Approximating Costly Functions Using Gaussian Processes

Hossein Mohammadi

Living Systems Institute, University of Exeter, UK

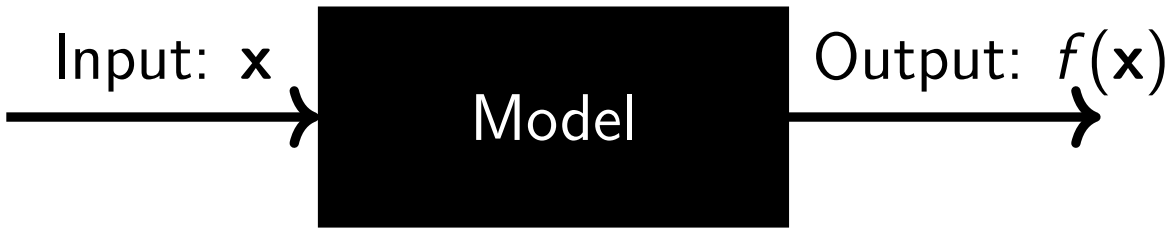
Developments in Healthcare Imaging

2 May 2018

Issac Newton Institute, Cambridge, UK

Computer models (numerical simulators)

- Mathematical representations of real-world phenomena
- Widely used in many applications
  - Help to better understand physical processes
  - Conducting physical experiments are too costly, time-consuming or even impossible
- Examples in biology: cardiac models, brain models, etc.
- The input/output relation is govern by a (unknown) function f

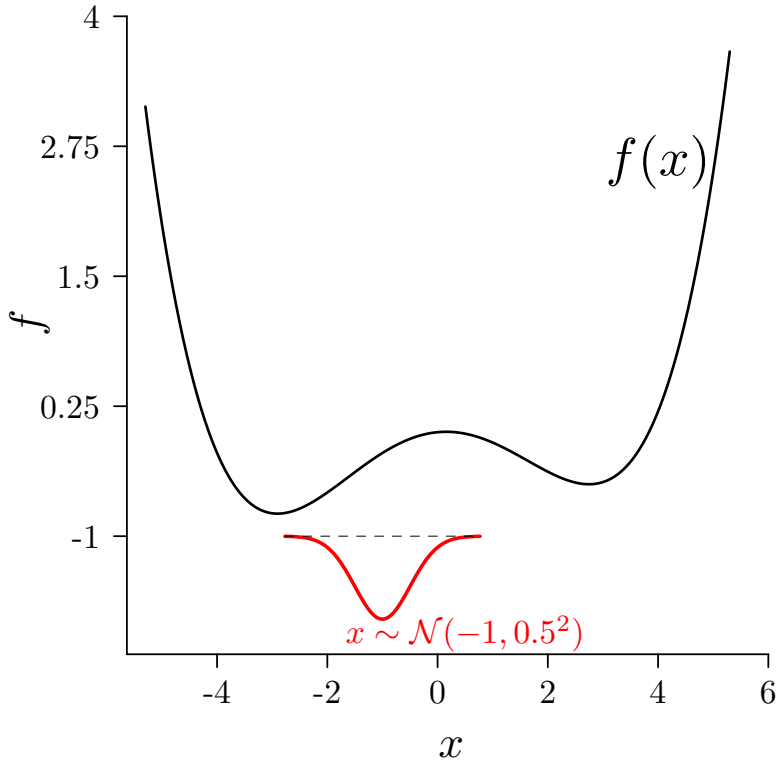


How good are these models?

- All models have inputs
- These inputs are often uncertain
- Uncertain inputs leads to uncertain outputs
- Decision makers need an estimation of the uncertainty propagated through a model $\rightarrow$ uncertainty analysis

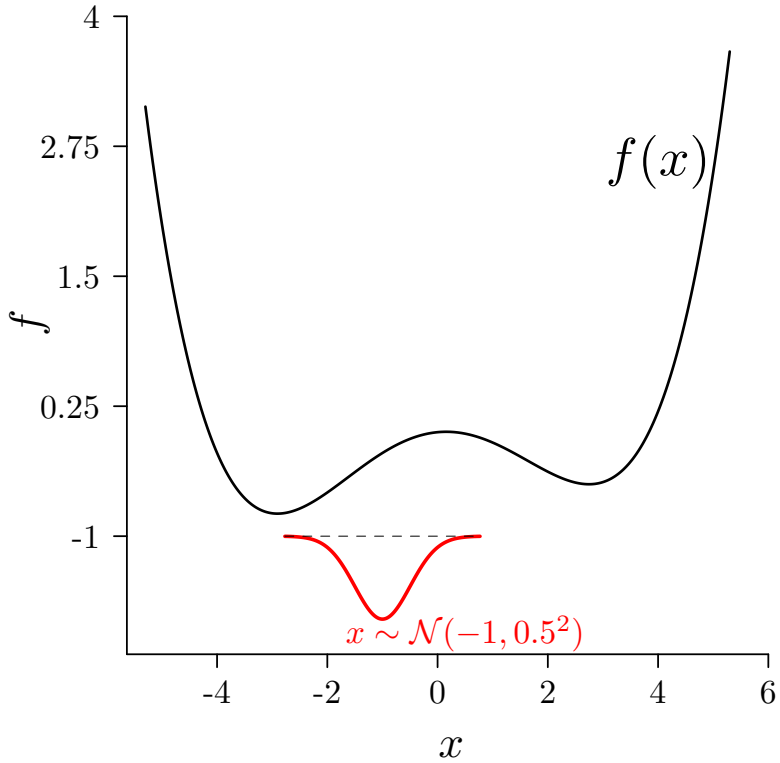
Uncertainty analysis

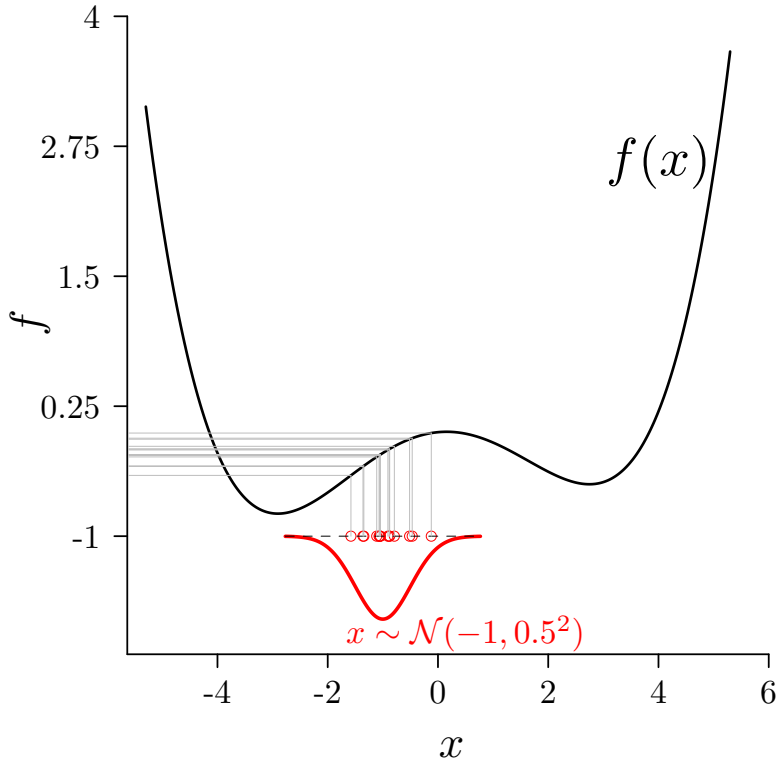
Given the input distribution, what is the output distribution?



Uncertainty analysis

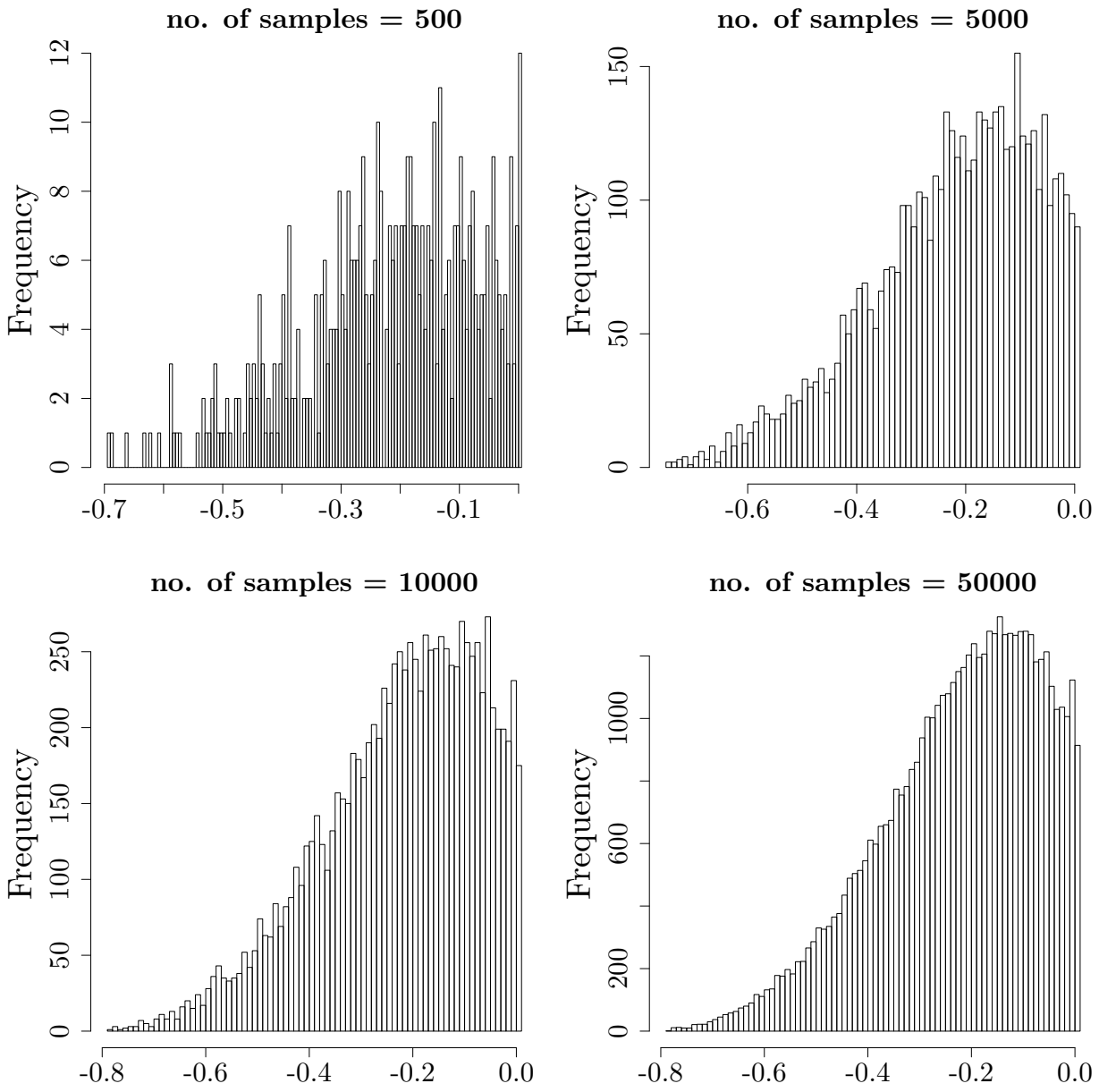
Given the input distribution, what is the output distribution?





Monte-Carlo method

Monte-Carlo results



If each evaluation (run) takes only one minute, you need to wait $50000 \times 1\ min \approx 34\ days$.

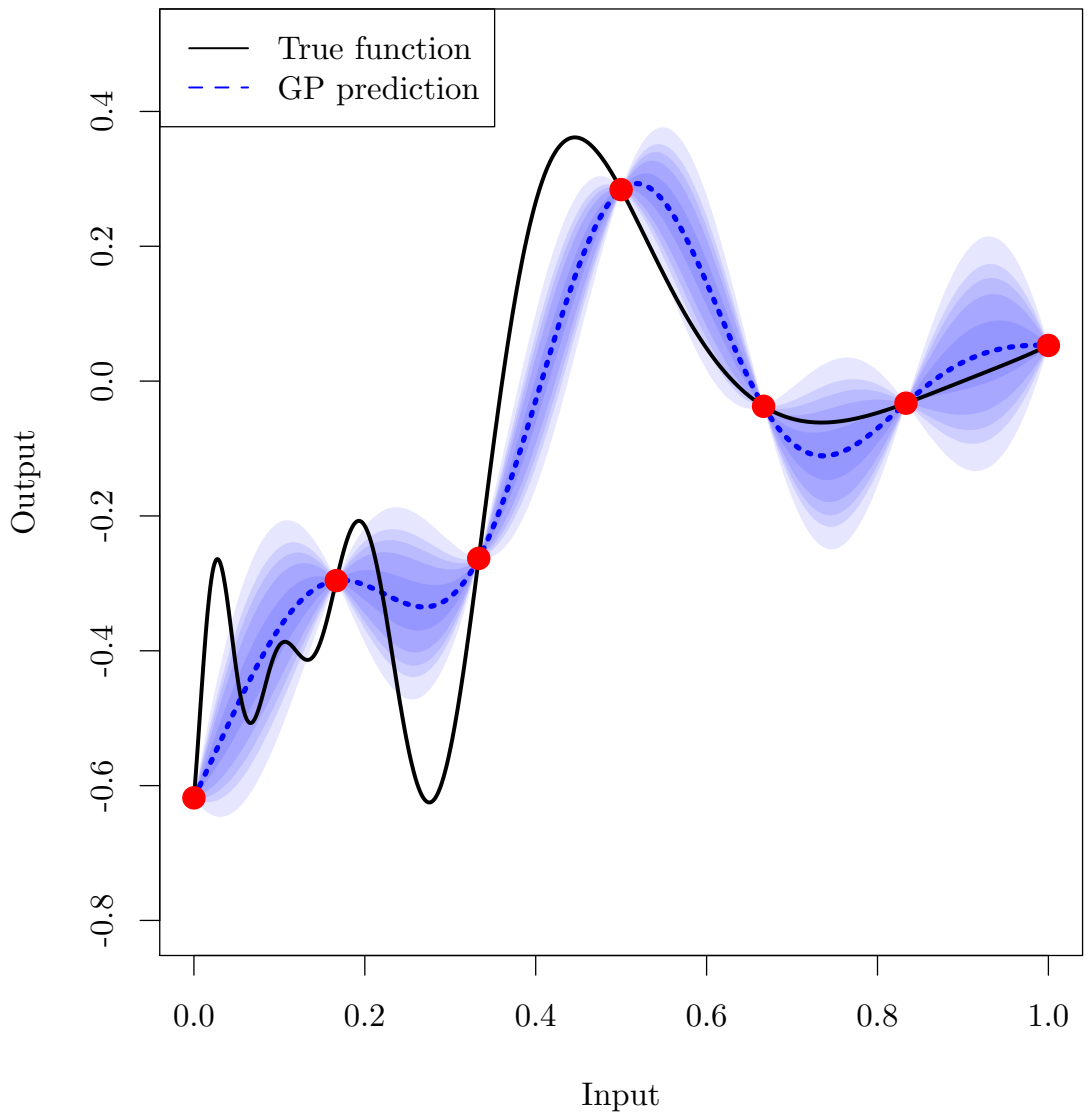
If each evaluation (run) takes only one minute, you need to wait $50000 \times 1\ min \approx 34\ days$.

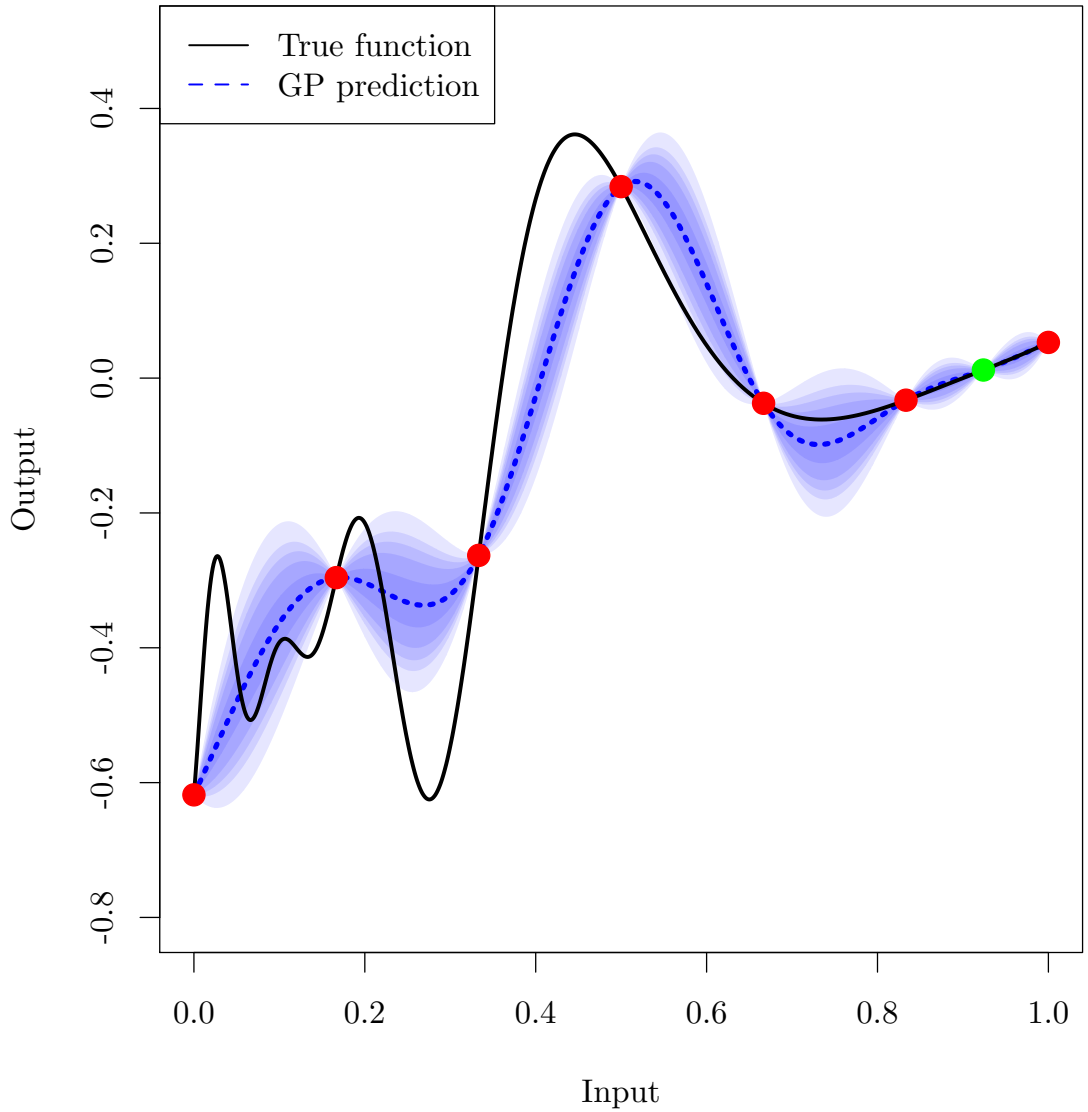
Solution: Replace f with $\hat{f}$ using surrogate techniques such as regression models, neural networks, **Gaussian processes**, etc.

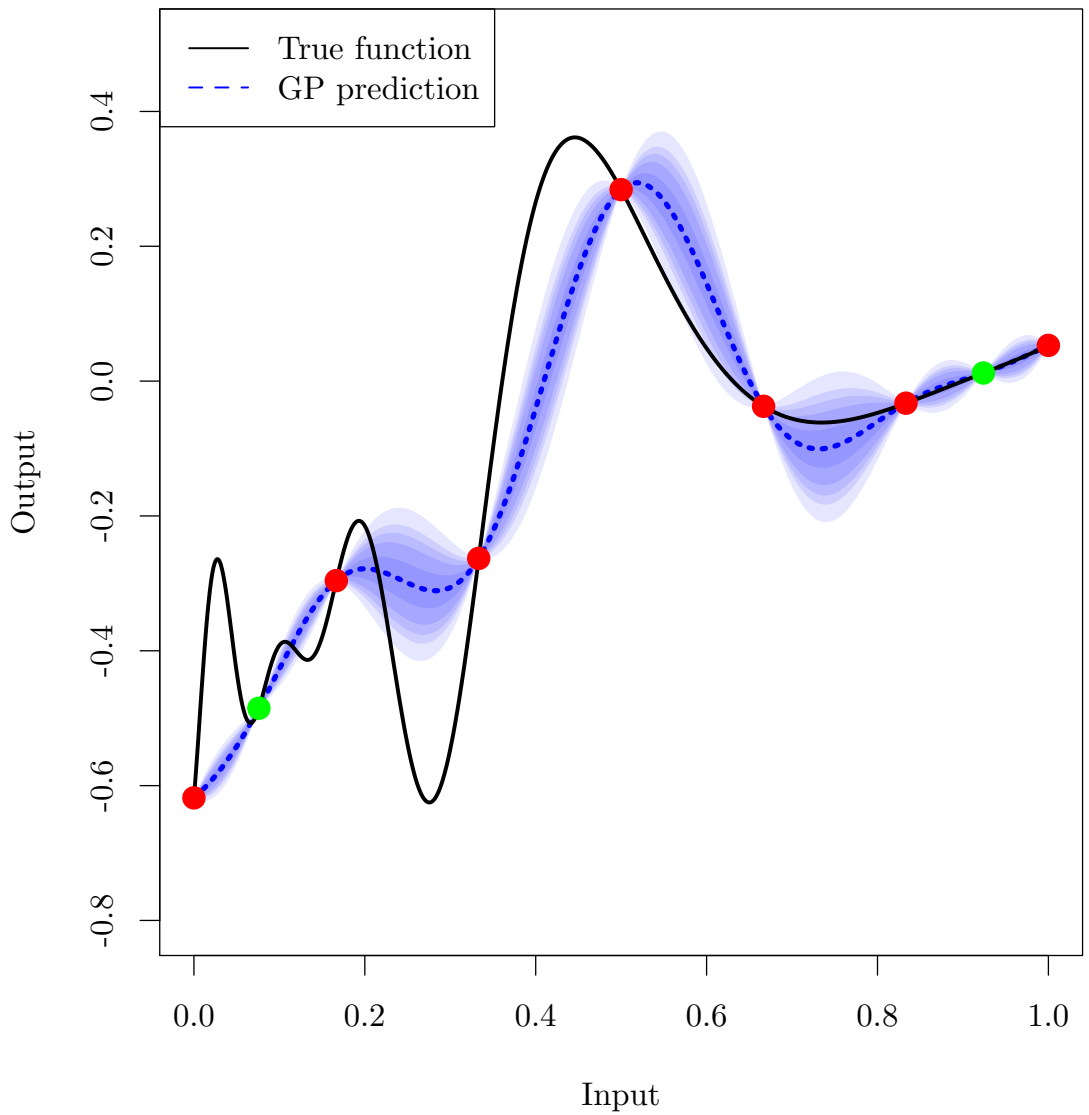
Why Gaussian processes (GPs)?

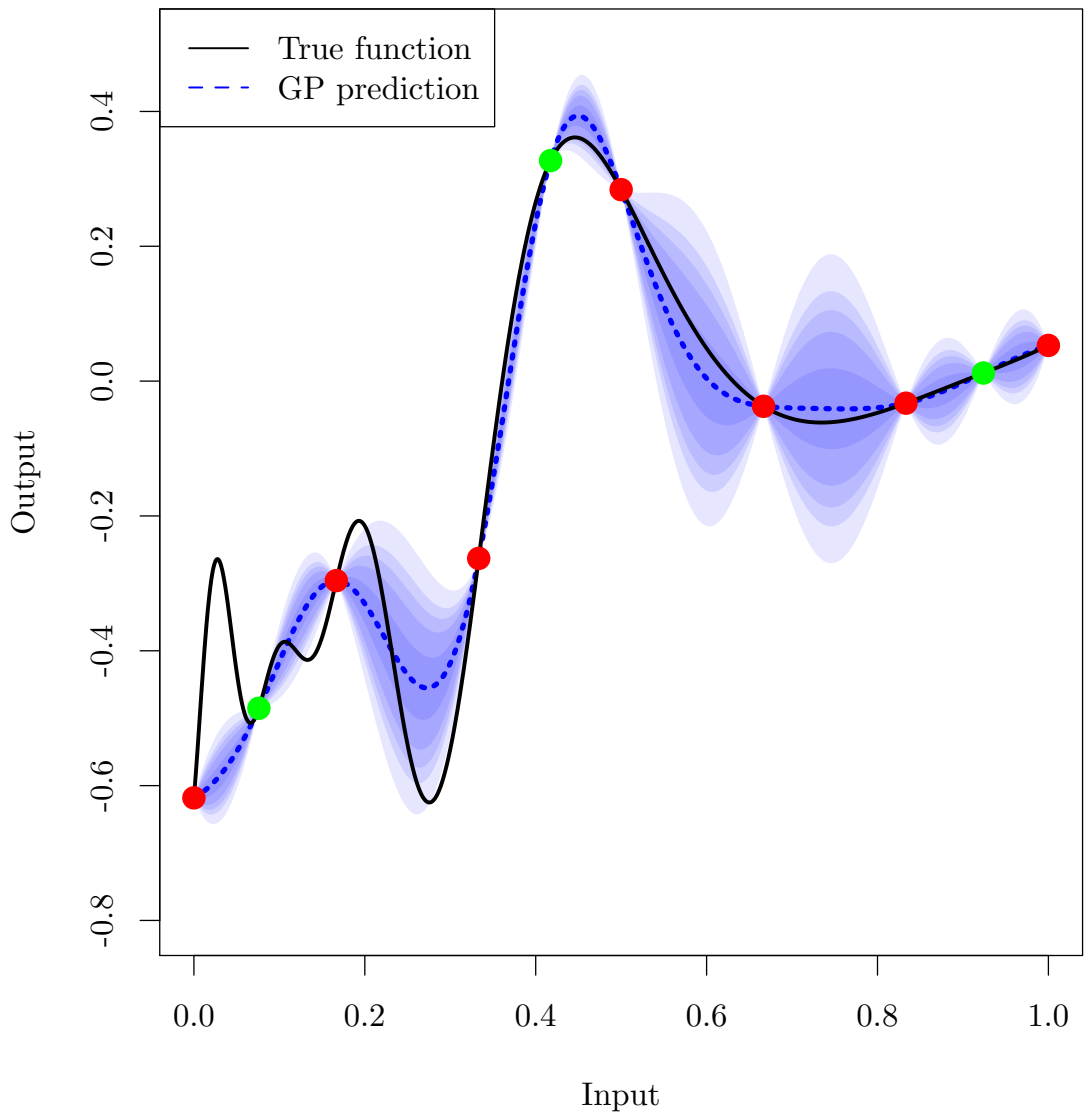
- Nonparametric models, no strong assumptions about the form of the function
- Require a small number of training data
- Comparable to (if not better than) other approaches like neural networks [1]
- Equipped with the uncertainty in their prediction (see next slides)

[1] Carl Edward Rasmussen. *Evaluation of Gaussian Processes and Other Methods for Non-linear Regression*. Ph. D. thesis, University of Toronto, 1996.

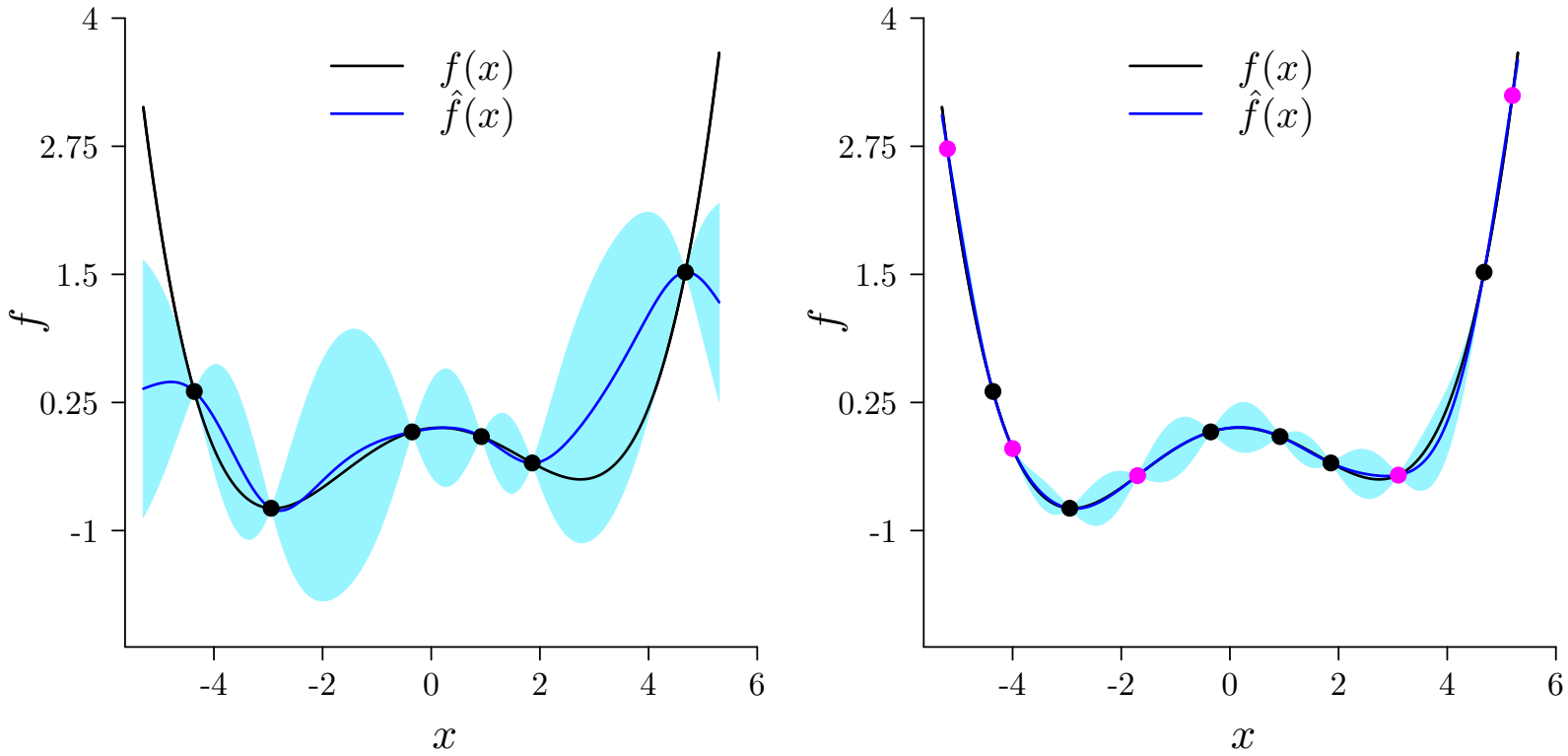

True function
GP prediction
Output
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
0.0
0.2
0.4
0.6
0.8
1.0
Input


True function
GP prediction
Output
Input
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
0.0
0.2
0.4
0.6
0.8
1.0


True function
GP prediction
Output
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
0.0
0.2
0.4
0.6
0.8
1.0
Input


True function
GP prediction
Output
Input
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
0.0
0.2
0.4
0.6
0.8
1.0

Back to the uncertainty analysis example



Left: GP prediction with six training data. Right: The predictive performance is improved by adding five new samples (cyan).

Cardiac Modelling (jointly with KCL)

- Diastolic heart disease is an incurable condition that affects 450k people in the UK
- Can we distinguish between healthy and unhealthy patients using a model
- Heart simulator takes 6 hours to run
- Emulate heart model
- Use inverse methods ('history matching') with MRI scan data
- Can separate patients onto two classes of model parameters (so far only one of each!).

Thank you for your attention!