COMPLEX SYSTEMS, NETWORKS AND SPACE

Cazabet Rémy

gopro2017 : Spatial Optimized Designs: from Cities to Natural systems

WHO AM I

- Maître de conférences (since September) in Computer Science (LIRIS, Lyon 1)

- Traditional topics:
  - "Network Science"
    - Network Analysis, Network Mining, "Community Detection", Dynamic Networks, etc.
  - "Data Science"
    - Learning knowledge from data
  - "Complex Systems"
    - Systems composed of multiple parts in interaction, non-linear behaviour, cannot be studied by reductionism:
    - Interactions between entities => Network

CITIES ? NATURAL SYSTEMS ?

- Cities and natural systems are complex systems
- 2015-2016: Working on Vél'innov ANR Project: understanding, characterising activity in Bike Sharing Systems (BSS)
- Currently: starting a collaboration with Claire Lesieur on the organisation of proteins

ORGANISATION OF COMPLEX SYSTEMS

- Usually, there is not **one** network of a complex system:
  - In cities:
    - Network of proximity between buildings
    - Network of trips using public transportations
    - Network of trips using bicycle
    - Network of roads
    - Network of socio-demographic similarities
    - Network of phone calls between neighbourhoods
    - …
  - Each dataset can be modelled by countless networks, using thresholds, temporal aggregations, etc.
  - I'm not working on a network in particular, just networks as models of interactions inside complex systems (Complex networks ?)

ORGANISATION OF COMPLEX SYSTEMS

- Complex Networks are not random. They have a particular **organisation**.
- My ultimate goal is to understand this organisation.
  - Find underlying rules explaining difference between observed and random networks
  - **Spatial organisation** is one potential **candidate** to explain the structure of networks

NETWORK & SPACE

- What is a spatial network model:
  - Nodes are characterised by a position, i.e an x-dimensional vector
  - The probability of observing an edge depends on the distance between nodes

NETWORK & SPACE

- Networks and spatial organisation have a complex history
  - First network models were often spatial-like:
    - Regular grids (nodes on a grid, edges at fix distance)
    - Watts-strogatz (nodes on a circle, most edges depends on distance)
  - Later models often have no spatial structure
    - Community-based
    - Dynamic models (preferential attachment, forest fire, …)
  - Most network representations are based on 2D projections
  - The come back of spatial organisation: network embedding.
    - Given a network, which position of nodes better explain its organisation ?

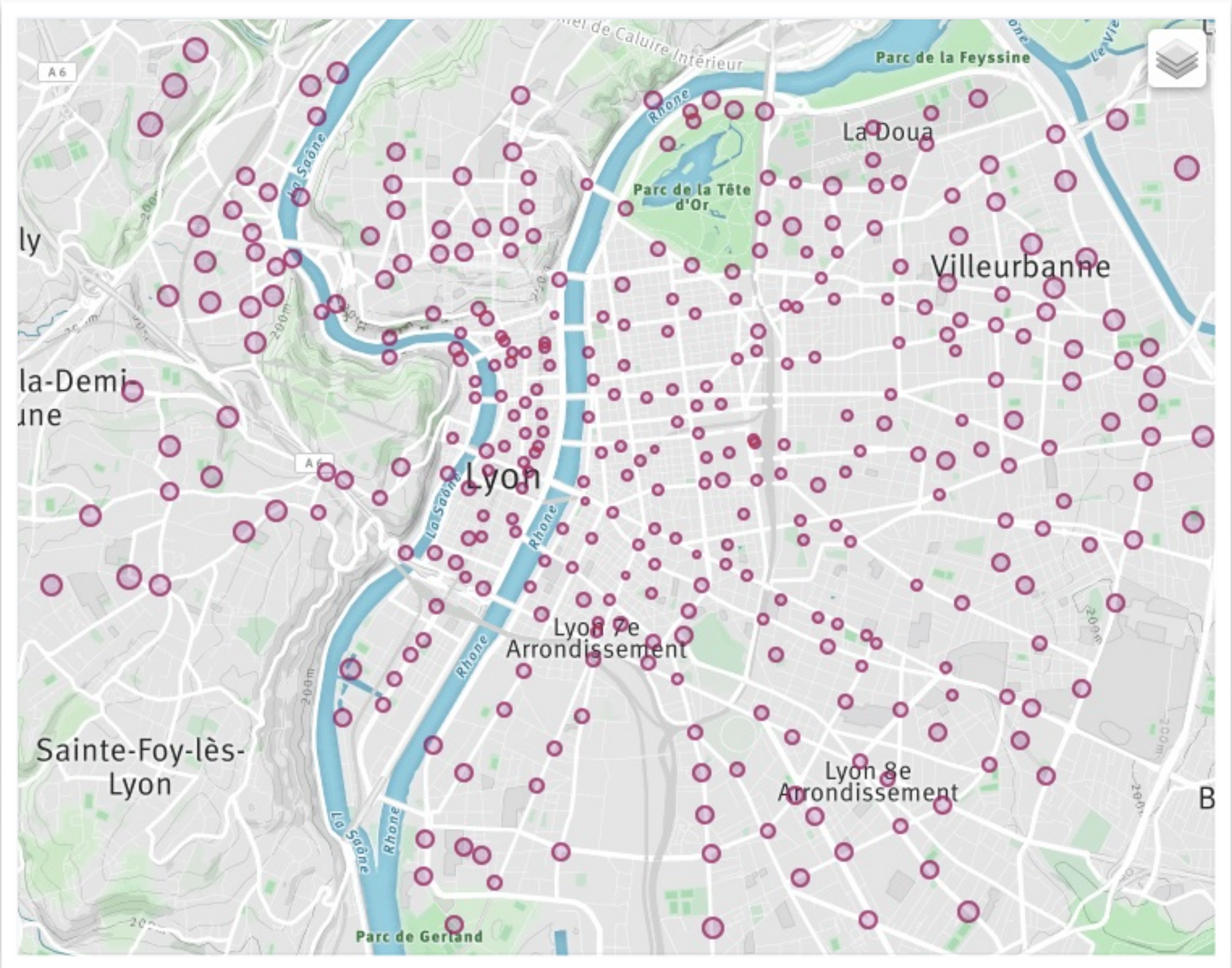
AN EXAMPLE: VÉLO'V

VÉLO'V

Bicycle Sharing System (BSS) in Lyon

Dataset: trips (5y) + sociodemographic around stations

VÉLO'V



Nodes: station (2D position)

Edges: number of trips over a period

NETWORK MODEL 1

- Random network
- #trips between any pair of station is the same

$$p_{ij}^{RR} = \frac{l}{m}$$

Model complexity:

1

Model precision:

+

Evaluation of the model :
diff between observed network and model

NETWORK MODEL 2

- Configuration model
- #trips between any pair of station depends on their "popularity"

$$p_{ij}^{Conf} = k_i k_j W$$

Model complexity:
n

Model precision:
++

NETWORK MODEL 3

- Simple Gravity
- #trips between any pair of station depends on their "popularity" and their distance

$$p_{ij}^{Grav2} = W \frac{k_i k_j}{d_{ij}^2}$$

Model complexity:
n+2n

Model precision:
++++

NETWORK MODEL 4

- Gravity with custom deterrence function
- #trips between any pair of station depends on their "popularity" and their distance.
- Distance influence learnt from data

Model complexity:
n+2n+a

Model precision:
++++

$$p_{ij}^{Grav2} = W k_i k_j f(d_{ij})$$

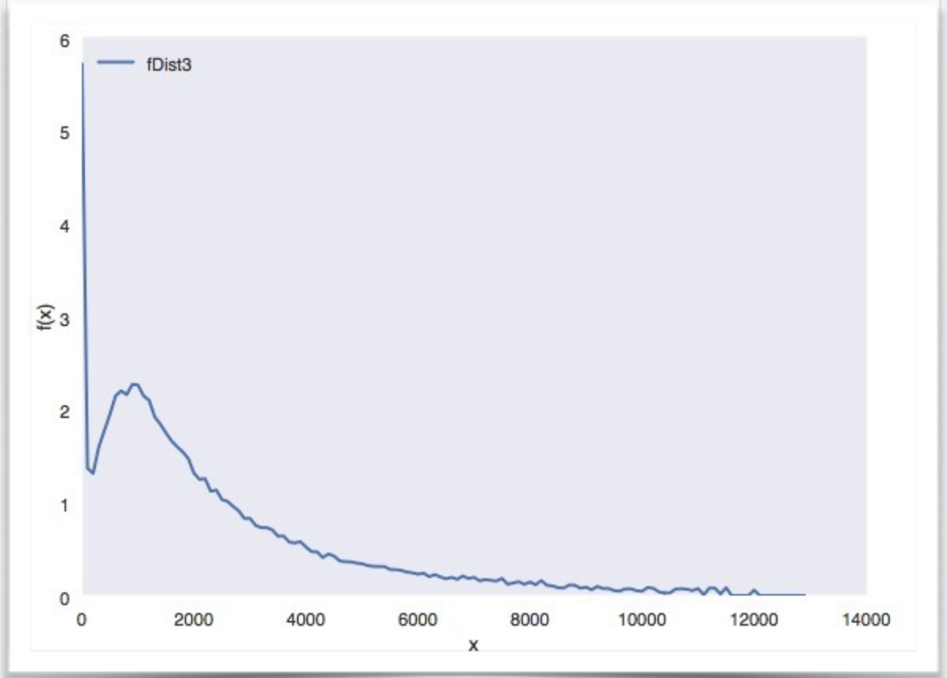
DETERRENCE FUNCTION

Computation of a deterrence function:
Impact of distance on edge probability

(Comparing observation with Configuration Model)

$$f(d) = \frac{\sum_{i,j|d_{ij}=d} A_{ij}}{\sum_{i,j|d_{ij}=d} N_i N_j}$$

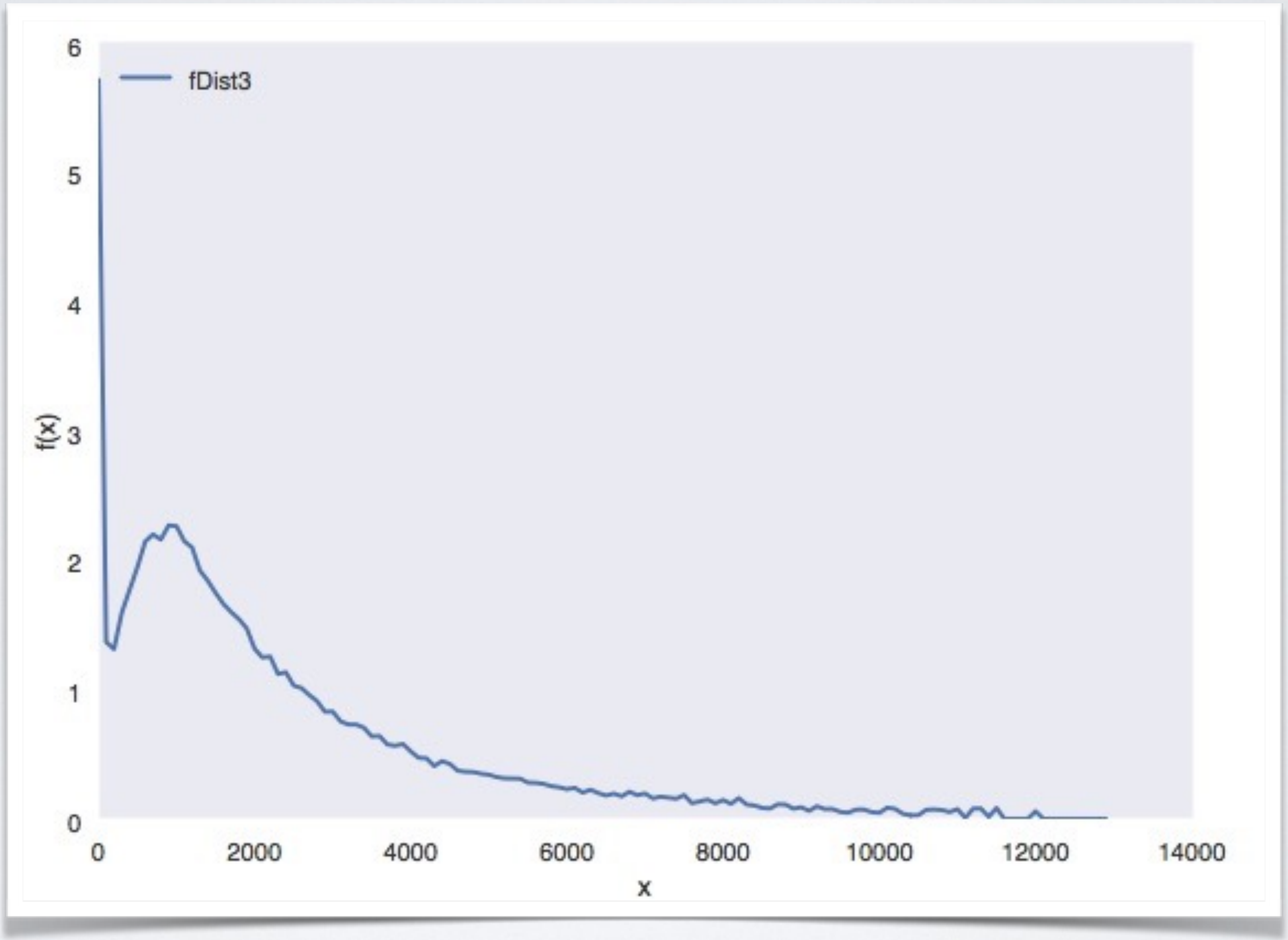
f(d)



Distance *d (meters)*

DETERRENCE FUNCTION

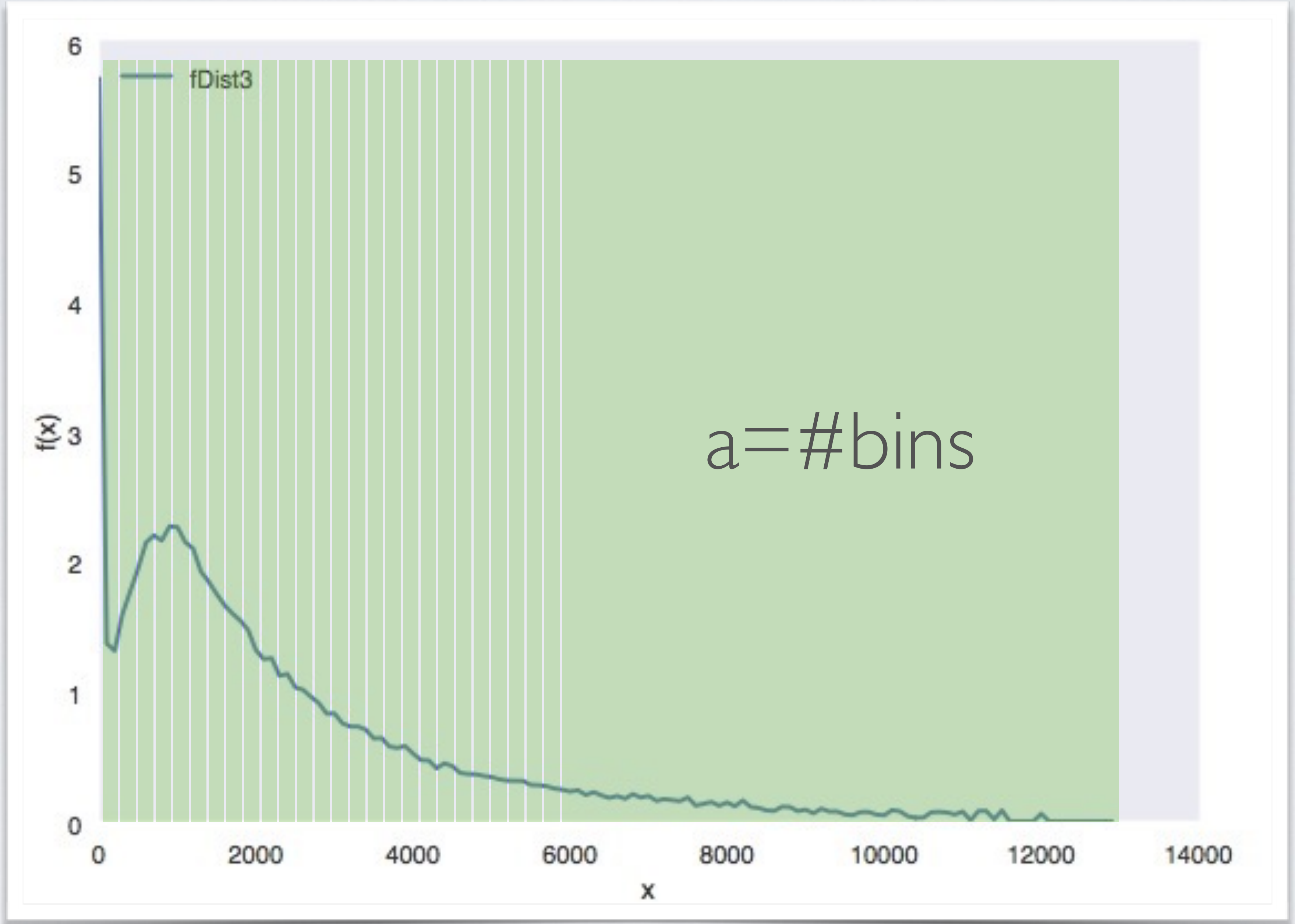
f(d)



Distance *d* *(meters)*

DETERRENCE FUNCTION

f(d)



Distance *d* *(meters)*

NETWORK MODEL 4

- Gravity with custom deterrence function
- #trips between any pair of station depends on their "popularity" and their distance.
- Distance influence learnt from data

Model complexity:
n+2n+a

Model precision:
++++

$$p_{ij}^{Grav2} = W k_i k_j f(d_{ij})$$

NETWORK MODEL 5

- Gravity with custom deterrence function and conservation of degrees

- Same as before, but constraint to conserve node degrees

Model complexity:

n+2n+a

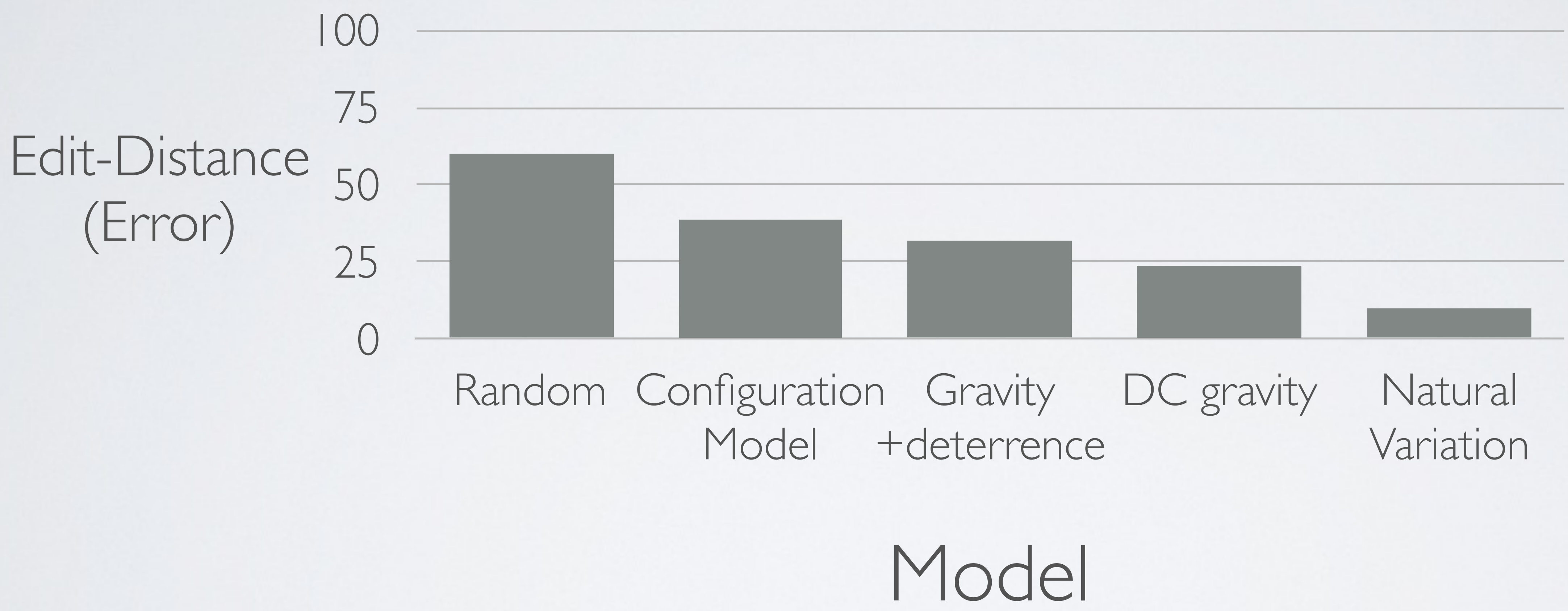
Model precision:

+++++

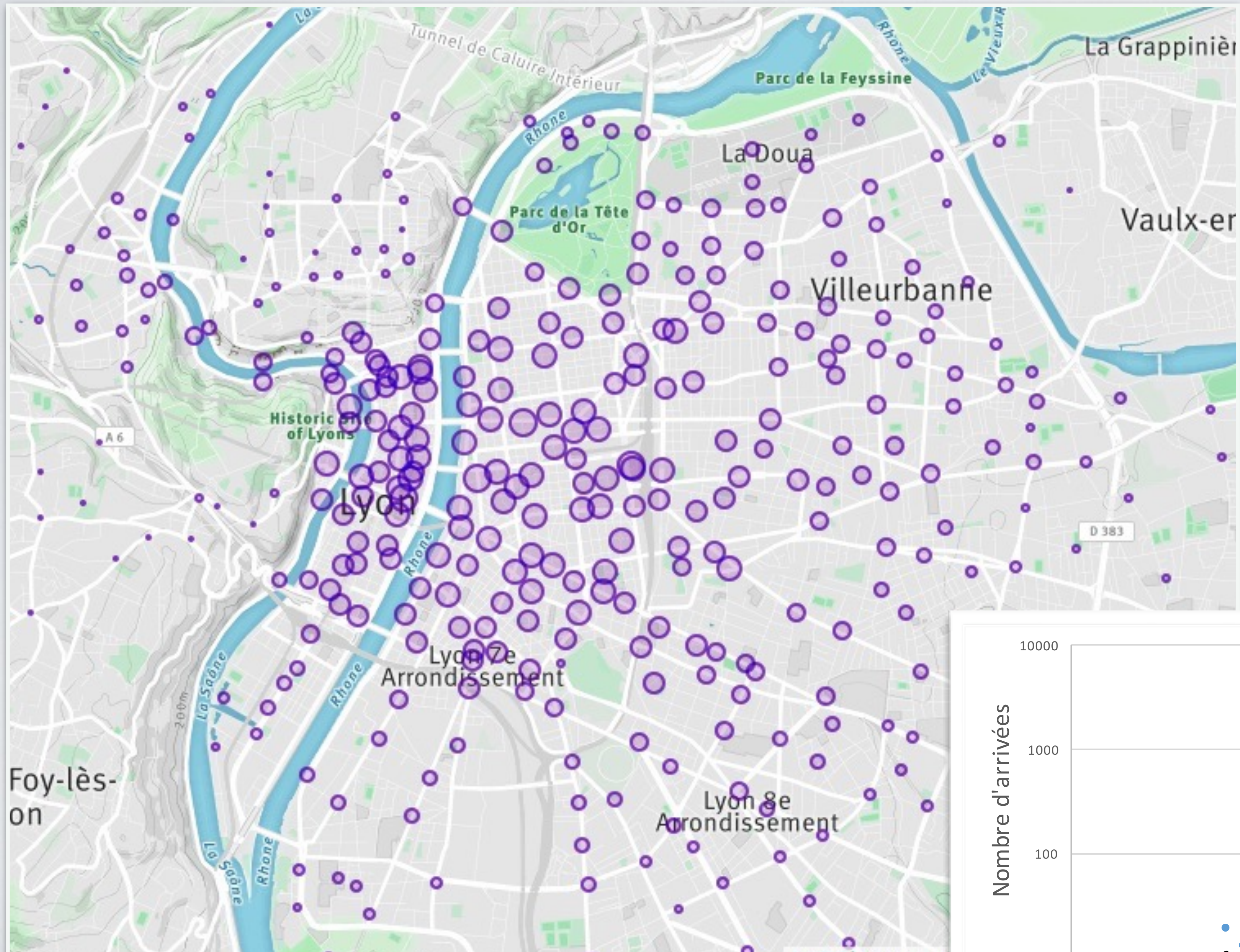
$$P_{ij}^{DCgrav} = n^{Oeis} n^{Ieis} f(d_{ij})$$

$$n^{Ieis} = \frac{deg^{out}(i)}{\sum_i n^{Oeis} f(d_{ij}))}, n^{Oeis} = \frac{deg^{in}(i)}{\sum_i n^{Ieis} f(d_{ij})}$$

MODEL EVALUATION

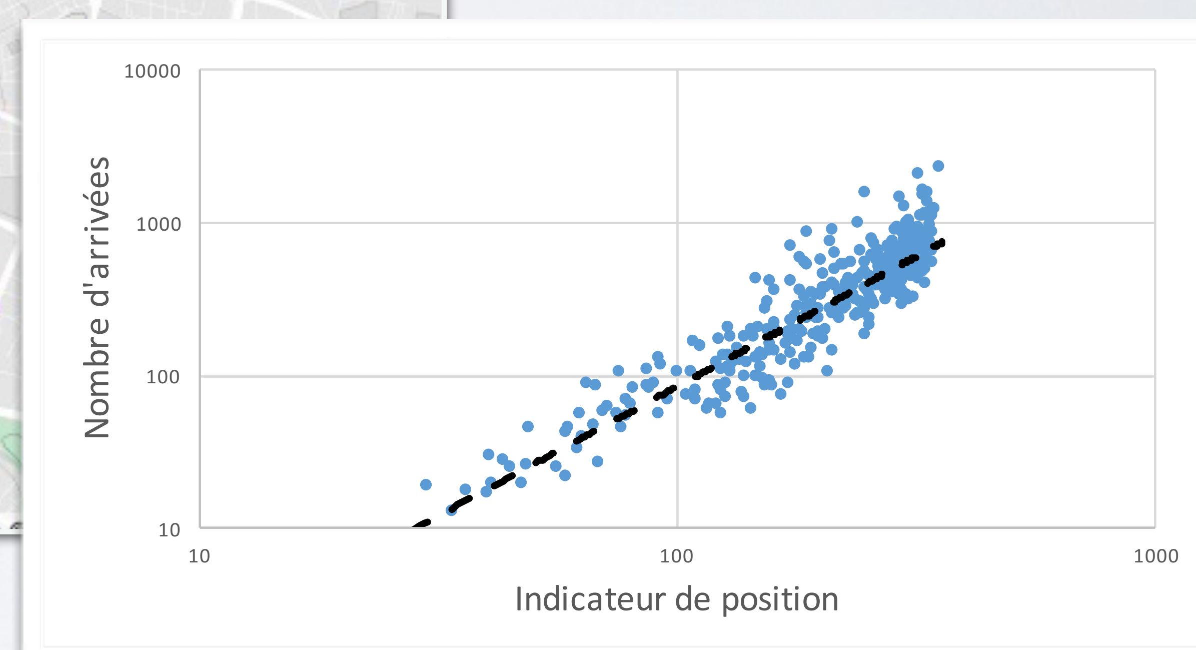


USEFUL MODEL ?



Geographic Potential Pi

$$P_i = \sum_j f(d(i,j))$$

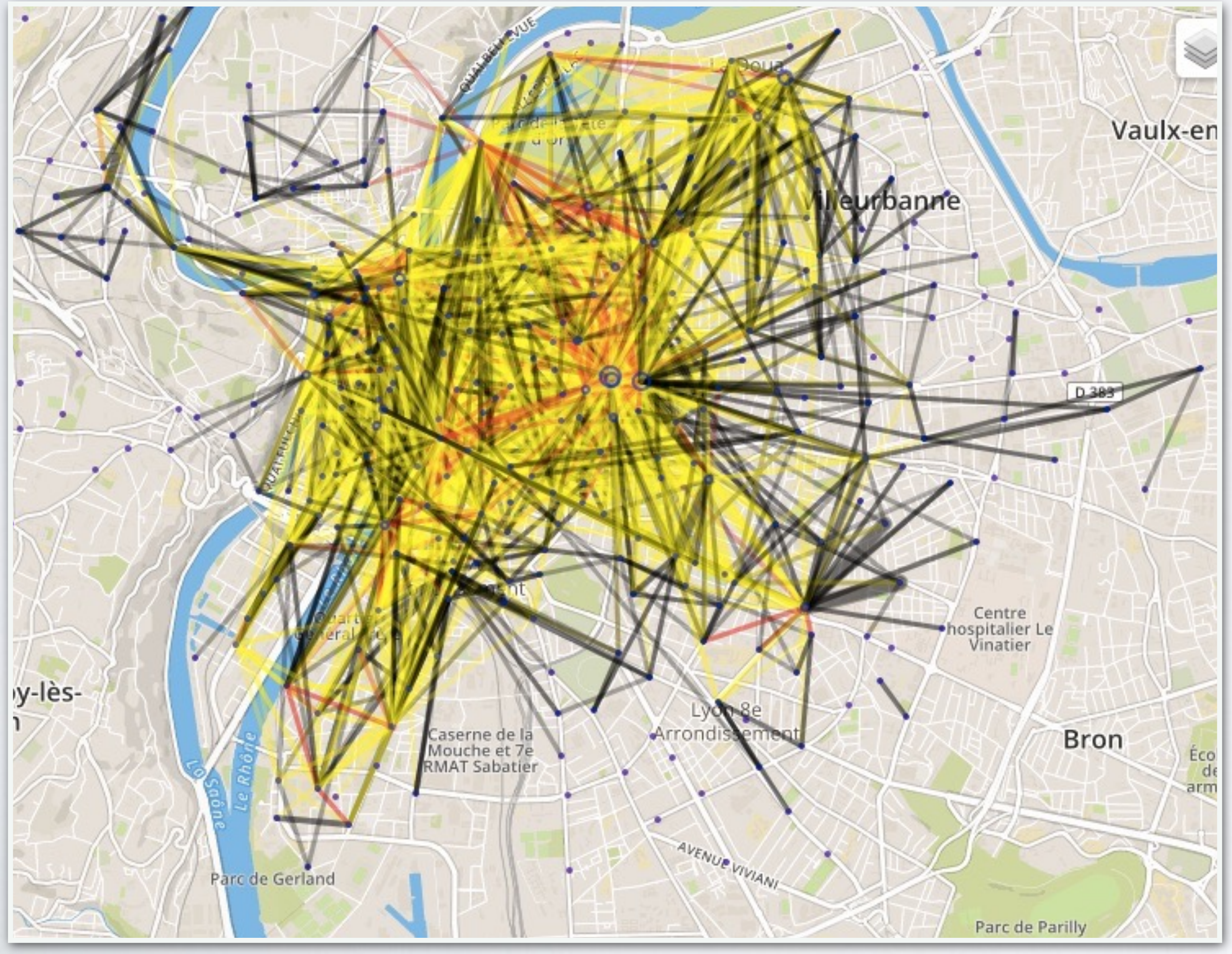


IMPROVING THE MODEL ?

- Difference between observed network and model. Random errors ?

IMPROVING THE MODEL ?

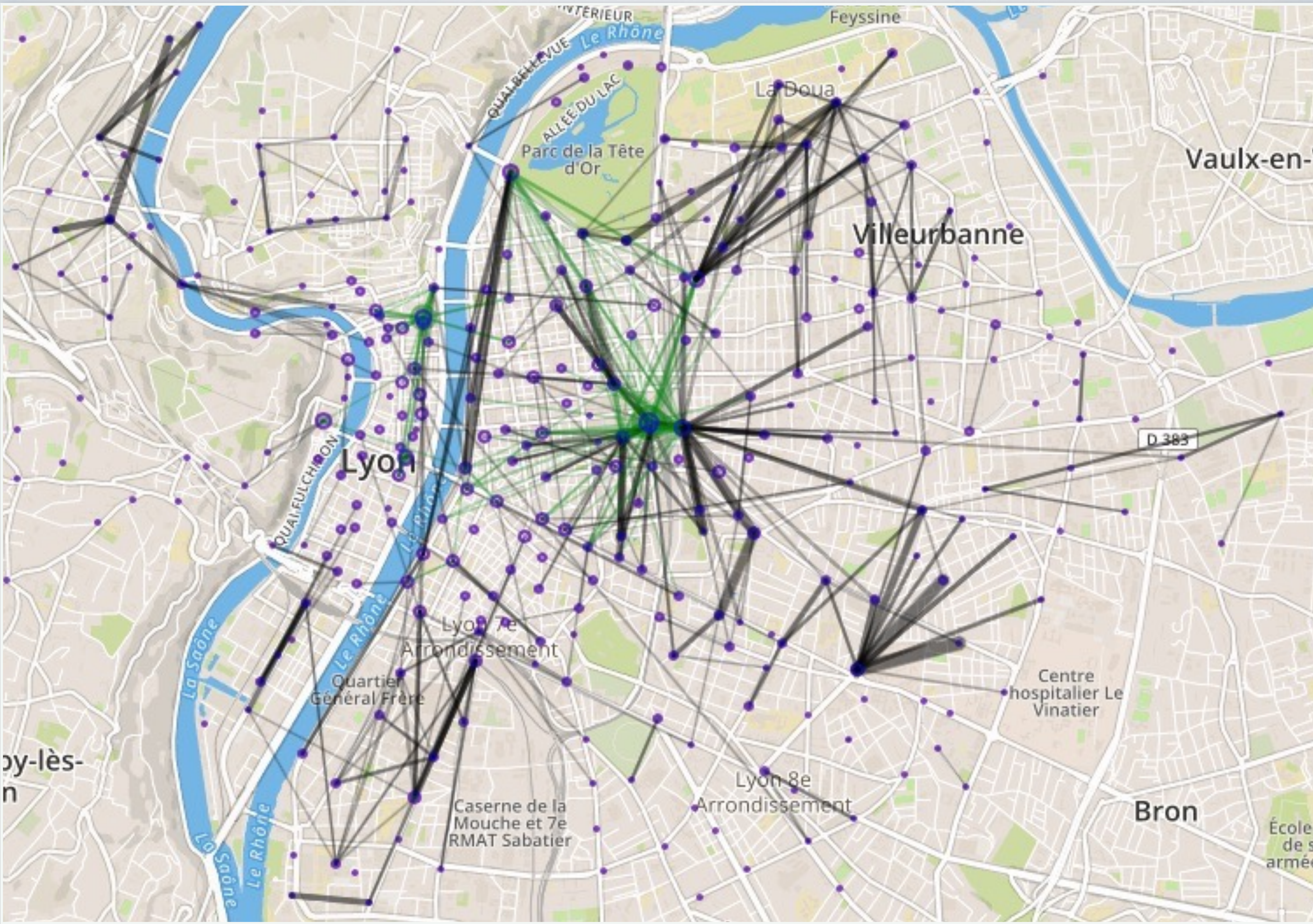
- Difference between observed network and model. Random errors ?



Yellow: well predicted

Red: overestimated

Black: Underestimated



Feyssine
Le Rhône
QUAI BELLEVUE
ALLÉE DU LAC
Parc de la Tête d'Or
La Doua
Vaulx-en-V
Villeurbanne
Lyon
QUAI FULCHIRON
D 383
Lyon 7e Arrondissement
Quartier Général Frère
Centre hospitalier Le Vinatier
La Saône
Le Rhône
Caserne de la Mouche et 7e RMAT Sabatier
Lyon 8e Arrondissement
Bron

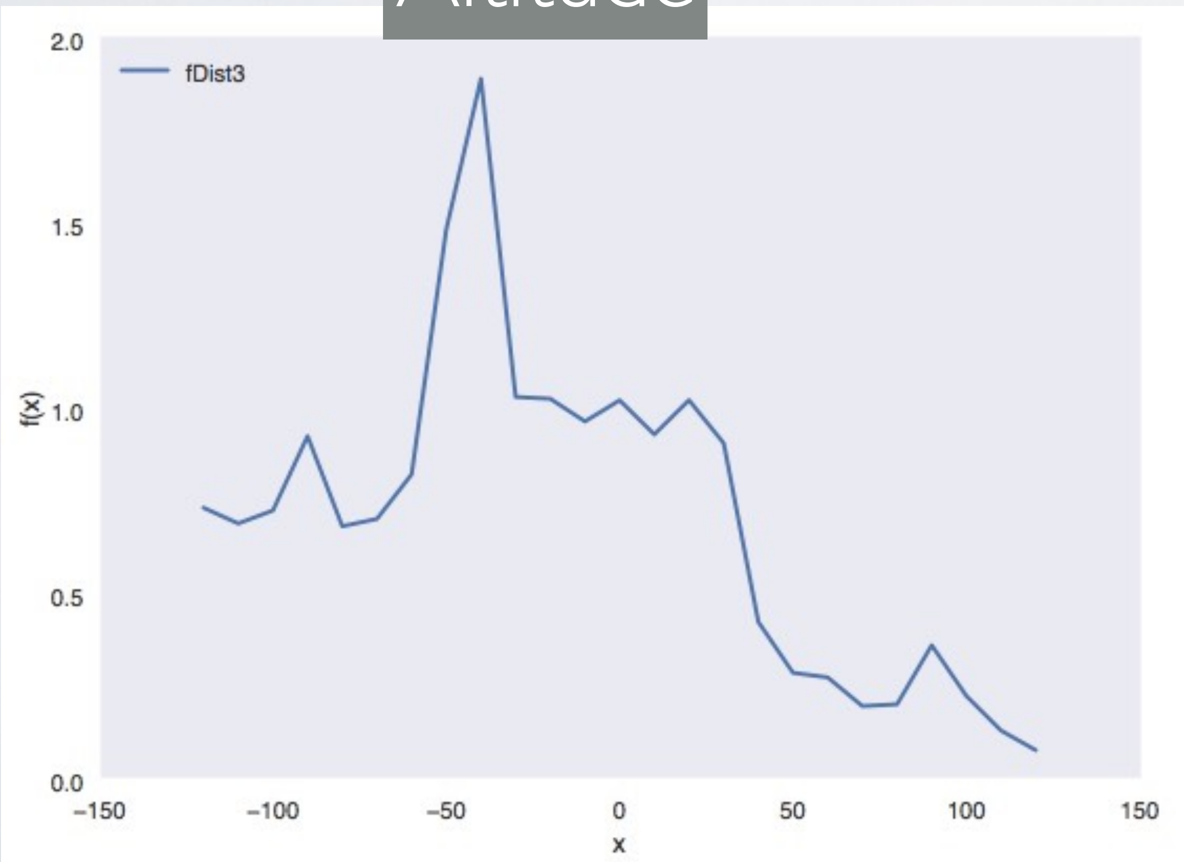
MODEL OF 1ST MODEL ERRORS ?

- Incorrectly predicted trips constitute a new network
  - Cannot be modelled by geographical gravity model (flat deterrence func)
  - Socio-demographic gravity model ?
  - Community structure ?

NON-GEOGRAPHICAL GRAVITY MODELS

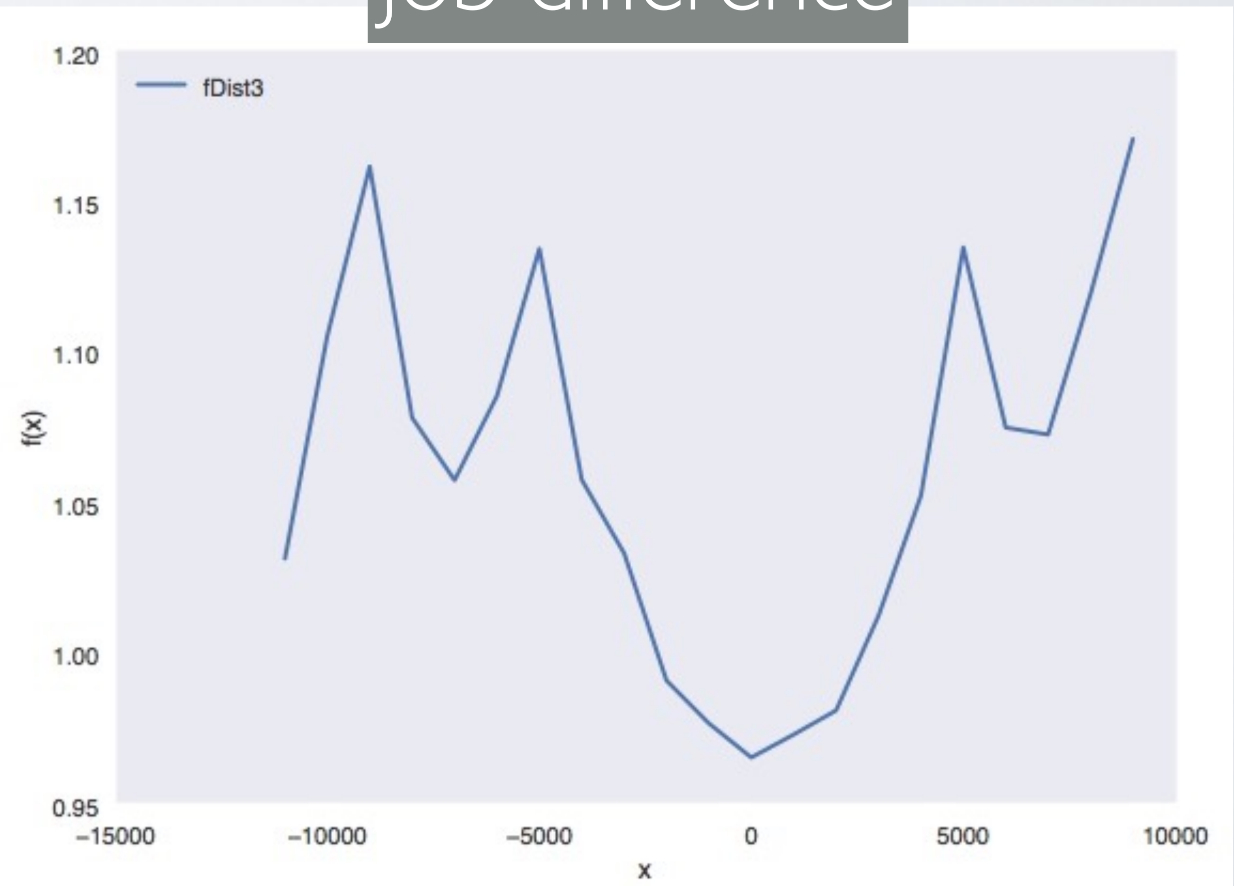
- Deterrence function can be computed on any distance function (here, on top of spatial effect)

Altitude



Altitude difference (meters)

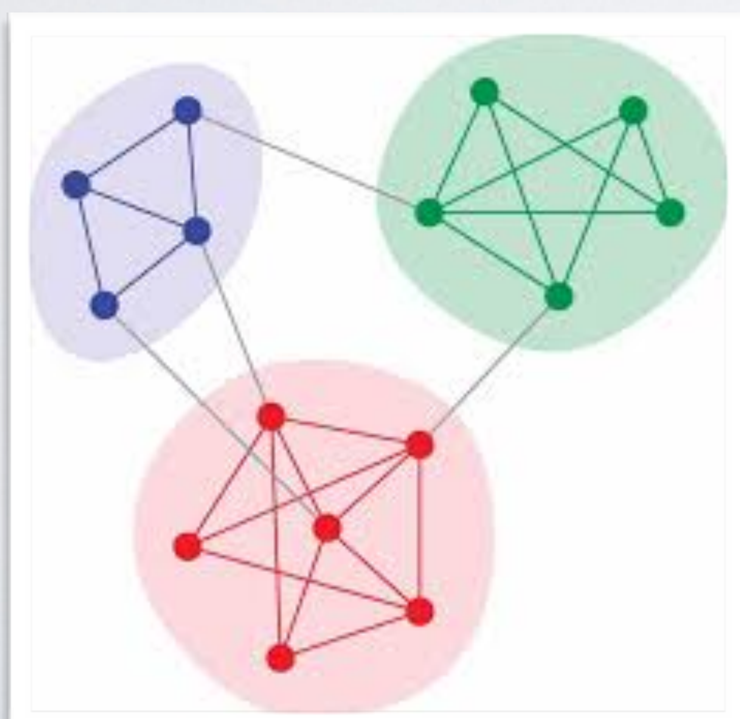
Job difference



Job difference (#job)

COMMUNITY STRUCTURE

- Community discovery (or graph clustering, SBM…)
  - Searching for groups of nodes with similar connections behaviours
  - Often dense groups less connected to the rest of the network, but not always

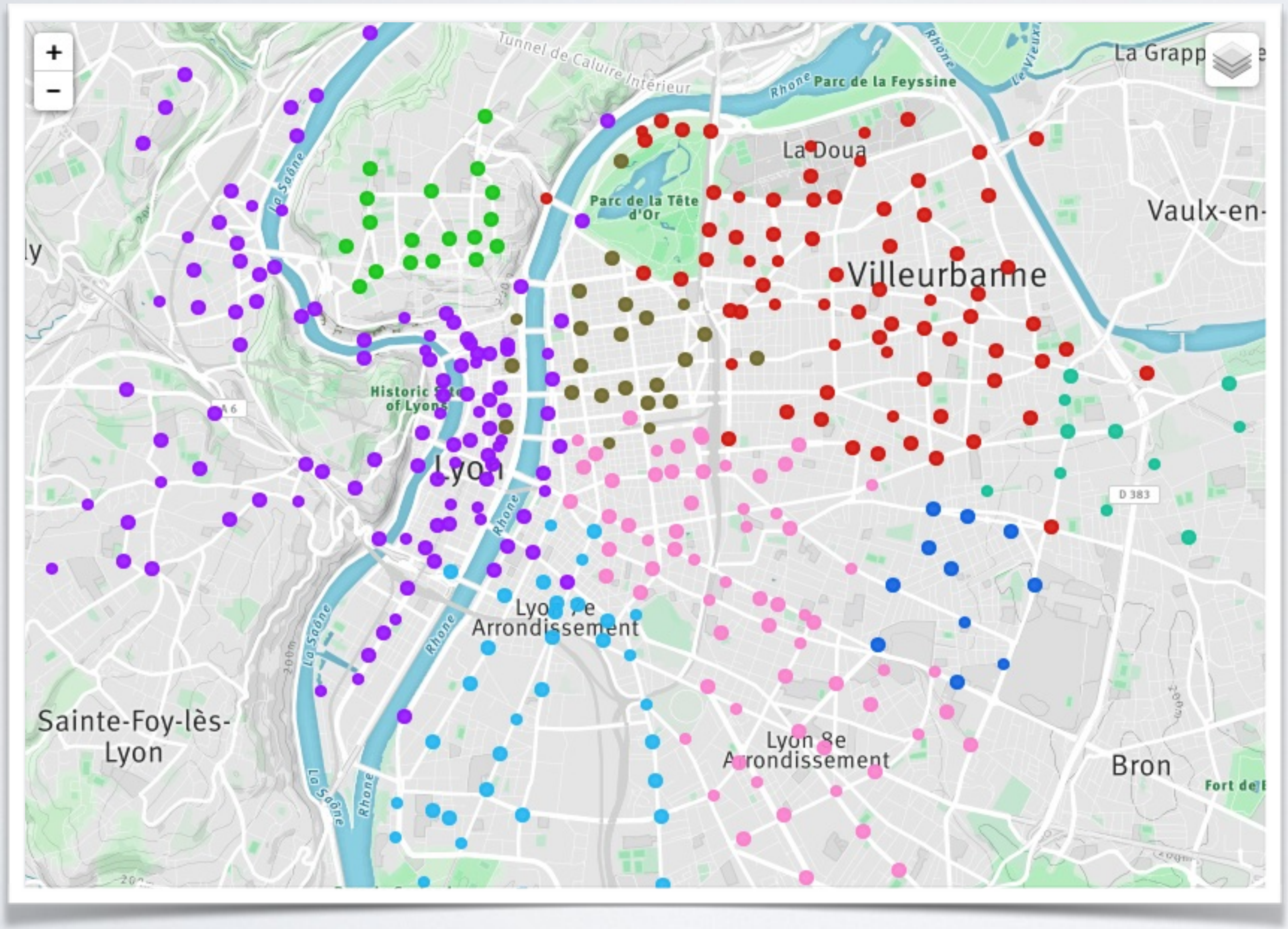


Community graph model:

Simple $p_{ij}^{SBM} = Wp^{C}(c_i, c_j)$

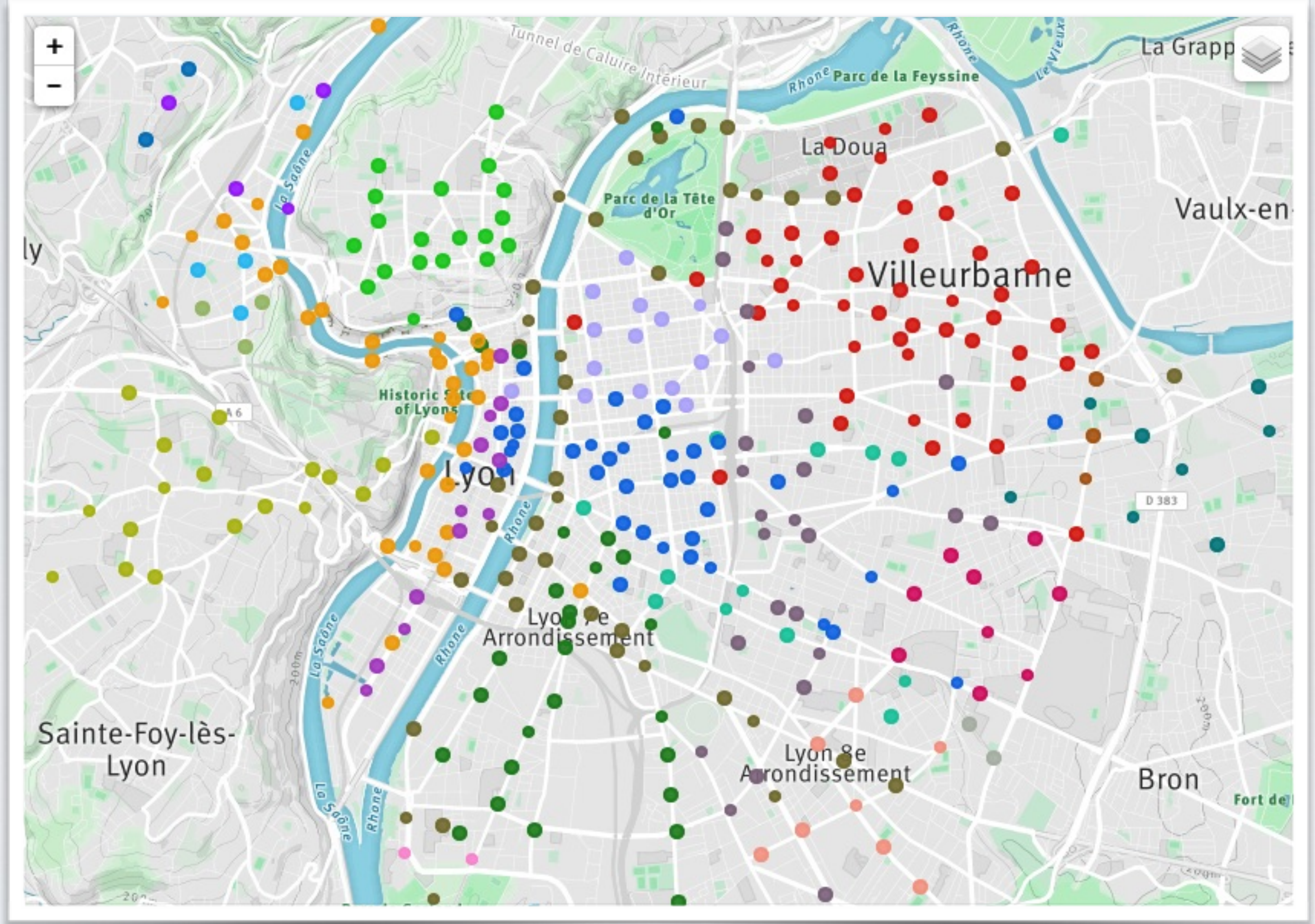
Degree-Corrected $p_{ij}^{DC\text{-}SBM} = Wk_i k_j p^{C}(c_i, c_j)$

COMMUNITY STRUCTURE



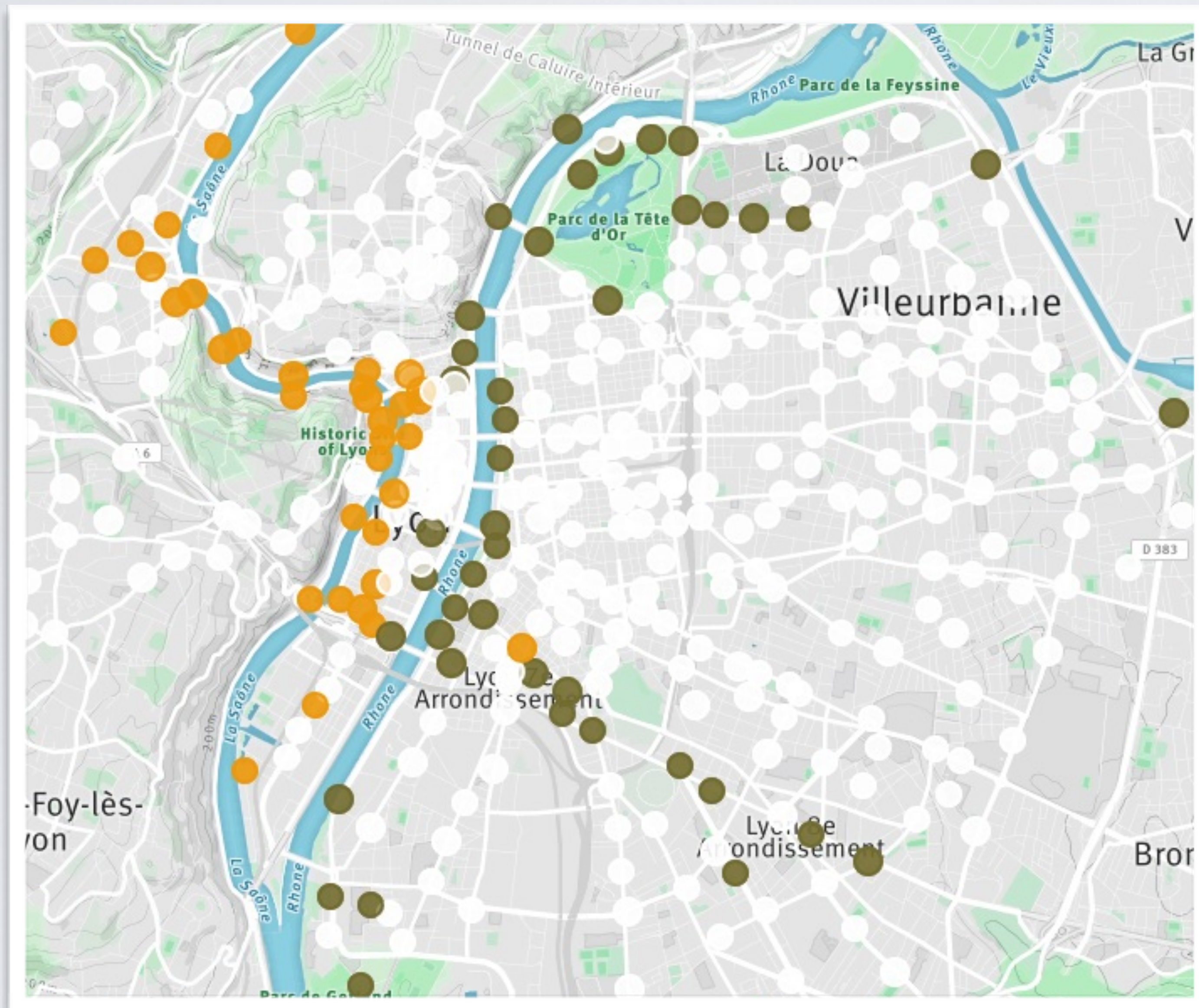
Community Structure of the Original Network

COMMUNITY STRUCTURE



Community
Structure
Of trips
Unexplained
By Spatial Model

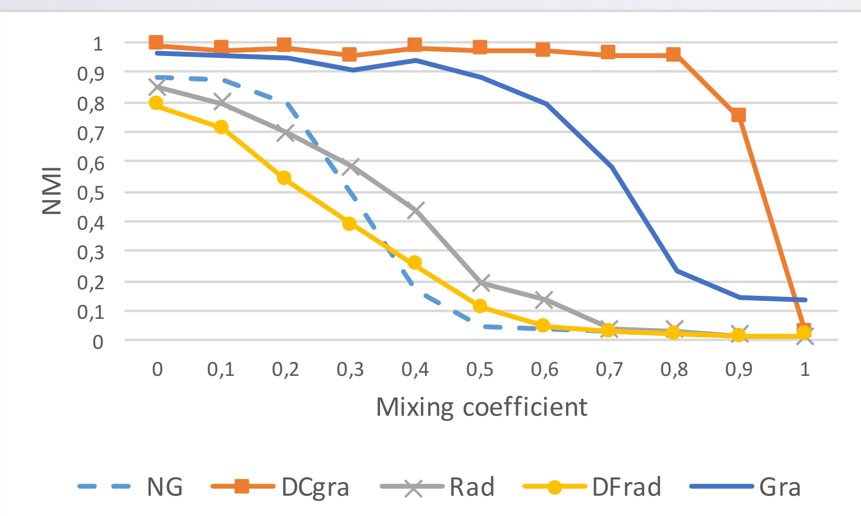
COMMUNITY STRUCTURE



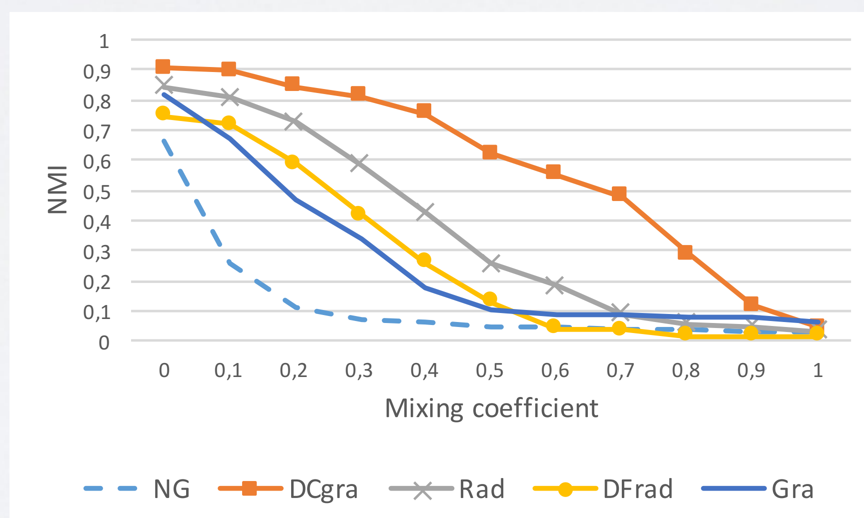
Community
Structure
Of trips
Unexplained
By Spatial Model

MIXED MODELS ?

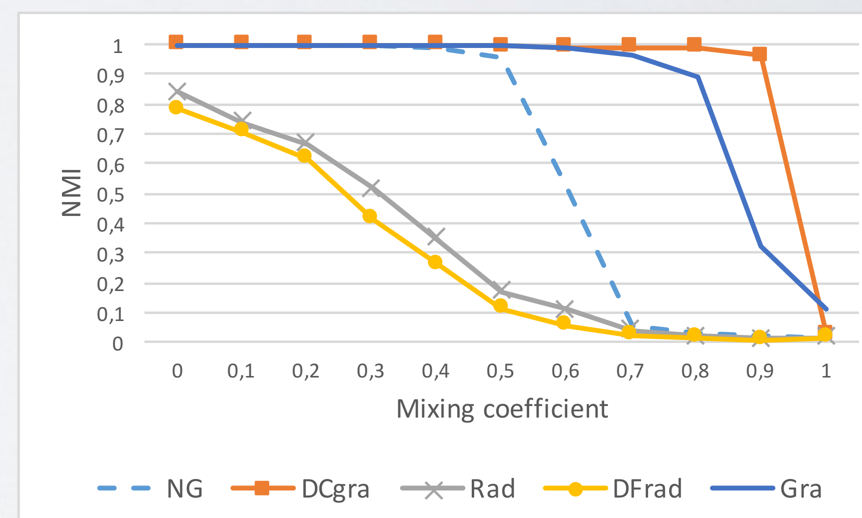
- Validation on synthetic cases:
  - Generation of networks with both a spatial and community structure.
  - Can we discover both of them ?



(b) $f(x) = 1/x$, N/P Only



(d) $f(x) = 1/x^2$, N/P Only

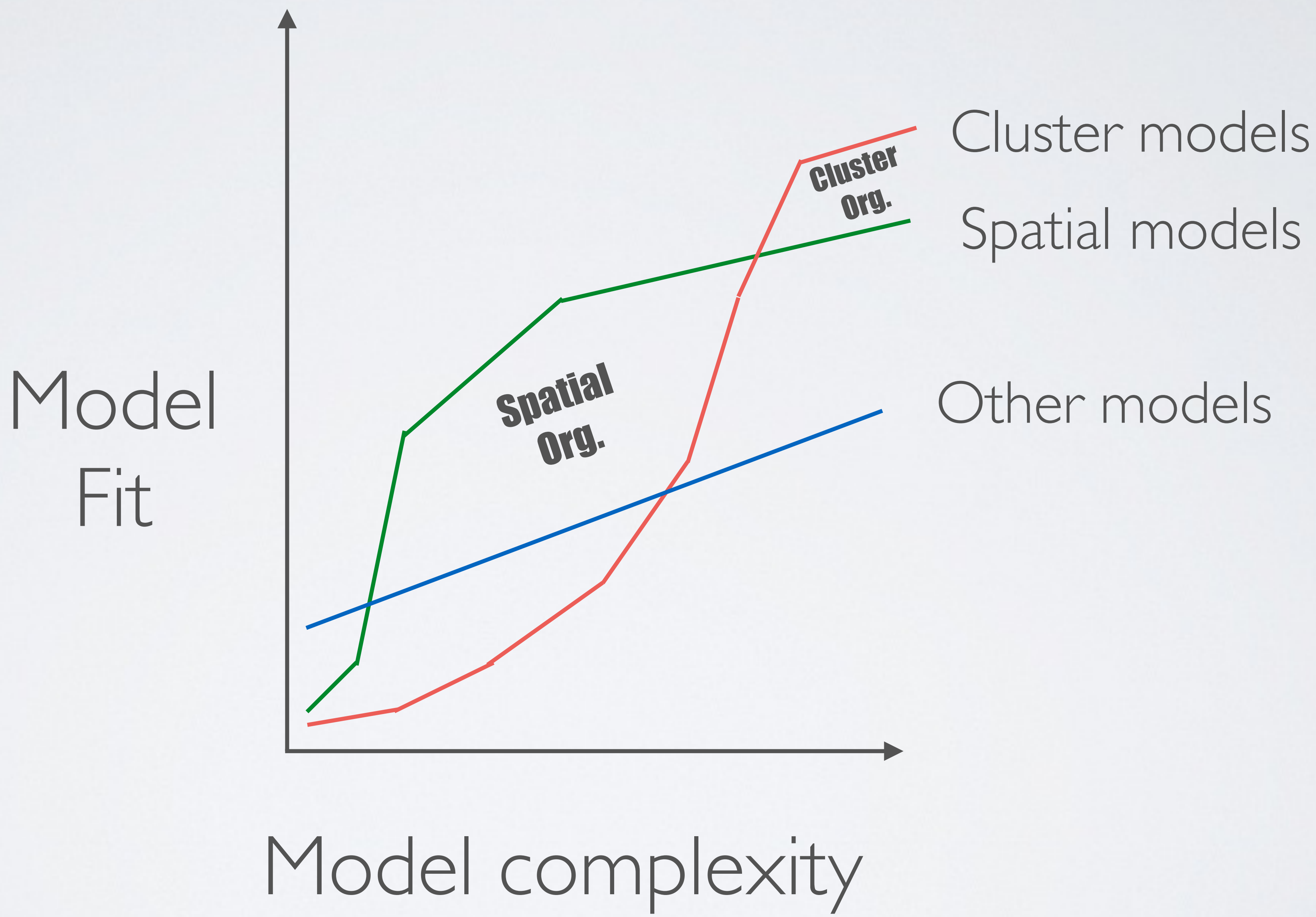


(f) $f(x) = 1/x^{0.5}$, N/P Only

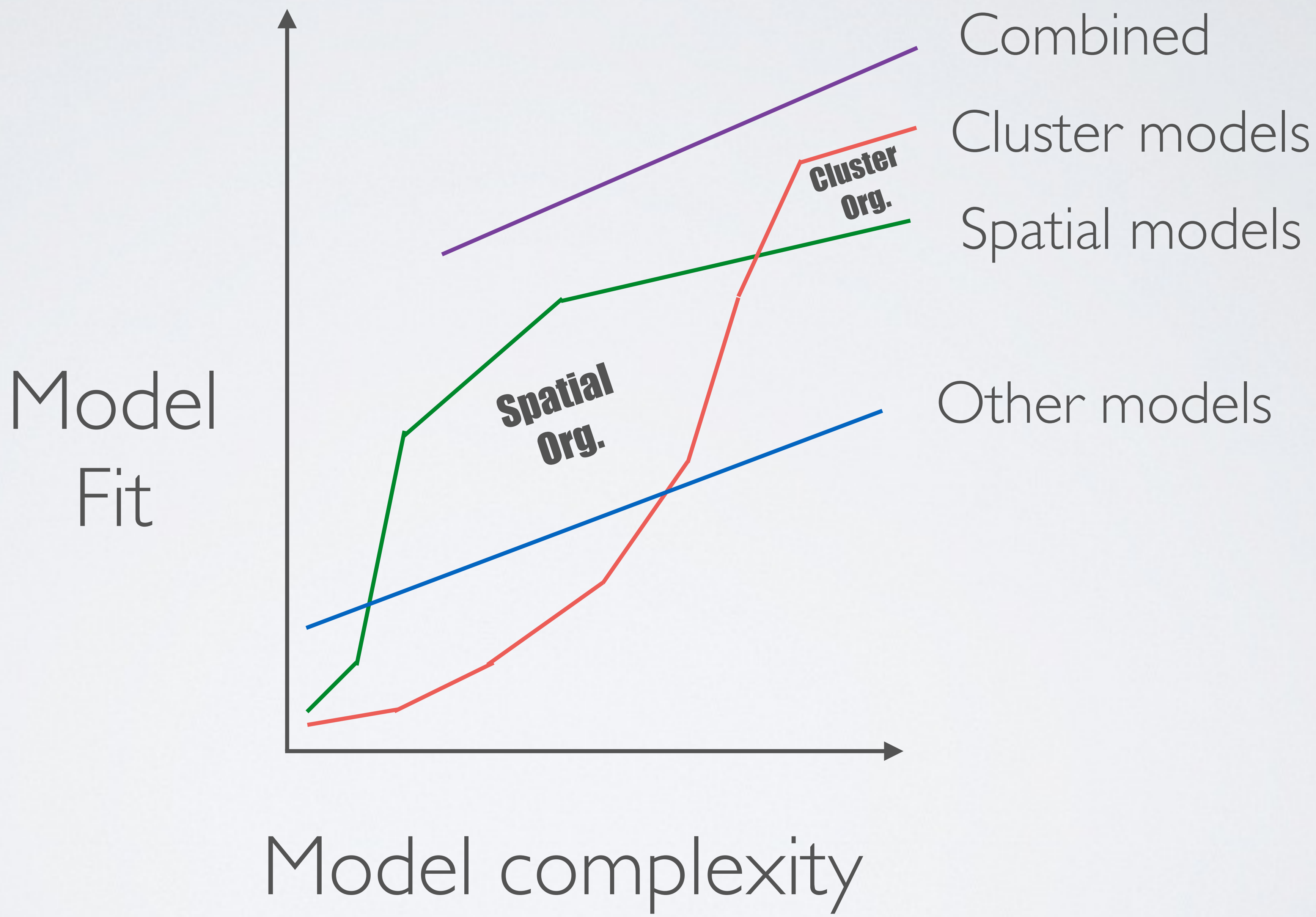
TAKE-HOME MESSAGE

- Complex systems organisation can often be modelled by networks
- Different network models exist
- Complex networks can probably be explained by a combination of factors, i.e. a combination of models

FUTURE WORK (HOPEFULLY)



FUTURE WORK (HOPEFULLY)



THANK YOU
FOR YOUR ATTENTION

PROPOSED NULL MODEL

- **Problem**: Does not conserve degrees !
  - Central nodes have higher degrees
  - Those at the periphery have lower ones.