



Development of empirical models with high accuracy for estimation of drag coefficient of flow around a smooth sphere: An evolutionary approach



Reza Barati^a, Seyed Ali Akbar Salehi Neyshabouri^{b,*}, Goodarz Ahmadi^c

^a Faculty of Civil and Environmental Engineering, Tarbiat Modares University, Tehran, Iran

^b Water Engineering Research Center, Department of Civil Engineering, Tarbiat Modares University, Tehran, Iran

^c Department of Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, NY, USA

ARTICLE INFO

Article history:

Received 23 October 2013

Received in revised form 12 February 2014

Accepted 14 February 2014

Available online 22 February 2014

Keywords:

Particle motion

Sphere drag

Reynolds number

Multi-gene Genetic Programming

ABSTRACT

An accurate correlation for the smooth sphere drag coefficient with wide range of applicability is a useful tool in the field of particle technology. The present study focuses on the development of high accurate drag coefficient correlations from low to very high Reynolds numbers (up to 10^6) using a multi-gene Genetic Programming (GP) procedure. A clear superiority of GP over other methods is that GP is able to determine the structure and parameters of the model, simultaneously, while the structure of the model is imposed by the user in traditional regression analysis, and only the parameters of the model are assigned. In other words, in addition to the parameters of the model, the structure of it can be optimized using GP approach. Among two new and high accurate models of the present study, one of them is acceptable for the region before drag dip, and the other is applicable for the whole range of Reynolds numbers up to 10^6 including the transient region from laminar to turbulent. The performances of the developed models are examined and compared with other reported models. The results indicate that these models respectively give 16.2% and 69.4% better results than the best existing correlations in terms of the sum of squared of logarithmic deviations (SSLD). On the other hand, the proposed models are validated with experimental data. The validation results show that all of the estimated drag coefficients are within the bounds of $\pm 7\%$ of experimental values.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The motion of particles in fluids is a key subject in many problems in the fields of chemical and metallurgical engineering as well as mechanical and environmental engineering. The solution of these problems generally involves determining the local behavior of flow and the interaction between solid and liquid phases through the knowledge of hydrodynamic forces such as drag. The drag force is the combination of the normal (i.e. pressure) and tangential (i.e. wall shear stress) forces on the body in the flow direction. However, the distributions of the pressure and wall shear stress are often very difficult to achieve, so the magnitude of the drag force can be determined only through the knowledge of drag coefficient. Analytical determination of the drag coefficient such as Stokes' law is only valid for Reynolds number, Re , less than 0.1 (Flemmer and Banks [1], Kreith [2]), although the drag coefficient can be ascertained using empirical and semi-empirical correlations based

on experimental data when inertial effects are significant (i.e. higher Reynolds numbers).

The drag coefficient of a smooth sphere in incompressible flow is a function of Re based on both theoretical investigations and numerous experimental data (Kreith [2]). The main classes of the dependence of drag coefficient on Reynolds number are (1) very low Reynolds number flow (i.e. creeping flow), (2) moderate Reynolds number flow (i.e. laminar boundary layer), and (3) very large Reynolds number flow (i.e. turbulent boundary layer) (Munson et al. [3]). In the first class ($Re < 1$), the flows reflect entirely the viscous effect of flow with no separation results. By increasing Reynolds numbers (i.e. increasing the particle size or flow velocity for a given Kinematic viscosity), the separation region can be observed at $Re \approx 10$, and the region increases until $Re \approx 1000$, where most of the drag is due to pressure drag rather than frictional drag. Parenthetically, it should be noted that the value of the drag coefficient decreases, as wake area becomes larger. At a sufficiently high Reynolds number ($10^3 < Re < 10^5$), the drag coefficient is relatively constant (Munson et al. [3]). When transition from laminar to turbulent flow occurs, a dramatic dip (up to almost 80%) in the drag coefficient appears at critical value $Re \approx 2 \times 10^5$ since the turbulent boundary layer travels further along the surface into the adverse pressure gradient on the rear portion of the sphere before the separation, so the wake is smaller, causing less pressure drag. After

* Corresponding author at: Tarbiat Modares University, Al-Ahmad Ave., Tehran I.R. Iran P.O. Box 14115-143. Tel.: +982182883316; fax: +982182883381.

E-mail addresses: r88barati@gmail.com, reza.barati@modares.ac.ir (R. Barati), salehi@modares.ac.ir (S.A.A.S. Neyshabouri), gahmadi@clarkson.edu (G. Ahmadi).

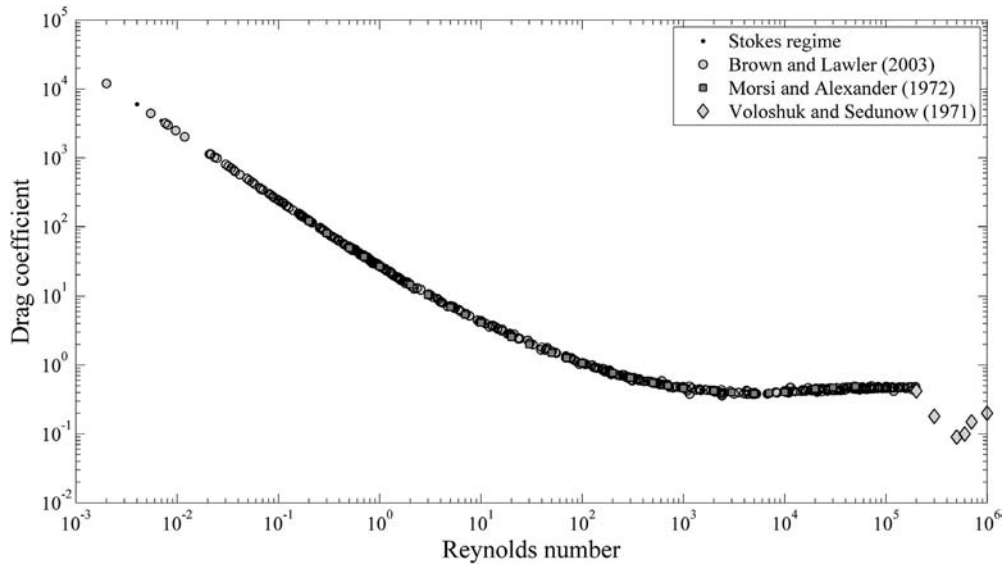


Fig. 1. Illustration of the variations of the drag coefficient with Reynolds numbers using reliable data points of Stokes regime and available experiments in the literature [5,8,11].

Table 1

Summary of some empirical relationships for $Re < 2 \times 10^5$.

Investigator	Model and Reynolds number range	Equation no.
Rouse [14]	$\hat{C}_D = \frac{24}{Re} + \frac{3}{Re^{0.5}} + 0.34$ for $Re < 2 \times 10^5$	(1)
Engelund and Hansen [13]	$\hat{C}_D = \frac{24}{Re} + 1.5$ for $Re < 2 \times 10^5$	(2)
Clift and Gauvin [17] ^a	$\hat{C}_D = \frac{24}{Re} \left(1 + 0.152Re^{0.677} \right) + \frac{0.417}{1 + 5070Re^{-0.994}}$ for $Re < 2 \times 10^5$	(3)
Morsi and Alexander [11]	$\hat{C}_D = \begin{cases} \frac{24}{Re} & \text{for } Re < 0.1, \\ \frac{22.7300}{Re} + \frac{0.0903}{Re^2} + 3.6900 & \text{for } 0.1 < Re < 1, \\ \frac{29.1667}{Re} - \frac{3.8889}{Re^2} + 1.2220 & \text{for } 1 < Re < 10, \\ \frac{46.5000}{Re} - \frac{116.6700}{Re^2} + 0.6167 & \text{for } 10 < Re < 100, \\ \frac{98.3300}{Re} - \frac{27.78}{Re^2} + 0.3644 & \text{for } 100 < Re < 1000, \\ \frac{148.6200}{Re} - \frac{4.75 \times 10^4}{Re^2} + 0.3570 & \text{for } 1000 < Re < 5000, \\ -\frac{490.5460}{Re} + \frac{57.87 \times 10^4}{Re^2} + 0.4600 & \text{for } 5000 < Re < 10,000, \\ -\frac{1662.5000}{Re} + \frac{5.4167 \times 10^6}{Re^2} + 0.5191 & \text{for } 10,000 < Re < 50,000 \end{cases}$	(4)
Graf [36]	$\hat{C}_D = \frac{24}{Re} + \frac{7.3}{1 + Re^{0.5}} + 0.25$ for $Re < 2 \times 10^5$	(5)
Flemmer and Banks [1] ^a	$\hat{C}_D = \frac{24}{Re} 10^E$ where $E = 0.383Re^{0.356} - 0.207Re^{0.396} - \frac{0.143}{1 + (\log Re)^2}$ for $Re < 2 \times 10^5$	(6)
Khan and Richardson [20] ^a	$\hat{C}_D = \left(2.49Re^{-0.328} + 0.34Re^{0.067} \right)^{3.18}$ for $Re < 2 \times 10^5$	(7)
Swamee and Ojha [21]	$\hat{C}_D = 0.5 \left\{ 16 \left[\left(\frac{24}{Re} \right)^{1.6} + \left(\frac{130}{Re} \right)^{0.72} \right]^{2.5} + \left[\left(\frac{40,000}{Re} \right)^2 + 1 \right]^{-0.25} \right\}^{0.25}$ for $Re < 1.5 \times 10^5$	(8)
Yen [37]	$\hat{C}_D = \frac{24}{Re} \left(1 + 0.15\sqrt{Re} + 0.017Re \right) - \frac{0.208}{1 + 10^4 Re^{-0.5}}$ for $Re < 2 \times 10^5$	(9)
Haider and Levenspiel [19] ^a	$\hat{C}_D = \frac{24}{Re} \left(1 + 0.150Re^{0.681} \right) + \frac{0.407}{1 + 8710Re^{-1}}$ for $Re < 2 \times 10^5$	(10)
Cheng [6]	$\hat{C}_D = \frac{24}{Re} \left(1 + 0.27Re \right)^{0.43} + 0.47 \left[1 - \exp \left(-0.04Re^{0.38} \right) \right]$ for $Re < 2 \times 10^5$	(11)
Terfous et al. [22]	$\hat{C}_D = 2.689 + \frac{21.683}{Re} + \frac{0.131}{Re^2} - \frac{10.616}{Re^{0.1}} + \frac{12.216}{Re^{0.2}}$ for $0.1 < Re < 5 \times 10^4$	(12)
Mikhailov and Freire [7]	$\hat{C}_D = \frac{3808 \left[(1.617, 933/2030) + (178, 861/1063)Re + (1219/1084)Re^2 \right]}{681Re \left[(77, 531/422) + (13, 529/976)Re - (1/71, 154)Re^2 \right]}$ for $Re < 118, 300$	(13)

^a These models were improved by Brown and Lawler [5].

Table 2Empirical relationships for Reynolds numbers up to 10^6 .

Investigator	Model and Reynolds number range	Equation no.
Clift et al. [16]	$\hat{C}_D = \begin{cases} \frac{24}{Re} + \frac{3}{16} & \text{for } Re < 0.01 \\ \frac{24}{Re} \left(1 + 0.1315Re^{0.82-0.05 \log Re } \right) & \text{for } 0.01 < Re \leq 20 \\ \frac{24}{Re} \left(1 + 0.1935Re^{0.6305} \right) & \text{for } 20 < Re \leq 260 \\ 10^{1.6435-1.1242 \log Re +0.1558 \log Re ^2} & \text{for } 260 < Re \leq 1500 \\ 10^{-2.4571+2.5558 \log Re -0.9295 \log Re ^2+0.1049 \log Re ^3} & \text{for } 1500 < Re \leq 1.2 \times 10^4 \\ 10^{-1.9181+0.6370 \log Re -0.0636 \log Re ^2} & \text{for } 1.2 \times 10^4 < Re \leq 4.4 \times 10^4 \\ 10^{-4.3390+1.5809 \log Re -0.1546 \log Re ^2} & \text{for } 4.4 \times 10^4 < Re \leq 3.38 \times 10^5 \\ 29.78-5.3 \log Re & \text{for } 3.38 \times 10^5 < Re \leq 4 \times 10^5 \\ 0.1 \log Re -0.49 & \text{for } 4 \times 10^5 < Re \leq 10^6 \end{cases}$	(14)
Ceylan et al. [9]	$\hat{C}_D = 1 - 0.5 \exp(0.182) + 10.11Re^{-2/3} \exp(0.952Re^{-1/4}) - 0.03859Re^{-4/3} \exp(1.30Re^{-1/2}) + 0.037 \times 10^{-4}Re \exp(-0.125 \times 10^{-4}Re) - 0.116 \times 10^{-10}Re^2 \exp(-0.444 \times 10^{-5}Re) \quad \text{for } 0.1 < Re < 10^6$	(15)
Almedeij [10]	$\hat{C}_D = \left[\frac{1}{(\varphi_1 + \varphi_2)^{-1} + (\varphi_3)^{-1} + \varphi_4} \right]^{0.1}$ where $\varphi_1 = (24Re^{-1})^{10} + (21Re^{-0.67})^{10} + (4Re^{-0.33})^{10} + (0.4)^{10}$, $\varphi_2 = \left[(0.148Re^{0.11})^{-10} + (0.5)^{-10} \right]^{-1}$, $\varphi_3 = (1.57 \times 10^8 Re^{-1.625})^{10}$, $\varphi_4 = \left[(6 \times 10^{-17} Re^{2.63})^{-10} + (0.2)^{-10} \right]^{-1}$ for $Re < 10^6$	(16)
Morrison [23]	$\hat{C}_D = \frac{24}{Re} + \frac{2.6(\frac{Re}{5})}{1 + (\frac{Re}{5})^{1.52}} + \frac{0.411(\frac{Re}{250000})^{-7.94}}{1 + (\frac{Re}{250000})^{-8.00}} + \left(\frac{Re^{0.8}}{461000} \right) \quad \text{for } Re < 10^6$	(17)

this abrupt descent, the value of the drag coefficient increases by increasing Reynolds numbers. Finally, for $Re > 10^6$ a constant value of the drag coefficient (≈ 0.2) is acceptable (Potter et al. [4]).

Most of the information pertaining to drag force on the sphere arises from numerous experiments with wind tunnels, water tunnels, towing tanks, and other ingenious devices (Munson et al. [3]). Experimental data of the drag coefficient of spherical particles have been presented in the literature having a wide range of Re . However, some of the available experimental data are not accurate, adequately. Brown and Lawler [5] reviewed the experimental studies of sphere drag coefficient for $Re < 2 \times 10^5$. They assembled 606 data points which were originally presented in tabular form. By excluding some experimental data for various reasons, Brown and Lawler [5] presented 480 very high quality data points by considering wall effects. This data set seems acceptable among other researchers for developing correlations (Cheng [6], Mikhailov and Freire [7]). On the other hand, Voloshuk and Sedunow [8] presented the experimental data for higher Reynolds numbers with good quality. This data set was also used in several studies such as Ceylan et al. [9] and Almedeij [10]. Fig. 1 illustrates all of the mentioned data. The variations of the drag coefficient with Reynolds numbers can follow as explained in the previous paragraph by considering Fig. 1.

In the previous studies, the regression analyses were applied to obtain a correlation for the estimation of the drag coefficient of spherical particles. Several forms and procedures such as multi-segment polynomial, exponential function, piecewise matched, power function and rational fraction were used in these studies. These forms of the correlations were developed by imposing general arithmetic operations (i.e. plus, minus, multiplication and division), and/or some function set (e.g. logarithm, and exponential functions) without considering other popular functions such as sin, cos, tan, tanh, and natural logarithm. Therefore, the performance of existing correlations is less than perfect. In the present study, multi-gene Genetic Programming (GP) is adopted

to develop high accurate models for the estimation of the drag coefficient of the free falling smooth sphere. Unlike traditional regression analysis in which the structure of the model must be specified, GP automatically evolves both the structure and parameters of the drag coefficient estimation model. Therefore, both parameters and structure of the model will be optimized. The experimental data points of Voloshuk and Sedunow [8], and Brown and Lawler [5] will be used to develop empirical models. Seventeen popular correlations will be reviewed for comparison purposes, and experimental data points of Morsi and Alexander [11] along with analytical solution of Stokes regime will be considered for the validation of the developed models.

2. Literature review

Many empirical or semi-empirical correlations that vary somewhat in form have been developed to estimate the standard drag curve of smooth sphere using regression techniques. Seventeen of them which are allocated in two groups based on range of applicability are presented in Tables 1 and 2. The first group covers Reynolds numbers up to 2×10^5 while the second covers Reynolds numbers up to 10^6 . A critical discussion about these models will be presented in the next paragraph.

Rubey [12] suggested a simple equation for estimation of the drag coefficient of sphere based on combining Stokes' law with the impact formula. Engelund and Hansen [13] revised Rubey's equation to improve it for natural sands and gravels. Rouse [14] proposed a relation that has been widely used in the environmental field. Morsi and Alexander [11] developed a model based on the multi-segment polynomial equations that are valid for the range of $Re < 5 \times 10^4$ (Nasrollahi et al. [15]). Clift et al. [16] analyzed the available experimental data by considering wall effects, compressibility effects, non-continuum effects and support interference. They established another multi-segment model using several polynomial equations. Two recent models have good

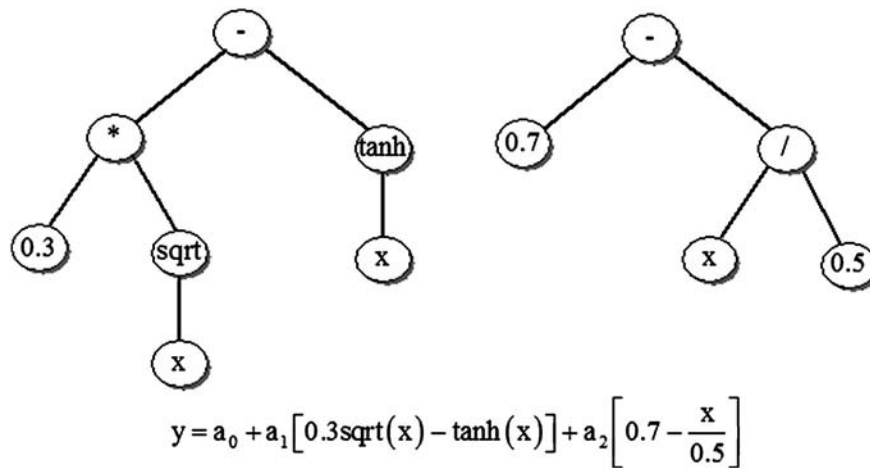


Fig. 2. A typical multi-gene tree-based model.

accuracy, but these complicated regression equations suffer from gradient discontinuities between subintervals. Clift and Gauvin [17] proposed a simple five-constant correlation to estimate the drag coefficient. After them, this correlation has been improved by Turton and Levenspiel [18], Haider and Levenspiel [19], and Brown and Lawler [5] based on regression analysis by minimizing the sum of squared of logarithmic deviations (SSLD). Flemmer and Banks [1] used a regression technique to generate another expression of an exponential form for the range of $Re < 3 \times 10^5$. Brown and Lawler [5] improved this expression both in range and accuracy using local minimizations. Khan and Richardson [20] compiled experimental results of various researchers. They proposed a drag equation of a power form using nonlinear regression on 300 data points. This equation was also improved by Brown and Lawler [20]. Swamee and Ojha [21] presented another model by combining four expressions that were obtained for different ranges of Reynolds numbers (i.e. a piecewise matched procedure). The authors claimed that the resulting equation faithfully described the drag coefficient-versus-Reynolds number variation of a spherical particle for $R < 1.5 \times 10^5$. Ceylan et al. [9] developed a model in the form of the series by analyzing several transfer phenomena for the ranges of $0.1 < Re < 10^6$. The effects of the form drag and skin friction drag were separately considered using two series in the model. However, the accuracy of this model is low, as will be shown. Almedeij [10] presented a model, which is valid for the ranges of $Re < 10^6$, using the concept of the matching asymptotically, and employing regression techniques. He developed the model by using four functions that can be combined into an overall relationship. However, this model suffers from the complicated form. Cheng [6] recommended a five-parameter correlation,

and compared it with seven previously-developed formulas. This correlation consists of two terms (1) an extended Stokes' law applicable approximately for $Re < 100$, and (2) an exponential function accounting for slight deviations from the Newton's law for high Re . Cheng [6] claimed that the proposed correlation gives the best representation of the historical data for the region before drag dip. Terfous et al. [22] developed an equation for $Re < 5 \times 10^4$ by using the simple series function and applying the least-squares method. However, this model is not applicable for $Re < 0.1$ because they used the term $21.683/Re$ instead of the traditional term $24/Re$ in the equation. Morrison [23] used a piecewise matched procedure to develop a model that covers the transition from laminar to turbulent flow region. Recently, Mikhailov and Freire [7] presented an expression for the estimation of the drag coefficient of a falling sphere for the ranges of $Re < 118,300$ based on Shanks transformation which is a non-linear rational fractional transformation of the series of Goldstein to Oseen's equation.

Briefly, it can be said that the above equations of the drag coefficient estimation suffer from the complicated correlation, bounds of applicability, and/or low accuracy. Therefore, the development of a simple and accurate model with a high range of applicability is vital.

3. Genetic Programming

Genetic Programming (GP) is a random-based procedure for automatically learning the most "fit" computer programs by means of artificial evolution (Johari et al. [24]). Recently, GP has been successfully applied in many applications such as the prediction of the soil–water characteristic of soils (Johari et al. [24]), the estimation of the bridge pier scour (Azamathulla et al. [25]), and the prediction of the outflow hydrograph from earthen dam breach (Hakimzadeh et al. [26]).

GP, which is a branch of the conventional genetic algorithm (GA; Holland [27]), was initially developed by Koza [28]. The significant superiority of GP over GA is that in GP there is no need to define the structure of the model a priori. In other words, GP can determine not only the coefficients and parameters of the model, but also and more importantly, the form of the model itself (Johari et al. [24]). GP randomly generates a population of equal or unequal length computer programs (or symbolic expression) with a high level of diversity. The programs are represented by tree-based structures using variables (terminal) and several mathematical operators (function) that can be selected from sets of terminals and functions, respectively. The terminal set contains the numerical constants and external inputs of the program while the function set contains the basic arithmetic operators such as $+$, $-$, $*$, and $/$ and function calls such as e^x , $\sin$, $\cos$, $\tan$, $\tanh$, $\log$, $\ln$, $\sqrt{x}$

Table 3
Range of parameters of multi-gene GP.

Parameter	Range
Population size	200–10,000
Number of generations	400–10,000
Maximum number of genes	1–10
Maximum number of nodes per tree	1–15
Maximum depth of trees	3–15
Probability of multi-gene GP tree mutation ^a	0.1–0.2
Probability of multi-gene GP tree cross over ^a	0.75–0.85
Probability of multi-gene GP tree direct copy ^a	0.05
Size of the tournament	2–3
Arithmetic operations and function set	($+$, $-$, $*$, $/$, square, $\tanh$, $\sin$, $\cos$, $\tan$, $\log$, $\ln$, $\exp$)

^a Sum must be equal to 1.

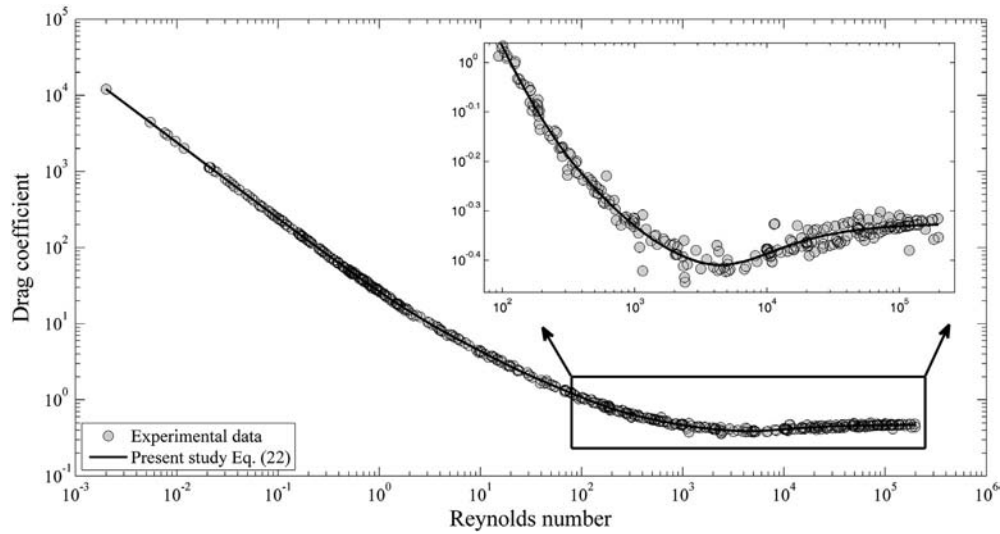


Fig. 3. Developed drag coefficient model for the region before drag dip along with experimental data of Brown and Lawler [5].

and power. After choosing the best symbolic expressions through defining a fitness function, an improved population will be created using crossover and mutation operators. The iterations proceed until they meet stopping criteria (e.g. reaching the specified number of generations). It is informative to express that GP is often known as symbolic regression, when an empirical model is developed using collected data from a process or system. More details of GP procedure can be found in Koza [28].

In order to improve fitness for non-linear behavior procedures, the multi-gene GP which is a branch of standard GP can be used (Hinchliffe et al. [29], and Hiden [30]). In the traditional GP, the model expresses as a single tree while, in the multi-gene GP, several trees may define the model through a weighted linear combination of each gene plus a bias term as

$$y = a_0 + a_1 \times \text{gene}_1 + a_2 \times \text{gene}_2 + \dots + a_n \times \text{gene}_n \quad (18)$$

where y is the output variable, a_0 is the bias term and a_i is weight of the i th gene.

An example of multi-gene tree-based model is presented in Fig. 2, where an output variable y is predicted using an input variable x . As can be seen in Fig. 2, multi-gene GP is a linear combination of nonlinear terms, and this feature makes it possible to recognize the pattern of engineering problems in a highly precise manner (Hinchliffe et al. [29]).

In the present study, GPTIPS, which is a new “Genetic Programming and Symbolic Regression” code (Searson [31]), was adapted to perform a multi-gene GP for the development of the sphere drag coefficient models. GPTIPS which uses lexicographic tournament selection as an effective technique for controlling the bloating of the model has the capability of setting some restrictions on initial parameters such as the maximum number of genes, maximum depth of trees and genes and maximum number of nodes per tree to avoid the most recent problems of bloat in GP. There are several successful applications of GPTIPS in the different fields such as predicting the liquefaction resistance of sand–silt

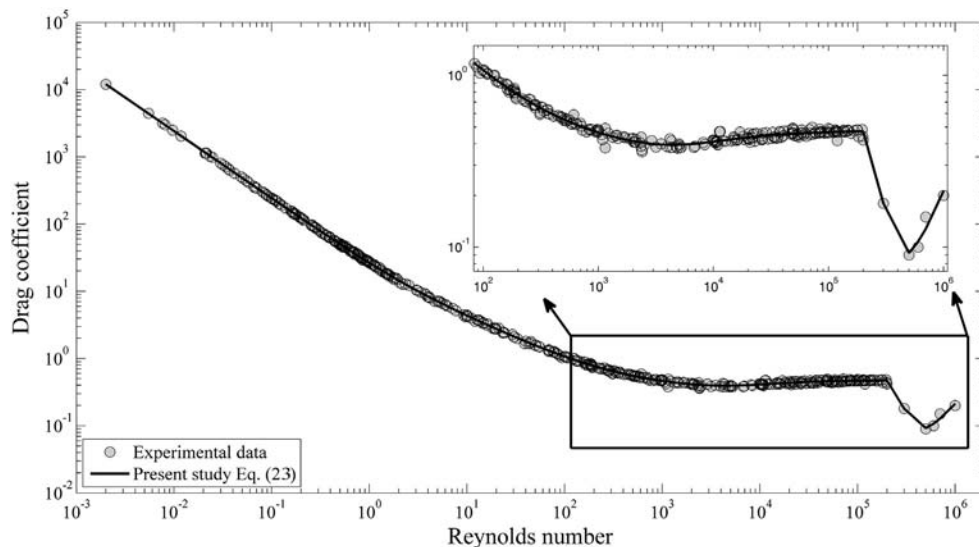


Fig. 4. Developed drag coefficient model for the wide range of particle Reynolds numbers along with experimental data of Brown and Lawler [5], and Voloshuk and Sedunov [8].

Table 4Comparison of the performance evaluation criteria of the relationships for $Re < 2 \times 10^5$.

Order of accuracy	Reference	SSLD	RMSLD	SRE
1	Present study Eq. (22)	0.0881	0.0136	10.7171
2	Cheng [6]	0.1051	0.0148	11.8567
3	Morsi and Alexander [11]	0.1314	0.0176	13.0465
4	Clift and Gauvin [17] ^a	0.1502	0.0177	15.4402
5	Haider and Levenspiel [19] ^a	0.1510	0.0177	15.5260
6	Flemmer and Banks [1] ^a	0.1978	0.0203	17.4739
7	Khan and Richardson [20] ^a	0.2091	0.0209	17.4464
8	Swamee and Ojha [21]	0.2778	0.0242	20.4107
9	Terfous et al. [22]	0.4670	0.0322	23.1950
10	Mikhailov and Freire [7]	0.4703	0.0319	25.9766
11	Yen [37]	0.8475	0.0420	38.7227
12	Rouse [14]	2.1313	0.0666	51.3652
13	Graf [36]	5.2163	0.1042	63.1865
14	Engelund and Hansen [13]	61.5283	0.3580	545.4005

^a These models were improved by Brown and Lawler [5].

mixtures (Baziar et al. [32]), and predicting the ultimate bearing capacity of shallow foundations on cohesionless soils (Shahnazari and Tutunchian [33]). The ranges of the initial parameters of multi-gene GP for runs of this study are summarized in Table 3. More descriptions about the initial parameters and specifications of multi-gene GP can be found in Searson [31].

4. Developed models based on multi-gene GP

The sum of the squared deviations is a good objective function to minimize the errors between experimental data and calculated results by considering the previous studies (Turton and Levenspiel [18]; Haider and Levenspiel [19]; Brown and Lawler [5]; Barati [34], Barati [35]). Therefore, the fitness function of multi-gene GP is to minimize the sum of squared of logarithmic deviations (SSLD) between the estimated drag coefficient and experimental data points

$$\text{Minimize SSLD} = \sum_{i=1}^N (\log C_D - \log \hat{C}_D)^2 \quad (19)$$

where C_D is the experimental value of the drag coefficient, $\hat{C}_D$ is the estimated drag coefficient, and N is the total number of the data.

For assessing the performance of the proposed models compared to existing correlations, root-mean-square of logarithmic deviation (RMSLD) and sum of the relative errors (SRE), as well as SSLD are used as

$$\text{RMSLD} = \sqrt{\sum_{i=1}^N (\log C_D - \log \hat{C}_D)^2 / N} \quad (20)$$

$$\text{SRE} = \sum_{i=1}^N \left(\frac{|C_D - \hat{C}_D|}{C_D} \right) \quad (21)$$

Table 5Comparison of the performance evaluation criteria of the relationships for Reynolds numbers up to 10^6 .

Order of accuracy	Reference	SSLD	RMSLD	SRE
1	Present study Eq. (23)	0.0982	0.0142	11.2853
2	Almedeij [10]	0.3210	0.0257	23.2094
3	Clift et al. [16]	0.4467	0.0303	17.4004
4	Morrison [23]	0.5672	0.0342	29.1768
5	Ceylan et al. [9]	0.9186	0.0449	33.2925

It is notable that SSLD and RMSLD were calculated based on logarithmic values of the drag coefficients while SRE can be used to evaluate the original value of the drag coefficients.

In order to select appropriate models, GPTIPS is run over 100 times, because multi-gene GP is a stochastic procedure. As will be shown, the best representation of available experimental data of the smooth sphere drag coefficient in the literature for Reynolds numbers up to 2×10^5 and up to 10^6 can be estimated by Eqs. (22) and (23), respectively.

$$\begin{aligned} \hat{C}_D = & 5.4856 \times 10^9 \tanh(4.3774 \times 10^{-9}/Re) + 0.0709 \tanh(700.6574/Re) \\ & + 0.3894 \tanh(74.1539/Re) - 0.1198 \tanh(7429.0843/Re) \\ & + 1.7174 \tanh[9.9851/(Re + 2.3384)] + 0.4744 \quad \text{for } Re < 2 \times 10^5 \end{aligned} \quad (22)$$

$$\begin{aligned} \hat{C}_D = & 8 \times 10^{-6} \left[(Re/6530)^2 + \tanh(Re) - 8 \ln(Re) / \ln(10) \right] \\ & - 0.4119 e^{-2.08 \times 10^{43}/[Re + Re^2]^4} - 2.1344 e^{-\{[\ln(Re^2 + 10.7563)/\ln(10)]^2 + 9.9867\}/Re} \\ & + 0.1357 e^{-[Re/1620]^2 + 10370/Re} - 8.5 \times 10^{-3} \{2 \ln[\tanh(\tanh(Re))]/ \\ & \ln(10) - 2825.7162\}/Re + 2.4795 \quad \text{for } Re < 10^6 \end{aligned} \quad (23)$$

The parameters of the above models were rounded through the sensitivity analysis. In other words, the decimal point accuracy of the parameters of the models was considered with sufficient digits in order to fade the rounding off errors.

The standard drag curves represented by Eqs. (22) and (23) in comparison with the experimental data are illustrated in Figs. 3 and 4, respectively. First of all, it can be seen that these models truthfully follow the experimental data, and they can represent the mean values of the experimental data with high accuracy. Furthermore, as shown in Figs. 3 and 4, both developed models converge to Stokes solution at very small Reynolds numbers, and the models can capture a minimum for the drag coefficient in a Reynolds number value around 5000 which agree with the experimental data. It should be noted that most of existing correlations cannot capture this minimum, accurately. On the other hand, the existing correlations of the second group, which are listed in Table 2, suffer from less accuracy in either moderate or large Reynolds numbers while Eq. (23) can estimate the experimental data in all of the creeping flow, transitional, and turbulent regions, precisely.

5. Test of the developed equations

In order to examine the developed models, the results of the drag coefficient obtained from the proposed models were compared with those from models developed by other researchers in terms of SSLD, RMSLD and SRE.

Performance evaluation criteria for the correlation of the first and second groups together with the corresponding values of the developed models are listed in Tables 4 and 5, respectively. It should be stated that the models are ranked in the increasing order of SSLD in Tables 4 and 5. Generally, other criteria (i.e. RMSLD and SRE) also increase in Tables 4 and 5 by increasing order of accuracy. The ranges of using the models in the calculation of the criteria were its applicability range, which is presented in Tables 1 and 2. Therefore, it should be noted that some of the correlations such as recently developed models of Terfous et al. [22] and Mikhailov and Freire [7] have large values of SSLD (i.e. low accuracy) although SSLD of them were calculated for $0.1 < Re < 5 \times 10^4$ and $Re < 118,300$, respectively, instead of whole range of Reynolds numbers up to 2×10^5 .

As shown in Tables 1 and 2, both proposed models have the lowest values of SSLD, RMSLD and SRE (higher accuracy). Eqs. (22) and (23) improved 16.2% and 69.4% to match with the experimental data points

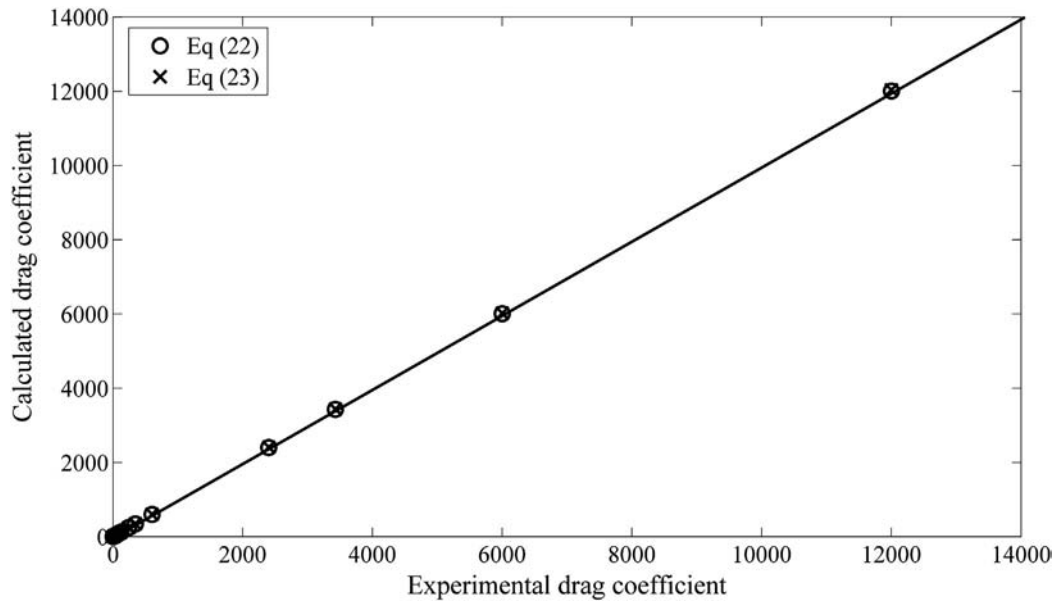


Fig. 5. Comparison between the validation data of the drag coefficient and the associated calculation results.

than the best existing correlations of the first and second groups in terms of SSLD, respectively. The new models substantially improve the fit to experimental data. Consequently, for the accurate estimation of the smooth sphere drag coefficient of engineering problems, Eqs. (22) and (23) can be utilized for problems with Reynolds numbers in the

Table 6

Validation of the developed models using the experimental data of Morsi and Alexander [11], and the analytical solution of Stokes regime.

Re	C_D	−5% of C_D	Eq. (22)	Eq. (23)	+5% of C_D
0.002	12,000	11,400.00	12,008.86	12,034.71	12,600.00
0.004	6000	5700.00	6005.70	6017.32	6300.00
0.007	3429	3257.14	3432.91	3438.94	3600.00
0.01	2400	2280.00	2403.80	2407.74	2520.00
0.04	600	570.00	602.85	603.54	630.00
0.07	343	325.71	345.57	345.88	360.00
0.1	240	228.00	242.66	242.84	252.00
0.2	120	114.00	122.59	122.63	126.00
0.3	80	76.00	82.57	82.57	84.00
0.5	49	46.55	50.56	50.53	51.45
0.7	36.5	34.68	36.83	36.80	38.32
1	26.5	25.18	26.54	26.50	27.82
2	14.4	13.68	14.50	14.48	15.12
3	10.4	9.88	10.46	10.44	10.92
5	6.9	6.56	7.12	7.10	7.24
7	5.4	5.13	5.60	5.58	5.67
10	4.1	3.90	4.36	4.36	4.31
20	2.55	2.42	2.73	2.76	2.68
30	2	1.90	2.12	2.14	2.10
50	1.5	1.42	1.58	1.57	1.58
70	1.27	1.21	1.31	1.30	1.33
100	1.07	1.02	1.08	1.07	1.12
200	0.77	0.73	0.77	0.77	0.81
300	0.65	0.62	0.65	0.66	0.68
500	0.55	0.52	0.56	0.55	0.58
700	0.5	0.48	0.51	0.50	0.52
1000	0.46	0.44	0.47	0.46	0.48
2000	0.42	0.40	0.41	0.41	0.44
3000	0.4	0.38	0.40	0.40	0.42
5000	0.385	0.36	0.39	0.39	0.40
7000	0.39	0.37	0.40	0.40	0.41
10,000	0.405	0.38	0.41	0.41	0.42
20,000	0.45	0.43	0.44	0.44	0.47
30,000	0.47	0.44	0.45	0.45	0.49
50,000	0.49	0.46	0.46	0.46	0.51

region before drag coefficient dip and with Reynolds numbers greater than $Re > 2 \times 10^5$, respectively. Finally, it must be stated that a constant value of the drag coefficient (≈ 0.2) is acceptable for higher Reynolds numbers (i.e. $Re > 10^6$) (Almedeij [10], Potter et al. [4]).

As mentioned previously, the developed equations were calibrated using the data of Voloshuk and Sedunow [8], and Brown and Lawler [5]. The data of Morsi and Alexander [11] along with analytical solution of Stokes regime will be used to validate the proposed models.

Fig. 5 represents the comparison between the validation data of the drag coefficient and the associated calculation results. The figure shows that the results of both models are comparable with the measurement data.

The experimental values of the drag coefficient and $\pm 5\%$ of them are compared with the estimated values by the proposed models in Table 6. The results indicated that the models have no significant error compared to the experimental data. All of the estimated drag coefficients by both models are within the bounds of $\pm 5\%$ of validation data except for 3 data points. If the bounds are increased until $\pm 7\%$, 3 data points will also be covered.

6. Discussion

In this section, two issues about the multi-gene GP procedure will be discussed: 1) Compatibility of multi-gene GP approach with the natural of the problem and 2) evaluation of the level of the accuracy of the developed models.

For the first issue, as mentioned previously, over 100 models were developed with different forms of the equations using multi-gene GP procedure. Although Eqs. (22) and (23) have the lowest errors than the others, most of the developed models are better than the best existing correlations. For example, Eq. (24) presents another model for the estimation of the drag coefficient:

$$\begin{aligned} \hat{C}_D = & 24.0285/Re + 9.4104Re/(Re + 2.8620)^2 + 0.0277[\ln(Re + 32.6397)]^2 \\ & - 0.0128[\ln(Re + 52269)]^2 - 0.4320 \ln(Re + 2.5064) + 3.5991 \quad \text{for } Re < 2 \times 10^5. \end{aligned} \quad (24)$$

The values of SSLD, RMSLD and SRE of Eq. (24) are 0.0960, 0.0141, and 11.5100, respectively. By considering Table 4, Eq. (24) has better accuracy than the best existing model of other studies. The key issue is that Stokes' law (i.e. $24/Re$) can be observed in this model. This fact

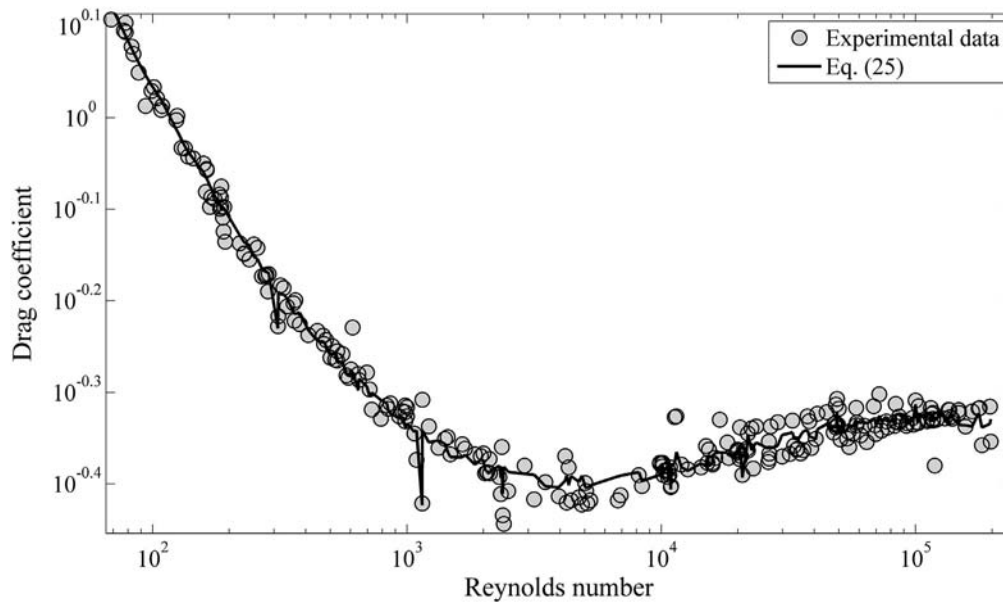


Fig. 6. Illustration of non-smooth curve of Eq. (25).

upholds that multi-gene GP procedure can adapt itself with the natural of the problem. However, for Eqs. (22) and (23), other functions appear instead of $24/Re$. For example, a form of tangent hyperbolic function plays the same role for the calculation of the drag coefficients in low Reynolds number (Stokes' law region).

For the second issue, interestingly, it is observed that the value of SSLD can be decreased [less than the values of Eqs. (22) and (23)] by increasing the maximum depth of trees in the multi-gene GP procedure. However, the obtained correlations by higher number of depth of trees have a non-smooth curve. For example, Eq. (25) presents one of these models

$$\begin{aligned} \hat{C}_D = & 24.03/Re + 1.3255e^{\tanh(Re)} + 0.5756 \ln(Re + 12.3504) - 0.3910 \ln(Re + 30.2029) \\ & - 0.0062 \cos(0.13993Re) - 3.3243 \{ \tanh(Re) + \tanh[\sinh(Re)] \} \\ & + [237119 \ln(Re) + 164358] / [5 \times 10^{10} \sin(Re) + 4.8396 \times 10^{10}] \\ & - \{ 1.5 \times 10^9 \sin[0.4343 \ln(Re)] \} / \{ 10^{14} [\cos(Re) + \tanh(Re)] \} \\ & + \sin[\ln(Re)/\ln(10)] (1.04 \times 10^{-6} Re + 1.60 \times 10^{-5}) \\ & + 14.91 \sin[\tanh(Re)] / \log(Re + 9.24) + 0.3996 \quad \text{for } Re < 2 \times 10^5. \end{aligned} \quad (25)$$

The values of SSLD, RMSLD and SRE of Eq. (25) are 0.0777, 0.0127, and 10.3246, respectively. By considering Table 4, Eq. (25) has the lower errors than Eq. (22) and other correlations. However, as it can be seen in Fig. 6, the curve of Eq. (25) is non-smooth which is not acceptable. In other words, for each model with a lower error than Eq. (22), the curve of the model becomes non-smooth. Therefore, by considering available experimental data of drag coefficient, it can be said that Eqs. (22) and (23) are almost the best possible smooth standard curves which can be constructed.

7. Conclusions

In the present study, the existing correlations of drag coefficient were discussed, critically. Then, a reliable and complete set of historical data was collected for the development and validation of correlations for the estimation of the smooth sphere drag coefficient. An effective procedure (i.e. multi-gene Genetic Programming) was used to develop drag coefficient models through optimizing both parameters and structure of models. Because the procedure is stochastic, the multi-gene GP was run over 100 times for selecting appropriate models.

Results of the numerical and experimental validations indicated that two new correlations which are accurate, simple, and explicit illustrated the best representation of available experimental data for both sub-critical and turbulent regions. The developed models by multi-gene GP substantially (up to almost 70%) improved the fit to experimental data than the best existing correlations in terms of SSLD. The correlations were also verified using another set of experimental data points. The results showed that all of the estimated drag coefficients were within the bounds of $\pm 7\%$ of experimental values.

Finally, it can be said that the proposed models are useful tools for computer programs because of the high accuracy and simplicity of them. Moreover, for future researches, it is hoped that the procedure of this study can be successfully adopted for other problems in the field of particle transport.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.powtec.2014.02.045>.

References

- [1] R.L.C. Flemmer, C.L. Banks, On the drag coefficient of a sphere, *Powder Technol.* 48 (1986) 217–221.
- [2] F. Kreith, *Mechanical Engineering Handbook*, CRC Press LLC, 1999.
- [3] B.R. Munson, D.F. Young, T.H. Okiishi, W.W. Huebsch, *Fundamentals of Fluid Mechanics*, John Wiley & Sons, Inc., New York, 2009.
- [4] M.C. Potter, D.C. Wiggert, B.H. Ramadan, *Mechanics of Fluids*, SI Version, Cengage Learning, 2011.
- [5] P.P. Brown, D.F. Lawler, Sphere drag and settling velocity revisited, *J. Environ. Eng.* 129 (2003) 222–231.
- [6] N.S. Cheng, Comparison of formulas for drag coefficient and settling velocity of spherical particles, *Powder Technol.* 189 (2009) 395–398.
- [7] M.D. Mikhailov, A.P. Freire, The drag coefficient of a sphere: an approximation using Shanks transform, *Powder Technol.* 237 (2013) 432–435.
- [8] V.M. Voloshuk, J.S. Sedunow, *The Processes of Coagulation in Dispersed Systems*, Nauka, Moscow, 1971.
- [9] K. Ceylan, A. Altunbaş, G. Kelbaliyev, A new model for estimation of drag force in the flow of Newtonian fluids around rigid or deformable particles, *Powder Technol.* 119 (2001) 250–256.
- [10] J. Almedeij, Drag coefficient of flow around a sphere: matching asymptotically the wide trend, *Powder Technol.* 186 (2008) 218–223.
- [11] S.A. Morsi, A.J. Alexander, An investigation of particle trajectories in two-phase flow systems, *J. Fluid Mech.* 55 (1972) 193–208.
- [12] W.W. Rubey, Settling velocity of gravel, sand, and silt particles, *Am. J. Sci.* 148 (1933) 325–338.

- [13] F. Engelund, E. Hansen, Monograph on Sediment Transport in Alluvial Streams, Monograph Denmark Technical University, Hydraulic Lab, Denmark, 1967.
- [14] H. Rouse, Fluid Mechanics for Hydraulic Engineers, Dover, New York, N.Y., 1938.
- [15] A. Nasrollahi, S.A.A. Salehi Neyshabouri, G. Ahmadi, M.M. Namin, Numerical simulation of particle saltation process, Part. Sci. Technol. 26 (2008) 529–550.
- [16] R. Clift, J.R. Grace, M.E. Weber, Bubbles, Drops, and Particles, Academic, New York, 1978.
- [17] R. Clift, W.H. Gauvin, The motion of particles in turbulent gas streams, Proc. Chemeca, 70, 1970, pp. 14–28.
- [18] R. Turton, O. Levenspiel, A short note on the drag correlation for spheres, Powder Technol. 47 (1986) 83–86.
- [19] A. Haider, O. Levenspiel, Drag coefficient and terminal velocity of spherical and non-spherical particles, Powder Technol. 58 (1989) 63–70.
- [20] A.R. Khan, J.F. Richardson, The resistance to motion of a solid sphere in a fluid, Chem. Eng. Commun. 62 (1987) 135–150.
- [21] P.K. Swamee, C.S.P. Ojha, Drag coefficient and fall velocity of nonspherical particles, J. Hydraul. Eng. 117 (1991) 660–667.
- [22] A. Terfous, A. Hazzab, A. Ghenaim, Predicting the drag coefficient and settling velocity of spherical particles, Powder Technol. 239 (2013) 12–20.
- [23] F.A. Morrison, An Introduction to Fluid Mechanics, Cambridge University Press, New York, 2013.
- [24] A. Johari, G. Habibagahi, A. Ghahramani, Prediction of soil–water characteristic curve using genetic programming, J. Geotech. Geoenviron. 132 (2006) 661–665.
- [25] H.M. Azamathulla, A.A. Ghani, N.A. Zakaria, A. Guven, Genetic programming to predict bridge pier scour, J. Hydraul. Eng. 136 (2009) 165–169.
- [26] H. Hakimzadeh, V. Nourani, A.B. Amini, Genetic programming simulation of dam breach hydrograph and peak outflow discharge, J. Hydrol. Eng. (May 18 2013), [http://dx.doi.org/10.1061/\(ASCE\)HE.1943-5584.0000849](http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000849) (Posted ahead of print).
- [27] J.H. Holland, Adaptation in Natural and Artificial System, University of Michigan Press, Ann Arbor Mich, 1975.
- [28] J.R. Koza, Genetic Programming: On the Programming of Computers by Means of Natural Selection, MIT Press, Cambridge, Mass, 1992.
- [29] M.P. Hinchliffe, M.J. Willis, H.G. Hiden, M.T. Tham, B. McKay, G. Barton, Modelling chemical process systems using a multi-gene genetic programming algorithm, Late Breaking Paper, GP'96, Stanford: USA, 1996.
- [30] H.G. Hiden, Data-Based Modelling using Genetic Programming, PhD Thesis Dept. Chemical and Process Engineering, University of Newcastle, UK, 1998.
- [31] D.P. Searson, GPTIPS: Genetic Programming & Symbolic Regression for MATLAB, User Guide, 2009.
- [32] M.H. Baziar, Y. Jafarian, H. Shahnazari, V. Movahed, M. Amin Tutunchian, Prediction of strain energy-based liquefaction resistance of sand–silt mixtures: an evolutionary approach, Comput. Geosci. 37 (2011) 1883–1893.
- [33] H. Shahnazari, M.A. Tutunchian, Prediction of ultimate bearing capacity of shallow foundations on cohesionless soils: an evolutionary approach, KSCE J. Civ. Eng. 16 (2012) 950–957.
- [34] R. Barati, Parameter estimation of nonlinear Muskingum models using Nelder–Mead simplex algorithm, J. Hydrol. Eng. ASCE 16 (2011) 946–954.
- [35] R. Barati, Application of Excel Solver for parameter estimation of the nonlinear Muskingum models, KSCE J. Civ. Eng. 17 (2013) 1139–1148.
- [36] W.H. Graf, Hydraulics of Sediment Transport, Water Resources Publications, Littleton, Colorado, 1984.
- [37] B.C. Yen, Sediment Fall Velocity in Oscillating Flow, University of Virginia, Department of Civil Engineering, 1992..