

# OPTIMIZATION OF A CYCLONE USING MULTIPHASE FLOW COMPUTATIONAL FLUID DYNAMICS

**Justin Weber, first author<sup>1</sup>**

US Department of Energy, National Energy Technology Laboratory

3610 Collins Ferry Road, Morgantown, WV, 26505

[justin.weber@netl.doe.gov](mailto:justin.weber@netl.doe.gov)

**William Fullmer, second author**

National Energy Technology Laboratory, Leidos Research Support Team,

3610 Collins Ferry Road, Morgantown, WV, 26505

[William.fullmer@netl.doe.gov](mailto:William.fullmer@netl.doe.gov)

**Aytekin Gel, third author**

National Energy Technology Laboratory, ALPEMI Consulting, LLC

8205 S Priest Dr #13951

[aike@alpemi.com](mailto:aike@alpemi.com)

**Jordan Musser, fourth author**

US Department of Energy, National Energy Technology Laboratory

3610 Collins Ferry Road, Morgantown, WV, 26505

[jordan.musser@netl.doe.gov](mailto:jordan.musser@netl.doe.gov)

## ABSTRACT

*The US Department of Energy (DOE) National Energy Technology Laboratory's (NETL) 50 kW<sub>th</sub> chemical looping reactor was determined to have an underperforming cyclone, which was designed primarily using empirical correlations. To improve the performance of this cyclone using computational fluid dynamics (CFD) based modeling simulations, four critical design parameters including the vortex tube radius and length, barrel radius, and the inlet width and height were optimized. For this work, NETL's open source Multiphase Flow with Interphase eXchange (MFix) CFD code has been used to model a series of cyclones by*

---

<sup>1</sup> Corresponding author; [justin.weber@netl.doe.gov](mailto:justin.weber@netl.doe.gov), 304-285-5270

*systematically varying the geometric design parameters. To perform the optimization process, the surrogate modeling and analysis toolset inside Nodeworks was used. The basic methodology for the process is to use a design of experiments method (optimal Latin Hypercube) to generate samples that fill the design space. CFD models are then created, executed, and post-processed. A response surface (Gaussian process model) is created to characterize the relationship between input parameters and the Quantities of interest (QoI). Finally, the CFD-surrogate is used by an optimization method (differential evolution) to find the optimal design condition. The resulting optimal cyclone has a larger diameter and longer vortex tube, a larger diameter barrel, and a taller and narrower solids inlet. The improved design has a predicted pressure drop 11-times lower than the original design while reducing the mass loss by a factor of 2.3. Keywords: optimization, computational fluid dynamics, MFIX-PIC, Nodeworks, response surface methods, design of experiments, Latin Hypercube sampling.*

## **INTRODUCTION**

Experience of the design engineer and availability of adequate experimental testing capability used to play crucial role in successful engineering designs and shorter time-to-market deployments. However, in the last several decades, modeling and simulation (M&S) has become one of the key enablers of robust and rapid design of engineering products, especially with the tremendous growth and ubiquitous availability of high performance computing (HPC). These developments have enabled compute resource intensive analysis like uncertainty quantification and/or optimization through simulation models to become more mainstream. The ensembles of sampling simulations capturing complex physics for such analysis could now be performed routinely through simulation campaigns that involve hundreds of sampling simulations on a HPC to characterize system response under a diverse range of design parameter conditions to achieve optimization more quantitatively. This has opened the door to optimization of

complex and dynamic multiphase flow reactors and their components, including cyclones.

Cyclones are widely encountered in industrial processes involving gas & solid phases such as for separating dust particles from gas. One of the design objectives when selecting or constructing a cyclone is to maximize separation efficiency while simultaneously minimizing pressure drop, which in turn enables the maximization of the overall process efficiency. Cyclones are often designed based on experience and empirical correlations. If the correlation is based on experimental data that is reasonably close to the targeted application process parameters, then the correlation generally performs well. However, if the range of targeted application process parameters differ from the range of experimental process parameters, the likelihood of a poorly designed cyclone increases significantly.

One such underperforming cyclone has been identified as part of the US Department of Energy (DOE) National Energy Technology Laboratory's (NETL) 50 kW<sub>th</sub> chemical looping reactor (CLR) [1]. The CLR circulates approximately 0.08 kg/s of oxygen carrier between two reactors in a process that facilitates the reaction of fossil fuels while simultaneously capturing carbon dioxide. In such processes, it is imperative that the primary cyclone be as efficient as possible while keeping the pressure drop to a minimum, to achieve high oxygen carrier recycle rates and maintain a delicate pressure balance. High oxygen carrier recycle rates are critical to keep the oxygen carrier in the system to minimize the amount of new oxygen carrier that needs to be added to maintain system inventory. Reduction in the oxygen carrier make-up helps to improve

the overall system economics. The current cyclone was determined to achieve only an efficiency of 95-98% over the entire particle size distribution, which is well under the typical cyclone efficiencies of > 99%. The CLR typically operates with a solid circulation rate of 288 kg/hr and an inventory of 50 kg. With a 95% efficient cyclone, the entire inventory will be lost out the cyclone exit in 3.5 hrs. Increasing the efficiency by only 4.9 percentage points to 99.9% drastically increases this time to 174 hrs; from less than a day to over a week.

To improve the performance of this cyclone, a CFD based modeling & simulation campaign was performed to assist the optimization procedure because previous attempts using experimentally derived correlations were not successful in improving cyclone performance significantly. The basic methodology for the process is to use statistical design of experiments principles to generate simulation samples that cover the parameter design space through the use of a space filling sampling method like Latin Hypercube sampling. Ensembles of CFD models are then created, executed, and post-processed. A response surface (surrogate model) is created to characterize the relationship between input parameters and the Quantities of interest (QoI) to avoid the necessity to perform costly simulations during optimization phase. Finally, the CFD-surrogate is used for objective function evaluations by the optimization method selected (differential evolution for this problem) to find the optimal design condition. The optimization process was demonstrated to achieve a cyclone design that is more efficient than the original design with a lower pressure drop. The rest of the paper is organized to as follows, after a brief overview of the computational model employed,

we present the details of the simulation campaign and the open-source toolkit, Nodeworks employed to generate and manage the sampling simulations. The problem configuration and the baseline results presented for comparison with the respect to optimized results. Next the process to generate the surrogate model is described, which replaces the actual CFD model to reduce computational cost. Prior the optimization results, the results of the sensitivity study are presented to demonstrate how such study can guide the optimization process by identifying the most influential parameters. Finally, the results of the optimization are presented by comparing against the baseline results to document the improvements.

## METHODS

### **MFix: Particle-in-Cell model**

MFix is an open source computational fluid dynamics code specifically developed for modeling gas-solid flows often found in the energy and chemical processing industries [2]. MFix has been used to model fluidized beds, circulating fluidized beds, cyclones, and hoppers—all of which are commonly found in chemical processes such as coal gasification, fluid catalytic cracking, waste treatment, and chemical production. MFix has three solids models including the two-fluid model (TFM), particle-in-cell (PIC) model, and the discrete element method (DEM). This work utilized the recently released MFix version 19.1 which includes a rigorous re-write of the PIC model.

PIC is a Eulerian-Lagrangian multiphase flow modeling approach that represents the fluid as a continuum while using parcels, or clouds, to represent groups of real particles with similar physical characteristics. Solids volume fraction is used to calculate

a solids pressure gradient on the Eulerian grid which in turn is used to approximate collisional stresses acting on parcels and prevent overpacking. This method avoids the high computational costs associated with CFD-DEM, specifically collision detection and small collision time scales, allowing PIC to be significantly faster. The current PIC implementation in MFIX closely follows the method of Snider [3].

### **Nodeworks: Surrogate Modeling and Analysis toolset**

Nodeworks is an open source graphical programming interface library and application where users can add, delete, and connect nodes to create customized visual workflows [4]. Nodes perform prescribed operations on data that are passed between nodes using connections. The library has been specifically developed in the Python programming language to be very flexible, portable and support a wide variety of applications with several collections of default nodes to assist deployment of commonly used workflows very quickly, even for novice users. Users can also create and add custom nodes for specific applications such as machine learning.

This work leverages a collection of nodes known as the Surrogate Modeling and Analysis toolset, which has been developed for implementation of workflows to construct and use data-fitted surrogate models, or response surfaces. The surrogate modeling and analysis toolset provides access to specialized nodes like optimization, sensitivity analysis, forward propagation of uncertainty, and Bayesian calibration. Nodeworks is directly embedded into the MFIX's graphical user interface (GUI), allowing Nodeworks to create input decks involving parametrically varying inputs and run the simulation campaigns with ease, Fig. 1. Nodeworks version 19.1 was used in this work.

Nodeworks can be also employed by other modeling tools to create similar workflows with ease.

## RESULTS

### Problem Definition and Baseline Simulation Setup

The cyclone constructed and installed in NETL's chemical looping reactor was modeled and used as the baseline case for the simulation campaigns. Fig. 2 shows an illustration of the cyclone modeled for this study. The problem configuration and the model itself is already provided as a tutorial case within the MFix distribution due to illustrative nature of the problem for typical multiphase flow applications. However, several changes were made to the tutorial to facilitate this optimization study, mainly the addition of more parametric geometry parameters.

The Particle-in-Cell (PIC) method was used to model the solids phase, which consists of monodisperse, high density polyethylene (HDPE) particles with a diameter of  $871\mu\text{m}$  and a density of  $860 \text{ kg/m}^3$ . The inlet boundary conditions for the cyclone was a gas mass flow rate of  $0.02 \text{ kg/s}$  and solids flow rate of  $0.08 \text{ kg/s}$ . Two pressure outlets are used, one for the vortex tube outlet at the top of the cyclone and a second for the cyclone outlet at the bottom. A constant pressure of  $101.32 \text{ kPa}$  is set for both pressure outlets. Since the cyclone is part of a larger unit, there is a standpipe located at the cyclone (bottom) outlet which prevents gas from leaving the bottom of the cyclone. To represent this in the cyclone model, a semi-impermeable surface is placed at the bottom outlet. The large resistance of the semi-impermeable surface prevents gas flow

through the cyclone outlet while allowing the solids to leave, i.e., the solid parcels are not affected by the semi-impermeable surface. All wall sections of the cyclone are treated as no-slip boundaries for the gas-phase. The wall geometry is defined by a stereolithography (STL) file, generated for each case from the parametrized geometry variables. A uniform CFD grid was applied of 5 mm in each direction or 5.7 times the particle diameter in the x-, y-, and z- directions. For the base geometry, the CFD grid resolves the geometry with four cells across the inlet and 16 cells along the length of the cyclone. A statistical weight of  $W = 1.0$  was applied in the PIC model, i.e., each parcel represents a single particle. No turbulence model was used.

A point monitor was placed at the inlet of the cyclone to measure the absolute pressure and write the transient data as a comma separated value (CSV) formatted ASCII file at a frequency of 100 Hz. The resulting time-averaged pressure drop of the cyclone can then be calculated by subtracting the vortex tube outlet pressure (101.32 kPa) from this point monitor pressure data. The amount of solids leaving the cyclone through the vortex tube is captured and recorded by a User Defined Function (UDF) subroutine that was compiled into the MFix code prior to the runs. This value was also saved to a CSV file at a frequency of 100 Hz. Due to the transient nature of the flow in the cyclone, the time-dependent solution of the flow field was simulated for a total time duration of 30 seconds, which required an execution time of 2.5 wall-clock hours on 8 cores for the base case. The transient data is time averaged from 5 to 30 s to produce scalar quantities of interest.

### **Design of Experiments (DOE)**

Optimization typically involves many evaluations of the objective function. In the context of the current problem, this implies the need for performing many CFD simulation corresponding to each updated evaluation of the optimizer. Considering the wall-clock time and computational resources required to perform many hundreds or even thousands of evaluations, performing direct simulations as part of optimization process is (potentially) computationally intractable. Instead, a surrogate model can be constructed to characterize the relationship between input parameters and quantities of interest using fewer samples of CFD simulations. The surrogate model is then used in lieu of the CFD simulations for optimization because it is significantly computationally cheaper to evaluate. The surrogate model (response surface) can be constructed through carefully designed simulation campaign using the principles of statistical design of experiments [5].

For the purposes of this study, a space-filling sampling method is utilized to cover the five-input parameter phase-space as much as possible with 100 samples. Here, a genetically optimized Latin hypercube design was used, which was already implemented in Nodeworks. We note that 20 samples per parameter exceeds the commonly used guideline for space-filling designs [6], which recommends at least 10 samples per parameter. The primary reason was the anticipated likelihood of high failure rate of the simulation without convergence due to the statistically generated variations in the geometry and accompanying automated mesh generation for each case. Hence, the employed sampling size enabled a margin that affords nearly 50% failure rate to still satisfy the guideline [6]. The systematically varied input parameters

were the five geometric dimensions of the cyclone, i.e., cyclone barrel radius ( $r_{\text{barrel}}$ ), the vortex tube radius ( $r_{\text{vortex}}$ ), and vortex tube height ( $h_{\text{vortex}}$ ), as well as the inlet height ( $h_{\text{inlet}}$ ) and inlet width ( $w_{\text{inlet}}$ ), which were all varied continuously within the lower and upper bounds presented in **Error! Reference source not found..**

Using 100 sampling simulations resulted in a balanced space-filling design with no noticeable correlation between the samples, as shown in Fig. 3. The quality of the space filling design property can be assessed through various statistical measures. For this simulation campaign a wrap-around L2-discrepancy measure, calculated as described by Eq. 5 from [7], of 0.00295 was computed and considered adequate for the constructed samples. The smaller the wrap-around L2-discrepancy measure, the better the samples are at filling the space.

### **Model Creation and Dispatch**

As the variable input parameters are geometrical, each sample design point requires a unique set of code input files, e.g., mfix.dat and geometry.stl files. Hence, the input decks for all 100 unique simulations were generated automatically using Nodeworks' design of experiments feature through the MFIX GUI. The MFIX executable was compiled using GCC 8.2 and OpenMPI 3.1.3. Finally, the DOE node was also used to launch all simulations to a queueing system on Joule 2. After launching the jobs, an MFIX-Nodeworks monitoring panel becomes active, from which the progress of the simulation campaign was monitored.

All simulations were carried out on NETL's high-performance computing cluster, Joule 2. Each sampling simulation was executed in distributed parallel mode using 8

MPI processes per simulation on 20 core Intel Xeon Gold 6148 series processors clocked at 2.4 GHz. Since the cell size was fixed at 5mm, the resulting grid resolutions varied from 40,320 to 169,764 cells, depending on the diameter of the barrel,  $r_{\text{barrel}}$ . All simulations were executed for a total simulated time duration of 30 seconds. Consequently, the wall-clock time required for completion of the simulations ranged from 21 minutes to 7 hours depending mainly on the number of grid cells. Fig. 4 shows an agglomerate snapshot of the simulation campaign at 30 seconds.

### **Response Surface**

The quantity of interest was calculated by temporal averaging of the results from sampling simulation output files (i.e., pressure and mass logs), discarding the first 5 seconds to allow the startup transient to pass and a steady operational state to be achieved. These averages were then normalized between 0 and 1 based on the observed minimum and maximum values. A single quantity of interest (QoI) was determined by adding these two values together to calculate a composite scalar value, so that the pressure drop and mass loss were equally weighted. Three simulations (run numbers 50, 51, and 74) were removed from the analysis because the QoI for these three outlier cases was significantly different than the remaining 97 simulation's QoI. For the three removed cases, the gap between the vortex tube and the barrel was on the order of one cell, resulting in a poor-quality fluid mesh, which caused solver convergence problems. A combined failure and outlier removal rate of just 6% was surprisingly good and much lower than the allowable 50% loss threshold built in to the samples by oversampling.

A strong correlation between the pressure drop and the radius of the vortex tube,  $r_{vortex}$ , can be observed from the scatter matrix plots, as shown in Fig. 5, which visually illustrates the relationships between input parameters and quantity of interest. The strong correlation observed is expected since the mass flow rate of gas is held constant and the radius of the vortex tube effects the area for that gas to leave. It is hard to distinguish other clear trends in terms of correlations.

The GaussianProcessRegressor (GPR) in the scikit-learn toolkit was used to construct a data-fitted surrogate model based on Gaussian Process Models (GPM) [8]. This method was selected over other surrogate model methods due to favorable unique features of GPM based surrogate models such as inherent uncertainty estimation. The default radial basis function (RBF) kernel was used for the GPM based surrogate. The GPR automatically fits a variety of hyper parameters by using the default *fmin\_l\_bfgs\_b* optimizer with 9 restarts. A non-negligible dependence on alpha, the value added to the diagonal of the kernel matrix during fitting, was observed. The alpha parameter controls the noise level or smoothing of the data. Larger values of alpha correspond to increased noise level in the observations and is similar to adding white noise to the kernel.

To pick an alpha that does not result in an over-fitted surrogate model, but still represents the variability of response well, the GPM was fit over a range of alphas. At each alpha, 10% of the samples were removed from the training set and used to calculate a mean squared error (MSE) of the fitted model. Since the MSE depends on which samples are withheld, the cross-validation procedure is repeated 100 times, each time randomizing the 10% holdout samples. The resulting distributions of MSE for each

alpha is shown in Fig. 6. An alpha of  $5 \times 10^{-9}$  displayed the best trade-off between overfitting and underfitting the data.

The GPM was then refitted with all of the samples (0% hold-out) and an alpha of  $5 \times 10^{-9}$  for use in the optimization routine. Although the scatter is significant, the GPM represents the surface reasonably well. Fig. 7 displays the fitness of the surrogate as a parity plot between predicted by the surrogate model (i.e., GPM) and true response (i.e., results from MFIX simulation) QoIs. It is observed that the GPM was determined to have difficulty in fitting the samples with high QoI values, most likely due to the extreme geometry configurations at the boundary of the input parameter ranges. Since the optimization routine will be minimizing the QoI, it is more important to have a better fit closer to a QoI of 0.

Using the GPM based surrogate model, the relationships between the variables and the quantity of interest can be better visualized by evaluating the model over the ranges of the variables, while keeping the unvaried variables at their nominal setting (i.e. midpoint of their range), as shown in Fig. 8. A strong relationship is observed between the QoI and the radius of the vortex tube,  $r_{\text{vortex}}$ . The other four variables do not exhibit strong relationships, especially the inlet width,  $w_{\text{inlet}}$ , and height,  $h_{\text{inlet}}$ , having an almost flat response.

### **Sensitivity Analysis**

Prior to the optimization of the geometry, a sensitivity analysis was performed using the Sensitivity Analysis node in Nodeworks to better understand which input parameters have the most influence on the quantities of interest. Using the surrogate

model constructed in the previous section and the Python library SALib, version 1.2 [9], variance decomposition based Sobol' Indices method [10] was employed for the global sensitivity analysis. The results of the sensitivity indices are shown in Fig. 9 in terms of first order indices (which indicate the main effects such as standalone  $r_{barrel}$ ), second order indices which aims to illustrate the effect of interactions between main effects and finally the total indices that capture all. As seen from Fig. 9, the most influential input parameter was determined to be  $r_{vortex}$ , which was followed by  $r_{barrel}$ . This also agrees with the trends observed in the surrogate model, Fig. 8. Hence, the optimization process is expected to be driven primarily by variation in  $r_{vortex}$ , and  $r_{barrel}$ , which can be also confirmed qualitatively by the final optimized geometric configuration visualization in Fig. 10 when compared to the original design.

## Optimization

Using the GPM based surrogate model and the differential evolution optimization routine in `scipy` [11], an optimal design was found by minimizing the QoI, a normalized sum of cyclone pressure drop and mass loss. The differential evolution algorithm is a global optimization routine that does not use gradients [12]. The technique works by randomly generating trials and evaluating the model at those locations. The *best1bin* strategy is then used to mutate the best member, creating a new set of trials. Considering the fit GPR shows smooth continuous functionality with no local minima or maxima and re-running the optimization routine multiple times with random initial seeds, the authors are confident that a global minimum has been found.

**Error! Reference source not found.** compares the optimized design to the original design parameters.

The optimal design has a larger diameter vortex tube and barrel with a narrower and taller inlet, Fig. 10. According to Fig. 8, the QoI has a strong functionality with the vortex tube radius, mainly due to the pressure drop. It makes sense that the optimal design has a larger vortex tube, reducing the pressure drop from the original design by 11 times (596 Pa for the original compared to 55 Pa for the optimal design).

It is more difficult to elucidate the most influential factor affecting the mass loss, although it appears to be related to the inlet geometry. Having a narrower, taller inlet keeps the solids entering the cyclone barrel closer to the wall. The optimal cyclone does reduce the average amount of mass loss by a factor of 2.3 compared to the original cyclone design.

Unfortunately, two of the parameters optimal points are at the edge of the design space. The height of the inlet,  $h_{inlet}$ , has an optimal value of 0.12, which is the maximum of the tested range. Similarly, the inlet width,  $w_{inlet}$ , has an optimal value of 0.015, which is the minimum of the tested range. Future work should include resampling with additional points extending beyond these two extrema to determine if the identified optimal is global. However, the optimal width of the inlet may converge to the particle diameter such that all the particles are forced to enter the barrel against the wall, reducing mass loss, with a corresponding increase in inlet height, to preserve the cross-sectional area, preventing an increase in the pressure drop.

## CONCLUSIONS

MFIX, in combination with the Surrogate Modeling and Analysis toolset of Nodeworks, was employed to optimize the cyclone in the 50 kW<sub>th</sub> Chemical Looping Reactor at NETL. A simulation campaign with 100 sampling simulations utilizing MFIX-PIC for cyclone modeling was carried out by systematically varying the five critical design parameters; vortex tube radius and vortex tube length, barrel radius, inlet width, and inlet height. Out of the 100 simulations, only 3 were removed from the analysis due to poor mesh generation resulting in extraneous QoI values (which were adversely affecting the quality of the data fitted surrogate before being rejected from the analysis). It is important to construct the best surrogate model for the given dataset as the surrogate model replaces the actual CFD simulations required for function evaluations during the optimization procedure. A Gaussian Process Model based data-fitted surrogate model was determined to give the best fit. Considering the computational resource requirements for 100 sampling simulation for a complex fluid flow, future study will aim to find the minimum number of samples that will facilitate the construction of an adequate surrogate model and yield to same analysis results for optimization.

Prior to the optimization procedure, it is beneficial to perform a global sensitivity analysis once an adequate quality surrogate model is constructed. Sensitivity analysis will provide valuable insight on which input parameters have the most influence on the quantities of interest. Hence, for situations where there are more than a couple input parameters, the optimization process can be driven by the insight gained from

sensitivity analysis through varying only the most influential input parameters identified and setting the remaining ones at nominal values, to reduce the burden during the optimization. Nodework's Sensitivity Analysis node was utilized to effortlessly calculate Sobol' Total Indices based global sensitivity analysis, which showed that  $r_{vortex}$ , and  $r_{barrel}$  are the most influential parameters affect the quantity of interest.

A differential evolution optimization routine, in the Nodework's general optimizer node, was used to find the minimum of the surrogate model or the optimal cyclone design. The optimal design has a larger diameter and longer vortex tube, a larger diameter barrel, and a taller and narrower solids inlet. The design is predicted to have a 11 times lower pressure drop and 2.3 times lower mass loss than the original cyclone.

The next phase of the study will be confirmation of the optimal design by implementing the optimal design in the chemical looping reactor to quantify if the new design configuration can demonstrate an improved performance (I.e., a lower pressure drop and higher separating efficiency) to the same extent observed in simulation results.

## **ACKNOWLEDGMENT**

This work was performed in support of the US Department of Energy's Fossil Energy Advanced Reaction Systems Program. The Research was executed through NETL Research and Innovation Center's Advanced Reaction Systems effort.

## NOMENCLATURE

$h_{inlet}$	height of the cyclone inlet, m
$h_{vortex}$	height of the vortex finder, m
$r_{barrel}$	radius of the cyclone barrel, m
$r_{vortex}$	radius of the vortex finder, m
$w_{inlet}$	width of the cyclone inlet, m

## REFERENCES

- [1] S. Bayham, D. Straub and J. Weber, "Operation of the NETL Chemical Looping Reactor with Natural Gas and a Novel Copper-Iron Material," National Energy Technology Lab. (NETL), Pittsburgh, PA, and Morgantown, WV, 2017.
- [2] US Department of Energy National Energy technology Laboratory, "MFiX," [Online]. Available: [mfix.netl.doe.gov/mfix](http://mfix.netl.doe.gov/mfix). [Accessed 2019].
- [3] D. M. Snider, "An Incompressible Three-Dimensional Multiphase Particle-in-Cell Model for Dense Particle Flows," *Journal of Computational Physics*, vol. 170, pp. 523-549, 2001.
- [4] US Department of Energy National Energy Technology Laboratory, "Nodeworks," [Online]. Available: [mfix.netl.doe.gov/nodeworks](http://mfix.netl.doe.gov/nodeworks). [Accessed 2019].
- [5] R. H. Myers and D. C. Montgomery, Response surface methodology: process and product optimization using designed experiments, vol. 4, New York: Wiley, 1995.
- [6] J. L. Loeppky, J. Sacks and W. J. Welch, "Choosing the sample size of a computer experiment: A practical guide," *Technometrics*, vol. 51, no. 4, pp. 366-376, 2009.
- [7] K. T. Fang and C. X. Ma, "Wrap-Around L2-Discrepancy of Random Sampling, Latin Hypercube, and Uniform Designs," *Journal of Complexity*, vol. 17, pp. 608-624, 2001.
- [8] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.
- [9] W. Usher, J. Herman, T. Iwanaga, N. Cellier, C. Whealton, D. Hadka, Xantares, F. Rios, Bernardoc, C. Mutel, J. van Engelen, H. Kranas and Antlford, "SALib/SALib: SALib v1.2," November 2018. [Online]. Available: <https://doi.org/10.5281/zenodo.1509666>.

- [10] A. Saltelli, M. Ratto, A. Terry, F. Campolongo, J. Cariboni and D. Gatelli, *Global Sensitivity Analysis: The Primer*, John Wiley & Sons, 2008.
- [11] E. Jones , T. Oliphant , P. Peterson and others, "SciPy: Open source scientific tools for Python," 2001. [Online]. Available: <http://www.scipy.org/>. [Accessed 22 January 2019].
- [12] R. Storn and K. Price, " Differential Evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization*, vol. 11, pp. 341-359, 1997.
- [13] T. Oliphant, *A guide to NumPy*, USA: Trelgol Publishing, 2006.

### Figure Captions List

Fig. 1 Screenshot of the Nodeworks workflow used in this work

Fig. 2 Cyclone schematic, all units in meters

Fig. 3 Scatterplot matrix of all five input parameters considered for the cyclone optimization study

Fig. 4 A snapshot of the simulations from 96 of the 100 cyclones

Fig. 5 Variation of the mass loss, pressure drop, and resulting quantity of interest (QoI) with parameter

Fig. 6 Mean Square Error distributions of 100 cross-validation runs for each alpha

Fig. 7 Parity plot comparing the simulation value, denoted as “True” in x-axis to the model prediction in y-axis

Fig. 8 Gaussian process model prediction evaluated at the center of the ranges (dashed line) compared to the QoI (circles), and QoI values within +/- 30% of the center (triangles)

Fig. 9 Sobol' Total Indices based global sensitivity analysis results for the quantity of interest showing the most influential parameters in terms of main effects (first order indices) and their interactions (second order indices) from left to the right

Fig. 10 Original cyclone (left) compared to the optimal cyclone (right)

**Table Caption List**

Table 1      Input parameters and lower/upper bounds used in designing the sampling simulation

Table 2      The original design compared to the optimal design

Fig. 1

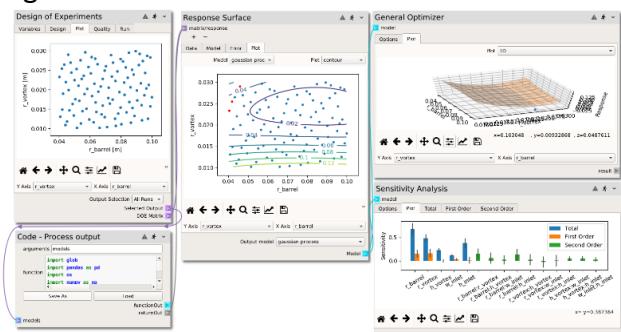


Fig. 2

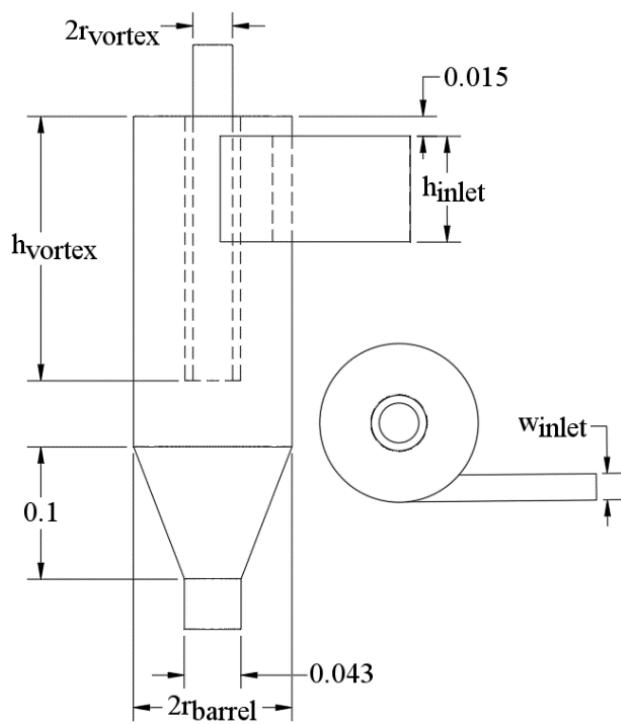


Fig. 3

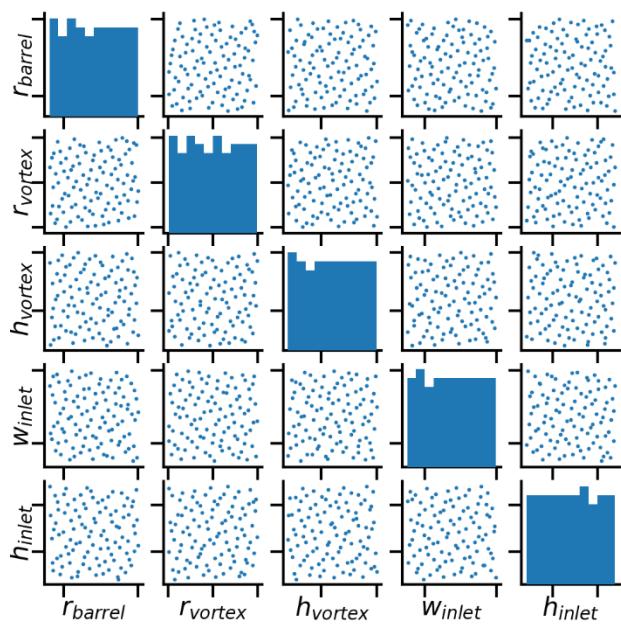


Fig. 4



Fig. 5

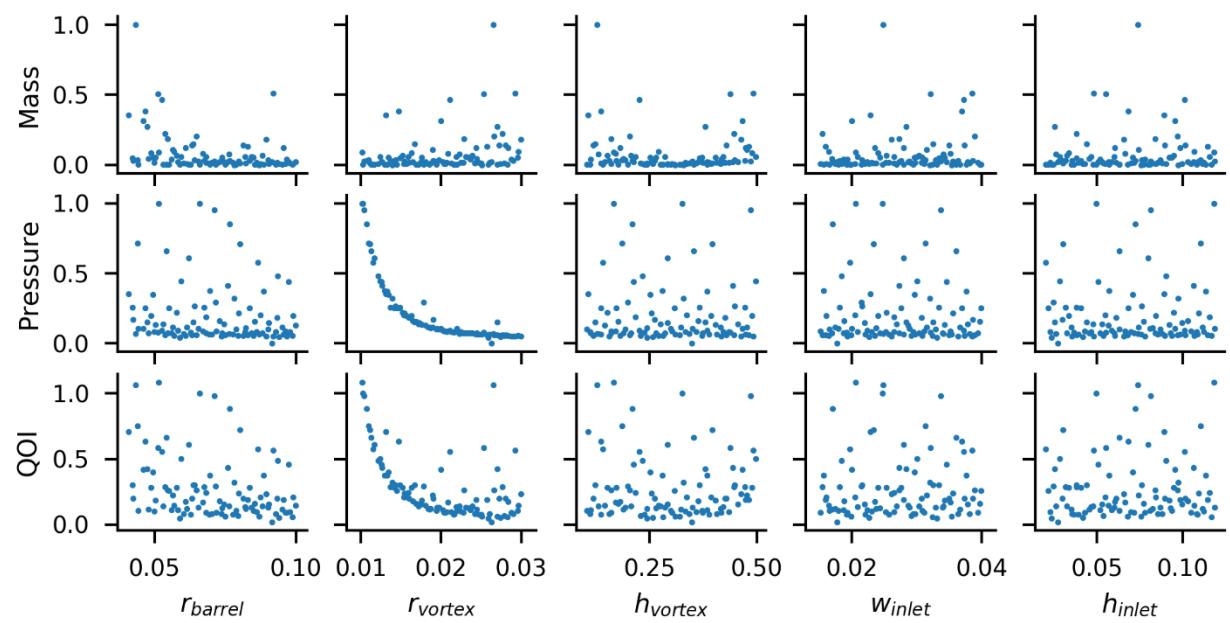


Fig. 6

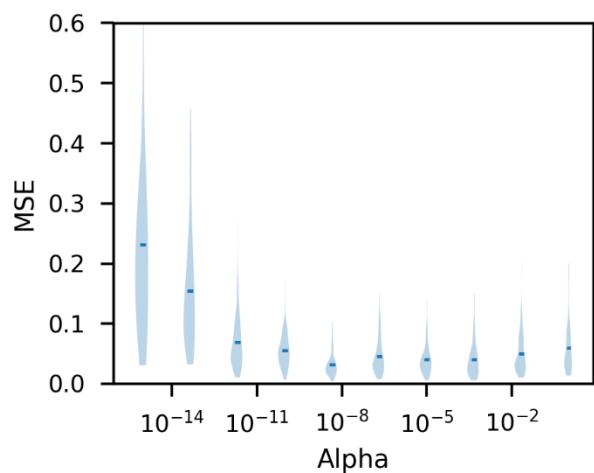


Fig. 7

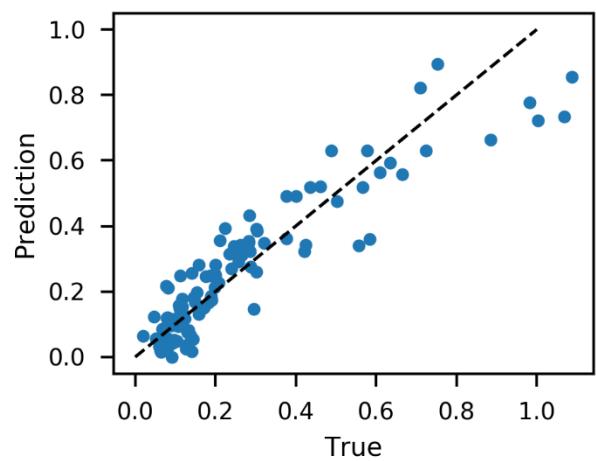


Fig. 8

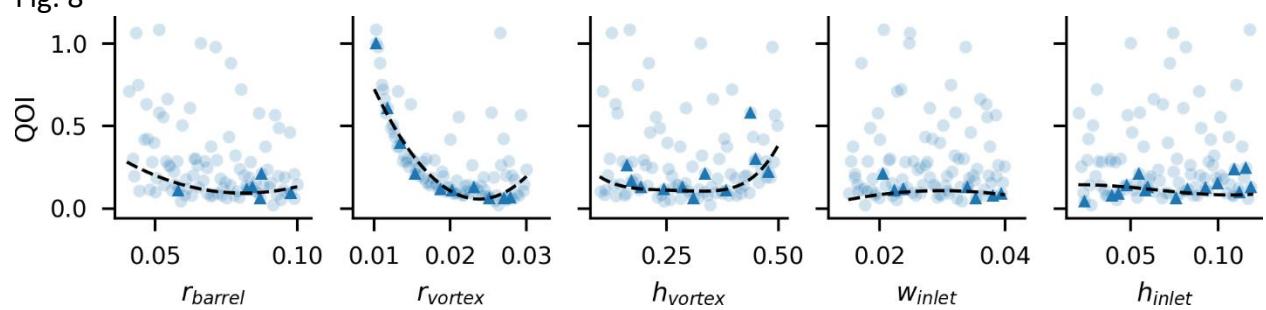


Fig. 9

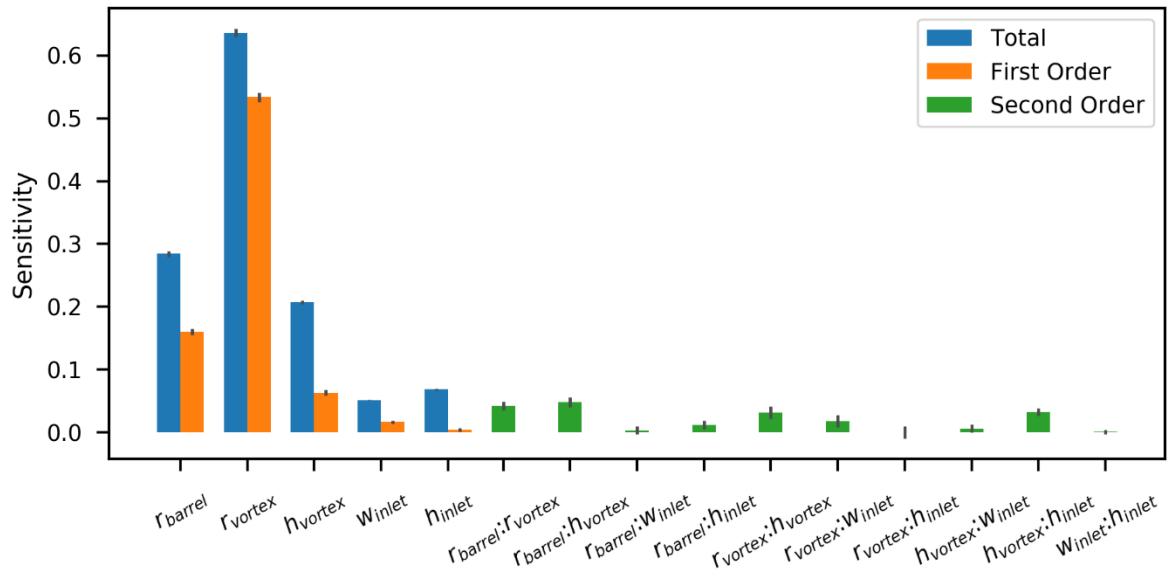


Fig. 10

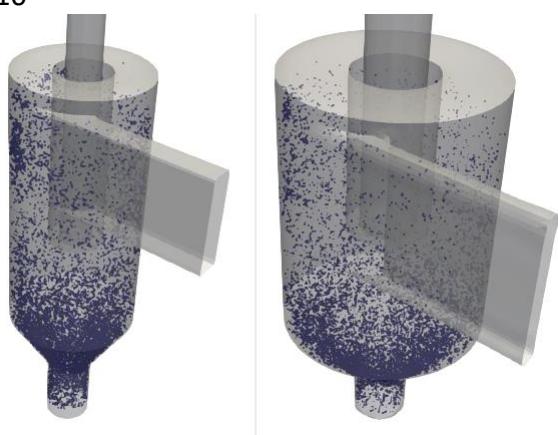


Table 1

Variable	min (m)	max (m)
$r_{\text{barrel}}$	0.04	0.1
$r_{\text{vortex}}$	0.01	0.03
$h_{\text{vortex}}$	0.1	0.5
$h_{\text{inlet}}$	0.02	0.12
$W_{\text{inlet}}$	0.015	0.04

Table 2

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