CHAPTER 4

RESULTS AND DISCUSSIONS

Chapter 4

Results and Discussions

Table 4.1 shows the best rule set for dissolve oxygen input data with minimal testing error in 100 runs.

<table>
<tr><th>Lake</th><th>Best Rule Sets</th><th>Condition</th><th>Training error</th><th>Testing error</th></tr>
<tr><td rowspan="2">Putrajaya (R2=0.97)</td><td>RULE SET 1:
IF (ln(|TSS|)>=92.530)
THEN DO=(Ecoli/((Ecoli-10000000000.000)*Chla))</td><td>DO1</td><td rowspan="2">0.943</td><td rowspan="2">0.993</td></tr>
<tr><td>ELSE DO=ln(|(((Cond*(-9.655))+exp(pH))+((pH-(28.402))*Chla))|)</td><td>DO2</td></tr>
<tr><td rowspan="2">Bera (R2=0.90)</td><td>RULE SET 2:
IF (Ecoli<=265.945)
THEN DO=(pH+(TEMP/((TEMP-Ecoli)+(-39.292))))</td><td>DO3</td><td rowspan="2">0.582</td><td rowspan="2">0.417</td></tr>
<tr><td>ELSE DO=ln(|((DS*(-17.562))-((ColiF*NO3)+(-322.330)))|)</td><td>DO4</td></tr>
</table>

Table 4.1 Best rule set generated for both Putrajaya and Bera Lake.

Look at Figure 4.1. The predicted D.O. value for Putrajaya Lake calculated based on RULE SET 1 corresponds very well with the observed data (R^2 = 0.97) even though the magnitude is slightly under estimate at the beginning and over-estimated from September onwards. This minor difference may be due to the average point taken for graph plotting. As for Bera Lake, the predicted value calculated based on RULE SET 2 compared well with the observed value (R^2 = 0.90) (Figure 4.2). The prediction was slightly off from August to October due to the seasonal heavy rainfall reported to occur between September to January every year. Heavy rainfall and massive cloud reduce the potentiality of photosynthesis to occur thus causing the depletion of dissolved oxygen and restoration of carbon dioxide.

Weather factor were not taken into consideration while running this algorithm and this explains the minor disparity when predicting the dissolved oxygen value.

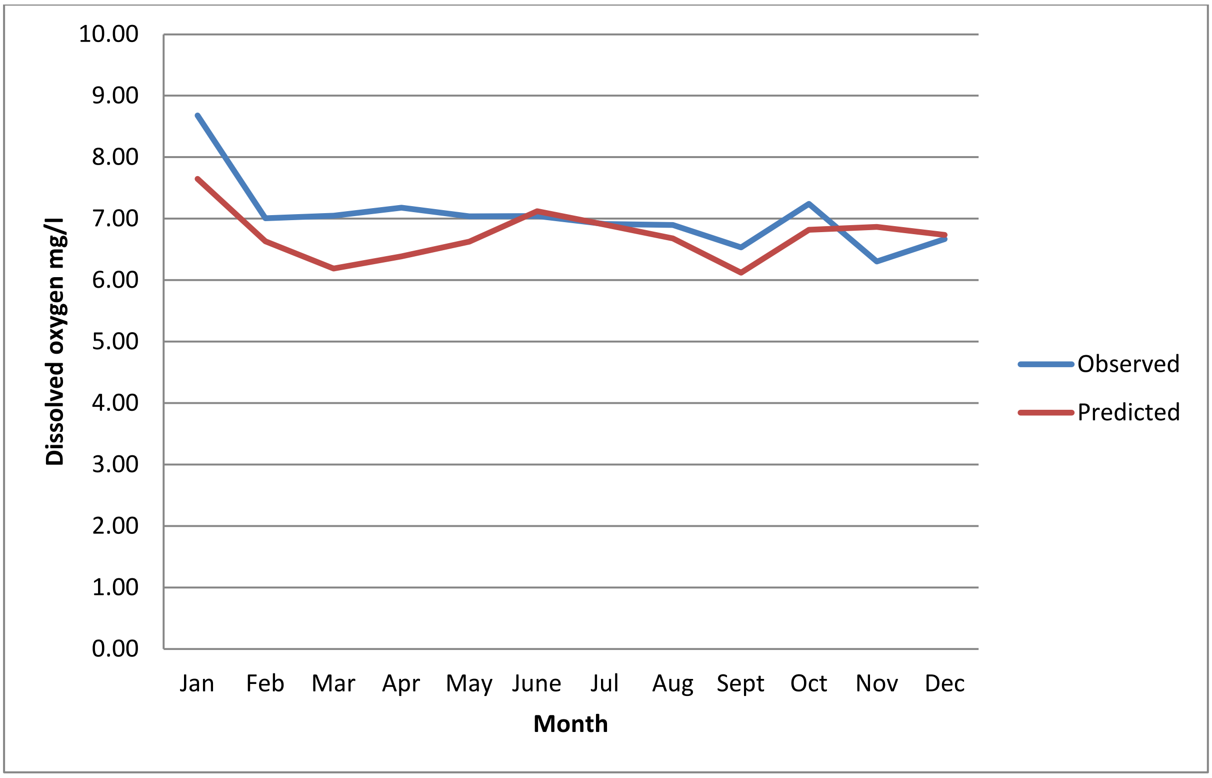

Figure 4.1 Observed and predicted dissolved oxygen value for Putrajaya Lake in 2009.

The forecasting of D.O. dynamics in both lakes is almost similar as the complexity of their rule set does not differ much. Both can be represented as a vector binary tree (T_{IF1}, T_{THEN1}, T_{ELSE1}). The difference was only in terms of parameters manipulated to determine the rule set and the comparison and arithmetic function set generated.

None of data set from Putrajaya Lake fulfils the condition of D01, it can be stated clearly that *E.coli* does not give a significant effect towards the quantity of dissolved oxygen in Putrajaya Lake. Therefore further analysis were done on the T_{ELSE1} (ln(|TSS|)<92.530) branch for condition DO2. Since all the data sets for Putrajaya Lake satisfy the DO2 condition, it indicates that population of dissolved oxygen in this lake are very sensitive to water conductivity, pH and the concentration of chlorophyll-*a*.

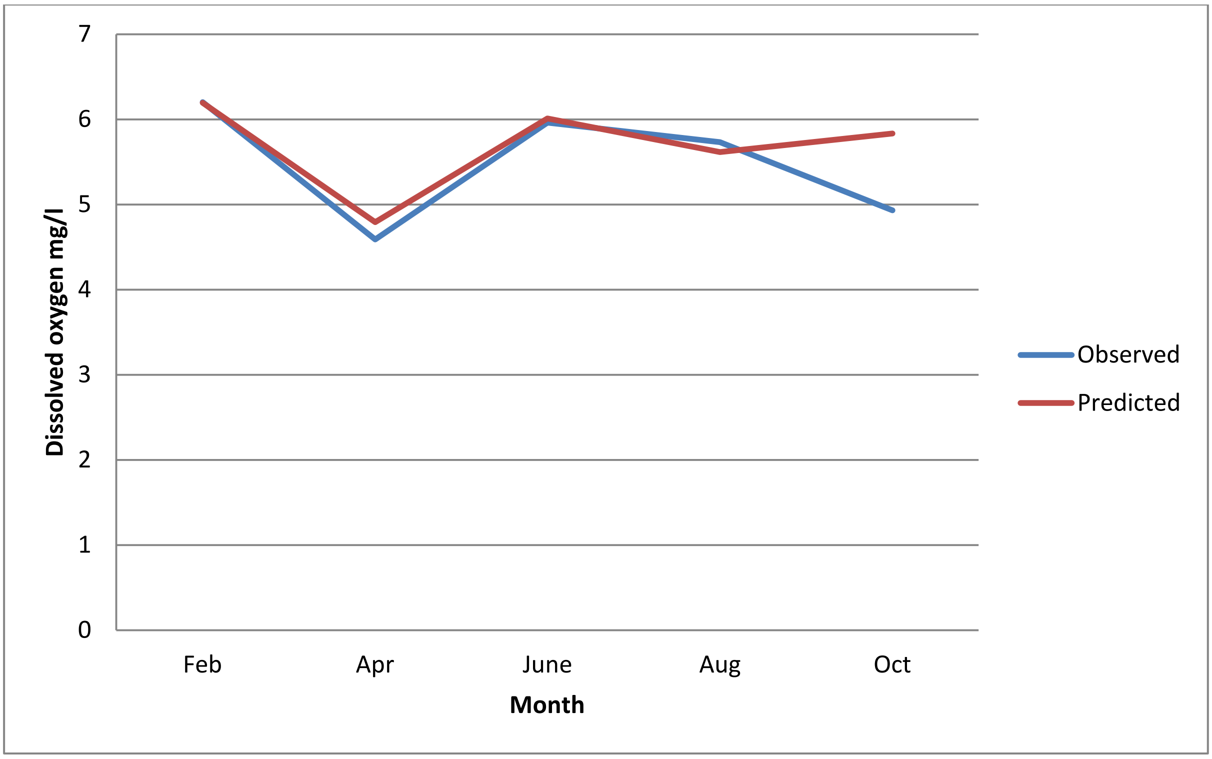

Figure 4.2 Observed and predicted dissolved oxygen value for Bera Lake in 2009.

Each parameter generated in RULE SET 1 and RULE SET 2 was tested individually on how they reflect to the variation of dissolved oxygen within the lake. The way it was tested is such that we normalized the data for parameters that we are looking into. Each of this data was inserted into the rule set generated by HEA and all other parameters were defined by an average value.

4.1 RULE SET 1 (ELSE-BRANCH)

4.1.1 *DO vs. Conductivity*

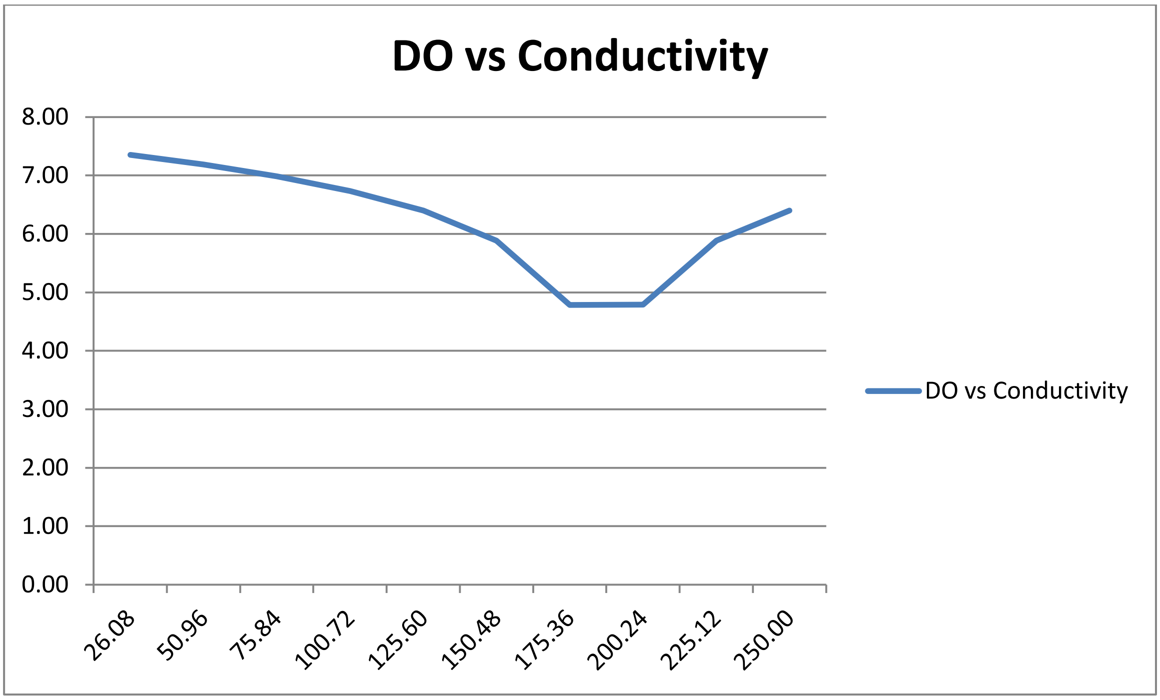

Figure 4.3 Dissolved oxygen vs. Conductivity

Average conductivity = 80.69 µS/cm (Class I INWQS)

Figure 4.3 shows the effect of water conductivity towards the dissolved oxygen concentration. There was reduction of oxygen concentration as the conductivity increased. However the concentration maintained between 175 to 225 µS/cm and rise again forward.

This event can be explained by the ion interaction with the dissolved oxygen in water. Poor conductivity signifies a small amount of ions interact with the dissolved oxygen and dissolved in water. Figure 4.3 illustrate as the conductivity gets higher, the concentration of dissolved oxygen reduced. This is due to more ions interact with the oxygen ions thus diminishing the concentration of dissolved oxygen.

Conductivity is dependent on the activities of ions, therefore we can consider that there is no direct relation between water conductivity and dissolved. Strong bonding between 2 oxygen atoms in oxygen gas allows no ionization of the dissolved oxygen in water bodies.

Conductivity of 80.69 µS/cm average for Putrajaya Lake is acceptable and falls into Class I of Malaysia Interim National Water Quality Standards (INWQS). Existence of ions in the water bodies may be due to the soil-containing nutrients (i.e. ammonium-nitrogen, nitrate-nitrogen and phosphate from fertilizers) and pesticides (insecticides and herbicides) from agricultural lands, parks, and residential areas.

4.1.2 *DO vs. pH*

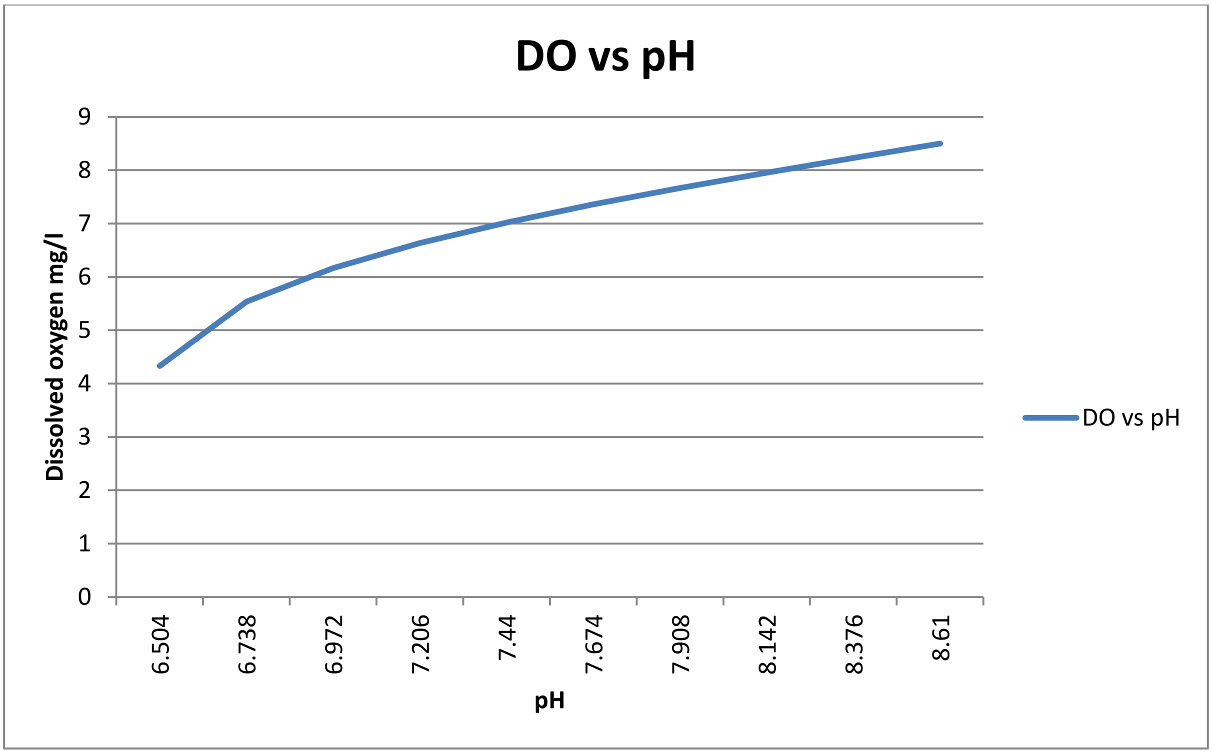

Figure 4.4 Dissolved oxygen vs. pH

pH values of Putrajaya Lake ranging between 6.27 to 8.61 indicating the water to be slightly acidic or alkaline at times. The difference in pH may be due to the variation of water depth designed for each station for Putrajaya Lake where samples were taken. Prescott (1986) stated that as the depth increases, the pH decreases.

Figure 4.4 illustrates the D.O-pH relationship in Putrajaya Lake ecosystem. A report by R. Swaminathan (2005) shows similar result where the oxygen concentration increases steadily as the pH increases. Generally, bacteria and plants can endure alkaline pH better than acidic pH considering the alkalinity was not too extreme. pH between 6.27 and 8.61 are still in range to keep the enzyme for chemical reaction turned on and allow photosynthesis to occur. Photosynthesis activities consequently effect on the high abundance of dissolved oxygen.

4.1.3 *DO vs. Chlorophyll-a*

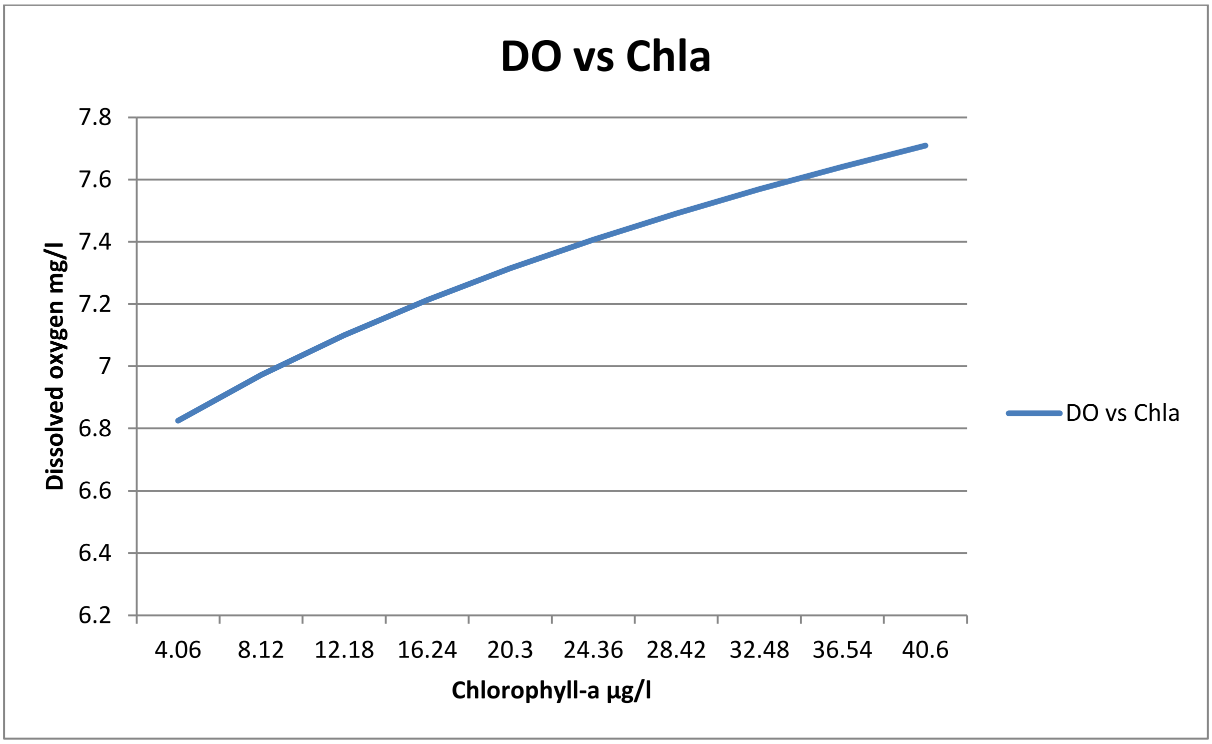

Figure 4.5 Dissolved oxygen vs. Chlorophyll-*a*

The same positive correlation was shown between dissolved oxygen and chlorophyll-*a* concentration. See Figure 4.5. Monitoring Chlorophyl-*a* concentration is a direct way to track the algal growth. High concentration of chlorophyll-*a* indicates high abundance of algae and phytoplankton in the water bodies. Both algae and phytoplankton are photosynthesizing to maintain the dissolved oxygen concentration in water. Dissolved oxygen levels are expected to deplete if chlorophyll conditions are typically high because nutrients such as phosphorus and nitrogen are unexpectedly high under such condition.

4.2 RULE SET 2 (THEN-BRANCH)

As for Bera Lake, most the data sets in this lake satisfy the DO3 condition in the Rule Set 2. We are going to discuss further on how these parameters affected the dissolved oxygen abundance in this lake.

.

4.2.1 *DO vs pH*

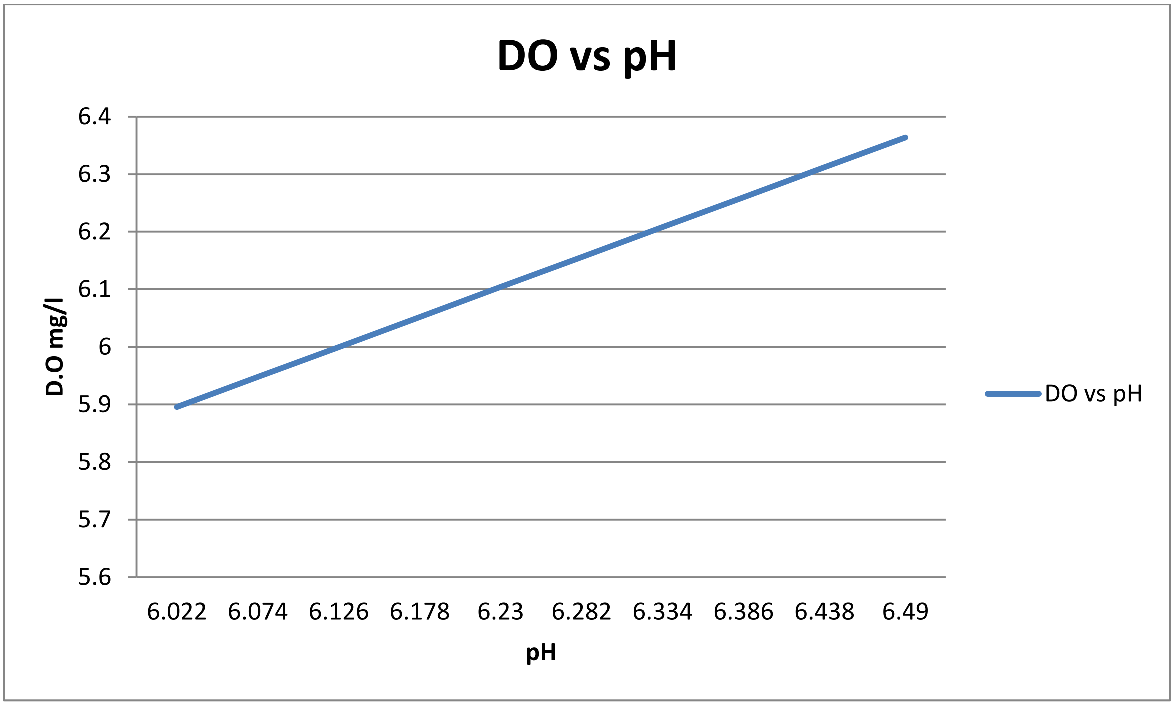

Figure 4.6 Dissolved oxygen vs. pH

pH of range 4.57 and 6.83 was reported by Furtado and Mori (1982) while in 1997, Phang and Murugadas stated the range to be between 5.9 to 6.7 for Bera Lake. For this study a reasonable pH was determined to be from 5.97 to 6.49 verifying slight acidity of Bera Lake's water. Allan (1995) stated that natural causes and anthropogenic inputs as to be the factor of variation in the acidity and alkalinity of the water. The acidic characteristic of Bera Lake was caused by the abundance of decaying plant matter in swamps and peaty area which releases humic acid into the water. Figure 4.6 shows the correlation between pH levels and oxygen concentrations in Bera Lake. This correlation is acceptable as discussed previously for condition DO2.

4.2.2 *DO vs. Temperature*

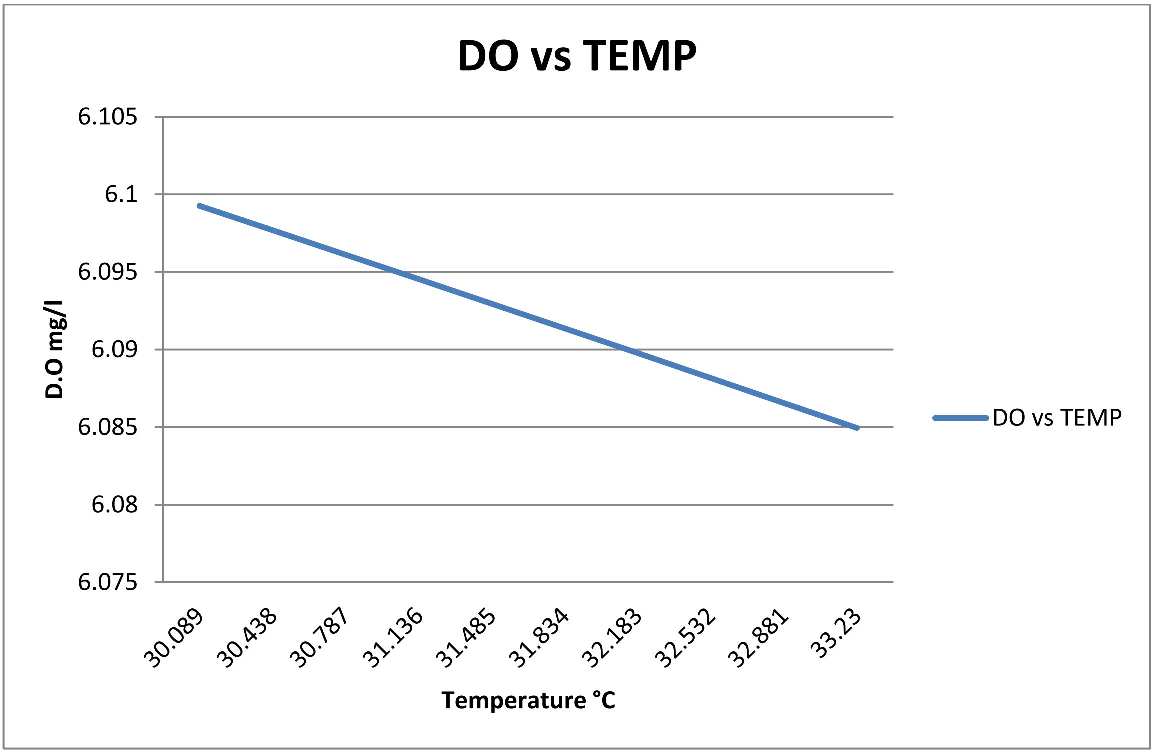

Figure 4.7 Dissolved oxygen vs. water temperature

Study by Phang and Murugadas (1997) recorded that water temperature of Bera Lake is ranging from 26.0^{o}C and 30.0^{o}C for the respective year. This range is slightly lower as compared with recent study which recorded the temperature range between 29.74^{o}C and 33.23^{o}C.According to Allan (1995), temperature of running waters varies on seasonal, daily time scales and among locations. This is due to the climate, elevation, extent of streamside vegetation and the relative groundwater inputs. Maintaining the water temperature is vital because many cellular processes are temperature dependant. Growth of phytoplankton doubles as the temperature rises by 10^{o}C.

Figure 4.7 shows that the concentration of dissolved oxygen lowers as the water temperature increases. This is caused by higher competition of dissolved oxygen among water inhabitants. As the temperature increases, bacteria and fish require more oxygen. However, warmer water is not able to dissolve as much oxygen gas.

4.2.3 *DO vs. E.coli abundance*

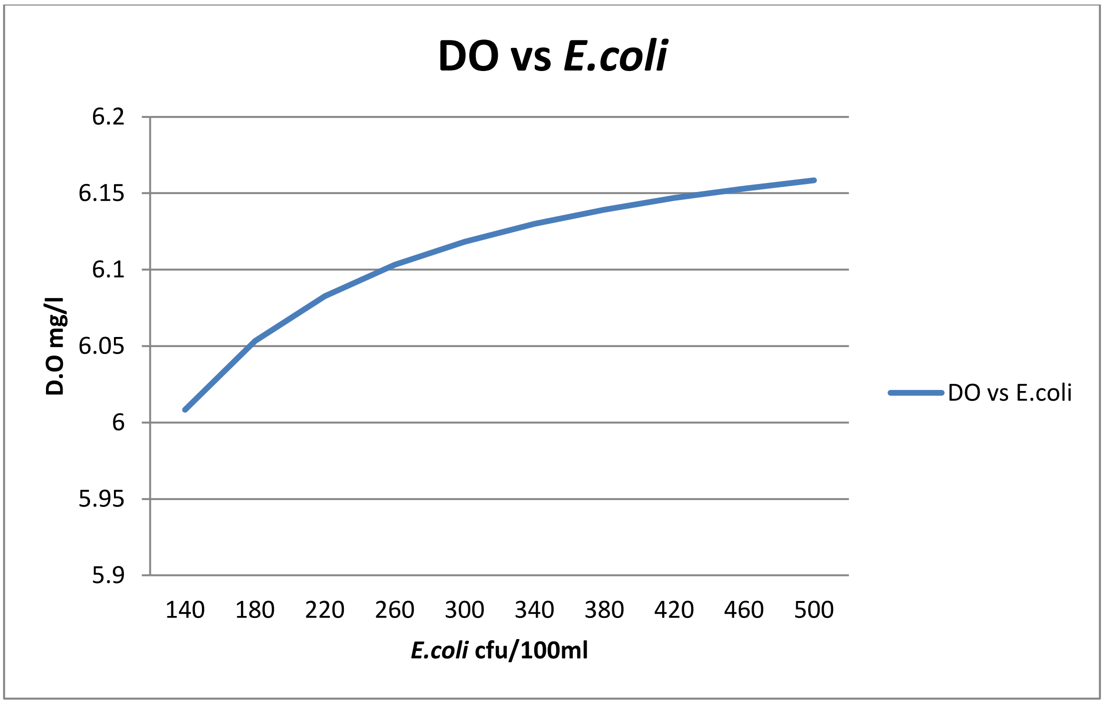

Figure 4.8 Dissolved oxygen vs. abundance of *E.coli*

The results appear to show that there were relationship between oxygen concentration and abundance of *Escherichia coli* in Bera Lake. See Figure 4.8. When the number of *E.coli* increases, the dissolved oxygen concentration shows relative increment. The result is by far a coincidence as stated in an environmental health study conducted at Tamaki Estuary, Auckland - the presence of oxygen should not affect the abundance of *E.coli* as they are facultative anaerobic organisms and therefore a relationship are not expected.

4.3 Comparison with Artificial Neural Network

A different approach was found for similar study of predicting the dissolved oxygen dynamics using the Levenberg-Marquardt algorithm, a Multi Layer Perceptron (MLP) neural network. The study was done in Dusit, one of the district canal in Bangkok, Thailand.

Water quality data used for training were provided by the Department of Drainage and Sewage, Bangkok Administration, for year 2006 until 2008. Parameters involved were almost the same as the parameters used in our research, which has been discussed in the earlier chapter. There are a total of 11 parameters consisting of pH value (pH), biochemical oxygen demand (BOD), chemical oxygen demand (COD), substance solid (SS), total kjeldahl nitrogen (TKN), ammonia nitrogen (NH3N), nitrite nitrogen (NO2N), nitrate nitrogen (NO3N), total phosphorous (TP), total coliform and dissolved oxygen (DO).

These parameters were taken into account in developing the model of the Levenberg-Marquardt algorithm. Performance of the model was evaluated by determining the correlation coefficient (R), mean absolute error (MAE) and a similar measurement used in our research, the root mean square error (RMSE).

This experiment resulted in a correlation which is slightly lower than the HEA technique, R=0.84, the mean square error, MSE=0.78 and the mean absolute error, MEA =

0.7. See the comparison of actual DO value and predicted DO value obtained by the ANN training in Figure 4.9.

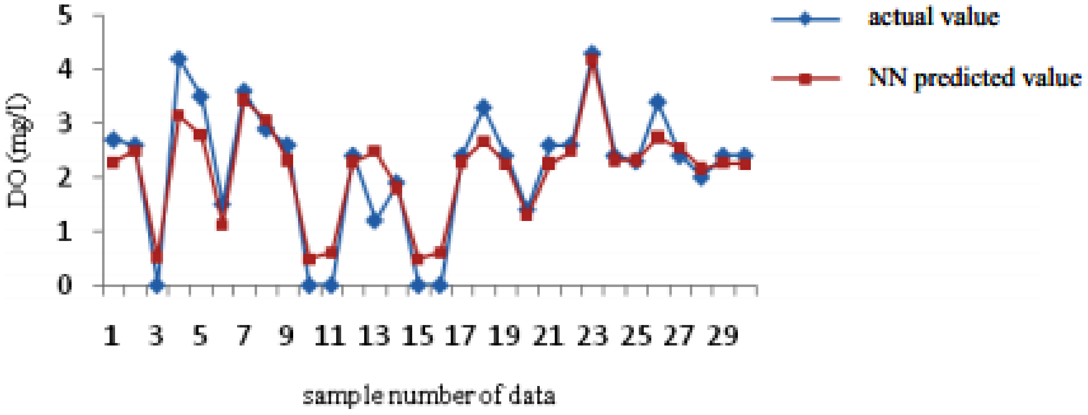

Figure 4.9 Comparison of actual DO value and predicted DO value from the ANN training.

Results obtained from the ANN are almost the same as the results obtained by HEA. However HEA offers some extra advantages in terms of reliability and accuracy, as the R value is much higher in HEA indicating the wellness of data fits to the curve for predicted and observed dissolved oxygen value.

HEA also offers the generation of rule set and the calculation were based only on parameters that are closely related for population under studies, after a series of parameter optimization over generation. These parameters were generated randomly in ANN application thus it's not applicable for real problems which are known not to occur in random. ANN is lack of explicit representation and did not produce a rule set for predicting dynamics evolutionary data. Even if it does, the rule was relatively simple with attributes being associated with constant parameters only. ANN may be used for data sets which are lower in complexity but for higher complexity HEA approach should be taken into consideration.