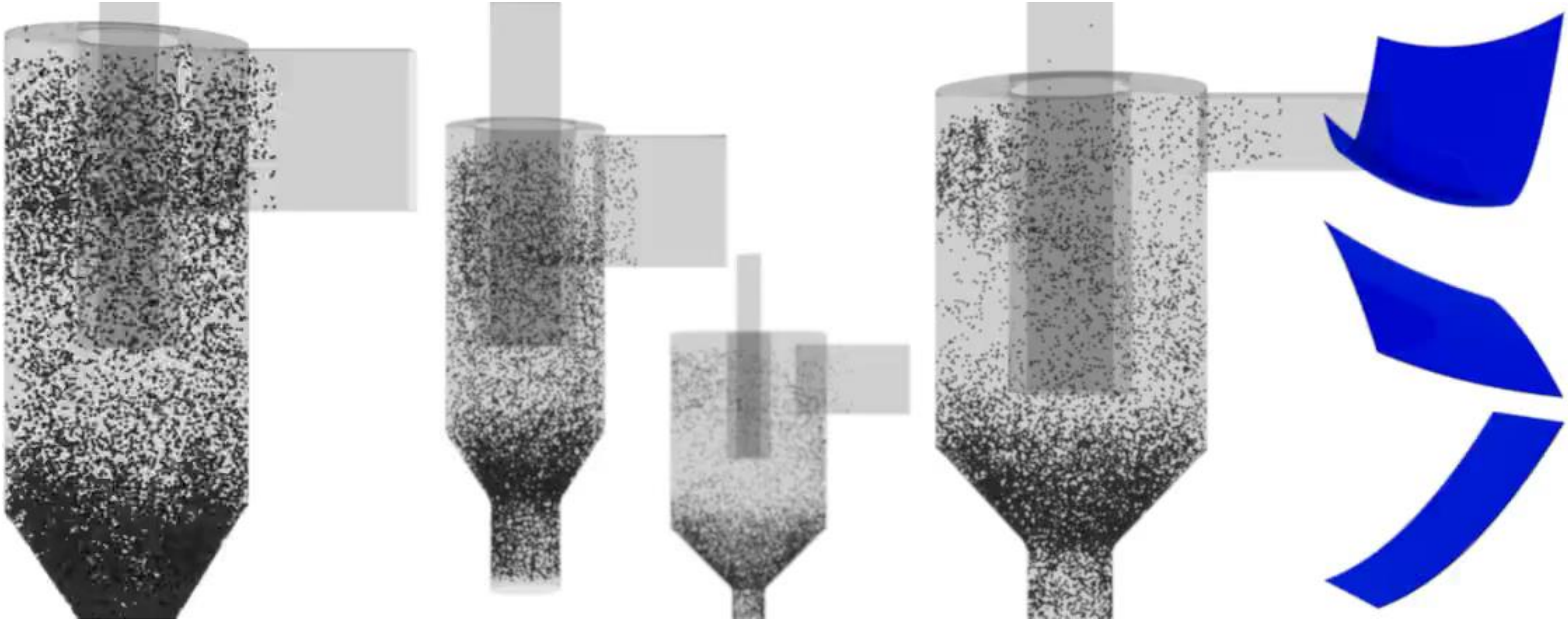
Nodeworks

NATIONAL ENERGY TECHNOLOGY LABORATORY

# Surrogate modeling and analysis toolset

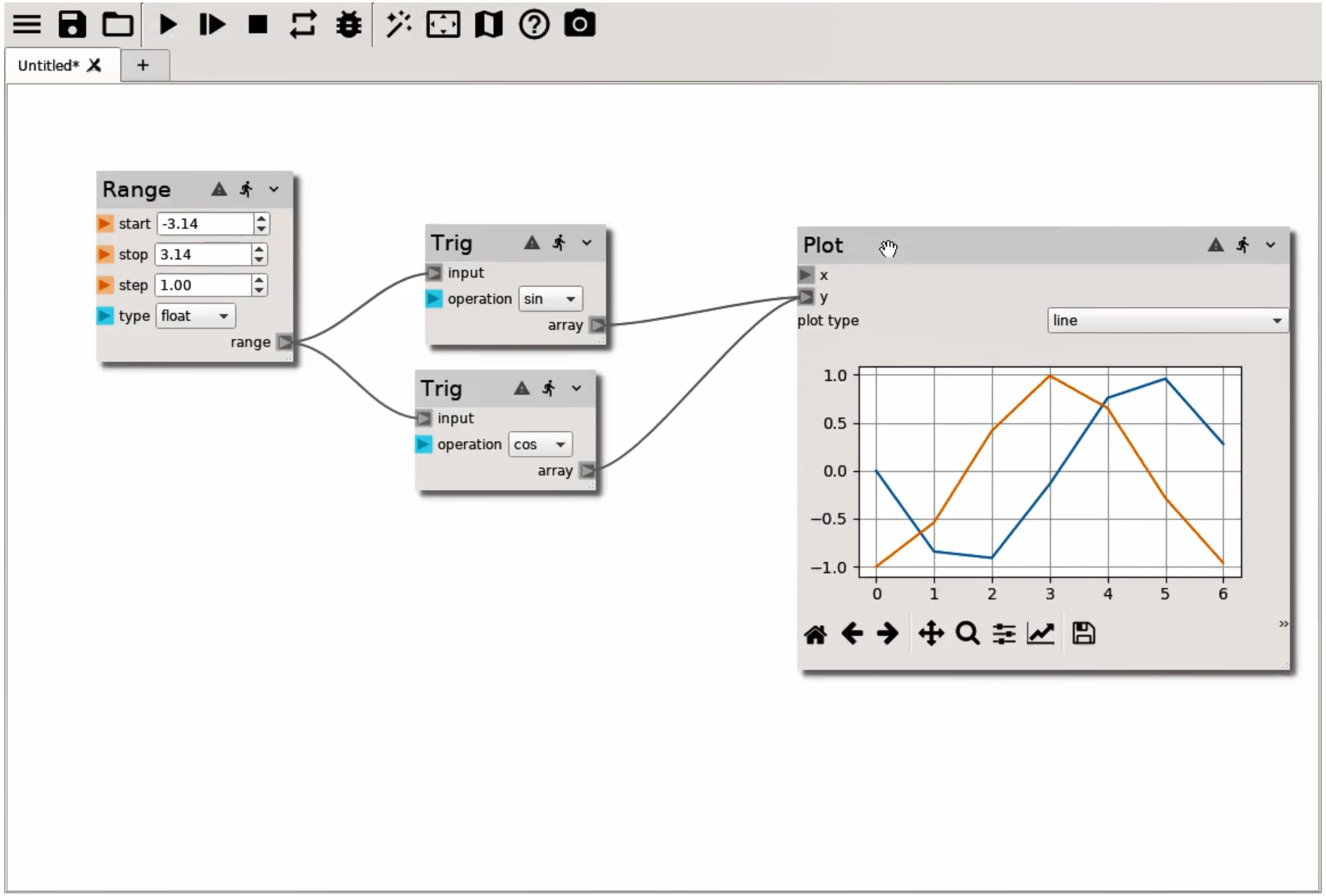
***Justin Weber, William Fullmer, Aytekin Gel***

*Research and Innovation Center (RIC)*



# Nodeworks

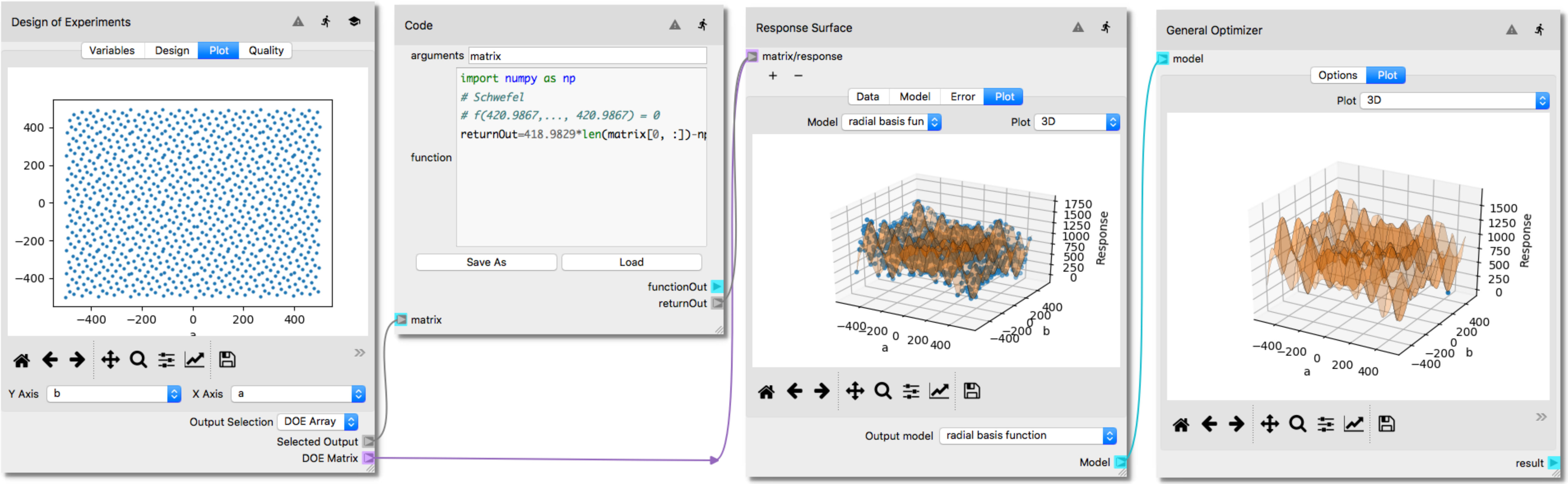
- Application and framework for graphical programing through the use of nodes and connections
- Underlying library for the optimization/UQ work.
- Integrates with the MFiX GUI



# Surrogate modeling and analysis toolset

NATIONAL ENERGY TECHNOLOGY LABORATORY

Design of Experiments → Model evaluation → Response Surface Construction → Optimization

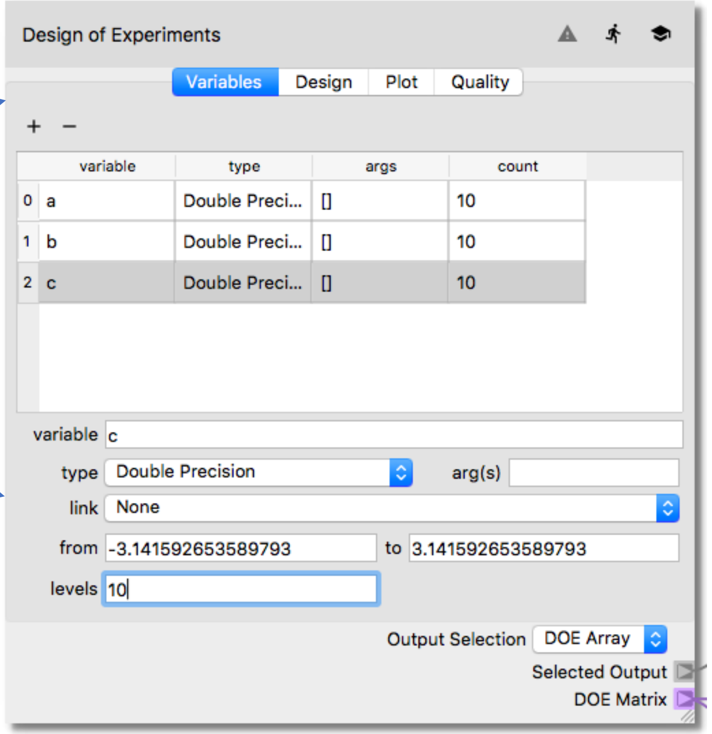


# Design of Experiments | Variables

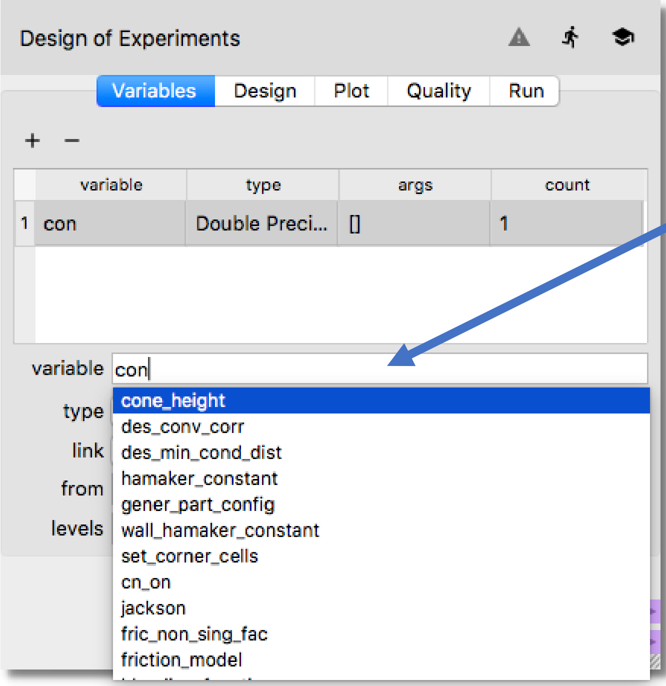
Inside MFiX = MFiX Aware

Add variables

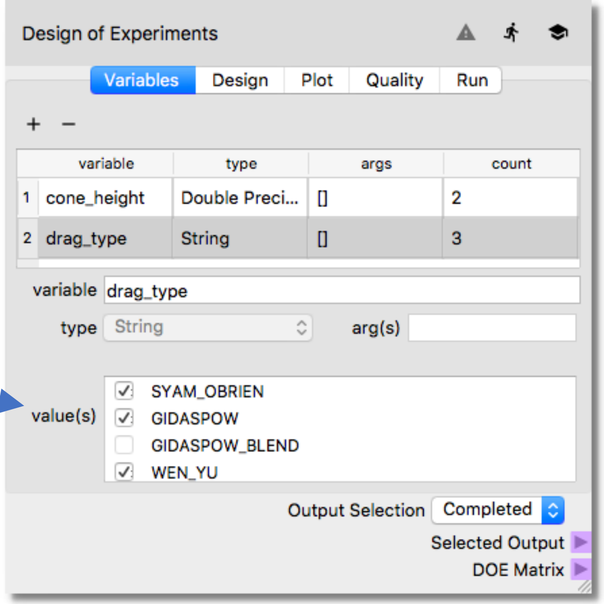
Select variable parameters




Fuzzy search of Parameters and MFiX Keywords



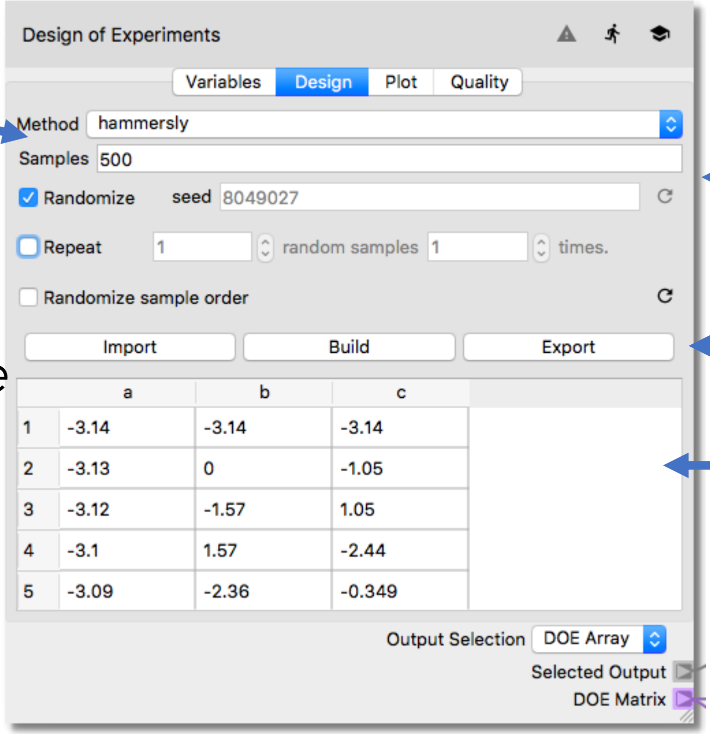

Automatic population of Categorical Variables



# Design of Experiments | Methods

Method

*Factorial*
*Covary*
*Montecarlo*
*Latin hypercube*
*Central composite*
*Sobol*
*Hammersly*
*Halton*



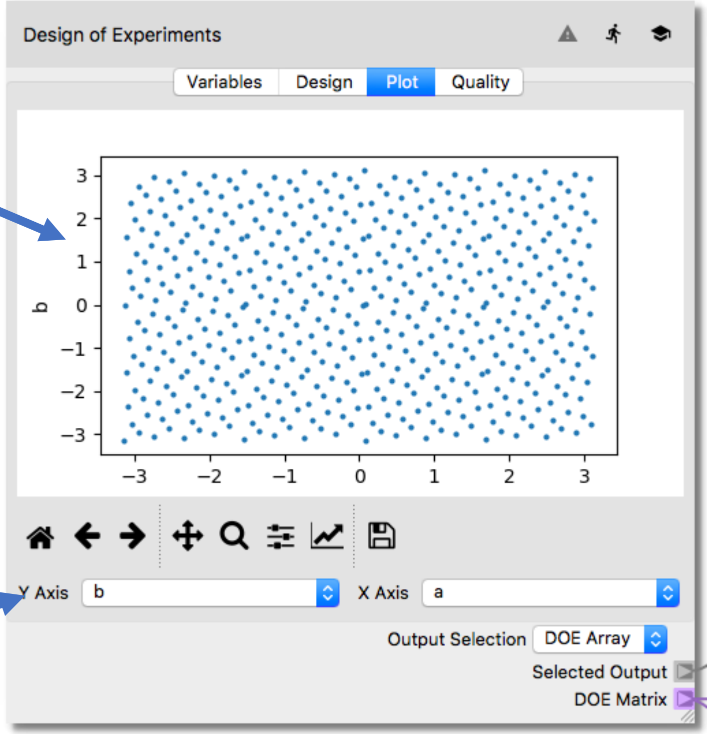

Method Options

Import, Build, Export

Samples

2D plot of samples

Change variables

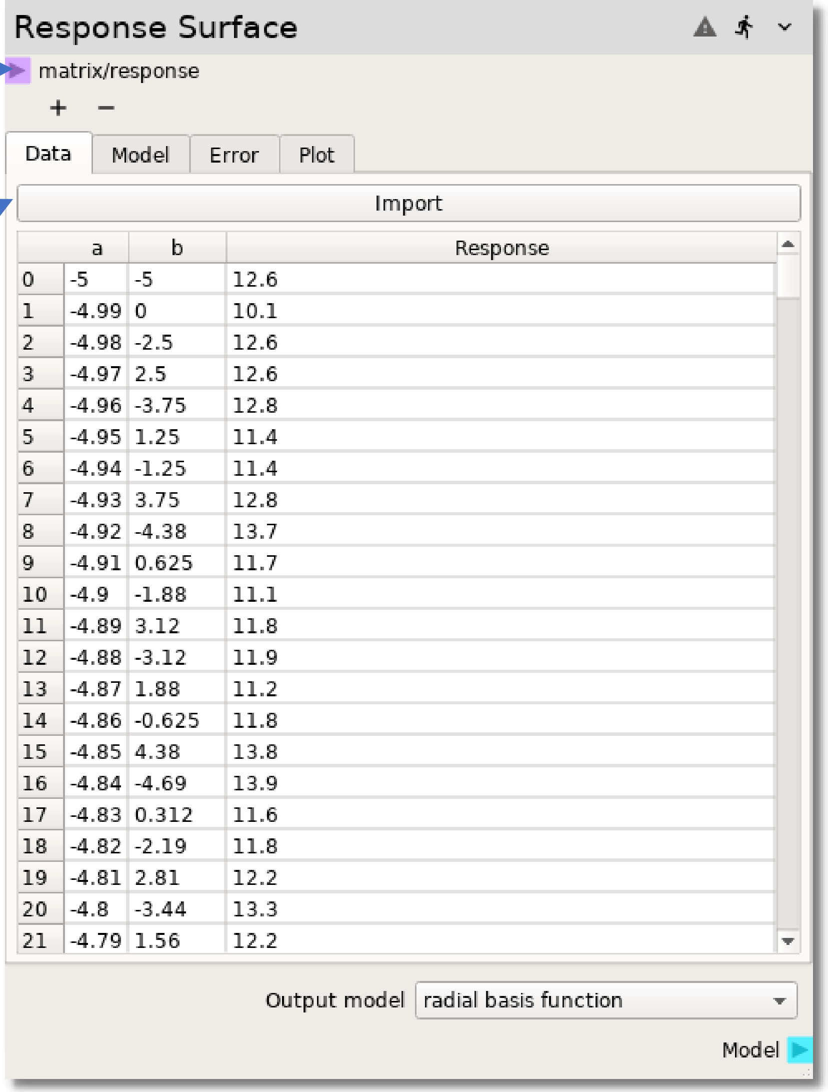


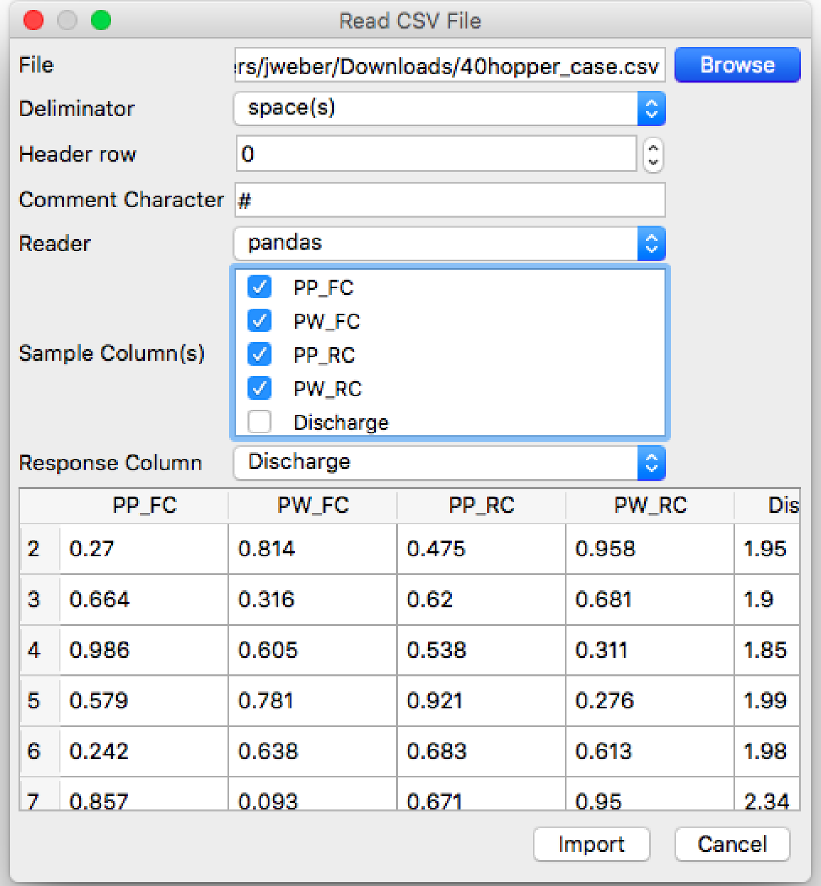
# Response Surface | Samples

Samples + Response

-or-

Read CSV



# Response Surface | Models

Response Surface

matrix/response

+ −

Data | Model | Error | Plot

Points to remove for cross validation → Cross validation points 10 %

Select models to fit →

| 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 |

Error metrics

Model parameters →

Max terms 100

Max degree 1

Penalty 3.0

Fit model → Refit Model(s)

Output model radial basis function

Model

scikit learn:
- gaussian process
- polynomial
- multilayer perceptron
- Support vector machine
- Decision tree
- Random forest

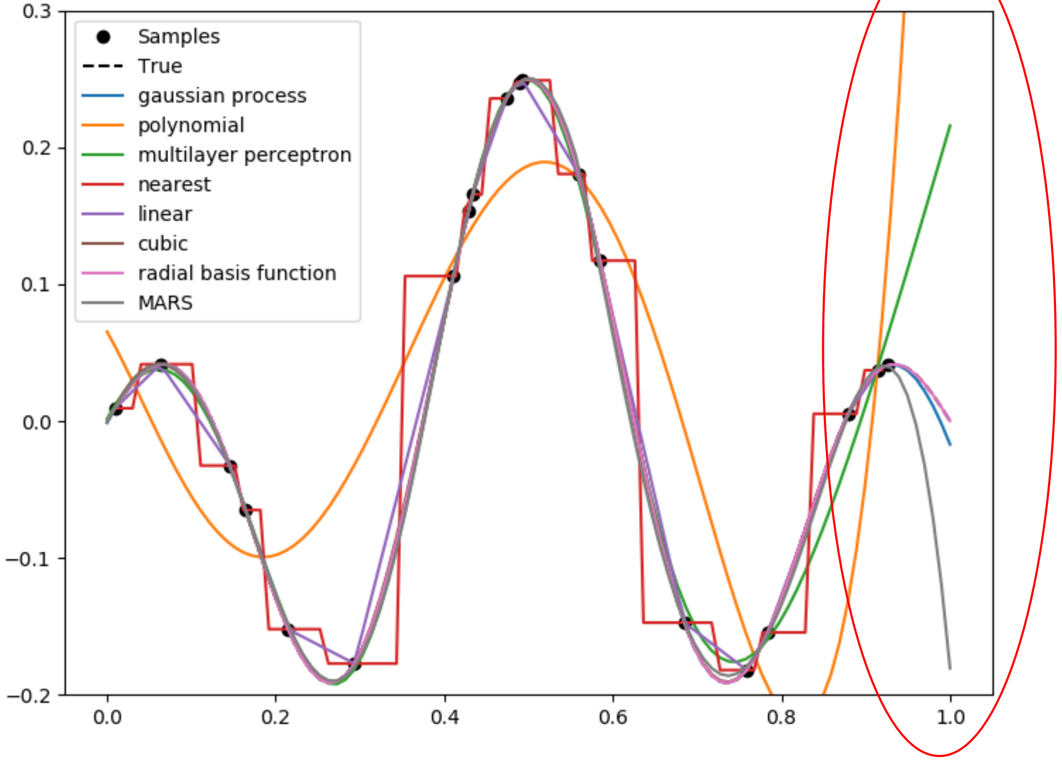
- nearest
- linear
- cubic (d<=2)
- radial basis function

py-earth MARS

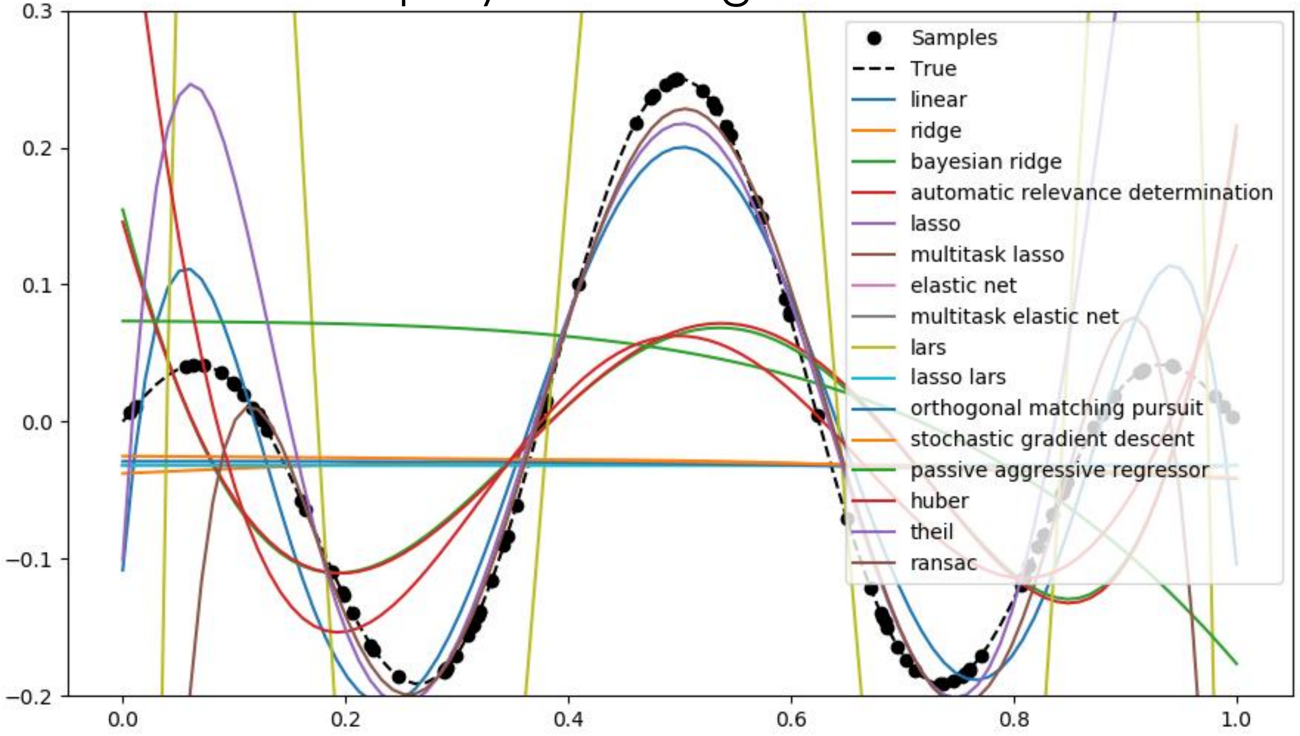
# Response Surface | Models

NATIONAL ENERGY TECHNOLOGY LABORATORY

1D model test




polynomial regressors

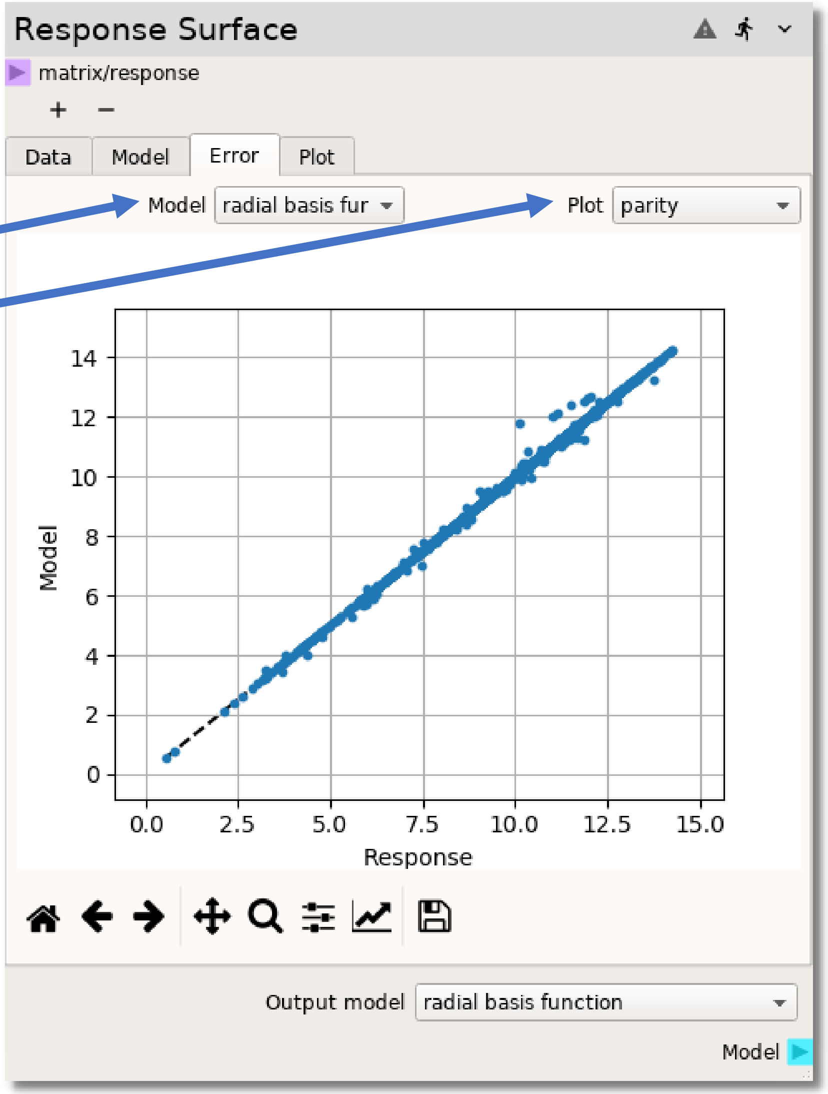



watch the edge!

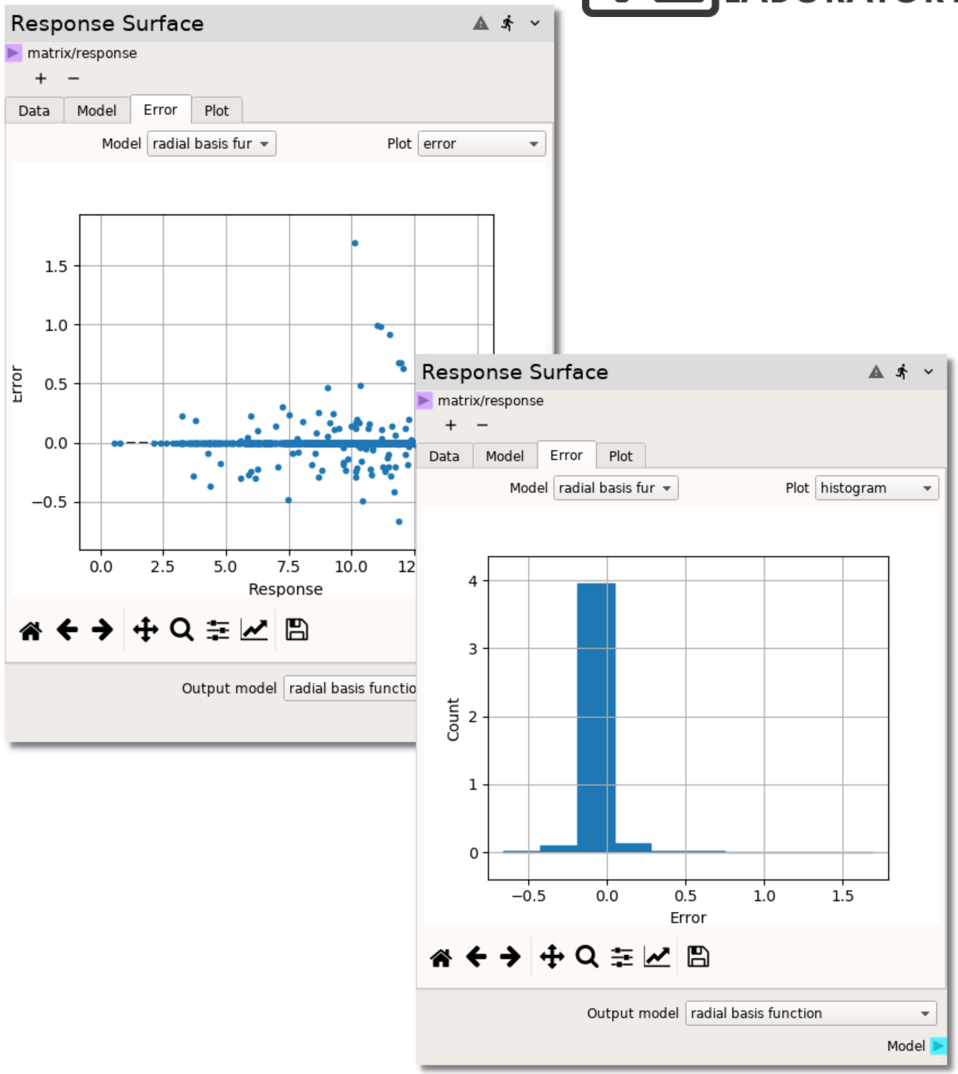
# Response Surface | Error Plots

Select model

Select plot



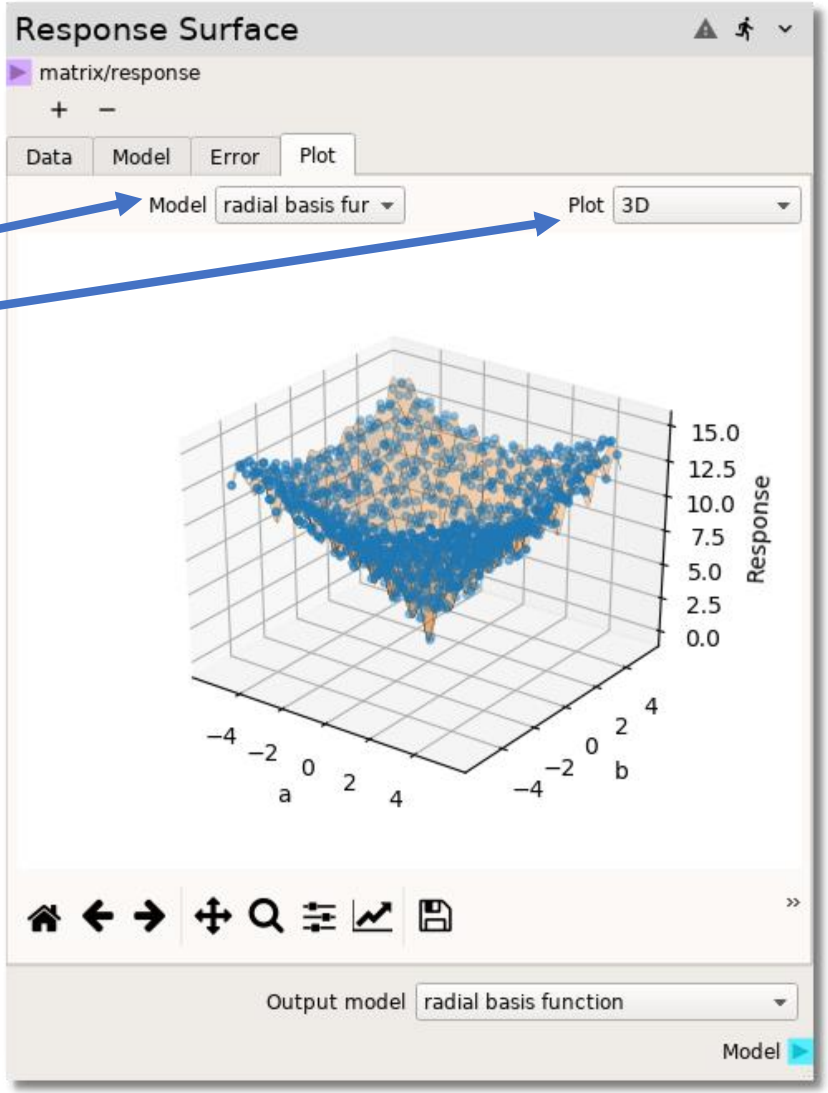

Save/manipulate plot



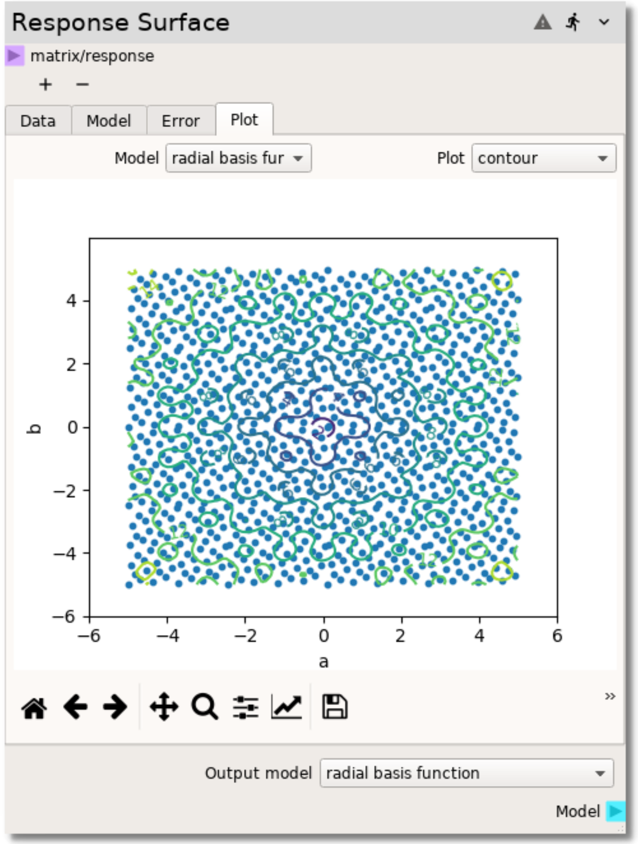
# Response Surface | Plots

Select model

Select plot

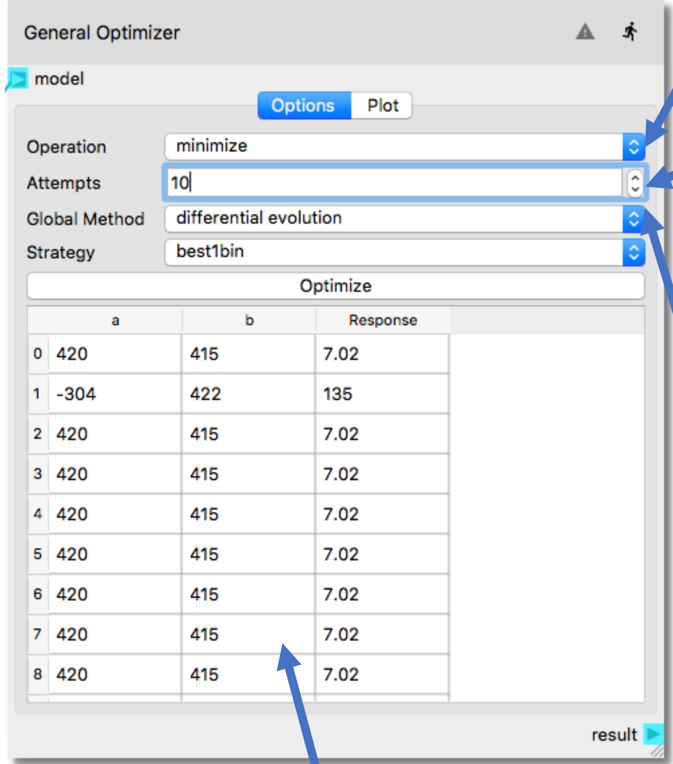



Save/manipulate plot



# Optimization





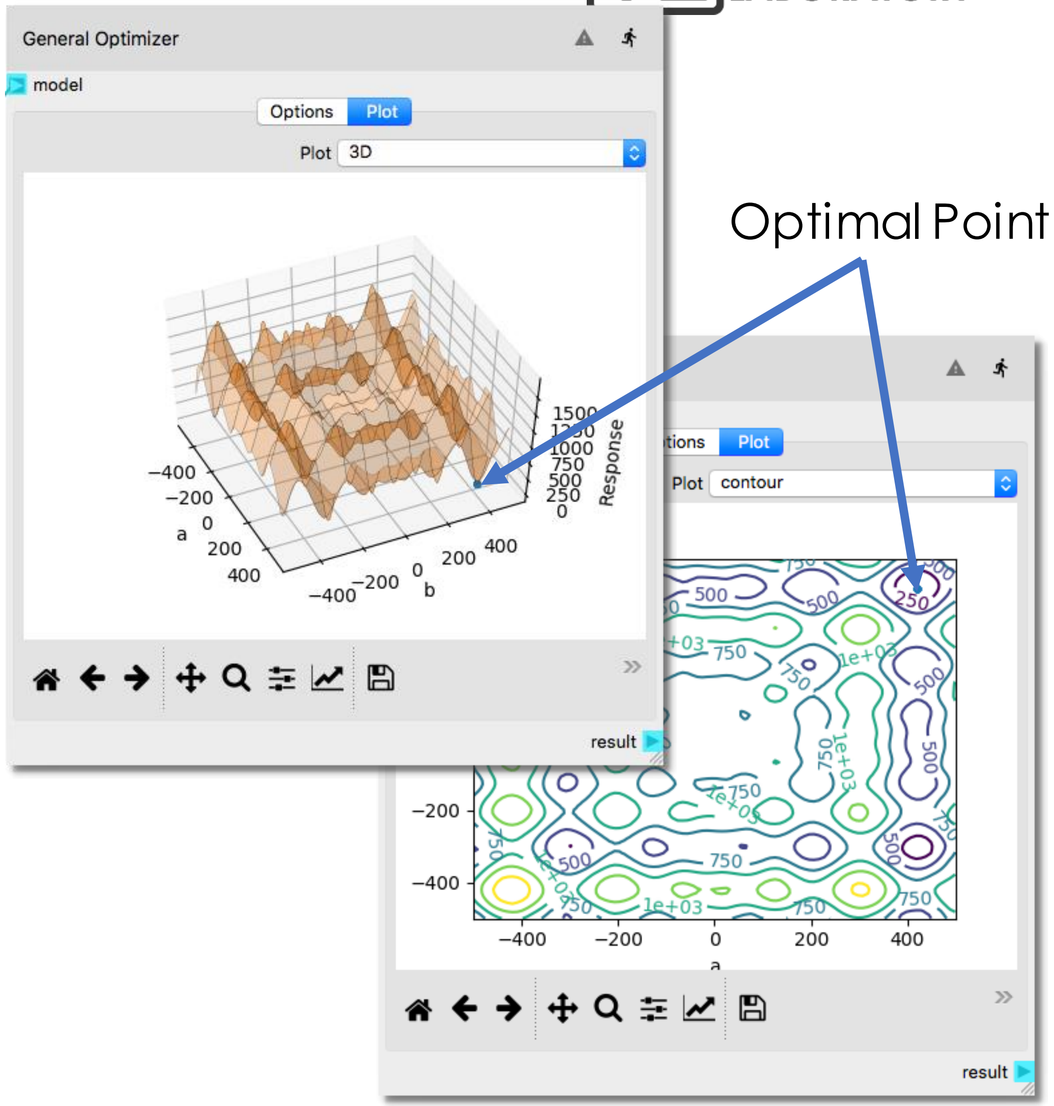

| | a | b | Response |
|---|---|---|---|
| 0 | 420 | 415 | 7.02 |
| 1 | -304 | 422 | 135 |
| 2 | 420 | 415 | 7.02 |
| 3 | 420 | 415 | 7.02 |
| 4 | 420 | 415 | 7.02 |
| 5 | 420 | 415 | 7.02 |
| 6 | 420 | 415 | 7.02 |
| 7 | 420 | 415 | 7.02 |
| 8 | 420 | 415 | 7.02 |

minimize
maximize
find value (root)

number of attempts

**differential evolution**
basin hopping
*Nelder-Mead*
*Powell*
*CG*
*BFGS*
*L-BFGS-B*
*TNC*
*COBYLA*
*SLSQP*

Results of optimization attempts



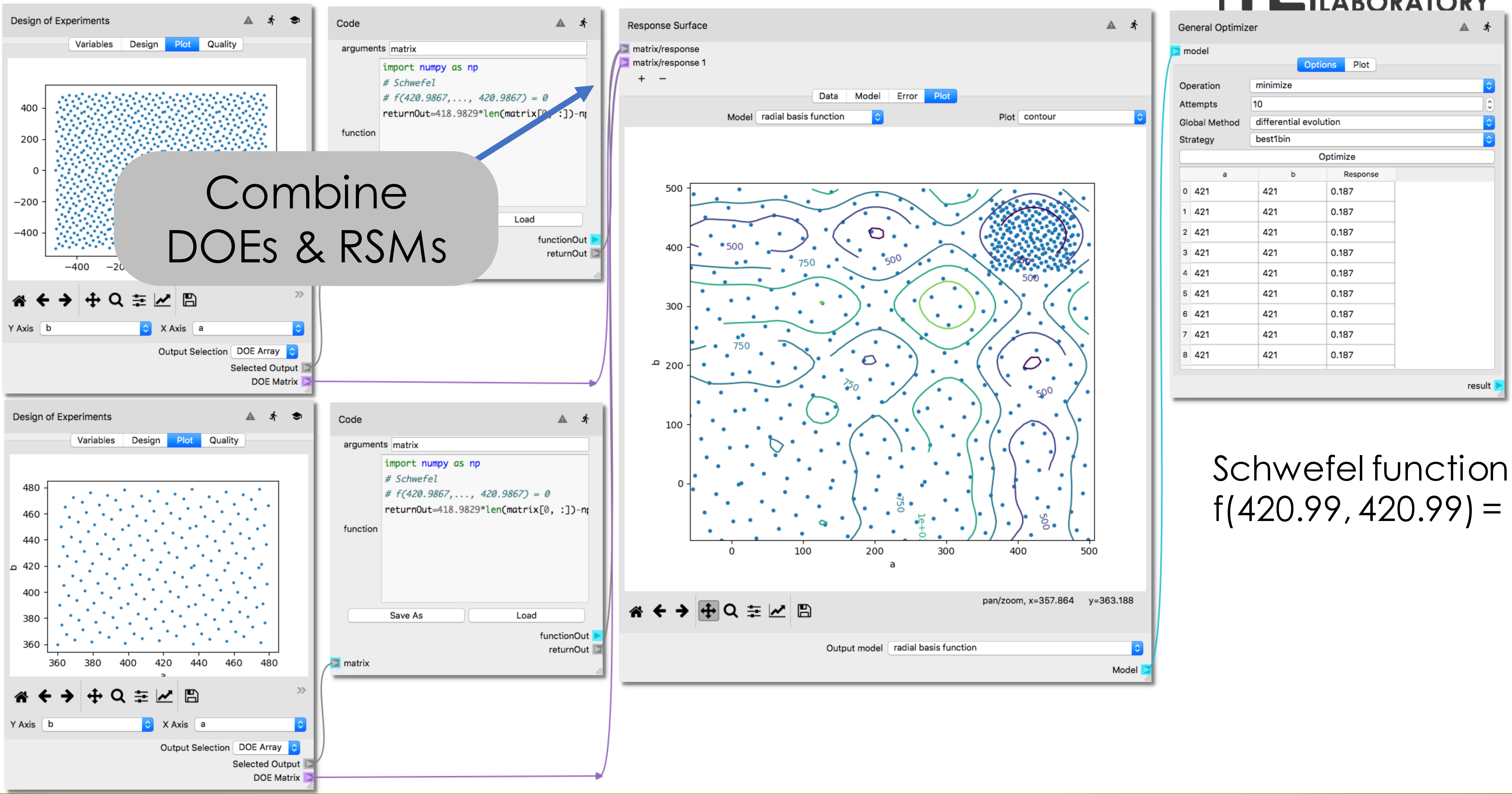

Optimal Point



Schwefel function    f(420.99, 420.99) = 0

# Response Surface | Refinement

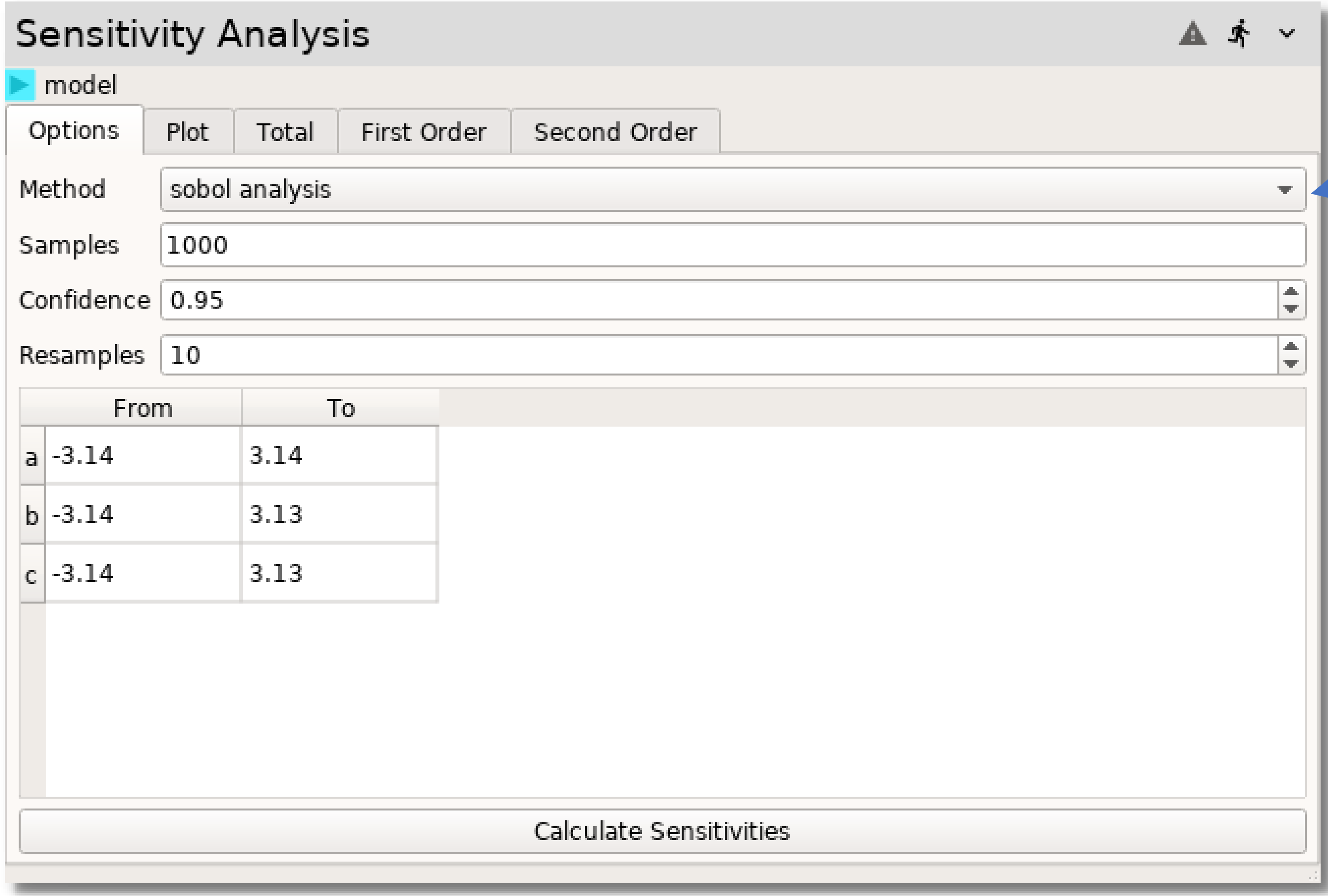





Schwefel function
f(420.99, 420.99) = 0

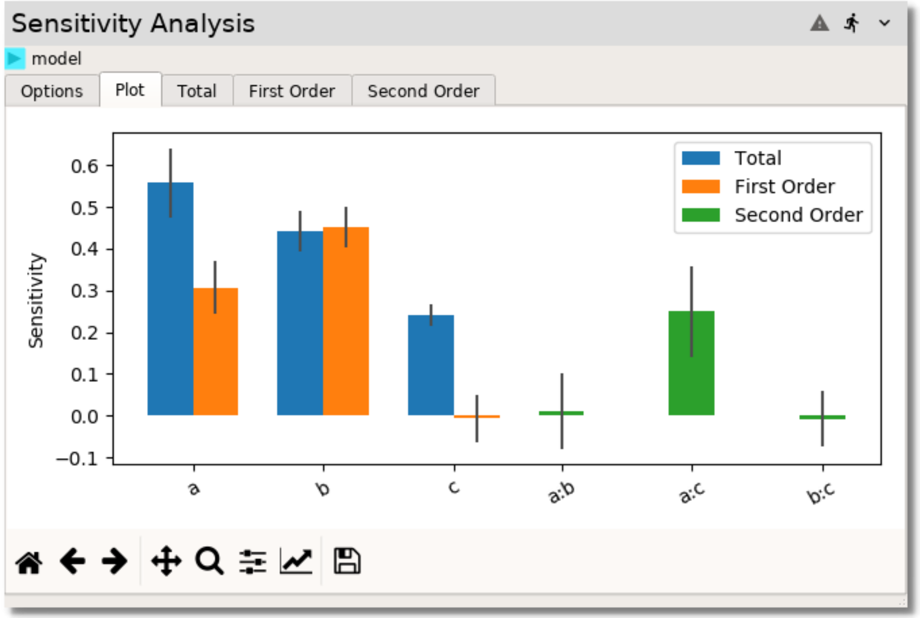
# Sensitivity Analysis



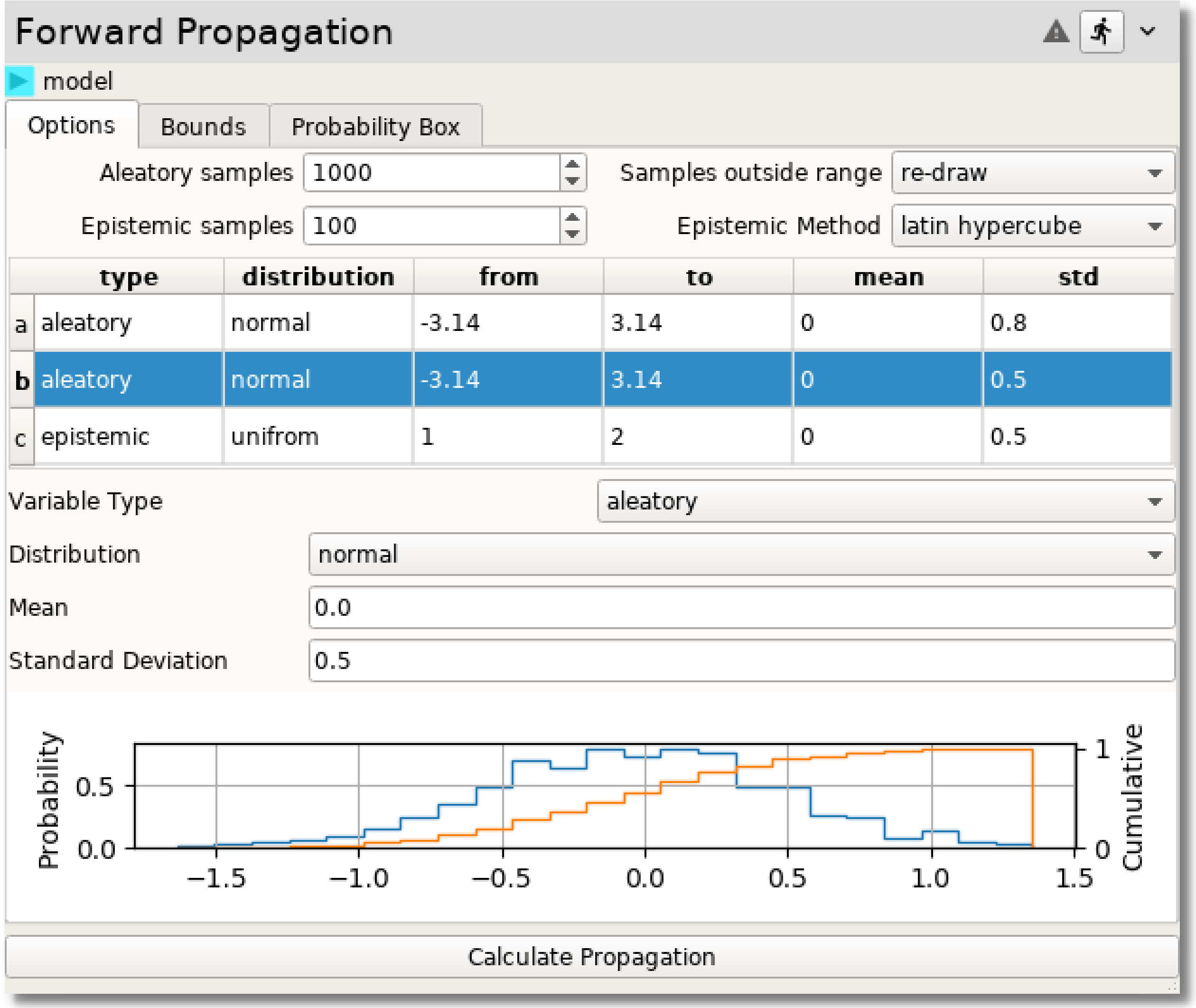



**SALib**

Sobol
Method of Morris
Fourier Amplitude
Delta Moment-independent
Random balance Fourier Amplitude



# Forward Propagation | Variables

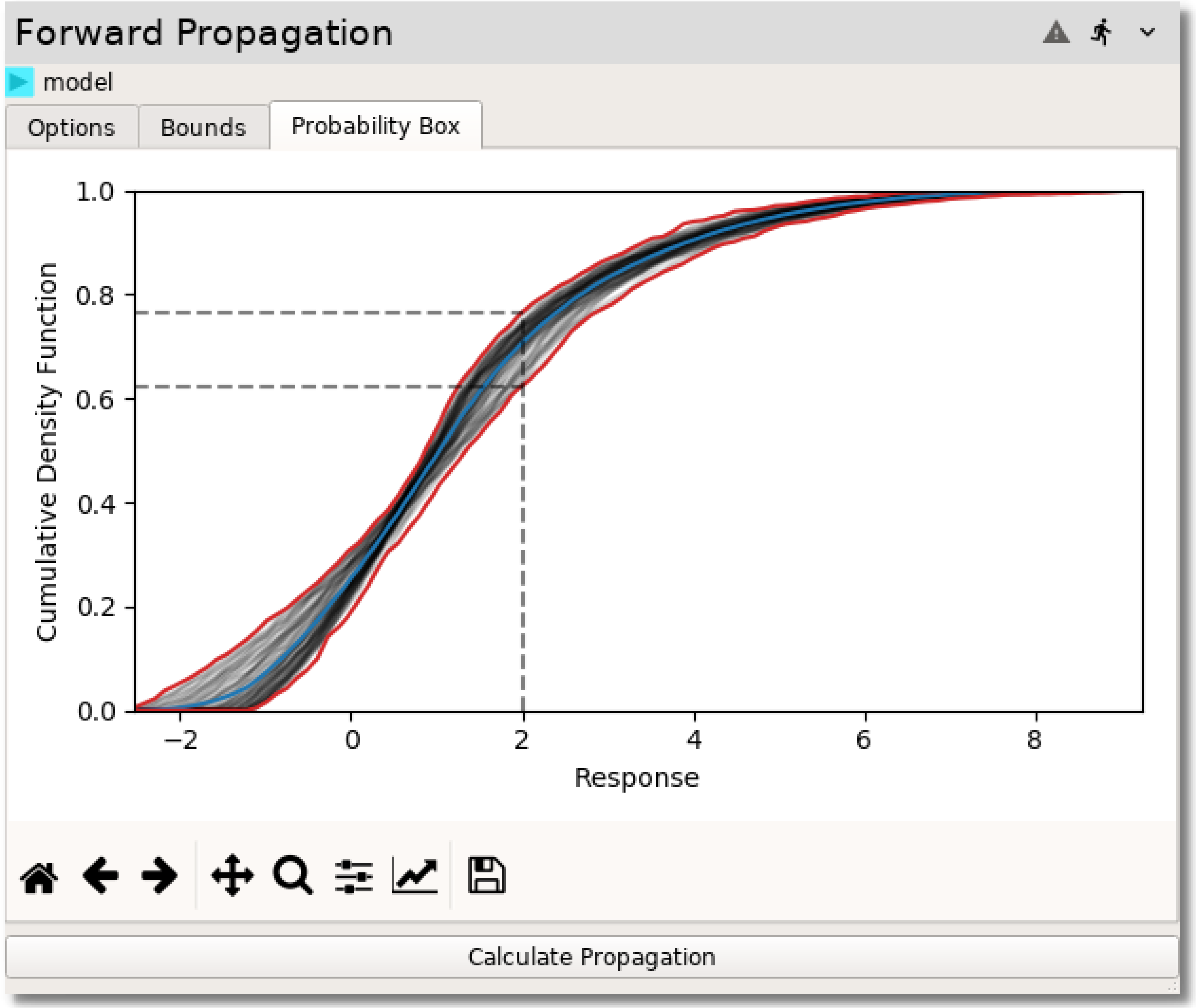


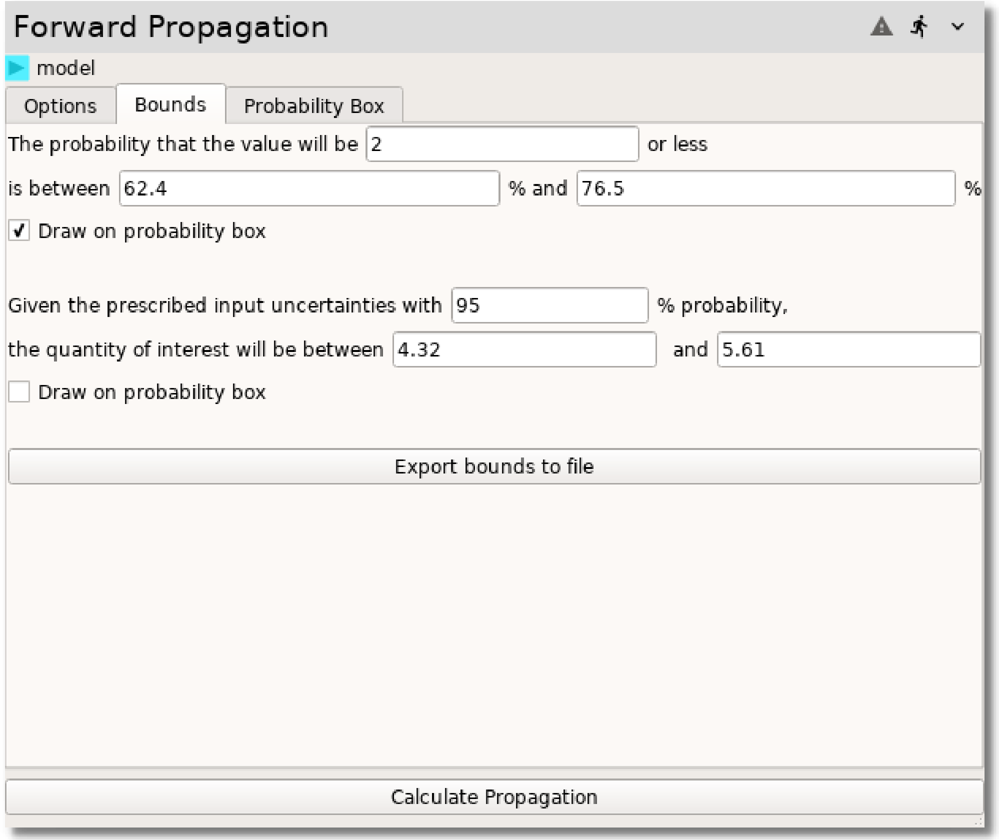

Samples

Variables

Variable options

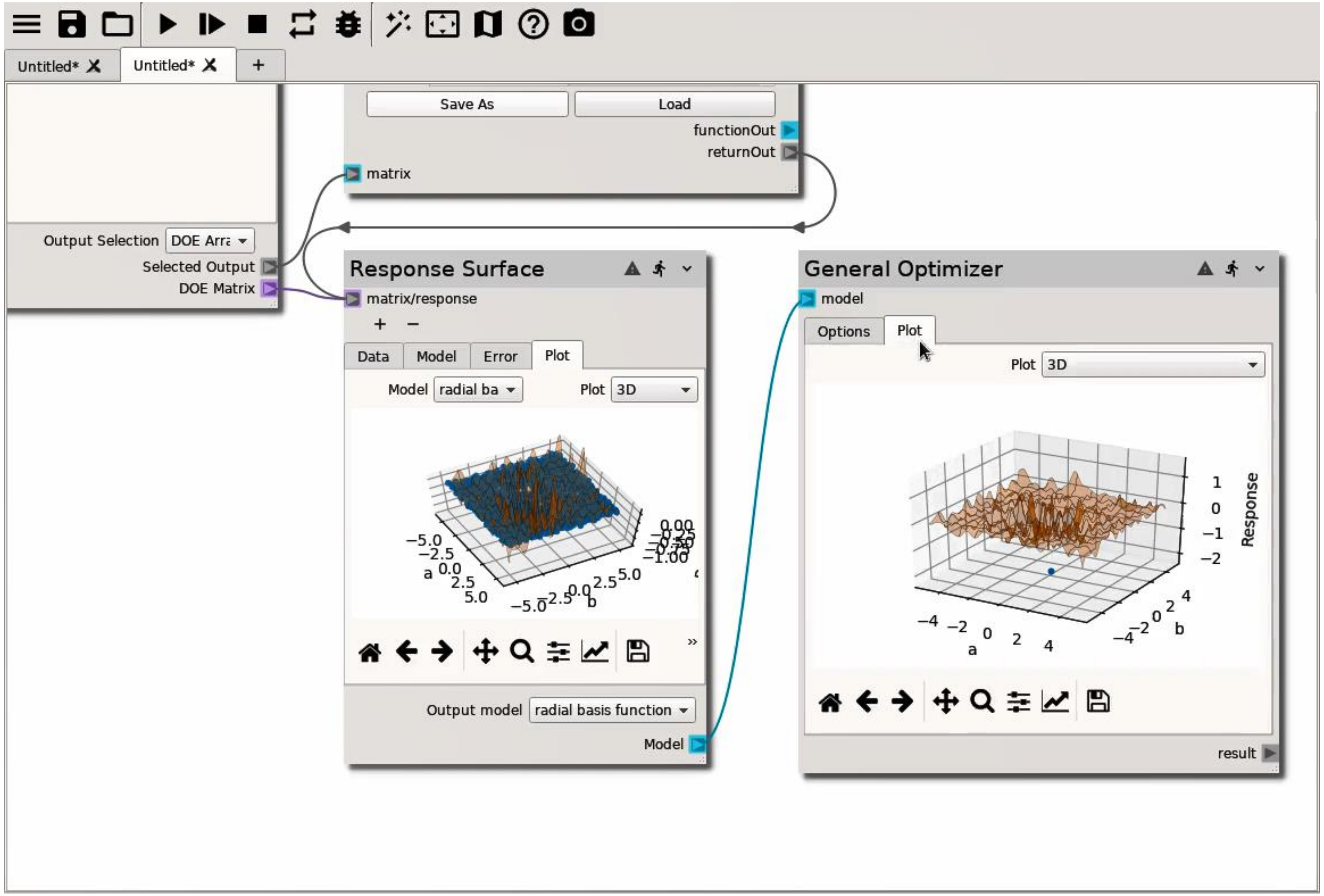
# Forward Propagation | P-Box



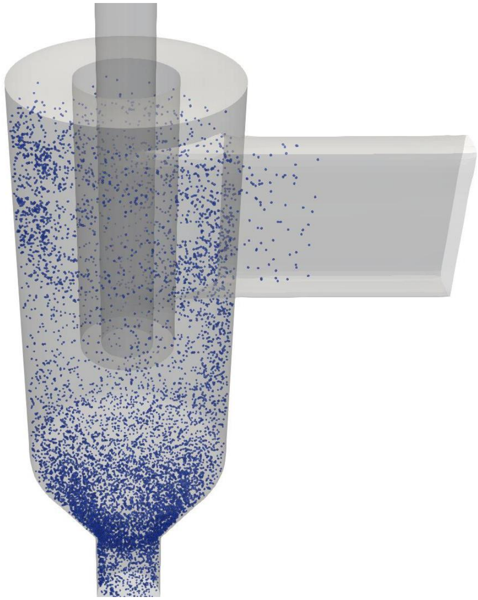



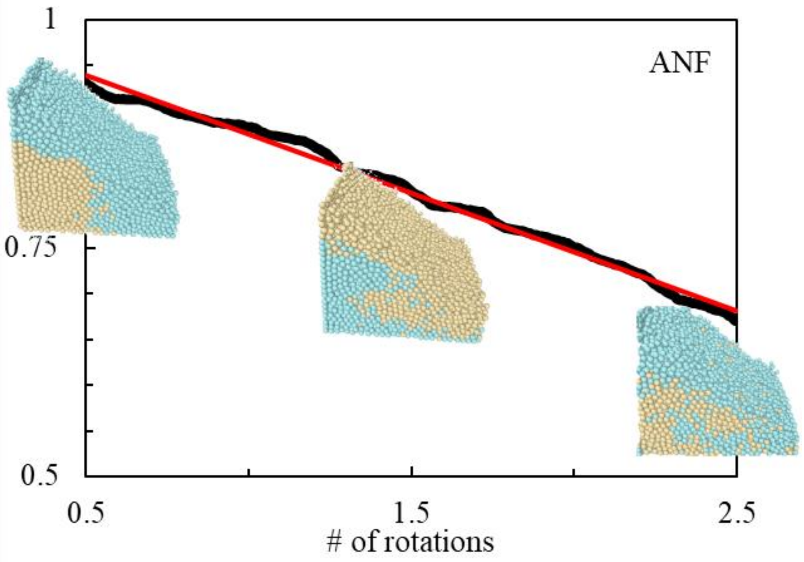
# Demo | Wizard

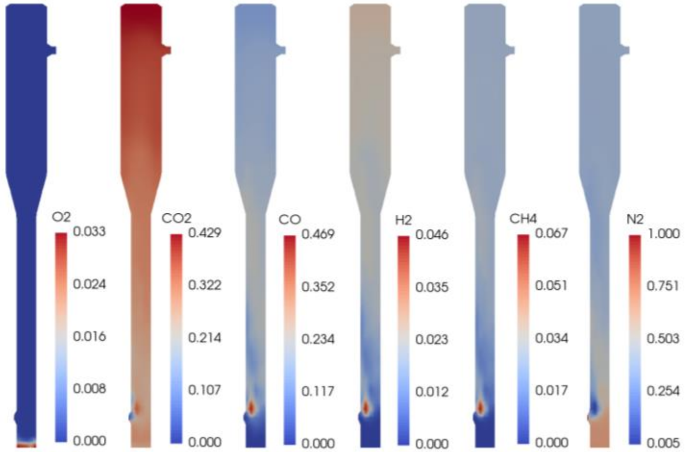
NATIONAL ENERGY TECHNOLOGY LABORATORY



# Examples




ANF
1
0.75
0.5
0.5
1.5
2.5
# of rotations

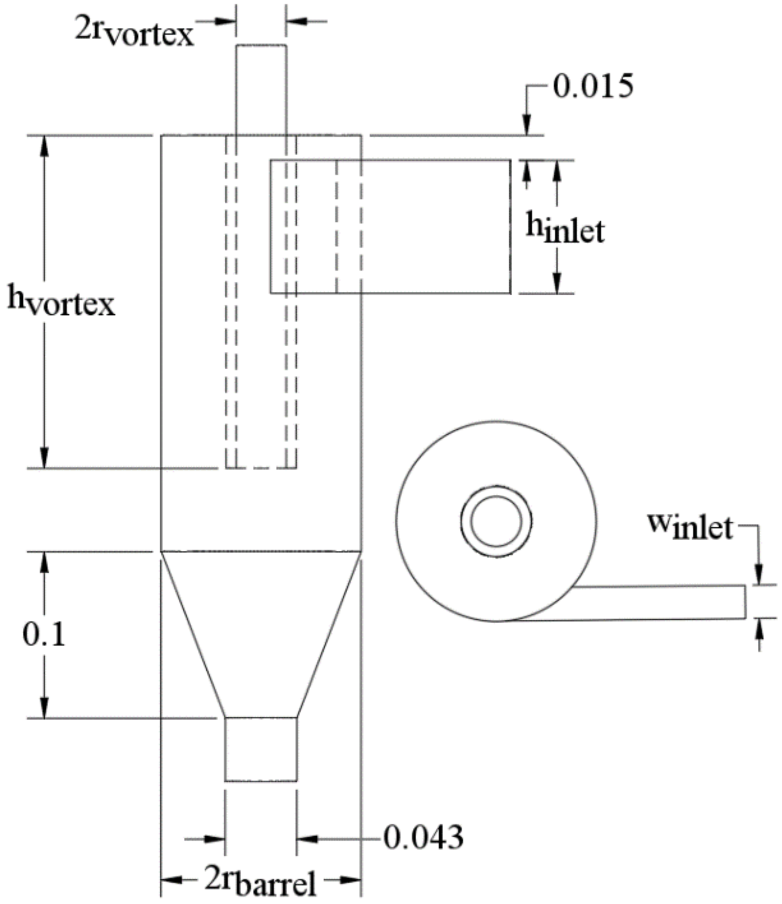

O2
0.033
0.024
0.016
0.008
0.000
CO2
0.429
0.322
0.214
0.107
0.000
CO
0.469
0.352
0.234
0.117
0.000
H2
0.046
0.035
0.023
0.012
0.000
CH4
0.067
0.051
0.034
0.017
0.000
N2
1.000
0.751
0.503
0.254
0.005

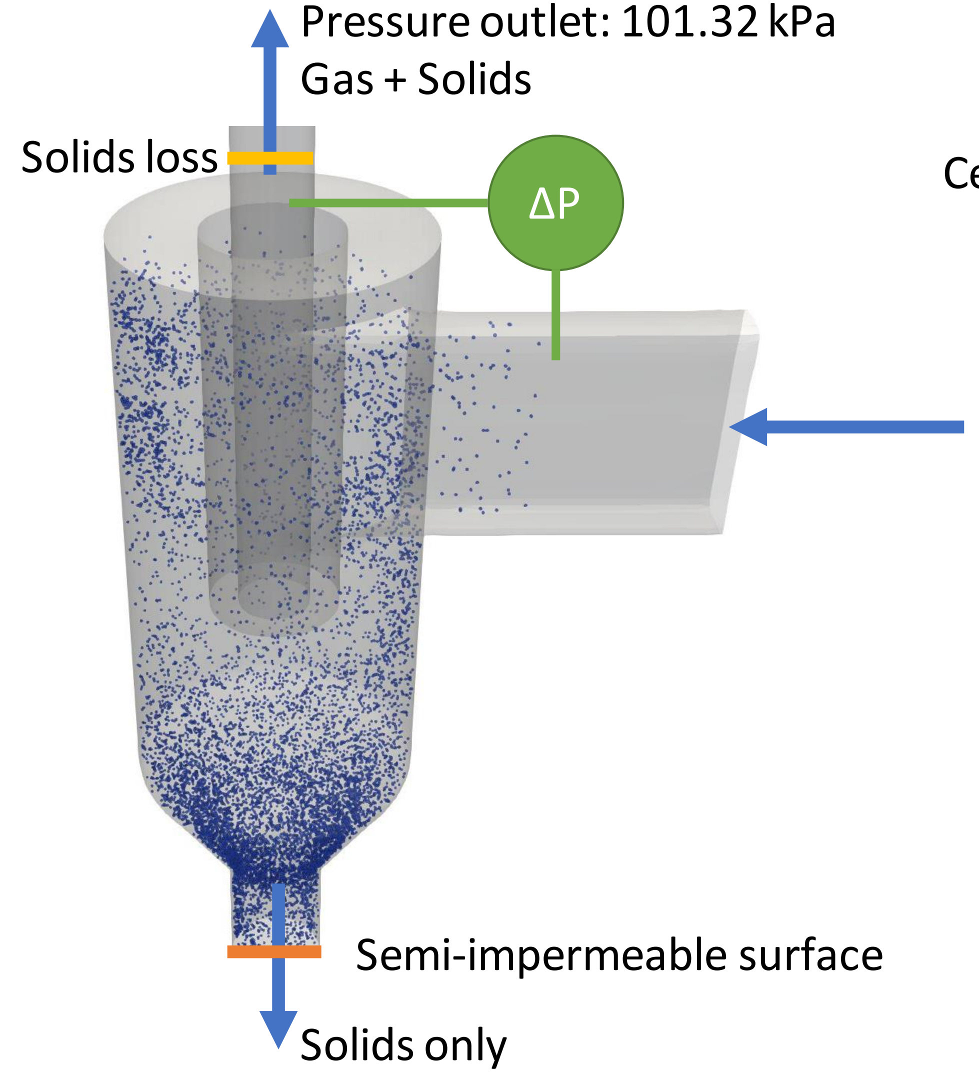
# Example | Cyclone Optimization

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 |

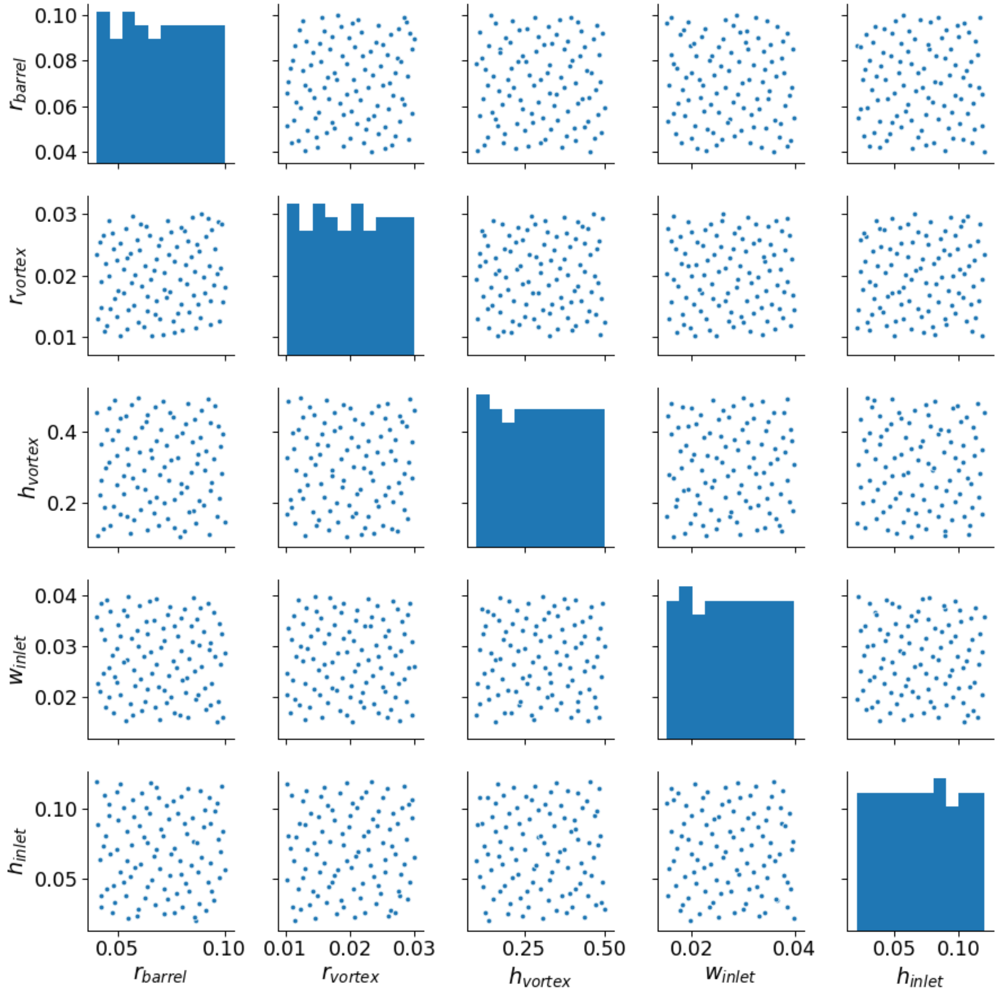
**HDPE**
Diameter: 871 μm
Density: 860 kg/m$^3$

~1 particles/parcel

# Design of experiments

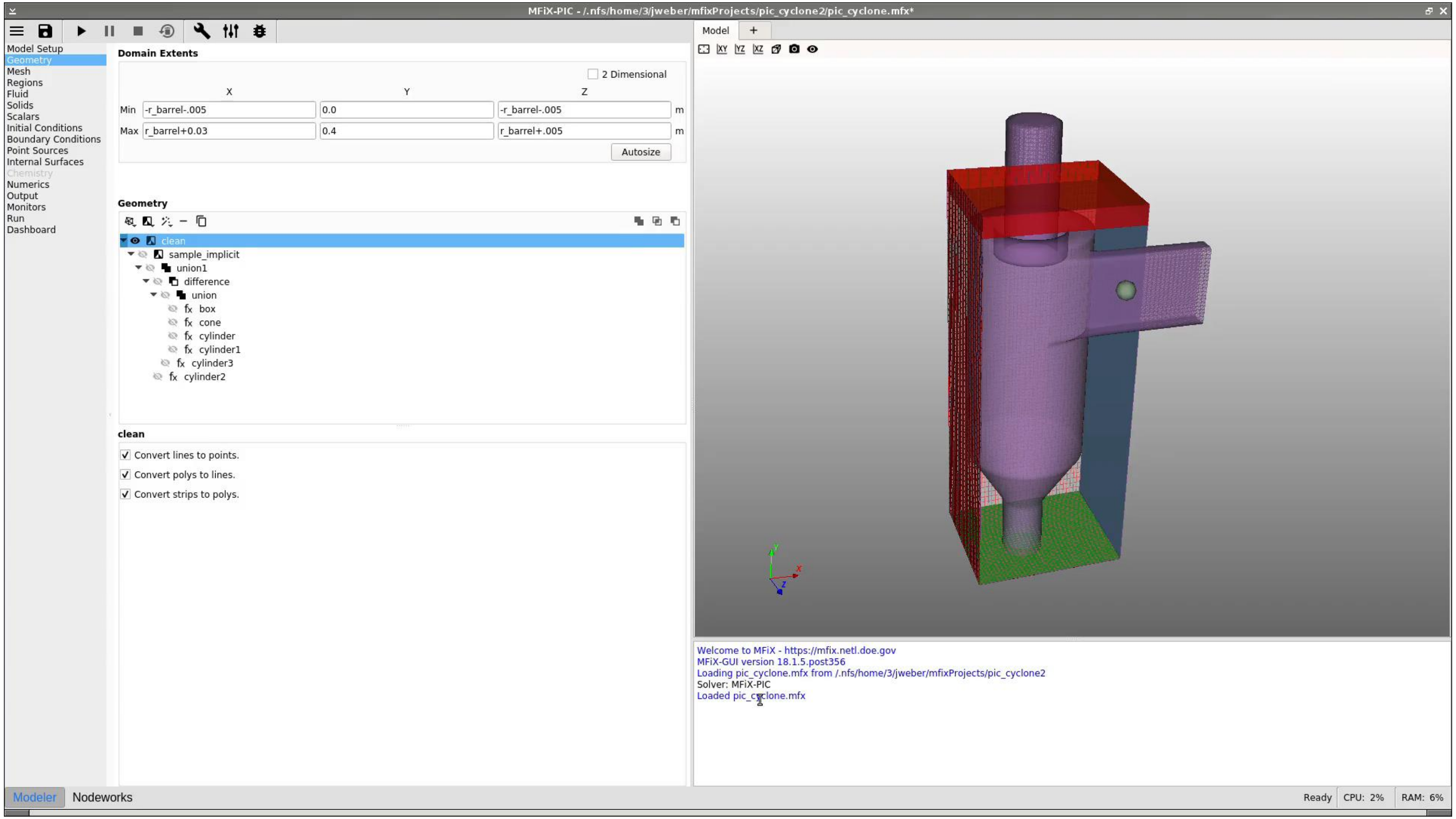
| Variable | min (m) | max (m) |
|---|---|---|
| $R_{BARREL}$ | 0.04 | 0.1 |
| $R_{vortex}$ | 0.01 | 0.03 |
| $H_{vortex}$ | 0.1 | 0.5 |
| $H_{inlet}$ | 0.02 | 0.12 |
| $W_{inlet}$ | 0.015 | 0.04 |

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



# Model creation

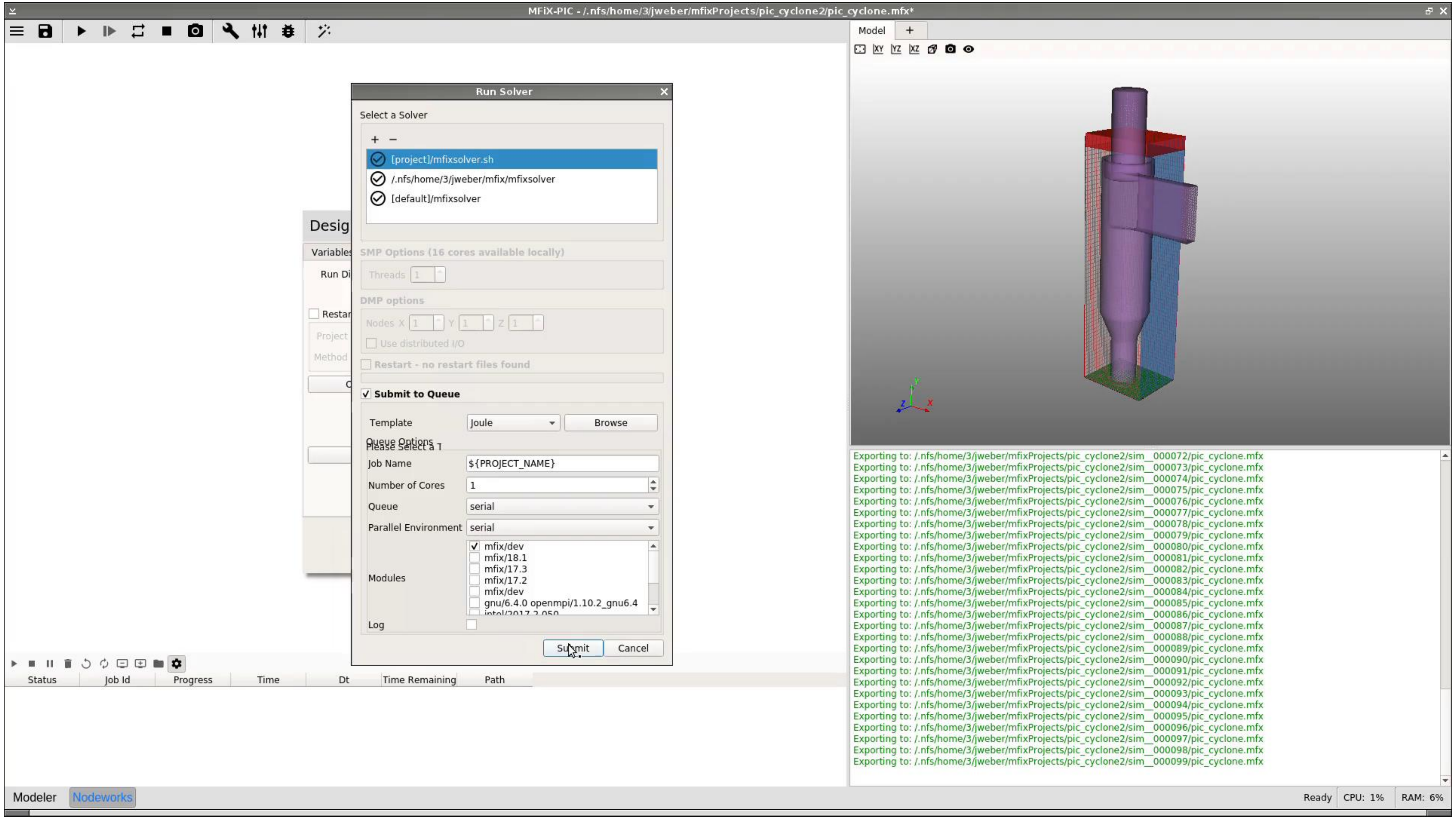
Models created using Nodeworks and MFiX



# Dispatch

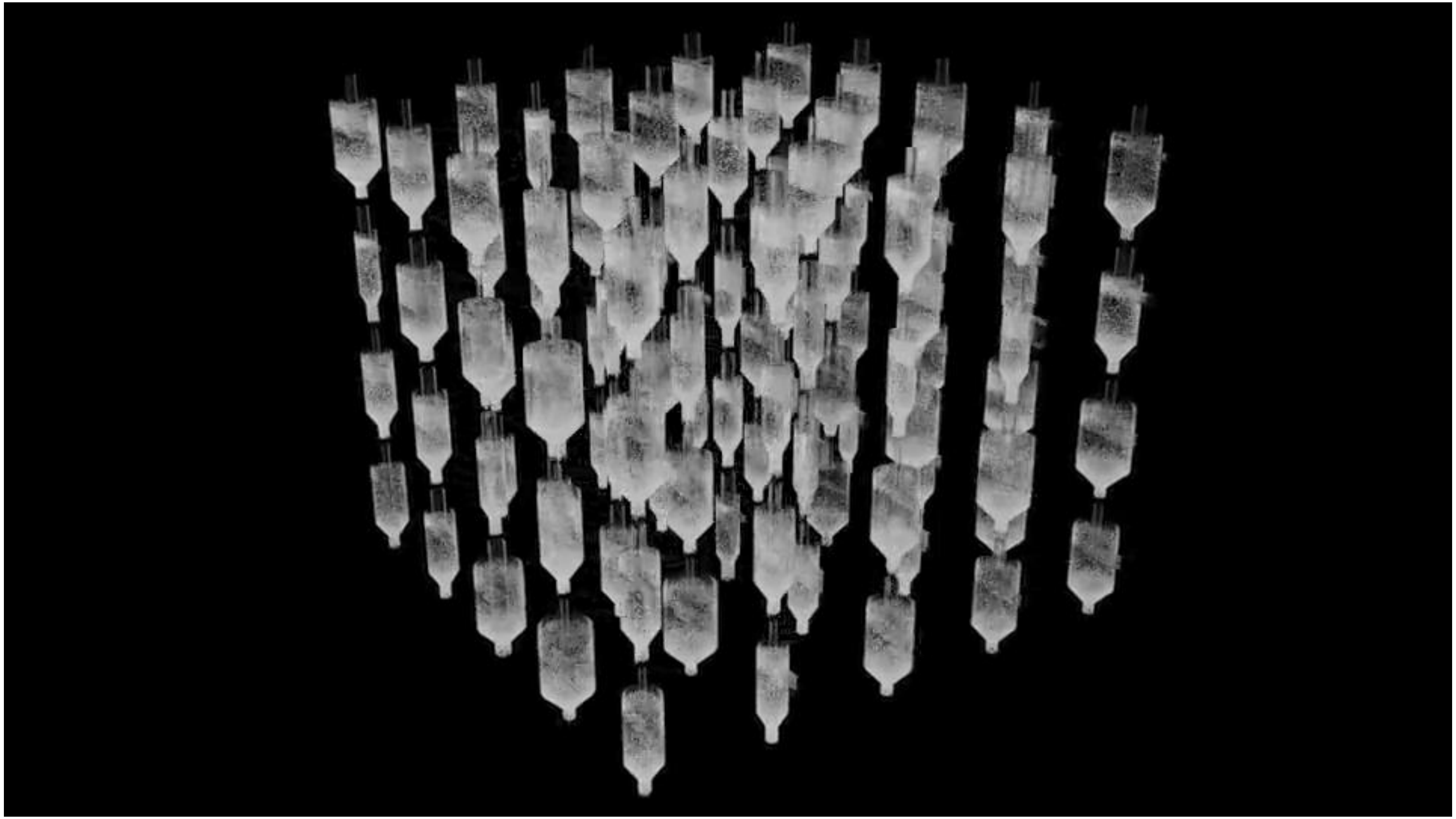
NATIONAL ENERGY TECHNOLOGY LABORATORY

Using Nodeworks and MFiX, Dispatch all models to the queue

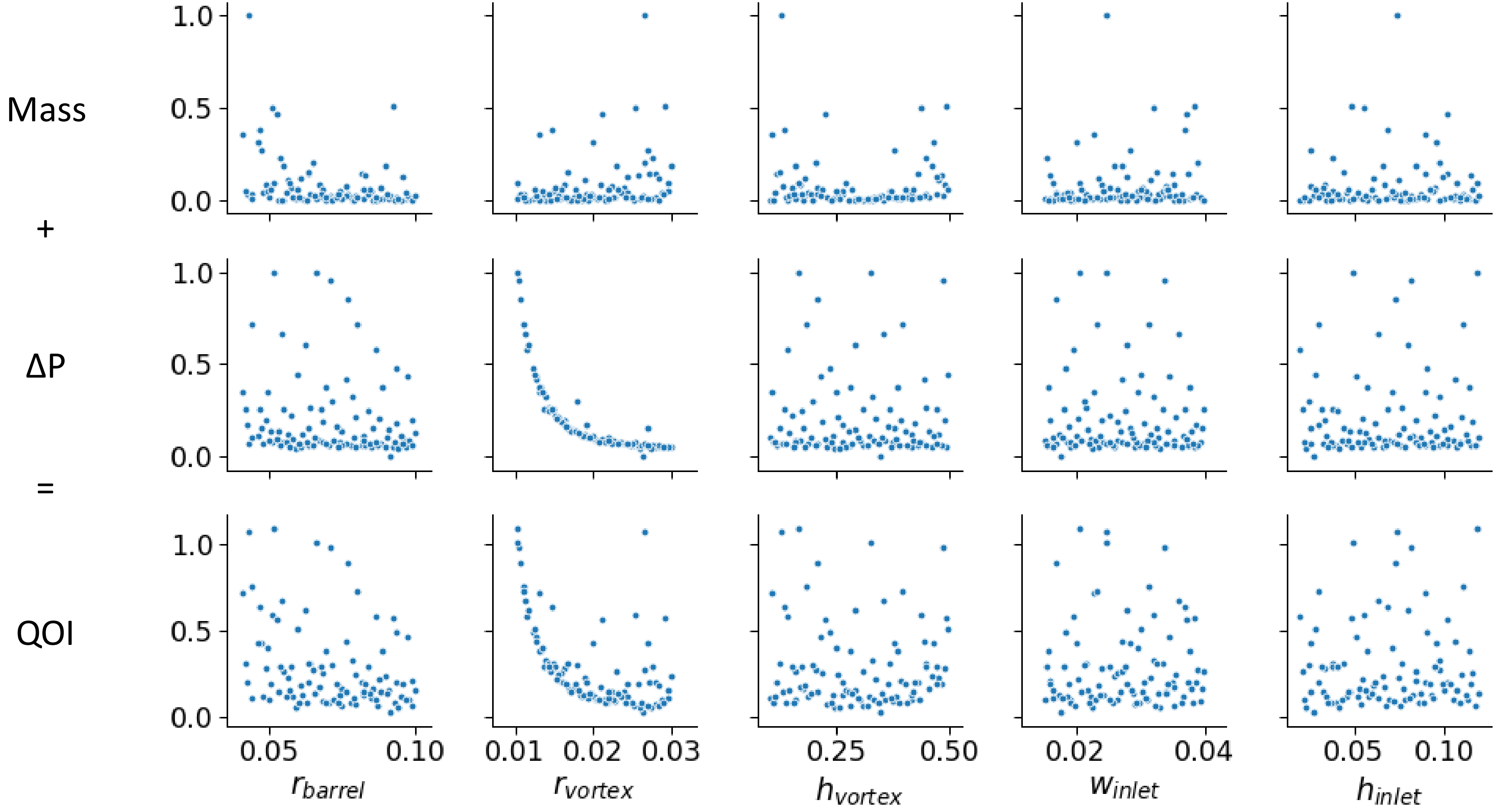


# Run the models!

- 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



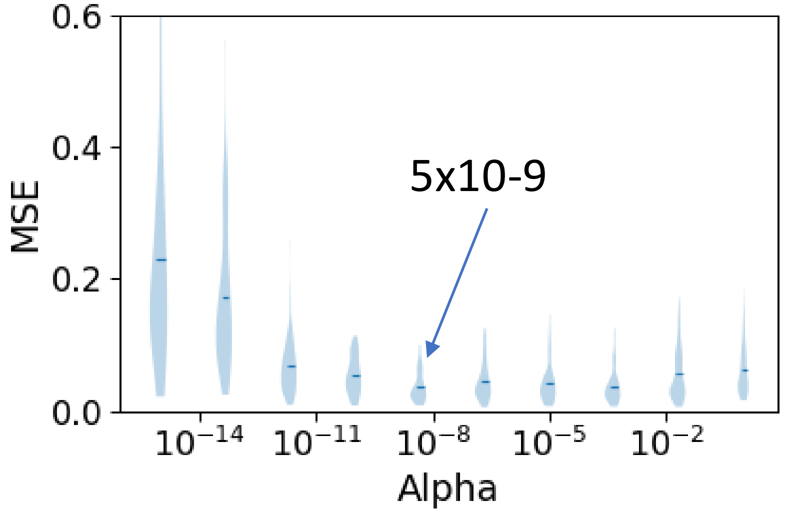
# Quantity of interest

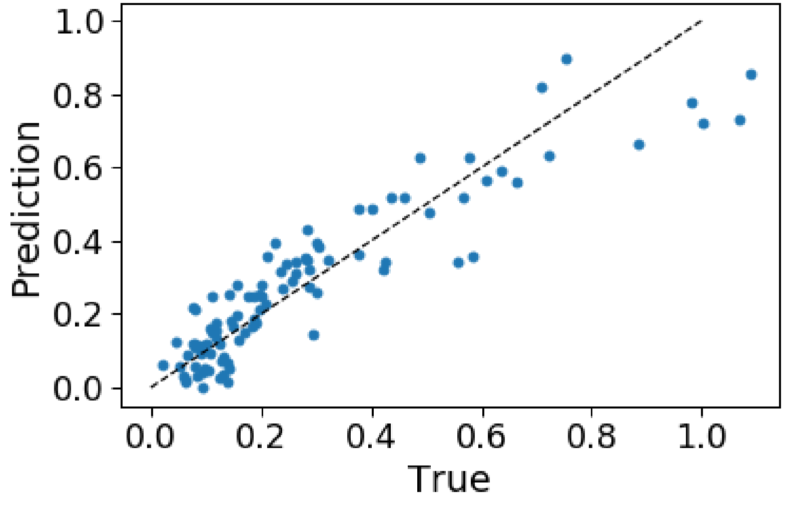


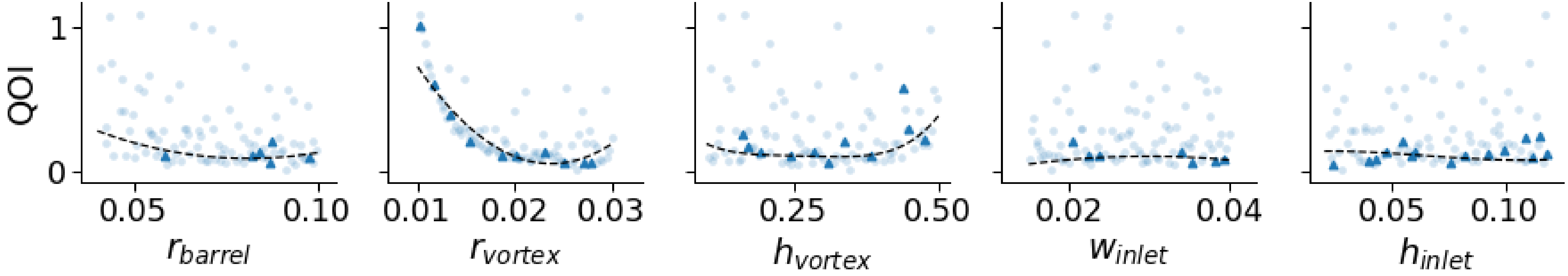
# Surrogate model: Gaussian Process

Alpha: noise level or smoothing of the data

10% hold out CV

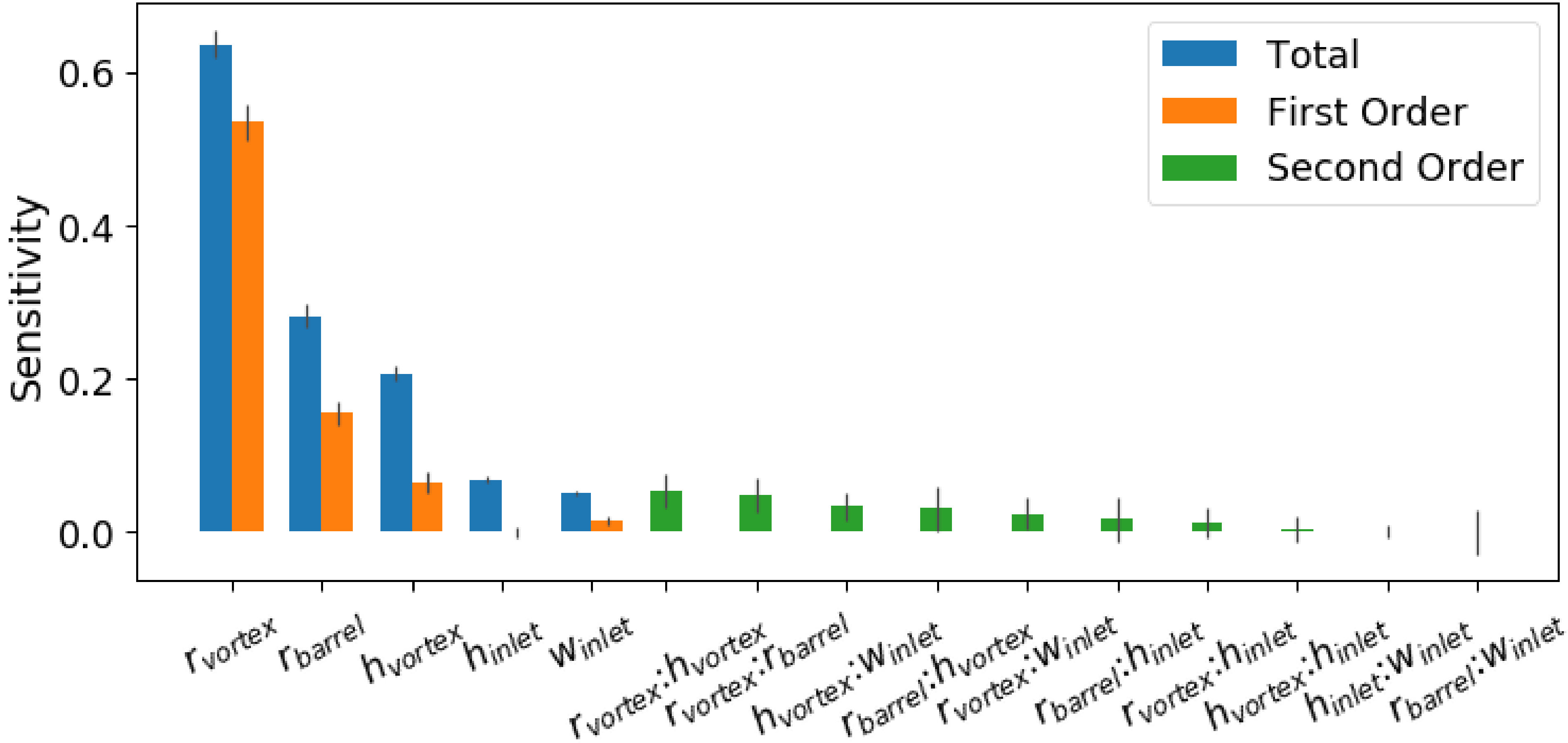


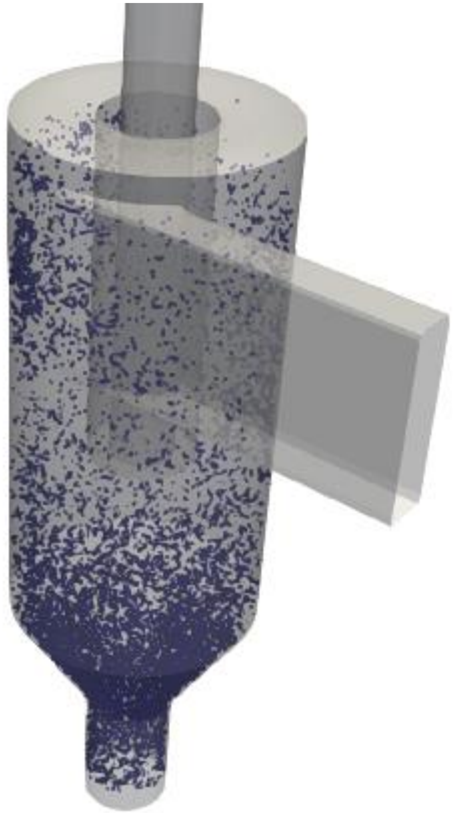
# Sobol Sensitivity

Using SALib



# Optimization

Original



Optimal

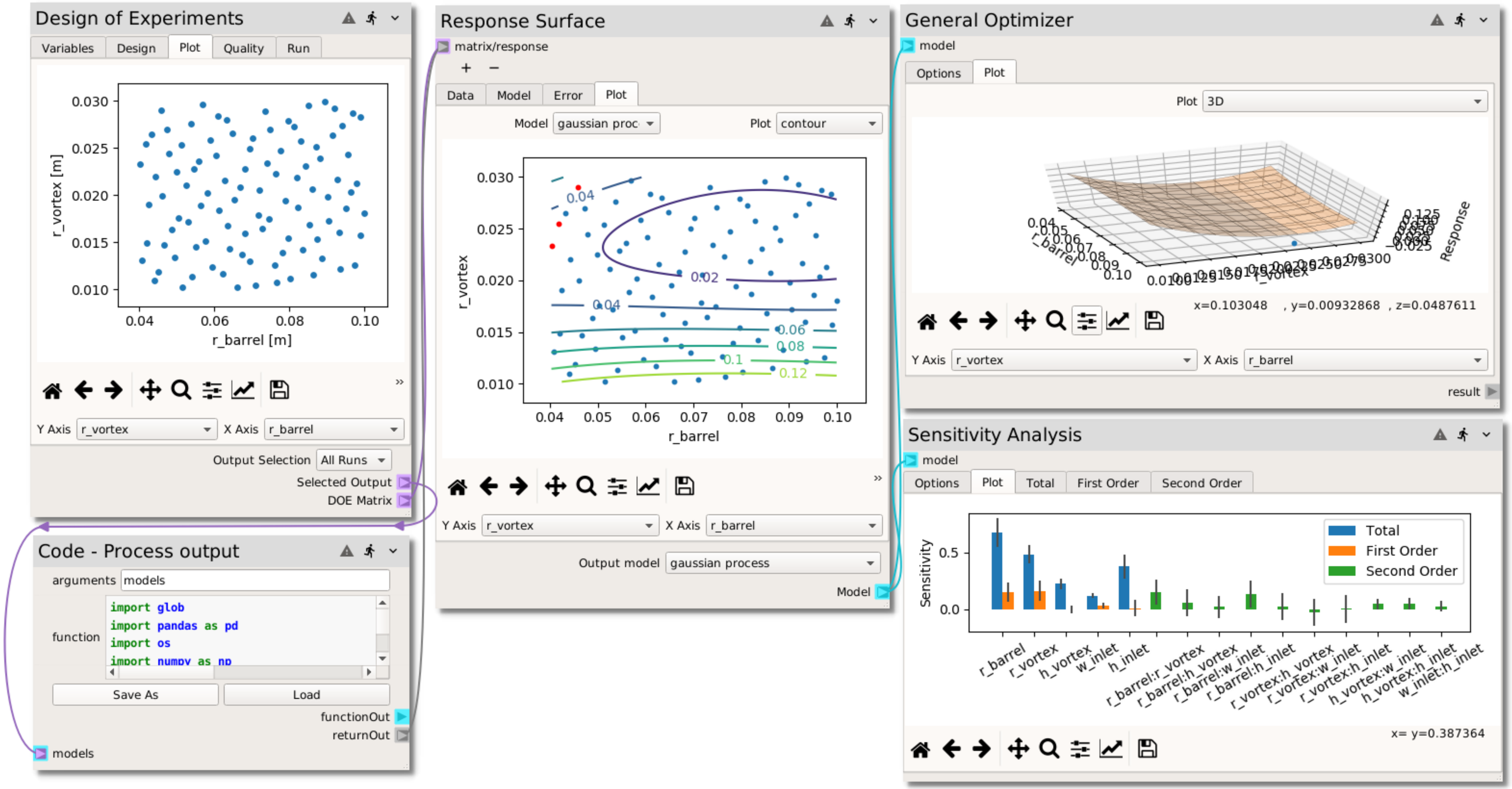
Using differential evolution

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

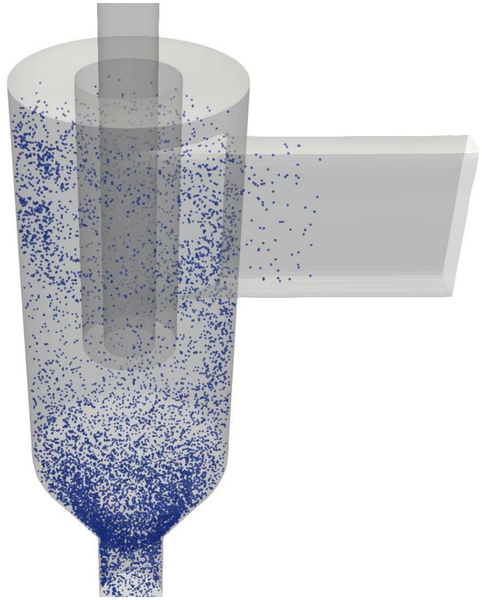
| Variable | Original (m) | Optimal (m) |
|---|---|---|
| $r_{barrel}$ | 0.06 | 0.096 |
| $r_{vortex}$ | 0.015 | 0.026 |
| $h_{vortex}$ | 0.4 | 0.373 |
| $h_{inlet}$ | 0.08 | 0.12 |
| $w_{inlet}$ | 0.02 | 0.015 |

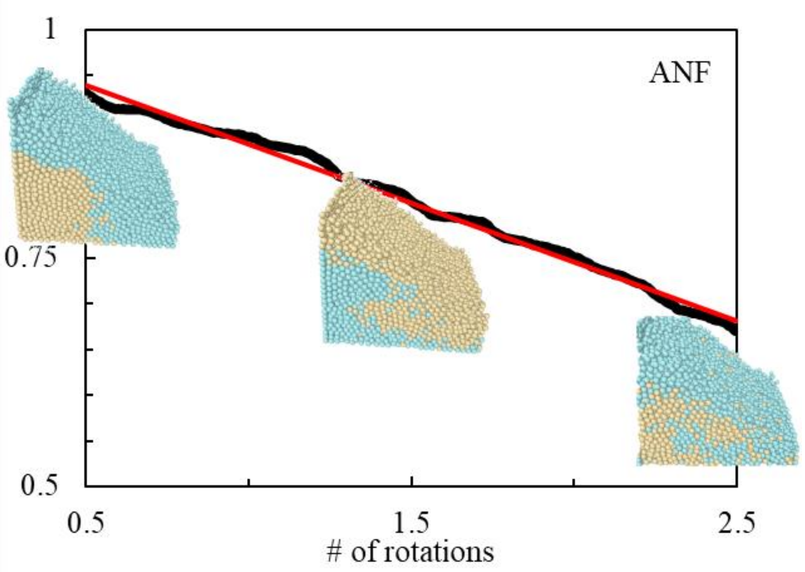
Edge of design space

# Putting it all together



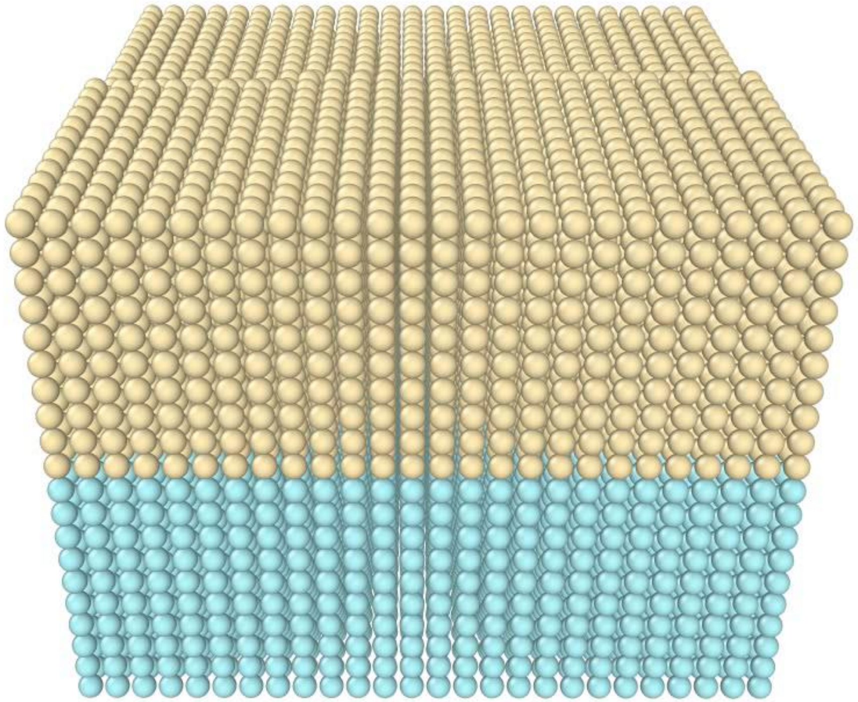
# Examples

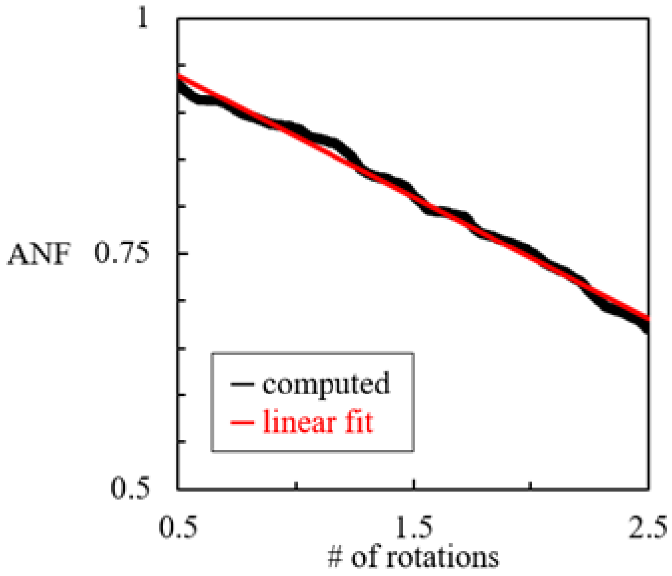




# Example | DEM Mixing UQ

Quantify mixing as the rate of decay of the **A**like **N**eighbor **F**raction (ANF)






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

# Problem definition

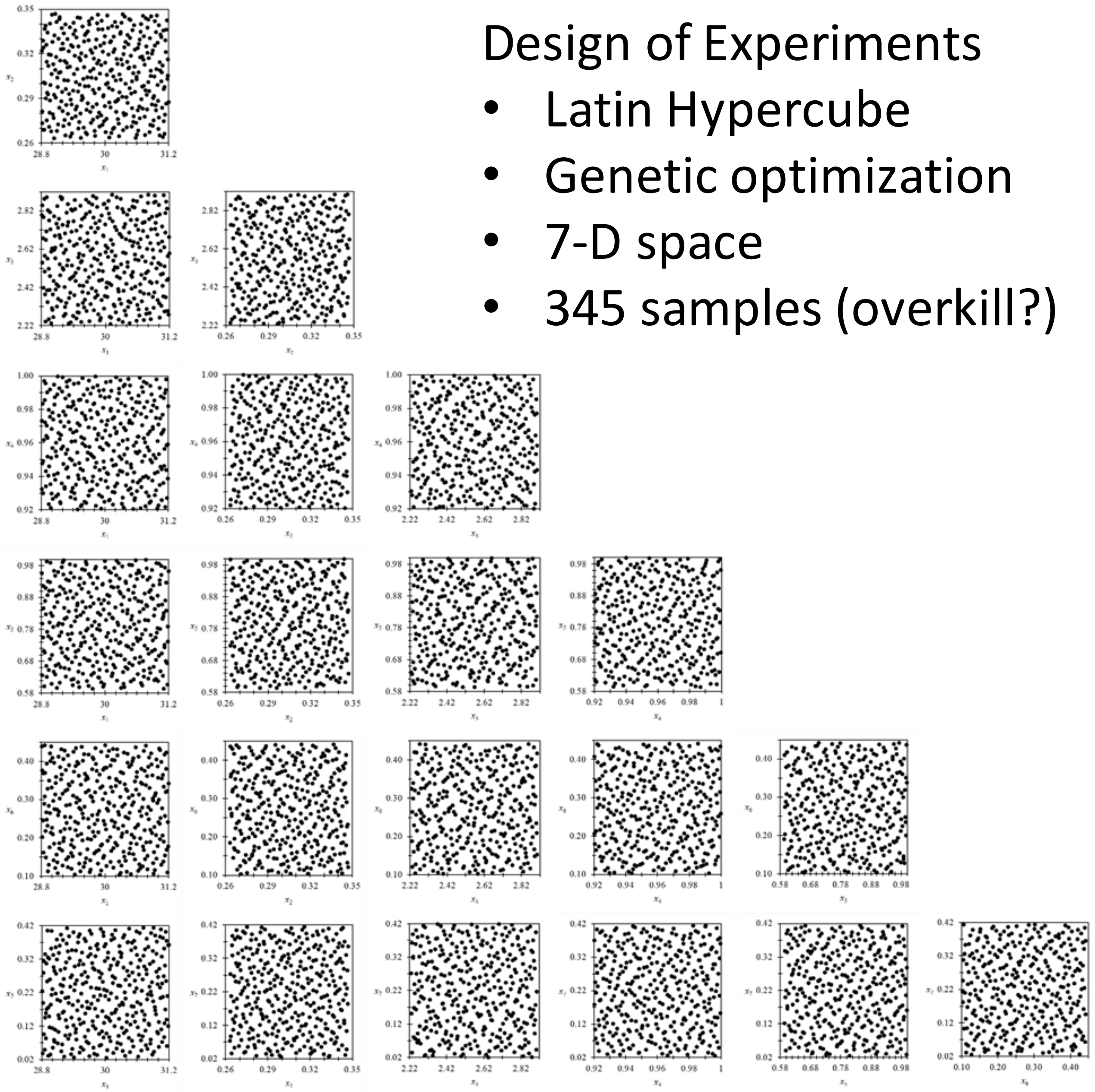
Model: MFiX-DEM
Rotation induced by angular gravity
Geometry considered fixed/known
Seven model parameters considered as unknown quantities

- 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 |
| $\rho_p$ | (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 |
| $\mu_{pp}$ | - | $x_6$ | 0.1 | 0.45 |
| $\mu_{pw}$ | - | $x_7$ | 0.02 | 0.42 |

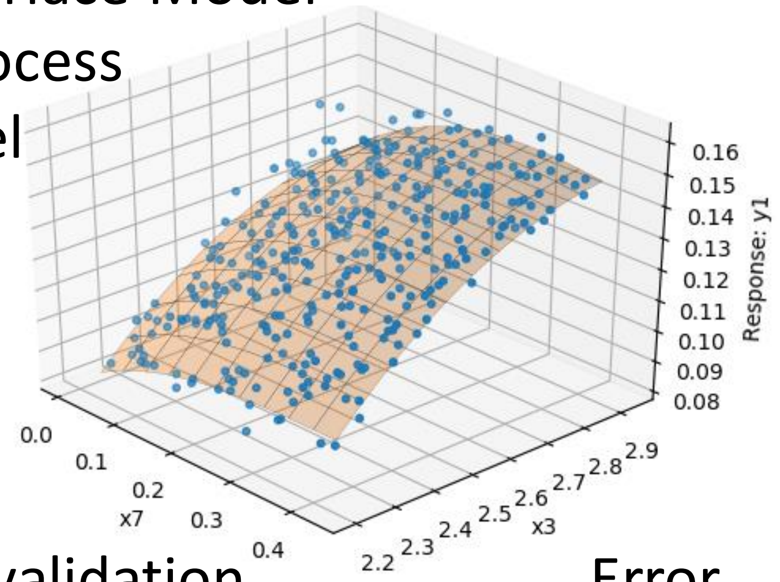
# DOE → Simulations → Surrogate



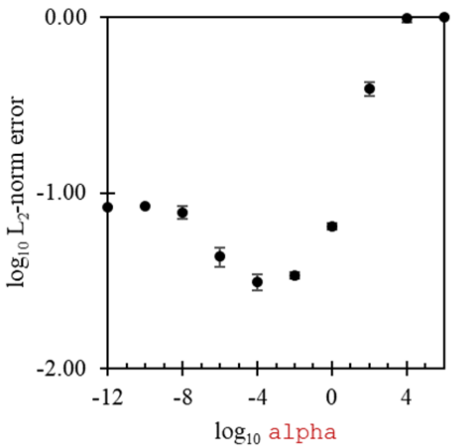
Design of Experiments
- Latin Hypercube
- Genetic optimization
- 7-D space
- 345 samples (overkill?)

Response Surface Model

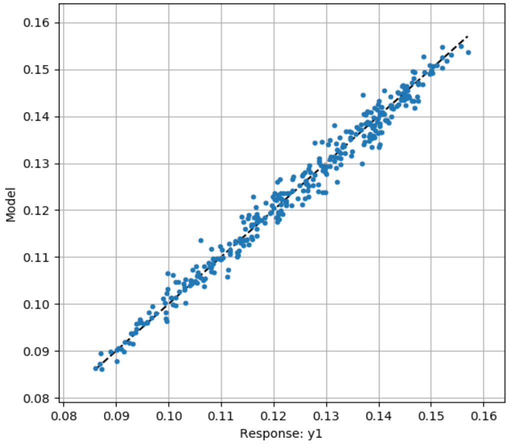
Gaussian Process
- RBF kernel




Cross-validation

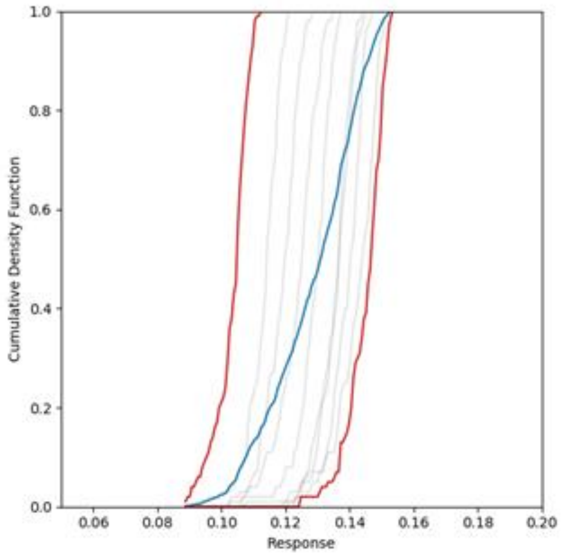


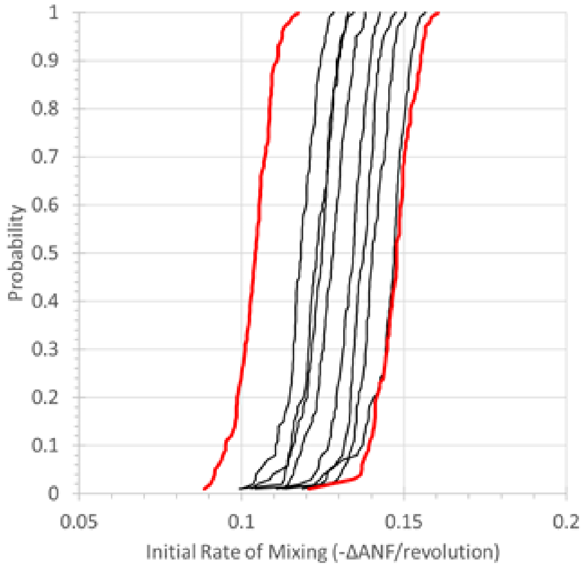

Error

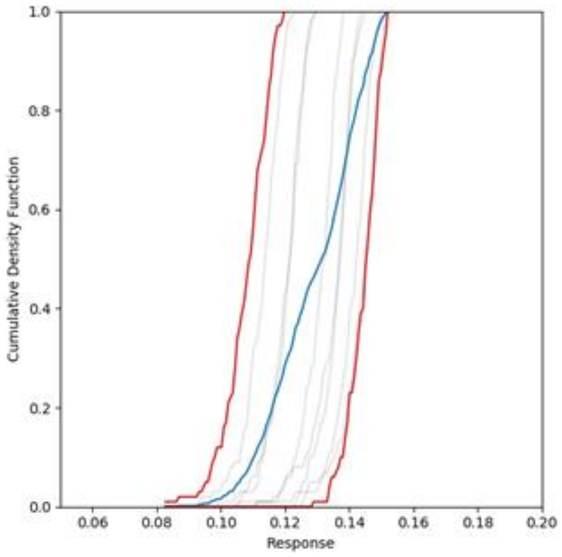


# Forward Propagation (of input uncertainties)

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



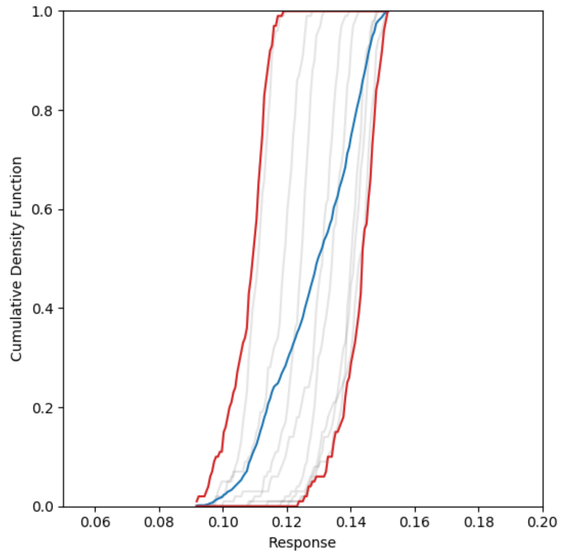




Original direct sample p-box of Dahl et al (2019)
Examples of surroage model propagated p-boxes

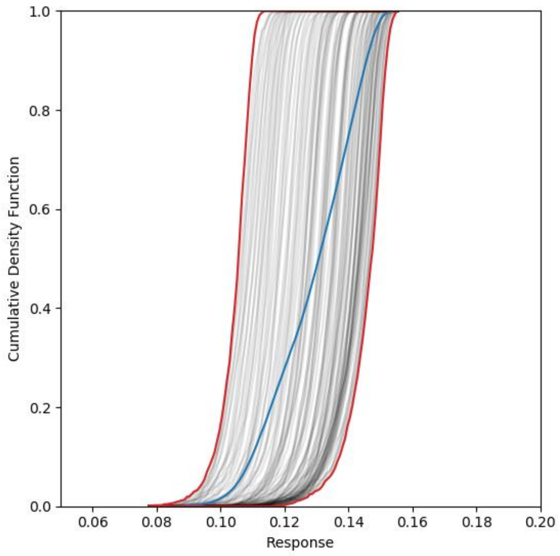
# Forward Propagation (of input uncertainties)

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

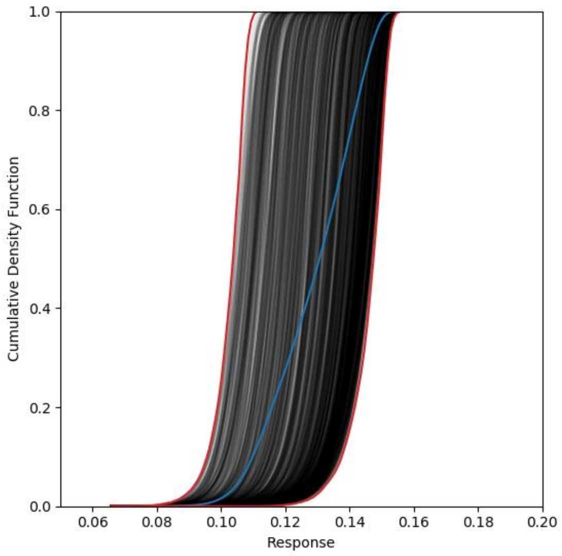
- Direct/full model: expensive
- Surrogate model: (once constructed) cheap




10 × 100

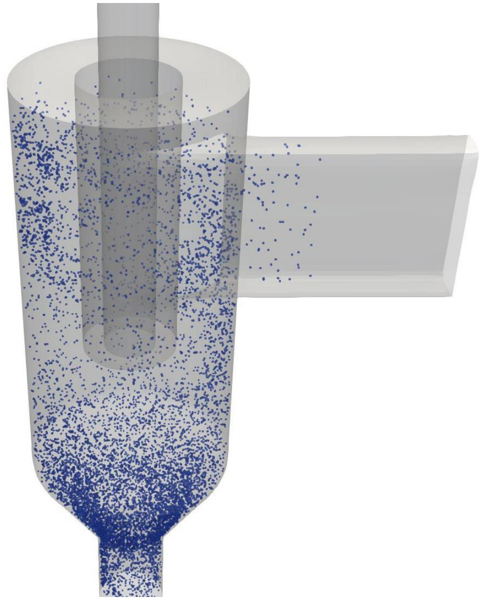


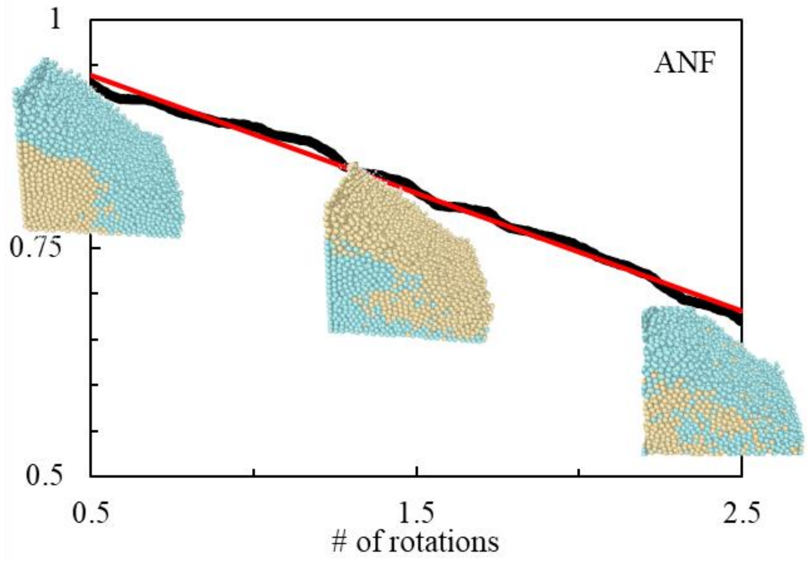

100 × 1000



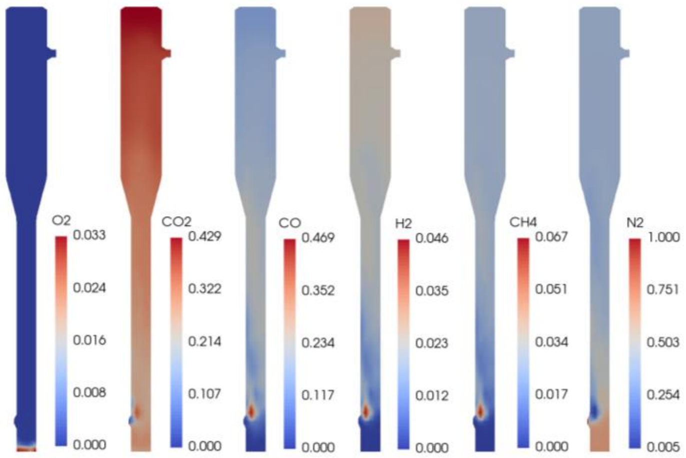

1000 × 10000

# Examples



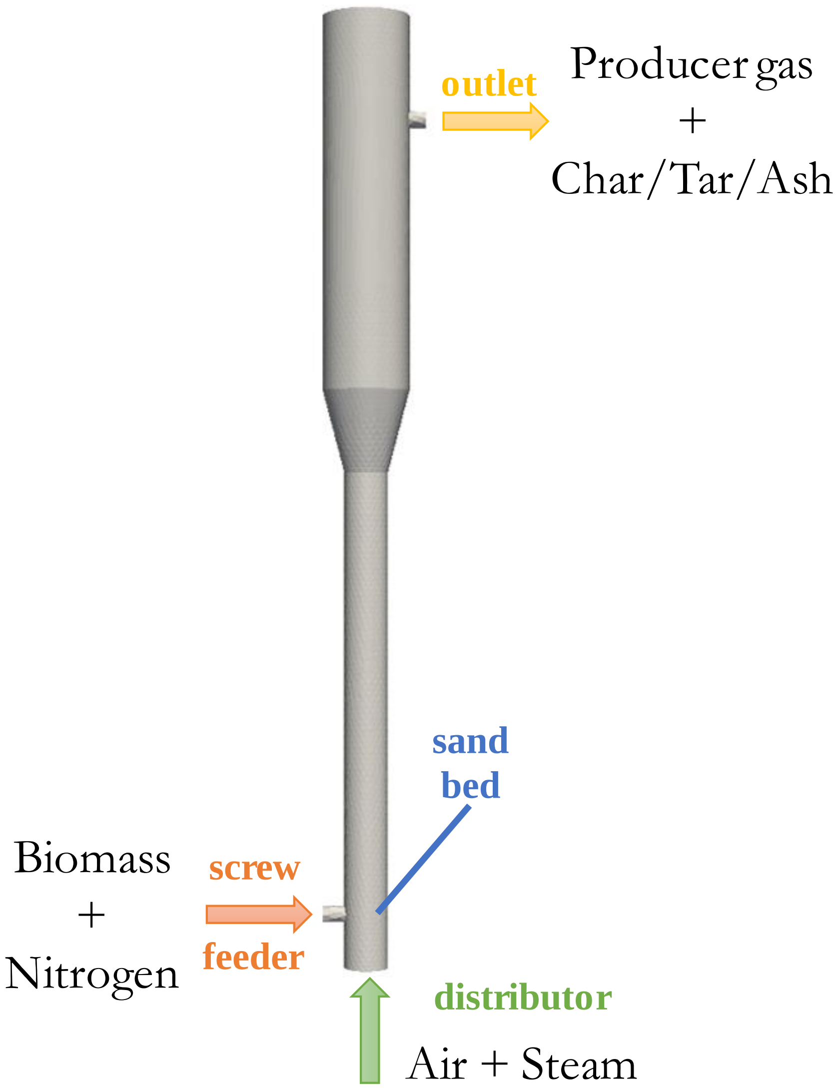


1
0.75
0.5
ANF
0.5
1.5
2.5
# of rotations

O2
0.033
0.024
0.016
0.008
0.000
CO2
0.429
0.322
0.214
0.107
0.000
CO
0.469
0.352
0.234
0.117
0.000
H2
0.046
0.035
0.023
0.012
0.000
CH4
0.067
0.051
0.034
0.017
0.000
N2
1.000
0.751
0.503
0.254
0.005

# 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)

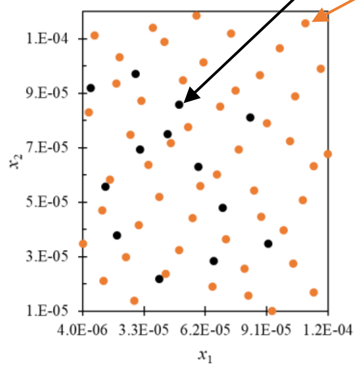
Objective Function:

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

# DOEs and Results

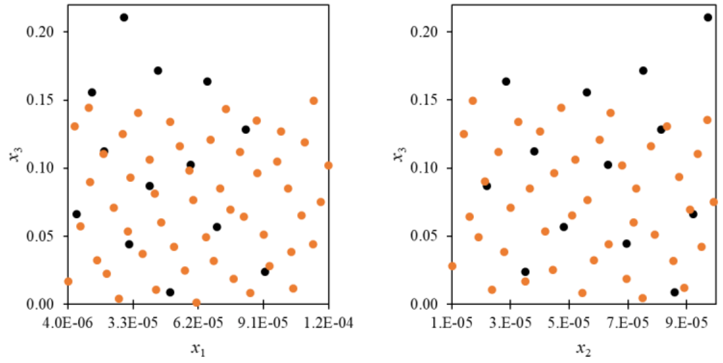
Sub-set of preliminary, scoping DOE

Secondary, refined DOE

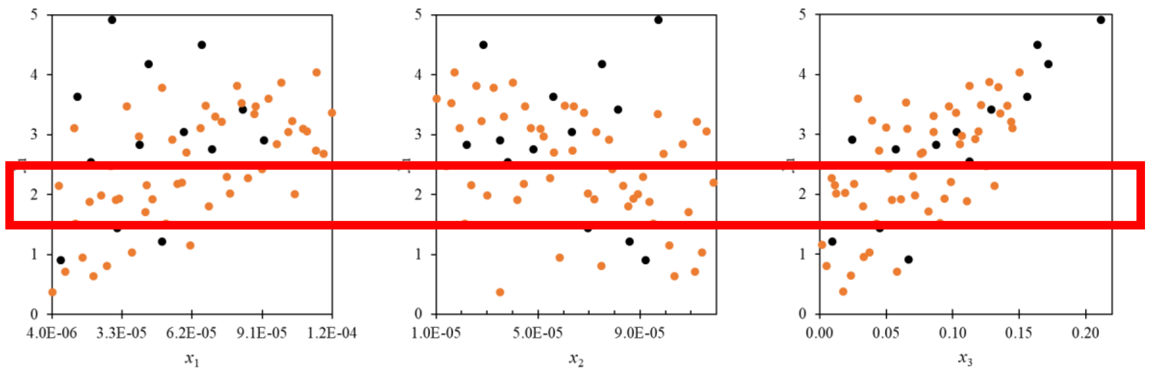



## Nodeworks

- Latin Hypercube
- Genetic Optimization
- Composite DOE not LH




## Results for the QoI, $H_2/CO$



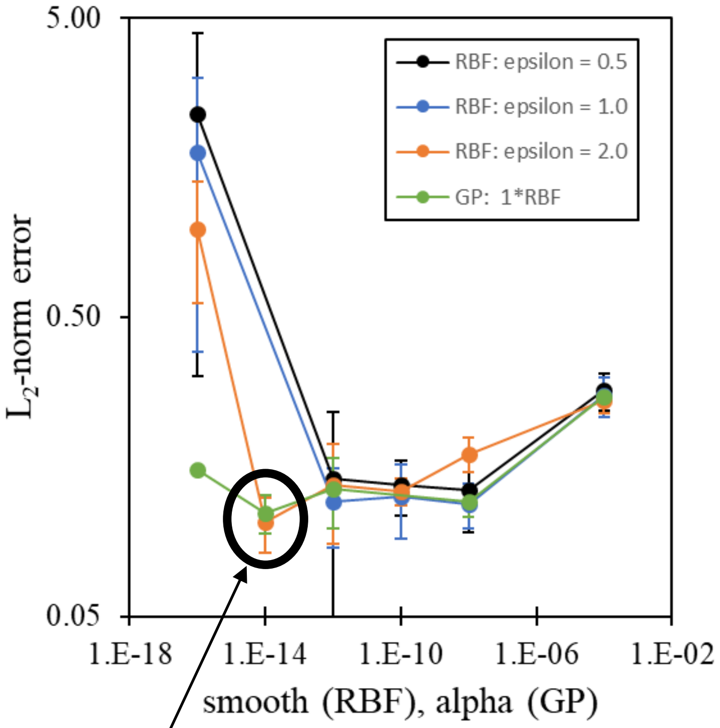

## Region of interest

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

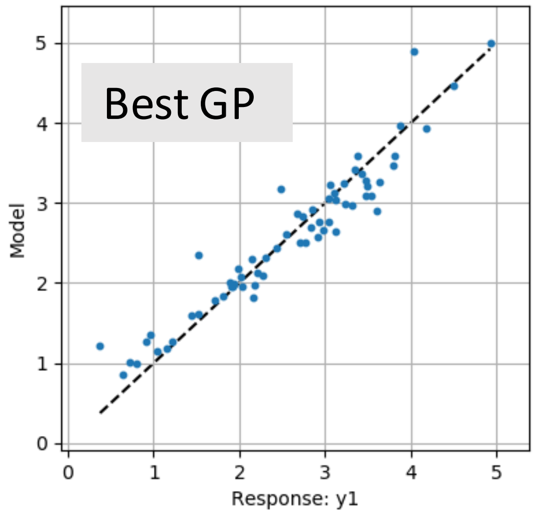
A: Construct a (4-D) response surface surrogate model and extract the (3-D) iso-surface characterizing $y_1$=2

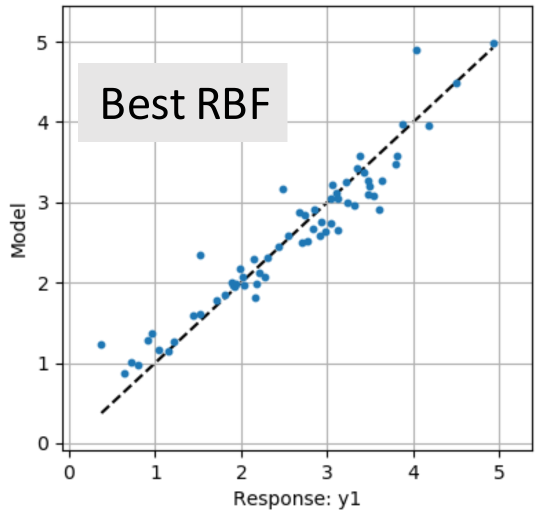
# Surrogate Modeling

Cross-Validation for the QoI, $H_2/CO$




Full Model Error



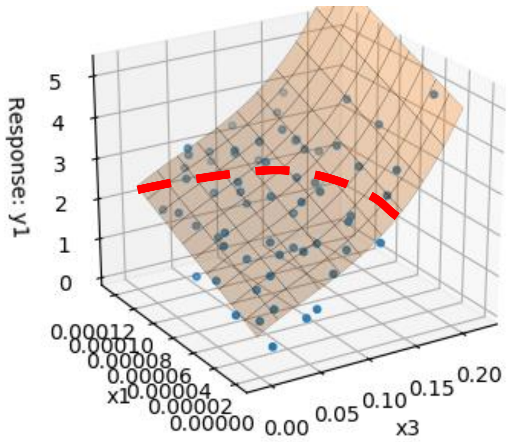

**Selection: GP (more consistent)**

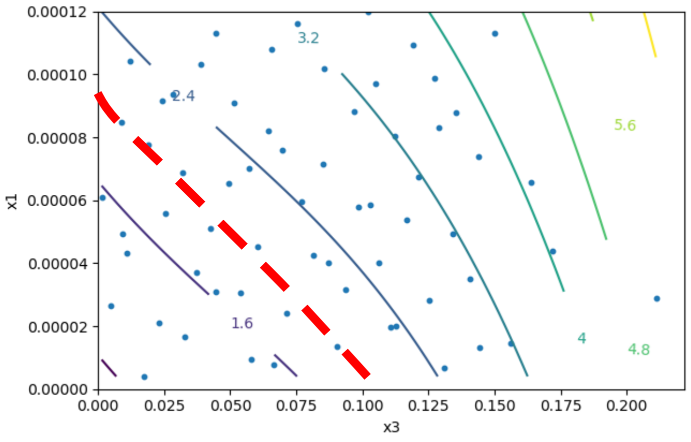
Best surrogate models:

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

# Optimization

Optimize: RSM == 2



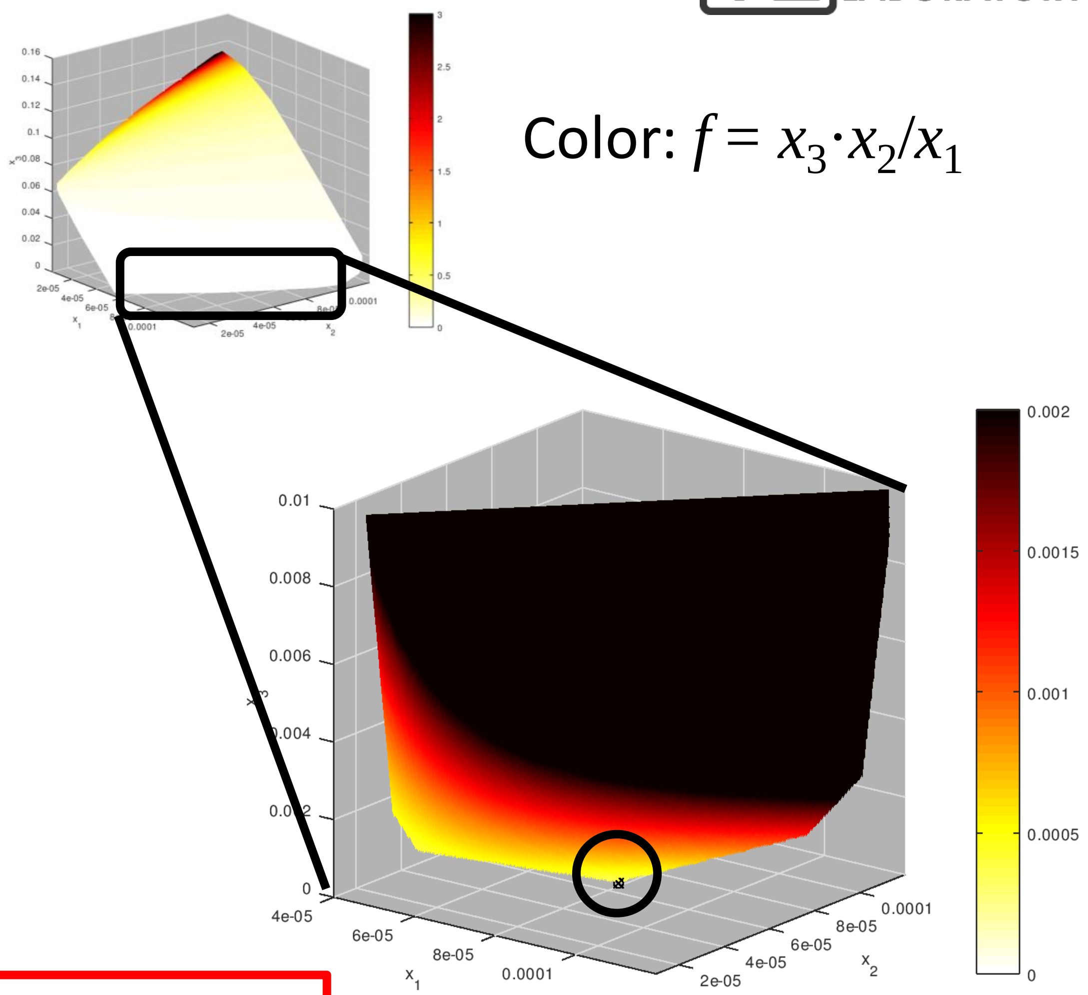

Iso-surface

| $x_1$ | $x_2$ | $x_3$ | $y_1$ |
|---|---|---|---|
| # | # | # | 2.000.. |
| # | # | # | 1.999.. |
| # | # | # | 1.999.. |
| # | # | # | 2.000.. |

…

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

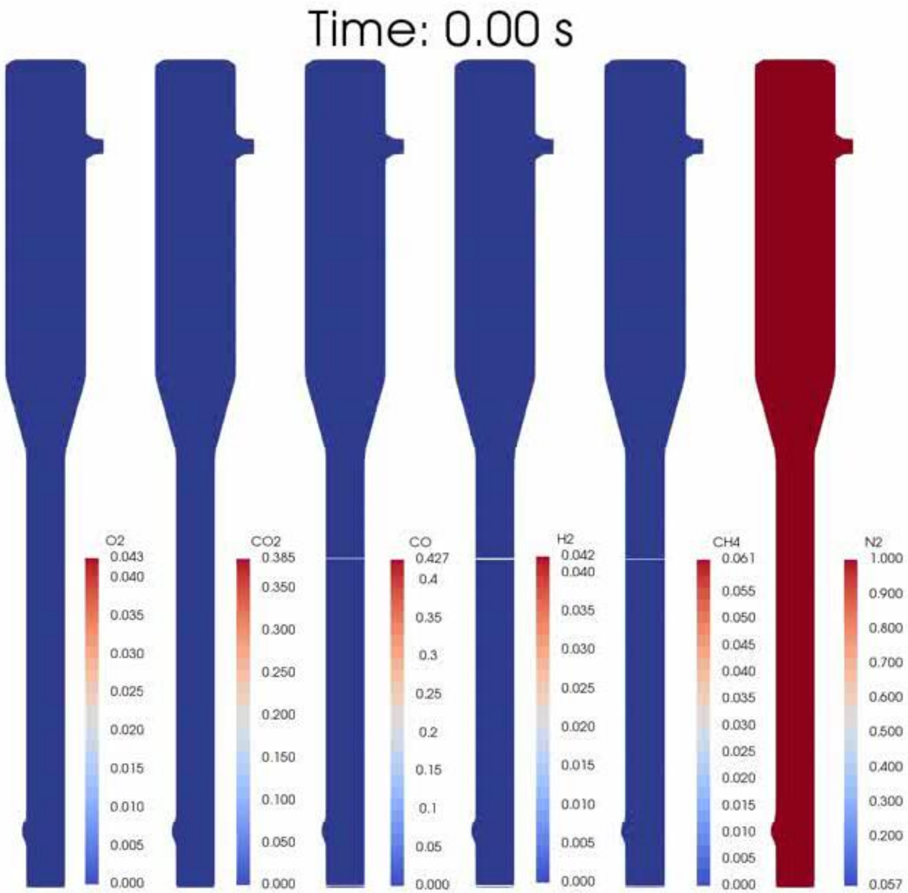
Color: $f = x_3 \cdot x_2 / x_1$

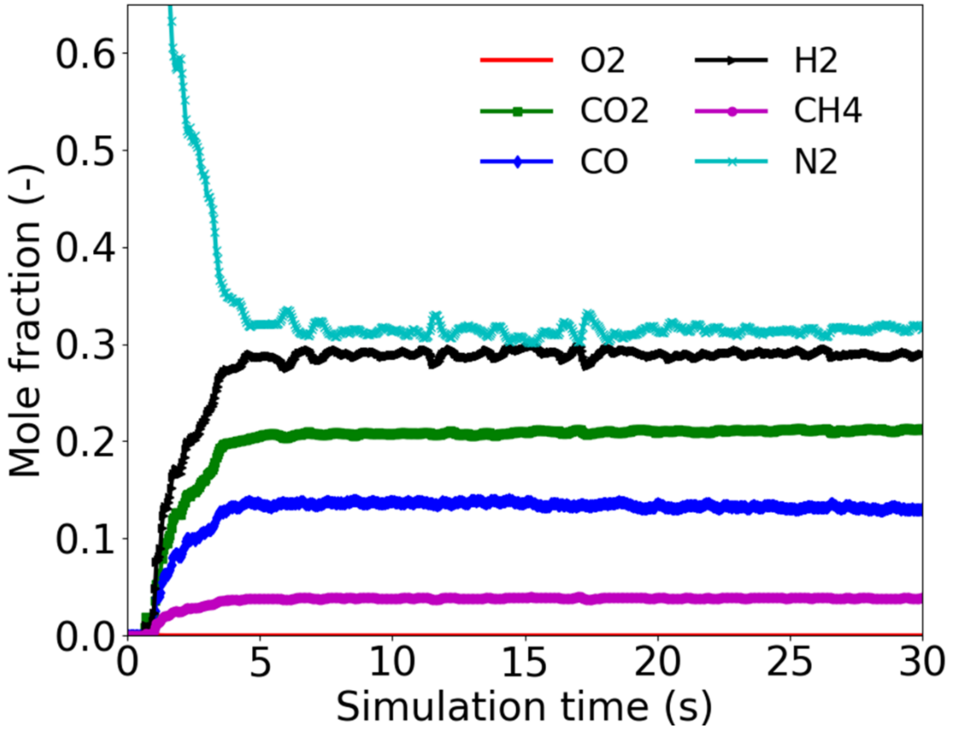



Optimal condition

# Validation of (surrogate) Optimum






$x_1$ = 0.086 (g/s), $x_2$ = 0.054 (g/s), $x_3 = 4.8\times10^{-4}$, $\hat{y}_1$ = 2, $y_1$ = 2.2

(within expected error from cross-validation test)