

Modelling Vector-borne Disease as a Complex Adaptive System: Challenges and Potential Solutions

Chris Stephens

C3-Centro de Ciencias de la Complejidad y Instituto de Ciencias Nucleares, UNAM Mexico-US Forum Arbovirus Disease Research: Priorities for Collaboration and Partnership: Nov. 28th-30th 2017



The Challenges



They are dynamical and adaptive

Number of cases x 1000

Scenario 2 Scenario 1 160 Scenario 3 Scenario 4 O O Quarantine 120 Interventions ecisions 80 Vaccination program 40 0 10 30 40 50 70 0 20 60 Time in weeks Preventative **Treatment** vector control

We want to predict and understand "histories" but nothing is written in stone. What is the "space" of interventions?







What do we know...? Known cases Known vectors Known hosts Known risk factors





How do we infer what's under the water?



What do we know...? Known cases Known vectors Known hosts Known risk factors **Known interactions** We need new mathods to Integrating multiple sources model these detained ata... and multiple formats Adaptive Systems What don't we know...?

What don't we know...? Unknown cases Unknown vectors Unknown hosts Unknown risk factors **Unknown interactions**



What is the niche of an emerging disease?

What do we want to predict? C - the presence, or abundance, of cases of a disease,... Probability to find disease P(C|X(t))

Probability to find disease given the niche factors **X**

Characterizes niche and "anti-niche" S(C|X) Risk score



What affects it? The "niche" **X** = (X1, X2, X3, ..., XM)

A large part of the complexity is in the multi-factoriality of both C and X. Adaptation is inherent in the fact that P(C | X) can change in time.

 $\mathbf{X} = X(sd) + X(se) + X(n) + X(ev) + X(g) + X(af) + X(hm) + X(i) + X(sp) + \dots$

Macro-Climactic factors

Micro-Climatic factors
Behavioural
characteristics
Phenotypic

Hydrography

Host species

Human activity

Competitor species

Predator species

Problems of co-dependence and causality

characteristics

What is a community of emerging diseases?





The Potential Solutions

One model = one question -> one answer



Where's the Complex Adaptive Systems modelling?

Models heterogeneity, but not naturally data based Google flu, REPORTA, influenzanet,...





Standard SIR models



Figure 3: Structure of the SIR model.



Complication of answering just one question - the number of infected, based on a small number of inputs without/with mosquitoes



Figure 4: Structure of the SIR/SI model.

$$\begin{aligned} \frac{dS_H}{dt}(t) &= \mu_H N_H - \beta_H S_H(t) I_v(t) / N_v - \mu_H S_H(t), \\ \frac{dI_H}{dt}(t) &= \beta_H S_H(t) I_v(t) / N_v - \gamma_H I_H(t) - \mu_H I_H(t) \\ \frac{dR_H}{dt}(t) &= \gamma_H I_H(t) - \mu_H R_H(t) \\ \frac{dI_v}{dt}(t) &= \beta_v S_v(t) I_H(t) / N_H - \mu_v I_v(t), \\ \frac{dS_v}{dt}(t) &= -\beta_v S_v(t) I_H(t) / N_H + \mu_v I_v(t), \\ \frac{dC_i}{dt} &= p\beta_H S_H(t) I_v(t) / N_v, \end{aligned}$$

but without...

mosquito type, hosts (known/unknown), habitat, socio-economic/

socio-demographic factors, climate/weather, previous infection/co-infection (human/vector host), interventions, etc. etc.

But, for multi-factorial solutions we need multi-factorial data

Ecological modelling from a "Data Science" perspective



The Data Revolution has provided a unique opportunity to construct multi-factorial ecological models for emerging diseases. However, most ecological data is spatio-temporal at multiple scales. Spatial data mining is much less developed than standard data mining.

Spatial data

P(C | X(t))

Co-occurrence between arbovirus occurrence, C, and niche variables, X(t) ecological interaction

Problems with data:

Different sources

Different location, data base, access,...

Different data types

categorical, metric, continuous, discrete,...

Different spatial resolution

Explicit – e.g., pixel by pixel in environmental layers Implicit - 30,000,000 data points versus 30 Abiotic versus biotic

- Collection data
- Ecological niche data
- Ecological niche model data Model/niche construction
- Socio-demographic data
- Phenotypic data
- Vegetable and crop cover
- Geographical data
- Medical and public health data...

The data are represented in space and time – spatial data mining

iche descrij

Predictive Model for potential hosts of ZIKV

Mammals with most statistically significant geographic overlap with Aedes Aegypti

Rank	Mammal	epsilon	Rank	Mammal	epsilon		Chri
1	Glossophaga soricina	12.78	38	Dasypus novemcinctus	7.11		Risk man for
2	Molossus rufus	11.99	39	Sigmodon hispidus	7.02		Misk map 101 –
3	Artibeus jamaicensis*	11.68	40	Uroderma bilobatum	6.82		A adapt A agreeti
4	Liomys pictus	11.06	41	Leptonycteris curasoae	6.75		Aedes Aegypti
5	Oryzomys couesi	11.04	42	Carollia perspicillata	6.71		
6	Carollia subrufa	10.49	43	Centurio senex	6.61		trom a biotic
7	Sturnira lilium	10.28	44	Sciurus colliaei	6.59	have by the proved of	•
8	Artibeus lituratus*	9.91	45	Lontra longicaudis	6.49		model for
9	Choeroniscus godmani	9.42	46	Didelphis marsupialis	6.49	Co-occurrence	
10	Liomys salvini	9.33	47	Cratogeomys bulleri	6.35	A DI SOLO -	$\mathbf{D}(\mathbf{C} \mid \mathbf{V})$
11	Oligoryzomys fulvescens	9.15	48	Carollia sowelli*	6.27	measure	$\Gamma(C X)$
12	Dermanura phaeotis	9.12	49	Myotis elegans	6.12	e shi ka a	
13	Rhogeessa tumida	9.06	50	Myotis nigricans*	6.06		
14	Pteronotus personatus	9.05	51	Sigmodon arizonae	6.00	Ligh High	
15	Baiomys musculus	8.97	52	Rhynchonycteris naso	5.95		
16	Glossophaga commissarisi	8.80	53	Tlacuatzin canescens	5.87	2000 00 00 00 00 00 00 00 00 00 00 00 00	
17	Didelphis virginiana	8.58	54	Leopardus pardalis	3.84	Low	
18	Pteronotus parnellu*	8.58	55	<i>Caluromys derbianus</i> Geoportal Leafl	et 5.78		3/ 1/2 1/2 1/2
19	Crinogeomys nispiaus	8.53	50 57	Molossus molossus	5./6	• Localities of A. Aegypti	30
20	Sciurus aureogasier	8.32	57	Oryzomys rostratus	5.70		
21	Molossus sinuloae	8.31	50	Myotis agreeni	5.66	Positives for	⊙ ► 🎇
22	Sacconterve hilineata	8.23	33 60	Myous carteris microtis	5.00		
23	Lasiurus intermedius	8.15	61	Subvilagus brasiliansis	5.52	DENGUE QŢP	
25	Phyllostomus discolor	8.12	62	Sylvilagus floridanus	5 37		
26	Philander opossum	8 10	63	Spermonhilus annulatus	5 36		Complete Information
27	Peromyscus gymnotis	7.90	64	Peromvscus leucopus	5.30		Complex Inference
28	Balantioptervx plicata	7.81	65	Conepatus leuconotus	5.30		
29	Eptesicus furinalis	7.69	66	Chaetodipus pernix	5.27		Network for
30	Pteronotus davyi	7.55	67	Sciurus yucatanensis	5.23		
31	Dermanura tolteca	7.48	68	Sigmodon mascotensis	5.13		Aedes acounti and
32	Sciurus variegatoides	7.48	69	Eira barbara	5.12		ricaes acgypti and
33	Mormoops megalophylla	7.45	70	Ateles geoffroyi	5.11		A a dag alle anistra
34	Oryzomys melanotis	7.42	71	Neotoma phenax	5.07		Aedes aldopictus
35	Artibeus intermedius	7.40	72	Noctilio leporinus	5.06		•
36	Chaetodipus artus	7.20	73	Reithrodontomys fulvescens	4.95		
37	Nasua narica	7.18					

306

De

Predictive Model for bird mortality from WNV



Relation between bird mortality and mosquito genera by geographic co-occurrence between them



Risk map for presence of pathogenic strains of WNV



Risk map for presence of pathogenic strains of WNV in bird conservation areas





To Link Data-Predictions-Experiment The Emerging Disease (Arbovirus) "production line"



Data-Predictions



Bienvenido a la Plataforma de exploración de datos ecológicos del C3 y la CONABIO.



Data-Predictions Test zoonosis - Leishmaniasis



3

- <i>P</i>						(95)
Carollia sowelli	8.83	43	2	45	4.4	-1 - 14
Heteromys gaumeri*	8.8	5	0	5	0	-15 - 29
Peromyscus mexicanus	8.79	115	6	121	5	2 - 11
Heteromys desmarestianus*	8.72	30	0	30	0	-2 - 16
Molossus rufus	8.63	1	0	1	0	-42 - 56
Glossophaga soricina	8.57	19	7	26	26.9	-3 - 16
Carollia perspicillata	8.5	8	0	8	0	-11 - 24
Pteronotus parnellii	8.16	4	0	4	0	-18 - 31
Desmodus rotundus	8.15	13	1	14	7.1	-6 - 20
Autibaua abaaatia	8.05	25	1	05	2.8	1 - 15
Arubeus phuebus	7.73	35	0	30	2.0	-1 - 13
Ototylomys phyllotis*	7.56	9	1	10	10	-20 - 41
Sigmodon hispidus*	7.28	36	4	40	10	-1 - 14
Peromyscus vucatanicus*	7.25	3	0	3	0	-22 - 35
Didelphis virginiana	7.12	3	0	3	0	-22 - 30
Didelphis marsupialis	6.44	11	0	11	0	-8 - 21
Philander opossum	6.25	6	1	7	14.3	-12 - 2
Centurio senex	6.01	1	0	1	0	-42 - :
Artibeus jamaicensis	5.98	81	5	86	5.8	1 - 1
Artibeus lituratus	5.84	38	3	41	7.3	-1 - 1
Myotis keaysi	5.61	2	0	2	0	-28 - 4
Chiroderma villosum	5.56	5	0	5	0	-15 - 2
Saccopteryx bilineata	5.3	1	0	1	0	-42 - :
Sciurus aureogaster	5.23	71	8	79	7.3	1 - 1
Baiomys musculus	5.21	2	0	2	0	-28 - 4
Artibeus watsoni	5.13	2	0	2	0	-28 - 4
Choeroniscus godmani	5.05	10	3	13	23.1	-7 - 2
Pteronotus personatus	5.03	3	1	4	25	-18 - 2
Reithrodontomys mexicanus	4.91	1	0	1	0	-42 - :
Oryzomys rostratus	4.87	22	1	23	4.3	-4 - 1
Micronycteris microtis	4.23	1	0	1	0	-42 - :
Oligoryzomys fulvescens	4.2	6	0	6	0	-13 - 2
Peromyscus leucopus	3.8	22	4	26	15.4	-3 - 1
Sturnira ludovici	3.79	24	1	25	4	-3 - 1
Vampyrodes caraccioli	3.69	1	0	1	0	-42 - :
Liomys pictus	3.61	47	1	48	2.1	0 - 1
Glossophaga commissarisi	3.49	2	6	8	75	-11 - 1
Lonchorhina aurita	3.48	1	0	I	0	-42 - :
Phyllostomus discolor	3.48	0	1	1	100	-42 - :
Lucdown a bilobatum	2.24	5	0	3	0	-22
Uroaerma bilobalum	2.07	4	0	4	0	-10
Drocyon cinereourgenieus	2.97	1	0	1	0	-42
Myotis velifer	2.53	1	0	1	0	-42
Microtus mericanus	2.50	16	0	16	0	-10 - 1
Myotis nigricans	2.55	2	0	2	0	-28 - 1
Lentonycteris verhabuenae	2.47	1	1	2	50	-20
Reithrodontomys fulvescens	2.08	20	0	20	0	-4 - 1
Neotoma mexicana	1.99	5	0	5	0	-15 - 1
Eptesicus fuscus	1.82	1	0	1	0	-42 - :
Peromyscus levipes	1.34	1	0	1	0	-42 - :
Sorex saussurei	1.29	3	0	3	0	-22 - 2
Osgoodomys banderanus	1.21	9	0	9	0	-10 - 2
Liomys irroratus	1.16	8	0	8	0	-11 - 2
Myotis auriculus	0.22	2	0	2	0	-28 - 4
Tadarida brasiliensis	-0.09	1	0	1	0	-42 - :
Peromyscus hylocetes	-0.28	2	0	2	0	-28 - 4
Antrozous pallidus	-0.34	1	0	1	0	-42 - :
Peromyscus zarhynchus	-0.46	2	0	2	0	-28 - 4
Chaetodipus hispidus	-0.71	4	0	4	0	-18 - 2
Peromyscus pectoralis	-0.73	2	0	2	0	-28 - 4
Neotomodon alstoni	-0.9	17	0	17	0	-5 - 1
Baiomys taylori	-1.16	10	3	13	23.1	-7 - 2
Chaetoatpus nelsoni	-1.24	5	0	5	0	-22 - 3
Neotoma micropus	-1.27	16	0	16	0	-6 - 1
Peromyscus maniculatus	-1.5/	80	2	00	3.5	0 - 1.
Perognathus flavus	-1.41	1	0	1	0	-42 - 3
Dine domas marine i	-1.52	1	0	1	0	-42 - 3
Dipoaomys merriami	-2.01	1	0	1	0	-42 - 5

Predictions-Experiment Test zoonosis - Leishmaniasis



- Only about 50 (2.5%) of mammals on the American continent have been identified as hosts of Leishmania
- In Mexico only 8 out of 419 (2.1%) had been identified as hosts
- We collected 922 individuals from 70 species
- Predicted and confirmed 21 new species of mammal as carriers of Leishmania in Mexico
- 13 of them are bats, identified for the first time in Mexico
- Squirrels identified as carriers
- 33% of collected species were confirmed as hosts
- Overall infection rate was 6.7%
- No species could be rejected as a host at this infection rate at the 95% confidence level
- Changes the picture for control of Leishmania totally;
- Leishmania and Lutzomyias are eclectic in their host source.
- Linnean classification is NOT ecologically relevant

Baiomys taylori Chaetodinus nelsoni Neotoma micropus Peromyscus maniculatus

10

-1.16

-1.24 -1.27

-1.37

-1.41

-1.52 -2.01

Peromyscus eremicus Data-Predictions-Experime Perognathus flavus Test zoonosis - Leishmaniasis

_			_
	Mammals	Epsilon	Conf.
1	Eira barbara	10.1683	
2	Rhogeessa aeneus	9.3649	
3	Artibeus intermedius	9.1628	
4	Reithrodontomys gracilis	8.8921	Yes
5	Carollia sowelli	8.8303	
6	Heteromys gaumeri	8.8000	Yes
7	Peromyscus mexicanus	8.7859	
8	Heteromys desmarestianu	8.7164	Yes
9	Molossus rufus	8.6277	
10	Glossophaga soricina	8.5713	
11	Carollia perspicillata	8.5030	
12	Orthogeomys hispidus	8.3468	
13	Pteronotus parnellii	8.1632	
14	Desmodus rotundus	8.1519	
15	Dasyprocta mexicana	8.1128	
16	Sturnira lilium	8.0290	
17	Dermanura phaeotis	8.0055	
18	Dasyprocta punctata	7.9678	
19	Oryzomys couesi	7.7253	
20	Potos flavus	7.7246	
21	Conepatus semistriatus	7.6879	1000
22	Ototylomys phyllotis	7.5587	Yes
23	Ateles geoffroyi	7.4787	
24	Cryptotis magna	7.4207	
25	Cuniculus paca	7.3220	
26	Lampronycteris brachyotis	7.2852	
27	Sigmodon hispidus	7.2805	Yes
28	Peromyscus yucatanicus	7.2486	Yes
29	Oryzomys chapmani	7.1242	
30	Didelphis virginiana	7.1150	
31	Peromyscus melanocarpus	7.0260	
32	Microtus umbrosus	6.9630	
33	Thyroptera tricolor	6.9630	
34	Nasua narica	6.8953	
35	Megadontomys cryophilus	6.6830	State State
36	Oryzomys alfaroi	6.6816	
37	Sorex veraepacis	6.6797	
38	Carollia subrufa	6.6316	
39	Peromyscus aztecus	6.6173	
40	Didelphis marsupialis	6.4390	Yes
41	Sciurus yucatanensis	6.3865	
42	Philander opossum	6.2546	
43	Habromys ixtlani	6.1120	
44	Microtus waterhousii	6.1120	Constant .
45	Pteronotus rubiginosus	6.1120	
46	Reithrodontomys microdor	6.0967	
47	Coendou mexicanus	6.0268	
48	Centurio senex	6.0076	
49	Artibeus jamaicensis	5.9786	
50	Glossophaga morenoi	5.8847	

	Mammals	Epsilon	Conf
51	Molossus sinaloae	5.8518	
52	Artibeus lituratus	5.8422	
53	Mormoops megalophylla	5.8374	
54	Habromys lepturus	5.7848	
55	Myotis keaysi	5.6148	
56	Chiroderma villosum	5.5562	
57	Tamandua mexicana	5.4845	
58	Tylomys nudicaudus	5.4510	
59	Saccopteryx bilineata	5.2984	
60	Macrotus mexicanus	5.2472	
61	Sciurus aureogaster	5.2267	
62	Baiomys musculus	5.2092	
63	Rhogeessa tumida	5.1950	
64	Sciurus deppei	5.1414	
65	Dermanura watsoni	5.1338	
66	Otonyctomys hatti	5.1338	
67	Orthogeomys grandis	5.0556	
68	Alouatta palliata	5.0457	
69	Choeroniscus godmani	5.0457	
70	Peropteryx macrotis	5.0457	
71	Pteronotus personatus	5.0266	
72	Lontra longicaudis	4.9330	
73	Reithrodontomys mexicanu	4.9120	
74	Oryzomys rostratus	4.8681	
75	Mimon cozumelae	4.8327	
76	Pteronotus davyi	4.7943	
77	Herpailurus yagouaroundi	4.7100	
78	Glossophaga leachii	4.6849	
79	Rhogeessa gracilis	4.6317	
80	Sylvilagus brasiliensis	4.6317	
81	Hodomys alleni	4.5155	
82	Leopardus wiedii	4.4420	
83	Peromyscus simulatus	4.4195	
84	Sigmodon alleni	4.3707	
85	Bassariscus sumichrasti	4.3110	
86	Oryzomys fulvescens	4.3110	
87	Diphylla ecaudata	4.3013	
88	Oryzomys melanotis	4.2907	Yes
89	Micronycteris microtis	4.2338	
90	Mazama americana	4.2274	
91	Microtus oaxacensis	4.2061	
92	Rheomys thomasi	4.2061	
93	Oryzomys saturatior	4.2061	
94	Myotis elegans	4.2024	11/3
95	Oligoryzomys fulvescens	4.1984	
96	Natalus stramineus	4.0626	
97	Balantiopteryx io	4.0522	
98	Nyctinomops laticaudatus	4.0522	1.119%
99	Tlacuatzin canescens	4.0119	
100	Odocoileus virginianus	3 9265	

	Mammals	Epsilon	Conf.
101	Balantiopteryx plicata	3.8590	
102	Peromyscus leucopus	3.7994	
103	Sturnina ludovici	3.7888	
104	Enchisthenes hartii	3.6929	
105	Vampyrodes caraccioli	3.6929	
106	Eptesicus furinalis	3.6453	
107	Liomys pictus	3.6107	
108	Glossophaga commissaris	3.4861	
109	Lonchorhina aurita	3.4781	
110	Phyllostomus discolor	3.4781	
111	Peromyscus gymnotis	3.4516	100
112	Anoura geoffrovi	3.4201	
113	Platvrrhinus helleri	3.3586	
114	Eumops bonariensis	3.3398	1.11
115	Sciurus variegatoides	3.3398	
116	Uroderma bilobatum	3.3373	
117	Lasiurus intermedius	3,2197	
118	Lasiurus ega	3,1739	
119	Peromyscus megalops	3,1410	
120	Eumops glaucinus	3.0564	
121	Urocyon cinereoargenteus	2 9697	
122	Procyon lotor	2 9502	1. S. J. R. S.
123	Hylonycteris underwoodi	2 9343	
124	Rhynchonycteris naso	2 8580	
125	Entesicus brasiliensis	2 8106	
126	Myotis albescens	2.8106	
127	Lophostoma evotis	2 8106	
128	Tapirus bairdii	2 8106	
129	Vampyrum spectrum	2 8106	
130	Marmosa mexicana	2 7731	Yes
131	Peromyscus furvus	2 7731	100
132	Myotis velifera	2 5757	
133	Spilogale putorius	2 5411	
134	Microtus mexicanus	2 5268	
135	Dasynus novemeinetus	2 4725	
136	Myotis nigricans	2 4704	
137	I onhostoma brasilianse	2 4407	
138	Diclidurus albus	2 4407	
130	Sciurus niger	2 4407	
140		2 4 2 6 8	
1/1	Nyctomys sumichraeti	2 4026	
142	Sigmodon mascotensis	2 3815	
1/3	Alouatta pigra	2 3374	
143	Peromyscus melanophrys	2 2204	
1/15	Dermanura tolteca	2 1020	
140	Trachons cirrhosus	2.1920	
140	Bauerus dubiaguercus	2.1003	
147	Spilogale pygmaea	2.1012	
140	L optopyctoris pivalis	2.1012	
149	Sylvilague floridanue	2.1402	
100	lo vivilagus ilo lualius	2.1002	Contraction of the second



Biotic facilitation seems to be the norm. Species are not distributed randomly





Conclusions

- Prediction of arbovirus disease risk factors and interventions is of huge public health importance
- Arbovirus diseases are Complex Adaptive Systems
 - * Multi-factorial, multi-scale, multi-discipline —> multi-interaction
 - Many arbovirus are multi-host, multi-vector
 - There are too many interactions to observe directly
 - Standard mathematical techniques model only a few factors
- The Data Revolution has made available large amounts of data with which their complex, adaptive nature may be better modelled
 - Spatio-temporal data about organisms, relative to each other (biotic) and relative to the environment (abiotic), can be used to deduce the nature of their interactions
 - * This can be done at the niche level (one to many) and at the community level (many to many)
 - * Obtaining and integrating data is a huge challenge political and technical
- The optimal use of this data requires innovation in modelling using multiple techniques from SIR-type models to agent-based modelling and the use of advanced machine learning and AI techniques.
- Our work on various zoonosis show the utility of innovative approaches that use data of arbitrary spatial resolution and format, such as predicting host range.
 - * Importance of a Data-Predictions-Experiment production line approach to emerging diseases
 - Importance of a multi-pathogen, multi-vector, multi-host approach



δώς μοι πά στώ καὶ τὰν γάν κινάσω Give me a place to stand on and I'll move the earth Give me enough data and I'll predict anything

The Data Revolution will revolutionise our ability to model and understand ecology