Desalination and Water Treatment
www.deswater.com

doi:10.5004/dwt.2017.20357

73 (2017) 90–100
April

Criteria for improving the traditional artificial neural network methodology applied to predict COP for a heat transformer

E. Martínez-Martínez[a], B.A. Escobedo-Trujillo[a], D. Colorado[b,*], L.I. Morales[c], A. Huicochea[d], J.A. Hernández[d], J. Siqueiros[e]

[a]*Facultad de Ingeniería, Universidad Veracruzana, Campus Coatzacoalcos, Av. Universidad km 7.5 Col. Santa, Isabel, C.P. 96535, Coatzacoalcos, Veracruz, México, email: escorpion410@live.com.mx (E. Martínez-Martínez); bescobedo@uv.mx (B.A. Escobedo-Trujillo)*
[b]*Centro de Investigación en Recursos Energéticos y Sustentables, Universidad Veracruzana. Av. Universidad km 7.5 Col. Santa Isabel, C.P. 96535, Coatzacoalcos, Veracruz, México, Tel. +52(921)2115700 ext. 59230, email: dcolorado@uv.mx,*
[c]*Escuela de Estudios Superiores de Xalostoc, Universidad Autónoma del Estado de Morelos (UAEM), email: laura.morales@uaem.mx*
[d]*Centro de Investigación en Ingeniería y Ciencias Aplicadas (CIICAp), Universidad Autónoma del Estado de Morelos (UAEM), Av. Universidad No. 1001 Col. Chamilpa, C.P. 62209, Cuernavaca, Morelos, México, email: huico_chea@uaem.mx (A. Huicochea); alfredo@uaem.mx (J.A. Hernández)*
[e]*Secretaria de Innovación, Ciencia y Tecnología, Av. Atlacomulco no. 13, esq. Calle de la ronda, Col. Acapantzingo, Cuernavaca, Morelos, C.P. 62440, email: javier.siqueiros@morelos.gob.mx*

Received 11 July 2016; Accepted 16 November 2017

ABSTRACT

This paper introduces three valuable criteria to reduce the number of input variables while predicting the coefficient of performance (COP) (of an absorption heat transformer with duplex components, using an artificial neural network (ANN) model developed in [1], with an experimental database of 1310 pieces of data, in which the experimental COP ranged from 0.10 to 0.36, considering 127 coefficients of adjustment (weights and bias), assuming 16 input variables and a coefficient of determination (r^2) of 0.9969. The database and COP range described above were used in this research considering 50% of data for training and 50%for testing, to present the following criteria: i) creating a correlation matrix to select the input variables in the ANN, ii)performing a residual analysis to validate the ANN models, and if there are several ANN models iii) this criterion could be used to choose the best model. These criteria were studied and included in the traditional ANN methodology proposed by authors [1], however, according to our criteria the best models only used 5 and 6 operation variables in the input layer of ANN architecture, with 33 and 37 coefficients of adjustment, respectively, besides a coefficient of determination (r^2) $r^2 \geq 0.9984$.

Keywords: Modeling; Absorption; Correlation matrix; Homoscedasticity; Normal distribution

1. Introduction

This study suggests three criteria to be considered when working with the traditional methodology of artificial neural network models: i) to create a correlation matrix to select the input variables in the ANN, ii) to validate the ANN models through a residual analysis, and if there are several ANN models iii) this criterion could be applied to select the best model;the preceding criteria with the main purpose of reducing the number of variables involved in the COP prediction by following a rigorous validation process.

The traditional ANN methodology is applied to the experimental database of an absorption heat transformer with duplex components as it has been reported by [1], who

*Corresponding author.

Presented at the EDS conference on Desalination for the Environment: Clean Water and Energy, Rome, Italy, 22–26 May 2016.

have predicted the coefficient of performance assuming 16 input variables, considering 127 coefficients of adjustment (weights and bias) in their ANN model, and presenting a coefficient of determination (r^2) of 0.9969 for the validation process. Our criteria have been applied on [1].

A traditional artificial neural network model has been proposed because of its great capability of estimation when assuming hyperbolic tangent or logarithmic tangent functions between dependent and independent variables. Several authors have previously applied the traditional artificial neural network model to estimate the coefficient of performance of absorption heat transformers; such as, Sozen et al [2] who have determined the performance of an ejector-absorption heat transformer using this kind of network, including four variables in the input layer with the purpose of calculating the coefficient of performance, the exergy coefficient of performance and the circulation ratio, while assuming a log-sigmoid transfer function in the hidden layer. Sozen et al. [3] have presented a traditional artificial neural network model to calculate the exergetic destruction in each component of an absorption cycle, obtaining values of the coefficient of determination higher than 0.98 in the validation analysis. There are other studies related to modeling of heat pumps presented by [4,5]. According to [6], optimizing the ANN configuration will be the object of future research.

To the extent of our knowledge, only the results obtained by [1,7] are comparable to ours. Morales et al. [1] have developed an artificial neural network model that takes into account the input and output temperature of each duplex component, concentration of solution and the pressure and mass flow measurements in the input layer of the ANN architecture. Hernandez et al. [7] have presented an artificial neural network model to calculate the coefficient of performance for a water purification process integrated to an absorption heat transformer.

The differences and novelties found when the present work was compared to the studies developed in [1] and [7] are:

- While model [1] takes into consideration 127 coefficients of adjustment (weight and bias) and 16 input variables to predict the coefficient of performance, our study includes a maximum number of six input variables, this, according to the correlation matrix suggestion without sacrificing accuracy.
- Hernandez et al. [7] presents the coefficient of determination ($r > 0.99$) as the only evidence of the model reliability, yet in our research a residual analysis is applied to select the best model.

This research simplifies the traditional artificial neural network architecture by adding three criteria that translate to novel outcomes:

- Reduced number of the input variables of ANN models
- Enhancements to the ANN models validation
- Improvements to the ANN model developed in [1]

From an instrumentation point of view and following the new criteria proposed in this work, fewer sensors will be required to estimate the coefficient of performance with high confidence.

The remainder of the paper is organized into 6 sections: Section 2 displays the experimental system (absorption heat transformer), Section 3 presents the traditional artificial neural network methodology used to predict the coefficient of performance, and it also shows the traditional validation procedure for ANN models. The implementation of the proposed criteria and the traditional ANN methodology to predict the coefficient of performance for absorption heat transformer is presented in Section 4. Main results, consisting of the three criteria to improve the traditional ANN methodology, are described in Section 5. Finally, Section 6 states some concluding remarks.

2. System description and experimental system

This section displays the experimental system.

Fig. 1 illustrates the experimental system, a compact absorption heat transformer (CAHT) for water purification, built with two duplex units (Generator-Condenser and Absorber-Evaporator) installed in the main equipment, with the purpose of reducing momentum and heat transfer losses.

Both duplex units are falling film helical coil heat exchangers. The generator-condenser consists of a double concentric coil where the working solution circulates through the generator. The condenser is placed in the upper part, in a separate chamber, connected to a cover that lets steam flow, while preventing the condensate to return into the generator. The absorber-evaporator is a chamber where one of the concentric helical coil exchangers works as an evaporator and the other as an absorber. Two independent distributors are installed; the condensed water flows down as a falling film on the wall of the evaporator and the concentrated working solution flows on the wall of the absorber coil. Each duplex unit has an eye hole to monitor the level of absorption of the mixture. Pressure transducers, temperature and flow meters are calibrated and installed in the equipment. Temperatures and pressures are monitored with a computer running a data acquisition system.

The CAHT utilities are electric heating and cooling water systems. Therefore, to simulate waste heat, water is heated and supplied through two different inlets; meanwhile chilled water is supplied by a cooler. There is a magnetic pumping system, with a stainless steel body and a head of pump of 0.07 hp, where a vacuum pump reduces the pressure in the circuit. The thermal load design for the CAHT is 2 kW. The approximate dimensions of the absorption heat transformer are 2.3 m x 2 m x 2 m; its main pieces/parts of equipment and piping are made of stainless steel 316 L, and to reduce heat losses the equipment is covered with foam insulation.

2.1. Experimental proofs

Experimental proofs were carried out considering several concentrations of a lithium bromide solution, while adhering to the next procedure:

- Internal pressure was reduced using a vacuum pump and when it was equal to 4 kPa we supplied the solution.
- While the waste heat auxiliary system was launched, the working solution circulated between the generator and

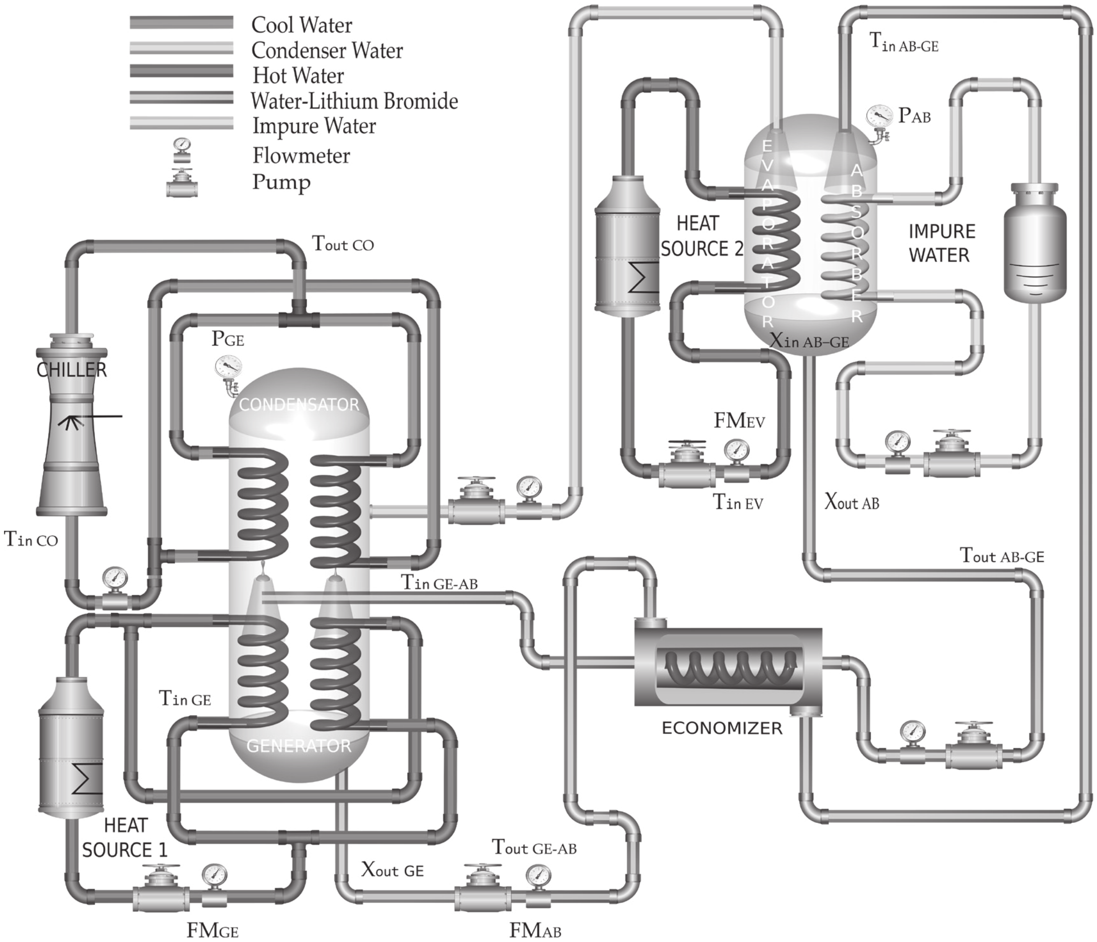

Fig. 1. Analyzed experimental absorption heat transformer system.

absorber and we monitored the levels of solution in both components.

- When the working solution reached saturation temperature, the generator produced steam that went into the condenser. Then, the condensed steam was pumped into the evaporator.
- In the evaporator, the refrigerant (water) changed into vapor; to later join in the absorber the strong solution coming from the generator. Then, an exothermic reaction was attained and the heat was transferred to the water purification system.
- The steady state of the system was reached when the deviations in each temperature meter was ±0.1°C. The operating parameters such as pressure in the absorber and generator, flow-meter lectures, and samples of the solution in the generator and absorber outlets were registered.

3. Traditional artificial neural network methodology

The traditional artificial neural network is one of the most common methods to perform predictive data analysis according to [8]. An artificial neural network consists of a set of highly interconnected processing units called neurons, where each of them accepts a weighted set of inputs and responds with an output. This data modeling tool is able to capture and represent complex input and output relationships because of its ability to learn these directly from the data pattern being modeled.

Let y and $x_1,x_2,x_3,\ldots,x_n$, be the dependent and independent variables, respectively. ANN seeks a relationship between y and $x_1,x_2,x_3,\ldots,x_n$, which can be written as the following mathematical equation:

$$y = ANN\left(x_1, x_2, x_3 \ldots x_n\right) + \varepsilon \tag{1}$$

where ε is a random error.

In the internal architecture of ANN, the first layer is the input layer, where the independent variables are introduced to the network; the last layer is the output layer that shows the dependent variables based on the network computations. Between these two, there is one or more layers, called hidden layer(s), which could be located [6]. Each element of the hidden layers is connected to each input neuron through the weight matrix. As suggested by [9], the architecture of a standard network for a function approximation is the multilayer perceptron (or feed-forward artificial neural network).

Several authors have proposed the following method for ANN training:

- First, gather an experimental database based on the operational variable range of the system and equipment available. However, a clear and effective procedure to select the most important operational variables before the training process is still lacking.
- Second, analyze the database and in some cases eliminate the noise in the acquisition process without discarding variables.
- Third, define the percentage of data for training and validation, also calculate all necessary parameters: architecture, number of hidden layers, optimization algorithm, activation function, number of iterations, and if it is possible, the most important factor, the number of hidden neurons per layer.
- At the end, calculate the difference between the target output and network output, which is the error and it should be minimized.

In the literature, tangent sigmoid transfer and logarithmic sigmoid transfer functions are mostly suggested in the hidden layer by [2,7], because these functions have presented high ability to predict a dependent variable. The artificial neural network model with a tangent sigmoid transfer function in a hidden layer and a linear function in the output layer is given by:

$$y_k = \sum_{j=1}^{J}\left[Wo(k,j)\cdot\left(\frac{2}{1+\exp\left(-2\sum_{r=1}^{\tilde{R}}\left(\left(Wi\{j,r\}p_r\right)+b1_j\right)\right)}-1\right)\right]+b2_{\{k\}} \qquad (2)$$

where J is the number of neurons in the hidden layer, Wi represents the weights in the input hidden layer, $b1_j$ is the j-th of bias in the hidden layer, Wo denotes the weights in the hidden output layer, R is the input neuron number, k is the output neuron number and $b2_{\{k\}}$ is the k-th value of bias in the output layer.

Fig. 2 shows the prediction procedure of ANN.

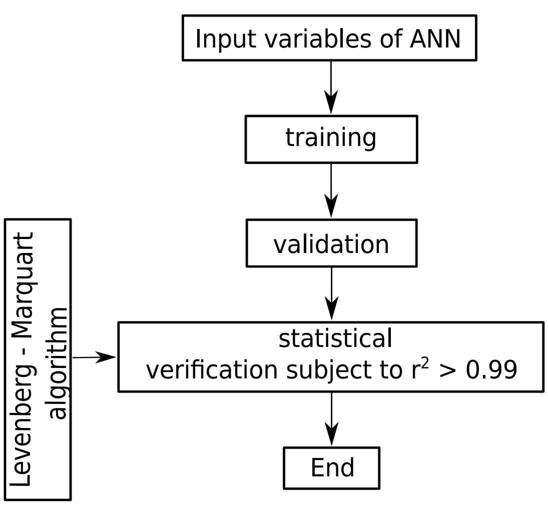

Fig. 2. Prediction procedure of an ANN.

3.1. Validation of the ANN model

Traditionally, the process in which the calculation of the dependent variable with the ANN model is compared to experimental results is called validation. According to [2] and [3] the coefficient of determination is enough to validate the ANN model and it is generally assumed as $r^2 > 0.98$. Colorado et al. [9] corroborated the ANN model with coefficient of determination and linear regression between experimental and simulated COP values. As stated by [9], linear regression focuses on the slope and intercept values considered close to 1 and 0 respectively; as a result,obtaining the correct model validation. Nevertheless, the error analysis in the linear regression was not performed.

4. Main result: Criteria to improve the ANN

This section describes the proposed criteria to improve the traditional artificial neuronal network (Section 3) consisting of three stages: 1) correlation analysis to select input variables of ANN, 2) residual analysis in the validation of ANN models and 3) selection of the best ANN model. These steps are explained below:

1. Correlation analysis

   Supposing that there are n input variables for the artificial neuronal network, $x_1, x_2, \ldots, x_n$ and we want to study the correlations among them, the correlation coefficient is calculated between each pair of variables (x_i, x_j), $i,j = 1, 2, \ldots, n$. The correlation coefficient r measures the linear relationship between paired values of two variables X and Y. The mathematical definition of is given by:

$$r = \frac{\sum_{i=1}^{i=n}(x_i-\bar{x})(y_i-\bar{y})}{(n-1)s_x s_y} \qquad (3)$$

where x_i (resp. y_i) is the i-th observation of a random variable X (resp. Y), $\bar{x}$, $\bar{y}$ and S_x, S_y are the means and sample standard deviations of X and Y respectively, n is the number of pairs of observations.

Because there are many coefficients, it is convenient to arrange them in an orderly and systematic fashion; like in a *correlation matrix*, see [10] for more details.

As other authors have previously mentioned in [10,11] there are numerous applications for correlation analysis. In this research we used the correlation coefficient in order to select those variables that are the most highly correlated with each other in order to reduce the number of input variables in the traditional neural network presented in Section 3. With regard to the prediction of the COP, this was selected as an output variable because it was the most representative estimation of the system.

2. Residual analysis in the validation of the obtained ANN models

   Assuming that in the first step we selected the following set of input variables $x_3, x_{11}, x_7, x_{16}, x_1, x_{12}$ there is a possibility of generating different ANN models (with high coefficient of determination) by changing the number of input variables in ANN.

The certainty of the neural network models is traditionally confirmed by calculating:

- The coefficient of determination
- The linear regression between experimental and simulated data obtained from the neural network model, as previously mentioned at the end of Section 3.

A simple linear regression is a statistical method that allows us to study and summarize with an equation the nature of the relationship between two variables, such as: Variable X, considered the predictor or independent variable and variable Y, considered the response, outcome, or dependent variable. The equation for the best fit line is expressed as:

$$Y = \beta_0 + \beta_1 X + \varepsilon \tag{4}$$

The parameters, β_0 and β_1 are usually unknown and must be estimated using data, whereas ε is a random variable.

Key assumptions of linear regression were formulated, for example, see [10],

Assumption 1.

a) The error has a mean of zero,
b) The error normally distributed,
c) The error has constant variance (homoscedasticity) and,
d) The error is independent.

As presented by authors [10,11], these assumptions can be verified with an analysis of residual errors which are defined as follows.

Definition 2. *Residual errors are defined as:*

Residual errors = $Y_i = \hat{Y}_i$.

where is the i-th observation of the variable Y, $\hat{Y}_i$ is the i-th estimated parameter of Y_i, i.e., $\hat{Y}_i = \beta_0 + \beta_1 x_i$ when $X = x_i$, $i = 1,2, ...,n$.

The residual analysis verifies that the Assumption 1 is true using the residual errors instead of a random error ε. In this work, this confirmation was made in graphic form [10,11].

The criteria developed above is summarized in Fig. 3.

3. Selection of the best ANN model.
   In case of obtaining several trained and validated ANN models, we select the best model following the next criterion.
   Given the frequency histogram of residual errors, the best ANN model must satisfy:
   a) The arithmetic mean of residual errors close to zero and,
   b) Low variance

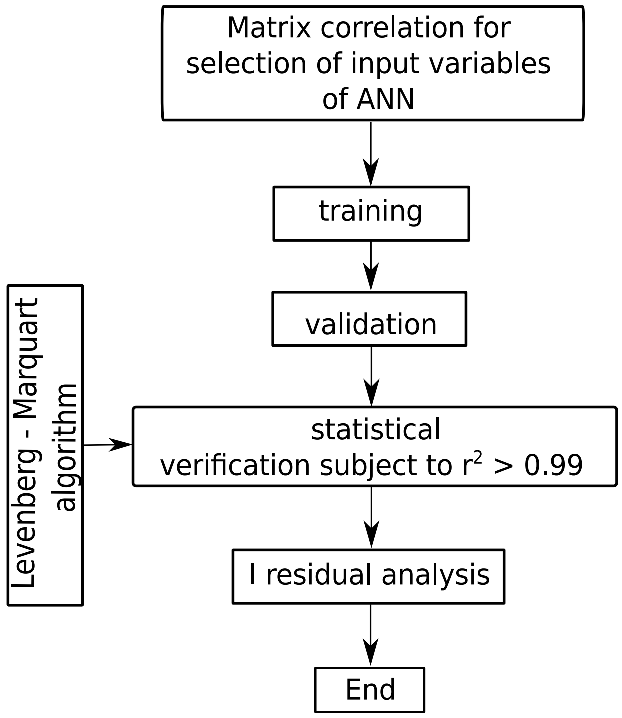

Fig. 3. Mathematical selection of input variables and improved methodology.

Properties a) and b) stand out from the following fact: If there is a density function for the random error given in Eq. (2), we could consider the expected value with respect to density, E_f, in both sides of equality (2) to obtain

$$Y = \beta_0 + \beta_1 \mathrm{X} + E_f[\varepsilon] \tag{5}$$

which implies that

$$E_f[\varepsilon] = \mathrm{Y} - (\beta_0 + \beta_1 \mathrm{X}) \tag{6}$$

Thus, the best ANN model must satisfy $E_f[\varepsilon] = 0$. As the density function is unknown, we do not have the expected exact value $E_f[\varepsilon]$, hence the arithmetic mean of residual errors is a good estimator.

5. Results and discussion

From the research that has been conducted improvements to the traditional ANN methodology for COP prediction of an absorption heat transformer are: 1) correlation matrix to select input variables in order to build ANN models following traditional methodology, 2) residual analysis for artificial neural networks models validation, considering from three to six input variables and 3) criterion to select the best ANN model.

The 16 input variables considered in this analysis are: T_{inGE}, $T_{inGE\text{-}AB}$, $T_{outAB\text{-}GE}$, $T_{inAB\text{-}GE}$, $T_{outGE\text{-}AB}$, T_{inCO}, T_{outCO}, T_{inEV}, $X_{inAB\text{-}GE}$, X_{outAB}, X_{outGE}, P_{AB}, P_{GE}, FM_{GE}, FM_{EV} and FM_{AB}, where T is temperature, X is concentration of solution, P is pressure, FM is mass flow rate, GE is the generator, AB is the absorber, CO is the condenser and EV is the evaporator.

1. Correlation matrix for selecting input variables for ANN.

First, the correlation matrix of the measured variables was calculated, allowing us to identify a possible relationship among them. This correlation matrix is shown in Tables 1 and 2.

Based on a visual inspection of the correlation matrix, the following can be concluded:

- T_{inGE}, T_{inEV} and P_{AB} are the operating variables that are strongly correlated with the COP and they could be used to build a simple model.
- If there is more experimental information available: $T_{inGE\text{-}AB}$, $T_{inAB\text{-}GE}$, $T_{outGE\text{-}AB}$ and P_{AB} could be registered to simulate the COP.

Table 1
Correlation matrix

	$T_{in\ GE}$	$T_{in\ GE–AB}$	$T_{out\ AB–GE}$	$T_{in\ AB–GE}$	$T_{out\ GE–AB}$	$T_{in\ CO}$	$T_{out\ CO}$	$T_{in\ EV}$	$X_{in\ AB–GE}$
$T_{in\ GE}$	1.0000	0.8485	0.7255	0.7972	0.9175	0.3502	0.1423	0.8914	0.1524
$T_{in\ GE–AB}$	0.8485	1.0000	0.5916	0.9547	0.9372	0.0688	–0.1978	0.7120	–0.4073
$T_{out\ AB–GE}$	0.7288	0.5916	1.0000	0.5580	0.6713	–0.1136	–0.1732	0.7235	0.1693
$T_{in\ AB–GE}$	0.7972	0.9547	0.5580	1.0000	0.9528	0.0476	–0.2885	0.7007	–0.4041
$T_{out\ GE–AB}$	0.9175	0.9372	0.6713	0.9528	1.0000	0.1568	–0.1573	0.8273	–0.2561
$T_{in\ CO}$	0.3503	0.0688	–0.1136	0.0476	0.1568	1.0000	0.8273	0.3132	–0.0157
$T_{out\ CO}$	0.1423	–0.1978	–0.1732	–0.2885	–0.1573	0.8273	1.0000	0.0752	0.1006
$T_{in\ EV}$	0.8914	0.7120	0.7235	0.7007	0.8273	0.3132	0.0752	1.0000	–0.0150
$X_{in\ AB–GE}$	–0.1524	–0.4073	0.1693	–0.4041	–0.2561	–0.0157	0.1006	–0.0150	1.0000
$X_{out\ AB}$	–0.0996	0.1386	0.0789	0.0453	–0.0131	–0.5191	–0.3590	–0.2868	0.1379
$X_{out\ GE}$	–0.0346	–0.1100	0.2951	–0.2078	–0.1322	–0.3832	–0.1000	–0.1864	0.5723
P_{AB}	–0.8058	–0.6406	–0.7191	–0.6213	–0.7644	–0.2273	–0.0278	–0.9268	–0.0827
P_{GE}	–0.4399	–0.3422	–0.1743	–0.3265	–0.4311	–0.2790	–0.0665	–0.3986	0.0899
FM_{GE}	0.0286	0.0087	0.4281	0.0137	0.0056	–0.4418	–0.2811	0.0998	0.3806
FM_{EV}	–0.0602	–0.1165	–0.0418	–0.0730	–0.0750	–0.0898	–0.0357	–0.0028	0.5296
FM_{AB}	–0.1271	0.1258	–0.2990	0.0373	–0.0537	–0.2317	–0.2562	–0.2449	–0.3567
COP	0.3213	0.2229	0.4307	0.2642	0.3965	–0.0876	–0.2035	0.3123	0.3021

Table 2
Correlation matrix (Continuation)

	$X_{out\ AB}$	$X_{out\ GE}$	P_{AB}	P_{GE}	FM_{GE}	FM_{EV}	FM_{AB}	COP
$T_{in\ GE}$	–0.0996	–0.0346	–0.8058	–0.4399	0.0286	–0.0602	–0.1271	0.3213
$T_{in\ GE–AB}$	0.1386	–0.1100	–0.6406	–0.3422	0.0087	–0.1165	0.1258	0.2229
$T_{out\ AB–GE}$	0.0789	0.2951	–0.7191	–0.1743	0.4281	–0.0418	–0.2990	0.4307
$T_{in\ AB–GE}$	0.0453	–0.2078	–0.6213	–0.3265	0.0137	–0.0730	0.0373	0.2642
$T_{out\ GE–AB}$	–0.0131	–0.1322	–0.7644	–0.4311	0.0056	–0.0750	–0.0537	0.3965
$T_{in\ CO}$	–0.5191	–0.3832	–0.2273	–0.2790	–0.4418	–0.0898	–0.2317	–0.0876
$T_{out\ CO}$	–0.3590	–0.1000	–0.0278	–0.0665	–0.2811	–0.0357	–0.2562	–0.2035
$T_{in\ EV}$	–0.2868	–0.1864	–0.9268	–0.3986	0.0998	–0.0028	–0.2449	0.3123
$X_{in\ AB–GE}$	0.1379	0.5723	–0.0827	0.0899	0.3806	0.5296	–0.3567	0.3021
$X_{out\ AB}$	1.0000	0.7287	0.2223	0.1429	0.3824	0.2448	0.1966	0.1293
$X_{out\ GE}$	0.7287	1.0000	0.1459	0.1638	0.5939	0.3640	–0.1719	0.1988
P_{AB}	0.2223	0.1459	1.0000	0.3743	–0.0616	–0.0160	0.1197	–0.3875
P_{GE}	0.1429	0.1638	0.3743	1.000	0.1832	–0.0389	–0.1642	–0.6491
FM_{GE}	0.3824	0.5939	–0.0616	0.1832	1.0000	0.2970	–0.3489	0.0752
FM_{EV}	0.2448	0.3640	–0.0160	–0.0389	0.2970	1.0000	0.0924	0.0898
FM_{AB}	0.1966	–0.1719	0.1197	–0.1642	–0.3489	0.0924	1.0000	–0.0880
COP	0.1293	0.1988	–0.3875	–0.6491	0.0752	0.0898	–0.0880	1.0000

- Input variables previously selected only include experimental information for the absorption process (Generator-Absorber). Therefore, in order to include the vapor production in the evaporator, the correlation matrix reveals that we can associate T_{inEV}, T_{inGE}, $T_{inGE\text{-}AB}$, $T_{inAB\text{-}GE}$ and $T_{outGE\text{-}AB}$ to COP.
- Finally, the COP is correlated with T_{inGE}, $T_{inGE\text{-}AB}$, $T_{inAB\text{-}GE}$, $T_{outGE\text{-}AB}$, T_{inEV} and P_{AB}, which means that we are able to measure practically the entire system.

According to the matrix inferences, shown in Tables 1 and 2, to predict the coefficient of performance of our system, the proposed number of neurons in the input layer can only be within 3 and 6.

For each suggestion of the correlation matrix of the input variables described above, we developed four ANN models considering the following conditions: For the artificial neural network, the experimental database consisted of 1310 pieces of data, considering 50% of them for learning and the other 50% for testing. The coefficients of the network, weights and bias, the number of iterations of the optimization algorithm were calculated during the training stage, thus minimizing a root mean square error (RMSE) between simulated and experimental data. The best weights and biases were determined to obtain the highest possible value of the regression coefficient. In this work, for the learning process, we fixed the number of neurons in the hidden layer as four. Four ANN models were developed with a hyperbolic tangent transfer function in one hidden layer and a linear function in the output layer, see Fig. 4. Throughout our analysis the Levenberg-Marquardt method was selected and used for the network optimization following the suggestion by [12].

The four ANN models are described as follows:

ANN model with 6 neurons in the input layer. Fig. 4(a) displays the ANN model considering six operation variables (T_{inGE}, $T_{inGE\text{-}AB}$, $T_{inAB\text{-}GE}$, $T_{outGE\text{-}AB}$, T_{inEV} and P_{AB}) in the input layer of ANN architecture. ANN Eq. (2) was developed assuming: $J = 4$, $R = 6$, $k = 1$ and $y_k = COP_{sim}$. Thus, ANN model for this case is given as:

$$COP_{sim} = 2\left[\frac{0.6578}{1+e^{\varphi_1}} - \frac{0.4038}{1+e^{\varphi_2}} - \frac{0.2596}{1+e^{\varphi_3}} - \frac{0.0655}{1+e^{\varphi_4}}\right] - \left(0.6578 - 0.4038 - 0.2596 - 0.0655\right) + 0.1822 \quad (7)$$

where:

$$\varphi_1 = -2(-8.4816\mathrm{Tin}_{GE} + 6.2576\mathrm{Tin}_{GE-AB} + 2.7682\mathrm{Tin}_{AB-GE} - 4.6990\mathrm{Tout}_{GE-AB} - 7.1527\mathrm{Tin}_{EV} + 1.0384P_{AB} + 8.1859)$$

$$\varphi_2 = -2(-14.2334\mathrm{Tin}_{GE} + 12.4429\mathrm{Tin}_{GE-AB} + 8.7167\mathrm{Tin}_{AB-GE} - 13.6168\mathrm{Tout}_{GE-AB} - 7.5706\mathrm{Tin}_{EV} + 2.7109P_{AB} + 10.5137),$$

$$\varphi_3 = -2(-6.5909\mathrm{Tin}_{GE} + 3.1316\mathrm{Tin}_{GE-AB} + 0.6686\mathrm{Tin}_{AB-GE} - 2.5646\mathrm{Tout}_{GE-AB} - 16.6884\mathrm{Tin}_{EV} + 3.7124P_{AB} + 13.7787)$$

and

$$\varphi_4 = -2(-2.5631\mathrm{Tin}_{GE} + 8.8186\mathrm{Tin}_{GE-AB} - 1.5703\mathrm{Tin}_{AB-GE} - 16.2158\mathrm{Tout}_{GE-AB} + 10.8729\mathrm{Tin}_{EV} - 1.7951P_{AB} + 3.1875)$$

A comparison was made between experimental and simulated COP values. As a result, the next linear regression was obtained:

$$COP_{sim} = 0.00043121 + 0.9978\ COP_{exp} \quad (8)$$

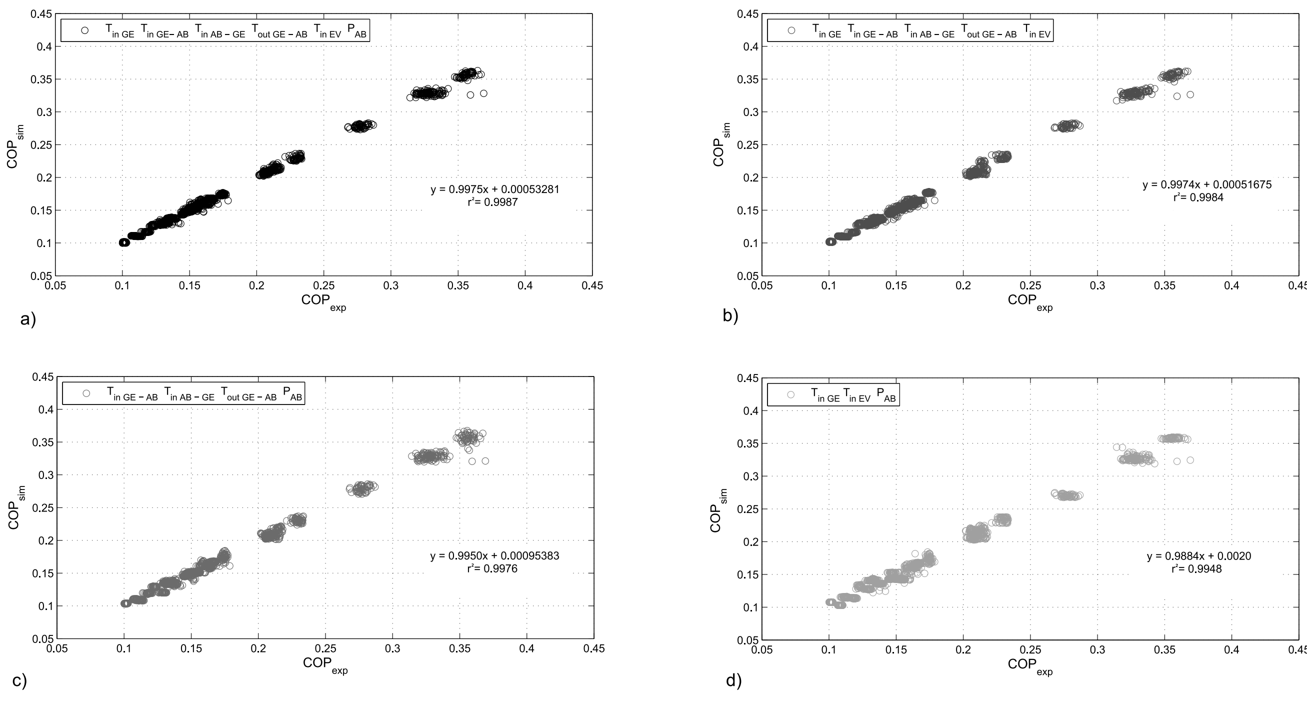

Fig 4. Comparison between the COP_{EXP} and the predicted COP_{ANN}, considering a) 6 neurons, b) 5 neurons, c) 4 neurons and d) 3 neurons;in the input layer.

From Eq. (8) it can be seen that the intercept and the slope approximated to 0 and 1, respectively. Indicating that COP_{exp} could be predicted with the ANN model, the coefficient of determination (r^2) was calculated, resulting in a value equal to 0.9987.

ANN model with 5 neurons in the input layer. Fig. 4b presents the proposed model. The input variables selected by the correlation matrix for ANN model involve the absorption process and the input temperature of the evaporator. The five operating variables were: T_{inEV}, T_{inGE}, $T_{inGE-AB}$, $T_{inAB-GE}$ and $T_{outGE-AB}$.

The ANN model considering 5 operating variables as we developed it is:

$$COP_{sim} = 2\left[\frac{-0.8681}{1+e^{\varphi_1}} + \frac{0.1671}{1+e^{\varphi_2}} - \frac{0.0876}{1+e^{\varphi_3}} - \frac{0.7073}{1+e^{\varphi_4}}\right] - \left(-0.8681 + 0.1671 - 0.0876 - 0.7073\right) + 0.2047 \quad (9)$$

where:

$$\varphi_1 = -2(5.3875\mathrm{Tin}_{GE} - 5.0064\mathrm{Tin}_{GE-AB} - 2.6359\mathrm{Tin}_{AB-GE} + 6.1035\mathrm{Tout}_{GE-AB} + 8.1112\mathrm{Tin}_{EV} - 8.6445),$$

$$\varphi_2 = -2(34.0036\mathrm{Tin}_{GE} - 6.5302\mathrm{Tin}_{GE-AB} - 32.6466\mathrm{Tin}_{AB-GE} + 19.2428\mathrm{Tout}_{GE-AB} + 30.2894\mathrm{Tin}_{EV} - 33.2274)$$

$$\varphi_3 = -2(0.2642\mathrm{Tin}_{GE} + 5.8341\mathrm{Tin}_{GE-AB} + 4.8368\mathrm{Tin}_{AB-GE} - 21.4879\mathrm{Tout}_{GE-AB} + 11.6259\mathrm{Tin}_{EV} + 1.2319),$$

and

$$\varphi_4 = -2(-3.7847\mathrm{Tin}_{GE} + 5.6966\mathrm{Tin}_{GE-AB} + 2.8396\mathrm{Tin}_{AB-GE} - 8.8353\mathrm{Tout}_{GE-AB} - 9.8910\mathrm{Tin}_{EV} + 9.7929).$$

The coefficient of determination between experimental and simulated COP (COP_{sim}) value was equal to 0.9984 and the linear regression obtained was:

$$COP_{sim} = 0.00051675 + 0.9984\ COP_{exp} \quad (10)$$

ANN model with 4 neurons in the input layer. Fig. 4c depicts the comparison between experimental and simulated COP values when the input operating variables were: $T_{inGE-AB}$, $T_{inAB-GE}$, $T_{outGE-AB}$ and P_{AB}. The coefficient of determination decreased to 0.9976, which indicates that knowing the inlet temperature of the evaporator contributes significantly to the COP prediction.

The mathematical model to predict COP as function of 4 input operating variables is as follows:

$$COP_{sim} = 2\left[\frac{0.5677}{1+e^{\varphi 1}} + \frac{+3.0483}{1+e^{\varphi 2}} + \frac{-2.6140}{1+e^{\varphi 3}} + \frac{-3.0089}{1+e^{\varphi 4}}\right] - 0.5677 + 3.0483 - 2.6140 - 3.0089) - 1.8245, \quad (11)$$

where:

$$\varphi_1 = -2(6.3642\mathrm{Tin}_{GE-AB} - 0.8439\mathrm{Tin}_{AB-GE} + 1.5006\mathrm{Tout}_{GE-AB} - 19.9528\mathrm{P}_{AB} + 2.3051,$$

$$\varphi_2 = -2(0.1968\mathrm{Tin}_{GE-AB} + 2.0206\mathrm{Tin}_{AB-GE} - 1.6507\mathrm{Tout}_{GE-AB} - 7.3422\mathrm{P}_{AB} + 1.7520),$$

$$\varphi_3 = -2(1.1529\mathrm{Tin}_{GE-AB} + 2.0711\mathrm{Tin}_{AB-GE} - 1.5560\mathrm{Tout}_{GE-AB} - 10.2042\mathrm{P}_{AB} + 2.1190)$$

and

$$\varphi_4 = -2(-0.2428\mathrm{Tin}_{GE-AB} + 1.5574\mathrm{Tin}_{AB-GE} - 1.5029\mathrm{Tout}_{GE-AB} - 3.4281\mathrm{P}_{AB} + 0.1387).$$

The linear regression obtained for simulated and experimental values of COP is given by:

$$COP_{sim} = 0.00095383 + 0.9950\ COP_{exp} \quad (12)$$

ANN model with 3 neurons in the input layer. Fig. 4d shows the proposed experimental test as a case with a few measurement instruments. According to the previous models, the COP prediction assuming 3 operating variables as input neurons is expressed as:

$$COP_{sim} = 2\left[\frac{-0.1644}{1+e^{\varphi_1}} - \frac{0.1659}{1+e^{\varphi_2}} - \frac{0.1232}{1+e^{\varphi_3}} - \frac{0.0186}{1+e^{\varphi_4}}\right] - \left(-0.1644 - 0.1659 - 0.1232 - 0.0186\right) + 0.2456 \quad (13)$$

where:

$$\varphi_1 = -2\left(7.2530\mathrm{Tin}_{GE} + 18.5387\mathrm{Tin}_{EV} - 5.6165\mathrm{P}_{AB} - 18.1581\right),$$
$$\varphi_2 = -2\left(-11.2980\mathrm{Tin}_{GE} - 45.0073\mathrm{Tin}_{EV} - 22.2448\mathrm{P}_{AB} + 48.3099\right),$$
$$\varphi_3 = -2\left(5.4868\mathrm{Tin}_{GE} + 61.0514\mathrm{Tin}_{EV} + 64.6052\mathrm{P}_{AB} - 64.4270\right)$$

and

$$\varphi_4 = -2\left(-10.2878\mathrm{Tin}_{GE} - 30.3397\mathrm{Tin}_{EV} - 22.7721\mathrm{P}_{AB} - 33.3396\right).$$

The comparison between experimental and simulated data of the COP was done, obtaining a coefficient of determination of 0.9948 and the linear regression model is given by:

$$COP_{sim} = 0.0020 + 0.9884\ COP_{exp} \quad (14)$$

Remark. Table 3 shows the results of coefficient of determination and linear regression coefficients for the models previously described. As can be seen in Table 3, the four ANN model presented the intercept close to 0 and the slope close to 1. In accordance with [2], [7] and [9] the four models predicted a COP with high confidence.

From Table 3, it is important to emphasize that, the artificial neural networks models developed in this work had higher values of r^2, and they presented better values of slope and intercept in comparison with the model proposed by Morales et al [1]. It is clear that, the r^2 value and the statistical information such as, slope and intercept in the linear regression justify the suggested criteria to improve the traditional methodology presented in this work.

2. Residual analysis of ANN models

The ANN models presented in the second step were validated with the coefficient of determination and linear regression between experimental and simulated data as it has traditionally been demonstrated by several authors, such as [2,7,9]. However, these authors haven't taken into

Table 3
Results of each ANN model presented in this work

Input operation variable	Number of hidden layer neurons	r^2	Slope	Intercept
16 input operation variables Morales et al. [1]	7	0.9969	0.9787	0.0014
T_{inGE}, $T_{inGE\text{-}AB}$, $T_{inAB\text{-}GE}$, $T_{outGE\text{-}AB}$, T_{inEV} and P_{AB}, Eq. (3) in this work.	4	0.9988	0.9978	0.00043121
T_{inEV}, T_{inGE}, $T_{inGE\text{-}AB}$, $T_{inAB\text{-}GE}$ and $T_{outGE\text{-}AB}$, Eq. (5) in this work.	4	0.9984	0.9984	0.00051675
$T_{inGE\text{-}AB}$, $T_{inAB\text{-}GE}$, $T_{outGE\text{-}AB}$ and P_{AB}, Eq. (7) in this work.	4	0.9976	0.9950	0.00095383
T_{inGE}, T_{inEV} and P_{AB}, Eq. (9) in this work.	4	0.9948	0.9884	0.00200000

account if the assumptions (see Assumption 1) on linear regression models have been satisfied. As it has been reported by others [10,11], if the previous assumptions are not supported, then, the linear regression model could not be suitable for the data.

The points on the graph of the residual errors with at least 3 standard deviations from the mean are known as outliers, which may affect the adequacy of the linear regression, therefore, in this work we performed residual analyses to verify the assumptions of the linear regression models (3)–(6) with and without outliers, with the aim of discarding any violation to Assumption 1.

a)–b) Mean of zero and normal distribution. Histograms of the residual errors with and without outliers are presented in Figs. 5 and 6, where it can be observed that residual errors for the ANN models with 6, 5 and 4 input variables have a normal distribution with a mean of zero. But, the ANN model with 3 input variables does not satisfy the assumption of normality.

To simplify this study, we only demonstrate the analysis of assumptions 1 *c*) and 1 *d*) for ANN model with 6 input variables, although, the analysis was done for all ANN models and it was found that they satisfy assumptions 1 *c*) and 1 *d*).

c)–d) independent and has constant variance (homoscedasticity). Fig. 7 illustrates the predicted coefficient of performance against standardized residual with and without outliers for ANN model with 6 input variables. It can be seen that the average of standardized residual is zero, the variance is constant and the residuals are independent.

Remark. As it can be noted, ANN models with 6, 5, and 4 input variables satisfy the linear regression assumptions (Assumption 1), but ANN model with 3 input variables does not satisfy the assumption of normality. So, for authors [2, 7] and [9] this model has been approved, though, in our study it has been discarded.

3. Selection of the better ANN model

Based on the selection criterion established in Section 4, the best ANN model must satisfy:

a) The arithmetic mean of residual errors close to zero and,
b) Low variance.

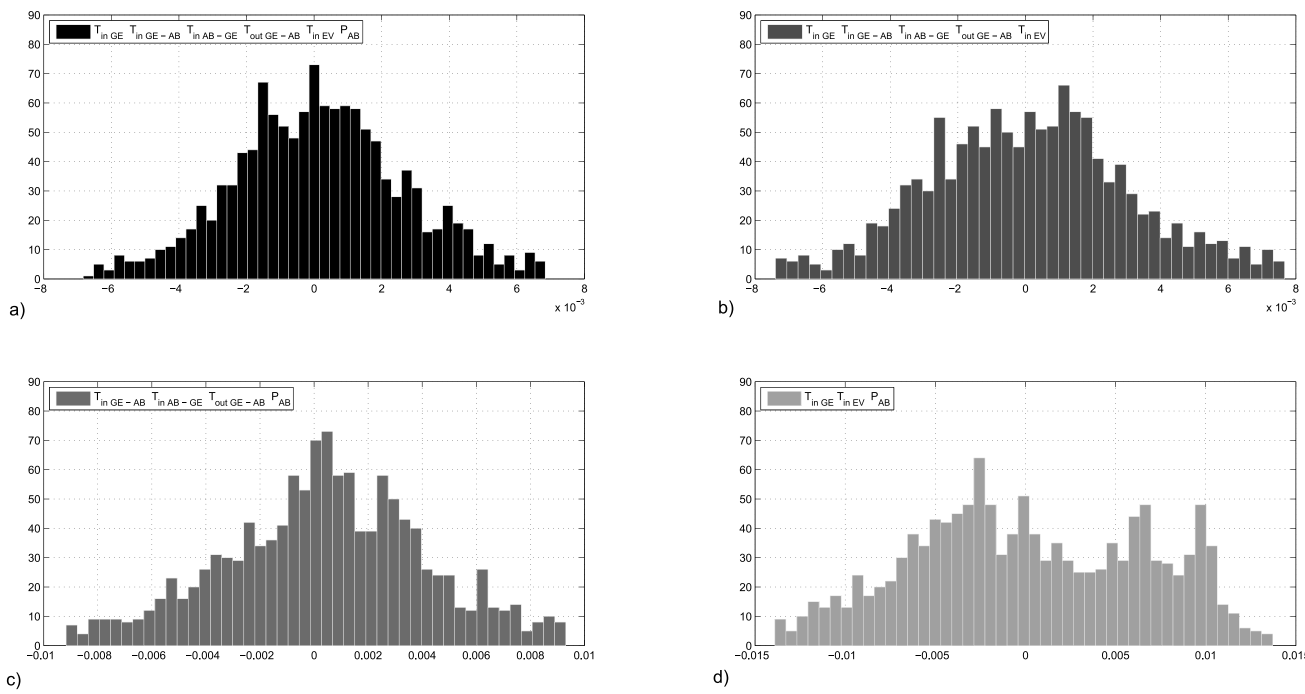

Fig. 5. Histogram of the standardized residuals without outliers.

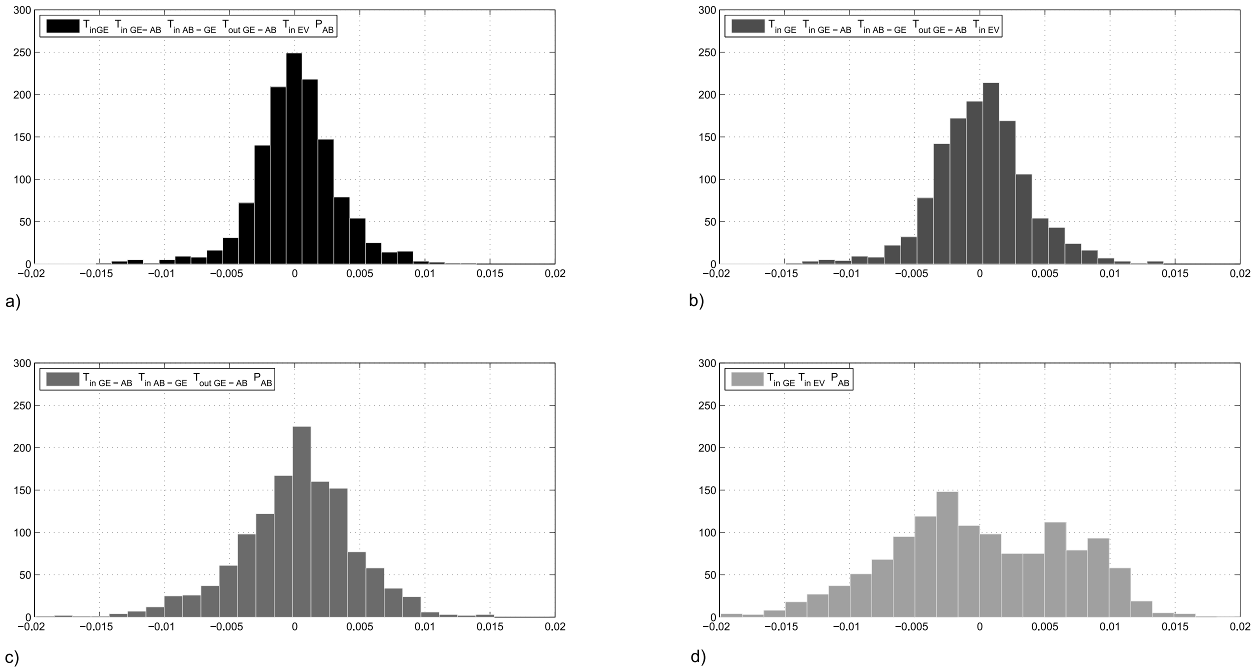

Fig. 6. Histogram of the standardized residuals with outliers.

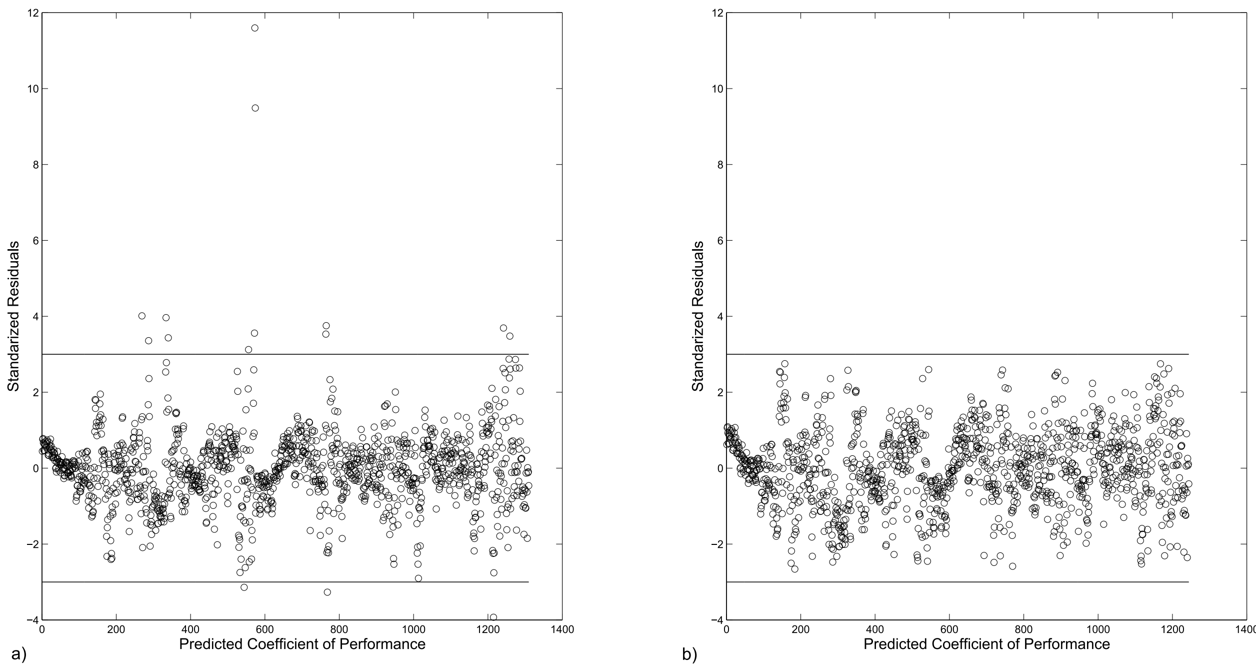

Fig. 7. Predicted coefficient of performance against standardized residual with and without outliers for the ANN model with 6 input variables.

Under this criterion ANN models with 4 ($T_{inGE-AB}$, $T_{inAB-GE}$, $T_{outGE-AB}$ and P_{AB}) and 3 (T_{inGE}, T_{inEV} and P_{AB}) input variables must to be discarded, see Table 4.

Table 4 shows that the ANN model with 6 input variables ($T_{inEV,}$ T_{inGE}, $T_{inGE-AB}$, $T_{inAB-GE}$ and $T_{outGE-AB}$) has the lowest mean and the ANN model with 4 input variables (T_{inGE}, $T_{inGE-AB}$, $T_{inAB-GE}$, $T_{outGE-AB}$, T_{inEV} and P_{AB}) has the lowest standard deviation. Thus, none of the two ANN models simultaneously satisfies both requirements given in the selection criterion. But, because the validation of these ANN models were confirmed with the coefficient of determination and linear regression analysis, besides a residual analysis to verify that the linear regressions were correct (adequacy of fit), consequently, it can be concluded that these ANN models are valid. Therefore, both ANN models were selected.

Table 4
Comparison of means and standard deviations for each ANN models obtained considering the residual errors with and without outliers

Input variables to the ANN	Mean with outliers	Mean without outliers	Standard deviation with outliers	Standard deviation without outliers
T_{inGE}, $T_{inGE-AB}$, $T_{inAB-GE}$, $T_{outGE-AB}$, T_{inEV} and P_{AB}	0.000077913	0.00015703	0.0035	0.0025
T_{inEV}, T_{inGE}, $T_{inGE-AB}$, $T_{inAB-GE}$ and $T_{outGE-AB}$	0.000042659	0.000092035	0.0039	0.0030
$T_{inGE-AB}$, $T_{inAB-GE}$, $T_{outGE-AB}$ and P_{AB}	0.000061085	0.00039466	0.0048	0.0037
T_{inGE}, T_{inEV} and P_{AB}	−0.000058619	0.00019679	0.0071	0.0064

6. Conclusions

Traditional methodology of ANN models has been improved by adding the following criteria: i) correlation matrix to select the input variables in the ANN, ii) residual analysis to validate the ANN models, in our case there were several ANN models and these were validated in a traditional way, therefore the following criterion was applied iii) criteria to select the best model, all previously stated, in order to simplify the architecture of ANN models, by adding an improved validation and considering a mathematical property to select the best model (only if there are several). This last criterion was applied to predict the coefficient of performance of an absorption heat transformer with duplex components.

The main contributions of this work are:

1. Simplifying the traditional artificial neural network model by selecting input operating variables of ANN, based on a correlation matrix suggestion. The accuracy to predict the coefficient of performance has not been sacrificed, the new artificial neural models have shown a coefficient of determination higher than 0.98. Thus, the models presented in this research, considering five and six input operating variables are simpler than the ANN model presented by [1].
2. The residual analysis has been presented as tool with the objective of improving the methodology of traditional validation in the artificial neural network models.
3. The best ANN models fulfilled the requirements such as, the arithmetic mean of residual errors close to zero and, low variance.
4. Even though, for some authors [2,7] and [9] ANN model with 3 input variables could be good. In this study, it has been discarded, because it does not satisfy the assumption of normality.
5. The ANN models with 3 and 4 input variables have been discarded because do not satisfy the selection criterion given in Section 4.
6. Artificial neural networks with four neurons in the hidden layer, five and six neurons in the input layer were successfully trained and validated for coefficient of performance prediction. These models have shown that measuring T_{inGE}, $T_{inGE-AB}$, $T_{inAB-GE}$, $T_{outGE-AB}$, T_{inEV} and P_{AB} is enough to predict the coefficient of performance $r^2 = 0.9988$ with high confidence and complying with strict residual analysis.

Acknowledgements

This work is a collaborative efforts between UV-CA-412 and UAEMOR-CA-70 research groups. The authors of research group UV-CA-412 would like to thank PRODEP-SEP project number 23688 to sponsor the present research.

References

[1] L.I. Morales, R.A. Conde-Gutierrez, J.A. Hernandez, A. Huicochea, D. Juarez-Romero, J. Siqueiros, Optimization of an absorption heat transformer with two-duplex components using inverse neural network and solved by genetic algorithm, Appl. Therm. Eng., 85(25) (2015) 322–333.
[2] Sozen, E. Arcaklioglu, M. Ozalpa, S. Yucesu, Performance parameters of an ejector-absorption heat transformer, Appl. Energ., 80 (2005) 273–289.
[3] Sozen, E. Arcaklioglu, Exergy analysis of an ejector-absorption heat transformer using artificial neural network approach, Appl. Therm. Eng., 27 (2007) 481–491.
[4] N. Gunasekar, M. Mohanraj, V. Velmurugan, Artificial neural network modeling of a photovoltaic-thermal evaporator of solar assisted heat pumps, Energy, 93 (2015) 908–922.
[5] M. Aktas, S. Sevik, M. Bahadir-Ozdemir, E. Gonen, Performance analysis and modeling of a closed-loop heat pump dryer for bay leaves using artificial neural networks, Appl. Therm. Eng., 87 (2015) 714–723.
[6] M. Mohanraja, S. Jayaraj, C. Muraleedharan, Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systems. A review, Renew. Sust. Energ. Rev., 16 (2012) 1340–1358.
[7] J.A. Hernandez, A. Bassam, J. Siqueiros, D. Juarez-Romero, Optimum operating conditions for a water purification process integrated to a heat transformer with energy recycling using neural network inverse, Renew. Energ., 34 (2009) 1084–1091.
[8] Qazi, H. Fayaz, A. Wadi, R.G. Raj, N.A. Rahim, W.A. Khan, The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review, J. Clean Prod., 104 (2015) 1–12.
[9] D. Colorado, J.A. Hernández, Y. El Hamzaoui, A. Bassam, J. Siqueiros, J. Andaverde, Error propagation on COP prediction by artificial neural network in a water purification system integrated to an absorption heat transformer, Renew. Energ., 36 (2011) 1315–1322.
[10] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to linear regression analysis, 3rd edition, John Wiley and Sons Inc, 2001.
[11] S.K. Kachigan, Multivariable statistical analysis: A conceptual introduction, 2nd edition, Radius Press, 1991.
[12] H. Demuth, M. Beale, 1998. In: Natrick, M.A. (Ed.), Neural Network Toolbox for Matlab-User's Guide Version 3, The MathWorks Inc.