

*Bayesian emulation for
complex model calibration -
the first waves*

Helen Johnson

Centre for the Mathematical Modelling of Infectious Diseases
London School of Hygiene & Tropical Medicine



Overview

- Motivation & Objective
- Introduction to Bayesian Emulation
 - What is an emulator?
 - How can they help us?
- An emulator-based calibration algorithm
 - Developing a sequential emulator-based calibration algorithm
 - Multivariate bimodal example
 - Application to an individual-based model of HIV in Uganda
 - Preliminary results
- Future work

Motivation

- Use of complex individual-based stochastic models is becoming increasingly common in many disciplines including infectious disease modelling
- Calibration methods for such models have lagged behind their development
- Likelihood is often intractable and even distance-based measures require many model runs, which may be prohibitively costly in computational time

Objective

- To combine elements of ABC SMC-like algorithms with the efficiency of Bayesian Emulation to develop a method that can realistically calibrate such complex models

What is an emulator?

- Consider that a computer model encodes a function, which takes inputs (θ) and produces outputs $x = f(\theta)$
- An emulator is a statistical approximation of that function

$$f(\theta) = \sum_i \beta_i g_i(\theta) + u(\theta)$$

linear regression term

uncertainty term (Gaussian process)

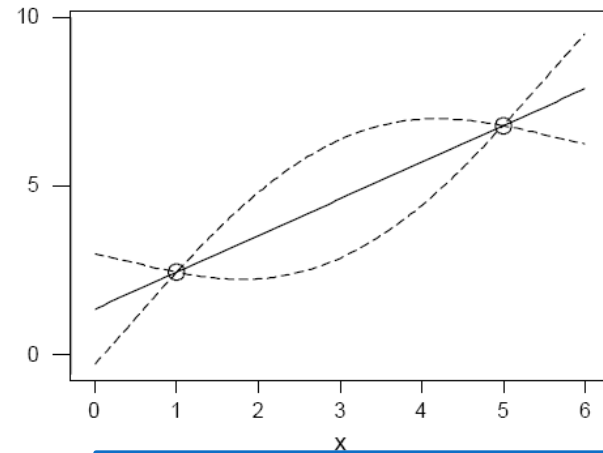
where $g_i(\theta)$ are known deterministic functions of θ
and β_i is a vector of scalar coefficients

(Kennedy and O'Hagan, 2001, Conti and O'Hagan, 2010)

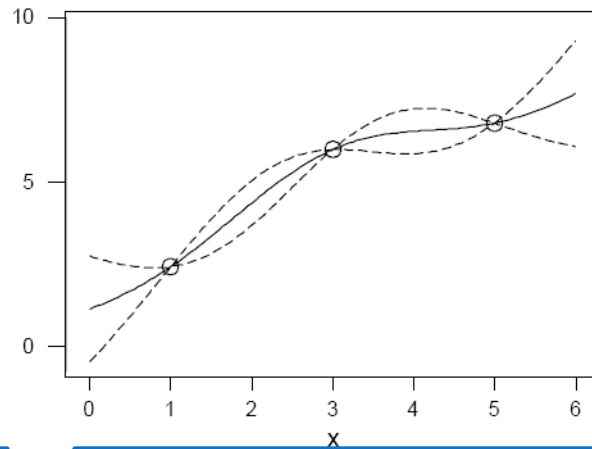
What is an emulator?

- Essentially we consider the Gaussian process form as a prior for the function $f(\theta)$
- This prior is then updated using as data a small ‘training sample’ of n selected complex model runs
- Formally, the emulator is the resultant posterior distribution of $f(\theta)$
- Conditioning upon the training set ensures that realisations from the emulator interpolate between complex model evaluations
- We can then use the emulator to estimate complex model output for given inputs, along with statistical measure of estimation error

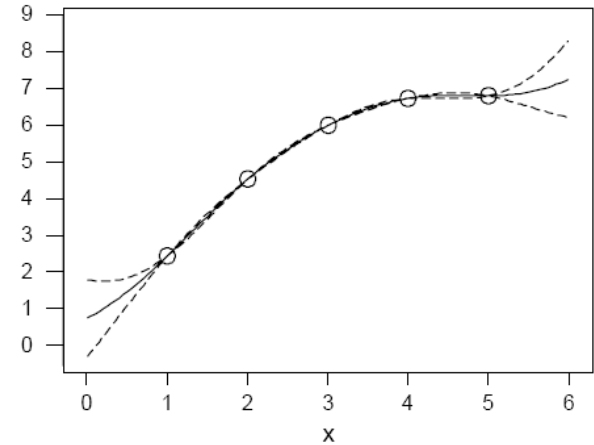
What is an emulator?



- Consider 1 input and 1 output
- Emulator estimate interpolates the two training points
- Uncertainty increases between training points



- Adding additional training points alters the estimate
- Uncertainty between training points is reduced



- and so on...

What is an emulator?

- A well-designed emulator:
 - estimates the complex model accurately
 - with small uncertainty
 - and runs “instantly” (1 prediction < 0.1 second)

Advantages:

- Short run time -> can cover large areas of parameter space efficiently
- Useful for model fitting and sensitivity and uncertainty analysis

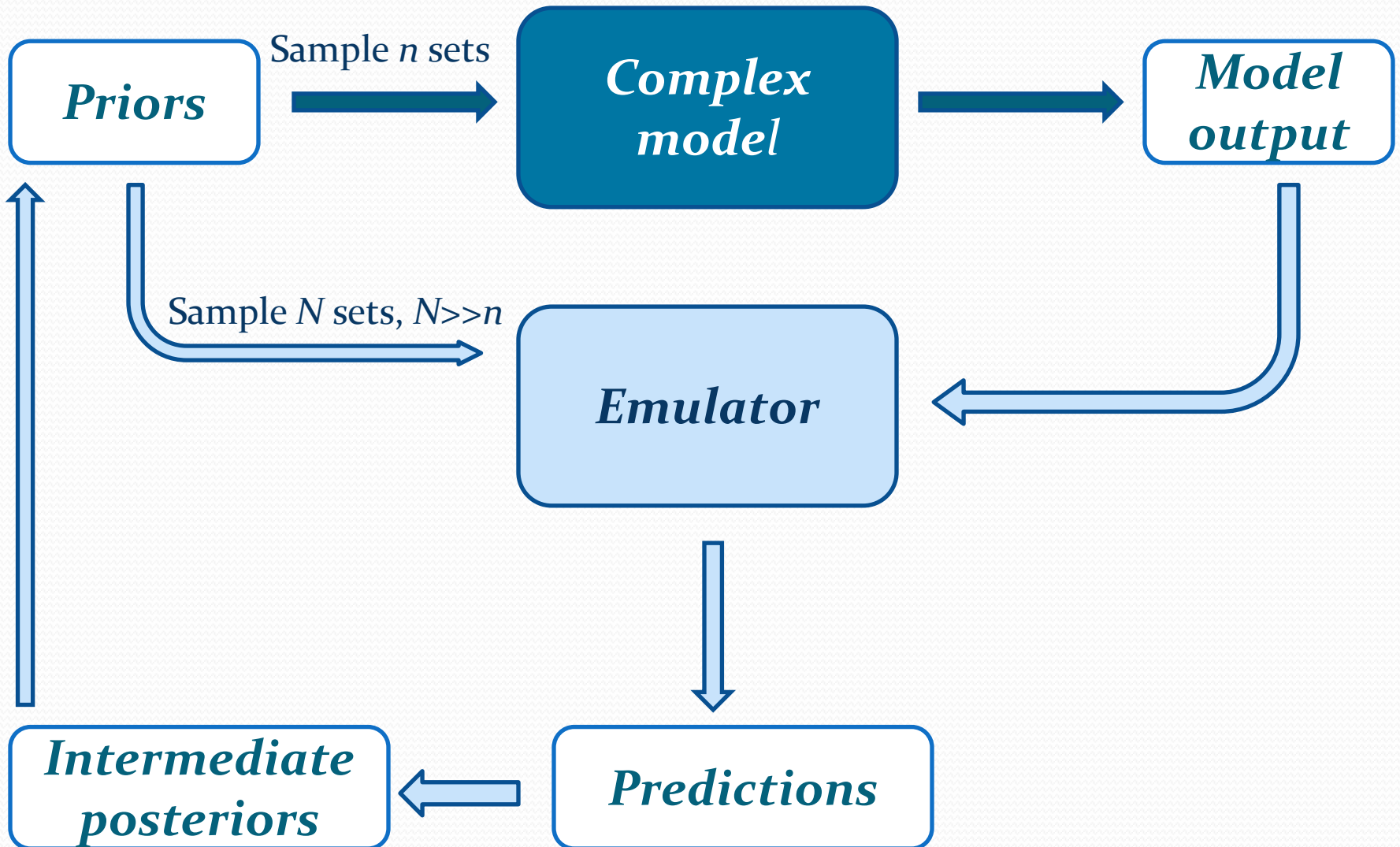
Limitations:

- Currently used almost exclusively for deterministic models

Developing an emulation-based calibration algorithm

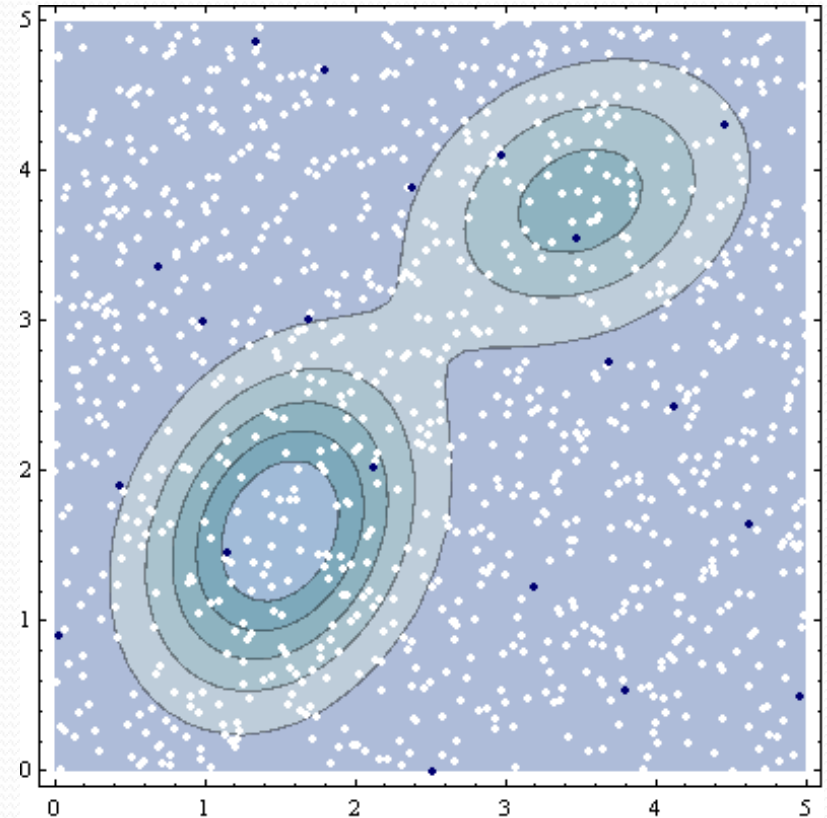
1. Generate an emulator to approximate complex model
2. Use emulator to predict model output for many points in parameter space
-> allows us to identify many local maxima
3. Exclude regions of parameter space where the emulator predictions are distant from observed data i.e. where a model fit is unlikely
4. Evaluate complex model for most promising parameter sets as identified by the emulator
5. Re-train emulator with this additional model output
-> more accurate emulator
6. Identify new parameter sets for emulator evaluation by re-sampling according to weight and perturbing.
7. Repeat steps 2-6 until sufficient fits are found

An emulation-based calibration algorithm

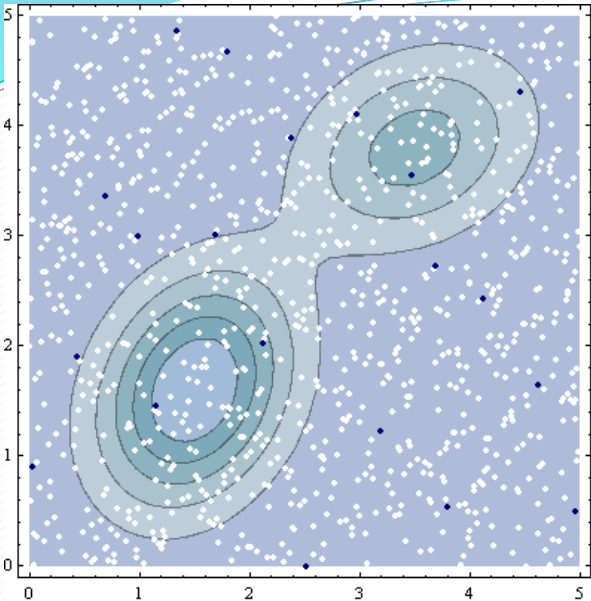


A simple toy example

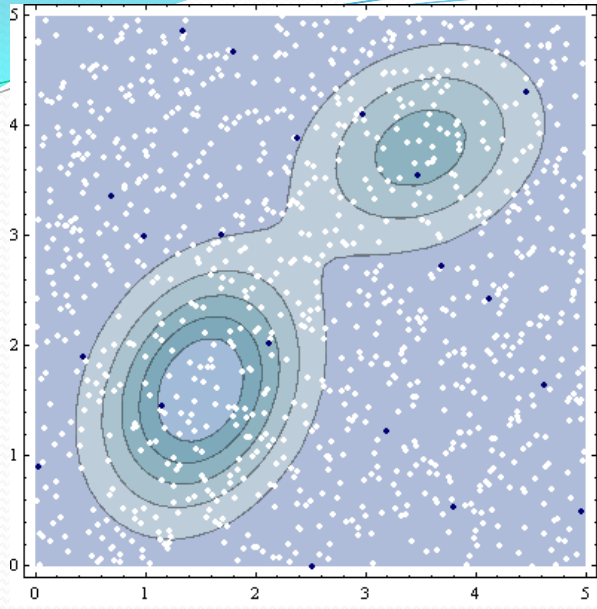
- Apply the method to a simple bivariate bimodal surface
- We know the height of the highest point on the surface and want to find the corresponding parameters
- Use LHS maximin sampling to sample 20 parameter sets for simulator evaluation and 1000 sets for emulator evaluation



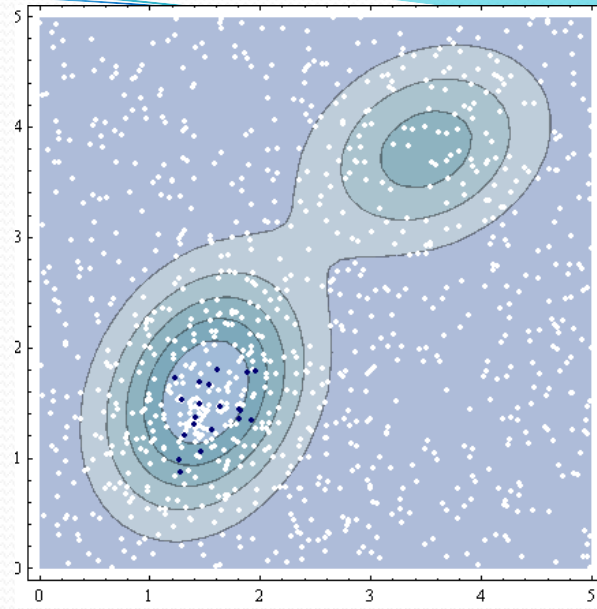
Wave 1



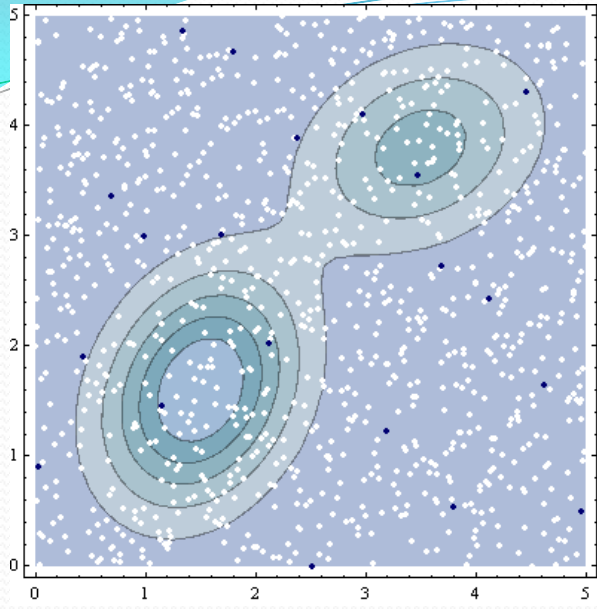
Wave 1



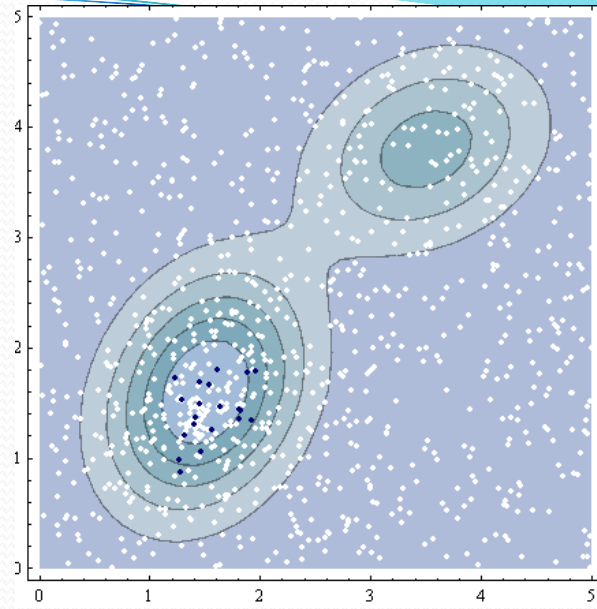
Wave 2



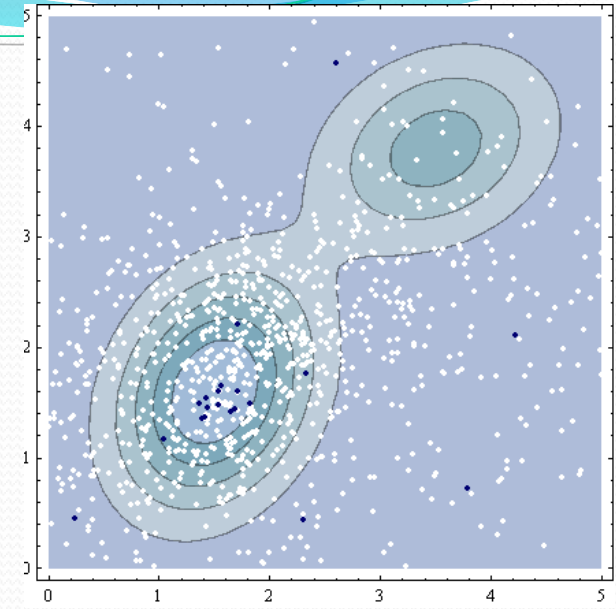
Wave 1



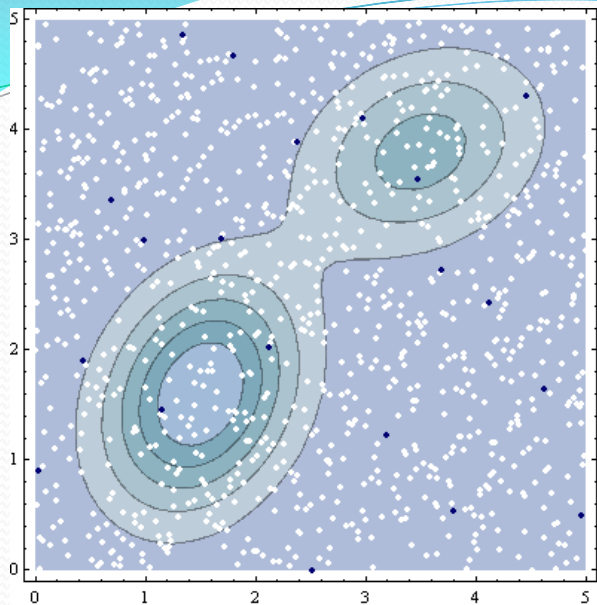
Wave 2



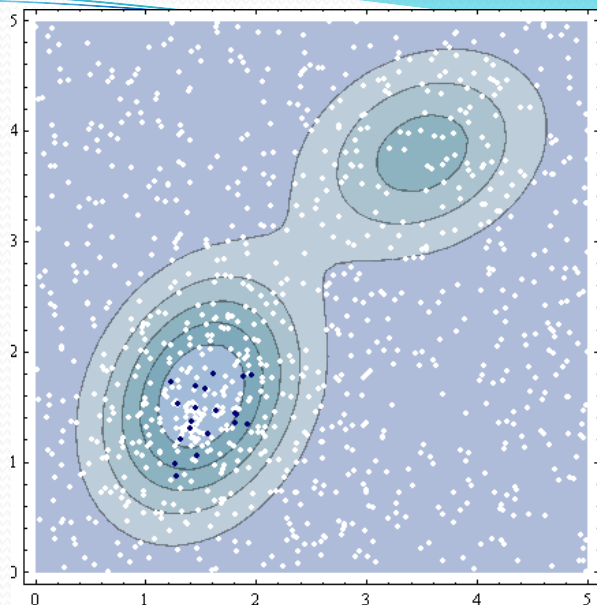
Wave 3



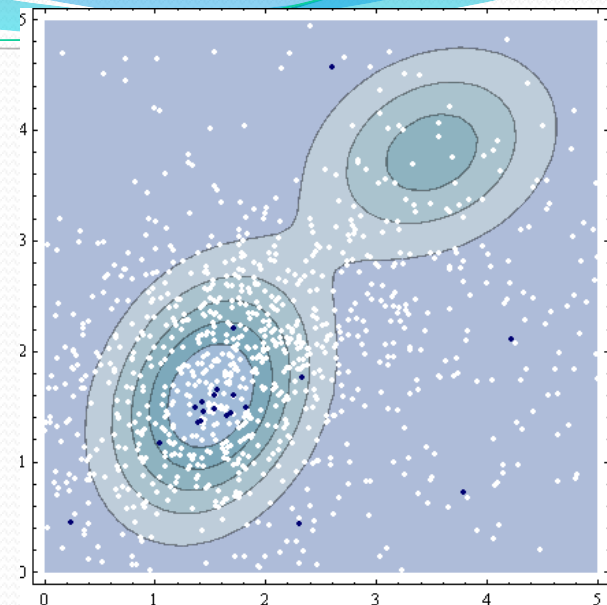
Wave 1



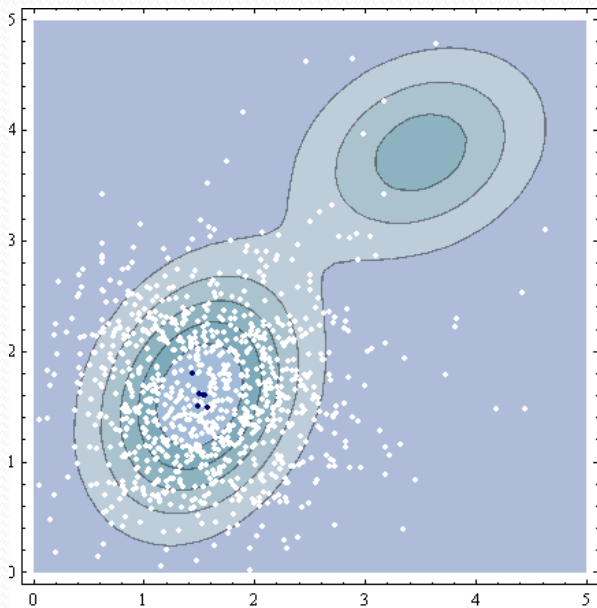
Wave 2



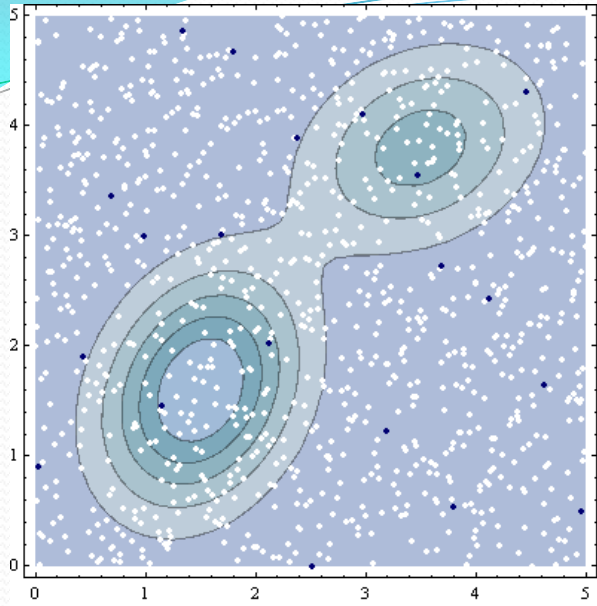
Wave 3



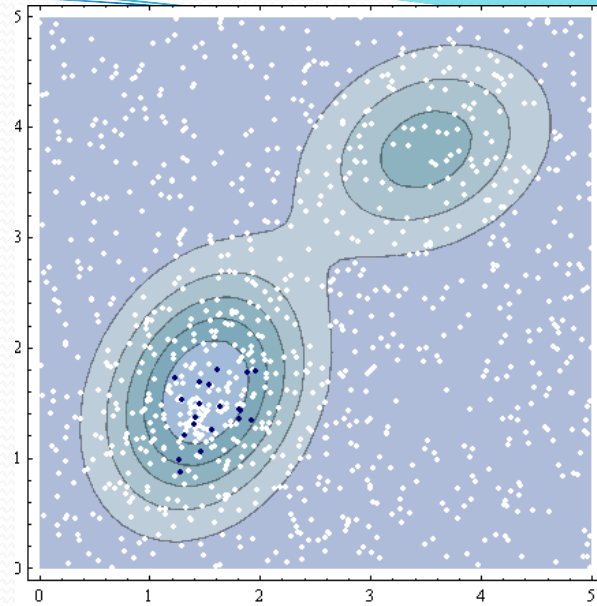
Wave 4



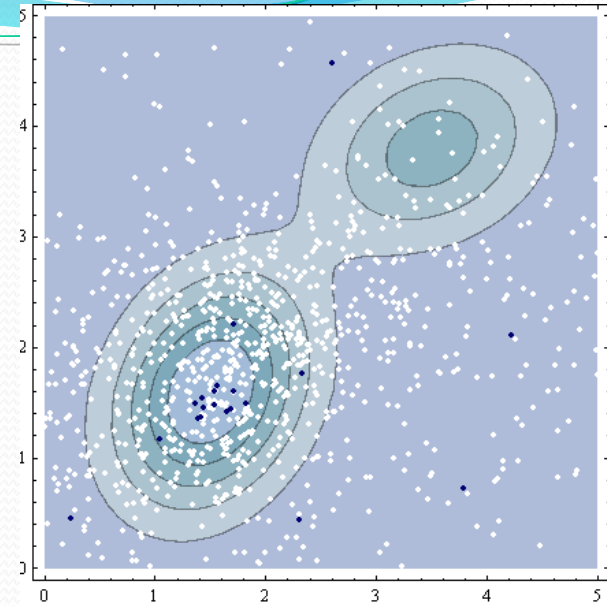
Wave 1



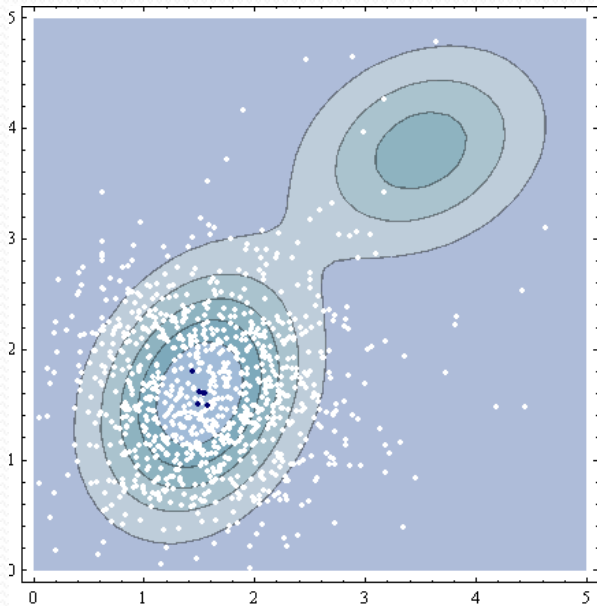
Wave 2



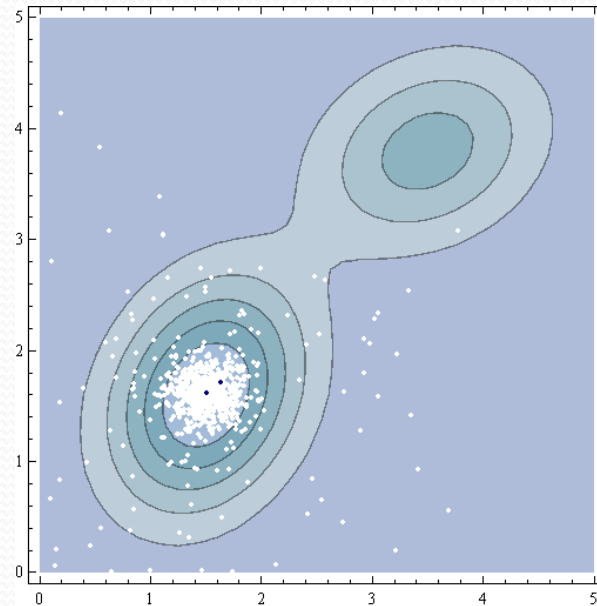
Wave 3



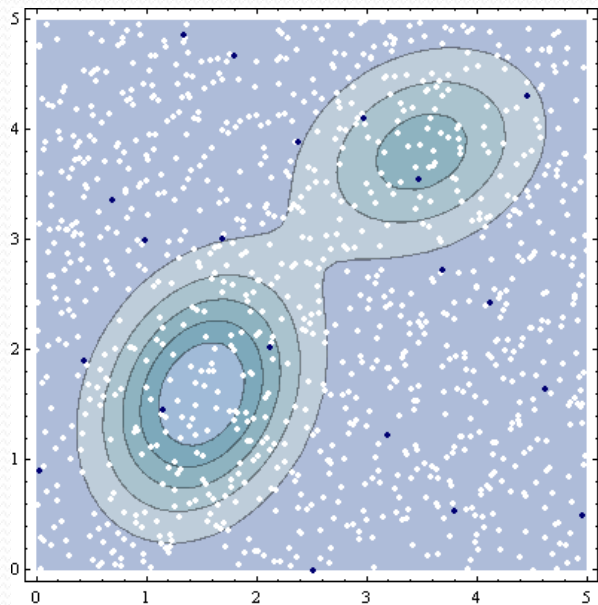
Wave 4



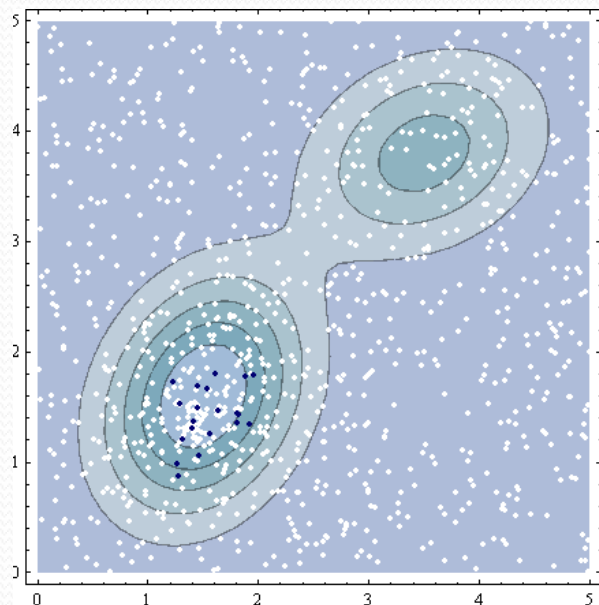
Wave 5



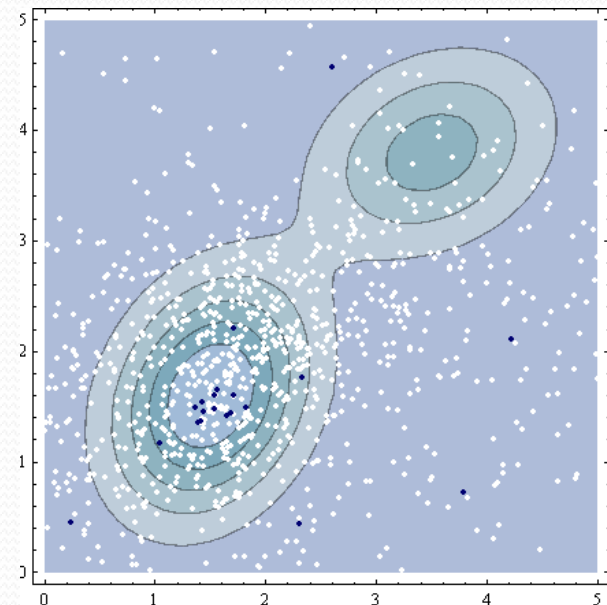
Wave 1



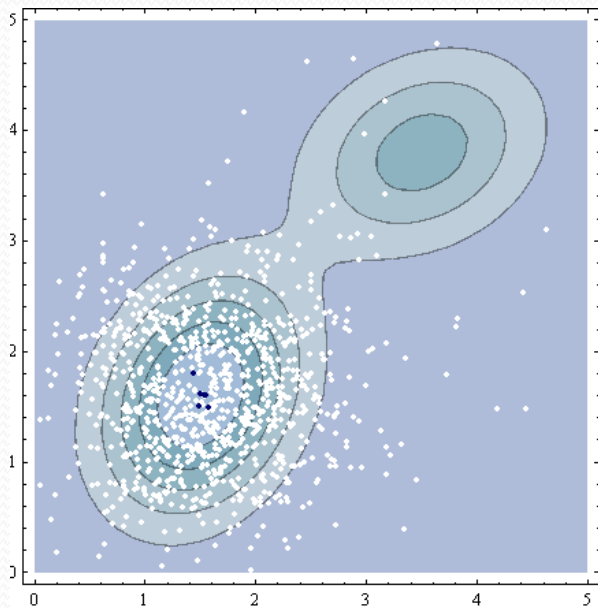
Wave 2



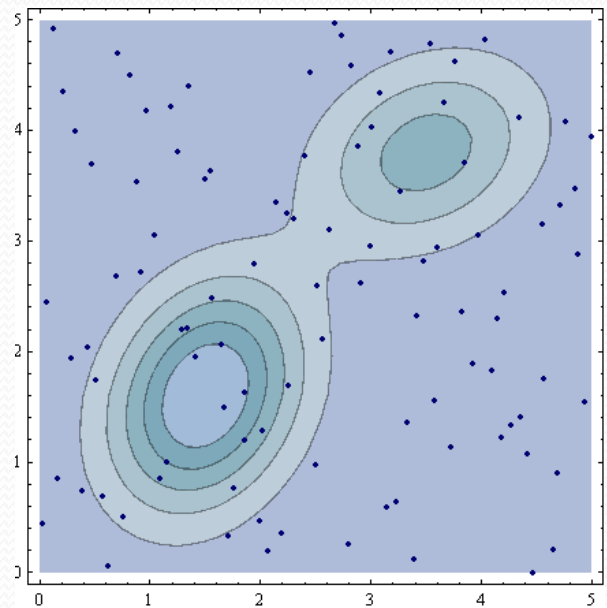
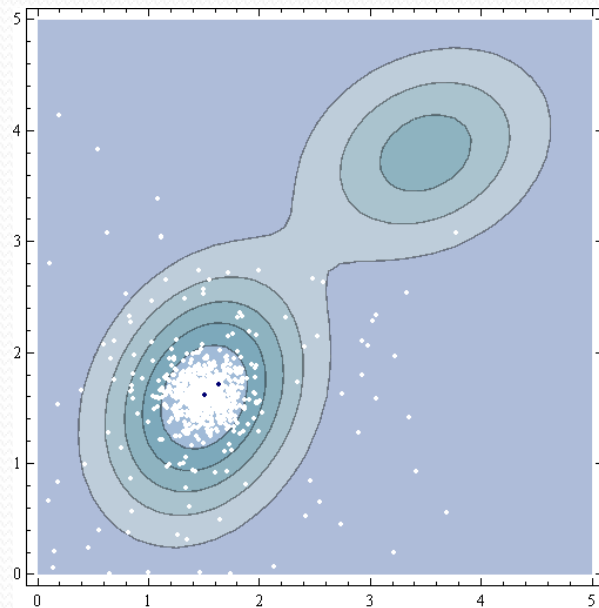
Wave 3



Wave 4

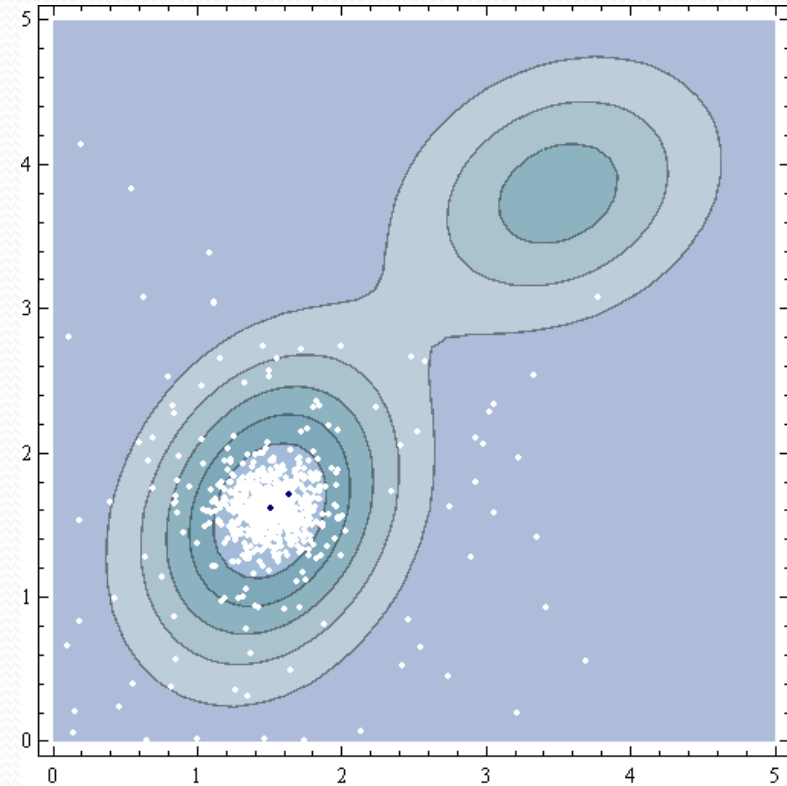


Wave 5



A simple toy example

- The method swiftly and accurately identified the parameter values that corresponded to the fitting point
- For a given number of simulator runs (computational time), cover far more parameter space
- However, 'simple' is an understatement!
Consider next fitting to a ridged surface or attempting to hit two peaks



Into the deep end!

- Application to a realistic, if dauntingly over-specified, model
- Consider a complex individual-based model of HIV and concurrency in Uganda incorporating:
 - Age-stratified structure
 - Different sexual activity groups
 - Differing preferences for concurrent partnerships
 - Different types of partnership (long-term vs. casual)
 - Behavioural interventions
- Model has been fitted 'by eye' to HIV prevalence and population data from Uganda. We are aiming to find the same fits and to identify other promising areas of parameter space.

Preliminary work: Emulating a complex individual-based stochastic model of HIV in Uganda

- Input parameters sampled(23!), including:
 - Year of introduction of HIV to population
 - Scaling factors for HIV transmission probability by stage (3)
 - Scaling factor for increased duration of HIV stage while on ART
 - Mean partnership duration for long-term partnerships
 - New partner acquisition rate (p.a.) for casual partnerships
 - Proportion of males / females in high sexual activity group
 - Proportion of males who are non-monogamous
 - Tendency of males/females who are non-monogamous to take additional partners (2)
 - Years of behavioural interventions (2)
 - Effectiveness of behavioural intervention (4)
- Output, fitted to Ugandan data:
 - HIV prevalence (9 points)
 - Population size (2 points)
 - Ratio males / females
 - Average population growth rate
 - Average male partner acquisition rate
 - Prevalence of concurrency (9 points)

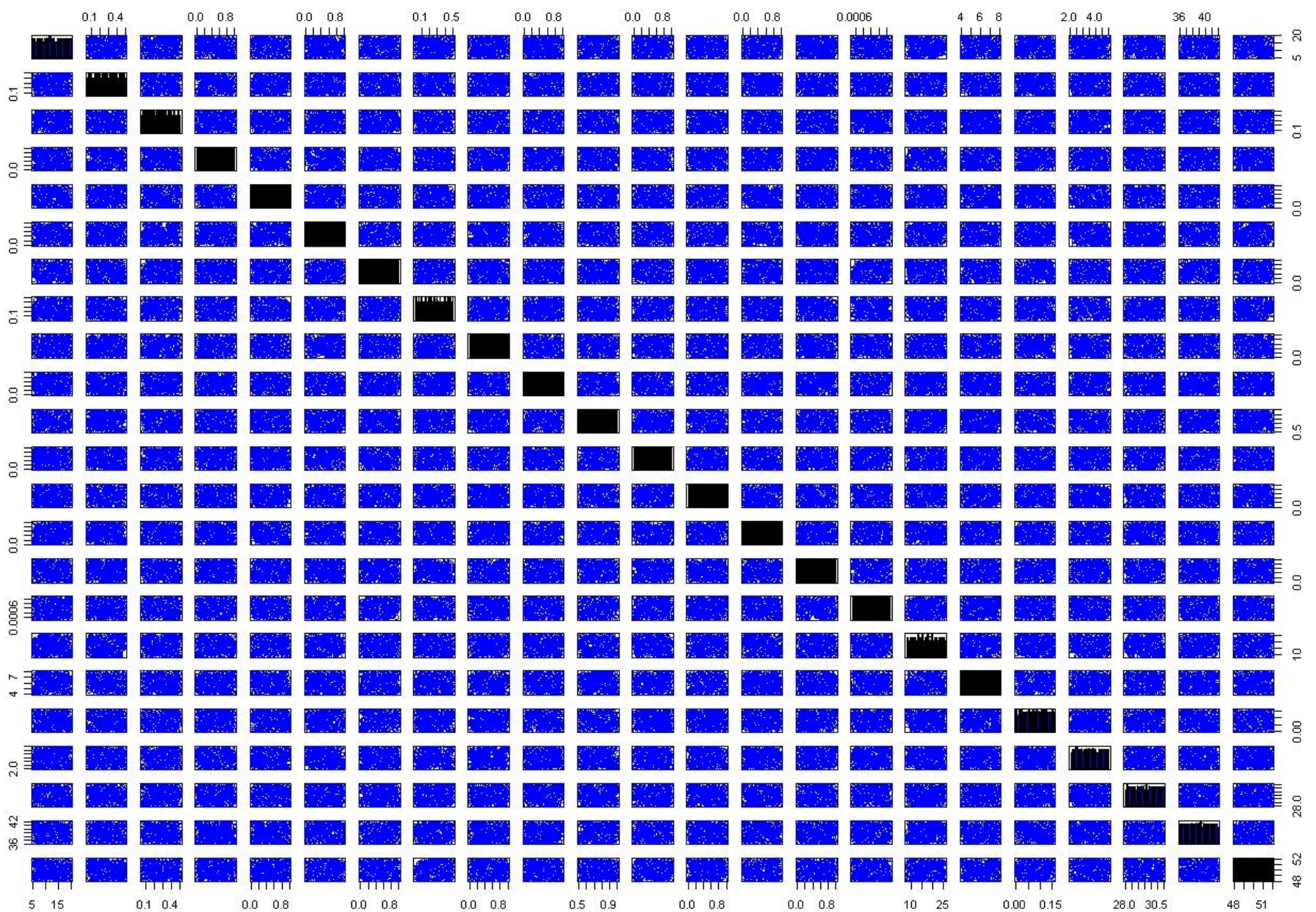
Preliminary work: fitting methods

- **Wave 1:** Sampled 200 sets of 23 biological and behavioural input parameters from relatively uninformative priors -> complex model -> point estimates for HIV prevalence population size etc. (calculated from mean of 100 runs)
- Trained an emulator and made predictions for 100,000 further input parameter sets
- Ranked parameter sets according to Mahalanobis distance from fitting points
- Weighted each parameter set relative to the best fit

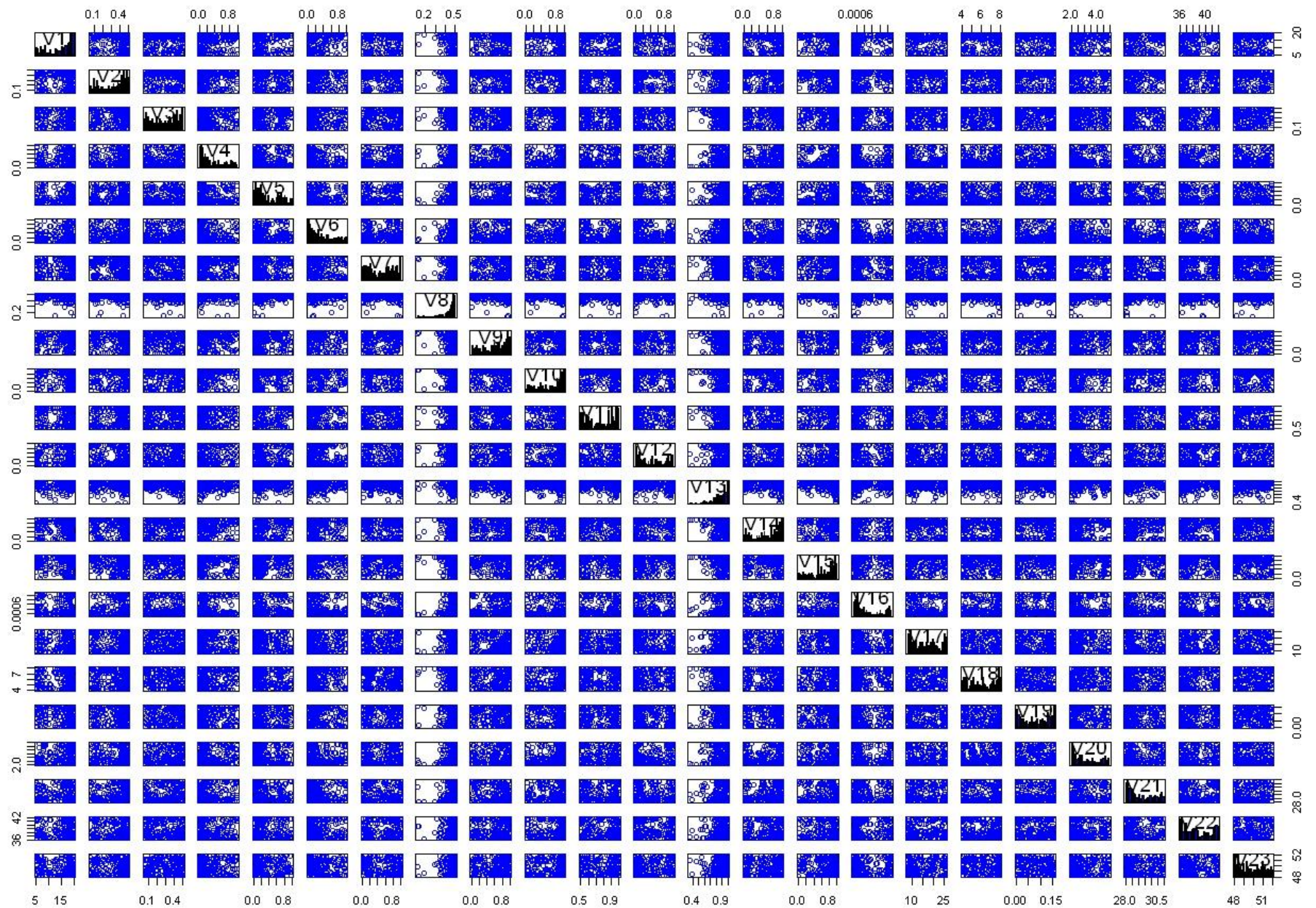
Preliminary work: fitting methods

- **Wave 2:** Sampled 200 parameter sets from the ranked and weighted wave 1 output using rejection sampling -> complex model
- Re-trained emulators incorporating these additional 200 points
- Made predictions for 100,000 further input parameter sets. These sets had been re-sampled according to the weights and then perturbed by a distance inversely proportional to their weight (using a Gaussian random walk)
- Again, parameter sets were ranked according to Mahalanobis distance from fitting points
- Weighted each parameter set relative to the best fit ready for re-sampling in wave 3

Preliminary work: wave 1 input



Preliminary work: wave 2 'posteriors'



Preliminary work: wave 2 results

- Early days! Reassuringly, we are still investigating most of parameter space.
- However, we can already gain some insight into the effects of one or two parameters. For example, it is not possible to fit the observed data if:
 - partner acquisition rate in the high sexual activity groups is very low
 - rates of partner acquisition are dramatically reduced after the first behavioural intervention
- It is easier to fit the observed data if a low proportion of women are in the high concurrency group.

Preliminary work: conclusions

- Hopeful that further waves will start to infer more detail about the regions of parameter space that will yield fits
- We are considering a vast parameter space but hope for an improvement on fitting 'by eye'
- Bayesian Emulation provides a promising approach to the calibration of complex models, greatly reducing computational time:
 - 100 model runs (1 parameter set): 30 mins on a high-performance cluster
 - 100,000 emulator predictions: 15 mins on a desktop PC

Preliminary work: limitations and questions

- No attempt to incorporate complex model variance into emulator training points
- Pitfalls of the emulator input re-sampling method. If the distribution of weights is biased, we will exacerbate this bias by homing in on highly weighted regions
- Invite comments on finding balance between re-sampling in promising regions and exploring elsewhere
- Is perturbing by a distance inversely proportional to weight reasonable?

Further work

- Continue!
- Currently evaluate most promising emulator parameter sets with complex model at the end of each wave. Consider using emulator 'sub-waves' to explore more parameter space before returning to complex model
- Refinement of re-sampling and perturbation of emulator inputs between waves
- Try training an emulator with the goodness of fit of the complex model, rather than the 'outputs' themselves -> a simpler univariate emulator?

Acknowledgements

- Supervisors: John Edmunds and Richard White: LSHTM
- Also Zaid Chalabi, Andy Cox, Pete Dodd, Toby Ealden, Ken Eames, Nicky McCreech and Katie O'Brien: LSHTM
- Stefano Conti, Daniela de Angelis: Health Protection Agency
- Ellen Brooks-Pollock, TJ McKinley and James Wood: University of Cambridge
- Ian Vernon & Michael Goldstein: University of Durham
- Jeremy Oakley: University of Sheffield