A comprehensive insight into the application of machine learning approaches in predicting the separation efficiency of hydrocyclones

Mohammad Zandieh, Alireza Kazemi*, Mohammad Ahmadi

Department of Petroleum Engineering, School of Petroleum Engineering, Amirkabir University of Technology, Tehran, Iran, emails: kazemia@aut.ac.ir (A. Kazemi), m.a.zandieh@aut.ac.ir (M. Zandieh), m.ahmady@aut.ac.ir (M. Ahmadi)

ABSTRACT

The applications of hydrocyclones have been reputed due to their versatile performance, robustness, and easiness in design. In this study, the separation efficiency of hydrocyclones is intended to be determined by means of machine learning approaches, namely radial basis function, least-squares support-vector machine, multi-layer perceptron, and adaptive network-based fuzzy inference system. A data set entailing diverse features of either geometrical features of hydrocyclones and operating conditions are taken as inputs, and the separation efficiency is considered as the target element. Sensitivity analysis is also carried out to eliminate less effective inputs before applying the algorithms to the collected data set to improve the accuracy of algorithms and avoid inevitable noises and unnecessary extra measurements. Since the performance of utilized artificial neural networks is considerably dependent on the optimal network configurations, an optimization scheme using particle swarm optimization is used to meet the objective. The results authenticated the applicability of the respective algorithms in predicting the separation efficiency of hydrocyclone as a function of 14 different features. Good agreements between experimental data and applied models have been ascertained. Lastly, in-depth analyses into the behavioral performance of employed algorithms and their potential effects on the prediction capability have also been presented.

Keywords: Hydrocyclones; Sensitivity analysis; Artificial neural networks; Machine learning algorithms; Data science

1. Introduction

Hydrocyclone is an instrument that causes the feed phases to separate using swirling centrifugal forces, in which the separation driving force lies within the transformation of the static energy of the fluid (fluid pressure) into dynamic energy (fluid velocity). The high desirability of utilizing hydrocyclone can be attributed to its simple design, high sustainability, high separation efficiency, and small occupying area. The exclusive purpose of hydrocyclones used to be solid–liquid separations [1]; however, with the burgeon of technology, there have been a plethora of researches so far in the quest of achieving an enhancement in its characteristics to boost the overall functionality and enhance the separation process [2].

Although the above-mentioned advantages associated with this apparatus have led it to be utilized extensively in a myriad of fields, the conceptual mechanism recognition of this apparatus is still an underrated area. Also, there are still some struggles that hinder the hydrocyclones from performing effectively. Of the main problems that are accountable for this downturn in the functionality of hydrocyclone, the high viscosity of the continuous phase, the inevitable impacts of possible phase inversion, long term changes in field conditions, unfavorable interfacial behavior, adverse

* Corresponding author.

physical properties, small inlet drop size distributions, and sub-optimal de-oiling hydrocyclone geometries can be mentioned [3]. Moreover, other concerns such as the presence of free gas and variation in feed phase content have put restrictions on hydrocyclones separation efficiency in some cases [2]. Therefore, resolving these negative sides are of paramount importance and necessity.

To this end, there have been numerous attempts, either experimentally and computationally, to delve into the above-mentioned difficulties associated with hydrocyclones. These examinations have brought about helpful insights and clearer perception into the separation process in conjunction with proposing better calibrations that are shown potential to alleviate the earlier concerns to some extent, however, there are still some shortcomings that must be pointed out. To begin with, bearing in mind the complications, uncertainties, and cost- and time-concerned difficulties associated with investigating this apparatus experimentally has led the researchers to use alternative approaches, particularly computational examinations. Similarly, although computational approaches have exhibited robust capabilities in simulating complex matters, given the inherent complication of hydrocyclones' functionality, computational approaches are considered time-consuming too. To be more precise, simulating hydrocyclones using software viz. computational fluids dynamics (CFD) not only comes with several assumptions and simplifications in the governing mechanisms, but also, it is usually hand in hand with cumbersome run-times if a precise setup is chosen throughout the simulations [2].

Given the raised concerns as above and bearing in mind that analyzing hydrocyclones' adjustable characteristics to tackle the declared obstacles may be considerably cumbersome, having a reliable tool to analyze, forecast, and optimize following projects prior to modeling and testing the hydrocyclone prototypes in a convenient manner is imperative. In other words, there is an obvious need for an elaborated technique that can predict the performance of hydrocyclones under diverse operating and geometrical conditions. To this end, machine learning approaches are introduced worldwide, which are currently being applied extensively in many fields including, statistics, physics, mathematics, and neuroscience. To shed light on some of their well-reputed applications, pattern recognition, signal processing, and time series analysis can be cited. The conceptual idea lying behind the basics of this intelligence is an ability programmed to learn from a system of inputs, with or without human intervention. The main benefit of this technology is saving costs and time while giving a thorough survey of what is about to be produced. Also, these tools are easy to implement and can avoid time-consuming experimental procedures and sophisticated interpretation processes.

Karr et al. [4] sought an optimization of hydrocyclone performance using genetic algorithms. They stated that genetic algorithms could be used to model hydrocyclones. They implemented a genetic algorithm to ascertain proper membership functions that can be utilized in various industrial approaches. Karimi et al. [5] analyzed the possibility of using artificial neural networks (ANNs) to predict underflow and overflow flow rates. Their results revealed the high capability of ANNs in the precise estimation of variable

hydrocyclone components. van Loggenberg et al. [6] studied the estimation of hydrocyclone cut size using an artificial neural network. They stated that the combination of the selected features is crucial to have a favorable outcome in predicting the cut size. In a major work, Fung et al. [7] sought a new technique to utilize the combination of ANN and fuzzy logic systems to design a neural-fuzzy hydrocyclone model. Their suggested system showed great abilities in learning and interpreting the functionality of hydrocyclones. They also recommended further studies to find an optimum configuration of the fuzzy system. Eren et al. [8] compared two conventional models of predicting hydrocyclone's cut size with their proposed ANN systems. Their outcomes portrayed excellent correspondence with original data. They indicated that a well-taught assortment of input variables and trained data could better fit perditions of ANN systems with experimental data. Hybrid fuzzy modeling was introduced by Wong et al. [9] as a substitute for hydrocyclones automatic control systems. They also stated that their proposed model could work more effectively compared to Sugeno and Yasukawa's qualitative models (SY models) under specific circumstances. Mohanty et al. [10] attempted to estimate a particular hydrocyclone's dewatering performance as a function of structural and operating variables including spigot diameter, vortex finder diameter, and inlet pressure using neural network models. Their proposed model employed the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) to lower the load of measurements and make the convergence of the utilized algorithm swifter.

On the other hand, since in most machine learning algorithms, the training process may take a considerable duration of time when the data size is large, the metaheuristic algorithms can be utilized to lower down the losses and boost up the accuracy of the models. There are numerous metaheuristic algorithms of which each come with different approaches and applicability. They are categorized into four types, for example, biology-based, physics-based, sociology-based, and mathematics-based algorithms [11]. A comprehensive review of the metaheuristic algorithms has been presented in the study of Yang et al. [11]. Of the common metaheuristic algorithms, genetic algorithms [12], differential evolution [13], artificial bee swarm optimization [14], whale optimization [15], improved ant lion optimizer [16], bird mating optimization [17], grey wolf optimization [18], bacterial foraging algorithm [19], artificial immune system algorithm [20], Salp swarm algorithm [21], mutative scale parallel chaos optimization algorithm [22], Lozi map-based chaotic optimization algorithm [23], harmony search algorithm [24], pattern search algorithm [25], shuffled complex evolution [26], JAYA algorithm [27], and simulated annealing algorithm [28] can be cited.

Furthermore, it is indisputable that an accurate prediction of hydrocyclone performance is highly dependent on a thorough consideration of all involved parameters viz. geometrical and dynamical characteristics, which is lacked in previous investigations. Besides, another critical milestone throughout this process is deemed as the utilized data bank. Concisely, since the prediction ability of machine learning approaches mainly goes hand in hand with the provision of a comprehensive data bank of previously conducted investigations, there is also a dearth of an elaborated data set in

previous works. Henceforth, in the present study, it is tried to meet both of these two requirements; as such, an all-embracing data bank of various hydrocyclone features and operating conditions are employed to examine the ability of employed machine learning approaches in assessing the performance of hydrocyclones in terms of separation efficiency (as the target) as a function of fourteen variables (as inputs). Additionally, it is also attempted to prepare a comprehensive data bank comprised of inputs and their corresponding targets to facilitate a more precise perdition of the separation efficiency and lower the uncertainties.

Also, it is attempted to improve the perdition ability of the utilized machine learning techniques and escalating their accuracy by narrowing down the inputs into the most dominant elements using a reliable tool known as the sensitivity analysis. In addition, the collected data might carry data points with high levels of inaccuracy and inconsistency because of the inevitable errors throughout experimental measurements. To overcome this obstacle, the outlier detection method is used to remove irrelevant data points. In the first phase of this study, the general characteristics of hydrocyclones, and the clarification on the rationale in preparing the inputs, are presented. Then, the outlier technique is applied in combination with the sensitivity analysis. Afterward, the proposed machine learning approaches are demonstrated. Subsequently, the paper culminates with analyzing the performance of each algorithm from different angles and hypothesize the behavior of each algorithm using various benchmarks. Last but not least, conclusions about the outcomes of this paper and suggestions for future research orientations have been presented.

2. Methodology

In this section, the applied methodologies including the integrated hydrocyclone design, data gathering process, outlier detection and sensitivity analysis are presented.

Furthermore, the rationales behind the results of such analyses are compared, followed by a thorough demonstration into the scientific reasons associated with the sensitivity analysis results.

2.1. Hydrocyclone design

According to experimental and computational researches that have been carried out in this area, there are numerous features that must be considered to analyze the performance of hydrocyclones. These parameters can be categorized as operational and structural types. Operating parameters concerns mostly with the properties of fluids, and as for the structural factors, physical elements are taken into account. The firstly introduced types of hydrocyclones were generally typical, however, up and coming projects have sought to scrutinize into more variables in the quest of maximizing the ability to predict the performance of hydrocyclones, which led the main structure to be more developed. A general schematic of typical hydrocyclones is given in Fig. 1. In this study, 14 input variables, namely, inlet width, overflow and underflow diameters, cylindrical part diameter, cylindrical part length, tail length, inlets number, dispersed phase droplet diameter, flowrate, cone and tail angles, initial pressure, initial temperature, phases density difference, have been taken into account. Table 1 portrays the specifications of inputs.

2.2. Data gathering

The references from which the data set is collected are as follows: [3,29–56]. Since the pivotal reliability of any estimation is having an adequate data set, in this study, a comprehensive data set has been gathered based on their frequency. In other words, the data set concerns those parameters which have been studied greatly in former investigations. In addition, it must be noted that since each

Db: Cylindrical part diameter

Do: Over flow diameter

Du: Under flow diameter

Di: Inlet width

α: Cone angle

β: Tail angle

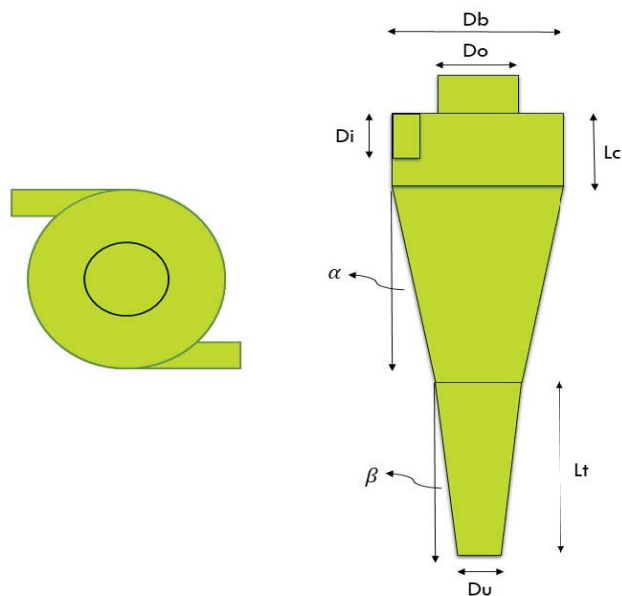Lc: Cylindrical part length

Lt: Tail length

Fig. 1. Typical schematic of hydrocyclone.

previous examination that is used for data gathering must have mentioned all the inputs illustrated in the present paper, there are some parameters (which may have been proven to be important) that have not been selected due to lack of adequate consideration in earlier examinations. To mention some features that can be deemed important, the volume fraction of feed fluids and their viscosities can be cited. Although there are some techniques to measure the missing data, it must be declared that samples with missing data points are not taken into account. The number

of collected samples that are applied in this research was 4,560 single data or 304 samples, as such a column of data sets containing all features assumed as a sample.

2.3. Outlier detection

Since the utilized data are mostly based on experimental measurements and bearing in mind that inaccuracy in calculations throughout the test procedure is unavoidable, there must be an analysis to evaluate the precision of gathered data. One of the most frequent and well-known method to do so is known as the Leverage mathematical technique. In this method, two important items are required to be calculated. Firstly, the residual values and then a matrix named Hat matrix, which is the matrix of input data. This matrix can be measured as follows [57]:

$$H = X\left(X^T X\right)^{-1} X^T \tag{1}$$

In which, X refers to the $K*P$ matrix, where K and P stand for the number of inputs, and the training points, respectively. It should be noted that critical leverage value is calculated using the formula below, and data points crossing this value would be considered outliers [58]:

$$H* = \frac{3(P+1)}{K} \tag{2}$$

Lastly, in order to portray a schematic of imprecise data points, William's plot can be drawn for the residual values vs. the hat values. The hat values can be achieved from the Hat matrix's main diagonal. The William's plot is presented in Fig. 2. Accordingly, 10 suspected inexact data

Table 1
The inputs specification in terms of range and unit

Inputs	Range
Flowrate, L/min	[40, 183606]
Dispersed phase droplet diameter, μ	[0.4167, 185]
Inlet width, mm	[10.5, 105]
Cylindrical part diameter, mm	[20, 205]
Over flow diameter, mm	[1.5, 105]
Underflow diameter, mm	[12.5, 79.8]
Cylindrical part length, mm	[60, 413]
Cone angle	[1.5, 20]
Tail angle	[0, 20]
Initial pressure, kPa	[100, 634.318]
Initial temperature, °C	[15.6, 800]
Inlets number[a]	[1, 4]
Phases density difference, g/cm^{3b}	[0.160, 2.7960]
Tail length, mm	[50, 1804.4]

[a]Number of selected inlets are: 1, 2 and 4;
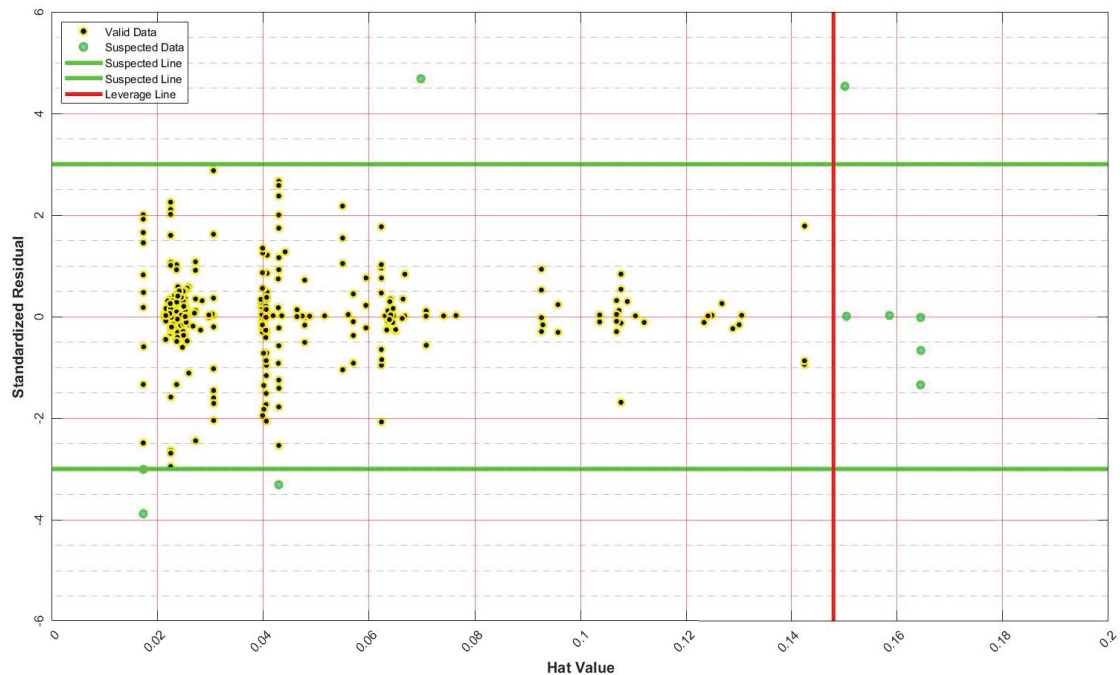[b]The absolute difference of densities are chosen.

Fig. 2. Outlier detection for probably suspected data points using William plot.

have been recognized. The following results are conducted after elimination of suspected data, and it was found that the precision of algorithms outcomes was improved up to approximately 12% after retraining of models.

2.4. Sensitivity analysis

Sensitivity analysis (or relevancy analysis) is a method that demonstrates the target parameters' degree of dependency towards the fluctuation in each input, individually. This analysis leads to cost-effective examinations of modeling and experimental explorations since it can reveal the most effective parameter that plays significant roles in overall functionally associated with the intended structure (depending on the target parameter); This technique is formulated as follows [59]:

$$R = \frac{\sum_{i=1}^{N}\left(X_{j,i} - \overline{X}_J\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^{N}\left(X_{j,i} - \overline{X}_J\right)^2 \sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}} \tag{3}$$

where $X_{j,i}$, $\overline{X}_J$, Y_i, $\overline{Y}$ and N are the input value, average value of 'i' value, output value, the average value of output, and total numbers of data points, respectively. Each parameters' degree of influence is represented by a factor denoted as 'R', in which the more the 'R' value, the greater the influence. Bedsides, the positive or negative sign of 'R' portrays the direct or inverse impacts of the intended parameter. The result of the relevancy analysis is presented in Fig. 3.

Another important analysis that can reveal a good clarification into the effects of each input is input combination sets [60]. To be more exact, different grouping assortment of inputs are applied to scrutinize the influences of the variables; as such, groupings of 1 input up to all 14 inputs are considered separately, and the performance (in terms of R^2 which is explained in the following sections) of the various models are investigated under these setups. Firstly, it was found that the LSSVM model exhibited more suitable precisions relatively, so it is chosen to analyze this part. Fig. 4 illustrates the average of R^2 for each setup of inputs combinations, which means each number on the x-axis represents all the combinations for that particular number of variables. It is discovered that as the number of inputs in the grouping setups increments, the capability of the model to recognize the intended target is strengthened (the reason behind this happening is discussed thoroughly in section 5.2). Also, as it can be perceived from Fig. 4, the accuracy of the model prediction tends to slightly decrement for the grouping of 13 and 14 inputs. This decline in the performance is ascribed to the consideration of less-effective inputs (as will be pointed out late in this section) in the grouping setup. This is in-line with the founding of the sensitivity analysis.

To demonstrate the sensitivities, the outcomes of this analysis must be validated primarily. To begin with, the increase in flow rate and initial pressure intensities the Reynolds number, consequently, the generated unstable air-core within the hydrocyclone becomes thinner and ultimately choking at the overflow will be diminished, and an enhancement in separation efficiency can be achieved [30]. The initial temperature is responsible for the diameter of the dispersed phase, feed viscosity as well as density difference. In a way, that increase in temperature mainly causes the interfacial tension between the particles to reduce, and as a result, the dispersed phase droplets have the chance to coalesce, which leads to bigger droplets formation that is
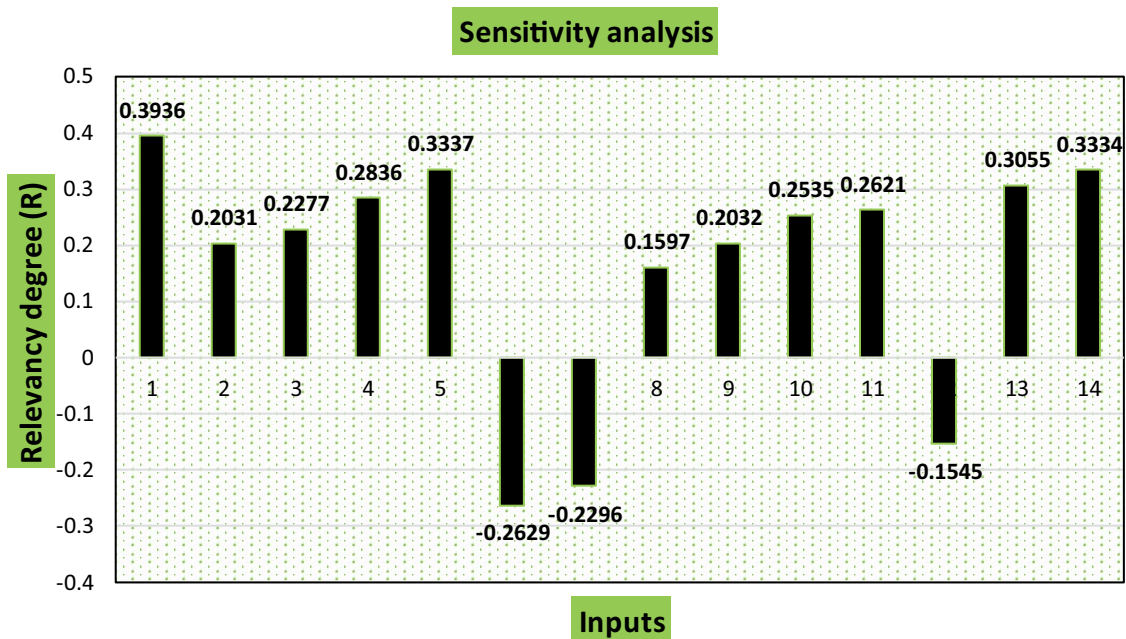

Fig. 3. The results of sensitivity analysis.

The input labels in the diagram are as follows: 1-Flowrate, 2-Dispersed phase droplet diameter, 3-Inlet width, 4-Overflow diameter, 5-Underflow diameter, 6-Cone angle, 7-Tail angle, 8-Initial pressure, 9-Cylindrical part diameter, 10-Initial temperature, 11-Inlets number, 12-Cylindrical part length, 13-Phases density difference and 14-Tail length.
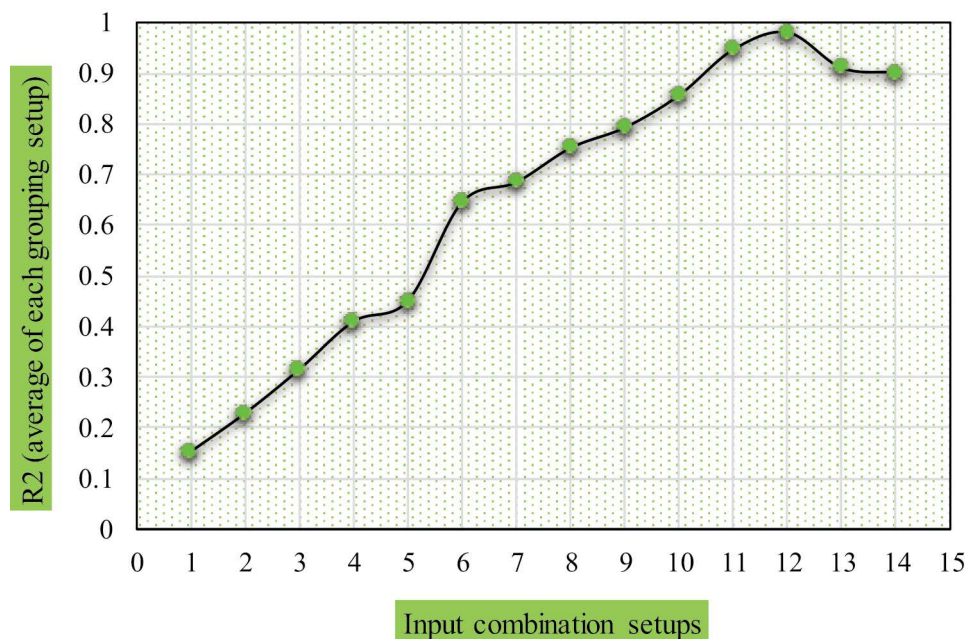
Fig. 4. The performance of the LSSVM model under various input setups.

easier to separate [49]. Moreover, the greater the phase density difference, the more effective the drag forces to form the primary and secondary vortexes, which ultimately leads the particles to be displaced much faster towards the central vortex. Also, provided that the particle diameters are big enough, higher phase density discrepancies result in higher tangential velocities commensurately, which improves the separation procedure and increases the separation efficiency [61]. The increase in inlet diameter is beneficial for either having higher capacities of feed fluid and higher swirling velocity [44].

The effects of underflow and overflow diameters are analogous. They allow the separated fluid to be discharged, so the wider the diameters, the more the discharged separated fluid [32]. However, these two parameters are effective as far as the flowrate is not high enough. If the flowrate is sufficient, the effects of these diameters will no longer be as important as earlier [41]. The tail length is accountable for the back pressure effects. Backpressure at the underflow applies an upwards force in terms of a vortex in order to direct the dispersed phase to the overflow. If this factor is high enough, the separation of very small droplets is promising [32]. The number of entrances used to be assumed as one specific entry. But it has been proven that double entrances cause the flow and dispersed phase within the central axis to be uniform which brings about high rates of separation increment. The increase of inlets to 4 was found to have positive effects of separation efficacy, but the contribution was not substantial [62]. On the contrary to the previous inputs that were directly proportional, the rise in tail angle allows the dispersed phase core to be more developed axially, which leads to smaller droplets being discharged through underflow and ultimately plunging the separation efficiency [32]. Meanwhile, it lessens the back pressure, which is imperative for performance stability.

The body length (cylindrical part length) is affecting the separation negatively due to increasing the reduction of primarily generated momentum. If it lengthens, the friction of walls, which is in contact with the main vortex, causes the fluid to lose energy, and consequently, the fluid would face a lack of necessary driving force (centrifugal force) to take part in the separation process [32]. Though, a particular length is required in order to let the natural turning length necessity to be fulfilled [31]. In the case of body diameter (cylindrical part diameter), it is not easy to examine the total contribution, due to the dependency on various aforementioned parameters, but it is proved that the increase in its diameter may enlarge the capacity occupied by both phases in which under specific circumstances, positive attributes may be perceived, and in some occasions, provided that there are sufficient flow rate and optimum overflow diameter, this factor may not be as influential as before [52]. The cone angle is accountable for strengthening the swirling intensity and as a consequence enhancing the applied drag forces, inner vortex formation, and upward movement of particles, however, it must be stated that depending on the feed fluids specifications, this parameter is delicate in which there is an optimum value that cyclone functionality is enhanced and subsequent to that value, the performance declines severely [46].

As it was highlighted before, since the convergence and accuracy of algorithms are dependent on a proper selection of inputs, cylindrical part length and initial pressure have been omitted since they carried the least relative effects on the separation efficiency based on the sensitivity analysis. However, it is worth mentioning that exclusion of the aforementioned parameters does not imply their complete lack of competency in affecting the separation efficiency. To validate this decision, it is attempted to run the algorithms under two circumstances of with and without the elimination

of less effective parameters. It was found that the results of the algorithms in the second scenario faced a reduction in accuracy. Hence, the elimination of the two aforementioned inputs turned out to enhance the algorithms performance.

3. Model development

In this study, four types of machine learning algorithms, particularly, adaptive network-based fuzzy inference system (ANFIS), multi-layer perceptron (MLP), radial basis function (RBF), and least-squares support-vector machine (LS-SVM) have been implemented to predict the separation efficiency of hydrocyclones under various chosen features. MATLAB R2019a was used in order to carry out the tests. The data are shuffled prior to further examinations in order to perform the analyses unbiasedly. Since the units and the range of inputs variability are widely different, the collected data must be normalized so they can be judged uniformly [63]. Data normalization is one of the pre-processing methods to either scale or transform the input data so they can be uniformly treated as of equal contributions. In other words, assuming two different inputs with huge dissimilar values, 1E-2 and 1E3 as an example, may cause the models to assess them erroneously by misinterpreting inputs with higher values as of higher influence. Furthermore, it is found that the machine learning algorithms performance is boosted commensurately, in terms of runtime when this input data pretreatment is applied [63]. Alternative approaches analogous to normalization techniques are scrutinized in the work of Singh and Singh [64]. The normalization formula is given below [65]:

$$X_N = 2\frac{X - X_{\min}}{X_{\max} - X_{\min}} - 1 \tag{4}$$

where the X_n, $X_{\min}$, $X_{\max}$ and X refer to the normalized value of the parameter, maximum and minimum of the point and the raw data point, respectively.

3.1. Artificial neural network

Artificial neural networks are mathematical algorithms that are inspired by biological neural concepts. These algorithms can perfectly predict non-linear functions [66]. They are also capable of machine learning as well as pattern recognition. The main components of the ANN are the nodes which are called processing elements [67]. Each node entails the inputs and a function to transform the received data into the outputs. Throughout this procedure, the learning process adopts and modifies the inputs to obtain the intended data structure. Also, the RBF and MLP algorithms are of the well-known artificial neural networks which are implemented in this research.

MLP structure encompasses three layers which are illustrated as input, output, and lastly hidden layers. In order to solve nonlinearly separable problems, a multi-layer perceptron is constructed by numbers of neurons. Each perceptron is used to classify small linearly separable sections of the received inputs. Then, outputs of the perceptron are combined into another perceptron to produce the final output. The hard-limiting (step) function used for producing the output prevents information on the inputs flowing on to inner neurons. It is demonstrated that a two-layer perceptron should be sufficient as an universal approximation of any non-linear functions and a three-layer network suffices to separate any polyhedral decision region (convex or non-convex) [68]. Furthermore, the 'Tansig' and 'Purelin' functions were employed for hidden and output layers of MLP algorithm, respectively; which stand for Tansigmoidal and Purely linear, respectively. The MLP model uses the Levenberg Marquardt algorithm to train the data. On the contrary to MLP conceptual structures, the RBF is designed using a simpler configuration and consequently, a swifter training process. The theoretical idea of radial basis function is driven from the theory of function approximation. Analogous to the MLP, the RBF utilizes a hidden layer learning process but with a slight difference in approach [69]. The RBF used in this research is a two-layer feed-forward network. The hidden nodes employ radial basis functions, Gaussian functions as an example. Similar to MLP, the output nodes apply linear summation functions. The training of the network can be classified into two sections; first, the weights from the inputs to the hidden layer are regressed, and then, the regression of weights from the hidden layer to the output layer. Presuming the inputs as x, the output of the RBF is as follows [70]:

$$y_i(x) = \sum_{k=1}^{h} w_{ki}\varnothing\left(\left\|x - x_k\right\|\right) \tag{5}$$

where w is the weight from the kth hidden unit to the ith output, x_k is the center of the kth hidden node, and $\|\ \|$ indicates the Euclidean function. As it was aforementioned, the Gaussian function ($\varnothing$) is a common radial basis function which can be illustrated as below [70]:

$$\varnothing(x) = \exp\left(-\frac{(x - c)^2}{r^2}\right) \tag{6}$$

where r and c stand for the radius and the center of the Gaussian function, respectively. Additionally, it is recommended that the number of neurons in the hidden layer for the RBF model must not exceed one-tenth of data points on aggregation [71]. Despite all the advantages the RBF model present, there are a couple of setbacks with the interpolation networks implemented in RBF. Firstly, it is not desirable to have the network's outputs gone through all the data points when the data are noisy, because generalization will be performing poorly. Secondly, since the hidden layer requires one particular basis function, in the presence of large data sets, the evaluation process will not be computationally efficient. However, there are some tactics that are promising to boost the performance [69]. For instance, the centers of the basis functions do not necessarily need to be defined as the training data input vectors. They can be determined by a training algorithm.

3.2. Particle swarm optimization

The particle swarm optimization (PSO) algorithm was primarily introduced by Kennedy and Eberhart [72]. It was

utilized to optimize continuous nonlinear functions [73]. The PSO algorithm maintains multiple potential solutions at one time [74]. Throughout each iteration, each solution is determined by an objective function to evaluate its fitness. Each solution is represented by a particle in the fitness space (search space). Then the particles either fly or swarm through the search space to discover the maximum value retrieved by the objective function. The algorithm encompasses three steps; (1) evaluation of each particles' fitness; (2) updating individual and global bests positions; (3) updating velocity and each particle's positions. The process goes on until one of these requirements is met. The relevant formulation is illustrated as follow:

Firstly, the position of individual particles updates using the formula below [75]:

$$x_{k+1}^i = x_k^i + v_{k+1}^i \tag{7}$$

And the velocity is calculated as follows [76]:

$$v_{k+1}^i = \omega v_k^i + c_1 r_1 \left(p_k^i - x_k^i \right) + c_2 r_2 \left(p_k^g - x_k^i \right) \tag{8}$$

where x_k^i is the particle position, v_k^i is the particle velocity, p_k^i is the best-remembered individual particle position, p_k^g is the best-remembered swarm position, c_1, and c_2 are cognitive and social parameters that aid particles to go through more adequate space of solution area, and r_1, r_2 are random numbers between 0 and 1. The ω is the inertia weight that keeps the particle moving alongside the same direction as it was formerly heading. The lower value of ωv_k^i increases the convergence, while higher values motivate reconnoitering the search space. The renewal of inertia weight is being carried out using the formula below [77]:

$$\omega_i = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{i_{max}} \times i \tag{9}$$

In which the ω_i is the inertia weight in each iteration, ω_{max} and ω_{min} are the maximum and minimum inertia weights, respectively. This factors fluctuates mostly between 0.8 and 1.2 [74].

3.3. Adaptive network fuzzy-based inference system

The pivotal involvement of fuzzy logic deals with inaccuracy and granularity [78]. It is also capable of numerical calculations by getting help from the membership function. The main concepts that lie behind fuzzy inference systems (FIS) are if-then rules, fuzzy set theory, and fuzzy cognition. Nowadays, FIS is extensively applied in various fields, viz. data classification, automatic control, robotics, and pattern classification. The FIS analysis requires several steps to be conducted including fuzzification, aggregation, activation, accumulation, and ultimately defuzzification. Different types of fuzzy systems can be classified as Mamdani, Singleton, and Takagi-Sugeno fuzzy systems.

The fuzzy logic takes imprecision and uncertainty into account, whilst, the neural network seeks a system of adaptability. A new class of adaptive networks that are functionally equivalent to fuzzy inference systems have been introduced as ANFIS which stand for adaptive network-based fuzzy inference system. This system is an integration of both previously mentioned models, neural network and fuzzy inference system. The parameters in ANFIS are estimated using both Sugeno and Tsukamoto fuzzy models [79], and hybrid learning algorithms are being used throughout the measurement procedures [80]. PSO, genetic algorithms (GA), and Imperialistic Cognitive Algorithm (ICA) can be favorable approaches to pursue optimal models while training ANFIS [81]. In this paper, the PSO option is utilized to meet the purpose. However, it is worth mentioning that without proper usage of MATLAB toolbox for ANFIS configuration, it might be almost impossible to run an ANFIS algorithm with inputs exceeding 5 numbers. A schematic of the ANFIS structure is given in Fig. 5.

For simplicity purposes two inputs and one output are assumed. After then the layers can be formulated as follows:

First layer: To begin with, the first layer can be generated as [82]:

$$O_i^1 = \mu_{Ai}(x) \tag{10}$$

where A_i is the linguistic variable, x is the ith node input and μ_{Ai} is the membership of function associated with A_i. The membership function can be described as [83]:
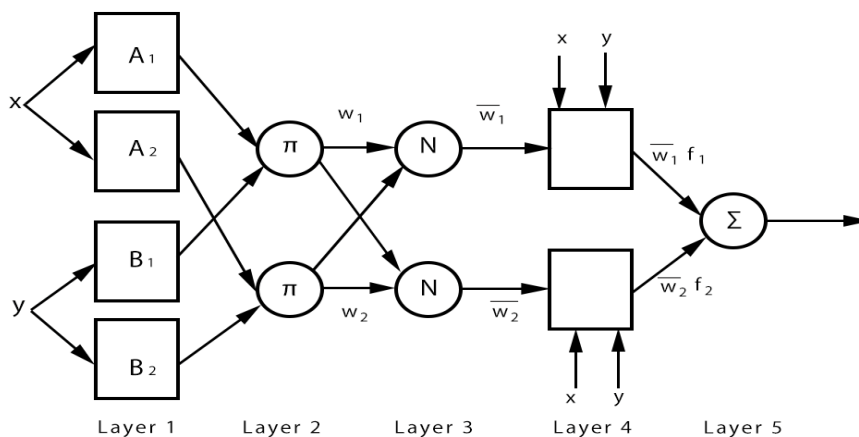

Fig. 5. A schematic of ANFIS structure.

$$\mu_{Ai}(x) = \frac{1}{1 + \left[\left(\dfrac{x - c_i}{a_i}\right)^2\right] b_i}$$ (11)

In which α_i, b_i, c_i are the premise parameter sets [78].

Second layer: In the second layer firing strength μ_i is calculated and the result of each node is a product of received indicators [82]. Here by the result is presented:

$$O_i^1 = \omega_i = \mu_{Ai}(x) \times \mu_{Bi}(y) \quad i = 1, 2$$ (12)

Third layer: In this layer, each node is a stable node where the ith rule's firing strength per aggregation of all firing strengths is measured in each ith node [82]. This ratio is given below:

$$O_i^3 = \frac{\omega_i}{\omega_1 + \omega_2} \quad i = 1, 2$$ (13)

Fourth layer: In this section, every node can be described as an adaptive node with a node function given as follows [57]:

$$O_i^4 = \varpi_i f_i = \varpi_i (p_i x + q_i y + r_i) \quad i = 1, 2$$ (14)

where ϖ_i is the result of the third layer and p_i, q_i, r_i are the subsequent parameter set [83].

Fifth layer: This layer consists of one solely node which measures aggregation of preceding outputs as the accumulation of all received indications [84].

$$O_i^5 = \sum_i \varpi_i f_i = \frac{\sum_i \omega_i f_i}{\sum_i \omega_i}$$ (15)

3.4. Least squares support vector machine

LS-SVM algorithm was suggested by Suykens and Vandewalle [85], which is a modified version of a previously introduced model, named support vector machine (SVM) which was originally presented by Vapnik [86]. The main theory of SVM was grounded on statistical learning and was used for classifications and regressions purposes primarily. However, the chief impediment of SVM was that extensive non-linear systems of equations were employed, on the contrary, LS-SVM makes a substantial significant benefit by applying a linear approach system of equations. Swifter performance of LS-SVM in comparison with SVM is the vital consequence of this upgrade. LS-SVM model presents a function of regression approximation as follows [75]:

$$y_k = (\omega, \varnothing(k)) + b + e_k \quad k = 1, 2, \ldots, N$$ (16)

Presuming the data set contains twelve variables, x stands for the inputs, y refers to the intended target in

which is a function of inputs, $\varnothing(k)$ refers to functions that can come in linear, polynomial, sigmoid, and radial forms, b stands for a scalar threshold or bias factor, ω denotes the weight factor and N conveys the set of data points. The first step is evaluating optimal values of ω and b. In order to determine function estimation in LS-SVM, the equations below are formulated [75]:

$$\min_{\omega, b, e} J(\omega, e) = \frac{1}{2}\|\omega\|^2 + \gamma \frac{1}{2}\sum_{k=1}^{N} e_k^2$$ (17)

Subjected to:

$$y_k = \left\{\omega, \varnothing(x_k) + b + e_k\right\} \quad k = 1, \ldots, N$$ (18)

In which e_k conveys the error, γ stands for regularization variable. Furthermore, the Lagrange equation is as follows [75]:

$$L = \frac{1}{2}\|\omega\|^2 + \gamma \frac{1}{2}\sum_{k=1}^{N} e_k^2 - \sum_{k=1}^{N} \alpha_i \left\{\left(\omega, \varnothing(x_k)\right) + b + e_i - y_i\right\}$$ (19)

In which α_i stands for the Lagrange multiplier. This equation must be constrained using following circumstances [87]:

$$\frac{\partial L}{\partial \omega} = 0 \rightarrow \omega = \sum_{k=1}^{N} \alpha_k \theta(x_k)$$ (20)

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{k=1}^{N} \alpha_k = 0$$ (21)

$$\frac{\partial L}{\partial e_k} = 0 \rightarrow \alpha_k = \gamma e_k \quad k = 1, 2, \ldots, N$$ (22)

$$\frac{\partial L}{\partial \alpha_k} = 0 \rightarrow \left(\omega \cdot \theta(x_k)\right) + b + e_k - y_k = 0 \quad k = 1, 2, \ldots, N$$ (23)

Assuming $Y = [\gamma_1, \ldots, \gamma_N]^T$, $\alpha = [\alpha_1, \ldots, \alpha_N]^T$, $l_v = [1, \ldots, l]^T$, bearing in mind that Kernel function is $\Omega_{ij} = \theta(x_i)\theta(x_j) = K(x_i \cdot x_j)$, and considering that l_n is the identity matrix [75], then:

$$\begin{bmatrix} 0 & l_n^T \\ l_n & \Omega + \gamma^{-1}l_n \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$ (24)

Kernel function can be categorized as linear, polynomial, sigmoid, and RBF functions [75]. Of all functions presented above, RBF is a well-reputed and a common Kernel function that is utilized in order to run the LS-SVM module and can be formulated as below [88]:

$$K(x_i \cdot x_j) = e^{\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right)}$$ (25)

Among all the parameters involved during LS-SVM optimization, two parameters of γ and σ^2 must be altered. Since the PSO has a great capability in ascertaining LS-SVM parameters, it is implemented throughout the LS-SVM optimization procedure.

4. Results and discussion

In this part, the outcomes of tests are presented in terms of diagrams and they have been validated under different criteria. Firstly, it must be highlighted that throughout the learning procedure, ascertaining the prediction function and asserting its validity using the same data set is a methodological misconception, that would ultimately lead to a phenomenon identified as overfitting. Thus, it is attempted to use a practical method known as cross-validation in order to use an extracted set of data points for the validation phase of the prediction process. Moreover, the particle swarm optimization technique is employed to optimize the hyper-parameters associated with the applied algorithms since they directly control the performance of algorithms.

4.1. Models' predictions outputs

The details of algorithms parameters are compiled in Table 2. The main criteria used to justify the validity of the retrieved data are statistical benchmarks including the standard deviations (STDs), the coefficient of determination (R^2), average absolute relative deviations (AARDs), and relative mean squared errors (RMSE). Formulation of each parameter is given below [75]:

$$AARD(100\%) = \frac{100}{N}\sum_{k=1}^{N}\left|\left(X_k^{predicted} - X_k^{actual}\right)\right| \tag{26}$$

$$RMSE = \left(\frac{1}{N}\sum_{k=1}^{N}\left(X_k^{predicted} - X_k^{actual}\right)^2\right)^{1/2} \tag{27}$$

$$R^2 = 1 - \frac{\sum_{k=1}^{N}\left(X_k^{predicted} - X_k^{actual}\right)^2}{\sum_{k=1}^{N}\left(X_k^{actual} - X_k^{\overline{actual}}\right)^2} \tag{28}$$

$$STD_{erro} = \sqrt{\left(\frac{1}{N-1}\right)\sum_{k=1}^{N}\left(X_k^{predicted} - X_k^{actual}\right)^2} \tag{29}$$

where predicted and actual denote the models predictions and the real extracted data.

Fig. 6 demonstrates the membership functions of the ANFIS algorithm. The membership function is defined as a distance between a given element and a standard member [89]. It is also defined as a tool to illustrate the degree of truth or false of how a crisp value is mapped to an input space [90]. Each membership function is normalized within the range of [–1,1] and they entail curvy styles, in which are named based on the shape they possess. The dependency parameter can be calculated by multiplying cluster numbers, membership function numbers, and the aggregation

Table 2
The algorithms configurations and specifications

| Specifications	Values/parameters	Specifications	Values/parameters
LS-SVM Algorithm		ANFIS Algorithm	
Kernel function	RBF-Kernel	Membership function	Triangular
Training data points percent	70	Training data points percent	80
Testing data points percent	30	Testing data points percent	20
Γ	233,084.45297	No. of MF parameters	12
$(\sigma)^2$	0.12298	No. of clusters	10
Maximum iterations	1,000	Maximum iterations	1,000
c_1	1	c_1	1
c_2	2	c_2	2
Population size	200	Population size	25
MLP Algorithm		RBF Algorithm	
No. input neuron layer	8	No. of input neuron layer	116
No. of first hidden neuron layer	4	No. of output neuron layer	1
No. of second hidden neuron layer	2	Training data points percent	85
No. of output neuron layer	1	Testing data points percent	15
Training data points percent	85	Hidden layer activation function	RBF
Testing data points percent	15	Output layer activation function	Purelin
Hidden layer activation function	Tansig[a]	No. of maximum iteration	1,000
Output layer activation function	Purelin[b]		
No. of maximum iterations	652		

[a]Tansigmoidal; [b]Purely linear
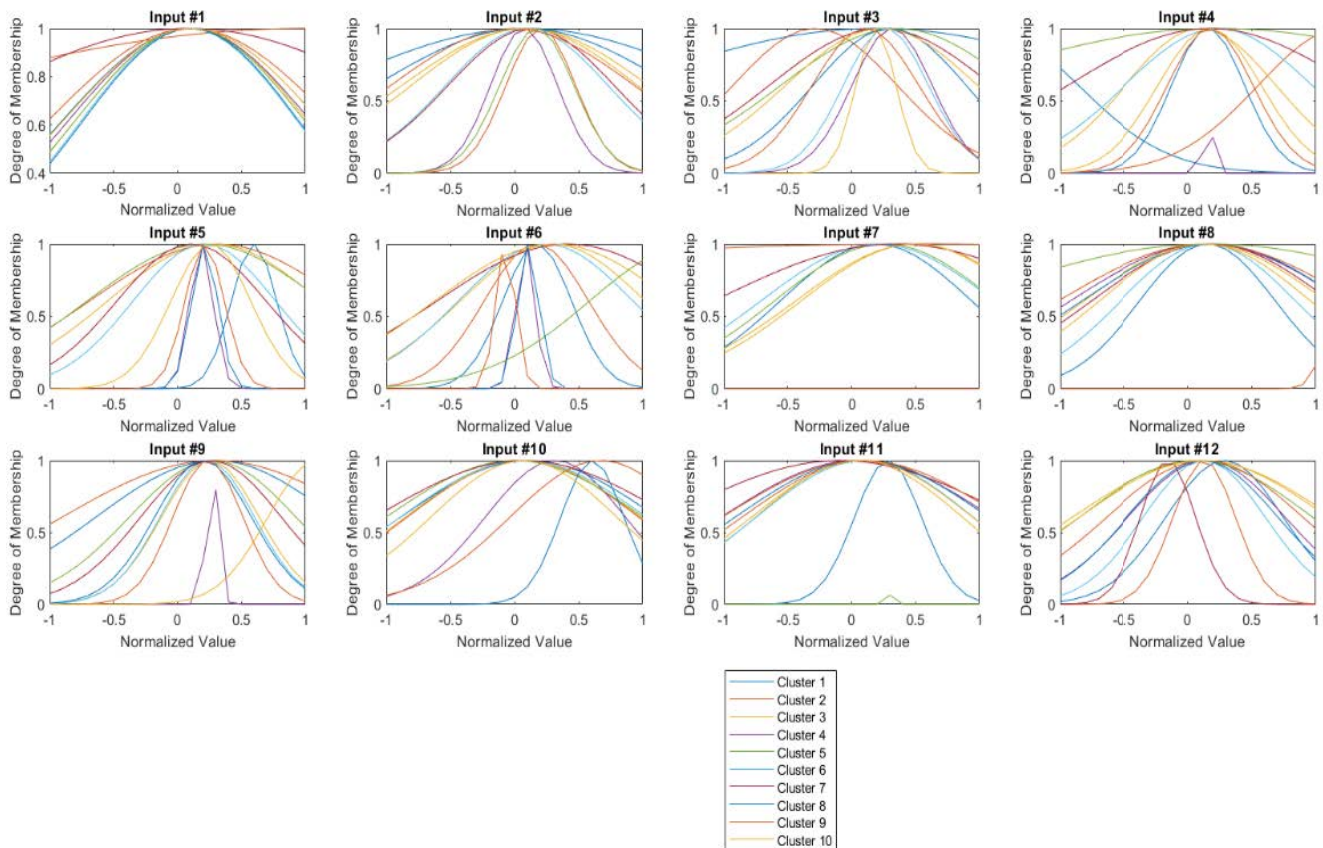
Fig. 6. Corresponding membership functions for different inputs: 1-Flowrate, 2-Dispersed phase droplet diameter, 3-Inlet width, 4-Overflow diameter, 5-Underflow diameter, 6-Cone angle, 7-Tail angle, 8-Cylindrical part diameter, 9-Initial temperature, 10-Inlets number, 11-Phases density difference and 12-Tail length.

of output and input parameters, in which for this proposed ANFIS algorithm, this dependency is 1800. Furthermore, among the available membership functions of Anfis_trimf, Anfis_trapmf, Anfis_gbellmf, Anfis_gaussmf, Anfis_gauss2mf, Anfis_pimf, Anfis_dsigmf, Anfis_psigmf, the first four types are utilized in the applied ANFIS algorithm.

Also, the integration of Anfis_gaussmf membership function was adequate enough to be used for input numbers of [1,2,7,8] (refer to Fig. 6 for the name of inputs), however, as for other inputs a combination of above-mentioned membership functions were needed to precisely incorporate the respective input, due to the complexities associated with the considered inputs [91].

Moreover, the regression plots for proposed algorithms are illustrated in Fig. 7.

The regression plots depict the experimental separation efficiencies vs. separation efficiencies measured by proposed algorithms. According to the diagrams, both trained and tested data sets are in great agreement with experimental data and they are scattered predominantly in the vicinity of the equivalent line. This feature is shown statistically in the correlated line of data points for both tested and trained data. Furthermore, the relative derivation error diagrams are presented in Fig. 8. Accordingly, the measured relative errors for all the employed algorithms are within an acceptable range. As it can be perceived from the figure,

the separation efficiencies below 0.1 tend to deviate sharply from the acceptable limits (except for the case of ANFIS model) for both trained and the tested data points. To clarify this phenomenon, it must be pointed out that hydrocyclones separation efficiency possesses complex dependency to integrated features; as such, more characteristics are required to precisely predict the separation efficiency for this particular range of separation efficiency.

Moreover, ANFIS model's outputs seem to portrays lower relative deviations compared to other models, particularly for lower values of separation efficiency. It is worthy to mention that although the number of extracted data points for the ANFIS algorithms in this range of separation efficiency is relatively lower than other models (as it can also be perceived from Fig. 7), however, to the best knowledge of the authors the ability of ANFIS algorithm to allocate different and also combined membership functions to better model the given inputs is accountable for this superior feature.

Another common scheme to assess the precision of proposed models is the simultaneous presentation of both experimental and predicted separation efficiency values. Fig. 9 portrays a graphical comparison between the two aforementioned values, in which the outputs stand for experimental values and targets represent the predicted data points. To make the comparison more convenient, the target and output data are sorted analogously to have
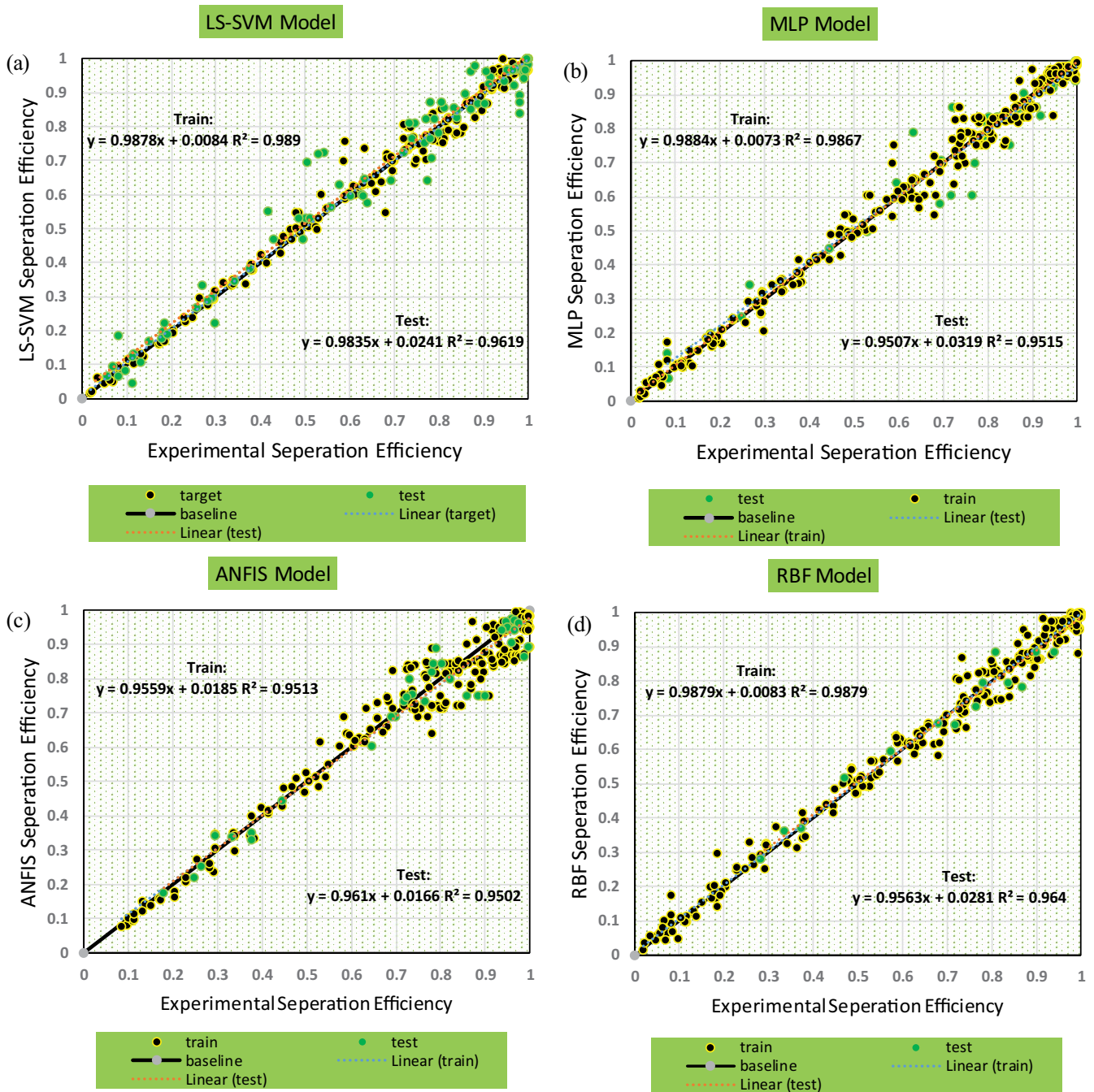
LS-SVM Model

(a)
Train:
y = 0.9878x + 0.0084 R² = 0.989

Test:
y = 0.9835x + 0.0241 R² = 0.9619

LS-SVM Seperation Efficiency / Experimental Seperation Efficiency

Legend: target, test, baseline, Linear (target), Linear (test)

MLP Model

(b)
Train:
y = 0.9884x + 0.0073 R² = 0.9867

Test:
y = 0.9507x + 0.0319 R² = 0.9515

MLP Seperation Efficiency / Experimental Seperation Efficiency

Legend: test, train, baseline, Linear (test), Linear (train)

ANFIS Model

(c)
Train:
y = 0.9559x + 0.0185 R² = 0.9513

Test:
y = 0.961x + 0.0166 R² = 0.9502

ANFIS Seperation Efficiency / Experimental Seperation Efficiency

Legend: train, test, baseline, Linear (train), Linear (test)

RBF Model

(d)
Train:
y = 0.9879x + 0.0083 R² = 0.9879

Test:
y = 0.9563x + 0.0281 R² = 0.964

RBF Seperation Efficiency / Experimental Seperation Efficiency

Legend: train, test, baseline, Linear (train), Linear (test)

Fig. 7. Regression plot for (a) LSSVM algorithm, (b) MLP algorithm, (c) ANFIS algorithm, and (d) RBF algorithm.
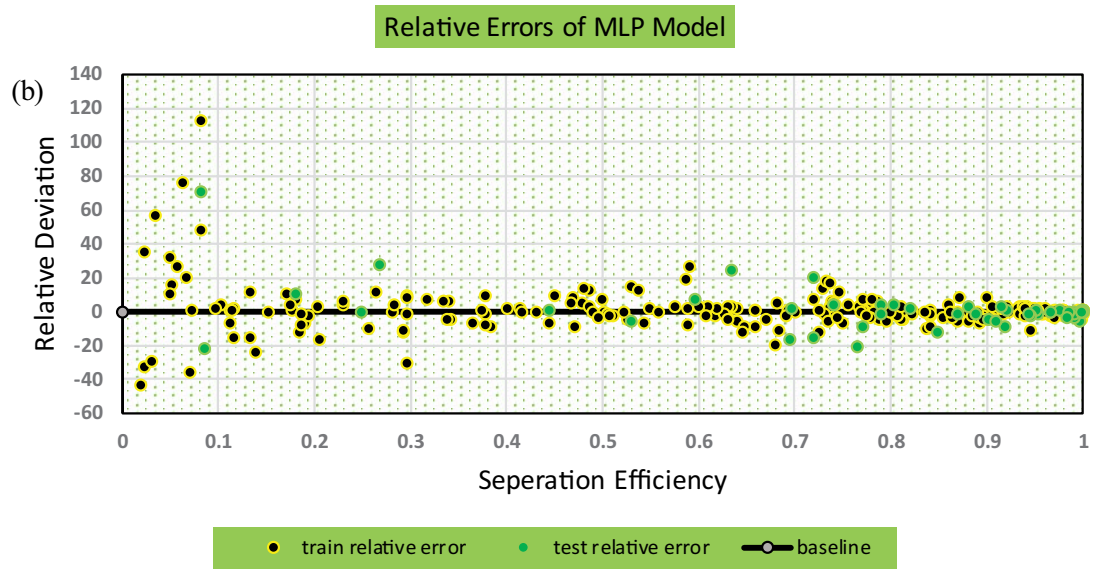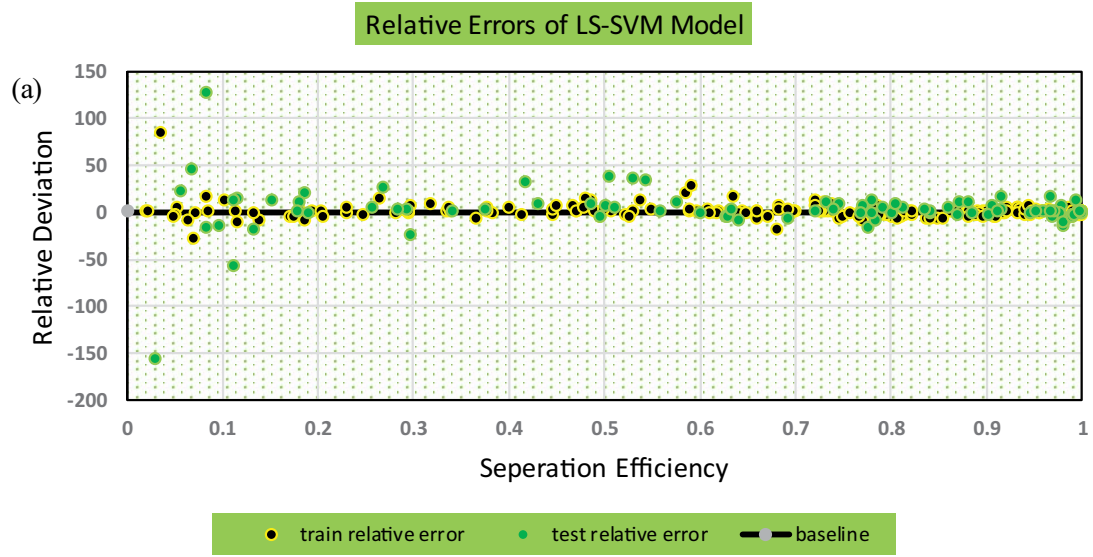
an illustrative same trend line. These graphs can come in handy in finding what data indexes contain deviated results. As such, each data index refers to a particular column in the dataset category.

To summarize the utilized criteria to analyze the performance of each algorithm, AARDs, RMSE, and the coefficient of determination (R^2) are given in Table 3. Concisely, it is indisputable that all the algorithms performed closely. However, further scrutiny reveals that RBF and MLP algorithms showed generally an incredible ability in predicting the separation efficiency. Also, as it was ascertained earlier,

although all the algorithms performed satisfactory enough, in the case if the occurrence of low separation efficiencies are possible, it is recommended to utilize the developed ANFIS model since it was previously perceived to give more reliable outcomes in these domains.

4.2. Comparative performance analysis of the proposed models

In this part, the performances of the applied models are compared to further scrutinize the behavioral functionality of each model in respect to different benchmarks.
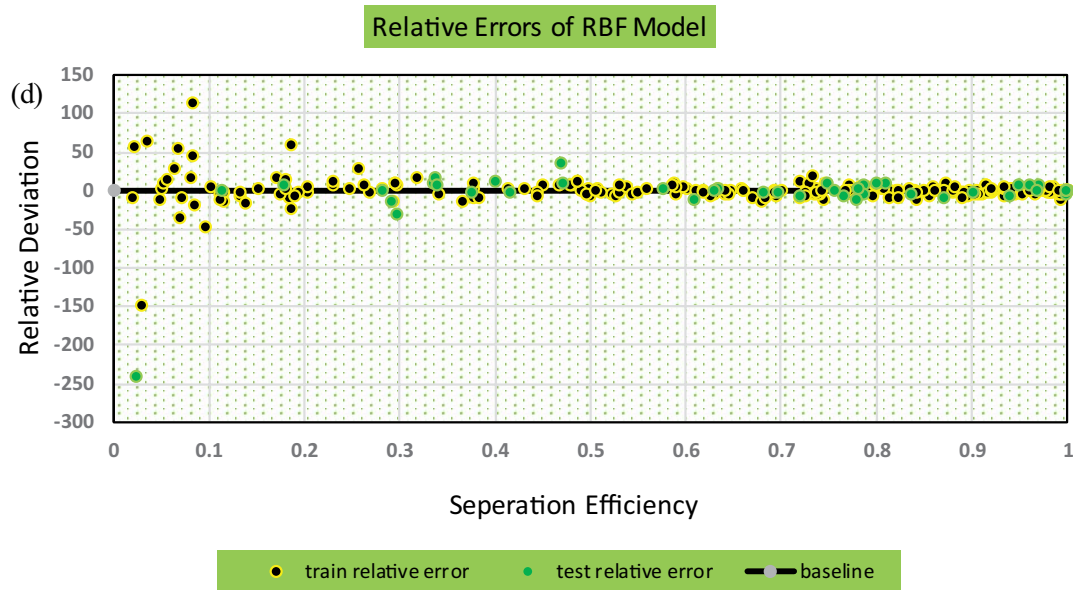
Relative Errors of LS-SVM Model

(a)

train relative error   test relative error   baseline

Relative Errors of MLP Model

(b)

train relative error   test relative error   baseline

Relative Errors of ANFIS Model

(c)

train relative error   test relative error   baseline

Continued

Relative Errors of RBF Model
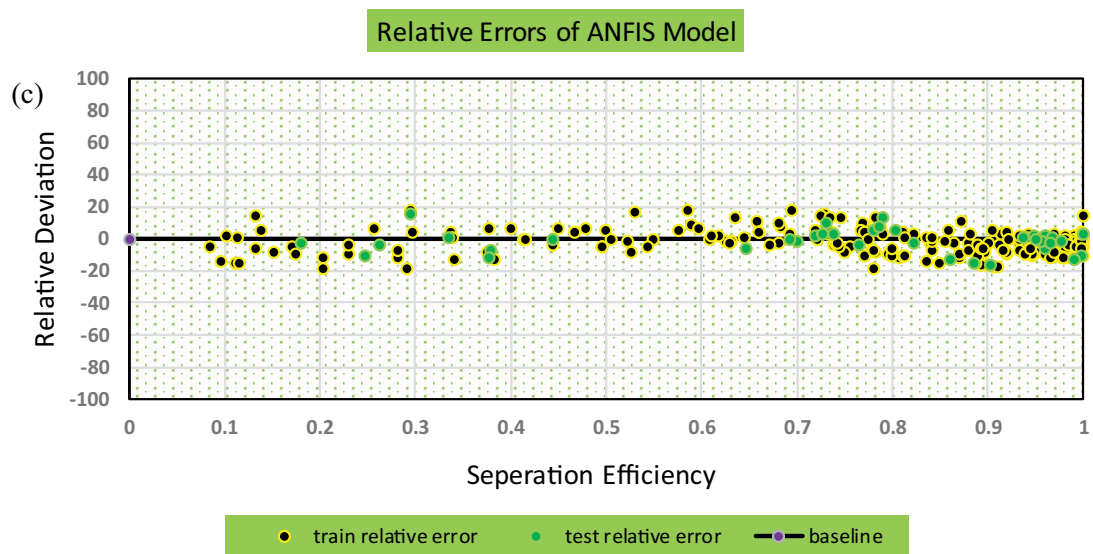
Fig. 8. Relative error diagram of (a) LS-SVM algorithm, (b) MLP algorithm, (c) ANFIS algorithm, and (d) RBF algorithm.

As such, an in-depth analysis into the Akaike Information Criterion as well as the time complexity is illustrated.

4.2.1. Akaike Information Criterion

In addition to the previous performance criteria, Akaike Information Criterion (AIC) is another useful metric that can provide a suitable insight into the accuracy of the proposed models while taking the complexity associated with the models into account [92]. In other words, as far as the accuracy of models is concerned, a model with lower required parameters to optimize for obtaining the high level of accuracy is more demanded [93]. Furthermore, the lower the measured AIC value, the more convenient the model performs in terms of accuracy and complexity. AIC is calculated as follows [92]:

$$\text{AIC} = n \times \log\left(\sqrt{\text{RMSE}}\right) + 2 \times k \tag{30}$$

where n, k, and RMSE stand for the number of samples, number of optimized parameters, and the root mean square error, respectively. The AIC values of each proposed model is depicted in Fig. 10. As it can be perceived form the figure, the LSSVM and the RBF models revealed better performance relatively by owing lower AICs followed by MLP and ANFIS models. This is because LSSVM algorithm deals with only two hyper-parameters as shown in Table 2. However, in case of other algorithms, more hyper-parameters are involved throughout the training and testing process. This would lead to a more complex structure comparatively.

4.2.2. Time complexity

The runtime can be used as a comparative criterion to further analyze the performance of the proposed algorithms. As such, the time required for the models to meet the specified limit is deemed as the runtime. The runtime of each model is presented in Fig. 11. First of all, it must be hinted that the models failed to meet the defined criterion ($R^2 = 0.97$) for input data sizes of less than 10. This is in line with Fig. 4, as for most of the combination of inputs with data sizes of lower than 10, the accuracy of the results possessed noticeably lower values. To the best of author's knowledge, the plausible explanation for this traces back to the complex behavior of hydrocyclones separation efficiency and its dependency to the selected features. Furthermore, although feeding 10 inputs would finally meet the satisfactory limit, yet again, considerable duration of time is needed.

On the other hand, as the inputs size increments, the runtime is reduced dramatically. To begin with, generally the runtime is expected to increase commensurably by the increase in the data size. This is because more time and space are required to load, process (train, test, and validate), and optimize the data. However, in some cases the inputs may carry specific features with high or low significance, which lead the algorithms to behave differently.

In other words, those features can help the algorithm in a swifter recognition of the target specification. Thus, as the inputs increases to 12, the runtime still drops, nonetheless, as for further increment in data size, the runtime seems to slightly raise again. Bearing in mind the results of the sensitivity and inputs combination analyses (Figs. 3 and 4), it can be inferred that apparently the previously detected less-effective features (cylindrical part length and initial pressure) have led the models to iterate for a longer time to meet the criterion limit. Other than this, comparing the algorithms runtime reveals that for lower data size, LSSVM algorithm requires higher duration of time (please note the time is reported in logarithmic scale). This can be attributed to the lower number of hyper-parameters (as given in Table 2) that are adjusted throughout the LSSVM model to optimize the outputs. Which means the complexity of hydrocyclone separation efficiency behavior for low-given inputs

Total Data Set-LS-SVM Model

(a)

Total Data Set-MLP Model

(b)

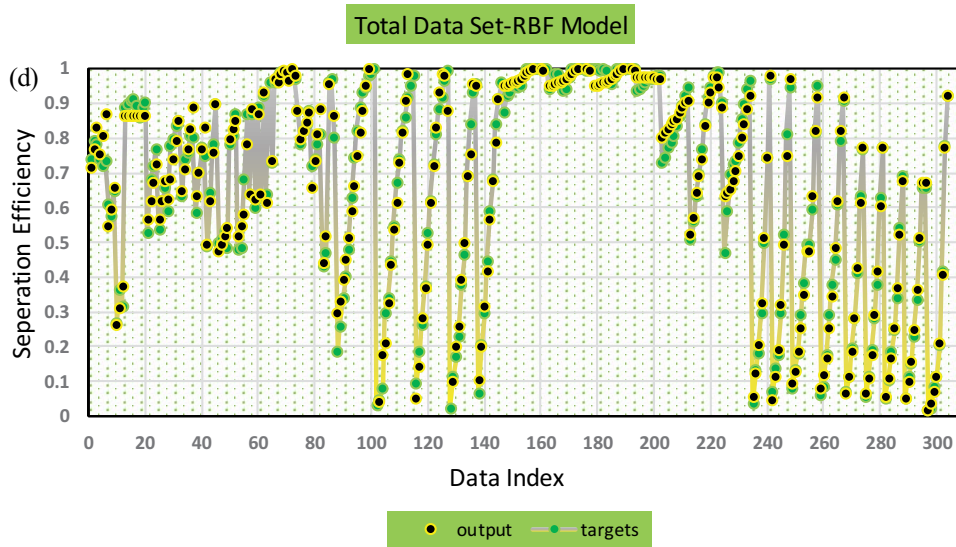**Total Data Set-ANFIS Model**

(c)
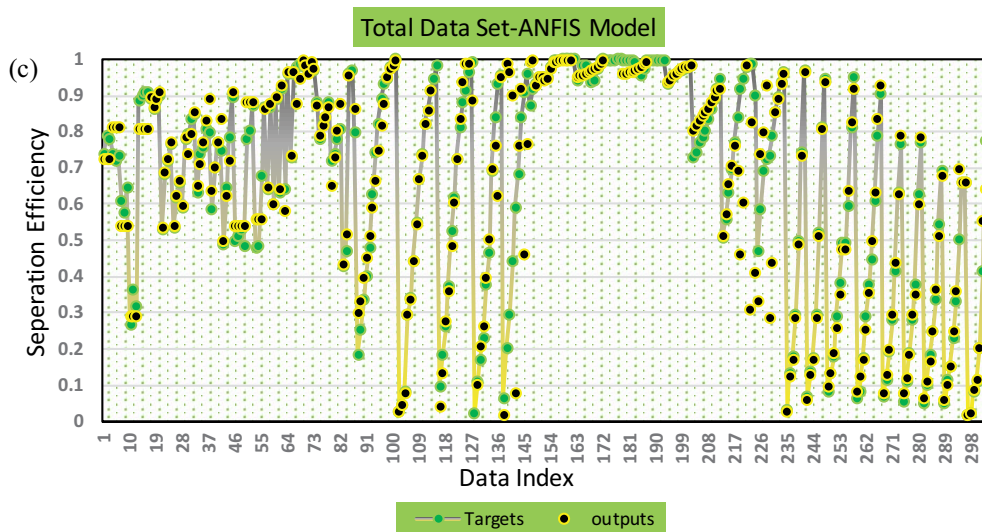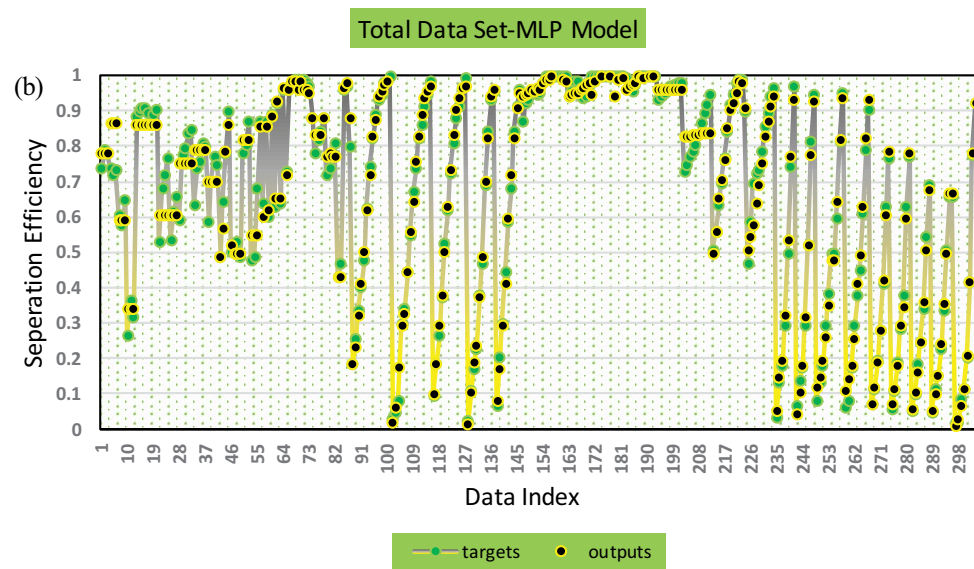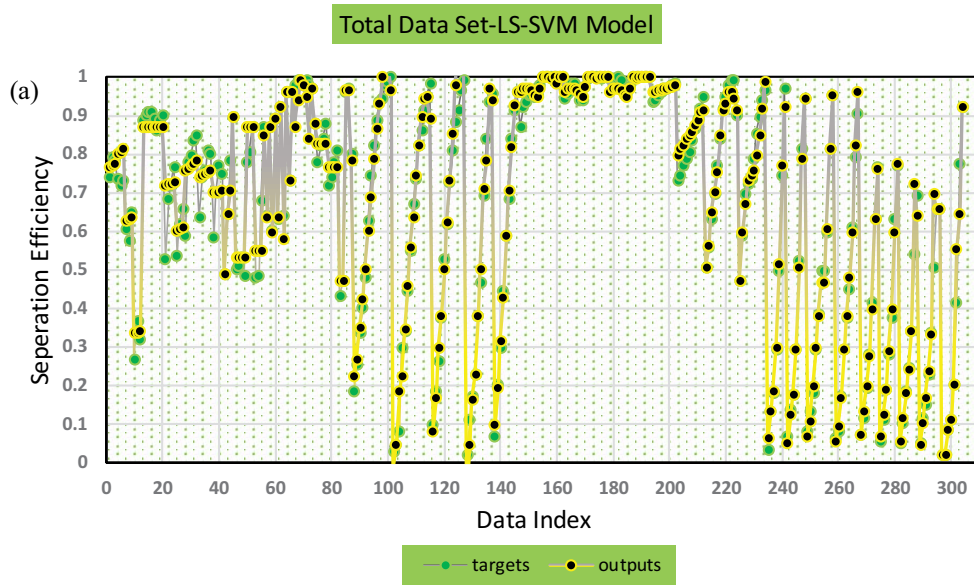
Continued

Fig. 9. Correspondence of experimental and predicted separation efficiencies of (a) LS-SVM model, (b) MLP model, (c) ANFIS model, and (d) RBF model.

Table 3
Summary of utilized algorithms statistical results

Algorithm	Data	R^2	AARD	RMSE
	Total	0.97914	7.764	0.043461
LS-SVM	Train	0.989	3.6124	0.030566
	Test	0.9596	11.7807	0.06336
	Total	0.98316	6.3588	0.039052
MLP	Train	0.9867	6.2114	0.03514
	Test	0.95076	7.1856	0.056154
	Total	0.94795	6.1277	0.058444
ANFIS	Train	0.94795	6.2233	0.058733
	Test	0.94772	5.5648	0.056712
	Total	0.98492	7.4802	0.036953
RBF	Train	0.98794	6.7421	0.033601
	Test	0.96339	11.6202	0.051884

is time-consuming to be predicted well-enough using only 2 hyper-parameters to optimize. Contrarily, other models portrayed better performance owing to higher number of hyper-parameters that can effectively contribute to find the best perdition promptly. However, this shortcoming tends to fade away for LSSVM model as the input size increases (in comparison to other models) since it revealed good controlling ability over the noises associated with the less-effective input data sets (data sizes of 13 and 14, which carry the cylindrical part length and initial pressure).

Another way of reporting the time complexity for machine leaning algorithms is known as big-o-notation, which illustrates the runtime as a function of $O(x)$. X is commonly allocated to the worst part of the algorithms in terms of runtime that can come in various formats vs. x. A conceivable deduction for the LS-SVM algorithm runtime trend for data sizes of further than this study is $O(1)$, which denotes a constant limit of runtime by the increase of data size. Nevertheless, in case of other algorithms, it is hard to reach to a conclusion about the general trend of the runtime since they fluctuate erratically for this particular data set. Yet again, since they have been proven sensitive to further data size, it is reasonable to hypothesize that eventually the data size increment would cause the runtime to increase.

Moreover, the big-o-notation usually takes all the feature inputs at once, and the number of samples will vary throughout the analysis. However, it is also attempted to conduct the test with this calibration as well, but the runtime trend remained the same with very minimal difference compared to Fig. 11.

5. Practical applications of the proposed models

The practical applications of the present study can be taken up on two levels. First and foremost, as specified earlier, hydrocyclones are being applied in various industries, and there are numerous surveys to calibrate or adjust some of its features to either use its apex of functionality or adapt it to the intended operating conditions. However, these examinations could bring about cumbersome measurements and cost- and time-consuming studies, either experimentally and computationally.

Moreover, the chief intention behind executing such investigations is to probe the efficiency of hydrocyclones separation process, as this is the core utility of hydrocyclones; also, in most ongoing projects, the geometrical and operating conditions are already known, but it can be highly inconvenient to apply all these diverse calibrations to find the most favored structure and applied circumstances. Additionally, it is worth mentioning that hydrocyclones mainly tend to perform quite erratically by a slight change in their structures or operating conditions due to severe turbulent intensities they inherit. This obstacle makes it even harder to reach an optimized hydrocyclone model and applied condition. Therefore, the proposed models of this study can contribute effectively by taking the previously
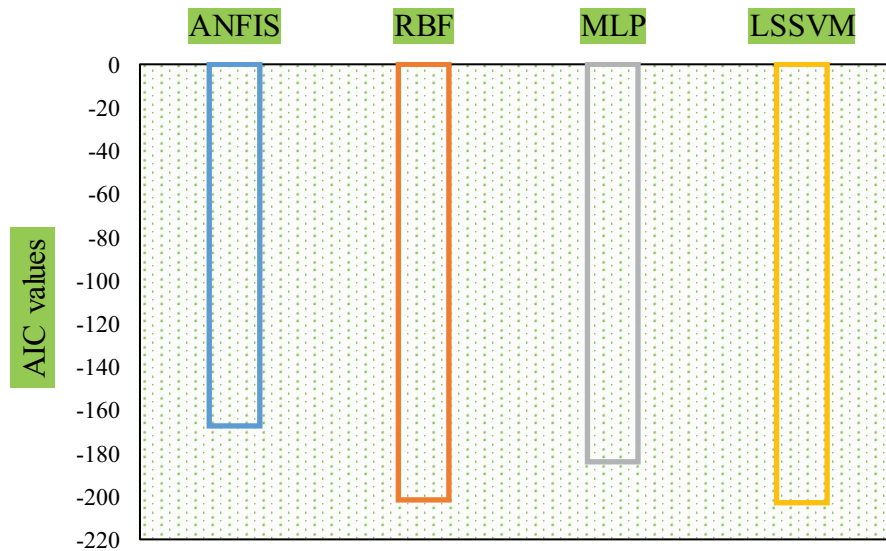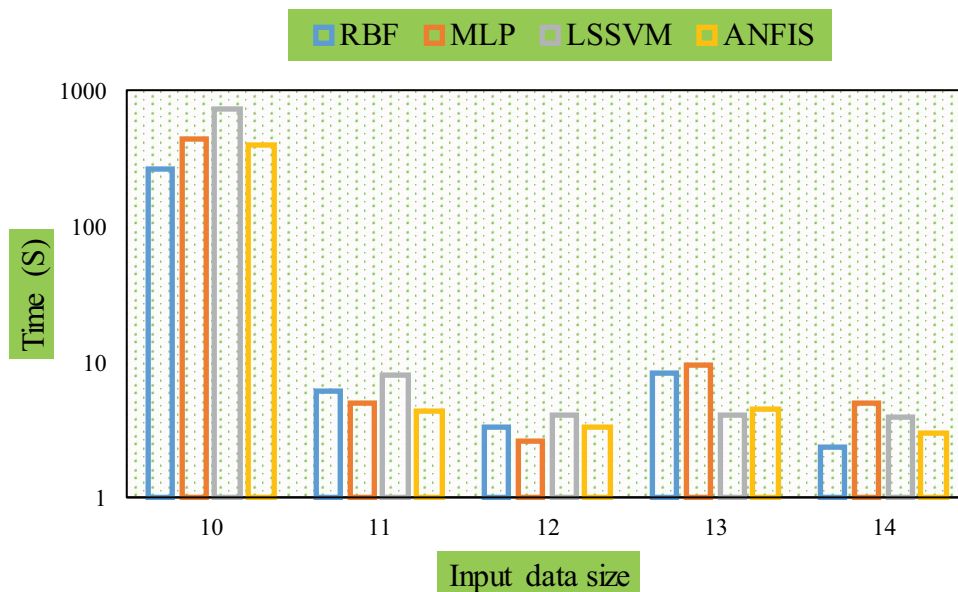
Fig. 10. The AIC values of applied algorithms.

Fig. 11. Total runtime of proposed models vs. the increment in data size.

mentioned parameters as inputs and present a reliable prediction of the separation efficiency without the need to actually build the instrument. This allows the researchers to save a considerable amount of time and resources, particularly for the manufacturing industry, while narrowing down their case studies into those specific conditions that their intended hydrocyclone is viable to perform satisfactorily.

On the other hand, of the main industries that utilize this apparatus, the petroleum industry is considered the most interested sector. The main motives to use this instrument were discussed in the previous parts, however, their simple installation, high efficiency, portability, easier cleaning process, cost convenient manufacturing, and more importantly, long durability under server conditions are the main targeted ambitions in petroleum industry. To date, there is

an emerging type of hydrocyclones that are installed in oil-gas wells (for both on-shore and off-shore rigs) known as downhole de-oiling hydrocyclones [2,29,94–99]. This newly introduced approach is proven to save huge amount of expenses and lower down the run-time of downstream sectors (petroleum refineries). A detailed insight into this novel approach is demonstrated in the work of Zandie et al. [2]. However, experience to date has proven that the down holes operating conditions (which are hard to control) are highly fluctuating because of natural intrinsic behaviors of petroleum reservoirs.

On the other hand, since hydrocyclones are mainly connected to the main production string, any failure in hydrocyclone functionality can disrupt the whole production process. For example, if the hydrocyclone is working with

very low separation efficiencies, choking (severe instability in separation process) is very likely to occur, and the production system will be affected as a consequence. For this reason, the provision of a tool to predict the performance of these hydrocyclones under different downhole operating conditions in a very instantaneous manner (to allow prompt reactions of involved inspectors) are of utmost necessities. Of the important downhole operating conditions, flowrate, dispersed phase droplet diameter, temperature, and phases density difference can be cited, which are incorporated in the present study. Henceforth, it is plausible to conclude that the present suggested models can help gaining a better control over the above-mentioned concerns as well.

6. Computational limitations

Certainly, the shortcomings of this study in terms of practical or computational limitations must also be pointed out for suitable utilization of these models as well as presetting an unbiased analysis. To begin with, bearing in mind the parameters taken up as inputs it can be inferred that there are still some other features that can be integrated. Of the most important properties, surface tension between the phases, internal roughness of hydrocyclone walls, vortex finder dimensions, inlets angle, and more importantly phases fractions can be mentioned. Unquestionably, the results of the present study could have been more robust provided that these elements were also integrated in the data bank. However, as it was declared in the data gathering section, the data bank must focus on those features that have been explored extensively, so a reliable amount of data can be extracted from literature. Otherwise, that parameter must be omitted from the list. For this reason, the present study in exclusively focused on the aforementioned geometrical and dynamical parameters.

Moreover, the size of the employed data set can also be concerned. To be more exact, it is generally believed that the number of data samples must increment commensurately as the number of inputs and targets increments. Having considered 14 variables for this paper might increase the requisite for higher number of data points, however, the lack of adequate data is the main impediment to do so. Admittedly, two important hints can be suggested to address this matter. Firstly, the present data bank can be supposed sufficient enough if less number of variables is chosen. Nevertheless, to the best knowledge of authors (as it is acknowledged in data gathering part), there is a strong dependence between the chosen parameters as such the exclusion of each one lowers down the prediction capability of applied machine learning approaches, and increases the run-time considerably. Thus, any utilization of the proposed models is dependent on the provision all the stated variables, otherwise, the models fail to present any form of data.

In addition, as it has been pointed out in previous parts, the applied algorithms results tend to deviate noticeably from the acceptable range for separation efficiencies lower than 0.1. This is crucial to be resolved in further researches because this range of separation efficiency is claimed to be practically important (please refer to section 5) and the failure of the proposed models in accurately predicting this range of separation efficiency might be considered a deficiency. Henceforth, according to the results discussed earlier, the integration of ANFIS algorithm is viable to resolve this concern, however, the feasibility of other models in that particular range of separation efficiency can be questionable.

7. Conclusions

In this investigation, four versions of machine learning approaches including ANFIS, LS-SVM, MLP, and RBF are employed to predict the separation efficiency of hydrocyclones as a function of 14 features. To boost the reliability of the analyses, an all-embracing data set entailing 304 data samples, 4,560 data points to be more exact, is collected. Furthermore, sensitivity analysis is carried out to present a comprehensive insight into the parameters' degree of impacts on the separation efficiency. The outlier detection technique is also applied for imprecise data points removal.

The present paper attempts to provide an insight and a clear perception into the application of machine learning approaches in predicating hydrocyclones separation efficiency as well as offering the following researches a suitable orientation in practical utilization of hydrocyclones, meanwhile saving considerable amount of time and costs. An important suggestion for following researches in extending the application of machine learning algorithms in predicting hydrocyclones performance can be exemplified under two major scenarios. Firstly, since the hidden layers and nodes in the machine learning algorithms are vital to gain reliable results, introducing algorithms that can predict the overall performance of hydrocyclones with an automatic adoption of hidden layer configuration in terms of layer arrangement, associated number of nodes, in addition to hyper- and primary-parameters automatic optimization is highly demanded. Secondly, the provision of a new data set containing other features viz. feed phases viscosities, surface tension between the phases, internal roughness of hydrocyclone walls, vortex finder dimensions, inlets angle, and more importantly dispersed phase volume fraction integrated with more target elements such as pressure drop, and testing it on the proposed algorithms (major adjustment in algorithms coding is required) are viable to bring about substantial contributions since they are deemed as important parameters in separation efficiency of hydrocyclones.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

References

[1] K.U. Bhaskar, Y.R. Murthy, M.R. Raju, S. Tiwari, J.K. Srivastava, N. Ramakrishnan, CFD simulation and experimental validation studies on hydrocyclone, Miner. Eng., 20 (2007) 60–71.

[2] M. Zandie, A. Kazemi, M. Ahmadi, M.K. Moraveji, A CFD investigation into the enhancement of down-hole de-oiling hydro cyclone performance, J. Pet. Sci. Technol., 199 (2021) 108352, doi: 10.1016/j.petrol.2021.108352.

[3] A. Belaidi, M. Thew, The effect of oil and gas content on the controllability and separation in a de-oiling hydrocyclone, Chem. Eng. Res. Des., 81 (2003) 305–314.

[4] C.L. Karr, D.A. Stanley, B. McWhorter, Optimization of hydrocyclone operation using a geno-fuzzy algorithm, Comput. Methods Appl. Mech. Eng., 186 (2000) 517–530.

[5] M. Karimi, A. Dehghani, A. Nezamalhosseini, S. Talebi, Prediction of hydrocyclone performance using artificial neural networks, J. S. Afr. Inst. Min. Metall., 110 (2010) 207–212.

[6] S. van Loggenberg, G. van Schoor, K. Uren, A. van der Merwe, Hydrocyclone cut-size estimation using artificial neural networks, IFAC-PapersOnLine, 49 (2016) 996–1001.

[7] C. Fung, K. Wong, H. Eren, Developing a Generalised Neural-Fuzzy Hydrocyclone Model for Particle Separation, IMTC/98 Conference Proceedings. IEEE Instrumentation and Measurement Technology Conference. Where Instrumentation is Going (Cat. No.98CH36222), IEEE, St. Paul, MN, USA, 1998.

[8] H. Eren, C.C. Fung, K.W. Wong, A. Gupta, Use of Artificial Neural Networks in Estimation of Hydrocyclone Parameters with Unusual Input Variables, Quality Measurement: The Indispensable Bridge between Theory and Reality (No Measurements? No Science! Joint Conference – 1996: IEEE Instrumentation and Measurement Technology Conference and IMEKO Tec, IEEE, Brussels, Belgium, 1996.

[9] K.W. Wong, Y.S. Ong, H. Eren, C.C. Fung, Hybrid Fuzzy Modelling Using Memetic Algorithm for Hydrocyclone Control, Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826), IEEE, Shanghai, China, 2004.

[10] S. Mohanty, S.K. Das, A.K. Majumder, Artificial neural network modeling and experimental investigation to characterize the dewatering performance of a hydrocyclone, Miner. Process. Extr. Metall., (2019) 1–13, doi: 10.1080/25726641.2019.1680177.

[11] B. Yang, J. Wang, X. Zhang, T. Yu, W. Yao, H. Shu, F. Zeng, L. Sun, Comprehensive overview of meta-heuristic algorithm applications on PV cell parameter identification, Energy Convers. Manage., 208 (2020) 112595.

[12] J.A. Jervase, H. Bourdoucen, A. Al-Lawati, Solar cell parameter extraction using genetic algorithms, Meas. Sci. Technol., 12 (2001) 1922.

[13] K. Ishaque, Z. Salam, An improved modeling method to determine the model parameters of photovoltaic (PV) modules using differential evolution (DE), Sol. Energy, 85 (2011) 2349–2359.

[14] A. Askarzadeh, A. Rezazadeh, Artificial bee swarm optimization algorithm for parameters identification of solar cell models, Appl. Energy, 102 (2013) 943–949.

[15] S. Mirjalili, A. Lewis, The whale optimization algorithm, Adv. Eng. Software, 95 (2016) 51–67.

[16] Z. Wu, D. Yu, X. Kang, Parameter identification of photovoltaic cell model based on improved ant lion optimizer, Energy Convers. Manage., 151 (2017) 107–115.

[17] M.A. Awadallah, Variations of the bacterial foraging algorithm for the extraction of PV module parameters from nameplate data, Energy Convers. Manage., 113 (2016) 312–320.

[18] B. Yang, X. Zhang, T. Yu, H. Shu, Z. Fang, Grouped grey wolf optimizer for maximum power point tracking of doubly-fed induction generator based wind turbine, Energy Convers. Manage., 133 (2017) 427–443.

[19] K. Passino, Biomimicry of bacterial foraging for distributed optimization and control, IEEE Control Syst. Mag., 22 (2002) 52–67.

[20] A. Diabat, D. Kannan, M. Kaliyan, D. Svetinovic, An optimization model for product returns using genetic algorithms and artificial immune system, Resour. Conserv. Recycl., 74 (2013) 156–169.

[21] B. Yang, L. Zhong, X. Zhang, H. Shu, T. Yu, H. Li, L. Jiang, L. Sun, Novel bio-inspired memetic salp swarm algorithm and application to MPPT for PV systems considering partial shading condition, J. Cleaner Prod., 215 (2019) 1203–1222.

[22] X. Yuan, Y. Yang, H. Wang, Improved parallel chaos optimization algorithm, Appl. Math. Comput., 219 (2012) 3590–3599.

[23] N. Pourmousa, S.M. Ebrahimi, M. Malekzadeh, M. Alizadeh, Parameter estimation of photovoltaic cells using improved Lozi map based chaotic optimization algorithm, Sol. Energy, 180 (2019) 180–191.

[24] X. Yuan, J. Zhao, Y. Yang, Y. Wang, Hybrid parallel chaos optimization algorithm with harmony search algorithm, Appl. Soft Comput., 17 (2014) 12–22.

[25] M. AlRashidi, M. AlHajri, K. El-Naggar, A. Al-Othman, A new estimation approach for determining the I–V characteristics of solar cells, Sol. Energy, 85 (2011) 1543–1550.

[26] R.C.M. Gomes, M.A. Vitorino, M.B. de Rossiter Correa, D.A. Fernandes, R. Wang, Shuffled complex evolution on photovoltaic parameter extraction: a comparative analysis, IEEE Trans. Sustainable Energy, 8 (2016) 805–815.

[27] K. Yu, J. Liang, B. Qu, X. Chen, H. Wang, Parameters identification of photovoltaic models using an improved JAYA optimization algorithm, Energy Convers. Manage., 150 (2017) 742–753.

[28] K.M. El-Naggar, M. AlRashidi, M. AlHajri, A. Al-Othman, Simulated annealing algorithm for photovoltaic parameters identification, Sol. Energy, 86 (2012) 266–274.

[29] S.M. Hosseini, K. Shahbazi, M.R. Khosravi Nikou, A CFD simulation of the parameters affecting the performance of downhole de-oiling hydrocyclone, Iran. J. Oil Gas Sci. Technol., 4 (2015) 77–93.

[30] M. Bennett, R.A. Williams, Monitoring the operation of an oil/water separator using impedance tomography, Miner. Eng., 17 (2004) 605–614.

[31] A. Hoffmann, M. De Groot, W. Peng, H. Dries, J. Kater, Advantages and risks in increasing cyclone separator length, AlChE J., 47 (2001) 2452–2460.

[32] G. Young, W. Wakley, D. Taggart, S. Andrews, J. Worrell, Oil-water separation using hydrocyclones: an experimental search for optimum dimensions, J. Pet. Sci. Eng., 11 (1994) 37–50.

[33] S. Bernardo, M. Mori, A. Peres, R. Dionisio, 3-D computational fluid dynamics for gas and gas-particle flows in a cyclone with different inlet section angles, Powder Technol., 162 (2006) 190–200.

[34] C. Gomez, J. Caldentey, S. Wang, L. Gomez, R. Mohan, O. Shoham, Oil-Water Separation in Liquid–liquid Hydrocyclones (LLHC)-experiment and Modeling, SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers, New Orleans, Louisiana, 2001.

[35] J.A. Delgadillo, R.K. Rajamani, A comparative study of three turbulence-closure models for the hydrocyclone problem, Int. J. Miner. Process., 77 (2005) 217–230.

[36] G. Patra, S. Chakraborty, B. Meikap, Role of vortex finder depth on pressure drop and performance efficiency in a ribbed hydrocyclone, S. Afr. J. Chem. Eng., 25 (2018) 103–109.

[37] K. Elsayed, C. Lacor, Analysis and Optimisation of Cyclone Separators Geometry Using RANS and LES Methodologies, M.O. Deville, J.-L. Estivalezes, V. Gleize, T.-H. Lê, M. Terracol, S. Vincent, Eds., Turbulence and Interactions, Proceedings of the TI 2012 Conference, Springer, Berlin, Heidelberg, 2014, pp. 65–74.

[38] T. Monredon, K. Hsieh, R.K. Rajamani, Fluid flow model of the hydrocyclone: an investigation of device dimensions, Int. J. Miner. Process., 35 (1992) 65–83.

[39] M. Bohnet, Influence of the gas temperature on the separation efficiency of aerocyclones, Chem. Eng. Process. Process Intensif., 34 (1995) 151–156.

[40] S.M. Vahedi, F. Parvaz, R. Rafee, M. Khandan Bakavoli, Computational fluid dynamics simulation of the flow patterns and performance of conventional and dual-cone gas-particle cyclones, Int. J. Heat Mass Transfer Res., 5 (2018) 27–38.

[41] D.A. Colman, M.T. Thew, Cyclone Separator, Google Patents, 1980.

[42] W.H. Koch, W. Light, New design approach boosts cyclone efficiency, Chem. Eng., 84 (1977) 80–88.

[43] M.R. Jadhav, Design of cyclone and study of its performance parameters, Int. J. Mech. Eng. Rob. Res., 3 (2014) 247.

[44] B.R.R. Dere, G.M. Babu, A.D. Sree, S.R. Rao, Design and analysis of cyclone dust separator, Int. J. Eng. Res. Technol. (IJERT), 3 (2014) 2278–0181.

[45] B. Zhao, H. Shen, Y. Kang, Development of a symmetrical spiral inlet to improve cyclone separator performance, Powder Technol., 145 (2004) 47–50.

[46] R. Xiang, S. Park, K. Lee, Effects of cone dimension on cyclone performance, J. Aerosol Sci., 32 (2001) 549–561.

[47] A. Gil, L.M. Romeo, C. Cortes, Effect of the solid loading on a PFBC cyclone with pneumatic extraction of solids, Chem. Eng. Technol., 25 (2002) 407–415.

[48] P. Soison, P. Supachart, P. Wongsarivej, Effect of Feed-Flow Rate in a Solid-Liquid Hydrocyclone Based on Total Solid Recovery Equation, P. Pinwanich, A. Soisungval, Eds., Key Engineering Materials (Vol. 751, pp. 173–179), Trans Tech Publications Ltd., Switzerland, 2017.

[49] A.B. Sinker, M. Humphris, N. Wayth, Enhanced Deoiling Hydrocyclone Performance Without Resorting to Chemicals, SPE Offshore Europe Oil and Gas Conference and Exhibition, Society of Petroleum Engineers, Aberdeen, United Kingdom, 1999.

[50] A.C. Hoffman, L.E. Stein, A.C. Hoffmann, L.E. Stein, Gas Cyclones and Swirl Tubes: Principles, Design, and Operation, Springer-Verlag, Berlin, Heidelberg, 2002.

[51] N. Kharoua, L. Khezzar, Z. Nemouchi, Hydrocyclones for de-oiling applications—a review, J. Pet. Sci. Technol., 28 (2010) 738–755.

[52] A. Lynch, T. Rao, K. Prisbrey, The influence of hydrocyclone diameter on reduced-efficiency curves, Int. J. Miner. Process., 1 (1974) 173–181.

[53] W. Wei, Y. Jiu-yang, Z. Xiao-tao, L. Xia, L. Wei, A new method for predicting the hydrocyclone efficiency with the light dispersed phase, Energy Procedia, 105 (2017) 4428–4435.

[54] A.C. Stone, Oil/Water Separation in a Novel Cyclone Separator, School of Engineering (SoE) (2001–July 2014), Library of University for Ph.D. and Masters Theses (SoE), Cranfield University, England, 2007.

[55] J. Martinez-Benet, J. Casal, Optimization of parallel cyclones, Powder Technol., 38 (1984) 217–221.

[56] J.-Y. Lin, R.-M. Wu, Three output membrane hydrocyclone: classification and filtration, Molecules, 24 (2019) 1116, doi: 10.3390/molecules24061116.

[57] R. Razavi, A. Sabaghmoghadam, A. Bemani, A. Baghban, K.-w. Chau, E. Salwana, Application of ANFIS and LSSVM strategies for estimating thermal conductivity enhancement of metal and metal oxide based nanofluids, Eng. Appl. Comput. Fluid Mech., 13 (2019) 560–578.

[58] A. Bemani, Q. Xiong, A. Baghban, S. Habibzadeh, A.H. Mohammadi, M.H. Doranehgard, Modeling of cetane number of biodiesel from fatty acid methyl ester (FAME) information using GA-, PSO-, and HGAPSO-LSSVM models, Renewable Energy, 150 (2020) 924–934.

[59] E. Khamehchi, A. Bemani, Prediction of pressure in different two-phase flow conditions: machine learning applications, Measurement, 173 (2021) 108665, doi: 10.1016/j.measurement.2020.108665.

[60] H. Azimi, H. Bonakdari, I. Ebtehaj, Sensitivity analysis of the factors affecting the discharge capacity of side weirs in trapezoidal channels using extreme learning machines, Flow Meas. Instrum., 54 (2017) 216–223.

[61] Y.-F. Chang, Study of the Flow in a Hydrocyclone Using Positron Emission Particle Tracking and Computational Fluid Dynamics Simulation, Research Institution in Bergen, The University of Bergen, Faculty of Mathematics and Natural Sciences, Department of Physics and Technology, Thesis Libraries, Norway, 2016.

[62] K.A. Hashmi, H.A. Hamza, J.C. Wilson, CANMET hydrocyclone: an emerging alternative for the treatment of oily waste streams, Miner. Eng., 17 (2004) 643–649.

[63] J. Sola, J. Sevilla, Importance of input data normalization for the application of neural networks to complex industrial problems, IEEE Trans. Nucl. Sci., 44 (1997) 1464–1468.

[64] D. Singh, B. Singh, Investigating the impact of data normalization on classification performance, Appl. Soft Comput., 97 (2020) 105524, doi: 10.1016/j.asoc.2019.105524.

[65] A. Bemani, A. Baghban, A. Mosavi, Estimating CO_2-brine diffusivity using hybrid models of ANFIS and evolutionary algorithms, Eng. Appl. Comput. Fluid Mech., 14 (2020) 818–834.

[66] T. Hill, L. Marquez, M. O'Connor, W. Remus, Artificial neural network models for forecasting and decision making, Int. J. Forecasting, 10 (1994) 5–15.

[67] E. Grossi, M. Buscema, Introduction to artificial neural networks, Eur. J. Gastroenterol. Hepatol., 19 (2007) 1046–1054.

[68] B. Lang, Monotonic Multi-Layer Perceptron Networks as Universal Approximators, International Conference on Artificial Neural Networks, Springer, 2005.

[69] J.A. Bullinaria, Radial Basis Function Networks: Introduction, Neutral Computation: Lecture, 2015.

[70] M. Abdi-Khanghah, A. Bemani, Z. Naserzadeh, Z. Zhang, Prediction of solubility of N-alkanes in supercritical CO_2 using RBF-ANN and MLP-ANN, J. CO_2 Util., 25 (2018) 108–119.

[71] S. Chen, S. Billings, P. Grant, Non-linear system identification using neural networks, Int. J. Control, 51 (1990) 1191–1214.

[72] R. Eberhart, J. Kennedy, Particle Swarm Optimization, Proceedings of ICNN'95 - International Conference on Neural Networks, IEEE, Perth, WA, Australia, 1995.

[73] J. Kennedy, Particle Swarm Optimization, C. Sammut, G.I. Webb, Eds., Encyclopedia of Machine Learning, Springer, Boston, MA, 2010, pp. 760–766.

[74] A. Tharwat, T. Gaber, A.E. Hassanien, B.E. Elnaghi, Particle Swarm Optimization: A Tutorial, Handbook of Research on Machine Learning Innovations and Trends, IGI Global: International Academic Publisher, Handbook of Research on Machine Learning Innovations and Trends (2 Volumes), Pennsylvania, United States, 2017, pp. 614–635.

[75] R. Razavi, A. Bemani, A. Baghban, A.H. Mohammadi, S. Habibzadeh, An insight into the estimation of fatty acid methyl ester based biodiesel properties using a LSSVM model, Fuel, 243 (2019) 133–141.

[76] J.-S. Chiou, S.-H. Tsai, M.-T. Liu, A PSO-based adaptive fuzzy PID-controllers, Simul. Modell. Pract. Theory, 26 (2012) 49–59.

[77] R. Poli, J. Kennedy, T. Blackwell, Particle swarm optimization, Swarm Intell., 1 (2007) 33–57.

[78] M. Buragohain, Adaptive Network Based Fuzzy Inference System (ANFIS) as a Tool for System Identification with Special Emphasis on Training Data Minimization, Doctor of Philosophy, Department of Electronics and Communication Engineering, Indian Institute of Technology, Guwahati, India, 2009.

[79] Y. Tsukamoto, Advances in Fuzzy Set Theory and Applications, Netherland, Amsterdam: North-Holland, DA, 1979, pp. 137–149.

[80] A. Cruz, N. Mestrado, ANFIS: Adaptive Neuro-Fuzzy Inference Systems, IM, UFRJ, Mestrado, NCE, 2009.

[81] M. Afshar, A. Gholami, M. Asoodeh, Genetic optimization of neural network and fuzzy logic for oil bubble point pressure modeling, Korean J. Chem. Eng., 31 (2014) 496–502.

[82] H. Moeeni, H. Bonakdari, I. Ebtehaj, Integrated SARIMA with neuro-fuzzy systems and neural networks for monthly inflow prediction, Water Resour. Manage., 31 (2017) 2141–2156.

[83] I. Ebtehaj, H. Bonakdari, Performance evaluation of adaptive neural fuzzy inference system for sediment transport in sewers, Water Resour. Manage., 28 (2014) 4765–4779.

[84] F. Moradi, H. Bonakdari, O. Kisi, I. Ebtehaj, J. Shiri, B. Gharabaghi, Abutment scour depth modeling using neuro-fuzzy-embedded techniques, Mar. Georesour. Geotechnol., 37 (2019) 190–200.

[85] J.A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Process. Lett., 9 (1999) 293–300.

[86] V. Vapnik, V. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.

[87] T. Van Gestel, J.A. Suykens, B. Baesens, S. Viaene, J. Vanthienen, G. Dedene, B. De Moor, J. Vandewalle, Benchmarking least squares support vector machine classifiers, Mach. Learn., 54 (2004) 5–32.

[88] K.-R. Muller, S. Mika, G. Ratsch, K. Tsuda, B. Scholkopf, An introduction to Kernel-based learning algorithms, IEEE Trans. Neural Networks, 12 (2003) 181–201.

[89] S. Medasani, J. Kim, R. Krishnapuram, An overview of membership function generation techniques for pattern recognition, Int. J. Approximate Reasoning, 19 (1998) 391–417.

[90] N. Talpur, M.N.M. Salleh, K. Hussain, An investigation of membership functions on performance of ANFIS for solving classification problems, IOP Conf. Ser.: Mater. Sci. Eng., 226 (2017) 012103.

[91] M. Babanezhad, A.T. Nakhjiri, S. Shirazian, Changes in the number of membership functions for predicting the gas volume fraction in two-phase flow using grid partition clustering of the ANFIS method, ACS Omega, 5 (2020) 16284–16291.

[92] H. Bonakdari, H. Moeeni, I. Ebtehaj, M. Zeynoddin, A. Mahoammadian, B. Gharabaghi, New insights into soil temperature time series modeling: linear or nonlinear?, Theor. Appl. Climatol., 135 (2019) 1157–1177.

[93] M. Zeynoddin, H. Bonakdari, I. Ebtehaj, F. Esmaeilbeiki, B. Gharabaghi, D.Z. Haghi, A reliable linear stochastic daily soil temperature forecast model, Soil Tillage Res., 189 (2019) 73–87.

[94] S. Liu, Y. Yan, Y. Gao, Optimization of geometry parameters with separation efficiency and flow split ratio for downhole oil-water hydrocyclone, Therm. Sci. Eng. Prog., 8 (2018) 370–374.

[95] S. Qiu, G. Wang, S. Zhou, Q. Liu, L. Zhong, L. Wang, The downhole hydrocyclone separator for purifying natural gas hydrate: structure design, optimization, and performance, Sep. Sci. Technol., 55 (2020) 564–574.

[96] M. Liu, J. Chen, X. Cai, Y. Han, S. Xiong, Oil–water pre-separation with a novel axial hydrocyclone, Chin. J. Chem. Eng., 26 (2018) 60–66.

[97] J.E. Hamza, H.H. Al-Kayiem, T.A. Lemma, Experimental investigation of the separation performance of oil/water mixture by compact conical axial hydrocyclone, Therm. Sci. Eng. Prog., 17 (2020) 100358, doi: 10.1016/j.tsep.2019.100358.

[98] H. OSEI, Experimental study of a hydrocyclonic oil-water separator for downhole separation, Ghana J. Technol., 4 (2019) 57–64.

[99] Y.-l. Chang, W.-q. Ti, H.-l. Wang, S.-w. Zhou, Y. Huang, J.-p. Li, G.-r. Wang, Q. Fu, H.-t. Lin, J.-w. Wu, Hydrocyclone used for in-situ sand removal of natural gas-hydrate in the subsea, Fuel, 285 (2021) 119075, doi: 10.1016/j.fuel.2020.119075.