

# Modeling of Rainfall-Runoff by Artificial Neural Network for Micro Hydro Power Plant: A Case Study in Cameroon

Kengne Signe Elie Bertrand<sup>(1&2)\*</sup>, O. Hamandjoda<sup>1</sup>, Fanyep Nana Antoine<sup>2</sup>,  
Gubong Takam Charles Bertrand<sup>2</sup>

Laboratory of Water, Energy and Environment, National Advanced School of Engineering of University of Yaoundé I,  
P.O Box 8390 Yaoundé, and Cameroon<sup>1</sup>

Laboratory of Energy Research, Institute for Geological and Mining Research, P.O Box 4110 Yaoundé, Cameroon<sup>2</sup>

**ABSTRACT:** In Cameroon, there are lots of small waterfalls suitable for the development of micro-hydroelectric power plant (MHPP) in that country. But, the development of these MHPP faces a deficit in terms of data on small watersheds which are not gauged for the majority. Since rainfall data are more available, their exploitation to cover the deficit of hydrological data is possible. The objective of this work is to develop a rainfall-runoff model to simulate and produce the runoff curves ranked (classified) necessary for the design of MHPP. The watershed test selected is the Mefou Nkol-Bisson watershed. Then, we have designed a model using artificial neural networks (ANN) with the method of back-propagation of the quadratic error to allow the prediction of flow rates on small watersheds from the rainfall data, for hydroelectric production. Our model has a very good performance because the NASH criteria and coefficient of correlation calculated are ( $R^2 = 0.983$  and NASH = 0.912 in calibration), ( $R^2 = 0.842$  and NASH = 0.875 in validation) and (NASH = 0.9 for an application). We believe that the application of this model on others small watersheds under different conditions can present irregularities. But for the MHPP design needs, the requirement is to make a more interesting hydrological forecasting than the method often used in similar projects (for our environment, it is the mean annual flow that is often used).

**KEYWORDS:** Rainfall-runoff modeling, ANN, micro hydroelectric power plant.

## I. INTRODUCTION

Considering the emissions of greenhouse gases in the choice of energy sources, systems of new and renewable energies are prioritized because of their good contributions to the fight against climate change. The access rate to electricity in Cameroon was less than 50% in 2012, (Lighting Africa, 2012). To improve that access rate to electricity; the State of Cameroon also promotes the development of renewable energies. Thus, any work in order to make a contribution to its ease is beneficial. The hydraulic energy, so the resource is very important in Cameroon, is one of the most sought renewable sources for its multiple benefits. But in that country, hydroelectric energy is currently produced on large watersheds which are the only gauged. The small watersheds which are not gauged could also be exploited by small isolated communities for the implementation of MHPP.

To exploit sustainably MHPP, hydrological data for long periods are most appreciated and requested. This data almost don't exist in Cameroon for small watersheds. But the rainfall data are the most available (M.Tsalefac et al, 2003).

Experience shows that an insufficient knowledge of hydrology is a frequent cause of failure in the implementation of MHPP. It is therefore, recommended to seriously reflect on the rates at the time of the project feasibility the study (G. I. Economiques, ORSTOM – EDF, HYDRO CONSULT, 1992). Sometime in Cameroon, the flow rate were used for the design of MHPP is the annual average runoff of the river selected. Also globally, feasibility studies were not done

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 8, August 2017

properly because of the lack of data and local skills. The consequence is that, almost all MHPP setting up by local peoples cannot be used effectively. Then, they abandon very early their equipment because of dysfunction.

There are lots of rainfall-runoff models in the world. These models vary strongly from a region to another and, there is not yet a particular model adapted to small watersheds in Cameroon. We would like to start something in that area.

The main objective of this work is to develop a hydrological rainfall-runoff model for restoring the runoff curves classified from rainfall data for a contribution of to the perspective of hydroelectric production. The method choice is based on artificial neural network (ANN). We will first present the design method of neural networks, followed by an application on the small catchment basin test Mefou to Nkol-bisson in Yaoundé in the Centre Region in Cameroon, from the available data. We will use the back propagation algorithm because of their efficiencies which has been demonstrated. The capability of the Artificial Neural Network with different input dimensions have been attempted and demonstrated especially in the simulation of runoff (M.P. Rajurkar et al, 2004).

## II. MATERIAL AND METHOD

The rainfall-runoff modeling in our case concern more small watersheds in the objective to give a contribution to the design of MHPP. In this approach, the daily evapotranspiration losses are not taken into account because only the raw rainfall data are available. Thus, the model assumes that all measured rain is converted to output. (Papazafiriou, 1976), (Kachroo et al. 1992), and (Labat et al. 2000) advocated that simplification and indicate that it does not alter the model results significantly.

### a. The artificial neural network

An artificial neural network employs a mathematical simulation approach, which adopts a biological system in order to process the acquire output(s) after the network has been trained properly for pattern recognition. The main theme of ANN model is, it considers the brain as a parallel computational device for various computational tasks that were performed relatively poorly by traditional serial computers. The neural network structure in the present study possessed adopts a three-layer learning network consisting of an input layer, a hidden layer and an output layer consisting of output variable(s) as shown in Figure I.

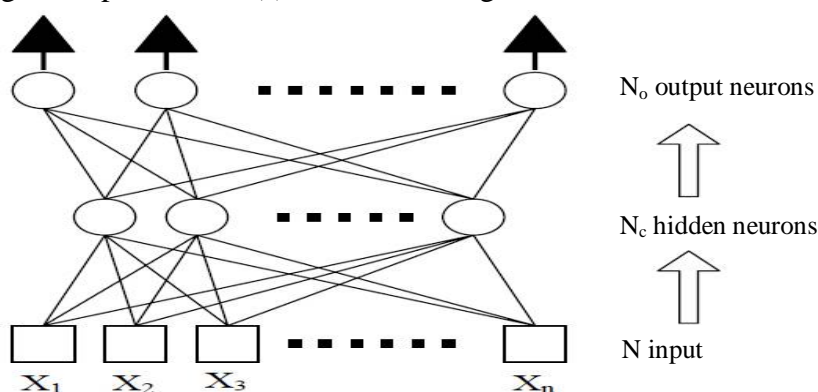


Figure I: Architecture of the Neural Network Model used in this Study

The advantage of using the RNA in hydrology is its ability to reflect the inherent non-linearity in the rainfall-runoff relationship without entering to the 'black box' process. This is an interesting property to consider only the meteorological parameters, especially as the hydrogeological parameters are almost nonexistent for our case.

Neural models have a wide variety of architecture. But the most network type used for modeling hydrological phenomena is the Multilayer Perceptron (MLP). The inputs of each neuron are the variables on which operates the neuron, and the output is the value of the activation function  $f$  often sigmoid. An example of activation function is usually:

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 8, August 2017

$$f(\mathbf{x}) = \frac{1}{1 + \exp(\mathbf{x})} \quad (1)$$

The setting of the function is frequently done by parameters that are attached to the inputs of the neuron: the output of the neuron is a nonlinear function of a combination of inputs  $\{x_i\}$  weighted parameters  $\{w_i\}$ , which are then often referred to as weight name. The most frequently used potential is the weighted sum, to which is added a constant  $W_0$  term or "bias". It is often given by the formula (2):

$$y = \text{th} \left[ W_0 + \sum_{i=1}^{n-1} W_i X_i \right] \quad (2)$$

The parameters are attached to the non-linearity of the neuron. In architecture layered networks, neurons of the same layer are not connected together. More connections between two layers of non-consecutive neurons are eliminated. Such architecture is historically used, mainly because of its relevance classification. The output of the neural network is then represented by the formula (3):

$$y = f \left[ \sum_{j=1}^{N_c} w_{kj} g \left( \sum_{i=1}^n w_{ji} x_j \right) \right] \quad (3)$$

In equation (3)  $g$  is the sigmoid activation function of each neuron in the hidden layer and  $f$  for the linear output neuron;  $n$  is the number of inputs  $N_c$  the number of hidden layers.

Initially networks with lot of hidden layers are operated, but it turns out that the network with a single layer is able to approximate any non-linear function to a degree of accuracy precision (Hornik. K et al, 1989). A conventional learning phase involves assigning initial values to the set of weights and biases, and then minimizes a squared error function  $S(w)$  using the Feed Forward Back Propagation (FFBP) algorithm with Levenberg–Marquardt method for training. This optimization technique is reported to be more powerful than the conventional gradient descent techniques. The study showed that the Marquardt algorithm is very efficient when training networks which has few hundred weights. Although the computational requirements are much higher in iterations of the Marquardt algorithm its efficiency is higher. This is especially true, when high precision is required. The Feed Forward Back Propagation (FFBP) distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or hidden units. The function of hidden neurons is to intervene between the external input and the network output in useful manner (Rumelhart.D.E., 1986).

$$S(w) = \frac{1}{2} \sum_{i=1}^N [y(x_i, w) - z_i]^2 \quad (4)$$

## b. Model performance

There are lots of tests which can be used to analyze the model performance. In this study, the performance tests used are: root mean square error (RMSE), coefficient of determination ( $R^2$ ), Mean absolute error (MAE) and the NASH criteria, which have been used from the following equations. Considering that  $Q_{obs,i}$ , et  $Q_{sim,i}$  are respectively the observed and simulated runoff at time  $i$ , and  $n$  is the total number of time steps on which the test is calculated, then:

- Root mean square error (RMSE): It is expressed as:

$$RSME = \sqrt{\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{n}} \quad (5)$$

- The formula for correlation coefficient can be obtained by the following relation:

$$R^2 = \left[ \frac{\sum_{i=0}^n (Q_{obs,i} - \overline{Q_{obs}}) \cdot (Q_{sim,i} - \overline{Q_{sim}})}{\sqrt{\sum_{i=0}^n (Q_{obs,i} - \overline{Q_{obs}})^2 \cdot \sum_{i=0}^n (Q_{sim,i} - \overline{Q_{sim}})^2}} \right]^2 \quad (6)$$

- Mean absolute error (MAE): It is expressed as:

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 8, August 2017

$$MAE = \frac{1}{n} \sum_{i=1}^n |Q_{obs,t} - Q_{sim,t}| \quad (7)$$

- Nash-Sutcliffe coefficient efficiency. In assessing the performance of the model after generalization, we used this criterion, which is most effective and widely used in hydrology (Nash. J.E. and Sutcliffe. J.V., 1970).

$$NASH = 1 - \frac{\sum_{i=1}^n (Q_{obs,t} - Q_{sim,t})^2}{\sum_{i=1}^n (Q_{obs,t} - \bar{Q}_{obs})^2} \quad (8)$$

The results will be even better than the Nash criterion is close to 1:

If Nash = 1: perfect match between Qobs and Qsim.

If Nash < 0: the average of the observed flow rates gives a better estimate of the range of rates observed than the simulated series (J.E. Nash and J.V. Sutcliffe, 1970).

### c. Data and study area

The test watershed is the Mefou to Nkol-Bisson, a portion of the Mefou River Basin is located in Yaoundé, Cameroon in Central Region, with an area of 109.36 km<sup>2</sup> (Figure 2). Its geographical coordinates are 3°51'50" North latitude and 11°27'9" East longitude. The altitude at the outlet is about 700 m, while the highest peak rises to 1229 m. The shape characteristics were determined from the Yaoundé 3D map 1:50 000 of the National Institute of Cartography published in 1974: Scope of 60.42 km, compactness coefficient (C = 0,28P/√A) C = 1.61 and equivalent rectangle has length of 25.4 km and 4.81 km wide.

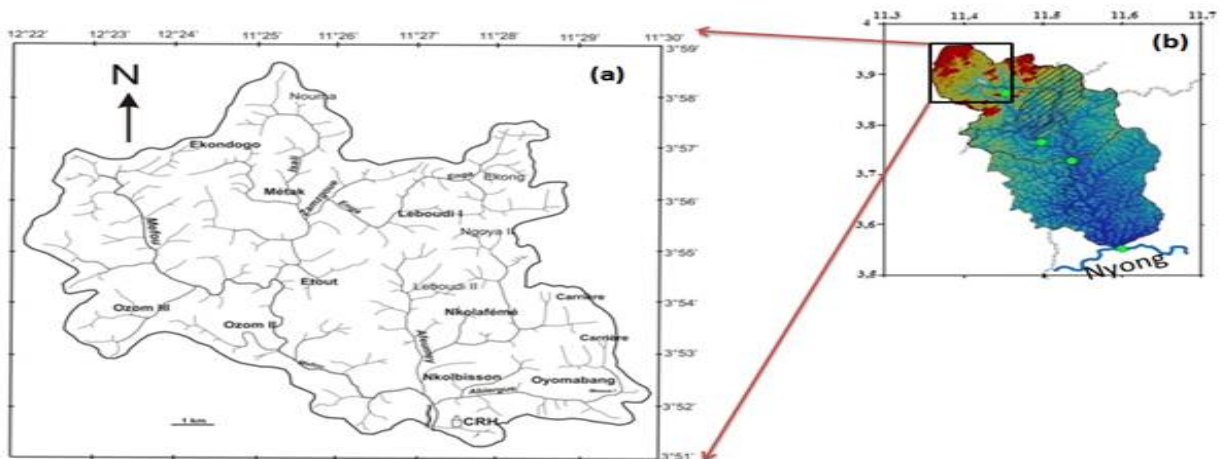


Figure II: Localization of Mefou at Nkolbisson watershed (a), on the big watershed of Mefou (b)

In that watershed, there are three hydrological station installed belonging to the Center for Hydrology Research of the Institute of Geological and Mining Research (CRH-IRGM). The choice of Mefou River was done because, not only it is a good representative small watershed of equatorials Regions, but also, we have 47 months (1425 days) of daily data of rainfall and runoff collected from September 2012 to July 2016. The runoff are taken only at a fixed time of day because of the recommendation in hydrology and the rain collected during the 24 hours before (Shamseldin. A.Y, 1997). In our case this time is set to 17 hours. All the networks selected were calibrated with different combinations of input. Daily rainfall and runoff values are used as input to ANN models.

### III.RESULTS AND DISCUSSION

The length of the data used for calibration is 715 days (from 01/09/2012 to 16/08/2014) and for validation and test is 715 days (from 17/08/2014 to 31/07/2016). Networks are tested with different number of hidden neurons and the model structures with least root mean square error (RMSE) are

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 8, August 2017

considered as the best structure. The ANN structure was tested for lots of hidden neurons. From 5 to 150 hidden neurons, looking every time cases which convergence after little iterations. Accordingly, selection of hidden neurons is done by comparing RMSE of the network. In the intention to find the more appropriate model, a total of eight combinations of input variables were kept for the Mefou river basin.

## a. Results

### i. Performances of models

During the training (calibration) of the eight developed ANN models, predicted daily Runoff are compared with observed daily discharges through statistics performances. The results are presented in the following table I.

Table I: Performances of models during the calibration step

N° of model	RMSE	MAE	R <sup>2</sup>	NASH
A	0,122	0,120	0,997	0,986
B	1,695	0,180	0,984	0,967
C	0,872	0,347	0,996	0,880
D	4,63E-07	0,0043	0,998	0,990
E	1,725	0,142	0,981	0,969
F	1,670	0,297	0,983	0,912
G	1,231	0,168	0,997	0,970
H	2,282	0,136	0,988	0,978

During the validation of the eight developed ANN models, as in the training step, predicted daily Runoff are compared with observed daily discharges through statistics performances. The results are presented in the following table II.

Table II: performances of models during the validation step

N° of model	RMSE	MAE	R <sup>2</sup>	NASH
A	12,754	0,596	0,359	0,0866
B	6,450	0,327	0,844	0,669
C	6,861	0,281	0,877	0,812
D	9,983	0,220	0,445	0,806
E	18,151	0,278	0,649	0,797
F	5,051	0,234	0,842	0,875
G	6,072	0,230	0,719	0,841
H	9,288	0,066	0,661	0,930

The two most efficiency of these performances criterion of the developed ANN models are RMSE and Nash-Sutcliffe efficiency. The following figure III (for calibration) and figure (for validation) give a good illustration.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 8, August 2017

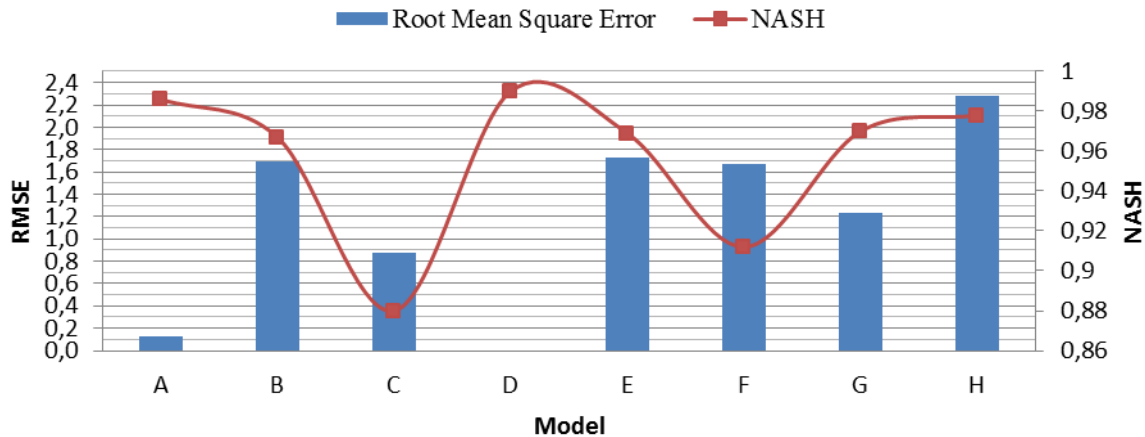


Figure III: Visualization of RMSE and NASH during calibration

During the validation of the model, as in the calibration step, these performances criterion RMSE and Nash-Sutcliffe efficiency are recorded. Their visualization is presented in the following figure IV.

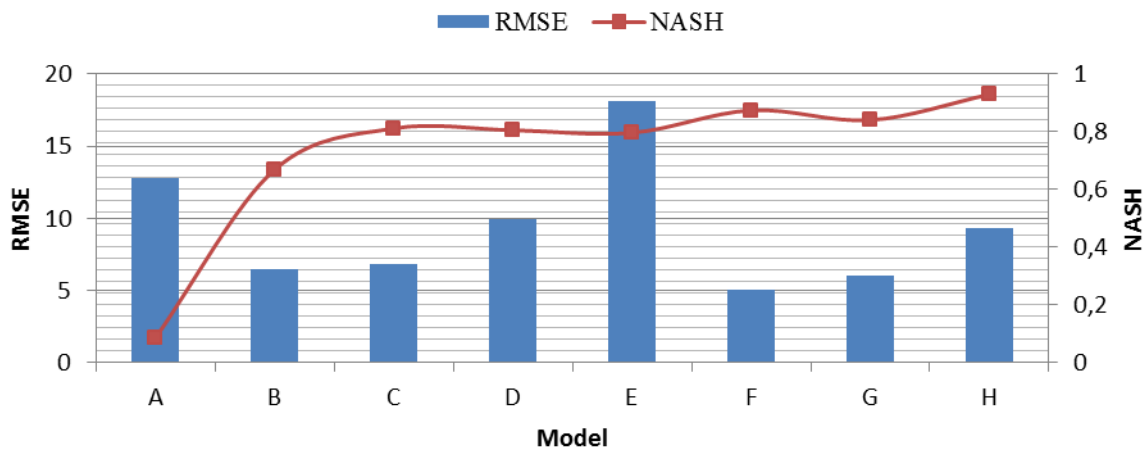


Figure IV: Visualization of RMSE and NASH during validation

## ii. Considered model and comparison of Correlation Coefficients

The objective is to choose the most effective. Let's observe that: the model (D) appears as the best one according to the training step. Best its performances in validation are not good (RMSE = 9,983;  $R^2 = 0,445$ ). It is a case of overtraining. It is the same situation for (A).

The other very good performance is achieved for model (H) during calibration. This model (H) resulted in Nash Sutcliffe efficiency of 0,978 during calibration and 0,930 during validation. But this model (H) has a not good RMSE (9,288) and  $R^2$  (0,661).

Let's observed that model F resulted in Nash Sutcliffe efficiency of 0,912 during calibration and 0,875 during validation. Its RMSE,  $R^2$  and MAE are better than all the others in validation. Then, the generalization of model (F) is more effective. Therefore Model (F) was considered for daily runoff prediction of Mefou Nkol-Bisson watershed. Then the graph of figure v below, illustrate the comparison of Correlation Coefficients between Observed Runoff and Predicted Runoff.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 8, August 2017

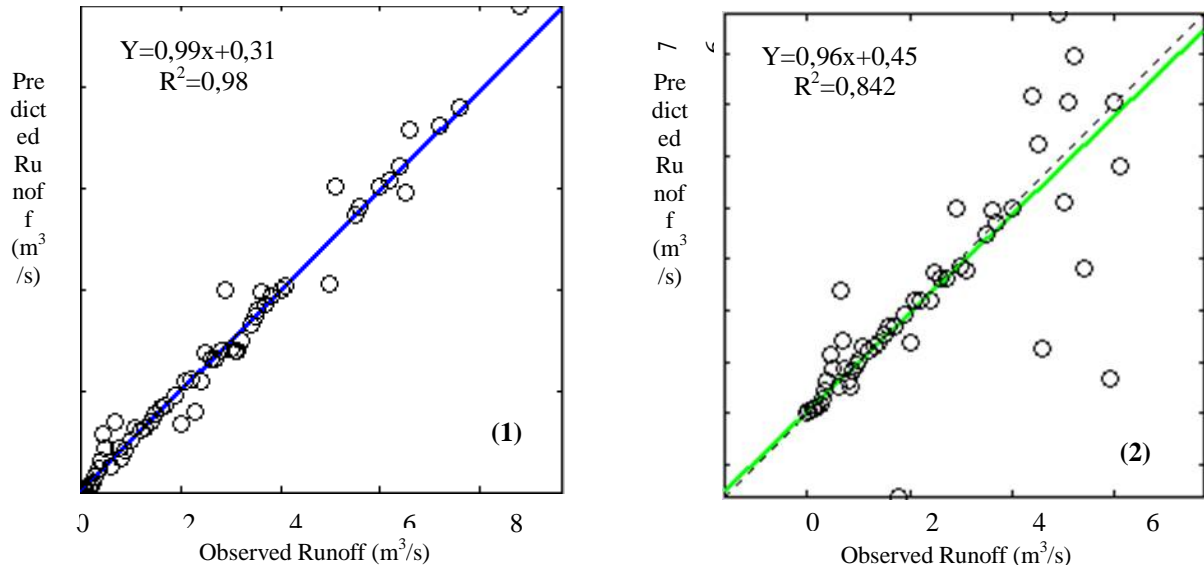


Figure V: Comparison of Correlation Coefficients between Observed Runoff and Predicted Runoff for calibration (1) and validation (2)

### iii. Application of the model

After the validation of that model, an application considered as a test of model generalized has been done on around one year of samples (20% of data, from 20/10/2016 to 31/07/2016). During that simulation, predicted runoff was compared to observed runoff. The NASH criteria obtained is 0,902. The following figure VI shows an illustration.

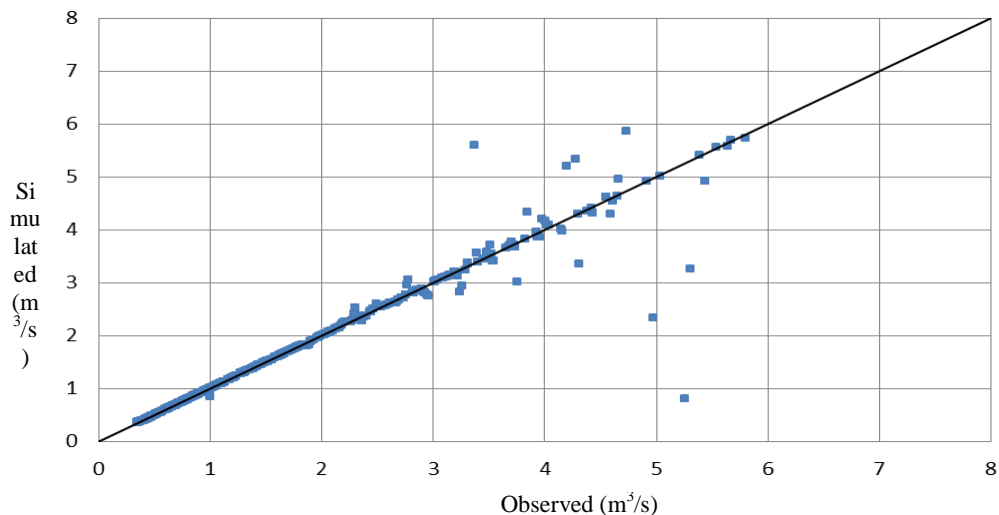


Figure VI: Comparison of runoff (observed/simulated)

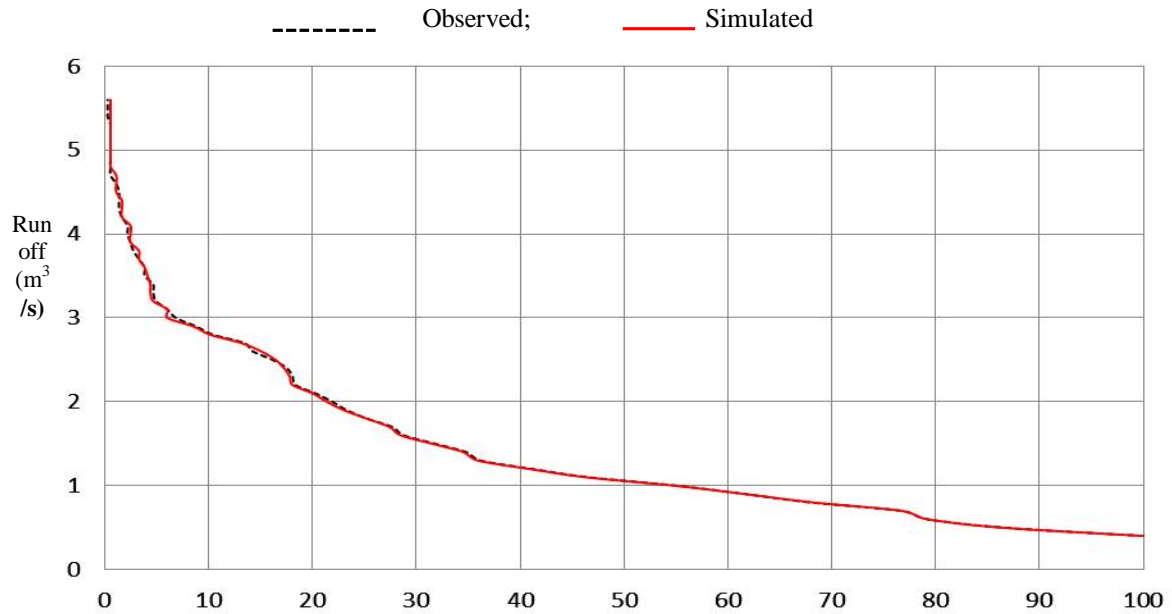
### iv. Classified runoff graph

For using the results of the prediction of runoff for the MHPP project, the classified runoff graph, defined as the graph giving the historical average daily flow of a site classified by the percentage of time where the runoff is greater than or equal to this flow, is proposed for the last year of the sample in the following Figure VII.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)



Frequency of time '%' (August 2015 – July 2016)

Figure VII: Classified discharges graph

## v. Hydrograph of Daily means rainfalls

The hydrograph based on observed and simulated data and the average daily rainfall of 2015 (a year selected in the sample of data) are represented in the following Figure III.

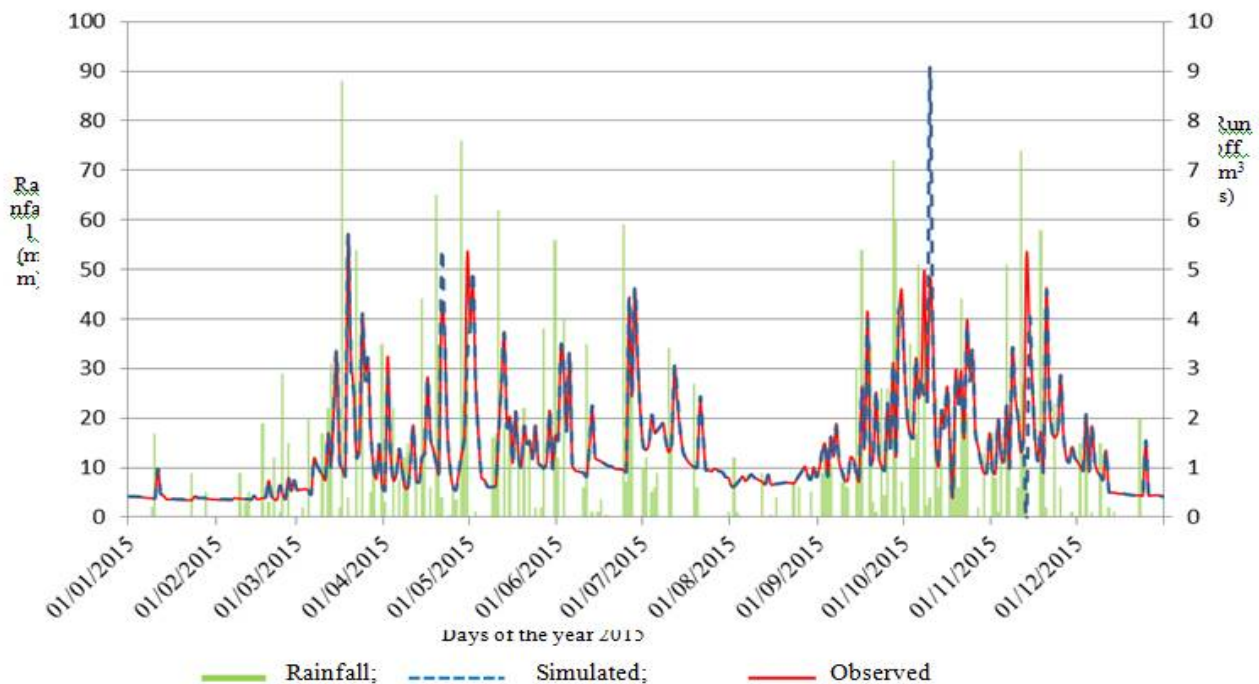


Figure VIII: Hydrograph and daily mean rainfall of the year 2015



# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 8, August 2017

## b. Discussion

Looking for the figures III & IV, it can be observed that on adding hidden neurons RMSE decreases up to a certain value and again increases.

Model (F) resulted in Nash Sutcliffe efficiency of 0,912 during calibration and 0,875 during validation. Its RMSE, R2 and MAE are better than all the others in validation. (F) Predicted daily runoff very close to observed daily runoff. Therefore Model (F) was considered for daily runoff prediction of Mefou Nkol-Bisson watershed.

Globally after the generalization of the model, from de figure 6 it can be observed that for runoffs less than 2.4 m<sup>3</sup>/s, predicted runoffs are almost identical to the observed runoffs. For the runoffs between 2.4 and 4m<sup>3</sup>/s, there are some significant differences. For the runoff exceeding 4m<sup>3</sup>/s, we note the larger gaps even have aberrations. The decline in performance is predictable for high runoff on small watersheds. So, the best performance of the model for the low runoff positively influences the overall performance of the model.

Moreover, the criterion of NASH after the generalization is around 0.9. This is a model of good performance in predicting runoff from rainfall on the watershed Mefou to Nkol-Bisson. This result shows that the model can predict the annual regime of the river significantly better than the mean annual flow. Therefore, it can be necessary for the effective design of micro-hydro; compared to the use of the mean annual flow regularly used for the same cause.

Looking at Figure 8, the presence of seasonality is always well marked. The fluctuation of the curves is due to the fact that the daily prediction of flow is dictated by the rain. The model underestimates the flood flows.

## IV. CONCLUSION

It appears from the analysis made that, in order to promote the development of MHPP in Cameroon in general and in the equatorial regions precisely, the need to solve the problem of lack of hydrological data on small watersheds no gauged is necessary. Rainfall data being more available, a rainfall-runoff model was developed on Mefou Nkol-Bisson watershed using artificial neural networks to allow the prediction of runoff from the rain fall data. The model allows subsequently simulating the runoffs classified graph very necessary for MHPP.

Since the performance criteria NASH and coefficient of correlation calculated are ( $R^2 = 0.983$  and  $NASH = 0.912$  in calibration), ( $R^2 = 0.842$  and  $NASH = 0.875$  in validation) and ( $NASH = 0.9$  for an application), the results from our model are very interesting. Its implementation can be done on other small watersheds with similar climates for the MHPP design needs. Thus, the model can still be recommended for any application in the objective of MHPP where the criterion of NASH is greater than 0.6.

## REFERENCES

- [1] Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C., "Image inpainting", Proc. SIGGRAPH, pp. 417-424, 2000.
- [2] Criminisi, A., Perez, P., and Toyama, K., "Region filling and object removal by exemplar-based image inpainting.", IEEE Transactions on Image Processing, Vol.13, N°9., pp. 1200-1212, 2004.
- [3] Marcelo, B., Luminita, V., Guillermo, S., Stanley, O., "Simultaneous Structure and Texture Image Inpainting", IEEE Transactions On Image Processing, Vol. 12, N°8, 2003.
- [4] Yassin, M. Y. H., and Lina, J. K., "Morphological Text Extraction from Images", IEEE Transactions On Image Processing, Vol. 9, N°11., 2000.
- [5] Eftychios, A., Pnevmatikakis, Petros, M., "An Inpainting System For Automatic Image Structure-Texture Restoration With Text Removal", IEEE trans, Vol.978,N°1, pp.1744-1764, 2008.
- [6] Bhuvaneshwari, S., Subashini, T.S., "Automatic Detection and Inpainting of Text Images", International Journal of Computer Applications, Vol. 61, N°7, 2013.
- [7] Aria, P., and Richard, L., Tutwiler, "Automatic Feature Extraction and Text Recognition from Scanned Topographic Maps", IEEE Transactions on geosciences and remote sensing, Vol. 49, N° 12, 2011.
- [8] Nobuo, E., Marius, B. L., and Schomaker., "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons" , Proc. of 17th Int. Conf. on Pattern Recognition (ICPR), IEEE Computer Society, Vol.2, pp. 683-686, 2004.
- [9] Rajesh, H. D., and Noor, M., "Text Detection, Removal and Region Filling Using Image Inpainting", International Journal of Futuristic Science Engineering and Technology, ISSN. 2320 – 4486, Vol.1, N° 2, 2013.
- [10] Uday, Modha, and Preeti, D., " Image Inpainting-Automatic Detection and Removal of Text From Images", International Journal of Engineering Research and Applications (IJERA), ISSN. 2248-9622, Vol. 2, N° 2, 2012.
- [11] Muthukumar, S., Krishnan, N., Pasupathi, P., and Deepa, S., "Analysis of Image Inpainting Techniques with Exemplar, Poisson, Successive Elimination and 8 Pixel Neighborhood Methods", International Journal of Computer Applications, ISSN. 0975 – 8887, Vol. 9, N°11, 2010.