COBEM 2023 27TH INTERNATIONAL CONGRESS OF MECHANICAL ENGINEERING

ABCM Brazilian Society of Mechanical Sciences and Engineering
27th ABCM International Congress of Mechanical Engineering
December 04-08, 2023, Florianópolis-SC, Brazil

COB-2023-1505
ARTIFICIAL NEURAL NETWORKS FOR PREDICTION OF THERMOSYPHON PERFORMANCE

Thomas Siqueira Pereira
Yara de Souza Tadano
Hugo Valadares Siqueira
Thiago Antonini Alves
Federal University of Technology - Parana (UTFPR), Ponta Grossa/PR, 84.017-220, Brazil
thomaspereira@alunos.utfpr.edu.br, yaratadano@utfpr.edu.br, hugosiqueira@utfpr.edu.br, antonini@utfpr.edu.br

Abstract. *Thermosyphons are heat exchangers known for being versatile, easy to construct, and highly efficient for small temperature gradients. These components have highly complex equations with high error percentages. Because of this, achieving the necessary results is often complicated and highly time-consuming. Artificial Intelligence methods like the Artificial Neural Networks (ANN) are an excellent option to overcome this issue. ANN are computer algorithms based on the animal neural system that allows solving complex problems using only simple mathematical operations such as additions and multiplications. This algorithm uses known data of the proposed approach or similar cases to "learn" the system's behavior. For this experimental investigation, data was collected from different systems using thermosyphons and used to evaluate the capacity of proposed ANN to predict the thermal performance of thermosyphons. The experiment used thermosyphons made of copper tubes filled with distilled water as the working fluid. The heat source was simulated by a metal electric ribbon wrapped in the evaporator and heated by the Joule's Effect, and the cooling was made by air forced convection in the condenser. Three ANN algorithms were used for the evaluation of the proposed systems: the Multilayer Perceptron (MLP), the Radial Basis Function (RBF), and the Extreme Learning Machine (ELM). The considered inputs were: the slop, the filling ratio, and the dissipated power, and as outputs the thermal resistance of each thermosyphon. The results showed that all ANN successfully predicted the experimental values with less than 15% error and that the ELM has better results with less computing time and no more than 10% error.*

Keywords: *thermal performance, Extreme Learning Machine, phase change, Artificial Neural Network.*

1. INTRODUCTION

With the increase in energy demand resulting from today's society's increasing modernization and industrialization, rising equipment efficiency is indispensable in managing energy demand and building an increasingly sustainable electrical matrix. Energy loss in the form of heat, whether during transport, through inefficient devices, or carried by combustion gases, significantly impacts the efficiency of energy use on a global scale. Cullen and Allwood (2010) estimate that the energy loss related to heat loss in primary energy consumption, which has not yet undergone conversion, reaches about 63% of the energy consumed. Even a portion of this heat, if recovered, either by increasing efficiency or reuse, can positively influence the energy demands inherent to the sector responsible for the waste, reducing the monetary and environmental impact of the energy production increase needed to supply these losses.

The need to improve efficiency in using heat leads us to thermosyphons, passive devices capable of transferring large amounts of heat even with minor temperature differences (Reay et al., 2014). These devices are constructed from tubes, evacuated, filled with working fluid, and sealed. These evacuated tubes are pressure controlled, allowing the working fluid to change phases more easily (Antonini Alves et al., 2018). The operation of thermosyphons begins with the heat exchange from the working fluid to the hot source, which causes the working fluid to change to the gas phase and carry the heat toward the cold source (Mantelli, 2021). Then, the working fluid loses heat and changes to the liquid phase, restarting the cycle. Figure 1 shows the working diagram of a thermosyphon (Dimbarre et al., 2021).

Even with the area's significant development, different thermosyphon models are still studied and developed. For many cases, theoretical modeling with great complexity is needed, mainly due to the convective and phase change characteristics of the involved processes. The difficulty in modeling them leads to limitations in the viability of using these devices. One option to facilitate the development of thermosyphons is using Artificial Neural Networks (ANN) (Chen et al., 2010).

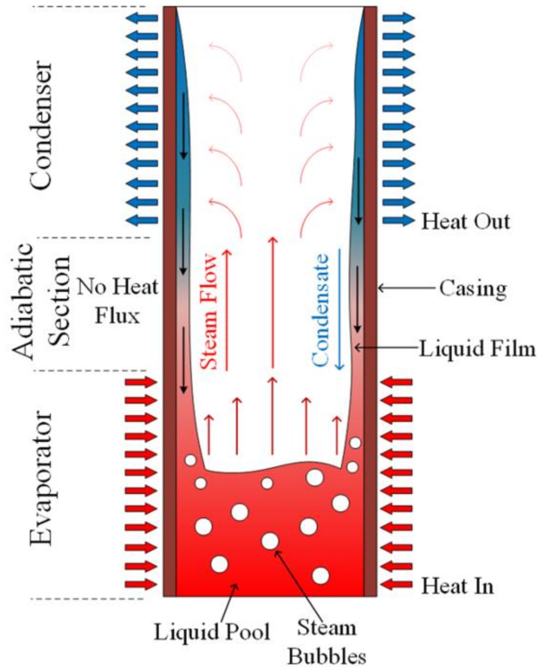

Figure 1. Schematic diagram of a thermosyphon

With the constant increase in the usable computational capacity in recent decades, it was possible, in parallel, to visualize a constant growth in the use and study of methods that require high processing power. Among these, Artificial Neural Networks stand out for their great range of applications and ease of implementation in complex problems (Araujo et al., 2020).

ANN are computational models based on the nervous system of animals, built to predict different behaviors of complex problems without the need for deep knowledge about the theory behind it. The operation of the ANN is based on using previously obtained experimental data and thus "teaching" the ANN about the problem's behavior. These algorithms are formed by connecting small modules, usually called neurons, distributed in different layers, as shown in Figure 2. These neurons simulate the behavior of organic neurons, receiving input data (on the left in Figure 2) and performing simple mathematical operations to obtain new values that will be transmitted to the following layers until finally obtaining results from the output layer (on the right in Figure 2). For the network to provide adequate results, it must first be trained, which can be done with several different methods, but usually involves an iterative algorithm that starts the network randomly and uses known data to modify the weights of the network at each iteration to approach the desired result (Haykin, 2008).

The Multilayer Perceptron (MLP) can be defined as an ANN of multiple feedforward layers (where there is no recurrence or return of information) with one or more hidden or intermediate layers in addition to the input and output layers. The MLP is one of the most used architectures of neural networks capable of working with non-linear problems and having universal approximation. Their training process uses supervised iterative methods. Among these, most known and used is backpropagation, where the resulting error at each iteration of the network is evaluated and used to modify the weights of each layer.

The Radial Base Network (RBF) is also a feedforward type of network, with only two layers since the input data is injected directly into the second layer without multiplication by weights. This model differs from the MLP mainly by its activation function, which is usually a radial basis function like the Gaussian one. For this same reason, training a RBF involves defining function parameters, such as center position and dispersion. For this, clustering algorithms are used, such as K-means.

Extreme Learning Machines (ELM) are ANN with feedforward architecture and structures like MLP but having a very different training method. During the training process, the values of the hidden layer weights are set randomly and not modified. Instead, it uses an analytical method to train only the hidden layer without using iterative methods, thus reducing the training time in many cases (Huang et al., 2006).

In this work, a study was carried out to estimate the thermal performance of thermosyphons using three different Artificial Neural Networks (MLP, RBF, and ELM). As input, results from Krambeck et al. (2019) were used. The author tested thermosyphons with different characteristics. The inputs were: slope, filling ratio, and dissipated power, and the output is the thermal resistance.

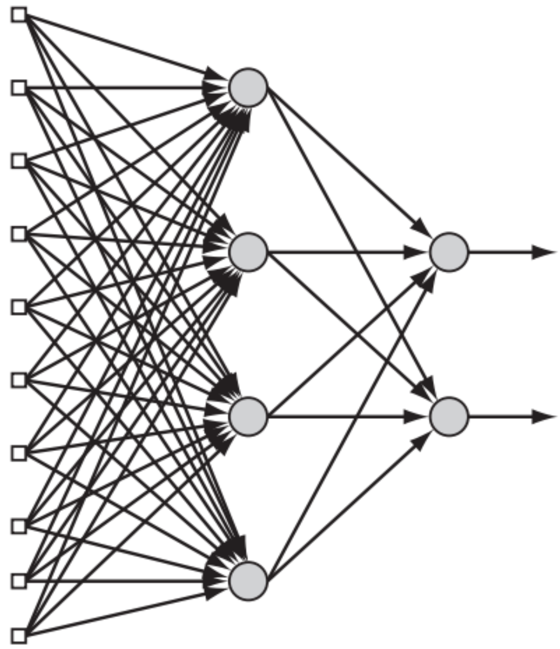

Figure 2. Representation of an ANN

2. EXPERIMENTAL DATA

The database was obtained from results by Krambeck et al. (2019) through an experimental investigation involving copper thermosyphons containing distilled water as the working fluid. The main characteristics of the thermosyphon are shown in Table 1. During the experiments, the temperature was maintained at approximately 20°C±1°C. The heat was generated from the Joule's Effect using a resistive ribbon to heat the evaporator, while air forced convection, generated by the fan, was used for cooling. Several thermocouples, represented by T_{cond}, T_{adiab}, and T_{evap}, were used to obtain the temperature at various points during the operation. The experiment diagram can be seen in Figure 3.

Table 1. Main characteristics of the used thermosyphon

Feature	Value				
Inner diameter [mm]	7.75				
Outer diameter [mm]	9.45				
Evaporator size [mm]	80				
Adiabatic Section size [mm]	20				
Condenser size [mm]	100				
Working fluid	Distilled water				
Filling ratio [%]	20	40	60	80	100
Working fluid volume [mL]	0.88	1.51	2.26	3.02	3.77

The tests were carried out with powers from 5 to 45W with steps of 5W, avoiding the critical temperature of 150ºC. Each power was maintained until the system reached a steady state so that the temperatures were taken. Two working slopes of the thermosyphons were also used: 45 and 90º with the horizontal.

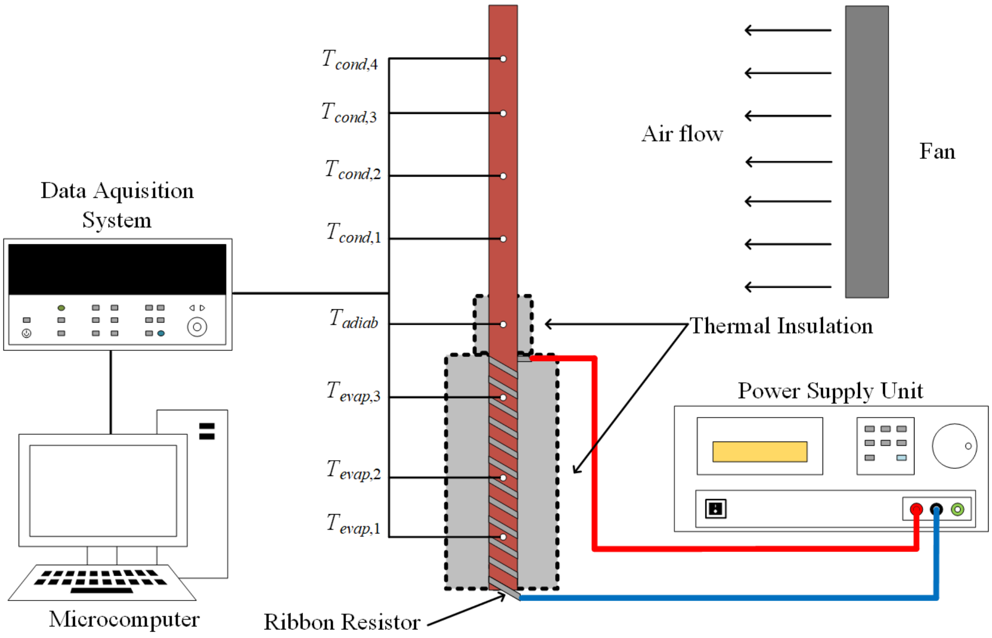

Figure 3. Diagram of the experimental procedure for the data acquisition

3. EXPERIMENTAL SETUP

The parameter chosen to evaluate the thermal performance of thermosyphons is the thermal resistance, R_{th}, which can be expressed by:

$$R_{th} = \frac{\left(T_{evap} - T_{cond}\right)}{q_{in}}, \tag{1}$$

where, T_{evap} and T_{cond} are the average temperatures of the evaporator and condenser, respectively, and q_{in} is the dissipated power. Thus, it was possible to determine the thermal resistance for each tested configuration.

The ANN were trained for a wide range of values for the number of neurons to obtain the configuration with the lowest Mean Squared Error (MSE) against the test data.

3.1 Multilayer Perceptron

For training the Multilayer Perceptron network, the data set was separated into three. The training group containing about 70% of the data, is the one effectively used in the network training cycles. The validation group contains 15% of the data, and is used as a stop criterion to prevent overtraining. The test group contains 15% of the data where the trained neural network is evaluated for unknown data. Initial weights were randomly obtained, and no moment term was used. Two steps were used to train this network.

The training started by varying the number of neurons to select the combinations that minimize the MSE between the network output and the expected result. The step used was 0.004. With this, the training process was carried out with a maximum of 50,000 cycles. At each iteration, a stopping criterion was tested. It occurred by comparing the squared error obtained when applying the validation data values, the network in the current state, and comparing them with the expected results. If the stopping criterion is not reached in the current cycle, the network weight values are updated using the backpropagation algorithm with the results obtained from the training data. At each iteration, it was also evaluated whether the current validation MSE was the smallest. If so, this group of weights was saved and would only be replaced if an even smaller error was found, ensuring that the best group of weights was saved even if the iterations continued. At the end of the search process, these saved weights were used for the test step. It consists of applying the test data, that are not yet "known" by the network, to the trained ANN and evaluate the resulting MSE. For each number of neurons, 30 tests were performed, and the average value between them was evaluated. The combinations with the lowest mean test errors were selected.

3.2 Radial Base Function

The RBF network has a strong relationship between the number of neurons in the hidden layer and the quality of the results. In some cases, the optimal number of neurons can be as high as the amount of input data available. For this reason, a higher number of neurons to be tested was used compared to the MLP. The dataset is divided in training, validation, and testing, as in the case of MLP. The activation function was the Gaussian Function.

For each network training, the centers of the Gaussian Functions must be found. The RBF is highly dependent on the initial condition that is randomly found, so each new test, even using the same data, can result in divergent results.

The second stage of RBF training follows the same parameters seen in the training of the MLP network. The number of neurons varied between 3 and 200, and the learning step used was 0.004, with a maximum number of 50,000 cycles and similar procedures during training. Thirty tests were also performed, and the number of neurons that resulted in the lowest mean squared error was chosen.

3.3 Extreme Learning Machine

The Extreme Learning Machine works with an analytical approach to the intermediate layer training, so there are no iterations. This reduces training time and allows for more significant variation in training parameters.

The ELM was trained by varying the number of neurons to select those that minimize the MSE between the network output and the expected result. Unlike the MLP and the RBF, the ELM does not use validation. Therefore, the dataset was divided into two groups. The training group contains about 85% of the data, and was effectively used in the network training. The test group containing 15% of the data was used to evaluate the network error at the end of training. Since the ELM is less time consuming it is possible to use a wider range of values for the number of neurons in the hidden layer. It was varied between 3 and 200.

At the end of the calculations, the best combinations of the network were selected, that is, the values that minimized the MSE. Thirty tests were performed for each combination to decrease the influence of the initial condition on the results, and the average value between them was evaluated.

A summary of the training can be seen in Table 2.

Table 2. Values used for training

Model	**No. of neurons**	**Step**	**Activation Function**	
			Hidden	**Output**
MLP	3 - 100	0.004	Logistic	Linear
RBF	3 - 200	0.004	Gauss Function	Linear
ELM	3 - 200	N/A	Logistic	Linear

4. RESULTS

For the evaluation of the results, the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE), and the Square Root of the Mean Squared Error (RMSE) described in Eqs. (2), (3), and (4), respectively, were considered.

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|d_t - y_t\right| , \tag{2}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{d_t - y_t}{d_t}\right|100 , \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(d_t - y_t\right)^2} , \tag{4}$$

where, d_t represents the experimental output of the database, y_t is the output of the ANN and N represents the number of data used.

Table 3 presents the results obtained for each model using the evaluated errors. The *NN* value represents the number of neurons in the hidden layer for the best results. Final values were obtained from the average of errors among 30 independent tests.

Table 3. Errors found for the best configuration of each ANN

Model	***NN***	**Hidden Layer Function**	**No. of Hidden Layers**	***MAE***	***RMSE***	***MAPE* [%]**
MLP	9	Logistic	1	0.214	0.258	8.24
RBF	90	Gauss Function	1	0.344	0.438	10.26
ELM	25	Logistic	1	0.205	0.248	6.71

Figure 4 presents the boxplot of the MAPE calculated for 30 simulations carried out for each of the three used ANN. As can be seen, the network with the lowest dispersion and MAPE error is the ELM which also has the lowest MAE and RMSE found in all tests. The MLP also presents results with high precision and low variability, as expected from an algorithm highly used and known for being remarkably consistent even though it does not always reach the best results. The results also demonstrate that although the RBF achieves good results for the problem, it presents a large variability between results since it depends significantly on the initial state randomly generated at each initialization.

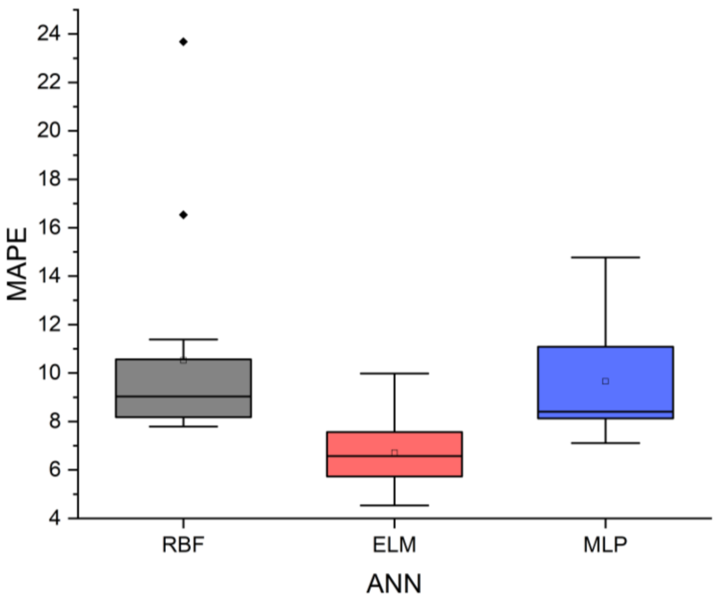

Figure 4. MAPE error boxplot for 30 iterations for each ANN

Figure 5 compares the thermal resistances of the thermosyphons obtained experimentally from the test database and the values calculated in a simulation using each of the evaluated networks. It is possible to see that not only the mean error values (Figure 4) show consistent results, but also the individual values obtained by the simulation demonstrate proximity to the experimental values used, with almost all of the test group within the area which represents an error of less than 20%. Once again, the values obtained with the ELM showed superiority, with considerably less dispersion, thus having concentrated results much closer to the ideal line.

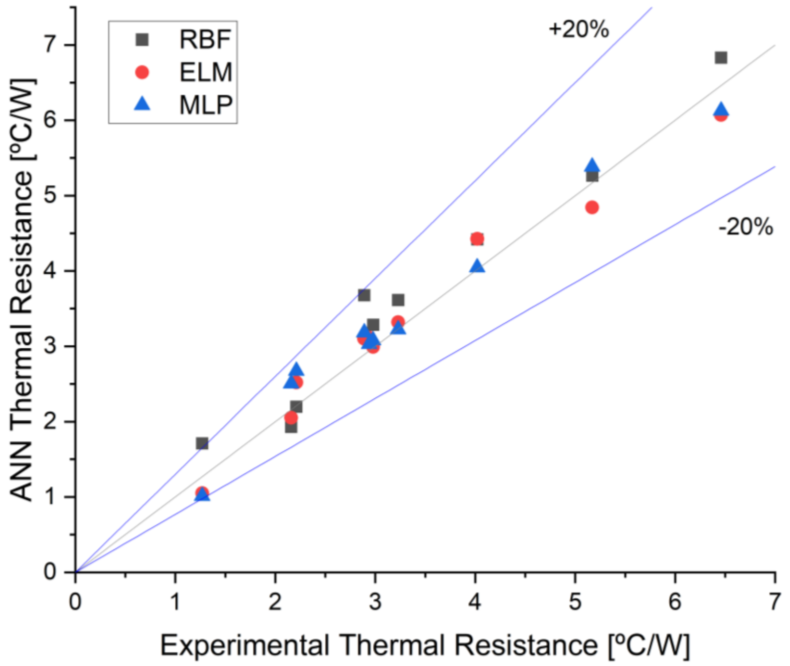

Figure 5. Experimental thermal resistance versus thermal resistance for each ANN

5. CONCLUSIONS

The definition of parameters for constructing a thermosyphon can be challenging, and often, rework and empirical methods are needed to make the desired thermosyphon. That said, developing methods such as Artificial Neural Networks are very important for the area, given its excellent capacity to simulate problems, often with complex equations, using only known experimental data. The results of this study show that all tested networks could successfully emulate the thermosyphon behavior and predict with low errors the thermal resistance of the thermosyphon tested under different operational conditions. It was also shown that the ELM network obtained the best results, having smaller errors and less dispersion of results. These results are essential for showing that this type of approximation can be beneficial for the area of thermosyphons and heat exchangers and show that new studies should be carried out.

6. ACKNOWLEDGEMENTS

We acknowledge the Capes, the CNPq, the DIRPPG-PG/UTFPR, the PPGEM-CP/PG/UTFPR, and the DAMEC-PG/UTFPR.

7. REFERENCES

Antonini Alves, T., Krambeck, L., Santos, P. H. D., 2018. *Heat Pipe and Thermosyphon for Thermal Management of Thermoelectric Cooling*. In: Aranguren, P. (Org.). *Bringing Thermoelectricity into Reality*. Intech, London/UK.

Araujo, L. N., Belotti, J. T., Antonini Alves, T., Tadano, Y. S., Siqueira, H., 2020, Ensemble method based on Artificial Neural Networks to estimate air pollution health risks, Environmental Modelling and Software, Vol. 123, 104567.

Chen, R. H., Su, G. H., Qiu, S. Z., Fukuda, K., 2010, Prediction of CHF in concentric-tube open thermosyphon using Artificial Neural Network and Genetic Algorithm, Heat and Mass Transfer, Vol. 46, No. 3, pp. 345-353.

Cullen, J. M., Allwood, J. M., 2010, Theoretical efficiency limits for energy conversion devices, *Energy*, Vol. 35, pp. 2059-2069.

Dimbarre, V. V., Bartmeyer, G. A., Machado, P. L. O., Turchiello, R. F., Santos, P. H. D., and Antonini Alves, T., 2021. "Experimental evaluation of thermal performance and internal pressure of thermosyphons using graphene oxide nanofluid". In *Proceedings of the 26th International Congress of Mechanical Engineering - COBEM 2021*. Florianópolis, Brazil.

Haykin, S., 2008. *Neural Networks and Learning Machines*. Prentice Hall, USA.

Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006, Extreme Learning Machine: Theory and applications, *Neurocomputing*, Vol. 70, pp. 489-501.

Krambeck, L., Nishida, F. B., Aguiar, V. M., Santos, P. H. D., Antonini Alves, T., 2019, Thermal performance evaluation of different passive devices for electronics cooling, *Thermal Science*, Vol. 23(2B), pp. 1151-1160.

Mantelli, M. B. H., 2021. *Thermosyphons and Heat Pipes: Theory and Applications*. Springer International Publishing, Switzerland.

Reay, D. A., Kew, P. A., McGlen, R. J., 2014. *Heat Pipe: Theory, Design and Applications*. Butterworth-Heinemann, Netherlands.

8. RESPONSIBILITY NOTICE

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