Gaussian Processes for Machine Learning

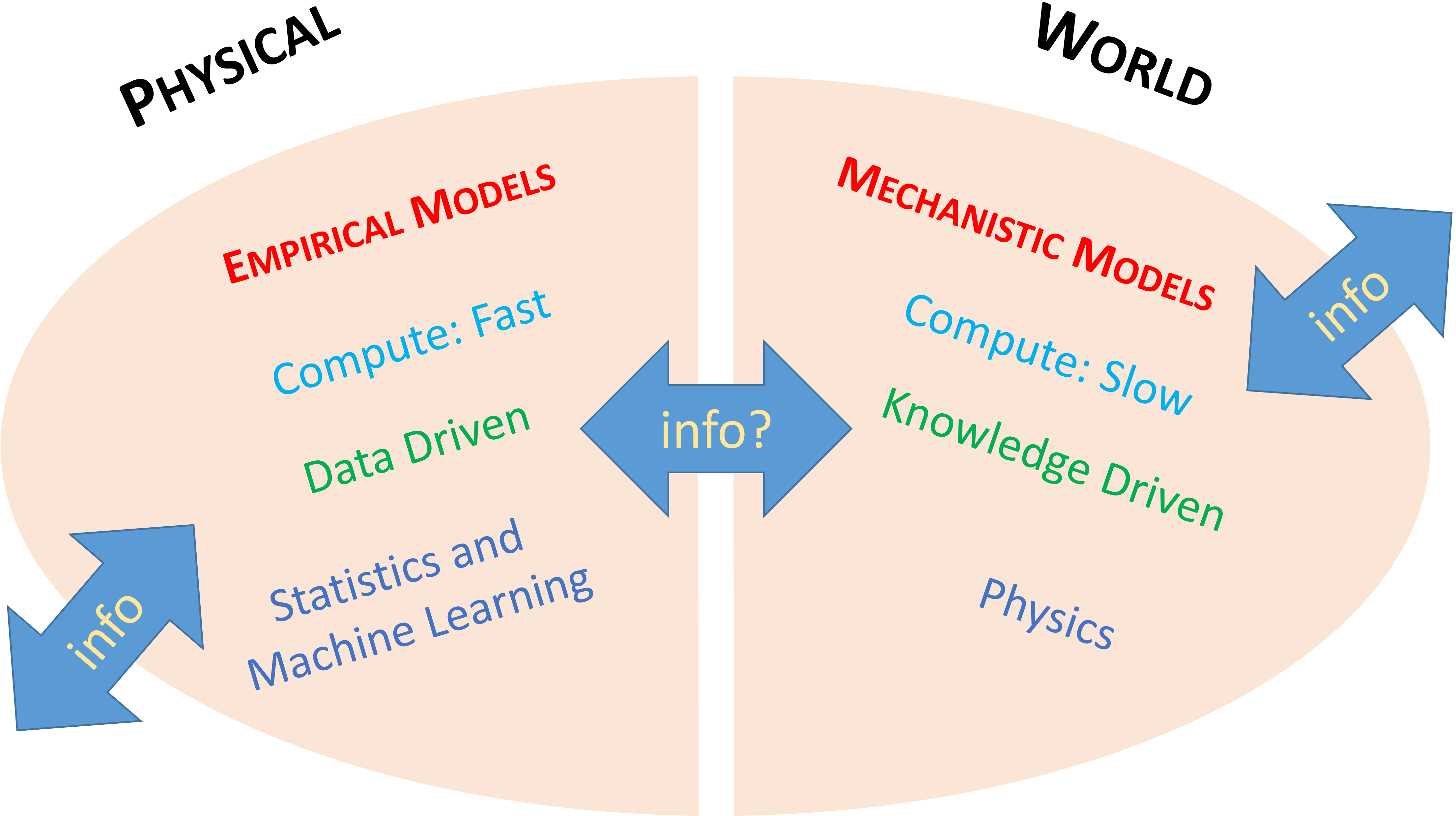
Neil Lawrence
University of Sheffield
29th January 2016
OxWaSP Symposium

choice is determined of itself and without motives.

We ought then to regard the present state of the universe as the effect of its anterior state and as the cause of the one which is to follow. Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. The human mind offers, in the perfection which it has been able to give to astronomy, a feeble idea of this intelligence. Its discoveries in mechanics and geometry, added to that of universal gravity, have enabled it to comprehend in the same analytical expressions the toward the beginning of April, 1759, which was actually verified by observation. The regularity which astronomy shows us in the movements of the comets doubtless exists also in all phenomena.

The curve described by a simple molecule of air or vapor is regulated in a manner just as certain as the planetary orbits; the only difference between them is that which comes from our ignorance.

Probability is relative, in part to this ignorance, in part to our knowledge. We know that of three or a greater number of events a single one ought to occur; but nothing induces us to believe that one of them will occur rather than the others. In this state of indecision it is impossible for us to announce their occurrence with certainty. It is, however, probable that one of these events, chosen at will, will not occur because we see several cases equally possible which exclude its occurrence, while only a single one favors it.



PHYSICAL
WORLD
EMPIRICAL MODELS
Compute: Fast
Data Driven
Statistics and
Machine Learning
info?
MECHANISTIC MODELS
Compute: Slow
Knowledge Driven
Physics
info
info

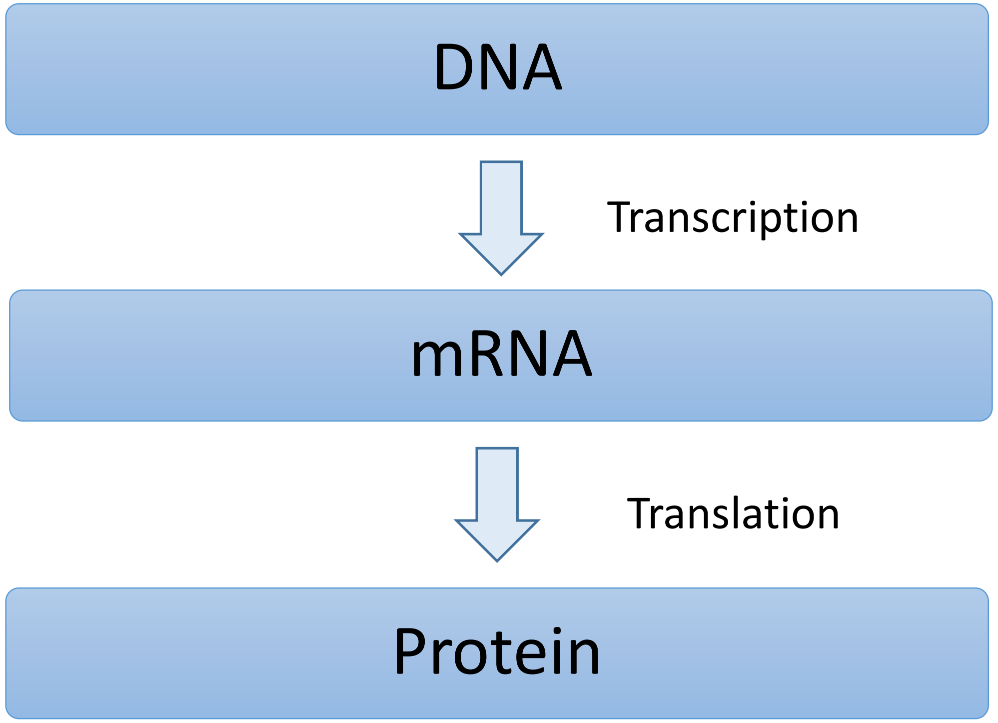
The Data are Not Enough

- Four pillars:
  - Deterministic/Stochastic
  - Mechanistic/Emipirical

- **Goal:** *model complex phenomena over time*
- **Problem:**
  - *Mechanistic* models are often inaccurate
  - Data is often not rich enough for an *empirical* approach

- Question 1: How do we combine *inaccurate physical model* with machine learning?

Central Dogma

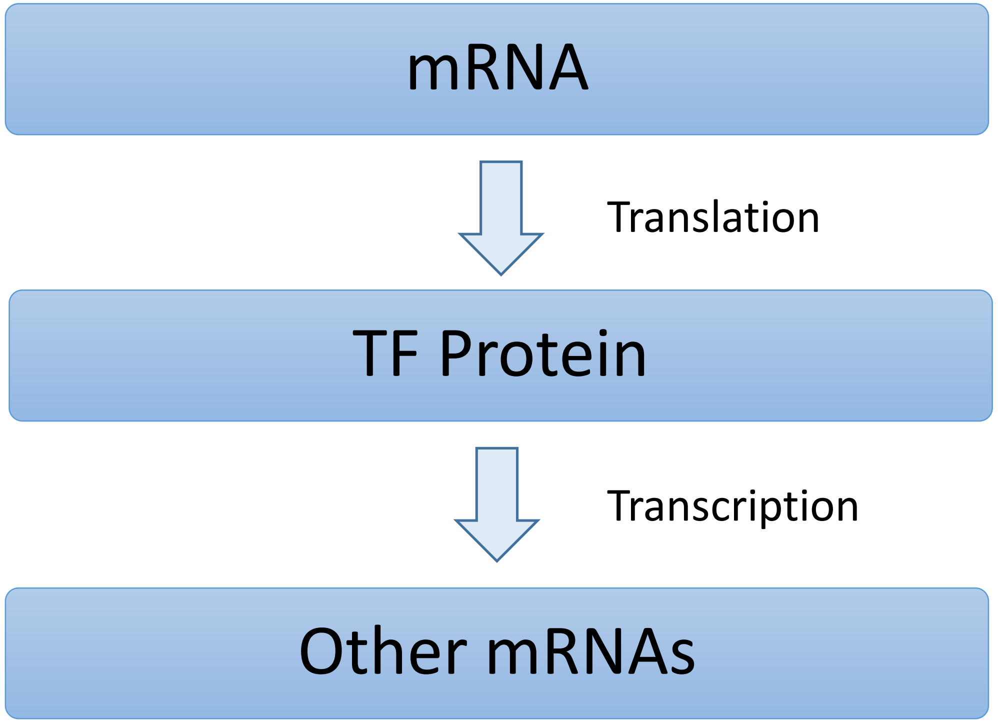


Decision: Transcription Factors

Measured using Microarray since 1998

Difficult to measure

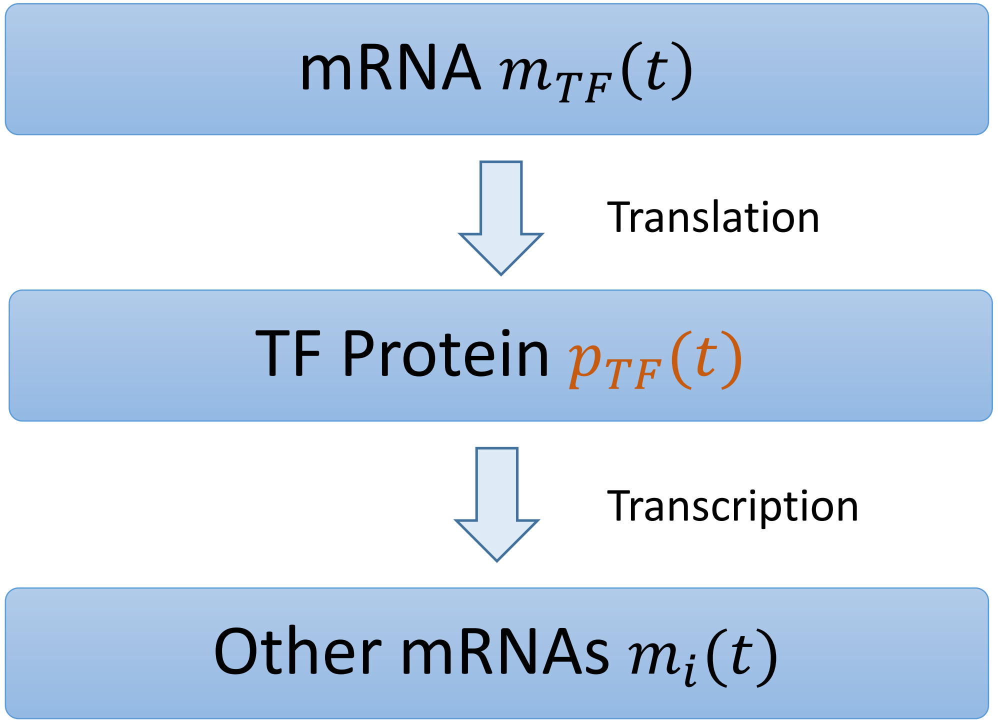
Measured using Microarray since 1998



Mechanistic Model

$$\frac{\mathrm{d}p_{TF}(t)}{\mathrm{d}t} = s_f m_{TF}(t) - d_f p_{TF}(t)$$

$$\frac{\mathrm{d}m_i(t)}{\mathrm{d}t} = s_i p_{TF}(t) - d_i m_i(t)$$



Need to Model $p_{TF}(t)$

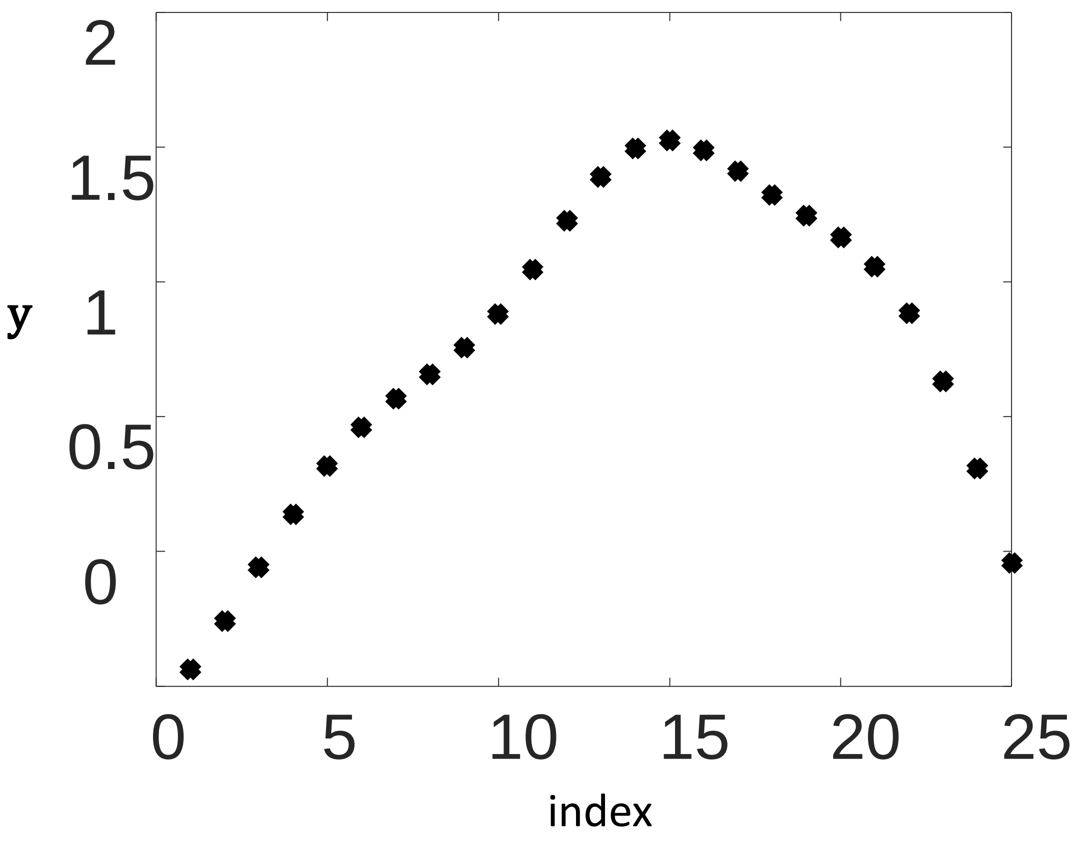
- Gaussian process: a *probabilistic* model for functions.
- Formally known as a *stochastic process*.
- Multivariate Gaussian is normally defined by a *mean vector*, $\boldsymbol{\mu}$, and a *covariance matrix*, $\mathbf{C}$.

$$y \sim N(\boldsymbol{\mu}, \mathbf{C})$$

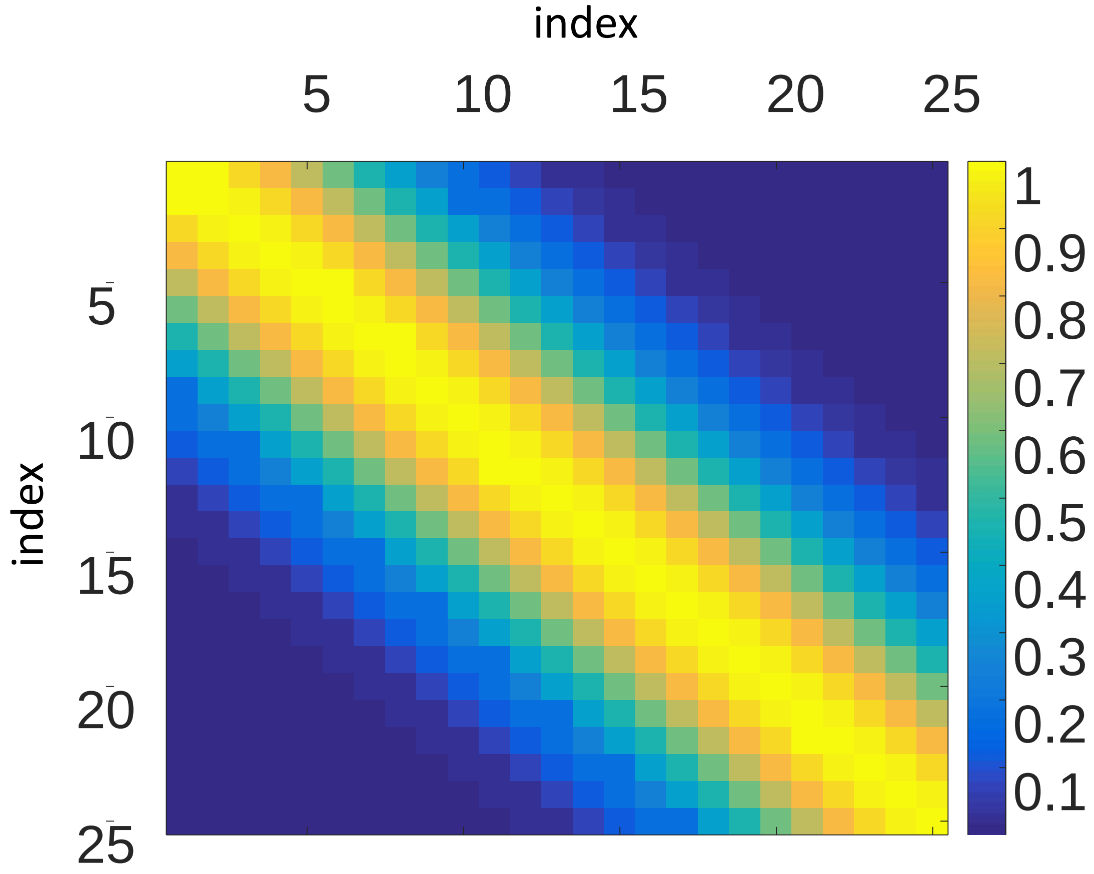
- Gaussian process defined by a *mean function*, $\mu(t)$, and a *covariance function*, $c(t, t')$.

$$y(t) \sim N(\mu(t), c(t, t'))$$

Zero Mean Gaussian Sample

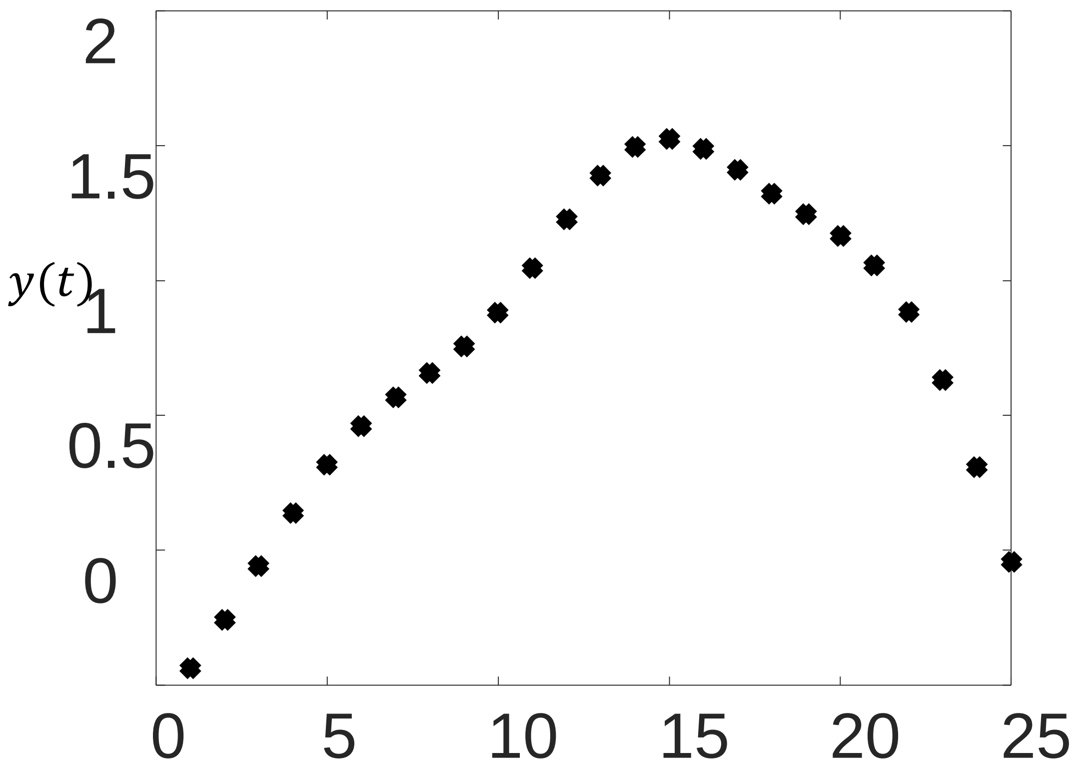


samples from Gaussian

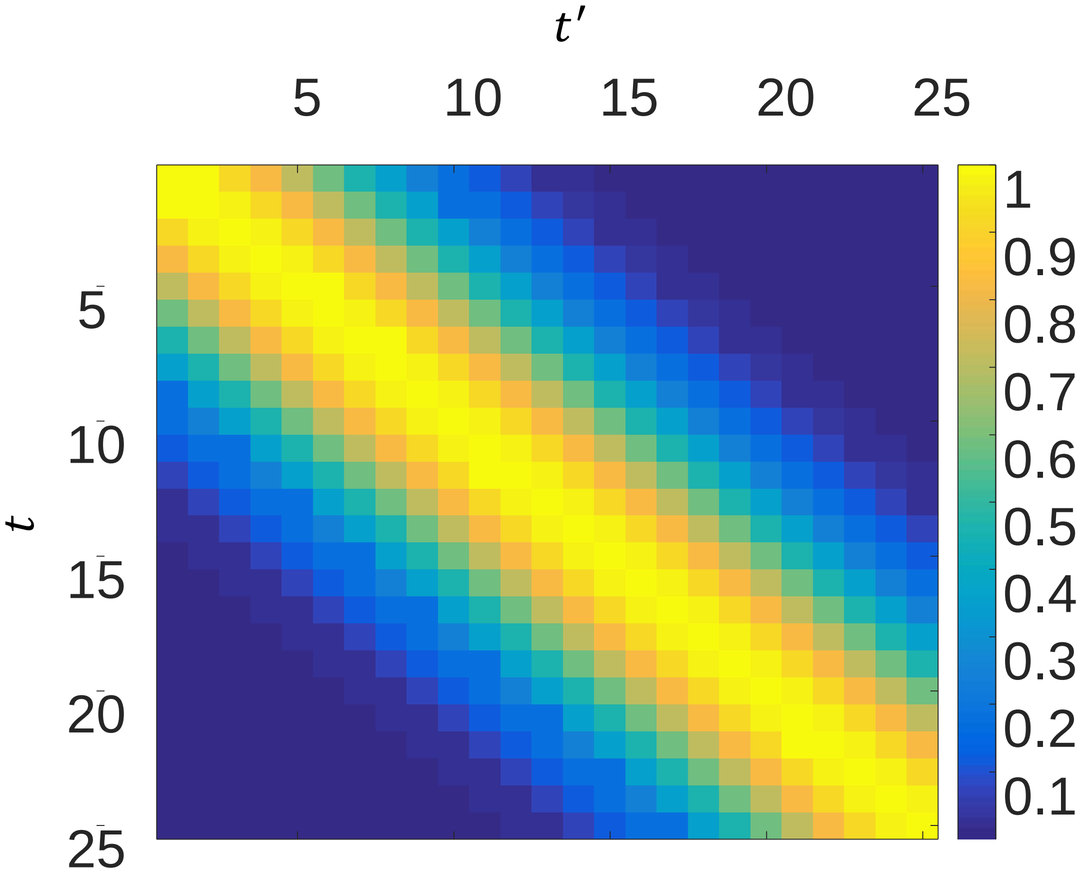


covariance $\mathbf{C}$

Zero Mean Gaussian Process Sample

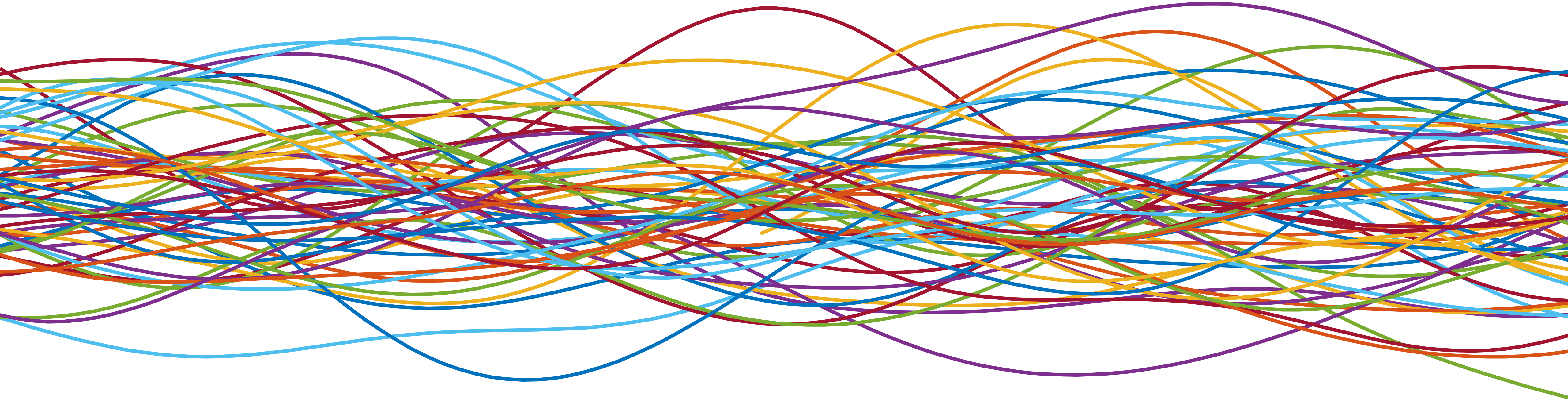


samples from Gaussian process

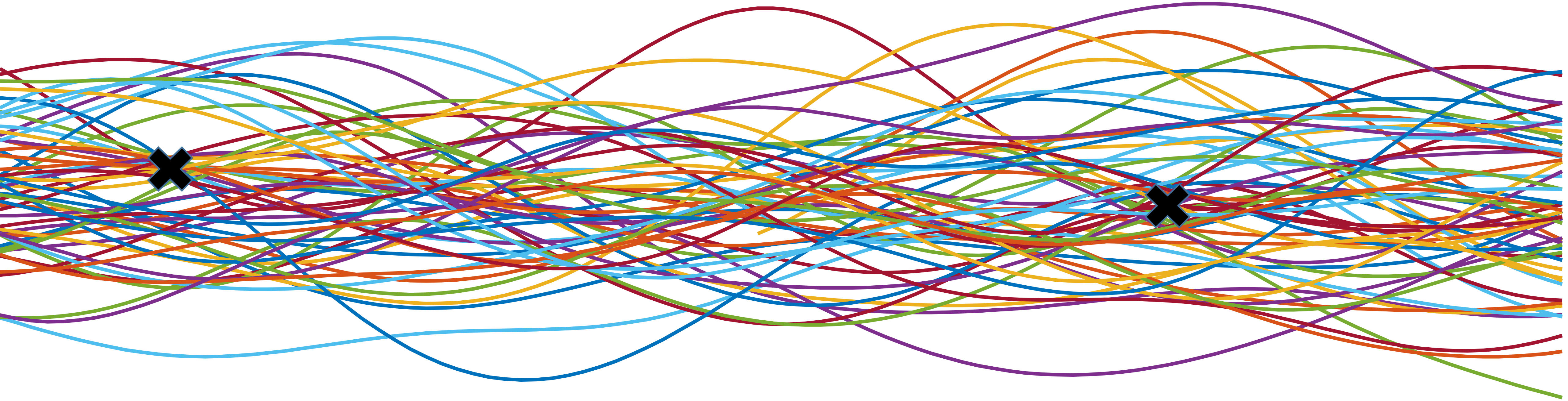


covariance function $c(t, t')$

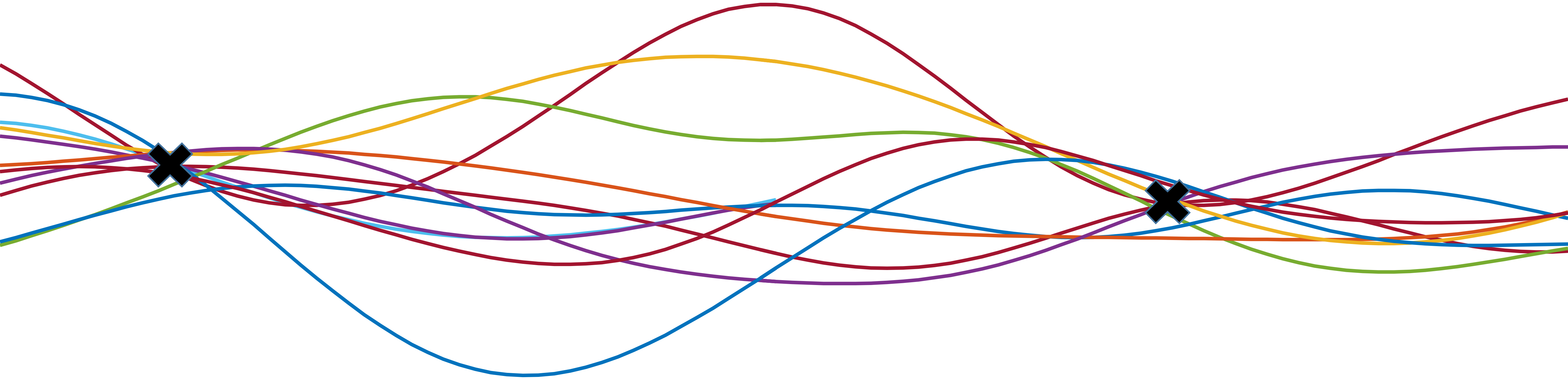
Gaussian Processes



Gaussian Processes



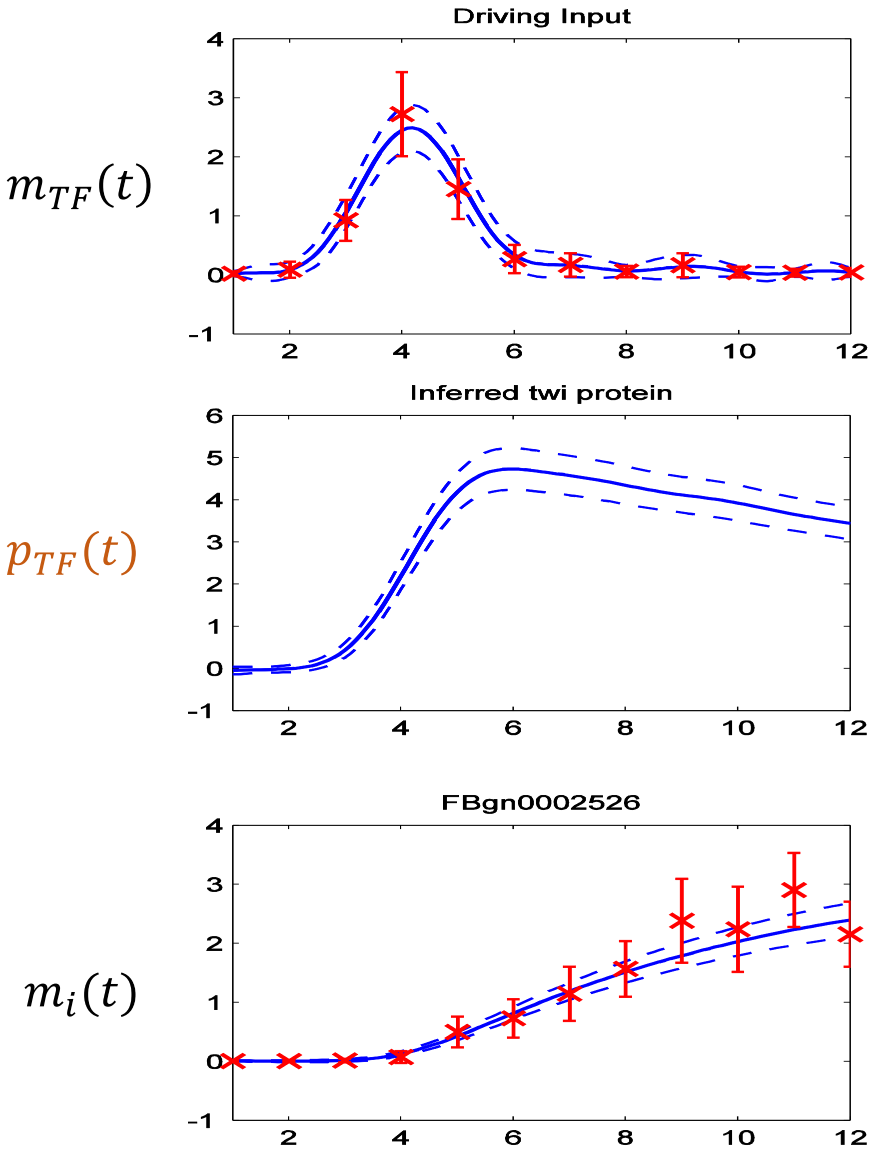
Gaussian Processes



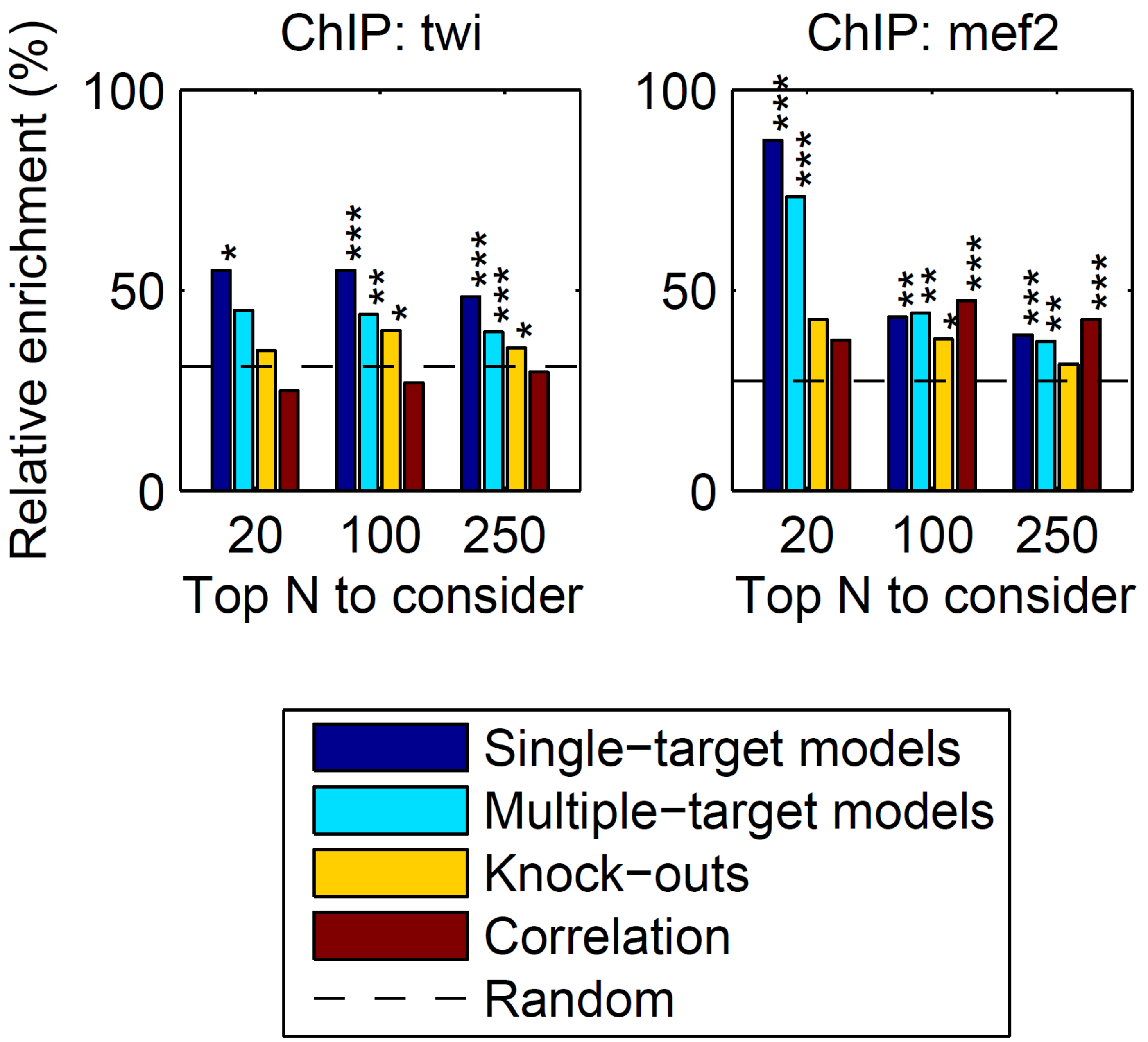
Results

$$\frac{\mathrm{d}p_{TF}(t)}{\mathrm{d}t} = s_f m_{TF}(t) - d_f p_{TF}(t)$$

$$\frac{\mathrm{d}m_i(t)}{\mathrm{d}t} = s_i p_{TF}(t) - d_i m_i(t)$$



TPAMI, 2 PNAS papers, 2 Comp Bio



ChIP: twi
ChIP: mef2
Relative enrichment (%)
100
50
0
20
100
250
Top N to consider
Single−target models
Multiple−target models
Knock−outs
Correlation
Random

MATLAB Demo

- demo_2016_01_29_OxWaSP.m

Further Challenge

- This model inter-relates different functions with mechanistic understanding.
- What if you need to inter-relate across different modalities of data at different scales.
- *E.g.* biopsy images + genetic test + mammogram for breast cancer diagnostics.

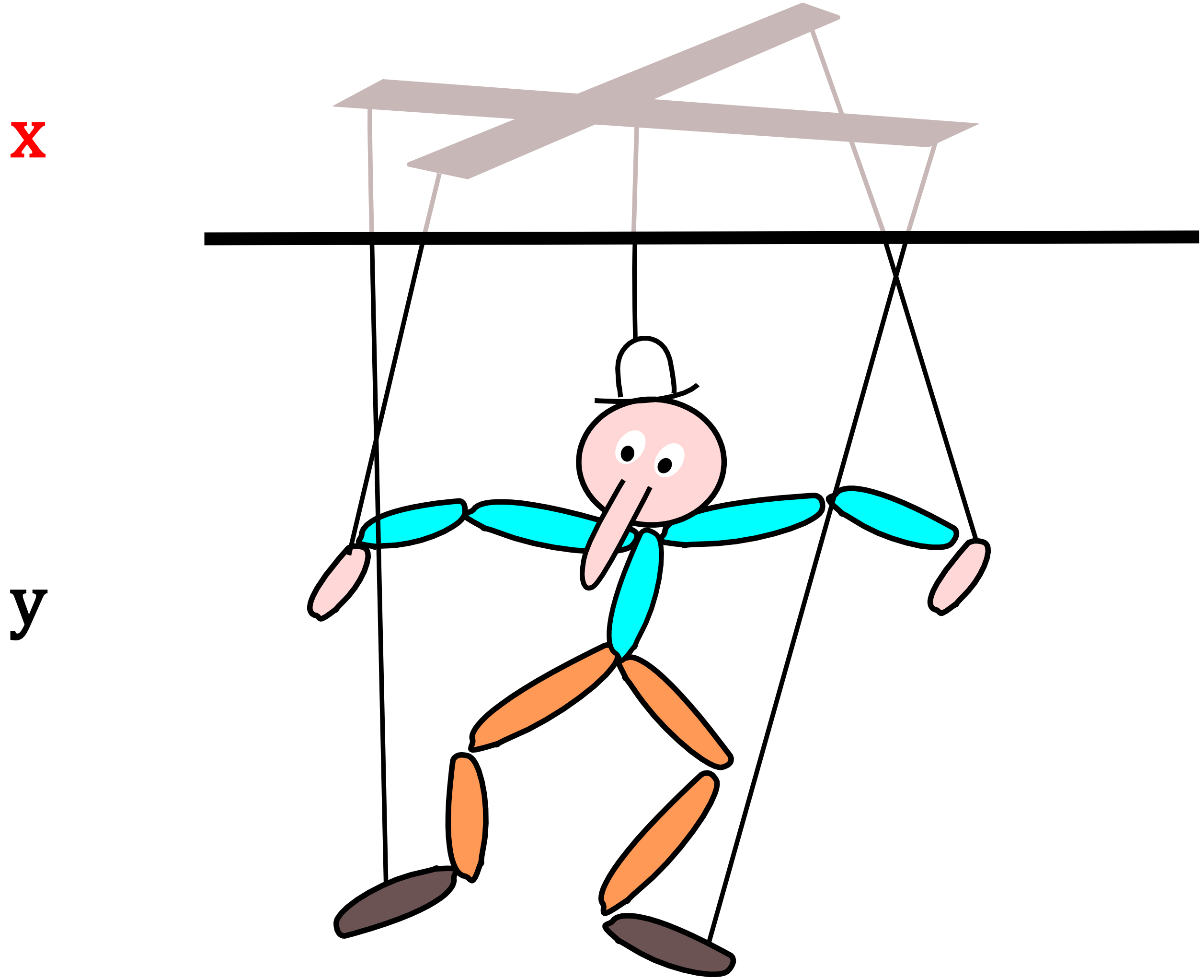
The Data are Not Enough

- Four pillars:
  - Deterministic/Stochastic
  - Mechanistic/Empirical

- **Goal:** *model complex phenomena over time*
- **Problem:**
  - *Mechanistic* models are often inaccurate
  - Data is often not rich enough for an *empirical* approach

- Question 2: How do we formulate the right representations to integrate different data modalities?

Classical Latent Variables



Classical Treatment

- Assume *a priori* that

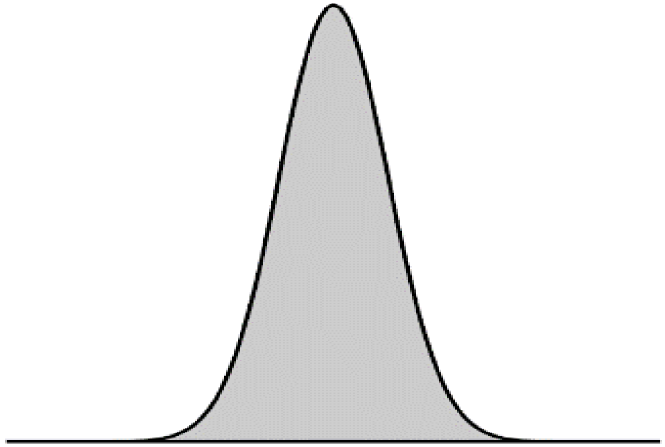
$$\mathbf{x} \sim N(\mathbf{0}, \mathbf{I})$$

- Relate linearly to $\mathbf{y}$

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \epsilon$$

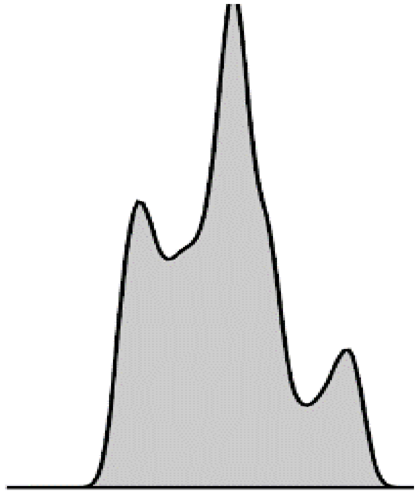
- Framework covers many classical models PCA, Factor Analysis, ICA

Render Gaussian Non Gaussian



x

$$y = f(x)$$



y

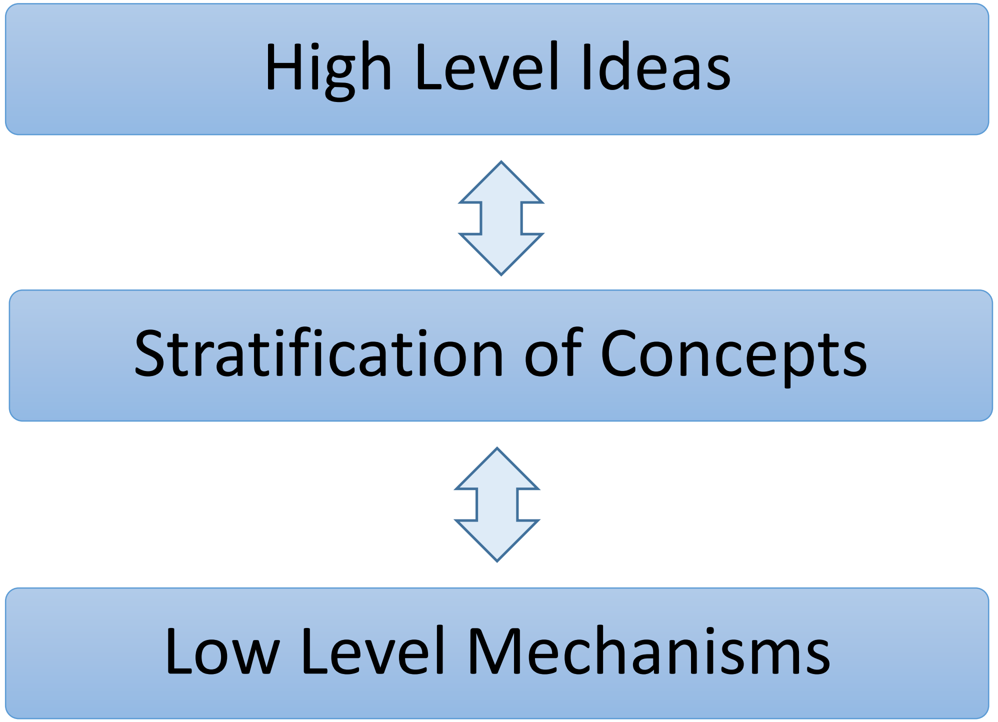
Stochastic Process Composition

- A new approach to forming stochastic processes
- Mathematical composition:

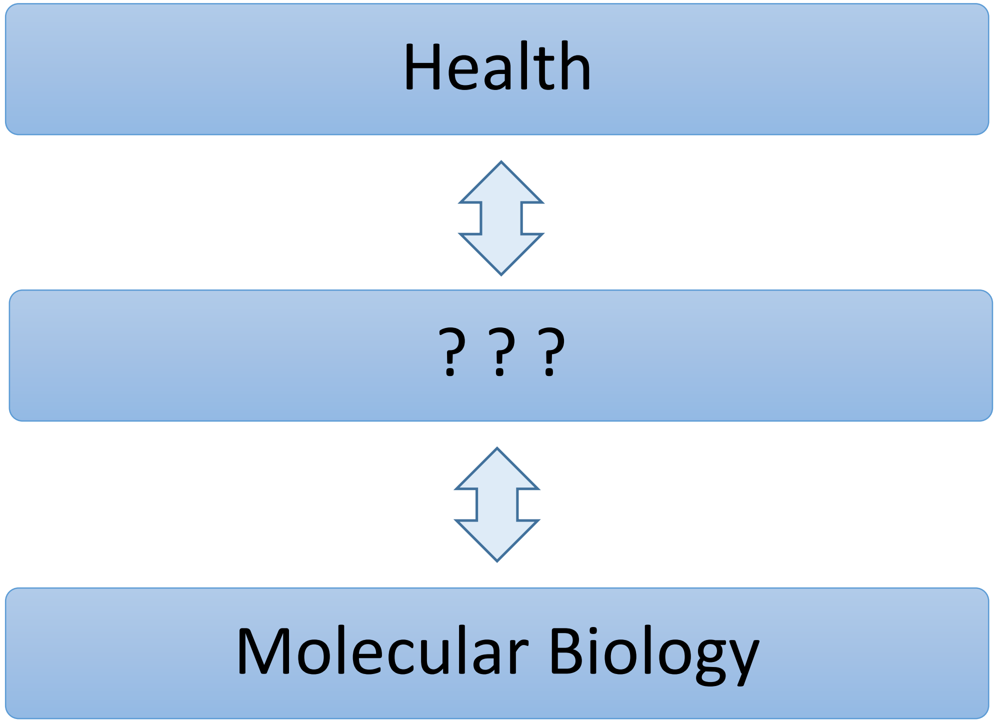
$$y(x) = f_1\left(f_2\left(f_3(x)\right)\right)$$

- Properties of resulting process highly non-Gaussian
- Allows for hierarchical structured form of model.
- Learning in models of this type has become known as: **deep learning**.

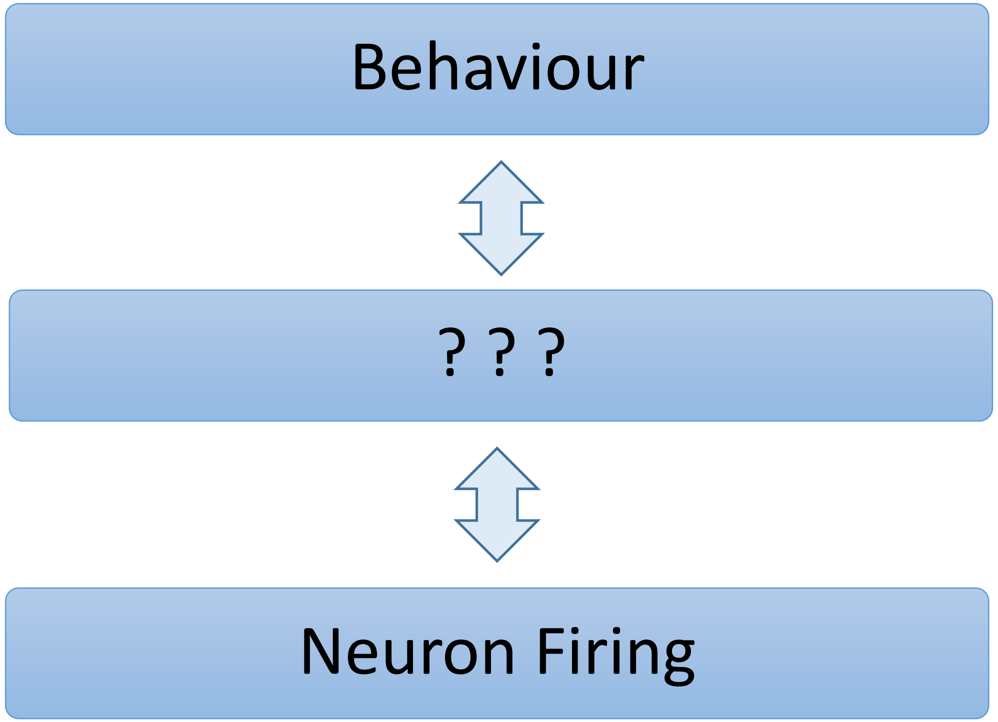
Use Abstraction for Complex Systems



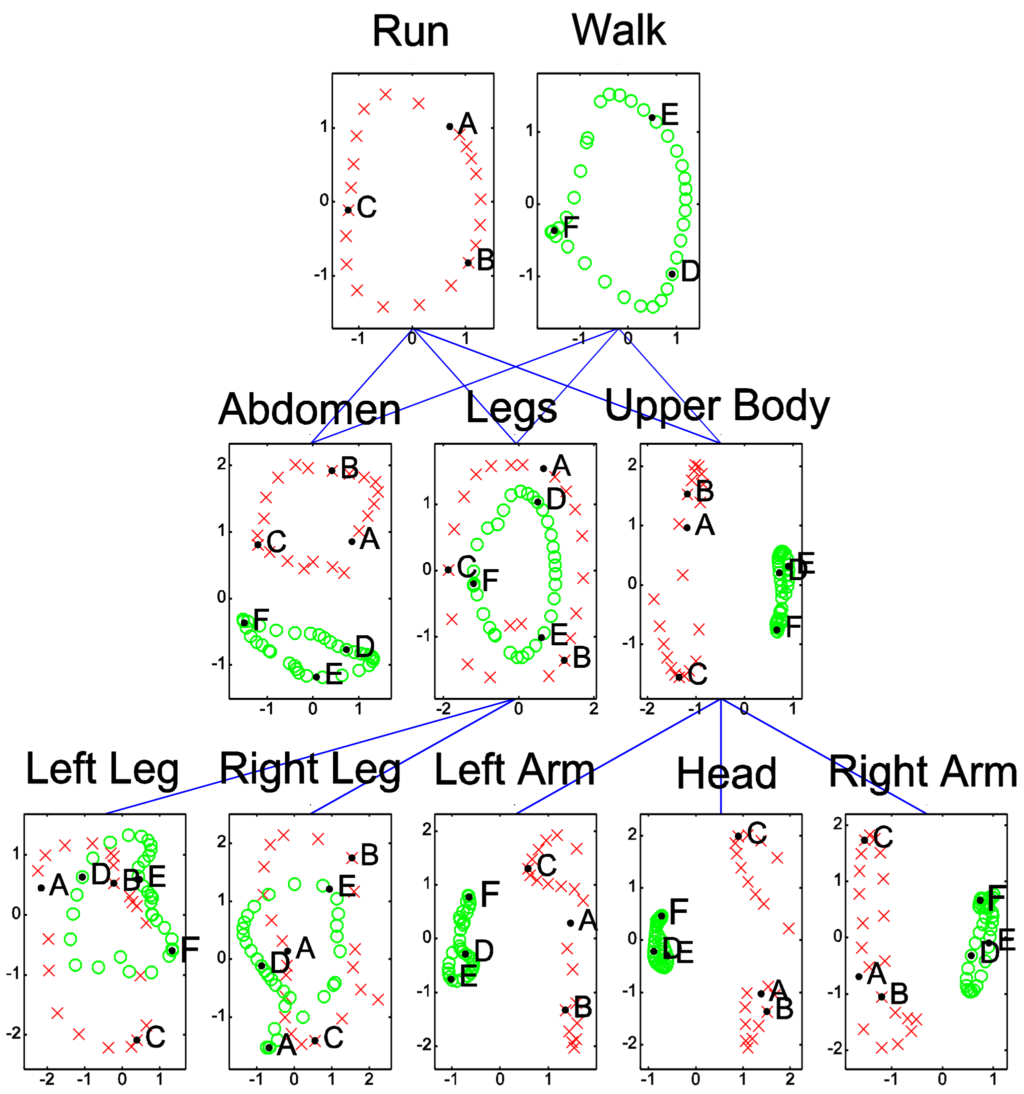
Biology and Health

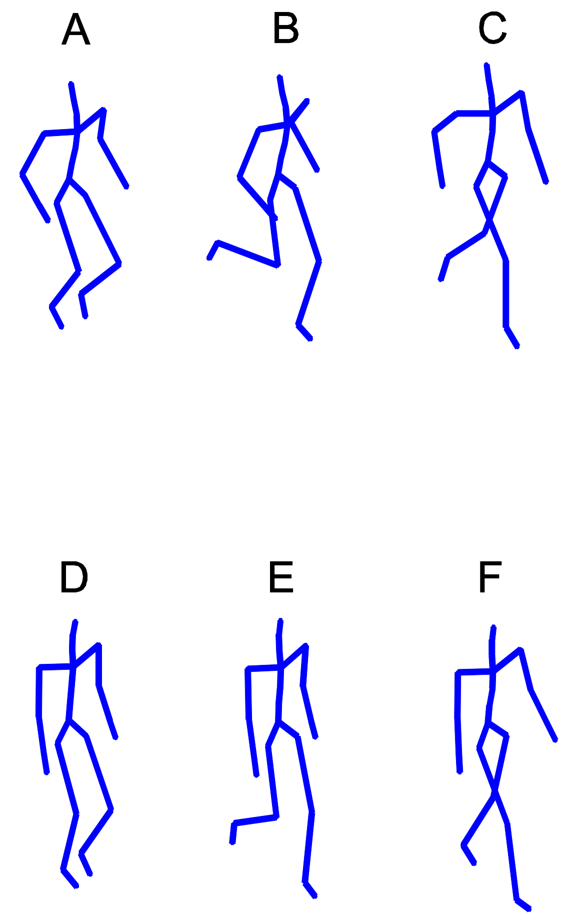


Neuroscience



Example: Motion Capture Modelling

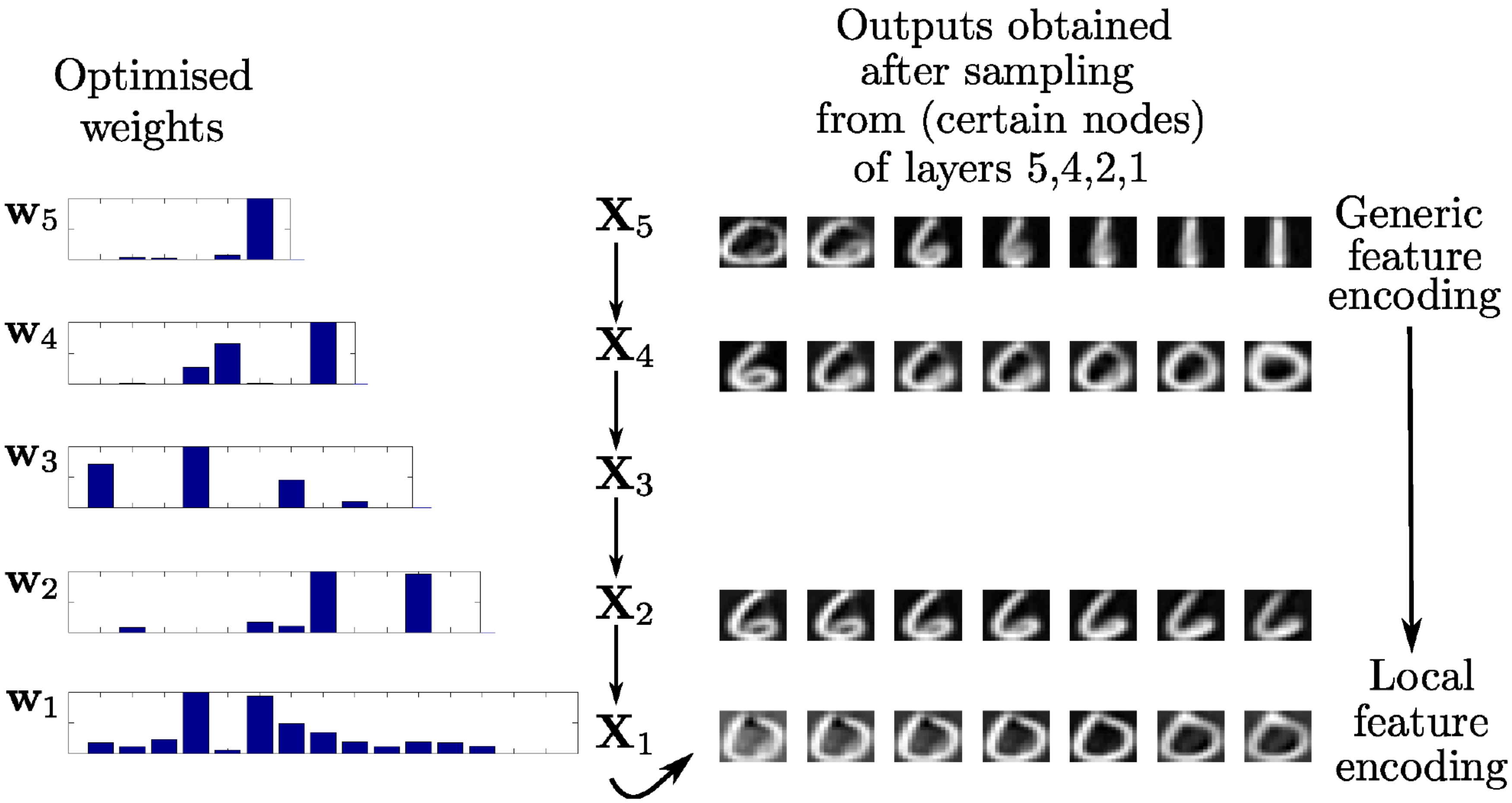




MATLAB Demo

- demo_2016_01_29_OxWaSP.m

Modelling Digits

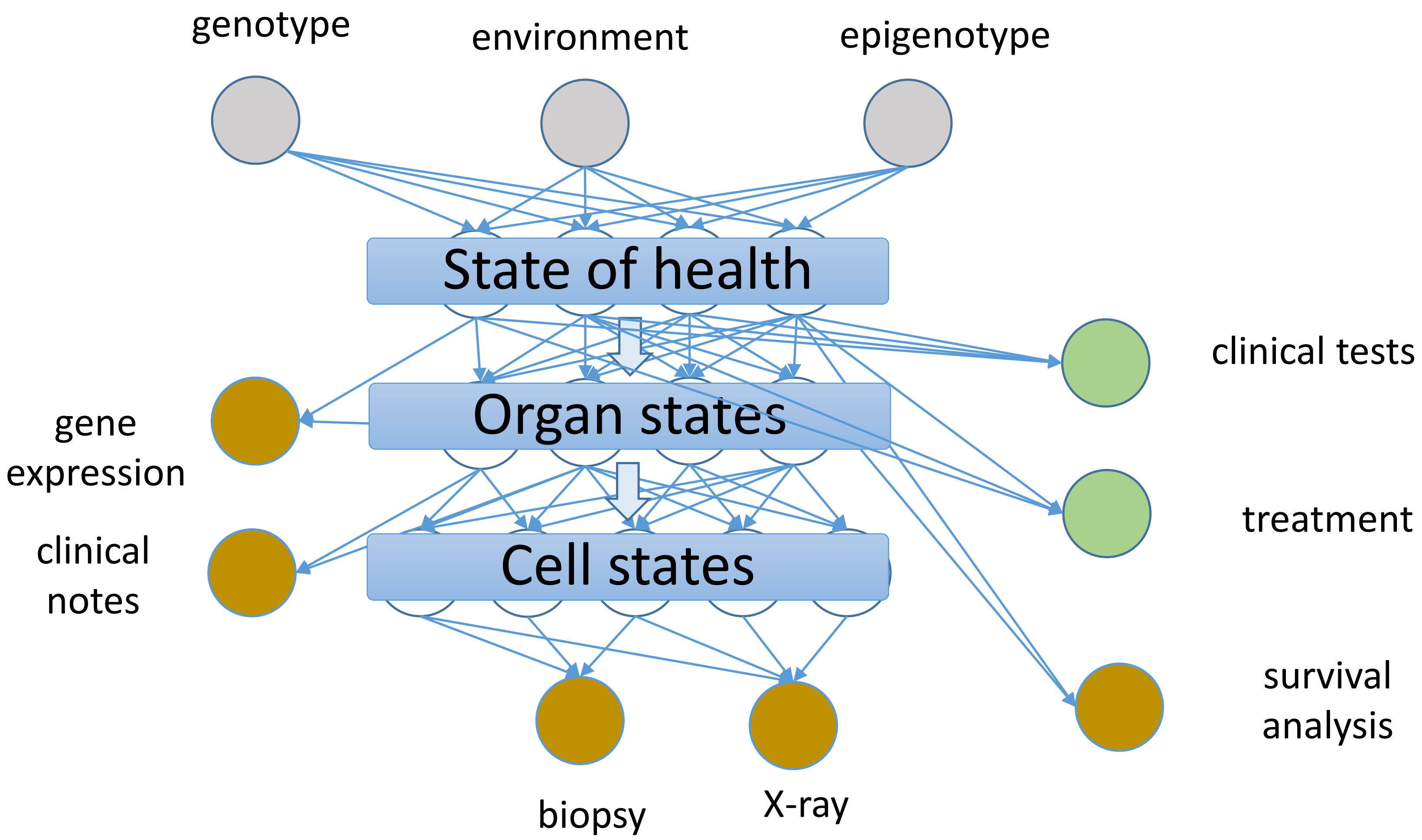


MATLAB Demo

- demo_2016_01_29_OxWaSP.m

Health

- Complex system
- Scarce data
- Different modalities
- Poor understanding of mechanism
- Large scale



PLoS Comp Bio, Nature Methods

To Find Out More

- Gaussian Process Summer School
    - 12th-15th September 2016 in Sheffield
    - This year in parallel with/themed as a UQ orientated school (co-organisation with Rich Wilkinson)
    - Occurring alongside ENBIS Meeting
    - http://gpss.cc/

Future

- Methodology
  - Deep GPs (also current)
  - Latent Force Models (current but dormant)
  - Latent Action Models and Stochastic Optimal Control (new)
  - Probabilistic Geometries (starting)

- Exemplar Applications
  - Health and Biology (existing)
  - Developing world (existing)
  - Robotics at different scales (starting)
  - Perception: vision (dormant) haptic (new)

Summary

- Complex systems:
  - ‘big data’ is too ‘small’.
    - The data are not enough.
  - Need data efficient methods
    - http://www.theguardian.com/media-network/2016/jan/28/google-ai-go-grandmaster-real-winner-deepmind

- Solutions:
  - Hybrid mechanistic-empirical models
  - Structured models for automated data assimilation

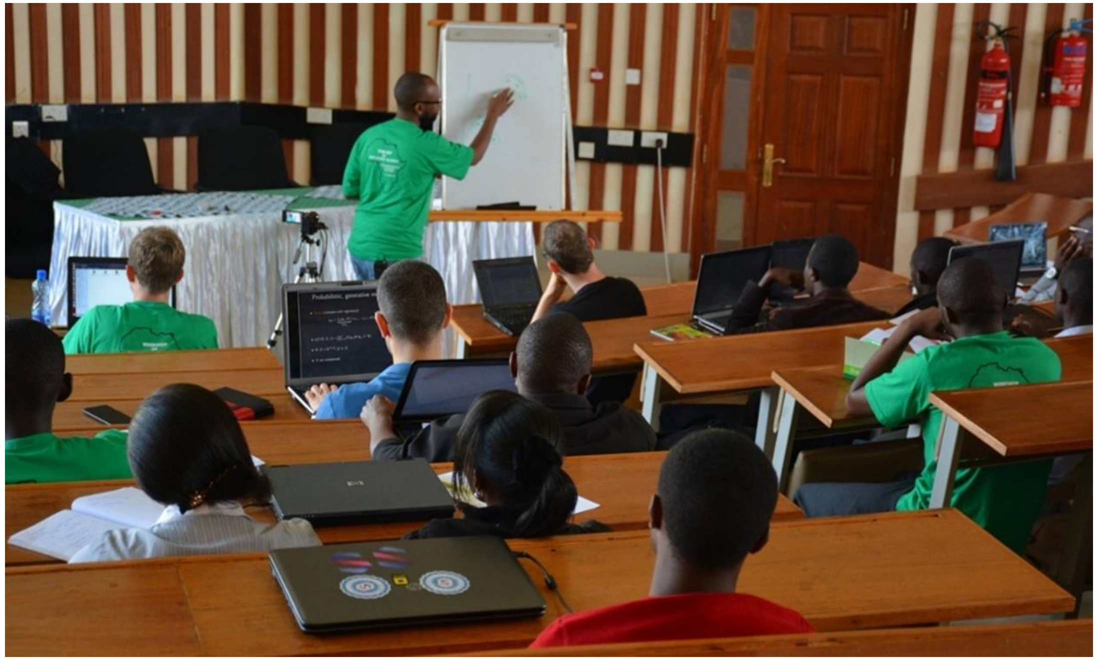
The Digital Oligarchy

- Response to concentration of power with data
- CitizenMe
  - London based start up
  - User-centric data modelling
- New challenges in ML
  - Integration of ML, systems, cryptography.

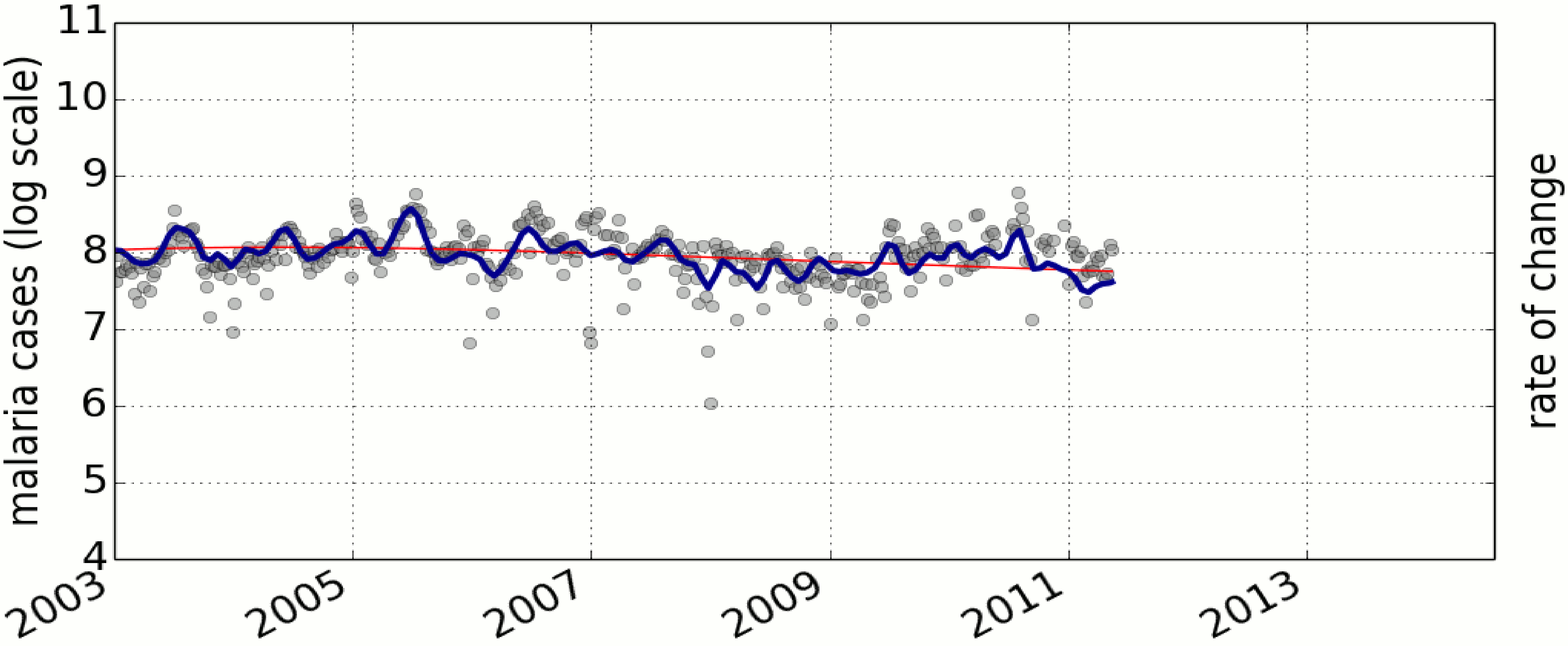
Open Data Science and Africa

Challenge

- "Whole pipeline challenge"
- Make software available
- Teach summer schools
- Support local meetings
  - Publicity in the Guardian
- Opportunities to deploy pipeline solution

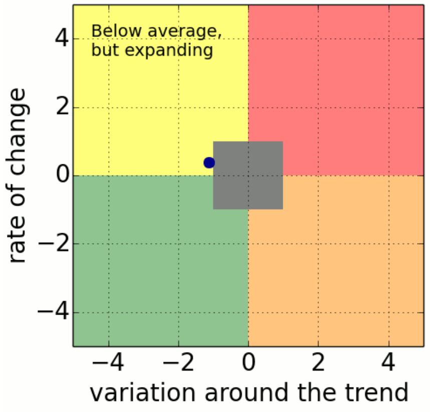


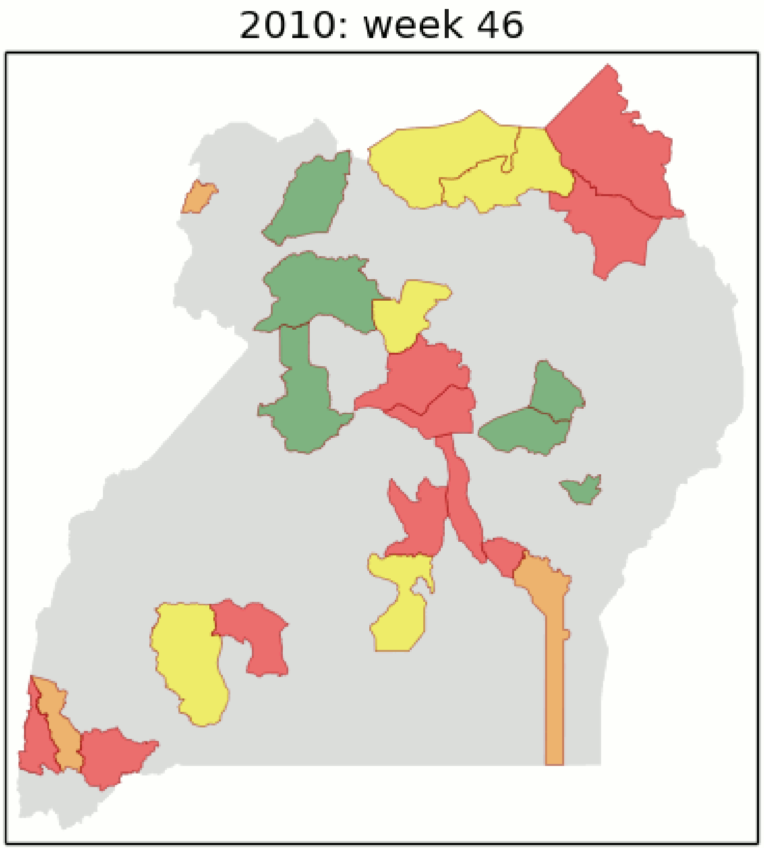
Disease Incidence for Malaria

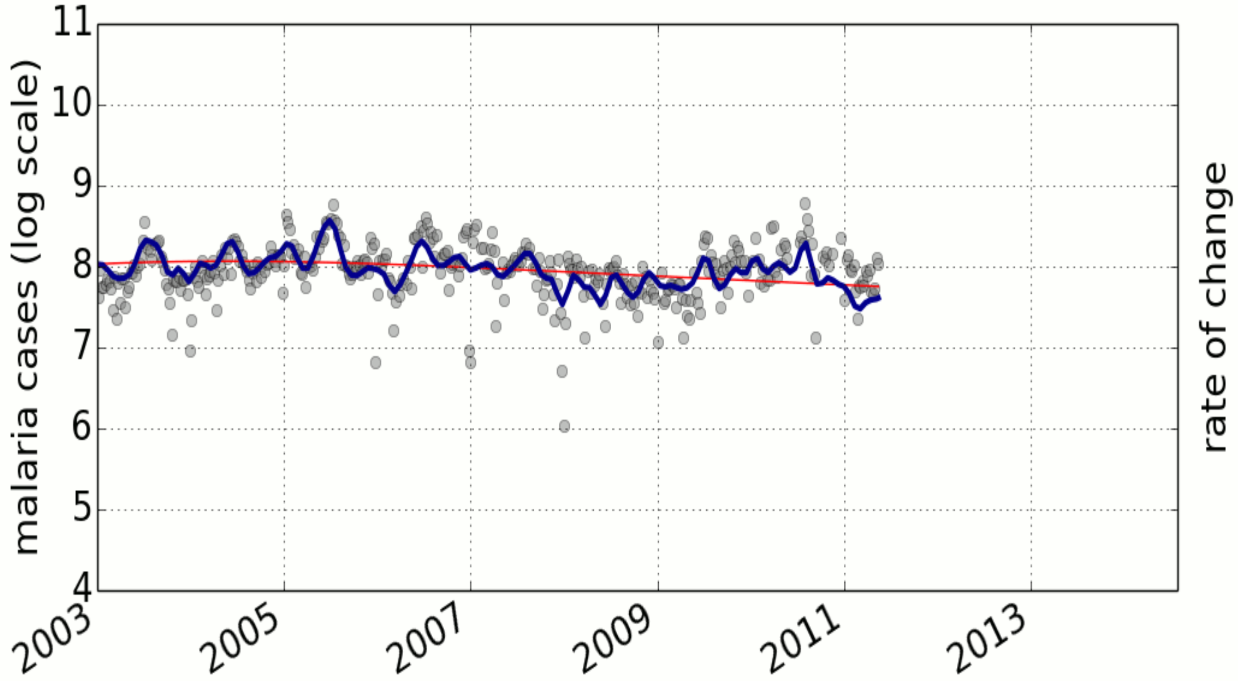


Uganda

- Spatial models of disease







Deployed with UN Global Pulse Lab

http://pulselabkampala.ug/hmis/