Chapter 4
Multi-Objective Surrogate Based Optimization of Gas Cyclones Using Support Vector Machines and CFD Simulations

Khairy Elsayed and Chris Lacor

Abstract In order to accurately predict the complex nonlinear relationships between the cyclone performance parameters (The Euler and Stokes numbers) and the four significant geometrical dimensions (the inlet section height and width, the vortex finder diameter and the cyclone total height), the support vector machines approach has been used. Two support vector regression surrogates (SVR) have been trained and tested by CFD datasets. The result demonstrates that SVR can offer an alternative and powerful approach to model the performance parameters. The SVR model parameters have been optimized to obtain the most accurate results from the cross validation steps. SVR (with optimized parameters) can offer an alternative and powerful approach to model the performance parameters better than Kriging. SVR surrogates have been employed to study the effect of the four geometrical parameters on the cyclone performance. The genetic algorithms optimization technique has been applied to obtain a new geometrical ratio for minimum Euler number and for minimum Euler and Stokes numbers. New cyclones over-perform the standard Stairmand design performance. Pareto optimal solutions have been obtained and a new correlation between the Euler and Stokes numbers is fitted.

K. Elsayed (✉)
Department of Mechanical Engineering, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Faculty of Engineering at El-Mattaria, Mechanical Power Engineering Department, Helwan University, Masaken El-Helmia P.O., Cairo 11718, Egypt
e-mail: kelsayed@vub.ac.be; kelsayed75@gmail.com

C. Lacor
Department of Mechanical Engineering, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
e-mail: chris.lacor@vub.ac.be

E. Iuliano, E.A. Pérez (eds.), *Application of Surrogate-based Global Optimization to Aerodynamic Design*, Springer Tracts in Mechanical Engineering,
DOI 10.1007/978-3-319-21506-8_4

4.1 Introduction

Cyclone separators are widely used in gas-solid separation for aerosol sampling and industrial applications [1] where both the gravity and centrifugal force are used to separate solids from a mixture of solids and fluids. Cyclones have the following advantages. Simplicity in construction, contain no moving parts, relatively maintenance free, can handle high pressure and temperature mixtures and corrosive gases, relative economy in power consumption. Due to these advantages, cyclone separators have become one of the most important particle removal devices in both engineering and process operation [1] such as cement industry, oil and gas, coal fired boiler, workshops and vacuum cleaners.

4.1.1 Cyclone Geometry

The cyclone geometry is described by seven geometrical parameters, viz. the inlet height a and width b, the vortex finder diameter D_x and length S, the cylinder height h, the cyclone total height H_t and cone-tip diameter B_c as shown in Fig. 4.1. It has been approved in previous studies by the authors that only four geometrical

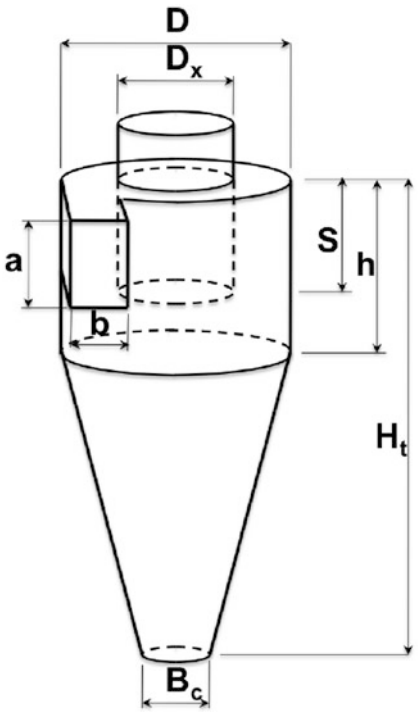

Fig. 4.1 Cyclone geometry. In this study, $h/D = 1.5, S/D = 0.5, B_c/D = 0.375$

Table 4.1 The definition of the cyclone performance parameters

The Euler number Eu	The Stokes number Stk_{50}
The Euler number is the dimensionless pressure drop ΔP. ΔP = (the area- and time-averaged static pressure at the inlet section) - (the area- and time-averaged static pressure at the gas exit section). $$Eu = \frac{\text{The pressure drop between the inlet and the gas exit}}{\text{The average kinetic energy at the inlet}}$$ $Eu = \frac{\Delta P}{\frac{1}{2}\rho V_{in}^2}$ where ρ is the gas density and V_{in} is the average inlet velocity. Eu is not affected by operating conditions in the high Reynolds number range ($R_e > 5 \times 10^4$, $Re = \frac{\rho V_{in} D}{\mu}$) [3, 10]	Stk_{50} is the dimensionless cut-off diameter x_{50}. x_{50} is the particle diameter that produces 50 % collection efficiency. $Stk_{50} = \rho_p x_{50}^2 V_{in}/(18\mu D) = \frac{\tau_p}{\tau_f}$ [11]. It is the ratio between the particle relaxation time (the time constant in the exponential decay of the particle velocity due to drag) $\tau_p = \rho_p x_{50}^2/(18\mu)$ and the gas flow integral time scale $\tau_f = D/V_{in}$ where ρ_p is the particle density and μ is the gas viscosity

parameters significantly affect the cyclone flow pattern and performance [2–9]. The four significant factors are the inlet section height a and width b, the vortex finder diameter D_x and the cyclone total height H_t.

4.1.2 Cyclone Performance

Besides the separation efficiency (or alternatively, the cut-off diameter for low mass particle loading), pressure drop is another major index for cyclone performance evaluation. Therefore, it is necessary to obtain an accurate model to determine the complex relationship between the performance parameters and the cyclone characteristics. Table 4.1 presents more details about the two performance parameters.

4.1.3 Literature Review

To estimate the cyclone performance parameters there are five approaches:

1. Experimental investigations [12]
2. Theoretical and semi-empirical models [13]
3. Statistical models [14]
4. Computational fluid dynamics (CFD) [2–4, 7, 8]
5. Surrogate models (e.g., Polynomial regression (PR), Kriging (KG) and artificial neural networks (ANN)) [5, 6]

The afore-mentioned prediction models (PR, KG and ANN) have numerous drawbacks, which include locally optimal solutions, low generalization, over-fitting and poor stability [15]. The support vector machine (SVM) surrogate can offer a

better alternative to these models. In the field of performance evaluation for cyclone separators, unfortunately, SVM does not receive a great deal of attention on its algorithmic advantages. There is only one study using the support vector regression (SVR) on cyclone separator performed by Zhao [1]. He approved the potential of SVR to model the effect of cyclone geometry on the pressure drop (based on experimental dataset collected from different sources) but he did not go further to use the fitted SVR to study the effect of each parameter on the performance or for optimization. Moreover, he used the traditional approach to find suitable values for SVR parameters.

4.1.4 Target of This Study

This study aims to:

- Apply the SVR surrogate to model the variation of the two cyclone performance parameters with the change in the most significant geometrical parameters based on CFD based dataset.
- Introduce a computationally cheap framework for SVR parameter optimization using a Python code.
- Compare the accuracy of the fitted SVRs models with the KG models.
- Study the effect of each significant geometrical parameter on the cyclone performance using the SVR models.
- Optimize the cyclone performance for minimum Euler number as well as for best performance using multi-objective optimization technique.

4.2 Least Squares: Support Vector Regression

The least squares support vector regression (LS-SVR) was introduced by Suykens et al. [16] as a reformulation to the standard SVR. LS-SVR simplified the standard SVR model to a great extent by applying linear least squares criteria to the loss function instead of a traditional quadratic programming method [16]. As excellent examples of the nonlinear dynamic system, LS-SVR based on the structured risk minimization principle has been successfully applied to many fields of function approximation and pattern recognition because of its high accuracy and generalization capabilities [17]. Compared with ANN, LS-SVR seeks to minimize an upper bound of the generalization error instead of the empirical error, and can provide more reliable and better generalization performance under the same training conditions [18].

In the LS-SVR model, the training dataset of l points is assumed to be x_k, y_k ($k = 1, 2, \cdots, l$), in which $x_k \in R^n$ is the input vector and $y_k \in R$ is the corresponding target vector. The regression problem can be transformed into the

following optimization problem [15, 19]:

$$\underset{\omega,b,e_k}{\text{minimize}}\, \Psi(\omega, e) = \frac{1}{2}\omega^T\omega + \frac{C}{2}\sum_{k=1}^{l} e_k^2 \tag{4.1}$$

$$\text{subject to}\, y_k = \omega^T\phi(x_k) + b + e_k, \quad (k = 1, 2, \ldots, l) \tag{4.2}$$

where e_k is the error between the predicted value and the true value of the system, $C > 0$ is the regularization parameter applied to minimize estimation error and control function smoothness, $\phi(\cdot)$ denotes the nonlinear mapping from input spaces to feature space, ω is an adjustable weight vector and b is the bias (scalar threshold). Equation (4.2) is the constraint.

The resulting LS-SVR model for function estimation is obtained as:

$$\hat{y} = f(x) = \sum_{k=1}^{l} \alpha_k K(x_k, x) + b \tag{4.3}$$

In Eq. (4.3), $K(x_k, x)$ is the kernel function which satisfies Mercers condition corresponding to a dot product in some feature spaces. Four common Mercer kernel functions are widely used [15]:

Linear kernel: $K(x_k, x) = x_k^T x$
Polynomial kernel: $K(x_k, x) = (x_k^T x/\sigma^2 + \gamma)^d$
RBF kernel: $K(x_k, x) = exp(-\gamma \parallel x_k - x \parallel^2)$
Sigmoid kernel: $K(x_k, x) = tanh(\gamma x_k^T x + r)$

where d, γ and σ are constants.

Because RBF kernels map samples into high dimensional space in a nonlinear fashion and have fewer parameters to set, and because this method handles the nonlinear relationship well and has an excellent overall performance, it is by far the most popular option for kernel function types [16]. This study consequently adopted an RBF kernel function, shown in Eq. (4.4) in order to contribute to the LS-SVR model's achieving optimal solution.

$$K(x_k, x) = exp(-\gamma \parallel x_k - x \parallel^2) \tag{4.4}$$

Generally, LS-SVR solves linear equations and will lead to important reductions in calculation complexity. Compared with SVR, *LS-SVR is characterized by faster training speed, higher stability and better control strategy* [15, 19].

4.2.1 LS-SVR Parameter Optimization

The LS-SVR performance heavily depends on the choice of several hyperparameters, which are necessary to define the optimization problem and the final LS-SVR model.

To design an LS-SVR, one must choose a kernel function, set hyperparameters such as the kernel parameters, and determine a regularization parameter C. The hyperparameters that should be optimized include the regularization parameter C and the kernel function parameters such as γ for the radial basis function (RBF) kernel. Thus, selecting appropriate model parameters has a crucial impact on prediction accuracy. Unfortunately, there is no exact method to obtain the optimal set of LS-SVR hyperparameters; consequently, a search algorithm must be applied to obtain the parameters.

For the nonlinear LS-SVR, its generalization performance depends on the proper setting of parameters C and kernel parameters γ. Inappropriate hyperparameters combinations in LS-SVR lead to over-fitting or under-fitting. One procedure to obtain the LS-SVR parameters follows the trial and error approach to minimize some generalized error measures such as the mean squared error. This procedure is time-consuming, tedious and unable, in many cases, to converge at the global optimum. Zhao [1] applied the two-step search technique to dynamically seek the optimal values for the LS-SVR parameters. The two steps are: First perform a coarse search to identify a better region in search field according to contour lines of MSE. Then perform a fine search over that region. The disadvantage of the multi-step search technique is that it will be more prone to be trapped in local optimum point especially if a limited number of points are used.

In this study, we propose an alternative approach. The proposed approach employs the simulated annealing optimization technique to heuristically seek the optimal values for the LS-SVR parameters that minimize the difference between the predicted and the true values.

The simulated annealing (SA) is used in this study to optimize the parameters of SVR: C and kernel parameter γ of RBF-kernel function. In the training and testing process of LS-SVR, the objective is to minimize the errors between the actual and predicted values of the testing samples. Therefore, the objective (fitness) function of SA is the mean squared error from the cross validation.

In the parameters optimization process, K-fold cross validation is employed to avoid the over-fitting and to calculate the fitness function. The original sample is randomly partitioned into K subsamples. In these subsamples, a single subsample is used as the validation data for testing the model while the other K-1 subsamples are used as training data. The cross validation process is then repeatedly performed K times, with each of the K subsamples selected exactly once as the validation data. The cross validation error is estimated as the average mean squared error (MSE) on test subsamples, as shown in Eq. (4.5). Commonly, fivefold and tenfold cross validation is the most widely used method. In this study, the fivefold cross validations are employed to estimate the MSE.

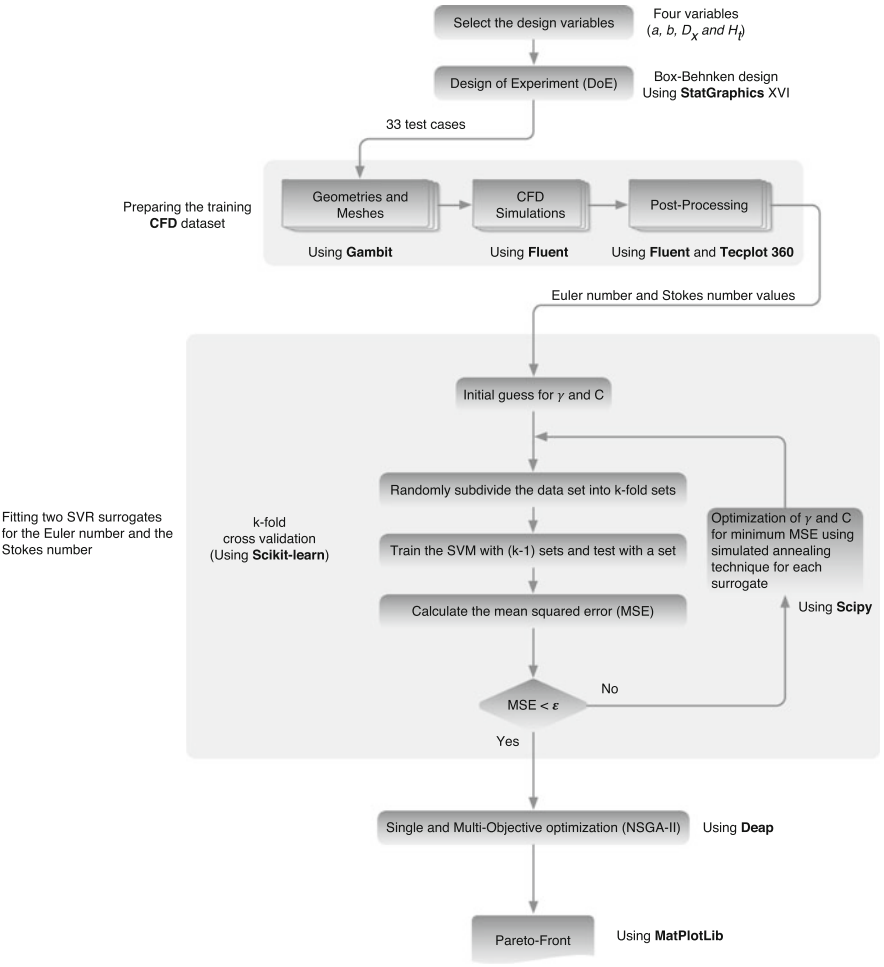

Fig. 4.2 Flow chart for the complete optimization framework

$$\text{MSE} = \frac{1}{K}\sum_{j=1}^{K}\left(\sum_{i=1}^{Nt_j}(\widehat{y_i} - y_i)^2\right) \tag{4.5}$$

where K is the number of folds (5 in this study), Nt_j is the number of testing points in fold j, y_i represents the actual values and $\widehat{y_i}$ represents the predicted values.

Figure 4.2 presents a flow chart for the complete optimization framework.

4.3 Results and Discussion

4.3.1 The Training Dataset

The training dataset has been created using the Box–Behnken design of experiment (DoE) and has been used in a previous study to optimize the cyclone geometry using the polynomial regression and RBF artificial neural network surrogates [7]. The minimum and maximum values for the four design parameters are listed in Table 4.2. To avoid scaling effect, all values are scaled (using the `preprocessing.MinMaxScaler` class from `Scikit-learn`) to be in range of 0 to 1 before being used in training the surrogates.

Table 4.3 presents the statistical descriptive parameters for the SVR (before and after parameter optimization) and Kriging surrogate. It is clear that the SVR with optimized parameter superior the performance of the Kriging model as is clear from the better matching between the statistical descriptive parameters of the input and the output results from the surrogate as well as the smaller value of MSE and the R^2 value close to unity.

4.3.2 Geometry Effect

One of the benefits of using surrogate models is to apply them to study the effect of each design variable on the response (performance parameters). The two optimized SVR models are used to study the effect of the four geometrical parameters on both the Euler number and the Stokes number. As is clear from Fig. 4.3, the SVR and KG models give the same trend of variation but the SVR models can predict more local variation than the KG model. It is worth to mention that the variation of the Euler number with the change in the vortex finder diameter D_x predicted by the SVR model is similar to that reported by the authors in previous studies [3, 7]. For the

Table 4.2 The values of the cyclone geometrical parameters used in the DoE (cf. Fig. 4.1)

Variables	Minimum	Center	Maximum
Inlet height, $x_1 = a/D$	0.4	0.55	0.7
Inlet width, $x_2 = b/D$	0.14	0.27	0.4
Vortex finder diameter, $x_3 = D_x/D$	0.2	0.475	0.75
Total cyclone height, $x_4 = H_t/D$	3.0	5.0	7.0
Cylinder height, h/D		1.5	
Vortex finder length, S/D		0.5	
Cone-tip diameter, B_c/D		0.375	

The values of the cylinder height, vortex finder length and the cone tip diameters are kept at the Stairmand design values, where $h - S = 1$ which is the optimum difference between the two dimensions as reported by many researchers [20]

Table 4.3 Statistical descriptive parameters for the SVR and KG models using the fivefold cross validation[a]

	SVR before parameter optimization (using the default values of C and γ)					
	Eu			$Stk_{50} \times 10^3$		
	x		y	x		y
Average	7.2436		4.3035	1.9496		1.9404
Standard deviation	6.9465		1.1746	0.9211		0.7573
Minimum	1.2110		2.0240	0.5880		0.6646
Maximum	27.2570		6.5702	3.8400		3.2680
Range	26.0460		4.5462	3.2520		2.6035
Coefficient of determination, R^2		0.131			0.900	
Mean squared error		47.750			0.076	
Intercept		3.2928			0.3480	
Slope		0.1395			0.8168	
The SVR regularization parameter, C		10.0			10.0	
The kernel parameter, γ		10.0			10.0	
	SVR after parameter optimization					
	Eu			$Stk_{50} \times 10^3$		
	x		y	x		y
Average	7.2436		7.2966	1.9496		1.9938
Standard deviation	6.9465		6.9108	0.9211		0.9271
Minimum	1.2110		1.3118	0.5880		0.4882
Maximum	27.2570		27.1573	3.8400		3.7398
Range	26.0460		25.8455	3.2520		3.2517
Coefficient of determination, R^2		0.9991			0.9941	
Mean squared error		2.579			0.00883	
Intercept		0.0935			0.0372	

(continued)

Table 4.3 (continued)

	SVR after parameter optimization					
	Eu			$Stk_{50} \times 10^3$		
	x		y	x		y
Slope		0.9944			1.0036	
The SVR regularization parameter, C		13502			283.151	
The kernel parameter, γ		0.154			0.041	
	Kriging (KG)					
	Eu			$Stk_{50} \times 10^3$		
	x		y	x		y
Average	7.2436		7.2436	1.9496		1.9496
Standard deviation	6.9465		6.9465	0.9211		0.9211
Minimum	1.2110		1.2110	0.5880		0.5880
Maximum	27.2570		27.2570	3.8400		3.8400
Range	26.0460		26.0460	3.2520		3.2520
Coefficient of determination, R^2 [21]		0.876			0.999	
Mean squared error		5.506			0.001	
Intercept		0.0			0.0	
Slope		1.0			1.0	

[a]x is the input to the SVR and y is the predicted value. Both x and y represent the Euler number and Stokes number.

Note that the values of *the average, the standard deviation, the minimum, the maximum and the range* have been calculated using all data for training the meta model. Consequently, for Kriging interpolation metamodel, these values are identical for x and y columns.

The average coefficient of determination R^2 using cross validation can be calculated as $R^2 = \frac{1}{K}\sum_{j=1}^{K}\left(1 - \frac{SS_{\text{res}_j}}{SS_{\text{tot}_j}}\right)$ where K is the number of folds, $SS_{\text{res}_j} = \sum_{i=1}^{Nt_j}(y_i - \widehat{y_i})^2$ is the sum of squares of residuals, $SS_{\text{tot}_j} = \sum_{i=1}^{Nt_j}(y_i - \bar{y}_j)^2$ is the total sum of squares, Nt_j is the number of testing points in fold j, y_i represents the actual values, $\bar{y}_j = \frac{1}{Nt_j}\sum_{i=1}^{Nt_j} y_i$ is the mean of the observed data in fold j and $\widehat{y_i}$ represent the predicted values

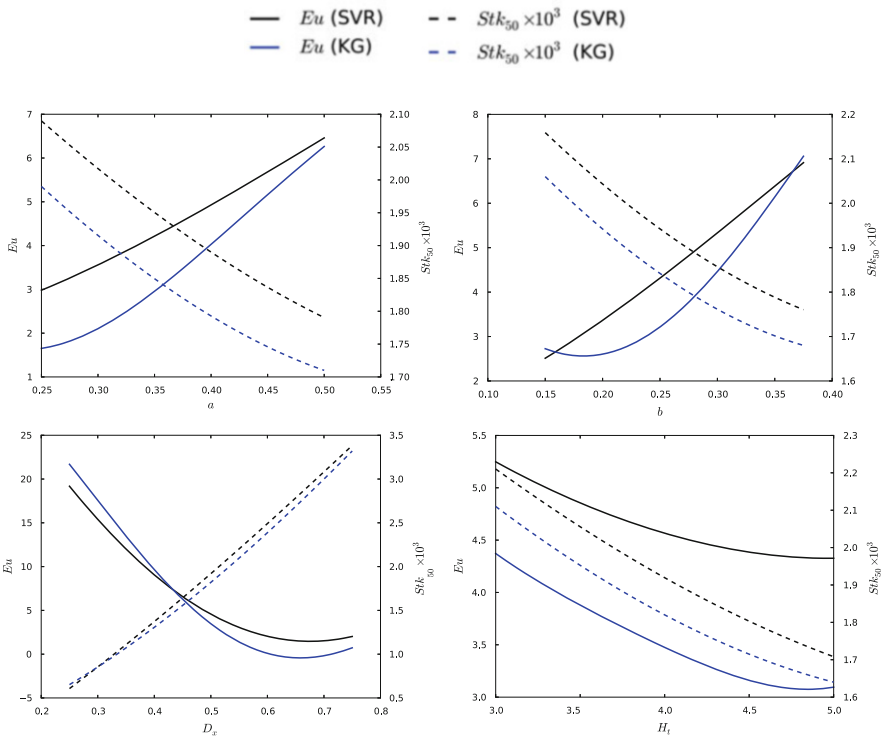

Fig. 4.3 Comparison between the effect of each geometrical parameter on the cyclone performance parameters using SVR (*black lines*) and KG (*blue lines*)

variation of the cyclone performance with the total height (H_t), the reduction in the Euler number (pressure drop) stops after $H_t = 4.625$ whereas the enhancement in the collection efficiency (reduction in the cut-off diameter) continue with lengthen the cyclone.

4.3.3 Geometry Optimization

Two optimization techniques have been applied to obtain new geometrical ratio set, namely the Nelder–Mead technique [22] and the genetic algorithms [6, 23]. Table 4.4 lists the new generated cyclone geometry ratios for minimum pressure drop (Euler number).

For single objective and one parameter optimization using the Nelder–Mead technique from `Scipy`. The total cyclone height H_t is optimized for minimum Euler number. The optimum value of H_t is 4.694, i.e., $h_c = 3.194$, where $h = 1.5$ which results in $Eu = 3.432$.

Table 4.4 Optimum geometrical ratios for minimum Euler number

	a	b	D_x	H_t	Eu
Optimum cyclone total height (Nelder–Mead technique)	0.25	0.15	0.5	4.694	3.432
Optimization of the four factors (Nelder–Mead technique)	0.499	0.15	0.658	3.0	0.667

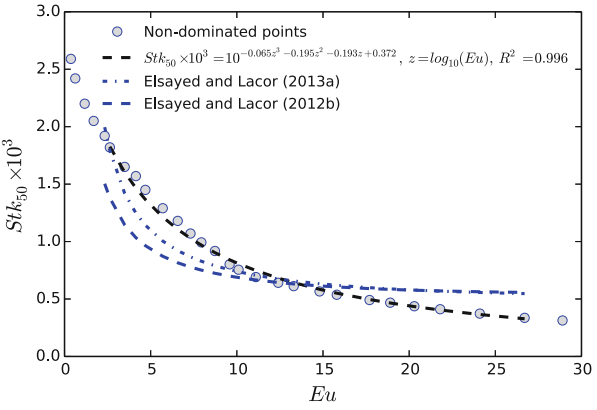

Fig. 4.4 Pareto front for NSGA-II optimization with polynomial fit

Since the cyclone performance has two major performance indices (the Euler and the Stokes numbers) a multi-objective optimization process is needed for optimum cyclone performance. In this study, NSGA-II [7, 24] available from `deep` (evolutionary toolbox for python) [25] has been used to obtain the Pareto front shown in Fig. 4.4.

The obtained Pareto front has been used to fit a correlation between the two performance parameters. The application of the `polyfit` function from `Numpy` Python package [26] results in the following correlation

$$Stk_{50} \times 10^3 = 10^{-0.065z^3 - 0.195z^2 - 0.193z + 0.372}$$

where $z = log_{10}(Eu)$. It is worth to mention that the accuracy of the new third order correlation is $R^2 = 0.996$ which is superior that proposed by the authors in a previous article [6]. This correlation can be used to predict the Stokes number (the cut-off diameter) given the pressure drop (Euler number).

Conclusions

In order to accurately predict the complex nonlinear relationships between the cyclone performance parameters and its geometrical dimensions, the support vector machines approach has been used and compared with the Kriging surrogate. Two SVR surrogate models have been trained and tested by 33 CFD datasets. The result demonstrates that SVR can offer an alternative and

(continued)

powerful approach to model the performance parameters. The SVR model parameters have been optimized to obtain the most accurate results from the cross validation steps. The parameters optimization has been optimized using the Simulated annealing technique. SVR (with optimized parameters) can offer an alternative and powerful approach to model the performance parameters better than Kriging. The SVR surrogates used to study the effect of the four geometrical parameters on the cyclone performance. The genetic algorithms optimization technique has been used to obtain a new geometrical ratio for minimum Euler number and for minimum Euler and Stokes numbers. The new cyclones over-perform the standard Stairmand design performance. A new correlation between the Stokes number and Euler number is provided which is more accurate than the existing correlations in the literature.

References

1. Zhao B (2009) Modeling pressure drop coefficient for cyclone separators a support vector machine approach. Chem Eng Sci 64:4131–4136
2. Elsayed K, Lacor C (2011) The effect of cyclone inlet dimensions on the flow pattern and performance. Appl Math Model 35(4):1952–1968
3. Elsayed K, Lacor C (2010) Optimization of the cyclone separator geometry for minimum pressure drop using mathematical models and CFD simulations. Chem Eng Sci 65(22):6048–6058
4. Elsayed K, Lacor C (2011) Numerical modeling of the flow field and performance in cyclones of different cone-tip diameters. Comput Fluids 51(1):48–59
5. Elsayed K, Lacor C (2011) Modeling, analysis and optimization of aircyclones using artificial neural network, response surface methodology and CFD simulation approaches. Powder Technol 212(1):115–133
6. Elsayed K, Lacor C (2012) Modeling and pareto optimization of gas cyclone separator performance using RBF type artificial neural networks and genetic algorithms. Powder Technol 217:84–99
7. Elsayed K, Lacor C (2013) 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
8. Elsayed K, Lacor C (2013) The effect of cyclone vortex finder dimensions on the flow pattern and performance using LES. Comput Fluids 71:224–239
9. Elsayed K, Lacor C (2014) CFD-based analysis and optimization of gas cyclones performance. In: International energy and environment foundation (IEEF), Chap 8, pp 223–276. ISBN 13: 978-1-49487-575-6
10. Hoffmann AC, Stein LE (2008) Gas cyclones and swirl tubes: principle, design and operation, 2nd edn. Springer, Berlin
11. Derksen JJ, Sundaresan S, van den Akker HEA (2006) Simulation of mass-loading effects in gas–solid cyclone separators. Powder Technol 163:59–68
12. Xiang R, Park SH, Lee KW (2001) Effects of cone dimension on cyclone performance. J Aerosol Sci 32(4):549–561
13. Stairmand CJ (1951) The design and performance of cyclone separators. Ind Eng Chem 29:356–383

14. Ramachandran G, Leith D, Dirgo J, Feldman H (1991) Cyclone optimization based on a new empirical model for pressure drop. Aerosol Sci Technol 15:135–148
15. Liu S, Xu L, Li D, Li Q, Jiang Y, Tai H, Zeng L (2013) Prediction of dissolved oxygen content in river crab culture based on least squares support vector regression optimized by improved particle swarm optimization. Comput Electron Agric 95:82–91
16. Suykens JAK, Gestel TV, Brabanter JD, Moor BD, Vandewalle J (2002) Least squares support vector machines. World Scientific, Singapore
17. Baylar A, Hanbay D, Batan M (2009) Application of least square support vector machines in the prediction of aeration performance of plunging overfall jets from weirs. Expert Syst Appl 36(4):8368–8374
18. Singh KP, Basant N, Gupta S, Sinha S (2011) Support vector machines in water quality management: a case study. Anal Chim Acta 703:152–162
19. Samui P (2011) Application of least square support vector machine (lssvm) for determination of evaporation losses in reservoirs. Engineering 3(4):431–434
20. Elsayed K, Lacor C (2013) Comparison of Kriging, RBFNN, RBF and polynomial regression surrogates in design optimization. In: Eleventh international conference of fluid dynamics (ICFD11), Alexandria
21. Elsayed K, Lacor C (2014) Robust parameter design optimization using Kriging, RBF and RBFNN with gradient-based and evolutionary optimization techniques. Appl Math Comput 236:325–344
22. Nelder JA, Mead R (1965) A simplex method for function minimization. Comput J 7(4):308–313
23. Holland JH (1975) Adaptation in natural and artificial systems. The University of Michigan Press, Ann Arbor
24. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197
25. Fortin F-A, De Rainville F-M, Gardner M-A, Parizeau M, Gagne C (2012) Deap: evolutionary algorithms made easy. J Mach Learn Res 13:2171–2175
26. Bressert E (2012) SciPy and NumPy: an overview for developers. O'Reilly Media, Sebastopol