Article

Interfacial Friction Prediction in a Vertical Annular Two-Phase Flow Based on Support Vector Regression Machine

Qiang Liu [1], **Xingya Feng** [1,2,*] **and Junru Chen** [1]

[1] Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China; liuq7@mail.sustech.edu.cn (Q.L.); 12131055@mail.sustech.edu.cn (J.C.)

[2] Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China

* Correspondence: fengxy@sustech.edu.cn

Abstract: Accurate prediction of interfacial friction factor is critical for calculation of pressure drop and investigation of flow mechanism of vertical annular two-phase flows. Theoretical models of interfacial friction factor based on physical insight have been developed; however, these are inconvenient in engineering practice as too many parameters need to be measured. Although many researchers have proposed various empirical correlations to improve computation efficiency, there is no generally accepted simple formula. In this study, an efficient prediction model based on support vector regression machine (SVR) is proposed. Through sensitivity analysis, five factors are determined as the input parameters to train the SVR model, relative liquid film thickness, liquid Reynolds number, gas Reynolds number, liquid Froude number and gas Froude number. The interfacial friction factor is chosen as the output parameter to check the overall performance of the model. With the help of particle swarm algorithm, the optimization process is accelerated considerably, and the optimal model is obtained through iterations. Compared with other correlations, the optimal model shows the lowest average absolute error (*AAE* of 0.0004), lowest maximum absolute error (*MAE* of 0.006), lowest root mean square error (*RMSE* of 0.00076) and highest correlation factor (*r* of 0.995). The analysis using various data in the literature demonstrates its accuracy and stability in interfacial friction prediction. In summary, the proposed machine learning model is effective and can be applied to a wider range of conditions for vertical annular two-phase flows.

Keywords: vertical annular two-phase flow; interfacial friction factor; support vector regression machine; particle swarm algorithm

1. Introduction

Vertical annular two-phase flow (as shown in Figure 1) is a type of gas-liquid two-phase flow in a vertical pipe [1]. A vertical annular flow occurs in the processes of steam generation, gas-liquid mixing and separation, oil and gas production and transportation. In particular, the gas occupies the bulk volume of the pipeline, and the liquid mainly distributes on the wall of pipeline in the form of liquid film. Entrained droplets could be formed in the liquid, and they flow in the gas stream along the pipeline [2]. In the process of a vertical annular flow, liquid droplets generated by the liquid film could enter the gas core and droplets embedded in the gas core would deposit into the liquid film. The entire complex process results in continuous exchange of mass, momentum and energy at the gas-liquid interface [3]. In order to explore in-depth the physics of vertical annular flows, many models have been proposed [4,5]. In these models, the pressure drop is dominated by friction force and gravity. Interfacial shear stress is a significant parameter for calculation of friction force, which remains difficult to measure. Experimental observations in the literature show that there were a few types of fluctuations at the gas-liquid interface, leading to the complex interfacial structure [6]. These fluctuation waves are generally classified into two categories, namely the large wave and the wavelet, according to parameters, such as relative wave height (relative average liquid film thickness), distribution continuity and

average liquid film flow rate [7,8]. Wavelets are waves of low velocity and small relative wave height. Large waves have relatively large wave height and long fluctuation distance, and they usually occur when the gas velocity is high and have dominant influence on the behavior at the gas-liquid interface, also known as disturbance wave. Because of the existence of these fluctuations, the flow structure of the liquid film is quite similar to the rough wall of a single-phase pipe flow. Therefore, many researchers study annular flow by analogy with turbulent flow of a single-phase rough pipe flow. Specifically, the liquid film is comparable to the rigid wall surface, the fluctuation to the tube wall roughness and the air core to the pipe flow core [9]. Based on this hypothesis, researchers have established a numerical model considering the interfacial shear stress. However, a large number of parameters have to be carefully measured and calibrated in order to apply the model.
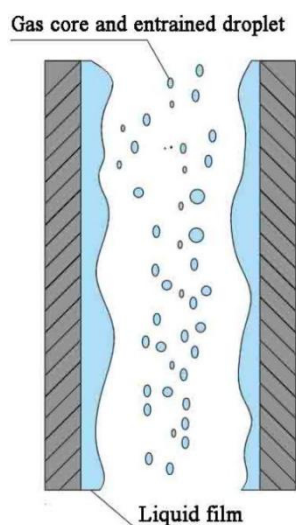

Figure 1. Flow pattern of vertical annular flow.

The interfacial shear stress is directly related to the interfacial friction factor. Accurate prediction of the interfacial friction factor could greatly simplify the calculation of the shear stress. Therefore, empirical formulas of the interfacial friction factor of an annular flow have been proposed by a number of researchers. Wallis et al. [10] proposed an empirical formula assuming that the interfacial friction factor was proportional to the relative liquid film thickness, and the Wallis formula is widely used in the literature. However, this formula is only applicable to cases of small relative liquid film thickness. Since then, researchers have made great improvements to the Wallis formula and developed different prediction formulas [1,11–22]. These extended formulas however have limitations for different ranges of superficial liquid and gas velocity. Therefore, it is necessary to develop a general and reliable method to more accurately predict the interfacial friction factor.

Machine learning (ML) techniques have been developed in recent years to make predictions for problems with a large number of influencing factors. One of the well-established ML techniques is the support vector machine (SVM) algorithm, which was firstly developed by Vapnik et al. at AT&T Bell laboratories [23–28]. As a matter of fact, it is difficult to find solutions of nonlinear problems directly, and they can be solved as linear separation problems. The SVM maps training samples to the high-dimensional characteristic space through nonlinear transformation, where a classification hyperplane is established and these samples can be separated linearly. As a general ML method, it follows the principle of structural risk minimization and shows a number of excellent advantages in solving problems of small sample size and nonlinear pattern recognition and can be applied to nonlinear regression problems. Based on applications of different problems, SVM is divided into support vector classification (SVC) and support vector regression (SVR). The application of SVM for regression prediction analysis is known as the support vector regression machine method. In order to find the best parameters for SVR, particle swarm

optimization (PSO) could be employed to optimize the searching process and improve the efficiency. PSO is an evolutionary algorithm developed by Kennedy et al. [29] and is of fast convergence and high precision. Similar to the simulated annealing algorithm, PSO starts from a random initial solution, searches for the optimal solution through iterations, evaluates the adaptability of solutions by a fitting function and updates the global optimal value after comparing with the current optimal value. Instead of applying empirical formulas, in this work, we develop a combined PSO-SVR model to compute the optimal friction factor for the vertical annular flow. In contrast to empirical formulas, the proposed PSO-SVR model has the capabilities of dealing with a wider range of flow conditions. It is able to analyze the influence of various factors on the change of interfacial friction factor due to the nature of the ML techniques.

We collect the empirical formulas for vertical annular flow available in the open literature and discuss the limitation of these formulas under different experimental conditions. Then, we analyze factors correlated with the friction factor and develop a prediction model for the interfacial friction factor in vertical annular flows using ML techniques. An examination of published models for the interfacial friction factor is carried out using data collected from the literature for vertical annular flows. By comparison with results from empirical formulas, we show that the proposed model is promising for calculating the friction factor for vertical annular flows in a wide range of parameters.

2. Methodology

2.1. Interfacial Friction Factor Calculation Model and Empirical Formulas

With the assumption that the droplet entrainment rate of the gas core and deposition rate in the gas core were equal and the axial flow rate was constant in the fully developed vertical film region, the momentum conservation equation for the vertical annular flow can be derived and be simplified to the interfacial shear stress formula [21,30] as

$$\tau_i = R_E U_E - R_D U_D - \frac{A_C}{S_C}\left(\frac{dp}{dz} + \frac{d}{dz}\left(\rho_C U_Z^2\right) + \rho_C g\right) \tag{1}$$

where dp/dz is the flow pressure difference, ρ_C is the air core density (when there are small droplets in the air core, it represents the mixture density of the air core and the droplets), R_E is the entrainment rate of droplets atomized from liquid film to gas core and R_D is the deposition rate of droplets to liquid film in the gas core. A_C is the cross section area of the gas core, S_C is the perimeter of the cross section of the gas core, U_E is the average axial velocity of liquid droplets and U_D is the droplet deposition velocity.

From the perspective of the gas flow, the liquid film can be taken as a rigid and rough wall with the velocity of $U_G + U_i$, where U_G is the gas flow velocity and U_i is the interface velocity. Compared with the gas flow velocity, the interface velocity is much smaller and can been neglected. Therefore, the friction factor [1,10] is expressed as:

$$f_i = \frac{2\tau_i}{\rho_G U_G^2} \tag{2}$$

where f_i is the interfacial friction factor of the vertical annular flow, τ_i is the interfacial shear stress, and ρ_G is the gas density. Substituting the shear stress in Equation (1) into the correlation, the interfacial friction factor could be obtained.

Through experimental analysis, researchers found that the friction factor correlates with relative liquid film thickness, gas Reynolds number, liquid Reynolds number and other parameters. Wallis [10] proposed an empirical formula to calculate friction factor using the relative liquid film thickness, which was then extended by a number of authors in a variety of forms for different conditions, as summarized in Table 1.

Table 1. Empirical formulas for interfacial friction factors.

Author	Formula
(Wongwises et al., 2001) [1]	$f_i = 17.172 Re_G^{-0.768}\left(\frac{t}{D}\right)^{-0.253}$
(Belt et al., 2009) [13]	$f_i = 1.158 t/D + 3.413 \times 10^{-4}$
(Henstock et al., 1976) [14]	$\frac{f_i}{f_G} = 1 + 212\sqrt{\frac{f_i}{f_G}}\frac{t}{D}, f_G = 0.046 Re_G^{-0.02}$
(Moeck, 1970) [15]	$f_i = 0.005\left[1 + 1458(t/D)^{1.42}\right]$
(Hori et al., 1978) [16]	$f_i = 1.13 Re_G^{-0.89} Re_L^{0.68} Fr_G^{0.25} Fr_L^{-0.45}(\mu_L/\mu_w)^{0.77}$ $Fr_G = u_{sg}/\sqrt{gD}, Re_G = \rho_g u_g D/\mu_g$
(Fukano et al., 1997) [19]	$f_i = 0.425(12 + v_l/v_w)^{-1.33}(1 + 12t/D)^8$
(Pan et al., 2015b) [22]	$f_i = 67.2 Re_G^{-0.91} Re_L^{0.30}$
(Wallis, 1969) [10]	$\frac{f_i}{f_G} = 1 + 300\frac{t}{D}, f_G = 0.005$
(Fore et al., 2000) [21]	$\frac{f_i}{f_G} = 0.005\left\{1 + 300\left[\left(1 + \frac{17500}{Re_G}\right)\frac{t}{D} - 0.0015\right]\right\}$

2.2. Machine Learning Method

2.2.1. Support Vector Regression Machine

Given a set of training data $(x_1, y_1), \ldots, (x_i, y_i), y \in [+1, -1], i = 1, 2, \ldots, l$, linear regression function can be used to fit the data as

$$f(x) = \omega x + b \tag{3}$$

where ω represents a multidimensional vector and b represents a real number. For a nonlinear regression problem, samples will be analyzed in a feature space of higher dimension (Hilbert space) through the non-linear mapping function φ: $X \to \varphi(X)$, as shown in Figure 2.

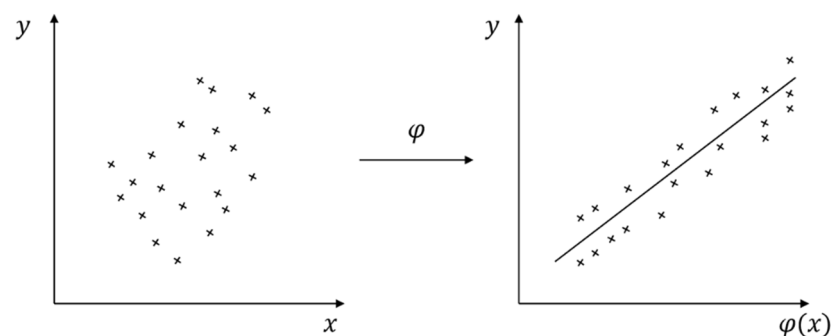**Figure 2.** A simplified diagram for converting a nonlinear problem to a linear problem through the mapping function φ: $X \to \varphi(X)$.

According to the structural risk minimization criterion, the regression estimation problem can be transformed into an optimization problem with a ε-insensitive loss function as

$$\begin{aligned} min \quad & \tfrac{1}{2}||\omega||^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i'\right) \\ s.t. \quad & \begin{cases} y_i - \omega x_i - b \leq \varepsilon + \xi_i \\ \omega x_i + b - y_i \leq \varepsilon + \xi_i' \\ \xi_i, \xi_i' \geq 0 \end{cases} \end{aligned} \tag{4}$$

where the constant C is positive and represents the penalty degree of samples outside the ε interval illustrated in Figure 3a. The ε is used to control the complexity of the regression

functions and the ε-insensitivity loss function is explained in Figure 3b. The slack variables ξ_i, ξ_i' are used to consider the fitting error.
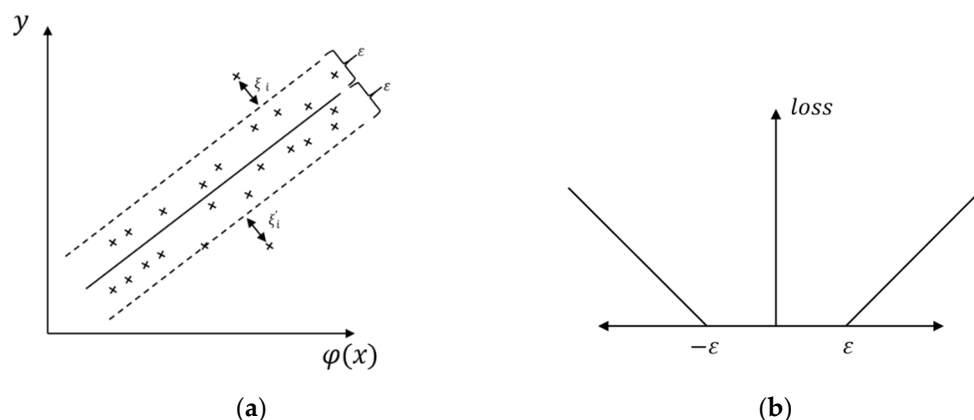

(**a**)　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 3. Diagram showing the ε-insensitive loss function and corresponding loss function. (**a**) feature space with ε interval (the points on or outside of the interval margin are support vectors). (**b**) the ε-insensitive loss function (the slope is determined by *C*).

Kernel functions, as the projecting function, are introduced to make the calculation directly on the input space, which is a real symmetric positive function and satisfies the Mercer condition [31–34]

$$\iint k(x,x')g(x)g(x')dxdx' \geq 0 \quad g(x) \in L_2\left(R^N\right),\ k(x,x') \in L_2\left(R^N \times R^N\right) \tag{5}$$

Common kernel functions include linear kernel function, polynomial kernel function, Gaussian kernel function, etc. Among them, Gaussian kernel function has a strong learning ability and is selected as kernel function of the model developed in this study. The formula is as follows:

$$(x_i, x_j) = exp\left(-\gamma||x_i, -x_j||\right) \tag{6}$$

2.2.2. Particle Swarm Optimization

The PSO algorithm was firstly proposed by Kennedy and Eberhart in 1995, which was originated from the study on bird behaviors [35]. Each particle in the algorithm represents a potential solution to the problem, which corresponds to a "fitness" determined by an optimization function. The function has no mass and volume but only speed and position. The velocity of a particle determines its flight direction and distance. The target function is gradually optimized through the updated information of velocity and position. The initial state of a particle swarm is a group of random particles, and iterations are carried out to find the optimal position of the whole population, that is, the global optimum. At every time step when a particle position is updated, a 'fitness' value needs to be calculated. The fitness values of the individual optimal position and the group optimal position are compared with that of the new position, and updated accordingly. After multiple iterations, the whole group gradually moves to the optimal solution of the target function. The process is shown in Figure 4.

Suppose a population of m particles is distributed in the D-dimensional search space, which is $X = (X_1, X_2, \ldots X_m)$. At the time t, the position of the ith particle is $x = [x_{i1}(t), x_{i2}(t), \ldots, x_{iD}(t)]$, the velocity is $v = [v_{i1}(t), v_{i2}(t), \ldots, v_{iD}(t)]$, the optimal position of each particle in iteration is $P_i(t) = [P_{i1}(t), P_{i2}(t), \ldots, P_{iD}(t)]$, and the optimal position of the population is $P_g(t) = [P_{g1}(t), P_{g2}(t), \ldots, P_{gD}(t)]$. In order to better balance the global and local search ability and ensure convergence, Shi and Eberhart [36] introduced an inertia weight ω into the velocity equation and obtained
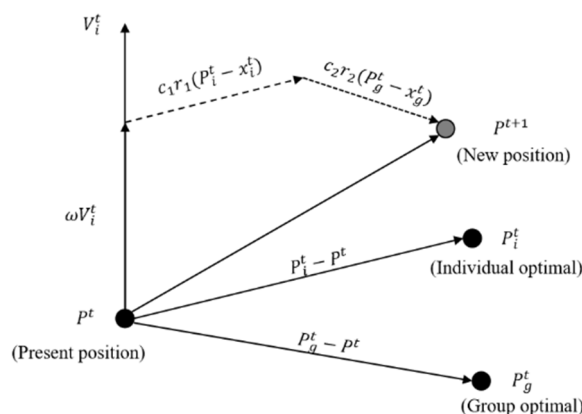
Figure 4. The iterative process of particle velocity and position.

$$\omega = \omega_{max} - \frac{t}{t_{max}}(\omega_{max} - \omega_{min}) \tag{7}$$

where ω_{max} is the initial inertia weight, ω_{min}. the final inertia weight, and t_{max} the maximum iteration number.

$$v_i^{t+1} = \omega v_i^t + c_1 r_1\left(p_i^t - x_i^t\right) + c_2 r_2\left(p_g^t - x_g^t\right)$$
$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{8}$$

where c_1 is cognitive coefficient and c_2 is social coefficient; r_1, r_2 are random numbers in (0,1). To avoid blind search of particles, the position and the speed are generally limited to the interval $[-X_{max}, X_{max}]$ and $[-V_{max}, V_{max}]$, respectively.

3. Model Development and Parametric Analysis

The SVR algorithm is employed to establish the model, with parameters optimized by a PSO algorithm. The model development process involves data processing, SVR input variable analysis and parameter optimization.

3.1. Data Processing

Data processing is needed before model training and prediction, which includes data collection and data normalization and denormalization. The database is established by collecting the data from available literature. Data normalization helps to accelerate numerical convergence, and denormalization is used to obtain the real values from predicted results.

3.1.1. Data Collection

For our analysis of vertical annular two-phase flow, 268 data points were collected from 4 sources, with pipe diameter D ranging from 16–50.8 mm. Most of the data sets were recorded at atmospheric pressure, with gas speed ranging from 9–132 m/s and liquid speed from 0.001–0.200 m/s. Details of the data are shown in Table 2.

Table 2. Experimental interfacial friction factor of vertical annular flow.

Researchers	U_g	U_L	D (mm)	No. Points
(Wongwises et al., 2001) [1]	9–34	0.05–0.20	29	39
(Fore and Dukler, 1995) [30]	16–37	0.006–0.066	50.8	44
(Asali, 1984) [37]	14–96	0.001–0.128	22.9; 42	161
(Shearer and Nedderman, 1967) [38]	28–132	0.002–0.017	16; 32	24

The selected data points have a wide range of distribution. The Hewitt & Roberts (HR) flow pattern graph has been the most widely used flow pattern diagram, where different

flow regimes are determined by the characterisics of gas and liquid. We see in Figure 5 that our selected data points all fall in the annular flow regime.
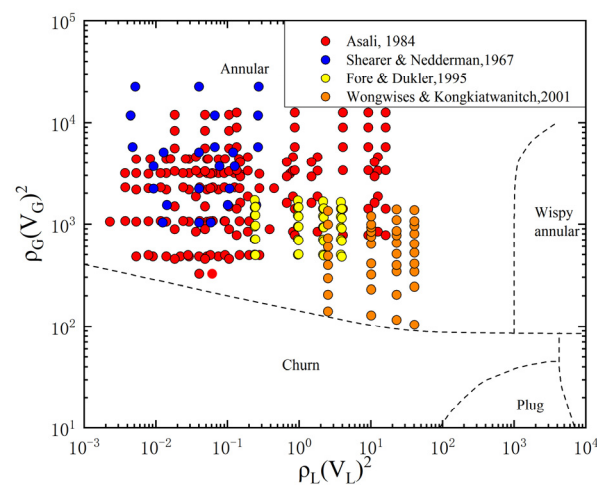

Figure 5. Distribution of experimental data against Hewitt and Roberts flow regime map [39] showing all data are in the annular flow regime.

In our ML model, the collected data are grouped into a training set and a testing set. The training set accounts for 90% (241) of the data bank, which is utilized to optimize the developed ML model by searching the best parameters. Normally, multiple-fold validation approach is employed to optimize the training process. The remaining 10% (27) is applied as the testing set to validate the accuracy of the optimized model. For each of the sources in Table 2, data points from the same source are randomly mixed and allocated into the training set and the testing set in proportion. For instance, with the 161 data points from Asali [37], 145 data points were randomly selected as the training set and the rest 16 as the testing set. With the splitting of each data source, the training sets and testing sets from different sources are merged respectively to form the final training set and testing set for our model.

3.1.2. Data Normalization and Denormalization

Before model training, all data should be normalized to eliminate large prediction errors caused by the order difference. In this paper, the normalization formula is given as follows

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{9}$$

where x is the original input data, x_{min} the minimum value of the original data, x_{max} the maximum value of the original input data, and $x_{original}$ the normalized solution. The model training and prediction are performed after data normalization. With normalization, the input data and the calculated results are all in the range of 0 to 1. In order to obtain the actual results, the predicted data shall be recovered by the following equation

$$y = y_{predicted} \times (x_{max} + x_{min}) \tag{10}$$

where $y_{predicted}$ is the output predicted by the ML model and y is the denormalized output.

3.2. SVR Input Variable Analysis

To ensure dimensional consistency, the input variables of the SVR should be dimensionless. However, there are no definite and general models to determine the factors that influence the interfacial friction factor. Based on the existing empirical formulas, five most relevant parameters are selected, i.e., relative film thickness, gas Reynolds number, liquid Reynolds number, gas Froude number and liquid Froude number. To investigate the

relationship between these five parameters and the interfacial friction factor, sensitivity analysis is conducted.

As can be seen from Figure 6, the overall trend shows that the interfacial friction factor is gradually enlarged with the increase of the relative liquid film thickness (the ratio of film thickness to pipe diameter). There is a clear correlation between the interfacial friction factor and the relative film thickness, which has been presented by Wallis et al. [10], Belt et al. [13] and Moeck [15]. However, for the data obtained under different experimental conditions, the exact relationship might be different. Therefore, it will not be sufficient to predict a wide range of interfacial friction factors based on this correlation alone.
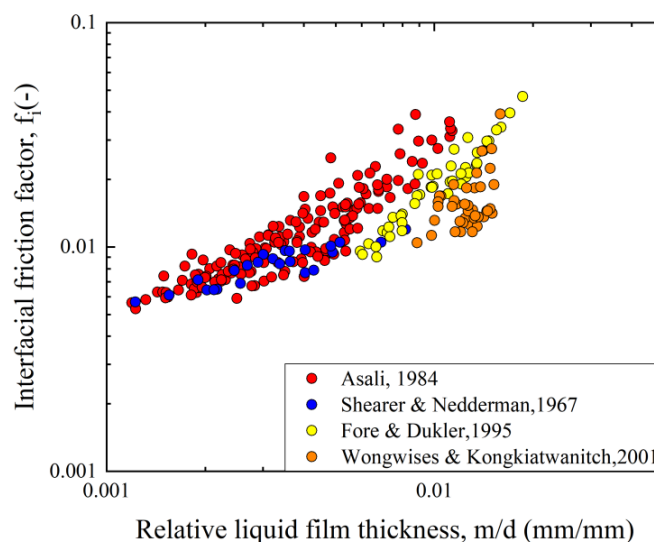

Figure 6. Variation of interfacial friction factor with relative liquid film thickness.

Normally, the vertical annular flow is modelled as a single-phase flow through a rough pipe as seen in Figure 1. In the analogy, the uneven liquid film can be regarded as the rough wall surface where the film surface motion represents the variation of the wall roughness. The difference between the gas and liquid velocity would affect the motion of the interface, such that the interfacial "friction" factor can be changed at different flow status. In addition, the fluid density and viscosity are important physical quantities for the flow velocities. Fukano and Furukawa [19] proposed an empirical formula of interfacial friction factor with consideration of fluid viscosity. In order to describe the common influence of fluid density, viscosity and flow rate, the Reynolds numbers of gas and liquid can be utilized to characterize the flow condition. In a single-phase flow, the friction factor of the transitional rough zone is affected by the roughness height and the Reynolds number. With the reference to single-phase flow analysis, the Reynolds number is included in many empirical formulas for annular flow. Fore et al. [21] noted that the Wallis formula could overpredict the friction factor at very small values of the relative film thickness and modified the interfacial friction factor formula with the introduction of the gas Reynolds number. Compared with the Wallis formula, the suggested formula by Fore et al. [21] showed better prediction accuracy in a wider range of film thickness. Similarly, the modification was adopted in the formula of Wongwises and Kongkiatwanitch [1]. Liquid Reynolds number was also included in the formulas proposed by Hori et al. [16] and Pan et al. [22]. Here, the Reynolds numbers for liquid and gas are defined as $Re_G = \rho_G v_G D / \mu_G$ and $Re_L = \rho_L v_L D / \mu_L$, respectively, where the subscripts G and L represent gas and liquid, respectively, and ρ, v, μ are the density, velocity and viscosity of gas and liquid.

The vertical annular flow is mainly governed by gravity, inertial force and friction force. When calculating the interfacial friction factor, it is necessary to consider the influence of gravity and inertial force on the flow. The fluid Froude number is defined as the ratio of inertial force to gravity of fluids in the vertical pipe. The dimensionless gas and liquid Froude numbers can properly represent the influence of gravity and inertial force on the

change of friction. The empirical formula proposed by Hori et al. [16] has introduced gas and liquid Froude numbers into the calculation of the interfacial friction factor and produces good predictions. Similarly, we introduce the Froude numbers into our prediction model. The gas and liquid Froude numbers can be written as $Fr_G = v_G/\sqrt{gD}$ and $Fr_L = v_L/\sqrt{gD}$. With the collected data from available studies, plots of f_i versus Re_G, Re_L, Fr_G and F_L values are illustrated in Figure 7, respectively. It is clear that the f_i magnifies as the gas Reynolds number increases, which shows the same trend as that of f_i with the gas Froude number. However, it goes conversely for the relationship between f_i and liquid Reynolds number, as well as liquid Froude number. As can readily be seen, differences of viscosity and diameter could only influence the slope and point distribution, while the variance of the interfacial friction factor mainly depends on gas velocity and liquid velocity.
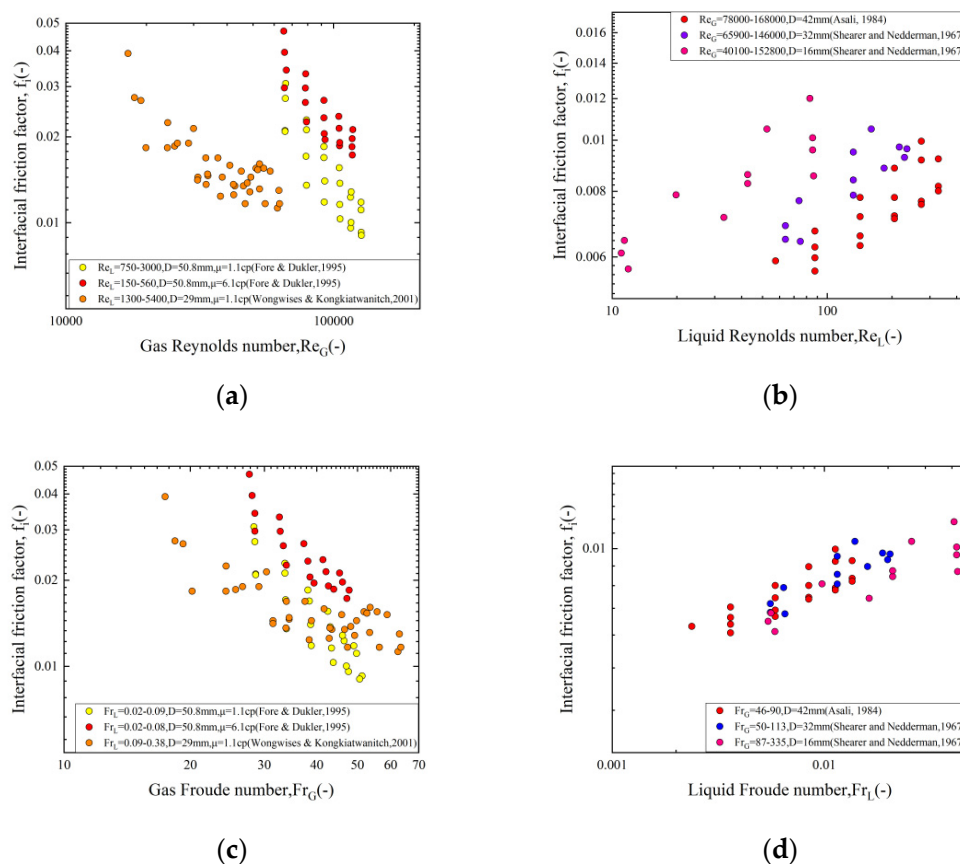

Figure 7. Variation of the interfacial friction factor with (**a**) Gas Reynolds number, (**b**) liquid Reynolds number, (**c**) gas Froude number and (**d**) liquid Froude number.

The above analysis indicates that f_i shows apparent relevance with relative thickness, Re_G, Re_L, Fr_G and F_L. Thus, we select these five parameters as the input variables of our PSO-SVR model.

3.3. Parameter Optimization of PSO-SVR Model

The prediction accuracy of the SVR model is determined by the model's control parameters. In this paper, the PSO algorithm is used to determine the optimal control parameters of the SVR model. Firstly, we construct the PSO-SVR model structure by programming, and we set the control parameters of the PSO algorithm, including weight ω, global search coefficient c_1, local search coefficient c_2, maximum evolution algebra max_{gen} and population $size_{pop}$. Then, we initialize the SVR model operation parameters, including the penalty coefficient C, g of the kernel function and the loss coefficient P. A larger C means it is easier for the model to be overfitted. g is a parameter of the RBF kernel, which implicitly determines the distribution of data mapped to a new feature space. A larger g

indicates a smaller support vector. In this paper, we set C $\in [10^{-3}, 10^3]$, $g \in [10^{-3}, 10^3]$ and P as 0.01 and use the four-order cross validation method to obtain the mean square error of solutions from each generation of the PSO-SVR model. Under different SVR parameter combination, the prediction accuracy of the SVR model is calculated and taken as the fitness for the PSO iteration process. As discussed before, the relative liquid film thickness, gas Reynolds number, liquid Reynolds number, gas Froude number and liquid Froude number are chosen as input parameters, and the interfacial friction factor is the desired output solution. After sufficient training, the built model reaches its optimal prediction performance with the best SVR parameter combination. Finally, the performance of the optimal SVR model is assessed by comparing the testing data and the result predicted from the optimal model. The procedure to make prediction of the interfacial friction factor using the PSO-SVR model is demonstrated in Figure 8.
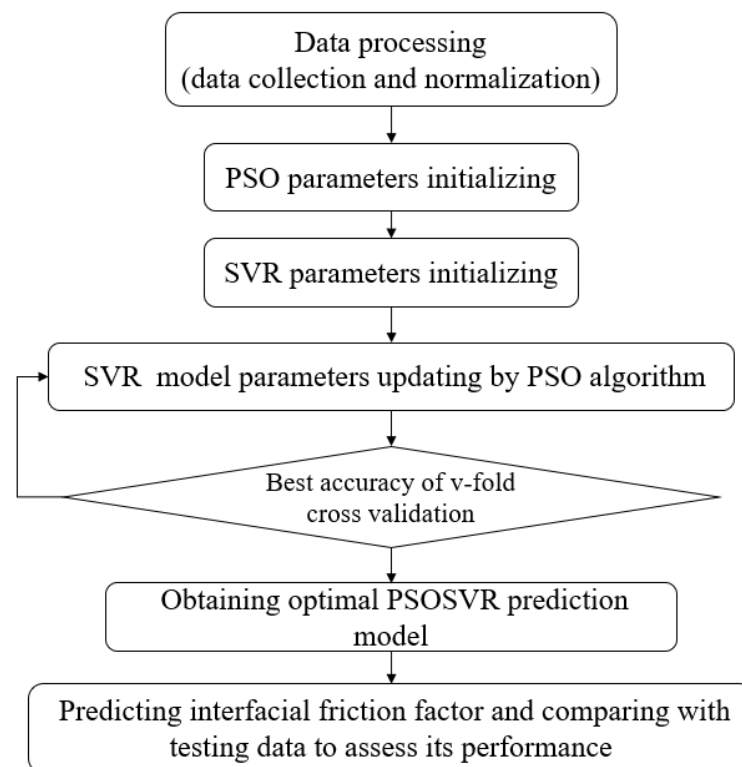

Figure 8. Flow chart of SVR optimization by PSO algorithm.

4. Result

In order to achieve better prediction, we assess the model performance under different combinations of parameters, including parameter combination 1 (h/D, Re_G, Re_L) and parameter combination 2 (h/D, Re_G, Re_L, Fr_G, F_L). The prediction accuracy of our PSO-SVR model is compared with that of formulas collected from previous studies shown in Table 1. Several parameters are selected as the evaluation indexes to numerically report the performance of the model, containing the correlation coefficient (r), coefficient of determination (R^2), root mean square error ($RMSE$), maximum absolute error (MAE), maximum relative error (MRE) and average relative error (AAE) as defined in Table 3. Among them, r represents correlation between the experimental f_i and the predicted f_i. $RMSE$ is used to quantify the stability of prediction performance, and MAE, AAE and MRE are used to quantify the overall prediction accuracy [40]. The PSO algorithm is applied to optimize the SVR model with control parameters shown in Table 4.

Table 3. Formulas of evaluation index.

Evaluation Index	Value
Maximum absolute error, *MAE*	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i \rvert$
Maximum relative error, *MRE*	$MRE = max\frac{\lvert y_i - \hat{y}_i \rvert}{y_i}$
Average absolute error, *AAE*	$AAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i \rvert$
Correlation coefficient, *r*	$r = \frac{\sum_{i=1}^{n}(y_i - \overline{y}_i)(\hat{y}_i - \overline{y}_i)}{\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \sum_{i=1}^{n}(\hat{y}_i - \overline{y}_i)^2}}$
Coefficient of determination, R^2	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(\hat{y}_i - \overline{y}_i)^2}$
Root mean square error, *RMSE*	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - \overline{y}_i)^2}$

Table 4. PSO control parameters.

Control Parameter	Value	Control Parameter	Value
Original inertia weight ω	0.9	Final inertia weight ω	0.4
Global search factor c_1	1.5	max_{gen}	200
Local search factor c_2	1.5	$size_{\text{pop}}$	50

4.1. Comparison of Different Input Parameter Combinations

Under different input parameter combinations, the optimal control parameters of the SVR model are set separately. The numerical analysis of the built model on training, testing and total data sets are listed in Table 5. The model based on parameter combination 2 (h/D, Re_G, Re_L, Fr_G, F_L) performed much better with higher correlation coefficient ($r = 0.995$) and coefficient of determination ($R^2 = 0.989$), lower maximum absolute error ($MAE = 0.006$),average absolute error ($AAE = 0.0015$) and lower root mean square error ($RMSE = 0.00074$) compared with the other model. The agreement between experimental f_i and predicted f_i is significantly improved, since the maximum absolute error (MAE) and average absolute error (AAE) decline from 0.019 to 0.006 and 0.0015 to 0.0004, respectively. It is concluded that the introduction of gas and liquid Froude number significantly improves the accuracy and stability of the model. Thus, the optimal model should be determined on parameter combination 2, whose control parameters are optimized as $C = 83.78$, $g = 1$ in this study. The searching process of the PSO algorithm is illustrated in Figure 9, and we see that the SVR model eventually becomes stable after 120 iterations. Figure 10 shows the comparison of the experimental f_i and the predicted f_i by the optimal model. We see that the prediction has only noticeable errors at certain large values of the interfacial friction factor.

Table 5. Model evaluation results.

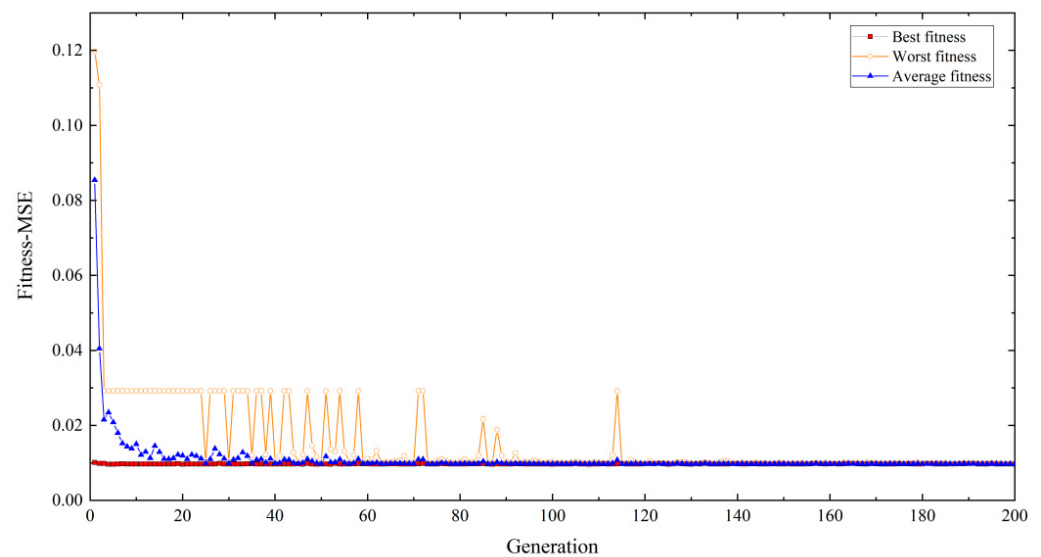| Input Parameter Combination	Dataset	Statistical Parameters					
		RMSE	*MAE*	*AAE*	*MRE*	*r*	R^2
1:h/D, Re_G, Re_L	Training set	0.00295	0.019	0.0016	9.63%	0.922	0.851
	Testing set	0.00126	0.004	0.0010	6.96%	0.969	0.939
	Total	0.00283	0.019	0.0015	9.36%	0.924	0.853
2:h/D, Re_G, Re_L, Fr_G, F_L	Training set	0.00074	0.006	0.0004	3.09%	0.995	0.990
	Testing set	0.00087	0.003	0.0007	4.91%	0.987	0.975
	Total	0.00076	0.006	0.0004	3.27%	0.995	0.989

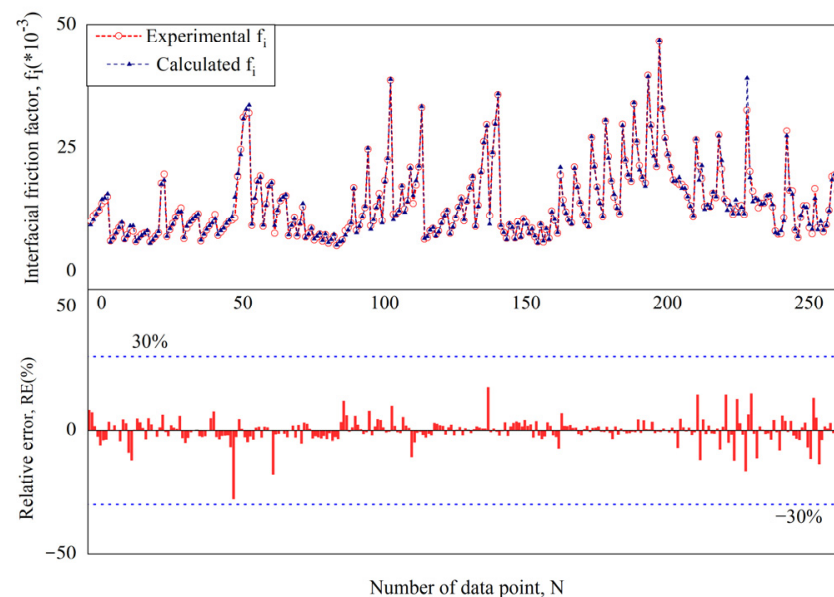**Figure 9.** Fitness optimization process of the built PSO-SVR model.

Figure 10. Comparison of the experimental f_i and predicted f_i by the optimal model.

4.2. Comparison of Optimal Model with Empirical Formulas

To test the reliability of the optimal model, a comparison between the existing empirical formulas and the optimal PSO-SVR model is conducted. The available empirical formulas from 9 sources are gathered, with detailed information listed in Table 1. Scatter diagrams in Figure 11 show the comparison between the calculated f_i and the experimental f_i. The evaluation results of each formula have been summarized in Table 6.

In Figure 11, the points distributed on the diagonal line means the predicted results are perfectly consistent with the actual results in experiments, and the points away from the diagonal tell the increasing prediction error. The interval between dash lines represents a 50% error band, directly displaying the prediction accuracy of each method, with visualization of points falling in this region.
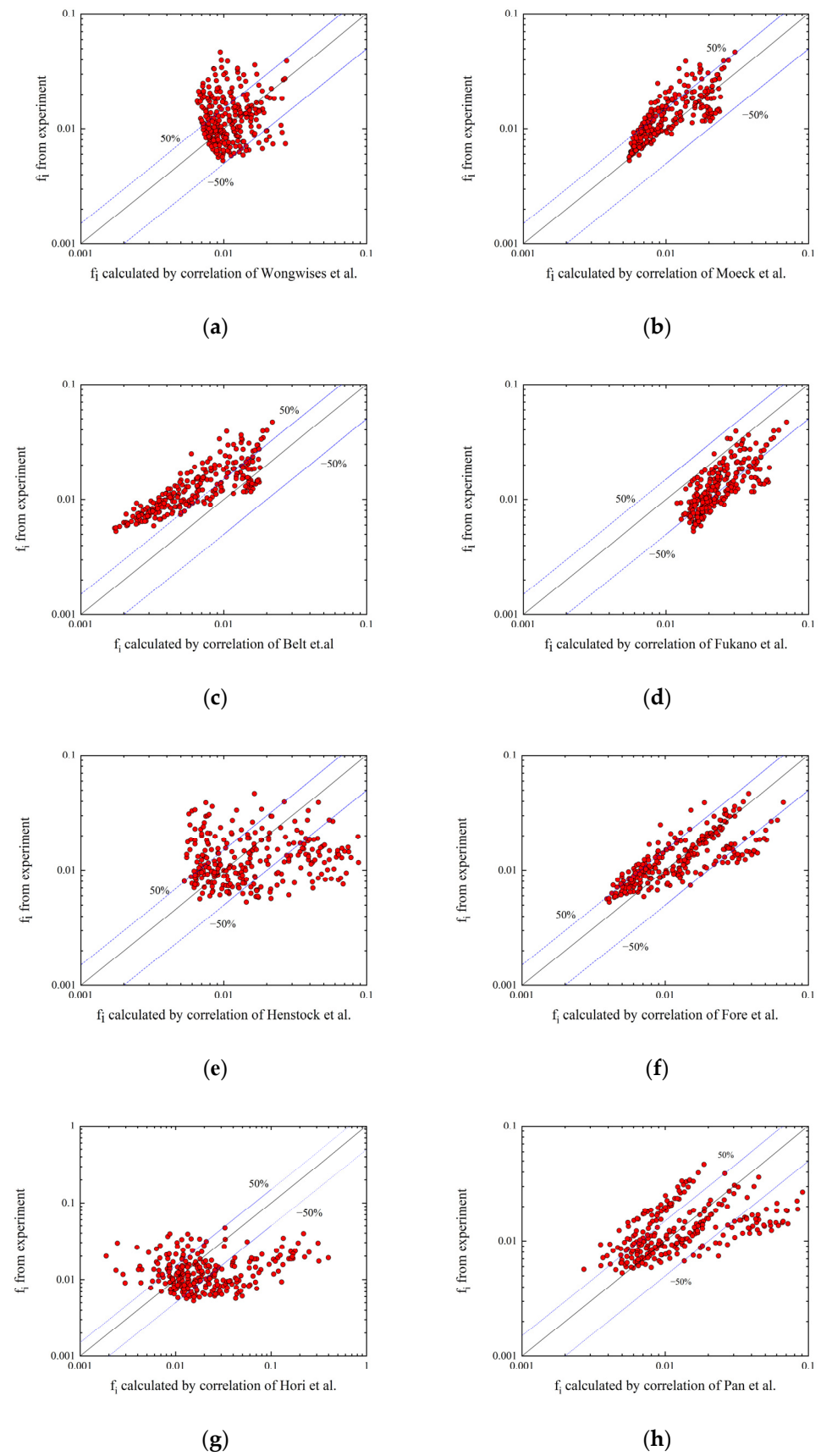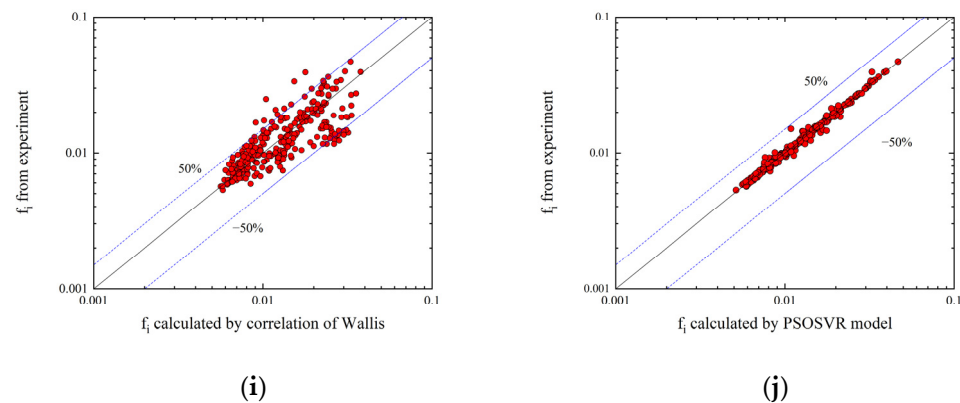
(**a**)

(**b**)

(**c**)

(**d**)

(**e**)

(**f**)

(**g**)

(**h**)

Figure 11. *Cont.*

(i)   (j)

Figure 11. Comparison between the calculated f_i and experimental f_i for empirical formulas and PSO-SVR model. figures (**a–j**) represent the results calculated by correlation of (**a**) Wongwises et al. (**b**) Moeck et al. (**c**) Belt et al. (**d**) Fukano et al. (**e**) Henstock et al. (**f**) Fore et al. (**g**) Hori et al. (**h**) R Pan et al. (**i**) Wallis (**j**) PSOSVR et al.

Table 6. Evaluation results of empirical formulas and PSO-SVR model.

No.	Formula	$N_{re<50\%}$	RMSE	MAE	AAE	ARE	R^2
1	Wongwises et al.	184	0.00830	0.0374	0.0057	40.34%	0.0134
2	Moeck et al.	243	0.00560	0.0252	0.0039	24.49%	0.5382
3	Belt et al.	126	0.00780	0.0284	0.0064	47.19%	0.5543
4	Fukano et al.	61	0.01320	0.0381	0.0113	95.88%	0.5412
5	Henstock et al.	130	0.02040	0.0760	0.0134	108.51%	0.0034
6	Fore et al.	218	0.00860	0.0334	0.0054	38.44%	0.4318
7	Hori et al.	101	0.05540	0.3784	0.0267	173.65%	0.0855
8	Pan et al.	179	0.01730	0.0849	0.0099	64.77%	0.1543
9	Wallis	232	0.00570	0.0211	0.0038	26.54%	0.4926
10	PSO-SVR	268	0.00076	0.0060	0.0004	3.27%	0.9890

$N_{re<50\%}$ represents the number of points in 50% error band.

As shown in Figure 11i, the classic formula developed by Wallis [10] makes good predictions, with *AAE* equaling 0.0038, and most of the data points uniformly lie in the 50% error band. However, with the expansion of the interfacial friction factor range, its prediction error increases gradually. Additionally, the calculated f_i of formulas proposed by Moeck and Fore et al. display good agreements with experimental results (*AAE* lower than 0.0060 and more than 80% of points in the 50% error band) with a better coefficient of determination (R2 over than 0.50). Nevertheless, others in Figure 11a–h predict a smaller number of points in the 50% error band and show higher *AAE* (over 0.0060). For our PSO-SVR model, it can be seen from Figure 11j that the prediction is stable with almost all points lying along the diagonal line. The accuracy is not sensitive to the change of f_i, which demonstrates its robustness in application. Compared with other formulas, the *AAE* of the present model remains the lowest in both the testing data set (0.0007) and the total data set (0.0004), with *MAE* lower than 0.01. To date, there is no widely used formula to calculate interfacial friction factor in all ranges of flow conditions. Nevertheless, with the development of computer science, it is possible to develop ML-based numerical models without any explicit formula. The above analysis suggests that the proposed PSO-SVR model is stable and can be applied in a wider range of parameters for prediction of the interfacial friction factor.

In order to further demonstrate the prediction stability of the model, 4 extra cases are established with the same sampling method described before. In Cases 1 and 2, 90% of the database is used as the training set, and 10% of the database is used as the testing set. In Cases 3 and 4, 80% of the database is the training set, and 20% of the database is the testing set. The prediction results in each case are shown in Figure 12.
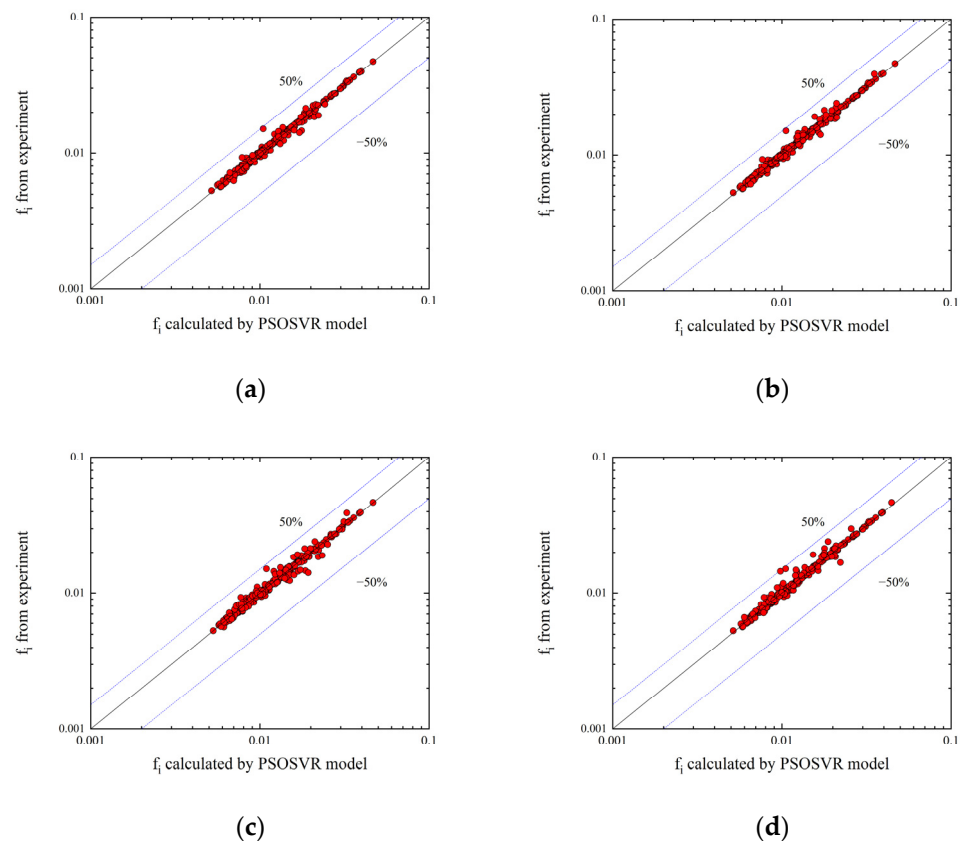
(**a**)

(**b**)

(**c**)

(**d**)

Figure 12. Comparison between the calculated f_i and experimental f_i for different cases: (**a**) Case 1, (**b**) Case 2, (**c**) Case 3 and (**d**) Case 4.

As shown in Figure 12, all data points for each case fall into the 50% error band and are distributed closely to the diagonal line. The *AAE*s for Cases 1 and 2 are the same (*AAE* = 0.0004), and it increases slightly to 0.0005 and 0.0006 for Case 3 and Case 4. The *MAE*s are 0.0046 and 0.0045 for Case 1 and Case 2, and 0.0054 and 0.0066 for Case 3 and Case 4. Due to limited amount of data, we have to reduce the number of training data when allocating more data for prediction. Nevertheless, the prediction error increases slightly, suggesting that reducing slightly the training data does not affect the overall prediction.

5. Conclusions

In this paper, we develop a machine learning model based on the SVR-PSO algorithms to predict the interfacial friction factor of a vertical annular two-phase flow.

From analysis of the input variables, five parameters (the relative film thickness, the Reynolds numbers of gas and liquid, and the Froude numbers of gas and liquid) show significant correlation with the interfacial friction factor and are considered as the input parameters of our model. By comparisons of different parameter combinations, it is concluded that the model considering the five parameters performs the best, with high accuracy and stability. With the well-trained model, we compare the predicted results with those by the empirical formulas. The ability of the model learning and generalization is shown to be satisfactory, and the prediction accuracy is obviously improved compared with the empirical formulas. In order to confirm the stability and sensitivity of the model, four different cases are generated based on random sampling of the database. All predicted points of the four cases fall into the 50% error band, and error analysis shows that the prediction accuracy is insensitive to the variation of the dataset.

It has to state that the method proposed in this paper only considers the influence of liquid film and gas core. The entraining and deposition of droplets corresponding to energy transfer at the gas-liquid interface could have an important impact on the liquid film thickness and the gas-liquid velocities. In order to further improve the prediction

accuracy of the interfacial friction factor, the influence of entrained droplets in the gas core could be taken into account in the following work in the near future.

References

1. Wongwises, S.; Kongkiatwanitch, W. Interfacial friction factor in vertical upward gas-liquid annular two-phase follow. *Int. Commun. Heat Mass Transf.* **2001**, *28*, 323–336. [CrossRef]
2. Mcquillan, K.W.; Whalley, P.B. Flow patterns in vertical two-phase flow. *Int. J. Multiph. Flow* **1985**, *11*, 161–175. [CrossRef]
3. Aliyu, A.M.; Baba, Y.D.; Lao, L.; Yeung, H.; Kim, K.C. Interfacial friction in upward annular gas–liquid two-phase flow in pipes. *Exp. Therm. Fluid Sci.* **2017**, *84*, 90–109. [CrossRef]
4. Alves, I.M.; Caetano, E.F.; Minami, K.; Shoham, O. Modeling annular flow behavior for gas wells. *SPE Prod. Eng.* **1991**, *6*, 435–440. [CrossRef]
5. Cioncolini, A.; Thome, J.R.; Lombardi, C. Algebraic turbulence modeling in adiabatic gas–liquid annular two-phase flow. *Int. J. Multiph. Flow* **2009**, *35*, 580–596. [CrossRef]
6. Han, H.; Zhu, Z.; Gabriel, K. A study on the effect of gas flow rate on the wave characteristics in two-phase gas–liquid annular flow. *Nucl. Eng. Des.* **2006**, *236*, 2580–2588. [CrossRef]
7. Hewitt, G.F. *Annular Two-Phase Flow*; Pergamon Press: Oxford, UK, 1970.
8. Zhao, Y. Wave Behaviour in Vertical Multiphase Flow. Ph.D. Thesis, Imperial College London, London, UK, 2014.
9. Wang, Z.; Gabriel, K.S.; Manz, D.L. The influences of wave height on the interfacial friction in annular gas–liquid flow under normal and microgravity conditions. *Int. J. Multiph. Flow* **2004**, *30*, 1193–1211. [CrossRef]
10. Wallis, G.B. *One Dimensional Two-Phase Flow*; McGraw-Hill: New York, NY, USA, 1969.
11. Ambrosini, W.; Andreussi, P.; Azzopardi, B.J. A physically based correlation for drop size in annular flow. *Int. J. Multiph. Flow* **1991**, *17*, 497–507. [CrossRef]
12. Asali, J.C.; Hanratty, T.J.; Paolo, A. Interfacial drag and film height for vertical annular flow. *AIChE J.* **2010**, *31*, 895–902. [CrossRef]
13. Belt, R.J.; Westende, J.M.C.V.; Portela, L.M. Prediction of the interfacial shear-stress in vertical annular flow. *Int. J. Multiph. Flow* **2009**, *35*, 689–697. [CrossRef]
14. Henstock, W.H.; Hanratty, T.J. The interfacial drag and the height of the wall layer in annular flows. *AIChE J.* **1976**, *22*, 990–1000. [CrossRef]
15. Moeck, E.O. *Annular-Dispersed Two-Phase Flow and Critical Heat Flux*; Atomic Energy of Canada Ltd.: Joe Cleaver, ON, Canada, 1970.
16. Hori, K.; Nakasatomi, M.; Nishikawa, K.; Sekoguchi, K. Study of ripple region in annular two-phase flow (third report, effect of liquid viscosity on gas-liquid interfacial character and friction factor). *Trans. Jpn. Soc. Mech. Eng.* **1978**, *44*, 3847–3856. [CrossRef]
17. Fukano, T.; Ito, A.; Miyabe, K.; Takamatsu, Y. Liquid Films Flowing Concurrently with Air in Horizontal Duct: 6th Report, Generation of a Disturbance Wave and its Role in the Breakdown of Liquid films. *Plant Physiol.* **1985**, *28*, 2302–2309. [CrossRef]
18. Fukano, T.; Ousaka, A.; Kawakami, Y.; Tominaga, A. Interfacial shear stress and holdup in an air–water annular two-phase flow. In Proceedings of the ASME-JSME Thermal Engineering Joint Conference, Reno, NV, USA, 17–22 March 1991; Volume 2, p. 217.
19. Fukano, T.; Furukawa, T. Prediction of the effects of liquid viscosity on interfacial shear stress and frictional pressure drop in vertical upward gas–liquid annular flow. *Int. J. Multiph. Flow* **1998**, *24*, 587–603. [CrossRef]
20. Holt, A.J.; Azzopardi, B.J.; Biddulph, M.W. Calculation of two-phase pressure drop for vertical upflow in narrow passages by means of a flow pattern specific model. *Chem. Eng. Res. Des.* **1999**, *77*, 7–15. [CrossRef]
21. Fore, L.B.; Beus, S.G.; Bauer, R.C. Interfacial friction in gas–liquid annular flow: Analogies to full and transition roughness. *Int. J. Multiph. Flow* **2000**, *26*, 1755–1769. [CrossRef]
22. Pan, L.M.; He, H.; Ju, P.; Hibiki, T.; Ishii, M. The influences of gas–liquid interfacial properties on interfacial shear stress for vertical annular flow. *Int. J. Heat Mass Transf.* **2015**, *89*, 1172–1183. [CrossRef]

23. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the 5th Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; pp. 144–152.
24. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
25. Smola, A.J. Regression Estimation with Support Vector Learning Machines. Master's Thesis, Technische Universität München, München, Germany, 1996.
26. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. *Advances in Neural Information Process. Syst.* **1997**, *9*, 155–161.
27. Vapnik, V.; Golowich, S.E.; Smola, A. Support vector method for function approximation, regression estimation, and signal processing. *Adv. Neural Inform. Process. Syst.* **1997**, *9*, 281–287.
28. Burges, C.J. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]
29. Eberhart, R.C.; Kennedy, J. A new optimizer using particle swarm theory. In Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995; pp. 39–43.
30. Fore, L.B.; Dukler, A.E. Droplet deposition and momentum transfer in annular flow. *AIChE J.* **1995**, *41*, 2040–2046. [CrossRef]
31. Müller, K.R.; Smola, A.J.; Rätsch, G.; Schölkopf, B.; Kohlmorgen, J.; Vapnik, V. Predicting time series with support vector machines. In *Lecture Notes in Computer Science, Proceedings of the Artificial Neural Networks—ICANN'97, Lausanne, Switzerland, 8–10 October 1997*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 999–1004.
32. Smola, A.J.; Schölkopf, B.; Müller, K.R. General cost functions for support vector regression. In Proceedings of the 8th International Conference on Artificial Neural Networks, Skövde, Sweden, 2–4 September 1998.
33. Schölkopf, B.; Smola, A. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. *Publ. Am. Stat. Assoc.* **2002**, *98*, 489.
34. Shang, Y. Limit of a nonpreferential attachment multitype network model. *Int. J. Mod. Phys. B* **2017**, *31*, 1750026. [CrossRef]
35. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
36. Shi, Y.; Eberhart, R. A Modified Particle Swarm Optimizer. Evolutionary Computation. In Proceedings of the IEEE World Congress on Computational Intelligence, Anchorage, AK, USA, 4–9 May 1998; pp. 69–73.
37. Asali, J.C. Entrainment in Vertical Gas-Liquid Annular Flows. Ph.D. Thesis, University of Illinois at Urbana-Champaign, Ann Arbor, MI, USA, 1984.
38. Shearer, C.J.; Nedderman, R.M. Pressure gradient and liquid film thickness in co-current upwards flow of gas/liquid mixtures: Application to film-cooler design. *Chem. Eng. Sci.* **1965**, *20*, 671–683. [CrossRef]
39. Hewitt, G.F.; Roberts, D.N. *Studies of Two-Phase Flow Patterns by Simultaneous X-ray and Flast Photography*; Atomic Energy Research Establishment: Harwell, UK, 1969.
40. Zhong, Z.; Carr, T.R. Application of mixed kernels function (MKF) based support vector regression model (SVR) for CO_2–Reservoir oil minimum miscibility pressure prediction. *Fuel* **2016**, *184*, 590–603. [CrossRef]