Journal of Engineering and Applied Sciences 1 (4): 406-409, 2006

Application of Artificial Neural Network for Predicting Microbiological Pollution in Fresh Water

[1,3]S. Bouharati, [1]K. Benmahammed, [2]Y.M. El-Assaf and [3]D. Harzallah
[1]Department of Electronics, Faculty of Engineering,
Laboratory of Intelligent Systems, UFA Setif, Algeria
[2]Faculty of Engineering, American University of Sharjah, UAE
[3]Department of Biology, Faculty of Sciences,
Microbiological Laboratory, UFA Setif, Algeria

Abstract: The classical methods for detecting the micro biological pollution in fresh water are based on the detection of the *Escherichia coli* bacteria which indicators of contamination. Some of them are based on simple and easy-to-handle concepts like laboratory methods. But to check each water supply for these contaminants would be a time-consuming job and a qualify operators. In this study an attempt was made to develop a new approach for the detection of *Escherichia coli* bacteria in fresh water is proposed. This is done by the Artificial Neural Networks (ANNs) in which the performance was measured in two ways, training and testing. The artificial neural networks prediction is proposed based on effect of the variations of the physical and chemical parameters occurred during bacteria growth-temperature, pH, electrical potential and electrical conductivity-of many varieties of water (surface water, well water, drinking water and used water) on the number *Escherichia coli* in water. The instantaneous result obtained by measurements of the inputs parameters of water from sensors. The superiority of the ANNs method is due to high prediction accuracy and the ability to compute combined effects of physical and chemical factors on the fresh water induced during the bacteria growth. This tool can be used in the first time for predicting the micro biological pollution in water.

Key words: Testing water, *E. coli* detection, artificial neural network, ANNs prediction

INTRODUCTION

Fresh water is the medium of bacteria growth. The institution of water quality standards has aided in the prevention of these types of outbreaks. The WHO currently uses a fecal *Escherichia coli* standard to measure bacterial water health. Because of their communal properties and prevalence they are used as indicator organisms. *Escherichia coli* is a member of the coliform family shown to more strongly correlate to the presence of pathogenic organisms. In the United States, the EPA strongly encourages the use of *E. coli* to measure water quality instead of fecal coliform. (Stevenson, WA, USA) The bacteriological quality of water, gagged by the presence of coliform bacteria, is easily tested in the laboratory[1]. But in medical and industrial microbiology is now a need for « Rapid methods » that will shorten the time between receipt of specimen or samples and the issue of a report. In hygienic water testing, emphasis is mainly placed on the other organisms, but further hygienically important bacteria can be present in water[2]. The mechanical, electrical or electronic procedures alone or in combination, can be removing the need for direct human action[3]. Conductance or impedance measurements are used, e.g. for rapid measurement of total microbial activity in many products, recently, an alternative to the traditional direct conductimetry has become available, namely the indirect conductimetry in which the evolution of pH and electrical potential from culture media as a result of growth of micro-organisms can be studied by Owens *et al.*[2]. In this study, modeling methods are used for predicting behavior of a complex system, such as *E. coli* concentration in water, which have a high non-linearity and a complex structure. The data used in this study were pH of the water; temperature; electrical potential; and electrical conductivity. The difficulties are embedded in these measurements especially in the values of parameters and their origins. Whereas, the Artificial Neural Network (ANN) has been increasingly in various fields of science and technology by Brion *et al.*[4].

Corresponding Author: Bouharati, S., Department of Electronic, Faculty of engineering, Laboratory of intelligent systems, UFA Setif-Algeria

The predictive microbiology becomes an emerging and important new field with broad applications to the food processing industry[5-8]. Recently, ANNs have found a number of applications in the area of water quality modeling. The Artificial neural network is a well-defined function that maps real-value inputs to real-value output. The inputs are the physical and chemical parameters variations of bacterial growth. The estimation of such variables is often a complex and non-linear problem, making it suitable for ANN application. The objective of this research was to report the detection of the presence or absence of the *Escherichia coli* in water. ANN computations were applied for developing predictive models as a result of combined physical and chemical parameters. of fresh water. Modeling ability of ANN algorithm was very adequate to predict microbiological contamination in fresh water.

MATERIALS AND METHODS

Varieties of water samples and bacterial analysis: 1000 samples of water are used. They were collected from different varieties of water (well water-250 samples surface water-250 samples used water-250 samples and drinking water-250 samples). The physical and chemical changes, pH, Electrical Potential (EP), Electrical Conductivity (EC) were measured at different Temperatures (T) according the number of bacteria LogN). The bacterial analysis concerned the *E. coli* bacteria (fecal coliforms) measured by the classical methods NPP in each study (Table 1).

Structure of the artificial neural network: Artificial Neural Networks can be defined as highly connected arrays of neurons[9]. The internal structure of a neuron is shown in Fig. 1.

The internal activity of a single neuron computes the weighted sum of the inputs e_i(net) and passes this sum through a non-linear function f, according to:

$$\begin{cases} net = \sum_{i=1}^{n} w_i e_i + w_b \\ o = f(net) \end{cases}$$

Another term called the bias term w_b is associated with this sum. The function w_b used as a non-linear function f is, for example, a sigmoid function given by:

$$f(net) = \frac{1}{1+e^{-net}}$$

Table 1: Variations of the Output LogN according each output T°; pH; EP; EC and the origin of water, surface water, used water, well water and drinking water

Log N	T°	pH	E.P	E.C	Origine
4.8	15	8.45	0.85	3.01	1
.	.	.	.	.	.
6.9	24	6.98	0.04	4.32	1
5.12	16	6.42	0.3	11.01	2
.	.	.	.	.	.
9.94	26	7.46	0.5	15.51	2
4.1	24	8.11	0.61	2.10	3
.	.	.	.	.	.
5.72	32	6.91	0.02	4.25	3
3.11	12	8.20	0.51	2.08	4
.	.	.	.	.	,
3.56	32	7.41	0.07	2.60	4

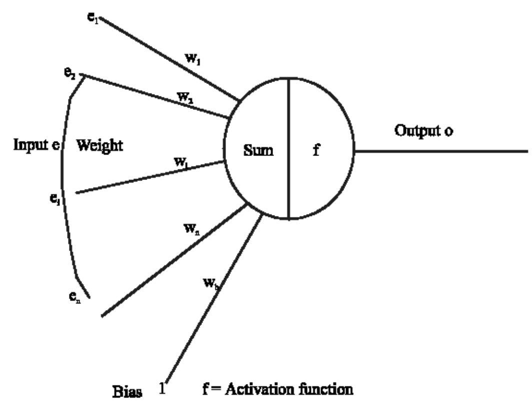

Fig. 1: Neuron model

A layer is a set of elementary neurons. The neural networks used here are basically layers of neurons connected in cascade, with one input layer, one or more hidden layers and one output layer. The input layer is the sensory organ for the Artificial Neural Networks (ANN). Each neuron in a layer is connected to an adjacent neuron layer with different weights. Each neuron, except for the neurons of the input layer, receives signals from the neurons of the previous layer, weighted by the interconnect values between neurons. Consequently the output layer produces an output signal. The calculation of weights is performed with the learning algorithm, which is presented in the next section.

Neural network model: The study consisted of training and testing of the ANN for prediction of outcome in collected water samples. The data on all measured parameters were prospectively collected in a computer database (Microsoft Excel). ANN was applied to provide a nonlinear relationship between inputs variables (Temperature, pH, electrical potential, electrical conductivity and origin of water) and the output variable (number of bacteria).

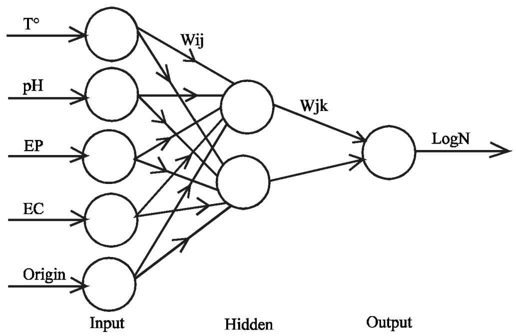

Fig. 2: Topology of five-input

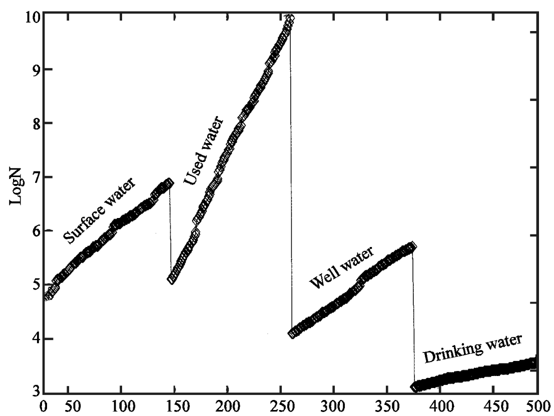

Fig. 3: The bacteria number variations in different varieties of water the values of testing steps are totally confused with the training values

In our study, a network with one hidden layer was selected. The output layer comprised a single neuron corresponding to the value of the dependent variable to be fit to or predicted the number of bacteria. The model used was standard, three layer, back-propagation, neural network with N input nodes, L hidden nodes and K output nodes, which can be mathematically expressed as:

$$O_{pk} = f_1\left(\sum_{j=1}^{L} W^{o}_{jk} f_2\left(\sum_{i=1}^{N} W^{h}_{ij} x_{PI} + b^{l}_{1}\right) + b^{k}_{2}\right)$$
$$\forall k \in 1, 2, \ldots k,$$

were O_{pk} is the output from the k^{th} node of the output layer of the network for the P^{th} vector (data point). X_{pi} the inputs to the network for the P^{th} vector (data point). W^{o}_{jk} the connection weight between the f^{th} node of the hidden layer and the k^{th} node of the output layer (Fig. 1).

W^{h}_{ij} the connection weight between the i^{th} node of the input layer and the j^{th} node of the hidden layer, b^{l}_{1} and b^{l}_{2} are bias terms and $f_1(.)$ and $f_2(.)$ are activation functions. The activation function used was a logistic sigmoid function which produces output in the range 0-1.

Suitable activation function for the hidden units, $f_2(.)$ is needed to introduce non-linearity into the network, which gives the power to capture non-linear relationship between input and output. For the output units, the selection of the activation function $f_1(.)$, is based ont the distribution of the target values[4].

Figure 2 describes the topology of a five-input, one-hidden and one-output (5-1-1 in neural network terminology) neural network model. The four inputs are the temperature (T°), pH, Electrical potential (EP), Electrical conductivity (EC), Origin of the water, surface water, used water, well water, drinking water. Wij and Wjk are weights, which represent the link between the inputs and the output. The weights contain all the information about the network. Therefore, the objective is to train the network to find a series of weights that yield an output signal that has a small error relative to the observed output described by Chen *et al.*[3].

RESULTS AND DISCUSSION

The Artificial neural network system shows that the output-number of bacteria in water corresponding to any variation in the input parameters. The result of the ANN program so far, is a numeric and symbolic terms of number of bacteria in water, using the input data from sensors and output data in the universe of discourse (safety; contaminated; more contaminated). if the number of bacteria in output is around 5 (LogN = 5), the water is contaminated; if it is less than 4, the water is safety; if it is more than 6 it is more contaminated.

Figure 3 presents the plot of the ANN number of bacteria model. These Figure help visualize the

Architecture of program:

```
Neural Network object:
   architecture:

       numInputs: 1
       numLayers: 2
     biasConnect: [1; 1]
    inputConnect: [1; 0]
    layerConnect: [1 0; 1 0]
   outputConnect: [0 1]
   targetConnect: [0 1]

      numOutputs: 1  (read-only)
      numTargets: 1  (read-only)
  numInputDelays: 0  (read-only)
  numLayerDelays: 1  (read-only)

   subobject structures:
          inputs: {1x1 cell}of inputs
          layers: {2x1 cell}of layers
         outputs: {1x2 cell}containing 1 output
         targets: {1x2 cell}containing 1 target
          biases: {2x1 cell}containing 2 biases
    inputWeights: {2x1 cell}containing 1 input weight
    layerWeights: {2x2 cell}containing 2 layer weights
```

```
functions:
        adaptFcn: 'trains'
         initFcn: 'initlay'
      performFcn: 'mse'
        trainFcn: 'traingdx'

   parameters:
      adaptParam: .passes
       initParam: (none)
    performParam: (none)
      trainParam: .epochs, .goal, .lr, .lr_dec,
                  .lr_inc, .max_fail, .max_perf_inc, .mc,
                  .min_grad, .show, .time

   weight and bias values:
            IW: {2x1 cell}containing 1 input weight matrix
            LW: {2x2 cell}containing 2 layer weight matrices
             b: {2x1 cell}containing 2 bias vectors

    other:
        userdata: (user stuff)
```

synthesized from of the predicted model bacteria number. Of the total 1000 observations (500 training and 500 testing), the result shows that the error of measurement is negligible.

CONCLUSION

In acute microbiological detection in fresh water, there are based on simple and easy-to-handle concepts like laboratory methods. But to check each water supply for these contaminants would be a time-consuming job and a qualify operators. In this study, we used ANN modeling techniques in an attempt to predict the microbiological pollution in fresh water. For any predictive instrument to be useful in making a triage decision, an important feature is that only data are readily available from sensors. From the results obtained by this study, ANN model for microbiological contamination of fresh water appear to be a useful tool for future water-testing on pathogenic risk identification, quantification and development of early warning systems for fresh water quality. There room for improvement in parameter selection. We emphasize that our ANN model is not meant to replace or to substitute for an experienced microbiologist; or laboratory investigations; on the contrary, we envisage that the ANN should be viewed as a decision aid for the busy emergency-department, microbiologist, particularly in times of epidemics or water diseases. Moreover, ANN are dynamic and can potentially learn from cumulative experience as the model encounters more and more input variables than conventional logistic regression-based model; this feature supports the hypothesis that ANN are able to process more information that may be inherent in the complex nature of biophysical processes.

REFERENCES

1. Franken, A.A.J.M. and P.S. Van der Zouwen, 1993. Direct and indirect conductimetry for identification and detection of plant pathogenic bacteria. J. Applied Bacteriology, 74: 234-242.
2. Owens, J.D., D.S. Thomas, P.S. Thompson and J.W. Timmerman, 1989. Indirect conductimerty enumeration of microbial populations. Lett. Applied Microbiol., 9: 245-249.
3. Chen, D.G., N.B. Haregreaves, D.M. Ware and Y. Liu, 2000. A Fuzzy logic model with genetic algorithm for analyzing fish stock-recruitment relationships. Can. J. Fish. Aquat. Sci., 57: 1878-1887.
4. Brion, G.M., T.R. Neelakantan and S. Lingireddy, 2002. A neural-network-based classification scheme for sorting sources and ages of fecal contamination in water. Water Res., 36: 3765-3774.
5. Bouchanan, R.L., H.G. Stahi and R.C. Whiting, 1989. Effects and inactivation of temperature, pH, atmosphere, sodium chloride and sodium nitrate on the growth of Listeria monocytogenes. J. Food Prot., 32: 844-851.
6. Cerf, O., K.R. Davey and A.K. Sadoudi, 1996. Thermal inactivation of bacteria. A new predictive model for the combined effects of three environmental factors temperature, pH and water activity. Food Res. Intl., 29: 219-226.
7. Davey, K.R., 1989. A predictive model for combined temperature and water activity on microbial growth during the growth phase. J. Applied Bacteriol., 67: 483-488.
8. Davey, K.R. and B.J. Daughtry, 1995. Validation of a model for predicting the combined effect of three environmental factors on both exponential and log phase of bacterial growth: Temperature, salt concentration and pH. Food Res, Intl., 28: 233-237.