

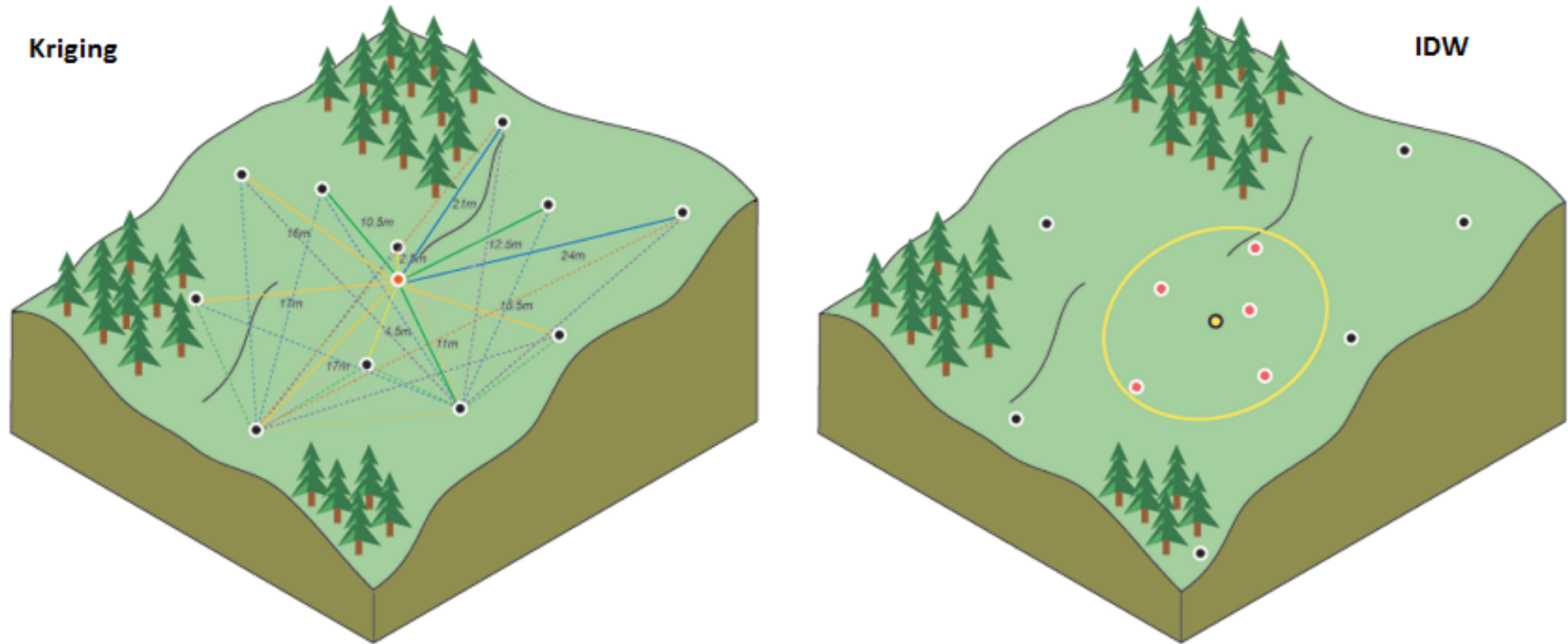
Deep learning and spatial mapping

Dr Samir Bhatt

Imperial College London

bhattsamir@gmail.com

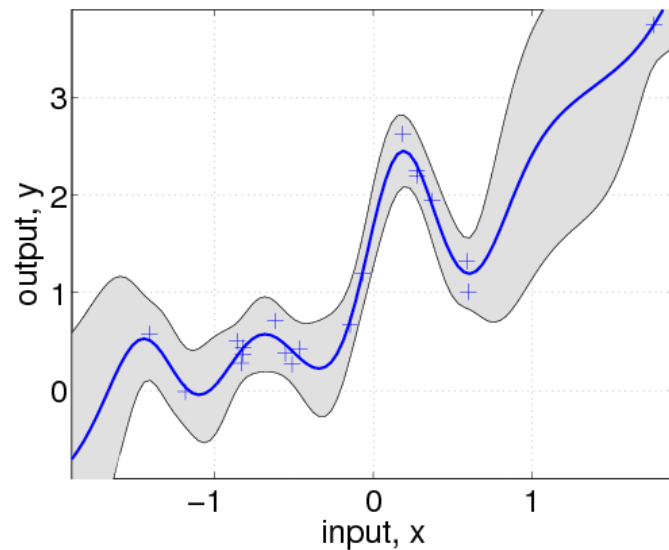
Fundamentally what is spatial statistics



A weighting of nearby points

$$f(\cdot) = \sum_{i=1}^n \alpha_i k(\cdot, x_i)$$

Expensive



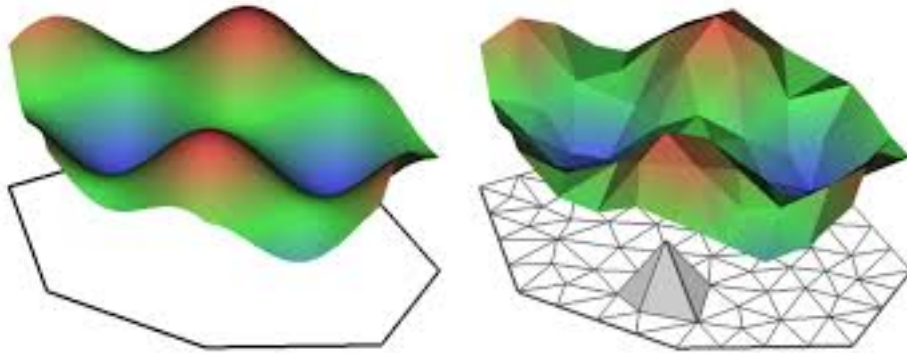
Either inverting a matrix or solving a linear system of equations

$$X|Y \sim \mathcal{N}(\mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - \mu_Y), \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX})$$

$$Y|X \sim \mathcal{N}(\mu_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(X - \mu_X), \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY})$$

$$Y = \beta X^T + \epsilon$$

Spatial Statistics - lots of innovation



Spatial Analysis Made Easy with Linear Regression and Kernels

Philip Milton

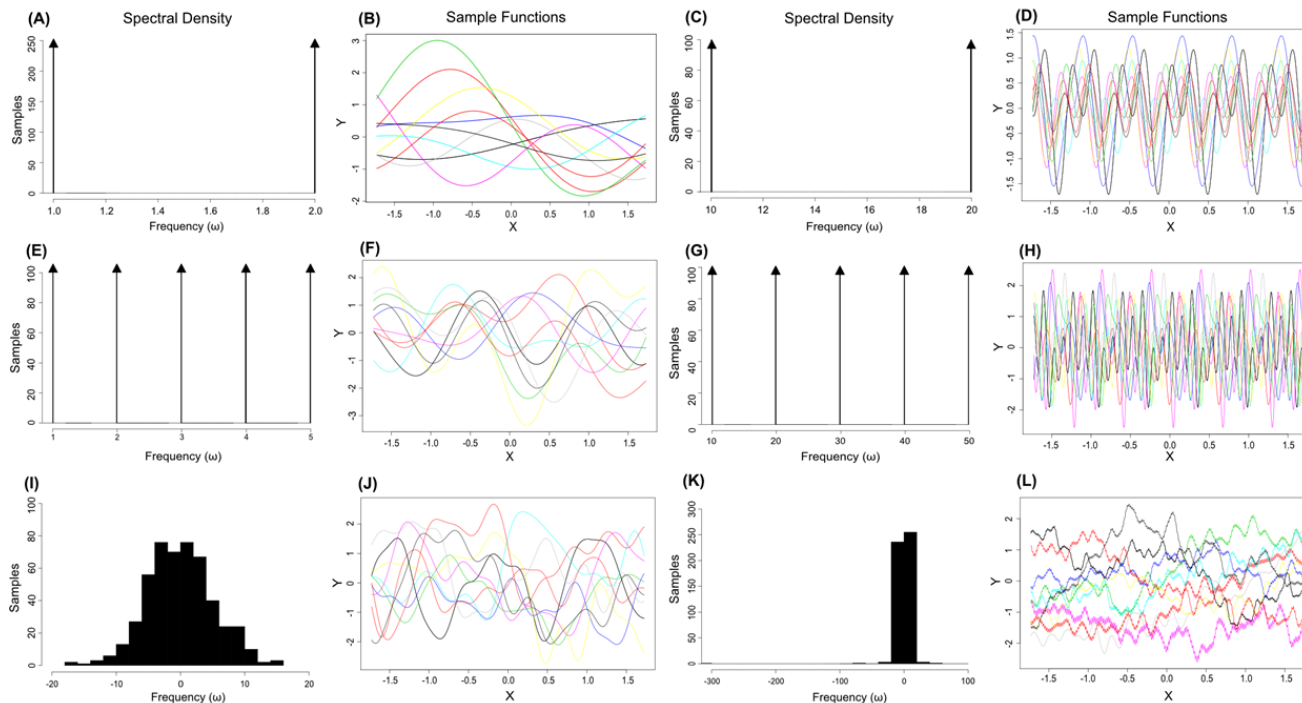
MRC Centre for Outbreak Analysis and Modelling
Department of Infectious Disease Epidemiology
Imperial College London, London, UK
PM5215@ic.ac.uk

Emanuele Giorgi

CHICAS, Lancaster Medical School
Lancaster University, Lancaster, UK
e.giorgi@lancaster.ac.uk

Samir Bhatt

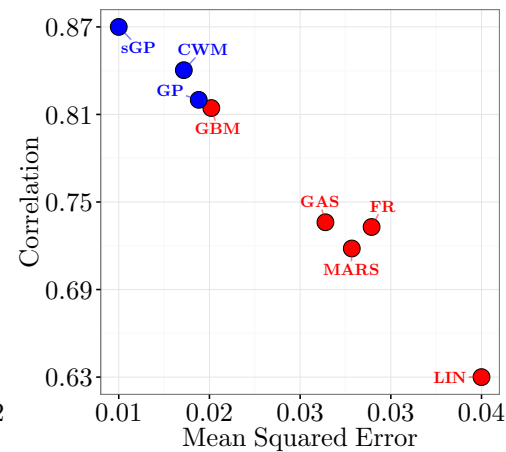
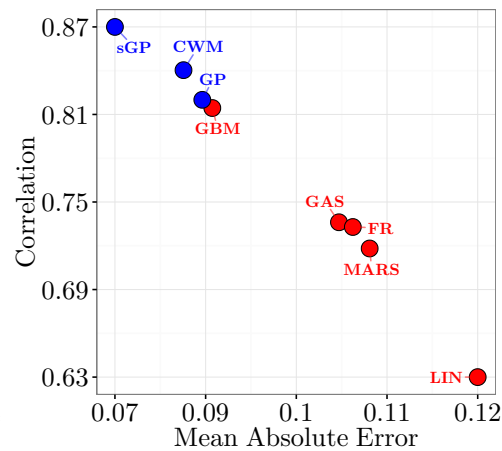
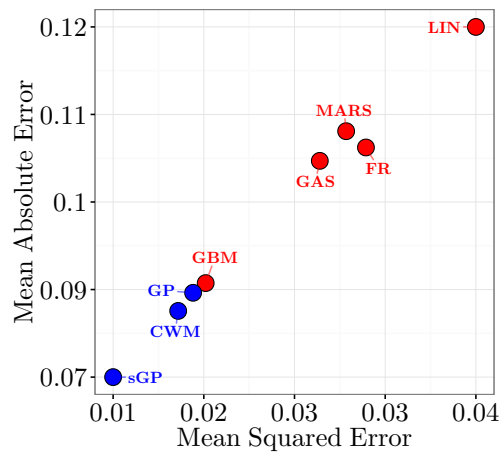
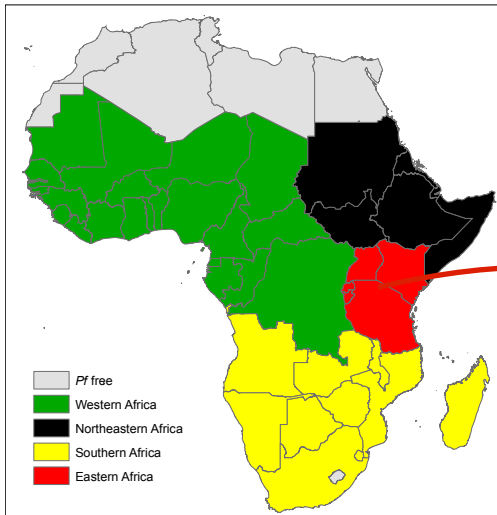
MRC Centre for Outbreak Analysis and Modelling
Department of Infectious Disease Epidemiology
Imperial College London, London, UK
s.bhatt@ic.ac.uk



Improving accuracy

Improved prediction accuracy for disease risk mapping using
Gaussian Process stacked generalisation

Samir Bhatt^{1,*}, Ewan Cameron², Seth R Flaxman⁴, Daniel J Weiss², David L Smith³,
and Peter W Gething²



Improving flexibility

LETTER

OPEN

<https://doi.org/10.1038/s41586-019-1050-5>



Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta



Mapping changes in housing in sub-Saharan Africa from 2000 to 2015

Lucy S. Tusting^{1*}, Donal Bisanzio^{2,3}, Graham Alabaster⁴, Ewan Cameron⁵, Richard Cibulskis⁶, Michael Davies⁷, Seth Flaxman⁸, Harry Gibson⁵, Jakob Knudsen⁹, Charles Mbogo¹⁰, Fredros Okumu^{11,12,13}, Lorenz von Seidlein¹⁴, Daniel Weiss⁵, Steve W. Lindsay¹⁵, Peter W. Gething⁵ & Samir Bhatt^{5,16}

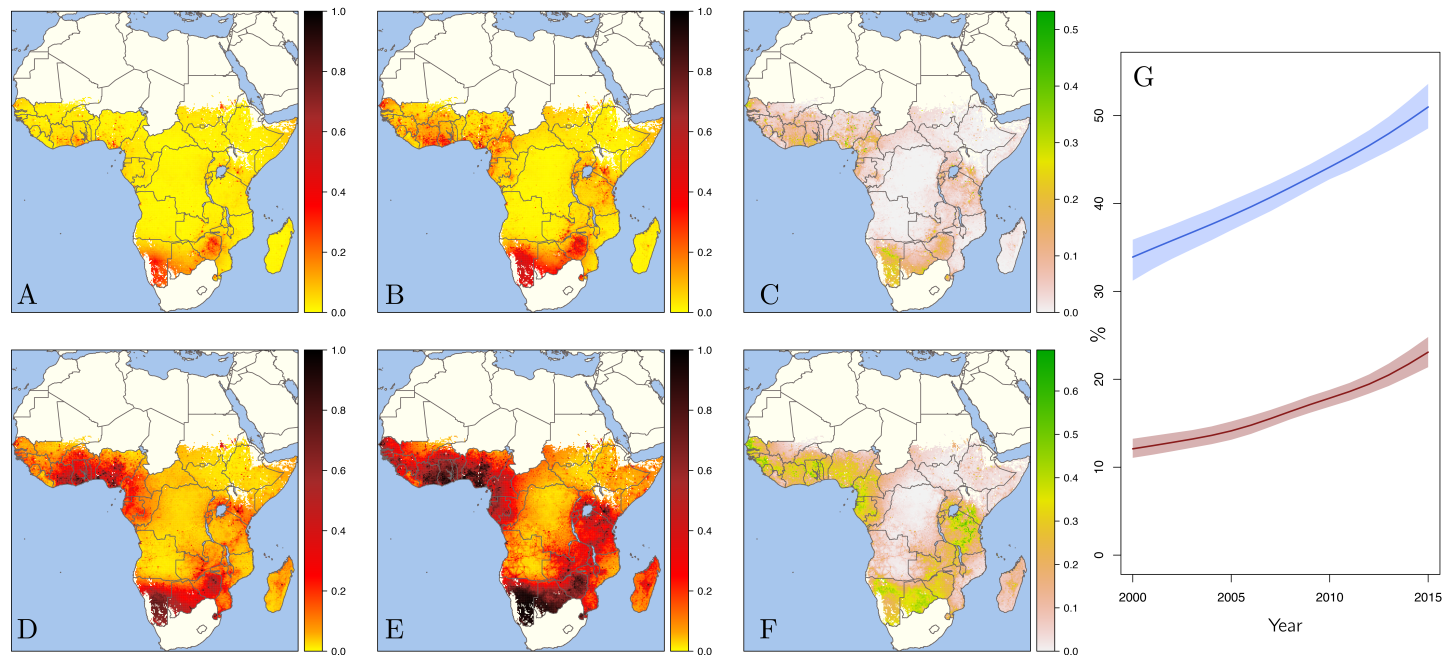
Spatial mapping with Gaussian processes and nonstationary Fourier features

Jean-Francois Ton^a, Seth Flaxman^b, Dino Sejdinovic^a, Samir Bhatt^{c,*}

^a Department of Statistics, University of Oxford, Oxford, OX1 3LB, UK

^b Department of Mathematics and Data Science Institute, Imperial College London, London, SW7 2AZ, UK

^c Department of Infectious Disease Epidemiology, Imperial College London, London, W2 1PG, UK

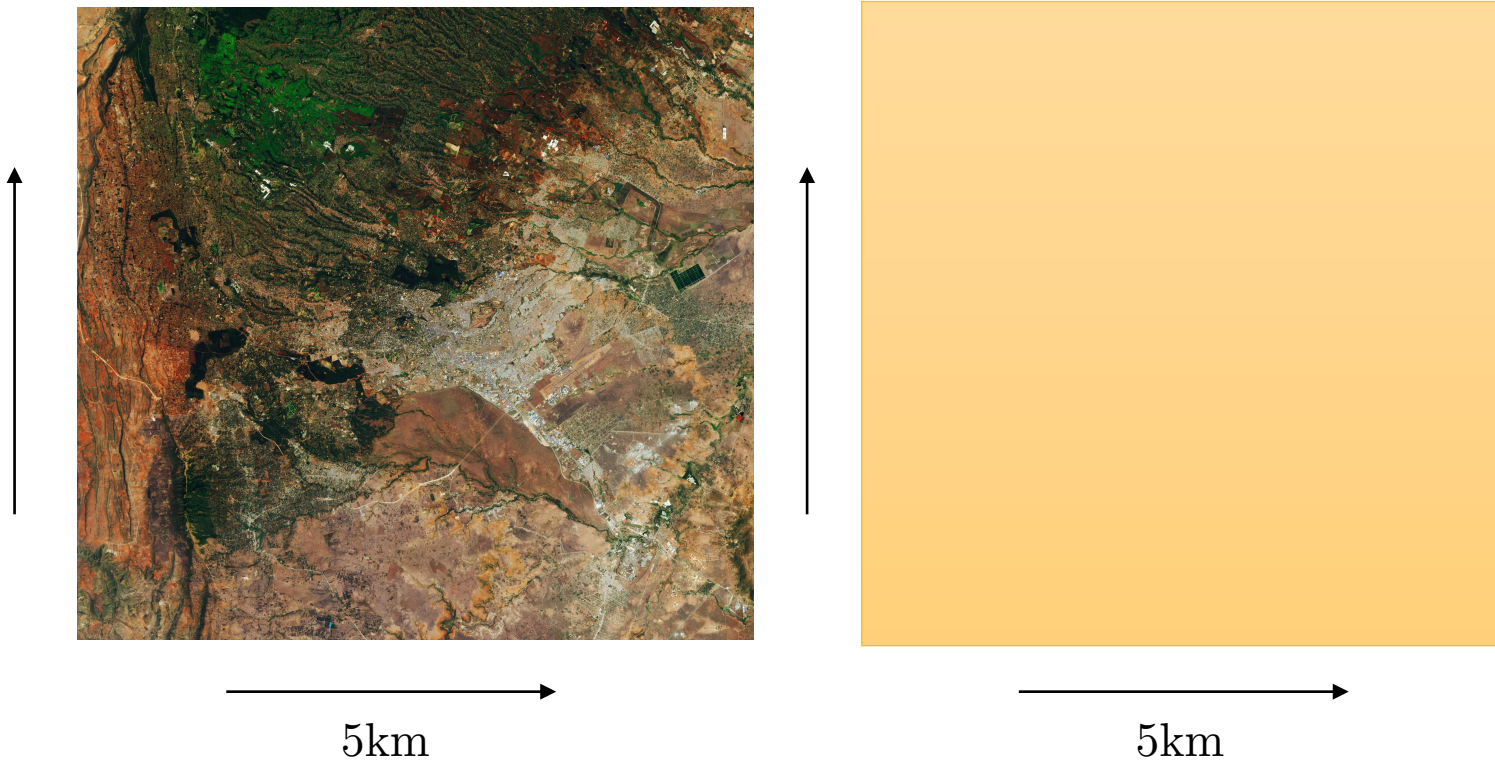


How can we shake things up?



What covariates to use?

1. Modifiable areal unit issues

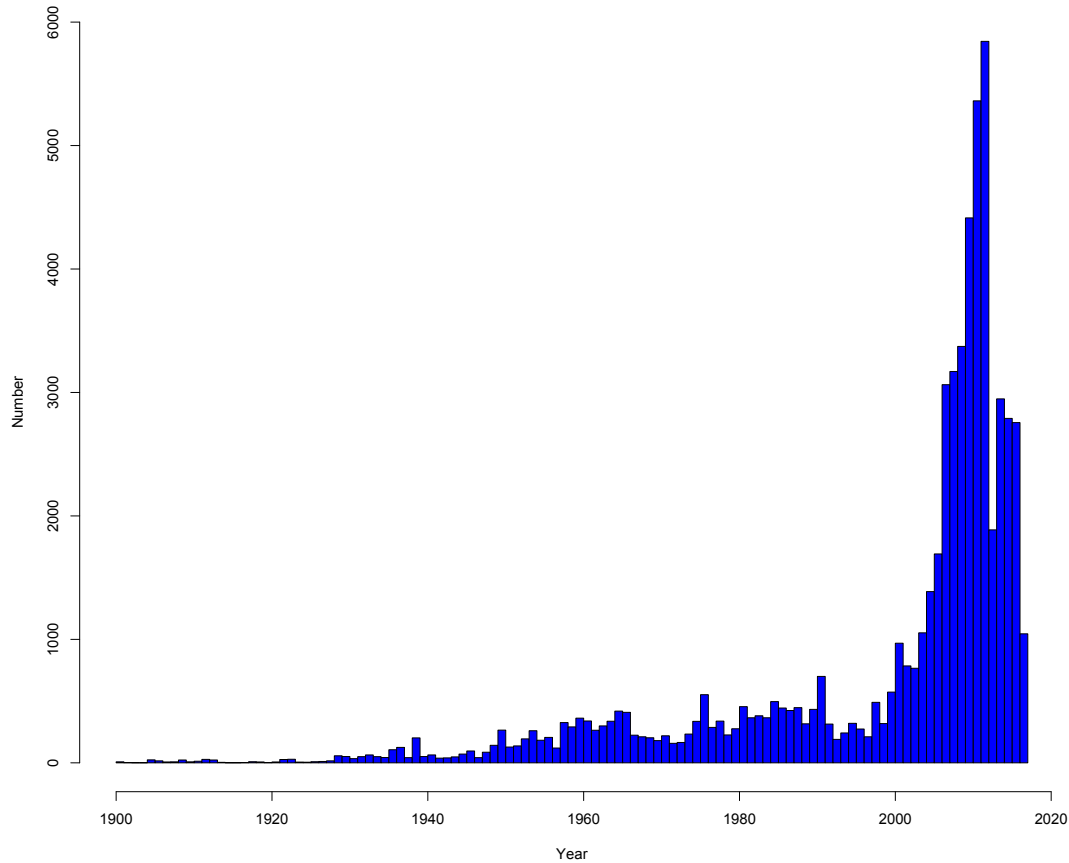


Displacement for *pseudonymising*

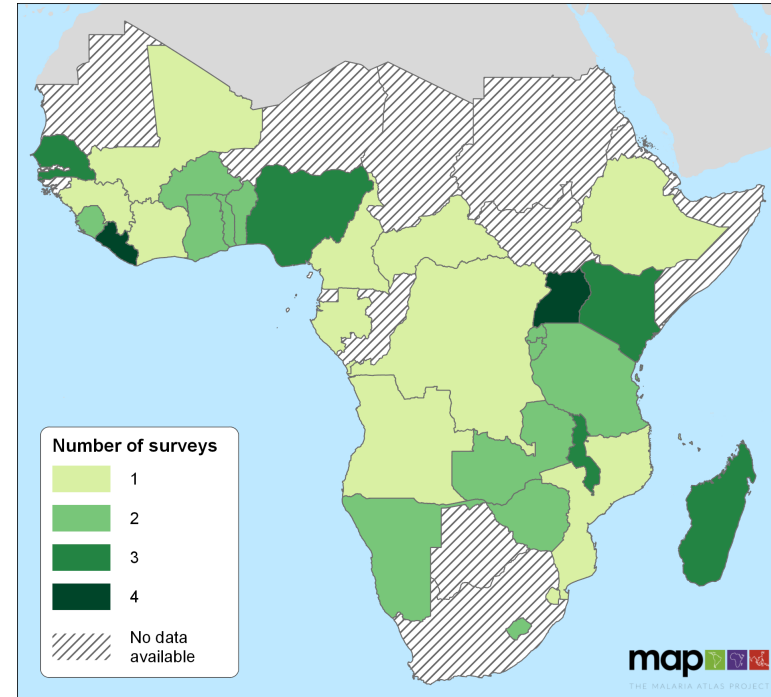


Extrapolation

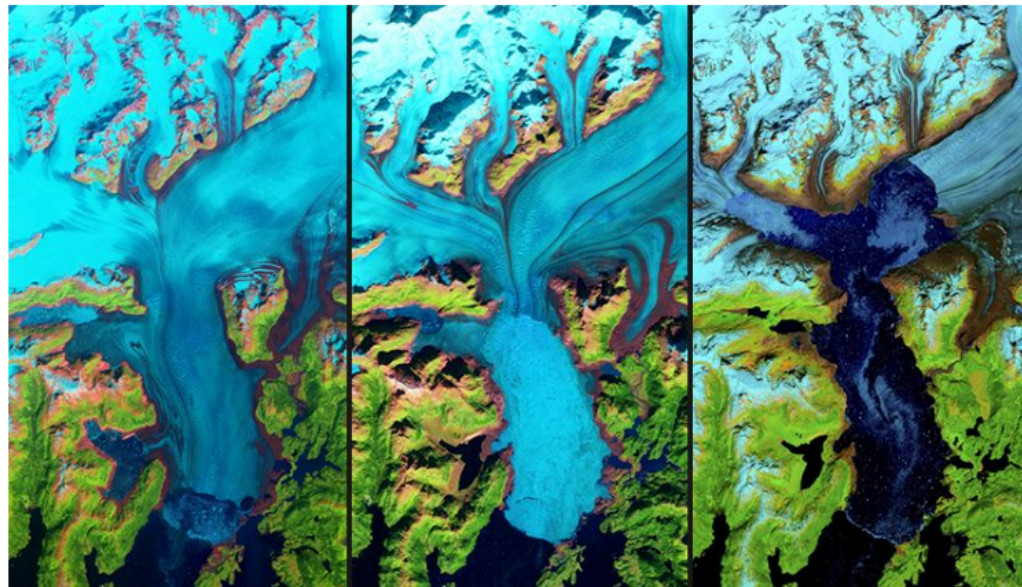
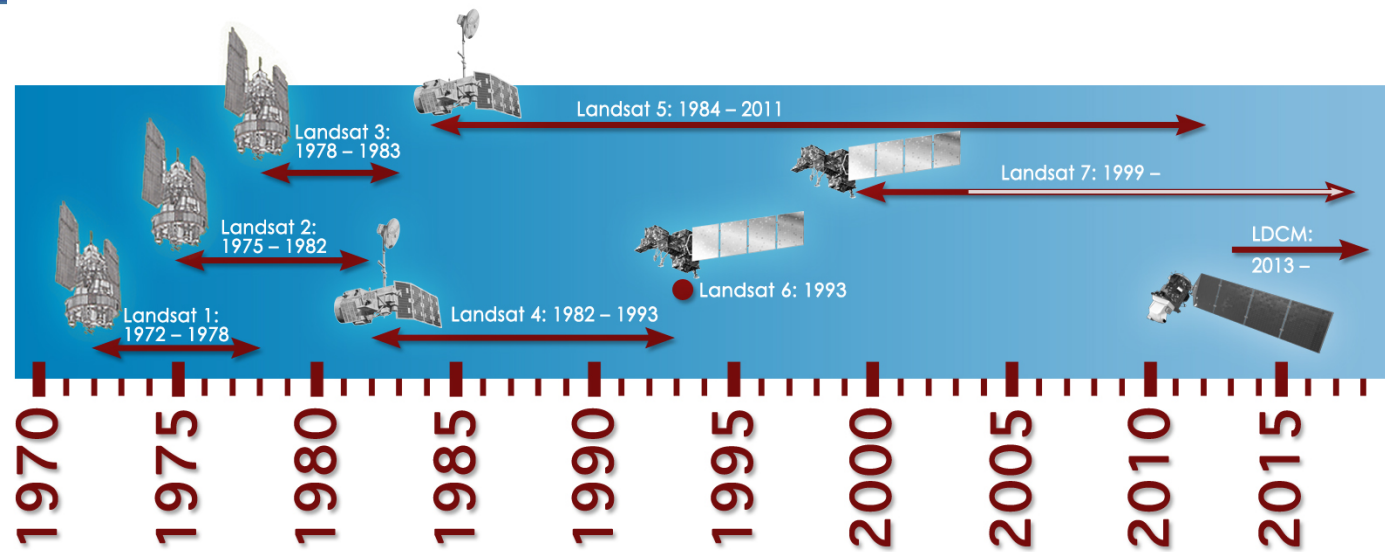
Malaria MAP data



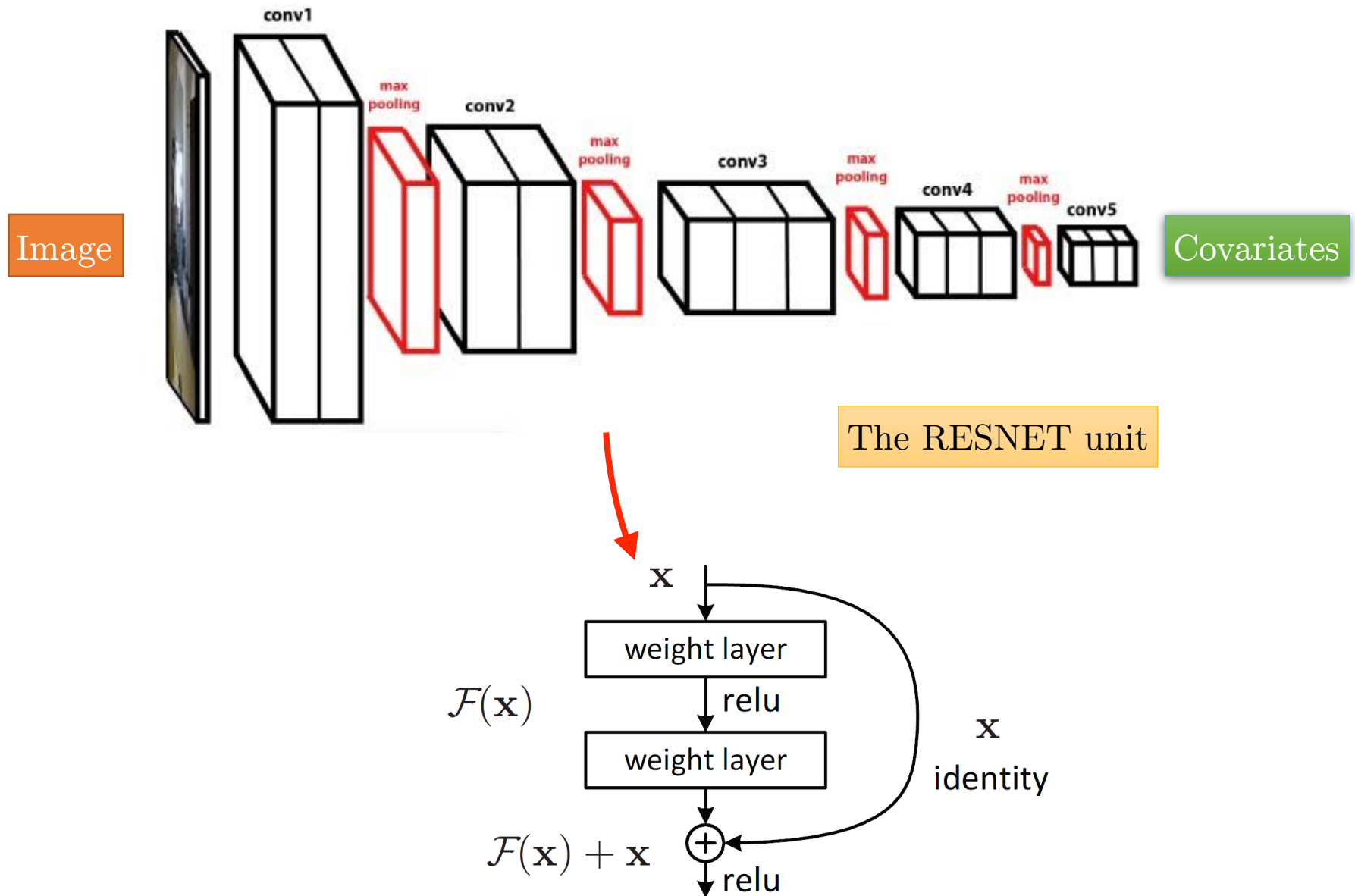
Housing



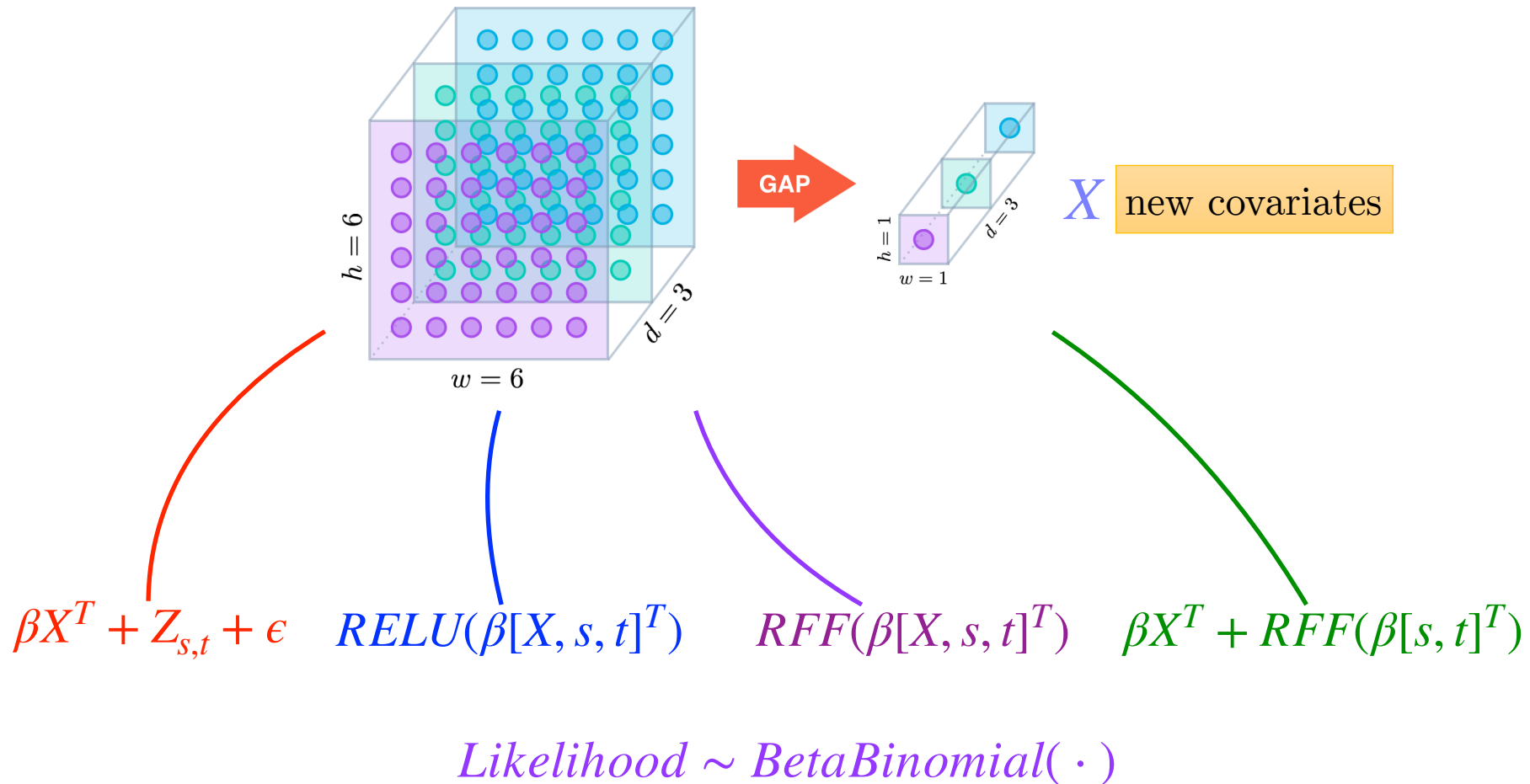
Tracking changes through time



CNNs + Landsat



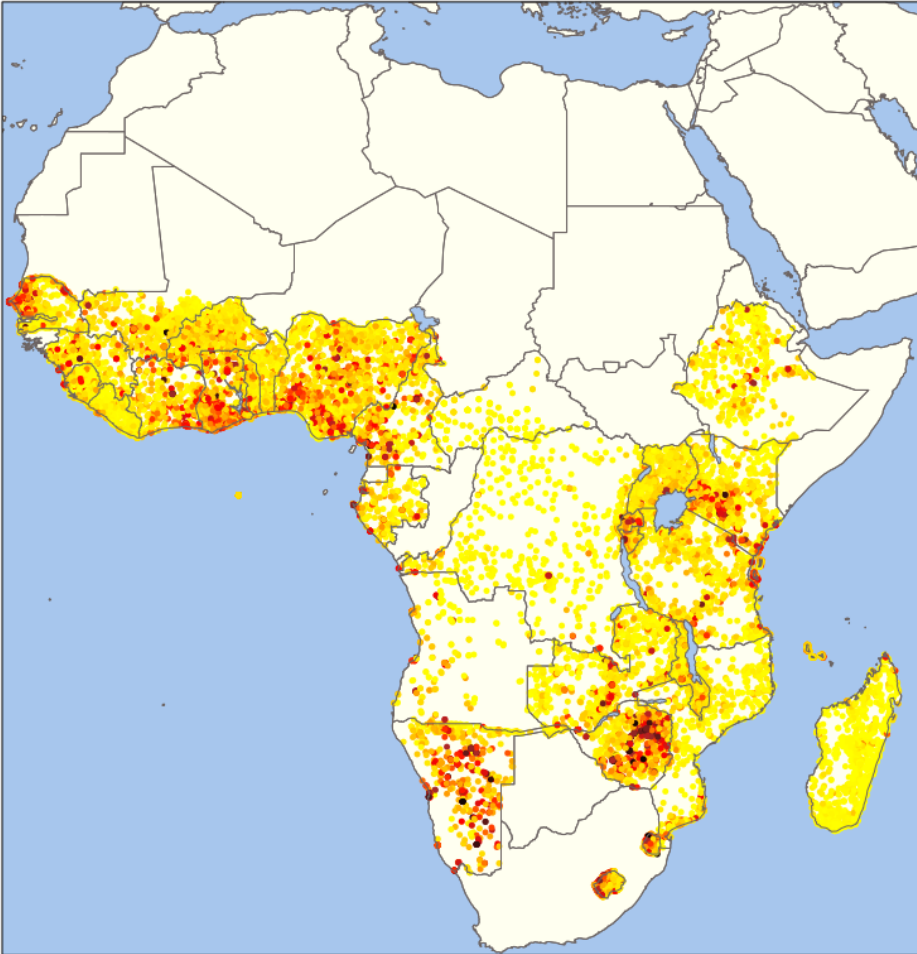
Adding space-time back in



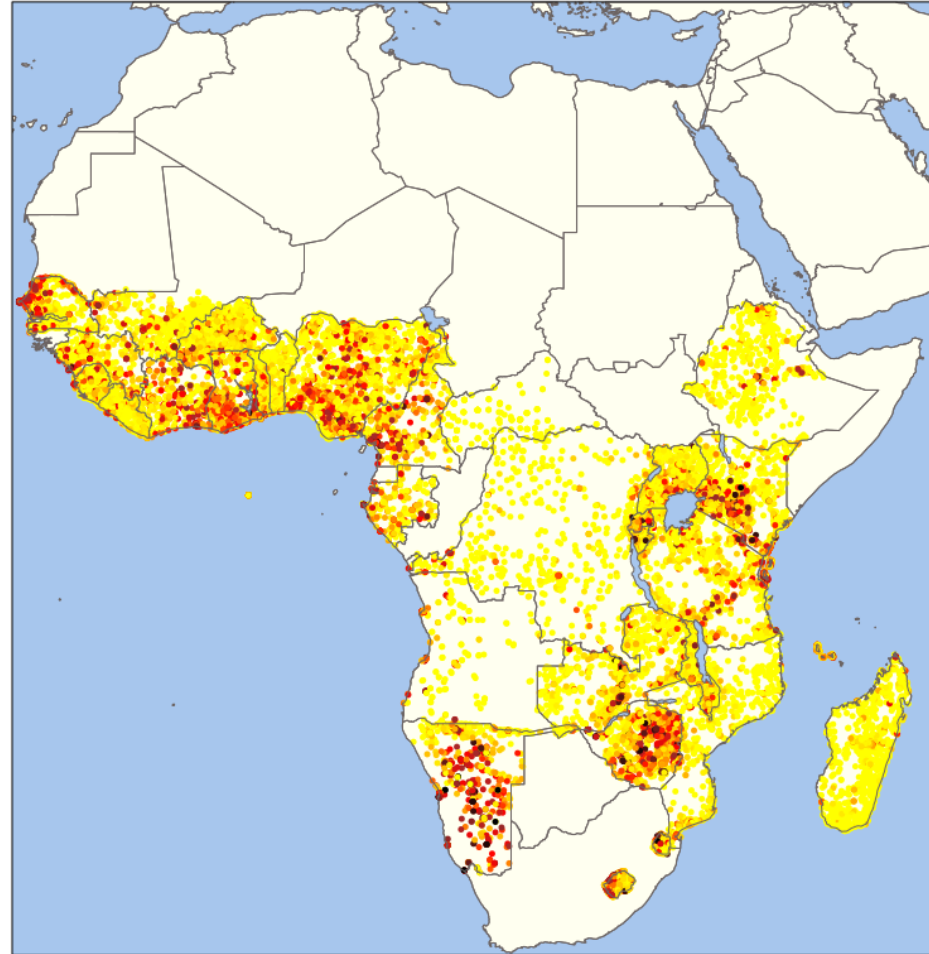
The use of wavelets?

Fit performance

Predicted



Actual



Validation performance

Mean absolute error

$RFF \sim 10 \%$

$CNN \sim 11 \%$

Full bells and whistles

Just RGB images

No time

No space

Mean absolute error - *leave country out*

$RFF \sim 15 \%$

$CNN \sim 11/12 \%$

Full bells and whistles

Just RGB images

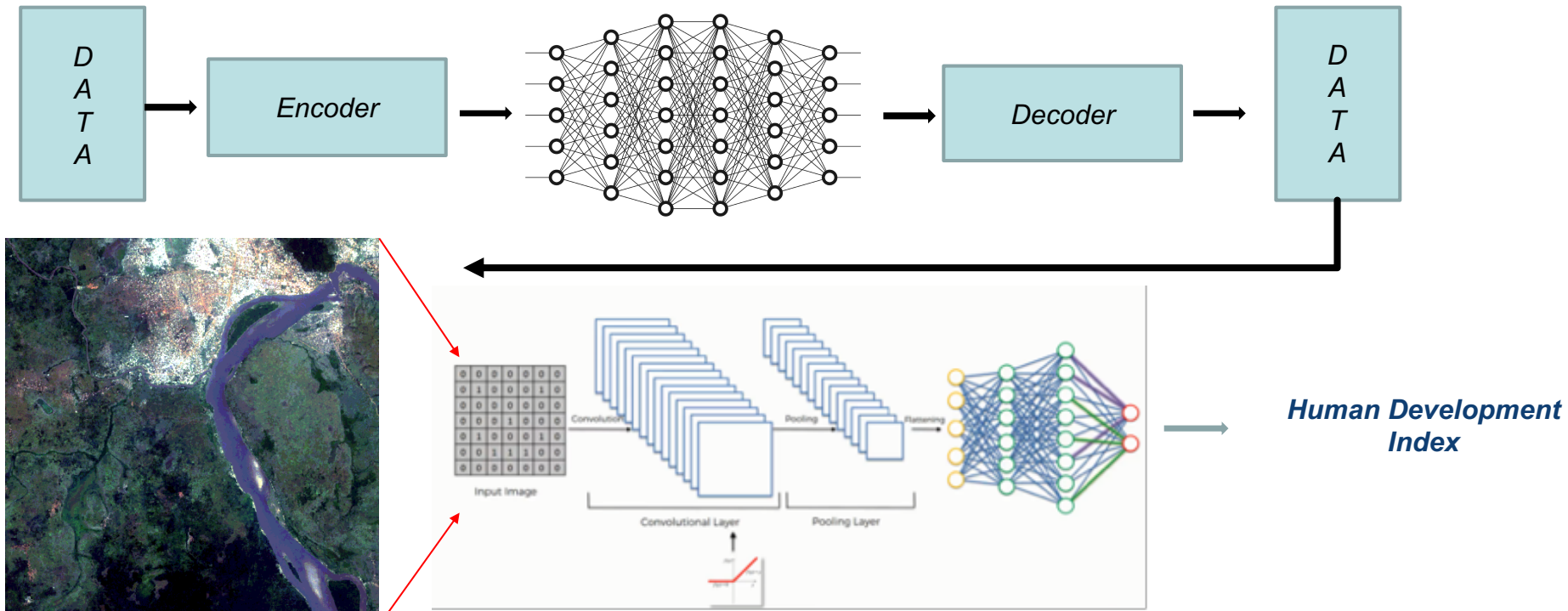
No time

No space

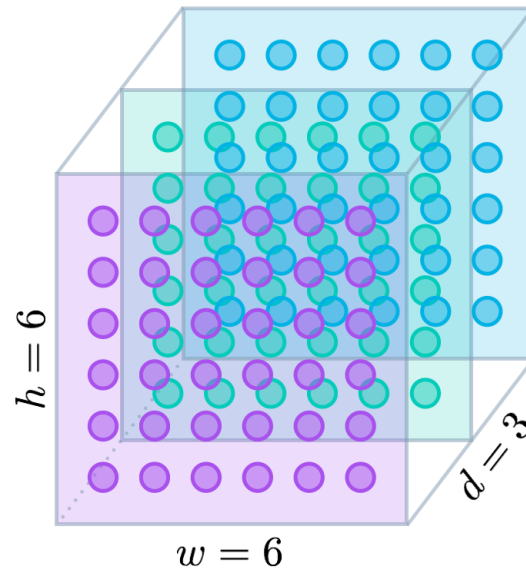
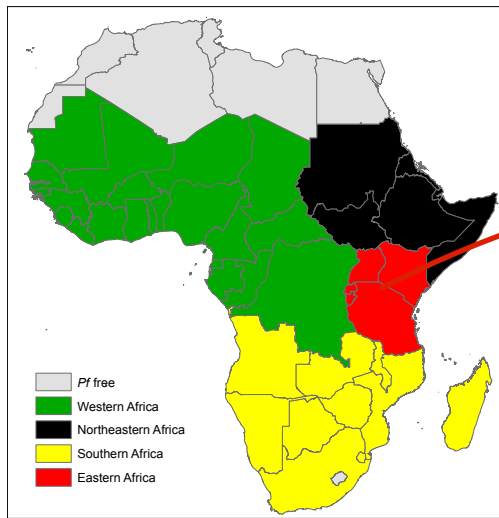
Next steps

1. Get terabytes of Landsat data
2. Investigate various temporal composites
3. Tune CNN hyperparameters
4. Explore multidimensional parameterisations
5. Include DEM and other high res products
6. Explore implications of covariate shift

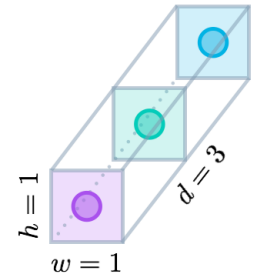
Long term vision (Phase 2 GC)



The human development net



GAP



Collaborators

1. MAP - Harry Gibson
2. Christl Donnelly
3. Seth Flaxman