# ROBUST DESIGN UNDER UNCERTAINTY
## via counterfactual Bayes

Carl Ehrett

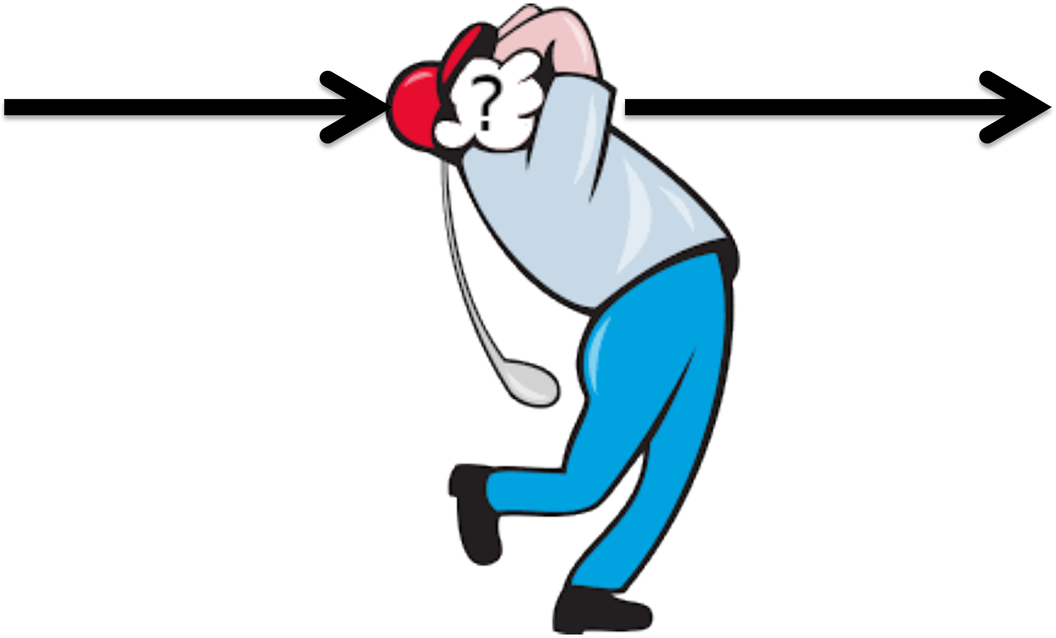
2019-10-28

Club properties:

- Loft angle
- Head material
- Shaft flexibility
- Shaft length

Determines

# Dave's secret weapon





Club properties:

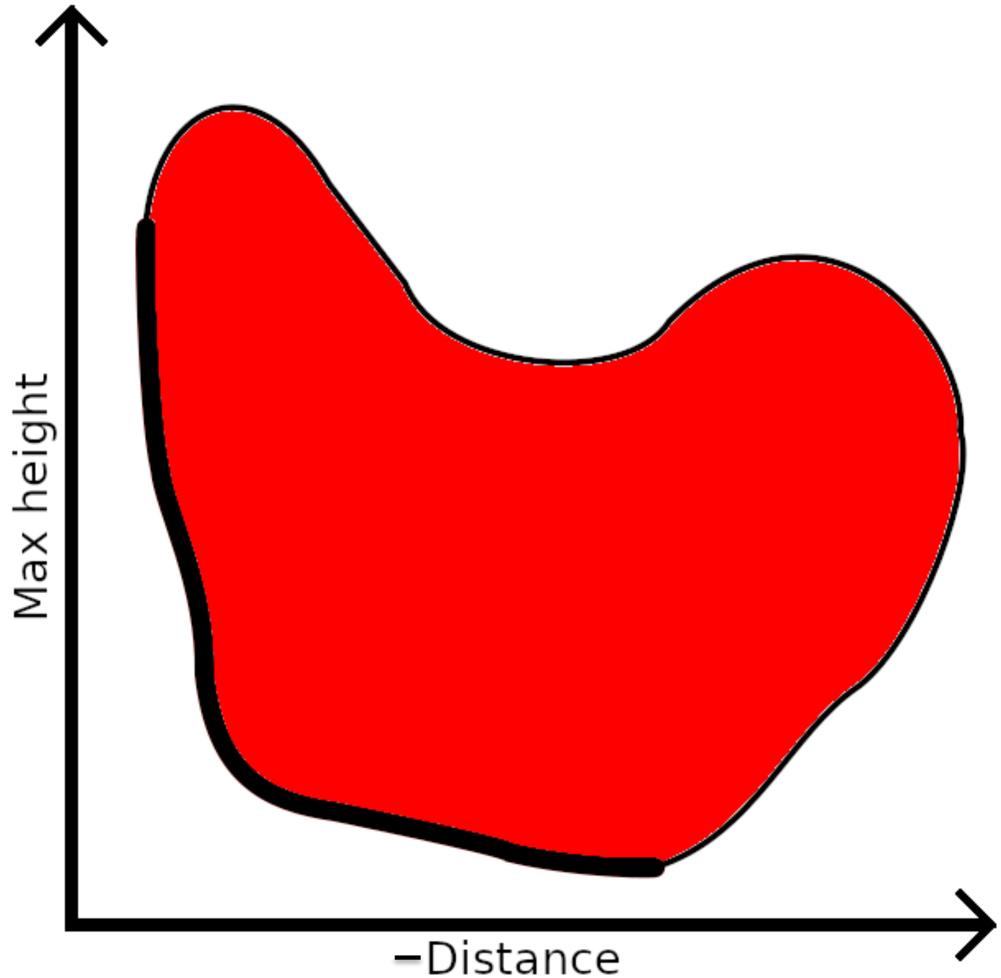
- Loft angle
- Head material
- Shaft flexibility
- Shaft length

## Multi-objective optimization

- Often, no particular solution is objectively best.
- We want then to identify some point on the Pareto front, or the whole front.
- The Pareto front is the set of solutions such that improving them in some outcome requires degrading them in some other outcome.

## Uncertainty quantification

- We also know our solution will be uncertain, and we want to quantify this.
- Uncertainty can come from uncertain inputs, observation error, model bias, nondeterministic models, and more.



# Model calibration

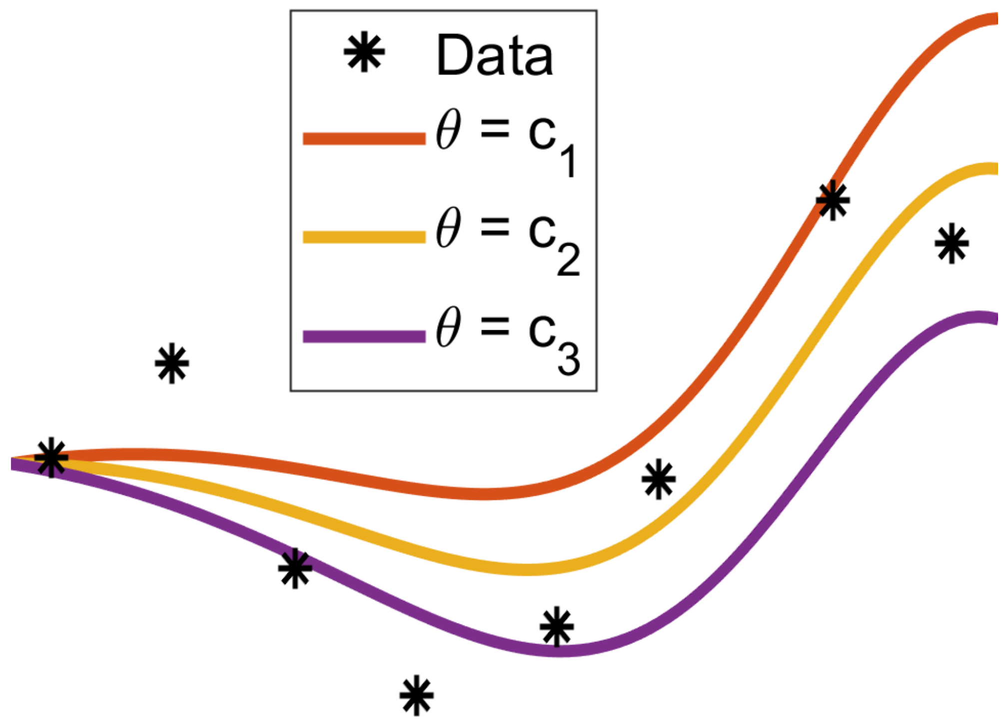
- In computer model calibration, we have:

$$\boldsymbol{y}(\boldsymbol{x}) = f(\boldsymbol{x}) + \epsilon(\boldsymbol{x}) = \eta(\boldsymbol{x}, \boldsymbol{\theta}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

where $\boldsymbol{y}(\boldsymbol{x})$ is the observed value at $\boldsymbol{x}$, $f(\boldsymbol{x})$ is the value of the true system at $\boldsymbol{x}$, $\epsilon(\boldsymbol{x})$ is the observation error, $\eta(\boldsymbol{x}, \boldsymbol{\theta})$ is the value of the computer model and $\delta(\boldsymbol{x})$ captures the bias of the computer model.

- Here, $\boldsymbol{x}$ is the known and/or controllable input settings, and $\boldsymbol{\theta}$ is the unknown parameters which must be given as inputs to the computer model.

- The purpose of computer model calibration is to use a set of observations $\boldsymbol{y}$ to estimate $\boldsymbol{\theta}$.



# Calibration to target outcomes (CTO)

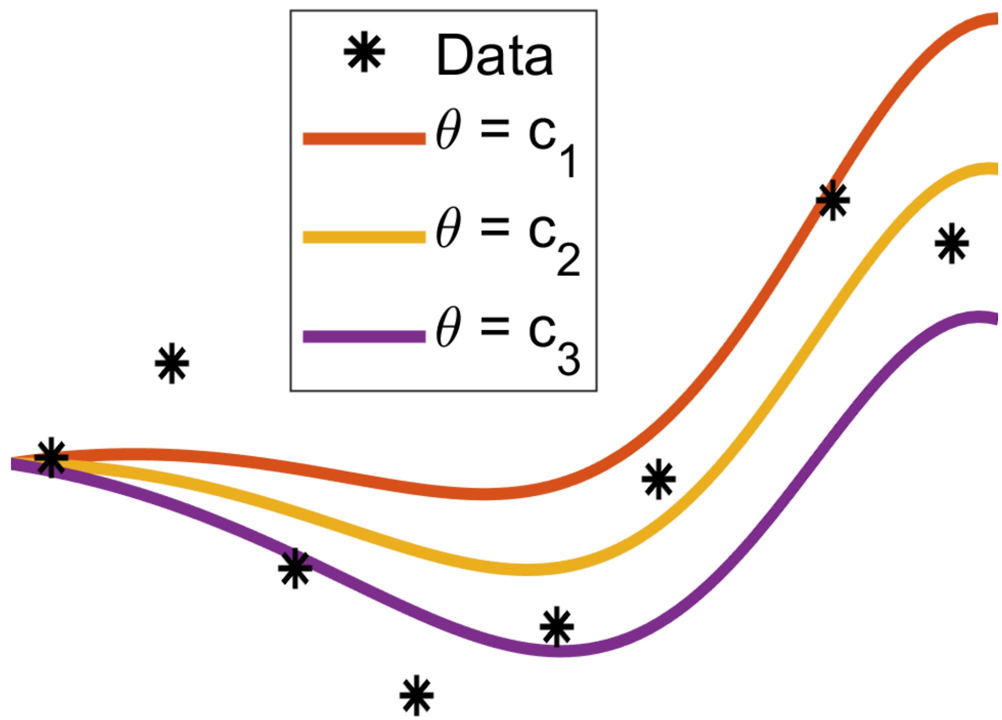
- When we consider this same framework:

$$\boldsymbol{y}(\boldsymbol{x}) = f(\boldsymbol{x}) + \epsilon(\boldsymbol{x}) = \eta(\boldsymbol{x}, \boldsymbol{\theta}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

as a method for design, $\boldsymbol{\theta}$ is now also a controllable input, as $\boldsymbol{x}$ is.

- The difference between $\boldsymbol{x}$ and $\boldsymbol{\theta}$ now is that the latter is the input for which we want to find optimal settings.

- Instead of observed data, we have ***target outcomes*** that describe the way we *want* the system to behave.

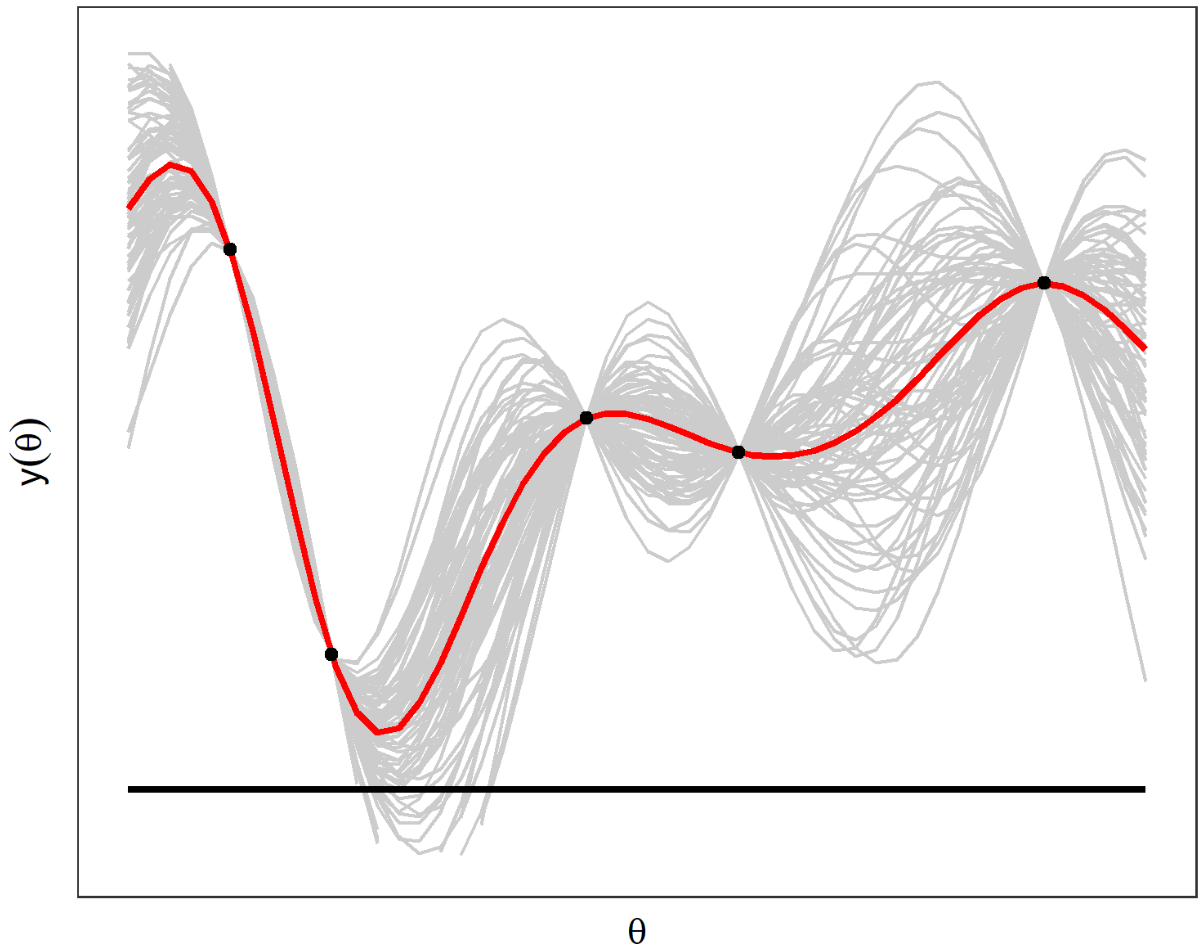


# Gaussian process (GP) meta-models

- For computationally expensive models, GPs are good meta-model.
- A GP is a random function characterized by a mean function $m(\cdot)$ and a covariance function $C(\cdot,\cdot)$.
- Let $\mathbf{\mu_Z}$ be the value of $m(\cdot)$ at each location in $\mathbf{Z} = (\mathbf{X}, \mathbf{\Theta})$, and $\mathbf{C_{Z,Z'}}$ a matrix formed by $C(\cdot,\cdot)$ on each pair of locations from $\mathbf{Z}$ and $\mathbf{Z}'$. Updated mean and covariance at $\mathbf{Z}'$ from training points $\mathbf{\eta_Z} = \eta(\mathbf{Z})$ :

$$\mathbf{\mu}^*_{\mathbf{Z}'} = \mathbf{\mu}_{\mathbf{Z}'} + \mathbf{C}_{\mathbf{Z}',\mathbf{Z}}\mathbf{C}^{-1}_{\mathbf{Z},\mathbf{Z}}(\mathbf{\eta_Z} - \mathbf{\mu_Z})$$
$$\mathbf{C}^*_{\mathbf{Z}',\mathbf{Z}'} = \mathbf{C}_{\mathbf{Z}',\mathbf{Z}'} - \mathbf{C}_{\mathbf{Z}',\mathbf{Z}}\mathbf{C}^{-1}_{\mathbf{Z},\mathbf{Z}}\mathbf{C}_{\mathbf{Z},\mathbf{Z}'}$$

- Provides quantification of remaining uncertainty.

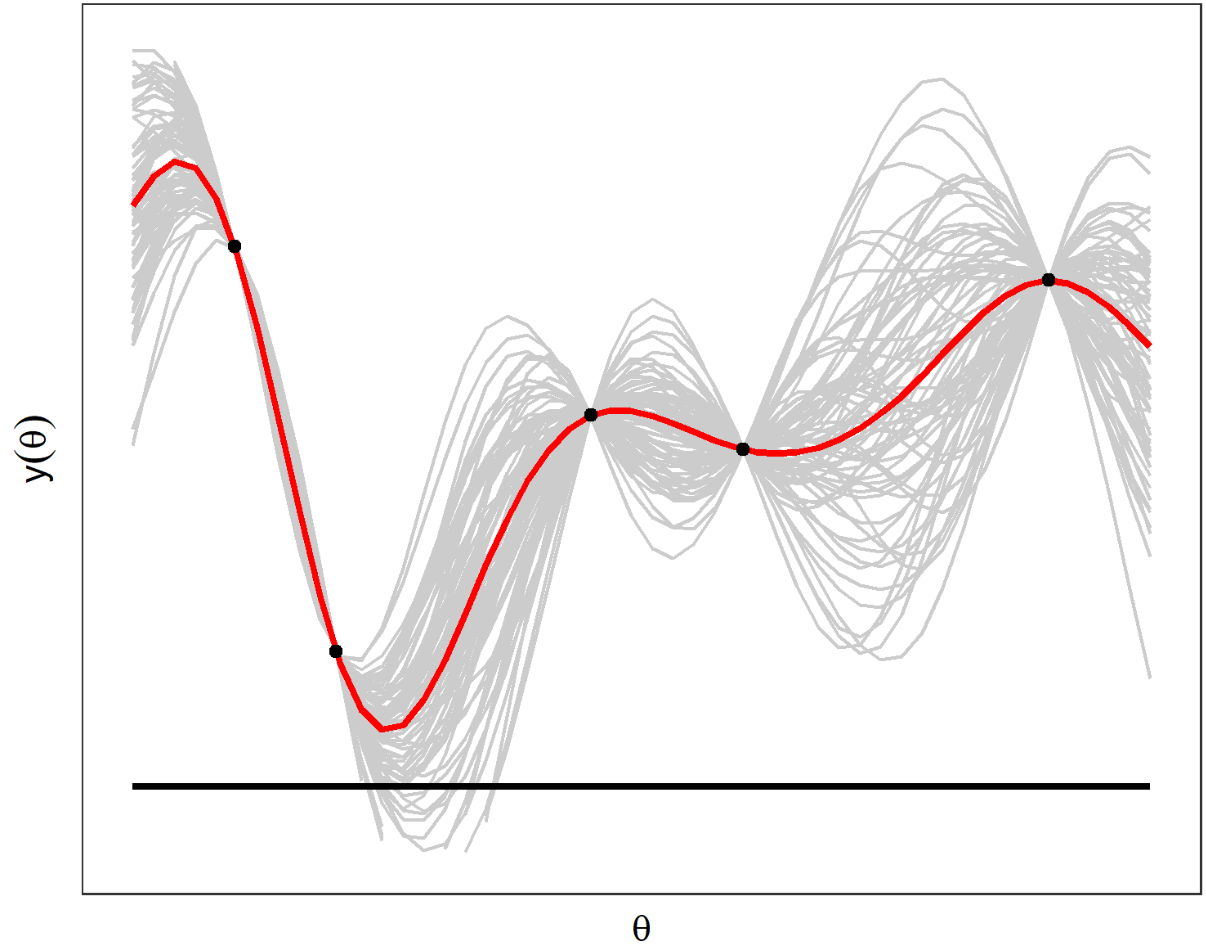


# Bayesian analysis

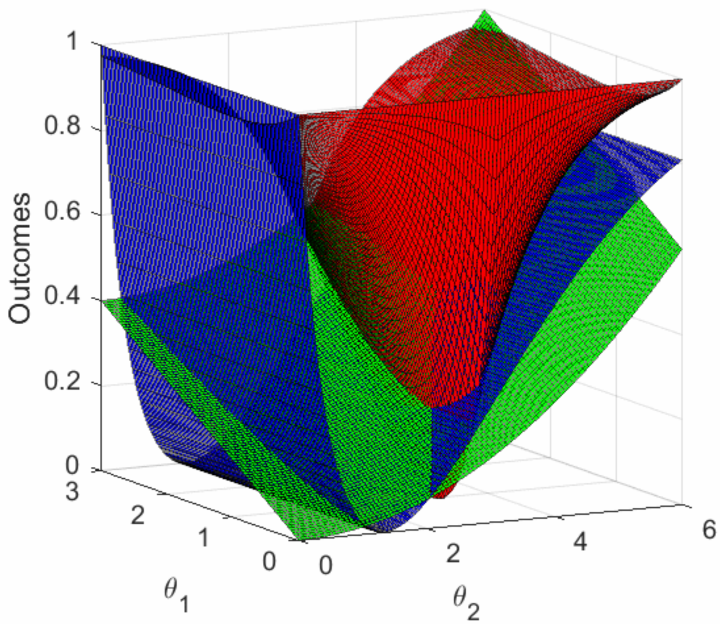
- The GP meta-model gives us $\pi(\mathbf{D}|\boldsymbol{\theta}, \boldsymbol{\phi}_\eta, \boldsymbol{\phi}_\delta)$ where $\mathbf{D}$ is an array including both the observed and target outcomes, and $\boldsymbol{\phi}_\eta, \boldsymbol{\phi}_\delta$ are the hyperparameters of the GP priors for the meta-model and the discrepancy $\delta()$.
- After setting priors on $\boldsymbol{\theta}, \boldsymbol{\phi}_\eta, \boldsymbol{\phi}_\delta$, we have the posterior:

$$\pi(\boldsymbol{\theta}, \boldsymbol{\phi}_\eta, \boldsymbol{\phi}_\delta|\mathbf{D}) \propto \pi(\mathbf{D}|\boldsymbol{\theta}, \boldsymbol{\phi}_\eta, \boldsymbol{\phi}_\delta)\times \pi(\boldsymbol{\theta})\times\pi(\boldsymbol{\phi}_\eta)\times\pi(\boldsymbol{\phi}_\delta)$$

- We use Markov chain Monte Carlo (MCMC) to explore this distribution.



# Simulated example



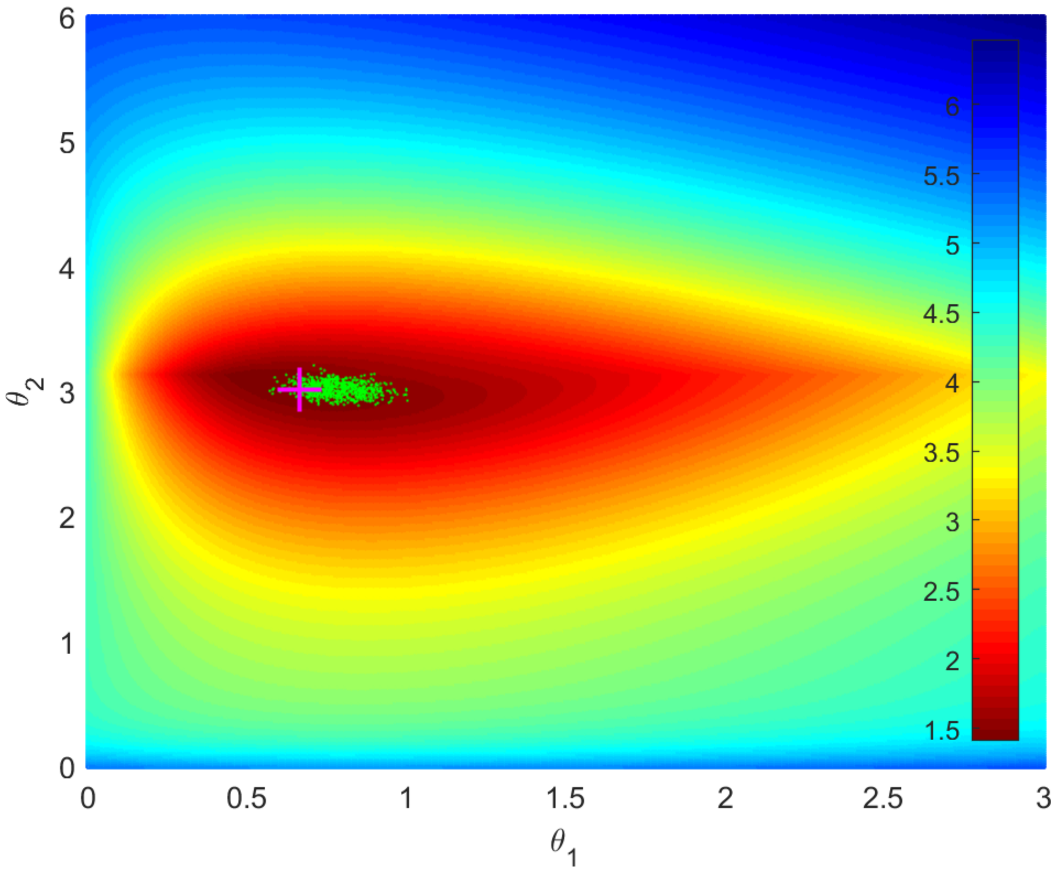

$$y_1(x, \theta_1, \theta_2) = \left(\theta_1 \exp\left(-\theta_1 - \left|\theta_2 - \frac{\pi x}{2}\right|\right) + 1\right)^{-1}$$

$$y_2(x, \theta_1, \theta_2) = \left(\theta_2^{x-1} \exp\left(-\frac{3\theta_2}{4}\right) + 1\right)^{-1}$$

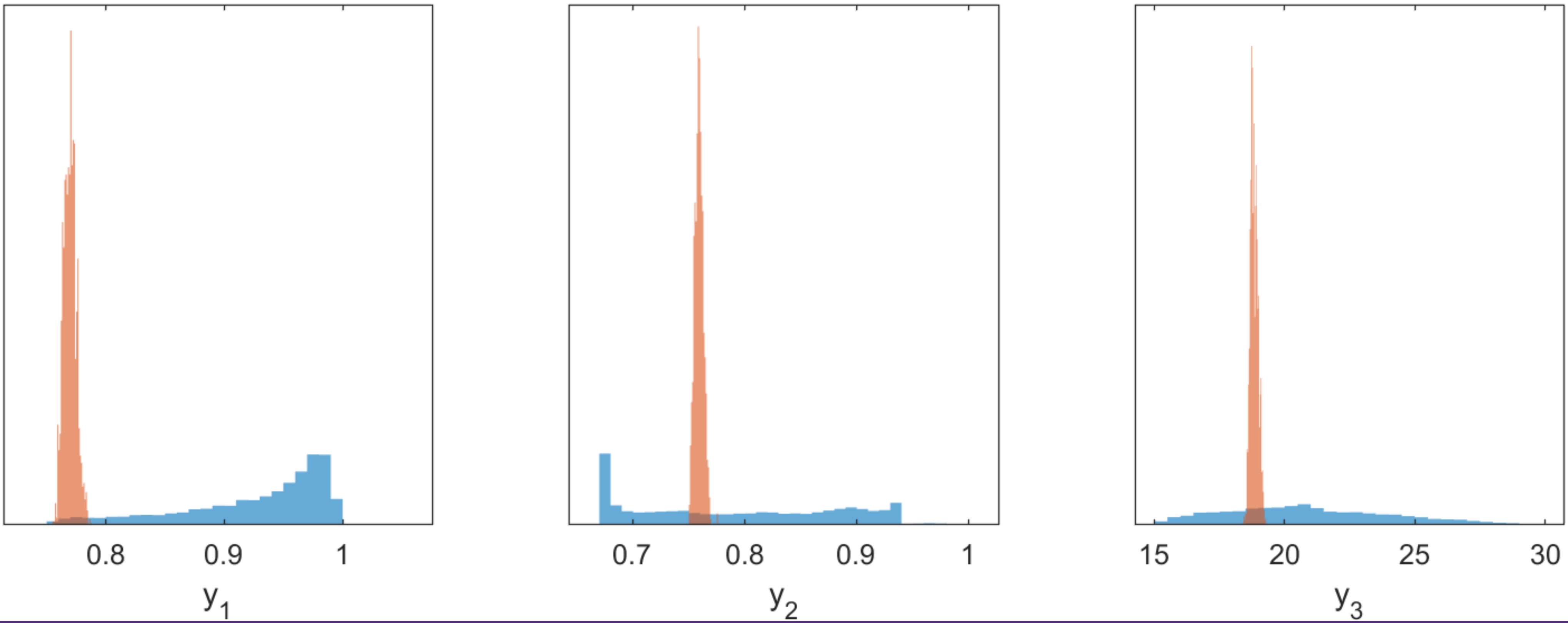
$$y_3(x, \theta_1, \theta_2) = 15 + 2\theta_1 + \frac{\theta_2^2}{4}$$

# Results of simulated example

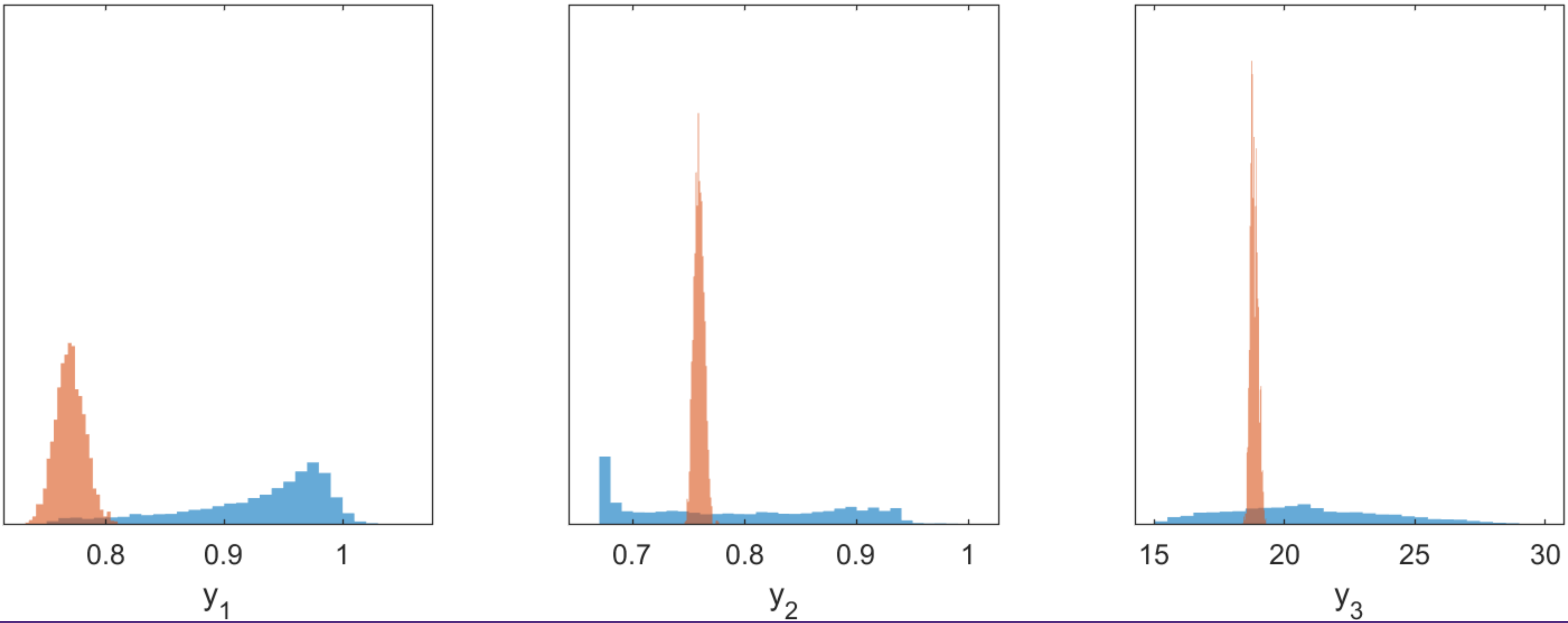
- The target outcome was chosen to be the "utopia point" of the model observations, constant w.r.t. the control input $x$.
- The heatmap represents, for each point in the design space, the distance of the model output from the target outcome (averaged over $x$).
- The posterior distribution approximates the contours of the optimal region.
- The posterior contains information about the sensitivity of the output in the optimal region to the two inputs.



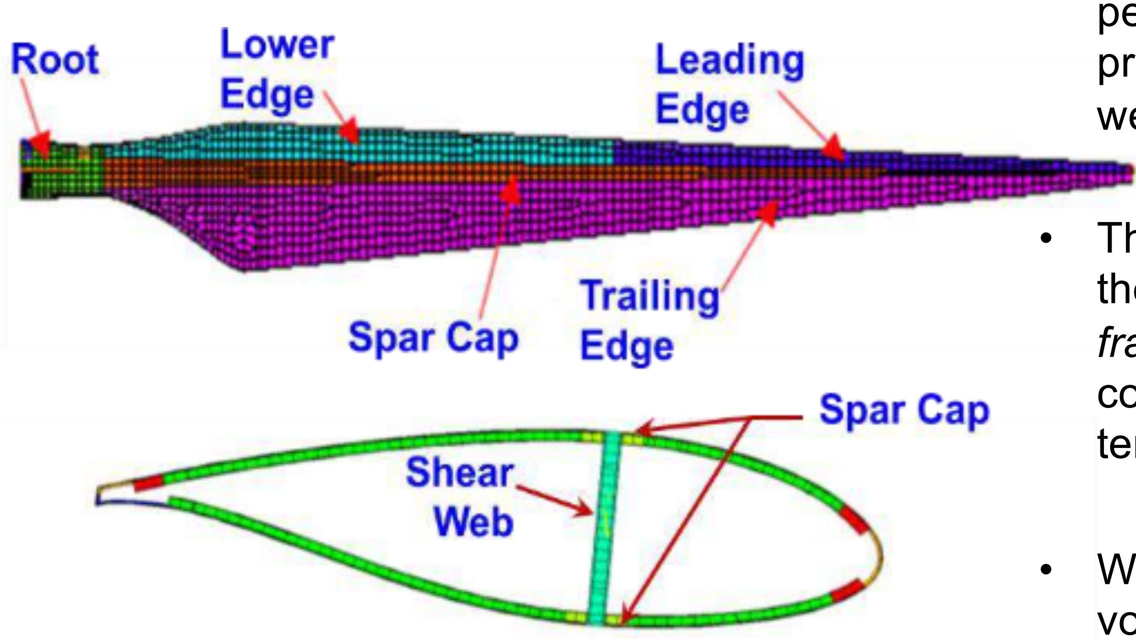
# Prior (blue) and posterior (red) predictive distributions of model output



# Prior (blue) and posterior (red) predictive distributions of model output (including GP uncertainty)



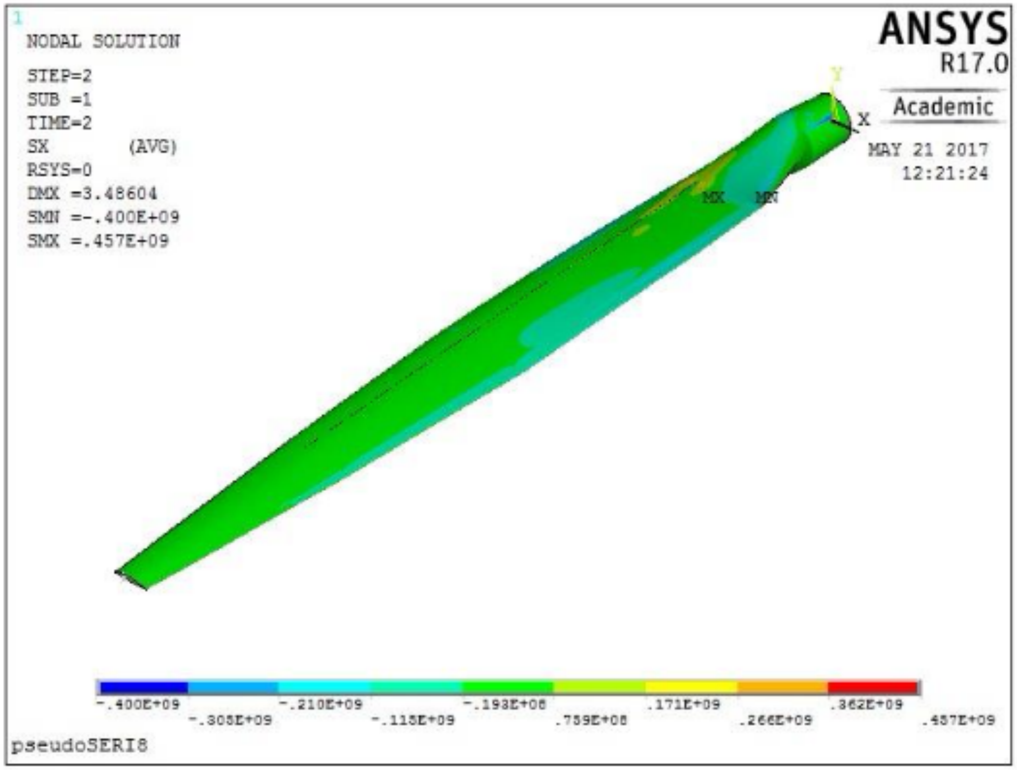
# Application: Material design for a wind turbine blade




- For a wind turbine blade of fixed geometry, its performance and cost depend on the material properties of the composite forming the shear web.
- These material properties in turn depend on the shear web thickness and the *volume fraction* (ratio of the two materials in the composite), as well as on the operating temperature.
- We wish to select a thickness (in mm) and volume fraction to minimize blade rotation (in radians), tip deflection (in meters), and cost per square meter (USD).

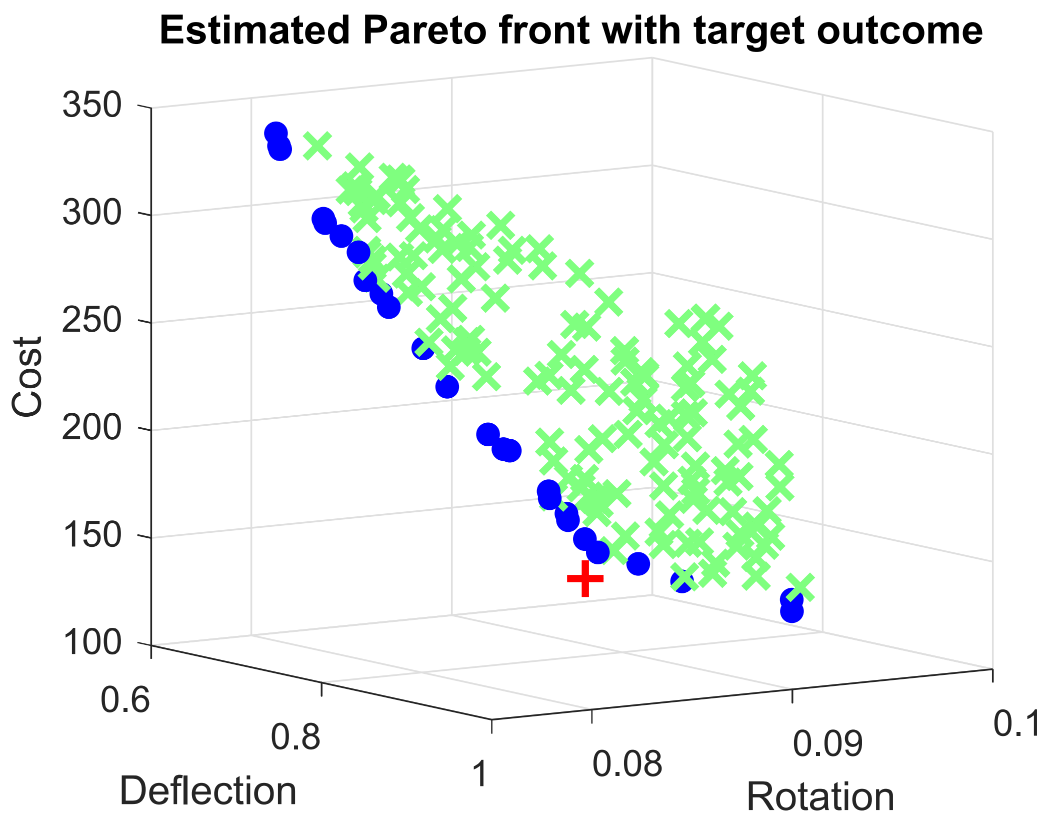
# Finite element model

- The basis for the design process is a finite element model (written by Evan Chodora).
- Due to the FE model's computational expense, we form a GP meta-model using a space-filling design of 504 observations.

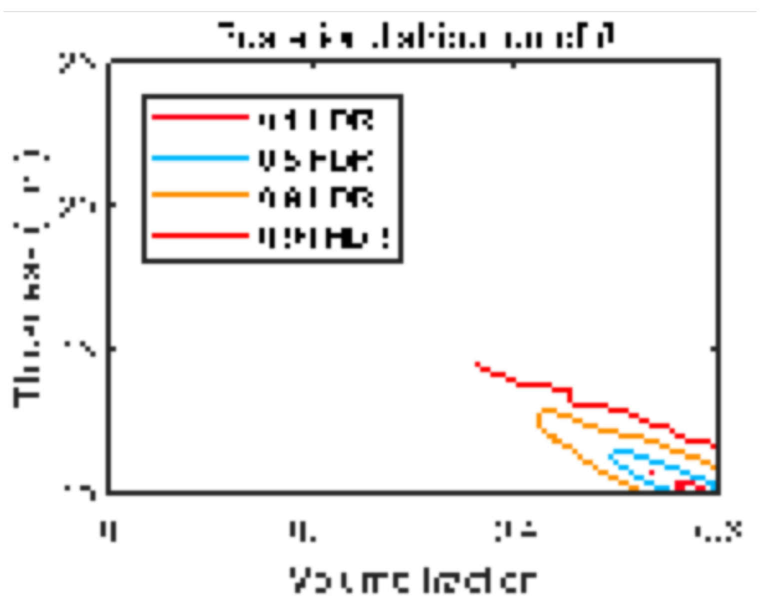


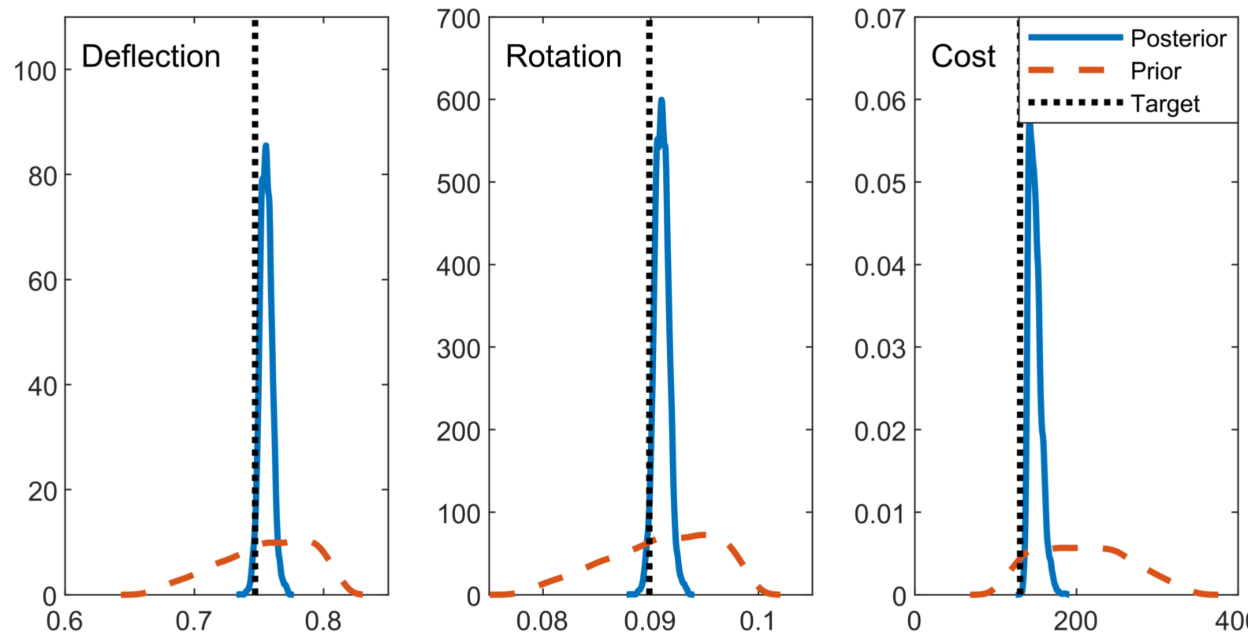
# Selecting a target outcome via preliminary CTO

- To select a target outcome, we first performed CTO with a deliberately vague prior on $\delta()$ and initial target [0, 0 ,0] to get a rough estimate of the Pareto front.
- Doing so, we find an “elbow” representing a point of diminishing returns in trading performance for cost.
- We select a target near this elbow.

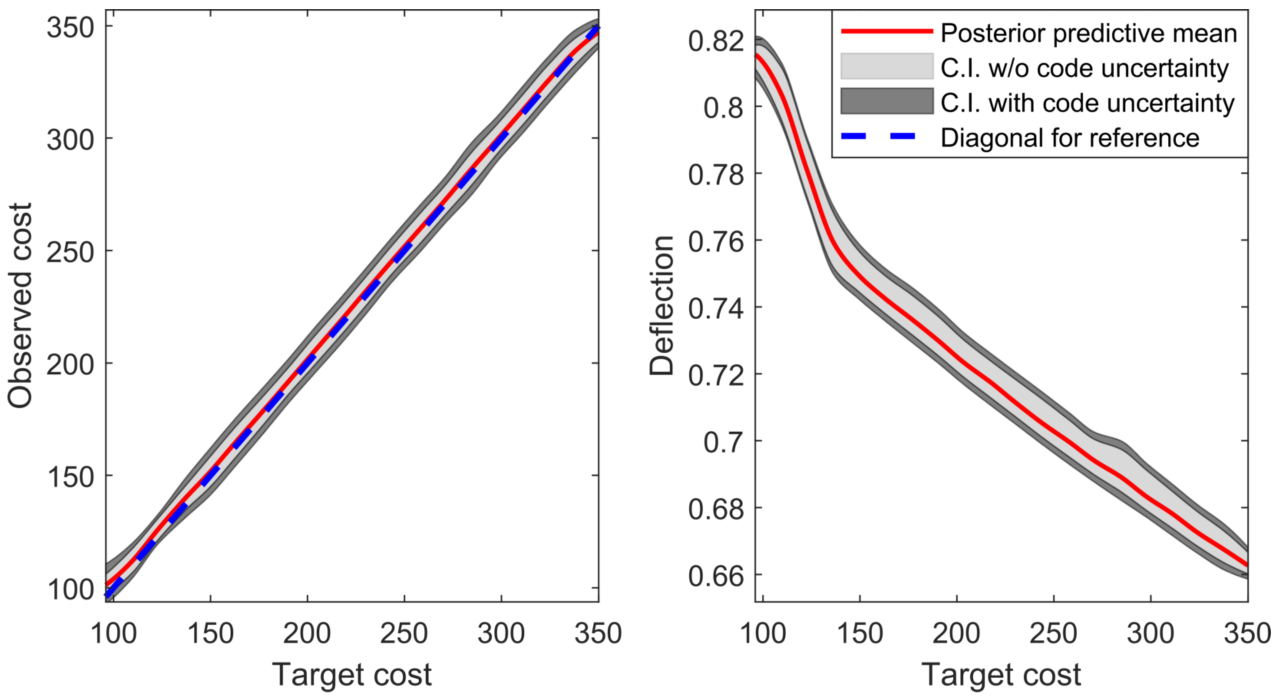


# Prior and posterior distributions of design settings and outcomes for the wind turbine application





# Pareto bands




- By performing CTO many times on a grid of targets, one can form a comprehensive estimate of the Pareto front.
- The estimate here is based on 20 iterations of CTO over a grid of costs.
- The resulting estimate quantifies remaining uncertainty about the location of the front.