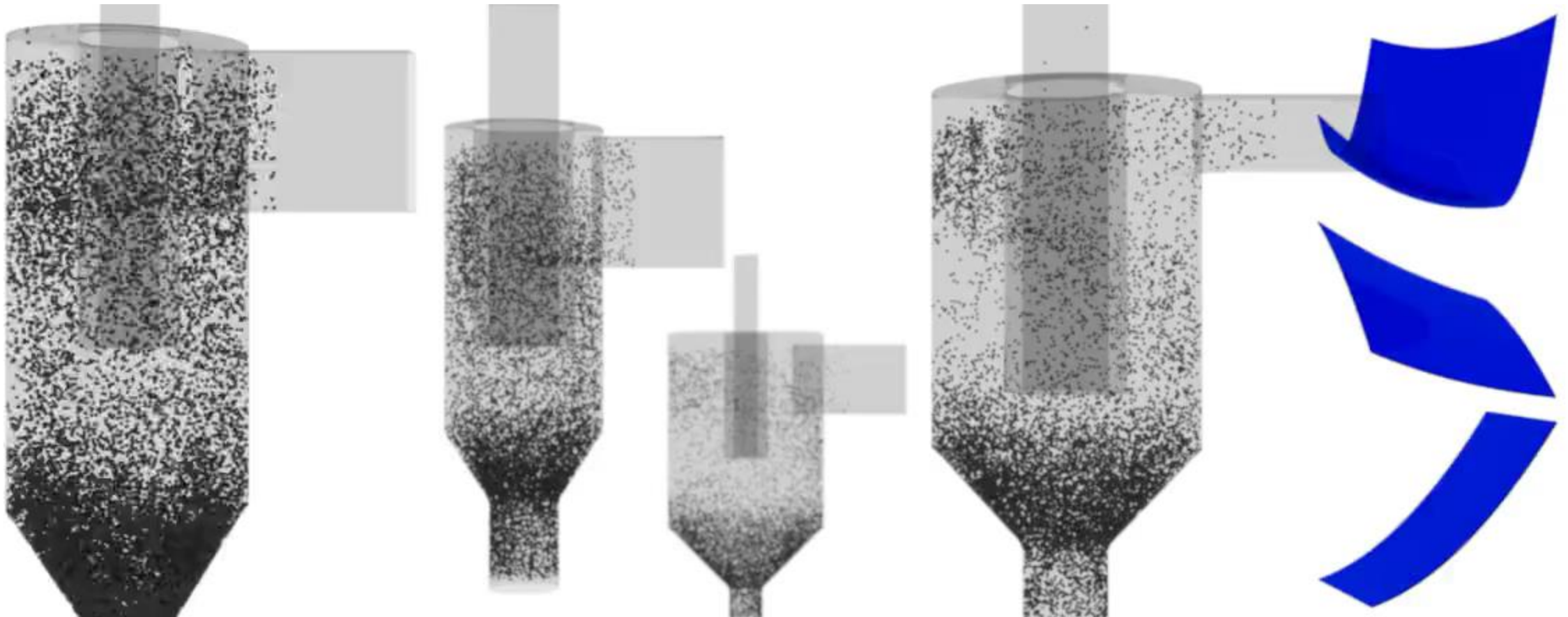




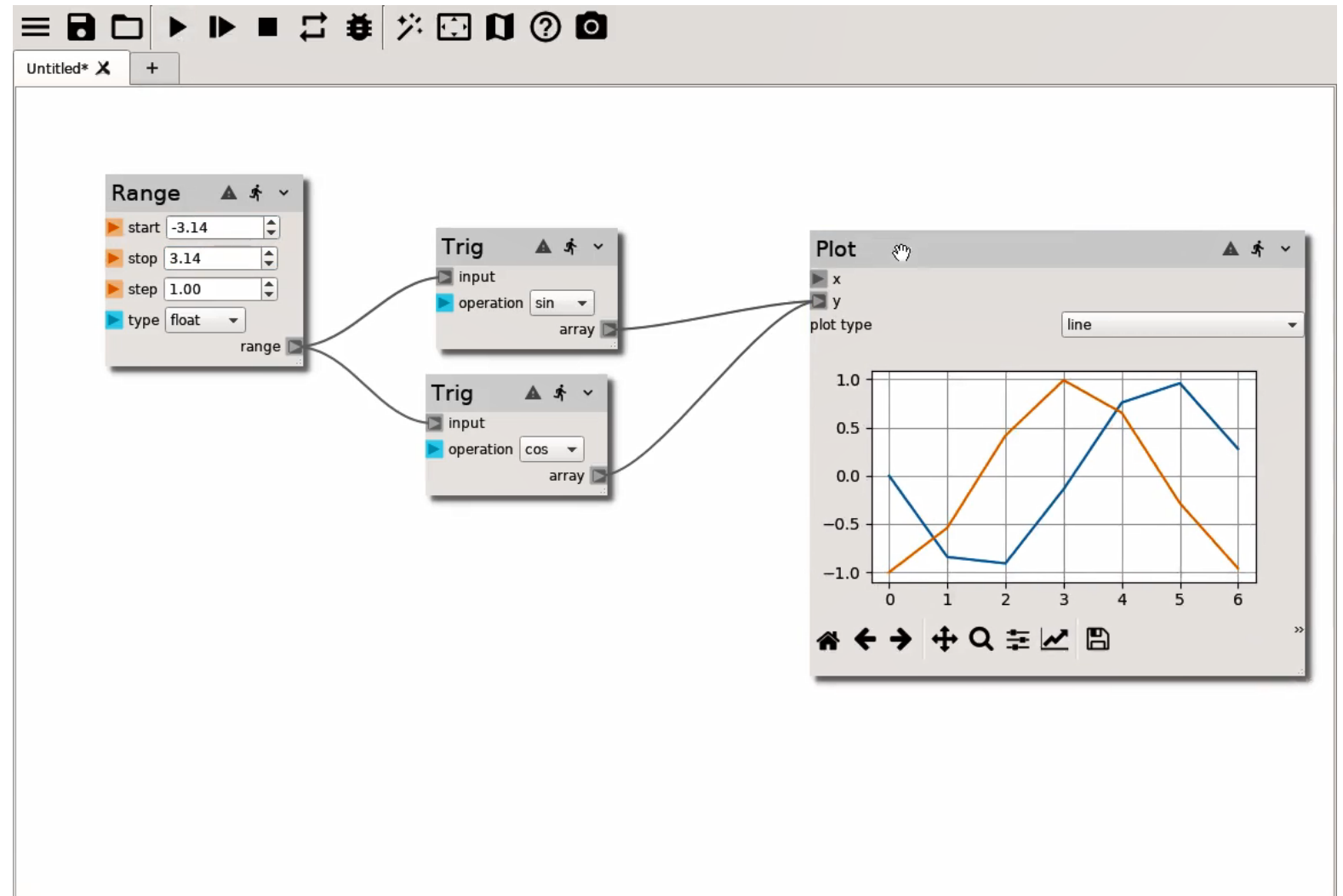
Surrogate modeling and analysis toolset

Justin Weber, William Fullmer, Aytakin Gel

Research and Innovation Center (RIC)



- Application and framework for graphical programming through the use of nodes and connections
- Underlying library for the optimization/UQ work.
- Integrates with the MFiX GUI



Surrogate modeling and analysis toolset

Design of
Experiments



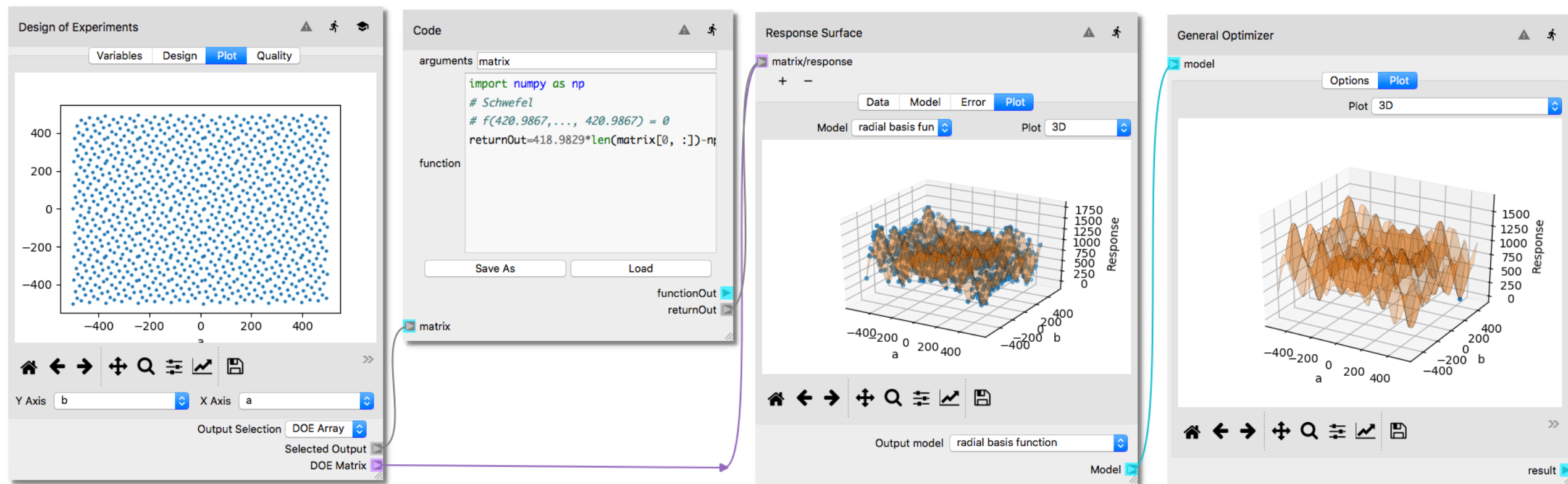
Model evaluation



Response Surface
Construction



Optimization



Design of Experiments | Variables

Inside MFiX = MFiX Aware

Fuzzy search of
Parameters and
MFiX Keywords

Automatic
population
of
Categorical
Variables

Add
variables

Select
variable
parameters

Design of Experiments

Variables Design Plot Quality

	variable	type	args	count
0	a	Double Preci...	[]	10
1	b	Double Preci...	[]	10
2	c	Double Preci...	[]	10

variable c

type Double Precision

link None

from -3.141592653589793 to 3.141592653589793

levels 10

Output Selection DOE Array

Selected Output DOE Matrix

Design of Experiments

Variables Design Plot Quality Run

	variable	type	args	count
1	con	Double Preci...	[]	1

variable con

type cone_height

link des_conv_corr

from des_min_cond_dist

levels hamaker_constant

gener_part_config

wall_hamaker_constant

set_corner_cells

cn_on

jackson

fric_non_sing_fac

friction_model

Design of Experiments

Variables Design Plot Quality Run

	variable	type	args	count
1	cone_height	Double Preci...	[]	2
2	drag_type	String	[]	3

variable drag_type

type String

value(s)

☒ SYAM_OBRIEN

☒ GIDASPOW

☐ GIDASPOW_BLEND

☒ WEN_YU

Output Selection Completed

Selected Output DOE Matrix

Design of Experiments | Methods

Method

Factorial
Covary
Montecarlo
Latin hypercube
Central
composite
Sobol
Hammersly
Halton

The screenshot shows the 'Design of Experiments' software interface with the 'Design' tab selected. The 'Method' dropdown is set to 'hammersly'. The 'Samples' field is set to 500. The 'Randomize' checkbox is checked, and the 'seed' is 8049027. The 'Repeat' checkbox is unchecked, and the 'random samples' field is set to 1. The 'Randomize sample order' checkbox is unchecked. Below these options are 'Import', 'Build', and 'Export' buttons. A table displays the generated samples:

	a	b	c
1	-3.14	-3.14	-3.14
2	-3.13	0	-1.05
3	-3.12	-1.57	1.05
4	-3.1	1.57	-2.44
5	-3.09	-2.36	-0.349

At the bottom, the 'Output Selection' dropdown is set to 'DOE Array', and the 'Selected Output' dropdown is set to 'DOE Matrix'.

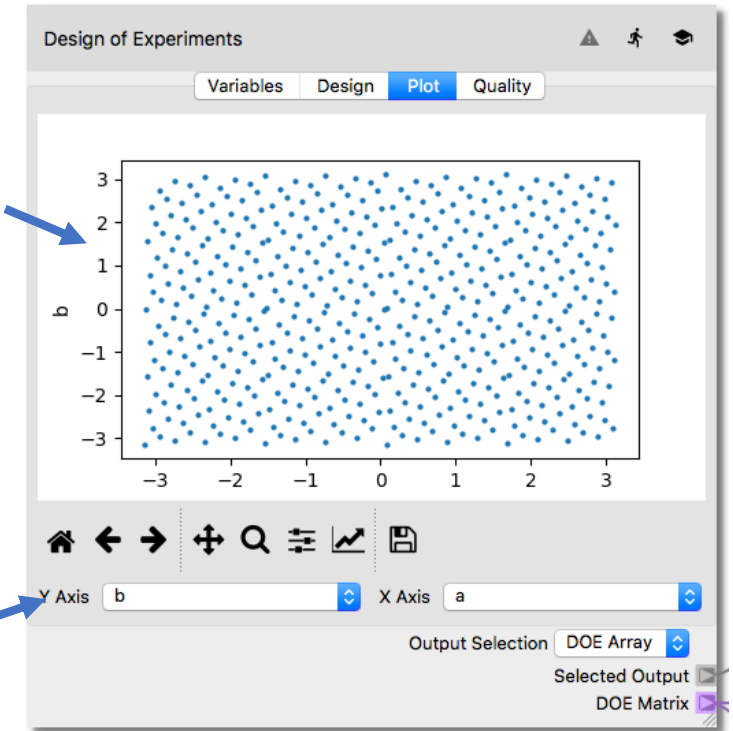
Method
Options

Import, Build,
Export

Samples

2D plot
of
samples

Change
variables



Response Surface | Samples

Samples +
Response

-or-

Read CSV

Response Surface

matrix/response

+ -

Data Model Error Plot

Import

	a	b	Response
0	-5	-5	12.6
1	-4.99	0	10.1
2	-4.98	-2.5	12.6
3	-4.97	2.5	12.6
4	-4.96	-3.75	12.8
5	-4.95	1.25	11.4
6	-4.94	-1.25	11.4
7	-4.93	3.75	12.8
8	-4.92	-4.38	13.7
9	-4.91	0.625	11.7
10	-4.9	-1.88	11.1
11	-4.89	3.12	11.8
12	-4.88	-3.12	11.9
13	-4.87	1.88	11.2
14	-4.86	-0.625	11.8
15	-4.85	4.38	13.8
16	-4.84	-4.69	13.9
17	-4.83	0.312	11.6
18	-4.82	-2.19	11.8
19	-4.81	2.81	12.2
20	-4.8	-3.44	13.3
21	-4.79	1.56	12.2

Output model radial basis function

Model

Read CSV File

File rs/jweber/Downloads/40hopper_case.csv Browse

Delimiter space(s)

Header row 0

Comment Character #

Reader pandas

Sample Column(s)

- ☒ PP_FC
- ☒ PW_FC
- ☒ PP_RC
- ☒ PW_RC
- ☐ Discharge

Response Column Discharge

	PP_FC	PW_FC	PP_RC	PW_RC	Dis
2	0.27	0.814	0.475	0.958	1.95
3	0.664	0.316	0.62	0.681	1.9
4	0.986	0.605	0.538	0.311	1.85
5	0.579	0.781	0.921	0.276	1.99
6	0.242	0.638	0.683	0.613	1.98
7	0.857	0.093	0.671	0.95	2.34

Import Cancel

Response Surface | Models

Points to
remove for
cross validation

Select
models to fit

Model
parameters

Fit model

Response Surface

matrix/response

+ -

Data Model Error Plot

Cross validation points 10 %

fit	Model	MSE	R ²	L _{inf}	L ₁	L ₂
✓	radial basis function	0.113	0.982	0.123	0.0231	0.0332
✓	cubic	0.14	0.978	0.099	0.0265	0.0368
✓	linear	0.236	0.963	0.121	0.0384	0.0477
✓	random forest	0.243	0.961	0.103	0.0383	0.0486
✓	support vector machine	0.372	0.941	0.135	0.0482	0.0603
✓	decision tree	0.433	0.931	0.13	0.052	0.065
✓	nearest	0.468	0.926	0.131	0.0538	0.0676
✓	multilayer perceptron	0.505	0.92	0.171	0.0594	0.0701
✓	MARS	0.678	0.892	0.161	0.065	0.0813
✓	polynomial	0.956	0.848	0.209	0.0807	0.0966
✓	gaussian process	1.07	0.83	0.643	0.0405	0.102

Max terms 100

Max degree 1

Penalty 3.0

Refit Model(s)

Output model radial basis function

Model

Error
metrics



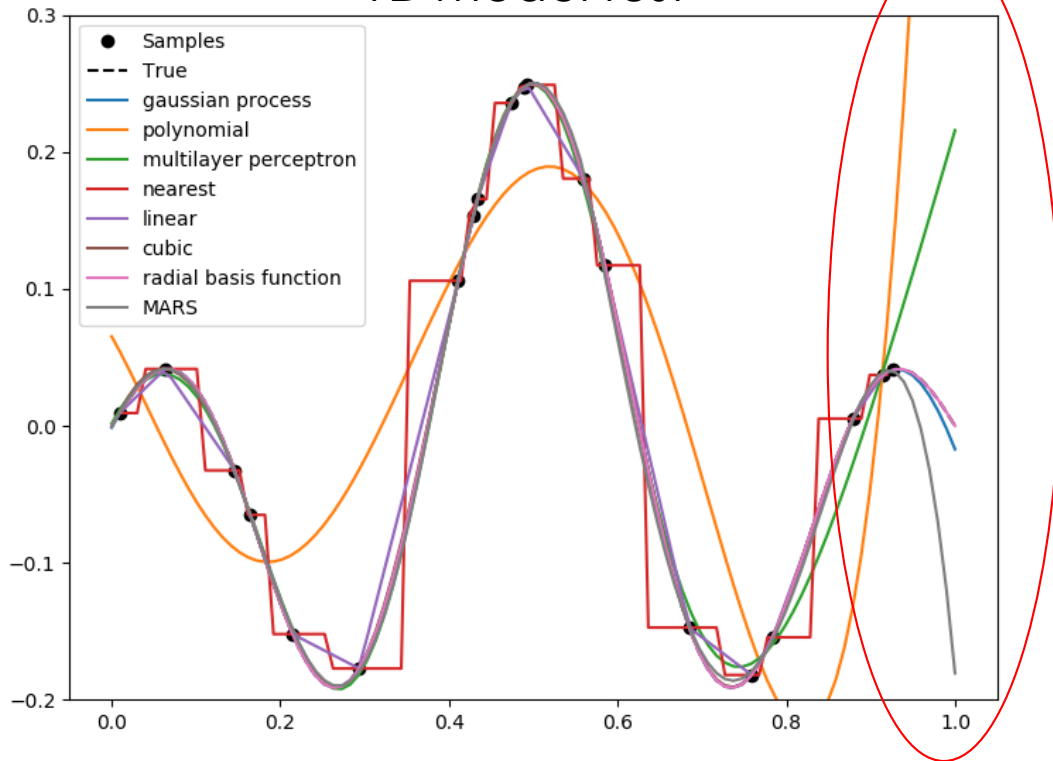
gaussian process
polynomial
multilayer perceptron
Support vector machine
Decision tree
Random forest

nearest
linear
cubic ($d \leq 2$)
radial basis function

py-earth MARS

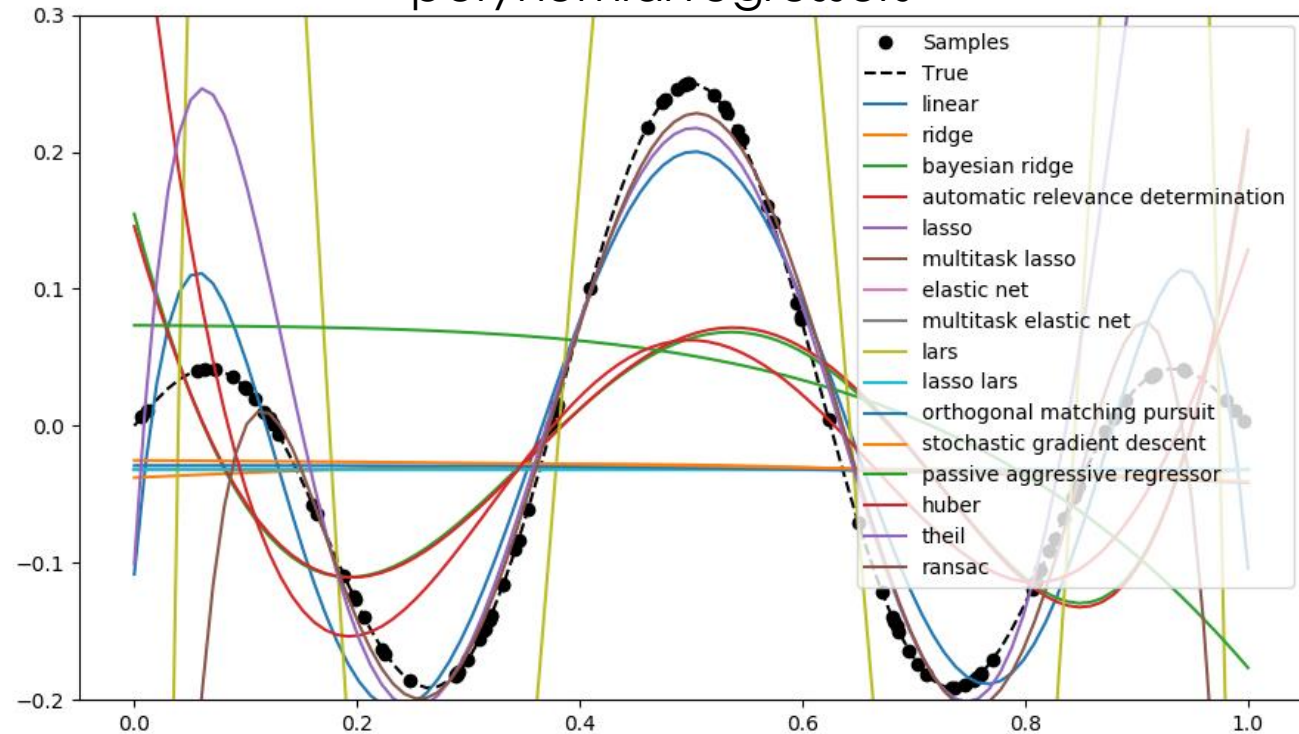
Response Surface | Models

1D model test



watch the edge!

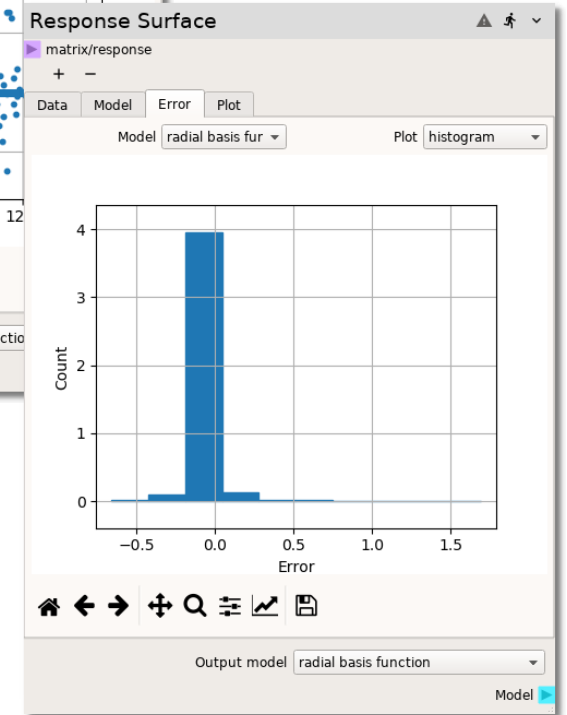
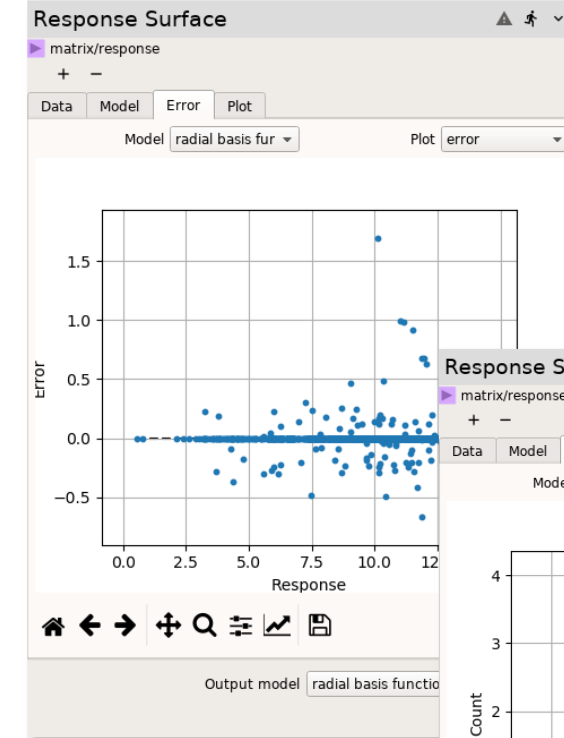
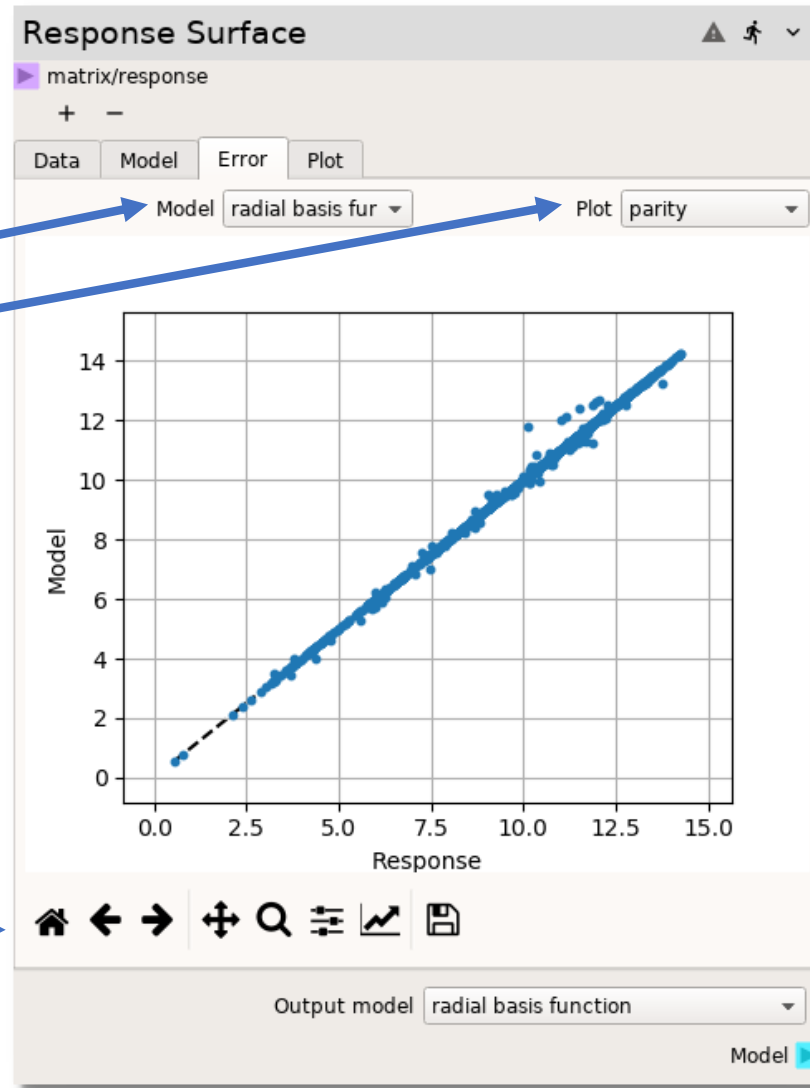
polynomial regressors



Response Surface | Error Plots

Select model

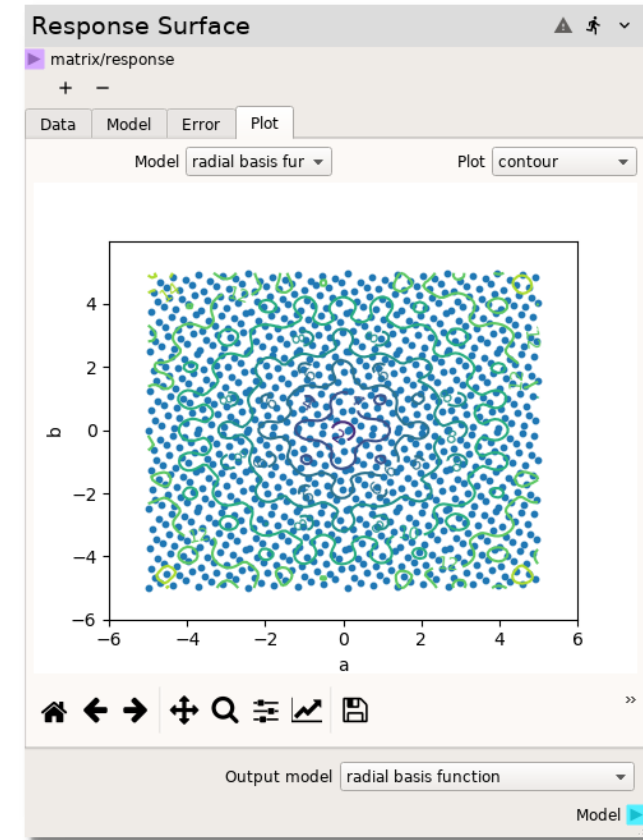
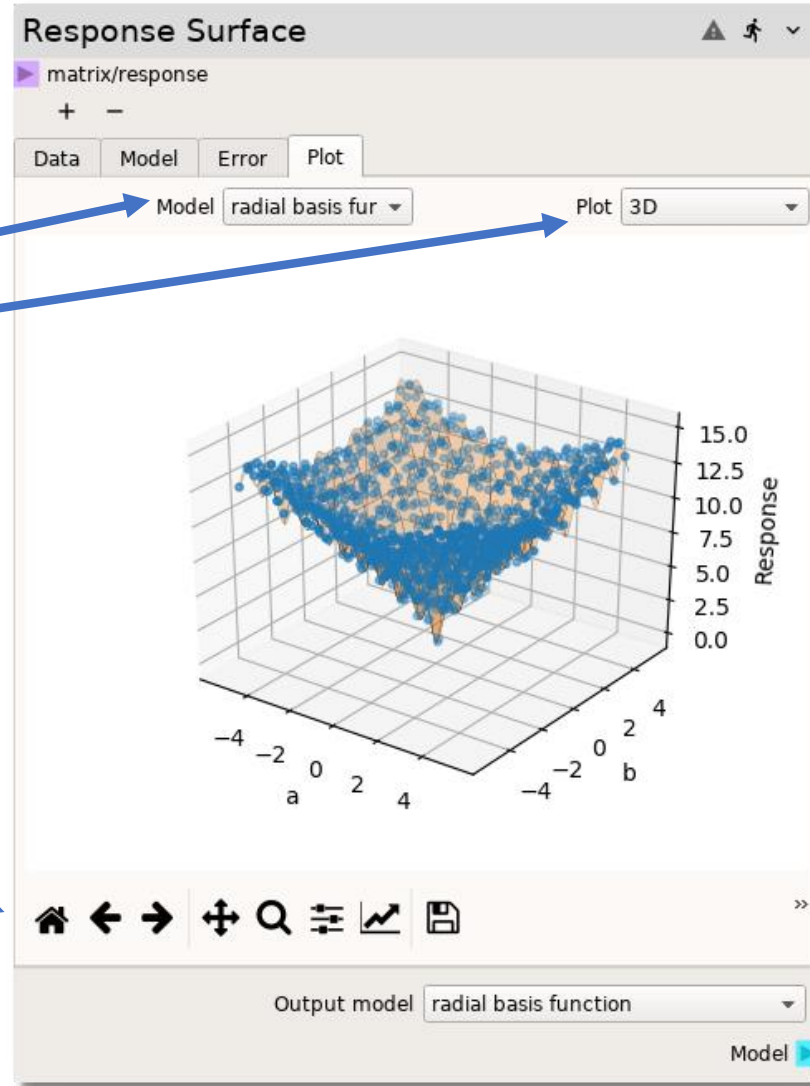
Select plot



Response Surface | Plots

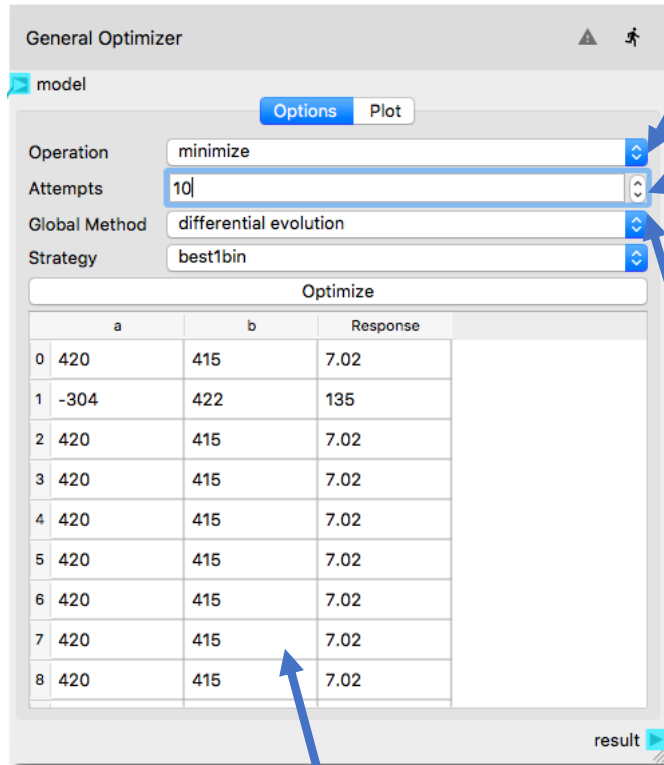
Select model

Select plot



Save/manipulate
plot

Optimization



minimize
maximize
find v value (root)

number of attempts

differential evolution

basin hopping

Nelder-Mead

Powell

CG

BFGS

L-BFGS-B

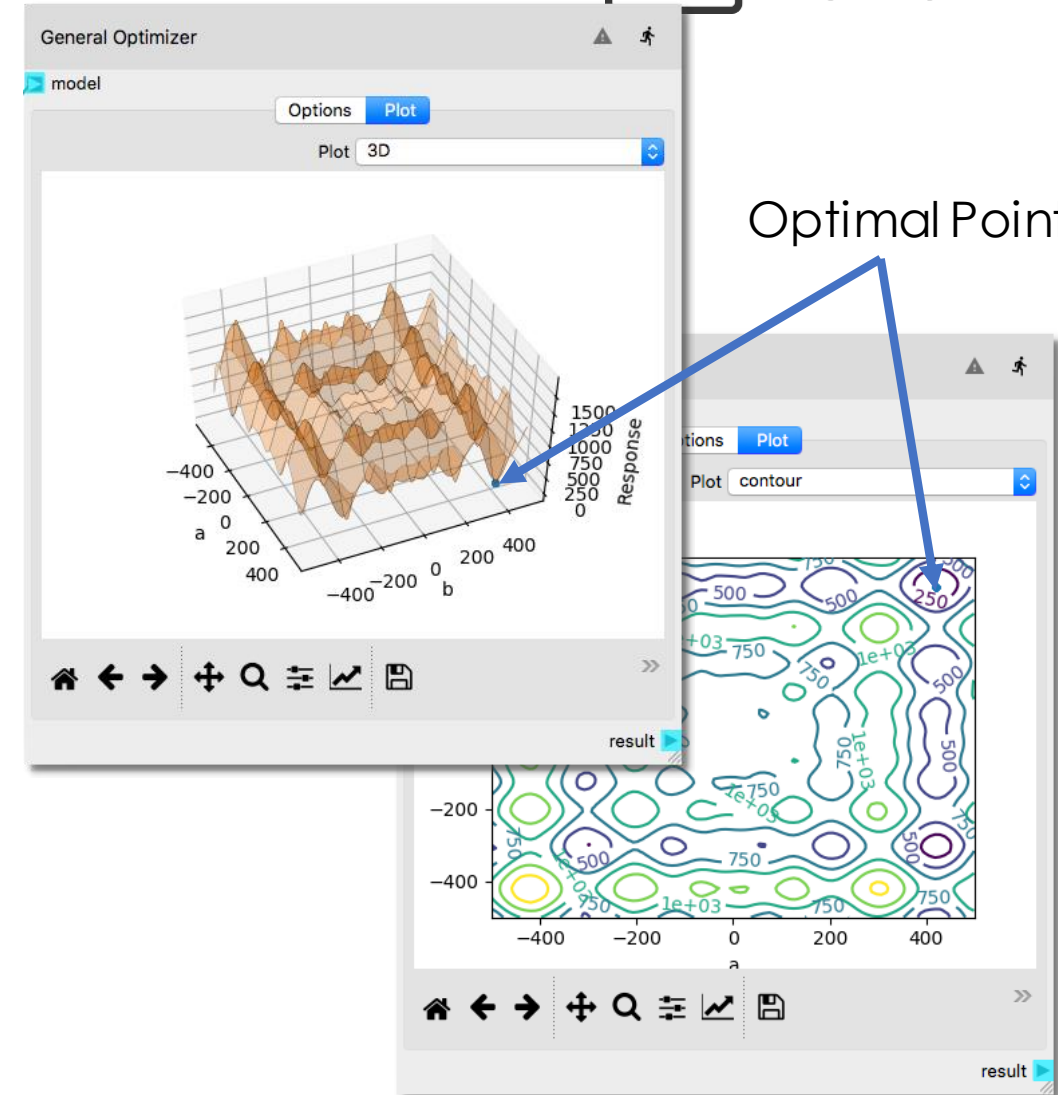
TNC

COBYLA

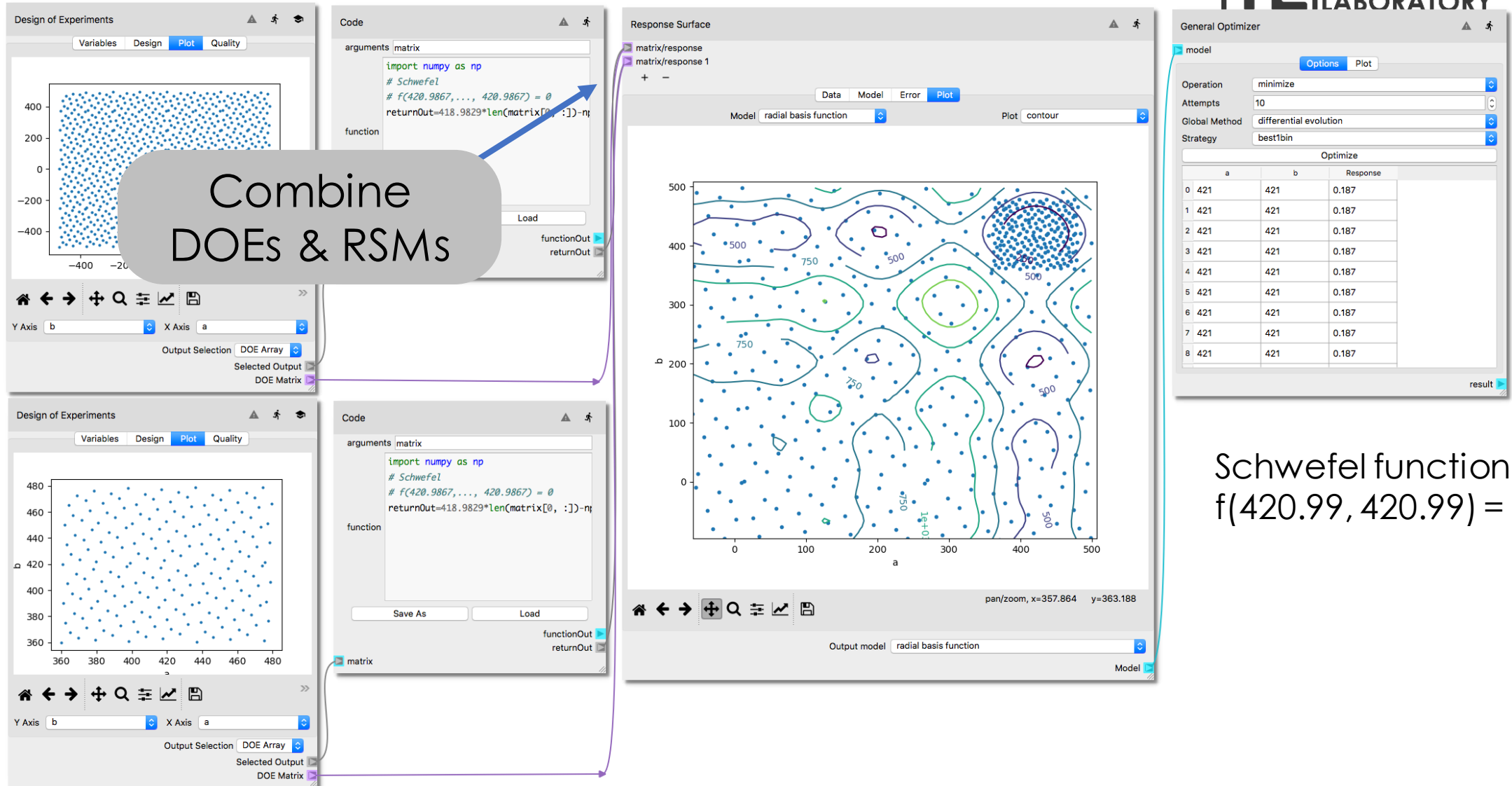
SLSQP



Results of optimization attempts



Response Surface | Refinement



Sensitivity Analysis

Sensitivity Analysis

model

Options Plot Total First Order Second Order

Method **sobol analysis**

Samples 1000

Confidence 0.95

Resamples 10

	From	To
a	-3.14	3.14
b	-3.14	3.13
c	-3.14	3.13

Calculate Sensitivities

SALib

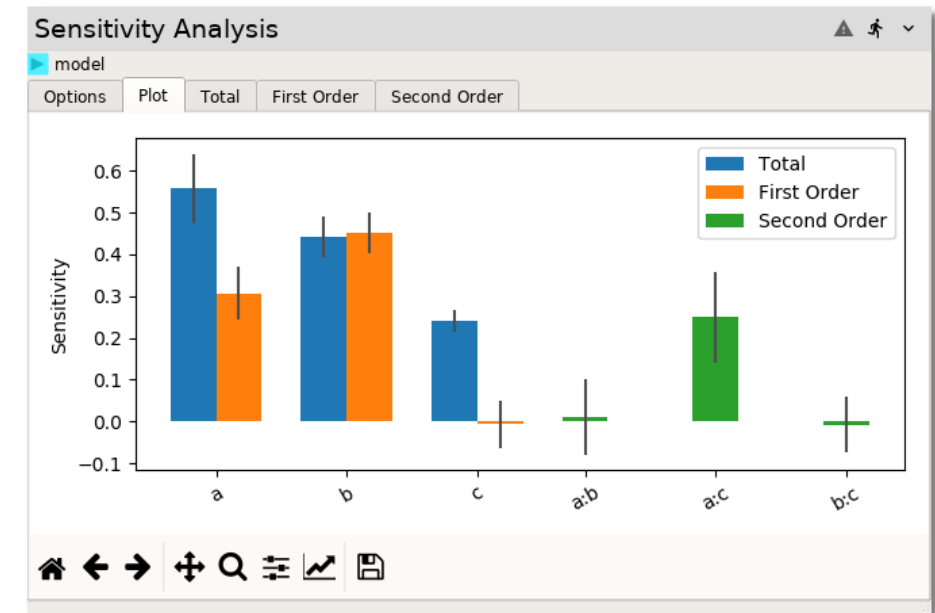
Sobol

Method of Morris

Fourier Amplitude

Delta Moment-independent

Random balance Fourier Amplitude



Forward Propagation | Variables

Forward Propagation

model

Options Bounds Probability Box

Aleatory samples 1000 Samples outside range re-draw

Epistemic samples 100 Epistemic Method latin hypercube

	type	distribution	from	to	mean	std
a	aleatory	normal	-3.14	3.14	0	0.8
b	aleatory	normal	-3.14	3.14	0	0.5
c	epistemic	unifrom	1	2	0	0.5

Variable Type aleatory

Distribution normal

Mean 0.0

Standard Deviation 0.5

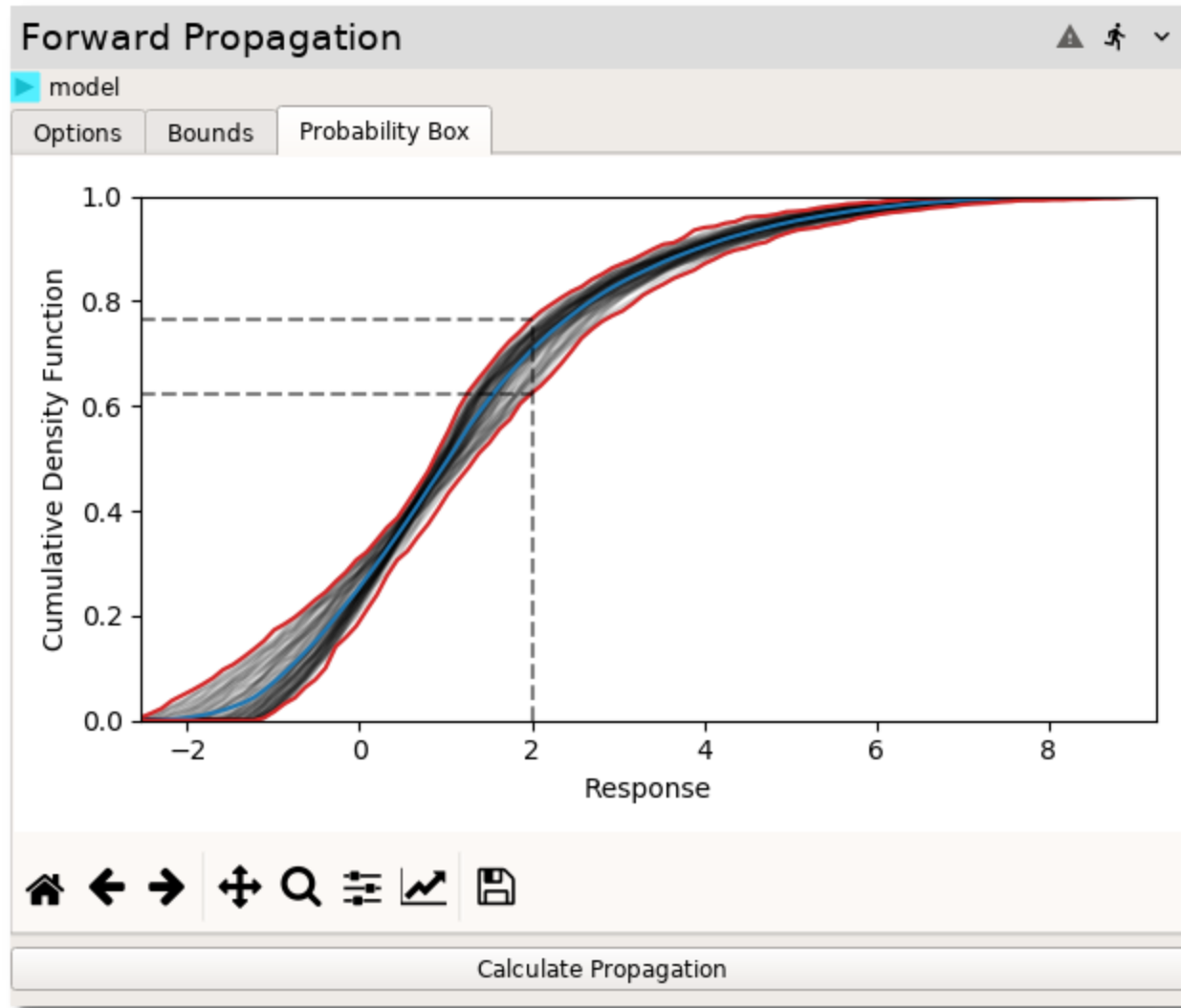
Calculate Propagation

Samples

Variables

Variable options

Forward Propagation | P-Box



Forward Propagation

model

Options Bounds Probability Box

The probability that the value will be or less
is between % and %

☒ Draw on probability box

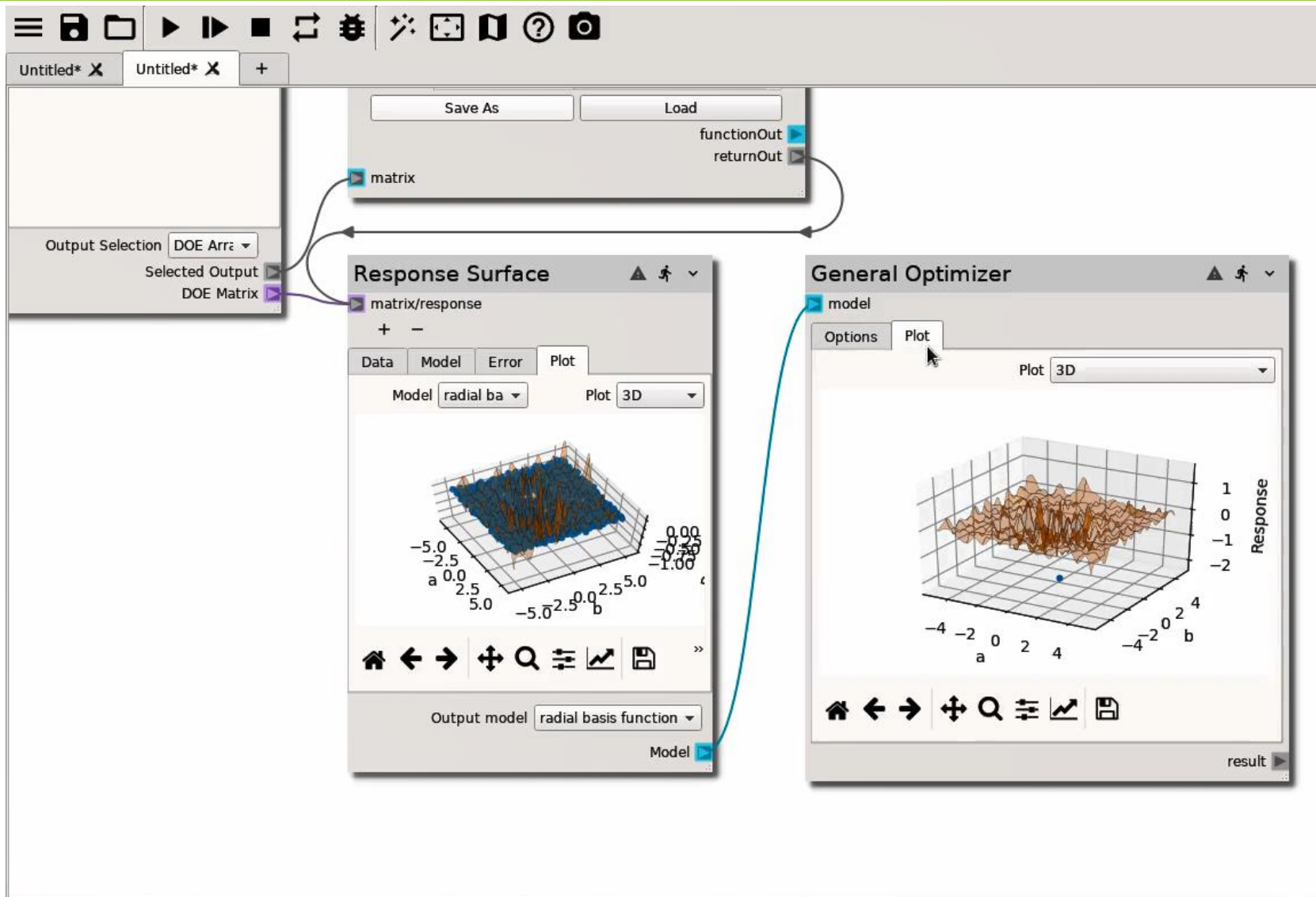
Given the prescribed input uncertainties with % probability,
the quantity of interest will be between and

☐ Draw on probability box

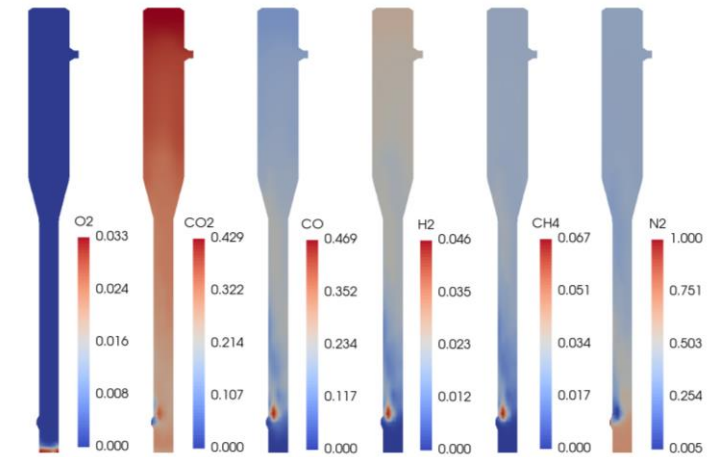
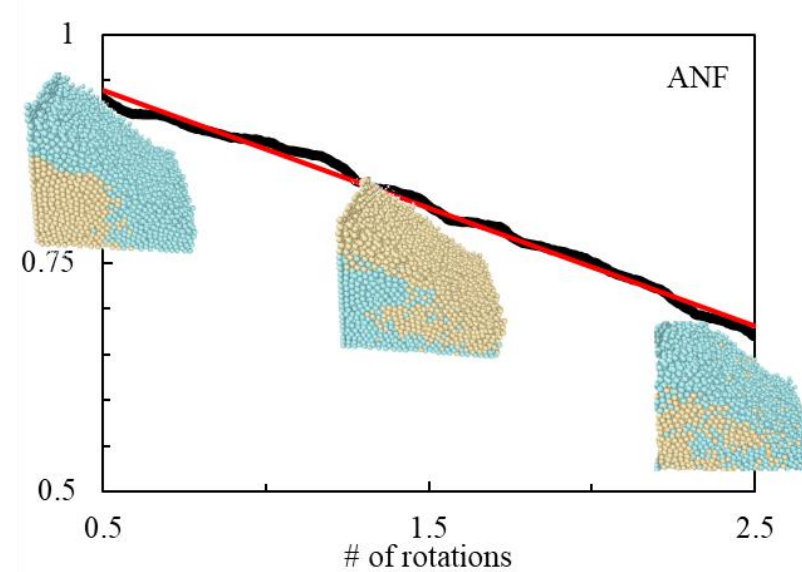
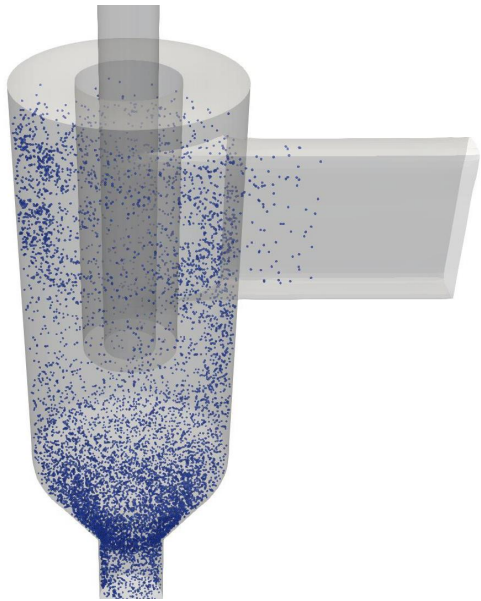
Export bounds to file

Calculate Propagation

Demo | Wizard



Examples

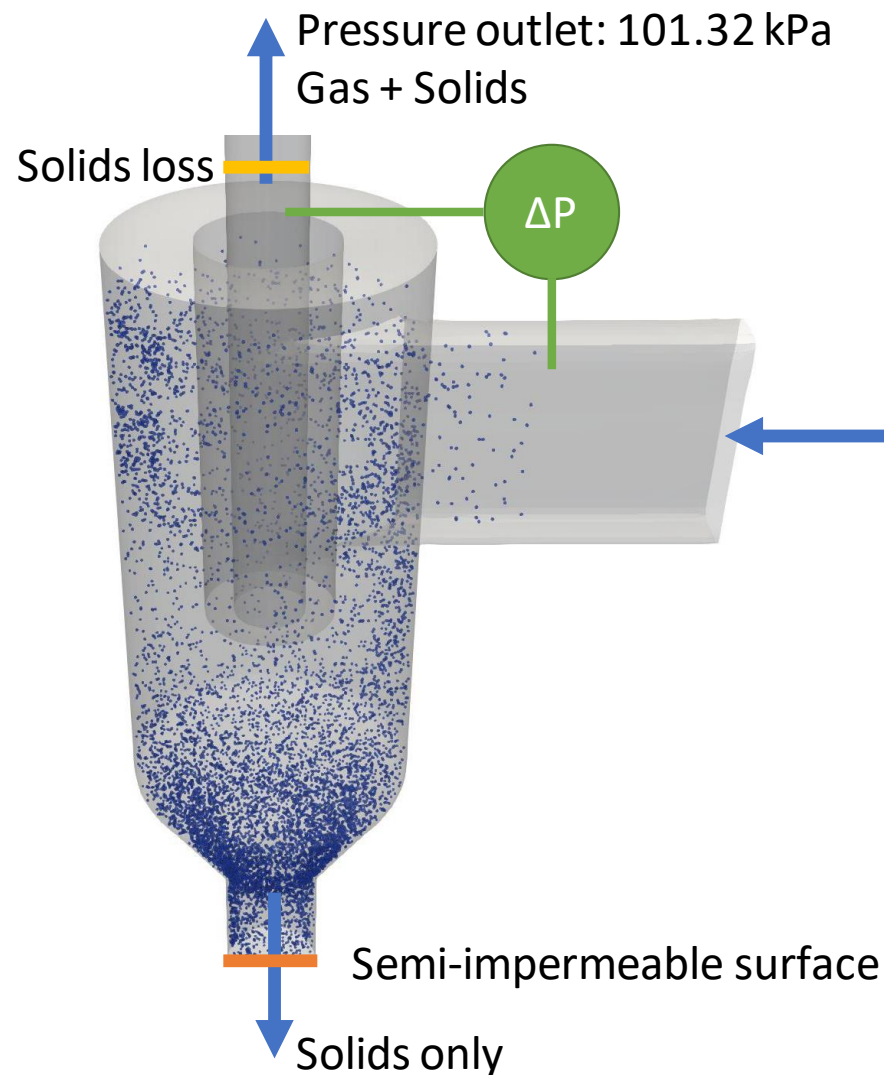
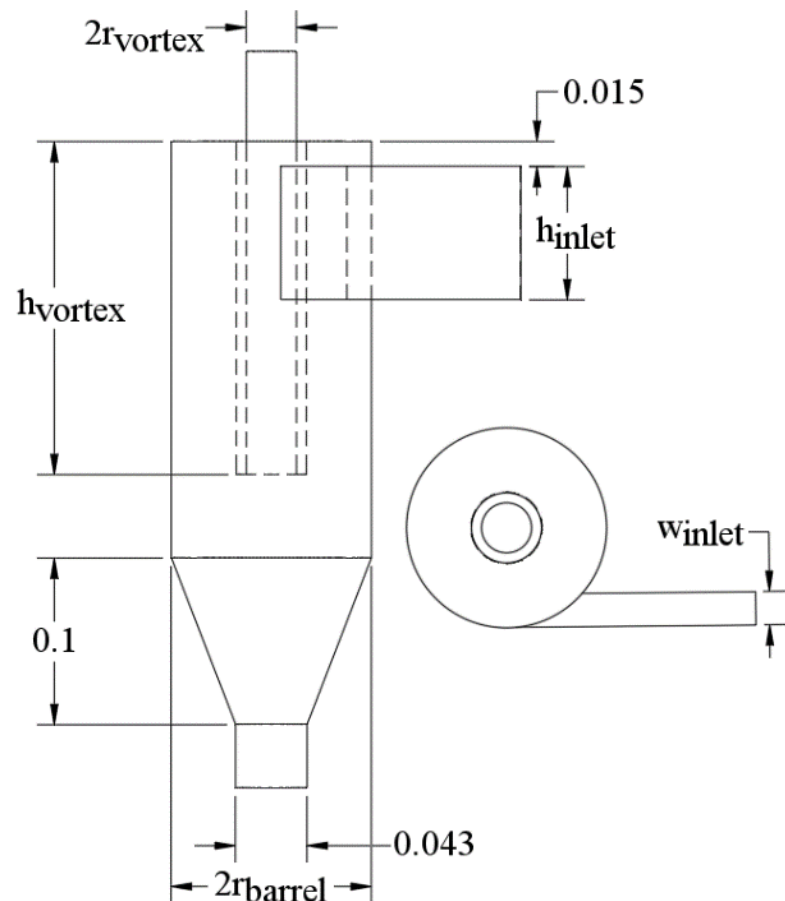




We have an underperforming cyclone on 50 kWth Chemical Looping Reactor

- Increase efficiency
- Maintain or lower pressure drop

Base cyclone



Cell size 5 x 5 x 5 mm, uniform

Gas 0.02 kg/s
Solids 0.08 kg/s

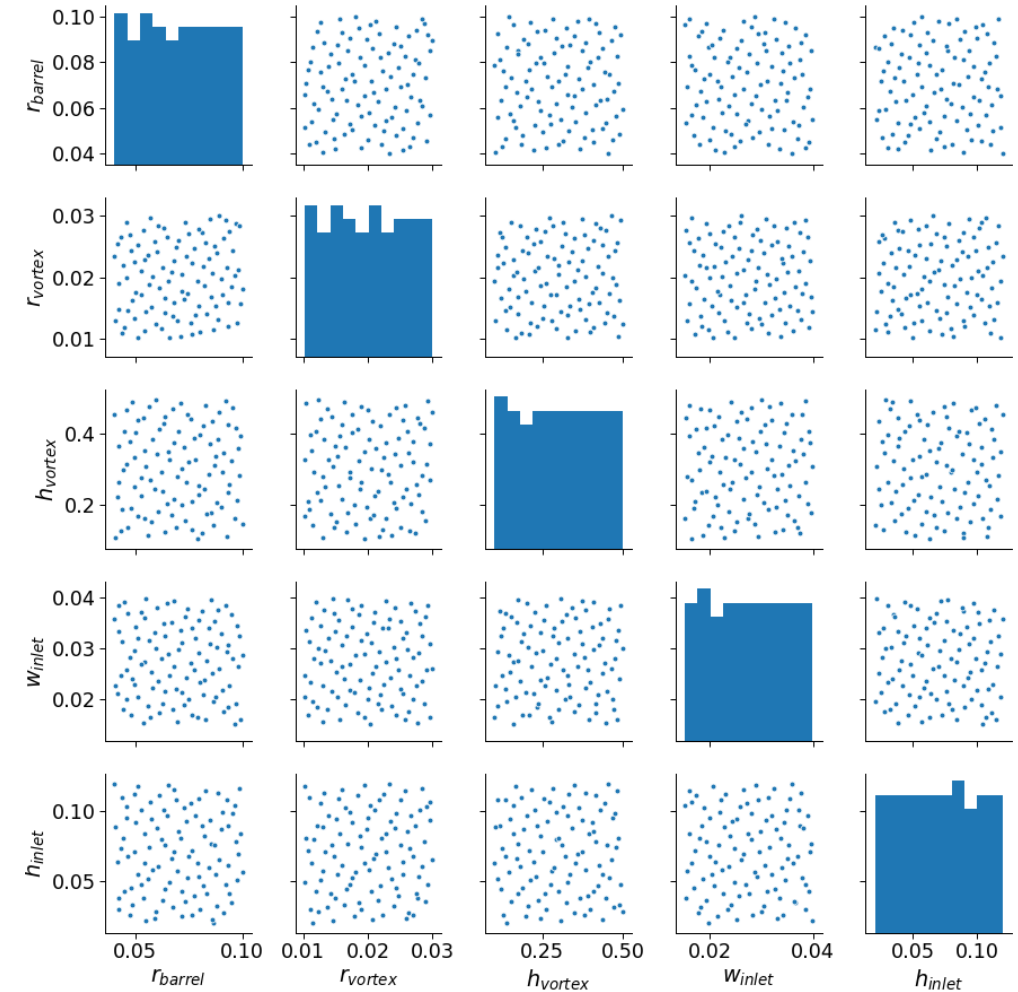
HDPE
Diameter: 871 μm
Density: 860 kg/m³

~1 particles/parcel

Design of experiments

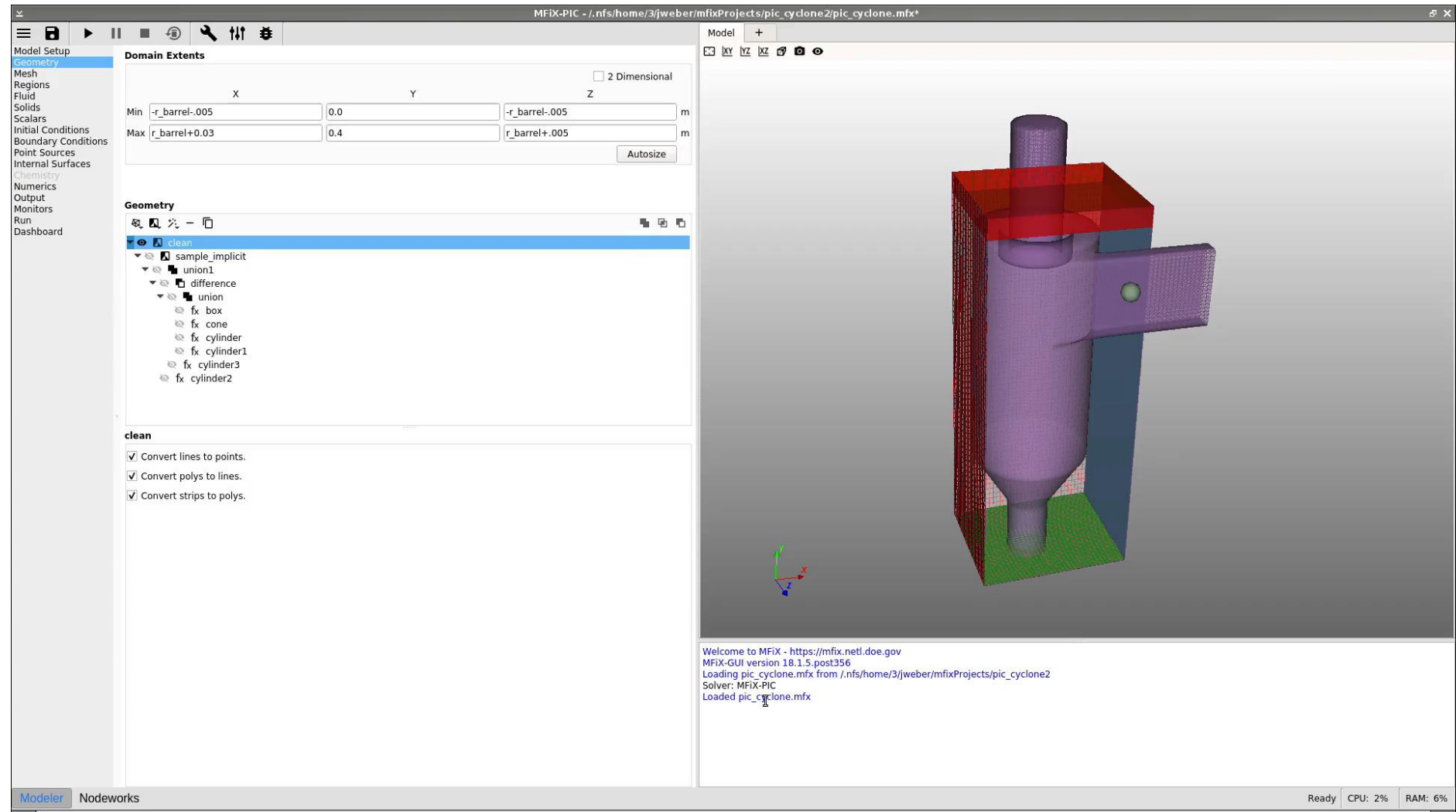
Variable	min (m)	max (m)
R_{BARREL}	0.04	0.1
R_{vortex}	0.01	0.03
H_{vortex}	0.1	0.5
H_{inlet}	0.02	0.12
W_{inlet}	0.015	0.04

- genetically optimized Latin hypercube
- 100 samples (2x recommended)
- L_2 -discrepancy measure of 0.00295



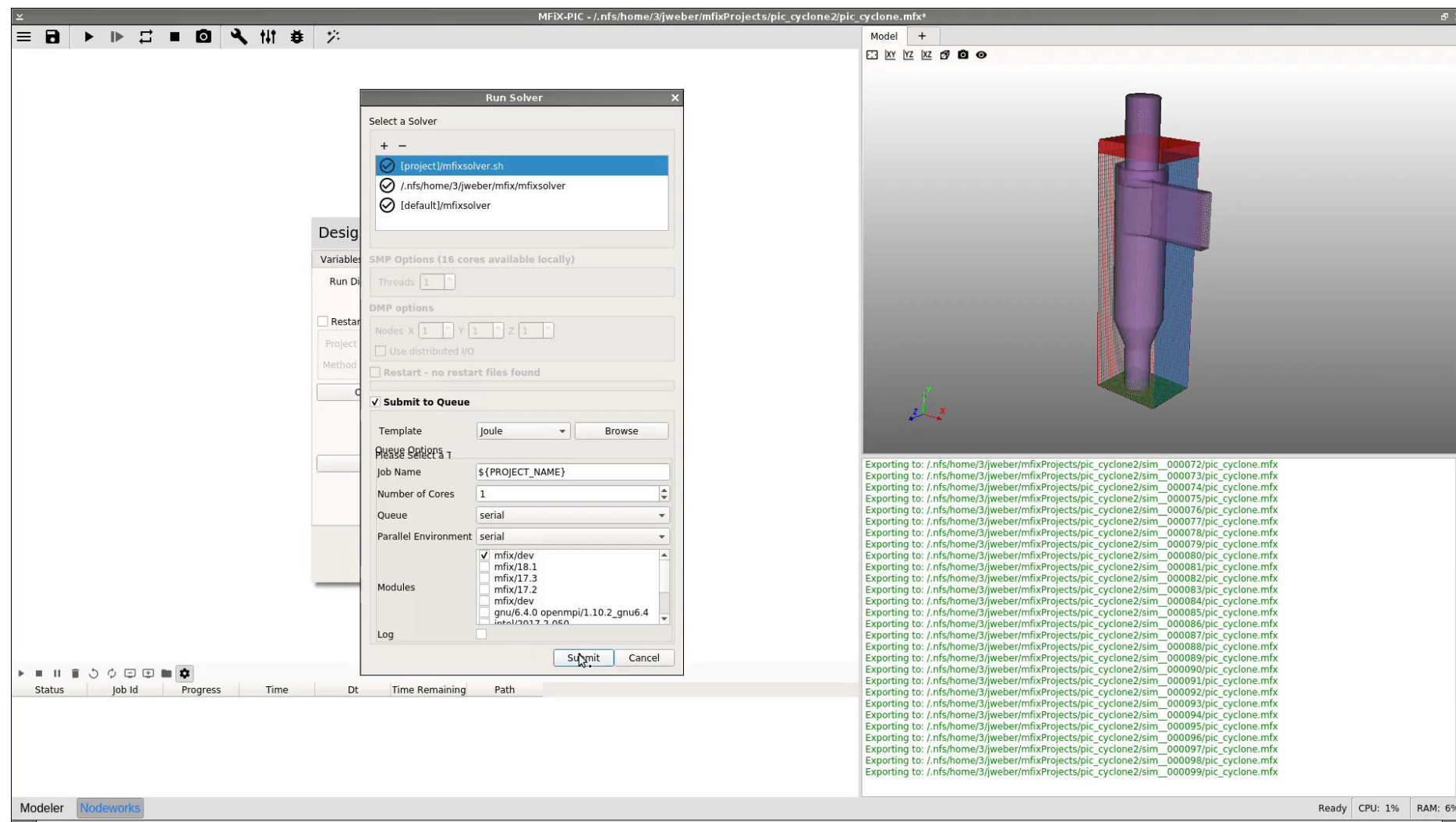
Model creation

Models created using
Nodeworks and MFiX



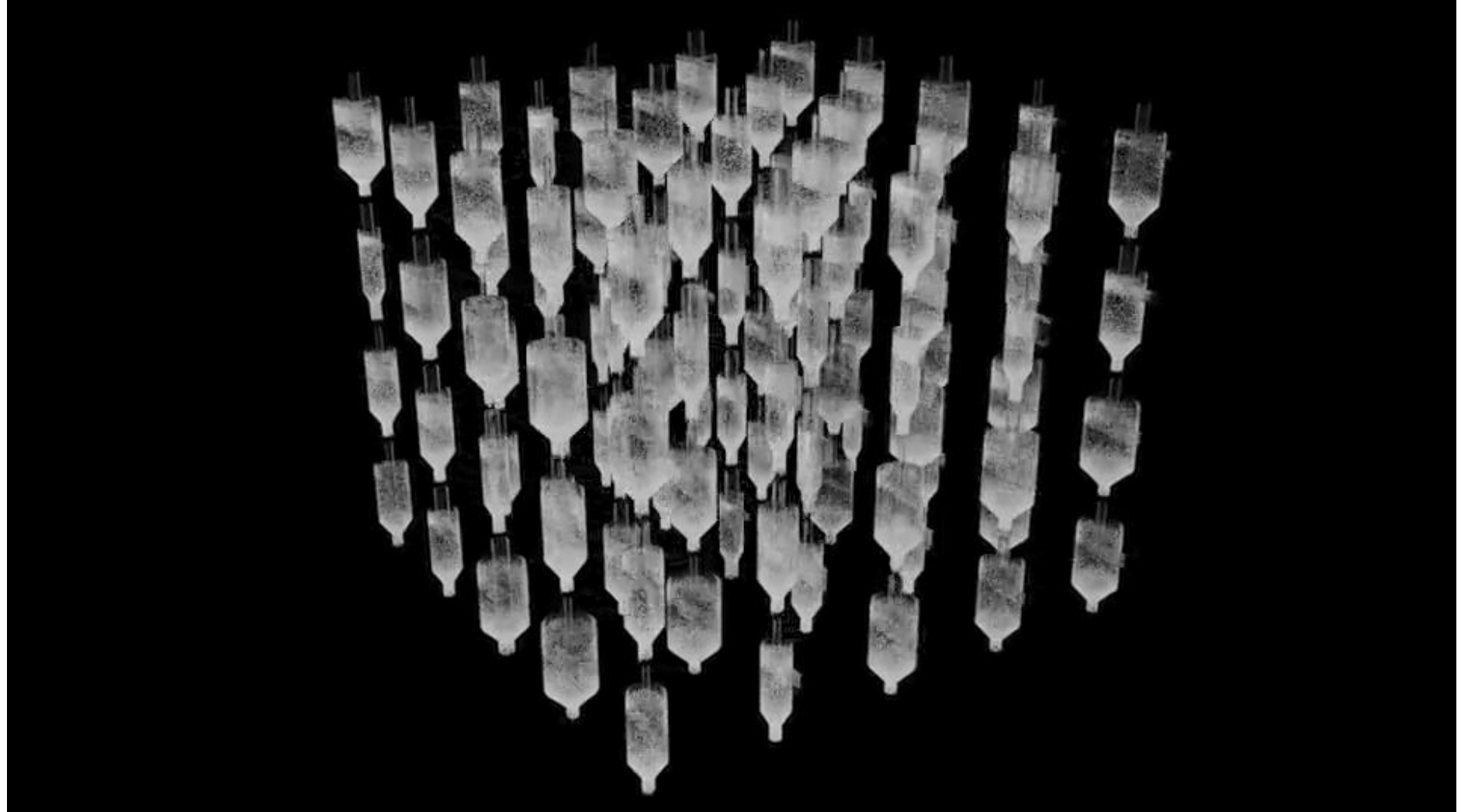
Dispatch

Using Nodeworks and
MFiX, Dispatch all models
to the queue

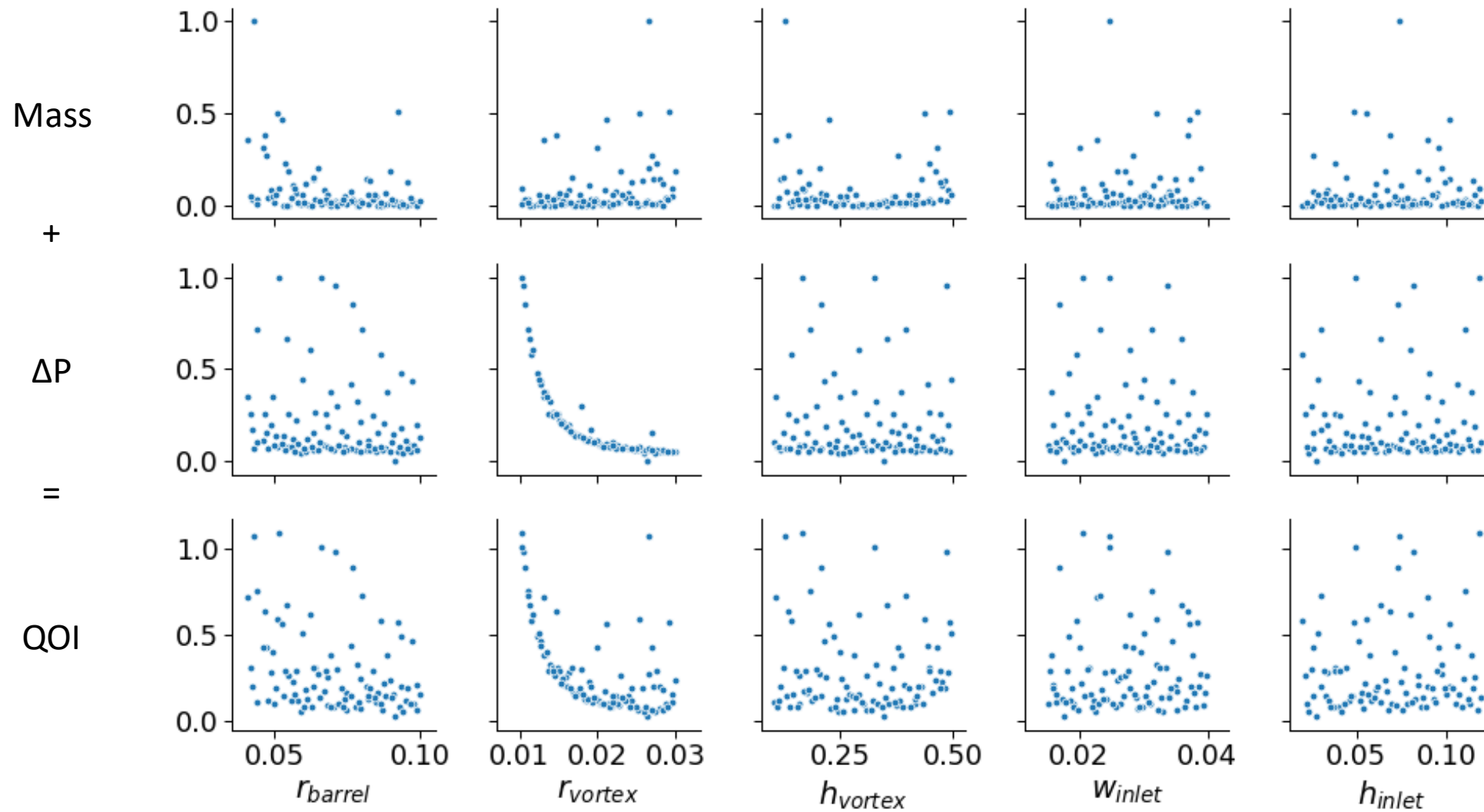


Run the models!

- All models ran simultaneously
- Took 21 minutes to 7 hours per model
- Cell count varied from 40,320 to 169,764
- Three models failed (6%), due to bad mesh



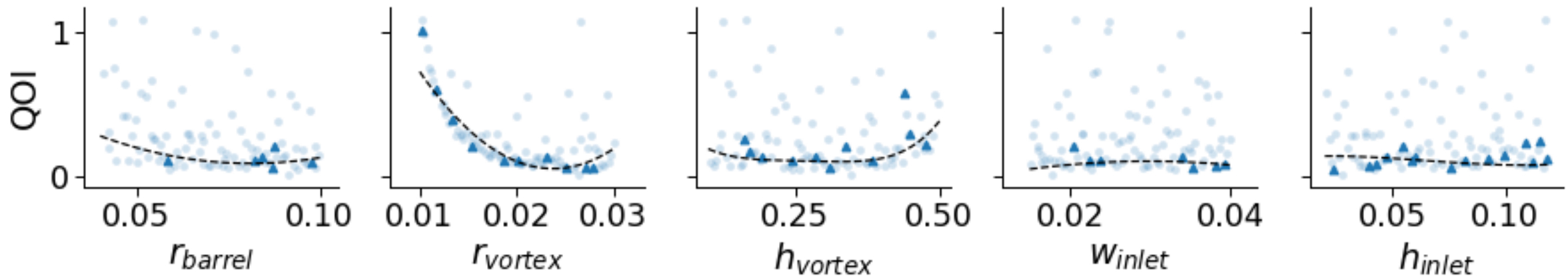
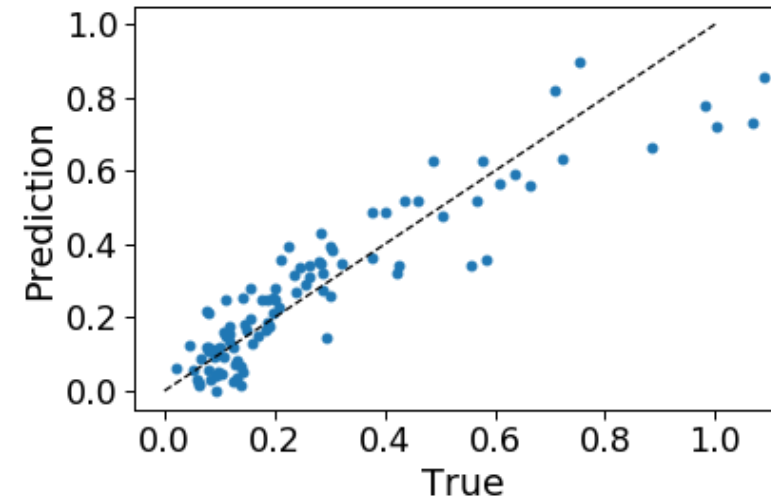
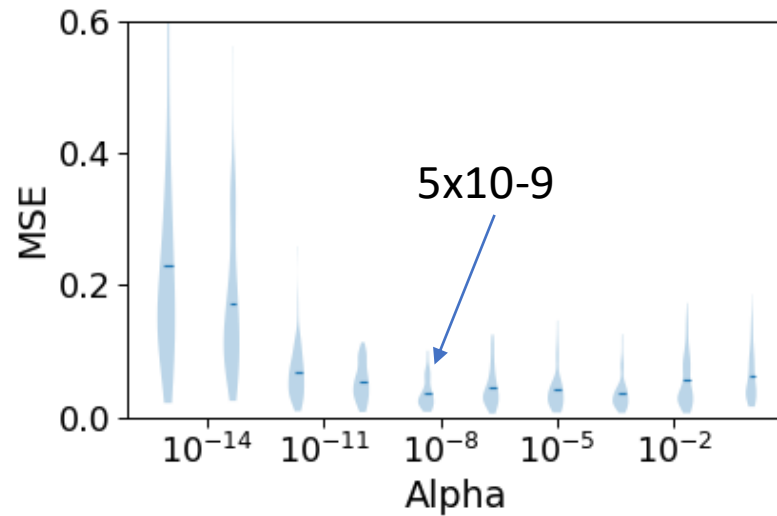
Quantity of interest



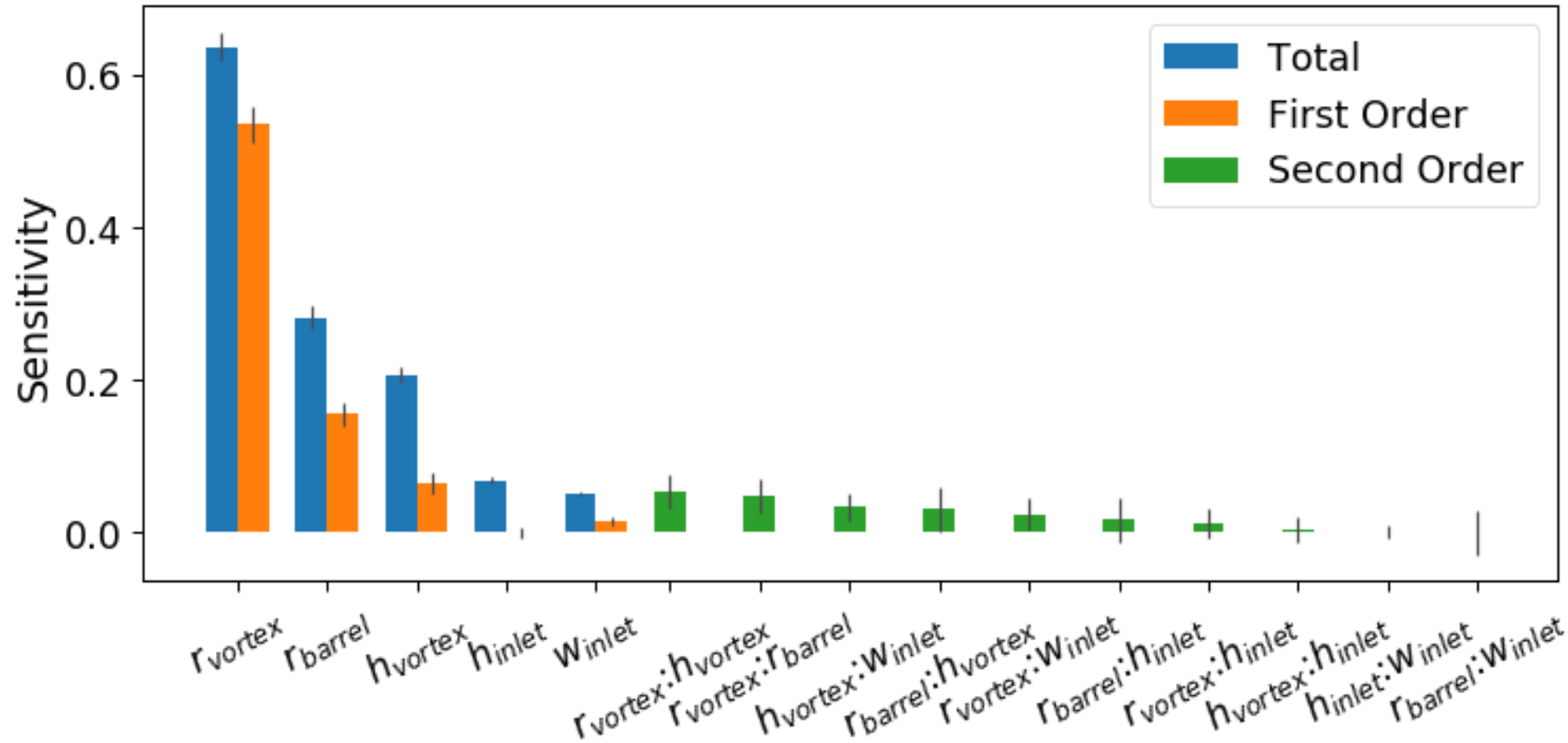
Surrogate model: Gaussian Process

Alpha: noise level
or smoothing of
the data

10% hold out CV

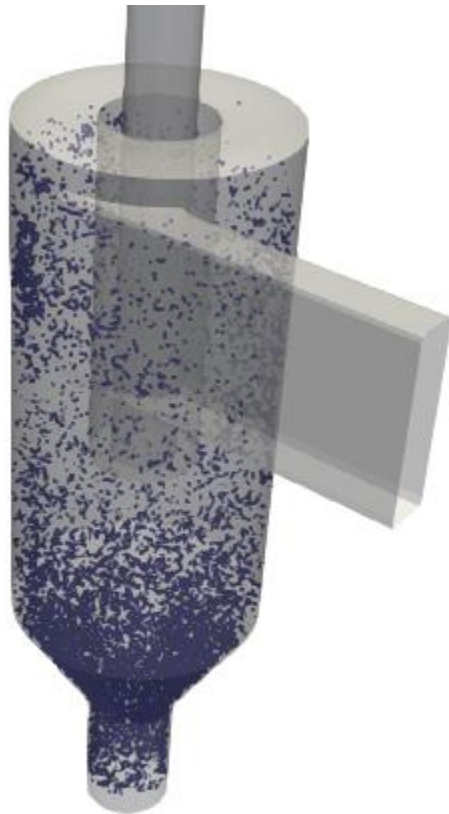


Using SALib



Optimization

Original



Optimal



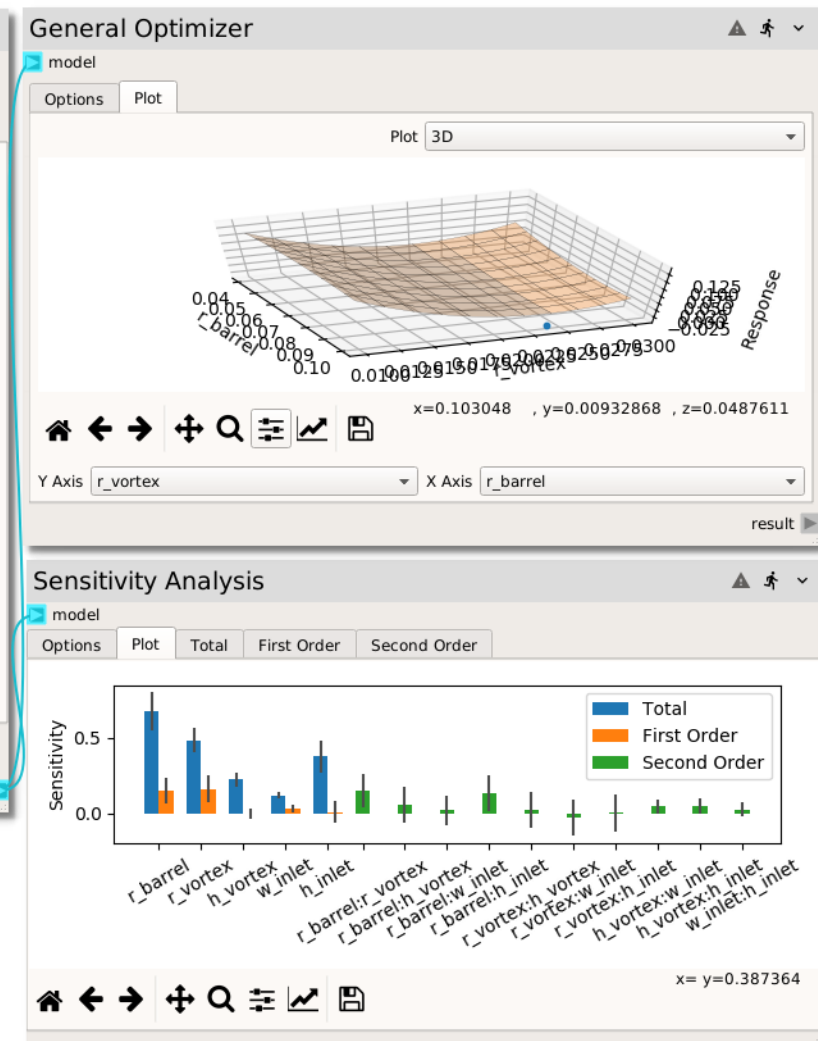
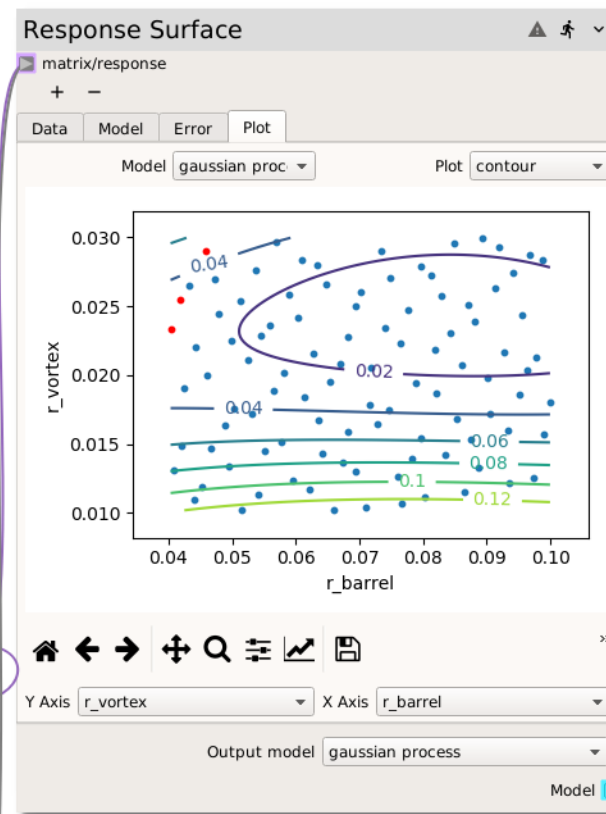
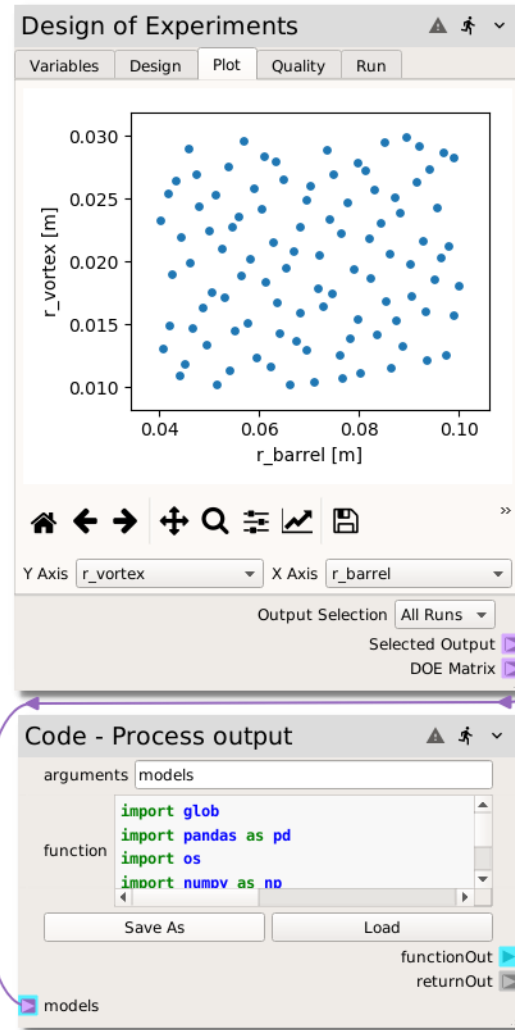
Using differential evolution

- 11 times lower pressure drop
- 2.3 times lower mass loss

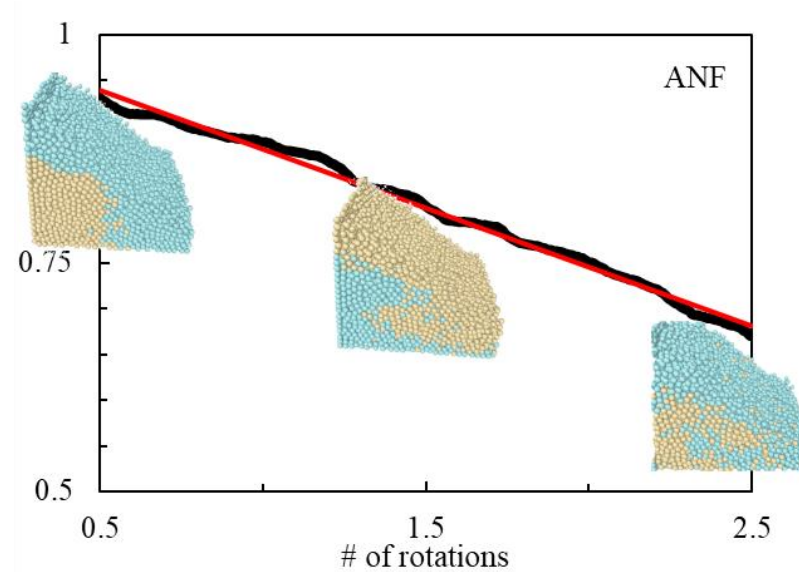
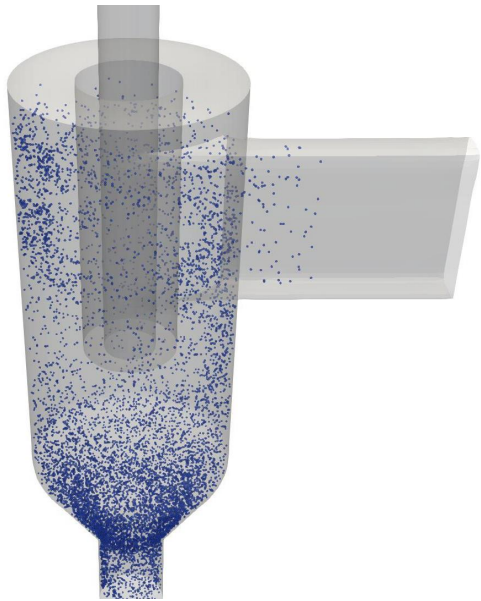
Variable	Original (m)	Optimal (m)
r_{barrel}	0.06	0.096
r_{vortex}	0.015	0.026
h_{vortex}	0.4	0.373
h_{inlet}	0.08	0.12
w_{inlet}	0.02	0.015

Edge of design space

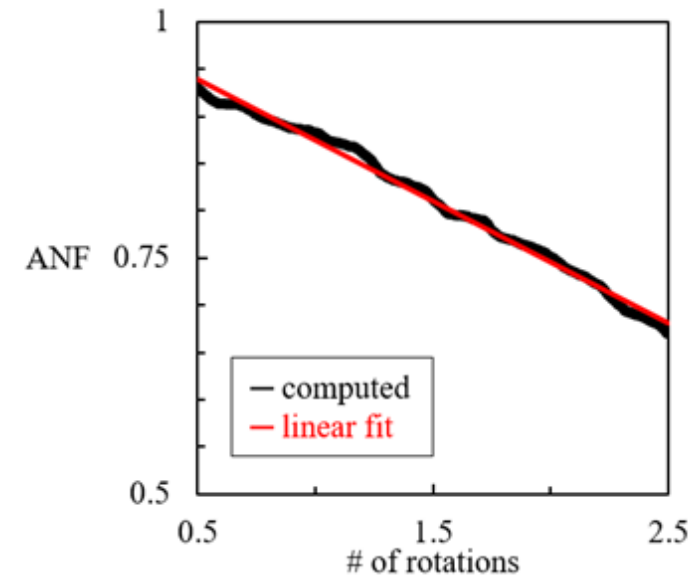
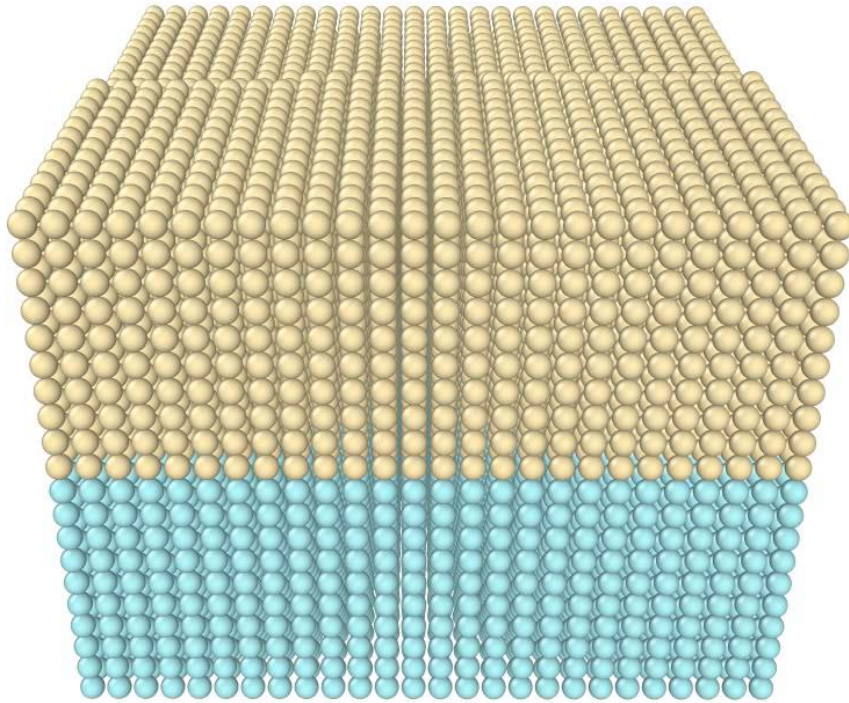
Putting it all together



Examples



Quantify mixing as the rate of decay of the **Alike Neighbor Fraction (ANF)**



ANF = fraction of particles within $2.5r_p$ -radius of a given particle with the same color (averaged over all particles)

Model: MFiX-DEM

Rotation induced by angular gravity

Geometry considered fixed/known

Seven model parameters

considered as unknown quantities

- Six of which are taken from measurements of real particles

Model uncertainties considered:

DEM Model Parameter	Units	DOE Input Variable	Min	Max
f	(rpm)	x_1	28.8	31.2
d_p	(cm)	x_2	0.26	0.35
ρ_p	(g/cm ³)	x_3	2.22	2.92
e_{pp}	-	x_4	0.92	0.9999
e_{pw}	-	x_5	0.58	0.9999
μ_{pp}	-	x_6	0.1	0.45
μ_{pw}	-	x_7	0.02	0.42

DOE → Simulations → Surrogate

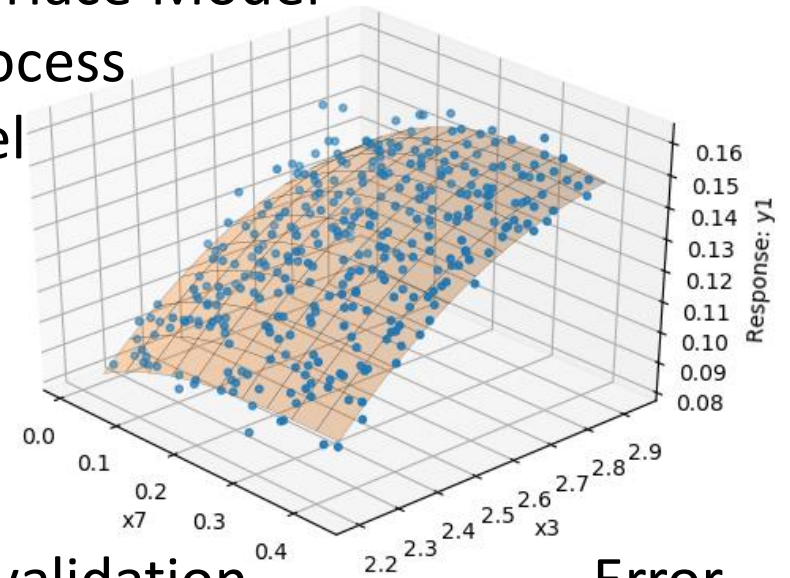
Design of Experiments

- Latin Hypercube
- Genetic optimization
- 7-D space
- 345 samples (overkill?)

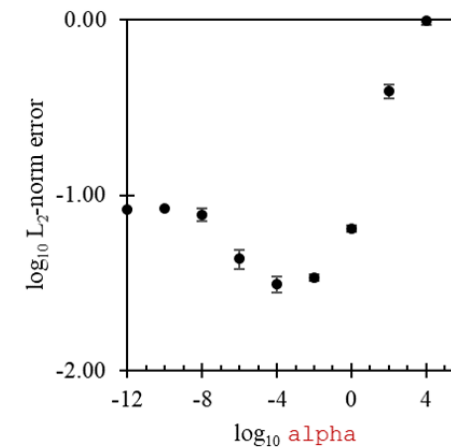
Response Surface Model

Gaussian Process

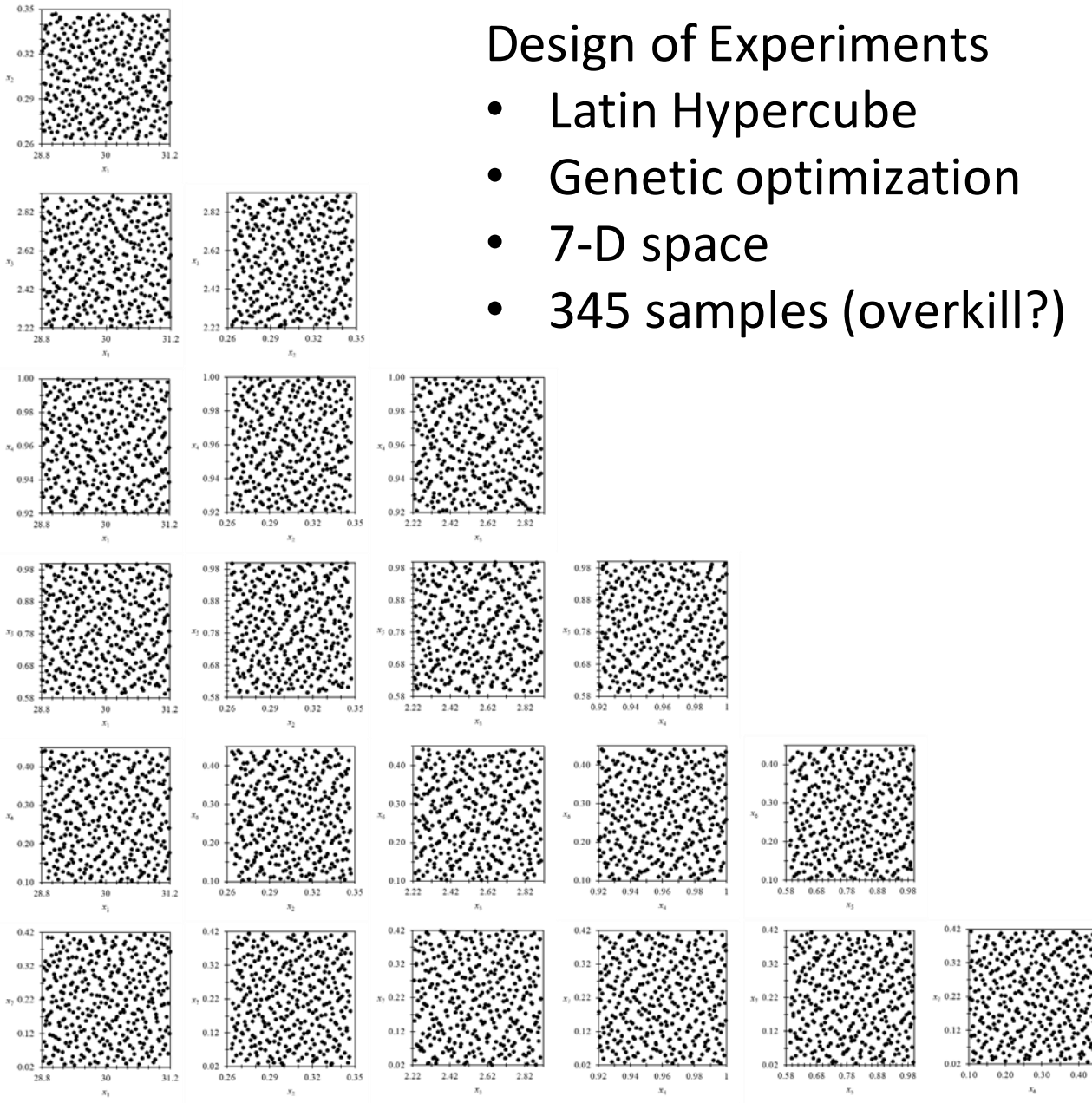
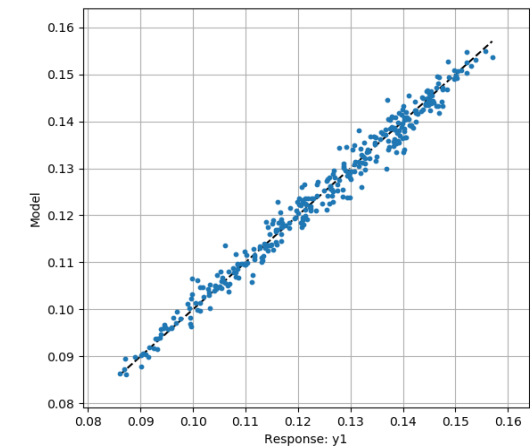
- RBF kernel



Cross-validation

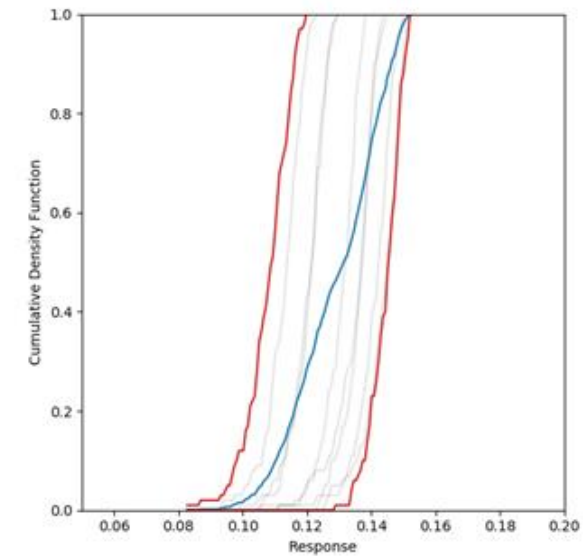
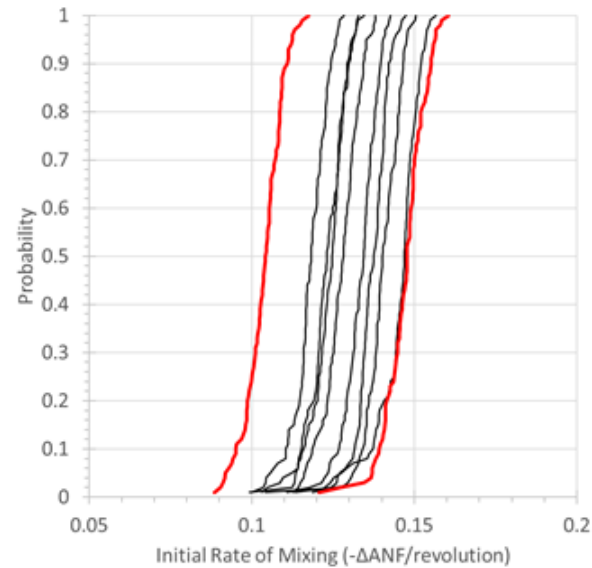
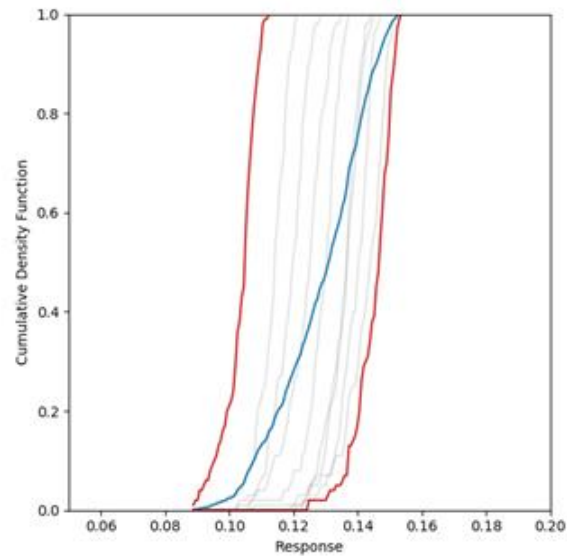


Error



Forward Propagation (of input uncertainties)

Hybrid/nested sampling approach of Roy & Obekampf
10 epistemic samples, each with 100 aleatory samples

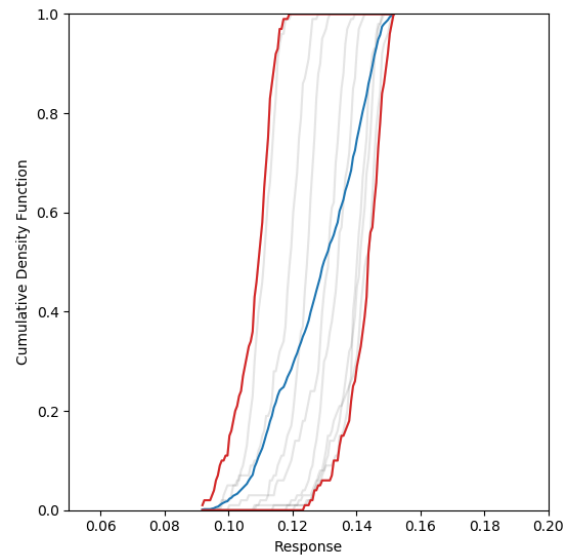


Original direct sample p-box of Dahl et al (2019)
Examples of surrogate model propagated p-boxes

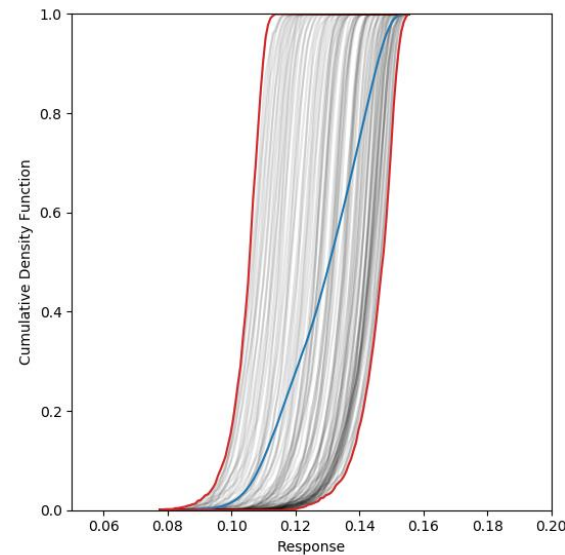
Forward Propagation (of input uncertainties)

What if... we decide the p-box is too coarse for our use purpose and we need to increase the number of samples?

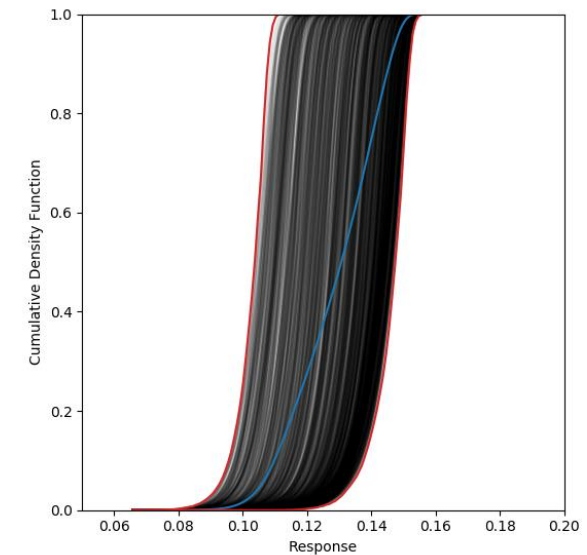
- Direct/full model: **expensive**
- Surrogate model: (once constructed) **cheap**



10×100

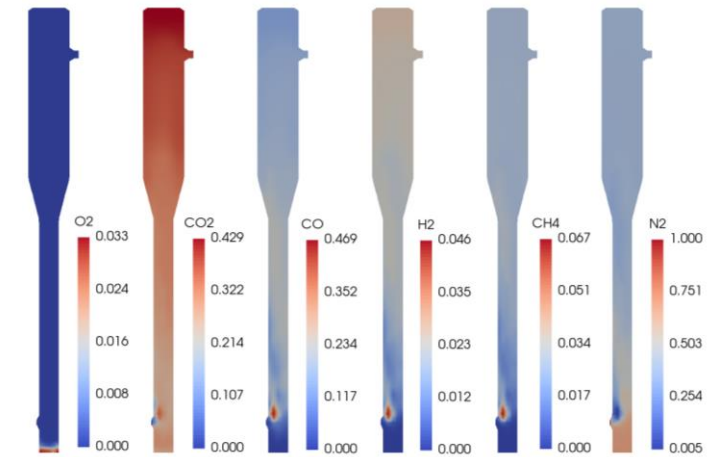
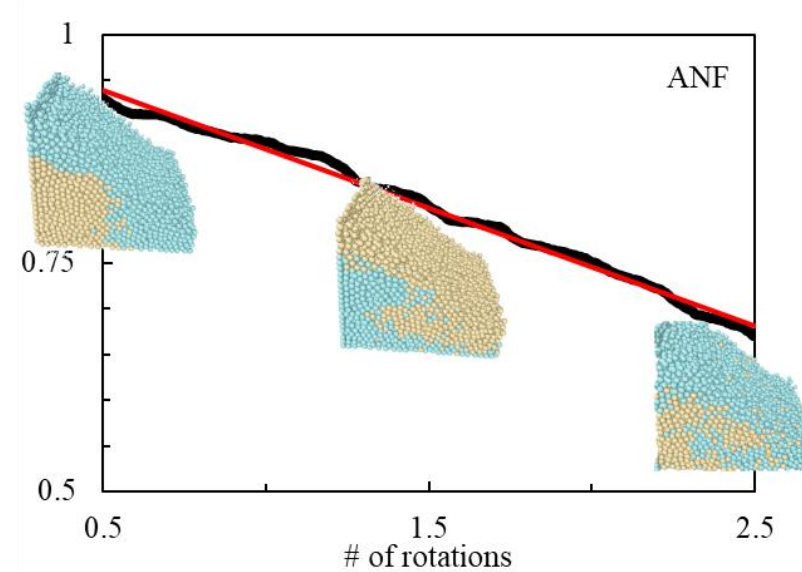
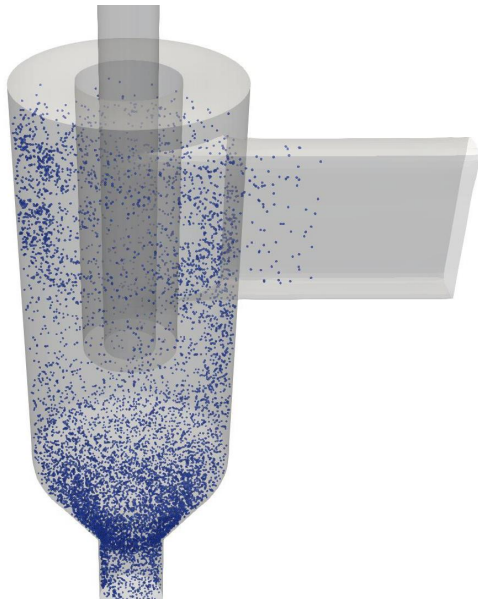


100×1000

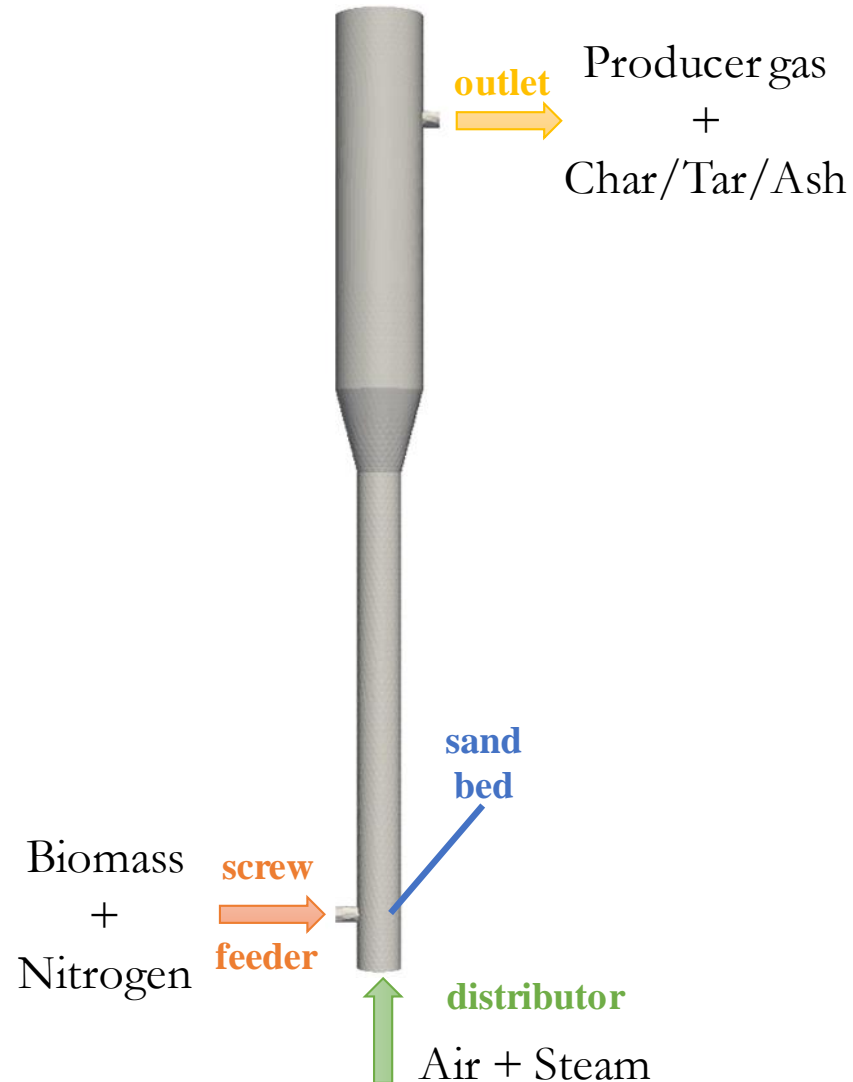


1000×10000

Examples



Example | Biomass Gasifier



Control Variables:

- x_1 = **biomass** mass flow rate
- x_2 = **inlet gas** mass flow rate
- x_3 = **inlet gas** steam mass fraction

System Response/QoI:

- y_1 = H_2/CO molar ratio of **product syngas** (time-averaged from 25 to 30s)

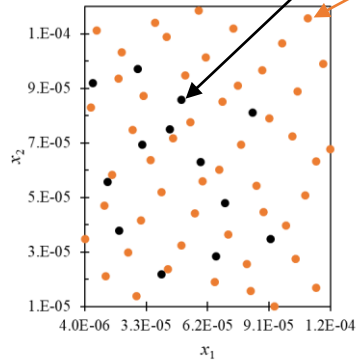
Objective Function:

- $\min \frac{x_3 x_2}{x_1} \Big|_{y_1=2}$, minimize the amount of steam required to produce a syngas with a 2:1 hydrogen to carbon monoxide ratio

DOEs and Results

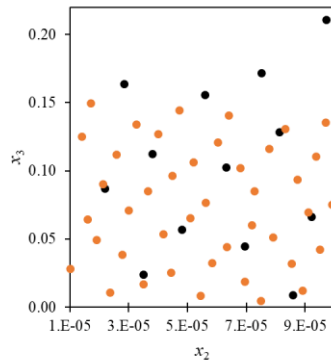
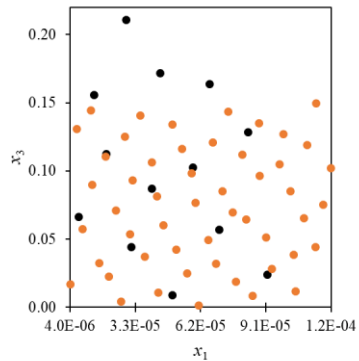
Sub-set of preliminary,
scoping DOE

Secondary, refined DOE

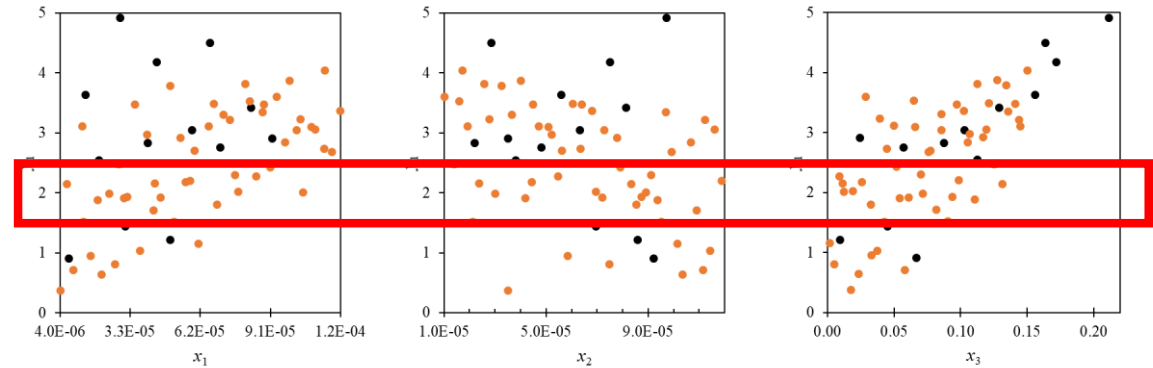


Nodeworks

- Latin Hypercube
- Genetic Optimization
- Composite DOE not LH



Results for the QoI, H₂/CO

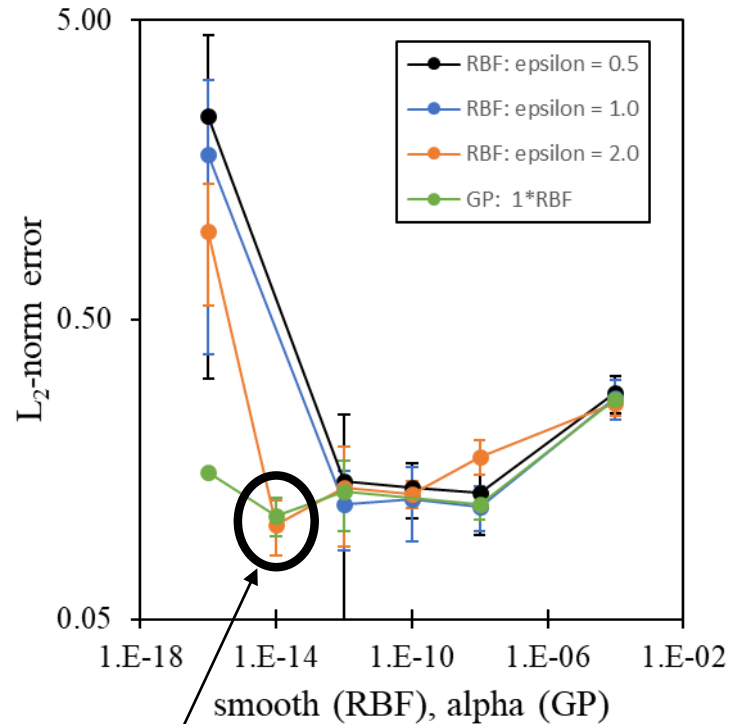


Region of interest

Q: How do we get a continuous surface of $y_1 = 2$

A: Construct a (4-D) response surface surrogate model
and extract the (3-D) iso-surface characterizing $y_1=2$

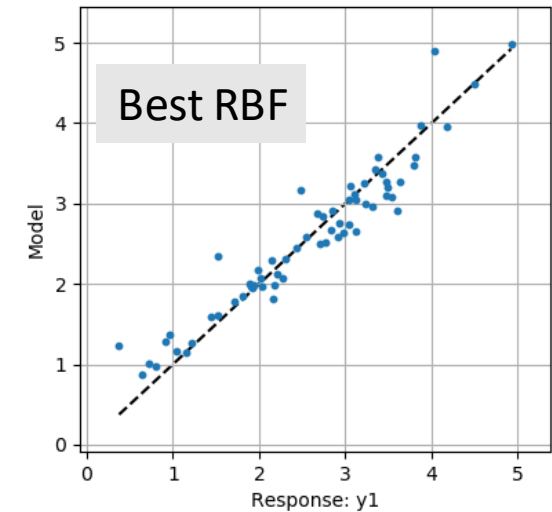
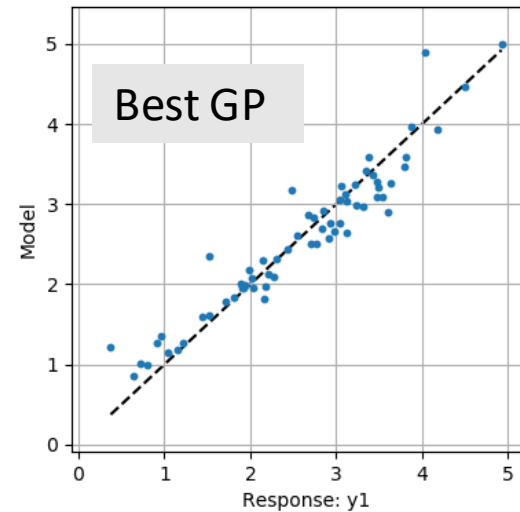
Cross-Validation for the QoI, H_2/CO



Best surrogate models:

- Radial basis function (RBF) with smoothing parameter of $1e-14$
- Gaussian process (GP) with RBF kernel and noise parameter of $1e-14$

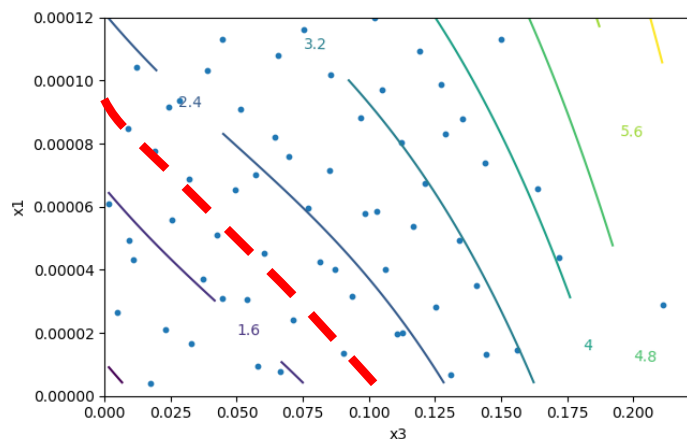
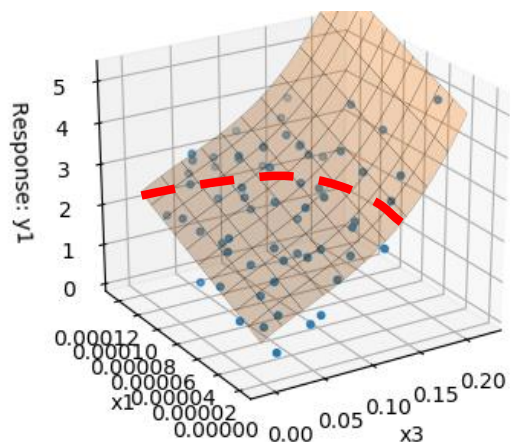
Full Model Error



Selection: GP (more consistent)

Optimization

Optimize: RSM == 2

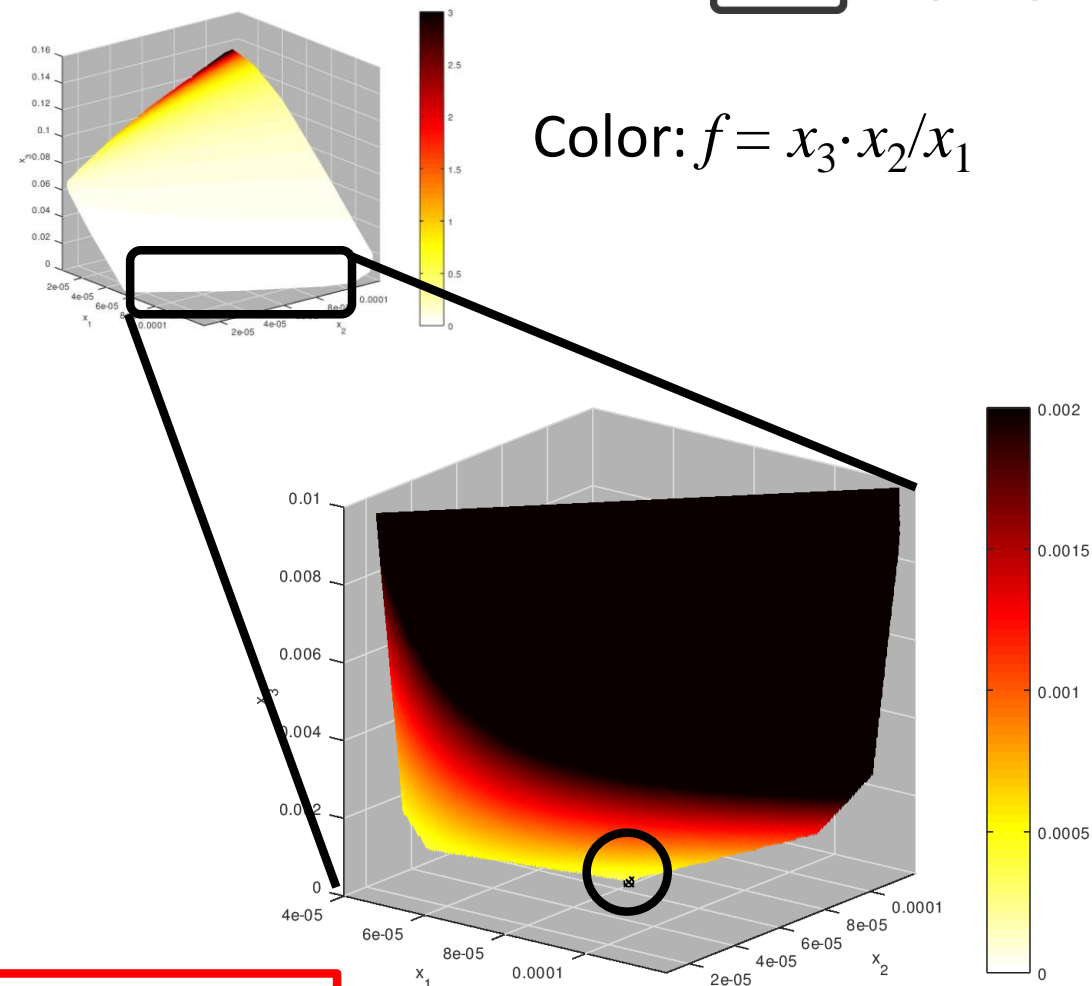


Iso-surface

x_1	x_2	x_3	y_1
#	#	#	2.000..
#	#	#	1.999..
#	#	#	1.999..
#	#	#	2.000..

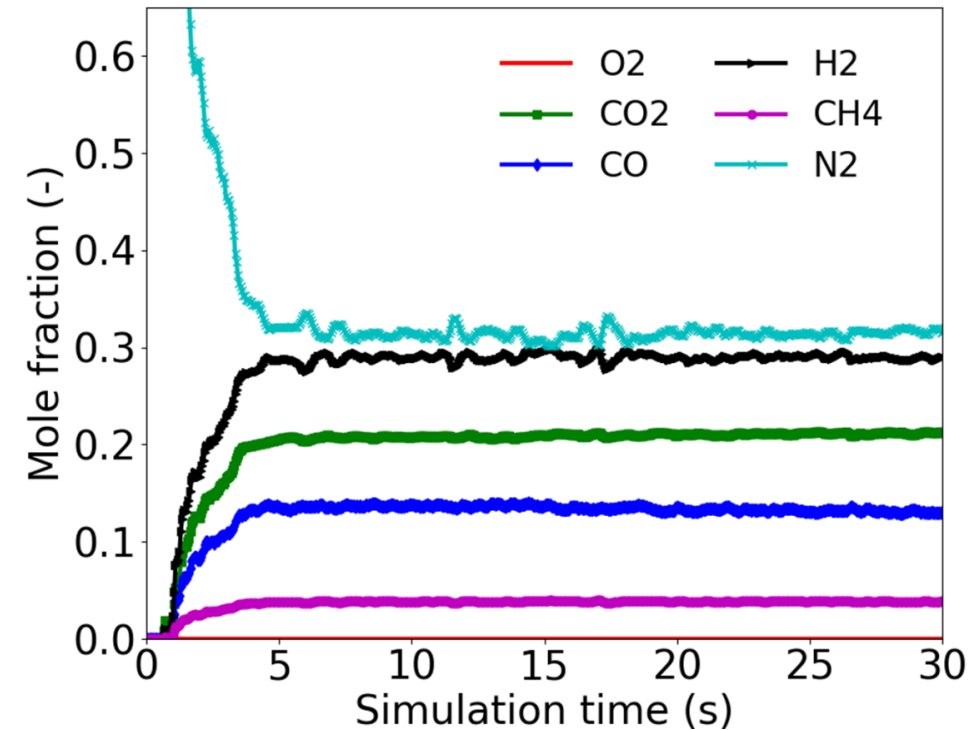
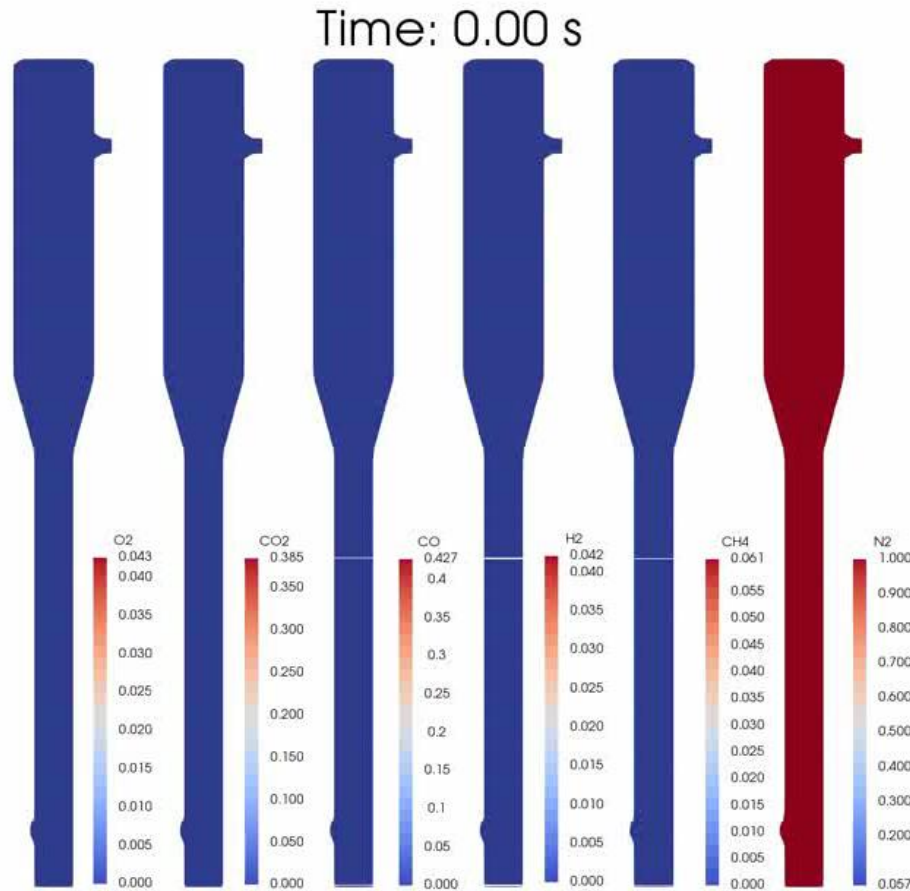
...

Another surrogate?
Nah, the GP is cheap.
Just iterate many times
and interpolate.



Optimal condition

Validation of (surrogate) Optimum



$x_1 = 0.086$ (g/s), $x_2 = 0.054$ (g/s), $x_3 = 4.8 \times 10^{-4}$, $\hat{y}_1 = 2$, $y_1 = 2.2$
(within expected error from cross-validation test)