# 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





What do we know...?

## Known cases

Known vectors Known hosts
Known risk factors


# How do we infer what's under the water? 

 8We 17 What do we know...? Inteoratither Known cases Known vectors Known hosts Known risk factors Known 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 given the niche factors $\mathbf{X}$

Characterizes niche and "anti-niche"

What affects it?
The "niche"
$\mathbf{X}=(\mathrm{X} 1, \mathrm{X} 2, \mathrm{X} 3, \ldots, \mathrm{XM})$

A large part of the complexity is in the multi-factoriality of both C and X . Adaptation is
inherent in the fact that $\mathrm{P}(\mathrm{C} \mid \mathrm{X})$ of both C and X . Adaptation is
inherent in the fact that $\mathrm{P}(\mathrm{C} \mid \mathrm{X})$ can change in time.

$$
X=X(\mathrm{sd})+X(\mathrm{se})+X(\mathrm{n})+X(\mathrm{ev})+X(\mathrm{~g})+X(\mathrm{af})+X(\mathrm{hm})+X(\mathrm{i})+X(\mathrm{sp})+\ldots
$$



Micro-Climatic factors

Risk score

## $\mathrm{P}(\mathrm{C} \mid \mathrm{X}(\mathrm{t}))$

Phenotypic characteristics

Competitor species
Predator species

Problems of co-dependence and causality

## What is a community of emerging diseases?



# The Potential Solutions 

## One model = one question

$\rightarrow$ one answer

## Mathematical Modeling of Infectious Disease Dynamics



## Standard SIR models



Figure 3: Structure of the SIR model.

$$
\begin{aligned}
\frac{d S}{d t}(t) & =\mu N-\beta S(t) I(t) / N-\mu S(t) \\
\frac{d I}{d t}(t) & =\beta S(t) I(t) / N-\gamma I(t)-\mu I(t) \\
\frac{d R}{d t}(t) & =\gamma I(t)-\mu R(t) \\
\frac{d C_{i}}{d t} & =p \beta S(t) I(t)
\end{aligned}
$$

Complication of answering just one question - the number of infected, based on a small


Figure 4: Structure of the SIR/SI model.

$$
\begin{aligned}
\frac{d S_{H}}{d t}(t) & =\mu_{H} N_{H}-\beta_{H} S_{H}(t) I_{v}(t) / N_{v}-\mu_{H} S_{H}(t), \\
\frac{d I_{H}}{d t}(t) & =\beta_{H} S_{H}(t) I_{v}(t) / N_{v}-\gamma_{H} I_{H}(t)-\mu_{H} I_{H}(t), \\
\frac{d R_{H}}{d t}(t) & =\gamma_{H} I_{H}(t)-\mu_{H} R_{H}(t) \\
\frac{d I_{v}}{d t}(t) & =\beta_{v} S_{v}(t) I_{H}(t) / N_{H}-\mu_{v} I_{v}(t), \\
\frac{d S_{v}}{d t}(t) & =-\beta_{v} S_{v}(t) I_{H}(t) / N_{H}+\mu_{v} I_{v}(t), \\
\frac{d C_{i}}{d t} & =p \beta_{H} S_{H}(t) I_{v}(t) / N_{v},
\end{aligned}
$$

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.
$\mathrm{P}(\mathrm{C} \mid \mathbf{X}(\mathrm{t}))$
Co-occurrence between arbovirus occurrence, C, and niche variables, $\mathrm{X}(\mathrm{t})$
 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 dat
- Socio-economic data
- Socio-demographic data
- Phenotypic data
- Vegetable and crop cover
- Geographical data
- Medical and public health data...



# Predictive Model for potential hosts of ZIKV 

Mammals with most statistically significant geographic overlap with Aedes Aegypti


# 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

Culiseta
Culex

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 Mining: Predictive models, Risk factors

Field work : samples

Five groups distributed throughout Mexico DF, Chiap., NL, Jal., Tab. collected from over 70 species

Close relationship with public health authorities and private sector

Over 1200 mammals
Laboratory analysis
Two laboratories

## Data-Predictions

Bienvenido a la Plataforma de exploración de datos ecológicos del


## Data-Predictions Test zoonosis - Leishmaniasis




# Data-Predictions-Experiment 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 |  |
| 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 melanocarpu | 7.0260 |  |
| 32 | Microtus umbrosus | 6.9630 |  |
| 33 | Thyroptera tricolor | 6.9630 |  |
| 34 | Nasua narica | 6.8953 |  |
| 35 | Megadontomys cryophilus | 6.6830 |  |
| 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 |  |
| 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 mexican | 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 |  |
| 95 | Oligoryzomys fulvescens | 4.1984 |  |
| 96 | Natalus stramineus | 4.0626 |  |
| 97 | Balantiopteryx io | 4.0522 |  |
| 98 | Nyctinomops laticaudatus | 4.0522 |  |
| 99 | Tlacuatzin canescens | 4.0119 |  |
| 100 | Odocoileus virginianus | 3.9265 |  |
|  |  |  |  |
| 10 |  |  |  |


|  | 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 |  |
| 112 | Anoura geoffroyi | 3.4201 |  |
| 113 | Platyrrhinus helleri | 3.3586 |  |
| 114 | Eumops bonariensis | 3.3398 |  |
| 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 |  |
| 123 | Hylonycteris underwoodi | 2.9343 |  |
| 124 | Rhynchonycteris naso | 2.8580 |  |
| 125 | Eptesicus 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 |  |
| 132 | Myotis velifera | 2.5757 |  |
| 133 | Spilogale putorius | 2.5411 |  |
| 134 | Microtus mexicanus | 2.5268 |  |
| 135 | Dasypus novemcinctus | 2.4725 |  |
| 136 | Myotis nigricans | 2.4704 |  |
| 137 | Lophostoma brasiliense | 2.4407 |  |
| 138 | Diclidurus albus | 2.4407 |  |
| 139 | Sciurus niger | 2.4407 |  |
| 140 | Leptonycteris curasoae | 2.4268 |  |
| 141 | Nyctomys sumichrasti | 2.4026 |  |
| 142 | Sigmodon mascotensis | 2.3815 |  |
| 143 | Alouatta pigra | 2.3374 |  |
| 144 | Peromyscus melanophrys | 2.2204 |  |
| 145 | Dermanura tolteca | 2.1920 |  |
| 146 | Trachops cirrosus | 2.1663 |  |
| 147 | Bauerus dubiaquercus | 2.1612 |  |
| 148 | Spilogale pygmaea | 2.1612 |  |
| 149 | Leptonycteris nivalis | 2.1402 |  |
| 150 | Sylvilagus floridanus | 2.1002 |  |
|  |  |  |  |



## 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 $\rightarrow$ 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

$\delta \omega ̂ ৎ \mu o ı ~ п a ̂ ~ \sigma t \omega ̂ ~ k a i ̀ ~ t a ̀ v ~ ү a ̂ v ~ k ı v a ́ \sigma \omega ~$
Give me a place to stand on and I'll move the earth Give me enough data and I'1l predict anything

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

