

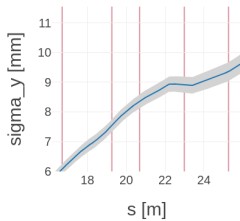
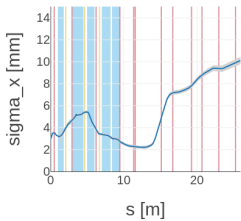
Surrogate model

Renato Bellotti

March 5, 2020

Dashboard: Uncertainties

Beam sizes



IBF [A]	450 ▾	[450, 550]	IM [A]	150 ▾	[150, 260]
GPHASE [°]	-50 ▾	[-50, 10]	ILS1 [A]	0 ▾	[0, 250]
ILS2 [A]	0 ▾	[0, 200]	ILS3 [A]	0 ▾	[0, 200]
Bunch charge [nC]	1 ▾	[1, 5]	lambda [ps]	0.3 ▾	[0.3, 2]
Laser radius [mm]	1.5 ▾	[1.5, 12.5]			

Quantile of the residuals to use as uncertainty:

0.95

cavity

YAG

solenoid

$$\sigma_{pred} \pm r_{95\%}$$

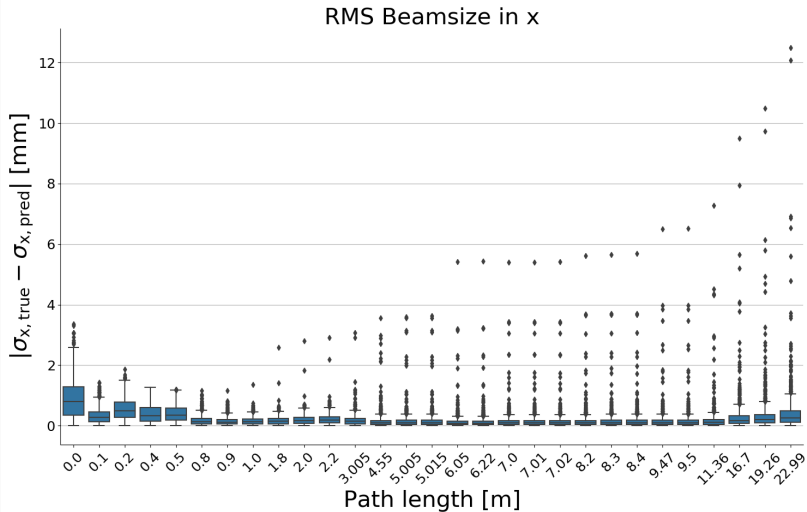
Design variables

Bound	IBF [A]	IM [A]	GPHASE [°]	ILS1 [A]	ILS2 [A]	ILS3 [A]	charge [nC]	λ [ps]	SIGXY [mm]
Lower	450	150	-50	0	0	0	1	0.3	1.5
Upper	550	260	10	250	200	200	5	2	12.5

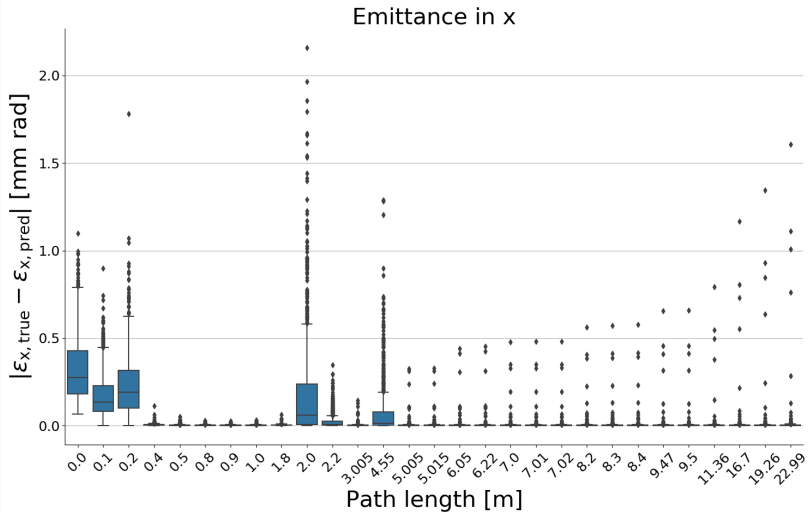
Cavities 1, 2, 3, 5: 25MV

Cavities 4 and 6: Off

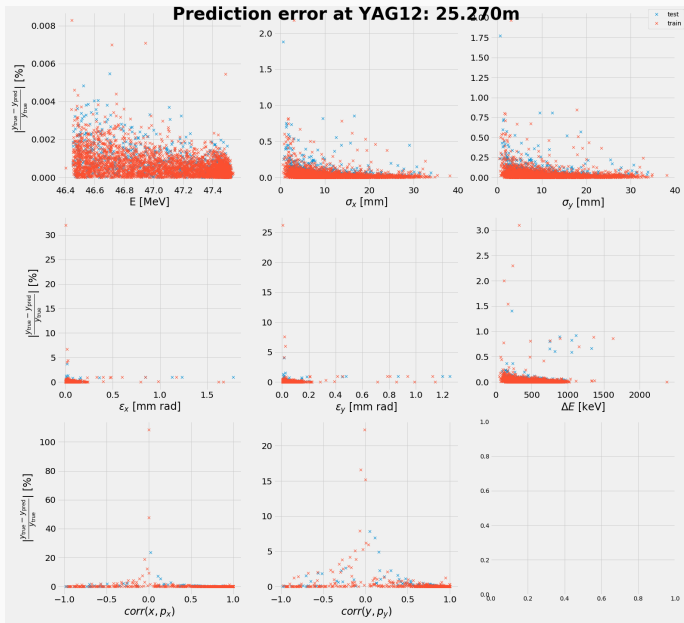
Model performance: σ_x



Model performance: ϵ_x



Model performance: Error at the last YAG screen



Backup slides

No analytical formula for the surrogate \rightarrow No CI

Residuals are not Gaussian distributed \rightarrow No $\mu \pm \sigma$ error bars

Use descriptive statistics!

Uncertainties

- $\vec{x}_i = \begin{pmatrix} \text{IBF}_i \\ \text{IM}_i \\ \vdots \end{pmatrix}$ i-th DVAR config in the test set
- $\sigma(\vec{x}, s)$: Beamsize at position s for DVAR configuration \vec{x}
 $\sigma_i(s)$: Beamsize at position s for DVAR configuration \vec{x}_i .
 $\vec{\sigma}(s) := (\sigma_i(s))$

- $\vec{r}_{\sigma,s}$: **Residuals** of $\vec{\sigma}$ at position s

$$r_{\sigma,s,i} := \sigma_{\text{OPAL},i}(s) - \sigma_{\text{pred},i}(s)$$

Each row of $\vec{r}_{\sigma,s}$ is the deviation of the prediction from the simulation for a specific DVAR configuration.

- **p%-Quantile** for the abs. residuals at position s :

$$q_p(s) := |\vec{r}_{\sigma,s}|_{p\%}$$

" p percent of the residuals at position s have an absolute value $\leq q_p(s)$."

- **Empirical estimate of the model precision:**

$$\sigma_{\text{upper,lower bound}}(\vec{x}, s) = \sigma_{\text{pred}}(\vec{x}, s) \pm q_p(s)$$

- **Note:**

$\vec{r}_{\sigma,s}$ is specific to the samples in the test set

$q_p(s)$ is a robust estimate for the error on the underlying data distribution!

(assuming that the test set is representative)