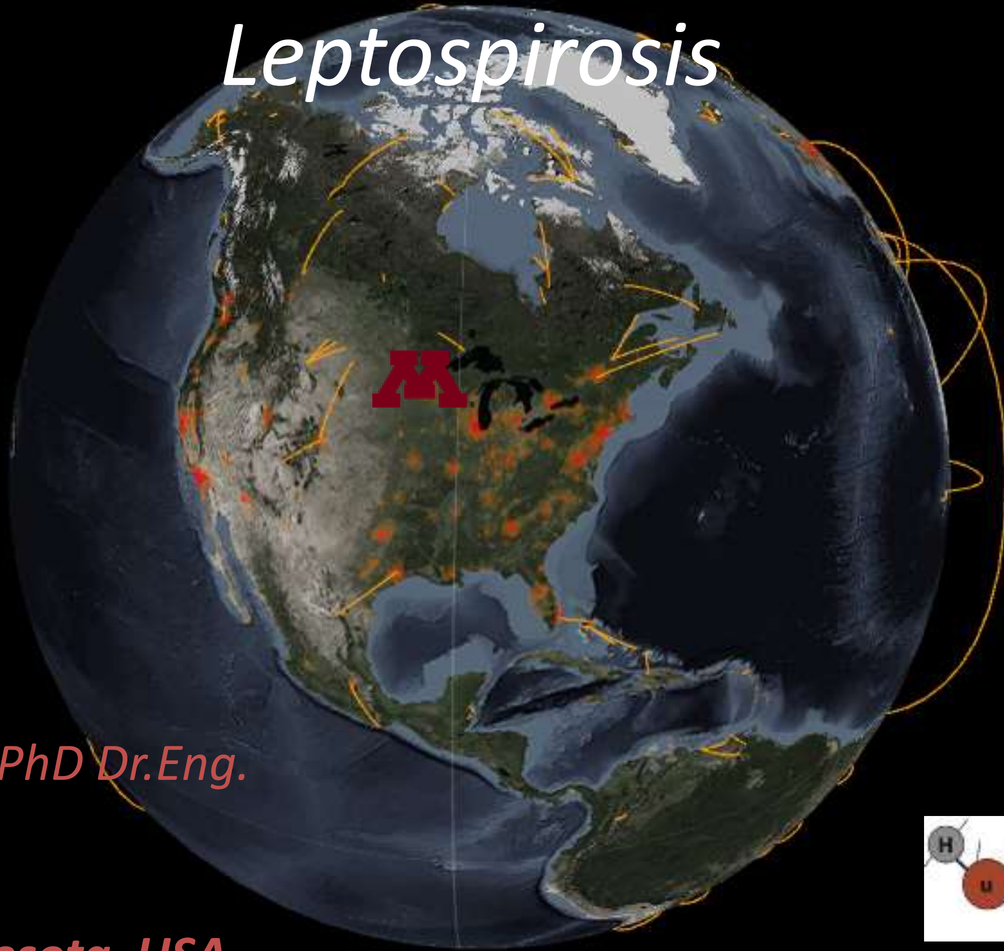
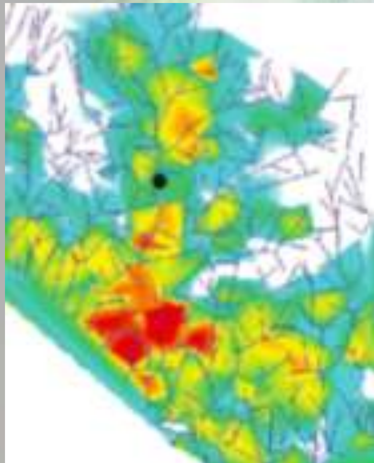
Dynamics and Computational Technology for Leptospirosis

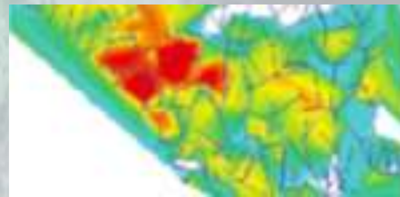


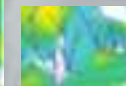
Matteo Convertino, PhD Dr.Eng.
Yang Liu, MSc
Abe Reddy, PhD
University of Minnesota, USA

Presentation conducted during the *International Workshop of the Oswaldo Cruz Institute/FIOCRUZ for Leptospirosis Research Based on Country Needs & the 5th Global Leptospirosis Environmental Action Network (GLEAN) Meeting* on November 10-12, 2015, in Rio de Janeiro, Brazil .

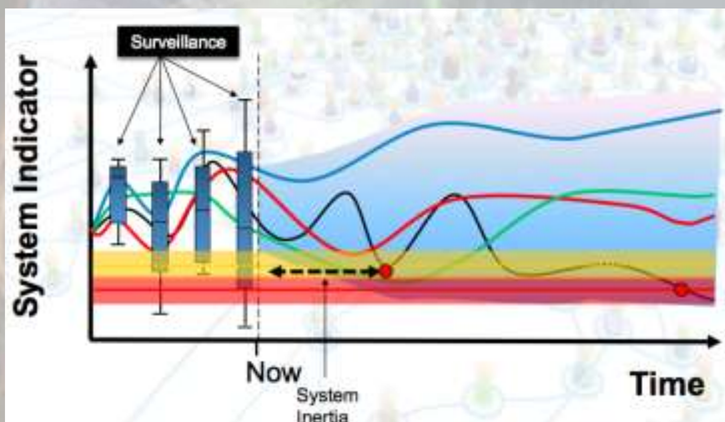
Three Pillars







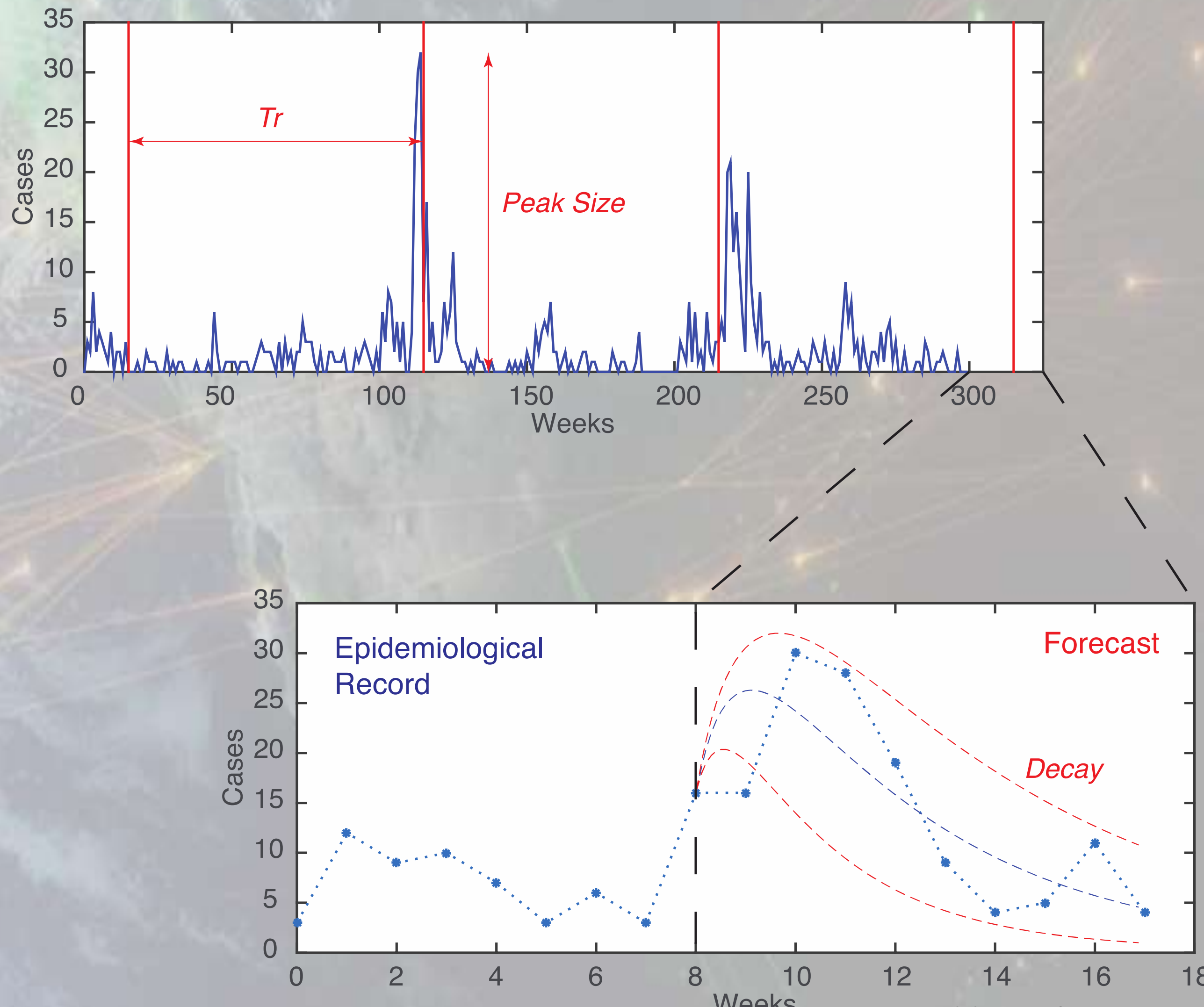
Scalability and Universality
(Endemic/Epidemic Characterization & Macro Prediction)



Environmental Dynamics
(Early Warning and Real-time Fine Scale Forecasting)

Systemic and Value-based Optimal Ecosystem Design
(Portfolio Decision Model)

From Emerging (Invariant) Patterns to Stochastic Early Warning System Models



Reddy et al., 2015, submitted to JRS-I

Example return periods

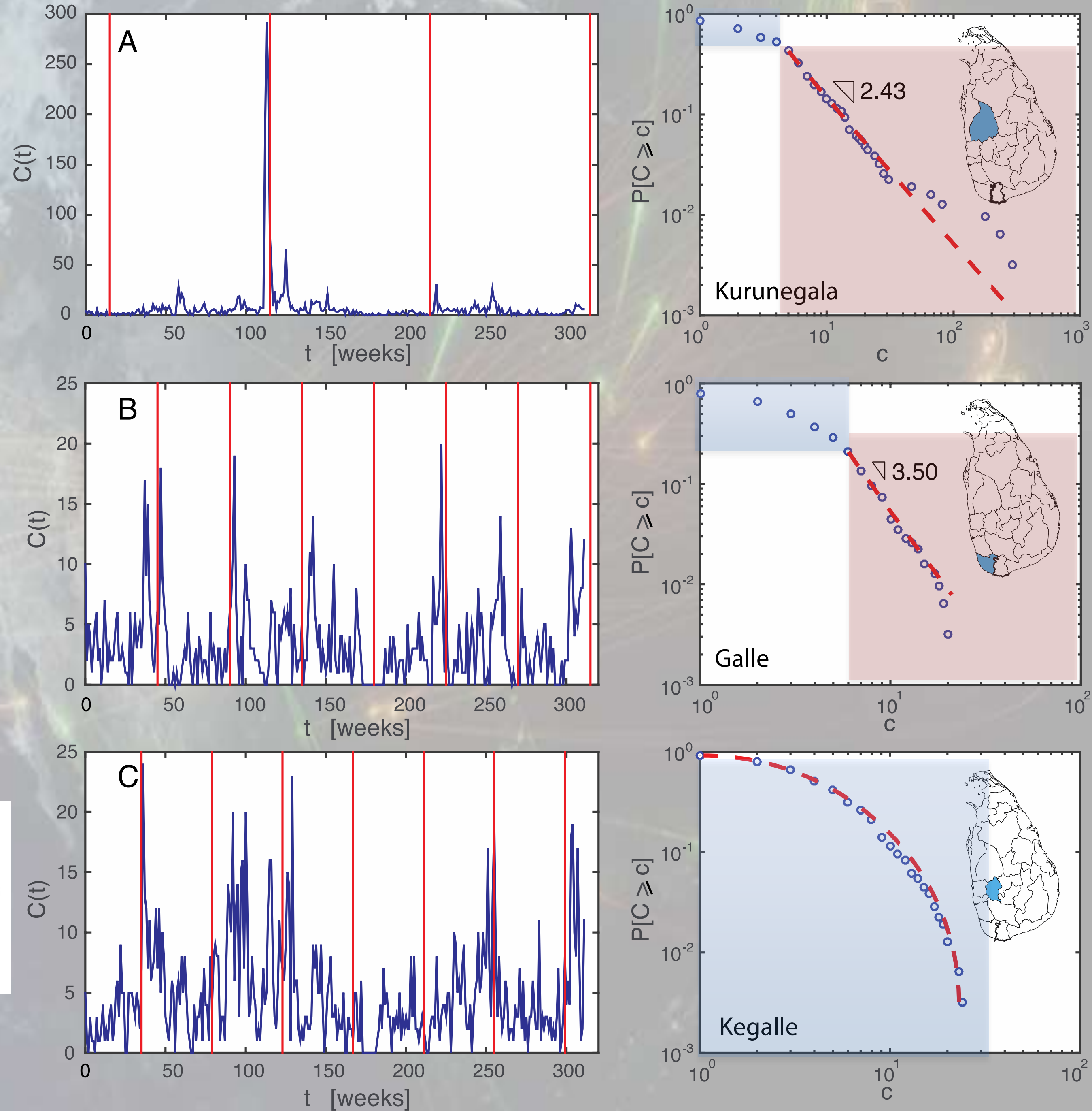
A) 115 weeks (>30)

B) 45 weeks (>15)

C) 43 weeks (>17)

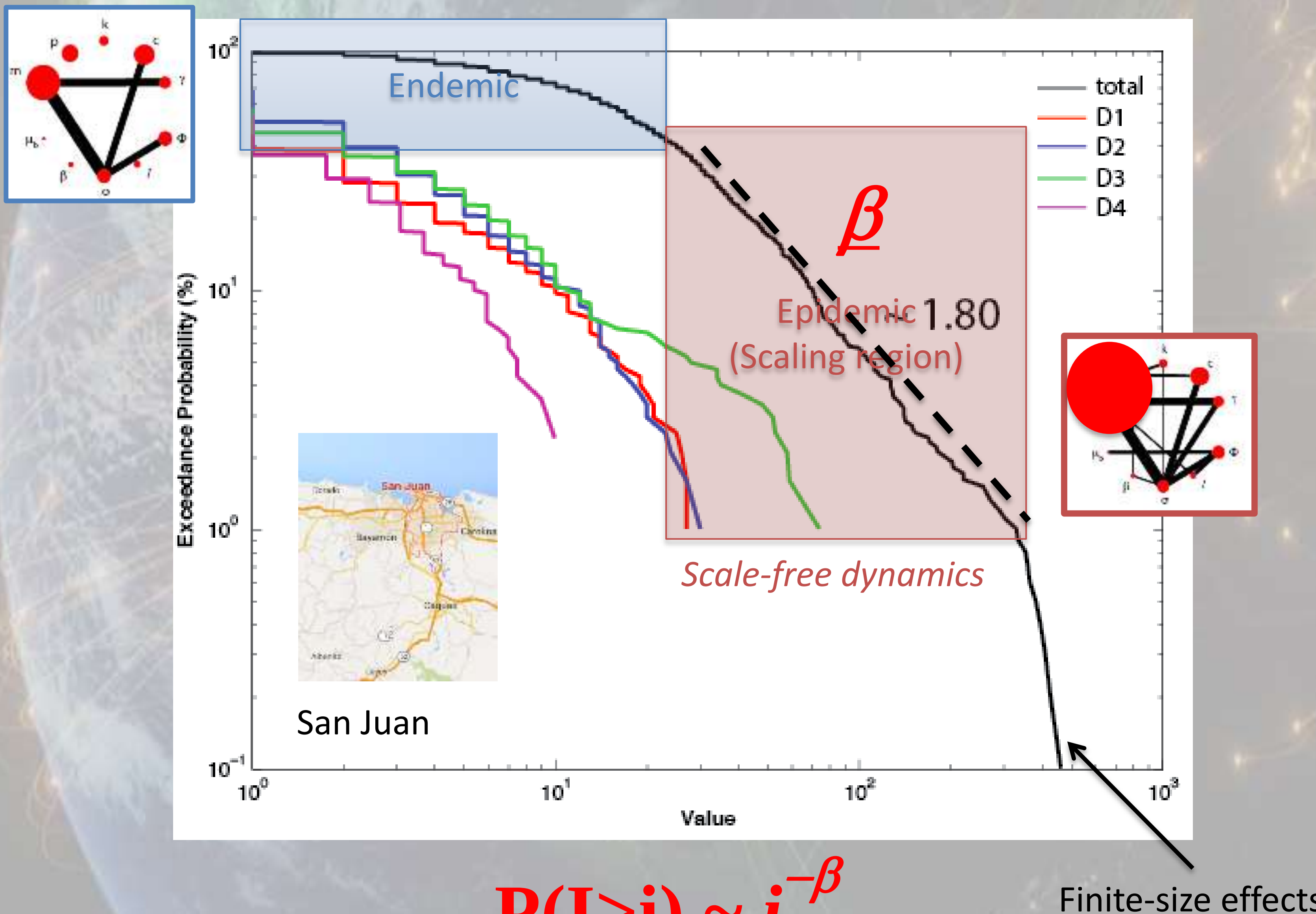
$$P(C \geq c)$$

$$T(C) = \frac{1}{P(C \geq c)}$$

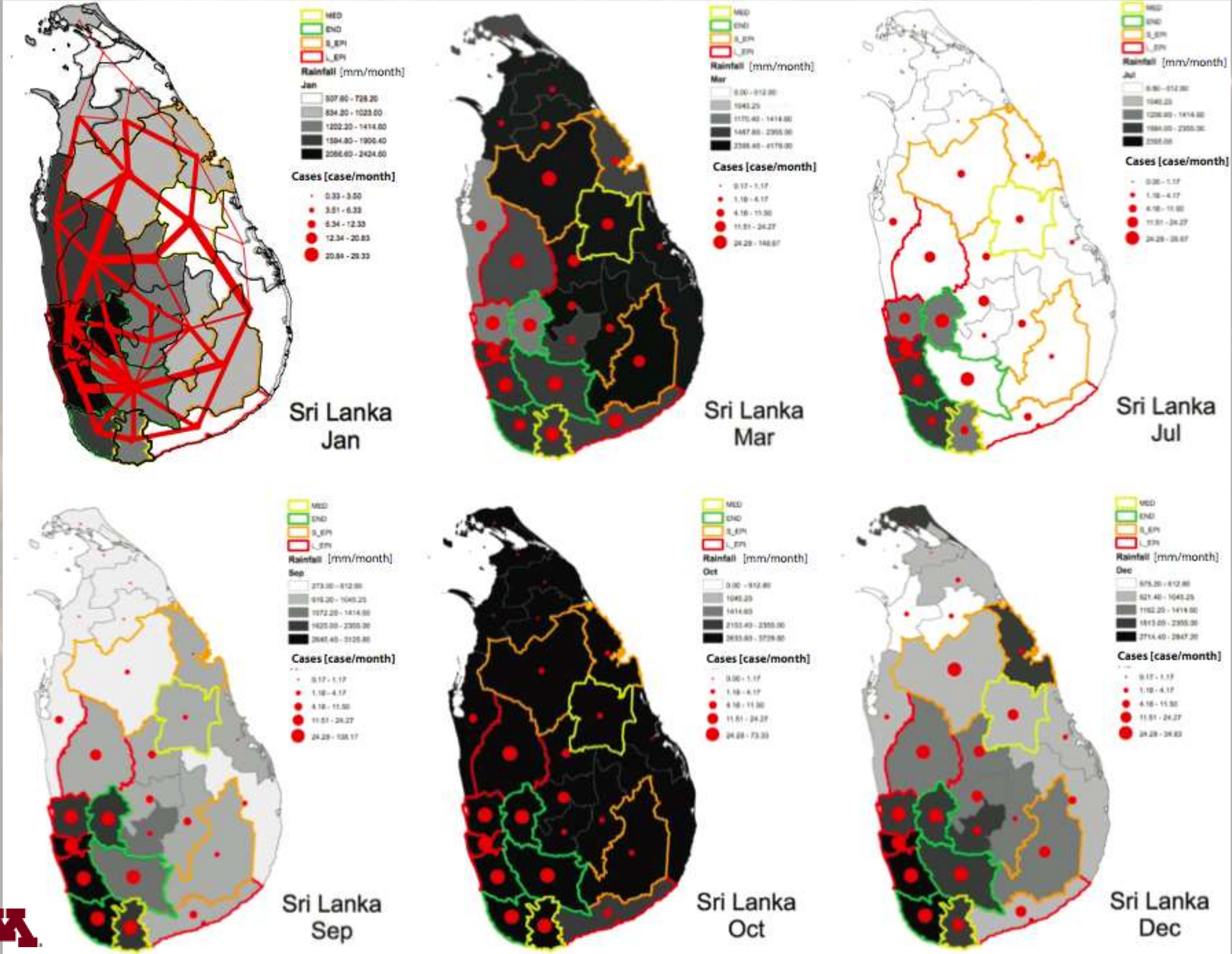


Distribution of EHL Outcomes

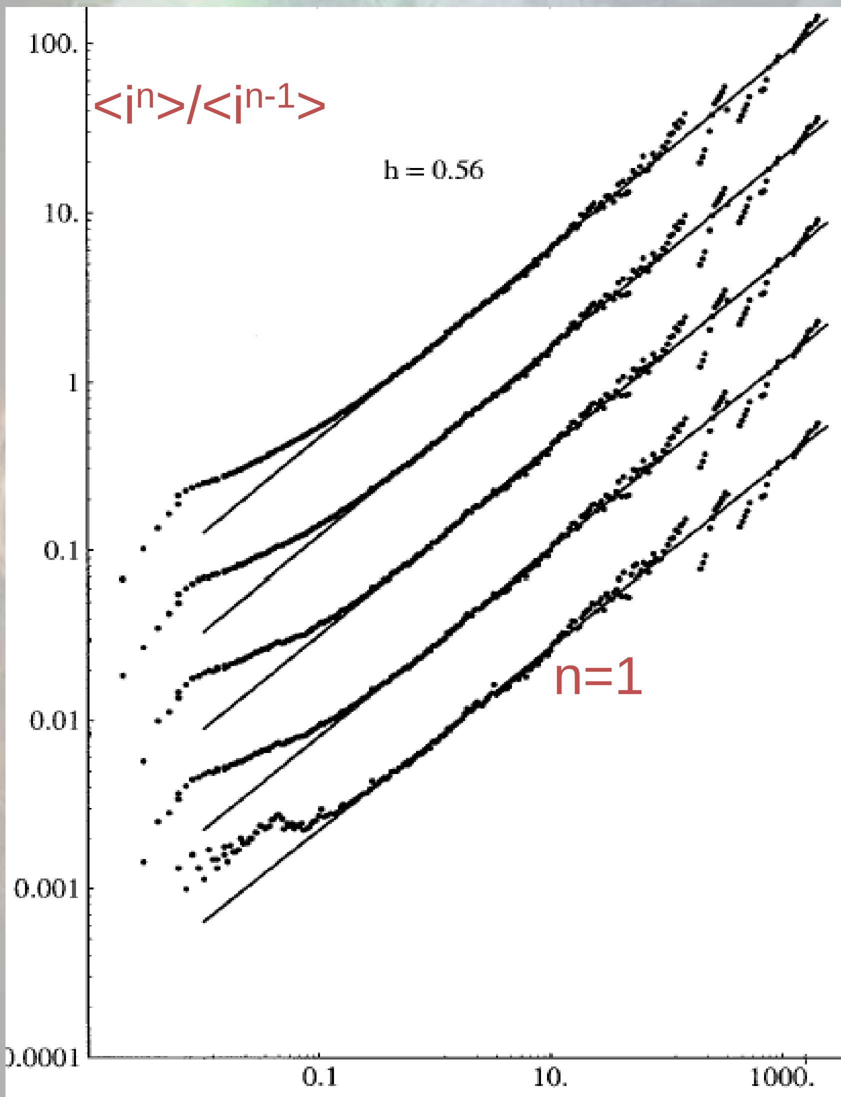
Exponential dynamics

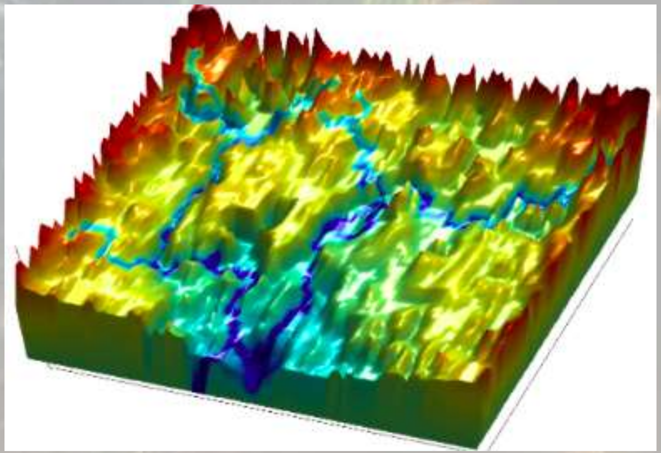


$$P(I>i) \sim i^{-\beta}$$


MED
END
S_EPI
L_EPI
Rainfall [mm/month]
Jan
Cases [case/month]
Sri Lanka Jan
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Mar
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Jul
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Sep
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Oct
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Dec

HydroEpi Networks and Scaling



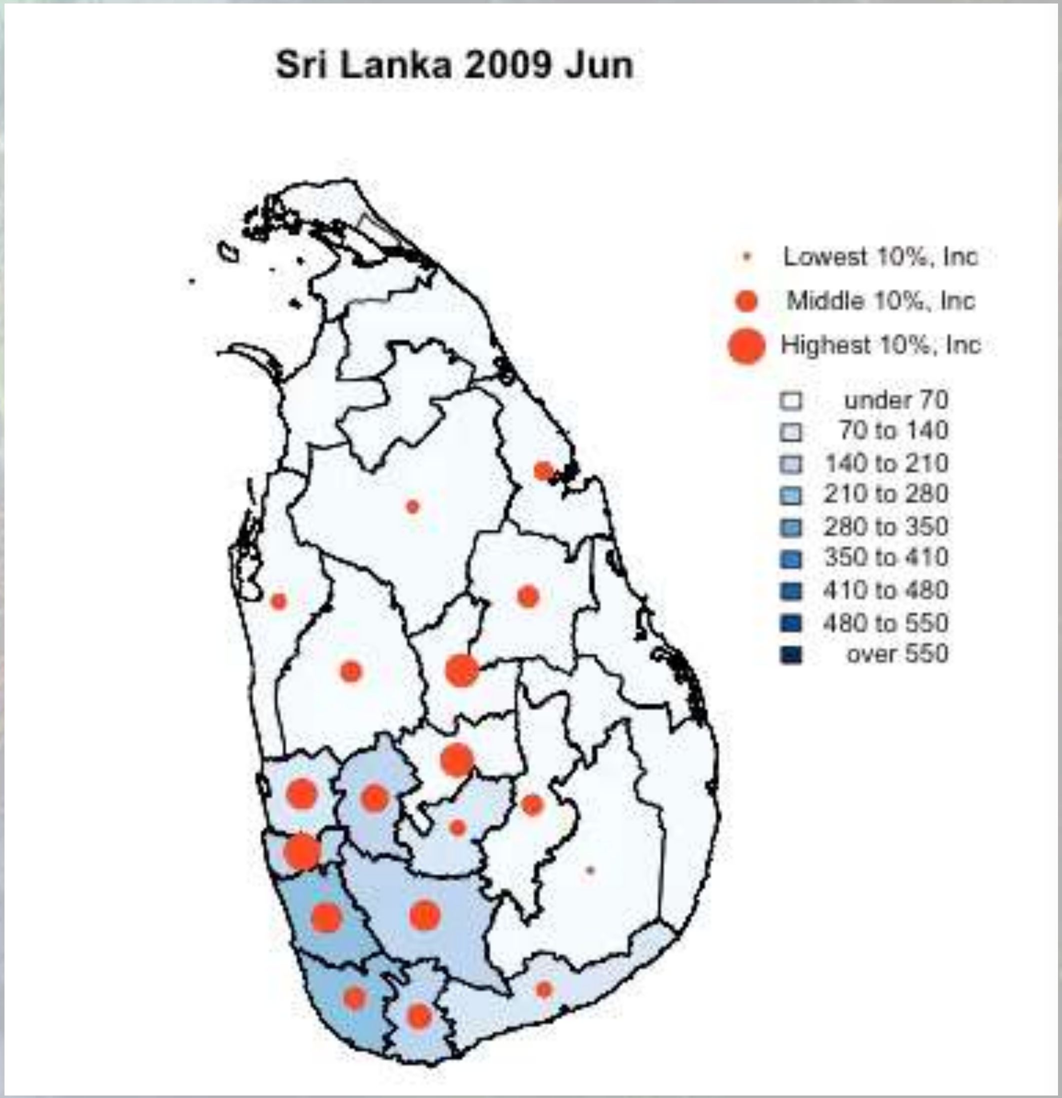


I = Cases
A = drainage area
h = scaling exponent

Multi-moment Scaling (!)

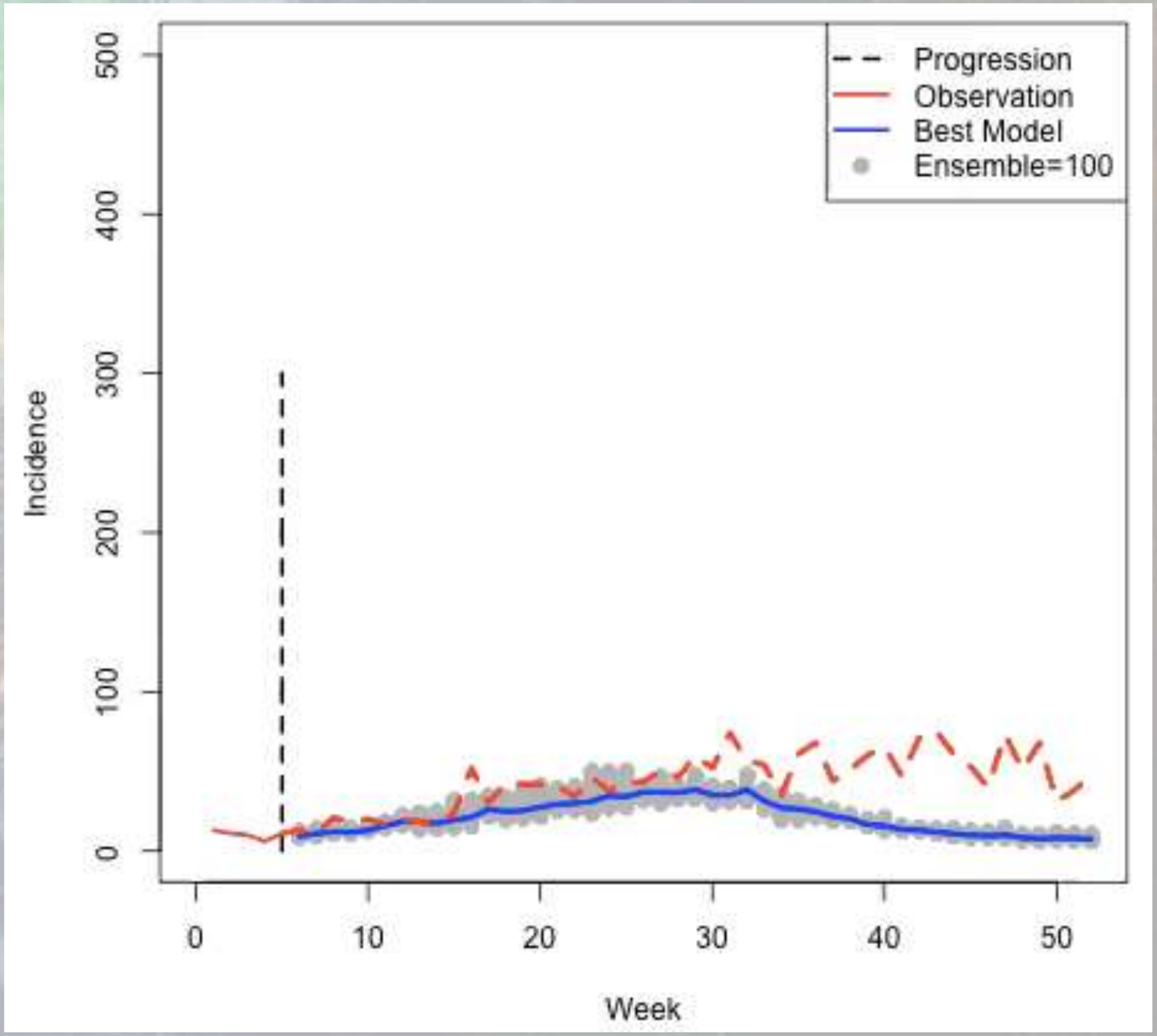
$$< \ell^{n} > / < \ell^{n-1} > \propto A^{h}$$

A Predictive Model



Reddy et al., 2015, submitted to JRS-Interface

The How



Progression
Observation
Best Model
Ensemble=100
Incidence
Week
0
100
200
300
400
500
0
10
20
30
40
50

Early Warning System (EWS) Model Analytics

- Model introduced by Azaele (PRL, 2010) to describe cholera
- Langevin equation with Gaussian white noise

$$\dot{I}(t) = b - \frac{I(t)}{\tau} + \sqrt{DI(t)}\xi(t)$$

$$\langle \xi(t)\xi(t') \rangle = 2\delta(t-t')$$

EWS Model Factors

- b ~ (inter community cases) immigration rate of infected hosts and contaminated water flow; proportional to j_e (the stressors)
- D ~ (intra community case fluctuations) stochasticity of disease incidence; proportional to W (the connections)
- τ ~ characteristic time scale of disease decay

b, D, tau found using a least squares optimization for peaks across Sri Lanka

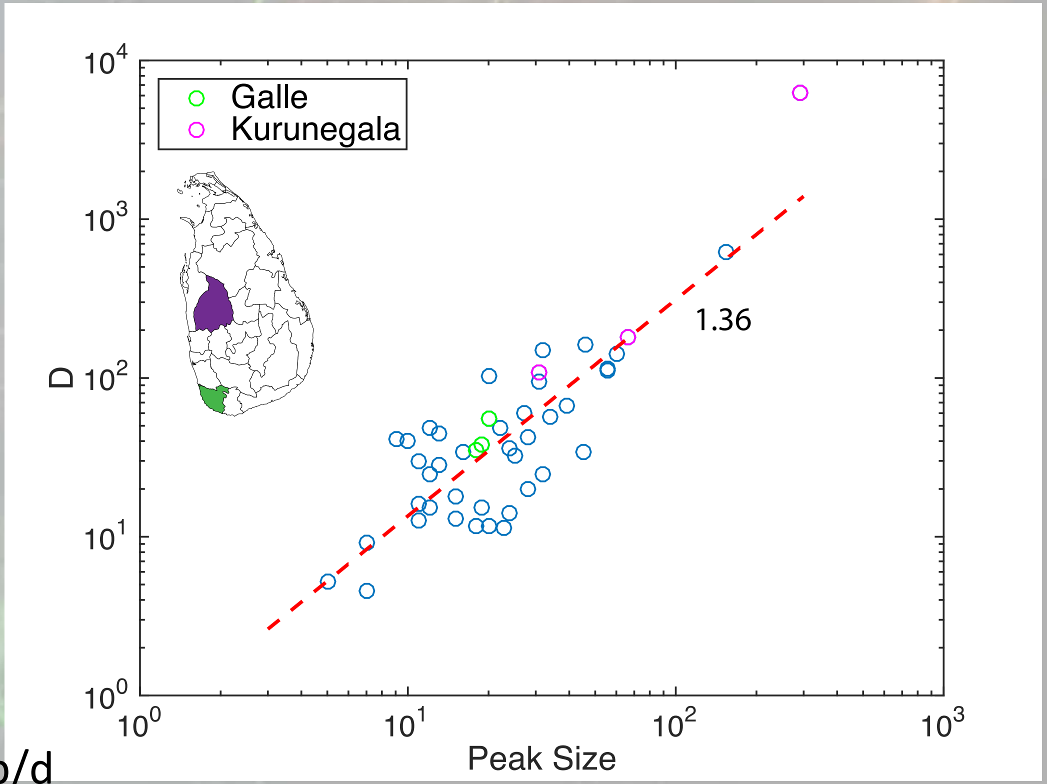
Scaling Model Factors

- D varies with peak size

$$D = 0.59x^{1.36}$$

$$b/D = .663 \pm .116$$

$$\tau = 1.301 \pm .496$$



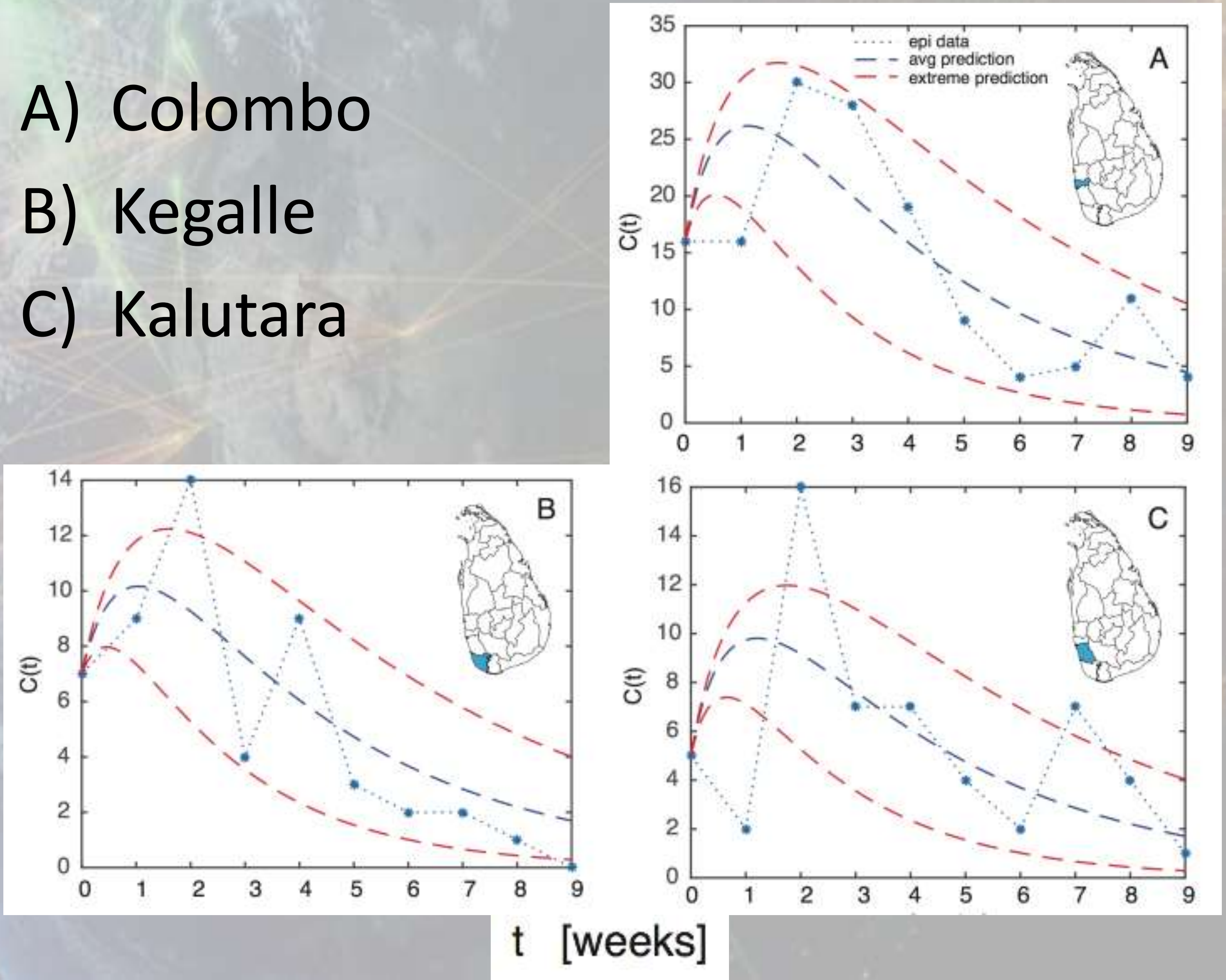
tau dependent on the disease as well as b/d

D is prop. to the peak size

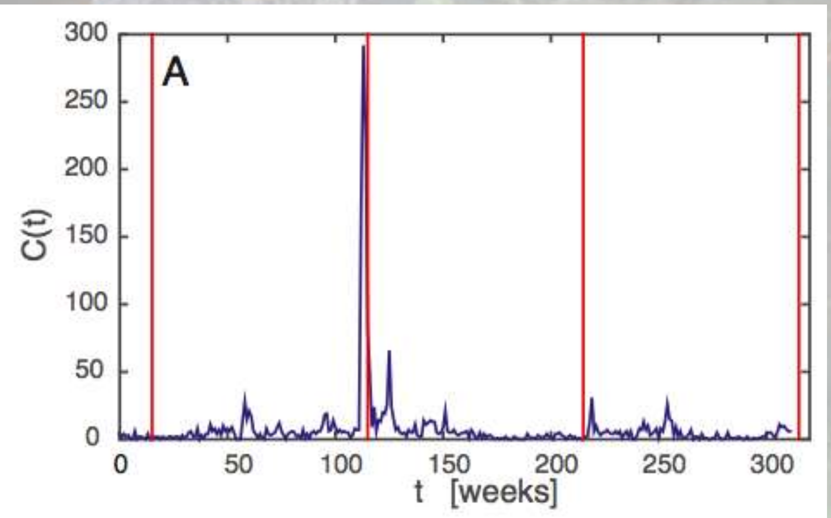
Reddy et al., 2015, submitted to JRS-I

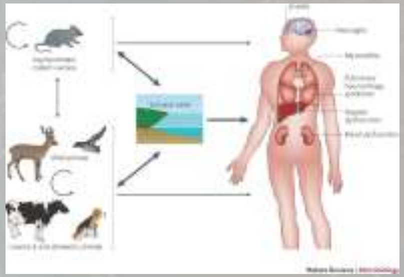
Model Prediction

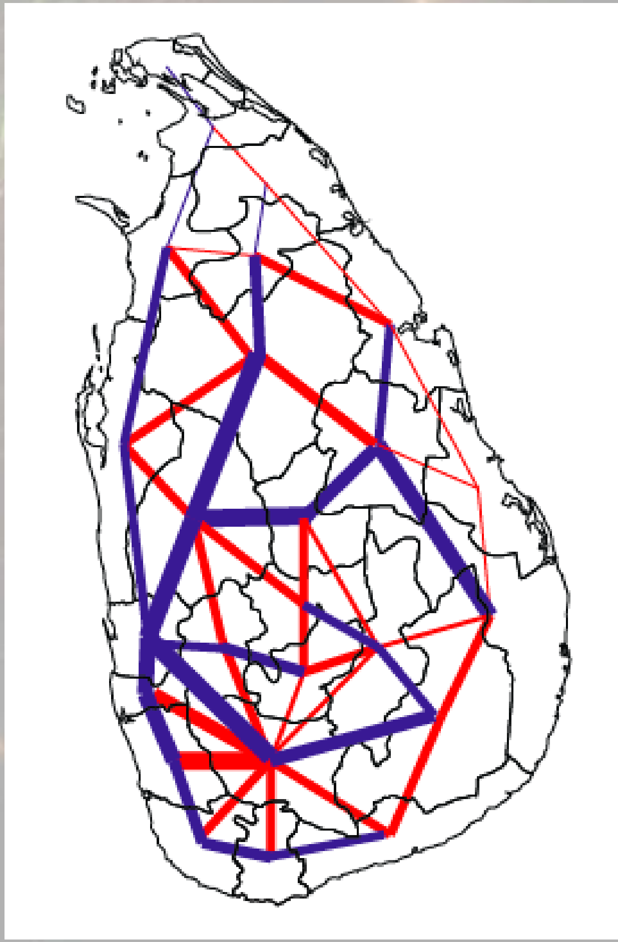
A) Colombo
B) Kegalle
C) Kalutara



Scaling and Early Warning Models: Application to Leptospirosis in Sri Lanka

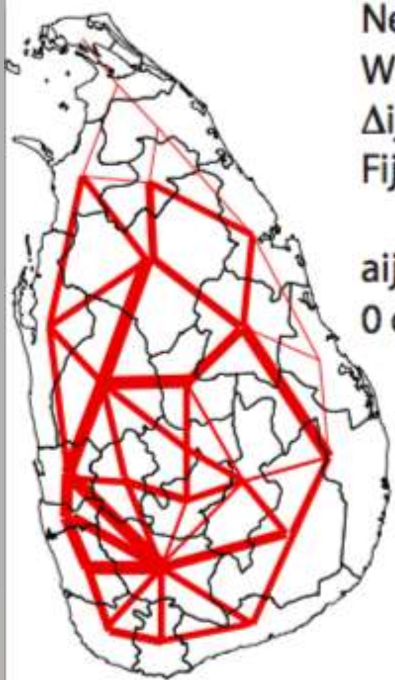






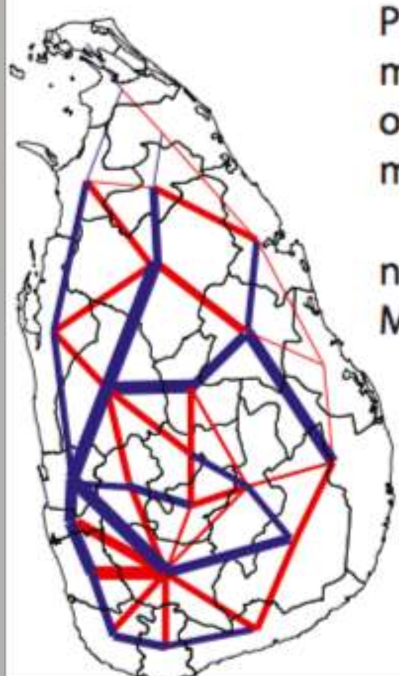
Reddy et al., 2015, submitted to JRS-I

Optimal Transmission Networks (OTN)



Nearest Neighborhood Model
Wij ~ 1/Δij (width)
Δij=|Ci-Cj| (gradient of cases)
Fij=1/Δij (fluxes)

aij=1 if j adjacent to j
0 otherwise



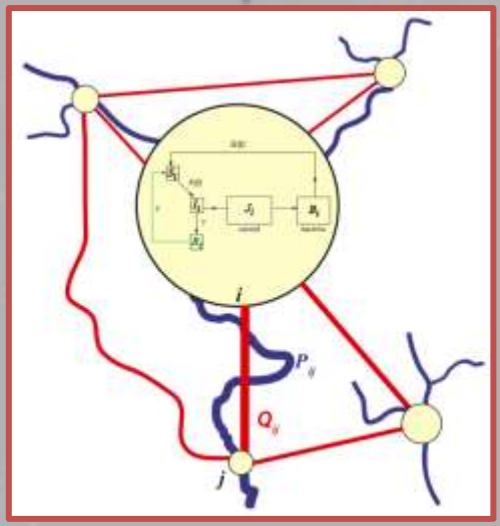
Preferential Case Pathways (at a closure point)
max Wij ~ 1/Δij
or min Δij=|Ci-Cj|
max(Fij)=1/min(Δij)

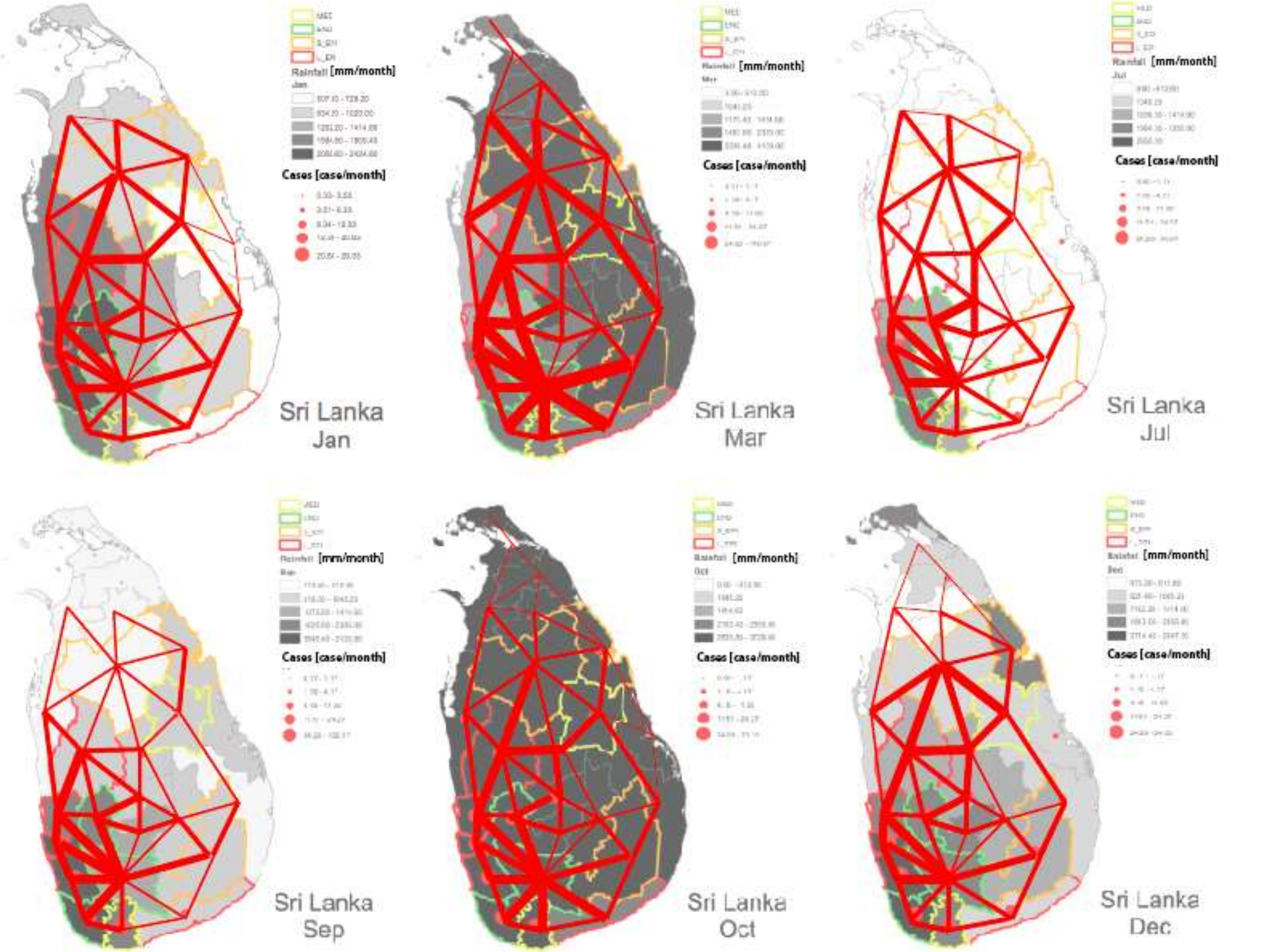
non looping network
Minimum Spanning Tree

W(x) only dependent on
Network Topology
(travel time distribution)

Width Function

$$c(t) = \int_0^\infty r(t) W(x)\, dP(x,t)$$




Rainfall [mm/month]
Cases [case/month]
Sri Lanka Jan
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Mar
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Jul
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Sep
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Oct
Rainfall [mm/month]
Cases [case/month]
Sri Lanka Dec

Connectopathies, Factorgenicity and Population Outcomes: A Morphological Effective Systemic EpiGraph model (MESE)

EPI (STATIC; RISK) TRANSPORT (DYNAMICS; OUTCOME)

$$I(\tau) = A \int j_e(\tau) [p_\gamma (f_{\gamma_1} * f_{\gamma_n})]_{t-\tau} d\tau$$

$$I(\tau) = A \int j_e(\tau) W(t-\tau) d\tau$$

$$f_\gamma = pdf(L_\gamma) pdf(T_{L_\gamma})$$

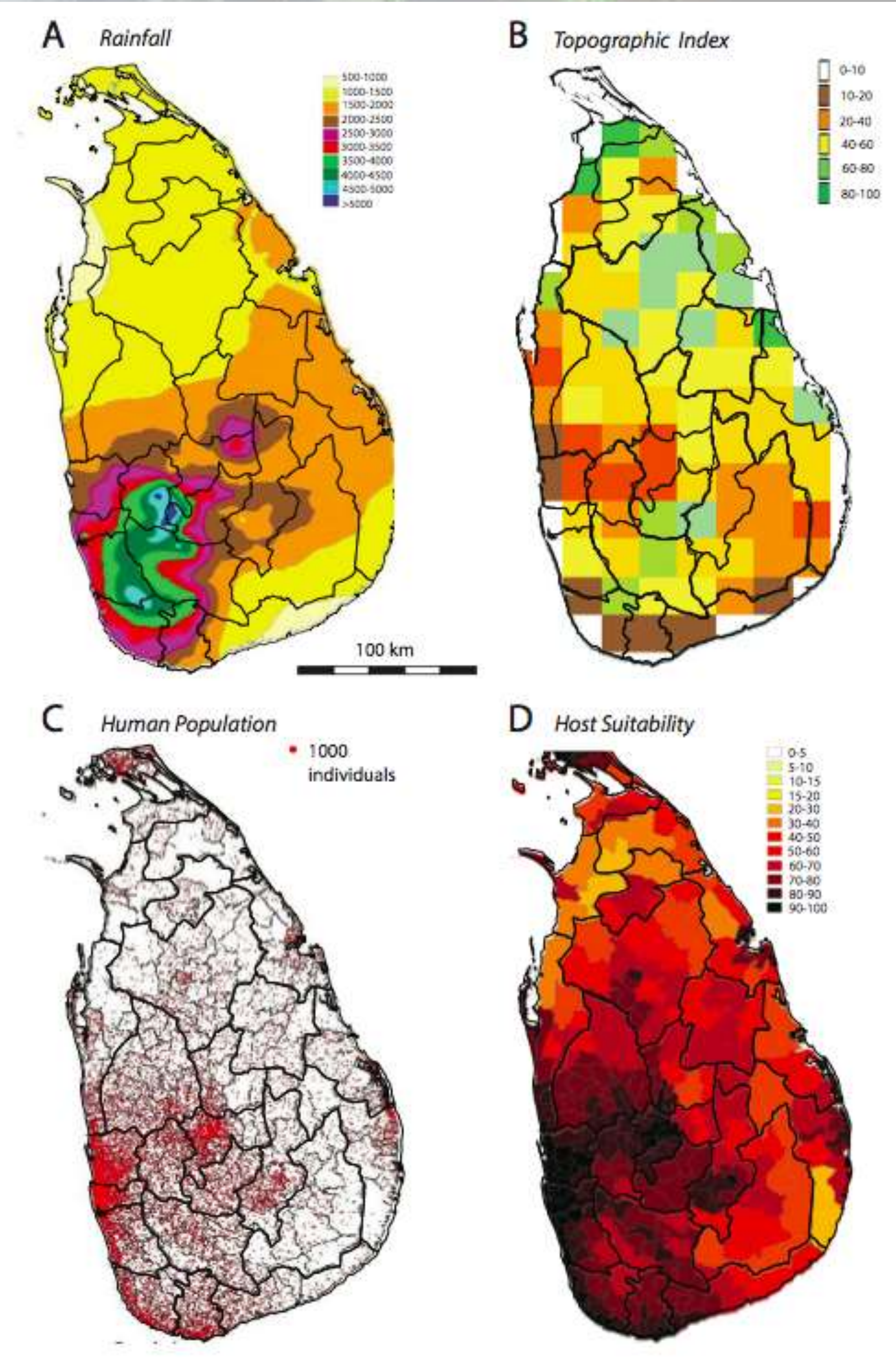
L=network length

Travel Time distribution ~ Arrival Time distribution (of Cases) ~ (Residence Time)$^{-1}$

Convertino, Huang, Liu, 2014, WRR, submitted
… GIUH from Rodriguez-Iturbe and Valdes, 1979

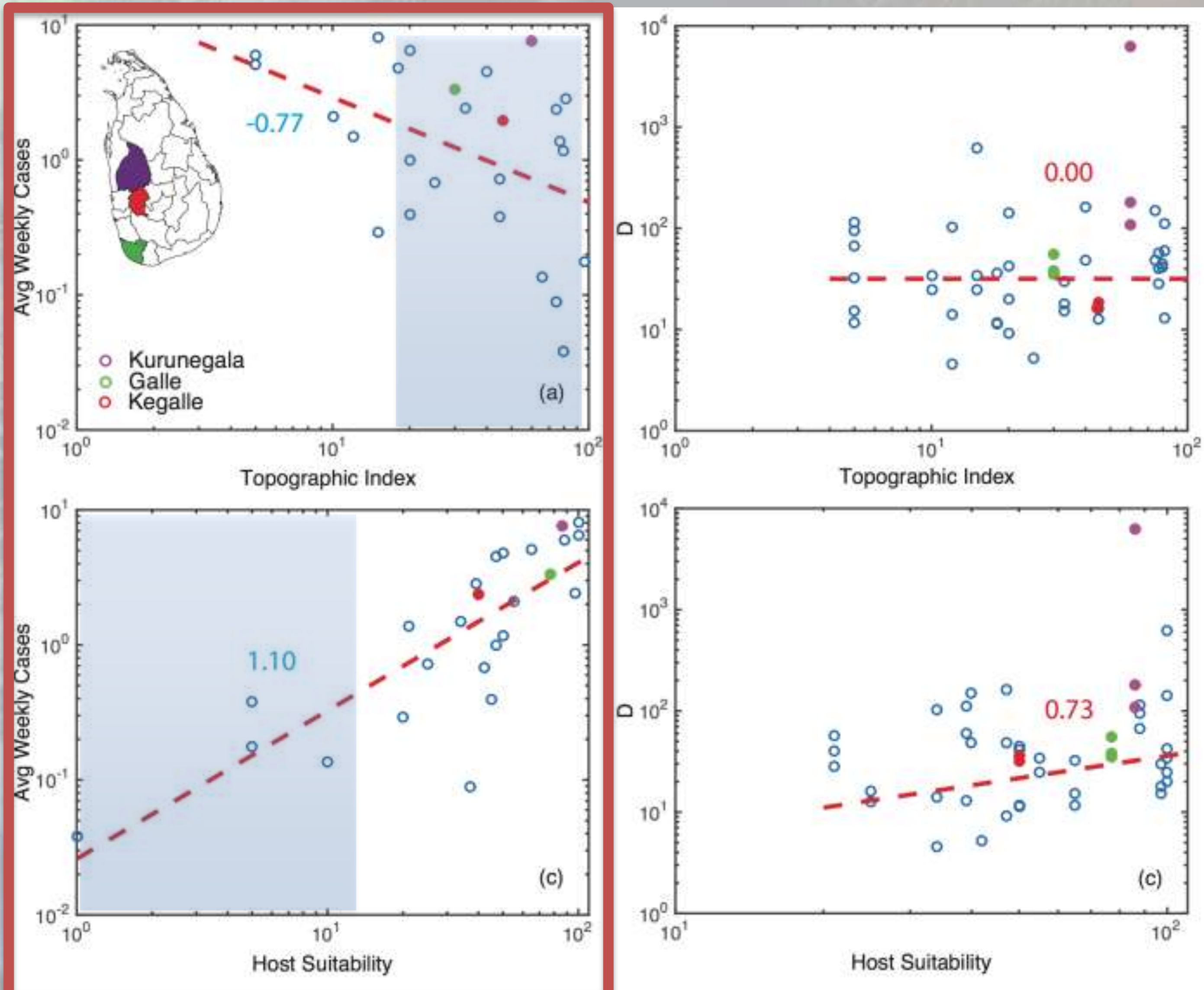
Socio-environmental factors (after metamodeling)

- Topographic Index
- Host Suitability
- Population
- Rainfall

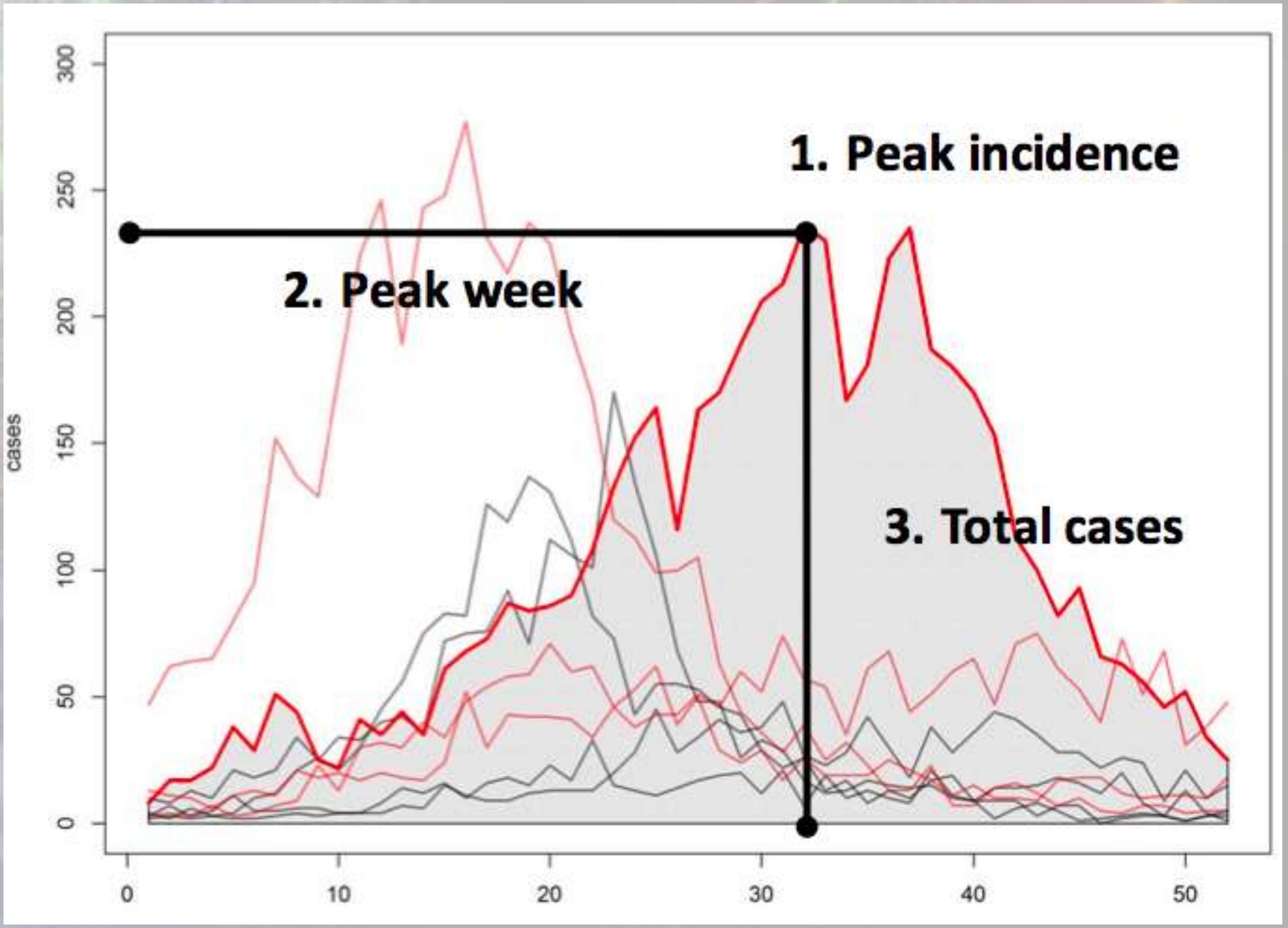


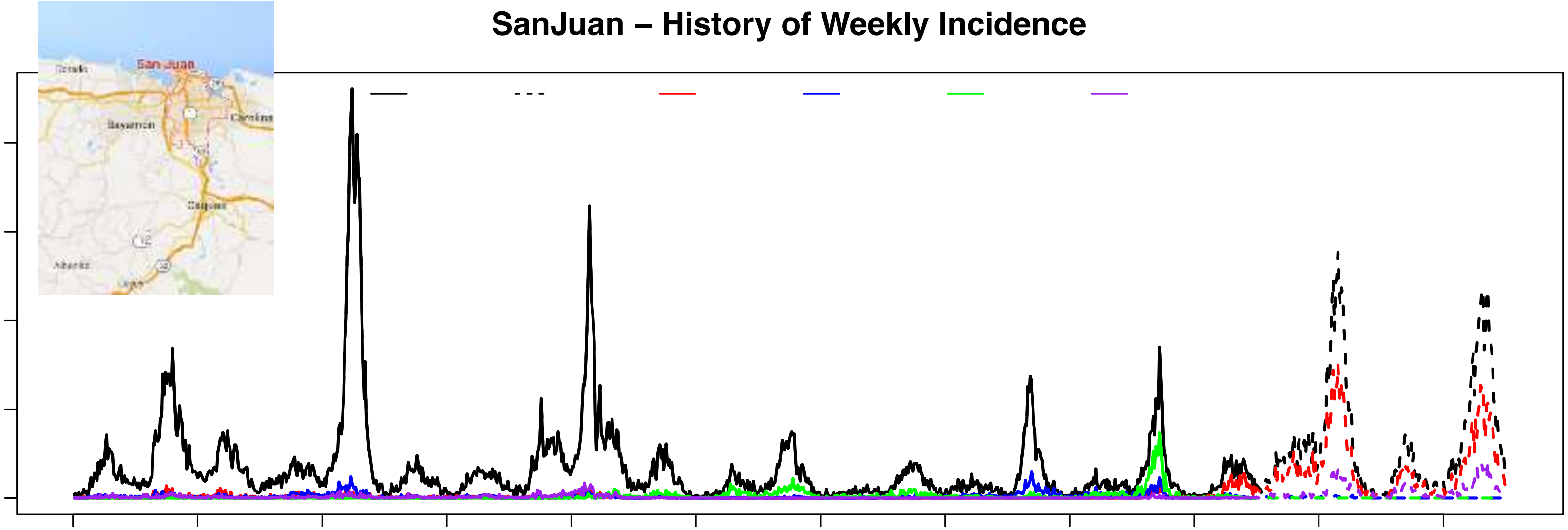
Reddy et al., 2015, submitted to JRS-I

Eco-epidemiological Scaling



Forecasting Model

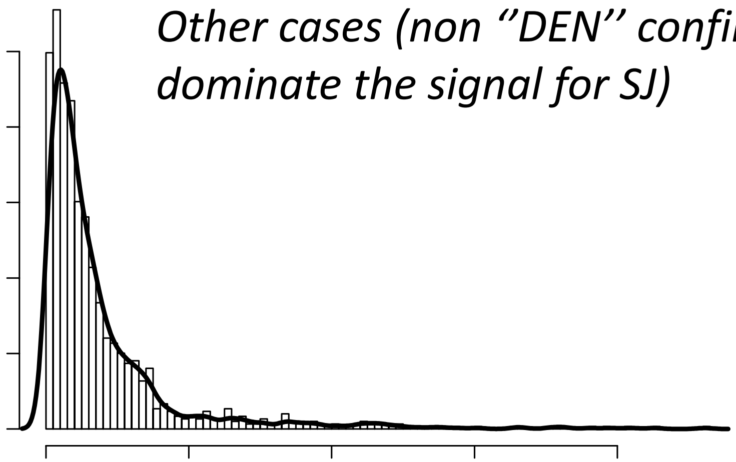




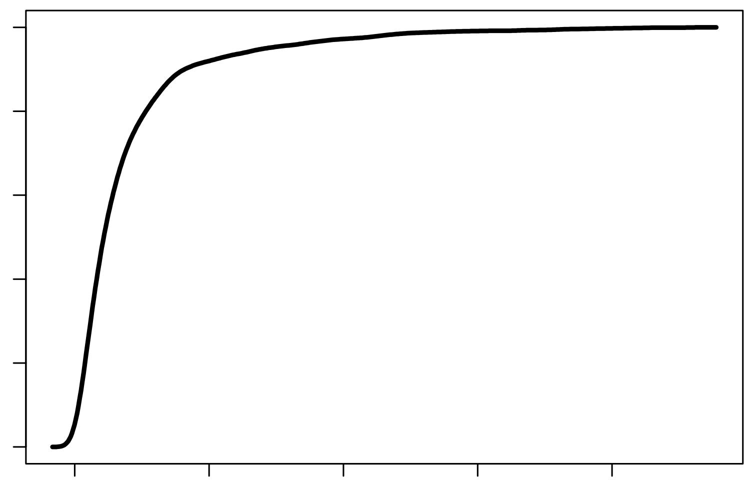
SanJuan – History of Weekly Incidence

SanJuan, PDF of Weekly Incidence

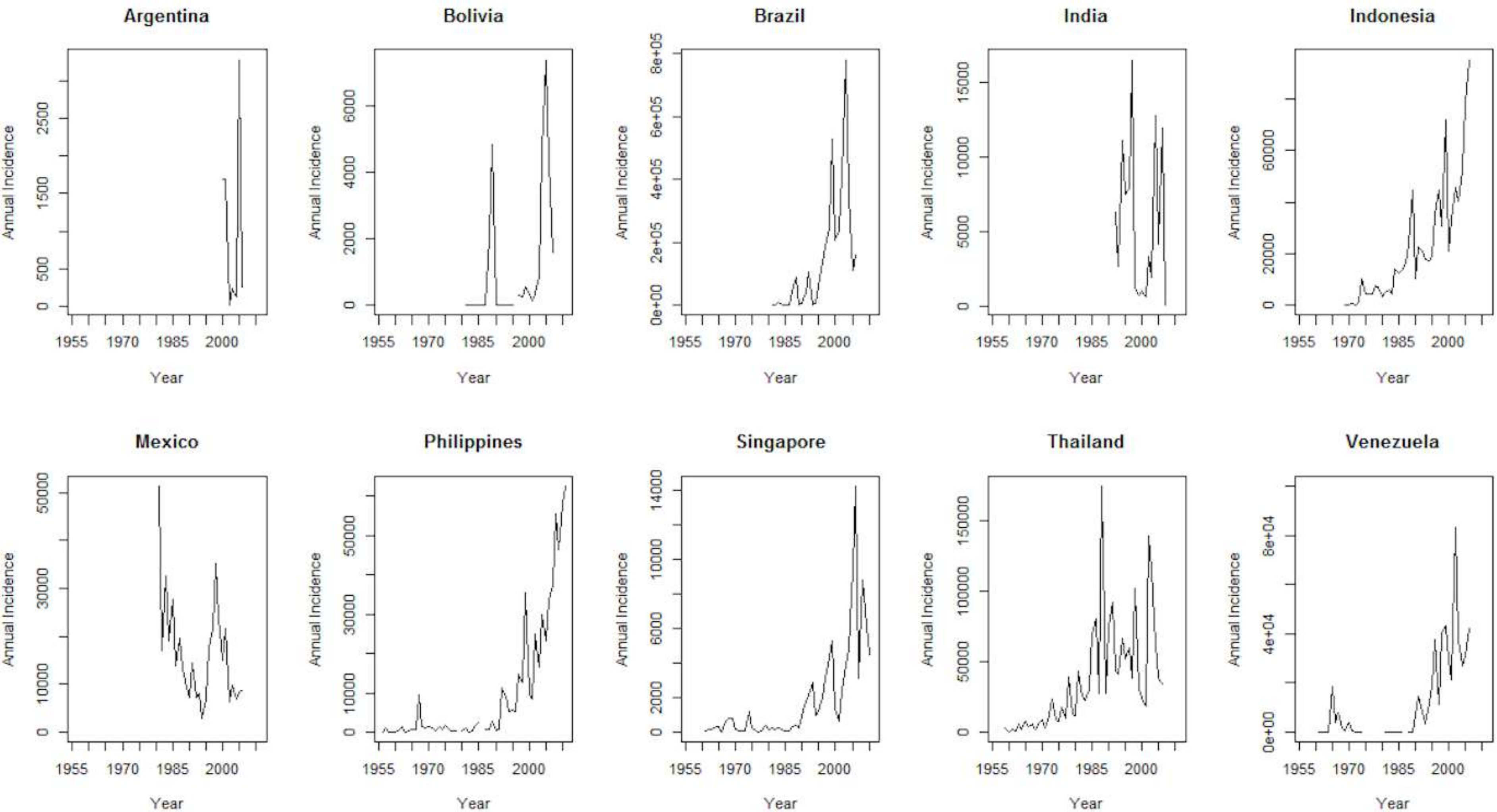
Other cases (non "DEN" confirmed dominate the signal for SJ)



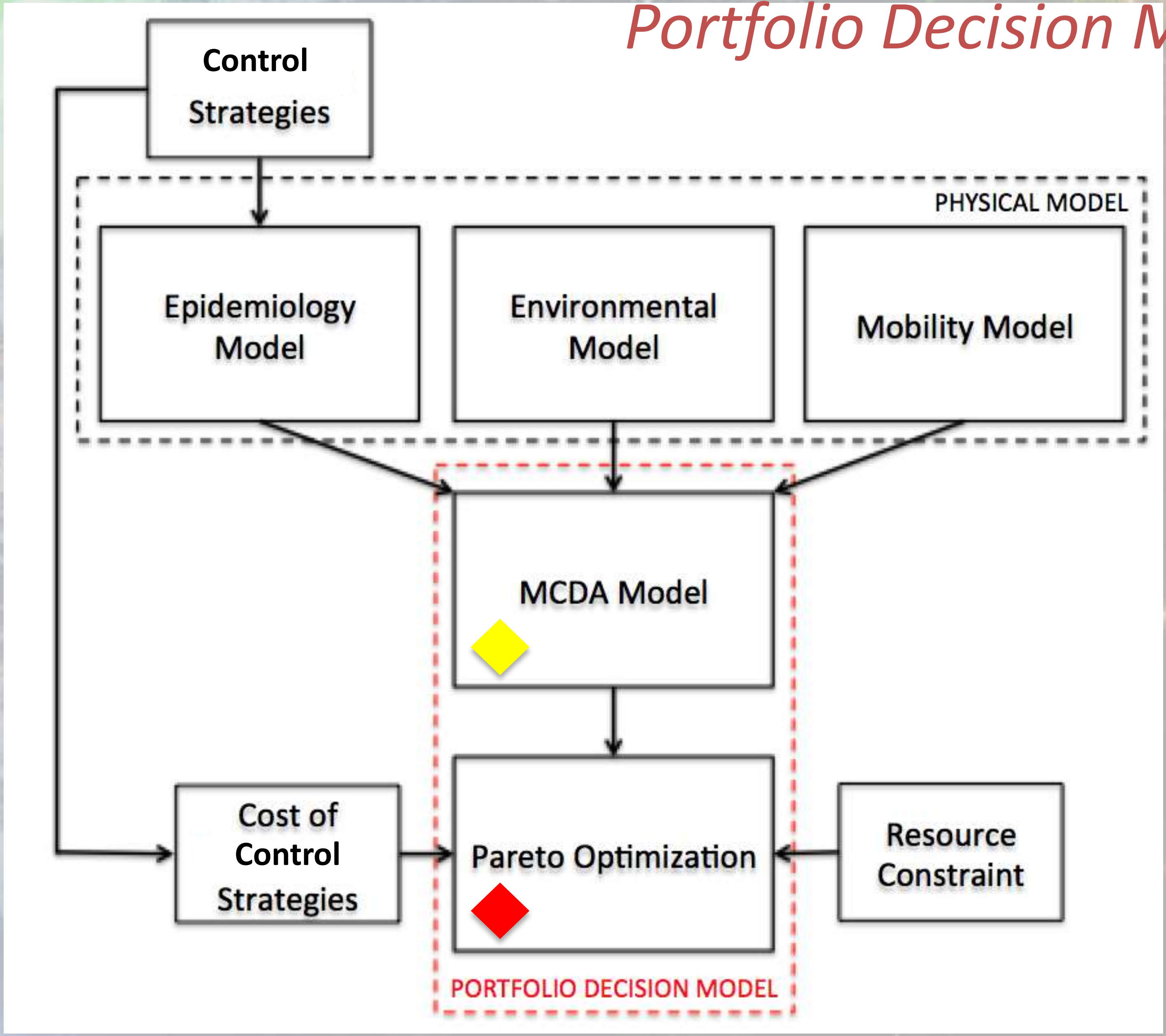
SanJuan, CDF of Weekly Incidence

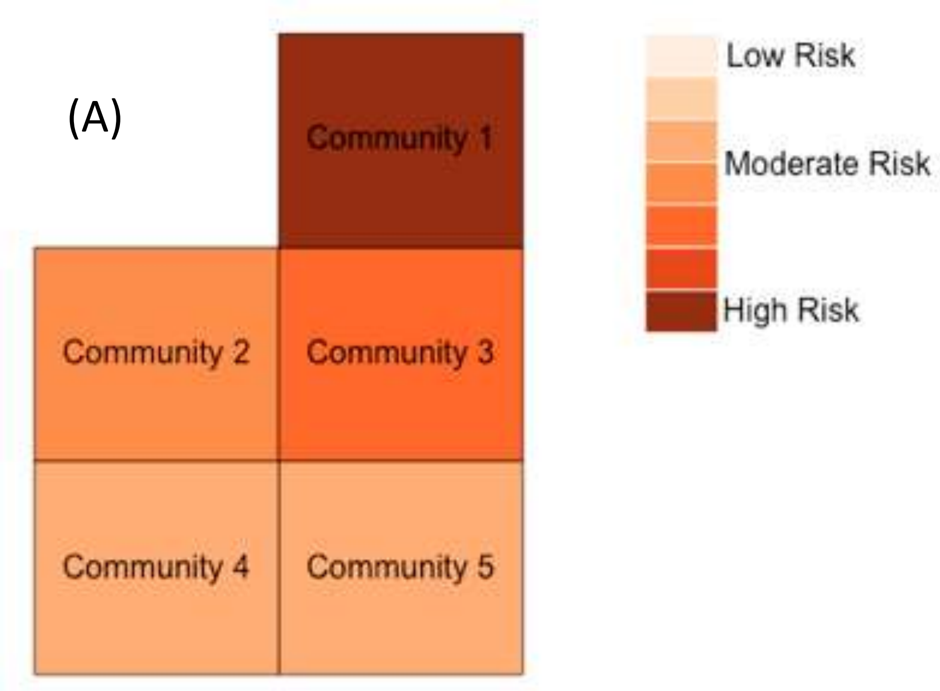


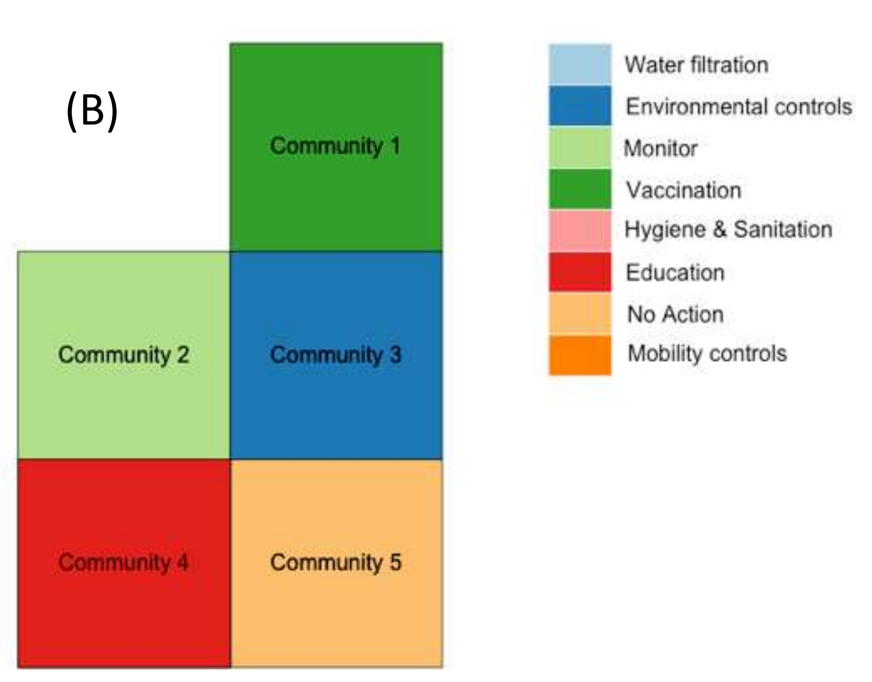
Worldwide country-scale forecasts

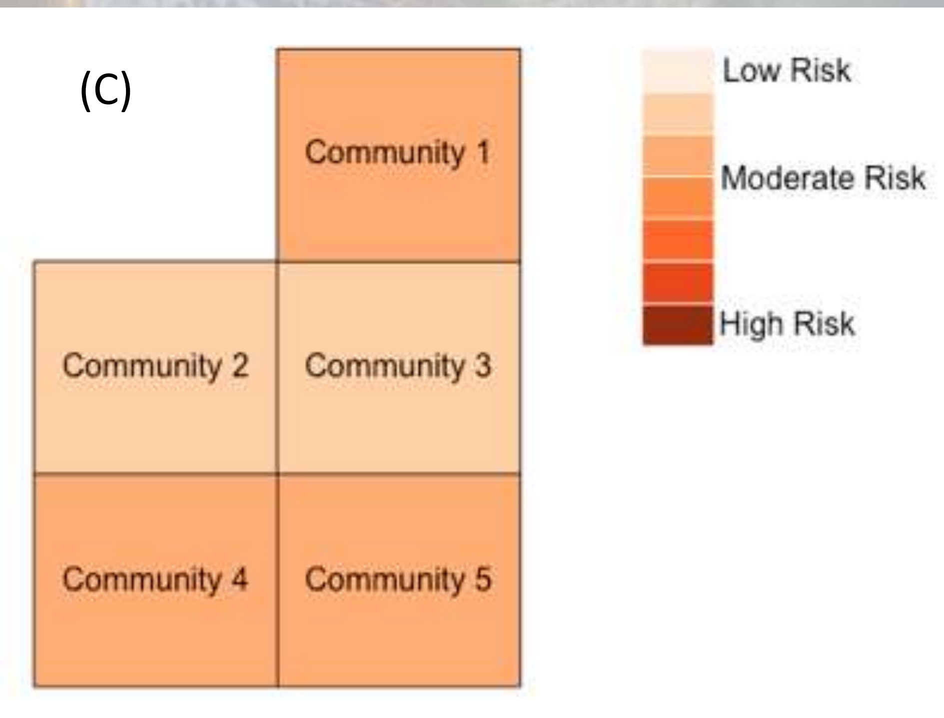


Portfolio Decision Model









Concept

(A) Output of the physical model: expected system outcome based on the epidemiology, environmental and mobility model
(B) Output of the portfolio decision model, selection of the optimal control set at the community scale
(C) Portfolio controlled solution: lowest systemic outcome (e.g. incidence).

Local Population-adjusted Risk

Alternative Effectiveness ~Efficacy

$$V^*_{m,j}(\underline{R}) = (1 - \boxed{v_j(\underline{R})})\ \boxed{f_{i(j)}} R_{i(j),m}\ V_{m,j}(\underline{R})$$

Population Vulnerability ~Urgency

(if available and meaningful)

Systemic Risk

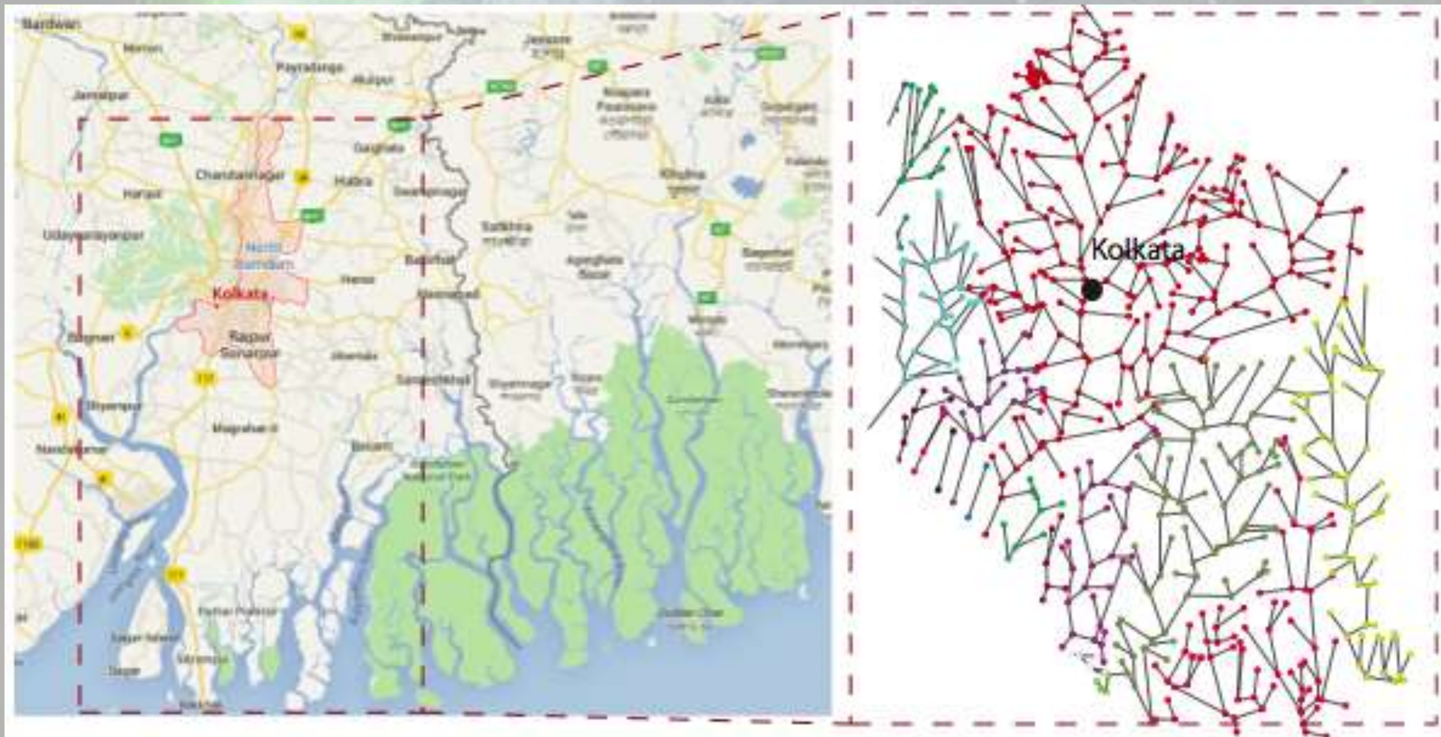
Stakeholder Preferences

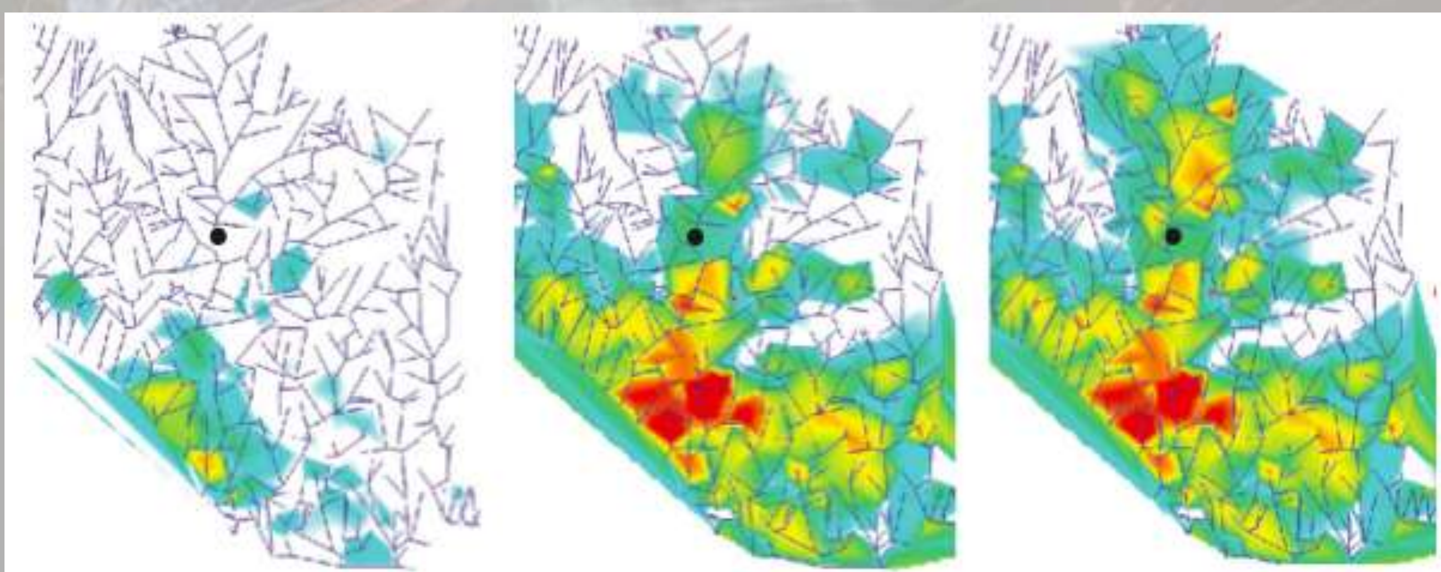
$$V_T(\underline{R}) = \sqrt{\sum_{m=1}^{M} \sum_{j=1}^{J} (V^*_{m,j}(\underline{R})\ \boxed{w_j})^2} = $$

$$= \sqrt{V_N(\underline{R})^2 + V_H(\underline{R})^2}$$

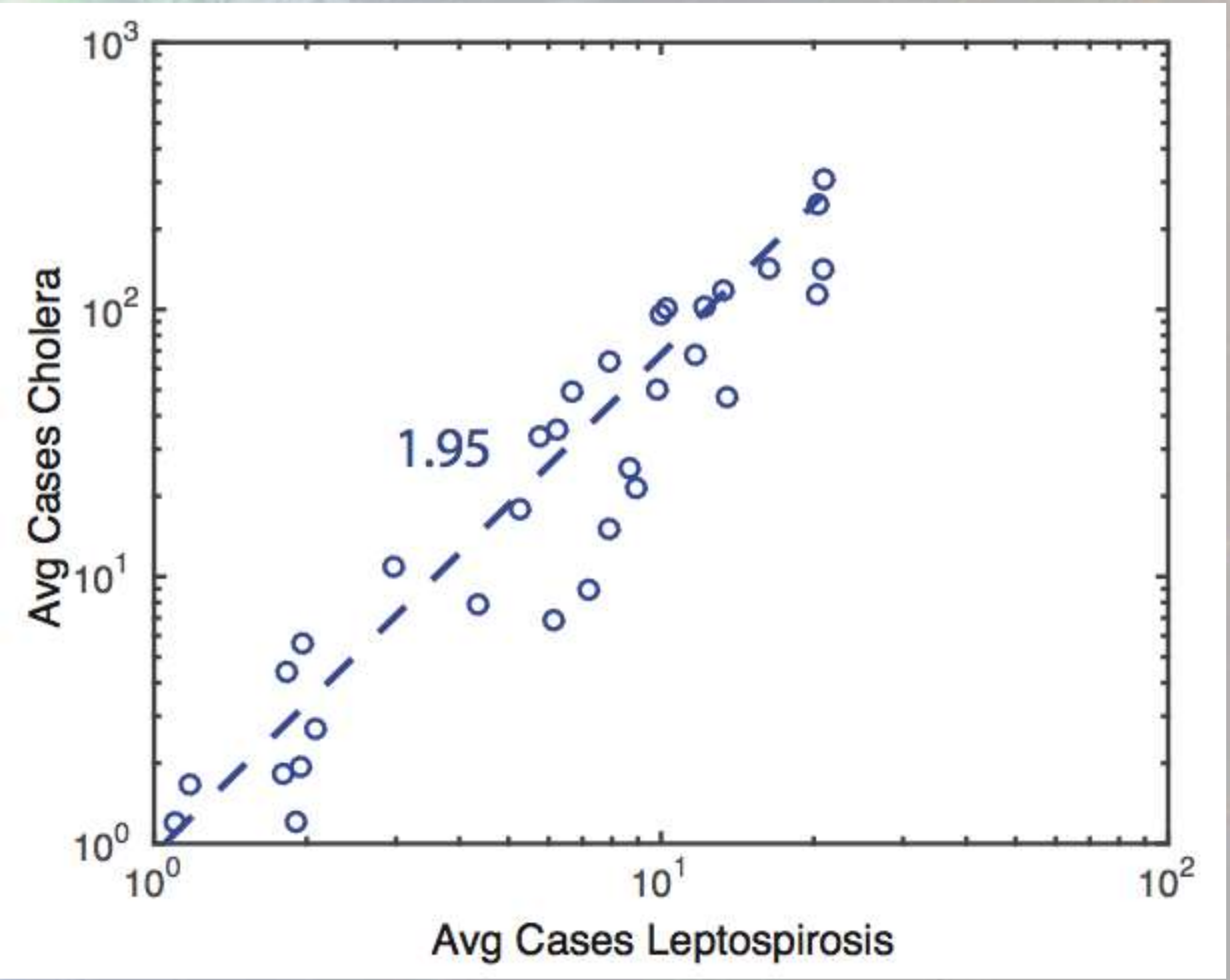
i=disease management alternative
j=target population
m=site

Domain

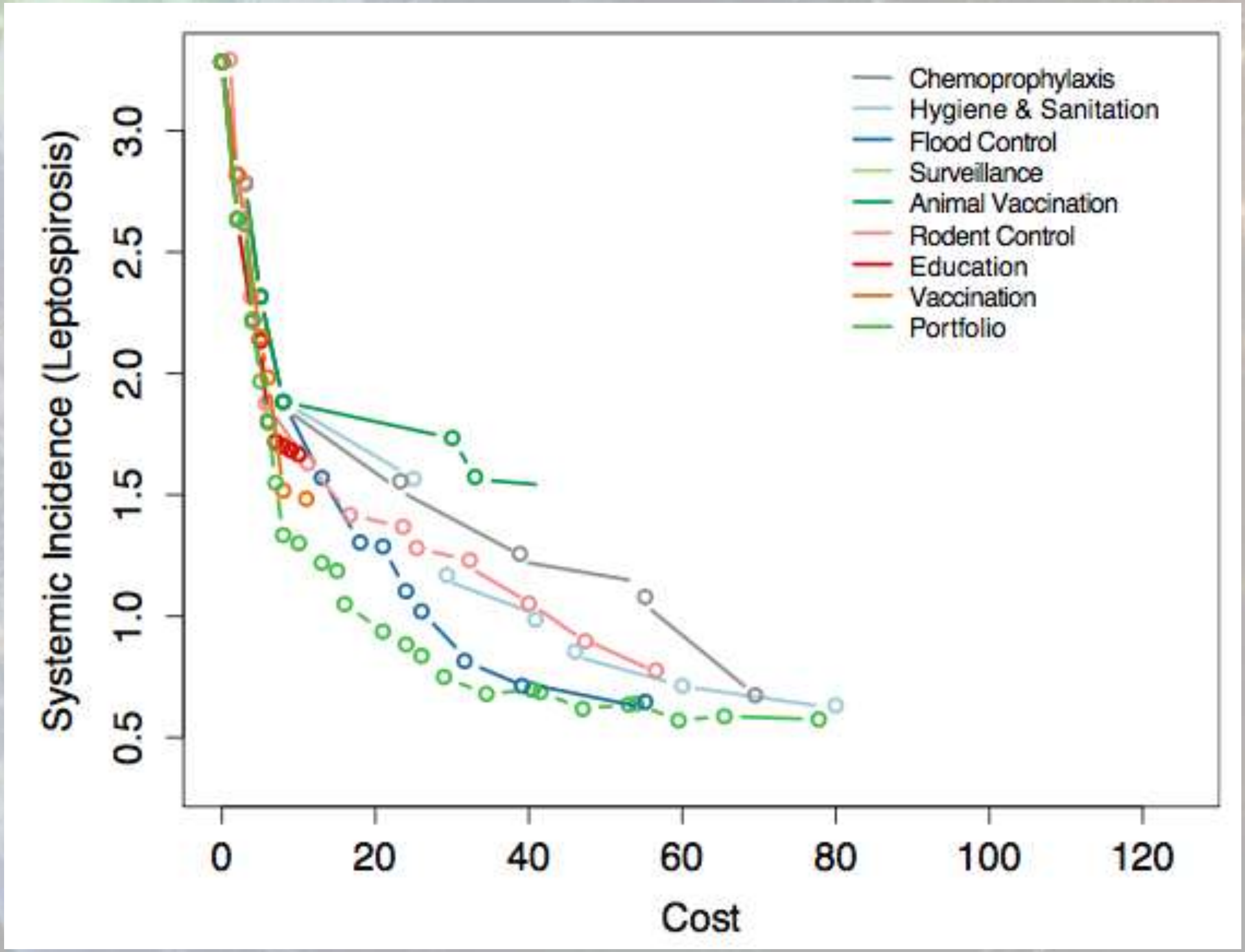




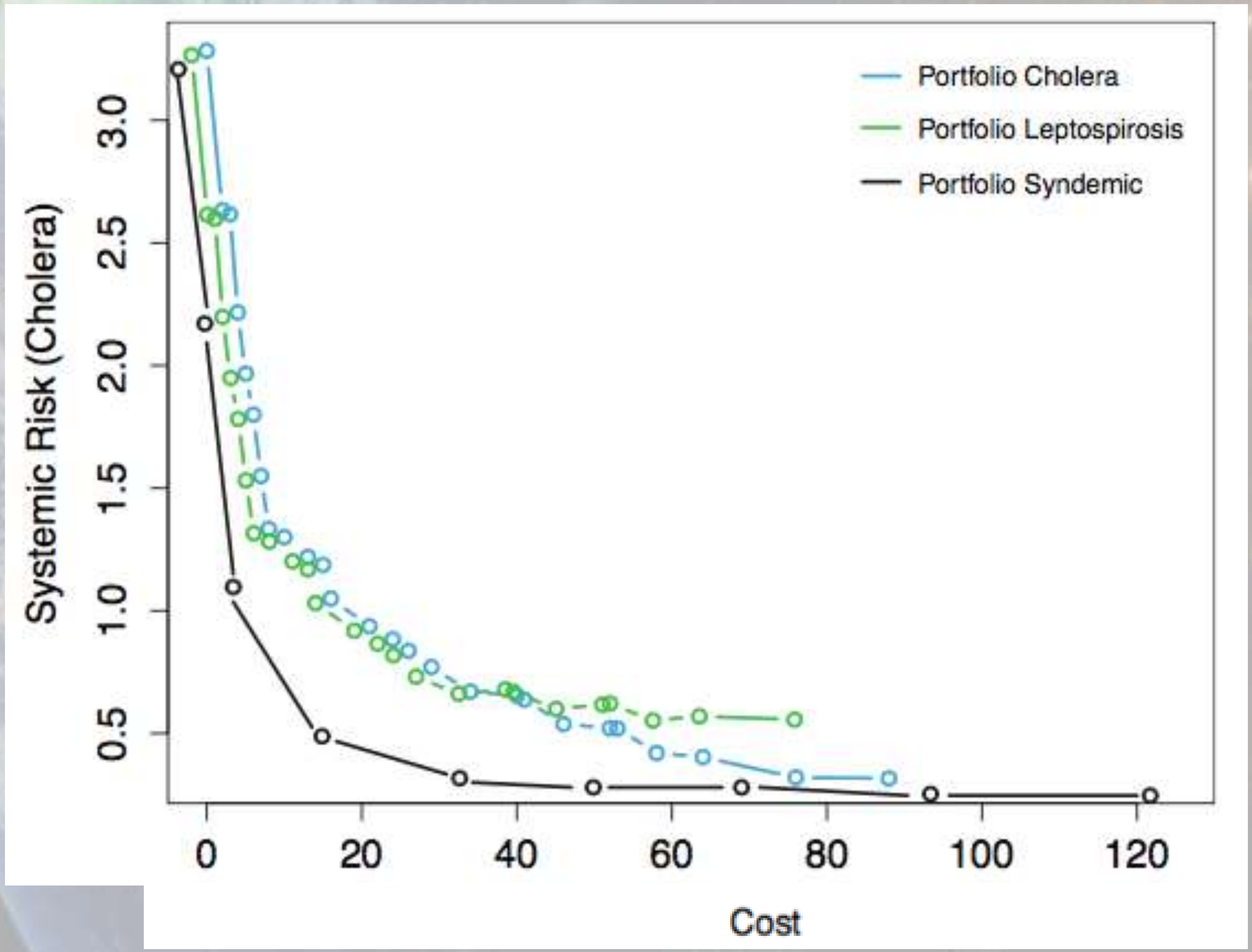
Syndemic Scaling



Optimizing One Disease Management



Optimizing Syndemic Management



Optimization of …

Analysis, with Theoretical Models as Macroscopes that look at *scaling and universality* of emerging disease patterns lead by the interaction of fundamental factors (Finite Topologies/Attractors exist in Nature!)

Predictive/Forecasting Models which should be as simple and accurate as possible – tight to the objective of study – versus fully mechanistic, complex and demanding models

Predictability of ''Unknowns'' (e.g., Low Probability High Consequence Events) for the identification of tipping points and potential future states

Models as technology to design the future (the environment) rather than just predicting the most likely one because it is more likely (and useful) to design an optimal future by *embracing the full uncertainty* of the status quo and the range of possibilities *-> the best way to predict the future is to design it*

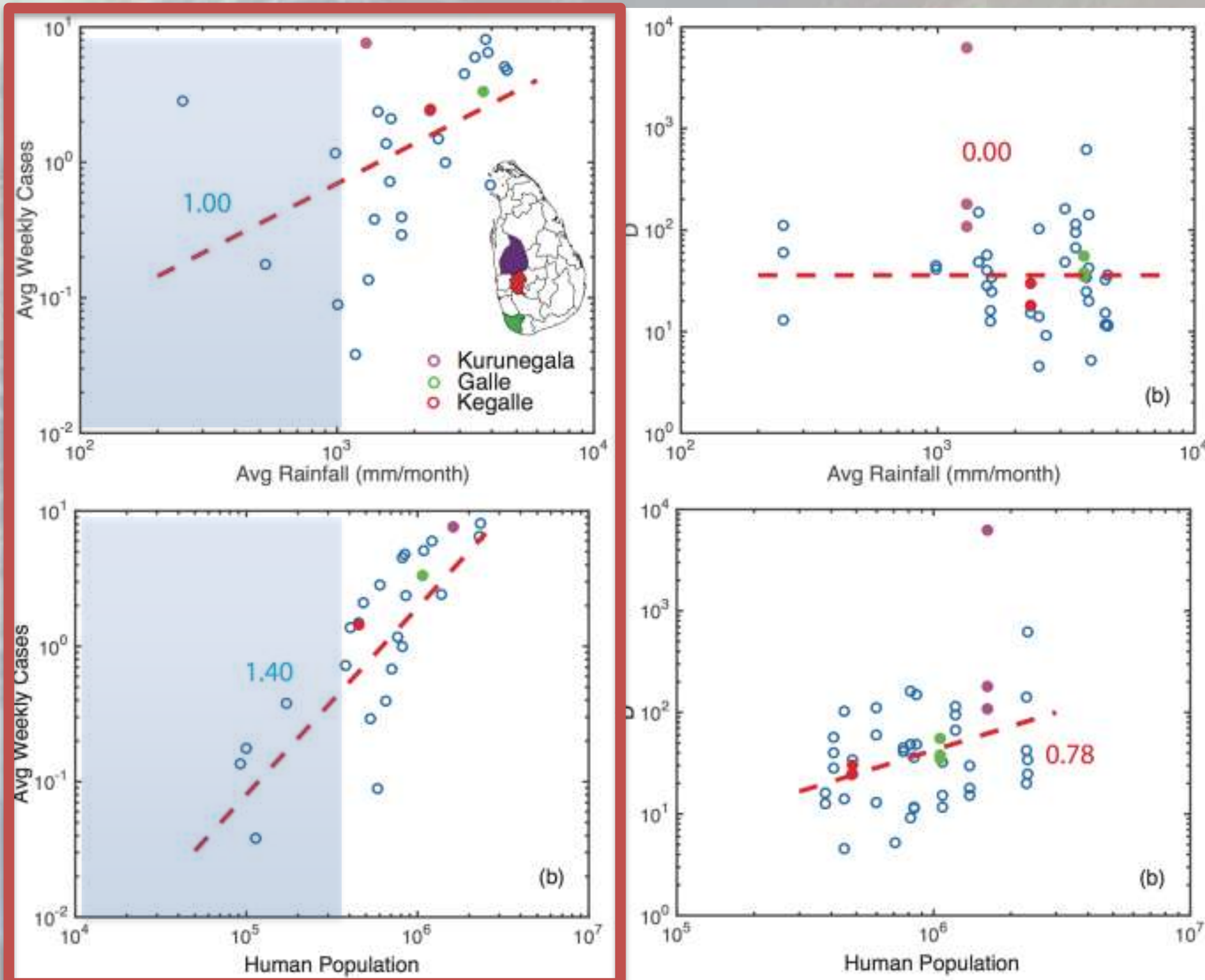
... and Simplicity

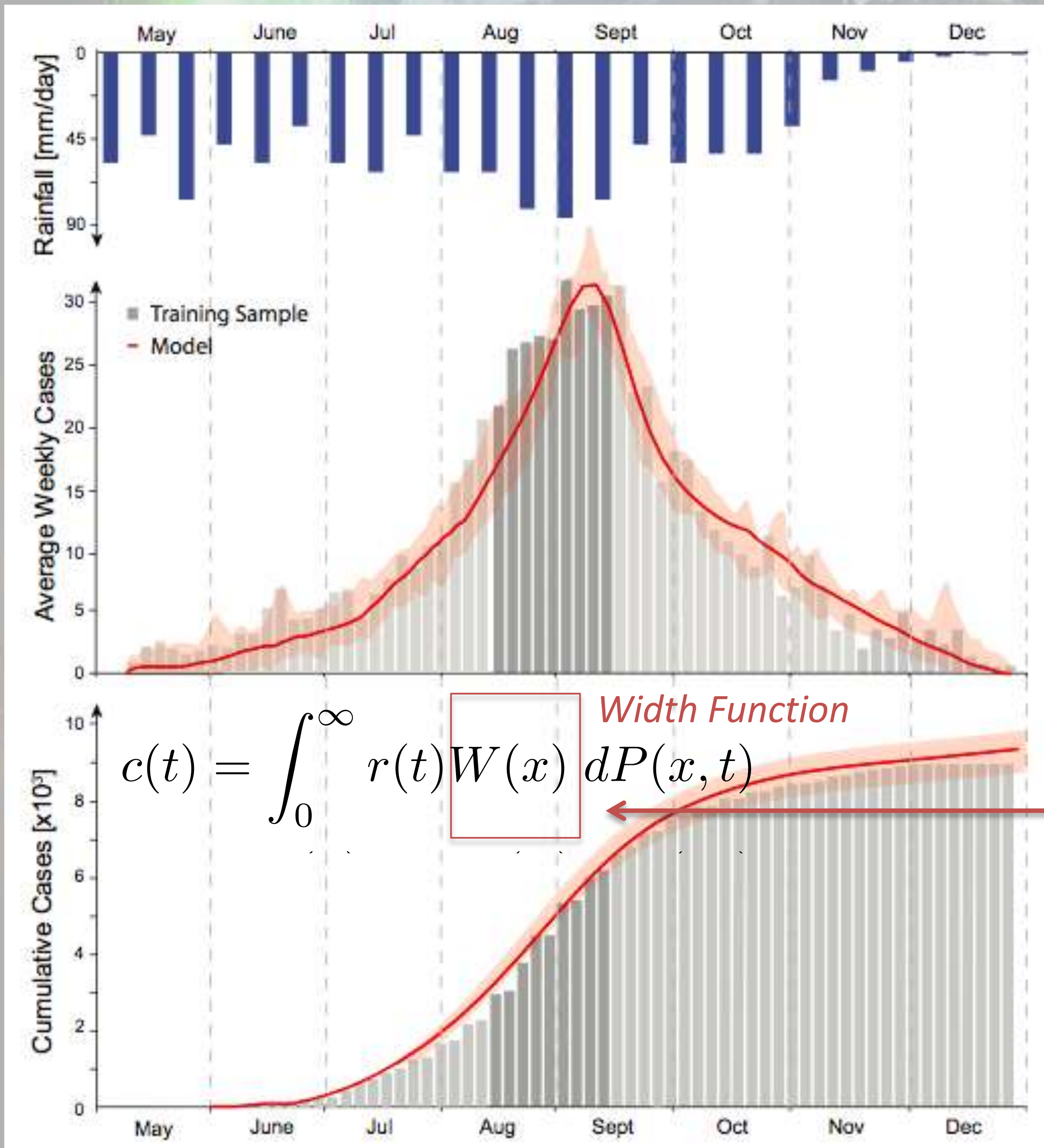
''Simplicity is the ultimate sophistication … and the solution of the complex nature''
Leonardo Da Vinci

Thanks!

Yang Liu, PhD Candidate
Matteo Convertino, PhD Dr.Eng.

Eco-epidemiological Scaling

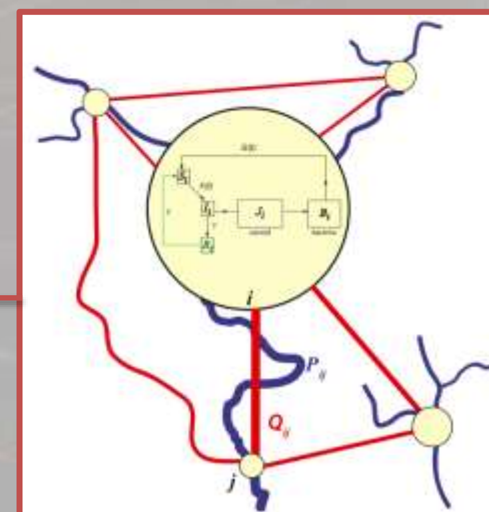




Width Function

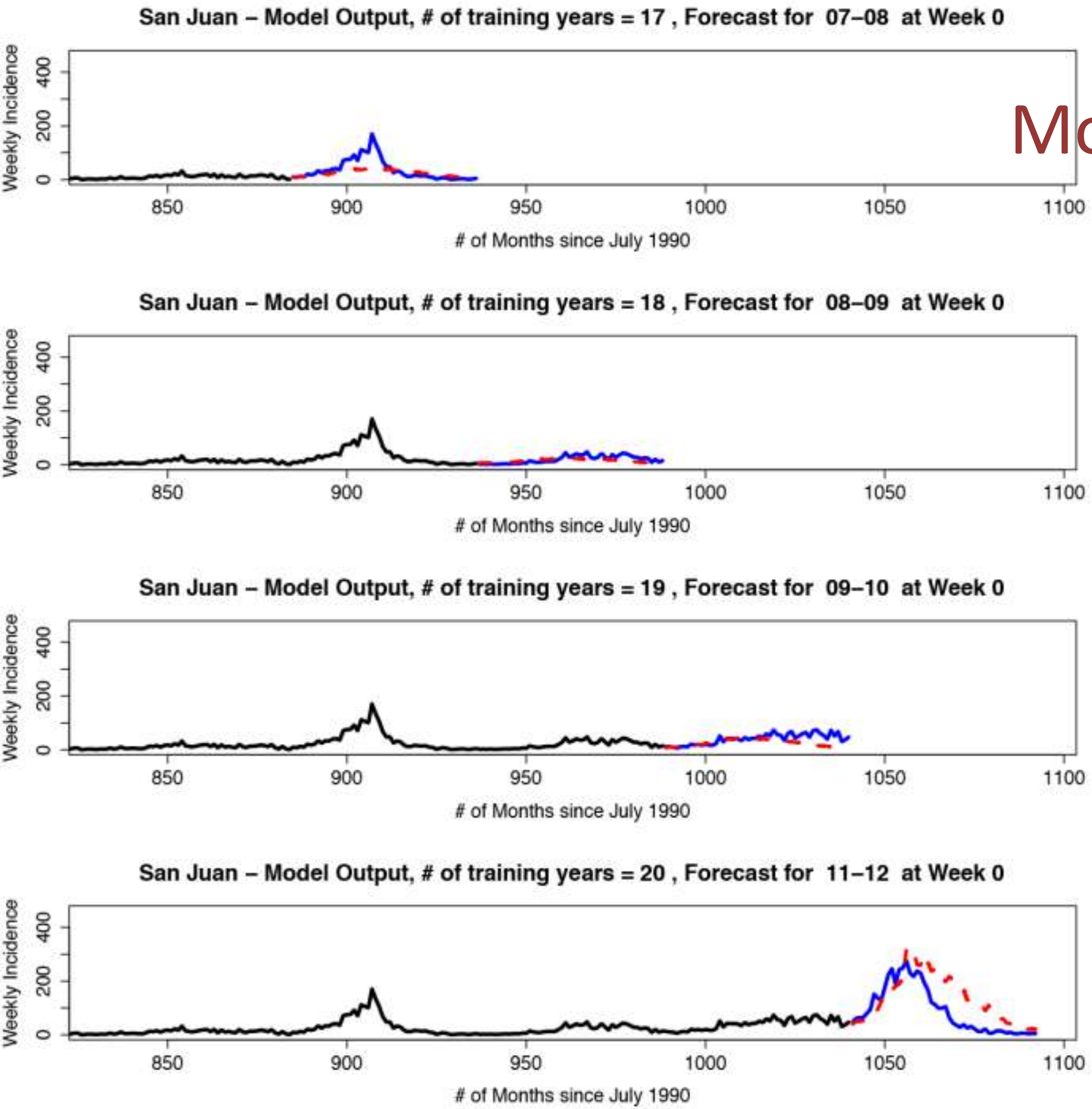
$$c(t) = \int_0^\infty r(t)W(x)\,dP(x,t)$$

W(x) only dependent on Network Topology (travel time distribution)

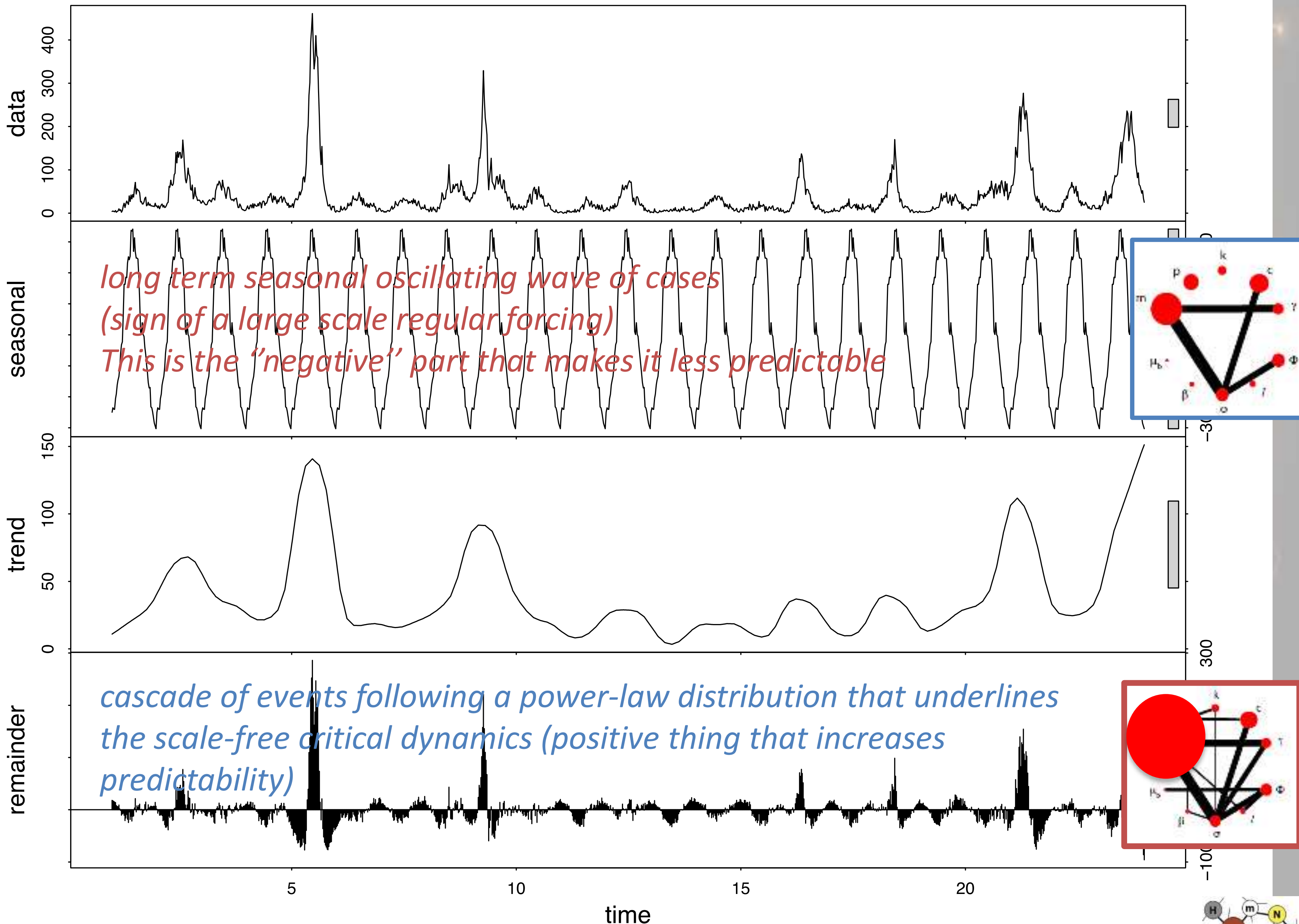


Multislice Network

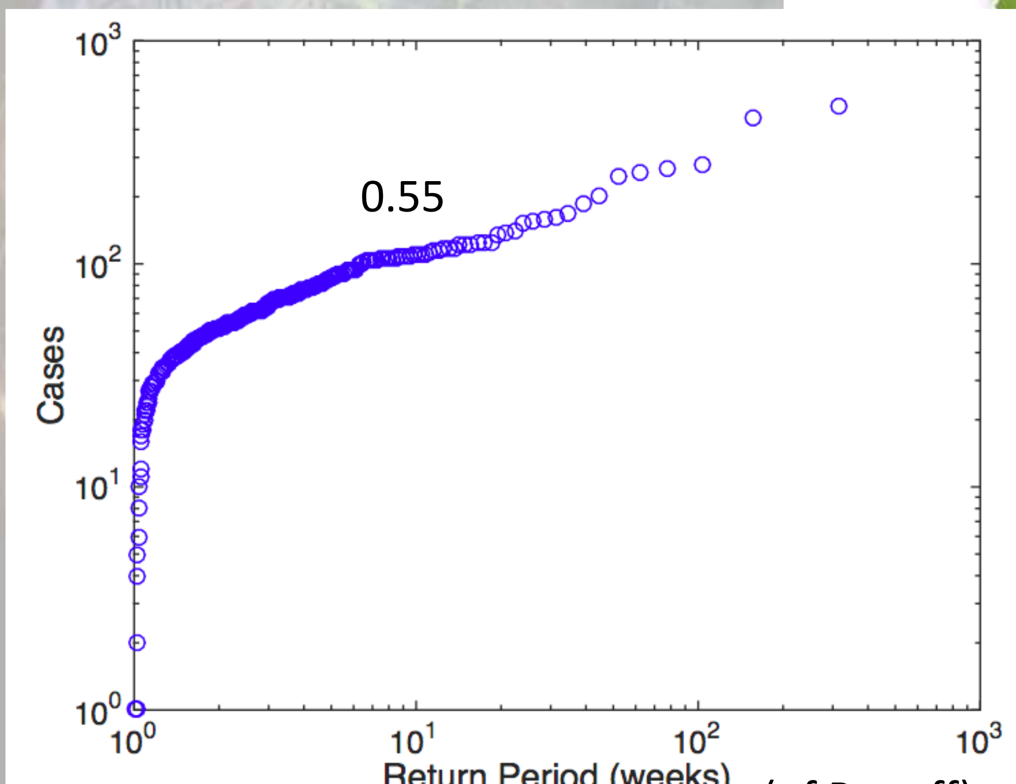
Zoomed Model Forecast



Supporting Information


data
seasonal
trend
remainder
time
long term seasonal oscillating wave of cases
(sign of a large scale regular forcing)
This is the ''negative'' part that makes it less predictable
cascade of events following a power-law distribution that underlines
the scale-free critical dynamics (positive thing that increases
predictability)

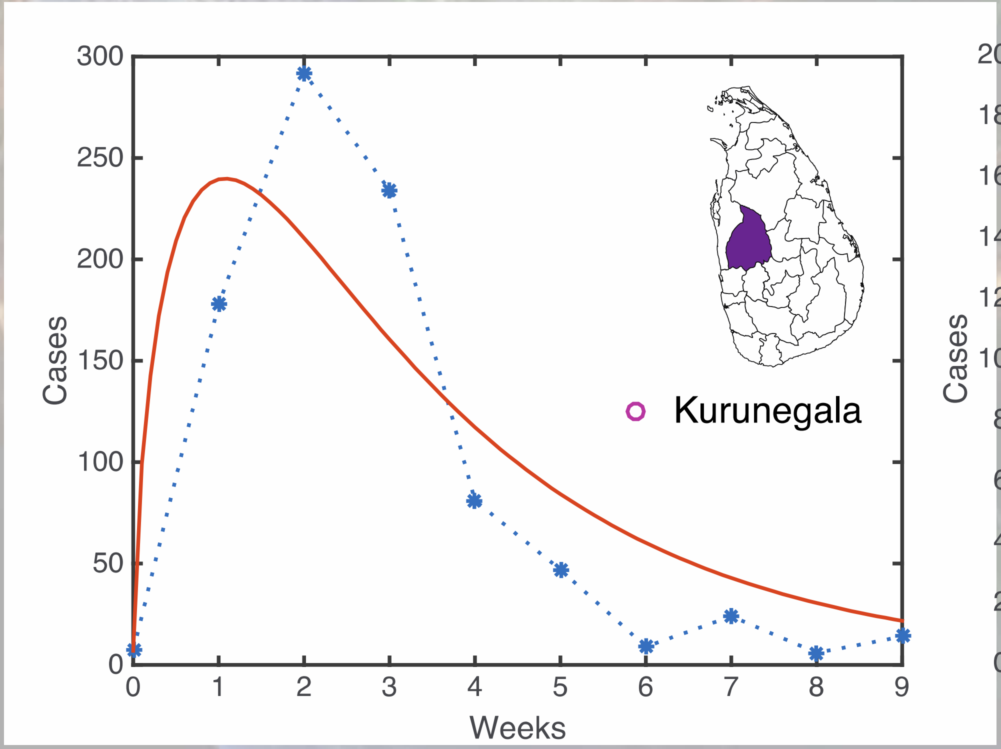
Scaling Epidemiology

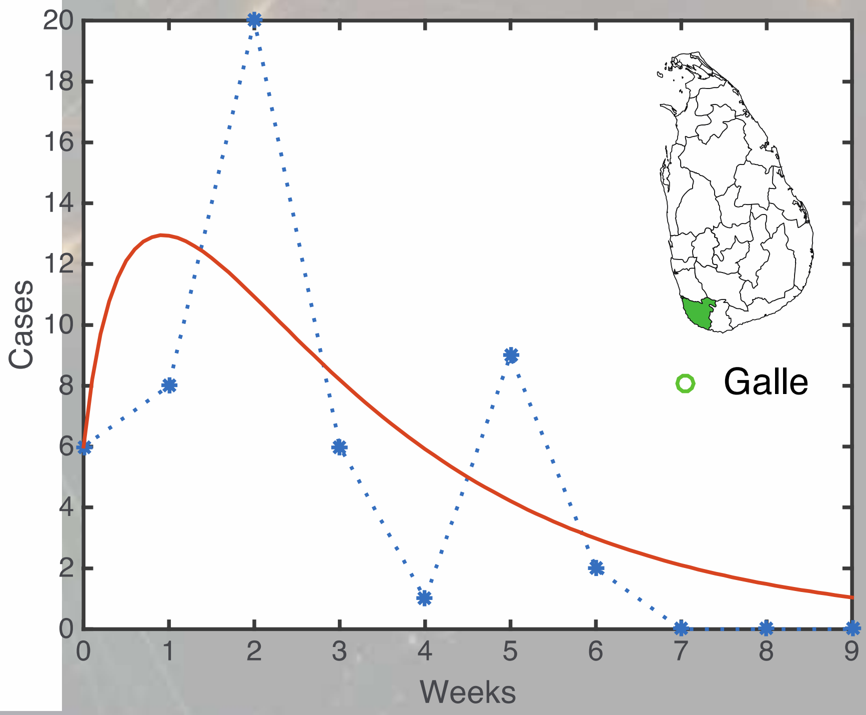


(of Runoff)

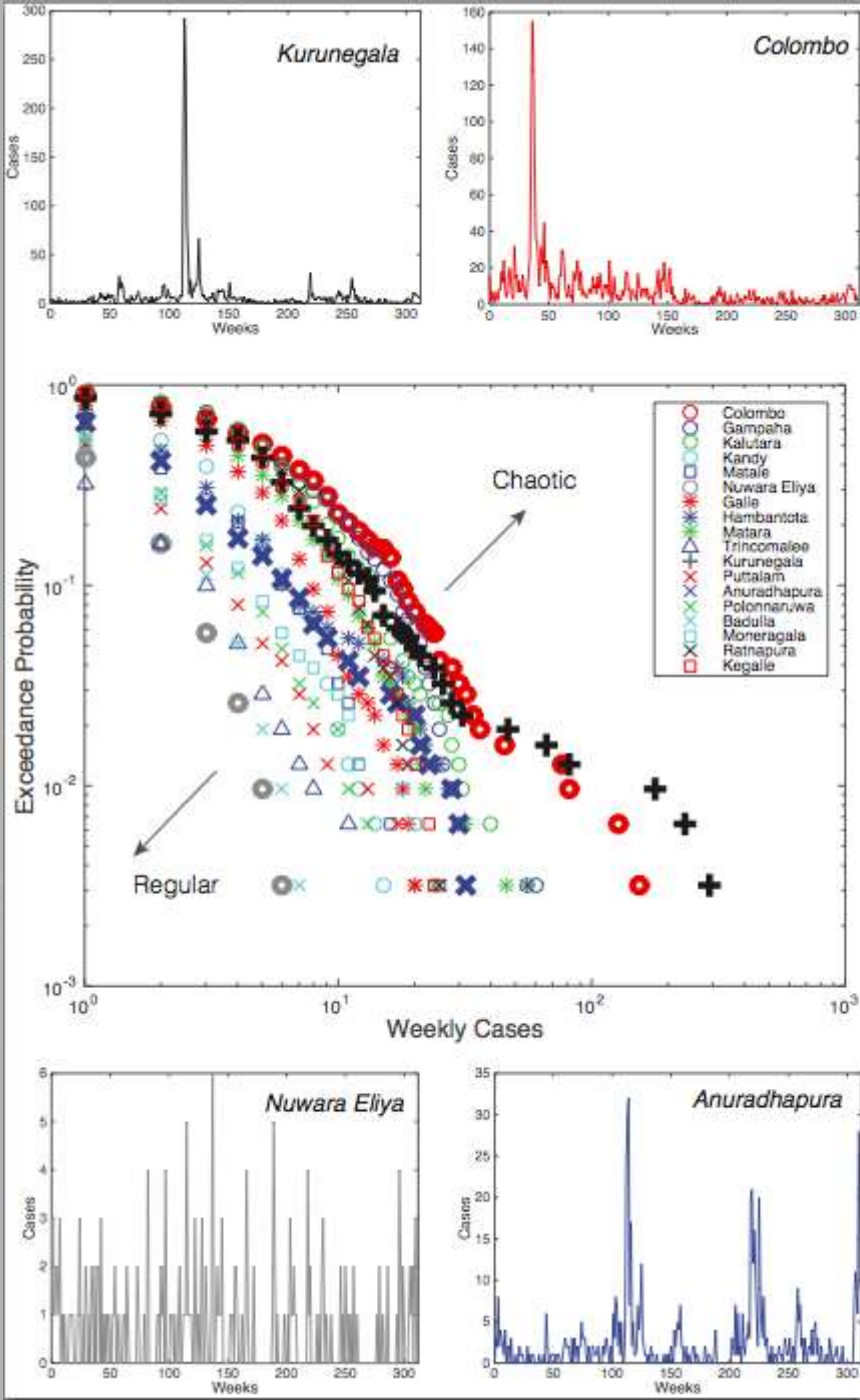
Model Calibration

- b, D, tau found using a least squares optimization for peaks across Sri Lanka

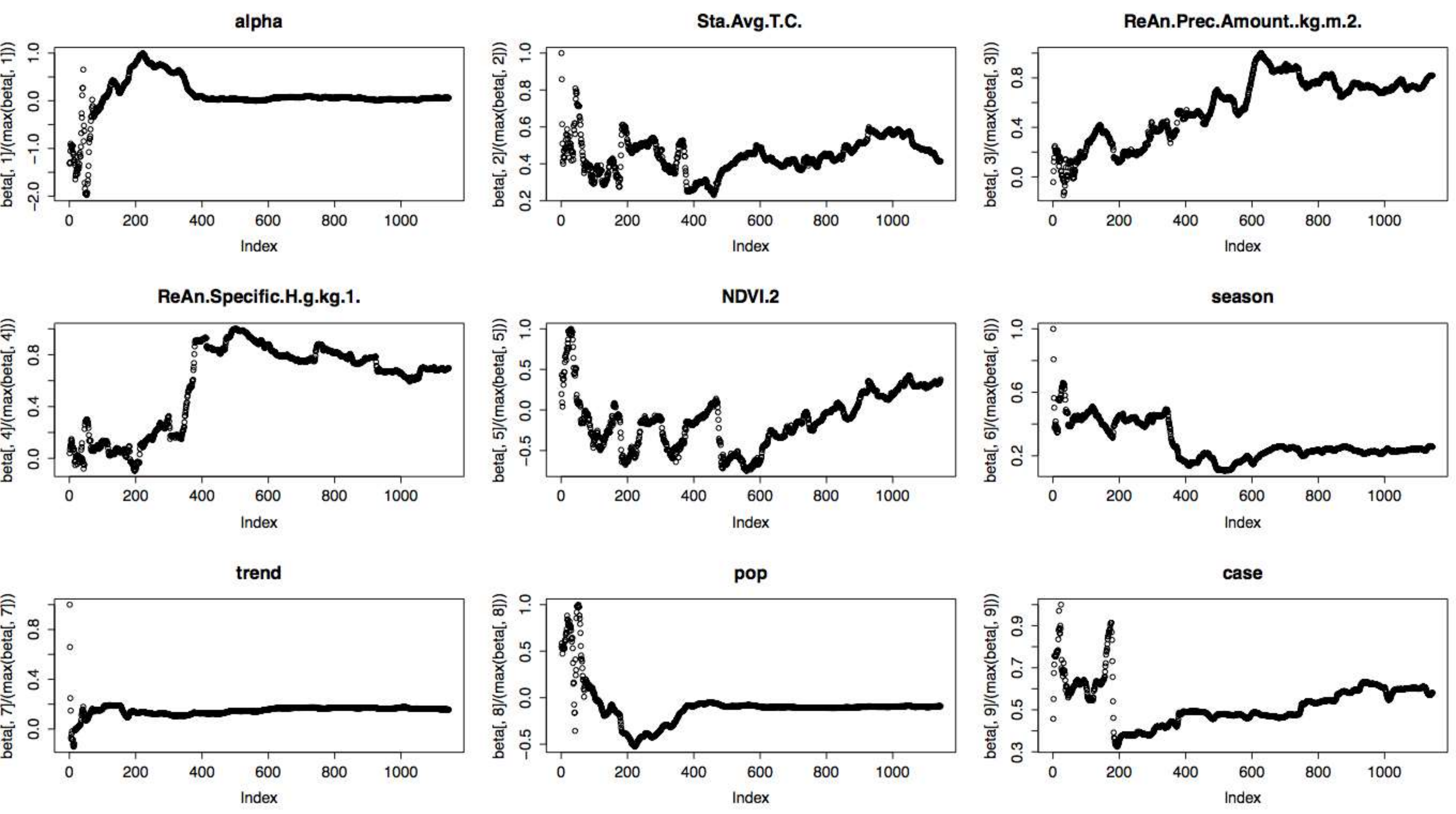




Distinct Signature of Complex System Dynamics



Time Dependent Importance of Variables



Topographic Index

- Steady state wetness index characterizing the ecohydrology of ecosystems

$$TI = \log\left(\frac{A_i/b_i}{\tan\beta_i}\right)$$

A_i=drainage area upstream a point

b_i=area per unit width orthogonal to the flow direction

β_i=slope

On the Return Time of Cases

- Exceedance probability is the likelihood to have and event greater than or equal to C

$$P(C \geq c).$$

- Return Period

$$T(C) = \frac{1}{P(C \geq c)}$$

Example return periods

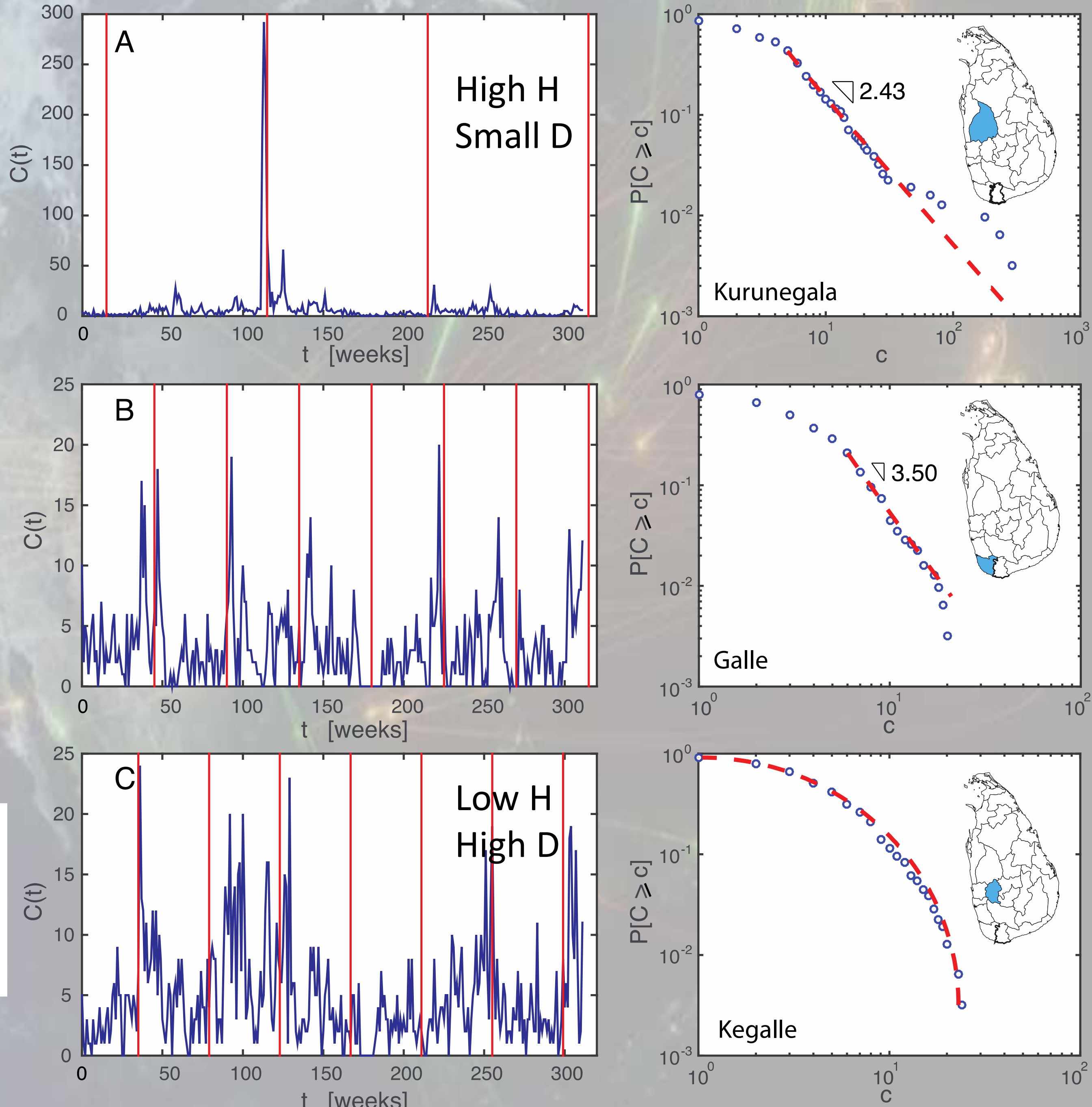
A) 115 weeks (>30)

B) 45 weeks (>15)

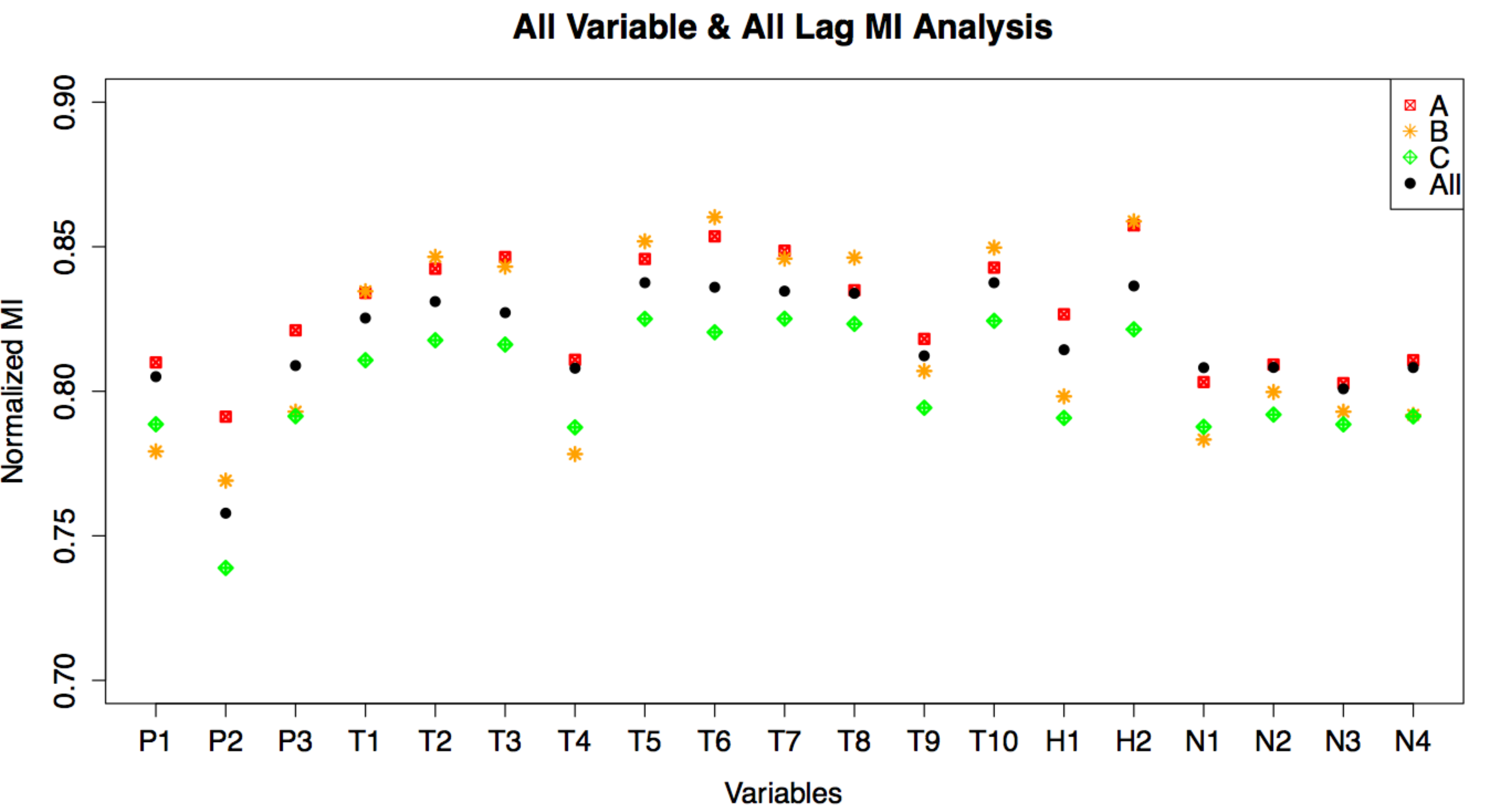
C) 43 weeks (>17)

$$P(C \geq c)$$

$$T(C) = \frac{1}{P(C \geq c)}$$

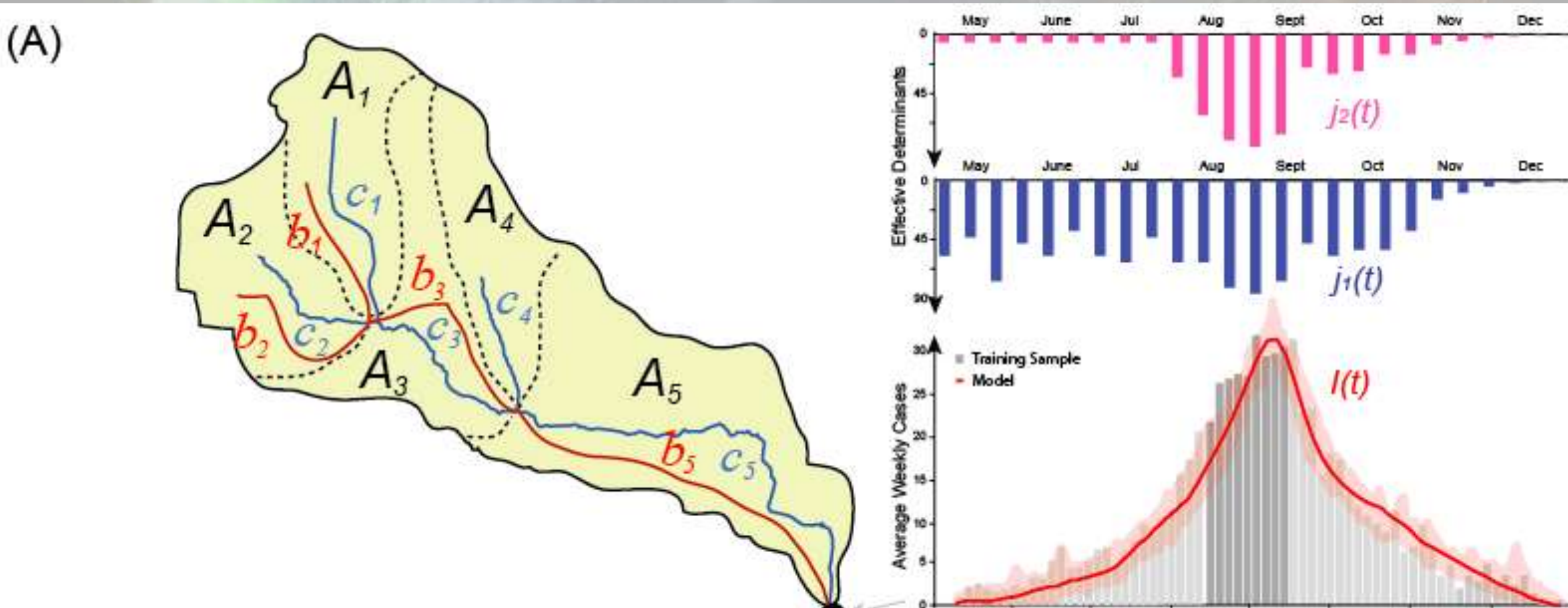


The How



All Variable & All Lag MI Analysis
Normalized MI
0.90
0.85
0.80
0.75
0.70
A
B
C
All
P1
P2
P3
T1
T2
T3
T4
T5
T6
T7
T8
T9
T10
H1
H2
N1
N2
N3
N4
Variables

On the Morphological Effective Systemic EpiGraph (MESE)



Novelty in the approach:

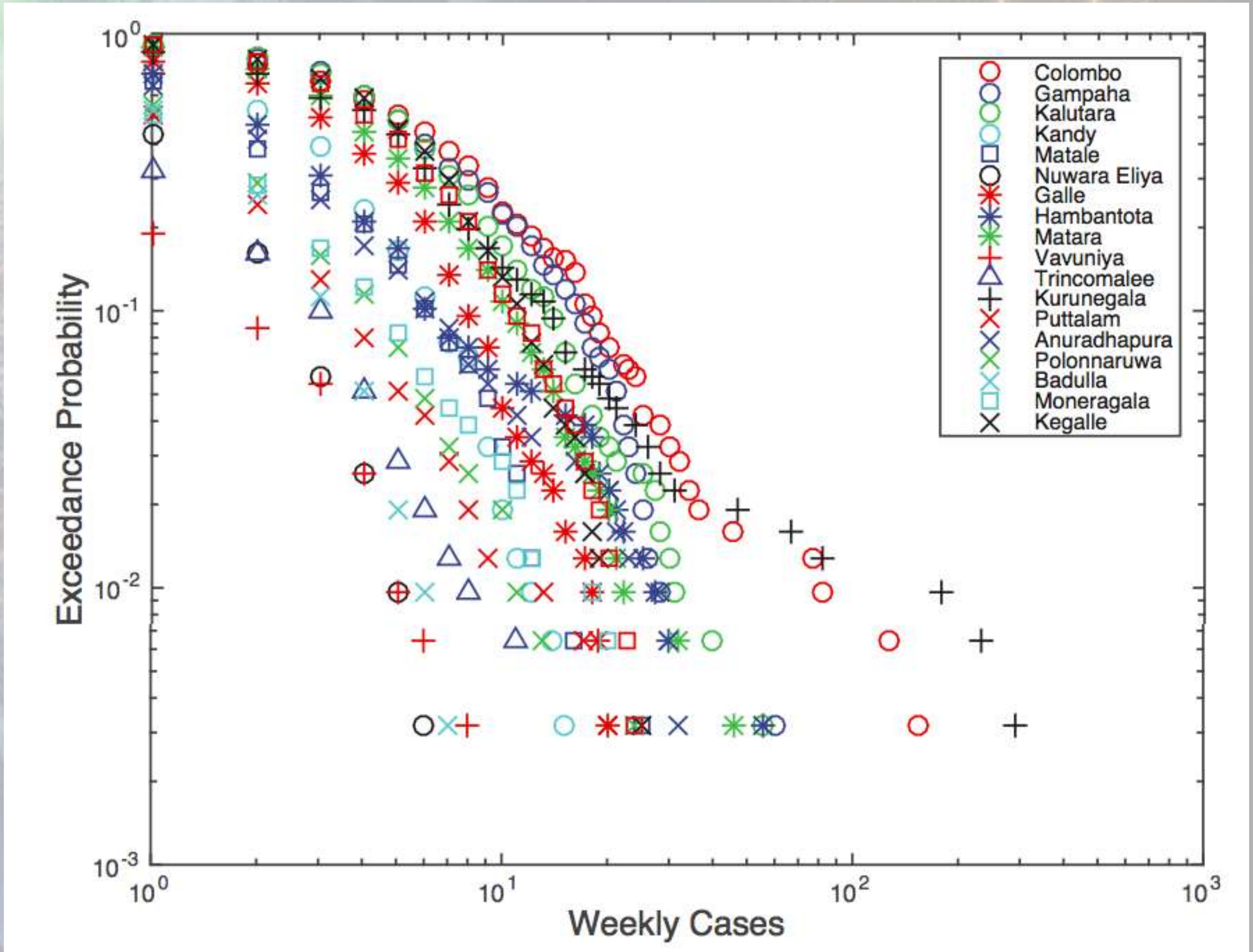
- Uncertainty and multiplicity in transmission routes and disease determinants
- Bidirectional fluxes on transmissions
- Effective distances (related to effective velocities)

Novelty in the Epi:

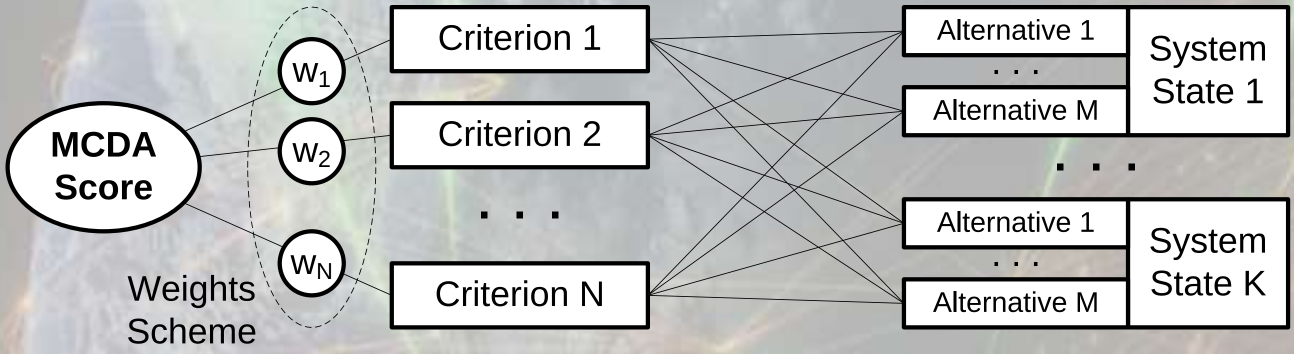
- Morphology contribution of disease production
- Time delay
- Factor interactions

Convertino, Huang, Liu, 2014, WRR, submitted

Epdf reflecting disease dynamics and transitions



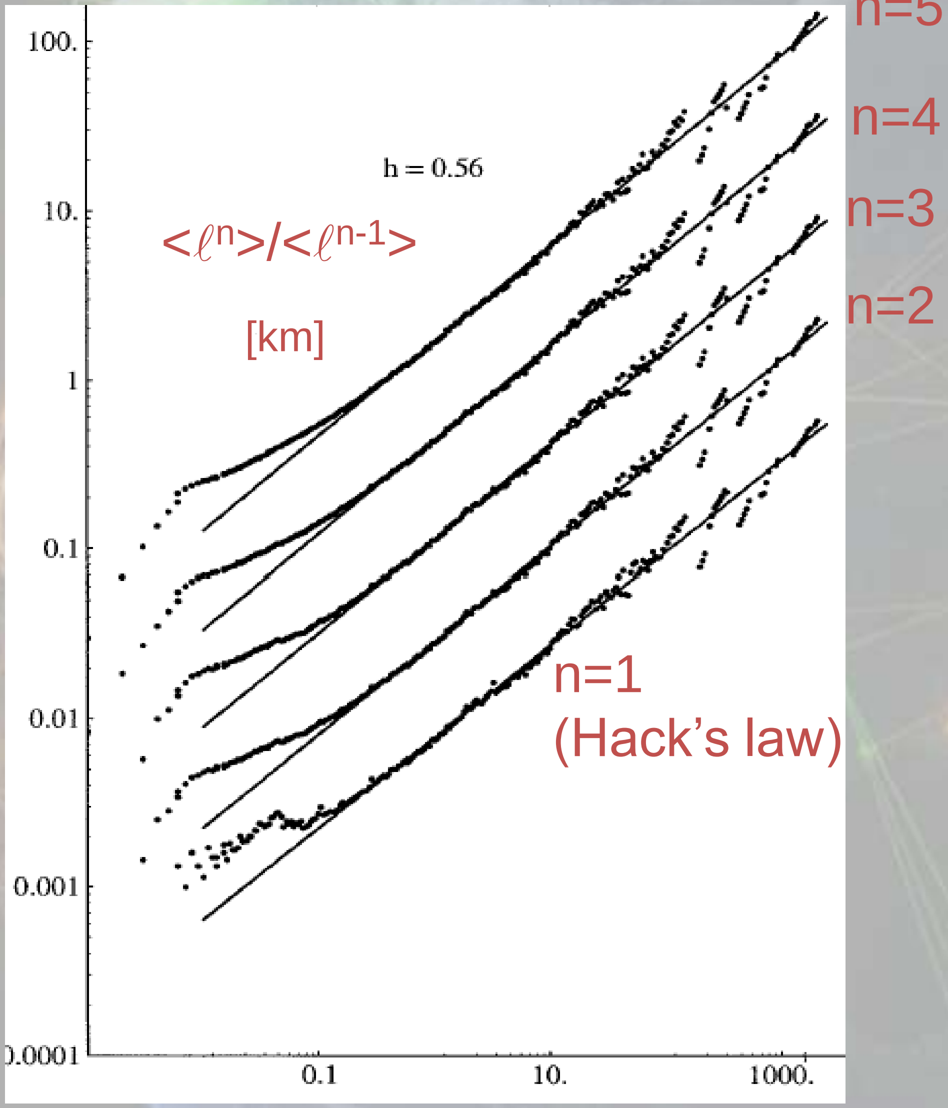
MCDA



Alternatives	Criteria	System States			
		State 1	State 2	. . .	State K
Alternative 1	Criterion 1	$X_{1,1,1}$	$X_{2,1,1}$	. . .	$X_{K,1,1}$
	Criterion 2	$X_{1,1,2}$	$X_{2,1,2}$	. . .	$X_{K,1,2}$
	. . .	. . .	. . .	. . .	. . .
	Criterion N	$X_{1,1,N}$	$X_{2,1,N}$	. . .	$X_{K,1,N}$
. . .	. . .	. . .	. . .	$X_{k,m,n}$	. . .
Alternative M	Criterion 1	$X_{1,M,1}$	$X_{2,M,1}$	. . .	$X_{K,M,1}$
	Criterion 2	$X_{1,M,2}$	$X_{2,M,2}$	. . .	$X_{K,M,2}$
	. . .	. . .	. . .	. . .	. . .
	Criterion N	$X_{1,M,N}$	$X_{2,M,N}$	. . .	$X_{K,M,N}$

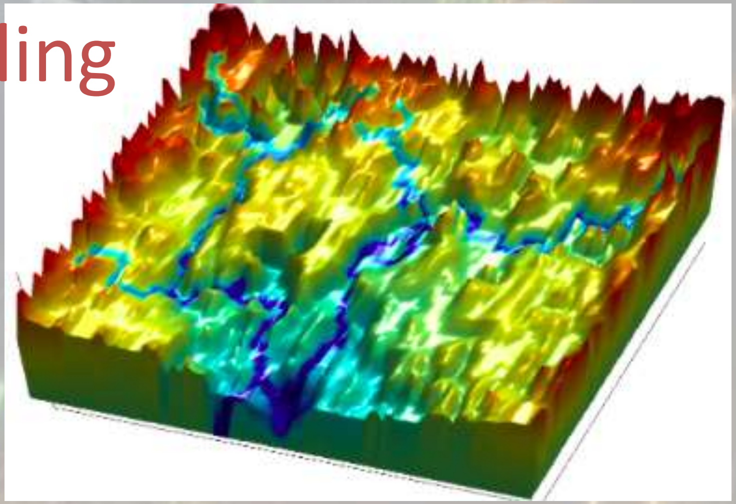
Supporting Information

Hydrological Networks and Scaling



Rigon et al., WRR, 1996

area A [km^2]



L = river's length
A = area
h = Hack's exponent

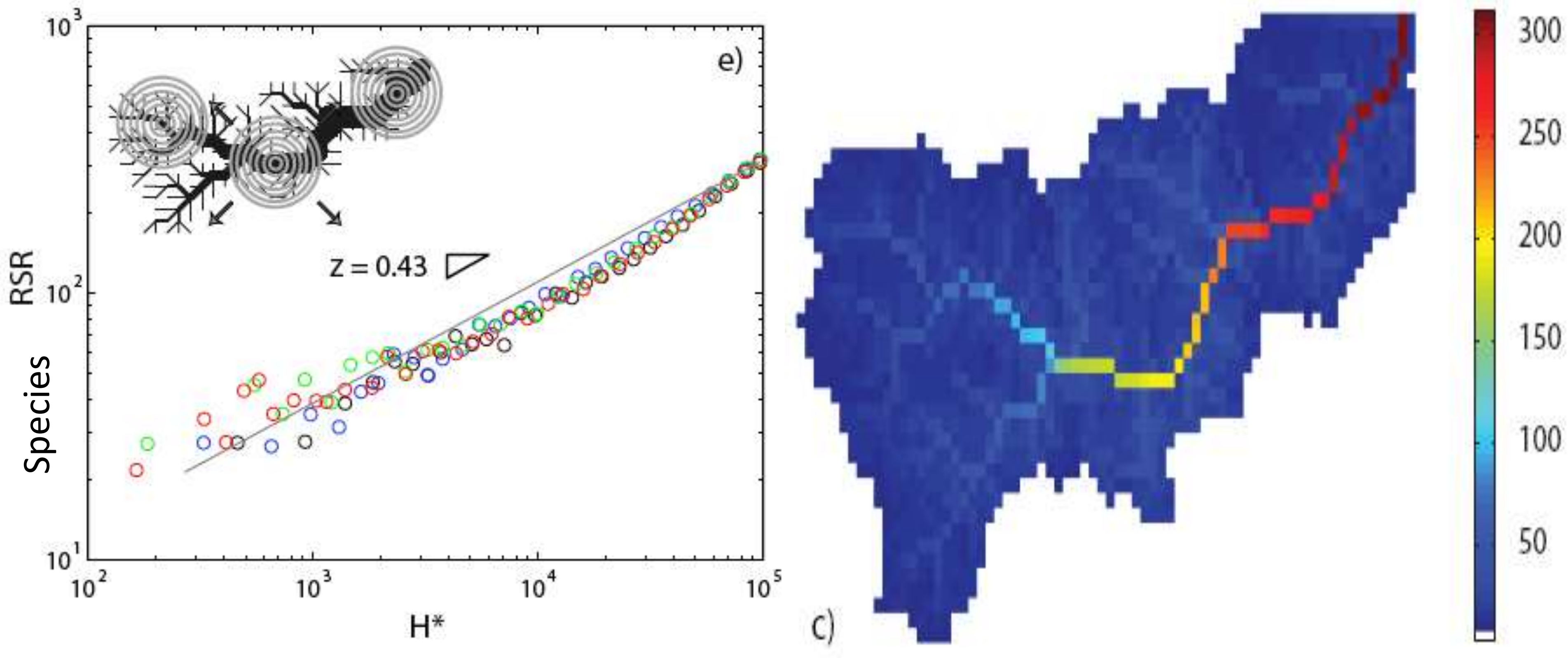
$< \ell^{n} > / < \ell^{n-1} > \propto A^{h}$

wonderful data show
that this is indeed the case over
several orders of magnitude

From Hydrology to Ecology

Species ~ Area ^{z}

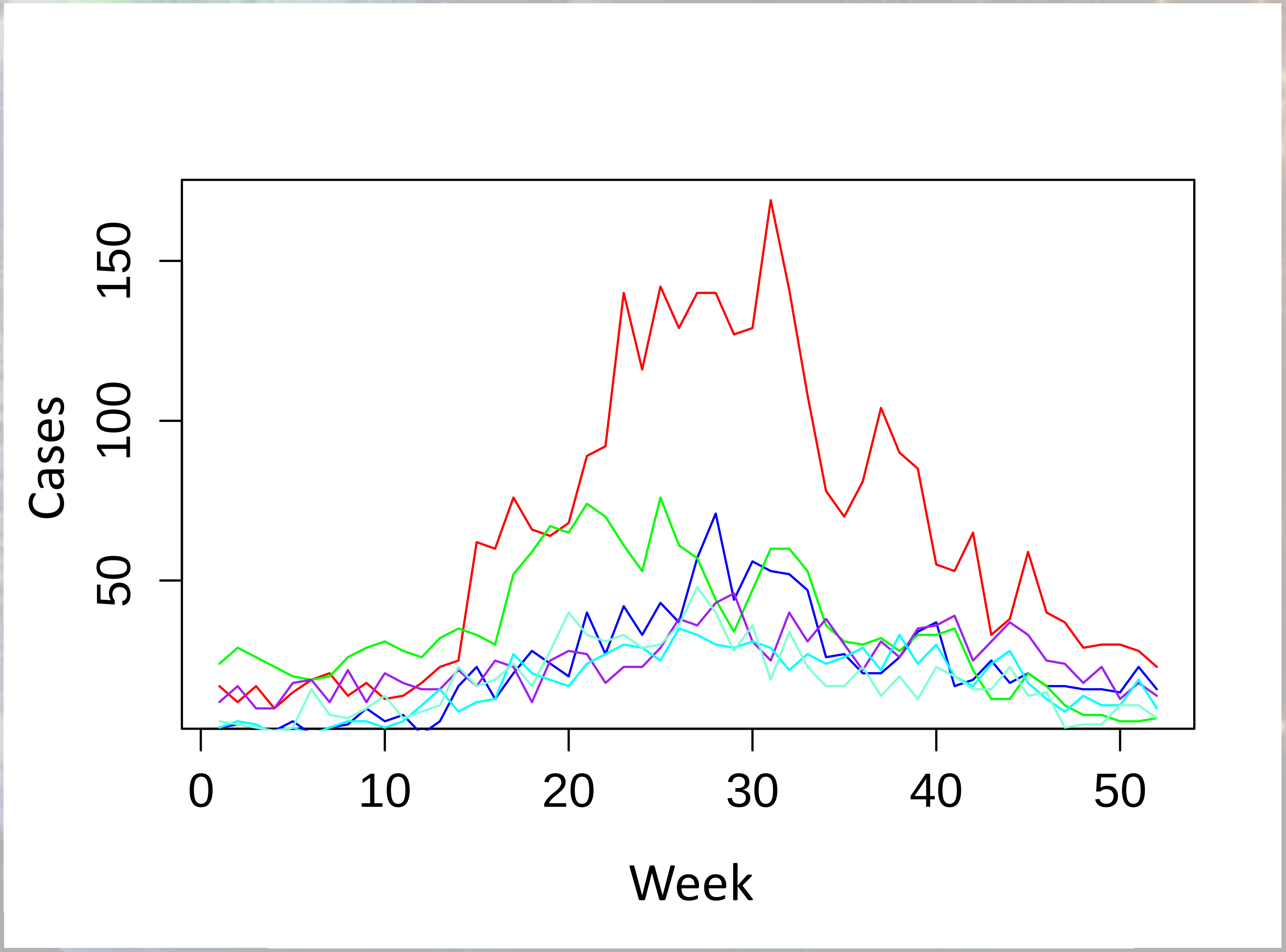
z = universal exponent independent of details of socio-ecological systems! (at stationarity)

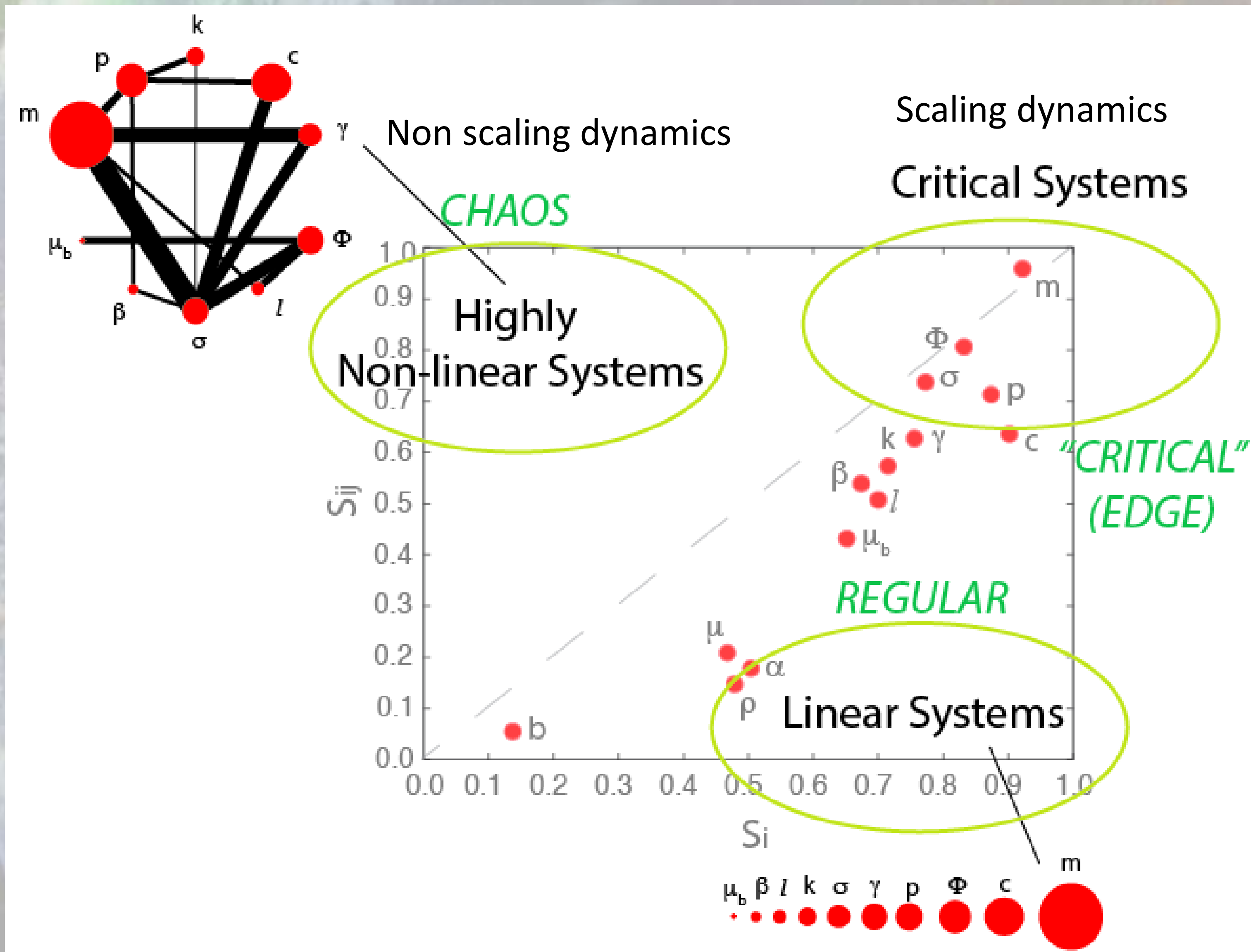


Species

Area

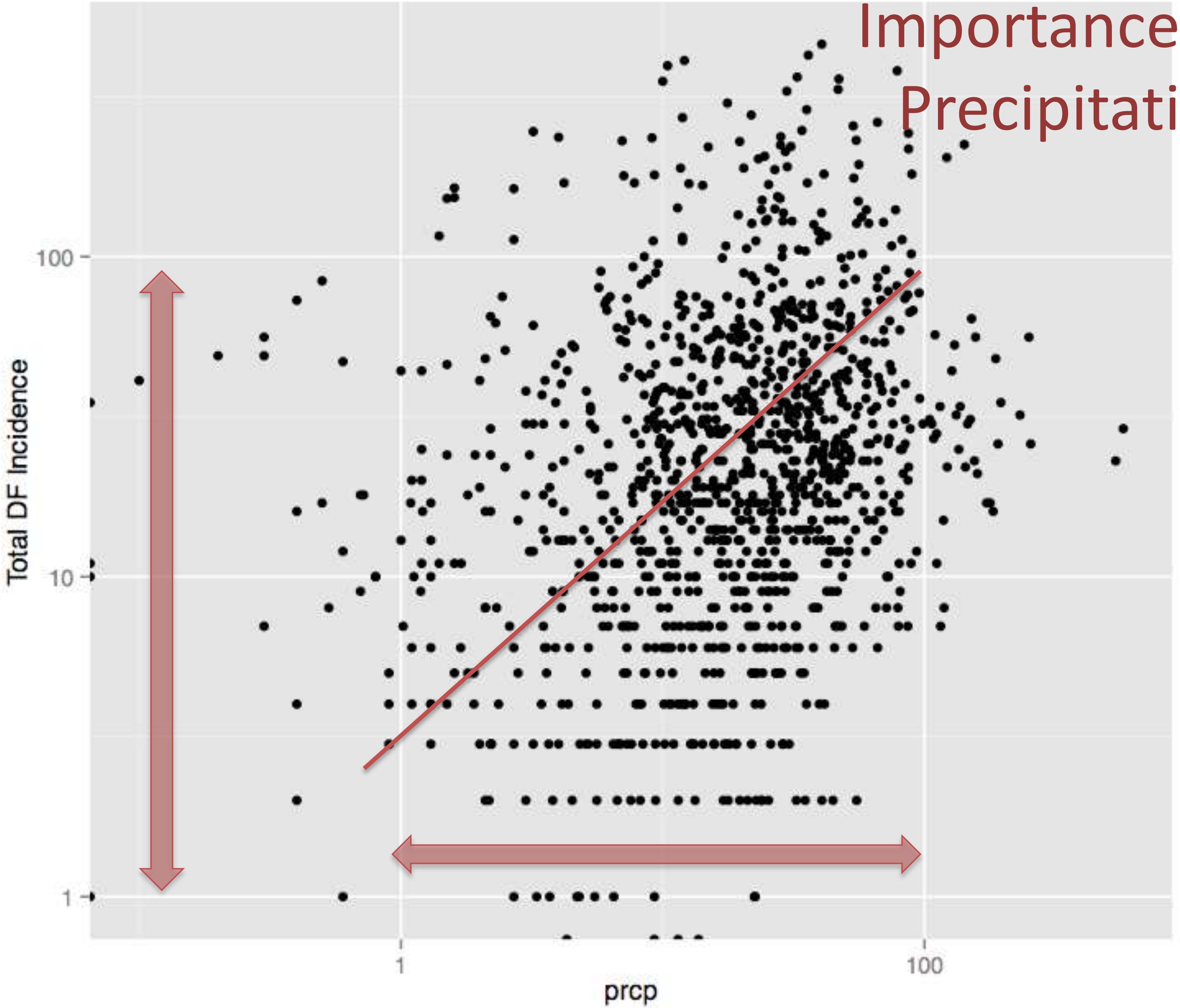
Overlap of Peaks as a Sign of Scale Invariance



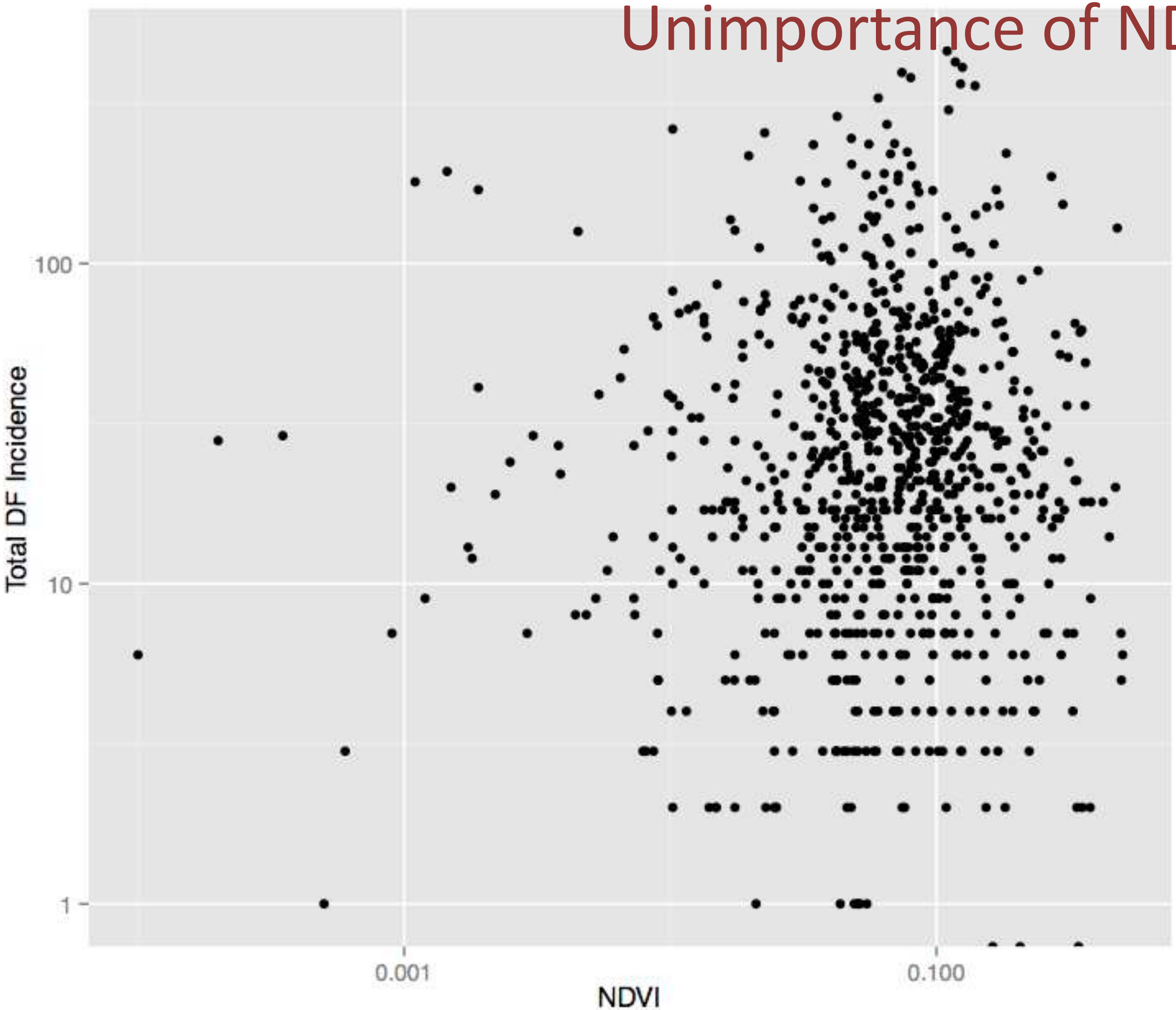

Non scaling dynamics
Scaling dynamics
Critical Systems
CHAOS
Highly
Non-linear Systems
"CRITICAL"
(EDGE)
REGULAR
Linear Systems
Sij
Si
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
m
Φ
σ
p
c
k
γ
β
l
μb
μ
α
ρ
b

Linearity

Importance of Precipitation



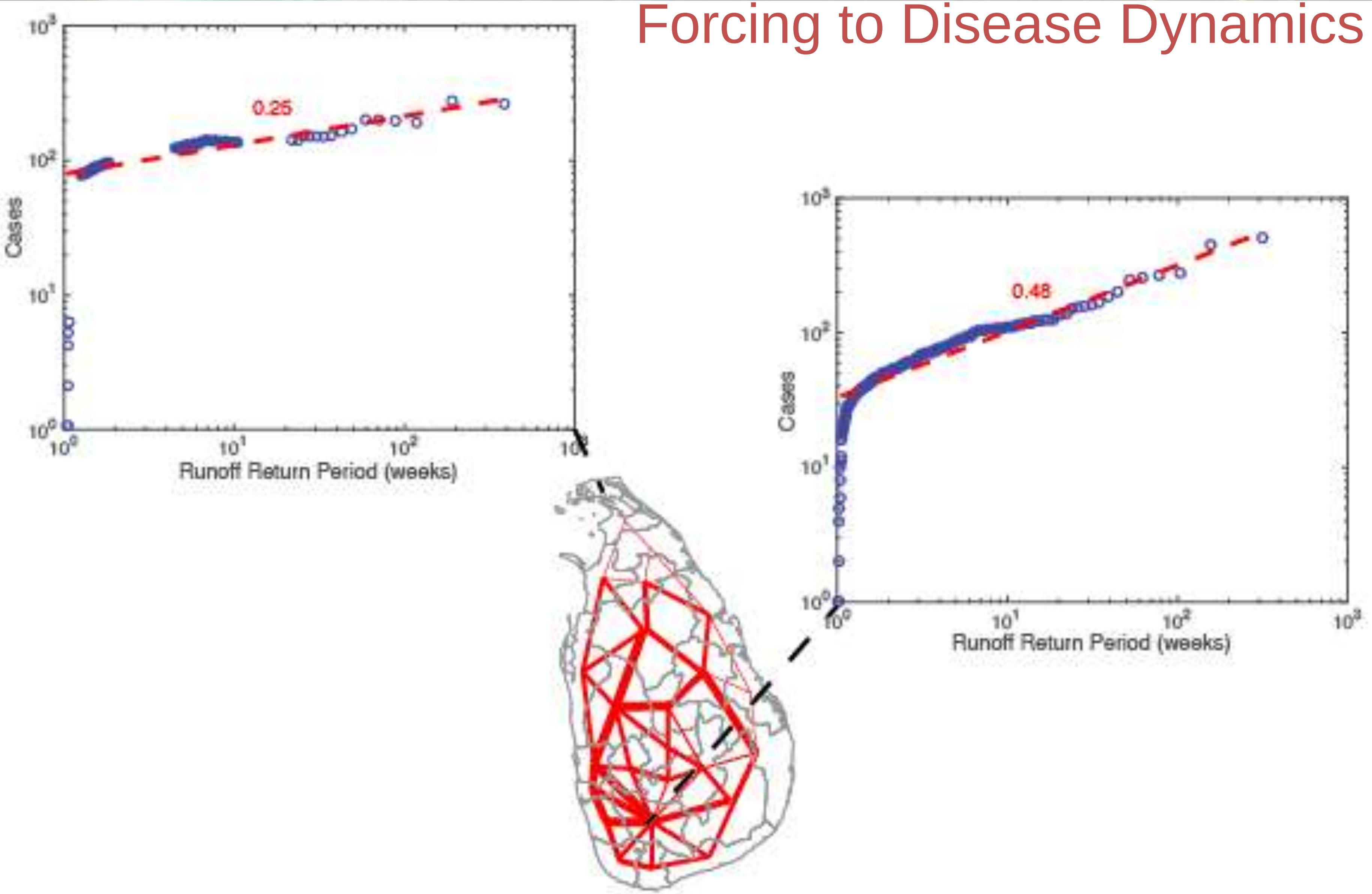
Unimportance of NDVI



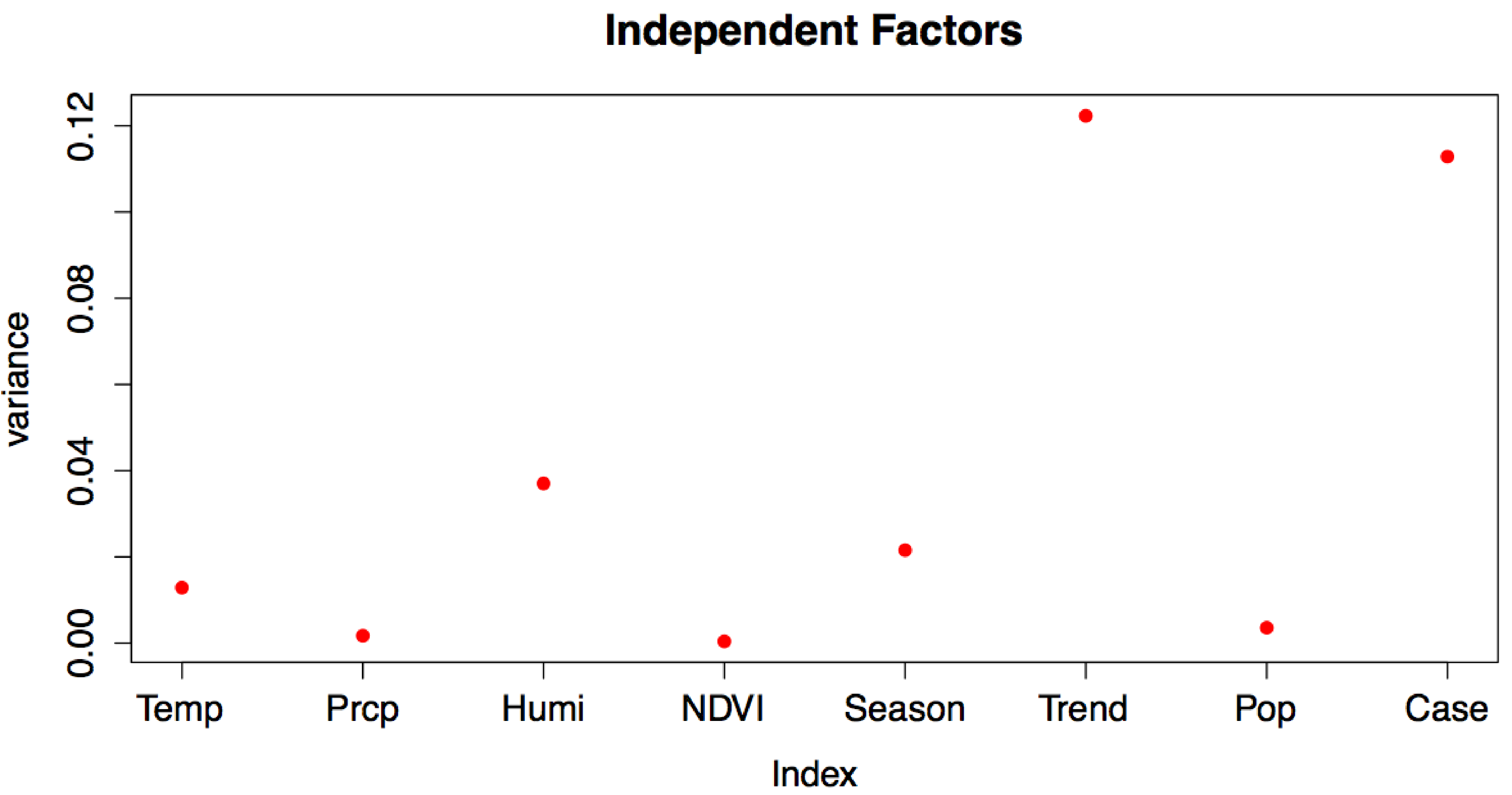
Disease Dynamics Classification

- Extreme events are often described via Pareto or power law distribution using what’s know as the 80-20 rule or Pareto principle
- “80/20 rule” - 80% of outcomes(cases) come from top 20% of causes(events)

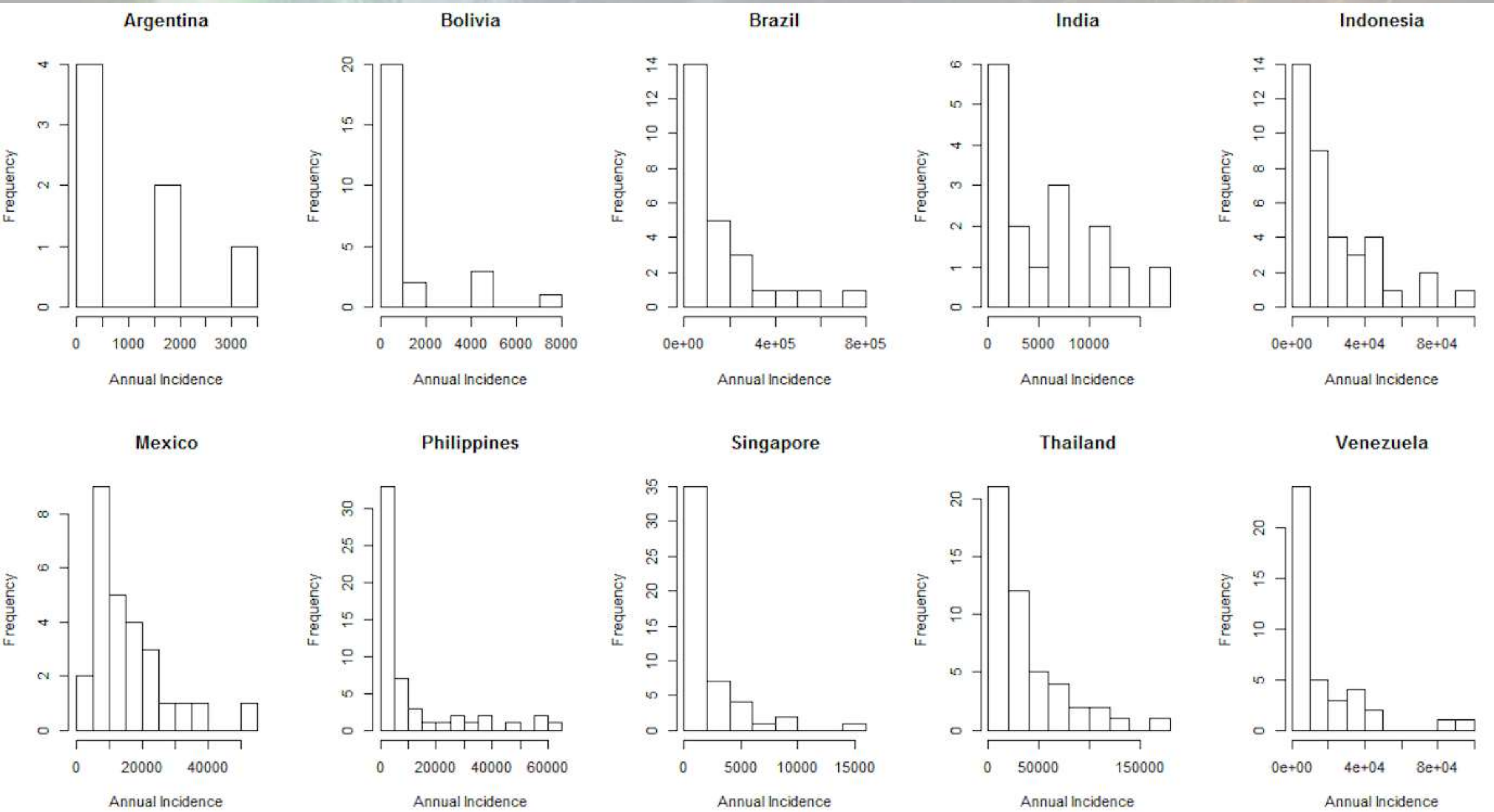
From Large Scale Forcing to Disease Dynamics



The How

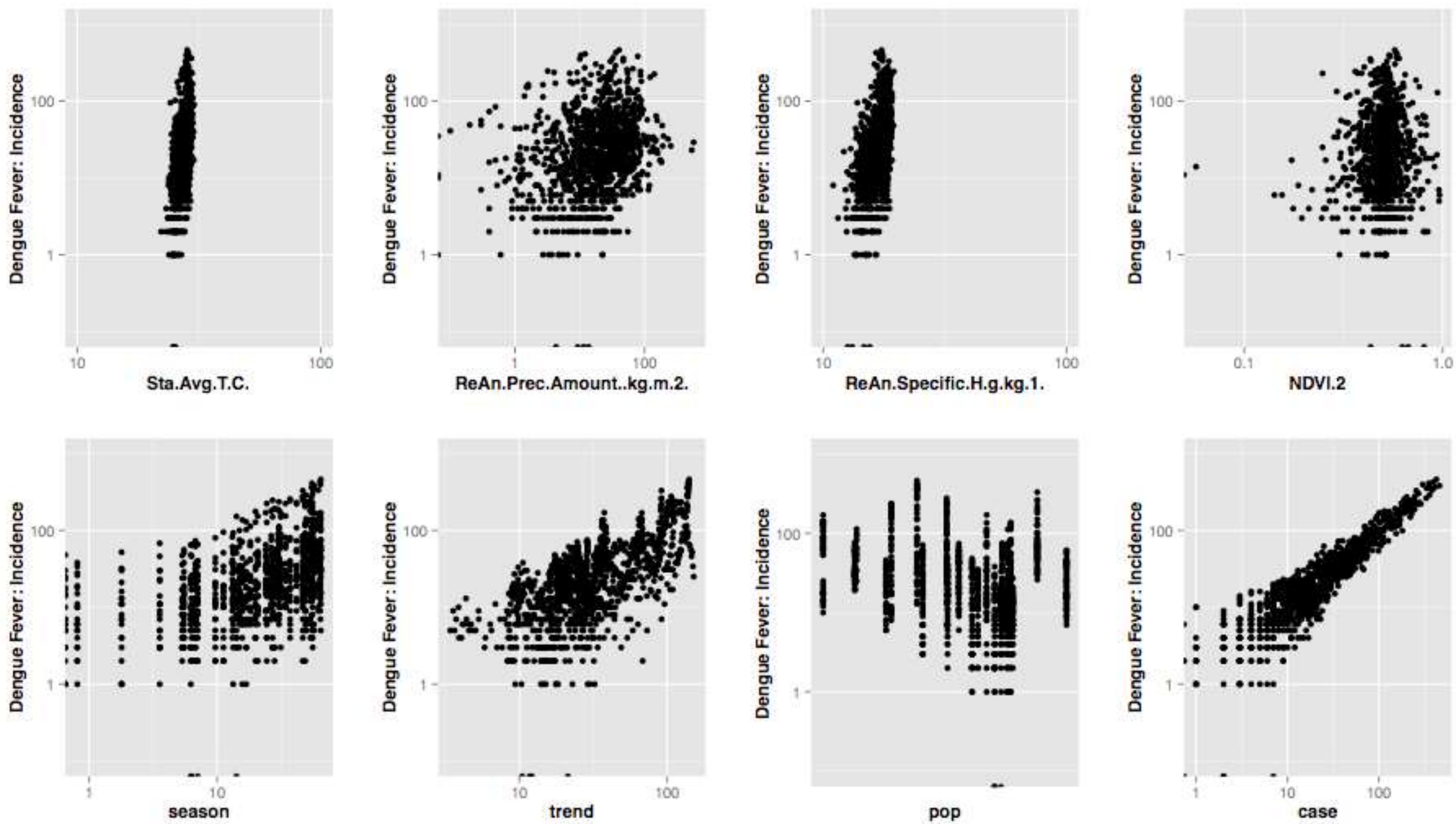

Independent Factors
variance
0.00
0.04
0.08
0.12
Temp
Prcp
Humi
NDVI
Season
Trend
Pop
Case
Index

The How

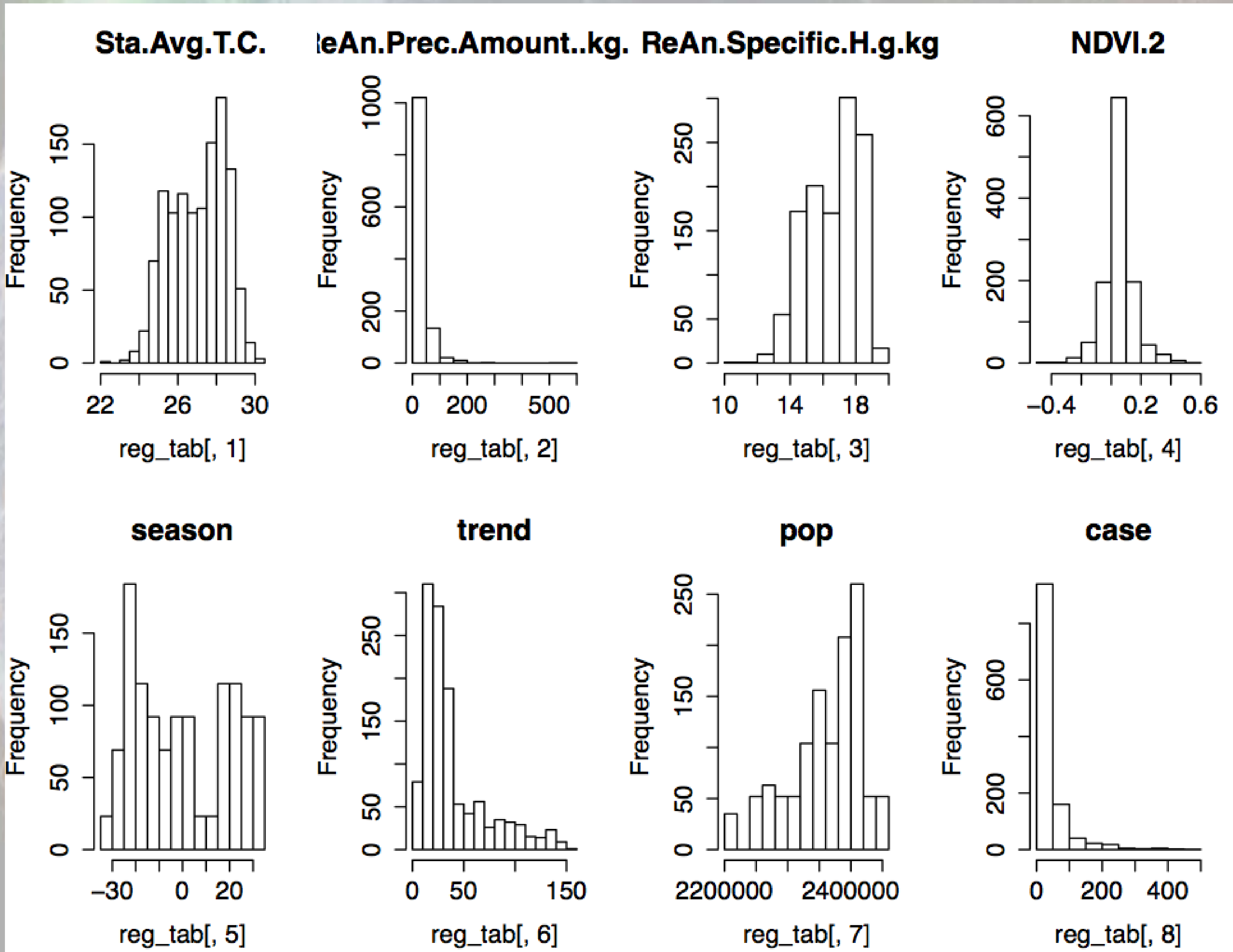


Argentina
Bolivia
Brazil
India
Indonesia
Mexico
Philippines
Singapore
Thailand
Venezuela
Frequency
Annual Incidence

The How

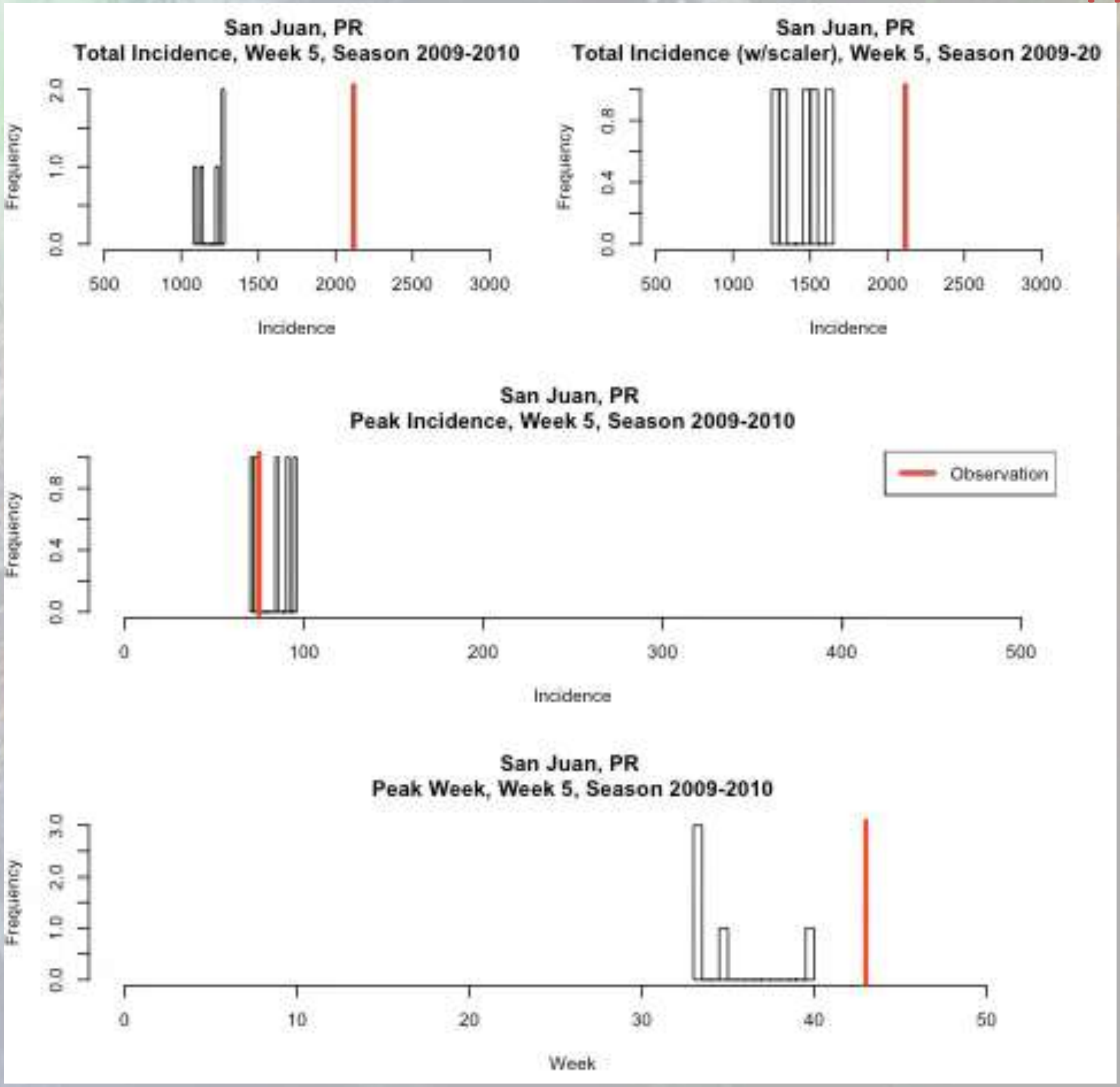

Dengue Fever: Incidence
100
1
10
100
Sta.Avg.T.C.
ReAn.Prec.Amount..kg.m.2.
ReAn.Specific.H.g.kg.1.
0.1
1.0
NDVI.2
season
trend
pop
case

The How



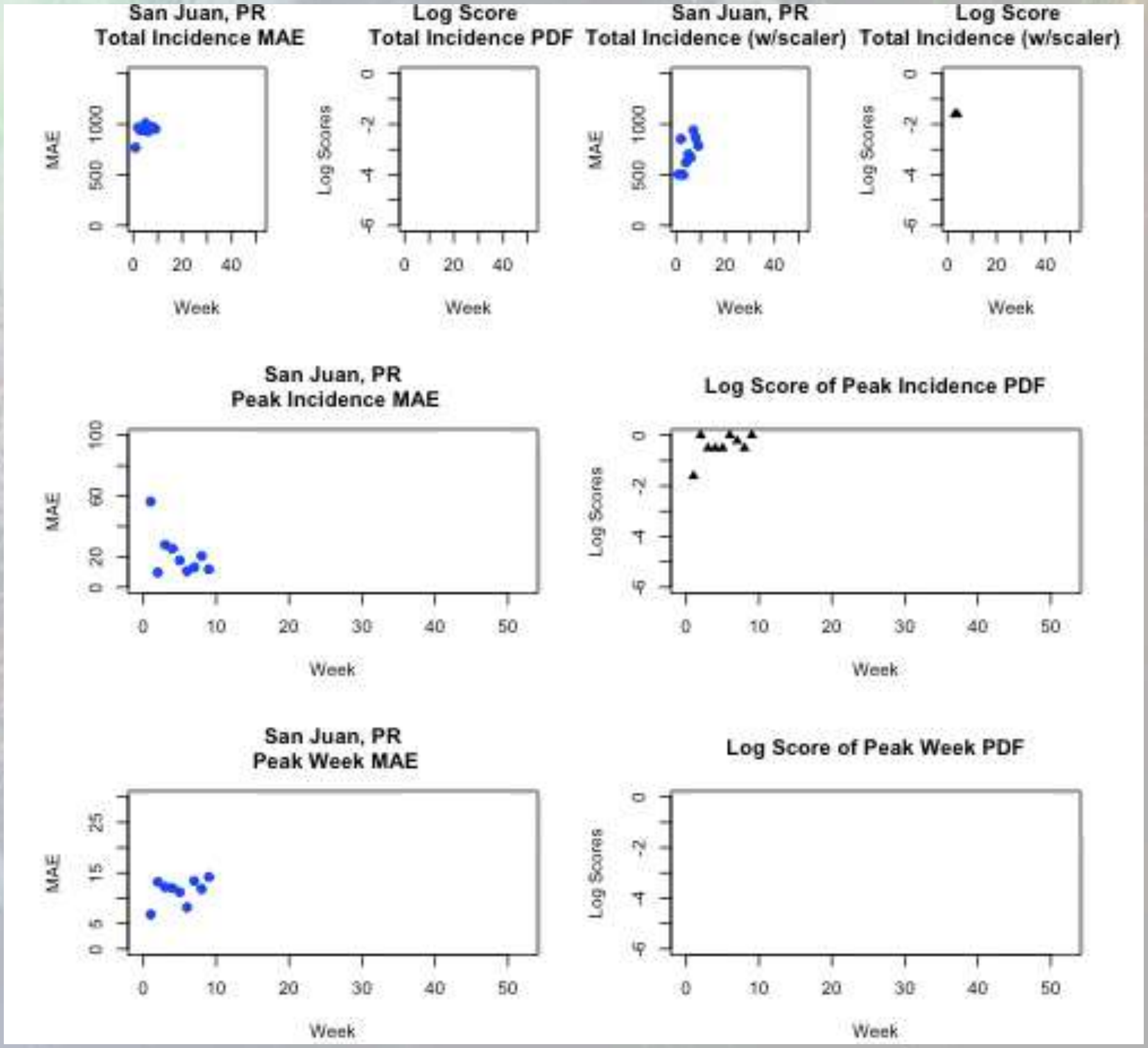
Sta.Avg.T.C.
eAn.Prec.Amount..kg.
ReAn.Specific.H.g.kg
NDVI.2
season
trend
pop
case
Frequency
reg_tab[, 1]
reg_tab[, 2]
reg_tab[, 3]
reg_tab[, 4]
reg_tab[, 5]
reg_tab[, 6]
reg_tab[, 7]
reg_tab[, 8]

The How



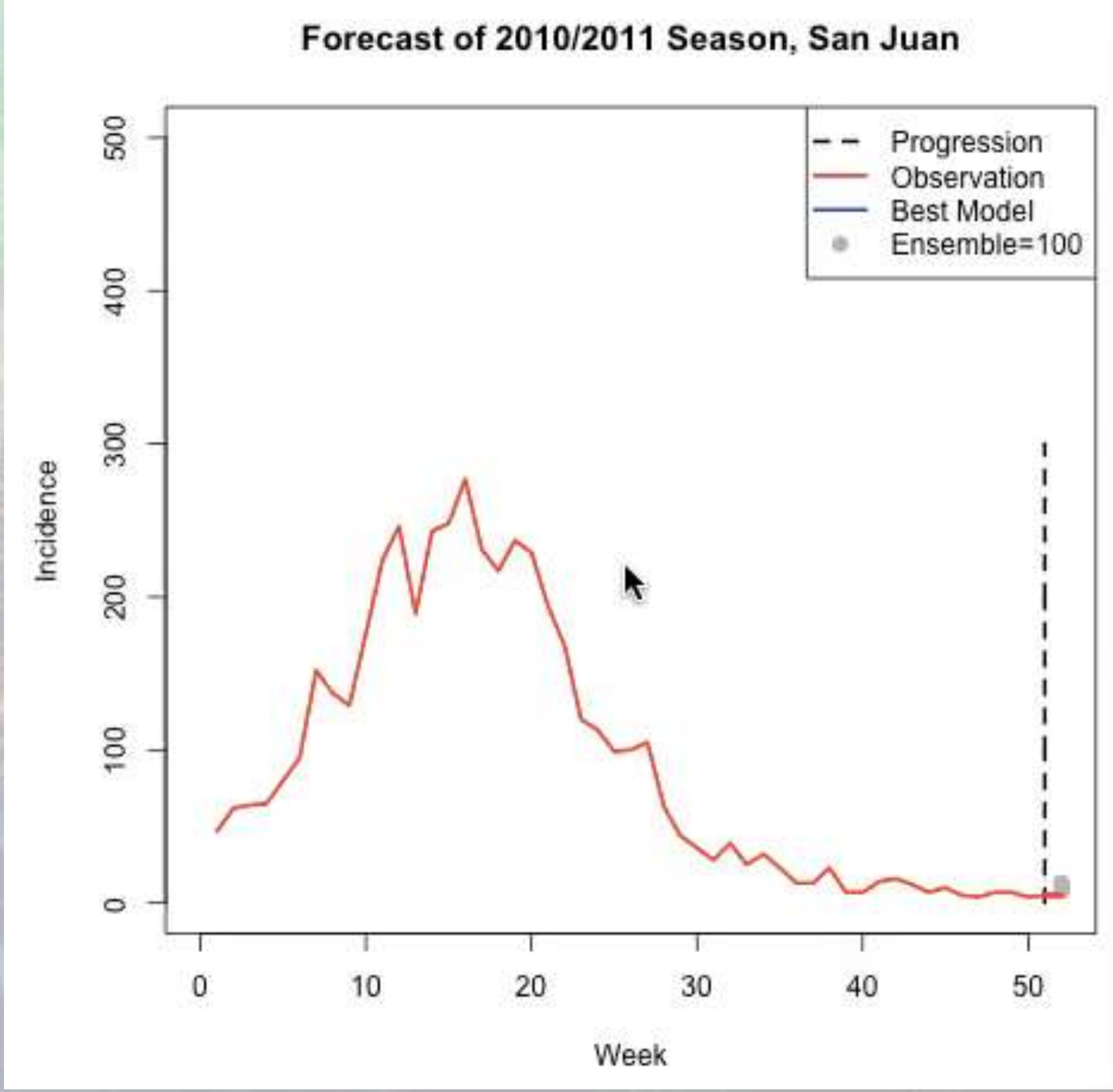
San Juan, PR
Total Incidence, Week 5, Season 2009-2010
Frequency
Incidence
San Juan, PR
Total Incidence (w/scaler), Week 5, Season 2009-20
Frequency
Incidence
San Juan, PR
Peak Incidence, Week 5, Season 2009-2010
Observation
Frequency
Incidence
San Juan, PR
Peak Week, Week 5, Season 2009-2010
Frequency
Week

The How



San Juan, PR
Total Incidence MAE
MAE
Week
Log Score
Total Incidence PDF
Log Scores
Week
San Juan, PR
Total Incidence (w/scaler)
MAE
Week
Log Score
Total Incidence (w/scaler)
Log Scores
Week
San Juan, PR
Peak Incidence MAE
MAE
Week
Log Score of Peak Incidence PDF
Log Scores
Week
San Juan, PR
Peak Week MAE
MAE
Week
Log Score of Peak Week PDF
Log Scores
Week

The How



Forecast of 2010/2011 Season, San Juan
Progression
Observation
Best Model
Ensemble=100
Incidence
0
100
200
300
400
500
Week
0
10
20
30
40
50

Models as Macroscopes that look at the *scaling and universality* of emerging patterns lead by fundamental collective factors interacting together

Importance to See Problems ''at Distance'': **Commonality of Systems' Dynamics** and *Methods for Analogous Solution* (systemic/inductive (complex systems) vs. reductionist/deductive approaches)

High **Predictability of ''Unknowns''** (e.g., Low Probability High Consequence Events) because Finite Topologies/Attractors exist in Nature. Tipping points can be predicted

Models as **technology to design the future** rather than just predicting the most likely one because it is more likely (and useful) to design an optimal future by *embracing the full uncertainty* of the status quo and the range of possibilities *-> the best way to predict the future is to design it*

''Simplicity is the ultimate sophistication … and the solution of the complex nature''
Leonardo Da Vinci

On the Morphological Effective Systemic EpiGraph (MESE)