PART 1

SCOPE

AND

BACKGROUND MATERIAL

The **objectives** of Part I of the book are to explain the important roles that time series modelling has to play in environmental decision making and to provide definitions for some basic statistical concepts that are used in time series modelling. As can be seen in the Table of Contents given at the start of the book, Part I consists of the following two chapters which are entitled:

CHAPTER 1 - ENVIRONMETRICS, SCIENCE AND DECISION MAKING

CHAPTER 2 - BASIC STATISTICAL CONCEPTS

The first chapter furnishes the basic **motivations** for writing a book on time series modelling of water resources and environmental systems as well as pointing out the import of **time series modelling in science and decision making**. Chapter 2 presents a variety of basic **statistical definitions** that are utilized in the subsequent chapters in the book.

Consider now in more detail some of the main contributions of the first two chapters, starting with Chapter 1. As explained in Section 1.1, this book on time series modelling is actually a document about **environmetrics** - the development and application of statistical methods in the environmental sciences. Because environmental data sets usually consist of observations measured over time, **time series models** constitute important statistical tools for use in environmetrics. In fact, the time series and other statistical methods presented in the book draw upon research developments from two areas of environmetrics called **stochastic hydrology** and **statistical water quality modelling** as well as research contributions from the field of **statistics**. As pointed out in Section 1.2, the use of statistical techniques can enhance the **scientific method** which in turn means that pressing environmental problems can be more efficiently and expeditiously solved. When carrying out a scientific data analysis study using environmental data such as hydrological and water quality time series, one can employ both exploratory data analysis and confirmatory data analysis tools. The purpose of **exploratory data analysis** is to use simple graphical methods to uncover the basic statistical characteristics of the data which can be modelled formally at the **confirmatory data analysis** stage utilizing time series models and other kinds of statistical methods. For example, in an **environmental impact assessment** study, exploratory graphs may clearly indicate the presence of **trends** in a water quality time series due to land use changes. The trends can then be modelled and their magnitudes estimated at the confirmatory data analysis stage using the **intervention model** of Part VIII. Section 1.3 outlines how a time series model, such as an intervention model, can be systematically fitted to a data set by following the identification, estimation and diagnostic check steps of **model construction**. By keeping in mind the basic **physical system** within which a data analysis is being carried out, one can put the overall environmental problem into proper perspective. Section 1.4 explains why the **hydrological cycle** provides a good physical structure for the types of environmental systems studied in the applications in this book. By executing a **scientific data analysis** study based upon a sound environmental system framework, one can improve environmental decision making. Section 1.5 provides a description of **engineering decision making** and explains how it can be enhanced using proper data analysis studies. Next, Section 1.6 describes the **organization** of the book and suggests various sequences of chapters that can be followed according to the needs and backgrounds of the readers. Finally, Section 1.7 describes a **decision support system,** called the **McLeod-Hipel Time Series Package,** that permits a user to take immediate advantage of the many statistical techniques presented in the book. The book should be useful for **teachers, students, researchers and practitioners** who are interested in confronting challenging data analysis problems arising in water resources and environmental engineering.

In Chapter 2, **basic statistical definitions** that are needed in time series modelling are presented. First, the different **kinds of time series** that can arise in practice are discussed. After briefly explaining what is meant by a **stochastic process**, the concepts of **stationarity** and **nonstationarity** are described. This is followed by a variety of specific statistical definitions including the **autocorrelation function** for describing linear dependence among observations in a time series. Although the time series modelling and analysis carried out in this book are mainly done in the **time domain**, some contributions from **spectral analysis** are discussed in Section 2.6.

CHAPTER 1

ENVIRONMETRICS,

SCIENCE

AND

DECISION MAKING

1.1 THE NEW FIELD OF ENVIRONMETRICS

The **overall objectives** of this book are to present flexible statistical methodologies for scientifically carrying out data analysis studies of environmental time series and to describe a broad variety of useful statistical tools for implementing these methodologies. The methodologies include general procedures for systematically executing a data analysis study as well as the main steps required for fitting a specific statistical model to a data set. Because environmental data are almost always available as observations measured over time, most of the particular tools presented in the book consist of different kinds of time series models. However, other statistical methods such as informative graphical techniques, regression analysis and nonparametric tests are also discussed. Finally, the main types of environmental time series that are used for demonstrating how to apply the procedures and techniques consist of hydrological observations such as riverflows, precipitation and temperature series, as well as many different kinds of water quality series measured in rivers and lakes.

In fact, the contents of this book fall within a relatively new and dynamic academic discipline called **Environmetrics**. The term Environmetrics was first coined by J.S. Hunter on January 27, 1976, at a meeting of the Committee on National Statistics held at the National Academy of Sciences in Washington, D.C., and it is defined as the development and application of statistical methodologies and techniques in the environmental sciences (Hunter, 1990). The environmetrics approaches and techniques given in this book are based upon research results developed largely in the areas of statistics, stochastic hydrology and statistical water quality modelling. This book presents pertinent developments from these fields in a systematic and coherent fashion under the unifying umbrella of environmetrics. Furthermore, the title of the book reflects the fact that the time series modelling and other procedures given in the book should be especially useful for scientists, engineers and applied statisticians studying water resources and environmental systems. Nonetheless, students, teachers, practitioners, and researchers working in many other fields where time series models are applied may find much of the material to be quite helpful for addressing many different kinds of data analysis problems.

As brief illustrations of the usefulness and importance of environmetrics, consider an application from **statistical water quality modelling** and another one from stochastic hydrology. Figure 1.1.1 displays a graph of 72 average monthly phosphorous observations (in milligrams per litre) from January, 1972, until December, 1977, for measurements taken by the Ontario Ministry of the Environment downstream from the Guelph sewage treatment plant located on the Speed River in the Grand River basin, Ontario, Canada. Notice in this figure that the abscissae along the X axis represent time, where the monthly observations are numbered sequentially from 1 to 72. The ordinates along the Y axis give the values of the phosphorous concentrations in mg/l. For easier interpretation of the graph, the measurements are plotted as small circles at discrete points in time and are joined by straight lines. In February, 1974, a pollution abatement procedure was brought into effect by implementing conventional phosphorous treatment at the Guelph station. Observe in Figure 1.1.1 the manner in which the man-made intervention of phosphorous removal has dramatically decreased the mean level of the series after the intervention date. Moreover, as indicated by the blackened circles in this figure, there are missing data points both before and after the intervention date. For displaying a missing value on the graph, the missing observation is replaced by its monthly average across all of the years. However, estimating a missing monthly value by a specified monthly mean may not be an accurate procedure since the autocorrelation or dependence structure inherent in the time series is ignored and the influence of the intervention is not considered. It is explained in Chapter 19 how the **intervention model** can be used not only to estimate the missing observations where the autodependence structure among measurements is automatically taken into account but also to model statistically the effects of the tertiary phosphorous treatment for reducing the mean level of the series. In Section 19.4.5, intervention analysis is employed for realistically modelling the water quality time series of Figure 1.1.1 by constructing an appropriate intervention model. The intervention model fitted to the series in Figure 1.1.1 shows that there is a 75% drop in the mean level of the series where the 95% confidence interval is from 71% to 78%. Rigorous statistical statements like this are extremely useful in **environmental impact assessment** studies. Besides Chapter 19, other **trend analysis** procedures and applications are given in Chapters 22 to 24 of Part X.

As a second demonstration of the efficacy of environmetrics, an application from stochastic hydrology is utilized. **Stochastic hydrology** arose in the early 1960's in the field of water resources and it deals with the application of stochastic and time series models to hydrological time series (Maas et al., 1962). Because simulated sequences from time series models fitted to riverflow series are used in the design and operation of systems of reservoirs, stochastic hydrology is also referred to as **synthetic or operational hydrology.** The water quantity application involves an interesting water resources systems problem that was solved in Brazil by Silva et al. (1984). In 1984, hydroelectric plants accounted for 85% of Brazil's installed electrical capacity of 40,000 MW. In order to optimize the generation of power from a vast complex of hydroelectric plants, various types of linear multivariate time series models were used to model and simulate flows into the reservoirs (Pereira et al., 1984). To coordinate the most economical operation of both the hydrothermal generating system and the hydroelectric system, stochastic dynamic programming was used (Pereira and Pinto, 1985). Silva et al. (1984) clearly demonstrated that their use of time series models as well as other systems science techniques for optimally operating the huge electrical system clearly saved the country about $87 million U.S. in five years. Because of this great practical accomplishment, in 1985 the Institute of Management Science (TIMS) awarded the authors second prize at the 14th Annual Competition for the Edelman Award for Management Science Achievement. Time series models similar to those used in the Brazilian study are presented in Parts VII to IX in this book.

There are many specific reasons why one may employ time series models in environmental engineering. For instance, in the first application referred to above, the intervention model is utilized for quantifying the magnitude of a step trend in a water quality time series. In the

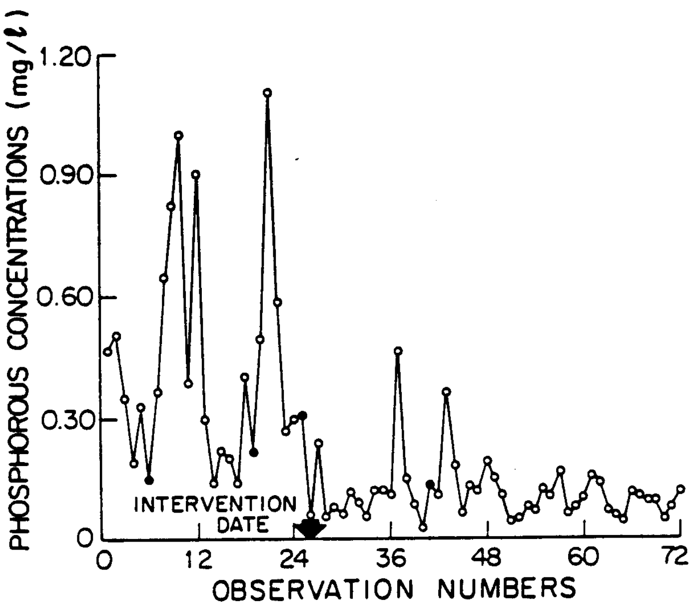

Figure 1.1.1. Monthly phosphorous data for the Speed River near Guelph, Ontario, Canada.

second application, simulated sequences and forecasts from multivariate time series models are used for maximizing profits by economically operating a complex system of hydroelectric power plants. However, there are also some very general advantages and uses for employing time series models, and, for that matter, most other types of mathematical models. Firstly, a model provides a common **communication medium** by which scientists, engineers, statisticians and other decision makers can realistically discuss an environmental problem. For instance, the graph in Figure 1.1.1 geometrically displays the obvious step drop in the phosphorous series due to the tertiary treatment. Additionally, as shown in Section 19.4, the intervention model fitted to this series accurately models this trend as well as other characteristics of the data such as autocorrelation and a pure random component. By examining the estimate for one of the parameters in the model, interested parties can see how the magnitude of the improvement in phosphorous levels is quantified. This type of representation of information and accompanying communication lead to a second important benefit of formal modelling - **understanding**. By discussing a problem with others and using mathematical models as a means of communication, the ultimate result is a better understanding of the problem by everyone concerned. In spite of the fact that the data in Figure 1.1.1 contain some randomness, missing values and a step trend, one can sort out these components using graphs and an intervention model, and, thereby, better understand what is happening. Moreover, a clearer understanding of the problem ultimately leads to **improved decision making**. Suppose, for instance, environmental authorities require that there be a 90% drop in the phosphorous levels in the river. The results of the intervention study clearly demonstrate that more intensive tertiary treatment would be required to reduce the phosphorous concentrations. Alternatively, if the environmental regulators want only an 77% reduction, one could argue that this level is almost achieved, since the 95% confidence interval for the estimated 75% reduction is from 71% to 78% and 77% is contained within this interval.

In summary, the modelling process can lead to better communication and understanding which eventually can result in improved decisions being made. However, inherent in this argument is that the mathematical model being used properly models the physical phenomena that are being studied. When modelling nature within the context of the scientific method, one should always employ a mathematical model that is realistically designed for capturing the key characteristics of the physical process being examined. One should always strive to **design a mathematical model to fit the physical problem** and never try to distort the physical process to fit a given mathematical model.

In the next section, it is explained how statistical modelling can enhance the **scientific method**. Subsequently, the main components of a **statistical scientific investigation** are pointed out and general approaches to **data analysis** are discussed. For systematically fitting a time series model to a given data set, an overall systems design approach to **model construction** is presented in Section 1.3. The **hydrological cycle** is discussed in Section 1.4 as a basic physical structure for describing the kinds of **environmental systems** studied in this book. The importance and role of environmetrics in **environmental decision making** are pointed out in Section 1.5. The **organization** of the book is thoroughly explained in Section 1.6, along with suggestions of various routes that can be followed when studying the rich variety of environmetrics techniques that are presented. Before the conclusions, a **flexible decision support system** is described in Section 1.7 for permitting a user to take full and immediate advantage of the environmetrics technologies given in the book.

1.2 THE SCIENTIFIC METHOD

1.2.1 Spaceship Earth

During the first weekend of December, 1989, the Cold War between the two superpowers came to an official end. On December 2 and 3, Soviet President Mikhail Gorbachev and American President George Bush held friendly bilateral meetings on ships anchored off the island of Malta in the Mediterranean Sea. Besides establishing a good working relationship between the two leaders, the summit's main achievement was the prospect of achieving an early agreement on decreasing by fifty per cent the superpowers' long-range nuclear arsenals. Massive nuclear and conventional weapon systems had been developed by both the Americans and Soviets during the Cold War period which lasted from the end of World War II right up until the end of the 1980's. At last, the threat of the destruction of the entire human race by a **global thermonuclear war** between the two superpowers seemed to be waning. Henceforth, the total number of nuclear weapons would subside and, hopefully, this would take place as quickly as possible.

Although the threat of extinction by a nuclear war has lessened, the citizens of the world are now well aware of an even more ubiquitous and deadly menace to human survival. This is the continuing ruination by mankind of the natural environment which supports all life forms on the planet earth. This **environmental devastation of the air, land and water** is being brought about by human activities such as cutting down forests, releasing untreated industrial and human wastes into the environment, widespread spraying of improperly tested insecticides on crops, draining too many wetlands, driving too many cars, excessive energy consumption, and, of course, overpopulation. Although proper management of the environment was not on the agenda at the Malta Summit, it seemed that nature gave Gorbachev and Bush a timely omen by flexing its muscles. On December 2, 1989, a violent storm with gusting winds of up to 100 km per hour created waves up to 4 metres high which played havoc with arrangements for the first day of the meeting. Because the Soviet missile cruiser Slava was bobbing wildly up and down like a cork in the storm, it was impossible to hold the first meeting aboard the Soviet cruiser as planned. Rather, the two Presidents were forced to meet for their first round of talks on the larger Soviet ship Maxim Gorky, which was moored in much calmer waters alongside a dock in the Maltese port of Marsaxlokk. For those who were thinking about the urgent environmental issues that these two world leaders as well as others must address, the message was clear - **nature is the key player here on earth and it should be treated with respect.**

For a dramatic example of an adverse environmental change caused by man, consider the so-called **greenhouse effect**. As reported by Levine (1990), during the past ten years, the trace gas composition of the atmosphere has been changing significantly over time. The buildup of atmospheric greenhouse gases including carbon dioxide, methane, nitrous oxide, chlorofluorocarbons and tropospheric ozone, could lead to global warming. This in turn could cause many undesirable aftermaths such as the melting of the polar icecaps and the ensuing flooding of coastal regions. Additionally, related climatic changes could turn fertile regions into deserts and thereby trigger huge migrations of populations to more hospitable regions. One major factor for the increase in CO_2 levels in the atmosphere is the conversion of forests to agricultural land through burning. Because the carbon incorporated in the trees is not balanced by carbon accumulated in crops or grasses, the burning constitutes a net release of carbon to the atmosphere. Tropical deforestation through burning is especially serious in the Amazon rainforests of Brazil. For example, during an American space shuttle flight in 1988, the astronauts photographed a biomass burn smoke cloud over the Amazon region which covered 3,000,000 km^2. The size of this cloud was second in size to the largest smoke cloud of 3,500,000 km^2 which was photographed by astronauts in 1985. Prior to 1985, the larger biomass clouds covered areas of only 300,000 km^2 (Levine, 1990). Other environmental problems caused by man-made changes to the atmosphere include depletion of stratospheric ozone which absorbs biologically lethal solar ultraviolet radiation, and acid rain which is insidiously decimating forests in the Northern Hemisphere.

The earth, in fact, has often been compared to a spaceship containing a valuable and fragile environment. It is the only known spaceship in the universe within which humans and other life-forms can live. Therefore, its natural resources should not be squandered, ruined or destroyed. As an illustration of what could happen to the entire planet, consider what may have taken place on **Easter Island** many centuries ago. Easter Island, located in the South Pacific ocean 2,700 km west of Chile, has an area of 163 km^2 and holds the distinction of being the most isolated piece of inhabited land in the world. On Easter Sunday in 1722, Jacob Roggereen, a Dutch explorer, discovered this remote island which he aptly named Easter Island. Today, the island is controlled by Chile. Easter Island is best known for its large stone statues called moai. More than 600 of these statues of grotesquely-shaped humans are scattered on the island and some of them are as high as 12 meters and weigh as much as 82 metric tons. However, most of the statues range from 3.4 to 6 meters in height. It is believed that most of these giant statues were sculpted in the period from about 1400 to 1680 A.D. What is not known is why the great moai culture that designed, built and erected these monoliths suddenly collapsed around 1680.

One hypothesis of what took place on Easter Island is provided by Flenley and King (1984) and Dransfield et al. (1984) in articles published in the journal Nature. More specifically, pollen records (Flenley and King, 1984) and shells from palm fruits coming from an extinct type of Chilean wine palm (Dransfield et al., 1984) suggest the existence of forests on the island and their decline during the last millennium. The authors of the two papers conjecture that the deforestation of Easter Island by the moai people caused their own cultural disintegration. In other words, the self-imposed environmental destruction of Easter Island led to the extinction of a great culture. Imagine what went through the minds of the remaining moai inhabitants as they cut down the last of the date palms and thereby severed their umbilical cord with nature. When one contemplates the analogy of Easter Island to the current treatment of the environment on spaceship earth by the world civilizations, the ultimate result is frightening.

The famous statistician George Box believes that the root cause of the present sorry state of the world is the **scientific method**. As explained by Box (1974), the scientific method provides the secret of learning fast and allows the normally very slow process of learning by chance experiences to be greatly accelerated. The scientific method furnished the fuel for the **industrial revolution** which started in Great Britain in the early 1700's and spread quickly to most of continental Europe and America in the 1800's. Today, the industrial revolution is a world-wide phenomenon along with the expansion into the present information age or, as it is also called, the second industrial revolution. In all of the changes brought about by the scientific method there are both advantages and drawbacks. For example, scientific medicine is responsible for fewer deaths at birth and by disease along with longer life expectancies. The disadvantage is that populations can grow too large for the environment to support properly. Scientific agricultural methods result in higher crop yields but at the expense of massive deforestation, the addition of chemical fertilizers and poisonous insecticides to the natural environment, as well as overpopulation. The scientific method furnishes the key for massive industrial expansion in order to produce great numbers of motor vehicles, lawn mowers, televisions, packaged foods and many other products that are in high demand. Unfortunately, the by-products that are endlessly dumped into the environment during the manufacture, utilization and ultimate disposal of these products are seriously polluting the air, earth and water. In short, the life support system for humanity is seriously ill because of the scientific method and it may never recover if drastic action is not taken now.

What can help to save humanity from its present dilemma? Well, mankind used the scientific method to create the current predicament and mankind can **utilize the scientific method to assist in restoring and properly managing the environment**. However, extremely quick and decisive action is required before it is too late. Box (1974) believes that **statistical methods** can act as a catalyst to further accelerate the scientific method for solving pressing environmental problems. In the next section, the scientific method is defined and the manner in which statistics can improve this powerful philosophy on the learning process is explained. Moreover, the potential influence of scientific studies of environmental problems upon the overall decision making process is described in Section 1.5. With knowledgeable and committed people at the helms of government and industry, as well as widespread public awareness, hopefully the current environmental mess can be rectified.

1.2.2 Description of the Scientific Method

Professor John Polanyi, winner of the 1986 Nobel Prize for Chemistry, presented a seminar on science, technology and society at Wilfrid Laurier University located in Waterloo, Ontario, Canada, on September 26, 1990. In his speech, he pointed out that the overall purpose of science is to **search for truth.** The general methodology which is employed to try to discover truths about nature is called the **scientific method.** Box (1974, 1976) considers the scientific method to be a process of **controlled learning.** By employing appropriate statistical methods in conjunction with the scientific method, the learning process can be made as efficient as possible. Furthermore, because this formal learning process can result in the discovery of thought-provoking and unforeseen truths, Groen et al. (1990) call science **the discipline of curiosity.**

A mandate of environmental agencies is to monitor the natural environment by taking measurements of various natural phenomena. For example, Environment Canada possesses massive statistical records on variables such as riverflows, temperature, precipitation, barometric pressures, and a wide range of water quality variables, from the east to west coasts of Canada. These huge accumulations of data on their own do not allow scientists and engineers to reach a better understanding of how various components of the environment function. On the other hand, speculative theoretical models or hypotheses about how the environment works, will not in the absence of data verification shed insight into what is taking place either. To reach a better understanding of nature through science, one must **consider both the available data and proposed theories** in order to be able to explain the behaviour of the phenomenon being studied.

Following the research of Box (1974, 1976), Figure 1.2.1 displays graphically how iterative learning between theory and data is carried out in science. In this figure, the theoretical realm of models and ideas is called **hypotheses** while the real world of facts and observations is referred to as **data.** Starting at the top left part of Figure 1.2.1, an initial hypothesis, H_1, about how nature behaves leads by a process of **deduction** to direct consequences of the theory which can be compared to the measured data. If these consequences fail to agree with the data, one can exploit the differences or errors in order to revise the hypothesis or theory by a process called **induction.** Notice in Figure 1.2.1 that induction goes from the data to the theory or, in other words, from specific facts to the general hypothesis, which is appropriately modified based upon the above stated discrepancies. Using the revised hypothesis, H_2, the learning cycle is repeated by employing deduction to go from the general to the specific. If the consequences of the new hypothesis are not in accordance with the data, one can utilize induction again as guidance for modifying the theory. These **learning cycles** consisting of deduction and induction are repeated as often as necessary until an acceptable hypothesis or theory is found. Eventually, a theory may be discovered which cannot be refuted by the available data. Further, this entire process of iterative learning leads to a much deeper understanding of what is occurring in the real world.

Figure 1.2.1 depicts the scientific learning process as an **iterative procedure.** One can also envisage the scientific method as a **feedback system.** As shown in Figure 1.2.2, the initial idea for a scientific study is stated as a hypothesis, H_1, which is then subjected to the ultimate test as to whether or not it describes what is happening in nature. More specifically, the discrepancies or errors between the data and the consequences of hypothesis H_1 lead to a modified hypothesis H_2. This feedback loop can continue so that H_2 leads to H_3 and, in general, H_i becomes H_{i+1}, until the data no longer refutes the hypothesis.

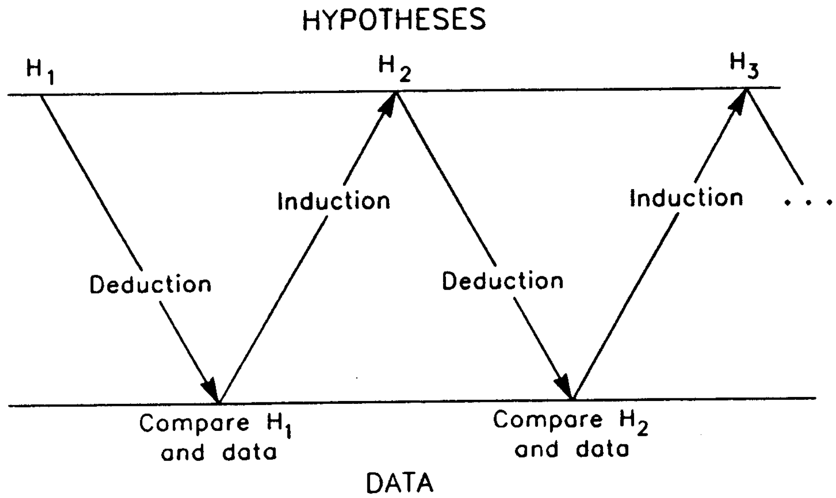

Figure 1.2.1. The iterative learning process used in science (Box, 1974, 1976).

At the bottom of Figure 1.2.2 is the **data input to the scientific method.** The data are placed here because they constitute the foundations of this whole learning procedure. As noted earlier, any hypothesis or mathematical model must reflect what is happening in the real world as represented by the observations. In the types of environmental problems used in applications in this book, the data were usually collected over a relatively long time period. For instance, average monthly riverflows are often available over a time span of 50 to 100 years. Weekly water quality data may be measured over a time period of 5 to 10 years. Whatever the case, when one is carrying out many environmental scientific studies, one can only use the available data, even though there may be many problems with these observations. There may simply not be enough time and money to obtain more data before the completion of the study. As a matter of fact, in many environmental agencies throughout the world, the scientists analyzing the natural data sets did not take part in designing the data collection procedure in the first place. Nonetheless, whenever possible, scientists are advised to assist actively in the design of the scheme for collecting the data which they will eventually analyze within the framework of the scientific method.

The methodology for efficiently collecting data for use in a scientific study is called **experimental design.** To obtain observations from the real world one must carry out experiments that are well designed. As just noted, for the case of environmental sciences such as hydrology and environmental engineering, one may have to collect data in the field over a fairly long time period. This is also the situation for areas like economics and history. However, in traditional and more basic sciences such as chemistry and physics, one can often obtain appropriate data within a fairly brief interval of time using experiments set up in the laboratory.

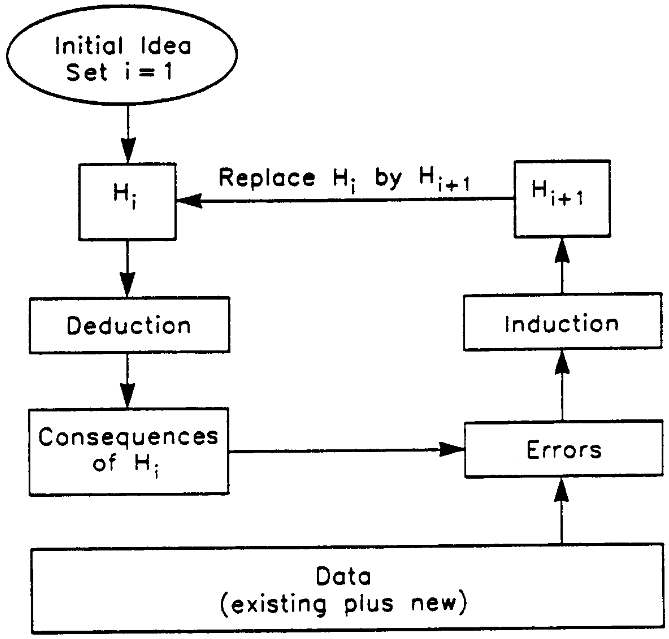

Figure 1.2.2. Feedback loop in the scientific method (Box, 1974, 1976).

To explain more specifically how experimental design plays a key role in the scientific method, refer to the expanded version of Figure 1.2.2 shown in Figure 1.2.3. At the base of the entire approach is nature depicted as a tree. An **experimental design forms a filter or window on nature** for efficiently obtaining the most appropriate observations for testing hypotheses. The available data may have been obtained from a formally designed experiment or data that were collected in an empirical fashion over the years. At each iteration in the scientific method, a current hypothesis H_i about the state of nature leads to specific consequences that are compared with facts obtained from the analysis of the available data. Differences between the consequences and facts can suggest how H_i can be modified to produce H_{i+1}. However, when it is not obvious as to what changes should be made to an unsatisfactory H_i or when further data may be required to confirm with more confidence a good hypothesis, further data should be obtained. One can see in the bottom right portion of Figure 1.2.3 that experimental design can be employed to obtain new observations. Notice that all data contain noise that must be taken into account in any data analysis. When experimental design is used, it is often possible to keep the noise to a minimum level, compared to when good data collection procedures are not utilized. Because the data represent the true state of nature, the scientific method leads to a convergence on the truth. If the noise level or experimental error is kept smaller, it will certainly be quicker and easier to discover the hypothesis that represents the true state of the component of nature being studied. Furthermore, even if two scientists who are separately studying the same problem start with different hypotheses and follow different routes, they will ultimately converge to the same destination when using the scientific method. For a good description of experimental design, readers may wish to refer to the textbook of Box et al. (1978) as well as references

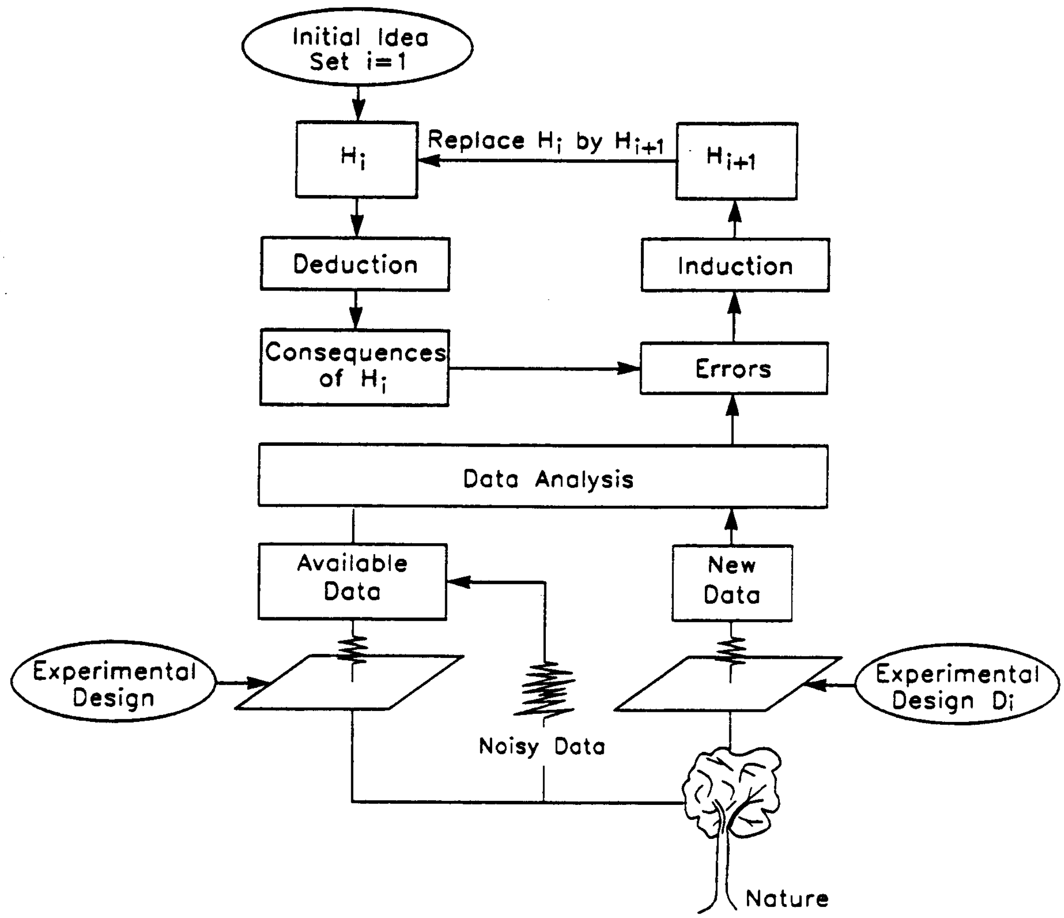

Figure 1.2.3. Data collection and analysis in the scientific method (Box, 1974, 1976; Box et al., 1978).

contained therein.

In the scientific method, one must test a hypothesis using real data. An explanation of **hypothesis testing** is given in Section 23.2 of the book just before the introduction of nonparametric statistical trend tests. The mathematical model underlying a given hypothesis is often expressed using some type of probabilistic **model**. For instance, a stochastic differential equation or a time series model may be employed to describe a given hydrological system where one might be testing a hypothesis about the output of the system given certain inputs. By definition, however, a mathematical model can never be the phenomenon it is describing, but only an approximation thereof. Nonetheless, if the mathematical model reflects well the key characteristics of the system, it may form a good basis for formally structuring the hypothesis and carrying out related data analyses. In Section 1.4.3, probabilistic models are classified according to informative criteria.

Another point to keep in mind is that the scientific method is purposefully designed to find out where one is wrong. In this way, one can **learn from experience** and come up with even a better hypothesis and underlying model (McPherson, 1990). As noted by Box (1974), there is no place in science for the man who wants to demonstrate that he has always been right.

1.2.3 Statistics in a Scientific Investigation

To carry out a **systematic scientific investigation** for discovering truths about nature, scientists and engineers employ the scientific method discussed in the previous subsection. From the outline of the main components of the scientific method, one can see that **statistics plays a key role in the scientific method** displayed in Figures 1.2.1 to 1.2.3. In fact, one can argue that by definition the scientific method must always involve statistics since one must use real data in order to refute or verify the current hypothesis. As noted by Box (1974), two main tasks in a given scientific investigation are:

1. the **design problem** where one must decide upon the appropriate data to obtain at each stage of an investigation.
2. the **analysis problem** where models are employed for determining what the data entitles the investigator to believe at each stage of the investigation.

In order to execute comprehensive analyses of the data, it is absolutely essential to determine properly the relevant data to obtain at the design phase by using appropriate techniques from **experimental design**. No amount of skill and experience in data analysis can extract information which is not contained in the data to begin with. Accordingly, **suitable data collection schemes are needed** for carrying out a time series analysis investigation. Within the statistical and engineering literature, extensive research has been published about designing optimal data collection schemes across a network of stations. For example, Moss (1979) wrote an introductory paper for a sequence of twenty-four papers published in Volume 15, Number 6, 1979, of Water Resources Research. For use in environmental pollution monitoring, Gilbert (1987) presents statistical methods from experimental design. Researchers at an international symposium that took place in Budapest, Hungary, delivered papers on how to design monitoring systems in order to detect changes in water quality variables (Lerner, 1986). At an international symposium held in Fort Collins, Colorado, on the design of water quality information systems (Ward et al., 1989), authors presented papers on topics ranging from data collection and network design to the roles of an information system within an overall water quality management system. Harmancioglu and Alpaslan (1992) describe water quality monitoring network design within a multiple objective framework. Other research regarding the proper design of water quality collection schemes for meeting a range of goals includes contributions by Ward et al. (1986), Ward and Loftis (1986), Whitfield (1988), and Loftis et al. (1991). Lettenmaier et al. (1978) suggest data collection schemes to use when one intends to employ the intervention model (see Section 19.7) to ascertain the effects of an intervention upon the mean level of a time series. Because most time series models must be used with a sufficient number of observations separated by equal time intervals, proper sampling is of utmost importance in time series analysis. If available measurements are not evenly spaced, appropriate **data filling** techniques can be utilized to estimate a series of evenly spaced data from the given information. As explained in Section 19.3.2, the particular technique to employ for data filling depends upon the type and amount of missing data. An advantage of employing nonparametric tests for detecting trends in time series is that they can usually be used with unequally spaced observations (see Chapter 23).

As already mentioned in Section 1.2.2, when dealing with time series studies in water resources and other environmental sciences, often the data were collected over a long period of time and the people analyzing the collected data did not take part in designing the data collection procedure in the first place. Of course, wherever possible, practitioners are advised to actively take part in the design of the scheme for collecting the data which they will analyze. Nevertheless, because environmental scientists are often confronted with analyzing data that were already collected, this book concentrates on data analysis while keeping in mind the great import of efficient data collection.

1.2.4 Data Analysis

When analyzing a given set of data within an overall scientific investigation, Tukey (1977) recommends adhering to the following two steps:

1. **exploratory data analysis,**
2. **confirmatory data analysis.**

The main purpose of the exploratory data analysis phase of data analysis is to discover important statistical properties in the given observations by carrying out simple graphical and numerical studies. The major objective of the confirmatory data analysis stage is to confirm statistically in a rigorous fashion the absence or presence of certain properties in the data.

In Part X of the book, it is explained how useful exploratory and confirmatory data analysis tools can be effectively employed for studying environmental data. Hydrological time series, such as seasonal riverflows, temperature and precipitation, are usually quite suitable for analysis purposes since, for example, they possess few missing values and outliers. However, other types of environmental series like water quality time series are often quite messy due to many factors. For instance, water quality time series may possess many missing observations among which there are long periods of time for which there are no measurements. Moreover, the data may have many extreme values and be affected by external interventions such as industrial development and other land use changes in a river basin. Fortunately, the **data analysis procedures of Part X are designed to handle both well behaved and messy time series.**

Many of the exploratory data analysis tools are presented in Part X, although **graphical procedures** are used throughout the book for explanation purposes. A wealth of **time series models** that can be used in data analysis studies are presented in Parts II to IX in the book. Additionally, **nonparametric trend tests** and **regression analysis methods** are discussed in Chapters 23 and 24, respectively. These latter two types of confirmatory data analysis techniques are especially well designed for use with messy environmental data.

The graph in Figure 1.1.1 of the average monthly phosphorous data demonstrates how a useful exploratory data analysis tool can usually convey a wealth of information. For instance, as already pointed out in Section 1.1, one can clearly see the drop in the mean level of the series due to the introduction of phosphorous treatment. Moreover, the location of the missing values marked as filled-in circles can be clearly seen. Within a scientific study, one may wish to test the hypothesis that there is a significant step drop in the mean level of the phosphorous and also to estimate its magnitude. The **intervention model** of Part VIII in the book can be employed for these purposes. More specifically, the details of this trend analysis assessment for the phosphorous data are presented in Section 19.4.5. From a qualitative viewpoint, the intervention model for the phosphorous series possesses the components shown below:

Phosphorous Series	=	Intervention Component	+	Missing Value Terms	+	Correlated Noise

Notice that this flexible model can simultaneously model the effect of the intervention, estimate the four missing values, and handle a correlated noise term. The estimate of the parameter in the intervention component, along with its standard error of estimation, allows one to carry out a hypothesis or significance test to see if there is a drop in the mean level and to quantify the magnitude of the drop. Indeed, as would be expected from the results of the exploratory graph in Figure 1.1.1, there is a significant step trend in the mean level. The best estimate for this drop is a 76% decrease compared to the previous mean, where the 95% confidence interval spans from 68% to 84%.

The exploratory and confirmatory data analysis study for the phosphorous time series points out a major contribution of this book - **the use of statistical methods in environmental impact assessment.** In a statistical environmental impact assessment, one is often required to detect and model trends in time series. In Section 22.3, a variety of **useful graphical techniques** are presented that can be employed as exploratory data analysis tools for confirming the presence of suspected trends as well as discovering unknown trends. Additionally, in Section 24.2.2, a regression analysis technique is described for tracing trends on the graph of a data set. The confirmatory data analysis tools that can be employed in trend assessment are:

1. **intervention analysis** (Chapter 19 and Section 22.4),
2. **nonparametric tests** (Chapter 23),
3. **regression analysis** (Chapter 24).

Additionally, **overall approaches for analyzing messy environmental data** that fall under the paradigm of exploratory and confirmatory data analysis are presented in Chapter 22, Section 23.5 and Section 24.3. In all three cases, detailed environmental applications are given to explain clearly how the methodologies work.

Each of the case studies in statistical environmental impact assessment presented in the book deal with the modelling and analysis of trends caused by one or more external interventions that have already taken place. For example, the step decrease in the monthly phosphorous level in the time series shown in Figure 1.1.1 from January, 1972 to December, 1977, was created by the tertiary phosphorous treatment which started in February, 1974. Consequently, the formal trend analysis of the phosphorous observations constitutes a **posterior environmental impact assessment.** In some situations, it may be required to find out the potential effects upon the environment of a planned project, before permission is granted for commencing construction. For instance, when designing a series of reservoirs for the production of hydroelectric power and other benefits, decision makers may wish to know the potential impacts of the scheme upon a range of hydrological and other environmental variables. Hence, scientists would perform an a **priori environmental impact assessment** using appropriate scientific tools. Although this book does not consider a priori environmental impact assessments, some of the statistical tools used in the many posterior environmental impact studies that are presented could provide some guidance in a priori studies. For example, knowledge gained from intervention investigations of reservoirs that have already been built could be employed for simulating possible environmental scenarios for planned reservoir systems.

1.3 PHILOSOPHY OF MODEL BUILDING

1.3.1 Occam's Razor

To better understand and control his environment, **mankind uses models.** In order to be sufficiently accurate and realistic, a model must be able to capture mathematically the key characteristics of a system being studied. At the same time, a model must be designed in a simple straightforward manner so that it can be easily understood, manipulated and interpreted. Because of the great complexity of water resources and other natural systems, models are extensively developed and applied in water resources and environmental engineering.

Time series models constitute an important class of models which can be used for addressing a wide range of challenging problems in the environmental sciences. In fact, **time series models** are a type of stochastic model designed for fitting to observations available at discrete points in time. An example of a time series is the graph of the average monthly phosphorous measurements for the Speed River shown in Figure 1.1.1. The specific type of time series model that is fitted to this series in Section 19.4.5 is the intervention model, briefly referred to in Section 1.2.4.

As explained in the previous subsection, a comprehensive **data analysis** study can be carried out by following the two main steps consisting of exploratory and confirmatory data analysis. Subsequent to employing informative graphical methods at the exploratory data analysis stage to appreciate the main statistical properties of the observations being studied, formal modelling can be done at the confirmatory data analysis stage to ascertain more precisely how the data are behaving. For instance, one may wish to use an intervention model to test the hypothesis that there is a step drop in the mean level of the series in Figure 1.1.1 and to estimate the magnitude of this decrease (see Section 19.4.5).

In the next subsection, a systematic procedure is outlined for determining the most appropriate time series model to fit to a given data set. This general model building approach is, in fact, adhered to for applying all of the time series models presented in this book to measurements taken over time. The basic idea underlying the **model construction** procedure of Section 1.3.2 is to identify a simple model which has as few model parameters as possible in order to provide a good statistical fit to the data.

The **principle of model parsimony** has historical roots that go back far into the past. Aristotle, for example, postulated that nature operates in the shortest possible way. A 14*th* century English Franciscan monk by the name of William of Occam (1280-1349) developed a principle now known as **Occam's razor**. One version of his principle states that when faced with competing explanations choose the most simple one. This is analogous to using a sharp thin razor to make a clean cut through some material. Another equivalent statement for Occam's razor is entities are not to be multiplied without necessity. Bertrand Russel (1946), the famous 20*th* century British mathematician, found Occam's razor to be very informative in logical analysis. This is because there is only one explanation or description of something which is minimum, whereas there can be an infinity of explanations which bring in other entities. Russell went on to claim that adherence to the minimum necessary explanation or description ensures that the examination of hypotheses and evidence for and against them will remain coherent. Checkland (1981) provides a good explanation of Occam's principle in his book on the theory and practice of systems engineering.

In fact, model parsimony is a key assumption embedded in the scientific method of Section 1.2. One should always strive to model nature and postulate hypotheses thereof in the most straightforward and simple manner, while still maintaining an accurate description of the phenomenon being modelled. Further discussions on modelling philosophies in time series analysis are presented in Section 5.2 as well as Section 6.3.

1.3.2 Model Construction

In time series modelling and analysis, one wishes to determine the most appropriate stochastic or time series model to fit to a given data set at the confirmatory data analysis stage. No matter what type of stochastic model is to be fitted to a given data set, it is recommended to follow the identification, estimation, and diagnostic check stages of model construction (Box and Jenkins, 1976). At the **identification stage**, the more appropriate models to fit to the data can be tentatively selected by examining various types of graphs. Some of the identification information may already be available from studies completed at the exploratory data analysis step discussed in Section 1.2.4. Because there may be a range of different families of stochastic models which can be fitted to the time series under consideration, one must choose the one or more families of models which are the most suitable to consider. The family selections can be based upon a sound physical understanding of the problem, output from the identification stage, and exploratory data analyses. Although sometimes it is possible to choose the best model from one or more families based solely upon identification results, in practice it is often not obvious which model is most appropriate and hence two or three models must be tentatively entertained. At the **estimation stage**, maximum likelihood estimates can be obtained for the model parameters and subsequently the fitted model can be subjected to **diagnostic checks** to ensure that the key modelling assumptions are satisfied. When considering linear stochastic models such as the ARMA (autoregressive-moving average) models of Chapter 3, one should check that the model residuals are not correlated, possess constant variance (i.e., homoscedasticity) and are approximately normally distributed. If the residuals are not white noise, the model should be redesigned by repeating the three phases of model construction. In practice, it has been found that a suitable Box-Cox power transformation (Box and Cox, 1964) (see Section 3.4.5) can rectify anomalies such as heteroscedasticity and non-normality. The specific tools utilized at the three stages of model construction are dependent upon the particular family of models being entertained. The logic underlying the traditional approach to model construction is displayed as a flowchart in Figure 1.3.1.

In Part III of this book, **a wide variety of model building tools** are presented for use with ARMA (defined in Chapter 3) and ARIMA (autoregressive integrated moving average) (Chapter 4) models that can be fitted to nonseasonal stationary and nonstationarity time series, respectively. Most of the identification, estimation and diagnostic check methods described in Chapters 5, 6, and 7, respectively, in Part III, can be expanded for employment with the many other kinds of time series models presented in this book.

1.3.3 Automatic Selection Criteria

As noted in Section 1.3.1, a basic tenet of model building is to keep the model as simple as possible but at the same time provide a good fit to the data being modelled. **Automatic selection criteria** (ASC) are now available for balancing the apparently contradictory goals of good statistical fit and model simplicity. In Section 6.3, a number of ASC are defined and it is

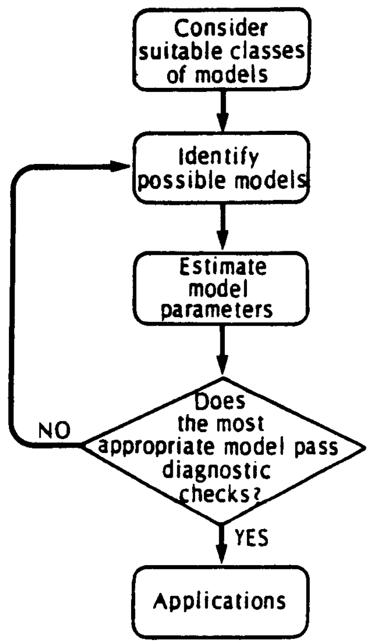

Figure 1.3.1. Model construction.

explained how they can enhance the three stages of model construction portrayed in Figure 1.3.1. One example of an ASC is the **Akaike information criterion** (AIC) of Akaike (1974) which is used throughout this book.

In general, an ASC is defined as follows:

ASC = Good Statistical Fit + Complexity

The first term on the right hand side is written as some function of the value of the maximized likelihood function for the model fitted to the data (see Section 6.2 for a discussion of maximum likelihood estimation). This term is defined in such a way that the smaller the value the better the statistical fit. One would expect that a more complex model would furnish a more accurate description of the data. The purpose of the second entry in the ASC formula is to guard against having a model which is too complex and to abide by the principle of Occam's razor of Section 1.3.1. The complexity component in the ASC is a function of the number of model parameters, where a smaller value means that there are fewer parameters in the model. Hence, overall one would like to select the model which has the lowest value for the ASC. In Chapter 6, Figure 6.3.1 depicts how an ASC can be used in model construction.

1.4 THE HYDROLOGICAL CYCLE

1.4.1 Environmental Systems

The basic structure of the scientific method described in Section 1.2 is depicted in Figure 1.2.3. Because the scientific method seeks to discover truths about nature, the natural world as represented by the tree in Figure 1.2.3 is shown as the foundation of the scientific method at the bottom of the figure. In order for the scientific method to work, one must **employ mathematical models that reflect the important physical characteristics of the system being studied.**

When carrying out a scientific study, one can better appreciate what is taking place if one envisions a conceptual framework of the physical world within which the mathematical modelling is being done. An ideal paradigm for accomplishing this is the systems design approach to modelling. In **systems modelling**, one thinks of an overall system composed of subsystems that interact with one another in some sort of hierarchical manner. Individual subsystems can be studied and analyzed in detail using powerful mathematical models. By properly connecting the subsystems together, one can synthesize the overall system behaviour, given the initial conditions and operating rules within the subsystems.

A question that naturally arises is how one should define the boundaries of the system and its subsystems given the type of problem being studied. As explained by authors such as White et al. (1984) and Bennett and Chorley (1978), there are many ways in which one can define **environmental systems**. At the solar system level, the system consists of the sun and each of the nine planets which rotate around the sun. The yearly rotation of the earth around the sun along with the tilting of the earth's axis is the main cause of the seasons on earth. The rotation of the moon about the earth creates tides in the oceans and seas. Besides gravitational forces, direct solar energy from the sun constitutes the major input for environmental systems contained on, in and around each planet. If one is studying the overall environmental system for the planet earth on its own, then one must subdivide this system into finer subsystems which may not be explicitly considered when looking at the earth and the other planets at the solar system level. More specifically, at this level one may wish to subdivide the overall global environmental system into a number of major subsystems which include the atmosphere, hydrosphere and lithosphere (earth's crust). The biosphere subsystem, in which plant and animal life exist, occurs at the transition zone between the lithosphere and atmosphere as well as within the hydrosphere. Different kinds of ecosystems are contained within the biosphere system.

Any of the aforementioned global subsystems can, of course, be further subdivided into finer subsystems. For instance, the atmosphere can be vertically categorized from the earth's surface outward, into subsystems consisting of the troposphere, stratosphere, mesosphere and thermosphere. Within the lower part of the atmosphere, one can examine the various circulation subsystems around the world.

In summary, one can **define environmental systems in a hierarchical fashion** from overall large systems at the planetary level to very detailed subsystems at much lower levels. The system definitions to be entertained depend upon the particular problem being studied. For example, if one is examining water pollution problems within the Great Lakes in North America, the largest environmental system to consider may be the drainage basin for the Great Lakes. Within this overall system, one could examine river subsystems through which pollutants carried by water can flow into the Great Lakes. The definitions of subsystems could be made as detailed as required for the problem at hand. Within and among the subsystems, appropriate mathematical models could be used to model precisely the physical, chemical and biological interactions.

One particular environmental system which may form a basis for many problems examined by water resources and environmental engineers is the well known **hydrological cycle**. This important environmental system is described in the next subsection. Subsequently, a range of mathematical models that can be used for modelling natural phenomena within the hydrological cycle, as well as other environmental systems, are classified according to informative criteria. Based on these discussions, one can appreciate the role of time series models for describing and analyzing important environmental phenomena.

1.4.2 Description of the Hydrological Cycle

Hydrology is the science of water. In particular, hydrology deals with the distribution and circulation of water on the surface of the land, underground and in the atmosphere. Additionally, hydrology is concerned with the physical and chemical properties of water and its relationships to living things. Similar definitions for hydrology can be found in hydrological books written by authors such as Eagleson (1970), Linsley et al. (1982) and McCuen (1989).

The environmental system which hydrologists employ to describe the components of their science is called the **hydrological cycle**. Figure 1.4.1 displays a schematic of the hydrological cycle which is based upon the figure provided by Eagleson (1970, p. 6). The **throughput** to the hydrological system in Figure 1.4.1 is water which can occur in a liquid, solid or vapour phase. Because the hydrological cycle does not allow water to escape, it forms a **closed system** with respect to water. The **main forces** which propel the water through the hydrological cycle are solar energy and gravity. As pointed out by Eagleson (1970, p. 5), the dynamic processes of vapour formation and transport are powered by solar energy while precipitation formation and the flow of liquid water are driven by gravity. The transformation of water from one phase to another as well as the transportation of water from one physical location to another are the main features of the hydrological cycle. To understand the possible routes and phases a water molecule passes through in the hydrological cycle in Figure 1.4.1, one can start at any point in the cycle. Note that all water is returned from the atmosphere to land or surface bodies of water through the process of **precipitation** by going from vapour to liquid form. Liquid water can **infiltrate** into the soil and flow overland via streams, rivers and lakes to the oceans. **Evaporation** is the dynamic process which returns water from its liquid phase in streams, rivers, lakes and oceans to its vapour phase in the atmosphere. Water molecules can also be released to the atmosphere from the surfaces of plants through a process called **transpiration. Evapotranspiration** consists of the total water transferred to the atmosphere by transpiration from plants plus evaporation from the soil on which the plants are growing. It is interesting to note that in land areas having a temperate climate, approximately 70% of the precipitation returns to the atmosphere via evapotranspiration while the remaining 30% mainly appears as riverflows (McCuen, 1989, p. 668). **Sublimation** takes place when water is transformed from its solid form directly to its gaseous state in the atmosphere. Notice also in Figure 1.4.1 the various ways in which water can enter and leave the two subsystems consisting of soil and underground aquifers.

Interest in ideas related to hydrology can be traced back to the ancient Egyptians, Greeks and Romans. For instance, as early as 3,000 B.C., the Egyptians had gauges called nilometers to measure the stages or depths of the Nile River. A **nilometer** consists of a stone pillar on which markings indicate the depth of the Nile, especially during flooding. However, it was not until

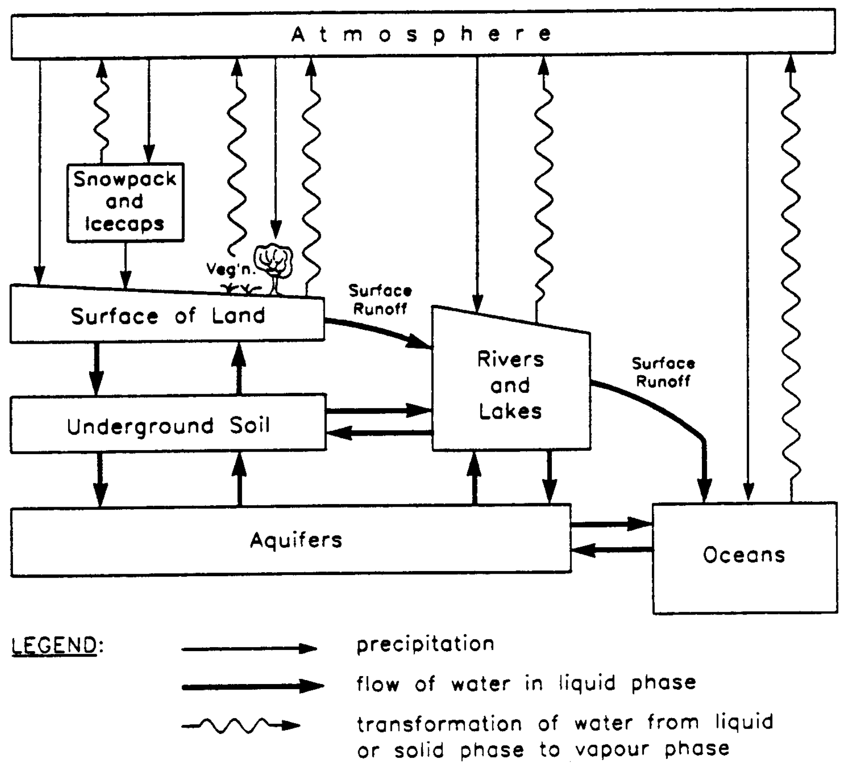

Figure 1.4.1. The hydrological cycle (Eagleson, 1970, p. 6).

the **Renaissance** that the first essentially correct picture of the hydrological cycle was produced by **Leonardo da Vinci**. Hence, this first accurate description of an important environmental system is now more than 400 years old. Besides providing a good conceptual description about how the transformation and transportation of quantities of water takes place on earth, the hydrologic cycle can be used for other purposes. Specifically, the hydrologic cycle furnishes a framework for understanding how man-made pollution can enter the hydrologic system at any point and pollute our entire environment by following the ancient pathways traced out by water in all of its forms. **Therefore, the science of hydrology provides solid foundations upon which many other environmental sciences can build and interact.**

During the past century, the development of hydrology has been mainly led by civil and agricultural engineers working on traditional engineering problems such as water supply and flood control. Consequently, the field of hydrology has been pragmatic in its outlook and narrowly focused upon only a few aspects of the overall hydrological cycle displayed in Figure 1.4.1. **However, in order to make wise and timely decisions regarding solutions to pressing environmental problems occurring throughout the hydrologic cycle and right up to the global level, a fundamental scientific understanding of hydrology is required.** Accordingly, in 1987 the National Research Council of the United States established a panel to conduct an assessment of hydrology that appeared as a report (National Research Council, 1991) which is summarized by the panel chairperson P.S. Eagleson (1991). The authors of this report, as well as Falkenmark (1990), describe many key research areas in which the science of hydrology should be expanded so that sound environmental policies can be properly devised and implemented. Indeed, they correctly point out that hydrology should be developed into a comprehensive and distinct geoscience. Moreover, they note that human activity and decision making can have dramatic influences upon the hydrological cycle and hence are an integral part of that cycle.

1.4.3 Classifying Mathematical Models

A general framework in which to envision how various components from nature interact with one another is to employ the environmental systems approach outlined in Section 1.4.1. A particularly informative environmental system for use as a sound scientific structure in water resources and environmental engineering is the hydrological cycle of Section 1.4.2 and Figure 1.4.1. Within a given environmental systems framework such as the hydrological cycle, one can use the scientific method of Section 1.2 to develop a range of specific models to describe natural phenomena occurring in the system. In order to abide by the principle of Occam's razor of Section 1.3.1, the simplest models to describe nature are almost always formulated in terms of appropriate mathematical equations. Hence, one can argue that **mathematics is the language of science**. Indeed, some mathematicians feel that mathematics should be considered as a separate scientific discipline. In reality, mathematics constitutes an interdisciplinary scientific field which supports all areas of science.

One should keep in mind that **real problems in the natural and social world inspired scientists to develop the most useful and dramatic contributions to scientific mathematics.** For example, to describe properly his revolutionary laws of nature, Sir Issac Newton developed the mathematics of calculus. In order to solve practical problems in agriculture, Sir Ronald A. Fisher invented experimental design and many other branches of statistics (Box, J., 1978). Worthwhile areas of game theory were formulated for modelling and analyzing actual social disputes (see discussion in Section 1.5.2).

Since the time of Newton, great progress has been accomplished in building a treasure house of many different kinds of mathematical models. To appreciate the general types of mathematical models that are available for application to scientific problems, it is informative to classify these models according to useful criteria. Two overall categories into which mathematical models can be placed are deterministic and stochastic models. When a mathematical model can be employed for determining exactly all the states of a system, the model is said to be **deterministic**. For instance, when plotting on a graph all values of an algebraic function representing the states of a system, the precise locations of all possible points on the curve are known because the equation is deterministic.

If a state of a system can only be described using probabilistic statements and hence its precise value is not known, the mathematical equations describing the system are said to be **stochastic or probabilistic**. For example, when forecasting tomorrow's weather conditions using an appropriate stochastic model along with the latest meteorological information, a meteorologist may state that there is an 80% chance of snowfall tomorrow. This means that there is still a 20% chance that no precipitation may occur. The meteorologist may go on to forecast that he or she is 95% confident that the amount of snow accumulation will be between 15 and 20 cm.

Most natural phenomena occurring in environmental systems appear to behave in random or probabilistic ways. In other words, it is almost impossible to say exactly how nature will behave in the future, although one can often make reasonable predictions about occurrences using probabilistic statements. Consequently, the mathematical tools presented in this book fall within the realm of stochastic models.

As explained in Section 2.3, **stochastic models** are mathematical models for describing systems which evolve over time according to probabilistic laws. When discussing the theoretical aspects of stochastic models, the term stochastic processes is often used in place of stochastic models. Following Cox and Miller (1965), Table 1.4.1 describes a method for categorizing stochastic models according to the two criteria of time and state space. Notice that **time** can be either discrete or continuous. The **state space** or values of the variables describing the system being modelled can also be subdivided according to discrete and continuous values. Examples of the four kind of models that can be categorized using the above criteria are given in Table 1.4.1. Markov chains, for instance, fall under the subdivision of stochastic models which incorporate discrete time and discrete values of the state space in their mathematical structure. Stochastic differential equations can handle continuous time and continuous values of the state space (Kloeden and Platen, 1992). Point processes, such as Poisson processes, model discrete values over continuous time.

Table 1.4.1. Classifications of stochastic models.

		STATE SPACE	
		Discrete	Continuous
TIME	Discrete	Markov Chains	Time Series Models
	Continuous	Point Processes	Stochastic Differential Equations

This book deals with stochastic models that model continuous observations measured at discrete points in time. Because these models formally describe measurements available over discrete time in the form of a time series, they are usually referred to as **time series models.** The application of time series models to actual data is popularly referred to as **time series analysis.**

Why are time series models of such great import in the environmental sciences? The answer is quite simple. **In order to understand how a natural system is behaving, scientists take measurements over time,** hopefully according to a proper experimental design (see Section 1.2.3). An example of a water quality time series is displayed graphically in Figure 1.1.1. Time series models are specifically designed for formally modelling this type of information which occurs frequently in practice. Furthermore, techniques for fitting time series models to data are now highly developed (see Sections 1.3.2 and 1.3.3) and, hence, these time series models can be immediately employed for modelling, analyzing and better understanding pressing scientific problems.

In 1970, Box and Jenkins dramatically launched time series modelling into the realm of real world applications with the publication of their seminal book entitled "Time Series Analysis: Forecasting and Control" (the second edition was published in 1976). Besides presenting a wide variety of useful time series models, they showed how these models can be applied to practical problems in a wide range of disciplines. One should keep in mind that both **Box and Jenkins considered themselves foremost to be scientists** and not mathematicians. In other words, these scientists developed and used mathematical tools for scientifically studying actual problems.

Some of the time series models presented in this book are also discussed by Box and Jenkins (1976). However, **in this book the latest developments in time series modelling are given,** including many new procedures for allowing models to be conveniently fitted to data within the framework of model construction outlined in Section 1.3.2. Moreover, this book presents other kinds of models, such as many of those given in Parts V to X, which are especially useful in water resources and environmental engineering. As mentioned in Section 1.1, this book is a document falling within the challenging field of environmetrics and contains a range of useful time series models which are currently completely operational.

When using a mathematical model to describe a natural system, one would like the model to have a **sound physical basis,** in addition to possessing attractive mathematical properties. As pointed out in various sections of this book, certain time series models are well designed for various kinds of applications in hydrology and water quality modelling. For example, as explained in Section 3.5, ARMA models possess good physical justifications for use in modelling annual streamflows (Salas and Smith, 1981).

1.5 DECISION MAKING

1.5.1 Engineering Decision Making

The **scientific method** of Section 1.2.2 provides a solid foundation upon which solutions to the physical aspects of environmental problems can be properly designed and tested. As pointed out in Section 1.2.3, **statistical and stochastic models** have a key role to play for enhancing scientific investigations in terms of accuracy, speed and better understanding. Furthermore, by keeping in mind the **hydrological cycle** of Figure 1.4.1, the overall physical relevance of environmental problems being studied can be kept in correct perspective.

The **physical characteristics of environmental investigations** can involve physical (ex. waterflow), chemical and biological factors. For instance, one may wish to examine how industrial chemical pollutants discharged into rivers affect certain populations of fish and suggest correct measures for overcoming any serious problems. Scientists may discover why and how fish populations are dwindling and be able to design pollution controls to rectify the situation. However, to implement corrective measures, finances are required and political decisions must be made. Therefore, in addition to generating physical solutions to a given environmental problem by use of the scientific method, scientists must also take into account the **socio-economic aspects of decision making.** In other words, **both the physical realm of nature as well as the social world created by mankind's ability to think, must be properly accounted for in real world decision making.**

As an example of a planned large-scale engineering project which could adversely affect the environment consider the case of the **Garrison Diversion Unit (GDU).** As explained by Hipel and Fraser (1980) and Fraser and Hipel (1984), the GDU is a partially constructed multipurpose water resources project in the United States which involves the transfer of water from the Missouri River basin to areas in central and eastern North Dakota that are mainly located within the Hudson Bay drainage basin. Figure 1.5.1, which is taken from Hipel and Fraser (1980) and also Fraser and Hipel (1984, p. 26), shows the major regions affected by this large project located in the geographical centre of the North American continent. After the system becomes operational, water will be pumped from Lake Sakakawea on the Missouri River via the McClusky canal to the Lonetree Reservoir located in the Hudson Bay drainage basin. From the Lonetree Reservoir water will flow along the Velva and New Rockford canals to major irrigation areas. Additionally, water from the Lonetree Reservoir will augment flow in the James River for downstream irrigation. The resulting runoff from the irrigated fields would flow via the Red and Souris Rivers into the Canadian province of Manitoba. Adverse environmental effects from the GDU include high pollution levels of the irrigation waters, increased chances of flooding in the Souris River, and the possibility of catastrophic environmental damage caused by foreign biota from the Missouri River basin destroying indigenous biota, such as certain fish species, in the Hudson Bay drainage basin.

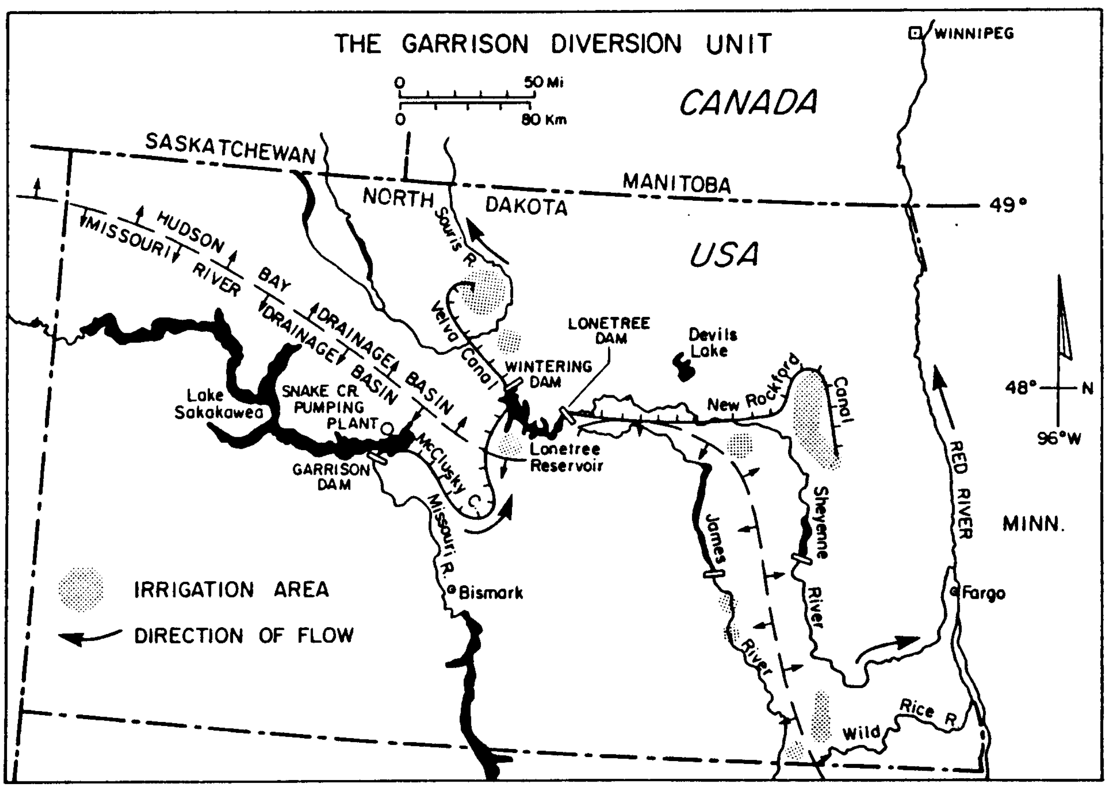

Figure 1.5.1. Map of the Garrison Diversion Unit (Hipel and Fraser, 1980; Fraser and Hipel, 1984, p. 26).

Because the GDU involved a **variety of interest groups** which interpreted the problem from different perspectives, the GDU project escalated into a **serious international environmental conflict.** The American proponents of the GDU wanted the full project, as approved by the U.S. Congress in 1965, to be built. However, Canada was afraid of potentially disastrous environmental consequences within its own borders while American environmentalists did not like some adverse environmental effects which could take place in North Dakota. The GDU controversy also involved the International Joint Commission (IJC), an impartial body initiated by the Boundary Waters Treaty of 1909 between the U.S. and Canada for investigating conflicts arising over water quantity and quality.

In order for a large-scale project like the GDU to be eventually implemented and brought into operation, the following factors must be adequately satisfied:

1. **Proper Physical Design** - For instance, physical structures such as dams, pumping stations and irrigation channels for the GDU must be correctly designed so that natural physical laws are not violated and the project is safe.
2. **Environmentally Sound Project** - If, for example, the GDU project were to be built and come into operation, adverse environmental consequences must be less than agreed upon levels. The ability to meet environmental standards must be incorporated into the basic design of the project.
3. **Economical and Financial Viability** - For the case of the GDU, the project must be economically feasible and sufficient financial resources must be available to pay for the project. It is interesting to note that some benefit-cost ratios for the GDU produced ratios much less than one.
4. **Socially and Politically Feasibility** - A politically feasible solution to the GDU project must be found before it can come into operation.

Unfortunately, for the case of the GDU only the first factor of the four listed above was ever properly satisfied. The Garrison dam on the Missouri River (see Figure 1.5.1) was completed by the Bureau of Reclamation of the U.S. Department of the Interior in 1955. Other physical facilities, such as the canals shown in Figure 1.5.1, were designed but never completely constructed. Environmental effects of the project were almost entirely ignored in the initial design of the project. **The social repercussions caused by the ensuing political controversy over environmental problems as well as suspect economic studies eventually prevented the completion of the project,** even though hundreds of millions of dollars had already been spent.

The main lesson garnered from this GDU fiasco is that all of the factors given above must be properly taken into account by scientists and other decision makers in any engineering project. Otherwise, the project may never be completed as first envisioned or it may be cancelled altogether. The scientific method and related mathematical modelling are especially important for ensuring **physical and environmental soundness.** Indeed, one should also follow a scientific approach in socio-economic modelling and analyses. Because the last two factors given above involve activities that relate directly to mankind as distinct from his natural environment, the models for studying these activities are sometimes referred to as **decision making tools.** As explained in Section 1.5.2, many of these decision making methods were developed within a field called **operational research.**

Figure 1.5.2 summarizes a **systems design approach to decision making** in engineering, which must properly consider the important factors mentioned before. To keep in mind that the entire activity takes place within the environment or natural world, the flowchart is enclosed by a wavy line. Notice that the physical, environmental, economical and financial considerations provide background information that can affect the preferences and actions of the decision makers who are included in the social and political modelling and analyses. If, for example, a first class environmental impact assessment is carried out beforehand for the project, there is a higher probability that decision makers will approve a design which abides by the suggested environmental standards. Additionally, if this project is economically and financially viable, the project has an even higher chance of being accepted by the decision makers. The political game that could take place among the decision makers who can influence the final decision can be modelled using techniques from **conflict analysis** (Fang et al., 1993; Hipel, 1990; Fraser and Hipel, 1984). The results of all of these formal studies provide background information upon which the actual decision makers can base their decisions. As shown by the feedback loops in Figure 1.5.2, additional information can be obtained as required and appropriate changes can be made. Moreover, some decision makers may obtain some of their information directly from their own observations of real world events. Hence, in Figure 1.5.2 there is an arrow going from the real world to the box labelled information for decision makers to indicate that people do not have to rely entirely upon results generated from formal studies. Finally, the design problem referred to in Figure 1.5.2 is not restricted to the construction of a new project such as the GDU. It can also represent situations such as a change in operating policy of a system of reservoir and the design of pollution control devices for installation in existing industrial facilities.

The remainder of Section 1.5, deals mainly with mathematical models that can be employed for modelling and analyzing decision making. In the next subsection, decision making models from the field of operational research are classified according to useful criteria. Subsequently, the use of conflict analysis for modelling and analyzing the GDU dispute is described and the importance of **sound scientific modelling within the overall decision making process** is once again emphasized.

1.5.2 Decision Making Techniques in Operational Research

The field of **operational research** consists of some general methodologies and many specific techniques for studying decision making problems. The British initiated operational research just prior to World War II when they performed **research** studies into the **operational** aspects of radar systems for detecting incoming enemy aircraft to the United Kingdom. Throughout the war, the British employed OR in all of their military services for successfully solving large scale military problems involving the movement of great numbers of military personnel and huge quantities of war materials (Blackett, 1962; Waddington, 1973). The American military also used this systems science approach to problem solving during the second world war but called it **operations research**. Many practitioners now simply refer to operational research or operations research as **OR**. Since the war, OR has been extensively expanded and utilized for looking at operational problems in many different fields outside of the military such as management sciences, transportation engineering, water resources and industrial engineering. Operational research societies have sprung up in most industrialized countries along with the publication of many OR journals.

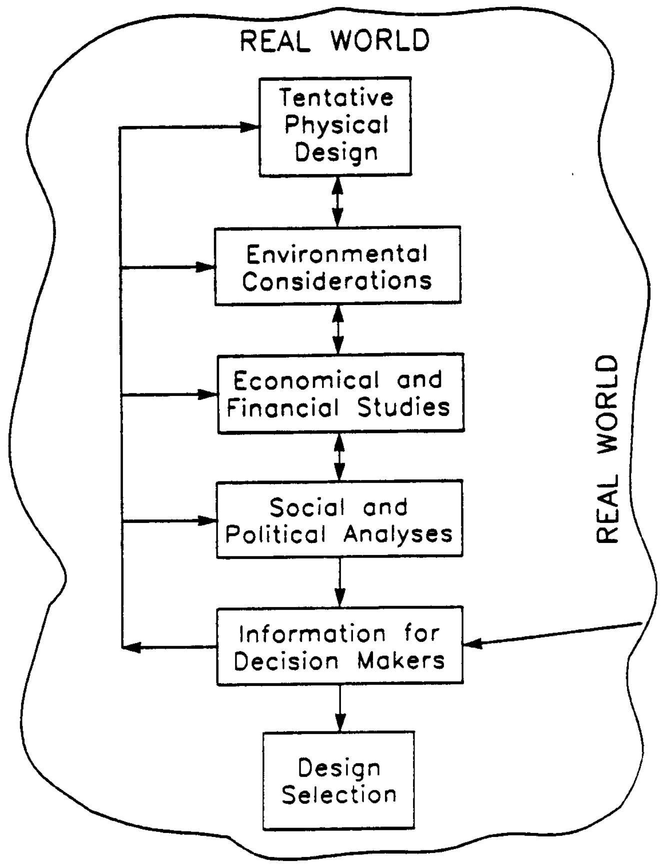

Figure 1.5.2. Engineering decision making.

The discipline of OR is both an art and a craft. The art encompasses general approaches for solving complex operational problems while the craft consists of a great variety of mathematical techniques which are meant to furnish reasonable results when properly applied to specific problems. Following Hipel (1990), Table 1.5.1 shows how **OR methods can be categorized according to the criteria of number of decision makers and number of objectives.** As shown in that table, most OR techniques reflect the viewpoint of one decision maker having one objective. Optimization techniques including linear and nonlinear programming, fall under this category because usually they are employed for minimizing costs in terms of dollars or maximizing monetary benefits from one group's viewpoint subject to various constraints. Often both economical and physical constraints can be incorporated into the constraint equations in optimization problems. Many of the probabilistic techniques like queueing theory, inventory theory, decision theory and Markov chains, fall under the top left cell in Table 1.5.1. An example of a technique designed for handling multiple objectives for a single decision maker is multicriterion modelling (see, for instance, Vincke (1992), Radford (1989), Roy (1985) and Goicoechea et al. (1982)). This method is designed for finding the more preferred alternative solutions to a problem where discrete alternatives are evaluated against criteria ranging from cost (a quantitative criterion) to aesthetics (a qualitative criterion). The evaluations of the criteria for each alternative reflect the objectives or preferences of the single decision maker. In Table 1.5.1, team theory is categorized according to multiple decision makers and one objective because in a sporting event, for instance, each team has the single objective of winning.

Conflict analysis (Fraser and Hipel, 1984), as well as an improved version of conflict analysis called the **graph model for conflict resolution** (Fang et al., 1993), constitute examples of techniques in Table 1.5.1 which can be used for modelling and analyzing disputes in which there are two or more decision makers, each of whom can have multiple objectives. Conflict analysis is a branch of game theory which was specifically designed and developed for studying problems in **multiple objective-multiple participant decision making.** In fact, conflict analysis constitutes a significant expansion of metagame analysis (Howard, 1971) which in turn is radically different from classical game theory (von Neumann and Morgenstern, 1953). A comparison of game theory techniques, their usefulness in OR as well as a description of present and possible future developments are provided by Hipel (1990). Additionally, Hipel (1990) explains how appropriate OR methods can be employed for studying both tactical and strategic problems which arise in decision making. In particular, conflict analysis is especially well designed for handling decision making at the strategic level where often compromise solutions must be reached in order to satisfy a wide range of different interest groups. Within the next subsection, the GDU environmental conflict introduced in Section 1.5.1 is employed to demonstrate how conflict analysis and other techniques like statistical methods from environmetrics can be used for systematically studying complex environmental problems.

Table 1.5.1 Classifications of decision making techniques.

		OBJECTIVES	
		One	Two or More
DECISION MAKERS	One	Most OR Methods	Multicriterion Modelling
	Two or More	Team Theory	Conflict Analysis

Operational research is probably the largest and most widely **known field** within which formal decision making techniques have been developed. Nonetheless, since World War II, other **systems sciences** fields have been started for efficiently solving well structured problems in order to satisfy specified objectives. Besides OR, the systems sciences include **systems engineering** (see, for instance, Checkland (1981) and references contained therein) and **systems analysis** (Miser and Quade, 1985, 1988). A large number of standard textbooks on OR are now available, such as contributions by Hillier and Lieberman (1990) and Wagner (1975). An interesting book on the application of OR methods to various water resources problems, is provided by Loucks et al. (1981). Unfortunately, none of the classical OR texts satisfactorily deal with problems having multiple decision makers (i.e., the second row in Table 1.5.1). However, Rosenhead (1989) and Hipel (1990) clearly point out directions in which OR and other systems sciences fields should be expanded so that more complex problems having many decision makers, unclear objectives and other difficult characteristics, can be properly modelled. Indeed, fields outside of OR, such as **artificial intelligence and expert systems,** as well as **information and decision technologies,** are already tackling challenging research problems in decision making at the strategic level where situations are usually not well structured. A monograph edited by Hipel (1992) on **multiple objective decision making in water resources** contains a sequence of eighteen papers regarding some of the latest developments in tactical and strategic OR techniques along with their application to challenging water resources systems problems.

1.5.3 Conflict Analysis of the Garrison Diversion Unit Dispute

The GDU dispute is employed in this section to explain how the engineering decision making procedure of Figure 1.5.2 can be carried out in practice. As pointed out in Section 1.5.1, the GDU is an important environmental dispute between Canada and the United States. Figure 1.5.1 depicts the location of the GDU irrigation scheme along with the physical facilities such as dams, canals and irrigation fields.

Hipel and Fraser (1980) and Fraser and Hipel (1984) carried out a metagame analysis and conflict analysis, respectively, of the GDU conflict as it existed in 1976, while Fang et al. (1988, 1993) also performed an analysis of the dispute for the situation that took place in 1984. To explain the type of engineering decision making used in the GDU conflict, only the results for 1976 are utilized.

Figure 1.5.3 depicts the general procedure for applying conflict analysis to an actual dispute. Initially, the real world conflict may seem to be confusing and difficult to comprehend. Nonetheless, by systematically applying conflict analysis according to the two main stages of **modelling** and **analysis** the controversy can be better understood in terms of its essential characteristics and potential resolutions. The **modelling stage** consists of ascertaining all the **decision makers** as well as each decision maker's **options** and **relative preferences.** At the **analysis stage,** one can calculate the **stability** of every possible **state** (also called an outcome or scenario) from each decision maker's viewpoint. A state is stable for a decision maker if it is not advantageous for the decision maker to move unilaterally away from it. **Equilibria** or compromise resolutions are states that are stable for every decision maker. The results of the stability analysis can be studied and interpreted by actual decision makers or other interested parties in order to understand their meaning in terms of the actual conflict. These findings may suggest types of **sensitivity analyses** that can be carried out, for example, by seeing how appropriate preference changes affect the overall equilibria. Moreover, the feedback arrows in Figure 1.5.3 indicate that the procedure for applying conflict analysis is done in an **iterative fashion.**

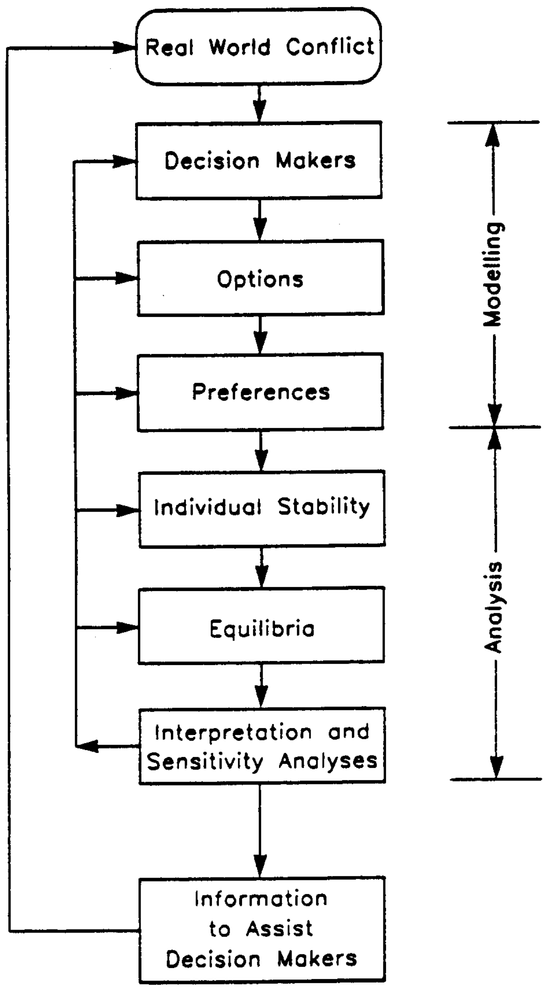

Figure 1.5.3. Applying conflict analysis.

To explain briefly the modelling and analysis of the GDU conflict using the approach of Figure 1.5.3, consider the dispute as it existed in 1976. Table 1.5.2 lists the **decision makers** involved in the conflict along with the courses of actions or **options** available to each decision maker. Briefly, the **U.S. Support** for the project consists of the U.S. Bureau of Reclamation of the U.S. Department of the Interior, the State of North Dakota and support groups within North Dakota. As shown in Table 1.5.2, the U.S. support has three mutually exclusive options available to it. The first one is to proceed to complete the GDU project as approved by Congress while the other two options constitute reduced versions of the full project to appease the Canadian Opposition (option 2) or the U.S. Opposition (option 3).

The **U.S. Opposition** consists mainly of environmental organizations such as the National Audubon Society and the Environmental Protection Agency. It has the single option of taking legal action against the project based upon American environmental legislation (option 4).

Table 1.5.2. Conflict model for the GDU dispute.

Decision Makers and Options	Representative State	
U.S. Support		
1. Complete full GDU.	N	Strategy
2. Building GDU modified to reduce Canadian impacts.	N	for U.S. Support
3. Construct GDU modified to appease U.S..	Y	
U.S. Opposition		
4. Legal action based on environmental legislation.	N	Strategy for U.S. Opposition
Canadian Opposition		
5. Legal action based on the Boundary Waters Treaty of 1909.	Y	Strategy for Canadian Opposition
International Joint Commission (IJC)		
6. Support full GDU.	N	
7. Recommend GDU modified to reduce Canadian impacts.	N	Strategy
8. Support suspension of GDU except for the Lonetree Reservoir.	Y	for IJC
9. Recommend cancellation of the GDU.	N	

The main **Canadian organizations opposed to the GDU** are the Federal Government in Ottawa, the Manitoba Provincial Government and Canadian environmental groups. The single course of action available to the Canadian Opposition is the ability to take legal action based upon the Boundary Waters Treaty of 1909 between the United States and Canada (option 5). This treaty confers legal rights to Canadians citizens for taking action in American courts when water quantity or quality is infringed upon by the Americans. The Americans also have the same rights under this treaty for entering Canadian courts.

The **International Joint Commission** (IJC) was formed under Article VI of the Boundary Waters Treaty as an impartial body to investigate water and other disputes arising between the two nations. Three Canadians and three Americans form the IJC. Whenever the IJC is called upon by the U.S. and Canada to look at a problem, it employs the best scientists from both countries to carry out rigorous scientific investigations in order to come up with a proper environmental impact assessment. As a matter of fact, the quality of the work of the IJC is so well respected that its findings usually significantly influence the preferences of decision makers involved in a given dispute. Because the GDU project is concerned with water quality, the IJC can only give a recommended solution to the problem. As shown in Table 1.5.2, the IJC has basically four mutually exclusive recommendations it could make after its study is completed. One option is to support completion of the full GDU project (option 6) while the other three (options 7 to 9) are reduced versions thereof.

Given the decision makers and options, one can determine strategies for each decision maker and overall states. A **strategy** is formed when a decision maker decides which of his or her options to select and which ones he or she will not take. To explain this further, refer to the column of Y's and N's in Table 1.5.2. A "Y" means that "yes" the option opposite the Y is taken by the decision maker controlling it while a "N" indicates "no" the option is not selected. Notice from Table 1.5.2, that the strategy taken by the U.S. Support is where it takes option 3 and not options 1 and 2. Possible strategies for each of the other three decision makers are also indicated in Table 1.5.2.

After each decision maker selects a strategy, a **state** is formed. Writing horizontally in text the vertical state listed in Table 1.5.2, state (NNY N Y NNYN) is created by the U.S. Support, U.S. Opposition, Canadian Opposition and IJC choosing strategies (NNY), (N), (Y) and (NNYN), respectively, in order to form the overall state.

In the GDU conflict, there is a total of 9 options. Because each option can be either selected or rejected, there is a total of 2^9=512 possible states. However, many of these states cannot take place in the actual conflict because they are **infeasible** for a variety of reasons. For instance, options numbered 1 to 3 are mutually exclusive for the U.S. Support because it can only build one alternative project. Hence, any state which contains a strategy in which the U.S. selects more than one option is infeasible. Likewise, options 6 to 9 are mutually exclusive for the IJC. When all of the infeasible outcomes are removed from the game, less than 50 out of 512 outcomes are left.

From Figure 1.5.3, the final step in modelling a conflict is to obtain relative **preferences** for each decision maker. The most precise type of preference information needed in a conflict study is **ordinal** where states are ranked from most to least preferred. This could include sets of states for which states are equally preferred within each set. Conflict analysis can also handle more general types of preferences such as **intransitive** preferences where a decision maker prefers state x to y, y to z, but z to x. When preferences are **transitive**, a decision maker prefers state x to y, y to z, and x to z. Cardinal utility functions are not used to represent preferences in conflict analysis because in practice they are almost impossible to obtain.

Preferences are often directly expressed in terms of options. For example, the **U.S. Support** most prefers states in which it selects its first option while the IJC chooses option 6. Assuming transitivity, algorithms are available to transform option preferences to state preferences such that the states are ranked from most to least preferred.

To give a general idea of the preferences in the GDU dispute, a preference description for the U.S. Support is now continued. Compared to states for which the full GDU is built, the U.S. Support prefers less states where the full GDU project is not built and the IJC also recommends this.

The **U.S. Opposition** prefers to take legal action if the full project were built. On the other hand, the U.S. Opposition would prefer not to press legal action if the U.S. were to select its third option.

The **Canadian Opposition** most prefers that no project be built. If option 2 were not chosen by the U.S. Support, the Canadian Opposition would prefer to take legal action based on the Boundary Waters Treaty. However, if the IJC recommended a given alternative which the U.S. Support decided to follow, the Canadian Opposition would prefer not to oppose it by going to court. This is because the IJC always carries out first class scientific and economic studies which are greatly respected by the Canadian government as well as others. Additionally, if the dispute were to end up in an American court or perhaps at the international court in the Hague, the court would probably follow the recommendations of the IJC in its ruling. Consequently, this application clearly demonstrates how **good science can dramatically affect the strategic decision making.** In Figure 1.5.2 proper scientific studies are used for the basic design of the project as well as the environmental considerations. Sections 1.2 to 1.4 outline how scientific investigations and related mathematical modelling can be carried out in practice.

Because the IJC is an impartial body, all of the states are equally preferred for it prior to the release of its comprehensive International Garrison Diversion Study Board Report in October, 1976. The report, which was commissioned by the IJC, consists of an overall report plus five detailed reports given as appendices. The five appendices are entitled Water Quality, Water Quantity, Biology, Uses, and Engineering Reports. All of the reports, and especially the first two given above, make extensive use of statistics. Indeed, **the GDU study constitutes an excellent example of how an environmental impact assessment should be executed.**

The GDU conflict model is now fully calibrated in terms of decision makers, their options and their preferences. This game model provides the basic structure within which the **possible strategic interactions** among the decision makers can be studied. More specifically, the systematic examination of the possible moves and counter moves by the decision makers during the possible evolutions of the conflict and the calculation of the most likely resolutions are referred to as the **stability analysis** stage (see Figure 1.5.3). The results of a stability analysis, including sensitivity analyses, can be used, for example, to help support decisions made by people having real power in a conflict. In practice, one would, of course, use a **decision support system** to carry out all the calculations and provide requested advice to a decision maker.

The details of the stability calculations are not given here but can be found in Chapter 2 of the book by Fraser and Hipel (1984) as well as Chapter 6 in the text of Fang et al. (1993). The stability analysis for the GDU was carried out for the situation that existed just prior to the release of the study board reports for the IJC in October, 1976. The state given in Table 1.5.2 is one of the **equilibria** predicted by the conflict analysis study and the one that occurred historically. Notice in Table 1.5.2 that the U.S. Support is going to build a project to appease the American environmentalists and, hence, the U.S. Opposition is not going to court. However, the Canadian Opposition will go to court under the Boundary Waters Treaty because the IJC is recommending a reduced version of the project (option 8).

In the **1984 conflict analysis study,** further reductions were made to the GDU project (Fang et al., 1988). As a matter of fact, all portions of the project that could adversely affect Canada were cancelled due to increased political pressures.

The GDU conflict vividly emphasizes the importance of following the main steps of the engineering decision making procedure of Figure 1.5.2. **Competent scientific and economic studies of the project by the IJC affected directly the political decision making taking place at the strategic level.** For instance, the Canadian Opposition and other interested groups put great faith in the IJC study board reports and this in turn influenced their strategic behaviour in 1976, especially in terms of preferences. By 1984, others were also influenced by **good science** and this led directly to the cancellation of most of this irrigation project which was clearly shown to be environmentally unsound.

This book deals exclusively with environmetrics. However, one should always keep in mind how the results of environmetric and other related scientific studies fit into the overall decision making process of Figure 1.5.2. Even though it may sometimes take a long time, **good science can have highly beneficial effects on decision making.**

1.6 ORGANIZATION OF THE BOOK

1.6.1 The Audience

As defined in Section 1.1, **environmetrics** is the development and application of statistical methods in the environmental sciences. This book focuses upon useful developments in environmetrics coming from the fields of **statistics, stochastic hydrology and statistical water quality modelling.** Of particular interest are time series models that can be employed in the design and operation of large-scale water resources projects, as well as time series models, regression analysis methods and nonparametric methods that can be used in trend assessments of water quality time series. In other words, this book deals with the statistical analyses of both water quantity and water quality problems. Moreover, it clearly explains how these problems can be jointly considered when carrying out environmental impact assessment studies involving trend assessment of water quality variables under the influence of riverflows, seasonality and other complicating factors.

Who will wish to study, apply and perhaps further develop the environmetrics technologies presented in this book? For sure, the environmetrics techniques should be of direct interest to **teachers, students, practitioners and researchers working in water resources and environmental engineering.** However, people from other fields who often consider environmental issues, should also find the contents of this book to be beneficial for systematically investigating their environmental problems. For example, **geographers, civil engineers, urban planners, agricultural engineers, landscape architects** and many others may wish to apply environmetrics methods given in this book to specific environmental problems that arise in their professions. Moreover, keeping in mind the great import of environmetrics and other scientific approaches in the overall decision making process (see Section 1.5), there may be many other professionals who may wish to better understand environmetrics and use it to improve their decision making capabilities. This group of professionals includes **management scientists, operational research workers, business administrators** and **lawyers** who may be employed by government agencies or industry. Finally, because most of the time series models can also be applied to data that are not environmental, there are other professionals who may find this book to be a valuable reference. **Economists**, for instance, who apply time series models to different kinds of economic time series may find useful results in this book that can assist them in their field of study. As a matter of fact, econometrics is defined as the development and application of statistical methods in economics. **Chemical engineers** may also discover useful ideas in the book for application to chemical processes involving input-output relationships.

From a **teaching viewpoint**, this book is designed for use as a course text at the upper undergraduate and graduate levels. More advanced theoretical topics and greater depth of topics could be used in graduate courses. If all of the chapters in the book are covered in depth the book could be used in a two semester (i.e. eight month) course in environmetrics. However, as explained in the next section, there are various routes that can be followed for studying a useful subset of the chapters. Hence, the book could be used in a variety of specially designed one semester environmetrics courses. Exercises are presented at the end of all of the chapters.

Virtually all of the tools presented in the book are highly developed from a theoretical viewpoint and possess appropriate algorithms that permit them to be applied in practice. Therefore, the **methods are completely operational** and can be used now for solving actual problems in environmetrics. **Practitioners** who are studying specific environmental problems can refer to appropriate environmetrics techniques given in the book that are immediately useful to them. Moreover, the **McLeod-Hipel Time Series Package** (McLeod and Hipel, 1992) described in Section 1.7 is a computerized **decision support system** that can be employed by practitioners for applying appropriate environmetrics technologies to their problems and obtaining useful information upon which optimal decisions can be made.

The book holds a treasure house of ideas for **researchers** in environmetrics and time series modelling. More specifically, besides defining useful statistical tools as well as informative applications of the methods to actual data in order to explain clearly how to use them, the book puts the relative importance of environmetrics techniques into proper perspective. Furthermore, based upon both practical and theoretical needs, the book provides guidance as to where further worthwhile research is required.

In the next subsection, the overall **layout** of the book is described and **possible routes** for exploring the countryside of ideas contained in the chapters are traced out. Subsequently, in Section 1.6.3 the book is compared to other available books that deal with specific areas of environmetrics.

1.6.2 A Traveller's Guide

There are many different kinds of environmetrics techniques which can be applied to a wide variety of environmental problems. Consequently, one could envision a substantial number of sequences in which to present topics in environmetrics. For example, one could subdivide topics in environmetrics according to types of application problems. Another approach is to present the statistical methods from simpler to more complex and then to present applications later. The motivation for the **sequence of presentation of topics in this book is pedagogical.** First, it is assumed that a reader of this book has the background acquired after completing one introductory course in probability and statistics. Next, the order of topics in the book is for a reader who is studying environmetrics for the first time. Accordingly, the topics in time series modelling are arranged from simpler to more complex. Throughout the book, practical applications are given so the reader can appreciate how the methods work in practice and the insights that can be gained by utilizing them. After the reader has accumulated a variety of environmetrics tools in the earlier chapters, later in the book more complex environmental impact assessment studies are examined and general methodologies are described for systematically applying appropriate statistical methods from the toolbox of ideas that are available. Indeed, in Part X, general approaches are presented for trend assessment of "messy" environmental data.

Of course, many readers of this book may already have some background in environmetrics and, more specifically, time series modelling. These readers may wish to skip some of the earlier topics and immediately start with subjects presented in later sections in the book. Because readers may have varying backgrounds and reasons for studying environmetrics, there are **many different routes** through which one could tour the territory of environmetrics topics given in this book. The purposes of this section are to outline the topics covered in the book and suggest some good itineraries for the environmetrics traveller, depending upon what types of interesting tourist spots he or she would like to discover.

As can be seen from the Table of Contents, the **24 Chapters** in the book are divided into **10 main Parts**. The titles of the Chapters within each Part provide guidance about the topics given in each Part. For the convenience of the reader, Table 1.6.1 furnishes a tabulation of the titles of each of the ten Parts, along with a listing of the Chapter titles. Readers may also wish to refer to the brief one or two page descriptions given at the start of each Part where it first appears in the book. Finally, an overview of the book is also provided in the Preface.

Within **Part I**, labelled Scope and Background Material, Chapter 1 puts the overall objectives of the book into proper perspective and points out its main contributions to environmetrics. In particular, as explained in Section 1.2, statistics provides a powerful means for enhancing the **scientific method** in the search for solutions to pressing environmental problems. This book emphasizes the use of time series and other statistical methods for carrying out systematic data analysis studies of environmental time series. As pointed out in Section 1.2.4 and described in detail in Part X, a data analysis study consists of the two main stages of **exploratory data analysis** plus **confirmatory data analysis**. When fitting a specific time series model to a sequence of observations at the confirmatory data analysis stage, one can follow the identification, estimation and diagnostic check stages of **model construction** (Section 1.3). Moreover, in a given environmetrics study, one should keep in mind the overall **physical aspects** of the environmental problem (Section 1.4) as well as the influence of scientific studies upon the **decision making process** (Section 1.5). In the second chapter of Part I, some **basic statistical concepts** that are particularly useful in time series modelling are presented.

In order to provide the reader with specific tools that can be used in a data analysis study, **Part II** of the book describes some simple, yet well designed, time series models for fitting to yearly time series. Chapter 3 defines **AR** (autoregressive), **MA** (moving average) and **ARMA** (autoregressive-moving average) models for fitting to **stationary nonseasonal time series**. As explained in Section 2.4, stationarity means that the basic statistical properties of the series do not change over time. In Chapter 4, the **ARIMA** (autoregressive integrated moving average) family of models is defined for describing **nonstationary yearly time series**, where, for example, the mean levels may increase with time. As is the case with all of the models presented in the book, each model in Part II is clearly defined, its main theoretical properties that are useful in practical applications are pointed out, and illustrative examples are employed to explain how to fit the model to real data.

Table 1.6.2 gives a list of all of the time series models presented in the book, acronyms used to describe the models, locations in the book where model definitions, model construction and applications of the models can be found, as well as brief descriptions of the domains of applicability of the models. **The nonseasonal time series models of Part II actually form the theoretical foundations upon which more complicated models of later Parts can be defined.** However, before defining more models after Part II, the manner in which the models of Part II

Table 1.6.1. Parts and chapters in the book.

Part Numbers	Part Titles	Chapter Numbers	Chapter Titles
I	Scope and Background Material	1. 2.	Environmetrics, Science and Decision Making Basic Statistical Concepts
II	Linear Nonseasonal Models	3. 4.	Stationary Nonseasonal Models Nonstationary Nonseasonal Models
III	Model Construction	5. 6. 7.	Model Identification Parameter Estimation Diagnostic Checking
IV	Forecasting and Simulation	8. 9.	Forecasting with Nonseasonal Models Simulating with Nonseasonal Models
V	Long Memory Modelling	10. 11.	The Hurst Phenomenon and Fractional Gaussian Noise Fractional Autoregressive-Moving Average Models
VI	Seasonal Models	12. 13. 14. 15.	Seasonal Autoregressive Integrated Moving Average Models Deseasonalized Models Periodic Models Forecasting with Seasonal Models
VII	Multiple Input-Single Output Models	16. 17. 18.	Causality Constructing Transfer Function-Noise Models Forecasting with Transfer Function-Noise Models
VIII	Intervention Analysis	19.	Building Intervention Models
IX	Multiple Input-Multiple Output Models	20. 21.	General Multivariate Autoregressive-Moving Average Models Contemporaneous Autoregressive-Moving Average Models
X	Handling Messy Environmental Data	22. 23. 24.	Exploratory Data Analysis and Intervention Modelling in Confirmatory Data Analysis Nonparametric Tests for Trend Detection Regression Analysis and Trend Assessment

are systematically fitted to data is explained in Part III. Specifically, the identification, estimation and diagnostic check stages of model construction are clearly explained using practical applications. Later in the book, these basic model building approaches of Part III are extended and modified for use with all the other time series models presented in Parts V to IX.

Two major types of application of time series models are forecasting and simulation. In **Part IV**, procedures are presented for forecasting (Chapter 8) and simulating (Chapter 9) with the linear nonseasonal models of Part II. The objective of **forecasting** is to obtain the best possible estimates or forecasts of what will happen in the future based upon the time series model fitted to the historical time series as well as the most recent observations. When operating a system of reservoirs for producing hydroelectrical power, forecasts of the inflows to the reservoirs are used for developing an optimal operating policy which maximizes profits from the sale of the electricity. In **simulation**, time series models are utilized to produce possible future sequences of the phenomenon being modelled. Simulation can be used for designing a large scale engineering project and for studying the theoretical properties of a given time series model.

All of the time series models presented in this book can be used for forecasting and simulation. As a matter of fact, extensive experiments in forecasting and simulation given in many parts of the book demonstrate the great import of forecasting and simulation in environmetrics as well as the ability of ARMA-type models to perform better than their competitors. Table 1.6.3 lists the locations in the book where the forecasting and simulation procedures and applications are given for many of the models given in Table 1.6.2.

Long memory models were developed within the field of stochastic hydrology in an attempt to explain what is called the **Hurst Phenomenon**. Within Chapters 10 and 11 of **Part V**, two long memory models called **FGN** (Fractional Gaussian noise) and **FARMA** (fractional autoregressive-moving average) models, respectively, are defined for fitting to annual geophysical time series. Additionally, the Hurst phenomenon is defined in Chapter 10 and a proper explanation for the Hurst phenomenon is put forward. More specifically, it is demonstrated using extensive simulation experiments in Chapter 10, that **properly fitted ARMA models can statistically preserve historical statistics related to the Hurst phenomenon.**

By the end of Parts IV or V, the reader has a solid background in nonseasonal time series modelling. He or she knows the basic definitions of some useful yearly models (Part II and also Part V), understands how to apply these models to actual data sets (Part III), and knows how to calculate forecasts and simulated sequences with these models (Part IV). The reader is now in a position to appreciate how some of these nonseasonal models can be extended for use with other kinds of data. As indicated in Tables 1.6.1 and 1.6.2, models for fitting to seasonal data are described in Part VI immediately after the Part on long memory models. The three types of **seasonal models** described in Chapters 12 to 14 are the **SARIMA** (seasonal autoregressive integrated moving average), deseasonalized, and **PARMA** (periodic autoregressive-moving average) models, respectively. A special case of the PARMA models of Chapter 14 is the set of **PAR** (periodic autoregressive) models. In order to reduce the number of parameters in a PAR model, the **PPAR** (parsimonious periodic autoregressive model) can be employed. The deseasonalized and periodic models are designed for fitting to natural time series in which statistical characteristics such as the mean and variance must be accounted for in each season. Furthermore, periodic models can describe autocorrelation structures which change across the seasons. **Forecasting experiments** in Chapter 15 clearly demonstrate that PAR models are well suited for forecasting seasonal riverflows. SARIMA models can be used for forecasting nonstationary socio-economic time series such as water demand and electricity consumption.

In many natural systems, a single output or response variable is driven by one or more input or covariate series. The **TFN** (transfer function-noise) model of Part VII is designed for stochastically modelling the dynamic relationship between the input series and the single

Table 1.6.2. Time series models presented in the book.

Categories	Model Names	Acronyms	Definitions	Model Construction	Applications	Domains of Applicability
Linear Nonseasonal Models (Parts II and III)	Autoregressive	AR	Section 3.2	Part III	Introduced in Part II and completed in Part III.	Stationary annual time series.
	Moving average	MA	Section 3.3			
	Autoregressive - moving average	ARMA	Section 3.4			
	Autoregressive integrated moving average	ARIMA	Section 4.3		Section 4.3.3 Part III	Nonstationary yearly time series.
Long Memory Models (Part V)	Fractional Gaussian noise	FGN	Section 10.4.2	Sections 10.4.3 & 10.4.4	Section 10.4.7	Stationary annual series having long memory.
	Fractional autoregressive-moving average	FARMA	Section 11.2.2	Section 11.3	Section 11.5	
Seasonal Models (Part VI)	Seasonal autoregressive integrated moving average	SARIMA	Section 12.2	Section 12.3	Sections 12.4 and 14.6	Seasonal time series having non-stationarity within each season.
	Deseasonalized	DES	Section 13.2	Section 13.3	Sections 13.4 and 14.6	Seasonal time series for which mean and/or variance are preserved in each season.
	Periodic autoregressive	PAR	Section 14.2.2	Section 14.3	Section 14.6	Seasonal time series for which correlation structure varies across the seasons.
	Periodic autoregressive-moving average	PARMA	Section 14.2.3	Section 14.7		
	Parsimonious periodic autoregressive	PPAR	Section 14.5.2	Section 14.5.3	Section 14.6	

Cont'd...

Table 1.6.2. Time series models presented in the book (continued).

Categories	Model Names	Acronyms	Definitions	Model Construction	Applications	Domains of Applicability
Single Output and Multiple Input Models (Part VII)	Transfer function – noise	TFN	Sections 17.2 (single input) and 17.5.2 (multiple inputs)	Sections 17.3 (single input) and 17.5.3 (multiple inputs)	Sections 17.4 (single input) and 17.5.4 (multiple inputs)	Nonseasonal time series and seasonal data which are usually first deseasonalized.
Single Output Having Multiple Interventions, Missing Data, and Multiple Input Series (Part VIII)	Intervention		Section 19.2.2 (multiple interventions)	Section 19.2.3 (multiple interventions)	Sections 19.2.4 and 19.2.5 (multiple interventions)	Nonseasonal data and seasonal time series which are usually first deseasonalized. Mean level of output series can be affected by one or more intervention series and missing values in the output series can be estimated. Can also handle multiple input series.
			Section 19.3.3 (missing data)	Section 19.3.4 (missing data)	Sections 19.3.5 and 19.3.6 (missing data)	
			Section 19.4.2 (multiple interventions and missing data)	Section 19.4.3 (multiple interventions and missing data)	Sections 19.4.4 and 19.4.5 (multiple interventions and missing data)	
			Section 19.5.2 (multiple interventions, missing data and input series)	Section 19.5.3 (multiple interventions, missing data and input series)	Section 19.5.4 (multiple interventions, missing data and input series)	
Multiple Ouput and Multiple Input Models (Part IX)	General multivariate autoregressive-moving average	Multivariate ARMA	Section 20.2	Section 20.3		Nonseasonal and and deseasonalized time series for which there is feedback between series.
	Contemporaneous autoregressive-moving average	CARMA	Section 21.1	Section 21.3 and Appendix A21.1	Section 21.5	

Table 1.6.3. Forecasting and Simulation.

<table>
<tr><th rowspan="2">Model Names</th><th colspan="2">Forecasting</th><th colspan="2">Simulation</th></tr>
<tr><th>Algorithms</th><th>Applications</th><th>Algorithms</th><th>Applications</th></tr>
<tr><td>AR</td><td rowspan="4">Minimum mean square error forecasts defined in Section 8.2</td><td rowspan="4">Section 8.3</td><td rowspan="3">Sections 9.2, 9.3, 9.4, 9.6, and 9.7</td><td rowspan="3">Sections 9.8, 10.5 and 10.6</td></tr>
<tr><td>MA</td></tr>
<tr><td>ARMA</td></tr>
<tr><td>ARIMA</td><td>Sections 9.2 to 9.7</td><td>Section 9.8</td></tr>
<tr><td>FGN</td><td>Section 10.4.5</td><td>Section 8.3.4</td><td>Section 10.4.6</td><td>Section 10.5.3</td></tr>
<tr><td>FARMA</td><td>Section 11.4.3</td><td>Section 8.3.4</td><td>Section 11.4.2</td><td></td></tr>
<tr><td>SARIMA</td><td>Section 15.2.2</td><td rowspan="3">Sections 15.3 and 15.4</td><td>Section 9.5</td><td rowspan="2"></td></tr>
<tr><td>DES</td><td>Sections 13.5 and 15.2.3</td><td>Section 13.5</td></tr>
<tr><td>PAR</td><td rowspan="3">Sections 14.8 and 15.2.4</td><td rowspan="3">Section 14.8</td><td>Section 14.8.2</td></tr>
<tr><td>PARMA</td><td></td><td></td></tr>
<tr><td>PPAR</td><td>Section 15.4</td><td>Section 14.8.2</td></tr>
<tr><td>TFN</td><td>Section 18.2</td><td>Sections 18.3 and 18.4</td><td>Section 18.5.3</td><td></td></tr>
<tr><td>CARMA</td><td></td><td></td><td>Section 21.4 and Appendix A 21.2</td><td></td></tr>
</table>

response. In **stochastic hydrology**, one may wish to model formally the manner in which precipitation and temperature cause riverflows. Qualitatively, a TFN model for describing this dynamic situation is written as:

Riverflows = Precipitation + Temperature + Noise

where the noise term is modelled as an ARMA model from Chapter 3. An example of the design of a TFN model for use in **stochastic water quality modelling** is:

Water Quality Variable (ex. Phosphorous)	=	**Riverflows**	+	**Other Water Quality Variables (ex. Water Temperatures and Turbidity)**	+	**Noise**

Because the above qualitative equation contains both water quantity and water quality time series, **the TFN model, as well as the related intervention model of Part VIII, provide a formal means for connecting stochastic hydrology (i.e., water quantity modelling) with statistical water quality modelling.**

As shown in the Table of Contents and Table 1.6.1, there are three chapters in Part VII. Chapter 16 explains how the **residual cross-correlation function** can be used to detect different kinds of **causality** between two variables. In Chapter 17, the **TFN model** is formally defined and flexible model building procedures along with illustrative applications are presented. The **forecasting experiments** of Chapter 18 clearly demonstrate that TFN models provide more accurate forecasts than their competitors. In fact, one of the forecasting experiments shows that **a simple TFN model forecasts better than a very complicated and expensive conceptual model.**

The **intervention model of Part VIII** constitutes a special type of TFN that is especially well suited for use in **environmental impact assessment**. Qualitatively, the intervention model has the following form:

Output Variable	=	**Multiple Inputs**	+	**Multiple Interventions**	+	**Missing Data**	+	**Noise**

In addition to describing the effects of multiple input series upon a single response variable, the intervention model can simultaneously model the effects of one or more external interventions upon the mean level of the output series, estimate missing observations and handle correlated noise through an ARMA noise component. For the case of a **water quality application**, an intervention model could be written as:

Water Quality Variable	**=**	**Riverflows**	**+**	**Other Water Quality Variables**	**+**	**Multiple Interven-ions**	**+**	**Missing Data**	**+**	**Noise**

The formal modelling and analysis of a data set using the intervention model is popularly referred to as **intervention analysis.** As pointed out in Section 1.2.4 and explained in detail in Section 19.4.5, a special form of the above intervention model can be designed for formally describing the drop in the mean level of the phosphorous series shown in Figure 1.1.1. Applications of a variety of intervention models to stochastic hydrology and environmental engineering data sets in Chapter 19 and Section 22.4, demonstrate that **the intervention model is one of the most comprehensive and flexible models available for use in environmetrics.** A sound and flexible theoretical design coupled with comprehensive model construction methods permit the intervention model to be conveniently and expeditiously applied in practice.

When there is **feedback** in a system, the input affects the output but the output can in turn have a bearing upon the input. For example, consider the situation where rivers drain into a large lake. The flow in a given river is caused by precipitation. However, evaporation from the large lake which is filled by the rivers causes precipitation that once again creates riverflows. To model formally this type of situation, one can employ the **multivariate ARMA** family of models of **Part IX** that has the form:

Multiple Outputs	**=**	**Multiple Inputs**	**+**	**Noise**

The terminology multivariate is employed because there are multiple output series. A drawback of the multivariate ARMA models is that they contain a great number of parameters. To reduce significantly the number of model parameters, one can employ the CARMA (contemporaneous ARMA) multivariate model of Chapter 21 for special types of applications. Good model construction methods, including a parameter estimation method that is efficient both statistically and computationally, are available for use with CARMA models.

A major strength of this book is the presentation of a variety of useful techniques that can be employed in complex environmental impact assessment studies. The objective of **Part X** is to explain how a variety of statistical methods can be used for **detecting and modelling trends,** as well as other statistical properties, in both water quantity and water quality time series. This is carried out within the overall framework of exploratory and confirmatory data analyses referred to in Sections 1.2.4 and 22.1. Within Section 22.3 a variety of informative **graphical tools** are described for use as **exploratory data analysis** tools. Additionally, in Section 24.2.2 it is explained how a technique called **robust locally weight regression smoothing** can be employed for tracing a trend that may be present in a time series. At the **confirmatory data analysis** stage, the three statistical approaches that are employed consist of **intervention analysis** (Chapter 19 and Section 22.4), **nonparametric tests** (Chapter 23) and **regression analysis** (Section 24.2.3).

Within a given data analysis study, one must select the most appropriate exploratory and confirmatory data analysis tools in order to discover, model and analyze the important statistical properties of the data. In fact, **data analysis** is composed of both an art and a craft. The **craft** consists of a knowledge and understanding of the main types of statistical tools that are available. This book, for example, describes and explains the capabilities of a wide variety of statistical methods. The **art** of data analysis is using the most appropriate statistical methods in an innovative and efficient manner for solving the data analysis problems currently being addressed. **The best way to explain how the art and craft of data analysis are carried out in practice is through the use of comprehensive real world case studies.**

Three major data analysis studies are presented in **Part X** of the book for carrying out trend assessments of water quality and water quantity time series. Each of the studies requires the development of a methodological approach within the encompassing structure of exploratory and confirmatory data analyses. Table 1.6.4 provides a list of the **three trend analysis methodologies** presented in Part X, the types of data analysis problems each procedure is applied to, brief descriptions of the methodologies, and the section numbers where explanations are provided. Notice that the first and third approaches in Table 1.6.4 deal with trend assessment of water quality and water quantity time series measured in rivers while the second procedure is concerned with water quality observations taken from a large lake. In all three studies, informative graphical techniques are employed as **exploratory data analysis** tools. At the **confirmatory data analysis** stage, different statistical methods are employed for trend modelling. Consider the first trend assessment methodology listed in Table 1.6.4. After filling in missing observations using the approach of Section 22.2, the time series approach of **intervention analysis** (Part VIII) is used to model trends in water quality series trends due to cutting down a forest in the Cabin Creek river basin of Alberta, Canada. As a matter of fact, the general form of one of the intervention models developed to describe the problem studied in Section 22.4.2 is:

Cabin Creek Water Quality Series	=	**Monthly Interventions**	+	**Cabin Creek Flows**	+	**Middle Fork Water Quality Series**	+	**Noise**

Because the trees were not cut down in the nearby Middle Fork river basin, the water quality series from this river can account for changes in the same Cabin Creek water quality series that are not due to cutting down the trees. The estimate for the intervention parameter for a given month in the intervention component in this general intervention model provides an estimate for the change in the mean level of the water quality variable for the Cabin Creek for the month. Moreover, this model also stochastically accounts for the influence of riverflows upon the water quality variable in the Cabin Creek.

The environmental data used in the second and third studies in Table 1.6.4 are very "**messy**" because they possess undesirable properties such as having many missing values and possessing a large number of outliers. Consequently, **nonparametric methods** are used as confirmatory data analysis tools in these two studies for trend detection as well as other purposes. As explained in Chapter 23, nonparametric techniques have fewer underlying assumptions than competing parametric approaches and, therefore, are better designed for use with messy data. In fact, because nonparametric methods can be used for investigating a wide range of statistical characteristics that may be embedded in messy water quality data, Chapter 23 as well as some parts of Section 24.3 are dedicated to explaining how nonparametric tests can be effectively employed in environmental impact assessment studies. The main emphasis in the discussions of nonparametric methods is their use in **trend detection and modelling.** A review of **hypothesis testing** using nonparametric or parametric test statistics is presented in Section 23.2.

The third approach in Table 1.6.4 employs the regression analysis method called **robust locally weight regression smooth** for clearly tracing trends in time series plots of messy water quality data. Besides being used in graphical procedures at the exploratory data analysis stage, **regression analysis** can also be employed as a confirmatory data analysis method for a wide variety of purposes, including the estimation of the shapes and magnitudes of trends. For example, the last entry in Table 23.5.1 summarizes how regression analysis is employed in the Lake Erie water quality study of Chapter 23. Even though regression analysis usually constitutes a parametric technique, it can be used with data that are unevenly spaced.

The intervention model as well as the other types of time series analysis models presented in this book assume that observations are available at evenly spaced time intervals. If there are gaps in the data, then one must estimate the missing observations before fitting a time series model to the data set. Table 1.6.5 lists the **data filling techniques** described in this book as well as the main types of situations in which they can be used. A general discussion of estimating missing values is presented in Section 19.3 along with a detailed description of the intervention analysis approach to data filling.

Table 1.6.1 provides a summary of the ten main topics or Parts into which the book is divided. Each Part, in turn, is subdivided into a number of chapters. The **contents of each chapter** consist of the **introduction, main sections, conclusions, appendices, problems, and references.** Within the main sections of each chapter, any statistical methods that are described are usually accompanied by **practical environmental applications** so the reader can appreciate their usefulness in environmetrics.

Notice in Table 1.6.1 that the **model construction methods** for the linear time series models of Part II are presented in Part III. However, for each of the other time series models of Table 1.6.2, the model building procedures are usually given just after the model is defined. **Illustrative applications** are always used for demonstrating how a given time series model is fitted to a data set using appropriate model building techniques. The **types of environmental time series** used in the applications include water quantity, water quality, precipitation, ambient temperature, as well as miscellaneous series such as tree ring indices and mud varve thicknesses.

Depending upon the background of the reader, there are a **variety of different routes** he or she can follow when travelling through the environmetrics terrain given in the book. A neophyte in environmetrics may wish to follow sequentially the entire itinerary summarized in Table 1.6.1 in a **two semester course.** A **one semester course** could cover Part I, Chapter 3, Part III, Chapter 8, Chapters 13 and 14 in Part VI, Chapters 16 and 17 in Part VII, plus some Parts of Chapter 19. Someone who already has a background in basic ARMA modelling may wish to extend his knowledge by studying the more complex types of time series models given in Parts VII to IX and listed in Table 1.6.2. A course on **statistical environmental impact assessment** should concentrate on Parts VII, VIII and X. Someone who is mainly interested in **stochastic hydrology** should not miss studying Part V. Courses that emphasize **forecasting and simulation** can cover the topics listed in Tables 1.6.3.

Table 1.6.4. Trend assessment methodologies.

Methodologies	Kinds of Data	Descriptions	Sections
Trend Assessment Using Intervention Analysis (Chapter 22)	Water quality and water quantity time series measured in rivers.	Intervention analysis is used to describe the effects of cutting down a forest upon the mean levels of water quality and quantity time series. A seasonal adjustment data filling method is used to estimate missing values prior to fitting the intervention models.	Section 22.2 describes the seasonal adjustment data filling method. The exploratory data analysis is done in Section 22.3 while the intervention modelling is carried out in Section 22.4.
Trend Analysis of Water Quality Data Measured in Lakes (Chapter 23)	Water quality observations from a large lake.	Nonparametric trend tests and other statistical methods are used to detect trends in water quality variables in a lake that may be affected by nearby industrial developments.	Exploratory and confirmatory data analysis results are presented in Section 23.5. Table 23.5.1 lists all of the statistical methods used in the study.
Trend Analysis of Messy Water Quality Time Series Measured in Rivers (Chapter 24)	Messy water quality data and water quantity time series measured in rivers.	Procedures are presented for accounting for the effects of flow upon a given water quality series and eliminating any trend in the flow before its effect upon the water quality series is removed. The Spearman partial rank correlation test is used to detect trends in water quality time series when the effects of seasonality are partialled out.	Section 24.2.2 describes the robust locally weighted regression smooth for tracing trends. In Section 24.3, the overall trend analysis methodology is presented and applied.

Table 1.6.5. Data filling methods described in the book.

Techniques	Purposes	Locations
Intervention Analysis	The intervention model can be used to estimate missing observations simultaneously with other parameters in the intervention model. No more than about 5% of the observations in a time series should be missing when using the intervention model for data filling.	Chapter 19 is dedicated to describing and applying various versions of the intervention model. Section 19.3 explains in detail how the intervention model is used for filling in data. Besides Chapter 19, other applications of intervention analysis are presented in Section 22.4.2.
Seasonal Adjustment	The seasonal adjustment algorithm is designed for estimating missing values in seasonal time series when there is a great number of missing observations and there may be long time periods over which no measurements were taken.	The seasonal adjustment algorithm is defined and applied in Section 22.2.
Back Forecasting	A TFN model is used to connect two or more time series where the series overlap in time. Then the calibrated TFN model is used to "back forecast" the unknown values of the shorter series given as the response variable.	The TFN model along with model building techniques are presented in Chapter 17, while the technique of back forecasting is described in Section 18.4.

1.6.3 Comparisons to Other Available Literature

As mentioned in Section 1.1, this is a book about **environmetrics**. The techniques and methodological approaches presented in this book draw upon developments in the fields of **statistics, stochastic hydrology** and **statistical water quality modelling**. One purpose of this section is to point out some of the main books, journals and other literature from these three areas that can provide valuable **complementary reference material** for the reader. Within each chapter, **comprehensive references** are provided for the specific topics contained in the chapter. A second objective of this section is to **compare** the main contents of this book to other available literature. As is clearly explained, **this book constitutes a unique contribution to environmetrics.**

From the **field of statistics**, the major type of models used in this book is a wide variety of **time series analysis models.** Because time series analysis is employed extensively in many fields, a vast body of literature has evolved, especially during the past three decades when the advent of the electronic computer made it possible for both simple and complex time series models to be conveniently applied to large data sets. The **seminal textbook publication** which furnishes a systematic and comprehensive presentation of many time series models is the book of Box and Jenkins called "Time Series Analysis: Forecasting and Control." The first edition of their book appeared in 1970 while the second one was published in 1976. Besides defining useful time series models such as ARMA, ARIMA, SARIMA and TFN models, Box and Jenkins explain how to fit time series models to data sets by following the identification, estimation and diagnostic check stages of model construction. As a result, time series modelling became widely accepted for applying to practical problems in many different fields. For example, time series analysis has been widely used in **economics** for forecasting economic time series (Nelson, 1973; Montgomery and Johnson, 1976; Granger and Newbold, 1977; Firth, 1977; Makridakis and Wheelwright, 1978; Granger, 1980; Abraham and Ledolter, 1983; Pankratz, 1983) and in **electrical engineering** for estimating the state of a system in the presence of additive noise (Ljung, 1987; Haykin, 1990). In fact, because the time series analysis work of Box and Jenkins (1976) has become so widely adopted, many books, including this one, employ the notation of Box and Jenkins when defining time series models. Besides the book of Box and Jenkins (1976), **textbooks by statisticians** such as Jenkins and Watts (1968), Hannan (1970), Anderson (1971), Kendall (1973), Brillinger (1975), Chatfield (1975), Fuller (1976), Jenkins (1979), Priestley (1981), Pandit and Wu (1983), McLeod (1983), Vandaele (1983), Young (1984), and Brockwell and Davis (1987) furnish in a pedagogical fashion well explained accounts of developments in time series analysis. Within the statistical literature, the major **journals** to refer to for research and review articles include the Journal of the Royal Statistical Society (Series A, B and C), Biometrika, Journal of the American Statistical Association, the Annals of Statistics, Journal of Time Series Analysis, International Journal of Forecasting, and Communications in Statistics. Additionally, **proceedings from conferences** hosted by statistical societies and also by individuals, provide other valuable sources for research material on time series analysis. McLeod (1987), for example, edited a book of conference papers on stochastic hydrology. Anderson has edited 12 conference proceedings since 1976 (see, for instance, Anderson (1979) and Anderson et al. (1985)).

A number of **books** on **time series analysis in hydrology and water resources** have been written by authors such as Fiering (1967), Yevjevich (1972), Clarke (1973), Kottegoda (1980), Salas et al. (1980), Bras and Rodriguez-Iturbe (1985), and McCuen and Snyder (1986). A monograph edited by Hipel (1985) contains many original contributions on time series analysis in water resources. Both the water resources systems book of Loucks et al. (1981) and the stochastic modelling text of Kashyap and Rao (1976) contain some chapters on the modelling of hydrological time series. Moreover, the book of Helsel and Hirsch (1992) on statistical methods in water resources has chapters on exploratory data analysis, regression analysis and trend tests. Although the measurement and subsequent analysis of water resources time series started a long time ago, the proliferation of time series analysis research in water resources commenced in the early 1960's. In fact, the **journals** Water Resources Bulletin and Water Resources Research, founded at that time by the American Water Resources Association and the American Geophysical Union, respectively, have published papers on time series analysis throughout their history. Other water resources journals that present applications and theoretical developments in time series analysis, include Stochastic Hydrology and Hydraulics, the Journal of Hydrology, Journal of the Water Resources Planning and Management Division of the American Society of Civil Engineers (ASCE), Journal of the Hydraulics Division of ASCE, Advances in Water Resources and Hydrological Sciences Bulletin. Moreover, proceedings from water resources conferences, such as the ones edited by Shen (1976), McBean et al. (1979a,b) and Shen et al. (1986), provide a rich variety of papers on time series analysis in water resources. In addition to time series models, the international conference on **Stochastic and Statistical Methods in Hydrology and Environmental Engineering** held at the University of Waterloo from June 21 to 23, 1993, had many paper presentations regarding the other kinds of stochastic models shown in Table 1.4.1. This conference was held in the honour and memory of the late Professor T. E. Unny.

In addition to time series analysis methods from statistics and stochastic hydrology, this book also deals with ideas from **statistical water quality modelling.** Some of the models described in the stochastic hydrology books referred to in the previous paragraph can be applied to water quality time series. However, the **intervention model** of Chapters 19 and 22 in this book as well as the **TFN model** of Part VII, constitute time series models that are especially well designed for simultaneously modelling both water quality and quantity series. When the environmental data are quite messy and there are a great number of missing values, one may have to employ **nonparametric tests** (Kendall, 1975) in an environmental impact assessment study. In his book on statistical methods for environmental pollution monitoring, Gilbert (1987) makes extensive use of nonparametric techniques. A monograph edited by Hipel (1988) on nonparametric approaches to environmental impact assessment contains papers having applications of nonparametric trend tests to water quality time series. Papers regarding the use of time series models and nonparametric tests in water quality modelling appear in many of the water resources journals referred to in the above paragraph. In Chapter 23 and Section 24.3 of this book a number of useful nonparametric trend tests are presented along with water quality applications. Other **environmental journals** having statistical papers on water quality modelling include Environmetrics, Journal of the Environmental Engineering Division of ASCE, Environmental Management, and Environmental Monitoring and Assessment. Additionally, a number of **conference proceedings** on statistical water quality modelling are available. For instance, El-Shaarawi and Esterby (1982), El-Shaarawi and Kwiatowski (1986) and Chapman and El-Shaarawi (1989) have edited conference proceedings on water quality monitoring and assessment.

Given the impressive array of literature available in time series analysis, stochastic hydrology and statistical water quality modelling, one is tempted to ask what are the **relative contributions** of this book. In fact, there are many ways in which this book combines, enhances and extends previous research. First, the current book **combines the aforementioned three areas in a coherent and systematic manner in order to have a comprehensive book on time series methods in environmetrics.** As a result, stochastic hydrology and statistical water quality modelling are no longer considered to be separate fields. Rather, techniques and approaches from both areas are used in combination with other statistical methods to confront challenging environmental problems. Secondly, this book contains the most **recent and useful developments in time series modelling.** As a matter of fact, virtually all of the time series and other statistical methods presented in the book are well developed theoretically and a range of flexible model construction algorithms are available to allow them to be immediately applied to real world data sets. Thirdly, the book contains **many theoretical contributions and applications** that have been developed by the authors and their colleagues. Although much of the research has appeared during the past fifteen years in various water resources and statistical journals, this is the first time that **it is published in a pedagogical fashion within a single document.** Fourthly, this book puts great emphasis on the use of both time series analysis and nonparametric methods in **environmental impact assessment** studies. Using challenging real world applications, Parts VII, VIII and X in the book clearly explain how this is accomplished in practice. Finally, as is explained in Section 1.6.2, the book is designed to be as flexible as possible so that it **can satisfy the specific needs of individual readers.** It could, for example, be used as a one semester course in environmetrics. A person who is solely interested in stochastic hydrology or statistical water quality modelling could refer to the appropriate chapters in the book, as pointed out in Section 1.6.2. The book could also be used as a main text on time series modelling within a statistics department.

1.7 DECISION SUPPORT SYSTEM FOR TIME SERIES MODELLING

To carry out a proper data analysis study, one requires the employment of a flexible **Decision Support System** (Sage, 1991). The **McLeod-Hipel Time Series (MHTS) Package** (McLeod and Hipel, 1992) constitutes a comprehensive decision support system for performing extensive data analysis investigations in order to obtain easily understandable results upon which sound decisions can be made.

As explained in Section 1.2.4, **data analysis** consists of both exploratory and confirmatory data analyses. A wide variety of **informative graphical methods** are contained in the MHTS package as **exploratory data analysis** tools to allow a user to clearly visualize the key statistical characteristics of the time series that he or she is studying. Some of the many graphical methods contained in the MHTS system are described in Section 22.3 as well as elsewhere in the book. Figure 1.1.1, for example, is a simple yet informative plot created by the MHTS package that clearly shows the step drop in the phosphorous levels in a river due to the introduction of tertiary treatment in upstream sewage plants.

At the **confirmatory data analysis** stage, the MHTS package can, for instance, fit a model to a data set in order to provide a more precise mathematical description of the main statistical properties of the data. For example, in Section 19.4, the MHTS package is employed for fitting the most appropriate **intervention model** to the time series displayed in Figure 1.1.1. Besides mathematically modelling the stochastic structure of the time series, the intervention model provides an efficient estimate of the magnitude of the drop in the mean level of the phosphorous series in Figure 1.1.1 after the intervention of introducing tertiary sewage treatment.

The MHTS package can be employed for fitting virtually all of the different kinds of time series models listed in Table 1.6.2 to time series. For each type of time series model, the three stages of **model construction** depicted in Figure 1.3.1 are adhered to when using the MHTS package. Moreover, the most appropriate and up-to-date tools are used by the MHTS package for each kind of model during the identification, estimation and diagnostic check stages of the model development.

The MHTS package allows a practitioner or researcher to employ calibrated time series models for carrying out **applications** such as **forecasting and simulation experiments.** Table 1.6.3, for example, summarizes a wide variety of forecasting and simulation investigations described in the book.

Each data analysis study usually contains unique problems and challenges that require specialized attention. The MHTS package furnishes the user with a **comprehensive array of tools** from which he or she can select the most appropriate ones in order to discover and build the most effective explanations and solutions to the problems being studied. This inherent comprehensive, yet flexible, design of the MHTS system is especially needed when dealing with the type of **messy environmental data** described in Part X of the book. Table 1.6.4, for instance, lists some general kinds of **trend assessment** studies that can be conviently executed in an iterative, yet logical and systematic fashion, using the MHTS package.

The MHTS package is extremely **user-friendly** and is designed for use by both novices and experts in time series modelling and analysis. Additionally, the MHTS package is **menu-driven** using attractive screen displays, operates on **personal computers** that use **DOS,** and can support a variety of dot-matrix and laser printers.

The MHTS package is probably the most advanced decision support system available which employs such a **rich range of ARMA-type, as well as other kinds of time series methods,** in statistical decision making. The MHTS package permits data analysis methods to be effectively utilized within the **scientific method** described in Section 1.2 and summarized in Figures 1.2.1 to 1.2.3, as well as the overall structure of **engineering decision making** explained in Section 1.5.1 and depicted in Figure 1.5.2. Finally, the MHTS package places a set of valuable statistical tools directly into the hands of decision makers working in **environmetrics who must make real decisions now about complex environmental problems.**

1.8 CONCLUDING REMARKS

The main purpose of this book is to present the art and craft of environmetrics for modelling water resources and environmental systems. The craft consists of a set of useful statistical tools while the art is composed of general procedures or methodologies for applying these tools to environmental time series. As pointed out in Section 1.6.2, the **tools of the trade** are stressed earlier in the book while **general methodologies** are presented after the reader has some tools to work with.

The **topics** on environmetrics given in this book and summarized in Table 1.6.1 are based upon **contributions from the fields of statistics, stochastic hydrology and statistical water quality modelling.** The main set of statistical methods used in the book is a range of families of **time series models.** Other statistical techniques include **graphical methods, nonparametric**

trend tests and regression models. Within the framework of the intervention model, for example, both water quality and water quantity data can be simultaneously considered in an environmental impact assessment study (see Chapters 19 and 22).

In Section 1.2, the **scientific method** is described and it is explained how statistics can enhance the scientific method so that good solutions to pressing environmental problems can be efficiently and expeditiously found. When carrying out a **scientific data analysis** study, informative exploratory and confirmatory data analysis tools can be employed. Graphical methods can be used as **exploratory data analysis** methods for discovering general statistical properties of a given data set. At the **confirmatory data analysis** stage, both parametric and nonparametric methods can be employed for rigorously modelling from a mathematical viewpoint the main statistical characteristics. For example, the intervention model from Part VIII can be utilized to model the shape and magnitude of the trend in the phosphorous data in Figure 1.1.1.

As listed in Table 1.6.2, a **wide variety of time series models** are presented in the book. In Section 1.3, it is explained how one should follow the identification, estimation and diagnostic check stages of **model construction** in order to design a parsimonious model that provides a reasonable statistical fit to the time series. In any statistical study, one should keep in mind the physical aspects of the environmental problem. Section 1.4 explains how the **hydrological cycle** forms a sound environmental system model for use in environmetrics.

In **environmental decision making**, one should remember how the results of an environmetrics study can influence the overall political decisions which are eventually made when a specific alternative solution to a given environmental problem is selected. As explained in Section 1.5, rigorous scientific studies can result in decisions that are environmentally acceptable. Moreover, active participation by scientists and engineers in both tactical and strategic decision making will allow better solutions to be reached for solving pressing environmental problems. An analogy with the success of Japanese industry will help to explain why this should be so. A majority of people at all levels of management, including the board of directors of most major Japanese corporations, have technical training as engineers and scientists (Reich, 1983). In addition, over a time period of a few decades these corporate leaders worked their way from the bottom of a given company to the top. Because they understand both the technical and social problems at all levels in the company, they are properly educated to make sound tactical and strategic decisions. The worldwide large sales of high quality Japanese products attest to the success of having wise leadership. Likewise, when it comes to solving complex environmental problems, properly educated decision makers are sorely needed. **The authors of this book feel that a good scientific education coupled with many years of solving tough environmental problems will produce leaders who will be capable of making optimal tactical and strategic decisions.** This environmental decision making may be carried out within a framework similar to that described in Section 1.5. One should always recall that good science can create imaginative solutions to environmental problems which in turn can influence the preferences of the decision makers who must eventually make the strategic decisions. Finally, by employing the **McLeod-Hipel Time Series Package** described in Section 1.7, a decision maker can use this flexible **decision support system** to immediately take full advantage of the rich array of environmetrics methods presented in the book.

Fertile fields of ideas in environmetrics await the curious reader in the upcoming chapters of this book. Depending upon the background and interests of the individual traveller, a **variety of touring routes** are suggested in Section 1.6.2. A reader who would like a review of some basic statistical concepts used in time series modelling may wish to start his or her journey by moving on to Chapter 2. Bon voyage!

PROBLEMS

1.1 What roles do you think science and statistics should play in developing sound environmental policies for restoring and preserving the natural environment?

1.2 Some serious types of environmental problems are referred to in Section 1.2.1. Discuss an important environmental problem that is of great concern to you. Outline realistic steps that could be taken to overcome this problem. Kindly provide the references from which you obtained your background information.

1.3 Many governments are deeply concerned with having human rights for each individual citizen. How do you think human rights and protection of the natural environment should be related? Discuss the situation in your own country.

1.4 The scientific method is explained in Section 1.2.2. Make a list of some of the key historical breakthroughs in the development of the scientific method. Wherever possible, point out when statistics played a key role in the improvement of the scientific method. Kindly provide a list of the references from which you obtained your material.

1.5 Tukey (1977) suggests a wide variety of graphical techniques for use as exploratory data analysis tools. Make a list of some of the main graphical procedures put forward by Tukey and briefly describe their purposes.

1.6 Explain in your own words why you think human beings like to develop models of natural and social systems. What is your opinion of the modelling procedure outlined in Section 1.3?

1.7 Chatfield (1988) describes a general approach for addressing real-life statistical problems. Summarize the main aspects of his approach and comment upon the advantages as well as the drawbacks of his methodology.

1.8 In the hydrological cycle displayed in Figure 1.4.1, the throughput to the system is water. Explain how energy could be used as the throughput in the hydrological cycle. Which throughput do you feel is more informative and easier to understand?

1.9 The Garrison Diversion Unit conflict of Sections 1.5.1 and 1.5.3 constitutes an example of an international environmental dispute for which good scientific and economic studies were carried out. Describe another planned large-scale project which is causing controversies to arise because of possible detrimental environmental effects. Who are the key decision makers involved in this environmental conflict and what are the main courses of action that each decision maker can follow. Discuss any scientific and/or economic studies that have been completed as well as how the results of the studies may influence any of the decision makers and the possible resolutions to the dispute. How are statistical methods used in any of these studies? If proper scientific or economic studies are not currently being executed, suggest how they could be done. Be sure to reference newspapers, magazines, journals and books from which you obtain your background information.

1.10 In Section 1.6.3, a number of water resources journals are mentioned. Select any two of the water resources journals and go to your library to obtain the last two years of publications. For each journal, make a list of categories under which the individual papers could be classified. What percentage of papers in each journal deal with environmetrics? Discuss the main types of environmetrics papers that are published.

1.11 A number of books on stochastic hydrology are mentioned in Section 1.6.3. Select one of these books and make a list of the major types of time series models discussed in the book. Compare this list of models to the one given in Table 1.6.2.

1.12 A decision support system for time series modelling and analysis is described in Section 1.7. By referring to appropriate literature, explain the basic design, function and user-friendly features of a decision support system in general. You may wish to read articles published in journals such as Decision Support Systems and Information and Decision Technologies, appropriate textbooks like the one by Sage (1991) and scientific encyclopediae such as the one edited by Sage (1990).

1.13 The McLeod-Hipel Time Series Package mentioned in Section 1.7 contains a wide variety of representative time series. Pick out a hydrological series that is of interest to you and use the package to produce a graph over time of the series. Write down some of the statistical characteristics of the time series that you can visually detect in the graph. If you feel adventurous, use the package to fit an appropriate time series model to the data and to produce forecasts 10 steps into the future. List some of the features that you like about the McLeod-Hipel package.

1.14 Why are education, in general, and science and engineering, in particular, so highly respected in Japan? What influence has this reverence for education had upon the economy of Japan? Have the Japanese adopted sound environmental policies within their own country and how are these policies connected to economic policies?

1.15 After you finish reading Chapter 1, write down your main reasons or objectives for studying environmetrics. After you complete taking an envionmetrics course using this book or reading the entire book on your own, refer to this list to see if your goals have been met.

REFERENCES

CONFLICT ANALYSIS

Fang, L., Hipel, K. W. and Kilgour, D. M. (1988). The graph model approach to environmental conflict resolution. *Journal of Environmental Management*, 27:195-212.

Fang, L., Hipel, K. W. and Kilgour, D. M. (1993). *Interactive Decision Making: The Graph Model for Conflict Resolution*. Wiley, New York.

Fraser, N. M. and Hipel, K. W. (1984). *Conflict Analysis: Models and Resolutions*. North-Holland, New York.

Hipel, K. W. (1990). Decision technologies for conflict analysis. *Information and Decision Technologies*, 16(3):185-214.

Hipel, K. W. and Fraser, N. M. (1980). Metagame analysis of the Garrison conflict. *Water Resources Research*, 16(4):629-637.

Howard, N. (1971). *Paradoxes of Rationality, Theory of Metagames and Political Behavior*. M.I.T. Press, Cambridge, Massachusetts.

von Neumann, J. and Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey, third edition.

DATA COLLECTION

Gilbert, R. O. (1987). *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Harmancioglu, N. B. and Alpaslan, N. (1992). Water quality monitoring network design: A problem of multi-objective decision making. *Water Resources Bulletin* 28(1):179-192.

Lerner, D. (Editor) (1986). *Monitoring to Detect Changes in Water Quality Series*. Proceedings of a symposium held during the 2nd Scientific Assembly of the International Association of Hydrological Sciences (IAHS) held in Budapest, Hungary, in July, 1986, IAHS Publication No. 157.

Lettenmaier, D. P., Hipel, K. W. and McLeod, A. I. (1978). Assessment of environmental impacts, part two: Data collection. *Environmental Management*, 2(6):537-554.

Loftis, J. C., McBride, G. B. and Ellis, J. C. (1991). Considerations of scale in water quality monitoring and data analysis. *Water Resources Bulletin* 27(2):255-264.

Moss, M. E. (1979). Some basic considerations in the design of hydrologic data networks. *Water Resources Research*, 15(6):1673-1676.

Ward, R. C. and Loftis, J. C. (1986). Establishing statistical design criteria for water quality monitoring systems: Review and synthesis. *Water Resources Bulletin* 22(5):759-767.

Ward, R. C., Loftis, J. C. and McBride, G. B. (1986). The data-rich but information-poor syndrome in water quality monitoring. *Environmental Management* 10:291-297.

Ward, R. C., Loftis, J. C. and McBride, G. B. (Editors) (1989). *Design of Water Quality Information Systems*. Proceedings of an International Symposium held at Fort Collins, Colorado from June 7 to 9, 1989, sponsored by Colorado State University and the U. S. Environmental Protection Agency, Information Series No. 61, Colorado Water Resources Research Institute.

Whitfield, P. H. (1988). Goals and data collection designs for water quality monitoring. *Water Resources Bulletin* 24(4):775-780.

ECONOMIC FORECASTING

Abraham, B. and Ledolter, J. (1983). *Statistical Methods for Forecasting*. John Wiley, New York.

Firth, M. (1977). *Forecasting Methods in Business and Management*. Edward Arnold Publishers, London.

Granger, C. W. J. (1980). *Forecasting in Business and Economics*. Academic Press, New York.

Granger, C. W. J. and Newbold, P. (1977). *Forecasting Economic Time Series*. Academic Press, New York.

Makridakis, S. and Wheelwright, S. (1978). *Interactive Forecasting: Univariate and Multivariate Methods*. Holden-Day, San Francisco.

Montgomery, D. C. and Johnson, L. A. (1976). *Forecasting and Time Series Analysis*. McGraw-Hill, New York.

Nelson, C. R. (1973). *Applied Time Series Analysis for Managerial Forecasting*. Holden-Day, San Francisco.

Pankratz, A. (1983). *Forecasting with Univariate Box-Jenkins Models*. John Wiley, New York.

HYDROLOGY AND ENVIRONMENTAL SYSTEMS

Bennett, R. J. and Chorley, R. J. (1978). *Environmental Systems: Philosophy, Analysis and Control*. Princeton University Press, New Jersey.

Eagleson, P. S. (1970). *Dynamic Hydrology*. McGraw-Hill, New York.

Eagleson, P. S. (1991). Hydrologic science: a distinct geoscience. *Reviews of Geophysics*, 29(2):237-248.

Falkenmark, M. (1990). Environmental management - what is the role of the hydrologist? *Paper presented at the 1990 Commemorative Symposium for 25 Years of IHD (International Hydrological Decade) and IHP (International Hydrological Program)*. UNESCO (United Nations Education, Scientific and Cultural Organization), held March 15-17, 1990, Paris, France.

Linsley, Jr., R. K., Kohler, M. A. and Paulhus, J. L. H. (1982). *Hydrology for Engineers*. McGraw-Hill, New York.

McCuen, R. H. (1989). *Hydrologic Analysis and Design*. Prentice-Hall, Englewood Cliffs, New Jersey.

National Research Council (1991). *Opportunities in the Hydrologic Sciences*. National Academy Press, Washington, D.C.

White, I. D., Mottershead, D. N. and Harrison, S. J. (1984). *Environmental Systems*. George Allen & Unwin, London.

OPERATIONAL RESEARCH AND SYSTEMS SCIENCES

Blackett, P. M. S. (1962). *Studies of War*. Oliver and Boyd, Edinburgh.

Checkland, P. (1981). *Systems Thinking, Systems Practice*. John Wiley, Chichester, United Kingdom.

Goicoechea, A., Hansen, D. R. and Duckstein, L. (1982). *Multiobjective Decision Analysis with Engineering and Business Applications*. Wiley, New York.

Hillier, F. S. and Lieberman, G. J. (1990). *Introduction to Operations Research*. McGraw-Hill, New York, fifth edition.

Hipel, K. W. (Editor) (1992). *Multiple Objective Decision Making in Water Resources*. AWRA Monograph #18 published by the American Water Resources Association, 5410 Grosvenor Lane, Suite 220, Bethesda, Maryland 20814-2192, U.S.A.

Loucks, D. P., Stedinger, J. R. and Haith, D. A. (1981). *Water Resources Systems Planning and Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey.

Maas, A., Hufschmidt, M. M., Dorfman, R., Thomas Jr., H. A., Marglin, S. A. and Fair, M. G. (1962). *Design of Water-Resource Systems*. Harvard University Press, Cambridge, Massachusetts.

Miser, H. J. and Quade, E. S. (Editors) (1985). *Handbook of Systems Analysis: Overview of Uses, Procedures, Applications, and Practice*. North-Holland, New York.

Miser, H. J. and Quade, E. S. (Editors) (1988). *Handbook of Systems Analysis: Craft Issues and Procedural Choices*. North-Holland, New York.

Radford, K. J. (1989). *Individual and Small Group Decisions*. Springer-Verlag, New York.

Reich, R. B. (1983). *The Next American Frontier*. Time Books, New York.

Rosenhead, J. (Editor) (1989). *Analysis for a Problematic World*. Wiley, Chichester, United Kingdom.

Roy, B. (1985). *Méthodologie Multicritère d'Aide a la Décision*. Economica, Paris.

Sage, A. P. (Editor) (1990). *Concise Encyclopedia of Information Processing in Systems and Organizations*. Pergamon, Oxford.

Sage, A. P. (1991). *Decision Support Systems Engineering*. Wiley, New York.

Vincke, P. (1992). *Multicriteria Decision-aid*. Wiley, Chichester, United Kingdom.

Waddington, C. H. (1973). *OR in World War 2*. Elek Science, London.

Wagner, H. M. (1975). *Principles of Operations Research*. Prentice-Hall, Englewood Cliffs, New Jersey, second edition.

SCIENCE

Box, J. (1978). *R.A. Fisher: The Life of a Scientist*. Wiley, New York.

Dransfield, J., Flenley, J. R., King, S. M., Harkness, D. D. and Rapu, S. (1984). A recently extinct palm from Easter Island. *Nature*, 312(5996):750-752.

Flenley, J. R. and King, S. M. (1984). Late quaternary pollen records from Easter Island. *Nature*, 307(5946):47-50.

Groen, J., Smit, E. and Eijsvoogel, J. (Editors) (1990). *The Discipline of Curiosity*. Elsevier, Amsterdam.

Levine, J. S. (1990). Global biomass burning: atmospheric, climatic and biospheric implications. *EOS, Transactions of the American Geophysical Union*, 71(37):1075-1077.

Russell, B. (1946). *History of Western Philosophy*. Allen and Unwin, London.

STATISTICAL WATER QUALITY MODELLING

Chapman, D. T. and El-Shaarawi, A. H. (Editors) (1989). Statistical methods for the assessment of point source pollution. In *Proceedings of a Workshop on Statistical Methods for the Assessment of Point Source Pollution*, held Sept. 12-14, 1988, at the Canada Centre for Inland Waters, Burlington, Ontario, Canada, Kluwer, Dordrecht, the Netherlands.

El-Shaarawi, A. H. and Esterby, S. R. (Editors) (1982). *Time Series Methods in Hydrosciences*. North-Holland, Amsterdam, The Netherlands.

El-Shaarawi, A. H. and Kwiatowski, R. E. (Editors) (1986). *Statistical Aspects of Water Quality Monitoring*. Elsevier, Amsterdam.

Hipel, K. W. (Editor) (1988). *Nonparametric Approaches to Environmental Impact Assessment*. American Water Resources Association, Bethesda, Maryland.

STATISTICS

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19:716-723.

Box, G. E. P. (1974). Statistics and the environment. *J. Wash. Acad. Sci.*, 64(2):52-59.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, 71(356):791-799.

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26:211-252.

Box, G. E. P., Hunter, W. G. and Hunter, J. S. (1978). *Statistics for Experimenters*. Wiley, New York.

Chatfield, C. (1988). *Problem Solving: A Statisticians Guide*. Chapman and Hall, London.

Helsel, D. R. and Hirsch, R. M. (1992). *Statistical Methods in Water Resources*. Elsevier, Amsterdam.

Hunter, J. S. (1990). Keynote address at the First International Conference on Statistical Methods for the Environmental Sciences, held April 4-7, Cairo, Egypt, 1989. *Environmetrics*, 1(1):1-6.

Kendall, M. G. (1975). *Rank Correlation Methods*. Charles Griffin, London, fourth edition.

McPherson, G. (1990). *Statistics in Scientific Investigations*. Springer-Verlag, New York.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, Massachusetts.

STOCHASTIC HYDROLOGY

Bras, R. L. and Rodriguez-Iturbe, I. (1985). *Random Functions and Hydrology*. Addison-Wesley, Reading, Massachusetts.

Clarke, R. T. (1973). *Mathematical Models in Hydrology*. Irrigation and Drainage Paper, Food and Agriculture Organization of the United Nations, Rome, 1973.

Fiering, M. B. (1967). *Streamflow Synthesis*. Harvard University Press, Cambridge, Massachusetts.

Hipel, K. W. (Editor) (1985). *Time Series Analysis in Water Resources*. American Water Resources Association, Bethesda, Maryland.

Kashyap, R. L. and Rao, A. R. (1976). *Dynamic Stochastic Models from Empirical Data*. Academic Press, New York.

Kottegoda, N. T. (1980). *Stochastic Water Resources Technology*. MacMillan Press, London.

McBean, E. A., Hipel, K. W. and Unny, T. E. (Editors) (1979a). *Inputs for Risk Analysis in Water Systems*. Water Resources Publications, Fort Collins, Colorado.

McBean, E. A., Hipel, K. W. and Unny, T. E. (Editors) (1979b). *Reliability in Water Resources Management*. Water Resources Publications, Fort Collins, Colorado.

McLeod, A. I. (Editor) (1987). Advances in statistical sciences. In *Festschrift in Honor of Professor V.M. Joshi's 70th Birthday, Volume IV, Stochastic Hydrology* (editors across all volumes are I.B. MacNeill and G.J. Umphrey), Reidel, Dordrecht, the Netherlands.

McCuen, R. H. and Snyder, W. M. (1986). *Hydrologic Modeling: Statistical Methods and Applications*. Prentice-Hall, Englewood Cliffs, New Jersey.

Pereira, M. V. F., Oliveira, G. C., Costa, C. C. G. and Kelman, J. (1984). Stochastic streamflow models for hydroelectric systems. *Water Resources Research*, 20(3):379-390.

Pereira, M. V. F. and Pinto, L. M. V. G. (1985). Stochastic optimization of a multireservoir hydroelectric system: A decomposition approach. *Water Resources Research*, 21(6):779-792.

Salas, J. D., Delleur, J. W., Yevjevich, V. and Lane, W. L. (1980). *Applied Modelling of Hydrologic Series*. Water Resources Publications, Littleton, Colorado.

Salas, J. D. and Smith, R. A. (1981). Physical basis of stochastic models of annual flows. *Water Resources Research*, 17(2):428-430.

Shen, H. W. (1976). *Stochastic Approaches to Water Resources, Volumes 1 and 2*. Colorado State University, Fort Collins, Colorado.

Shen, H. W., Obeysekera, J. T. B., Yevjevich, V. and Decoursey, D. G. (1986). Multivariate analysis of hydrological processes. In *Proceedings of the Fourth International Hydrology Symposium* held at Colorado State University, Fort Collins, Colorado, July 15-17, 1985. Engineering Research Center, Colorado State University.

Silva, L. F. C. A., Sales, P. R. H., Araripe Neto, T. A., Terry, L. A. and Pereira, M. F. V. (1984). *Coordinating the energy generation of the Brazilian system*. Technical report, Departamento de Operacao Energetica, Eletrobras, Rio de Janeiro, Brazil.

Yevjevich, V. (1972). *Stochastic Processes in Hydrology*. Water Resources Publications, Littleton, Colorado.

TIME SERIES ANALYSIS AND STOCHASTIC PROCESSES

Anderson, O. D. (Editor) (1979). *Proceedings of the International Time Series Meeting, Nottingham University, March 1979*. North-Holland, Amsterdam.

Anderson, O. D., Ord, J. K. and Robinson, E. A. (1985). Time series analysis: Theory and practice 6. In *Proceedings of the International Conference held at Toronto, Canada, August 10-14, 1983*. North-Holland, Amsterdam.

Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. John Wiley, New York.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, Oakland, California, revised edition.

Brillinger, D. R. (1975). *Time Series Data Analysis and Theory*. Holt, Rinehart and Winston, New York.

Brockwell, P. J. and Davis, R. A. (1987). *Time Series: Theory and Methods*. Springer-Verlag, New York.

Chatfield, C. (1975). *The Analysis of Time Series: Theory and Practice*. Chapman and Hall, London.

Cox, D. R. and Miller, H. D. (1965). *The Theory of Stochastic Processes*. Chapman and Hall, London.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*. John Wiley, New York.

Hannan, E. J. (1970). *Multiple Time Series*. John Wiley, New York.

Haykin, S. (1990). *Modern Filters*. MacMillan, New York.

Jenkins, G. M. (1979). *Practical Experiences with Modelling and Forecasting Time Series*. Gwilym Jenkins and Partners Ltd., Parkfield, Greaves Road, Lancaster, England.

Jenkins, G. M. and Watts, D. G. (1968). *Spectral Analysis and its Applications*. Holden-Day, San Francisco.

Kendall, M. G. (1973). *Time Series*. Hafner Press, New York.

Kloeden, P. E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey.

McLeod, A. I. and Hipel, K. W. (1992). The McLeod-Hipel Time Series Package. Copyright of this decision support system owned by A.I. McLeod and K.W. Hipel. Distributed by McLeod-Hipel Research, 121 Longwood Drive, Waterloo, Ontario, Canada, N2L 4B6. Tel: (519) 884-2089.

McLeod, G. (1983). *Box-Jenkins in Practice, Volume 1, Univariate Stochastic and Transfer Function/Intervention Analysis*. Gwilym Jenkins and Partners Ltd., Parkfield, Greaves Road, Lancaster, England.

Pandit, S. M. and Wu, S. M. (1983). *Time Series and System Analysis with Applications*. John Wiley, New York.

Priestley, M. B. (1981). *Spectral Analysis and Time Series, Volume 1: Univariate Series, Volume 2: Multivariate Series, Prediction and Control*. Academic Press, London.

Vandaele, W. (1983). *Applied Time Series and Box-Jenkins Models*. Academic Press, New York.

Young, P. C. (1984). *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, Berlin.