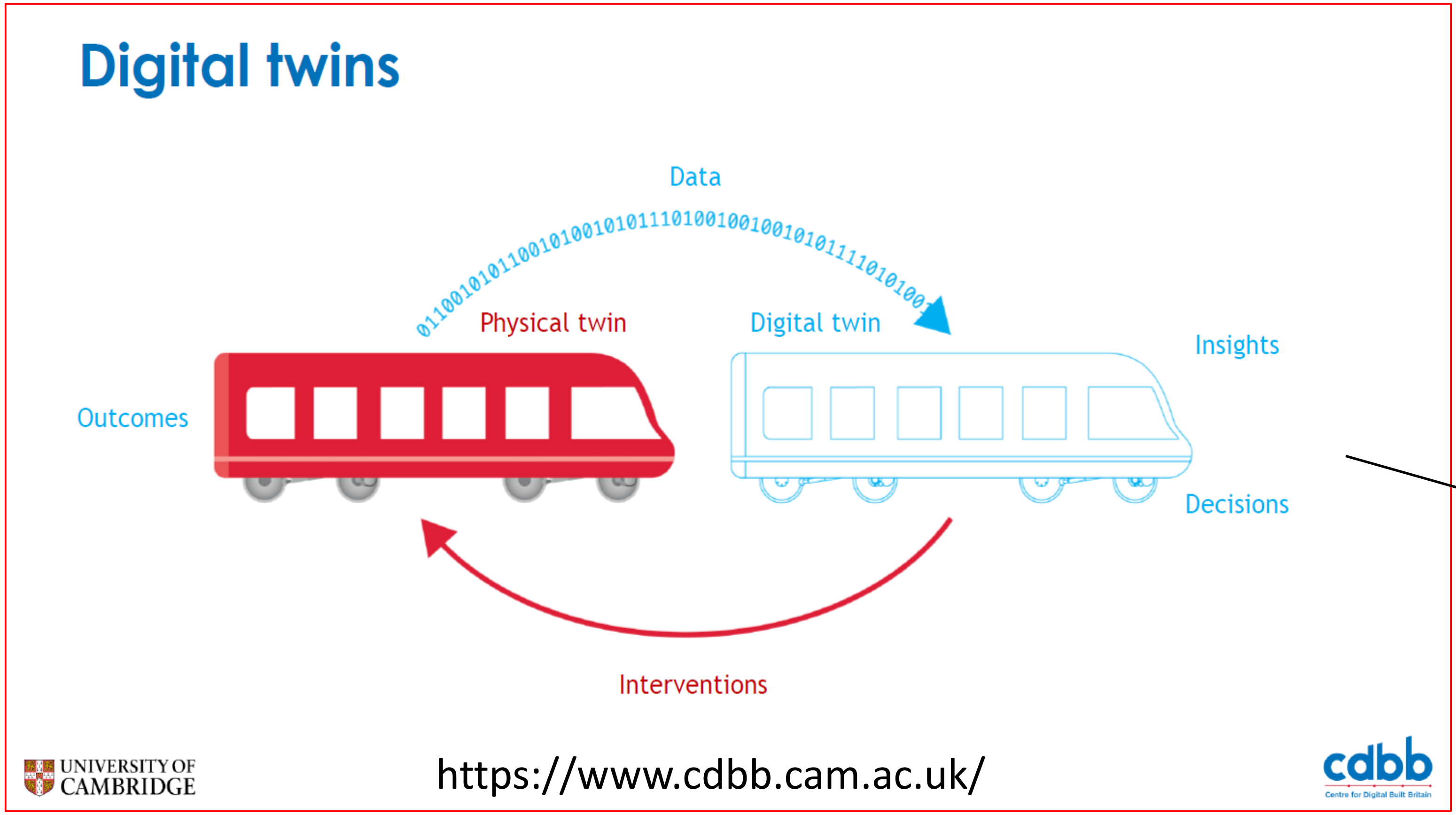
Toolbox for Engineering Design Sensitivity - TEDS

Jiannan Yang

jy419@cam.ac.uk

Dec 2020

Let’s start from Digital Twins …

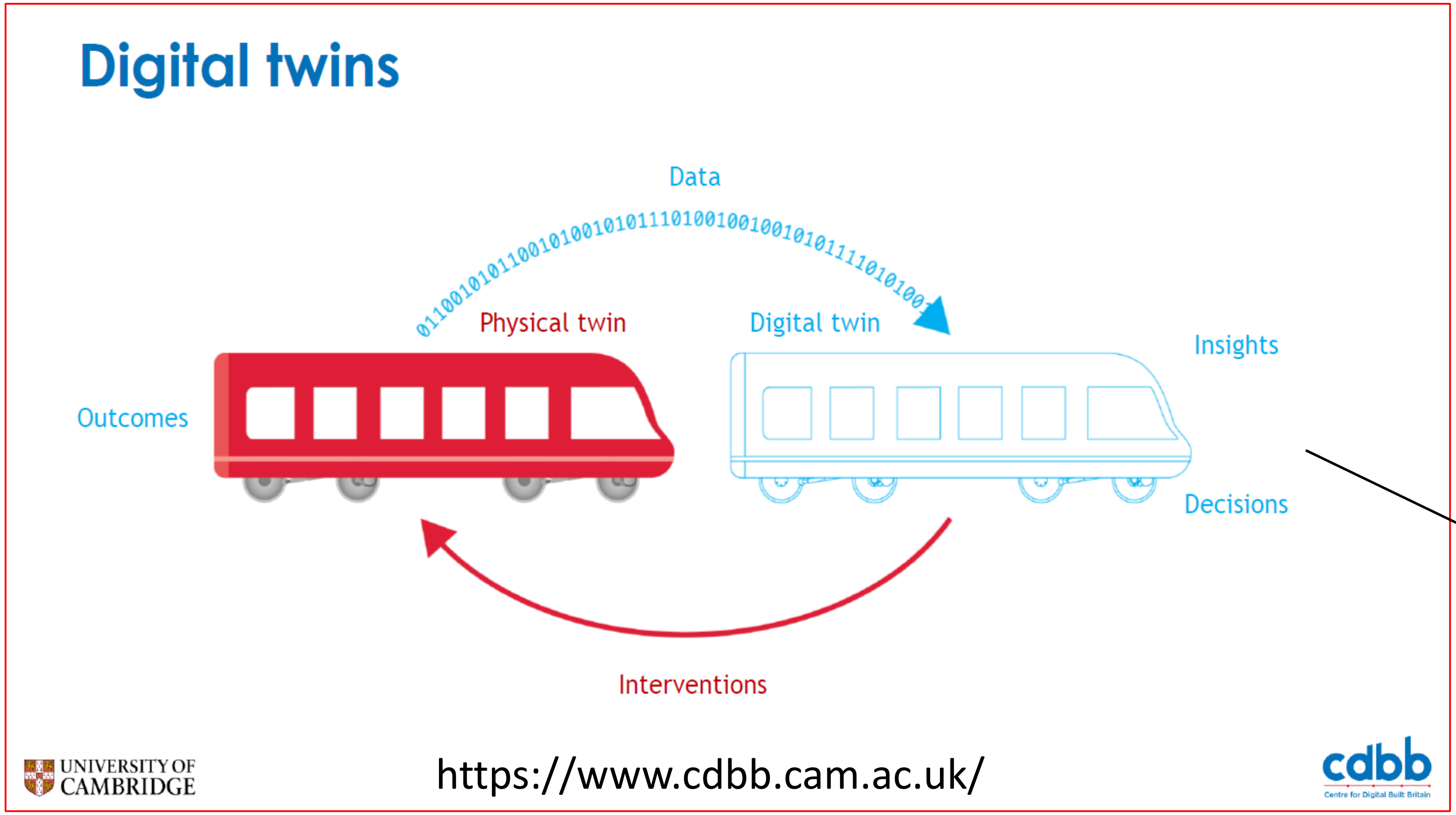


The definitions of digital twins are not unique. Here is one description of digital twins from Centre for Digital Build Britain. No matter what is the definition, one thing we agree is that DTs provide insights so that informed decisions can be made. But how to extract most useful information from digital twins? DTs are being built across different sectors, our focus is to develop **metric toolbox for digital twins** to support decision making

Metric Toolbox

Decisions

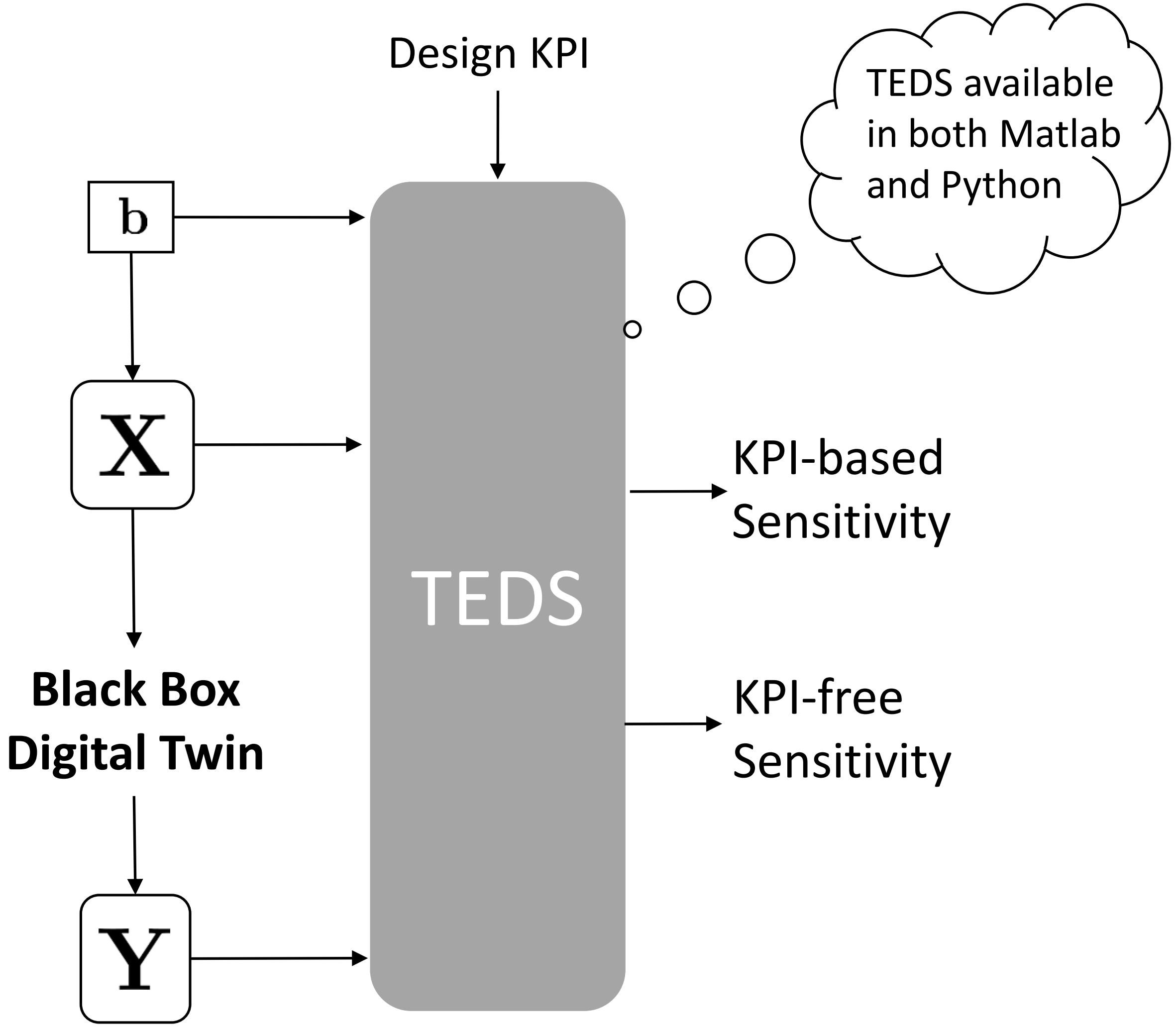
Toolbox for engineering design sensitivity (TEDS)



In this presentation, I discuss in details a sensitivity toolbox (TEDS) to identify most important parameters in design

TEDS → Design Sensitivity

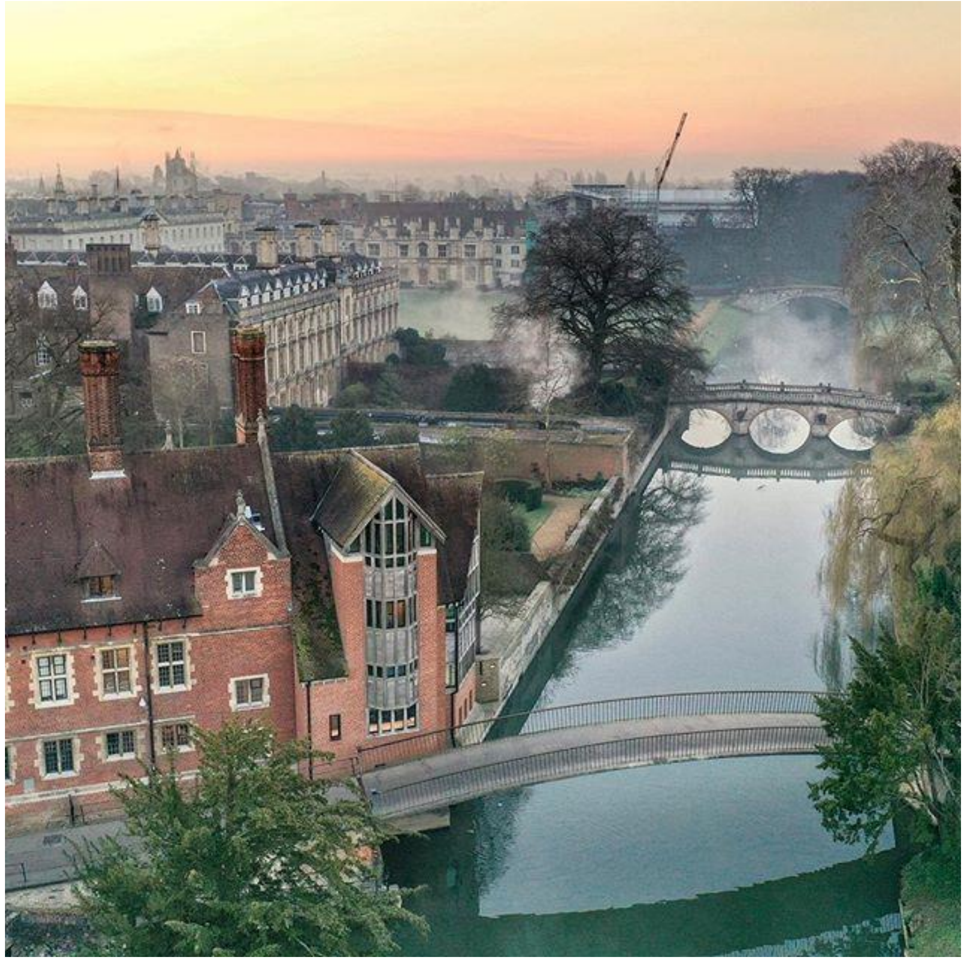
Toolbox for engineering design sensitivity (TEDS)



KPI: Key performance indicator

- A quick overview of the key features of TEDS
  - identifies most influential random design variables
  - non-intrusive toolbox because it wraps around black box digital twins and computational simulations
  - provides both KPI-based and KPI-free sensitivities

Overview of the contents



Some background of the project

https://digitwin.ac.uk/

DigiTwin consortium overview

The aim is to create a robustly-validated virtual prediction tool called a "digital twin".

Starting in Feb 2018 **Funded by EPSRC:** Engineering and Physical Sciences Research Council

Duration 5 years **Value £4.9M**

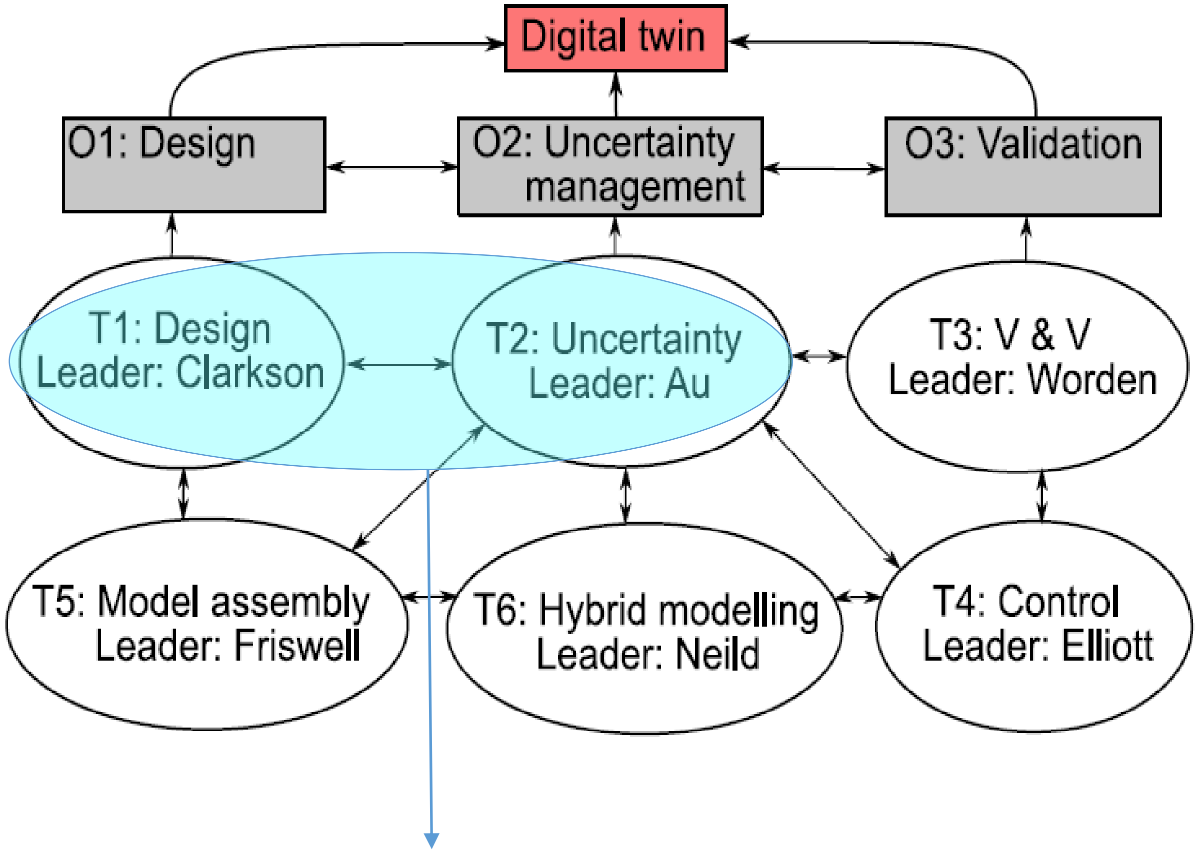
A consortium of 6 UK Universities and 10 Industry Partners

University of Bristol, University of Cambridge, The University of Sheffield, University of Liverpool, University of Southampton Institute of Sound and Vibration Research, Swansea University Prifysgol Abertawe

Airbus, EDF Energy, Leonardo, LOC Marine & Engineering Consultants, Romax Technology, Schlumberger, Siemens Gamesa Renewable Energy, Stirling Dynamics, Ultra Electronics, Siemens

Engineering and Physical Sciences Research Council

DigiTwin
Digital twins for improved dynamic design



At Cambridge, we look at design in the presence of uncertainties

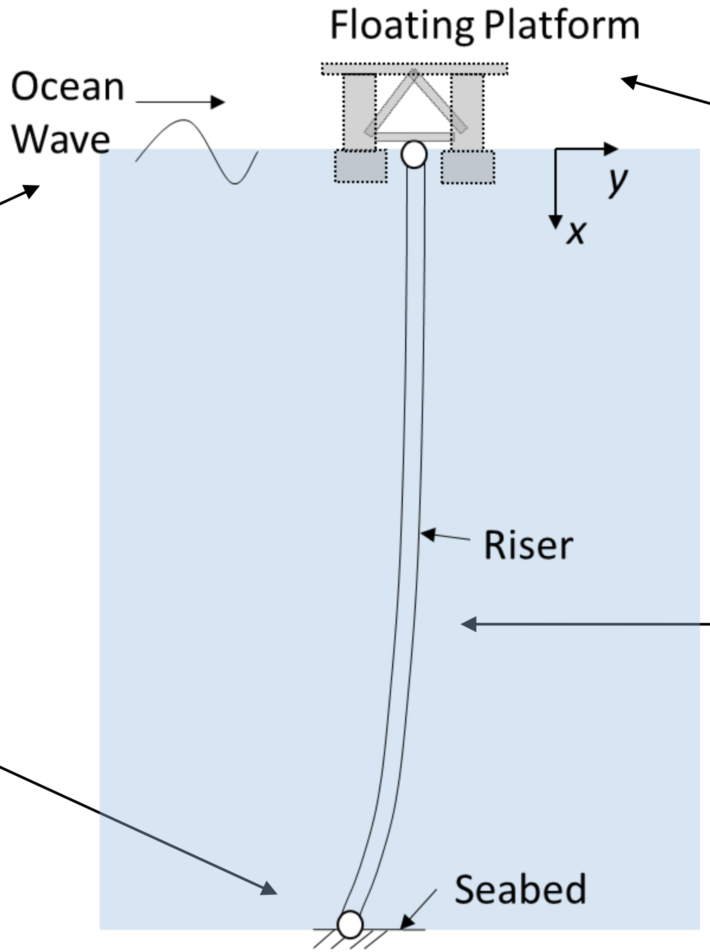
Design in the presence of uncertainties

Design in the presence of uncertainties is a really broad area. Although the proposed methodology can be potentially applied to different sectors, let’s use an example to have a more concrete idea of what we are trying to do here. We have chosen offshore structure design as our examples because uncertainty considerations are important in the industrial design process of offshore structures.

Design in the presence of uncertainties

<u>Uncertain</u> ocean wave heights which can be stochastic by nature, and <u>uncertain</u> structure/fluid interaction coefficients due to a lack of complete underlying physics

<u>Uncertain</u> soil conditions at the sea bed



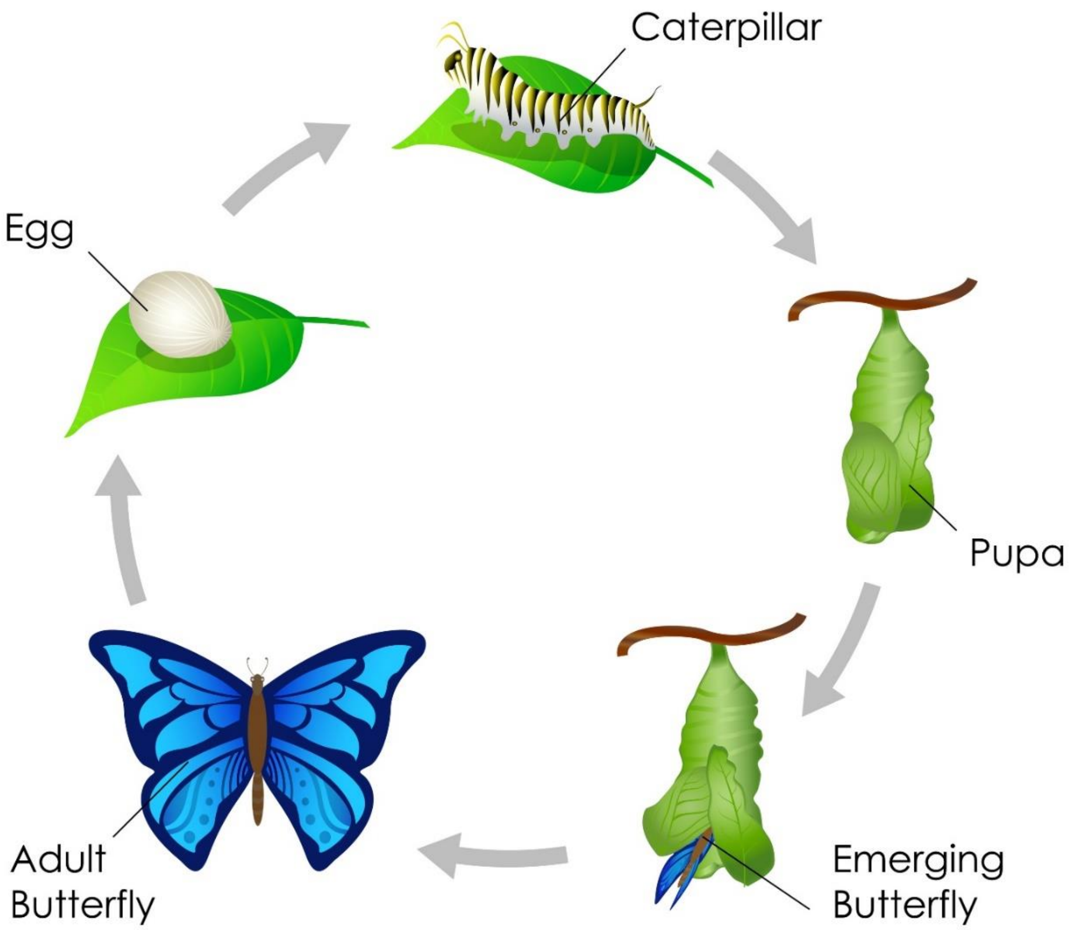
<u>Uncertain</u> force exerted by the floating platform onto the riser, depending on operation condition

<u>Uncertain</u> material property, probably due to manufacturing

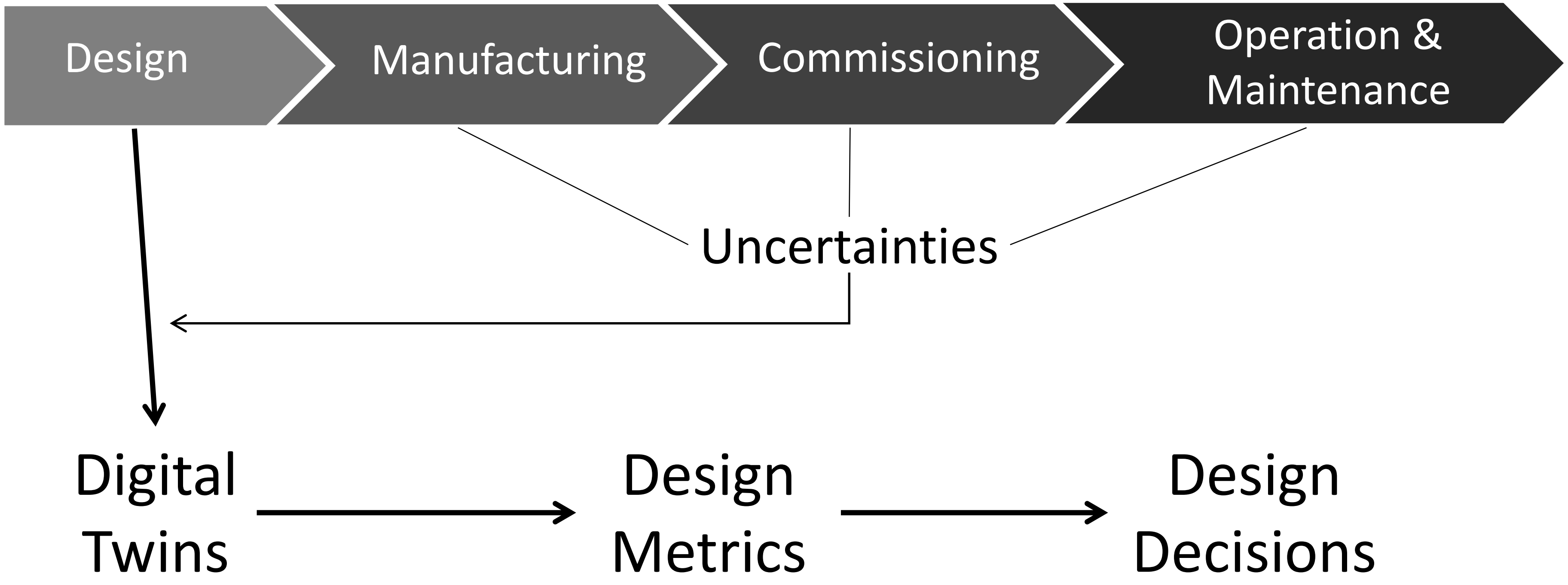
A marine riser is a conduit that extracts subsea oil to a surface platform. This example with a marine riser highlights the ubiquitous role of uncertainties for engineering design. I will use the marine riser as an example throughout the presentation (details of the dynamic model https://github.com/longitude-jyang/hydro-suite)

Design in the presence of uncertainties

Where are the uncertainties in the whole lifecycle?



Design in the presence of uncertainties

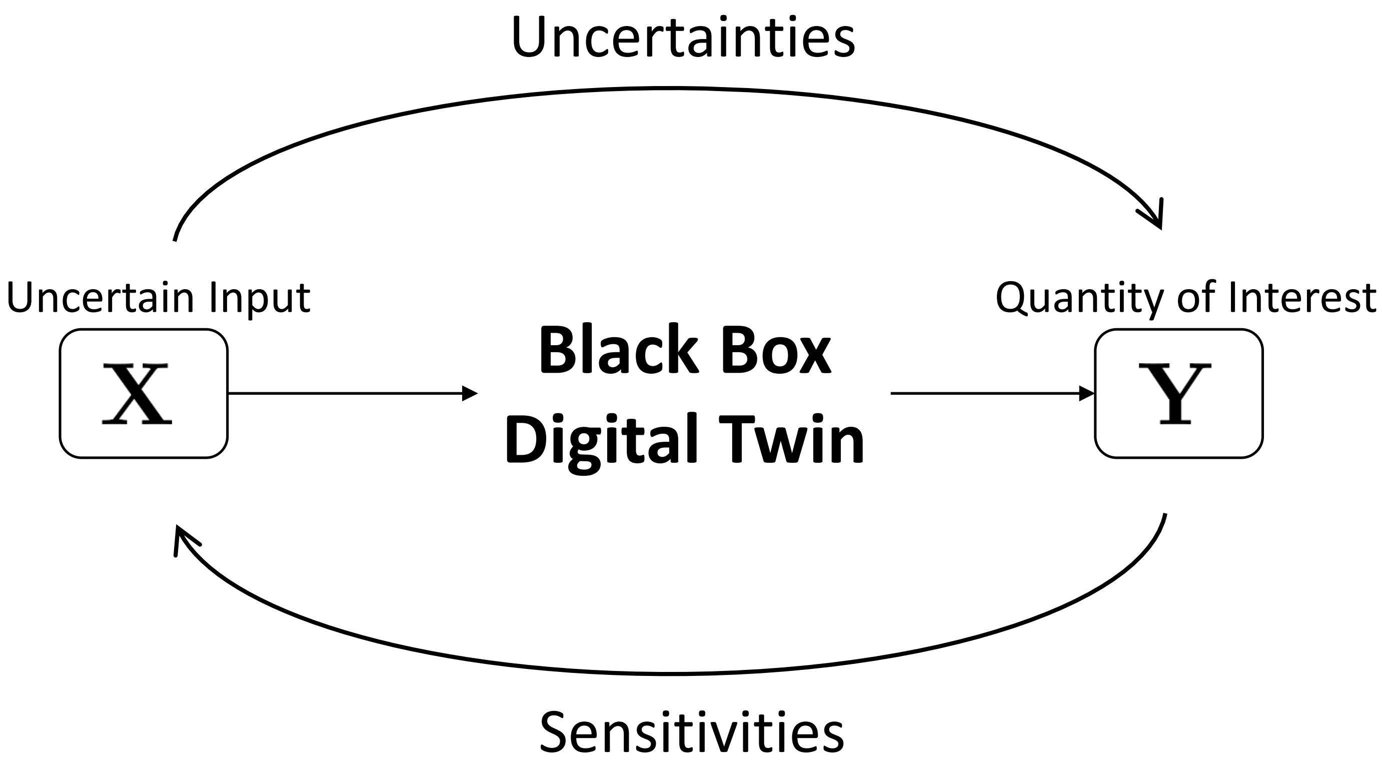


From the riser example, we could generalise the view a bit and look at the engineering process more abstractly. Taking account of uncertainties in the design process allows us to work towards ‘life cycle design’. Digital twins (robustly validated computational models) are increasingly used in the design. However, to make informed design decisions, the key is to provide suitable metrics of the design.

Design metric - sensitivity to uncertainties

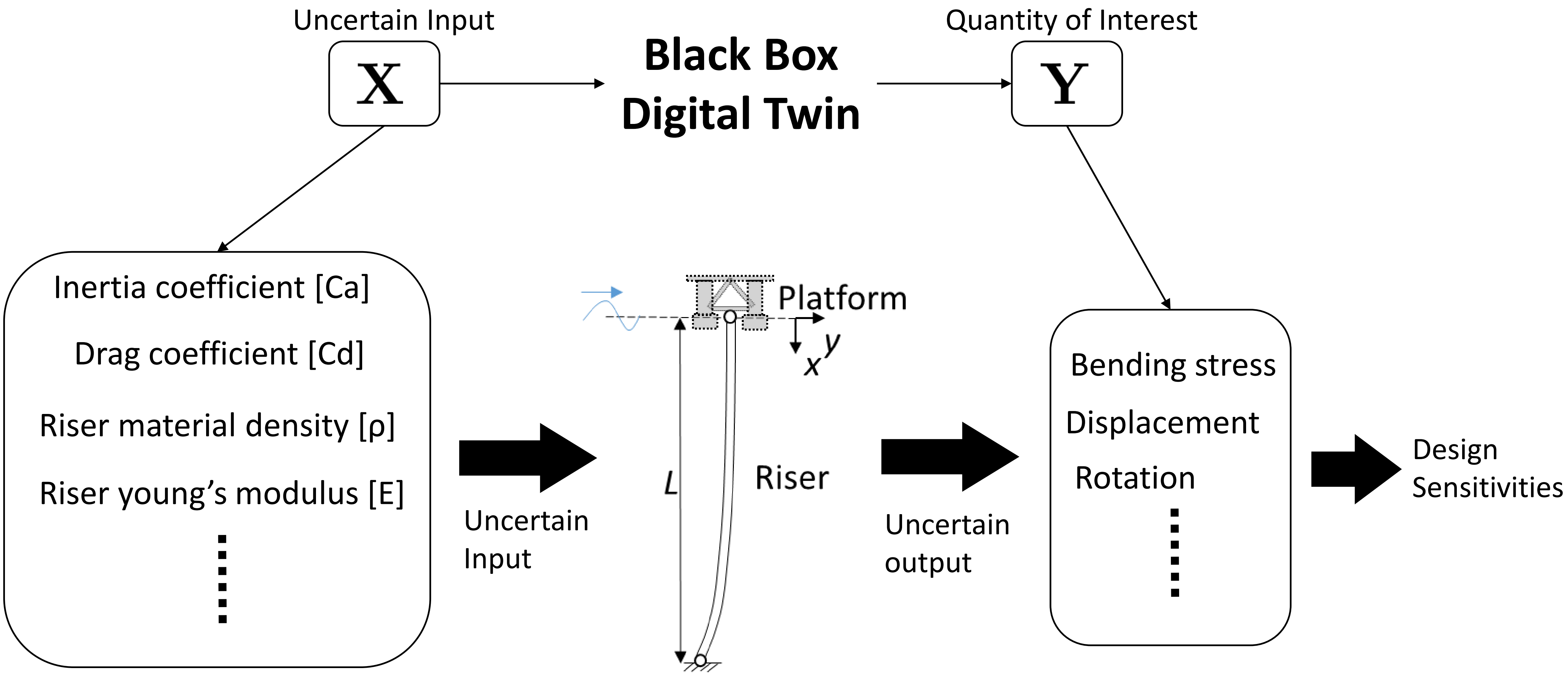
One of the most important design metric is the sensitivity to uncertainties

Design metric - sensitivity to uncertainties



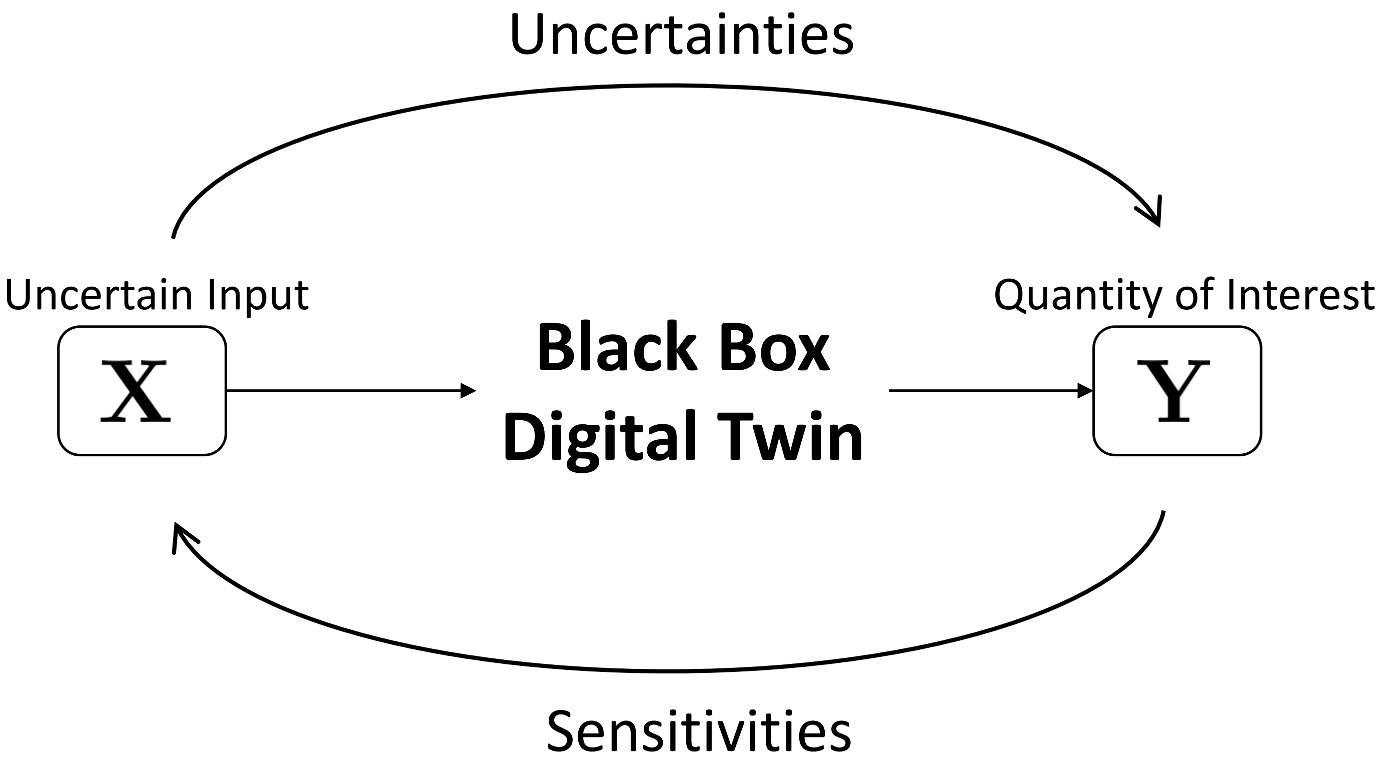
In the presence of uncertainty, the 1st step is to propagate the uncertainties through the digital twin and quantify the resulted uncertainties in the design output. This is often done using Monte Carlo methods. Once uncertainties are quantified in the design, it is desirable to conduct sensitivity analysis to understand the relative importance of the different sources of uncertainties.

Uncertainty and sensitivity analysis



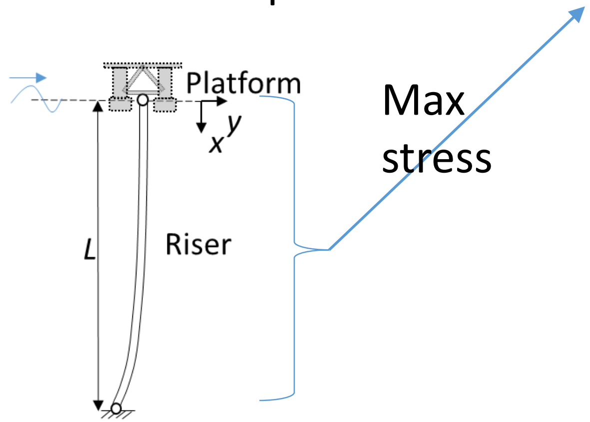
Uncertainty and sensitivity analysis

If Monte Carlo methods are used for uncertainty analysis where a large number of samples are generated , we can produce **scatter plots** between the uncertain inputs and outputs from the simulated samples. And this is often the most straightforward and effective way to look at sensitivities!

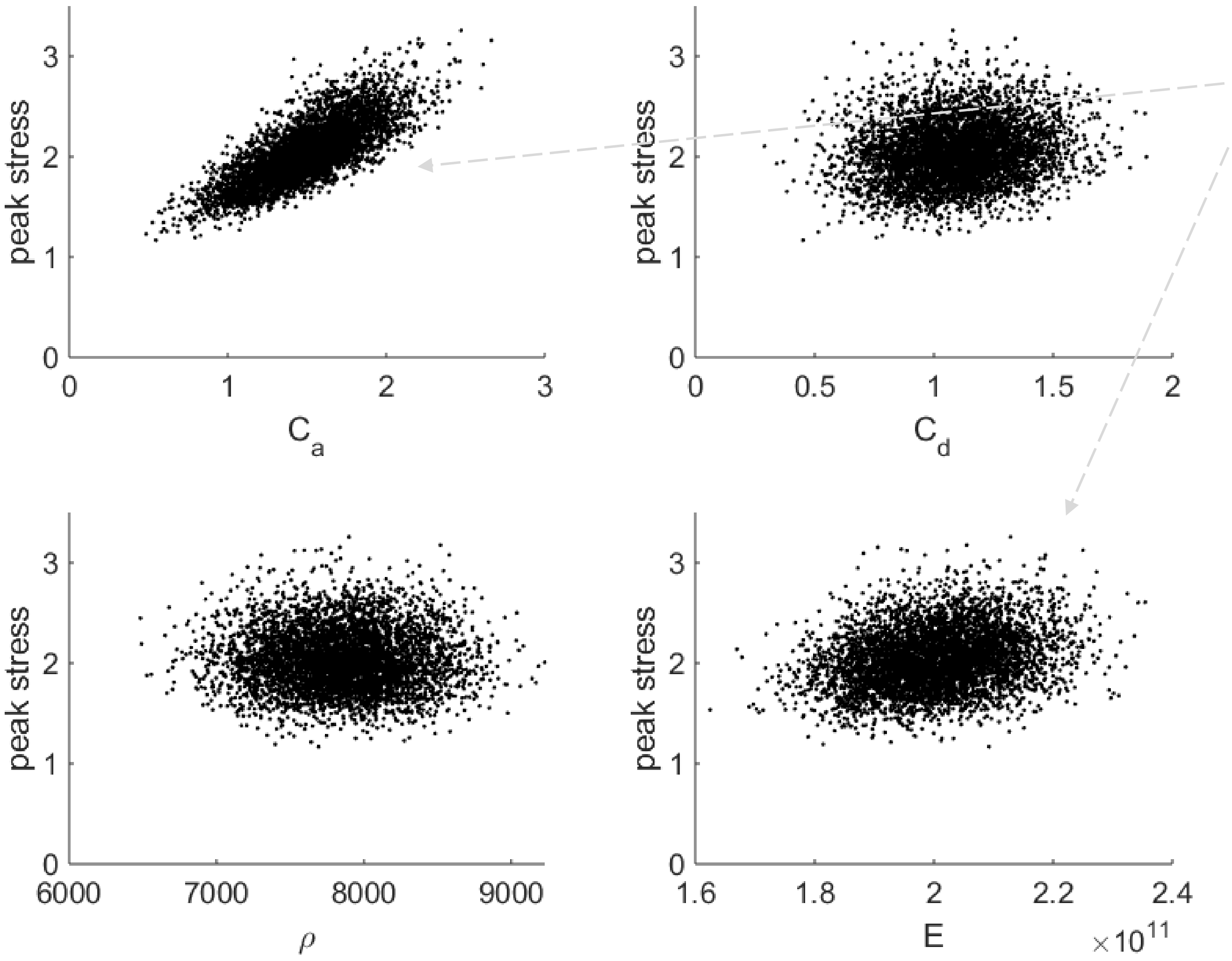


Uncertainty and sensitivity analysis

Take an example of scatter plot here for the max bending stress along the riser, against a few of the uncertain inputs.



The random bending stress is position dependent. It is of engineering interest to look at the **max stress** along the riser and how does it relate to the uncertain inputs.



There seems a *strong* dependency of the max stress along the riser on Ca and *weak* dependency on E

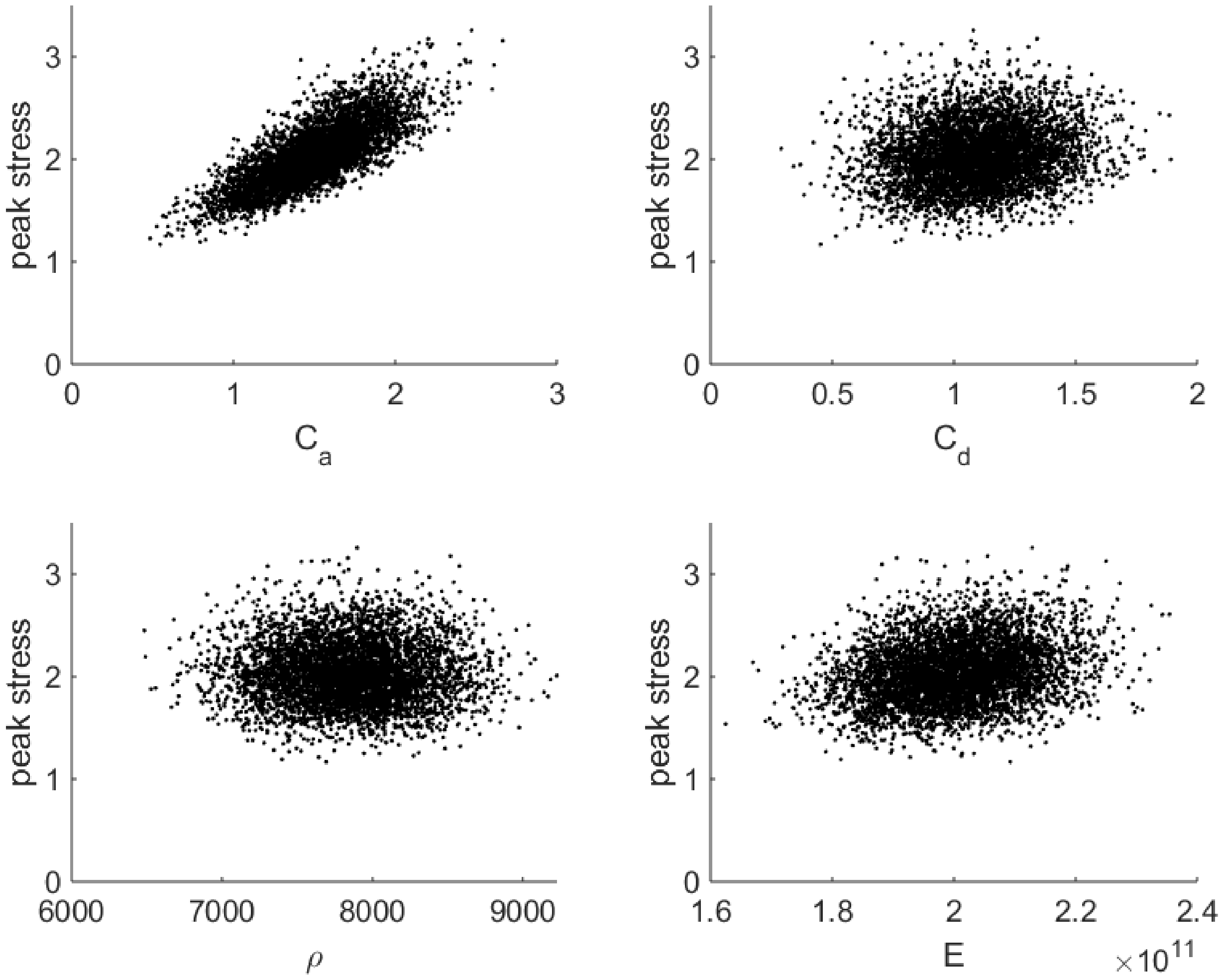
However, take a closer look, the range of Ca has more than doubled, while for E, there is only about 20% variable of its values. So which one is more important?

Uncertainty and sensitivity analysis

From this example, we see that if the scatter plots are available , it is worth to have a look at it.

The scatter plots provide rich information about the relationship between inputs and outputs!

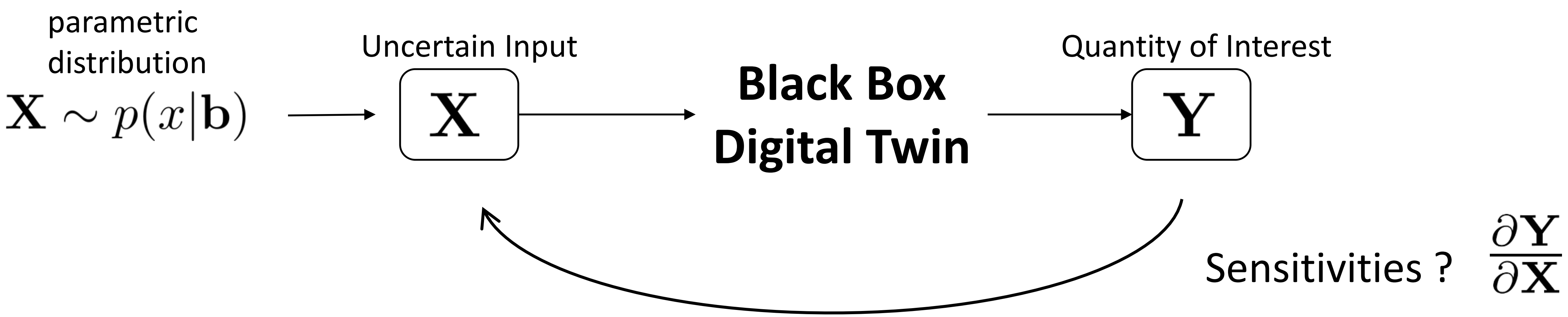
However, it can be overwhelming for large number of variables. Even for a few variables, as shown in the example here, a lot of analyses (e.g. regression analysis) are needed!



Uncertainty and sensitivity analysis

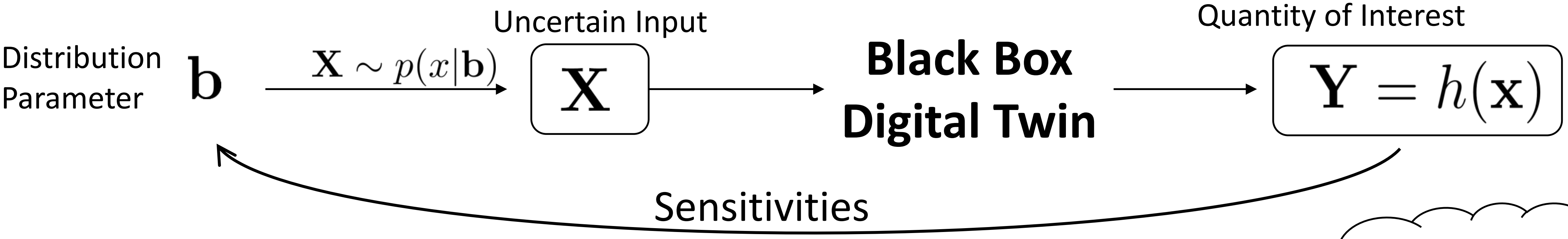
In design, it would be quite useful to have metrics that provide more compact information for designers!

Parametric sensitivity



Let's take a minute and think about what is the sensitivity to the random inputs? The most natural way is to compute the derivative of y to x. However, the function of interest is often not differentiable, not to mention black box models! To overcome this issue, let's make an assumption that the input uncertain variables can be described by parametric distribution models, in other words, we can use mean and variance to describe entire distribution if the underlying distribution is assumed to be Gaussian. This is not an unreasonable assumption because most statistical methods are parametric. When you collect more information about an uncertain variable, for example by doing some experiments, it is the distribution parameters like mean and variance that will change.

Parametric sensitivity



$\mathbb{E}$ represents mathematical average

So why not look at directly at the sensitivity of the output to distribution parameters of the uncertain variables? It turns out that taking this view, there is a numerical trick called *Likelihood Ratio* method that will make things much easier, because the gradient information is obtained basically free (in terms of computational cost). Also, distribution functions are more likely to be differentiable.

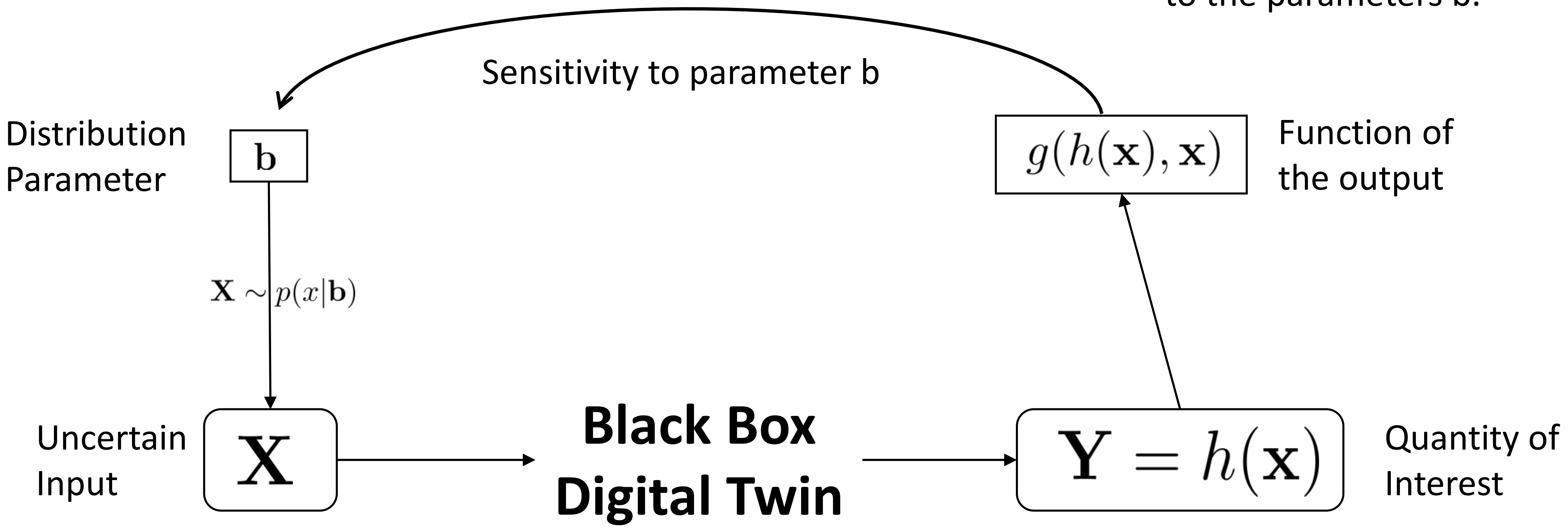
$$\mathbb{E}_X[h(\mathbf{b})] = \int h(\mathbf{x})p(\mathbf{x}|\mathbf{b})d\mathbf{x}$$

Free of charge to get gradient! Because this term is often available analytically

$$\frac{\partial \mathbb{E}_X[h(\mathbf{b})]}{\partial \mathbf{b}} = \int h \frac{\partial p(\mathbf{x})}{\partial \mathbf{b}} d\mathbf{x} = \mathbb{E}_X\left[h \frac{\partial \ln p(\mathbf{x})}{\partial \mathbf{b}}\right]$$

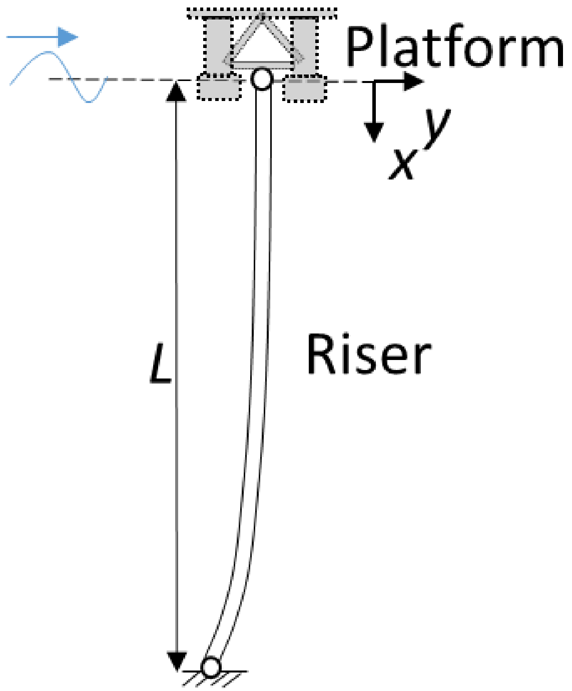
Parametric sensitivity

More generally, we can look at sensitivities of any function of the random design outputs to the parameters b.



$$L(\mathbf{b}) = \mathbb{E}_X[g(\mathbf{b})] = \int g(h(\mathbf{x}), \mathbf{x}) p(\mathbf{x}|\mathbf{b}) d\mathbf{x} \qquad \frac{\partial L}{\partial \mathbf{b}} = \int g \frac{\partial p(\mathbf{x})}{\partial \mathbf{b}} d\mathbf{x} = \mathbb{E}_X \left[g \frac{\partial \ln p(\mathbf{x})}{\partial \mathbf{b}} \right]$$

Sensitivity in the design context



Concept design → Detailed design

KPI-free
How is the distribution of the random bending stress affected by input uncertainties?

$$p(\mathbf{y}|\mathbf{b}) \text{ and } \frac{\partial p(\mathbf{y}|\mathbf{b})}{\partial \mathbf{b}}$$

KPI-based:
Probability of failure with 80 years design life and how sensitive of it to uncertainties?

$$P_f = \Pr(\mathbf{y} > \text{threshold}) \text{ and } \frac{\partial P_f}{\partial \mathbf{b}}$$

Now that we have found the appropriate sensitivity that can be easily computed, we can implement that for design. So first we take a look at what kind of sensitivity information are needed in design? In the design context, there are two different types of scenarios where we would be interested to understand the sensitivity. For existing design or final stages of the design process, we normally have a specified design target or KPI. What we are interested is to quantify the probability that the design would fail to meet the KPI, failure probability Pf(y>y0). On the other hand, for new designs or at early design stage, a clear design target is not normally fixed. What we are interested here is the general distribution of the designed response.

Mathematical framework for sensitivity

KPI-based sensitivity

$$\boxed{\mathbf{b}} \xrightarrow{\mathbf{X} \sim p(x|\mathbf{b})} \boxed{\mathbf{X}} \longrightarrow \boxed{\mathbf{Y} = \mathbf{h}(\mathbf{X})} \longrightarrow P_f = \Pr(\mathbf{y} > \text{threshold})$$

Given the response, what is the probability that the design will not satisfy the specified KP ? In other words, what is the probability of failure for the given KPI?

KPI Sensitivity $\boxed{\dfrac{\partial P_{\mathrm{f}}}{\partial \mathbf{b}}}$

A row vector with j^{th} element the partial derivative with respect to parameter b_j $\dfrac{\partial P_{\mathrm{f}}}{\partial b_j}$

Normalize

$\boxed{\mathbf{r}}$

Sensitivity is the normalised gradient vector, which is unitless so that we can compare different parameters, even their values are of many orders of magnitude different (e.g. E is in the order of 10^9, and Ca is in the order of 10^0)

$$r_j = \frac{\partial P_{\mathrm{f}}}{\partial b_j} \frac{b_j}{P_{\mathrm{f}}} \approx \frac{\Delta P_f}{\Delta b_j}$$

Mathematical framework for sensitivity

KPI-free sensitivity

$$\boxed{\mathbf{b}} \xrightarrow{\mathbf{X} \sim p(x|\mathbf{b})} \boxed{\mathbf{X}} \longrightarrow \boxed{\mathbf{Y} = \mathbf{h}(\mathbf{X})} \longrightarrow p(\mathbf{y}|\mathbf{b})$$

KPI-free Sensitivity $\dfrac{\partial p(\mathbf{y}|\mathbf{b})}{\partial \mathbf{b}}$

Normalize

$$\frac{\partial p(y)}{\partial b}\frac{b}{p(y)}$$

This is a random variable because *y* is random

mean

$$\mathbb{E}_Y\left[\frac{\partial p}{\partial b}\frac{b}{p}\right] = 0$$

Mean value is zero !

$$\mathbb{E}\left[\frac{\partial p}{\partial b}\frac{b}{p}\right] = \int \frac{\partial p}{\partial b}\frac{b}{p} p dy = b\frac{\partial}{\partial b}\int p dy = 0$$

variance

$$\mathbb{E}_Y\left[\left(\frac{\partial p}{\partial b}\frac{b}{p}\right)^2\right] = b^2 \mathbb{E}_Y\left[\left(\frac{\partial \ln p}{\partial b}\right)^2\right]$$

We need to take a closer look at the variance!

(scalar notation is used here to keep things simple)

Mathematical framework for sensitivity

KPI-free sensitivity

Let's take a closer look at the variance term. It turns out that this is **Fisher Information** that is widely used in statistics

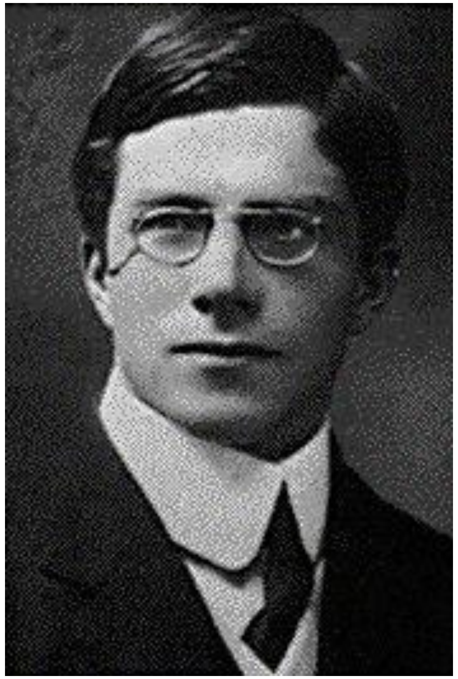
$$b^2 \mathbb{E}_Y \left[\left(\frac{\partial \ln p}{\partial b} \right)^2 \right]$$

Multiple variable case

$$b_j b_k \mathbb{E}_Y \left[\frac{\partial \ln p}{\partial b_j} \frac{\partial \ln p}{\partial b_k} \right] = F_{jk}$$

$$\mathbf{F}$$

Fisher Information Matrix, a square matrix of dimension *n*, where *n* is the number of parameter **b**



Ronald Fisher 1912.jpg

Mathematical framework for sensitivity

KPI-free sensitivity

Fisher matrix is a covariance matrix

$$\begin{array}{ccc}
\mathbf{F} & \xrightarrow{F_{jk}} & \mathbb{E}_Y\left[\frac{b_j \partial \ln p}{\partial b_j}\frac{b_k \partial \ln p}{\partial b_k}\right] \\
\big\downarrow & & \big\downarrow \text{Change to covariance notation} \\
\mathbf{F}\mathbf{q}_i = \lambda_i \mathbf{q}_i & \longleftarrow & cov\left(\frac{b_j \partial \ln p}{\partial b_j}, \frac{b_k \partial \ln p}{\partial b_k}\right)
\end{array}$$

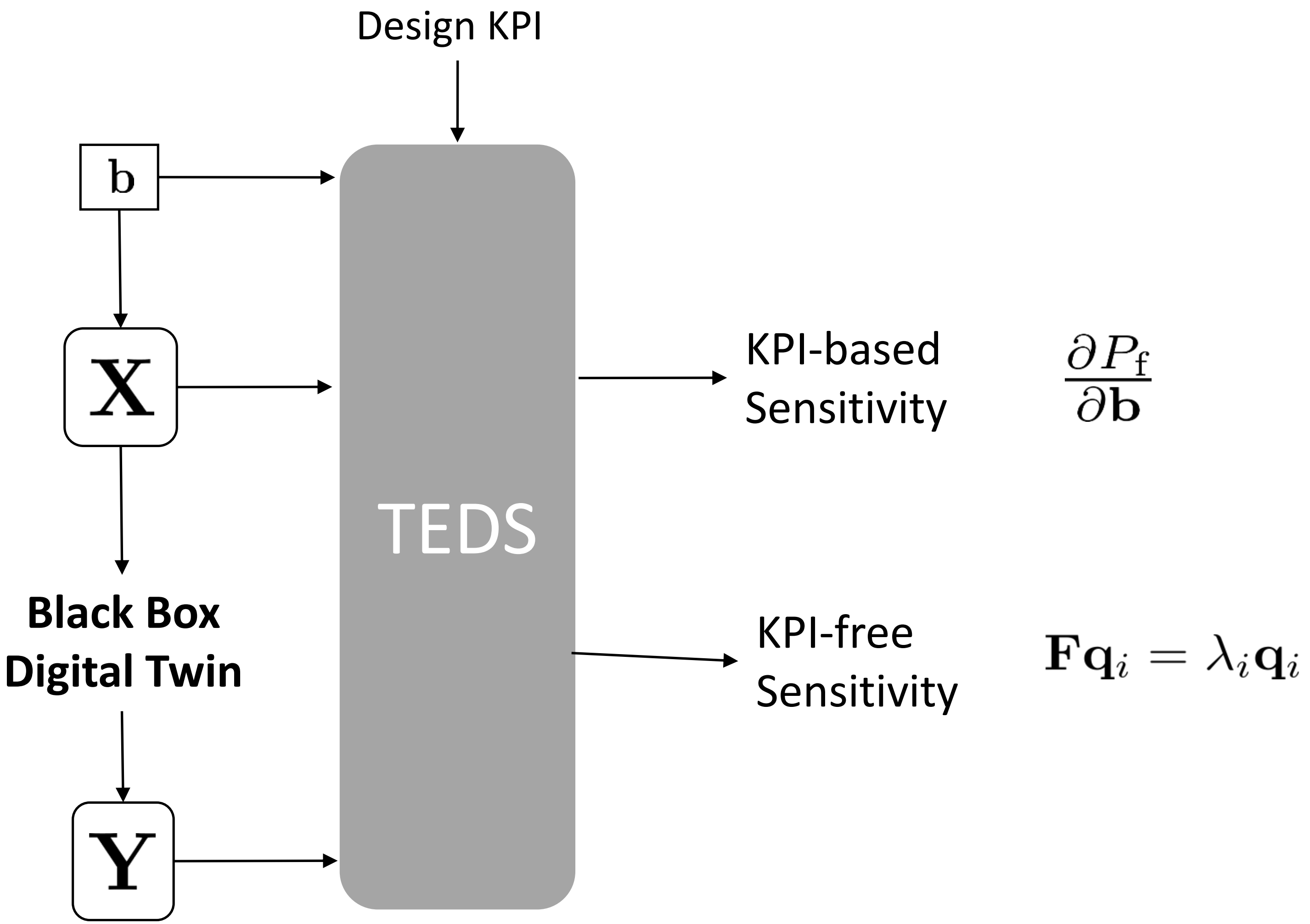
With the covariance matrix, we can look at the principal directions for the sensitivity (the most sensitive directions), by computing the eigenvectors of the Fisher matrix (**q** is the eigenvector and λ is the eigenvalue). The eigenvectors (directions) with largest eigenvalues then point out the most important parameters that we should focus on to reduce uncertainties.

What does this mean? It means that we try to measure the spread of the sensitivity at different realizations of our random output (note that the mean value of the sensitivity is zero)

Toolbox for engineering design sensitivity (TEDS)

In summary, we have developed a toolbox to calculate sensitivity to uncertainties in the design. It can be either KPI based or KPI free.

As to be shown in the example results, these two are correlated for the same quantity of interest. Therefore, design ‘surprises’ are minimized even KPIs only specified quite late in the design.

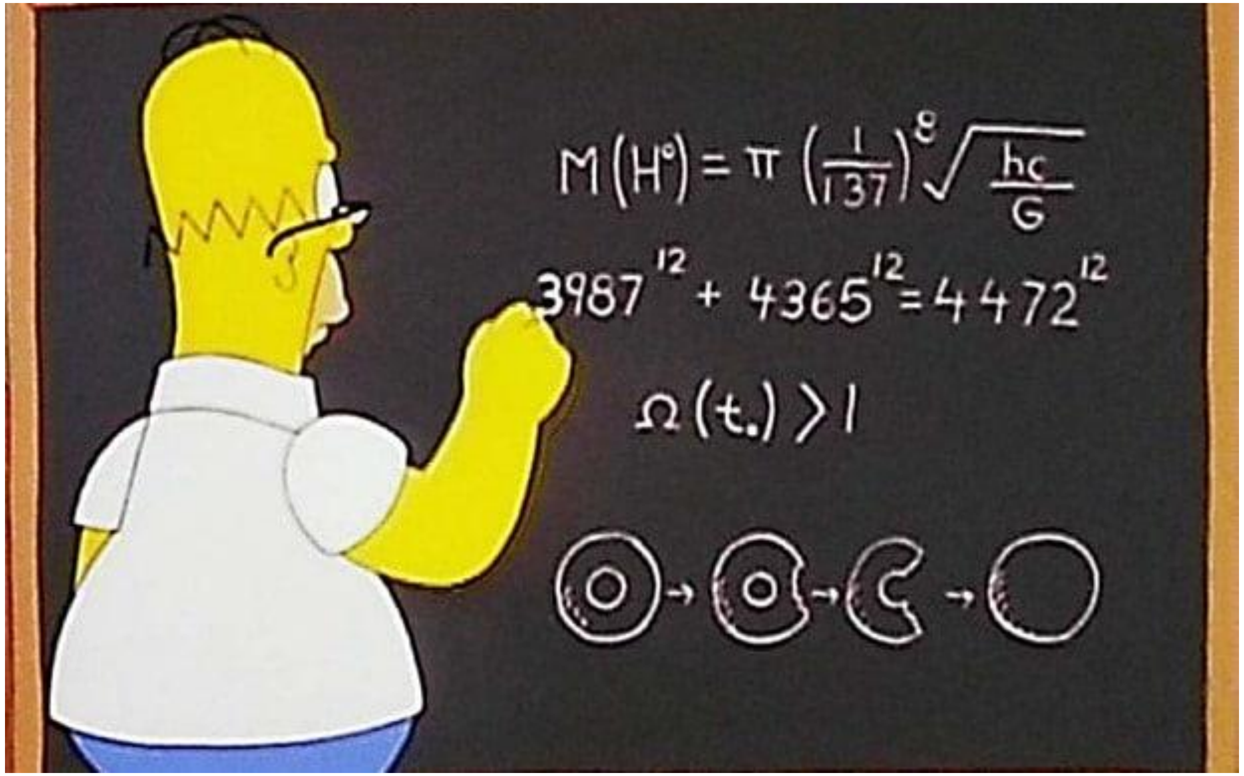


KPI: Key performance indicator

Example application of TEDS

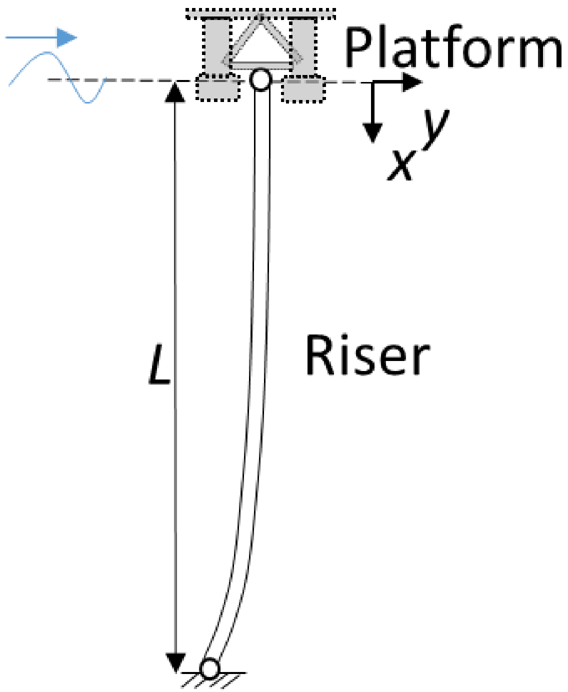
We have covered the details of what is inside the toolbox TEDS.

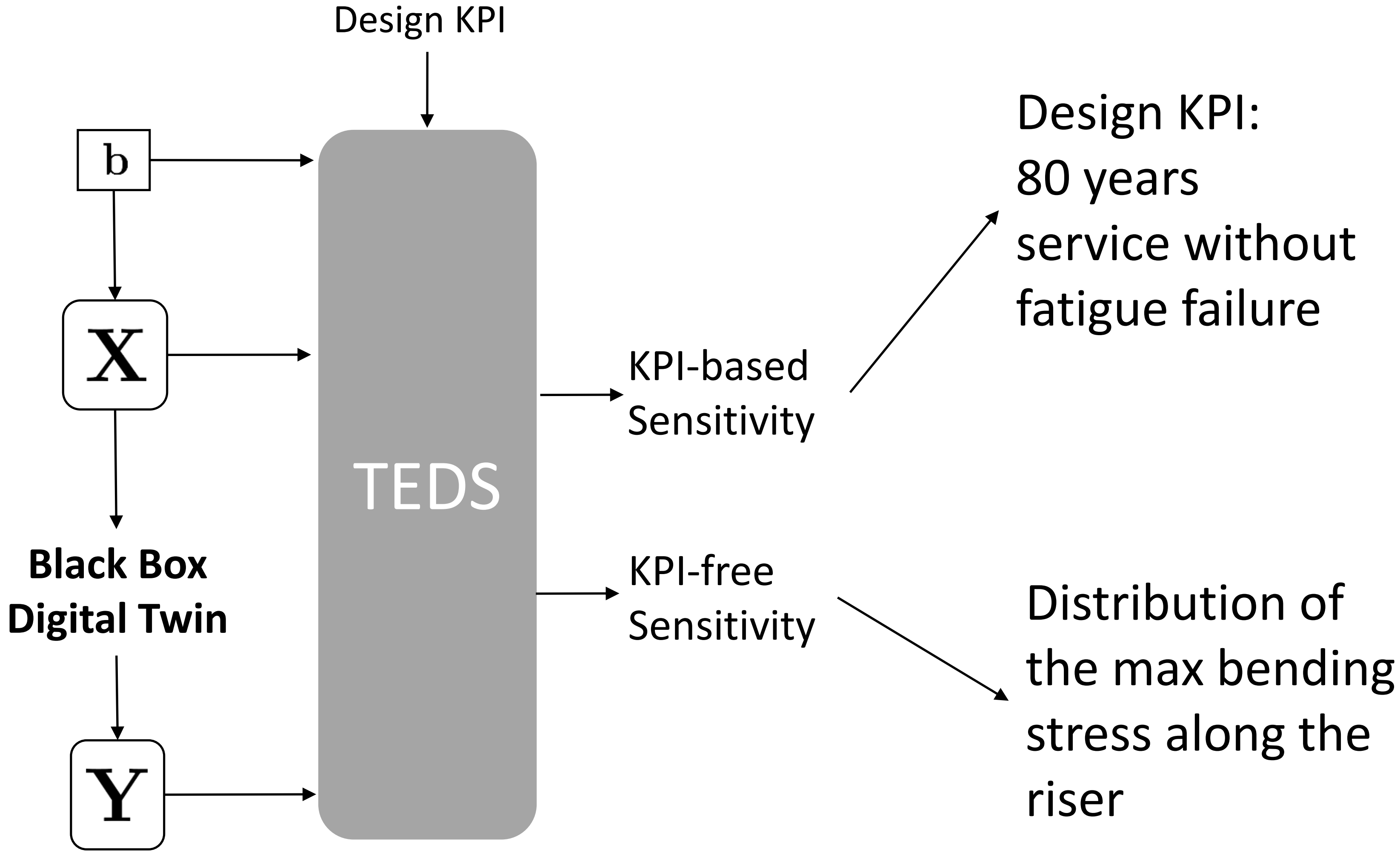
No more equations, let's have a look at some example results.



Example results

To have a better idea of how this toolbox is applied, we analyse an example on design of a marine riser.

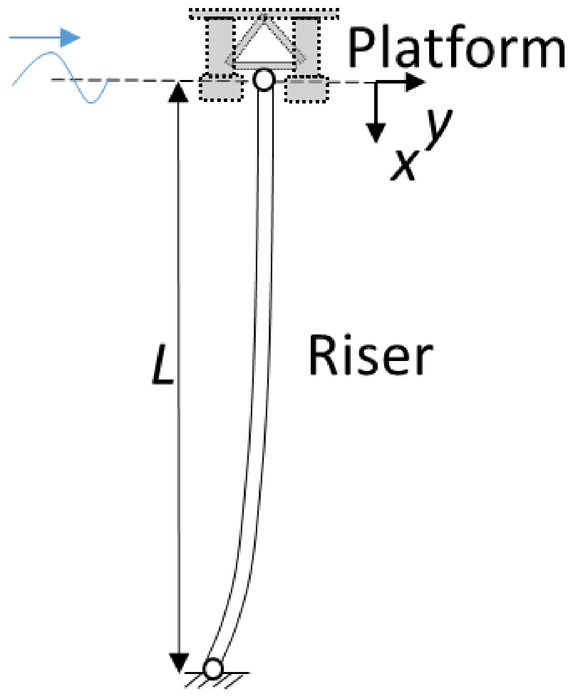




Example results

Mean and standard deviation values for the random input variables (Gaussian)

Random Variable		Mean	Standard deviation
Morison's equation added mass coefficient	C_a [-]	1.5	0.3
Morison's equation drag coefficient	C_d [-]	1.1	0.22
Marine riser steel density	ρ [kg/m^{-3}]	7840	392
Marine riser Young's modulus	E [GPa]	200	10
Riser internal oil density	ρ_o [kg/m^{-3}]	920	92
Marine riser top tension	T_0 [kN]	4905	490.5
Material S-N curve coefficients	α [GPa]	199	19.9
	δ [-]	3	0.3



S-N law $N(s) = \alpha s^{-\delta}$

Example results

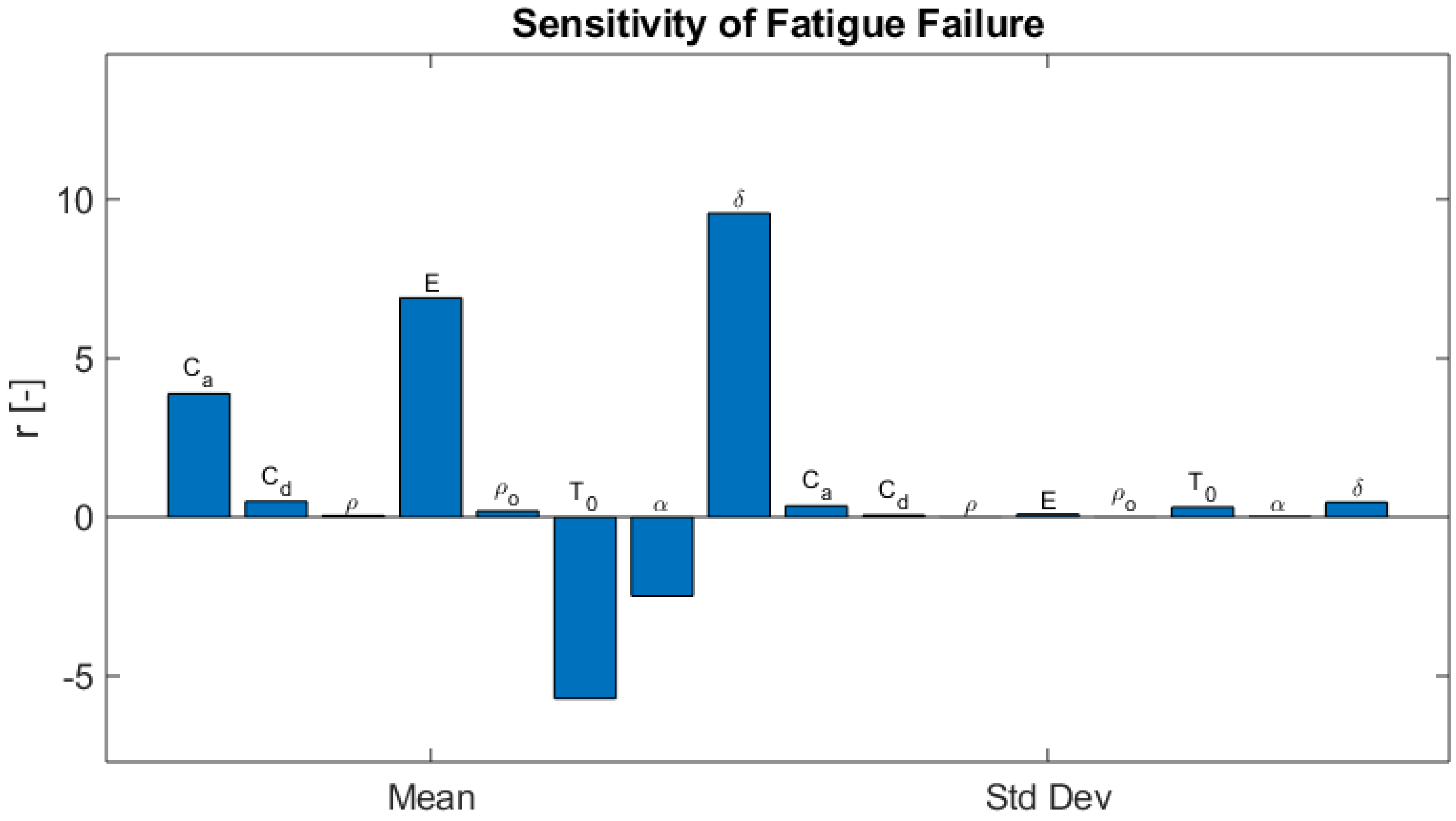
KPI-based sensitivity

What is the sensitivity of design KPI: 80 years service without fatigue failure

The sensitivity results tells us which parameter is important for the specified KPI: bigger value of r means the fatigue failure is more sensitive to it.

The results shown here do make sense because for example:

- δ is quite important because number of cycles to fatigue failure depends on stress to the power of δ
- E is important because bending stress is proportional to it

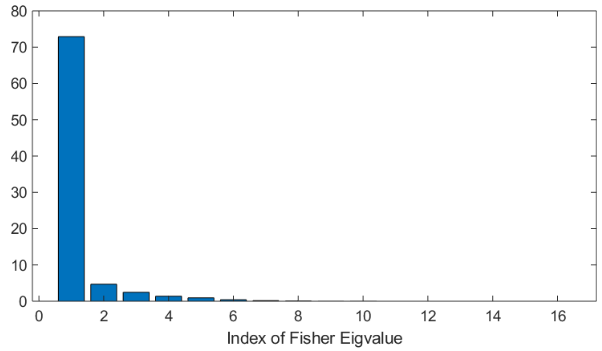


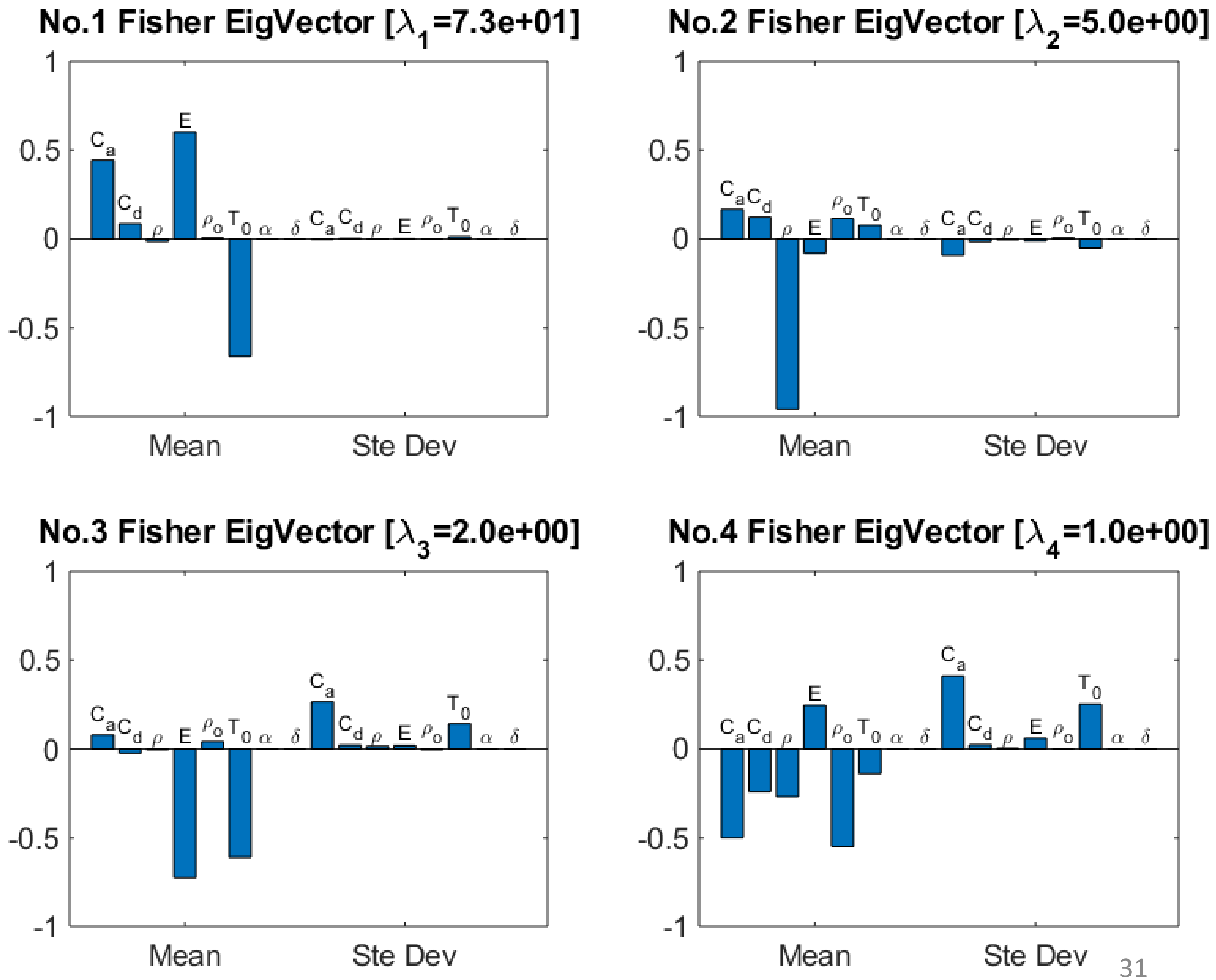
Example results

KPI-free sensitivity

What is the sensitivity of the general distribution of the max bending stress? Fisher Eigenvectors

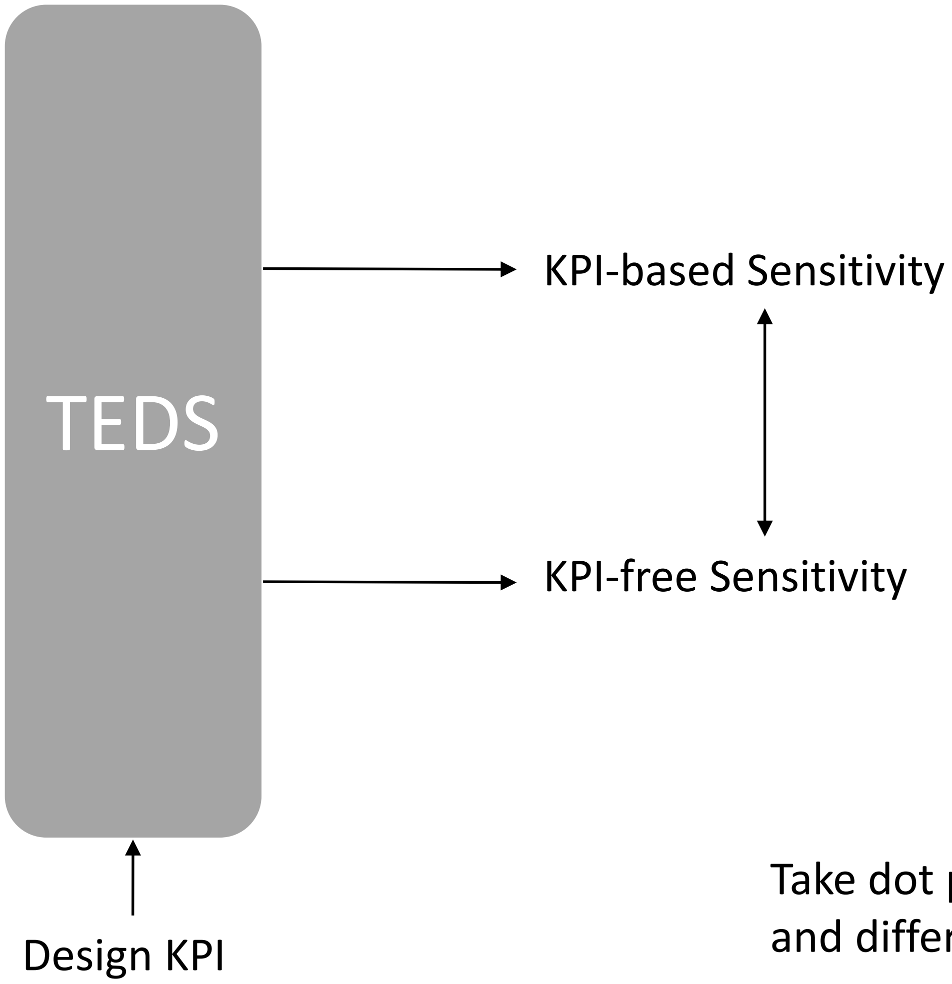
First, only one dominant eigenvector in this case (see magnitude of eigenvalues below) → this is the most sensitive direction



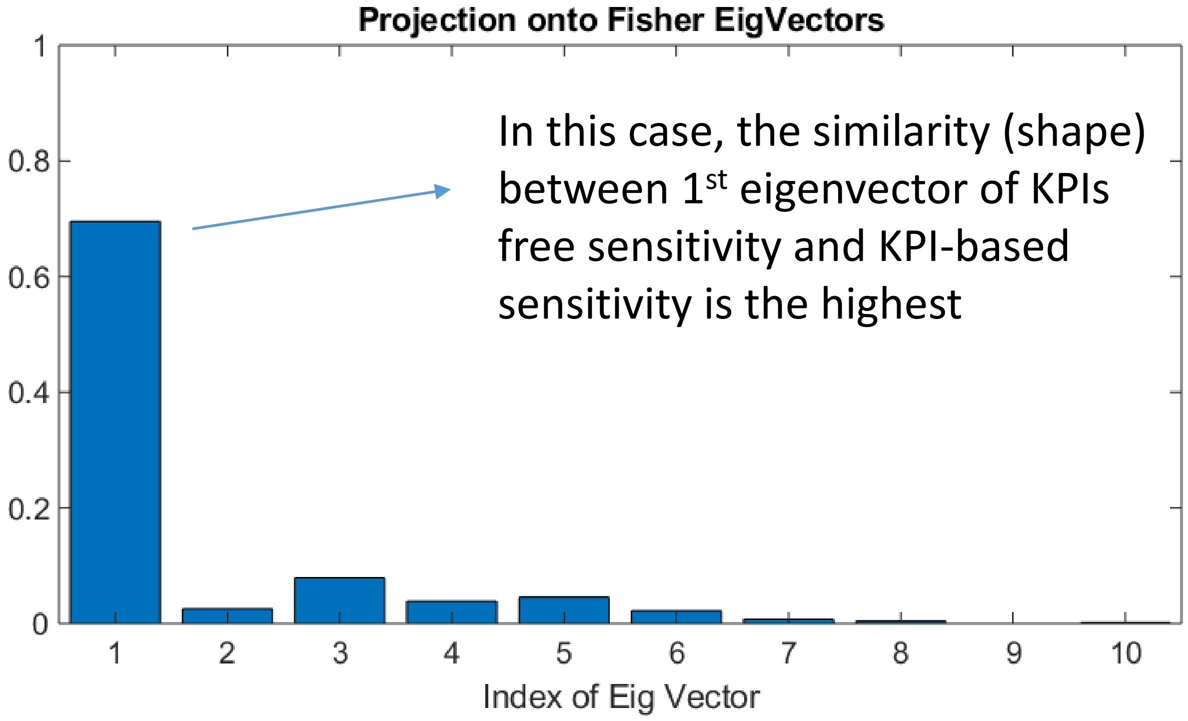


Second, as compared to the scatter plot at the beginning of the presentation, we have Ca and E as important ones, but E of similar importance as Ca, unlike the scatter plot where it seemed that Ca is dominant

Example results



How are these two types of sensitivity related in this case?



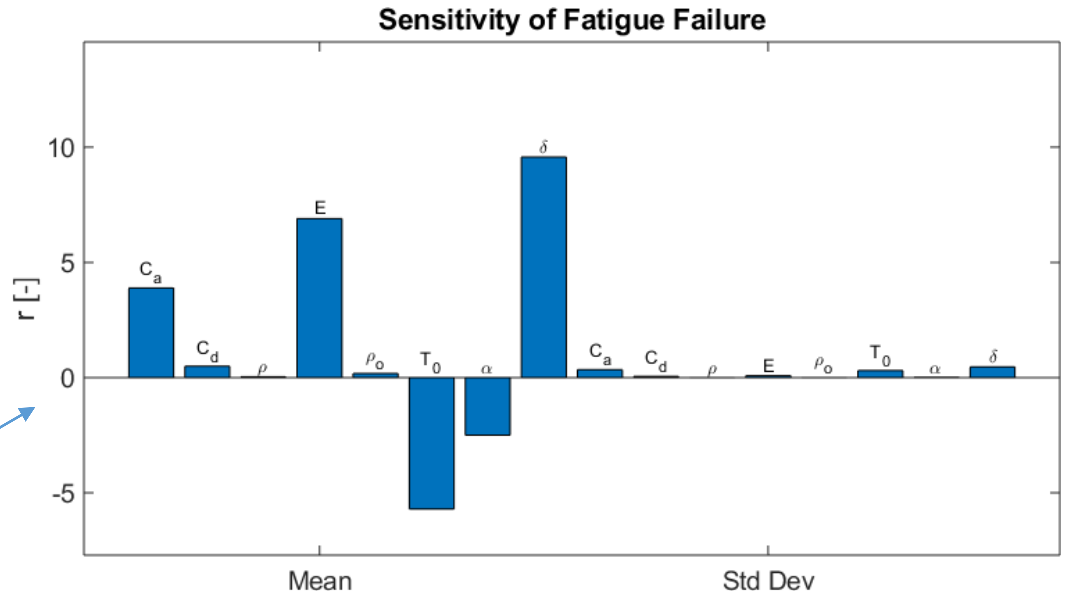
Take dot product between the failure sensitivity vector and different eigenvectors to compute the projection $\frac{\partial P_{\mathrm{f}}}{\partial \mathbf{b}} \cdot \mathbf{q}_i$

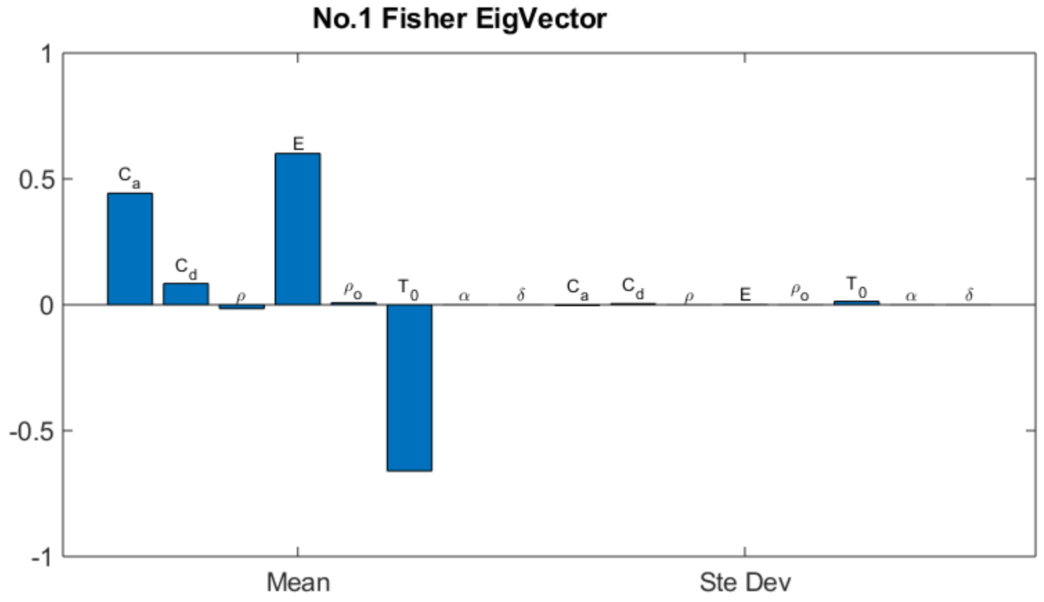
Example results

TEDS

KPI-based Sensitivity

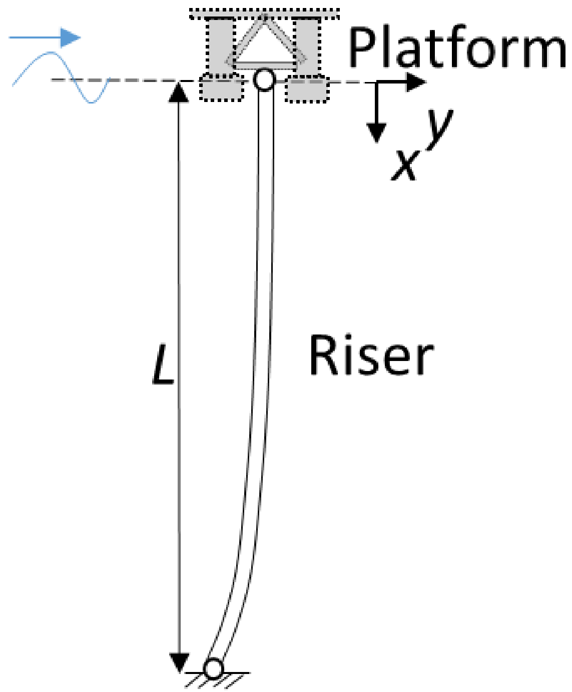
KPI-free Sensitivity

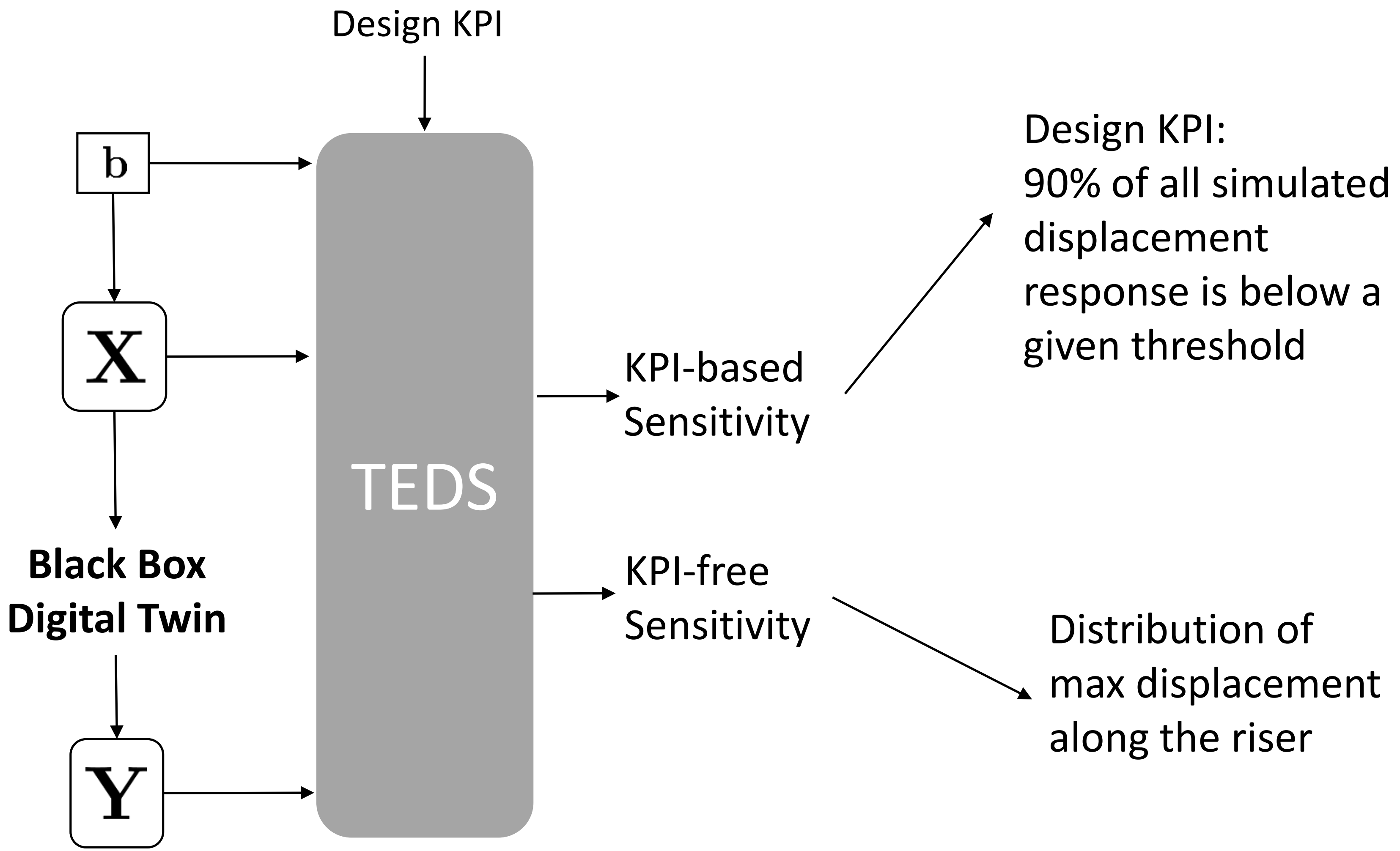




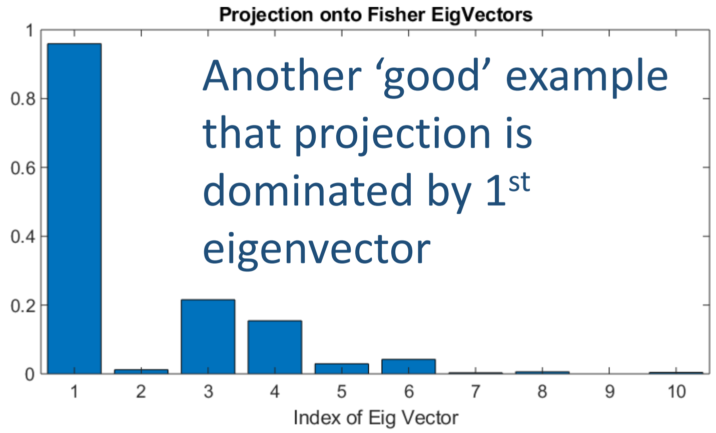
Clearly quite similar in this case! This is expected because both sensitivities are related to high stress levels. However, the level of similarity is case dependent.

More Example Results – Displacement Response





More Example Results – Displacement Response

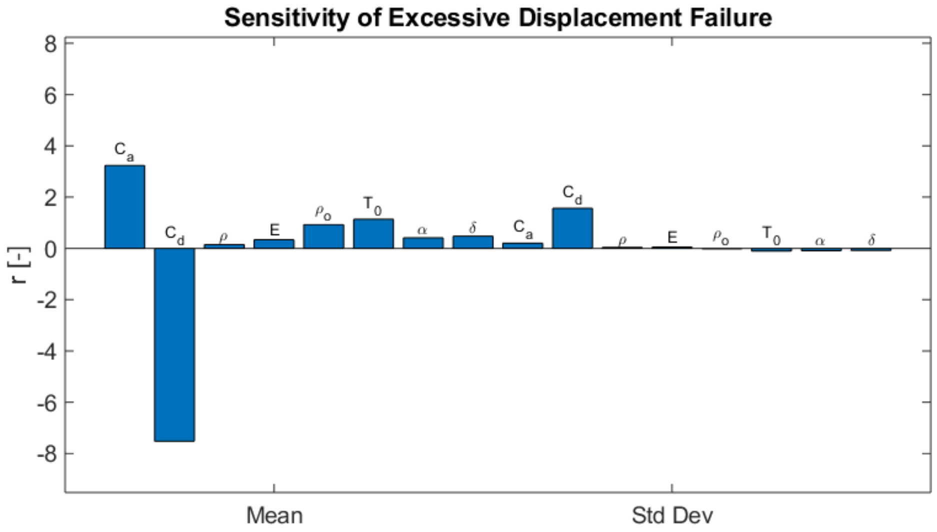


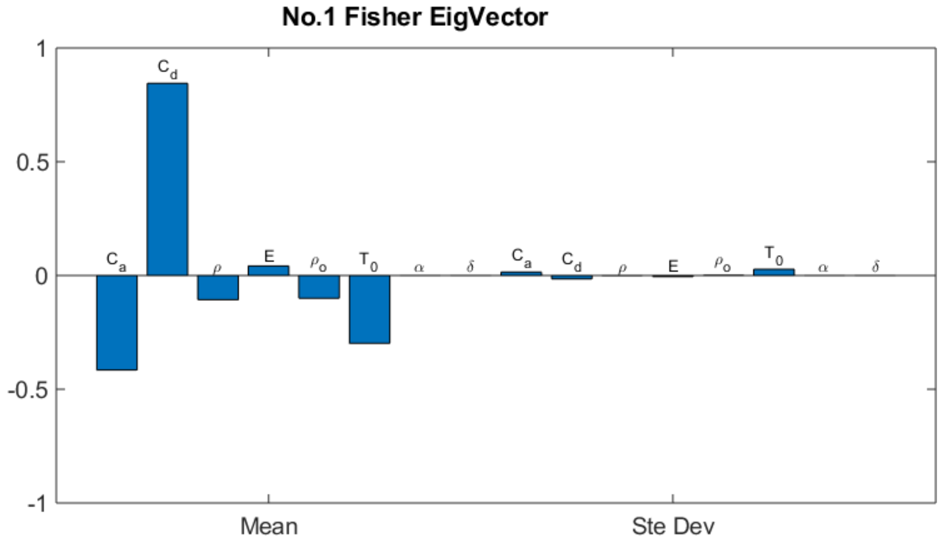
Another ‘good’ example that projection is dominated by 1st eigenvector

TEDS

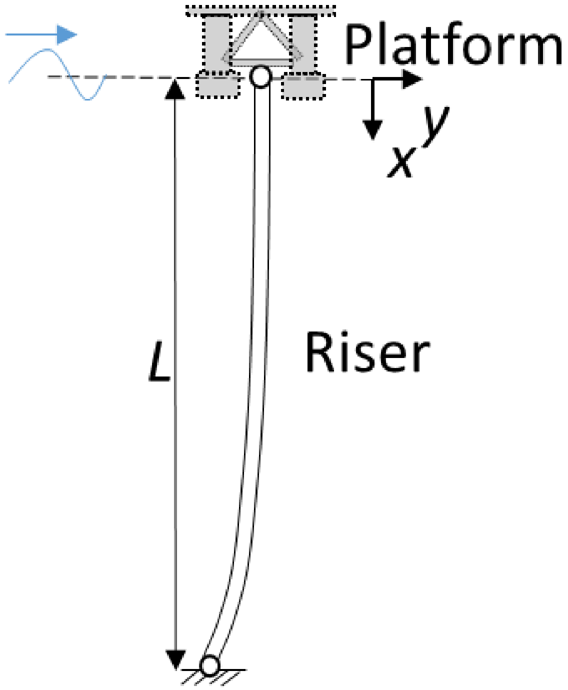
KPI-based Sensitivity

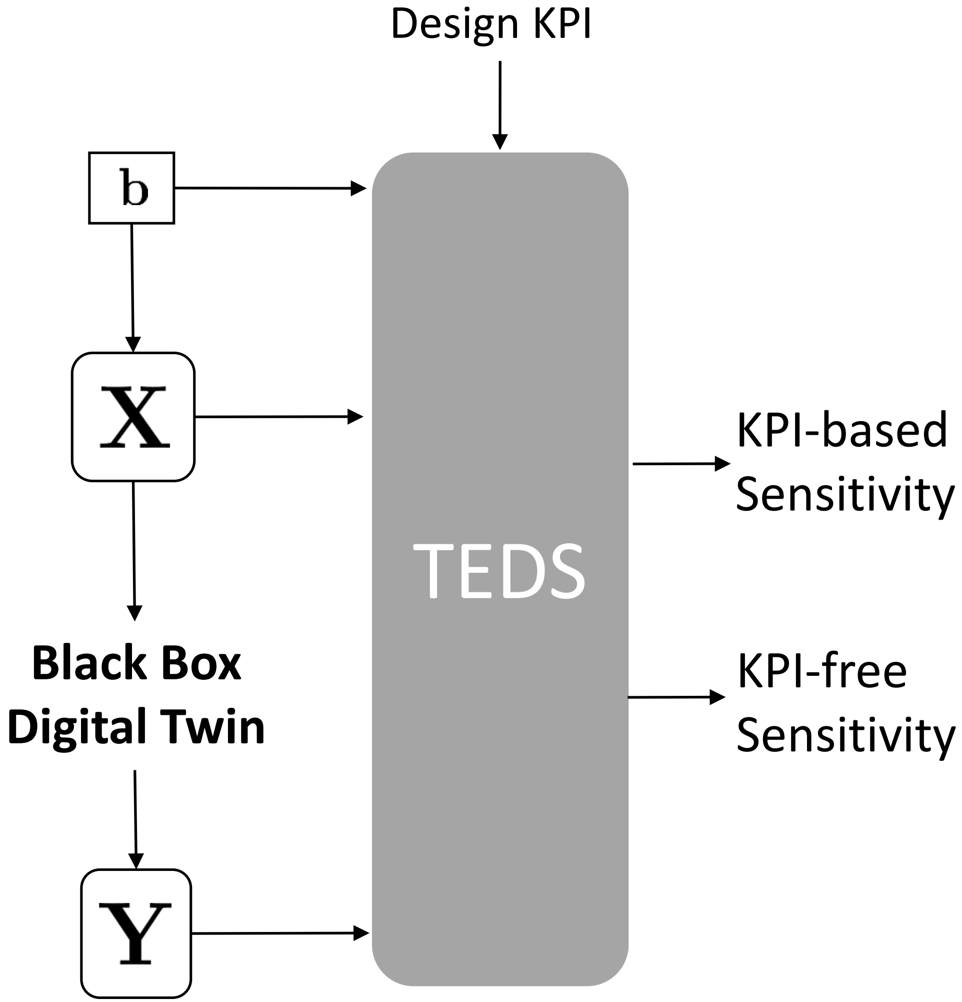
KPI-free Sensitivity





More Results – General Case





Three Design KPI:
- Fatigue failure
- Excessive displacement
- Excessive rotation

Joint distribution of max stress, displacement and rotation along the riser

More Results – General Case

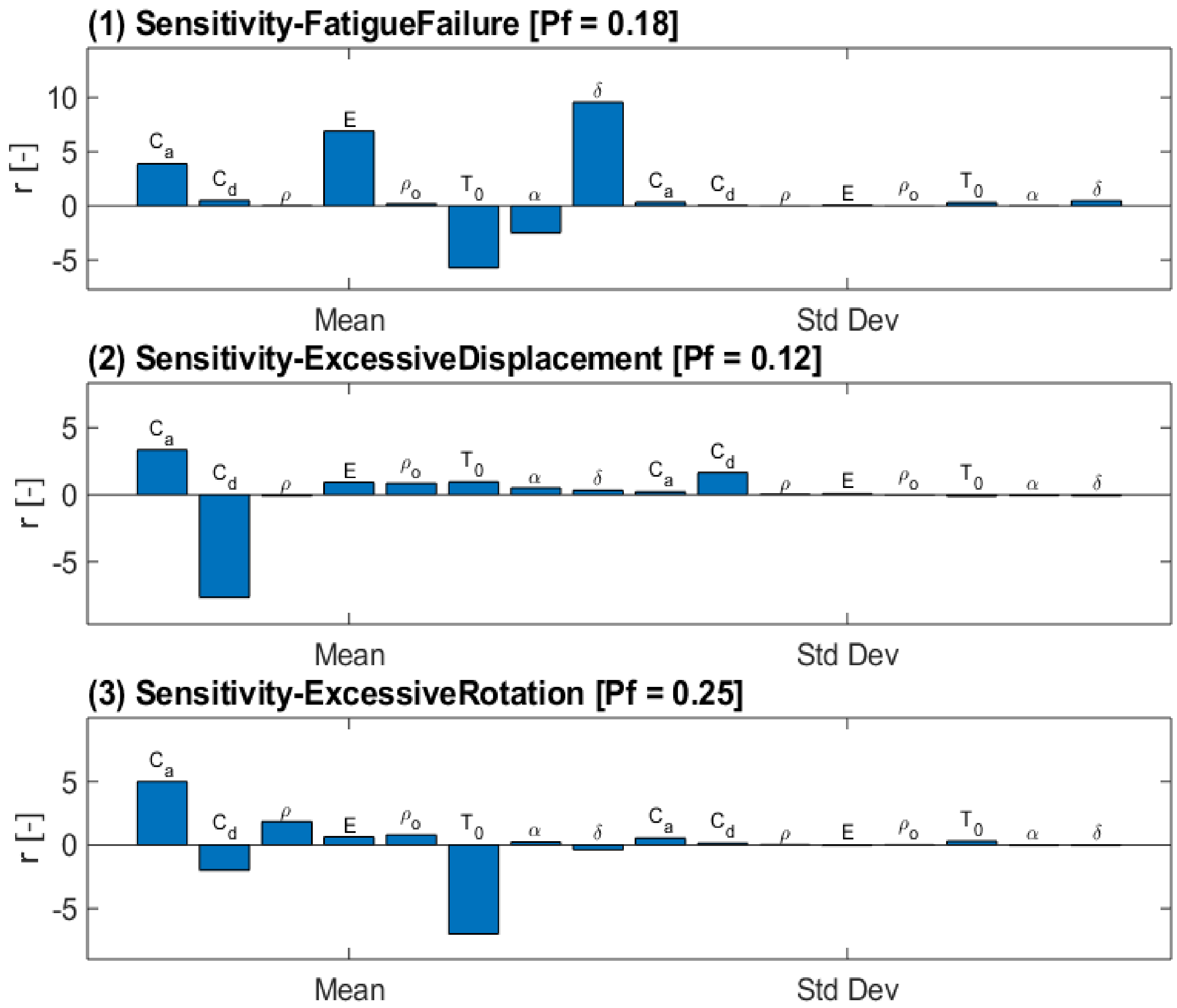
Different important parameters for the three failure modes

TEDS

KPI-based Sensitivity

Three Design KPI:

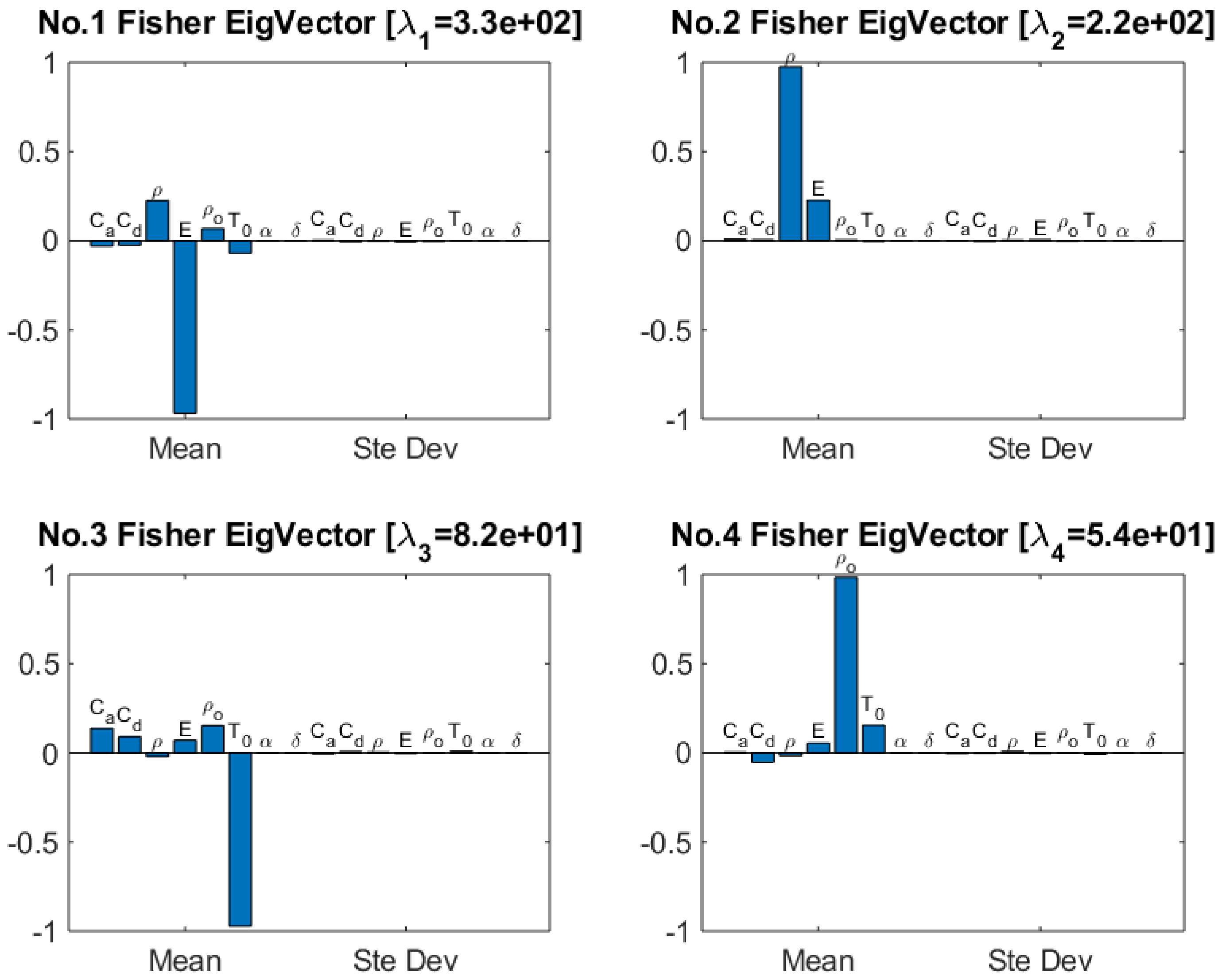
- Fatigue failure
- Excessive displacement
- Excessive rotation



More Results – General Case

TEDS

KPI-free Sensitivity



Joint distribution of max stress, displacement and rotation along the riser

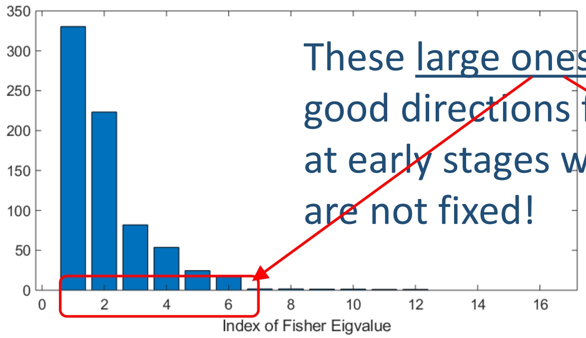
More Results – General Case

KPI-based Sensitivity

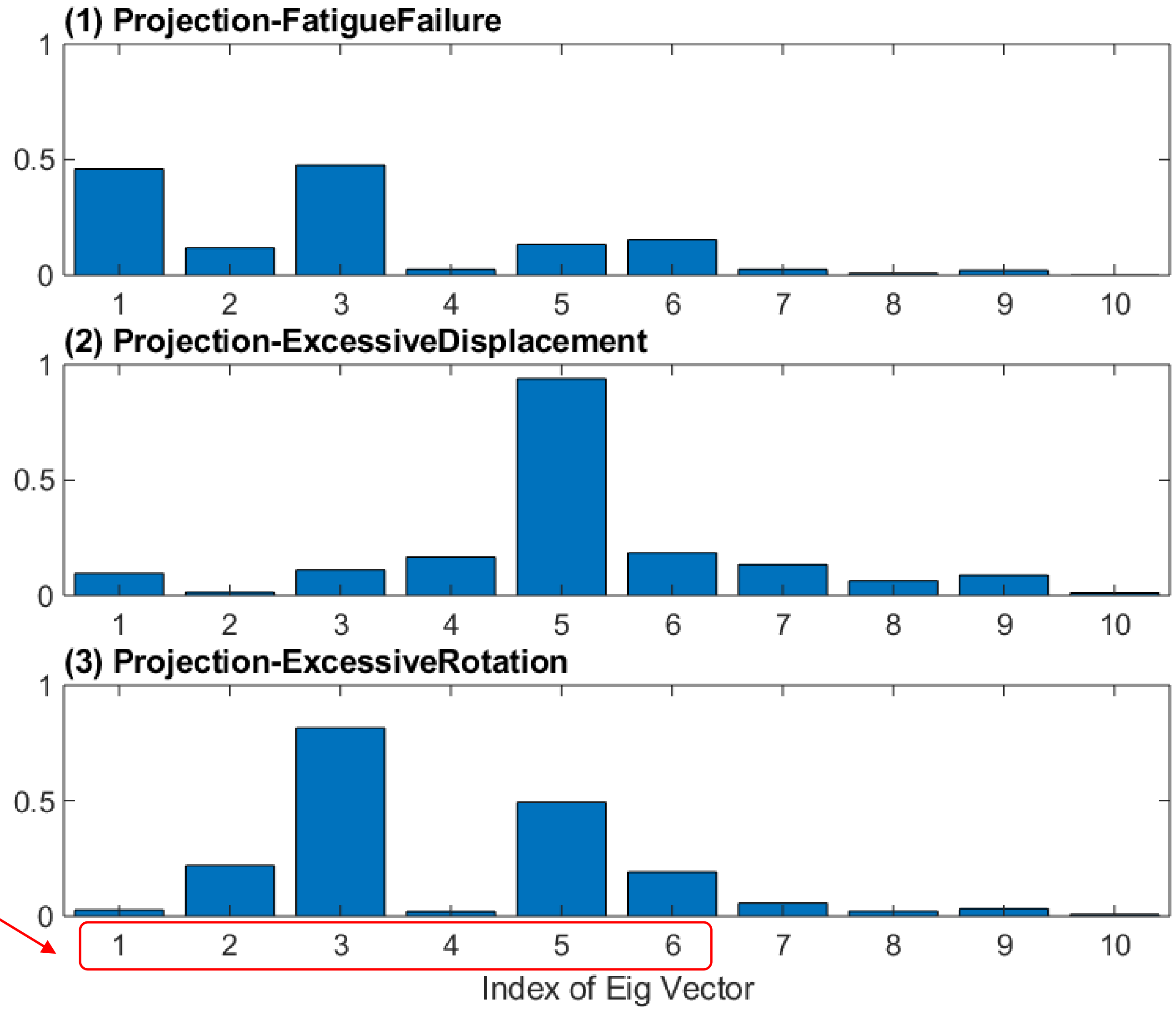
This is a counter example that the two sensitivities are not always pointing exactly to the same direction !

KPI-free Sensitivity

However, it can be seen that the projections of the KPI-based sensitivities are still dominated by the eigenvectors of Fisher matrix, with large eigenvalues !



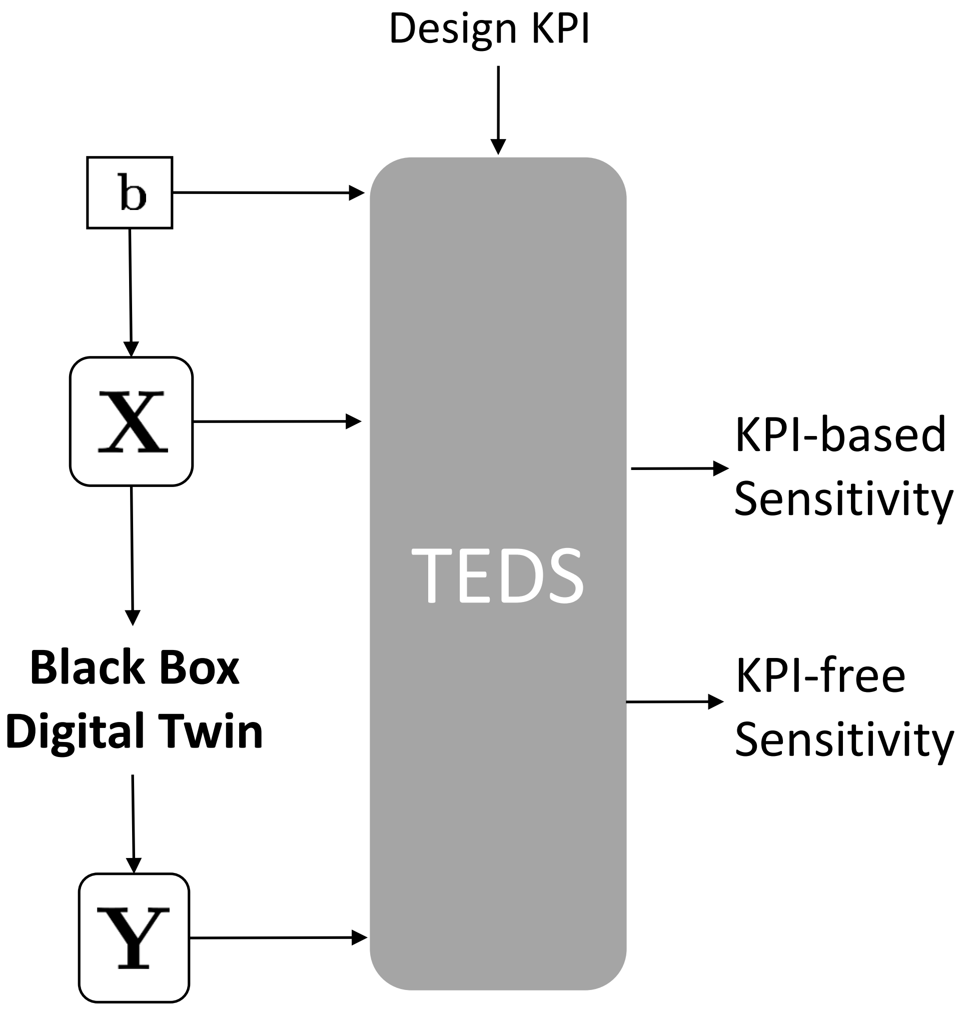
These large ones give you good directions for design at early stages where KPIs are not fixed!



Conclusions

- TEDS – a toolbox for engineering design sensitivities
  - a probabilistic model that quantifies sensitivity to design uncertainties
  - can be wrapped around black box digital twins and computational simulations
  - provides both KPI-based and KPI-free sensitivities

- Yet applied on real industrial cases, a few limitations expected:
  - Only parametric distribution for design input, where a variety of uncertainty types are needed for industrial applications
  - Still a gap to design decisions: is it worth investment to reduce uncertainties?



KPI: Key performance indicator

Any questions/comments:

jy419@cam.ac.uk