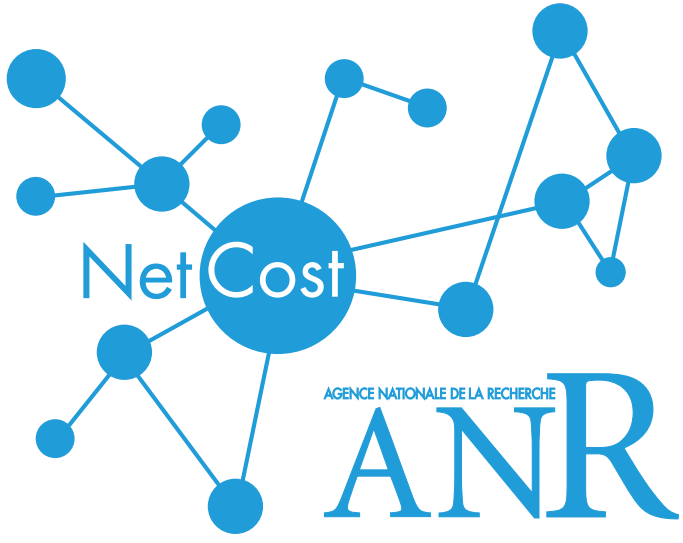
From individual spatio-temporal trajectories to spatial networks

Maxime Lenormand

From individual spatio-temporal trajectories to spatial networks



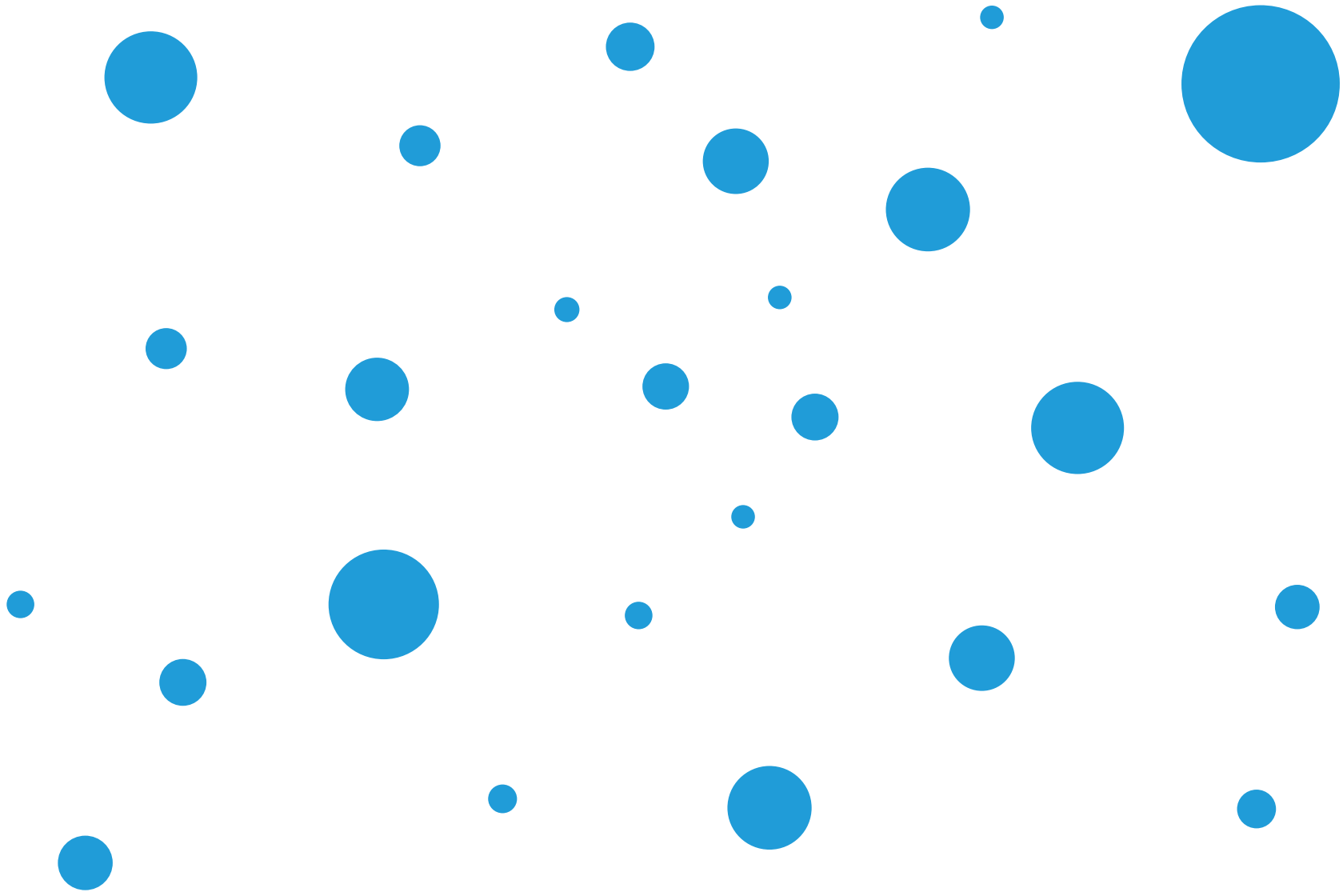
Maxime Lenormand
UMR TETIS, Irstea, France

XTerM2019 | Le Havre, France

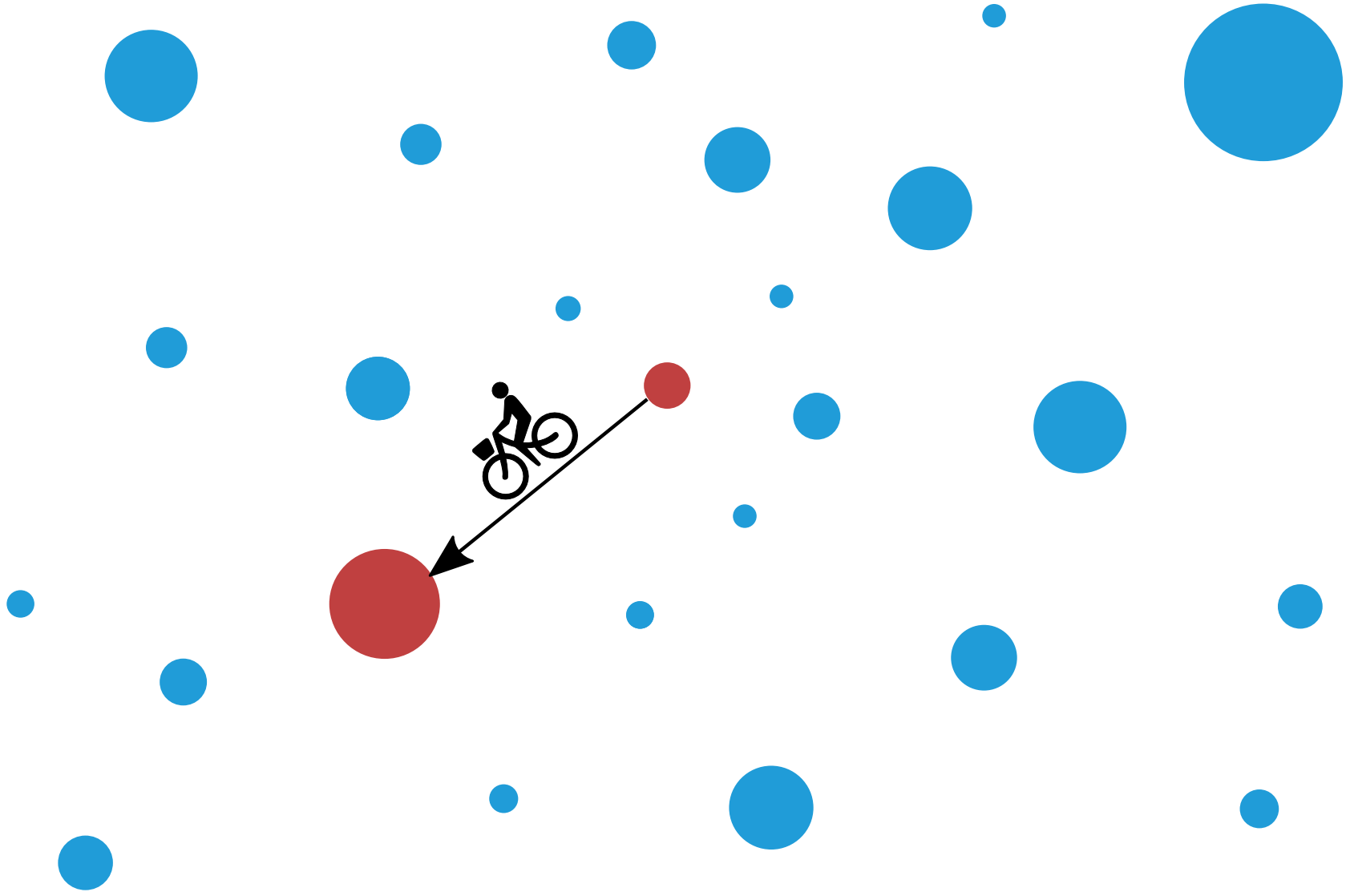
June 27, 2019

irstea

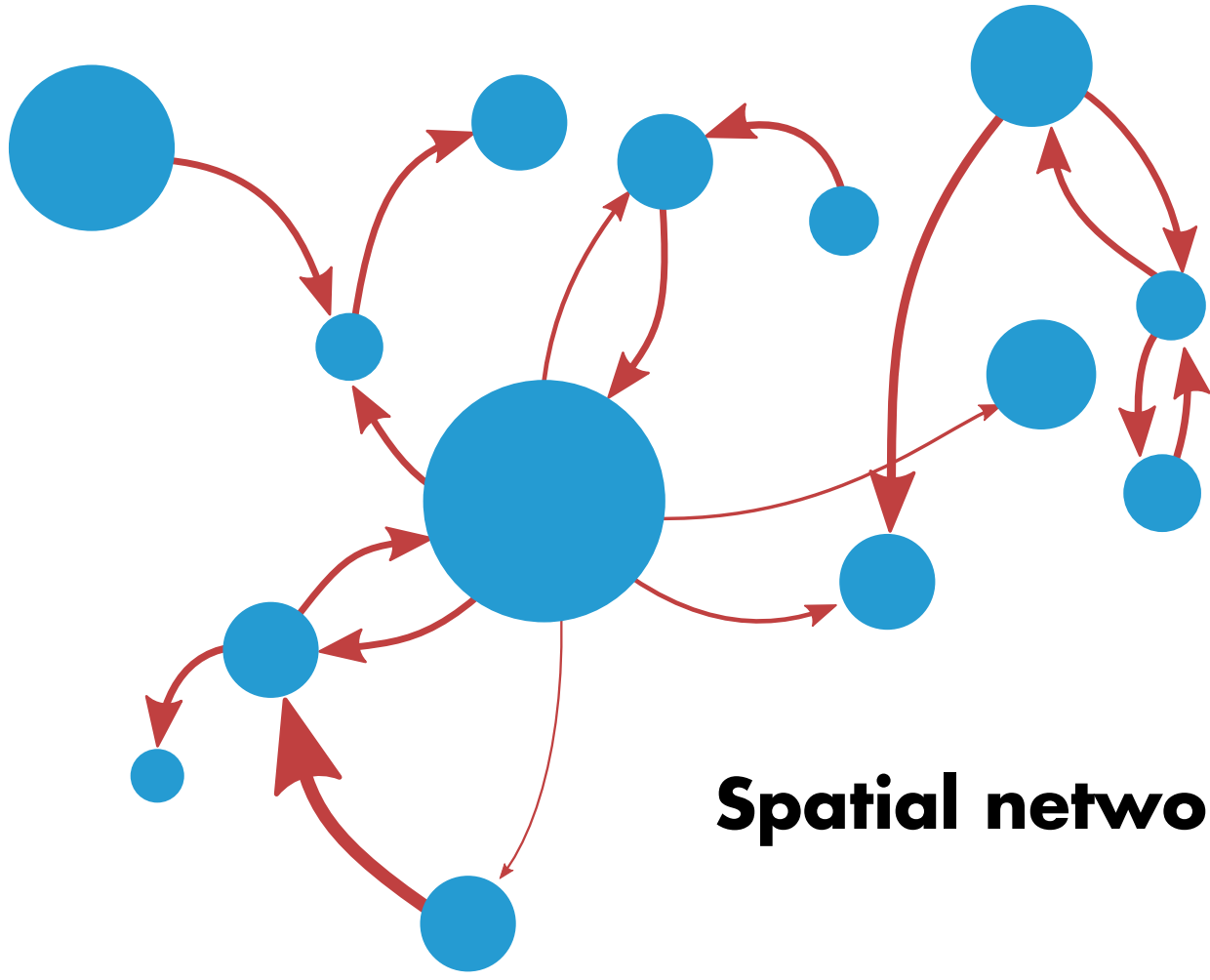
Motivation



Motivation



Motivation



Spatial network

Motivation

T_{ij}

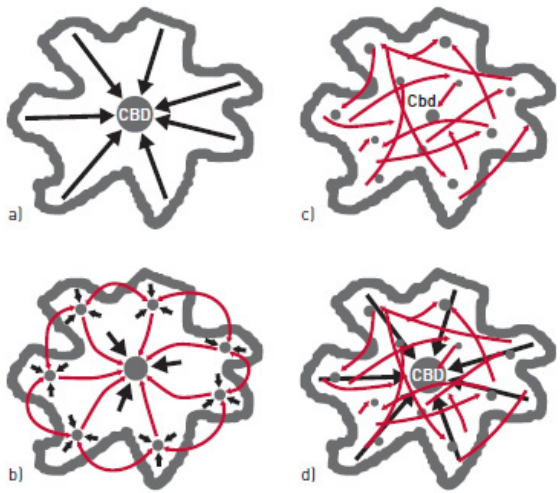
	1	2	3	4	5	
1	0	1	8	4	0	13
2	2	0	2	35	3	42
3	13	1	0	9	4	26
4	1	23	2	0	1	27
5	5	34	8	2	0	49
	20	59	20	50	8	

O_i

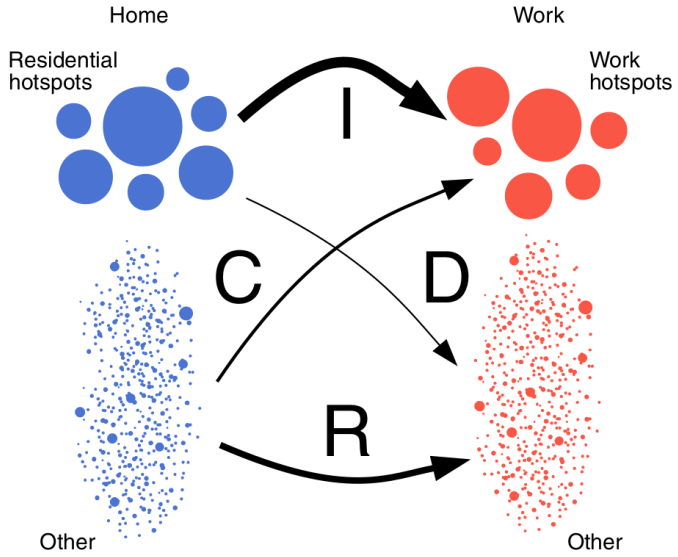
D_j

Origin-Destination matrix

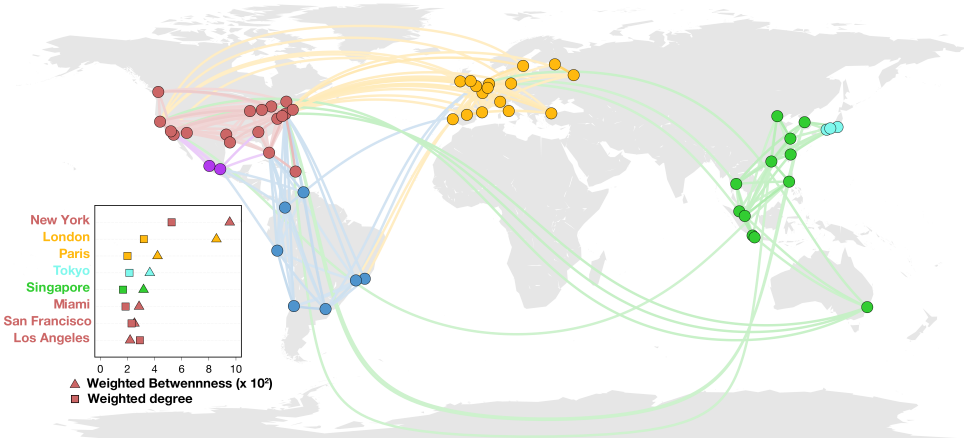
Applications



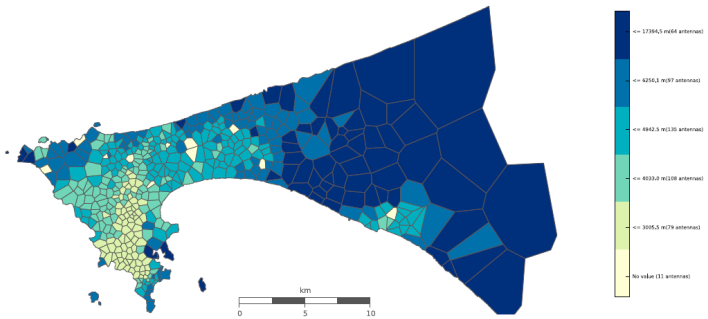
Bertaud & Malpezzi (2003)



Louail *et al.* (2015)

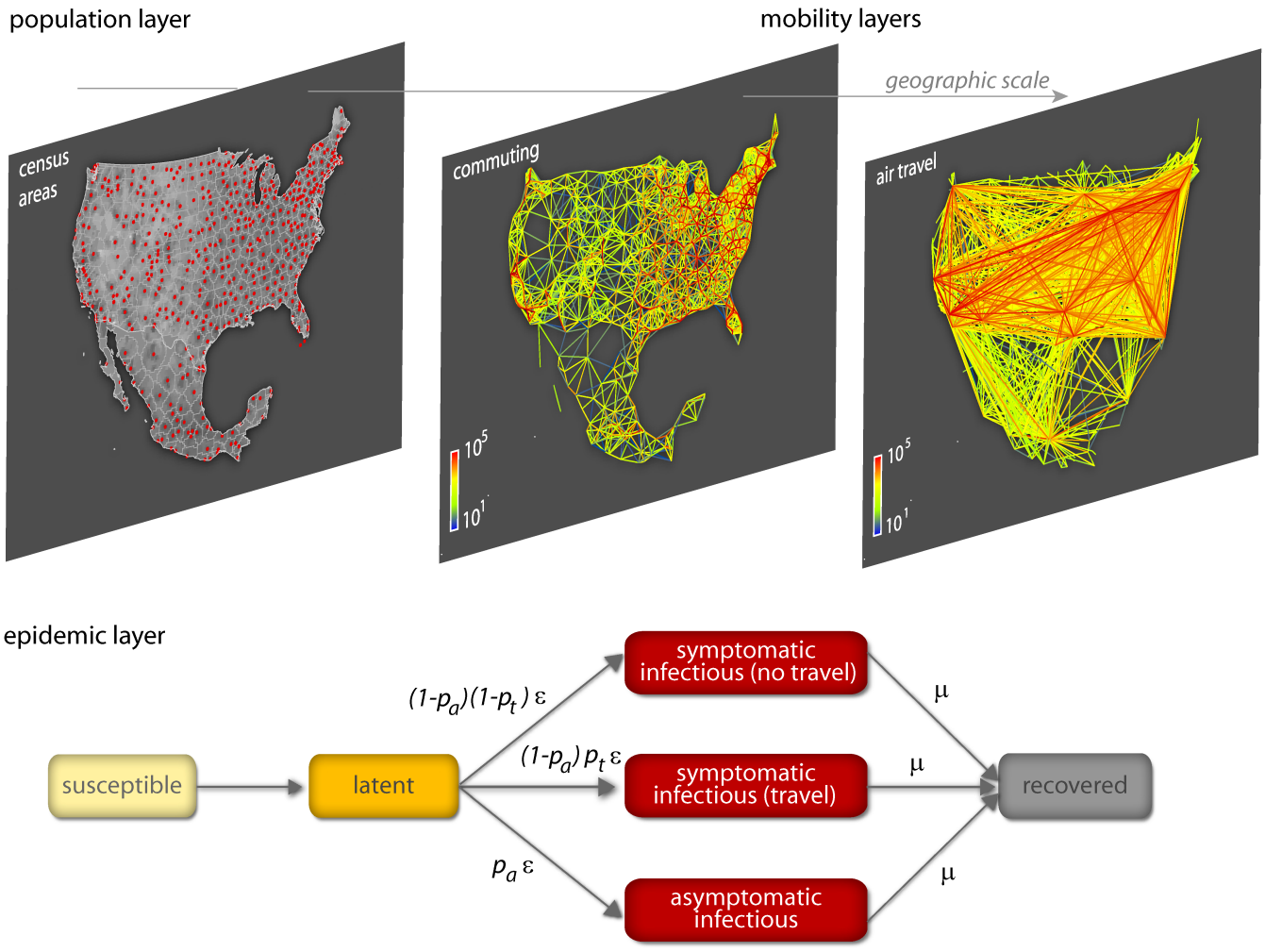


Lenormand *et al.* (2015)



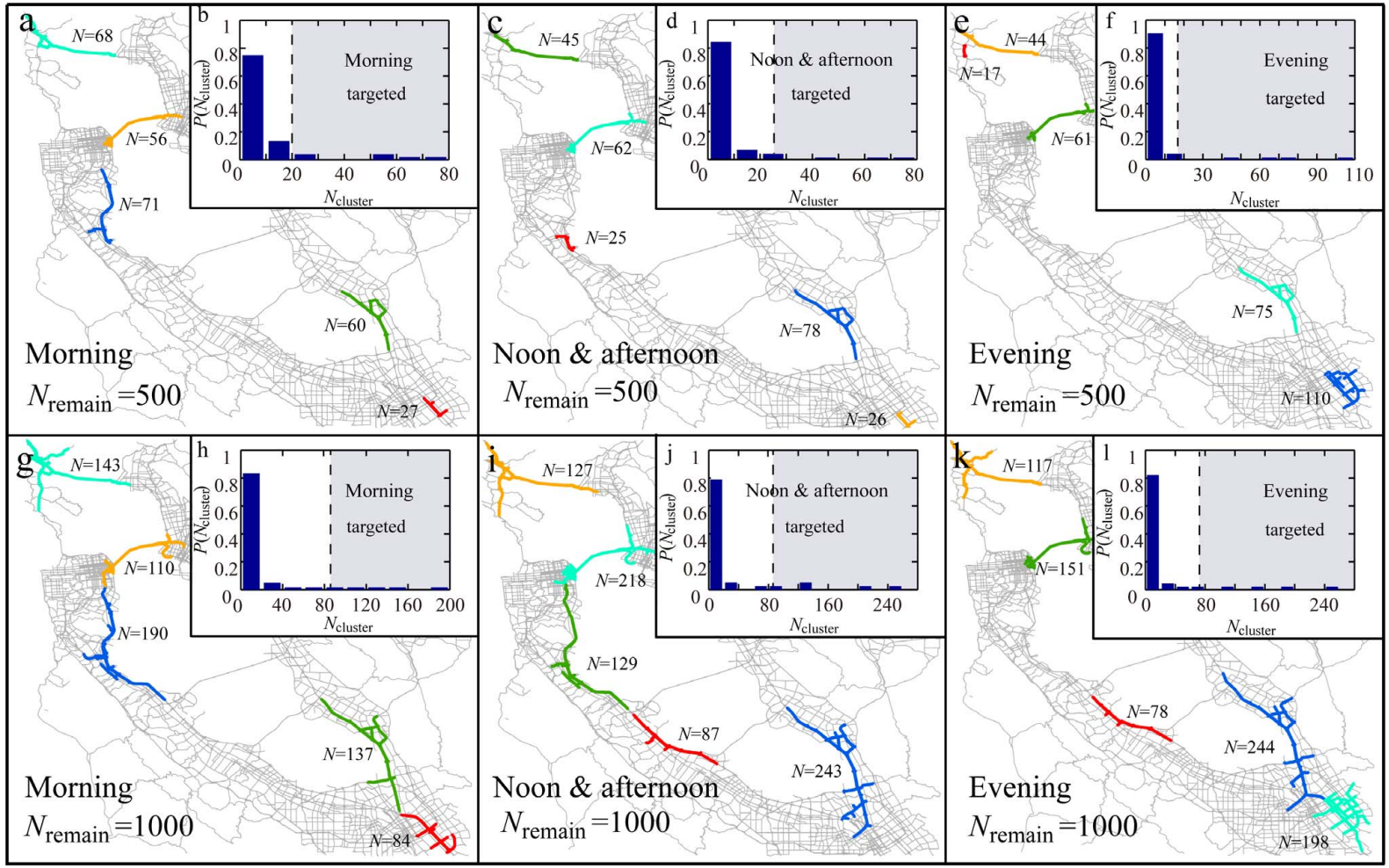
Louf et al. (2015)

Applications



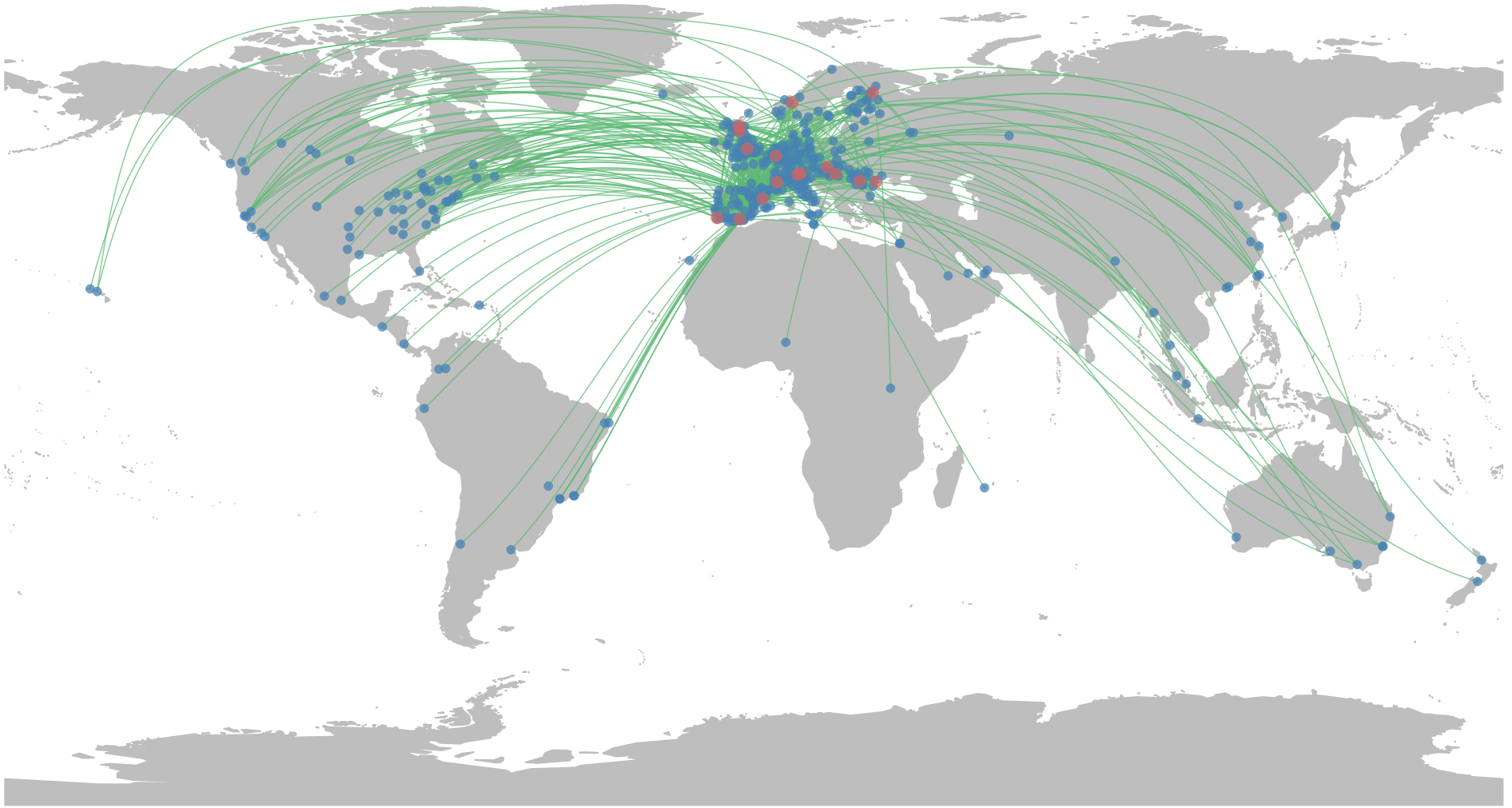
Balcan *et al.* (2009) Seasonal transmission potential and activity peaks of the new influenza. *BMC Medecine 7*, 15052015.

Applications



Wang *et al.* (2014) Encapsulating urban traffic rhythms into road networks. *Scientific Reports* 4, 4141.

Applications



Lenormand *et al.* (2018) Multiscale socio-ecological networks in the age of information. *PLoS ONE 13, e0206672.*

How to estimate these flows?

- Census & survey

How to estimate these flows?

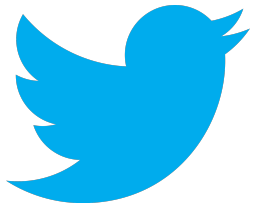
- Census & survey
- Spatial interaction models

How to estimate these flows?

- Census & survey
- Spatial interaction models
- Individual geolocalized data

How to estimate these flows?

- Census & survey
- Spatial interaction models
- **Individual geolocalized data**



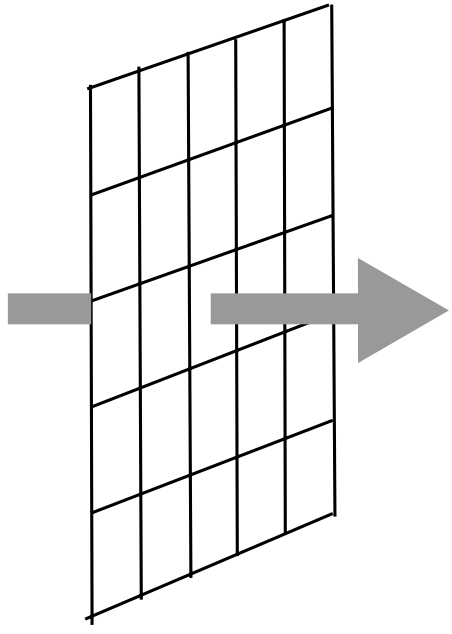
flickr

BBVA

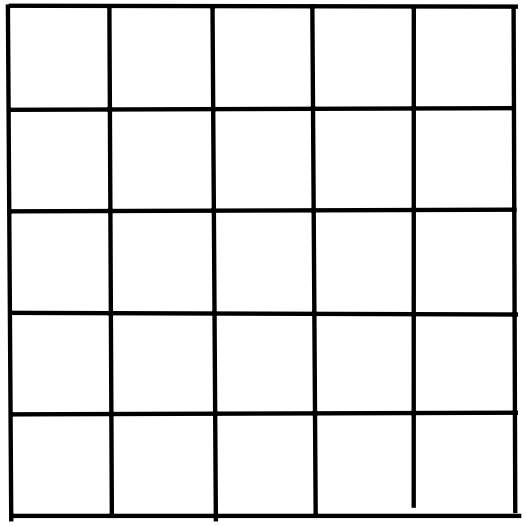
Telefónica

Sampling framework

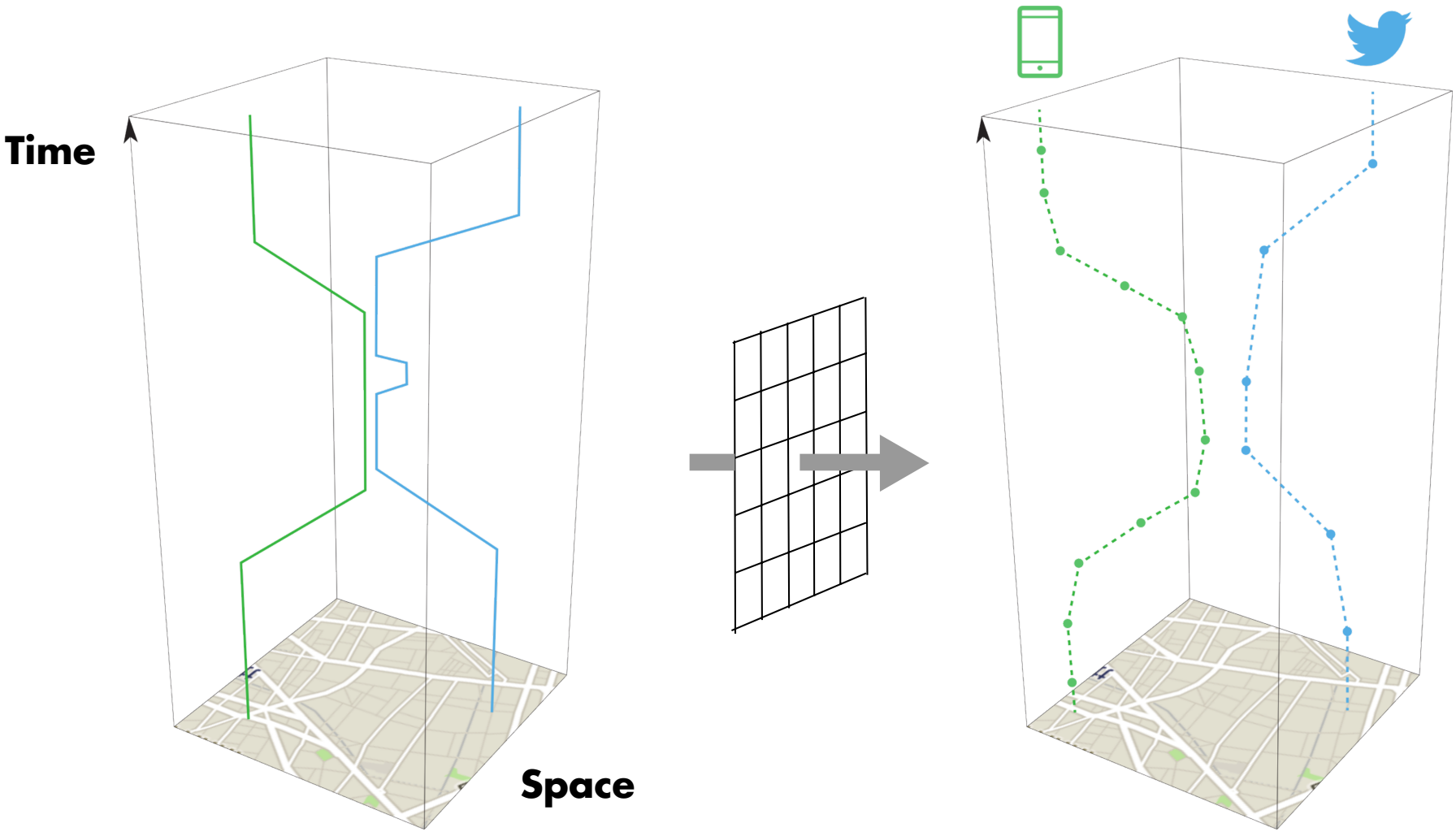
"Reality"



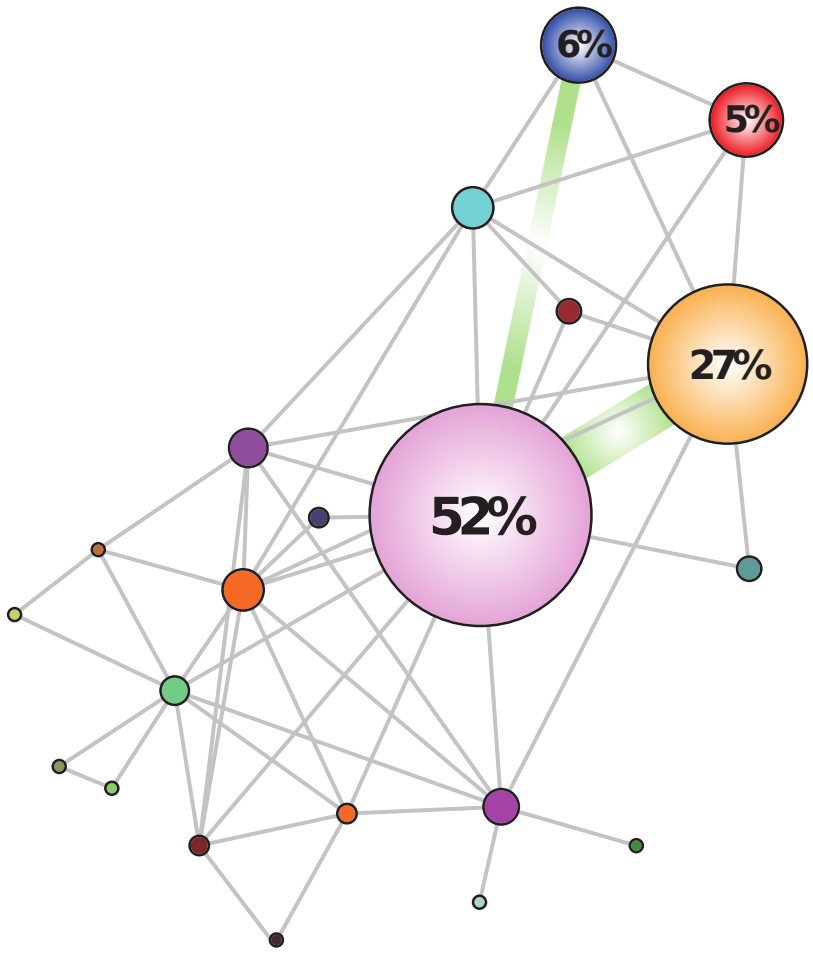
Data



Sampling framework



Most frequented locations



Home

Most frequented location between 7pm and 7am

Work

Most frequented location between 8am and 5pm on weekdays

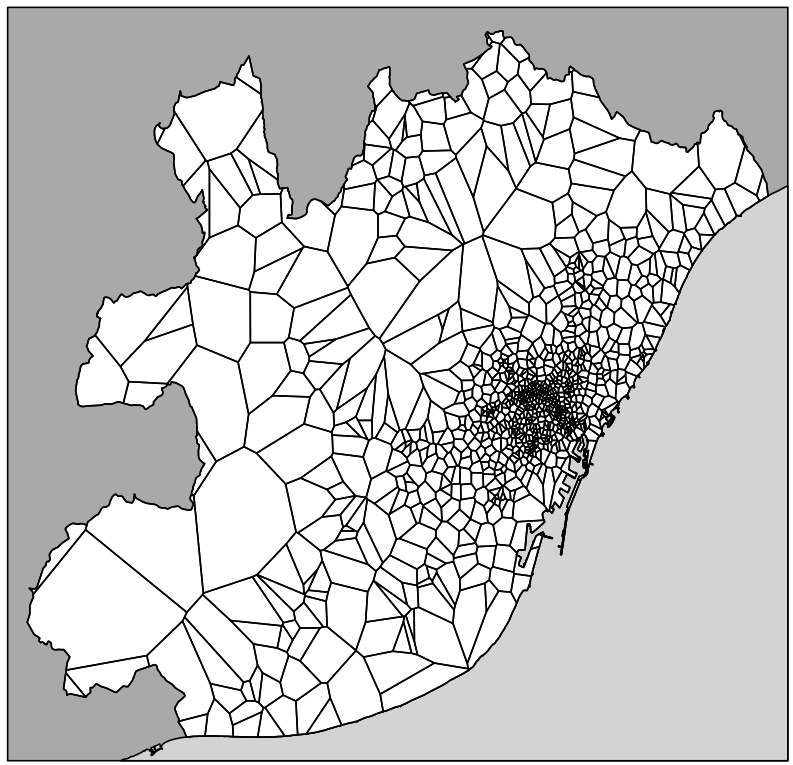
↓

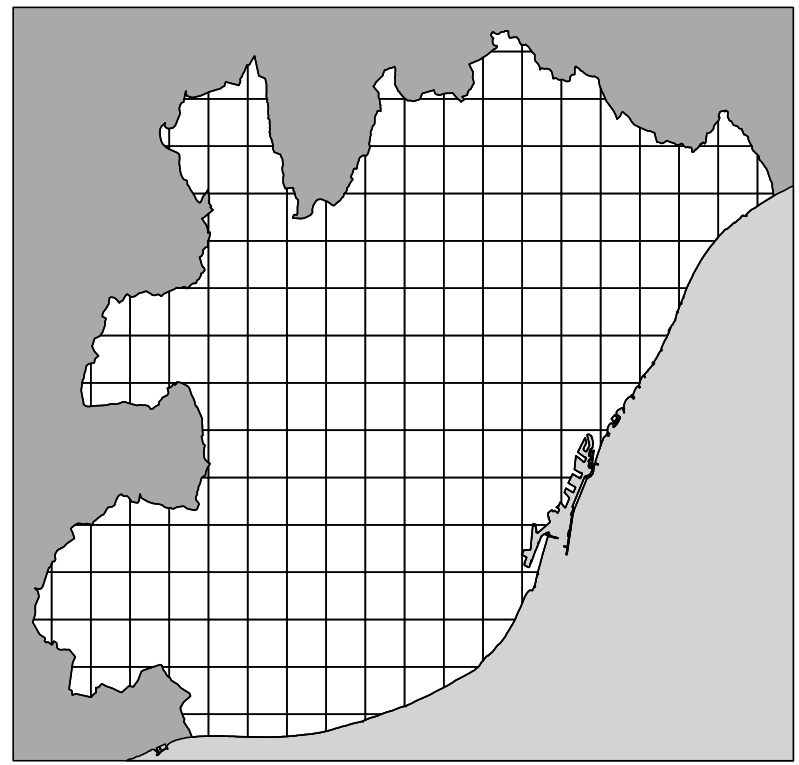
Origin-Destination Matrix

Tij: *number of individuals living in cell* ***i*** *and working in cell* ***j***

Song *et al.* (2010) Limits of predictability in human mobility. *Science* 327, 1018-1021.

Spatial discretization





Location?

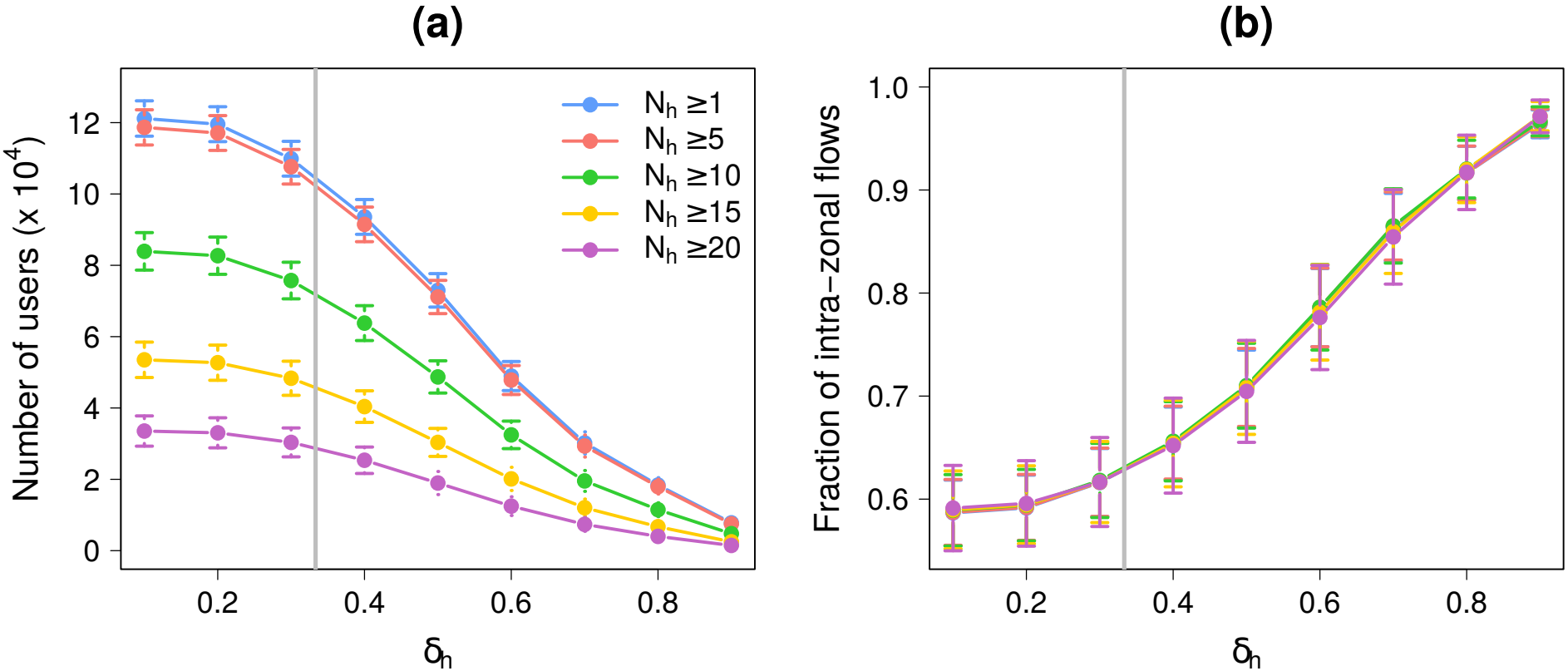
Extracting most frequented locations

- The hours of activity are divided into two groups, daytime hours and nighttime hours. Only days of the week from Monday to Thursday are taken into account.
- **First filter:** Consider only individuals "actives" during at least $\mathbf{N_h}$ hours spread over at least $\mathbf{N_d}$ days.
- For each hour of activity, the most frequently visited zone during this hour is identified.
- For both groups of hours (daytime and nighttime), we identify the zone in which the user has been localized the highest number of hours.
- **Second filter:** Select only users whose fraction of hours spent at "home" and "work" are larger than a fraction δ_h of the total number of locations visited during nighttime and daytime, respectively.

Lenormand et al. (2016) Is spatial information in ICT data reliable? In proceedings of the 2016 Spatial Accuracy Conference, 9-17, Montpellier, France.

https://gitlab.com/maximelenormand/Most-frequented-locations

Extracting most frequented locations



Lenormand et al. (2016) Is spatial information in ICT data reliable? In proceedings of the 2016 Spatial Accuracy Conference, 9-17, Montpellier, France.

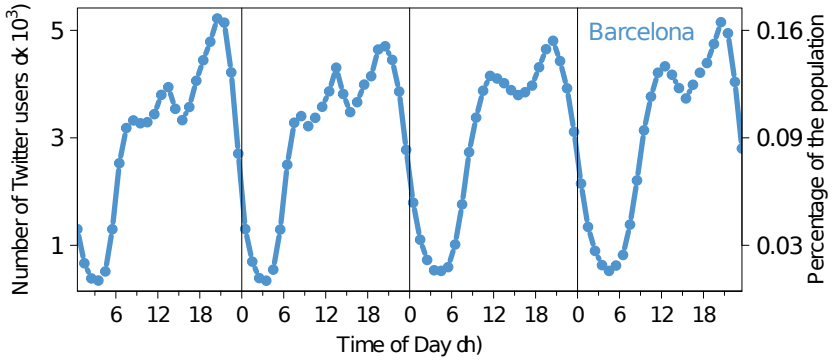
Cross-checking different sources of mobility information

Lenormand *et al.* 2014

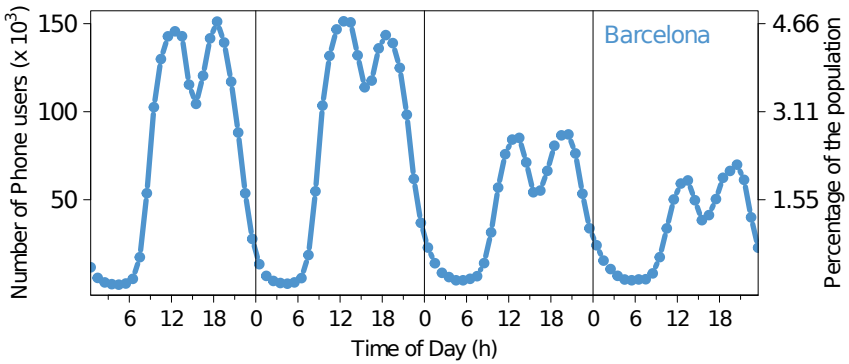
Data

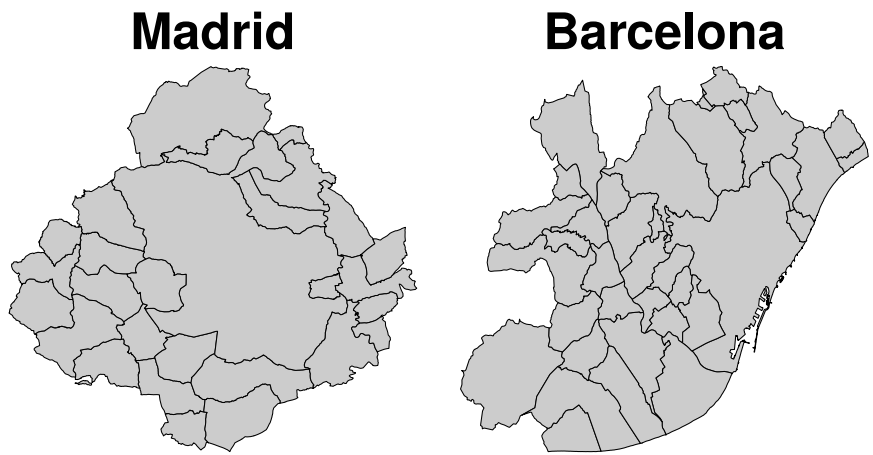
BBVA

INE Instituto Nacional de Estadística



Telefónica

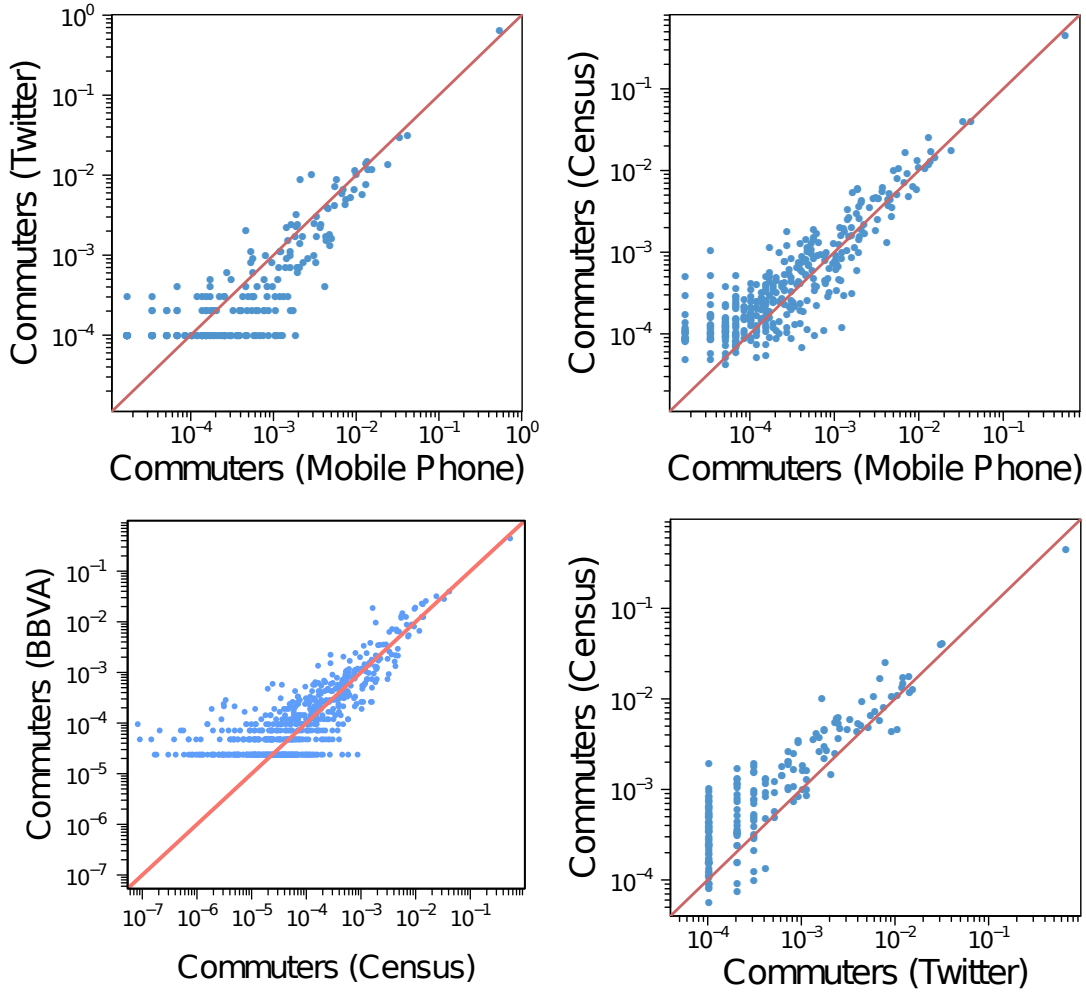




Lenormand *et al.* (2014) Cross-checking different sources of mobility information.*PlosOne*,9(8):e105407.

Louail *et al.* (2017) Crowdsourcing the Robin Hood effect in cities. Applied Network Science 2, 11.

Pairwise OD comparison



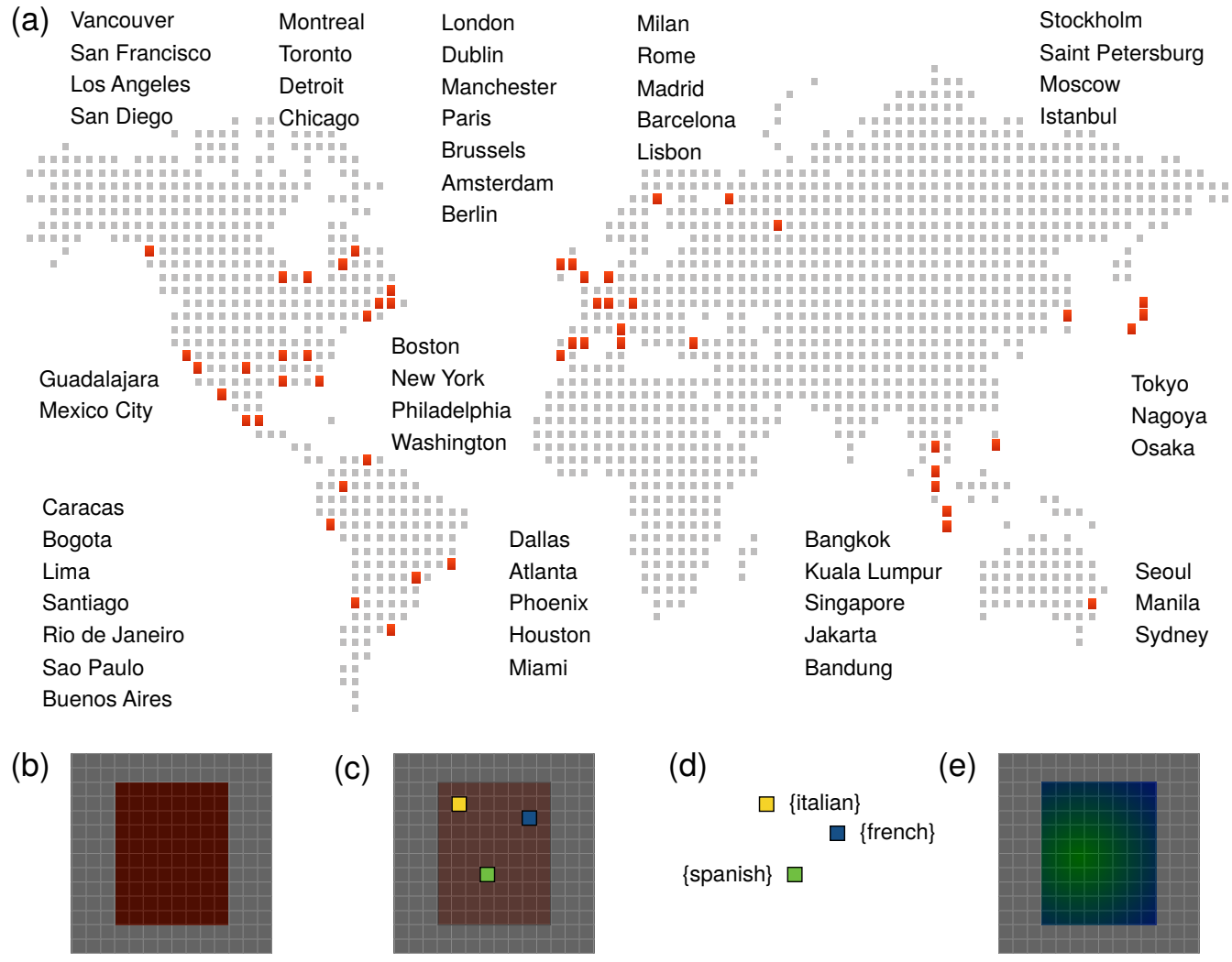
Lenormand *et al.* (2014) Cross-checking different sources of mobility information.*PlosOne*,9(8):e105407.

Louail *et al.* (2017) Crowdsourcing the Robin Hood effect in cities. Applied Network Science 2, 11.

Immigrant community integration in world cities

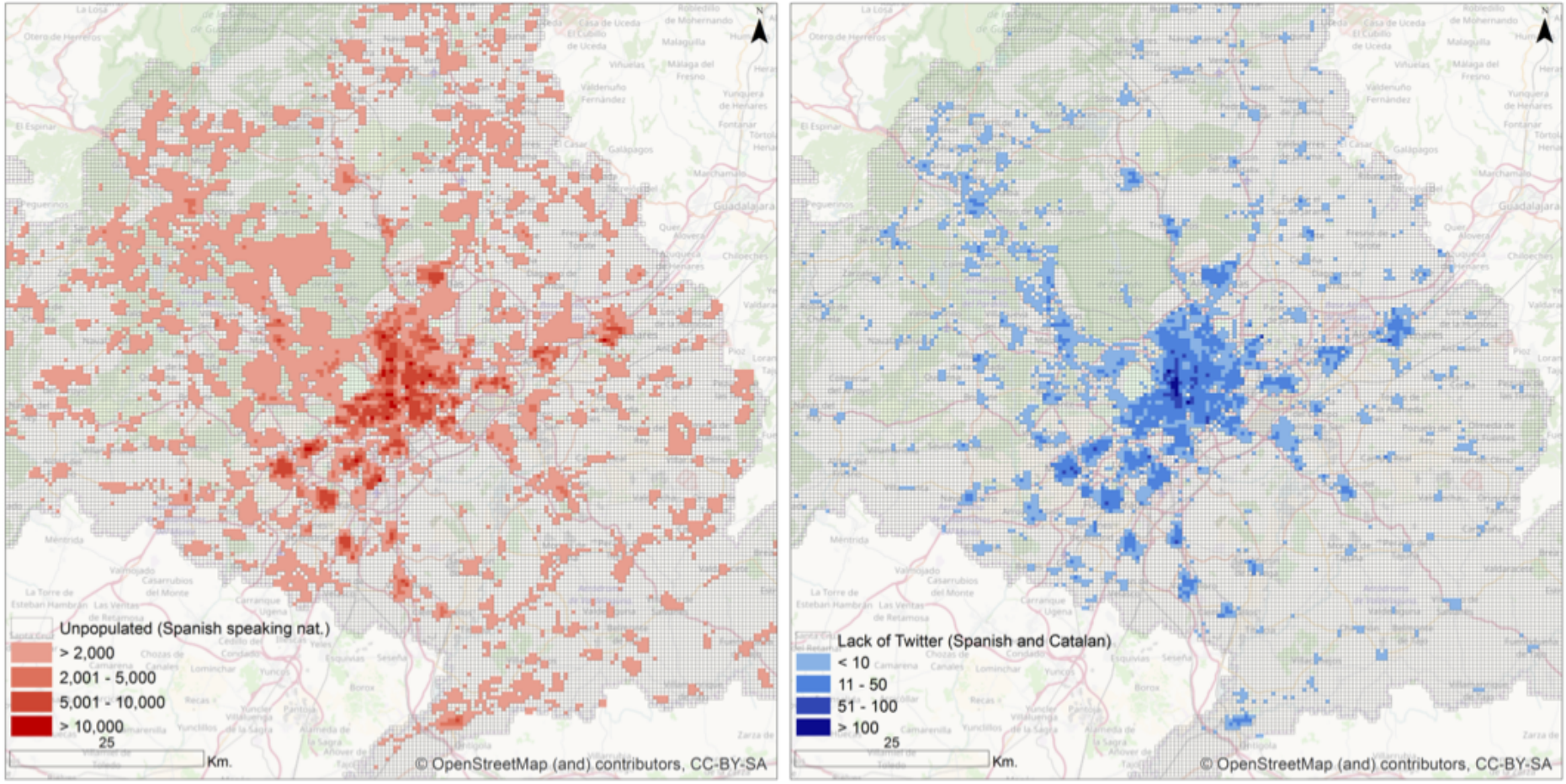
Lamanna *et al.* 2018

Data



Lamanna *et al.* (2018) Immigrant community integration in world cities. *Plos One 13, e0191612.*

Spatial distribution of residence

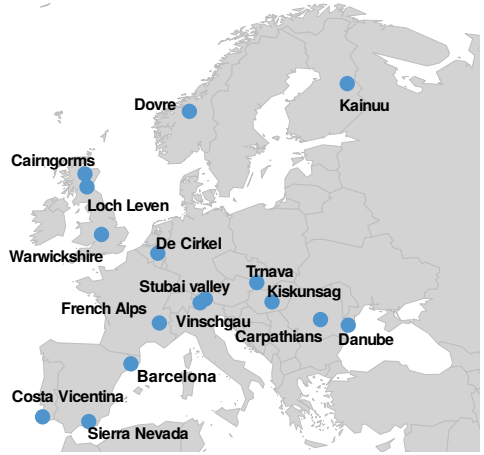


Lamanna *et al.* (2018) Immigrant community integration in world cities. *Plos One 13, e0191612.*

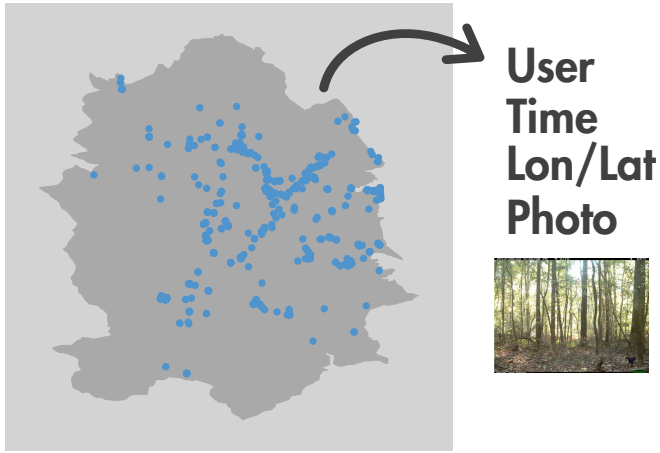
Multiscale socio-ecological networks in the age of information

Lenormand *et al.* 2018

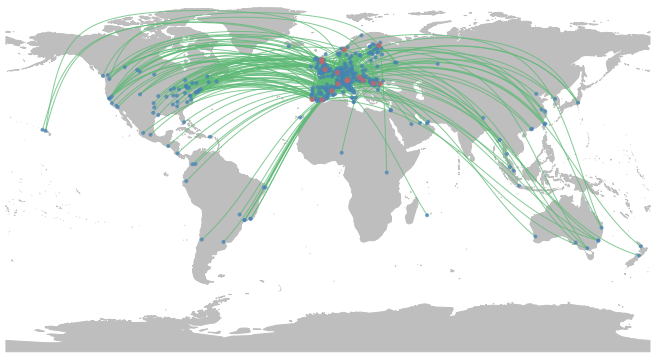
Data

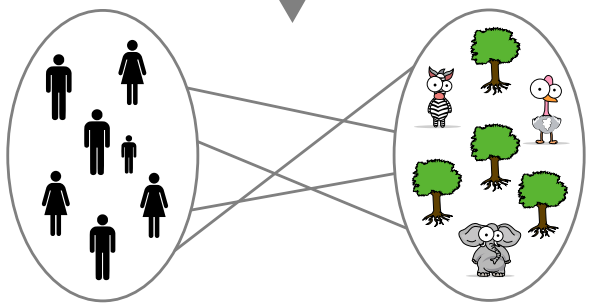


16 case study sites



flickr

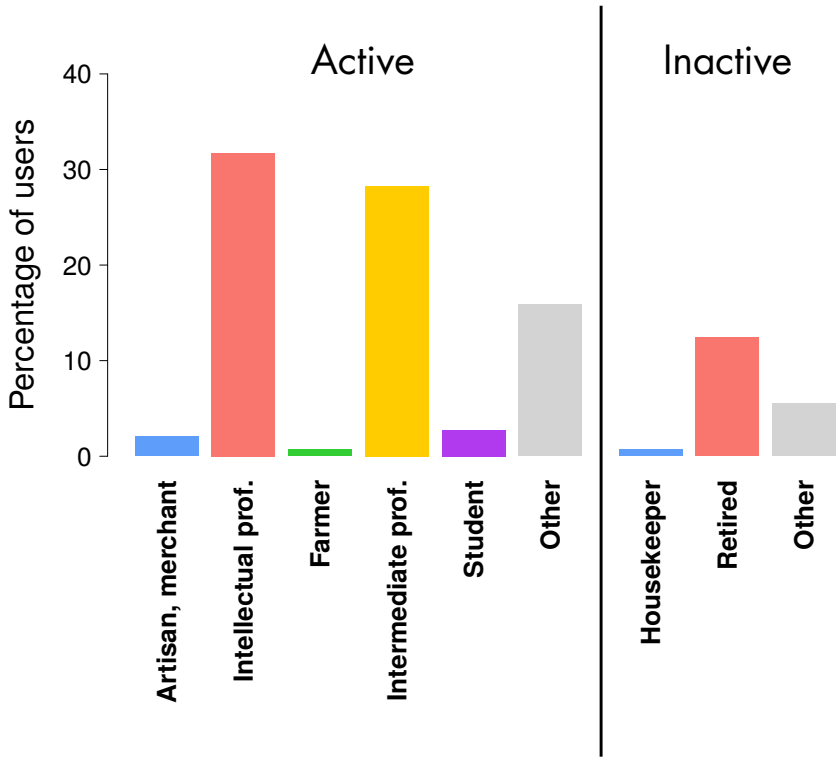


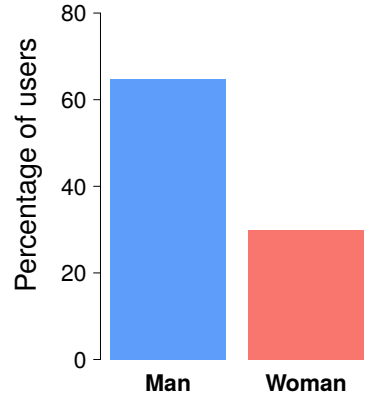


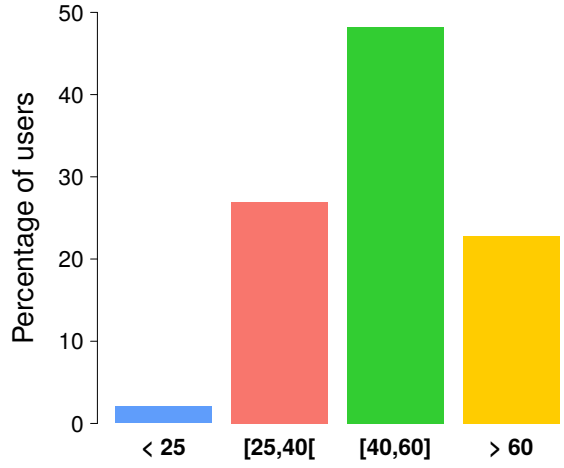
Lenormand *et al.* (2018) Multiscale socio-ecological networks in the age of information. *PLoS ONE 13, e0206672.*

Survey

11% response rate







90% of accuracy in the users' place of residence detection!

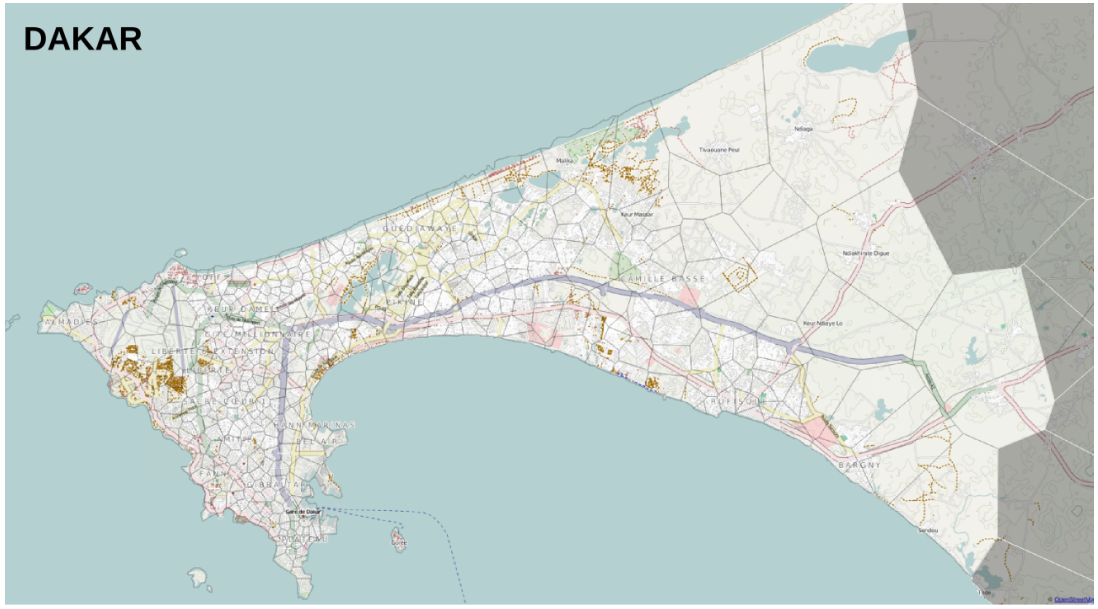
Lenormand *et al.* (2018) Multiscale socio-ecological networks in the age of information. *PLoS ONE 13, e0206672.*

Is spatial information in ICT data reliable?

Lenormand *et al.* 2016

Data

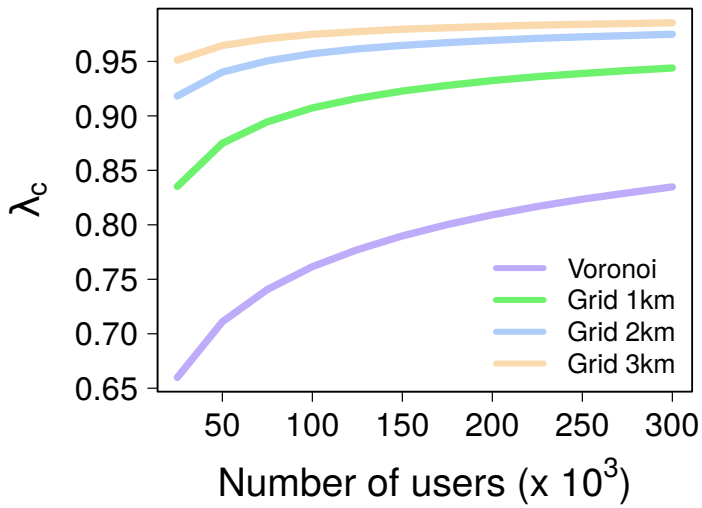
300,000 mobile phone users' trajectories **x 25** two-week periods

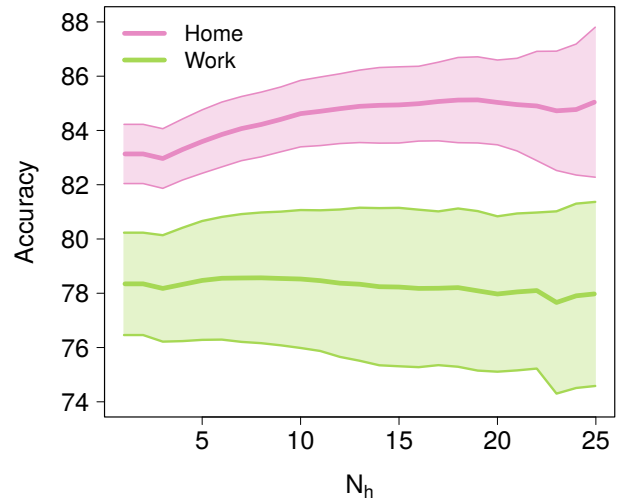


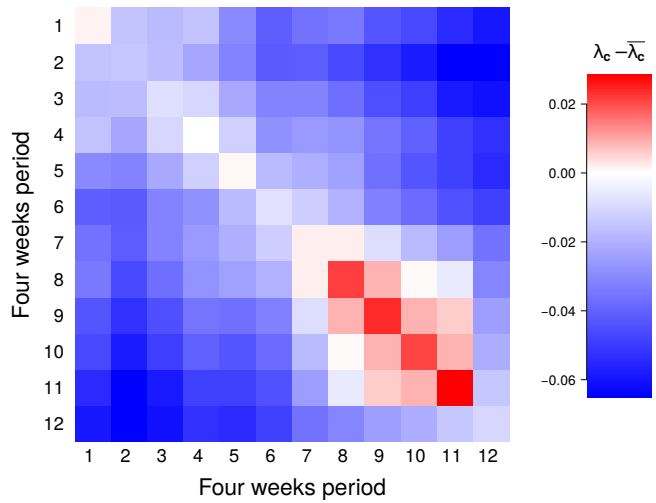
Identifying **home-work locations**
from mobile phone activity

Lenormand *et al.* (2016) Is spatial information in ICT data reliable? In proceedings of the 2016 Spatial Accuracy Conference, 9-17, Montpellier, France.

OD comparison







Lenormand *et al.* (2016) Is spatial information in ICT data reliable? In proceedings of the 2016 Spatial Accuracy Conference, 9-17, Montpellier, France.

Take home messages...

- Good agreement between the different data sources.
- Uncertainty & accuracy are highly dependent of the spatial resolution and sample size.
- More studies in this spirit need to be done to assess the biases and uncertainty associated with ICT data.
- It could be interesting to involve (more strongly and widely) the indiviual ICT data providers.

References

- **Caceres *et al.*** (2007) Deriving origin-destination data from a mobile phone network. Intell. Transport Syst.I ET1 1, 15–26.
- **Iqbal *et al.*** (2014) Development of origin-destination matrices using mobile phone call data, Transp. Res. Part C:EmergingTechnol. 40, 63–74.
- **Tizzoni *et al.*** (2014) On the use of human mobility proxies for modeling epidemics. PLoS Comput. Biol. 10, e1003716.
- **Toole *et al.*** (2015) The path most traveled: Travel demand estimation using big data resources Transp. Res. PartC: Emerging Technol. 58, PartB,162–177.
- **Lenormand *et al.*** (2016) Is spatial information in ICT data reliable? In proceedings of the 2016 Spatial Accuracy Conference, 9-17, Montpellier, France.
- **Barbosa-Filho *et al.*** (2018) Human Mobility: Models and Applications. Physics Reports 734, 1-74.

www.maximelenormand.com

Acknowledgement



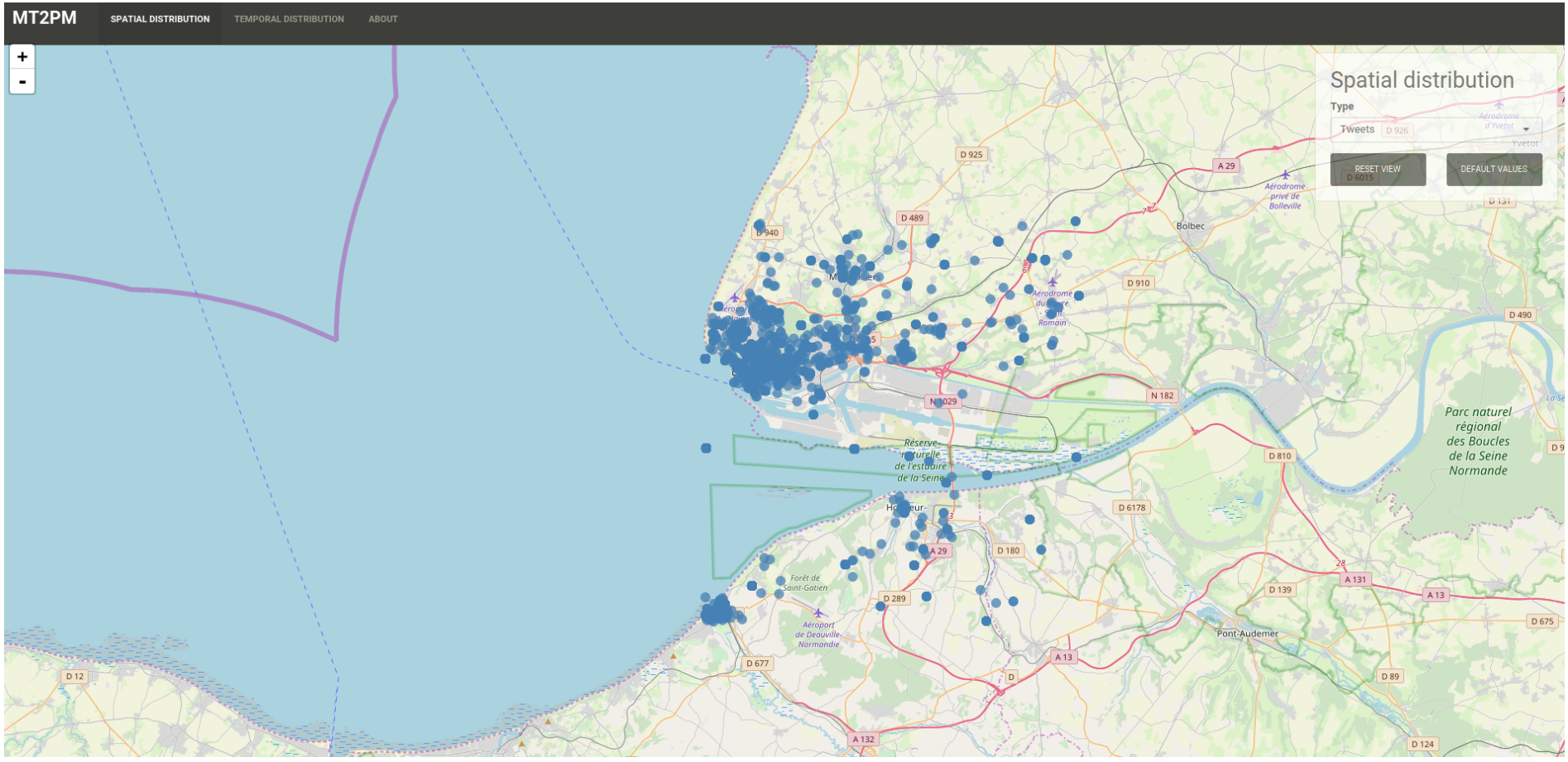


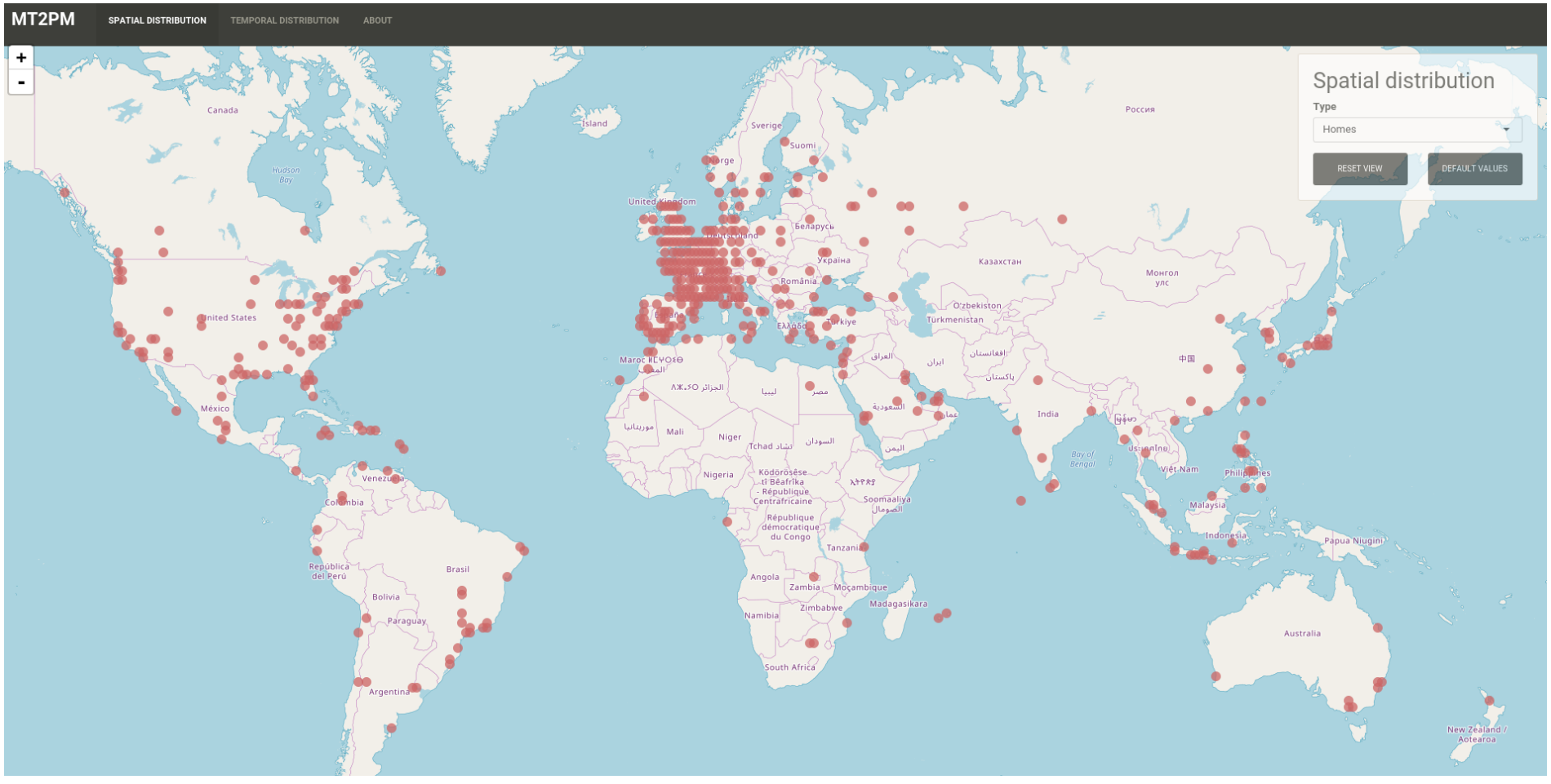









MT2PM
SPATIAL DISTRIBUTION
TEMPORAL DISTRIBUTION
ABOUT
+
-
Spatial distribution
Type
Tweets
RESET VIEW
DEFAULT VALUES
D 925
A 29
D 489
D 940
D 910
D 490
N 182
D 810
D 6178
D 180
A 131
A 13
D 139
D 675
D 289
D 677
D 12
A 132
D 89
D 124
Bolbec
Pont-Audemer
Parc naturel régional des Boucles de la Seine Normande
Réserve naturelle de l'estuaire de la Seine
Forêt de Saint-Gatien
Aéroport de Deauville-Normandie
Aérodrome privé de Bolleville
Aérodrome d'Yvetot


MT2PM
SPATIAL DISTRIBUTION
TEMPORAL DISTRIBUTION
ABOUT
+
-
Spatial distribution
Type
Homes
RESET VIEW
DEFAULT VALUES
Canada
Island
Sverige
Suomi
Россия
United Kingdom
Беларусь
Україна
România
Казахстан
Монгол улс
O'zbekiston
Türkmenistan
Türkiye
Ελλάδα
Maroc
الجزائر
ليبيا
مصر
العراق
ایران
افغانستان
پاکستان
中国
India
United States
México
Mali
Niger
Nigeria
Tchad
السودان
اليمن
Soomaaliya
الصومال
Venezuela
Colombia
República del Perú
Brasil
Bolivia
Paraguay
Argentina
Angola
Zambia
Moçambique
Zimbabwe
Namibia
South Africa
Madagasikara
Tanzania
Indonesia
Malaysia
Việt Nam
Philippines
Papua Niugini
Australia
New Zealand / Aotearoa