Revista Facultad de Ingeniería
ISSN: 0121-1129
ISSN: 2357-5328
Universidad Pedagógica y Tecnológica de Colombia

García-Ubaque, Cesar-Augusto; Ladino-Moreno, Edgar-Orlando; García-Vaca, María-Camila

Determination of the Inside Diameter of Pressure Pipes for Drinking Water Systems Using Artificial Neural Networks

Revista Facultad de Ingeniería, vol. 31, no. 59, e204, 2022, January-March
Universidad Pedagógica y Tecnológica de Colombia

DOI: https://doi.org/10.19053/01211129.v31.n59.2022.14037

Revista Facultad de Ingeniería

Journal Homepage: https://revistas.uptc.edu.co/index.php/ingenieria

Determination of the Inside Diameter of Pressure Pipes for Drinking Water Systems Using Artificial Neural Networks

Cesar-Augusto García-Ubaque[1]

Edgar-Orlando Ladino-Moreno[2]

María-Camila García-Vaca[3]

Received: January 16, 2022 **Accepted:** March 18, 2022 **Published:** March 25, 2022

Citation:

C.-A. García-Ubaque, E.-O. Ladino-Moreno, M.-C. García-Vaca, "Determination of the Inside Diameter of Pressure Pipes for Drinking Water Systems Using Artificial Neural Networks," *Revista Facultad de Ingeniería*, vol. 31 (59), e14037, 2022. https://doi.org/10.19053/01211129.v31.n59.2022.14037

[1] Ph. D. Universidad Distrital Francisco José de Caldas (Bogotá-Distrito Capital, Colombia). cagarciau@udistrital.edu.co. ORCID: 0000-0002-6959-6610

[2] M. Sc. Universidad Distrital Francisco José de Caldas (Bogotá-Distrito Capital, Colombia). eoladinom@correo.udistrital.edu.co. ORCID: 0000-0002-7770-452X

[3] M. Sc. Universidad Católica de Colombia (Bogotá-Distrito Capital, Colombia). mcgarciav@ucatolica.edu.co. ORCID: 0000-0002-1382-7387

Abstract

The fifth-degree polynomial equation determines the diameter in pressurized drinking water systems. The input variables are Q: flow (m^3/s), H: pressure drop (m); L: pipe length (m); ε: roughness (m), ϑ: kinematic viscosity (m^2/s), and Σk: sum of minor loss coefficients (dimensionless). After applying the energy equation for a hydraulic system composed of two tanks connected to a pipe of constant diameter and accepting the Colebrook-White and the Darcy-Weisbach equations, an undetermined expression is obtained since more unknowns than equations are established. This problem is solved by implementing a nested loop for the coefficient of friction and the diameter. This article proposes an Artificial Neural Network (ANN) implementing the Levenberg-Marquardt backpropagation method to estimate the diameter from the log-sigmoidal transfer function under stationary flow conditions. The training signals set consists of 5,000 random data that follow a normal distribution, calculated in Visual Basic (®Excel). The statistics used for the network evaluation correspond to the mean square error, the regression analysis, and the cross-entropy function. The architecture with the best performance had a hidden layer with 25 neurons (6-25-1) presenting an MSE equal to 5.41E-6 and 9.98E+00 for the Pearson Correlation Coefficient. The cross-validation of the neural scheme was carried out from 1,000 independent input signals from the training set, obtaining an MSE equal to 6.91E-6. The proposed neural network calculates the diameter with a relative error equal to 0.01% concerning the values obtained with ®Epanet, evidencing the generalizability of the optimized system.

Keywords: Artificial Neural Network; Colebrook-White; Darcy-Weisbach; Levenberg-Marquardt; pipeline hydraulics.

Determinación del diámetro interior de tuberías a presión para sistemas de agua potable utilizando redes neuronales artificiales

Resumen

El diámetro en sistemas a presión de agua potable es posible determinarlo mediante una ecuación polinómica de quinto grado. Como variables de entrada se tiene: Q: caudal (m^3/s), H: pérdida de carga (m); L: longitud de la tubería (m); ε: rugosidad (m), ϑ: viscosidad cinemática (m^2/s) y Σk: sumatoria de coeficientes de pérdidas menores (adimensional). Aplicado la ecuación de la energía para un sistema hidráulico compuesto por dos tanques conectados con una tubería de diámetro constante y aceptando la ecuación de Colebrook-White y la ecuación de Darcy-Weisbach se obtiene una expresión subdeterminada debido a que se establecen más incógnitas que ecuaciones. Este problema se soluciona implementando un *bucle* anidado para el coeficiente de fricción y el diámetro. Este artículo propone una Red Neuronal Artificial (RNA) implementando el método de Retropropagación Levenberg-Marquardt para estimar el diámetro a partir de la función de transferencia log-sigmoidal, esto bajo condiciones estacionarias de flujo. El conjunto de las señales de entrenamiento está conformado por 5,000 datos aleatorios que siguen una distribución normal, calculados en Visual Basic (®Excel). Los estadísticos utilizados para la evaluación de la red corresponden al error medio cuadrático, el análisis de regresión y la función de entropía cruzada. La arquitectura que demostró un mejor redimento correspondió a una capa oculta con 25 neuronas (6-25-1) presentando un MSE igual a 5.41E-6 y 9.98E+00 para el Coeficiente de Correlación de Pearson. La validación cruzada del esquema neuronal se realizó a partir de 1,000 señales de entrada independientes del conjunto de entrenamiento obteniendo MSE igual 6.91E-6. La red neuronal propuesta calcula el diámetro con un error relativo igual a 0.01% con respecto a los valores obtenidos a partir de ®Epanet, evidenciando la capacidad de generalización del sistema optimizado.

Palabras clave: Colebrook-White; Darcy-Weisbach; hidráulica de tuberías; Levenberg-Marquardt; red neuronal artificial.

Determinação do diâmetro interno de tubulações de pressão para sistemas de água potável usando redes neurais artificiais

Resumo

O diâmetro em sistemas de água potável pressurizada pode ser determinado por meio de uma equação polinomial de quinto grau. Como variáveis de entrada temos: Q: vazão (m3/s), H: perda de carga (m); L: comprimento do tubo (m); ε: rugosidade (m), ϑ: viscosidade cinemática (m2/s) e Σk: soma dos coeficientes de perdas menores (adimensional). Aplicando a equação de energia para um sistema hidráulico composto por dois tanques conectados por uma tubulação de diâmetro constante e aceitando a equação de Colebrook-White e a equação de Darcy-Weisbach, obtém-se uma expressão subdeterminada, pois se estabelecem mais incógnitas do que equações. Este problema é resolvido implementando um loop aninhado para o coeficiente de atrito e o diâmetro. Este artigo propõe uma Rede Neural Artificial (RNA) implementando o método Backpropagation de Levenberg-Marquardt para estimar o diâmetro a partir da função de transferência log-sigmoidal, isto sob condições de fluxo permanente. O conjunto de sinais de treinamento é composto por 5.000 dados aleatórios que seguem uma distribuição normal, calculados em Visual Basic (®Excel). As estatísticas utilizadas para a avaliação da rede correspondem ao erro quadrático médio, à análise de regressão e à função de entropia cruzada. A arquitetura que apresentou melhor rendimento correspondeu a uma camada oculta com 25 neurônios (6-25-1) apresentando um MSE igual a 5,41E-6 e 9,98E+00 para o Coeficiente de Correlação de Pearson. A validação cruzada do esquema neural foi realizada a partir de 1.000 sinais de entrada independentes do conjunto de treinamento, obtendo-se MSE igual a 6,91E-6. A rede neural proposta calcula o diâmetro com um erro relativo igual a 0,01% em relação aos valores obtidos do ®Epanet, mostrando a capacidade de generalização do sistema otimizado.

Palavras-chave: Colebrook-White; Darcy-Weisbach; Levenberg-Marquardt; rede neural artificial; tubulação hidráulica.

I. Introduction

In the design of hydraulic systems, the calculation of the diameter is fundamental. This parameter determines the behavior of the pressure along the pipe. Likewise, the average flow velocity will remain constant because there is no variation in the pipe cross-section. This effect causes the system to have a single value for the friction coefficient and Reynolds number. The design equation for single pipe diameters (9) is obtained by accepting the governing equation for pressure-flow (2) and setting the flow velocity in terms of the flow rate. This fifth-degree equation presents two unknowns, the diameter and the friction coefficient. Therefore, a nested loop must be established to solve both variables simultaneously. For the diameter, it is suggested starting from a seed value equal to 0.254 m, and for the friction coefficient, the recommended seed value is 0.015. These values accelerate the convergence processes of the numerical method used. Once the corresponding diameter is found, it should be approximated to the upper commercial diameter. Consequently, if a single pipe is considered, which connects two reservoirs (Figure 1) and applies the Energy Equation (1) on the surface of the reservoirs, it is obtained.

$$z_1 + \frac{P_1}{\gamma} + \frac{V_1^2}{2g} = z_2 + \frac{P_2}{\gamma} + \frac{V_2^2}{2g} + H \qquad (1)$$

For a constant incompressible one-dimensional flow, the energy per unit weight (Nm/N) generated by the hydraulic system (Figure 1) is defined by the Energy Equation (1); where tank 1 has: z_1: position head (m), $\frac{P_1}{\gamma}$: pressure head (m), $\frac{V_1^2}{2g}$: velocity head (m), H: head loss (m). The head loss H is determined by the losses caused by the friction between the pipe and the fluid; the losses generated by the fittings. The Darcy-Weisbach equation (2) rules the head loss.

$$h_f = f\frac{L}{D}\frac{V^2}{2g} \qquad (2)$$

Where, h_f: head loss (m); f: friction coefficient; L: piping length (m); D: diameter (m); V: average flow velocity (m/s); g: gravity (m/s^2). For local or minor losses, we have,

$$h_l = \sum k\frac{V^2}{2g} \qquad (3)$$

Where, h_l: minor losses (m); k: minor loss coefficient (dimensionless); V: average flow velocity (m/s); g: gravity (m/s^2). The principle of conservation of mass for a given

DOI: https://doi.org/10.19053/01211129.v30.n56.2021.14037

control volume, where the flow has an incompressible performance and there is no variation of the discharge as a function of time and space (steady-state), is determined by the principle of continuity from the following expression:

$$Q = \int_A v dA = V_1 A_1 = V_2 A_2 \quad (4)$$

Where, Q: Discharge (m^3/s); V: Average flow velocity (m/s); A: Cross-sectional area of the pipe (m^2).

Fig. 1. Simple piping scheme.

Thus, from the energy equation (1) and establishing the velocity in terms of the streamflow, we have:

$$z_1 - z_2 - f\frac{L}{D}\frac{V^2}{2g} - \sum k\frac{V^2}{2g} = 0 \quad (5)$$

$$z_1 - z_2 = f\frac{L}{A^2 D}\frac{Q^2}{2g} + \Sigma k\frac{Q^2}{A^2 2g} \quad (6)$$

$$z_1 - z_2 = f\frac{LQ^2}{2g0.25^2\pi^2 D^5} + \Sigma k\frac{Q^2}{2g0.25^2\pi^2 D^4} \quad (7)$$

$$z_1 - z_2 = H \quad (8)$$

Design equation:

$$f(D) = 12.1026HD^5 - \Sigma kQ^2 D - fLQ^2 = 0 \quad (9)$$

First derivative:

$$f'(D) = 60.513HD^4 - \Sigma kQ^2 \quad (10)$$

Newton-Raphson method:

$$D_{n+1} = D_n - \frac{f(D_n)}{f'(D_n)} \quad (11)$$

$$D_{n+1} = D_n - \frac{12.1026HD^5 - \Sigma kQ^2 D - fLQ^2}{60.513HD^4 - \Sigma kQ^2} \quad (12)$$

The nested loop development (Figure 2) determines the solution of the equation (9). For the single pipes diameter calculation, 75% of the total head loss of the system is proposed as a seed value for the friction losses [1]. For this study, the input signals correspond to Q: discharge (m^3/s), H: head loss (m), L: pipe length (m), ε: pipe roughness (m), ϑ: kinematic viscosity (m^2/s), and Σk: summation of minor loss coefficients (dimensionless). The output signal is determined by the diameter (m). The equation proposed by Colebrook-White (13) was implemented to calculate the friction coefficient (f).

$$\frac{1}{\sqrt{f}} + 2log\left[\frac{\varepsilon/D}{3.7} + \frac{2.51}{R_e\sqrt{f}}\right] = 0 \qquad (13)$$

Where *f*: coefficient of friction (dimensionless factor), ε: absolute roughness (m), D: diameter (m), Re: Reynolds number (dimensionless factor). Similarly, the Newton-Raphson method for the calculation of the friction coefficient is given by:

$$f_{n+1} = f_n - \frac{\frac{1}{\sqrt{f_n}} + 2log\left[\frac{\varepsilon/D}{3.7} + \frac{2.51}{R_e\sqrt{f_n}}\right]}{-\frac{1}{2}f_n^{-1.5} + \frac{2\left[\frac{-2.51}{2R_e}f_n^{-1.5}\right]\log(e)}{\left[\frac{\varepsilon/D}{3.7} + \frac{2.51}{R_e\sqrt{f_n}}\right]}} \qquad (14)$$

If two tanks connected by a pipe section are considered (Figure 1), and the total head loss and the discharge conveyed by the pipe are known, it is possible to determine the value of the internal diameter from equation (9). Likewise, [2] used the Fixed-Point iteration method to calculate the diameter in pressurized piping systems. Figure 2 presents the flowchart for the nested loop with two unknowns and six input signals; the process achieves convergence with an approximation equal to 1E-12.

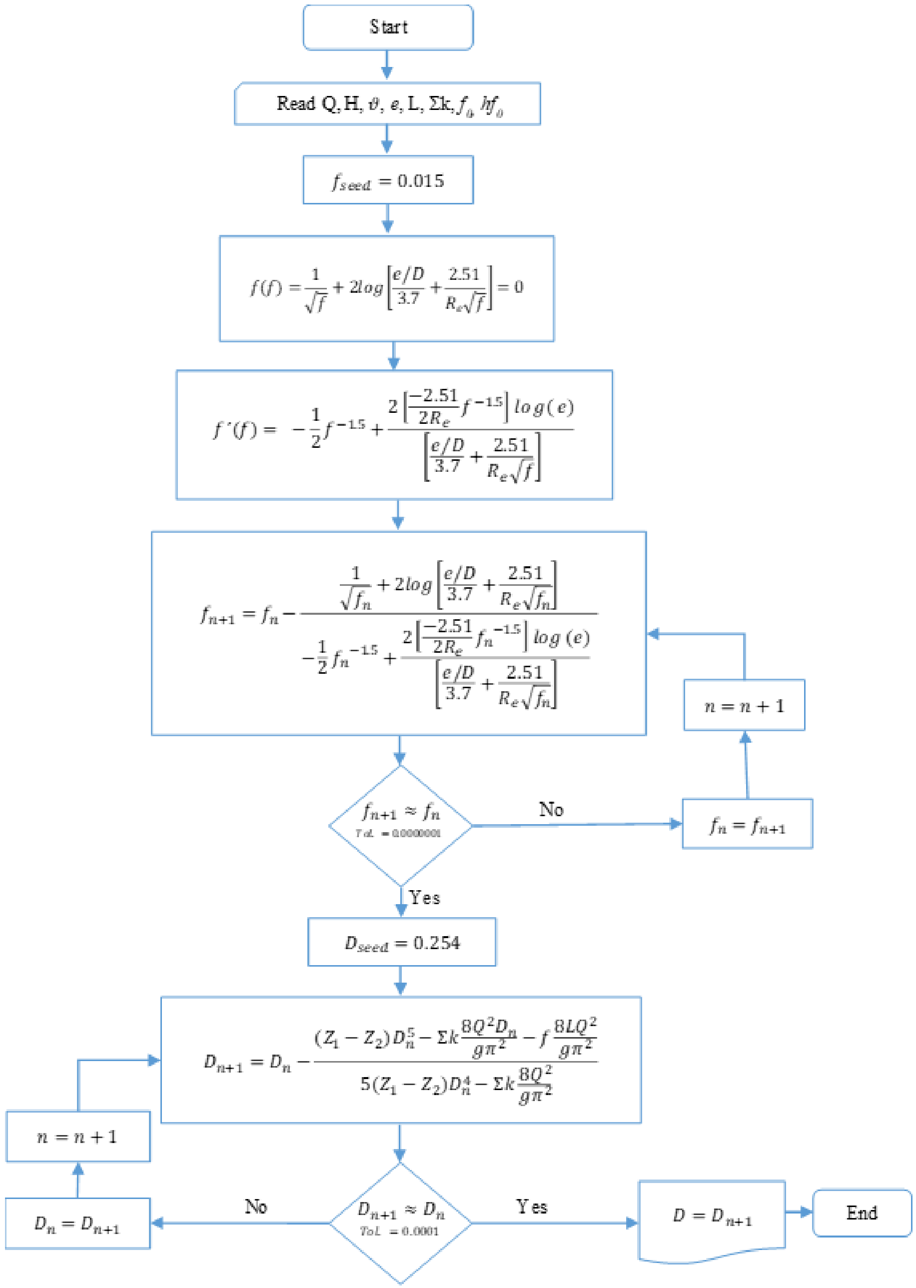

Fig. 2. Flow chart for diameter design [3].

II. Neuronal Structure

A. Data Processing

Using Visual Basic (®Excel), a routine is structured to calculate the diameter (D) from the input signals (Q, H, L, ε, ϑ, Σk). The code is created from equations (12) and (14). The iteration shown in Table 2 was repeated 5,000 times from random data

DOI: https://doi.org/10.19053/01211129.v30.n56.2021.14037

fitting a normal distribution, establishing the input-output matrix for this study. These data are available at the following link (*download data*). The Domain of the input variables varies from minimum to maximum values; Table 1 shows the ranges established.

Table 1. Input variable ranges (inputs).

	Discharge	**Head Loss**	**Length**	**Roughness**	**Kinematic viscosity**	**Coef. accessory**
	Q (m^3/s)	H (m)	L (m)	ε (m)	Vis (m^2/s)	Σk
Minimum	0.000096	10	100	0.0000015	0.000000661	0
Maximum	0.475	50	500	0.00045	0.000001519	10

The following code made in Visual Basic (®Excel) generates an iterative loop for the diameter and the friction coefficient. The seed values directly affect the convergence process. In this sense, 0.015 is the seed value for the friction coefficient and 0.254 m for the diameter. If the seed values are far from the solution value, there is a probability that the algorithm will diverge. Consequently, the proposed seed values guarantee the convergence of the iterative method. The code outputs 200 iterations for each diameter value and friction coefficient with an approximation of 1E-12 for the objective function. The loop stops when it does not detect a numerical value in the following grid cell to be iterated.

```
Dim i As Integer
Sub MacroD_f()
For i = 1 To 10
  Range("h11").Select
    Do Until ActiveCell = ""
      ActiveCell.Offset(0, 1).GoalSeek Goal:=0, ChangingCell:=ActiveCell
      ActiveCell.Offset(1, 0).Range("A1").Select
 Loop
    Range("o11").Select
      Do Until ActiveCell = ""
      ActiveCell.Offset(0, 1).GoalSeek Goal:=0, ChangingCell:=ActiveCell
      ActiveCell.Offset(1, 0).Range("A1").Select
 Loop Next i
End Sub
```

Table 2 shows the results generated for the diameter calculation from the nested loop. In order to explain the velocity as a discharge function and consider the input data for head loss, pipe length, discharge, and the seed friction coefficient, a fifth-degree equation is established. The initial equation solution determines the value of the cross-sectional pipe area. Consequently, it is possible to calculate the flow velocity. Once the speed is calculated, the Reynolds number and the estimated friction coefficient are obtained. Then, the loop is generated until convergence of the objective function is reached, both for the diameter and the friction coefficient simultaneously. Usually, this convergence is reached in the fourth iteration with an approximation equal to 1E-12. The method used to obtain the training signals for this study corresponds to the numerical approach proposed by Newton-Raphson (12), (14).

Table 2. Initial iteration of input signals.

f	Q (m^3/s)	H (m)	L (m)	Σk	D (m)	V (m/s)	υ (m^2/s)	Re	ε (m)	f´
1.500E-02	3.190E-01	5.0008E+01	1.887E+02	8E+00	2.40E-01	7.04E+00	1.416E-06	1.194E+06	3.24E-04	2.135E-02
2.135E-02	3.190E-01	5.0008E+01	1.887E+02	8E+00	2.52E-01	6.39E+00	1.416E-06	1.138E+06	3.24E-04	2.111E-02
2.111E-02	3.190E-01	5.0008E+01	1.887E+02	8E+00	2.51E-01	6.41E+00	1.416E-06	1.140E+06	3.24E-04	2.112E-02
2.112E-02	3.190E-01	5.0008E+01	1.887E+02	8E+00	2.51E-01	6.41E+00	1.416E-06	1.140E+06	3.24E-04	2.112E-02

B. Data Scale

We, as authors, chose to perform the neural model with the actual data without scaling. Table 3 establishes that the statistical criteria are more favorable for the unscaled data than the scaled data from the logarithm in base 10.

Table 3. Comparison of scaled data vs. actual data.

Scale	No. of hidden layers	No. of neurons per layer	Architecture	R^2	MAE	MSE	SSE	SAE	BCE
Logarithmic	1	25	6-25-1	9.98E+00	1.57E-03	1.79E-05	8.15E-02	7.1E+00	3.6E-01
Without scale	1	25	6-25-1	9.99E+00	9.71E-04	5.41E-6	2.71E-2	4.8E+00	1.2E+00

C. Artificial Neural Network (ANN) Architecture

Due to the nonlinearity of the functions that estimate the diameter (12) and the friction coefficient (14), it is feasible to estimate the output parameter through optimization algorithms. Artificial neural networks can approximate any continuous

DOI: https://doi.org/10.19053/01211129.v30.n56.2021.14037

nonlinear function independent of the function degree [4]. The implementation of artificial intelligence techniques contributes to finding the minimum error surface generated by the cost function. A neural structure created in Matlab (2021a) is proposed (Figure 3) corresponding to 6 input signals (Q, H, L, ε, ϑ, Σk), a hidden layer with 25 neurons, and an output signal for diameter estimation. The logsig transfer function showed optimal results in terms of MSE and computational time required to reach convergence. The Levenberg-Marquardt (trainlm) method is suited to both the training data and the test data sets. The network weights are iteratively adjusted from the error estimate [5] [6] describes the application of the Levenberg-Marquardt in neural network systems for training. This algorithm has demonstrated a higher training speed for the neural network [7]. Similarly, [8] proposes a neural architecture of 5 input variables, a hidden layer with 36 neurons, and 10 output parameters to classify the optimal commercial diameter for the hydraulic system.

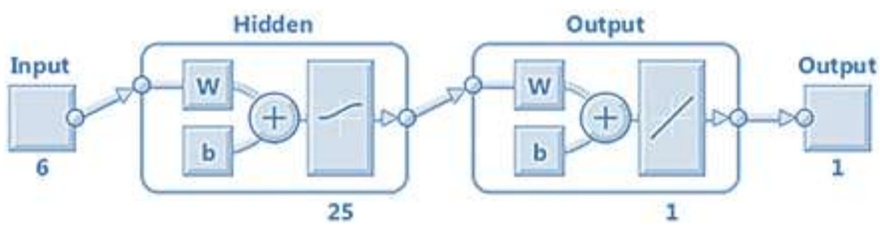

Fig. 3. Schematic diagram of neuronal architecture (6-25-1).

Table 4 indicates different neural structures tested for this study to obtain the lowest value for the MSE; this value was achieved for the architecture (6-25-1) with an MSE equal to 5.41E-6. It was found that increasing the number of hidden layers does not guarantee a decrease in the MSE, and as a consequence, it does increase the computational cost of the iterative process. The computational time for the (6-25-25-25-1) scheme was 3 hours and 25 minutes. In contrast, the time required for the (6-25-25-1) scheme was 15 minutes. Thus, neural models with several hidden layers tend to overfit so that the model can predict the training data. However, for the prediction of independent data, the overfitted neural model has shortcomings. The most important property of a neural network is its ability to generalize and model new input data [9].

Table 4. Network topology for diameter calculation [logsig, trainlm].

No. of hidden layers	No. of neurons per layer	Architecture	R^2	MAE	MSE	SSE	SAE	BCE
1	2	6-2-1	9.84E+0	6.16E-03	7.64E-5	3.82E-1	30.3E+00	15 E+00
1	10	6-10-1	9.99E+0	1.67E-03	9.74E-6	4.87E-2	4.95E-02	4.63E-01
1	20	6-20-1	9.99E+0	1.28E-03	7.68E-6	3.84E-2	6.41E+00	4.47E-01
1	25	6-25-1	9.99E+0	9.71E-04	5.41E-6	2.71E-2	4.84E+00	1.25E+00
1	30	6-30-1	9.99E+0	9.96E-04	5.65E-6	2.82E-2	4.97E+00	8.54E-01
2	10	6-10-10-1	9.99E+0	9.11E-04	5.85E-6	2.93E-2	4.55E+00	1.48E-01
2	30	6-30-30-1	9.99E+0	1.08E-04	6.39E-6	3.19E-2	5.39E+00	2.25E-01
3	25	6-25-25-25-1	9.99E+0	1.09E-03	6.63E-6	3.31E-2	5.46E+00	2.17E-01

D. Neural Network Training

The Levenberg-Marquardt training function (trainlm) uses the second derivatives of the cost function upgrading the convergence times [10]. The implementation of this algorithm is feasible as long as the second derivative of the neural network weights exists. Thus, the input signals are affected by the random weights and biases values (16). Once this value is obtained, the log-sigmoid activation function (logsig) (15) is implemented. This function implements the Jacobian matrix for the calculations. This matrix is formed by the first-order partial derivatives of the function. The performance for this function is measured through the MSE. The algorithm presents good performance for escaping moderate local minima and oscillation problems. The interval [0,1] determines the range of the corresponding function.

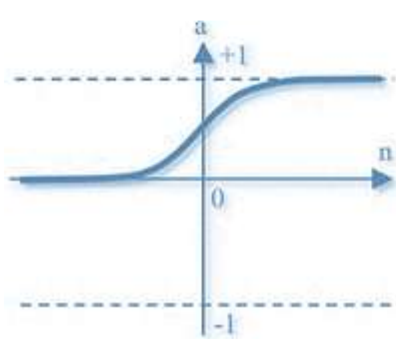

Fig. 4. Log sigmoid transfer function.

The following expression defines the log-sigmoid activation function (logsig):

$$a = \frac{1}{1+e^{-n}} \quad (15)$$

$$n = wp + b \quad (16)$$

Six neural architectures are proposed for the diameter in pressure piping systems calculation. The lowest Pearson Correlation Coefficient was obtained for the scheme (6-2-1) with an R equal to 0.99282; the computational time required for this scheme was 4 minutes. Figure 5a presents the dispersion of the outputs for ANN (6-2-1). However, for the arrangement (6-25-1), represented in Figure 5d, an R equal to 0.99939 indicates the best training and output signals fit; the computational time required was approximately 10 minutes. The scheme (6-30-30-1) presented a similar performance to the system (6-25-1). Nevertheless, this scheme composed of two hidden layers considerably increased the computational time required to reach convergence, approximately 3 hours and 25 minutes.

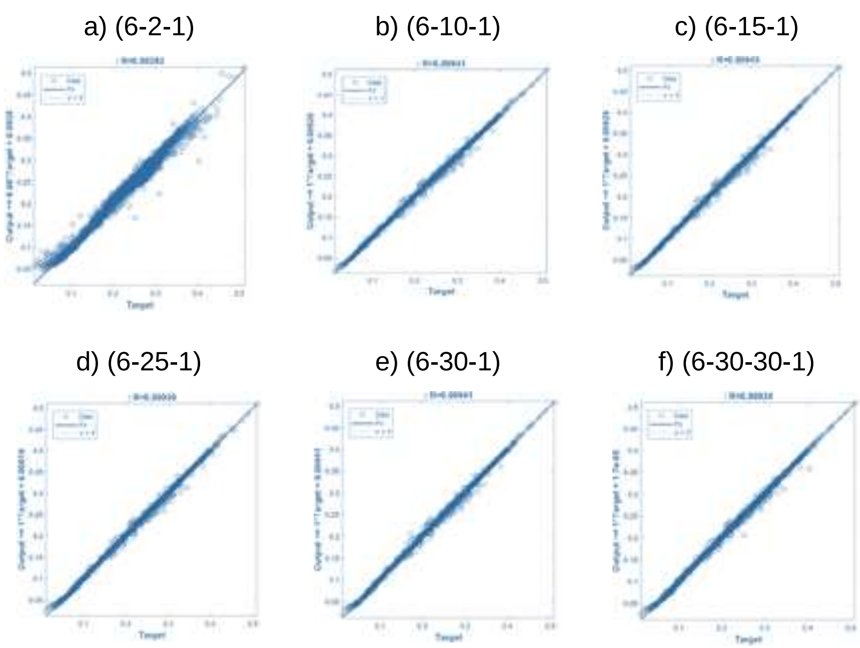

Fig. 5. ANN architectures.

Figure 6 shows the 3D irregular surface structured from the weights, the bias parameter, and the sum of squared errors. The objective of the cost function is to establish the minimum for this surface.

DOI: https://doi.org/10.19053/01211129.v30.n56.2021.14037

6-25-1 6-25-1 6-25-1

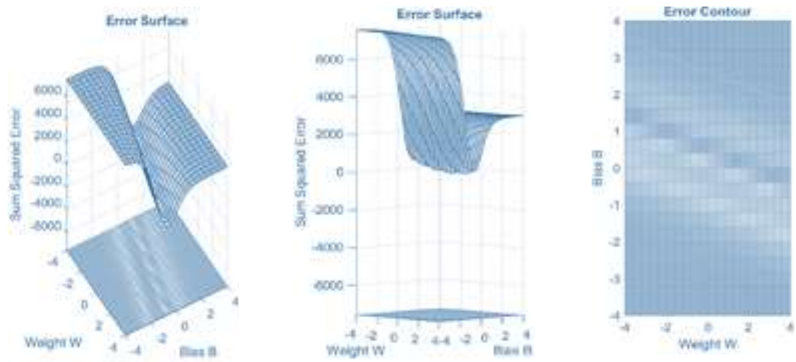

Fig. 6. Error surface for ANN (6-25-1).

Figure 7a indicates the performance of the validation and training curve; the error decreases as the number of epochs of the iterative process increases. If the MSE of the validation curve starts to increase and distance itself from the training curve, the model will overfit, which directly affects the generalization capability of the network. Figure 7b presents the adaptation parameter *mu* used in the Levenberg-Marquardt optimization process. For this study, the *mu* parameter corresponds to 1E-08 for season 68.

(a) Seasons: 68, (6-25-1) (b) Seasons: 68, (6-25-1) (c) Seasons: 68, (6-25-1)

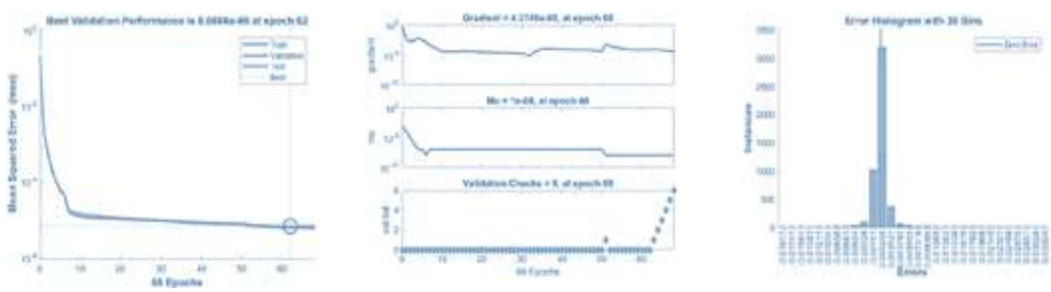

Fig. 7. Training status (6-25-1).

The data were divided into three sets using random indexes, training, validation, and test, as represented in Figure 8. The straight line at 45° represents a perfect fit, i.e., the values of the estimated outputs are equal to the target values. Figure 8c indicates the relationship between outcomes and targets for the test set. This study found a linear relationship between the output signals and the objectives with a Pearson Correlation Coefficient equal to 0.99946 for a neural scheme (6-25-1).

DOI: https://doi.org/10.19053/01211129.v30.n56.2021.14037

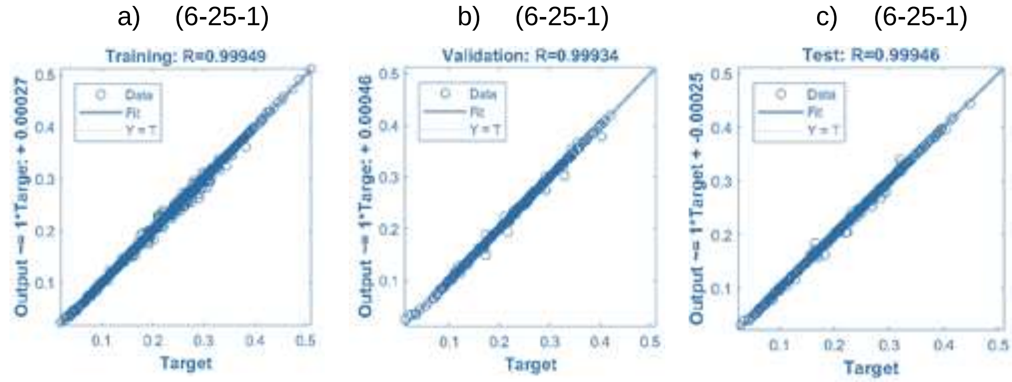

Fig. 8. Training regression (6-25-1).

E. Validation

For the analysis of the network output signal and to validate the independent data set, six statistics were considered: R (Pearson correlation coefficient), MAE (Mean Absolute Error), MSE (Mean Squared Errors), SSE (Sum of Squared Errors), SAE (Sum of Absolute Errors), BCE (Binary Cross Entropy). In order to establish a representative spectrum associated with the independent data set, a set of 1,000 random data was structured for the test signals. Cross-validation allows comparing the estimated signals with the target values from the independent data of the training set. The error histogram presented in Figure 7c indicates a minimum standard deviation with a mean tending to zero for the individual data; the histogram tends to be symmetric around the average. The percentage difference of the MSE between the training data and the independent data corresponds to 0.00015%, indicating the generalization capability of the neural model.

Table 5. Cross-validation

Data	No. of data	Architecture	R^2	MAE	MSE	SSE	SAE	BCE
Training	5,000	6-25-1	9.99E+00	9.71E-04	5.41E-6	2.71E-2	4.84E+00	1.25E+00
Independent	1,000	6-25-1	9.98E+00	1.27E-03	6.91E-06	6.91E-03	1.27E+00	0.52E-01

The U.S. Environmental Protection Agency developed ®Epanet for the calculation of pressurized piping systems. This system works with hydraulic simulation periods from the drinking water distribution systems. It also models water quality within a

DOI: https://doi.org/10.19053/01211129.v30.n56.2021.14037

pressurized network and can be used for any non-compressible fluid flowing under pressure analysis. ®Epanet determines the flow rates through the pipes and the values of the pressures at the nodes based on the principle of conservation of mass and energy by implementing the gradient method proposed by Todini and Pilati (1987). For the calculation of friction losses, three models are presented: Darcy-Weisbach (D-W), Hazen-Williams (H-W), Chézy-Manning (C-M).

Similarly, the neural network was evaluated from the data shown in Table 6. The results obtained were compared with the values calculated in ®Epanet, ®Excel, and the application for the calculation of pressure pipes of the hydraulics online website, accepting the Darcy Weisbach equation (2). For the first data in Table 6, the modeling was performed in ®MatLab software. Two tanks were established (Figure 1), the first tank with a height above the water surface equal to 36.712 m, a fitting at the inlet with a coefficient of 1, the main pipe with a diameter of 0.2428 m, a length equal to 104.31 m, a loss coefficient per fitting equal to 1 at the outlet, a tank at the end with zero height, the kinematic viscosity corresponds to 0.000001404 m2/s, and the roughness of the pipe equals to 0.0002574 m. The model in ®Epanet obtained a flow rate equal to 0.38109 m3/s, with a velocity of 8.23 m/s. The results obtained in ®Epanet, ®Excel, and the website validate the values calculated by ANN (6-25-1). According to the hydraulic calculation performed in ®Epanet, the velocity remains constant because the system does not present a variation of the pipe cross-section. This performance causes the Reynolds number and the friction coefficient to remain constant along the length of the pipe.

Table 6. ANN validation, Epanet, Excel, and web page.

Input							Output			
#	Q (m^3/s)	H (m)	L (m)	*e* (m)	υ (m^2/s)	Σk	Diameter (m)			
							RNA (6-25-2)	®Epanet	®Excel	Página Web (www.edgarladino.com)
1	3.81E-01	3.67E+01	1.04E+02	2.57E-04	1.40E-06	2E+00	2.42E-01	2.42E-01	2.42E-01	2.42E-01
2	4.6588E-01	3.19E+01	1.05E+02	1.46E-04	1.39E-06	3E+00	2.72E-01	2.71E-01	2.71E-01	2.71E-01
3	1.618E-01	1.31E+01	1.33E+02	4.74E-05	1.24E-06	4E+00	2.17E-01	2.16E-01	2.16E-01	2.16E-01
4	2.392E-01	1.10E+01	1.01E+02	1.83E-04	1.13E-06	2E+00	2.52E-01	2.52E-01	2.52E-01	2.52E-01
5	3.535E-01	3.03E+01	1.55E+02	2.35E-04	7.76E-07	9E+00	2.85E-01	2.84E-01	2.84E-01	2.84E-01

In addition, the generalization capability of the neural model is evidenced by the estimation of 50 independent data. These data show no dependence on the training, test, and validation data set. Figure 10 shows the actual data (circles) obtained from equations (12), (14), and the ANN estimated data (crosses). The Pearson Correlation Coefficient obtained was equal to 9.9742E-1 showing a linear relationship between the input signals and the estimated signals. The root average squared error for the 50 data corresponded to 3.65E-05. The cross-entropy obtained was equal to 5.504E-1, which establishes a lower uncertainty for the probability distribution.

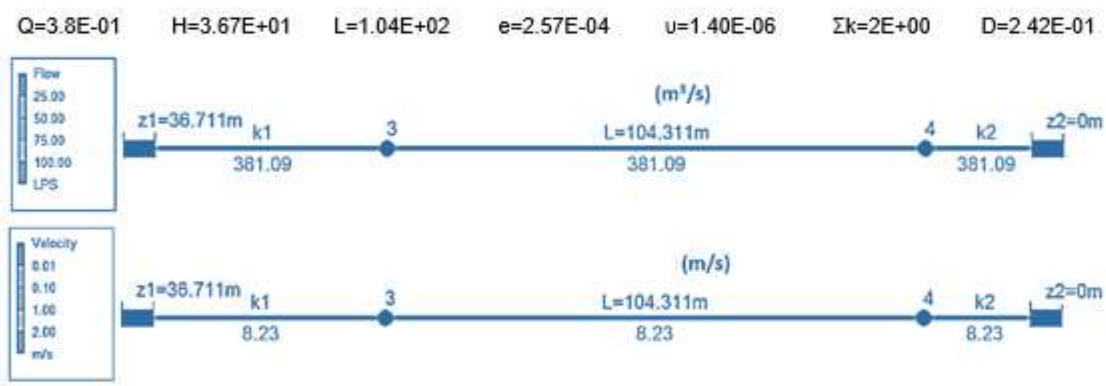

Fig. 9. Discharge – speed.

Finally, Figure 10 shows the generalization capability of the neural network (6-25-1) to estimate the theoretical diameter in pressurized pipes. This study demonstrated the potential of artificial neural networks to solve nonlinear systems.

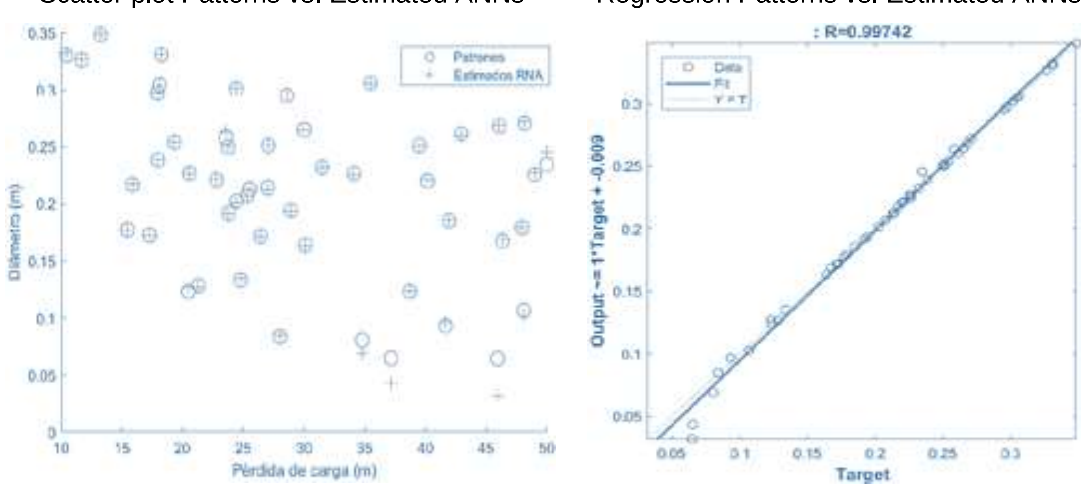

Fig. 10. Diameter – Head Loss. Patterns – Estimated ANNs.

III. Conclusions

The best performing architecture corresponded to a hidden layer with 25 neurons (6-25-1), presenting an MSE equal to 5.41E-6 and 9.98E+00 for the Pearson Correlation Coefficient. The cross-validation of the neural scheme was performed from 1,000 independent input signals of the training set, obtaining an MSE equal to 6.91E-6. This validation demonstrated the generalization capability of the proposed neural arrangement for the theoretical diameter in pressurized piping systems estimation.

It was found that increasing the number of hidden layers does not guarantee a decrease in the MSE and increases the computational cost of the iterative process. Similarly, increasing the number of hidden layers and neurons can generate overfitting of the input signals, limiting the model's capacity in the generalization process. Finally, this study demonstrated the potential of artificial neural networks to solve nonlinear systems.

Authors' Contribution

César Augusto García-Ubaque: Research, Supervision, Methodology, Validation, Writing - Review and editing.

Edgar Orlando Ladino-Moreno: Conceptualization, Data Curation, Formal Analysis, Research, Writing - Review and editing.

María Camila García-Vaca: Research, Supervision, Methodology, Validation, Writing - Review and editing.

References

[1] E. Ladino, C. García, M. García, "La implicancia económica mediante Newton Rapshon para el desarrollo de un aplicación Android para el diseño del diámetro de tuberías a presión," *Aglala,* vol. 11, no. 1, pp. 149-168, 2020

[2] N. Zaragoza, J. Baeza, "Determinación del diámetro de sistemas de tuberías mediante la utilización del Visual Basic para Aplicaciones y el Método de Aproximación de Punto Fijo," *Ingeniería,* vol. 7, no. 2, pp. 55-64, 2003

[3] E. Ladino, C. García, M. García, "Darcy-Weisbach resistance coefficient determination using Newton-Raphson approach for android 4.0," *Tecnura,* vol. 23, no. 60, p. 52–58, 2019. https://doi.org/10.1016/j.egypro.2016.11.077

[4] Z. Rao, F. Alvaruiz, "Use of an artificial neural network to capture the domain knowledge of a conventional hydraulic simulation model," *Q IWA Publishing Journal of Hydroinformatics,* vol. 1, no. 9, e14, 2007. https://doi.org/10.2166/hydro.2006.014

[5] Y. Najjar, *Quick manual for the use of ANN program TR-SEQ1*, Manhattan: Department of Civil Engineering, Kansas State University, 1999

[6] M. T. Hagan, M. Menhaj, "Training feedforward networks with the Marquadt algorithm," *IEEE Transactions on Neural Networks,* vol. 5, no. 6, pp. 989-993, 1994. https://doi.org/10.1109/72.329697

[7] K. Kipli, M. Mohd, S. Wan Masra, N. Zamhari, K. Lias, D. Awang , "Performance of Levenberg-Marquardt Backpropagation for Full Reference Hybrid Image," in *Proceeding of the international multiconference of engineers and computer scientists,* 2012, pp. 20-25

[8] J. Dawidowicz, "A Method for Estimating the Diameter of Water Pipes Using Artificial Neural Networks of the Multilayer Perceptron Type," *Technologies and Applications,* vol. 4, no. 1, pp. 26-30, 2018

[9] J. Dawidowicz, "Evaluation of a pressure head and pressure zones in water distribution systems by artificial neural networks," *Neural Computing and Applications,* vol. 30, pp. 2531–2538, 2018. https://doi.org/10.1007/s00521-017-2844-8

[10] A. Markopoulos, S. Georgiopoulos, D. Manolakos, "On the use of back propagation and radial basis function neural networks in surface roughness prediction," *Journal of Industrial Engineering International,* vol. 12, pp. 389–400, 2016

Nomenclature

Symbol	Unit	Parameter
D	m	Diameter
ε	m	Roughness
f	Dimensionless factor	Coefficient of friction
g	m/s^2	Gravity
H	m	Total head loss
h_f	m	Friction loss
hl	m	Accessory pressure drop
J	m/m	Unit loss
k	Dimensionless factor	Loss coefficient per fitting
L	m	Length
LGH	m	Hydraulic gradient line
LE	Nm/N	Energy line
m.c.a	m	Water column meter
P	m	Pressure
P_0	m	Atmospheric pressure
Re	Dimensionless factor	Reynolds number
V	m/s	Velocity
z	m	Height
γ	N/m^3	Specific gravity
[illegible]	m^2/s	Kinematic viscosity

Abbreviations

BCE	Binary Cross-Entropy
MAE	Mean Absolute Error
MSE	Mean Square Error
EL	Energy Line
HGL	Hydraulic Gradient Line
R	Pearson correlation coefficient
ANN	Artificial Neural Network
ASE	Absolute Sum of Errors
SSE	Sum of Squared Errors
Logsig	Log-sigmoid activation function

Appendix A. Coding of the ANN model for Matlab (Version 2021a).

```
close all; clear all; clc; format long
%============================================================
% Optimización | MatLab| Edgar O. Ladino M.| César A. García U. | Ingeniería Civil
% Universidad Distrital Francisco José de Caldas
% Facultad Tecnológica
% Bogotá | Colombia
%============================================================

% ====== Cálculo de diámetro en tuberías a presión =======
%================== Artificial neural networks ==================
%____________________________________________________________
% 1. Importar dataset 2,000 datos
d = csvread('DataSet_D_5000_RNA.csv'); %Archivo plano (csv)

%____________________________________________________________
% 2. Definición matriz p (Inputs); Vector t (Outputs)
p = d(:,1:6)'; %Matriz p: Transpuesta de la columna 1, 2, 3, 4, 5 y 6 de la matriz d
t = d(:,7)'; %Matriz t: Transpuesta de la columna 7 de la matriz d

%____________________________________________________________
% 3. Importar dataset matriz de prueba (Test) 1,000 datos
test = csvread('DataSet_D_RNA_Test_H35.csv'); %Archivo plano (csv)

%____________________________________________________________
% 4. Definición matriz test_p (Inputs) 5,000 datos
test_p = test(:,1:6)'; %Matriz p: Transpuesta de la columna 1 y 2 de la matriz test
test_pD = test(:,7)'; %Matriz p: Transpuesta de la columna 7 de la matriz test
test_pH = test(:,2)'; %Matriz p: Transpuesta de la columna 2 de la matriz test
test_pf = test(:,7)'; %Matriz p: Transpuesta de la columna 7 de la matriz test

%____________________________________________________________
% 5. Arquitectura de la red neuronal
net = fitnet(25); %feedforwardnet; patternnet; network; # neuronas
net.layers{1}.transferFcn = 'logsig'; %logsig; hardlim, tansig, purelin
net.performFcn= 'mse'; %mse; crossentropy; mae; msereg
net.trainFcn = 'trainlm'; %Entrenamiento: trainlm backpropagation; trainbr; traingd;
trainrp...
net.divideFcn = 'dividerand'; %División: dividerand; divideblock; divideint; divideind
net.trainParam.epochs = 6000; %Controla el número de epocas
[net, tr] = train(net,p,t); %Entrenamiento de la red
view(net) %Grafica esquema de la red
y = net(p); %Función de la red
y_test = net(test_p); %Función de la red test 2,100 datos
mse_test=1/1000*(test_pf-y_test).^2;
classes = vec2ind(y);
```

```
%_____________________________________________________________
% 6. Gráficas rendimiento de la red
figure;
plottrainstate(tr)
figure;
plotperform(tr)
figure;
plotregression(t,y)
figure;
plotregression(test_pf,y_test)

%_____________________________________________________________
% 7. Estimación del error
e = t-y; %y (entrenamiento) - y (Estimado)
figure;
ploterrhist(e,'bins',30) %Histograma de errores
R = corrcoef(t,y) %Coeficientes de correlación
MAE = mae(e) %mae: Error absoluto medio
MSE = immse(t,y) %Error medio cuadrático
SSE = sse(net,t,y,1) %sse: Error de suma cuadrada
SAE = sae(net,t,y) %Suma absoluta de errores
BCE = crossentropy(net,t,y,{1},'regularization',0.1)%Entropía cruzada

%_____________________________________________________________
% 8. Estimación del error datos de prueba (Test 2000)
e_test = test_pf-y_test; %y (entrenamiento) - y (Estimado)
figure;
ploterrhist(e_test,'bins',30) %Histograma de errores
R_test = corrcoef(test_pf,y_test); %Coeficientes de correlación
MAE_test = mae(e_test) %mae: Error absoluto medio
MSE_test = immse(test_pf,y_test) %Error medio cuadrático
SSE_test = sse(net,test_pf,y_test,1) %sse: Error de suma cuadrada
SAE_test = sae(net,test_pf,y_test) %Suma absoluta de errores
BCE_test = crossentropy(net,test_pf,y_test,{1},'regularization',0.1)%Entropía cruzada

%_____________________________________________________________
% 9. Pesos y bias
w1 = net.IW{1}; %Pesos de la capa de entrada a oculta
w2 = net.LW{2}; %Pesos de la capa oculta a salida
b1 = net.b{1}; %Sesgo de entrada a la capa oculta
b2 = net.b{2}; %Sesgo de la capa oculta a la salida

%_____________________________________________________________
% 10. Validación
input = [0.353547;30.337191;155.845532;0.0002357;0.000000776;9]% Datos de prueba
output_Diametro = sim(net, input)%Dato estimado
D_Real=0.2849252
D_Error_Absoluto=D_Real-output_Diametro
```

```
D_Error_Relativo=abs((D_Real-output_Diametro)./D_Real)*100

%____________________________________________________________
% 11. Grafica 3D: Superficie de error
figure;
wv = -4:0.4:4; %Límites de la grilla; Tamaño del cuadrante
bv = wv;
ES = errsurf(y,t,wv,bv,'tansig'); %y(Datos predecidos); t(Datos objetivos)
plotes(wv,bv,ES,[60 30])

%____________________________________________________________
% 12. Grafica 3D: Superficie de error MSE
x = test_pD;
y = test_pH;
z = mse_test;
figure;
scatter3(x',y',z','MarkerEdgeColor','k','MarkerFaceColor',[0 .75 .75])
view(-30,10)
xlabel('Perdida de carga (m)')
ylabel('Longitud tubería (m)')
zlabel('Error medio cuadrático')

%____________________________________________________________
% 13. Grafica puntos de disprsión H= 35 m
figure;
H50= test(:,2);
D50=test(:,7);
sz = 90;
scatter(H50,D50,sz,'o')
xlabel('Carga hidráulica (m)')
ylabel('Diámetro (m)')
hold on
H50_test= test(:,2);
D50_test=(y_test)';
sz = 70;
scatter(H50_test,D50_test,sz,'+')
hold on
```