

**COB-2023-1123**

# **PREDICTING FRICTION FACTORS IN TURBULENT FLOW OF HERSCHEL-BULKLEY FLUIDS: A RADIAL BASIS FUNCTION NEURAL NETWORK APPROACH**

**Gláucio K. Matoba**  
**Charlie van der Geest**  
**Marcelo S. Castro**

School of Mechanical Engineering, Universidade Estadual de Campinas (UNICAMP), Campinas-SP, Brazil  
g232009@dac.unicamp.br, charlievander@gmail.com, mscaastro@unicamp.br

**Daiane M. Iceri**

Center for Petroleum studies (CEPETRO), Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brazil, 13083-896.  
daiiceri@unicamp.br

**Roney L. Thompson**

Dep. of Mechanical Engineering, COPPE, Federal University of Rio de Janeiro, Centro de Tecnologia, Ilha do Fundão, 21945-970,  
Rio de Janeiro, RJ, Brazil  
rthompson@mecanica.coppe.ufrj.br

**Abstract.** *The accurate frictional pressure gradient is crucial for pipeline design systems in the industry. The prediction of frictional pressure gradient relies on factors such as pipeline diameter, mean velocity, fluid properties, and friction factor. However, predicting friction factors, especially for the turbulent flow of non-Newtonian fluids, remains challenging and prone to significant errors. In order to reduce these discrepancies, it is proposed an artificial neural network with a radial basis function (RBF) model to enhance friction factor prediction for non-Newtonian fluids with Herschel-Bulkley behavior. The RBF model was trained using a comprehensive database of 300 experimental data points. The model presented promising results, with an absolute error of approximately  $\pm 10\%$  compared to existing literature. This indicates the model's ability to provide reasonably accurate predictions for friction factors. Additionally, a statistical analysis was conducted to identify the most influential parameters of the friction factor, such as rheological parameters, Reynolds number, and pipe diameter. In summary, the proposed RBF model offers an effective approach to improve the estimation of friction factors for non-Newtonian fluids with Herschel-Bulkley behavior. It provides a valuable tool for pipeline system design, where accurate friction factor prediction is crucial. Moreover, the statistical analysis further enhances the understanding of the key parameters influencing friction behavior in non-Newtonian fluid flow.*

**Keywords:** *Herschel-Bulkley fluid, Friction factor, Turbulent flow, Radial Basis Function, Artificial Neural Network.*

## **1. INTRODUCTION**

The pressure gradient is essential in designing pipeline systems across various industries, including oil, food, and pharmaceutical sectors (Kumar *et al.*, 2020; Mishra *et al.*, 2021; Munir *et al.*, 2022). The pressure gradient is influenced by factors such as pipe diameter, mean velocity, friction factor, and fluid properties, all of which are crucial for calculating the pressure gradient along the pipeline.

Among these factors, fluid properties, especially non-Newtonian characteristics, significantly impact the pressure gradient. One important class of non-Newtonian fluids is viscoplastic fluids, which exhibit yield stress. A widely used model for viscoplastic fluids is the Herschel-Bulkley (HB) model (Herschel and Bulkley, 1926). In this model, a generalized Newtonian behavior is assumed, i.e.  $\tau = 2\eta D$ , where  $\eta$  is the viscosity of the fluid where parameters like the yield stress are inserted. A quite general viscoplastic model is given by:

$$\begin{cases} \tau = \tau_y + k\dot{\gamma}^n & \tau > \tau_y \\ \dot{\gamma} = 0 & \tau \leq \tau_y \end{cases}, \quad (1)$$

where  $\dot{\gamma}$  and  $\tau$  are scalar quantities that capture the intensity of the corresponding tensors,  $k$  and  $n$  are the consistency coefficient and the power-law index, respectively. Particular cases of the Herschel-Bulkley model are obtained when  $n = 1$ , which results in the Bingham constitutive equation, and when  $\tau_y = 0$ , where the power-law constitutive equation is recovered. In the case of  $\tau_y = 0$  and  $n = 1$ , the HB model reduces to the Newtonian constitutive equation.

The rheological characteristics of HB fluids present a considerable challenge in establishing reliable correlations for

friction factors and consequently affect pressure gradient prediction. Various studies in the literature have examined the friction factor, employing explicit and implicit analytical correlations, as well as machine learning-based approaches. Next, it is presented a brief literature review on the friction factor for both Newtonian and non-Newtonian fluids with HB behavior.

### 1.1 Friction factor for Newtonian fluids

The friction factor for a Newtonian fluid is well-established in the literature. The laminar regime is characterized by a clear relationship ( $f = 16/Re$ ), while in the turbulent regime, there are well-known references like Blasius (1913) and Colebrook (1939).

The analytical prediction of the friction factor for turbulent flow in Newtonian fluids has limitations, such as the imprecision of the Blasius equation and numerical resolution issues with the Colebrook correlation. As a result, alternative approaches using machine learning have been introduced. Researchers like Salmasi *et al.* (2012); Samadianfard *et al.* (2014) have utilized neural networks in their studies to address these challenges. It is worth mentioning that all of these authors relied on a database generated from the Moody diagram or Colebrook equation as a basis for their research.

Salmasi *et al.* (2012) developed a methodology utilizing an artificial neural network (ANN) and genetic algorithm, using the Colebrook equation as the basis for predicting the friction factor. The models provided explicit functions with up to 5 parameters and demonstrated average errors of 2.2% and 1.3% for the ANN and genetic algorithm, respectively.

Samadianfard *et al.* (2014) introduced three machine learning models: gene expression programming, adaptive neuro-fuzzy inference system, and artificial neural network (ANN). Alongside these models, they employed a data mining approach called M5 model tree. These models aimed to estimate the friction factor in a more accurate and straightforward manner compared to the Colebrook equation. Notably, among the proposed models, the ANN model achieved the best performance with an average error of 0.016%.

These results highlight the effectiveness of using artificial neural networks to predict the friction factor for Newtonian fluids, particularly considering the reliance on a comprehensive database generated from the Moody diagram or Colebrook equation. Such models provide reliable and user-friendly alternatives for friction factor estimation in practical engineering applications.

### 1.2 Friction factor for HB fluids

The friction factor estimation for non-Newtonian fluids presents challenges due to the influence of rheological properties. Simplified analytical correlations have shown promising results for laminar flow (Metzner and Reed, 1955; Peixinho *et al.*, 2005; Chhabra and Richardson, 2008). However, developing a generalized analytical correlation that accurately predicts the friction factor for HB fluids in turbulent flow remains an open issue.

In the literature, implicit and explicit correlations have been proposed for HB fluids in smooth pipes. The widely used implicit correlation by Dodge and Metzner (1959) can exhibit an error of up to 26.46% (Gul *et al.*, 2019). Explicit correlations, such as those by Virk (1975) and El-Emam *et al.* (2003), also show relatively high errors of 69.13% and 18.60%, respectively (Gul *et al.*, 2019). Modifications to the Colebrook (1939) correlation by Reed and Pilehvari (1993) for non-Newtonian fluids with yield stress still yield errors of up to 20% (Sorgun *et al.*, 2022). To address this, Sorgun *et al.* (2022) proposed an explicit correlation for turbulent friction factor in rough pipes, achieving satisfactory results with a mean absolute error of 6.07%.

Given the complexities of non-Newtonian fluids, machine learning techniques offer a promising alternative. Gul *et al.* (2019) proposed a methodology combining neural network and decision tree algorithms to estimate the friction factor for HB fluids. Their approach achieved a mean absolute error of 1.28% and an  $R^2$  value of 0.99, using a database constructed from their own experimental data. However, further validation and comparison with other databases are necessary for accurate predictions in complex fluid systems.

Based on these advancements, this paper proposes a methodology based on an artificial neural network (ANN) with radial basis functions (RBF) to predict the friction factor for turbulent flow in pipelines of non-Newtonian fluids with HB behavior. The approach utilizes a database specifically tailored for this study, extracted from the literature. The advantages of employing the RBF model within the neural network lie in its simplicity of implementation and ability to provide accurate predictions (Rooki and Rakhshkhorshid, 2017). This methodology holds great potential for accurately estimating the friction factor in turbulent flow, addressing the complexities associated with HB behavior.

## 2. METHODOLOGY

### 2.1 Friction factor

The friction factor is a crucial dimensionless parameter used in the pressure drop calculation. The Fanning friction factor for internal laminar flow in a pipe can be determined by the equation,  $f = 16/Re$ . This analytical and explicit

equation is applicable to both Newtonian and non-Newtonian fluids, as supported by studies conducted by Metzner and Reed (1955); Peixinho *et al.* (2005); Chhabra and Richardson (2008). However, the  $Re$  number plays a significant role between these two fluid types.

The friction factor for turbulent flow is more complex. Various empirical correlations have been proposed in the literature for estimating the friction factor in smooth and rough pipes. Notable examples include the works of Dodge and Metzner (1959); Torrance (1963); El-Emam *et al.* (2003); Gul *et al.* (2019) for smooth pipes, as well as Reed and Pilehvari (1993); Sorgun *et al.* (2022) for rough pipes. Among these correlations, the one proposed by Dodge and Metzner (1959) and modified by Reed and Pilehvari (1993) for Herschel-Bulkley fluids is given by:

$$\frac{1}{f^{\frac{1}{2}}} = -4.0 \log_{10} \left\{ 0.27 \left( \frac{\epsilon}{D} \right) + \frac{1.26 N^{(-1.2)}}{[N_{Re} f^{(1-\frac{N}{2})}]^{N(-0.75)}} \right\}, \quad (2)$$

where  $N_{Re}$  is the generalized  $Re$  number for Herschel-Bulkley fluid, given by,

$$N_{Re} = \frac{\rho v^{2-N} D^N}{k 8^{N-1}}, \quad (3)$$

and  $N$  is the generalized flow index behavior:

$$N = \frac{1}{A + B}, \quad (4)$$

$$A = \frac{(1 - 2n)\tau_w + 3n\tau_y}{n(\tau_w - \tau_y)} \quad \text{and} \quad B = \frac{2n(1 + n)[(1 + 2n)\tau_w^2 + n\tau_y\tau_w]}{n(1 + n)(1 + 2n)\tau_w^2 + 2n^2(1 + n)\tau_w\tau_y + 2n^3\tau_y^2}, \quad (5)$$

and the wall shear stress ( $\tau_w$ ) is given by:

$$\tau_w = \tau_y + k \left( \frac{8v}{D} \right)^n. \quad (6)$$

The friction factor correlation proposed by Reed and Pilehvari (1993) is widely used, but it has the disadvantage of requiring implicit calculations. In contrast, Sorgun *et al.* (2022) introduced a simpler and more user-friendly approach with an explicit friction factor correlation. This explicit correlation provides a straightforward method for calculating the friction factor, making it more convenient for practical use, given by:

$$f = \frac{0.06 N^{3.34}}{\left[ \log_{10} \left( \frac{\epsilon/D}{3.7} + 9.86 (N_{Re})^{-N} \right) \right]} \quad (7)$$

The existing methods based on correlations have limitations in accurately estimating the friction factor and, consequently, the frictional pressure drop. The equations proposed by Reed and Pilehvari (1993) and Sorgun *et al.* (2022) can have relative errors of up to 20%.

The methodology used in this work is based on a comprehensive database obtained from articles such as Vajargah *et al.* (2017); Johnson *et al.* (2018); Gul *et al.* (2019); Sorgun *et al.* (2022) for turbulent flow. The experimental Fanning friction factor is extracted from graphs of the pressure gradient ( $\Delta P/L$ ) versus  $Re$  number, by:

$$f = \frac{D}{2\rho v} \frac{\Delta P}{L} \quad (8)$$

where  $v$ ,  $D$ , and  $\rho$  are the mean velocity, internal diameter, and density, respectively. The  $Re$  number was obtained from the conditions and correlations provided by each author.

## 2.2 Data analysis

### 2.2.1 Radial Basis Function

The Radial Basis Function (RBF) is a regression model employed in machine learning. The implementation of this model in our study was based on the works of Penrose (1955); Broomhead and Lowe (1988); Haykin (1999); Rooki and Rakhshkhorshid (2017); Kumar *et al.* (2020). The RBF model is capable of predicting an output based on multiple input variables. In order to obtain accurate predictions, the model undergoes a training process where it learns from a given dataset, forming an artificial neural network (ANN). The dataset is typically divided into training and testing sets. The RBF layers are generated using the training data, and the predictions are validated using the test data.

The RBF model consists of three layers: the input layer (representing the input data), a hidden layer containing the radial basis function, and the output layer that provides the predicted data. The choice of the radial function for the hidden

layer depends on the specific problem. Common options include Gaussian, Multi-quadric, and Inverse Multi-quadric functions. During the study, an extensive exploration of these options was carried out, leading to the determination that the Inverse Multi-quadric function, represented by Eq. (9), provided the best fit for the data,

$$R_p = \frac{1}{\sqrt{1 + \sigma^2 \cdot x^2}} \quad (9)$$

where  $x = \|x_{data} - C_p\|$  is a Euclidean distance between the input data ( $x_{data}$ ) and a predetermined center point ( $C_p$ ) of the radial function. The  $\sigma$  is a shape factor, which allows an adjustment of the radial function. After that, the radial basis function is applied to the approximation function, given by,

$$F(x) = \sum_{p=1}^P a_p R_p + b \quad (10)$$

where  $a_p$  are the weights associated with the radial basis,  $R_p$  is the radial function,  $b$  is the bias, and  $P$  is the number of poles/radial functions applied in the model.

### 2.2.2 Database

The database used in this study was compiled from a scarce and limited number of experimental works found in the literature. Table 1 and 2 were constructed using data collected from multiple sources, including studies by Peixinho *et al.* (2005); Vajargah *et al.* (2017); Johnson *et al.* (2018); Gul *et al.* (2019); Sorgun *et al.* (2022). The selection criteria for these studies were based on studies involving flow inside circular pipelines with fluids exhibiting Herschel-Bulkley behavior. Specifically, only data points representing the transition and fully turbulent flow regimes *i.e.*, Reynolds numbers higher than 2300, were included in the analysis.

Table 1 provides general information about the database used to train the RBF model, including the range of rheological properties, diameter, and the quantity of training data for all tested conditions. However, it is important to note that this table only displays the minimum and maximum values of the experimental conditions for each paper. This database exclusively consists of experimental data points, no data was derived from author correlations.

With the exception of Peixinho *et al.* (2005), all articles involved the use of multiple concentrations of reagents (Xanthan gum or CMC), resulting in several fluids with different rheological properties. For each type of fluid ( $\tau_y, n, k$ ), tests were conducted across a range of flow rates to obtain values of pressure gradients. These data were then converted into Reynolds numbers and friction factors ( $f$ ) for subsequent model analysis. It is worth mentioning that not all authors published results with the same number of experimental points for each condition, ranging from 4 to 17 points.

Then, in order to separate the training and test data, a specific methodology was adopted. Only the conditions from each author that presented at least 15 experimental points were chosen, with 5 points designated for testing and the remaining data for training. This criterion resulted in a total of 300 training points and 20 test points. The conditions of the test database are depicted in Table 2. In the case of the experimental conditions described by Peixinho *et al.* (2005), the available data consisted of less than 15 experimental points, and therefore, all of these data points were considered only for the training process.

Table 1. Maximum and minimum values of experimental parameters for Herschel-Bulkley fluids of database training.

Ref	NN - Fluid	$\tau_y$ [Pa]	$k$ [Pa.s <sup><i>n</i></sup> ]	$n$	D [mm]	Re	Training data
Peixinho <i>et al.</i> (2005)	Carbopol	6.300	2.2	0.5	30	3753 - 5751	4
Vajargah <i>et al.</i> (2017)	XG <sup>(1)</sup>	0.792 - 3.244	0.032 - 0.560	0.468 - 0.961	7.74	2442 - 29246	42
Johnson <i>et al.</i> (2018)	XG <sup>(1)</sup>	0.170 - 5.850	0.110 - 1.380	0.46 - 0.68	7.74	2576 - 48488	70
Gul <i>et al.</i> (2019)	XG <sup>(1)</sup>	0.160 - 3.353	0.014 - 2.292	0.366 - 0.831	1.09	2323 - 52568	117
Sorgun <i>et al.</i> (2022)	CMC <sup>(2)</sup>	1.020 - 1.460	0.007 - 0.060	0.85 - 0.98	50, 80, 90	3246 - 85650	67

<sup>(1)</sup> Xanthan Gum, <sup>(2)</sup> Carboxymethyl Cellulose Solution

Furthermore, the database does not include fluid density due to the limited information available from the selected experimental data sources. However, the density variations between the fluids discussed in the articles are relatively insignificant. This can be attributed to the small concentrations of the additional reagents present in the fluids.

Additionally, the relative roughness parameter was not included in the database as it was not provided by the authors of the studies, with the exception of the study conducted by Sorgun *et al.* (2022), where relative roughness data was available. Nevertheless, it is important to acknowledge that these data alone were insufficient to conduct a comprehensive analysis of the entire dataset.

Even though the density and roughness data are not available in the current database, the proposed methodology remains valid. It is important to note that this approach is flexible and can be extended to include databases that take into account these factors in the future.

Table 2. Rheological parameters, diameter, and  $Re$  number range for Herschel-Bulkley fluids of the test database.

Ref	NN - Fluid	$\tau_y$ [Pa]	$k$ [Pa.s <sup><i>n</i></sup> ]	$n$	D [mm]	Re	Quantity of test data
Vajargah <i>et al.</i> (2017)	XG <sup>(1)</sup>	2.139	0.5992	0.4679	7.74	4053 - 27071	5
Johnson <i>et al.</i> (2018)	XG <sup>(1)</sup>	3.940	1.0300	0.4800	7.74	3900 - 23023	5
Gul <i>et al.</i> (2019)	XG <sup>(1)</sup>	3.228	1.5480	0.4020	7.74	4536 - 26660	5
Sorgun <i>et al.</i> (2022)	CMC <sup>(2)</sup>	1.460	0.0600	0.8500	80.00	6428 - 15873	5

<sup>(1)</sup> Xanthan Gum, <sup>(2)</sup> Carboxymethyl Cellulose Solution

### 2.2.3 Scaling the data

Once the database was divided into training and testing data, it became necessary to standardize the data. This step is crucial due to the disparate scales present in the dataset. For instance, the power-law index ranges from 0.37 to 0.98, while the Reynolds number spans from  $2.3 \times 10^3$  to  $3.0 \times 10^4$ . Consequently, it is essential to transform the data to a common scale to facilitate meaningful analysis without compromising information integrity or introducing distortions. In order to achieve this, the chosen transformation method was standardization, which can be expressed as follows:

$$x_{std} = \frac{x_{data} - \mu_{data}}{\sigma_{data}} \quad (11)$$

where  $x_{std}$  is the standardized variable,  $x_{data}$  is the variable to be standardized,  $\mu_{data}$  and  $\sigma_{data}$  are the database mean and standard deviation, respectively. By standardizing the entire database, the resulting transformed variables exhibit a mean of 0 and a standard deviation of 1. This standardization process not only ensures a consistent scale for analysis but also enhances the convergence of the RBF model.

### 2.2.4 Performance metrics

The performance of the model was assessed using several evaluation metrics, including the absolute error ( $AE$ ), mean absolute error ( $MAE$ ), and coefficient of determination ( $R^2$ ).

The percentage absolute error, given by Eq. (12), compares the measured friction factors against the predictions made by the RBF model. This metric provides an overall measure of the accuracy of the model's predictions.

$$AE = \frac{f_{mrd} - f_{prd}}{f_{mrd}} \cdot 100\% \quad (12)$$

where  $f_{mrd}$  denotes the measured friction factor from the test data, and  $f_{prd}$  represents the friction factor predicted by the model. The mean absolute error, calculated using Eq. (13), compares each individual reference value with the corresponding prediction from the fitted model. It provides an average measure of the magnitude of the errors.

$$MAE = \frac{100\%}{N} \cdot \sum_{i=1}^N \left| \frac{f_{mrd,i} - f_{prd,i}}{f_{mrd,i}} \right| \quad (13)$$

where  $N$  represents the number of data points. Lastly, the coefficient of determination, expressed by Eq. (14), is a statistical metric that assesses the regression performance of the model by comparing its predictions with the test data points. A higher  $R^2$  value indicates a better fit of the model to the test data.

$$R^2 = 1 - \frac{\sum_{i=1}^N (f_{mrd,i} - f_{prd,i})^2}{\sum_{i=1}^N (f_{mrd,i} - \bar{f}_{mrd,i})^2} \quad (14)$$

An optimal condition is achieved when the mean absolute error ( $MAE$ ) is equal to 0, indicating a perfect match between the predicted and measured values. Similarly, a coefficient of determination ( $R^2$ ) equal to 1 signifies that the model has fully converged and accurately captures the relationship between the predicted and measured values. These evaluation metrics collectively provide insights into the accuracy and performance of the RBF model in predicting the friction factors for the given test data.

### 2.2.5 Variance inflation factor

The variance inflation factor (VIF) is a statistical method and it was used to quantify the degree of multicollinearity between the model's input variables. Multicollinearity can be described as a situation where there is a strong correlation or interdependence between two or more independent variables in a regression model. It means that the predictor variables are not completely independent and can influence each other, which can add challenges in accurately determining the

unique contribution of each variable to the dependent variable. The VIF for a particular predictor variable is calculated by regressing that variable against all the other predictor variables in the model. The VIF value indicates how much the variance of the estimated coefficient for that variable is increased due to multicollinearity. That is, high VIF values indicate a higher level of multicollinearity. A mathematical definition for multicollinearity is demonstrated by Gunst (1984); Öztürk and Akdeniz (2000). The VIF calculation is given by (Neter *et al.*, 1983),

$$VIF_{var} = \frac{1}{1 - R_{var}^2} \quad (15)$$

where  $R^2$  is the coefficient of determination of the variable,  $var$ . A minimum VIF value of 1 means the absence of multicollinearity for a specific variable, indicating no correlation with other variables. Higher VIF values indicate increasing levels of multicollinearity, implying a higher correlation between the variable and others in the model.

A high VIF value means that the variable does not significantly influence the model response. However, it is not possible to isolate the variable's impact due to its correlation with other variables. It set the VIF threshold value higher than 5, as recommended by Thompson *et al.* (2017).

Conversely, a low VIF value suggests that the variable has a lower degree of multicollinearity and may have a direct influence on the model's response. VIF values were calculated for each input variable, including  $\tau_y$ ,  $n$ ,  $k$ ,  $D$ , and  $Re$ . In order to facilitate the interpretation of the VIF analysis, it was converted these values into percentages, representing their relative importance. It is worth noting that the VIF value and relative importance are inversely proportional. In other words, higher VIF values indicate the lower relative importance of the variable in the model.

### 3. RESULTS AND DISCUSSION

This section presents the results obtained from the described methodology. It is divided into two parts: a comparison between experimental friction factors from the literature, and a weighted analysis of the effects of rheological parameters, diameter, and Reynolds number on the friction factor. This last analysis helps to identify the influential factors and enhances our understanding of their impacts on friction behavior.

#### 3.1 Friction factor analysis

Figure 1 illustrates the comparison of friction factors between the RBF model and the test data from Table 2. The markers correspond to experimental data from the relevant articles, while the lines represent the predictions of the RBF model adjusted with the corresponding rheological data and diameter from each study.

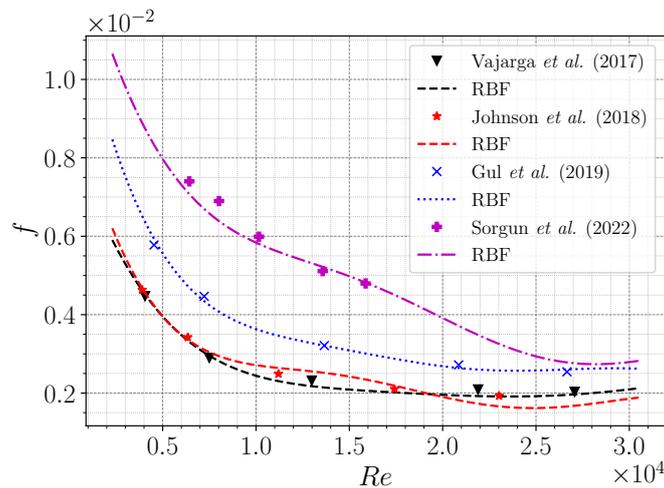


Figure 1. Friction factor vs  $Re$  for turbulent flow

The RBF model demonstrated accurate predictions and captured the expected curve behavior for most of the test data, except for the results reported in Johnson *et al.* (2018). In the case of Johnson *et al.* (2018), the model exhibited an unexpected oscillation between Reynolds numbers  $1 \times 10^4$  and  $3 \times 10^4$ . This oscillation contradicts the physical behavior of a continuous reduction in the friction factor as the Reynolds number increases, especially considering that a fully turbulent flow typically results in a constant friction factor. However, it is important to note that overall, the RBF model effectively fits the test data, resulting in tolerable errors that will be discussed in detail later. The observed consistency in the majority of cases highlights the model's ability to capture the behavior of turbulent flow.

Furthermore, the RBF model effectively captured the behavior of the data reported in Vajargah *et al.* (2017); Gul

*et al.* (2019). The model's curve closely approximated the test points and exhibited a consistent trend similar to the Colebrook (1939) function, where the friction factor remained constant as the Reynolds number increased. The RBF function also provided a robust fit for the data reported in Sorgun *et al.* (2022). However, it's crucial to note that the RBF model was initially trained using limited data from Sorgun *et al.* (2022), specifically for a Reynolds number range up to  $Re = 1.9 \times 10^4$ , with the fluid properties shown in Table 2. When the model was extrapolated to  $Re = 3.0 \times 10^4$ , the curve exhibited an abrupt reduction. This decrease can be attributed to the overfitting of the RBF model to the training data. To mitigate overfitting and enhance the model's reliability, additional experimental data is necessary to validate the RBF function's estimates and extend its capabilities for further improvement through higher-quality training.

Nevertheless, within the experimental Reynolds number range where training data available is wide, the RBF function demonstrated a superior fit to the experimental data when compared to the correlation presented in Eq. (7). The mean absolute error (*MAE*) for the RBF model was 3.50%, while the *MAE* for the correlation proposed by Sorgun *et al.* (2022) was approximately 6.22%. Consequently, the RBF model provides a significant advantage by offering a more precise estimation of the friction factor when the tested data falls within the established limits.

A 10-fold cross-validation test was conducted to assess overfitting and examine the consistency of *MAE* and  $R^2$  across the tests. In this 10-fold test, the mean absolute error averaged was 6.58%, with an  $R^2$  of 0.95. The coefficient of determination ranged from a minimum of 0.88 to a maximum of 0.98. These results indicate a slight degree of overfitting in the model, which aligns with the discrepancies arising from the data extrapolation by Sorgun *et al.* (2022).

### 3.1.1 Error analysis

The error analysis was conducted using the mean absolute error (*MAE*) and the coefficient of determination ( $R^2$ ), as presented in Table 3. This table compares the model predictions with the test data from each study. The articles with the highest coefficient of determination were Vajargah *et al.* (2017) and Gul *et al.* (2019), achieving  $R^2$  values of 0.98 and 0.99, respectively. This result aligns with the visual observation of the graph in Figure 1, where a clear overlap between the data and the RBF curve can be observed. Moreover, the mean absolute error for Gul *et al.* (2019) was the lowest among the authors, at 2.14%. The RBF function, adjusted by the inverse Multi-quadric function, effectively captured the training data and provided accurate predictions for the friction factor.

On the other hand, Johnson *et al.* (2018) and Sorgun *et al.* (2022) exhibited the lowest values of  $R^2$ , which is evident from the graph in Figure 1. The behavior of the RBF function with the Johnson *et al.* (2018) data showed oscillations, which showed unexpected behavior. Consequently, it resulted in a low coefficient of determination and a higher mean absolute error. The unsatisfactory coefficient of determination for Sorgun *et al.* (2022) can be attributed to the limited availability of data. Even though, within the training range, the model exhibited a good fit to the data, as indicated by an *MAE* value of 3.50%.

Table 3. MAE and  $R^2$  for friction factor predictions by the RBF model in each study

Flow regime	Ref	MAE [%]	$R^2$
Turbulent	Vajargah <i>et al.</i> (2017)	4.08	0.98
	Johnson <i>et al.</i> (2018)	5.38	0.97
	Gul <i>et al.</i> (2019)	2.14	0.99
	Sorgun <i>et al.</i> (2022)	3.50	0.92

Figure 2 presents a comparison between the measured friction factor values from Table 2 and the corresponding predicted values obtained from the RBF model. It can be observed that 95% of points fall within an absolute error range of  $\pm 10\%$ , with only one value with an error of 14%. This outcome demonstrates that when provided with experimental data, the RBF model is capable of generating friction factor estimates with an error below 10%. It is important to note that the RBF model attempts to fit the available data obtained from the literature, emphasizing the need for a substantial amount of data including various flow conditions and rheological properties to enhance its performance.

### 3.2 Variables influences on friction factor

The variables' influence on the friction factor prediction was analyzed using the VIF method. The VIF results for each variable are presented in Table 4. This method enables the assessment of the importance or influence of the following variables on the friction factor (*f*): Reynolds ( $Re$ ) number, diameter ( $D$ ), consistency coefficient ( $k$ ), power-law index ( $n$ ), and yield stress ( $\tau_y$ ).

In order to aid in the interpretation of the VIF analysis, a bar graph was created, as shown in Figure 3. This figure illustrates the relative importance of each variable in influencing the turbulent friction factor.

The power-law index demonstrates the highest VIF value, reaching 5.3, indicating a strong correlation with the other variables (multicollinearity). Consequently, it has a relatively low influence on the friction factor. Specifically, the relative

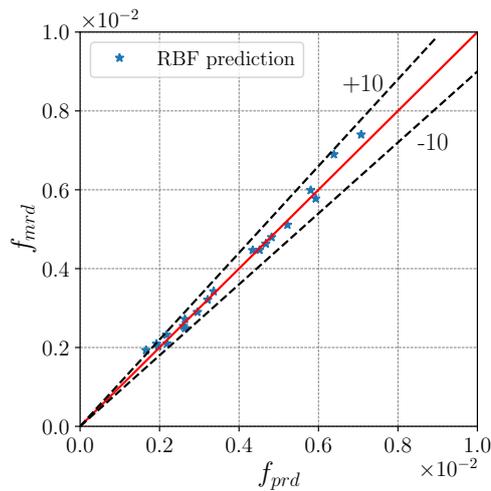


Figure 2. Friction factor comparison - Measured (Data Test) vs Predicted (RBF model).

Table 4. Reponse of VIF analysis

Variable	VIF
$\tau_y$	4.46
$n$	5.34
$k$	3.50
$D$	2.43
$Re$	2.63

importance of the power-law index on the friction factor is found to be  $n = 12.60\%$ . This statistical analysis reveals that the power-law index is influenced by the other input variables. In other words, its relative importance is distributed among the subsequent input variables, resulting in a diminished direct impact on the friction factor. A similar pattern is observed for the yield stress ( $\tau_y$ ) and consistency coefficient ( $k$ ), with relative importance values of  $\tau_y = 15.05\%$  and  $k = 19.20\%$ , respectively.

In contrast, the diameter ( $D$ ) exhibits the lowest VIF value of 2.43, indicating a low level of correlation with the other variables (low multicollinearity). As a result, it exerts a significant influence on the friction factor, signifying high importance. Specifically, the relative importance of the diameter in relation to the friction factor is determined to be  $D = 27.63\%$ . A similar trend is observed for the Reynolds number, which demonstrates a relative importance of 25.52%.

The statistical analysis provides compelling evidence highlighting the substantial role of rheological properties in influencing the turbulent friction factor, accounting for a significant 46.85% of relative importance. The remaining 53.15% importance is attributed to the Reynolds number and diameter, aligning well with the analytical friction factor prediction.

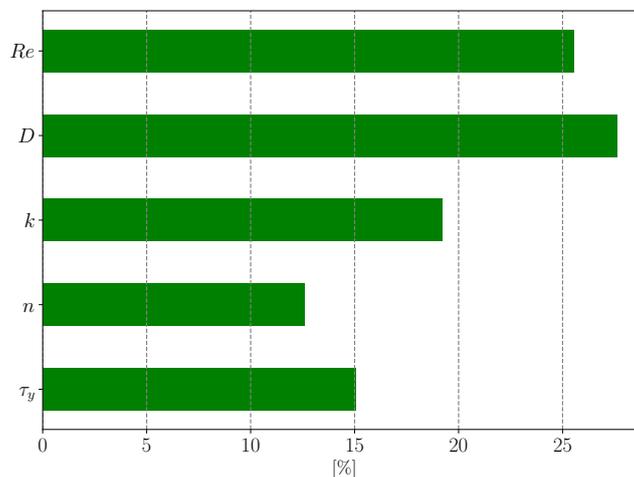


Figure 3. Relative variable importance on friction factor prediction.

Under these conditions, increasing the flow mean velocity induces increased velocity gradients near the wall, leading to heightened wall shear stress.

According to Eq. (6), the wall shear stress is directly influenced by rheological properties of the fluid ( $n$ ,  $k$ , and  $\tau_y$ ), as well as the diameter ( $D$ ) and flow velocity ( $v \propto Re$ ). Then, the wall shear stress correlates with the generalized  $Re$  number and influences on friction factor prediction.

Moreover, the Reynolds number prediction relies on the rheological properties of the fluid and the wall shear stress ( $Re = 8\rho v^2/\tau_w$ ). Consequently, all these variables exert an influence on the friction factor, and statistical analysis enables a quantitative assessment of their respective importance.

The VIF analysis was employed to quantify the relevance of variables that affect the friction factor in the turbulent flow regime. The statistical results lead to the conclusive inference that the rheological parameters, as well as the diameter and Reynolds number, are of fundamental importance in estimating the friction factor in a turbulent flow. In fact, all variables play a significant role in determining the friction factor, as demonstrated by the equations proposed by renowned researchers such as Dodge and Metzner (1959); Torrance (1963); Reed and Pilehvari (1993); El-Emam *et al.* (2003); Gul *et al.* (2019); Sorgun *et al.* (2022).

#### 4. CONCLUSION

In this study, a machine learning approach using a radial basis function (RBF) artificial neural network was employed to enhance the prediction of friction factors in turbulent regimes for Herschel-Bulkley fluids. The RBF model demonstrated its capability by achieving a minimum mean absolute error of 2.14% when compared with the data from Gul *et al.* (2019), and a maximum error of 5.38% with the data from Johnson *et al.* (2018). Moreover, the model exhibited an overall absolute error within  $\pm 10\%$  across the entire test database. In order to decrease the error, a further model improvement is adding a monotonic decrease behavior in the RBF function. This does not allow an oscillatory behavior of friction factor prediction as seen in the results from Johnson *et al.* (2018) and Sorgun *et al.* (2022). Nevertheless, the absolute error highlights the accuracy of the RBF model in predicting friction factors for the given conditions. One of the advantages of the RBF model is its ease of implementation, allowing for straightforward application in practical scenarios.

However, it is important to note that the RBF model's performance is limited to the available database and may not be extrapolatable to conditions beyond the training range. Besides, the model assumes that the data are combined based on the same diameter and fluid properties, making it applicable only when the combination matches those present in the database. These limitations emphasize the need for a comprehensive and diverse database encompassing various flow conditions and fluid properties to further enhance the model's accuracy and applicability. Therefore, accurate predictions cannot be provided for diameters not covered by the available data or when combining the properties of different authors.

In addition to the RBF model, a variance inflation factor (VIF) statistical approach was employed to assess the significance of input variables on the response variable, *i.e.*, the friction factor. This statistical analysis provides insights into the relative importance of different variables influencing friction behavior, aiding in the interpretation and understanding of the underlying mechanisms.

It is crucial to emphasize that this study serves as a valuable tool/methodology that can be further validated and enhanced with a larger and more diverse database. By evaluating the importance of parameters on friction factor, this approach can be utilized to predict pressure gradient and determine the optimal experimental matrix setup. Overall, the utilization of the RBF model and the VIF statistical model contributes to advancing its ability to accurately predict friction factors and gain valuable insights into the behavior of Herschel-Bulkley fluids in turbulent flow regimes.

#### 5. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of TotalEnergies EP Brazil and ANP. The results presented in this document were obtained through a Research, Development and Innovation (R,D&I) project carried out in partnership with the company TotalEnergies EP Brazil and financed with resources from the R,D&I Clause regulated by the National Agency of Petroleum, Gas and Biofuels - ANP. Acknowledgements are also extended to the Center for Energy and Petroleum Studies (CEPETRO) and the School of Mechanical Engineering (FEM), both at the University of Campinas (UNICAMP - Brazil), and to the Alberto Luiz Coimbra Institute (COPPE-UFRJ). Finally, we acknowledge ALFA (Artificial Lft and Flow Assurance) Research Group for the collaborative efforts to achieve the objectives of this publication.

#### 6. REFERENCES

- Blasius, H., 1913. "Das aehnlichkeitsgesetz bei reibungsvorgangen in flussigkeiten". In *Mitteilungen über Forschungsarbeiten auf dem Gebiete des Ingenieurwesens*, Springer Berlin Heidelberg, pp. 1–41.
- Broomhead, D.S. and Lowe, D., 1988. "Multivariable functional interpolation and adaptive networks". *Complex Syst.*, Vol. 2.
- Chhabra, R. and Richardson, J., 2008. *Non-Newtonian Flow and Applied Rheology: Engineering Applications*. Elsevier

Science.

- Colebrook, C.F., 1939. "Turbulent flow in pipes, with particular reference to the transition region between the smooth and rough pipe laws". *Journal of the Institution of Civil Engineers*, Vol. 11, No. 4, pp. 133–156.
- Dodge, D.W. and Metzner, A.B., 1959. "Turbulent flow of non-newtonian systems". *AIChE Journal*, Vol. 5, No. 2, pp. 189–204.
- El-Emam, N., Kamel, A., El-Shafei, M. and El-Batrawy, A., 2003. "New equation calculates friction factor for turbulent flow on non-newtonian fluids". *Oil & gas journal*, Vol. 101, No. 36, pp. 74–74.
- Gul, S., Johnson, M.D., Vajargah, A.K., Ma, Z., Hoxha, B.B. and van Oort, E., 2019. "A data driven approach to predict frictional pressure losses in polymer-based fluids". In *Day 1 Tue, March 05, 2019*. SPE.
- Gunst, R.F., 1984. "Comment: Toward a balanced assessment of collinearity diagnostics". *The American Statistician*, Vol. 38, No. 2, p. 79.
- Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*. International edition. Prentice Hall. ISBN 9780139083853.
- Herschel, W.H. and Bulkley, R., 1926. "Konsistenzmessungen von gummi-benzollösungen". *Kolloid-Zeitschrift*, Vol. 39, No. 4, pp. 291–300.
- Johnson, M., Gul, S., Vajargah, A.K., Hoxha, B. and van Oort, E., 2018. "Real-time friction factor monitoring: characterization of drag reduction in polymer-based fluids". In *AADE Fluids Conference and Exhibition*.
- Kumar, A., Ridha, S., Ganet, T., Vasant, P. and Ilyas, S.U., 2020. "Machine learning methods for herschel-bulkley fluids in annulus: Pressure drop predictions and algorithm performance evaluation". *Applied Sciences*, Vol. 10, No. 7, p. 2588.
- Metzner, A.B. and Reed, J.C., 1955. "Flow of non-newtonian fluids—correlation of the laminar, transition, and turbulent-flow regions". *AIChE Journal*, Vol. 1, No. 4, pp. 434–440.
- Mishra, K., Grob, L., Kohler, L., Zimmermann, S., Gstöhl, S., Fischer, P. and Windhab, E.J., 2021. "Entrance flow of unfoamed and foamed herschel-bulkley fluids". *Journal of Rheology*, Vol. 65, No. 6, pp. 1155–1168.
- Munir, I.D., Jaafar, N.A. and Shafie, S., 2022. "Analysis of unsteady solute dispersion in a blood flow of herschel-bulkley through a catheterized stenosed artery". *Malaysian Journal of Fundamental and Applied Sciences*, Vol. 18, No. 4, pp. 430–447.
- Neter, J., Wasserman, W. and Kutner, M., 1983. *Applied Linear Regression Models*. R.D. Irwin. ISBN 9780256025477.
- Peixinho, J., Nouar, C., Desaubry, C. and Théron, B., 2005. "Laminar transitional and turbulent flow of yield stress fluid in a pipe". *Journal of Non-Newtonian Fluid Mechanics*, Vol. 128, No. 2-3, pp. 172–184.
- Penrose, R., 1955. "A generalized inverse for matrices". *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 51, No. 3, p. 406–413.
- Reed, T. and Pilehvari, A., 1993. "A new model for laminar, transitional, and turbulent flow of drilling muds". In *SPE Production Operations Symposium*. OnePetro.
- Rooki, R. and Rakhshkhorshid, M., 2017. "Cuttings transport modeling in underbalanced oil drilling operation using radial basis neural network". *Egyptian Journal of Petroleum*, Vol. 26, No. 2, pp. 541–546.
- Salmasi, F., Khatibi, R. and Ghorbani, M.A., 2012. "A study of friction factor formulation in pipes using artificial intelligence techniques and explicit equations". *Turkish Journal of Engineering and Environmental Sciences*, Vol. 36, No. 2, pp. 121–138.
- Samadianfard, S., Sattari, M.T., Kisi, O. and Kazemi, H., 2014. "Determining flow friction factor in irrigation pipes using data mining and artificial intelligence approaches". *Applied Artificial Intelligence*, Vol. 28, No. 8, pp. 793–813.
- Sorgun, M., Muftuoglu, T.D. and Gucuyener, I.H., 2022. "Friction factor estimation for turbulent flow of herschel-bulkley and power law fluids in pipes". *Journal of Petroleum Science and Engineering*, Vol. 211, p. 110044. ISSN 0920-4105.
- Thompson, C.G., Kim, R.S., Aloe, A.M. and Becker, B.J., 2017. "Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results". *Basic and Applied Social Psychology*, Vol. 39, No. 2, pp. 81–90.
- Torrance, B., 1963. "Friction factors for turbulent non-newtonian fluid flow in circular pipes". *S. Afr. Mech. Eng*, Vol. 13, pp. 89–91.
- Vajargah, A.K., Sullivan, G., Johnson, M. and van Oort, E., 2017. "Transitional and turbulent flow of drilling fluids in pipes : An experimental investigation".
- Virk, P.S., 1975. "Drag reduction fundamentals". *AIChE Journal*, Vol. 21, No. 4, pp. 625–656.
- Öztürk, F. and Akdeniz, F., 2000. "Ill-conditioning and multicollinearity". *Linear Algebra and its Applications*, Vol. 321, No. 1-3, pp. 295–305.

## 7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.