Global Sensitivity Analysis for Complex Models

Jeremy Oakley
University of Sheffield

9 October 2018

Overview

Global sensitivity analysis: investigating how output of a model responds to changes in inputs, over some input range of interest

1. Notation / definitions
2. Variance-based methods for global sensitivity analysis
3. Computation

Set-up and notation

- We have a computer model $y = f(\boldsymbol{x})$, output y and input $\boldsymbol{x} = (x_1, \ldots, x_d)$.
- We consider a scalar output (e.g. one element or a univariate summary of a multivariate output)
- f usually not available in closed form.
- f constructed from modeller's understanding of the process.
  - there may be no physical input-output data.
- f assumed deterministic

Model inputs

- Model $y = f(\boldsymbol{x})$, input $\boldsymbol{x} = (x_1, \ldots, x_d)$.
- Different types of 'input'

1. 'Control input': value is known, no uncertainty, e.g., modeller chooses a value of x_i corresponding to some case of interest
2. Uncertain input or 'parameter': modeller needs to specify a value of x_i, but unsure what value to use

- Two categories:
  1. "Observable input": has meaning outside the model
  2. "Tuning parameter": artefact of the model, used to fit the model to data
- In this talk, interest is in uncertain, observable inputs.

Local and global sensitivity analysis

- Interest in understanding how model output responds to changes in individual inputs
- Local sensitivity analysis

$$\left.\frac{\partial f}{\partial x_i}\right|_{\boldsymbol{x}=\boldsymbol{x}^*}$$

- Global sensitivity analysis: how does output vary as input varies over some region?

Global SA

- Suppose each uncertain input has a **true value** X_i
- We assign a probability distribution to each X_i:
  - separate source of data
  - expert judgement
  - often assumed in the SA literature that (after scaling) $X_i \sim U[0, 1]$
- Now define $Y = f(X_1, \ldots, X_d)$
- Will analyse how X_i contributes to uncertainty in Y
- Analysis is of the *combination* of f and $p(\boldsymbol{X})$

Need to think carefully about input distributions

- Consider

$$f(x) = \exp(-x),$$

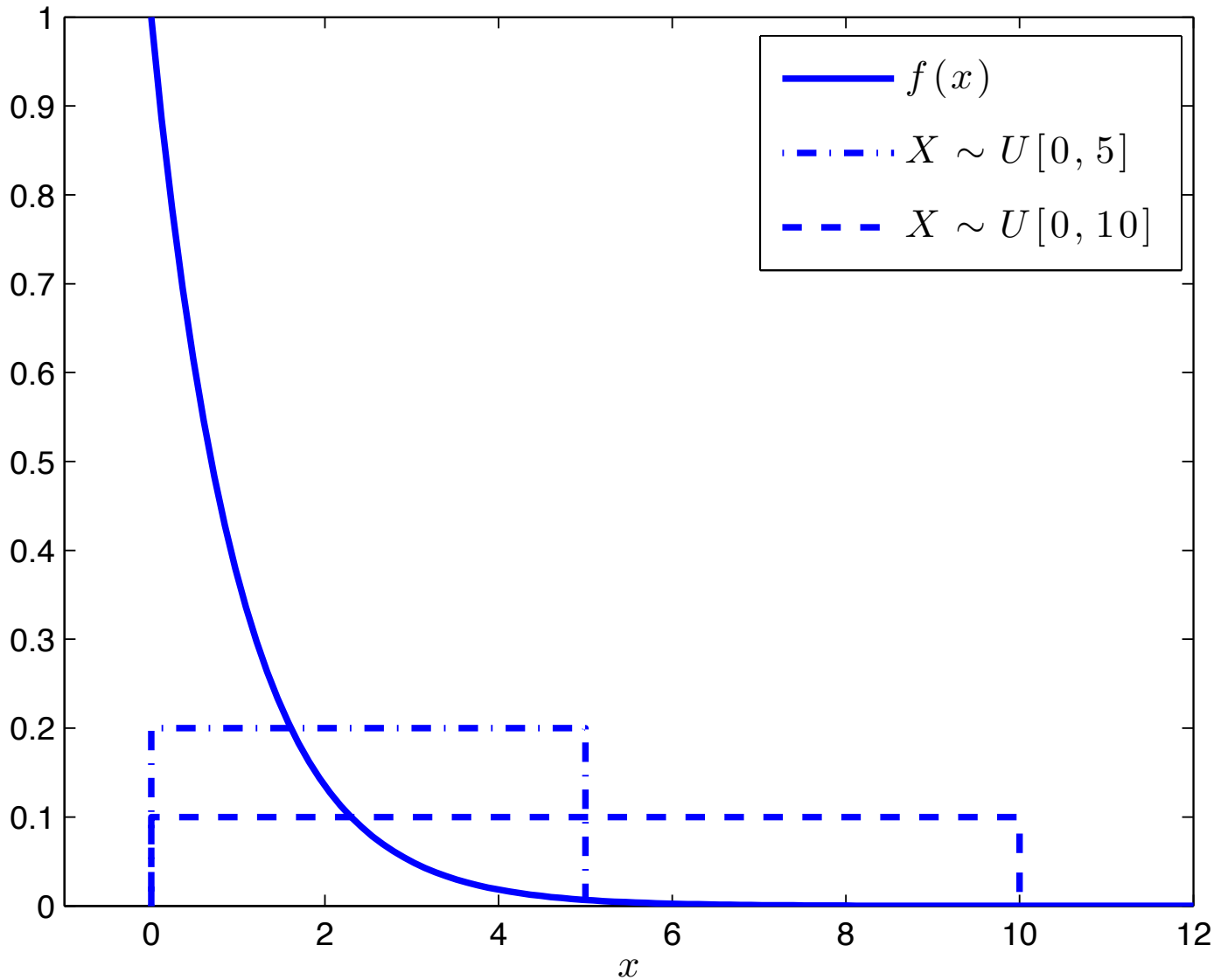
with $Y = f(X)$ and

$$X \sim U[0, b].$$

- In this case we have

$$Var(Y) = \frac{b - 2 + 4\exp(-b) - (b+2)\exp(-2b)}{2b^2},$$

- Increasing b increases the variance of X but *decreases* the variance of Y


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
x
$f(x)$
$X \sim U[0, 5]$
$X \sim U[0, 10]$

Variance based sensitivity analysis

(See, e.g., Saltelli et al., 2008)

- Consider $Y = f(X_1, \ldots, X_d)$ with $X_1, \ldots, X_d$ uncertain
- Is it worth reducing/eliminating uncertainty about a particular X_i (or subset of $X_1, \ldots, X_d$)?
- In variance-based approach, describe uncertainty about Y using $Var(Y)$
- We consider

$$Var_{X_i}\{E_{\boldsymbol{X}_{-i}}(Y|X_i)\}$$

with

$$E_{\boldsymbol{X}_{-i}}(Y|X_i) = \int f(X_1, \ldots, X_d) p(\boldsymbol{X}_{-i}|X_i) d\boldsymbol{X}_{-i}$$

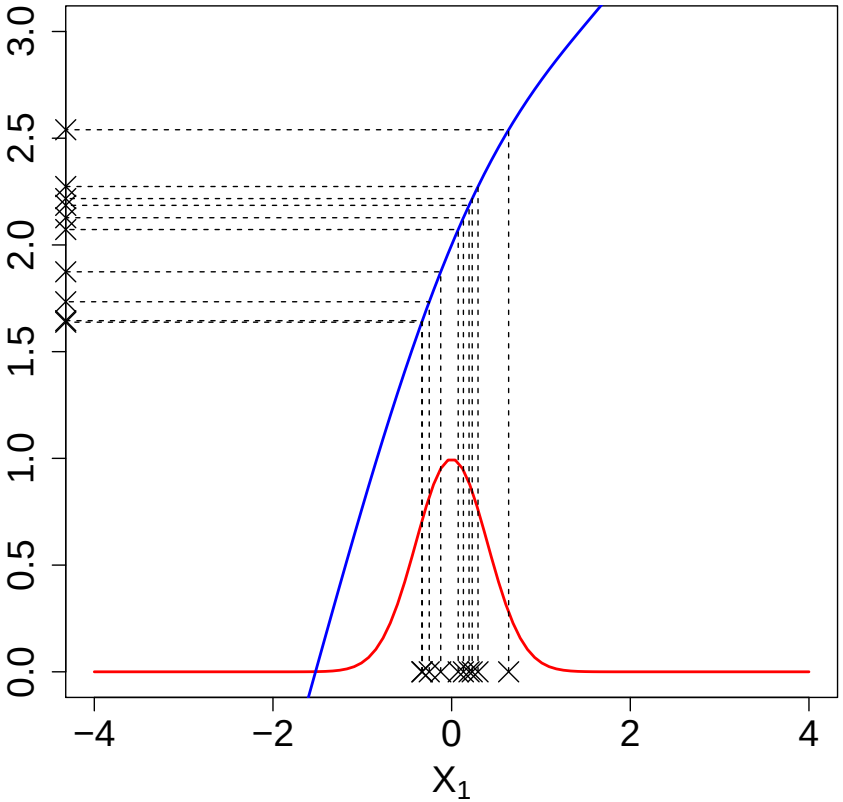
$$Var_{X_i}\{E_{\boldsymbol{X}_{-i}}(Y|X_i)\}$$

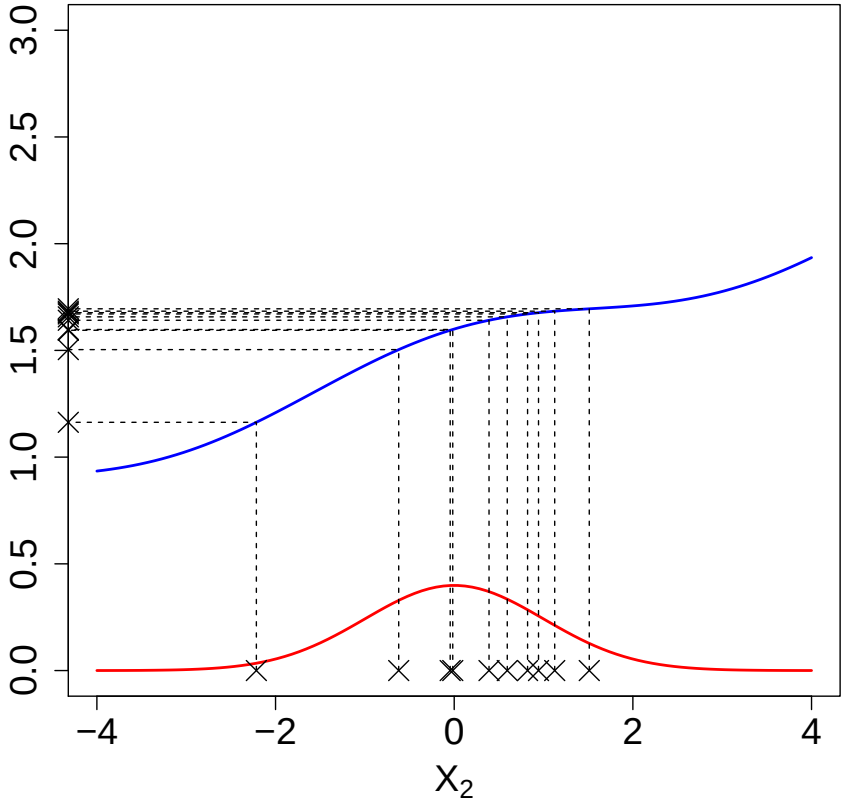
- This gives *expected* reduction in $Var(Y)$ achieved by learning true value of X_i:

$$Var(Y) - E_{X_i}\{Var_{\boldsymbol{X}_{-i}}(Y|X_i)\} = Var_{X_i}\{E_{\boldsymbol{X}_{-i}}(Y|X_i)\}$$

- Interpretation holds for independent or dependent inputs

Variance based SA





— $p_{X_1}(x_1)$

— $E(Y|X_1 = x_1)$

— $p_{X_2}(x_2)$

— $E(Y|X_2 = x_2)$

Variance-based sensitivity analysis

$$S_i = \frac{Var_{X_i}\{E_{\boldsymbol{X}_{-i}}(Y|X_i)\}}{Var(Y)}$$

- Various names: main effect index, Sobol' index, correlation ratio
- Small value does *not* imply 'unimportant': contribution to variance may be through interactions, e.g.

$$Y = X_1X_2 + X_3X_4,$$

with $X_1, \ldots, X_4 \overset{i.i.d}{\sim} N(0,1)$, then

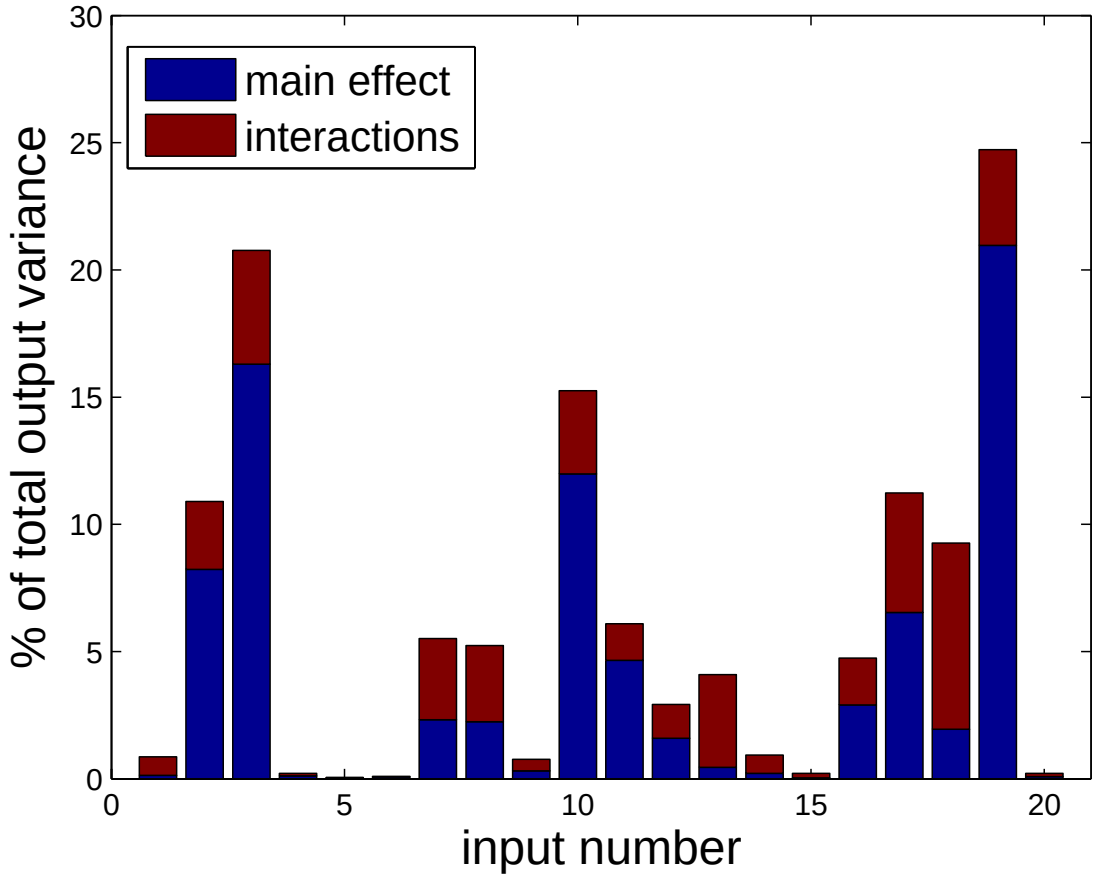
$$Var_{X_i}\{E_{\boldsymbol{X}_{-i}}(Y|X_i)\} = Var_{X_i}(0) = 0.$$

Total effect indices

$$T_i = \frac{E_{\boldsymbol{X}_{-i}}\{Var_{X_i}(Y|\boldsymbol{X}_{-i})\}}{Var(Y)}$$

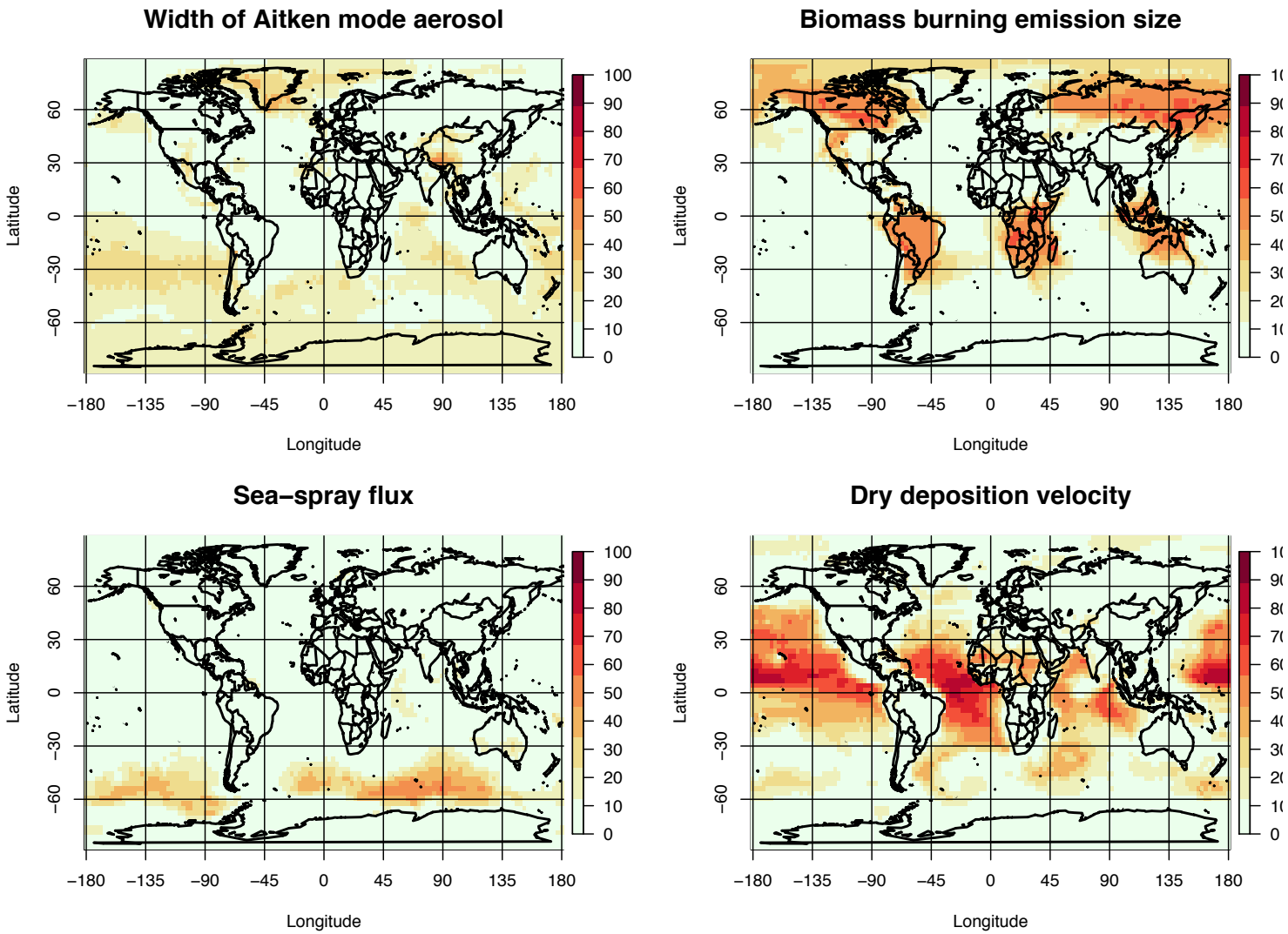
- Numerator is expected remaining variance if all inputs apart from X_i are learned
- Measures contribution to variance from X_i including interactions
- Small value *does* imply 'unimportant'
- But requires independence between $\{X_1, \ldots, X_d\}$

Example: SA for an infectious disease model



(Thanks to John Paul Gosling, Hugo Maruri-Aguilar, Alexis Boukouvalas)

SA for a global aerosol model: Lee et al. 2013



(Thanks to Lindsay Lee)

Main effect plots for exploratory analysis

- Measure importance of input X_i via

$$Var_{X_i}\{E(Y|X_i)\}$$

- Think of $E(Y|X_i)$ as a function of X_i: how does $E(Y|X_i)$ vary as X_i varies?
- Could plot $E(Y|X_i)$ against X_i to learn about input-output relationship

Main effect plot

Plot $E(Y|X_i) - E(Y)$ against X_i
$E(Y|X_i) - E(Y)$ referred to as main effect of X_i.

- Can plot $E(Y|X_i, X_j) - E(Y|X_i) - E(Y|X_j)$ against X_i and X_j to visualise interactions
- But need to watch for confounding of f with $p(\boldsymbol{X})$.

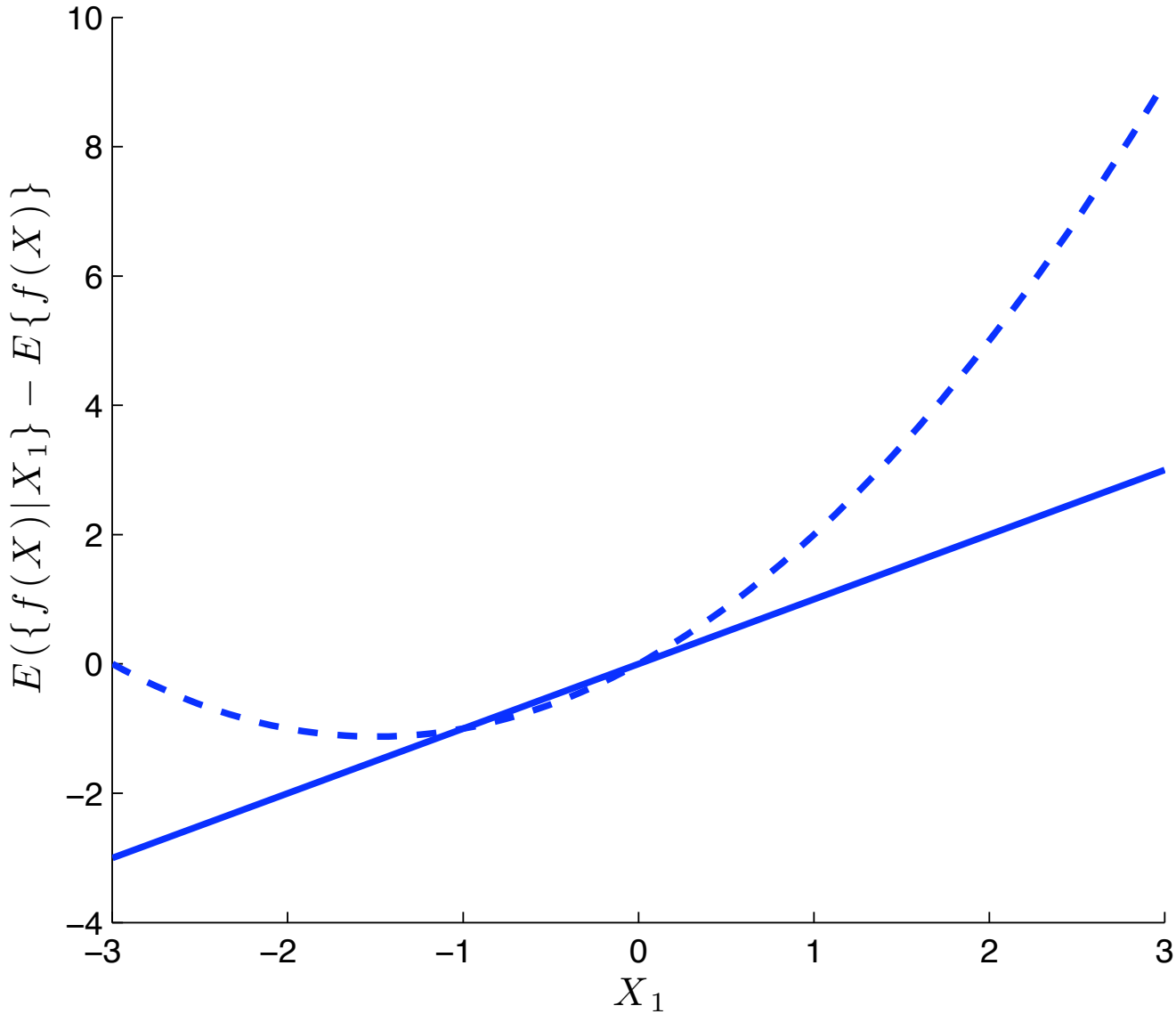
Example

- Consider $Y = X_1 + X_2 + X_1X_2$, with X_1, X_2 i.i.d $N(0,1)$.
- The main effect of X_1 is

$$E[Y|X_1] - E[Y] = X_1,$$

- Now suppose that we judge that X_1 and X_2 are have instead a bivariate normal distribution with covariance 0.5, and standard normal marginal distributions.
- The main effect of X_1 is now

$$E[Y|X_1] - E[Y] = 1.5X_1 + 0.5X_1^2,$$


10
8
6
4
2
0
−2
−4
−3
−2
−1
0
1
2
3
X_1
$E(\{f(X)|X_1\} - E\{f(X)\}$

Computation

- Typically requires evaluating $f(\boldsymbol{x})$ for large numbers of input values $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$
- For computationally expensive models, can speed things up with an emulator (Oakley and O'Hagan 2004)
- But suppose we have already done basic uncertainty propagation with Monte Carlo
  - we have a sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N$ from $p(\boldsymbol{X})$
  - we have evaluated $Y_1 = f(\boldsymbol{X}_1), \ldots, Y_N = f(\boldsymbol{X}_N)$
- For moderately large N (e.g. a few hundred), this gives us all we need to compute main effect indices (and low order interaction variances as well) (Strong and Oakley 2013)

Computation

We write

$$Var_{X_i}(E(Y|X_i)) = Var(g(X_i))$$

which we could estimate using

$$\frac{1}{N-1}\sum_{j=1}^{N}(g(X_{i,j}) - \bar{g})^2,$$

with

$$\bar{g} = \frac{1}{N}\sum_{j=1}^{N} g(X_{i,j})$$

- We already have the sample $X_{i,1}, \ldots, X_{i,N}$.
- We can use the same Monte Carlo sample to estimate $g(X_i)$

Example

- Model with three inputs

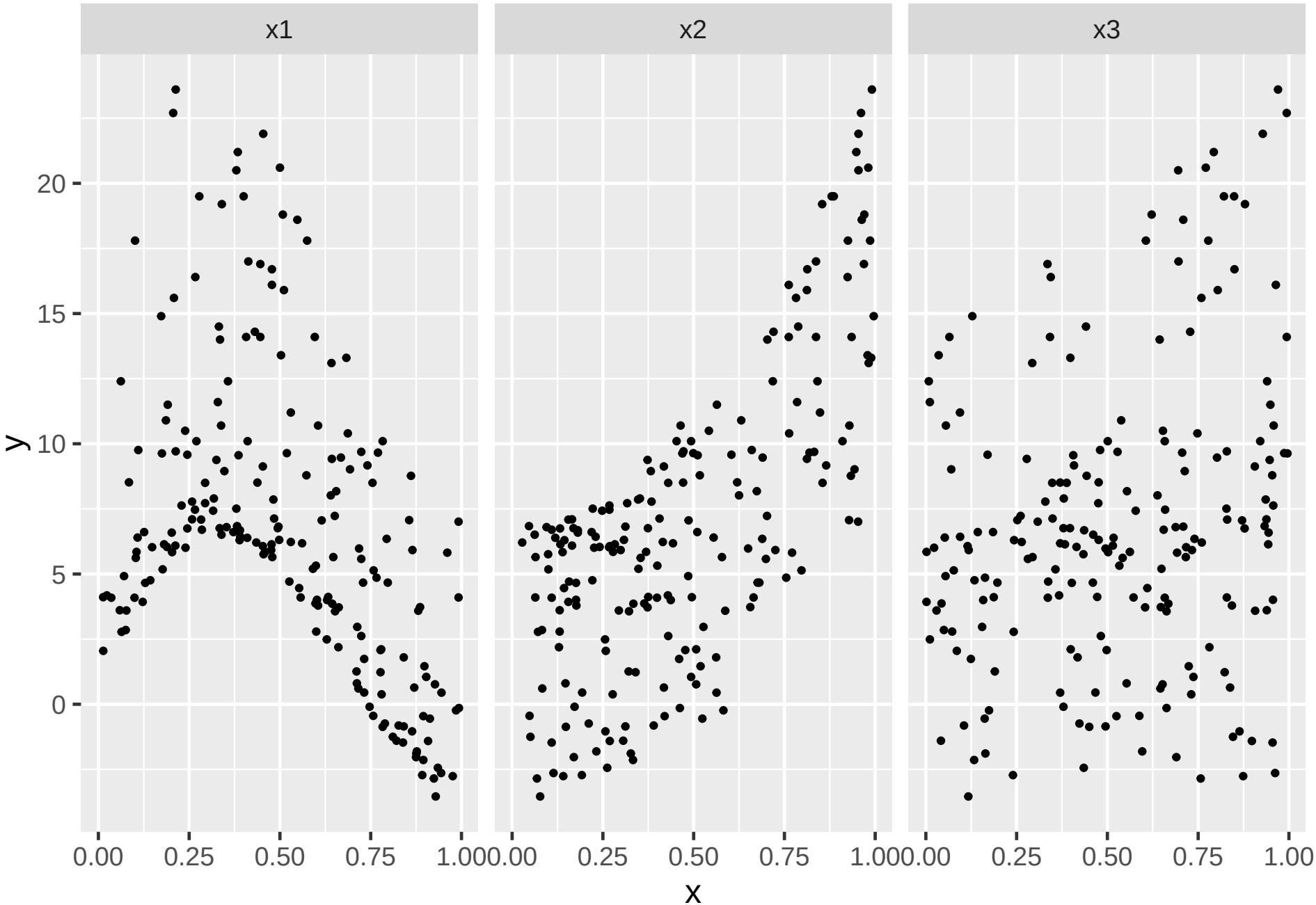
$$Y = f(X_1, X_2, X_3)$$

with $X_1, X_2, X_3 \overset{i.i.d}{\sim} U[0, 1]$

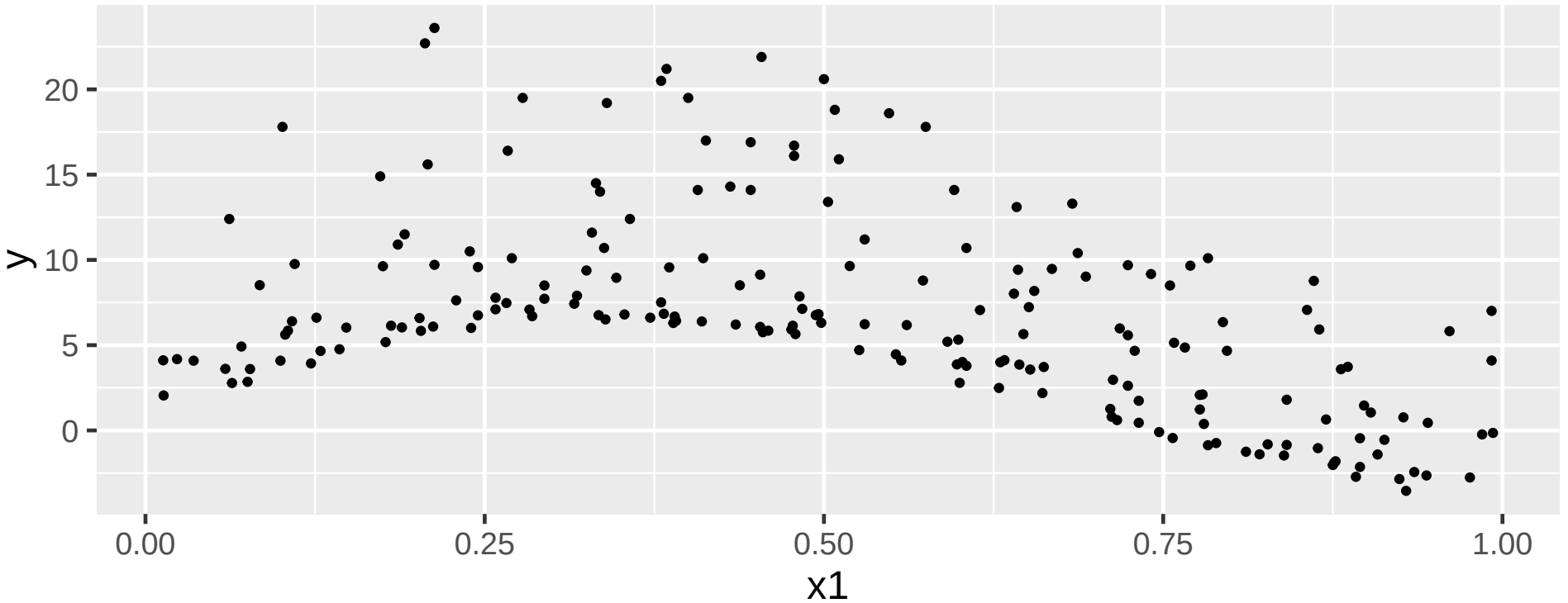
- Monte Carlo sample of model runs:

y	x1	x2	x3
7.47	0.266	0.268	0.659
6.61	0.372	0.219	0.185
8.79	0.573	0.517	0.954
-1.41	0.908	0.269	0.898
6.59	0.202	0.181	0.944
1.46	0.898	0.519	0.724
⋮	⋮	⋮	⋮

Plot the output against each input



Each plot contains the data needed to estimate $g(X_i) = E(Y|X_i)$:

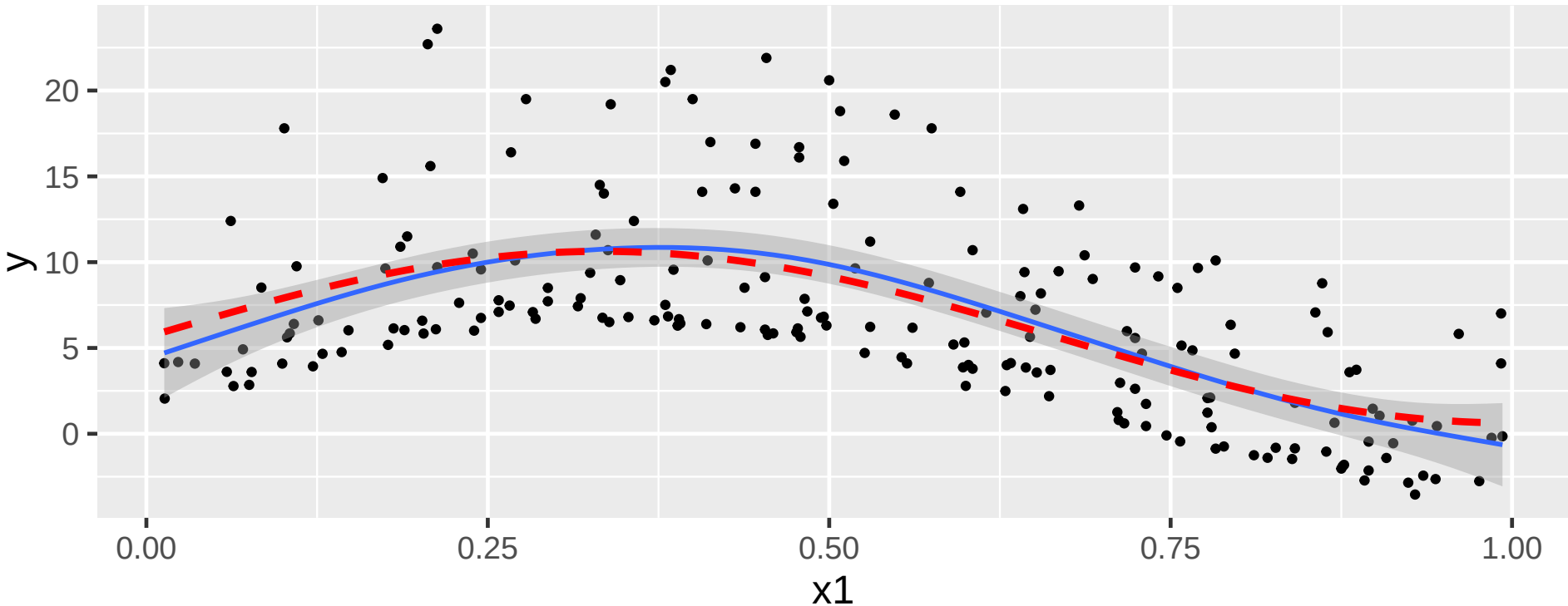


- We write

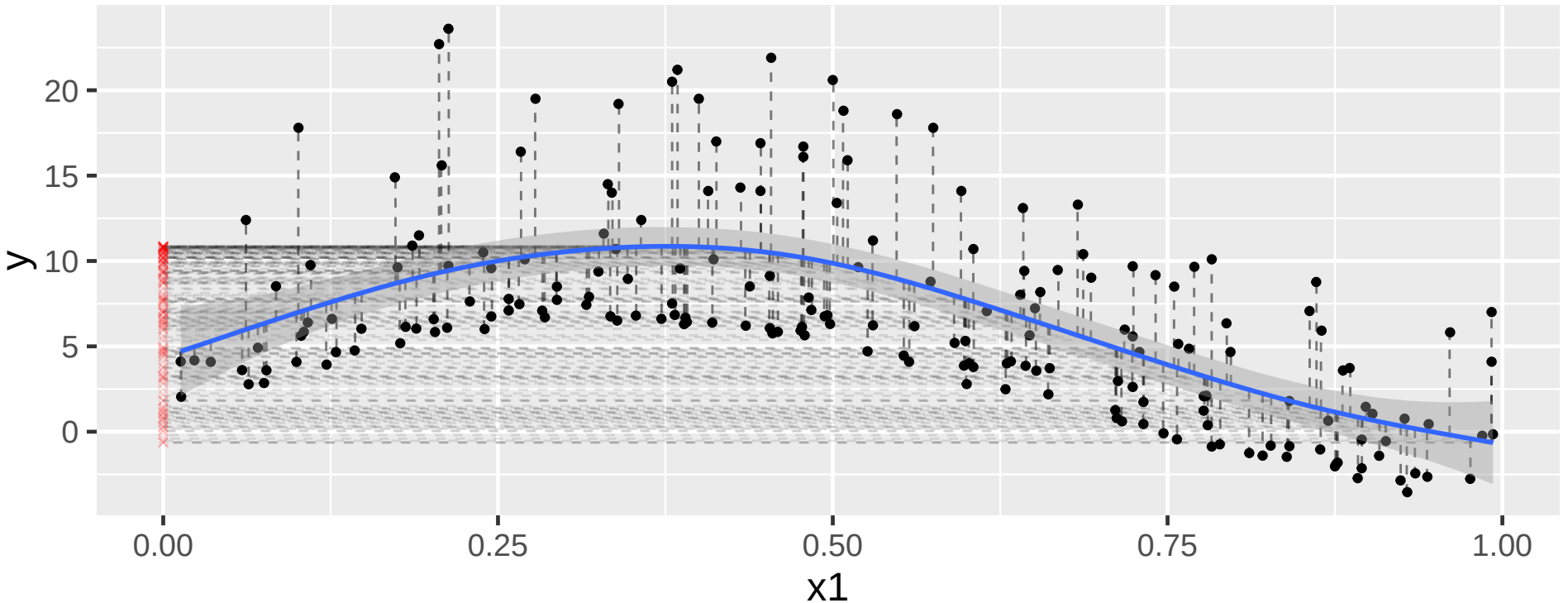
$$Y_j = f(X_{1,j}, X_{2,j}, X_{3,j}) = g(X_{1,j}) + \varepsilon_j$$

- ε_j has expectation zero (though variance may not be constant)

Estimate $g(X_1) = E(Y|X_1)$ with nonparametric regression (e.g. generalised additive models):



Variance of fitted values gives estimate of $Var_{X_1}(E(Y|X_1))$



Two lines of R code!

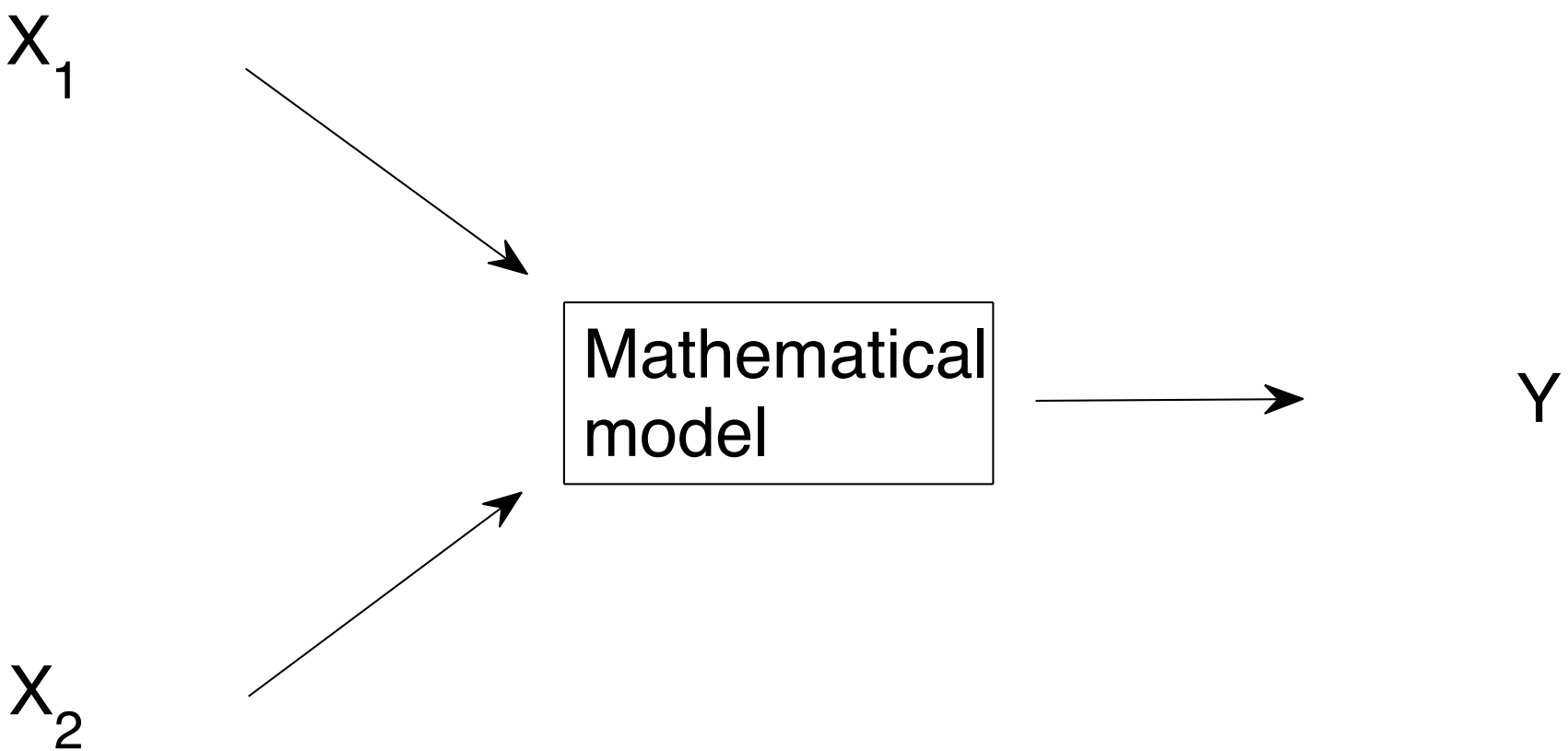
```
gam1 <- mgcv::gam(modelruns$y ~ s(modelruns$x1))
100 * var(fitted(gam1)) / var(modelruns$y)
```

```
## [1] 39.83785
```

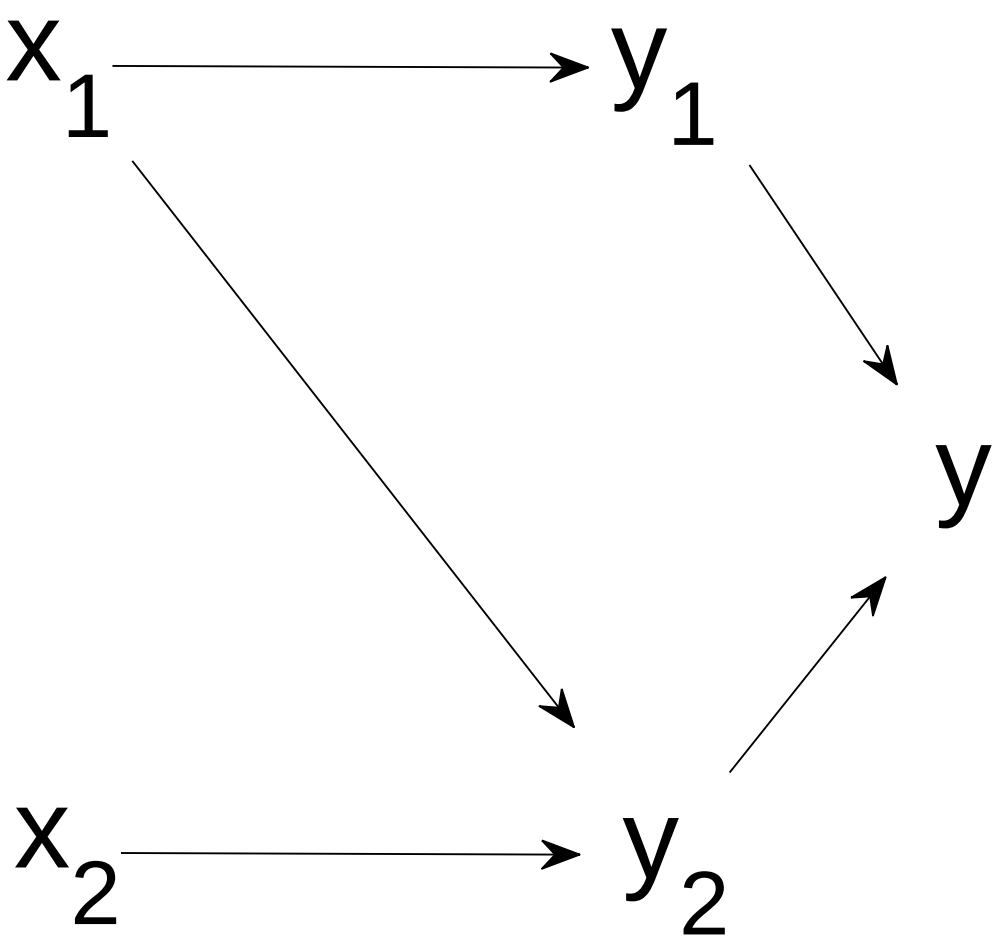
Extension: sensitivity analysis for model discrepancy

- Model $y = f(x)$, with uncertain true inputs X, and $Y = f(X)$.
- Target quantity Z, to be estimated using model.
- If model is not perfect then $Y = f(X) \neq Z$
- Potentially, different sources of error within a model structure: are they all equally important?

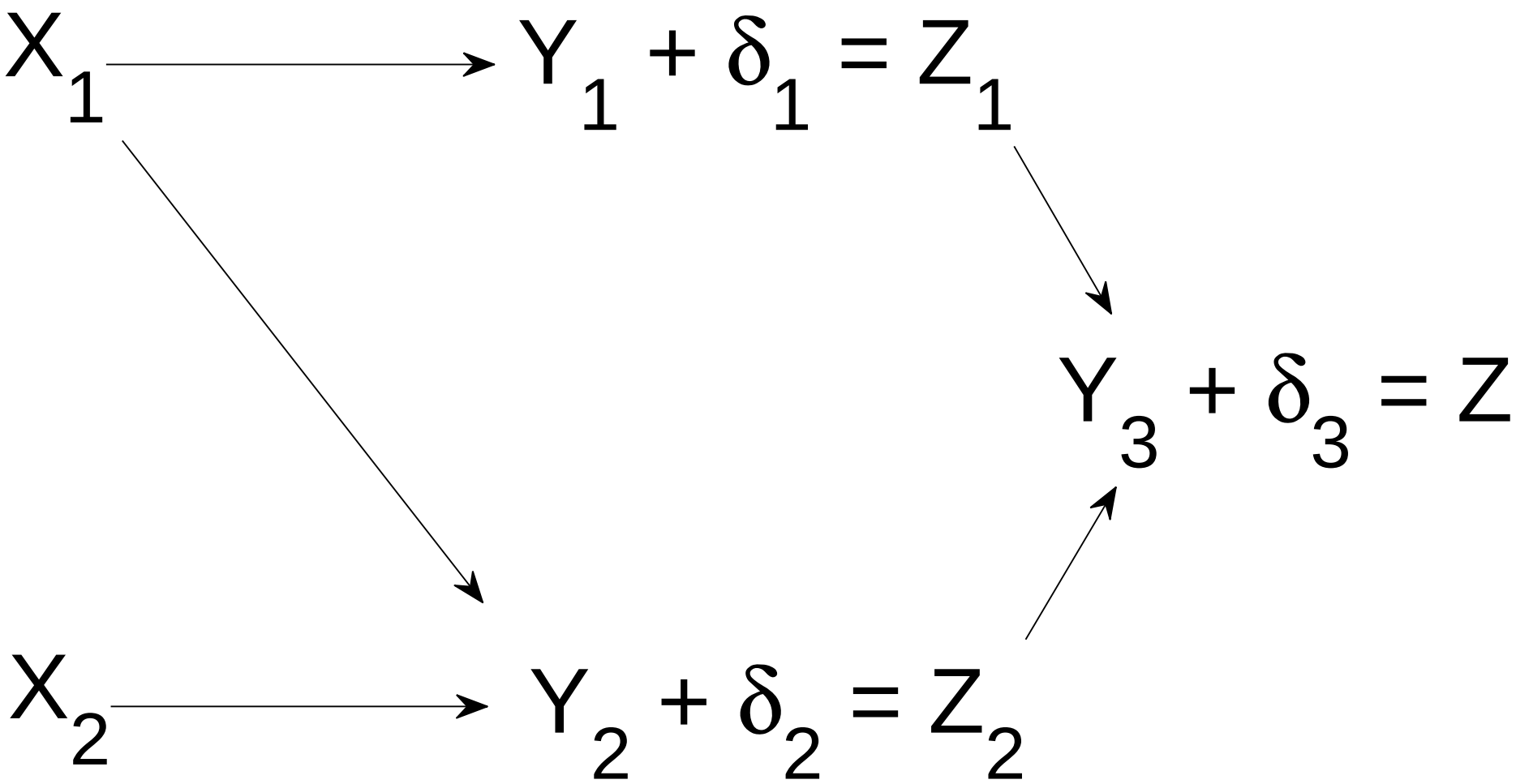
The model



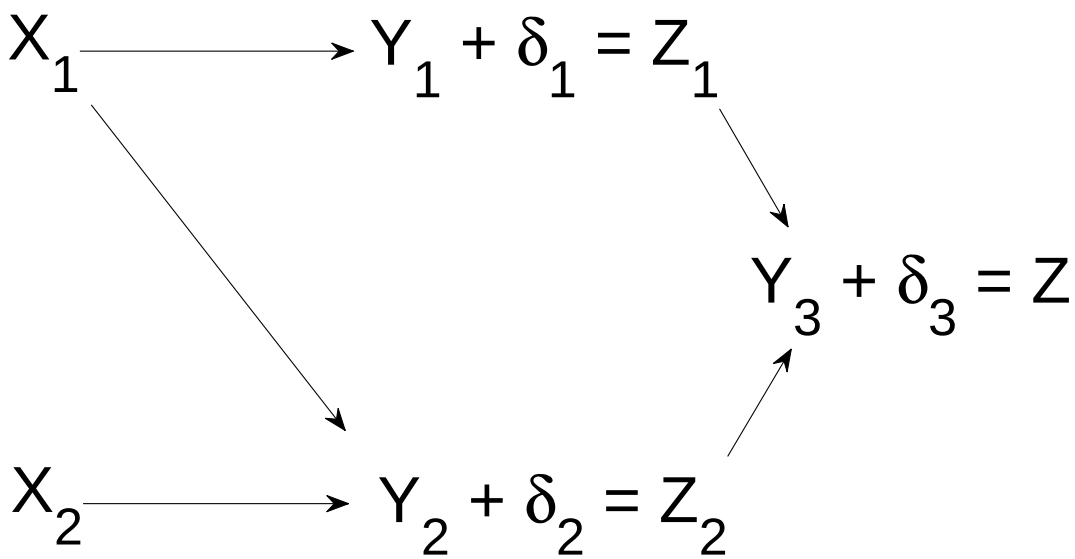
Opening the black box



Linking the model to reality



Using sensitivity analysis



- Specify joint distribution for $X_1, X_2, \delta_1, \delta_2, \delta_3$
- Use variance-based sensitivity analysis to investigate how learning δ_i would reduce variance of Z

$$\frac{Var_{\delta_i}\{E(Z|\delta_i)\}}{Var(Z)}$$

References

- Oakley, J. and O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society Series B, 66, 751-769.
- Saltelli, A. et al. (2008). Global Sensitivity Analysis: The Primer, New York: Wiley.
- Strong, M., Oakley J. E. and Chilcott, J. (2012). Managing structural uncertainty in health economic decision models: a discrepancy approach. Journal of the Royal Statistical Society, Series C, 61(1), 25-45.
- Lee, L. A. et al. (2013) The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei, Atmospheric Chemistry and Physics, 13, pp.8879-8914.
- Strong, M. and Oakley, J. E. (2013). An efficient method for computing single parameter partial expected value of perfect information. Medical Decision Making, 33, 755-766.