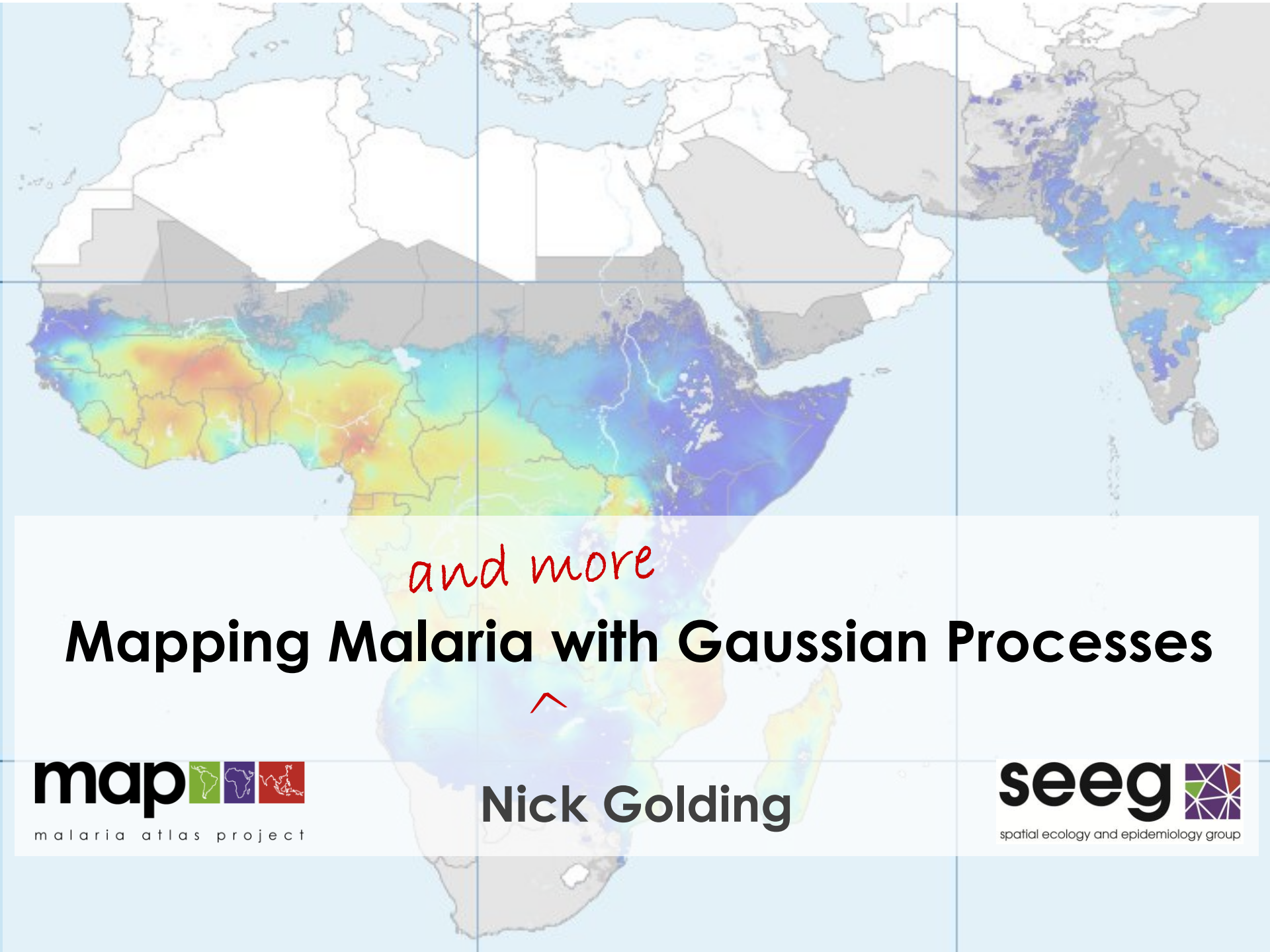


and more

Mapping Malaria with Gaussian Processes

Nick Golding

malaria atlas project

spatial ecology and epidemiology group

seeg

spatial ecology and epidemiology group

UNIVERSITY OF OXFORD

map

malaria atlas project

IDAMS dengue

abraid

atlas of baseline risk assessment for infectious disease

part 1:

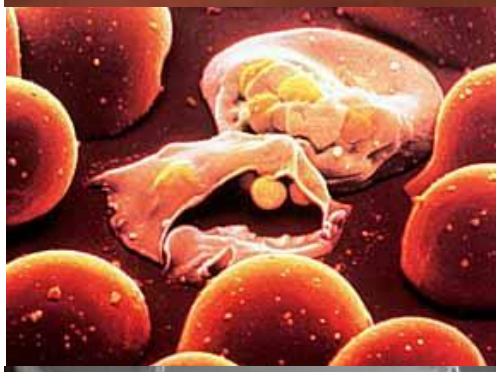
malaria

malaria

≈ **450,000,000 cases** &
≈ **1,000,000 deaths** per year

caused by: ***Plasmodium falciparum***
P. vivax
P. malariae
P. ovale
P. knowlesi

it is **controllable** and **treatable**



malaria atlas project

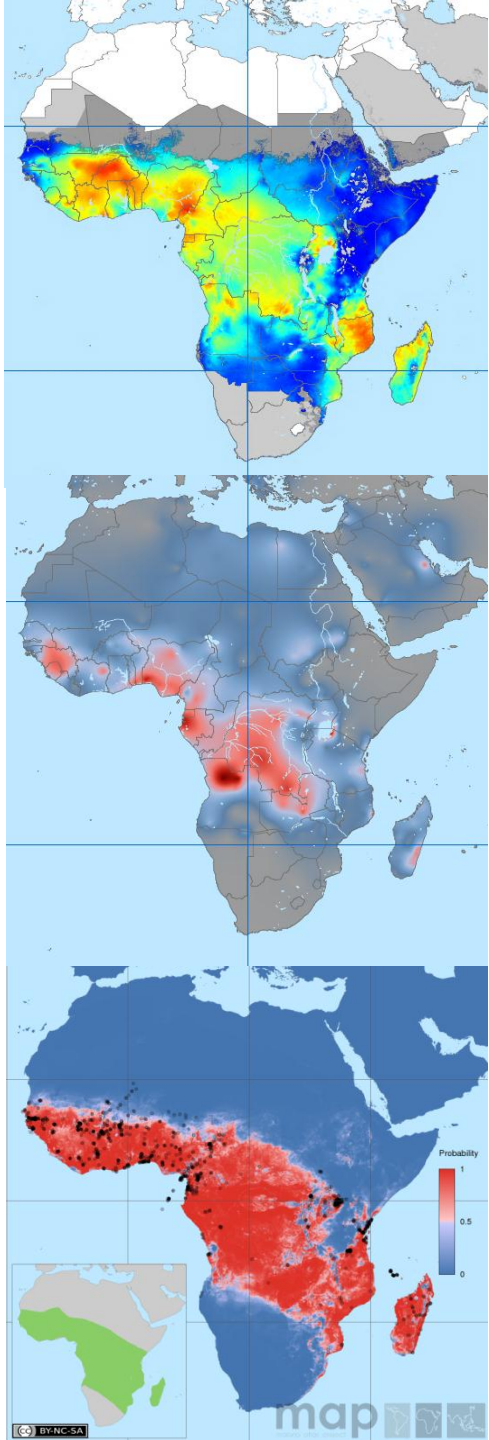
open-access spatial info
for control & elimination

global malaria **risk maps**

estimates of **clinical burden**

blood disorder maps

mosquito maps & more



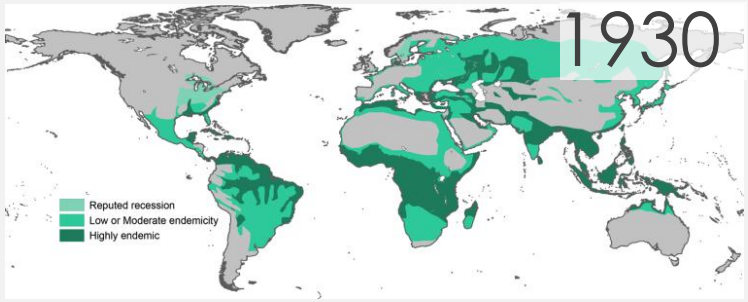
why maps?

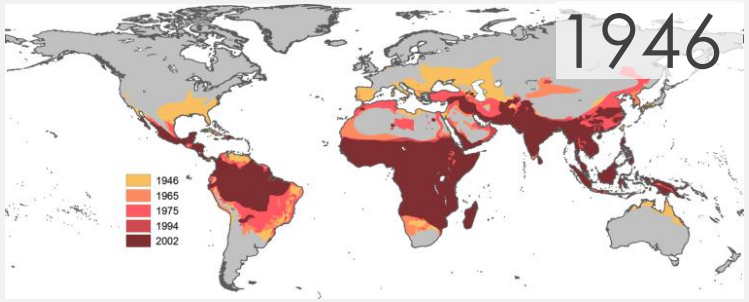
malaria is **spatially heterogeneous**

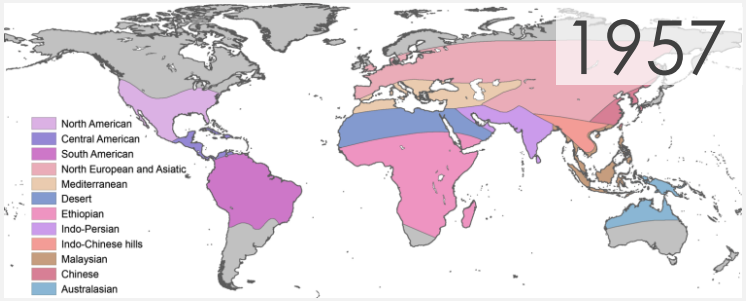
maps can:

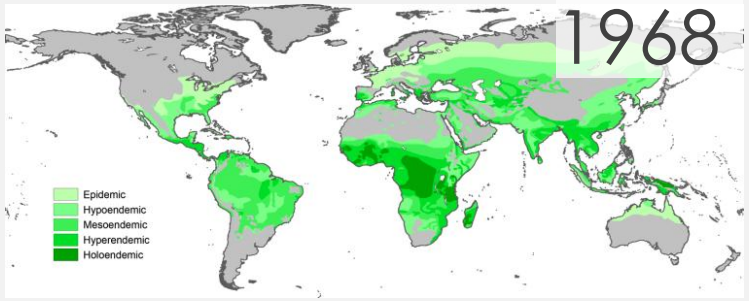
- identify **populations at risk**
- **evaluate impact** of interventions
- objectively **evaluate options** for control

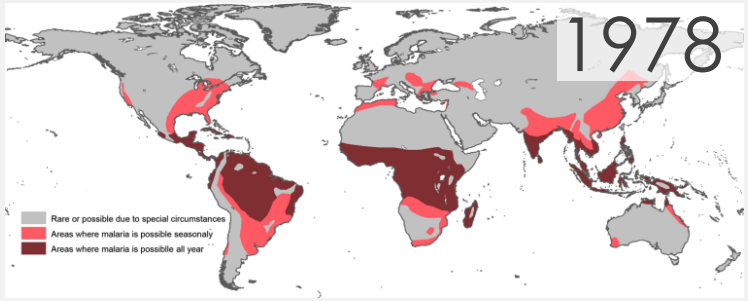
why new maps?

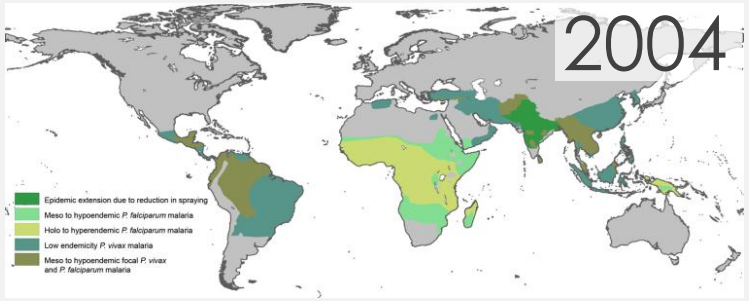

1930
Reputed recession
Low or Moderate endemicity
Highly endemic


1946
1946
1965
1975
1994
2002


1957
North American
Central American
South American
North European and Asiatic
Mediterranean
Desert
Ethiopian
Indo-Persian
Indo-Chinese hills
Malaysian
Chinese
Australasian


1968
Epidemic
Hypoendemic
Mesoendemic
Hyperendemic
Holoendemic


1978
Rare or possible due to special circumstances
Areas where malaria is possible seasonaly
Areas where malaria is possible all year


2004
Epidemic extension due to reduction in spraying
Meso to hypoendemic P. falciparum malaria
Holo to hyperendemic P. falciparum malaria
Low endemicity P. vivax malaria
Meso to hypoendemic focal P. vivax and P. falciparum malaria

why new maps?

previous maps:

- poorly described **evidence-base**
- poorly defined, subjective **methods**
- no **estimates of reliability**

mapping endemicity

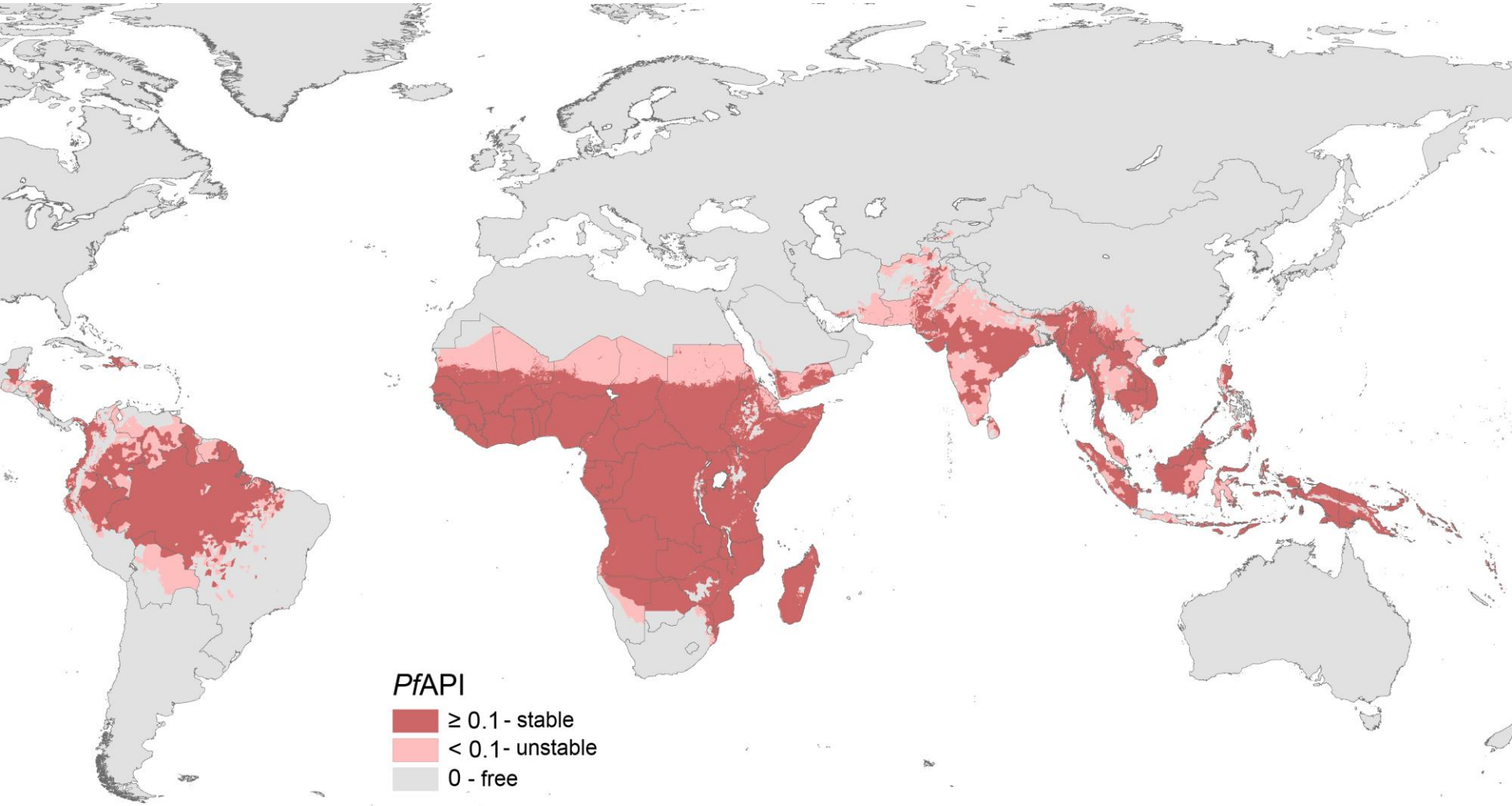
transmission limits

- regional **endemic status**
- **biological masks**

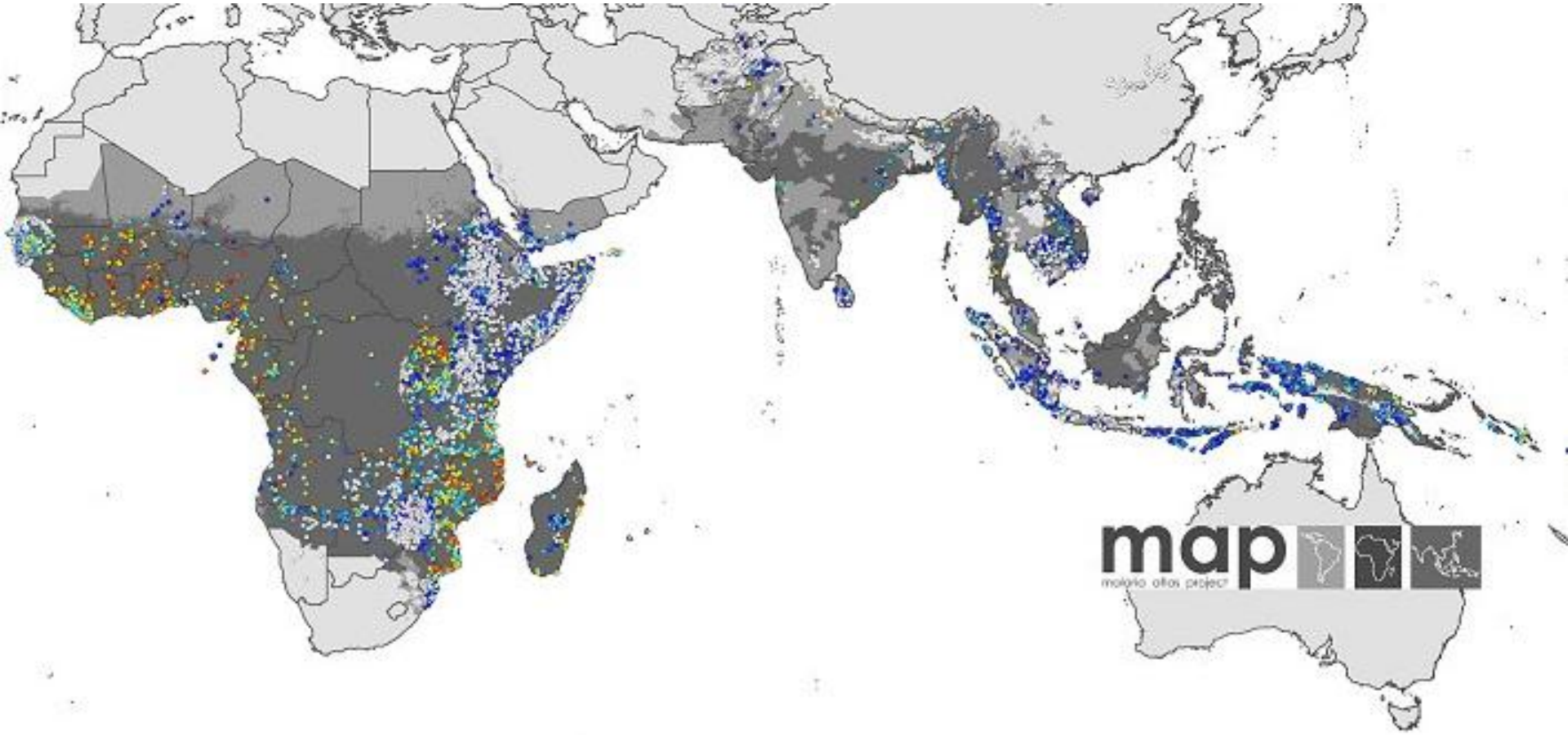
infection prevalence

- *P.f* parasite rate **surveys**
- formal **statistical model** (MBG)

P.f. transmission limits



P.f. prevalence survey data



22,212 surveys

P.f. model

$$N_i^+ \mid N_i, P(\mathbf{x}_i, \mathbf{t}_i) \sim \mathrm{Bin}\left(N_i, P(\mathbf{x}_i, \mathbf{t}_i)\right)$$

$$P(\mathbf{x}, \mathbf{t}) = \mathrm{logit}^{-1}\left(f(\mathbf{x}, \mathbf{t}) + \varepsilon\right)$$

$$f \mid \theta \sim \mathrm{GP}\left(M_\theta, C_\theta\right)$$

$$\varepsilon \sim N\left(0, \sigma^2\right)$$

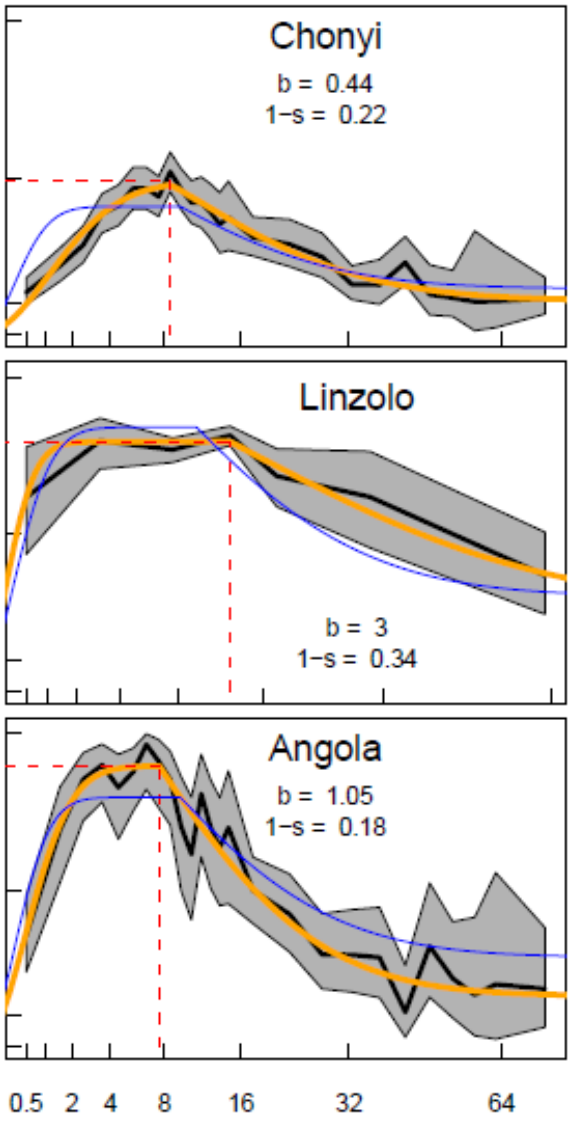
tweaks - age stratification

prevalence is **age-dependent**

surveys of different ages not equivalent

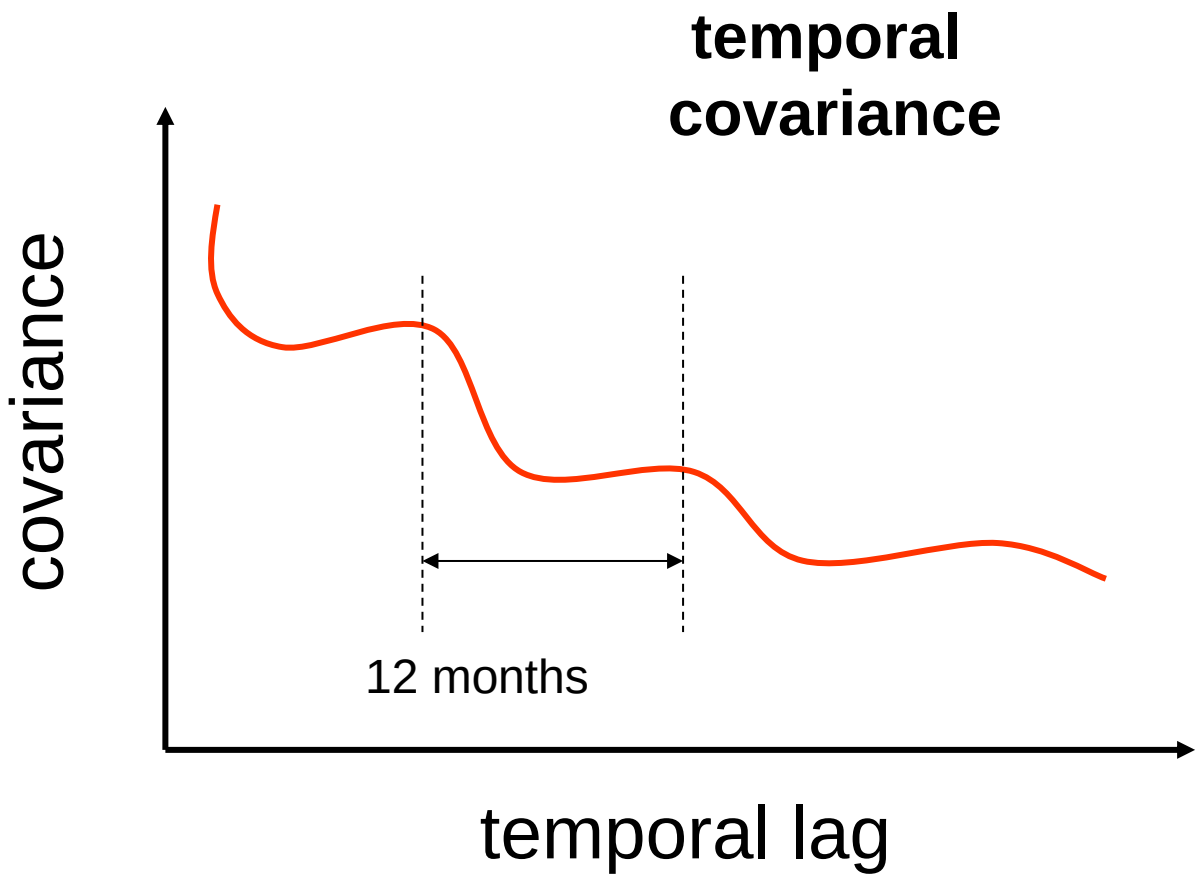
needed to **standardise** (2-10 year-olds)

used a **Bayesian sub-model**



Smith et al.(2007) Malaria Journal

tweaks - seasonality



space-time covariance

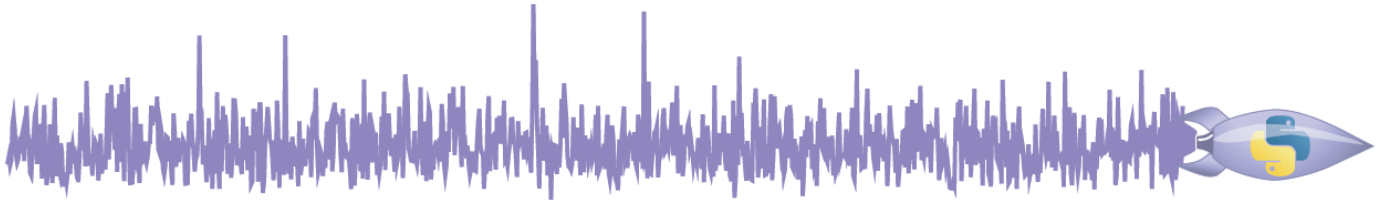
$$C(x_i, t_i; x_j, t_j) = \tau^2 \gamma(0) \frac{(\Delta x)^{\gamma(\Delta t)} K_{\gamma(\Delta t)}(\Delta x)}{2^{\gamma(\Delta t)-1} \Gamma(\gamma(\Delta t)+1)},$$

$$\gamma(\Delta t) = \frac{1}{2\rho + 2(1-\rho)\left[(1-\upsilon) e^{-|\Delta t|/\phi_t} + \upsilon \cos(2\pi \Delta t)\right]},$$

$$\Delta t = |t_i - t_j|$$

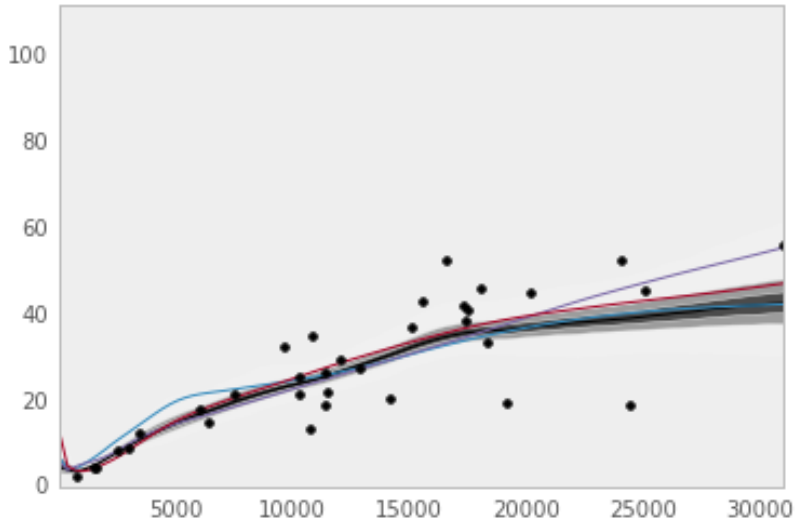
$$\Delta x = 2\sqrt{\gamma(\Delta t)} \frac{D_{GC}(x_i, x_j) \sqrt{1 - \psi^2 \cos^2(\theta(x_i, x_j) - \lambda)}}{\phi_x}$$

PyMC



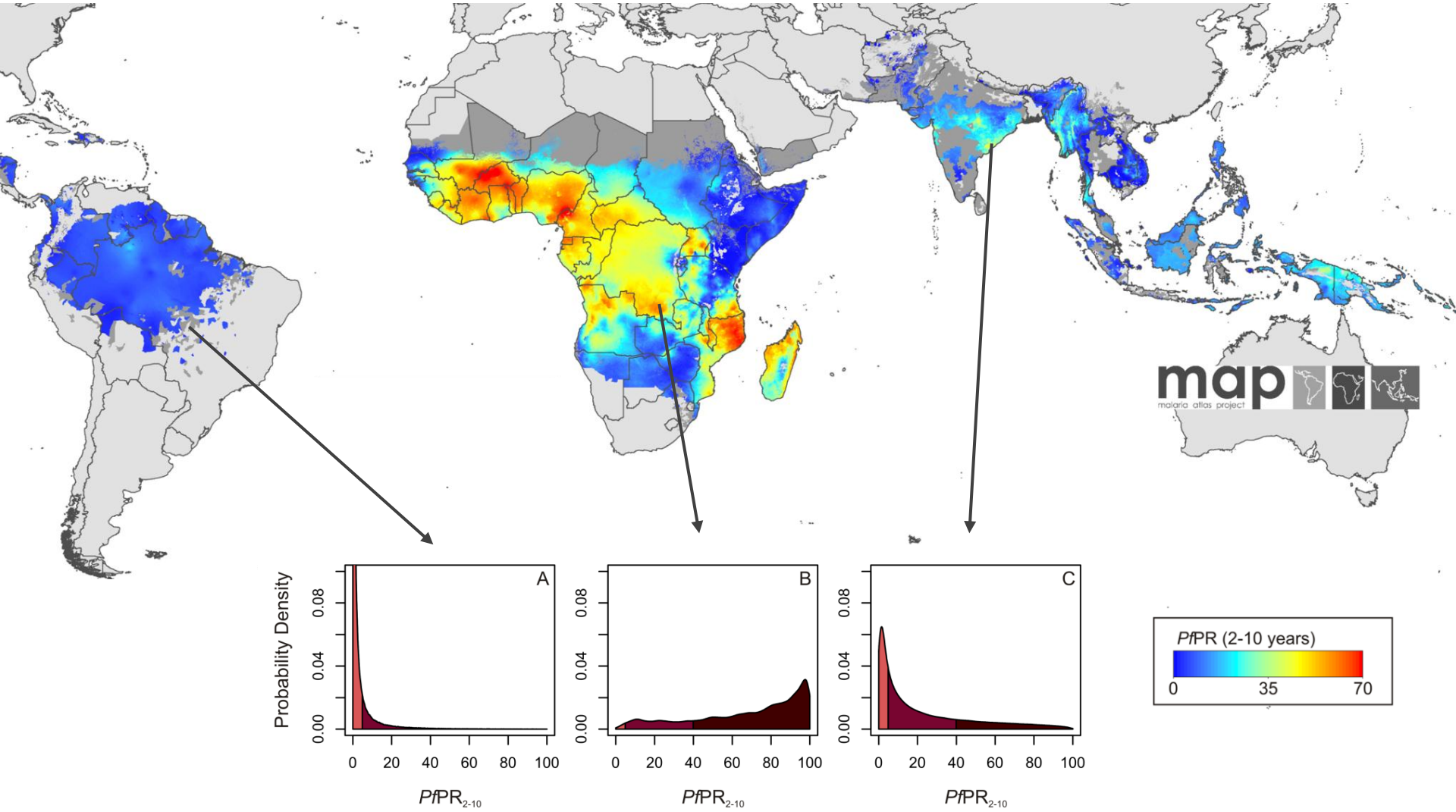
```
[-----------------96%----------------  ] 9684 of 10000 complete in 51.3 sec
[-----------------97%----------------- ] 9746 of 10000 complete in 51.8 sec
[-----------------98%----------------- ] 9810 of 10000 complete in 52.3 sec
[-----------------98%----------------- ] 9873 of 10000 complete in 52.8 sec
[-----------------99%----------------- ] 9937 of 10000 complete in 53.3 sec
[-----------------99%----------------- ] 9999 of 10000 complete in 53.8 sec
[-----------------100%-----------------] 10000 of 10000 complete in 53.8 sec
```

In [49]:
```
ss.plot_SR()
```



In [50]:
```
ss.plot_traces()
```

mapping endemicity



new work: more temporal

part 2:

all the other diseases

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

rstb.royalsocietypublishing.org

Review

Cite this article: Hay SI, Battle KE, Pigott DM, Smith DL, Moyes CL, Bhatt S, Brownstein JS, Collier N, Myers MF, George DB, Gething PW. 2013 Global mapping of infectious disease. Phil Trans R Soc B 368: 20120250.
http://dx.doi.org/10.1098/rstb.2012.0250

One contribution of 18 to a Discussion Meeting Issue 'Next-generation molecular and evolutionary epidemiology of infectious disease'.

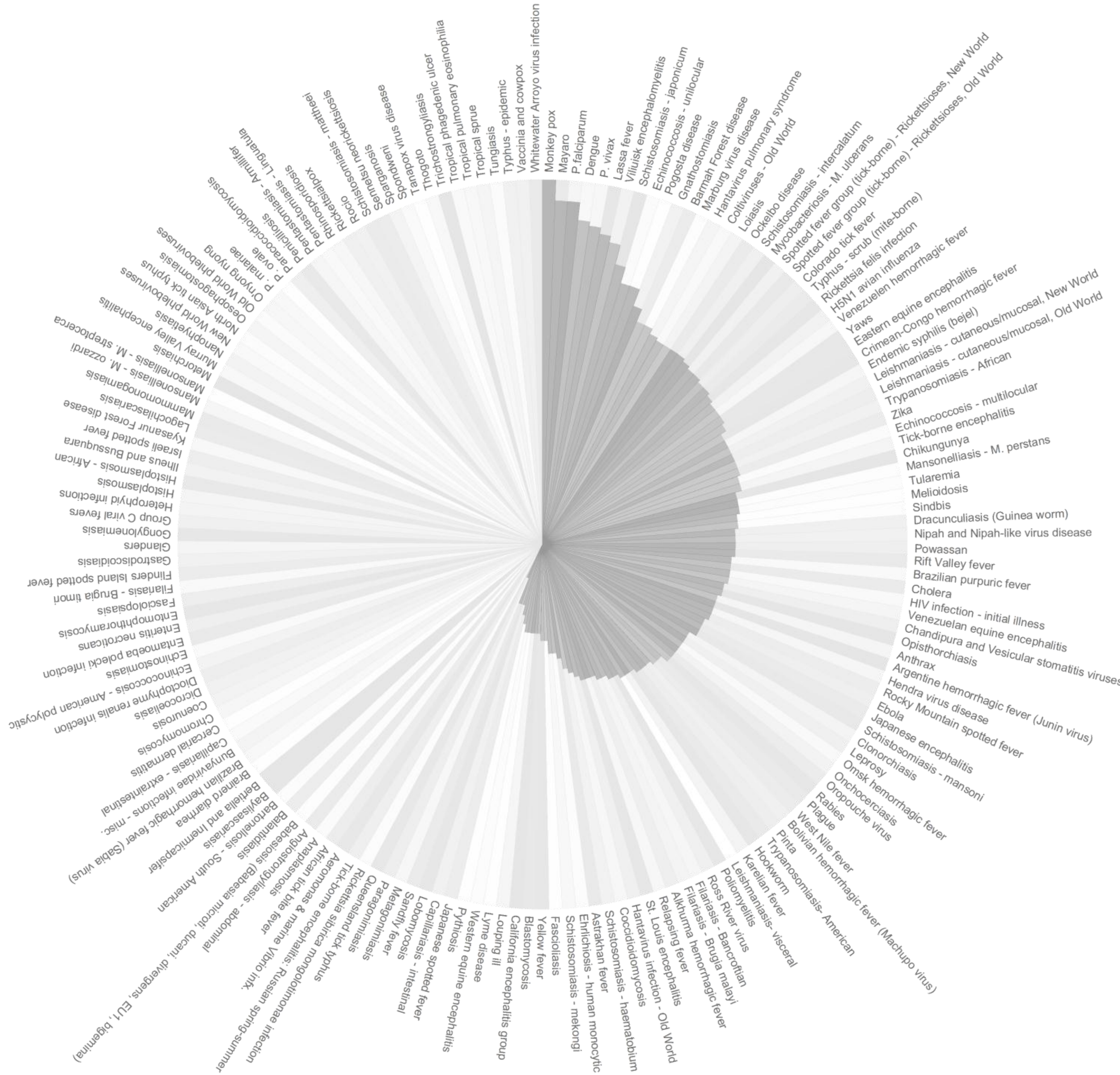
Subject Areas:
health and disease and epidemiology,

Global mapping of infectious disease

Simon I. Hay[1,2], Katherine E. Battle[1], David M. Pigott[1], David L. Smith[2,3], Catherine L. Moyes[1], Samir Bhatt[1], John S. Brownstein[4], Nigel Collier[5], Monica F. Myers[1], Dylan B. George[2] and Peter W. Gething[1]

[1]Spatial Ecology and Epidemiology Group, Department of Zoology, University of Oxford, Oxford, UK
[2]Fogarty International Center, National Institutes of Health, Bethesda, MD, USA
[3]Department of Epidemiology and Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[4]Department of Pediatrics, Harvard Medical School and Children's Hospital Informatics Program, Boston Children's Hospital, Boston, MA, USA
[5]National Institute of Informatics, Research Organization of Information and Systems, Tokyo, Japan

The primary aim of this review was to evaluate the state of knowledge of the geographical distribution of all infectious diseases of clinical significance to humans. A systematic review was conducted to enumerate cartographic progress, with respect to the data available for mapping and the methods currently applied. The results helped define the minimum information requirements for mapping infectious disease occurrence, and a quantitative framework for assessing the mapping opportunities for all infectious diseases. This revealed that of 355 infectious diseases identified, 174 (49%) have a strong rationale for mapping and of these only 7 (4%) had been comprehensively mapped. A variety of ambitions, such as the quantification of the global burden of infectious disease, international biosurveillance, assessing the likelihood of infectious disease outbreaks and exploring the propensity for infectious disease evolution and emergence, are limited by these omissions. An overview of the factors hindering progress in disease cartography is provided. It is argued that rapid improvement in the landscape of infectious diseases mapping can be made by embracing non-conventional data sources, automation of geo-positioning and mapping procedures enabled by machine learning and information technology, respectively, in addition to harnessing

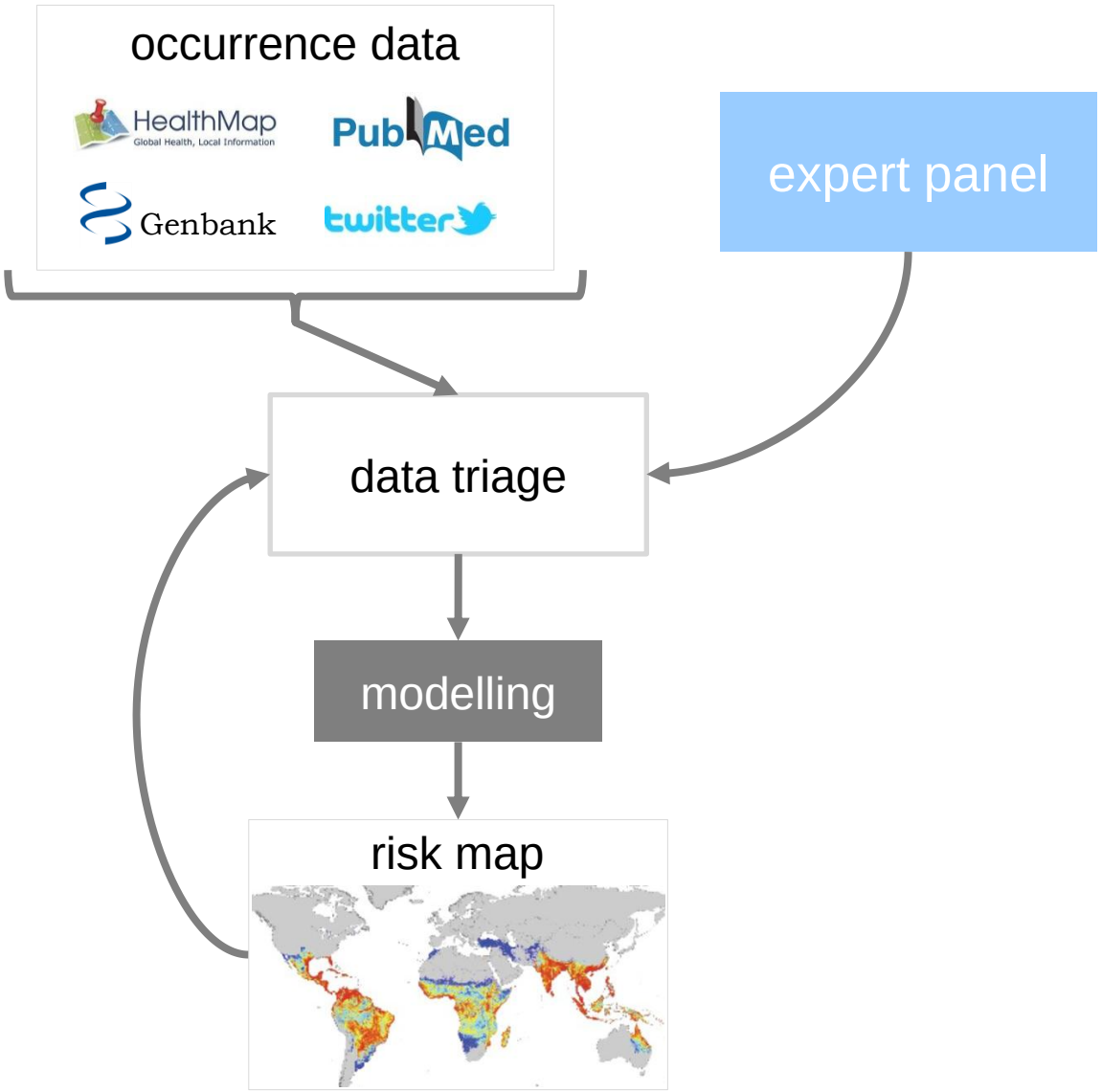


Monkey pox
Mayaro
P.falciparum
Dengue
P. vivax
Lassa fever
Viliuisk encephalomyelitis
Schistosomiasis - japonicum
Echinococcosis - unilocular
Pogosta disease
Gnathostomiasis
Barmah Forest disease
Marburg virus disease
Hantavirus pulmonary syndrome
Coltiviruses - Old World
Loiasis
Ockelbo disease
Schistosomiasis - intercalatum
Mycobacteriosis - M. ulcerans
Spotted fever group (tick-borne) - Rickettsioses, New World
Spotted fever group (tick-borne) - Rickettsioses, Old World
Colorado tick fever
Typhus - scrub (mite-borne)
Rickettsia felis infection
H5N1 avian influenza
Venezuelen hemorrhagic fever
Yaws
Eastern equine encephalitis
Crimean-Congo hemorrhagic fever
Endemic syphilis (bejel)
Leishmaniasis - cutaneous/mucosal, New World
Leishmaniasis - cutaneous/mucosal, Old World
Trypanosomiasis - African
Zika
Echinococcosis - multilocular
Tick-borne encephalitis
Chikungunya
Mansonelliasis - M. perstans
Tularemia
Melioidosis
Sindbis
Dracunculiasis (Guinea worm)
Nipah and Nipah-like virus disease
Powassan
Rift Valley fever
Brazilian purpuric fever
Cholera
HIV infection - initial illness
Venezuelan equine encephalitis
Chandipura and Vesicular stomatitis viruses
Opisthorchiasis
Anthrax
Argentine hemorrhagic fever (Junin virus)
Hendra virus disease
Rocky Mountain spotted fever
Ebola
Japanese encephalitis
Schistosomiasis - mansoni
Clonorchiasis
Leprosy
Omsk hemorrhagic fever
Onchocerciasis
Oropouche virus
Rabies
Plague
West Nile fever
Pinta
Bolivian hemorrhagic fever (Machupo virus)
Trypanosomiasis- American
Hookworm
Karelian fever
Leishmaniasis- visceral
Poliomyelitis
Ross River virus
Filariasis - Bancroftian
Filariasis - Brugia malayi
Alkhurma hemorrhagic fever
Relapsing fever
St. Louis encephalitis
Hantavirus infection - Old World
Coccidioidomycosis
Schistosomiasis - haematobium
Astrakhan fever
Ehrlichiosis - human monocytic
Schistosomiasis - mekongi
Fascioliasis
Yellow fever
Blastomycosis
California encephalitis group
Louping ill
Lyme disease
Western equine encephalitis
Pythiosis
Japanese spotted fever
Capillariasis - intestinal
Lobomycosis
Sandfly fever
Metagonimiasis
Paragonimiasis
Queensland tick typhus
Rickettsia sibirica mongolotimonae infection
Tick-borne encephalitis: Russian spring-summer
Aeromonas & marine Vibrio infx.
African tick bite fever
Anaplasmosis
Angiostrongyliasis - abdominal
Babesiosis (Babesia microti, ducani, divergens, EU1, bigemina)
Balantidiasis
Bartonellosis - South American
Baylisascariasis
Bertiella and Inermicapsifer
Brainerd diarrhea
Brazilian hemorrhagic fever (Sabia virus)
Bunyaviridae infections - misc.
Capillariasis - extraintestinal
Cercarial dermatitis
Chromomycosis
Coenurosis
Dicrocoeliasis
Dioctophyme renalis infection
Echinococcosis - American polycystic
Echinostomiasis
Entamoeba polecki infection
Enteritis necroticans
Entomophthoramycosis
Fasciolopsiasis
Filariasis - Brugia timori
Flinders Island spotted fever
Gastrodiscoidiasis
Glanders
Gongylonemiasis
Group C viral fevers
Heterophyid infections
Histoplasmosis
Histoplasmosis - African
Ilheus and Bussuquara
Israeli spotted fever
Kyasanur Forest disease
Lagochilascariasis
Mammomonogamiasis
Mansonelliasis - M. ozzardi
Mansonelliasis - M. streptocerca
Metorchiasis
Murray Valley encephalitis
Nanophyetiasis
New World phleboviruses
North Asian tick typhus
Oesophagostomiasis
Old World phleboviruses
O'nyong nyong
P. malariae
P. ovale
Paracoccidioidomycosis
Penicilliosis
Pentastomiasis - Armillifer
Pentastomiasis - Linguatula
Rhinosporidiosis
Rickettsialpox
Rocio
Schistosomiasis - mattheei
Schistosu neorickettsiosis
Semnetsu
Sparganosis
Spondweni
Spondpox virus disease
Tanapox
Thogoto
Trichostrongyliasis
Tropical phagedenic ulcer
Tropical pulmonary eosinophilia
Tropical sprue
Tungiasis
Typhus - epidemic
Vaccinia and cowpox
Whitewater Arroyo virus infection

abraid

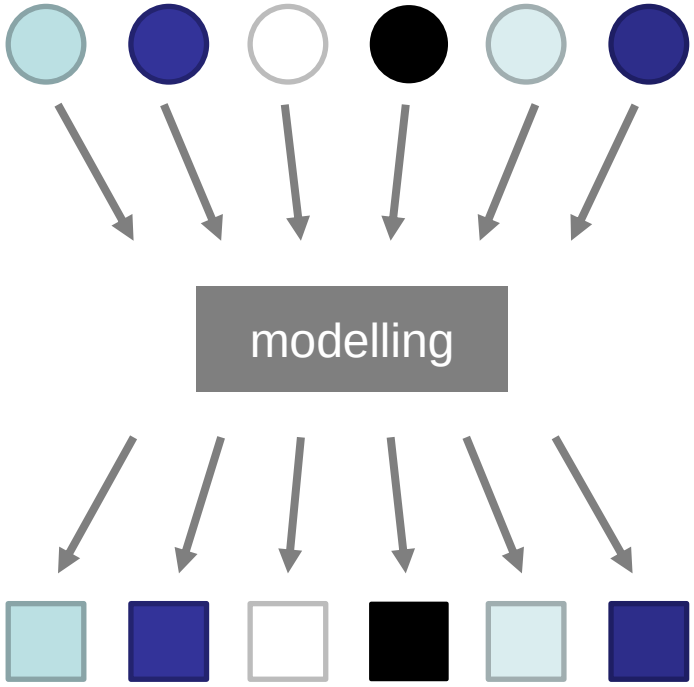
atlas of baseline risk assessment for infectious disease

abraid


occurrence data
HealthMap
Global Health, Local Information
PubMed
Genbank
twitter
expert panel
data triage
modelling
risk map

abraid

data on many diseases



many risk maps

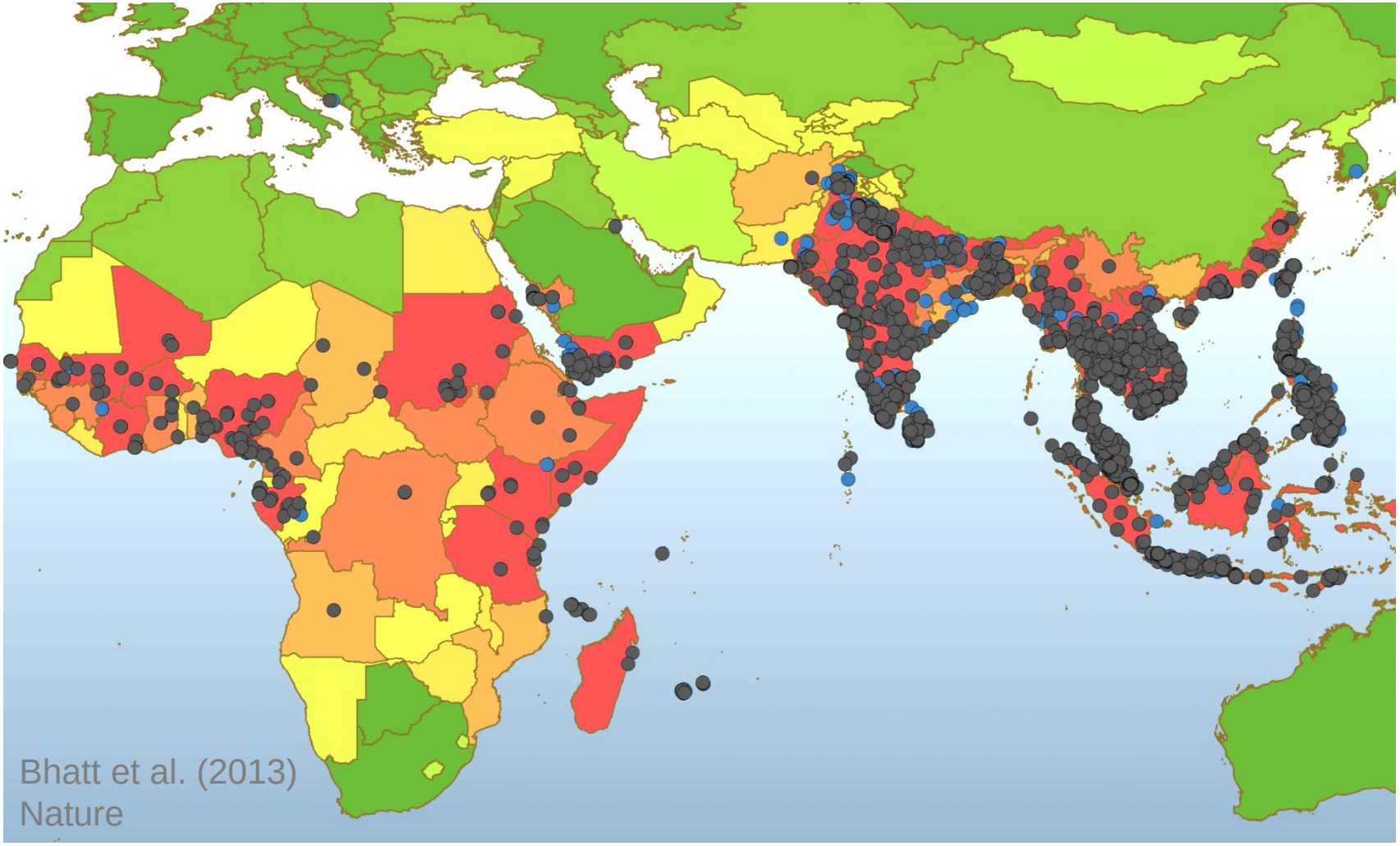
challenges

fitting **accurate** models with poor data

expressing **uncertainty**

‘one size fits all’ model

sparse, presence-only data



species distribution modelling

classification (species present/absent)

function of environmental space (**fundamental niche**)

require **pseudo-absence** data

well studied, **widely used** in ecology

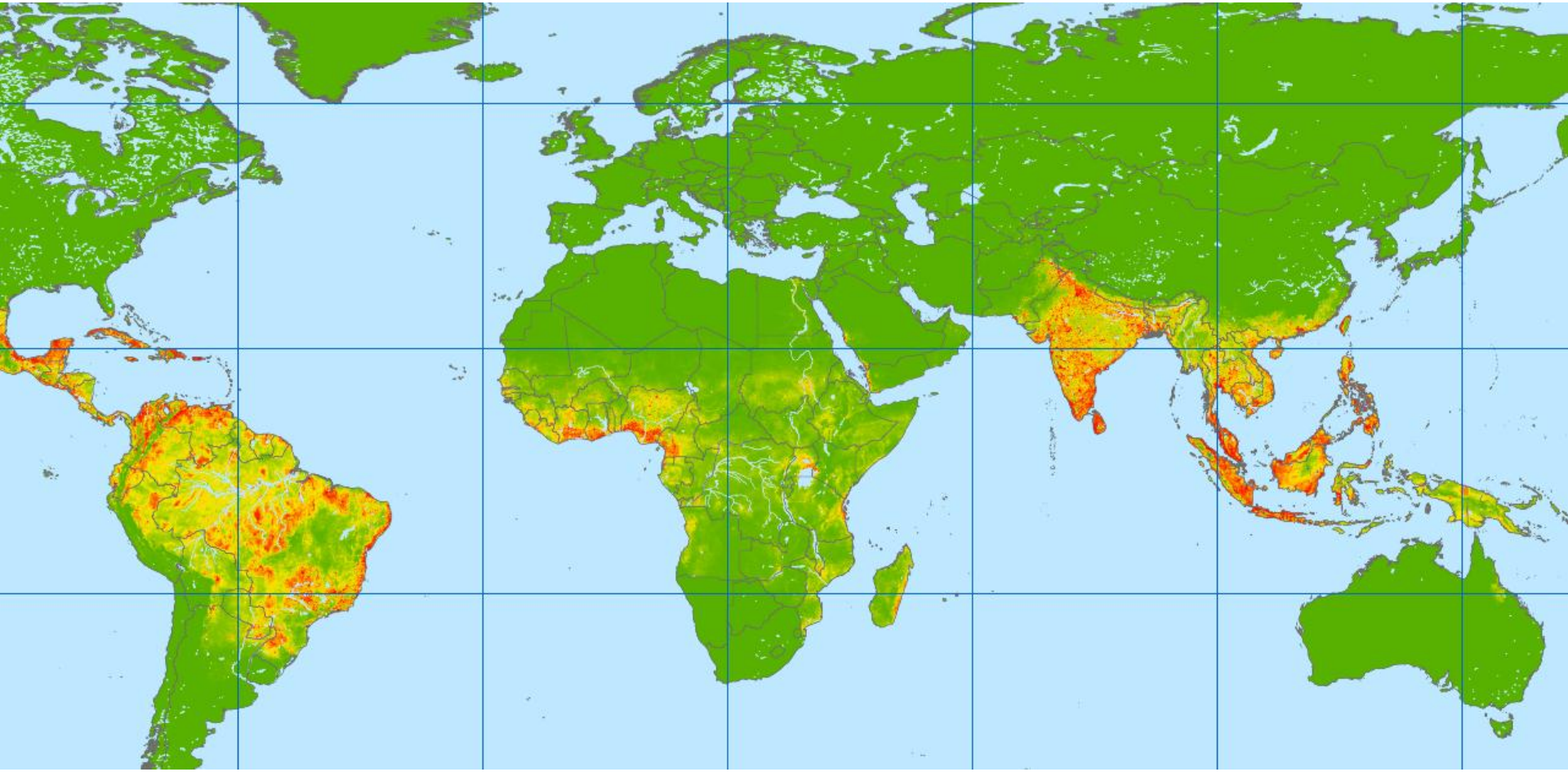
problems with SDM

assumes **fundamental niche** is key
- communicable diseases?

assumes equilibrium
- emerging diseases?

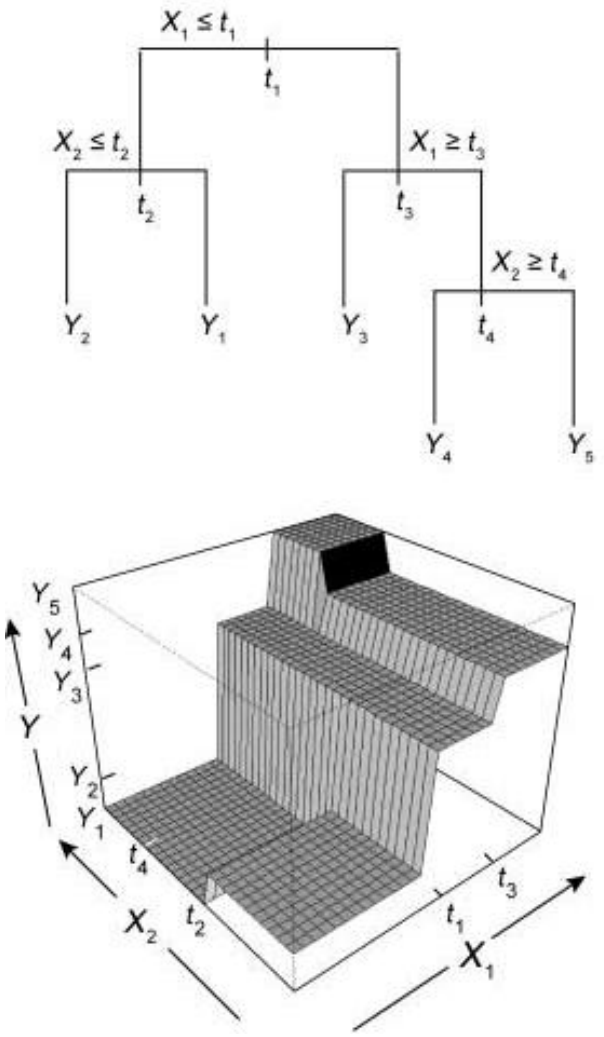
discards spatial info
- disease control?

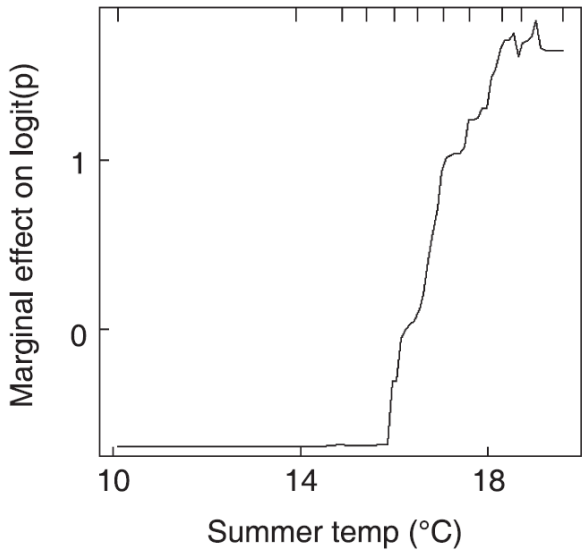
dengue



Bhatt et al. (2013) Nature

boosted regression trees

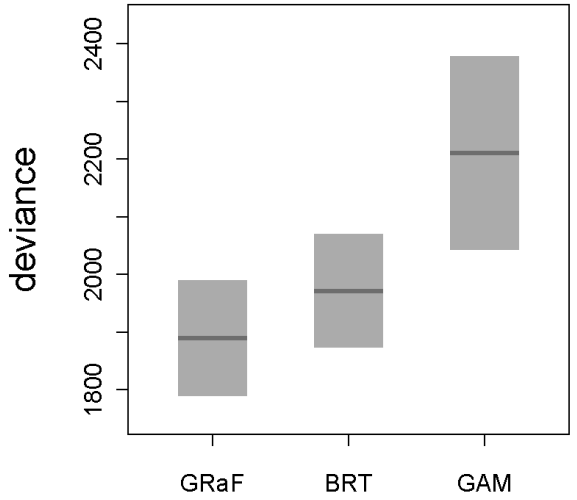




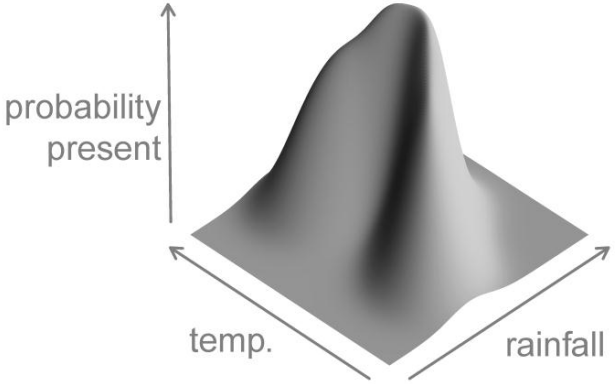
Elith *et al.* (2008)

Gaussian process SDM

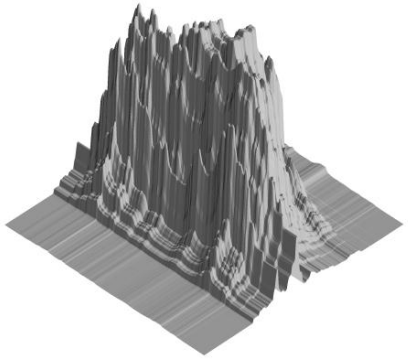
GP classification
outperforms best SDMs



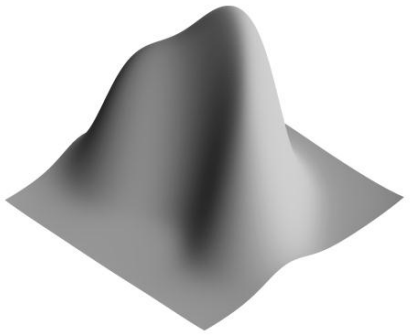
truth



BRT



GP



advantages of GPs

add **prior knowledge** (mean function)

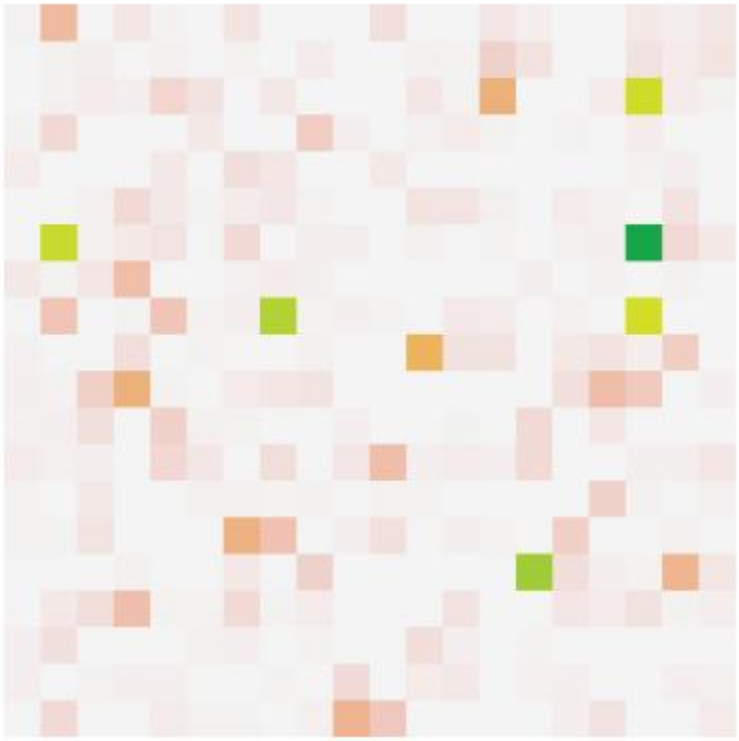
marginalise uncertainty from **polygon** occurrences

efficient (approximations)

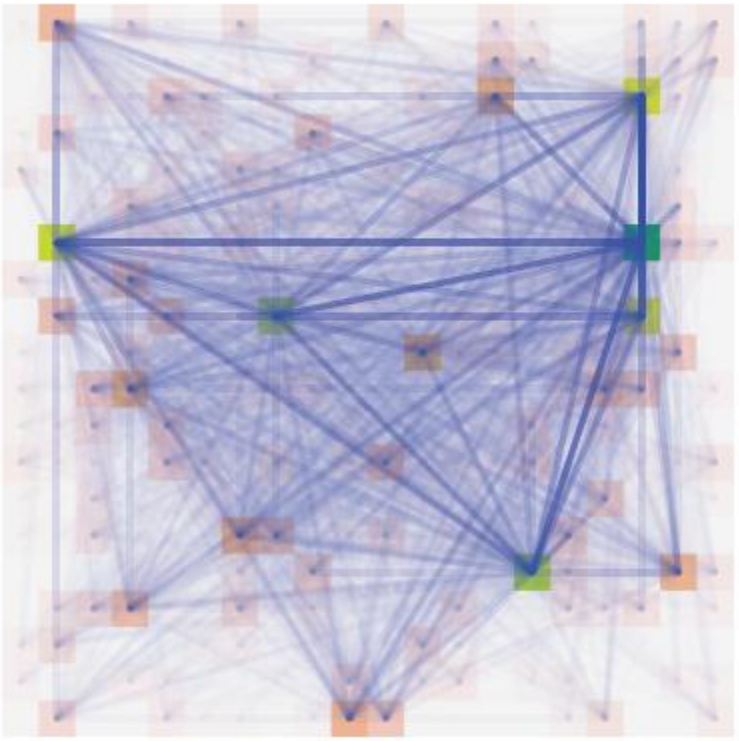
extendable (**spatial** models ++)

movement models

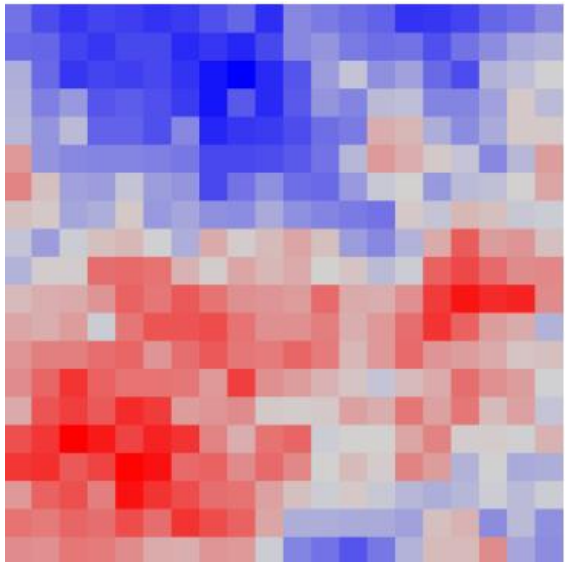
human population



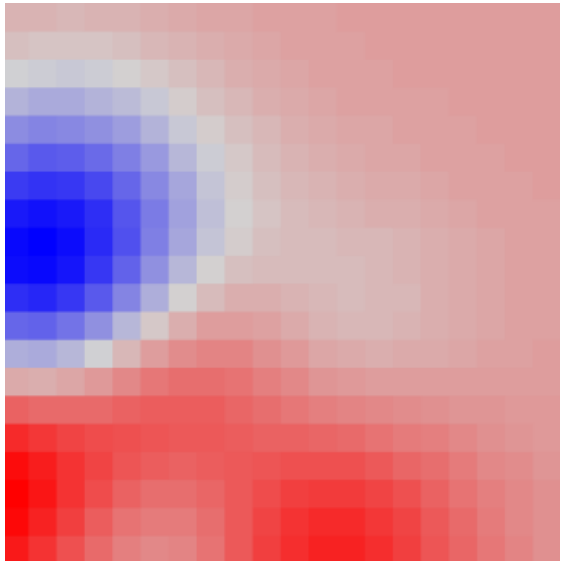
predicted movement



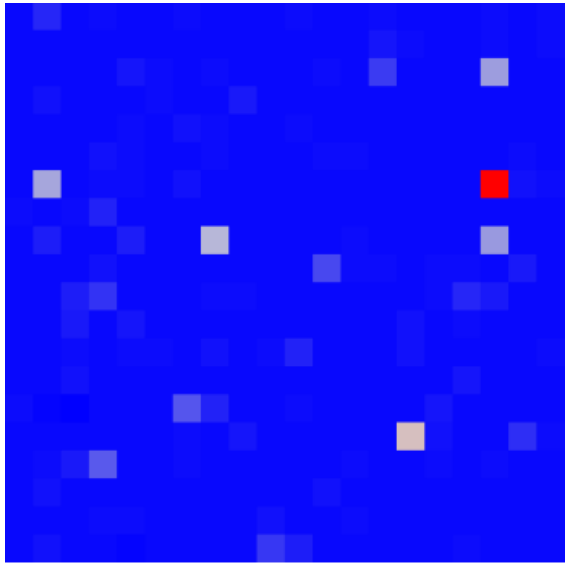
environmental



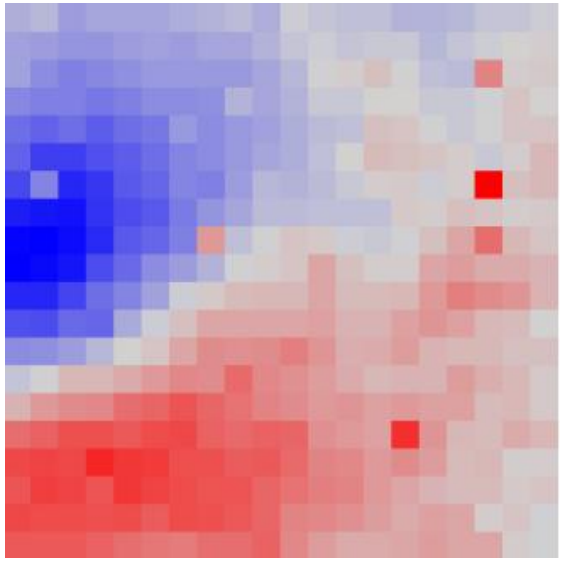
spatial



movement



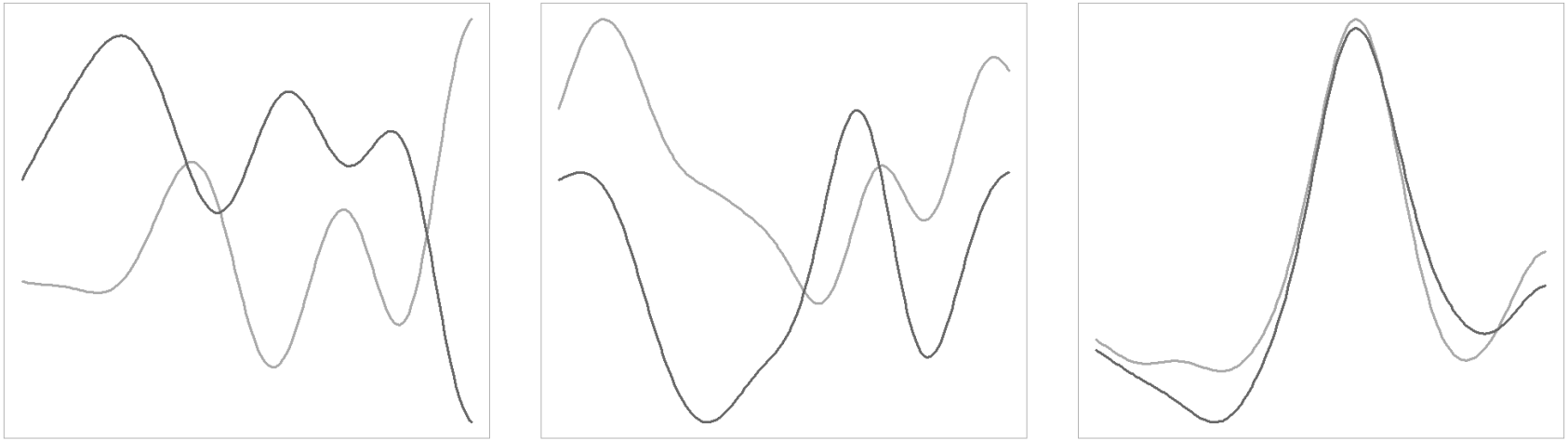
all



biotic interactions

correlation between diseases or control measures

convolved GPs



seeg.zoo.ox.ac.uk

map.ox.ac.uk

BILL & MELINDA GATES foundation

MRC Medical Research Council

wellcome trust

tweaks - anisotropy

GP captures **environmental drivers**

these are **not isotropic**

deformation of sphere

inclination **angle**

axis ratio