappalardo, L., Vanhoof, M., Gabrielli, L. *et al.* An analytical framework to nowcast well-being using mobile phone data. *Int J Data Sci Anal* **2**, 75–92 (2016)
Schlosser, Frank, et al. "COVID-19 lockdown induces disease-mitigating structural changes in mobility networks." *PNAS* 117.52 (2020): 32883-32890.

Mobile phone data bias

Urban-rural divide

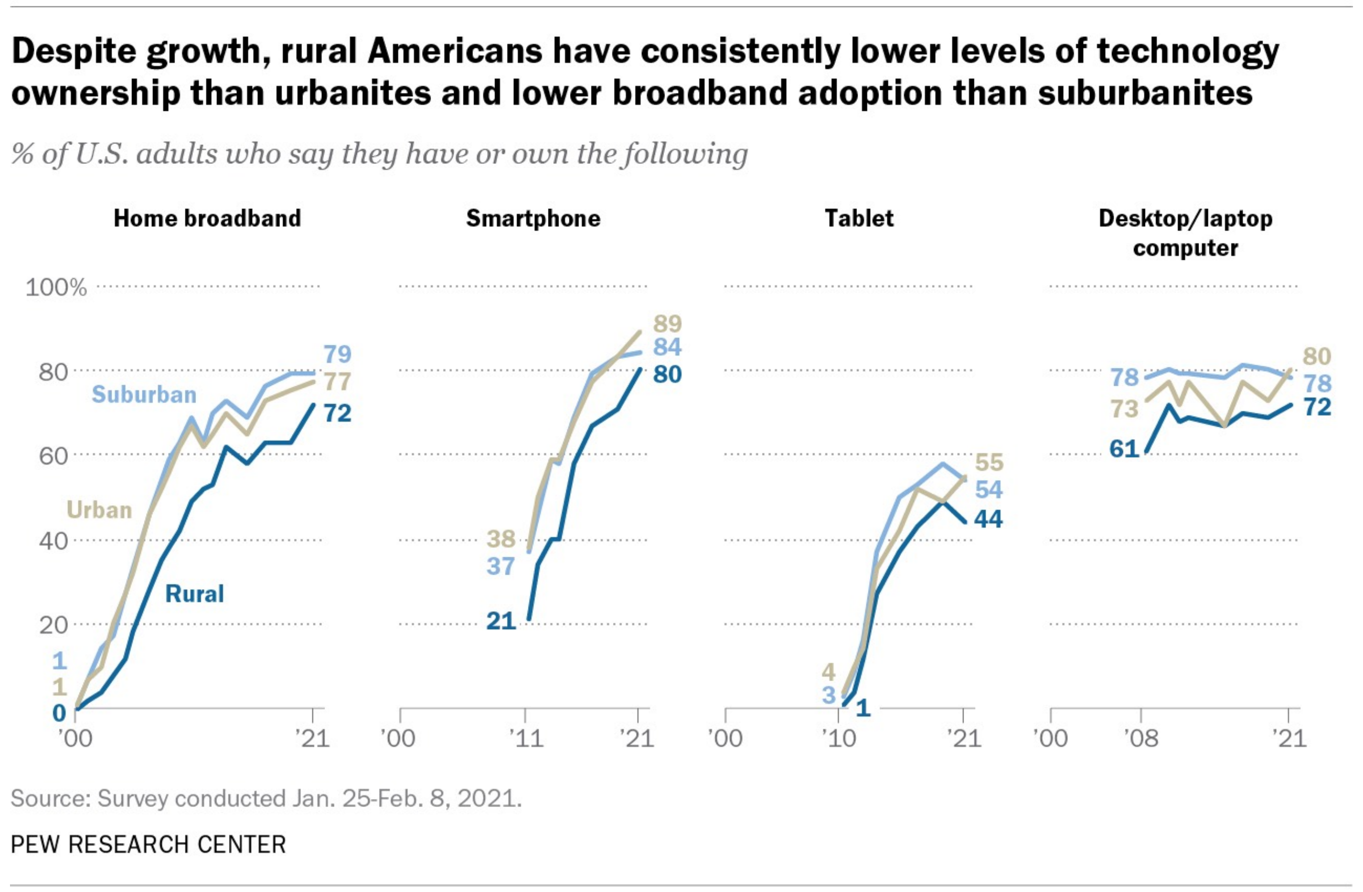


Heterogeneous cell towers spatial distribution (urban vs rural)

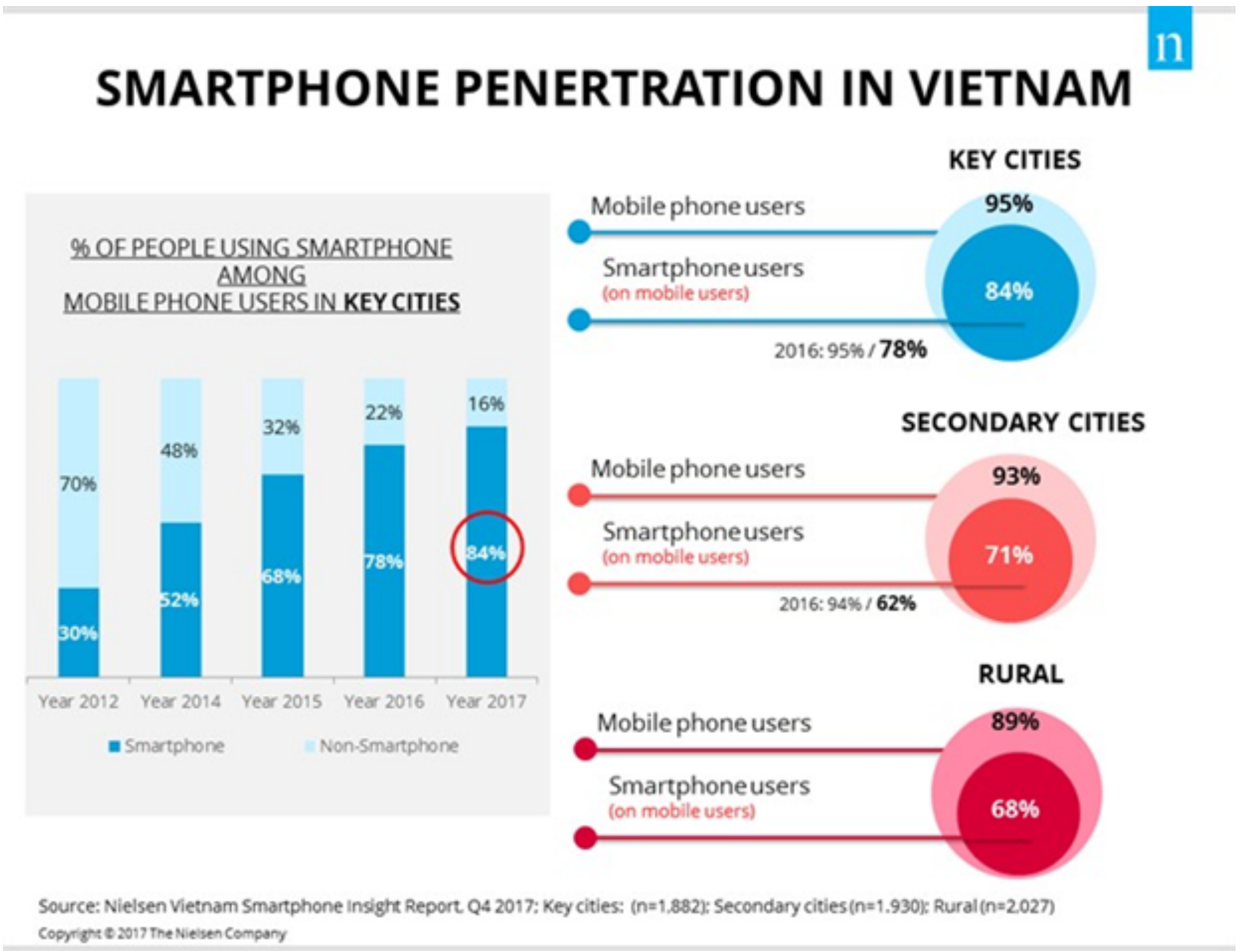
High resolution in densely populated areas
Low spatial resolution in rural areas

Examples of urban-rural divides

US

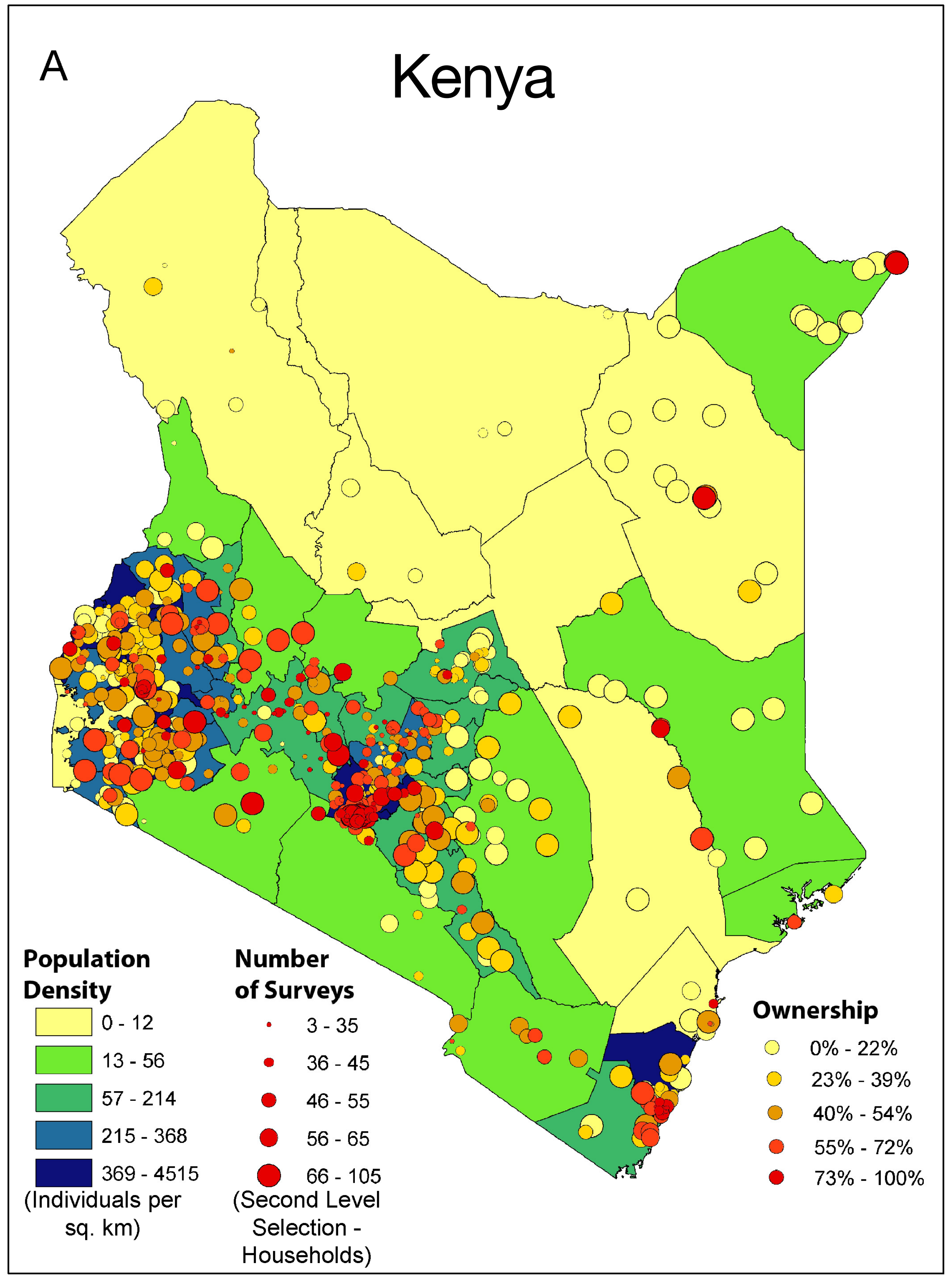


Vietnam



Chen, Guangshuo, et al. "Enriching sparse mobility information in call detail records." *Computer Communications* 122 (2018): 44-58.
https://www.pewresearch.org/short-reads/2021/08/19/some-digital-divides-persist-between-rural-urban-and-suburban-america/

Mobile phone data bias



Different levels of tech adoption translates to **BIAS**

The urban-rural divide in smartphones adoption is one type of bias, let's see more

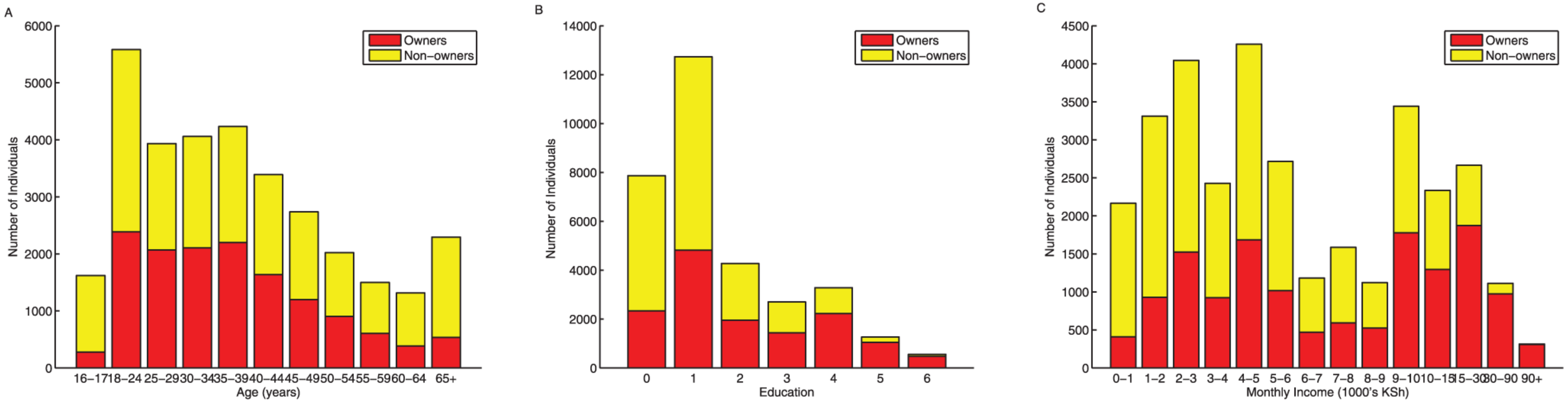
- Socio-economic status
- Gender
- Ethnicity
- Urban vs rural areas

Usually no information on owners' traits
No de-biasing possible without parallel dataset (surveys)
Dataset often shared already aggregated

Wesolowski, Amy, et al. "Heterogeneous mobile phone ownership and usage patterns in Kenya." *PloS one* 7.4 (2012): e35319.

Mobile phone data bias

Mobile phone ownership biases by socio-demographics in Kenya



The impact of biases in mobile phone ownership on estimates of human mobility

Amy Wesolowski[1], Nathan Eagle[2,3], Abdisalan M. Noor[4,5], Robert W. Snow[4,5] and Caroline O. Buckee[3,6]

"[…] Mobility estimates are surprisingly robust to the substantial biases in phone ownership across different geographical and socioeconomic groups."

Wesolowski, Amy, et al. "Heterogeneous mobile phone ownership and usage patterns in Kenya." *PloS one* 7.4 (2012): e35319.
Wesolowski, Amy, et al. "The impact of biases in mobile phone ownership on estimates of human mobility." *Journal of the Royal Society Interface* 10.81 (2013): 20120986.

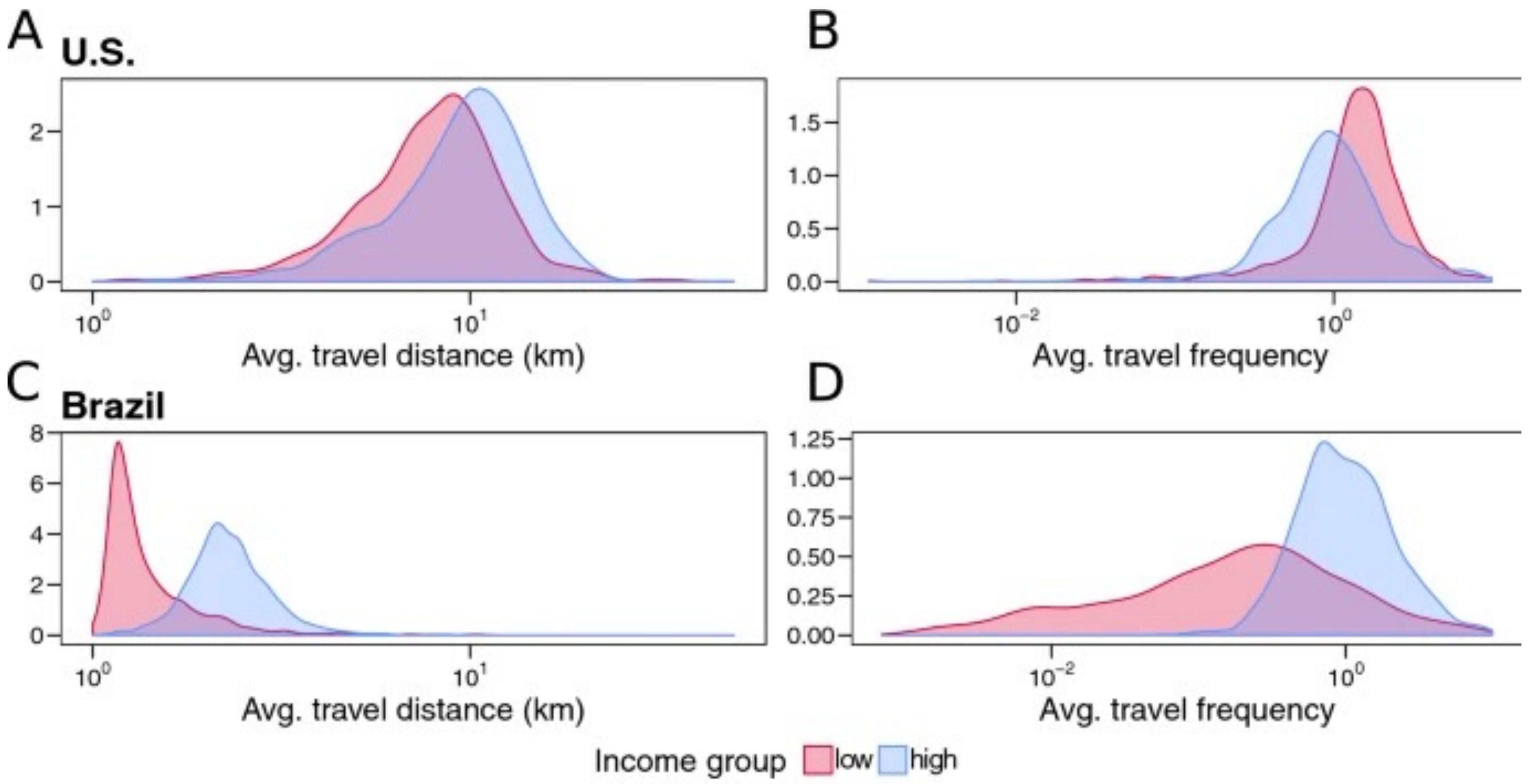
Inequalities in mobility

At the individual level: income, gender and age impact on the traveled distance and frequency of trips.

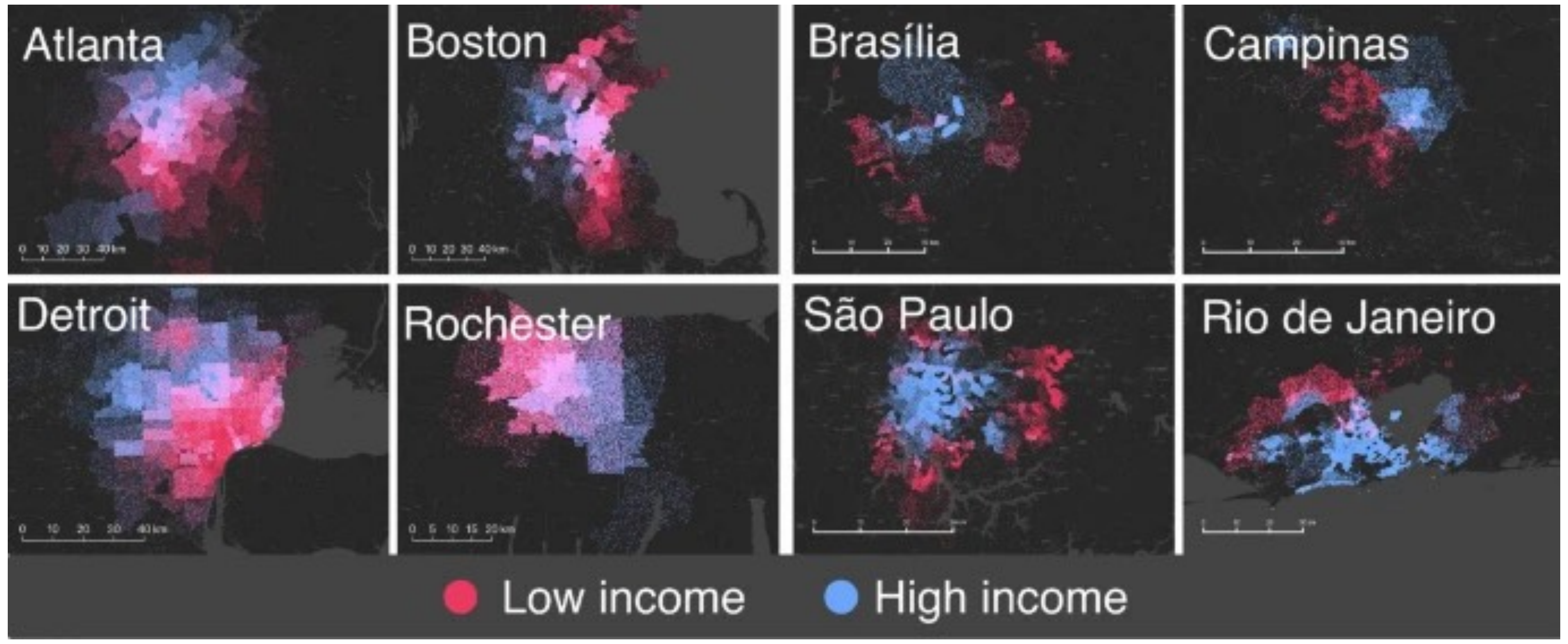
Mobile ownership skewed towards wealthier demographic strata:

- high income individuals, young adults, males are overrepresented in mobile phone data

OK to use the aggregate metrics for epi modeling, caution is required when using individual traces



Another example of inequalities in mobility

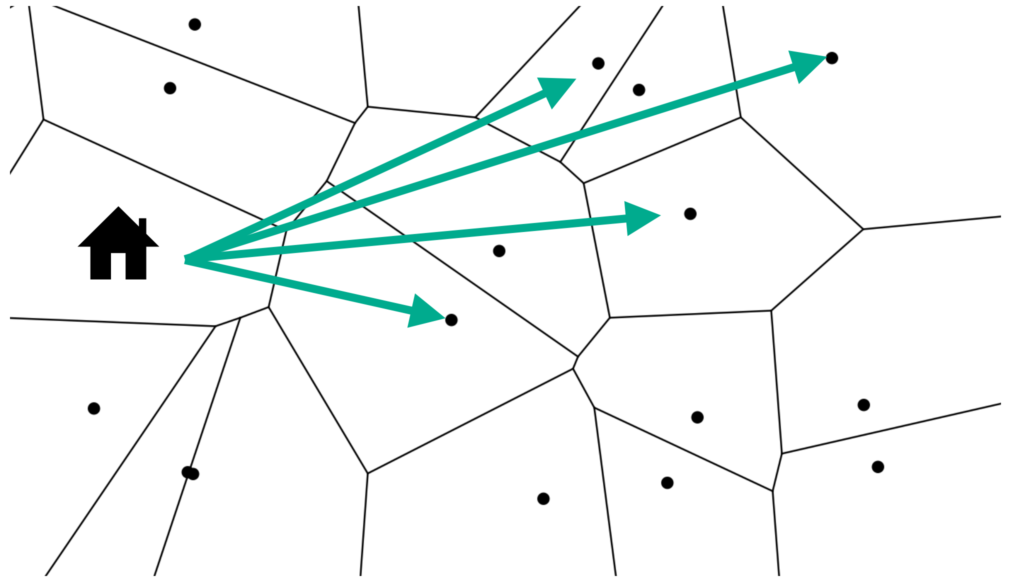


Barbosa, H., Hazarie, S., Dickinson, B. *et al.* Uncovering the socioeconomic facets of human mobility. *Sci Rep* **11**, 8616 (2021).

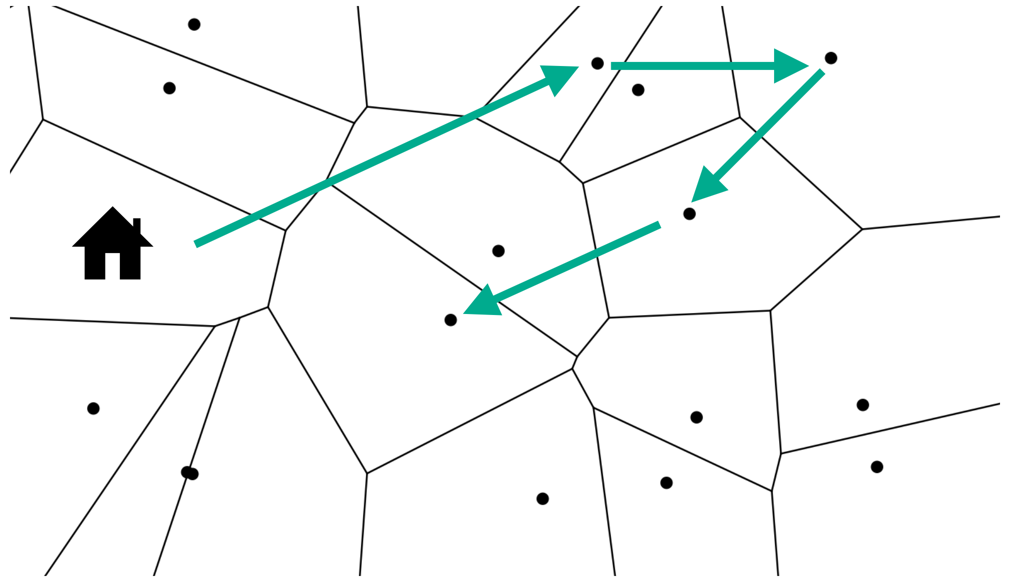
Trips aggregation to OD matrix

Individual data

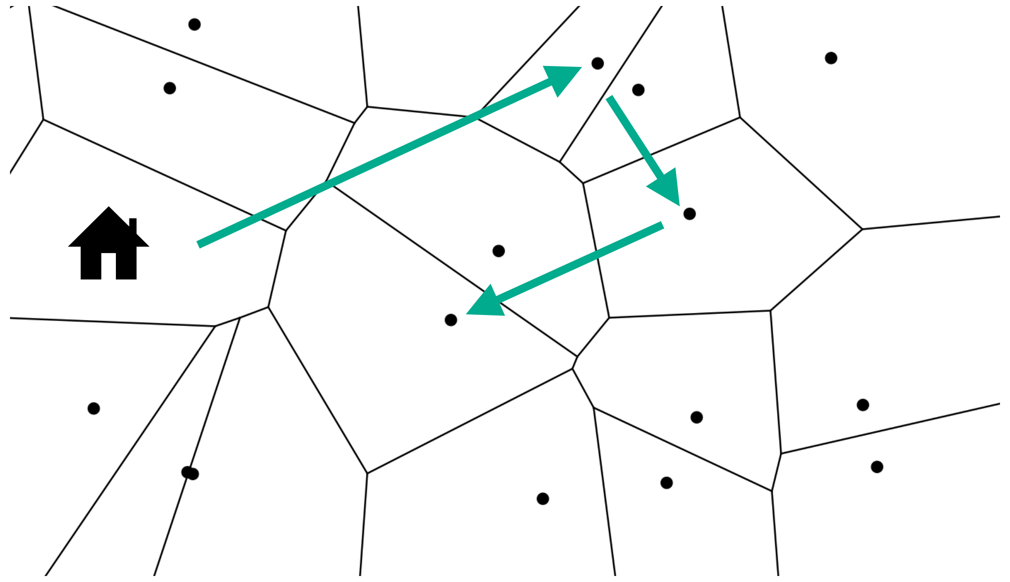
Home to stops



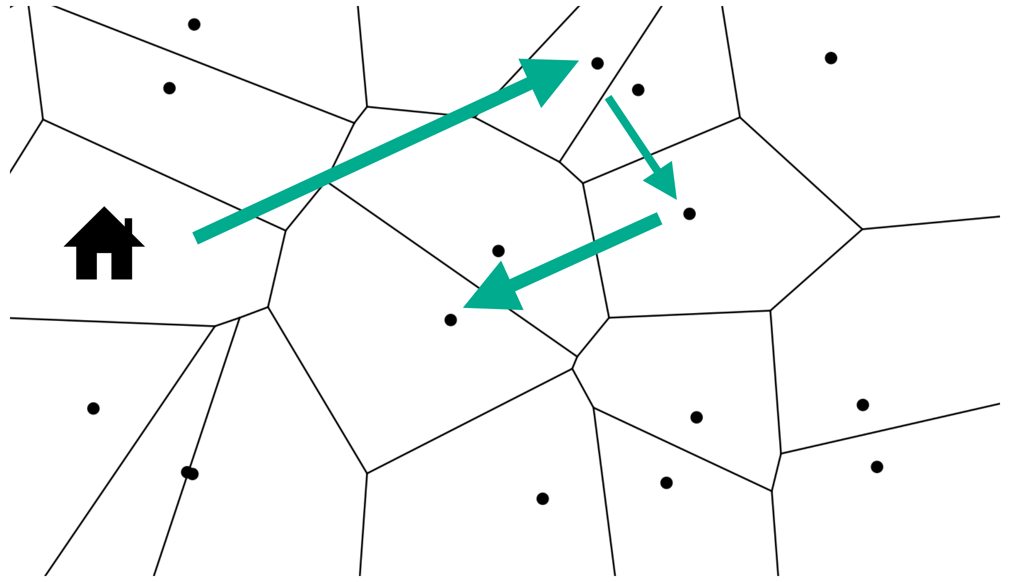
Consecutive trips



Consecutive stays

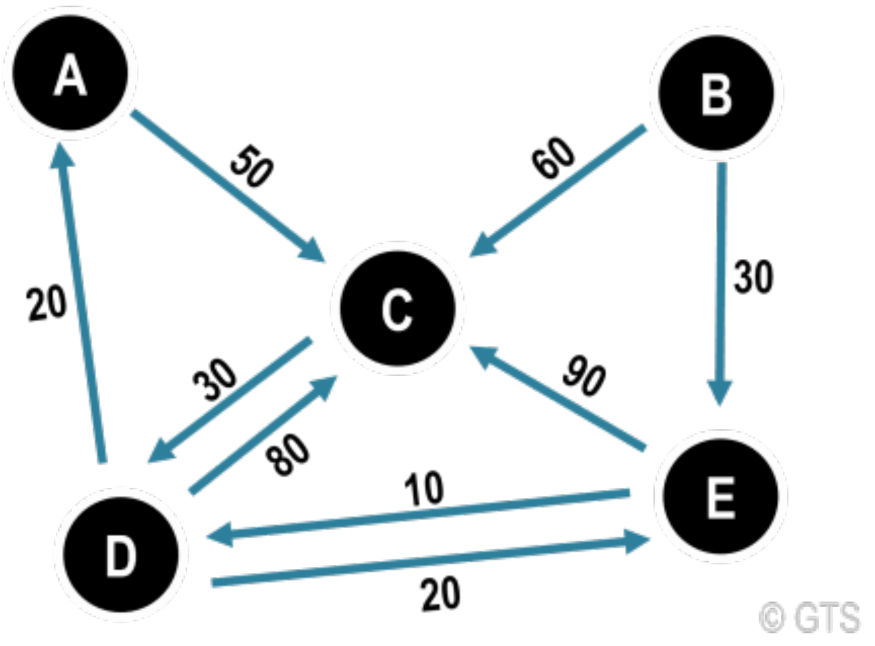


Stay time weighted trips



Population level

SPATIAL INTERACTIONS



O/D MATRIX

	A	B	C	D	E	Ti
A	0	0	50	0	0	50
B	0	0	60	0	30	90
C	0	0	0	30	0	30
D	20	0	80	0	20	120
E	0	0	90	10	0	100
Tj	20	0	280	40	50	**390**

OD matrix depends on type of aggregation

Provided data from Telcos comes already aggregated

Implications for epidemic modeling

Universal standards missing

Mobile phone data (CDR & XDR)

PROS

- High representativity of the population wrt other data sources
- High spatial resolution in urban areas
- Aggregated features robust to ownership bias
- Useful at multiple scales, e.g. municipality, provinces, regions
- High temporal resolution for XDR

CONS

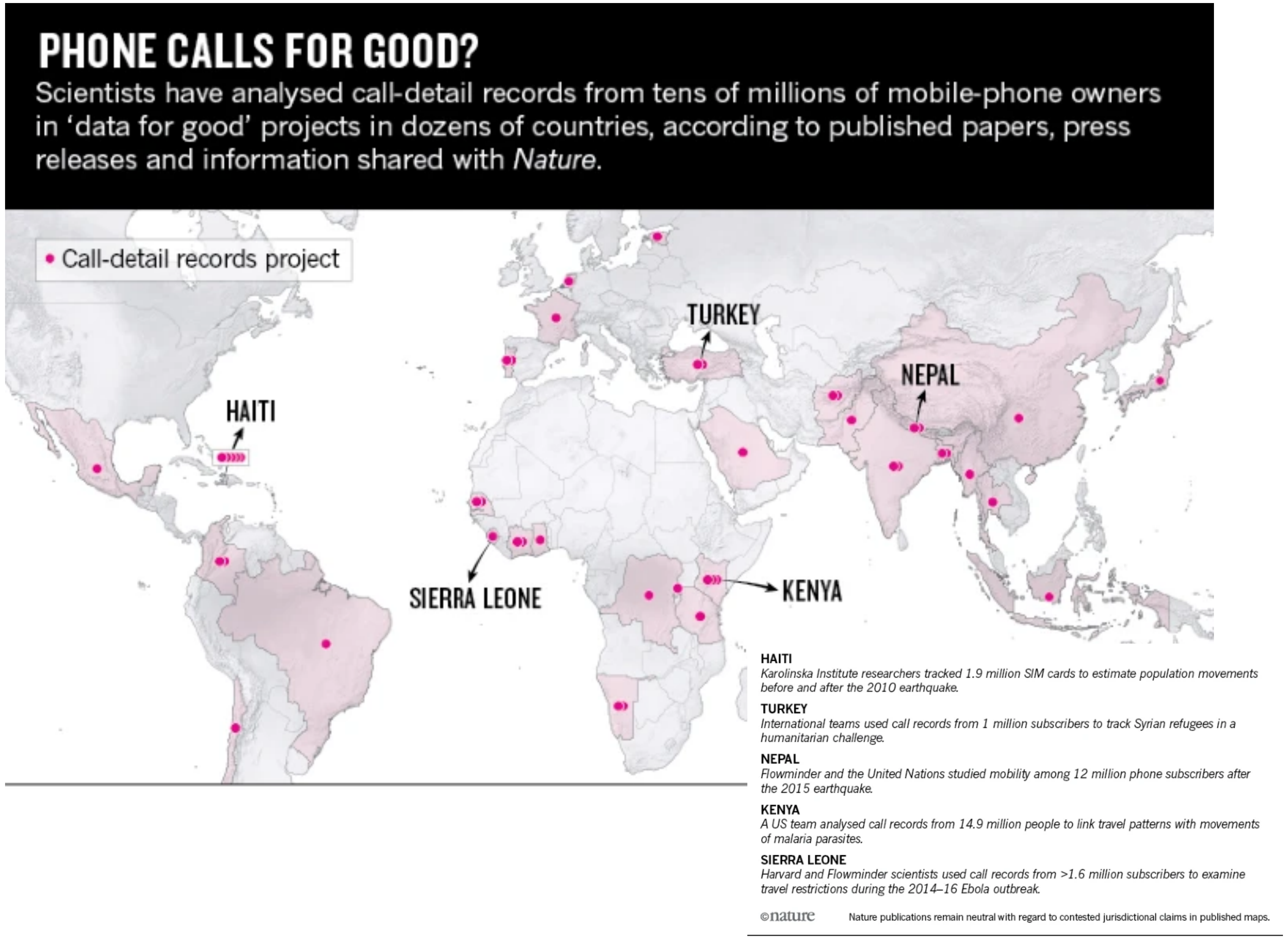
- Low spatial resolution in rural areas
- Low temporal resolution for CDR
- Country segmented market
- Not available in most non-Western countries
- Private, expensive
- Real-time data stream depends on previous collaboration
- Provided already aggregated, no power on type of aggregation
- Device ownership bias, especially in low & middle income countries

USAGE

- Inform sub-national spatial transmission models

SCALE

- Cell towers, municipalities, provinces, regions



Maxmen, Amy. "Can tracking people through phone-call data improve lives?." *Nature* 569.7758 (2019): 614-618.

International travel (IATA air travel data)

IATA: international air travel agency

PROS

- Highly representative of international mobility
- Almost all countries included

CONS

- Pre-defined spatial resolution (airports)
- Private, expensive
- Pre-set temporal scale (usually 1 month)
- Tech access bias (only air traffic)
- Not capturing cross-border mobility in neighbouring countries
- Trips estimated by revenues and sold tickets

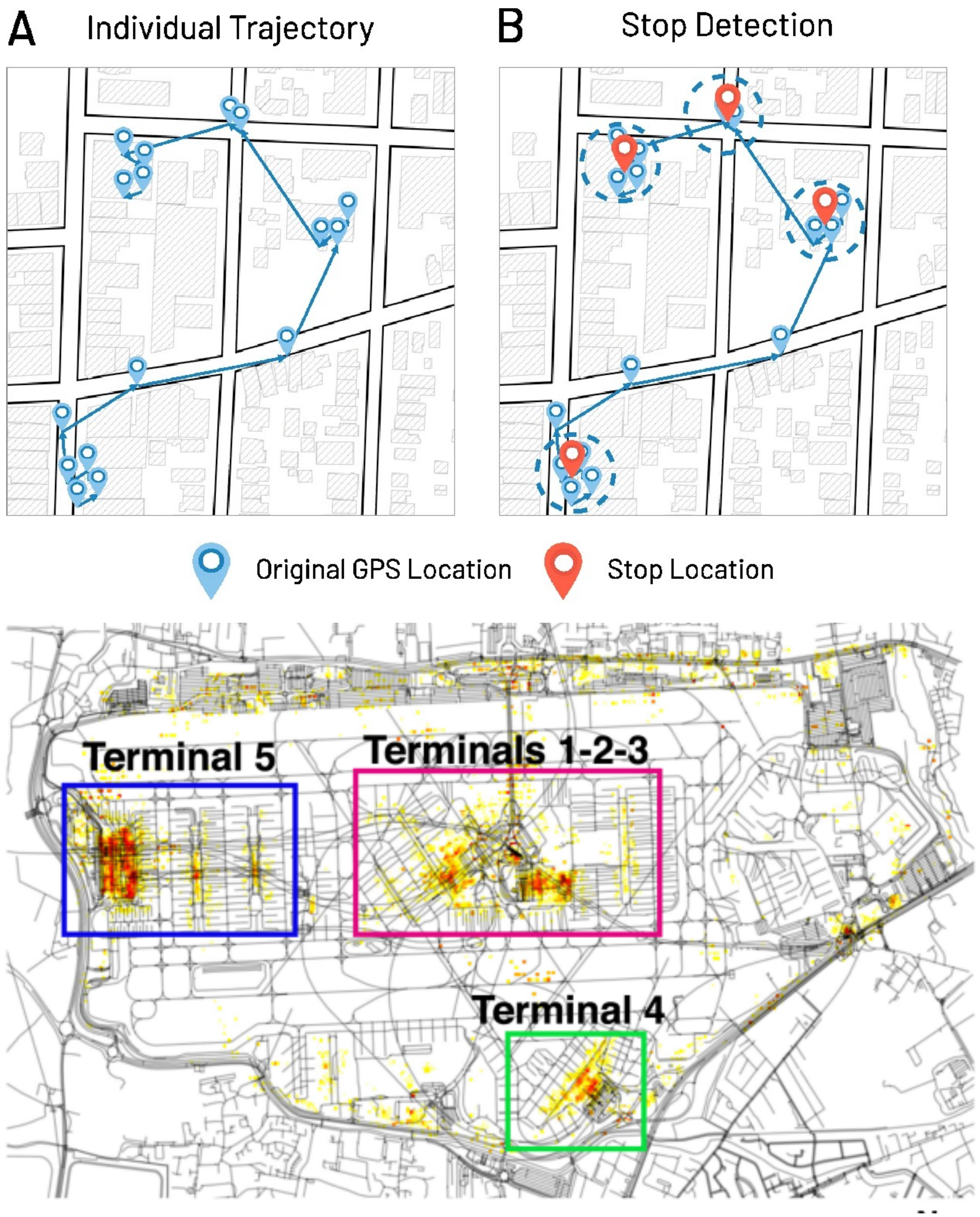
USAGE

- Inform epidemic importation models

SCALE

- Country, airports catchment areas

GPS traces (Cuebiq / Spectus)



PROS

- Highest spatial and temporal resolution (5m - 5')
- Free and accessible in pandemic period (Cuebiq Data for Good program)
- Include all modes of transportation

CONS

- Very low representativity of the population
- Tech adoption bias
- Available mostly in most Western countries
- Private, expensive (out of pandemic period)

USAGE

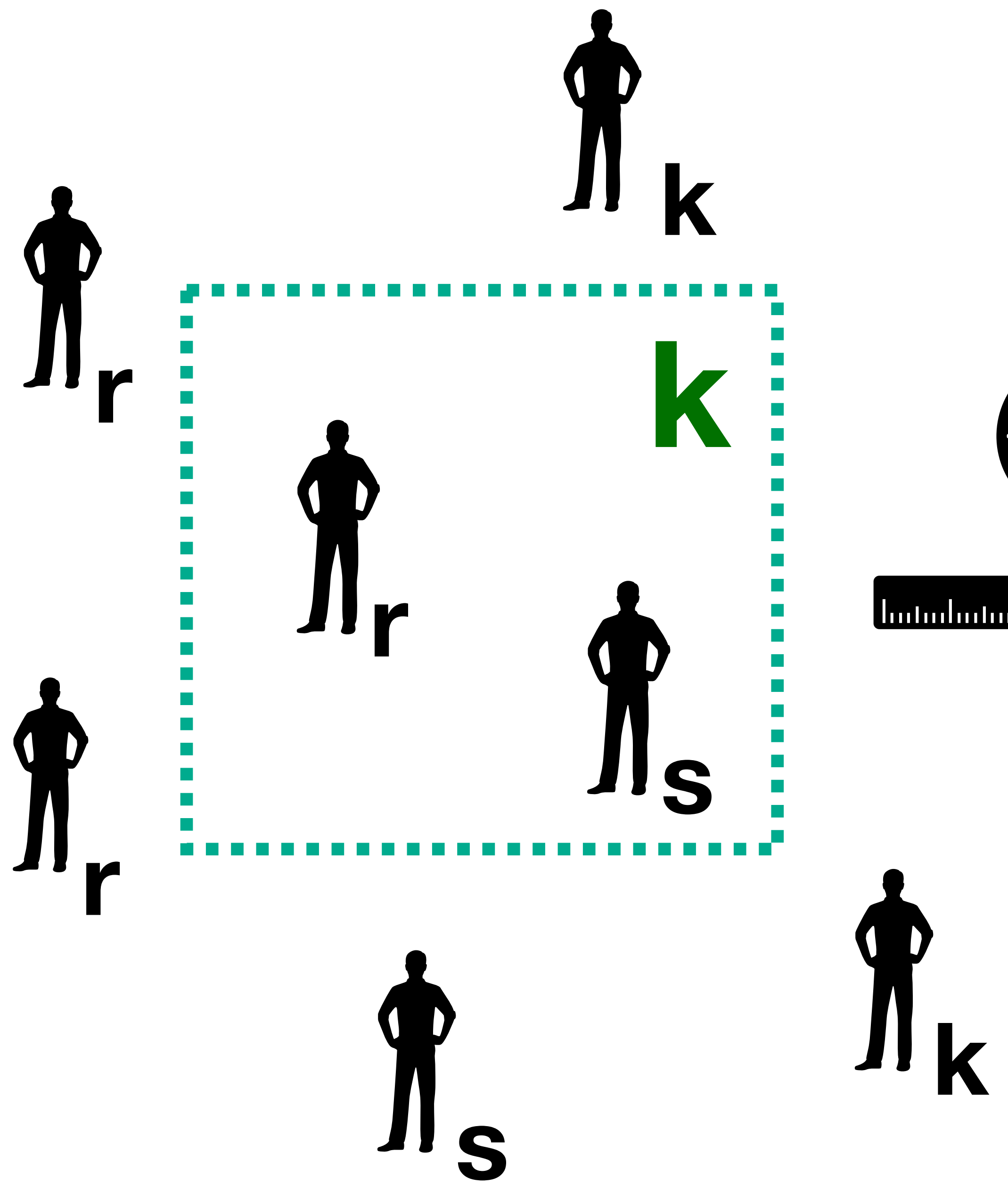
- Inform agent based models in specific settings

SCALE

- Latitude, longitude
- Scalable to lower spatial resolution

Lucchini, L., Centellegher, S., Pappalardo, L. *et al.* Living in a pandemic: changes in mobility routines, social activity and adherence to COVID-19 protective measures. *Sci Rep* **11**, 24452 (2021).
Mazzoli, M., Gallotti, R., Privitera, F. *et al.* Spatial immunization to abate disease spreading in transportation hubs. *Nat Commun* **14**, 1448 (2023).

GPS traces (META Co-location)



5 min

600m x 600 m

Co-location detected at 600x600 sq meters, 5 minutes duration
Provided at province/regional and weekly scale

$$m_{rs} = \sum_{ij} X_{ijr} X_{ijs}$$

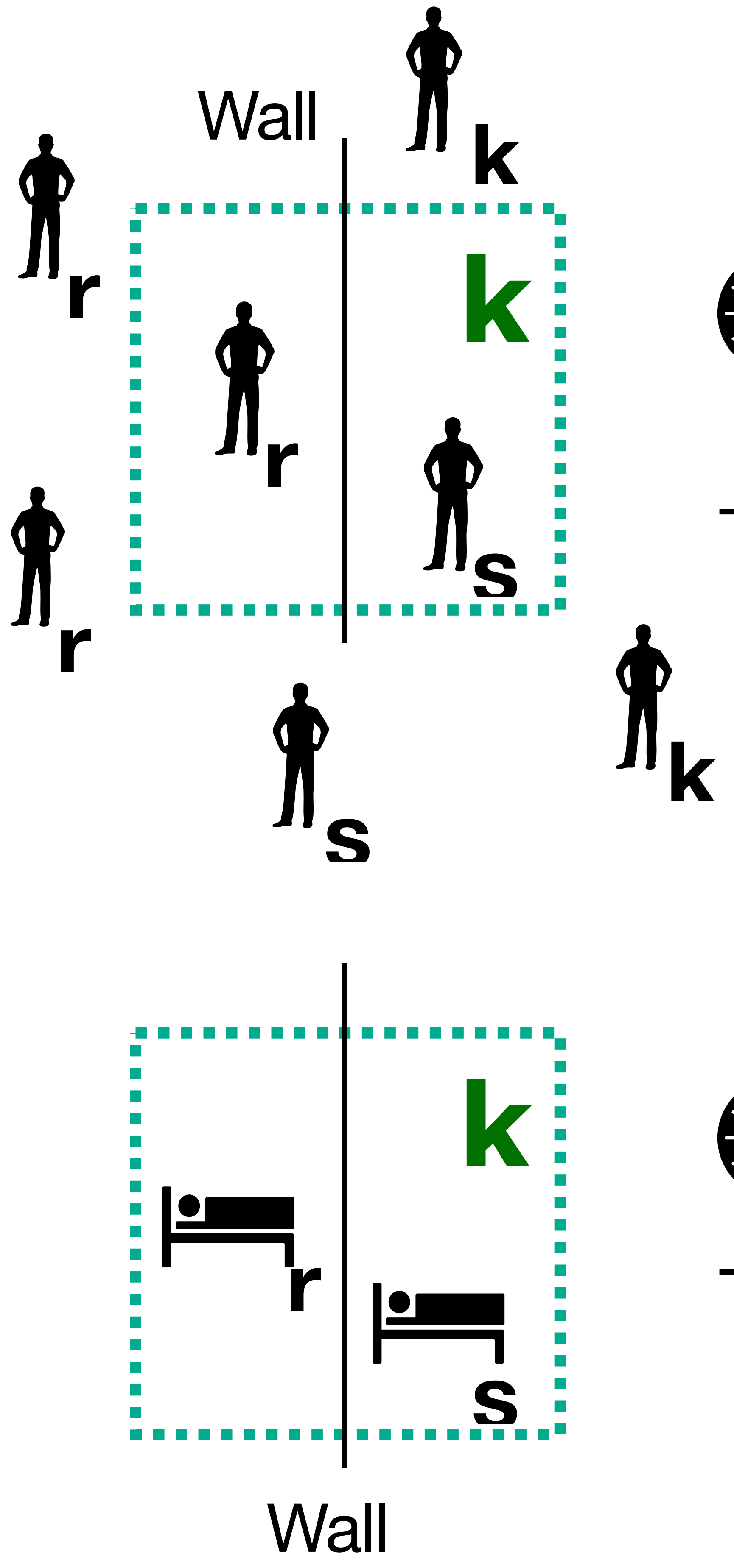
n of contacts between residents of r and s
occurring in one week, in place i at time slot j

$$p_{rs} = \frac{1}{2016} \frac{m_{rs}}{n_r n_s}$$

mixing probability of residents of r and s $\in [0,1]$
n_r, n_s = sample of residents in r and in s
2016 = n of 5min time slots in one week

Iyer, Shankar, et al. "Large-scale measurement of aggregate human colocation patterns for epidemiological modeling." *Epidemics* 42 (2023): 100663.

GPS traces (META Co-location)

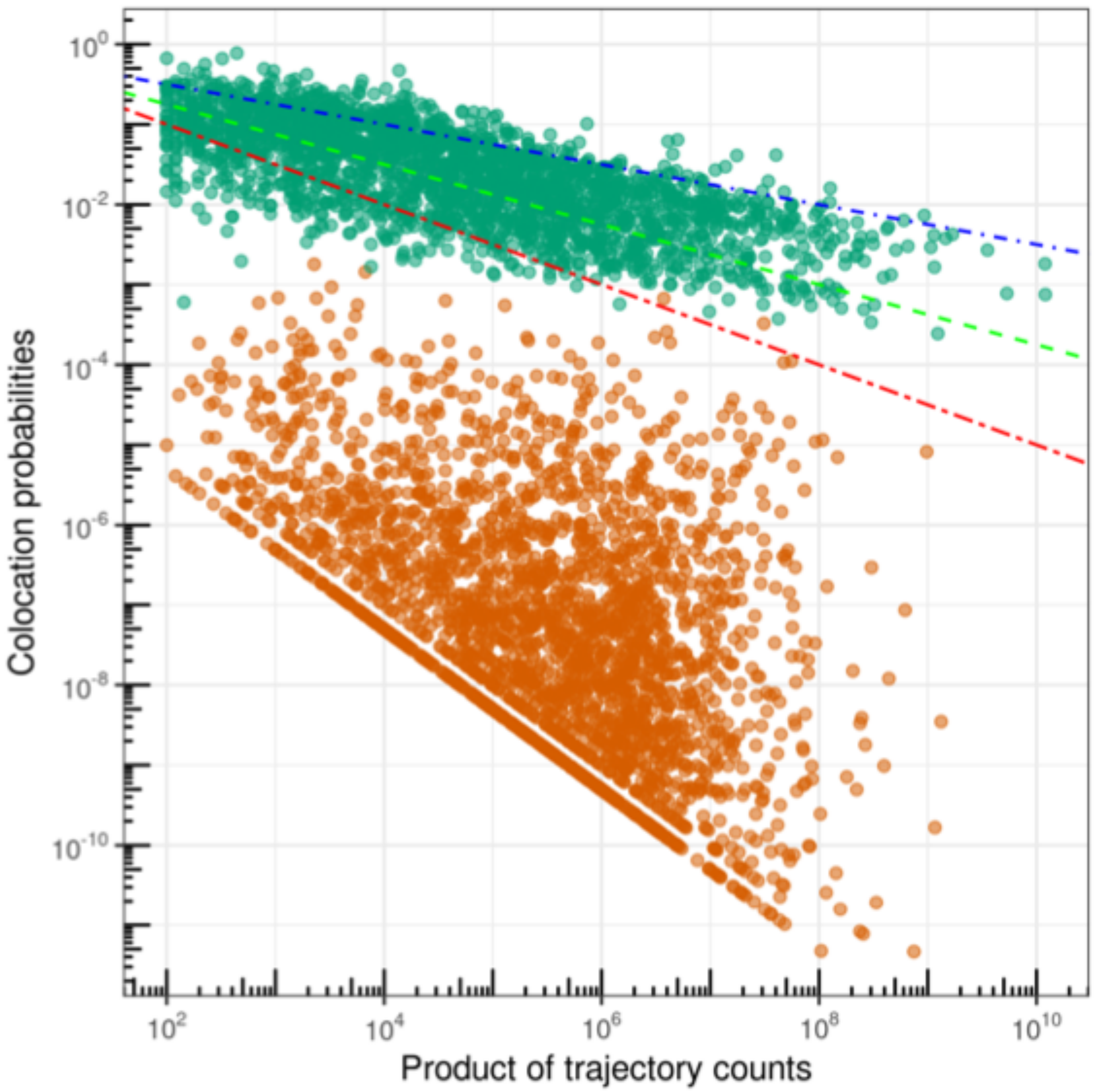


11 AM

This still counts as a co-location

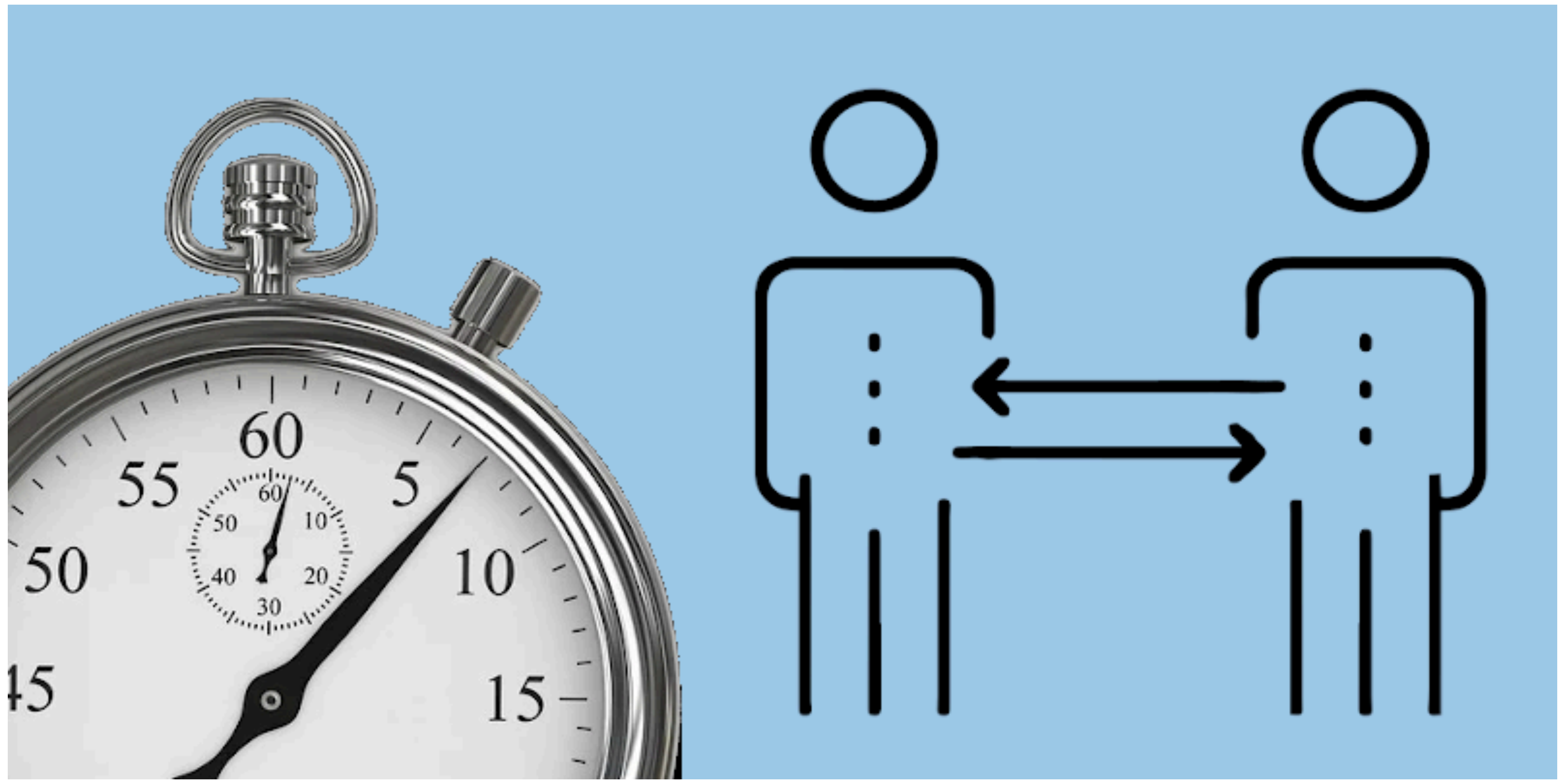
3 AM

This still counts as a co-location



Iyer, Shankar, et al. "Large-scale measurement of aggregate human colocation patterns for epidemiological modeling." *Epidemics* 42 (2023): 100663.

GPS traces (META Co-location)



PROS

- Free and accessible (Meta Data for Good program)
- Include all modes of transportation
- Proxy contacts among resident populations
- Available in many world regions

CONS

- Pre-aggregated data
- Co-location can occur anywhere
- Spatial resolution country dependent
- Pre-set temporal resolution (weekly level)
- Overestimated internal mixing, due to spurious co-locations

USAGE

- Inform sub-national spatial transmission models

SCALE

- Provinces or regions, depending on country

Iyer, Shankar, et al. "Large-scale measurement of aggregate human colocation patterns for epidemiological modeling." *Epidemics* 42 (2023): 100663.

Surveys

PROS

- Usually free and accessible
- Richness of metadata (age, gender, job, wealth, mode of transport)
- Optimal to target specific communities (medical conditions, migrants, etc)

CONS

- Scarse spatio-temporal resolution
- Non-representative of population
- Response bias

USAGE

- Determinants of mobility, transport mode adoption, sustainable mobility

SCALE

- Municipalities, census areas

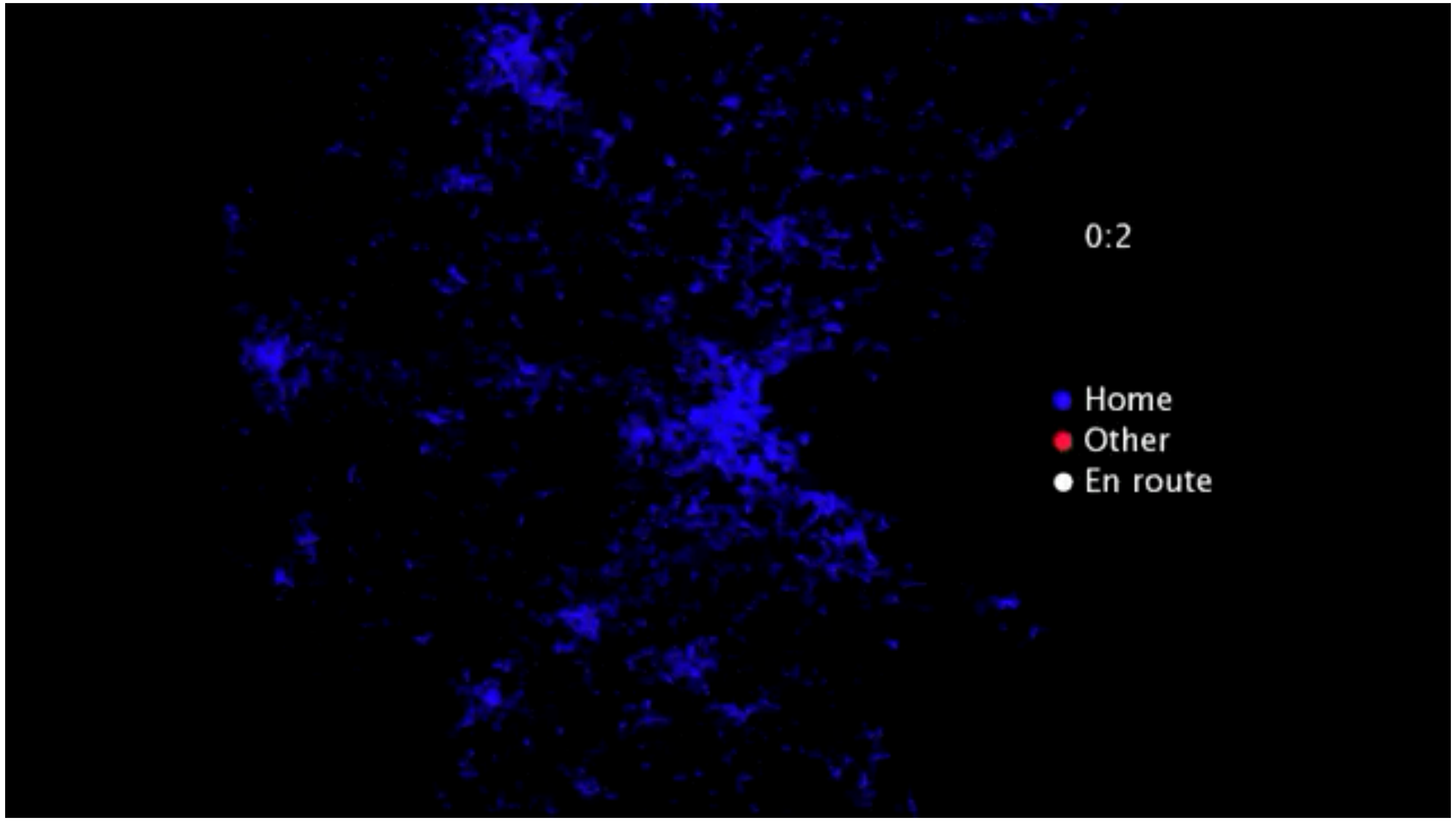
Human mobility modeling

Collective models

- Gravity model
- Radiation model

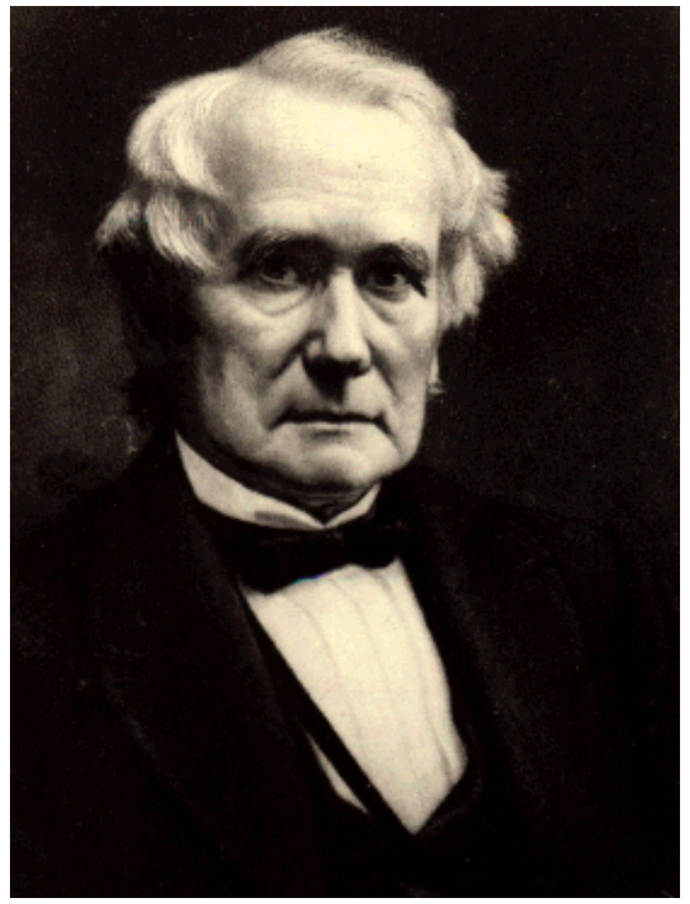
Individual models

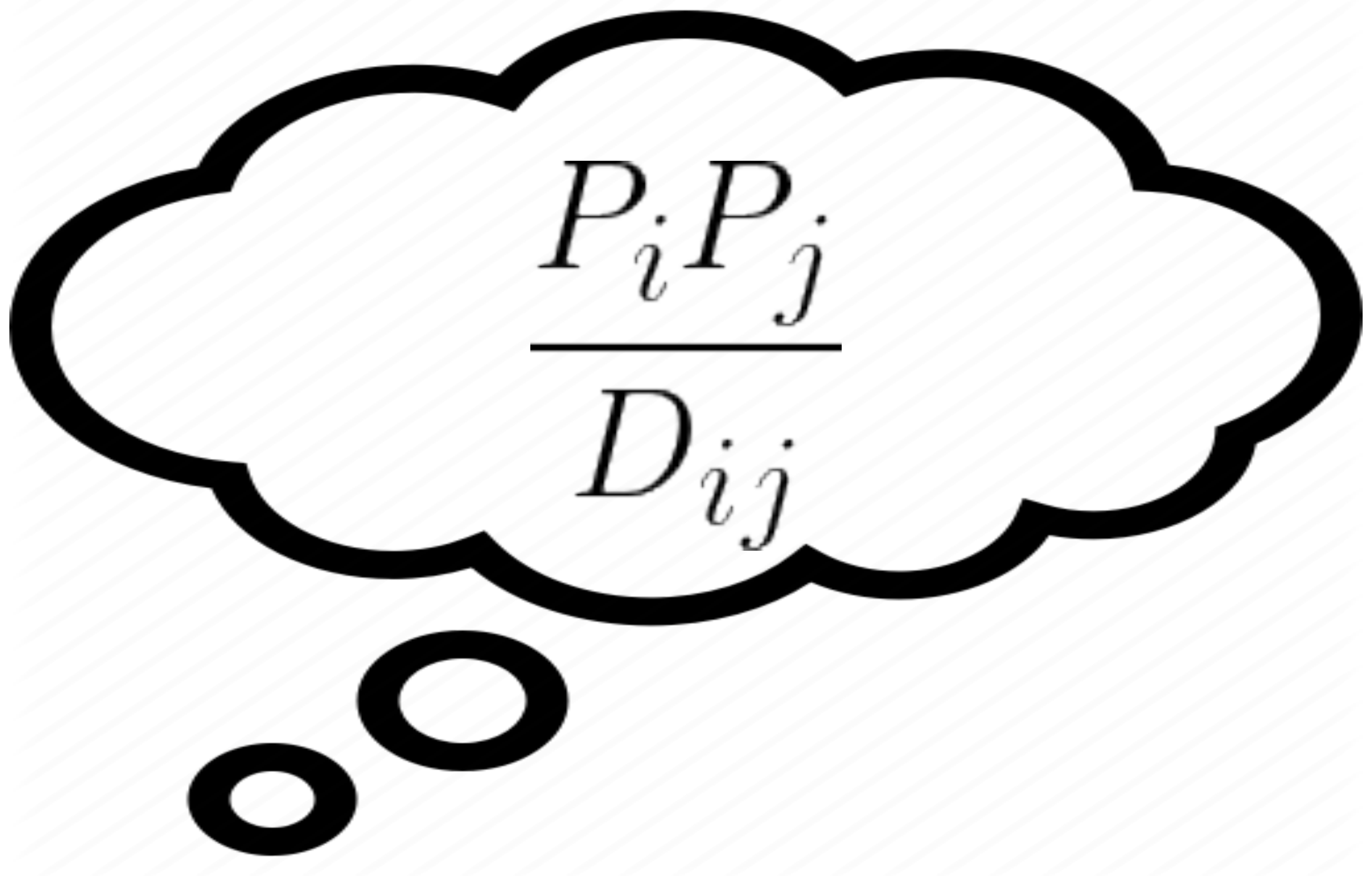
- EPR model (aka Song model)
- MATsim



Marta Gonzalez, YouTube

Gravity model

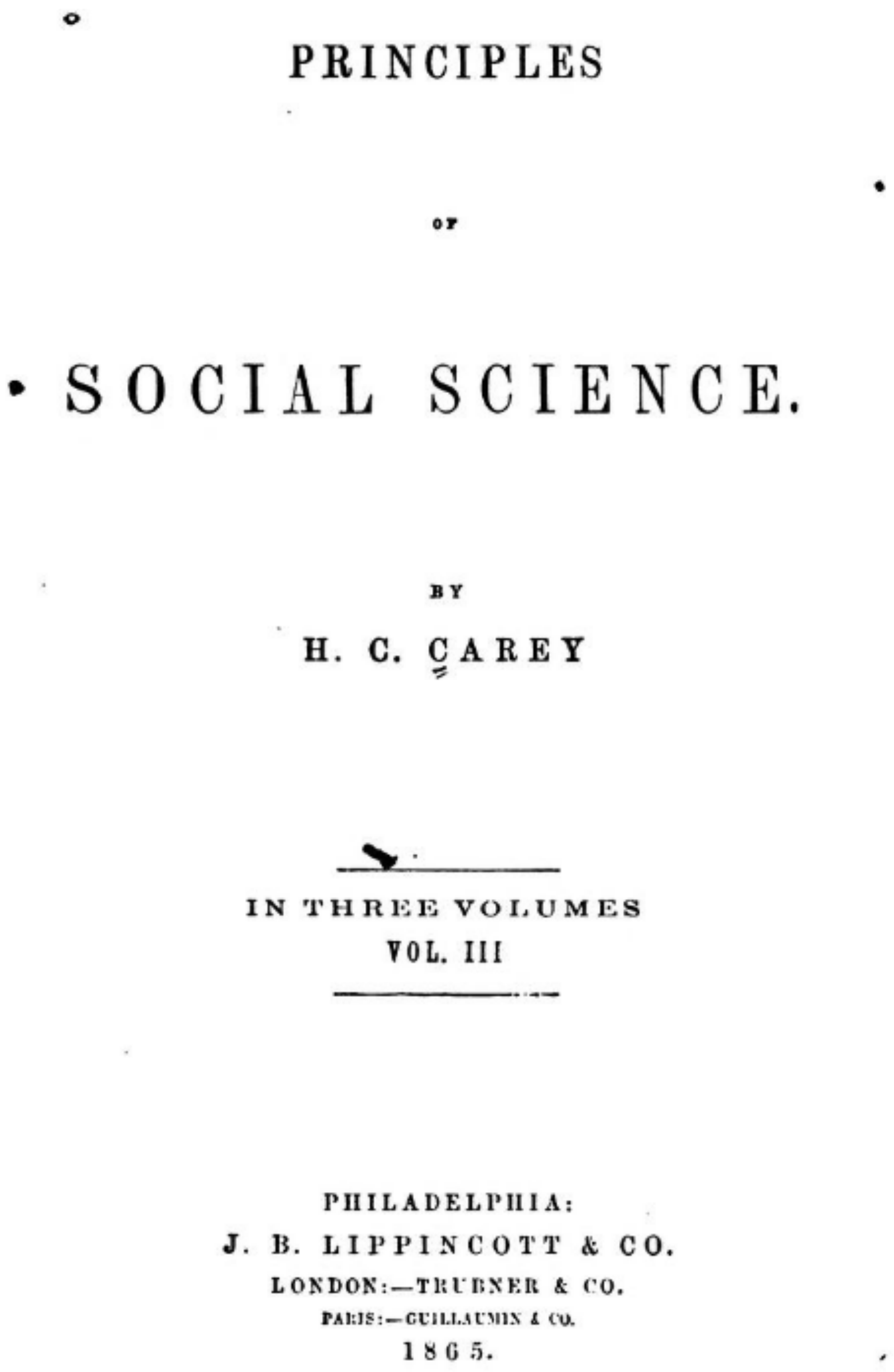




H.C. Carey (1865)
US economist & economic adviser of Abraham Lincoln

“Man tends of necessity to gravitate towards his fellow-man... and the greater the number collected (of man, ndr) in a given space the greater is the attractive force there exerted…"

“Gravitation is here, as everywhere else in the material world, in the direct ratio of the mass, and in the inverse one of the distance”



PRINCIPLES

OF

SOCIAL SCIENCE.

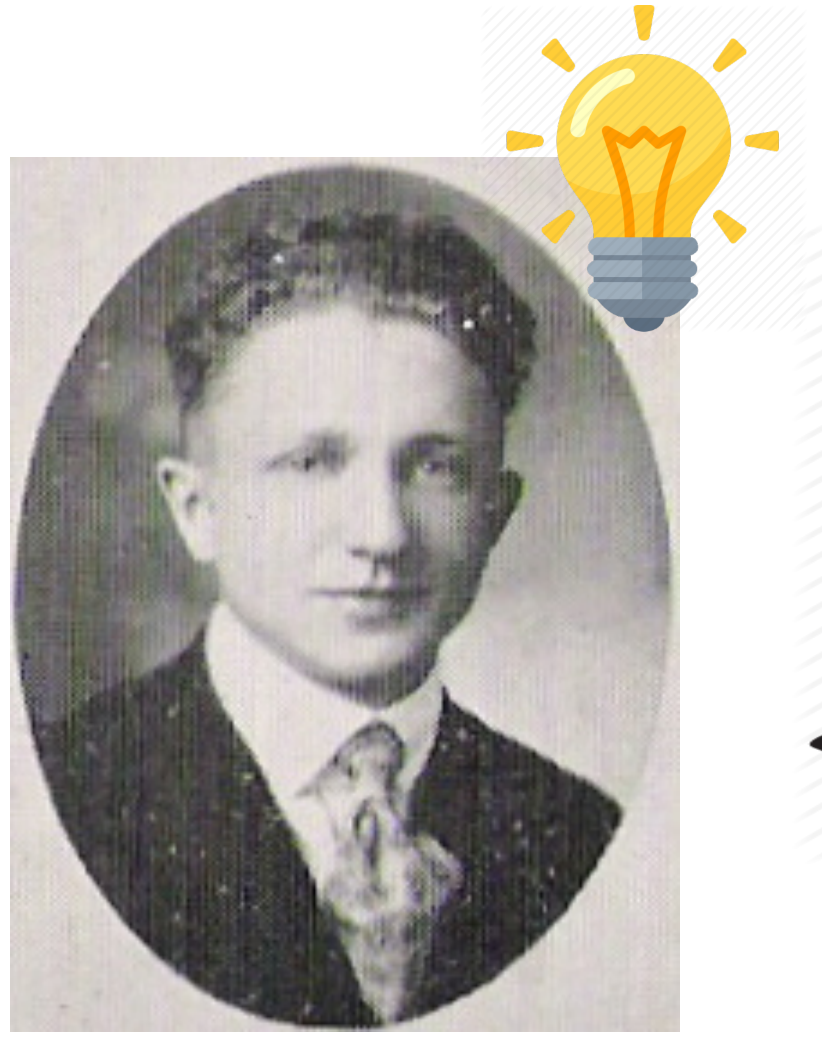
BY

H. C. CAREY

IN THREE VOLUMES
VOL. III

PHILADELPHIA:
J. B. LIPPINCOTT & CO.
LONDON:—TRUBNER & CO.
PARIS:—GUILLAUMIN & CO.
1865.

Gravity model

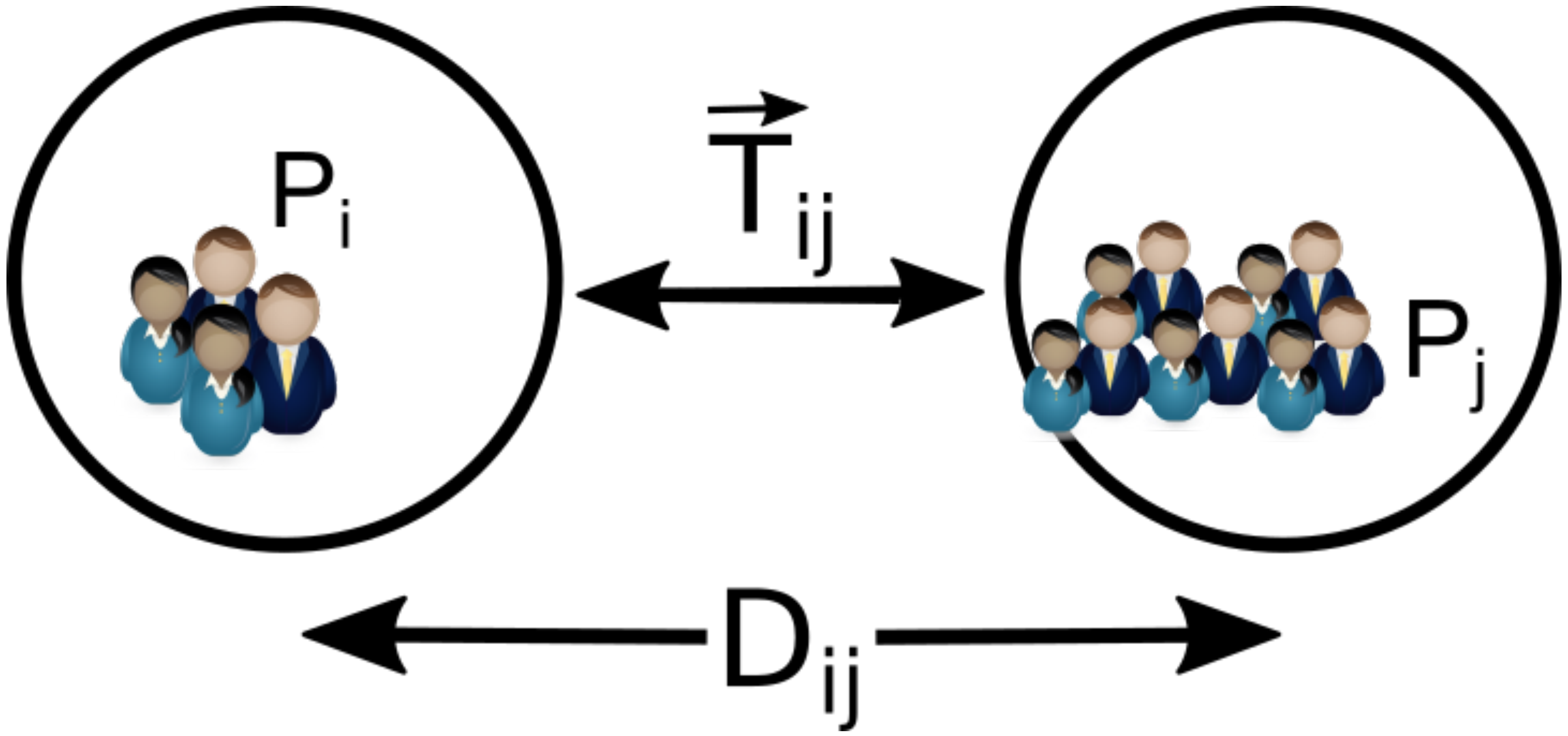


$$T_{ij} \propto \frac{P_i P_j}{D_{ij}}$$

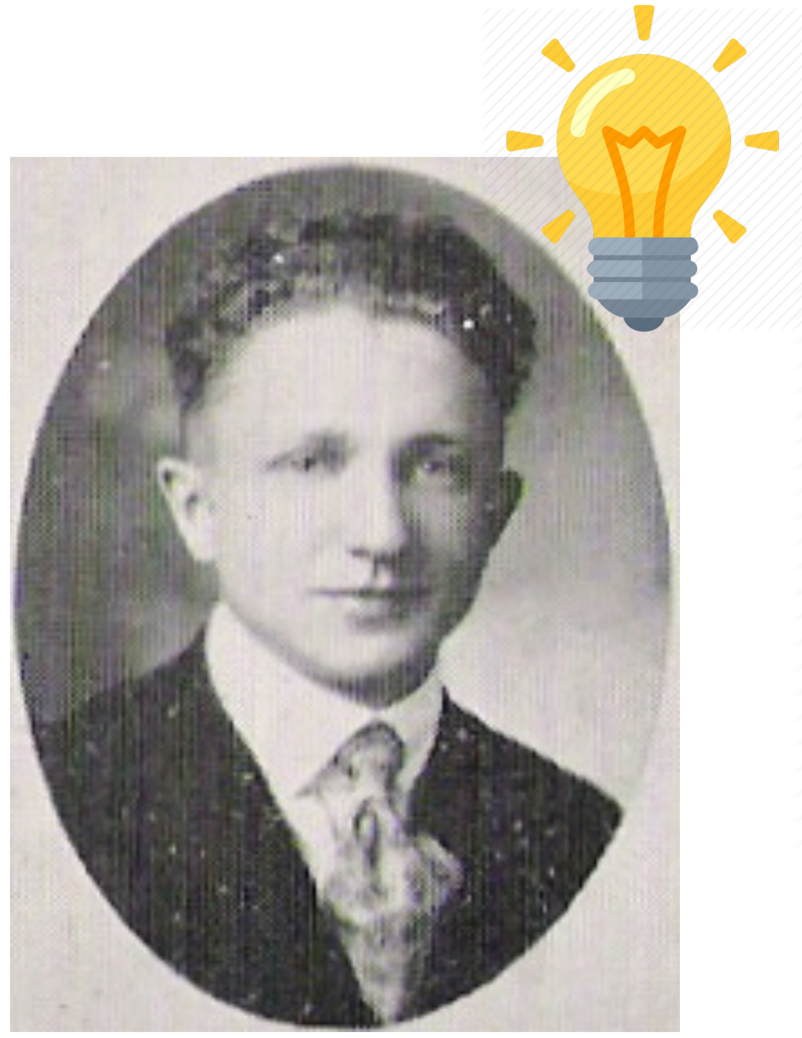
G.K. Zipf (1946)
US linguist and philologist

THE $\frac{P_1 P_2}{D}$ HYPOTHESIS: ON THE INTERCITY MOVEMENT OF PERSONS

GEORGE KINGSLEY ZIPF
Harvard University



Gravity model



$$T_{ij} \propto \frac{P_i P_j}{D_{ij}}$$

G.K. Zipf (1946)
US linguist and philologist

THE $\frac{P_1 \, P_2}{D}$ HYPOTHESIS: ON THE INTERCITY MOVEMENT OF PERSONS

GEORGE KINGSLEY ZIPF
Harvard University

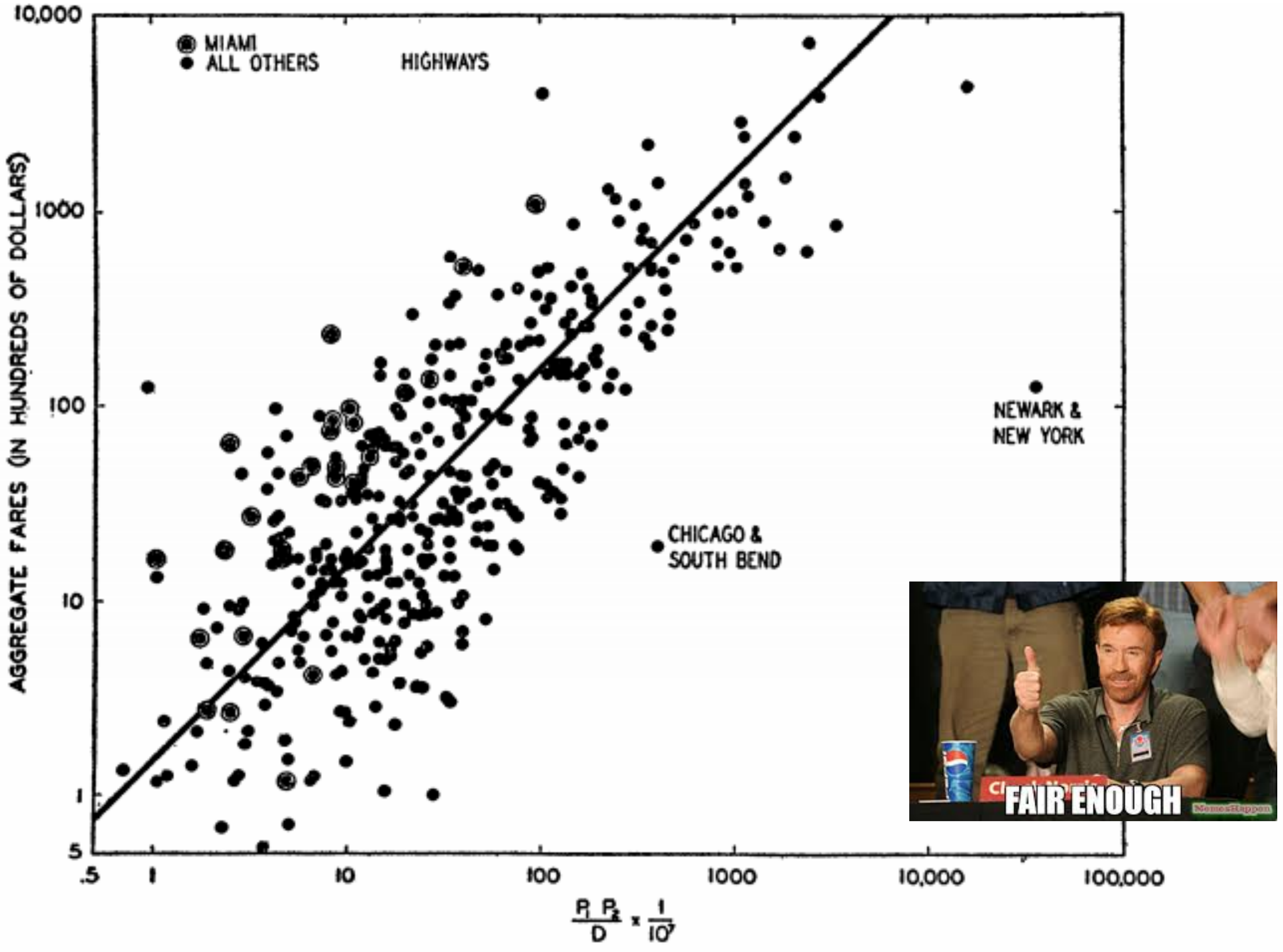
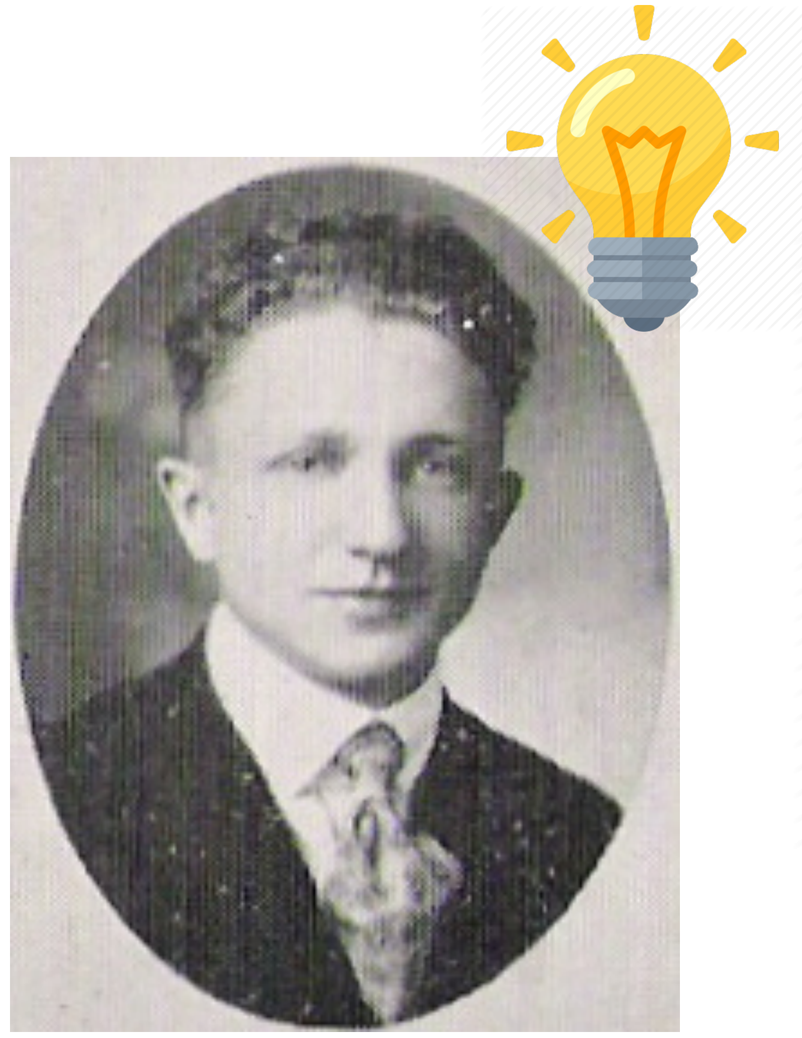


FIGURE 4. The aggregate fares (in hundreds of dollars) paid by the highway passengers reported in Figure 3. The ideal line has a slope of 1.

Gravity model

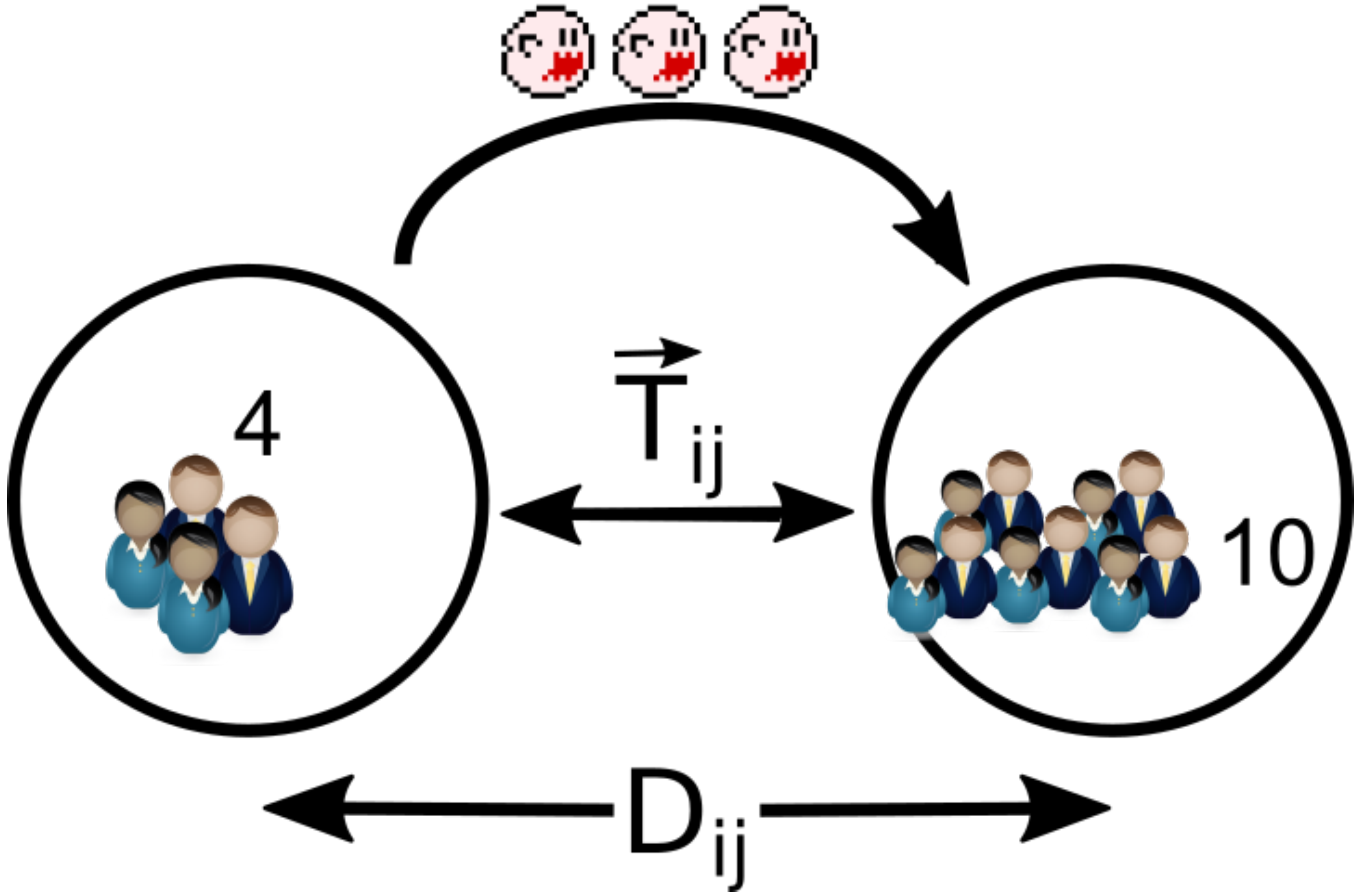


$$T_{ij} \propto \frac{P_i P_j}{D_{ij}}$$

G.K. Zipf (1946)
US linguist and philologist

THE $\frac{P_1 P_2}{D}$ HYPOTHESIS: ON THE INTERCITY MOVEMENT OF PERSONS

GEORGE KINGSLEY ZIPF
Harvard University



Gravity model

Unconstrained gravity model $T_{ij} = K m_i m_j \underline{f(r_{ij})}$ → deterrence function

- Power law: $r_{ij}^{-\alpha}$
- Exponential $e^{-r_{ij}/d'}$

Singly constrained gravity model
(Production constrained)

$$T_{ij} = K_i \underline{O_i} m_j f(r_{ij})$$

Doubly constrained gravity model

$$T_{ij} = K_i \underline{O_i} L_j \underline{A_j} f(r_{ij})$$

Needs data on outflows and inflows
Unfeasible without data

$$O_i = \sum_j T_{ij}$$

$$A_j = \sum_i T_{ij}$$

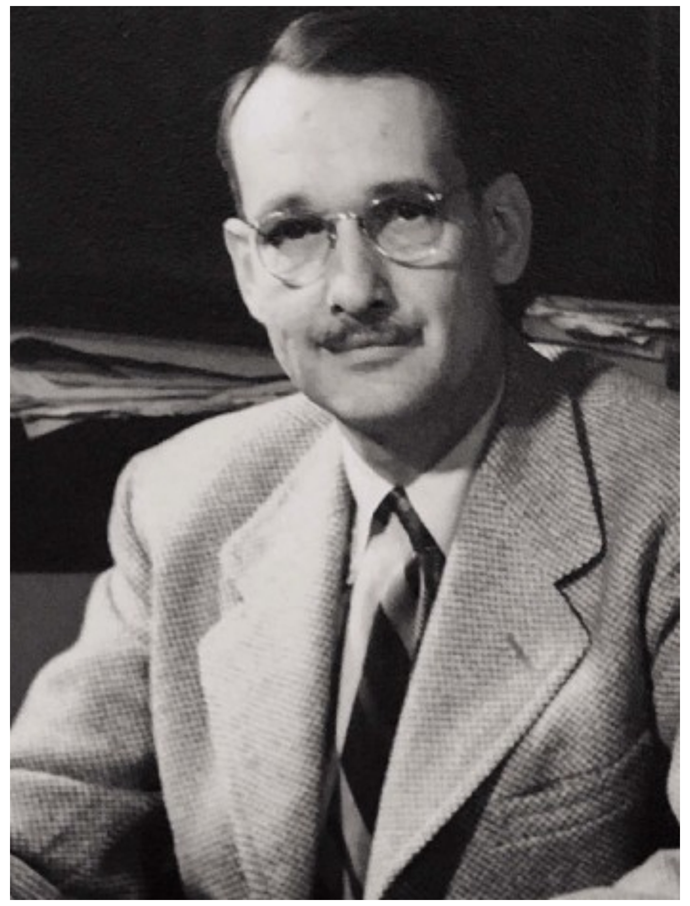
Barbosa, Hugo, et al. "Human mobility: Models and applications." Physics Reports 734 (2018): 1-74.

Radiation model

Meanwhile in Sociology...

Tired of looking at the stars, Professor Jenkins takes up sociology.

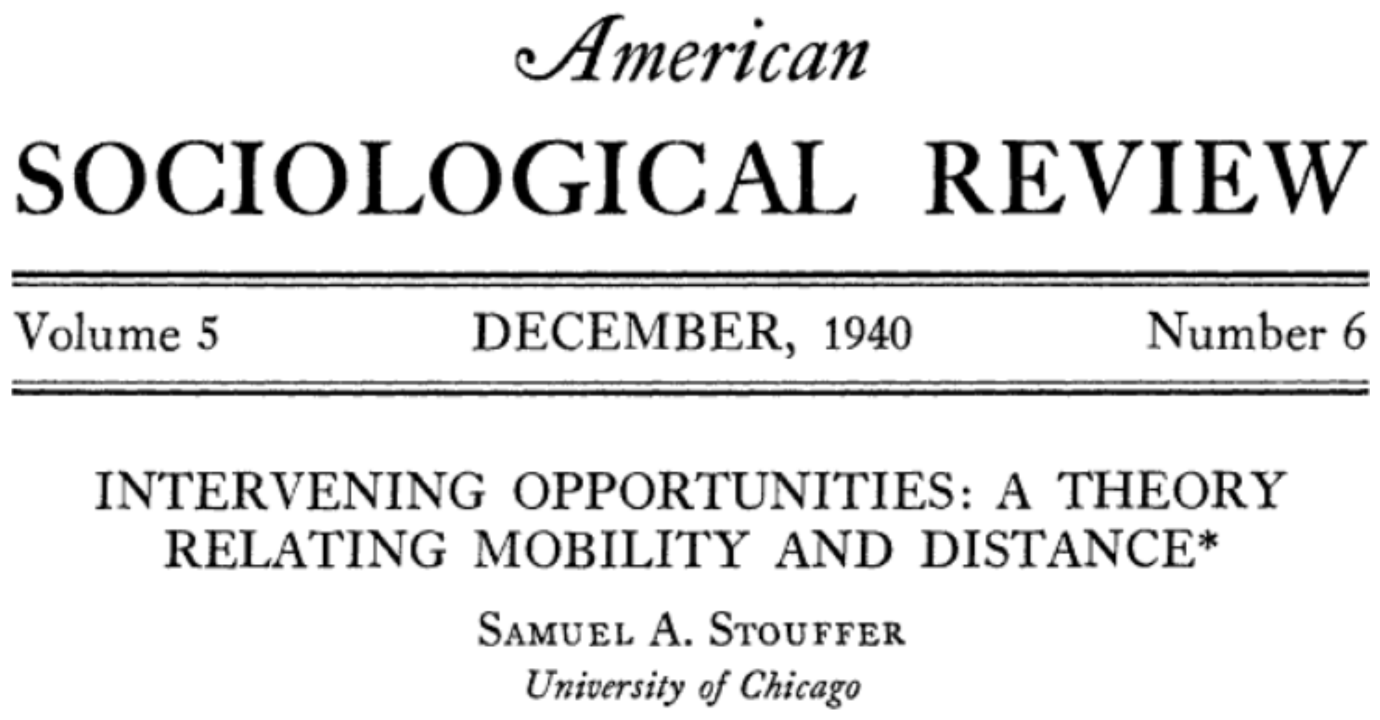
Intervening opportunities model



$$\frac{dy}{ds} = \frac{a}{x}\frac{dx}{ds}$$

S.A. Stouffer (1940)
US Sociologist

"The number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities."

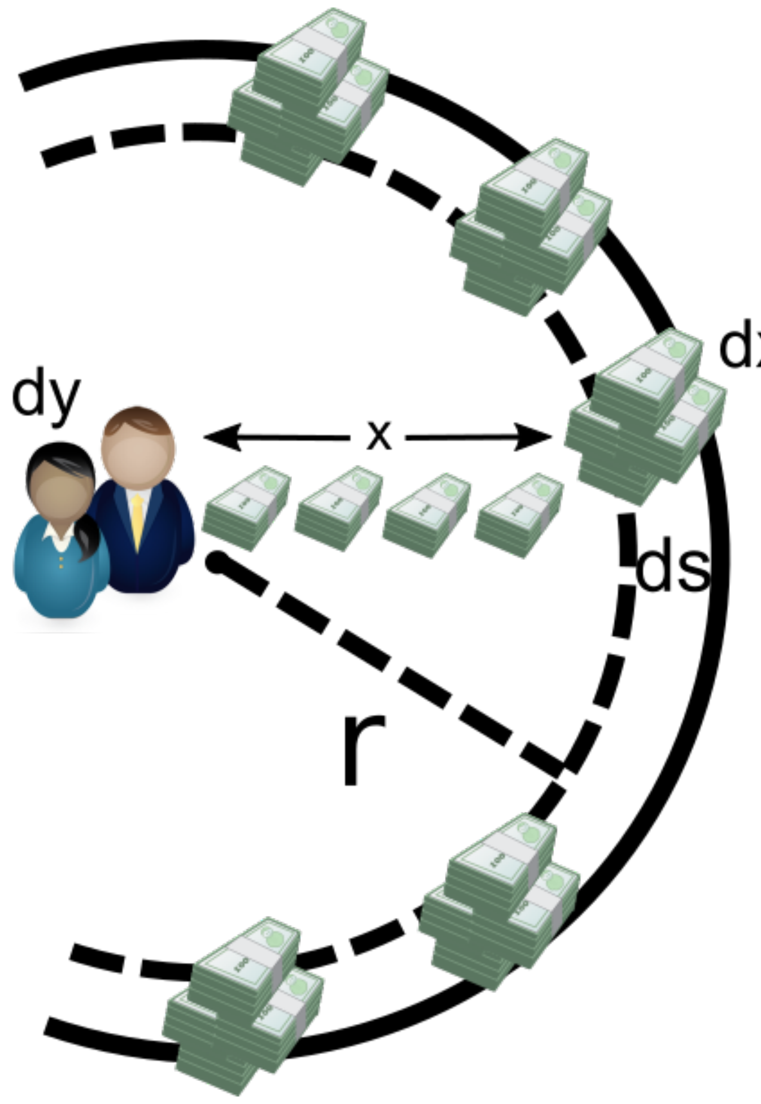


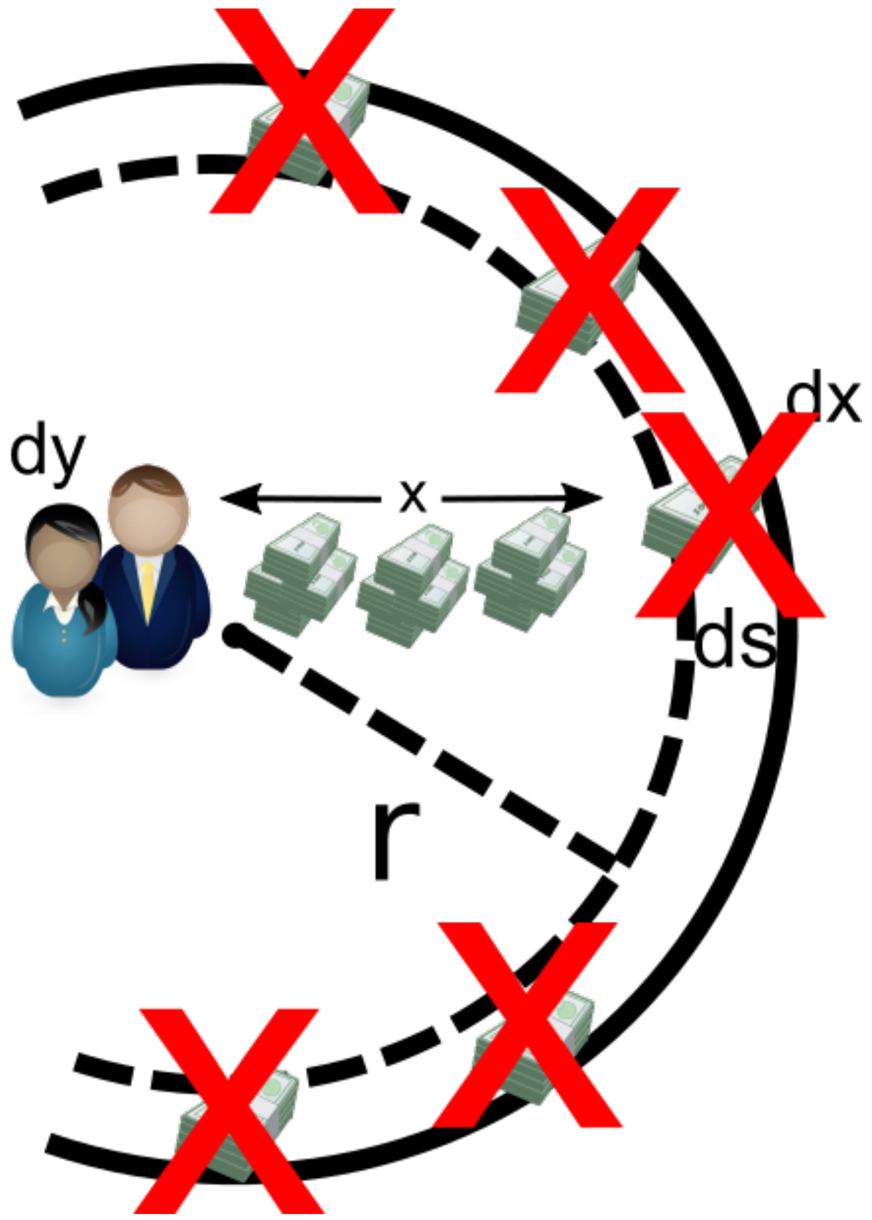
American
SOCIOLOGICAL REVIEW

Volume 5 DECEMBER, 1940 Number 6

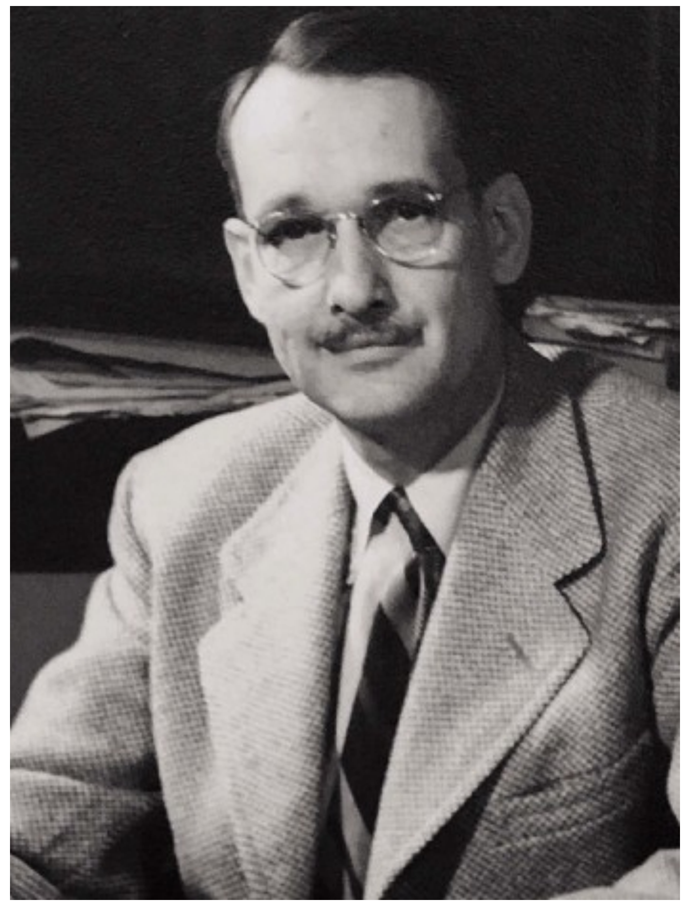
INTERVENING OPPORTUNITIES: A THEORY RELATING MOBILITY AND DISTANCE*

SAMUEL A. STOUFFER
University of Chicago





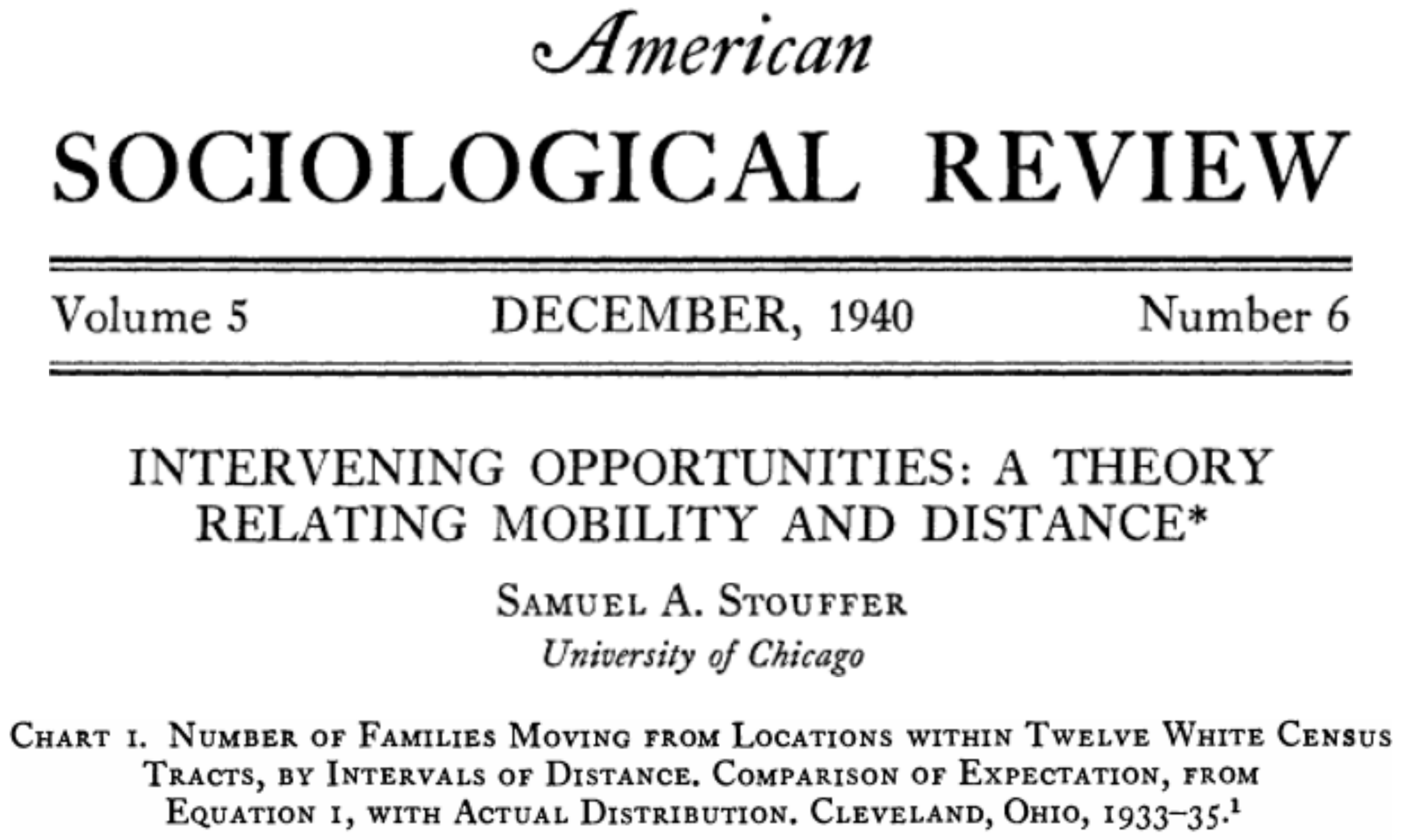
Intervening opportunities model



$$\frac{dy}{ds} = \frac{a}{x}\frac{dx}{ds}$$

S.A. Stouffer (1940)
US Sociologist

“The number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities.”



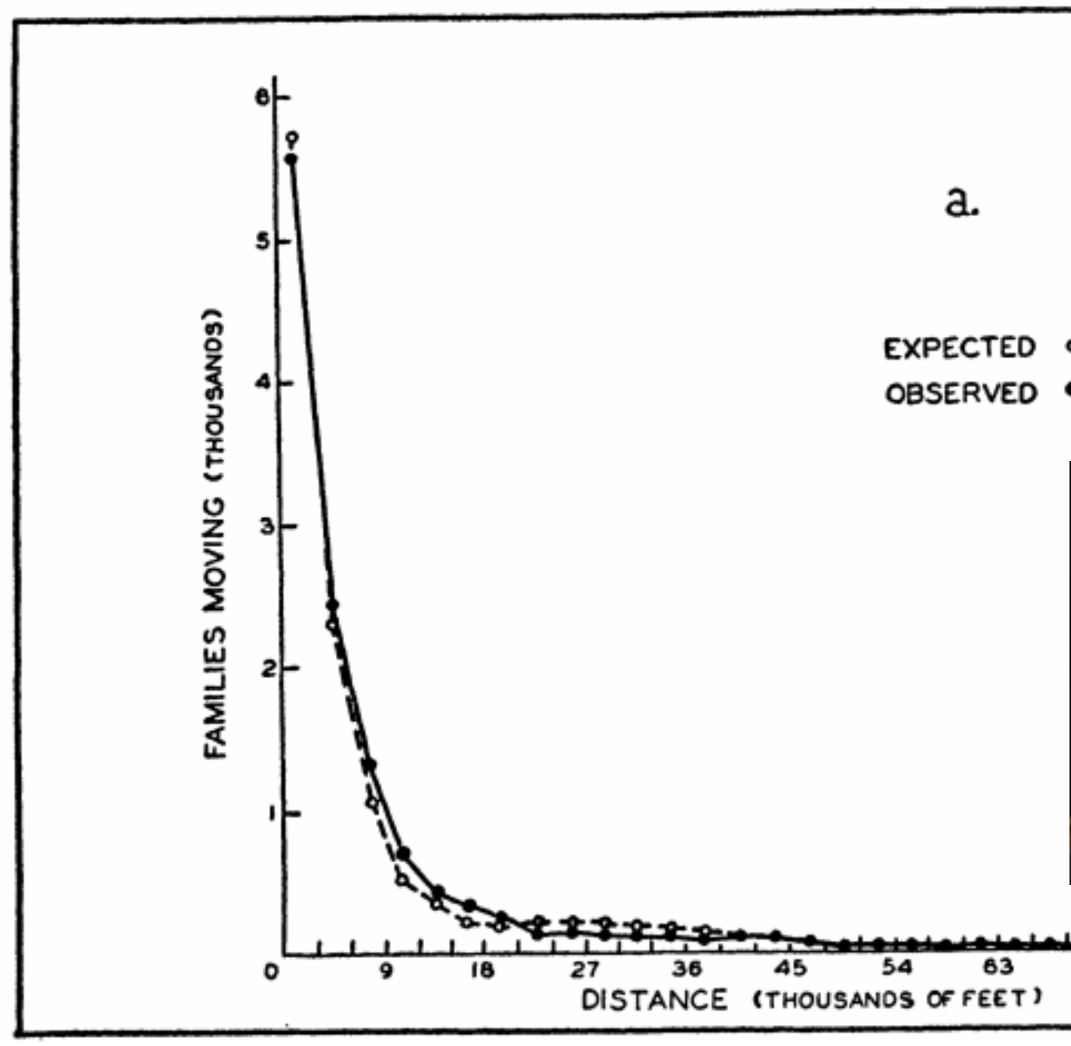
American
SOCIOLOGICAL REVIEW

Volume 5 DECEMBER, 1940 Number 6

INTERVENING OPPORTUNITIES: A THEORY RELATING MOBILITY AND DISTANCE*

SAMUEL A. STOUFFER
University of Chicago

CHART I. NUMBER OF FAMILIES MOVING FROM LOCATIONS WITHIN TWELVE WHITE CENSUS TRACTS, BY INTERVALS OF DISTANCE. COMPARISON OF EXPECTATION, FROM EQUATION I, WITH ACTUAL DISTRIBUTION. CLEVELAND, OHIO, 1933–35.[1]



Radiation model

LETTER

doi:10.1038/nature10856

A universal model for mobility and migration patterns

Filippo Simini[1,2,3], Marta C. González[4], Amos Maritan[2] & Albert-László Barabási[1,5,6]

(2012)

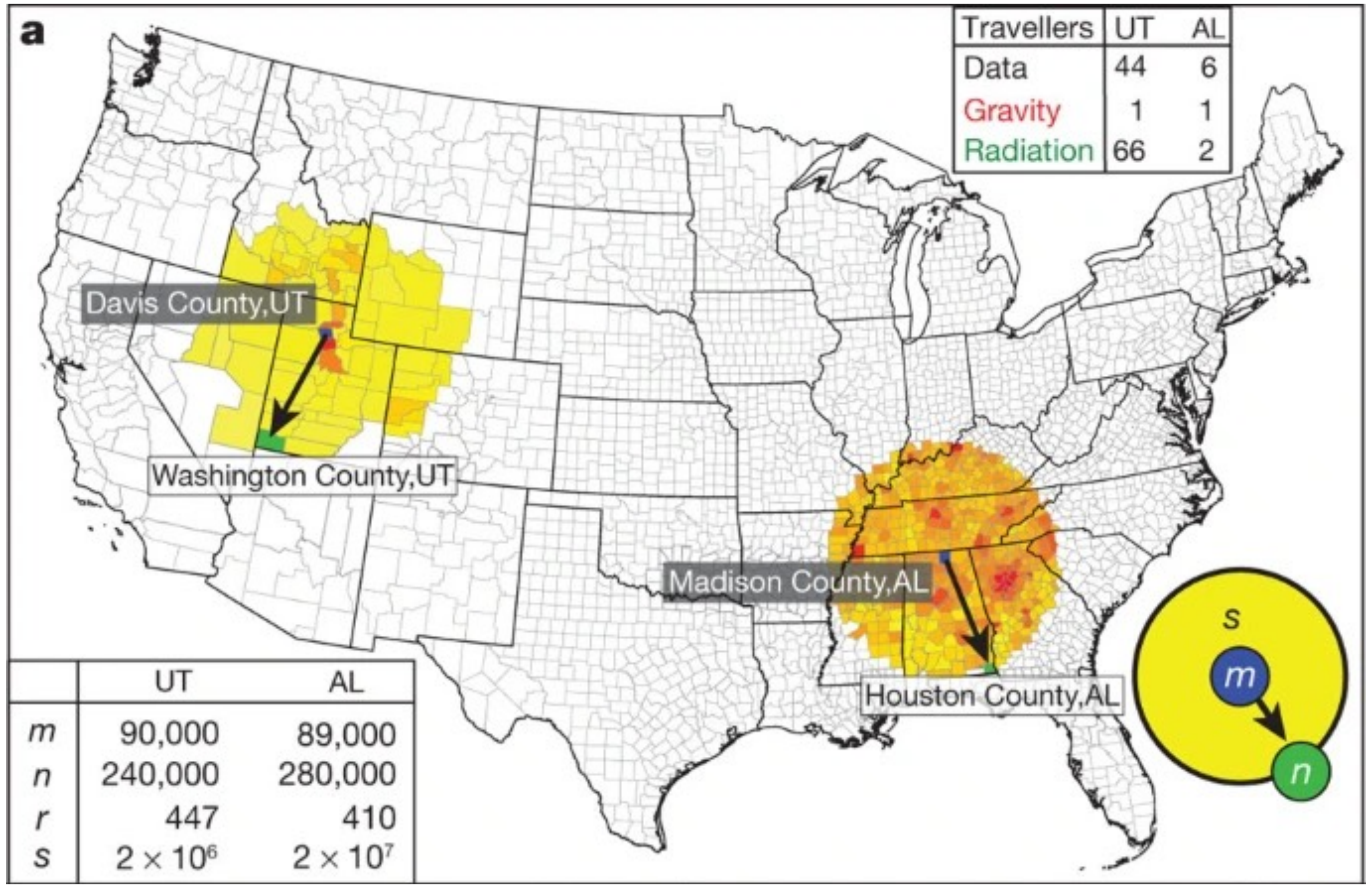
Inspired by the Intervening opportunities model

It mimics the radiation and absorption of particles
Particles emitted and absorbed proportional to local population

Opportunities = individuals

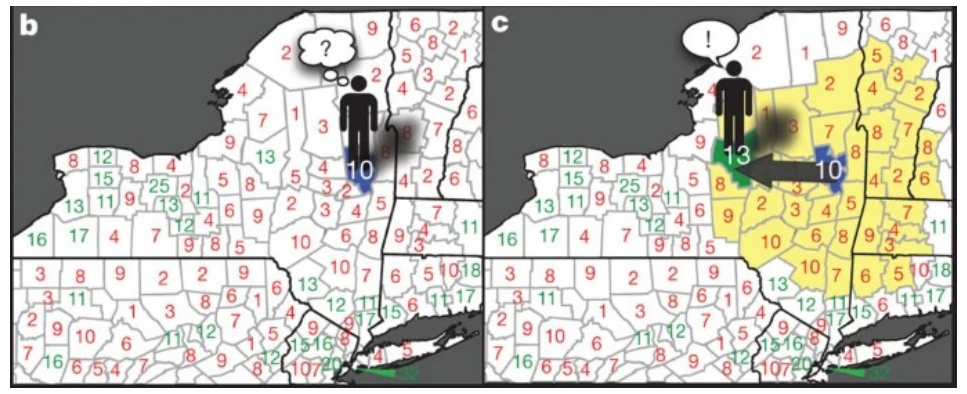
Parameter free

Requires knowledge on the outflows



$$T_{ij} = T_i \frac{P_i \, P_j}{(P_i + s_{ij})\,(Pi + P_j + s_{ij})}$$

$$T_i = \sum_j T_{ij}$$



Simini, Filippo, et al. "A universal model for mobility and migration patterns." *Nature* 484.7392 (2012): 96-100.
Barbosa, Hugo, et al. "Human mobility: Models and applications." Physics Reports 734 (2018): 1-74.

Agent based models (MATsim)

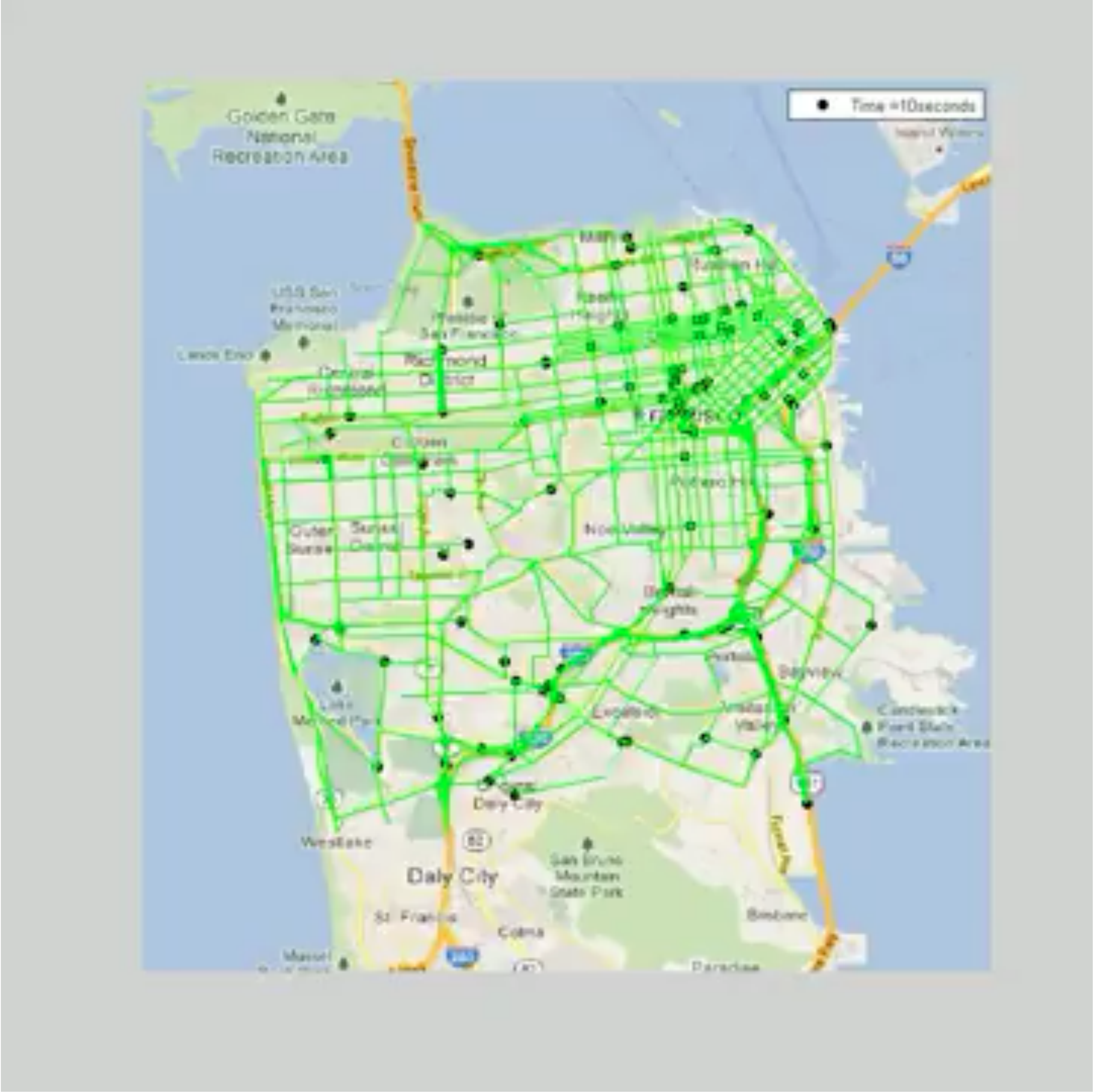
Open-source framework for implementing large-scale agent-based transport simulations.

Needs road and transportation network (OpenStreetMaps)

Useful to:

- Simulate traffic congestions
- Analyse transport demand
- Regulate transport offer
- Estimate car traffic emissions and exposure to pollutants
- propose traffic interventions, e.g. tolls

Make your own simulation!
https://www.matsim.org/about-matsim



Marta Gonzalez, YouTube

Mobility models performance metrics

… among others

Macro-scale metrics:

- CPC: common part of commuters
- Normalised root mean squared error
- Pearson correlation

Micro (individual) scale metrics:

- radius of gyration
- jump length

CPC: common part of commuters

$$CPC\left(T,\tilde{T}\right) = \frac{\sum_{i,j=1}^{n} min\left(T_{ij},\tilde{T}_{ij}\right)}{N} = 1 - \frac{1}{2}\frac{\sum_{i,j=1}^{n} |T_{ij} - \tilde{T}_{ij}|}{N}$$

0: no agreement, 1: full agreement

Normalised root mean squared error

$$NRMSE\left(T,\tilde{T}\right) = \frac{\sum_{i,j=1}^{n}\left(T_{ij}-\tilde{T}_{ij}\right)^2}{\sum_{i,j=1}^{n} T_{ij}^{\ 2}}$$

$\tilde{T}_{ij}$: simulated

0: full agreement

Pearson correlation

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

-1: anti correlated, 1: fully correlated

Lenormand, Maxime, Aleix Bassolas, and José J. Ramasco. "Systematic comparison of trip distribution laws and models." *Journal of Transport Geography* 51 (2016): 158-169.

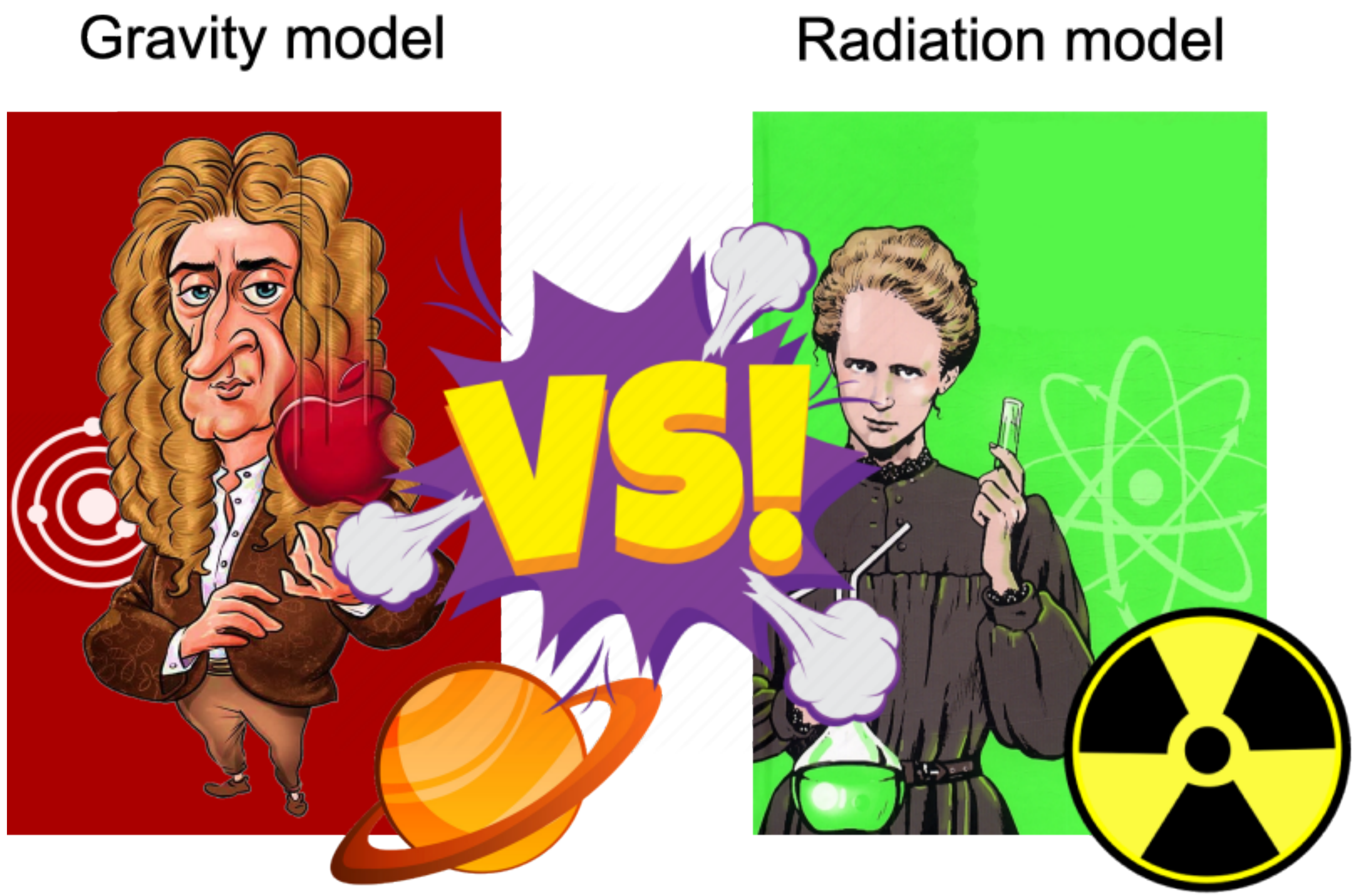
Hands on mobility models

Let's try gravity and radiation models on a real dataset and see which one works best

Data: New York City commuters
Scale: counties

Follow the instructions at
https://scikit-mobility.github.io/scikit-mobility/index.html

Go to mattiamazzoli.github.com/
Go to Teaching
Download the notebook "mobility models"
Follow the instructions



Mobility in urban spaces

Human Mobility: data, theory and models

Human mobility in urban spaces

- **The first study on mobility with telephone data**
  - Marchetti’s antropological invariants
- **The 15 minute city [GPS traces]**
  - From Marchetti’s law to urban isochrones
- **Urban land use [CDR data]**
  - Characterising how citizens live the city

The first study on telephone data and mobility

How were mobile phone data used before the big data revolution?

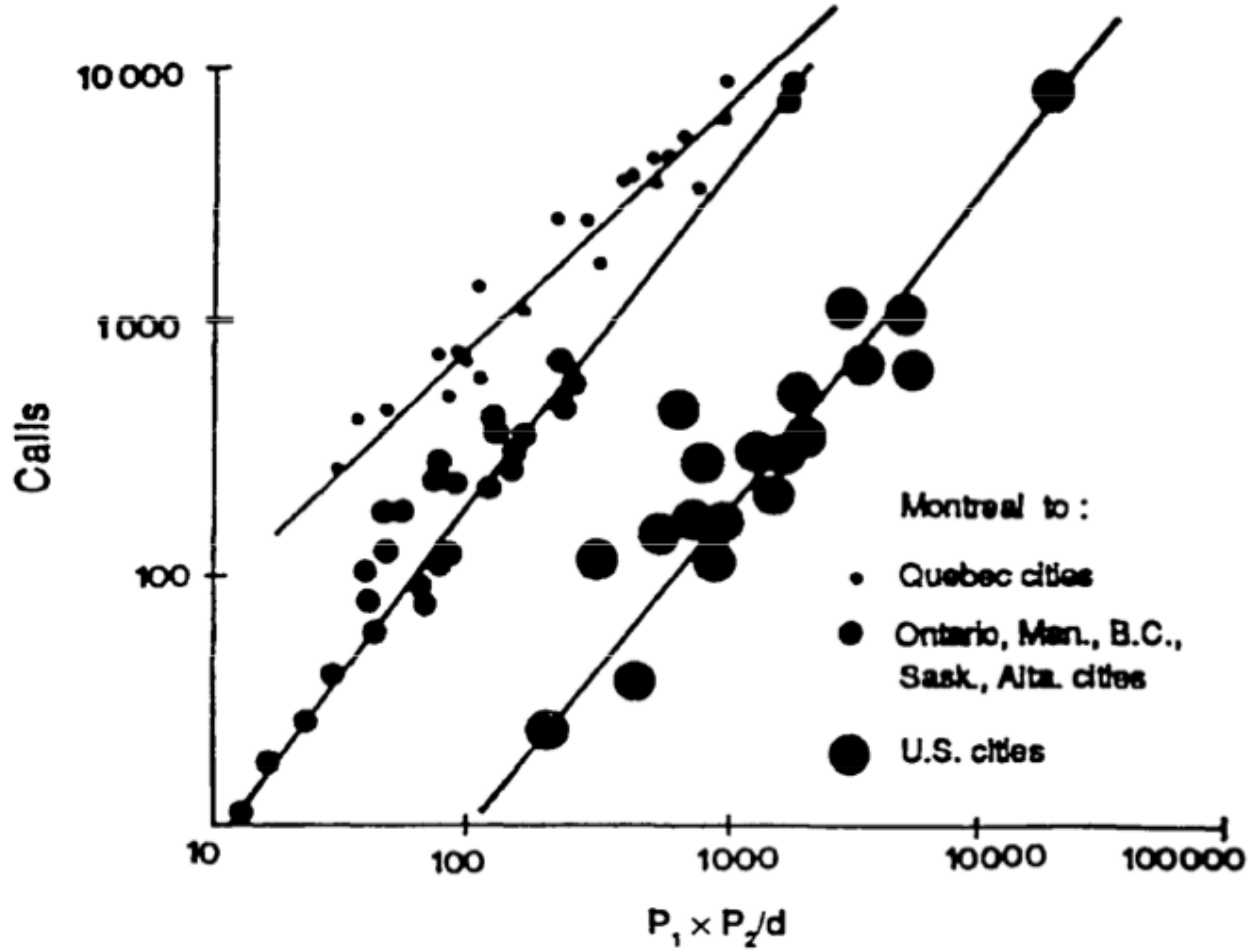
Cesare Marchetti
(Italian Physicist)
Anthropological invariants in travel behaviour (1994)

“Due to parallelism between message exchange by telephone and traveling, we may use the first as a proxy for the second […]”

TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE 47, 75-88 (1994)

Anthropological Invariants in Travel Behavior

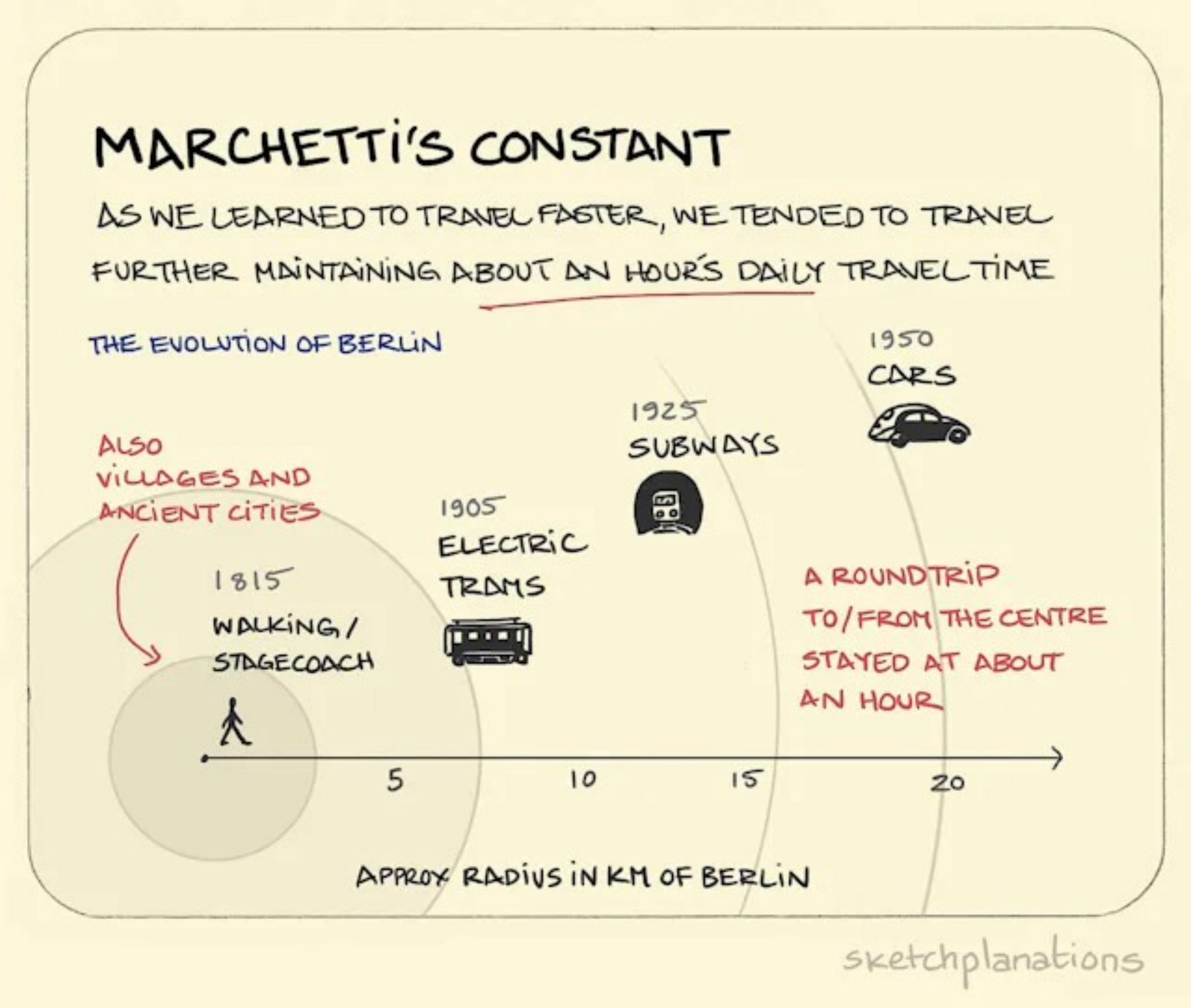
C. MARCHETTI



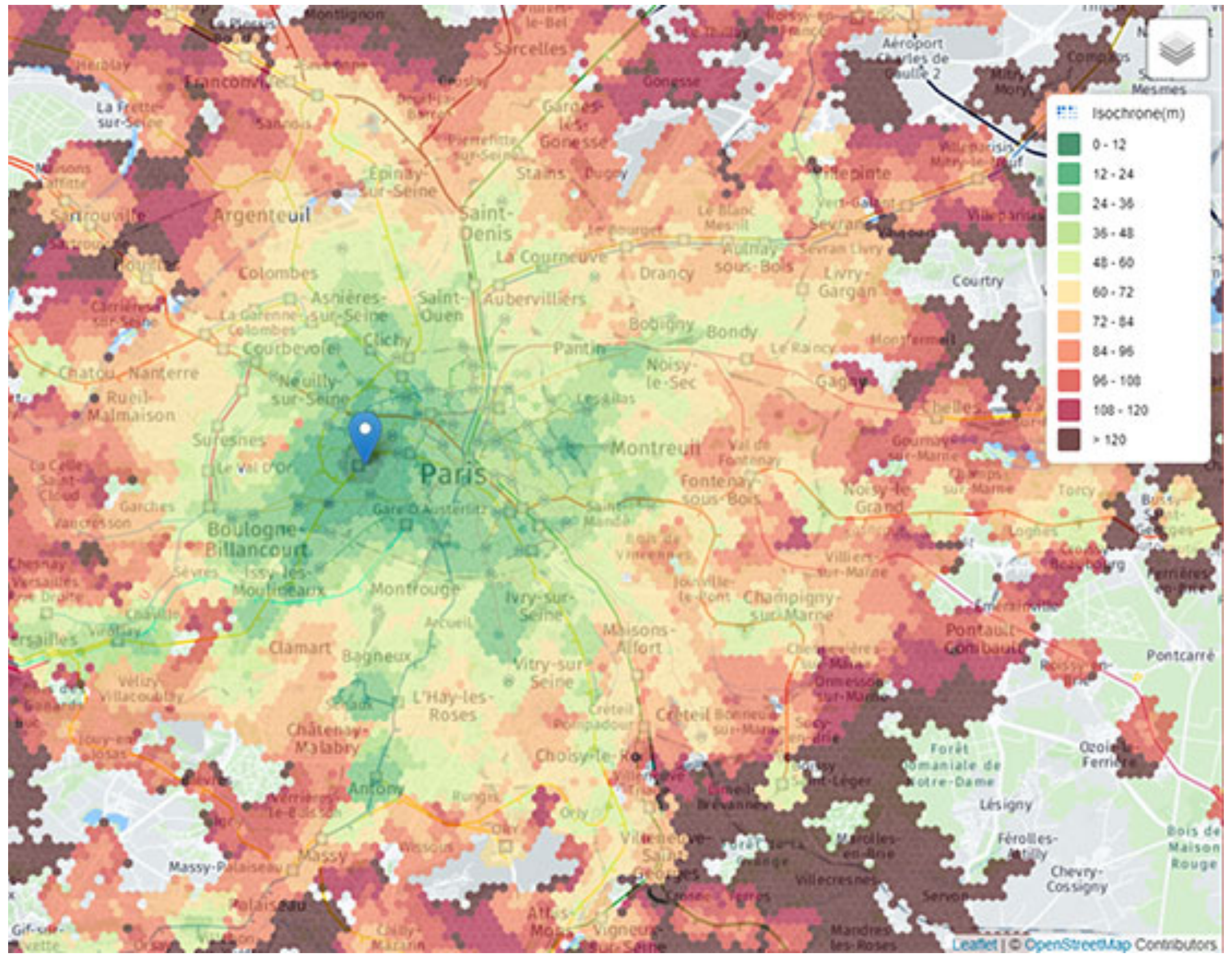
Marchetti, Cesare. "Anthropological invariants in travel behavior." Technological forecasting and social change 47.1 (1994): 75-88.

The 15 minutes city

The original theory

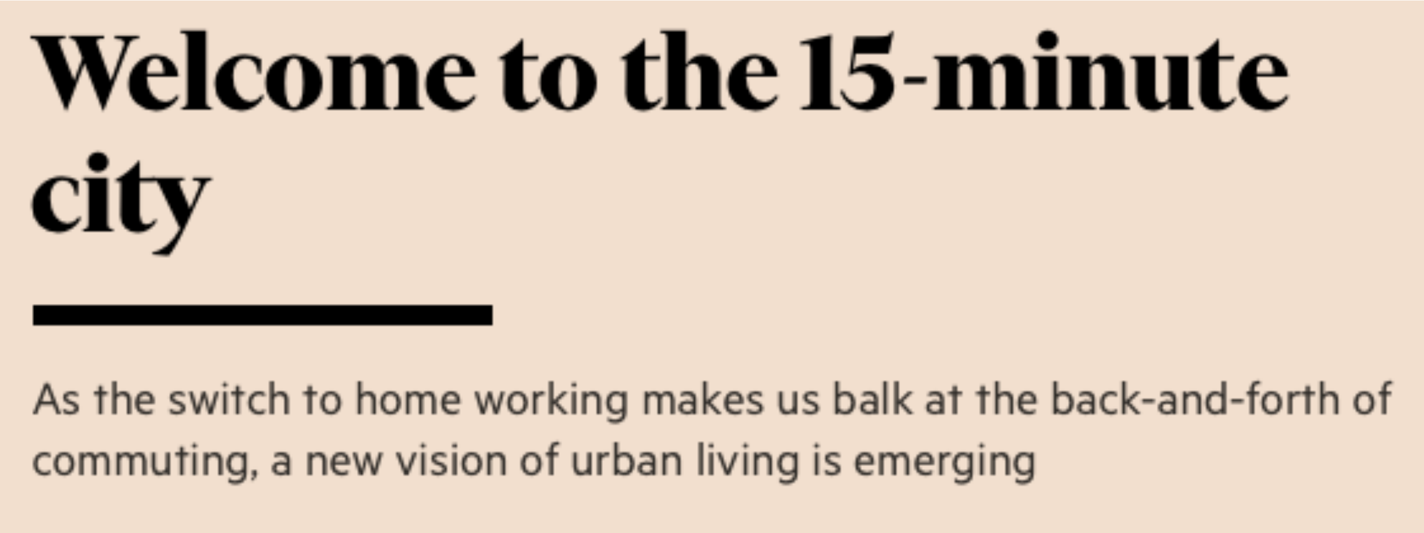


The big data application



http://whatif.sonycsl.it/citychrone/

The policy



Welcome to the 15-minute city

As the switch to home working makes us balk at the back-and-forth of commuting, a new vision of urban living is emerging

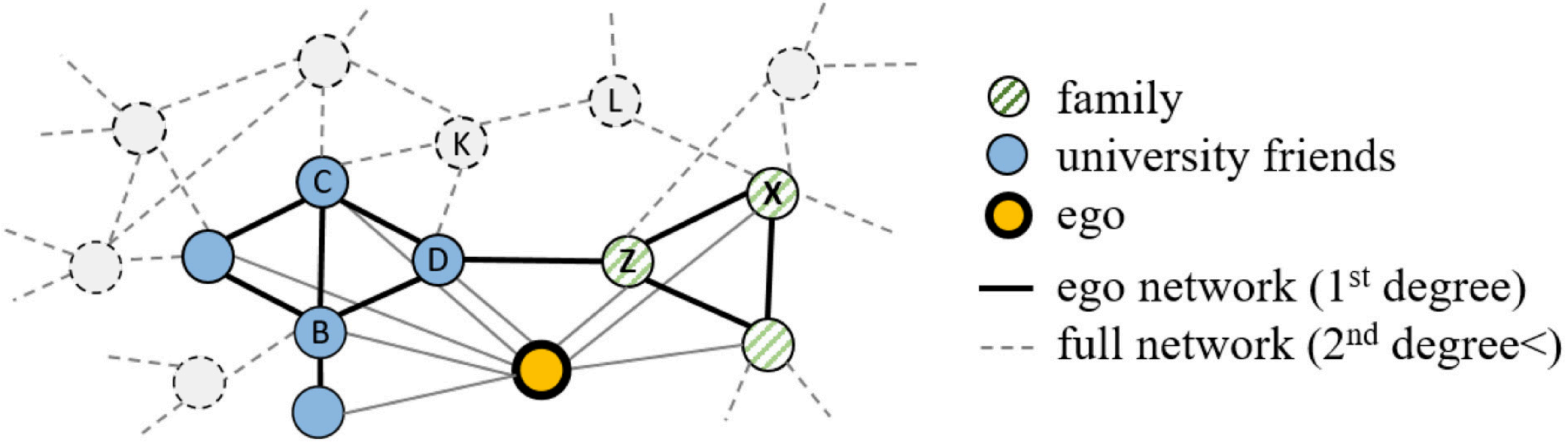
https://www.ft.com/content/c1a53744-90d5-4560-9e3f-17ce06aba69a

Marchetti, Cesare. "Anthropological invariants in travel behavior." Technological forecasting and social change 47.1 (1994): 75-88.
Biazzo, Indaco, Bernardo Monechi, and Vittorio Loreto. "General scores for accessibility and inequality measures in urban areas." *Royal Society open science* 6.8 (2019): 190979.
https://csl.sony.it/project/the-15-minutes-city/

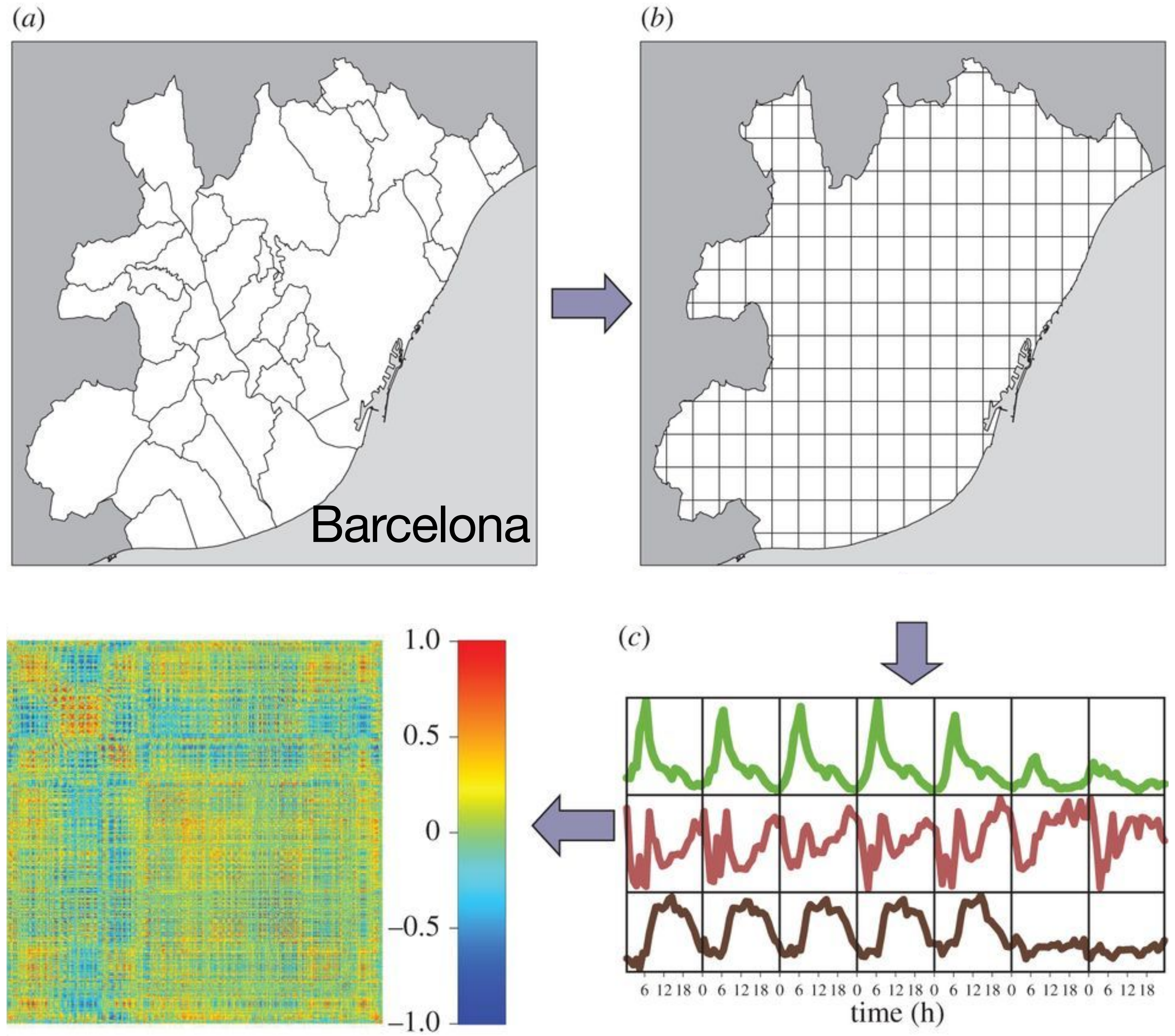
The ego-network

Sub-set of original network
Filtered on single node (ego) connectivity

It can be expanded to further degrees, e.g. friends of friends, etc



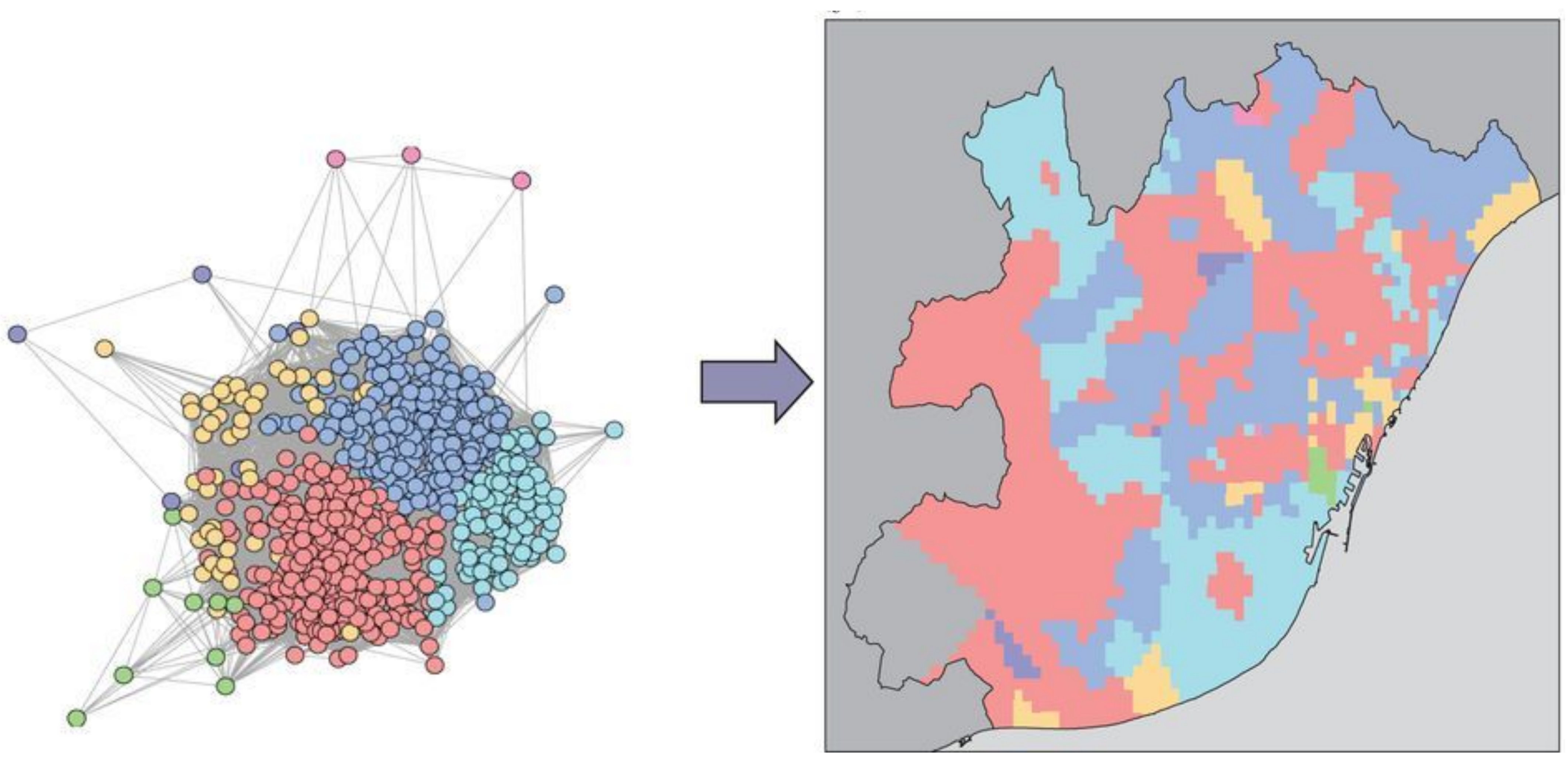
Urban land use



Pearson correlation of signals between each pair of cells

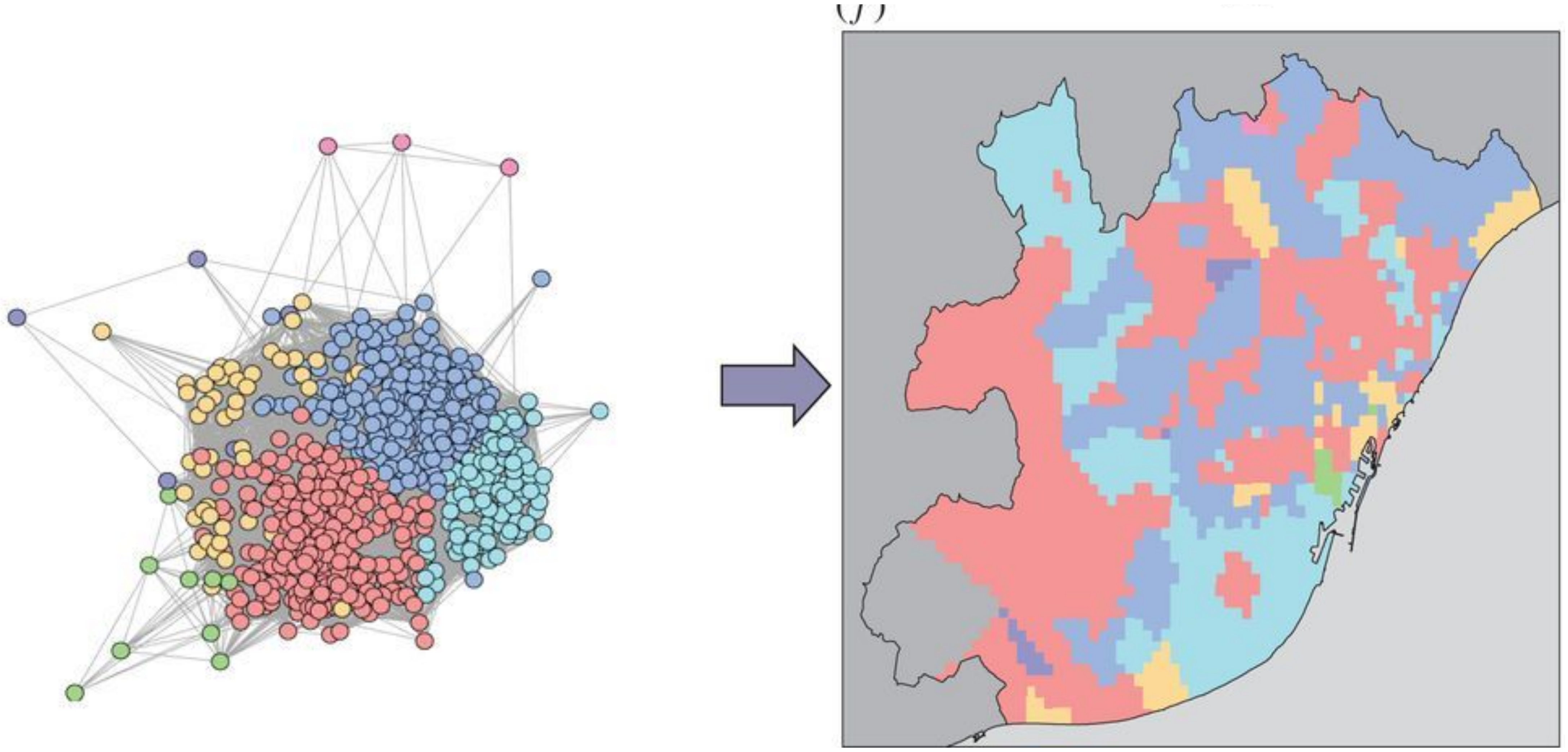
Activity profile in each cell

- Obtain a completely connected network of cells
- Nodes = areas
- Links = activity correlation
- Impose a threshold on the correlation, e.g. 0
- Run community detection algorithms on the network of cells



Lenormand Maxime, Picornell Miguel, Cantú-Ros Oliva G., Louail Thomas, Herranz Ricardo, Barthelemy Marc, Frías-Martínez Enrique, San Miguel Maxi and Ramasco José J. 2015Comparing and modelling land use organization in cities, R. Soc. Open Sci.2150449

Urban land use



Community represents a land use type

Compute typical (average) activity of communities

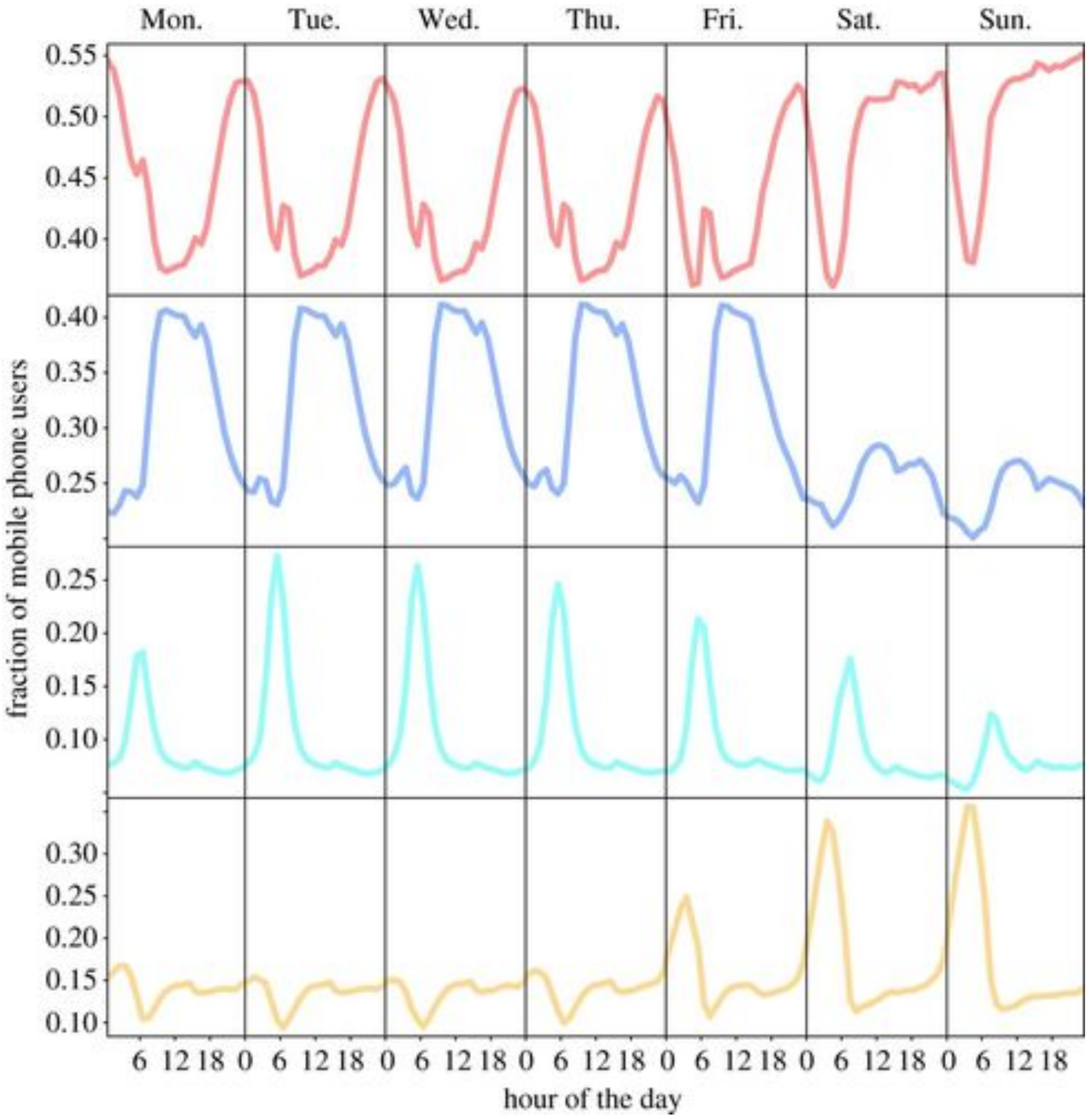
Temporal patterns associated with the four clusters of activity.

Red: residential cluster

Blue: business

Cyan: logistics/industry

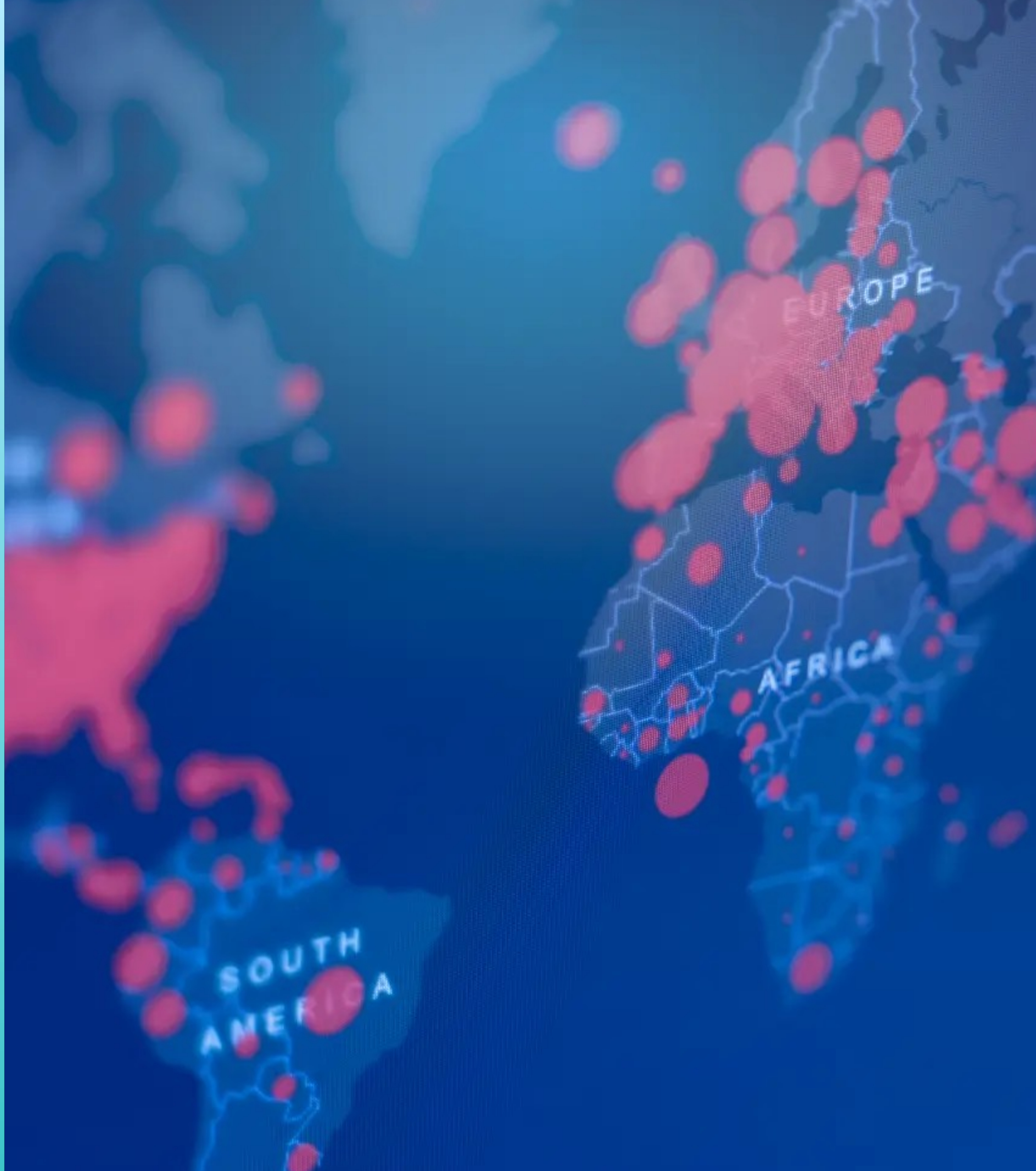
Orange: nightlife



Lenormand Maxime, Picornell Miguel, Cantú-Ros Oliva G., Louail Thomas, Herranz Ricardo, Barthelemy Marc, Frías-Martínez Enrique, San Miguel Maxi and Ramasco José J. 2015Comparing and modelling land use organization in cities, R. Soc. Open Sci.2150449

Mobility in epidemics

Human Mobility: data, theory and models



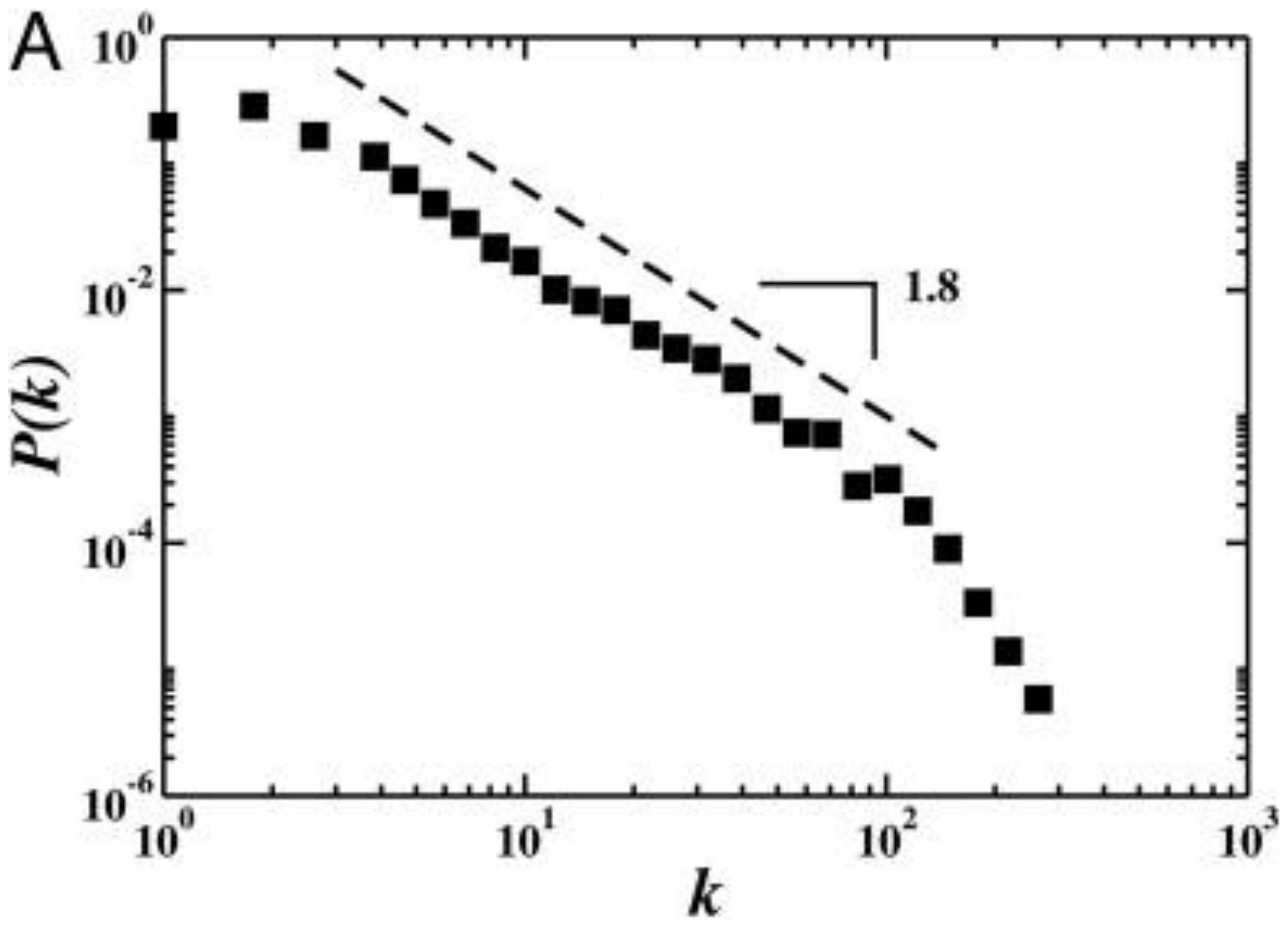
Mattia Mazzoli - UniTo

Human mobility impact on epidemics

Disease spreading on mobility network

- **Country scale: spatial networks [air traffic data]**
  - Country level arrival times: the hidden geometry of epidemic spread
  - Pathways of spatial invasion
- **Sub-national scale: invasion trees [CDR data]**
  - Disease invasion reconstruction: phylogeographic analyses
- **Individual scale: co-presence temporal networks [GPS traces]**
  - Hotspots of infection in airports
- **Location scale: bipartite networks [surveys]**
  - Spatial spread of disease across venues

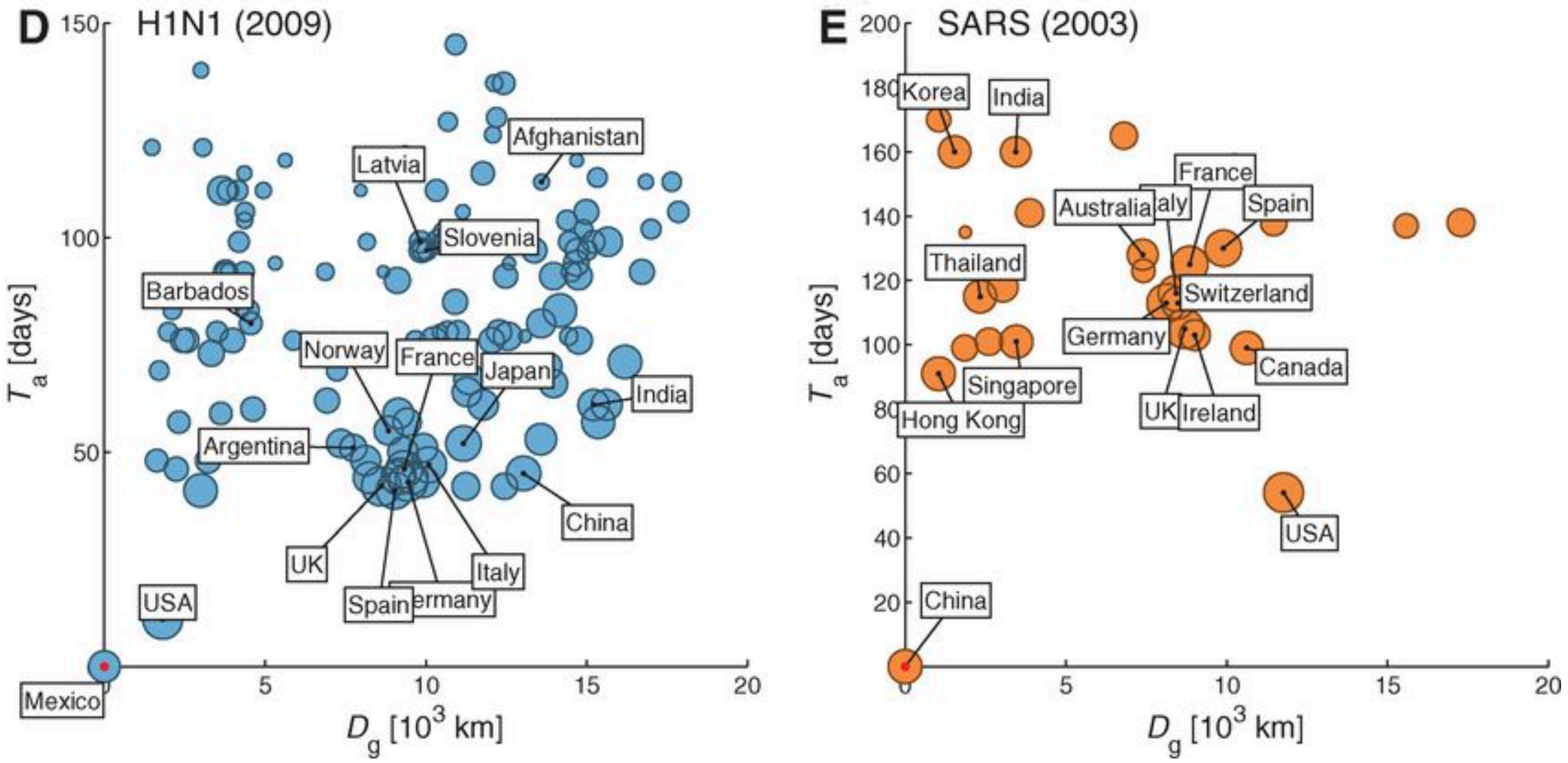
The air transportation network



The air transportation network is vulnerable to epidemics

The hidden geometry of epidemic spread, predicting arrival times

Predicting disease arrival times at country scale
From distance to the effective distance



Geographic distance

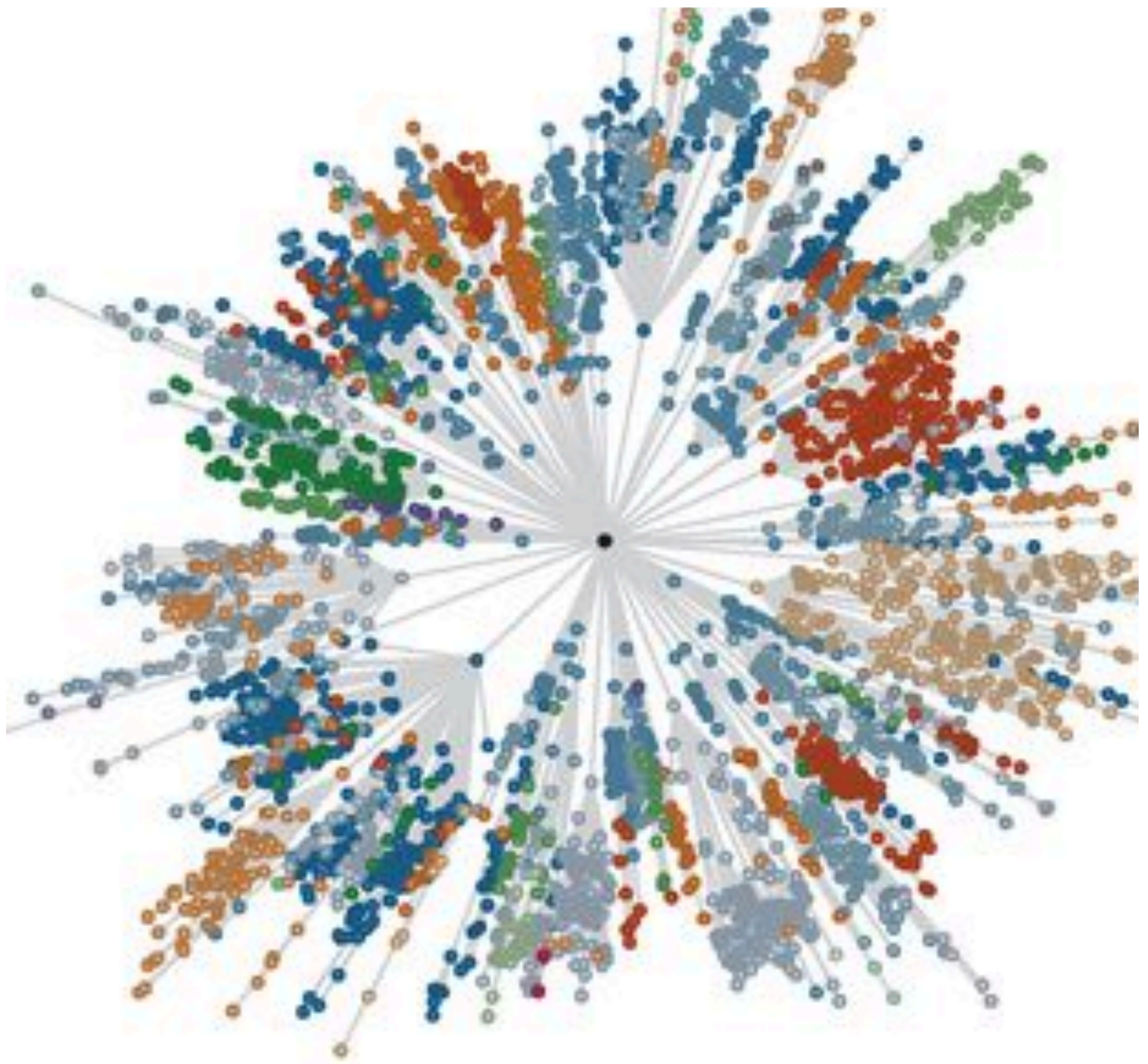
Dirk Brockmann, Dirk Helbing ,The Hidden Geometry of Complex, Network-Driven Contagion Phenomena.Science342,1337-1342(2013)

The hidden geometry of epidemic spread, predicting arrival times

$$d_{mn} = (1 - log P_{mn})$$

Flow fraction

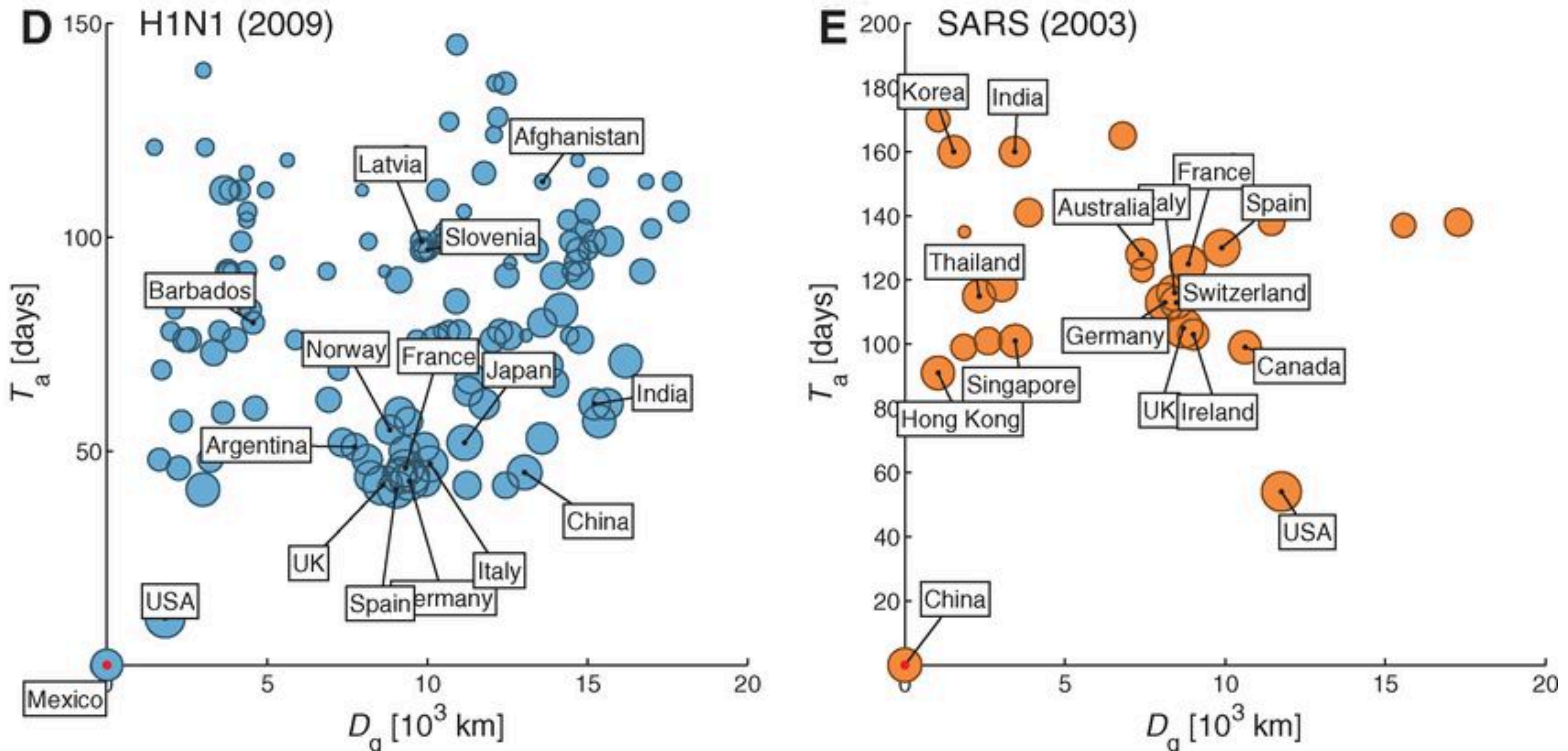
$$D_{mn} = min_{\Gamma} \lambda(\Gamma)$$



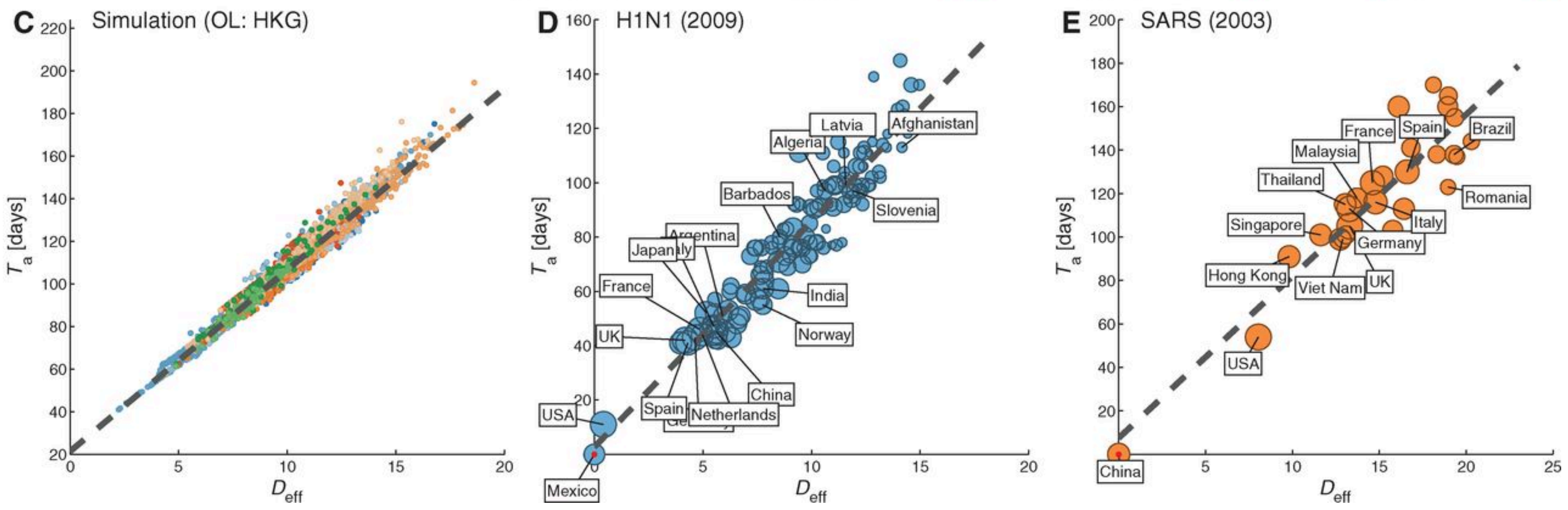
Dirk Brockmann, Dirk Helbing ,The Hidden Geometry of Complex, Network-Driven Contagion Phenomena.Science342,1337-1342(2013)

The hidden geometry of epidemic spread, predicting arrival times

Predicting disease arrival times at country scale
From distance to the effective distance



Geographic distance



Effective distance

$$d_{mn} = (1 - log P_{mn})$$

Flow fraction

$$D_{mn} = min_{\Gamma} \lambda(\Gamma)$$

Dirk Brockmann, Dirk Helbing ,The Hidden Geometry of Complex, Network-Driven Contagion Phenomena.Science342,1337-1342(2013).DOI:10.1126/science.1245200

The hidden geometry of epidemic spread, predicting arrival times

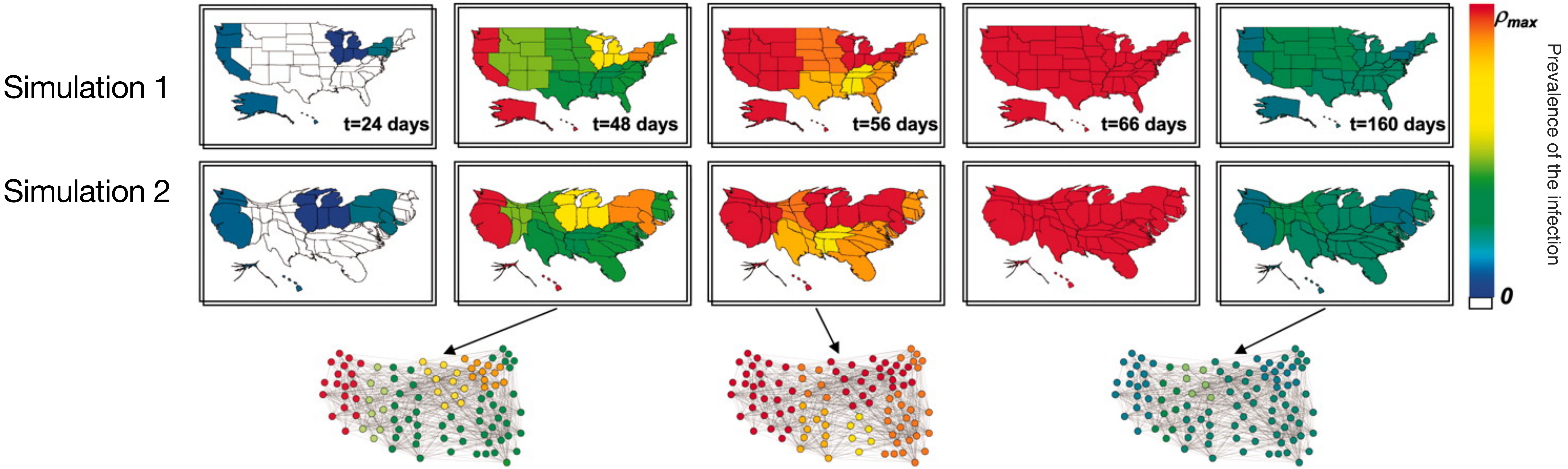
Dirk Brockmann, YouTube

HANDS ON SESSION
Go to https://github.com/mattiamazzoli/Teaching

Wave-like diffusion is still there, but now it is projected in another space!

Pathways of spatial invasion

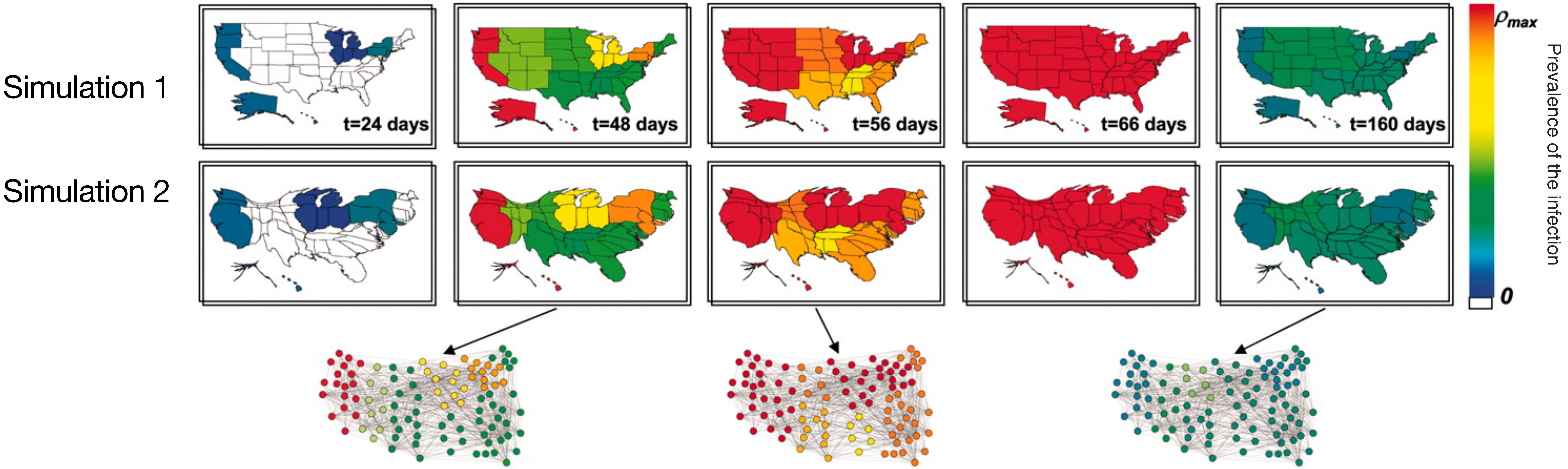
Simulating epidemic starting in Hong Kong on US States
Similarity of outcome affected by:
- stochasticity of the model
- heterogeneity of the air transportation network



Colizza, Vittoria, et al. "The role of the airline transportation network in the prediction and predictability of global epidemics." *PNAS* 103.7 (2006)

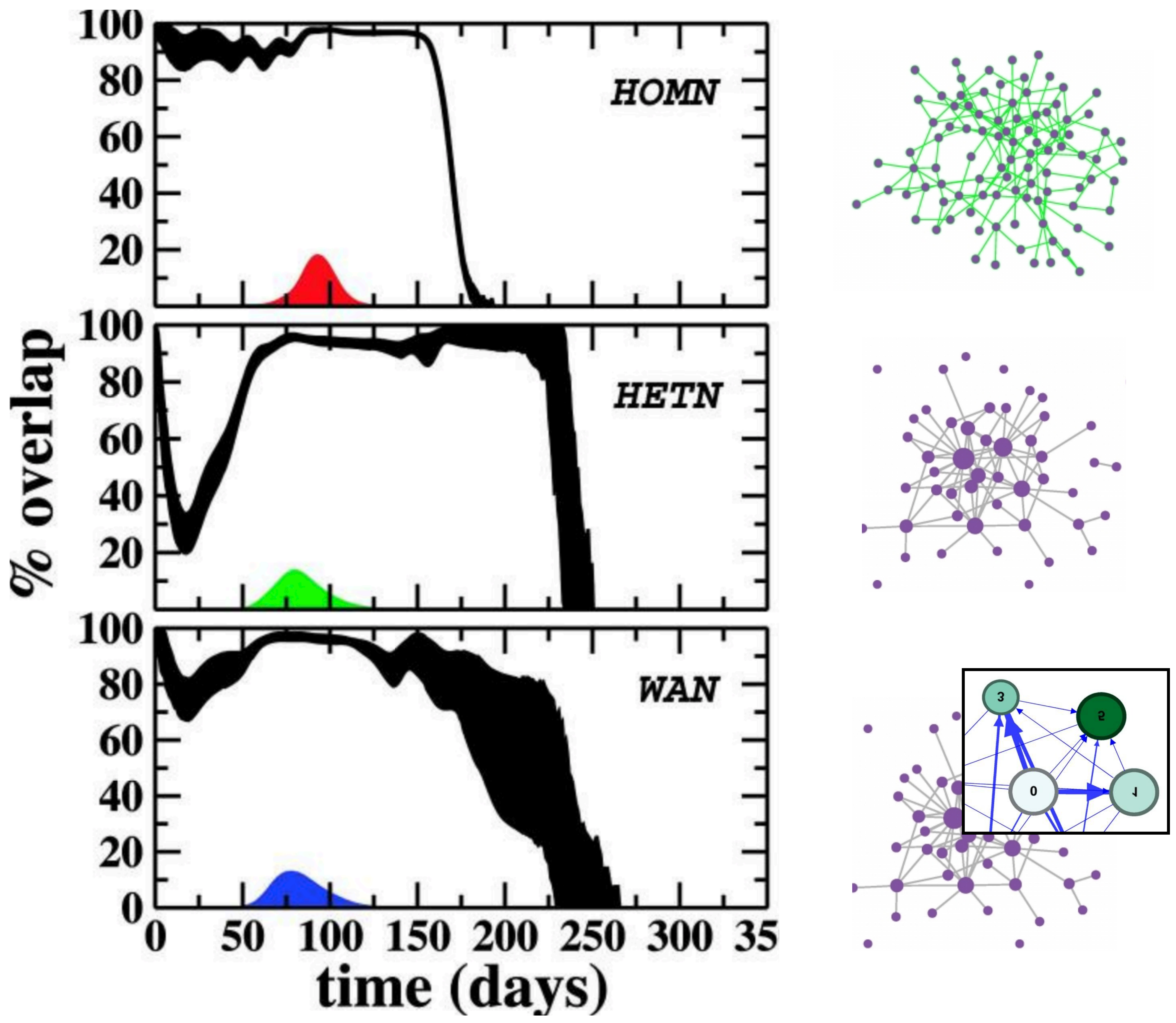
Pathways of spatial invasion

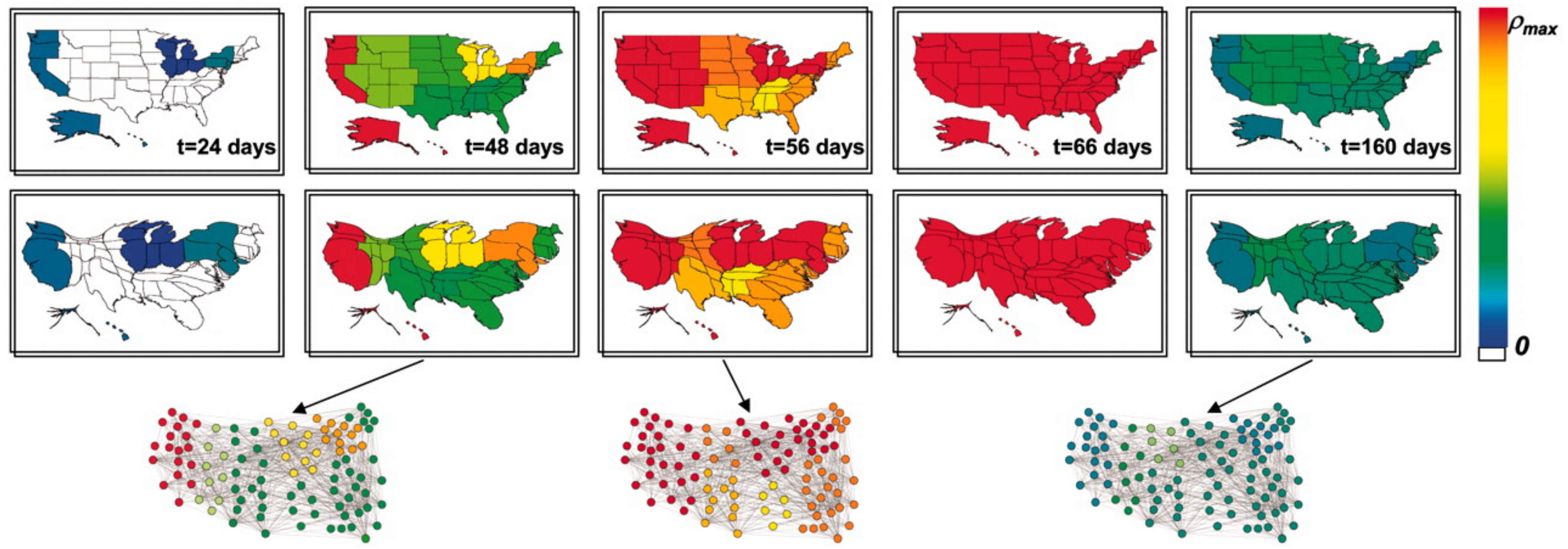
$\Theta(t)$ Overlap function: measures similarity between prevalence at time t in all cities i



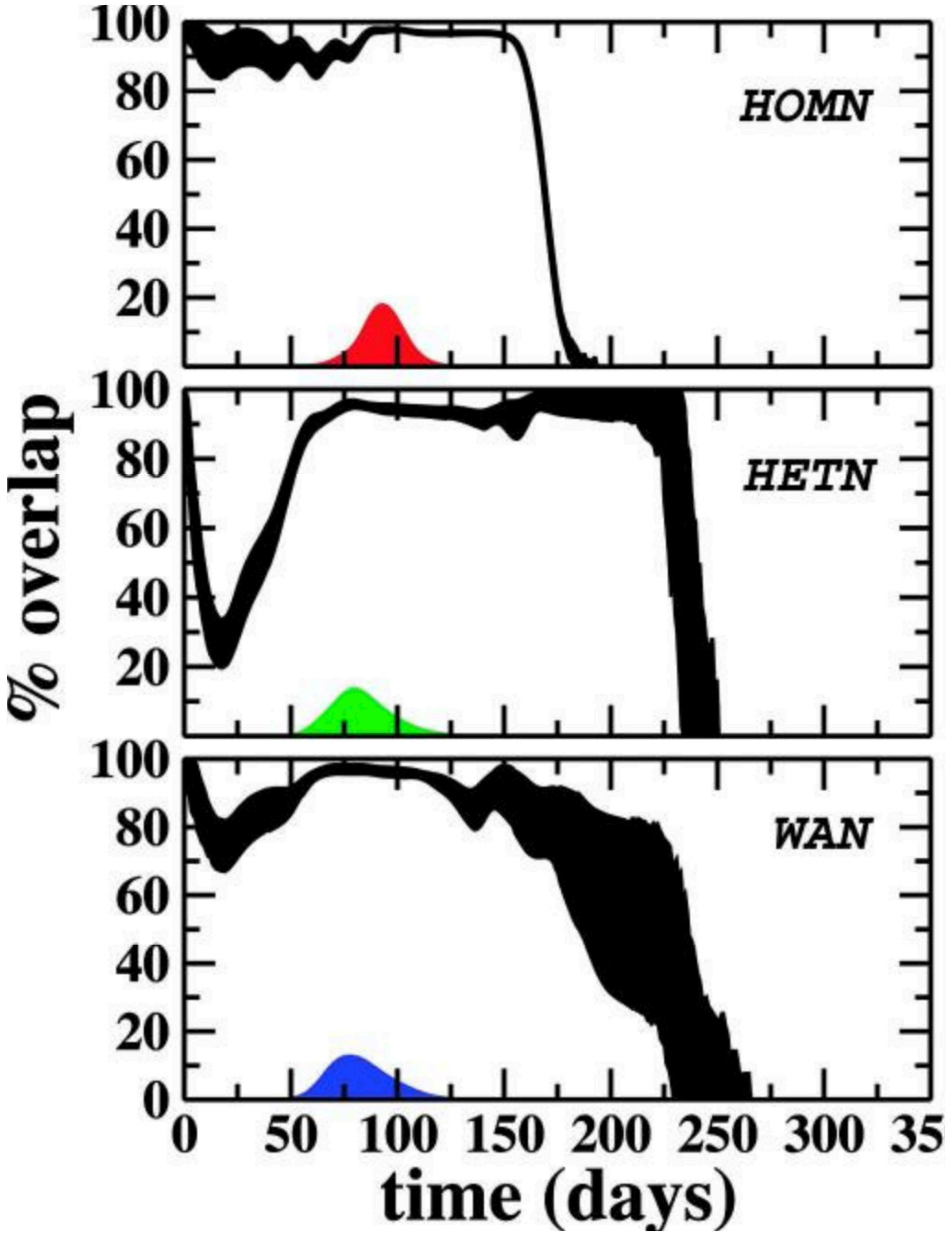
Colizza, Vittoria, et al. "The role of the airline transportation network in the prediction and predictability of global epidemics." *PNAS* 103.7 (2006)

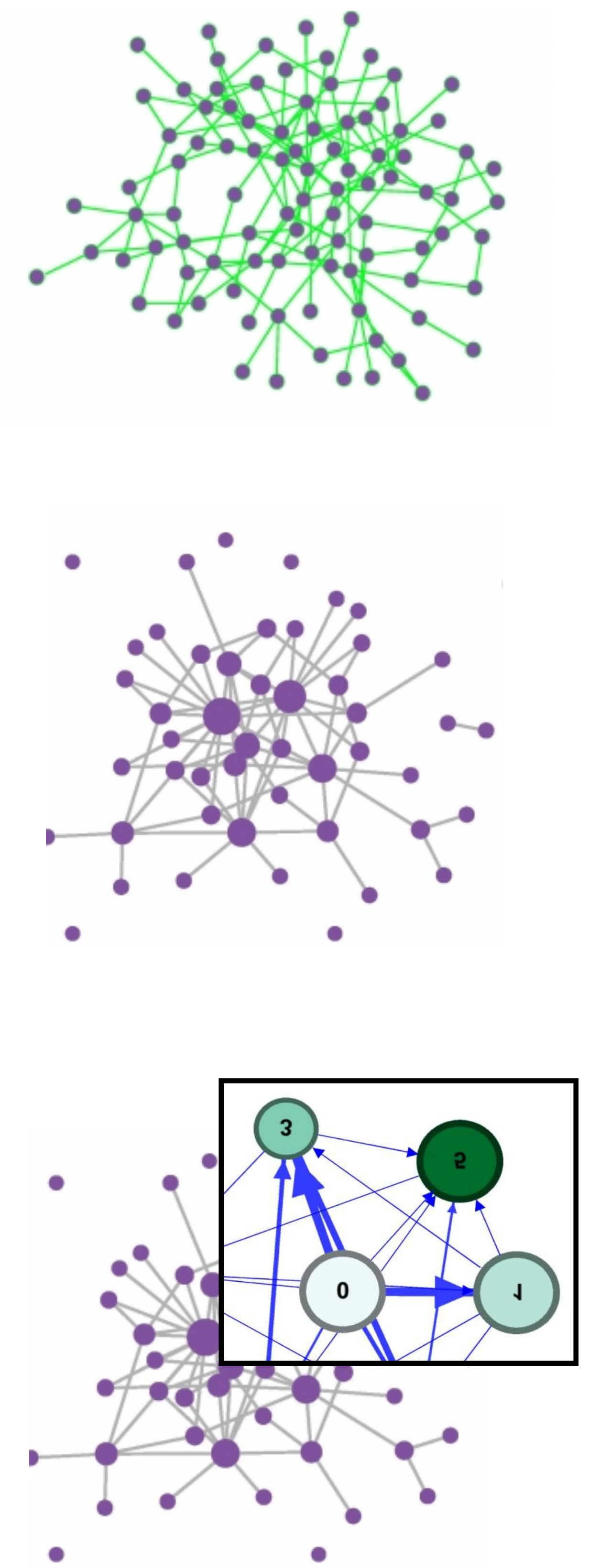
Pathways of spatial invasion

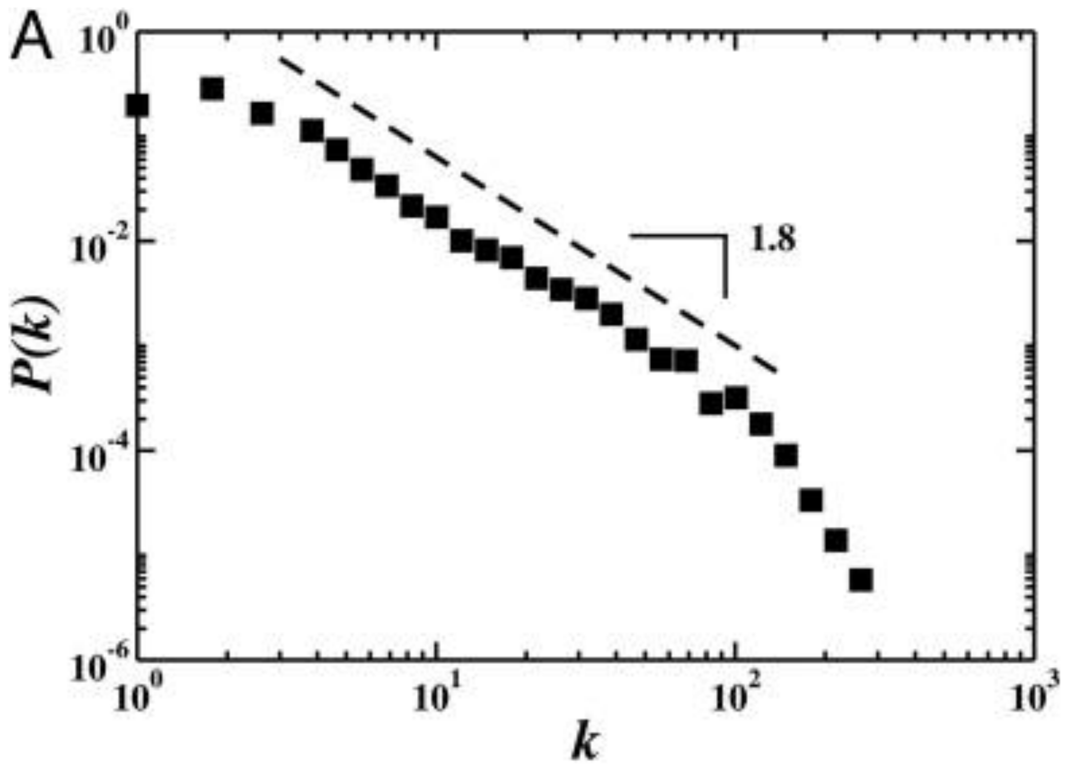


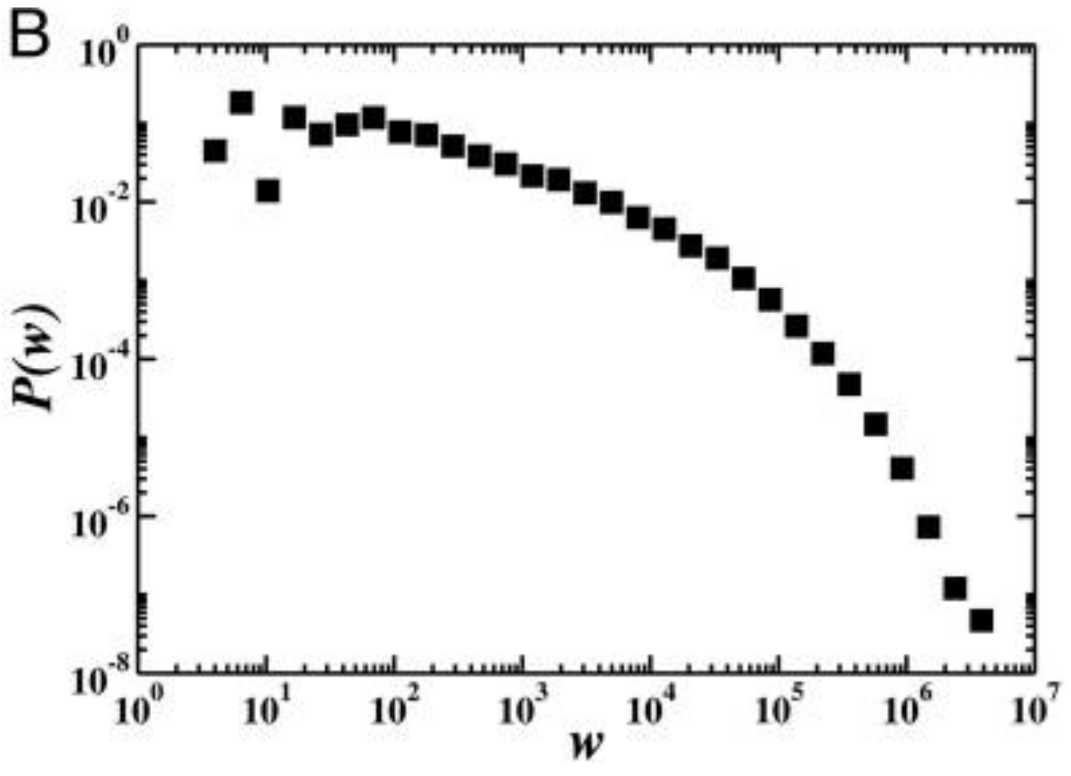


Pathways of spatial invasion









Two dynamics at play:

Degree heterogeneity
lowers predictability of spatial invasion

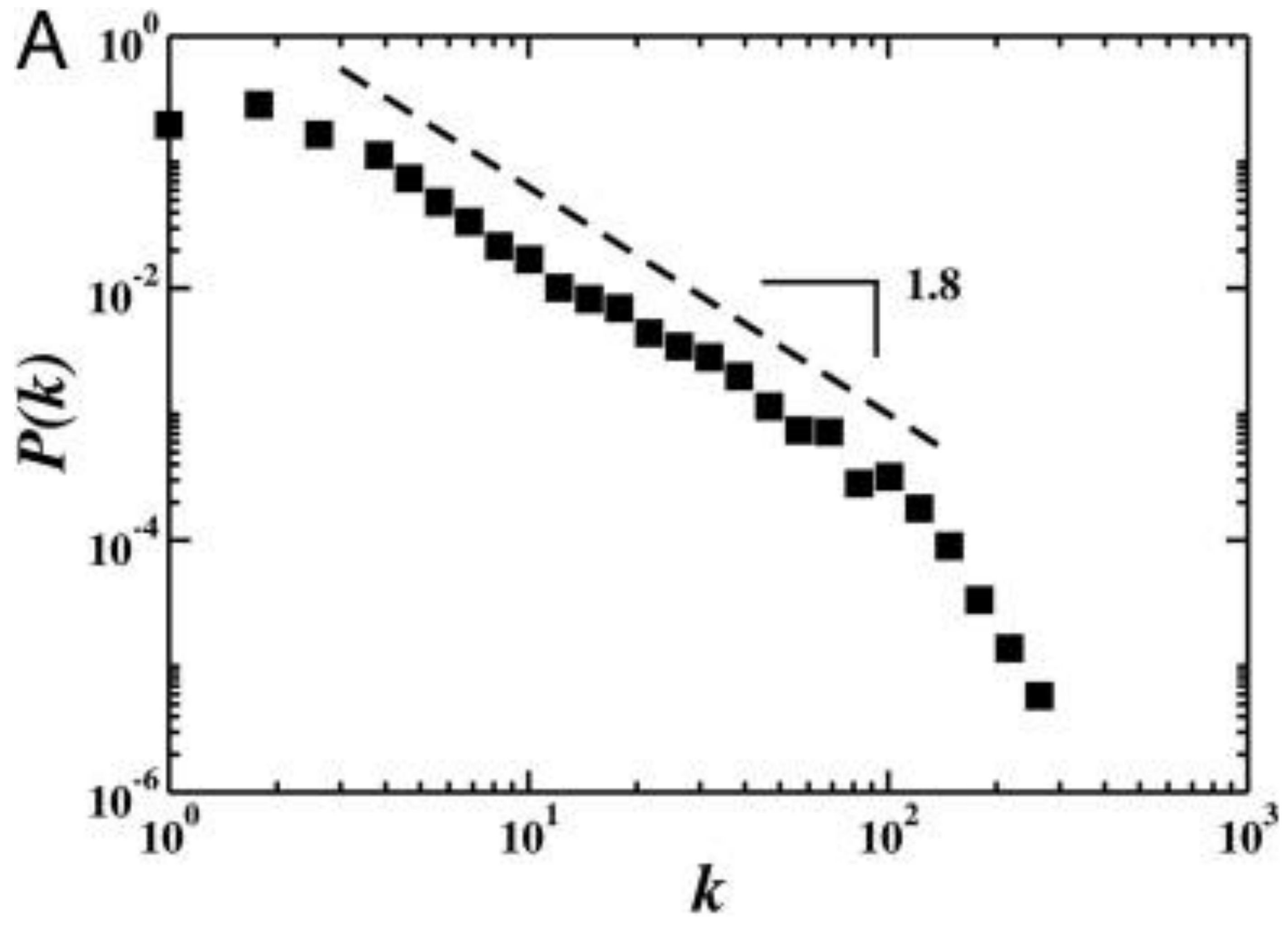
Weight heterogeneity
makes certain pathways more probable overall

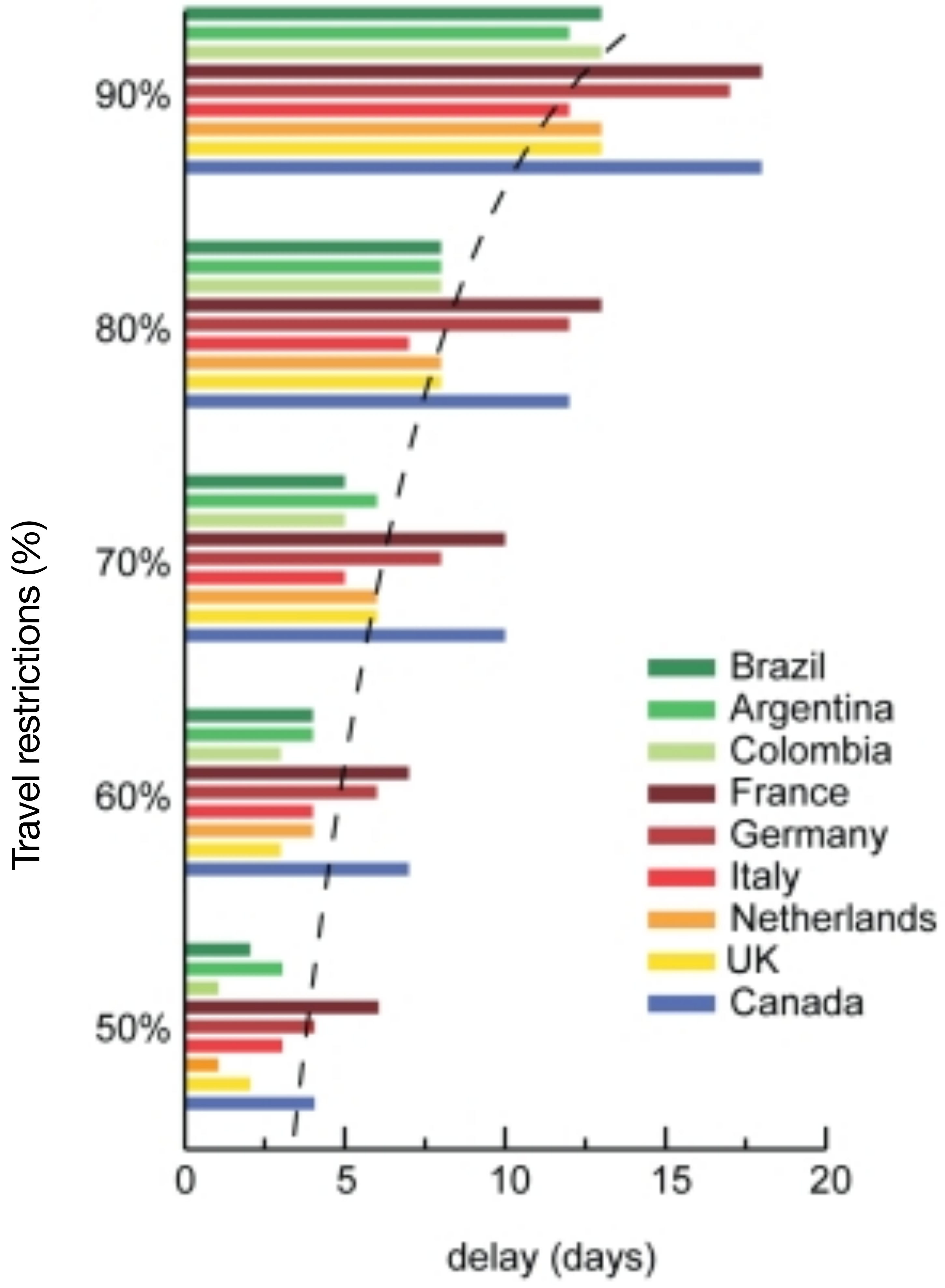
High heterogeneity at large times due to different lifetimes of epidemics in each simulation

Public health interventions on spatial invasion: air travel bans

Slowing down the international spread of diseases

Impossible to stop spread without suppressing > 99% of flights
Effect is due to the topology of the air travel network
Best you can do is delaying arrival of disease to anticipate and better prepare response



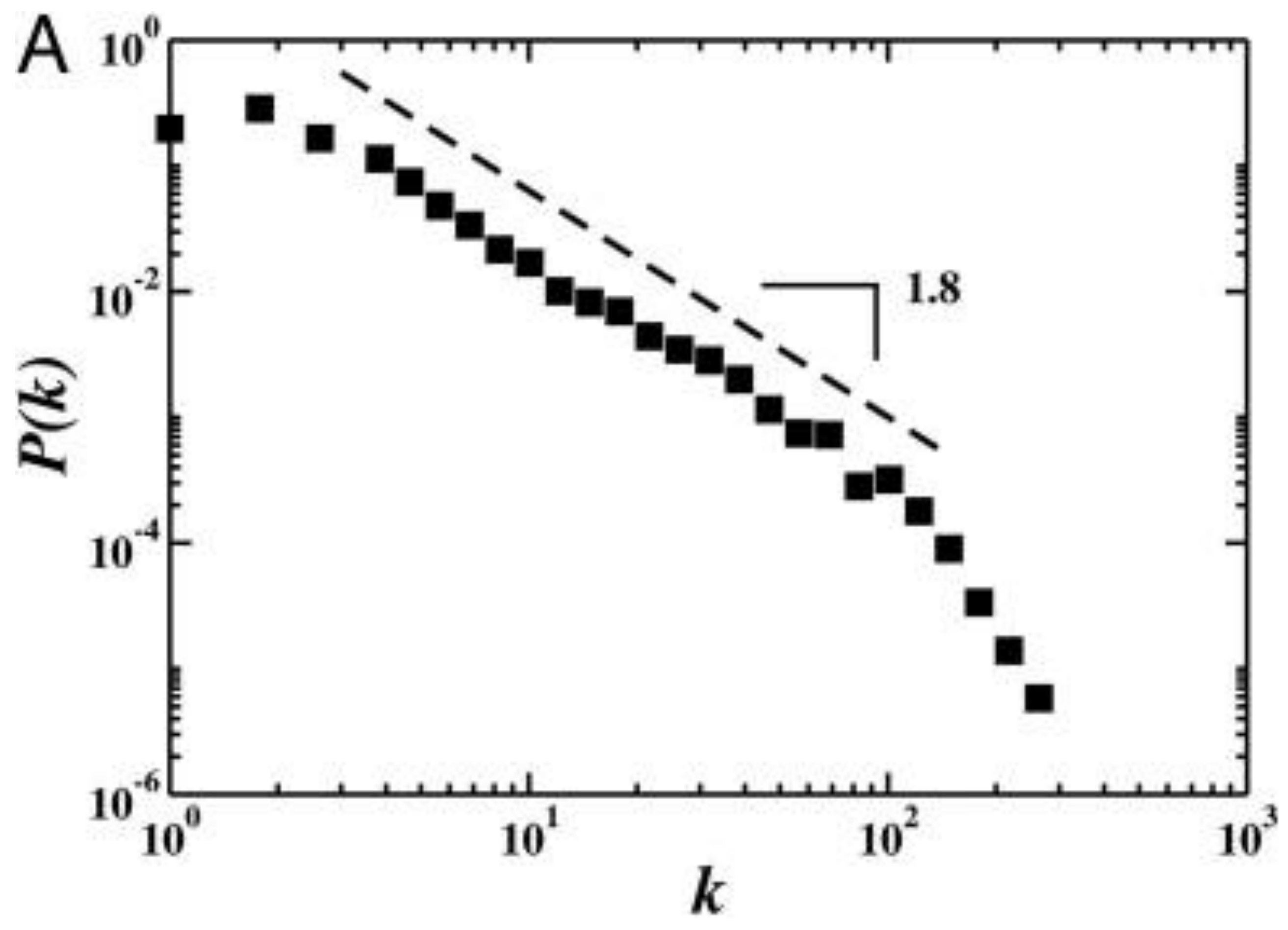


Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, Vespignani A. Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. PLoS One. 2011
Colizza V, Barrat A, Barthélemy M, Vespignani A. The role of the airline transportation network in the prediction and predictability of global epidemics. Proc Natl Acad Sci U S A. 2006

Public health interventions on spatial invasion: air travel bans

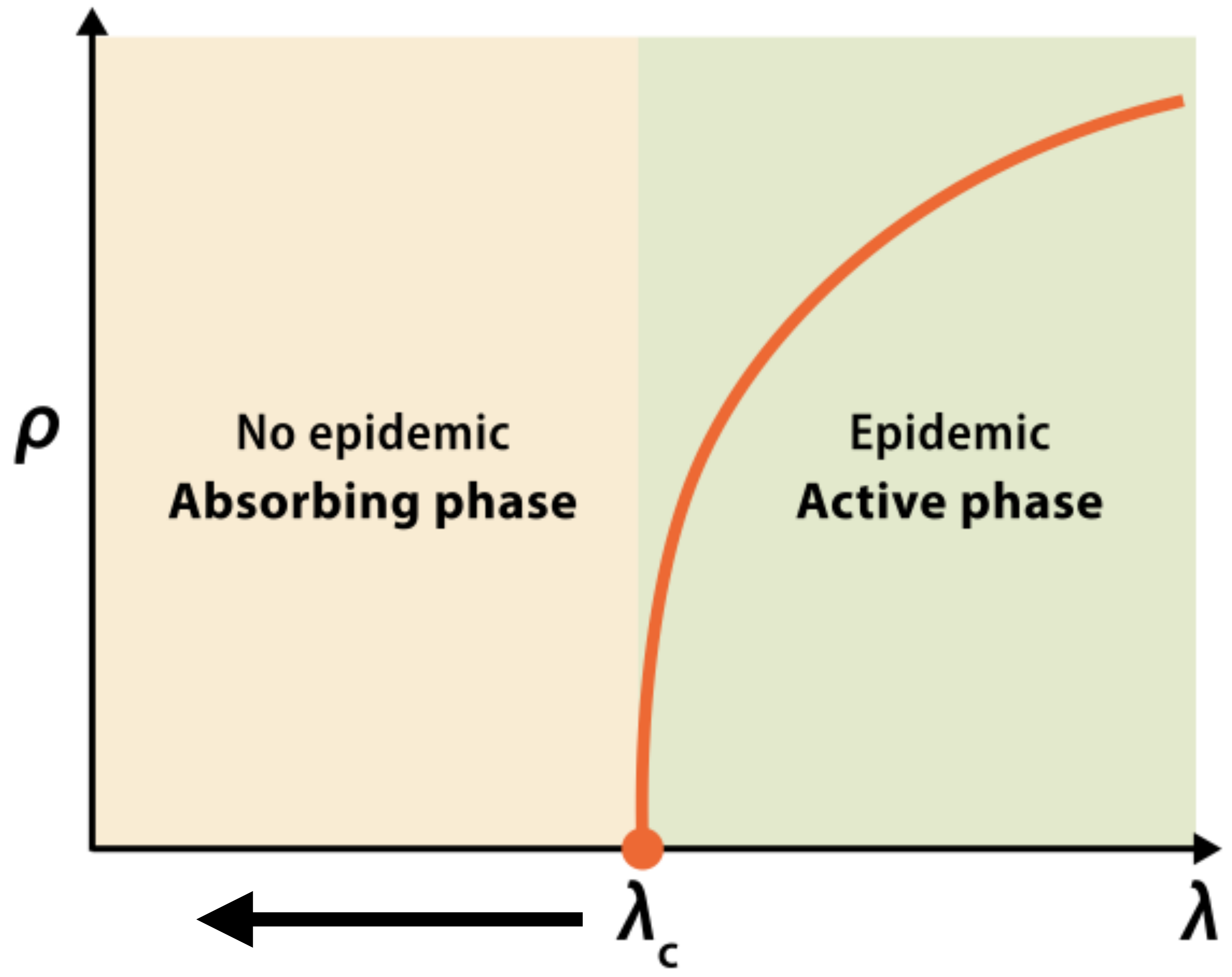
Why is it so hard?

Epidemic threshold affected by 2nd moment of degree distribution of the air travel network



$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

$$\begin{cases} \text{If } \gamma > 2 \\ \langle k^2 \rangle \to \infty \\ \lambda_c \to 0 \end{cases}$$



HANDS ON SESSION

Try this yourself! Go to page: https://epirisk.net/
Play with travel restrictions (reach ~99%), check the effect on the n of importated cases

Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, Vespignani A. Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. PLoS One. 2011
Colizza V, Barrat A, Barthélemy M, Vespignani A. The role of the airline transportation network in the prediction and predictability of global epidemics. Proc Natl Acad Sci U S A. 2006

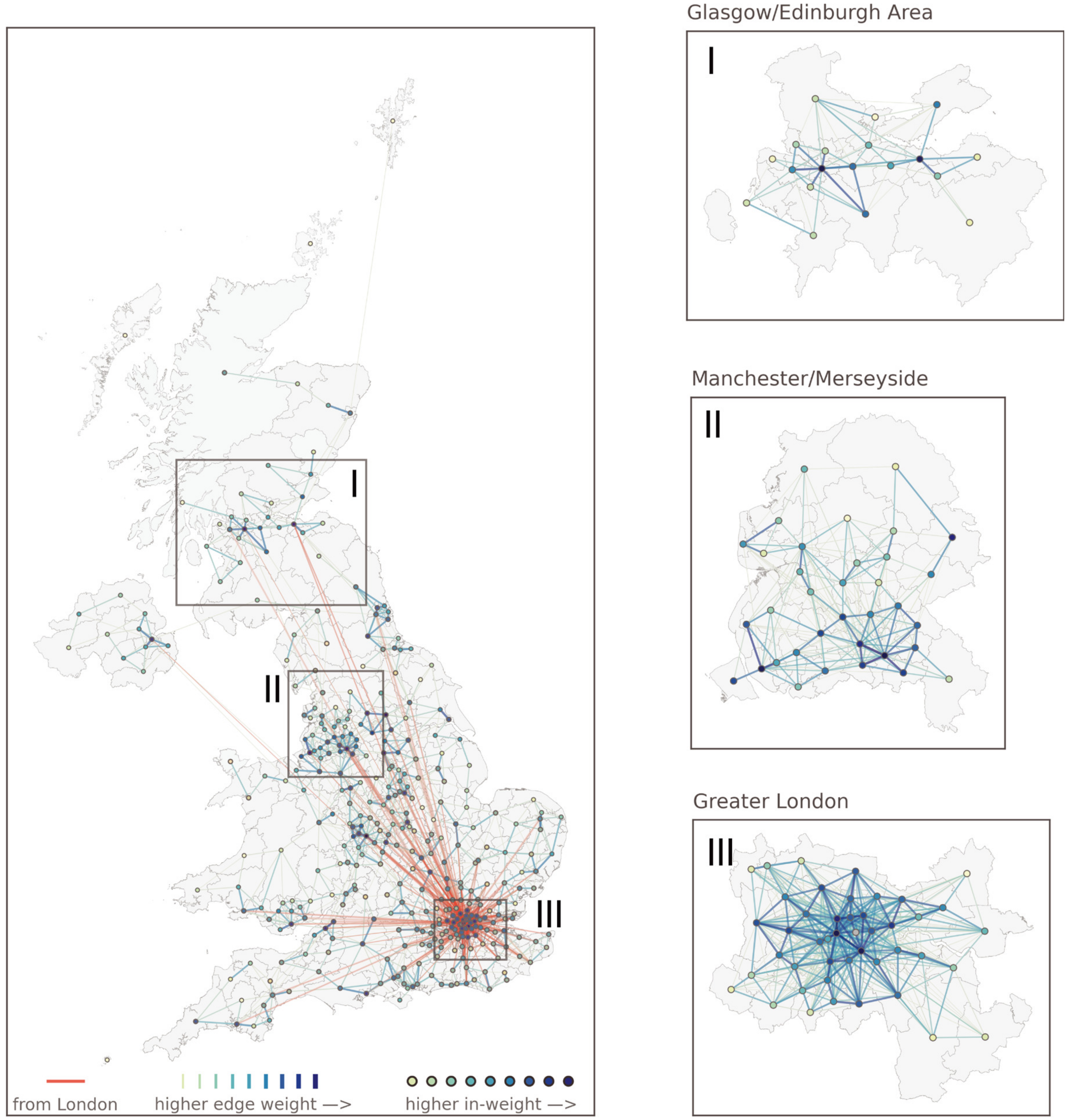
Sub-national scale: invasion trees

Mobility shapes spatial invasion of disease also at lower scales

SARS-Cov-2 Alpha variant in UK reconstructed through phylogenetic analyses

Phylogenetic models (BEAST) consider:

- mobility networks
- samples of viruses collected through tests
- tests report a location and date of collection

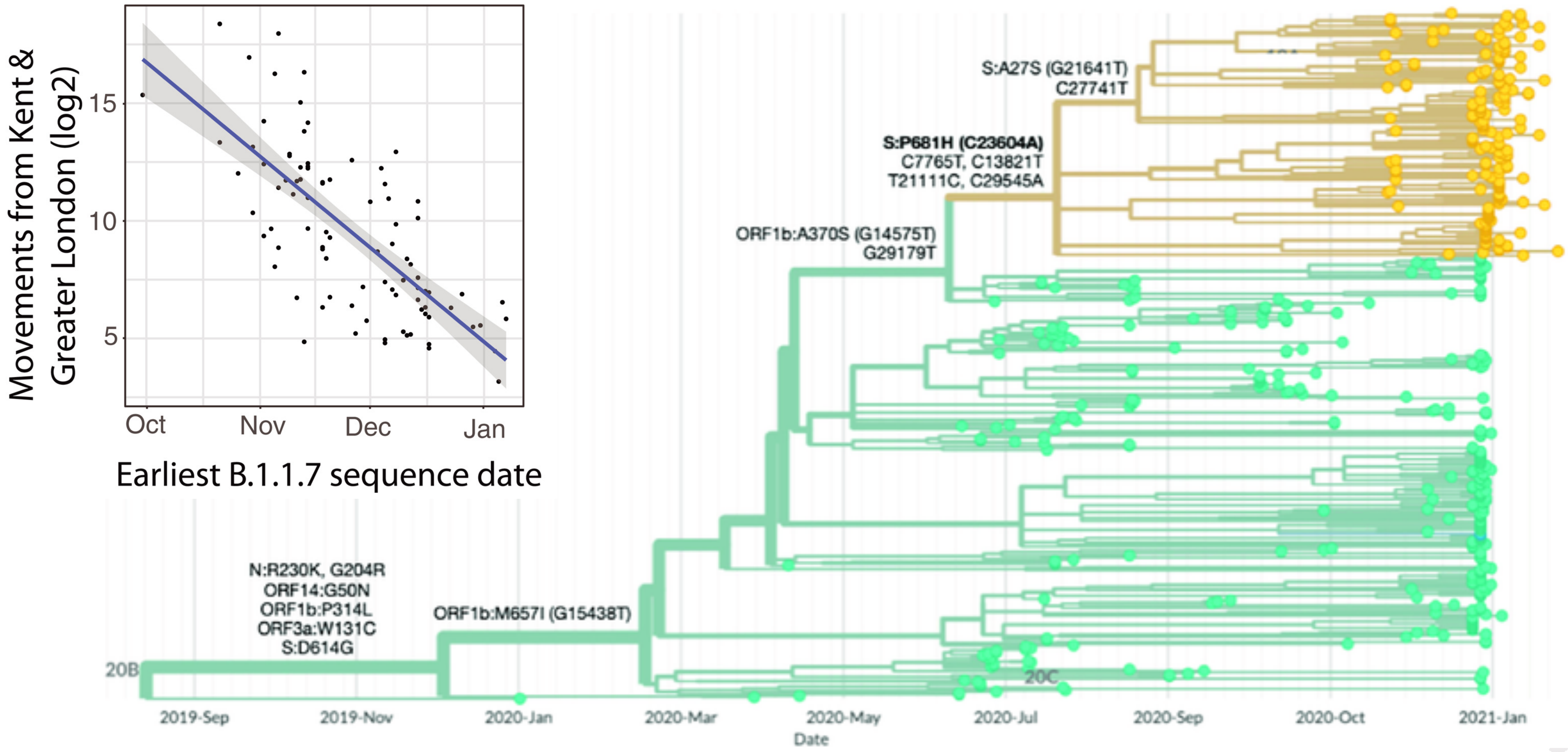


The model defines the most probable lineages (group of viruses with a common ancestor) and their spatial invasion dynamic in the country

Moritz U. G. Kraemer et al. ,Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence.Science 373,889-895 (2021)

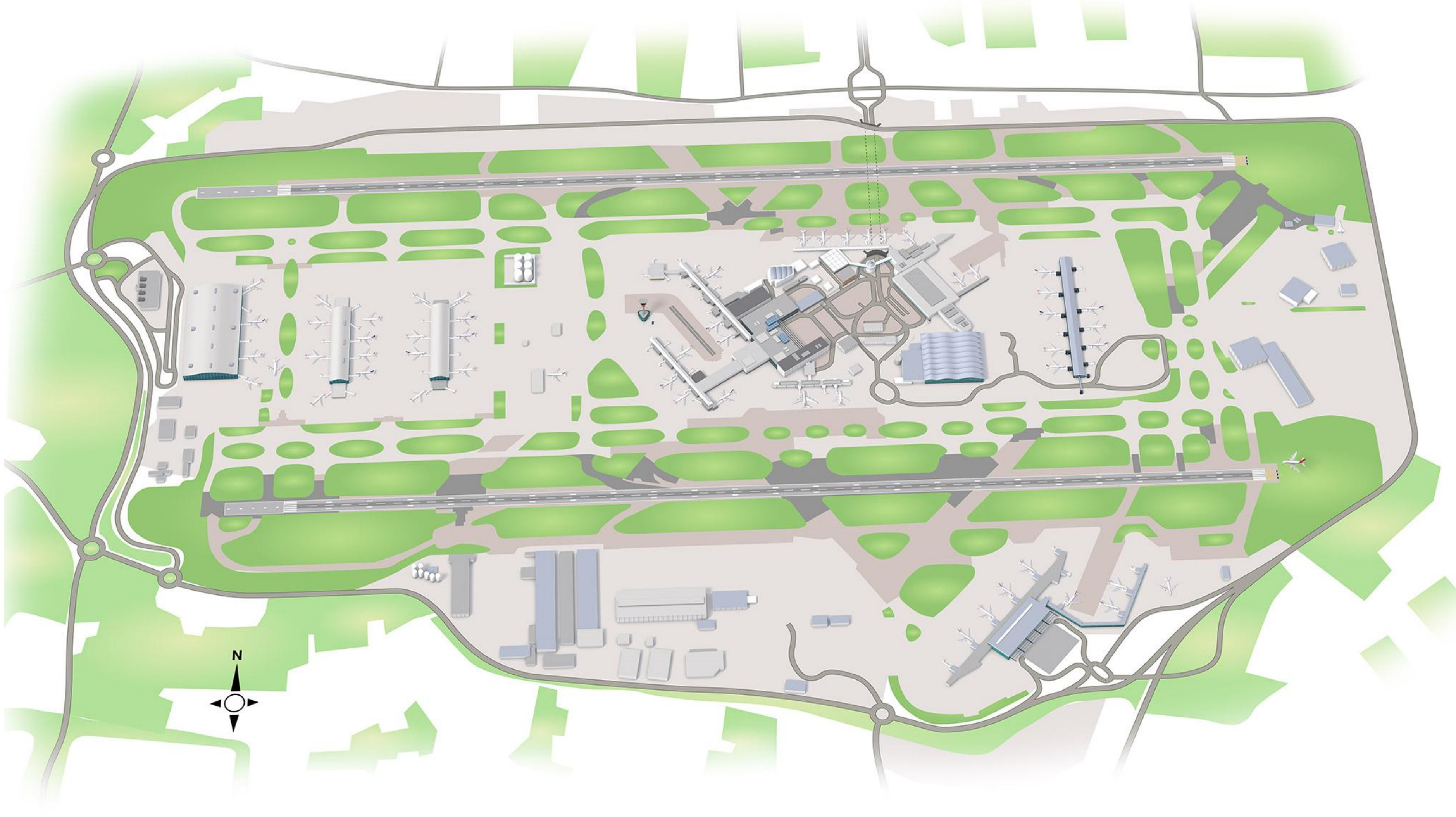
Sub-national scale: invasion trees

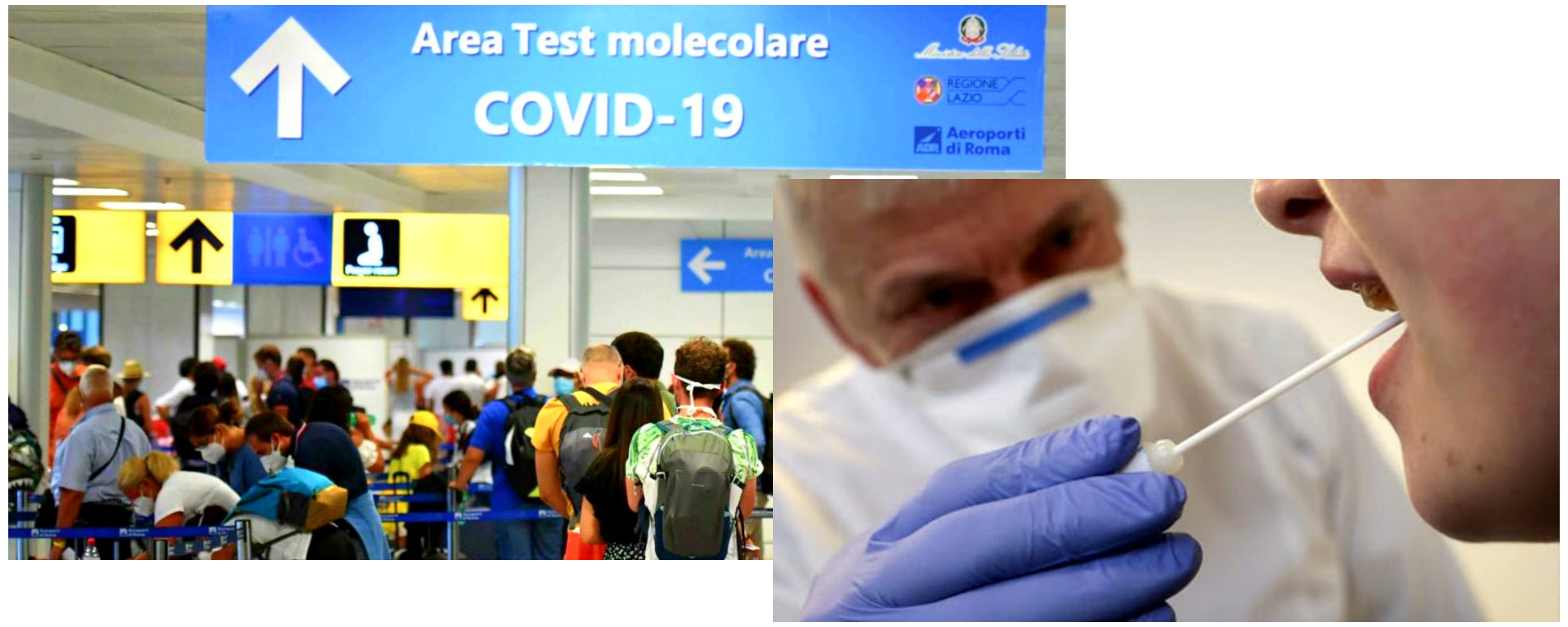
The result of the phytogeographic analysis

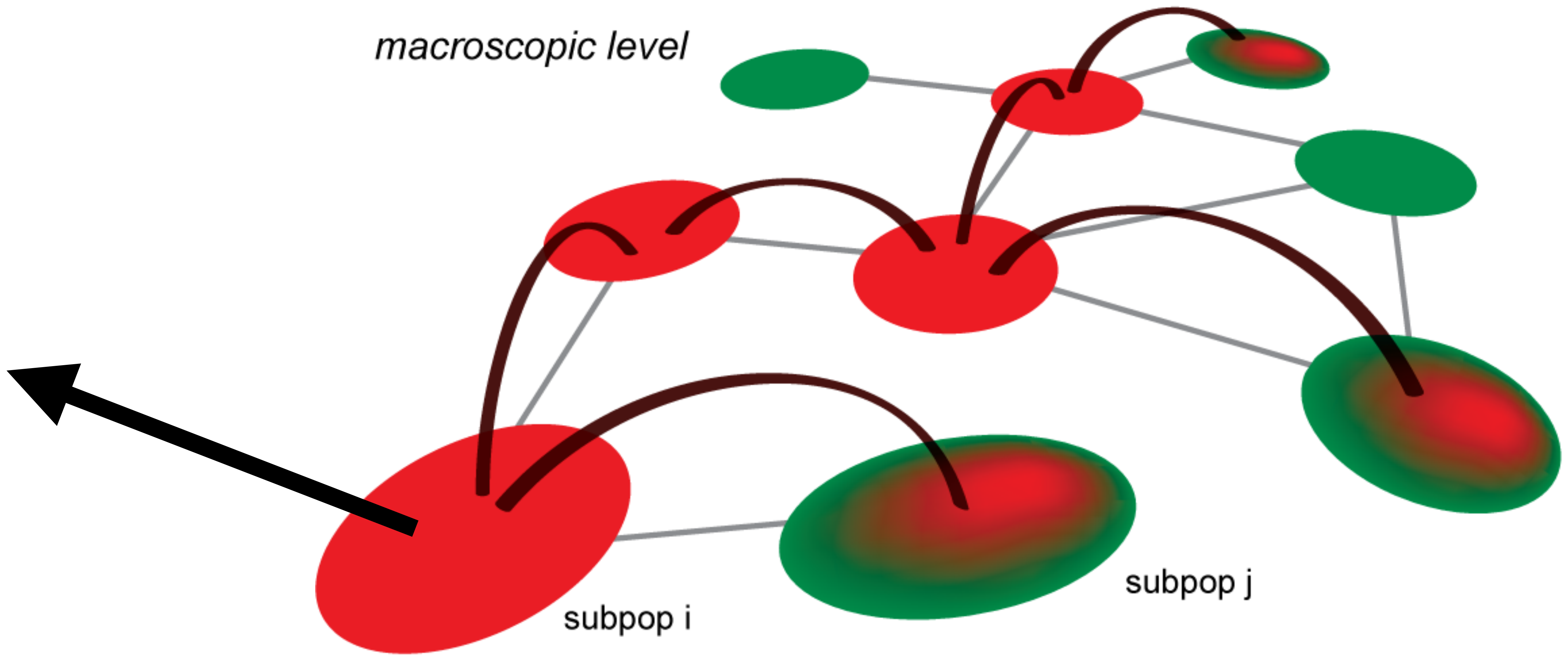


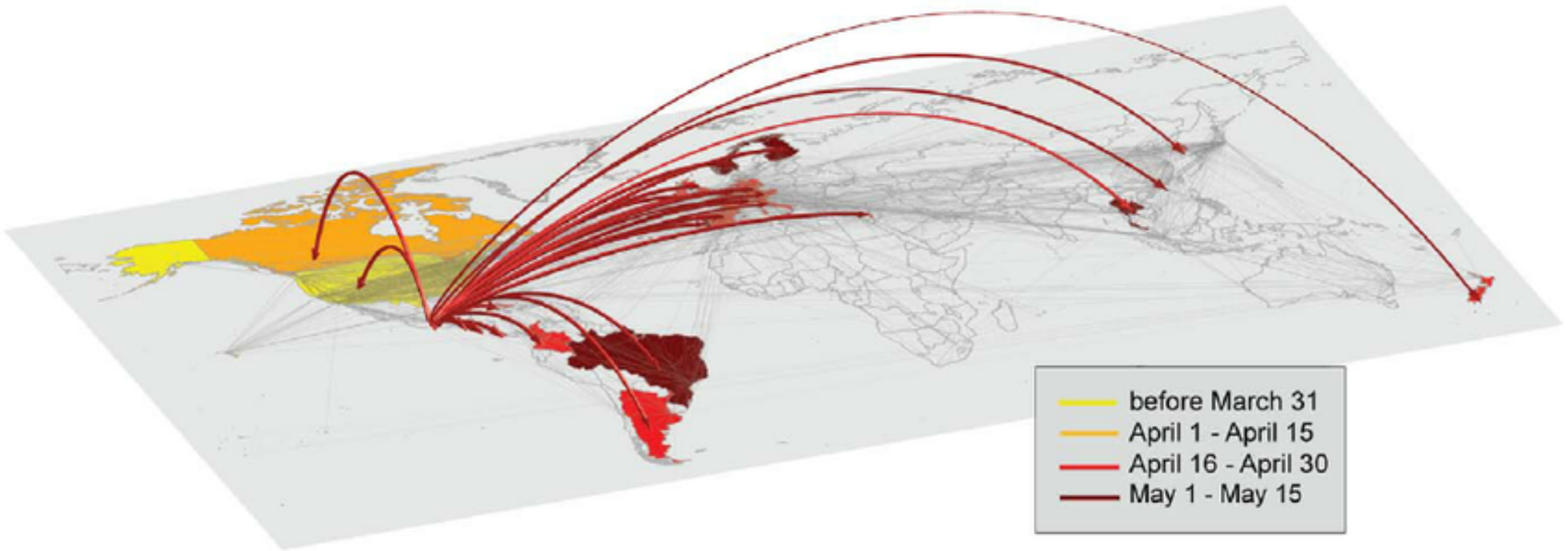
Moritz U. G. Kraemer et al. ,Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence.Science 373,889-895 (2021)
Zuckerman, N.S.et al. A Unique SARS-CoV-2 Spike Protein P681H Variant Detected in Israel. *Vaccines* (2021), *9*, 616.

Individual scale: co-presence network







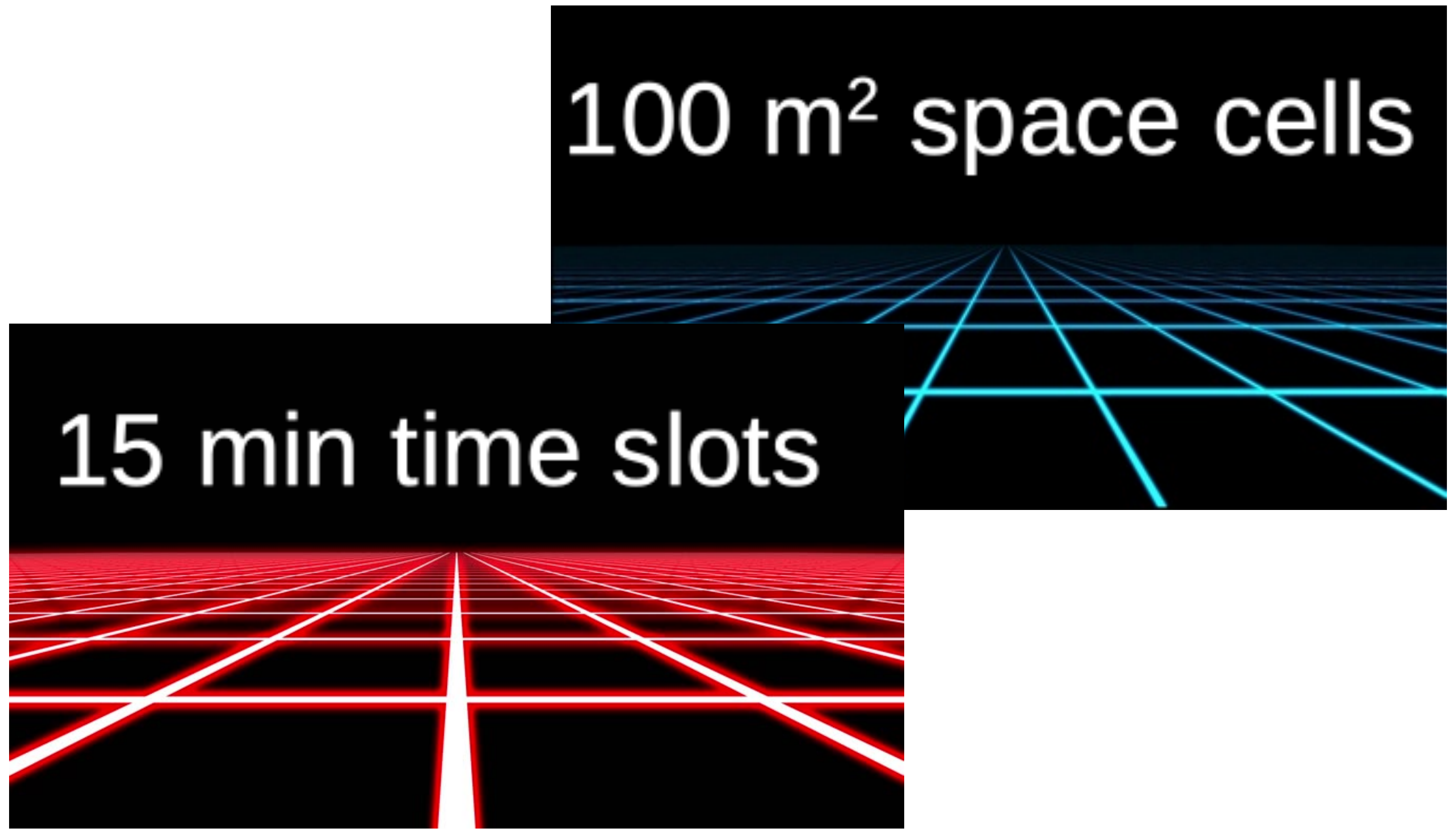


Mazzoli, M., Gallotti, R., Privitera, F. *et al.* Spatial immunization to abate disease spreading in transportation hubs. *Nat Commun* **14**, 1448 (2023)

Individual scale: co-presence network

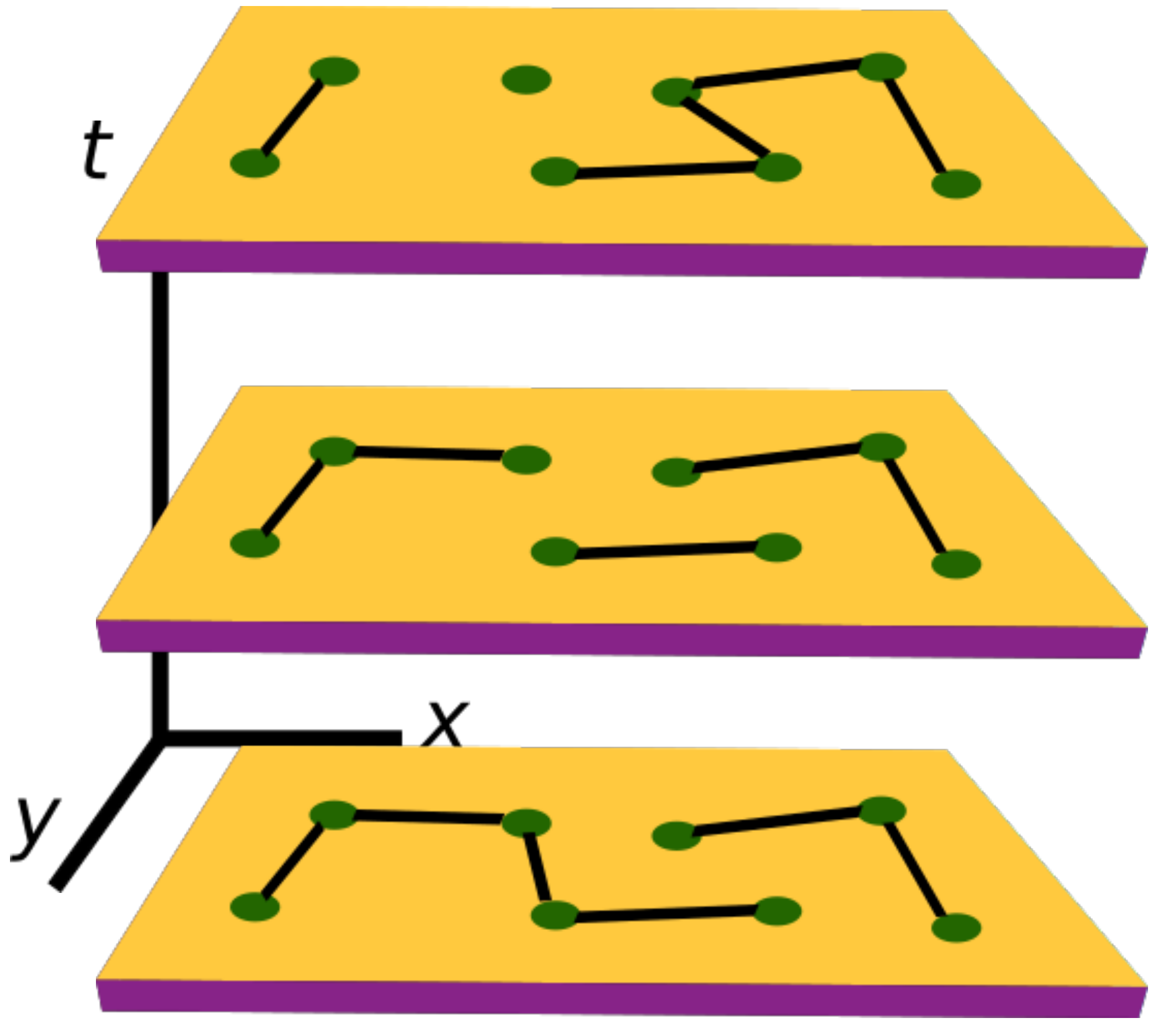
Data

Anonymized GPS trajectories of users who opted in to the service
Collected by Cuebiq from smartphone apps during 6 months in 2017
Structure: Hashed encrypted ID, (lon,lat), epoch time
Gps precision ~ 10m



Data treatment

- Trajectories discretization: discretize data points time and space
- Supersampling: move all the trajectories to one standard day
- Trajectories reconstruction: data driven optimal path
- Temporal contact network: contact defined as co-presence in 15’



Mazzoli, M., Gallotti, R., Privitera, F. *et al.* Spatial immunization to abate disease spreading in transportation hubs. *Nat Commun* **14**, 1448 (2023)

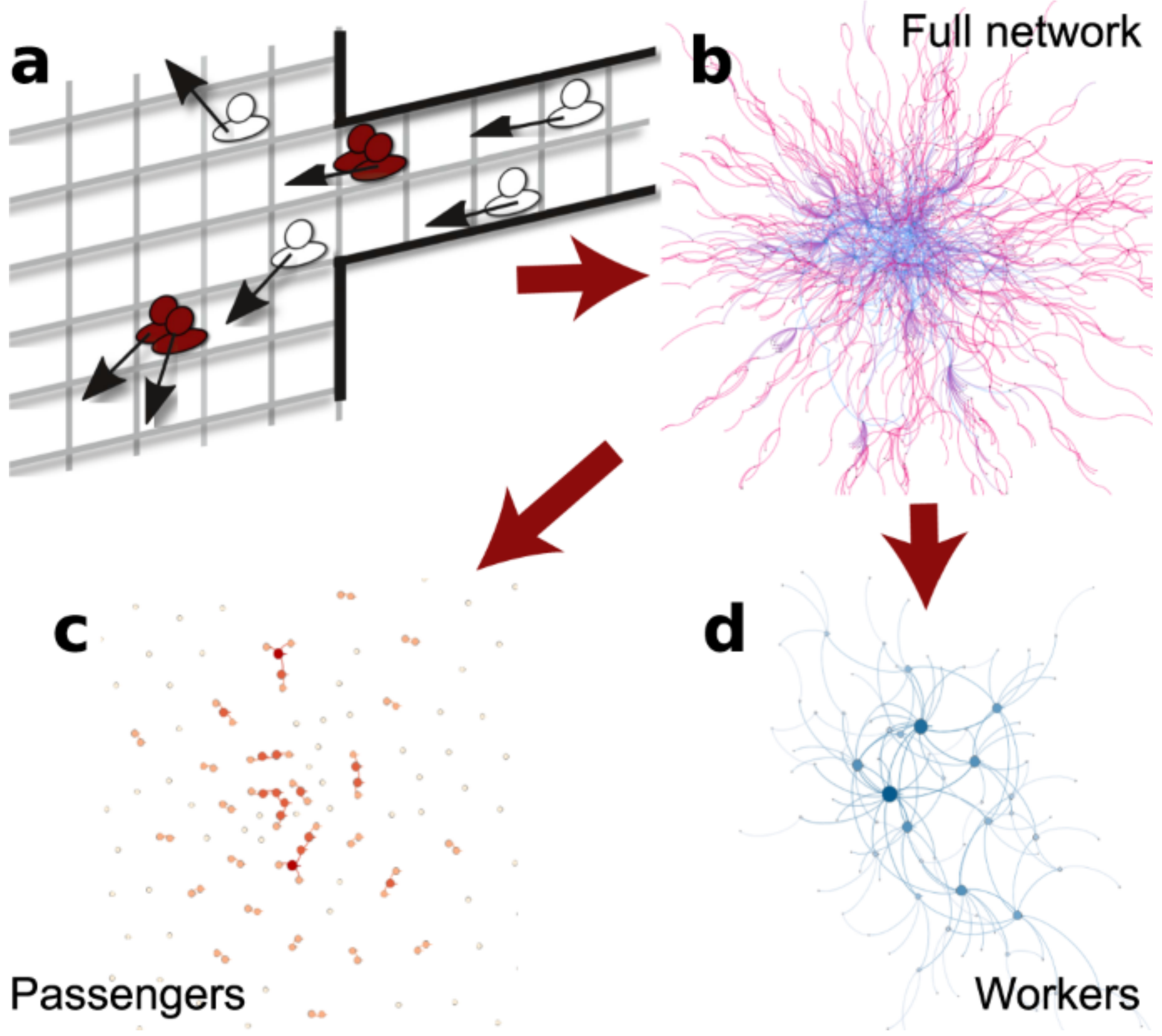
Individual scale: co-presence network

Far-UVC lamps -93/-99% activity of Corona-viruses and Influenza-viruses

Identify hotspots of contagion during the first day circulation

Apply immunization on identified hotspots \w 95% transmissibility reduction

Generalizable to other strategies (e.g. air filtering)
Results stable for lower efficacies

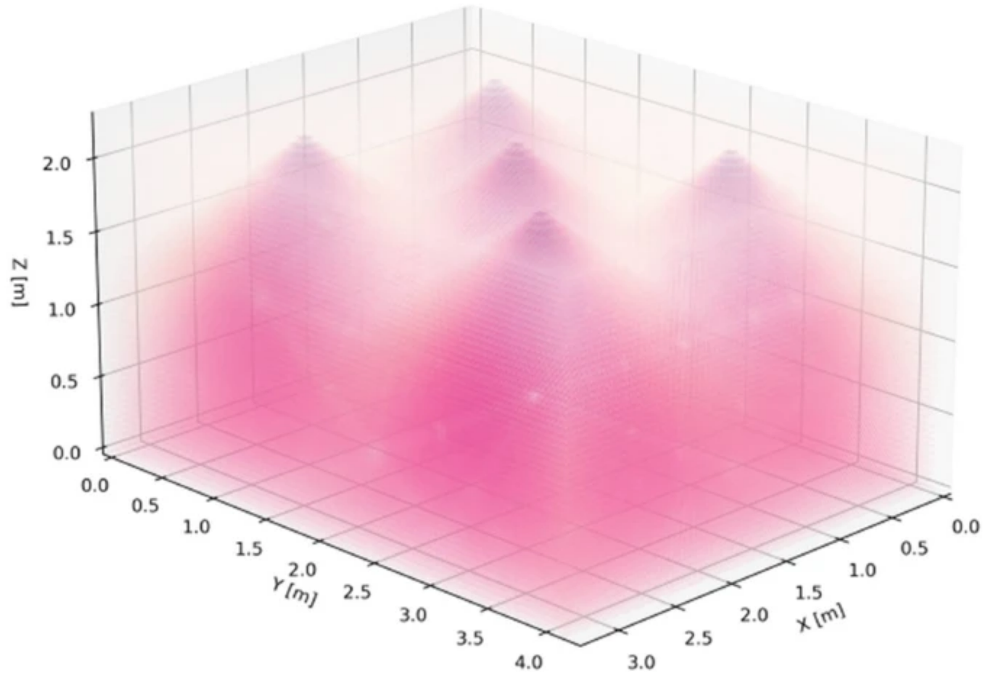


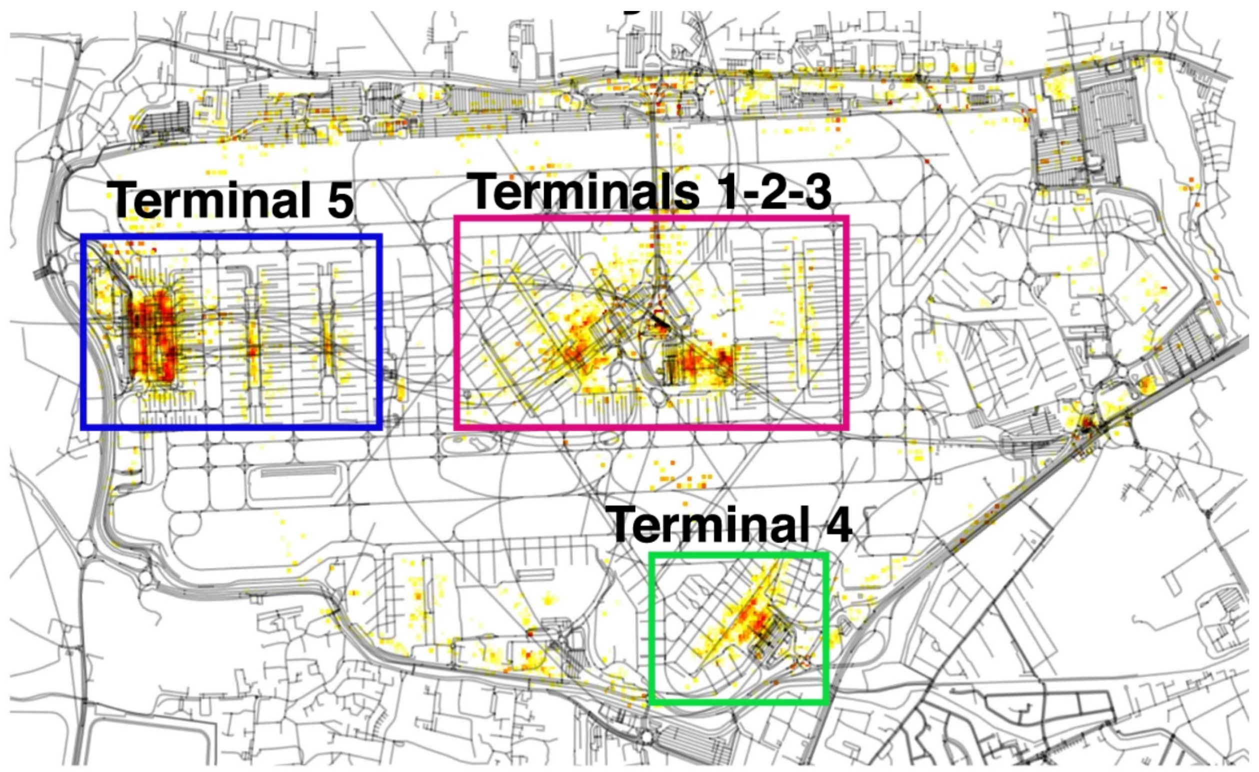
Model

ABM on temporal contact network

One seed arrives at the airport at 1:30 PM on the first day

One week simulations





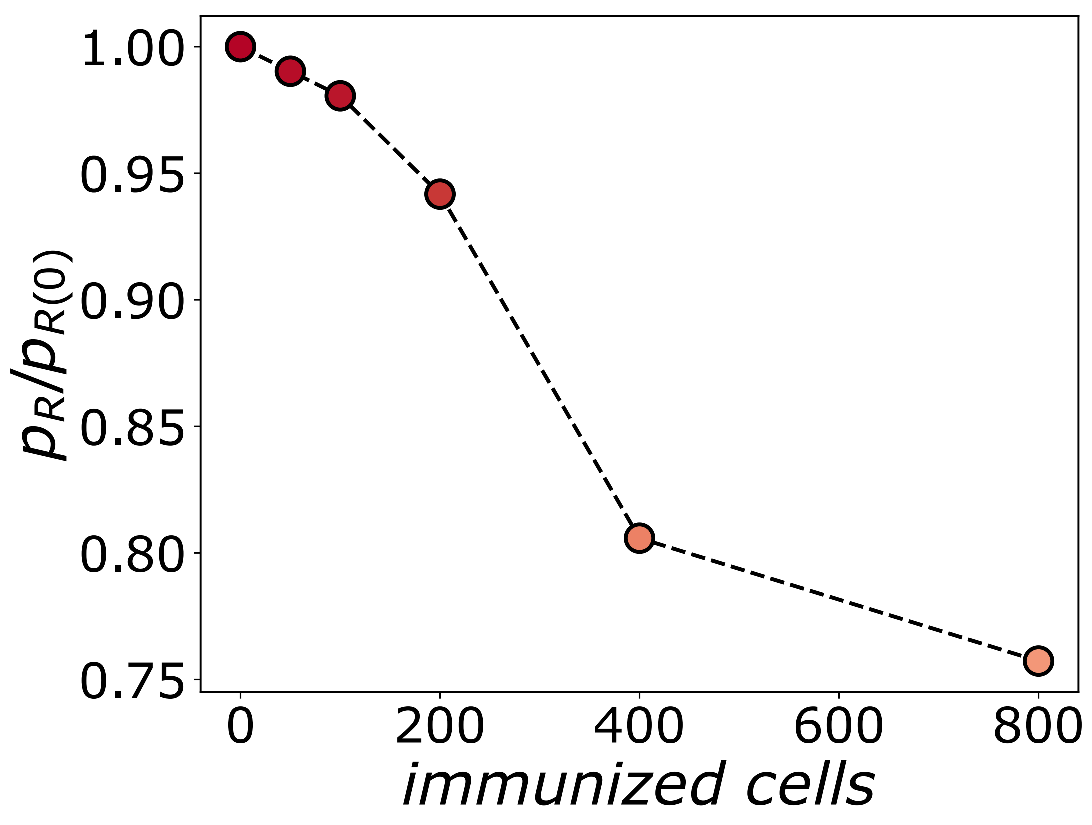
Eadie, E., Hiwar, W., Fletcher, L. *et al.* Far-UVC (222 nm) efficiently inactivates an airborne pathogen in a room-sized chamber. *Sci Rep* **12**, 4373 (2022).
Mazzoli, M., Gallotti, R., Privitera, F. *et al.* Spatial immunization to abate disease spreading in transportation hubs. *Nat Commun* **14**, 1448 (2023)

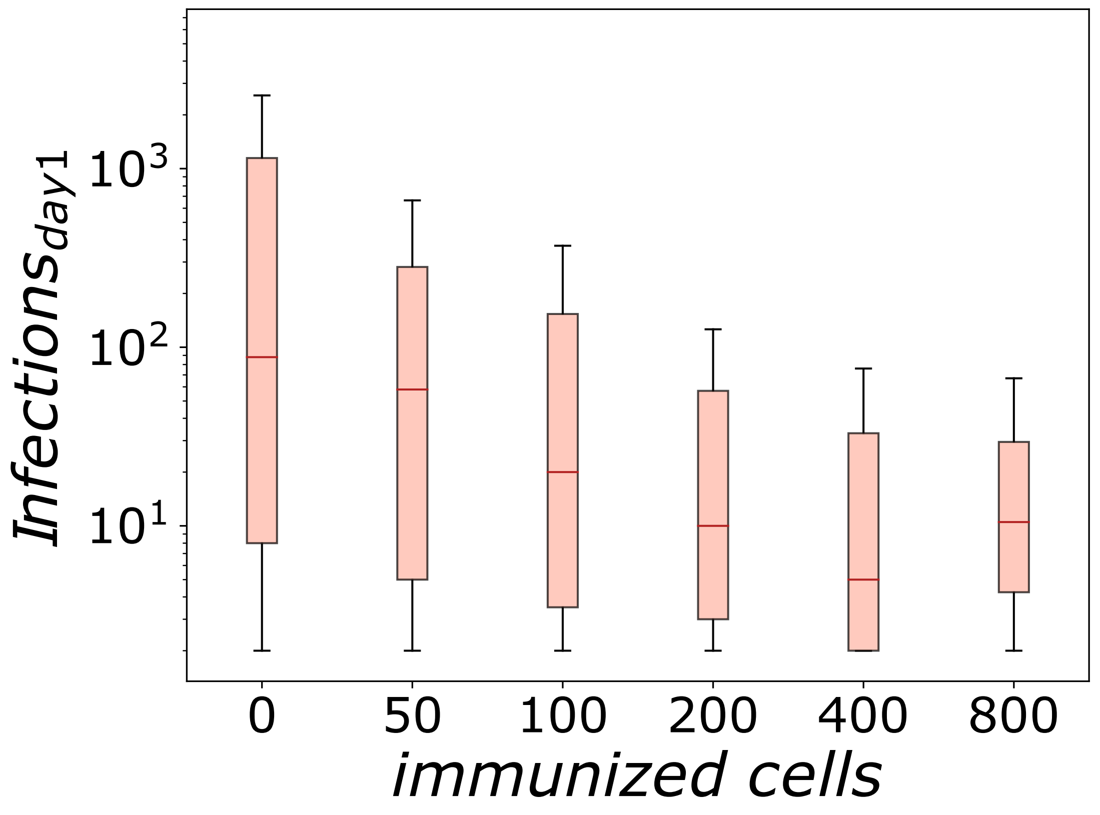
Individual scale: co-presence network

SIR model parameterized on **SARS**

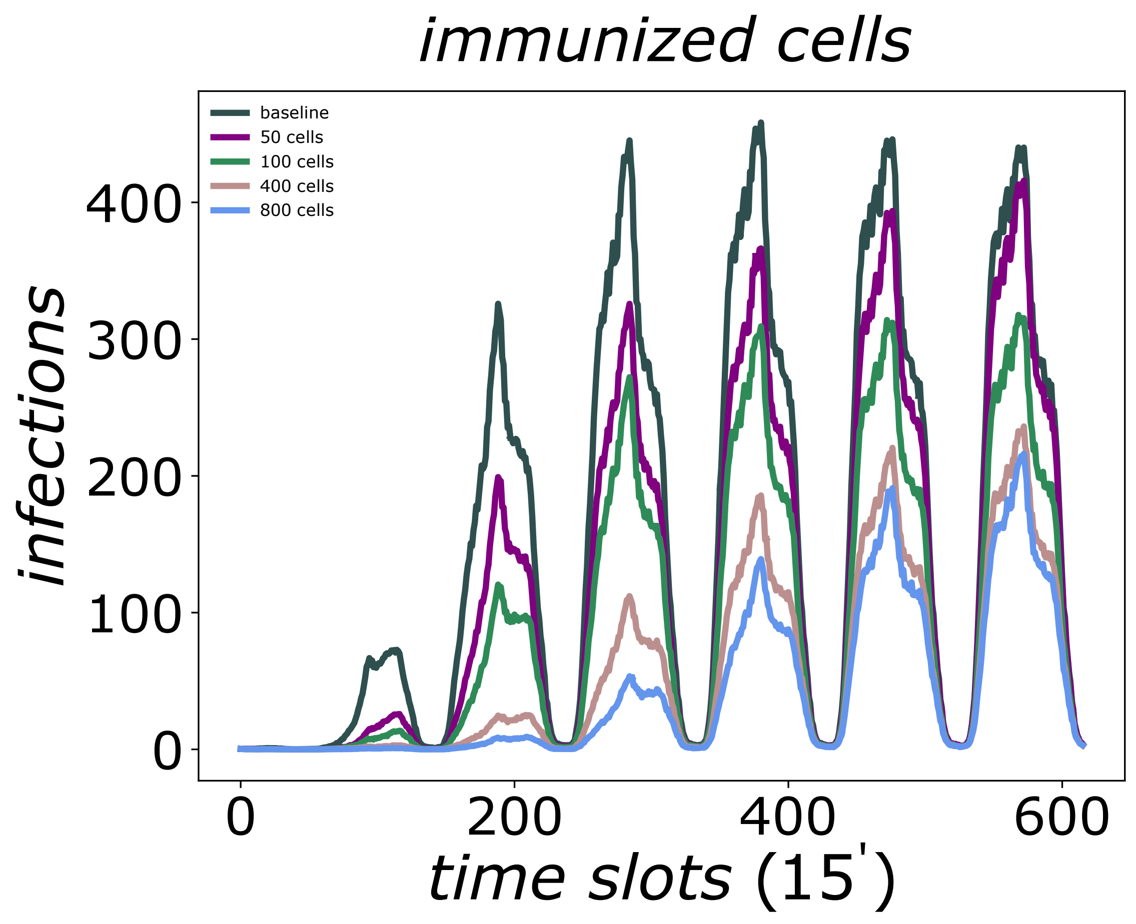
- -25% prob of developing a secondary case by acting on 2% of the available space

- Strong reduction in outbreak intensity in 24h





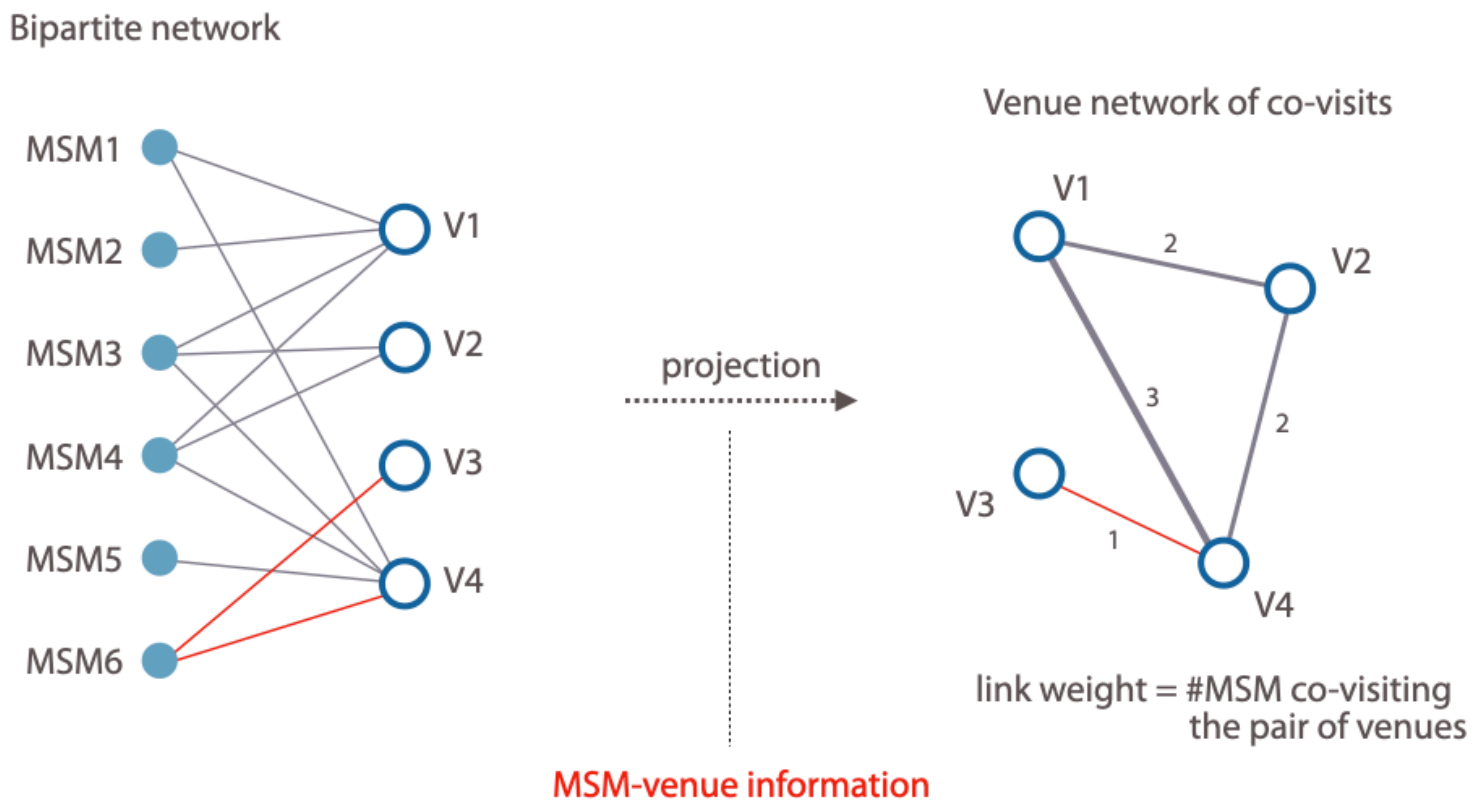
- Reduction in outbreak intensity holds on during the week
  —-> gain time for contact tracing of the first detected cases

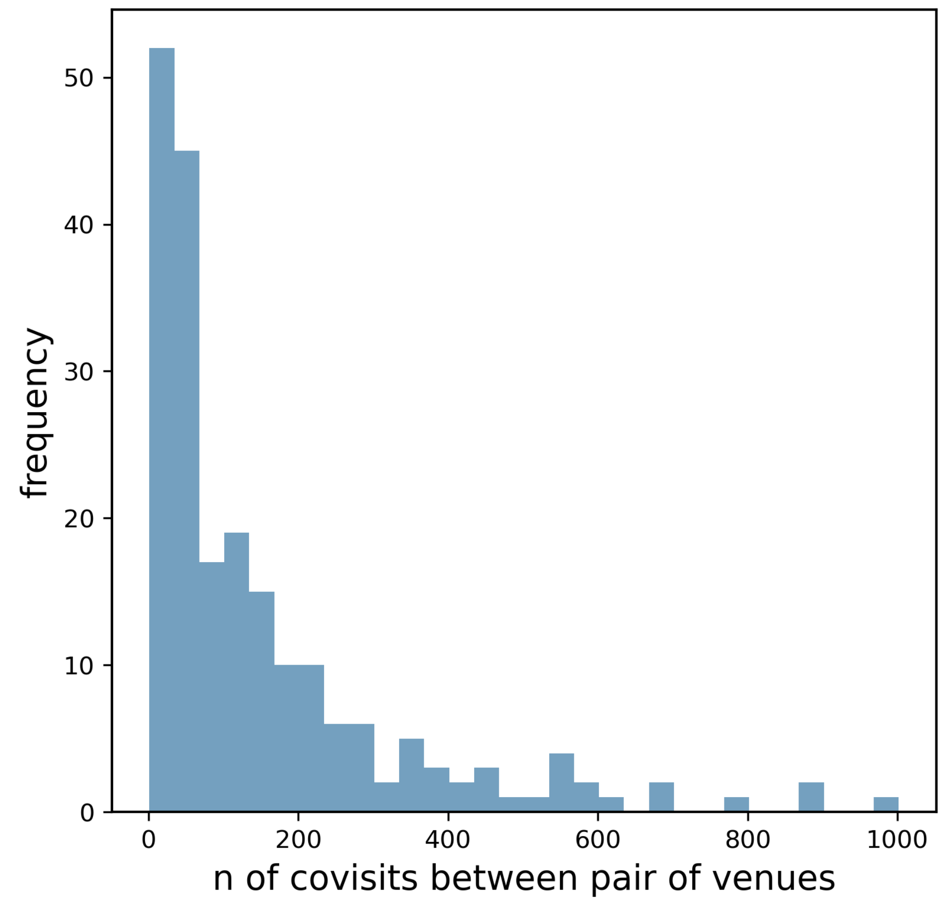


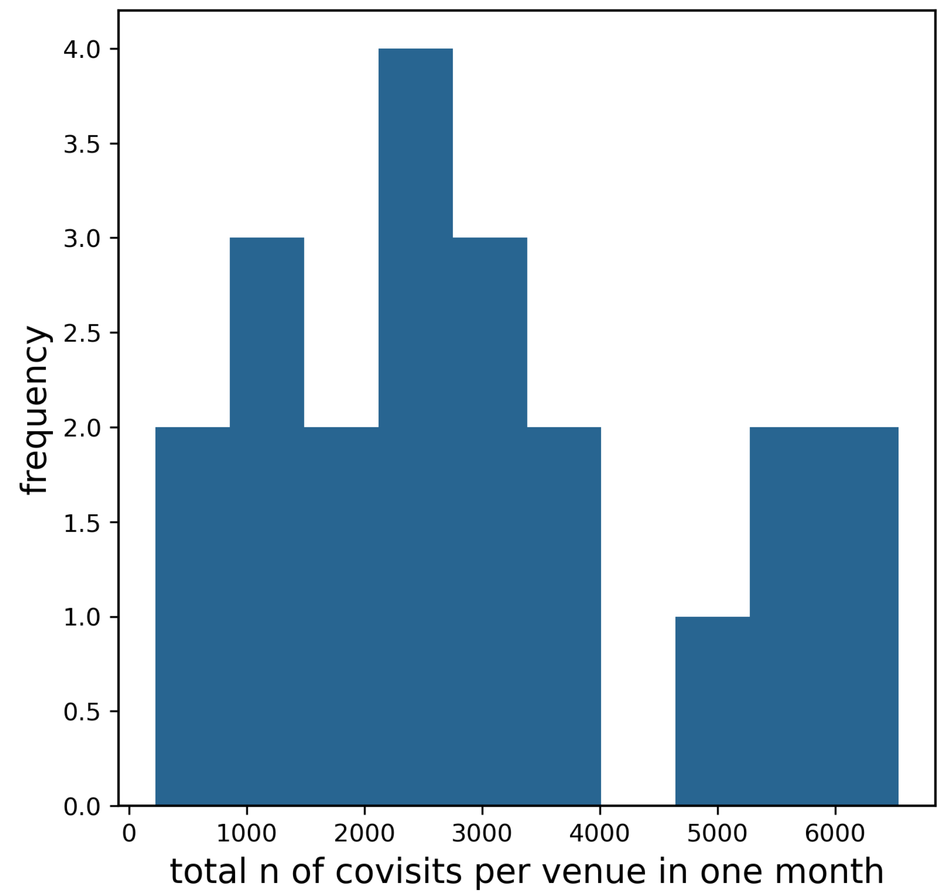
Mazzoli, M., Gallotti, R., Privitera, F. *et al.* Spatial immunization to abate disease spreading in transportation hubs. *Nat Commun* **14**, 1448 (2023)

Venue networks

60. Au cours du DERNIER MOIS, combien de fois avez-vous fréquenté les établissements suivants ?
Une réponse par item
Bornes : 0 à 50

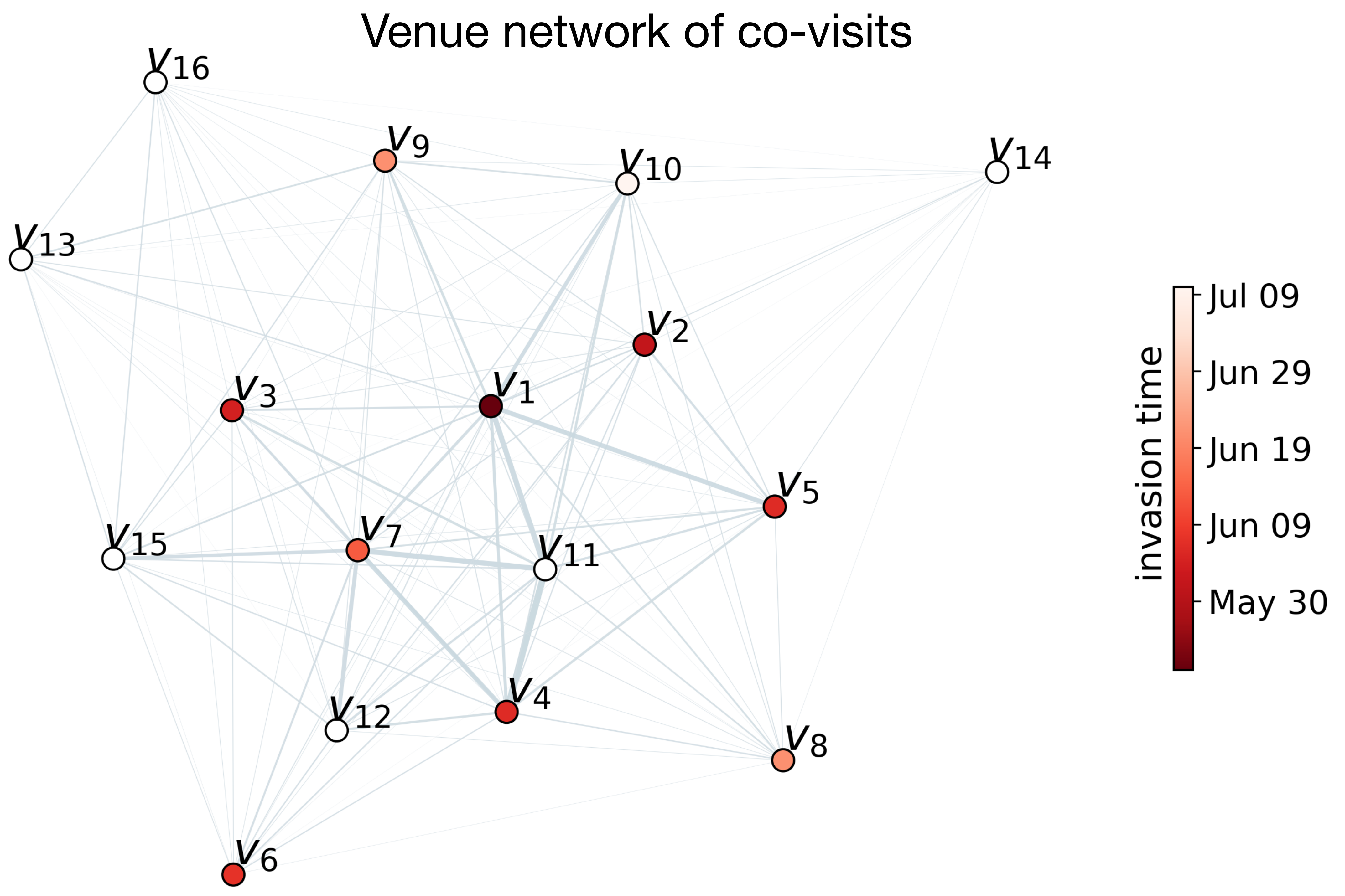






Venue networks

Préciser le(s) lieu(x), ~~date(s)~~ de ces rassemblements (dont les soirées dans des bars ou lieux privés) **durant les 3 semaines précédant le début des symptômes** (là où le cas a pu se contaminer)



From MPX post-infection survey —> venue 'invasion time'

- Venues dates of invasion assumed as earliest date of onset declared by any MPX case

Venue invasion time = date of symptoms onset of first case who declared visiting the venue in the previous 21 days

Venue networks

- Can we reproduce the dynamic of invasion among the venue network of co-visits?
- What feature of the co-visits network is the most important to describe it?

Venue networks

SI model of epidemic diffusion between venues

$\lambda_{i \to j} = \beta w_{ij}$ risk of transmission from venue *i* to venue *j*

$w_{ij} =$ (normalized) #co-visits between venue *i* and venue *j*

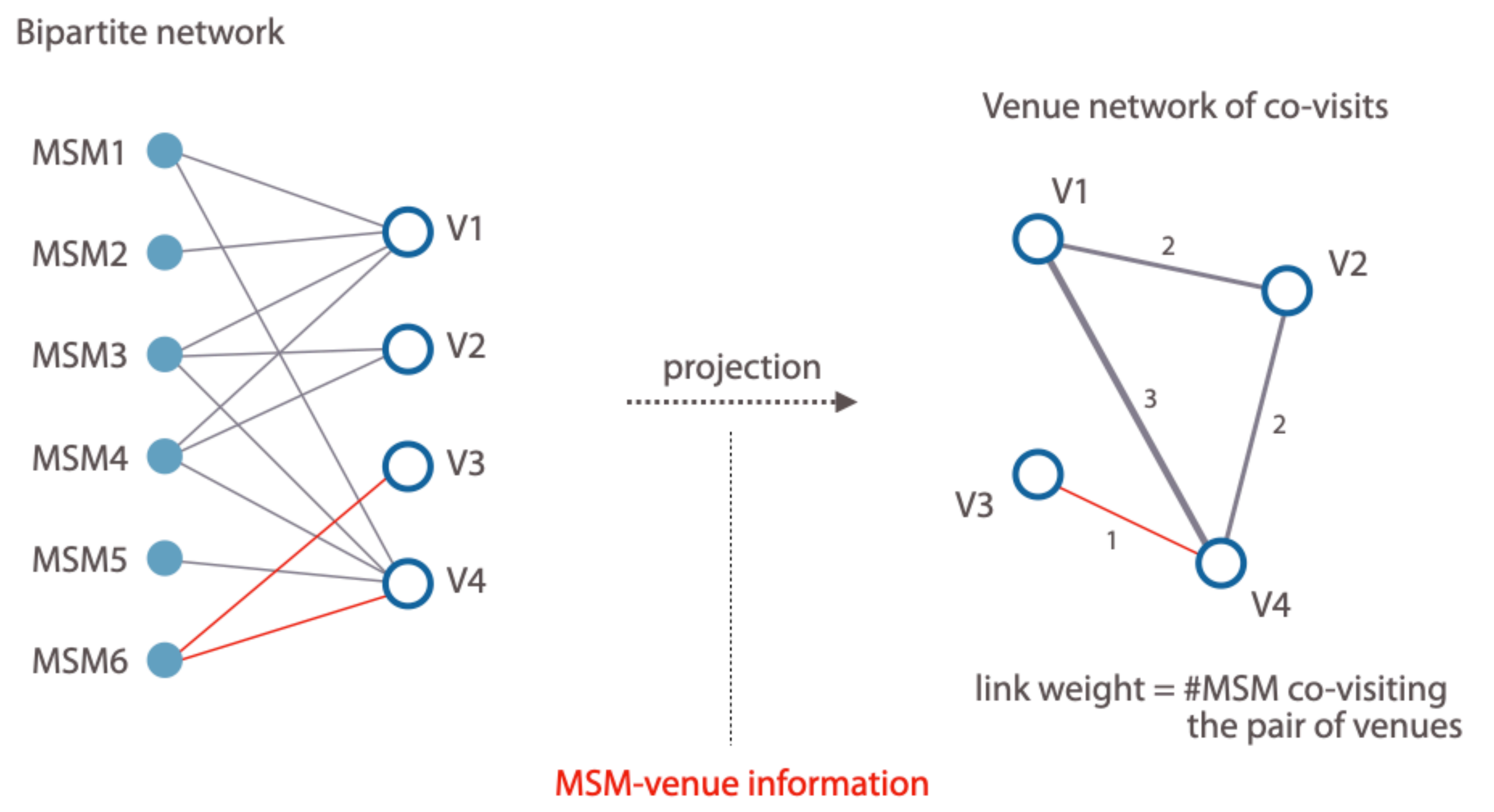
Eggo et al. J Roy Soc Interface 2011,
Cauchemez et al. Eurosurv 2014

Model of risk of invasion	Risk of transmission from venue *i* to venue *j*	Element tested in the permutation
Empirical venue network of co-visits from standard projection (MSM-venue info)	βw_{ij}^{venue}	--
Permutation #1 (preserving avg # visited venues per MSM and avg # visitors per venue)	βw_{ij}^{perm1}	• # co-visits of each venue • # co-visiting MSM between any two venues
Permutation #2 (preserving # co-visits of each venue, but randomizing # co-visiting MSM across pairs of venues)	βw_{ij}^{perm2}	• # co-visiting MSM between any two venues
Empirical venue network of co-visits from projection accounting for number of visits (MSM-visit-venue info)	$\beta w_{ij}^{venue,v}$	--

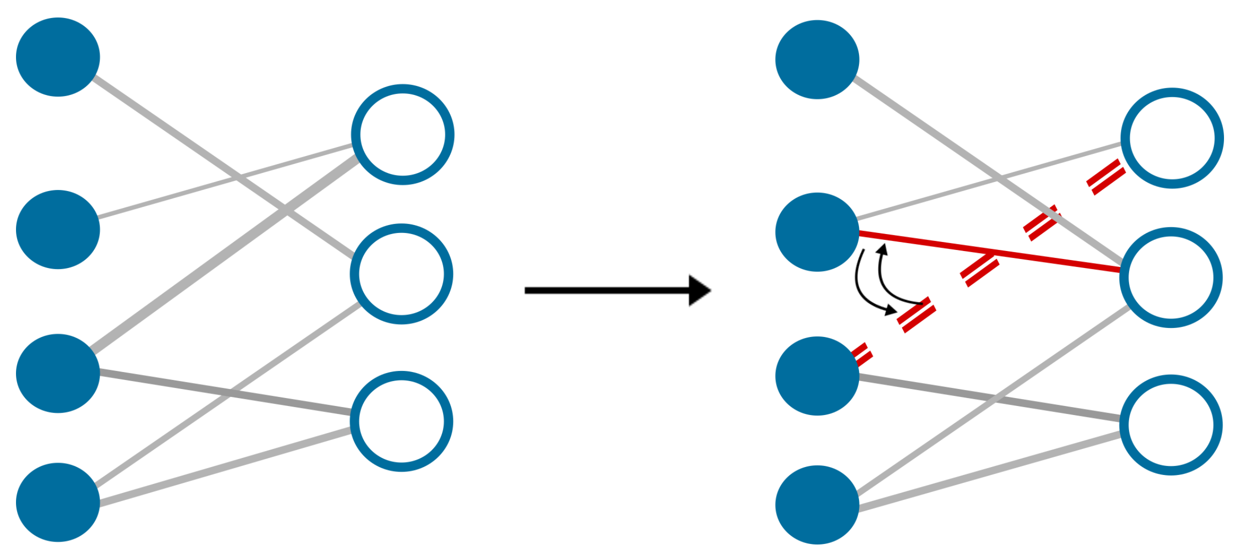
Maximum likelihood approach based on the observed times of invasion, MCMC to explore the parameters

Venue networks

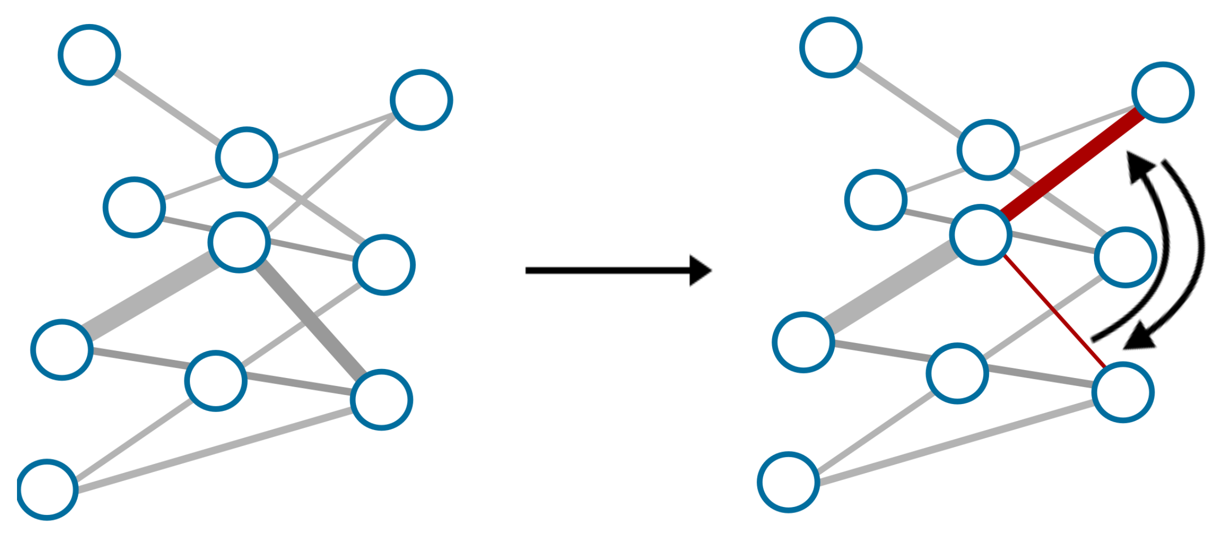
Permutations (null models)



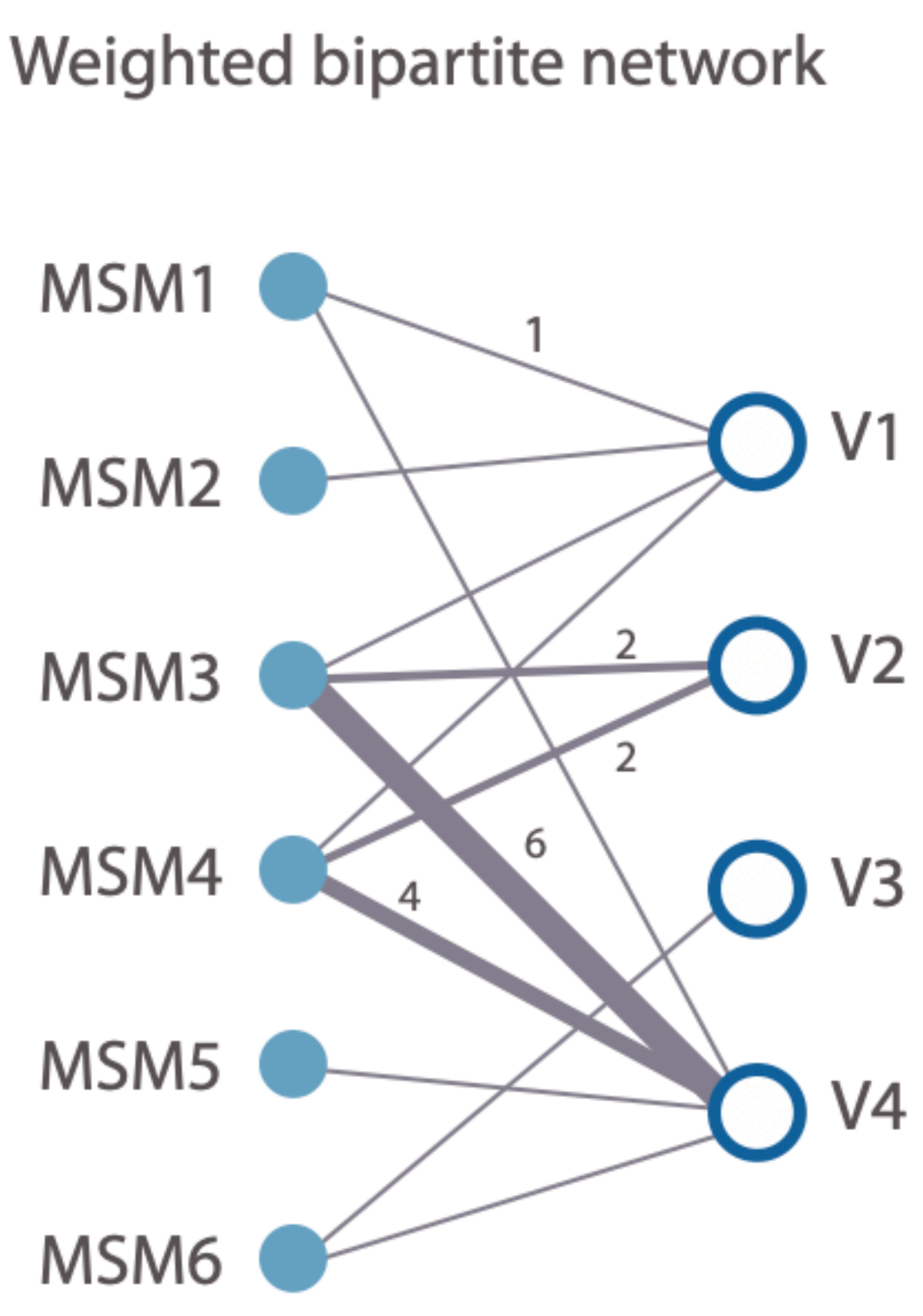
Permutation #1



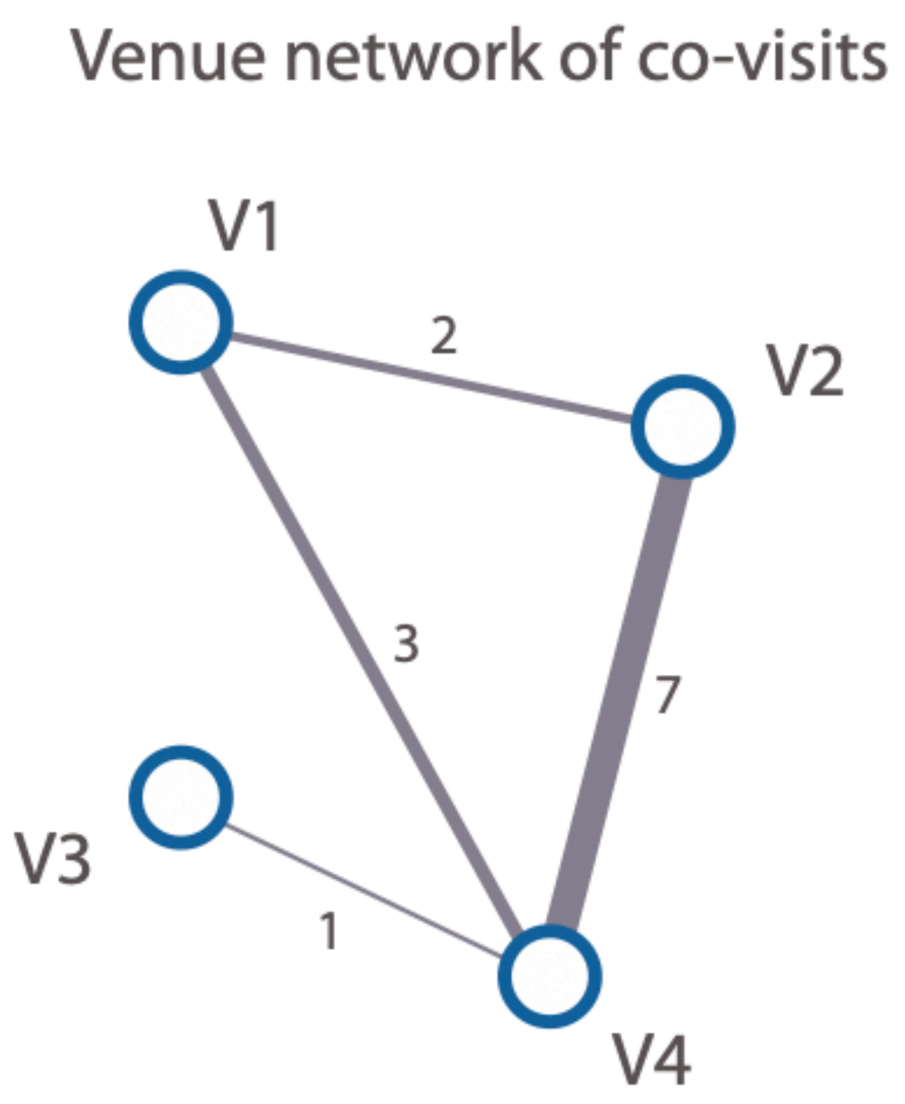
Permutation #2



Venue networks



projection accounting for weights (i.e. # visits to the venues)



link weight = average number of visits of the MSM attending both venues

MSM-visit-venue information

Venue networks

Models of risk of invasion	Risk of transmission from venue to venue	DIC	
		June 20	June 27
Empirical venue network of co-visits (MSM-venue info)	βw_{ij}^{venue}	62.36	81.1
Permutation #1	βw_{ij}^{perm1}	63.87	81.19
Permutation #2	βw_{ij}^{perm2}	63.18	81.65
Empirical venue network of co-visits (MSM-visit-venue info)	$\beta w_{ij}^{venue,v}$	62.63	82.04

- Venue network of co-visits (from standard projection of the bipartite network —> MSM-venue information) best describes the observed pattern of invasion among venues
- Statistically preferable wrt venue network of co-visits from projection accounting the #visits (MSM-visit-venue information) —> info on visits less important
- MSM co-visiting multiple venues are important (permutation #2 vs. empirical) maintains # total co-visits, but loses info on who visits what venue

- Visits are an information that does not provide any improvement to the accuracy of the model of reproducing the spread

Questions?