

Tracing epidemics with agent-based and network-based models (EpiAgeNet)

<http://www.prokopenko.net/epiagenet.html>

Prof. Mikhail Prokopenko
Centre for Complex Systems
Faculty of Engineering



THE UNIVERSITY OF
SYDNEY

EpiAgeNet Tutorial, ALIFE 2020
13 July 2020



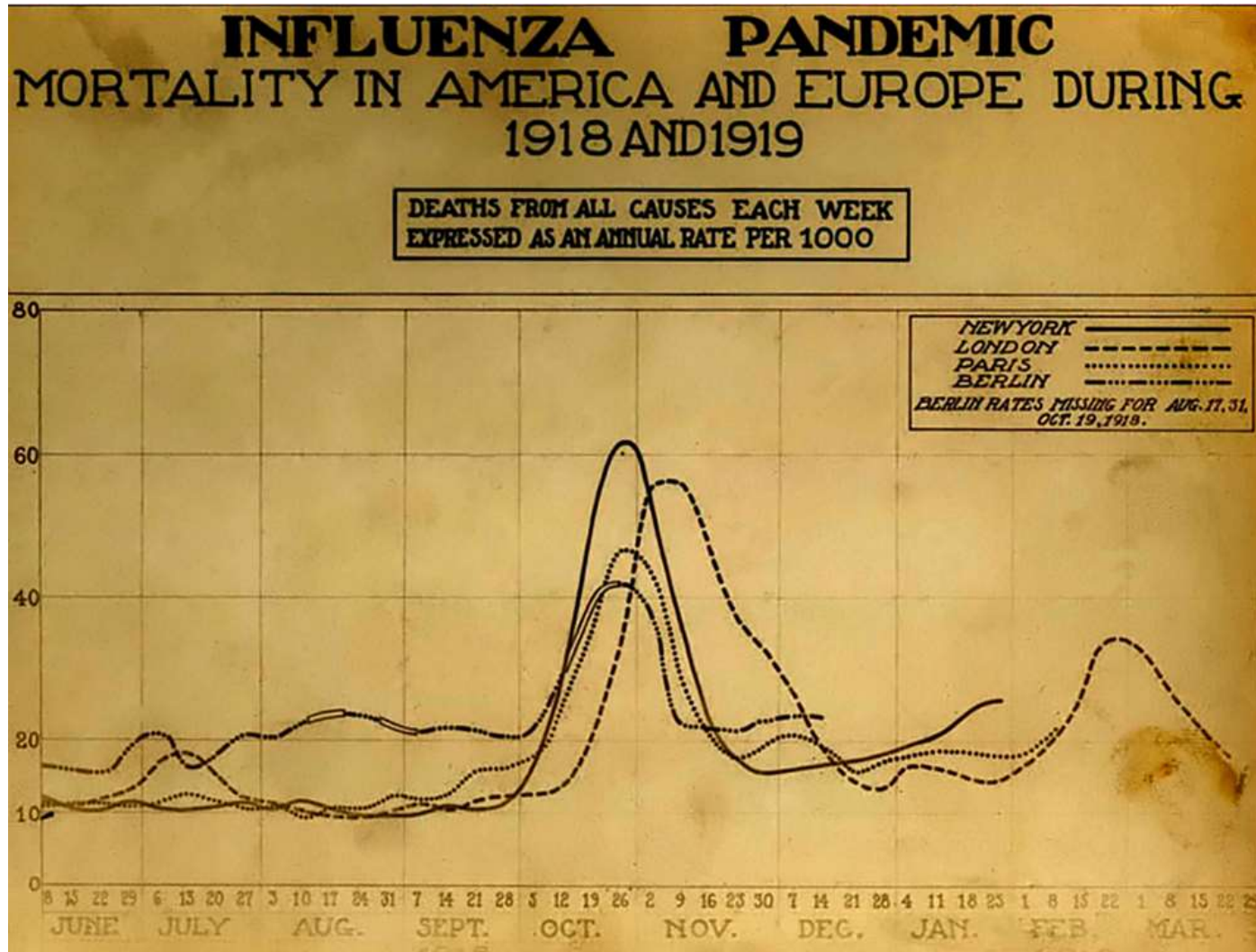
Spanish Flu 1918: 500 million infected, with deaths of three to five percent of the world's population



Soldiers from Fort Riley, Kansas, ill with Spanish influenza at a hospital ward at Camp Funston
Otis Historical Archives Nat'l Museum of Health & Medicine - NCP 1603



Spanish Flu 1918: a chart of deaths in major cities



Pandemic Influenza: The Inside Story. Nicholls H, *PLoS Biology* Vol. 4/2/2006, e50
courtesy of the National Museum of Health and Medicine

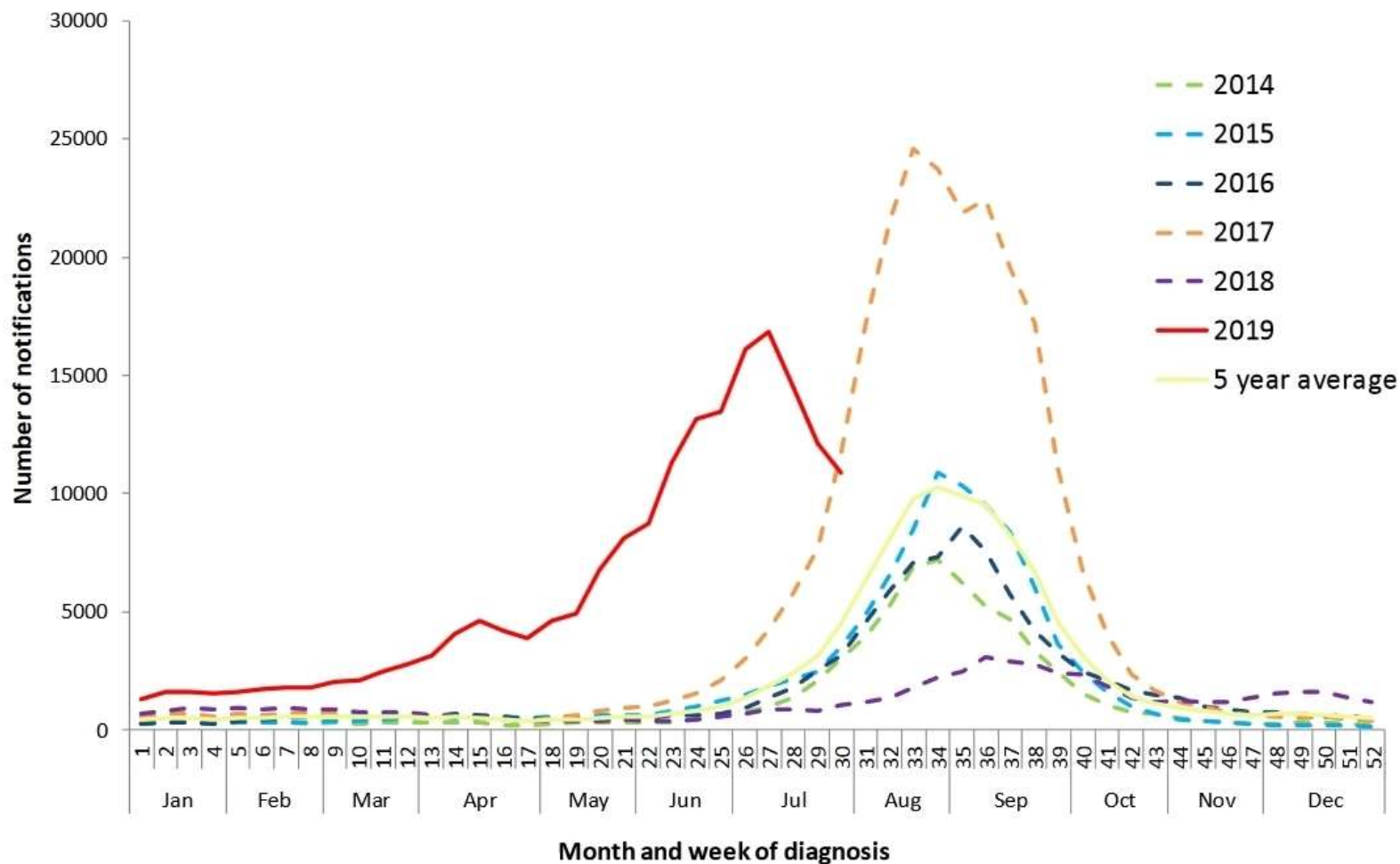
“I had hoped that hitting the 100th anniversary of this epidemic (Spanish flu) would spark a lot of discussion about whether we’re ready for the next global epidemic. Unfortunately, it didn't, and we still are not ready”

Bill Gates
Chair of Bill & Melinda Gates Foundation
2018



Australian Influenza Surveillance Report No 07 : 15 to 28 July 2019

Figure 5. Notifications of laboratory confirmed influenza, Australia, 1 January 2013 to 28 July 2019, by month and week of diagnosis.*



A Contribution to the Mathematical Theory of Epidemics.

By W. O. KERMACK and A. G. MCKENDRICK.

(Communicated by Sir Gilbert Walker, F.R.S.—Received May 13, 1927.)

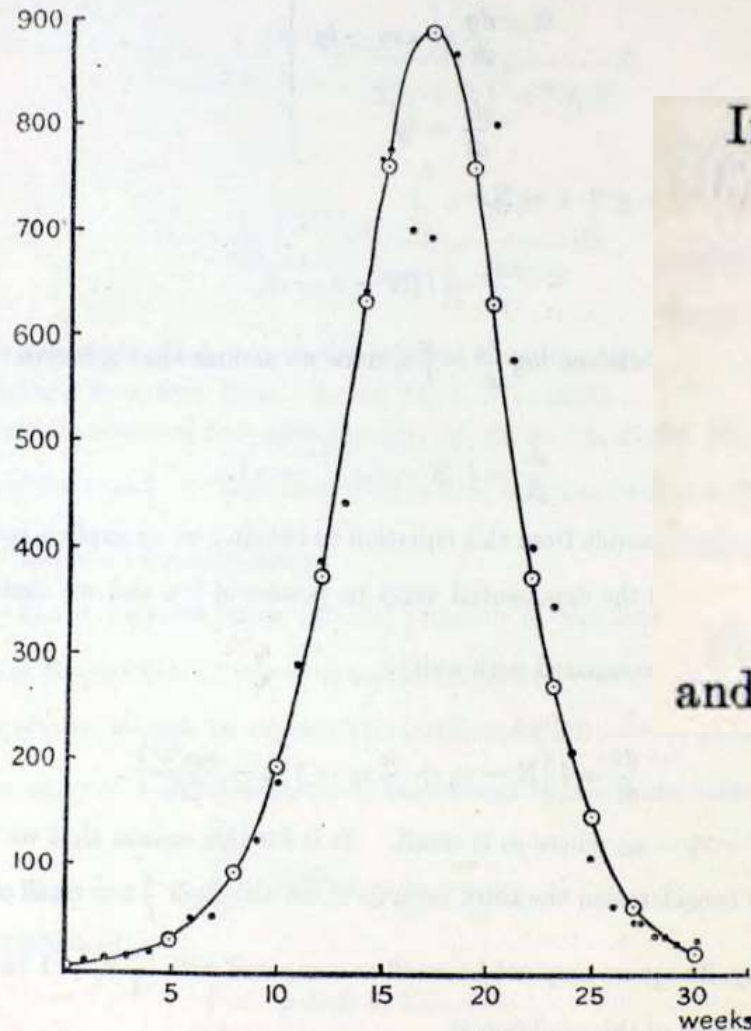
(From the Laboratory of the Royal College of Physicians, Edinburgh.)

Introduction.

(1) One of the most striking features in the study of epidemics is the difficulty of finding a causal factor which appears to be adequate to account for the magnitude of the frequent epidemics of disease which visit almost every population. It was with a view to obtaining more insight regarding the effects of the various factors which govern the spread of contagious epidemics that the present investigation was undertaken. Reference may here be made to the work of Ross



Compartmental models in epidemiology: Susceptible – Infectious – Recovered



The accompanying chart is based upon figures of deaths from plague in the island of Bombay over the period December 17, 1905, to July 21, 1906. The ordinate represents the number of deaths per week, and the abscissa denotes the time in weeks. As at least

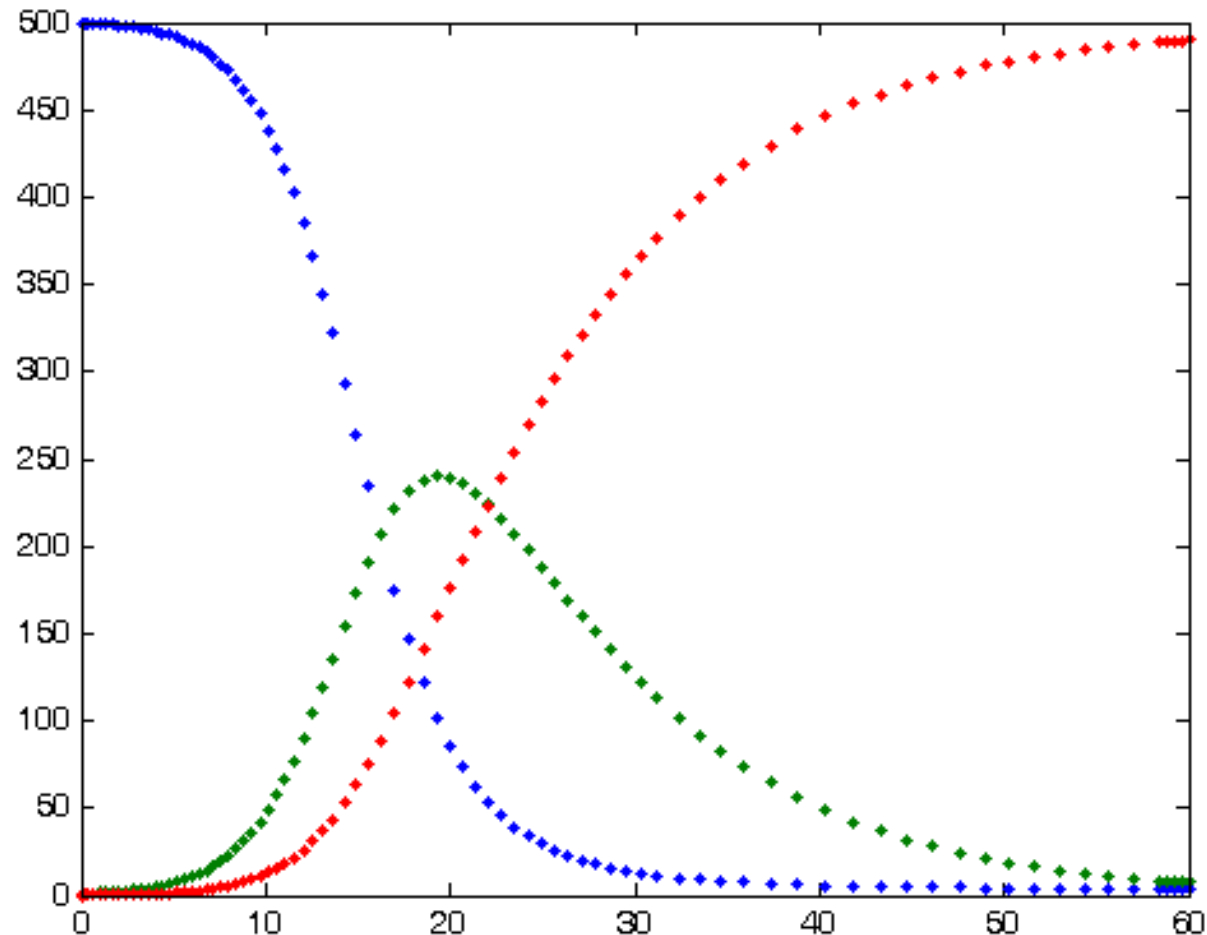
In this case the equations are

$$\left. \begin{aligned} \frac{dx}{dt} &= -\kappa xy \\ \frac{dy}{dt} &= \kappa xy - ly \\ \frac{dz}{dt} &= ly \end{aligned} \right\}$$

and as before $x + y + z = N$.

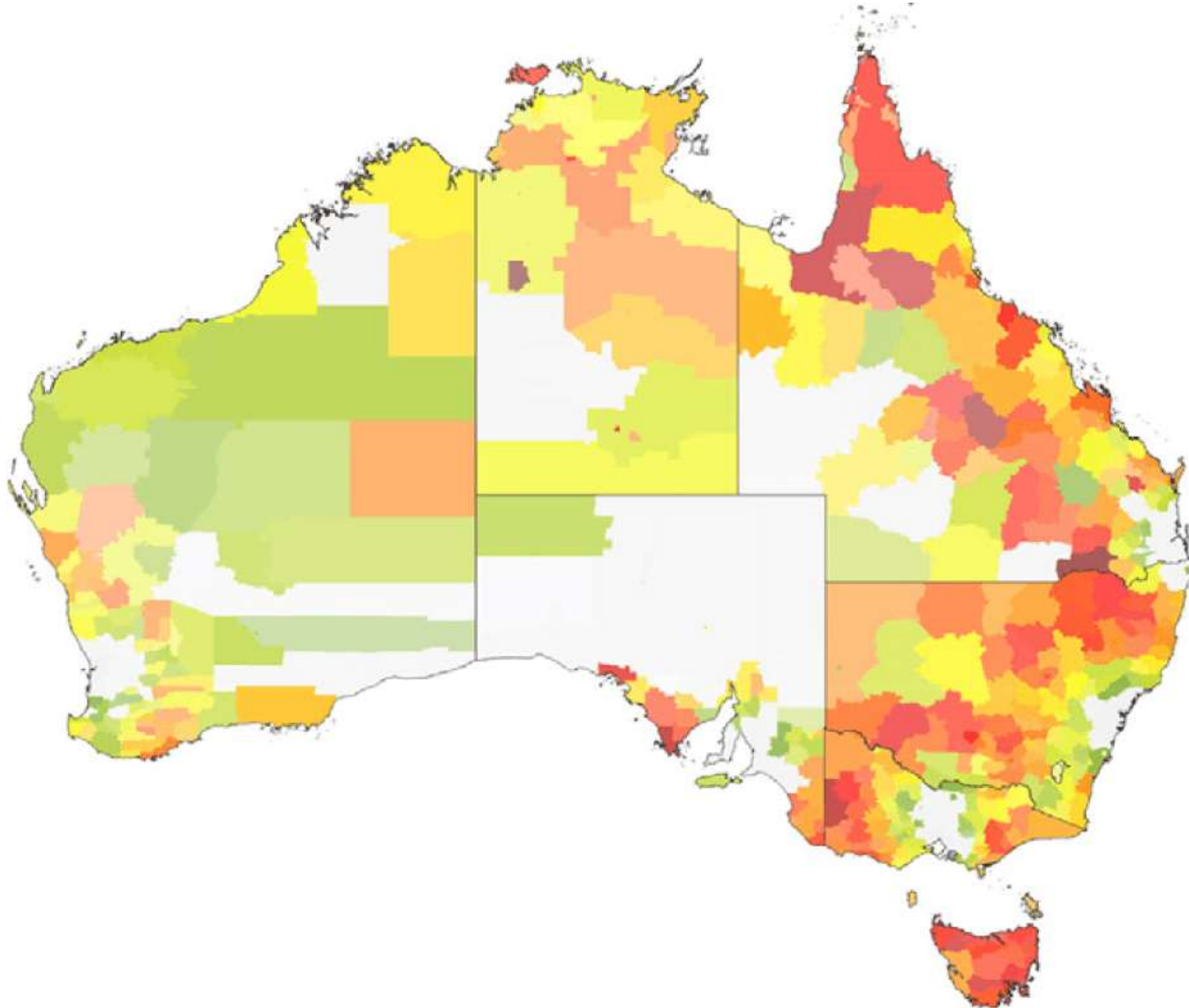


Compartmental models in epidemiology: Susceptible – Infectious – Recovered



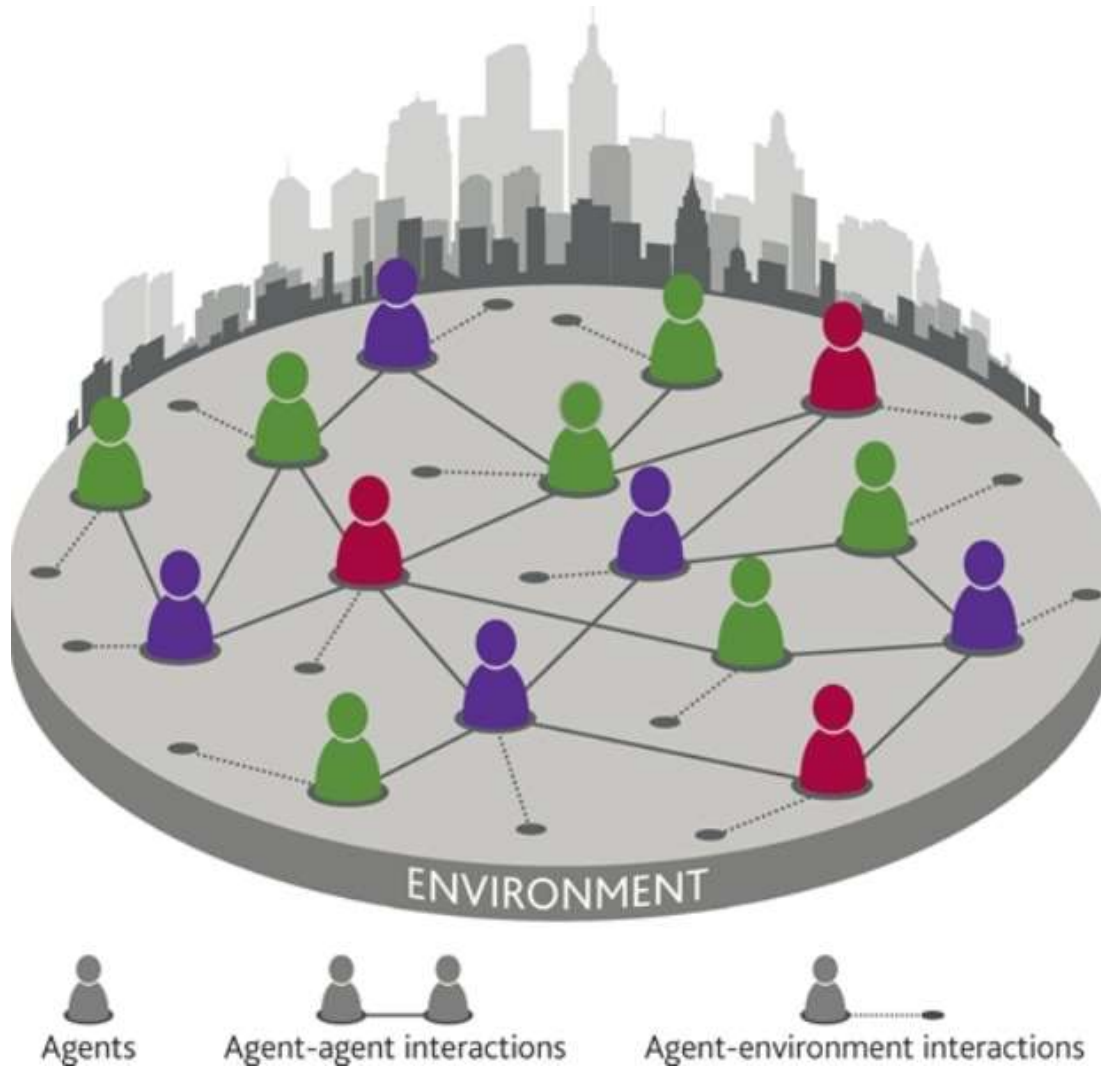


Different questions: How to “zoom in” ? Where to intervene?





Agent-based modelling (ABM)



Agent-based Modelling (ABM) of epidemics

- anonymous individuals (census based) → agents with attributes (e.g., age, gender, occupation, susceptibility and immunity to diseases)
- agent interactions: contacts and disease transmission over about 24M agents, grouped in social “contexts” (households, neighbourhoods, communities, workplaces, schools, classrooms, etc.)
- specific virus (transmission rates, natural history of the disease)
- outbreak modelling of pandemic scenarios (international air traffic)
- varying sources and intensity of infection, as well as population sets
- calibration to known data on reproductive ratio R_0 , attack rates (across “contexts”), growth rates, generation period, other parameters

International Journal of Modern Physics C
Vol. 15, No. 1 (2004) 193–201
© World Scientific Publishing Company

 **World Scientific**
www.worldscientific.com

LARGE-SCALE MOLECULAR-DYNAMICS SIMULATION OF 19 BILLION PARTICLES

KAI KADAU

*Theoretical Division, Los Alamos National Laboratory
MS B262, Los Alamos, New Mexico 87545, USA
kkadau@lanl.gov*

TIMOTHY C. GERMANN

*Applied Physics Division, Los Alamos National Laboratory
MS F699, Los Alamos, New Mexico 87545, USA
tcg@lanl.gov*

PETER S. LOMDAHL

*Theoretical Division, Los Alamos National Laboratory
MS B262, Los Alamos, New Mexico 87545, USA
pxl@lanl.gov*

Received 8 August 2003
Revised 10 August 2003



SPaSM (Scalable Parallel Short-range Molecular dynamics)

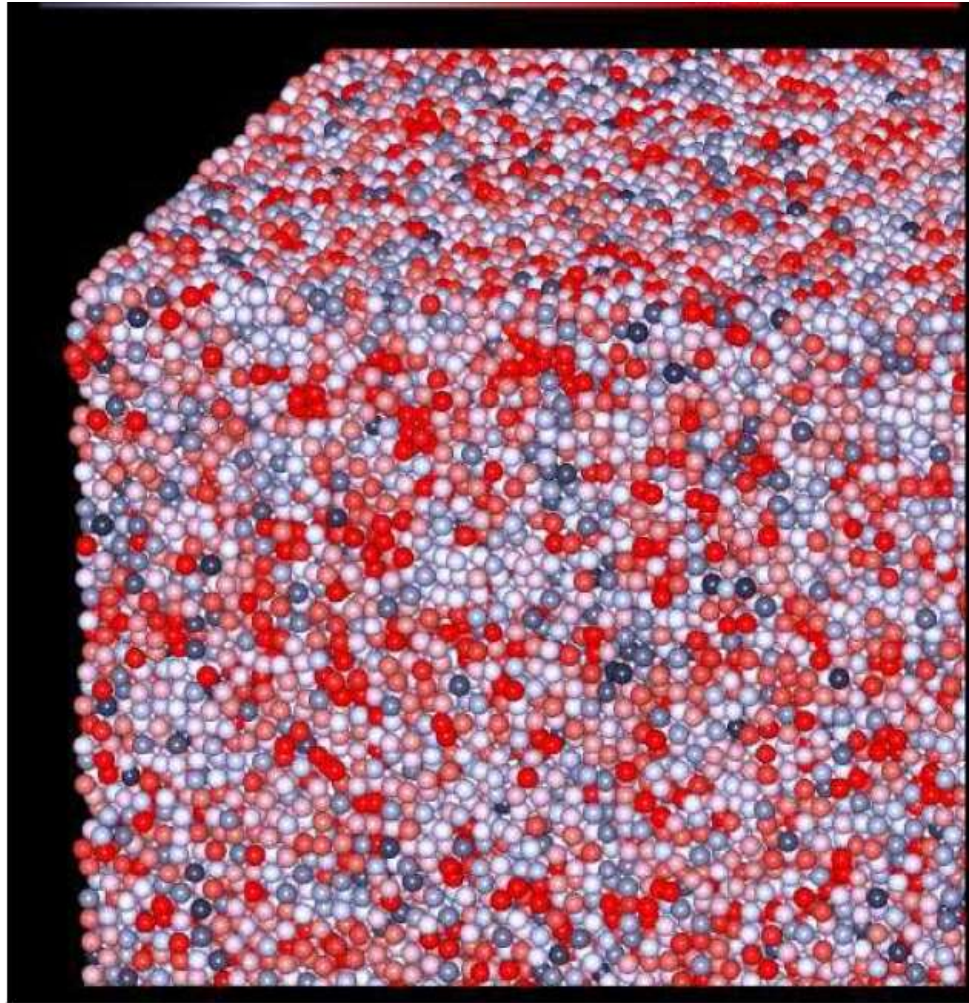


Fig. 3. ≈ 37 million particles rendered on four PN with a resolution of 5000 pixel by 5000 pixel (top). The bottom shows a close-up of the same picture file. The grayscale represents the potential energies of the atoms from -6 to -2 (grayscale version of the original color picture).

Mitigation strategies for pandemic influenza in the United States

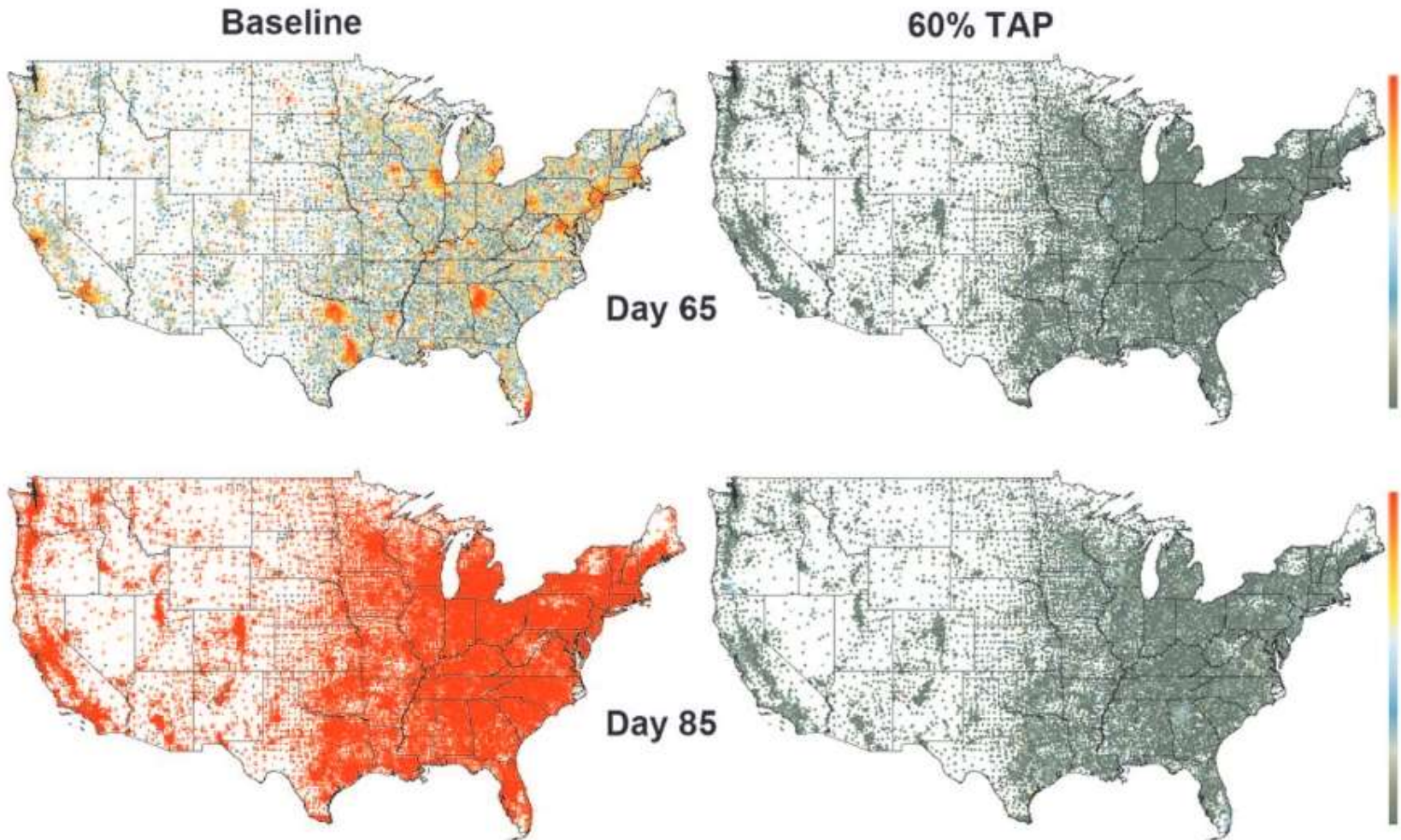
Timothy C. Germann^{*†}, Kai Kadau^{*}, Ira M. Longini, Jr.[‡], and Catherine A. Macken^{*}

^{*}Los Alamos National Laboratory, Los Alamos, NM 87545; and [†]Program of Biostatistics and Biomathematics, Fred Hutchinson Cancer Research Center and Department of Biostatistics, School of Public Health and Community Medicine, University of Washington, Seattle, WA 98109

Communicated by G. Balakrish Nair, International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Bangladesh, February 16, 2006
(received for review January 10, 2006)

Recent human deaths due to infection by highly pathogenic (H5N1) avian influenza A virus have raised the specter of a devastating pandemic like that of 1917–1918, should this avian virus evolve to become readily transmissible among humans. We introduce and use a large-scale stochastic simulation model to investigate the spread of a pandemic strain of influenza virus through the U.S. population of 281 million individuals for R_0 (the basic reproductive number) from 1.6 to 2.4. We model the impact that a variety of

resources to minimize the impact of the outbreak? Precise planning is hampered by several unknowns, most critically the eventual human-to-human transmissibility of the human-adapted avian strain (characterized by the basic reproductive number R_0 , the average number of secondary infections caused by a single typical infected individual among a completely susceptible population), and the supply of therapeutic agents. Manufacturers of neuraminidase inhibitors, such as oseltamivir,



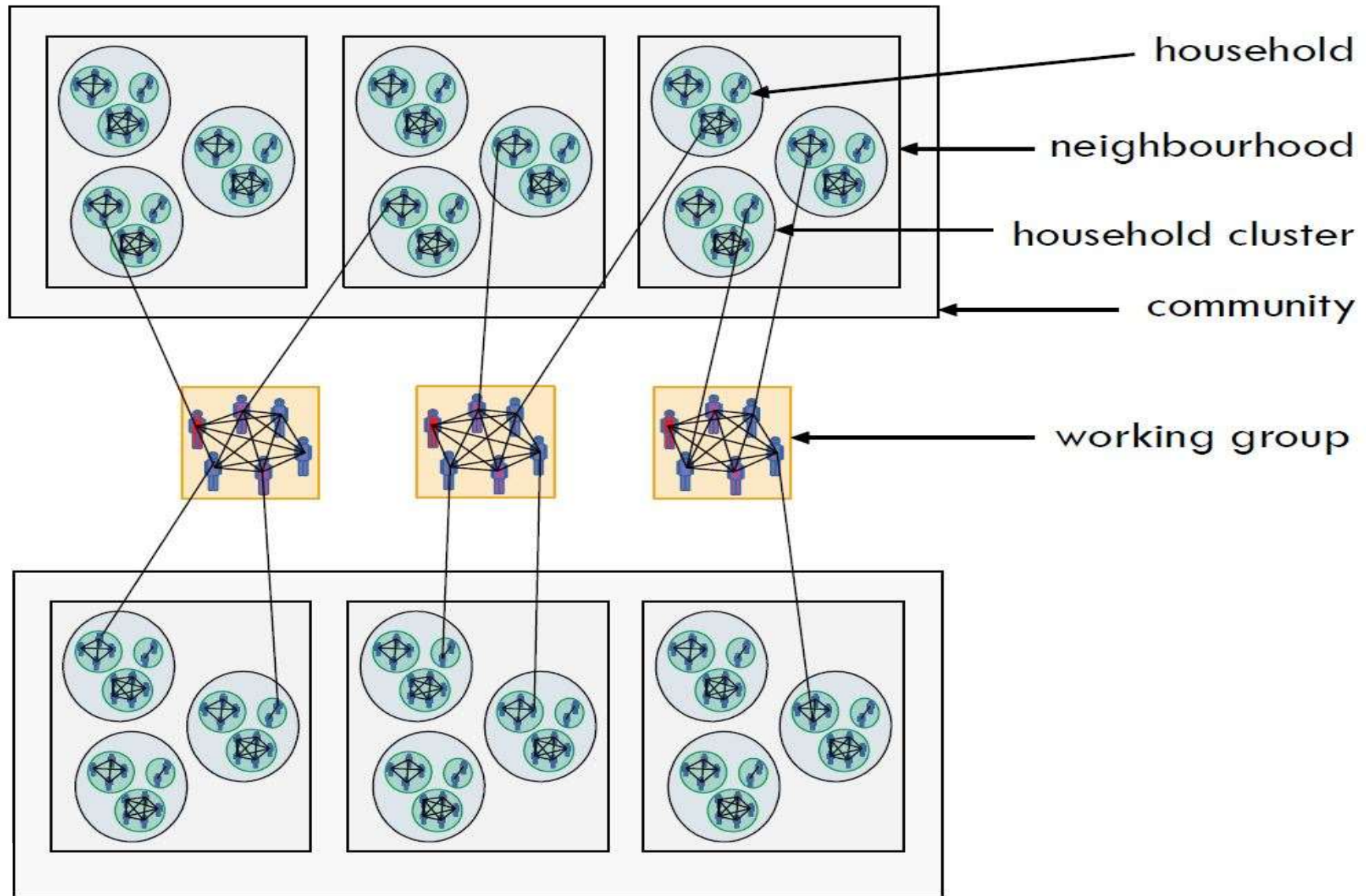
Our pandemic modelling research (since 2016)

- Modelling pandemics with large-scale high-resolution agent-based models
 - *demographics*: from census based data to agents
 - *mobility*: travel patterns including long-distance
 - *infection*: disease transmission and natural history models
 - ACEMod – Australian Census-based Epidemic Model
 - AMTraC-19 – Agent-based Model of Transmission and Control of the COVID-19 pandemic in Australia

- Influenza pandemics (H1N1):
 - pandemic trends (peaks, synchrony, bimodality, critical regimes)
 - effects of urbanisation
 - counter-factual analysis
 - efficiency of interventions: geographically-targeted anti-prophylaxis (GTAP), contact-targeted anti-prophylaxis (TAP), vaccination



“Same storm, different boats”: ABM mixing contexts



Population partitions: residential areas and destination zones

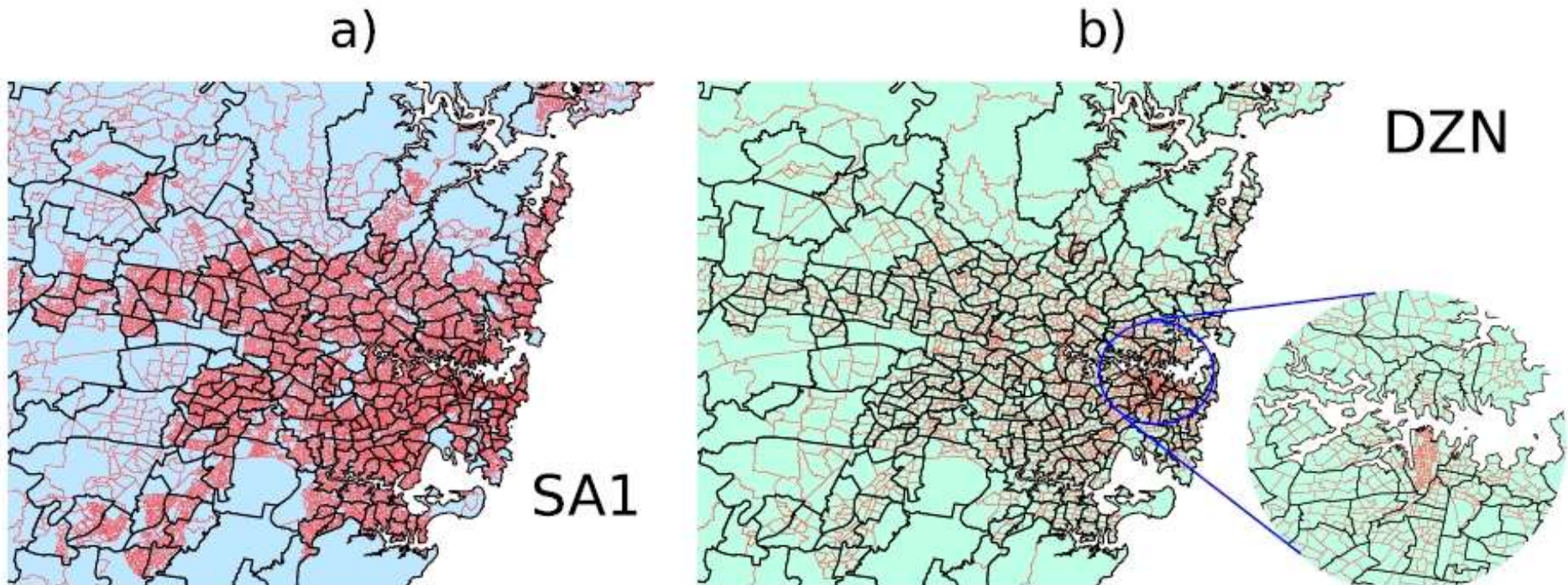


Fig. 1 Maps of the Greater Sydney region illustrating the distribution of population partitions. (a) A map of the Greater Sydney region showing SA2 (black) and SA1 (red) population partitions. (b) A map of the same area showing SA2 (black) and DZN (red) partitions. The inset in (b) zooms in on the Sydney central business district to illustrate the much denser packing of DZN partitions in that area.

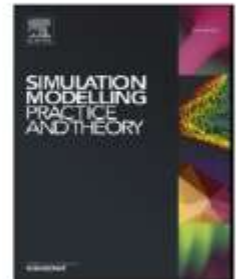
Simulation Modelling Practice and Theory 87 (2018) 412–431



Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat



Investigating spatiotemporal dynamics and synchrony of influenza epidemics in Australia: An agent-based modelling approach



Oliver M. Cliff^{a,c}, Nathan Harding^a, Mahendra Piraveenan^a, E. Yagmur Erten^{a,b},
Manoj Gambhir^c, Mikhail Prokopenko^{a,d}

^a Centre for Complex Systems, Faculty of Engineering and IT, University of Sydney, Sydney, NSW 2006, Australia

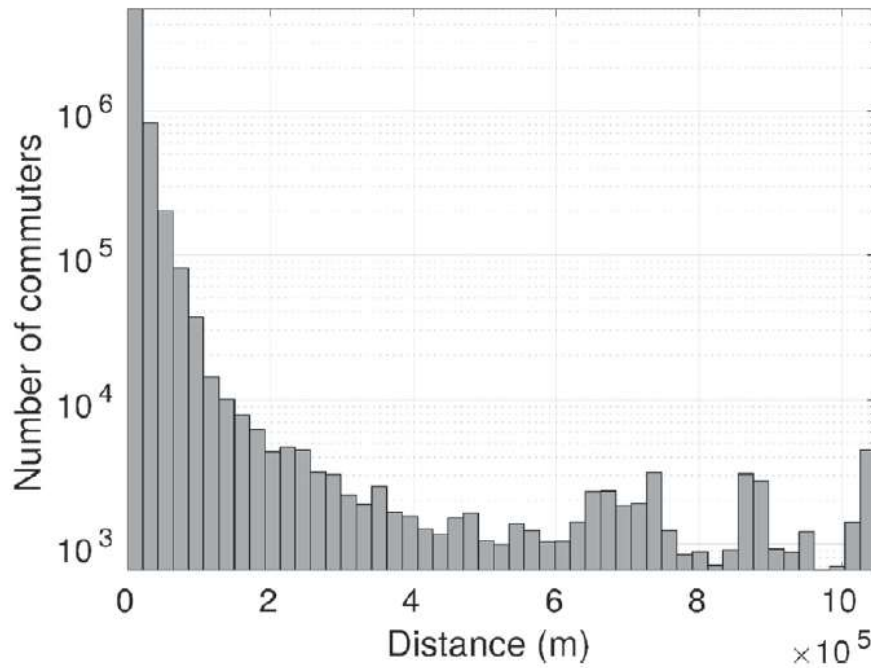
^b Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, Zurich, 8057, Switzerland

^c IBM Research, Melbourne, Australia

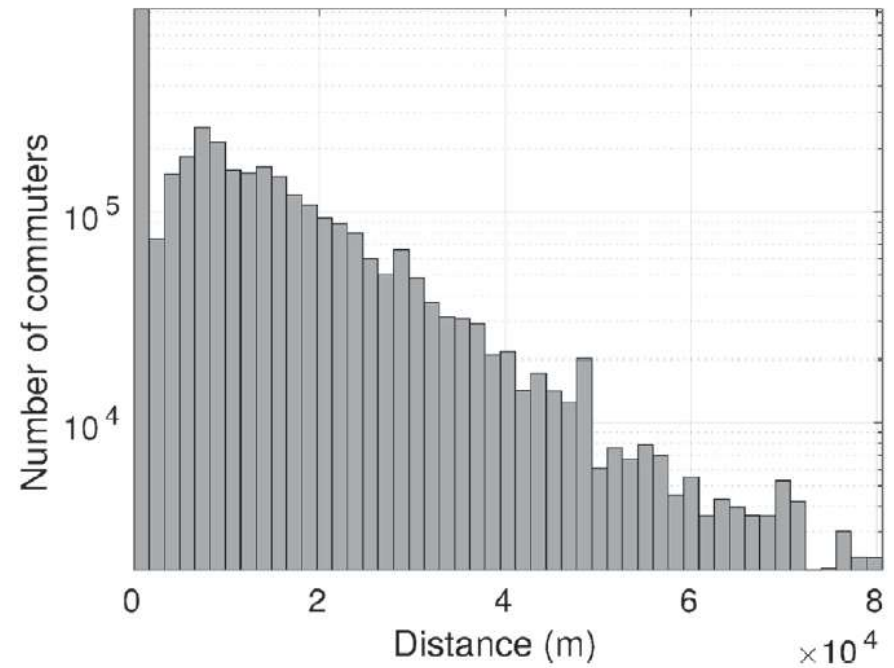
^d Marie Bashir Institute for Infectious Diseases and Biosecurity, University of Sydney, Westmead, NSW 2145, Australia



Australian Census: travel-to-work data (mobility)



(a) Workers



(b) Students

Fig. B1. Commute distance distributions.

Airport code	State	City	Passengers
SYD	NSW	Sydney	40884
MEL	VIC	Melbourne	25859
BNE	QLD	Brisbane	14250
PER	WA	Perth	11449
OOL	QLD	Gold Coast	3022
ADL	SA	Adelaide	2214
CNS	QLD	Cairns	1874
DRW	NT	Darwin	597
TSV	QLD	Townsville	105

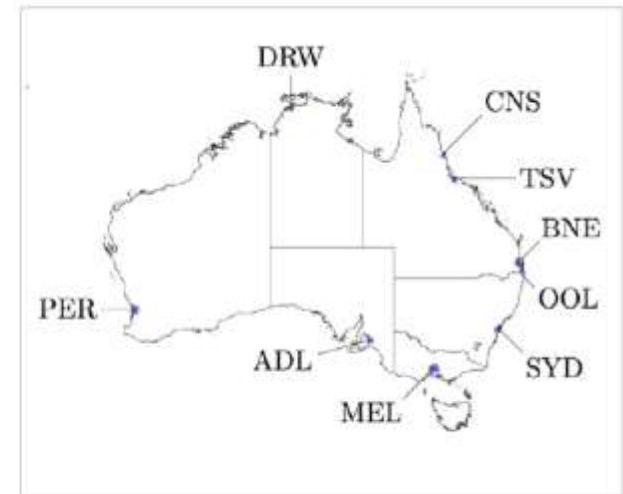
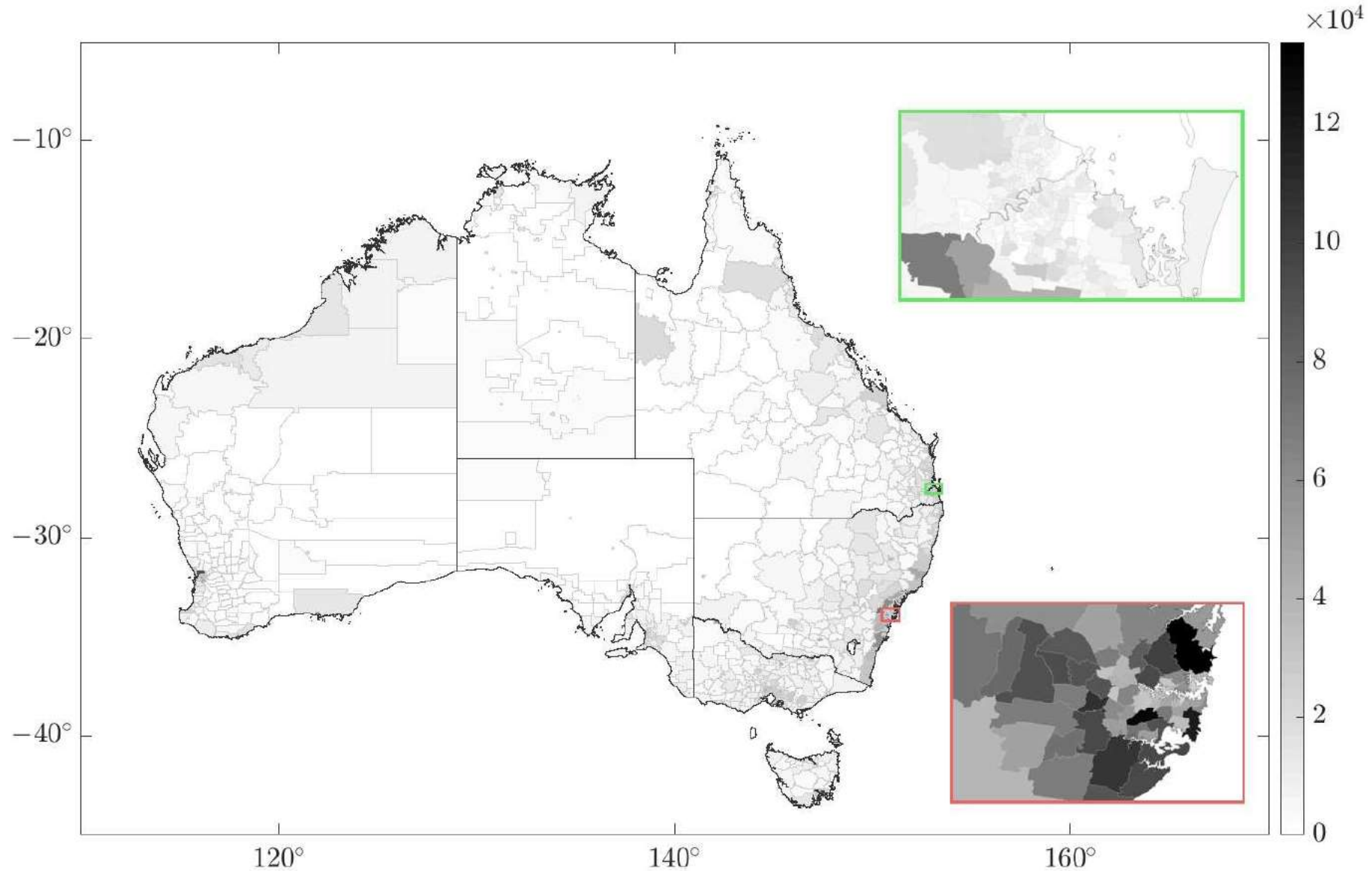


Fig. 3. Daily incoming passengers per Australian international airport obtained from BITRE [30] along with a map detailing the airport locations.

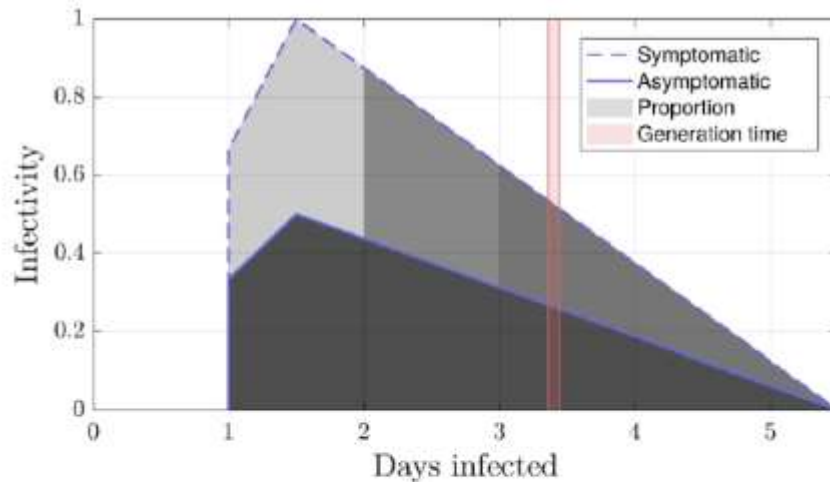


THE UNIVERSITY OF
SYDNEY

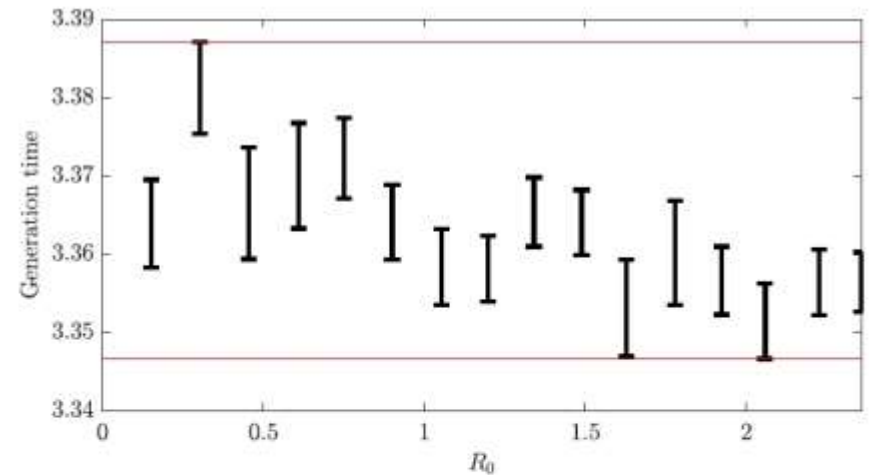
Australian Census based Epidemic Modelling: ACEMod



Epidemic modelling: natural history of the disease



(a) Natural history of the disease.



(b) Simulated generation time.

Fig. 2. Natural history of the disease and corresponding simulated generation time. The disease dynamics are modelled as having a linear increase followed by a linear decrease, as illustrated in Fig. 2 (a). In the figure, the area under the curve is shaded according to the proportion of people at least that infectious after disease onset (darker representing a higher proportion). If an agent becomes symptomatic, their infectiousness doubles (dashed blue line) from that day onward. Moreover, 67% of agents become symptomatic; of these agents, 30% start showing symptoms on day 1, 50% on day 2, and the remaining 20% on day 3. We obtain empirical generation times from simulations resulting from this model, shown in 2(b) for a number of R_0 values. The confidence intervals range from 3.35 to 3.39 days (also shown on Fig. 2(a)), depending on R_0 and, in general, the generation time has a slight downward trend as a function of disease severity.

Epidemic modelling: natural history of the disease

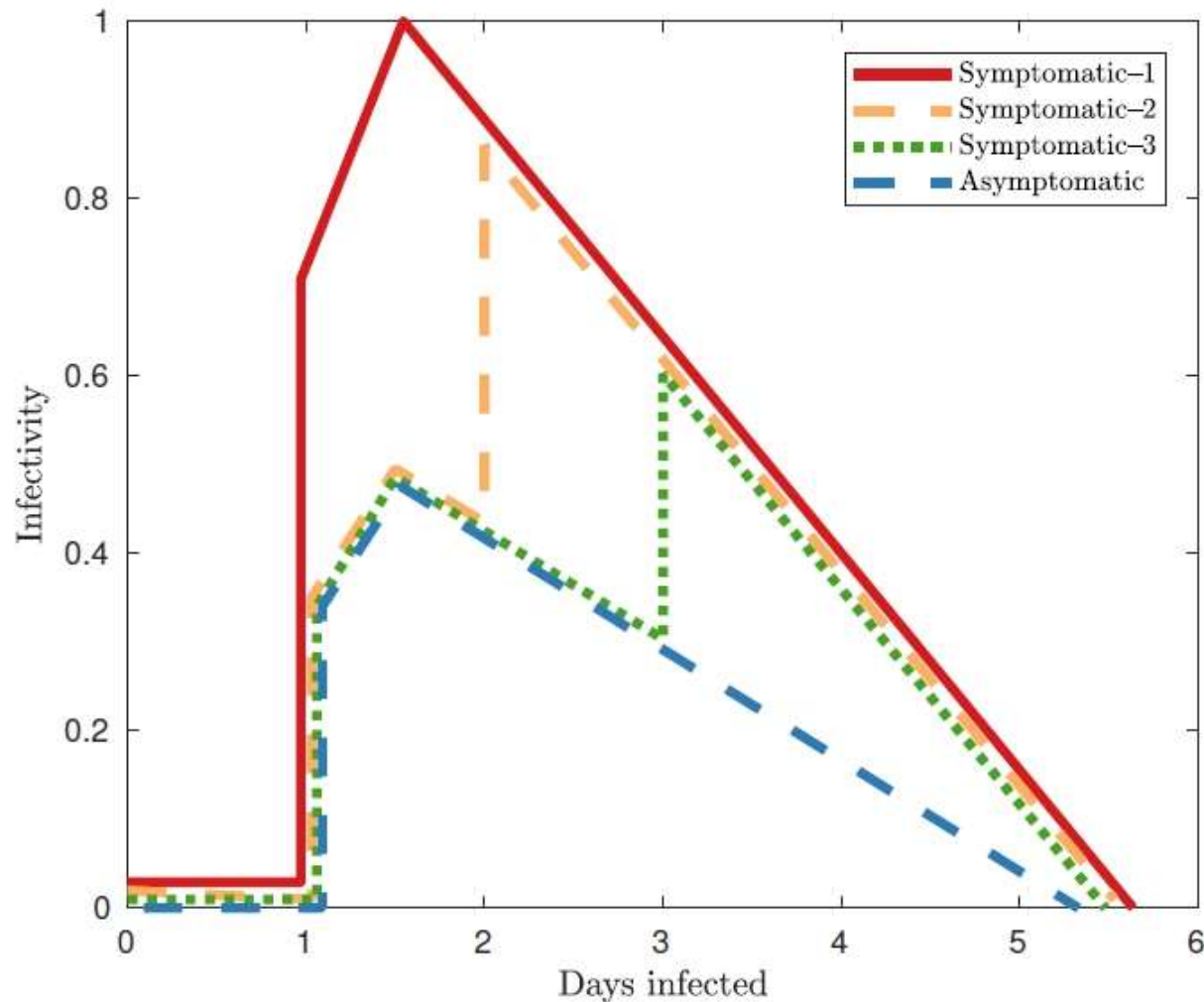


Table C2

Daily transmission probabilities $q_{j \rightarrow i}^g$ for different contact groups g , obtained by Eq. (4) where $\beta_{j \rightarrow i}^g$ are reported by [10].

Contact Group g	Infected Individual j	Susceptible Individual i	Transmission Probability $q_{j \rightarrow i}^g$
Household size 2	Any	Child (<19)	0.0933
	Any	Adult (>18)	0.0393
Household size 3	Any	Child (<19)	0.0586
	Any	Adult (>18)	0.0244
Household size 4	Any	Child (<19)	0.0417
	Any	Adult (>18)	0.0173
Household size 5	Any	Child (<19)	0.0321
	Any	Adult (>18)	0.0133
Household size 6	Any	Child (<19)	0.0259
	Any	Adult (>18)	0.0107
School	Child (<19)	Child (<19)	0.000292
Grade	Child (<19)	Child (<19)	0.00158
Class	Child (<19)	Child (<19)	0.035

$$p_{j \rightarrow i}^g(n) = \kappa f(n - n_j | j) q_{j \rightarrow i}^g$$

$$p_i(n) = 1 - \prod_{g \in G_i(n)} \left[\prod_{j \in \mathcal{A}_g \setminus i} (1 - p_{j \rightarrow i}^g(n)) \right]$$

global scalar

Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza

Simon Cauchemez^{a,1}, Achuyt Bhattarai^b, Tiffany L. Marchbanks^c, Ryan P. Fagan^b, Stephen Ostroff^c, Neil M. Ferguson^a, David Swerdlow^b, and the Pennsylvania H1N1 working group^{b,c,2}

^aMedical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London W2 1PG, United Kingdom; ^bCenters for Disease Control and Prevention, Atlanta, GA 30333; and ^cPennsylvania Department of Health, Harrisburg, PA 17120-0701

Edited by David Cox, Nuffield College, Oxford, United Kingdom, and approved December 22, 2010 (received for review June 22, 2010)

Evaluating the impact of different social networks on the spread of respiratory diseases has been limited by a lack of detailed data on transmission outside the household setting as well as appropriate statistical methods. Here, from data collected during a H1N1 pandemic (pdm) influenza outbreak that started in an elementary school and spread in a semirural community in Pennsylvania, we quantify how transmission of influenza is affected by social networks. We set up a transmission model for which parameters are estimated from the data via Markov chain Monte Carlo sampling. Sitting next to a case or being the playmate of a case did not significantly increase the risk of infection; but the structuring of

sylvania to investigate how social networks and population structures affect influenza transmission.

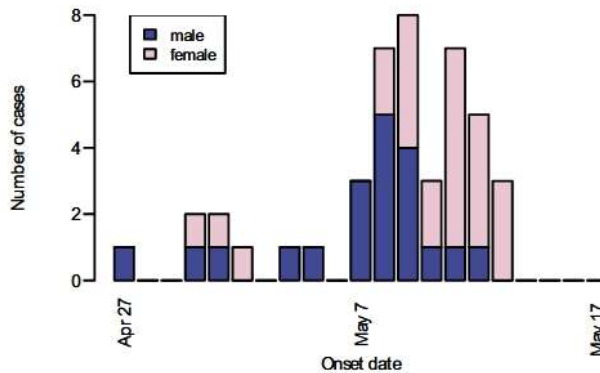
Results and Discussion

Outbreak Investigation. Fig. 1 presents the data that were collected during the outbreak investigation. Demographic and clinical information on 370 (81%) students from 295 (81%) households and their 899 household contacts was collected during two rounds of phone interviews (May 16–21 and May 26–June 2). One hundred twenty-nine (35%) students and 141 (16%) household contacts were reported to have had acute re-



Social interactions (Cauchemez et al., 2010)

C



D

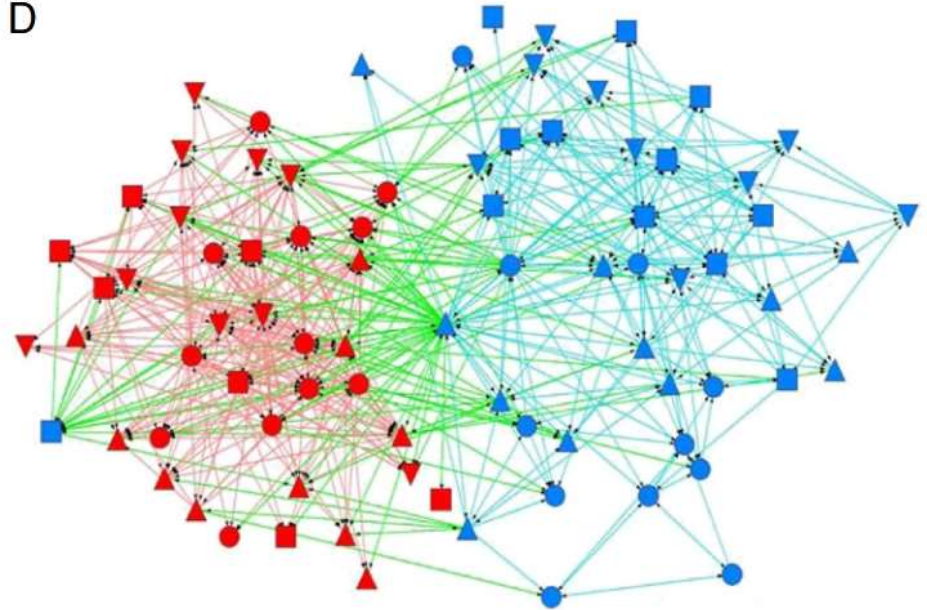


Fig. 1. Epidemiological data collected in the school. (A) Number of acute respiratory illness (ARI) cases by date of symptom onset for different types of individuals. (B–D) Survey of fourth graders with (B) seating charts and diagnosis for ARI in classroom C, (C) number of ARI cases by date of symptom onset and sex among fourth graders, and (D) social networking among fourth graders based on the question “Who are your playmates?” [color of the nodes, red, female; blue, male; color of the lines, red, girl–girl interaction; cyan, boy–boy interaction; green, boy–girl interaction (one symbol shape per class)]. The algorithm used to draw the network aims at (i) distributing nodes evenly, (ii) making edge length uniform, (iii) minimizing edge crossings, and (iv) keeping nodes from coming too close to edges (32, 33) (software: *Netdraw*). It does not use data on sex to position the nodes.

Table C1

Daily contact probabilities $c_{j \rightarrow i}^g$ for different contact groups g , reported by [22].

Mixing group g	Infected individual j	Susceptible individual i	Contact probability $c_{j \rightarrow i}^g$
Household cluster	Child (<19)	Child (<19)	0.08
	Child (<19)	Adult (>18)	0.035
	Adult (>18)	Child (<19)	0.025
	Adult (>18)	Adult	0.04
Working Group	Adult (19-64)	Adult (19-64)	0.05
Neighbourhood	Any	Child (0-4)	0.0000435
	Any	Child (5-18)	0.0001305
	Any	Adult (19-64)	0.000348
	Any	Adult (65+)	0.000696
Community	Any	Child (0-4)	0.0000109
	Any	Child (5-18)	0.0000326
	Any	Adult (19-64)	0.000087
	Any	Adult (65+)	0.000174

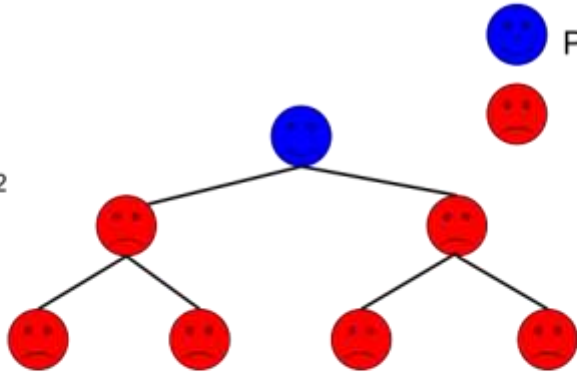


Epidemic modelling: reproductive ratio R_0

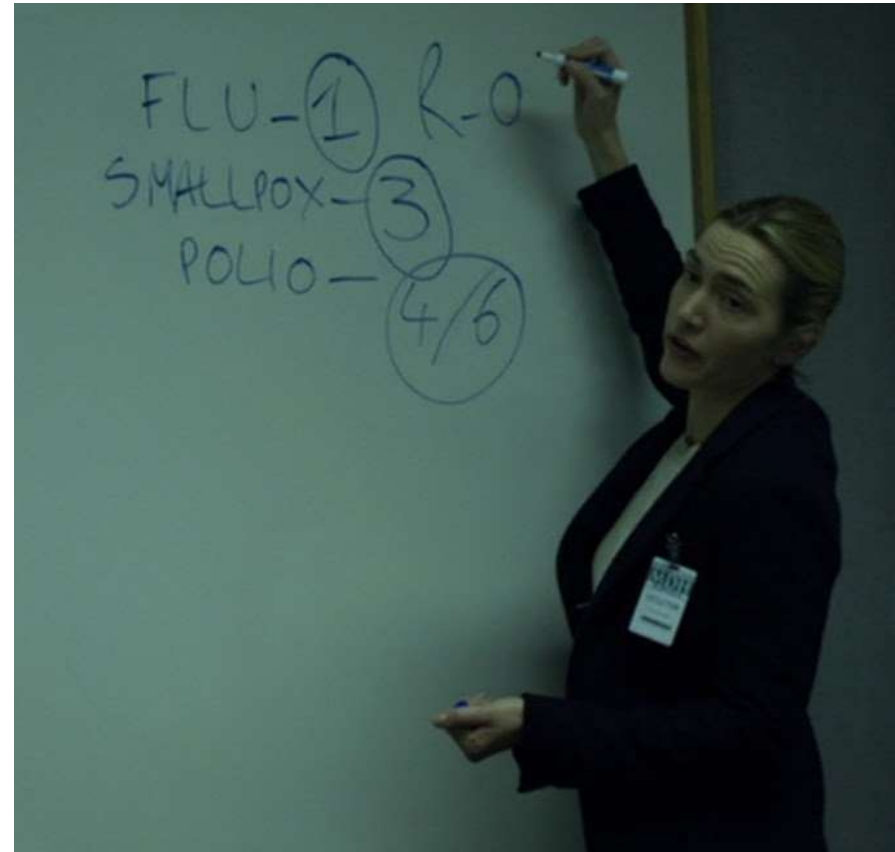
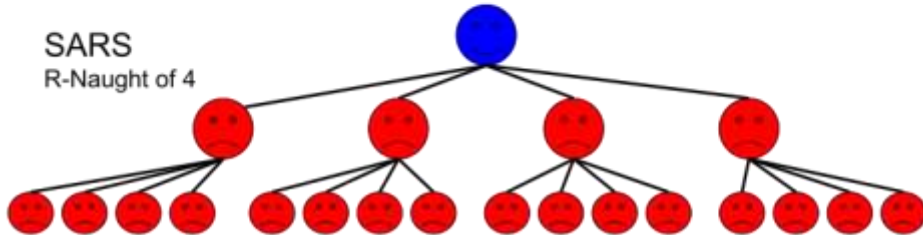
● Patient Zero

● Infected

Ebola:
R-Naught of 2



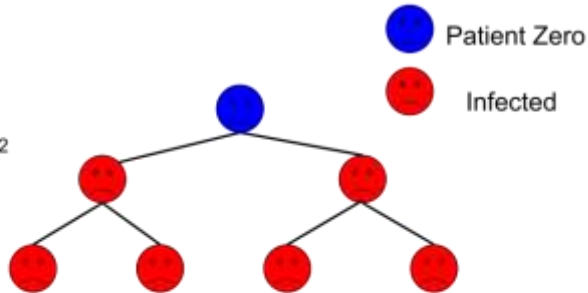
SARS
R-Naught of 4



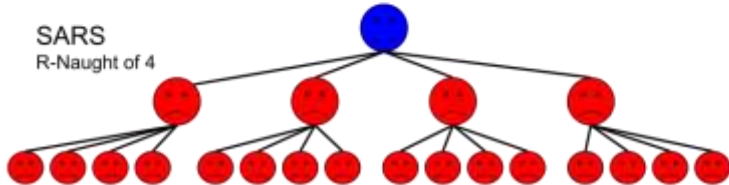


Epidemic modelling: reproductive ratio R_0

Ebola:
R-Naught of 2



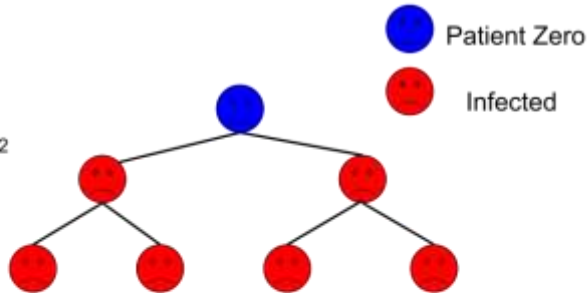
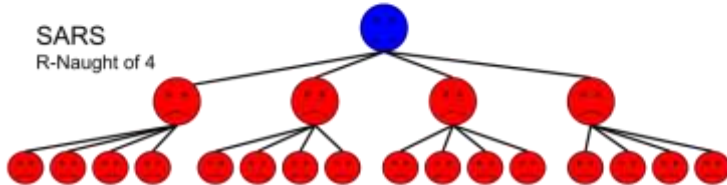
SARS
R-Naught of 4



$$\frac{dS}{dt} = \gamma I - \beta IS$$

$$\frac{dI}{dt} = \beta IS - \gamma I$$

$$\beta / \gamma = R_0$$

Epidemic modelling: reproductive ratio R_0 Ebola:
R-Naught of 2SARS
R-Naught of 4

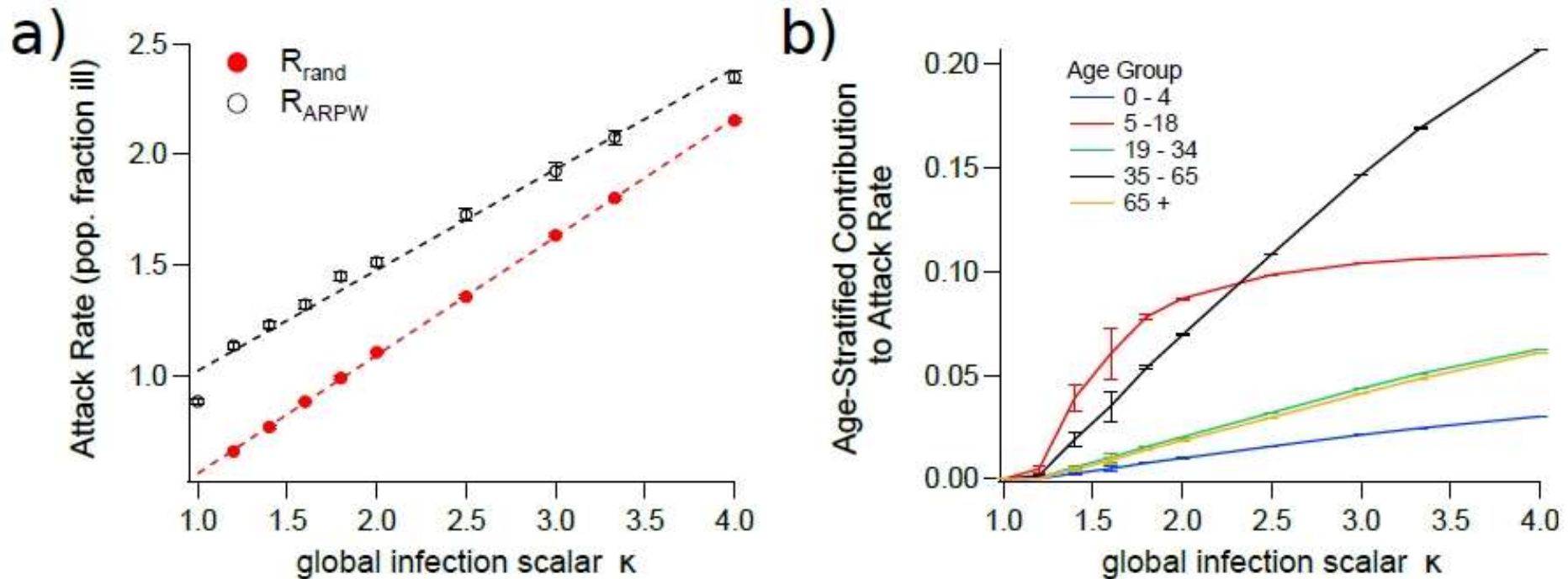
$$\frac{dS}{dt} = \gamma I - \beta IS$$

$$\frac{dI}{dt} = \beta IS - \gamma I$$

$$\beta / \gamma = R_0$$

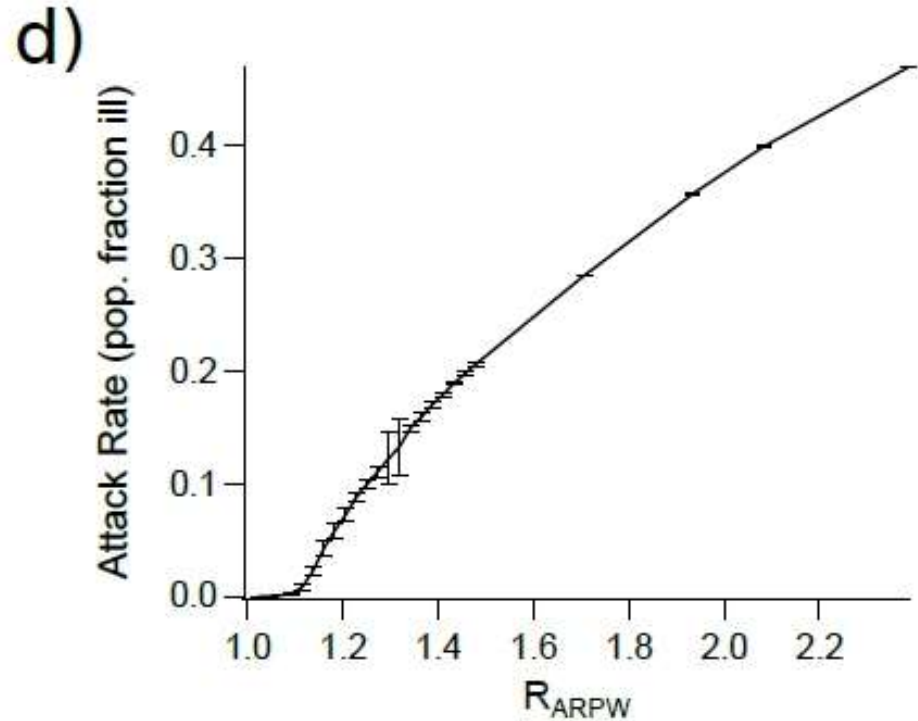
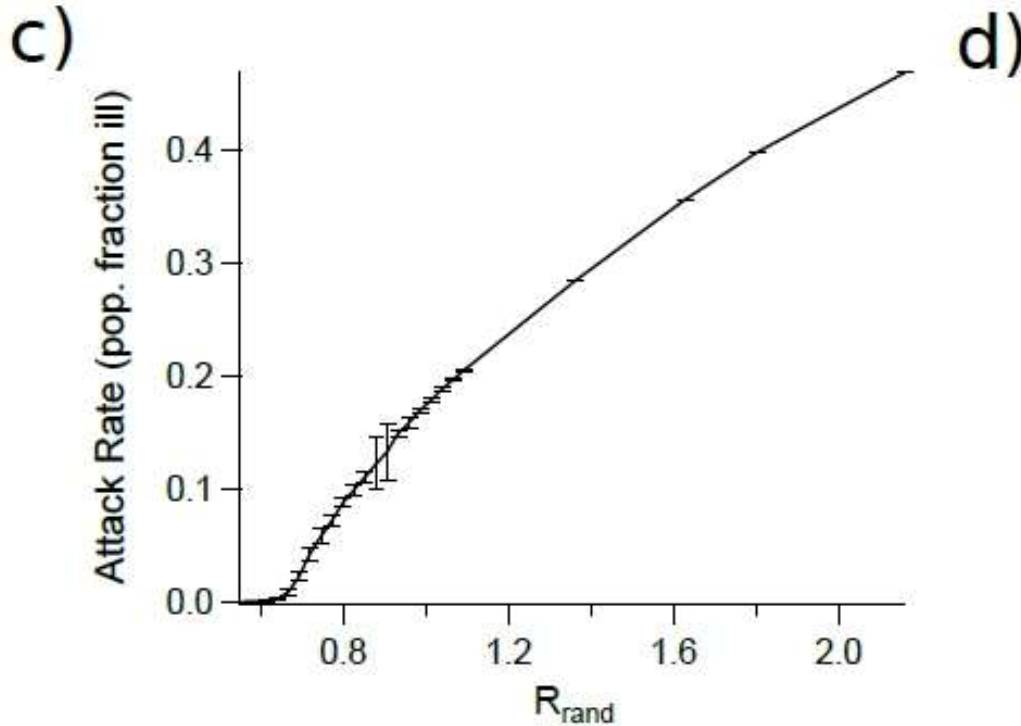
$$\begin{aligned}
 R_0 &= \mathbb{E} \left(\sum_i X_i(N) = \text{SYMPTOMATIC} \middle| X_J(0) = \text{SYMPTOMATIC} \right) \\
 &= \mathbb{E} \left(\mathbb{E} \left(\prod_n p_I(n) \middle| X_J(0) = \text{SYMPTOMATIC} \right) \right) \\
 &= \sum_j \sum_i \prod_n \left(1 - \prod_{g \in \mathcal{G}_j(n)} \left(1 - p_{j \rightarrow i}^g(n) \right) \right)
 \end{aligned}$$

Epidemic modelling: reproductive ratio R_0



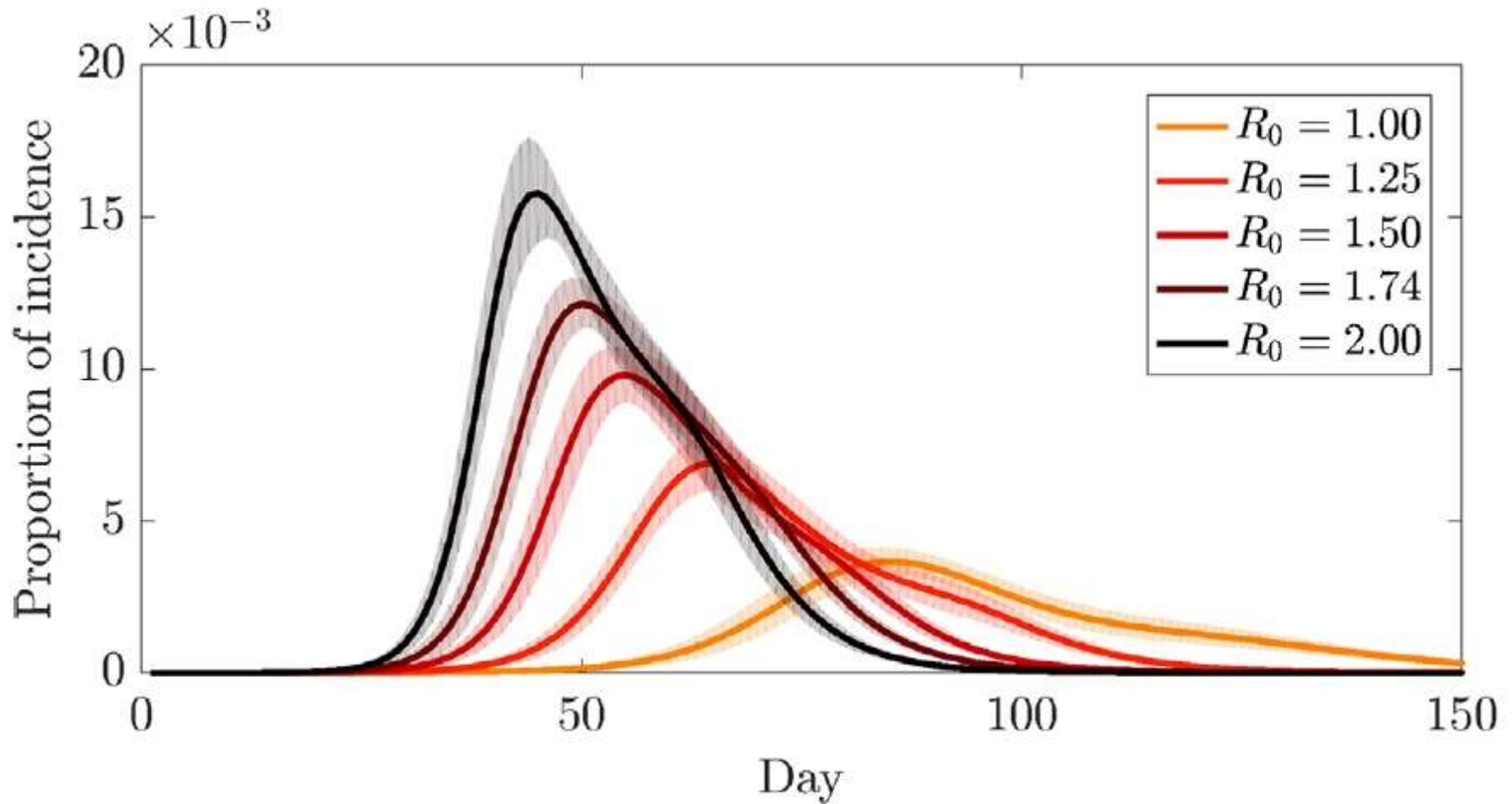
C. Zachreson, K. M. Fair, N. Harding, M. Prokopenko, Interfering with influenza: nonlinear coupling of reactive and static mitigation strategies, *J. Royal Society Interface*, 17(165): 20190728, 2020.

Epidemic modelling: reproductive ratio R_0



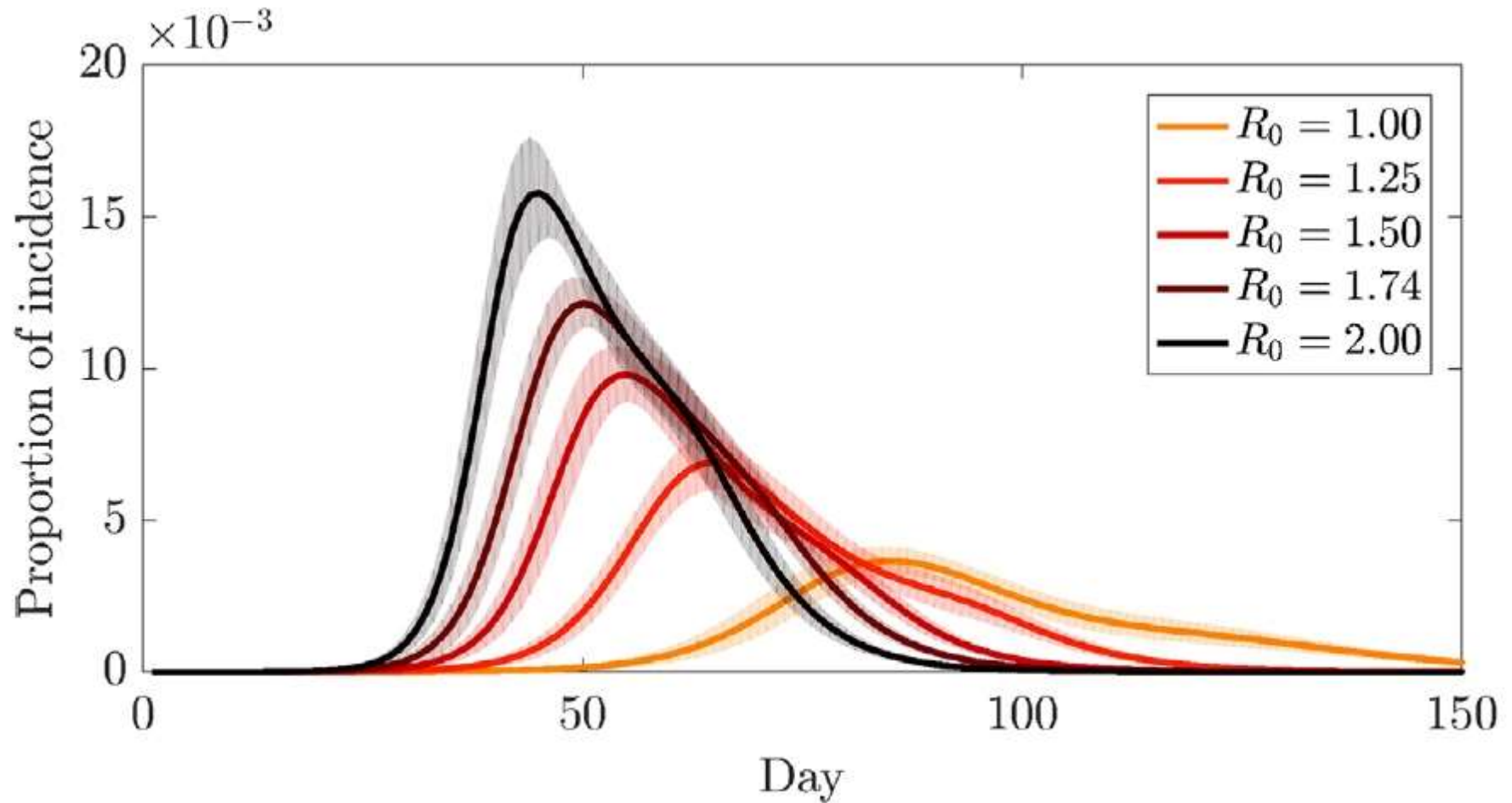
C. Zachreson, K. M. Fair, N. Harding, M. Prokopenko, Interfering with influenza: nonlinear coupling of reactive and static mitigation strategies, *J. Royal Society Interface*, 17(165): 20190728, 2020.

Epidemic modelling: reproductive ratio R_0

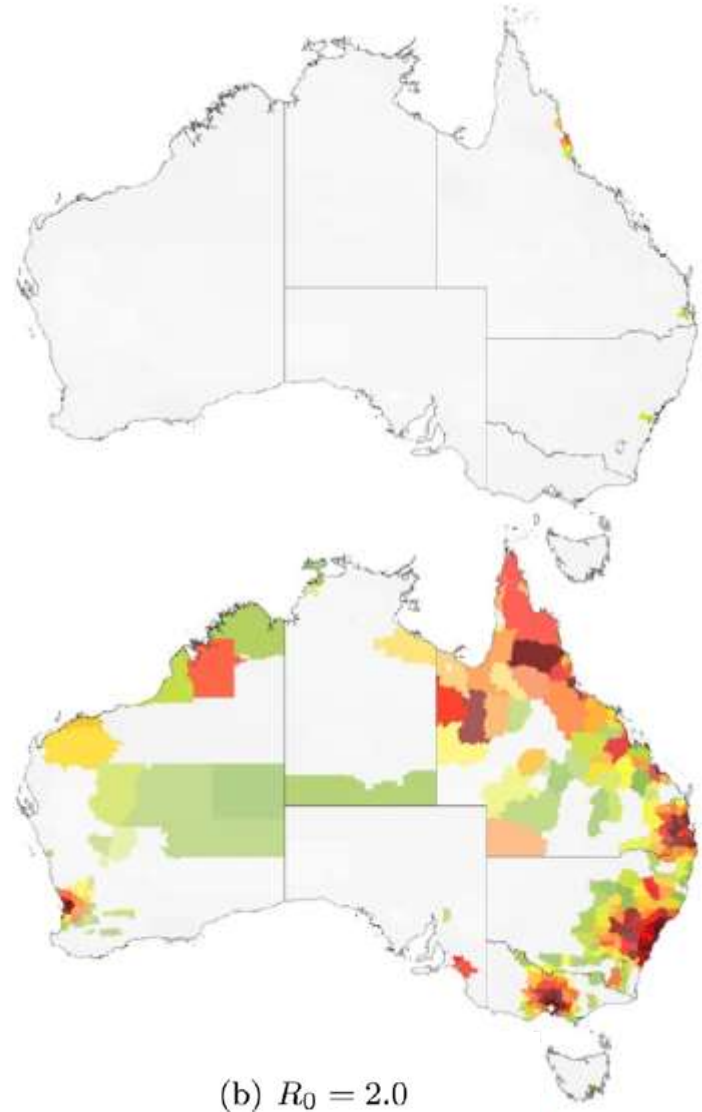
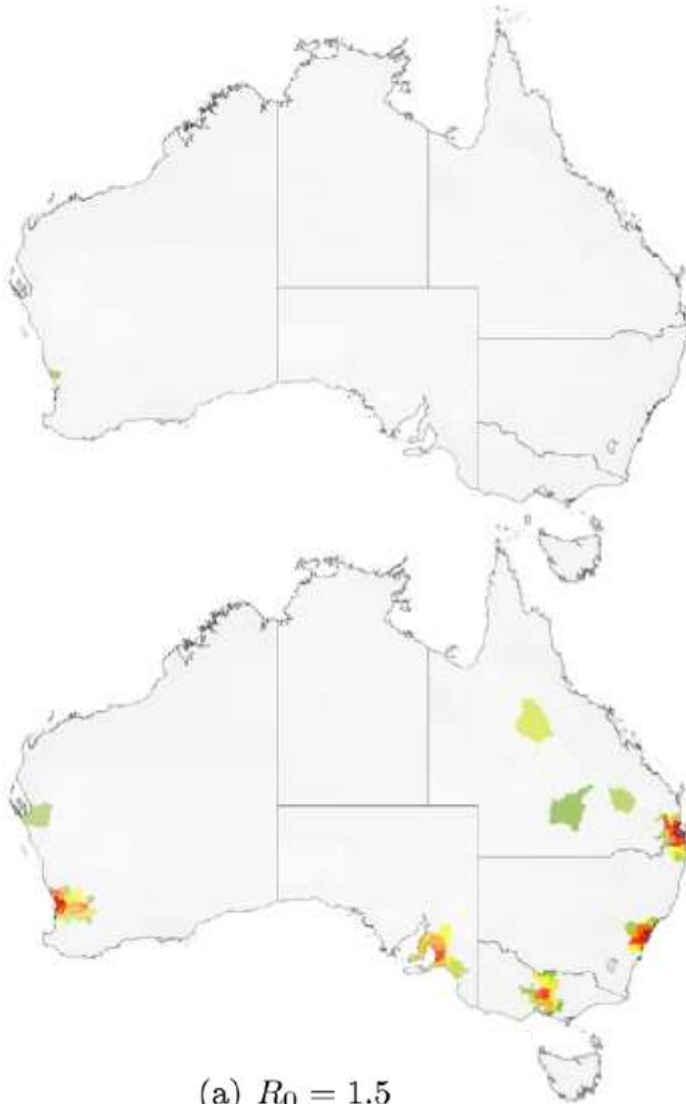


O. M. Cliff, N. Harding, M. Piraveenan, E. Y. Erten, M. Gambhir, M. Prokopenko, Investigating Spatiotemporal Dynamics and Synchrony of Influenza Epidemics in Australia: An Agent-Based Modelling Approach, *Simulation Modelling Practice and Theory*, 87, 412-431, 2018.

Epidemic modelling: reproductive ratio R_0

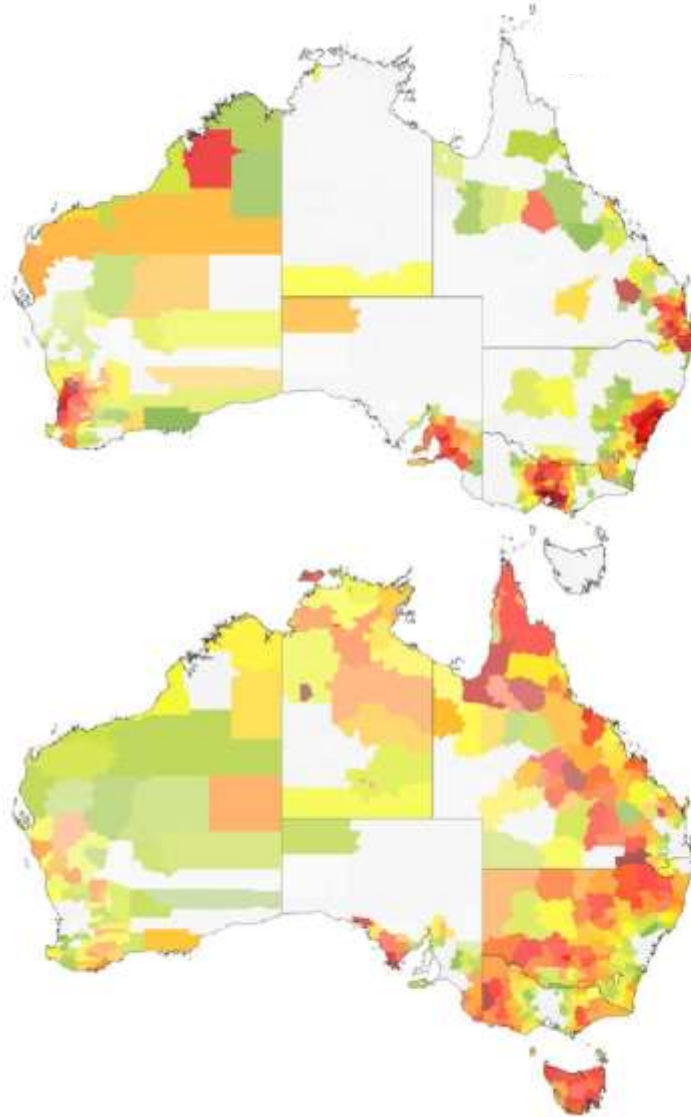


“herd immunity” threshold: $1 - 1/R_0$

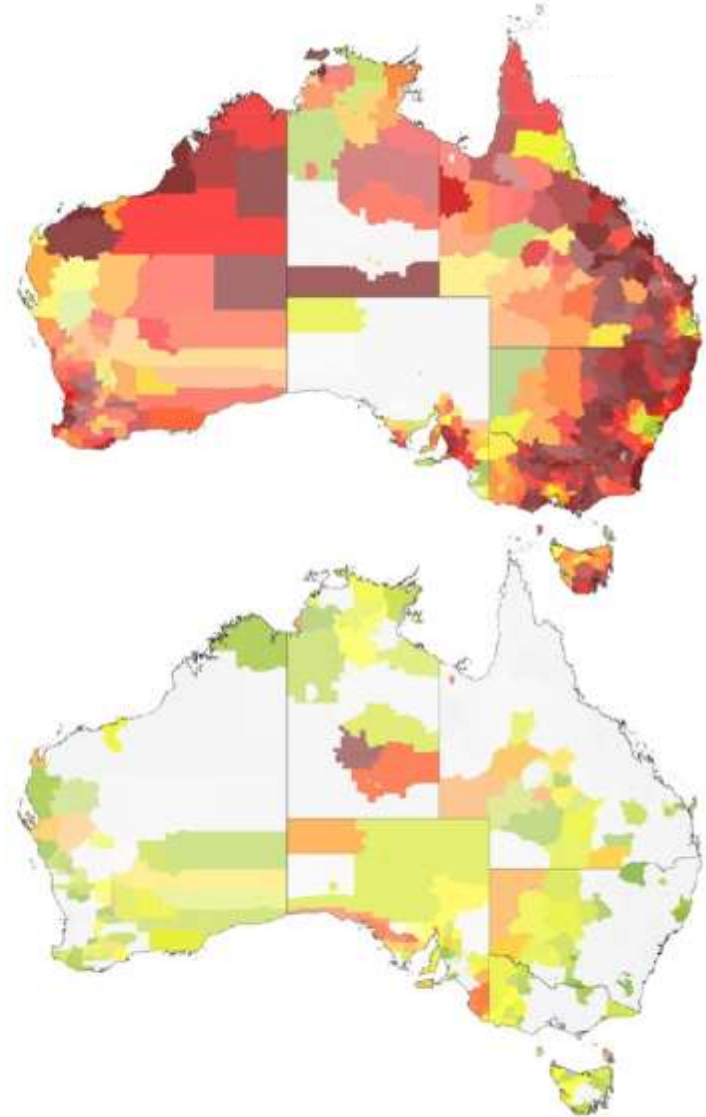




Spatiotemporal synchrony



(a) $R_0 = 1.5$



(b) $R_0 = 2.0$



Hierarchical spatial spread or wave-like diffusion??

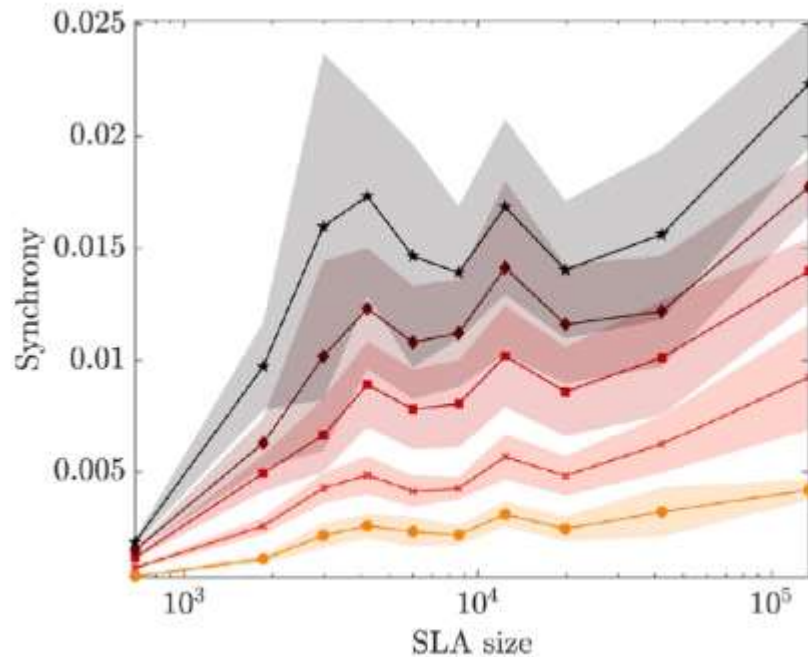
C. Viboud, O.N. Bjørnstad, D.L. Smith, L. Simonsen, M.A. Miller, B.T. Grenfell, Synchrony, waves, and spatial hierarchies in the spread of influenza, *Science* 312 (5772) (2006) 447–451.

The regional spread of infection correlates more closely with rates of movement of people to and from their workplaces (workflows) than with geographical distance.

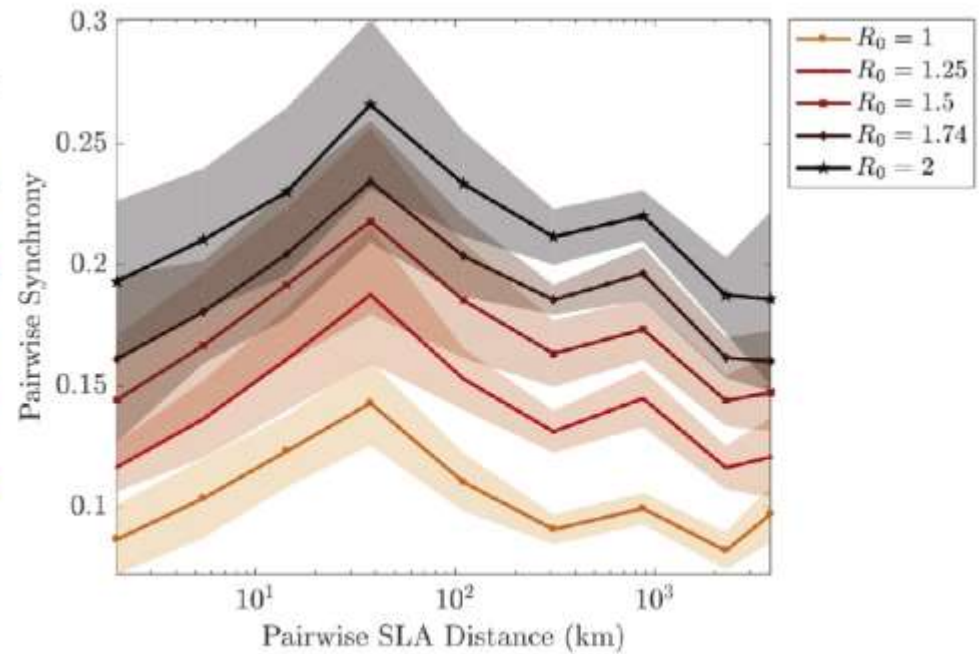
The hierarchy of spread is immediately apparent: The most populous states exhibit synchronized epidemics, whereas less populated states exhibit more erratic patterns, both relative to each other and to the continental norm.



Hierarchical spatial spread or wave-like diffusion??



(a)



(b)

RESEARCH ARTICLE | SOCIAL SCIENCES

Urbanization affects peak timing, prevalence, and bimodality of influenza pandemics in Australia: Results of a census-calibrated model

Cameron Zachreson^{1,*}, Kristopher M. Fair¹, Oliver M. Cliff¹, Nathan Harding¹, Mahendra Piraveenan¹ and Mikhail Prokopenko^{1,2}

¹Complex Systems Research Group, School of Civil Engineering, Faculty of Engineering and IT, The University of Sydney, Sydney, NSW 2006, Australia.

²Marie Bashir Institute for Infectious Diseases and Biosecurity, The University of Sydney, Westmead, NSW 2145, Australia.

✉*Corresponding author. Email: cameron.zachreson@sydney.edu.au

– Hide authors and affiliations

Science Advances 12 Dec 2018;
Vol. 4, no. 12, eaau5294
DOI: 10.1126/sciadv.aau5294

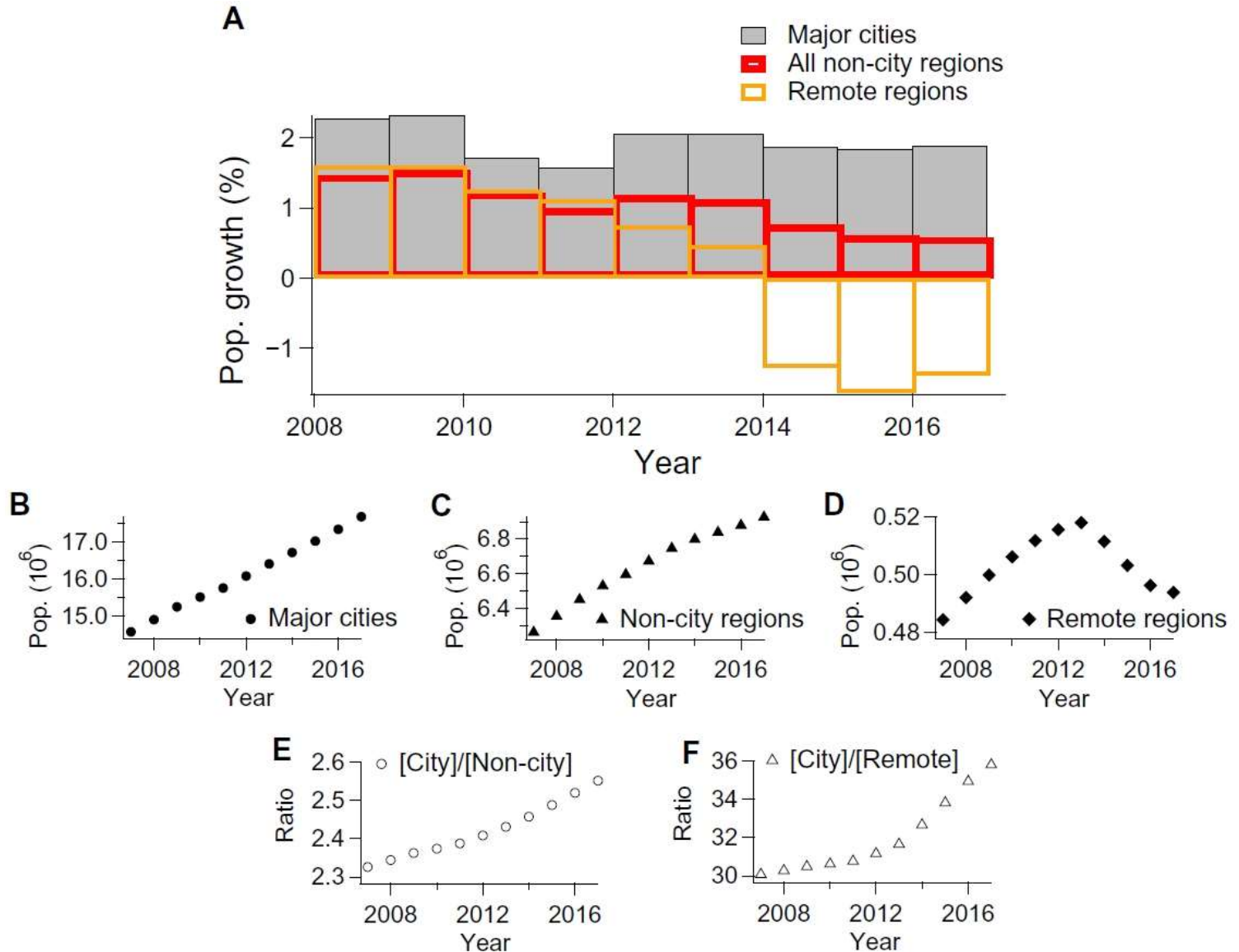
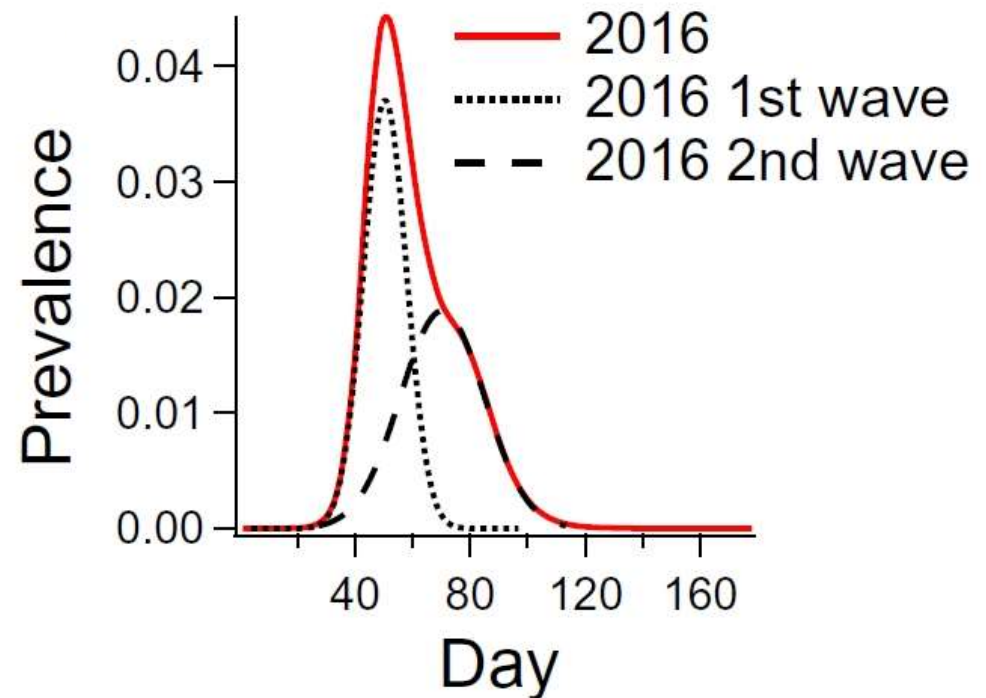
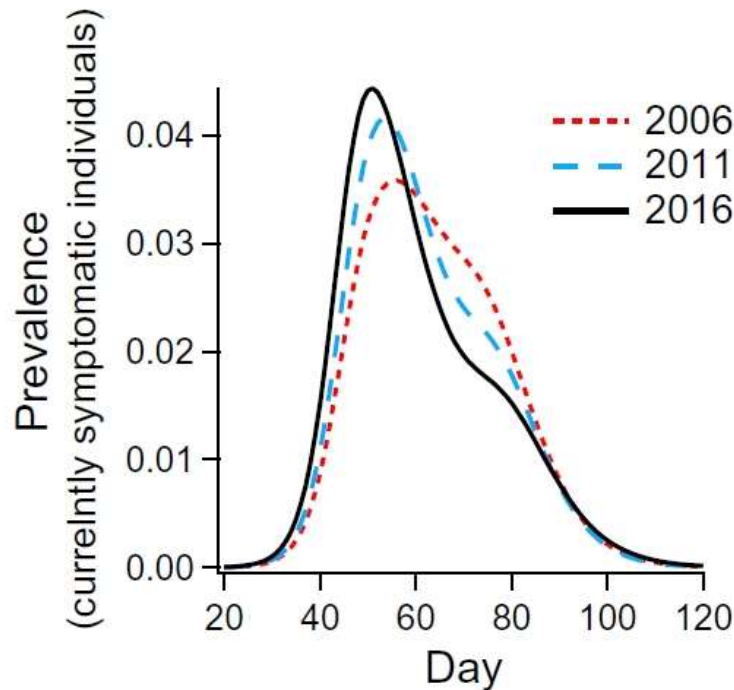




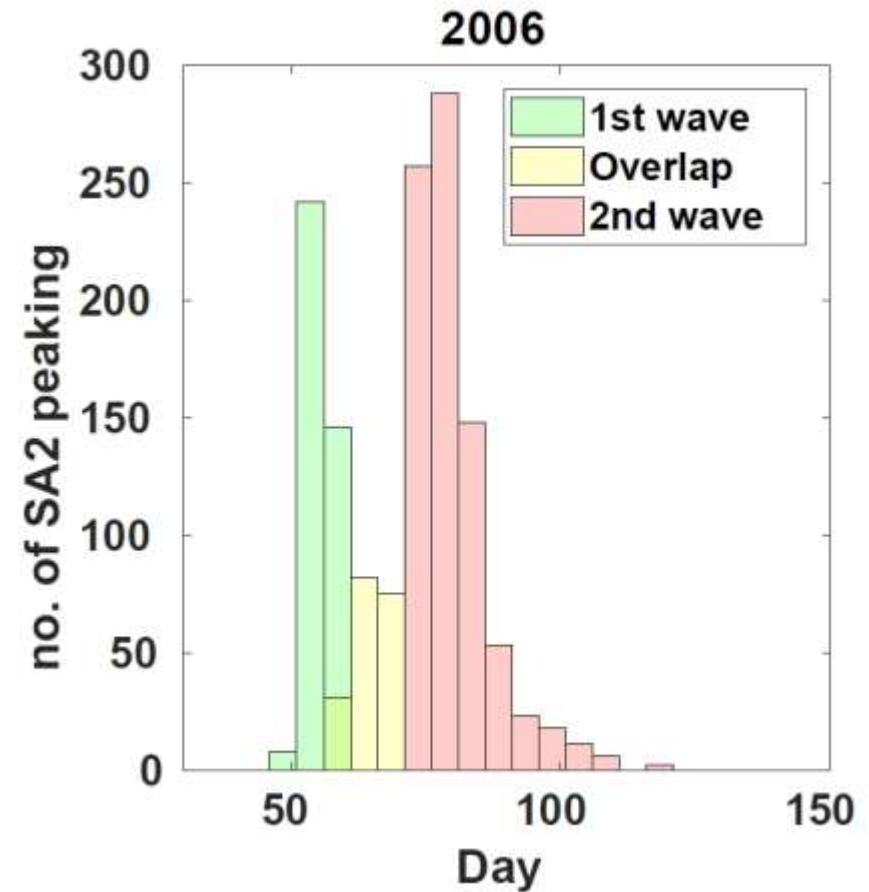
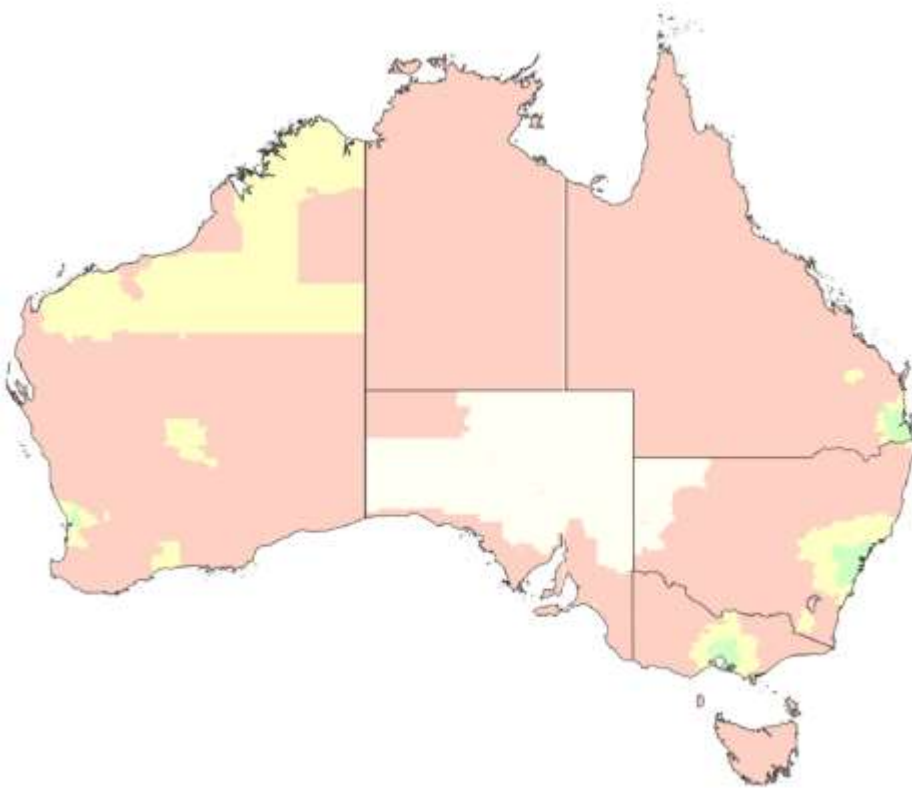
Table 1. Average daily incoming international air traffic.

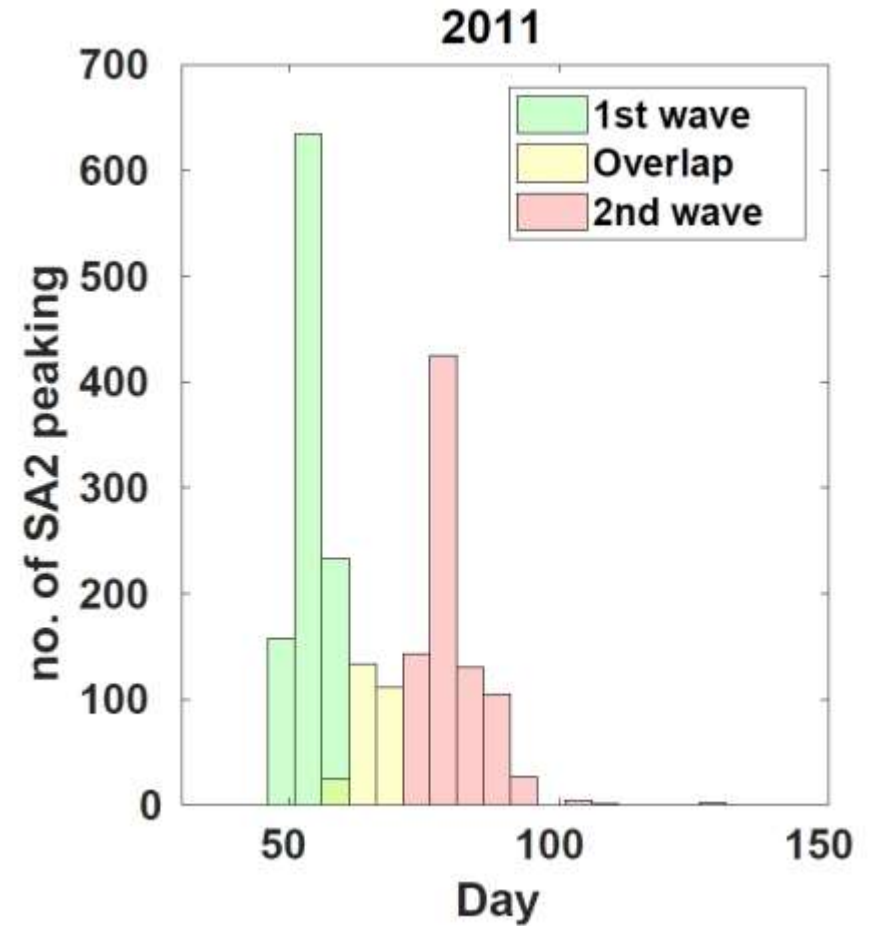
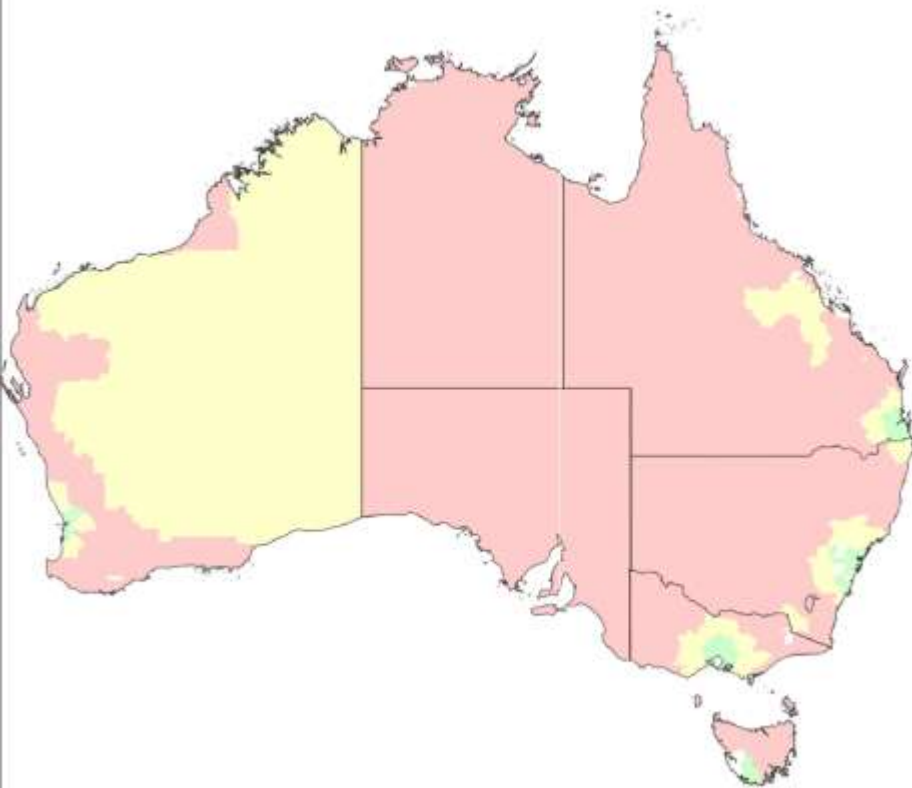
Airport	State	Year		
		2006	2011	2016
Sydney	New South Wales	13,214	15,995	19,991
Melbourne	Victoria	5,923	8,557	12,802
Brisbane	Queensland	5,053	5,946	7,299
Perth	Western Australia	2,766	4,512	5,906
Gold Coast	Queensland	285	1,044	1,435
Adelaide	South Australia	492	766	1,170
Cairns	Queensland	1,186	707	824
Darwin	Northern Territory	160	356	355
Townsville	Queensland	0	11	39

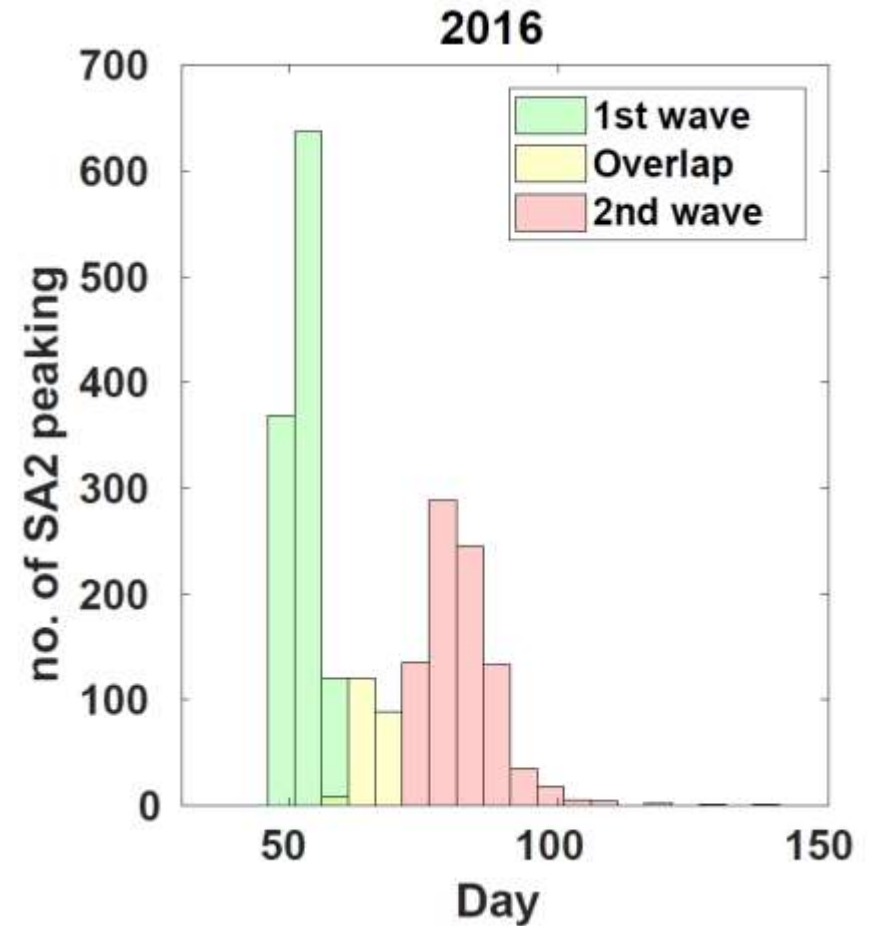
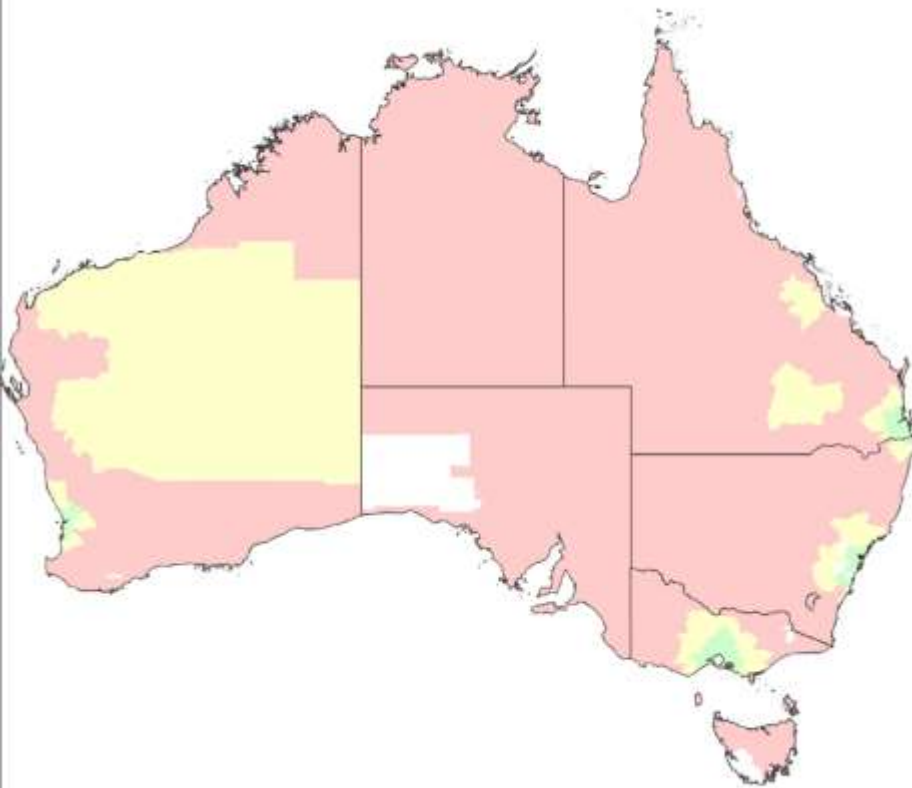
Prevalence and epidemic peaks: H1N1



C. Zachreson, K. M. Fair, O. M. Cliff, N. Harding, M. Piraveenan, M. Prokopenko, Urbanization affects peak timing, prevalence, and bimodality of influenza pandemics in Australia: Results of a census-calibrated model, *Science Advances*, 4(12), eaau5294, 2018.

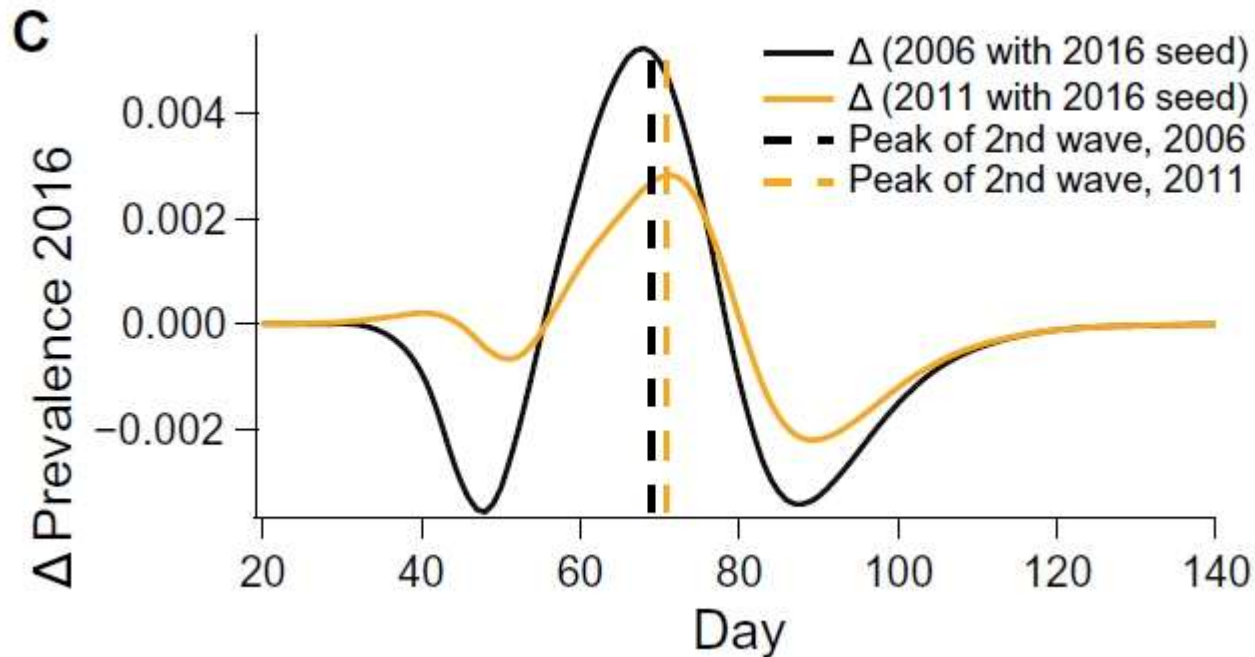




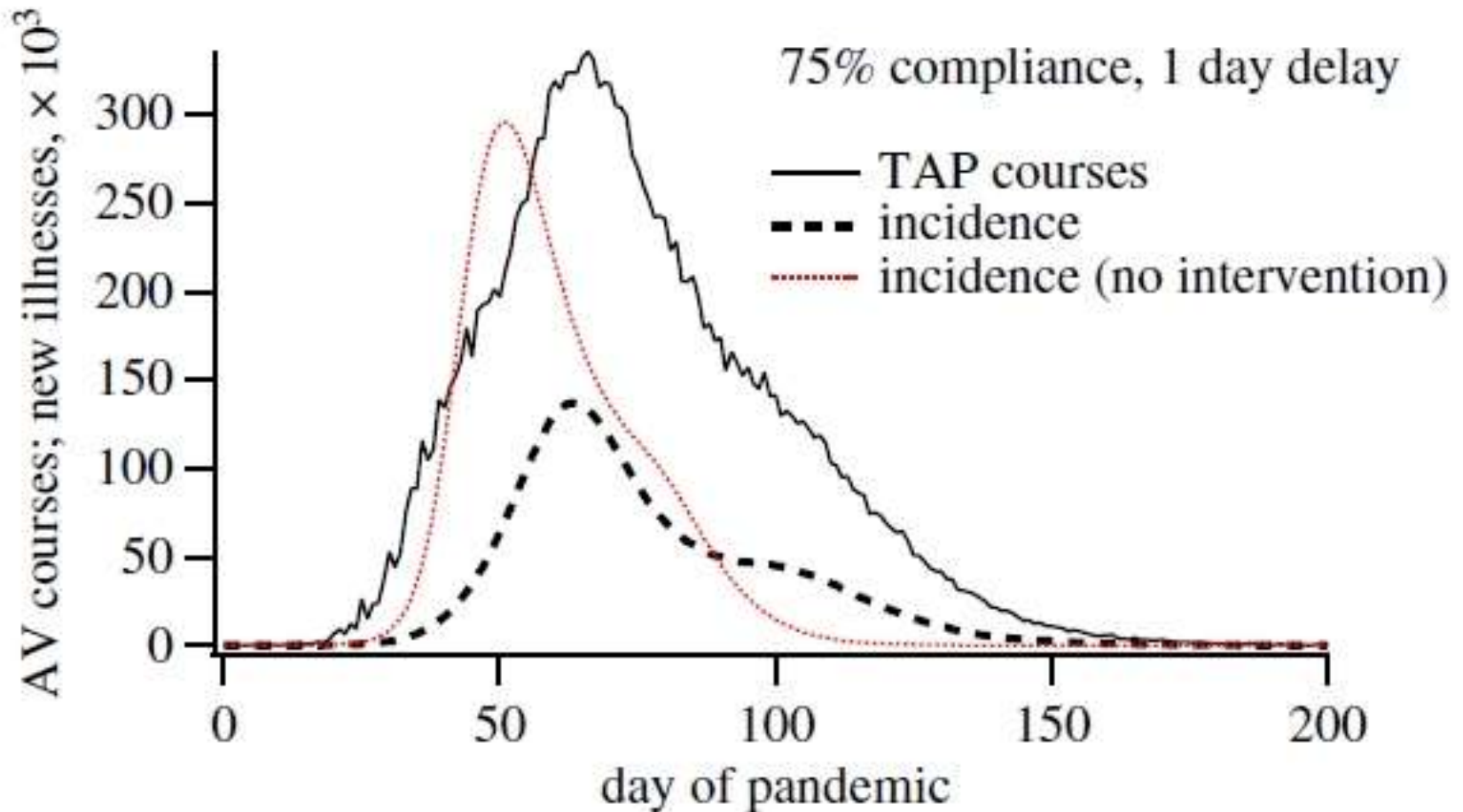




Key factors: higher urbanisation or more air traffic?



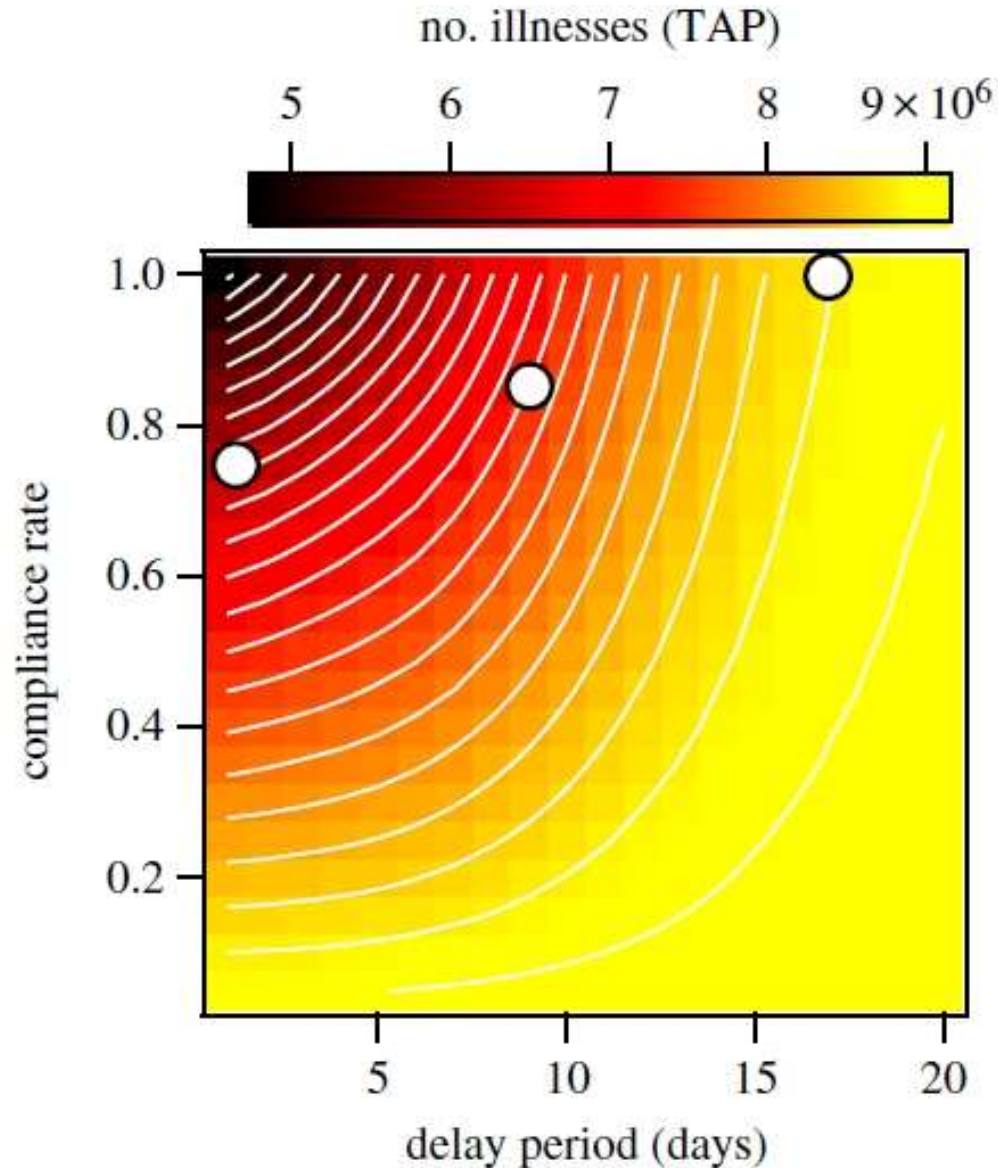
- seeding conditions have a larger impact on the first wave than on the second
- seeding does not account for the decrease in the intensity of the second pandemic wave from year to year, a trend that we ascribe to increased urbanisation



C. Zachreson, K. M. Fair, N. Harding, M. Prokopenko, Interfering with influenza: nonlinear coupling of reactive and static mitigation strategies, *J. Royal Society Interface*, 17(165): 20190728, 2020.

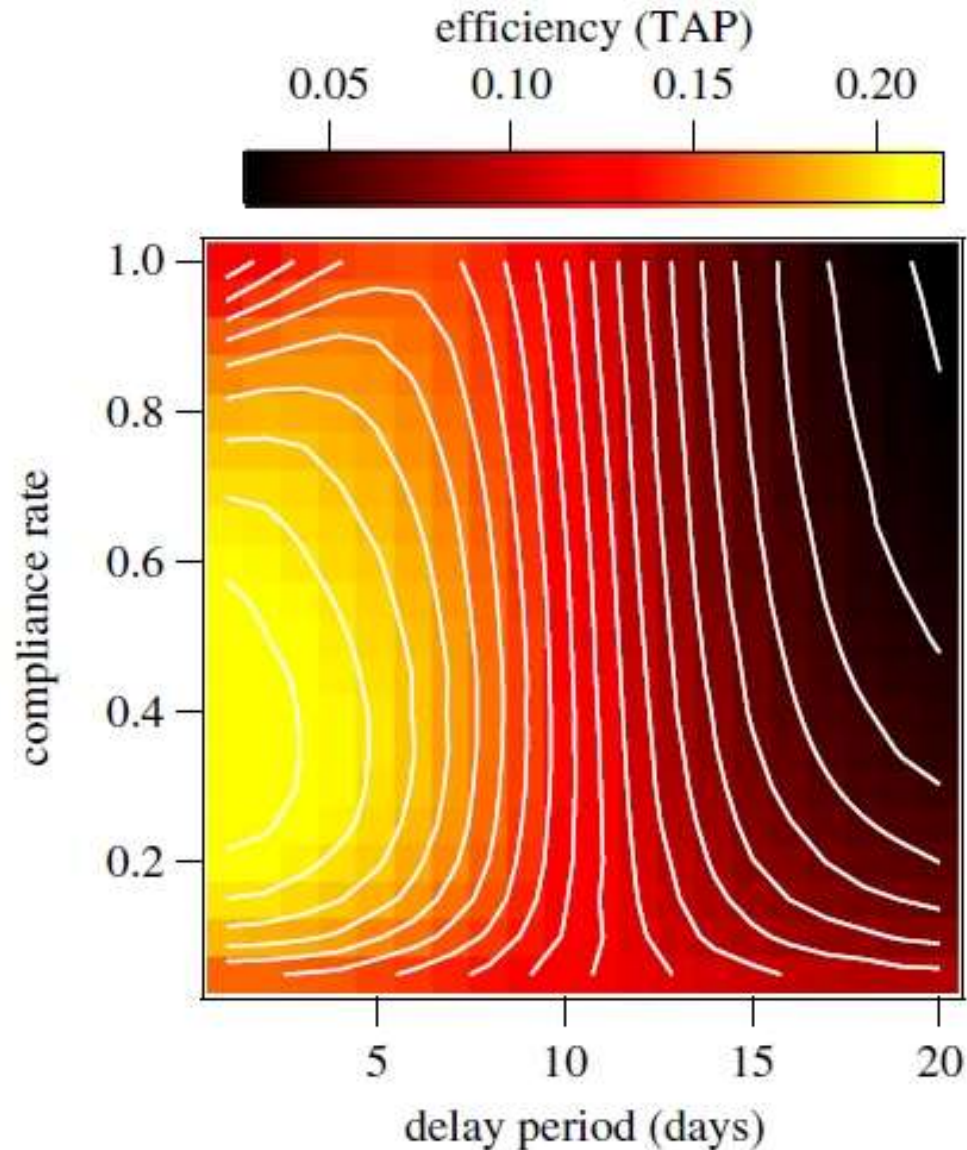


Phase space of intervention parameters: H1N1





Efficiency of interventions: H1N1




➤ Strengths:

- sensitivity analysis is embedded in heterogeneous agents
- individual-based rather than aggregate focus
- age-dependent epidemic characteristics
- spatial / geographic accuracy
- cross-jurisdictional impact
- time-dependent and context-dependent interventions
- counter-factual analysis (“what-if” scenarios: delays, scale, scope)
- critical phenomena analysis

➤ Weaknesses:

- need to calibrate multiple parameters
- reliance on high-performance computing

SCIENTIFIC REPORTS



OPEN

Network properties of salmonella epidemics

Oliver M. Cliff¹, Vitali Sintchenko^{2,3}, Tania C. Sorrell^{2,3}, Kiranmayi Vadlamudi¹, Natalia McLean¹ & Mikhail Prokopenko^{1,3}

Received: 7 January 2019

Accepted: 3 April 2019

Published online: 16 April 2019

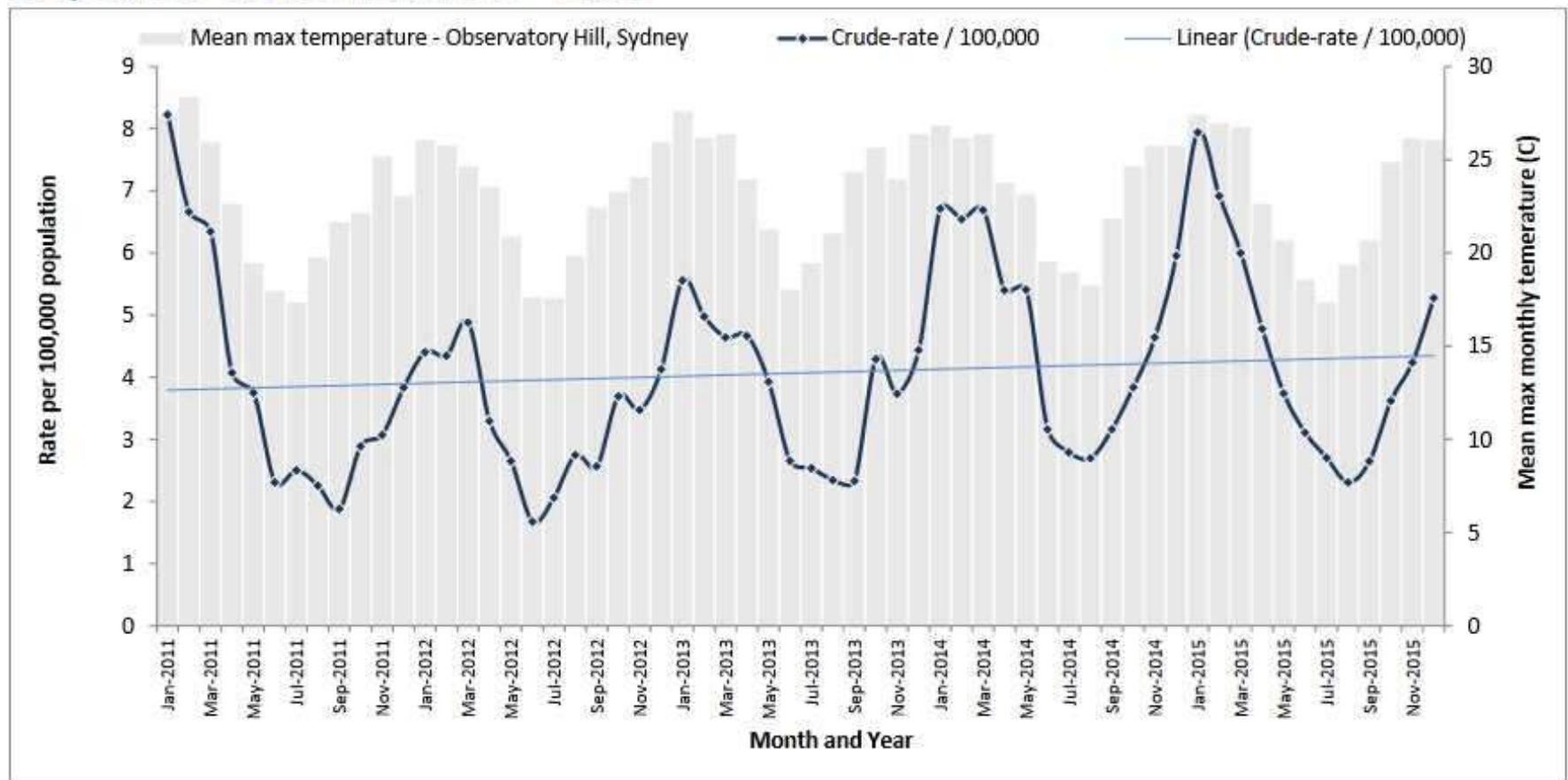
We examine non-typhoidal Salmonella (*S. Typhimurium* or STM) epidemics as complex systems, driven by evolution and interactions of diverse microbial strains, and focus on emergence of successful strains. Our findings challenge the established view that seasonal epidemics are associated with random sets of co-circulating STM genotypes. We use high-resolution molecular genotyping data comprising 17,107 STM isolates representing nine consecutive seasonal epidemics in Australia, genotyped by multiple-locus variable-number tandem-repeats analysis (MLVA). From these data, we infer weighted undirected networks based on distances between the MLVA profiles, depicting epidemics as networks of individual bacterial strains. The network analysis demonstrated dichotomy in STM populations which split into two distinct genetic branches, with markedly different prevalences. This distinction revealed the emergence of dominant STM strains defined by their local network topological properties, such as centrality, while correlating the development of new epidemics with global network features, such as small-world propensity.

- **93.8 million cases** and **155,000 deaths** each year globally
 - more cases per capita in Australia than anywhere else in the world
 - most fatal foodborne disease in Australia

- **S. Typhimurium (STM)** is the dominant subspecies of non-typhoidal salmonellosis

- Drivers of STM evolution remain poorly understood

Crude monthly salmonellosis notification rate per 100,000 population and mean maximum monthly temperature* in NSW from 2011 – 2015.



*Bureau of Meteorology, mean maximum monthly temperatures taken from Observatory Hill in Sydney

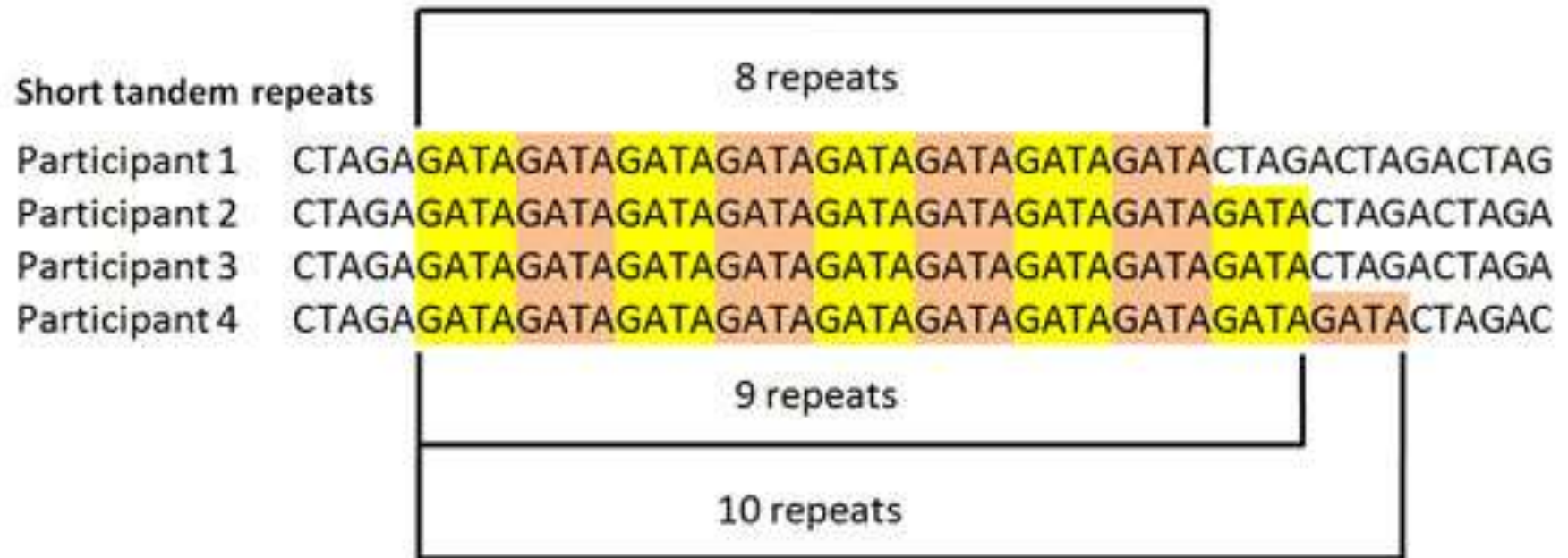
Dataset (NSW Salmonella Reference Laboratory)

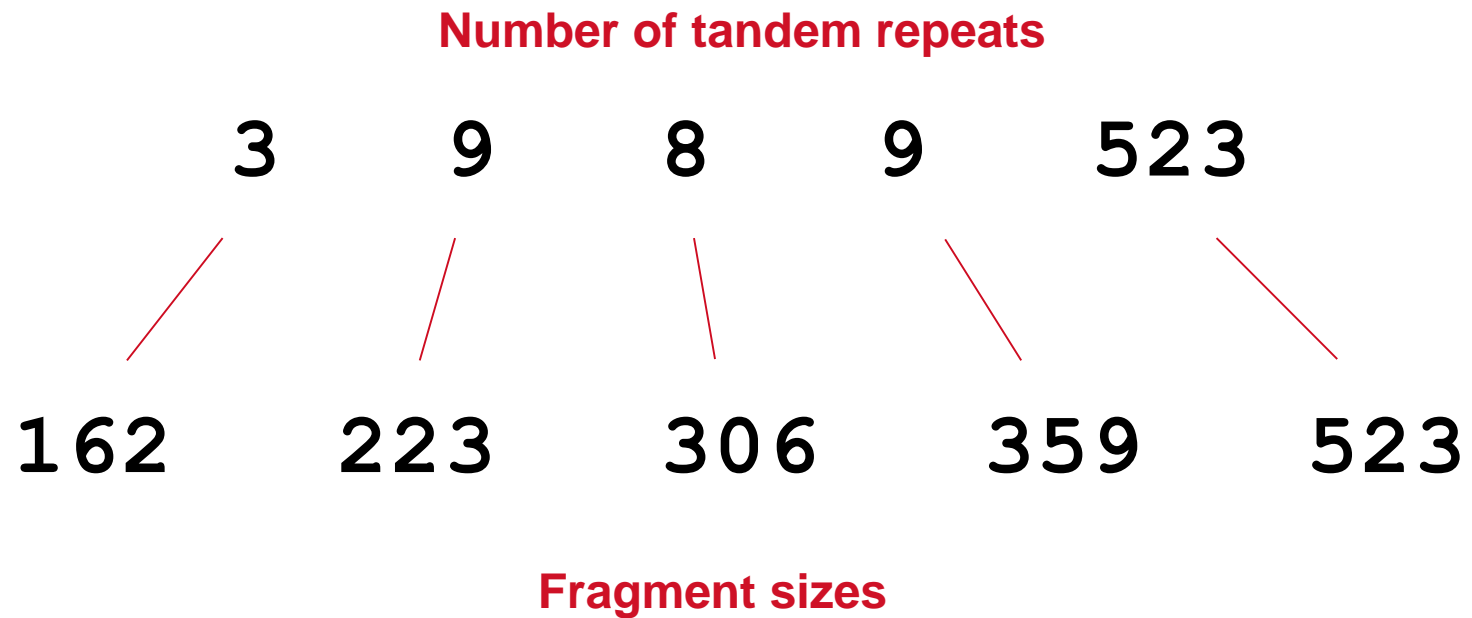
- Nine consecutive seasons of instances in NSW
 - 1st January 2008 to 31st December 2016

- **17,107 isolates** of STM
 - 99.3% of all STM found from human cases in NSW over 3,287 days

- Genotyped through **MLVA**
 - **M**ultiple-**L**ocus **V**ariable-number tandem repeat (VNTR) **A**alysis
 - 1675 unique MLVAs identified

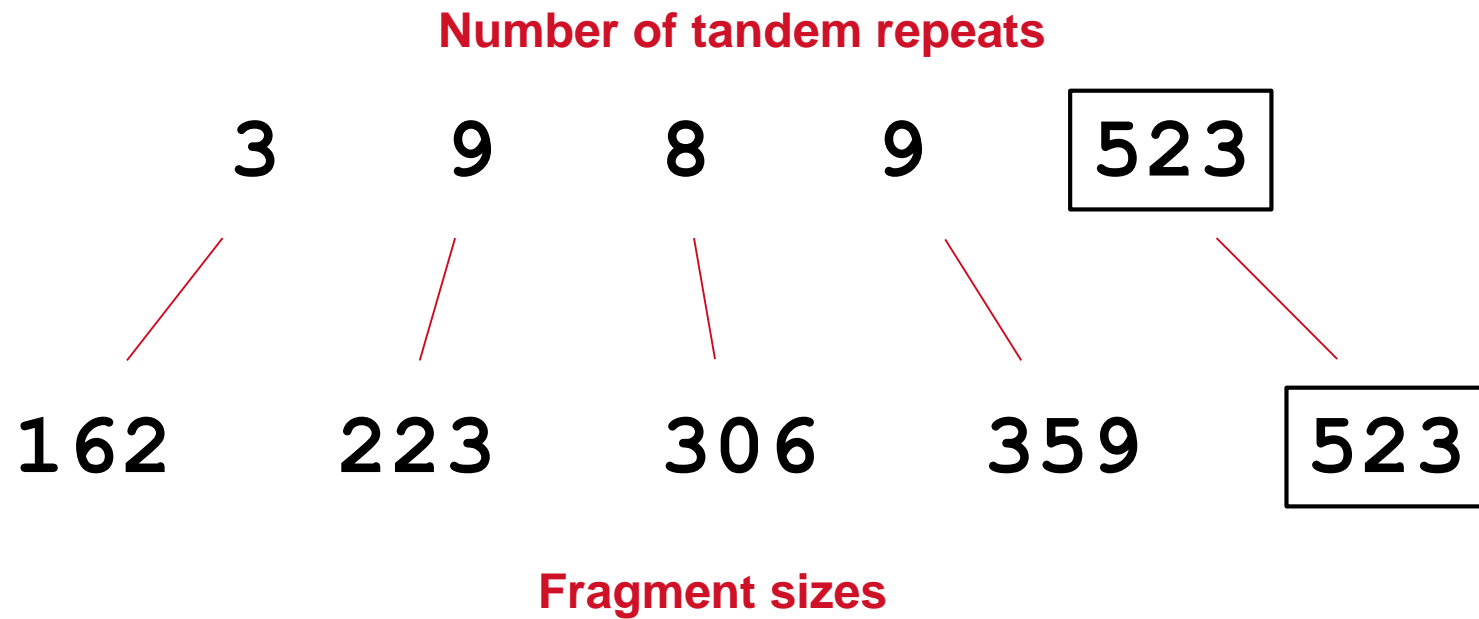
Tandem repeats

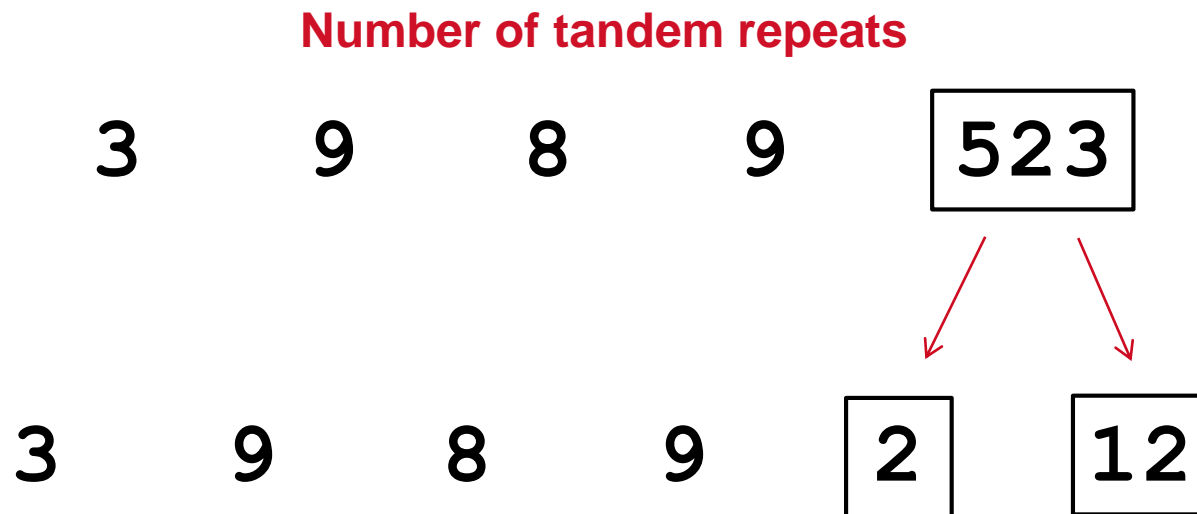






What is MLVA?





Larsson et al. (2009)

Larsson, J. T. et al. Development of a new nomenclature for *Salmonella* Typhimurium multilocus variable number of tandem repeats analysis (MLVA). *Eurosurveillance* 14, pii:19174 (2009).

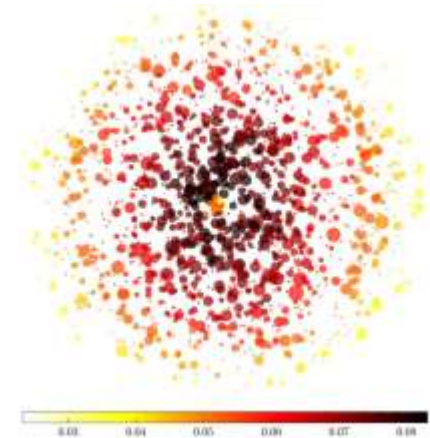
Top 10 *Salmonella* Typhimurium MLVA profiles and number of isolates from 2011-2015, NSW*

	2011	2012	2013	2014	2015
1	3-9-7-13-523 (259)	3-17-9-12-523 (150)	3-17-9-12-523 (159)	3-17-9-11-523 (210)	3-12-11-14-523 (100)
2	3-10-8-9-523 (149)	3-9-8-13-523 (124)	3-9-8-13-523 (83)	3-12-11-14-523 (149)	3-17-9-11-523 (91)
3	3-9-8-13-523 (113)	3-9-7-13-523 (100)	3-9-7-13-523 (74)	3-12-12-9-523 (141)	3-12-12-9-523 (82)
4	3-9-7-14-523 (92)	3-16-9-12-523 (66)	3-10-14-12-496 (61)	3-10-7-12-523 (99)	3-12-13-9-523 (54)
5	3-12-9-10-550 (76)	3-10-8-9-523 (50)	3-10-7-14-523 (55)	3-9-7-12-523 (98)	3-24-13-10-523 (53)
6	3-9-7-15-523 (59)	3-9-8-12-523 (38)	3-13-11-9-523 (48)	3-9-8-12-523 (97)	3-10-8-12-523 (42)
7	3-14-11-12-523 (50)	3-9-8-14-523 (38)	3-9-7-14-523 (45)	3-16-9-11-523 (94)	3-9-7-12-523 (40)
8	3-10-14-12-496 (48)	3-9-9-13-523 (37)	3-23-23-11-523 (43)	3-17-10-11-523 (94)	3-17-8-11-523 (39)
9	3-12-15-13-523 (46)	3-9-9-12-523 (34)	3-10-8-9-523 (39)	3-10-13-11-496 (52)	3-16-9-11-523 (37)
10	3-13-11-9-523 (30)	3-12-11-13-523 (29)	3-17-9-11-523 (39)	3-16-9-12-523 (50)	3-24-14-10-523 (33)

*Colour code indicates closely related MLVA patterns.

MLVA isolates as a complex network

- Construct a complete graph
 - 1675 nodes (unique MLVA profiles)
 - edge weights are inverse of pairwise MLVA distance
- Compute closeness centrality of MLVA profile in network
- Cluster nodes (MLVA profiles)
 - **partitioned** clusters
 - **overlapping** clusters
- Trace changes in the global network and individual clusters



➤ Edge weights

- Inverse of Manhattan distance (L1-norm)

x 3 9 8 9 2 12

➤ Edge weights

- Inverse of Manhattan distance (L1-norm)

<i>x</i>	3	9	8	9	2	12
<i>y</i>	3	10	8	8	2	12

➤ Edge weights

- Inverse of Manhattan distance (L1-norm)

x	3	9	8	9	2	12
y	3	10	8	8	2	12
$d(x, y)$	0	$ +1 $	0	$ -1 $	0	0

➤ Edge weights

- Inverse of Manhattan distance (L1-norm)

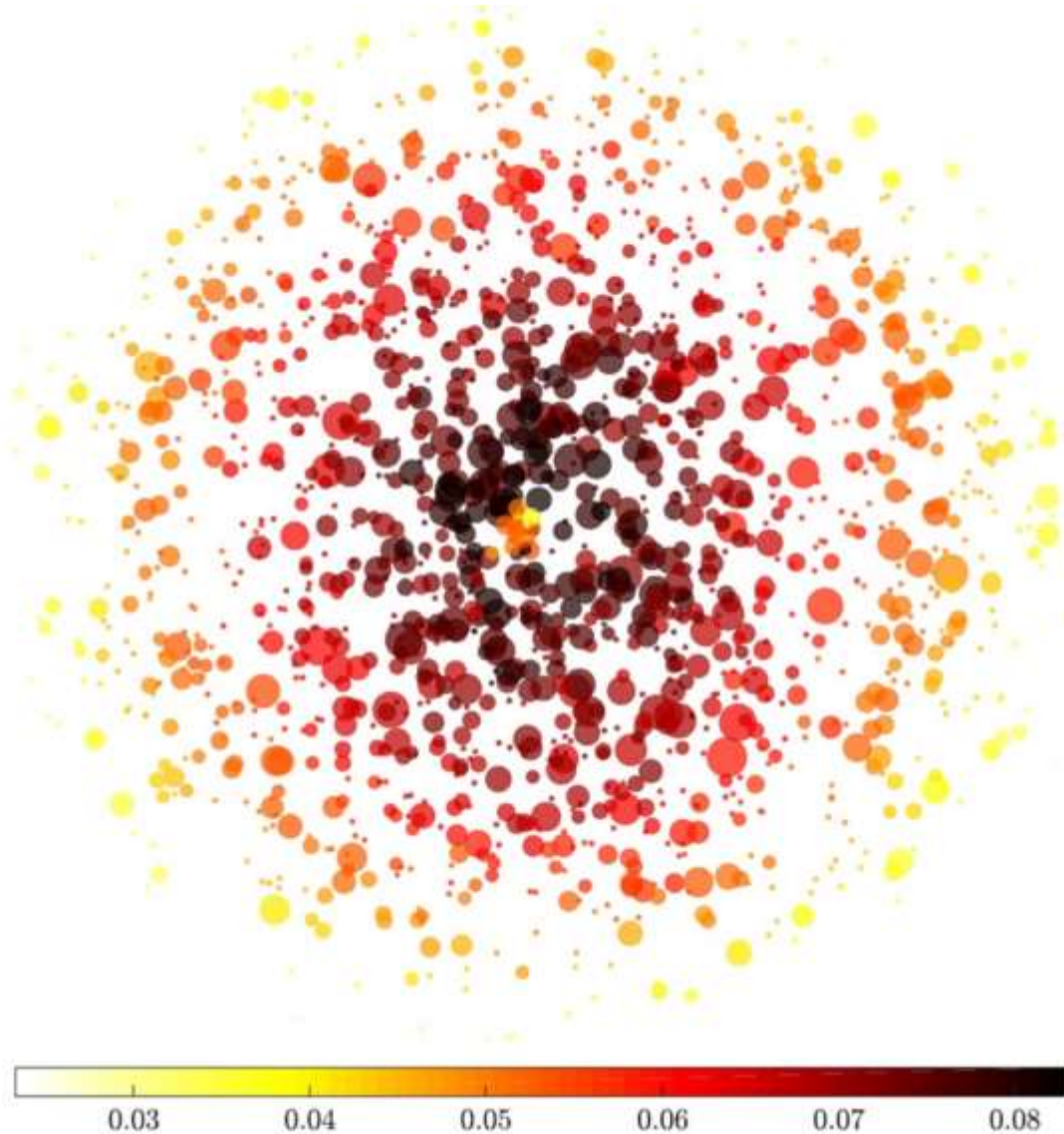
x	3	9	8	9	2	12
y	3	10	8	8	2	12
$d(x, y)$	0	$ +1 $	0	$ -1 $	0	0

➤ Closeness centrality $C(x) = \frac{1}{\sum_y d(x, y)}$

Path length: average distance to all other nodes

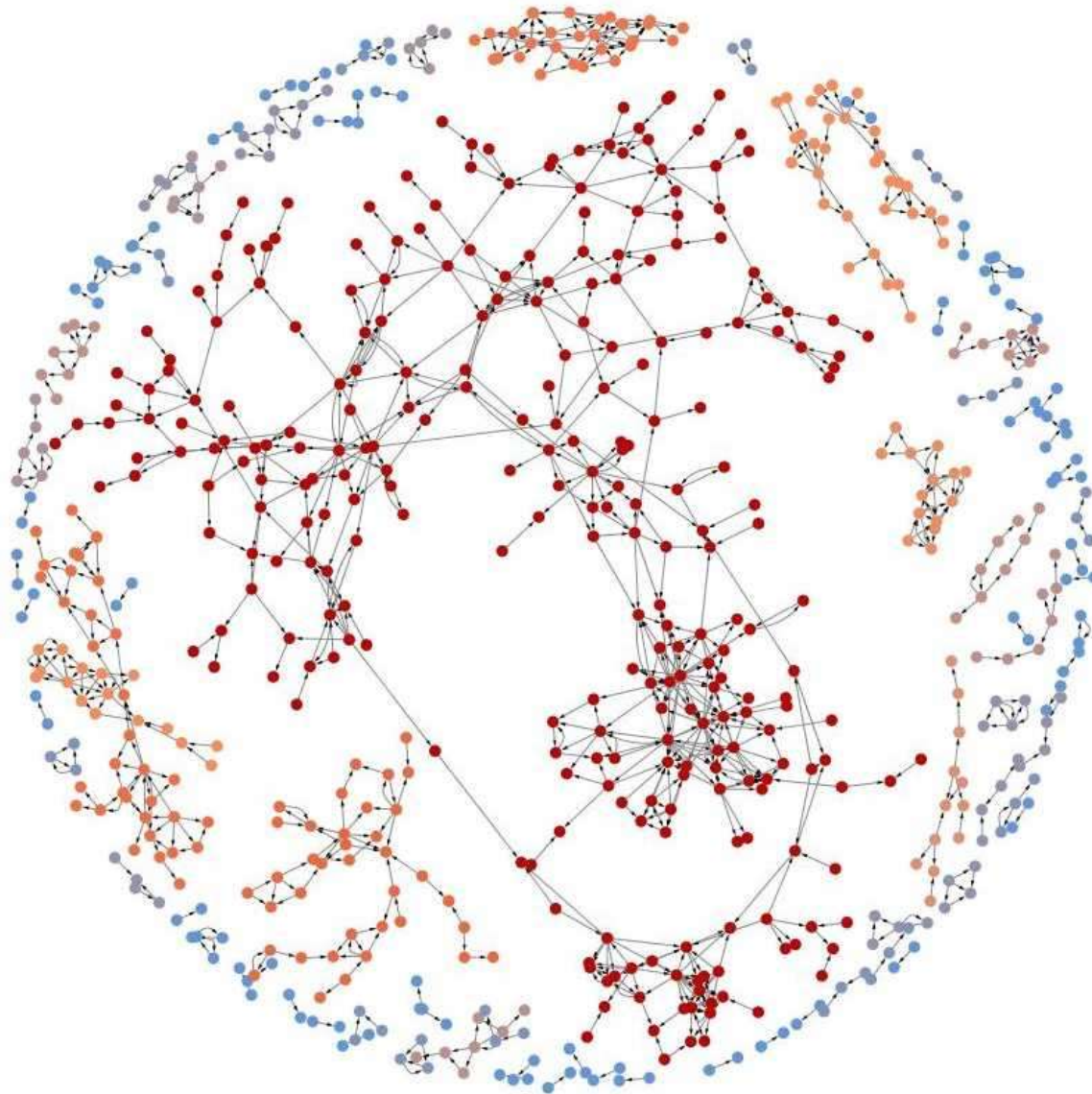


Closeness centrality of MLVA profiles



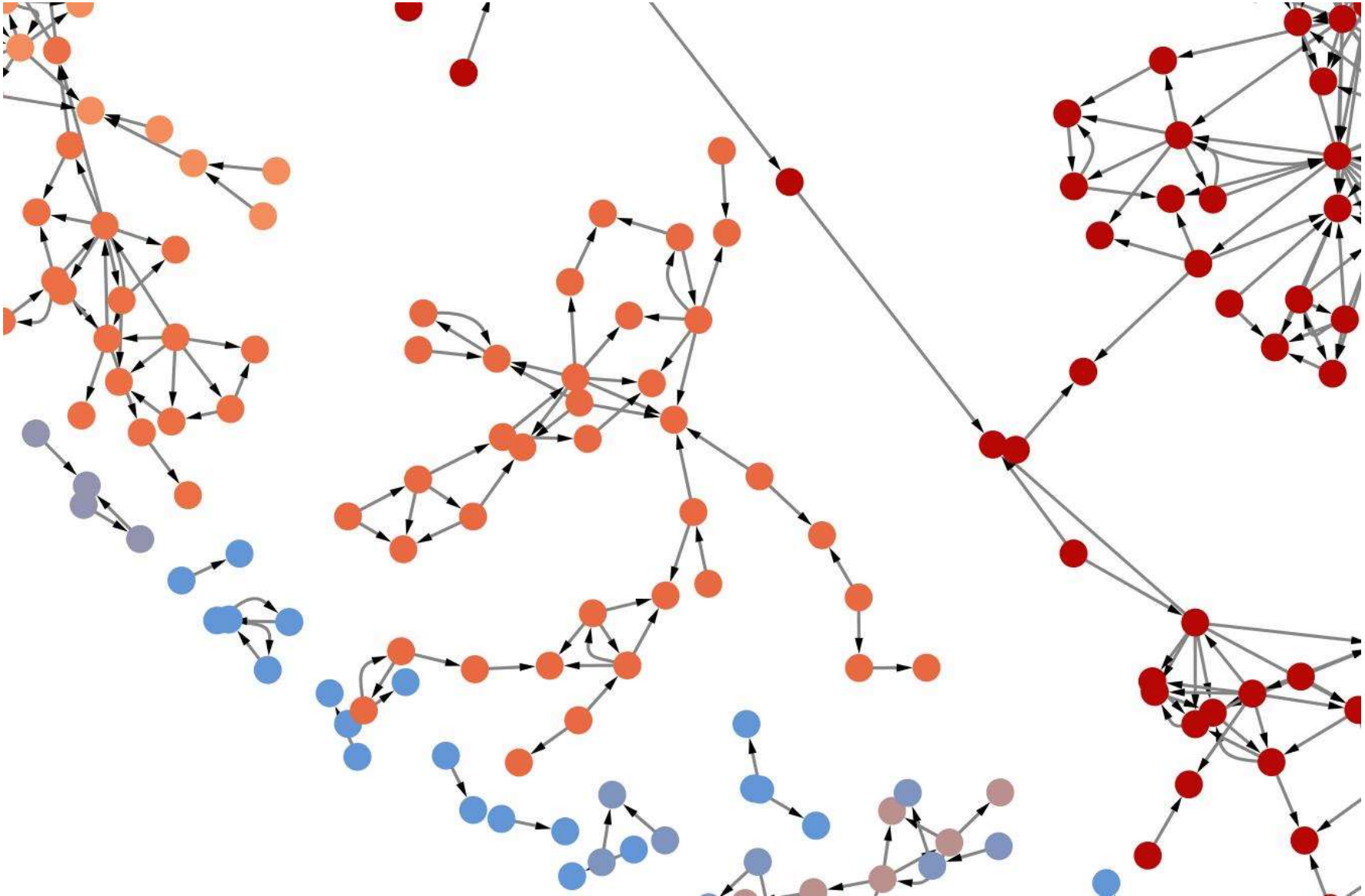


Network of MLVA profiles



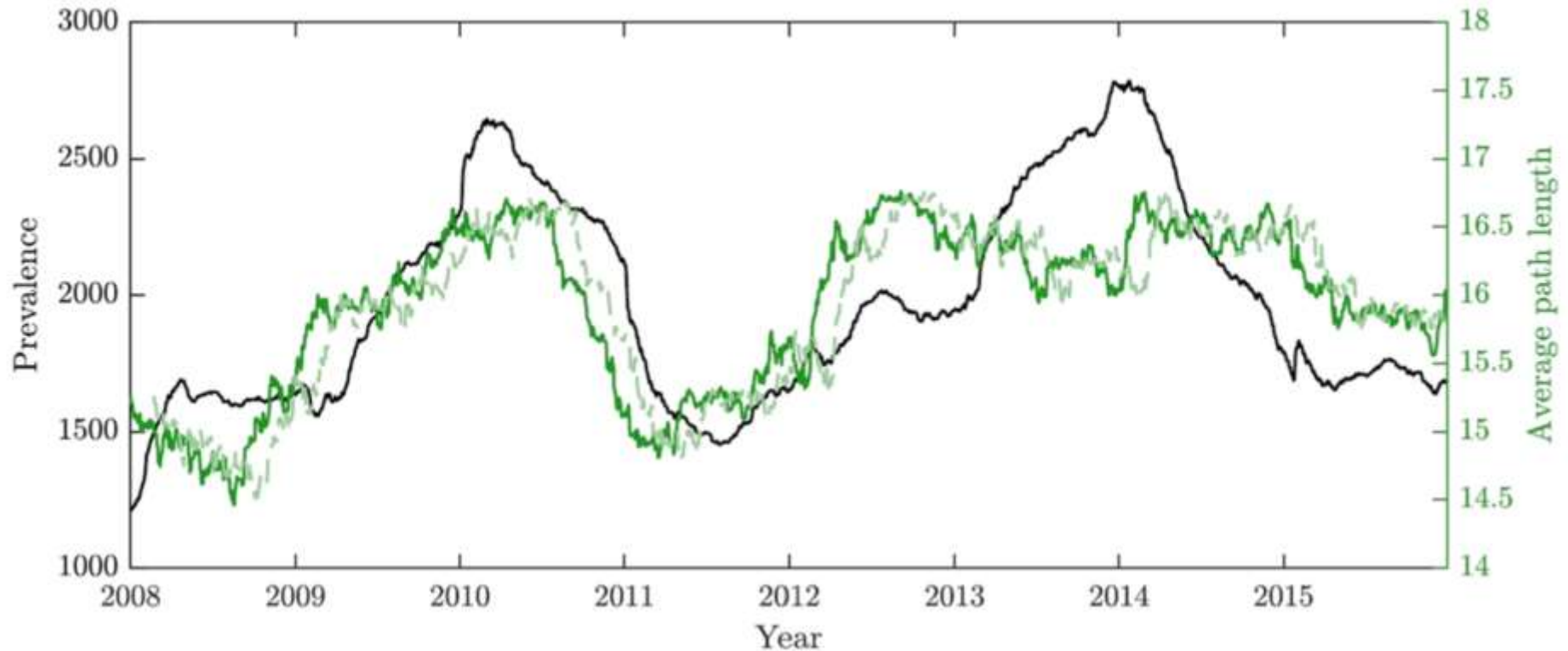


Network of MLVA profiles





Global network properties: average path length

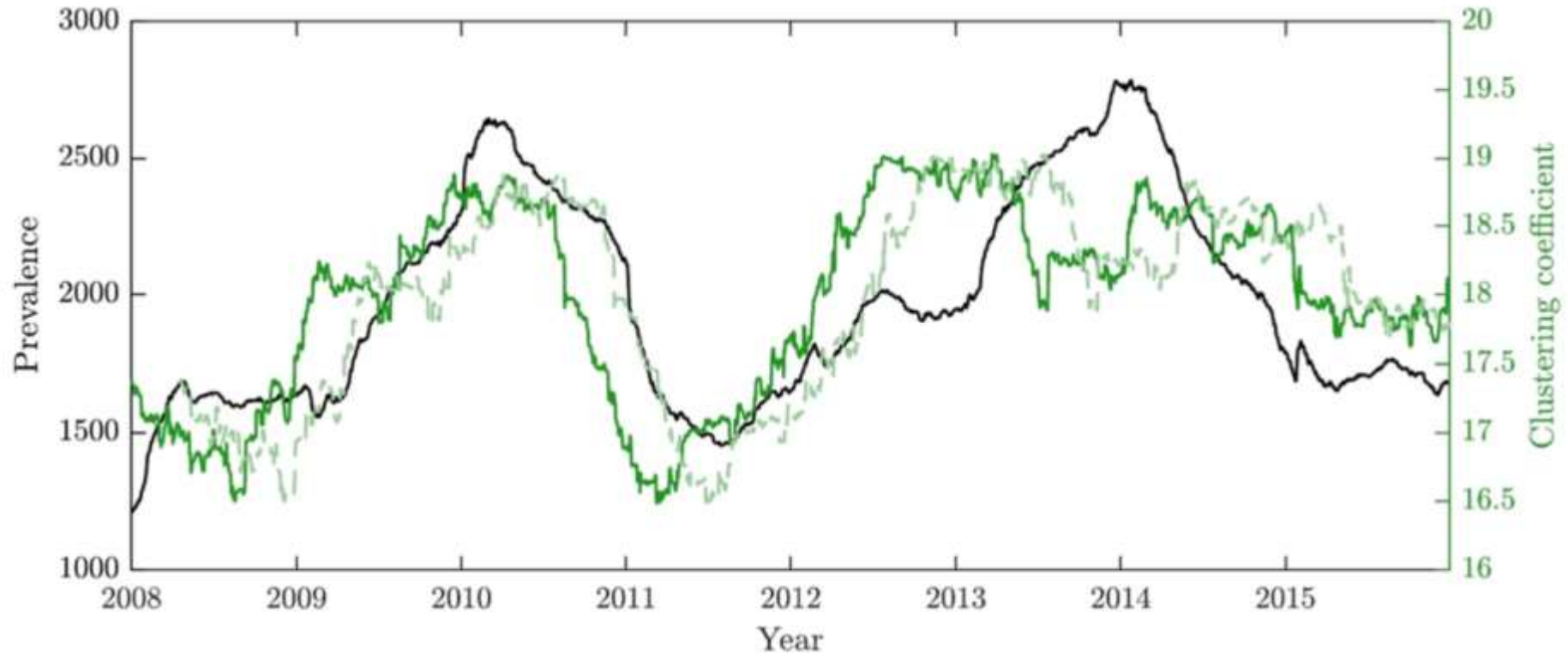


Average path length (average distance to all other nodes)

correlates with prevalence: $\rho \approx 0.7$ at ~ 100 days



Global network properties: clustering coefficient



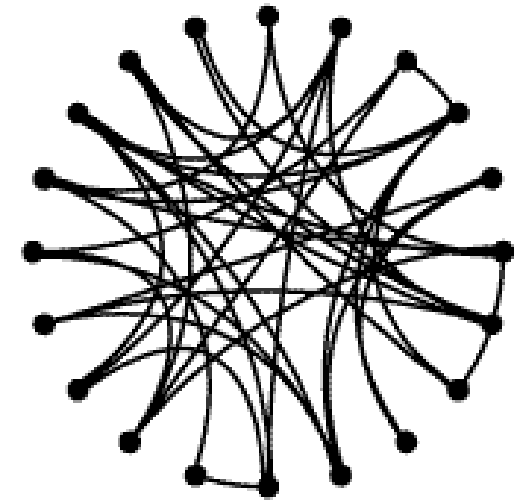
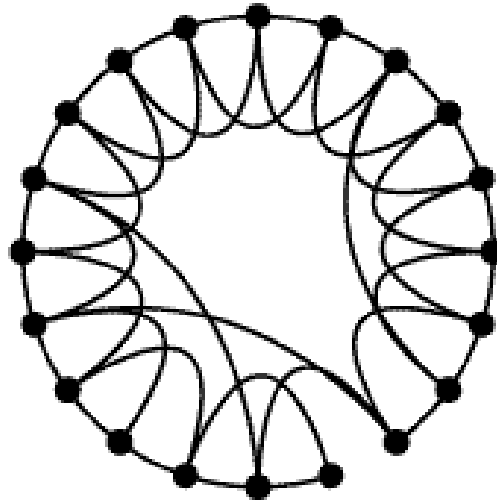
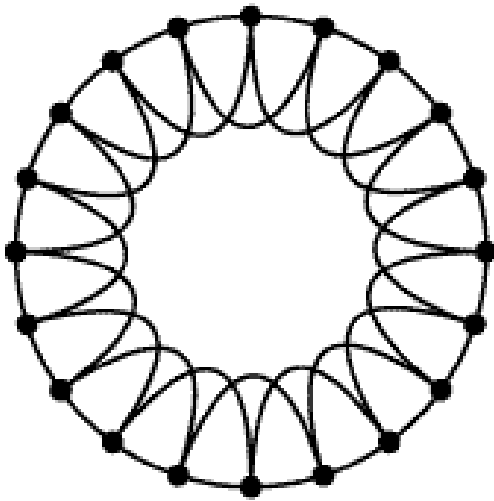
Clustering coefficient (how well node's neighbours are connected among themselves)
correlates with prevalence: $\rho \approx 0.7$ at ~ 50 days

Small-world networks

Regular

Small-world

Random



$p = 0$

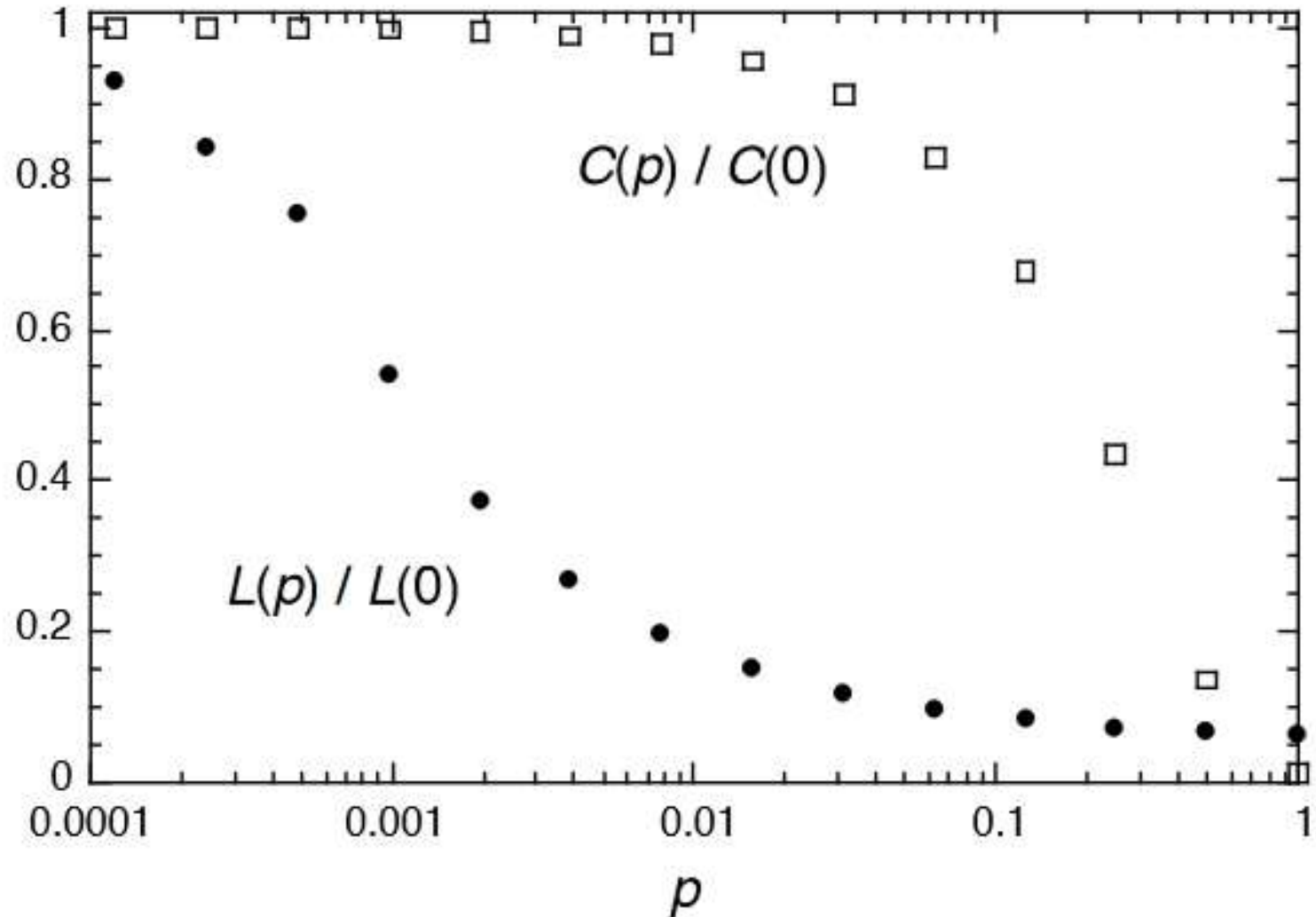
Increasing randomness



$p = 1$

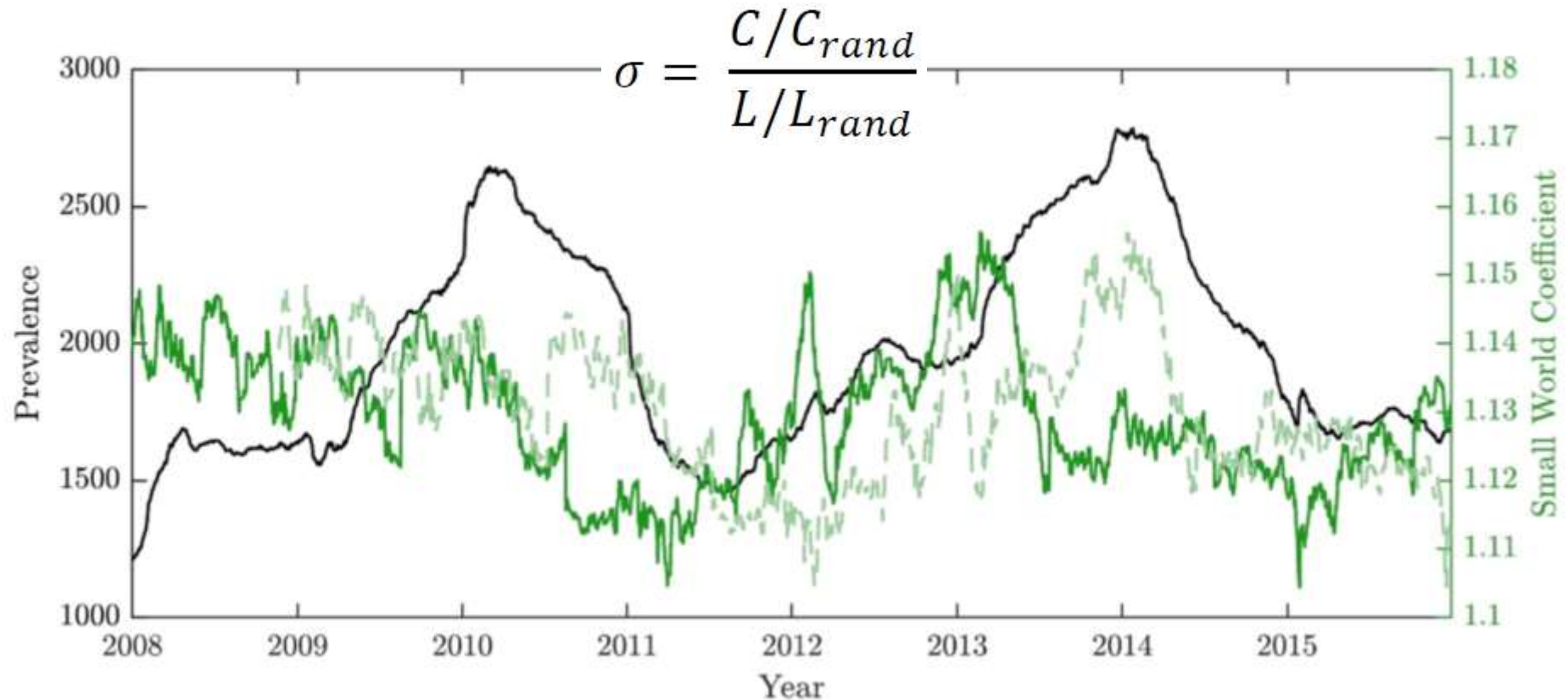


Small-world networks





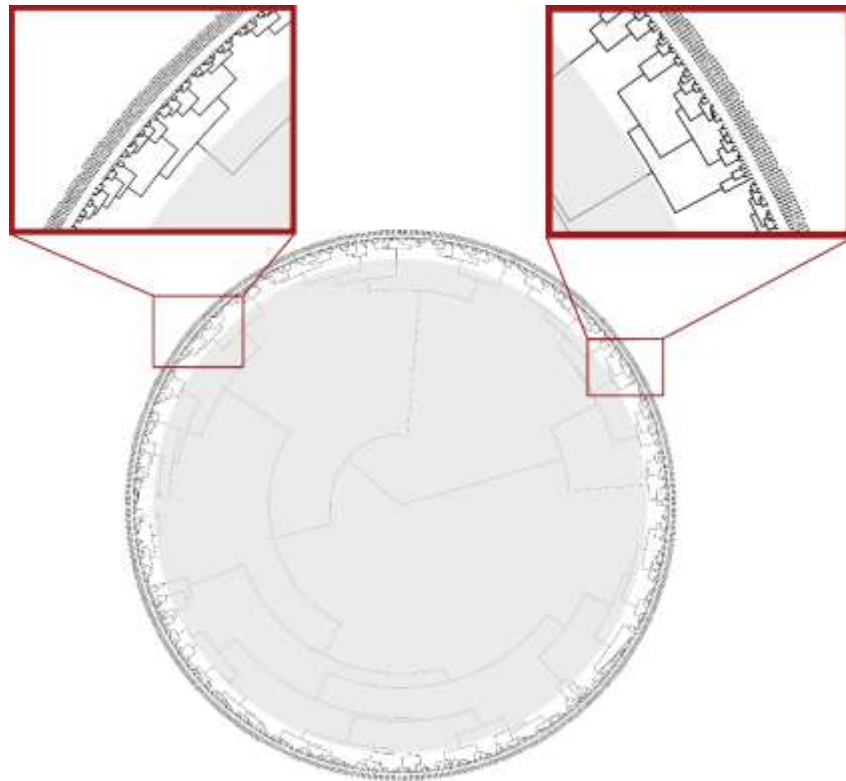
Global network properties: small-world coefficient



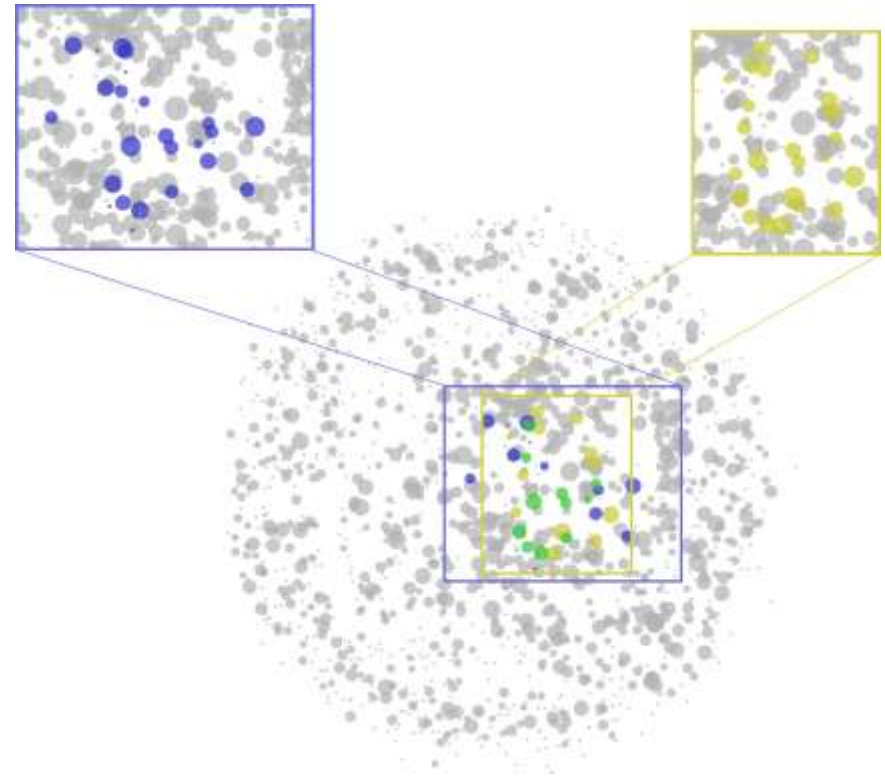
Small-world coefficient (ratio of clustering coefficient to path length)

correlates with prevalence: $\rho \approx 0.6$ at ~ 300 days

Linkage (partitioning) approach

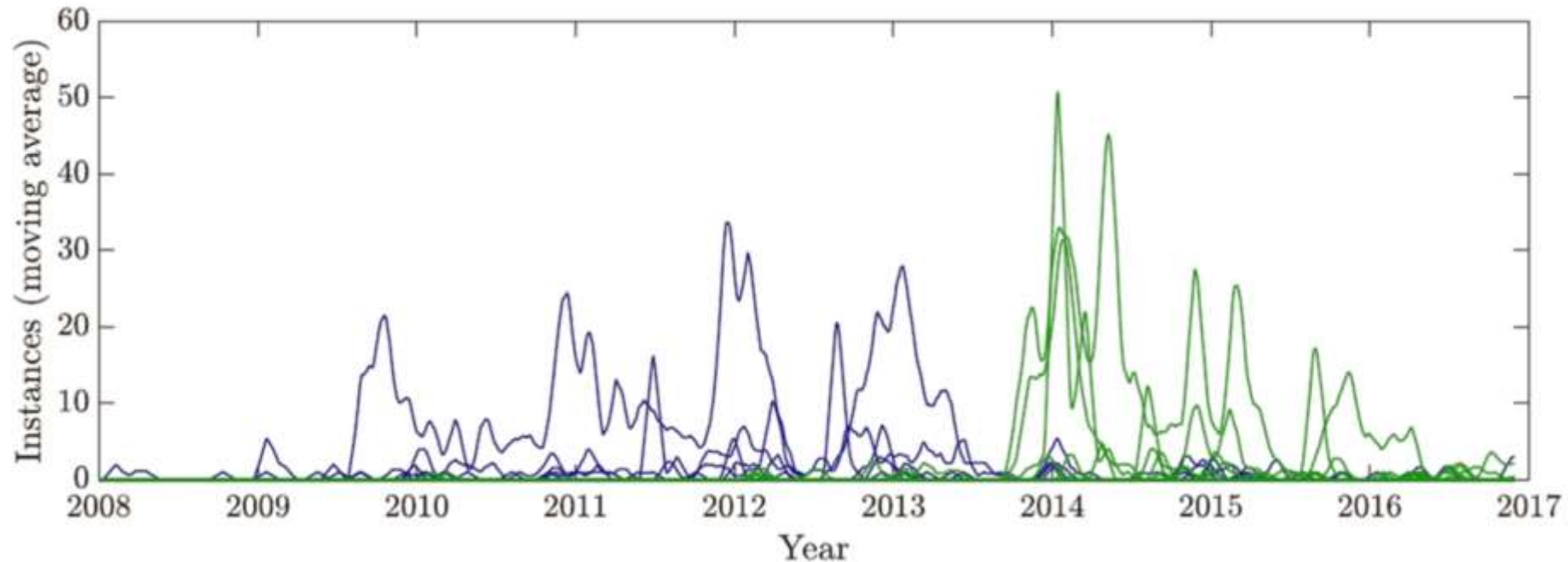


Overlapping approach



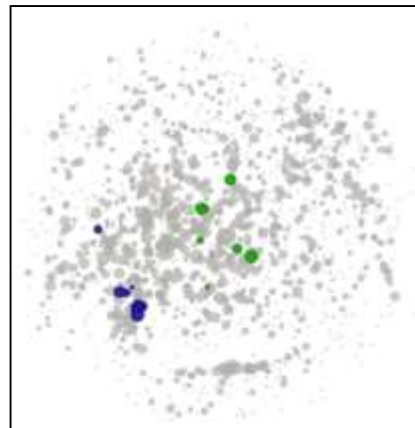


Temporal evolution of clusters



- 3-9-8-13-523
- 3-9-8-14-523
- 3-9-9-13-523
- 3-9-9-13-525
- 3-9-9-14-523

- 3-9-9-14-523
- 3-9-9-15-523
- 3-10-8-13-523
- 3-10-8-14-523
- 3-10-9-13-523
- 3-10-9-14-523

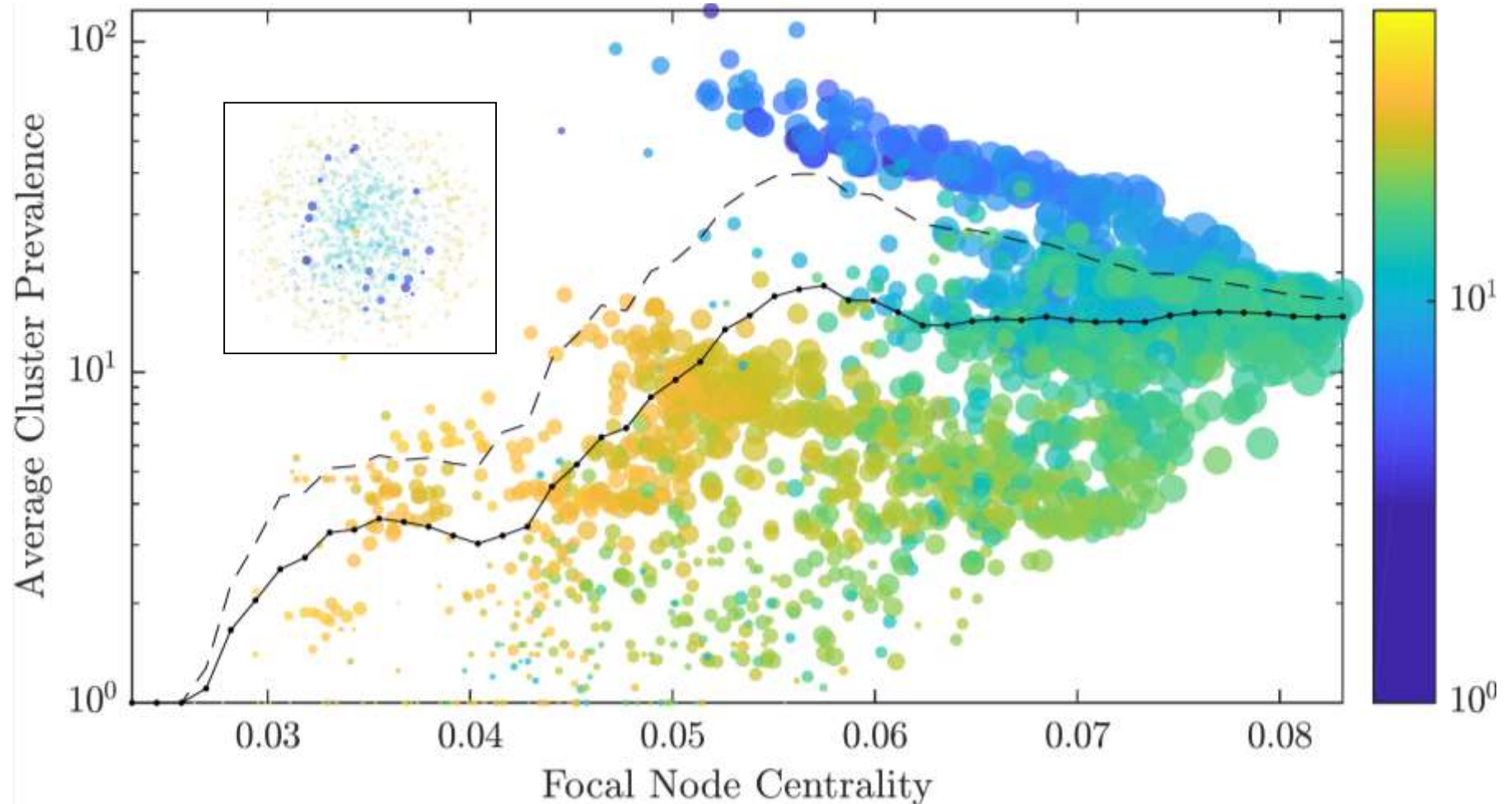


- 3-16-9-11-523
- 3-16-10-10-523
- 3-16-10-11-523
- 3-17-9-11-523
- 3-17-10-11-523
- 4-16-10-11-517

- 2-16-9-11-523
- 3-15-9-10-496
- 3-15-9-10-523
- 3-15-9-11-523
- 3-16-9-10-496
- 3-16-9-10-523



Prevalence-Centrality plot (overlapping clusters)



Colours denote distance to the node with the highest average cluster prevalence

- inferred undirected STM networks from surveillance and molecular genotyping data
- quantified diversity and variability of evolving STM networks
- correlated small-world network properties with the epidemic severity
- identified two distinct evolutionary branches in terms of centrality
- suggested to monitor ongoing STM population diversity and focus on new genotypes as reservoirs from which future epidemics might emerge



entropy

an Open Access Journal

IMPACT
FACTOR
2.419

New Advances in Biocomplexity

Guest Editor

Prof. Mikhail Prokopenko

Deadline

31 September 2020

Special Issue

mdpi.com/si/28540

Invitation to submit

- S. Cauchemez, A. Bhattarai, T. L. Marchbanks, R. P. Fagan, S. Ostroff, N. M. Ferguson, D. Swerdlow; Pennsylvania H1N1 Working Group, Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza, *PNAS*, 108, 2825–2830, 2011.
- O. M. Cliff, V. Sintchenko, T. C. Sorrell, K. Vadlamudi, N. McLean, M. Prokopenko, Network properties of Salmonella epidemics, *Scientific Reports*, 9, 6159, 2019.
- O. M. Cliff, N. Harding, M. Piraveenan, E. Erten, M. Gambhir, M. Prokopenko, Investigating spatiotemporal dynamics and synchrony of influenza epidemics in Australia: An agent-based modelling approach, *Simulation Modelling Practice and Theory*, 87, 412-431, 2018.
- K. M. Fair, C. Zachreson, M. Prokopenko, Creating a surrogate commuter network from Australian Bureau of Statistics census data, *Scientific Data*, 6, 150, 2019.
- T. C. Germann, K. Kadau, I. M. Longini Jr., C. A. Macken, Mitigation strategies for pandemic influenza in the United States, *PNAS*, 103, 5935–5940, 2006.
- N. Harding, R. E. Spinney, M. Prokopenko, Phase transitions in spatial connectivity during influenza pandemics, *Entropy*, 22(2), 133, 2020.
- K. Kadau, T. C. Germann, P. Lomdahl, Large-Scale Molecular-Dynamics Simulation of 19 Billion Particles, *International Journal of Modern Physics C*, 27(15): 193-201, 2004.
- C. Viboud, O. N. Bjørnstad, D. L. Smith, L. Simonsen, M. A. Miller, B.T. Grenfell, Synchrony, waves, and spatial hierarchies in the spread of influenza, *Science*, 312(5772): 447–451, 2006.
- C. Zachreson, K. M. Fair, O. M. Cliff, N. Harding, M. Piraveenan, M. Prokopenko, Urbanization affects peak timing, prevalence, and bimodality of influenza pandemics in Australia: Results of a census-calibrated model, *Science Advances*, 4(12), eaau5294, 2018.