FILTRATION MODELING USING ARTIFICIAL NEURAL NETWORK (ANN)

By Dr. Awatif Soaded Alsaqqar

University of Baghdad/ College of Engineering/ Civil Engineering Dept.

ABSTRACT

In this research Artificial Neural Network (ANN) technique was applied to study the filtration process in water treatment. Eight models have been developed and tested using data from a pilot filtration plant, working under different process design criteria; influent turbidity, bed depth, grain size, filtration rate and running time (length of the filtration run), recording effluent turbidity and head losses. The ANN models were constructed for the prediction of different performance criteria in the filtration process: effluent turbidity, head losses and running time. The results indicate that it is quite possible to use artificial neural networks in predicting effluent turbidity, head losses and running time in the filtration process, with a good degree of accuracy reaching 97.26, 95.92 and 86.43% respectively. These ANN models could be used as a support for workers in operating the filters in water treatment plants and to improve water treatment process. With the use of ANN, water systems will get more efficient, so reducing operation cost and improving the quality of the water produced.

KEY WORDS: Artificial Neural Network, modeling, water treatment, filtration, turbidity, head losses, running time.

الخلاصة

في هذا البحث جرى تطبيق تقنية الشبكات العصبية الاصطناعية لدراسة عملية الترشيح في معالجة المياه. بنيت ثمانية نماذج و اختبرت باستعمال النتائج العملية من منظومة ريادية للترشيح والتي تعمل تحت ظروف تصميمية: الكدرة، عمق المرشح، حجم مادة الترشيح، سرعة الترشيح و فترة الترشيح و تسجيل الكدرة للماء الراشح والفقدان في الشحنة. هذه النماذج بنيت للتنبؤ بعكورة الماء الراشح ،الفقدان في الشحنة و فترة الترشيح. وأظهرت النتائج بان هذه النماذج لها القابلية للتنبؤ بدرجة كبيرة تصل إلى 97.26% للكعورة ، 95.92% لفقدان بالشحنة و 86.43% لفترة الترشيح. هذه النماذج تساعد الفنيين في تشغيل المرشحات في محطات معالجة المياه و لتحسين عمليات المعالجة. استعمال تقنية الشبكات العصبية الاصطناعية ، تجعل هذه الأنظمة كفوءة وبذلك تنخفض كلفة التشغيل وتحسين نوعية الماء الناتج.

INTRODUCTION

Hard work has been done to predict the relationship between water quality parameters and the most efficient points in water treatment plants. The progress in such work was carried out using multi varied regressions. Nowadays artificial intelligence techniques such as Artificial Neural Networks (ANNs) and Genetic Algorithms can contribute for optimizing performance and operation in water treatment plants (Fernandez and Galvis, 2002).

Over the past two decades, there has been an increased interest in the new class of intelligent systems known as the Artificial Neural Networks (ANNs). These networks are found to be powerful tools for organizing and correlating information in ways, which have proved to be useful for solving problems too complex to understand, too poorly to analyze or too resource-intensive to tackle using more traditional computational methods (TRB, 1999). ANNs have a wide range of scientific applications among them is in different engineering aspects.

In river sanitation, Stewart used ANN for predicting dissolved oxygen concentrations in the Tualatin River, Oregon, USA. The ANN model was constructed using air temperature, solar radiation, rainfall and stream flow as input data. The model predicted the dissolved oxygen concentrations with acceptable accuracy and high correlation reaching 0.83 and absolute error less than 0.9 mg/l (Stewart, 2002). Kanani and his team used multi layer perception (MLP) neural network and input delay neural network (IDNN) for predicting total dissolved solid (TDS) in the Achechay River in Iran. The two models showed acceptable precision in predicting TDS in the study area and the results can be utilized in optimized management and planning of water recourses (Kanani et al., 2008).

Another application of ANN is the modeling of water distribution systems. In 2003, a team of engineers made use of linear regression models and multi layer perceptron artificial neural networks, to predict chlorine concentrations in the Hope Valley water distribution system, located in Adelaide, south of Australia. The inputs required to feed the models were, WTP chlorine concentration, WTP water temperature, WTP flow, chlorine concentration at the inlet tank and water temperature and chlorine concentration at the sampling points in the distribution system. The multi layer perceptron ANN was found to consistently outperform traditional linear regression models (Gibbs et al., 2003). The ANN can be an online tool to aid in the determination of the required chlorine dosing rates in water distribution network. Another case study was conducted by Memon and his group in 2008 for the prediction of the electrical conductivity (EC) in drinking water flowing in the network of Hyderabad, Pakistan. In this study a Radial Basis Function (RBF) network was applied. The model had the powerful looming of predicting EC in the flowing water, taking into account the parameters turbidity, pH, chloride and alkalinity at the sampling points as input data in the network (Memon et al. in 2008).

In water treatment, ANN was used in some aspects also. Nicolas and Thierry applied ANN techniques for coagulation control in the drinking water treatment plant of Viry-Chatillon near Paris, France. They developed a software sensor including a self organizing map (SOM) for

sensor data validation and missing data reconstruction, and a multi layer perceptron (MLP) for modeling the coagulation process. The important objective of this study is to automatically validate the sensor measurements to provide reliable inputs to the automatic coagulation control system (Nicolas and Thierry, 2001).In 2002, Fernandez and Galvis showed that ANNs are capable of identifying usable relationships between optimal doses of coagulants and inflow water quality in Puerto Mallarino water treatment plant in Cali. According to the results the online ANN model implementation, in real time, coupled with the alum ejector of the plant, could save around 10% in alum dosage applied in the coagulation process (Fernandez and Galvis, 2002).

In this study the ANN technique is used in modeling the filtration process in rapid sand filters in drinking water treatment plants. Filtration is considered a polishing process which removes the remaining colloidal particles from water. The main objective of this process is to reduce the turbidity of water to drinking water specifications providing long filtration runs with low head losses. These ANN models may help to control the filtration process and improve the performance of the system.

ARTIFICIAL NEURAL NETWORKS (ANNs)

ANN models are specified by network topology, node characteristics and training or learning rules. It is an interconnection set of weights that contains the knowledge generated by the model (Hafizan et al., 2004). Different types of ANNs exist; the most common type is the feed forward network termed the multilayer perceptron. In this network, the artificial neurons or processing units are arranged in a layered configuration as shown in Figure 1:

Input layer - connecting the input information to the network.

Hidden layer (one or more) – acting as the intermediate computational layer.

Output layer – producing the desired output.

Units in the input layer introduce normalized of filtered values of each input into the network. Units in the hidden and output layers are connected to all of the units in the preceding layer. Each connection carries a weighting factor. The weighted sum of all inputs to a processing unit is calculated and compared to a threshold value. An activation signal then is passed through a mathematical transfer function to create an output signal that is sent to processing units in the next layer. Training an ANN is a mathematical exercise that optimizes all of the network weights and threshold values, using some fraction of the available data. ANN learns as long as the input data set contains a wide range of patterns that the network can predict. The final model is likely to find those patterns and successfully use them in its prediction (Stewart, 2002).

CASE STUDY

Preliminary design of rapid filters in water treatment consists, of several tasks including (Crittenden et al., 2005):

- Performance criteria such as, effluent water quality mainly turbidity, filter run length and recovery (ratio of net to total water filtered).

-Selecting process design criteria such as, required level of pretreatment (influent turbidity), filter media type, size & depth, filtration rate and available head.

Filter efficiency is a function of these design criteria and the best way to determine the factors that influences this process is by a pilot plant study (Qasim et al., 2000). The data required as inputs to the ANN models were collected from filtration pilot plants. In such plants the filtration units are columns of different materials, glass or plastic that contain the filter media. In this study, Iraqi sand used in water treatment plants, was tested under different working conditions. Table 1 describes the case study which was selected as the input data for the ANN models.

A total number of 15 filtration runs were performed testing the filtration process, recording effluent turbidity and head losses under different process design criteria; influent turbidity, bed depth, grain size, filtration rate and running time (length of the filtration run). The operation of a filter must focuses on head loss, filter run length and effluent turbidity. Improper operation of filtration units can result in poor quality of the filtered water and may damage the filter bed.

ARTIFICIAL NEURAL NETWORKS (ANN) MODELS

Several neural network software are available; Neuframe 4 has been used in this study. Three ANN models were constructed for the prediction of different performance criteria in the filtration process: prediction of effluent turbidity, prediction of the head loss and prediction of the length of the filter run.

-Effluent water quality (turbidity)

The most common measurement of particulate matter in water is turbidity. Filtered turbidity can vary due to changes in raw water quality and utilities need some latitude to respond to these changes. Most utilities set their turbidity goal below 0.3 NTU with a typical goal being 0.1 NTU (Crittenden et al., 2005).

- Head loss

Throughout a filter run, solid particles are deposited in the pores of the filter bed and cause a decrease in the porosity and an increase in the head loss through the filter bed. The head loss at the end of the filter run is to be calculated when designing the filter. This could be done using some empirical equations depending on the porosity at the end of the filter run. The maximum head loss in rapid filters may reach 2 to 3 m. Accurate estimates of head loss values can be determined from pilot plant studies (Qasim et al.,2000).

-Filters run length (running time)

The length of the filter runs dictates how often backwashes must be performed and has an impact on recovery. The frequency of washing has a direct impact on the quantity of labor involved in filter operation. Typically the minimum filter run is one day, with some designs it is also common to reach four days (Crittenden et al., 2005).

The first step for the determination of ANN model is the selection of the data to be the input variables, they were provided from the pilot plant as mentioned in Table 1. The data for each model is shown in Table 2.

The data have to be divided to three sets, training, testing and validation. This step is achieved by trial and error to select the best division with respect to the lowest testing error followed by training error and high correlation coefficient of the validation set. The general strategy adopted for finding the optimal network architecture and internal parameters that control the training process is by trial and error using the default parameters of the software. In this step first the nodes of the hidden layer are increased until no significant improvement is gained in the model performance. Then the model is tested by changing the default parameters of the software, the momentum term which is 0.8 and the learning rate 0.2. Finally the transfer functions of the input and hidden layers are tested where the default functions of the software are, linear in the input layer and sigmoid in the hidden layer. The default alternatives of the software are to test the following functions: linear, sigmoid and hyperbolic tangent (tanh). The effect of the different combinations of these parameters is summarized in Table 3 which shows the best model performance recording the lowest testing error and the highest correlation coefficient. In each step the testing errors were recorded and the reduction percentages were calculated to see how the improvement is gained in the model performance, this is shown in Table 4.

RESULTS AND DISCUSSION

Prediction of Effluent turbidity

Three models were determined for the prediction of the effluent turbidity of the filtered water which is illustrated in Table 3. All models had one hidden layer but with different number of nodes. In model tur and tur3 one node was used and two nodes in tur2 were required. The momentum rate was high, reaching 0.98 and 0.92 in models tur and tur2 respectively; these values are higher than the default value of the software 0.8 (momentum rate represents the weight change which is in a direction that is a combination of the current gradient descent. This approach is beneficial when some training data are very different from the majority of the data. The momentum rate is added to obtain a faster convergence. Sivanandam and Paulraj, 2004). The learning rate values were around the default 0.2 only in model tur2 it reached 0.7 (weight changes are proportional to the negative gradient of the error. This guideline determines the relative changes that must occur in different weights when a training samples(s) is presented. These magnitudes changes are dependent on the learning rate. Sivanandam and Paulraj, 2004).The method for adopting the learning rate was increasing it in order to improve the performance of the model by decreasing the testing error and increasing the correlation coefficient. Finally the transfer functions were tested, for the three models, the function of the input layer was tanh and in the hidden layer it remained sigmoid as it is a default function in the software. The reduction percentages in the testing error, as shown in Table 4, were small when increasing the number of nodes in the hidden layer. When testing the effect of increasing the momentum rate, only in model tur it reduced the testing error by 4.58%. A reduction of 4.73% was recorded in model tur2 for a tanh function of the input layer. Finally the correlation coefficient in models tur and tur2 reached 97.26 and 92.84 % respectively.

Prediction of Head loss

For the prediction of the head loss at the end of the filtration run, three models were also constructed. Table 3 illustrates the architecture of each model. The number of nodes in the hidden layer reached five in model hl and two in models hl2 and hl3. The momentum rate ranged 0.68 to 0.92, where the learning rate kept near to the default value of 0.2. The reduction in the testing rate was, 2.18% in model hl at momentum rate 0.92 and 6.65% for being the two transfer functions tanh in the input and hidden layers in model hl2. The models hl and hl2 gave the highest correlation coefficients recording 95.92 and 93.06 %respectively.

Prediction of Running time (filtration length)

As for predicting the filtration length, two models were only constructed. As shown in Table 3, one or two nodes are required in the hidden layer where it reduced the testing error 10. 38 % in model hl-rt. The momentum rate was high in the two models, reaching 0.9 and 0.96. The learning rate kept near to the default value of 0.2. The transfer functions in the input and hidden layer were tanh where it reduced the testing error by 8.25 %. The model hl-rt predicted the running time with a correlation coefficient 86.43%.

A total data of 127 were exploited to build the different models of this study. In these models, 68-70% of the input data were used for training, 19-21% for testing and the remaining 10-13% for validation. The models developed were tested to show the best models to predict the effluent turbidity, head losses and running time. The model tur showed the best agreement between the predicted and measured effluent turbidity which is shown in figure 2. This figure illustrates the hourly variation of the predicted and measured turbidity for one of the experimental runs in the pilot filtration plant. These final predictions appear to be accurate enough to be useful. The models for predicting the head losses also gave good agreement between the predicted and measured head losses from the experimental work as shown in figure 3, the best model was hl2. Figure 4 shows the best model for predicting the running time (rt) or the length of the filtration run, which was hl-rt.

The results indicate that it is quite possible to predict effluent turbidity, head losses and running time in the filtration process using intelligent data driven methods such as artificial neural networks.

CONCLUSIONS

The results of this study indicate that it is quite possible to predict effluent turbidity, head losses and running time in the filtration process using intelligent data driven methods such as artificial neural networks, and the following conclusions are brought out:

-The data used to build the ANN models were distributed, 68-70% used for training, 19-21% for testing and the remaining 10-13% for validation.

-Models predicting effluent turbidity, had 1-2 nodes in the hidden layer, high momentum rate reaching 0.98, learning rate within 0.2 and the testing error was low 4.65% with a correlation coefficient of 97%.

- Models predicting the head losses, had 2-5 nodes in the hidden layer, high momentum rate reaching 0.92, learning rate within 0.2 and the testing error was low 4.46% with a correlation coefficient of 95%.

-The model for predicting the running time or the length of the filtration run had 2 nodes in the hidden layer, 0.96 momentum rate, 0.22 learning rate where the testing error was 5.93% which gave a correlation coefficient of 86%.

-All models gave good agreement between the predicted and measured data from the experimental work.

REFERENCES

- Crittenden J. C., R. R. Trussel, D. W. Hand, K. J. Howe and G. Tchobanogious, (2005), Water Treatment Principles and Design .2nd edition, USA.

- Fernandez E. and Galvis A .(2002), Artificial Neural Network Model used for Clear Water Treatment Plants. Seminario Internacional La Hidroinformatica en la Gestion Integrada de los Recursos Hidricos

- Gibbs M. S., N. Morgon, H. R. Maier, G. C. Dandy, M. Holmes and J. B. Nixon (2003),Use of Artificial Neural Networks for Modeling Chlorine Residuals in Water Distribution Systems.

- Hafizan Juahir, Sharifuddin M. Zain, Mohd Ekhwan, Mazlin Mokhtar and Hasfalina Man (2004), Application of Artificial Neural Network Models for Predicting Water Quality Index. Journal Kejuruteraan Awam 16(2): 42-55.

- Kanani,S.,G. Asadollahfardi and A. Ghanbari (2008), Application of Artificial Neural Network to Predict Total Dissolved Solid in Achechay River Basin. World Applied Science Journal 4 (5):646-654.

- Memon N., M. A. Unar, A. K. Ansari, G. B. Khaskheli and B. A. Memon (2008), Prediction of Parametric Value of Drinking Water of Hyderabad City by Artificial Neural Networks Modeling. WSEAS Transactions on Environment and Development, Issue 8, volume 4, August.

- Nicolas Valentin and Thierry Denceux (2001), A Neural Network Based Software Sensor for Coagulation Control in a Water Treatment Plant. Intelligent Data Analysis 5 23-39 IOS Press.

- Qasim S. R., E. M. Motley and G. Zhu (2000), Water Works Engineering Planning, Design and Operation. Prentice Hall PTR. USA.

- Sivanandam S. N. and M. Paulraj (2004), Introduction to Artificial Neural Networks. Vikas Publishing House PVT LTD New Delhi.

- Stewart A. Rounds (2002), Development of a Neural Network Model for Dissolved Oxygen in the Tualatin River, Oregon. Proceedings of the 2nd Federal Interagency Hydrologic Modeling Conference, Las Vegus, Nevada July 29- August 1.

- TRB (1999), Use of Artificial Neural Network in Geomechanical and Pavement Systems. Transportation Research Circular No. e-c 012.

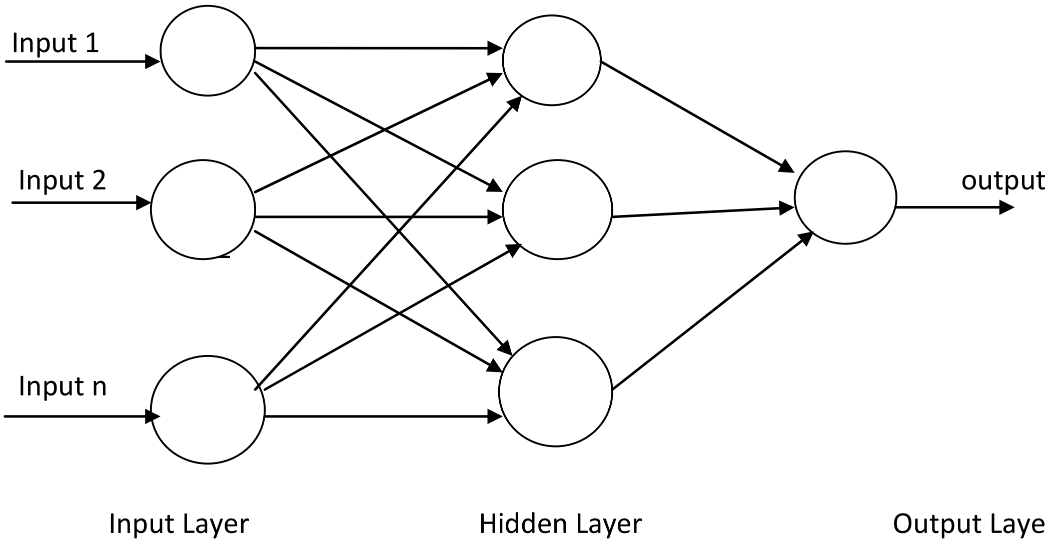

Fig. 1 Multi layer neural network

Table 1 Description of the case study

Filtration rate (m/hr)	Bed depth (cm)	Grain size (mm)	Influent	No. of runs	Running time (hr)
5, 7.5, 10, 15	70	0.72	The same water to the filters of Alwathba WTP with variable turbidity ranging 3.7 to 27 NTU	15	1 to 30

Table 2 Data used in ANN models

Model	Input data
	Prediction of Effluent turbidity (tur)
tur	Running time, filtration rate, bed depth, grain size and influent turbidity
tur-2	Influent turbidity, filtration rate, bed depth, grain size and head loss
tur-3	Running time, filtration rate, bed depth, grain size ,influent turbidity and head loss
	Prediction of Head loss (hl)
hl	Running time, filtration rate, bed depth, grain size , influent and effluent turbidity
hl-2	Running time, filtration rate, bed depth, grain size and influent turbidity
hl-3	Influent turbidity, filtration rate, bed depth, grain size and effluent turbidity
	Prediction of Running time (rt)
tur-rt	Influent turbidity, filtration rate, bed depth, grain size and effluent turbidity
hl-rt	Influent turbidity, filtration rate, bed depth, grain size and head loss

Table 3 Models Architecture, Optimization and Stopping Criteria

Model	Input layer	Hidden layer	Momentum rate	Learning rate	Testing error (%)	Training error (%)	Correlation coefficient (R%)
Prediction of Effluent turbidity (tur)							
tur	5 tanh	1 sigmoid	0.98	0.18	4.6509	5.6430	97.26
tur2	5 tanh	2 sigmoid	0.92	0.7	4.6132	5.7169	92.84
tur3	6 tanh	1 sigmoid	0.76	0.24	5.2271	5.3156	79.56
Prediction of Head loss (hl)							
hl	6 linear	5 linear	0.92	0.2	4.4668	5.5207	95.92
hl2	5 tanh	2 tanh	0.8	0.18	4.6637	5.4402	93.06
hl3	5 tanh	2 sigmoid	0.68	0.2	7.7716	9.1015	68.78
Prediction of Running time (rt)							
tur-rt	5 linear	1 tanh	0.9	0.22	11.9993	12.4373	69.86
hl-rt	5 tanh	2 tanh	0.96	0.22	5.9355	6.3205	86.43

Note: default values of the software, momentum rate= 0.8 & learning rate= 0.2

Table 4 Reduction in the testing error (%)

Model	Nodes in the hidden layer	Momentum rate	Learning rate	Transfer function
tur		4.58 (0.98)	0.21	0.05
tur2	0.08	0.84	3.99 (0.7)	4.73 (1^{st} tanh)
tur3	0.67	0.01	0.02	0.14
hl	0.73	2.18 (0.92)		0.72 (2^{nd} linear)
hl2	0.89		0.17	6.65 (1^{st}& 2^{nd} tanh)
hl3	0.15	0.09		4.81 (1^{st} tanh)
tur-rt		0.37	0.03	0.04 (2^{nd} tanh)
hl-rt	10.38 (2 nodes)	2.04 (0.96)	0.06	8.25(1^{st}& 2^{nd} tanh)

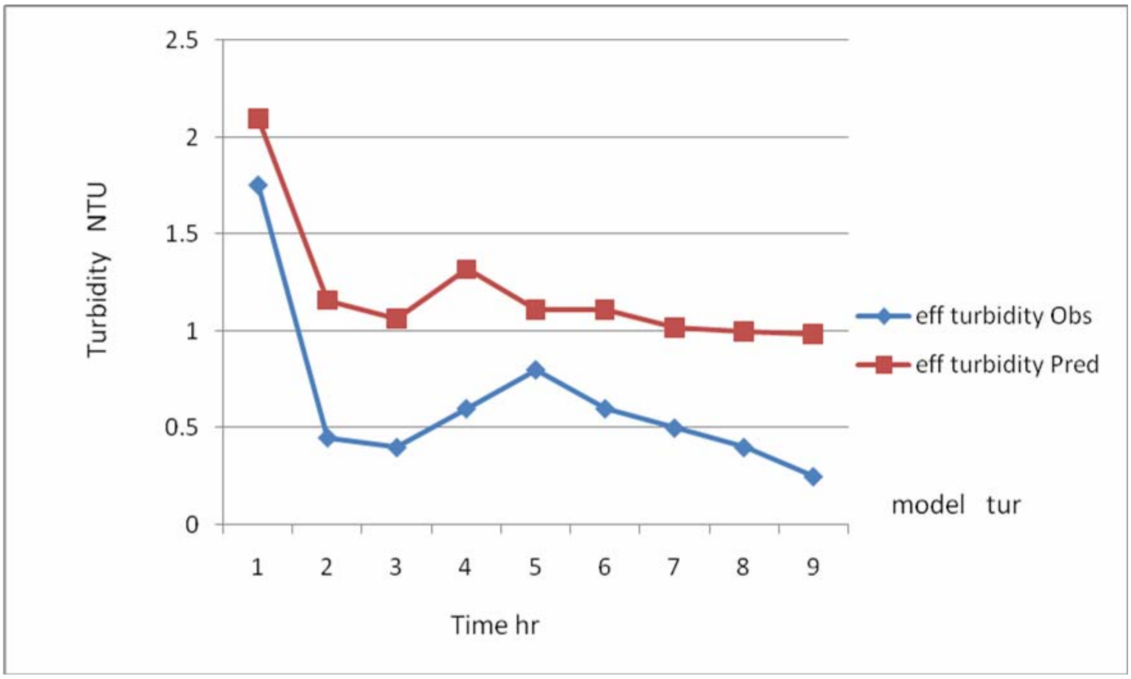

Fig.2 Hourly variation of the predicted and measured turbidity

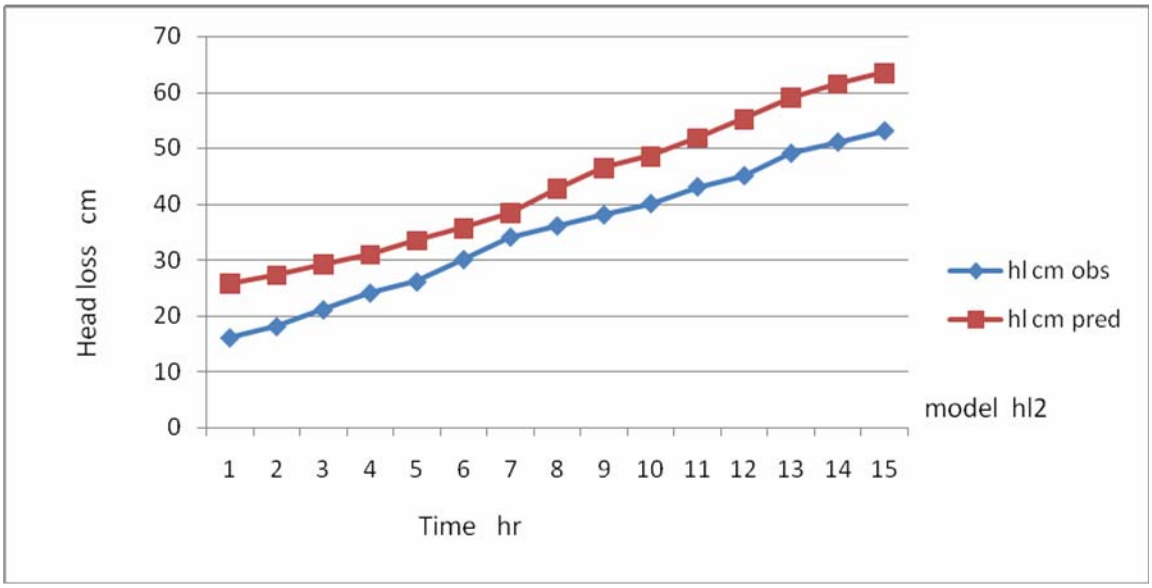

Fig.3 Hourly variation of the predicted and measured head losses

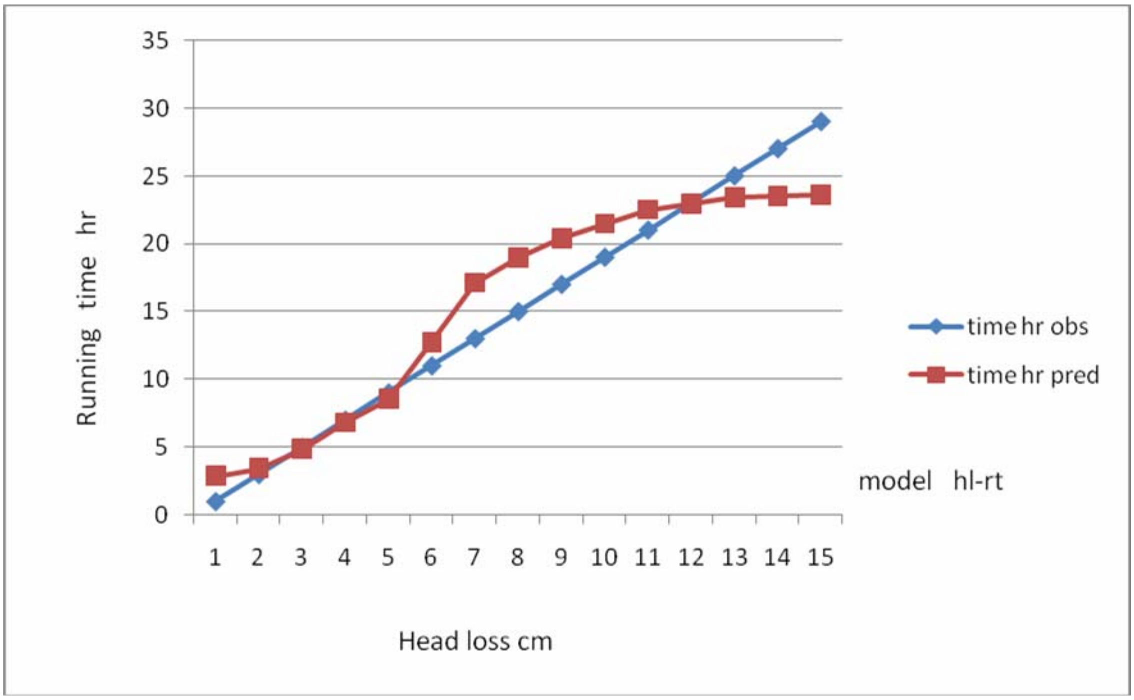

Fig.4 Variation of the predicted and measured running time