Spatio-temporal Modelling of Dengue Risk

Towards an Early Warning System for Brazil

Trevor Bailey

College of Engineering, Mathematics and Physical Sciences

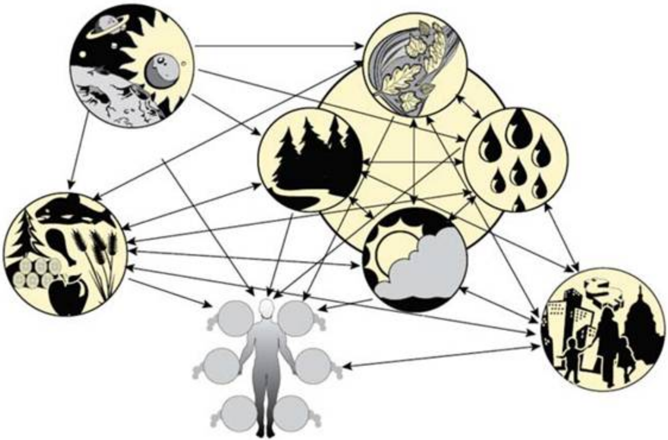
Exeter University, UK.

(t.c.bailey@exeter.ac.uk)

Piracicaba, Brazil, July 2015

R Lowe, X Rodó (IC3), D Stephenson, T Jupp (UoE), R Graham (Met Office), C Coelho (CPTEC), M Sá Carvalho, C Barcellos, G Coelho (FIOCRUZ), A Monteiro (INPE)

Modelling climate (or weather) impacts on health is tricky!



A very complex set of interacting systems is involved

Modelling climate (or weather) impacts on health is tricky!

Even if we just focus on the more ‘direct issues’ then:

Modelling climate (or weather) impacts on health is tricky!

Even if we just focus on the more 'direct issues' then:

- **Data is largely 'available' rather than from a 'designed' study** (confounding factors/latent structures are rife)

Modelling climate (or weather) impacts on health is tricky!

Even if we just focus on the more 'direct issues' then:

- **Data is largely 'available' rather than from a 'designed' study**
  (confounding factors/latent structures are rife)
- **Relationships are inevitably very noisy**
  (signals are weak)

Modelling climate (or weather) impacts on health is tricky!

Even if we just focus on the more ‘direct issues’ then:

- **Data is largely ‘available’ rather than from a ‘designed’ study**
  (confounding factors/latent structures are rife)
- **Relationships are inevitably very noisy**
  (signals are weak)
- **Data is usually mixture of spatial and temporal observations**
  (need to allow for complex correlation structures)

Modelling climate (or weather) impacts on health is tricky!

Even if we just focus on the more 'direct issues' then:

- **Data is largely 'available' rather than from a 'designed' study**
  (confounding factors/latent structures are rife)
- **Relationships are inevitably very noisy**
  (signals are weak)
- **Data is usually mixture of spatial and temporal observations**
  (need to allow for complex correlation structures)
- **Data is often multi-scale**
  (differentially aggregated or averaged over time/space)

Modelling climate (or weather) impacts on health is tricky!

Even if we just focus on the more 'direct issues' then:

- **Data is largely 'available' rather than from a 'designed' study** (confounding factors/latent structures are rife)
- **Relationships are inevitably very noisy** (signals are weak)
- **Data is usually mixture of spatial and temporal observations** (need to allow for complex correlation structures)
- **Data is often multi-scale** (differentially aggregated or averaged over time/space)
- **Relationships may involve multi-level (hierarchical) structure** (something leads (uncertainly) to something else which leads (uncertainly) to ... etc.)

Modelling climate (or weather) impacts on health is tricky!

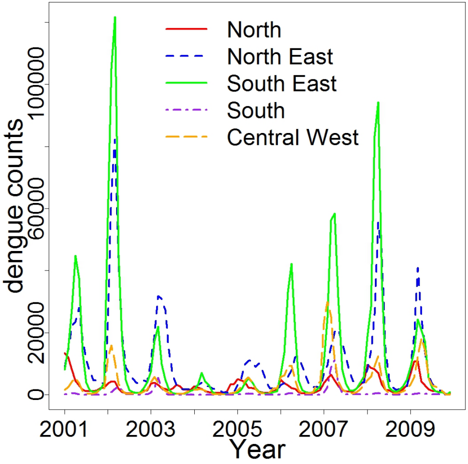
Even if we just focus on the more 'direct issues' then:

- **Data is largely 'available' rather than from a 'designed' study**
  (confounding factors/latent structures are rife)
- **Relationships are inevitably very noisy**
  (signals are weak)
- **Data is usually mixture of spatial and temporal observations**
  (need to allow for complex correlation structures)
- **Data is often multi-scale**
  (differentially aggregated or averaged over time/space)
- **Relationships may involve multi-level (hierarchical) structure**
  (something leads (uncertainly) to something else which leads (uncertainly) to ... etc.)
- **Relationships may exhibit threshold or extreme dependencies**
  (rather than average behaviour)

Dengue in Brazil

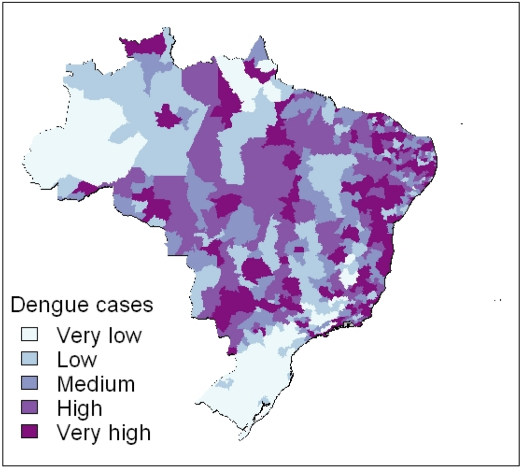
- Dengue transmitted by *Aedes aegypti* mosquitoes
- Severe joint and muscle pain (rarely fatal)
- Epidemics depend on mosquito density and distribution, virus circulation and human susceptibility
- Brazil has more cases of dengue than anywhere else in the world
- More than 3 million cases in Brazil 2001-2009
- 2008 epidemic: 787,726 cases, 448 deaths
- Seasonal pattern: increases in Jan-May when climate warmer/humid
- Early warning systems that account for multiple dengue risk factors, are required to implement timely control measures
- Seasonal climate forecasts provide potential to anticipate dengue epidemics several months in advance.

Temporal variability in dengue in Brazil



Monthly dengue counts for main regions of Brazil 2001-2009

Spatial variability in dengue in Brazil



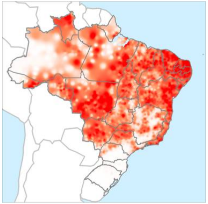
Total dengue cases in microregions (553) 2001-2009

Dengue transmission

- Epidemiological drivers, e.g.
  - Susceptible population
  - Sero-type circulation
- Human drivers, e.g.
  - population growth/urbanisation/poverty (substandard housing)
  - abundance of water-storage (containers/bad drainage)
- Environmental drivers, e.g.
  - Precipitation (filling of containers)
  - Temperature/humidity (mosquito development)

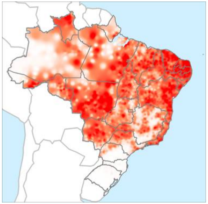
Some questions?

- Is it possible to develop a model to provide spatio-temporal probabilistic forecasts of dengue risk?

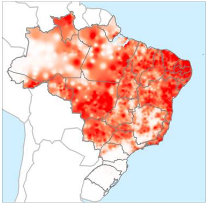


Some questions?

- Is it possible to develop a model to provide spatio-temporal probabilistic forecasts of dengue risk?
  - To what extent can variations in dengue risk be accounted for by climate variations?



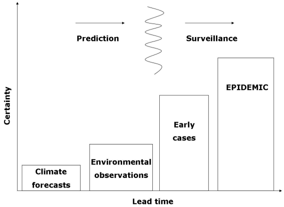
Some questions?



- Is it possible to develop a model to provide spatio-temporal probabilistic forecasts of dengue risk?
  - To what extent can variations in dengue risk be accounted for by climate variations?
  - Which observed and unobserved non-climatic confounding factors should be incorporated?

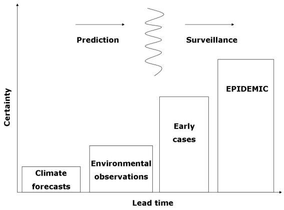
Some more questions?

- Is climate information useful in a dengue Early Warning System (EWS) for Brazil?



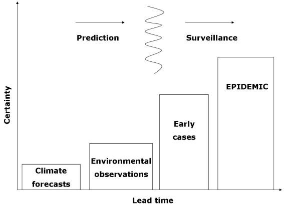
Some more questions?

- Is climate information useful in a dengue Early Warning System (EWS) for Brazil?
  - How well can the developed model predict future and geographically specific dengue epidemics?



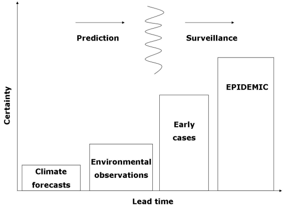
Some more questions?

- Is climate information useful in a dengue Early Warning System (EWS) for Brazil?
  - How well can the developed model predict future and geographically specific dengue epidemics?
  - How does this compare with current ‘surveillance and response’ approach in Brazil (observe early dengue cases Dec/Jan then estimate epidemic potential for late austral summer)



Some more questions?

- Is climate information useful in a dengue Early Warning System (EWS) for Brazil?
  - How well can the developed model predict future and geographically specific dengue epidemics?
  - How does this compare with current 'surveillance and response' approach in Brazil (observe early dengue cases Dec/Jan then estimate epidemic potential for late austral summer)
  - How can early warnings of dengue epidemics based on climate information be effectively communicated to public health decision makers?



Disease and Demographic Data

Disease data SINAN-DATASUS

- Monthly dengue cnts (originally Jan 2001 - Dec 2009, but now until 2013)
- Spatial unit: microregion

Census/cartographic data SIDRA-IBGE

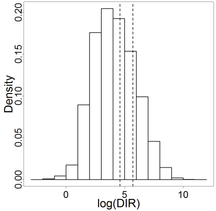
- % urban population
- Altitude
- Administrative region
- Zone or Biome (e.g. Atlantic/Amazon Rainforest)

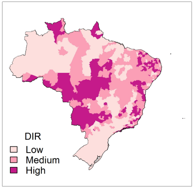
Original dataset: 108 months, 553 locations

DIR$=\frac{y_{st}}{p_{st}} \times 12 \times 100,000$

Low: DIR < 100
Med: $100 <$ DIR < 300

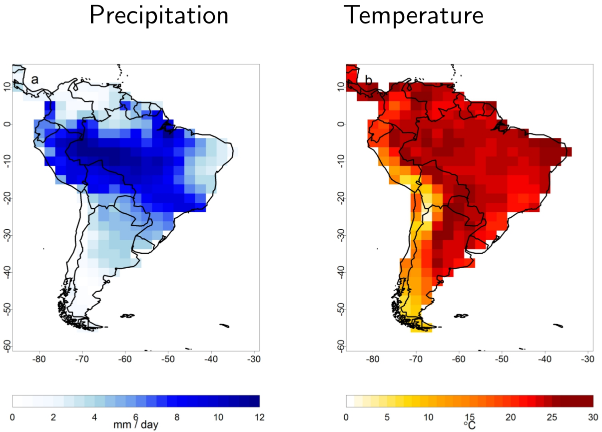
High: DIR > 300





Gridded climate data $(2.5^\circ \times 2.5^\circ)$

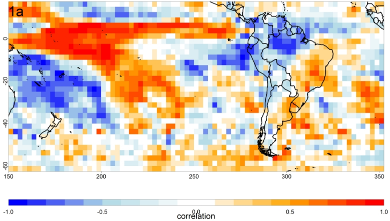
- Average precipitation rate (GPCP)
- Reanalysis average temperature (NCEP/NCAR)



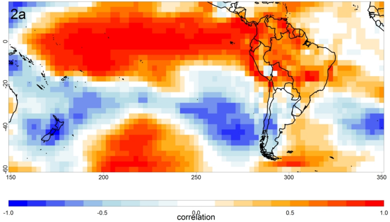
Dec-Feb climatology (2000-9)

ENSO

Precipitation



Temperature



Correlation Oceanic Niño Index (ONI) vs Dec-Feb precipitation & temperature

GLMM model framework

$$
\begin{aligned}
y_{st} | \phi_s, \nu_s, \omega_{t'(t)} &\sim \text{NegBin}(\mu_{st}, \kappa); \quad s = 1, \ldots, 553;\ t = 1, \ldots, 108 \\
\log \mu_{st} &= \underbrace{\log e_{st}}_{\text{offset}} + \alpha + \underbrace{\delta_{1t'(t)} + \delta_{2s'(s)} + \delta_{3s'(s)t'(t)}}_{\text{month+zone factors}} \\
&+ \underbrace{\gamma_1 w_{1st} + \gamma_2 w_{2s}}_{\text{non-climate vars: pop dens+altitude}} \\
&+ \underbrace{\beta_{1s'(s)} x_{1,s,t-2} + \beta_{2s'(s)} x_{2,s,t-2} + \beta_{3s'(s)} x_{3,t-6}}_{\text{climate vars: precip+temp+ONI}} \\
&+ \underbrace{\phi_s + \nu_s}_{\text{spatial random effects}} + \underbrace{\omega_{t'(t)}}_{\text{monthly random effects}}
\end{aligned}
$$

$$
\begin{aligned}
& t'(t) = 1, \ldots, 12 \\
& s'(s) = 1, \ldots, 8 \\
& \phi_s \sim \text{N}(0, \sigma^2_\phi); \quad s = 1, \ldots, 553 \\
& (\nu_1, \ldots, \nu_{553}) \sim \text{CAR}(\sigma^2_\nu) \\
& \omega_1 \sim \text{N}(\omega_{12}, \sigma^2_\omega) \\
& \omega_{t'(t)} \sim \text{N}(\omega_{t'(t)-1}, \sigma^2_\omega); \quad t'(t) = 2, \ldots, 12
\end{aligned}
$$

GLMM model conclusions

- Climate signal is weak but highly significant

GLMM model conclusions

- Climate signal is weak but highly significant
- Precipitation and temperature averaged over preceding 3 month period, 2 month lag with dengue. (particularly seems to help in accounting for spatial variation)

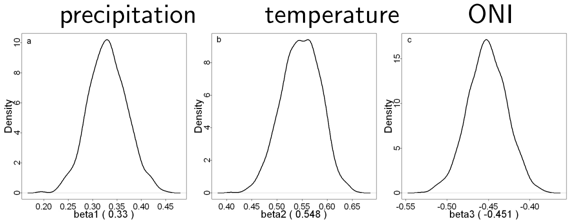
GLMM model conclusions

- Climate signal is weak but highly significant
- Precipitation and temperature averaged over preceding 3 month period, 2 month lag with dengue. (particularly seems to help in accounting for spatial variation)
- ONI lagged 6 months with dengue, 4 months with climate variables (particularly seems to help in temporal variation)

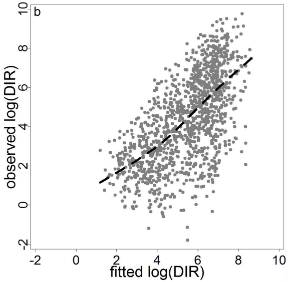
GLMM model conclusions

- Climate signal is weak but highly significant
- Precipitation and temperature averaged over preceding 3 month period, 2 month lag with dengue. (particularly seems to help in accounting for spatial variation)
- ONI lagged 6 months with dengue, 4 months with climate variables (particularly seems to help in temporal variation)
- Random effects are important
  - Unobserved confounding factors (population immunity to circulating serotype, health interventions/vector control measures)
  - Overdispersion
  - Temporal correlation and spatial clustering

Selected results - GLMM, SE Brazil



Climate coefficient posteriors

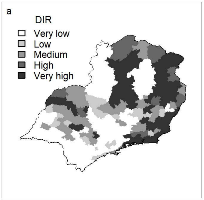


Observed log(DIR) vs model fit, FMA, 2001-2009

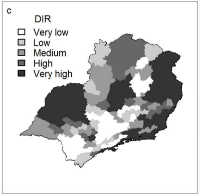
Selected results - GLMM, SE Brazil, FMA season

2008 (epidemic year)

Observed

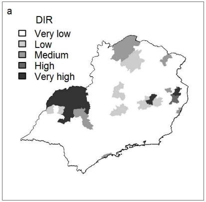


Predicted

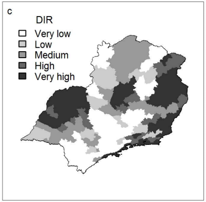


2005 (non-epidemic year)

Observed



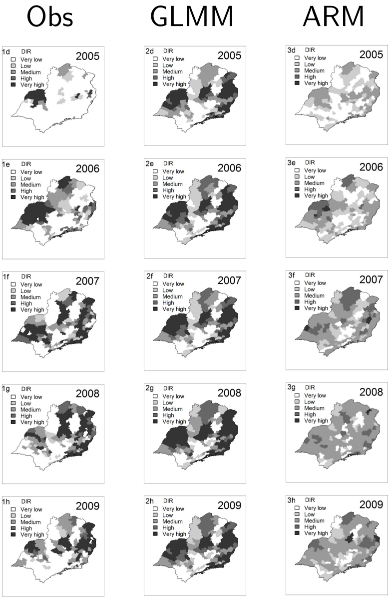
Predicted



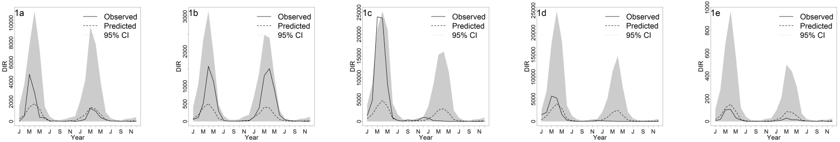
GLMM and current surveillance practice, SE Brazil, FMA

Current surveillance practice effectively equates to the auto-regressive model (ARM):

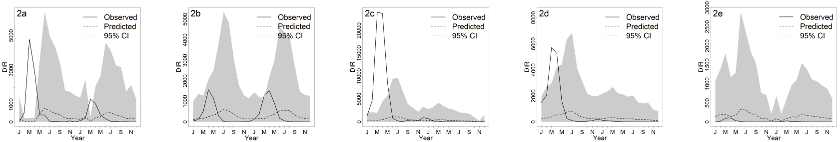
$$
\begin{aligned}
y_{st} &\sim \text{NegBin}(\mu_{st}, \kappa) \\
\log \mu_{st} &= \log e_{st} + \alpha + \beta \log\left(\frac{y_{s,t-3}}{e_{s,t-3}}\right)
\end{aligned}
$$



GLMM

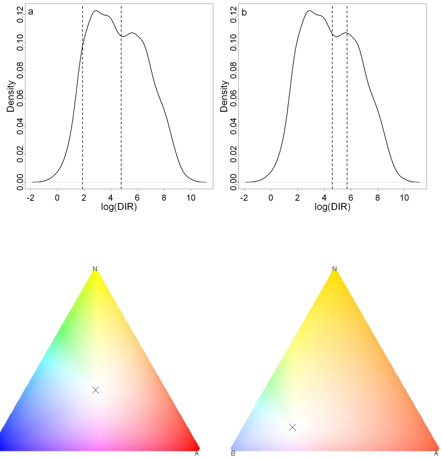


ARM



(a) Três Marias, (b) Belo Horizonte, (c) Baía de Ilha Grande, (d) Rio de Janeiro, (e) São Jose dos Campos

Defining and visualising epidemic risk



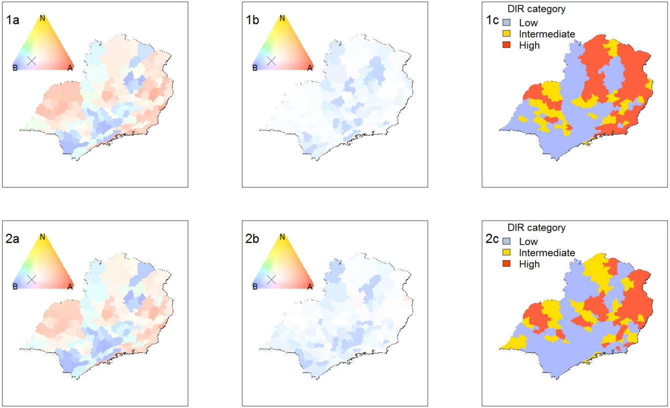
Symmetric (tercile) and non-symmetric (100 and 300 cases per 100,000) category boundaries of the observed distribution of DIR, FMA 2001-2007, SE Brazil

Visualising GLMM probabilistic forecasts

GLMM | ARM | Observed

SE, FMA 2008

SE, FMA 2009



Epidemic prediction: FMA 2008, SE Brazil, GLMM

Posterior predictive results in 160 microregions in SE for DIR exceeding 300 cases per 100,000 at probability decision thresholds (50%&30%)

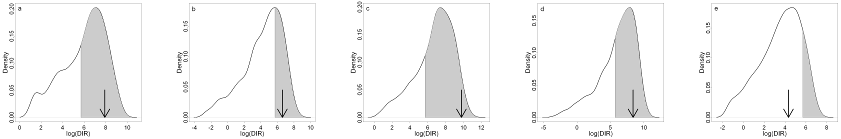
50%		Obs	
		Yes	No
Pred	Yes	31	13
	No	23	93

PC=78%,HR=57%,FAR=12%

30%		Obs	
		Yes	No
Pred	Yes	51	31
	No	3	75

PC=79%,HR=94%,FAR=29%

Posterior predictive distributions and prob of $>$ 300 per 100,000 in 5 selected regions (arrow indicates observed DIR)



(a) Três Marias, (b) Belo Horizonte, (c) Baía de Ilha Grande (d) Rio de Janeiro, (e) São Jose dos Campos

Combined GLMM model framework

$$y_{st}|\phi_s,\nu_s,\omega_{t'(t)} \sim \text{NegBin}(\mu_{st}\,,\kappa); \quad s=1,\ldots,553;\ t=1,\ldots,108$$

$$\log\mu_{st} = \underbrace{\log e_{st}}_{\text{offset}} + \alpha + \underbrace{\delta_{1t'(t)} + \delta_{2s'(s)} + \delta_{3s'(s)t'(t)}}_{\text{month+zone factors}}$$

$$+ \underbrace{\gamma_1 w_{1st} + \gamma_2 w_{2s}}_{\text{non-climate vars: pop dens+altitude}} + \underbrace{\lambda z_{s,t-4}}_{\text{lagged log SMR}}$$

$$+ \underbrace{\beta_{1s'(s)}x_{1,s,t-2} + \beta_{2s'(s)}x_{2,s,t-2} + \beta_{3s'(s)}x_{3,t-6}}_{\text{climate vars: precip+temp+ONI}}$$

$$+ \underbrace{\phi_s + \nu_s}_{\text{spatial random effects}} + \underbrace{\omega_{t'(t)}}_{\text{monthly random effects}}$$

$$t'(t) = 1,\ldots,12$$
$$s'(s) = 1,\ldots,8$$
$$\phi_s \sim \text{N}(0,\sigma^2_\phi); \quad s=1,\ldots,553$$
$$(\nu_1,\ldots,\nu_{553}) \sim \text{CAR}(\sigma^2_\nu)$$
$$\omega_1 \sim \text{N}(\omega_{12},\sigma^2_\omega)$$
$$\omega_{t'(t)} \sim \text{N}(\omega_{t'(t)-1},\sigma^2_\omega); \quad t'(t)=2,\ldots,12$$

Posterior predictive results in 160 microregions in SE for DIR exceeding 300 per 100,000 at probability decision thresholds (50%&30%)

50%		Obs Yes	Obs No
Pred	Yes	34	10
	No	20	96

PC=81%, HR=63%, FAR=9%

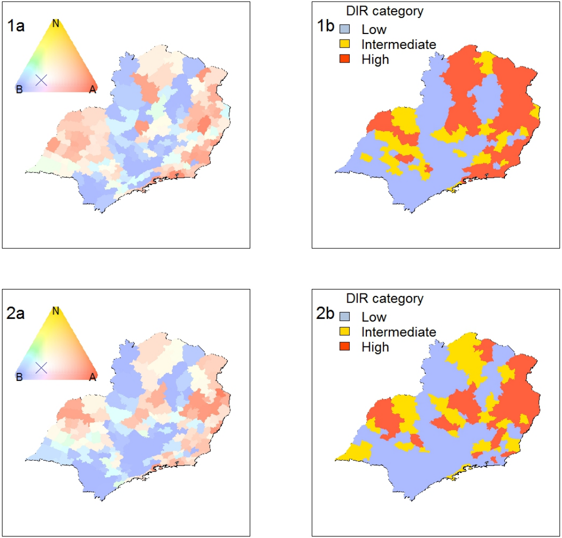
30%		Obs Yes	Obs No
Pred	Yes	47	24
	No	7	82

PC=81%, HR=87%, FAR=23%

Epidemic prediction combined model

Combined GLMM

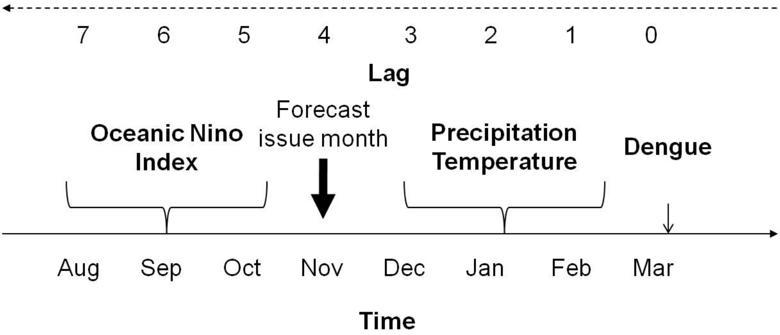
Observed



FMA 2008

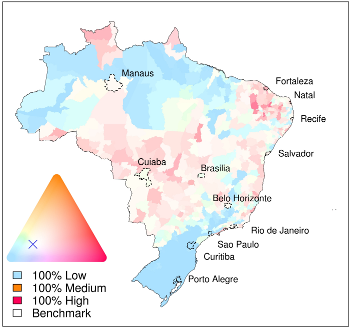
FMA 2009

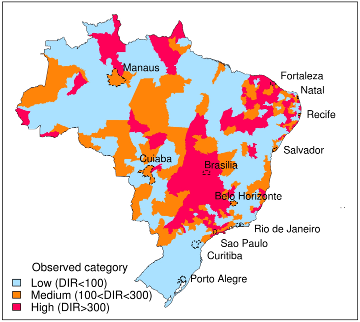
Extending prediction lead-time with forecast climate



Forecasting Dengue Risk Levels for the World Cup

Framework applied to predict dengue risk for June 2014 during the World Cup in Brazil, a mass gathering of more than 3 million local/international spectators.

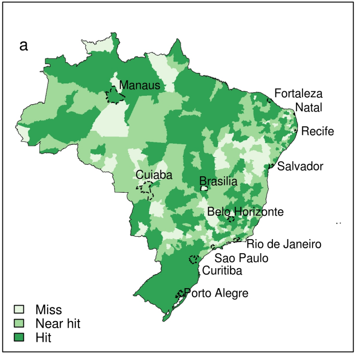




Evaluation of June 2014 Forecasts on National Basis

		Observed Category			
		Low	Medium	High	Total
Forecast Category	Low	193 (34.9%)	49 (8.9%)	40 (7.2%)	282
	Medium	50 (9.0%)	20 (3.6%)	26 (4.7%)	96
	High	38 (6.9%)	47 (8.5%)	90 (16.3%)	175
	Total	281	116	156	n=553

Hit: 54.8% **Near hit**: 31.1% **Miss**: 14.1%



Did it make a difference?

- This **timely** warning complimented the national dengue control programme action plan, implemented ahead of the World Cup.

- Results disseminated to the **general public** and visitors travelling to Brazil (European Centre for Disease Control health risk assessment, UK National Health Service, >18 international press outlets, e.g. BBC) raising general **awareness** about dengue for **travellers** to endemic regions.

- Case study in WHO/WMO and UNISDR publications.

- White House "Predict the Next Pandemic" Initiative – dengue model intercomparison project.

Conclusions and Future Work

Lowe, R., Bailey T. et al. (2010), Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil, *Computers and Geosciences*

Lowe, R., Bailey T. et al. (2012) The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil, *Statistics in Medicine*

Lowe, R., Bailey T. et al. (2014) Dengue outlook for the World Cup in Brazil: an early warning model framework driven by real-time seasonal climate forecasts *The Lancet: Infectious Diseases*

Conclusions and Future Work

Lowe, R., Bailey T. et al. (2010), Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil, *Computers and Geosciences*

Lowe, R., Bailey T. et al. (2012) The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil, *Statistics in Medicine*

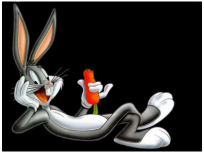
Lowe, R., Bailey T. et al. (2014) Dengue outlook for the World Cup in Brazil: an early warning model framework driven by real-time seasonal climate forecasts *The Lancet: Infectious Diseases*

- Ongoing collaboration between public health and climate institutions and experts, including data managers, mathematical modellers and policy makers (vocabulary and local knowledge)
- Timely access to data (disease, human/vector/host structure, socio-economic, climate-observations, hindcasts, forecasts).
- Incorporation of serotype information, disease transmission process, health intervention/prevention information and movement of human hosts
- Iterative evaluation of model assumptions and predictive performance
- Communication to decision makers and the general public
- Transformation of a case study into a sustainable service

Modelling climate (or weather) impacts on health is tricky!

Even ignoring the systems implications:

- **Data is largely 'available' rather than from a 'designed' study**
- **Relationships are inevitably very noisy**
- **Data is usually mixture of spatial and temporal observations**
- **Data is often multi-scale**
- **Relationships may involve multi-level (hierarchical) structure**
- **Relationships may exhibit threshold or extreme dependencies**



... *THANKS & THAT'S ALL FOLKS*! ...