Some practical problems about planar graphs: time evolution and typology

Marc Barthelemy

CEA, Institut de Physique Théorique, Saclay, France

EHESS, Centre d'Analyse et de Mathématiques sociales, Paris, France

marc.barthelemy@cea.fr
www.quanturb.com

20th Itzykson conference 2015

Importance of planar graphs

- Planar graphs are present in many practical applications
  - Transportation networks (roads, subway, …)
  - Infrastructure networks (power grids, water distribution, …)
  - Biological networks (veination pattern, …)
- Need for 'global' characterization of graphs
  - Characterize the structure
  - Compare different graphs

Note: All planar graphs are here embedded in the 2d plane and are maps (literally)

Outline

- Introduction
- I. Time evolution of planar graphs
- II. Typology of planar graphs
- Discussion and perspectives

I. Time evolution: Characterization ?

An old problem in quantitative geography

- Kansky (63-69)
Evolution
of the Sicilian
railroad network

- Morrill (1965)
Railway network
Growth

- New data sources:
Digitization of
old maps

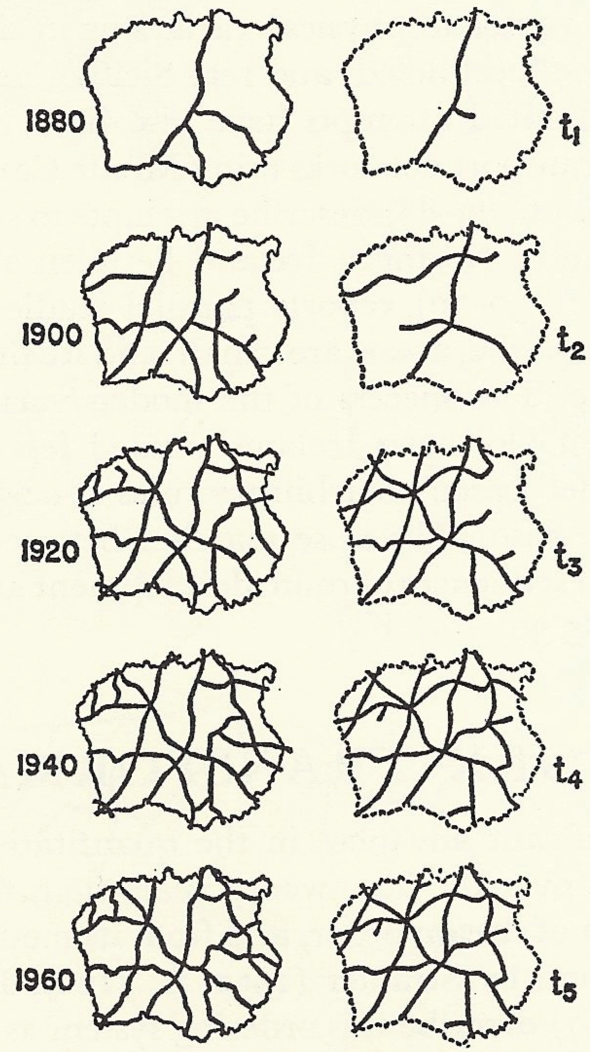
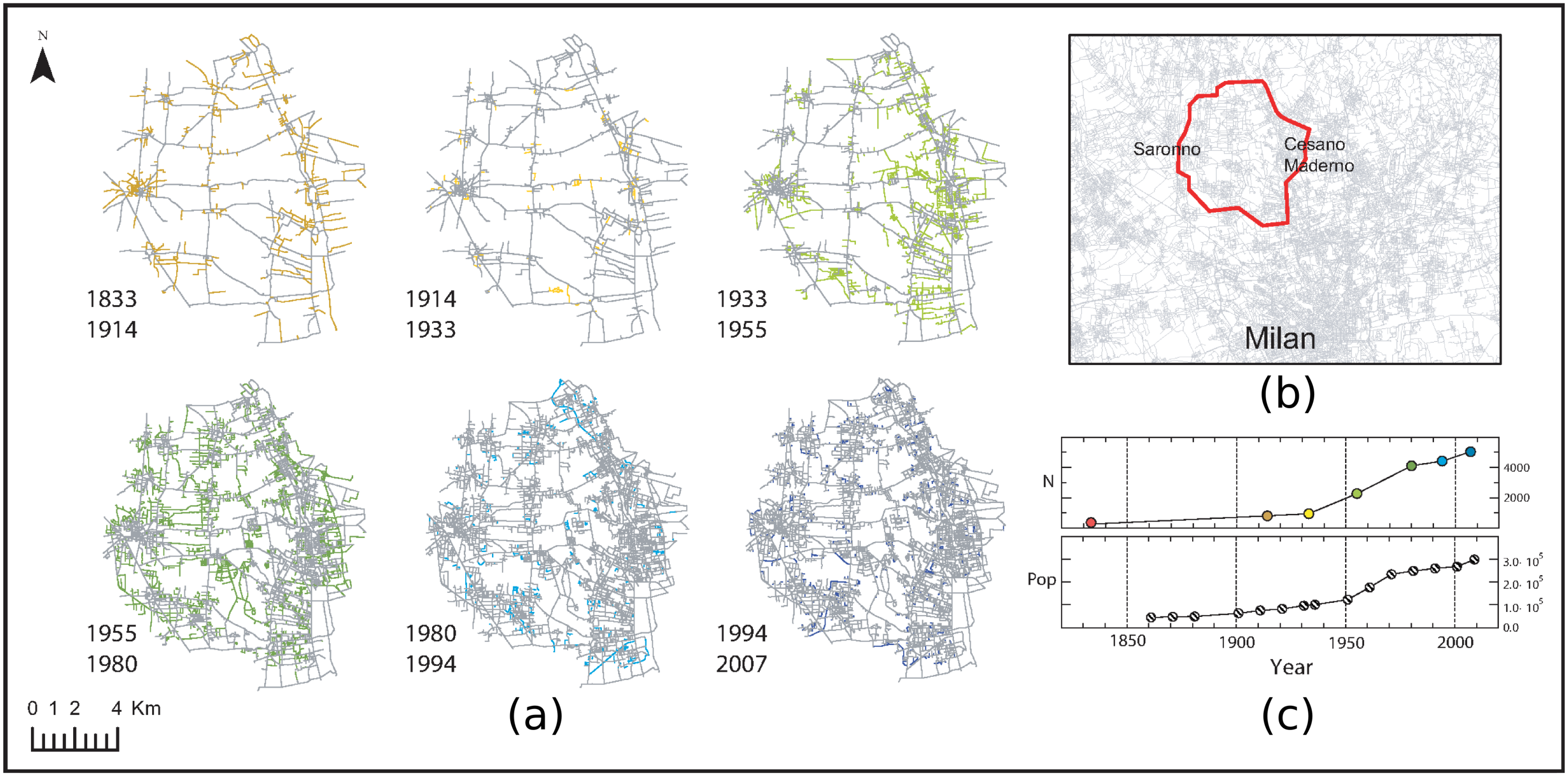


Fig. 5.30. Comparison of the (*A*) actual growth of railway network in central Sweden with (*B*) the simulated growth. Source: Morrill, 1965, pp. 130–70.

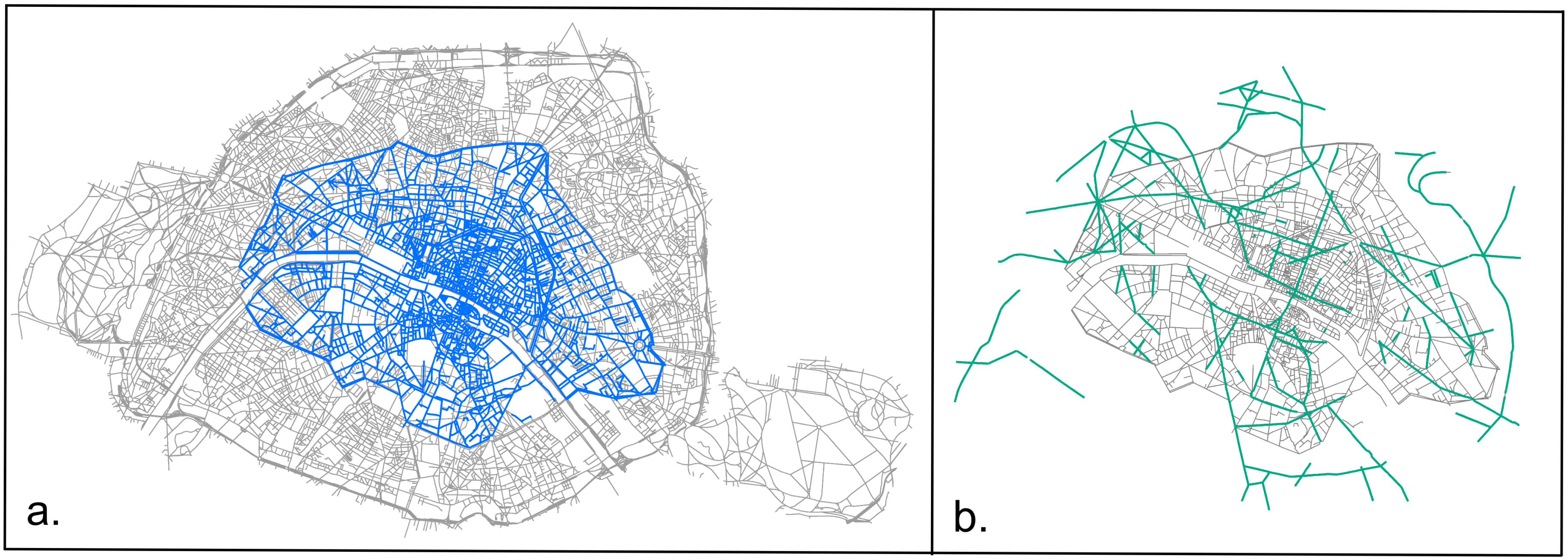
Road network evolution
Groane region, Italy 1833-2007



Strano, Nicosia, Latora, Porta, MB, Nature Scientific Reports (2012)

Road network evolution: Importance of central planning

- Evolution of the Paris street network 1789-2010
(1789, 1826, 1836, 1888, 1999, 2010-soon 1591, 1652, 1728)

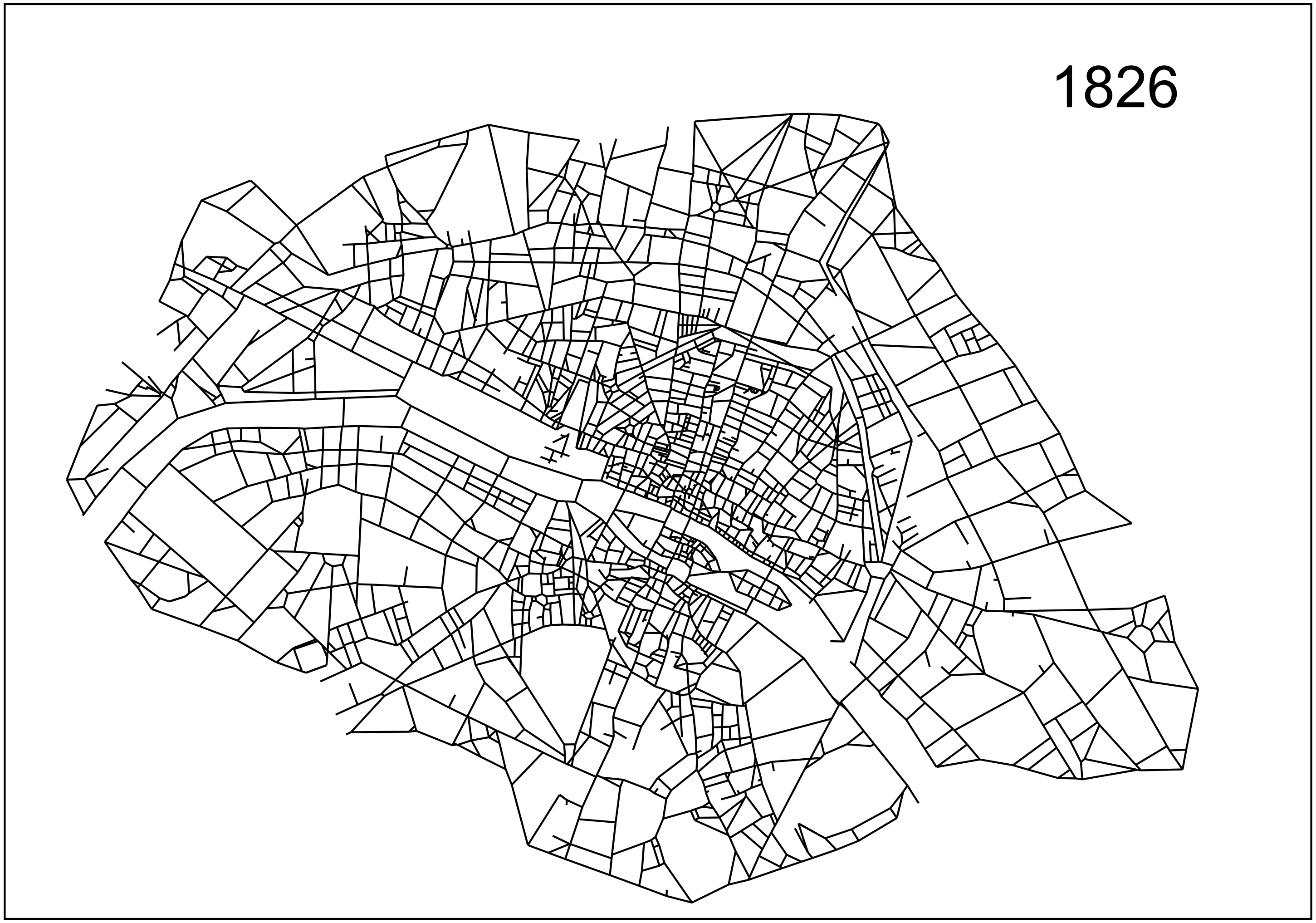


- Haussmann period (~1853-1870)

Road network evolution
Central Paris, France 1789-2010

MB, Bordin, Berestycki, Gribaudi (2013)

Road network evolution
Central Paris, France 1789-2010

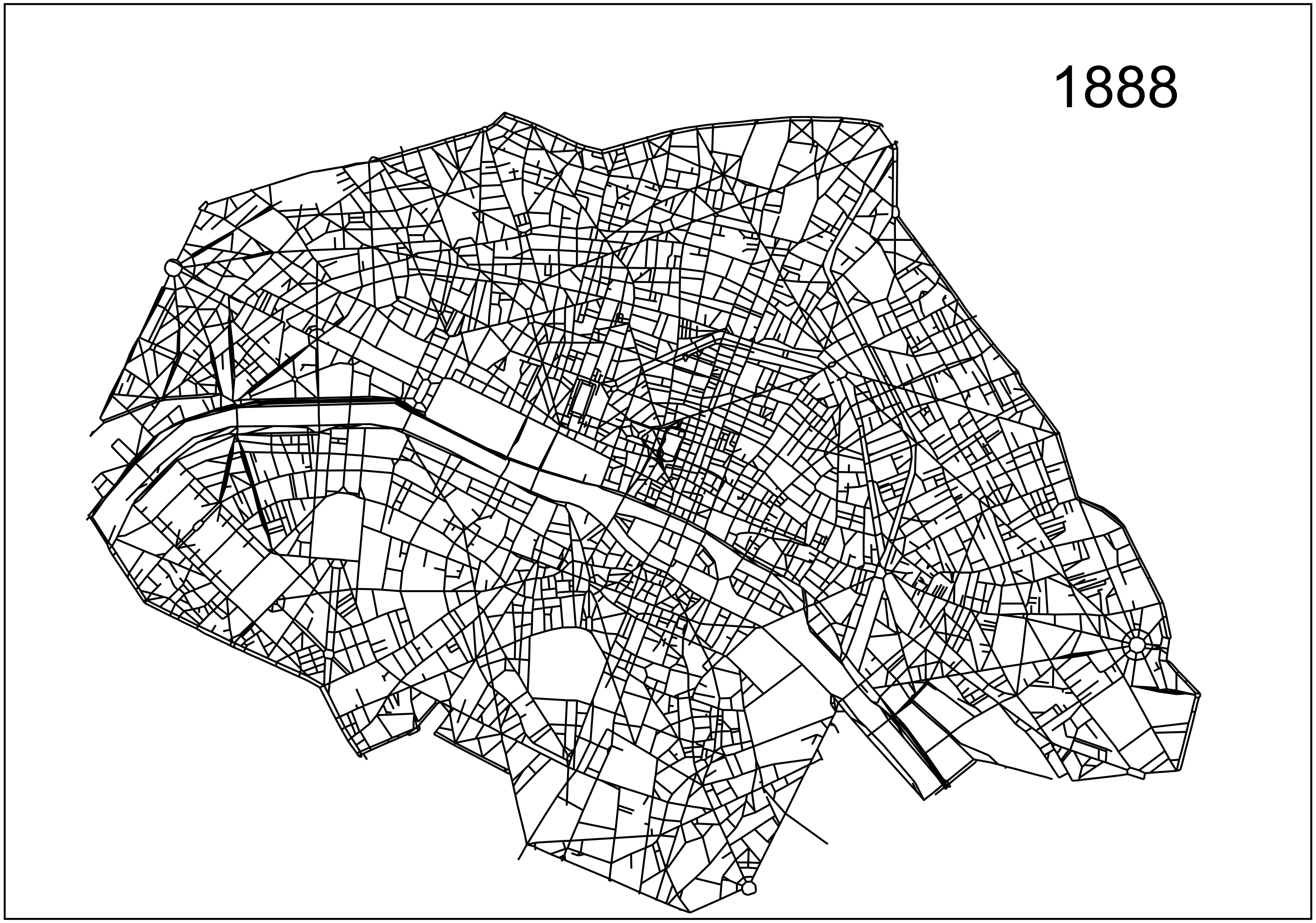


MB, Bordin, Berestycki, Gribaudi (2013)

Road network evolution
Central Paris, France 1789-2010

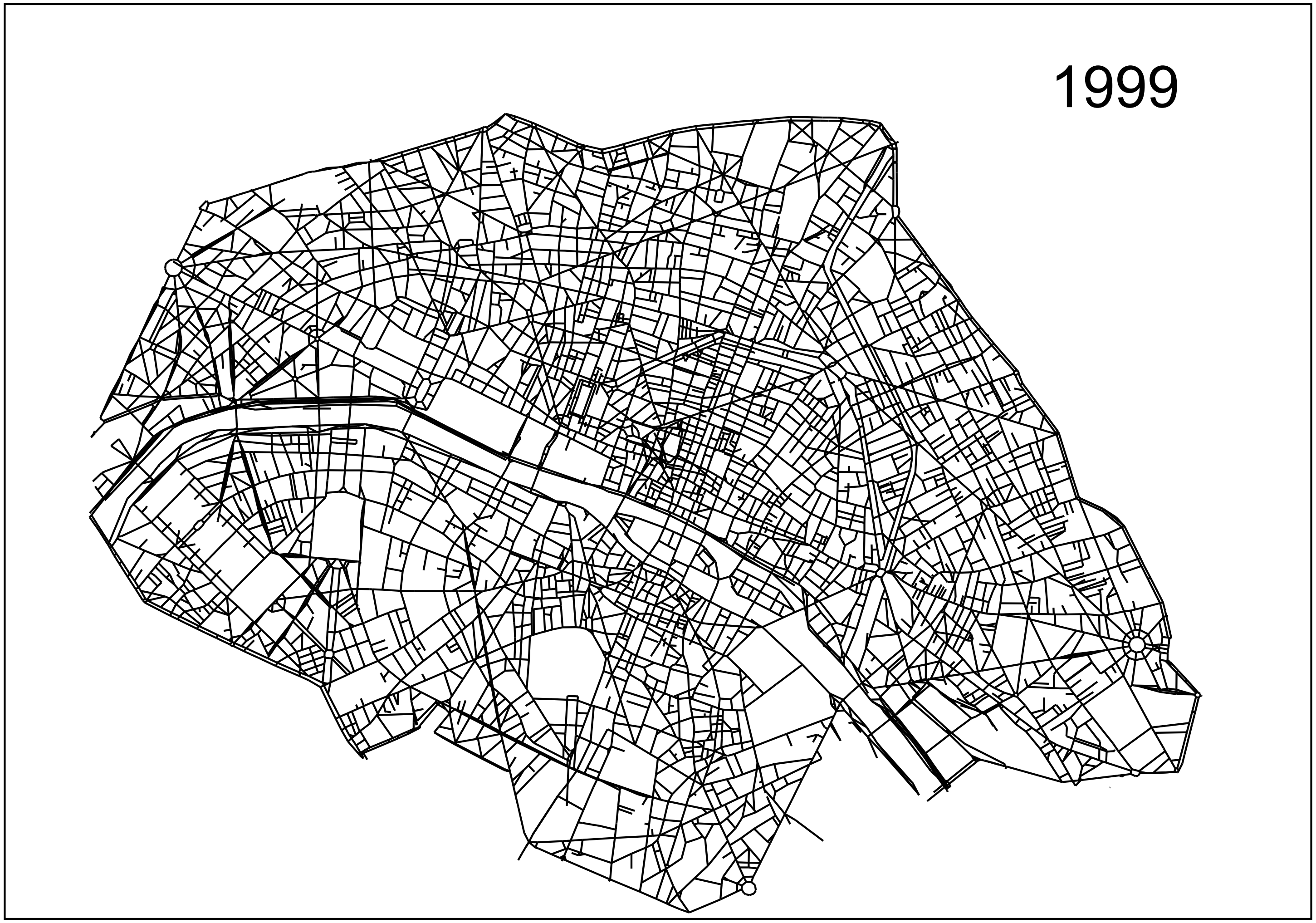
MB, Bordin, Berestycki, Gribaudi (2013)

Road network evolution
Central Paris, France 1789-2010



MB, Bordin, Berestycki, Gribaudi (2013)

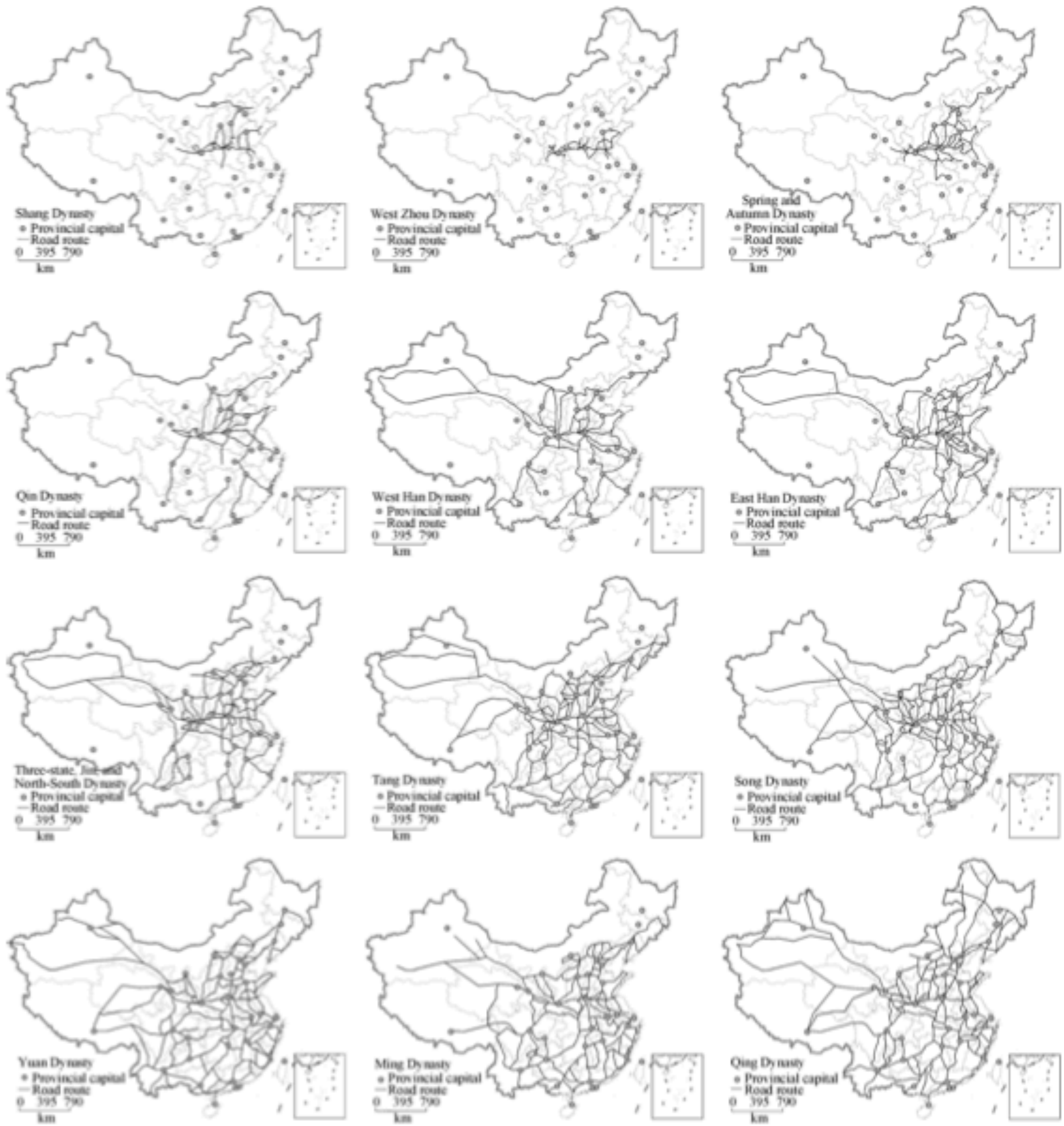
Road network evolution
Central Paris, France 1789-2010



MB, Bordin, Berestycki, Gribaudi (2013)

Road network China

1600(BC)- 1900 (AC)



Wang, Ducruet, Wang (2015)

1. Simple measures

Road network evolution (Groane region, Italy)

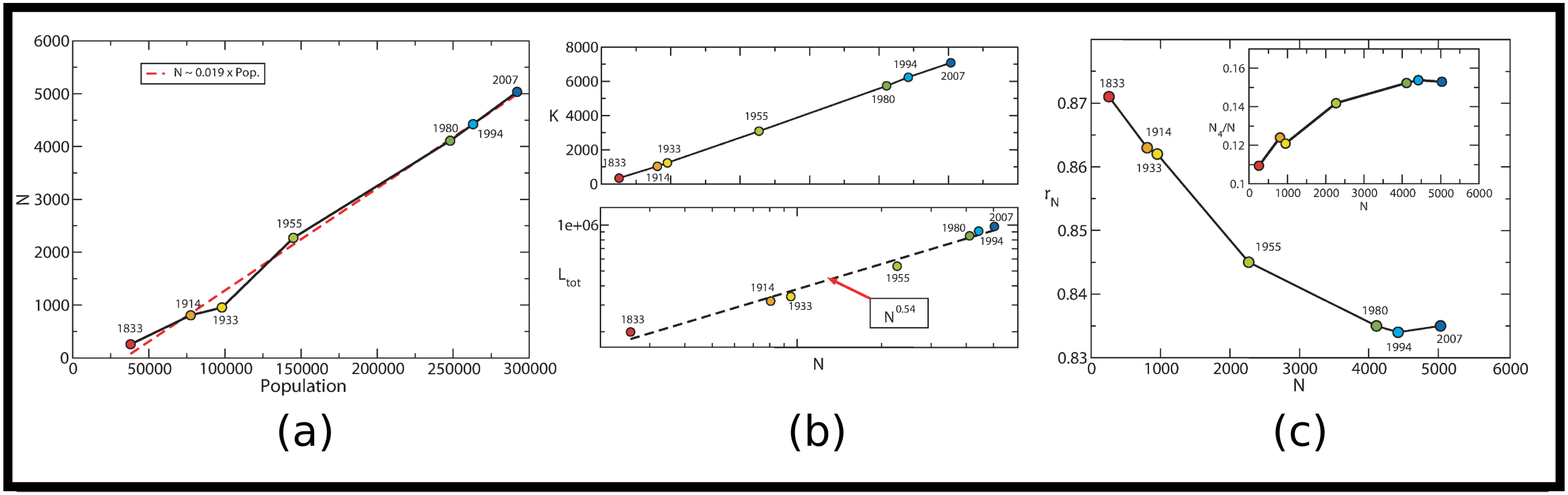


Figure 2 | (a) Number of nodes N versus total population (continuous line with circles) and its linear fit (red dashed line). (b) Total number of links E and total network length L_{tot}, as a function of the number of nodes N. The total network length increases as $N^{0.54}$. (c) Value of the ratio r_N between the number of nodes with degree $k = 1$ and $k = 3$, and the total number of nodes. In the inset we report the percentage of nodes having degree $k = 4$ as a function of N. Notice that the relative abundance of four-ways crossings increases by 5% in two centuries.

$$r_N = \frac{N(1) + N(3)}{\sum_{k \neq 2} N(k)}$$

Road network evolution

- Total length of the network of size N

$$L_{tot} \sim \sqrt{N}$$

- Density $\rho = \dfrac{N}{A}$

- Assumption: weakly perturbed lattice

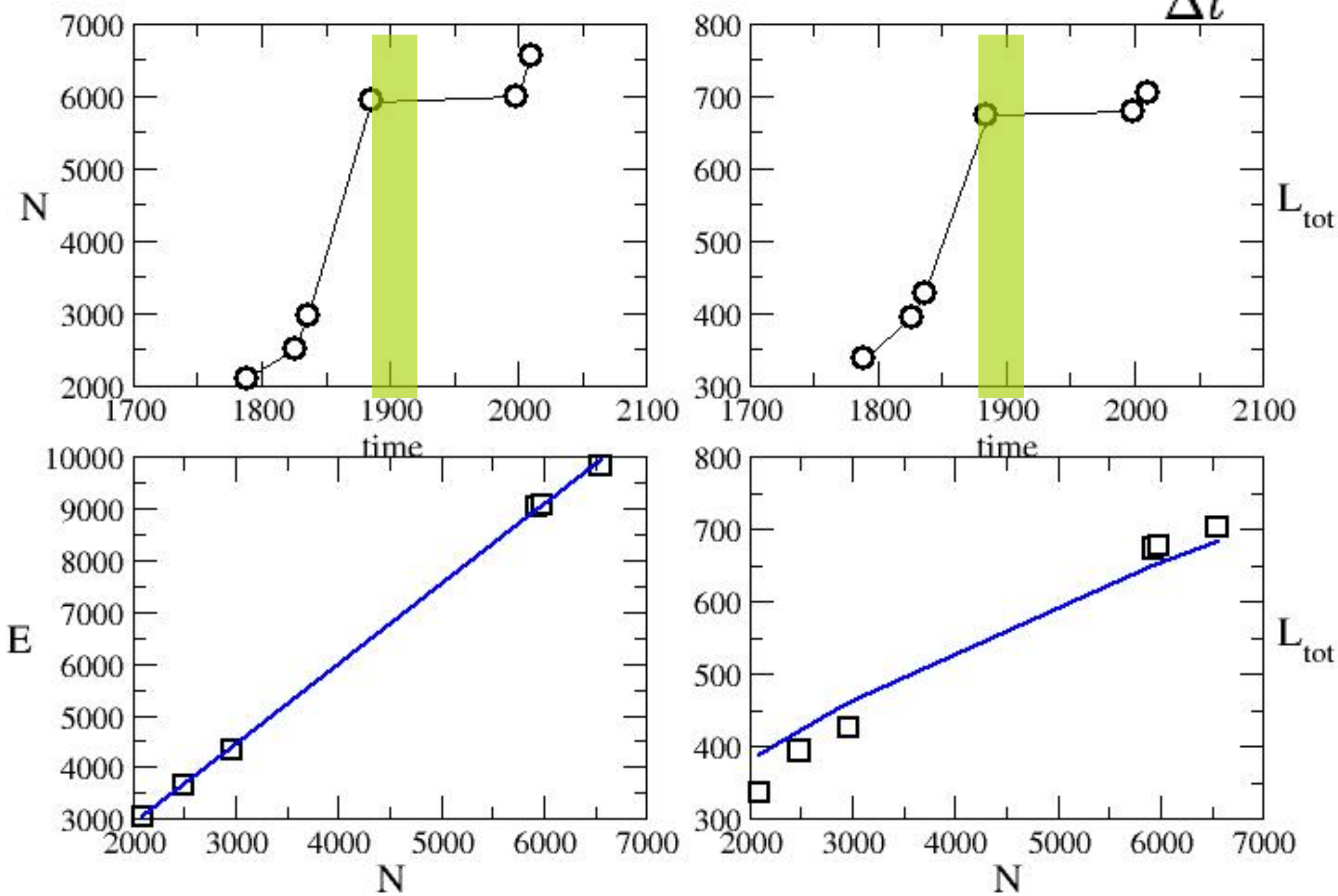
$$\ell_1 \sim 1/\sqrt{\rho}$$

$$\Rightarrow L_{tot} \sim N \times \ell_1 \sim \sqrt{N}$$

Time evolution (Paris 1789-2010)

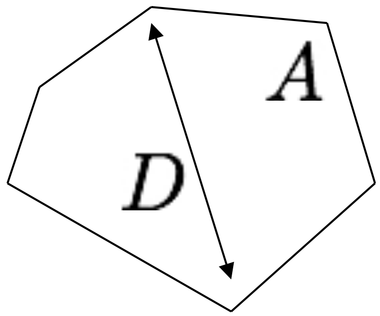
- Standard indicators versus time or N

$$\frac{\Delta L_{tot}}{\Delta t} \approx 1.6 \ \text{km/year}$$



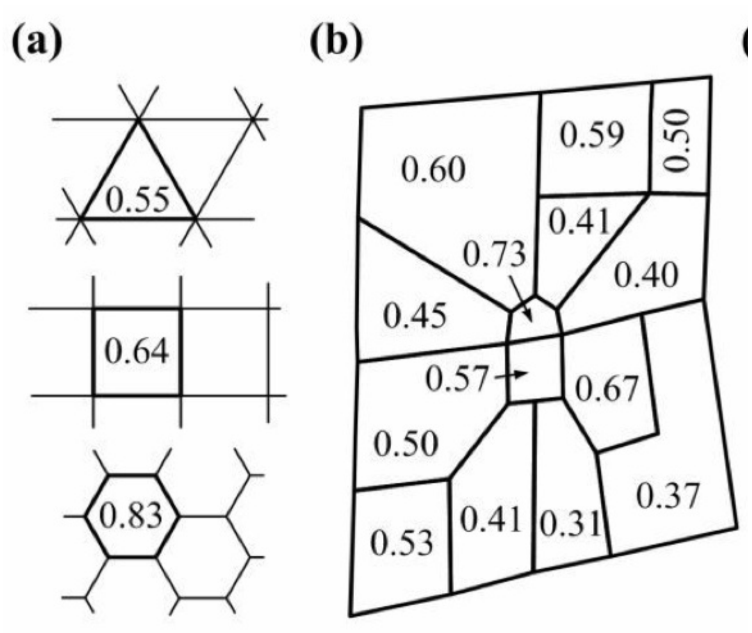
$$L_{tot} \sim \frac{\langle k \rangle}{2} \sqrt{AN}$$

Faces (blocks): shape and area

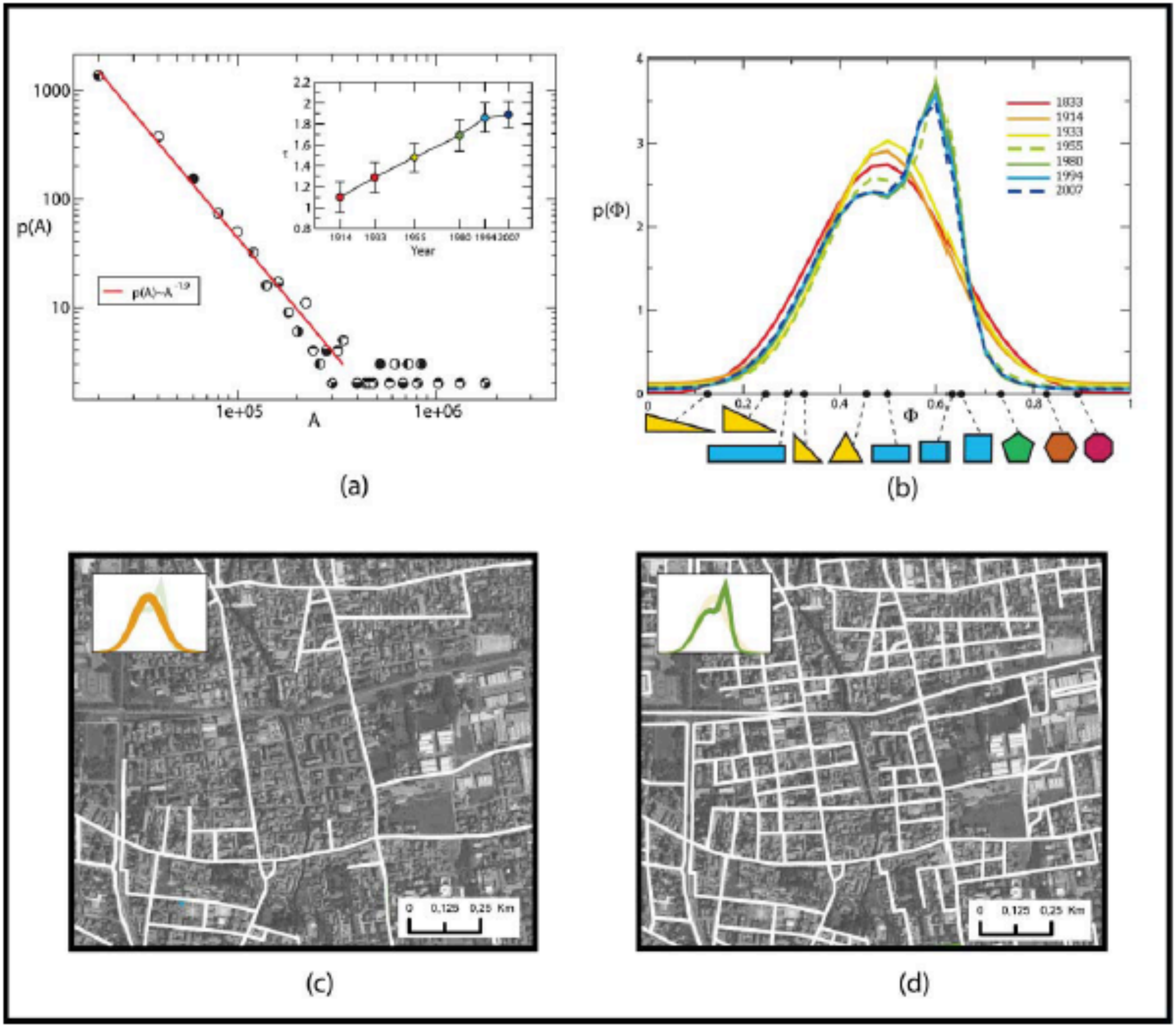


$$P(A) \sim ?$$

$$\phi = \frac{A}{(\pi D^2/4)}$$



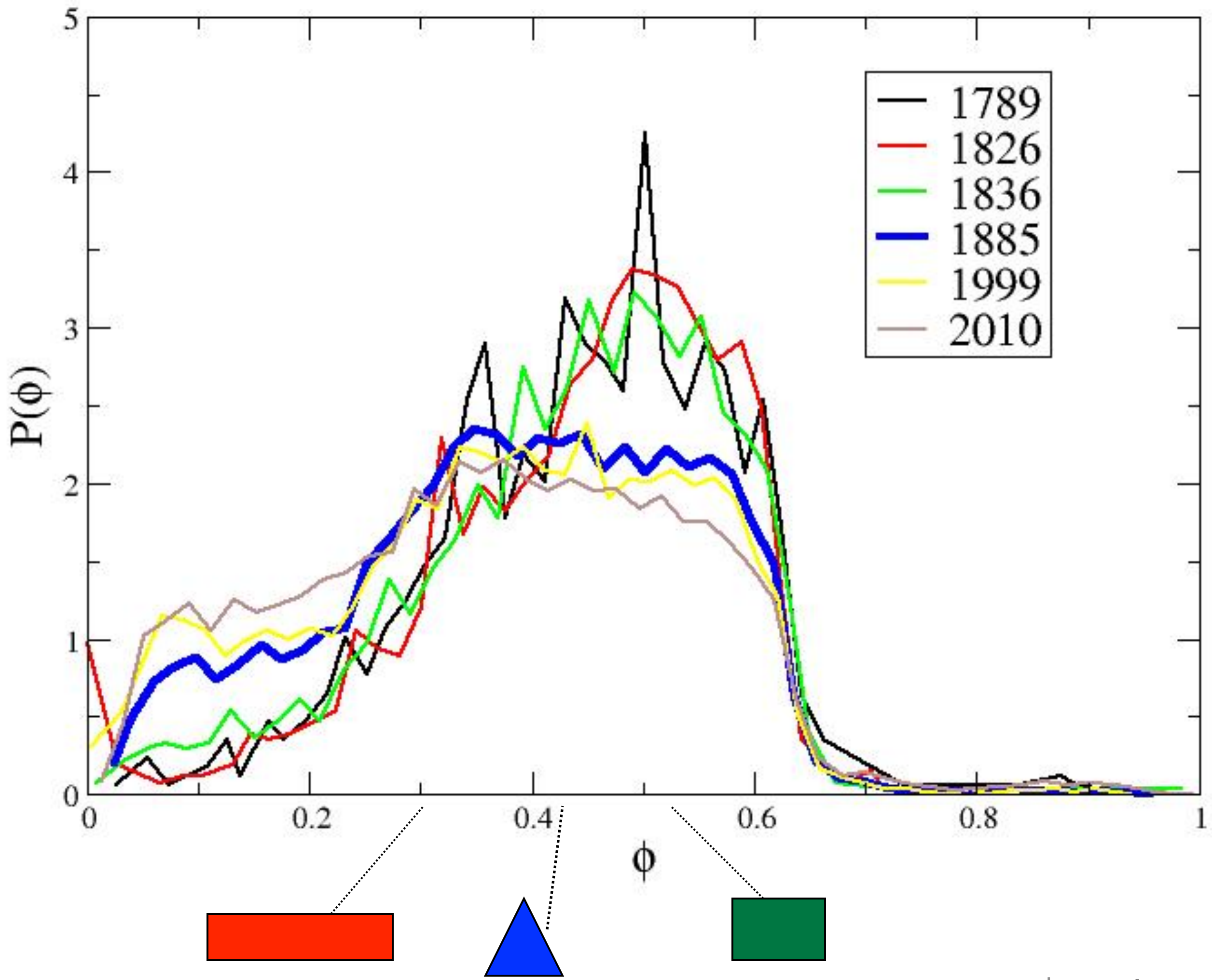
Lammer et al, Physica A (2006)



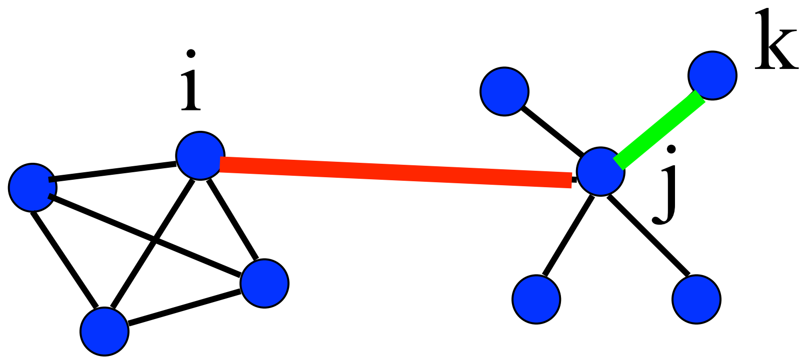
p(A)
1000
100
10
A
1e+05
1e+06
p(A)~A^-1.9
τ
Year
1914
1933
1955
1980
1994
2007
p(Φ)
Φ
1833
1914
1933
1955
1980
1994
2007
0 0,125 0,25 Km
0 0,125 0,25 Km
(a)
(b)
(c)
(d)

2015

Haussmann effect: shape factor



More interesting: Betweenness Centrality (Freeman '77)



ij: large centrality

jk: small centrality

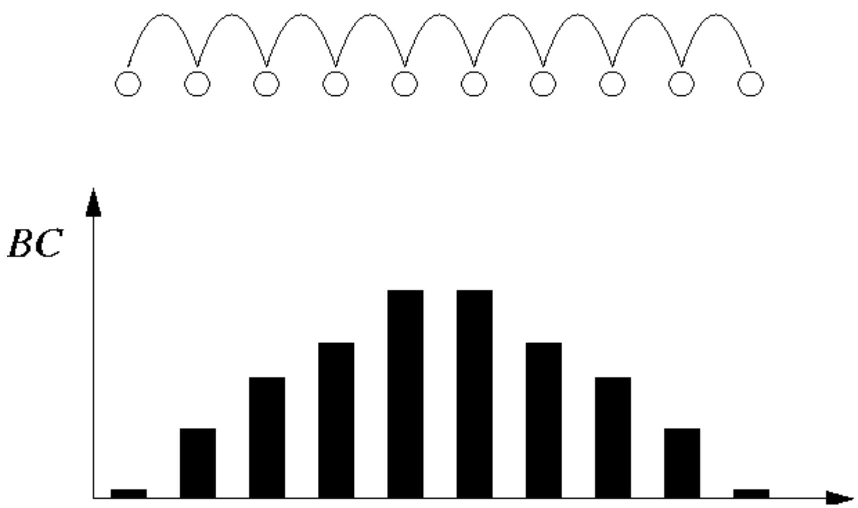
$$g(ij) = \sum_{s,t} \frac{\sigma_{st}(ij)}{\sigma_{st}}$$

σ_{st} = # of shortest paths from s to t

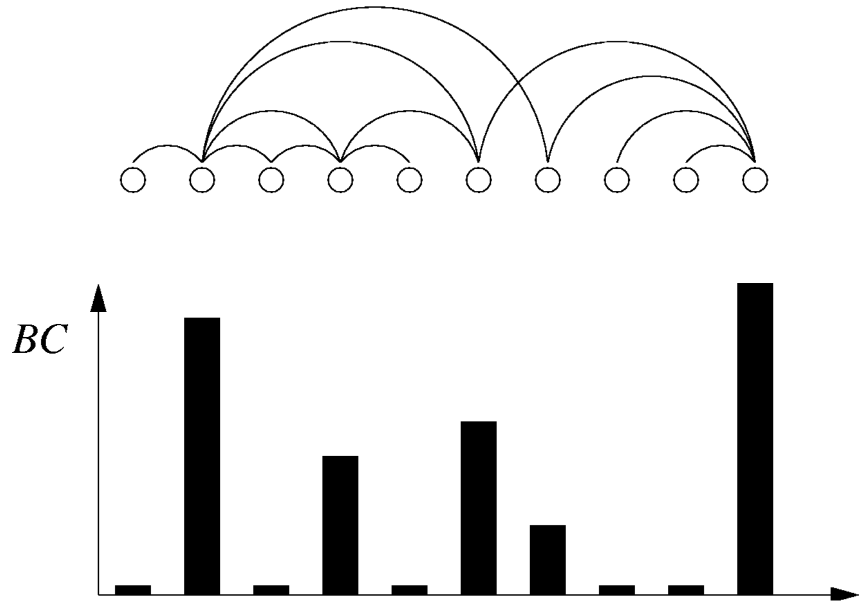
$\sigma_{st}(ij)$= # of shortest paths from s to t **via** (ij)

Measures the importance of a segment in the shortest paths flow

Betweenness centrality and space

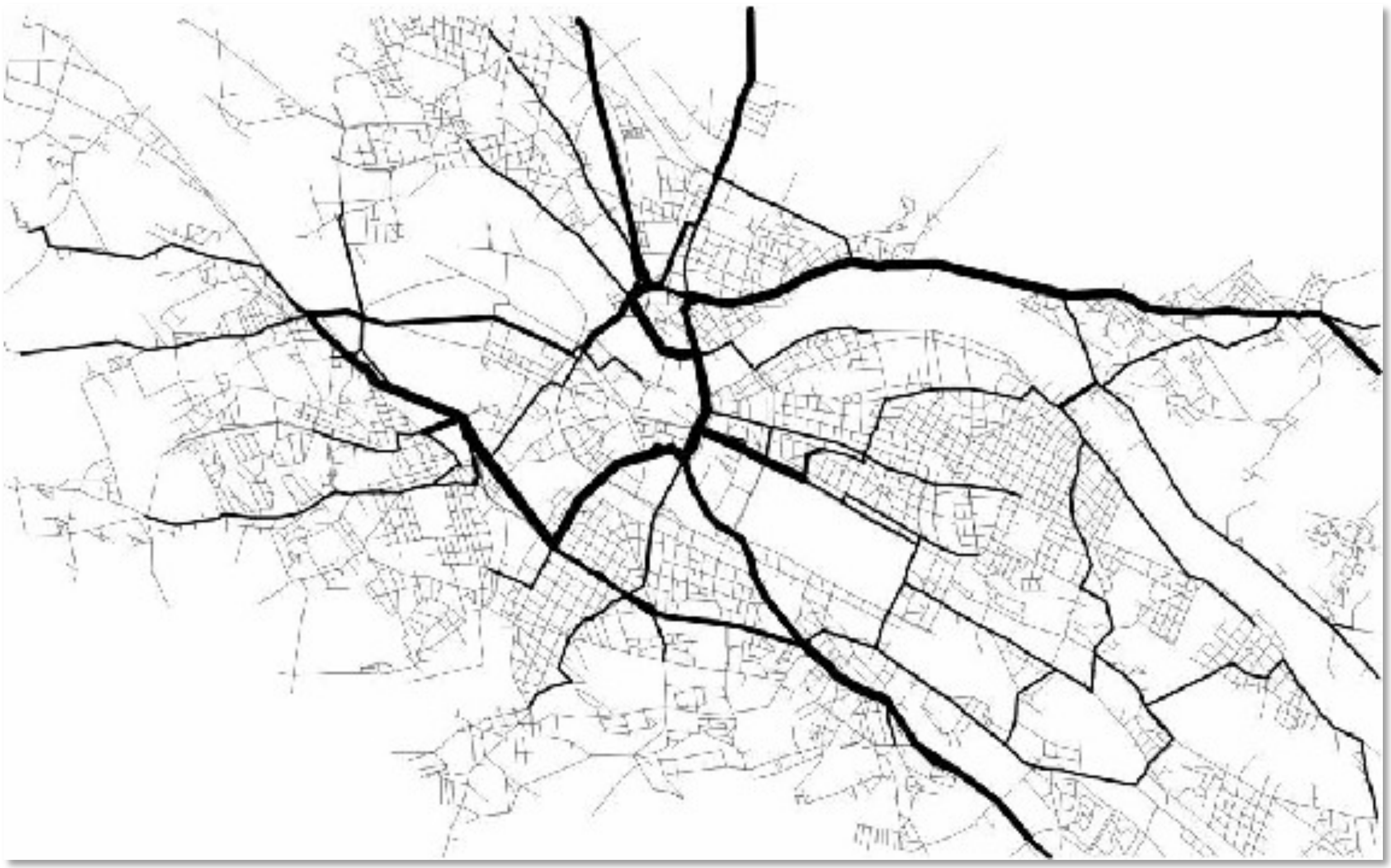


Large BC: distance to barycenter



Large BC: large degree

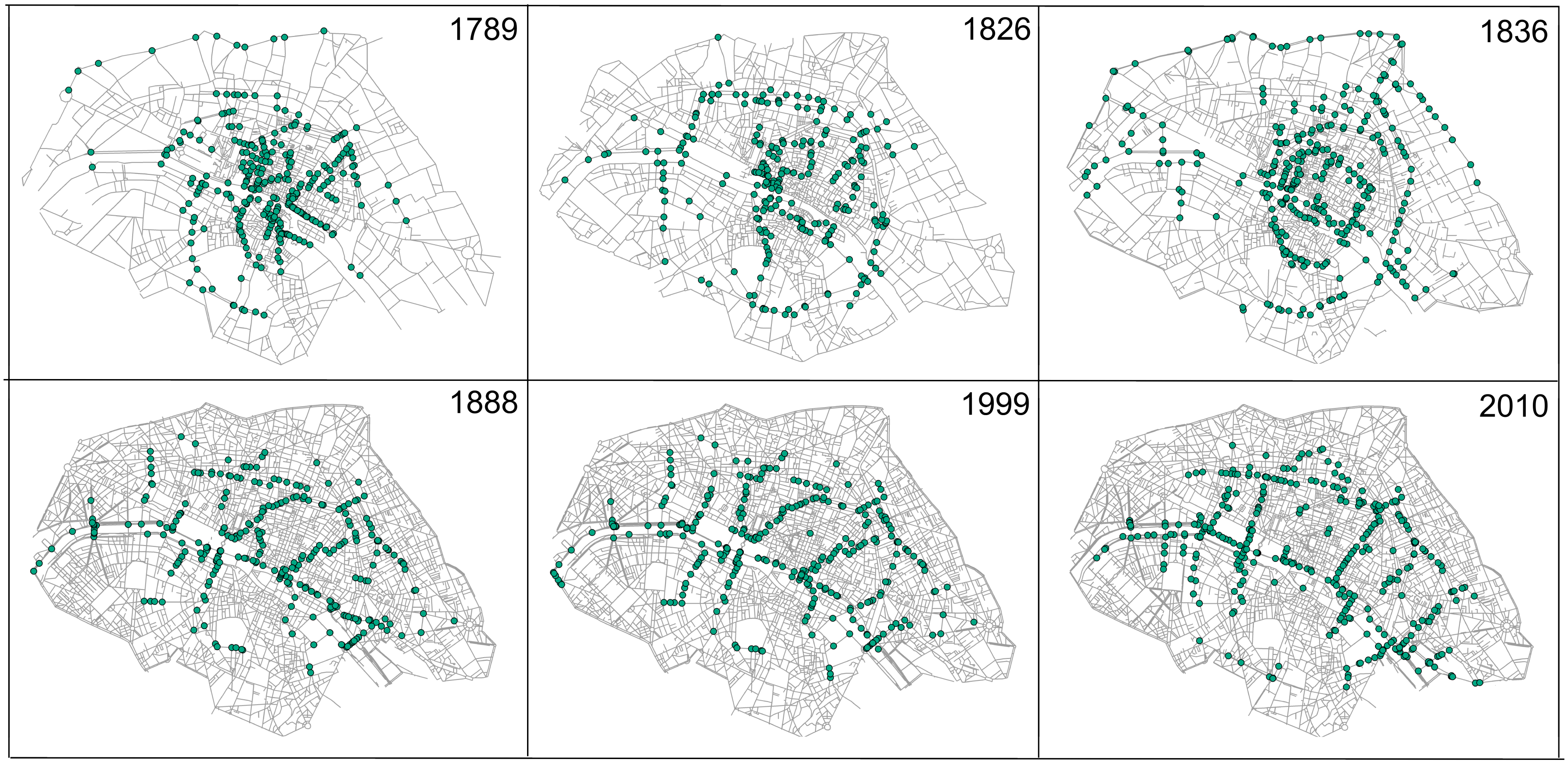
Betweenness centrality and space



Lammer et al, 2006

Haussmann effect

- Spatial distribution of centrality (most central nodes)



Betweenness centrality

- Backbone of stable central roads

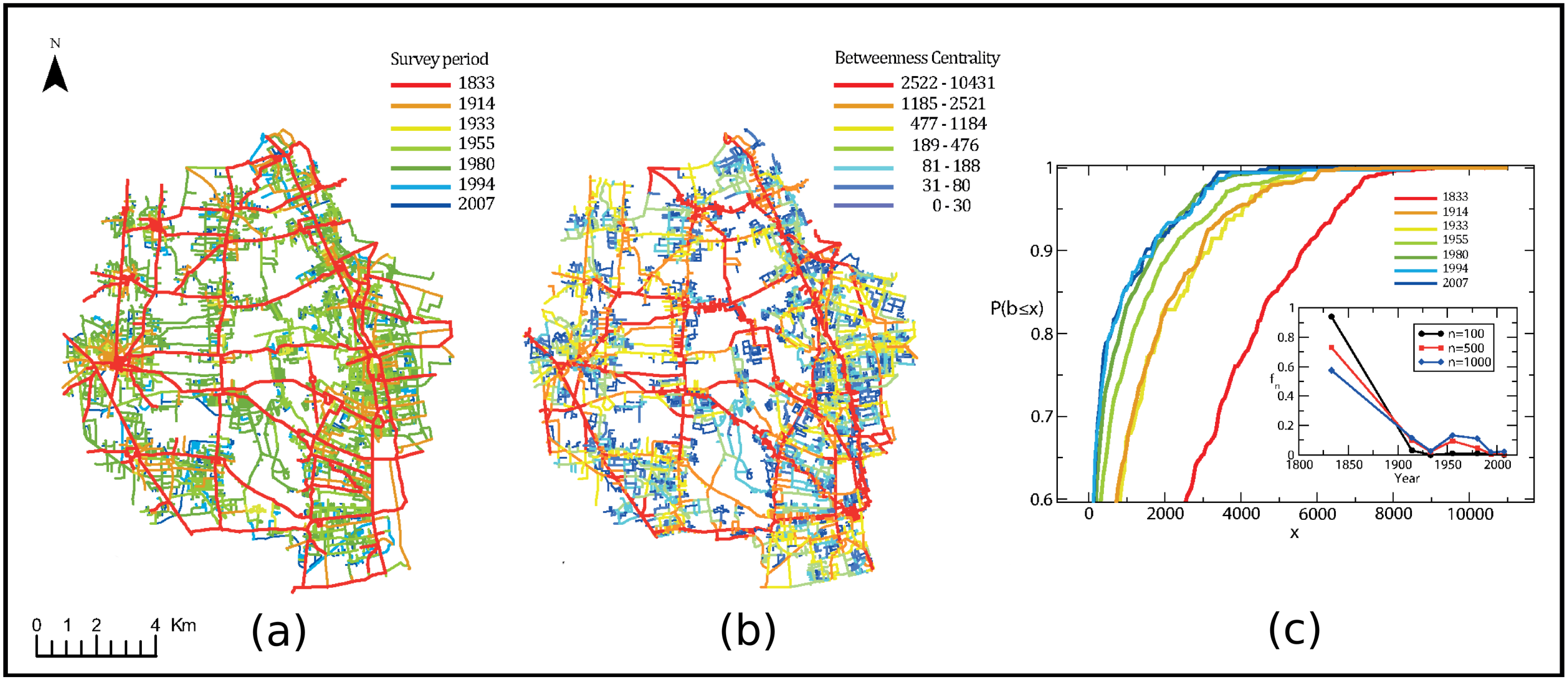


Figure 6 | Color maps indicating (a) the time of creation of each link and (b) its value of betweenness centrality (BC) at year 2007. (c) The cumulative distribution of BC of links added at different times. The inset reports the percentage of edges added at a certain time which are ranked in the top n positions according to the BC. Different curves correspond to n = 100, 500, 1000.

Characterization of new links: BC impact

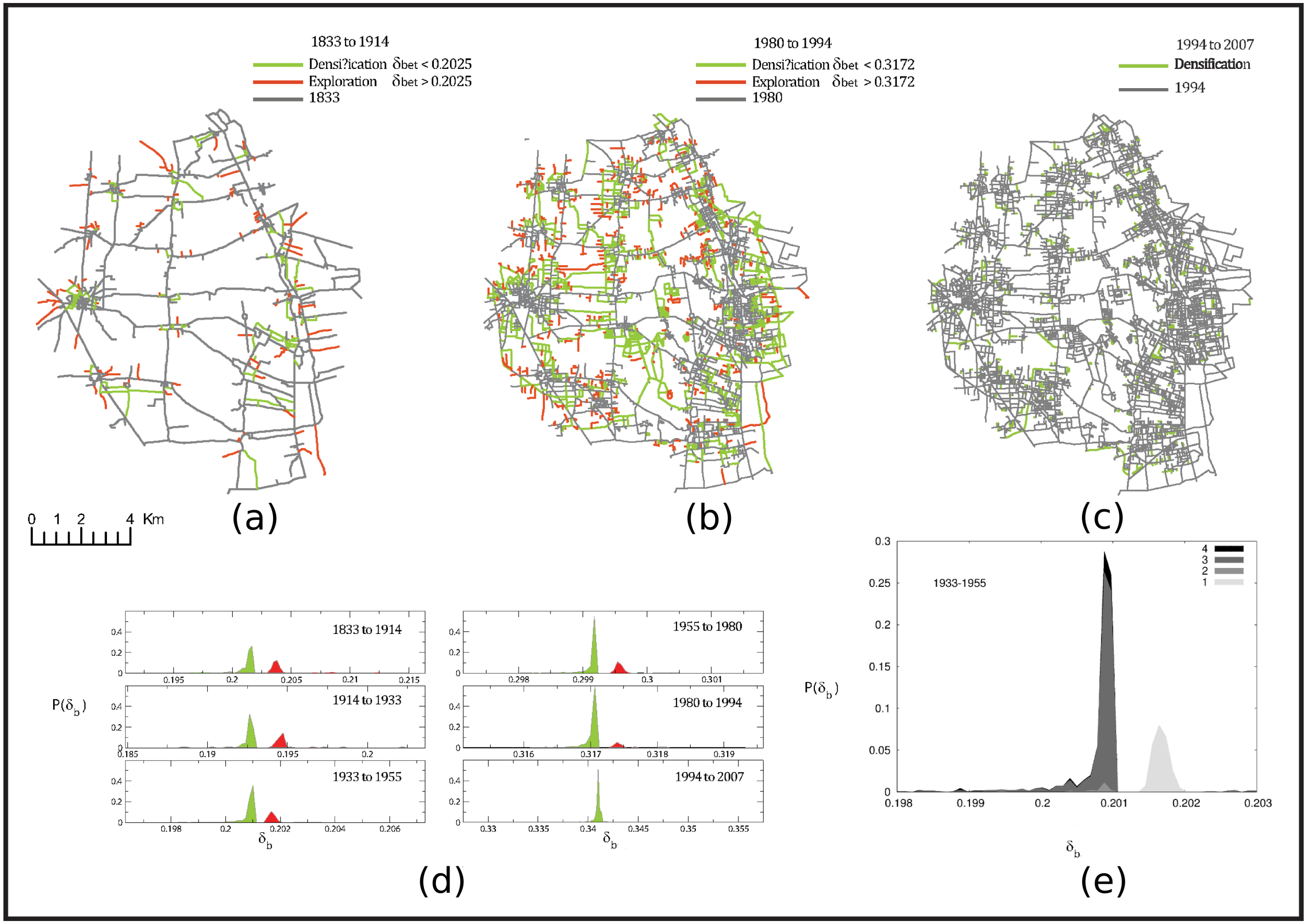
- Average BC of the graph at time t:

$$\bar{b}(G_t) = \frac{1}{(N(t)-1)(N(t)-2)} \sum_{e \in E_t} b(e)$$

- BC impact of new edge e*:

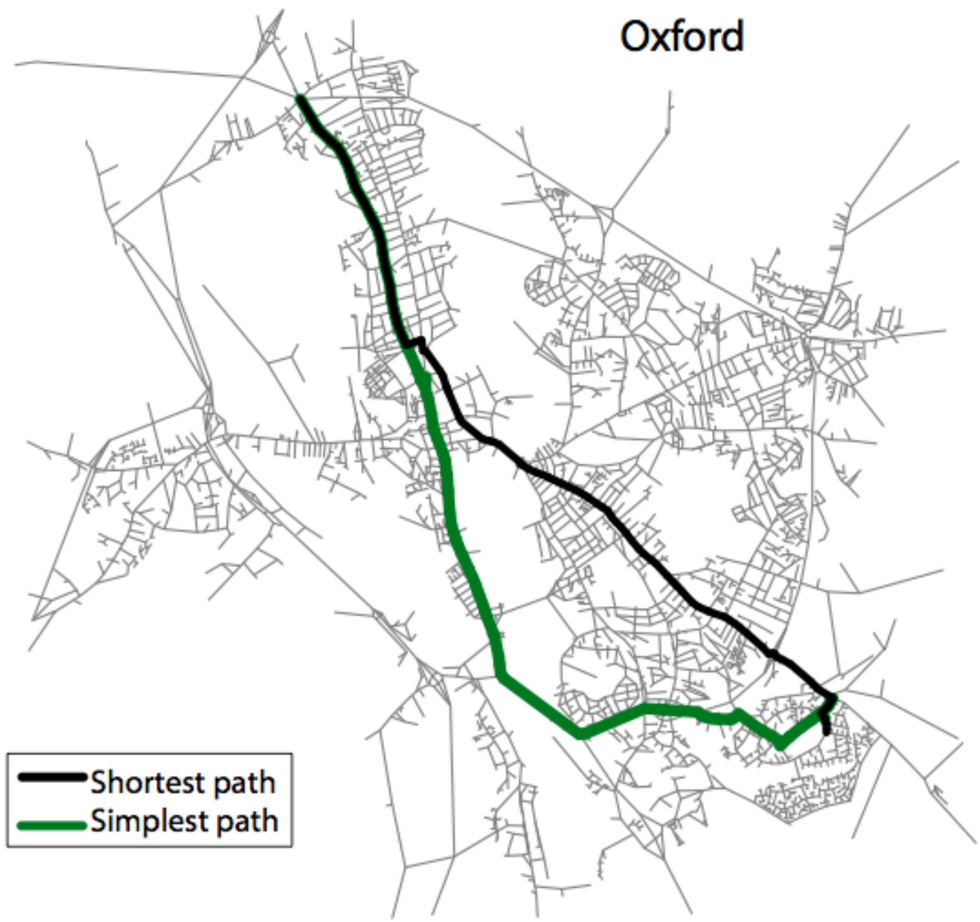
$$\delta b(e*) = \frac{\bar{b}(G_t) - \bar{b}(G_t \setminus \{e*\})}{\bar{b}(G_t)}$$

Evolution: two processes



- Two different categories of new links: ‘densification’ and ‘exploration’ clearly identified by the BC impact

Another measure: Simplicity of planar networks



Statistical comparison of the length of shortest and simplest paths (with the minimal number of turns)

Simplicity of paths

$$S(d) = \frac{1}{N(d)} \sum_{i,j/d(i,j)=d} \frac{\ell^*(i,j)}{\ell(i,j)}$$

$\ell(i,j)$ Length of shortest path

$\ell^*(i,j)$ Length of simplest path

For small d: $S(d \to 0) \approx 1$ and increases

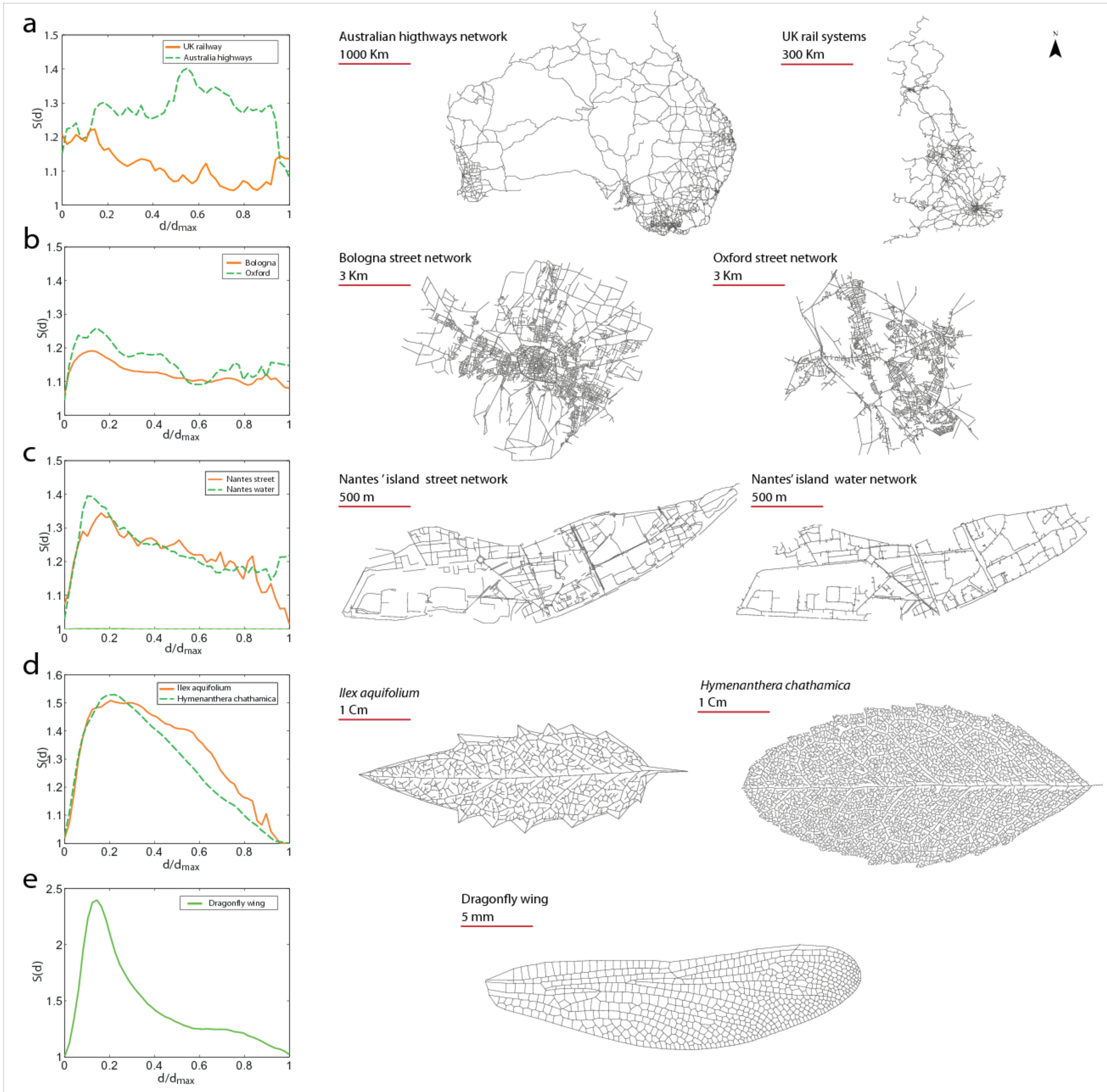
For large d: $S(d \to d_{\max}) \approx 1$

$\Rightarrow$ There is a (at least one) maximum at d=d*

Meaning of d*: typical size of 'domains' not crossed by long straight lines

Viana, Strano, Bordin, MB Scientific reports (2013)

Simplicity Spectrum



Viana, Strano, Bordin, MB Scientific reports (2013)

2. Template: the subway case

The subway evolution: not a new problem

- Cope (1967): Stages of the London underground rail

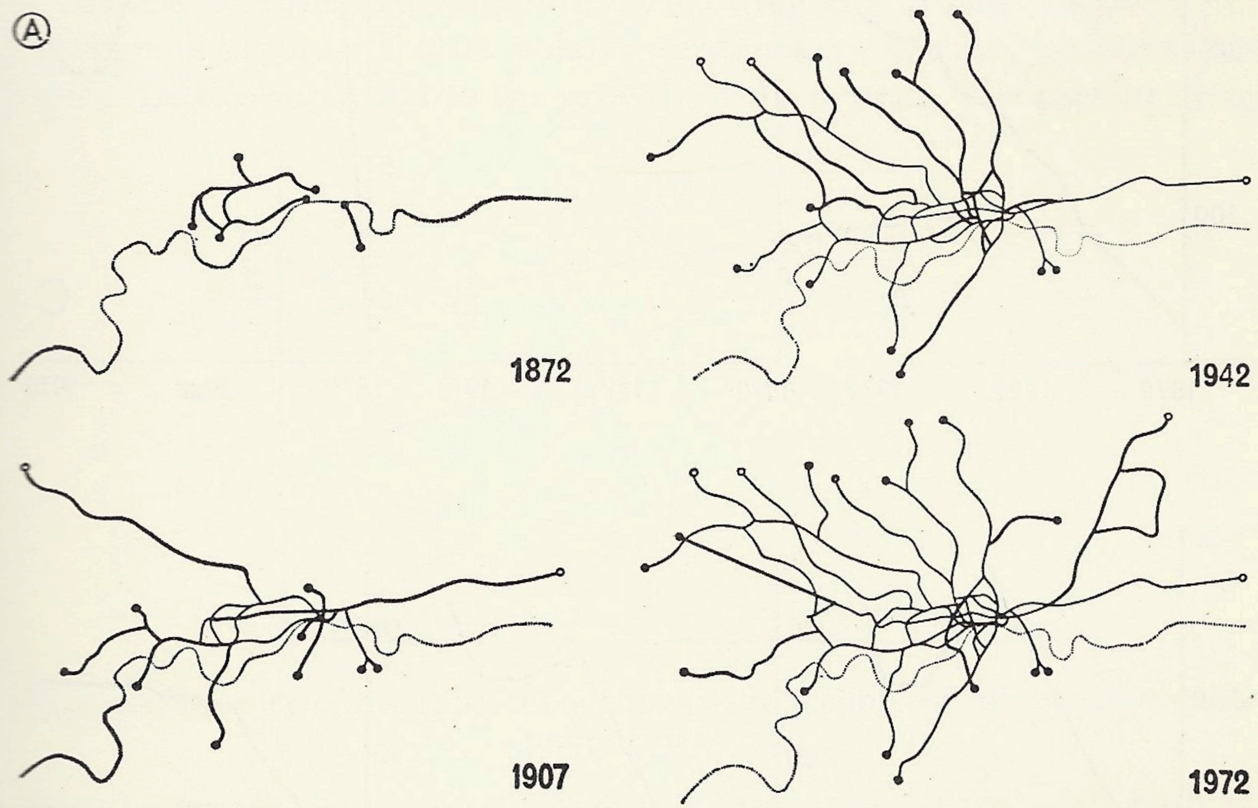


Fig. 5.7. Stages in the growth of the London underground rail network. Source: Cope, 1967.

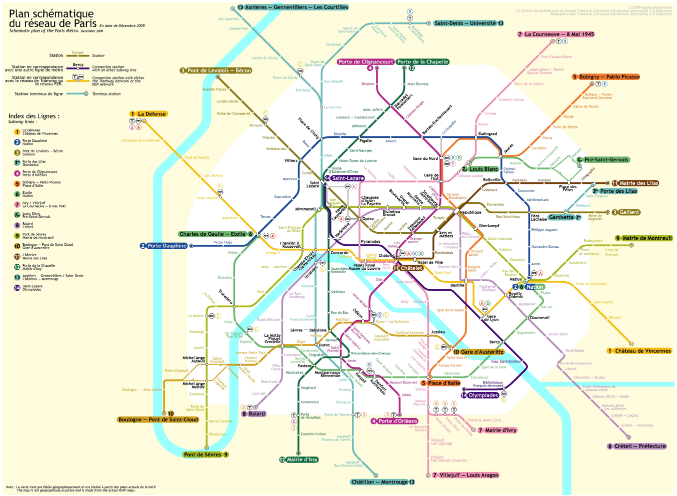
World subway networks

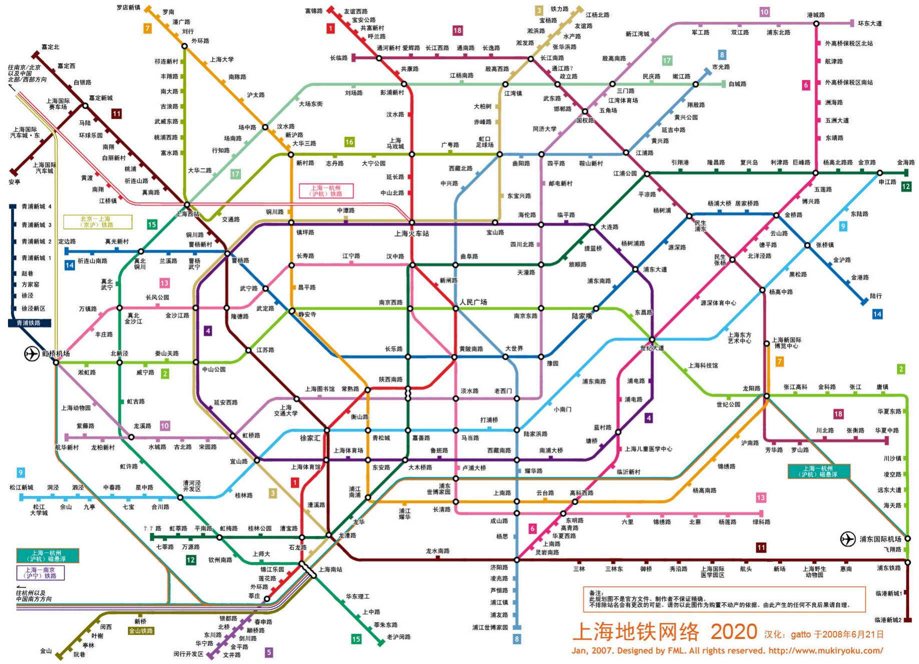
Most large cities have a subway network (50% for $P>10^6$)

We focus on large networks (N>100 stations)

city	P (millions)	N_L	N	$\overline{\ell_1}$ (km)	ℓ_T (km)	ℓ_T/ℓ_T^{reg}	β (%)
Beijing	19.6	9	104	1.79	204	0.14	39
Tokyo	12.6	13	217	1.06	279	0.13	43
Seoul	10.5	9	392	1.39	609	0.39	38
Paris	9.6	16	299	0.57	205	0.18	38
Mexico City	8.8	11	147	1.04	170	0.15	39
NYC	8.4	24	433	0.78	373	0.12	36
Chicago	8.3	11	141	1.18	176	0.08	71
London	8.2	11	266	1.29	397	0.20	47
Shanghai	6.9	11	148	1.47	233	0.21	61
Moscow	5.5	12	134	1.67	260	0.16	71
Berlin	3.4	10	170	0.77	141	0.30	60
Madrid	3.2	13	209	0.90	215	0.42	46
Osaka	2.6	9	108	1.12	137	0.88	43
Barcelona	1.6	11	128	0.72	103	0.32	38

Time evolving spatial networks: too many things to measure !

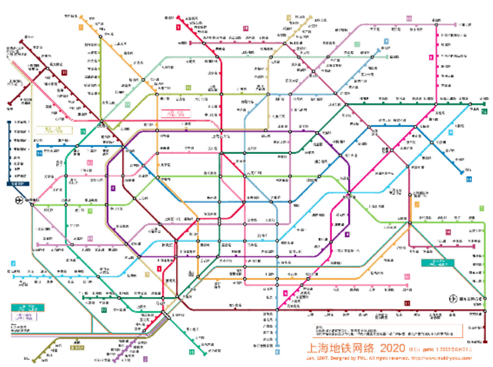

Plan schématique
du réseau de Paris
Index des Lignes :


上海地铁网络 2020
Jan, 2007. Designed by FML. All rights reserved. http://www.mukiryoku.com/

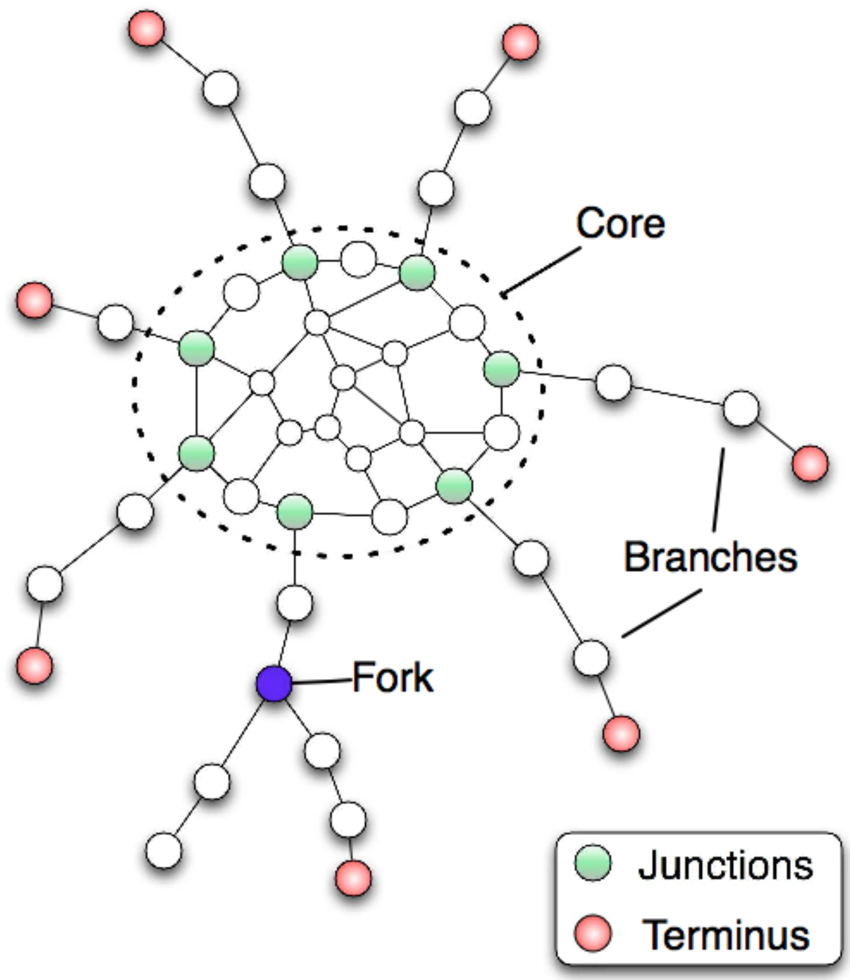

ЛИНИИ МОСКОВСКОГО МЕТРОПОЛИТЕНА


Subway Line
Nanko Port Town Line (New Tram)
Through-Operation on Other Railways
OSAKA Subway
2007 © UrbanRail.Net (Robert Schwandl)

"Universal" template



Algorithm to identify the core and the branches



Measures on this universal structure

- Characterizing the core

$$\langle k_{core} \rangle = \frac{2E_C(t)}{N_C(t)}$$

$$\alpha(t) = \frac{E_C(t) - N_C(t) + 1}{2N_C(t) - 5}$$

N_C: number of nodes in the core
E_C: number of links in the core

Measures on this universal structure

- Characterizing the branches

$$\beta(t) = \frac{N_B(t)}{N_B(t) + N_C(t)}$$

N_B: number of stations in branches
N_C: number of stations in the core

$$\eta(t) = \frac{\overline{D}_B(t)}{\overline{D}_C(t)}$$

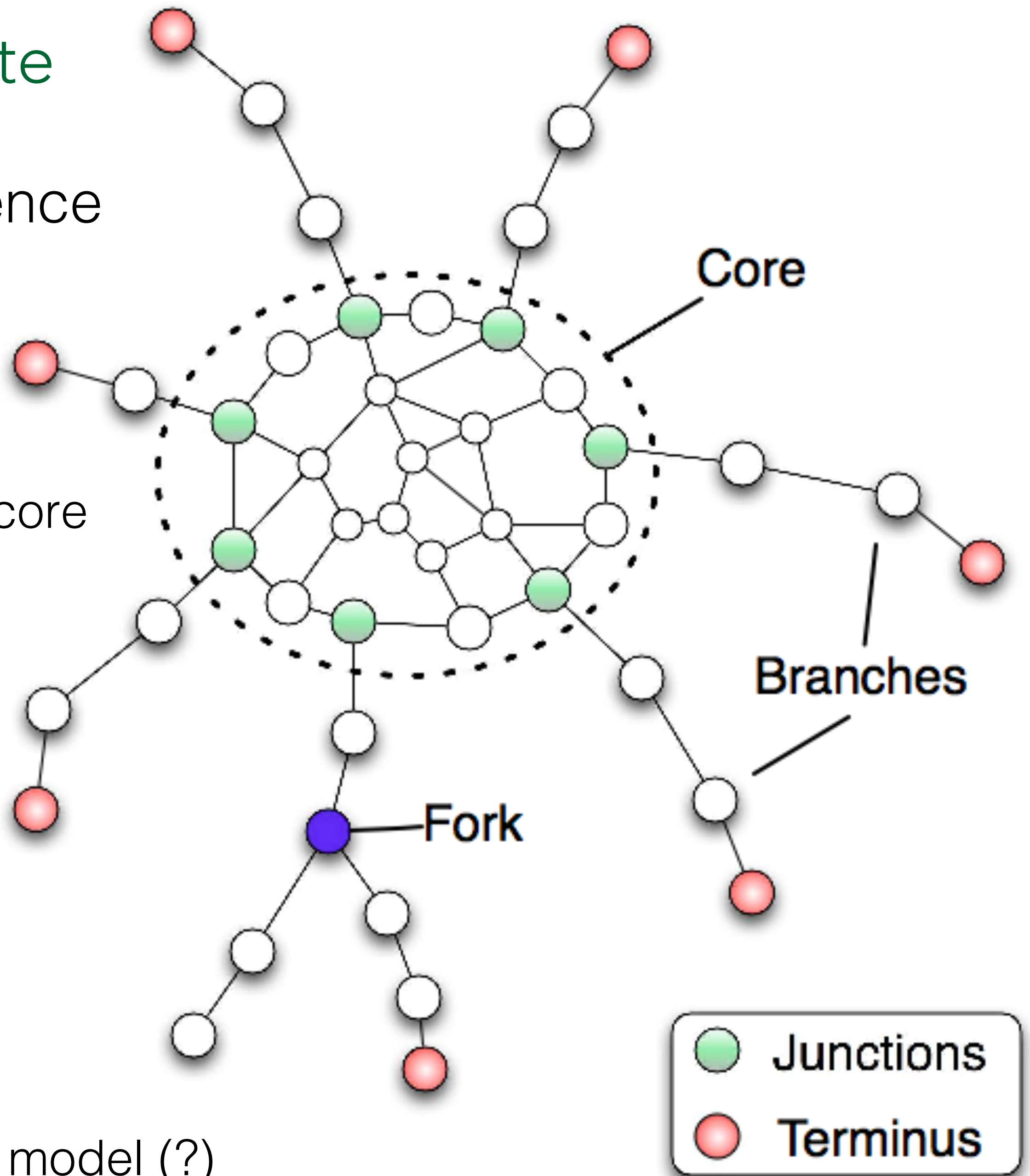
D_B: average distance from barycenter to branches stations
D_C: average distance from barycenter to core stations

"Universal" template

- Quantitative convergence
  - Fraction of branches stations of order 50%
  - Extension of branches/core extension of order 2
  - Average degree of core of order 2.5 and f_2>60%
  - Number of branches

$$\mathcal{N}_B \sim \sqrt{N}$$

=> Existence of a minimal model (?)



Number of branches

- If the spacing b between two branches is constant:

$$\mathcal{N}_B \sim \text{ring perimeter}/b$$

- For a lattice of size N

$$\mathcal{N}_B \sim \sqrt{N}$$

Number of branches

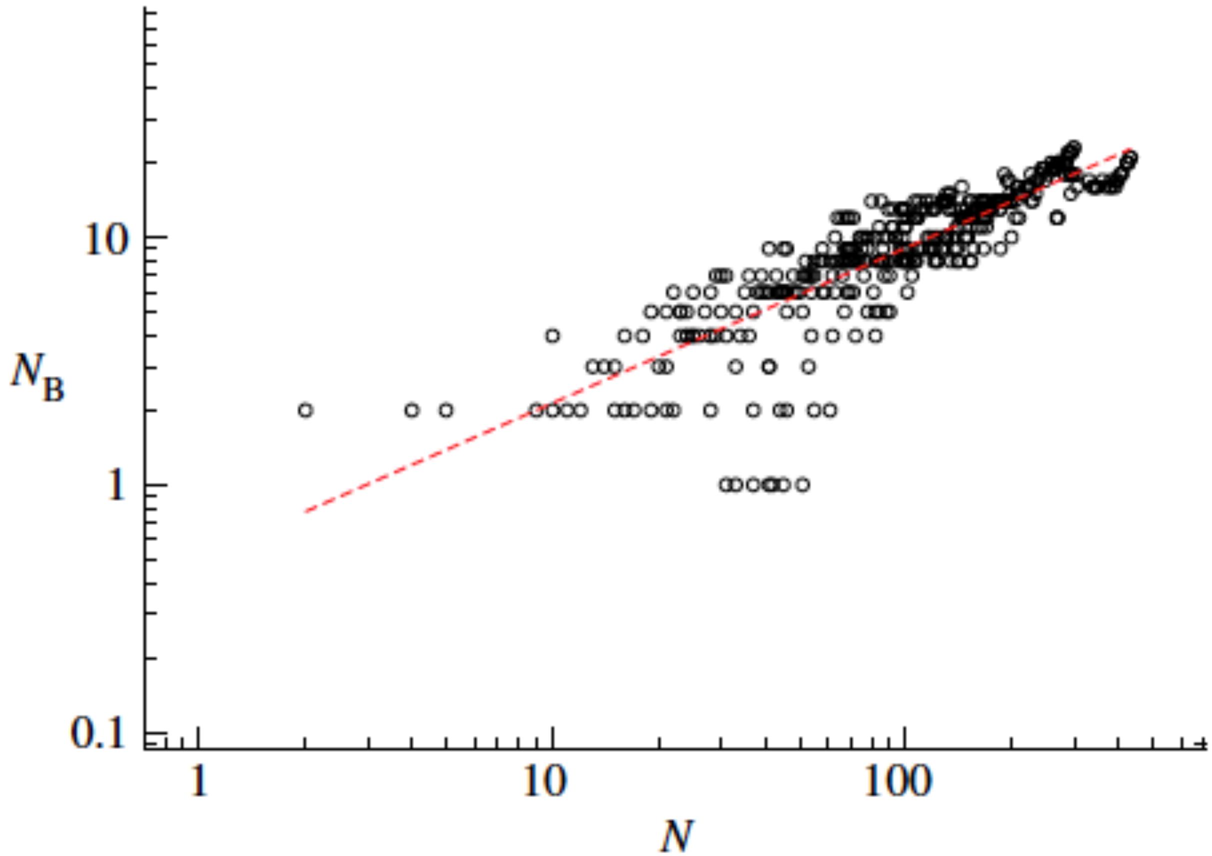


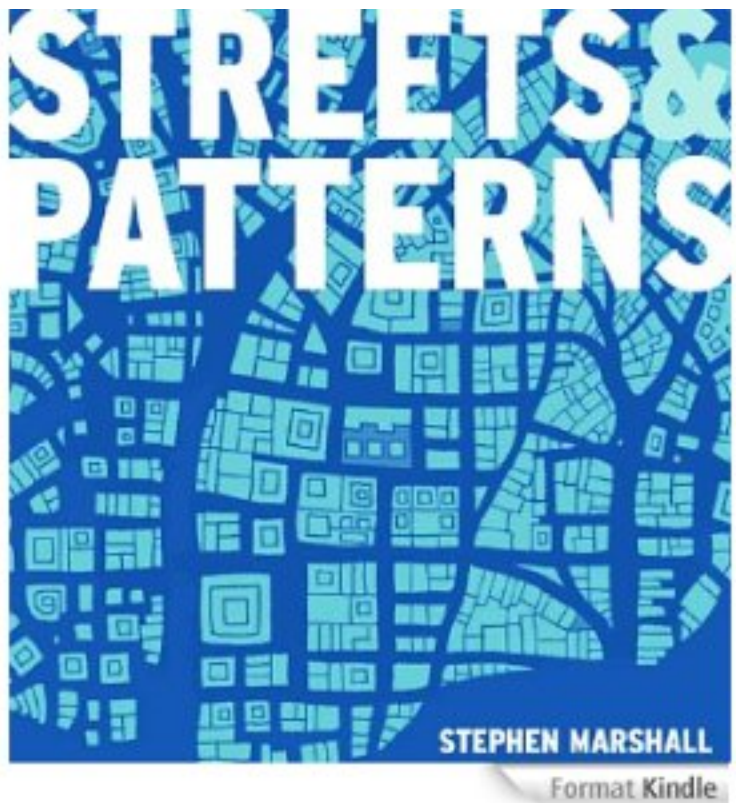
Figure 10. The log–log plot of the number of different branches versus the number of stations for the different subway networks considered here. The dashed line is a power law fit with exponent ≈0.6. (Online version in colour.)

II. Typology: "classification" of planar graphs

Typology of planar graphs

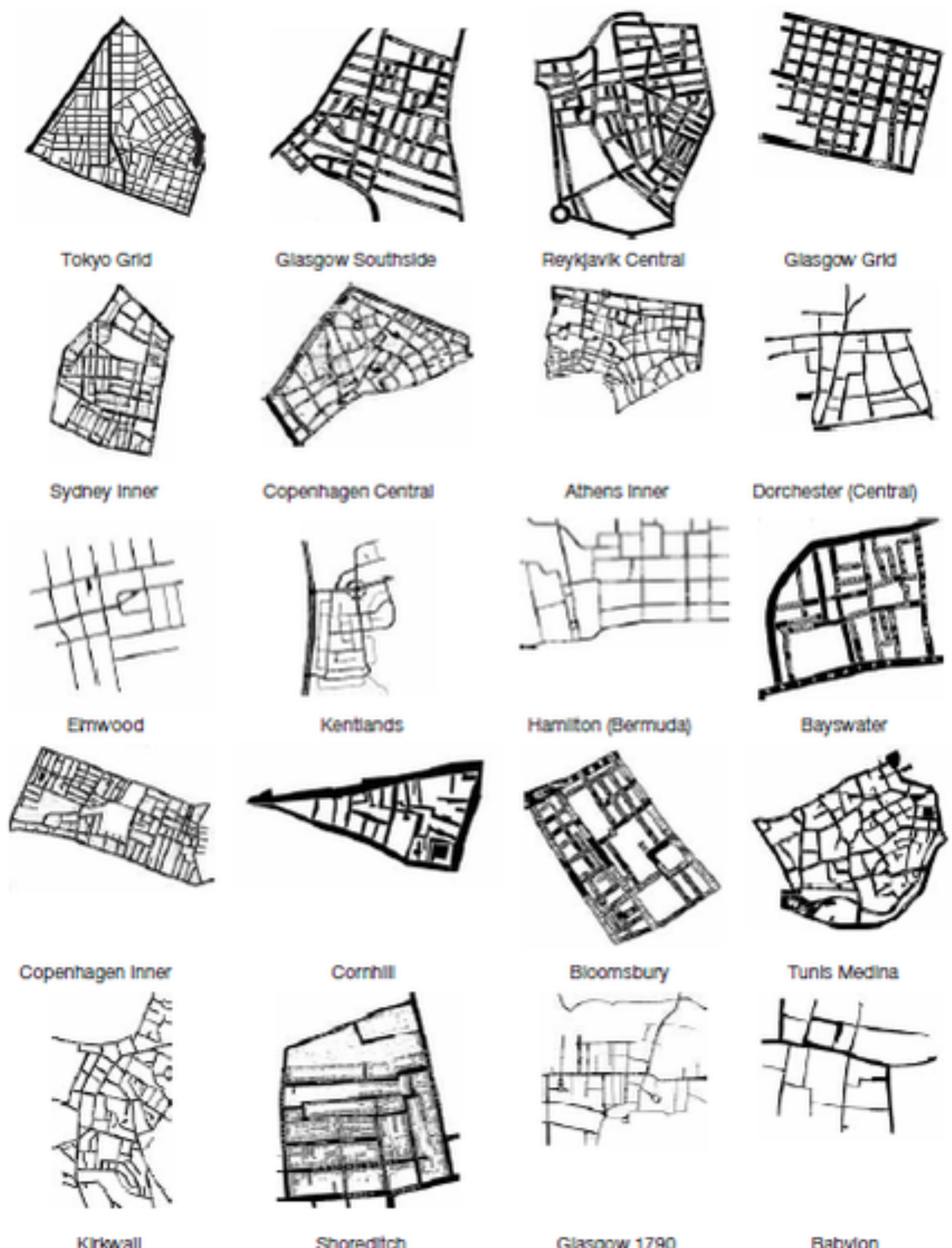
Many applications:

- Botanics (classification of leaves)
- Urban morphology: street network ("Space syntax")



Typology of street networks

136 STREETS & PATTERNS

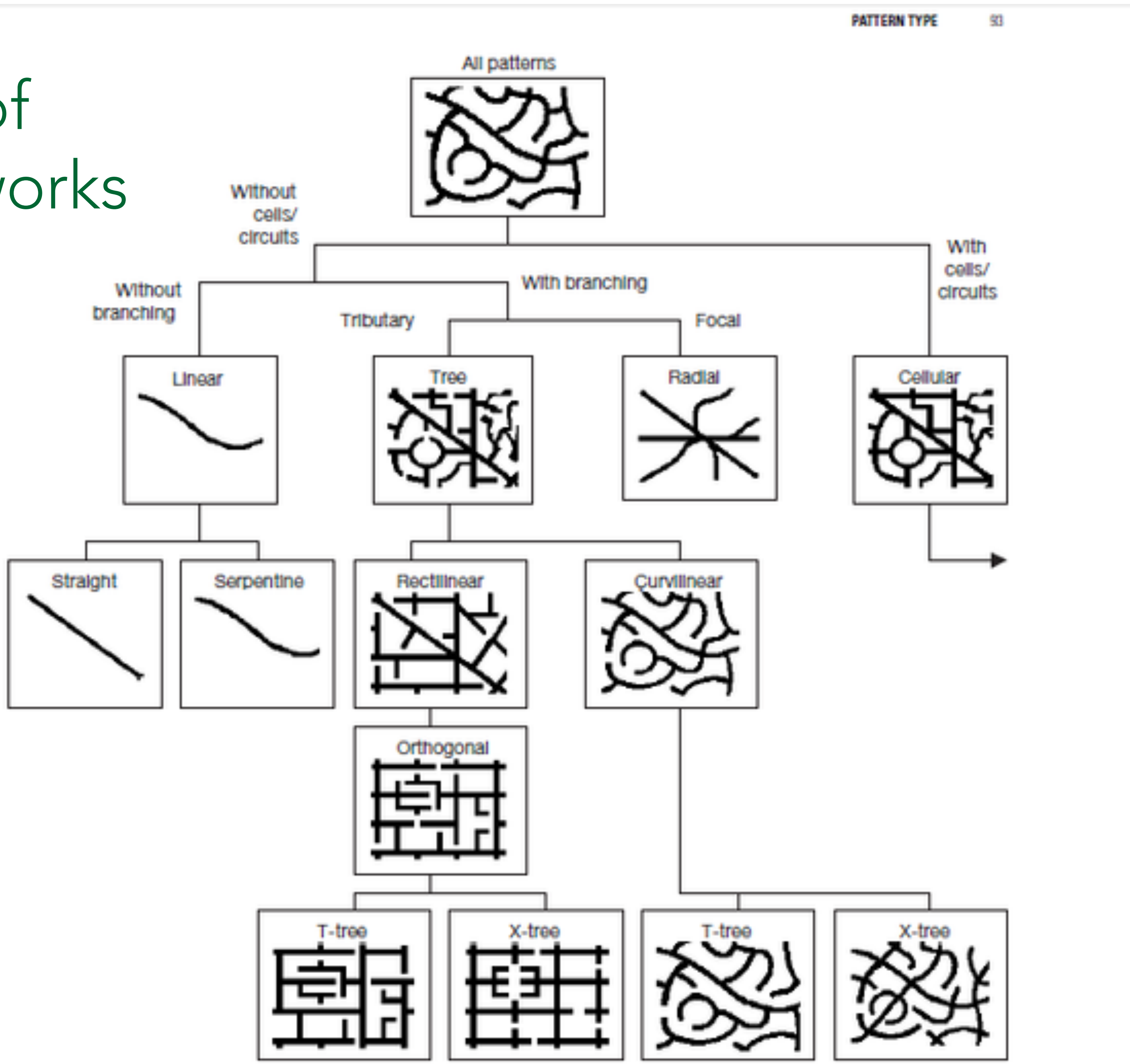


6.1 • Street patterns.

S. Marshall (2005)

Typology of street networks

Classification Attempt

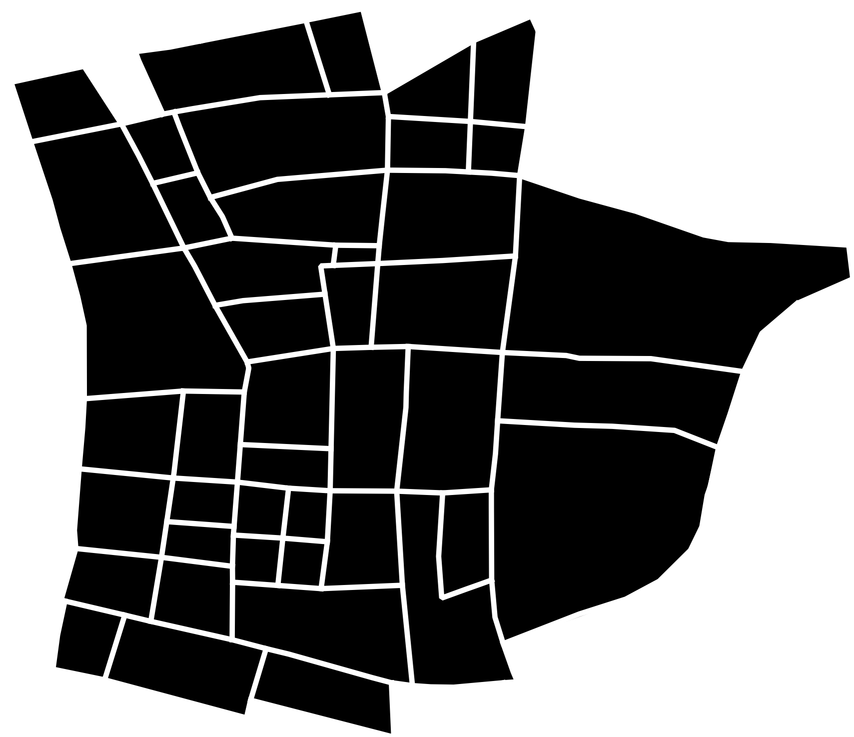


4.14 • Elaborated taxonomy of patterns.

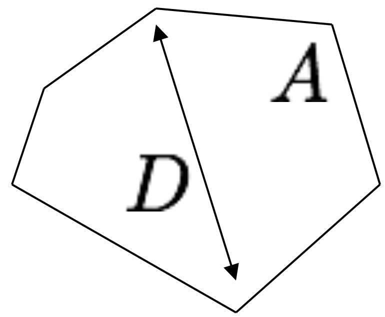
S. Marshall (2005)

Typology of planar graphs

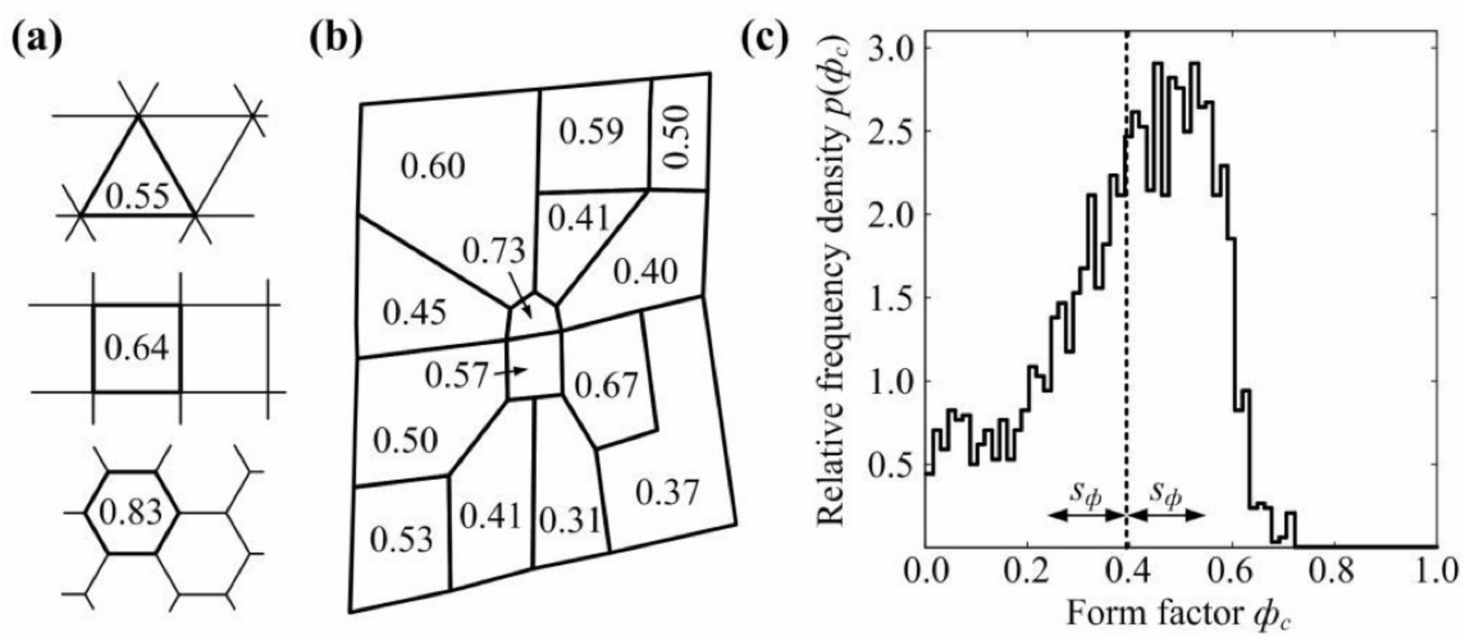
From the graph to the statistics of blocks



Shape of blocks



$$\phi = \frac{A}{(\pi D^2/4)}$$



Lammer et al, Physica A (2006)

Area of blocks

$$P(A) \sim A^{-\tau}$$

$$\tau \simeq 2.0$$

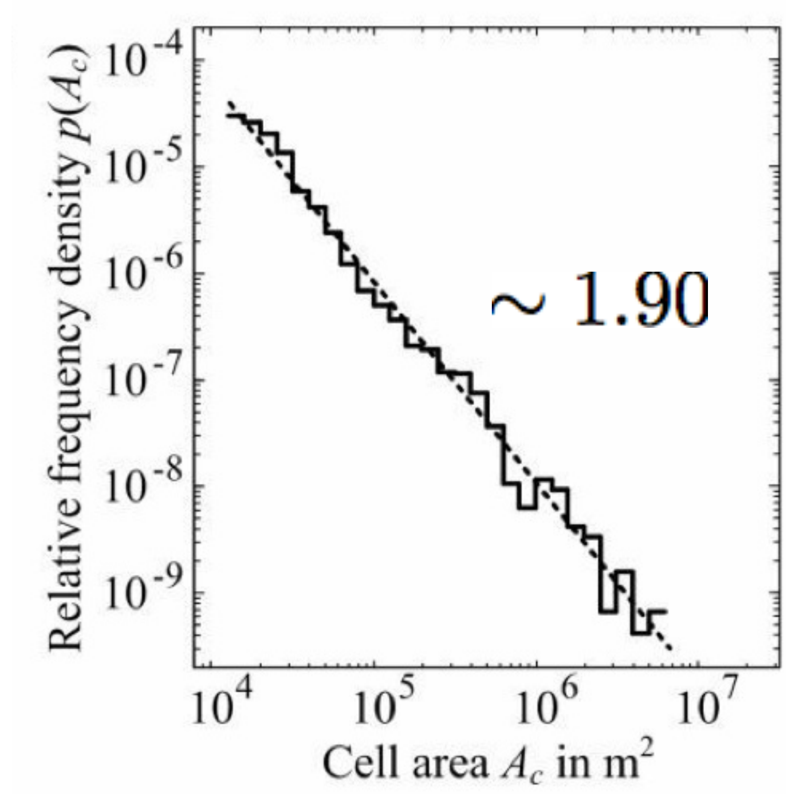
Simple argument:

$$A \sim \ell_1^2$$

$$\ell_1^2 \sim 1/\rho$$

Density distributed $f(\rho)$

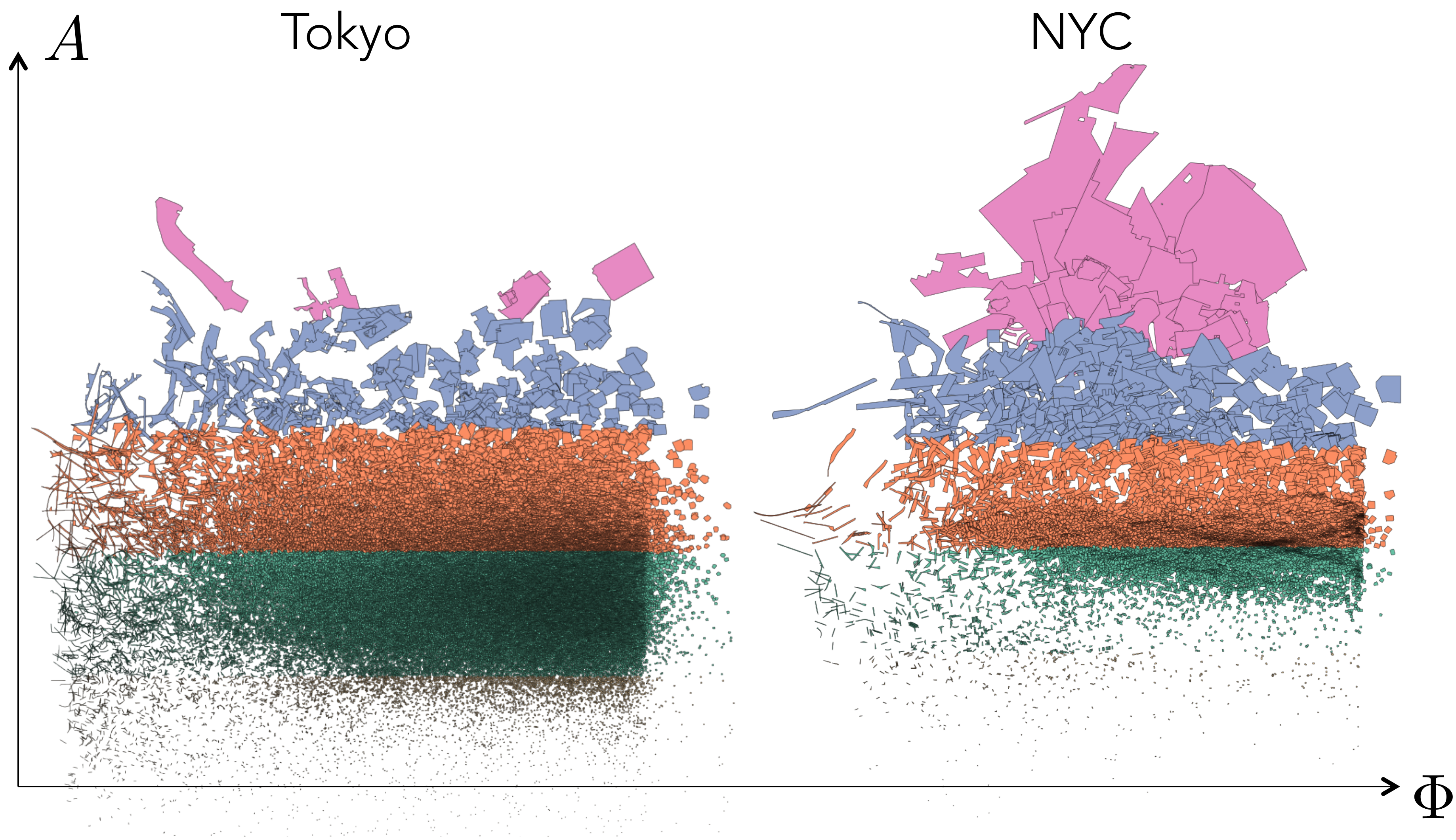
$$P(A) \sim \frac{1}{A^2} f(1/A)$$



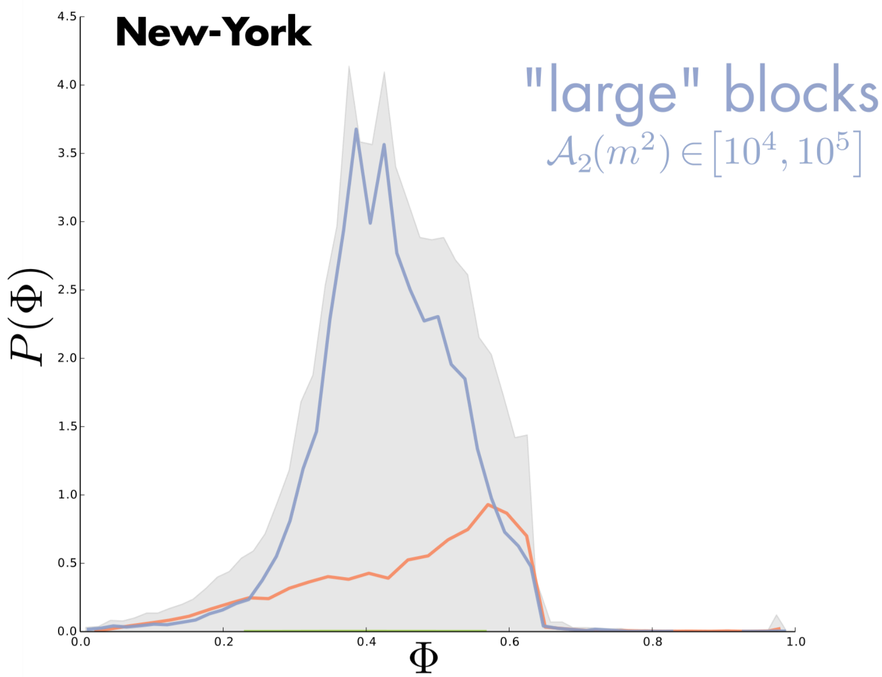
- Lammer et al, Physica A (2006)
- MB, Physics Reports (2011)

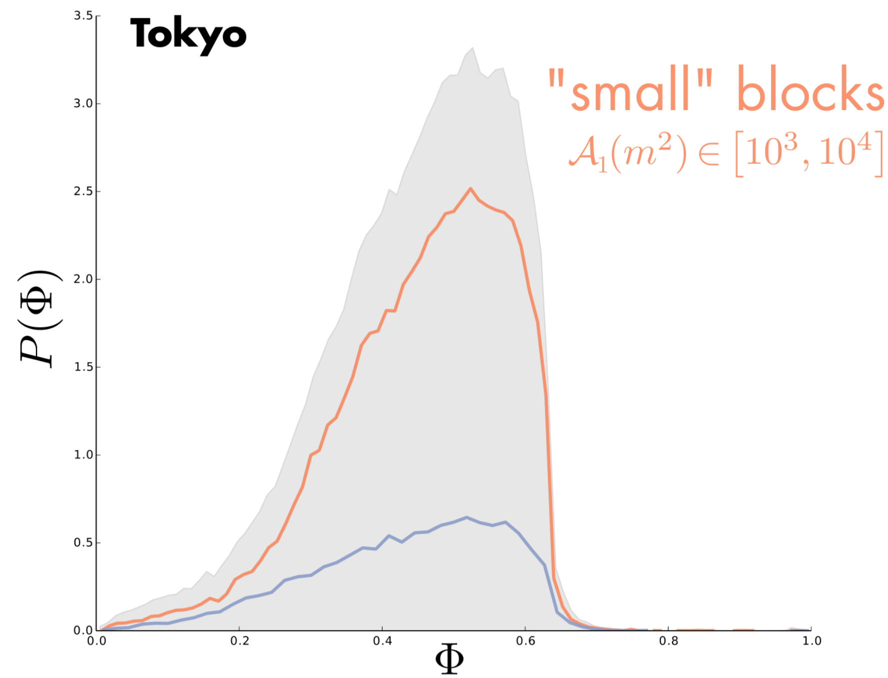
Shape and area distributions

Wide distribution of shapes and areas



Combining shape and area of blocks: a "fingerprint" of planar graphs: $P(\Phi|A)$

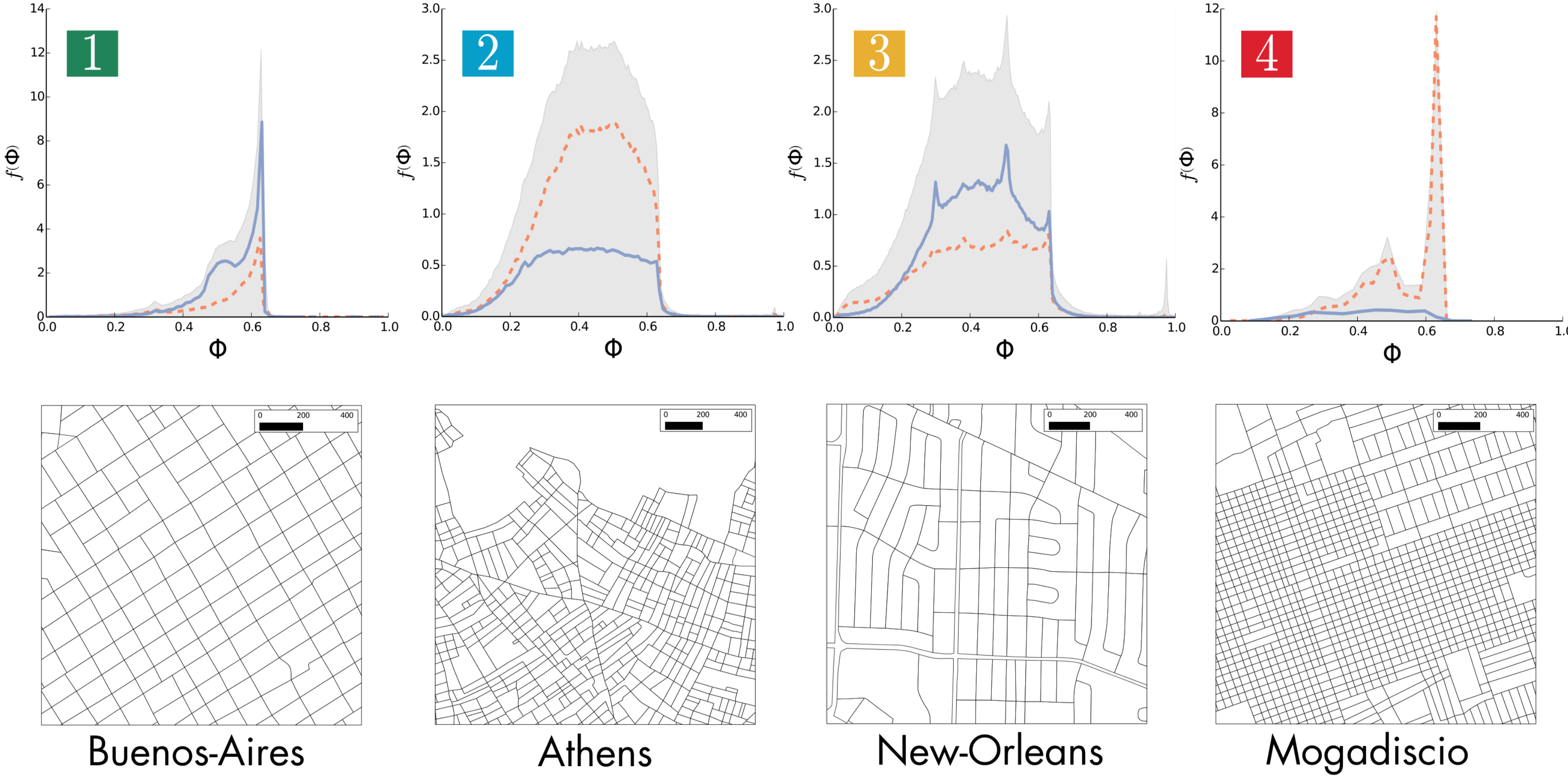




Louf & MB, RS Interface (2014)

Typology of street patterns

- Distance constructed on $P(\Phi|A)$
- Clustering -> classes of planar graphs (131 cities)



Louf & MB, RS Interface (2014)

Discussion

- New results on new datasets usually imply to have new tools !
- Evolution of planar graphs
  - Simple measures usually not very helpful
  - Important structural changes: betweenness centraliy distribution
  - Use of templates
  - Better characterization ?
  - Models ?
- Typology
  - Attempt to classify planar maps
  - Taking into account both topological and geometrical features
  - Correlations ?
- … we could use some help !

Thank you for your attention.

www.quanturb.com
marc.barthelemy@cea.fr

Students and Postdocs:

Giulia Carra (PhD student)
Riccardo Gallotti (Postdoc)
Thomas Louail (Postdoc)
Remi Louf (PhD student)
Emanuele Strano (PhD student)

Collaborators:

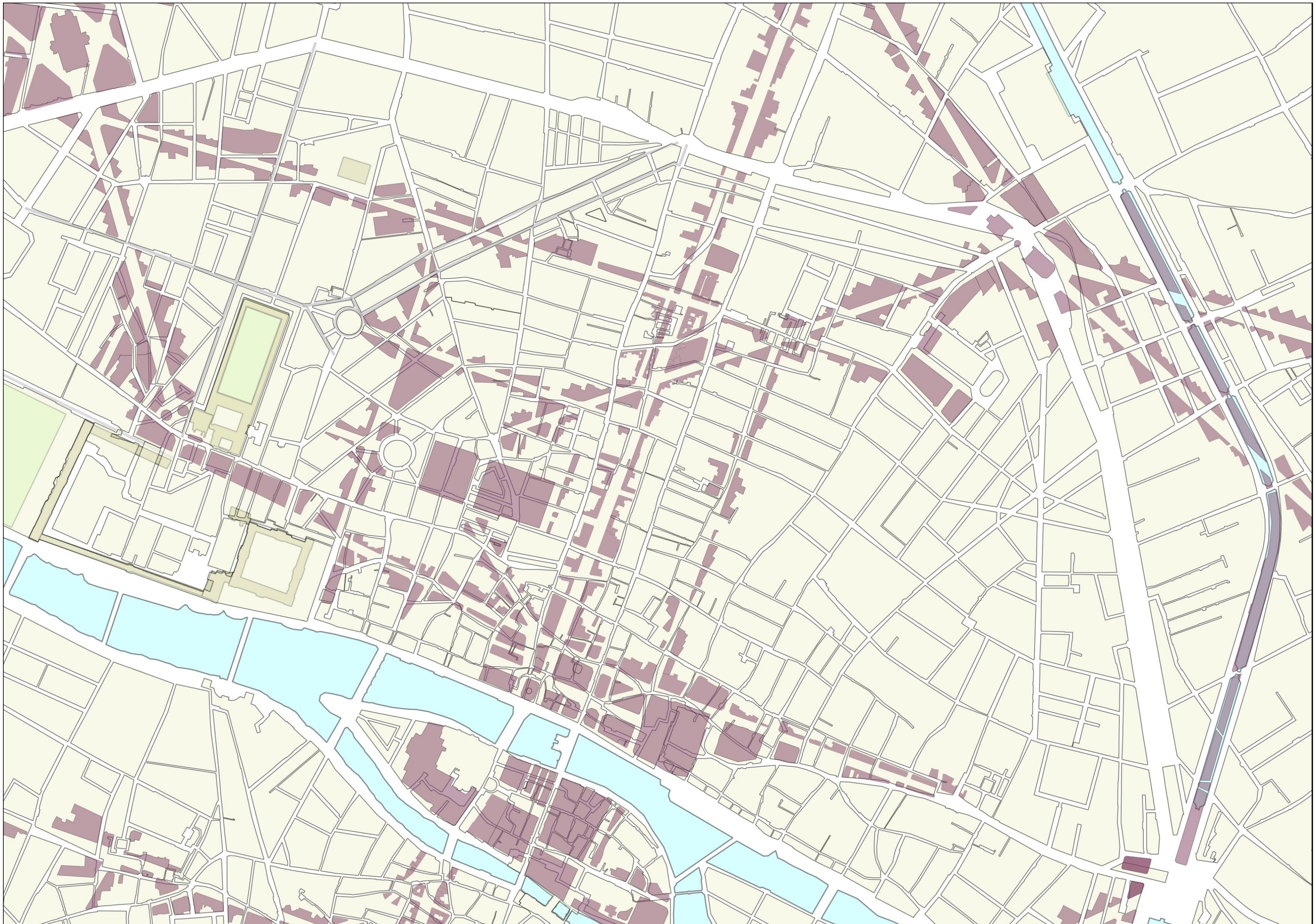
M. Batty	H. Berestycki	P. Bordin
S. Dobson	M. Gribaudi	P. Jensen
JP. Nadal	V. Nicosia	V. Latora
J. Perret	S. Porta	C. Roth
S. Shay	MP. Viana	

Funding:

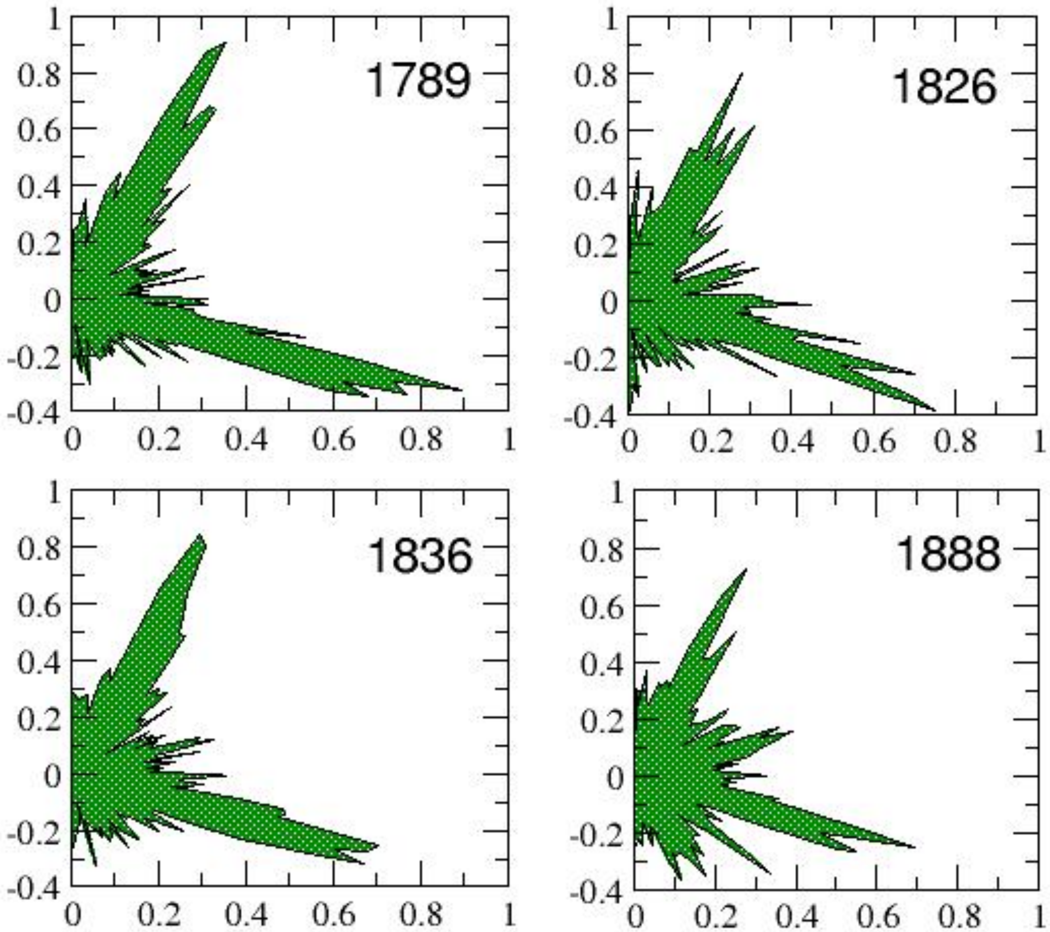
EUNOIA (FP7-DG.Connect-318367 European Commission)
PLEXMATH (FP7-ICT-2011-8 European Commission)

Additional slides

Importance of central planning

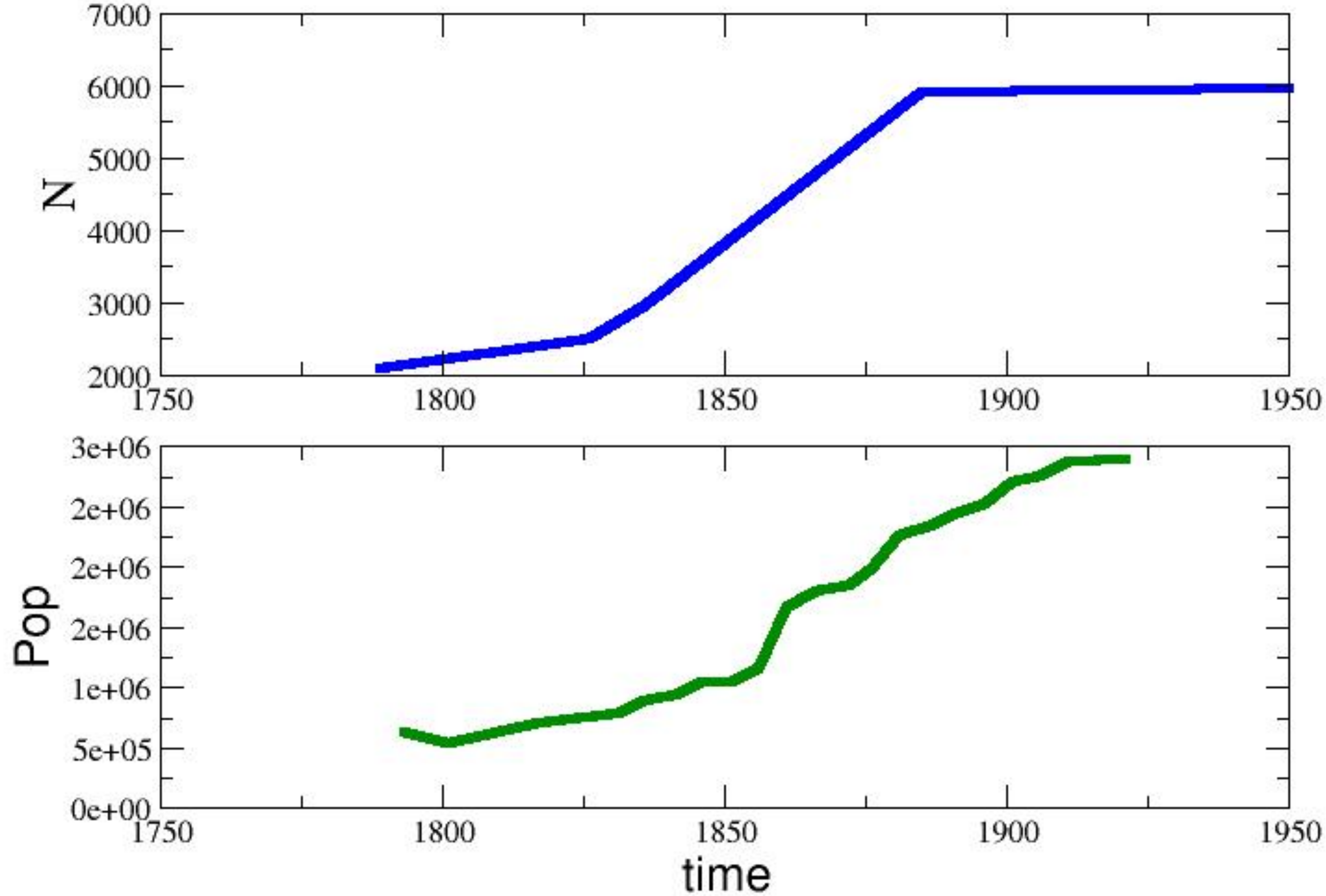


Haussmann effect: angle distribution



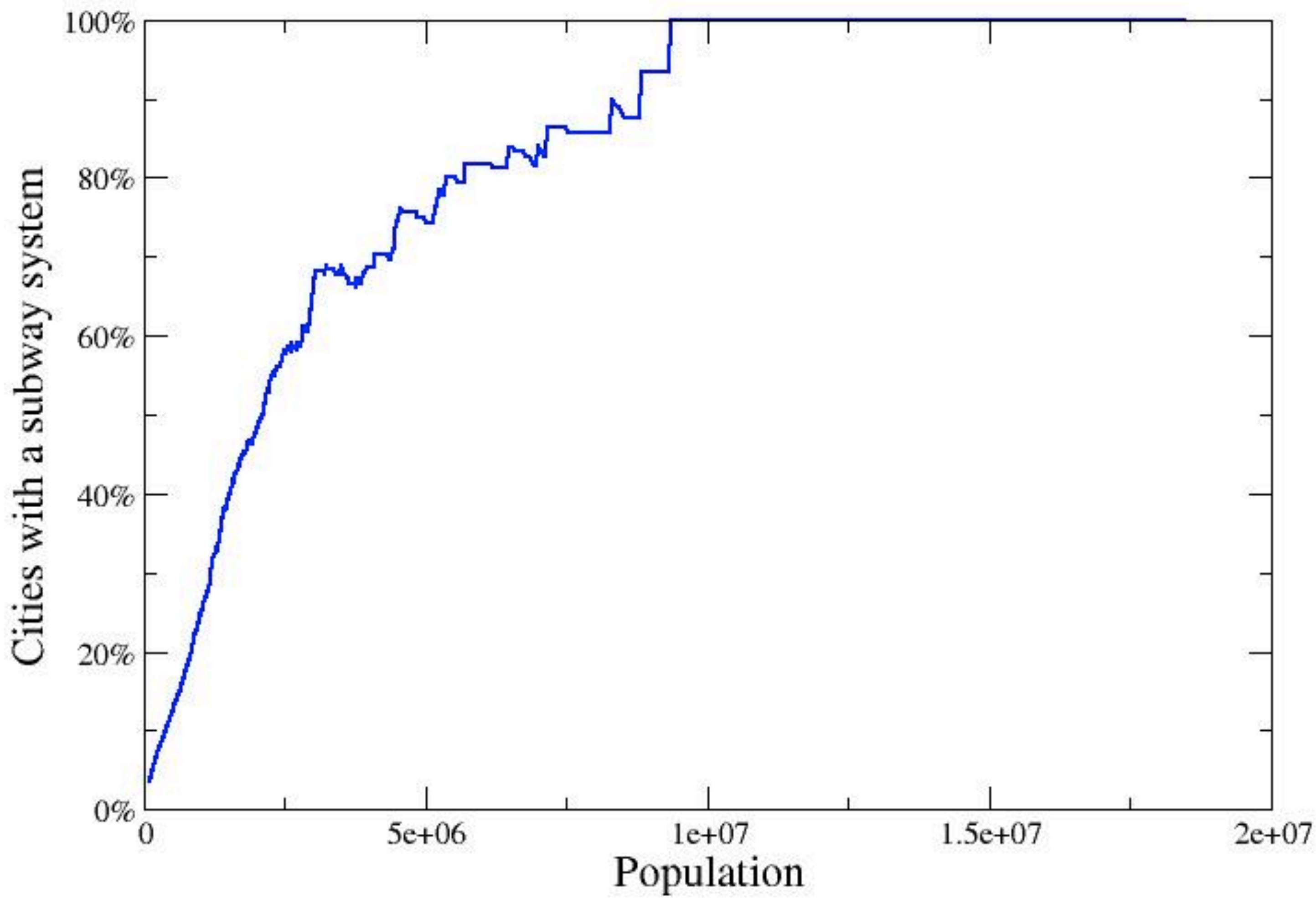
Importance of central planning

- N follows the population evolution



- Most indicators follow a smooth evolution…

All large cities have a subway system



Evolution fraction of branches stations

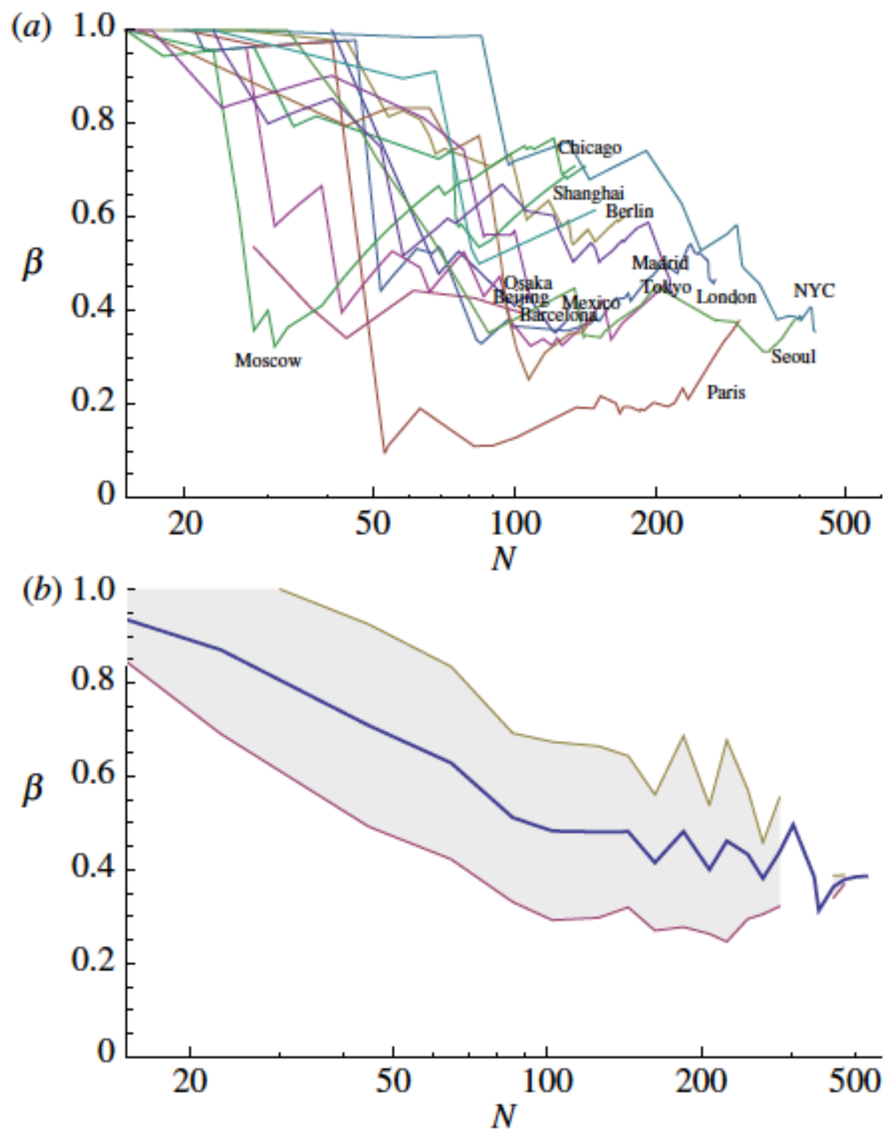


Figure 6. (a) Parameter β as a function of the number of stations N for the different world subways. (b) Same as (a) but averaged over 20 bins and showing the standard deviation. (Online version in colour.)

Average degree
Percentage f_2

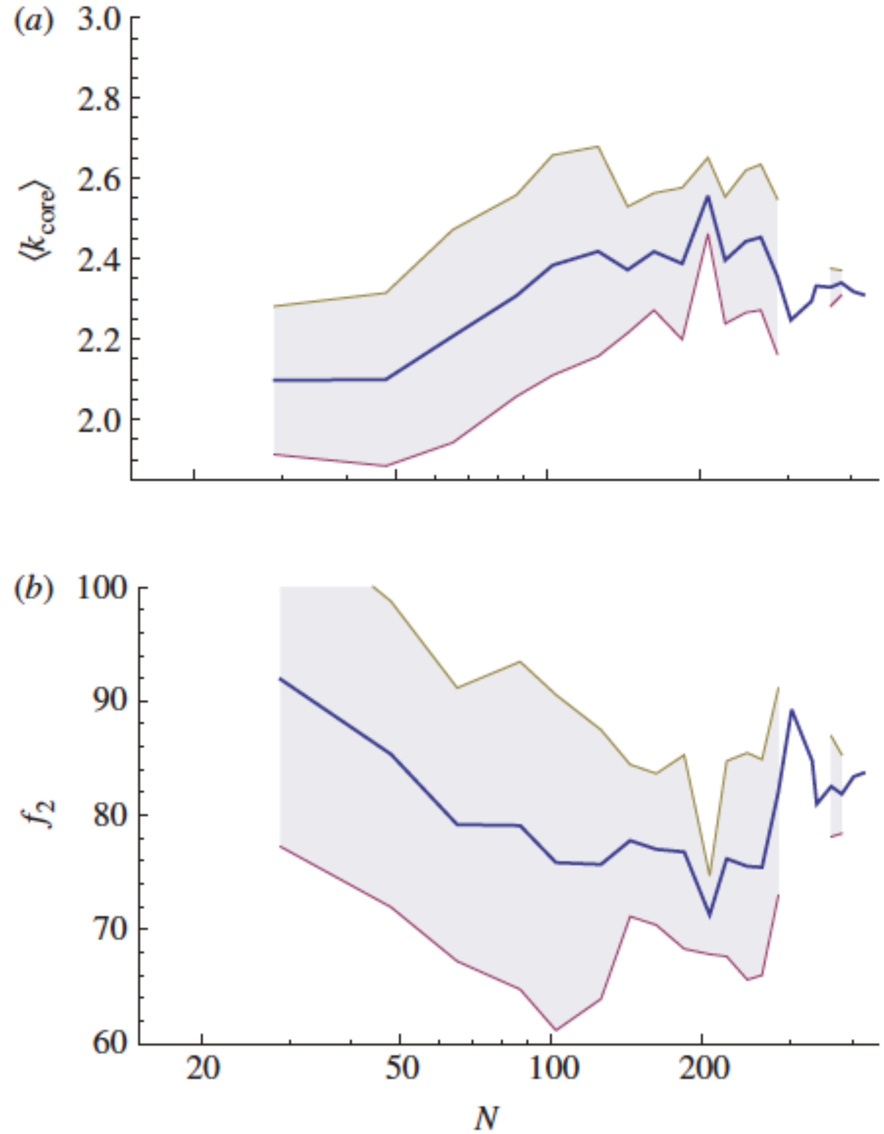


Figure 8. (a) Average degree of the core $\langle k_{\text{core}} \rangle$ (equation (3.3)) and its dispersion versus number of stations (averaged over 20 bins). (b) Evolution of the percentage f_2 of $k = 2$ core nodes (averaged over 20 bins). (Online version in colour.)

Spatial extension of branches

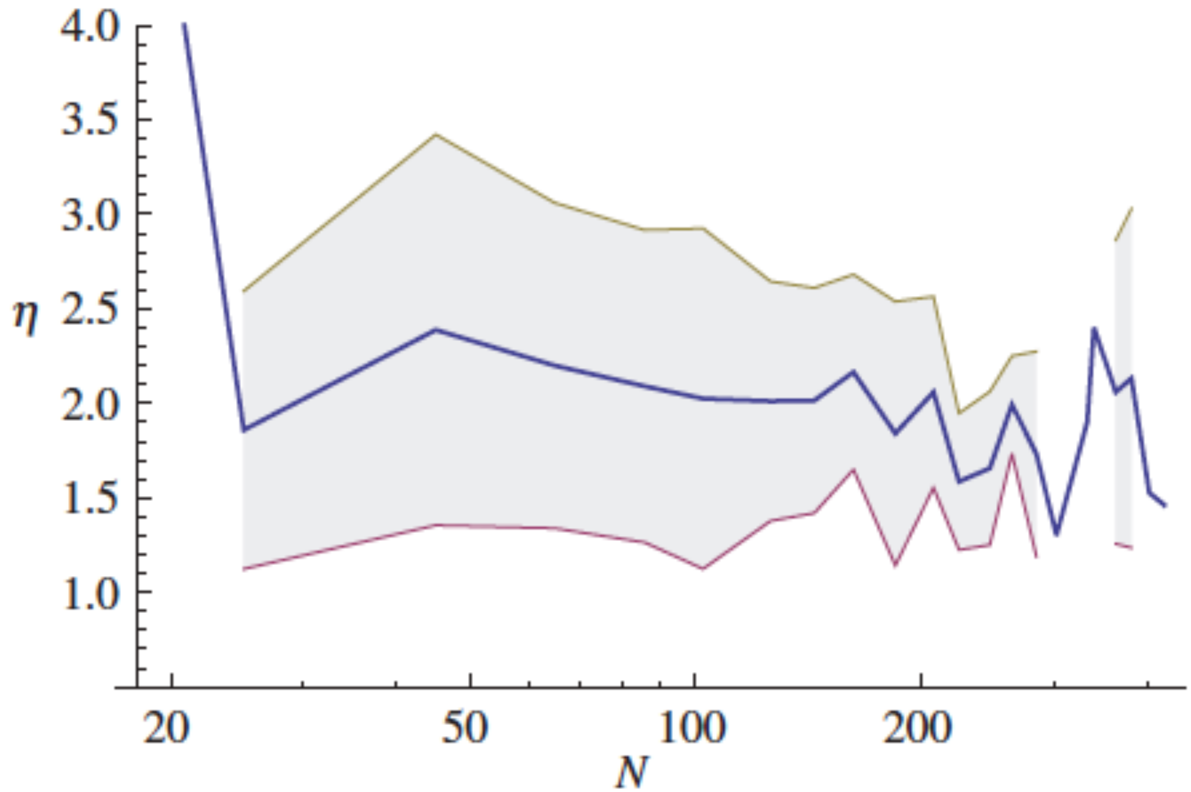


Figure 7. Evolution of the ratio η, which characterizes the spatial extension of branches relative to the core. (Online version in colour.)

Spatial organization of the core and branches

- Old result for Paris (Benguigui, Daoud 1991)

$$\text{Short scale: } N(r) \sim r^2$$
$$\text{Long scale: } N(r) \sim r^{0.5}$$

N(r): number of stations at distance less than r from barycenter

First regime: homogeneous distribution with d_f=2
Second regime ?

Spatial organization of the core and branches

- Natural explanation with the universal template

$$N(r) = \begin{cases} \rho_C \pi r^2 & \text{for } \; r \ll r_C \\ \rho_C \pi r_C^2 + N_b \int_{r_C}^{r} \frac{dr}{\Delta(r)} & \text{for } r_C \ll r < r_{max} \end{cases}$$

ρ_C : core density
N_b : number of branches
$\Delta(r)$: Interstation spacing at distance r

Spatial organization of the core and branches

- Interstation spacing at distance r

$$\Delta(r) \sim r^{\tau} \Rightarrow N(r \gg r_C) \sim r^{1-\tau}$$

$$\tau \simeq 0.05 \ (\mathrm{Moscow})$$

$$\tau \simeq 0.5 \ (\mathrm{Paris})$$

- Natural explanation of the Benguigui-Daoud result

An old problem in quantitative geography

- Kansky (1963-69)

Evolution of the Sicilian Road network

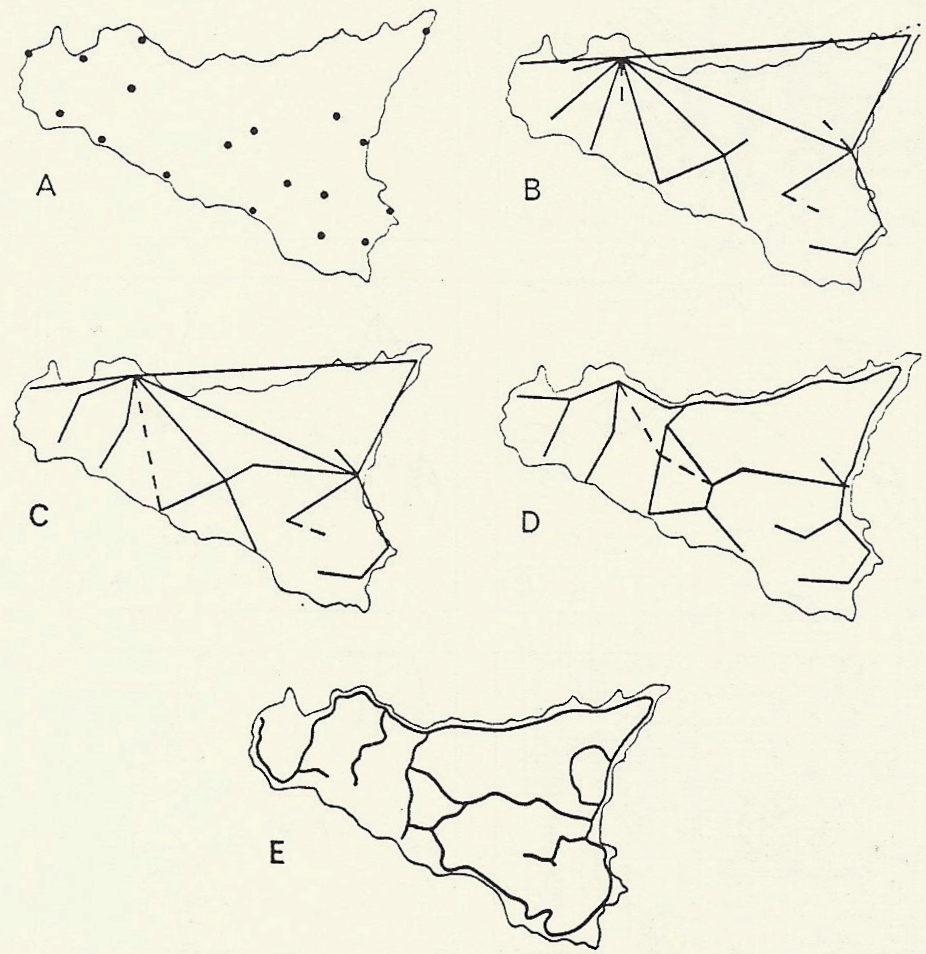
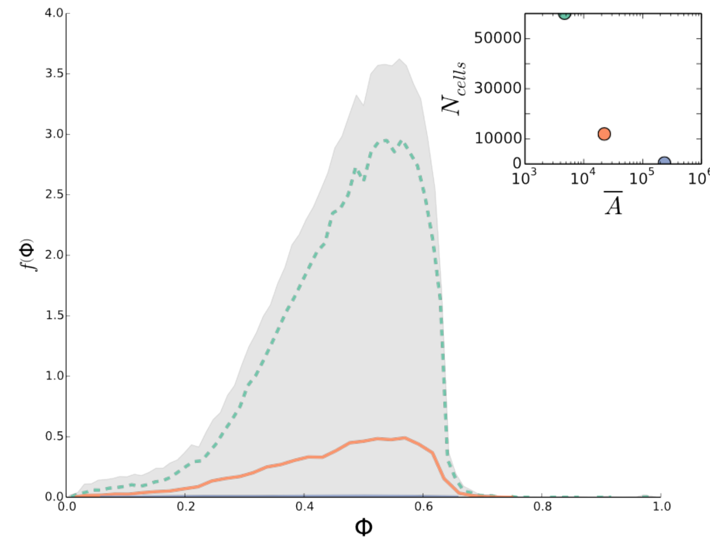


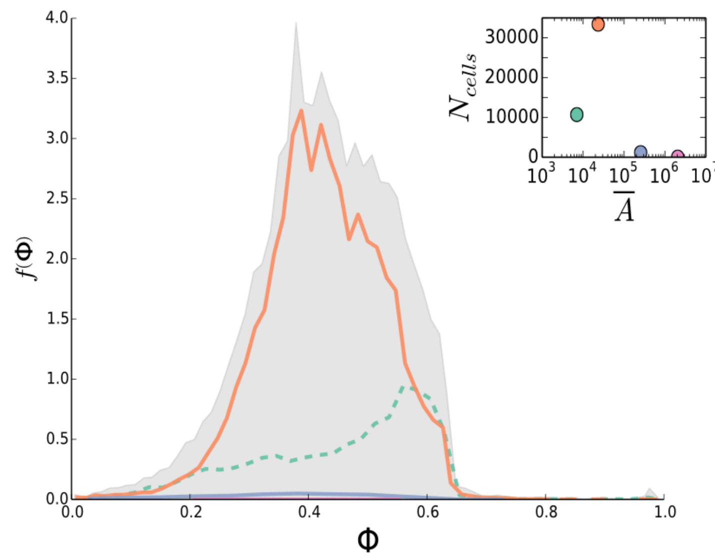
Fig. 5.29. Stages in link allocation for simulating the Sicilian 1908 railroad network. (*A*) Selected vertices. (*B–C*) Stages in link allocation. (*D*) Postdicted 1908 network. (*E*) Actual 1908 network. Source: Kansky, 1963, pp. 139–46.

A “fingerprint” of planar graphs

Tokyo

NYC





Evolution of road networks: Summary

- ‘Natural’ case:
  - Two elementary processes (densification and exploration)
  - Long time scale, short spatial scale
  - Perspective: modeling (fragmentation models ?)

- Central planning:
  - Does not respect the previous geometry
  - Acts on the spatial structure of flows
  - Short time scale and large spatial scale.
  - Possibility of modeling ?