SIMULATION AND FORECASTING OF SURFACE WATER QUALITY

A Thesis Presented to
The Faculty of the College of Engineering and Technology
Ohio University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

by
Rabah Y. Odeh
March, 1992
This thesis has been approved
for the Department of Civil Engineering
and the College of
Engineering and Technology

Associate Professor of Civil Engineering

Dean of the College of
Engineering and Technology
ACKNOWLEDGEMENTS

With gratitude and constancy, I praise the Almighty Allah for the grace and favors that He bestowed on me.

I wish to extend my genuine appreciation to my advisor, Dr. Tiao J. Chang, for his extensive guidance, assistance, encouragement and support. The assistance from the committee members, Dr. Joseph A. Recktenwald and Dr. David Keck of Ohio University are greatfully acknowledged. The financial support from the Civil Engineering Department is greatly appreciated.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>ACKNOWLEDGEMENTS</strong></td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td><strong>TABLE OF CONTENTS</strong></td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td><strong>LIST OF TABLES</strong></td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td><strong>LIST OF FIGURES</strong></td>
<td>viii</td>
</tr>
<tr>
<td>I</td>
<td><strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>I.1 Water Quality and Stochastic Process</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>I.2 Nature of the Problem</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>I.3 Objective of the Study</td>
<td>3</td>
</tr>
<tr>
<td>II</td>
<td><strong>LITERATURE REVIEW</strong></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II.1 Statistical Analysis of Water Quality</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>II.2 Water Quality Variables</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>II.2.1 Water Temperature</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>II.2.2 Specific Conductance</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>II.2.3 pH</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>II.2.4 Dissolved Oxygen</td>
<td>14</td>
</tr>
<tr>
<td>III</td>
<td><strong>METHODOLOGY</strong></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>III.1 Moving Average Plots</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>III.2 Correlation Analysis</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>III.3 Trend Analysis</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>III.3.1 Graphical Techniques</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>III.3.2 Regression Analysis</td>
<td>20</td>
</tr>
</tbody>
</table>
III.3.3 Simple Linear Regression .......... 21
III.4 Simulation ........................................ 25
   III.4.1 Definitions and Assumptions ...... 25
   III.4.2 Markovian Models ................. 27
III.5 Time Series Analysis ....................... 29
   III.5.1 Modelling Technique & Model Selection ................................. 30
   III.5.2 Autocorrelation, Autocovariance, & Partial Autocorrelation .......... 31
   III.5.3 ARMA Models ......................... 34
      III.5.3.1 Autoregressive (AR) Processes. 34
      III.5.3.2 Moving Average (MA) Processes. 35
      III.5.3.3 Mixed Autoregressive Moving-Average Processes (ARMA) ..... 36
   III.5.4 Diagnostic Checking ................... 37
      III.5.4.1 Test of Independence .......... 37
      III.5.4.2 Normality Test .................. 38
      III.5.4.3 Stationarity ...................... 39
      III.5.4.4 Parsimony of Parameters ...... 40
   III.5.7 Forecasting ......................... 41

IV. APPLICATIONS AND DISCUSSION ...................... 44
   IV.1 Data Acquisition ............................. 44
   IV.2 Graphical Analysis and Moving Average Plots ............................... 45
   IV.3 Correlation Analysis ...................... 49
   IV.4 Trend Analysis ............................. 53
   IV.5 Simulation ................................. 56
   IV.6 Time Series Analysis .......................... 57
IV.6.1 Autocorelation Function and Seasonality ...................... 57
IV.6.2 Standardization of Original Time Series ......................... 58
IV.6.3 Model Identification, Estimation and Checking .................. 59
IV.6.4 Best Model Selection ............................................. 61
IV.6.5 Forecasting ....................................................... 62

V. CONCLUSIONS .................................................................. 100

BIBLIOGRAPHY .................................................................. 103
### LIST OF TABLES

| Table IV.3.1a | Correlation matrix for dry season, Scioto River at Higby, Ohio, 1988 | 63 |
| Table IV.3.1b | Correlation matrix for wet season, Scioto River at Higby, Ohio, 1988 | 63 |
| Table IV.4.1a | Linear long-term trend fitted for different water quality variables, Scioto River at Higby, Ohio | 64 |
| Table IV.4.1b | Linear Short-term trend fitted for different water quality variables, Scioto River at Higby, Ohio | 65 |
| Table IV.5.1 | Average monthly mean, standard deviation, and correlation coefficients of maximum pH for both historical and simulated time series, Scioto River at Higby, Ohio | 66 |
| Table IV.5.2 | Average monthly mean, standard deviation, and correlation coefficients of minimum pH for both historical and simulated time series, Scioto River at Higby, Ohio | 67 |
| Table IV.5.3 | Average monthly mean, standard deviation, and correlation coefficients of maximum specific conductance for both historical and simulated time series, Scioto River at Higby, Ohio | 68 |
| Table IV.5.4 | Average monthly mean, standard deviation, and correlation coefficients of maximum water temperature for both historical and simulated time series, Scioto River at Higby, Ohio | 69 |
| Table IV.5.5 | Average monthly mean, standard deviation, and correlation coefficients of minimum dissolved oxygen for both historical and simulated time series, Scioto River at Higby, Ohio | 70 |
Table IV.5.6 Average monthly mean, standard deviation, and correlation coefficients of streamflow for both historical and simulated time series, Scioto River at Higby, Ohio ....... 71

Table IV.6.1 Characteristics of ARMA models used in simulation of stochastic maximum pH time series, Scioto River at Higby, Ohio ....... 72

Table IV.6.2 Characteristics of ARMA models used in simulation of stochastic minimum pH time series, Scioto River at Higby, Ohio ....... 73

Table IV.6.3 Characteristics of ARMA models used in simulation of stochastic maximum specific conductance time series, Scioto River at Higby, Ohio ................................. 74

Table IV.6.4 Characteristics of ARMA models used in simulation of stochastic maximum water temperature time series, Scioto River at Higby, Ohio ................................. 75

Table IV.6.5 Characteristics of ARMA models used in simulation of stochastic minimum dissolved oxygen time series, Scioto River at Higby, Ohio ................................. 76

Table IV.6.6 Characteristics of ARMA models used in simulation of stochastic streamflow time series, Scioto River at Higby, Ohio .... 77

Table IV.6.7 Selected ARMA models for different water quality and quantity variables and their estimates ................................. 78
LIST OF FIGURES

Figure IV.2.1 10-Day moving average of maximum pH and streamflow time series, Scioto River at Higby, Ohio, 1988

Figure IV.2.2 10-Day moving average of minimum pH and streamflow time series, Scioto River at Higby, Ohio, 1988

Figure IV.2.3 10-Day moving average of maximum specific conductance and streamflow time series, Scioto River at Higby, Ohio, 1988

Figure IV.2.4 10-Day moving average of maximum water temperature and streamflow time series, Scioto River at Higby, Ohio, 1988

Figure IV.2.5 10-Day moving average of minimum dissolved oxygen and streamflow time series, Scioto River at Higby, Ohio, 1988

Figure IV.2.6 10-Day moving average of maximum specific conductance and maximum pH time series, Scioto River at Higby, Ohio, 1988

Figure IV.2.7 10-Day moving average of minimum dissolved oxygen and maximum water temperature time series, Scioto River at Higby, Ohio, 1988

Figure IV.2.7 10-Day moving average of maximum specific conductance and maximum pH time series, Scioto River at Higby, Ohio, 1988

Figure IV.4.1a Linear long-term trend fitted to the average monthly maximum pH time series, Scioto River at Higby, Ohio

Figure IV.4.1b Linear Short-term trend fitted to the average monthly maximum pH time series, Scioto River at Higby, Ohio
Figure IV.4.2a Linear long-term trend fitted to the average monthly maximum Specific conductance time series, Scioto River at Higby, Ohio ................. 84

Figure IV.4.2b Linear short-term trend fitted to the average monthly maximum Specific conductance time series, Scioto River at Higby, Ohio ......................... 84

Figure IV.4.3a Linear long-term trend fitted to the average monthly maximum water temperature series, Scioto River at Higby, Ohio ......................... 85

Figure IV.4.3b Linear short-term trend fitted to the average monthly maximum water temperature time series, Scioto River at Higby, Ohio ......................... 85

Figure IV.4.4a Linear long-term trend fitted to the average monthly minimum dissolved oxygen time series, Scioto River at Higby, Ohio ......................... 86

Figure IV.4.4b Linear short-term trend fitted to the average monthly minimum dissolved oxygen time series, Scioto River at Higby, Ohio ......................... 86

Figure IV.4.5a Linear long-term trend fitted to the average monthly streamflow time series, Scioto River at Higby, Ohio ......................... 87

Figure IV.4.5b Linear short-term trend fitted to the average monthly streamflow time series, Scioto River at Higby, Ohio ......................... 87

Figure IV.5.1a Correlogram of maximum pH for historical time series, Scioto River at Higby, Ohio ............................................. 88

Figure IV.5.1b Correlogram of maximum pH for simulated time series, Scioto River at Higby, Ohio ............................................. 88

Figure IV.5.2a Correlogram of minimum pH for historical time series, Scioto River at Higby, Ohio ............................................. 89
Figure IV.5.2b Correlogram of minimum pH for simulated time series, Scioto River at Higby, Ohio ........................................... 89
Figure IV.5.3a Correlogram of maximum specific conductance for historical time series, Scioto River at Higby, Ohio ............... 90
Figure IV.5.3b Correlogram of maximum specific conductance for simulated time series, Scioto River at Higby, Ohio ............. 90
Figure IV.5.4a Correlogram of maximum water temperature for historical time series, Scioto River at Higby, Ohio .................... 91
Figure IV.5.4b Correlogram of maximum water temperature for simulated time series, Scioto River at Higby, Ohio ................... 91
Figure IV.5.5a Correlogram of minimum dissolved oxygen for historical time series, Scioto River at Higby, Ohio ....................... 92
Figure IV.5.5b Correlogram of minimum dissolved oxygen for simulated time series, Scioto River at Higby, Ohio ..................... 92
Figure IV.5.6a Correlogram of streamflow for historical time series, Scioto River at Higby, Ohio ........................................... 93
Figure IV.5.6b Correlogram of streamflow for simulated time series, Scioto River at Higby, Ohio ........................................... 93
Figure IV.6.1 Correlogram of maximum pH for standardized time series, Scioto River at Higby, Ohio ................................. 94
Figure IV.6.2 Correlogram of minimum pH for standardized time series, Scioto River at Higby, Ohio ................................. 94
Figure IV.6.3 Correlogram of maximum specific conductance for standardized time series, Scioto River at Higby, Ohio ......... 95
Figure IV.6.4 Correlogram of maximum water temperature for standardized time series, Scioto River at Higby, Ohio ............ 95
Figure IV.6.5 Correlogram of minimum dissolved oxygen for standardized time series, Scioto River at Higby, Ohio .................. 96
Figure IV.6.6 Correlogram of streamflow for standardized time series, Scioto River at Higby, Ohio .................. 96
Figure IV.6.7 12-month real time forecasts for maximum pH, Scioto River at Higby, Ohio ............ 97
Figure IV.6.8 12-month real time forecasts for minimum pH, Scioto River at Higby, Ohio ............ 97
Figure IV.6.9 12-month real time forecasts for maximum specific conductance, Scioto River at Higby, Ohio .................. 98
Figure IV.6.10 12-month real time forecasts for maximum water temperature, Scioto River at Higby, Ohio .................. 98
Figure IV.6.11 12-month real time forecasts for minimum dissolved oxygen, Scioto River at Higby, Ohio .................. 99
Figure IV.6.12 12-month real time forecasts for streamflow, Scioto River at Higby, Ohio. 99
CHAPTER I

INTRODUCTION

I.1 Water Quality and Stochastic Processes

Water quality management may be defined as an effort by society to control physical, chemical and biological characteristics of water. Water quality is mainly affected by two factors (Sanders et al., 1983): the activities of society, and the nature of the hydrologic cycle. Both factors can be described as stochastic processes since each factor is affected to a certain extent by the laws of chance or probability. This is particularly true for the effects introduced through the variations of the hydrologic cycle. Therefore, water quality variables can be considered as random variables governed by stochastic processes, where the outcome of any observation is governed by the laws of chance. Although some aspects of the outcomes such as periodicity with day, season or year in some water quality variables are regular or predictable, other parts of the outcomes are not. A stochastic process implies that one is interested in the value of the random variable as a function of time, space, or both.
I.2 Nature of the Problem:

Water Quality and Quantity

The science of hydrology includes studying the distributions of water quantity and water quality in space and time. It is important, therefore, to understand the interactions between water quantity and quality variables as an essential step towards the solution of water quality problems.

The development of water resources require substantial information on water quality and their relationship with water quantity. Expressing water quality in terms of water quantity variables is becoming increasingly important in the water resources area. In order to discuss a process of the water quality, Sanders et al. (1983) suggested the hierarchical division of variables. The water quantity sits at the top of the hierarchical classification of the water quality variables. A water quality variable studied jointly with water quantity would represent a multivariate stochastic process in which water quantity is the primary or basic random variable and water quality is the secondary or associated random variable. The second level in the hierarchical division includes water temperature, pH, BOD, dissolved oxygen, conductivity, turbidity, radioactivity and others. The third and fourth levels consist of a more detailed
I.3 Objective of the Study

In this study, water quality variables of pH, specific conductance, dissolved oxygen and water temperature are studied. An effort is made to relate these water quality variables to water quantity with an emphasis on different behaviors of water quality variables in both wet and dry seasons.

The method of moving average is used to graphically investigate the relationship between water quality and water quantity in addition to the correlation analysis. Trends of water quality during different periods are analyzed in conjunction with the trends of streamflow during the same periods.

In general, historical data of water quality are somewhat short. In order to overcome this problem, simulation techniques were used to generate long record water quality and quantity data. The simulation model preserved the mean, standard deviation and correlation coefficients of the historical data.
It is of interest to forecast the water quality for the management of water resources. The Box and Jenkins processes were used to analyze the time series of varied quality to find the best-fit model. Based on the best-fit models, the real time forecast method was developed to conduct the forecasts of water quality, where confidence limits are provided.
CHAPTER II

LITERATURE REVIEW

II.1 Statistical Analysis of Water Quality

In reviewing the recent studies of water quality in stochastic hydrology, there has been much emphasis on relating water quality variables to the corresponding water quantity. Moreover, statistical analysis of water quality is becoming acceptable as a routine part of water quality management programs. The applicability of the normal distribution, the presence of seasonal patterns and the presence of serial correlation in the water quality data have received the greatest attention (Loftis et al., 1991).

Damsleth (1986) studied the relationship between water acidity and river flow for three rivers in Norway for the period of 1972 - 1982. He used univariate Autoregressive Integrated Moving Average (ARIMA) models to describe the time series of the acidity. Damsleth (1986) concluded that the river acidity is affected by the discharge; the relationship between the two variables was described by a transfer
function, no trends in the water acidity were observed in that study.

In a different study by Morkoc et al. (1989), specific conductivity and total dissolved solids were related to discharge. Historical data of six years were analyzed and linear regression techniques were introduced in the analysis. Results indicated an inverse deterministic relationship between specific conductance and mean daily discharge.

The Autoregressive Moving Average (ARMA) models were used by Hipel et al. (1982), where transfer functions were used for linking a monthly river flow series to precipitation and temperature data, and the Akaike Information Criterion (AIC) was used for the selection of the best model. Furthermore, the cross correlation analysis used in the study indicated no significant relationship between the stream discharge and the water temperature (Hipel et al., 1982).

The Box-Jenkins models were used by Morkoc et al. (1982) for the analysis of water temperature time series. A first order ARIMA model was selected to fit the time series of water temperature. The model was used for forecasting the water temperature up to fifteen days in advance.

Haugh et al. (1986) used the average weekly concentration
data of the total suspended solids, total phosphorus, and the average weekly flow during the period of 1979-1983 at one station in Vermont. Based on a variety of statistical analysis, including time plots, autocorrelation function, partial autocorrelation function, consideration of differencing and transformation, a univariate time series model was selected to fit each set of data. The cross correlations of the univariate residuals from the fitting of water quality and flow series were used to identify the transfer function relating water quality variables to the corresponding flow (Haugh et al., 1986).

The trend analysis of water quality has received a fair amount of attention in the area of water quality research recently due to the associated environmental problems. Whitfield (1986) used the time series of weekly average temperature and Ph data in the period of 1959-1985 to conduct spectral analysis. The results of trend analysis for the data from three different rivers under investigation showed a downward trend of pH series over the study period, where the pH did not show any pronounced seasonality. On the other hand, water temperature data were highly seasonable, i.e., higher in summer and lower in winter, and no obvious trend in water temperature was detected.

Hirsch et al. (1991) suggested the use of the Kendall
non-parametric test for the detection of trends in seasonally varied water quality time series. The method computes the Kendall statistics for each season individually, then sums them over the seasons. For those water quality variables that are highly streamflow-dependent, Hirsch et al. (1991) recommended that the confounding effects of streamflow be removed, based on the analysis of the residuals from a flow-concentration relationship.

Lettenmaier et al. (1991) selected the period of 1978 - 1987 for the analysis of about 400 National Stream Quality Accounting Network (NASQAN) stations. It was found that the concentration was often correlated to the streamflow, and the apparent trends of concentration were in fact the result of streamflow variations. Most ion records were negatively correlated with the streamflow because of the dilution effect. The correlation analysis showed a mostly negative correlation between pH and streamflow. However, correlation coefficients between dissolved oxygen and streamflow were mostly insignificant. Trend analysis for common ions, particularly total dissolved solids showed a general upward trend. For dissolved oxygen deficit, there was a slight predominance of downward trend over upward trend. Finally, for pH and alkalinity, there was a strong predominance of upward trend over downward trend. The period of analysis for one decade was concluded as a reasonable compromise for the trend
In some cases, even if the entire period of water quality data does not show a pronounced trend, it may be of interest to study trends over a short period. Due to the concerns about possible differences in the analytical techniques, Loftis et al. (1991) analyzed the trend of water quality data by breaking the data into different segments, i.e., 1967 - 1970, 1971 - 1977, 1978 - 1983. Trends of different magnitudes and directions were identified during the varied time segments for each water quality variable. Variations of water quality trend for the shorter periods, namely for total phosphorus and total nitrogen data were consistent with those observed in the nutrient loadings during the corresponding time period. In this case, analyzing the entire 20-year period did not give the detailed trend information as those obtained by analyzing the short periods of data.

II.2 Variables of Water Quality

II.2.1 Water Temperature

Water temperature is a relatively new concern because of the increasing use of water for the purpose of cooling by
varied power plants. According to Krenkel and Novotny (1980), water temperature above 15° C are objectionable for drinking water.

Water temperature is one of the most important factors which affects to a certain extent other water qualities. Water temperature affects ion and phase equilibria and influences the rates of biochemical processes that accompany the changes of concentration and the content of organic and mineral substances. Generally, it is known that the rate of chemical reactions depend considerably on changes in the water temperature, on the average of two to three times for every ten degrees Celsius. The reaction rate is related to the water temperature in the following equation (Krenkel and Novotny (1980):

\[ K_T = K_{20} \theta^{(T-20)}, \]

where: \( K_T \) : Reaction rate at temperature \( T \),
\( K_{20} \): Reaction rate at temperature 20 °C,
\( \theta \): Thermal factor (1.029 - 1.106).

Water temperature data are also used to calculate the degree of its saturation by oxygen and other gases. The effect of water temperature on oxygen saturation is
approximated by the following equation (Krenkel and Novotny, 1980)

\[ C_o = 14.562 - 0.41022 T + 0.007991 T^2 - 0.000777 T^3 \quad \text{(II.2)} \]

The shifting of various dynamic equilibria such as concentrations of sulfide, carbonate, degree of alkalinity and electroconductivity are also affected by the changes in water temperature.

II.2.2 Specific Conductance

Specific conductance is determined by the presence of substances that dissociate in cations and anions. The unit of specific conductance is the reverse microhm or the micromhos per one centimeter. The specific conductance is defined as the reciprocal of the resistance of a column of water one centimeter long with a cross section of one squared centimeter. Most portable water has a specific conductance of 50 to 1500 \( \mu \) mhos/cm.

Because the electrical conductivity depends on the ionic concentration, it is logical to expect a relationship between conductance and total dissolved solids (TDS). Different researchers proposed different equations to relate the
specific conductance to the total dissolved solids content, Ledbetter and Gloyna (Krenkel and Novotny, 1980) proposed the following equation:

\[ y = aK + bK^2, \]  

(II.3)

where \( y \): mg/l of dissolved solids,  
\( K \): specific conductance,  
\( a, b \): constants.

Fair et al. (1968) derived the following relationship to relate specific conductance to the TDS content:

\[ \text{TDS (mg/l)} = 4.5 \times 10^{-5} \times (1.02)^{T-25} \times C, \]  

(II.4)

where \( T \): temperature, °C,  
\( C \): conductivity, \( \Omega^{-1} \text{ cm}^{-1} \).

The above equation is valid for pH values from 5 to 9 and water temperatures from 10 to 40 °C with a total dissolved solids content of less than 100 mg/l.

II.2.3 pH

The pH value of natural water is one of the primary
indicators in the evaluation of water suitability for various beneficial uses. The pH value expresses the molar concentration of the hydrogen ion as its negative logarithm. Under natural conditions, the pH value of surface water is specified for the protection of fish life and for the control of chemical reactions such as the dissolution of metal ions in acidic water (IHD-WHO, 1978). Many substances increase in toxicity with changes in pH. For instance, the ammonium ion is shifted to the much more poisonous form of un-ionized ammonia when the pH value of water increases. The criteria set by the U.S. Environmental Protection Agency (USEPA) for the pH values are 6.5 - 9.0 for fresh-water aquatic life, 6.5 - 8.5 for the marine aquatic life, and 5.0 - 9.0 for domestic water supplies (Hammer et al., 1985).

Natural water is usually buffered by a carbon dioxide - bicarbonate system. The acidity represents the amount of carbonic acid presence in water. If the pH value of water is below 8.5, some acidity is present. Though a small variation of the acidity is relatively of low importance from the public health point of view, the excess carbon dioxide may have adverse effects on humans and aquatic life in natural water. The excess acidity may also be harmful to concrete structures because of its corrosive and dissolving properties of free carbon dioxide (Tebbutt, 1983).
On the other hand, alkalinity of water is a measure of its buffering capacity or its ability to neutralize acids or bases. In the unpolluted natural water, the alkalinity is mainly due to the presence of bicarbonate of alkali earth metals. The pH value in such water generally does not exceed 8.3 unless carbonate ions are present (Tebbutt, 1983).

II.2.5 Dissolved Oxygen

Dissolved oxygen (DO) is an important element in the control of water quality. The oxygen-saturated water has a pleasant taste, while the water that lacks dissolved oxygen have an insipid taste (Tebbutt, 1983).

Oxygen presence is essential in maintaining the higher forms of biological life (Krenkel et al., 1980). Because the dissolved oxygen is required by aquatic organisms, its concentration has long been used as an indicator of water quality for aquatic life. Water quality standards have been used for the design of water sewerage plants based on the preservation of dissolved oxygen levels in the receiving water (Hammer et al., 1985).

The oxygen content of water depends on various physical, chemical, biological and microbiological processes. Natural
waters which are in contact with air contain a quantity of oxygen depending on the atmospheric pressure (partial oxygen pressure), temperature, light, and the content of dissolved salt (Tebbutt, 1983).

The content of oxygen is an important indicator of the pollution of a water body, indicating its biological state, the destruction of organic substance, and the intensity of self purification. The USEPA criteria requires that for warm water biological systems, the dissolved oxygen (DO) concentration should be above 5.0 mg/l, assuming its normal variations with time to be above this concentration. For those of cold water, the dissolved oxygen concentration should not be below 7.0 mg/l at any time (Krenkel and Novotny, 1980).
CHAPTER III

METHODOLOGY

III.1 Moving Average Plots

Graphical analysis is the first step in examining the relation between two different variables (McCuen, 1988). The moving average plots (MAP) were used by Chang (1988) to detect the nonhomogeneity of the annual water-loss time series. The same technique can be used to examine the relation between the water quality and quantity variables.

If \( N \) is the length of the time series \( \{Z_i\} \), \( m \) is an odd number \( (m \leq N) \), \( h=(m-1)/2 \), and \( k=(m+1)/2 \), then a new time series \( \{T_n\} \) can be formed, where

\[
T_n = b_1 Z_j + b_2 Z_{j+1} + \ldots + b_m Z_{j+m-1},
\]

(III.1)

\( n=k, \ k+1, \ldots, \ N-h; \ j=n-h; \) and \( b_1, b_2, \ldots, b_m \) are the weights whose sum is equal to one.
III.2 Correlation Analysis

Correlation is a measure of the degree of association between two variables. Correlation coefficients provide a quantitative index of the degree of association. A relationship between two variables can be either linear or nonlinear. The most frequently used correlation is the linear relationship. However, a low linear correlation may result from a nonlinear relationship even when the relationship is obvious (McCuen, 1989).

The correlation coefficient, \( \rho \), is defined as the ratio of the covariance of two different variables \( X \) and \( Y \), to the square root of the product of the variances of \( X \) and \( Y \) (Edwards, 1976), which can be expressed as:

\[
\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}, \tag{III.2}
\]

where \( Cov(X,Y) \) is the covariance of \( X \) and \( Y \); \( Var(X) \) and \( Var(Y) \) are the variances of \( X \) and \( Y \), respectively. The covariance of \( X \) and \( Y \) is defined as:
and the variances of $X$ and $Y$ can be expressed as:

$$Var(X) = E[(X - \mu_X)^2],$$  \hspace{1cm} (III.4)

and

$$Var(Y) = E[(Y - \mu_Y)^2],$$  \hspace{1cm} (III.5)

where $E$ is the expectation, $\mu_X$ and $\mu_Y$ are the population means of $X$ and $Y$.

For a sample of size $N$, the estimation of the correlation coefficient, $r$, is expressed by (Salas et al., 1988):

$$r = \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{X}) (Y_i - \bar{Y}) \frac{1}{S_X S_Y},$$  \hspace{1cm} (III.6)
where \( r \) is the correlation coefficient; \( N \) is the sample size; \( \bar{X} \) and \( \bar{Y} \) are the estimated means and \( S_x, S_y \) are the estimated standard deviations of the variables \( X \) and \( Y \).

### III.3 Trend Analysis

#### III.3.1 Graphical Techniques

Two primary types of trends are considered in this study. First, the step trend which assumes that the data observed before a specific time are from a distinctly different population than those observed after that time. The other trend assumes that the population of the observed data shifts monotonically so that there is no reversal of direction over time, but it does not specify if this occurs continuously, linearly, or in any other specific pattern. Therefore, the step trend is much more specific than the monotonic trend.

In analyzing water quality records, the step trend analysis can be used if there is a known event occurring at a specific time during the period of observation which might result in a definite change in the observed records of water quality (Smith et al., 1991).
Several known techniques can be used to determine trends in the water quality data. The simple way is to graphically plot the observed data. This method can provide the analyst with an eye-view of a possible trend, but cannot determine the degree of the trend. In order to quantify the degree of the trend for the data, it is necessary to express the trend by a mathematical function, (IHD-WHO, 1978).

III.3.2 Regression Analysis

Regression analysis consists of graphical and analytical methods in determining the relation between a dependent variable (response) $Y$ and one or more independent variables $x_1, x_2, \ldots, x_n$ (Mason, 1980). This relation expressed by a mathematical form is called a regression equation. In the case of the relationship between two different variables, it is a regression of $Y$ on $x$.

The simple description of a relationship between two different variables is the linear regression. Let a random sample of size $n$ be denoted by the paired set $\{(x_i, y_i); i=1, 2, \ldots, n\}$. The value of $y_i$ in the ordered pair $(x_i, y_i)$ is a value of some random variable $Y_i$. $Y/x$ is defined as the random variable $Y$ that corresponds to a fixed value $x$ (Myers,
The mean and variance of $Y/x$ are denoted by $\mu_{Y/x}$ and $\sigma^2_{Y/x}$, respectively. The linear regression term implies that $\mu_{Y/x}$ is linearly related to $x$ by the regression equation

$$\mu_{Y/x} = \alpha + \beta x \quad (\text{III.7})$$

The regression coefficients $\alpha$ and $\beta$ can be estimated from the sample. If their corresponding estimates from the sample are denoted by $a$ and $b$, then $\mu_{Y/x}$ can be estimated by $y$ from the sample regression equation

$$\hat{y} = a + bx, \quad (\text{III.8})$$

The estimation of $a$ and $b$ can be obtained by the intercept and the slope of the best-fit line on the $Y$ vs. $X$ plane. Note that $\hat{y}$ represents the estimated value given by the sample regression line while $y$ represents the actual observed value of $Y$ for some $x$ value.

### III.3.3 Simple Linear Regression

Simple linear regression deals with a single independent
variable x and a single dependent variable Y. Y_i may be described by the simple linear regression model (Myers et al., 1985)

\[ Y_i = \mu_{Y|x} + E_i = \alpha + \beta x_i + E_i, \]  

(III.9)

where \( E_i \) is the model error which has a mean of zero. Each observation \((x_i, y_i)\) satisfies the equation

\[ y_i = \alpha + \beta x_i + \delta_i, \]  

(III.10)

where \( \delta_i \) is the value assumed by \( E_i \) when \( Y_i \) takes on the value of \( y_i \). Using the fitted regression line \( \hat{y} = a + bx \), each pair of observations satisfies the relation

\[ y_i = a + bx_i + e_i, \]  

(III.11)

where \( e_i = y_i - \hat{y}_i \) is the residual that describes the error in the fitting of the model at the \( i^{th} \) data point.
The estimates of $\alpha$ and $\beta$, i.e., $a$ and $b$, are to be found such that the sum of the squares of the residuals or errors (SSE) is a minimum. The minimization procedure used is the method of least squares, i.e., to minimize the SSE:

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2, \quad (III.12)$$

By differentiating SSE with respect to $a$ and $b$, the results are:

$$\frac{\partial (SSE)}{\partial a} = -2 \sum_{i=1}^{n} (y_i - a - bx_i), \quad (III.13)$$

and

$$\frac{\partial (SSE)}{\partial b} = -2 \sum_{i=1}^{n} (y_i - a - bx_i)x_i \quad (III.14)$$

By setting the partial derivatives in Eqns. (III.13) and (III.14) equal to zero and rearranging, the following equations are obtained:
The above two equations are solved simultaneously to obtain \( a \) and \( b \) as follows:

\[
a = \frac{\sum_{i=1}^{n} x_i - \left( \sum_{i=1}^{n} x_i \right)^2}{n} \\
b = \frac{\sum_{i=1}^{n} x_i y_i - \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{i=1}^{n} y_i \right)}{n \sum_{i=1}^{n} x_i^2 - \left( \sum_{i=1}^{n} x_i \right)^2},
\]

and

\[
a = \frac{\sum_{i=1}^{n} y_i - b \sum_{i=1}^{n} x_i}{n}
\]

(III.15)  
(III.16)  
(III.17)  
(III.18)
III.4 Simulation

III.4.1 Definitions and Assumptions

Simulation is a powerful tool that has been used in many applications for long term studies. Particularly in water resources engineering, planners use the simulation technique to reduce the design risk that could have been involved if the design is based on the available historical data that are relatively short compared with the design life. A good example is the design of a dam structure to be used for storage or flood control. In such case, it is likely that the length of the historical record of the flow data is shorter than the proposed economic life of the structure. Furthermore, it is highly unlikely for the exact pattern of the record to recur during the life time of the structure. This leads to the conclusion that the worst flood (or drought) based on the records may not be the worst possible flood (or drought) in the design life of the structure. Therefore, it is desirable to use a longer synthetic data base. In such situations, hydrologic simulation may be used for generating synthetic data that preserve the statistics of the historical record.

Furthermore, water resources planners may be interested to know whether a certain water resources system will fail to
meet quality standards. For instance, in studying the pattern of the dissolved oxygen level in a river, the simulated data can be used for an estimation of the expected frequency that the dissolved oxygen fails to meet the standard (Fiering & Jackson, 1981).

It is important to emphasize the fact that simulation is a tool for planning when applied to the hydrologic data. The statistical methods used for generating synthetic data do not pretend to provide casual models for actual flows (Fiering & Jackson, 1981). Nevertheless, simulated data are realistic enough that their use will significantly improve the water resources planning process.

One of the basic assumptions behind the simulation technique is that the synthetic data are considered to be the results of a random process that depends somehow on the elements of chance and probability (Fiering & Jackson, 1981). Therefore, it is not assumed that the actual values of the variable studied can be predicted. However, what one expects is that the general variability of the data is preserved. For example, the basic assumption behind a flow simulation model is that high flows tend to be followed by high flows and low flows tend to be followed by low flows. Therefore, based on the historical data, a simulation model should be constructed by including a random component to reflect the uncertainty of
the process.

III.4.2 Markovian Models

A Markovian process can be used for the purpose of simulation, it takes the following form:

\[ q_i = d_i + e_i, \quad (\text{III.19}) \]

where \( x_i \) is the \( i^{th} \) value, \( d_i \) is the deterministic component and \( e_i \) is the \( i^{th} \) random component in the process, which is assumed to be independently distributed with a mean of zero and a constant variance. Furthermore, the deterministic component takes the linear autoregressive form as (Fiering & Jackson, 1971)

\[ d_i = \beta_0 + \beta_1 q_{t-1} + \beta_2 q_{t-2} + \ldots + \beta_m q_{t-m}, \quad (\text{III.20}) \]

For \( m = 1 \), the model assumes that the process is affected by only one preceding data point. It is reduced to be the first order Markovian model of the form:
\[ q_i = \beta_0 + \beta_1 q_{i-1} + e_i, \]  

\text{(III.21)}

where \( \beta_0 \) and \( \beta_1 \) are constants. For example, for a time series with the mean of \( \mu \), lag-one serial correlation coefficient of \( \rho \) and a variance of \( \sigma^2 \), the process of the following form can be used (Fiering & Jackson, 1971):

\[ q_i = \mu + \rho (q_{i-1} - \mu) + e_i, \]  

\text{(III.22)}

where \( e_i \) is assumed to be normally distributed. The variance of \( q_i \) is expressed by:

\[ \sigma^2 = \mathbb{E}[\mu + \rho (q_{i-1} - \mu) + e_i]^2 - \mu^2 = \rho^2 \sigma^2 - \sigma_e^2, \]  

\text{(III.23)}

where \( \sigma_e^2 \) is the variance of the random components \( e_i \), the variance of the \( q_i \) is related to \( \sigma_e^2 \) by

\[ \sigma_e^2 = \sigma^2 (1 - \rho^2), \]  

\text{(III.24)}
Assuming that \( \{t_i\} \) is a sequence normally distributed, serially independent random variable with a mean of zero and a standard deviation of one, then the process becomes (Fiering & Jackson, 1971):

\[
q_i = \mu + \rho (q_{i-1} - \mu) + t_i \sigma \sqrt{1 - \rho^2}, \quad (III.25)
\]

Such a process can be used for the purpose of simulation by preserving the mean, variance, and first order correlation coefficient of the historic time series.

**III.5 Time Series Analysis**

The analysis of time series is used to describe the characteristics of the historical data for the construction of a mathematical model to represent the time series. The modeling of a time series includes the selection of the type of the model, the identification of the form of the model, the estimation of the model parameters and the diagnostic checking of the model.

Hydrological time series modeling can be divided into two
approaches - deterministic and stochastic. For the deterministic approach, the hydrologic system is expressed by a deterministic model. On the other hand, the stochastic approach describes the system by a probabilistic method. For instance, the unit hydrograph method is a deterministic model for watershed modeling, while the autoregressive models is a stochastic approach in the analysis of hydrologic data (Salas et al., 1988). In the following, the Autoregressive Moving Average (ARMA) process is discussed for the purpose of model building.

III.5.1 Modeling Technique and Model Selection

The characteristics of historical data are important factors in the selection of the model type. The selection of the model type depends on the modeler (Salas et al., 1988). The modeler's knowledge of the advantages and limitations of the different types of models enables him to choose the type of the model. Once the model type has been selected, the form of the model is to be identified. Model identification involves the determination of the order of the model.

Following the model identification, the estimation of the parameters for the model is conducted. The most common parameter estimation techniques are: the method of moments,
the method of least squares and the method of maximum likelihood for the ARMA modeling.

The final step in the modeling process is the diagnostic checking, which includes testing the independence and normality of the residual series for a model. If the selected model does not pass tests of independence, normality and stationarity, the model form or the model type will be modified until a satisfactory model is reached.

III.5.2 Autocovariance, Autocorrelation and Partial Autocorrelation

Assuming that \( x_t, x_{t+k}, \ldots, x_n \) is a stationary time series, the statistical properties of \( n \) observations at the time origin \( t \) are identical to those of \( n \) observations at that of \( t+k \). The mean of the stationary series is defined as

\[
E(x_t) = \mu, \quad (III.26)
\]

The variance of the series is

\[
\gamma_0 = Var(x_t) = E[x_t - E(x_t)]^2, \quad (III.27)
\]
The autocovariance with a time lag \( k \) is defined as the cross product between the series with the time origin at \( t \) and that with the time origin at \( t+k \). The autocovariance function is used to measure the degree of linear auto-dependence of a time series and can be expressed as (Johnson et al., 1976)

\[
\gamma_k = \text{Cov}(x_t, x_{t+k}) = E[x_t - E(x_t)][x_{t+k} - E(x_{t+k})], \quad \text{(III.28)}
\]

The autocorrelation of lag \( k \) refers to the correlation coefficient between the time series with the time origin at \( t \) and that with the time origin at \( t+k \). The autocorrelation can be used as a tool for the identification of the basic pattern that describes the data (Wheelwright et al., 1980). While the correlation coefficient measures the association between two different variables, the autocorrelation coefficient describes the association among the sequences of the same variable at different times. The autocorrelation function is defined as

\[
\rho_k = \frac{\text{Cov}(x_t, x_{t+k})}{\sqrt{\text{Var}(x_t) \cdot \text{Var}(x_{t+k})}} = \frac{\gamma_k}{\gamma_0}, \quad \text{(III.29)}
\]
It is important to notice that the autocorrelation function is dimensionless and is independent of the scale of the observed time series. Since $\gamma_k = \rho_k \gamma_0$, the autocorrelation function reveals equivalent knowledge of the autocovariance function (Nelson, 1973).

The partial autocorrelation coefficient $\phi_{kk}$ is used in an autoregressive process to describe the infinite characteristics of the autocorrelation function. The partial autocorrelation coefficients of a stationary time series satisfy the Yule-Walker equation (Montgomery et al., 1976)

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \cdots + \phi_{kk}\rho_{j-k}, \quad j = 1, 2, \ldots, k \quad (III.30)$$

Hence, this can be used to estimate the partial autocorrelation function.

The knowledge of the autocorrelation and the partial autocorrelation functions plays a significant role in the ARMA modeling for tentatively identifying the model. The comparison made between the theoretical autocorrelation and partial autocorrelation functions and their counterparts of estimations can be used for the selection of model patterns in order to select tentative models.
III.5.3 ARMA Models

III.5.3.1 Autoregressive (AR) Processes

Consider a stationary time series \( \{y_t\} \) which is normally distributed with a mean \( \mu \) and a standard deviation \( \sigma^2 \). The autoregressive model of order \( p \), AR\( (p) \), is generally written as (Salas et al., 1988)

\[
y_t = \mu + \phi_1 (y_{t-1} - \mu) + \ldots + \phi_p (y_{t-p} - \mu) + \delta_t,
\]

or

\[
y_t = \mu + \sum_{j=1}^{p} \phi_j (y_{t-j} - \mu) + \delta_t,
\]

where \( y_t \) is the time dependent variable and \( \delta_t \) is the time independent series which is normally distributed with a mean of zero and a variance of one.

For \( p=1 \), it is the first-order autoregressive model AR\( (1) \), which can be expressed as
A general moving average process $MA(q)$ can be represented by the following equation (Box & Jenkins, 1970; Johnson et al., 1976)

$$y_t = \mu + \phi_1(y_{t-1} - \mu) + \delta_t$$  \hspace{1cm} (III.33)

For the example of $q=1$ which is a special case of $MA(q)$, it can be represented as

$$y_t = \mu + \delta_t - \theta_1 \delta_{t-1}$$  \hspace{1cm} (III.34)

For a second-order moving average process $MA(2)$, the representing equation is

$$y_t = \mu + \delta_t - \theta_1 \delta_{t-1} - \theta_2 \delta_{t-2}$$  \hspace{1cm} (III.35)
III.5.3.3 Mixed Autoregressive Moving-Average (ARMA) Processes

An autoregressive process of order $p$ mixed with a moving average process of order $q$ is denoted as, ARMA($p$, $q$) and can be expressed by (Nelson, 1973)

$$y_t = \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \delta_t - \theta_1 \delta_{t-1} - \ldots - \theta_p \delta_{t-p} \quad \text{(III.37)}$$

In many cases, the inclusion of both AR and MA terms in one model results in a model that has fewer parameters than what would be necessary for a satisfactory model of either AR or MA models alone (Nelson, 1973).

A simple mixed process is the ARMA($1$, $1$) which can be represented by

$$y_t = \phi_1 y_{t-1} + \delta_t - \theta_1 \delta_{t-1} \quad \text{(III.38)}$$
III.5.4 Diagnostic Checking

III.5.4.1 Test of Independence

If a time series \((x_t)\) of size \(N\) is described by an ARMA\((p,q)\) model, where \(p\) is the number of autoregressive parameters, and \(q\) is the number of moving average parameters. To check whether the model is adequate, the Porte Manteau lack of fit test (Box and Jenkins, 1970) is applied to the residual time series \(\{\epsilon_t\}\). The test statistic \(Q\) is represented by

\[
Q = (N) \sum_{k=1}^{L} r_k^2(\epsilon),
\]  

(III.39)

where \(r_k(\epsilon)\) is the correlogram of the residual series \(\{\epsilon_t\}\), \(N\) is the sample size, and \(L\) is the maximum lag considered. The statistic \(Q\) is compared with the chi-square value with a degree of freedom of \((p-q)\) (Salas et al., 1988); if \(Q\) is less than \(\chi^2(p-q)\), then the residual series are independent and the model is adequate; otherwise the hypothesis of independence is rejected.
III.5.4.2 Test of Normality

There are several statistical methods available for testing the hypothesis of normality, such as the skewness and the chi-square tests. The chi-square test is discussed in the following.

To test whether a residual series \( \{\epsilon_t\} \) is normally distributed, the series is arranged in an increasing order, and \( K \) classes are selected with an assumption of the probability of \( 1/K \) for each interval. The test statistic, \( \chi^2 \), is given by (Salas et al., 1988)

\[
\chi^2 = \sum_{i=1}^{K} \frac{(N_i - N/K)^2}{N/K}
\]

(III.40)

where \( N_i \) is the number of the ordered series which fall within the \( i \)th class intervals, and \( N/K \) is the expected number which falls in each interval. The above statistic has a \( \chi^2 \)-distribution with \( K-2 \) degrees of freedom.

For a significance level of \( \alpha \), the statistic \( \chi^2 \) is used to compare with \( \chi^2_{(1-\alpha)} \) (K-2). If \( \chi^2 < \chi^2_{(1-\alpha)} \) (K-2), the hypothesis of normality is accepted; otherwise, it is rejected.
### III.5.4.3 Stationarity

A time series is stationary in the mean if its expected values do not change with time, that is, \( E(X_1) = E(X_2) = \ldots = E(X_t) = \mu \), where \( \mu \) is the mean (Salas et al., 1988). The stationarity in the covariance implies that the covariance depends only on the time lag \( k \), but not on the position \( t \), meaning that \( \text{Cov}(X_t, X_{t-k}) = \text{Cov}(k) \) (Salas et al., 1988). If a time series is stationary in the mean and in the covariance, then it is defined as a stationary time series.

For an AR\((p)\) process to be stationary, the parameters \( \phi_1, \phi_2, \ldots, \phi_p \) must satisfy the stationarity conditions, i.e., the roots \( \{u_i; i=1, 2, \ldots, p\} \) of the characteristic equation,

\[
u^p-\phi_1u^{p-1}-\phi_2u^{p-2}-\ldots-\phi_p=0, \quad (\text{III.41})
\]

must lie within the unit circle (Salas et al., 1988). For instance, the case of an AR\((1)\) model, Eqn. (III.41) becomes

\[
u-\phi_1=0 \Rightarrow u=\phi_1 \quad (\text{III.42})
\]

That is \( |\phi_1| \) should be less than one for an AR\((1)\) model to
meet the stationarity condition. For the case of an AR(2), then \( \phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1, \) and \( -1 < \phi_2 < 1. \)

**III.5.4.4 Parsimony of Parameters**

If several competitive ARMA models pass the tests of diagnostic checking, the criterion of using the minimum number of parameters which is used for the selection of the final model is known as the principle of parsimony.

Akaike developed a mathematical presentation of the principal of parsimony, where the model which gives the minimum AIC value is the best. The Akaike Information Criterion (AIC) for the ARMA modeling is expressed by (Salas et al., 1988)

\[
AIC(p, q) = N \ln(\hat{\sigma}_e^2) + 2(p+q),
\]

where \( p, q \) are the ARMA model parameters, \( N \) is the sample size, \( \hat{\sigma}_e^2 \) is the maximum likelihood estimate of the variance of the residual series.
Forecasting techniques have been widely used in varied management and planning areas such as marketing, financial investment, production and inventory, quality control, as well as water resources engineering (Montgomery et al., 1976). It is of importance to forecast the coming hydrologic events such as high or low flow events in order to reach a decision for a better management of water resources.

All forecasting methods assume that there is a basic underlying pattern representing the historical data. This underlying pattern is used as the basis for future forecasts. Among many well known forecasting methods, the Box & Jenkins method is one of the most widely used in the water resources area (Wheelwright et al., 1980). The method is capable of handling complex time series such as characteristics of varied patterns in the series.

In the Box & Jenkins method, no specific pattern is assumed initially. Instead, tentative models are selected based on the method of identification so that the error resulting from the fitting is minimum. The iterative approach in the modeling procedure allows the modeler to obtain an optimal model in terms of the basic pattern and the minimization of errors (Wheelwright et al., 1980).
The concern of a forecast is to determine the value of \( z_{t+1} \), \( l \geq 1 \), while standing at time \( t \). The forecasting value of \( z_{t+1} \) by an ARMA model can be expressed by (Box & Jenkins, 1972)

\[
z_{t+1} = \phi_1 z_{t+1-1} + \cdots + \phi_p z_{t+1-p} - \theta_1 a_{t+1-1} - \cdots - \theta_q a_{t+1-q} + a_{t+1}, \quad (III.44)
\]

where \( z_{t+1} \) is the forecasting value, \( \phi \)'s are the autoregressive parameters, \( \theta \)'s are the moving average parameters, \( t \) is the origin time, \( l \) is the lead time, and \( p \) and \( q \) are the orders of the ARMA model. The minimum-square error forecast at origin time \( t \) for a lead time \( l \), \( \hat{z}(l) \), is the conditional expectation of \( z_{t+1} \) at time \( t \) (Box & Jenkins, 1972), i.e.,

\[
\hat{z}(l) = E_t[z_{t+1}] = E[z_{t+1} | z_t, z_{t-1}, \ldots], \quad (III.45)
\]

where \( E[z_{t+1} | z_t, z_{t-1}, \ldots] \) is the conditional expectation of \( z_{t+1} \) given the values of \( z_t, z_{t-1}, \ldots \)

For a certain significance level \( \alpha \), the limits of a mean-square-error forecast are provided with a probability of \( (1-\alpha) \) that the actual value of \( z_{t+1} \) will be within the bounds, i.e.,
where \( P \) is the probability that the forecasting values will be within the confidence bounds.

\[
P[z_{t+1}(-) < z_{t+1} < z_{t+1}(+)] = 1 - \alpha, \quad (III.46)
\]
CHAPTER IV

APPLICATIONS AND DISCUSSION

IV.1 Data Acquisition

Data of daily streamflow and water quality, namely, pH, water temperature, specific conductance, and dissolved oxygen, were obtained from the United States Geological Survey (USGS). There are two gaging stations in the Scioto River, a northern tributary of the Ohio River, used for this study. The Higby station has a drainage area of 5,131 square miles and Chillicothe station has a drainage area of 3,849 square miles. A fixed record length for the period of 1972 - 1988 was used for Higby station. Missing daily data were filled by the average value of the preceding and the following daily value. On the other hand, there was a four-year period of missing data for the Chillicothe station, where only the available data were analyzed. Their results are used for discussion only.

Minimum, average and maximum daily values were available for all of the water quality data. However, it is of interest
to investigate the worst scenarios, therefore, the following decisions were made regarding the data used for the study:

1. For specific conductance, maximum daily data were used since the conductivity usually increases with decreasing water quantity.

2. Maximum daily water temperature data were used since higher temperatures lead to more chemical reactions and usually to less dissolved oxygen content. Hence, higher temperature generally results in the deterioration of the water quality.

3. Minimum daily dissolved oxygen data were used since most regulations and standards from the environmental agencies set limitations on the minimum dissolved oxygen concentrations.

4. Both maximum and minimum pH values were used since both extreme situations would adversely affect the water quality.

IV.2 Graphical Analysis and Moving Average Plots

Graphical analysis provides first-hand information of the relationship between the water quality and it's corresponding flow. In order to smooth the variation of daily data, the moving average method was introduced (Chang, 1988). By setting
the weight $B_i$ equal to $(1/10)$ in Eqn. (III.1), the 10-Day moving average method was used to graphically examine the relation between the water quality and the corresponding streamflow. Figures IV.2.1 to IV.2.6 show the moving average plot (MAP) of the water quality and streamflow during the year 1988 for the gaging station at Higby in the Scioto River where a long dry spell occurred from April to September in 1988.

1. pH

(1) An increase in the pH values were noticeable from April until August, where the period from April to August was significantly dry in 1988 as shown in Figures IV.2.1 & IV.2.2.

(2) Peak streamflow discharge in the middle of July resulted in a significant drop of the maximum pH value as shown in Figure IV.2.1. A similar phenomenon was also shown for minimum pH in Figure IV.2.2. This implies a negative correlation between pH and streamflow in the dry season as also shown in Table IV.3.1a, which will be further discussed in the following section.

(3) During the period form January to March, which was relatively wet, the relationship between pH and streamflow was less conclusive.

2. Specific Conductance

The MAP of specific conductance series in Figure IV.2.3 show that there is a significant increase of specific
conductance during the period of low flow. Though there are variations of specific conductance, the decrease of specific conductance generally resulted from an increase of streamflow as shown in the middle of July and those of August, September and November (Figure IV.2.3). This is also further confirmed by the high negative correlation coefficient between the specific conductance and flow in the dry season given in Table IV.3.1a. On the other hand, the increase of streamflow is not always accompanied with a decrease in the specific conductance in the wet period.

3. Water Temperature

Figure IV.2.4 shows the MAP's of the maximum water temperature and the corresponding streamflow, where daily water temperatures were much higher in the dry season than in the wet season. Water temperature is considered as an important factor for its effects on other water quality. For instance, high water temperature will result in an increase of chemical reactions, a reduction in the solubility of gas, and an amplification of taste and odor.

4. Dissolved Oxygen

The MAP's of minimum dissolved oxygen and streamflow are given in Figure IV.2.5, where the low values of dissolved oxygen concentration are generally seen in the dry period. In the middle of July, there is a significant decrease of
dissolved oxygen concentration accompanied with an increase in the streamflow. The explanation of such a phenomenon may be due to the fact that dissolved oxygen concentration can be affected by many other factors such as sunshine and vegetation.

5. pH Vs. Specific Conductance

The MAP of the maximum specific conductance and maximum pH in Figure IV.2.6 shows that there is a fairly agreeable variation pattern between the two variables during the dry period. This can be further confirmed by their relatively high correlation coefficients given in Table IV.3.1a.

6. Minimum DO Vs. Water Temperature

As stated earlier, high water temperatures result in the reduction of the solubility of gases in the water. Such phenomenon is shown in Figure IV.2.7, where high water temperatures are accompanied with low dissolved oxygen concentration in the dry season. This conclusion is further supported by the high correlation coefficient between the water temperature and the dissolved oxygen content as given in Table IV.3.1a.
IV.3 Correlation Analysis

To investigate the relationship between the streamflow and water quality, the data during the year of 1988 from the streamflow gaging station at Higby in the Scioto River are used for the correlation analysis. The year is divided into the wet and dry periods, where the dry period is from May 1st to November 30th and the wet period is from January 1st to April 30th.

The matrix of correlation coefficients is constructed to show the relationship among the streamflow and different water quality variables, namely, maximum pH, minimum pH, maximum specific conductance, maximum water temperature, and minimum dissolved oxygen. Tables IV.3.1a and IV.3.1b are the correlation matrices for the dry and wet seasons, respectively, where the IMSL Subroutine BEMMI was used for the computation. It is shown that there are distinctive relations among varied water quality variables in the different seasons.

1. pH

A correlation coefficient of -0.476 was obtained when the maximum pH data were correlated with their corresponding streamflow in the dry period. This implies that the pH values increase with the decrease of flows. On the other hand, it shows an insignificant correlation coefficient of 0.052
between the maximum pH values and streamflow in the wet period. Minimum pH data show basically a similar pattern as that of the maximum pH with a coefficient of -0.576 and 0.030 in the dry and wet periods, respectively. This leads to a conclusion that the low flow has an inverse effect on pH values. Whitlatch (1989) conversely concluded that there was a good correlation between streamflow and pH.

2. Specific Conductance

When the specific conductance data were correlated to streamflow, they resulted in negative correlation coefficients in both wet and dry seasons. Tables IV.3.1a and IV.3.1b give the value of -0.864 for the dry period and -0.406 for the wet period. This suggests that there is a strong relation between specific conductance and its corresponding flow, especially in the dry period. Morkoc et al. (1989) stated that there was a strong deterministic relationship between streamflow and specific conductance: whenever discharge increases, specific conductance decreases.

3. Water Temperature

Water temperature had a correlation coefficient of 0.236 with discharge in the wet period and 0.021 in the dry season. This indicates no conclusive difference of water temperature related to discharge between the two different seasons. Hipel et al. (1982) also concluded that there was no significant
relationships between water temperature and river flow.

4. Dissolved Oxygen

When the dissolved oxygen data in the dry period were correlated to the corresponding flows, they resulted in a correlation coefficient of -0.307. This indicates an increase in the dissolved oxygen concentration with a decrease of flow. On the other hand, the correlation coefficient for the wet period is 0.011, which is insignificant.

5. pH and Specific Conductance

Both maximum and minimum pH data in the dry period have relatively high and positive correlation coefficients when they are correlated with the specific conductance data. On the other hand, their counterparts of the wet period have very small correlation coefficients. This implies that the pH and specific conductance are strongly correlated in the dry season.

6. pH and Water Temperature

Higher positive correlation coefficients were obtained for the dry season compared to those for the wet season when the maximum and minimum pH data were correlated with the corresponding water temperature. This leads to the conclusion that high water temperatures in the dry season are accompanied with high pH values, though the values of the coefficients are
relatively small.

7. pH and Dissolved Oxygen

Tables IV.1a and IV.1b show that there are lower positive correlations between the pH data and their corresponding dissolved oxygen in the dry season compared to those in the wet season. This concludes that there is a stronger relation between the pH and the dissolved oxygen concentration in the wet period than their counterparts in the dry period.

8. Water Temperature and Dissolved Oxygen

Water temperature has a high negative correlation coefficient with its corresponding dissolved oxygen in the dry season; on the other hand, the two variables have a positive correlation coefficient in the wet season. This implies that there is a very strong inverse relationship between the water temperature and the dissolved content oxygen in the dry season.

9. Water Temperature and Specific Conductance

The correlation coefficient between the specific conductance and the water temperature in the wet season is much greater than that in the dry season.
IV.4 Trend Analysis

To conduct the trend analysis, monthly data are used. Values of monthly water quality data from the Higby station in the Ohio River were plotted against time for the period of 1972-1988. A linear trend was best-fitted to each plot based on the method of least squares. Figures IV.4.1a, IV.4.2a, IV.4.3a, and IV.4.4a show that there are upward trends for the water quality data of pH, specific conductance, water temperature and dissolved oxygen. Figure IV.4.5a gives the plot of the corresponding monthly streamflow during the same time period and shows a strong downward trend as opposed to those of the water quality data.

To check the short term trends, the record period of 1972-1988 was divided into three time segments, i.e., 1972-1977, 1977-1981, and 1981-1988 based on graphical analysis of data. The results of the short term trend analysis are shown in Figures IV.4.1b, IV.4.2b, IV.4.3b, and IV.4.4b. On the other hand, the monthly streamflow data were divided into the same time periods for the short term trend analysis. Figure IV.4.5b further confirms those conclusions obtained in the long term trend analysis with more pronounced results.

Table IV.4.1a lists the best fit equations that resulted from the linear regression analysis for the long-term analysis
of water quality and quantity and their corresponding errors. It is noted that the slopes of water quality data are positive while that of the flow is negative. Regression equations for the short term trend analysis and regression errors are listed in Table IV.4.1b, which further strengthens the above conclusion from the long-term analysis.

For the monthly discharge, there was an overall downward trend for the record period of 1972-1988. On the other hand, pH, specific conductance and dissolved oxygen showed an upward trend during the same time period. This implies that the decrease of flow during the time period studied was accompanied with an increase in the pH and the specific conductance of water. There was also an upward trend in the water temperature over the same time period though it is less significant compared with other water quality variables.

Dividing the time period into shorter segments for the trend analysis reveals the same conclusion that the decrease of streamflow has an adverse effect on water quality. The results of specific conductance, pH, and water temperature show pronounced opposite trends compared to their counterparts of streamflow. Figures IV.4.1b, IV.4.2.b, IV.4.3b, and IV.4.4b are the results of short-term trend analysis and show that there are increased trends of pH, specific conductance and water temperature during the period of 1972-1977.
other hand, Figure IV.4.5b shows a significant decreased trend of flow in the same period. For the period of 1977-1981, the increase in streamflow was accompanied by a decrease in pH and specific conductance and temperature values. Similar results for the periods of 1972-1977 and 1981-1988 can be seen in Figures IV.4.1b, IV.4.2b, IV.4.3b, and IV.4.5b.

For the short-term trend analysis during the period of 1977-1981, the dissolved oxygen content shows a slightly increasing trend while there is an increase of streamflow during the same period. However, during the period of 1981-1988, the decrease of dissolved oxygen was accompanied by a decrease of streamflow.

IV.5 Simulation

Since water quality records are relatively short, it was desirable to obtain their synthetic data. Based on the available monthly records of water quality and quantity, a FORTRAN computer program was designed to generate synthetic data of monthly water quality, i.e., pH, dissolved oxygen, water temperature specific conductance, and streamflow, using the model described in section III.4.2 (Fiering & Jackson, 1981). The simulation model was designed to preserve the statistical characteristics of historical data including mean,
standard deviation, and correlation coefficients.

Monthly mean, standard deviations, correlation coefficients between successive months, and the autocorrelation coefficients of both original and simulated data are calculated for the purpose of comparison. Tables IV.5.1-IV.5.6 give the monthly means, standard deviations, and correlation coefficients for both historical and simulated data. These results show that the simulated time series of water quality and water quantity successfully preserve the statistical properties of the observed data.

Autocorrelation coefficients are plotted against time lag (month) for both historical and simulated data. The autocorrelation plots of the simulated maximum and minimum pH series show similar patterns as those of their corresponding historical records without seasonal effects (Figures IV.5.1a to IV.5.2b). On the other hand, the autocorrelation plots of simulated specific conductance, water temperature, and dissolved oxygen series show seasonal patterns similar to those of the historical data (Figures IV.5.3a to IV.5.6b).

By preserving the statistical characteristics of the historical time series, the simulation model show that it is adequate to generate long-term water quality data for further water quality studies in case of limited available records.
IV.6 Time Series Modeling

IV.6.1 Autocorrelation Function and Seasonality

The information of the autocorrelation function can serve for the purpose of identifying the order of the time series model by comparing the sample correlogram with their theoretical counterparts. For instance, if the autocorrelation coefficients ($\rho_k$) tails off in an exponential manner, it would indicate an AR(1) model as a potential model of the studied time series $y_t$ (Salas et al., 1988).

The autocorrelation function can also be used as a tool to detect the seasonality of a time series. The autocorrelation plots of the historical water quality data of specific conductance, dissolved oxygen, water temperature, and streamflow show pronounced seasonal patterns as shown in Figures IV.5.3a to IV.5.6a. On the other hand, the autocorrelation function plots (correlograms) of the maximum and minimum pH data in Figures IV.5.1a, IV.5.2a do not show significant seasonality.

IV.6.2 Standardization of Original Time Series

There is an obvious seasonal pattern in the time series
of water quality and flow data as described in the previous section. It is desirable to remove such seasonality from the original series before proceeding with the time series modeling. The seasonality can be removed by the transformation of standardization. The transformed series \( \{z_t\} \) are obtained by the following equation:

\[
z_t = \frac{x_t - \mu}{\sigma_t},
\]

(IV.1)

where \( \{x_t\} \) are the original time series, \( \mu \) is the periodic mean and \( \sigma_t \) is the periodic standard deviation (Salas et al., 1988).

The autocorrelation functions of the standardized water quality and quantity data are calculated and graphically shown in Figures IV.6.1 - IV.6.6. Compared with the autocorrelation plots of the original time series in Figures IV.5.1a - IV.5.6a, it can be seen that the seasonal effects in the transformed series of specific conductance, water temperature, dissolved oxygen, and streamflow have been successfully removed by the standardization.
IV.6.3 Model Identification, Estimation and Checking

ARMA models of different orders were identified based on the comparison between the autocorrelation plots of the standardized series and those of the theoretical models. For instance, the specific conductance correlogram in Figure IV.6.3 shows fluctuating positive values of the autocorrelation function which will indicate a positive time dependence, this pattern of autocorrelation results in a tentative model selection of either an AR(1) or an AR(2) model. A similar case can be seen in the correlogram of the water temperature (Figure IV.6.4). The correlogram for the minimum dissolved oxygen (Figure IV.6.5) shows a very strong autocorrelation which suggests a high order, positively dependent model, possibly an AR(2) model. On the other hand, streamflow has a relatively low autocorrelation values (Figure IV.6.6) which leads to a tentative selection of the AR(1) model.

Following the model identification, a computer program was developed to use the following subroutines of the International Mathematical and Statistical Library (IMSL): FTAUTO: estimation of the mean, variance, autocovariance, autocorrelation and partial autocorrelation; FTARPS: preliminary estimation of autoregressive parameters; FTMA: preliminary estimation of moving average parameters; FTML:
maximum likelihood estimation of autoregressive and moving average parameters; GFIT: chi-squared goodness of fit test.

After the preliminary estimation of parameters was carried out by the subroutines of FTARPS and FTMA, the final estimation of parameters was done using the FTML subroutine. Tables IV.6.1 to IV.6.6 list the estimated parameters for different ARMA models fitted to the standardized time series of maximum and minimum pH, maximum specific conductance, minimum dissolved oxygen, maximum water temperature, and streamflow. The white noise variances (variances of the residual series) are also shown in Tables IV.6.1 to IV.6.6. The diagnostic checking of parameters for the goodness of fit was done using the subroutine GFIT. These Tables give the results of the independence and normality tests for different ARMA models.

**IV.6.4 Best Model Selection**

If several ARMA models pass the step of diagnostic checking, i.e., the independence and normality tests, the Akaike Information Criterion (AIC) value is used to select the best model based on the least information. The AIC values of the competitive models that pass the diagnostic checking are calculated as given in Tables IV.6.1 to IV.6.6, where the
model with the least AIC value is selected for each time series. For instance, for the maximum pH time series (Table IV.6.1), the ARMA(1,1) model has the least AIC value among those models that pass both independence and normality tests. Therefore, the model ARMA(1,1) with an autoregressive parameter of 0.647 and a moving average parameter of 0.248 is selected as the best model for the maximum pH time series. On the other hand, for the streamflow time series (Table IV.6.6), though the ARMA(0,1) model has the least AIC value, the model does not pass the independence test, leading to the rejection of the ARMA(1,0). In the same manner, the best ARMA models among different competing models were selected for all other water quality time series and given in Table IV.6.7.

**IV.6.4 Forecasting**

Based on the best ARMA models selected for the time series of water quality and quantity as given in Table IV.6.7, the IMSL routine, FTCAST, was used to conduct the real-time forecasting. The original monthly time series of 1972-1987 (192 months) was used as an input for to forecast the 193rd month value, i.e., the value for January, 1988. Then the actual 193 months data were used to forecast the 194th month which is the value for February, 1988. This process was
continued until 12 monthly values were obtained for each variable of water quality and quantity. Furthermore, the 95% confidence limits for each forecasted value were computed based on the method discussed in section III.5.5.

Actual monthly values averages for the year of 1988 and their corresponding forecasted values and the upper and lower 95% confidence limits are plotted and are shown in Figures IV.6.7 - IV.6.12. It can be seen that the forecasted values are within the bounds of the confidence limits and preserve the patterns of the original series.
Table IV.3.1a Correlation matrix for dry season, 1988.
Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pHmax</td>
<td>1.000</td>
<td>0.956</td>
<td>0.505</td>
<td>0.407</td>
<td>0.178</td>
<td>-0.476</td>
</tr>
<tr>
<td>pHmin</td>
<td>0.956</td>
<td>1.000</td>
<td>0.626</td>
<td>0.267</td>
<td>0.351</td>
<td>-0.576</td>
</tr>
<tr>
<td>max. SC</td>
<td>0.505</td>
<td>0.626</td>
<td>1.000</td>
<td>-0.077</td>
<td>0.334</td>
<td>-0.864</td>
</tr>
<tr>
<td>Temp.</td>
<td>0.470</td>
<td>0.267</td>
<td>-0.077</td>
<td>1.000</td>
<td>-0.710</td>
<td>0.021</td>
</tr>
<tr>
<td>D.O.</td>
<td>0.178</td>
<td>0.351</td>
<td>0.334</td>
<td>-0.710</td>
<td>1.000</td>
<td>-0.307</td>
</tr>
<tr>
<td>SF</td>
<td>-0.476</td>
<td>-0.576</td>
<td>-0.864</td>
<td>0.021</td>
<td>-0.307</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table IV.3.1b Correlation matrix for wet season, 1988.
Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pHmax</td>
<td>1.000</td>
<td>0.965</td>
<td>-0.064</td>
<td>0.150</td>
<td>0.813</td>
<td>0.052</td>
</tr>
<tr>
<td>pHmin</td>
<td>0.965</td>
<td>1.000</td>
<td>0.069</td>
<td>-0.042</td>
<td>0.799</td>
<td>0.030</td>
</tr>
<tr>
<td>max. SC</td>
<td>-0.064</td>
<td>0.069</td>
<td>1.000</td>
<td>-0.885</td>
<td>-0.045</td>
<td>-0.406</td>
</tr>
<tr>
<td>Temp.</td>
<td>0.150</td>
<td>-0.042</td>
<td>-0.885</td>
<td>1.000</td>
<td>0.100</td>
<td>0.236</td>
</tr>
<tr>
<td>D.O.</td>
<td>0.813</td>
<td>0.799</td>
<td>-0.045</td>
<td>0.100</td>
<td>1.000</td>
<td>0.011</td>
</tr>
<tr>
<td>SF</td>
<td>0.052</td>
<td>0.030</td>
<td>-0.406</td>
<td>0.236</td>
<td>0.011</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table IV.4.1a Linear long-term trend fitted for different water quality variables, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Analysis Period</th>
<th>Linear Fit Equation</th>
<th>Regression Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum pH</td>
<td>1972-1988</td>
<td>$y=-33.22+0.0208x$</td>
<td>0.2466</td>
</tr>
<tr>
<td>Specific Conductance</td>
<td>1972-1988</td>
<td>$y=-17532+9.1884x$</td>
<td>92.5617</td>
</tr>
<tr>
<td>Water Temperature</td>
<td>1972-1988</td>
<td>$y=-162.29+0.089x$</td>
<td>8.3638</td>
</tr>
<tr>
<td>Dissolved Oxygen</td>
<td>1972-1988</td>
<td>$y=-284.67+0.148x$</td>
<td>2.1457</td>
</tr>
<tr>
<td>Streamflow</td>
<td>1972-1988</td>
<td>$y=377143-187.86x$</td>
<td>4172.8</td>
</tr>
</tbody>
</table>
Table IV.4.1b Linear Short-term trend fitted for different water quality variables, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Analysis Period</th>
<th>Linear Fit Equation</th>
<th>Regression Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum pH</td>
<td>1972-1977</td>
<td>$y = -107.26 + 0.058x$</td>
<td>0.2232</td>
</tr>
<tr>
<td></td>
<td>1977-1981</td>
<td>$y = 107.91 + 0.0505x$</td>
<td>0.2159</td>
</tr>
<tr>
<td></td>
<td>1981-1988</td>
<td>$y = -2.324 + 0.0053x$</td>
<td>0.2617</td>
</tr>
<tr>
<td>Specific Conductance</td>
<td>1972-1977</td>
<td>$y = -39681 + 20.411x$</td>
<td>66.847</td>
</tr>
<tr>
<td></td>
<td>1977-1981</td>
<td>$y = 93030 - 46.691x$</td>
<td>73.277</td>
</tr>
<tr>
<td></td>
<td>1981-1988</td>
<td>$y = -32168 + 16.563x$</td>
<td>97.893</td>
</tr>
<tr>
<td>Dissolved Oxygen</td>
<td>1972-1977</td>
<td>$y = -298.90 + 0.155x$</td>
<td>2.265</td>
</tr>
<tr>
<td></td>
<td>1977-1981</td>
<td>$y = -15.279 + 0.001x$</td>
<td>2.115</td>
</tr>
<tr>
<td></td>
<td>1981-1988</td>
<td>$y = 251.741 - 0.122x$</td>
<td>2.02</td>
</tr>
<tr>
<td>Water Temperature</td>
<td>1972-1977</td>
<td>$y = -382.73 + 0.202x$</td>
<td>8.364</td>
</tr>
<tr>
<td></td>
<td>1977-1981</td>
<td>$y = 780.88 - 0.388x$</td>
<td>8.752</td>
</tr>
<tr>
<td></td>
<td>1981-1988</td>
<td>$y = -1113.9 + 0.569x$</td>
<td>8.340</td>
</tr>
<tr>
<td>Streamflow</td>
<td>1972-1977</td>
<td>$y = 1.9E+06 - 966.6x$</td>
<td>4340.0</td>
</tr>
<tr>
<td></td>
<td>1977-1981</td>
<td>$y = -2.9E+06 + 1471x$</td>
<td>4230.4</td>
</tr>
<tr>
<td></td>
<td>1981-1988</td>
<td>$y = 687566 - 344.20x$</td>
<td>3866.7</td>
</tr>
</tbody>
</table>
Table IV.5.1 Average monthly mean, standard deviation, and correlation coefficients of maximum pH for both historical and simulated time series, Scioto River at Higby, Ohio, 1988.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Historical</td>
<td>7.98</td>
<td>8.06</td>
<td>7.99</td>
<td>8.01</td>
<td>8.09</td>
<td>8.14</td>
<td>8.23</td>
<td>8.03</td>
<td>7.94</td>
<td>7.99</td>
<td>8.26</td>
<td></td>
</tr>
<tr>
<td>St. Deviation</td>
<td>0.22</td>
<td>0.16</td>
<td>0.20</td>
<td>0.20</td>
<td>0.29</td>
<td>0.37</td>
<td>0.32</td>
<td>0.31</td>
<td>0.27</td>
<td>0.18</td>
<td>0.16</td>
<td>0.21</td>
</tr>
<tr>
<td>Historical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulated</td>
<td>0.16</td>
<td>0.16</td>
<td>0.19</td>
<td>0.20</td>
<td>0.26</td>
<td>0.32</td>
<td>0.32</td>
<td>0.31</td>
<td>0.27</td>
<td>0.19</td>
<td>0.16</td>
<td>0.19</td>
</tr>
<tr>
<td>Corr. Coefficient</td>
<td>0.67</td>
<td>0.58</td>
<td>0.26</td>
<td>0.64</td>
<td>0.73</td>
<td>0.68</td>
<td>0.41</td>
<td>0.61</td>
<td>0.25</td>
<td>0.45</td>
<td>0.64</td>
<td>0.26</td>
</tr>
<tr>
<td>Historical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulated</td>
<td>0.71</td>
<td>0.49</td>
<td>0.24</td>
<td>0.52</td>
<td>0.60</td>
<td>0.69</td>
<td>0.45</td>
<td>0.64</td>
<td>0.36</td>
<td>0.46</td>
<td>0.47</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Table IV.5.2 Average monthly mean, standard deviation, and correlation coefficients of minimum pH for both historical and simulated time series, Sioto River at Higby, Ohio, 1988.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>7.84</td>
<td>7.93</td>
<td>7.83</td>
<td>7.83</td>
<td>7.88</td>
<td>7.89</td>
<td>7.90</td>
<td>7.95</td>
<td>7.81</td>
<td>7.79</td>
<td>8.86</td>
<td>8.14</td>
</tr>
<tr>
<td>Simulated</td>
<td>8.14</td>
<td>8.14</td>
<td>7.96</td>
<td>7.90</td>
<td>7.93</td>
<td>7.94</td>
<td>7.94</td>
<td>7.97</td>
<td>7.82</td>
<td>7.79</td>
<td>7.86</td>
<td>8.15</td>
</tr>
<tr>
<td><strong>St. Deviation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>0.22</td>
<td>0.16</td>
<td>0.24</td>
<td>0.21</td>
<td>0.24</td>
<td>0.30</td>
<td>0.28</td>
<td>0.29</td>
<td>0.24</td>
<td>0.18</td>
<td>0.18</td>
<td>0.20</td>
</tr>
<tr>
<td>Simulated</td>
<td>0.16</td>
<td>0.16</td>
<td>0.21</td>
<td>0.21</td>
<td>0.23</td>
<td>0.26</td>
<td>0.27</td>
<td>0.29</td>
<td>0.24</td>
<td>0.20</td>
<td>0.18</td>
<td>0.19</td>
</tr>
<tr>
<td><strong>Corr. Coefficient</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>0.73</td>
<td>0.49</td>
<td>0.46</td>
<td>0.64</td>
<td>0.68</td>
<td>0.73</td>
<td>0.48</td>
<td>0.56</td>
<td>0.68</td>
<td>0.54</td>
<td>0.44</td>
<td>0.01</td>
</tr>
<tr>
<td>Simulated</td>
<td>0.70</td>
<td>0.64</td>
<td>0.45</td>
<td>0.67</td>
<td>0.77</td>
<td>0.75</td>
<td>0.48</td>
<td>0.53</td>
<td>0.57</td>
<td>0.53</td>
<td>0.54</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Table IV.5.3 Average monthly mean, standard deviation, and correlation coefficients of maximum specific conductance for both historical and simulated time series, Sioto River at Higby, Ohio, 1988.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical Mean</td>
<td>680</td>
<td>646</td>
<td>601</td>
<td>589</td>
<td>642</td>
<td>661</td>
<td>652</td>
<td>700</td>
<td>715</td>
<td>738</td>
<td>701</td>
<td>670</td>
</tr>
<tr>
<td>Simulated Mean</td>
<td>610</td>
<td>609</td>
<td>586</td>
<td>587</td>
<td>641</td>
<td>666</td>
<td>656</td>
<td>704</td>
<td>719</td>
<td>736</td>
<td>700</td>
<td>671</td>
</tr>
<tr>
<td>Historical St. Deviation</td>
<td>109</td>
<td>106</td>
<td>83</td>
<td>68</td>
<td>77</td>
<td>82</td>
<td>93</td>
<td>100</td>
<td>115</td>
<td>103</td>
<td>75</td>
<td>96</td>
</tr>
<tr>
<td>Simulated St. Deviation</td>
<td>108</td>
<td>107</td>
<td>86</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td>85</td>
<td>95</td>
<td>102</td>
<td>102</td>
<td>84</td>
<td>90</td>
</tr>
<tr>
<td>Historical Corr. Coefficient</td>
<td>0.53</td>
<td>0.34</td>
<td>0.36</td>
<td>0.56</td>
<td>0.52</td>
<td>0.81</td>
<td>0.66</td>
<td>0.63</td>
<td>0.57</td>
<td>0.68</td>
<td>0.65</td>
<td>0.48</td>
</tr>
<tr>
<td>Simulated Corr. Coefficient</td>
<td>0.59</td>
<td>0.41</td>
<td>0.45</td>
<td>0.56</td>
<td>0.49</td>
<td>0.77</td>
<td>0.62</td>
<td>0.52</td>
<td>0.57</td>
<td>0.78</td>
<td>0.55</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Table IV.5.4 Average monthly mean, standard deviation, and correlation coefficients of maximum water temperature for both historical and simulated time series, Sioto River at Higby, Ohio, 1988.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>3.17</td>
<td>4.06</td>
<td>8.92</td>
<td>14.1</td>
<td>19.9</td>
<td>23.7</td>
<td>26.0</td>
<td>26.1</td>
<td>23.1</td>
<td>16.6</td>
<td>11.1</td>
<td>5.91</td>
</tr>
<tr>
<td>Simulated</td>
<td>6.01</td>
<td>6.00</td>
<td>10.2</td>
<td>14.6</td>
<td>20.1</td>
<td>23.9</td>
<td>26.1</td>
<td>26.2</td>
<td>23.1</td>
<td>16.6</td>
<td>11.1</td>
<td>5.94</td>
</tr>
<tr>
<td>St. Deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>1.33</td>
<td>1.67</td>
<td>2.38</td>
<td>1.82</td>
<td>1.92</td>
<td>1.44</td>
<td>1.35</td>
<td>1.21</td>
<td>1.09</td>
<td>1.26</td>
<td>1.29</td>
<td>1.34</td>
</tr>
<tr>
<td>Simulated</td>
<td>1.70</td>
<td>1.70</td>
<td>2.09</td>
<td>1.84</td>
<td>1.90</td>
<td>1.60</td>
<td>1.48</td>
<td>1.35</td>
<td>1.05</td>
<td>1.21</td>
<td>1.22</td>
<td>1.34</td>
</tr>
<tr>
<td>Corr. Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>0.68</td>
<td>0.69</td>
<td>0.35</td>
<td>0.41</td>
<td>0.55</td>
<td>0.68</td>
<td>0.65</td>
<td>0.16</td>
<td>0.48</td>
<td>0.32</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td>Simulated</td>
<td>0.72</td>
<td>0.55</td>
<td>0.39</td>
<td>0.43</td>
<td>0.66</td>
<td>0.75</td>
<td>0.73</td>
<td>0.11</td>
<td>0.41</td>
<td>0.24</td>
<td>0.23</td>
<td>0.03</td>
</tr>
</tbody>
</table>
Table IV.5.5 Average monthly mean, standard deviation, and correlation coefficients of minimum dissolved oxygen for both historical and simulated time series, Sioto River at Higby, Ohio, 1988.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>10.8</td>
<td>11.0</td>
<td>10.0</td>
<td>8.82</td>
<td>7.59</td>
<td>6.74</td>
<td>6.43</td>
<td>6.29</td>
<td>7.28</td>
<td>8.80</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td>Simulated</td>
<td>12.1</td>
<td>12.1</td>
<td>10.6</td>
<td>9.07</td>
<td>7.75</td>
<td>6.94</td>
<td>6.59</td>
<td>6.44</td>
<td>7.30</td>
<td>8.82</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td><strong>St. Deviation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>1.29</td>
<td>1.27</td>
<td>1.31</td>
<td>0.90</td>
<td>1.07</td>
<td>1.45</td>
<td>1.58</td>
<td>1.71</td>
<td>1.32</td>
<td>1.44</td>
<td>1.37</td>
<td>1.20</td>
</tr>
<tr>
<td>Simulated</td>
<td>1.28</td>
<td>1.28</td>
<td>1.31</td>
<td>0.97</td>
<td>1.04</td>
<td>1.23</td>
<td>1.37</td>
<td>1.49</td>
<td>1.37</td>
<td>1.41</td>
<td>1.33</td>
<td>1.22</td>
</tr>
<tr>
<td><strong>Corr. Coeff.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>0.81</td>
<td>0.54</td>
<td>0.41</td>
<td>0.57</td>
<td>0.75</td>
<td>0.82</td>
<td>0.84</td>
<td>0.82</td>
<td>0.57</td>
<td>0.77</td>
<td>0.44</td>
<td>0.06</td>
</tr>
<tr>
<td>Simulated</td>
<td>0.83</td>
<td>0.52</td>
<td>0.55</td>
<td>0.56</td>
<td>0.63</td>
<td>0.75</td>
<td>0.79</td>
<td>0.84</td>
<td>0.56</td>
<td>0.77</td>
<td>0.42</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Table IV.5.6 Average monthly mean, standard deviation, and correlation coefficients of streamflow for both historical and simulated time series, Sioto River at Higby, Ohio, 1988.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>5387</td>
<td>7802</td>
<td>9623</td>
<td>8362</td>
<td>5303</td>
<td>4433</td>
<td>2965</td>
<td>2242</td>
<td>2179</td>
<td>2114</td>
<td>3596</td>
<td>6888</td>
</tr>
<tr>
<td>Simulated</td>
<td>9394</td>
<td>9377</td>
<td>10122</td>
<td>8768</td>
<td>6038</td>
<td>5257</td>
<td>3539</td>
<td>3067</td>
<td>3147</td>
<td>2858</td>
<td>4392</td>
<td>7642</td>
</tr>
<tr>
<td>St. Deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>3451</td>
<td>4215</td>
<td>5121</td>
<td>3790</td>
<td>3786</td>
<td>3064</td>
<td>1987</td>
<td>2532</td>
<td>2853</td>
<td>1746</td>
<td>3458</td>
<td>3877</td>
</tr>
<tr>
<td>Simulated</td>
<td>3931</td>
<td>3928</td>
<td>4782</td>
<td>3718</td>
<td>3113</td>
<td>2652</td>
<td>1938</td>
<td>1861</td>
<td>1905</td>
<td>1591</td>
<td>2617</td>
<td>3219</td>
</tr>
<tr>
<td>Corr. Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>0.36</td>
<td>0.26</td>
<td>0.28</td>
<td>0.36</td>
<td>0.36</td>
<td>0.48</td>
<td>0.48</td>
<td>0.54</td>
<td>0.63</td>
<td>0.31</td>
<td>0.60</td>
<td>0.08</td>
</tr>
<tr>
<td>Simulated</td>
<td>0.44</td>
<td>0.19</td>
<td>0.36</td>
<td>0.29</td>
<td>0.33</td>
<td>0.57</td>
<td>0.40</td>
<td>0.29</td>
<td>0.68</td>
<td>0.09</td>
<td>0.39</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Table IV.6.1 Characteristics of ARMA models used in simulation of stochastic maximum pH time series, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter Estimates</th>
<th>White Noise Var.</th>
<th>Independence Test</th>
<th>Normality Test</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\phi_1$ $\phi_2$ $\theta_1$ $\theta_2$</td>
<td></td>
<td>$Q$ Ch.sq P/F</td>
<td>$Q$ Ch.sq P/F</td>
<td></td>
</tr>
<tr>
<td>ARMA(1,0)</td>
<td>0.522 ----- -----</td>
<td>0.7169</td>
<td>10.865 19.68 Pass</td>
<td>5.412 16.92 Pass</td>
<td>-50.893</td>
</tr>
<tr>
<td>ARMA(2,0)</td>
<td>0.462 0.091 ------</td>
<td>0.6992</td>
<td>7.897 18.31 Pass</td>
<td>7.863 16.92 Pass</td>
<td>-70.993</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>0.647 ----- 0.248</td>
<td>0.6839</td>
<td>11.010 18.31 Pass</td>
<td>12.865 16.92 Pass</td>
<td>-75.500</td>
</tr>
<tr>
<td>ARMA(1,2)</td>
<td>0.660 ----- 0.185 0.019</td>
<td>0.6977</td>
<td>18.773 16.92 Fail</td>
<td>17.667 16.92 Fail</td>
<td>-67.416</td>
</tr>
<tr>
<td>ARMA(0,1)</td>
<td>------ ------ -0.414</td>
<td>0.7928</td>
<td>32.656 19.68 Fail</td>
<td>13.647 16.92 Fail</td>
<td>-45.366</td>
</tr>
</tbody>
</table>
Table IV.6.2 Characteristics of ARMA models used in simulation of stochastic minimum pH time series, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter Estimates</th>
<th>White Noise Var.</th>
<th>Independence Test</th>
<th>Normality Test</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>φ1    φ2    θ1    θ2</td>
<td></td>
<td>Q      Ch.sq  P/F</td>
<td>Q      Ch.sq  P/F</td>
<td></td>
</tr>
<tr>
<td>ARMA(1,0)</td>
<td>0.580  ------  ------  ------</td>
<td>0.6575</td>
<td>15.720  19.68 Pass</td>
<td>5.020  16.92 Pass</td>
<td>-83.550</td>
</tr>
<tr>
<td>ARMA(2,0)</td>
<td>0.478  0.162  ------  ------</td>
<td>0.6317</td>
<td>9.568   18.31 Pass</td>
<td>5.412  16.92 Pass</td>
<td>-91.711</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>0.786  ------  0.358  ------</td>
<td>0.5679</td>
<td>13.531  18.31 Pass</td>
<td>2.373  16.92 Pass</td>
<td>-113.425</td>
</tr>
<tr>
<td>ARMA(1,2)</td>
<td>0.841  ------  0.320  0.034</td>
<td>0.5650</td>
<td>19.091  16.92 Fail</td>
<td>12.961 16.92 Fail</td>
<td>-110.467</td>
</tr>
<tr>
<td>ARMA(0,1)</td>
<td>------  ------  -0.427  ------</td>
<td>0.7674</td>
<td>76.099  19.68 Fail</td>
<td>11.588 16.92 Fail</td>
<td>-52.020</td>
</tr>
</tbody>
</table>
Table IV.6.3 Characteristics of ARMA models used in simulation of stochastic maximum specific conductance time series, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter Estimates</th>
<th>White Noise Var.</th>
<th>Independence Test</th>
<th>Normality Test</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\phi_1$ $\phi_2$ $\theta_1$ $\theta_2$</td>
<td>Q Ch. sq P/F</td>
<td>Q Ch. sq P/F</td>
<td>Q Ch. sq P/F</td>
<td></td>
</tr>
<tr>
<td>ARMA(1,0)</td>
<td>0.581 ------ ------ ------ ------</td>
<td>0.6656</td>
<td>15.931 19.68 Pass</td>
<td>5.510 16.92 Pass</td>
<td>-81.055</td>
</tr>
<tr>
<td>ARMA(2,0)</td>
<td>0.604 -0.040 ------ ------ ------</td>
<td>0.6645</td>
<td>16.296 18.31 Pass</td>
<td>5.314 16.92 Pass</td>
<td>-79.387</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>0.529 ------ -0.070 ------ ------</td>
<td>0.6642</td>
<td>39.518 18.31 Fail</td>
<td>9.922 16.92 Pass</td>
<td>-79.461</td>
</tr>
<tr>
<td>ARMA(1,2)</td>
<td>0.694 ------ 0.093 0.144</td>
<td>0.6571</td>
<td>62.346 16.92 Fail</td>
<td>7.667 16.92 Pass</td>
<td>-79.662</td>
</tr>
<tr>
<td>ARMA(0,1)</td>
<td>------ ------ -0.519 ------ ------</td>
<td>0.7176</td>
<td>69.675 19.68 Fail</td>
<td>5.216 16.92 Pass</td>
<td>-65.685</td>
</tr>
</tbody>
</table>
Table IV.6.4 Characteristics of ARMA models used in simulation of stochastic maximum water temperature time series, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter Estimates</th>
<th>White Noise Var.</th>
<th>Independence Test</th>
<th>Normality Test</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\phi_1$ $\phi_2$ $\theta_1$ $\theta_2$</td>
<td></td>
<td>$Q$ Ch.sq P/F</td>
<td>$Q$ Ch.sq P/F</td>
<td></td>
</tr>
<tr>
<td>ARMA(1,0)</td>
<td>0.497</td>
<td>0.7481</td>
<td>5.565 19.68 Pass</td>
<td>9.431 16.92 Pass</td>
<td>-57.204</td>
</tr>
<tr>
<td>ARMA(2,0)</td>
<td>0.435 0.132</td>
<td>0.7315</td>
<td>2.131 18.31 Pass</td>
<td>3.255 16.92 Pass</td>
<td>-59.795</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>0.702</td>
<td>0.7391</td>
<td>24.453 18.31 Fail</td>
<td>8.353 16.92 Pass</td>
<td>-57.678</td>
</tr>
<tr>
<td>ARMA(1,2)</td>
<td>0.599 -0.069</td>
<td>0.7301</td>
<td>19.700 16.92 Fail</td>
<td>8.941 16.92 Pass</td>
<td>-58.164</td>
</tr>
<tr>
<td>ARMA(0,1)</td>
<td>-0.370</td>
<td>0.8172</td>
<td>30.260 19.68 Fail</td>
<td>18.843 16.92 Fail</td>
<td>-39.190</td>
</tr>
</tbody>
</table>
Table IV.6.5 Characteristics of ARMA models used in simulation of stochastic minimum dissolved oxygen time series, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter Estimates</th>
<th>White Noise Var.</th>
<th>Independence Test</th>
<th>Normality Test</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\phi_1$ $\phi_2$ $\theta_1$ $\theta_2$</td>
<td></td>
<td>$Q$ Ch.sq P/F</td>
<td>$Q$ Ch.sq P/F</td>
<td></td>
</tr>
<tr>
<td>ARMA(1,0)</td>
<td>0.687 --- --- ---</td>
<td>0.5370</td>
<td>18.737 19.68 Pass</td>
<td>14.137 16.92 Pass</td>
<td>-124.843</td>
</tr>
<tr>
<td>ARMA(2,0)</td>
<td>0.551 0.201 --- ---</td>
<td>0.5147</td>
<td>11.368 18.31 Pass</td>
<td>11.868 16.92 Pass</td>
<td>-131.489</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>0.839 --- 0.319 ---</td>
<td>0.5062</td>
<td>26.244 18.31 Fail</td>
<td>10.608 16.92 Pass</td>
<td>-134.890</td>
</tr>
<tr>
<td>ARMA(1,2)</td>
<td>0.908 --- 0.416 0.091</td>
<td>0.4978</td>
<td>19.716 16.92 Fail</td>
<td>9.039 16.92 Pass</td>
<td>-136.292</td>
</tr>
<tr>
<td>ARMA(0,1)</td>
<td>--- --- -0.514 ---</td>
<td>0.6853</td>
<td>189.404 19.68 Fail</td>
<td>12.078 16.92 Pass</td>
<td>-75.082</td>
</tr>
</tbody>
</table>
Table IV.6.6 Characteristics of ARMA models used in simulation of stochastic streamflow time series, Scioto River at Higby, Ohio.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter Estimates</th>
<th>White Noise Var.</th>
<th>Independence Test</th>
<th>Normality Test</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>φ1</td>
<td>φ2</td>
<td>θ1</td>
<td>θ2</td>
<td>Q</td>
</tr>
<tr>
<td>ARMA(1,0)</td>
<td>0.430</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>0.8160</td>
</tr>
<tr>
<td>ARMA(2,0)</td>
<td>0.489</td>
<td>-0.136</td>
<td>------</td>
<td>------</td>
<td>0.7969</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>0.171</td>
<td>------</td>
<td>-0.321</td>
<td>------</td>
<td>0.7962</td>
</tr>
<tr>
<td>ARMA(1,2)</td>
<td>0.302</td>
<td>------</td>
<td>-0.193</td>
<td>0.069</td>
<td>0.7929</td>
</tr>
<tr>
<td>ARMA(0,1)</td>
<td>------</td>
<td>------</td>
<td>-0.491</td>
<td>------</td>
<td>0.7963</td>
</tr>
</tbody>
</table>
Table IV.6.7  Selected ARMA models for different water quality and quantity variables and their estimates.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Selected Model</th>
<th>Model Parameters</th>
<th>( \phi_1 )</th>
<th>( \phi_2 )</th>
<th>( \theta_1 )</th>
<th>WNV</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>maximum pH</td>
<td>ARMA(1,1)</td>
<td></td>
<td>0.6476</td>
<td>----</td>
<td>0.2477</td>
<td>0.6839</td>
<td>-75.500</td>
</tr>
<tr>
<td>minimum pH</td>
<td>ARMA(1,1)</td>
<td></td>
<td>0.7861</td>
<td>----</td>
<td>0.3581</td>
<td>0.5680</td>
<td>-113.425</td>
</tr>
<tr>
<td>specific conductance</td>
<td>ARMA(1,0)</td>
<td></td>
<td>0.5806</td>
<td>----</td>
<td>----</td>
<td>0.6656</td>
<td>-81.055</td>
</tr>
<tr>
<td>dissolved oxygen</td>
<td>ARMA(2,0)</td>
<td></td>
<td>0.5509</td>
<td>0.2006</td>
<td>----</td>
<td>0.5147</td>
<td>-131.489</td>
</tr>
<tr>
<td>water temperature</td>
<td>ARMA(2,0)</td>
<td></td>
<td>0.4348</td>
<td>0.1324</td>
<td>----</td>
<td>0.7315</td>
<td>59.795</td>
</tr>
<tr>
<td>streamflow</td>
<td>ARMA(1,0)</td>
<td></td>
<td>0.4301</td>
<td>----</td>
<td>----</td>
<td>0.8160</td>
<td>-39.478</td>
</tr>
</tbody>
</table>
Figure IV.2.1 10-Day moving average of maximum pH and streamflow time series, Scioto River at Higby, Ohio, 1988.

Figure IV.2.2 10-Day moving average of minimum pH and streamflow time series, Scioto River at Higby, Ohio, 1988.
Figure IV.2.3 10-Day moving average of maximum specific conductance and streamflow time series, Scioto River at Higby, Ohio, 1988.

Figure IV.2.4 10-Day moving average of maximum water temperature and streamflow time series, Scioto River at Higby, Ohio, 1988.
Figure IV.2.5 10-Day moving average of minimum dissolved oxygen and streamflow time series, Scioto River at Higby, Ohio, 1988.

Figure IV.2.6 10-Day moving average of maximum specific conductance and maximum pH time series, Scioto River at Higby, Ohio, 1988.
Figure IV.2.7 10-Day moving average of minimum dissolved oxygen and maximum water temperature time series, Scioto River at Higby, Ohio, 1988.

Figure IV.2.8 10-Day moving average of maximum specific conductance and maximum pH time series, Scioto River at Higby, Ohio, 1988.
Figure IV.4.1a Linear long-term trend fitted to the average monthly maximum pH time series, Scioto River at Higby, Ohio.

Figure IV.4.1b Linear Short-term trend fitted to the average monthly maximum pH time series, Scioto River at Higby, Ohio.
Figure IV.4.2a Linear long-term trend fitted to the average monthly maximum Specific conductance time series, Scioto River at Higby, Ohio.

Figure IV.4.2b Linear short-term trend fitted to the average monthly maximum Specific conductance time series, Scioto River at Higby, Ohio.
Figure IV.4.3a Linear long-term trend fitted to the average monthly maximum water temperature series, Scioto River at Higby, Ohio.

Figure IV.4.3b Linear short-term trend fitted to the average monthly maximum water temperature time series, Scioto River at Higby, Ohio.
Figure IV.4.4a Linear long-term trend fitted to the average monthly minimum dissolved oxygen time series, Scioto River at Higby, Ohio.

Figure IV.4.4b Linear short-term trend fitted to the average monthly minimum dissolved oxygen time series, Scioto River at Higby, Ohio.
Figure IV.4.5a Linear long-term trend fitted to the average monthly streamflow time series, Scioto River at Higby, Ohio.

\[ y = 377143 - 187.862 \times x \]

Figure IV.4.5b Linear short-term trend fitted to the average monthly streamflow time series, Scioto River at Higby, Ohio.
Figure IV.5.1a Correlogram of maximum pH for historical time series, Scioto River at Higby, Ohio.

Figure IV.5.1b Correlogram of maximum pH for simulated time series, Scioto River at Higby, Ohio.
Figure IV.5.2a Correlogram of minimum pH for historical time series, Scioto River at Higby, Ohio.

Figure IV.5.2b Correlogram of minimum pH for simulated time series, Scioto River at Higby, Ohio.
Figure IV.5.3a Correlogram of maximum specific conductance for historical time series, Scioto River at Higby, Ohio.

Figure IV.5.3b Correlogram of maximum specific conductance for simulated time series, Scioto River at Higby, Ohio.
Figure IV.5.4a Correlogram of maximum water temperature for historical time series, Scioto River at Higby, Ohio.

Figure IV.5.4b Correlogram of maximum water temperature for simulated time series, Scioto River at Higby, Ohio.
Figure IV.5.5a Correlogram of minimum dissolved oxygen for historical time series, Scioto River at Higby, Ohio.

Figure IV.5.5b Correlogram of minimum dissolved oxygen for simulated time series, Scioto River at Higby, Ohio.
Figure IV.5.6a Correlogram of streamflow for historical time series, Scioto River at Higby, Ohio.

Figure IV.5.6b Correlogram of streamflow for simulated time series, Scioto River at Higby, Ohio.
Figure IV.6.1 Correlogram of maximum pH for standardized time series, Scioto River at Higby, Ohio.

Figure IV.6.2 Correlogram of minimum pH for standardized time series, Scioto River at Higby, Ohio.
Figure IV.6.3 Correlogram of maximum specific conductance for standardized time series, Scioto River at Higby, Ohio.

Figure IV.6.4 Correlogram of maximum water temperature for standardized time series, Scioto River at Higby, Ohio.
Figure IV.6.5 Correlogram of minimum dissolved oxygen for standardized time series, Scioto River at Higby, Ohio.

Figure IV.6.6 Correlogram of streamflow for standardized time series, Scioto River at Higby, Ohio.
Figure IV.6.7 12-month real time forecasts for maximum pH, Scioto River at Higby, Ohio.

Figure IV.6.8 12-month real time forecasts for minimum pH, Scioto River at Higby, Ohio.
Figure IV.6.9 12-month real time forecasts for maximum specific conductance, Scioto River at Higby, Ohio.

Figure IV.6.10 12-month real time forecasts for maximum water temperature, Scioto River at Higby, Ohio.
Figure IV.6.11 12-month real time forecasts for minimum dissolved Oxygen, Scioto River at Higby, Ohio.

Figure IV.6.12 12-month real time forecasts for streamflow, Scioto River at Higby, Ohio.
CHAPTER V

CONCLUSIONS

The relationship between the water quality and the streamflow was graphically investigated using the moving average plots (MAP), where the 10-Day moving average method has been used. The results showed that using the MAP's, different water quality variables and the corresponding discharge can be successfully related. It was found that there was a significant increase of the pH, specific conductance and water temperature whenever the corresponding discharge decreased in the dry season. Furthermore, a peak discharge would have resulted in a significant drop of the maximum pH, maximum specific conductance or maximum water temperature.

Correlation analysis was used to further investigate the relationship between the water quality and the streamflow mathematically. Correlation coefficients were estimated for both dry and wet seasons, respectively. The results showed that correlation coefficients between the water quality and the water quantity in the dry season are higher than those in
the wet season. For minimum pH, maximum pH, and maximum specific conductance, the correlation coefficients with the corresponding streamflow in the dry season were much higher than those in the wet season. On the other hand, water temperature and dissolved oxygen correlations with streamflow were not conclusive.

How the water quality changes with time is of interest. The monthly average time series of the water quality and streamflow where used in the trend analysis. Based on the graphical analysis shown in Figures IV.4.1a - IV.4.6a, it is found that the downward trend for the pH, specific conductance, and water temperature occur whenever an upward trend occurs in the streamflow. On the other hand, they also show that an upward trend in pH, specific conductance and water temperature is accompanied with a downward trend in the streamflow.

Furthermore, linear regression was used to quantify the trend analysis (Tables IV.4.1a, and IV.4.1b). It was found that the slopes of the pH, specific conductance, water temperature, and dissolved oxygen are positive while that of the flow is negative.

To solve the problem of the short length of water quality records which is usually encountered in most related studies,
the Markov simulation method was introduced to simulate the synthetic data of water quality. The simulated data preserved the characteristics of the historical records, i.e., means, standard deviations, and correlation coefficients. Based on the results, it is found that for the simulation of synthetic data for water quality, the Markovian simulation method is successful in accomplishing the preservation of the mean, standard deviation, and correlation coefficients of the historical data.

It is of interest to forecast the water quality for the purpose of water resources management. The real-time forecasting method was developed based on the selected best ARMA model for each water quality variable, where the four-step modeling procedure, Identification, Estimation, Diagnostic Checking, and Best model selection were used.

The forecasted values for varied water quality variables were plotted to compare with the actual values and the 95% confidence limits were provided in each case. Based on the comparison, it could be seen that the forecasting models successfully preserve the pattern of the data.
BIBLIOGRAPHY


