Teaching Systems Thinking in Applied Sciences: Should Students Build Their Own Models?
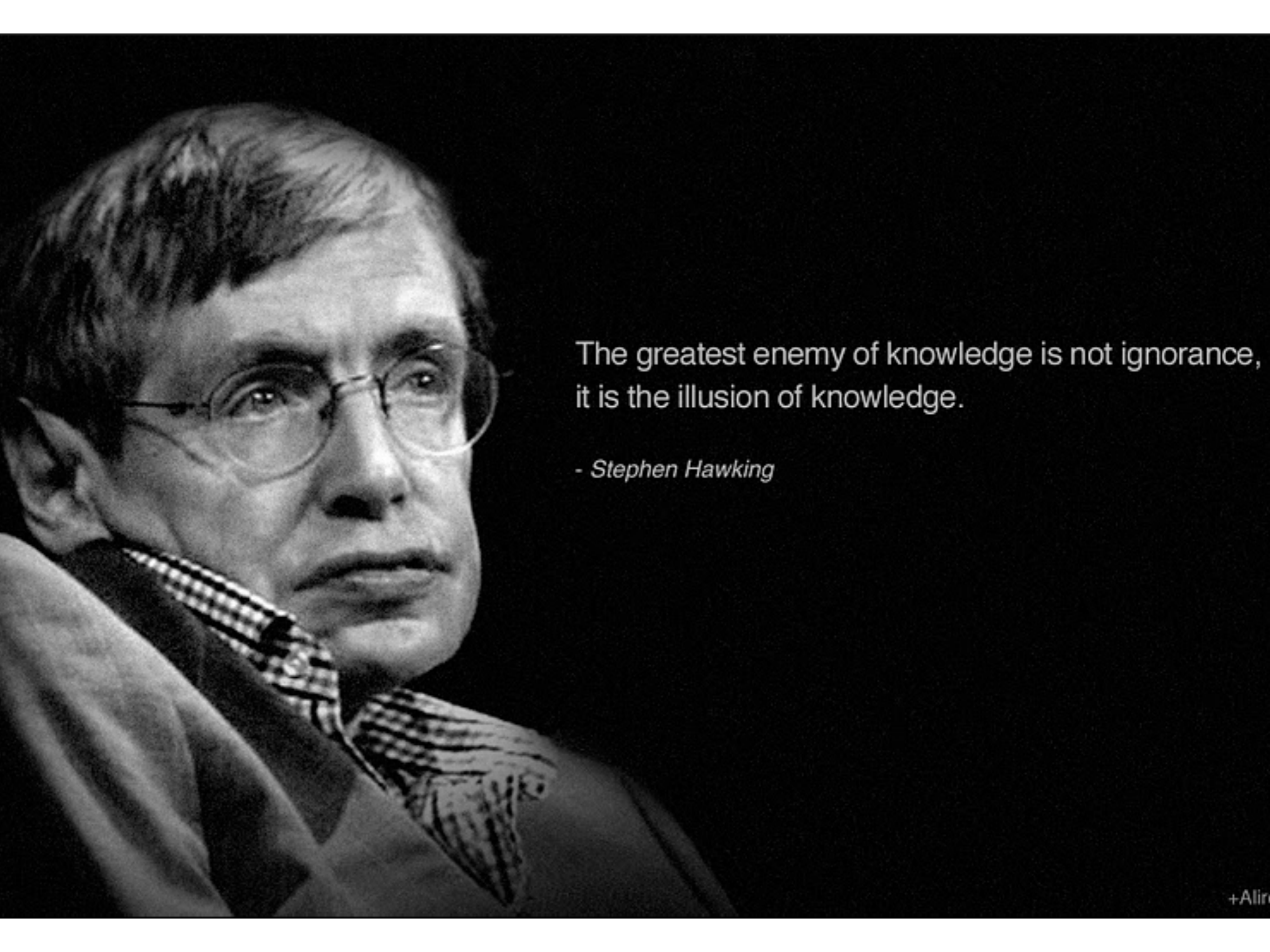
Wayne M. Getz
University of California at Berkeley

Berkeley CNR

The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge.

- Stephen Hawking

Catalyzing Inquiry at the Interface of **Computing** and **Biology**

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

A current challenge

5 Computational Modeling and Simulation as Enablers for Biological Discovery

While the previous chapter deals with the ways in which computers and algorithms could support existing practices of biological research, this chapter introduces a different type of opportunity. The quantities and scopes of data being collected are now far beyond the capability of any human, or team of humans, to analyze. And as the sizes of the datasets continue to increase exponentially, even existing techniques such as statistical analysis begin to suffer. In this data-rich environment, the discovery of large-scale patterns and correlations is potentially of enormous significance. Indeed, such discoveries can be regarded as hypotheses asserting that the pattern or correlation may be important—a mode of "discovery science" that complements the traditional mode of science in which a hypothesis is generated by human beings and then tested empirically.

For exploring this data-rich environment, simulations and computer-driven models of biological systems are proving to be essential.
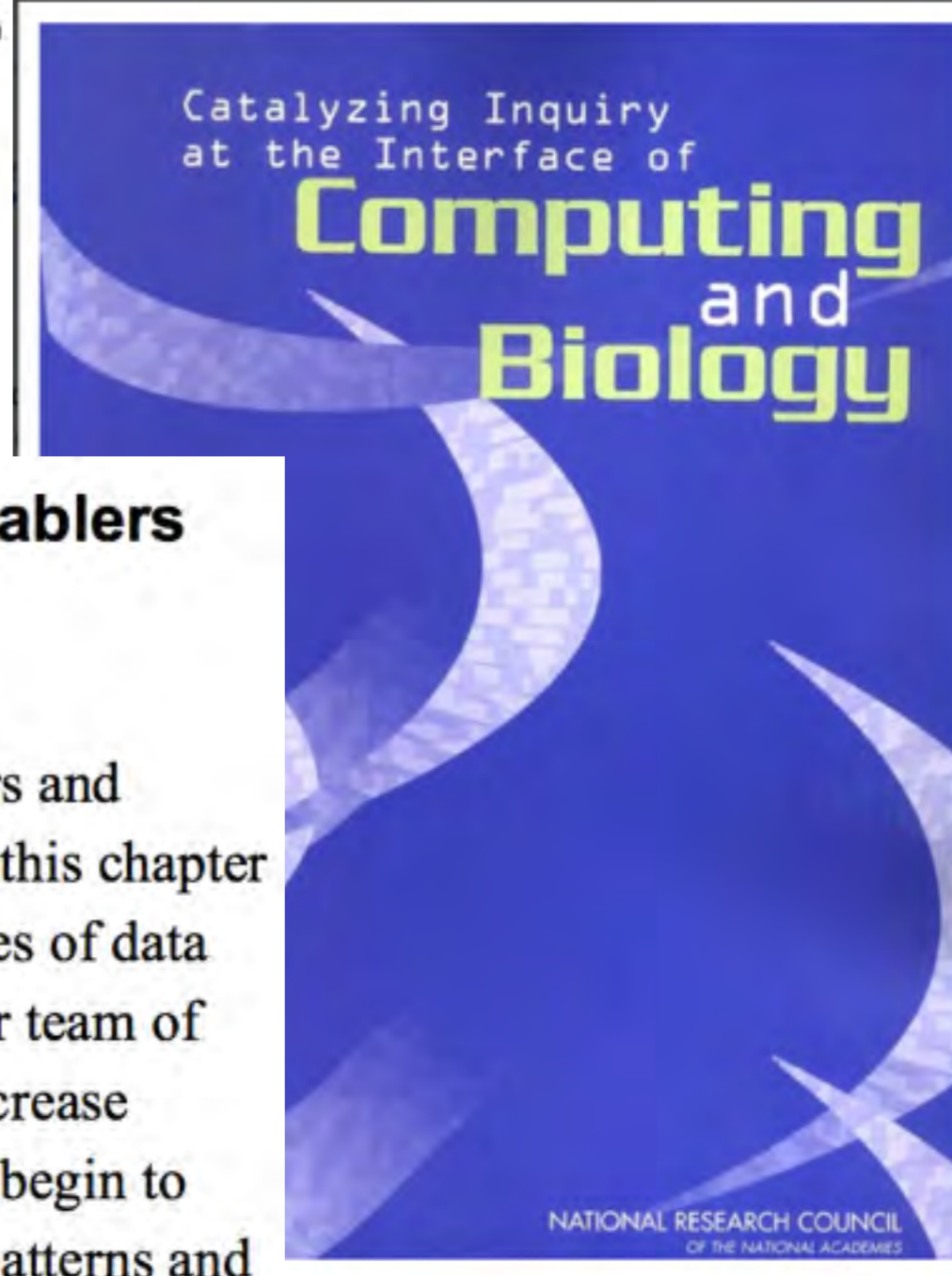
Why build models of systems

12 points from NRC text:

1. Provide coherent framework
2. Highlight basic concepts
3. Exploration uncovers new phenomena
4. Identify key factors/components
5. Link across scales (e.g. individual & populations)
6. Formalize intuitive understanding
7. Screen hypothesis
8. Inform experimental design
9. Predict values that cannot be easily measured
10. Link what is known to what is yet unknown
11. Generate accurate predictions
12. Expand range of questions that can be meaningfully asked

The conundrum

The utility of software applied to big data problems, such as identifying patterns, classifying or a categorizing objects tagged by data, is self-evident through the results produced: **either something useful emerges or the software fails to produce good results (weak vs strong methods)**

The validity of software used to evaluate the impacts of management actions on complex ecosystems is very hard to verify prior to such actions being implemented: **once implement actual outcomes can be compared to predicted outcomes, but this is often a *one-point comparison* with experiments hard to repeat and time horizons are generally large** (some times multigenerational) **(good vs bad models)**

Example of a platform that could work very well without the user knowing "the physics"

ZDOCK SERVER

ZDOCK   M-ZDOCK   Help   Links   References

Input Protein 1   PDB ID ÷   4gkz

HA1.7, a MHC class II restricted TCR specific for haemagglutinin

● Use biological assembly

○ Select chains manually   MOL1(**Alpha chain of Class II TCR**): A
                            MOL2(**Beta Chain of Class II TCR**): B

Input Protein 2   PDB ID ÷

Enter your email:

Optional:
Select ZDOCK version   ZDOCK 3.0.2 ÷
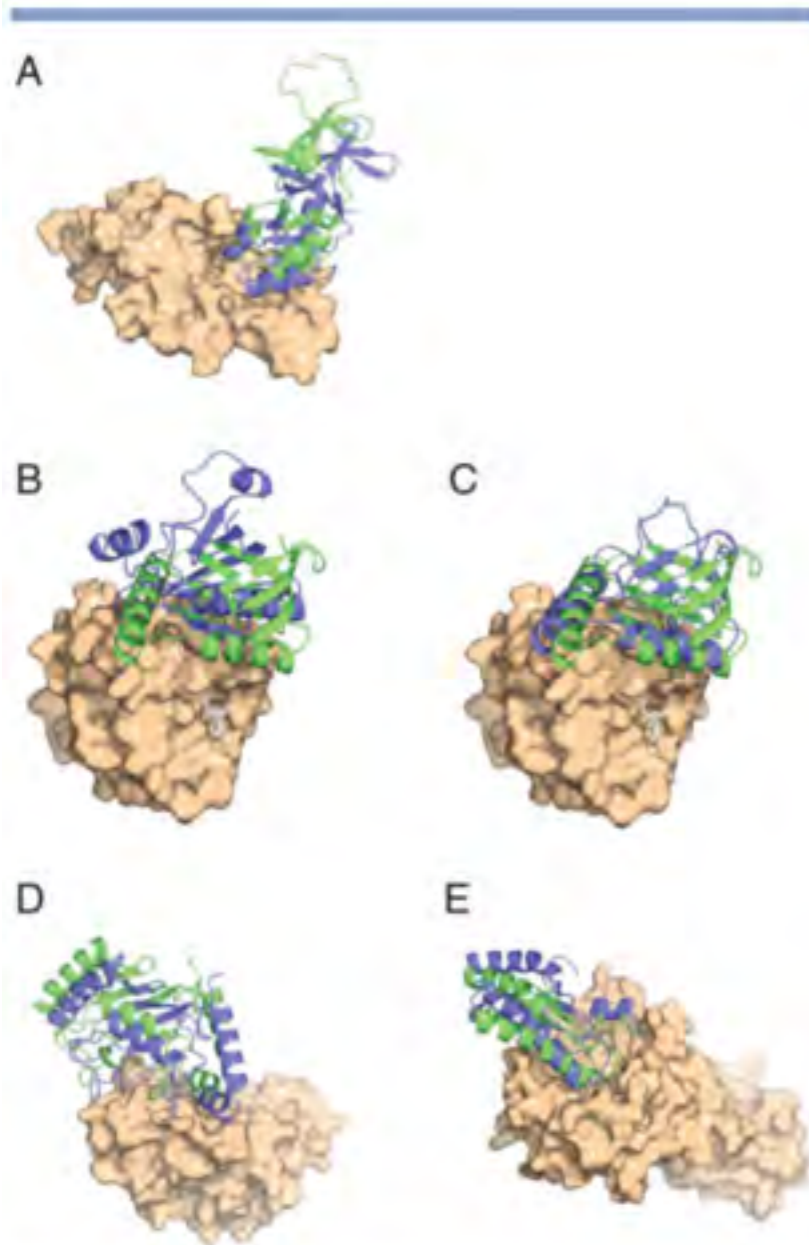Skip residue selection ☐

Figure 1

Docking predictions for Targets 21–27 of CAPRI Rounds 7–11. Orientations of predicted (blue) and crystal structure (green) ligands after superposition of their receptors (beige, only crystal structure receptor shown). (A) Target 21: Orc1p (receptor) and Sir1p (ligand); (B) Target 24: Arf1 (receptor) and homology model of ArfBD (ligand); (C) Target 25: Arf1 (receptor) and bound ArfBD (ligand); (D) Target 26: TolB (receptor) and Pal (ligand); (E) Target 27: Hip2 (receptor) and Ubc9 (ligand). Figures were created using PyMOL.

University of Massachusetts Medical School **UMASS**

Protein Docking

Home

Software

Benchmark

Performance

Decoy Sets

ZLAB

Docking sites

References

Contact

Among the many recent advances in bioinformatics is the rapid accumulation of 3D structures of protein complexes through X-ray crystallization. These structures provide many insights on protein-protein interactions, facilitating rational drug development and the treatment of disease. However, not all protein complexes have been crystallized, so various computational techniques have been

Example of a platform that provides a knowledgeable user with a very useful tool

What is Cytoscape?

Cytoscape is an open source software platform for **visualizing** molecular interaction networks and biological pathways and **integrating** these networks with annotations, gene expression profiles and other state data. Although Cytoscape was originally designed for biological research, now it is a general platform for complex network analysis and visualization. Cytoscape *core* distribution provides a basic set of features for data integration, analysis, and visualization. Additional features are available as **Apps** (formerly called *Plugins*). Apps are available for network and molecular profiling analyses, new layouts, additional file format support, scripting, and connection with databases. They may be developed by anyone using the Cytoscape open API based on Java™ technology and App community development is encouraged. Most of the Apps are freely available from Cytoscape App Store.

Example of a platform that may be very useful but could be dangerous in the hands of an unqualified user

Ecopath with Ecosim
No fish is an island

About Consortium | Publications | Models | Downloads | EwE life cycle | Support | Co-development | Courses | Conferences

Search this site:
[Search]

Follow us

Donate

Home

Ecopath with Ecosim (EwE) is a free ecological/ecosystem modeling software suite. EwE has three main components: *Ecopath* – a static, mass-balanced snapshot of the system; *Ecosim* – a time dynamic simulation module for policy exploration; and *Ecospace* – a spatial and temporal dynamic module primarily designed for exploring impact and placement of protected areas. The Ecopath software package can be used to

- Address ecological questions;
- Evaluate ecosystem effects of fishing;
- Explore management policy options;
- Analyze impact and placement of marine protected areas;
- Predict movement and accumulation of contaminants and tracers (Ecotracer);
- Model effect of environmental changes;
- Facilitate end-to-end model construction.

The latest release of Ecopath with Ecosim is version 6.4.3, released in October 2014.

News

- Post-doc position Spatial Marine Ecosystem Modeling at Stockholm University, deadling March 20, 2015
- Three new Consortium members
- 18 Nov: Ecopath developer back-end back alive
- 17 Nov: Ecopath developer back-end is

Preparing applied scientists to use models as an aid to managing complex environmental and ecological systems

Two examples:

- **California Department of Fisheries and Wildlife (CDFW) are charged with recommending dam release schedules to protect fall run Chinook salmon in the San Joaquin River system**

- **WHO/CDC/US Military healthcare professionals charged with assisting in-country management of the outbreak of Ebola Viral Disease in West Africa**

SalSim – Salmon Simulator (Version 2.0, June 11 2013)

SALSIM is a total life-history population simulation model for fall-run Chinook salmon originating from the San Joaquin River (SJR) and its salmon bearing tributaries. To use the model, specify the simulation settings and then press the "Run Simulation" button. Full model documentation is available for download.

Control Panel

General Settings   Load Saved Scenario

Run ID: [Baseline]   (Optional. Limits: 8 characters maximum; only letters and numbers. [About])

Simulation Time Span

Simulation Data is available from Brood Year 1994 to Escapement Year 2010. Specify a range of years to simulate by changing length of the simulation in years. The simulation begins with escapement in late summer and fall.

Start Brood Year: [1994]

Number of Years: [16]

Operations Management

You may either upload a temperature and flow scenario or configure the operation of reservoirs along the Stanislaus, Tuolumne and Merced rivers.

[Configure Flow Regime]   [Upload HEC-5Q Data]   [Configure Hatchery]

Run Simulation

SALSIM Documentation | Watch a video about Chinook Salmon

© 2011-2013
CALIFORNIA DEPARTMENT OF FISH and WILDLIFE

Running your simulation...

Expect roughly 2 to 4 minutes to perform the simulation.

Playing back the simulation results

The Ebola Crisis

DONATE NOW
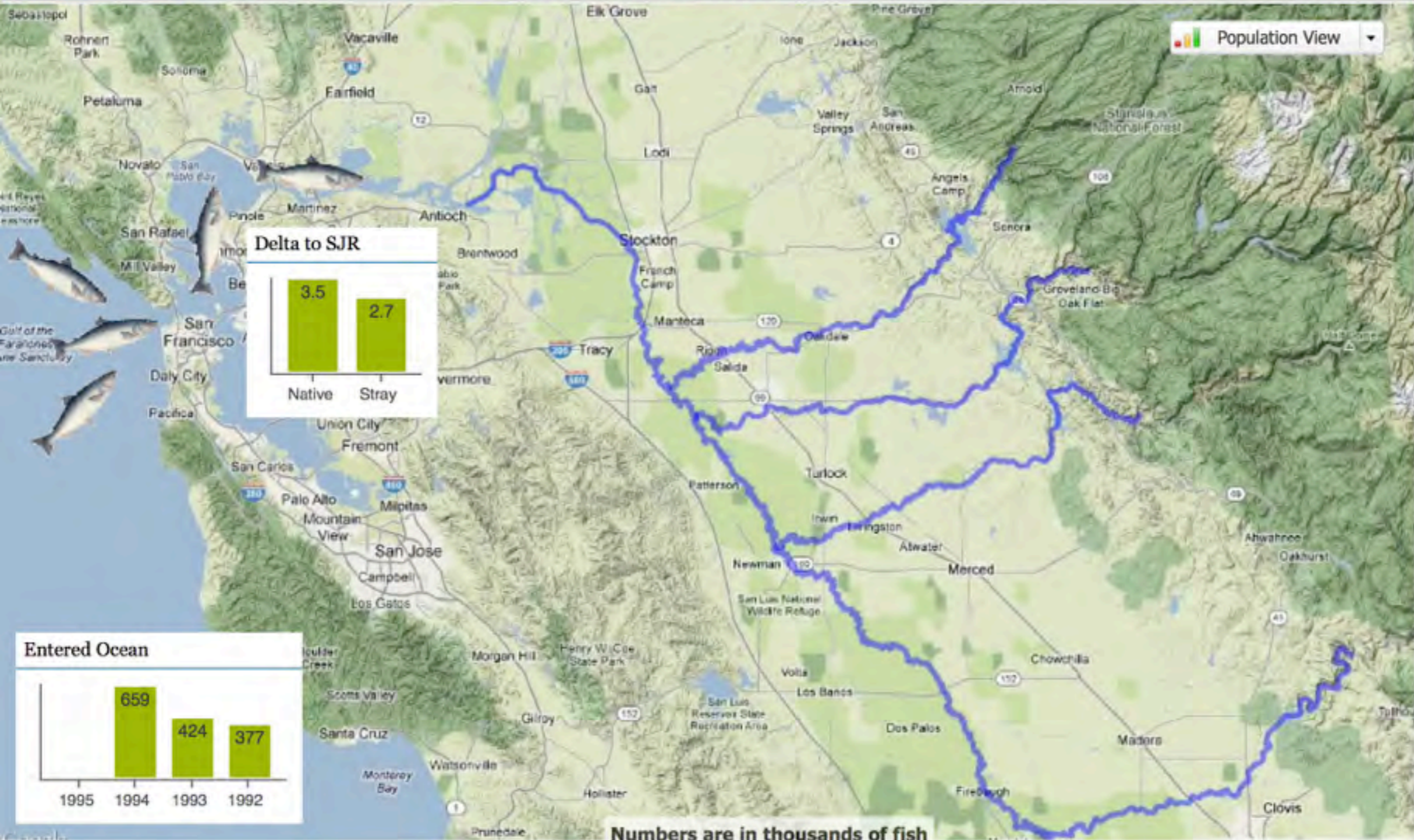
CONNECT WITH MSH

Stronger health systems. Greater health impact.

HOME   ABOUT US   OUR WORK   RESOURCES   NEWS & EVENTS   PARTNER WITH US   BLOG

EBOLA

The current Ebola epidemic in West Africa is unprecedented in its scope. The health systems of Liberia, Guinea, and Sierra Leone were not adequate to the task of containing the outbreak. Thousands of people have been infected and thousands have lost their lives. We are applying our over 40 years of expertise in fragile nations to help restore essential health services, rebuild and strengthen health systems, and prevent future epidemics.

Media Center   Research & Publications   Our Catholic Values   Connect   CRSEspañol.org   DONATE

CRS faith. action. results.
CATHOLIC RELIEF SERVICES

ABOUT   OUR WORK OVERSEAS   WAYS TO GIVE   GET INVOLVED

EBOLA OUTBREAK: CALL FOR ROBUST RESPONSE

- By September 30, 2014, CDC estimates that there will be approximately 8,000 cases, or as high as 21,000 cases if corre made.

  ○ Without additional interventions or changes in community behavior, CDC estimates that by January 20, 2015, the approximately 550,000 Ebola cases in Liberia and Sierra Leone or 1.4 million if corrections for underreporting are

- Cases in Liberia are currently doubling every 15-20 days and those in Sierra Leone and Guinea are doubling every 30-4

CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

SEARCH 🔍

CDC A-Z INDEX ∨

Ebola (Ebola Virus Disease)

Ebola (Ebola Virus Disease)

About Ebola +

2014 West Africa Outbreak −

What's New Timeline

Case Counts +

Q&A: 2014 Ebola Outbreak

Communication Resources +
for West African Audiences

What CDC is Doing +

Q&A: Estimating the
Future Number of Cases
(Liberia and Sierra Leone)

Outbreaks +

Signs and Symptoms

Transmission +

CDC > Ebola (Ebola Virus Disease) > 2014 West Africa Outbreak

Questions and Answers: Estimating the Future Number of Cases in the Ebola Epidemic—Liberia and Sierra Leone, 2014–2015

Language: English ∨

Summary

The September 26, 2014, Morbidity and Mortality Weekly Report (MMWR), Estimating the Future Number of Cases in the Ebola Epidemic—Liberia and Sierra Leone, 2014–2015, estimates the future number of cases if current trends continue. The MMWR also adjusts the number of cases based on estimated underreported cases.

- By September 30, 2014, CDC estimates that there will be approximately 8,000 cases, or as high as 21,000 cases if corrections for underreporting are made.
  ○ Without additional interventions or changes in community behavior, CDC estimates that by January 20, 2015, there will be a total of approximately 550,000 Ebola cases in Liberia and Sierra Leone or 1.4 million if corrections for underreporting are made.
- Cases in Liberia are currently doubling every 15-20 days, and those in Sierra Leone and Guinea are doubling every 30-40 days.
- Halting the epidemic requires that approximately 70% of Ebola cases be cared for in Ebola Treatment Units or, if they are at capacity, at home or in a community setting in which there is a reduced risk of disease transmission and safe burials are provided.

EVD outbreak in West Africa still ongoing but, thankfully, numbers are much, much lower than predicted one year ago

Date	Total		Guinea		Liberia		Sierra Leone	
	Cases	Deaths	Cases	Deaths	Cases	Deaths	Cases	Deaths
13 Sept 2015	28,220	11,291	3,792	2,530	10,672	4,808	13,756	3,953
6 Sept 2015	28,147	11,291	3,792	2,530	10,672	4,808	13,683	3,953
30 Aug 2015	28,073	11,290	3,792	2,529	10,672	4,808	13,609	3,953
16 Aug 2015	27,952	11,284	3,786	2,524	10,672	4,808	13,494	3,952
26 July 2015	27,748	11,279	3,786	2,520	10,672	4,808	13,290	3,951
28 June 2015	27,514	1,220	3,729	2,482	10,666	4,806	13,119	3,932

Date	Total	
	Cases	Deaths
13 Sept 2015	28,220	11,291

Research Article

Tactics and Strategies for Managing Ebola Outbreaks and the Salience of Immunization

Wayne M. Getz,[1,2] Jean-Paul Gonzalez,[3] Richard Salter,[4] James Bangura,[5] Colin Carlson,[1] Moinya Coomber,[6] Eric Dougherty,[1] David Kargbo,[7] Nathan D. Wolfe,[3] and Nadia Wauquier[6,8]
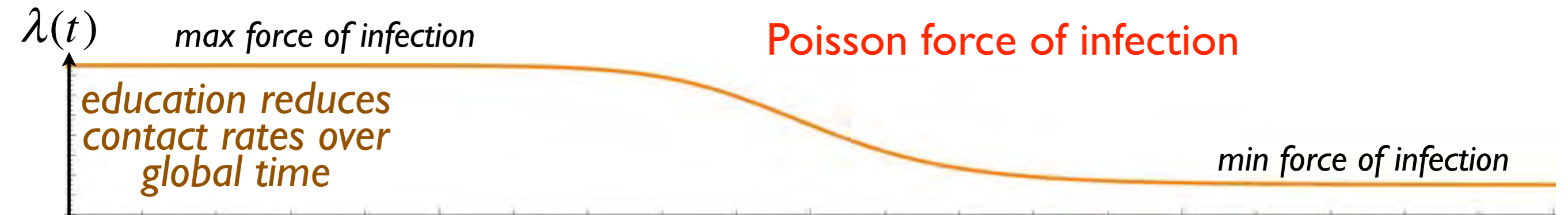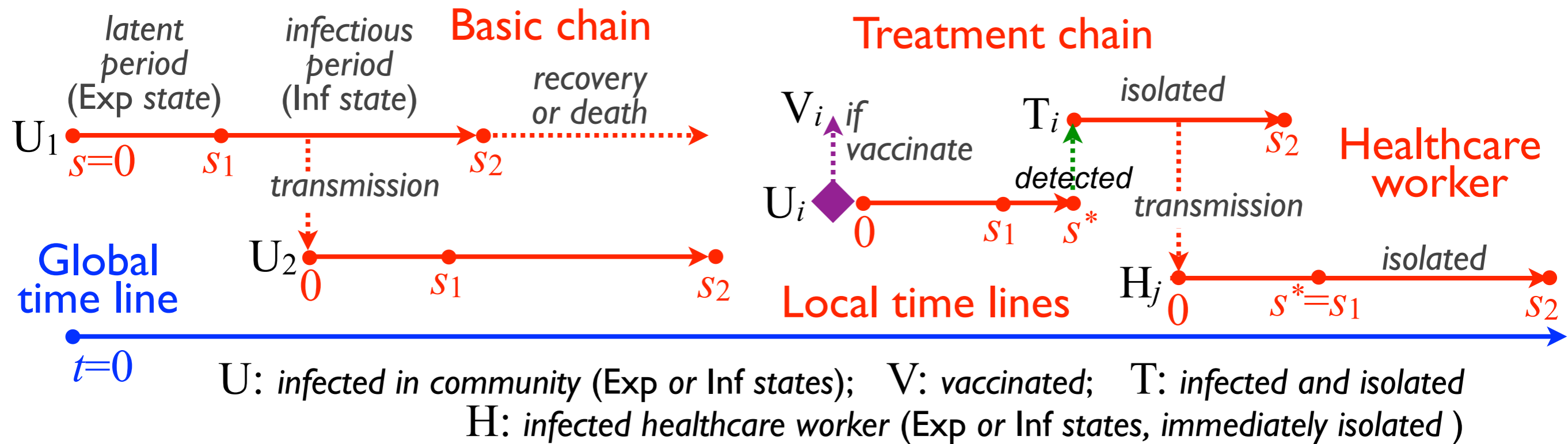
Simulation
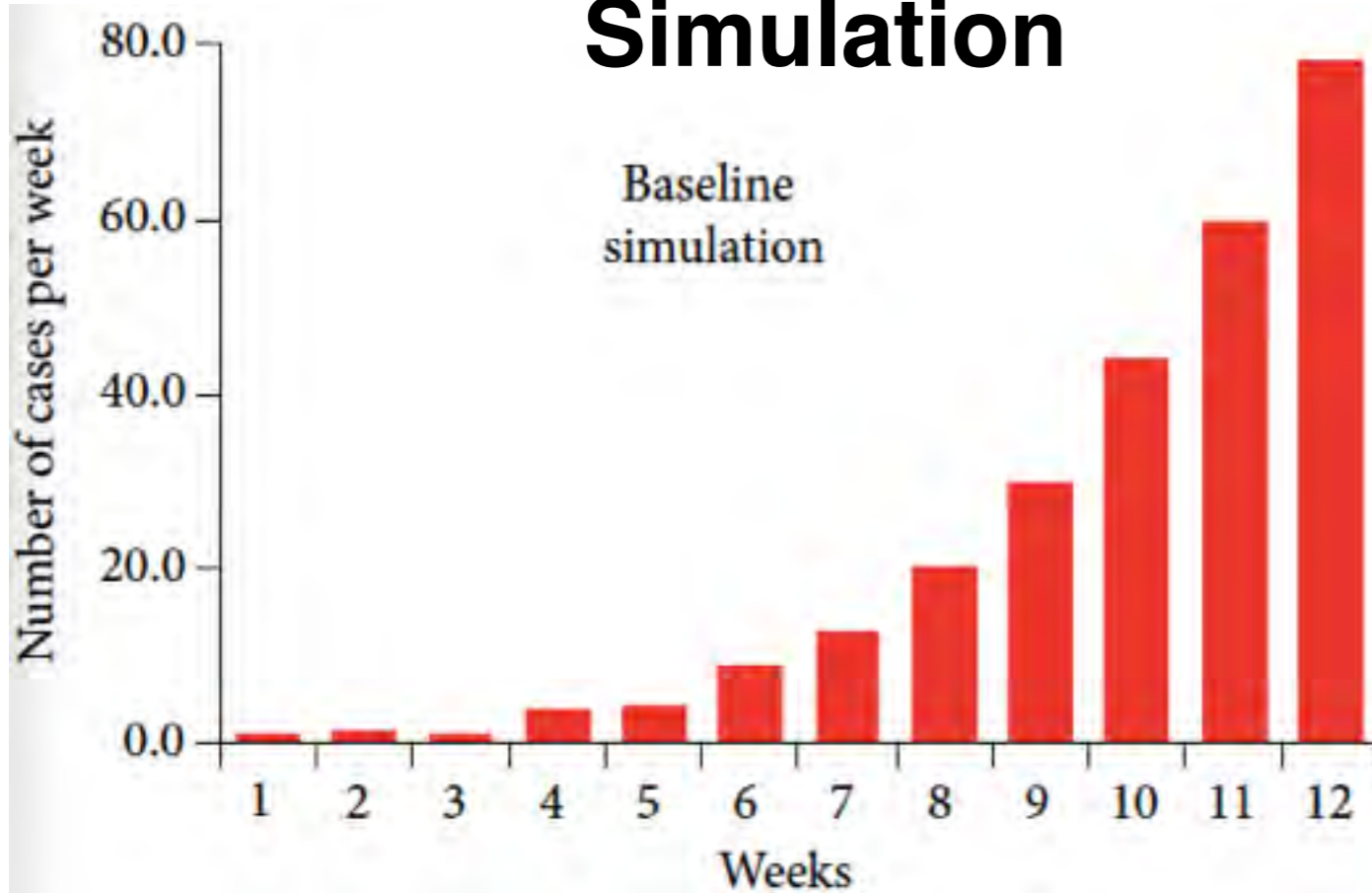
Data (cases per week)
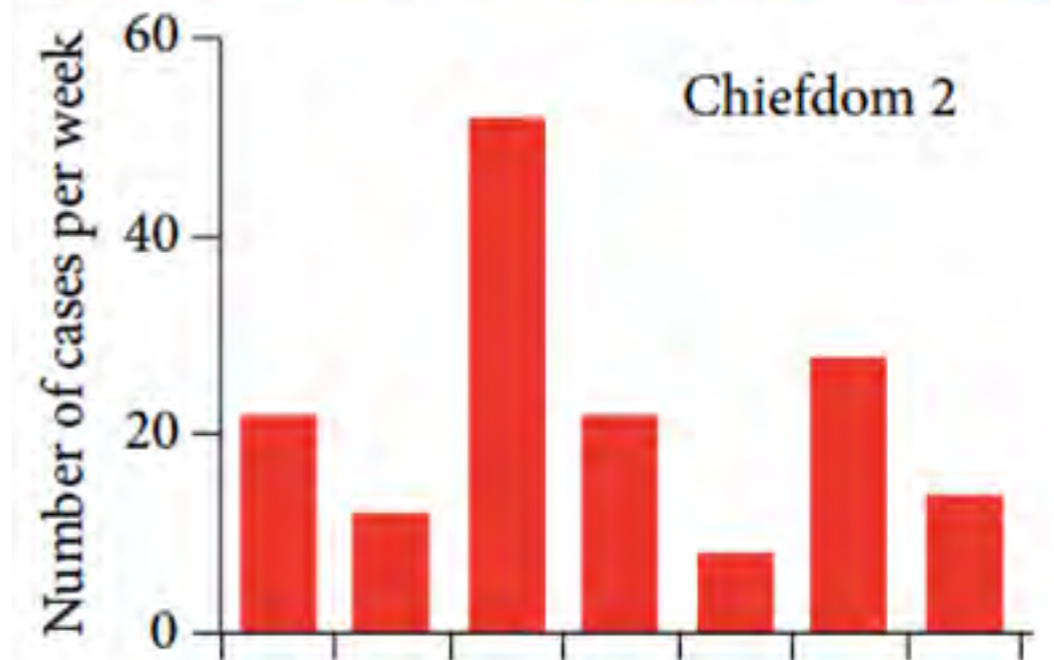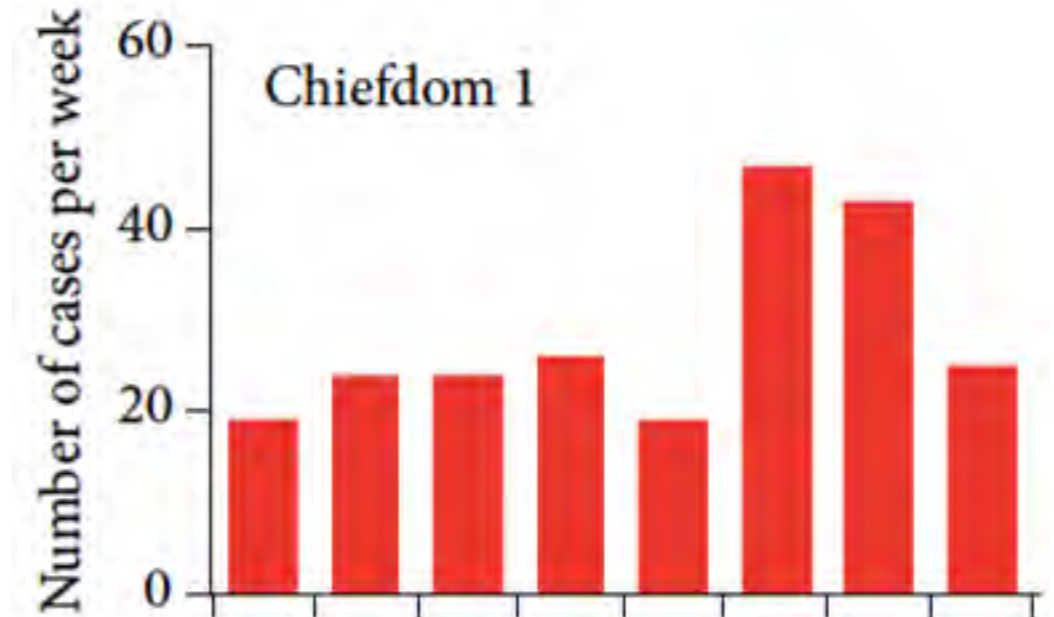
The data do not show a homogeneous outbreak pattern, but suggest the existence of considerable spatial inhomogeneity
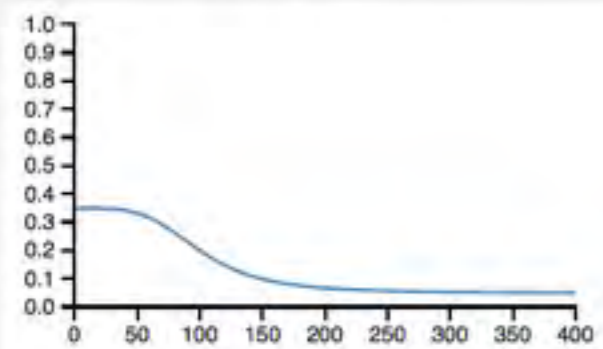
Ebola: transmission chain model

Basic chain

latent period (Exp state) *infectious period (Inf state)* *recovery or death*

U_1 $s=0$ s_1 s_2

transmission

U_2 0 s_1 s_2

Global time line

Local time lines

Treatment chain

V_i *if vaccinate* T_i *isolated* s_2 **Healthcare worker**

U_i 0 s_1 *detected* s^* *transmission*

H_j 0 $s^*=s_1$ *isolated* s_2

Global time line

$t=0$

U: *infected in community* (Exp *or* Inf *states*);   V: *vaccinated*;   T: *infected and isolated*
H: *infected healthcare worker* (Exp *or* Inf *states, immediately isolated*)

$\lambda(t)$ *max force of infection* **Poisson force of infection**

education reduces contact rates over global time

min force of infection

Probability of isolation

$\tau(s,t)$

probability of being isolated increases with global time $s^*(t_1)<s^*(t_2)$ *for* $t_1 > t_2$ *local time isolation starts sooner as global time progress*

Nova implementation: population of "victims"
creation of offspring distribution is critical

Ebola Epidemic Model

Nova online app

Max No. of Days: [slider] 400 | Reset | Run | Stop | Continue | Step | Days: 185 | [toggle] Save Data to File? | Filename: eboladata | About this file

Transmission Probability

Community Learning: [slider] 100

Transmission Min: [slider] 0.05

Transmission Max: [slider] 0.35

Isolation Probability

Detection Response: [slider] 1

Healthcare Gear-up: [slider] 500

Healthcare Transmission: [slider] 0.05

Vaccination Probability

Half Saturation (days): [slider] 100

Vaccination Min: [slider] 0.01

Vaccination Max: [slider] 0.15
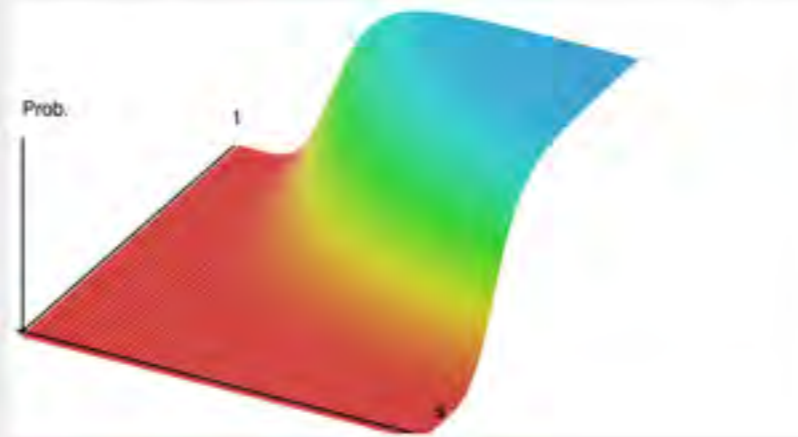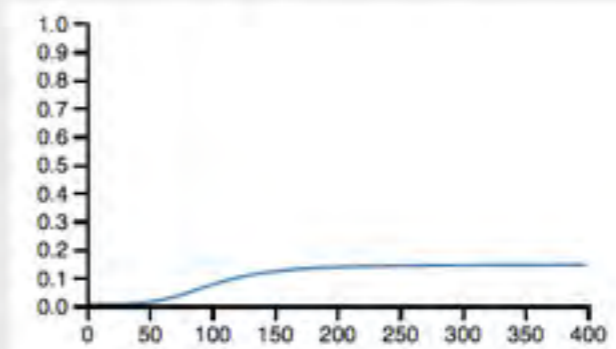
Mortality Rate:
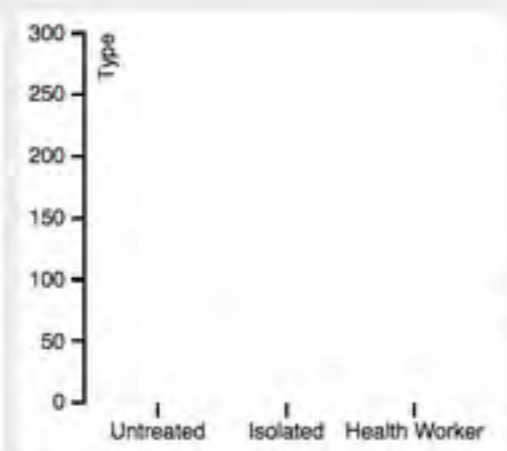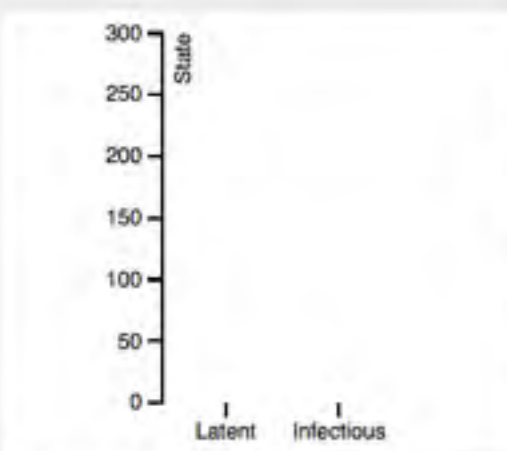
	Infected	Died	Recovered
Total	981	473	508
Untreated	32	12	20
Isolated	795	383	412
Health Worker	154	78	76
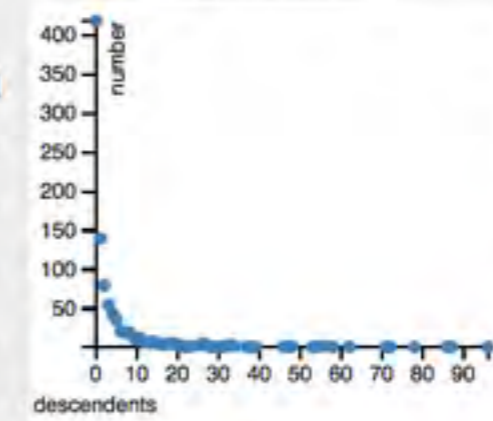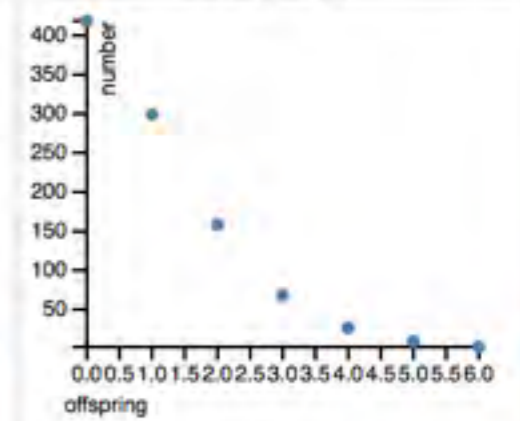
Latent: 0   Infectious: 0   U: 0   I: 0   H: 0

Histograms

Descendents
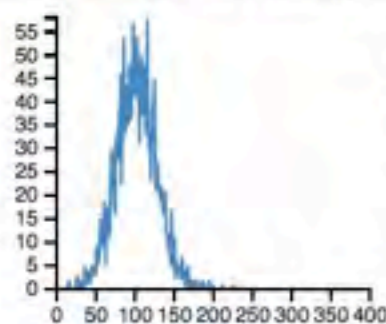

descendents   Save

Offspring

offspring   Save

Aggregated Dynamics

New Infections Per Day
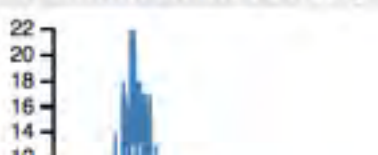
Accum Infections

New Isolations Per Day

Accum Isolations

Daily Isolations

Accum Daily Isolations

| Max No. of Days: | | 400 | Reset | Run | Stop | Continue | Step | Days: | 243 | | Save Data to File? | Filename: eboladata | About this file |

Transmission Probability

Community Learning:		100
Transmission Min:		0.05
Transmission Max:		0.35

Isolation Probability

Detection Response:		1
Healthcare Gear-up:		500
Healthcare Transmission:		0.05

Prob.

Vaccination Probability

Half Saturation (days):		100
Vaccination Min:		0.01
Vaccination Max:		0.15

Mortality Rate:

Histograms

Descendents

number

descendents

Save

Offspring

number

offspring

Save

State

Type
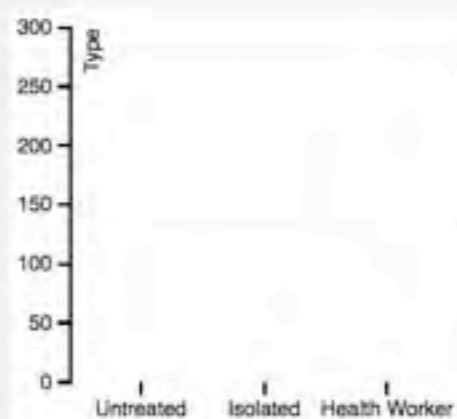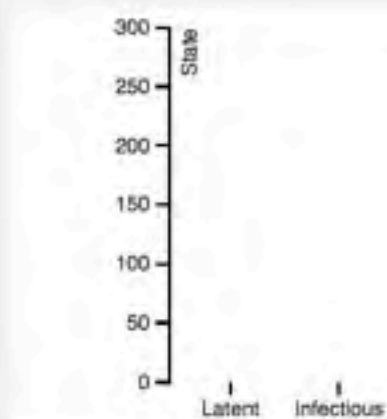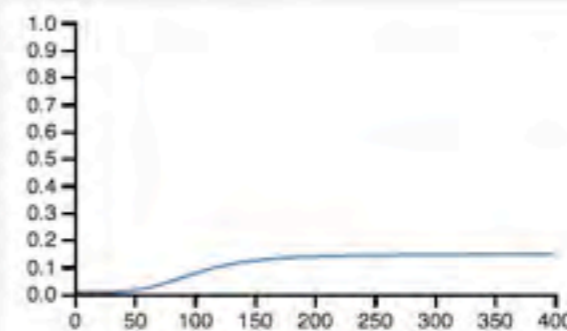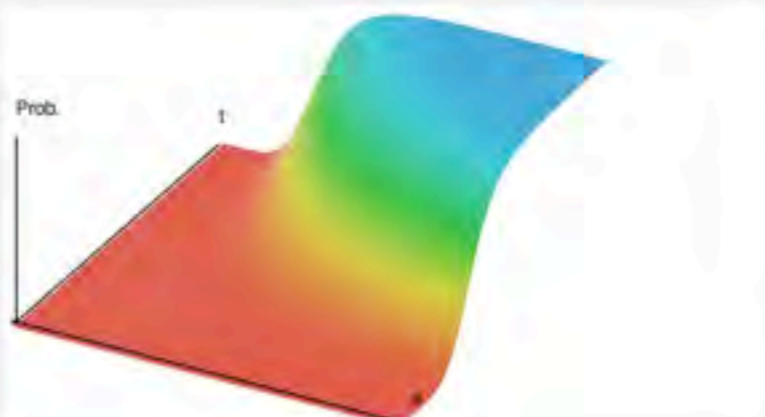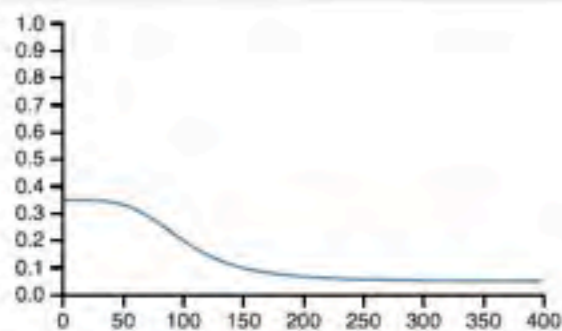
	Infected	Died	Recovered
Total	2937	1446	1491
Untreated	100	53	47
Isolated	2389	1166	1223
Health Worker	448	227	221
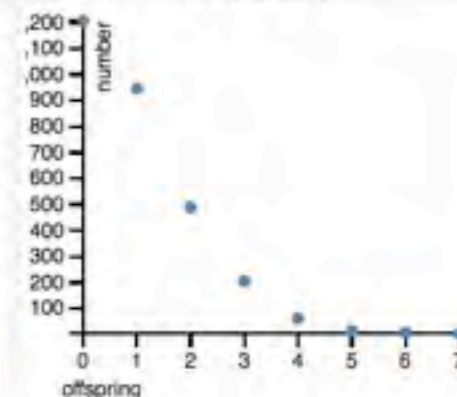
Latent | Infectious

Untreated | Isolated | Health Worker

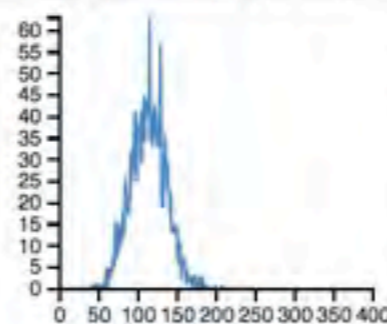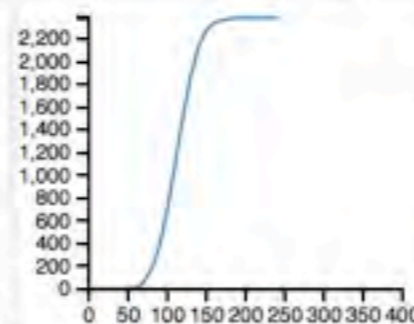Latent: 0   Infectious: 0   U: 0   I: 0   H: 0

Aggregated Dynamics

New Infections Per Day

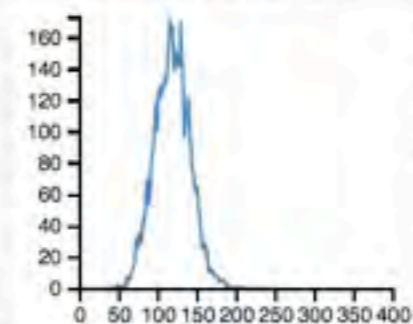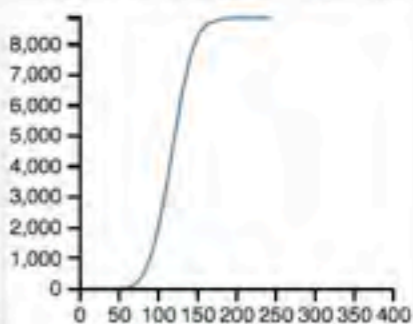### Accum Infections

New Isolations Per Day

Accum Isolations

Daily Isolations

Accum Daily Isolations

More and more, complex systems models implemented as "web apps" are becoming available for general use

Are we thereby enabling Hawking's aphorism?
"the greatest enemy of knowledge is not igorance, it is the illusion of knowledge"

If so, can we mitigate this through education?
Do we need to insist that anyone using some elses model should, at least, have some experience in building their own models of similar, albeit simpler, systems?

My feeling is that model building experience is critical when it comes to using the "canned" models of others!
Maybe we should even require individuals to have
"model user licenses"
equivalent to Masters level degrees in the health sciences

Thanks