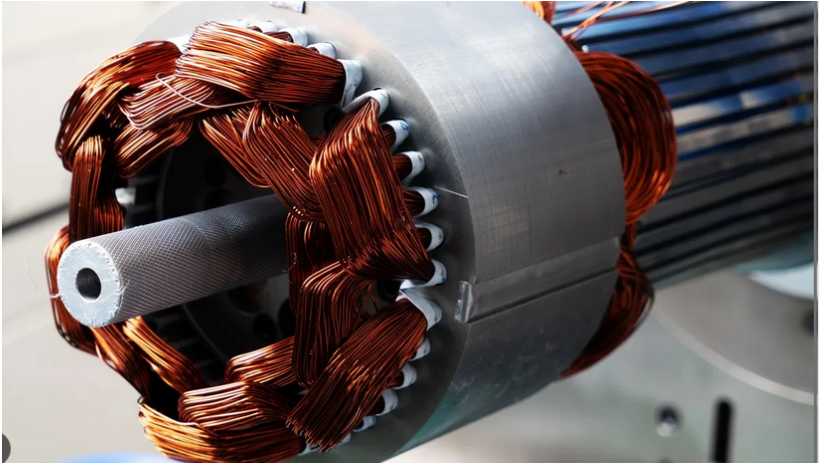
Bayesian Experiment Design for the Development of a Degradation Model

Jan Kaska, Václav Šmídl, Jan Leffler, Pavel Trnka

University of West Bohemia, Faculty of Electrical Engineering

May 26, 2023

Motivation

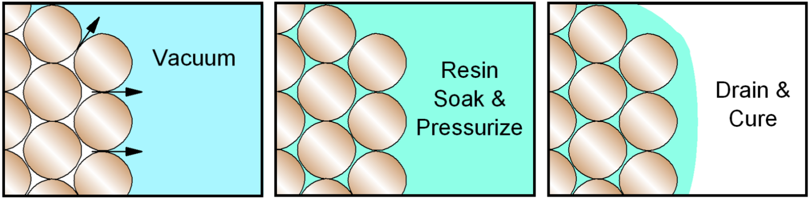


1

[1] Marshall Brain & Kristen Hall-Geisler. *How Electric Motors Work*. https://electronics.howstuffworks.com/motor.htm. 2021.

Motivation

Winding of electrical motors and transformers



2

[2] Teledyne Hastings Instruments. *Vacuum Pressure Impregnation (VPI) Systems*. https://www.teledyne-hi.com/resourcecenter/Application%20Notes/vacuum%20pressure%20impregnation%20system.pdf. 2021.

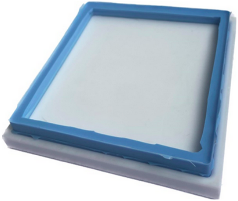
Motivation

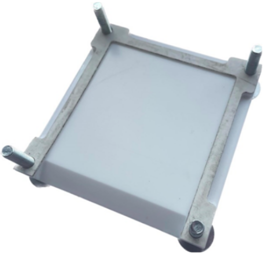
- Failure of the electrical insulation system (EIS) is one of the most common electrical machine faults[3]
- EIS ageing is caused by many factors including temperature, humidity, mechanical wear, dielectric stress, UV, radiation, ...
- There is need for a model of behaviour of the material under stress → hygrothermal aging

[3] Jan Leffler and Pavel Trnka. "Failures of Electrical Machines-Review". In: *2022 8th International Youth Conference on Energy (IYCE)*. IEEE. 2022, pp. 1–4.

Sample preparation

1. Pour resin into the casts (10 x 10 x 1.5)
2. Get rid of any impurities and bubbles
3. Cure the resin in the oven





Aging and Measurement

1. Age samples in a climate chamber for at least 14 days
2. Measure 5 electrical breakdowns on each of the 5 samples
3. Measure sample thickness in point of breakdown

Goal

- Measurement of one experimental setup takes at least 16 days
- There are many models that describe behaviour under stress
- We need to find a suitable model with as few measurements as possible $\rightarrow$ Design of Experiments

- Measured data:
  $\mathcal{D} = \{\boldsymbol{x}_i, y_i\}_{i=1}^{n}, \quad \boldsymbol{x} = [\vartheta, \mathrm{RH}], \quad y = \ln(E_{\mathsf{p}})$
- Models linear-in-parameters: $\boldsymbol{y} = \boldsymbol{\theta}^\top \Phi(\boldsymbol{x}) + \boldsymbol{e}$
- Basis functions $\Phi(\boldsymbol{x})$ are known

Models

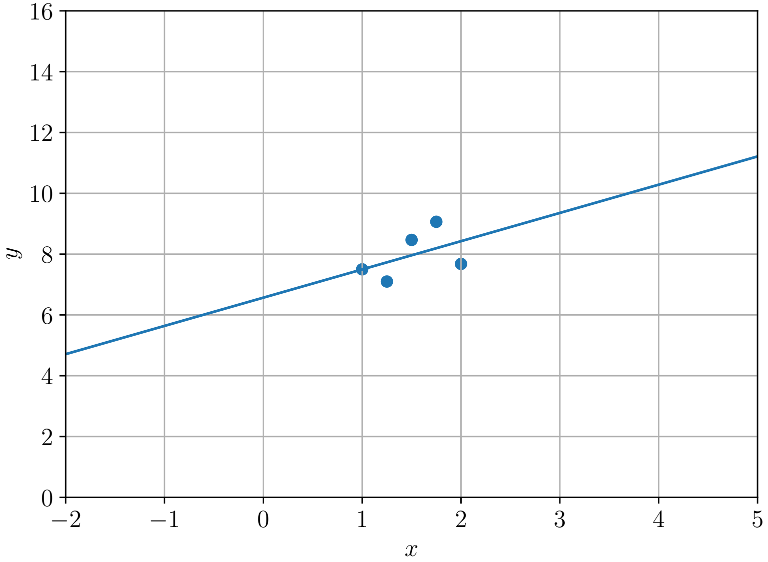
$$\mathcal{M}_1 : \boldsymbol{\phi}_1(\boldsymbol{x}) = [1,\ \vartheta,\ \mathrm{RH}]$$
$$\mathcal{M}_2 : \boldsymbol{\phi}_2(\boldsymbol{x}) = [1,\ \vartheta,\ \mathrm{RH},\ \vartheta\,\mathrm{RH}]$$
$$\mathcal{M}_3 : \boldsymbol{\phi}_3(\boldsymbol{x}) = \left[1,\ \vartheta,\ \mathrm{RH},\ \vartheta^2,\ \mathrm{RH}^2\right]$$
$$\mathcal{M}_4 : \boldsymbol{\phi}_4(\boldsymbol{x}) = \left[1,\ \vartheta,\ \mathrm{RH},\ \vartheta^2,\ \mathrm{RH}^2,\ \vartheta^2\,\mathrm{RH}^2\right]$$

$$\mathcal{M}_1 : E_{\mathrm{p}} = e^{\theta_0} e^{\theta_1 \vartheta} e^{\theta_2 \mathrm{RH}}$$

Ordinary least squares

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta} \|\boldsymbol{\theta}^\top \Phi(\boldsymbol{x}) - \boldsymbol{y}\|^2$$
$$= \left(\Phi(\boldsymbol{x})^\top \Phi(\boldsymbol{x})\right)^{-1} \Phi(\boldsymbol{x})^\top \boldsymbol{y}$$

OLS



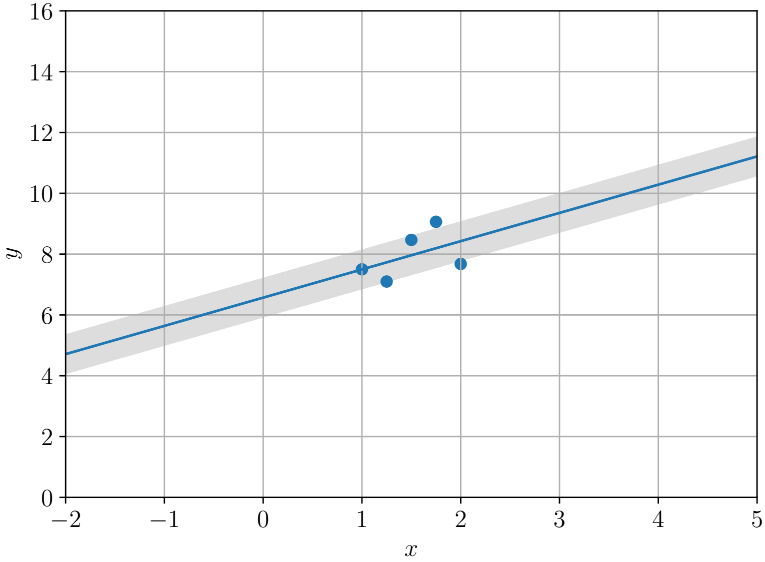
OLS

Ordinary least squares - probabilistic perspective

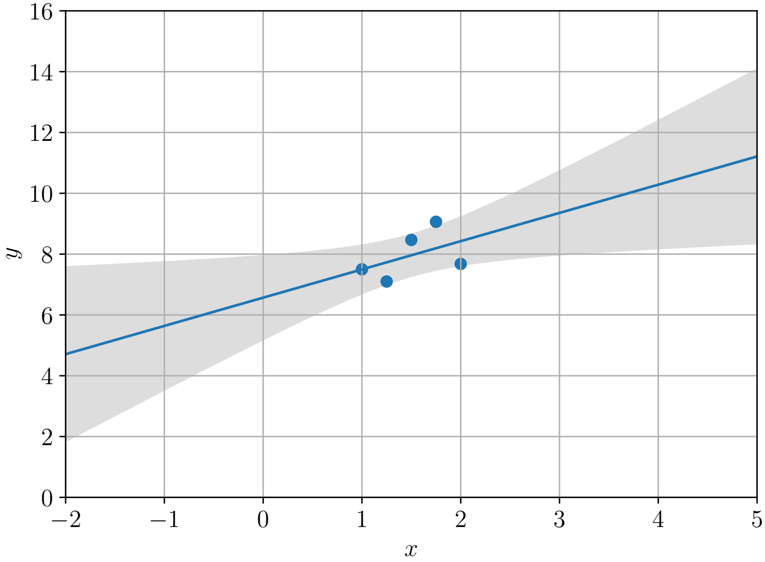
$$p(y|\boldsymbol{\theta}, \boldsymbol{x}) = \mathcal{N}(\boldsymbol{\theta}^\top \Phi(\boldsymbol{x}), \sigma_{\mathcal{D}}^2 I)$$

$$\sigma_{\mathcal{D}}^2 = \frac{1}{N} \sum_{n=1}^{N} (y_n - \boldsymbol{\theta}^\top \boldsymbol{\phi}(\boldsymbol{x}_n))^2$$

OLS



OLS?



Bayes rule

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

$$p(\theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)p(\theta)$$

Prior distribution of the parameters

$$p(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}_0, \sigma_0^2 I)$$

The posterior has Gaussian form $p(\boldsymbol{\theta}|\mathcal{D}) = \mathcal{N}(\bar{\boldsymbol{\theta}}, \boldsymbol{S})$

$$\bar{\boldsymbol{\theta}} = \frac{1}{\sigma_{\mathcal{D}}^2} \boldsymbol{S} \Phi(\boldsymbol{x})^\top \boldsymbol{y}$$

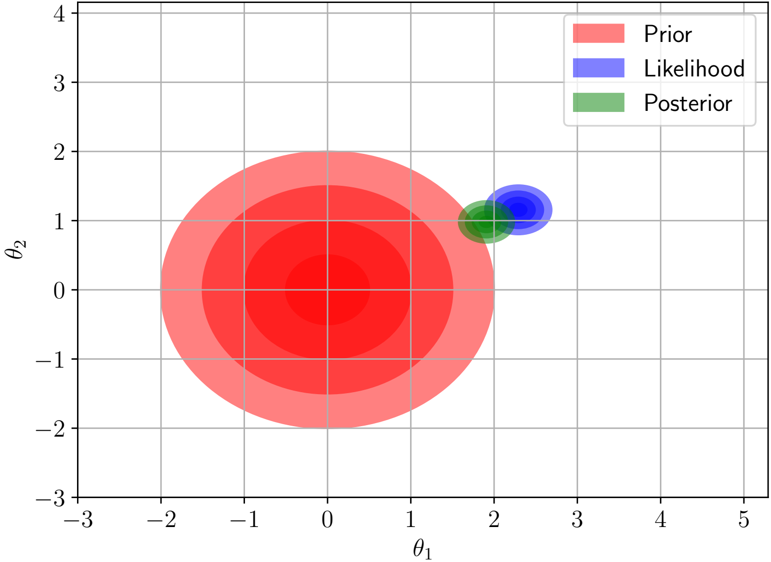
$$\boldsymbol{S}^{-1} = \frac{1}{\sigma_0^2} \boldsymbol{I} + \frac{1}{\sigma_{\mathcal{D}}^2} \Phi(\boldsymbol{x})^\top \Phi(\boldsymbol{x})$$

For $\sigma_0 \to \infty$ the mean $\bar{\boldsymbol{\theta}}$ corresponds to OLS solution

$$\boldsymbol{S}^{-1} = \frac{1}{\sigma_0^2}\boldsymbol{I} + \frac{1}{\sigma_{\mathcal{D}}^2}\Phi(\boldsymbol{x})^\top\Phi(\boldsymbol{x}) \quad \to \quad \frac{1}{\sigma_{\mathcal{D}}^2}\Phi(\boldsymbol{x})^\top\Phi(\boldsymbol{x})$$

$$\bar{\boldsymbol{\theta}} = \frac{1}{\sigma_{\mathcal{D}}^2}\boldsymbol{S}\Phi(\boldsymbol{x})^\top\boldsymbol{y} \quad \to \quad \Big(\Phi(\boldsymbol{x})^\top\Phi(\boldsymbol{x})\Big)^{-1}\Phi(\boldsymbol{x})^\top\boldsymbol{y}$$

Bayes rule

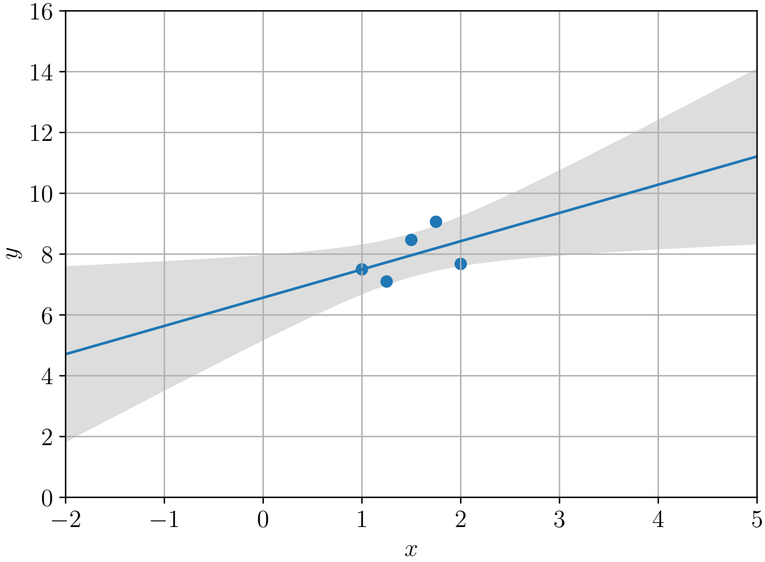

Prior
Likelihood
Posterior
θ_2
θ_1
4
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3
4
5

Predictions for new $\boldsymbol{x}$

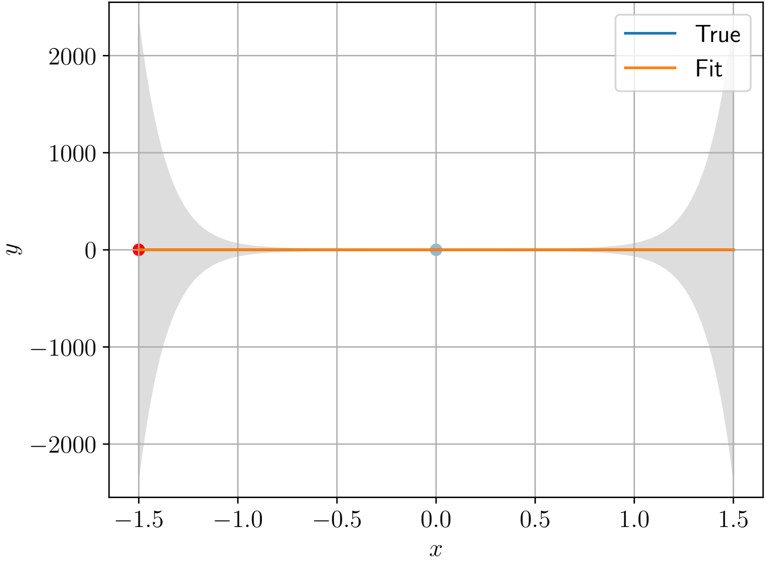
$$p(y|\boldsymbol{x}, \mathcal{D}) = \mathcal{N}\left(\bar{\boldsymbol{\theta}}^{\top}\Phi(\boldsymbol{x}), \boldsymbol{\Sigma}(\boldsymbol{x})\right)$$

$$\boldsymbol{\Sigma}(\boldsymbol{x}) = \sigma_{\mathcal{D}}^{2} + \Phi(\boldsymbol{x})^{\top}\boldsymbol{S}\Phi(\boldsymbol{x})$$

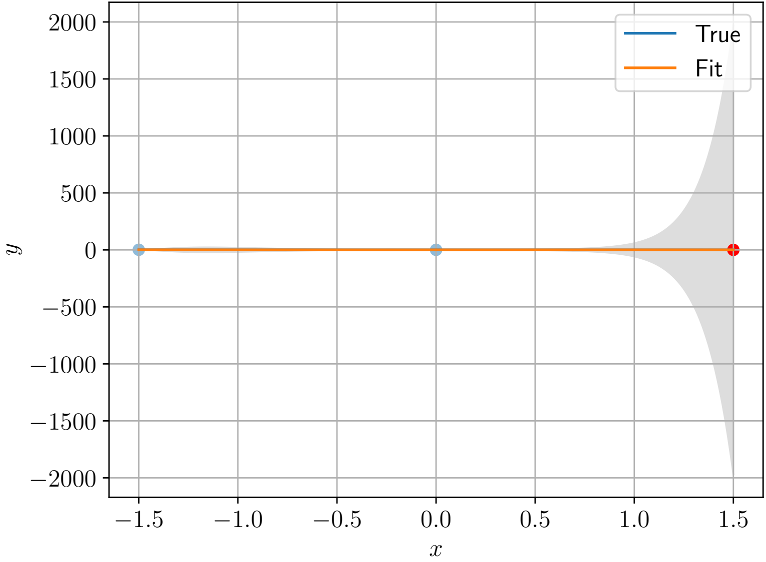
Bayes rule



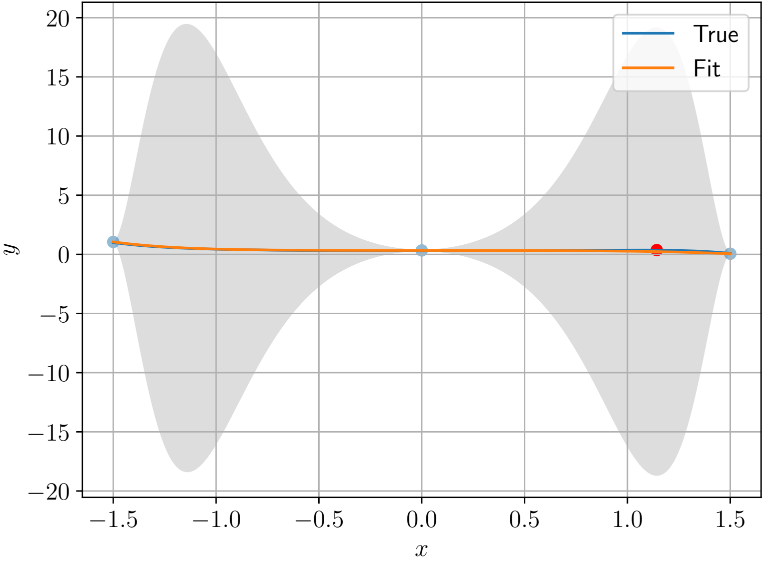
Uncertainty



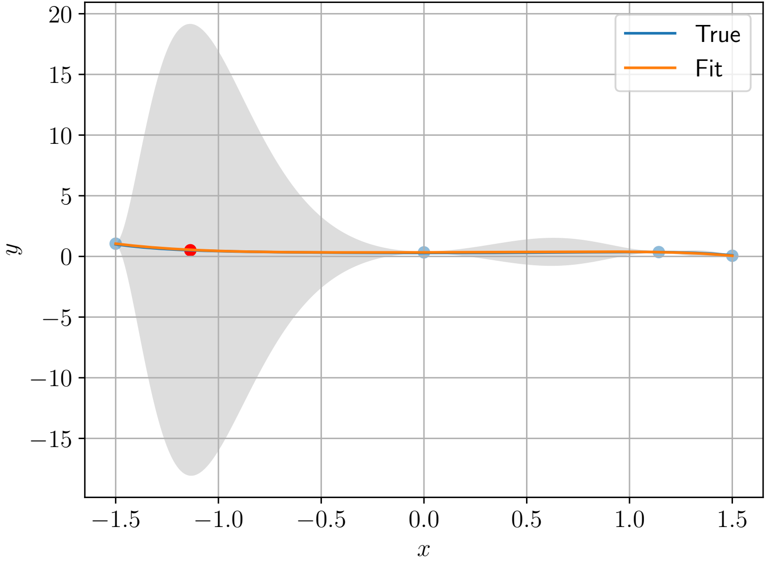
Uncertainty



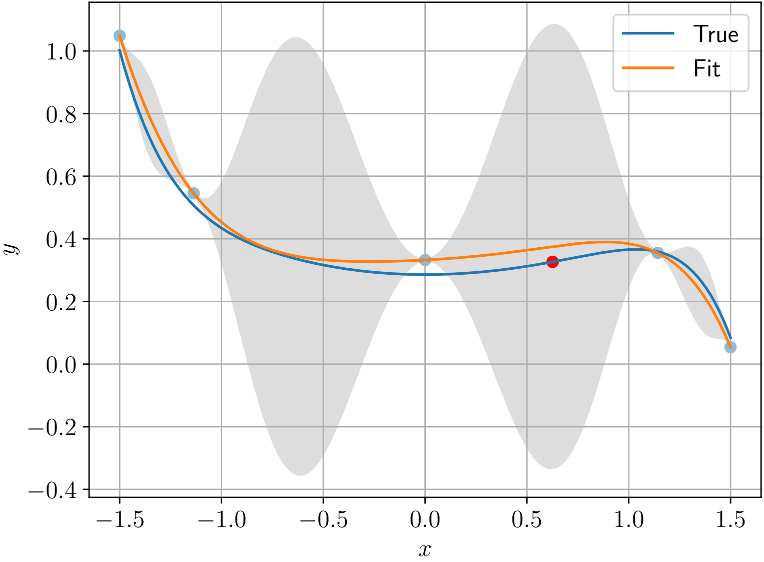
Uncertainty



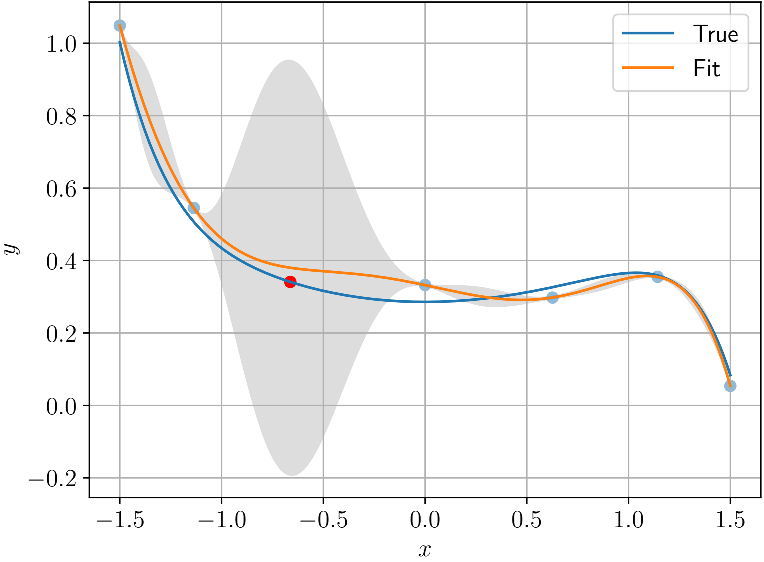
Uncertainty



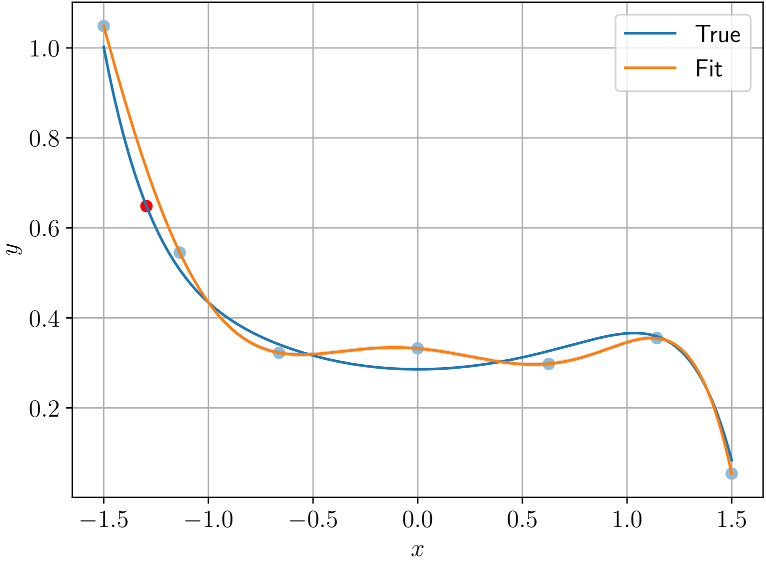
Uncertainty



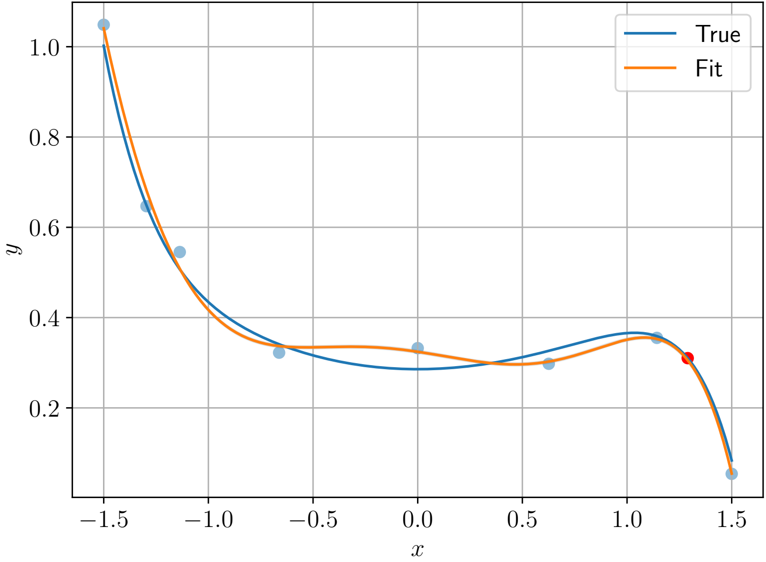
Uncertainty



Uncertainty



Uncertainty



Model selection

Multiple models

Model selection

Probability that data $\mathcal{D}$ is generated from model $\mathcal{M}_m$

$$p(\mathcal{M}_m|\mathcal{D}) \propto p(\boldsymbol{y}|\mathcal{M}_m, \Phi_m(\boldsymbol{x}))p(\mathcal{M}_m) \ \forall m = 1 \ldots 4$$

where $p(\mathcal{M}_m) = 1/4$ is the prior probability of each model

Model selection

$$\begin{aligned}\ln p(\boldsymbol{y}|\mathcal{M}_m, \Phi(\boldsymbol{x})) = & -d_m \ln \sigma_{0,m} + n \ln \sigma_{\mathcal{D},m} \\ & - \mathrm{E}(\bar{\boldsymbol{\theta}}_m) - \frac{1}{2} \ln \left(\boldsymbol{S}_m^{-1} \right) - \frac{n}{2} \ln(2\pi) \\ \mathrm{E}(\bar{\boldsymbol{\theta}}_m) = & \frac{1}{2\sigma_{\mathcal{D},m}^2} \|\bar{\boldsymbol{\theta}}_m^\top \Phi_m(\boldsymbol{x}) - \boldsymbol{y}\|^2 + \frac{1}{2\sigma_{0,m}^2} \bar{\boldsymbol{\theta}}_m^\top \bar{\boldsymbol{\theta}}_m\end{aligned}$$

Model selection

$$w_m = \frac{p(\boldsymbol{y}|\mathcal{M}_m, \Phi(\boldsymbol{x}))}{\sum_{m=1}^{4} p(\boldsymbol{y}|\mathcal{M}_m, \Phi(\boldsymbol{x}))}$$

Bayesian model averaging (BMA)

$$p(\hat{y}|\boldsymbol{x}, \mathcal{D}) = \sum_{m=1}^{4} w_m p(\hat{y}|\boldsymbol{x}, \mathcal{D}, \mathcal{M}_m)$$

$$\begin{aligned}
\mathrm{var}(\boldsymbol{y}'|\boldsymbol{x}', \mathcal{D}) &= \mathrm{E}(\boldsymbol{y}'\boldsymbol{y}') - \mathrm{E}(\boldsymbol{y}')^2 \\
&= \sum_{m=1}^{M} w_m \left(\hat{y}_m \hat{y}_m + \boldsymbol{\Sigma}_m\right) - \left(\sum_{m=1}^{M} w_m \hat{y}_m\right)^2
\end{aligned}$$

Algorithm 1 Summary of the Bayesian experiment design for for selection of optimal linear model

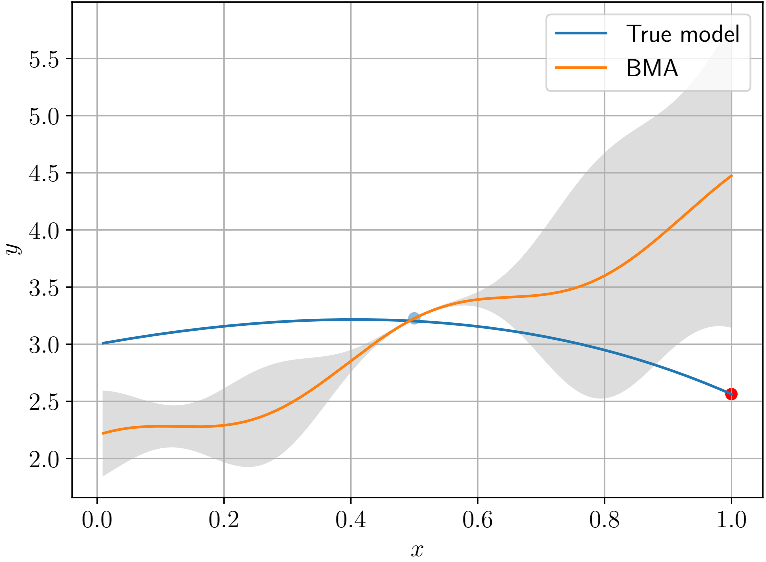
Init: Choose initial samples $\boldsymbol{x}_{(0)} = \boldsymbol{x}$ and measure output variable $\boldsymbol{y}^{(0)}$. Set iteration counter $l = 0$.

Iterate until convergence:

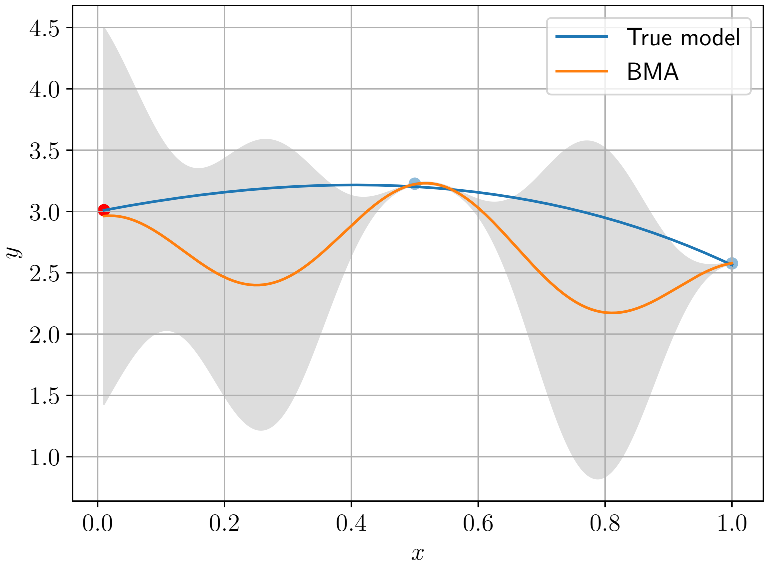
1. For each model $\mathcal{M}_m, m = 1 \ldots 4$ calculate coefficients $\bar{\boldsymbol{\theta}}_m$ and probability w_m for data $\mathcal{D}^{(l)}$
2. On the grid of candidate points $\boldsymbol{x}'$ compute predictive variance
3. Select point $\boldsymbol{x}'$ with highest variance for the next measurement,
4. Perform measurement to obtain y' and extend the data sets: $\mathcal{D}^{(l+1)} = [\mathcal{D}^{(l)}, \boldsymbol{x}', y']$

Return: most likely model and its parameters $\boldsymbol{\theta}_m$

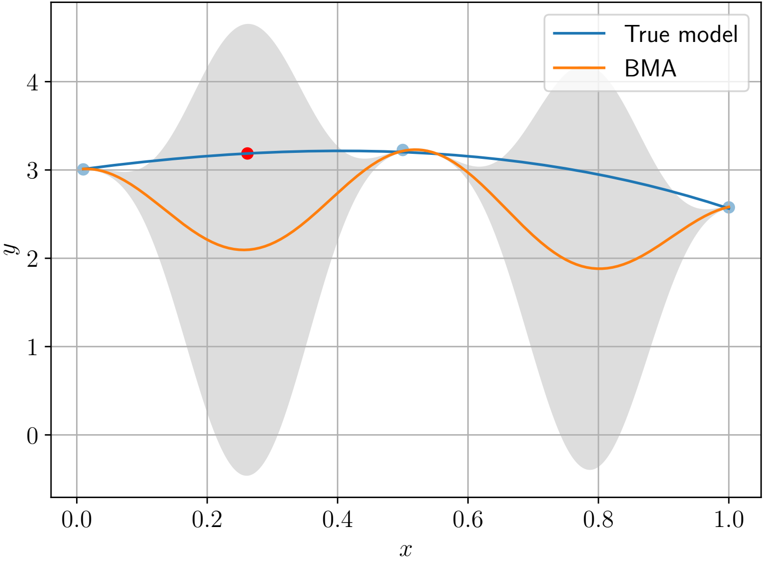
Model selection



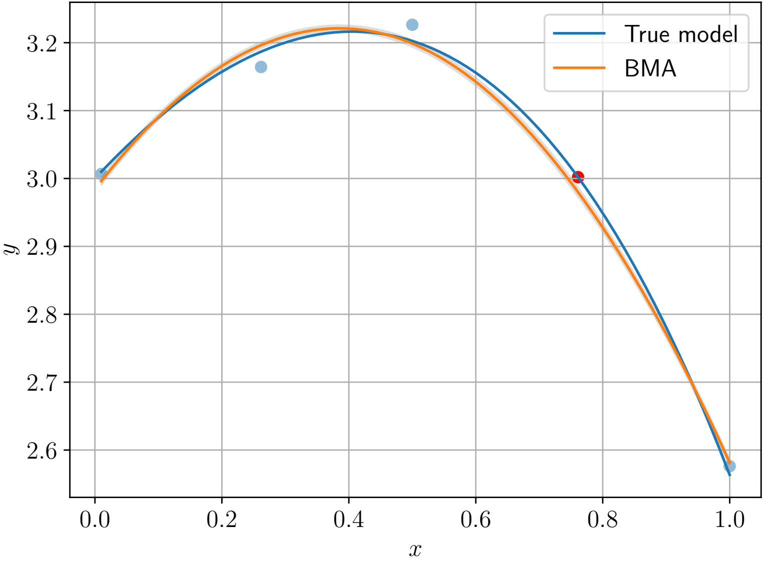
Model selection



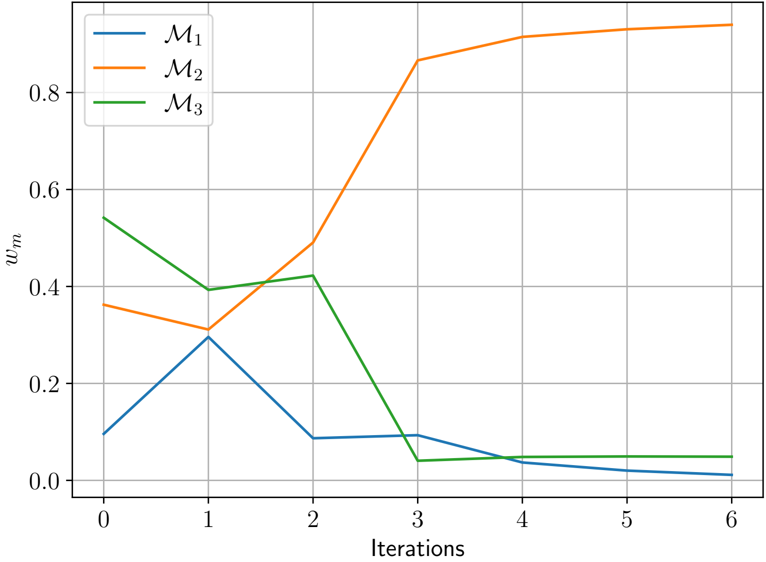
Model selection



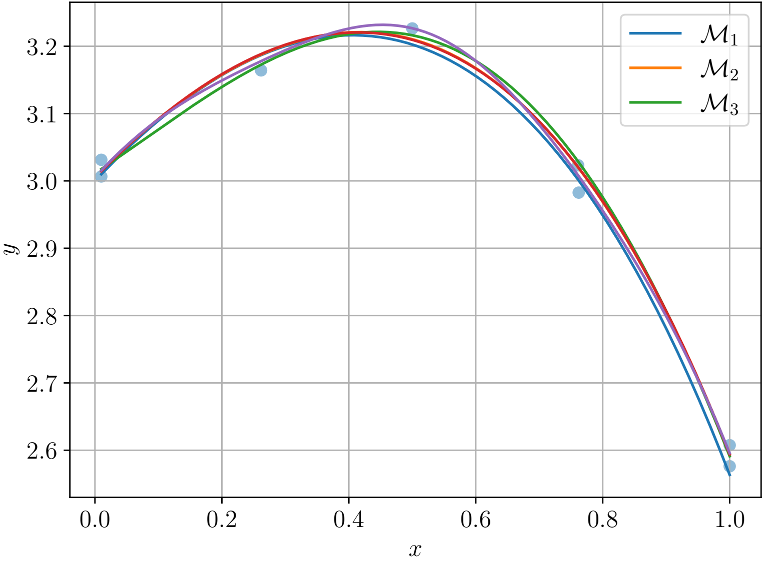
Model selection



Model selection



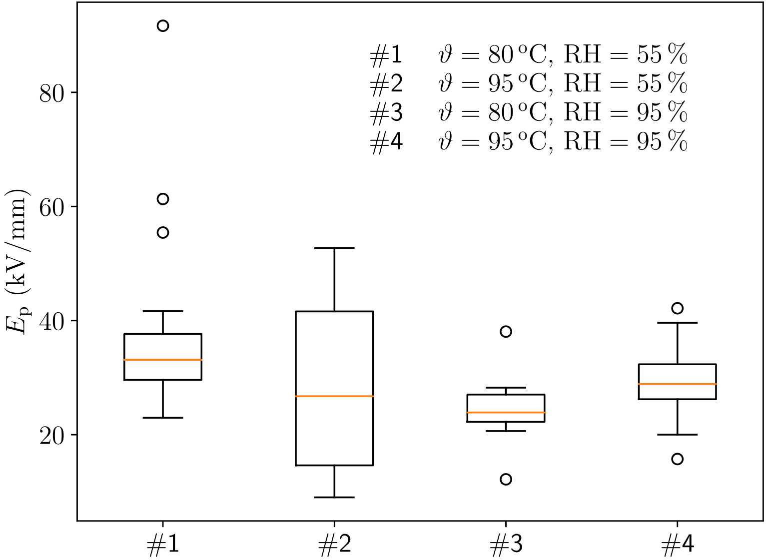
Model selection

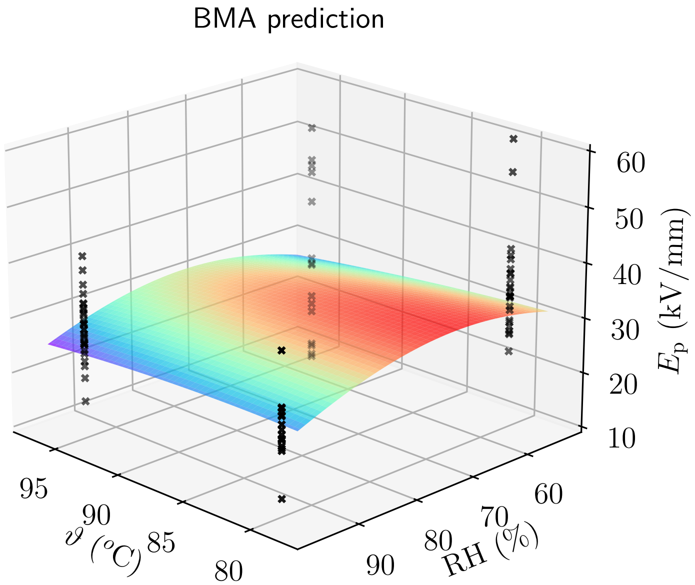


Model selection

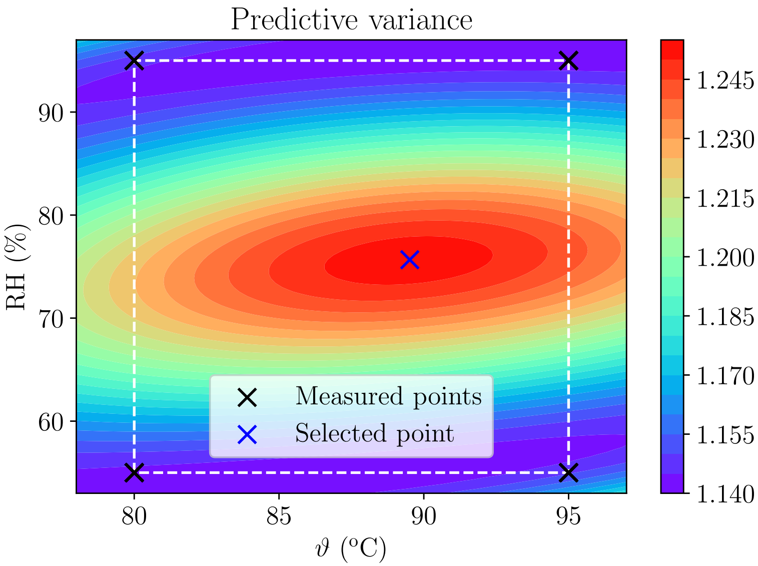
$$\mathcal{M}_1 : \boldsymbol{\phi}_1(\boldsymbol{x}) = [1,\ \vartheta,\ \mathrm{RH}]$$
$$\mathcal{M}_2 : \boldsymbol{\phi}_2(\boldsymbol{x}) = [1,\ \vartheta,\ \mathrm{RH},\ \vartheta\,\mathrm{RH}]$$
$$\mathcal{M}_3 : \boldsymbol{\phi}_3(\boldsymbol{x}) = \left[1,\ \vartheta,\ \mathrm{RH},\ \vartheta^2,\ \mathrm{RH}^2\right]$$
$$\mathcal{M}_4 : \boldsymbol{\phi}_4(\boldsymbol{x}) = \left[1,\ \vartheta,\ \mathrm{RH},\ \vartheta^2,\ \mathrm{RH}^2,\ \vartheta^2\,\mathrm{RH}^2\right]$$

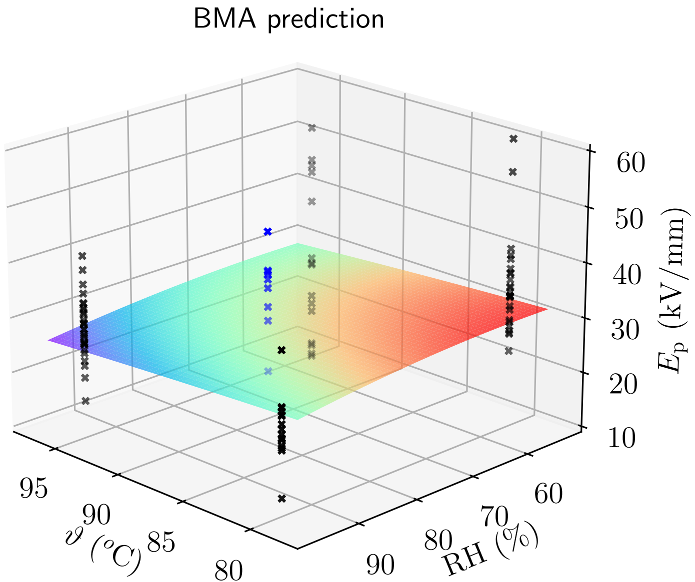
Model selection



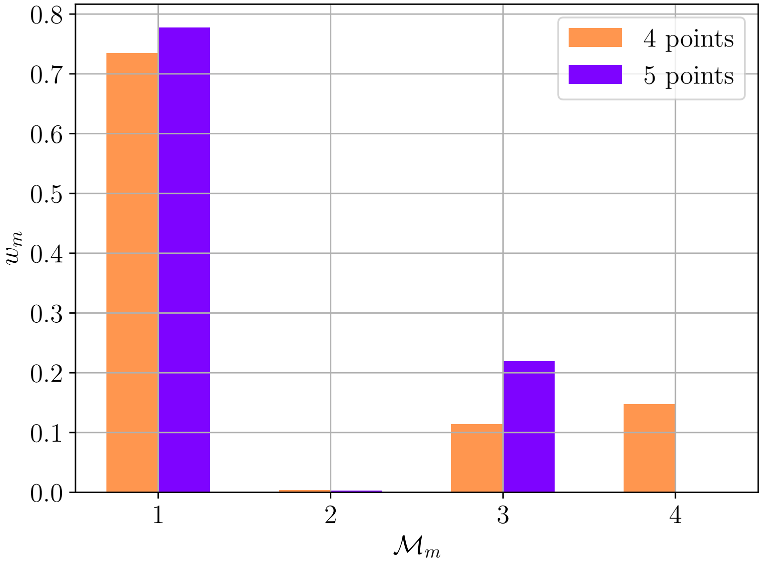

BMA prediction
60
50
40
30
20
10
E_{p} (kV/mm)
95
90
85
80
ϑ ($^o C$)
60
70
80
90
RH (%)

Model selection




BMA prediction
60
50
40
30
20
10
E_p (kV/mm)
95
90
85
80
ϑ (^oC)
60
70
80
90
RH (%)

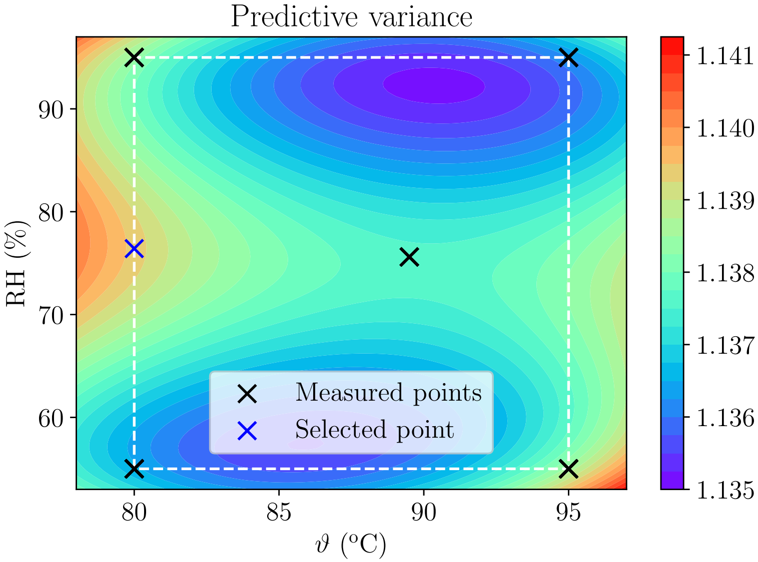
Model selection



Model selection

Predicted E_p from initial experiment					Validation
BMA	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	measurement
32.52	28.16	29.03	32.16	67.55	**33.31**

Model selection



QA