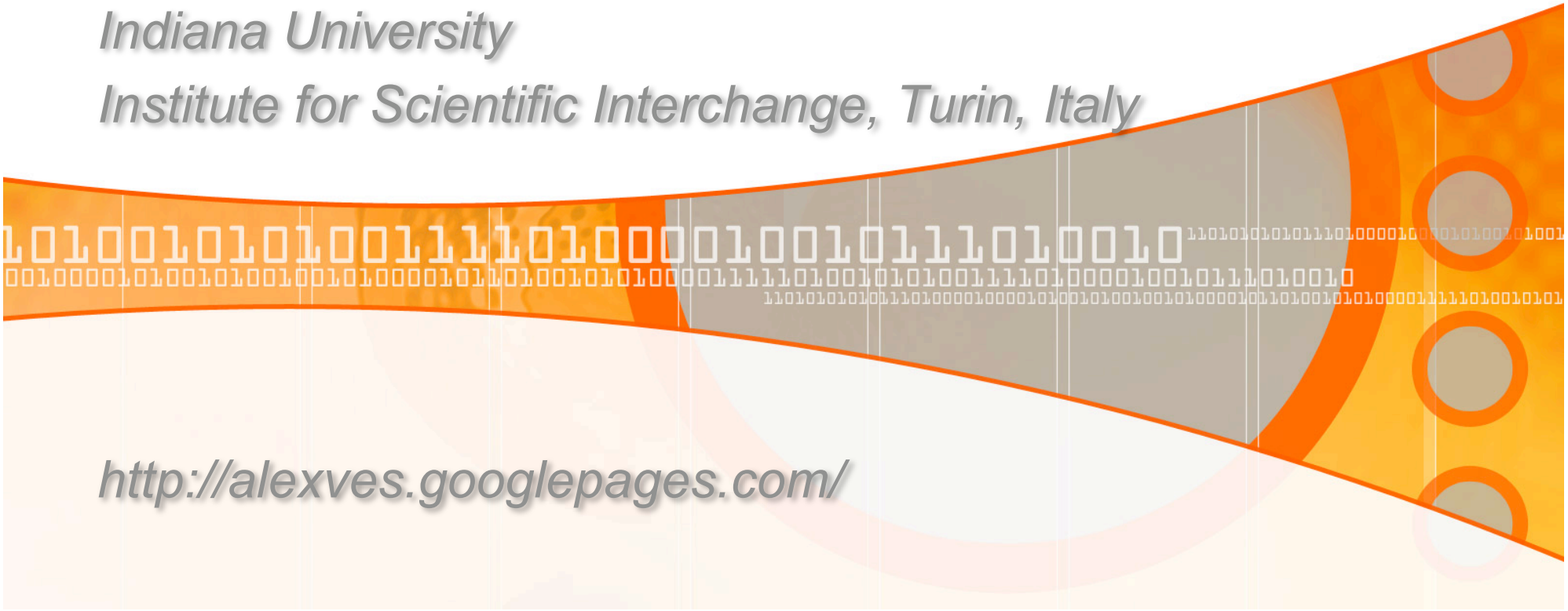
Planning for Pandemic Outbreaks with Large Scale Computational Models

A.Vespignani,
Center for Complex Networks and Systems Research (CNetS) and Pervasive Technology Institute, Indiana University
Institute for Scientific Interchange, Turin, Italy

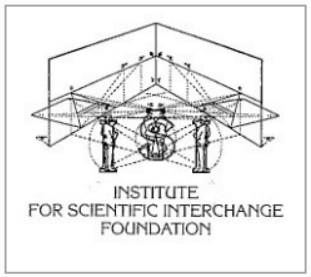


http://alexves.googlepages.com/

Collaborators....

CNetS -Center for Complex Networks and Systems Research, Indiana University

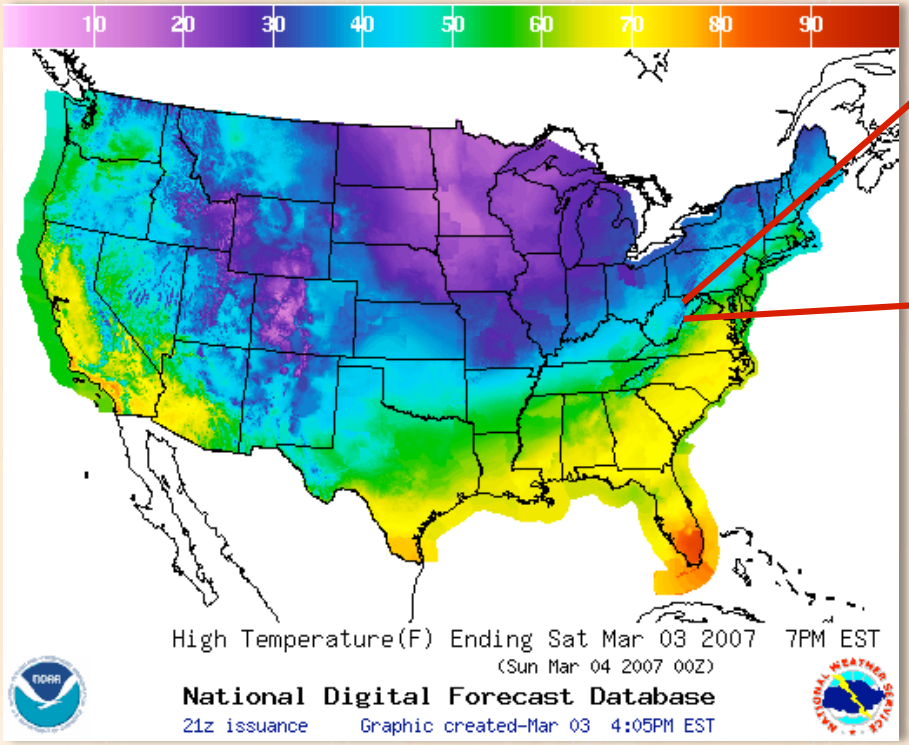
- *D. Balcan*
- *B. Goncalves*
- *H. Hu*
- *K.Borner*
- *J.Sherman*

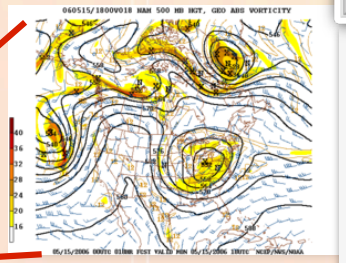


- *V. Colizza*
- *C. Cattuto*
- *D.Paolotti*
- *J. Ramasco*
- *V. Van den Broeck*

- *M. Ajelli*
- *A. Barrat*
- *M. Barthelemy*
- *S. Merler*
- *R. Pastor Satorras*
- *A.J.Valleron*

Weather forecast





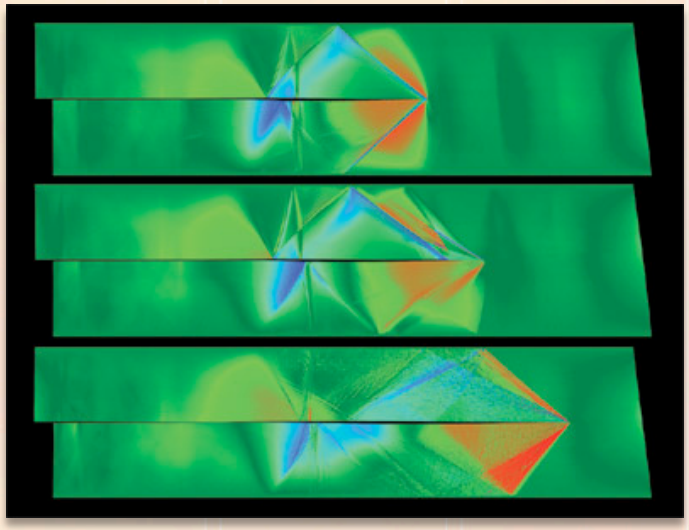
Parameters

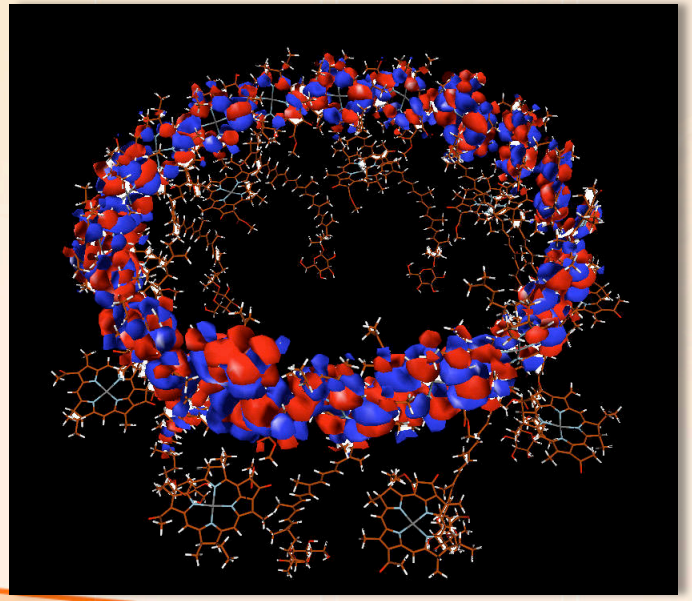
u is the zonal velocity (velocity in the east/west direction tangent to the sphere).
v is the meridional velocity (velocity in the north/south direction tangent to the sphere).
ω is the vertical velocity
T is the temperature
φ is the geopotential
f is the term corresponding to the Coriolis force, and is equal to 2Ωsin(φ), where Ω is the angular rotation rate of the Earth (2π / 24 radians/hour), and φ is the latitude.
R is the gas constant
p is the pressure
cp is the specific heat
J is the heat flow per unit time per unit mass
π is the exner function
θ is the potential temperature

Numerical weather prediction uses mathematical models of the atmosphere to predict the weather. Manipulating the huge datasets with the most powerful supercomputers in the world.

The primitive equations can be simplified into the following equations:
Temperature: $\partial T/\partial t = u\,(\partial Tx/\partial X) + v\,(\partial Ty/\partial Y) + w\,(\partial Tz/\partial Z)$
Wind in E-W direction: $\partial u/\partial t = \eta v - \partial \Phi/\partial x - Cp\,\theta\,(\partial \pi/\partial x) - z\,(\partial u/\partial \sigma) - [\partial(u2 + y)/2]/\partial x$
Wind in N-S direction: $\partial v/\partial t = -\eta(u/v) - \partial \Phi/\partial y - Cp\,\theta\,(\partial \pi/\partial y) - z\,(\partial v/\partial \sigma) - [\partial(u2 + y)/2]/\partial y$
Precipitable water: $\partial W/\partial t = u\,(\partial Wx/\partial X) + v\,(\partial Wy/\partial Y) + z\,(\partial Wz/\partial Z)$
Pressure Thickness: $\partial(\partial p/\partial \sigma)/\partial t = u\,[(\partial p/\partial \sigma)x/\partial X] + v\,[(\partial p/\partial \sigma)y/\partial Y] + z\,[(\partial p/\partial \sigma)z/\partial Z]$

Super-computer simulations

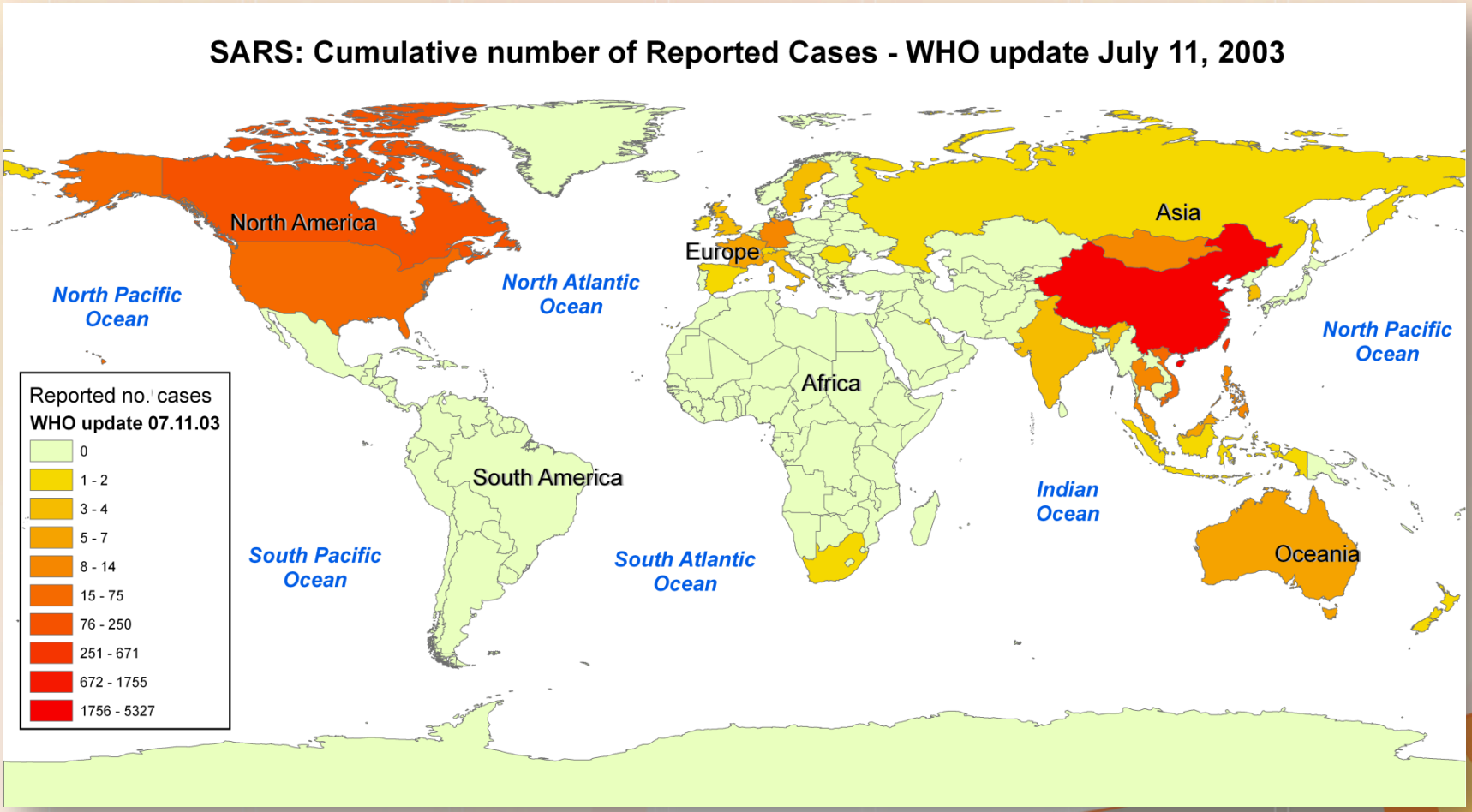




- Fracture in 1.6 millions atoms material
- 6.8 billion finite elements plasma
- Ab initio simulations of thousand of atoms pico-second scale
- .....

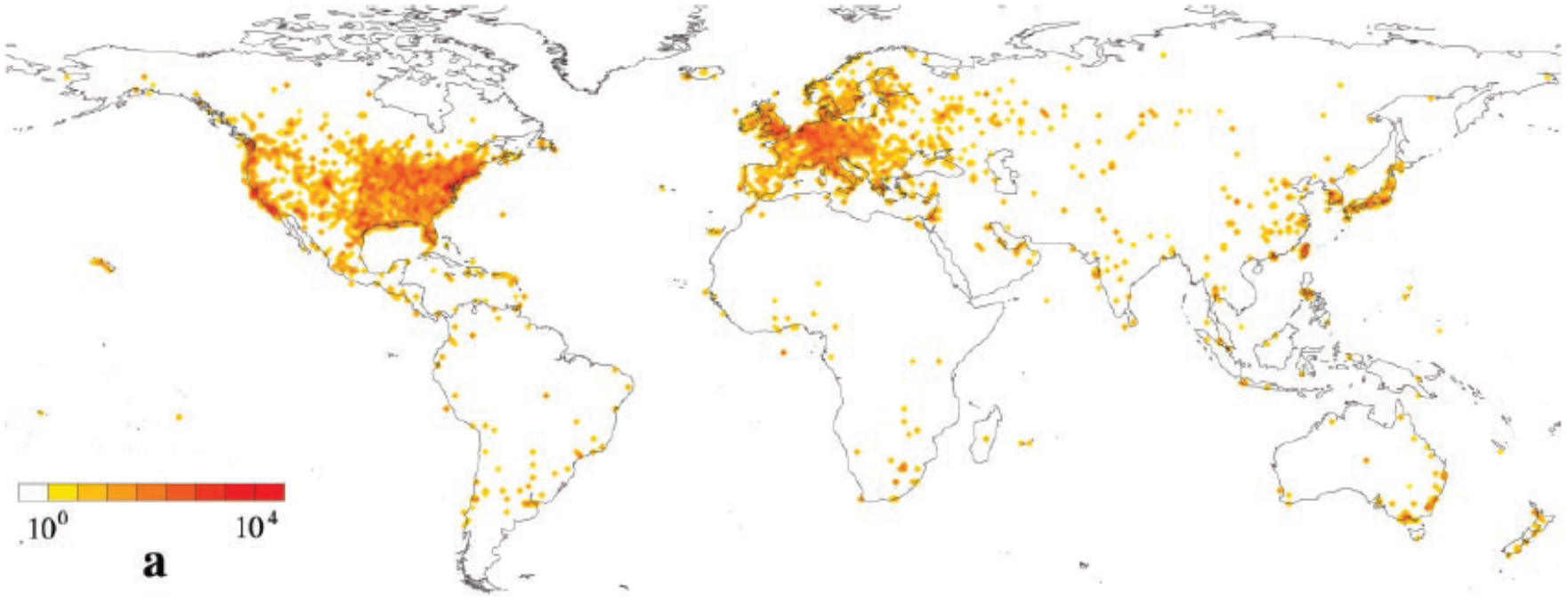
Why not forecast on…

Emerging disease spreading evolution



WHY NOT FORECAST ON…

Router density worldwide in ten years



- What will happen when 1.5 billion people from china will scream connectivity??
- Will BGP collapse ??

Collective Bio-Techno-social problems ....

In other words

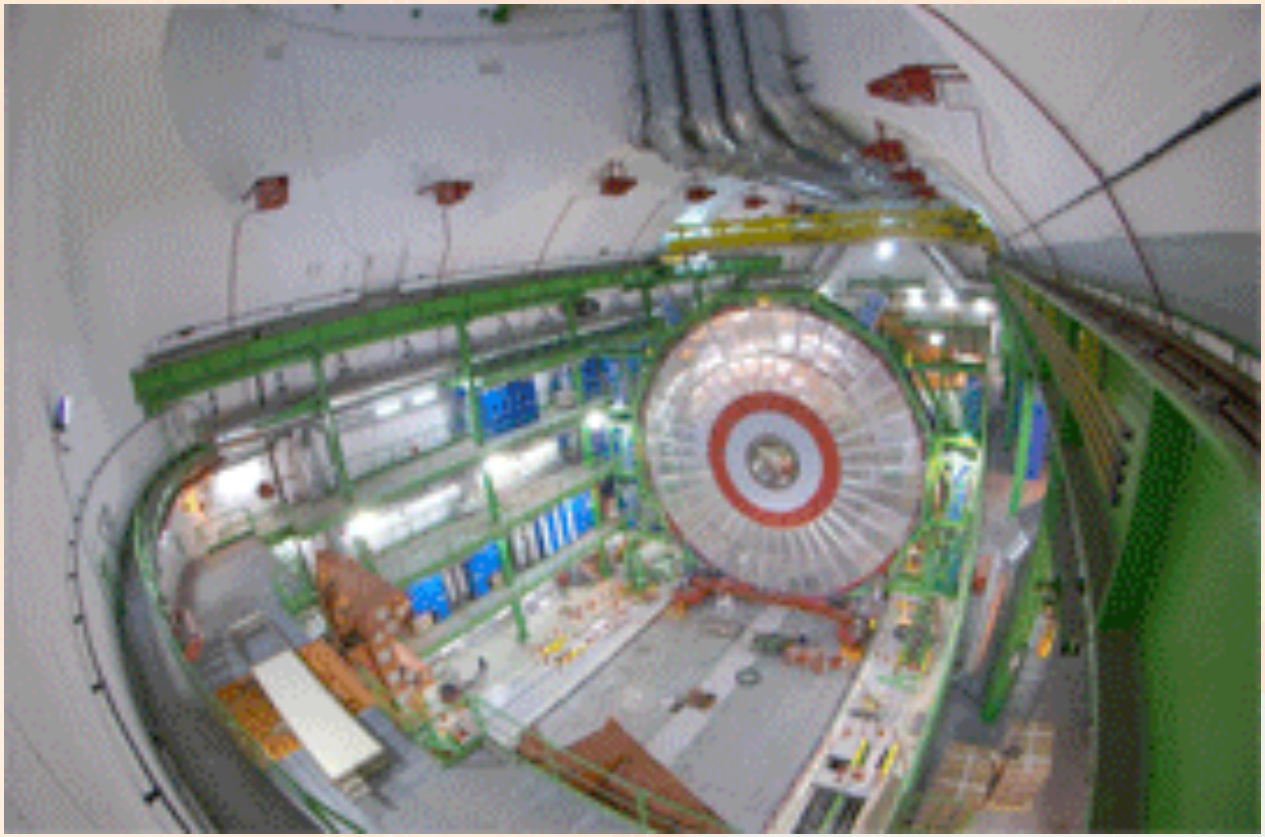
The complete temperature analysis of the sea surface, and satellite images of atmospheric turbulence are easier to get than the large scale knowledge of commuting patterns or the quantitative measure of the propensity of a certain social behavior, or the spreading rate of a given pathogen.

Complexity+complications

- *large numbers of heterogeneous individuals*
- *over multiple time and size scales*
- *Non-linearity, threshold effects, discreteness, cooperation*
- *huge richness of cognitive/social science problems*
- *Technological infrastructures*
- *Bio-medical understanding*

Hubble telescope >6.5 Billion

Large Hadron Collider >8 Billion



Large Social collider

SOCIAL SCIENCE

Computational Social Science

A field is emerging that leverages the capacity to collect and analyze data at a scale that may reveal patterns of individual and group behaviors.

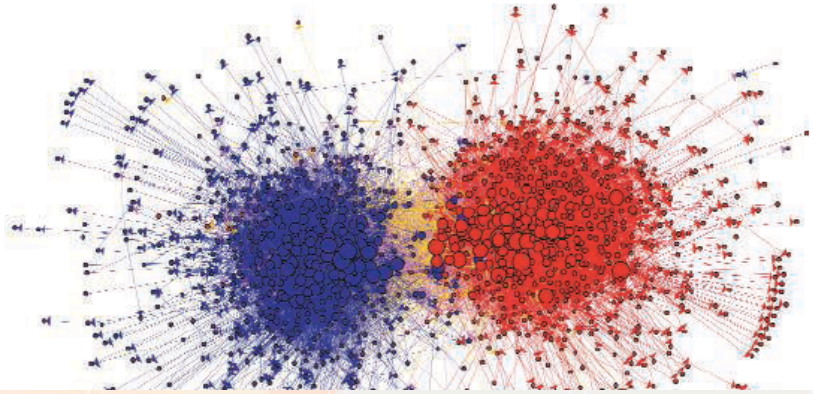
David Lazer,[1] Alex Pentland,[2] Lada Adamic,[3] Sinan Aral,[2,4] Albert-László Barabási,[5] Devon Brewer,[6] Nicholas Christakis,[1] Noshir Contractor,[7] James Fowler,[8] Myron Gutmann,[3] Tony Jebara,[9] Gary King,[1] Michael Macy,[10] Deb Roy,[2] Marshall Van Alstyne[2,11]

We live life in the network. We check our e-mails regularly, make mobile phone calls from almost any location, swipe transit cards to use public transportation, and make purchases with credit cards. Our movements in public places may be captured by video cameras, and our medical records stored as digital files. We may post blog entries accessible to anyone, or maintain friendships through online social networks. Each of these transactions leaves digital traces that can be compiled into comprehensive pictures of both individual and group behavior, with the potential to transform our understanding of our lives, organizations, and societies.

The capacity to collect and analyze massive amounts of data has transformed such fields as biology and physics. But the emergence of a data-driven "computational social science" has been much slower. Leading journals in economics, sociology, and political science show little evidence of this field. But computational social science is occurring—in Internet compa-

ment agencies such as the U.S. National Security Agency. Computational social science could become the exclusive domain of private companies and government agencies. Alternatively, there might emerge a privileged set of academic researchers presiding over private data from which they produce papers that cannot be critiqued or replicated. Neither scenario will serve the long-term public interest of accumulating, verifying, and disseminating knowledge.

What value might a computational social science—based in an open academic environment—offer society, by enhancing understanding of individuals and collectives? What are the



IT AND THE SOCIETY "TOMOGRAPHY"

- Data integration and assimilation
- Mobile telephone and devices..
- Web 2.0, Proxy networks
- Pervasive and embedded technology/ sensors

UNPRECEDENTED AMOUNT 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)

COMPLEX MOBILITY AND TRAFFIC NETS….

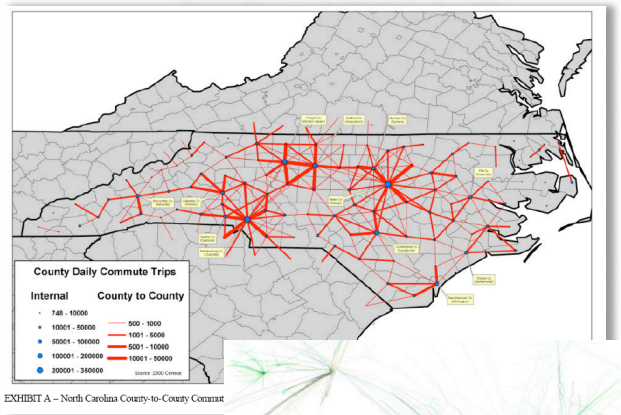
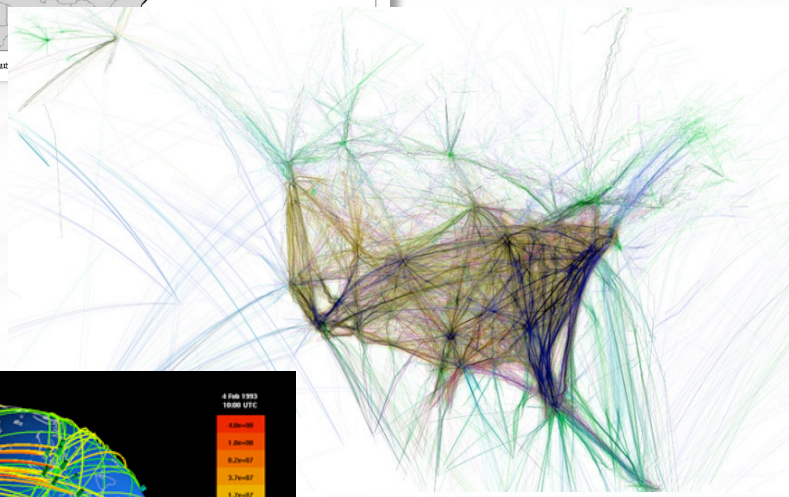
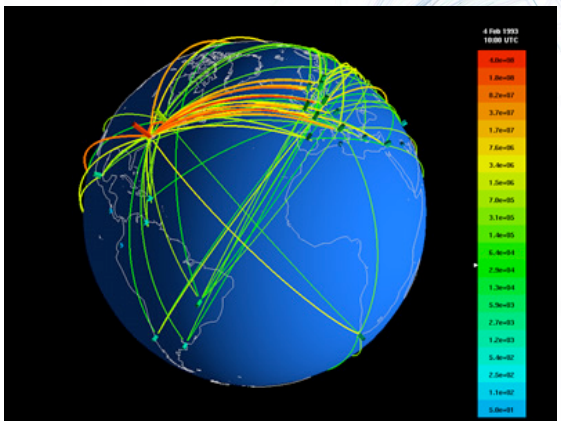


EXHIBIT A – North Carolina County-to-County Commut

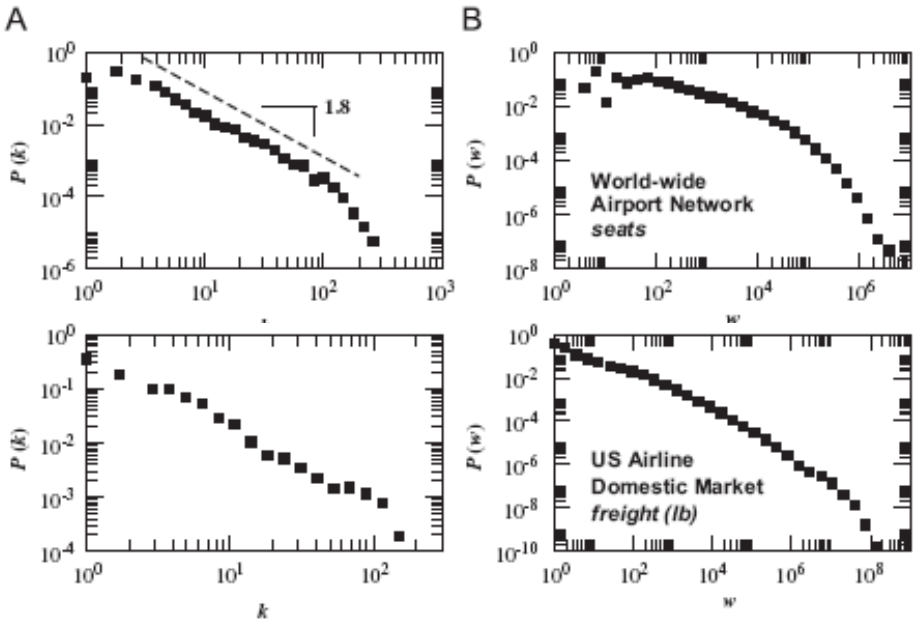




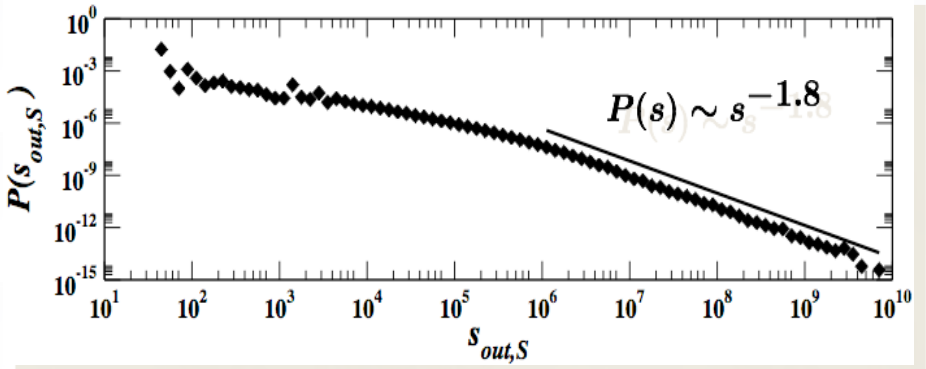
- Which are the topology and traffic/flows statistical properties of real world networks.

CHARACTERIZATION OF WEIGHTED NETWORKS

Transportation and freight

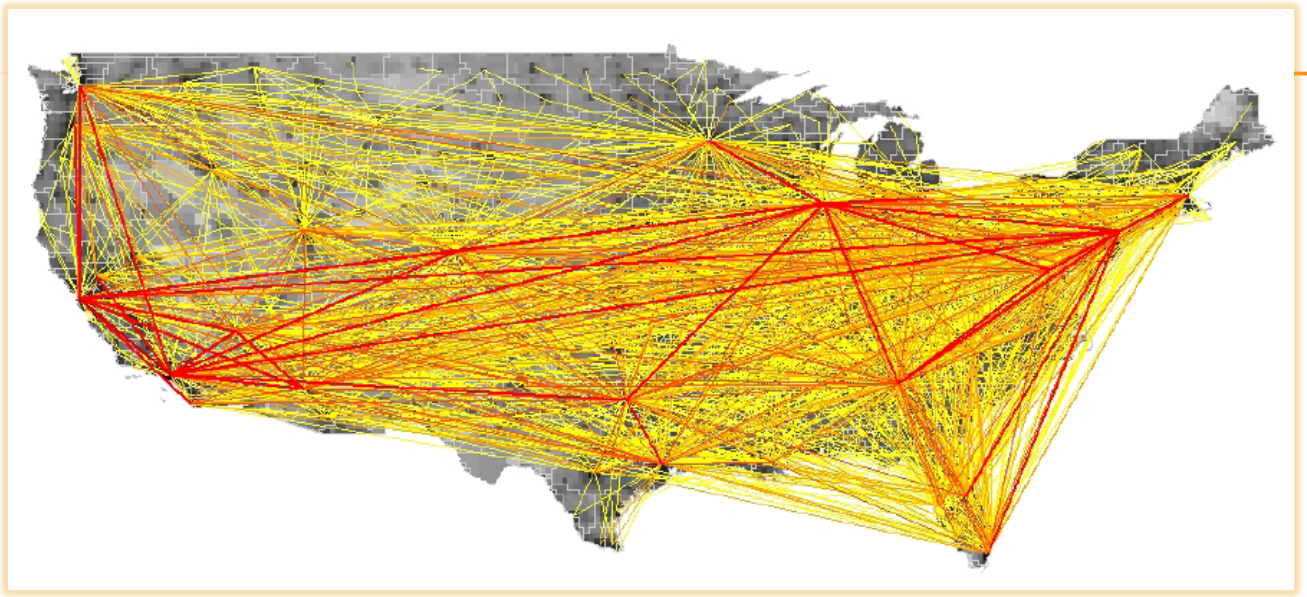


Web traffic

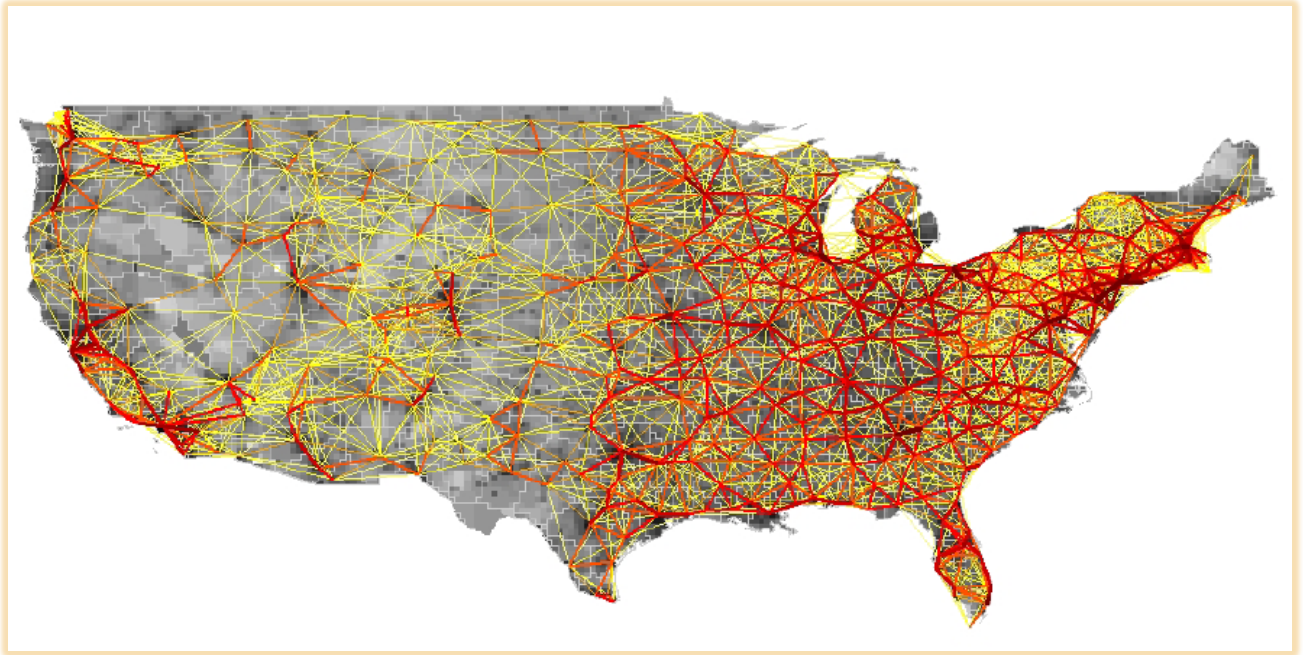


- General characterization of weighted complex networks:
  - + Definition of appropriate quantities
  - + Correlation measure
  - + Assortativity/disassortativity
  - + Backbone identification
  - + Traffic topology relation
- Dynamical models of growing weight networks

A Multiscale problem

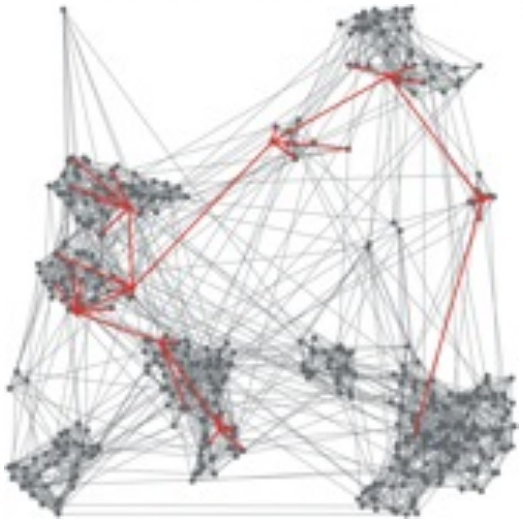


Airline transportation network



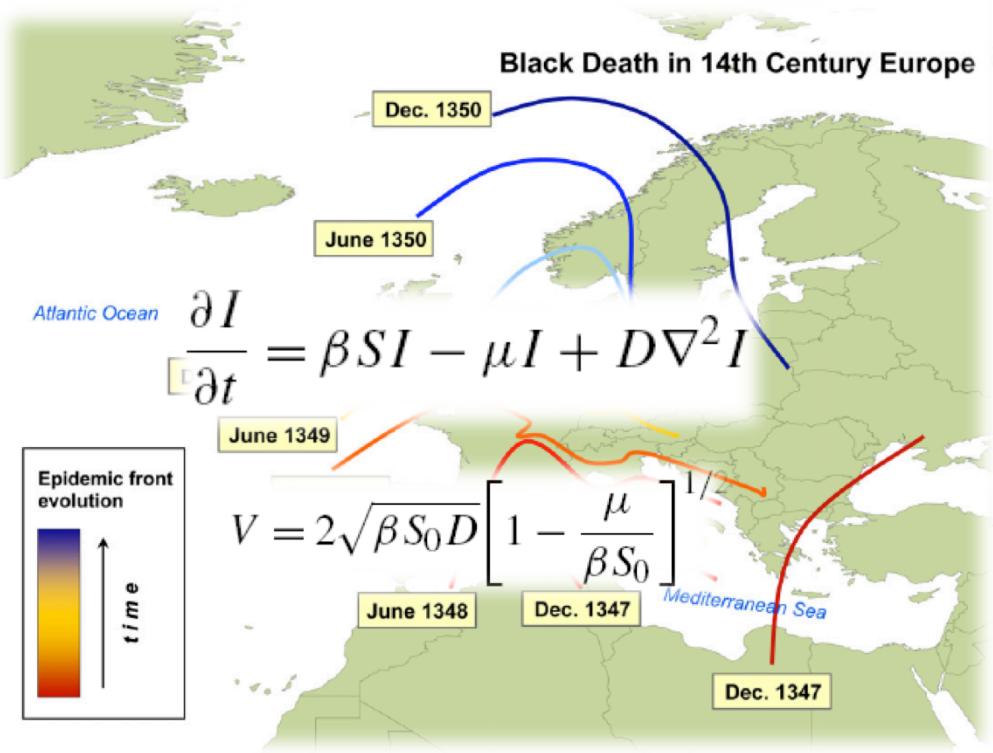
Daily commuting network

OPEN ISSUES….

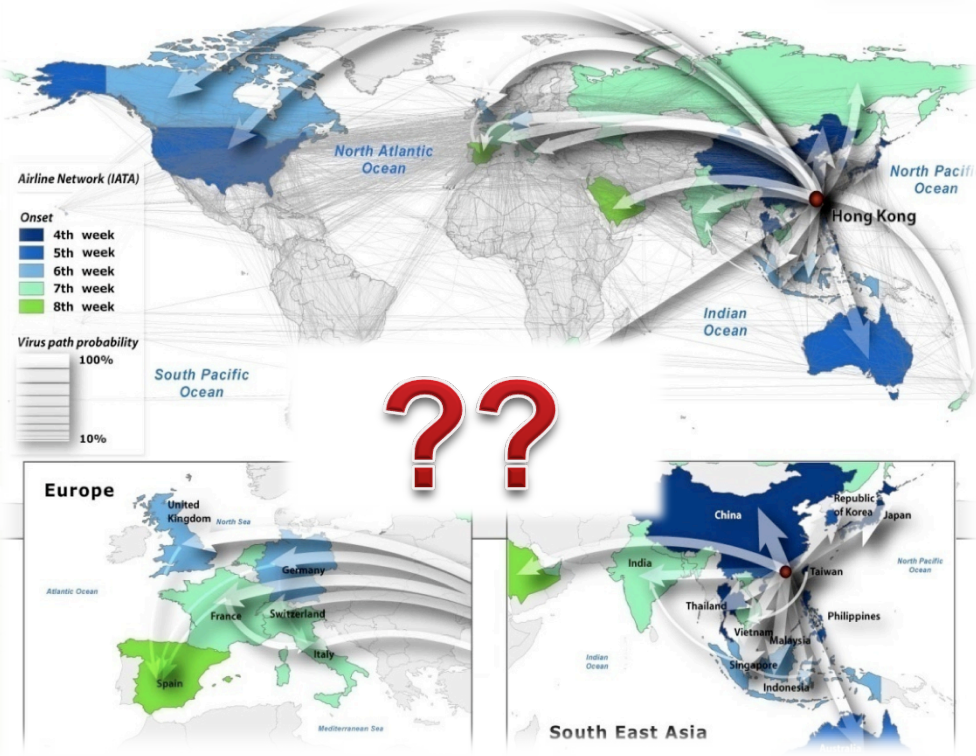


- spatially structured models depend on the notion of identifiable populations in a meta-population.
- For instance, cities, town and villages (on a smaller scale schools, workplaces) and homes, cell tower, WiFi base station area range, etc.
- intuitive segmentations is often fuzzy, ill defined, arbitrary and lacks a systematic backbone.
- How much geography, cultural diversity, spatial and technological variability is encoded in the topology and mobility/diffusion of multi-scale networks

BLACK DEATH IN1347: A CONTINUOUS DIFFUSION PROCESS



SARS EPIDEMICS: A DISCRETE NETWORK DRIVEN PROCESS



NETWORKS AND NETWORK THINKING

Epidemic forecast infrastructure

biology
individuals
society
environment

...complexity...

epidemic processes

predictions ??

predictability ??

Data

demographic data

transportation data

geographic data

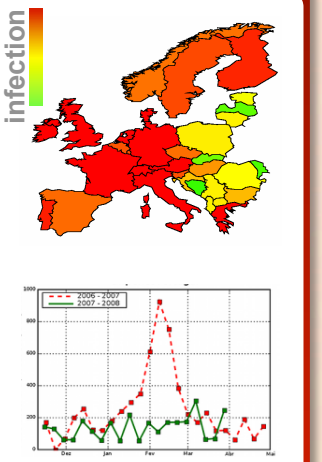
Models

contact network models

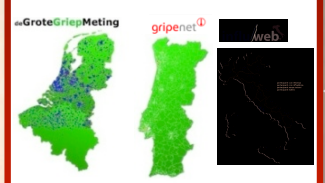
Multi-scale models

hospital
school
home
work
Agent based models

Forecasts



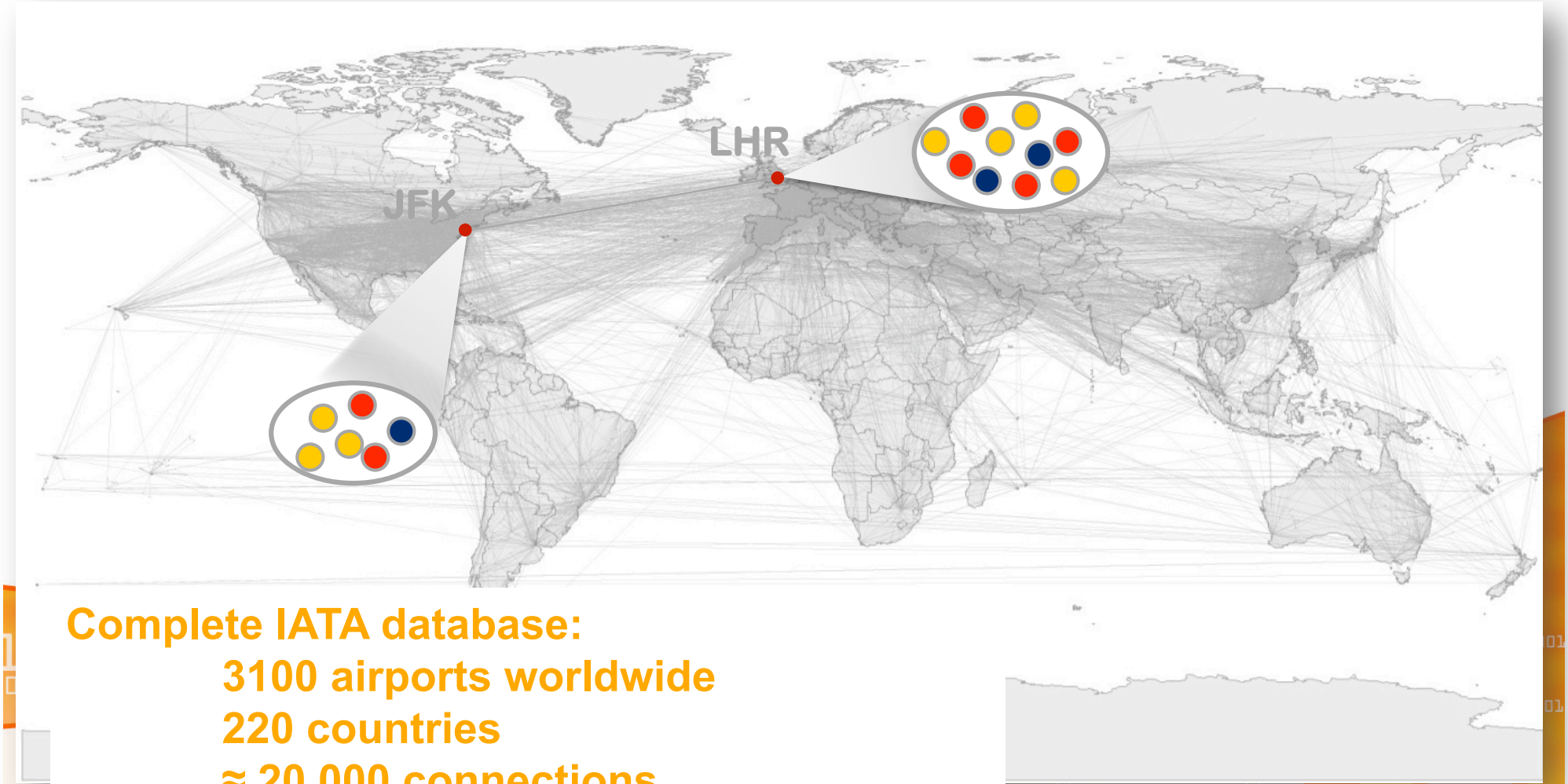
Monitoring



- **Epidemic Marketplace Platform**
- **Epidemic Modelling Platform**

Developing the framework for an epidemic forecast infrastructure

Global Epidemic Modeler (GLEaM) platform: Ab-initio modeling of the global spread of epidemics



Complete IATA database:

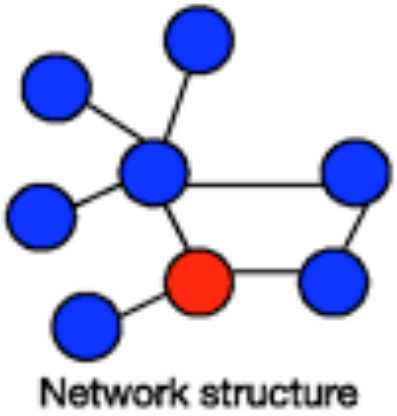
- 3100 airports worldwide
- 220 countries
- ≈ 20,000 connections
- w_{ij} #passengers on connection i-j
- >99% total traffic

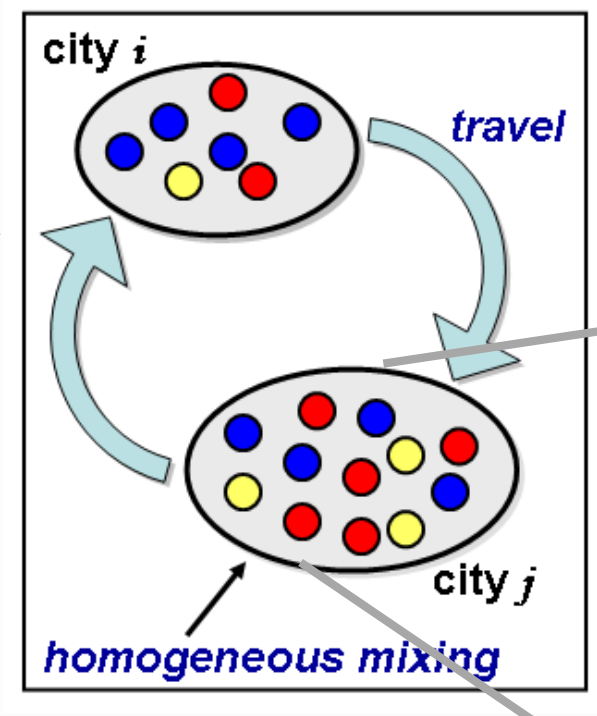
Colizza, Barrat, Barthelemy & Vespignani, PNAS (2006)

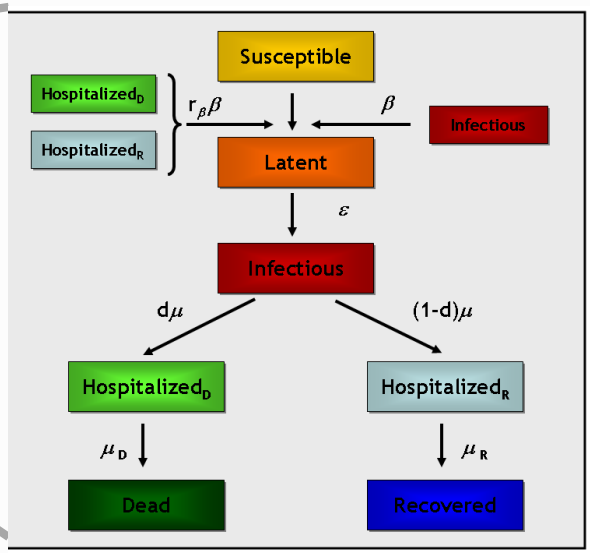
What's in GLEaM...

- Refined census data (Up to 2.5 arc/min resolution)
- IATA/OAG airline database
- Commuting data for about ~40 countries in 4 continents (definition of a world-wide commuting networks)
- Disease structure and population heterogeneity.
- Multiscale force of infection
- Intervention scenarios (vaccination, contact tracing, secondary infections, quarantine..)

Mechanistic metapopulation models...

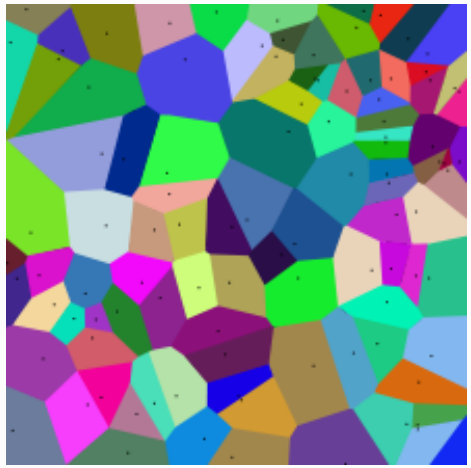






Intra-population infection dynamics

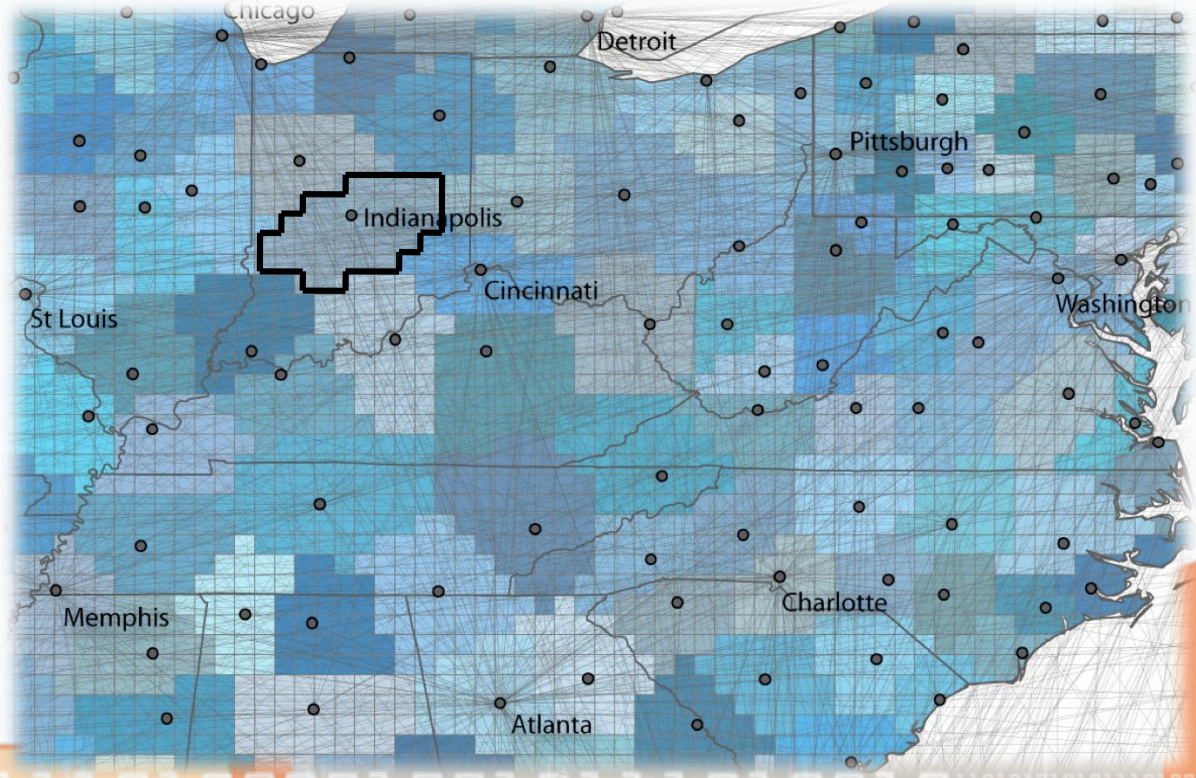
Metapopulation structure via a Voronoi decomposition



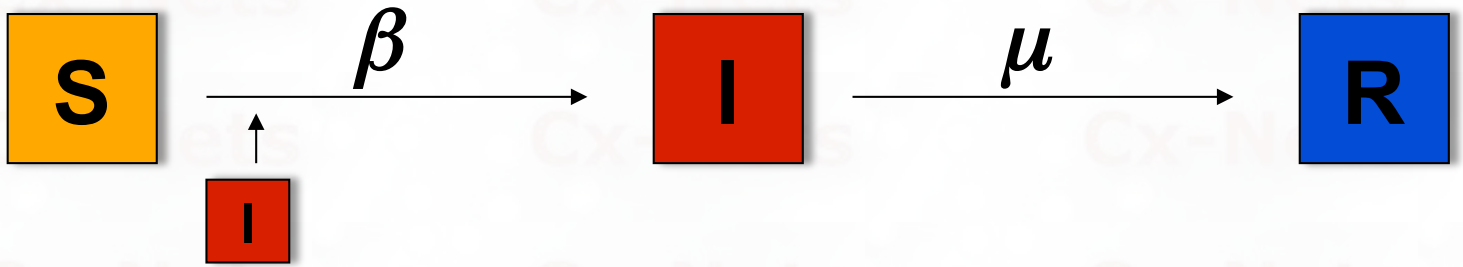
Given a set of points *S* in the plane, which are the Voronoi sites, each site *s* has a Voronoi cell *V(s)* consisting of all points closer to *s* than to any other site.

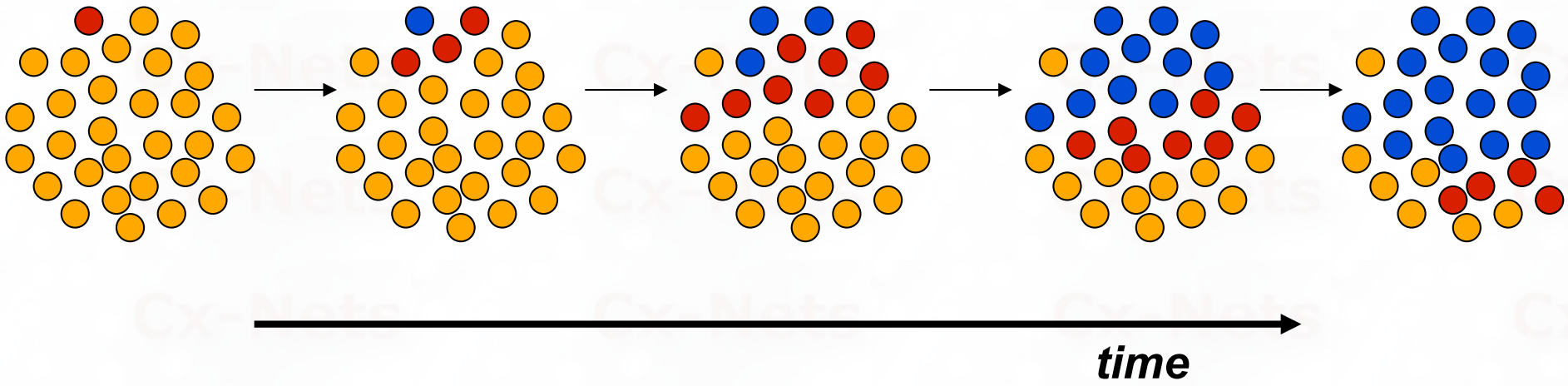
The segments of the Voronoi diagram are all the points in the plane that are equidistant to two sites. The Voronoi nodes are the points equidistant to three (or more) sites.

Voronoi sites
=
airports location



INTRA-POPULATION: THE SIR MODEL



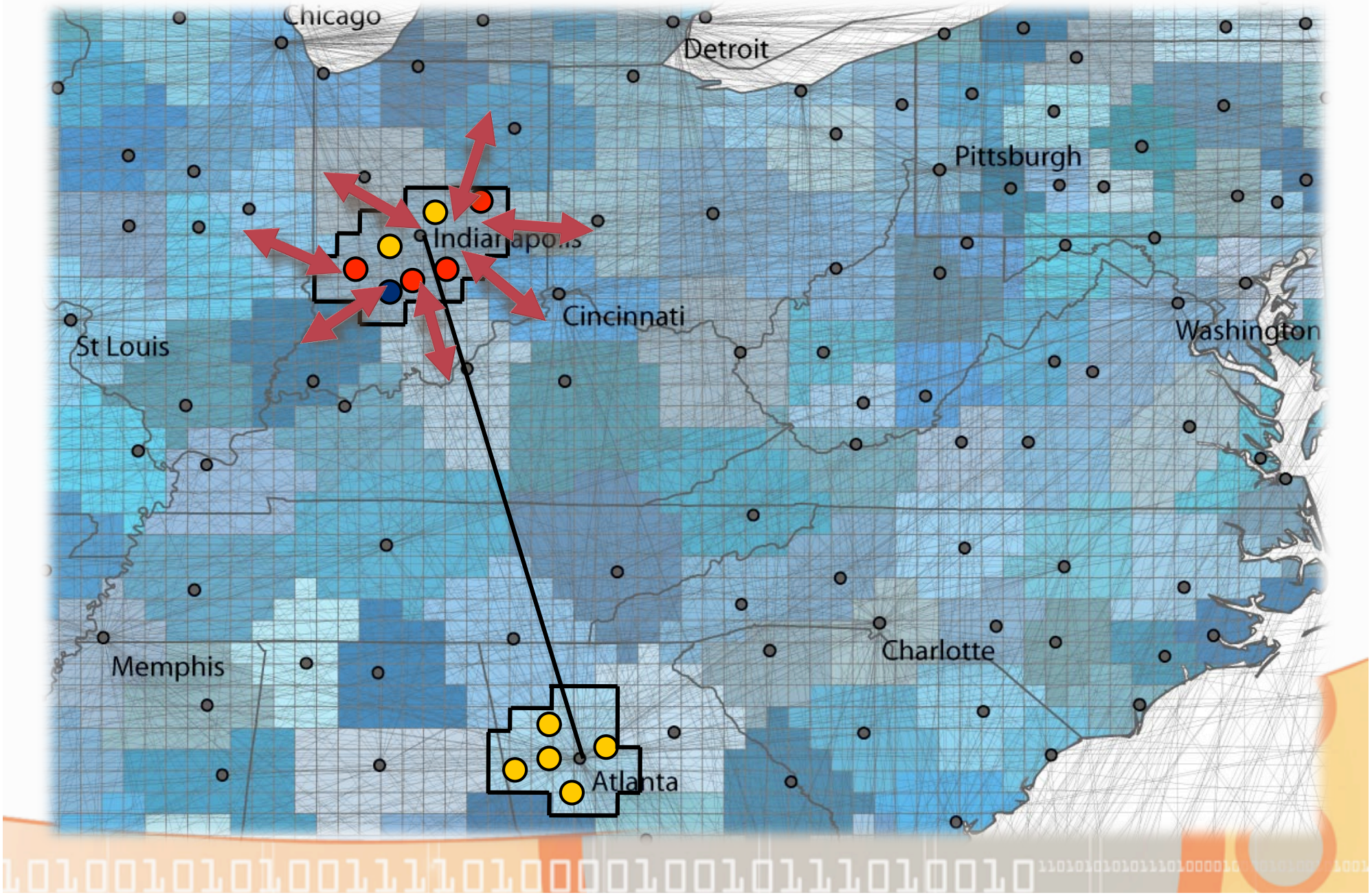


Mobility modeling (0th order)

- Probability that any individual in the class X travel from $j \rightarrow l$
  - Proportional to the traffic flow
  - Inversely proportional to the population

$$p_{jl} = \frac{w_{jl}}{N_j} \Delta t$$

Long and Short range mobility (I)



Computational model

$$S_{j,t+\Delta t} = S_{j,t} - Binom_j(S_{j,t}, \beta \Delta t\, I_{j,t}/N) + \Omega_j(S)$$

$$I_{j,t+\Delta t} = I_{j,t} + Binom_j(S_{j,t}, \beta \Delta t\, I_{j,t}/N) - Binom_j(I_{j,t}, \mu \Delta t) + \Omega_j(I)$$

$$R_{j,t+\Delta t} = R_{j,t} + Binom_j(I_{j,t}, \mu \Delta t) + \Omega_j(R)$$

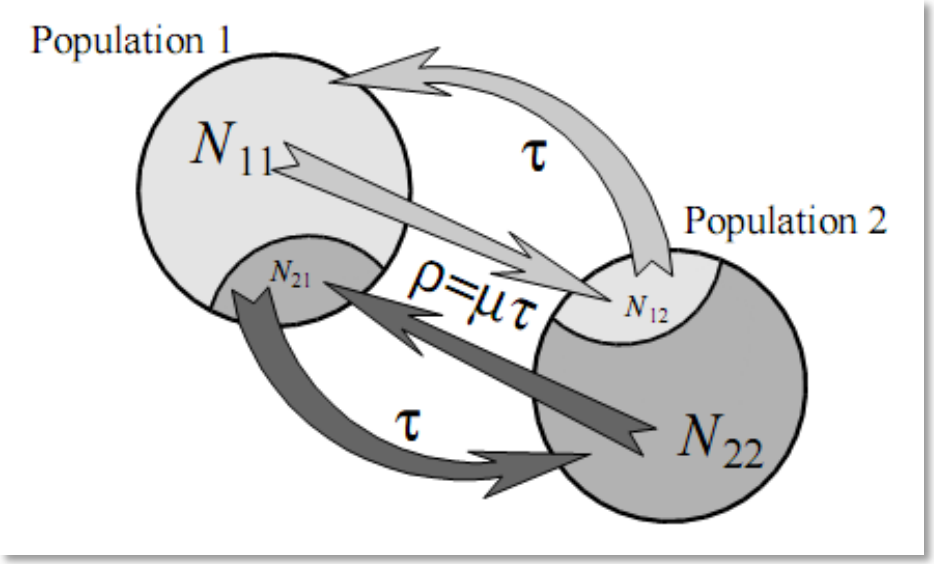
Stochastic coupling terms

=

3,000 -20,000 coupled stochastic equations to integrate

Time scale separation technique

- *Theoretical description of reaction diffusion processes with memory and "return rates".*
- *define μ= ρ / τ (leaving rate / returning rate) ,*
- *N_{ij}: population in i who is visiting region j*
- *N_{ii}: population in i who is actually present in the home region*



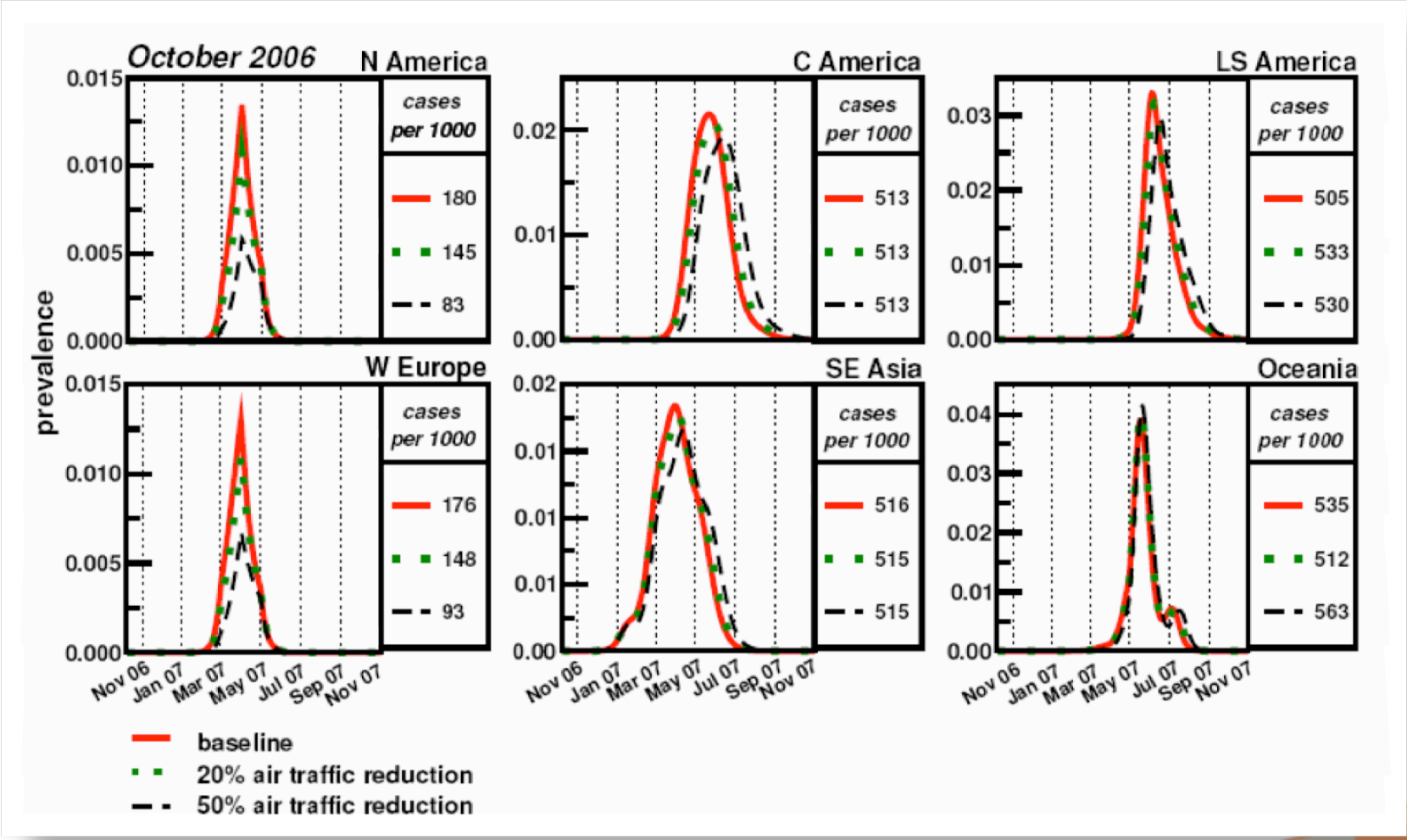
Time scale separation *μ* <<1 ; heterogeneous population sizes

$$p_\ell^{\mathcal{D}}(S \to L) = \frac{\beta_\ell}{(1+\widetilde{\mu}_\ell)N_\ell^*}\left[\frac{r_\beta I_\ell^a + I_\ell^t}{1+\widetilde{\mu}_\ell} + I_\ell^{nt} + \sum_k \frac{\widetilde{\mu}_k \widetilde{\nu}_{k\ell}}{1+\widetilde{\mu}_k}(r_\beta I_k^a + I_k^t)\right]\Delta t$$

$$+ \sum_k \frac{\beta_k \widetilde{\mu}_\ell \widetilde{\nu}_{\ell k}}{(1+\widetilde{\mu}_\ell)N_k^*}\left[\frac{r_\beta I_k^a + I_k^t}{1+\widetilde{\mu}_k} + I_k^{nt}\right]\Delta t$$

GEM applications

- *Travel restriction inefficacy*



GEM applications (I)

Colizza & Vespignani, 2008
Journal of Theoretical Biology

$$p_c \bar{N} \geq \frac{\langle k \rangle^2}{\langle k^2 \rangle} \frac{\mu R_0^2}{2(R_0 - 1)^2}$$

	Airport network
Average population N	$\approx 10^5$
$\langle k \rangle^2 / \langle k^2 \rangle$	$\approx 10^{-1}$ -10^{-2}
μ	1/3
R_0	1.1
p_c	$\approx 10^{-5}$-10^{-6}
p_{real}	$\approx 10^{-3}$ -10^{-4}

Tenfold traffic reduction is needed to achieve noticeable effects

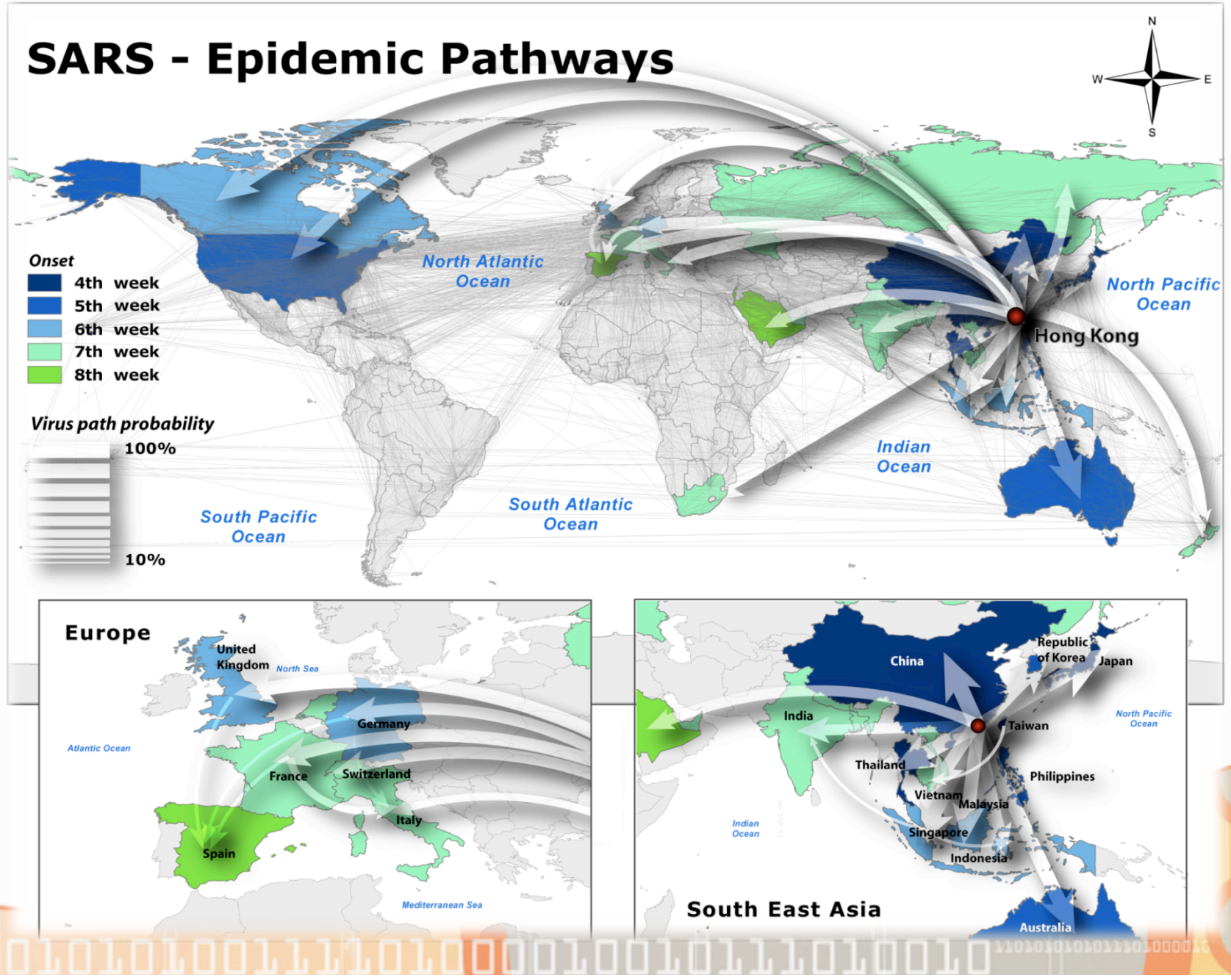
Predictability: Taking advantage of complexity…

- *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*

GEM applications

*Colizza eta al.BMC Medicine **5**, 34(2007)*



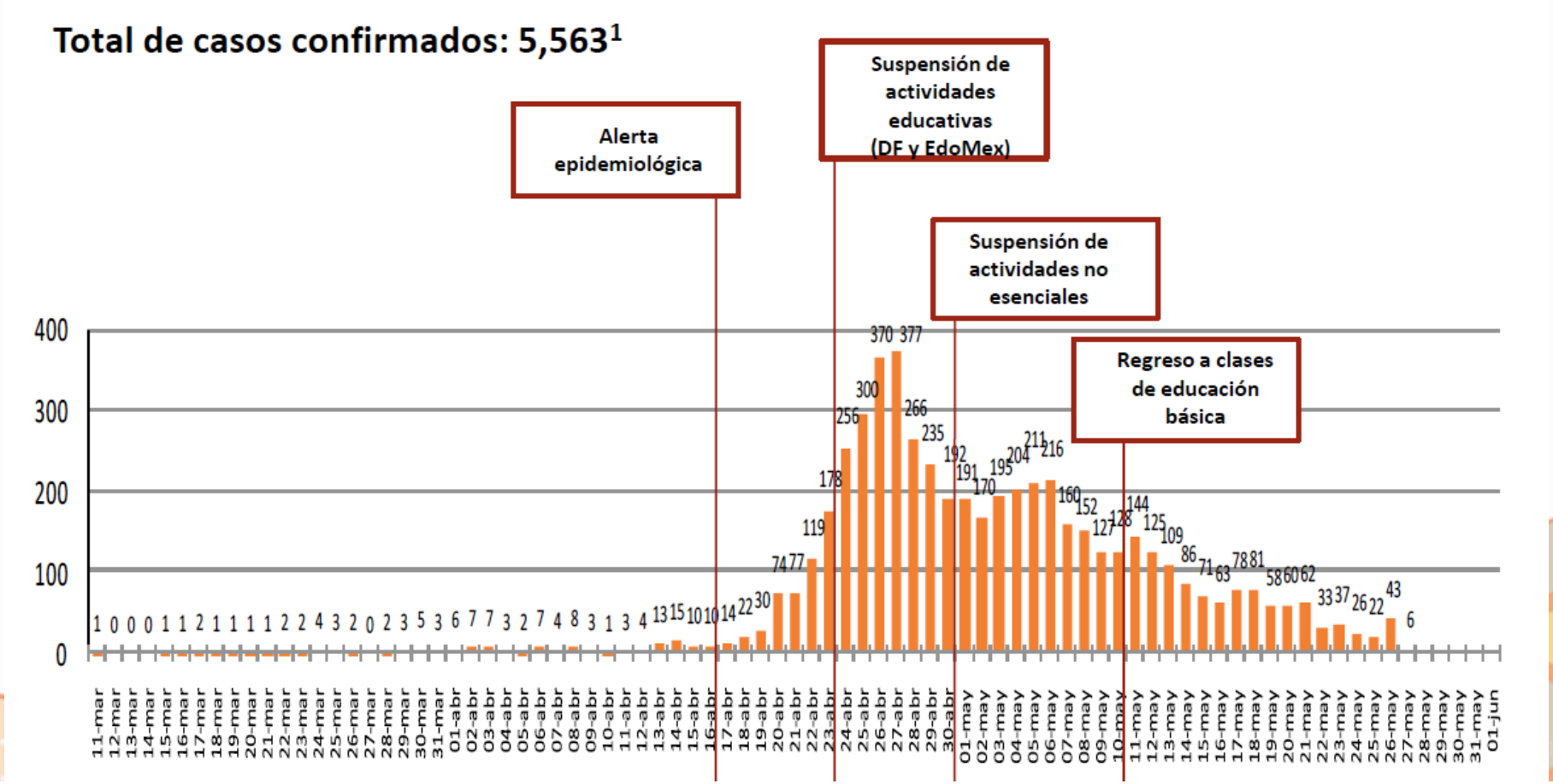
H1N1 epidemic (Feb. 2009, Mexico):

Can we do forecasts in real time ?

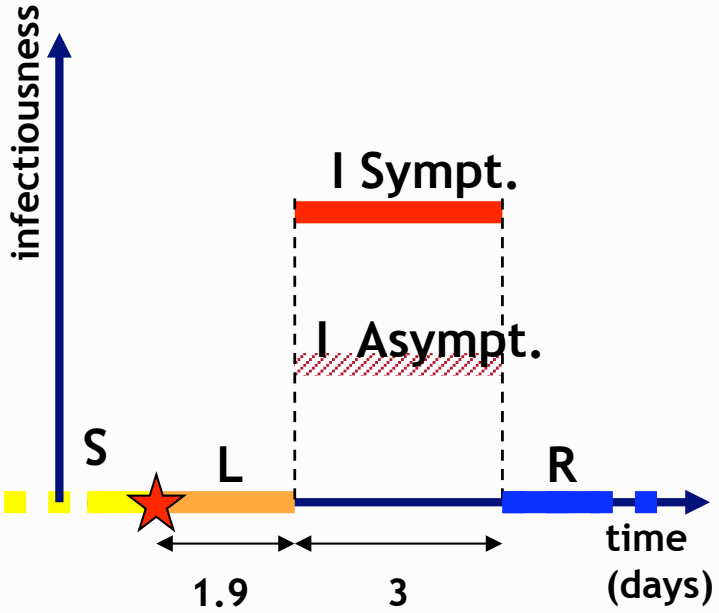
What kind of forecasts? (short-term; long-term)

What we need to do forecast?

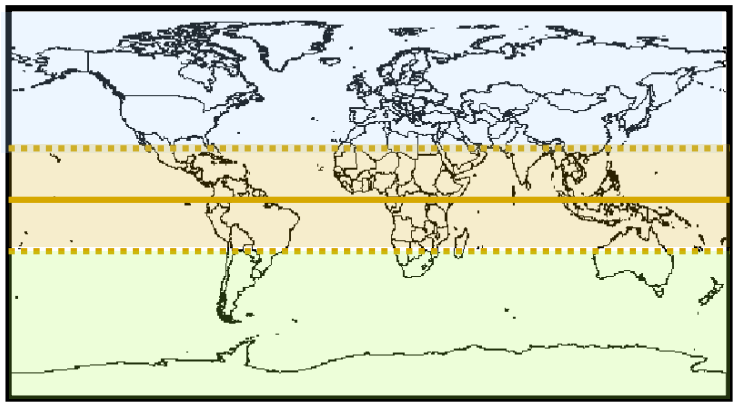
Are forecasts reliable?

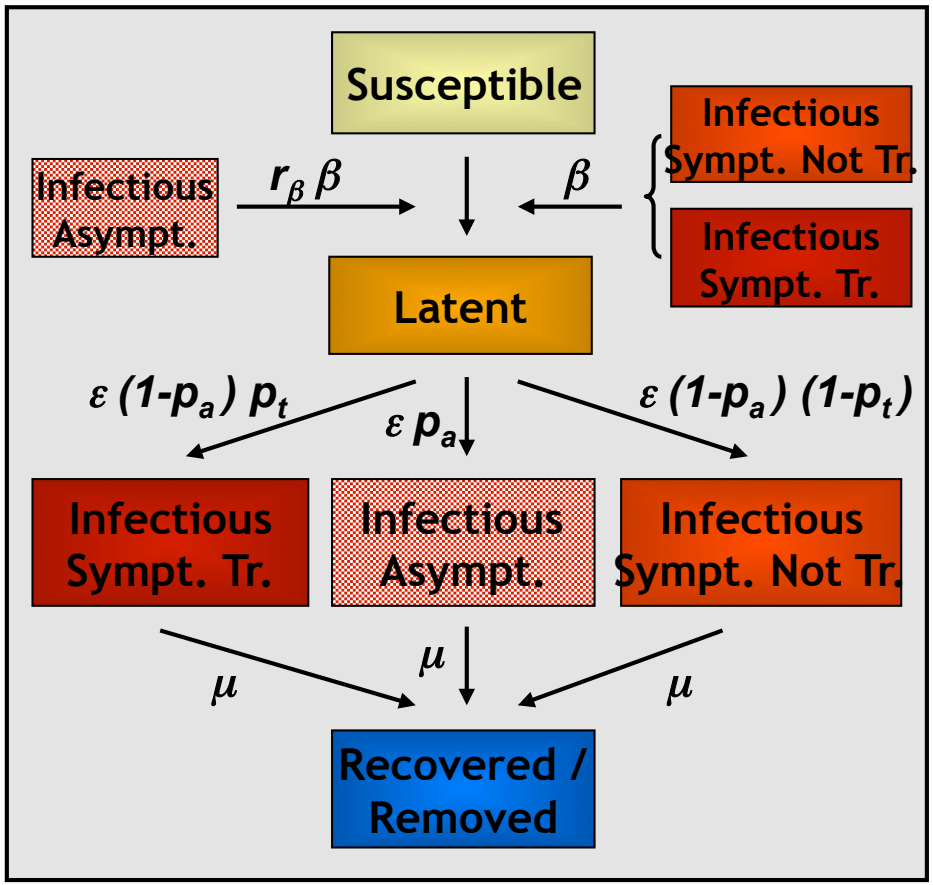


Scenarios and policy making



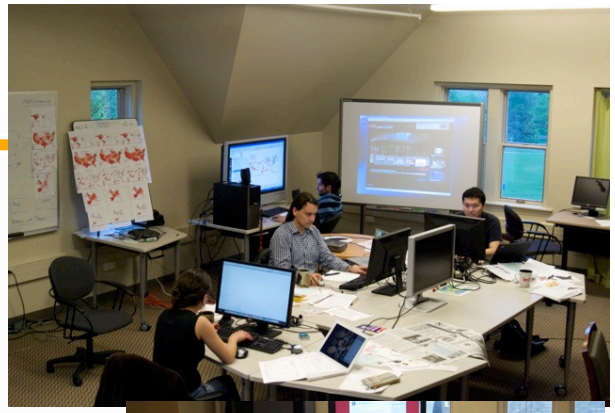
Seasonality





Longini *et al. Am. J. Epid.* (2004)

Strategy:



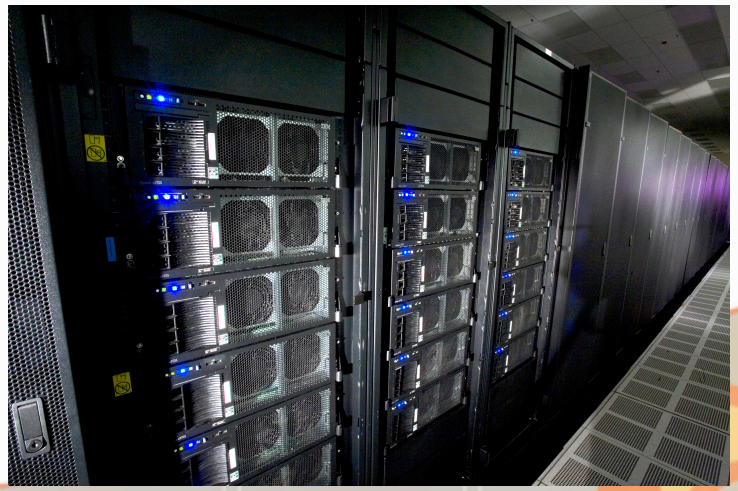
Fit the basic parameters of the global model
Two kinds: a) #of cases
b) arrival time

Sensitivity analysis

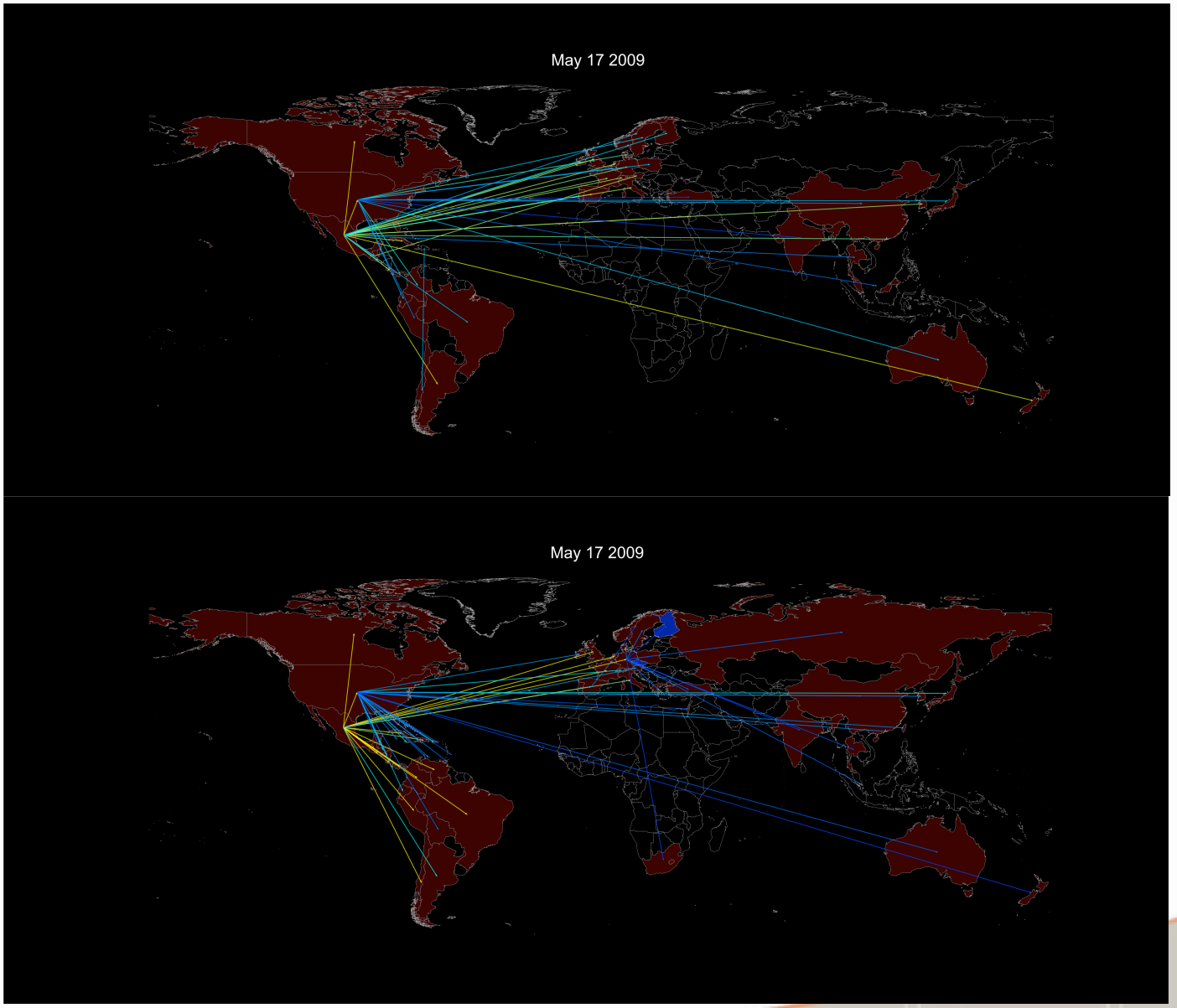
Constant recalibration with new data

The exploration of each point of the parameter phase space amount at 5,000 stochastic realizations for two or more months of the outbreak.

Indiana University's supercomputer system -- named **"Big Red"** -- is the fastest supercomputer owned and operated by a U.S. university and the 23rd fastest supercomputer in the world



Innovative calibration by arrival times

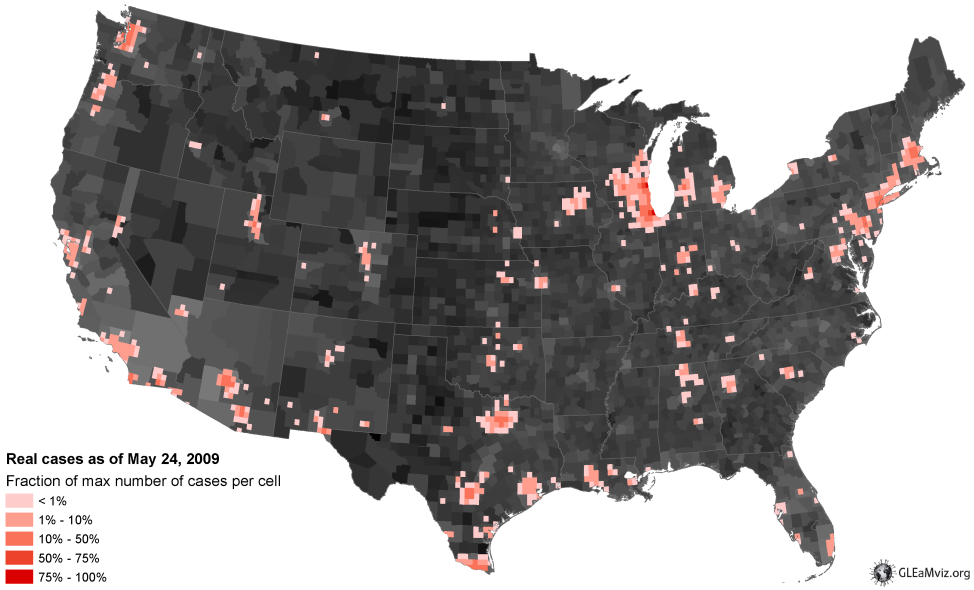


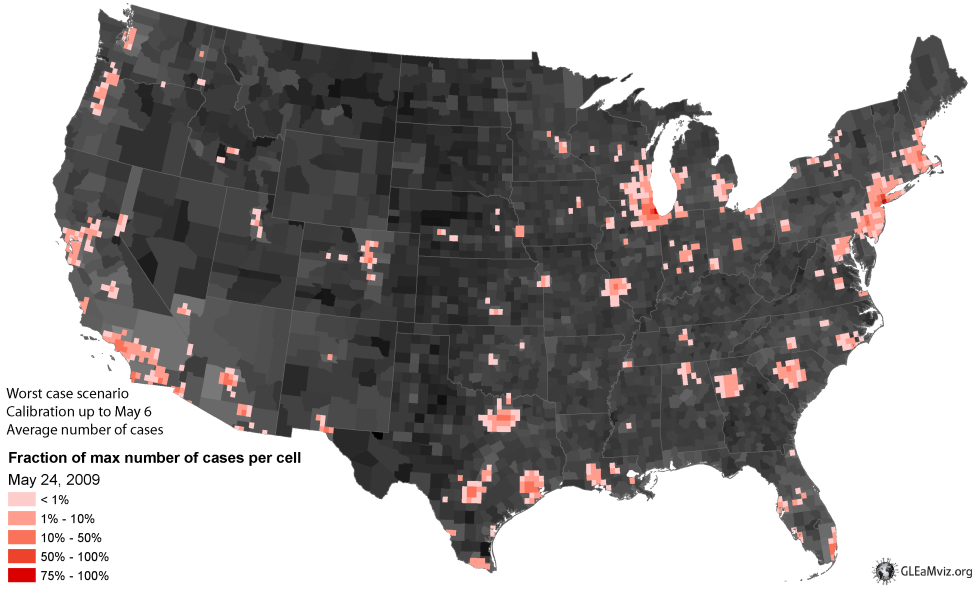
Calibrations (as published on GLEaMviz.org)

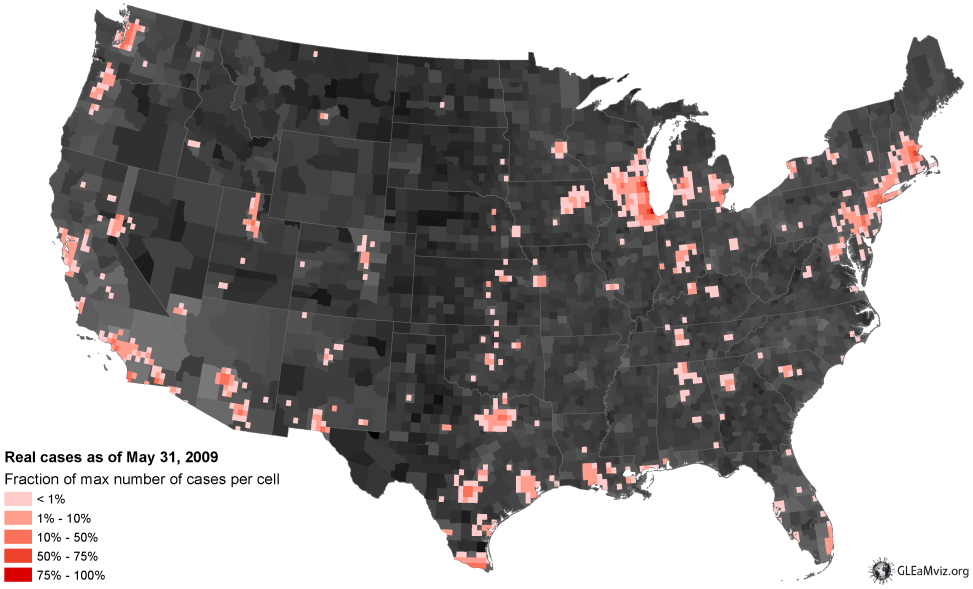
- April the 30th (# of cases)
- May 3rd (# of cases)
- May 6th (# of cases) and (arrival time)
- May 11th (# of cases) and (arrival times)
- Model with seasonality (space exploration enlarged)

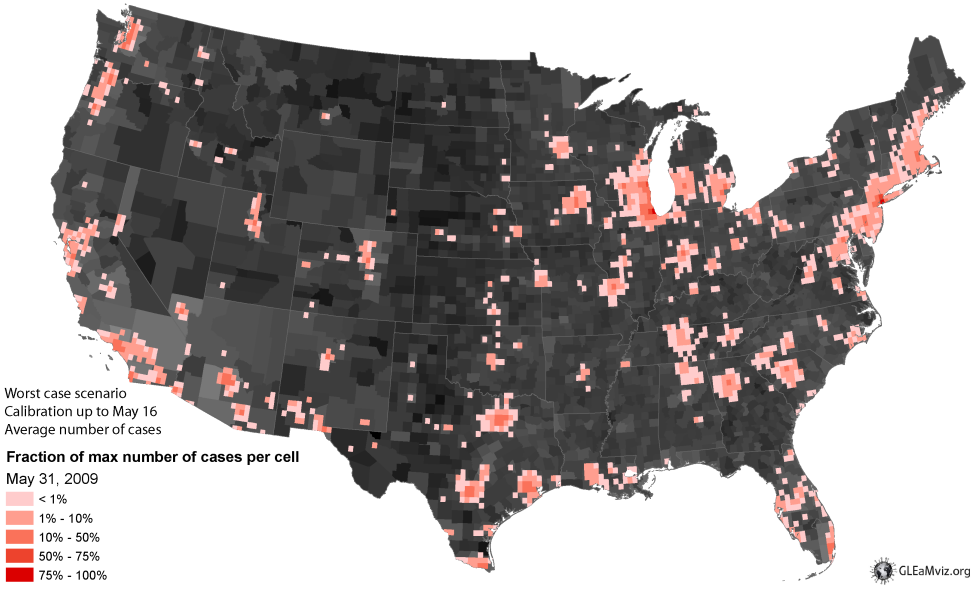
Cumulative number of cases in USA				
Source	**Date**			
	May 10th	**May 17th**	**May 24th**	**May 31st**
Calibration up to April 30th	**567-3,179**	**NA**	**NA**	**NA**
Calibration up to May 3rd	**876-3,101**	**2,603-8,861**	**NA**	**NA**
Calibration up to May 6th	**NA**	**4,092-11,371**	**11,772-31,734**	**NA**
Calibration up to May 11th	**NA**	**9,392-16,244**	**27,462-46,475**	**50,200-110,000**
Seasonality (Apr. 30th)	**2,315-32,226**	**8,025-74,854**	**22,519-143,737**	**50,356-240,684**
Unlimited AV treatment (1)	**2,034-27,737**	**5,845-57,116**	**14,535-98,708**	**30,262-148,865**
Unlimited AV treatment (2)	**1,900-27,844**	**5359-56,616**	**13,483-97,248**	**28,376-146,050**
CDC reports	**2,381**	**4,592**	**7,021**	**~12,000 (x 20??)**

Geographical distribution of cases



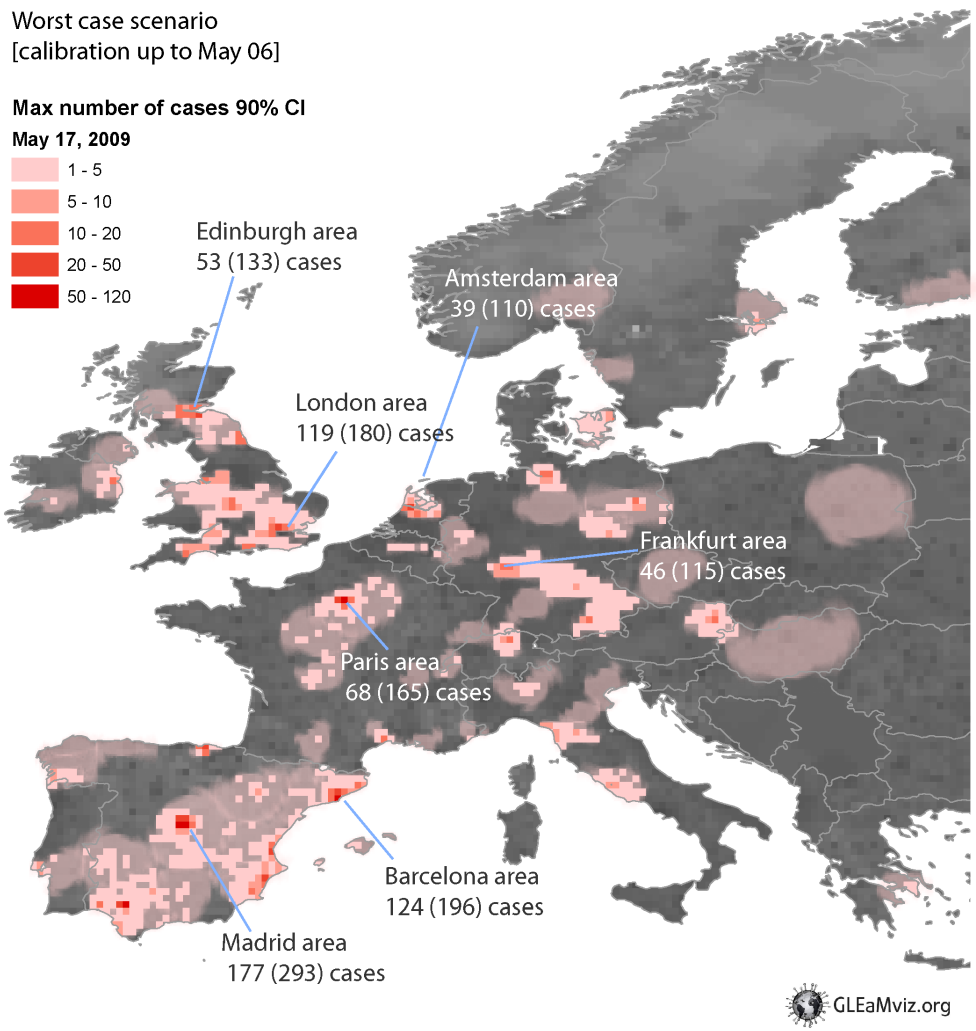




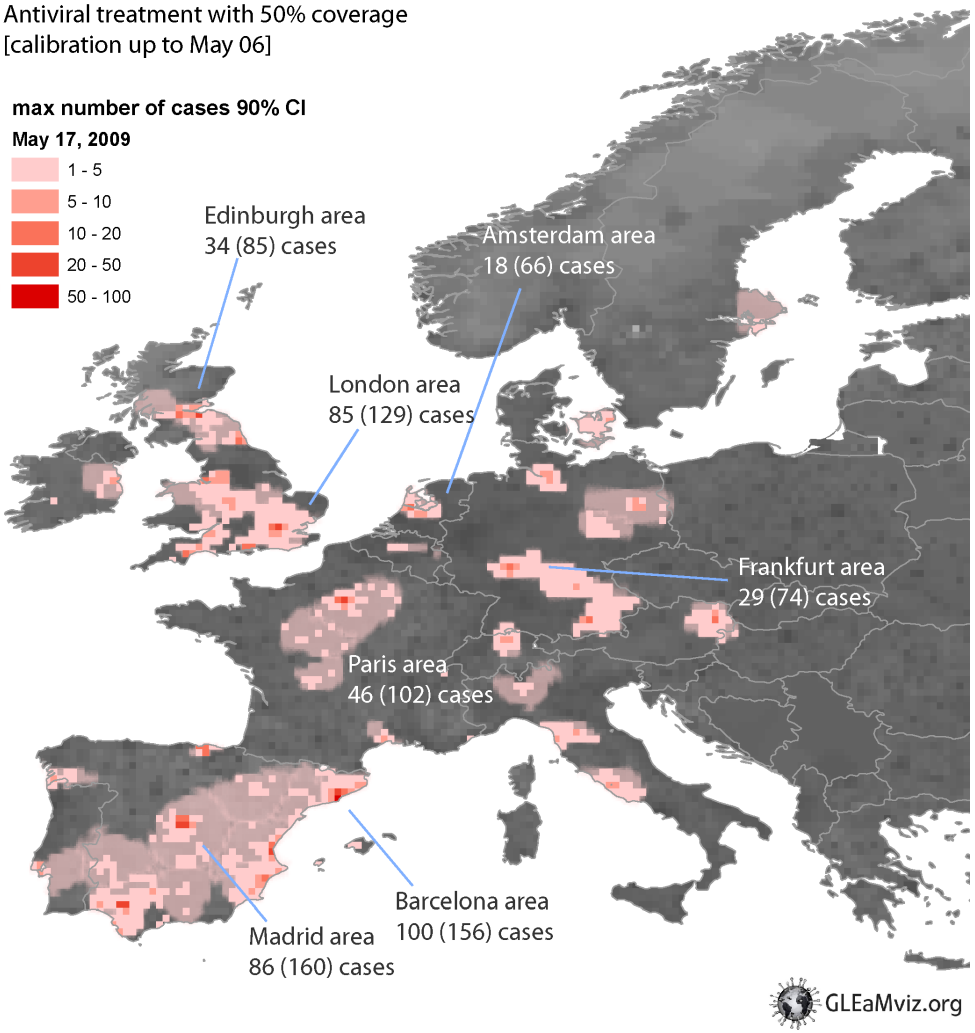


Use of interventions

Worst case scenario
[calibration up to May 06]



Antiviral treatment with 50% coverage
[calibration up to May 06]

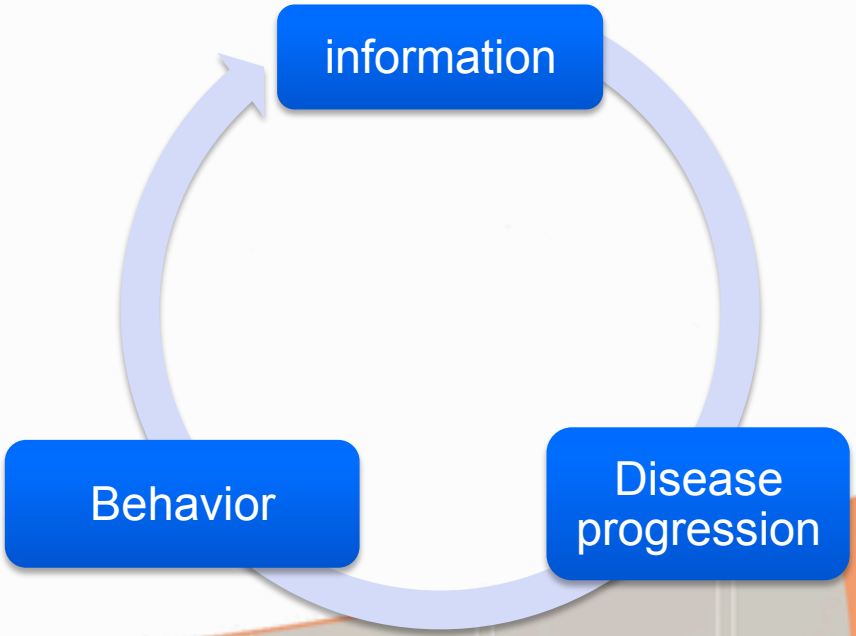


What's next for GEM…

•GemViz: user friendly interface for exploration/experiment + Visualization

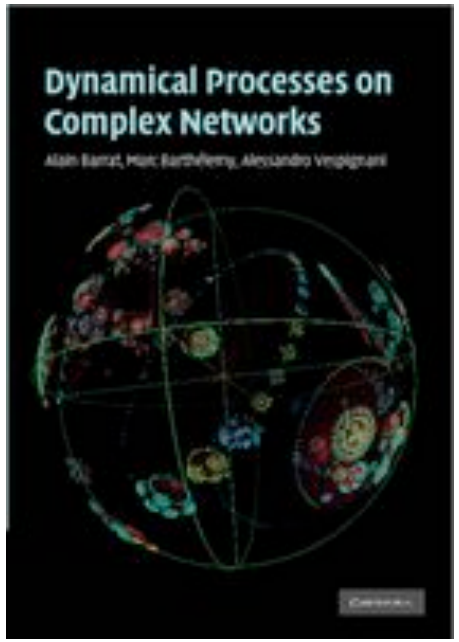
•Age structure (transmissibility/traveling)

•Risk perceptions…….



Recent papers

- Dynamical Processes on complex networks, Cambridge University Press 2008.
- *Theoretical Computer Science* ***355***, 6-24 (2006***)***
- PNAS, **103**, 2015 (2006)
- Bull. Math. Bio. **68**, 1893 (2006)
- Plos *Medicine*, **4**, e13 (2007)
- Nature Physics, **3,** 276-282 (2007)
- Physical Review Letters **99**, 148701 (2007).
- BMC-*Medicine* 5, e34 (2007)
- Journal of Theoretical Biology 251, 450-467 (2008)



More Information/paper/data

Cx-Nets

http://cxnets.googlepages.com