25th ABCM International Congress of Mechanical Engineering
October 20-25, 2019, Uberlândia, MG, Brazil

COB-2019-1821
REGRESSION MODELS AND NEURAL NETWORKS IN ESTIMATION OF THE WORKING FLUID TEMPERATURE OF PARABOLIC SOLAR CONCENTRATORS

Lino Wagner Castelo Branco Portela
Mechanical Engineering Department, Federal University of Ceará – UFC, Campus of Pici, CEP 60020-181, Fortaleza – CE, Brazil
linowcbp@gmail.com

Erilson de Sousa Barbosa
Federal Institute of Education, Science and Technology of Ceará - IFCE, CEP 62500-000, Itapipoca – CE, Brazil
erillson@yahoo.com.br

Ana Fabíola Leite Almeida
Mechanical Engineering Department, Federal University of Ceará – UFC, Campus of Pici, CEP 60020-181, Fortaleza – CE, Brazil
anfaleal@yahoo.com.br

Abstract. *This work proposes the verification of the application of the mathematical models Simple Linear Regression, Multiple Linear Regression and Artificial Neural Networks (ANNs) in the prediction of the working fluid outlet temperature a parabolic solar concentrator (CSP) with tracking system. The concentrator is aided by sensors in the measurement of temperature and solar radiation. With the R programming language, it was possible to analyze the collected data and perform statistical calculations, where the Artificial Neural Networks model presented 5.23 in the Root Mean Square Error (RMSE) and 78.89 % in the R^2 statistic.*

Keywords: *Simple Linear Regression, Multiple Linear Regression, Artificial Neural Networks, Parabolic Solar Concentrator, RMSE.*

1. INTRODUCTION

In Brazil, in recent years, there have been major investments for the energy transition between conventional and renewable sources. Among them, the creation of the Solar Park Nova Olinda, in the state of Piauí, whose capacity of electric power generation characterizes it with the largest park in Latin America (Vieira, 2018). Despite this, according to the National Electric Energy Agency (ANEEL), the country's main source of energy still comes from hydroelectric plants (60.7 %). As well as fossil fuels, the most used are natural gas (7.7 %), petroleum (5.8 %) and coal (2.2 %).

The commercial value of renewable sources, such as solar energy, has been shown in recession, favoring its development. While consolidated industries like the declining oil, due to the high volatility of its price. Over the next decades, clean sources - biomass, wind, solar - tend to represent a large part of the Brazilian energy matrix (Schaeffer, 2016).

In this context, and considering the great solar energy potential of Brazil, the parabolic solar concentrators (CSP) can represent a considerable contribution in the use of this type of energy. These concentrators are collectors of solar radiation whose operation basically consists of heating a fluid, usually water or thermal oil, which passes inside a tube located in the focus of the parabolic trough (Kalogirou, 2009). They have numerous applications of solar energy, ranging from a simple domestic water heating for personal use to a broad chain of power generation for industrial processes.

Given that the Northeast region is the one with the highest rates of direct solar irradiation and that prolonged periods of water scarcity are common, CSP technology can also be used for water desalination. However, for this to occur effectively, it is necessary for the concentrators to function as efficiently as possible (Pereira, 2017).

The development of models for testing the temperature of the working fluid can also contribute to the increase of the production of these systems, allowing to monitor their efficiency.

From this, this article aims to verify the Simple Linear Regression, Multiple Linear Regression and Artificial Neural Networks (ANNs) methods in the analysis of temperature prediction of working fluid output. Observing the influence of variables such as time, solar radiation, working fluid inlet temperatures and parabolic trough tube. The validation of these models is done by means of data obtained experimentally from a parabolic solar concentrator with tracking system.

2. THEORY ANALYSIS OF THE METHODS USED

The linear regression model can be considered one of the best known and applied. It consists of a dependent variable that is related to other independent variables, individually or collectively (Medeiros, 2009). When related to only one independent variable, the regression is interpreted as simple. It is represented mathematically according to equation (1):

$$Y = \beta_0 + \beta_1 X + \epsilon \qquad (1)$$

Where: $\boldsymbol{Y}$ is the dependent variable to be predicted; $\boldsymbol{X}$ is the independent variable; $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$ are the constants that represent the terms of intercept and slope in the linear model; $\boldsymbol{\varepsilon}$ the random error term.

On the other hand, the multiple regression model is understood as the extension of simple regression and is considered a better interpretable approach when relating multiple variables. Assuming that there are n variables, it can take the form of equation (2):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon \qquad (2)$$

Where: $\boldsymbol{Y}$ is the dependent variable to be predicted; $\boldsymbol{X_1}$, $\boldsymbol{X_2}$ and $\boldsymbol{X_n}$ are the independent variables; $\boldsymbol{\beta_0}$, $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$, and $\boldsymbol{\beta_n}$ the constants representing the intercept and slope terms of the model; $\boldsymbol{\varepsilon}$ the random error term.

The neural network model, multilayer perceptron, performs computational simulation based on the behavior of neurons of the human brain, simulating the patterns of the brain, in order to produce results from the perceived actions. This model is based on learning a set of training data. Among the characteristics of the neural network, the main one is the ability to learn by example, it creates the input and output relationship by a learning algorithm (Bocco, 2010).

Regression models and neural networks are evaluated, as their quality, according to two related quantities: the Root Mean Square Error (RMSE) and the R^2 statistic. The RMSE can be understood as an average estimate of how much the forecast deviates from the true regression line. The R^2 statistic is a measure that quantifies the linear relationship between the variable to be predicted and the independent one, being represented by a number that is between 0 to 1. When calculating these quantities, it is possible to quantify the influence of each variable analyzed in the system (James, 2013).

3. PROPOSED MATERIALS AND METHODOLOGY

The values used in this research were acquired from a parabolic solar concentrator aided by a tracking system in operation in the city of Teresina, in the state of Piauí. It has a structure in carbon steel, and its reflective parabolic plate in mirrored stainless steel (thickness 0.8 mm). At the center of the gutter, is a copper tube with a diameter of 5/8".The opening of the concentrator is 1800 mm in length and 900 mm in width.Your tracking system is controlled by a microcontroller Arduino Mega 2560 R3, coupled with a module of two NORPS-12 LDR (Light Depedent Resistors) light sensors. The Fig. 1 shows the image of the concentrator.

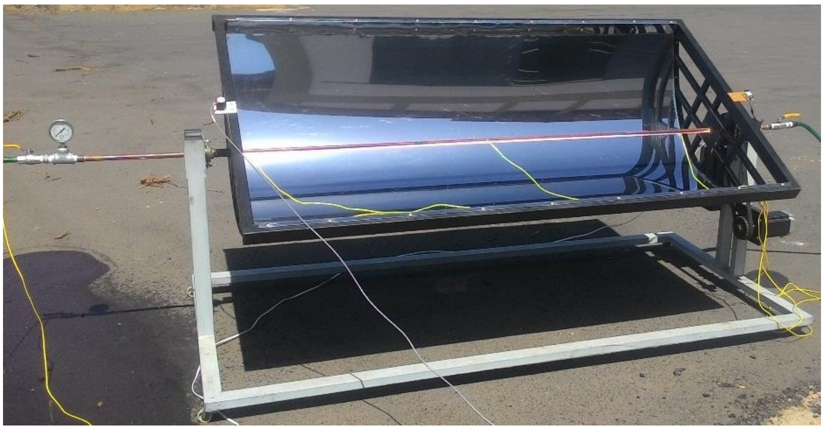

Figure 1. Solar Parabolic Concentrator.

As for the sensors, five K-type thermocouple sensors were used to monitor the temperature in the outer wall of the copper tube and the inlet and outlet temperature of the water used in the system. The first measuring the inlet temperature of the water supplying the system. Another three were attached along the tube. By reading these, the average temperature along the length of the tube was subsequently calculated. And the last one in the system output measuring the final temperature. On the other hand, the acquisition of solar radiation data (W/m²) was performed using a KIMO™ pyrometer sensor - model CR 110. The Fig. 2 illustrates a location of these sensors.

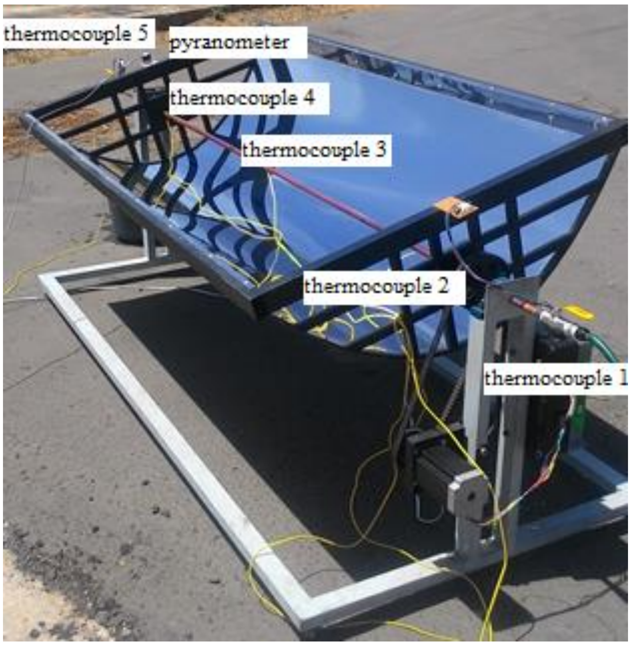

Figure 2. Location of thermocouple sensors and pyranometer.

The data considered were recorded, monitored and stored by means of a data logger. In his system it was possible to characterize the readings of the thermocouples and the pyranometer. The reading was performed at intervals of 10 seconds.

As for the statistical analysis, the variables time, temperature and solar radiation were grouped and analyzed with the aid of RStudio software on a computer with Windows 10 Operating System. RStudio uses the R language, which is applied to statistical and graphical calculations. In it, the RMSE and R^2 statistics were calculated, as well as the graphical plotting of the relationships between the predictors and the temperature to be predicted.

After data acquisition, so that the Simple Linear Regression, Multiple Linear Regression and Artificial Neural Networks (ANNs) methods could be applied, some preliminary procedures were followed. Initially, from the entire data set, 60 % of them were applied in the algorithm training process. After that, 40 % remaining was used to be tested. Finally, to analyze the accuracy of the models, the RMSE and R^2 statistics were calculated. The Fig. 3 shows a flowchart illustrating this process.

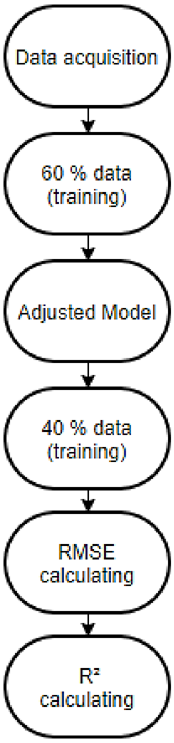

Figure 3. The flowchart representing the process.

4. RESULTS

After collecting data and inserting them into the RStudio platform, it was possible to verify the relationship between solar radiation and other variables. The mean temperature in the tube was the one that had the greatest influence on the forecast of the exit temperature. Table 1 shows the correlation matrix of the variables.

Table 1. Correlation Matrix Between Variables

Variables	Hor	Hr	Min	Seg	T0	T1	T2	T3	Tm	Ts	Rds
Hor	1.0	0.98	0.02	0.05	0.36	0.15	0.28	0.13	0.23	0.42	-0.24
Hr	0.98	1.0	-0.15	0.05	0.34	0.16	0.28	0.16	0.24	0.41	-0.22
Min	0.02	-0.16	1.0	-0.01	0.1	-0.02	-0.02	-0.13	-0.07	0.04	-0.05
Seg	0.05	0.05	-0.01	1.0	0.03	0.02	0.02	0.04	0.03	0.04	0.01
T0	0.4	0.35	0.05	0.03	1.0	0.31	0.42	0.32	0.42	0.24	-0.01
T1	0.15	0.16	-0.02	0.02	0.31	1.0	0.87	0.37	0.88	0.64	0.33
T2	0.28	0.28	-0.02	0.02	0.42	0.87	1.0	0.41	0.91	0.65	0.32
T3	0.13	0.16	-0.13	0.04	0.32	0.37	0.41	1.0	0.71	0.38	0.28
Tm	0.23	0.24	-0.07	0.03	0.42	0.88	0.91	0.71	1.0	0.67	0.37
Ts	0.42	0.41	0.04	0.04	0.24	0.64	0.65	0.38	0.67	1.0	0.26
Rds	-0.23	-0.22	-0.05	0.01	-0.01	0.33	0.32	0.28	0.38	0.26	1.0

Hor: time; Hr: hours; Min: minutes; Seg: seconds; T0: initial water temperature; T1 temperature of the first thermocouple in the tube; T2: temperature of the second thermocouple in the tube; T3: temperature of the third thermocouple in the tube; Tm: mean temperature in the tube; Ts: water outlet temperature; Rds: solar radiation.

The results of the modeling process allowed the calculation of the statistical parameters RMSE and R^2 between the observed and estimated values of the exit temperature of the water. Table 2 shows the result of these parameters.

Table 2. Root of the Mean Square Error (RMSE) and R^2 of the different models

MODEL	RMSE	R^2
Simple Linear Regression	8.27	0.4451
Multiple Linear Regression	6.97	0.5361
Artificial Neural Networks (ANNs)	5.23	0.7889

The Figures 3, 4 and 5 show, respectively, the graphical representation of simple, multiple and neural network linear regression.

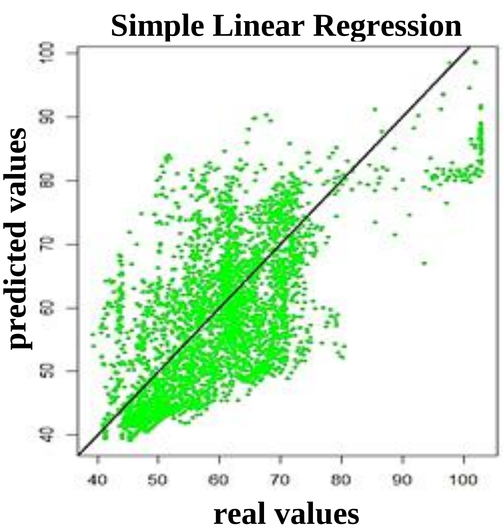

Figure 4. Model Simple Linear Regression.

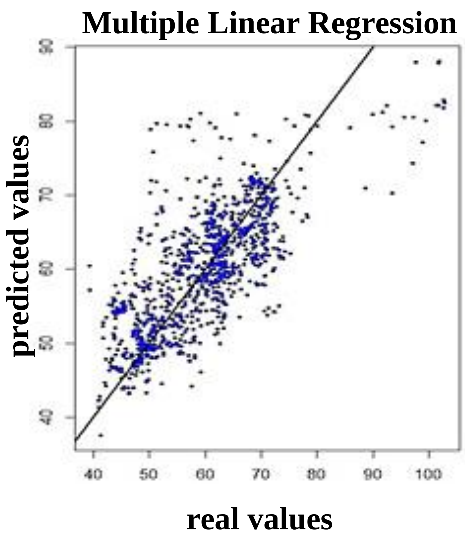

Figure 5. Model Multiple Linear Regression.

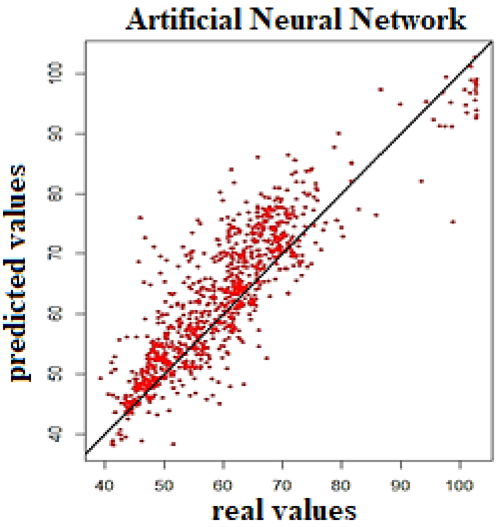

Figure 6. Model Artificial Neural Network.

In Fig. 4, it is possible to notice that in the application of the Simple Linear Regression model there is a large data dispersion in relation to the adjustment line. This in turn is demonstrated by its high RMSE value, being the largest among the other models presented, which characterizes the largest difference between the predicted values in relation to the real value. Similarly, in Fig. 5, with the application of the Multiple Linear Regression model, there is still considerable data dispersion in relation to the fit line, not as much as the model in Fig. 4, but still considered a high RMSE.

The Artificial Neural Networks model, in Fig. 6, was modeled with an input layer of four neurons and a hidden layer of 10 neurons, a parameter used by Bocco et al. (2010), who also achieved better results using neural network modeling. In this model, the RMSE value was the smallest and consequently the largest R^2, evidenced by the proximity of the data to the adjustment line, which characterizes the higher precision of the predicted values in relation to the real values. Therefore, being considered the one that presented the best representativeness for data prediction.

5. CONCLUSIONS

In this work, an approach was presented on the application of mathematical models for the prediction of the temperature of the exit of parabolic solar concentrator fluid. It was possible to notice that the simple and multiple linear regression models presented a small difference in the prediction of the results. Despite the small interval between the data variation (10 seconds), the model using neural network was more efficient in the prediction.

6. ACKNOWLEDGEMENTS

The authors thank the Coordination for the Improvement of Higher Education Personnel (CAPES) for the financial support in the form of a scholarship and the National Council for Scientific and Technological Development (CNPq) for the promotion of the research through public notice Nº 18/2013 MCTI / CNPq / SPM-PR / Petrobras - Girls and Young People Making Exact Sciences, Engineering and Computing.

7. REFERENCES

O.L. Vieira, F.J.R. Silva, L.A. Machado, A.N.F. Correia, C.T. Vitoriano, A.H. Oliveira, R.P.D. Franklin, J.C.F. Menezes, M.F. Apolinário, J.W.F. Nascimento and P.C.M. Carvalho. Comparison of Photovoltaic Plant Performance Estimation Models: A Case Study for Fortaleza, Brazil. IEEE Latin America Trans. Vol. 16, No. 7, pp. 1898-1906, 2018.

BRASIL. Agência Nacional de Energia Elétrica. Matriz de Energia Elétrica. Disponível em: <http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/OperacaoCapacidadeBrasil.cfm>. Acesso em: 20 out. 2018.

R. Schaeffer. [R]evolução Energética: rumo a um Brasil com 100% de energias limpas e renováveis. Disponível em: <iei-brasil.org/wp-content/uploads/2017/08/Revolução-Energética-2016.-Greenpeace-Brasil.pdf>. Acesso em: 28 set. 2018.

S. A. Kalogirou. Solar energy engineering: processes and systems. EUA: Elsevier, Academic Press, 2009.

E. Pereira. et al. Instituto Nacional de Pesquisas Espaciais (INPE). Atlas Brasileiro de Energia Solar. Disponível em: <http://ftp.cptec.inpe.br/labren/publ/livros/Atlas_Brasileiro_Energia_Solar_2a_Edicao.pdf>. Acesso em: 20 set. 2018.

F. S. B Medeiros and R. C. Bianchi. A aplicação do método regressão linear simples na demanda de produtos sazonais: um estudo de caso. Disciplinarum Scientia. Vol. 5, No. 1, pp. 35-53, 2009.

M. Bocco, E. Willington and M. Arias. Comparison of Regression and Neural Networks Models to Estimate Solar Radiation. Chilean journal of agricultural research. Vol. 70, No. 3, pp. 428-435, 2010.

G. James, D. Witten and T. Hastie. An Introduction to Statistical Learning. New York: Springer, 2013.

ARDUINO. Arduino Mega 2560 Rev3. Disponível em: <https://store.arduino.cc/usa/arduino-mega-2560.rev3>. Acesso em: 20. out. 2018.

KIMO. CR 110: Capteur/transmetteur de rayonnement solaire. Disponível em: <http://www.kimo.fr/fr/capteur/cr-110>. Acesso em: 12 set. 2018.

RStudio. Disponível em: < https://www.rstudio.com/products/rstudio/download/>. Acesso em: 22. out. 2018.

R: The R Project for Statistical Computing. Disponível em: <https://www.r-project.org>. Acesso em: 15. set. 2018.

8. RESPONSIBILITY NOTICE

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