

DYNAMICAL PROCESSES on Complex Networks

Resilience, vulnerability

Random Walks

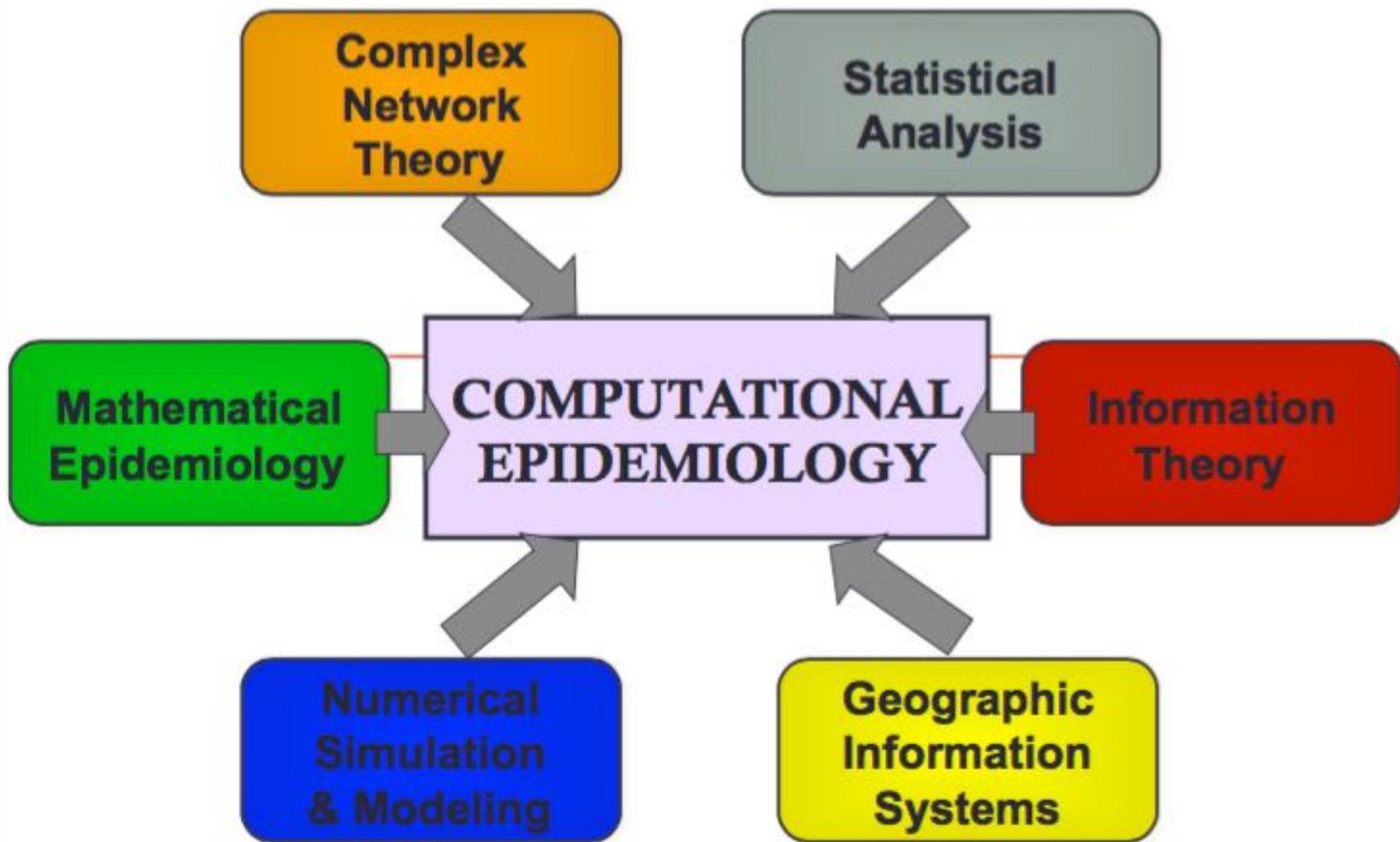
Epidemic processes

Social phenomena

Epidemiology: is the scientific study of factors affecting the health of populations: we will focus on infectious disease epidemiology, i.e. diseases transmitted through contacts among individuals

Two levels:

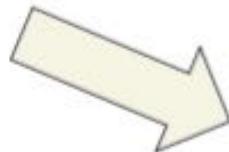
- Microscopic: researchers try to disassemble and kill new viruses => quest for vaccines and medicines
- **Macroscopic:** statistical analysis and modeling of epidemiological data in order to find information and policies aimed at lowering epidemic outbreaks => macroscopic prophylaxis, vaccination campaigns...



Modeling: What is it ?



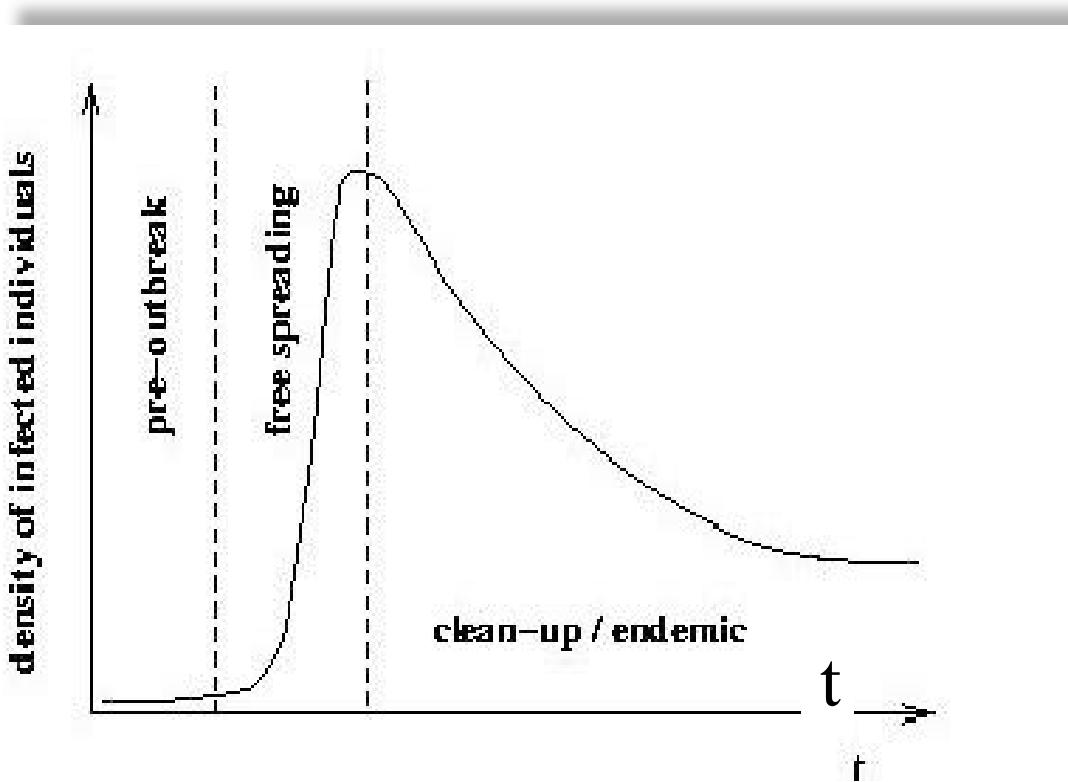
← reality



abstraction,
conceptualization



Stages of an epidemic outbreak

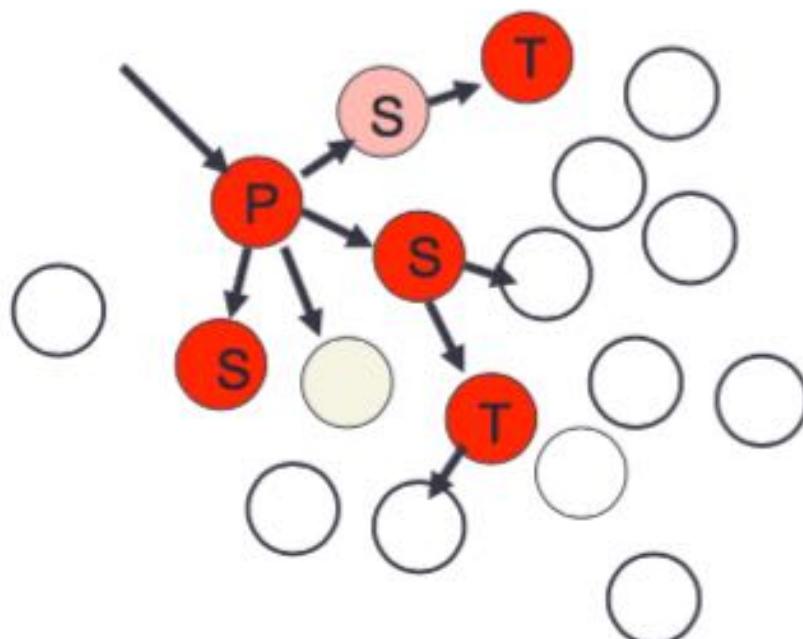
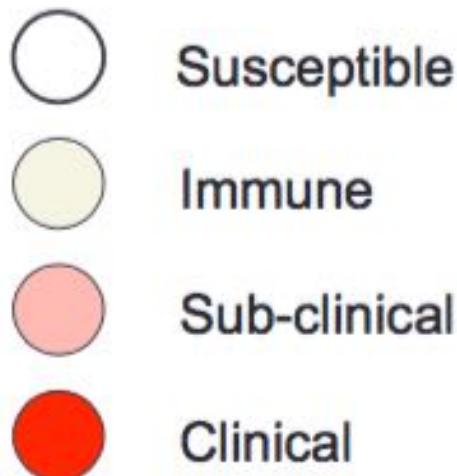


Infected individuals => prevalence/incidence

Transmission

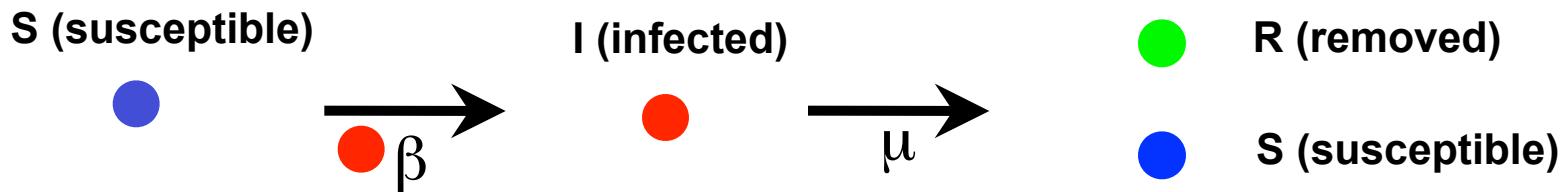
Cases

- ❖ Index – the first case identified
- ❖ Primary – the case that brings the infection into a population
- ❖ Secondary – infected by a primary case
- ❖ Tertiary – infected by a secondary case

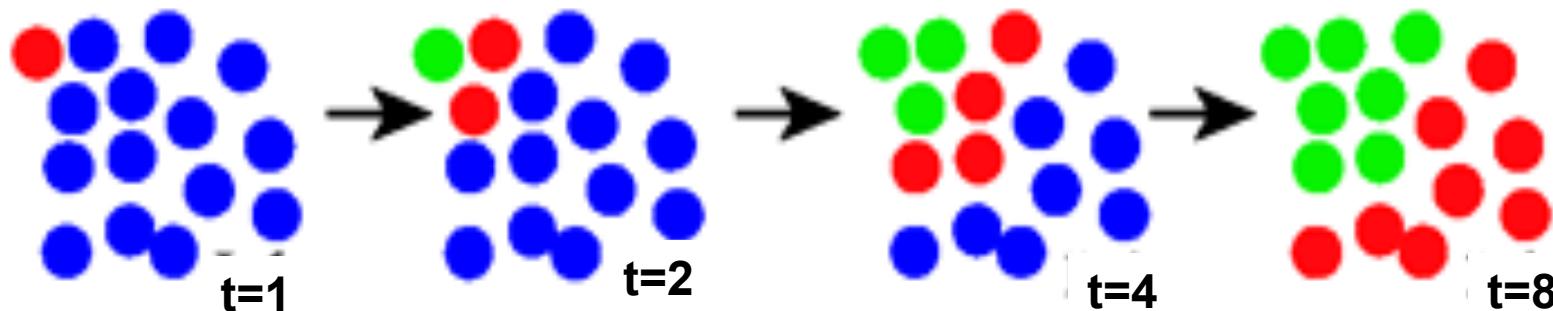


Standard epidemic modeling

Compartments: S, I, R...



Homogeneous mixing assumption (Mean-field)



SIR model



S number of susceptible individuals

I number of infectious individuals

R number of recovered individuals

N total population

$$N = S + I + R$$

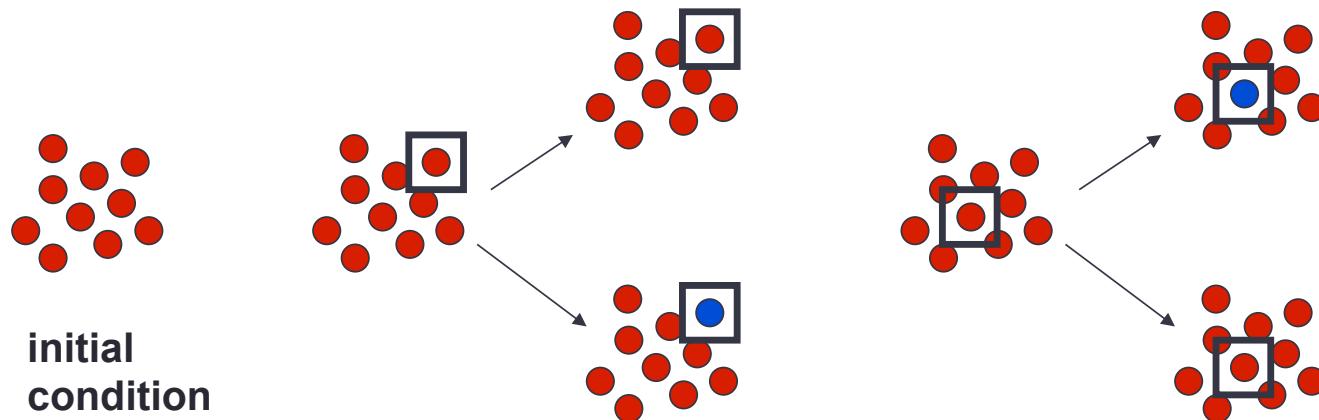
- ~~▪ age
▪ gender
▪ health
▪ job
▪ severity of disease
▪ environment
▪ social status
▪ latency
▪~~

SIR model: $I \rightarrow R$

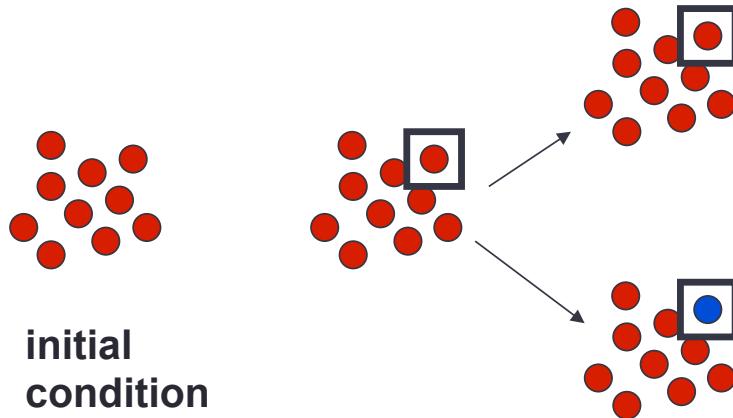


- transition rate $I \rightarrow R$ constant: μ **WHY ???**

τ : average infectious time; e.g. $\tau = 3\text{d}$



SIR model: $I \rightarrow R$

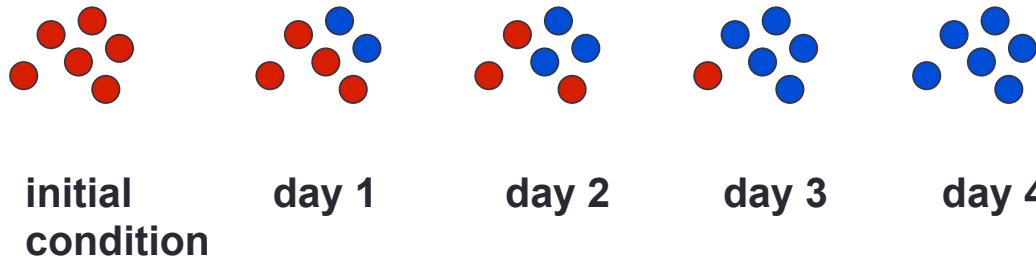


τ : average infectious time; e.g. $\tau = 3\text{d}$

probability of $I \rightarrow R$ at each day?

average number of $I \rightarrow R$ at each day?

SIR model: $I \rightarrow R$



$$\mu = \frac{1}{\tau}$$

rate of recovery = inverse of average infectious period

$$\Delta I = \mu I \Delta t$$

average number of infected recovering in the time interval Δt

compartments: *one-body interactions*



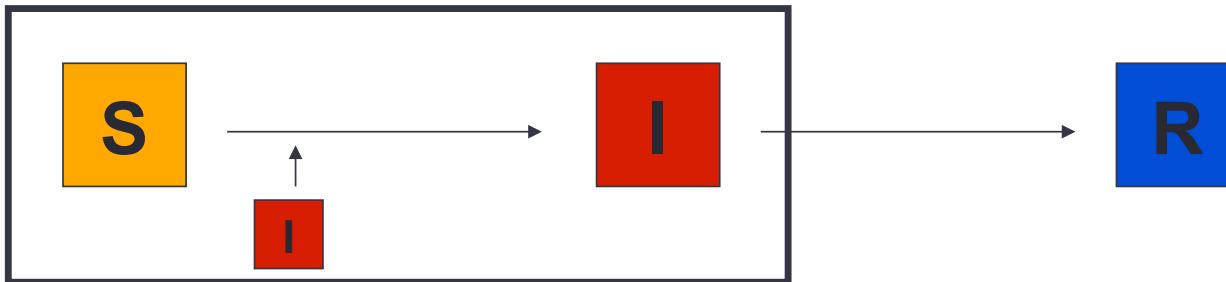
$$\gamma = \frac{1}{\tau_\gamma}$$

rate of transition = inverse of average period spent in C_1

$$\Delta C_1 = \gamma \cdot C_1 \Delta t$$

average number of transitions from C_1 to C_2 in the time interval Δt

SIR model: $S+I \rightarrow 2I$



- ***homogeneous mixing***: net rate of new infections proportional to

$$\frac{S \cdot I}{N}$$

WHY ???

→ Course of an epidemic depends on the rate of contacts between susceptible and infectious individuals (Hamer 1906 – Ross 1908)

SIR model: $S+I \rightarrow 2I$

force of infection for each susceptible:

transmission rate

x *effective number of contacts per unit time*

x *proportion of contacts infectious* $\longrightarrow I/N$

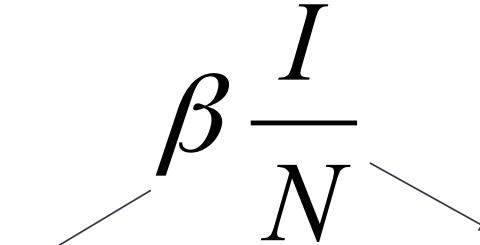
β

SIR model: $S+I \rightarrow 2I$

force of infection for each susceptible:

$$\beta \frac{I}{N}$$

number of contacts & transmission probability: contact = I



average number of $S \rightarrow I$???

SIR model: $S+I \rightarrow 2I$

average number of $S \rightarrow I$

$$\beta \cdot S \frac{I}{N}$$

The term $\beta \cdot S \frac{I}{N}$ is shown in a red-bordered box. Two arrows point towards it from below: one to the left pointing to the text "number of contacts & transmission", and one to the right pointing to the text "probability: contact = I ".

number of contacts
& transmission

probability: contact = I

compartments: *two-body interactions*



$$\beta \frac{I}{N}$$

rate of transition

$$\Delta S = \beta \cdot S \frac{I}{N} \Delta t$$

average number of transitions
from S to I in the time interval
 Δt

Deterministic SIR model

$$\langle S_{t+\Delta t} \rangle = \langle S_t \rangle - \beta \Delta t \langle S_t \rangle \langle I_t \rangle / N$$

$$\langle I_{t+\Delta t} \rangle = \langle I_t \rangle + \beta \Delta t \langle S_t \rangle \langle I_t \rangle / N - \mu \Delta t \langle I_t \rangle$$

$$\langle R_{t+\Delta t} \rangle = \langle R_t \rangle + \mu \Delta t \langle I_t \rangle$$

Constant population

$$\langle S_{t+\Delta t} \rangle + \langle I_{t+\Delta t} \rangle + \langle R_{t+\Delta t} \rangle = \langle S_t \rangle + \langle I_t \rangle + \langle R_t \rangle$$

Deterministic SIR model continued...

$$\frac{S(t + \Delta t) - S(t)}{\Delta t} = -\beta \frac{I(t)S(t)}{N}$$

$$\frac{I(t + \Delta t) - I(t)}{\Delta t} = \beta \frac{I(t)S(t)}{N} - \mu I(t)$$

$$\frac{R(t + \Delta t) - R(t)}{\Delta t} = \mu I(t)$$

$$\lim_{\Delta t \rightarrow 0} \frac{S(t + \Delta t) - S(t)}{\Delta t} \rightarrow \partial_t S(t)$$

Discrete deterministic SIR model

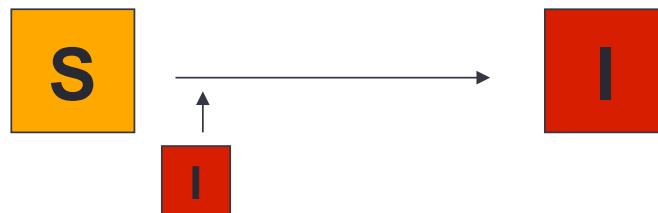
$$\Delta t = 1$$

$$S(t+1) = S(t) - \beta \frac{I(t)S(t)}{N}$$

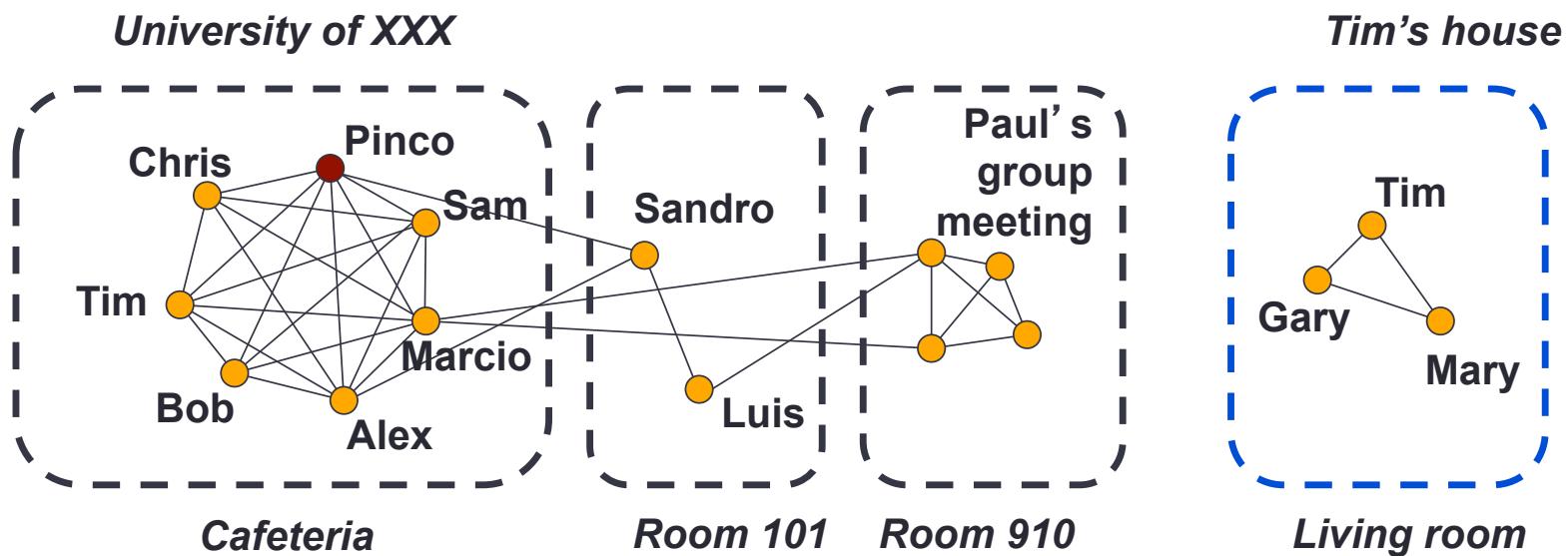
$$I(t+1) = I(t) + \beta \frac{I(t)S(t)}{N} - \mu I(t)$$

$$R(t+1) = R(t) + \mu I(t)$$

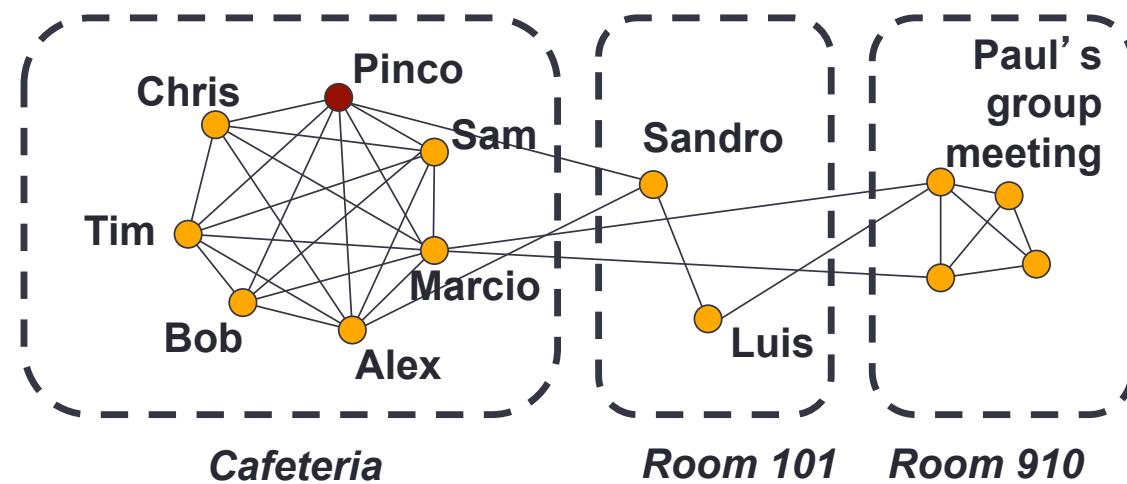
Homogeneous Mixing



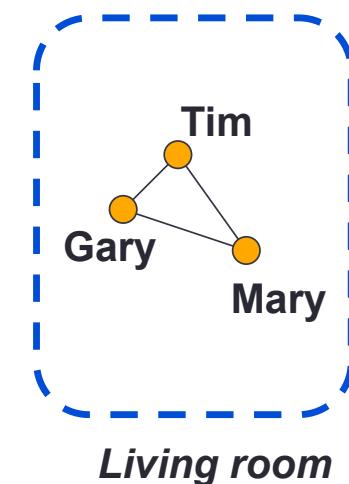
$$\Delta S = \beta \cdot S \frac{I}{N} \Delta t$$



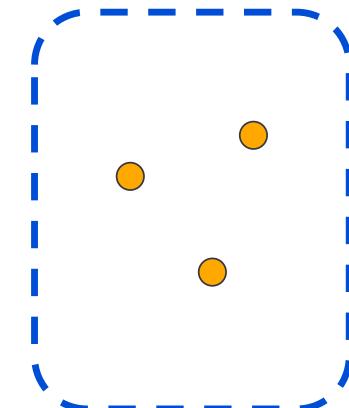
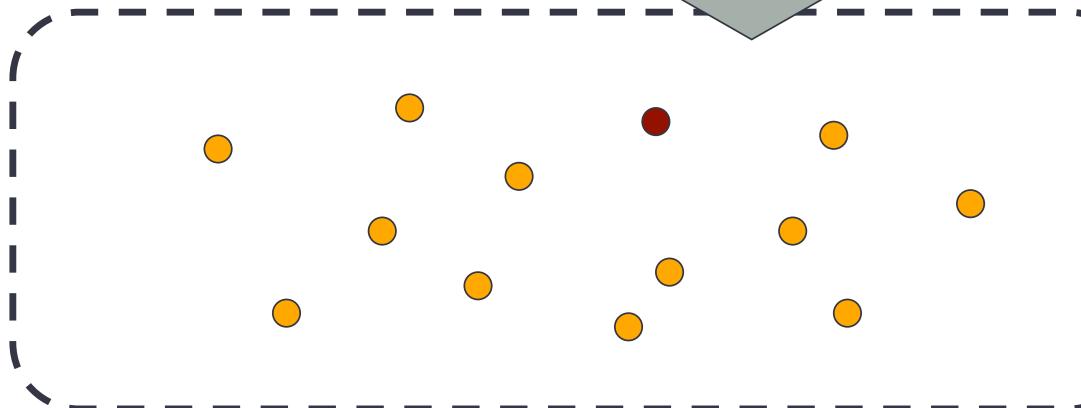
University of XXX



Tim's house



HOMOGENEOUS MIXING



SIR solution: Early stage approximation

$$S(t) = N - R - I$$

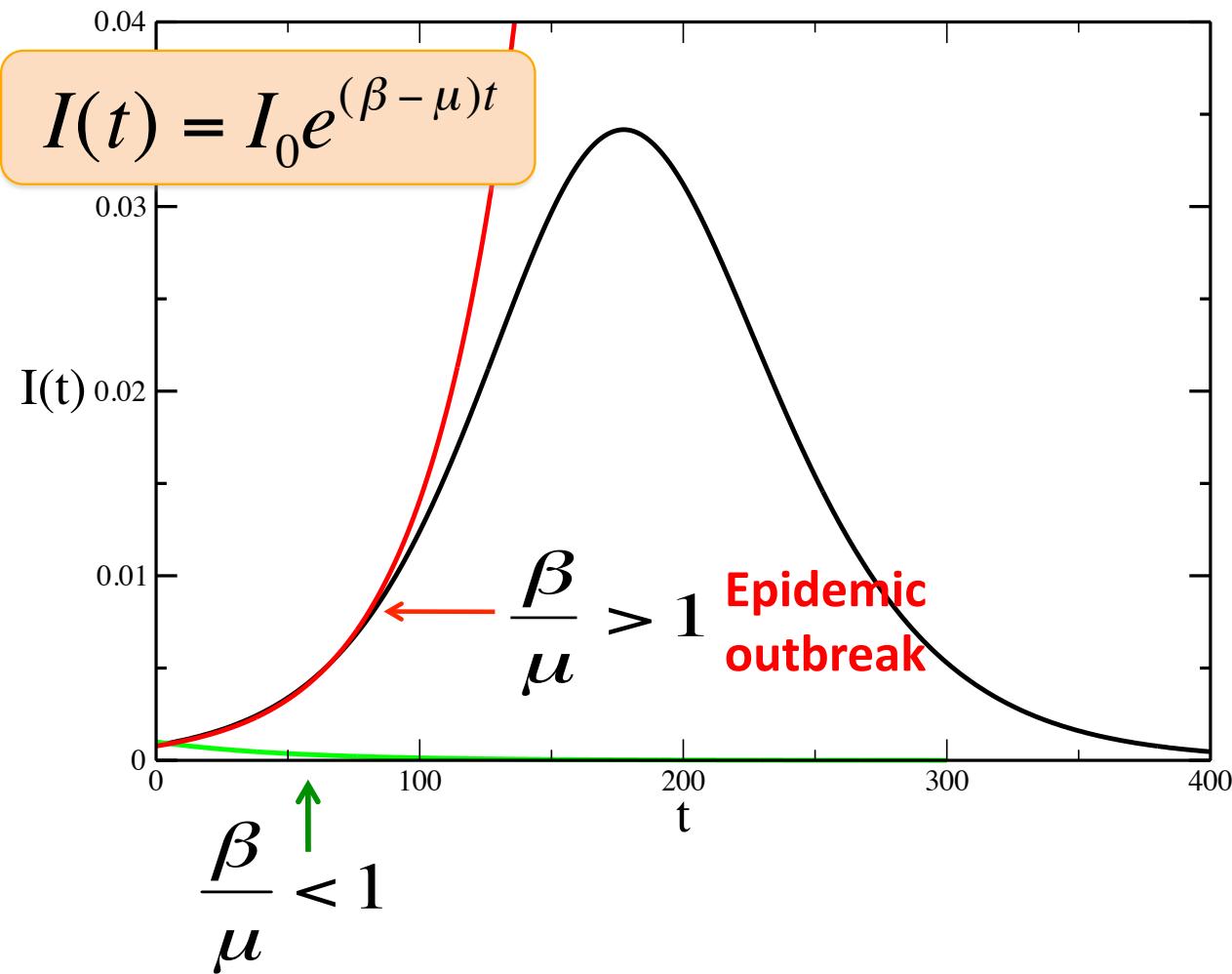
$$\partial_t S(t) = -\beta \frac{I(t) S(t)}{N} \approx N$$

$$\partial_t I(t) = \beta \frac{I(t) S(t)}{N} - \mu I(t) \approx N$$

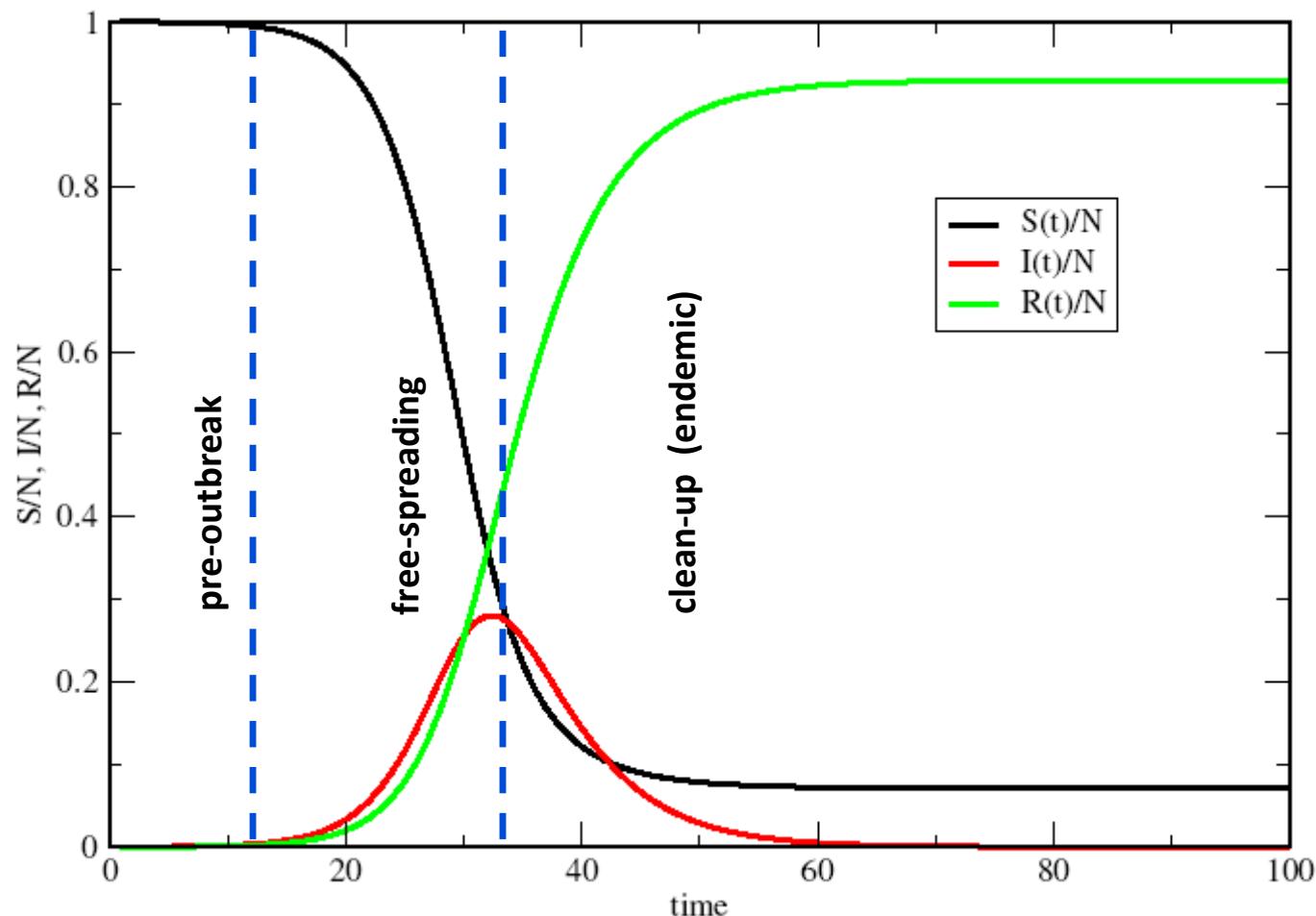
$$\partial_t R(t) = \mu I(t)$$

$$\boxed{\partial_t I(t) = (\beta - \mu)I(t)}$$

SIR solution: Early stage approximation



SIR dynamics



Condition to have an epidemic

$$\beta - \mu > 0$$

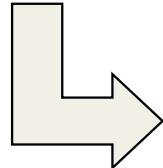


$$\frac{\beta}{\mu} > 1$$

Epidemic
Threshold

Reproductive number

$$\beta - \mu > 0$$

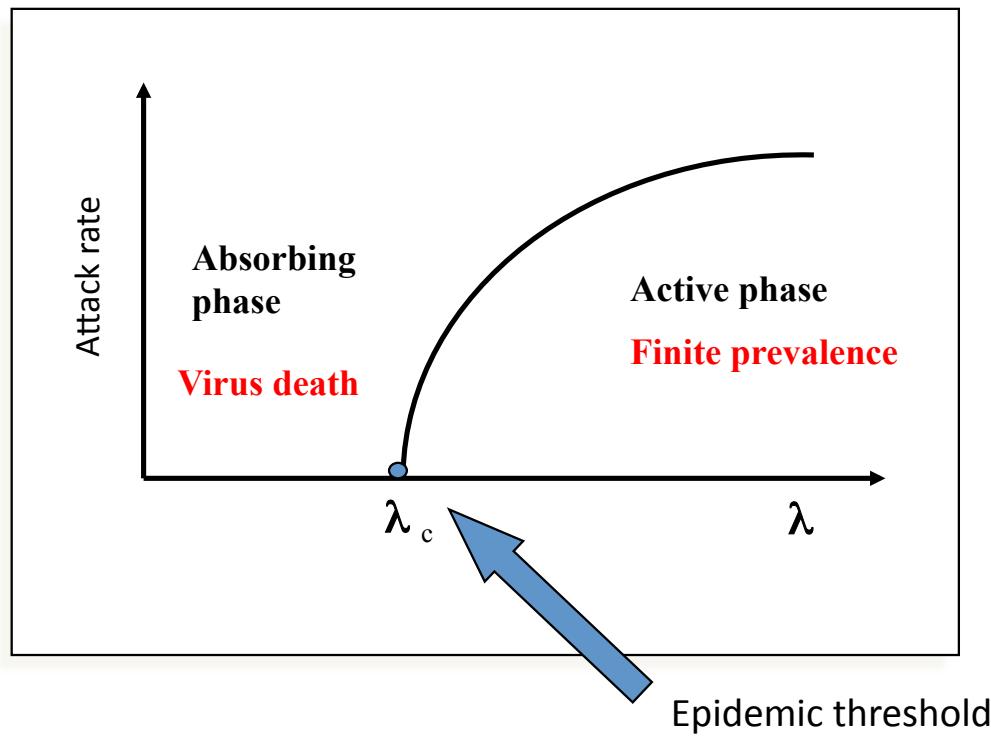


$$\mu \left(\frac{\beta}{\mu} - 1 \right) > 0 \Rightarrow \mu(R_0 - 1)$$

$$R_0 = \frac{\beta}{\mu}$$

Basic reproductive number

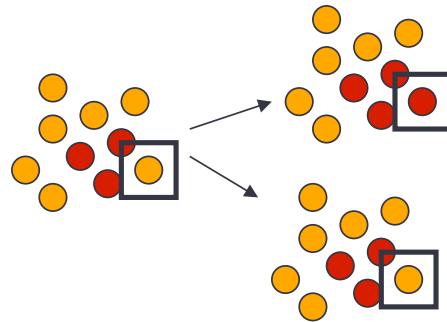
Phase diagram



In the SIR
 $\lambda=\beta/\mu$

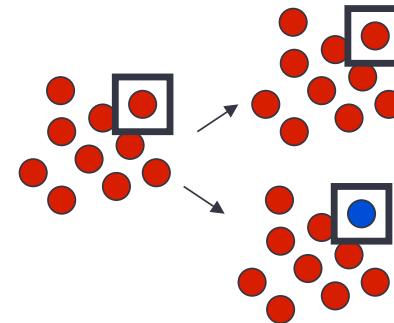
In general is the condition
on the set of disease
parameters
such that an outbreak
occur

Chance: how to model?



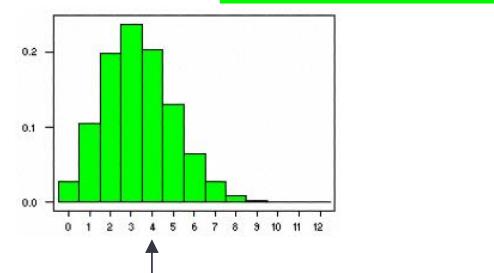
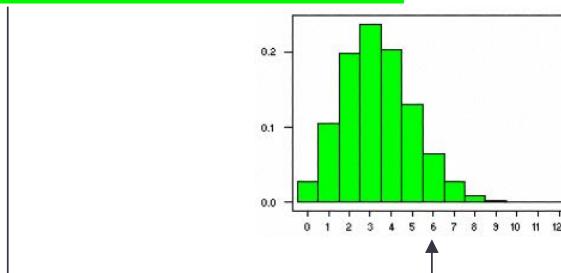
- number of trials S_t
- prob. success $\beta \Delta t I_t / N$

$$S_{t+\Delta t} = S_t - \Delta S$$
$$\Delta S = \text{Binom}(S_t, \beta \Delta t I_t / N)$$



- number of trials I_t
- prob. success $\mu \Delta t$

$$I_{t+\Delta t} = I_t + \Delta S - \Delta R$$
$$\Delta R = \text{Binom}(I_t, \mu \Delta t)$$



Stochastic SIR model

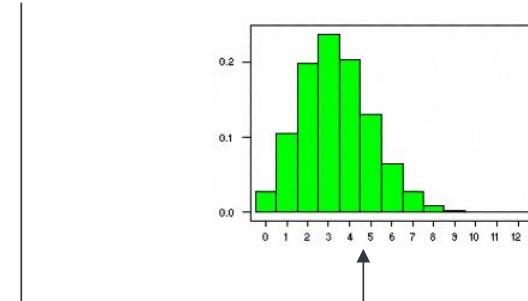
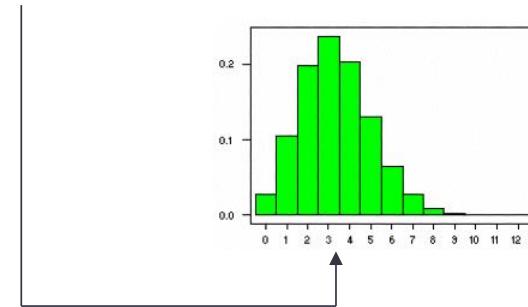
$$S_{t+\Delta t} = S_t - \Delta S$$

$$\Delta S = \text{Binom}(S_t, \beta \Delta t I_t / N)$$

$$I_{t+\Delta t} = I_t + \Delta S - \Delta R$$

$$R_{t+\Delta t} = R_t + \Delta R$$

$$\Delta R = \text{Binom}(I_t, \mu \Delta t)$$



Chance: how to code?

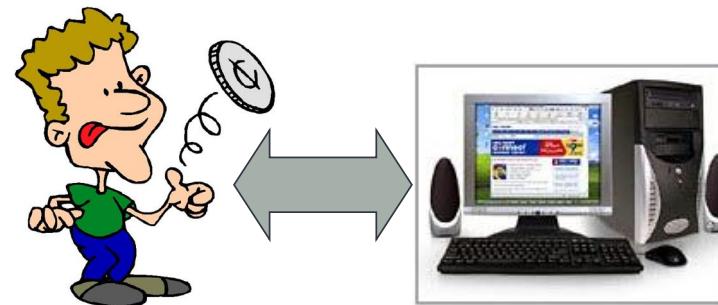
Algorithm for pseudo random numbers:

formula + random seed

- set `random_seed` value
- `random_number = F(random_seed)`

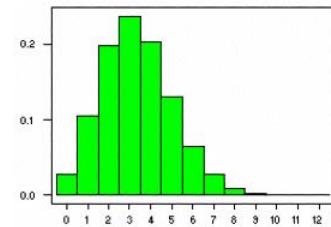
`F()` → integer random number [0,RAND_MAX]

`F()` → real random number [0,1]

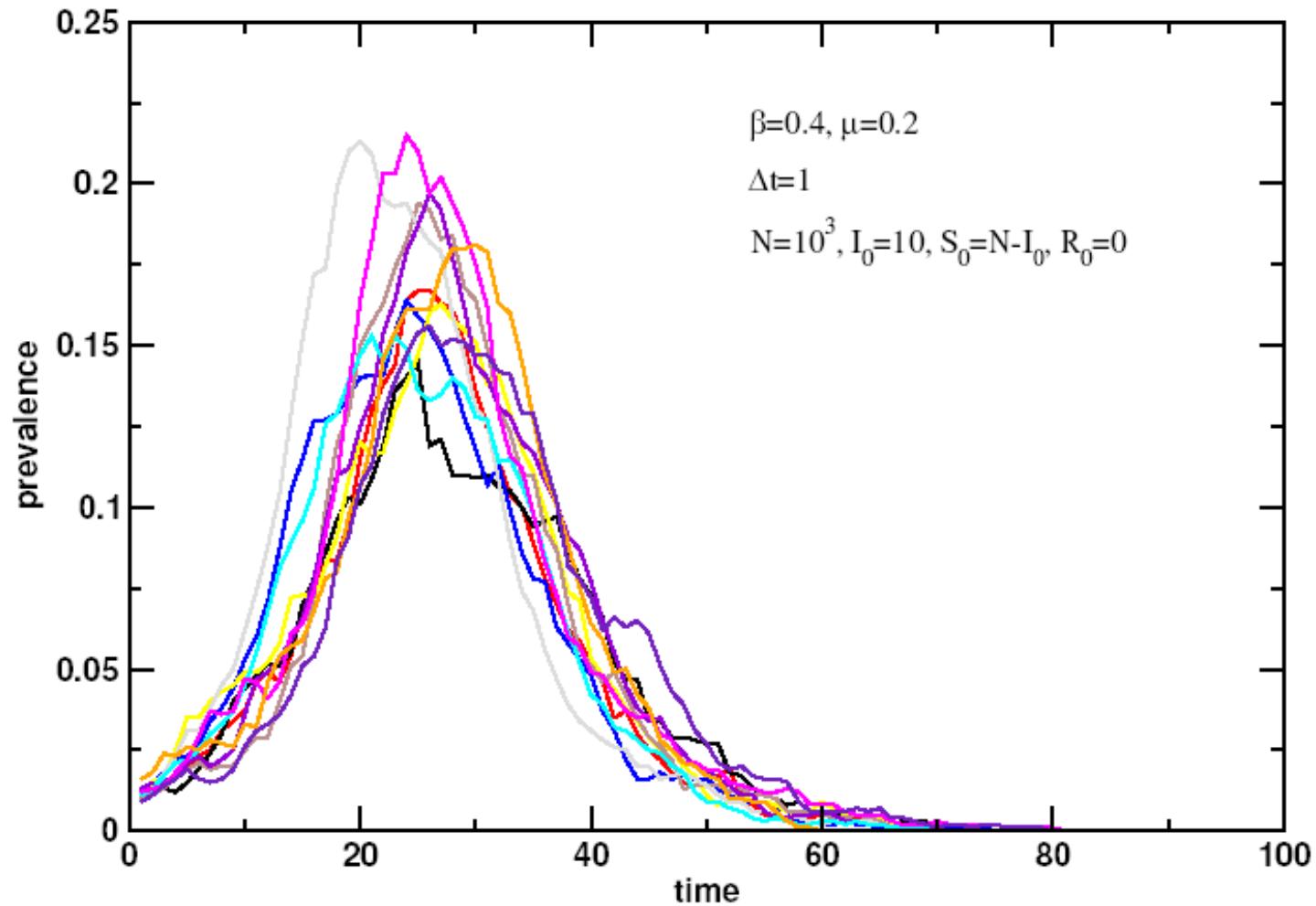


stochastic SIR model: *pseudo-code*

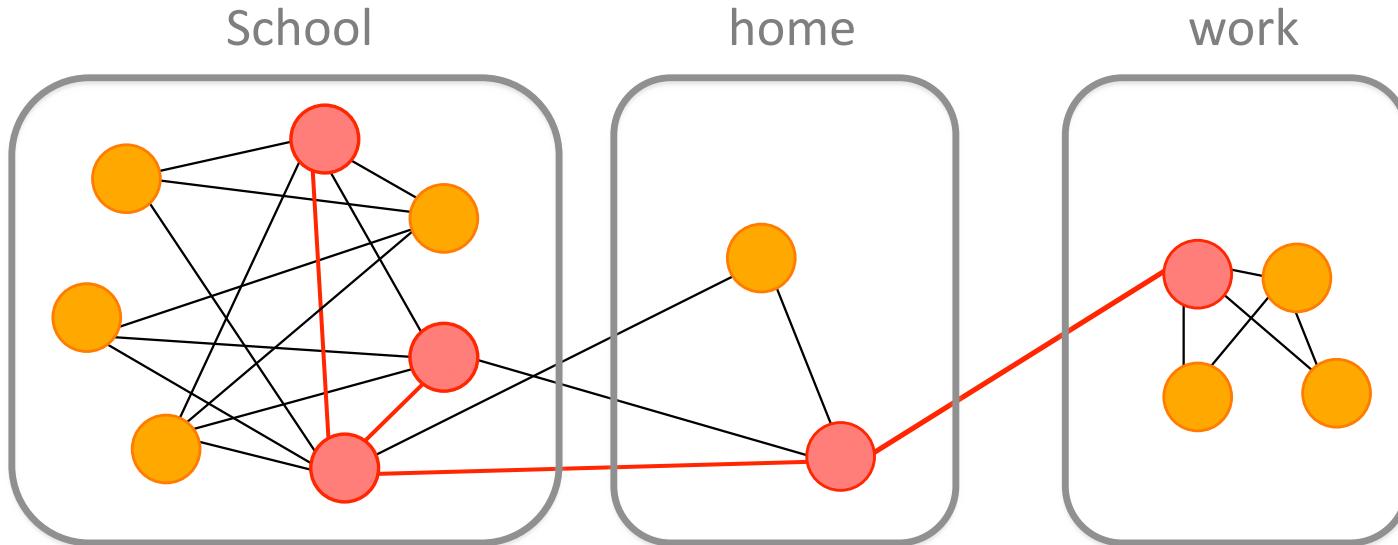
- set disease parameter values: β, μ
- set initial conditions: N, S_0, I_0, R_0
- set time step Δt
- define variables S, I, R and auxiliary variables S', I', R'
- set variables: $S=S'=S_0, I=I'=I_0, R=R'=R_0$
- loop on time t :
 - $\Delta S = \text{Binom}(S, \beta \Delta t I/N)$
 - $\Delta R = \text{Binom}(I, \mu \Delta t)$
- $S' = S - \Delta S$
- $I' = I + \Delta S - \Delta R$
- $R' = R + \Delta R$
- $S = S'$
- $I = I'$
- $R = R'$
- print results at time t



stochastic SIR model: *many runs*



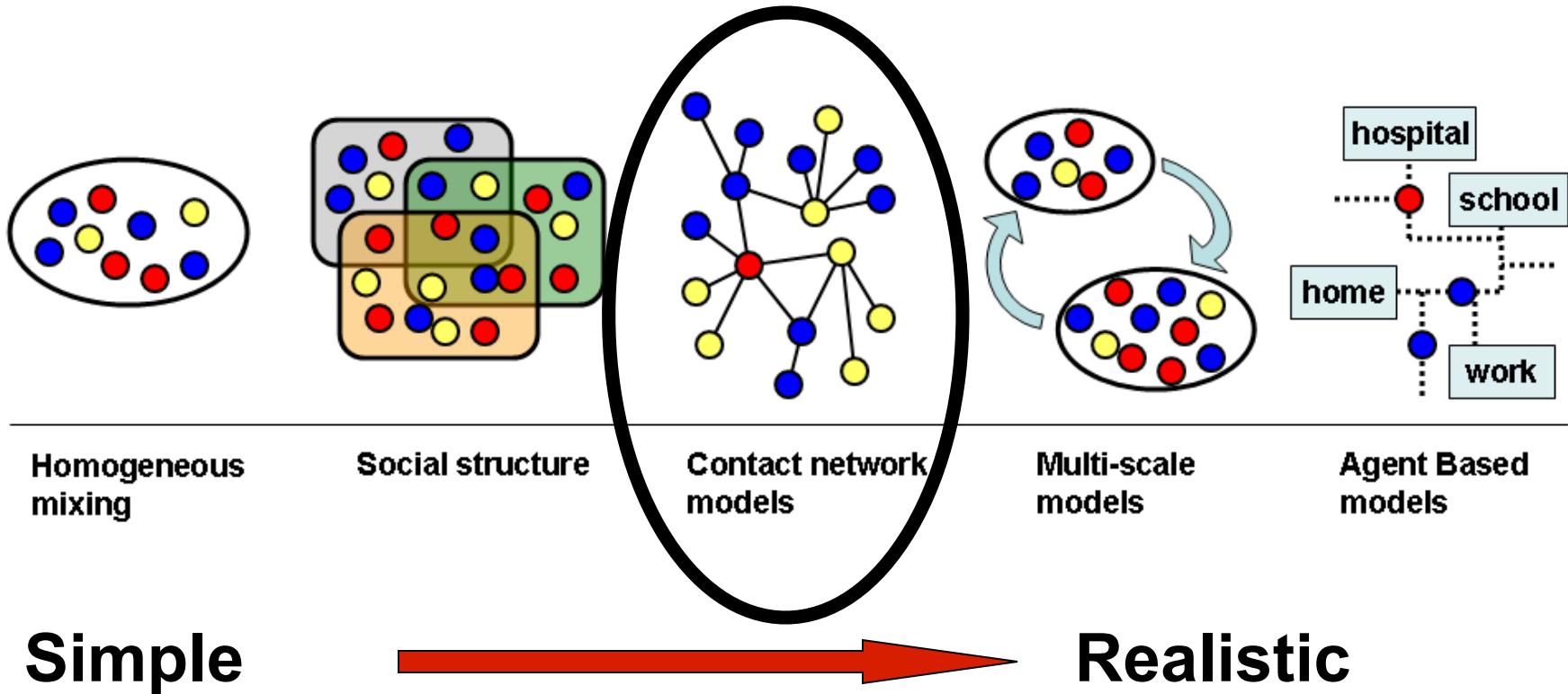
Epidemics in more realistic systems



Two strategies

- Agent based model approach
- Statistical description by means of a network theory approach

Wide spectrum of complications and complex features to include...



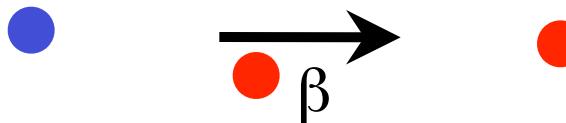
Ability to explain trends at a population level

Model realism loses in transparency.
Validation is harder.

Transmission

S (susceptible)

I (infected)



HOMOGENEOUS MIXING ASSUMPTION

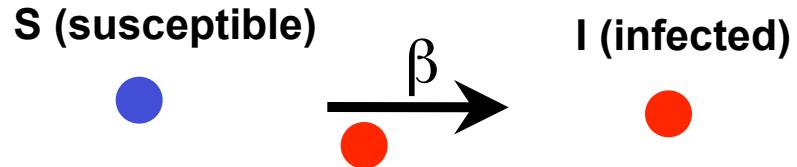
Individual in state S, with k contacts, among which n infectious: in the homogeneous mixing approximation, the probability to get the infection in each time interval dt is:

$$\begin{aligned}\text{Proba}(S \rightarrow I) &= 1 - \text{Proba}(\text{not to get infected by any infectious}) \\ &= 1 - (1 - \beta dt)^n \\ &\approx \beta n dt \quad (\beta dt \ll 1) \\ &\approx \beta k i dt \quad \text{as } n \sim k i \text{ for homogeneous mixing}\end{aligned}$$

Hypothesis of mean-field nature:

every individual sees the same density of infectious among his/her contacts, equal to the average density in the population

The SI model



N individuals

$I(t)$ =number of infectious, $S(t)=N-I(t)$ number of susceptible

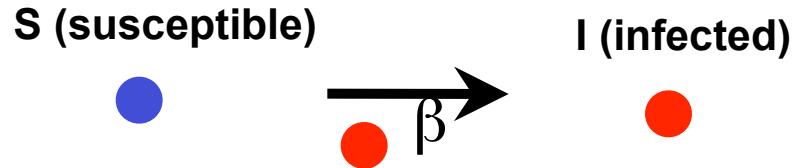
$$i(t)=I(t)/N, s(t)=S(t)/N = 1 - i(t)$$

If $k = \langle k \rangle$ is the same for all individuals (homogeneous network):

$$\begin{aligned} \frac{dI}{dt} &= S(t) \times \text{Proba}(S \rightarrow I) \\ &= \beta k S(t) i(t) \end{aligned}$$

$$\boxed{\frac{di}{dt} = \beta k i(t)(1 - i(t))}$$

The SI model



N individuals

$I(t)$ =number of infectious, $S(t)=N-I(t)$ number of susceptible

$i(t)=I(t)/N$, $s(t)=S(t)/N$

If k is the same for all individuals (homogeneous network):

$$\frac{di}{dt} = \beta \langle k \rangle i(1 - i)$$

The SI model

S (susceptible)



I (infected)

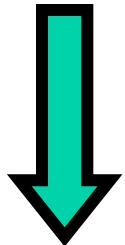


N individuals

$I(t)$ =number of infectious, $S(t)=N-I(t)$ number of susceptible

$i(t)=I(t)/N$, $s(t)=S(t)/N$

$$\frac{di}{dt} = \beta \langle k \rangle i(1 - i)$$



$$i(t) = \frac{i_0 \exp(t/\tau)}{1 + i_0(\exp(t/\tau) - 1)} \quad \tau = 1/(\beta \langle k \rangle)$$

The SIS model



N individuals

$I(t)$ =number of infectious, $S(t)=N-I(t)$ number of susceptible
 $i(t)=I(t)/N$, $s(t)=S(t)/N$

Homogeneous network:

$$\frac{di}{dt} = \beta \langle k \rangle i(1 - i) - \mu i$$



Competition of two time scales: $1/\mu$ and $1/(\beta \langle k \rangle)$

The SIR model

N individuals

$I(t)$ =number of infectious, $S(t)$ number of susceptible, $R(t)$ recovered

$$i(t)=I(t)/N, s(t)=S(t)/N, r(t)=R(t)/N=1-i(t)-s(t)$$

Homogeneous network:

$$\frac{ds}{dt} = -\beta \langle k \rangle i(t) s(t)$$

$$\frac{di}{dt} = \beta \langle k \rangle i(t) s(t) - \mu i(t)$$

$$\frac{dr}{dt} = \mu i(t)$$

SIS and SIR models: linear approximation

Short times, $i(t) \ll 1$ (and $r(t) \ll 1$ for the SIR)

$$\frac{di}{dt} \approx (\beta\langle k \rangle - \mu)i(t)$$

Exponential evolution $\exp(t/\tau)$, with

$$1/\tau = \beta\langle k \rangle - \mu$$

If $\beta\langle k \rangle > \mu$: exponential growth

If $\beta\langle k \rangle < \mu$: extinction

Epidemic threshold condition: $\beta\langle k \rangle = \mu$

Long time limit, SIS model

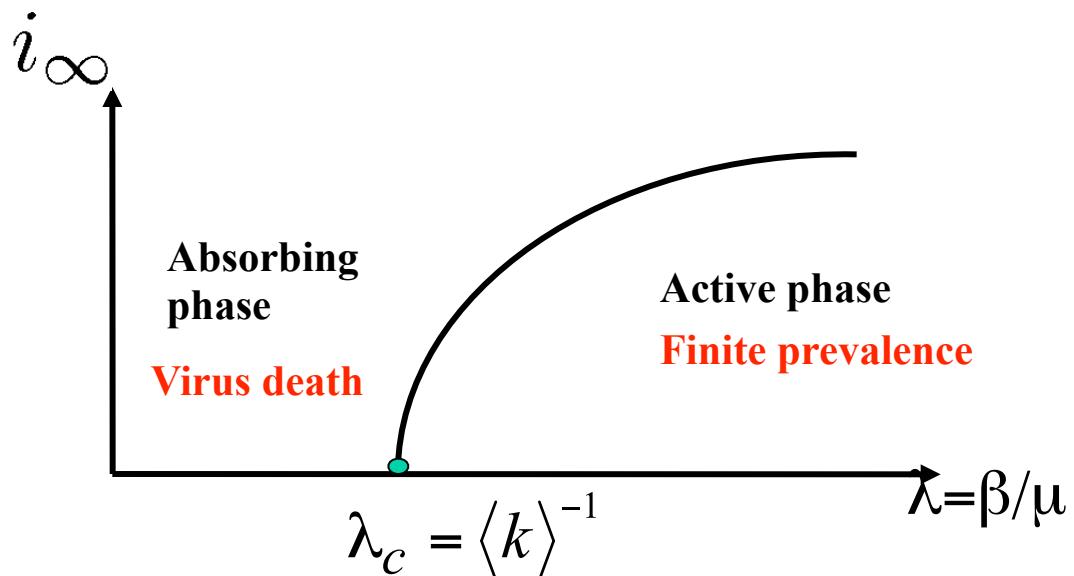
Stationary state: $di/dt = 0 \quad \mu i_\infty = \beta \langle k \rangle i_\infty (1 - i_\infty)$

If $\beta \langle k \rangle < \mu$: $i_\infty = 0$

Epidemic threshold condition: $\beta \langle k \rangle = \mu$

If $\beta \langle k \rangle > \mu$: $i_\infty = 1 - \mu / (\beta \langle k \rangle)$

Phase diagram:



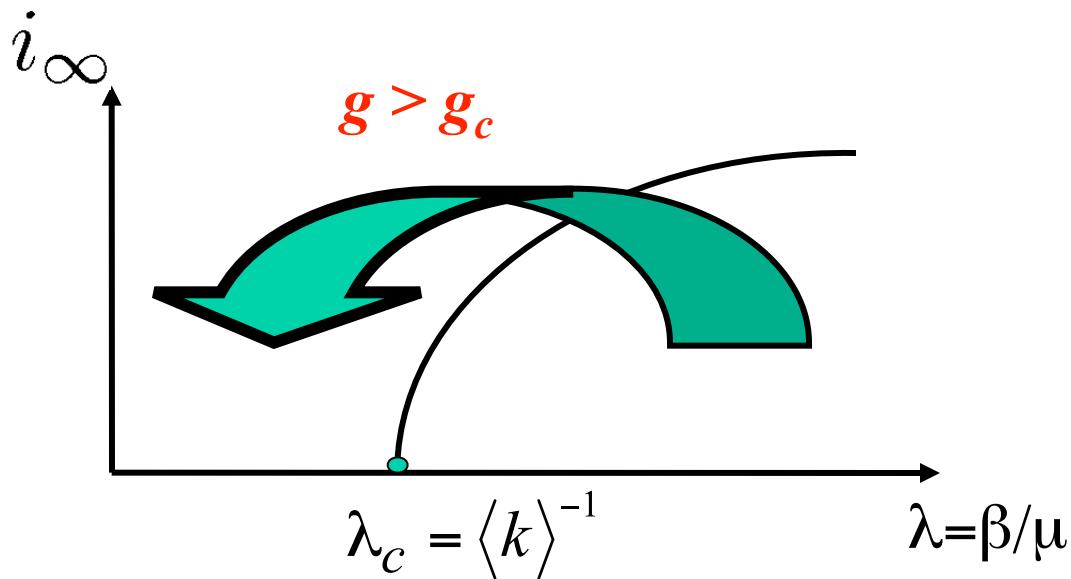
Immunization

Fraction g of immunized (vaccined) individuals: reduce the possibility of finding susceptible individuals

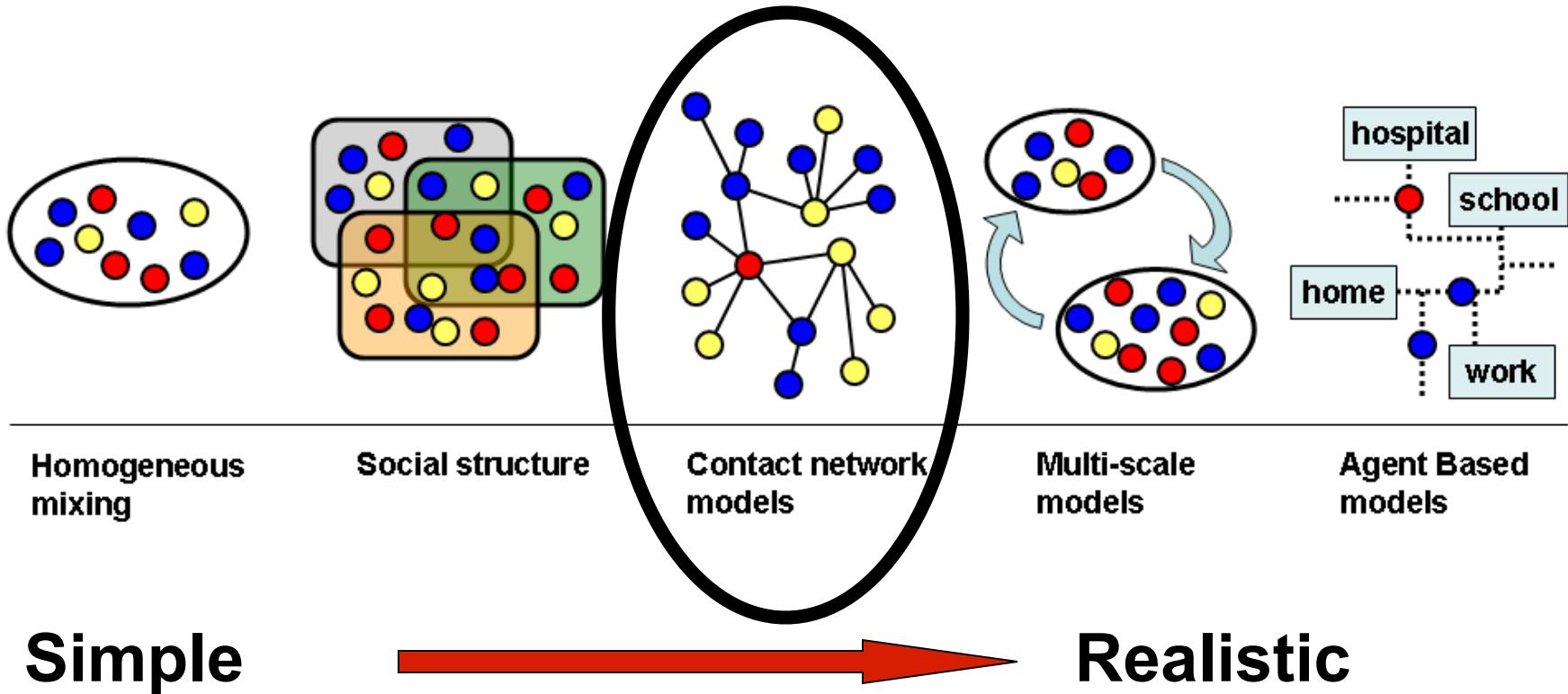
$$\beta \rightarrow \beta(1-g)$$

=> critical immunization threshold

$$g_c = 1 - \mu / (\beta \langle k \rangle)$$



Wide spectrum of complications and complex features to include...



Ability to explain trends at a population level

Model realism loses in transparency.
Validation is harder.

Complex networks

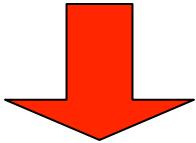
Viruses propagate on networks:

- Social (contact) networks
- Technological networks:
 - Internet, Web, P2P, e-mail...

...which are **complex, heterogeneous networks**

Epidemic spreading on heterogeneous networks

Number of contacts (degree) can vary a lot
huge fluctuations ($\langle k^2 \rangle \gg \langle k \rangle$)



Heterogeneous mean-field: density of

- Susceptible in the class of degree k , $s_k = S_k / N_k$
- Infectious in the class of degree k , $i_k = I_k / N_k$
- Recovered in the class of degree k , $r_k = R_k / N_k$

$$s(t) = \sum P(k) s_k, \quad i(t) = \sum P(k) i_k, \quad r(t) = \sum P(k) r_k$$

Epidemic spreading on heterogeneous networks

Relative density of infected nodes with given degree k : i_k

SIR model:

$$\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k - \mu i_k$$

Θ_k =Probab. that any given link points to an infected node,
i.e. density of infected neighbours

$$\Theta_k = \sum_{k'} P(k'|k) i_{k'}$$

Mean-Field

$P(k'|k)$ = the probability that a link originated in a node
with connectivity k points to a node with connectivity k'

SIS model on heterogeneous networks

$$\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k - \mu i_k \quad \Theta_k = \sum_{k'} P(k'|k)i_{k'}$$

In uncorrelated networks: $\Theta_k = \Theta = \sum_{k'} \frac{k'}{\langle k \rangle} P(k')i_{k'}$

Short times, $i_k(t) \ll 1$

$$\frac{d\Theta}{dt} = \left(\beta \frac{\langle k^2 \rangle}{\langle k \rangle} - \mu \right) \Theta$$

Epidemic threshold condition

$$\frac{\beta}{\mu} = \frac{\langle k \rangle}{\langle k^2 \rangle}$$

Long time limit, SIS model on het. networks

$$\frac{di_k}{dt} = 0 \rightarrow i_k(\infty) = \frac{\beta k \Theta_k(\infty)}{\beta k \Theta_k(\infty) + \mu}$$

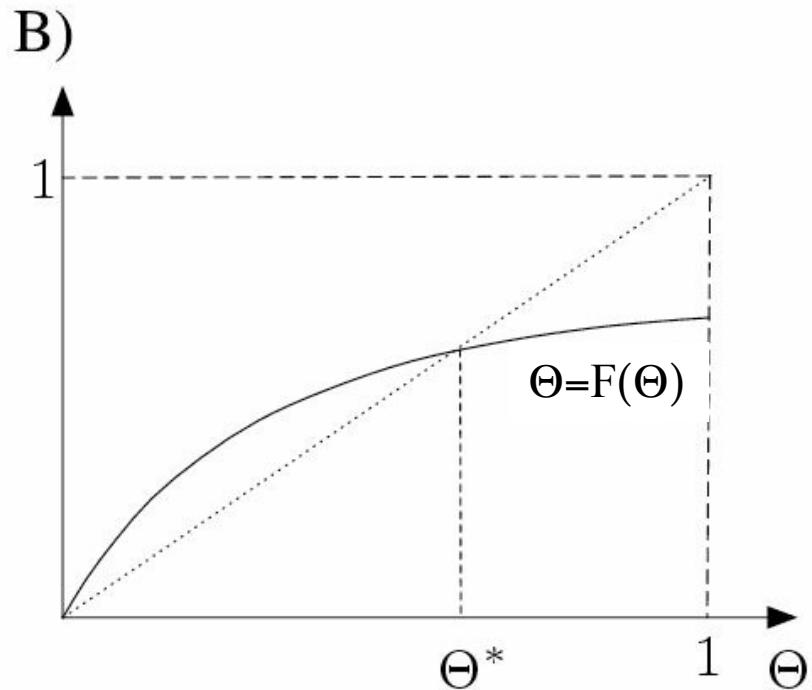
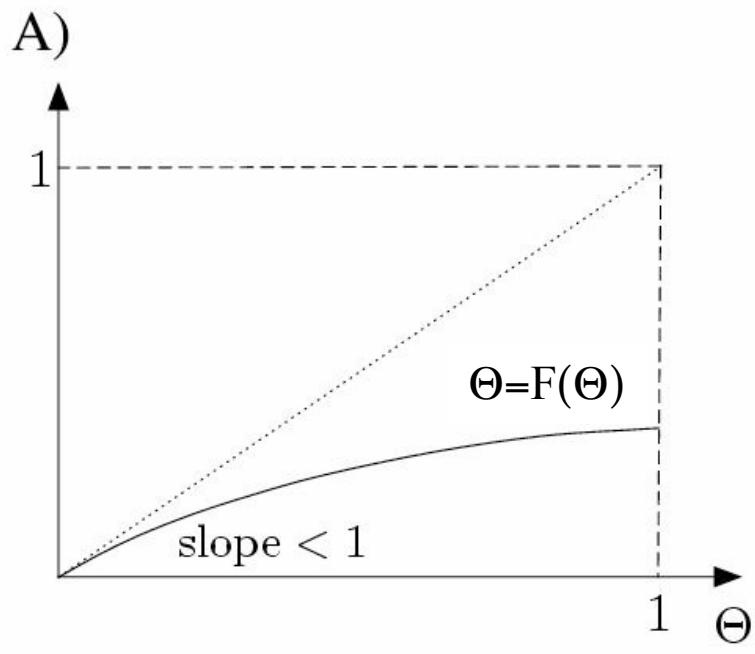
$$\Theta_k = \sum_{k'} P(k'|k) i_{k'}$$

In uncorrelated networks: $\Theta_k = \Theta = \sum_{k'} \frac{k'}{\langle k \rangle} P(k') i_{k'}$

 $\Theta(\infty) = \frac{1}{\langle k \rangle} \sum_k \frac{\beta k^2 P(k) \Theta(\infty)}{\beta k \Theta(\infty) + \mu}$

Self-consistent equation of the form $x=F(x)$
with $F(0)=0$, $F' > 0$, $F'' < 0$

Graphical solution



Epidemic threshold:

existence of a non-zero solution for $\Theta \Leftrightarrow F'(0) > 1$:

$$\sum_k \frac{\beta k^2 P(k)}{\mu \langle k \rangle} > 1 \quad \Leftrightarrow \quad \frac{\beta \langle k^2 \rangle}{\mu \langle k \rangle} > 1$$

Epidemic threshold in uncorrelated networks

$$\frac{\beta}{\mu} = \frac{\langle k \rangle}{\langle k^2 \rangle}$$

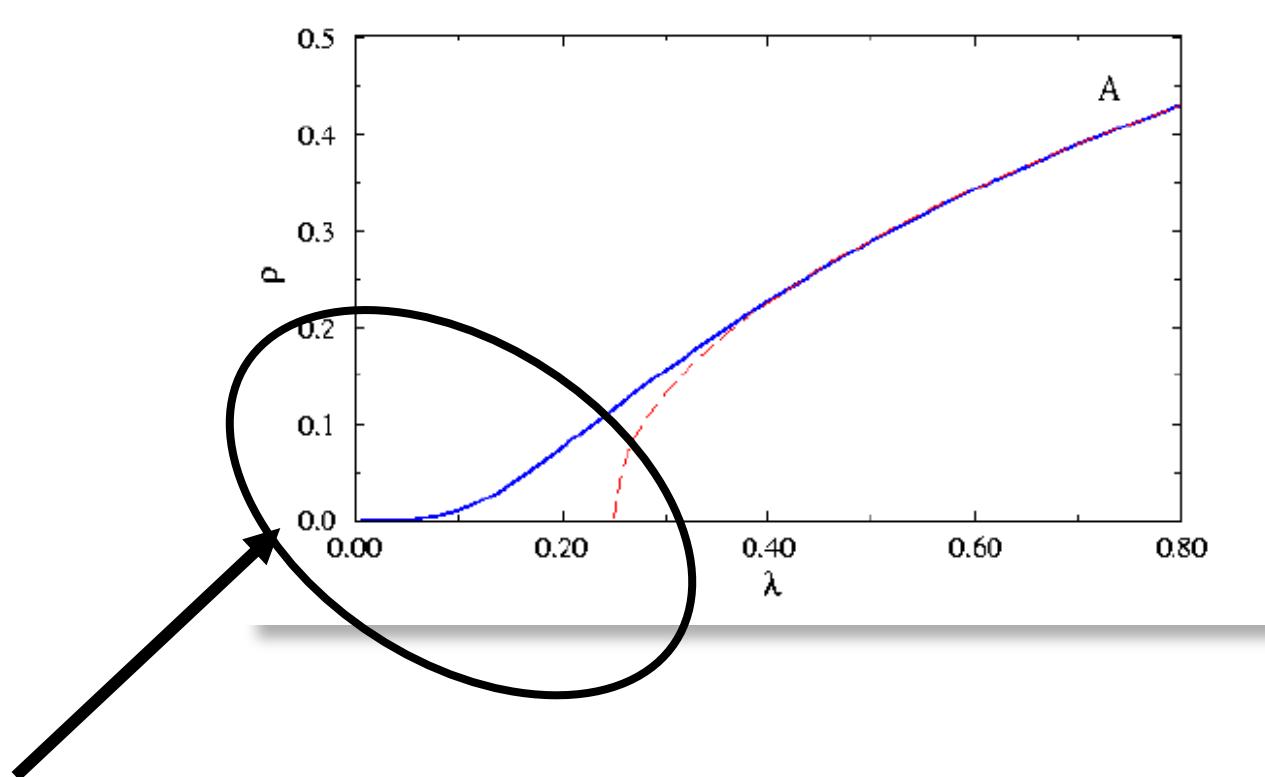
Heterogeneous, infinite network:

$$\langle k^2 \rangle \rightarrow \infty$$

Condition always satisfied

Finite prevalence for any spreading parameters

Epidemic phase diagram in heterogeneous networks



- Wide range of spreading rate with low prevalence
- Lack of healthy phase = standard immunization cannot drive the system below threshold!!!

Case of correlated networks

SIS model: $\frac{di_k}{dt} = \beta k(1 - i_k)\Theta_k - \mu i_k \quad \Theta_k = \sum_{k'} P(k'|k)i_{k'}$

Short times:

$$\frac{di_k}{dt} \sim \sum_{k'} L_{kk'} i_{k'} \quad L_{kk'} = -\mu\delta_{kk'} + \beta k P(k'|k)$$

Solution $i_k=0$ unstable if there exists at least one positive eigenvalue

Λ_m largest eigenvalue of $C_{kk'} = kP(k'|k)$ Connectivity matrix

Epidemic threshold:

$$\frac{\beta}{\mu} = \frac{1}{\Lambda_m}$$

diverges in Markovian scale-free networks

MODEL	CONTINUUM EQUATION	τ	λ_c
SI	$\frac{di_k}{dt} = \beta[1 - i_k]k\theta_k$	$\frac{\langle k \rangle}{\beta(\langle k^2 \rangle - \langle k \rangle)}$	0
SIS	$\frac{di_k}{dt} = \beta[1 - i_k]k\theta_k - \mu i_k$	$\frac{\langle k \rangle}{\beta\langle k^2 \rangle - \mu\langle k \rangle}$	$\frac{\langle k \rangle}{\langle k^2 \rangle}$
SIR	$\frac{di_k}{dt} = \beta s_k \theta_k - \mu i_k$ $s_k = 1 - i_l - r_k$	$\frac{\langle k \rangle}{\beta\langle k^2 \rangle - (\mu + \beta)\langle k \rangle}$	$\frac{1}{\langle k^2 \rangle} - 1$

What does HMF neglect

1. Structural correlations in the network

(HMF equivalent to an annealed network approximation)

2. Dynamical correlations

(emerging during the spreading process)

Beyond HMF

Chatterjee, S. & Durrett, R. Contact processes on random graphs with power law degree distributions have critical value 0. Annals of Probability 37, 2332–2356 (2009).

R. Durrett, Some features of the spread of epidemics and information on a random graph, Proc. Natl. Acad. Sci. USA 107, 4491 (2010)

=> *rigorous proof* that, for strictly infinite system size, the epidemic threshold, in networks with power-law degree distribution, is exactly 0 for *any exponent γ* of the degree distribution

(HMF: epidemic threshold > 0 if $\gamma > 3$)

Beyond HMF

Chakrabarti et al., Epidemic Thresholds in Real Networks, ACM Trans. Inf. Syst. Secur. 10, 1 (2008)

Gomez et al., Discrete-time Markov chain approach to contact-based disease spreading in complex networks, EPL 89 38009 (2010)

“Quenched” mean field theory:

- write the evolution equation for the probability that a node i is infected
- take into account the real connections of the network as given by the adjacency matrix
- check for the stability of the absorbing state of zero infection

=> epidemic threshold for SIS given by

$$\lambda_c = 1/\Lambda_m$$

where Λ_m is the largest eigenvalue of the adjacency matrix

Beyond HMF

For random scale-free networks, it is possible to obtain the scaling of Λ_m (Chung, Lu, Vu, Proc. Natl. Acad. Sci. USA 100, 6313 (2003))

=> epidemic threshold for SIS given by

$$\lambda_c = \begin{cases} 1/\sqrt{k_{max}} & \text{if } \gamma > 5/2 \\ \langle k \rangle / \langle k^2 \rangle & \text{if } 2 < \gamma < 5/2 \end{cases}$$

=> epidemic threshold vanishes in the thermodynamic limit in power-law distributed networks

- for any value of γ , even larger than 3,
- as long as k_{max} is a diverging function of the network size N

=> role of the hubs

Beyond HMF

Kitsak et al., Identification of influential spreaders in complex networks. Nat. Phys. 6, 888–893 (2010)

Numerical study => most efficient spreaders are located at the innermost, dense core of the network, as identified by means of a k-core decomposition

Beyond HMF

Castellano & Pastor-Satorras, Competing activation mechanisms in epidemics on networks, Scientific Reports 2, 371 (2012)

=> numerical investigation, measure of:

- density of infected vertices in the whole network
- density of infected when the dynamics takes place (in isolation) on the **k-core of highest index** (maximum k-core)
- density of infected when the dynamics takes place (in isolation) on the star-graph centered around the **hub** of the network, with largest degree

Immunization strategies

Uniform immunization:

Fraction g of randomly chosen immunized (vaccined) individuals:

$$\beta \rightarrow \beta(1-g)$$

=> inefficient: need

$$(1 - g) \frac{\beta}{\mu} < \frac{\langle k \rangle}{\langle k^2 \rangle}$$

$$g > g_c = 1 - \frac{\mu}{\beta} \frac{\langle k \rangle}{\langle k^2 \rangle}$$

close to 1

Proportional immunization

g_k fraction of immunized individuals of degree k , such that:

$$\beta k(1 - g_k) = \beta' = cst$$



$$\frac{di_k}{dt} = \beta'(1 - i_k)\Theta_k - \mu i_k$$

Short times (uncorr. nets):

$$\frac{d\Theta}{dt} = (\beta' - \mu) \Theta \quad \text{Epidemic threshold recovered!}$$

Efficient immunization: need

$$g_k > 1 - \frac{\mu}{\beta k}$$

Targeted immunization

=> immunize fraction g of individuals with largest connectivity

need: $\frac{\langle k \rangle_g}{\langle k^2 \rangle_g} > \frac{\beta}{\mu}$

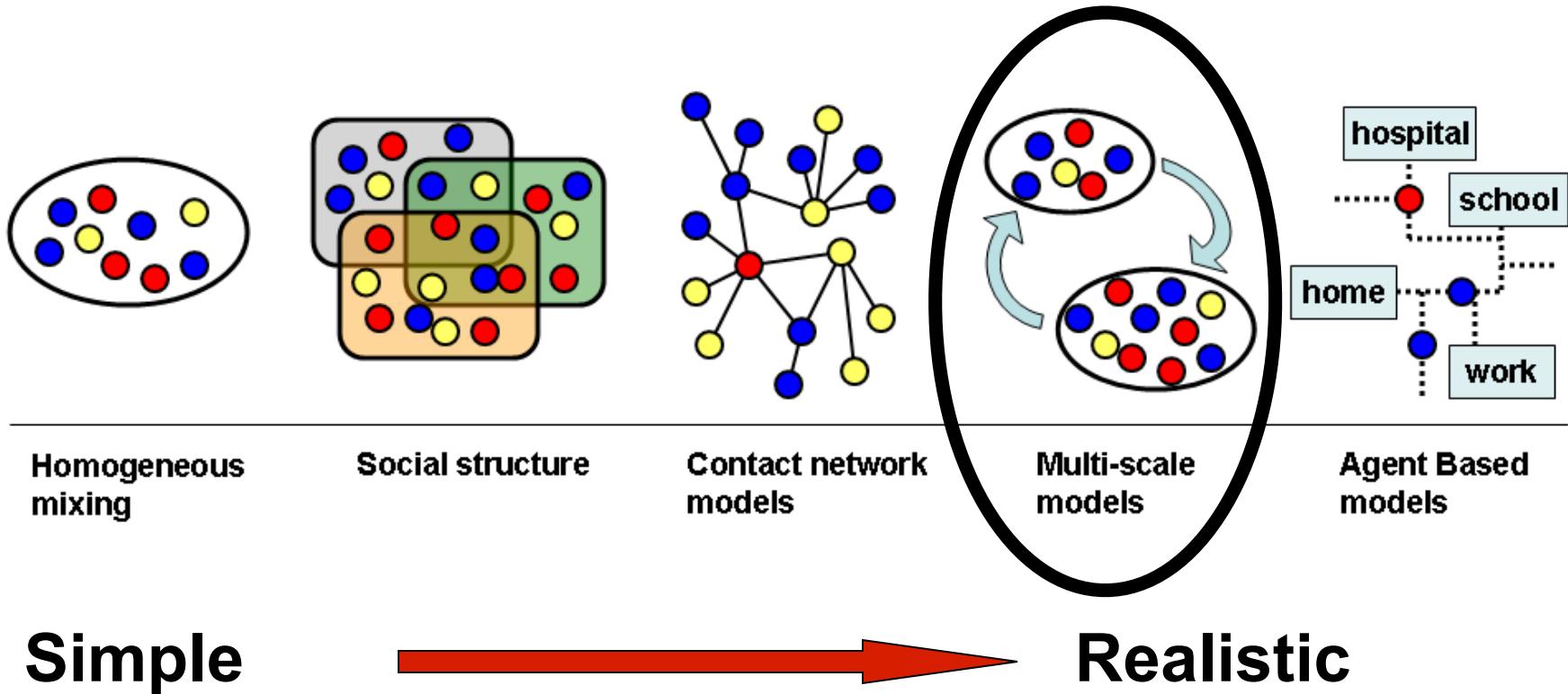
cf targeted attacks!!!

immunizing \Leftrightarrow removing nodes and links

Ex of explicit calculation for BA network:

$$g_c \propto \exp(-2\mu/m\beta)$$

Wide spectrum of complications and complex features to include...



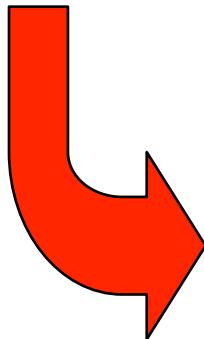
Ability to explain trends at a population level

Model realism loses in transparency.
Validation is harder.

General framework:

Bosonic reaction-diffusion processes

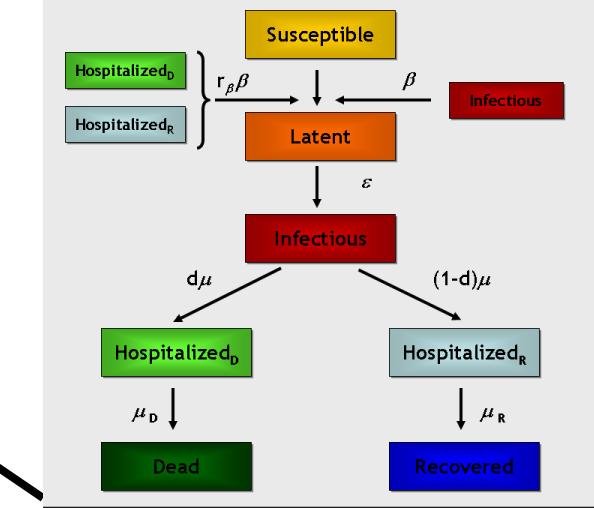
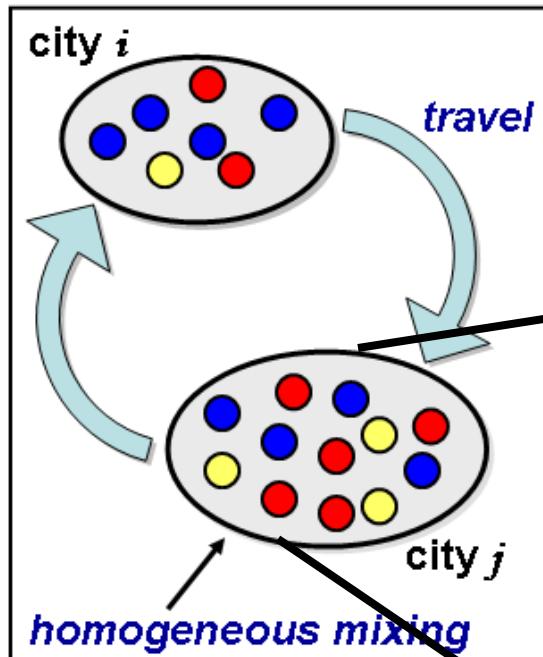
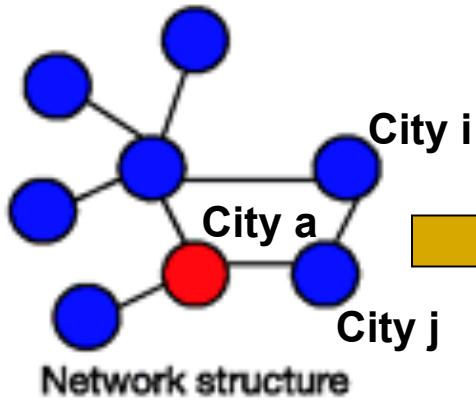
- Previous cases: (at most) one particle/individual *per site*
- In general: reaction-diffusion processes on networks
=> no restriction on the number of particles per site



“Particles”

- diffusing along edges
- reacting in the nodes

Example: Meta-population models



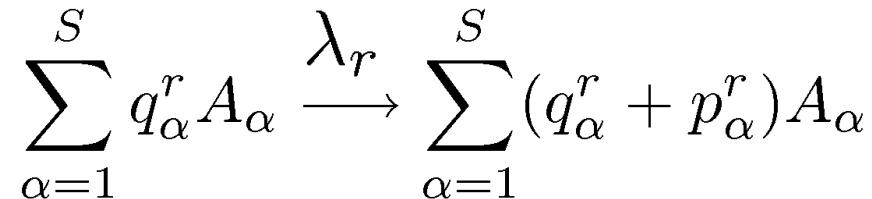
Intra-population infection dynamics by stochastic compartmental modeling

Baroyan *et al.* (1969)

Ravchev, Longini (1985)

Bosonic RD processes on complex networks

- A_α , $\alpha=1,\dots,S$ types of particles
- Diffusion coefficients D_α
- Reactions ($r=1,\dots,R$):



Bosonic RD processes on complex networks

Heterogeneous mean-field formalism:

Densities

$$\rho_{\alpha,k} \equiv \frac{n_{\alpha,k}}{N_k} = \sum_{i \in k} \frac{n_{\alpha,i}}{N_k}$$

$$\rho_{\alpha}(t) = \sum_k \rho_{\alpha,k}(t) P(k)$$

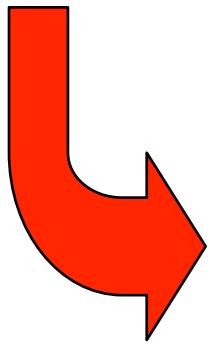
$$\partial_t \rho_{\alpha,k}(t) = \mathcal{D}_{\alpha} + \mathcal{R}_{\alpha}$$



Bosonic RD processes on complex networks

Diffusion term:

$$\text{For site } i: -D_\alpha n_{\alpha,i}(t) + D_\alpha \sum \frac{x_{ji}}{k_j} n_{\alpha,j}(t)$$



$$\mathcal{D}_\alpha = -D_\alpha \rho_{\alpha,k}(t) + D_\alpha k \sum_{k'} \frac{P(k'|k)}{k'} \rho_{\alpha,k'}(t)$$

Bosonic RD processes on complex networks

$$\sum_{\alpha=1}^S q_\alpha^r A_\alpha \xrightarrow{\lambda_r} \sum_{\alpha=1}^S (q_\alpha^r + p_\alpha^r) A_\alpha$$

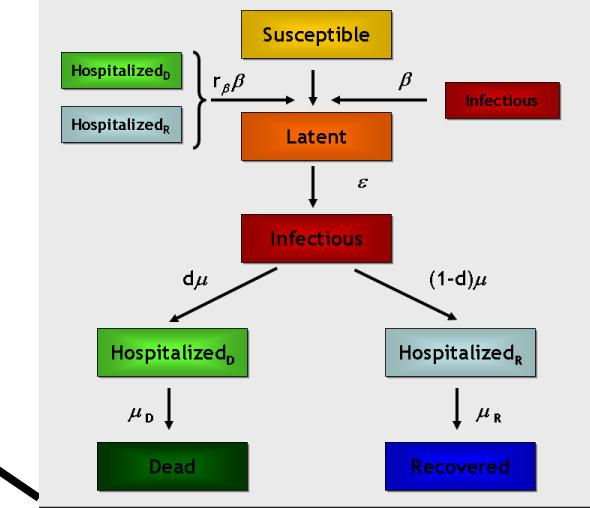
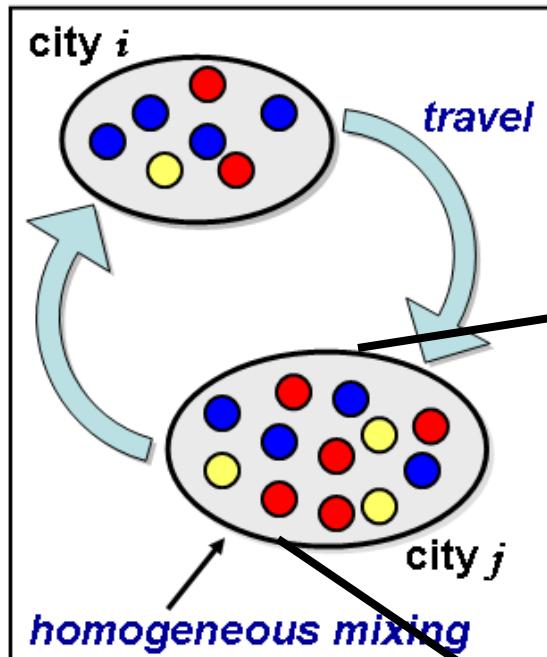
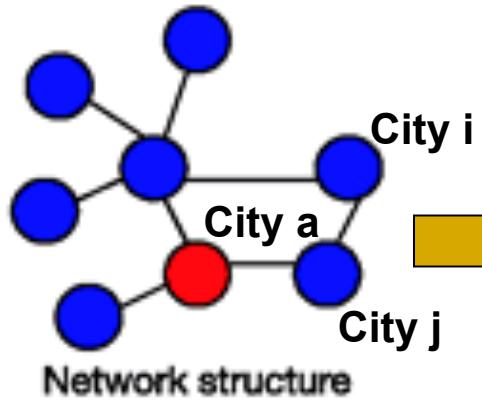


Reaction term: $\mathcal{R}_\alpha = \sum_r p_\alpha^r \lambda_r \prod_r (\rho_{\beta,k})^{q_\beta^r}$

See Baronchelli et al, Phys. Rev. E 78, 016111 (2008),
Colizza, Pastor-Satorras, Vespignani, Nature Phys (2007)
for various examples of further computations

A concrete example:
epidemic meta-population models

Meta-population models



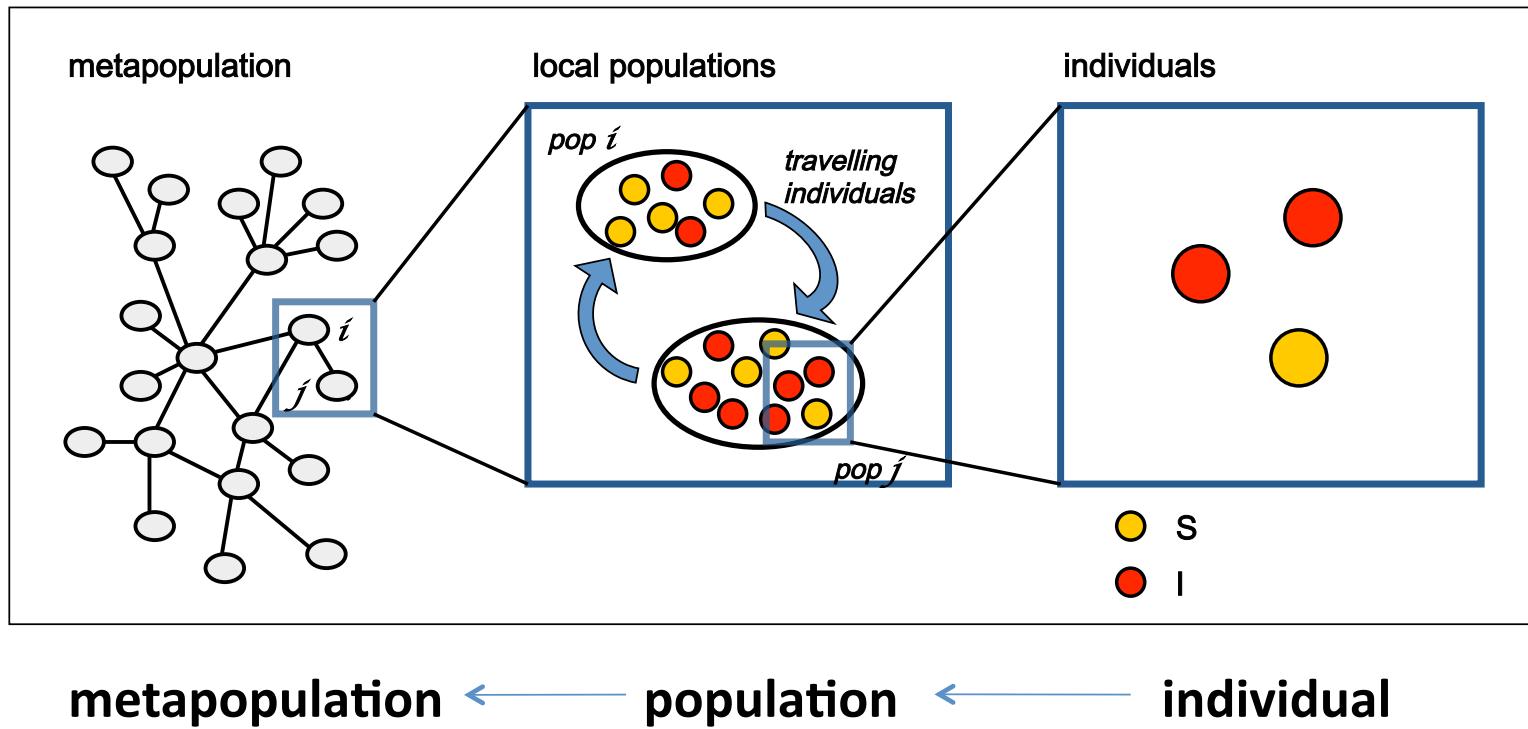
Intra-population infection dynamics by stochastic compartmental modeling

Baroyan *et al.* (1969)

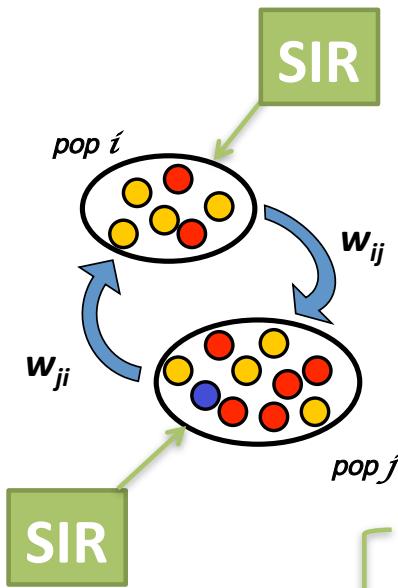
Ravchev, Longini (1985)

Multi-scale models

...or metapopulation models



Metapopulation models; SIR



$$S(t)$$

$$I(t)$$

$$R(t)$$

$$N(t)$$

V: # populations

$$S_i(t)$$

$$I_i(t)$$

$$R_i(t)$$

$$N_i(t) = S_i(t) + I_i(t) + R_i(t)$$

Global variables:

$$S(t) = S_1(t) + S_2(t) + S_3(t) + \dots + S_V(t) = \sum_i S_i(t)$$

$$I(t) = I_1(t) + I_2(t) + I_3(t) + \dots + I_V(t) = \sum_i I_i(t)$$

$$R(t) = R_1(t) + R_2(t) + R_3(t) + \dots + R_V(t) = \sum_i R_i(t)$$

$$N(t) = N_1(t) + N_2(t) + N_3(t) + \dots + N_V(t) = \sum_i N_i(t)$$

Metapopulation model. SIR

$$\Delta t = 1$$

$$S_i(t+1) = S_i(t) - \beta \frac{I_i(t)S_i(t)}{N_i} + \dots$$

$$I_i(t+1) = I_i(t) + \beta \frac{I_i(t)S_i(t)}{N_i} - \mu I_i(t) + \dots$$

$$R_i(t+1) = R_i(t) + \mu I_i(t) + \dots$$

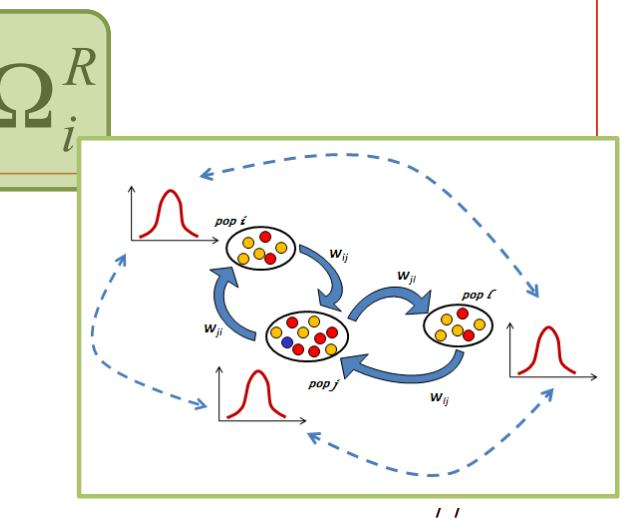
Metapopulation model. SIR

$$S_i(t+1) = S_i(t) - \beta \frac{I_i(t)S_i(t)}{N_i} + \Omega_i^S$$

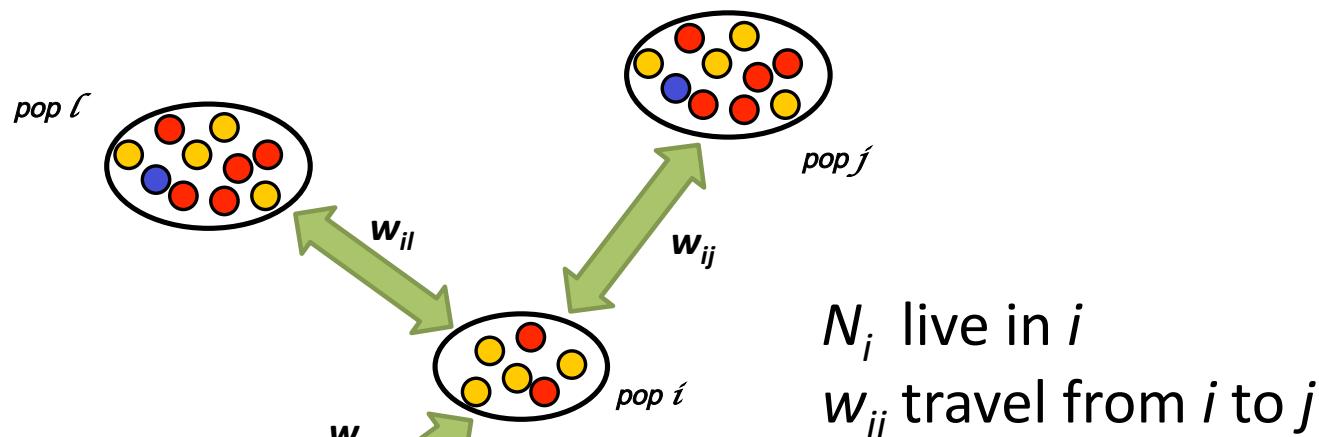
$$I_i(t+1) = I_i(t) + \beta \frac{I_i(t)S_i(t)}{N_i} - \mu I_i(t) + \Omega_i^I$$

$$R_i(t+1) = R_i(t) + \mu I_i(t) + \Omega_i^R$$

Ω_i^X Measure of *in-flow* and *out-flow* of people in compartment X



Metapopulation model. Coupling

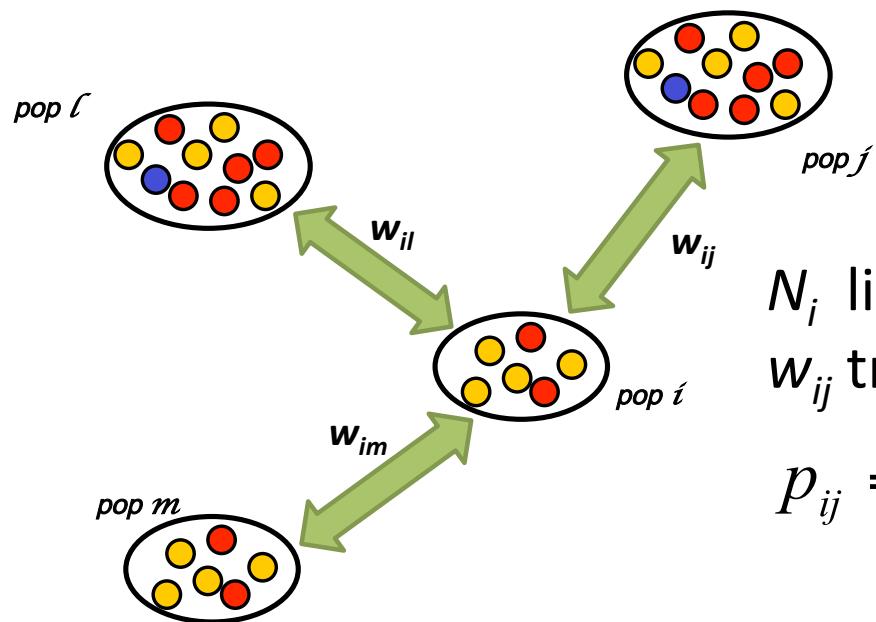


probability for an individual in i
to travel from i to j ???

$$p_{ij} = \frac{w_{ij}}{N_i}$$

for all compartments

Metapopulation model. Coupling



N_i live in *i*

w_{ij} travel from *i* to *j*

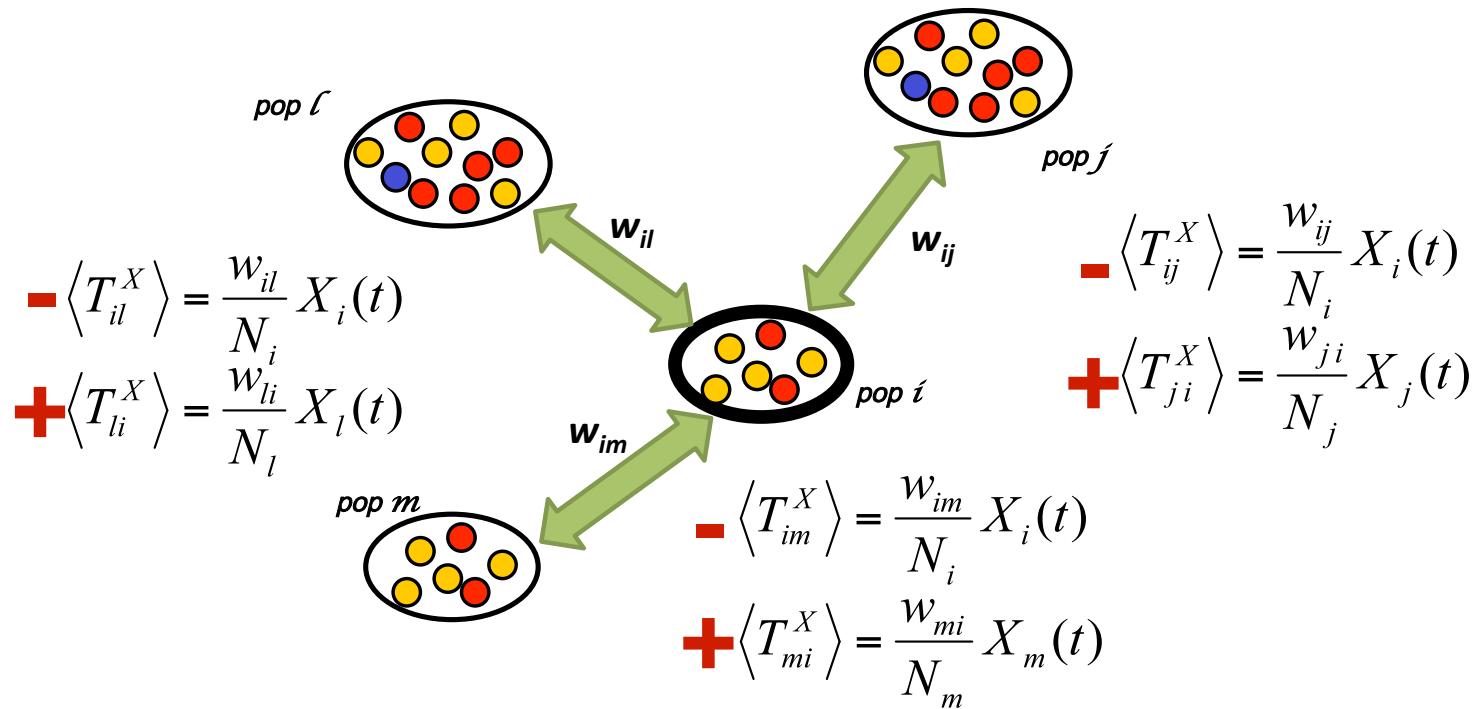
$$p_{ij} = \frac{w_{ij}}{N_i} \quad \text{probability travel from } i \text{ to } j$$

average number of individuals in compartment *X* in *i* traveling from *i* to *j* ???

travelers

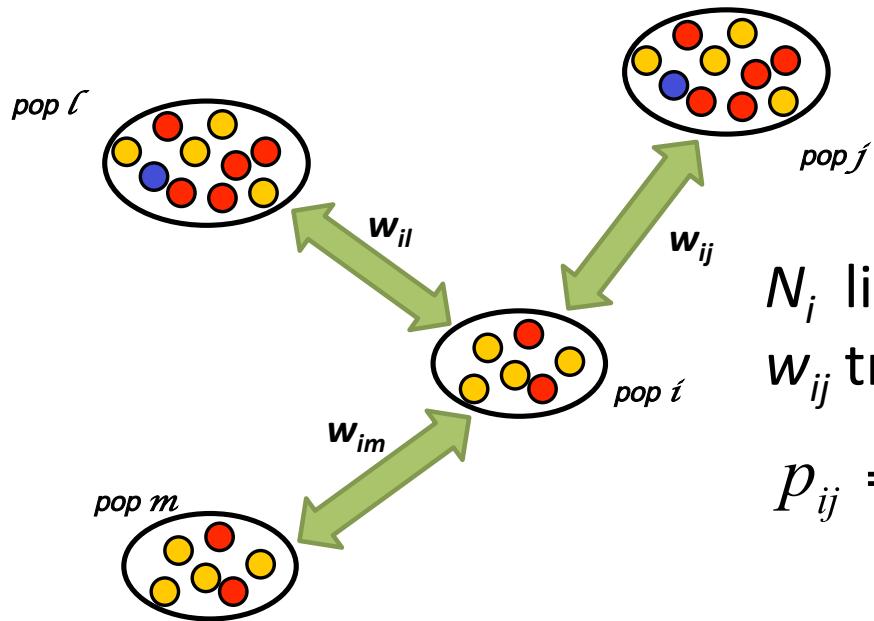
$$\langle T_{ij}^X \rangle = p_{ij} X_i(t) = \frac{w_{ij}}{N_i} X_i(t)$$

Metapopulation model. Coupling



$$\Omega_i^X = \sum_j \left(\frac{w_{ji}}{N_j} X_j - \frac{w_{ij}}{N_i} X_i \right)$$

Deterministic → stochastic metapop.



N_i live in *i*

w_{ij} travel from *i* to *j*

$$p_{ij} = \frac{w_{ij}}{N_i} \quad \text{probability travel from } i \text{ to } j$$

~~average number of individuals in compartment *X* in *i* traveling from *i* to *j* ???~~

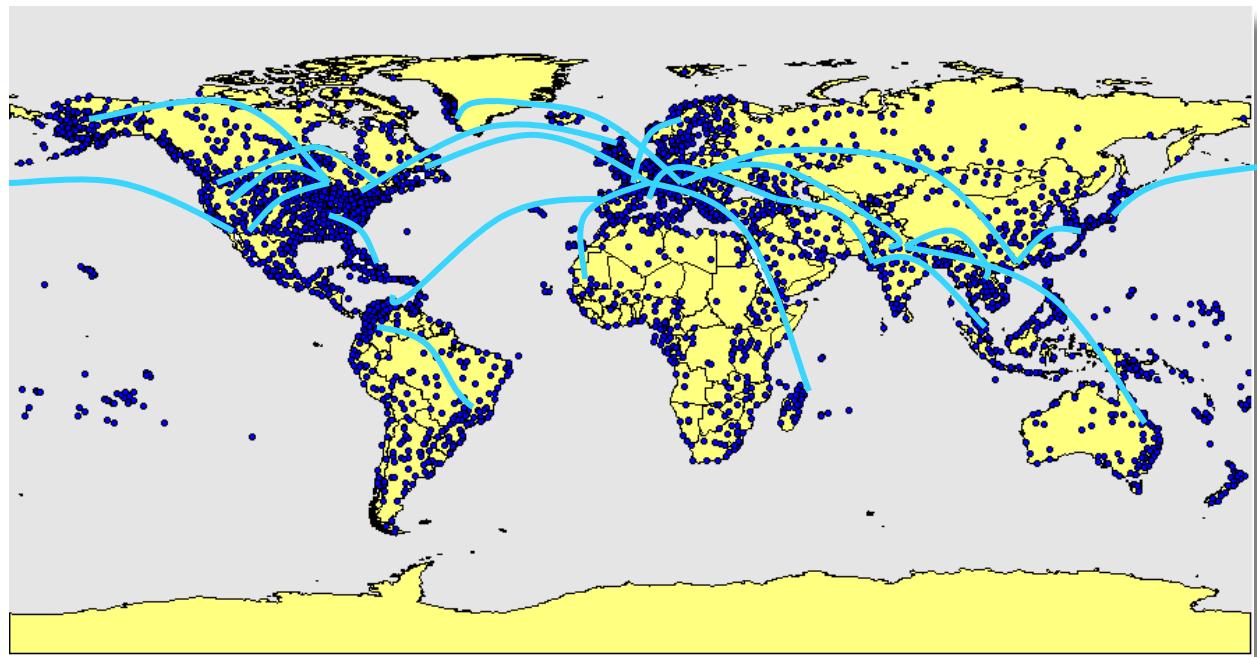
travelers

$$T_{ij}^X = \text{binom}\left(X_i, \frac{w_{ij}}{N_i}\right)$$

Modeling of global epidemics propagation

multi-level description :

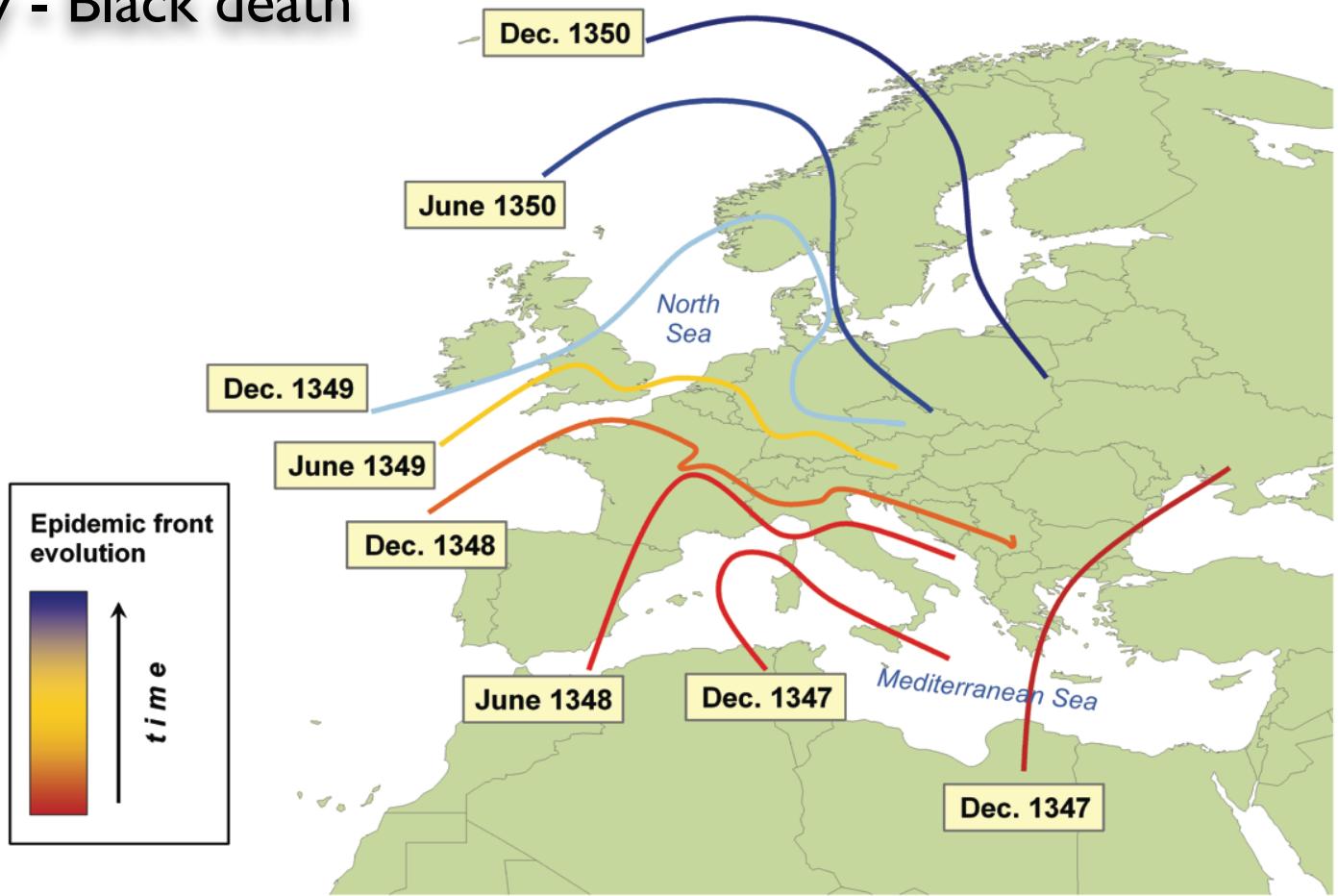
- intra-city
epidemics
- inter-city
travel



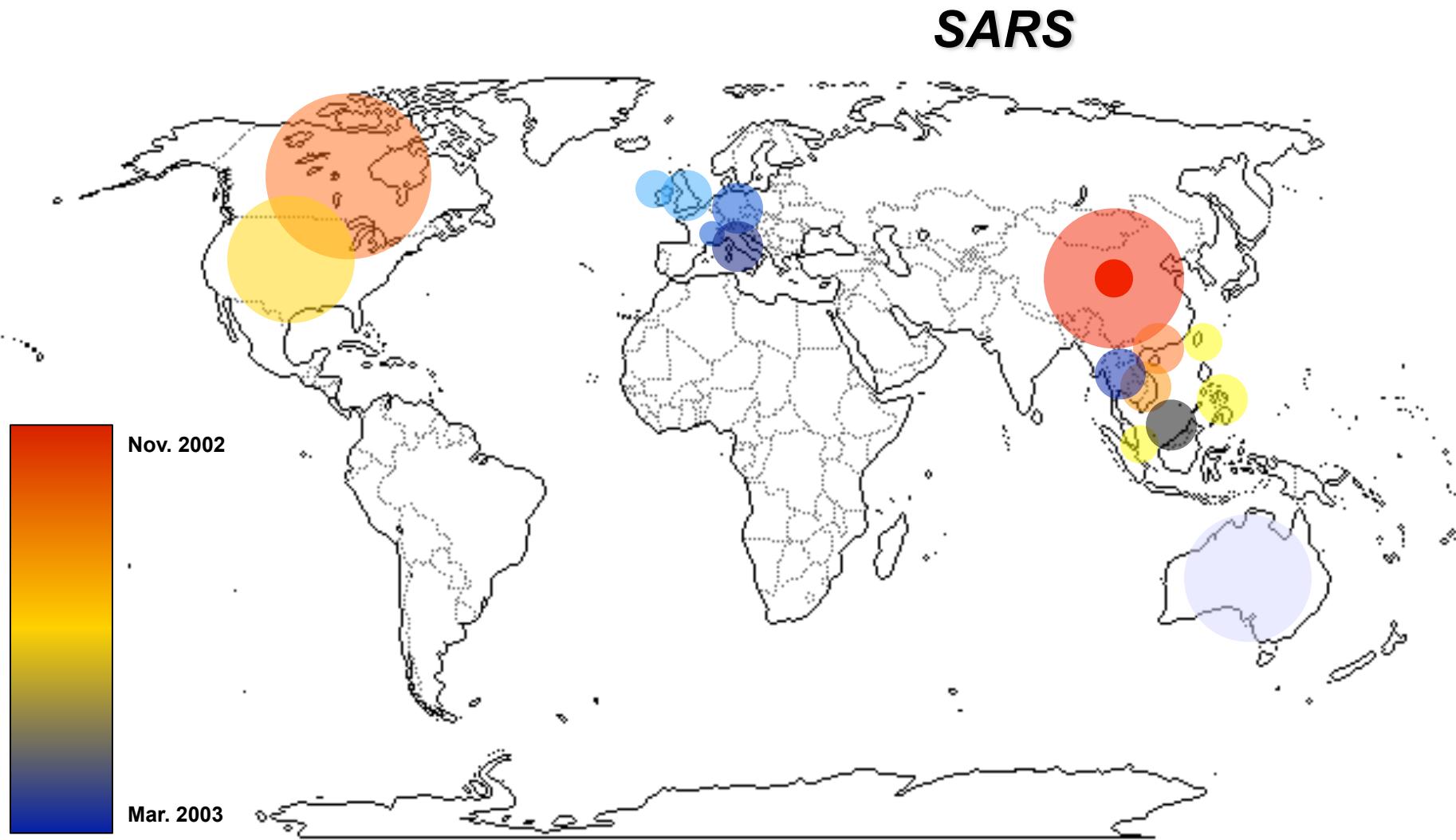
Baroyan *et al.* (1969)
Ravchev, Longini (1985)

Why is a large-scale approach needed?

14th century - Black death

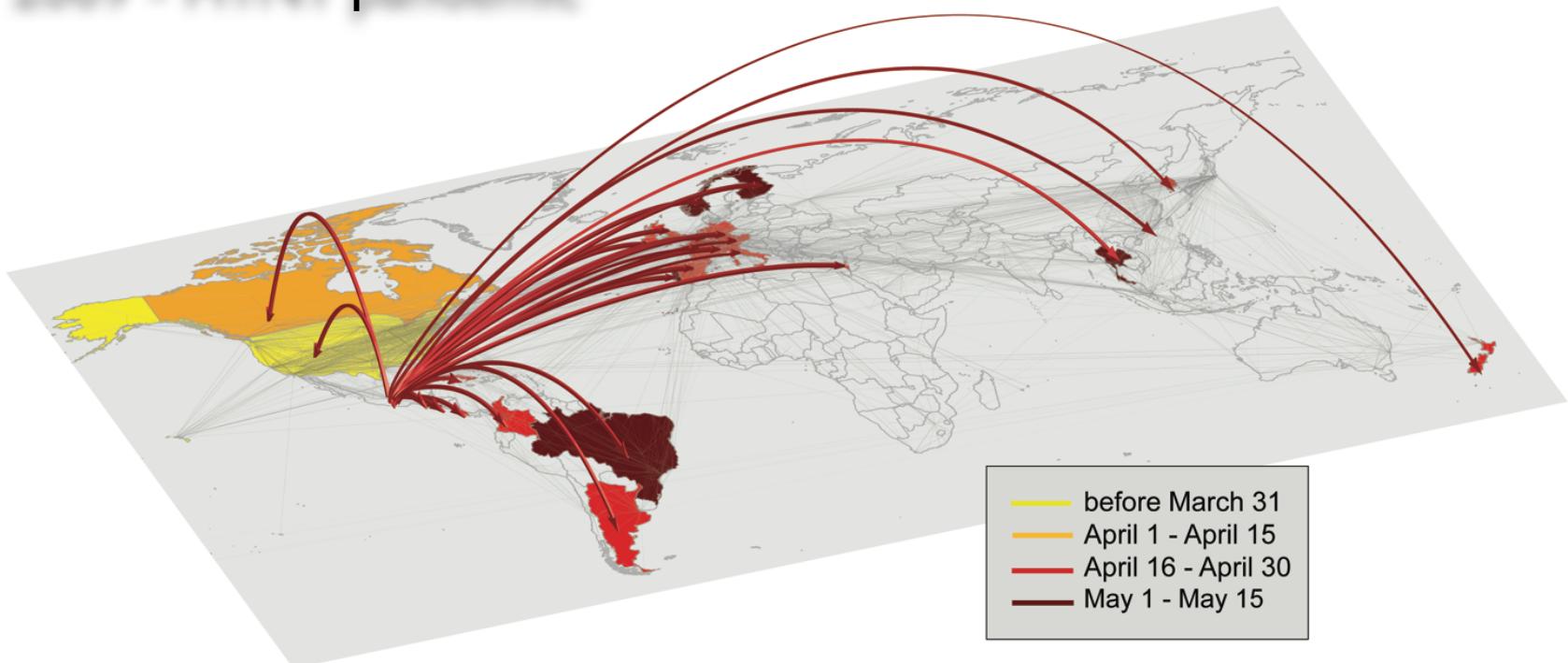


Why is a large-scale approach needed?



Why is a large-scale approach needed?

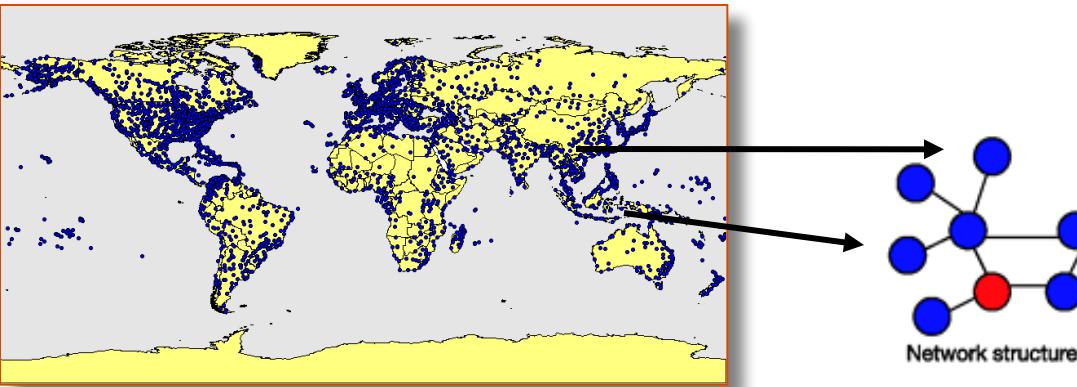
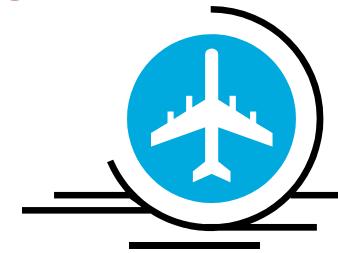
2009 - HINI pandemic



What has changed: availability of unprecedented amounts of data.....

- Transportation infrastructures
 - Behavioral Networks
 - Census data
 - Commuting/traveling patterns
-
- Different scales:
 - International
 - Intra-nation (county/city/municipality)
 - Intra-city (workplace/daily commuters/individuals behavior)

Global spread of epidemics on the airport infrastructure



Urban areas
+
Air traffic flows

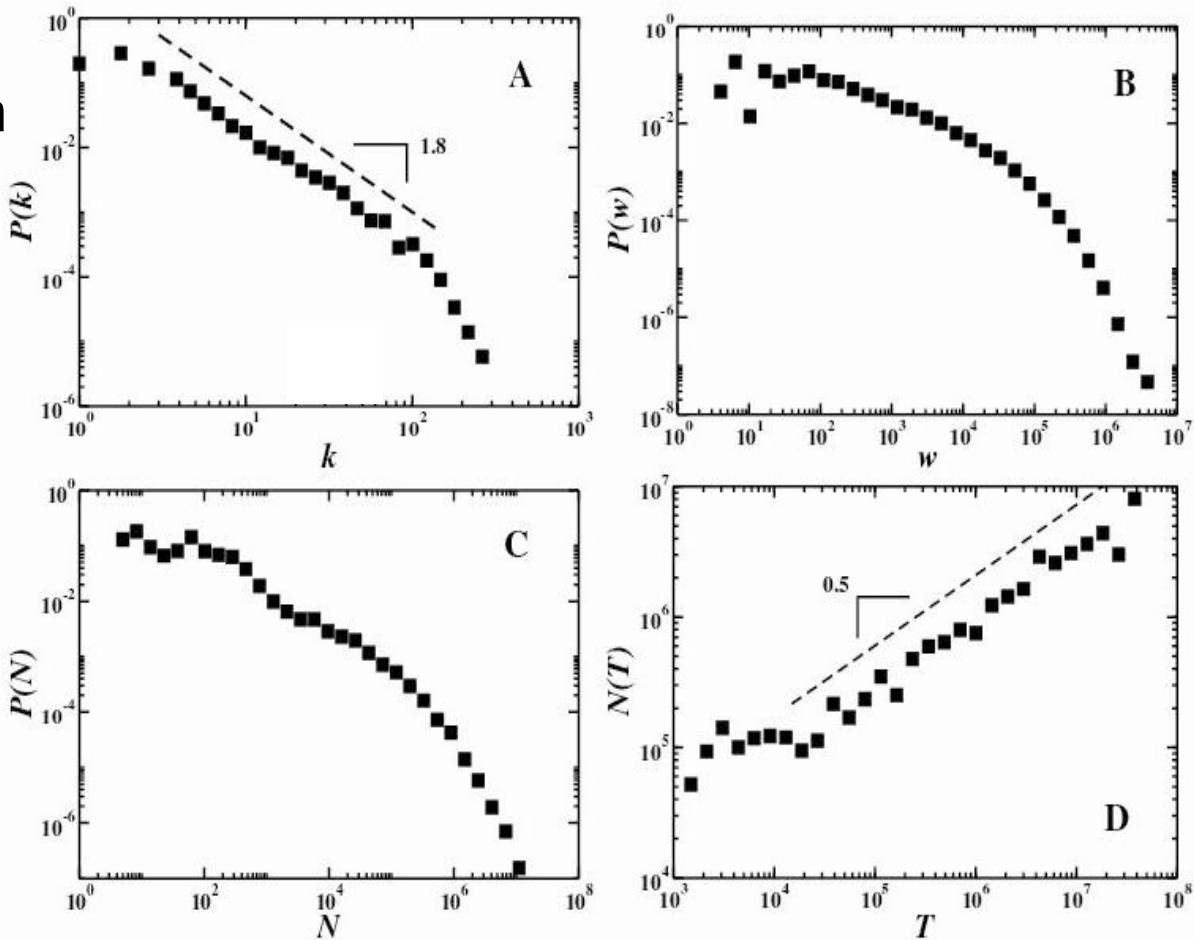
World-wide airport network

- complete IATA database
 - $V = \underline{3100}$ airports
 - $E = \underline{17182}$ weighted edges
 - w_{ij} #seats / (different time scales)
- N_j urban area population
(UN census, ...)

> 99% of total traffic

Statistical distributions...

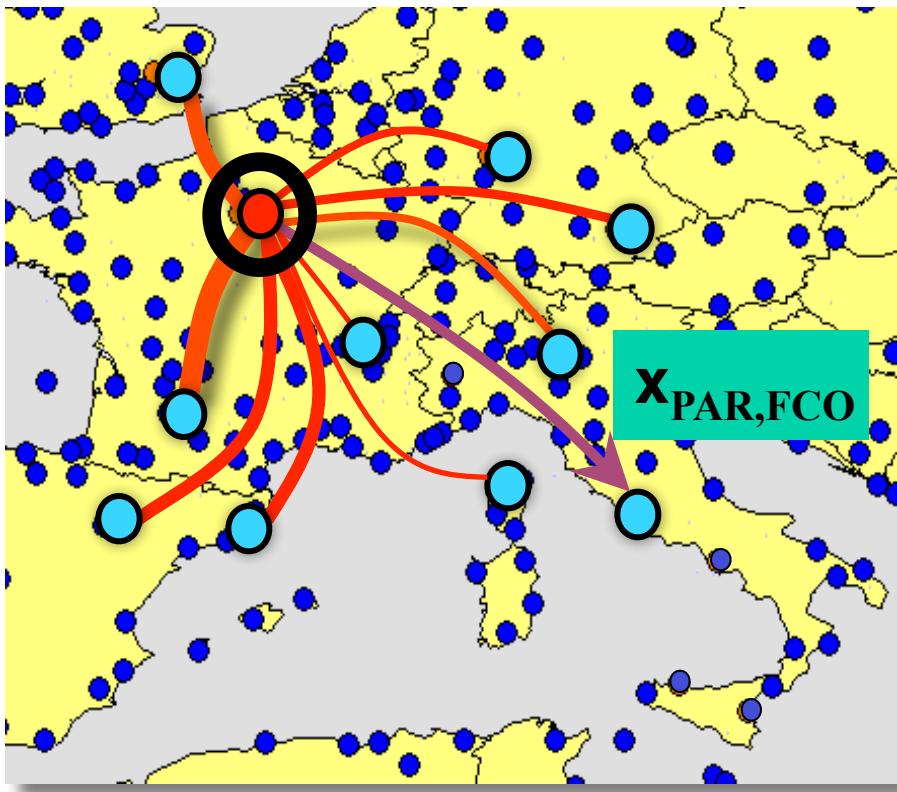
- Skewed
- Heterogeneity and high variability
- Very large fluctuations (variance>>average)



Barrat, Barthélémy, Pastor-Satorras, Vespignani. *PNAS* (2004)

Colizza, Barrat, Barthélémy, Vespignani. *PNAS* (2006)

Stochastic model: travel term

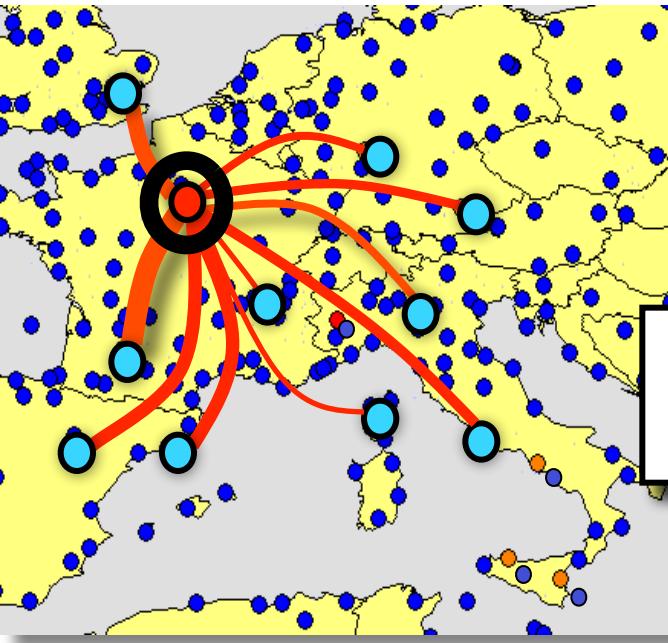


Travel probability
from PAR to FCO:

$$P_{PAR,FCO} = \frac{w_{PAR,FCO}}{N_{PAR}} \Delta t$$

$w_{PAR,FCO}$ # passengers
from PAR to FCO:
Stochastic variable,
multinomial distr.

Stochastic model: travel term



Transport operator:

$$\Omega_{\text{PAR}}(\{X\}) = \sum_l (\Omega_{l,\text{PAR}}(\{X_l\}) - \Omega_{\text{PAR},l}(\{X_{\text{PAR}}\}))$$

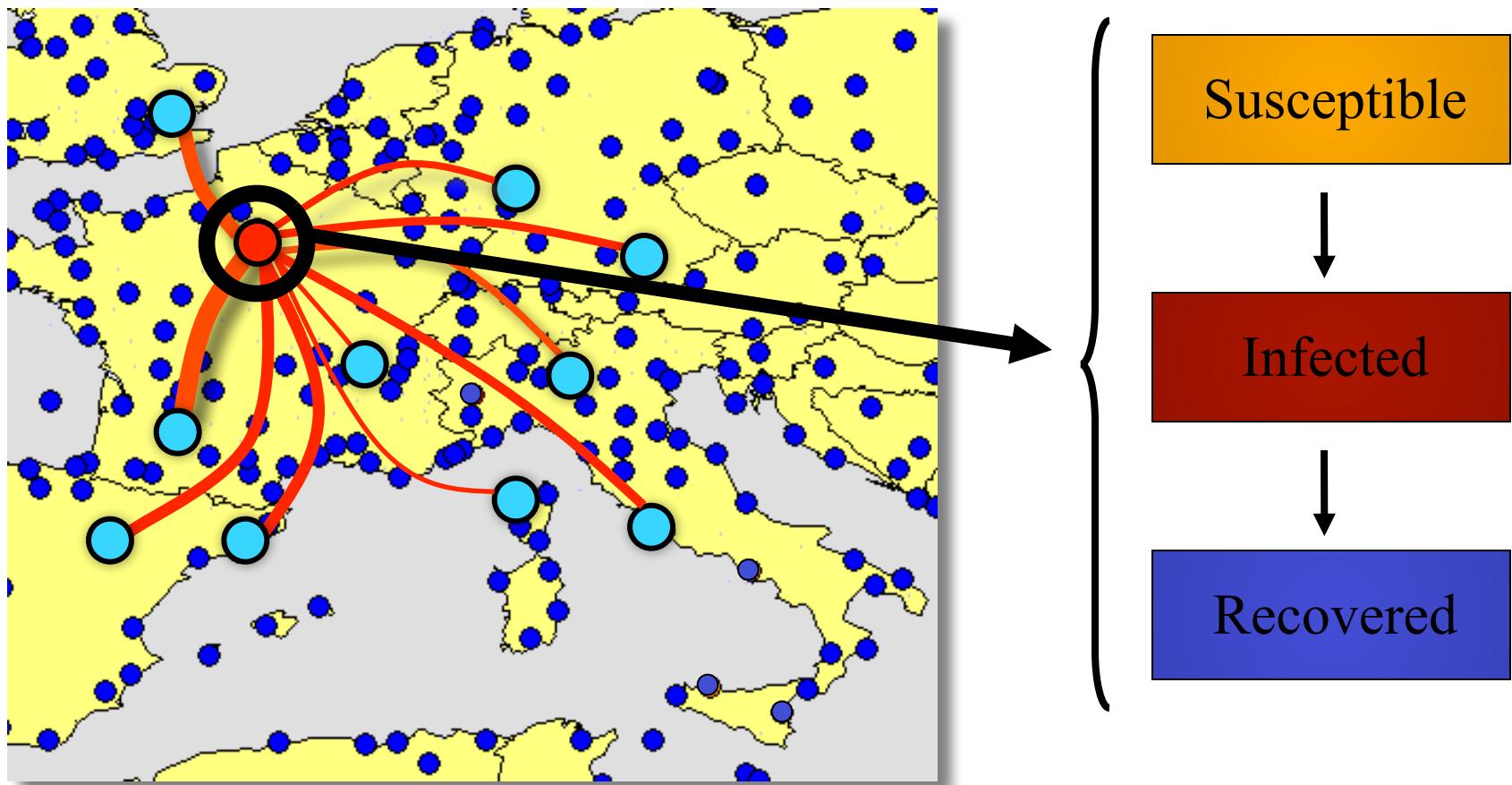
ingoing

outgoing

- other source of **noise**: $w_{jl}^{\text{noise}} = w_{jl} [\alpha + \eta(1-\alpha)]$ $\alpha=70\%$
- **two-legs travel**: $\Omega_j(\{X\}) = \Omega_j^{(1)}(\{X\}) + \Omega_j^{(2)}(\{X\})$

Stochastic large-scale model

compartmental model + air transportation data

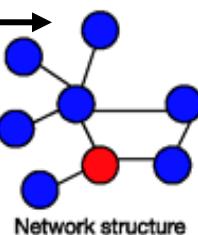
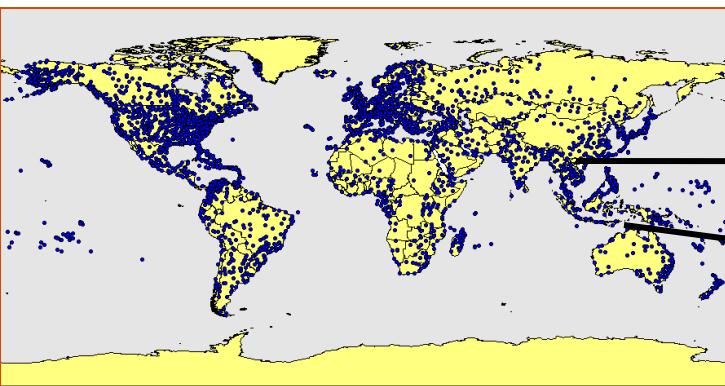


Large-scale model

Stochastic evolution equations describing the disease evolution at a mean-field level in each city

...

Coupled through transport operators



- complete IATA database
 - $V = \underline{3100}$ airports
 - $E = \underline{17182}$ weighted edges
 - W_{ij} #seats / (different time scales)
- N_j urban area population (UN census, ...)

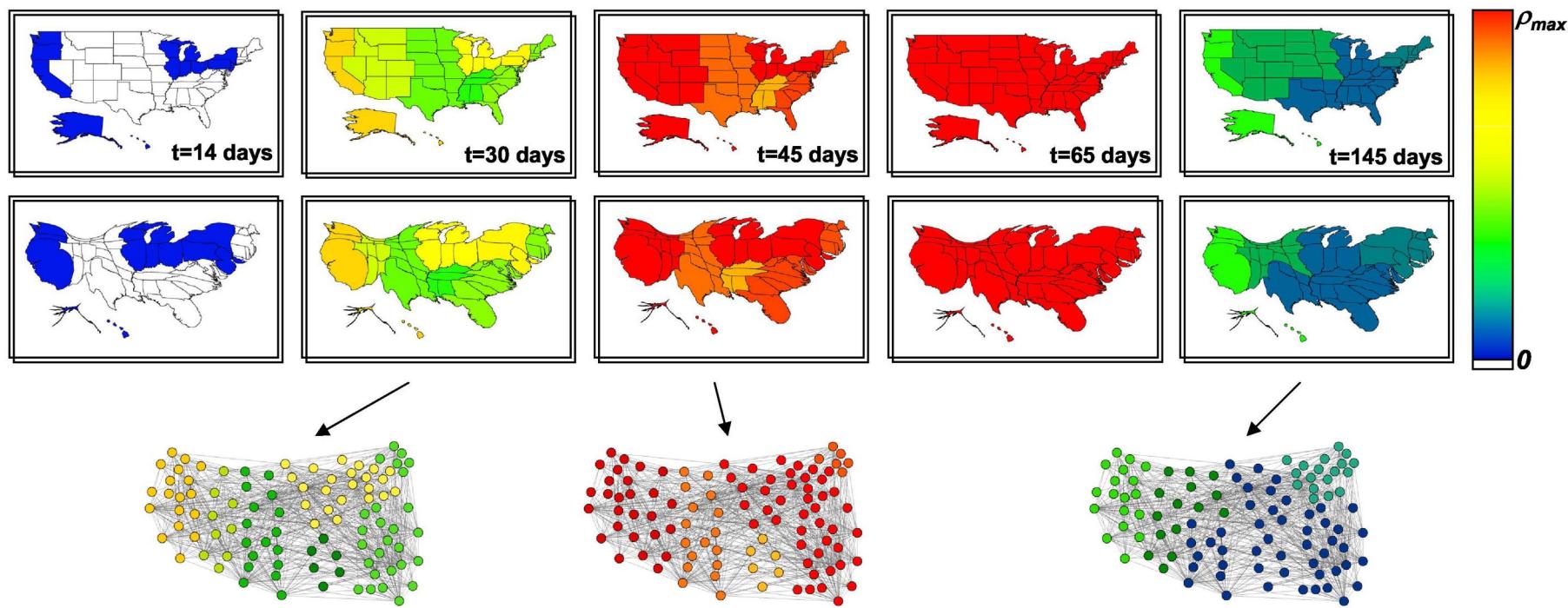
> 99% of total traffic

Directions.....

- Applications...
 - Historical data
 - Scenarios forecast
 - complicated realistic disease models
- Basic **theoretical** questions...
 - simple SIS, SIR models
 - features determining **propagation pattern?**
 - issue of **predictability**, epidemic forecasting?

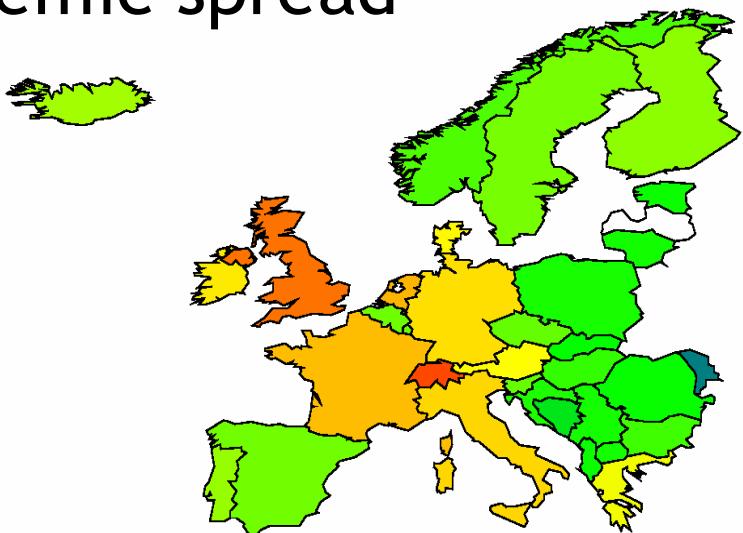
Propagation pattern

Epidemics starting in Hong Kong (SIR model)



Heterogeneity

- maps → heterogeneity epidemic spread
- appropriate measure ?
- role of specific structural properties:
topology, traffic, population ?
- comparison with null hypothesis



Heterogeneity: quantitative measure

$$i_j(t) = \frac{I_j(t)}{N_j} \quad \text{prevalence in city } j \text{ at time } t$$

$$\rho_j(t) = \frac{i_j(t)}{\sum_l i_l(t)} \quad \text{normalized prevalence}$$

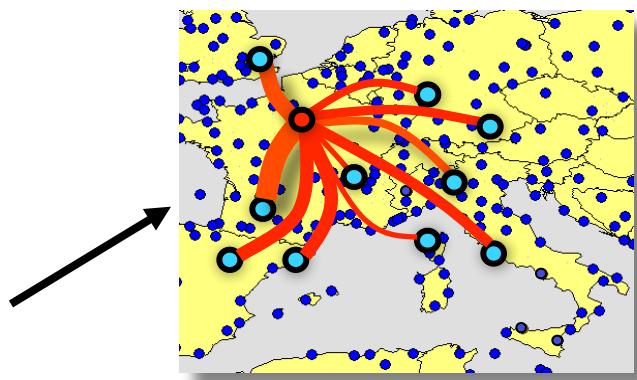
Entropy:

$$H(t) = -\frac{1}{\ln V} \sum_j \rho_j \ln \rho_j$$

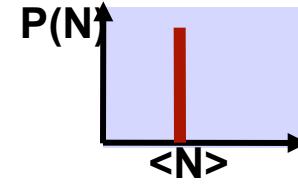
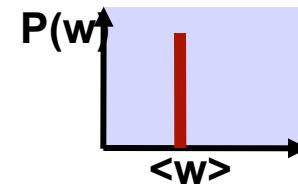
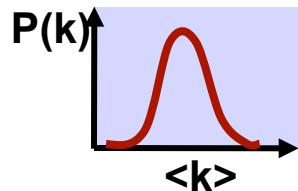
$H \in [0, 1]$
 $H=0$ most het.
 $H=1$ most hom.

...and: compare with null hypothesis!

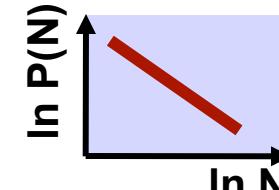
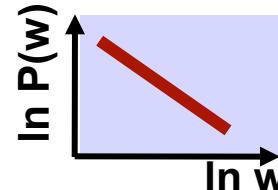
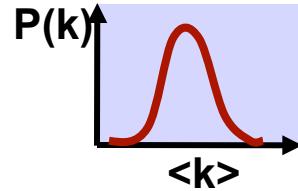
WAN



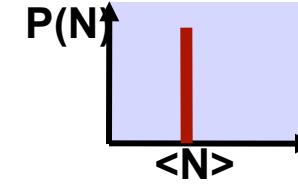
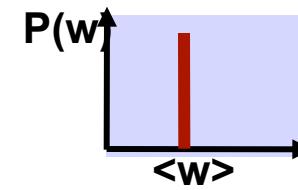
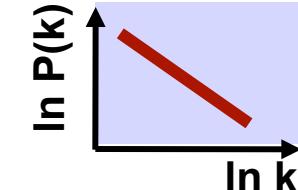
■ **HOM**



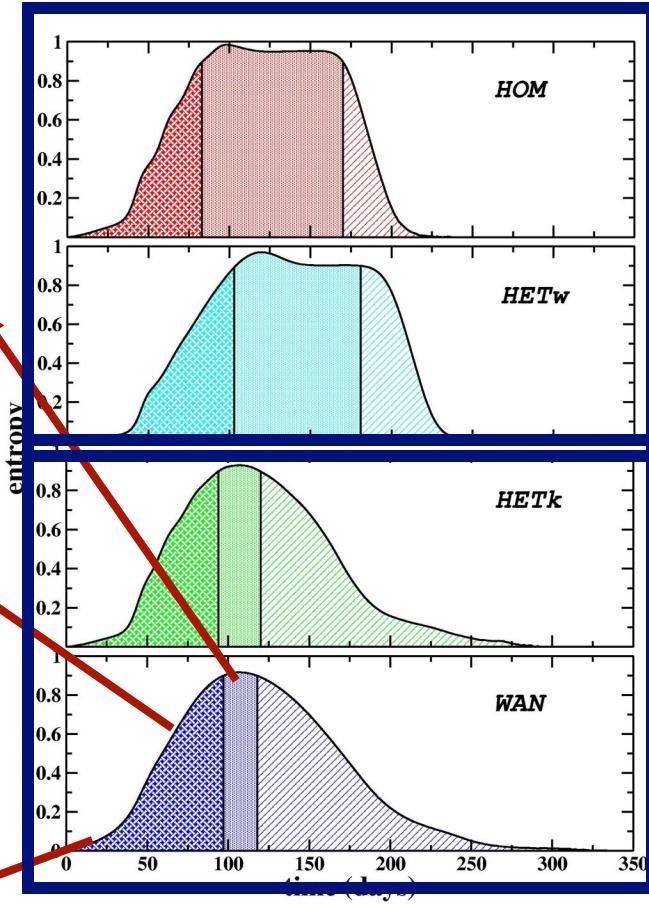
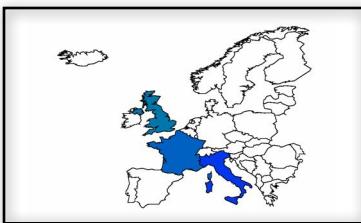
■ **HET_W**



■ **HET_k**



Results: Heterogeneity



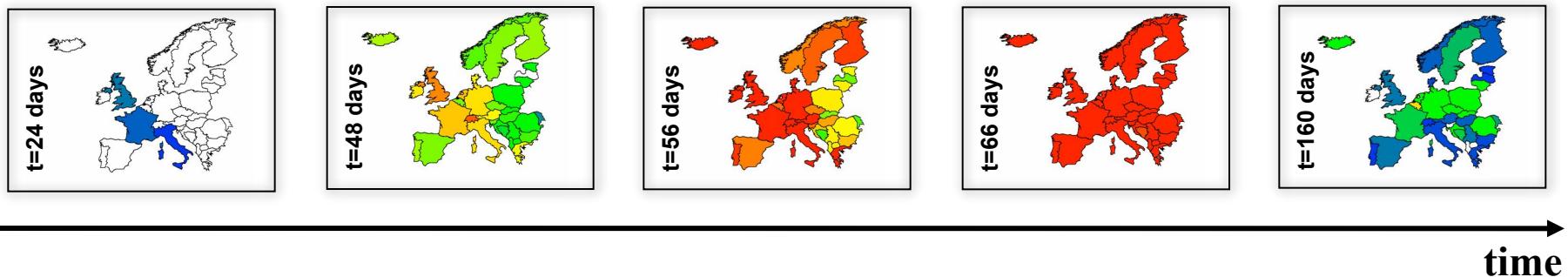
- global properties
- average over initial *seed*
- central zone: $H > 0.9$
- $HETk \cong WAN$
→ importance of $P(k)$

Prediction and predictability

- Do we have consistent scenario with respect to different stochastic realizations?
- What are the network/disease features determining the predictability of epidemic outbreaks
- Is it possible to have **epidemic forecasts?**

Predictability

One outbreak realization:



Another outbreak realization ?



- epidemic forecast
- containment strategies

Quantitative characterization of epidemic predictability

Observable: vector whose components are normalized probability that an infected individual is in city j

Statistical similarity of two outbreaks (*I* and *II*) with the same initial conditions subject to different noise realizations

The similarity between two outbreaks realizations is quantitatively measured by the statistical similarity of two realizations of the global epidemic characterized by the vectors I and II, respectively.

$$\delta(t), \quad \delta_j(t) = \frac{I_j(t)}{\sum_l I_l(t)}$$



$$sim(\delta^I, \delta^{II}) = \sum_j \sqrt{\delta_j^I \delta_j^{II}}$$

Hellinger affinity

Quantitative characterization of epidemic predictability

NB: The normalized distribution similarity is the same in the case of different total prevalence

Normalised similarity measures do not account for the difference in the total epidemic prevalence, and we have to consider also $\text{sim}(i^I, i^{II})$:

$$\vec{i}(t) = (i, 1-i), \quad i = \sum_j I_j / \sum_j N_j$$

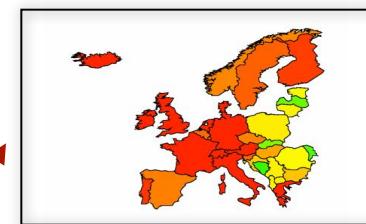
Worldwide epidemic prevalence

Overlap function:

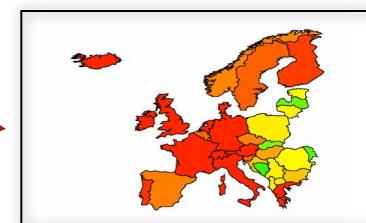
$$\Theta(t) = \text{sim}(\vec{i}^I, \vec{i}^{II}) \times \text{sim}(\vec{\delta}^I, \vec{\delta}^{II})$$

Quantitative characterization of epidemic predictability

$\Theta(t)$ in $[0,1]$

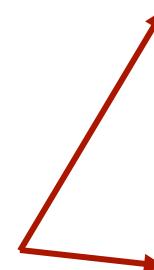


time t

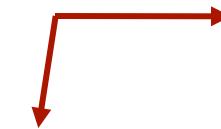


time t

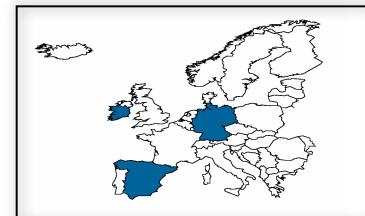
$\Theta(t)=1$: 2 identical outbreaks



$\Theta(t)=0$: 2 distinct outbreaks

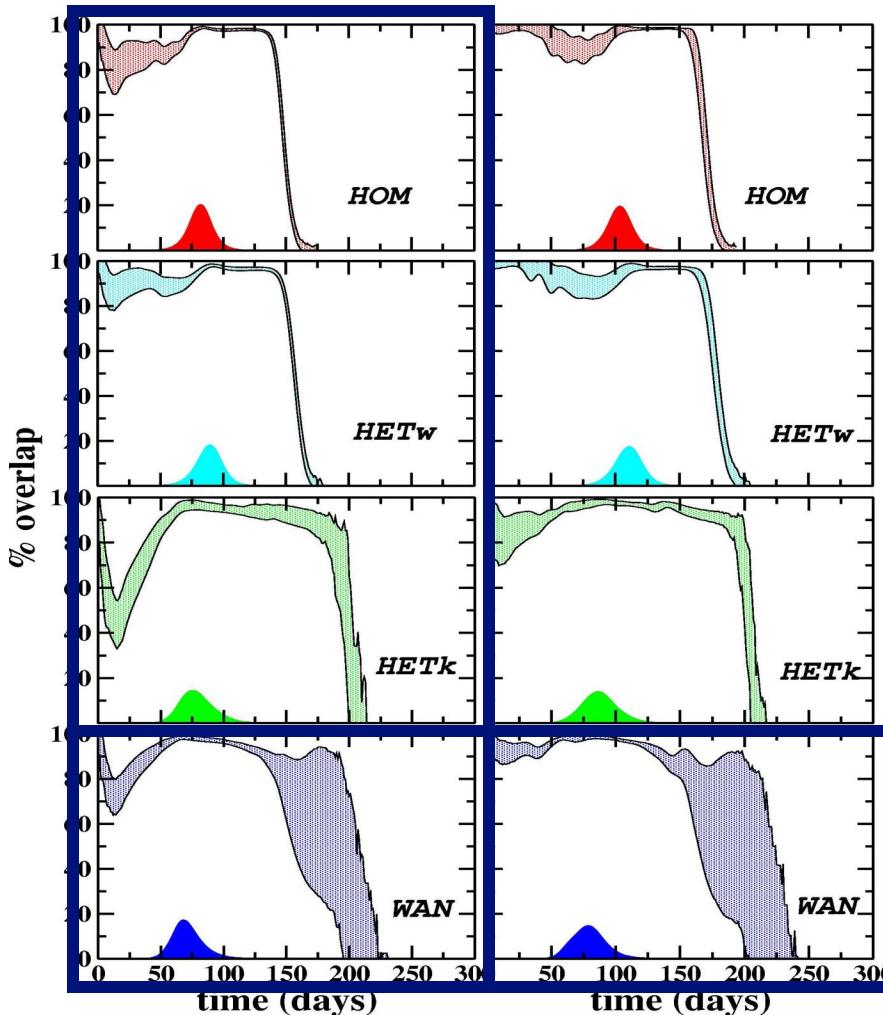


time t



time t

Results: predictability



left: *seed* = airport hubs
right: *seed* = poorly connected airports

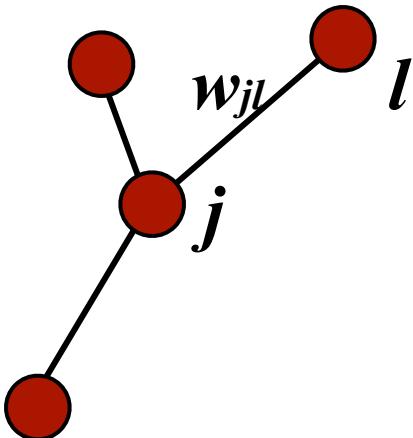
HOM & HET_w high overlap

HET_k low overlap

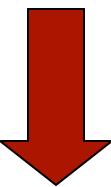
WAN increased overlap !!

The shaded area corresponds to the standard deviation obtained with 5·10³ couples of different realizations of the global spreading model based on a SIR dynamics within each city.

Results: predictability

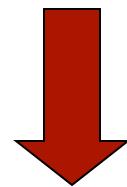


$HOM: k_i \approx \langle k \rangle$ few channels
high overlap

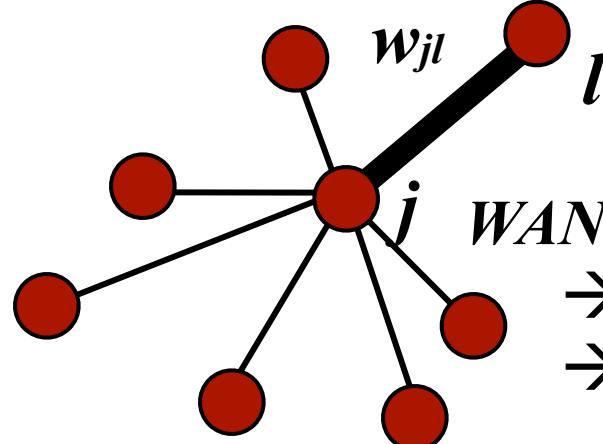


+ degree heterog.

$HETk$: broad $P(k)$, lots of channels,
low overlap



+ weight heterog.



WAN : broad $P(k), P(w)$ lots of channels, but...
→ emergence of **preferred** channels
→ increased overlap !!!

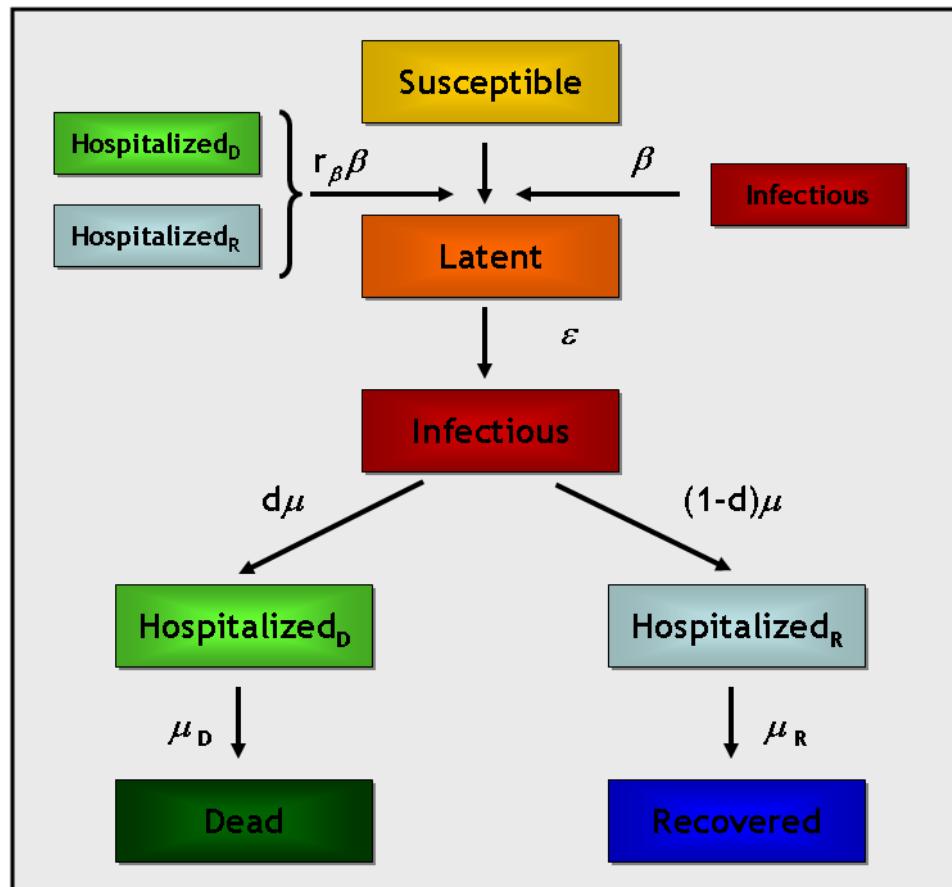
Taking advantage of complexity...

- Two competing effects
 - Paths degeneracy (connectivity **heterogeneity**)
 - Traffic selection (**heterogeneous** accumulation of traffic on specific paths)
- Definition of ***epidemic pathways*** as a backbone of dominant connections for spreading

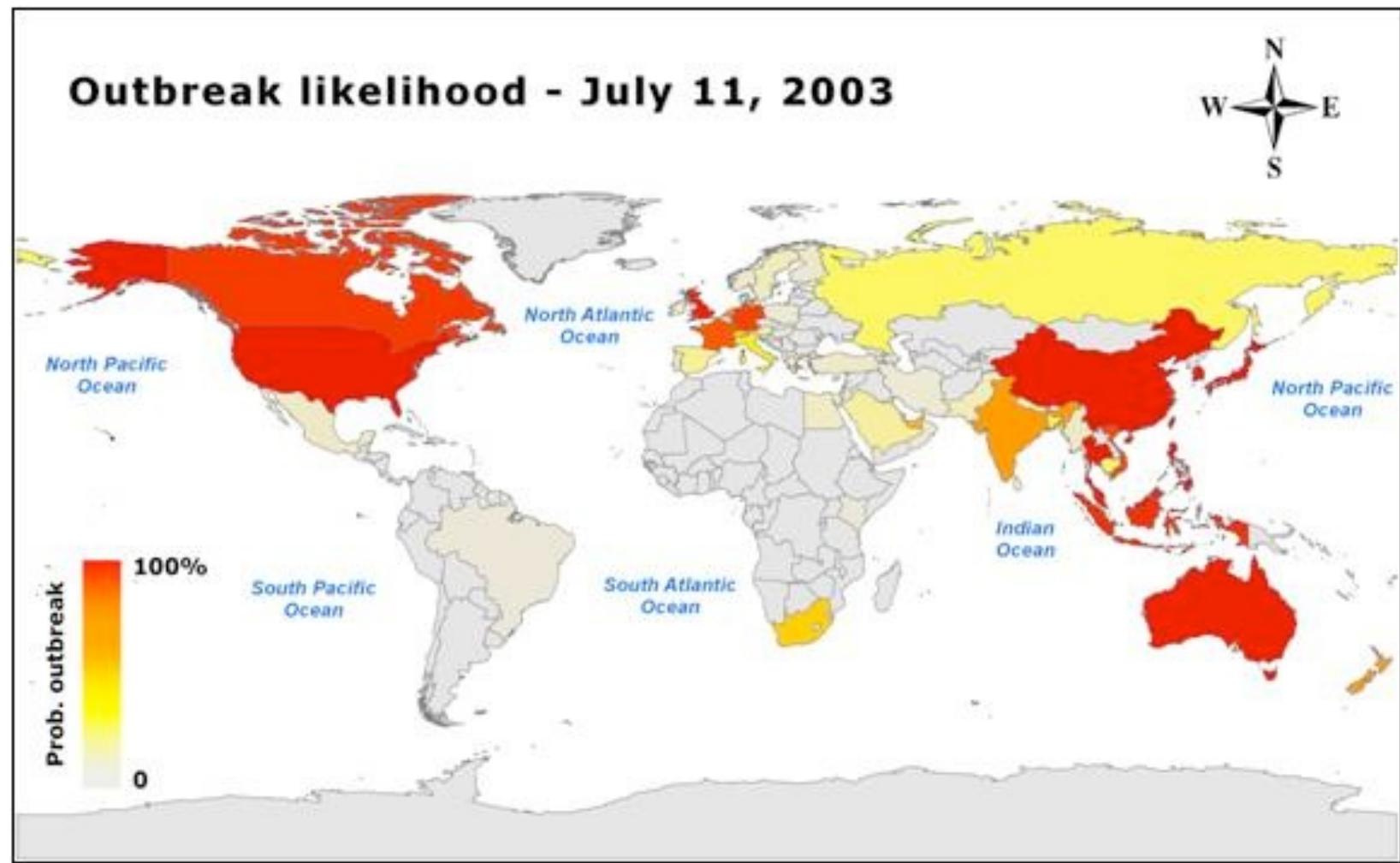
Applications

- Historical data: SARS
 - *more involved model*
 - *validation vs real data*
- Pandemic forecast
 - *effect of travel limitations*
 - *scenario evaluation*

Historical data : The SARS case...

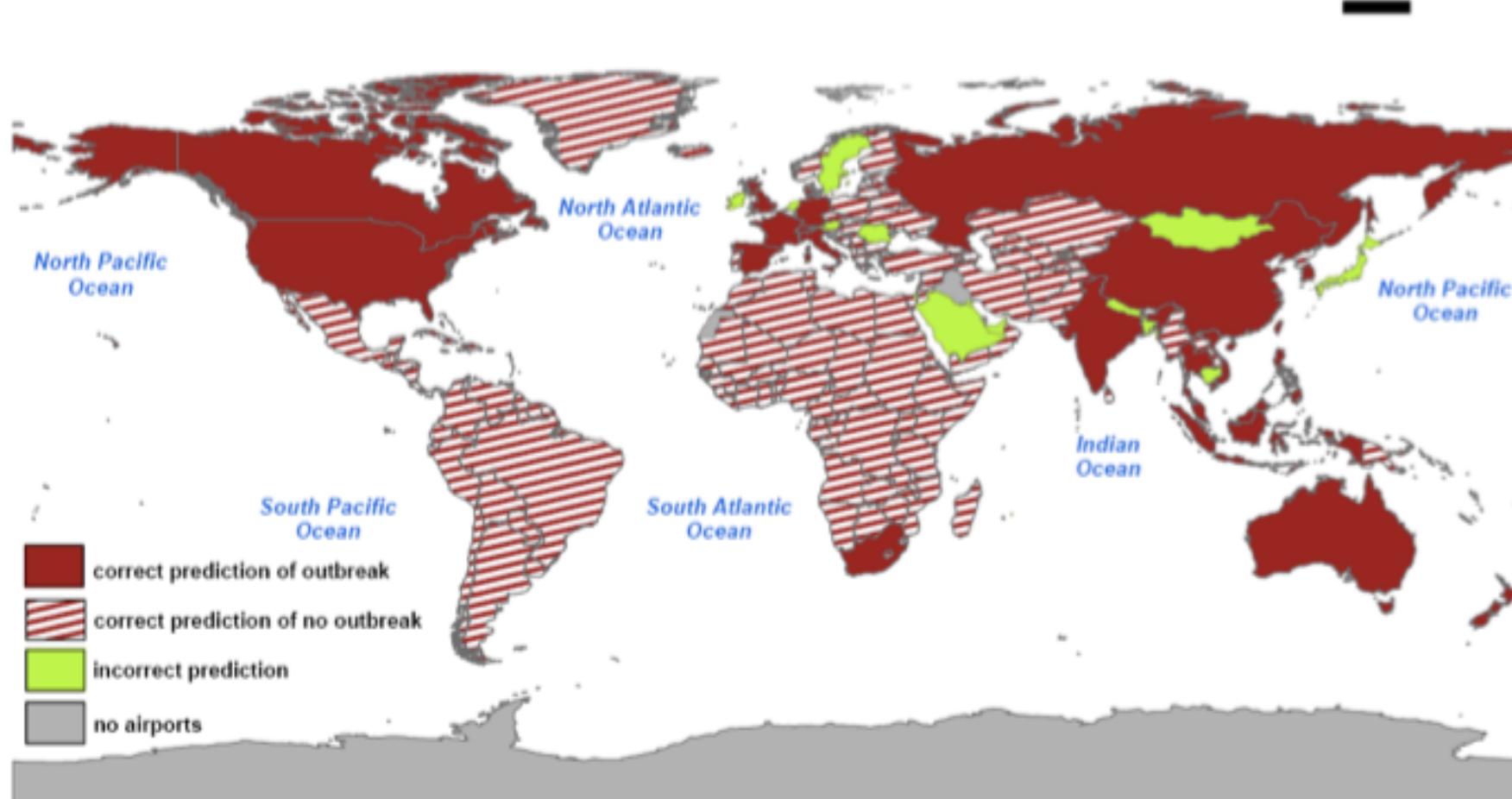


Predictions...

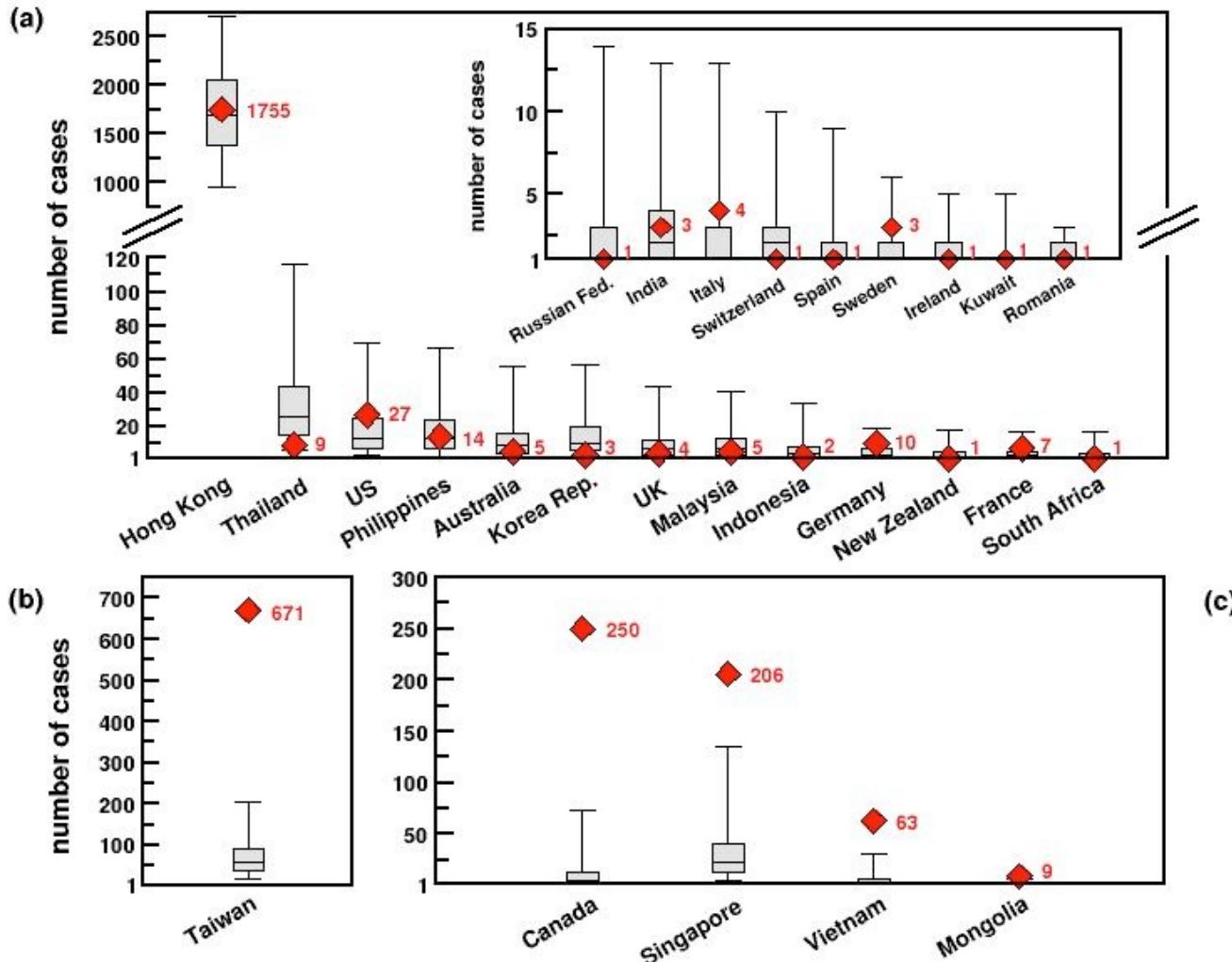


Predictions...

SARS - July 11, 2003

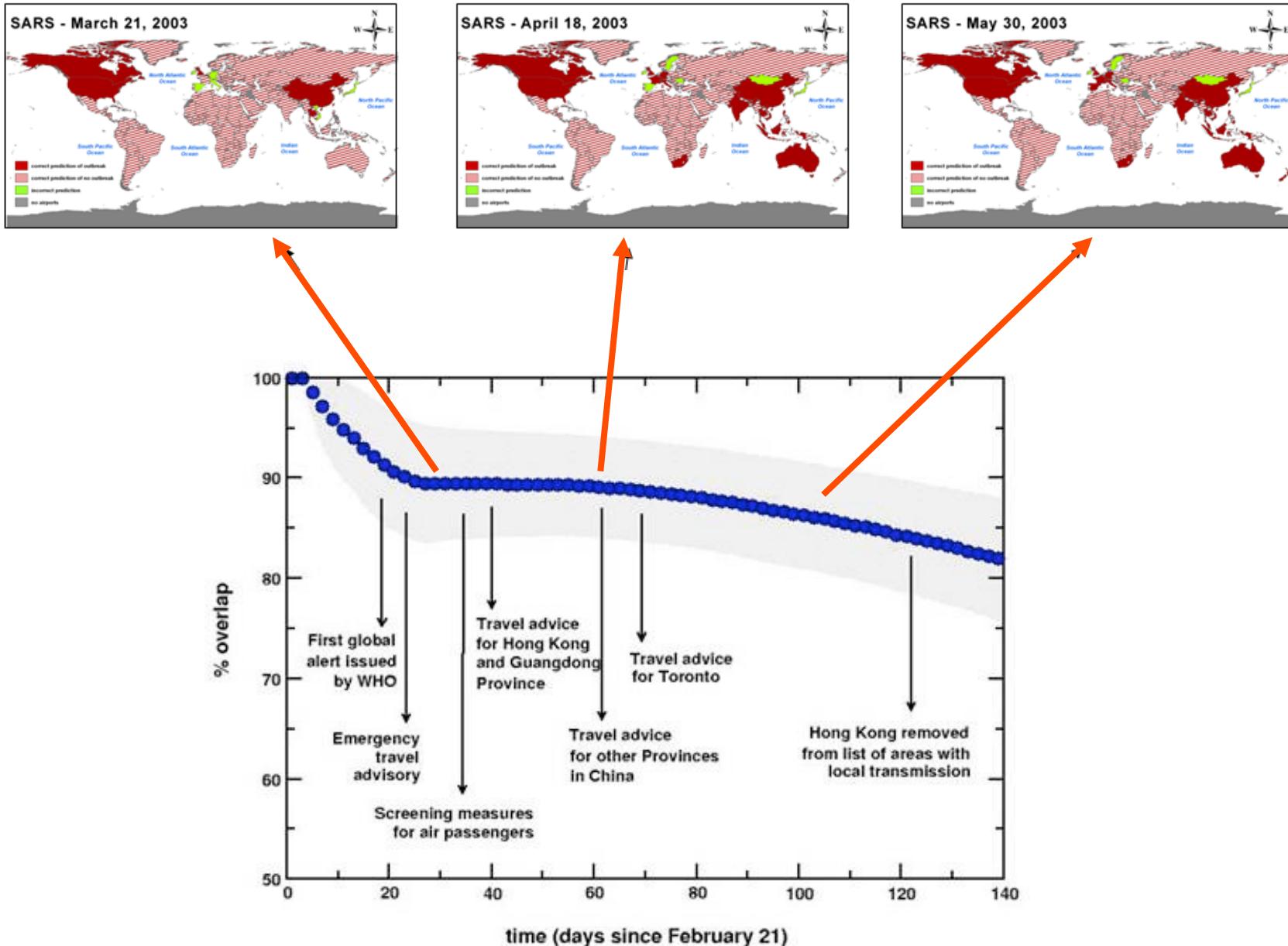


Quantitatively speaking

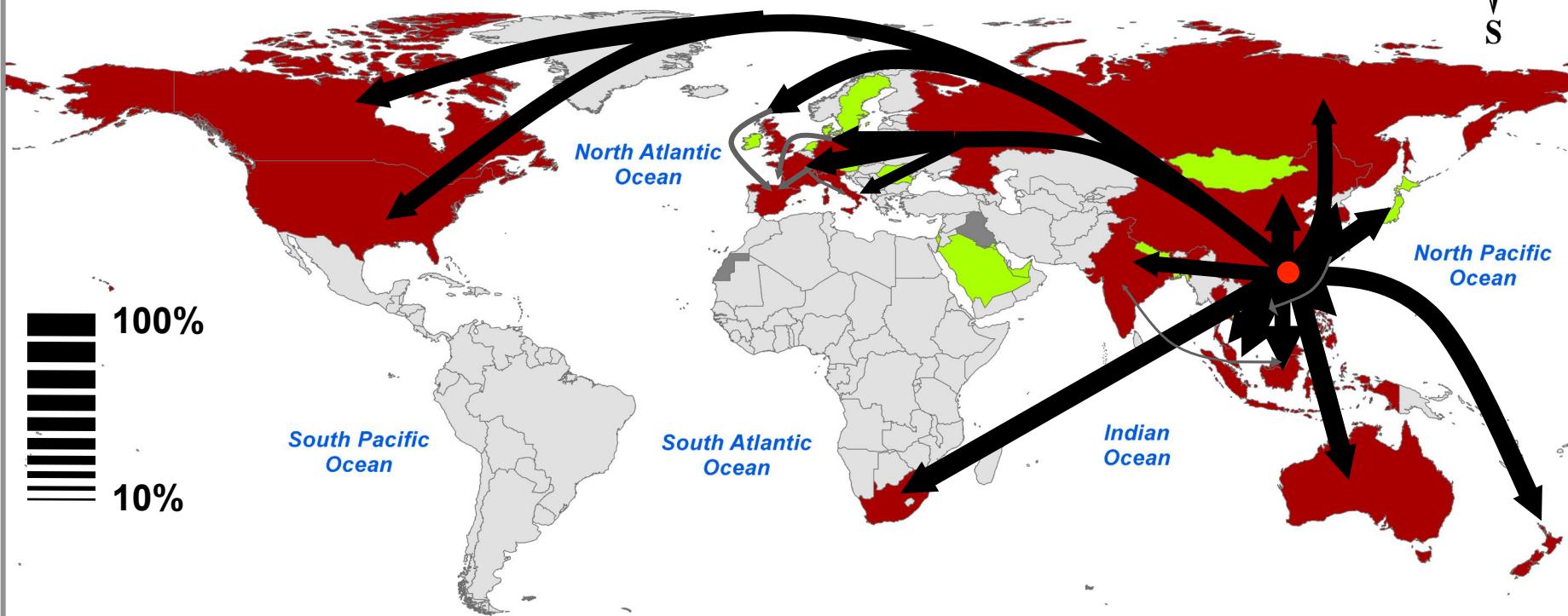


NB: populations of millions of individuals!!

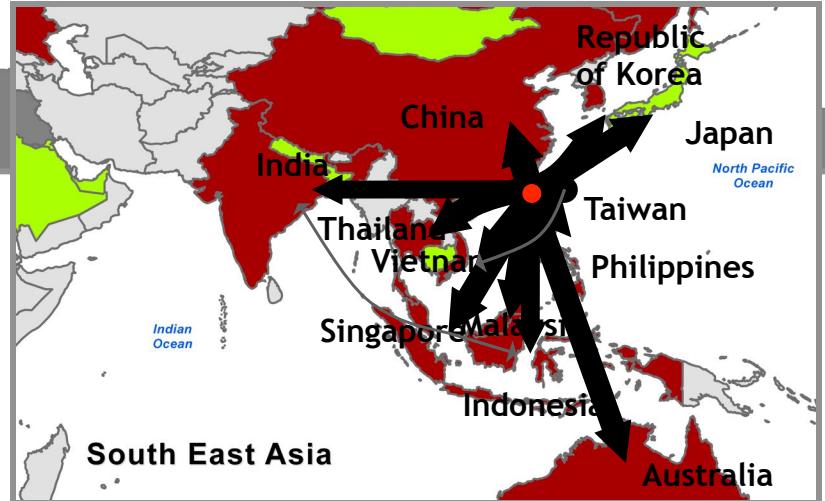
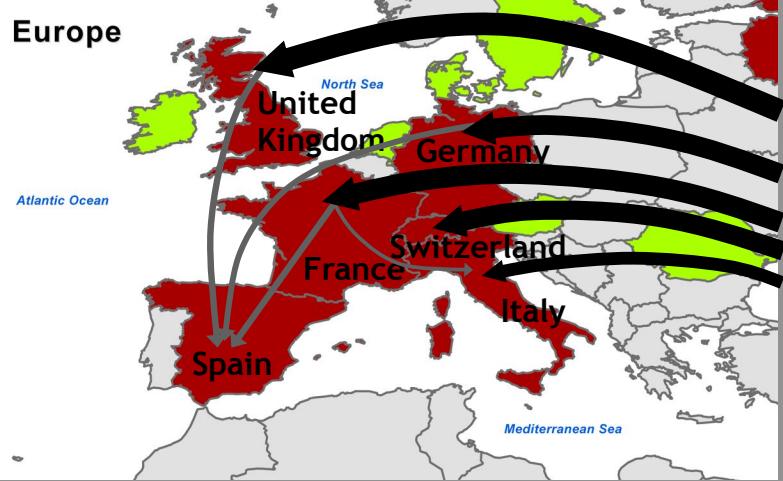
Very good results because...



Epidemic Pathways



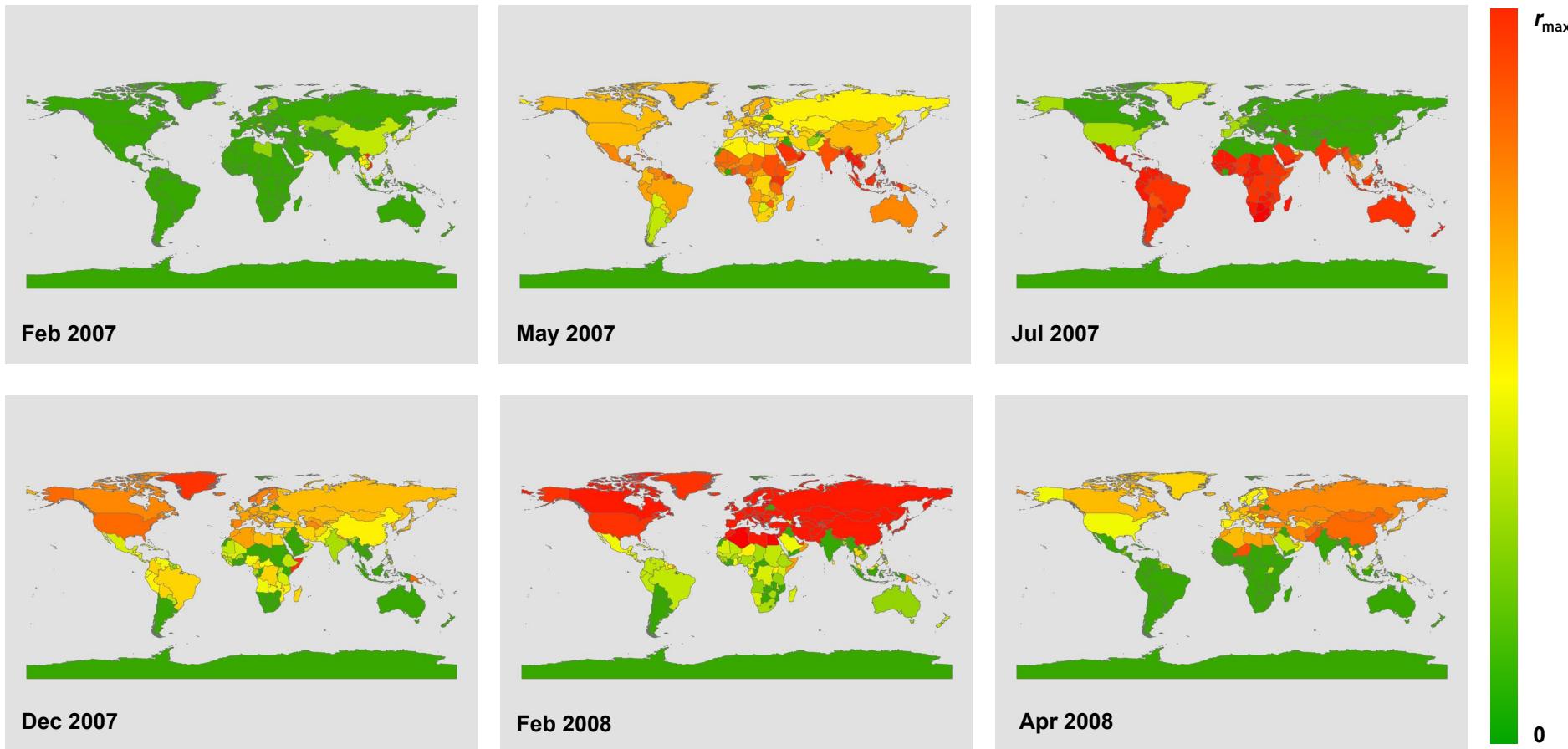
Europe



Scenario evaluation: pandemic forecast

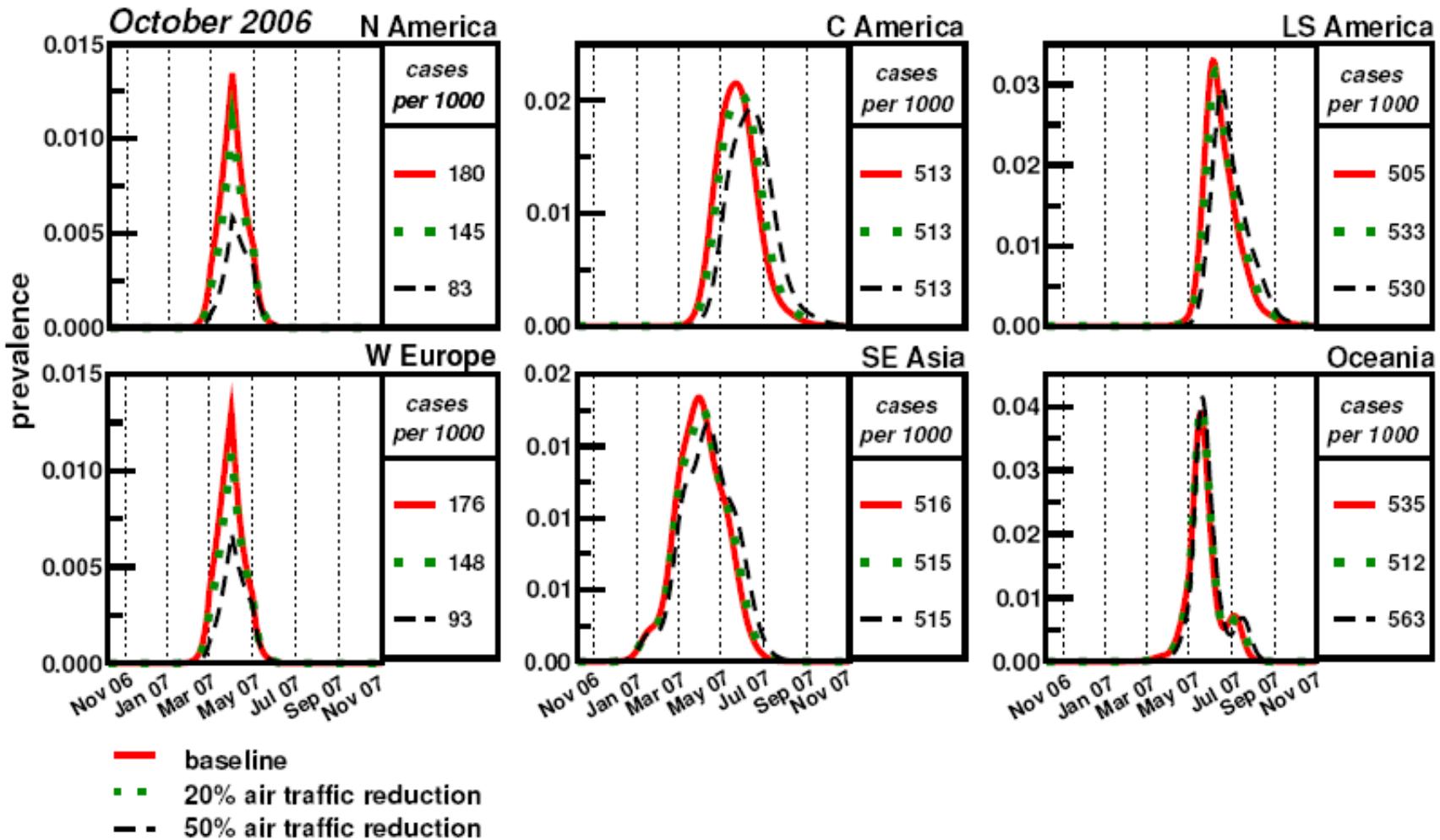
- Estimation of the “dangerous” range of epidemiological parameters
- Inefficiency of **travel restrictions**
- Efficiency of **antivirals**
- Necessity of **cooperative strategies**, which benefit also to the prepared countries

Pandemic forecast...



Pandemic flu with $R_0=1.6$ starting from Hanoi (Vietnam) in October (2006)
Baseline scenario

Travel limitations....

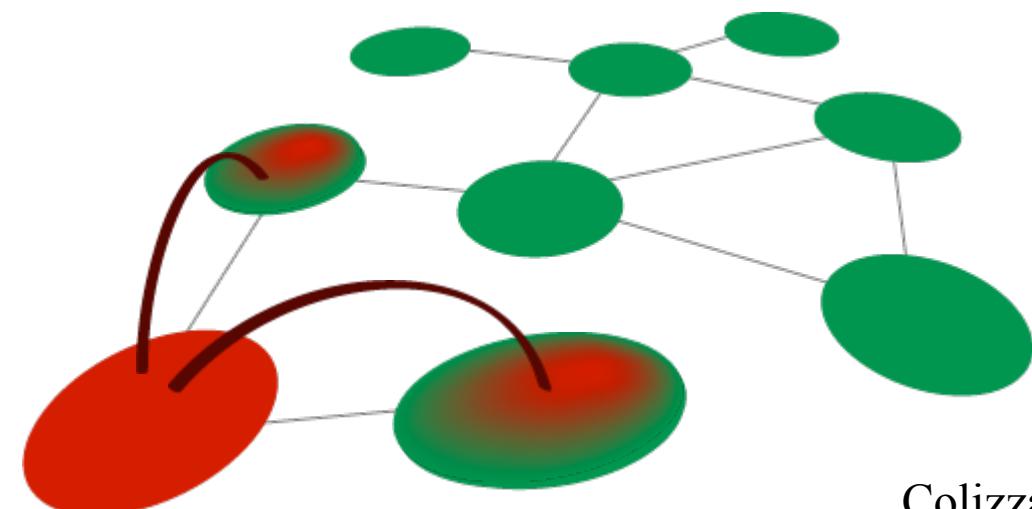


Inefficiency of travel limitations?

Back to theory...: heterogeneous mean-field

$$\partial_t I_k = -p_k I_k + (1 - p_k) [-\mu I_k + \beta \Gamma_k] + k \sum_{k'} P(k' | k) d_{k'k} [(1 - \mu) I_{k'} + \beta \Gamma_{k'}]$$

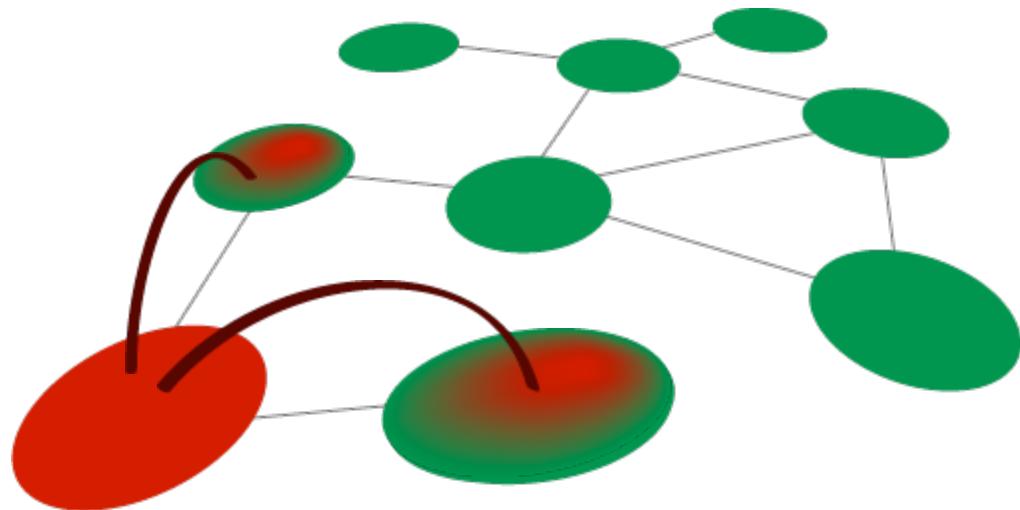
leaving *staying and reacting*
incoming



(SIR model, parameters β and μ)

Can be very small: travel of *fractions* of individuals

$$\begin{aligned}\partial_t I_k = & -p_k I_k + (1 - p_k) [-\mu I_k + \beta \Gamma_k] \\ & + k \sum_{k'} P(k' | k) d_{k'k} [(1 - \mu) I_{k'} + \beta \Gamma_{k'}]\end{aligned}$$



continuum approximation
NOT VALID



discreteness
stochasticity

Invasion: branching process

D_k^0, D_k^1, \dots # of **diseased** nodes (i.e., with at least one infected individual) of degree k , at generation $n=0, 1, \dots$

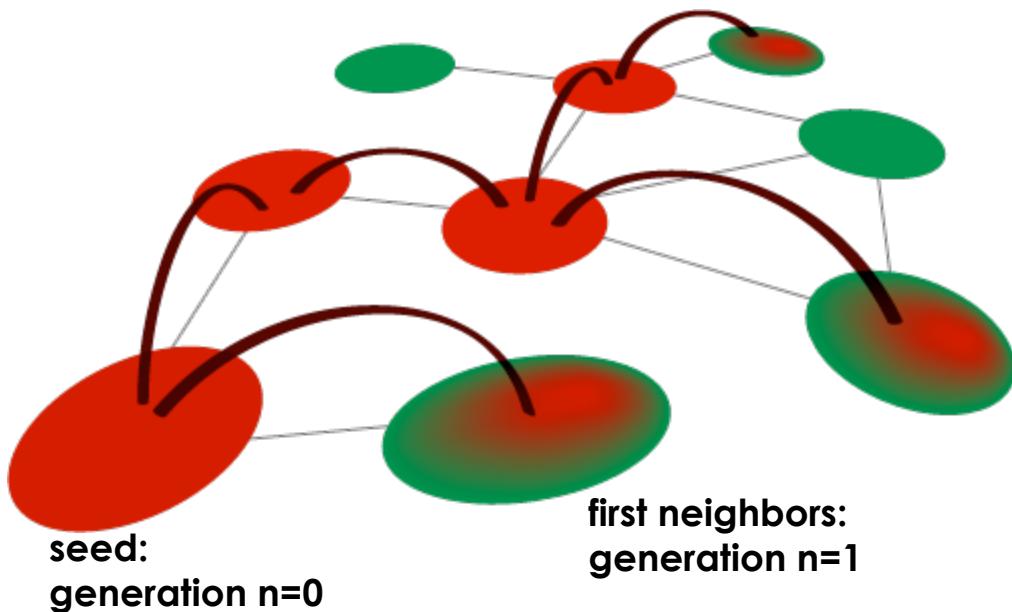
$$D_k^n = \sum_{k'} D_{k'}^{n-1} P(k | k')(k'-1) \left(1 - \frac{D^{n-1}}{V} \right) \left[1 - R_0^{-\lambda_{kk'}} \right]$$

*possible paths
of infection*

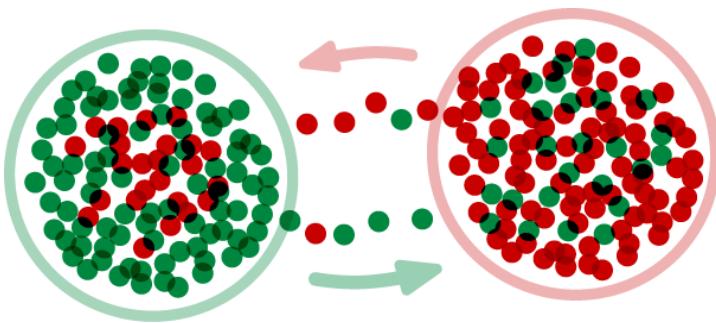
connection $k'-k$

from micro to macro

*probability of finding
uninfected
subpopulation*



from micro to macro scale



of new seeds from k' to k
(time-scale separation):

$$\lambda_{k'k} = d_{k'k} \frac{\alpha N_{k'}}{\mu}$$

total # of infected generated
diffusion rate *rate* $\mapsto \mathbb{R}$

Ex of diffusion rates:

$$d_{k'k} = p w_0 (k'k)^\theta / T_k,$$

$$d_{k'k} = w_0 (kk')^\theta / N_k,$$

(stationary populations ind. of diffusion process))

For R_0 close to 1, and at short times:

$$D_k^n = (R_0 - 1) \frac{k^{1+\theta} P(k)}{\langle k \rangle} \frac{w_0 \alpha}{\mu} \sum_{k'} D_{k'}^{n-1} k'^\theta (k' - 1)$$



$$\Theta^n = (R_0 - 1) \frac{\langle k^{2+2\theta} \rangle - \langle k^{1+2\theta} \rangle}{\langle k \rangle} \frac{w_0 \alpha}{\mu} \Theta^{n-1}$$


$> 1 \Leftrightarrow$ global invasion

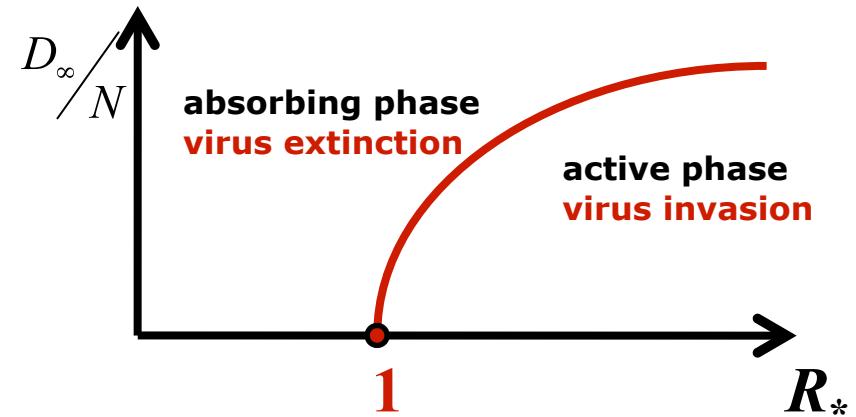
Global invasion threshold

$$R_* = (R_0 - 1) \frac{\langle k^{2+2\theta} \rangle - \langle k^{1+2\theta} \rangle}{\langle k \rangle} \frac{w_0 \alpha}{\mu}$$

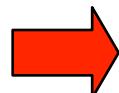
Ex: SIR, $\alpha \sim 2(R_0 - 1)/R_0^2$

$$w_{0c} = \frac{\mu R_0^2}{2(R_0 - 1)^2} \frac{\langle k \rangle}{\langle k^{2+2\theta} \rangle - \langle k^{1+2\theta} \rangle}$$

phase transition in mobility



Real-world network: w_0 100 times larger than w_{0c} !!!



Explains empirical results!!

Further recent developments

- Integration of additional details in the model
 - Mobility patterns at various scales (commuting)

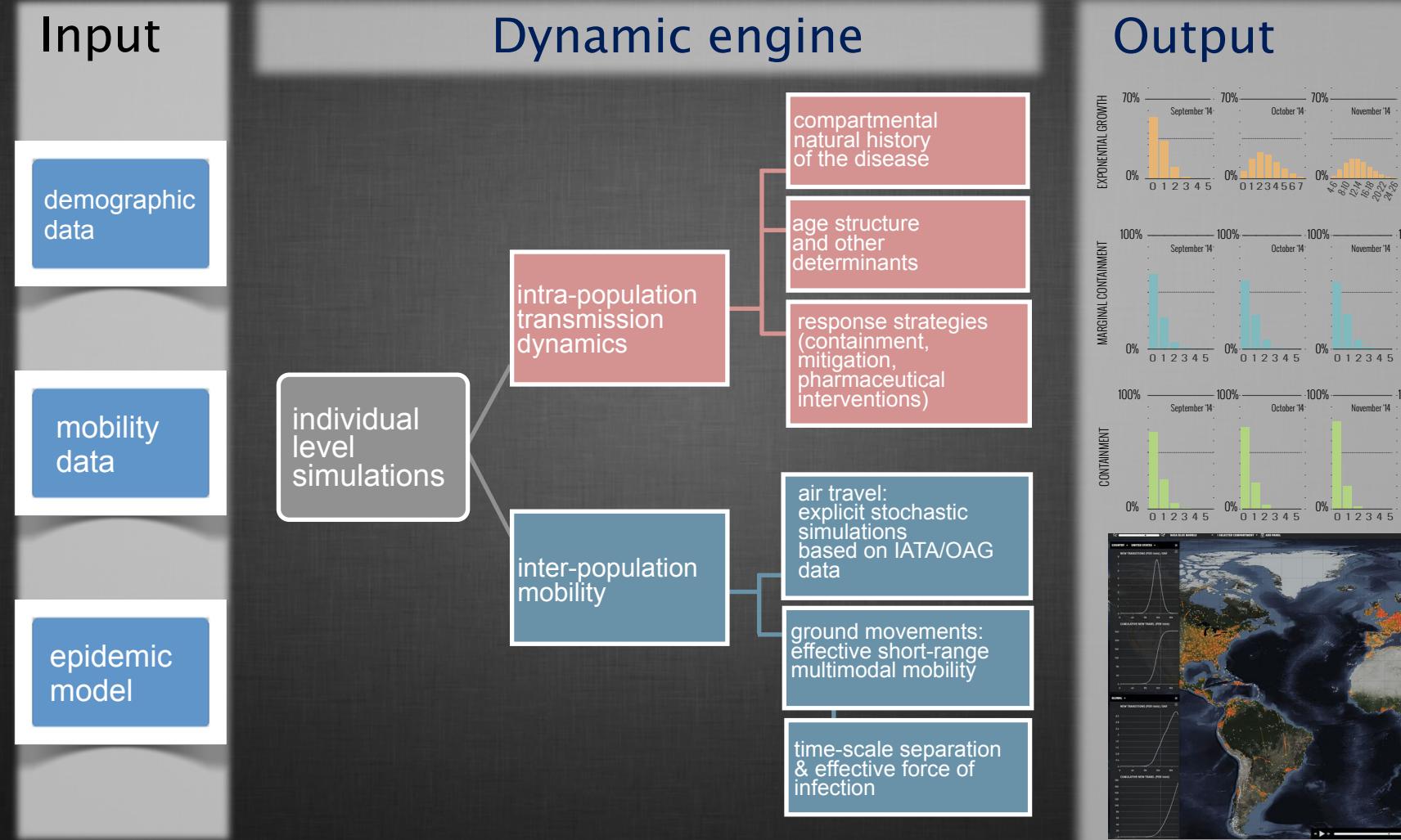
<http://www.gleamviz.org>

simulation platform for the worldwide propagation of diseases, used in real time during the H1N1 pandemic

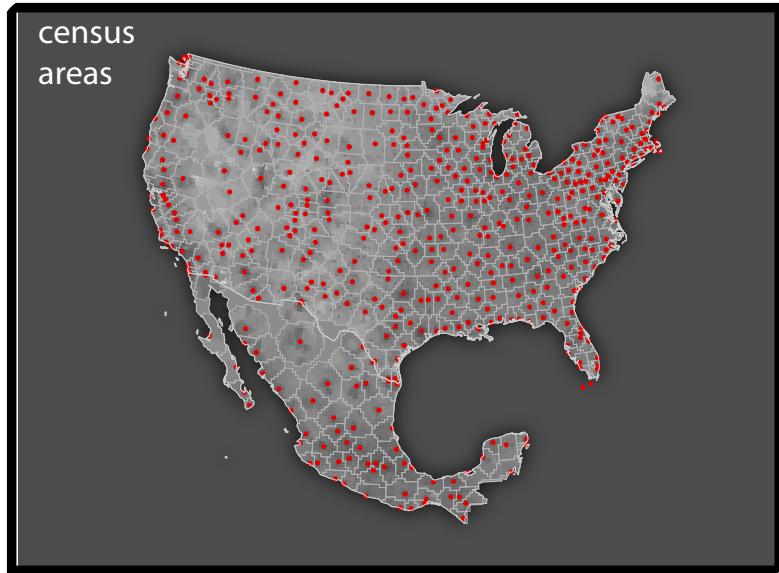
- Heterogeneity of travel behaviors
- Social response to a crisis...????

D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J.J. Ramasco, A. Vespignani
Proc. Natl. Acad. Sci. USA 106, 21484-21489 (2009)

MODEL ARCHITECTURE



GLEaM in brief

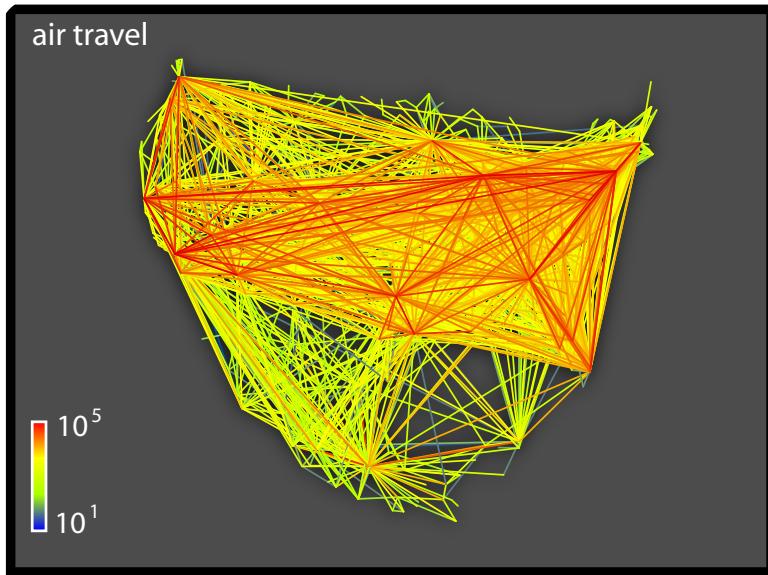


Population distribution:
detailed population data from
1/4x1/4 degree tesselation.

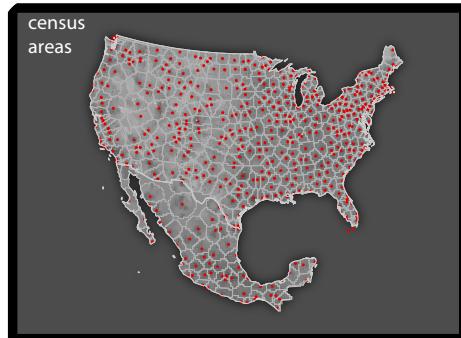
GLEaM in brief



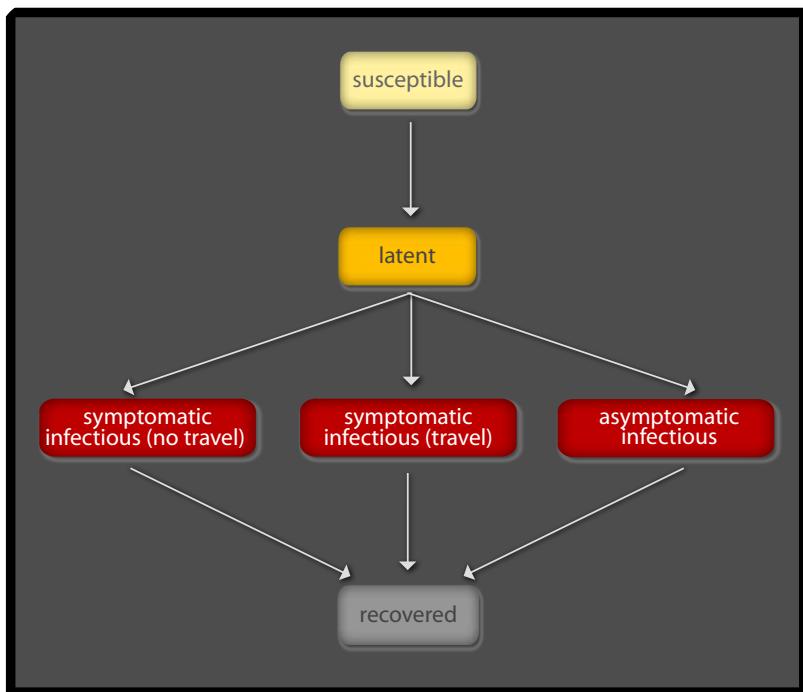
Local mobility:
census data from about 30 countries in the 5 continents extended to all countries.



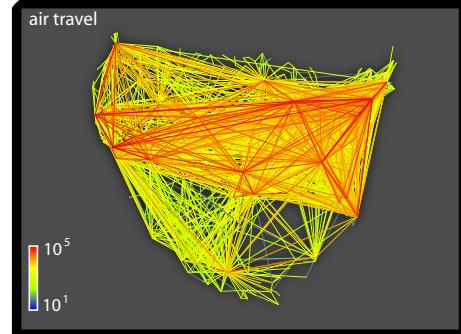
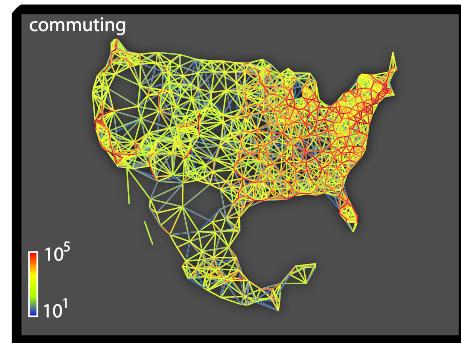
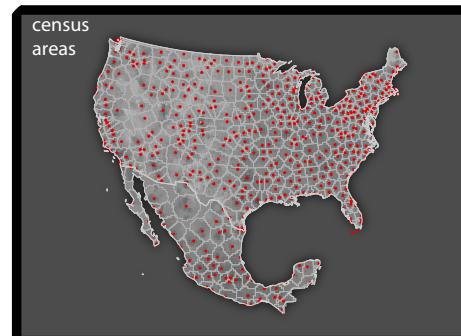
Long range travel:
3362 cities in 220 countries.
More than 16000 connections with travel flows.



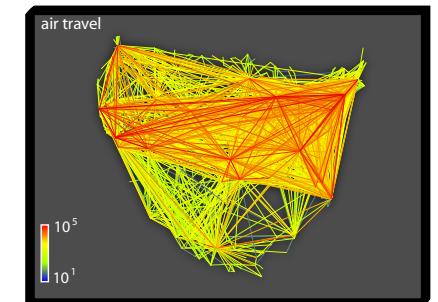
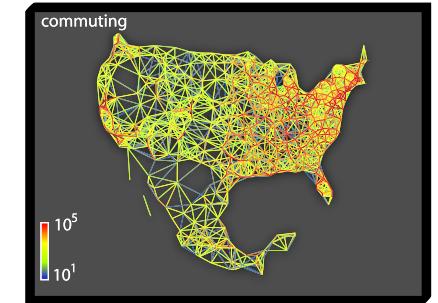
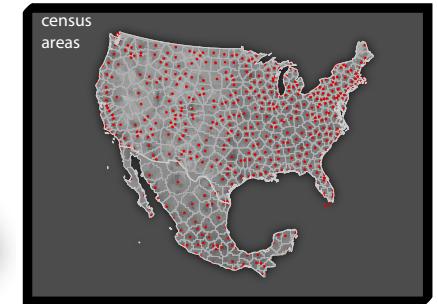
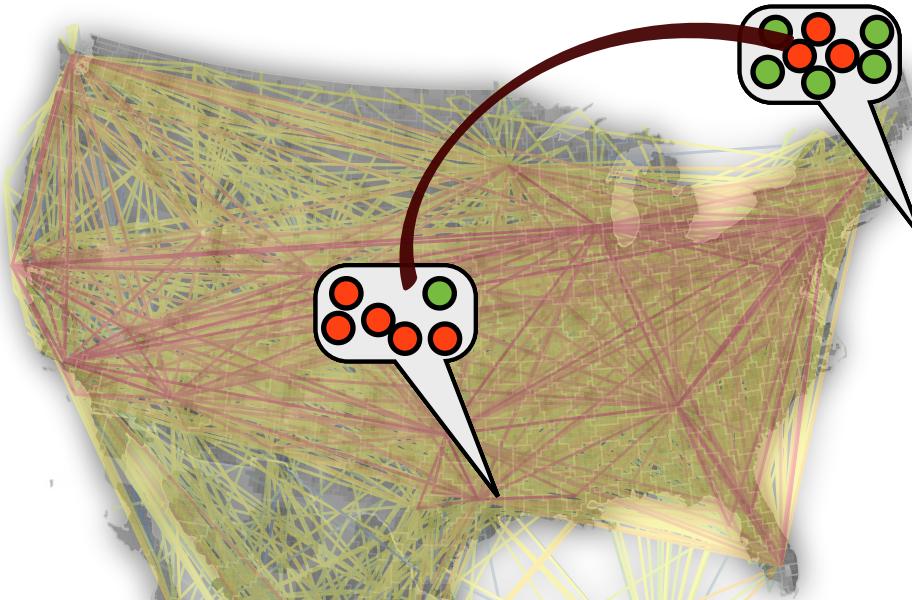
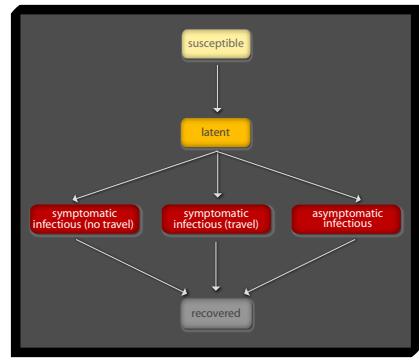
GLEaM in brief



Epidemic compartmental model
Metapopulation model with homogeneous mixing assumption.



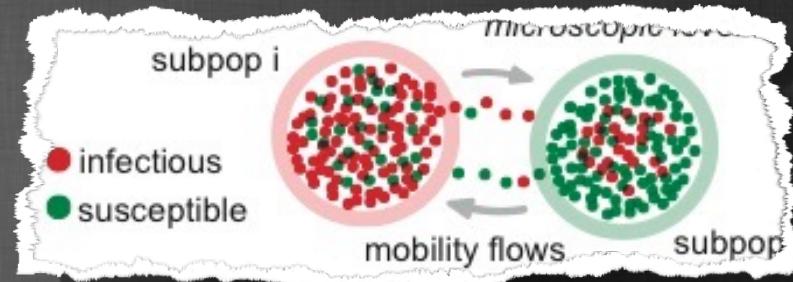
GLEaM in brief



WHAT IS UNDER THE HOOD

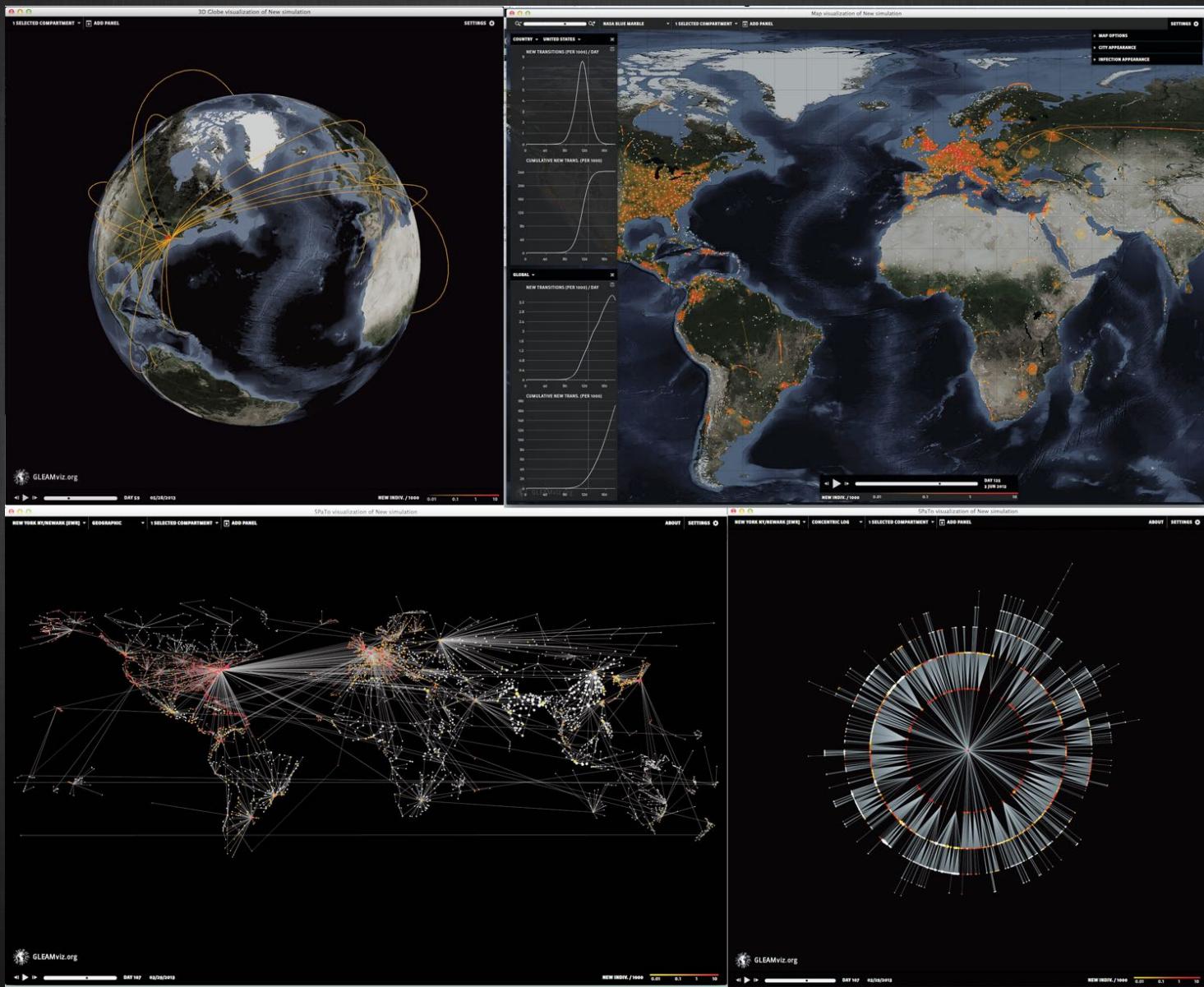
Stochastic Intra population disease evolution: chain binomial process

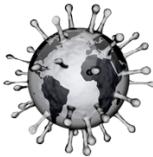
- REPEAT
 - CALL RANBin($S, \beta I/N$) and RANBin(I, μ)
 - $S=S-\text{RANBin}(S, \beta I/N)$
 - $I=I+\text{RANBin}(S, \beta I/N)-\text{RANBin}(I, \mu)$
 - $R=R+\text{RANBin}(I, \mu)$
 - $t = t + \Delta t$
 - PRINT S, I, R, t
- UNTIL $I = 0$



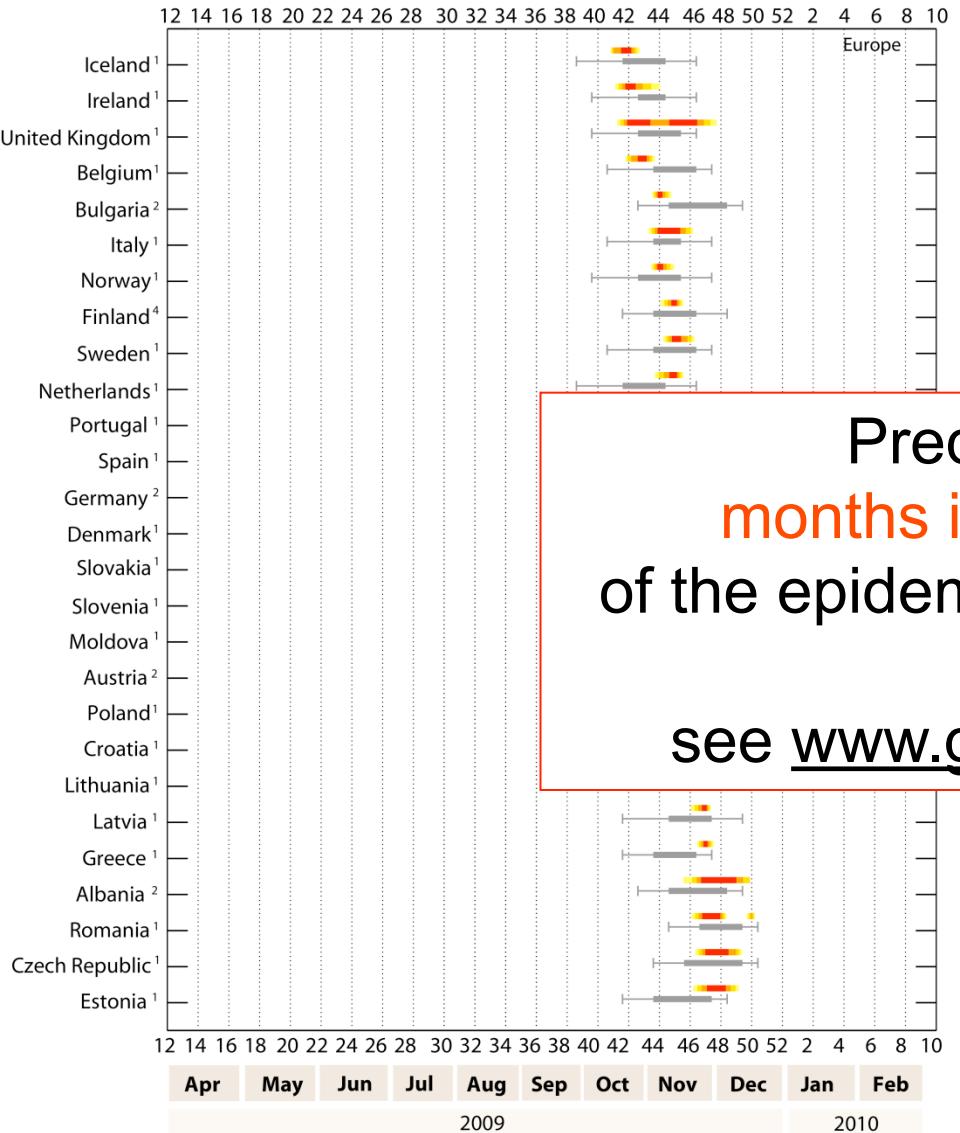
Stochastic Inter population dynamics:

- data driven Explicit stochastic simulation of slow mode traveling patterns.
- Time-Scale separation through effective force of infection for fast commuting patterns
 - > 3,500 single populations coupled models
 - > 40,000-100,000 stochastic discrete equations



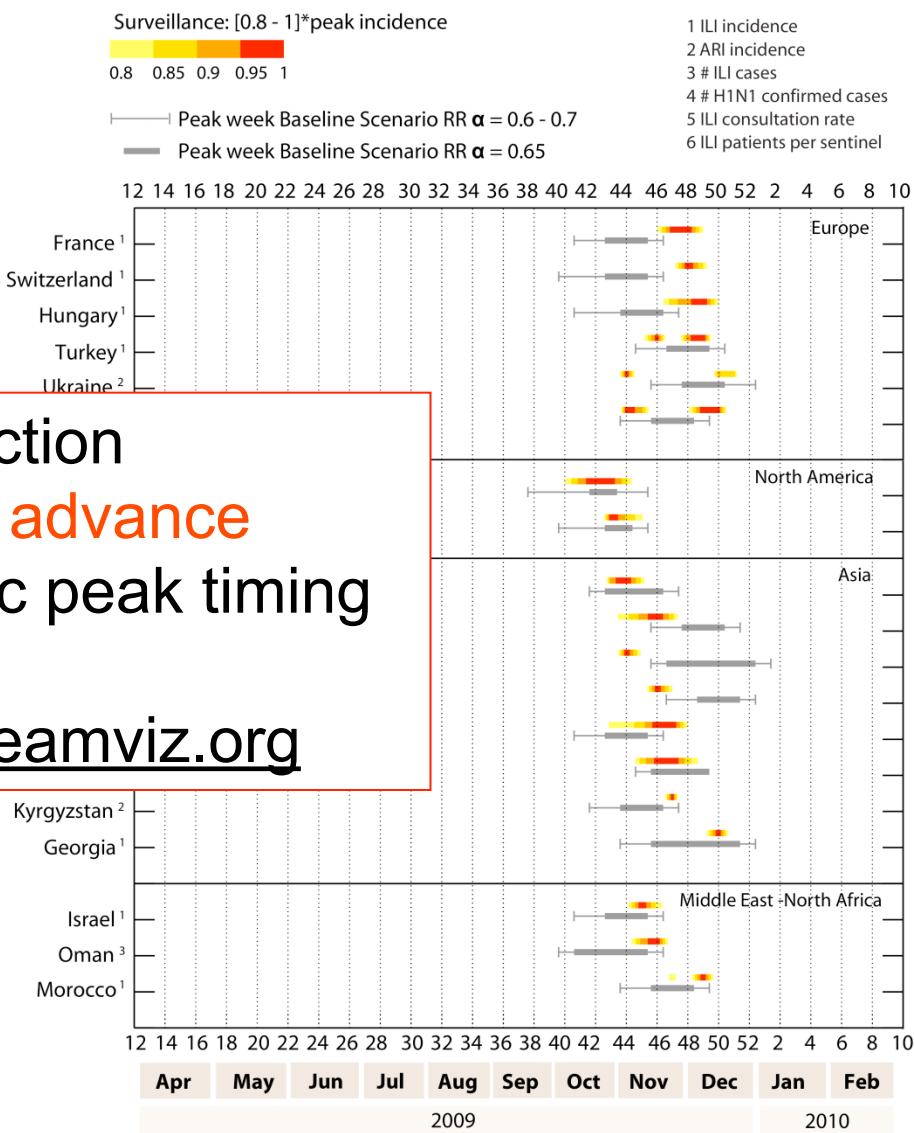


comparing with data



Prediction
months in advance
of the epidemic peak timing

see www.bleamviz.org



REAL TIME FORECAST FOR THE H1N1PDM (2009)

Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility

Duygu Balcan^{†1,2}, Hao Hu^{†1,2,3}, Bruno Goncalves^{†1,2}, Paolo Bajardi^{†4,5}, Chiara Poletto^{†4}, Jose J Ramasco⁴, Daniela Paolotti⁴, Nicola Perra^{1,6,7}, Michele Tizzoni^{4,8}, Wouter Van den Broeck⁴, Vittoria Colizza⁴ and Alessandro Vespignani *^{1,2,4}

Published: 10 September 2009

BMC Medicine 2009, 7:45 doi:10.1186/1741-7015-7-45

Received: 31 July 2009

Accepted: 10 September 2009

Key parameters

- Transmissibility
- Generation time
- Seasonality scaling

Monte-Carlo Likelihood estimate

Structural data

- Transportation/mobility
- Demographic/census

DATA

Plausible parametrization

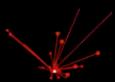
Parameter	Description	Value	Sensitivity Analysis Range
r_β	Relative infectiousness of asymptomatic individuals	0.5	0.2 – 0.8
p_a	Probability of becoming an asymptomatic individual	0.33	0.33 – 0.5
p_t	Probability of becoming a traveling symptomatic individual	0.5	0.4 – 0.6
β	Transmission rate	$\mu^{-1} R_0 / (1 - p_a - r_\beta p_t)$	
α_{\max}	Maximal seasonality rescaling	1.1	1.0 – 1.1

Monte Carlo likelihood parameters' estimate

Chicago
New York
Los Angeles
Houston
Toronto
Vancouver
Calgary
Indianapolis

La Gloria
Sao Paulo
Mexico City
Rio De Janeiro
San Juan
Bogota

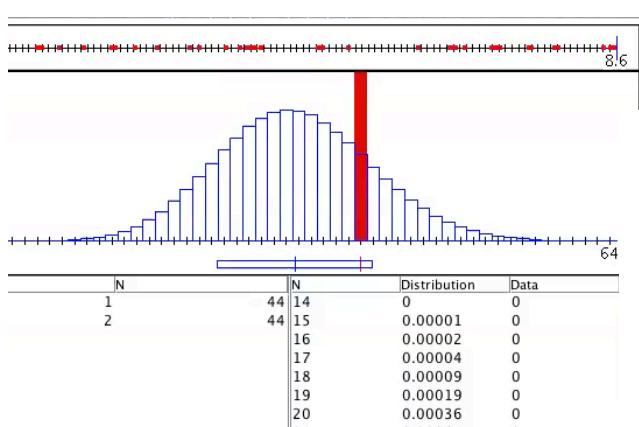
Johannesburg
Cairo
Cape Town
Nairobi



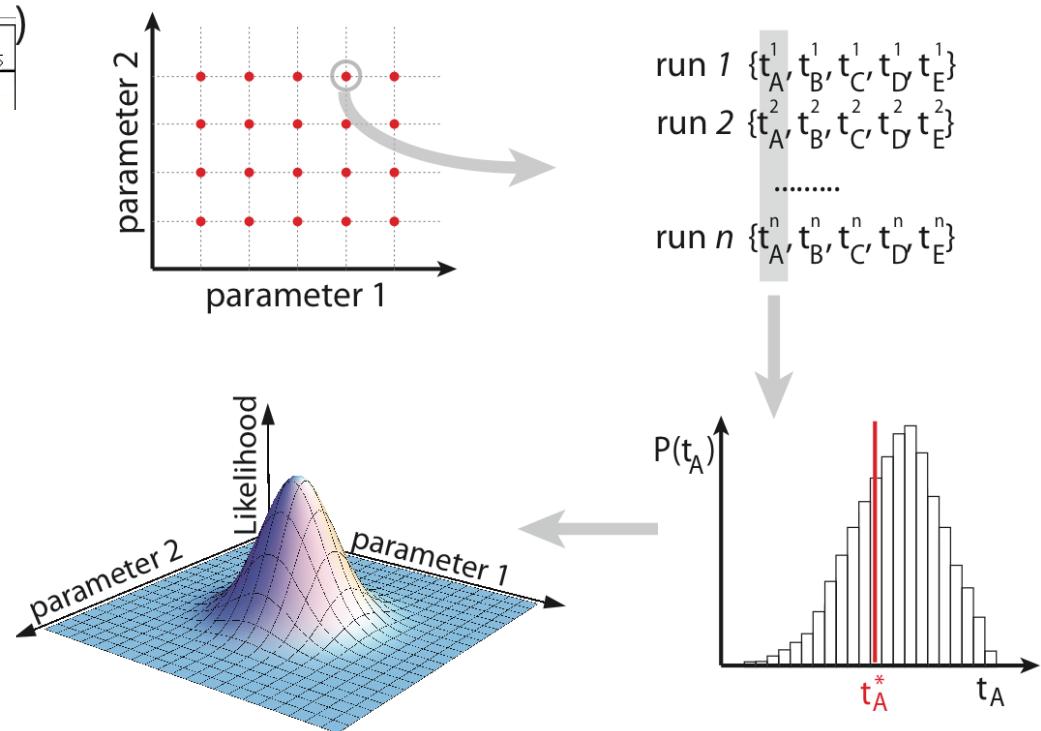
00:00:09 ◀ ▶ -0:00:41

Backtrack of the number of infections in Mexico from case importation and transportation data.
(Fraser et al. Science 2009, 324, 1557)

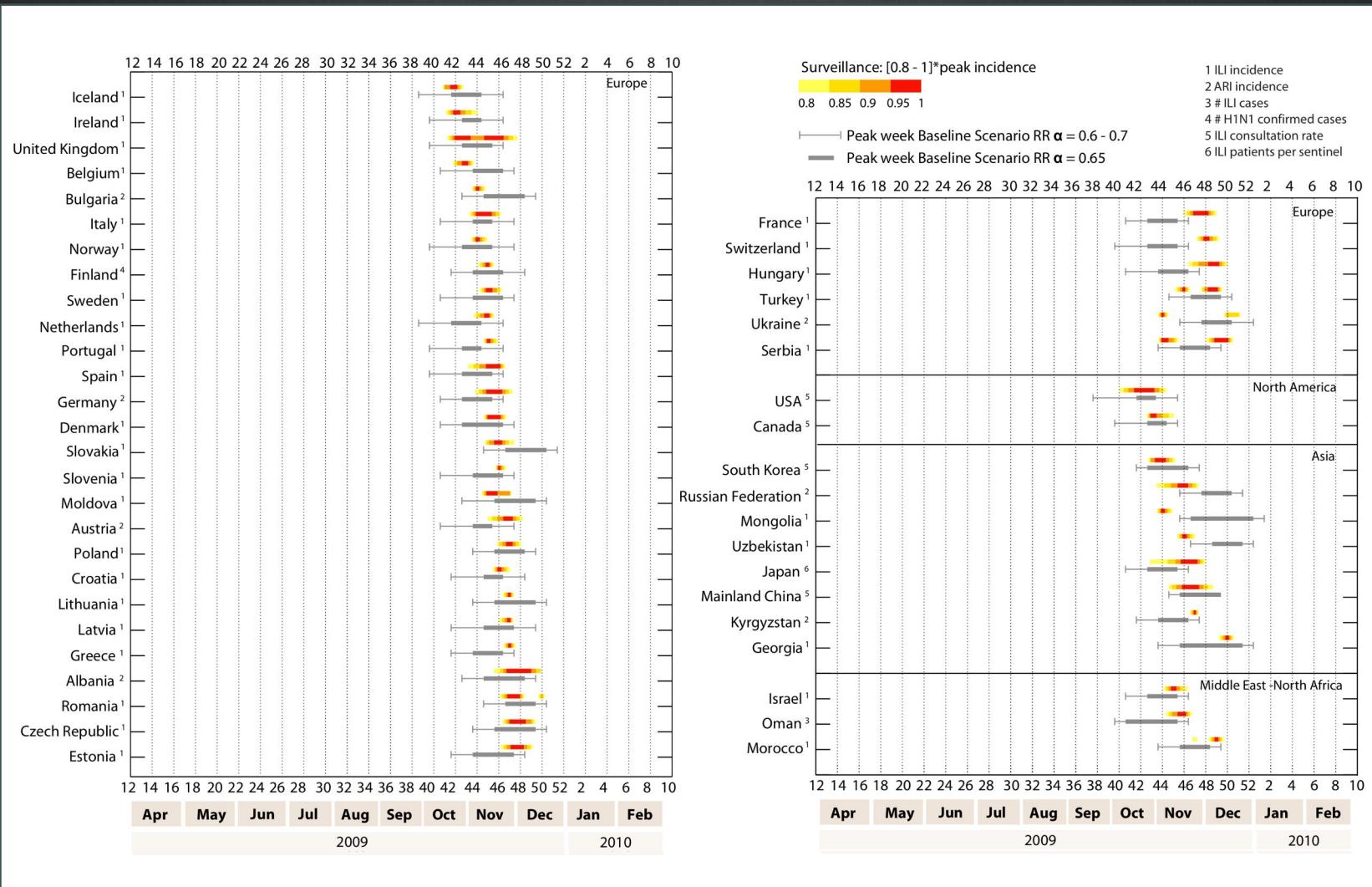
Numerical generation of thousands of infection trees (Importation/generation of the first symptomatic infectious in a given subpopulation).



Statistical distribution of the seeding time after $>10^3\text{-}10^4$ numerical stochastic realizations for each set of the parameters.



Real time forecast for the H1N1pdm (2009)



The spread of a pathogen, as predicted by GLEAM, from three initial outbreak locations. While the geographic spreading pattern is difficult to interpret, in the effective distance representation the pandemic follows a regular radial pattern (**Figure 10.31**).

The observed spreading patterns prompt us to ask: What is the speed of a typical pathogen as it spreads around the globe? The speed depends on three key parameters:

1. The basic reproduction number R_0 , which is in the vicinity of 2 for influenza type viruses (**Table 10.2**).
2. The recovery rate, which is approximately 3 days for influenza.
3. The mobility rate, which represents the total fraction of the population that travels during a day. This parameter is in the range of 0.01-0.001.

Running GLEAM (**Figure 10.26**) with these parameters we can compute the correlation between the arrival time and the geographic distance to the source of the epidemic, obtaining a speed of about 250-300 km/day. **Therefore an influenza virus moves through a continent with the speed of a sports car or of a smaller airplane.**

•

TIPPING POINT AS A FUNCTION OF HUMAN MOBILITY

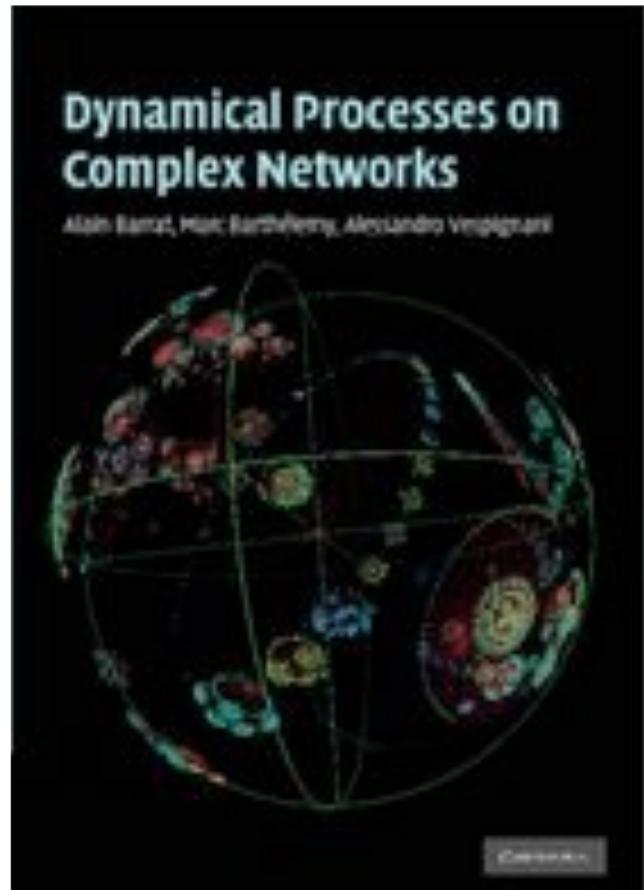
$$R_* = (R_0 - 1) \frac{\langle k^{2+2\theta} \rangle - \langle k^{1+2\theta} \rangle}{\langle k \rangle} \frac{w_0 \alpha}{\mu}$$

Tenfold traffic reduction is needed to achieve containment effects

More Info

<http://mobs-lab.org>

@alexvespi



LABORATORY FOR THE MODELING OF BIOLOGICAL
AND SOCIO-TECHNICAL SYSTEMS