Sensitivity and uncertainty analysis of cardiac cell models with Gaussian process emulators

Richard Clayton, Sam Coveney, Richard Wilkinson, Jeremy Oakley

INSIGNEO Institute *in silico* Medicine, Department of Computer Science, and School of Mathematics and Statistics,

University of Sheffield

INSIGNEO
Institute for *in silico* Medicine

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

Outline

- Why we care about uncertainty in cardiac models.
- Approaches to sensitivity and uncertainty analysis.
- Worked example with a cardiac cell model.

Electrical activity in the heart

- The heart is an electromechanical pump.
- Electrical excitation initiates and synchronises contraction

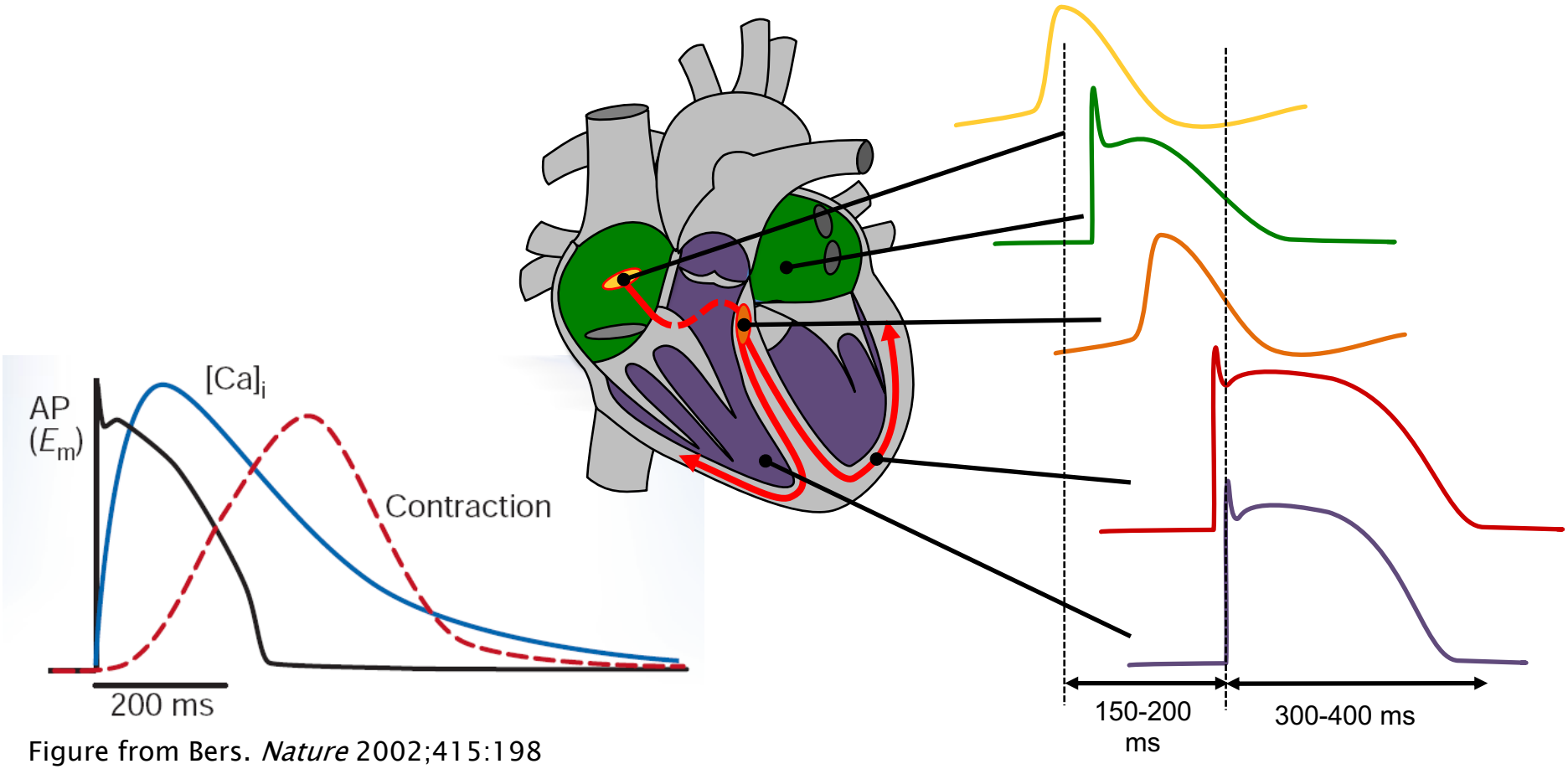


Figure from Bers. *Nature* 2002;415:198

Clinical translation

A possible use case for cardiac models:

- Atrial fibrillation (AF) is a common arrhythmia, and can be abolished by RF ablation.
- But ~40% of patients who have RF ablation require an additional procedure for atrial tachycardia, which is expensive and bad for the patient.
- Can we use patient-specific models of electrical activation to:
  - identify additional ablation sites?
  - Identify additional ablation sites ***in-procedure***?
  - Given noisy and low resolution, and measurements of electrical activation that are sparse and noisy.

Addressing this use case requires us to understand how our confidence in model predictions depends on uncertainty and variability in model inputs (parameters, geometry, numerics).

Sensitivity and uncertainty analysis

Sensitivity analysis – quantify how model outputs depend on each input (parameter) and their interactions.

- Insights into model itself.
- Identify model inputs that need to be controlled or measured precisely.
- Quantify effect of uncertainty and variability in model inputs.

Uncertainty analysis – quantify uncertainties in model outputs arising from uncertain model inputs.

- Estimate confidence in model output given uncertain inputs.
- May enable probabilistic workflows.

Worked example

Concentrate on model of the action potential.

- Set of ODEs, each ODE represents magnitude and kinetics of current through the membrane.
- Current carried by Na^+, Ca^{2+}, and K^+ through ion channels, pumps and exchangers.
- ODEs calibrated against experimental data.

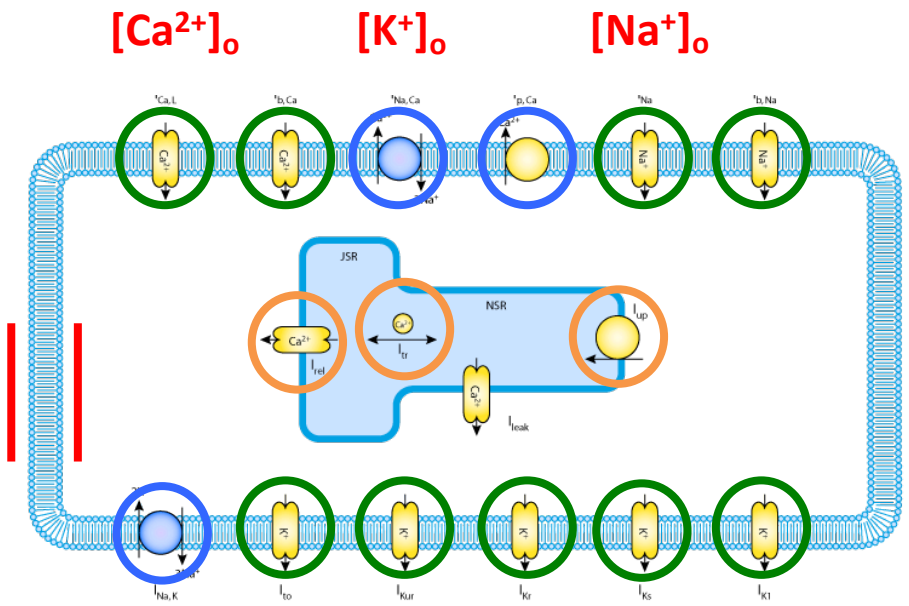
Sensitivity analysis – How do current magnitudes and boundary conditions affect the simulated action potential?

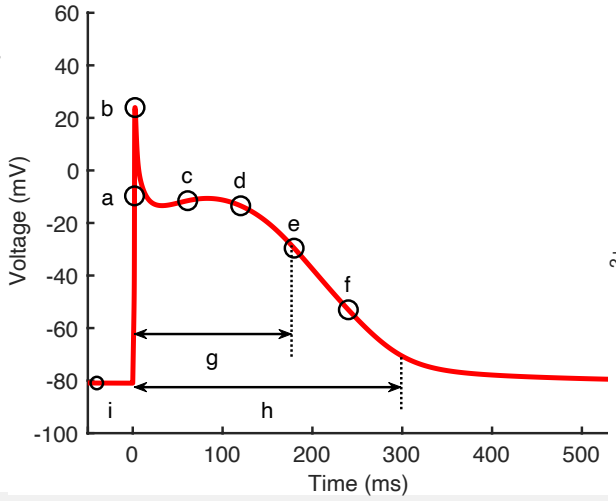
Uncertainty analysis – How do uncertain current magnitudes affect uncertainty in the action potential?

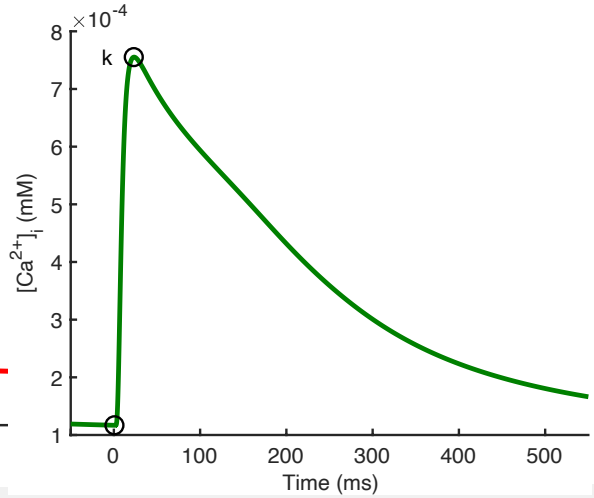
A cardiac cell model (CRN)

Minimal list of 20 model parameters:

- **9 ion channel maximal conductances**,
- **3 pump exchanger maxima**,
- **4 external boundary conditions (C_m, $[Na^+]_o$, $[K^+]_o$, $[Ca^{2+}]_o$).**
- **4 Ca^{2+} handling parameters**.
- Kinetic parameters not included.







Curse of dimensionality

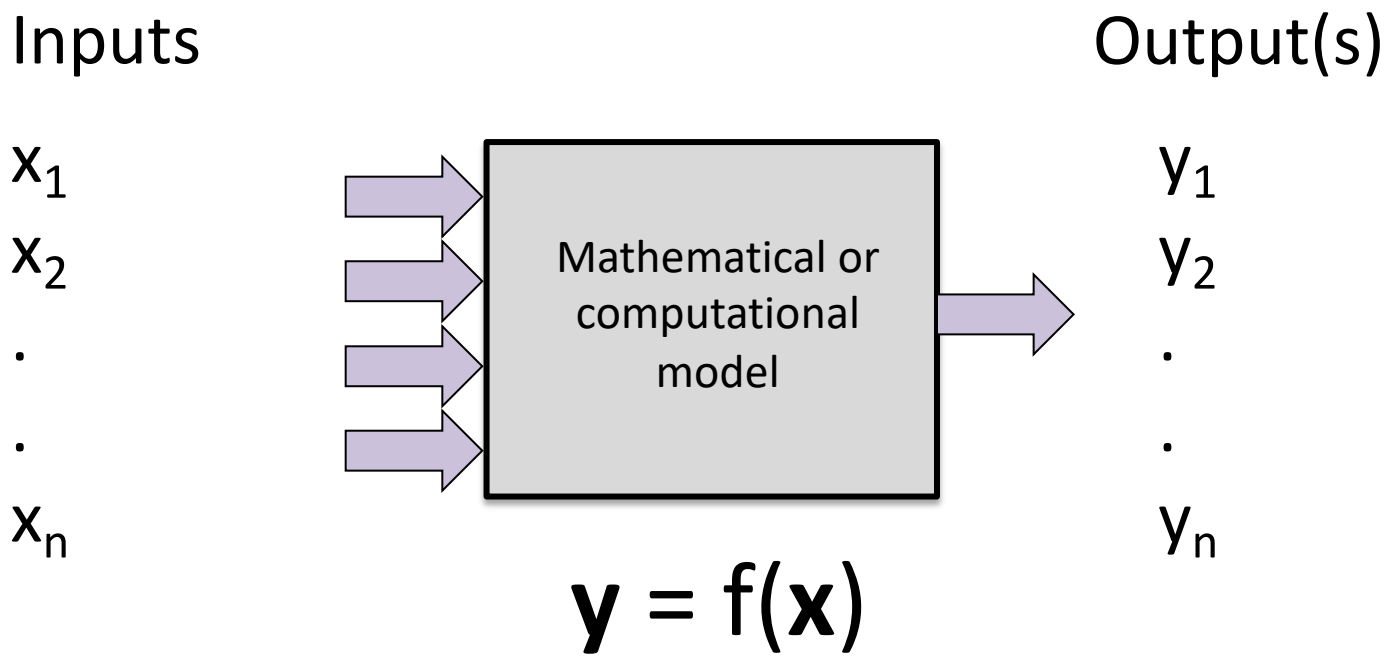
For a cardiac cell model, the number of model parameters (inputs) is at least *O*(10), sometimes O(100).

- To sample 1-D input space on [0 .. 1] evenly with spacing 0.5 requires 3^1 model evaluations.
- For 2-D input space, we require 3^2 evaluations.
- For N-D input space, we require 3^N evaluations.

The CRN model has 20 minimal inputs, so comprehensive sampling requires 3^{20} evaluations. **3.4×10^9 (110 yr @1s per run)**

Even sampling lower and upper limits requires 2^{20} evaluations (12 days).

Model description



- Sample **x** from a distribution, and evaluate **y** for each sample
- Replace **y** = f(**x**) with an emulator **y** ≈ f'(**x**)

Gaussian process emulators

A Gaussian process (GP) is a statistical model that is a distribution over functions.

It effectively interpolates an output surface **y** = f'(**x**).

A GP can be used to calculate expected output E[f'(**x'**)] (and its variance Var[f'(**x'**)]), for ***any*** input vector **x'**.

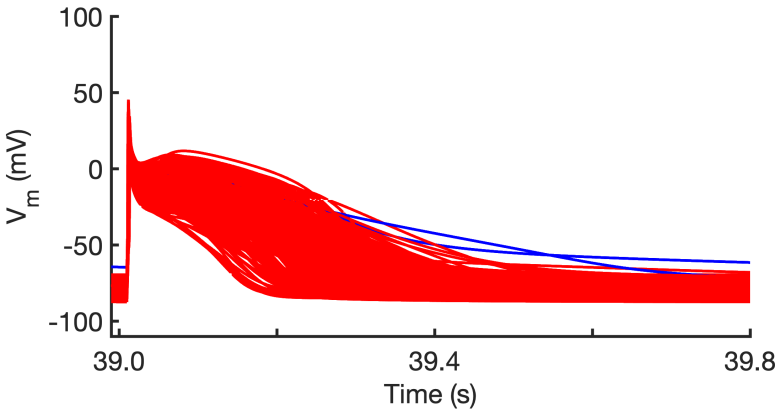
- The GP is trained using a set of ***design data*** **y** = f(**x**).
- No assumptions except that f(**x**) is smooth (-ish).
- Evaluating the emulator is fast, so can explore input space thoroughly.
- Any input can be uncertain, so **x'** can have a variance.
- If inputs and outputs are normally distributed, then ***we can calculate distributions on outputs directly***.

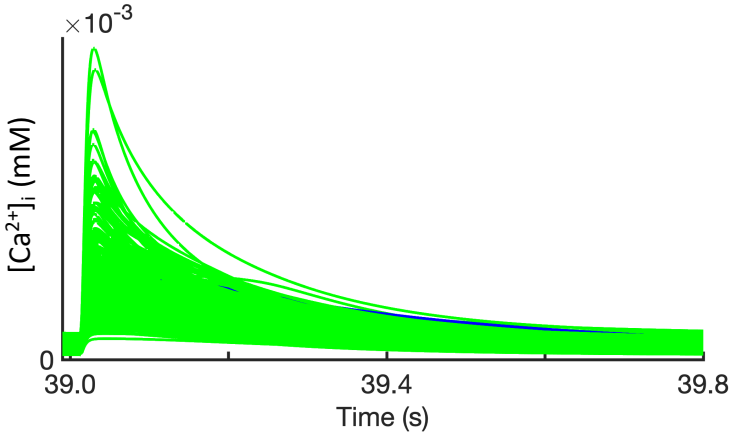
GP training and validation

Design data for training the emulators:

- Separate emulator for each output.
- Code auto-generated from CellML.
- 300 model runs (10 x number of inputs is rule of thumb).
- Each input varied from 0.5 x to 1.5 x default (±50%), except G_{K1} and C_m (±25%), and ion concentrations (±10%).
- Model runs with pacemaking, APD>500 ms, APD alternans, or other problems removed (5/300).

Validation against 150 additional runs using mean average percentage error, <10%.





Sensitivity indices

First order index – expectation of the reduction in output variance if we learn input x_i exactly:

$$S_i = \frac{Var^*[\mathbf{y}] - E[Var^*(\mathbf{y}|x_i)]}{Var^*[\mathbf{y}]} = \frac{Var^*[E(\mathbf{y}|x_i)]}{Var^*[\mathbf{y}]}$$

Total effect index – reduction in output variance if we know every input except x_i ($x_{\sim i}$) exactly:

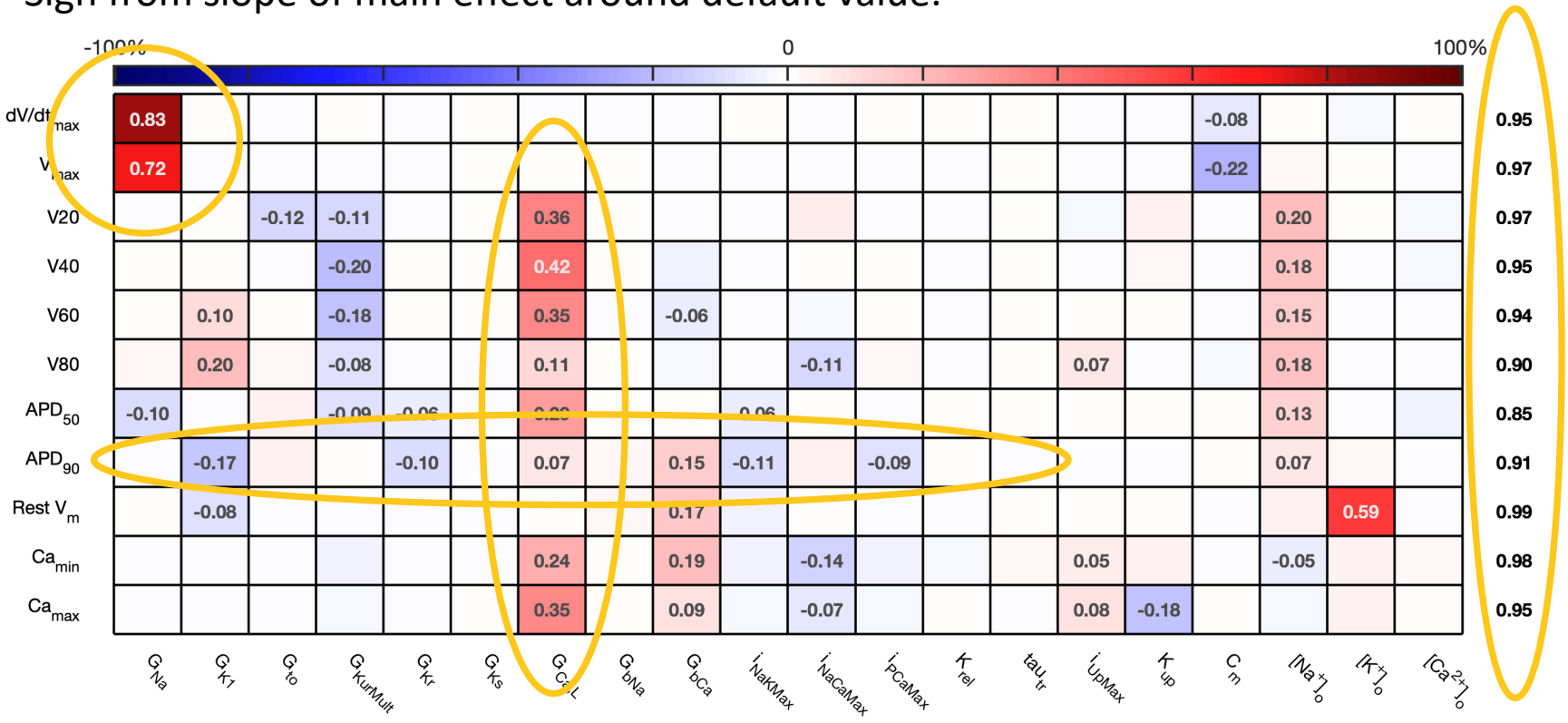
$$S_{Ti} = \frac{Var^*[\mathbf{y}] - Var^*[E(\mathbf{y}|x_{\sim i})]}{Var^*[\mathbf{y}]}$$

These are enough to understand both sensitivity at first order, and interactions.

Saltelli, Chan, Scott, 2000, Oakley and O'Hagan 2004

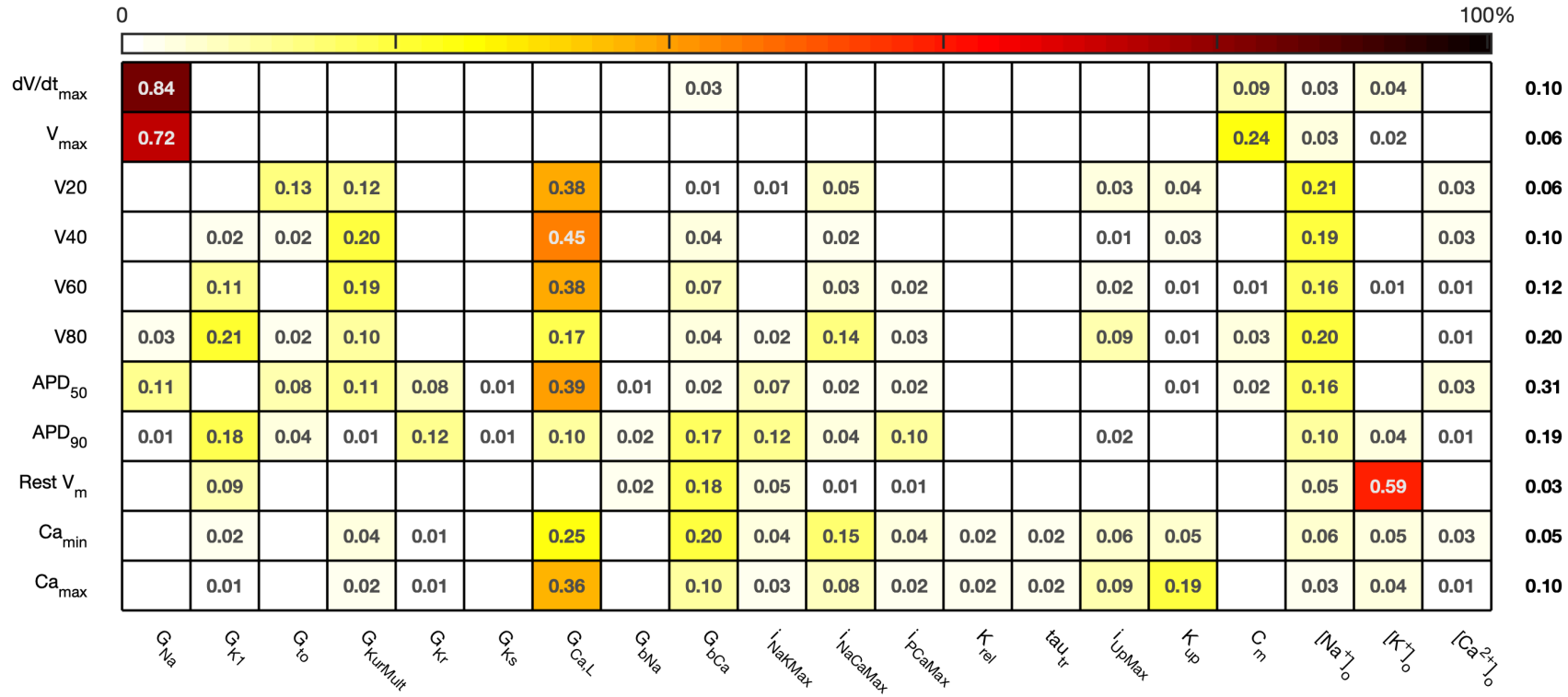
First order indices – CRN model

- Variance based sensitivity indices: proportion of output variance accounted for by each uncertain input – ***directly calculated*** from GP emulators.
- Inputs set to default value, with standard deviation of 0.2 in normalized units.
- Sign from slope of main effect around default value.



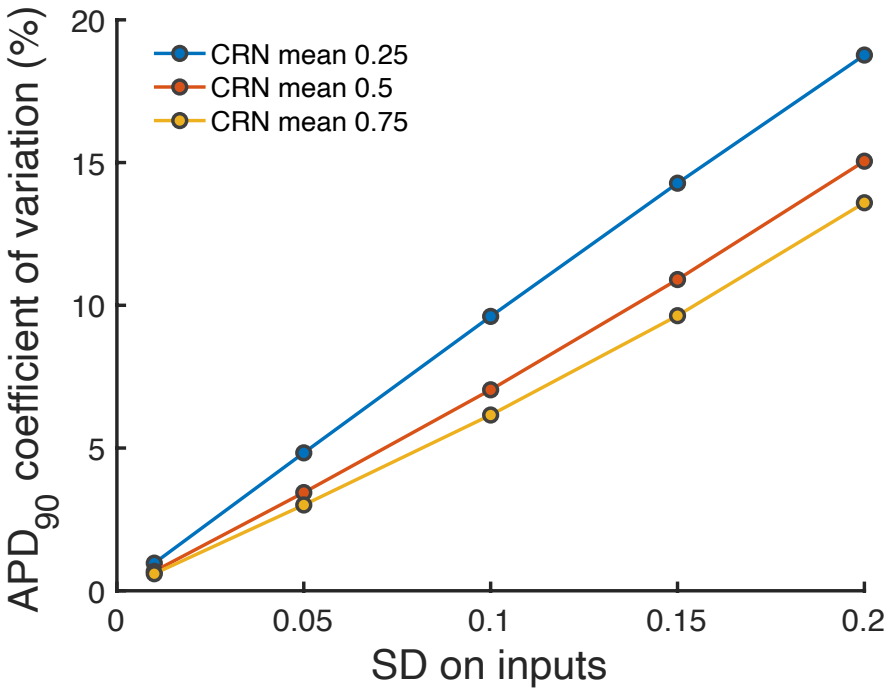
Total effect indices – CRN model

- Numbers on right are sum of differences between first order index and total effect index – indication of interactions.
- Most differences <0.05 – effect of interactions is small.

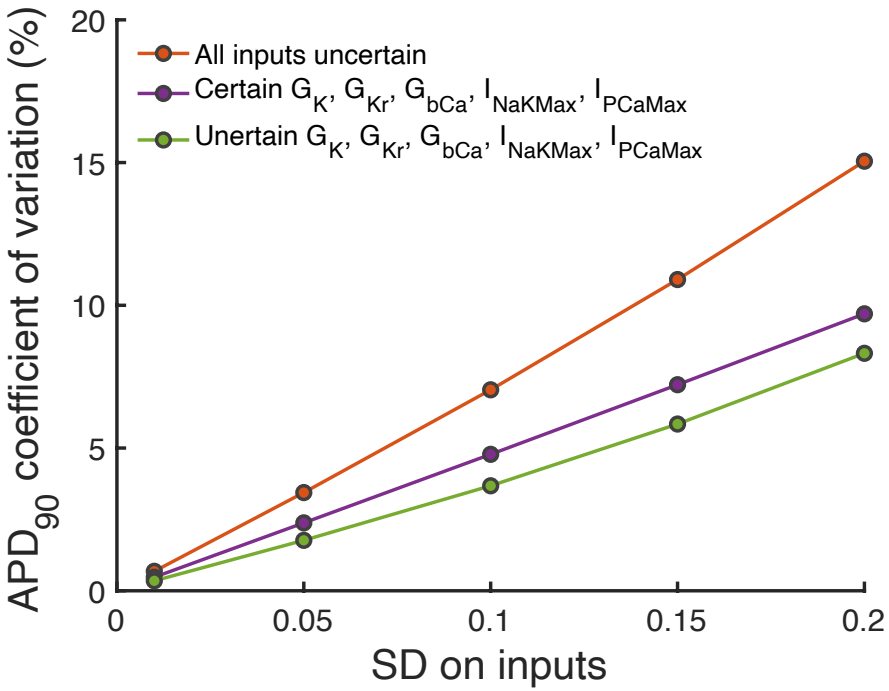


Uncertainty analysis for APD90

All inputs uncertain



Some inputs uncertain



Concluding remarks

- As modellers, we must account for uncertainty and variability more carefully than we do at the moment – single model runs for a single set of parameters are not good enough.
- GPs are effective, but other tools exist and it is not yet clear what approaches work best.
- Variability and uncertainty in Physiome models is interesting, topical, difficult, and absolutely crucial for models that will have credibility.
- Challenges include:
  - Non-uniqueness and identifiability – different combinations of parameters can produce identical action potentials.
  - Incorporating dynamic behaviour into analysis.
  - Extending to propagation in tissue.
- Want to have a go with GPs? Download our Python implementation from https://github.com/samcoveney/maGPy

EPSRC
Engineering and Physical Sciences Research Council

Thanks for listening

Ongoing work

Model analysis and comparison:

- GPs enable sensitivity of different models to the same inputs to be compared.
- Investigate mechanisms of variability, compare to real cardiac cells.

Model calibration:

- Evaluating a GP is very fast – ~10^6 emulator evaluations in ~20 mins on single core, i.e. ~10^3 per second. Model evaluation takes ~10 s, so 10^4 speedup.
- Enables model calibration using history matching – see https://doi.org/10.1016/j.pbiomolbio.2018.08.001

Challenges:

- Non-uniqueness – different combinations of parameters can produce identical action potentials.
- Incorporating dynamic behaviour into analysis – variable diastolic interval as another input.

GP emulators - 2

Emulator

$$f'(\boldsymbol{x}') = h(\boldsymbol{x}')^T \beta + c(\boldsymbol{x}', \boldsymbol{x})$$

Vector of inputs

Mean function

Covariance function

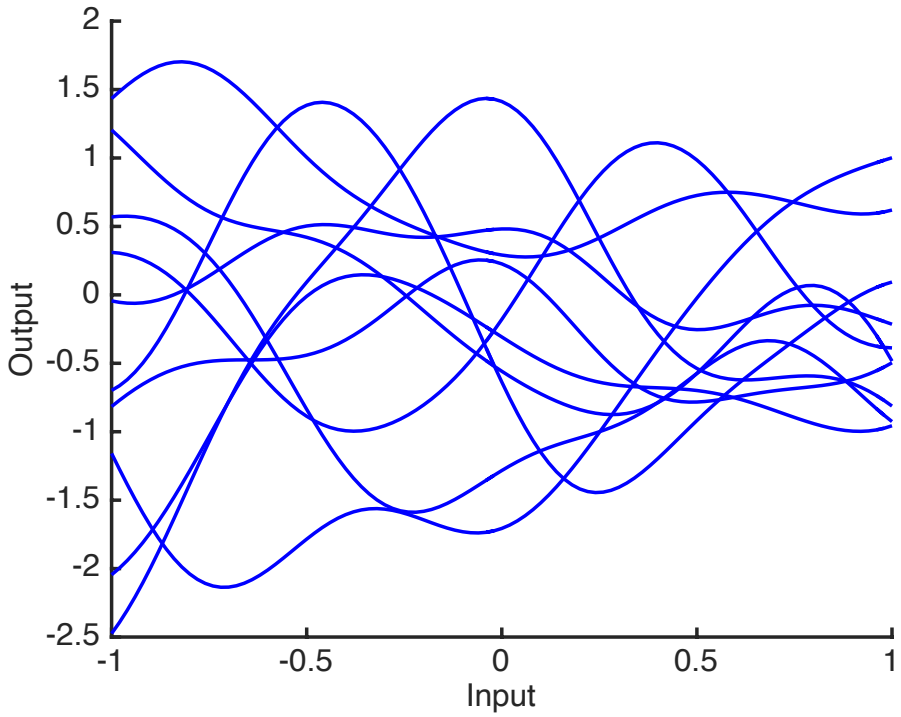
Hyperparameters

$$h(\boldsymbol{x}')^T \beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_P x_P$$

$$c(\boldsymbol{x}', \boldsymbol{x}) = \sigma^2 \, exp\left[-\sum_{p=1}^{P} \left\{ \frac{\left(x'_p - x_p\right)}{\delta_p} \right\}^2 \right]$$

- Hyperparameters can be optimised using a set of ***design data*** – a set of simulator input and output data.
- Maximise log-likelihood of hyperparameters given design data – Bayesian trick.
- Posterior mean and variance of emulator are then conditional on design data – Chang et al PloS ONE 2015

Simple example



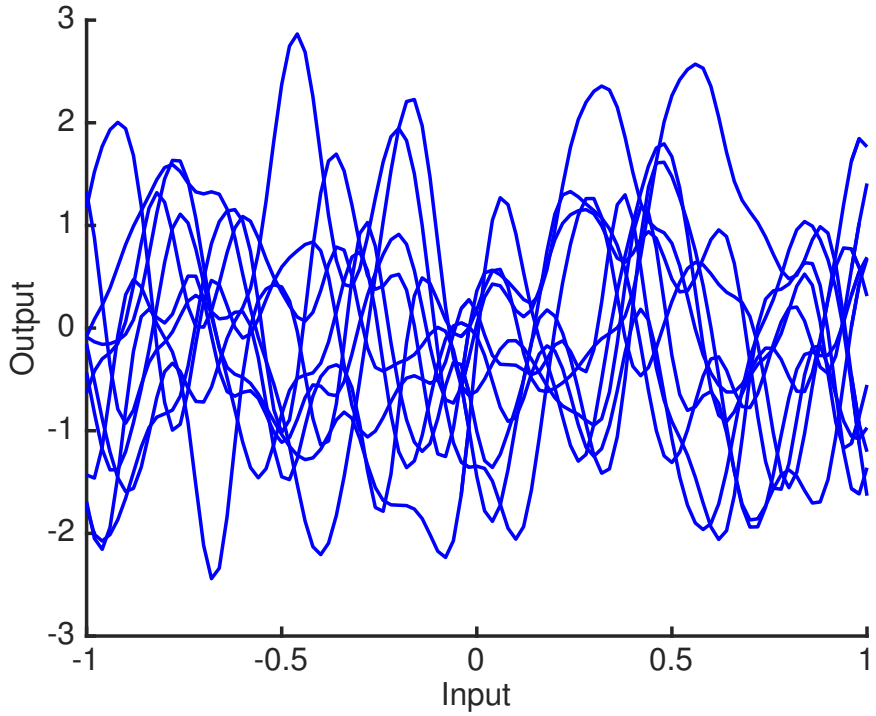
$\sigma^2 = 1.0, \delta = 0.5$

10 samples from a GP.

Here x and x' random numbers, mean is zero.

- σ^2 – how far f(x) deviates from mean.
- δ – length-scale (wiggliness) of f(x).

Simple example



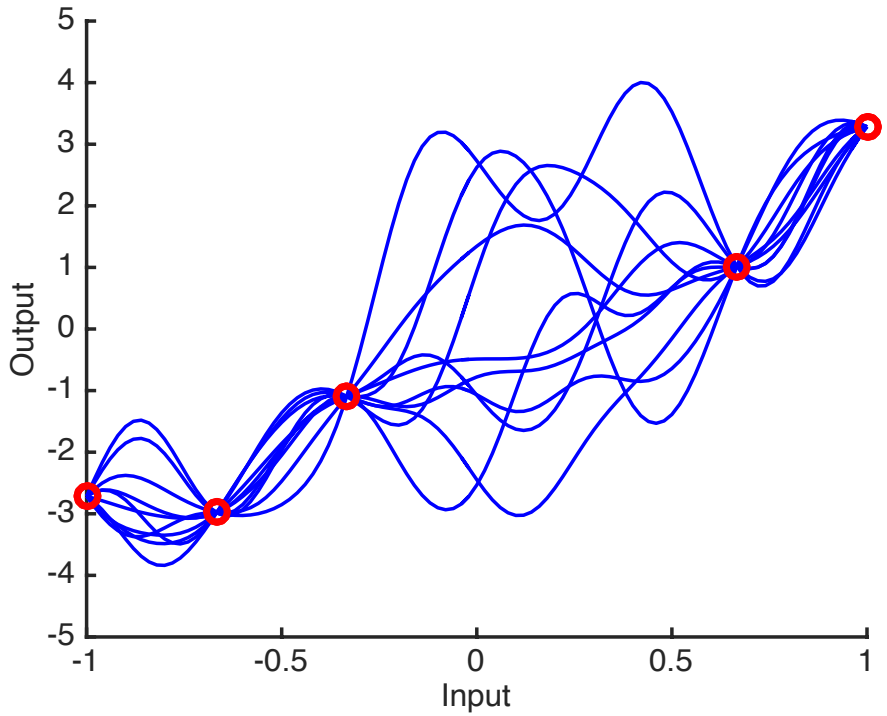
$\sigma^2 = 1.0$, $\delta = 0.1$

10 samples from a GP.

Here x and x' random numbers, mean is zero.

- σ^2 – how far f(x) deviates from mean.
- δ – length-scale (wiggliness) of f(x).

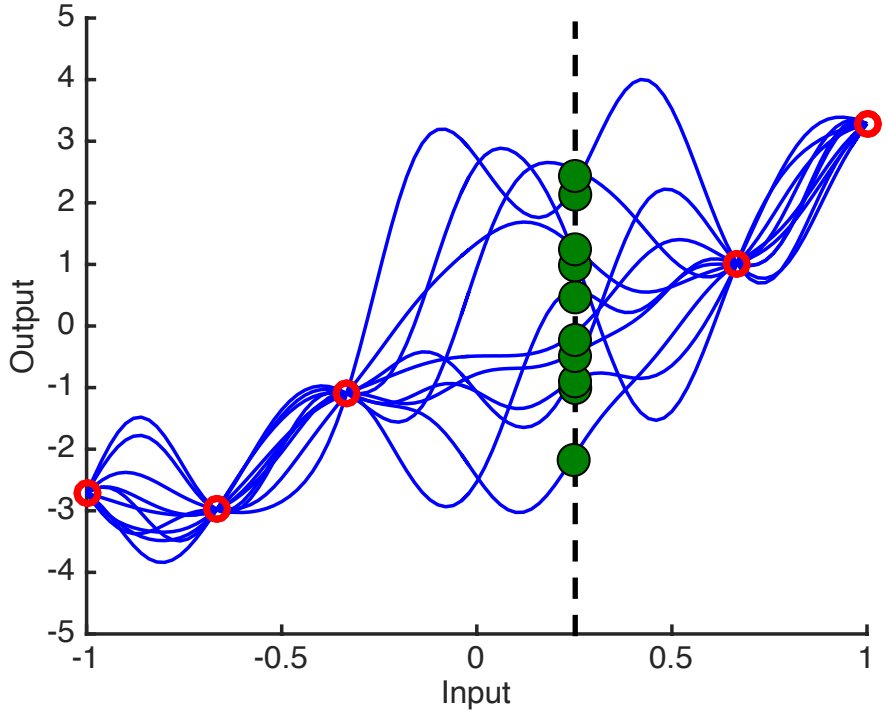
Simple example



GP hyperparameters have been fitted using ***design data*** $y_1 = f(x_1)$ – red points.

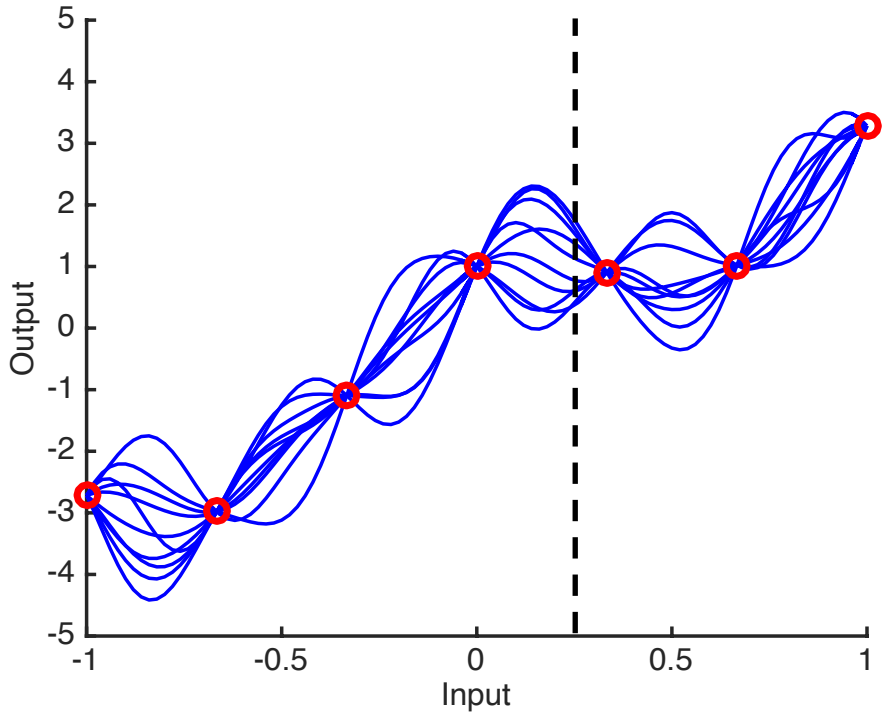
Process gives maximum likelihood of GP hyperparameters given design data

Gaussian process



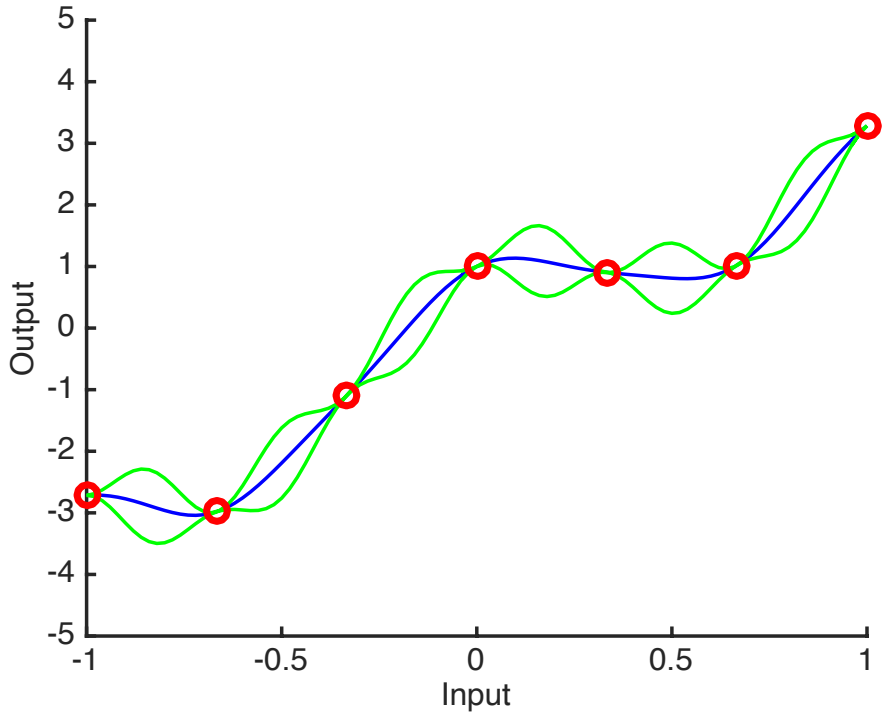
GP can then provide estimates of $y_2 = f(x_2)$ given the design data.

Gaussian process



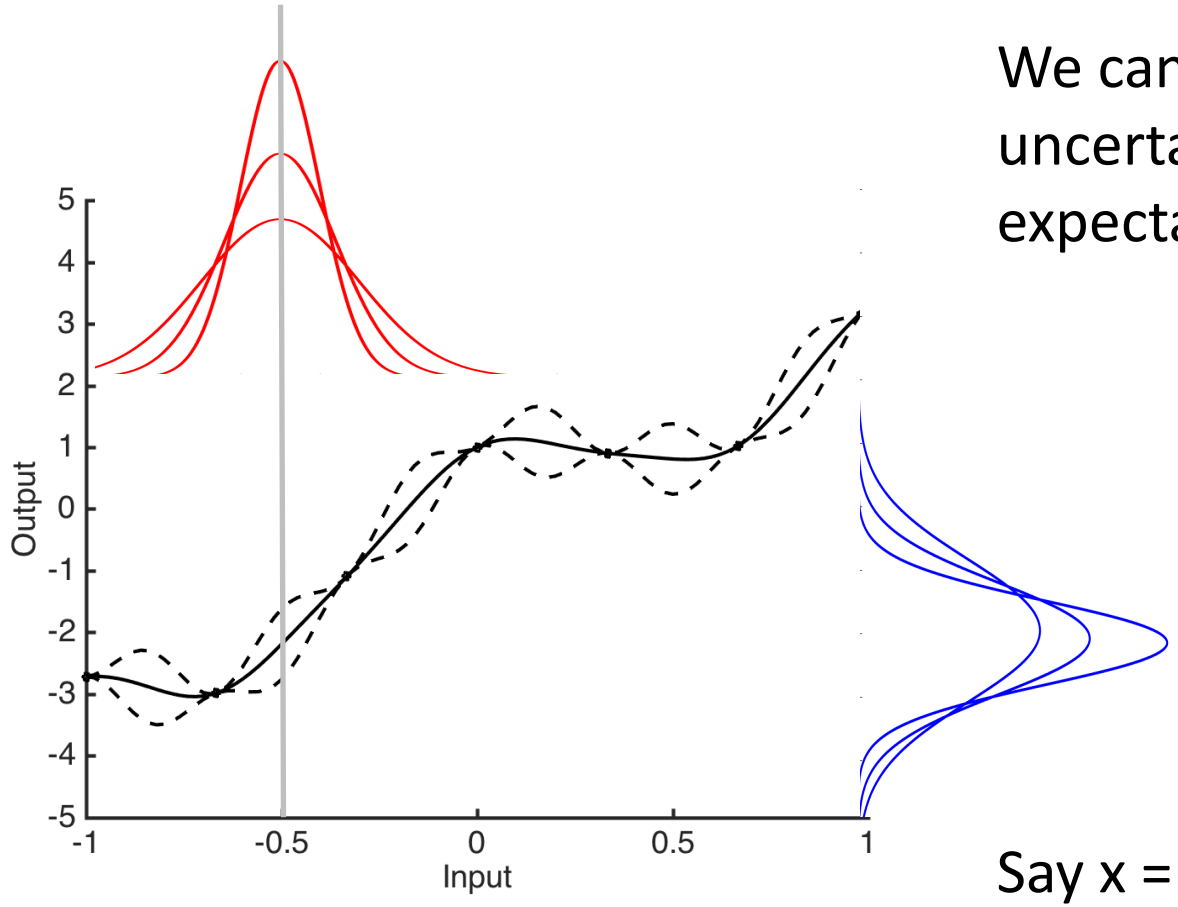
GP fitting takes place in Bayesian setting, so we can update GP parameters with extra data.

Gaussian process



For any value of x, we can directly calculate the expected output y, and its variance.

Gaussian process



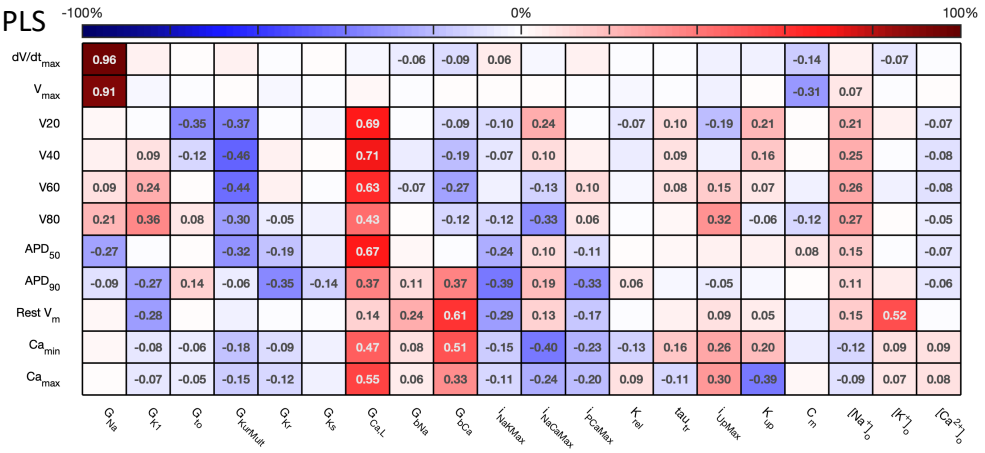
We can also treat x as uncertain, with an expectation and variance.

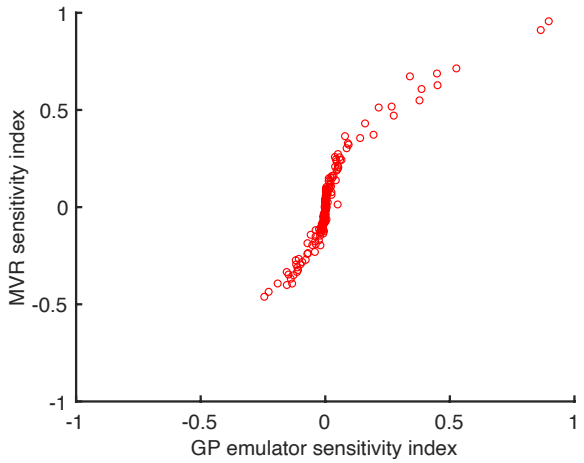
Say x = 0.5, with variance of 0.01, 0.02, and 0.04.

Comparison with PLS B-indices

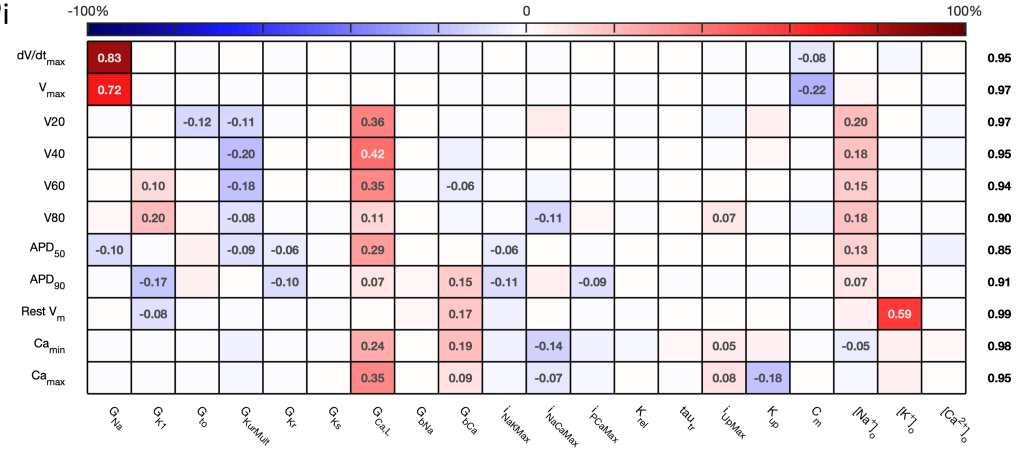
B_{PLS} obtained by minimising |**Y'**-**Y**| where **Y'** = **XB**, **X** and **Y** are based on design data, with each x and y regularised by subtracting mean and dividing by standard deviation.

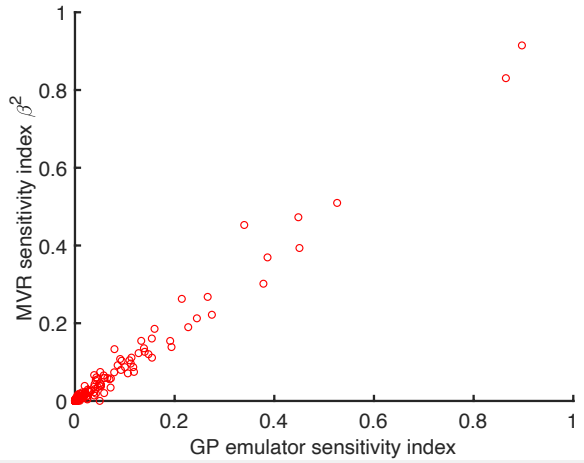
B_{PLS}



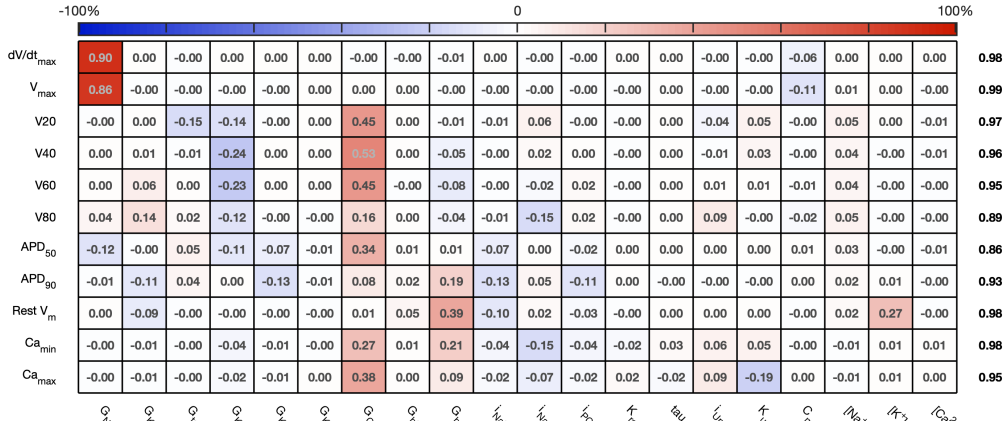


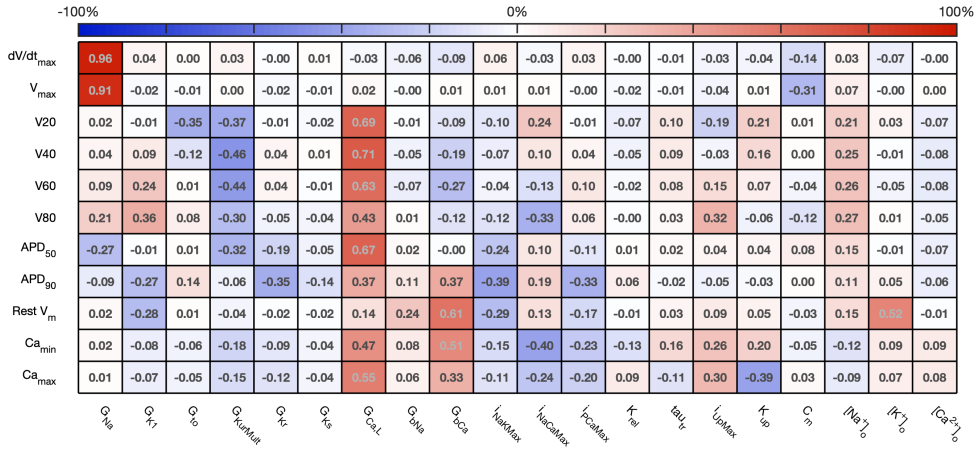
S_i

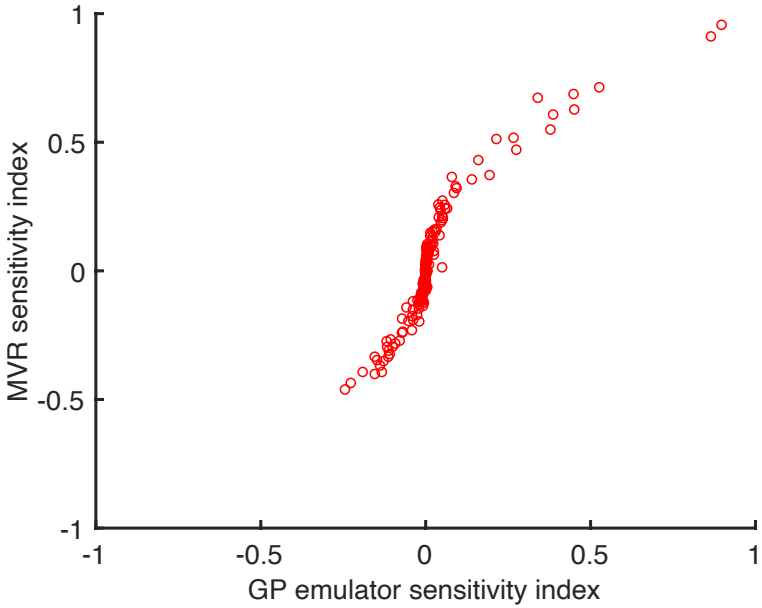




SI comparison







Uncertainty in cardiac models

Sources of uncertainty in cardiac models include:

- ***Intrinsic variability*** – within and between cells.
- ***Measurement uncertainty*** – in experiments used to construct and calibrate models.
- ***Lack of information/knowledge*** – some quantities are impossible to measure.
- ***Parameter uncertainty*** – models calibrated from variable and uncertain data.
- ***Condition uncertainty*** – in initial and boundary conditions.
- ***Geometry uncertainty*** – in a computational mesh.

Useful models must strike a balance between model complexity and model uncertainty.

(Mirams et al, *J Physiol* 2016, Eck et al, *Int J Numer Meth Biomed Eng* 2015)

Main effects