Hybrid Machine Learning for Design of Experiments

Dorina Weichert | 05.07.2022

Agenda

Main Ideas

1. Design of Experiments
2. Bayesian Models and sequential Design of Experiments
3. A real example: estimating fatique strength
4. Conclusion

Fraunhofer IAIS

Design of Experiments

Generating costly data

The industrial problem

In R&D, real life experiments are taken out, where each experiment costs a lot of time and money (~ k€).

Can we be more flexible and reduce the costs using ML-methods?

How we cannot solve the problem:

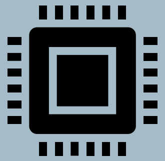
Refining Traditional DoE-Methods:

✖ strong assumptions and/or crazy heuristics

Using pure heuristic Machine Learning:

✖ Data is costly.

A good solution uses both:

expert knowledge + 

machine learning = 

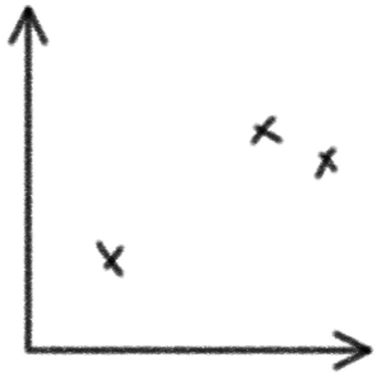
experiment

Fraunhofer IAIS

A solution

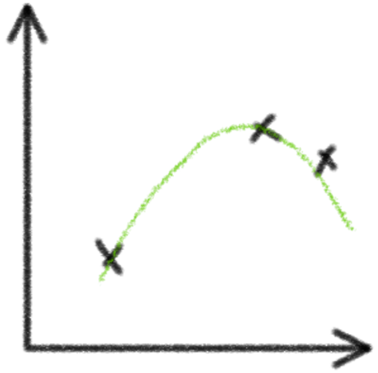
Think Bayes!

Experimental data



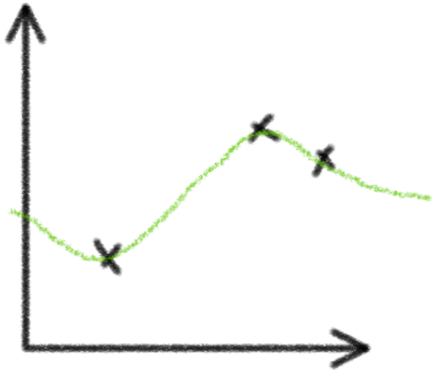
Initialisation
What the engineers in the lab measured once.

ML Model



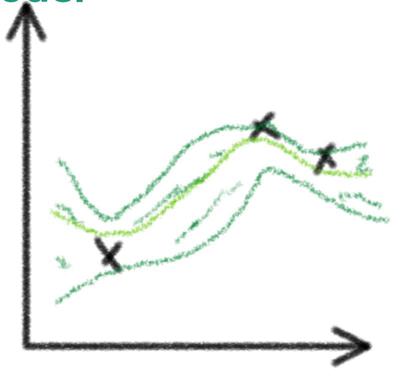
Pure Machine Learning
What a data scientist could do with the data.

Informed ML Model



Informed Machine Learning
An expert told the data scientist that the response always falls back to a constant value.

Probabilistic Informed ML Model

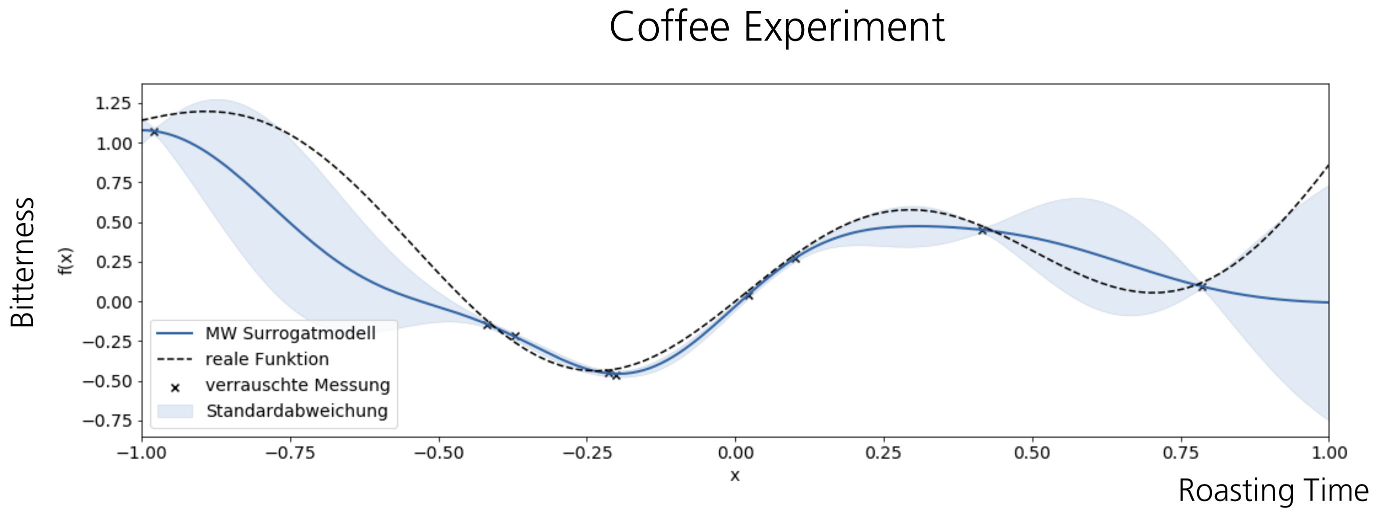


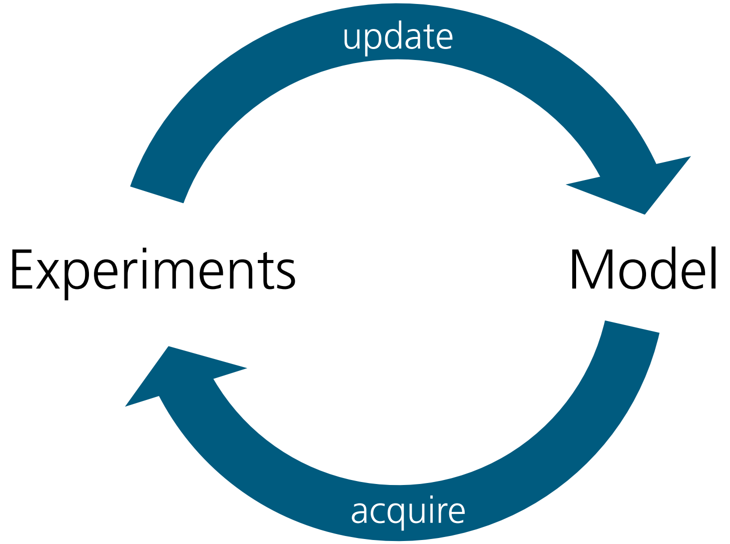
Probabilistic Machine Learning
A data scientist that does not fully trust the expert.

Active Learning, Bayesian Optimization and the sequential DoE Process

Think Bayes!

Use acquisition functions to maximize information in relevant areas of your process.
Here: search for the minimum.





 07.05.2022

Fraunhofer IAIS

A real project: overview and goal

Experiments to determine component and material properties are **complex, time-consuming and expensive**.

We have already successfully used **machine learning for data-driven prediction** of fatigue properties of high-strength steels.

In this project, we use uncertainty-aware ML methods to **predict the 50% fatigue strength and its scatter** from material and component properties.

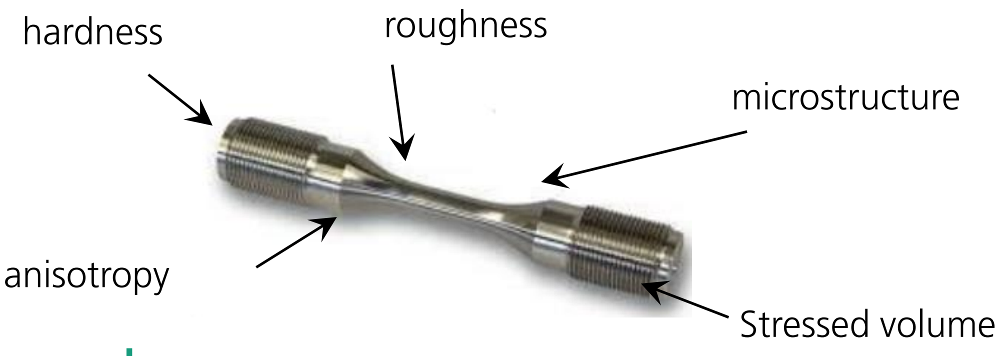
We use this prediction as a prior distribution and derive a recommendation for the load to be tested in the following experiment (**design-of-experiments**).

This recommendation is chosen in such a way that it optimally improves the fatigue strength estimation and **minimizes the number of necessary experiments**.

With each experiment, the prediction is refined.

Fraunhofer IAIS

The problem



Task:

For a specific material, the fatique strength is log-normally distributed. Find a method to estimate the mean of this distribution at a given precision for a new material.

Given:

Mean of the distributions for different materials. (Some) full experimental series. (material -> loads -> failure/survivor)

Approach:

Material parameters

(Probabilistic) Regression Model

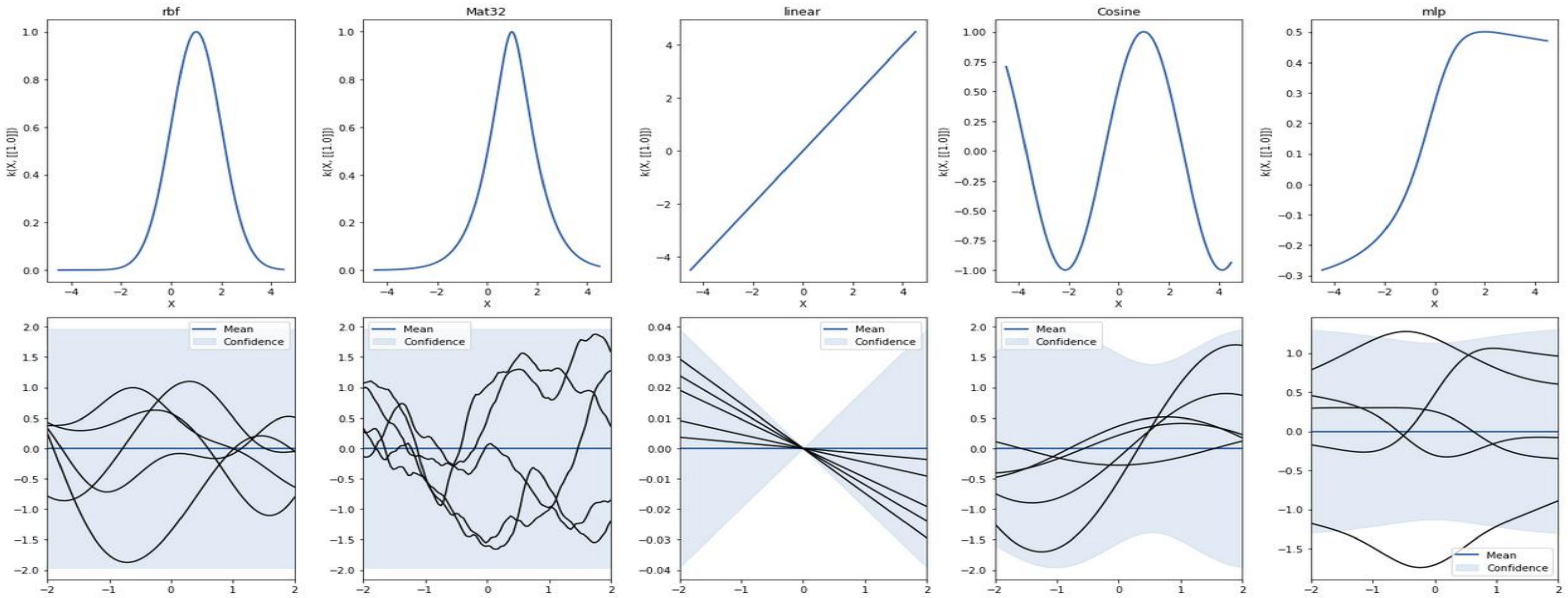
Prior: normal distribution

MAP estimate of mean fatique strength

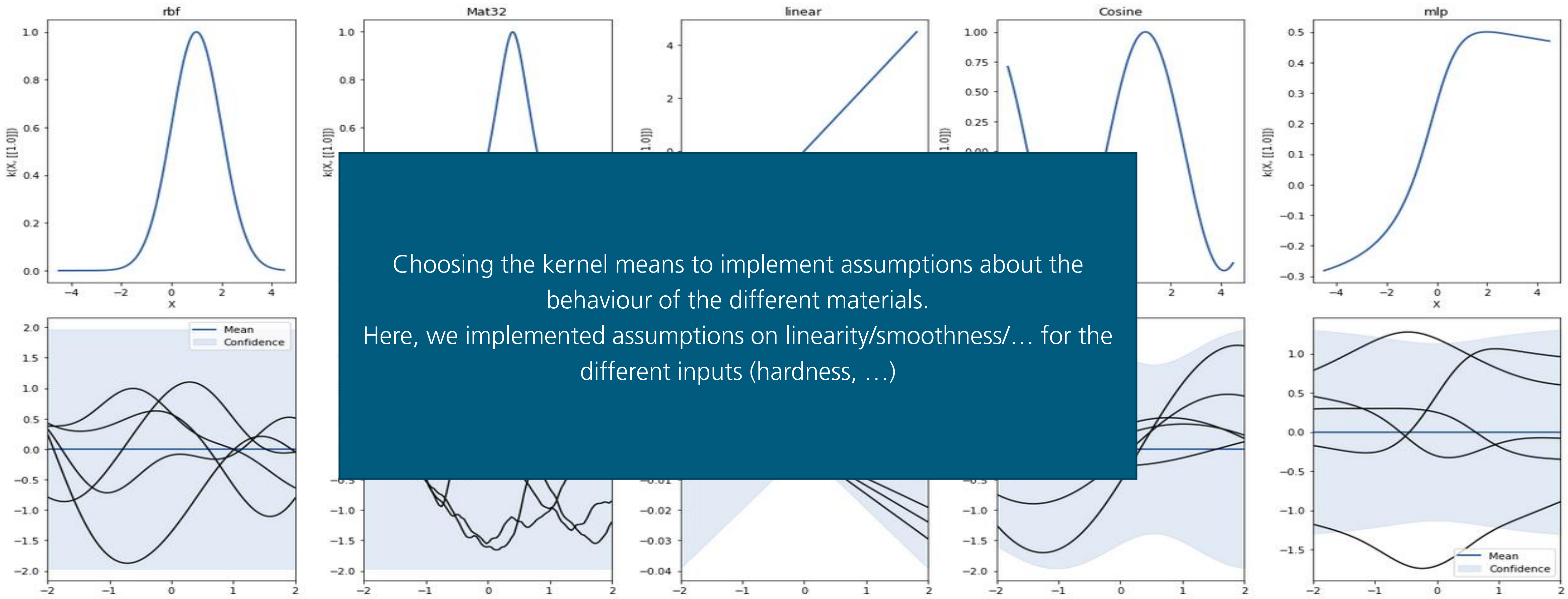
Get point with highest influence on mean fatique strength, update

Fraunhofer IAIS

Covariance functions (kernels)
Examples of covariance functions [Duvenaud, 2014]



Covariance functions (kernels)
Examples of covariance functions [Duvenaud, 2014]

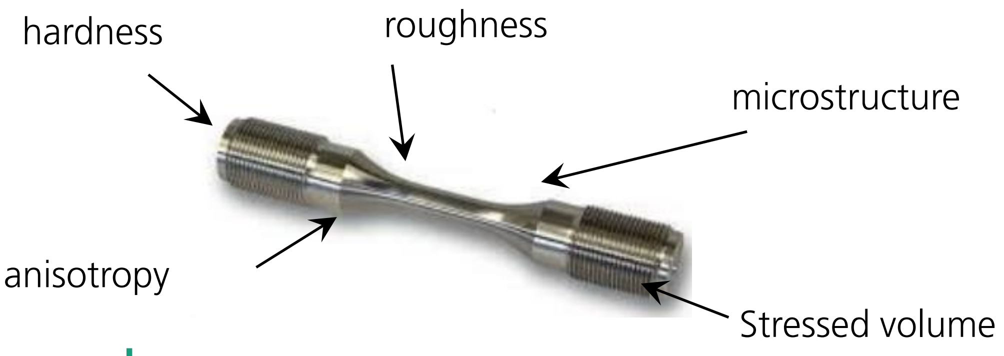


Choosing the kernel means to implement assumptions about the behaviour of the different materials.
Here, we implemented assumptions on linearity/smoothness/… for the different inputs (hardness, …)

offen

Fraunhofer IAIS

The problem



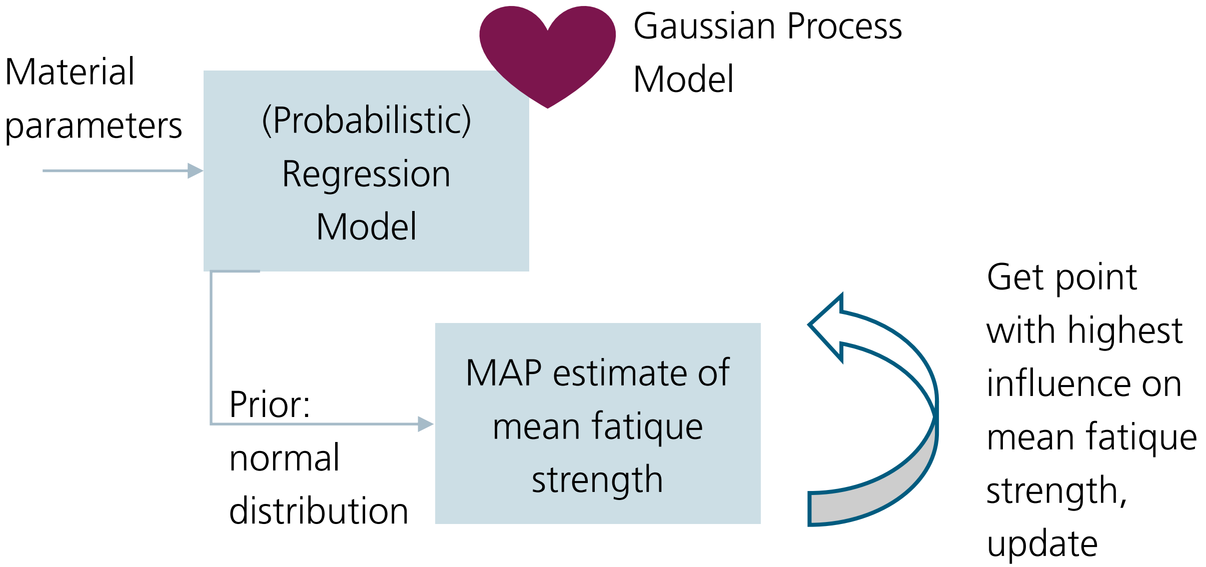
Task:

For a specific material, the fatique strength is log-normally distributed. Find a method to estimate the mean of this distribution at a given precision for a new material.

Given:

Mean of the distributions for different materials. (Some) full experimental series. (material -> loads -> failure/survivor)

Approach:



offen

Fraunhofer IAIS

A real Project Setup

Material parameters → Prob. Regression Model (GP) → Fatique strength μ_m, σ_m

classical Bayes:
Use μ_m, σ_m GPs as prior for the fit of the (log)normally distributed fatique strength in a MAP approach,
search for experiment with largest influence on μ_{MAP}

To convergence

Posterior

Prior $\mu \sim N(\mu_m, \sigma_m)$

Likelihood

$$g(\mu) = f(\mu) \cdot f(s_i, s_j | \mu) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-{}^1/_2 \left(\frac{\mu - \mu_m}{\sigma_m}\right)^2\right) \cdot \left(\prod_i \Phi_{\mu,\sigma}(s_i) \cdot \left(\prod_j 1 - \Phi_{\mu,\sigma}(s_j)\right)\right)$$

MAP estimate $\quad \mu_{MAP} = argmax_\mu \, g(\mu)$

New experiment $\quad s^* = argmax_s \frac{dg}{ds} = \mu_{MAP}$

Uncertainty of MAP $\quad d\mu_{MAP} = Std[g(\mu)]$

failures *i*, survivors *j*
load ***s***
Assumption: normal distribution of fatique strength

Fraunhofer IAIS

A real Project Setup

Material parameters → Prob. Regression Model (GP) → Fatique strength μ_m, σ_m

classical Bayes:
Use μ_m, σ_m GPs as prior for the fit of the (log)normally distributed fatique strength in a MAP approach,
search for experiment with largest influence on μ_{MAP}

To convergence

Posterior

$g(\mu) = f(\mu) \cdot f(s_i, s$ … $\prod_j \left[1 - \Phi_{\mu,\sigma}(s_j)\right]\Big)\Big)$

…od

The posterior arised from the set-up: being a failure or not is a bernoulli-distributed variable, the fatique strength is (like assumed in the GP) normally distributed.

MAP esti…

failures *i*, survivors *j*
load *s*
Assumption: normal distribution of fatique strength

New experiment $s^* = argmax_s \frac{dg}{ds} = \mu_{MAP}$

Uncertainty of MAP $d\mu_{MAP} = Std[g(\mu)]$

Fraunhofer IAIS

The problem

Approach

Material parameters

(Probabilistic) Regression Model

Gaussian Process Model

Prior: normal distribution

MAP estimate of mean fatique strength

Likelihood

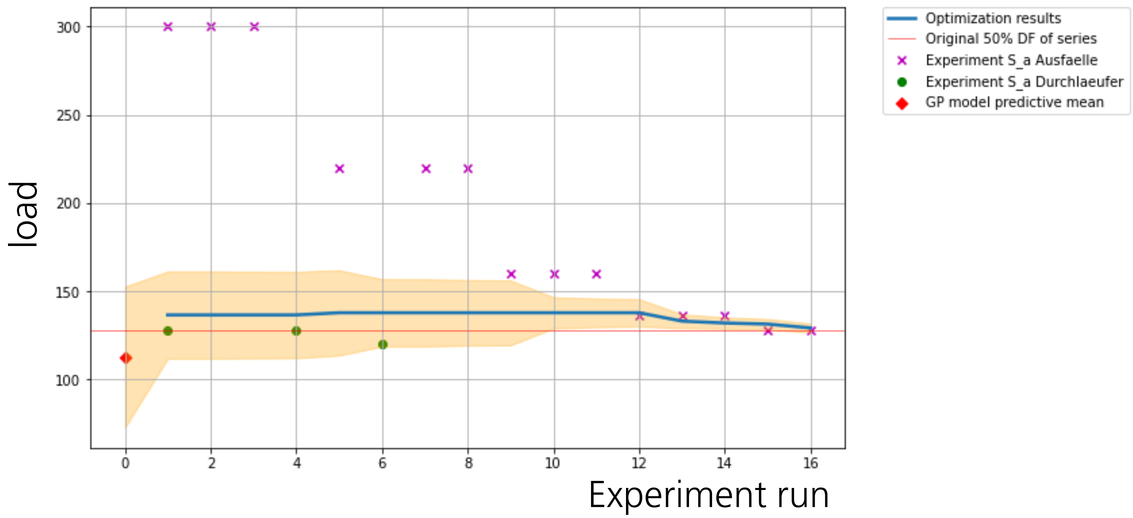
Get point with highest influence on mean fatique strength, update

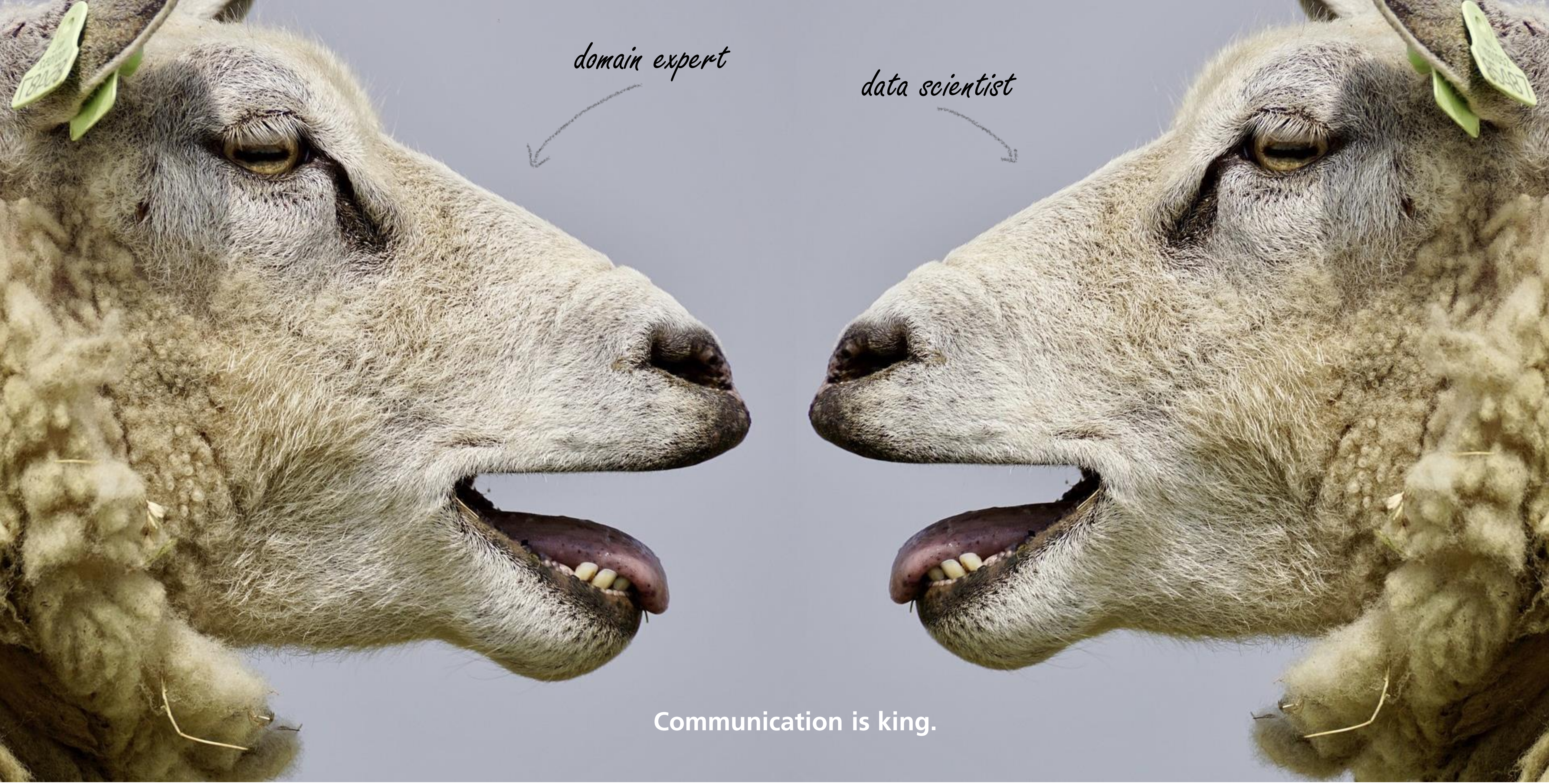
Results

With historical data, we could estimate that our approach reduces the amount of necessary experiments by > 25 %.

In real life, that would mean a reduction of ~ 60.000 €.

A feasibility study will follow.




domain expert
data scientist
Communication is king.

Contact

Dorina Weichert
Data Scientist
Tel. +49 2241 14-2284
dorina.weichert@iais.fraunhofer.de

Fraunhofer-Institut für intelligente
Analyse- und Informationssysteme IAIS
Schloss Birlinghoven 1
53757 Sankt Augustin

www.iais.fraunhofer.de