



# Multi-objective Geometry Optimization of a Gas Cyclone Using Triple-Fidelity Co-Kriging Surrogate Models

Prashant Singh<sup>1</sup>  · Ivo Couckuyt<sup>2</sup> · Khairy Elsayed<sup>3,4</sup> · Dirk Deschrijver<sup>2</sup> · Tom Dhaene<sup>2</sup>

Received: 15 October 2016 / Accepted: 24 April 2017 / Published online: 10 May 2017  
© The Author(s) 2017. This article is an open access publication

**Abstract** Cyclone separators are widely used in a variety of industrial applications. A low-mass loading gas cyclone is characterized by two performance parameters, namely the Euler and Stokes numbers. These parameters are highly sensitive to the geometrical design parameters defining the cyclone. Optimizing the cyclone geometry therefore is a complex problem. Testing a large number of cyclone geometries is impractical due to time constraints. Experimental data and even computational fluid dynamics simulations are time-consuming to perform, with a single simulation or experiment taking several weeks. Simpler analytical models are therefore often used to expedite the design process. However, this comes at the cost of model accuracy. Existing techniques used for cyclone shape optimization in literature do not take multiple fidelities into account. This work combines cheap-to-evaluate well-known mathematical models of cyclones, available data from computational fluid dynamics simulations and experimental data to build a triple-fidelity recursive co-Kriging model. This model can be used as a surrogate with a multi-objective optimization algorithm

---

Communicated by Zenon Mroz.

---

✉ Prashant Singh  
prashant.singh@it.uu.se

- <sup>1</sup> Division of Scientific Computing (TDB), Department of Information Technology, Uppsala Universitet, Box 337, 751 05 Uppsala, Sweden
- <sup>2</sup> Department of Information Technology (INTEC), Ghent University-imec, Technologiepark-Zwijnaarde 15, 9052 Gent, Belgium
- <sup>3</sup> Mechanical Power Engineering Department, Faculty of Engineering at El-Mattaria, Helwan University, Masaken El-Helmia P.O., Cairo 11718, Egypt
- <sup>4</sup> Department of Mechanical Engineering, Research Group Fluid Mechanics and Thermodynamics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

to identify a Pareto set of a finite number of solutions. The proposed scheme is applied to optimize the cyclone geometry, parametrized by seven design variables.

**Keywords** Cyclone separator · Surrogate-based optimization (SBO) · Multi-fidelity data · Sparse Kriging · Recursive co-Kriging

**Mathematics Subject Classification** 68T37 · 68T05 · 68U20 · 97M10 · 93A30 · 49Q10 · 80M50 · 76Txx · 76Dxx · 68Uxx · 90-08

## Nomenclature

$a$	Cyclone inlet height (m)
$b$	Cyclone inlet width (m)
$B_c$	Cyclone cone-tip diameter (m)
$d$	Number of design variables (-)
$D$	Cyclone body diameter (m)
$D_x$	Cyclone vortex finder diameter (m)
$h$	Cylindrical part height (m)
$H_t$	Cyclone total height (m)
$\mathcal{K}$	Kriging model
$\mathcal{K}_A$	Kriging model with training data from analytical models
$\mathcal{K}_{CFD}$	Kriging model with training data from CFD simulations
$\mathcal{K}_{Exp}$	Kriging model with training data from Experiments
$n$	Number of sample points (-)
$N_s$	Number of sample points in each dimension of gridded data
$\mathcal{P}$	Pareto set
$s$	Standard deviation of model prediction
$S$	Vortex finder length (m)
$V_{in}$	Area-average inlet velocity (m/s)
$\mathbf{x}$	Data instance
$X$	Training set of data
$X_A$	Training set of data from analytical models
$X_{CFD}$	Training set of data from CFD simulations
$X_{Exp}$	Training set of data from experiments
$\hat{y}$	Mean of model prediction
$\mathbf{y}$	Vector of training set targets
$\mathbf{y}_A$	Vector of training set targets for data from analytical models
$\mathbf{y}_{CFD}$	Vector of training set targets for data from CFD simulations
$\mathbf{y}_{Exp}$	Vector of training set targets for data from experiments
$Y$	Random variable

## Greek letters

$\alpha$	Regression function coefficient
$\gamma$	Scaling parameter for sub-Kriging models

$\Delta P$	Pressure drop (N/m <sup>2</sup> )
$\epsilon$	Distance threshold
$\mu$	Dynamic viscosity (kg/(m s))
$\phi(\cdot)$	Transformation function
$\Phi\cdot$	Normal cumulative distribution function
$\psi^{(i)}$	Correlations between $Y(\mathbf{x})$ at the point to be predicted and at the sample data points
$\Psi$	Correlation matrix of given data samples
$\rho$	Gas density (kg/m <sup>3</sup> )
$\sigma^2$	Variance
$\theta$	Kriging hyper-parameter variables

### Dimensionless numbers

$E_u$	Euler number
$Re$	Reynolds number
$Stk_{50}$	Stokes number corresponding to 50% separation efficiency

### Abbreviations

BEEQ	Bayesian estimation error quotient
CFD	Computational fluid dynamics
EGO	Efficient global optimization
EMO	Efficient multi-objective optimization
MLE	Maximum likelihood estimation
MOEA	Multi-objective evolutionary algorithm
MSE	Mean squared error
NSGA-II	Non-dominated sorting genetic algorithm 2
RBF	Radial basis function
SBO	Surrogate-based optimization
SMS-EMOA	$S$ -Metric selection-evolutionary multi-objective optimization algorithm
SVR	Support vector regression

## 1 Introduction

Performing physical experiments to test a cyclone geometry is not practical as it is very time-consuming and expensive. Computational fluid dynamics (CFD) simulations offer a faster alternative, but they are still computationally too expensive for integration with optimization algorithms as each simulation takes several weeks. Mathematical models are very cheap to evaluate but less accurate in comparison with CFD simulations.

Due to the prohibitory computational cost of identifying the complete Pareto Front of solutions, the problem setting considered in this work involves finding a

Pareto-optimal set (Pareto set) of a finite number of solutions instead of the complete continuous Pareto front. Multi-objective evolutionary algorithms (MOEAs) are a popular choice for solving such problems. However, MOEAs typically require a large number of objective function evaluations to converge and offer viable solutions. Cheap-to-evaluate mathematical models can be used with MOEAs in practice, but solutions cannot be guaranteed to be accurate. Surrogate-based optimization (SBO) algorithms reduce the number of objective function evaluations required during the optimization process and can be used in conjunction with mathematical models and CFD simulations.

Given the expensive nature of high-quality data, it is important to utilize all available fidelities. SBO algorithms typically select data points during the optimization process, which are then evaluated and used to update the surrogate. This acts as a continuous validation mechanism and requires the availability of a simulation code. When only precomputed (old) data are available (e.g., from previous physical experiments or old CFD simulations), time constraints and practicality do not offer the freedom to request additional data at arbitrary locations. Therefore, there is a need for an optimization scheme which can utilize multiple fidelities of available data to optimize the shape of a cyclone separator.

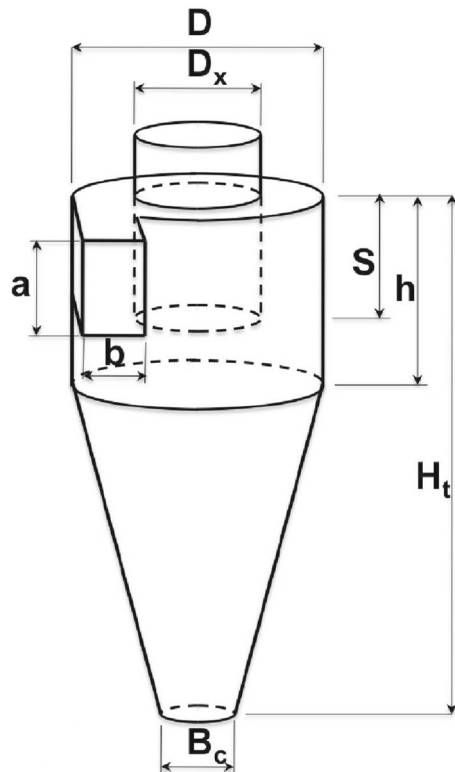
This paper presents an algorithm that trains a surrogate model using multiple fidelities of available data. The surrogate model can then be used as a replacement of mathematical models or CFD simulations by MOEAs to obtain a Pareto set of solutions. Recursive co-Kriging formulation is used to train the model, with the scheme being demonstrated using three fidelities of data (physical experiments, CFD simulations and analytical models in order of decreasing quality). The novelties of this work include triple-fidelity surrogate modeling (not many works can be found in the literature that attempt it), training a sub-Kriging model from a million-point dataset, and handling missing output dimensions.

The paper is organized as follows. Section 2 presents the detailed description of the problem of cyclone shape optimization. Section 3 introduces Kriging models, and the recursive co-Kriging formulation. Section 4 explains the proposed algorithm for training a multi-fidelity recursive co-Kriging model. Sections 5 and 6 list the experimental setup and discuss the results, respectively. Section 7 concludes the paper.

## 2 Geometry Optimization of a Gas Cyclone Separator

Gas cyclones are separation devices that apply a swirl to separate out the dispersed phase (solid particles) from the continuous phase (gas) [1,2]. They have been widely used in many industrial applications for many years. The advantages of cyclones over other separation devices include simple construction, low maintenance and no moving parts. The tangential inlet Stairmand high-efficiency cyclone [3] shown in Fig. 1 is widely used to separate particles from a gaseous stream. The gas-solid mixture enters tangentially which produces a swirling motion. The centrifugal forces throw the particles to the wall, and they are eventually collected at the bottom of the cyclone. The gas (with some small particles) changes its direction and exits from the top of the cyclone.

**Fig. 1** Schematic diagram of the gas cyclone separator



Although the geometry of the gas cyclone is simple, the flow is three-dimensional and unsteady. The swirling flow in the cyclone makes the study of gas cyclones more complicated. A wide range of industrial applications of gas cyclones has stimulated many experimental research studies toward the flow pattern and the performance of gas cyclones [1]. These experiments are generally costly because they involve manufacturing of gas cyclones and using expensive measurement techniques such as laser Doppler anemometry, and particle image velocimetry [4,5]. With the availability of computing resources, CFD has become an alternative approach to study cyclone separators [2,6,7]. CFD is currently used to study the flow pattern [8–11], as a simulator for CFD based geometry optimization with different surrogate models [7,12–16] as well as for shape optimization using the adjoint method [17]. Analytical models based on simplified theoretical assumptions are still in use for prediction of cyclone performance [18–23]. For further details regarding the most widely used analytical models in gas cyclone performance prediction, the reader is referred to Hoffmann and Stein [23] and Elsayed [24].

The dominant factor influencing the cyclone performance and flow pattern is the cyclone geometry [24]. The cyclone geometry is described by seven basic dimensions, namely the cyclone inlet height  $a$  and inlet width  $b$ , the vortex finder diameter  $D_x$  and length  $S$ , barrel height  $h$ , total height  $H_t$  and cone-tip diameter  $B_c$  [24] as shown in

Fig. 1. All the parameters are given as the respective ratios of cyclone body diameter  $D$ .

The two performance indicators widely used in low-mass loading gas cyclones are the pressure drop and the cutoff diameter  $x_{50}$  [12,24]. The cutoff diameter  $x_{50}$  is defined as the particle diameter which produces a 50% separation efficiency [23]. In some other studies, the overall collection efficiency is used instead of the cutoff diameter. The literature has some empirical formulae that can use  $x_{50}$  to generate the grade efficiency curve (particle diameter versus the collection efficiency) [23].

The problem of geometry optimization for gas cyclones has been well-studied in the literature. Ravi et al. [25] carried out a multi-objective optimization of a system of  $N$  identical cyclone separators in parallel. The two objectives were the overall collection efficiency and the pressure drop. Salcedo and Candido [26] conducted two optimization studies to maximize cyclone collection and an efficiency/cost ratio. They applied the Mothes and Loffler [27] analytical models.

Swamee et al. [28] optimized the number of parallel cyclones, the barrel diameter and vortex finder diameter at a certain gas flow rate. The two cost functions (the pressure drop and cutoff diameter) were blended into a single objective problem [17].

Safikhani et al. [29] performed a multi-objective optimization of cyclone separators using an artificial neural network surrogate model. Pishbin and Moghiman [30] applied genetic algorithms with seven geometrical parameters using 2-D axisymmetric CFD simulations.

There are many other optimization studies such as [7,13–16] where the source of training data for surrogate models is either experimental data, CFD simulations or analytical models. Elsayed [13] applied co-Kriging in the geometry optimization of gas cyclone using two sources of training data; CFD and an analytical model.

It is preferable to use dimensionless numbers in the design variables and the two performance parameters. The dimensionless pressure drop (Euler number,  $Eu$ ) is defined as [23],

$$Eu = \frac{\Delta p}{\frac{1}{2}\rho V_{in}^2}, \tag{1}$$

where  $\Delta p$  is the pressure difference between the inlet and the exit sections,  $\rho$  is the gas density, and  $V_{in}$  is the average inlet velocity [24].

The Euler number can be modeled using different models. Elsayed [24] recommends the Ramachandran et al. [19] model for prediction of the Euler number. It is reported that the Ramachandran model is superior to Shepherd and Lapple [31], and Barth [32] models and results in a better agreement with experimental results. Consequently, this study considers only the Ramachandran model to estimate the Euler number.

The Stokes number [33]  $Stk_{50}$  is a dimensionless representation of cutoff diameter [24],

$$Stk_{50} = \rho_p x_{50}^2 V_{in} / (18\mu D), \tag{2}$$

where  $\rho_p$  is the particle density and  $\mu$  is the gas viscosity.

The Iozia and Leith model [18] exhibits good agreement with experimental data, [23,24] and will be used in this study to estimate the Stokes number values.

In the calculation of the cyclone performance parameters, the following constants have been used. The barrel diameter equals  $D = 31 \times 10^{-3}$  m. This value represents a sampling cyclone (suited to a number of different applications in health-related aerosol sampling [34]). The volume flow rate is 50L/min. The air and particle densities are 1.225, 860 kg/m<sup>3</sup>, respectively. The air viscosity equals  $1.7894 \times 10^{-5}$  Pa s.

Finally, the optimization problem is formulated over  $\mathbf{x} = (a, b, D_x, H_t, h, S, B_c)$  as,

$$\begin{array}{lll} \min_{\mathbf{x}} & Eu(\mathbf{x}), & Stk_{50}(\mathbf{x}) \\ \text{such that} & 0.4 \leq a \leq 0.7, & 0.14 \leq b \leq 0.4 \\ & 0.4 \leq D_x \leq 0.75, & 3 \leq H_t \leq 7 \\ & 1 \leq h \leq 2, & 0.4 \leq S \leq 2 \\ & 0.2 \leq B_c \leq 0.4. & \end{array} \quad (3)$$

### 3 Surrogate Models

Surrogate-based optimization (SBO) methods have proven themselves to be effective in solving complex optimization problems and are increasingly being used in different fields [35–39]. SBO methods may directly solve the optimization problem (e.g., the EGO [40] or EMO [41] algorithms) or may train a surrogate model to be used in lieu of expensive simulators with traditional optimization algorithms. This work concerns the second scenario: training a surrogate model to replace an expensive simulator.

Evolutionary multi-objective algorithms like NSGA-II [42], SMS-EMOA [43] and SPEA2 [44] are popular but typically consume a large number of function evaluations to converge. This limits their use in computationally expensive multi-objective optimization problems. A possible solution is to use the previously trained cheap-to-evaluate surrogate model instead of the expensive simulator.

Popular surrogate model types include Kriging, radial basis function (RBF), support vector regression (SVR), Splines. This work uses co-Kriging models, which are explained below.

#### 3.1 Kriging

Kriging models are very popular in engineering design optimization [45]. This is partly due to the fact that Kriging models are very flexible and provide the mean and prediction variance which can be exploited by statistical sampling criteria. Their popularity also stems from the fact that many implementations are widely available [46–49].

Assume a set of  $n$  samples  $X = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$  in  $d$  dimensions having the target values  $\mathbf{y} = (y_1, \dots, y_n)'$ . A Kriging model consists of a regression function  $h(x)$  and a centered Gaussian process  $Z$ , having variance  $\sigma^2$  and a correlation matrix  $\Psi$ ,

$$Y(\mathbf{x}) = h(\mathbf{x}) + Z(\mathbf{x}). \quad (4)$$

The  $n \times n$  correlation matrix  $\Psi$  of the Gaussian process is defined as,

$$\Psi = \begin{pmatrix} \psi(\mathbf{x}_1, \mathbf{x}_1) & \dots & \psi(\mathbf{x}_1, \mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ \psi(\mathbf{x}_n, \mathbf{x}_1) & \dots & \psi(\mathbf{x}_n, \mathbf{x}_n) \end{pmatrix}, \tag{5}$$

with  $\psi(\mathbf{x}_i, \mathbf{x}_j)$  being the correlation function having hyperparameters  $\theta$ . The correlation function greatly affects the accuracy of the Kriging model used for experiments, the smoothness of the model and the differentiability of the surface. This work uses the squared exponential (SE) correlation function,

$$\psi(\mathbf{x}_a, \mathbf{x}_b) = a^2 \exp\left(-0.5\|\mathbf{x}_a - \mathbf{x}_b\|^2/l^2\right),$$

with  $l = \sqrt{\sum_{i=1}^d \theta_i (x_a^i - x_b^i)^2}$ . The hyperparameters  $\theta$  are obtained using maximum likelihood estimation (MLE).

The prediction mean and prediction variance of Kriging are then derived, respectively, as,

$$\hat{y}(\mathbf{x}) = \alpha + r(\mathbf{x}) \cdot \Psi^{-1} \cdot (\mathbf{y} - \mathbf{1}\alpha), \tag{6}$$

$$s^2(\mathbf{x}) = \sigma^2 \left( 1 - r(\mathbf{x})\Psi^{-1}r(\mathbf{x})^\top + \frac{(1 - \mathbf{1}^\top\Psi^{-1}r(\mathbf{x})^\top)}{\mathbf{1}^\top\Psi^{-1}\mathbf{1}} \right), \tag{7}$$

where  $\mathbf{1}$  is a vector of ones,  $\alpha$  is the coefficient of the constant regression function, determined by generalized least squares (GLS),  $r(\mathbf{x})$  is a  $1 \times n$  vector of correlations between the point  $\mathbf{x}$  and the samples  $X$ , and  $\sigma^2 = \frac{1}{n}(\mathbf{y} - \mathbf{1}\alpha)^\top\Psi^{-1}(\mathbf{y} - \mathbf{1}\alpha)$  is the variance.

### 3.2 Recursive Co-Kriging

An autoregressive co-Kriging formulation was proposed by Kennedy and O’Hagan [50]. Le Gratiet [51] reformulated the standard autoregressive co-Kriging (introduced by Kennedy and O’Hagan) in a recursive way. This recursive co-Kriging model scales better with the increasing number of fidelities and is used in this work.

Let there be two fidelities of data available—cheap and expensive. The objective is to combine predictions from Kriging models trained using each of the two datasets in order to maximize available training data and increase the accuracy of predictions. The rationale behind recursive co-Kriging is that the two separately trained Kriging models (and indeed the two datasets) are correlated, and there exists a *scaling factor*  $\gamma$  that can be used to calibrate the predictions from the two Kriging models.

Let  $\hat{y}_c(\mathbf{x})$  and  $\hat{y}_e(\mathbf{x})$  be the function estimates from Kriging models trained using cheap and expensive data  $(X_c, \mathbf{y}_c)$  and  $(X - e, \mathbf{y}_e)$ , respectively. Let  $\hat{y}_d(\mathbf{x})$  be the estimate from a Kriging model trained from residuals of the scaled cheap and expensive data  $(X_e, \mathbf{y}_d = \mathbf{y}_e - \gamma\hat{y}_c(X_e))$ . Then the co-Kriging estimate  $\hat{y}_{CoK}(\mathbf{x})$  is defined regressively as,



$$\hat{y}_{CoK}(\mathbf{x}) = \gamma \hat{y}_c(\mathbf{x}) + \hat{y}_d(\mathbf{x}), \quad (8)$$

where  $\gamma$  is the scaling hyper-parameter estimated using MLE.

The following section describes the training algorithm used to train a triple-fidelity recursive co-Kriging model using the method described above.

## 4 The Training Algorithm

Let  $X_A$ ,  $X_{CFD}$  and  $X_{Exp}$  represent the three fidelities of available data, namely data from analytical models, existing data available from CFD simulations and existing experimental data, respectively. The following subsections explain how each fidelity of data is used for training the recursive co-Kriging model.

### 4.1 Analytical Models

A one-shot Kriging model  $\mathcal{K}_A$  is trained on the dataset  $(X_A, \mathbf{y}_A)$  using the GPatt framework [52]. The learning process corresponding to the formulation of Kriging described in the previous section involves a complexity of  $\mathcal{O}(n^3)$  for  $n$  training instances. The computationally expensive part requires the calculation of the inverse of a correlation matrix  $\Psi$ . This usually involves computing the Cholesky decomposition of  $\Psi$  requiring  $\mathcal{O}(n^3)$  operations and  $\mathcal{O}(n^2)$  storage. This limits the applicability of the usual formulation of Kriging to relatively smaller datasets.

Assuming data are present on a  $N_s^d$  grid, GPatt exploits the grid structure to represent the kernel as a Kronecker product of  $d$  matrices. This allows exact inference and hyper-parameter optimization in  $\mathcal{O}(dn^{\frac{d+1}{d}})$  time with  $\mathcal{O}(dn^{\frac{2}{d}})$  storage (for  $d > 1$ ).

Let  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_d$  be multidimensional inputs along a Cartesian grid with  $\mathbf{x} \in \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_d$ . Assuming a product correlation function,

$$\psi(\mathbf{x}_a, \mathbf{x}_b) = \prod_{p=1}^d \psi(\mathbf{x}_a^{(p)}, \mathbf{x}_b^{(p)}), \quad (9)$$

the  $n \times n$  correlation matrix  $\Psi$  can be represented by a Kronecker product  $\Psi = \Psi_1 \otimes \Psi_2 \otimes \dots \otimes \Psi_d$ . This allows efficient calculation of the eigendecomposition of  $\Psi = QVQ^\top$  by separately calculating the eigendecomposition of each of  $\Psi_1, \Psi_2, \dots, \Psi_d$ . Similarly, Kronecker structure can also be exploited to allow fast computation of matrix vector products [53]. For a detailed explanation, the reader is referred to Wilson et al. [52].

Since mathematical or analytical models are cheap to evaluate and allow the freedom of choice of parameter combinations, it makes them an ideal choice to use with the GPatt framework. To this end, a Gaussian process (GP) model is trained on parameter combinations lying on a  $7D$  grid of  $N_s$  points. This essentially means that  $N_s$  points span the range of each parameter, and the grid consists of all possible combinations of input parameters, each consisting of  $N_s$  points. This results in a total of  $(N_s)^7$  points

comprising the grid and form the set of samples  $X_A$ . The analytical model(s) are used to evaluate  $X_A$  to obtain  $y_A$ . The model  $\mathcal{K}_A$  provides estimates as  $\hat{y}_A(\mathbf{x})$ .

### 4.2 CFD Data

Several validated CFD simulations results conducted by Elsayed and Lacor [14] have been used in the fitting of the multi-fidelity surrogate models in this study. The available CFD data  $X_{CFD}$  consist of  $n = 33$  points with four of seven dimensions varying ( $X_1$  to  $X_4$ , corresponding to the inlet width and height, the vortex finder diameter and the total cyclone height), while the remaining are fixed at  $h = 1.5$ ,  $S = 0.5$ ,  $B_c = 0.3750$ , respectively. Therefore, only a 4D Kriging model  $\mathcal{K}_{CFD}$  was trained on the CFD data using the ooDACE toolbox [48]. The 4D model is extended to 7D by concatenating the hyperparameters of the three fixed dimensions from the model  $\mathcal{K}_A$  learned above. Model training and description follows the explanation in Sect. 3.1.

Let  $\hat{y}_A(\mathbf{x})$  be the estimates obtained on a given point  $\mathbf{x}$  using the model  $\mathcal{K}_A$ . Let  $\hat{y}_{d_1}$  be the estimate from the Kriging model trained from residuals of analytical and CFD data scaled using parameter  $\gamma_1$ , ( $X_{CFD}, \mathbf{y}_{d_1} = y_{CFD} - \gamma_1 \hat{y}_A(X_{CFD})$ ). Then estimates on the point  $\mathbf{x}$  are obtained recursively as,

$$\hat{y}_{CFD}(\mathbf{x}) = \gamma_1 \hat{y}_A(\mathbf{x}) + \hat{y}_{d_1}(\mathbf{x}), \tag{10}$$

where  $\gamma_1$  is the scaling parameter described in Sect. 3.2.

### 4.3 Experimental Data

Experimental data  $X_{Exp}$  of  $n = 96$  points available in the literature for the Euler number values [16] are used to train the Kriging model  $\mathcal{K}_{Exp}$  using the ooDACE toolbox [48]. The model itself is trained as explained in Sect. 3.1.

However, objective function values for only one objective (the Euler number) are available. Consequently, a double-fidelity co-Kriging model is trained in lieu of a triple-fidelity co-Kriging model for the second objective (the Stokes number).

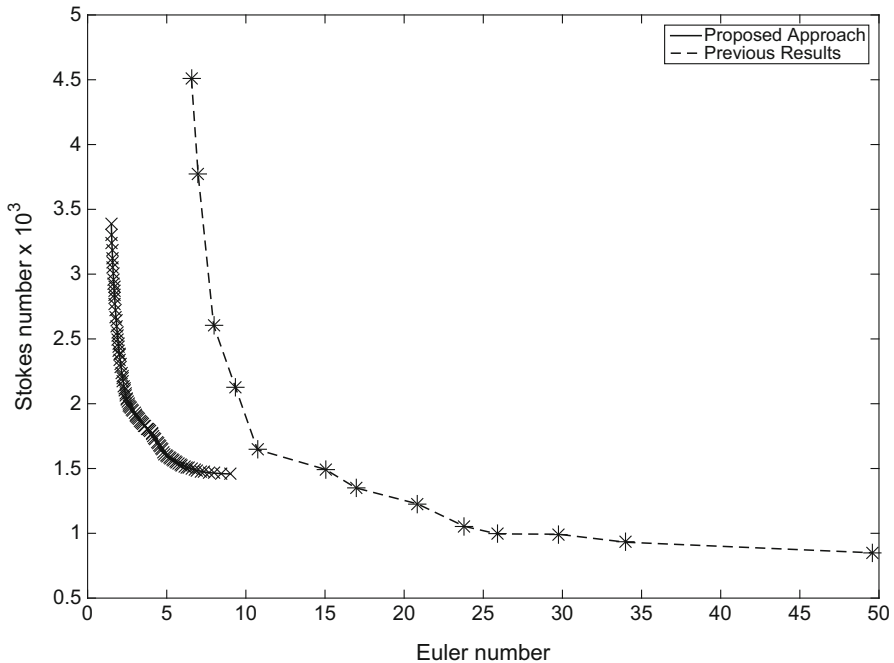
Let  $\gamma_2$  define the scaling between CFD and experimental data. Let  $\hat{y}_{d_2}$  be the estimate from the Kriging model trained from residuals of scaled CFD and experimental data ( $X_{Exp}, \mathbf{y}_{d_2} = y_{Exp} - \gamma_2 \hat{y}_{CFD}(X_{Exp})$ ). The estimate of the model  $\mathcal{K}_{Exp}$  on the point  $\mathbf{x}$  is given by,

$$\hat{y}_{Exp}(\mathbf{x}) = \gamma_2 \hat{y}_{CFD}(\mathbf{x}) + \hat{y}_{d_2}(\mathbf{x}). \tag{11}$$

The process of estimation of scaling parameters  $\gamma_1$  and  $\gamma_2$  is described below.

### 4.4 Estimation of the Scaling Parameter

The scaling parameters  $\gamma_1$  and  $\gamma_2$  are estimated using cross-validation using the scheme described by Le Gratiet [51]. Essentially, the cross-validation procedure was performed in two stages. The first stage involved cross-validation the double-fidelity co-Kriging model trained using analytical models and CFD data, with CFD data divided into



**Fig. 2** The Pareto set obtained using the SMS-EMOA algorithm

partitions for training and testing. This yielded the value of  $\gamma_1$  to be used for scaling between the sub-Kriging models trained using CFD and analytical data. Similarly, the second stage involved partitioning the experimental data into folds, while the sub-Kriging models obtained using CFD data and analytical models were not modified. This yielded the value of  $\gamma_2$  for scaling between the sub-Kriging models trained using experimental data, and CFD data.

## 5 Numerical Settings

For the purpose of obtaining  $X_A$ , the number of points in each dimension of the grid  $N_s = 7$ . This results in  $n = 7^7 = 823,543$  training points in  $X_A$  for  $\mathcal{K}_A$ . The recursive co-Kriging model is used as the objective function to be optimized using the SMS-EMOA algorithm [43]. The size of the initial population is set to 100 individuals allowed to evolve until the total number of objective function evaluations reaches 10,000.

## 6 Results and Discussion

The SMS-EMOA algorithm evolved the population of size 100 into a Pareto set of 100 solutions, which are shown in Fig. 2 and listed in Table 1. The solutions are well-spread and show an improvement over past approaches. A comparison with previous results obtained using multi-objective surrogate-based optimization with the EMO algorithm

**Table 1** Pareto-optimal solutions obtained from SMS-EMOA

$a$	$b$	$D_x$	$H_t$	$h$	$S$	$B_c$	$E_u$	$\log(Stk_{50} \times 10^3)$
0.40004	0.15715	0.75	4.0001	1.3557	0.40738	0.2	0.40443	1.22
0.4	0.16176	0.73095	4	1.3189	0.40838	0.2	0.41671	1.1955
0.4	0.15256	0.72908	3.9999	1.0695	0.4	0.20581	0.42892	1.1761
0.40007	0.15135	0.72449	4	1.0813	0.44006	0.22046	0.44005	1.1593
0.40003	0.1606	0.71554	4	1.6365	0.4	0.20269	0.45397	1.1375
0.40013	0.15861	0.71143	4.0001	1.4385	0.4	0.2	0.46092	1.1283
0.40001	0.15592	0.71123	4	1.801	0.40726	0.20985	0.46782	1.117
0.40018	0.16281	0.70747	4	1.38	0.4	0.21525	0.48541	1.1001
0.4	0.15797	0.70472	4	1.661	0.4038	0.21363	0.49291	1.0833
0.40003	0.14595	0.70693	4	1.6023	0.41446	0.20931	0.50501	1.0749
0.40007	0.16209	0.70053	3.9999	1.538	0.40965	0.20601	0.51229	1.0647
0.40018	0.15611	0.69984	4.0002	1.6118	0.45373	0.22077	0.51963	1.0542
0.40041	0.14954	0.6998	4	1.4998	0.43039	0.20327	0.52814	1.0446
0.40024	0.15996	0.69533	3.9999	1.475	0.41972	0.20315	0.53618	1.035
0.40012	0.15655	0.69294	4.0001	1.4752	0.46542	0.21307	0.54647	1.02
0.40004	0.15839	0.68958	4.0002	1.1425	0.4	0.2	0.56708	1.001
0.4	0.15525	0.68588	4.0001	1.1711	0.41902	0.21259	0.58206	0.98043
0.4	0.14859	0.68588	4.0001	1.1499	0.4	0.23004	0.59408	0.97093
0.40005	0.14906	0.68226	4	1.1343	0.42343	0.20589	0.61056	0.95322
0.40005	0.16184	0.67597	4.0002	1.5304	0.474	0.22231	0.62878	0.93647
0.40027	0.15635	0.6749	4	1.38	0.4	0.2	0.63374	0.92651
0.40008	0.15688	0.67273	3.9999	1.1197	0.4173	0.21683	0.64742	0.91514
0.40005	0.15407	0.67064	3.9999	1.1498	0.42418	0.21335	0.65941	0.90267

Table 1 continued

$a$	$b$	$D_x$	$H_t$	$h$	$S$	$B_c$	$E_u$	$\log(\text{Stk}_{50} \times 10^3)$
0.40004	0.16168	0.66696	3.9999	1.274	0.42074	0.21847	0.67041	0.89324
0.40009	0.15928	0.66452	4	1.2249	0.42003	0.21288	0.68274	0.88017
0.40006	0.15688	0.66263	4	1.3032	0.4	0.2196	0.69212	0.87013
0.40027	0.15592	0.66237	4.0003	1.5304	0.4	0.20717	0.70577	0.86465
0.4	0.16123	0.65704	4.0002	1.2293	0.42044	0.2	0.71954	0.84867
0.4002	0.15239	0.65539	3.9999	1.2072	0.40431	0.20785	0.73378	0.83488
0.40012	0.16411	0.65082	4.0002	1.5387	0.4	0.2	0.74789	0.82593
0.40001	0.1598	0.64535	3.9999	1.2958	0.4	0.20444	0.76988	0.80381
0.40012	0.15847	0.64324	4.0002	1.2126	0.4	0.2	0.7824	0.79492
0.40025	0.15121	0.64351	3.9997	1.2054	0.43532	0.2	0.79647	0.78854
0.4	0.16211	0.63642	3.9999	1.2628	0.41617	0.20168	0.80622	0.77609
0.4002	0.15479	0.63216	4	1.2557	0.40836	0.20682	0.82862	0.75936
0.4	0.15279	0.62981	4.0001	1.3806	0.4	0.2	0.83879	0.75208
0.40006	0.16194	0.62471	4.0001	1.3069	0.41982	0.22455	0.85032	0.74459
0.40003	0.15936	0.61951	3.9999	1.2616	0.40712	0.2	0.86871	0.73182
0.40004	0.15839	0.61648	4.0002	1.1425	0.4	0.2	0.88222	0.72462
0.40001	0.1598	0.60891	3.9999	1.2958	0.4	0.20444	0.90043	0.71304
0.40007	0.16123	0.60489	4.0001	1.5279	0.40284	0.2185	0.91061	0.70797
0.4	0.16357	0.59856	3.9999	1.1442	0.4216	0.22549	0.92762	0.70039
0.40012	0.1545	0.59395	4.0002	1.1656	0.4	0.2	0.94502	0.69318
0.40025	0.14679	0.58965	3.9998	1.4753	0.4	0.2	0.96532	0.68718
0.4	0.14	0.58417	3.9999	1.0717	0.42783	0.2	0.98963	0.68101
0.40021	0.14	0.58476	3.9994	1.0042	0.43841	0.2	1.0112	0.6767

**Table 1** continued

$a$	$b$	$D_x$	$H_t$	$h$	$S$	$B_c$	$E_u$	$\log(Stk_{50} \times 10^3)$
0.4	0.14	0.58286	3.9991	1.289	0.4	0.2	1.0374	0.6714
0.40005	0.20357	0.56283	4.0002	1.2293	0.41505	0.2	1.0549	0.66669
0.4	0.21929	0.56825	3.9998	1.0045	0.4216	0.22957	1.0865	0.6576
0.40004	0.22392	0.56063	4.0001	1.2495	0.4	0.2	1.1076	0.65059
0.4	0.22912	0.55877	4	1.3393	0.40305	0.21076	1.1239	0.64576
0.4	0.23108	0.55709	3.9998	1	0.42054	0.22023	1.1366	0.6423
0.40007	0.23777	0.55493	3.9999	1.1069	0.41807	0.21599	1.1558	0.63644
0.40003	0.24418	0.55504	3.9999	1.0875	0.41106	0.20291	1.1746	0.6314
0.40005	0.24314	0.54104	3.9999	1.1069	0.41807	0.20293	1.1943	0.62501
0.40079	0.24544	0.53749	4.0004	1	0.4	0.2221	1.2187	0.6197
0.4	0.24949	0.53355	4.0004	1.321	0.40356	0.21437	1.2334	0.61527
0.40025	0.25762	0.53796	3.9999	1.2128	0.4	0.20726	1.2503	0.61114
0.40005	0.25738	0.52501	3.9999	1.0748	0.42343	0.20234	1.2751	0.60441
0.40025	0.26041	0.52832	4.0006	1.3603	0.4	0.20198	1.2941	0.60171
0.4	0.2611	0.51079	3.9999	1.0276	0.42903	0.21024	1.3258	0.59267
0.4	0.26416	0.5104	4	1.5491	0.41018	0.21531	1.3391	0.59018
0.4	0.2496	0.48731	4	1	0.4	0.22096	1.3602	0.58459
0.40002	0.26053	0.49385	3.9999	1.3227	0.40817	0.21715	1.3762	0.58079
0.4001	0.15939	0.43287	4	1.3031	0.42918	0.20511	1.3994	0.57171
0.40012	0.14	0.42908	3.9999	1.4098	0.40175	0.20085	1.4091	0.56741
0.4	0.14462	0.42621	4.0001	1.0434	0.4	0.20454	1.4253	0.55881
0.40007	0.14688	0.42384	4	1	0.43116	0.20628	1.4399	0.55172
0.40026	0.14	0.42047	3.9999	1.1442	0.4216	0.22549	1.4559	0.54453
0.4	0.14136	0.41679	4.0001	1.078	0.4	0.20369	1.4758	0.53403

Table 1 continued

$a$	$b$	$D_x$	$H_t$	$h$	$S$	$B_c$	$E_u$	$\log(\text{Stk}_{50} \times 10^3)$
0.40001	0.14	0.41406	4	1	0.42626	0.24685	1.4909	0.52681
0.40014	0.14129	0.41193	3.9999	1.1316	0.41222	0.2	1.5031	0.52026
0.40002	0.1477	0.41022	3.9999	1.3385	0.40554	0.2	1.518	0.51239
0.40009	0.14744	0.40779	3.9999	1.0218	0.4	0.20041	1.5321	0.50511
0.4	0.14	0.40431	4.0003	1.4764	0.41515	0.2	1.5459	0.49834
0.4	0.14	0.4	3.9999	1.305	0.4	0.20013	1.5694	0.48497
0.40009	0.15127	0.4	4	1.3358	0.4031	0.20149	1.5839	0.47802
0.4	0.16111	0.4	4.0001	1.2489	0.42174	0.20191	1.6012	0.47139
0.40003	0.16976	0.4	4.0001	1.0461	0.40668	0.20219	1.6197	0.46523
0.4	0.17713	0.4	4	1.1216	0.4	0.2	1.6381	0.45954
0.4	0.18395	0.4	3.9999	1.2628	0.41617	0.20254	1.6574	0.45407
0.4	0.19009	0.4	4	1	0.41098	0.20254	1.6747	0.44935
0.40005	0.19585	0.4	4.0001	1.2423	0.40407	0.21815	1.6936	0.44446
0.40002	0.20161	0.4	4.0001	1.0588	0.40399	0.204	1.7111	0.4399
0.4	0.20655	0.4	4	1.0671	0.4	0.20265	1.7274	0.43584
0.4	0.21274	0.4	4	1	0.41098	0.20254	1.748	0.43092
0.40002	0.21729	0.4	4	1.1848	0.42279	0.20841	1.7649	0.42703
0.40005	0.2229	0.4	4	1.2989	0.4	0.21815	1.785	0.42253
0.40002	0.23178	0.4	3.9998	1.377	0.4	0.2	1.816	0.41597
0.4	0.23788	0.4	4	1	0.4	0.20735	1.8362	0.41189
0.4	0.24305	0.4	3.9999	1.2628	0.41617	0.20168	1.8568	0.40798

Table 1 continued

$a$	$b$	$D_x$	$H_t$	$h$	$S$	$B_c$	$E_u$	$\log(Stk_{50} \times 10^3)$
0.40004	0.25018	0.4	3.9999	1.161	0.41413	0.20032	1.8824	0.40355
0.4	0.25712	0.4	3.9999	1.305	0.4	0.20013	1.9092	0.39923
0.40025	0.26275	0.4	4	1.3507	0.41727	0.20818	1.9323	0.39599
0.4	0.27122	0.4	3.9999	1.4972	0.42614	0.2	1.9656	0.39174
0.40015	0.27981	0.4	4	1.2286	0.40738	0.2	1.9985	0.38811
0.40022	0.28864	0.4	4	1.1709	0.41328	0.2	2.0344	0.3849
0.40019	0.29787	0.4	3.9999	1.2919	0.46601	0.2	2.0746	0.38196
0.4	0.3115	0.4	4.0001	1.5041	0.4	0.21113	2.131	0.37884
0.40012	0.32981	0.4	4.0001	1.4847	0.46542	0.22444	2.2025	0.3772



**Table 2** Error estimates obtained using cross-validation

Model	MSE		BEEQ	
	<i>Eu</i>	<i>Stk</i> <sub>50</sub>	<i>Eu</i>	<i>Stk</i> <sub>50</sub>
Double-fidelity: CFD + analytical	3.1254	0.0773	3.1393	0.8637
Triple-fidelity: experimental + CFD + analytical	0.6699	–	0.6776	–

[12] can also be seen in Fig. 2. Admittedly, the comparison is not completely fair since the previous approach only utilized data from analytical models, while the proposed approach incorporated experimental and simulated data in addition to data from an analytical model. Nevertheless, the results confirm that a more accurate surrogate model for the optimization problem is obtained when all available data are incorporated.

Since the trained model is used as a surrogate model to be treated as the objective function by SMS-EMOA, its accuracy is crucial. Toward this end, cross-validation of the multi-fidelity recursive co-Kriging model was performed using the scheme described in Sect. 4.4.

Table 2 lists the mean squared error (MSE) and Bayesian estimation error quotient (BEEQ) estimates obtained using cross-validation, as described above. BEEQ [54] quantifies the improvement of error of a Bayesian estimator  $\hat{y}$ , over the prior mean  $\bar{y}$ , or of the updated estimate  $\hat{y}$  of a recursive estimator over the predicted estimate  $\hat{y}$ . It is defined as  $BEEQ(\hat{y}) = \left( \prod_{i=1}^n \beta_i \right)^{1/n}$ , where:

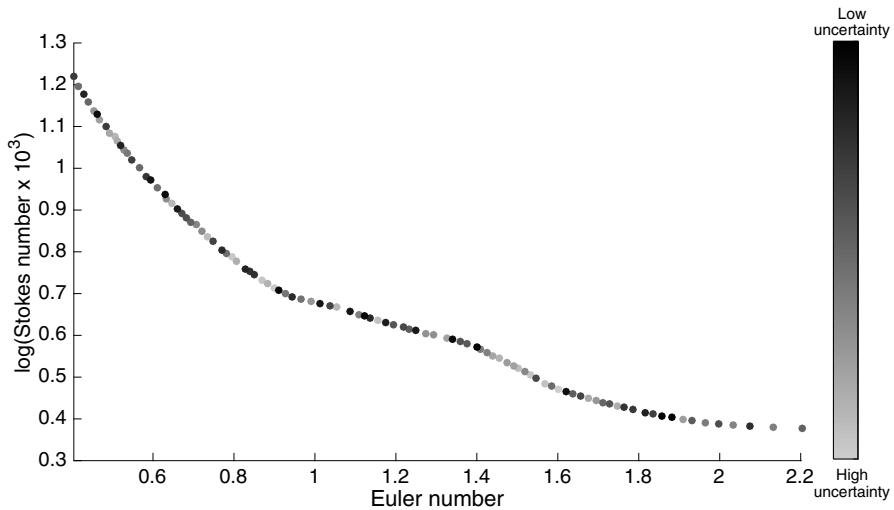
$$\beta_i = \frac{\|y_i - \hat{y}(\mathbf{x}_i)\|}{\|y_i - \bar{y}(\mathbf{x}_i)\|}.$$

The closer the value of BEEQ to 0, the better the model. BEEQ brings the advantage of nullifying the effect of very large or small magnitudes of values on error estimates.

The double-fidelity recursive co-Kriging model obtained using CFD data and analytical models is very accurate for the Stokes number with MSE=0.0773 and BEEQ=0.8637. Considering the fact that the absolute range,

$$(\max(\mathbf{y}_{\text{CFD}}) - \min(\mathbf{y}_{\text{CFD}})),$$

for Stokes number is 1.8765, a mean error of  $\sqrt{0.0073} = 0.0854$  is a very good result. Similarly, the absolute range for the Euler number is 3.1139. A mean variation of  $\sqrt{3.1254} = 1.7679$  units and a BEEQ score of 3.1393 is not an excellent result, but considering the model is a 7D model trained using only 33 data points, the results are acceptable. Moreover, using the sub-Kriging model trained from experimental data improves the accuracy of the final triple-fidelity model for the Euler number. The absolute range of the Euler number in the experimental dataset is 4.2124. A mean variation of  $\sqrt{0.6699} = 0.8185$  and a BEEQ score of 0.6776 is a much better result compared to using only the double-fidelity co-Kriging model and serves to capture the general trend of the problem under consideration.



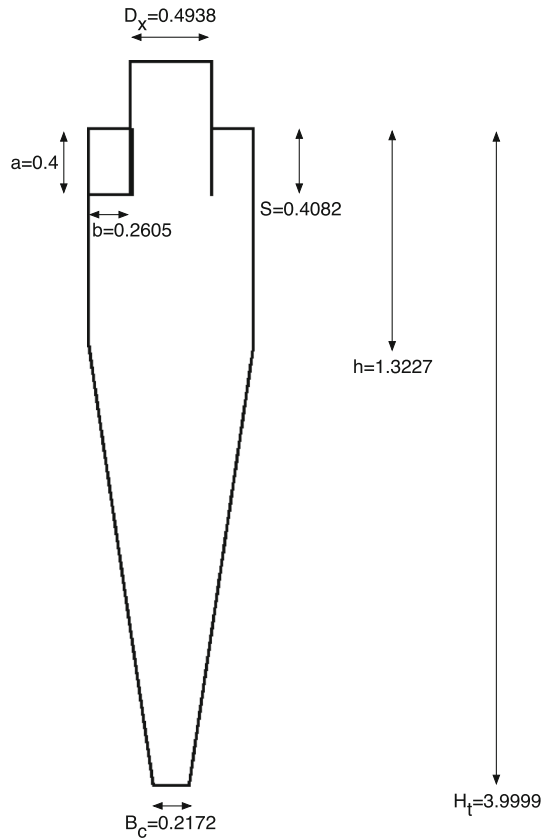
**Fig. 3** The resulting Pareto set. Each Pareto-optimal solution corresponds to a point in the figure representing the mean of variance (uncertainty) of prediction for both objectives, according to the co-Kriging models

Another advantage of using the co-Kriging model is that it can be used to obtain the variance of prediction of all Pareto-optimal points. Figure 3 shows the variance of prediction for each Pareto-optimal solution as an ellipse. The major axis of a particular ellipse corresponds to the variance of prediction for the Euler number, and the minor axis corresponds to the variance of prediction for the Stokes number. The availability of such information can aid the practitioner in selecting a particular Pareto-optimal solution over another, since it may be impractical to validate all Pareto-optimal solutions using CFD simulations, or experiments. A representative solution is shown in Fig. 4. The solution having the least variance of prediction obtained using the co-Kriging surrogate model was selected to be evaluated as a potential design, and plotted.

## 7 Conclusions

The problem of cyclone shape optimization involves computationally expensive simulations, and data from time-consuming experiments. Existing approaches do not take multi-fidelity data into account. This work proposes a novel approach to combine existing data from experiments, CFD simulations and mathematical models to train a tripe-fidelity recursive co-Kriging surrogate model. The novelties include handling missing output dimensions and performing co-Kriging with sub-Kriging models trained using million-point datasets. The surrogate model is used as a replacement for the expensive data in a conventional optimization algorithm like SMS-EMOA. Experimental results indicate that the proposed approach leverages the advantages of multiple fidelities of data to result in an accurate surrogate model which leads to a good Pareto set of possible solutions. The accurate surrogate model allows thousands of evaluations to be computed cheaply, enabling optimization algorithms like SMS-EMOA to find

**Fig. 4** Cyclone shape corresponding to a representative Pareto-optimal solution chosen according to minimum variance of prediction of the co-Kriging surrogate model. The design parameters  $(a, b, D_x, H_t, h, S, B_c) = (0.4, 0.2605, 0.4938, 3.9999, 1.3227, 0.4082, 0.2172)$



high-quality solutions in a very short time. Comparison of the proposed approach to state-of-the-art approaches shows substantial improvement in the obtained solutions.

**Acknowledgements** This work was supported by the Interuniversity Attraction Poles Programme BEST-COM initiated by the Belgian Science Policy Office, and the Research Foundation Flanders (FWO-Vlaanderen). Ivo Couckuyt is a Postdoctoral Researcher of FWO-Vlaanderen.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix A

### A.1 Ramachandran Model

The Ramachandran et al. [19] model has been constructed through a statistical analysis. The Euler number  $Eu$  is given,

$$Eu = 20 \left\{ \frac{a b}{D_x^2} \right\} \left\{ \frac{\frac{S}{D}}{\frac{H}{D} \frac{h}{D} \frac{B_c}{D}} \right\}^{1/3} \tag{12}$$

### A.2 The Iozia and Leith Model

The Iozia and Leith model [18] is based on the equilibrium-orbit theory (Force balance).

$$x_{50} = \left\{ (9 \mu Q) / \left( \pi H_{CS} \rho_p V_{\theta_{max}}^2 \right) \right\}^{1/2} \tag{13}$$

where  $\rho_p$  is the particle density..

Where  $H_{CS}$  is the core height and  $V_{\theta_{max}}$  is the maximum tangential velocity that occurs at the edge of the control surface CS. In this model, the values of the core diameter  $d_c$  and the tangential velocity at the core edge  $V_{\theta_{max}}$  are calculated via a regression of the experimental data as follows,

$$V_{\theta_{max}} = 6.1 v_{in} \left\{ \left( ab/D^2 \right)^{-0.61} \left( D_x/D \right)^{-0.74} \left( H_t/D \right)^{-0.33} \right\} \tag{14}$$

$$d_c = 0.52 D \left( ab/D^2 \right)^{-0.25} \left( D_x/D \right)^{1.53} \tag{15}$$

### References

1. Funk, P., Elsayed, K., Yeater, K., Holt, G., Whitelock, D.: Could cyclone performance improve with reduced inlet velocity? *Powder Technol.* **280**, 211–218 (2015)
2. Brar, L.S., Sharma, R., Elsayed, K.: The effect of the cyclone length on the performance of Stairmand high-efficiency cyclone. *Powder Technol.* **286**, 668–677 (2015)
3. Stairmand, C.J.: The design and performance of cyclone separators. *Ind. Eng. Chem.* **29**, 356–383 (1951)
4. Hu, L.Y., Zhou, L.X., Zhang, J., Shi, M.X.: Studies on strongly swirling flows in the full space of a volute cyclone separator. *AIChE J.* **51**(3), 740–749 (2005)
5. Zhang, B., Hui, S.: Numerical simulation and PIV study of the turbulent flow in a cyclonic separator. In: *International Conference on Power Engineering*, Hangzhou, China (2007)
6. Gimbin, J., Chuah, T., Fakhru'l-Razi, A., Choong, T.S.Y.: The influence of temperature and inlet velocity on cyclone pressure drop: a CFD study. *Chem. Eng. Process.* **44**(1), 7–12 (2005)
7. Elsayed, K., Lacor, C.: Optimization of the cyclone separator geometry for minimum pressure drop using mathematical models and CFD simulations. *Chem. Eng. Sci.* **65**(22), 6048–6058 (2010)
8. Elsayed, K., Lacor, C.: Analysis and optimisation of cyclone separators geometry using RANS and LES methodologies. In: Deville, M.O., Estivaleres, J.L., Gleize, V., Le, T.H., Terracol, M., Vincent, S. (eds.) *Turbulence and Interactions, Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, vol. 125, pp. 65–74. Springer, Berlin (2014)
9. Elsayed, K., Lacor, C.: The effect of cyclone vortex finder dimensions on the flow pattern and performance using LES. *Comput. Fluids* **71**, 224–239 (2013)
10. Elsayed, K., Lacor, C.: The effect of the dust outlet geometry on the performance and hydrodynamics of gas cyclones. *Comput. Fluids* **68**, 134–147 (2012)
11. Elsayed, K., Lacor, C.: The effect of cyclone inlet dimensions on the flow pattern and performance. *Appl. Math. Model.* **35**(4), 1952–1968 (2011)
12. Singh, P., Couckuyt, I., Elsayed, K., Deschrijver, D., Dhaene, T.: Shape optimization of a cyclone separator using multi-objective surrogate-based optimization. *Appl. Math. Modell.* **40**(5), 4248–4259 (2015)

13. Elsayed, K.: Optimization of the cyclone separator geometry for minimum pressure drop using co-Kriging. *Powder Technol.* **269**, 409–424 (2015)
14. Elsayed, K., Lacor, C.: CFD modeling and multi-objective optimization of cyclone geometry using desirability function, artificial neural networks and genetic algorithms. *Appl. Math. Model.* **37**(8), 5680–5704 (2013)
15. Elsayed, K., Lacor, C.: Modeling and pareto optimization of gas cyclone separator performance using RBF type artificial neural networks and genetic algorithms. *Powder Technol.* **217**, 84–99 (2012)
16. Elsayed, K., Lacor, C.: Modeling, analysis and optimization of aircyclones using artificial neural network, response surface methodology and CFD simulation approaches. *Powder Technol.* **212**(1), 115–133 (2011)
17. Elsayed, K.: Design of a novel gas cyclone vortex finder using the adjoint method. *Sep. Purif. Technol.* **142**, 274–286 (2015)
18. Iozia, D.L., Leith, D.: Effect of cyclone dimensions on gas flow pattern and collection efficiency. *Aerosol Sci. Technol.* **10**(3), 491–500 (1989)
19. Ramachandran, G., Leith, D., Dirgo, J., Feldman, H.: Cyclone optimization based on a new empirical model for pressure drop. *Aerosol Sci. Technol.* **15**, 135–148 (1991)
20. Trefz, M.: Die verschiedenen abscheidevorgänge im hoher un hoch beladenen gaszyklon unter besonderer berucksichtigung der sekundarstromung. *Fortschritt-Berichte VDI; VDI-Verlag GmbH: Dusseldorf, Germany* **295** (1992)
21. Trefz, M., Muschelknautz, E.: Extended cyclone theory for gas flows with high solids concentrations. *Chem. Ing. Tech.* **16**, 153–160 (1993)
22. Cortés, C., Gil, A.: Modeling the gas and particle flow inside cyclone separators. *Prog. Energy Combust. Sci.* **33**(5), 409–452 (2007)
23. Hoffmann, A.C., Stein, L.E.: *Gas Cyclones and Swirl Tubes: Principle, Design and Operation*, 2nd edn. Springer, Berlin (2008)
24. Elsayed, K.: *Analysis and Optimization of Cyclone Separators Geometry Using RANS and LES Methodologies*. Ph.D. thesis, Vrije Universiteit Brussel (2011)
25. Ravi, G., Gupta, S.K., Ray, M.B.: Multiobjective optimization of cyclone separators using genetic algorithm. *Ind. Eng. Chem. Res.* **39**, 4272–4286 (2000)
26. Salcedo, R.L.R., Candido, M.G.: Global optimization of reverse-flow gas cyclones: application to small-scale cyclone design. *Sep. Sci. Technol.* **36**(12), 2707–2731 (2001)
27. Mothes, H., Löffler, F.: Prediction of particle removal in cyclone separators. *Int. Chem. Eng.* **28**(2), 231–240 (1988)
28. Swamee, P.K., Aggarwal, N., Bhojhiya, K.: Optimum design of cyclone separator. *Am. Inst. Chem. Eng. (AIChE)* **55**(9), 2279–2283 (2009)
29. Safikhani, H., Hajiloo, A., Ranjbar, M., Nariman-Zadeh, N.: Modeling and multi-objective optimization of cyclone separators using CFD and genetic algorithms. *Comput. Chem. Eng.* **35**(6), 1064–1071 (2011)
30. Pishbin, S.I., Moghiman, M.: Optimization of cyclone separators using genetic algorithm. *Int. Rev. Chem. Eng. (I.R.E.C.H.E.)* **2**(6), 683–690 (2010)
31. Shepherd, C.B., Lapple, C.E.: Flow pattern and pressure drop in cyclone dust collectors cyclone without intel vane. *Ind. Eng. Chem.* **32**(9), 1246–1248 (1940)
32. Barth, W.: Design and layout of the cyclone separator on the basis of new investigations. *Brennstow Wärme Kraft (BWK)* **8**(4), 1–9 (1956)
33. Derksen, J.J., Sundaresan, S., van den Akker, H.E.A.: Simulation of mass-loading effects in gas-solid cyclone separators. *Powder Technol.* **163**, 59–68 (2006)
34. Kenny, L., Gussman, R.: Characterization and modelling of a family of cyclone aerosol pre separators. *J. Aerosol Sci.* **28**(4), 677–688 (1997)
35. Badhurshah, R., Samad, A.: Multiple surrogate based optimization of a bidirectional impulse turbine for wave energy conversion. *Renew. Energy* **74**, 749–760 (2015)
36. Chen, X.M., Zhang, L., He, X., Xiong, C., Li, Z.: Surrogate-based optimization of expensive-to-evaluate objective for optimal highway toll charges in transportation network. *Comput. Aided Civ. Infrastruct. Eng.* **29**(5), 359–381 (2014)
37. Costas, M., Díaz, J., Romera, L., Hernández, S.: A multi-objective surrogate-based optimization of the crashworthiness of a hybrid impact absorber. *Int. J. Mech. Sci.* **88**, 46–54 (2014)
38. Kozziel, S., Ogurtsov, S.: Numerically Efficient Approach to Simulation-Driven Design of Planar Microstrip Antenna Arrays by Means of Surrogate-Based Optimization, pp. 149–170. Springer International Publishing, Cham (2014). doi:[10.1007/978-3-319-08985-0\\_6](https://doi.org/10.1007/978-3-319-08985-0_6)

39. Liu, B., He, Y., Reynaert, P., Gielen, G.: Global optimization of integrated transformers for high frequency microwave circuits using a Gaussian process based surrogate model. In: Design, Automation and Test in Europe Conference and Exhibition (DATE), 2011, pp. 1–6. IEEE (2011)
40. Jones, D.R., Schonlau, M., Welch, W.J.: Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **13**(4), 455–492 (1998)
41. Couckuyt, I., Deschrijver, D., Dhaene, T.: Fast calculation of multiobjective probability of improvement and expected improvement criteria for pareto optimization. *J. Glob. Optim.* **60**(3), 575–594 (2014)
42. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **6**, 182–197 (2002)
43. Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.* **181**(3), 1653–1669 (2007)
44. Zitzler, E., Laumanns, M., Thiele, L., et al.: Spea 2: improving the strength pareto evolutionary algorithm. *Eurogen* **3242**, 95–100 (2001)
45. Forrester, A.I.J., Keane, A.J.: Recent advances in surrogate-based optimization. *Prog. Aerosp. Sci.* **45**(1–3), 50–79 (2009)
46. Couckuyt, I., Koziel, S., Dhaene, T.: Surrogate modeling of microwave structures using Kriging, co-Kriging, and space mapping. *Int. J. Numer. Model. Electron. Netw. Devices Fields* **26**(1), 64–73 (2013)
47. Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T., Crombecq, K.: A surrogate modeling and adaptive sampling toolbox for computer based design. *J. Mach. Learn. Res.* **11**, 2051–2055 (2010)
48. Couckuyt, I., Dhaene, T., Demeester, P.: oodace toolbox: a flexible object-oriented Kriging implementation. *J. Mach. Learn. Res.* **15**(1), 3183–3186 (2014)
49. Lophaven, S.N., Nielsen, H.B., Søndergaard, J.: Aspects of the Matlab Toolbox DACE. Technical Report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby (2002)
50. Kennedy, M.C., O’Hagan, A.: Bayesian calibration of computer models (with discussion). *J. R. Stat. Soc. B* **63**, 425–464 (2001)
51. Le Gratiet, L., Garnier, J.: Recursive co-Kriging model for design of computer experiments with multiple levels of fidelity. *Int. J. Uncertain. Quantif.* **4**(5) (2014)
52. Wilson, A., Gilboa, E., Cunningham, J.P., Nehorai, A.: Fast multidimensional pattern extrapolation. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (eds.) *Advances in Neural Information Processing Systems*, vol. 27, pp. 3626–3634. Curran Associates, Inc. (2014). <http://papers.nips.cc/paper/5372-fast-kernel-learning-for-multidimensional-pattern-extrapolation.pdf>
53. Wilson, A.G., Nickisch, H.: Kernel interpolation for scalable structured gaussian processes (KISS-GP). In: *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6–11 July 2015*, pp. 1775–1784 (2015). <http://jmlr.org/proceedings/papers/v37/wilson15.html>
54. Li, X.R., Zhao, Z.: Relative error measures for evaluation of estimation algorithms. In: *2005 8th International Conference on Information Fusion*, vol. 1, p. 8. IEEE (2005)