

**ANALYSIS, MODELLING AND OPTIMAL CONTROL OF
WATER SUPPLY AND DISTRIBUTION SYSTEMS**

YUE-CHUN CHEN

**A thesis submitted in partial fulfilment of the
requirements of the Council for National Academic Awards
for the degree of Doctor of Philosophy**

December 1989

Leicester Polytechnic

**ANALYSIS, MODELLING AND OPTIMAL CONTROL
OF WATER SUPPLY AND DISTRIBUTION SYSTEMS**

by

YUE-CHUN CHEN

ABSTRACT

This thesis is concerned with the developments of analysis, modelling and optimization techniques and computer program algorithms, with the ultimate aim of control of water supply and distribution systems to lead to overall optimal operation.

Typical system features and operational conditions are analyzed, and the requirements for the overall objective are examined, to determine an overall control strategy which is subsequently developed and tested on real systems throughout this thesis. As a prerequisite, short-term water demand forecasting is extensively studied by employing time series analysis. Special consideration is given to improving the forecasting accuracy of the method and its on-line implementation. In order to speed up the solution time of optimal system operation, simplified system models -- namely, piecewise macroscopic model and equivalent network model -- are developed respectively. Then by employing the piecewise macroscopic model, a nonlinear programming method is developed to cater for the optimal operation of a class of multi-source systems without significant storage. The optimal operation policy obtained by this method is realized at two levels: the first level calculates the optimized apportioning of water to be delivered by different sources; the second level decides the least cost pump schedules to supply the optimized apportioning of water. Based on the equivalent network model, a linear programming method is developed for optimization of a class of multi-source, multi-reservoir systems with a mixture of fixed speed pumps and variable speed and/or variable throttle pumps. This method yields directly optimized pump schedules and reservoir trajectories in terms of least cost system operation.

The integration of the developments results in a scheme which can be applied to give overall dynamic control of a wide range of water supply and distribution systems. The application results presented in this thesis justify the theoretical developments and show that benefits can be obtained from these developments.

ACKNOWLEDGEMENTS

This thesis is a presentation of the research done by the author during the period from October 1985 to December 1989. I am indebted to my previous supervisor, Professor Qin Yang, since part of the work was done under his guidance at the Department of Environmental Engineering, Tongji University, China, during the first two years of the period. I wish to thank my present supervisors, Professor Bryan Coulbeck and Dr. Chun-Hou Orr, since the major work, as an academically collaborative project, was done under their guidance and support at the Water Control Centre, Leicester Polytechnic, England, during the final two and a quarter years of this period. In addition, I am grateful to them for their continual effort in seeking financial support in the form of a scholarship for me. I would also like to thank my technical adviser, Dr. Mietek Brdys', for his helpful suggestions.

I would also like to acknowledge with thanks the scholarship provided to me jointly by Leicester Polytechnic and the Water Control Centre during my two and a quarter years in the U.K.

Members of the Water Control Centre have provided much technical and linguistic support. Special thanks are due to Dr. Masud Parkar, Mr. Steve Tennant, Mr. John Rance and Mrs Andrea Mitchell for providing the above support.

Finally, I would like to express my deepest appreciation to my parents and my fiancée for their continual encouragement and patience.

CONTENTS	PAGE
ACKNOWLEDGEMENTS	i
CONTENTS	iii
LIST OF FIGURES	vii
LIST OF TABLES	ix
CHAPTER 1 INTRODUCTION	1
1.1 SIGNIFICANCE OF THE RESEARCH TOPIC	1
1.2 OVERVIEW OF THE THESIS	2
1.3 MAJOR CONTRIBUTIONS OF THE PRESENT RESEARCH	8
CHAPTER 2 GENERAL DESCRIPTION AND ANALYSIS OF WATER SUPPLY AND DISTRIBUTION SYSTEMS	10
2.1 INTRODUCTION	10
2.2 PUMPING STATION	11
2.2.1 Head Increase vs Flow Relationship	13
2.2.2 Flow vs Efficiency and Power Relationships	20
2.3 PIPES	22
2.4 PRESSURE REDUCING VALVES (PRV)	24
2.5 CHECK VALVES (OR NON-RETURN VALVES)	25
2.6 RESERVOIRS, TANKS, TOWERS	26
2.7 HYDRAULIC ANALYSIS OF WATER SUPPLY AND DISTRIBUTION SYSTEMS	29
2.7.1 Static Simulation	31
2.7.2 Dynamic Simulation	32
2.8 OPTIMAL OPERATION OF WATER SUPPLY AND DISTRIBUTION SYSTEMS	33
2.9 CONCLUSIONS	34

	PAGE
CHAPTER 3 ANALYSIS AND FORECASTING OF WATER DEMAND	36
3.1 INTRODUCTION	36
3.2 CHARACTERISTICS OF WATER DEMAND	37
3.3 STUDIES ON A CLASS OF WATER DEMAND FORECASTING	37
3.3.1 Estimation of Parameters φ_i	39
3.3.2 Selection of Order p	42
3.3.3 Forecasting	43
3.3.4 Application	43
3.4 BOX-JENKINS APPROACH TO GENERAL ARIMA MODELS	48
3.4.1 Identification	51
3.4.2 Estimation	55
3.4.3 Forecasting	63
3.4.4 Applications to Water Demand Forecasting	67
3.5 COMPARISON WITH THE PREDICTIONS FROM THE EXISTING DEMAND PREDICTION PROGRAM GIDAP	82
3.5.1 Introduction to GIDAP	82
3.5.2 Discussions and Comparisons	88
3.6 CONCLUSIONS	95
CHAPTER 4 NETWORK MODELLING AND SIMPLIFICATIONS FOR SYSTEM OPTIMAL OPERATIONS	98
4.1 INTRODUCTION	98
4.2 SUMMARY OF THE MACROSCOPIC MODEL	99
4.2.1 Pumping Station Relationships	99
4.2.2 Internal Network Nodal Pressure Relationship	100

	PAGE	
4.3	PIECEWISE MACROSCOPIC MODEL	102
4.3.1	Pumping Station Relationship	102
4.3.2	Internal Network Nodal Pressure Relationship	103
4.3.3	Model Validation	103
4.3.4	Division of a Day into Proportional Loading Periods	107
4.4	EQUIVALENT NETWORK MODEL	109
4.4.1	Theoretical Analysis and Deductions	109
4.4.2	Model Validations	121
4.5	CONCLUSIONS	135
CHAPTER 5	OPTIMAL SYSTEM OPERATIONS	139
5.1	INTRODUCTION	139
5.2	ALGORITHM FOR OPTIMAL OPERATIONS OF MULTI-SOURCE SYSTEMS CONTAINING INSIGNIFICANT WATER STORAGE	144
5.2.1	Formulation of Optimization Problem	144
5.2.2	Summary of the Optimization Technique	149
5.2.3	Application Results	152
5.3	ALGORITHM FOR OPTIMAL OPERATIONS OF A CLASS OF MULTI-SOURCE, MULTI-RESERVOIR SYSTEMS	156
5.3.1	Formulation of Optimization Problem	158
5.3.2	Summary of the Revised Simplex Method	168
5.3.3	Application Results	171
5.3.4	Studies on the Decomposition of Variable Speed Pumps	192

	PAGE
5.3.5 Comparison with the Modified GIPOS Application	206
5.4 CONCLUSIONS	223
CHAPTER 6 OVERALL COMMENTS AND CONCLUSIONS	227
6.1 OVERALL COMMENTS	227
6.2 GENERAL CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK	231
REFERENCES	236
APPENDIX PUBLISHED PAPERS	251

LIST OF FIGURES

- 2.1 A MULTI-SOURCE AND MULTI-RESERVOIR SYSTEM
- 2.2 COEFFICIENT FITTING PUMP DATA
- 2.3 COMPARISON OF MODEL AND FIELD DATA
- 2.4 HEAD VS FLOW RELATIONSHIP FOR TWO DIFFERENT PUMPS
- 2.5 OPTIMAL OPERATION OF A WATER SUPPLY AND DISTRIBUTION SYSTEM
- 3.1 MODELLING RESULTS USING THE SERIES FORECASTING METHOD
- 3.2 MODELLING RESULTS USING THE PARALLEL FORECASTING METHOD
- 3.3 MODELLING RESULTS OF THE COMBINED PREDICTION METHOD
- 3.4 FUNCTIONAL DIAGRAM OF THE BOX-JENKINS APPROACH TO
GENERAL ARIMA TIME SERIES MODELS
- 3.5 FORECASTING RESULTS FOR 3 JUNE (FRI)
- 3.6 FORECASTING RESULTS FOR 17 JUNE (FRI)
- 3.7 A COMPARISON BETWEEN TWO MODELS
- 3.8 THE PREDICTION PROCESS OF GIDAP
- 3.9 A COMPARISON OF FORECASTS FOR 3 JUNE 1988
- 3.10 A COMPARISON OF FORECASTS FOR 10 JUNE 1988
- 4.1 LAYOUT OF THE PUDONG SYSTEM, IN SHANGHAI
- 4.2 AN EXAMPLE WATER SUPPLY NETWORK
- 4.3 EQUIVALENT NETWORK OF FIG 4.2
- 4.4 CONFIGURATION OF THE GOLDTHORN ZONE
- 4.5 EQUIVALENT NETWORK OF FIG 4.4 (TO BE FURTHER SIMPLIFIED)
- 4.6 EQUIVALENT NETWORK OF FIG 4.4 (0.0 HR TO 8.0 HRS)
- 4.7 EQUIVALENT NETWORK OF FIG 4.4 (9.0 HRS TO 23.0 HRS)

- 5.1 STRUCTURE FOR OPTIMAL OPERATIONS
- 5.2A CONFIGURATION OF THE NURTON/BUSHBURY ZONE
- 5.2 TARIFF DATA (UNIT CHARGE)
- 5.3 EQUIVALENT NETWORK OF NURTON/BUSHBURY ZONE
- 5.4 OPTIMIZED PUMP SCHEDULES
- 5.5 RESERVOIR TRAJECTORIES
- 5.6 OPTIMIZED PUMP SCHEDULES
- 5.7 RESERVOIR TRAJECTORIES
- 5.8 BUSHBURY SUB-SYSTEM
- 5.9 NURTON SUB-SYSTEM
- 5.10 COMPARISON OF PUMP SCHEDULES
- 5.11 A COMPARISON OF RESERVOIR TRAJECTORIES
- 5.12 A COMPARISON OF PUMP SCHEDULES
- 5.13 A COMPARISON OF RESERVOIR TRAJECTORIES
- 6.1 AN INTEGRATED CONTROL SCHEME FOR A WATER SUPPLY
AND DISTRIBUTION SYSTEM

LIST OF TABLES

- 3.1 MODELLING RESULTS FOR THE SERIES METHOD
- 3.2 MODELLING RESULTS FOR THE PARALLEL METHOD
- 3.4 AUTOCORRELATIONS AND PARTIAL AUTOCORRELATIONS
- 3.5 COMPUTED RESULTS FOR THE MODEL ARIMA (0,1,0)(1,1,0)₂₄
- 3.6 FORECASTING RESULTS FOR 3 JUNE 1988 (FRI)
- 3.7 FORECASTING RESULTS FOR 17 JUNE 1988 (FRI)
- 3.8 COMPUTED RESULTS FOR THE MODEL ARIMA(0,1,0)(2,1,0)₂₄
- 3.9 A COMPARISON OF FORECASTING RESULTS BETWEEN MODEL(3.43)
AND MODEL(3.45)
- 3.10 A COMPARISON OF FORECASTS FROM A TIME SERIES MODEL AND
AN EXPONENTIAL SMOOTHING MODEL FOR 3 JUNE 1988
- 3.11 A COMPARISON OF FORECASTS FROM A TIME SERIES MODEL AND
AN EXPONENTIAL SMOOTHINGMODEL FOR 10 JUNE 1988
- 4.1 SIMULATION RESULTS OF PIECEWISE MACROSCOPIC MODEL
- 4.2 RESULTS OF CORRELATION ANALYSIS BETWEEN NODES
- 4.3 RESULTS FOR EQUIVALENT NETWORK MODELLING OBTAINED FROM
PROGRAM EQUINET
- 4.4 RESULTS FOR EQUIVALENT NETWORK MODELLING FROM
PROGRAM ENCQP (0.00 HR TO 8.00 HRS)
- 4.5 RESULTS FOR EQUIVALENT NETWORK MODELLING FROM
PROGRAM ENCQP (8.30 HRS TO 23.00 HRS)
- 4.6 VALIDATION OF EQUIVALENT NETWORK MODELLING
- 4.7 VALIDATION OF EQUIVALENT NETWORK MODELLING FOR DIFFERENT
OPERATIONAL CONDITIONS

- 4.8 VALIDATION OF EQUIVALENT NETWORK MODELLING FOR DIFFERENT OPERATIONAL CONDITIONS (THE INFLUENCE OF PUMP COMBINATIONS IS IGNORED)
- 5.1 OPTIMIZATION RESULTS OF PUDONG WATER SUPPLY SYSTEM
- 5.2 DATA FOR PUMPS IN FIGURE 5.2A
- 5.3 DATA FOR RESERVOIR IN FIGURE 5.2A
- 5.4 SIMULATION RESULTS FOR NURTON/BUSHBURY ZONE (ALL PUMPS ON)
- 5.5 SIMULATION RESULTS FOR NURTON/BUSHBURY ZONE (PUMP AT STABLEFORD OFF)
- 5.6 SIMULATION RESULTS FOR NURTON/BUSHBURY ZONE (PUMP AT COPLEY OFF)
- 5.7 DATA FOR THE EQUIVALENT NETWORK OF NURTON/BUSHBURY ZONE
- 5.8 PUMP AVERAGE VALUES OF FLOW AND POWER
- 5.9 RESULTS FOR THE OPTIMAL OPERATIONS OF NURTON/BUSHBURY ZONE
- 5.9A ADDITIONAL PUMP COMBINATIONS AT TETTENHALL
- 5.9B AVERAGE VALUES OF FLOW AND POWER OF ADDITIONAL PUMP COMBINATIONS (ADDITIONAL SPEED STEP 1300 RPM)
- 5.10 SENSITIVITY ANALYSIS FOR ADDITIONAL SPEED OF 1300 RPM OF THE VSP'S
- 5.11 RESULTS FOR THE OPTIMAL OPERATIONS OF NURTON/BUSHBURY ZONE
- 5.12 AVERAGE VALUES OF FLOW AND POWER OF ADDITIONAL PUMP COMBINATIONS (ADDITIONAL SPEED STEP 1525 RPM)
- 5.12A SENSITIVITY ANALYSIS FOR ADDITIONAL SPEED OF 1525 RPM OF THE VSP'S

- 5.13 AVERAGE VALUES OF PUMP FLOW AND POWER USED BY BOTH OF THE APPROACHES
- 5.14 RESULTS FROM OPPUS FOR THE OPTIMAL OPERATION OF NURTON/BUSHBURY ZONE
- 5.15 RESULTS FROM MODIFIED GIPOS FOR THE OPTIMAL OPERATION OF NURTON/BUSHBURY ZONE
- 5.16 RESULTS FROM OPPUS FOR NURTON/BUSHBURY ZONE (FINAL RESERVOIR LEVELS ARE KEPT THE SAME AS THOSE FROM THE MODIFIED GIPOS)

1.1 SIGNIFICANCE OF THE RESEARCH TOPIC

Water supply and distribution systems play an important part in modern cities by providing the vast quantities of purified water required for domestic, industrial and commercial consumers. With the growth of concentration of population, many water supply and distribution systems become so large and complicated that their efficient and proper operation can no longer be achieved manually and must thus rely on some degree of automatic monitoring and control. The rising cost of labour also increases the importance of releasing manpower from routine operating tasks to participate in more productive and creative activities. Further, and more importantly, there has been worldwide concern for the reduction of energy consumption while the cost of electricity for pumping water is huge: the bill for the UK is more than £100 million per annum (CREASEY, 1988). Therefore, it is considered inevitable that operation of water supply systems will eventually yield to fully automated control in order to achieve efficient operation of these systems of ever increasing complexities and costs.

From the viewpoint of modern control theory, a water supply distribution system is a highly non-linear, stochastic and large scale dynamic system, which is one of the most difficult problems in theory. Over the last fifteen years, many researchers from different disciplines in a number of countries have been devoting themselves to studying this problem. Various methods have been employed and great progress has been made. However, owing to the complexities of the practical systems, the limitations of related theories, and other factors, the stage has not yet been reached where the operation of water supply and distribution systems is possible under fully automated and optimal control; nor is there available a universal approach which can be satisfactorily applied to any water supply distribution system. The state-of-the-art of the research and applications seems to be that different approaches are effective and applicable to certain classes of water supply distribution systems, each of them having their own advantages and disadvantages. (BRDYS, 1988; CHEN, 1988a; COHEN, 1982; COULBECK, 1977; COULBECK, 1988; DEMOYER and HORWITZ, 1975a, 1976b; MOSS, 1979; PERRY, 1975; SABET and HELWEG, 1985).

This thesis presents the work by the author on the development of alternative methods of achieving the same objectives, and improvements or refinements to the existing methods for the overall control of water supply and distribution systems.

1.2 OVERVIEW OF THE THESIS

This section gives an overview of the thesis and the work performed. The thesis is arranged in such a way as to provide a coherent and

overall approach to the study of optimal operation of water supply and distribution systems.

CHAPTER 1 introduces the significance and the start-of-the-art of the research topic. This chapter also gives an overview of the thesis and a summary of the major contributions of the present research.

CHAPTER 2 describes and analyzes typical features of water supply distribution systems and their mathematical modelling and simulations, which forms a background for the succeeding studies.

CHAPTER 3 shows in detail how time series analysis is applied to short-term water demand forecasting. Firstly, AR (Autoregressive) models are applied to water demand forecasting, which have the advantage of simplicity in parameter estimation and convenience in forecasting formulation. The results of the application to the demand forecasting for Shanghai, China are quite satisfactory and show that AR models are at least suitable for a class of water demand forecasting. Secondly, in order to search for a method which can cater for more general water demand forecasting, the Box-Jenkins approach to general ARIMA (Autoregressive Integrated Moving Average) models is employed for further studies on demand forecasting. The satisfactory application results are shown in detail.

Theoretical and practical comparisons are also made and presented, between the performance of the time series models and an existing demand prediction program, GIDAP, developed at the Water Control Centre (TENNANT et al, 1986). The latter makes demand forecasts using

the exponential smoothing technique.

CHAPTER 4 discusses network modelling and simplification for optimal operation purposes. An ordinary water supply distribution system is composed of hundreds or even thousands of pipes and thus is a very complicated system. In conventional network simplification practice, networks are simplified by cancellation of pipes of small diameters, replacement of parallel or series pipes by "equivalents" and so on. This type of simplification is mainly applied to network hydraulic analysis and is very limited. The simplified network is still too complicated for optimal operation studies. Owing to this limitation and the reason of no better alternatives, some complicated networks are simplified by eliminating a large part of the network to leave only a few pipes. In the latter case, the characteristics of the original network might not be properly reflected and customer service requirements as well as some important components in the network, cannot be taken into account directly.

In this chapter, two independent network modelling and simplification methods are developed and investigated. The first of these methods extends the macroscopic model originated by R. DeMoyer, Jr. and his colleagues (DEMOYER and HORWITZ, 1975a, 1975b). This extension leads to the piecewise macroscopic model which can be successfully applied to those systems in which the loading patterns are not too far from proportional loading assumptions. For the validation of the piecewise macroscopic model, results, from both multiple regression and stepwise regression, respectively, are given, which confirm that this extension is worthwhile and practical.

The second method is equivalent network modelling, which has originated from the author's studies. This method is developed by introducing the concept of fictitious pipes. By employing matrix algebra, the detailed mathematical deduction leads to a least-squares estimation problem to minimize the discrepancies between the original detailed network model and the simplified equivalent network model. The application results show that the accuracy of the method is satisfactory and the method is very effective in the reduction of system complexity.

CHAPTER 5 examines the problem of optimal system operations. General formulation of the optimal operation of a system will lead to a nonlinear, large-scale dynamic optimization problem, which is one of the most difficult problems to solve. There is almost no method in the field of optimization that can cope with all the difficulties at the same time. Investigation into the various approaches by different researchers over the last fifteen years concludes that the stage has not yet been reached where a universal approach is available which can cope with all the difficulties and which is applicable to any water supply and distribution system.

In this chapter, by exploring the particular features of different classes of water supply and distribution systems, two algorithms are developed. The first algorithm caters for a class of multi-source systems without significant storage in the distribution part. Under such a premise, the large-scale nonlinear problem for the whole control period can be divided into many much smaller ones for each time interval which can be solved separately. These small problems

are easier to solve than the original single large-scale problem, and the computation time for all the small problems together is much shorter than that of the original single large-scale one. The formulation of the optimal operation problem for each time interval by this algorithm results in a constrained non-linear programming problem where the pumping costs together with the treatment costs comprise the objective function. The pressure requirements of some pressure monitoring nodes in the system as well as pumping capacities and other features are taken into account in the constraint set. Technically, the optimal operation problems are further simplified by employing the piecewise macroscopic model developed in Chapter 4. This model relates major variables in the system directly to avoid solving the simultaneous network equations iteratively, which is time consuming and is hardly feasible for on-line control purposes. The optimal operation produced by this algorithm is realized at two levels. The first level calculates the optimized apportioning of water to be delivered by different sources. The second level caters for the derivation of the optimized pump schedules to supply the optimized apportioning of water. The application results of this algorithm show that significant savings can be achieved from this optimal operation scheme. It is concluded that this algorithm is directly applicable to a general class of multiple source systems and can also be extended to solve the optimal operation problems of multi-source, multi-reservoir systems by incorporating piecewise macroscopic relationships for reservoirs into the constraint set. However, after the extension, much greater difficulties might be encountered in the non-linear programming computations, and global optima cannot be guaranteed.

The second algorithm is developed for a class of multi-source, multi-

reservoir systems in which pump flows, and especially pump consumed powers, are nearly constant during the whole control period. This condition can be satisfied in many systems with significant reservoir storage, since the significant storage can to a certain degree compensate for the fluctuations in water demand to keep the pump flows (and thus powers) from changing significantly. Based on the equivalent model developed in Chapter 4, and by rightly choosing time of pumping rather than, as in the usual approach, pump flow, as a decision variable, this algorithm eventually comes to a large-scale linear programming problem for which global optima can be usually guaranteed. This algorithm is also capable of dealing with the mixture of discrete variables and continuous variables by discretizing the continuous variables. The method of discretization is also studied systematically, based on a post-optimality analysis. The application results show that optimized pump schedules and reservoir trajectories can be obtained by this algorithm to lead to least cost system operations. It is concluded that this algorithm is also suitable to apply to those systems in which the pump flows are varying significantly, by taking the average values of the flows. The derived solution could provide a good indication or reference for the optimal operations of those systems.

Comparisons are made between the applications of this algorithm and the existing optimization program GIPOS (Graphical and Interactive Pump Optimization and Scheduling) developed at the Water Control Centre. It is shown that the former algorithm is more suitable for applying to that system, particularly for its future operation.

Computer programs are written in Fortran 77 by the author for both of the optimization algorithms for general purpose applications.

CHAPTER 6 reviews and draws conclusions on the developments, and limitations and indicates requirements for further work of the research in various aspects of the overall objectives. A proposal is given for an integrated scheme for the optimal dynamic control of water supply and distribution systems. This scheme will incorporate the mathematical models and program algorithms developed in this thesis. General conclusions are drawn and worthwhile further work is envisaged on the main aspects of the research.

1.3 MAJOR CONTRIBUTIONS OF THE PRESENT RESEARCH

The following major contributions all reflect original work by the author and are considered to represent significant advances over existing work. All of them are also established in the form of computer programs written in Fortran 77 and some of them have been published in periodicals, books and conference proceedings as listed in the References and given in the Appendix.

(i) Development of versatile time series water demand forecasting model (Chapter 3). This is an original application and covers the whole procedure from model building to forecasting, which is suitable

for immediate application to other water supply distribution systems.

(ii) Development of a piece-wise macroscopic modelling technique for network simplification (Chapter 4). This extends existing methods to allow for application to more general systems.

(iii) Development of an equivalent modelling technique for network simplification (Chapter 4). This is a new approach which, for the first time in network simplification studies and practice, introduces the concept of fictitious pipes resulting in a much simplified network which can reflect the main feature of the original network and is convenient for optimization manipulations.

(iv) Development of theory and computer program of nonlinear programming method using piecewise macroscopic model. This method caters for the optimal operation of a class of multi-source water supply and distribution systems (Chapter 5). This is an original application and has resulted in optimized apportioning of water, to be delivered, among different sources and least cost pump schedules to supply these optimized proportions. These are conducted at two levels.

(v) Development of theory and computer program of linear programming method using equivalent network model. This method is applicable to a class of multi-source, multi-reservoir systems with fixed speed pumps and/or variable speed pumps (Chapter 5). This is also an original application. The program evaluates directly optimized pump schedules and reservoir trajectories to lead to least cost system operations.

2.1 INTRODUCTION

Generally, an urban water supply and distribution system is composed of water sources, treatment works, pumping stations, control valves, reservoirs, and distribution networks. Water is abstracted from water sources (surface and/or underground water), and is purified at the treatment works. After purification, water is transported under the influence of pressure through conduits to service reservoirs and consumers. Pressurization is achieved either by utilizing the force of gravity along favourable topographical gradients, or by using pumps. Fig 2.1 shows the configuration of a multi-source, multi-reservoir water supply and distribution system.

This chapter describes in brief the functions of the important components of a water supply and distribution system together with appropriate mathematical models.

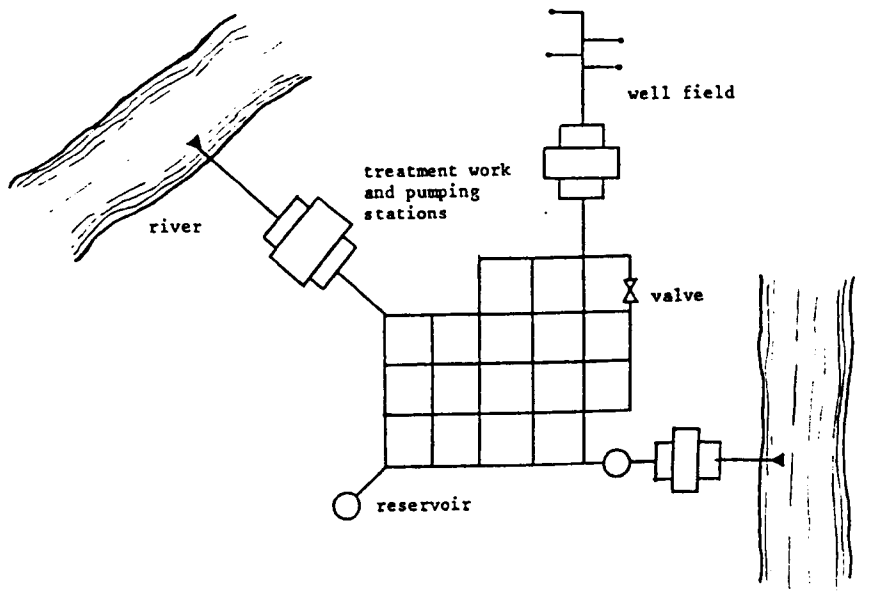


Fig 2.1 A Multi-Source and Multi-Reservoir System

2.2 PUMPING STATIONS

For water supply purposes, there are several types of pumps in use: 1) borehole pumps, which deliver water from wells or boreholes and are of submersible or vertical spindle type; 2) pumps after treatment works (usually with a contact tank) which are horizontally mounted of centrifugal type; and 3) in-line booster pumps, which serve to raise the pressure of water to deliver water from low to high pressure regions within the distribution system, these are also of centrifugal type. All these pumps may be either fixed or variable speed (COULBECK, 1977; FAIR et al, 1966; YANG, 1956).

Most pumps are driven by electric motors and consume large amounts of electrical energy. Pumping costs constitute the large part of operating costs for most water supply and distribution systems. From an operational point of view, they are the most important control components in the system operation, and thus are our main concern in optimal operations.

A general pump station may contain parallel groups of fixed speed and variable speed and/or variable throttle pumps, where each group can be composed of a number of pumps of similar characteristics. Individual pump group head/flow characteristics are typically non-linear, with combined pump flow being a function of the station head increase and the number of parallel pumps. For variable speed or variable throttle pumps, the flow is also a function of speed or throttle factors, respectively. Individual pump efficiency is a non-linear function of pump flow. Pump group power consumption is also a non-linear function of the station head increase, number of parallel pumps, pump flow and pump efficiency.

It is found that in general, pump characteristics can be adequately fitted by a family of quadratic functions (CHEN, 1985; ORR et al, 1988b; YANG, 1979).

2.2.1 Head Increase vs Flow Relationship

i) Fixed Speed Pumps (FSP):

single pump:

$$H = AQ^2 + BQ + C \quad (2.1)$$

where:

Q = pump flow. (e.g., l/s)

H = pump head increase. (e.g., m)

A, B, C = coefficients.

combined identical pumps in parallel (i.e., group equation):

$$H_f = A(Q_f/u)^2 + B(Q_f/u) + C \quad (2.2)$$

where:

$Q_f = uQ$ is the combined pump group flow. (e.g., l/s)

H_f = combined pump group head increase. (e.g., m)

u = number of identical pumps in operation.

ii) Variable Speed Pumps (VSP)

single pump:

$$H_s/s^2 = A(Q_s/s)^2 + B(Q_s/s) + C \quad (2.3)$$

where:

s = (actual speed / nominal speed) speed ratio; the nominal speed is the speed at which the hydraulic pump coefficients A, B, C are determined.

H_s = pump head increase for speed ratio s. (e.g., m)

Q_s = pump flow for speed ratio s . (e.g., 1/s)

A,B,C = coefficients.

combined identical pumps in parallel (group equation):

$$H_s/s^2 = A(Q_s/su)^2 + B(Q_s/su) + C \quad (2.4)$$

where:

s = as in eqn(2.3).

u = number of identical pumps in operation.

H_s = pump group head increase for speed ratio s and pump control u . (e.g., m)

Q_s = combined pump group flow. (e.g., 1/s)

iii) Variable Throttle Pumps (VTP)

single or combined identical throttle pumps:

$$H_v = A(Q_v/uv)^2 + B(Q_v/uv) + C \quad (2.5)$$

where:

v = throttle factor, $0.0 \leq v \leq 1.0$. Practically, $v=1.0$ represents a fully open valve and $v=0.0$ represents a fully closed valve.

u = number of identical pumps in operation.

Q_v = combined or single (when $u=1$) throttle pump group flow. (e.g. 1/s)

H_v = combined or single throttle pump group head increase. (e.g. m)

The A,B,C coefficients in eqn(2.1) to eqn(2.5), can be estimated using the least-squares method, with data from pump manufacturer's test-bed curves or from field tests. In practice, it is found that field test data are more reliable and accurate for coefficient estimation.

Fig 2.2 and Fig 2.3 show the modelling results of a variable speed pump, which is obtained by using the Water Control Centre program GIPADS (ORR et al, 1988b).

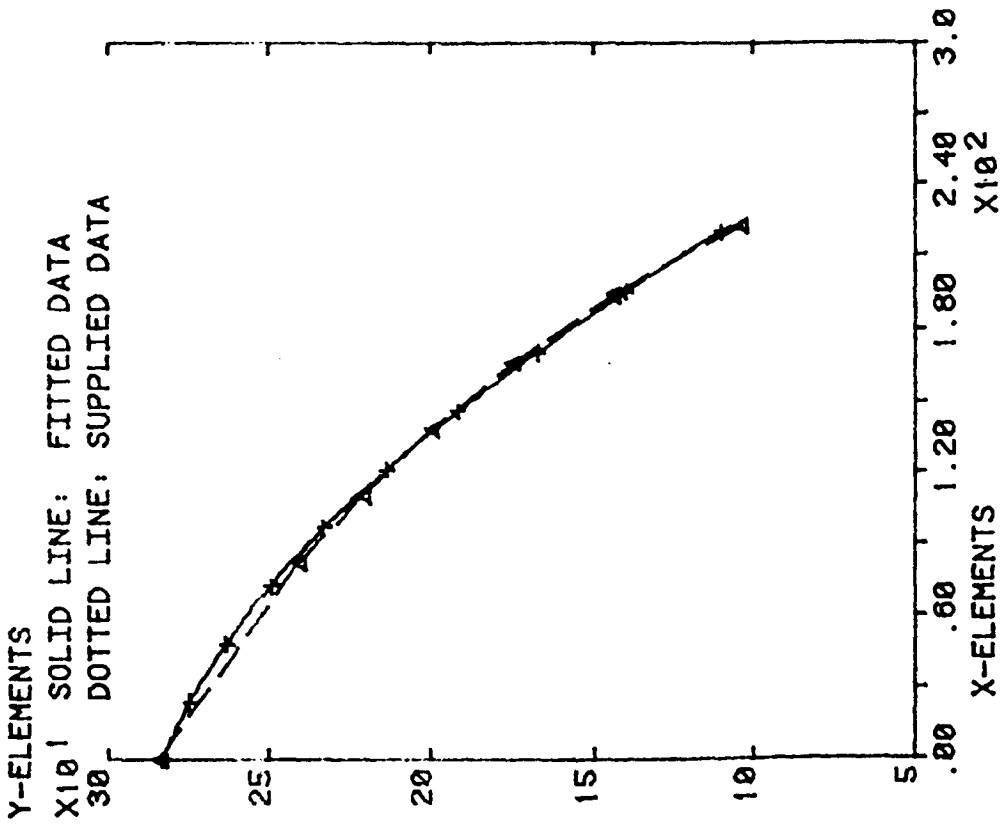
HYDRAULIC COEFFICIENTS

A-QUADRATIC TERM: -0.002256
 B-LINEAR TERM: -0.289421
 C-CONSTANT TERM: 282.93

RESULT COMPARISON		
FLOW-Q	HEAD-H	FITTED-H
0.00	284.00	282.93
63.10	241.00	243.30
111.00	221.00	223.01
138.90	200.00	199.21
166.70	175.00	172.00
194.50	144.00	141.30
222.20	104.00	107.24

X DEVIATION ERROR: 1.61

DATA ENTRY WITH		QUIT
FSP	VSP	VTP



ANALYSIS>HYD.COEFFICIENTS	DATA MODIFICATION	COEFFICIENT FITTING	RETURN
---------------------------	-------------------	---------------------	--------

Fig 2.2 Coefficient Fitting of Pump Data

REPEAT	HEAD	SET:	SPECIFIC
QUIT	% EFF	1288.	FLOW SCALE
	POWER	1165.	N/A
		▽	
		△	
		+	

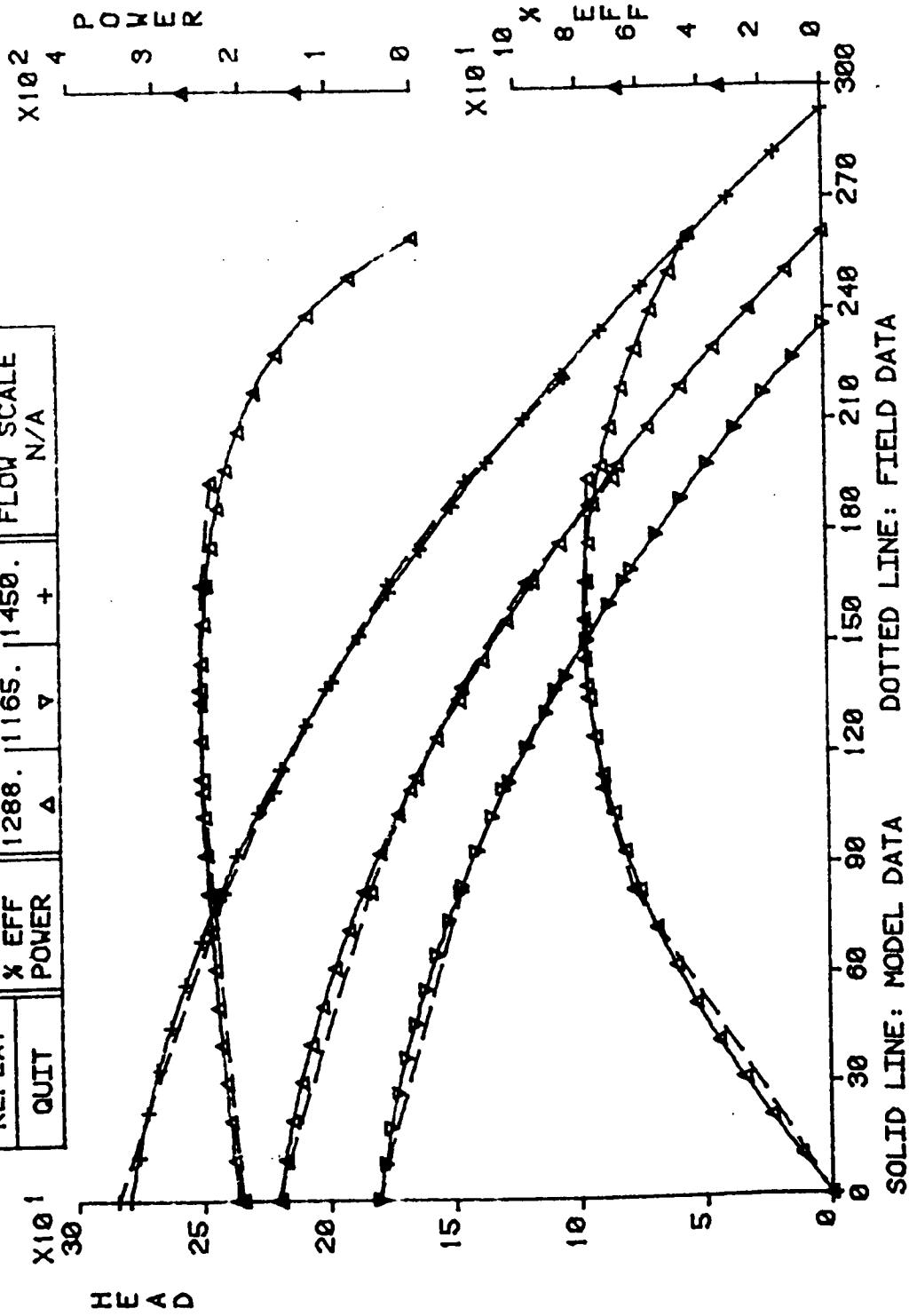


Fig 2.3 Comparison of Model and Field Data

iv) combinations of different pumps

For the case of combinations of different pumps, it is much more difficult to obtain the analytical head vs flow relationship than for that of combinations of identical pumps. In the author's studies, there are two methodologies. The first one is to carry out a field test for joint pump operations then fit the data into an approximate quadratic function as the head vs flow relationship for the pump combination:

$$H = AQ^2 + BQ + C \quad (2.6)$$

where:

H = combined pump head increase. (e.g., m)

Q = combined pump flow. (e.g., l/s)

A, B, C = coefficients for the combined pump.

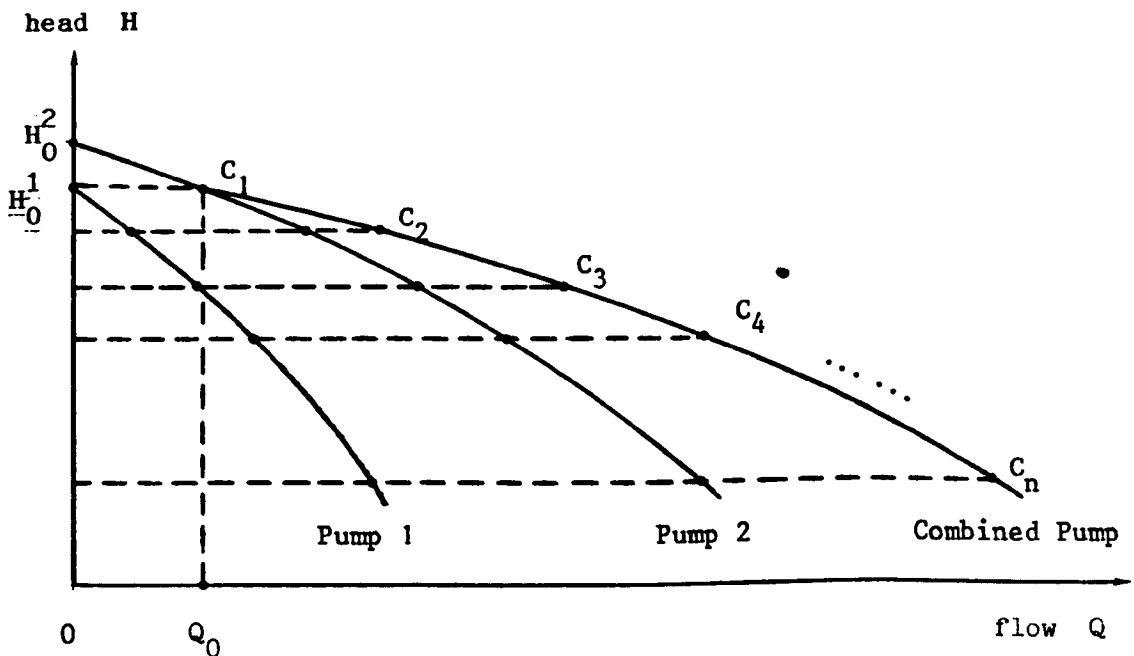


Fig 2.4 Head vs Flow Relationship for Two Different Pumps

The second methodology is shown through the following example:

Fig 2.4 shows the head vs flow relationships of two different fixed speed pumps.

Assume:

$$\text{pump 1 } H = A_1Q^2 + B_1Q + C_1$$

$$\text{pump 2 } H = A_2Q^2 + B_2Q + C_2$$

$$(H_0^2 \geq H_0^1, \text{ see fig 2.4})$$

When they are in operation simultaneously in parallel, the following holds:

$$H_1 = H_2 = H_c$$

$$Q_1 + Q_2 = Q_c$$

$$H_1 = A_1Q_1^2 + B_1Q_1 + C_1$$

$$H_2 = A_2Q_2^2 + B_2Q_2 + C_2$$

$$H = A_cQ_c^2 + B_cQ_c + C_c \dots\dots\dots (2.7)$$

where:

Q_1, Q_2 = pump 1 and pump 2 flow , respectively. (e.g., l/s)

H_1, H_2 = pump 1 and pump 2 head increase, respectively.(e.g., m)

H_c, Q_c = combined head increase and flow, respectively.

A_c, B_c, C_c = coefficients for the combined pumps to be determined.

With the above relationships, for given head vs flow relationships for pumps 1 and 2, a few data points (c_1, c_2, \dots, c_n) as shown in Fig 2.4 for the combined pump can be generated from eqn(2.7). These data

points can be used to fit an approximate quadratic function for the head vs flow relationship of the combined pumps as:

$$H = A_c Q^2 + B_c Q + C_c \quad (2.8)$$

In practice, the second approach is found to be more convenient to apply (CHEN, 1985).

The combinations of two different variable speed pumps (VSP's) and/or variable throttle pumps (VTP's) could be much more complex. It will be very difficult to obtain the head vs flow relationships of the combined pumps, since this has to be done repeatedly at a number of speed ratios for a VSP or at a number of throttle factors for a VTP.

2.2.2 Flow vs Efficiency and Power Relationships

1) Fixed Speed Pumps

a) efficiency

$$\begin{aligned} \eta_f &= \eta^* \{1 - [Q_f / (uQ^*) - 1]^2\} \\ &= \eta^* (Q_f / uQ^*) (2 - Q_f / uQ^*) \end{aligned} \quad (2.9)$$

where:

η_f = operating efficiency for single pump ($u=1$) or combined identical pumps. (%)

η^* = peak efficiency for an individual pump. (%)

Q^* = peak efficiency flow for an individual pump. (e.g., l/s)

Q_f = see eqn(2.2) and eqn(2.1).

u = see eqn(2.2).

Eqn(2.9) is a symmetrical curve about peak efficiency flow. It can be uniquely determined when peak efficiency and corresponding flow are known.

b) Pump Group Power Consumption

$$P_f = gH_f Q_f / \eta_f \quad (2.10)$$

where:

P_f = total pump group input power. (e.g., kW)

g = unit conversion factor for electrical power relating water quantities to electrical energy. (e.g., 0.98kWs/m/l)

Q_f, H_f, η_f = see eqn(2.2) and eqn(2.9), respectively

ii) Variable Speed Pumps

a) Efficiency

$$\begin{aligned} \eta_s &= \eta^* \{1 - [Q_s / (suQ^*) - 1]^2\} \\ &= \eta^* (Q_s / suQ^*) (2 - Q_s / suQ^*) \end{aligned} \quad (2.11)$$

where:

η_s = operating pump efficiency. (%)

Q^* = peak efficiency flow. (e.g., l/s)

s, Q_s, u = see eqn(2.4).

b) Total Pump Group Power Consumption

$$P_s = gH_s Q_s / \eta_s \quad (2.12)$$

A comparison between modelled and real values for pump efficiency and power characteristics for a variable speed pump are also shown in Fig

2.3. For fixed speed pumps or variable throttle pumps the results will be similar.

iii) Variable Throttle Pumps

a) Efficiency

$$\eta_f = \eta^* \{1 - [Q_f / (uQ^*) - 1]^2\} \quad (2.13)$$

b) Pump Group Power

pump group power

$$P_v = gH_f Q_v / \eta_f \quad (2.14)$$

where:

g = see eqn(2.11).

H_f = head increase of the fixed speed pump(s) in the group.(e.g., m)

Q_v = pump group flow.(e.g., l/s)

η_f = efficiency of the fixed speed pump(s) in the group.(%)

however, some of this power will be lost by dissipation in the throttle valve.

2.3 PIPES

These are the most commonly occurring elements in a network. There are several equations which are often used to evaluate the friction head loss (i.e., conversion of energy per unit weight of water into a non-recoverable form of energy) along a pipe. The most fundamentally sound method for computing such head losses is by means of the Darcy-

Weisbach equation (JEPPSON, 1979; FAIR et al, 1966; YANG, 1956):

$$H_f = f \frac{L}{D} \frac{v^2}{2g} \quad (2.15)$$

where:

H_f = friction head loss along a pipe. (e.g., m)

f = is a friction factor .

D = pipe diameter. (e.g., m)

L = pipe length. (e.g., m)

V = average velocity of flow. (e.g., m/s)

g = acceleration of gravity. (e.g., m/s²)

Whilst the Darcy-Weisbach equation is the most fundamentally sound method for determining head losses or pressure drops in closed conduit flow, empirical equations are widely used. Perhaps the most widely used of such equations are the empirical Hazen-Williams and Manning equations.

Hazen-Williams equation:

$$H_f = [10.70L / (C_{HW}^{1.852} D^{4.87})] Q^{1.852} \quad (2.16)$$

where:

H_f = head losses across pipe (m).

L = pipe length (m).

D = pipe diameter (mm).

Q = flow rate (l/s).

C_{HW} = Hazen-Williams roughness coefficient depending upon the pipe material, age, etc.

Manning equation

$$H_f = (10.29n^2L/D^{5.333})Q^2 \quad (2.17)$$

where:

H_f , L , D , Q = see eqn(2.16).

n = Manning roughness coefficient which depends upon pipe parameters.

Both eqn(2.16) and eqn(2.17) can be summarized as:

$$H_f = SQ^\alpha \quad (2.18)$$

or:

$$Q = GH_f^\beta \quad (2.19)$$

where:

S = general pipe resistance coefficient; for Hazen-Williams equation, $S = (10.70L)/(C_{HW}^{1.852}D^{4.87})$; for Manning equation, $S = (10.29n^2L)/D^{5.333}$.

$$G = 1/S^\beta$$

α = flow power coefficient, for Hazen-Williams equation, equals 1.852; for Manning equation, equals 2.0.

β = $1/\alpha$, for Hazen-Williams equation, equals 0.54; for Manning equation, it is 0.50.

2.4 PRESSURE REDUCING VALVES (PRV)

A pressure reducing valve is designed to maintain a constant pressure

(or head) downstream from the PRV when the upstream pressure (or head) is within the valve design limits. This can be modelled by assuming that between downstream node i and upstream node j , there is a valve with a setting of pressure to equal to H_{PRV} (COULBECK, 1977; JEPSON, 1979; RAO and BREE, 1977a, 1977b):

i) if $H_j \geq H_{PRV} \geq H_i$, the valve reduces the head to H_{PRV} to give a flow from node j to node i , given by:

$$Q_{ij} = G |H_{PRV} - H_i|^\beta \quad (2.20)$$

where:

Q_{ij} = flow from node j to node i . (e.g., l/s)

H_i, H_j = head at node i and node j , respectively. (e.g., m)

H_{PRV} = pressure setting at the pressure reducing valve. (e.g., m)

G = see eqn(2.19).

ii) if $H_i > H_{PRV}$, the PRV acts as a check valve, no reverse flow takes place and hence $Q_{ij} = 0$.

iii) if $H_j < H_{PRV}$ and $H_i < H_{PRV}$, the valve has no effect on flow conditions and acts as a pipe with a head drop of $(H_j - H_i)$.

2.5 CHECK VALVES (OR NON-RETURN VALVES)

A check valve is designed to allow only one direction of flow. Assume that between a downstream node i and upstream node j , there is a check valve and that the flow is only allowed from node j to node i .

1) when $H_j > H_i$, the check valve acts as an ordinary pipe,

$$Q_{ij} = G H_f^\beta \quad (2.21)$$

where:

G = see eqn(2.19).

H_f = head loss along the pipe. (e.g., m)

ii) when $H_j \leq H_i$,

$$Q_{ij} = 0 \text{ (a closed pipe)} \quad (2.22)$$

2.6 RESERVOIRS, TANKS AND TOWERS.

Reservoirs in the distribution system are sometimes called service reservoirs. The main functions of service reservoirs are to provide storage to cater for fluctuations in normal demand and a reserve for emergencies (such as fire fighting and mains bursts).

The dynamics of the i th reservoir in a system can be described as (COHEN, 1982; MCPHERSON, 1960, 1966):

$$S_i(x_i)\dot{x}_i = Q_i \quad (2.23)$$

where:

x_i = water level at time t in the i th reservoir. (e.g., m)

\dot{x}_i = time derivative of water level of i th reservoir. (e.g., m/s)

$S_i(x_i)$ = cross-sectional area of the i th reservoir at water level x_i . (e.g., m^2)

Q_i = the inflow or outflow (negative inflow) of the i th reservoir. (e.g., m^3/s)

Reservoir inflow is generally a non-linear function of pump flows, demands and reservoir water levels. Therefore, for a system with I reservoirs, M pump outflows and J demands, the dynamics of the reservoirs can be alternatively expressed in the following vector form:

$$\dot{\underline{X}} = \underline{f}(\underline{X}, \underline{QP}, \underline{D}) \quad (2.24)$$

where:

$\dot{\underline{X}} = [\dot{x}_1, \dot{x}_2, \dots, \dot{x}_I]^T$, \dot{x}_i ($i=1, 2, \dots, I$) is the time derivative of the i th reservoir level. (e.g., m/s)

$\underline{X} = [x_1, x_2, \dots, x_I]^T$, x_i ($i=1, 2, \dots, I$) is the i th reservoir level. (e.g., m)

$\underline{QP} = [QP_1, QP_2, \dots, QP_M]^T$, QP_m ($m=1, 2, \dots, M$) is the m th pump outflow. (e.g., m^3/s)

$\underline{D} = [d_1, d_2, \dots, d_J]^T$, d_j ($j=1, 2, \dots, J$) is the j th demand in the system. (e.g., m^3/s)

$\underline{f}(\cdot)$ = vector function.

Its discrete and linearized form could be (COULBECK, 1977; COHEN, 1982):

$$\underline{X}(k+1) = \underline{A}_x \underline{X}(k) + \underline{B}_x \underline{QP}(k) + \underline{C}_x \underline{D}(k), \quad (2.25)$$

and $\underline{X}(0) = \underline{X0}$

where:

$\underline{X}(k+1), \underline{X}(k)$ = reservoir level vector at time stage $k+1$ and k , respectively. (e.g., m)

$\underline{QP}(k)$ = pump outflow vector at stage k . (e.g., m^3/s)

$\underline{D}(k)$ = demand vector at stage k . (e.g., m^3/s)

$\underline{A}_x, \underline{B}_x, \underline{C}_x$ = coefficients matrices.

\underline{X}_0 = initial reservoir level vector, which is usually known or given in advance. (e.g., m)

Alternatively, from the principle of mass balance for a reservoir, the following holds:

$$\begin{aligned} V_i(k+1) &= V_i(k) + \sum Q_{1,i,k1} \Delta t_{i,k1} - \sum Q_{2,i,k2} \Delta t_{i,k2}, \\ \text{and } V_i(0) &= V_{0i} \\ i &= 1, 2, \dots, I \end{aligned} \quad (2.26)$$

where:

$V_i(k+1), V_i(k)$ = volume storage of water in reservoir i at stage $k+1$ and k , respectively (e.g., m^3)

$Q_{1,i,k1}$ = flow into reservoir i during period $k1$ of stage k . (e.g., m^3/s)

$Q_{2,i,k2}$ = flow out of reservoir i during period $k2$ of stage k . (e.g., m^3/s)

$\Delta t_{i,k1}$ = duration of period $k1$ of stage k when the inflow occurs at reservoir i . (e.g., s)

$\Delta t_{i,k2}$ = duration of period $k2$ when the outflow occurs at reservoir i . (e.g., s)

V_{0i} = initial volume storage of water in reservoir i , which is usually known or given in advance. (e.g., m^3)

The above relationship will be further explored in Chapter 5.

2.7 HYDRAULIC ANALYSIS OF WATER SUPPLY AND DISTRIBUTION SYSTEMS

The task of the hydraulic analysis of a water supply and distribution system is to calculate all pipe flows, node pressures, pumping flows, delivery heads, reservoir levels, etc. under given demands, system controls, and initial and network boundary conditions. In order to make the analysis practical, a network is first 'skeletonized' such that those pipes which are considered to be insignificant are deleted and all loads (demands) are aggregated at nodes. The static solution or network instantaneous balancing is obtained based upon the following laws (FAIR et al, 1966; JEPPSON, 1979; YANG, 1979):

1) Continuity law:

At each node:

$$\sum Q_{\text{inflow}} - \sum Q_{\text{outflow}} = 0 \quad (2.27)$$

If a pipe network contains J nodes (or junctions) and all external flows are known then there will be J-1 independent continuity equations in the form of eqn(2.27). The last, or the Jth continuity equation, is not independent, that is to say, it can be obtained from combination of the first J-1 equations.

2) Energy law:

In each loop:

$$\sum H = 0 \quad (2.28)$$

In addition to the continuity equations which must be satisfied, the energy law provides equations which must also be satisfied. These additional equations are obtained by noting that if one adds the head

losses around a closed loop, taking into account whether the head loss is positive (clockwise) or negative (counterclockwise), that upon arriving at the beginning point the net head losses equal zero. For a network with L non-overlapping (or natural) loops, the following holds:

$$\begin{aligned}
 & \text{I} \\
 & \sum_1 h_{f1} = 0 \\
 & \\
 & \text{II} \\
 & \sum_1 h_{f1} = 0 \qquad (2.29) \\
 & \\
 & : \\
 & : \\
 & : \\
 & \\
 & \text{L} \\
 & \sum_1 h_{f1} = 0
 \end{aligned}$$

where the summation on small l is over the pipes in the loops I, II, ..., L. By use of eqn(2.18), eqn(2.29) can be written in terms of the flow:

$$\begin{aligned}
 & \text{I} \\
 & \sum_1 S_1 Q_1^\alpha = 0 \\
 & \\
 & \text{II} \\
 & \sum_1 S_1 Q_1^\alpha = 0 \qquad (2.30) \\
 & \\
 & : \\
 & : \\
 & : \\
 & \\
 & \text{L} \\
 & \sum_1 S_1 Q_1^\alpha = 0
 \end{aligned}$$

If there is a pump or a valve in series with a pipe, the head increase

or head loss can be added to the relevant equation.

A pipe network consisting of J nodes (or junctions), L non-overlapping loops and N pipes will satisfy the equation:

$$N = (J - 1) + L \quad (2.31)$$

Since the flow in each pipe can be considered unknown, there will be N unknowns. The number of independent equations which can be obtained for a network are (J-1)+L. Consequently the number of independent equations is equal in number to the unknown flows in the N pipes. The (J-1) continuity equations are linear and the L energy (or head loss) equations are nonlinear.

With the advent of high speed computers, various techniques, such as linear theory method, Newton-Raphson method, finite element method have been employed to solve those simultaneous continuity equations and energy equations for instantaneous hydraulic analysis of water distribution systems. There are now many commercial, fully comprehensive graphical and interactive computer package available, which bring great convenience to researchers, engineers and operators (COULBECK, 1985; CREASEY, 1988; JEPPSON, 1979; RAO et al, 1974; SHAMIR, 1968, 1977; YANG, 1979).

2.7.1. Static Simulation

By static simulation, is meant that for fixed pump and valve controls, and fixed reservoir head and given demands, all pipe flows and nodal heads are solved and balanced according to eqn(2.27) and eqn(2.30).

This constitutes a basic hydraulic analysis of a water distribution system. There are various algorithms to solve these simultaneous equations. For technical details, consult (COULBECK, 1985; JEPPSON, 1979; SHAMIR, 1968; YANG, 1979).

2.7.2 Dynamic Simulation

Dynamic simulation or extended period simulation (COULBECK, 1977, 1985; RAO et al, 1974) consists of a sequence of static solutions which are performed successively at pre-specified intervals. Usually the time horizon is a 24-hour period, and by discretizing this period into a number of intervals, it is possible to transform the dynamic problem to a sequence of interconnected static problems. Successive time intervals are linked through the reservoir levels, which are calculated assuming reservoir flows to remain constant during each discrete time increment. At the end of each time increment, water levels in the reservoirs are updated. The flow into or out of the reservoir, as given by the flow solution, is integrated over the time interval to yield the total volume which entered or left the reservoir during the time interval. This can then be converted into a change in water level in the reservoir, using the known geometry of the reservoir. The only difference between solving a series of independent flow problems and the dynamic simulation over time is the automatic updating of water levels in the reservoir at the end of each time interval, which prepares the data for the next solution. From the computational viewpoint, there is also the difference that during a dynamic simulation, each static simulation uses the values of the previous solution as initial values for the dependent variables. Pre-

defined schedules of pump operations and valve settings and, time-varying load values are also necessarily used to update the inputs to the static solutions in each successive time interval.

Dynamic simulation is very useful and can be applied in many ways: it can be employed to investigate the operations of distribution systems under different pump combinations, valve settings and demand patterns; It can also be used in evaluating the effects of the implementation of optimized pump schedules; It can even be applied to generate data for simplified network modelling (which will be further discussed in Chapter 4) as well as other aspects.

2.8 OPTIMAL OPERATION OF WATER SUPPLY AND DISTRIBUTION SYSTEM

The aim of optimal operation of a water supply distribution system is to be fulfilled by adequately arranging pump scheduling such that the overall pumping cost of the system (sometimes also treatment cost etc.) is a minimum whilst meeting customers' service requirements and system operational constraints. Fig 2.5 shows a proposed scheme for the optimal operation of a water supply and distribution system. The objective of the studies presented in this thesis is to develop systematic methods and a suite of computer programs for a computer-based control system.

Since optimal pump scheduling is made for future operations, and also usually on a day to day basis, short-term water demand forecasting, particularly hourly demand forecasting (for the 24 hours of a day) is a prerequisite, which will be discussed in Chapter 3. Furthermore,

since water supply and distribution networks are highly non-linear, stochastic and large scale dynamic systems, network simplifications are necessary for system operation studies. Different simplification methods have been extensively studied and are summarized in Chapter 4. Finally, with the availability of demand forecasting and simplified system models, different optimization methods to suit different classes of practical systems and their validation and applications are reported in Chapter 5.

2.9 CONCLUSIONS

This chapter has given a brief description of the functions and mathematical models of important components in a water supply and distribution system. The methods and usefulness of hydraulic analysis of a water supply and distribution system are also stated. These should form a basis for the studies in the thesis. The main original contribution in this chapter is the formulation of the approximate models for the combinations of different pumps.

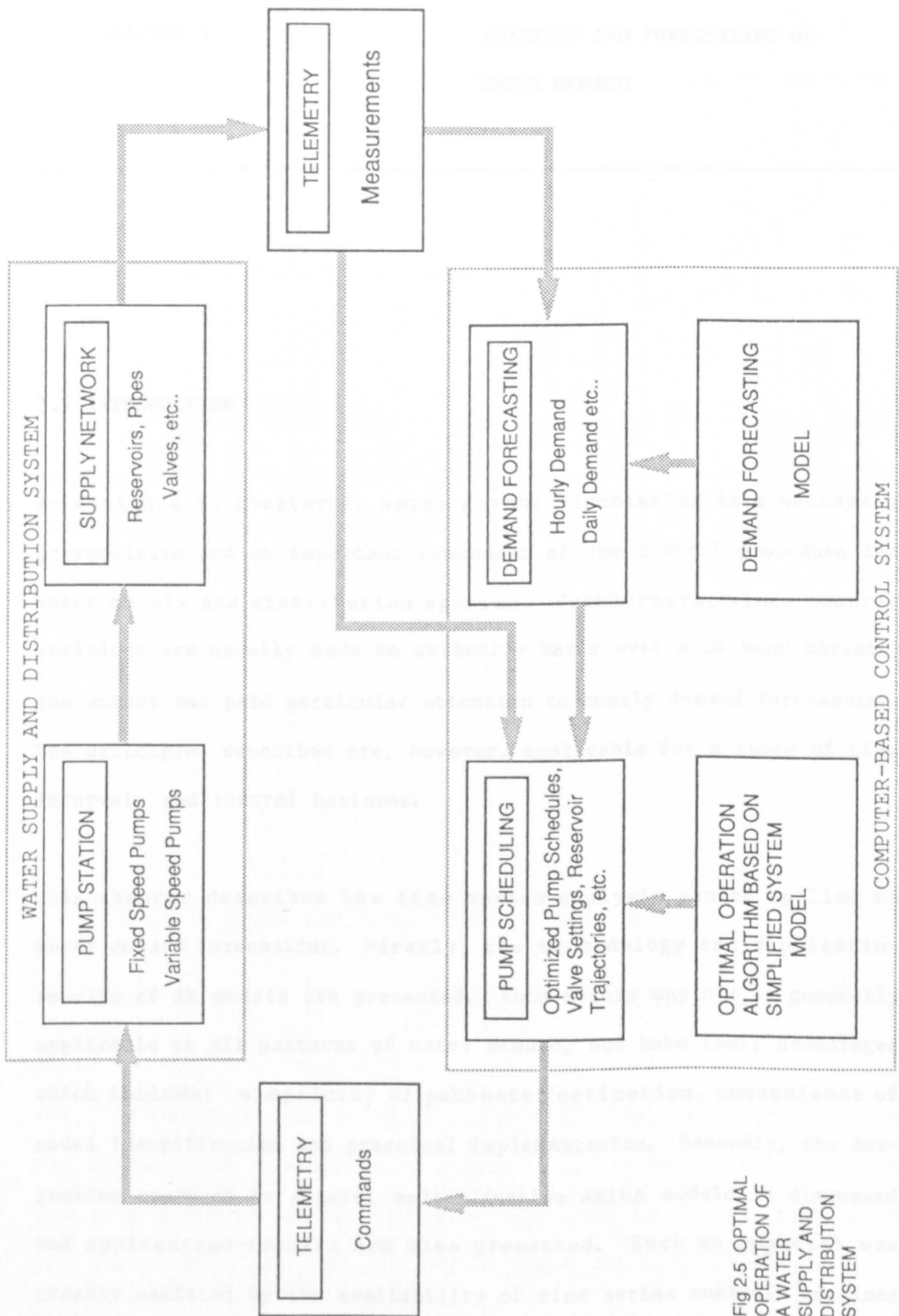


Fig 2.5 OPTIMAL OPERATION OF A WATER SUPPLY AND DISTRIBUTION SYSTEM

3.1 INTRODUCTION

As outlined in Chapter 2, water demand forecasting is a necessary prerequisite and an important component of the control procedure for water supply and distribution systems. Furthermore, since control decisions are usually made on an hourly basis over a 24 hour horizon, the author has paid particular attention to hourly demand forecasting. The principles described are, however, applicable for a range of time intervals and control horizons.

This chapter describes how time series analysis can be applied to water demand forecasting. Firstly, the methodology and application results of AR models are presented. Such models may not be generally applicable to all patterns of water demand, but have their advantages which include: simplicity of parameter estimation, convenience of model identification and practical implementation. Secondly, the Box-Jenkins approach to general multiplicative ARIMA models is discussed and application results are also presented. Such an approach was greatly assisted by the availability of time series analysis routines

from the NAG Fortran Library (NAG, 1983). Thirdly, a comparison is presented, between the performance of the time series models and an existing demand prediction program, 'GIDAP', developed at the Water Control Centre (TENNANT, 1987). The latter makes demand forecasts using the exponential smoothing technique.

3.2 CHARACTERISTICS OF WATER DEMAND

The instantaneous consumption of water in an urban distribution system is determined by the large number of industrial, commercial, public and domestic consumers distributed throughout the area supplied. This consumption is influenced by such factors as the weather, season, temperature, holidays and leaks. Thus the total demand in an urban water distribution system is a time-varying, periodic, nonstationary stochastic series, which can be modelled and predicted, by using time series analysis.

3.3 STUDIES ON A CLASS OF WATER DEMAND FORECASTING

A general class of time series models, for many time series when properly analyzed, can yield excellent fits and generates accurate forecasts. This broad class of models is called multiplicative ARIMA (Autoregressive Integrated Moving Average) models, and is formulated as follows (AN, 1983; BOX and JENKINS, 1976; VANDAELE, 1983):

$$\varphi(B) \phi(B^S) \nabla_S^D \nabla^d X_t = \theta(B) \Theta(B^S) a_t \quad (3.1)$$

where:

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad (3.2)$$

$$\phi(B^S) = 1 - \phi_1 B^S - \phi_2 B^{2S} - \dots - \phi_p B^{pS} \quad (3.3)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (3.4)$$

$$\Theta(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS} \quad (3.5)$$

∇^d = non-seasonal differencing

∇_S^D = seasonal differencing

d = order of non-seasonal differencing

D = order of seasonal differencing

S = seasonality

Alternatively, eqn(3.1) can be written as:

$$\text{ARIMA}(p,d,q)(P,D,Q)_S \quad (3.6)$$

Among this general class of multiplicative ARIMA models, there is a special class of models called AR (Autoregressive) models. For this sub-class of models, parameter estimation can be performed using linear least-squares the formulation of which is also very simple. In the author's studies, it has become evident that the consumption patterns for some water supply distribution systems can be described as an AR model or ARIMA(p,d,0)(0,0,0)₀, thus:

$$\nabla^d X_t = \varphi_1 \nabla^d X_{t-1} + \varphi_2 \nabla^d X_{t-2} + \dots + \varphi_p \nabla^d X_{t-p} + a_t \quad (3.7)$$

where:

$\nabla^d = (1-B)^d$, is the d-order backward differencing operator,

e.g. for $d = 1$, $\nabla^1 X_t = X_t - X_{t-1}$; for $d = 2$, $\nabla^2 X_t = X_t - 2X_{t-1} + X_{t-2}$, etc.

X_t = water demand at time t

φ_i = parameters to be estimated

a_t = residual at time t. a_t is a white noise series following an independent normal distribution with expectation value 0 and variance σ_a^2 .

Let $W_t = \nabla^d X_t = (1 - B)^d X_t$, then eqn(3.7) becomes:

$$W_t = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} + \dots + \varphi_p W_{t-p} + a_t \quad (3.8)$$

In fact, eqn(3.8) is an AR(p) model.

3.3.1 Estimation of Parameters φ_i

Note that eqn(3.8) is a linear model about the parameters:

$$\underline{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_p),$$

which can be estimated using the method of linear least-squares (AN, 1983; BOX and JENKINS, 1976):

The sum of squares of the residual a_t is:

$$\begin{aligned}
S(\underline{\varphi}) &= \sum_{t=p+1}^N a_t^2 \\
&= \sum_{t=p+1}^N (W_t - \varphi_1 W_{t-1} - \dots - \varphi_p W_{t-p})^2
\end{aligned}
\tag{3.9}$$

where:

N = length of the sample

In order to determine the conditions on the estimate $\hat{\underline{\varphi}}$ that minimizes $S(\underline{\varphi})$, differentiate $S(\underline{\varphi})$ with respect to $\underline{\varphi}$ and equate the result to zero, i.e.:

$$\left. \frac{\partial S(\underline{\varphi})}{\partial \underline{\varphi}} \right|_{\underline{\varphi} = \hat{\underline{\varphi}}_L} = 0$$

where:

$\hat{\underline{\varphi}}_L = (\hat{\varphi}_1^L, \hat{\varphi}_2^L, \dots, \hat{\varphi}_p^L)$, the estimate of $\underline{\varphi}$.

then, for $j= 1, 2, \dots, p$,

$$\sum_{t=p+1}^N W_{t-j} \times (W_t - \hat{\varphi}_1^L W_{t-1} - \dots - \hat{\varphi}_p^L W_{t-p}) = 0
\tag{3.10}$$

or

$$\begin{aligned} & \frac{\hat{\phi}_1^L}{N} \sum_{t=p+1}^N W_{t-1} \times W_{t-j} + \frac{\hat{\phi}_2^L}{N} \sum_{t=p+1}^N W_{t-2} \times W_{t-j} + \dots + \\ & + \frac{\hat{\phi}_p^L}{N} \sum_{t=p+1}^N W_{t-p} \times W_{t-j} = \frac{1}{N} \sum_{t=p+1}^N W_t \times W_{t-j} \end{aligned} \quad (3.11)$$

Now, if we let

$$\hat{r}_{ij}^L = \frac{1}{N} \sum_{t=p+1}^N W_{t-i} \times W_{t-j} = \hat{r}_{ji}^L$$

and

$$\begin{aligned} \underline{A} &= [\hat{r}_{ij}^L]_{p \times p} \\ \underline{B} &= [\hat{r}_{0j}^L]_{p \times 1} \end{aligned}$$

then, eqn(3.11) becomes:

$$\sum_{i=1}^p \hat{r}_{ij}^L \times \hat{\phi}_i^L = \hat{r}_{0j}^L \quad (3.12)$$

or

$$\underline{A} \underline{\hat{\phi}}_L^T = \underline{B} \quad (3.13)$$

Solving eqn(3.13) gives:

$$\underline{\hat{\phi}}_L^T = [\hat{\phi}_1^L, \hat{\phi}_2^L, \dots, \hat{\phi}_p^L]^T = \underline{A}^{-1} \underline{B}$$

then, for $t = p+1, \dots, N$

$$\hat{a}_t^L = W_t - \hat{\phi}_1^L W_{t-1} - \dots - \hat{\phi}_p^L W_{t-p} \quad (3.14)$$

The estimate of σ_a^2 can be computed from

$$\hat{\sigma}_a^2 = \frac{1}{N-p} \sum_{t=p+1}^N [\hat{a}_t^L]^2 = \frac{1}{N-p} S(\hat{\underline{\Phi}}_L) \quad (3.15)$$

3.3.2 Selection of Order p

It is suggested that the order p of eqn(3.8) be determined using an F-test (AN, 1983):

$$H_0 \text{ (hypothesis): } \varphi_{p+1} = 0$$

Let A_0 denote the residual sum of squares of AR(p+1), i.e. $S(\underline{\varphi})|_{AR(p+1)}$, and A_1 denote the residual sum of squares of AR(p), i.e., $S(\underline{\varphi})|_{AR(p)}$, then

$$F = \frac{A_1 - A_0}{1} / \frac{A_0}{N - (p+1)}$$

follows the F-distribution with the degrees of freedom 1 and N-(p+1), respectively.

For a given significance level α (generally $\alpha=0.05$ or $\alpha=0.01$), the value of $F_{\alpha}(1, N-(p+1))$ can be found from a table of F-distributions. If $F > F_{\alpha}$, this means the hypothesis is not tenable and the order of the model still needs to be increased, otherwise, AR(p) is a suitable model.

3.3.3 Forecasting

According to the theory of Minimum Mean Square Error Forecast (AN, 1983; BOX and JENKINS, 1976), it is easy to form the one-step ahead forecasts of model AR(p):

$$\hat{W}_t(1) = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} + \dots + \varphi_p W_{t-p} \quad (3.16)$$

where:

$$\hat{W}_t(1) = \text{one-step ahead forecast at time } t$$

Therefore, the one-step ahead forecast for X_t at time t , $X_t(1)$, can be obtained from $\hat{W}_t(1) = \nabla^d \hat{X}_t(1)$.

3.3.4 Application

The time series for hourly demands can be obtained by two methods: by considering each successive hour of the same day (also called series forecasting); or by considering the same hour for each successive day (also called parallel forecasting).

Table 3.1 lists the results of using the first method to model hourly demands for Shanghai, China, during 1985. Fig.3.1 is a graphical representation of these results. Table 3.2 and Fig. 3.2 shows results for the same data using the second of the above methods.

Comparing Table 3.1 and Fig 3.1 with Table 3.2 and Fig 3.2, it is evident that the forecasting precision of the second method is greater than that of the first method, under normal conditions. But under conditions which influence water demands, such as changes in weather, temperature and burst mains, large errors can occur in the second method. However, the forecasting curve of the first method is able to closely follow the changes of the actual curve. It is therefore evident that certain improvements are required to each approach in order to increase the accuracy of demand forecasts.

TABLE 3.1 MODELLING RESULTS FOR THE SERIES FORECASTING METHOD

ORDER OF DIFFERENCE	ORDER OF MODEL AND VALUES OF PARAMETERS		RESIDUAL SUM OF SQUARES		F	F_{α} $\alpha=.05$	PER CENT OF PREDICTIONS FOR WHICH THE RELATIVE ERRORS ARE NOT GREATER THAN 3%
	ϕ_1		A_1	A_2			
d=1	p=6	p=7	(p=6)	(p=7)			
	.59443	.58822					
	-.12656	-.13160	.4154	.4133	.40	3.9	85.36 %
	-.07543	-.06839					
	-.00659	-.00870					
	.04927	.05295					
	-.22292	-.23382					
		.02045					

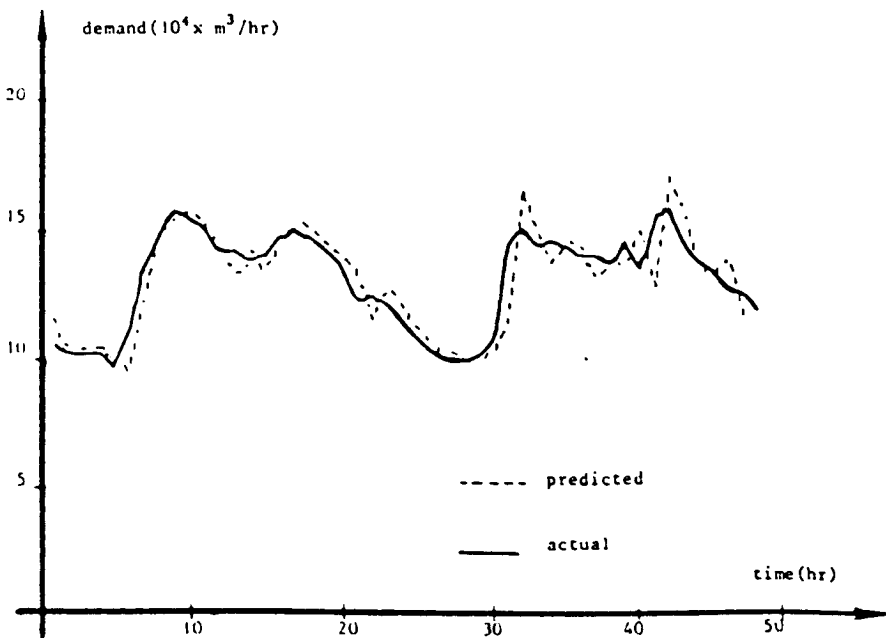


Fig 3.1 Modelling Results Using the Series Forecasting Method

TABLE 3.2 MODELLING RESULTS FOR THE SERIES FORECASTING METHOD

ORDER OF DIFFER- ENCE	ORDER OF MODEL AND VALUES OF PARAMETERS		RESIDUAL SUM OF SQUARES		F	F_{α} $\alpha=.05$	PER CENT OF PREDICTIONS FOR WHICH THE RELATIVE ERRORS ARE NOT GREATER THAN 3%
	ϕ_1	ϕ_2	A ₁ (p=6)	A ₂ (p=7)			
d=1	p=6	p=7	(p=6)	(p=7)			

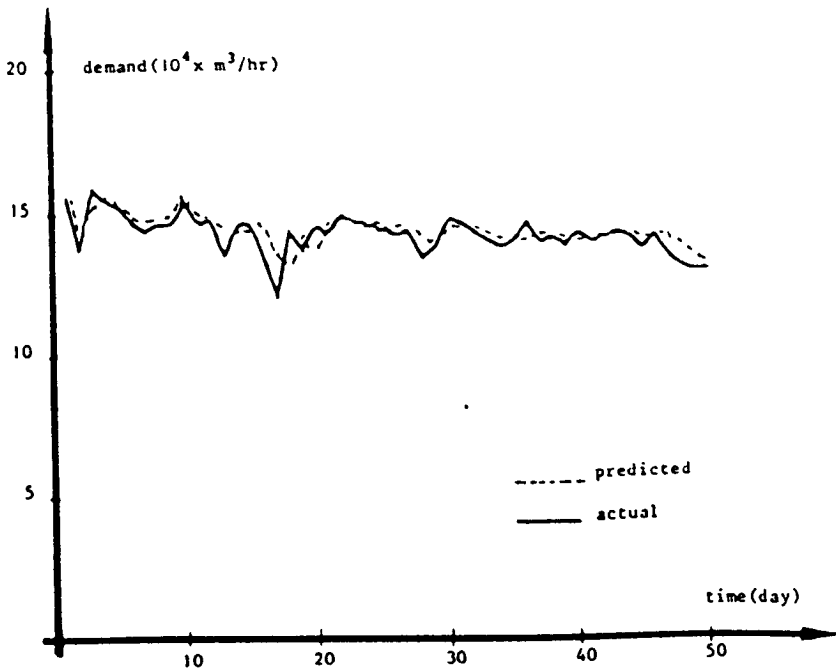


Fig 3.2 Modelling Results Using the Parallel Forecasting Method

One suggested way to improve the forecasting precision is to combine the two models together as follows:

$$Y_t(1) = c W_t^I(1) + (1-c)W_t^{II}(1) \quad (3.17)$$

where:

$Y_t(1)$ = comprehensive one-step ahead forecast at time t

$W_t^I(1)$ = one-step ahead forecast at time t by using the first method

$W_t^{II}(1)$ = one-step ahead forecast at time t by using the second method

c = weighted coefficient, $0 \leq c \leq 1$, which is determined according to the changes in weather, temperature, etc.

The result is as shown in Fig 3.3 (where $c = 0.25$).

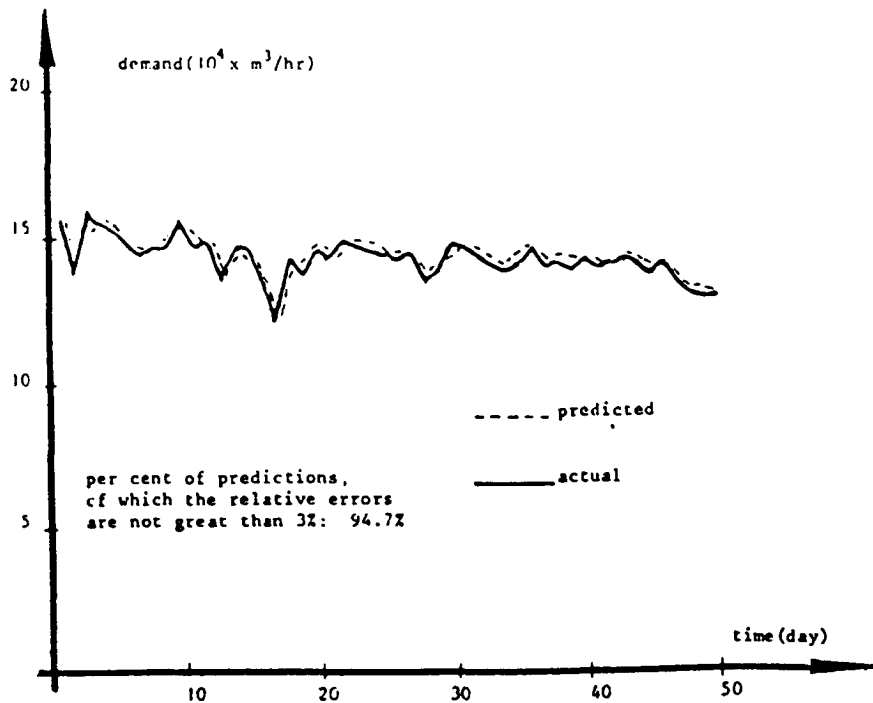


Fig 3.3 Modelling Results of the Combined Prediction Method

From the above studies, it should be noted that the estimation of parameters for AR models is solved by the linear least-squares method.

For general ARIMA models, the estimation of parameters is solved by the non-linear least-squares method, which will be discussed in detail later. It should also be noted that the selection of the order of the model describing the water demand of a particular water supply system is quite simple. Further, the one-step ahead forecasting formula is also straightforward, and the forecasting error is just white noise. From the author's experiments, it was found that AR models are quite suitable for the forecasting of water demand for Shanghai, China. The model and the algorithm originally produced by the author is successfully supporting water supply system operations in Shanghai. However, owing to the limitations on AR models in representing the characteristics of time series, this algorithm seems to be only applicable to the class of water demands which have similar characteristics to those of Shanghai.

With the availability of a series of subroutines for general time series analysis models in the NAG Fortran Library on Prime Computers in the UK, the author was therefore motivated to carry out further and more systematic studies, through time series analysis, on water demand forecasting as stated hereafter.

3.4 BOX-JENKINS APPROACH TO GENERAL ARIMA MODELS

The Box-Jenkins iterative approach for constructing multiplicative ARIMA models, eqn(3.1), basically consists of four steps (BOX and JENKINS, 1976; VANDAELE, 1983):

1. IDENTIFICATION of the specifications of the model;

2. ESTIMATION of the parameters of the model;
3. DIAGNOSTIC CHECKING of model adequacy; and
4. FORECASTING future realization.

This approach is represented as a flow chart in Fig. 3.4.

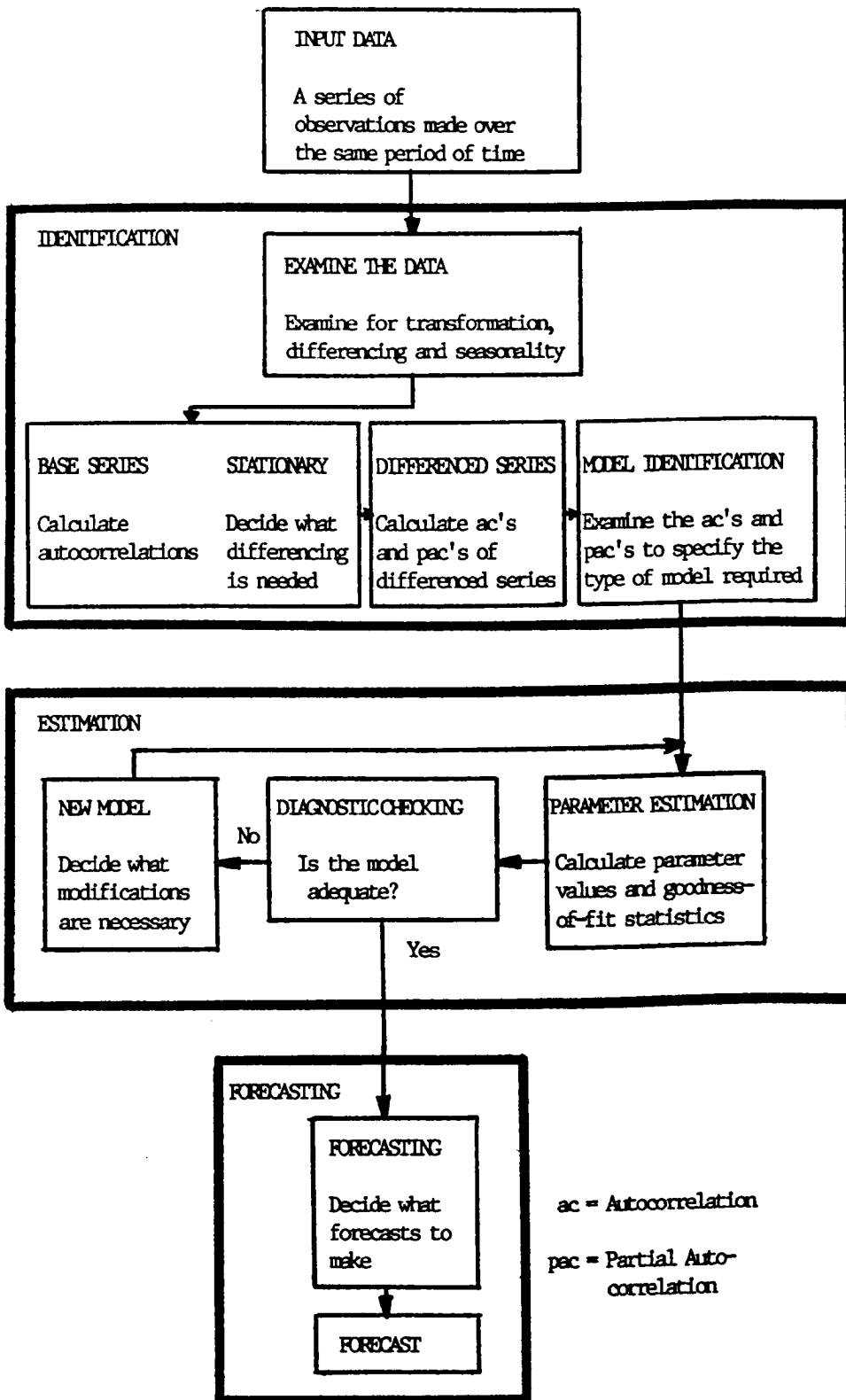


Fig 3.4 Functional Diagram of the Box-Jenkins Approach to General ARIMA Models

3.4.1. Identification

In the identification stage, one chooses a particular model from the general class of ARIMA models specified in eqn(3.1). This involves selecting the order of non-seasonal and seasonal differencing required to make the series stationary, as well as specifying the order of the non-seasonal and seasonal autoregressive and moving average polynomials necessary to adequately represent the time series model. In the identification stage, the sample autocorrelations and partial autocorrelations will be used as to judge the results.

3.4.1.1 Differencing Operations

Let $\nabla^d \nabla_s^D X_i$ be the i (th) value of a time series X_i , $i=1,2,\dots,n$ after non-seasonal differencing of order d and seasonal differencing of order D (with seasonality s). In general,

$$\nabla^d \nabla_s^D X_i = \nabla^{d-1} \nabla_s^D X_{i+1} - \nabla^{d-1} \nabla_s^D X_i \quad d > 0 \quad (3.18)$$

$$\nabla^d \nabla_s^D X_i = \nabla^d \nabla_s^{D-1} X_{i+s} - \nabla^d \nabla_s^{D-1} X_i \quad D > 0 \quad (3.19)$$

Non-seasonal differencing up to the required order d is obtained using

$$\begin{aligned}
\nabla^1 X_i &= X_{i+1} - X_i && \text{for } i = 1, 2, \dots, (n-1) \\
\nabla^2 X_i &= \nabla^1 X_{i+1} - \nabla^1 X_i && \text{for } i = 1, 2, \dots, (n-2) \\
&\vdots && \\
&\vdots && \\
&\vdots && \\
\nabla^d X_i &= \nabla^{d-1} X_{i+1} - \nabla^{d-1} X_i && \text{for } i = 1, 2, \dots, (n-d)
\end{aligned}
\tag{3.20}$$

Seasonal differencing up to the required order D is then obtained using

$$\begin{aligned}
\nabla^d \nabla_s^1 X_i &= \nabla^d X_{i+s} - \nabla^d X_i \\
&\text{for } i = 1, 2, \dots, (n-d-s) \\
\nabla^d \nabla_s^2 X_i &= \nabla^d \nabla_s^1 X_{i+s} - \nabla^d \nabla_s^1 X_i \\
&\text{for } i=1, 2, \dots, (n-d-2s) \\
&\vdots \\
&\vdots \\
&\vdots \\
\nabla^d \nabla_s^D X_i &= \nabla^d \nabla_s^{D-1} X_{i+s} - \nabla^d \nabla_s^{D-1} X_i \\
&\text{for } i=1, 2, \dots, (n-d-Ds)
\end{aligned}
\tag{3.21}$$

Mathematically, the sequence in which the differencing operations are performed does not affect the final resulting series of $m = n - d - DxS$ values. The subroutine 'G13AAF', part of the commercial NAG library (NAG, 1983), was employed in the programming. This routine performs non-seasonal and seasonal differencing for a time series.

The effect of differencing operations is to make the original non-stationary series become stationary.

3.4.1.2 Calculations of Sample Autocorrelation

Functions (ac)

The sample autocorrelation function of lag k is defined as:

$$r_k = \frac{\sum_{i=1}^{n-k} (X_i - \bar{X})(X_{i+k} - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$
$$k = 1, 2, \dots, K \quad (3.22)$$

Where

r_k = autocorrelation function of lag k .

X_i ($i = 1, 2, \dots, n$) = either original observations
or differenced values from a time series.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}, \text{ sample mean.}$$

K = number of lags required.

The actual structure of the $ARIMA(p,d,q) \times (P,D,Q)_s$ model for a specific time series is obtained by comparing the sample ac of a stationary series with the theoretical ac's (BOX and JENKINS, 1976; VANDAELE, 1983).

The subroutine 'G13ABF' in the commercial NAG library computes the sample autocorrelations of a time series.

3.4.1.3 Calculations of Sample Partial Autocorrelation

Function (pac)

Unlike the autocorrelations, the sample partial autocorrelations cannot be estimated using a simple, straightforward formula. In fact, the partial autocorrelations $P_{1,1}$ are calculated from a solution of the Yule-Walker equation system, which expresses the partial autocorrelations as a function of the autocorrelations (BOX and JENKINS, 1976), thus:

$$r_i = P_{1,1} r_{i-1} + P_{1,2} r_{i-2} + \dots + P_{1,1} r_{i-1},$$

$i=1,2,\dots,1$ (3.23)

taking $r_j = r_{|j|}$, when $j < 0$, and $r_0 = 1$

where:

1 = lag of the partial autocorrelation function.

r_i = autocorrelation function of lag i ($i=1,2,\dots,1$).

$P_{1,i}$ = parameter of the Yule-Walker equation ($i=1,2,\dots,1$),
in particular, $P_{1,1}$ is the partial autocorrelation
function required.

The partial autocorrelation function is an additional characteristic of ARIMA models. Such functions enable models to be distinguished one from another.

The subroutine 'G13ACF', part of the commercial NAG library,

calculates partial autocorrelations for a given set of autocorrelations.

A program called IDENT, which calls the NAG subroutines 'G13AAF', 'G13ABF' and 'G13ACF', was produced by the author for the identification purposes.

At this stage, by the comprehensive use of the information provided by differencing operations, ac's and pac's, several tentative models could be suggested, from the broad class of multiplicative ARIMA models. Such models need to be further specified in the estimation and, even, the forecasting stage.

3.4.2 Estimation

3.4.2.1 Parameter Estimation

After identifying a particular ARIMA model from the general class of multiplicative models,

$$\varphi(B) \phi(B^S) W_t = \theta(B) \Theta(B^S) a_t \quad (3.24)$$

where

$$W_t = \nabla_s^D \nabla^d Z_t$$

the next step is to estimate the vectors of parameters

$$\underline{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_p)^T, \quad \underline{\phi} = (\phi_1, \phi_2, \dots, \phi_p)^T, \\ \underline{\theta} = (\theta_1, \theta_2, \dots, \theta_d)^T \quad \text{and} \quad \underline{\Theta} = (\theta_1, \theta_2, \dots, \theta_Q)^T.$$

There are basically two methods available for this purpose. One such method is the least squares method; the other is known as the maximum likelihood method (BOX and JENKINS, 1976; NAG, 1983). The least squares method is the simpler one of the two methods and is thus more widely used. The methodology involves the selection of parameter values such that the sum of the squared residuals is made as small as possible. That is, $\hat{\varphi}$, $\hat{\phi}$, $\hat{\theta}$, and $\hat{\Theta}$ are chosen as estimates of φ , ϕ , θ , and Θ , respectively, such that the sum of squared residuals, SSR:

$$S(\hat{\varphi}, \hat{\phi}, \hat{\theta}, \hat{\Theta}) = \sum_{t=1}^n \hat{a}_t^2 \quad (3.25)$$

is a minimum. The NAG routine 'G13AEF' fits a seasonal ARIMA model to an observed time series, using a nonlinear least-squares procedure incorporating backforecasting. The general principle is summarized as follows:

The time series X_1, X_2, \dots, X_n supplied is assumed to follow an ARIMA model defined as:

$$\nabla^d \nabla_s^D X_t - c = W_t \quad (3.26)$$

where, $\nabla^d \nabla_s^D X_t$ is the result of applying non-seasonal differencing of order d and seasonal differencing of seasonality s and order D to the series X_t . The differenced series is then of length $N = n - d'$, where $d' = d + (Dxs)$ is the generalized order of differencing. The scalar c is the expected value of the differenced series, and the series W_1, W_2, \dots, W_N then follows a zero-mean ARMA model defined by a

pair of recurrence equations. The W_t is expressed in terms of an independent, and thus uncorrelated, series a_t , through an intermediate series e_t . The first equation describes the seasonal structure:

$$\begin{aligned}
 W_t &= \phi_1 W_{t-s} + \phi_2 W_{t-2xs} + \dots + \phi_p W_{t-pxs} + \\
 &e_t - \theta_1 e_{t-s} - \theta_2 e_{t-2xs} - \dots - \theta_q e_{t-qxs}
 \end{aligned}
 \tag{3.27}$$

If the model is purely non-seasonal the above equation becomes redundant and e_t above is equated with W_t . The second equation describes the non-seasonal structure:

$$\begin{aligned}
 e_t &= \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \dots + \varphi_p e_{t-p} + \\
 &a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}
 \end{aligned}
 \tag{3.28}$$

The estimates of the model parameters defined by

$$\begin{aligned}
 &\varphi_1, \varphi_2, \dots, \varphi_p, \theta_1, \theta_2, \dots, \theta_q, \\
 &\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q
 \end{aligned}$$

and (optionally) c are derived by minimizing a quadratic function of the vector, $\underline{W} = (W_1, W_2, \dots, W_N)^T$. This is:

$$QF = \underline{W}^T \underline{V}^{-1} \underline{W}
 \tag{3.29}$$

where:

\underline{V} = the covariance matrix of \underline{W} , and is a function of the model parameters.

When moving average parameters θ_i or Θ_i are present, so that the generalized moving average order $q' = q + s \times Q$ is positive, backforecasts $W_{1-q'}, W_{2-q'}, \dots, W_0$ are introduced as additional parameters. The 'sum of squares' function may then be written as:

$$S(\underline{pm}) = \sum_{t=1-q'}^N a_t^2 - \sum_{t=1-q'-p'}^{-q'} b_t^2 \quad (3.29a)$$

where:

\underline{pm} = a combined vector of parameters, consisting of the backforecasts followed by the ARMA model parameters.

$$p' = p + sP$$

The equations defining a_t and b_t are respectively:

$$\begin{aligned} e_t = & W_t - \phi_1 W_{t-s} - \phi_2 W_{t-2xs} - \dots - \phi_p W_{t-pxs} + \\ & \theta_1 e_{t-s} + \theta_2 e_{t-2xs} + \dots + \theta_Q e_{t-Qxs} \\ & \text{for } t = 1 - q', 2 - q', \dots, N \end{aligned} \quad (3.30)$$

$$\begin{aligned} a_t = & e_t - \varphi_1 e_{t-1} - \varphi_2 e_{t-2} - \dots - \varphi_p e_{t-p} + \\ & \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \\ & \text{for } t = 1 - q', 2 - q', \dots, N \end{aligned} \quad (3.31)$$

$$\begin{aligned} f_t = & W_t - \phi_1 W_{t+s} - \phi_2 W_{t+2xs} - \dots - \phi_p W_{t+pxs} + \\ & \theta_1 f_{t-s} + \theta_2 f_{t-2xs} + \dots + \theta_Q f_{t-Qxs} \\ & \text{for } t = (1 - q' - sxP), (2 - q' - sxP), \dots, (-q' + P) \end{aligned} \quad (3.32)$$

$$\begin{aligned} b_t = & f_t - \varphi_1 f_{t+1} - \varphi_2 f_{t+2} - \dots - \varphi_p f_{t+p} + \\ & \theta_1 b_{t-1} + \theta_2 b_{t-2} + \dots + \theta_q b_{t-q} \\ & \text{for } t = (1 - q' - p'), (2 - q' - p'), \dots, (-q') \end{aligned} \quad (3.33)$$

For all four of these equations, the following conditions hold:

$$\begin{aligned}
 W_i &= 0 \text{ if } i < 1-q' \\
 e_i &= 0 \text{ if } i < 1-q' \\
 a_i &= 0 \text{ if } i < 1-q' \\
 f_i &= 0 \text{ if } i < 1-q'-sxP \\
 b_i &= 0 \text{ if } i < 1-q'-p'
 \end{aligned}
 \tag{3.34}$$

An extension of the algorithm of Marquardt (MARQUARDT, 1963) is used for the minimization of S with respect to \underline{pm} . The first derivatives of S with respect to the parameters are calculated as:

$$2 \sum a_t \times a_{t,i} - 2 \sum b_t \times b_{t,i} = 2G_i
 \tag{3.35}$$

where:

$$\begin{aligned}
 a_{t,i} &= \text{derivatives of } a_t \text{ with respect to the } i\text{th parameter.} \\
 b_{t,i} &= \text{derivatives of } b_t \text{ with respect to the } i\text{th parameter.}
 \end{aligned}$$

The second derivative of S is approximated by

$$2 \sum a_{t,i} \times a_{t,j} - 2 \sum b_{t,i} \times b_{t,j} = 2H_{ij}
 \tag{3.36}$$

Successive parameter iterations are performed to calculate a vector of corrections \underline{dpm} by solving the equations:

$$(H + \rho D)\underline{dpm} = -G
 \tag{3.37}$$

where:

$$\begin{aligned}
 G &= \text{a vector with elements } G_i. \\
 H &= \text{a matrix with elements } H_{ij}
 \end{aligned}$$

ρ = a scalar controlling the search

D = the diagonal matrix of H.

The new parameter values are then $\underline{p}_m + \underline{dpm}$.

If a step results in new parameter values which give a reduced value of S, then ρ is reduced by a factor τ . If a step leads to new parameter values which produce an increased value of S, or in ARMA model parameters which break the stationarity and invertibility conditions (BOX and JENKINS, 1976), then the new parameters are rejected, ρ is increased by the factor τ , and the revised equations are solved for a new parameter correction. This action is repeated until either a reduced value of S is obtained, or ρ reaches a prescribed limit (e.g., 10^9). This limit is used to indicate a failure of the search procedure.

The estimated residual variance is

$$\text{erv} = S_{\min} / \text{df} \quad (3.37a)$$

where:

S_{\min} = the final value of S

The residual number of degrees of freedom is given by

$$\text{df} = \begin{cases} N - p - q - P - Q - 1, & \text{if } c \text{ is estimated;} \\ N - p - q - P - Q, & \text{if } c \text{ is not estimated.} \end{cases}$$

The covariance matrix of the vector of estimates \underline{p}_m is given by

$\text{erv} \times H^{-1}$

where:

H = is evaluated at the final parameter values.

From this expression the vector of standard deviations and the correlation matrix for the whole parameter set are derived.

The differenced series W_t , intermediate series e_t and residual series a_t are all available upon completion of the iterations over the range

$$t = 1 - q', 2 - q', \dots, N$$

For convenient application in forecasting, the following quantities constitute the 'state set', which contains the minimum amount of time series information needed to construct forecasts:

- (i) the differenced series W_t , for $(N - s \times P) < t \leq N$;
- (ii) the d' values required to reconstitute the original series X_t from the differenced series W_t ;
- (iii) the intermediate series e_t , for $(N - \max(p, Q \times s)) < t \leq N$;
- (iv) the residual series a_t , for $(N - q) < t \leq N$.

This state set is available upon completion of the iterations.

3.4.2.2 Diagnostic Checking

Once a model has been identified and its parameters estimated, it is

necessary to verify whether the model can be improved upon or not. This procedure is called diagnostic checking.

The diagnostic checking will be performed mainly in the following aspects:

a) Examination of the standard errors of the parameter estimates to see whether the parameter estimates are different from zero;

b) Inspection of the sampling correlations matrix of the parameter estimates. The correlation matrix expresses the degree of correlation that exists between the different parameter estimates. High correlation between the different parameter estimates points in the direction of model simplification;

c) Residual analysis: if a model adequately represents the ARIMA process governing the series being studied, then the residuals of the model should be white noise. This can be checked by examining the autocorrelations of the residuals. If the residuals are truly white noise, then their ac's should have no significant spikes. Further, the analysis of the residual autocorrelations can be based on the Ljung-Box Q- statistic or the Portmanteau test (VANDAELE,1983):

$$Q(K) = n(n+2) \sum_{k=1}^K [1/(n-k)] \times r_k^2(a) \quad (3.38)$$

If the fitted model is appropriate, i.e., if the residuals are white noise, Q is approximately distributed as a χ^2 (chi-squared) distributed variable with $K-p-q-P-Q$ degrees of freedom. The hypothesis that the residuals are white noise is rejected when values

of the Q statistics are large relative to the value from a χ^2 (chi-squared) distribution table.

d) Overfitting: a very useful check on the model adequacy is to determine whether the current model contains redundant parameters. Redundant parameters can be detected by a careful use of the estimate of the standard error of the parameter estimates (SE) and the estimate of the correlations between these parameter estimates.

e) Underfitting: in order to verify that the tentatively identified model contains the appropriate number of parameters to represent the data, one can include an additional parameter in the ARIMA model to see if this addition results in a significant improvement over the original model.

A program ESTIMA, which employs the commercially available NAG routines, 'G13AEF' and 'G13ABF' (NAG,1983), etc., has been written for the estimation and diagnostic checking.

3.4.3 Forecasting

Once a fitted model has been judged as adequately representing the process governing the series, it can be used to generate forecasts for further periods, which are based on the theory of Minimum Mean Square Error Forecast (BOX and JENKINS, 1976):

Let m_h be the expected value of Z_{n+h} forecasted at time period n , $m_h = EZ_{n+h}$. Also, let m be any other forecast of Z_{n+h} defined as:

$$m = m_h + d \quad (3.39)$$

where:

$d =$ is the difference between m and m_h .

Using the point forecast m , the expected value of the forecast error squared is then:

$$E[(Z_{n+h} - m)^2] = E\{[Z_{n+h} - (m_h + d)]^2\} \quad (3.40)$$

Rearranging, the right-hand side of eqn(3.40) becomes:

$$\begin{aligned} E[(Z_{n+h} - m)^2] &= E[(Z_{n+h} - m_h)^2] \\ &\quad - 2dE(Z_{n+h} - m_h) + d^2 \end{aligned} \quad (3.41)$$

Since $m_h = EZ_{n+h}$, the second term on the right-hand side of eqn(3.41) is equal to 0 and since $d^2 \geq 0$, the minimum of $E[(Z_{n+h} - m)^2]$ will be achieved only when $d=0$. However, the $E[(Z_{n+h} - m_h)^2]$ is the mean squared error of the forecast m_h . Therefore, the optimal mean squared error forecast of Z_{n+h} is obtained for $m = m_h = EZ_{n+h}$.

To calculate the mean of the forecast distribution, $E(Z_{n+h})$, let Z_t denote a stationary and invertible ARMA(p,q) process. For the time period $t = n + h$, this process can be expressed as

$$\begin{aligned} Z_{n+h} &= \varphi_1 Z_{n+h-1} + \dots + \varphi_p Z_{n+h-p} + a_{n+h} - \theta_1 a_{n+h-1} \\ &\quad - \dots - \theta_q a_{n+h-q} \end{aligned} \quad (3.42)$$

By utilizing information up to period n , the expected value of Z_{n+h} in eqn(3.42), is obtained as follows:

a) replace the current and past errors a_{n+j} , $j \leq 0$, with actual residuals;

b) replace each future error a_{n+j} , $0 < j \leq h$, with its expectation, which, since a_{n+j} is white noise, is 0;

c) replace current and past observations Z_{n+j} , $j \leq 0$, with the actual observed values;

d) replace each future value of Z_{n+j} , $0 < j < h$, with their forecast $Z_n(j)$. $Z_{n+1}, Z_{n+2}, \dots, Z_{n+h-1}$ should therefore be forecasted in advance in order to forecast Z_{n+h} .

The approach discussed above is not restricted to just ARMA models. It can easily be extended to obtain minimum mean squared error forecasts for any nonseasonal, as well as multiplicative seasonal model, stationary as well as nonstationary ARIMA model.

As soon as new observations become available, the forecasts for further periods beyond the new observations can be generated in two different ways (VANDAELE, 1983):

a) sequential updated forecasting. By using the new observations, the parameters of the current model can be re-estimated and then forecasts can be made in the usual manner.

b) adaptive forecasting. The parameters of the current model are

left unchanged but the origin for forecasting can be changed to incorporate the new observations.

A program FRCAST, which uses the commercially available NAG subroutines 'G13AGF' and 'G13AHF' (NAG, 1983), has been designed for forecasting.

3.4.4 Applications to Water Demand Forecasting

3.4.4.1 Demand Data

The demand data studied came from the telemetry meter M6 (chase view, 250 mm) within the Eastern Zone of the Wolverhampton region (TENNANT, 1987). Hourly demand data from 24 May 1988 to 21 Jun 1988 are available, although some of the telemetry data were lost during data transmission. The data is first processed by the data screening scheme in the demand prediction program GIDAP (TENNANT, 1987) in order to remove potential measurement errors in the data.

The data has been used to verify the Box-Jenkins approach. Typical results are presented below.

3.4.4.2 Identification

The initial set of hourly demand data, for model identification, is for 24 May (Tue) 1988 to 30 May (Mon) 1988 (one week).

Various differencing operations have been extensively applied to the data. Table 3.4 contains the autocorrelations and partial autocorrelations of several main series.

TABLE 3.4 AUTOCORRELATIONS AND PARTIAL AUTOCORRELATIONS

SERIES WITH D=1 DS=0
 MEAN = 0.168 VARIANCE = 485.96

AUTOCORRELATIONS

LAGS	ROW SE*												
1-12	.08	.67	.29	.06	-.08	-.22	-.33	-.31	-.19	-.10	-.07	-.06	-.09
13-24	.13	-.10	-.10	-.10	-.19	-.29	-.30	-.20	-.07	.03	.20	.54	.78
25-36	.18	.63	.33	.09	-.03	-.18	-.28	-.28	-.20	-.13	-.06	-.04	-.08
37-48	.20	-.10	-.07	-.06	-.13	-.23	-.23	-.16	-.07	.00	.12	.38	.60

PARTIAL AUTOCORRELATIONS (STANDARD ERROR = 0.08)

LAGS													
1-12	.67	-.29	.01	-.11	-.17	-.13	.01	.02	-.05	-.09	-.06	-.16	
13-24	-.03	-.06	-.08	-.29	-.25	-.21	-.11	-.17	-.18	-.04	.39	.33	
25-36	-.05	-.10	-.02	.10	-.02	.04	-.02	-.05	-.10	.02	-.02	-.05	
37-48	.02	.05	.03	.03	.06	.09	-.05	-.02	.03	-.06	-.08	.05	

SERIES WITH D=0 DS=1 S=24
 MEAN = -6.368 VARIANCE = 280.11

AUTOCORRELATIONS

LAGS	ROW SE												
1-12	.08	.82	.64	.51	.42	.39	.38	.38	.41	.43	.40	.39	.37
13-24	.21	.33	.28	.26	.26	.23	.18	.15	.12	.06	.05	.07	.08
25-36	.23	.14	.11	.05	.01	-.05	-.11	-.13	-.12	-.14	-.11	-.13	-.15
37-48	.23	-.15	-.12	-.12	-.12	-.10	-.08	-.09	-.11	-.14	-.18	-.20	-.22

PARTIAL AUTOCORRELATIONS (STANDARD ERROR = 0.08)

LAGS													
1-12	.82	-.07	.04	.02	.12	.08	.07	.12	.08	-.07	.10	.00	
13-24	-.02	-.08	.09	.01	-.09	-.11	.06	-.11	-.11	.07	.09	-.10	
25-36	.22	-.20	-.07	.01	-.05	-.09	.03	-.01	-.05	.03	-.10	-.04	
37-48	.10	.08	-.04	.01	.07	.06	-.10	.05	-.07	.01	-.05	-.09	

SERIES WITH D=1 DS=1 S=24
 MEAN = -0.13 VARIANCE = 102.33

AUTOCORRELATIONS

LAGS	ROW SE												
1-12	.08	-.02	-.13	-.08	-.17	-.07	-.03	-.06	.00	.13	-.03	.02	.03
13-24	.09	.03	-.09	-.02	.06	.07	-.08	.02	.05	-.12	-.10	.07	-.19
25-36	.10	.27	.12	-.10	.08	-.01	-.12	-.04	.04	-.11	.13	.01	-.08
37-48	.11	-.05	.07	-.02	-.04	-.03	.13	.00	.02	.03	-.05	-.03	-.02

PARTIAL AUTOCORRELATIONS (STANDARD ERROR = 0.08)

LAGS													
1-12	-.02	-.14	-.09	-.20	-.12	-.11	-.15	-.11	-.03	-.11	-.02	-.01	
13-24	.05	-.10	-.02	.06	.10	-.09	.06	.08	-.09	-.14	.08	-.23	
25-36	.21	.01	-.03	.03	.05	-.06	.02	.01	-.05	.08	.00	-.14	
37-48	-.08	.01	-.06	-.11	-.08	.07	-.09	.06	-.04	.01	.07	-.07	

* ROW SE is the standard errors for r_k under the null hypothesis that the cure process is a MA(0), MA(12), MA(24), etc. (one less than the first lag printed on each row) (VANDAELE, 1983).

The autocorrelation r_k whose sample value does not exceed twice the standard error may suggest that its theoretical autocorrelation is essentially zero. This is also applicable to partial autocorrelations (BOX and JENKINS, 1976).

After careful analysis of the autocorrelations and partial autocorrelations of each series, several tentative models are developed according to the properties of different ARIMA models revealed by their autocorrelation and partial autocorrelation functions:

a) For series ∇X_t , the tentative model is

Model 1: ARIMA (2,1,0)(2,0,0)₂₄

b) For series $\nabla_{24} X_t$, the tentative model is:

Model 2: ARIMA (1,0,0)(0,1,0)₂₄

c) For series $\nabla \nabla_{24} X_t$, the tentative model is:

Model 3: ARIMA (1,1,0)(1,1,0)₂₄

3.4.4.3 Estimation and Diagnostic Checking

For those tentative models stated above, parameters have been estimated and different diagnostic checking measures have been applied. Eventually, the most adequate model is shown to be the ARIMA (0,1,0)(1,1,0)₂₄ summarized in Table 3.5, which is based on the above tentative model 3.

TABLE 3.5. COMPUTED RESULTS
FOR THE MODEL ARIMA (0,1,0)(1,1,0)₂₄

Data used for estimation: 24 May to 30 May 88.

***** RESULTS FOR PRELIMINARY ESTIMATES *****

AUTOREGRESSIVE p=0
DIFFERENCING d=1
MOVING AVERAGE q=0
SEASONAL AUTOREGRESSIVE P=1
SEASONAL DIFFERENCING D=1
SEASONAL MOVING AVERAGE Q=0
SEASONALITY S=24

ARIMA MODEL PARAMETER PRELIMINARY ESTIMATE

-0.19000

RESIDUAL VARIANCE = 98.63878

***** RESULTS FOR LEAST SQUARES ESTIMATES *****

ITERATION 0 RESIDUAL SUM OF SQUARES = 0.1390E+05
ITERATION 1 RESIDUAL SUM OF SQUARES = 0.1388E+05

CONVERGENCE ACHIEVED AFTER 1 CYCLE

RESIDUAL SUM OF SQUARES IS 13878.328 WITH 141 DEGREES OF FREEDOM

RESIDUAL VARIANCE RMS = 98.428

LEAST SQUARES METHOD
PARAMETER ESTIMATES

	EST	SE	EST/SE
PAR(1)	-0.2368	0.0919	-2.5753
CONSTANT	-0.1380	0.6935	-0.1989

TABLE 3.5. (Continued)

***** RESULTS FOR DIAGNOSTIC CHECKS *****

SAMPLE MEAN OF RESIDUALS = 0.0152
 SAMPLE VARIANCE OF RESIDUALS = 98.4617

AUTOCORRELATIONS OF RESIDUALS

LAGS	ROW SE	AUTOCORRELATIONS							
1-12	0.08	0.06	-0.14	-0.15	-0.15	-0.08	-0.08		
		-0.05	0.03	0.11	-0.02	0.03	0.02		
13-24	0.09	0.03	-0.09	0.00	0.06	0.06	-0.06		
		0.00	0.02	-0.14	-0.15	0.06	0.02		
25-36	0.10	0.26	0.10	-0.11	0.03	-0.02	-0.10		
		-0.05	0.02	-0.08	0.12	0.02	-0.08		
37-48	0.11	-0.05	0.04	-0.01	-0.01	-0.02	0.11		
		0.00	0.03	-0.01	-0.06	-0.02	-0.05		

CHI-SQUARED TEST (PORTMANTEAU TEST)

	COMPUTED	DEG.FREE.	TABLE VALUE
Q(12)	14.9	11	21.0
Q(24)	26.3	23	36.4
Q(36)	49.8	35	49.8
Q(48)	54.6	47	66.0

3.4.4.4 Forecasting

With the well-defined ARIMA (0,1,0)(1,1,0)₂₄ model,

$$(1 + 0.2368 B^{24}) \nabla_{24} \nabla X_t = a_t \quad (3.43)$$

the demand forecasts up to 24 steps ahead can be made in the previous day, to form the hourly demand forecasts for the following day. Results for the 3rd June 1988 (Fri.) are given in Table 3.6 and Fig 3.5.

When the actual demand data of one day becomes available, the forecasts up to 24 steps ahead to form the hourly demand of the next day can be made by adaptive forecasting on a day to day basis. Fig 3.6 and Table 3.7 display the forecasting results for 17 June 1988(Fri), which show that forecasting precisions are still satisfactory, though the data used to build the model is for the period 24 May to 30 May 1988.

By using the hourly demand data from 17 June to 23 June 1988 for identification and estimation, the model is still

ARIMA (0,1,0)(1,1,0)₂₄:

$$(1 + 0.2923B^{24}) \nabla_{24} \nabla X_t = a_t \quad (3.44)$$

and the parameter -0.2923 is quite close to the parameter -0.2368 in model (3.43). This further indicates the adequacy and stability of the model.

In the building of the models discussed above, the data include the hourly demands of Saturday and Sunday (28 May and 29 May). When the model is used to produce forecasts for Saturdays and Sundays, it is noticed that large forecasting deviations can occur. Large deviations can also occur in forecasts for Mondays. This is because the demand patterns of Saturdays and Sundays are quite different from each other and from week days, and it is not appropriate to incorporate them in one time series. Additionally, the hourly demand data from 24 May to 1 June 1988, excluding Saturday (28 May) and Sunday (29 May), was used to build the ARIMA model. After identification, estimation and diagnostic checking, the most appropriate model is shown to be the ARIMA(0,1,0)(2,1,0)₂₄ model:

$$(1 + 0.6170 B^{24} + 0.3321B^{48}) \nabla_{24} X_t = a_t, \quad (3.45)$$

which is summarized in Table 3.8.

Using Model (3.45), better forecasting precisions are achieved. Fig 3.7 and Table 3.9 show the comparisons of forecasts for 20 June 1988 (MON) between Model (3.43) and Model (3.45).

As for the forecasts for Saturdays and Sundays, separate time series of hourly demands need to be formed which only include hourly demands for Saturdays or Sundays. It is anticipated that such an approach would improve the forecasts.

TABLE 3.6 FORECASTING RESULTS FOR 3 JUNE 1988 (FRIDAY)

	ACTUAL	FORECAST	95%CONF.LIMIT		DEVIATION FROM ACTUAL(%) *
			LOWER	UPPER	
1	160.3800	159.7420	140.2967	179.1873	0.40
2	142.7500	138.1060	110.6062	165.6058	3.25
3	134.0000	132.1388	98.4585	165.8191	1.39
4	134.2500	128.2944	89.4037	167.1850	4.44
5	136.7500	132.4744	88.9933	175.9555	3.13
6	142.5000	138.2308	90.5997	185.8619	3.00
7	179.5000	176.2636	124.8161	227.7111	1.80
8	235.7500	223.3296	168.3299	278.3293	5.27
9	279.5000	269.4012	211.0652	327.7372	3.61
10	288.5000	295.5988	234.1073	357.0903	2.46
11	296.0000	289.8484	220.3556	349.3412	3.77
12	294.5000	279.3556	211.9950	346.7162	5.14
13	276.5000	268.3952	198.2841	338.5063	2.93
14	259.0000	266.4816	193.7239	339.2394	- 2.89
15	254.2500	263.2772	187.9658	338.5886	- 3.55
16	249.5000	253.8956	176.1143	331.6769	- 1.76
17	256.8800	261.4952	181.3201	341.6703	- 1.80
18	250.6300	251.0336	168.5341	333.5331	- 0.16
19	259.7500	246.9540	162.1938	331.7142	4.39
20	266.5000	242.6584	155.6963	329.6205	8.95
21	230.0000	240.0212	150.9115	329.1309	- 4.36
22	218.6300	210.1976	118.9909	301.4043	3.86
23	210.0000	210.4328	117.1763	303.6893	- 0.21
24	187.6300	183.0528	87.7906	278.3150	2.11

Ave. 3.10⁺

* is calculated from $|(ACTUAL - FORECAST)/ACTUAL| \times 100$.

+ is calculated from $\sum |DEVIATION FROM ACTUAL| / 24$.

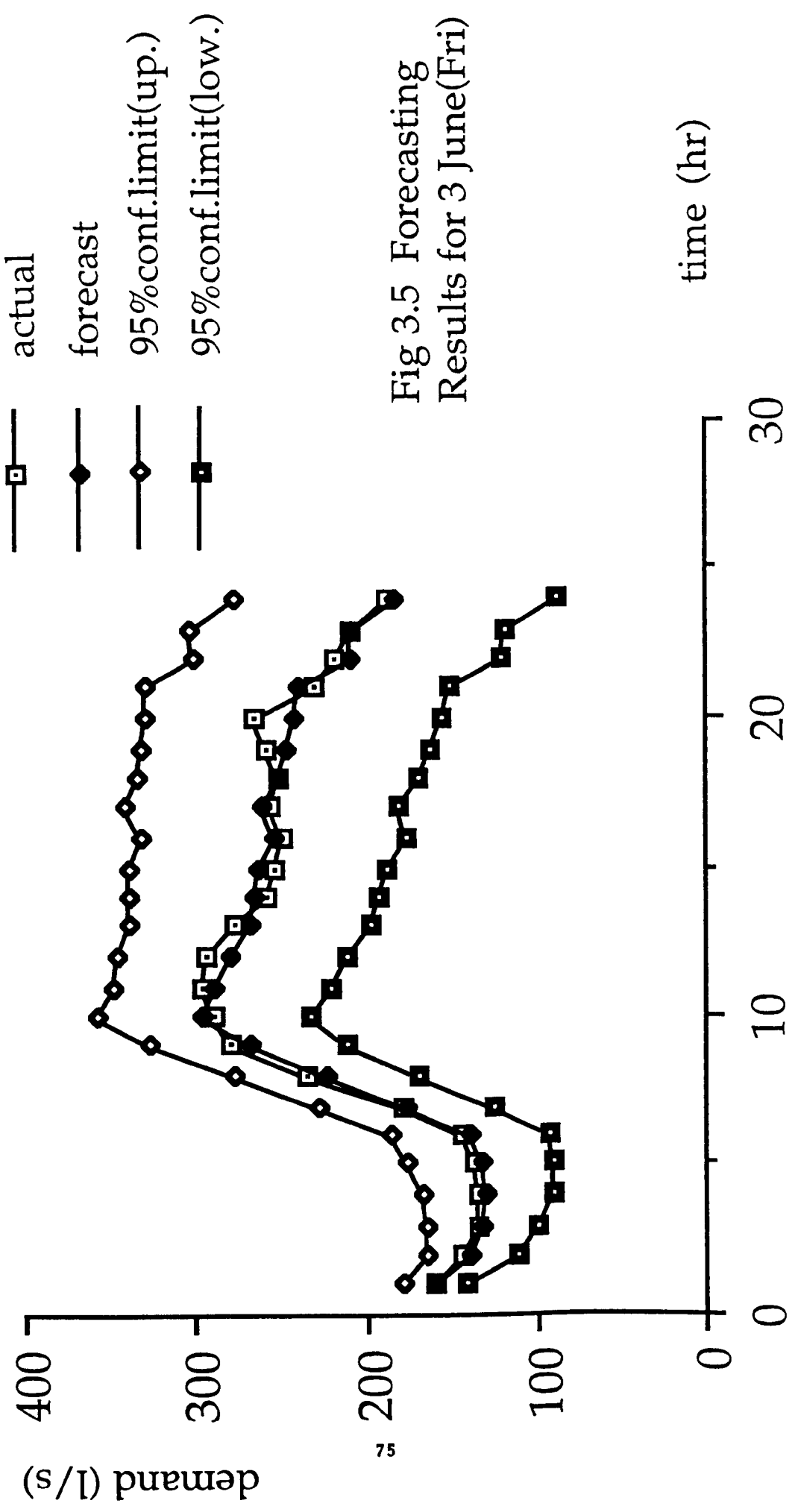


Fig 3.5 Forecasting Results for 3 June (Fri)

TABLE 3.7 FORECASTING RESULTS FOR 17 JUNE 1988 (FRIDAY)

	ACTUAL	FORECAST	95% CONF. LIMIT		DEVIATION FROM ACTUAL (%) *
			LOWER	UPPER	
1	158.7500	157.7103	138.2650	177.1556	0.65
2	144.9600	139.4444	111.9446	166.9443	3.80
3	145.7500	138.1024	104.4221	171.7827	5.25
4	140.0000	138.7140	99.8234	177.6047	0.92
5	142.5000	137.6654	94.1843	181.1465	3.39
6	155.7500	147.0304	99.3993	194.6616	5.60
7	204.7500	190.7759	139.3284	242.2234	6.82
8	287.5000	275.0969	206.0972	316.0965	4.32
9	306.3000	316.9056	258.5697	375.2416	-3.46
10	311.0000	304.5984	243.1069	366.0900	2.06
11	306.0000	300.8769	236.3841	365.3698	1.67
12	295.5000	294.0100	226.6494	361.3706	0.50
13	286.2500	303.2564	233.1453	373.3676	-5.94
14	272.0000	289.9536	217.1959	362.7114	-6.60
15	277.0000	281.1968	205.8854	356.5083	-1.52
16	269.3800	279.3352	201.5540	357.1165	-3.70
17	270.4600	283.7969	203.6217	363.9720	-4.93
18	265.1800	278.7368	196.2373	361.2364	-5.11
19	273.7100	286.7368	201.9766	371.4971	-4.76
20	264.7500	271.5659	184.6038	358.5280	-2.57
21	234.0000	258.7824	169.6728	347.8921	-10.59
22	227.5000	238.9737	147.7670	330.1803	-5.04
23	203.5000	213.2742	130.0177	316.5307	-4.80
24	191.7300	201.1865	105.9243	296.4488	-4.93
					Ave. 3.70 ⁺

* see the footnote of Table 3.6.

+ see the footnote of Table 3.6.

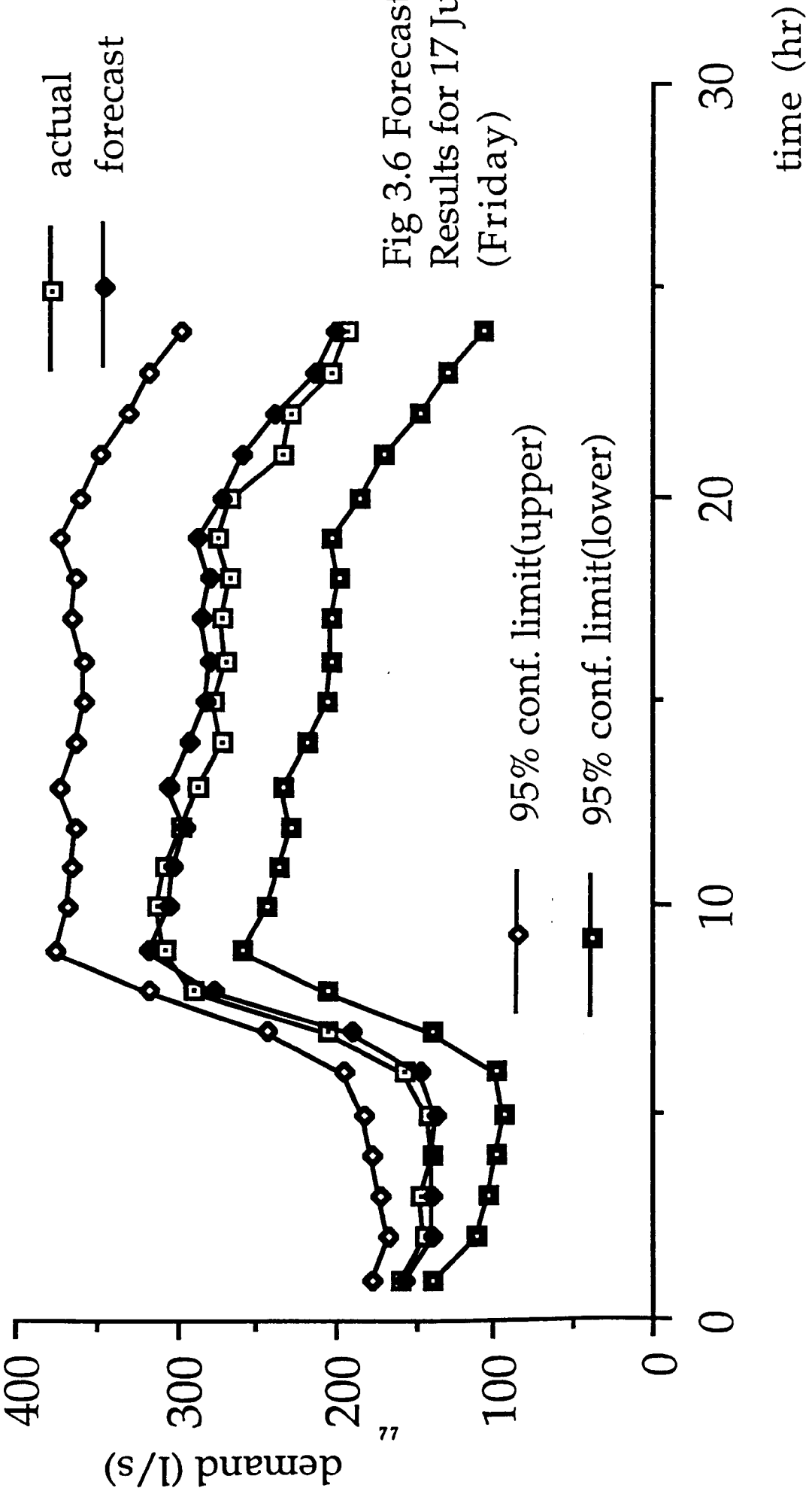


Fig 3.6 Forecasting Results for 17 June (Friday)

TABLE 3.8 COMPUTED RESULTS FOR THE MODEL ARIMA(0,1,0)(2,1,0)₂₄

Data used for estimation: 24 MAY to 1 JUN 1988, excluding Saturday and Sunday

***** RESULTS FOR PRELIMINARY ESTIMATES *****

AUTOREGRESSIVE p = 0

DIFFERENCING d = 1

MOVING AVERAGE q = 0

SEASONAL AUTOREGRESSIVE P = 2

SEASONAL DIFFERENCING D = 1

SEASONAL MOVING AVERAGE Q = 0

SEASONALITY S = 24

ARIMA MODEL PARAMETER PRELIMINARY ESTIMATES

-0.43297 -0.17020

RESIDUAL VARIANCE = 97.90221

***** RESULTS FOR LEAST SQUARES ESTIMATES*****

CONVERGENCE ACHIEVED AFTER 2 CYCLES

RESIDUAL SUM OF SQUARES IS 13030.42430 WITH 140 DEGREES OF FREEDOM

RESIDUAL VARIANCE RMS = 97.2816

LEAST SQUARES METHOD

PARAMETER ESTIMATES

	EST	SE	EST/SE
PAR (1)	-0.6170	0.0949	-6.5018
PAR (2)	-0.3321	0.1152	-2.8822
CONSTANT	-0.0851	0.4867	

TABLE 3.8 (Continued)

CORRELATION MATRIX

PAR (1)

PAR (2) 0.418

***** RESULTS FOR DIAGNOSTIC CHECKS *****

SAMPLE MEAN OF RESIDUALS = 0.0091

SAMPLE VARIANCE OF RESIDUALS = 97.3590

AUTOCORRELATIONS OF RESIDUALS

LAG ROW SE

1-12	0.08	0.11	-0.06	-0.09	-0.12	-0.09	0.00
		0.03	0.07	0.04	0.01	0.07	-0.01
13-24	0.09	-0.04	-0.04	-0.01	-0.01	0.02	-0.04
		-0.03	-0.01	-0.12	-0.05	0.06	0.08
25-36	0.09	0.13	-0.04	-0.10	0.06	0.01	-0.06
		-0.02	-0.02	0.01	0.09	0.01	-0.07
37-48	0.10	0.02	-0.02	-0.03	-0.03	0.02	0.03
		-0.04	-0.02	0.03	0.00	0.00	0.04

CHI - SQUARED TEST (PORTMANTEAU TEST)

	COMPUTED	DEG. FREE.	TABLE VALUE
Q(12)	16.3	10	21.0
Q(24)	21.9	22	36.4
Q(36)	31.2	34	49.8
Q(48)	33.0	46	66.0

TABLE 3.9 COMPARISON OF FORECASTING

PRECISIONS BETWEEN MODEL(3.43) AND MODEL(3.45)

	ACTUAL	FORECASTS BY MODEL(3.43)	DEVIATION FROM ACTUAL(%)*	FORECASTS BY MODEL(3.45)	DEVIATION FROM ACTUAL(%)*
1	150.15	147.99	1.43	148.08	0.98
2	137.07	133.36	2.70	134.40	2.67
3	122.50	133.79	-9.22	128.41	-4.80
4	118.25	129.23	-9.28	125.66	-6.26
5	120.25	131.67	-9.50	123.67	-2.85
6	135.25	144.03	-6.49	138.12	-2.12
7	184.75	191.36	-3.58	185.28	-0.29
8	266.78	271.05	-1.60	260.77	2.25
9	301.74	298.74	0.99	302.01	-0.09
10	316.75	299.99	5.29	304.31	3.92
11	316.00	294.67	6.75	292.16	7.54
12	305.00	284.55	6.71	285.52	6.39
13	301.13	280.92	6.71	283.54	5.84
14	279.75	266.67	4.67	270.62	3.26
15	281.25	268.12	4.67	268.10	4.67
16	273.70	262.19	4.21	262.97	3.92
17	288.00	264.82	8.05	273.65	4.98
18	268.94	259.22	3.61	259.73	3.42
19	272.79	267.62	1.89	267.94	1.78
20	263.80	254.88	3.38	262.80	0.38
21	246.50	230.09	6.66	237.48	3.66
22	232.25	219.63	5.34	225.46	2.92
23	230.00	197.95	13.94	210.42	8.51
24	189.25	184.37	2.58	185.17	2.16
			Ave. 5.4 ⁺		Ave. 3.6 ⁺

* see the footnote of Table 3.6.

+ see the footnote of Table 3.6.

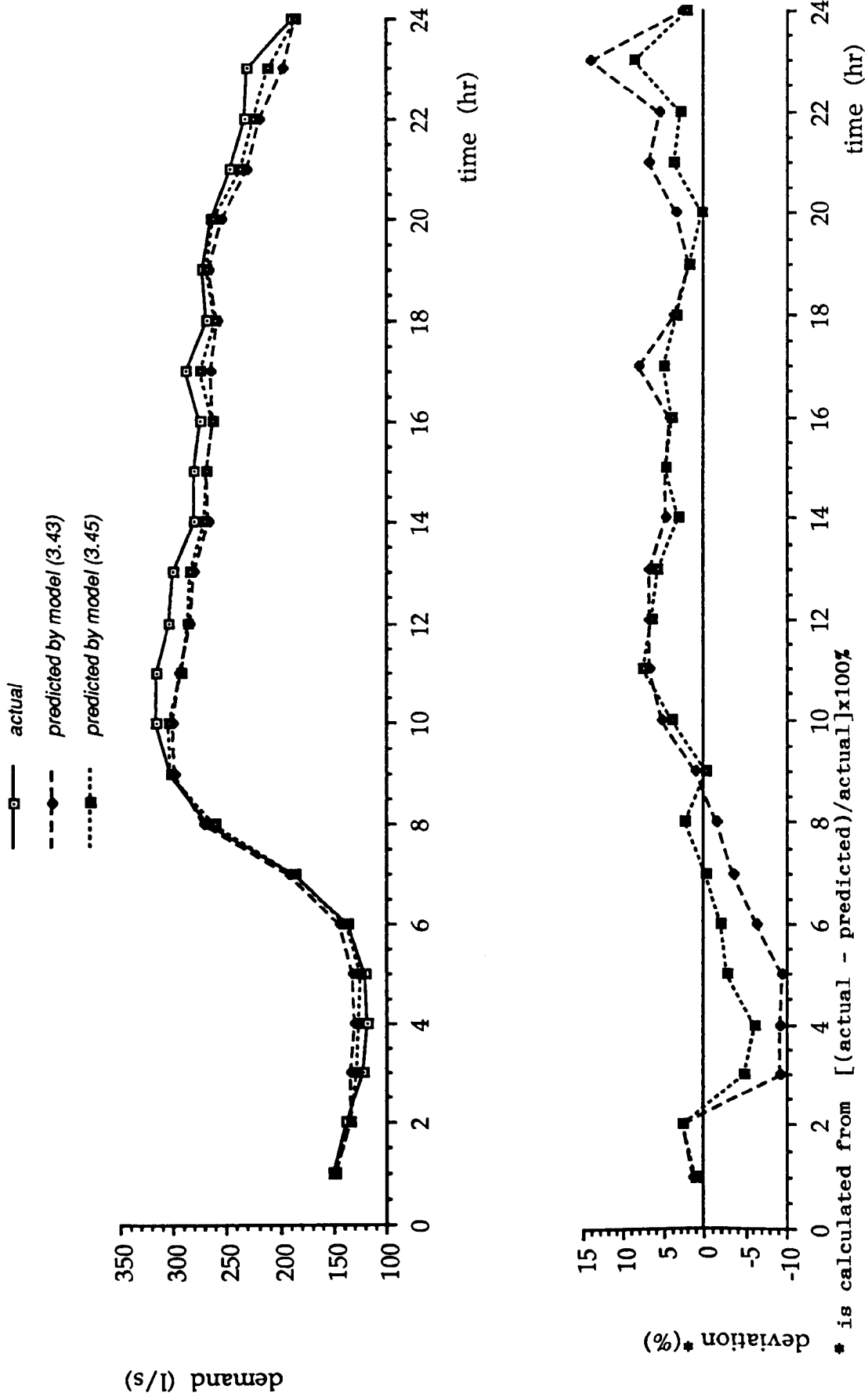


FIG 3.7 A COMPARISON BETWEEN TWO MODELS

3.5 COMPARISON WITH THE PREDICTIONS FROM THE EXISTING DEMAND PREDICTION PROGRAM GIDAP

3.5.1 Introduction to GIDAP

GIDAP is a Graphical Interactive Demand Analysis and Prediction Package developed in the Water Control Centre at Leicester Polytechnic, England. The prediction part of the program provides forecasts of daily or weekly demand profiles, while the analysis part enables data attributes, essential to prediction, to be determined. Predicted profiles can be replaced by pre-defined special profiles for unusual days (such as bank-holidays) and full account is taken of British-Summer-Time clock changes. GIDAP is fully configurable for a variety of telemetry types and settings and includes a facility to calibrate telemetry demand data. An extensive demand data and parameter display and modification facility is provided. (TENNANT et al, 1986; TENNANT, 1987).

The basic process of the prediction technique is outlined in Fig 3.8. Each step in the data processing is described below:

1) The Screening Process

The screening process detects and removes coarse measurement and transmission errors in the raw demand data. Upper and lower thresholds are applied to each data value to detect excessively high and low demands respectively. First and second difference thresholds are then applied to remaining data values to detect excessive fluctuations in demand. Rejected values are replaced by predicted values if they are available. If these are not available, linear

interpolation between non-rejected values is used to find suitable replacements.

For measured data, $X(t)$, with a sampling interval of Δt , and consecutive values of $X(i)$, $X(j)$ and $X(k)$, the second difference will be:

$$d^2X(t)/dt^2 \simeq [X(k) - 2X(j) + X(i)] / \Delta t^2 \quad (3.46)$$

The second difference of a demand value is a measure of the 'peakiness' of that value, or the acceleration of demand represented by that value. Whereas the first difference of a demand value defines the rate of change of demand represented by that value.

Analysis of sets of measured data will establish realistic thresholds. In the present application maximum, minimum, and first and second difference thresholds are evaluated and used.

2) The Smoothing Process

The smoothing process is designed to detect and reject fine random errors and effects in the consumption and measurement process. Evaluation and recombination of significant components will give a smoothed estimate of the measured data as follows:

$$\begin{aligned}
X(t, NH) = a_0 + \sum_{n=1}^{NH} a_n \cos(2 \pi n t / T) \\
+ \sum_{n=1}^{NH} b_n \sin(2 \pi n t / T)
\end{aligned}
\tag{3.47}$$

where:

t = the sample times.

T = 24 hour or 7 day period

a, b = component amplitudes determined by Fast-Fourier-Transform (FFT) analysis.

n = integer harmonic multiples of the fundamental frequency.

NH = highest significant harmonic, determined by statistical analysis.

3) The Trend Estimation Process

The triple exponential smoothing technique is used to provide trend estimates which define current demand patterns. These estimates are extrapolated to form predicted demand profiles. The estimates are continually updated according to discrepancies between the predicted profiles and the corresponding smoothed actual profiles.

The vector of prediction errors at the current daily or weekly period is defined by:

$$\underline{e}_t = \hat{\underline{X}}_{t-1}(1) - \underline{X}_t \quad (3.48)$$

where:

\underline{X}_t = vector of data sample values at the current period.

$\hat{\underline{X}}_{t-1}(1)$ = is the 1 period (step) ahead forecast at the previous period.

When the current demand profile becomes available the error vector can be used to correct trend estimates according to:

$$\hat{\underline{a}}_t = \underline{X}_t + (1 - sm)^3 \underline{e}_t \quad (3.49)$$

$$\hat{\underline{b}}_t = \hat{\underline{b}}_{t-1} + \hat{\underline{c}}_{t-1} - 1.5 sm^2(2 - sm) \underline{e}_t \quad (3.50)$$

$$\hat{\underline{c}}_t = \hat{\underline{c}}_{t-1} - sm^3 \underline{e}_t \quad (3.51)$$

where:

sm = a smoothing parameter with a typical value of 0.1.

$\hat{\underline{a}}, \hat{\underline{b}}, \hat{\underline{c}}$ = estimates of position, velocity, and acceleration trend components at periods t and t-1. Initialization procedures set $\hat{\underline{a}}_{t-1}$ to \underline{X}_{t-1} and $\hat{\underline{b}}_{t-1}, \hat{\underline{c}}_{t-1}$ to zero.

The 1 period ahead prediction will then be given by:

$$\hat{\underline{X}}_t(1) = \hat{\underline{a}}_t + \hat{\underline{b}}_t + 0.5 \hat{\underline{c}}_t \quad (3.52)$$

Prediction can be produced on a daily basis where, on Monday, Tuesday's demand profile is predicted, then on Tuesday, Wednesday's demand profile is predicted and so on. Days of the week are placed into categories such that days in the same category have statistically similar demand profiles. The trend estimates for one day in a category can then be used to predict the profile for the next day in the same category. Consequently one set of trend estimates is

maintained for each category. These categories are determined by the appropriate data analysis.

More details can be found in (COULBECK et al, 1985 ; TENNANT et al, 1986; TENNANT, 1987).

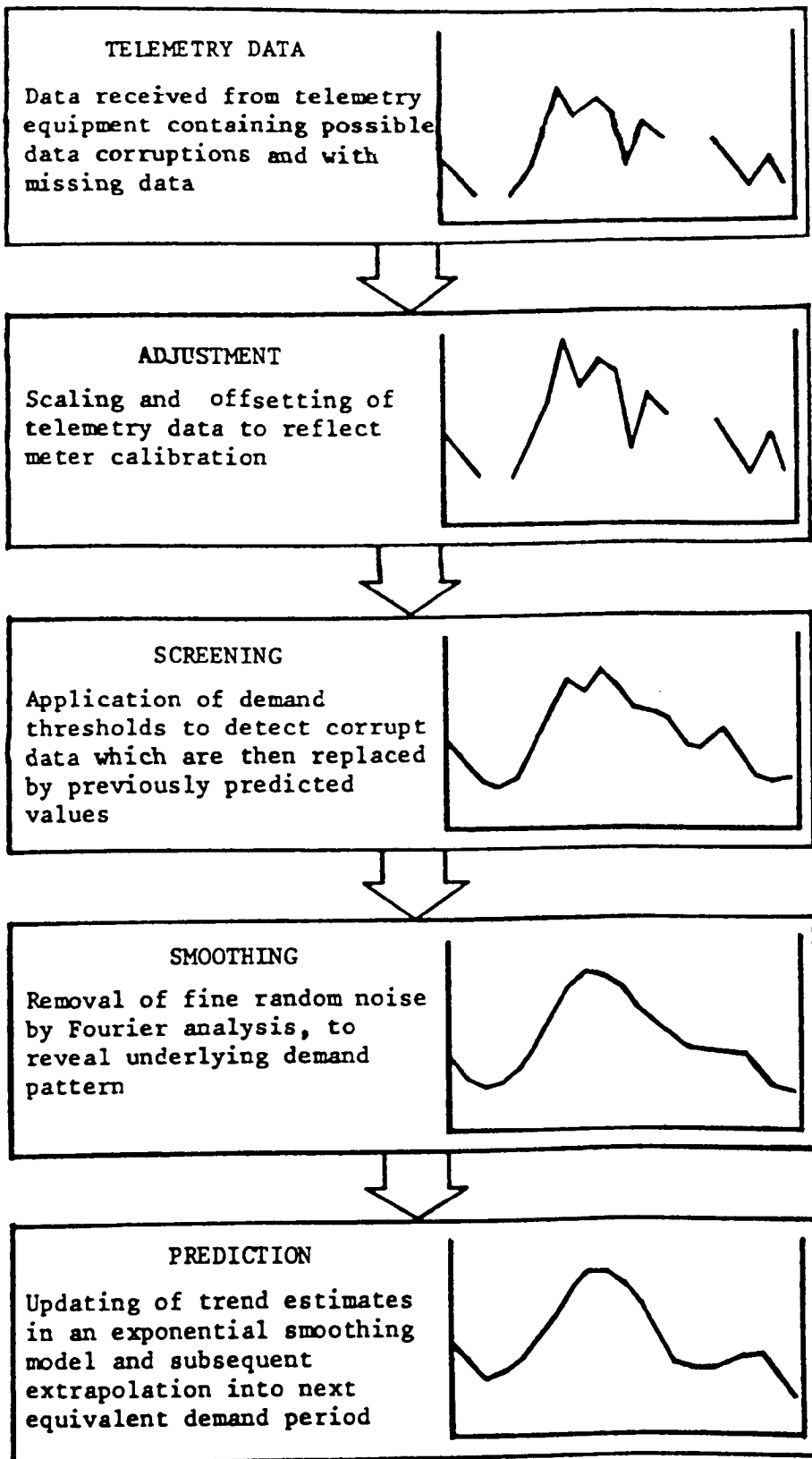


Fig 3.8 the Prediction Process of GIDAP

3.5.2 Discussions and Comparisons

Exponential smoothing methods have a number of advantageous features which include: recursive formulation, trend following, error correction, and minimum data requirements (COULBECK et al, 1985). The screening process used in GIDAP is a good method for the removal of measurement noise. However, once data has been smoothed, it is evident that the original random process of demand is considered to be deterministic. From the theoretical point of view, exponential smoothing is only correct for a class of ARIMA models (VANDAELE, 1983). This is demonstrated as follows:

The IMA(1,1) or ARIMA(0,1,1)(0,0,0) with zero mean μ can be represented as,

$$Z_t - Z_{t-1} = a_t - \theta_1 a_{t-1} \quad (3.53)$$

Similarly, this model can be rewritten for time period $n+1$ and earlier as,

$$\begin{aligned} Z_{n+1} - Z_n &= a_{n+1} - \theta_1 a_n \\ Z_n - Z_{n-1} &= a_n - \theta_1 a_{n-1} \\ Z_{n-1} - Z_{n-2} &= a_{n-1} - \theta_1 a_{n-2} \\ &\vdots \\ &\vdots \end{aligned} \quad (3.54)$$

On multiplying both sides of the second equation in (3.54) by θ_1 , both sides of the third equation by θ_1^2 , and so on, and then adding

all these transformed equations together, we can solve for Z_{n+1} to obtain:

$$Z_{n+1} = a_{n+1} + (1 - \theta_1)Z_n + \theta_1(1 - \theta_1)Z_{n-1} + \theta_1^2(1 - \theta_1)Z_{n-2} + \theta_1^3(1 - \theta_1)Z_{n-3} + \dots \quad (3.55)$$

The minimum mean squared error forecast for Z_{n+1} is obtained from eqn(3.55) after replacing the future error a_{n+1} by its mean value 0, i.e.,

$$Z_n(1) = (1 - \theta_1)Z_n + \theta_1(1 - \theta_1)Z_{n-1} + \theta_1^2(1 - \theta_1)Z_{n-2} + \dots \quad (3.56)$$

Eqn(3.56) represents the exponential weighted moving average process, from which the various exponential smoothing techniques originate. This equation shows that for an IMA(1,1) process, recent observations receive larger weights than more distant observations in the past. In other word, Z_n, Z_{n-1}, \dots are weighted in an esponentially decreasing manner, hence the name of exponential smoothing.

The above proves that exponential smoothing is only correct for an IMA(1,1) model, which is only one of many ARIMA models. Since the natural processes that can be represented by IMA(1,1) are limited, the applicability of exponential smoothing techniques are consequently also limited. The multiplicative time series model, however, can represent any class of demand, although the identification of the most appropriate model for a particular demand series is rather complex and the data requirement for parameter estimation is comparatively large,

which brings some difficulties for on-line implementation.

The author's limited investigations have revealed that accurate predictions can be achieved using the exponential smoothing method used in GIDAP. This is especially true when the demand series is relatively stationary and regular. The accuracy of such forecasts are compared with those from the model of eqn(3.43) in Table 3.10 and Fig 3.9.

However, it is evident that the response of GIDAP to fluctuations in demand is somewhat slower than for the time series model. This manifests itself as larger prediction errors from GIDAP, at those times when large changes in demand occur.

Table 3.11 and Fig 3.10 compare typical results for demands on 10 June, 1989. Telemetry data for a few days prior to the 10 June, 1989 were unavailable. The forecasts for the 10 June, 1989 were then based on their corresponding forecasts for the previous days.

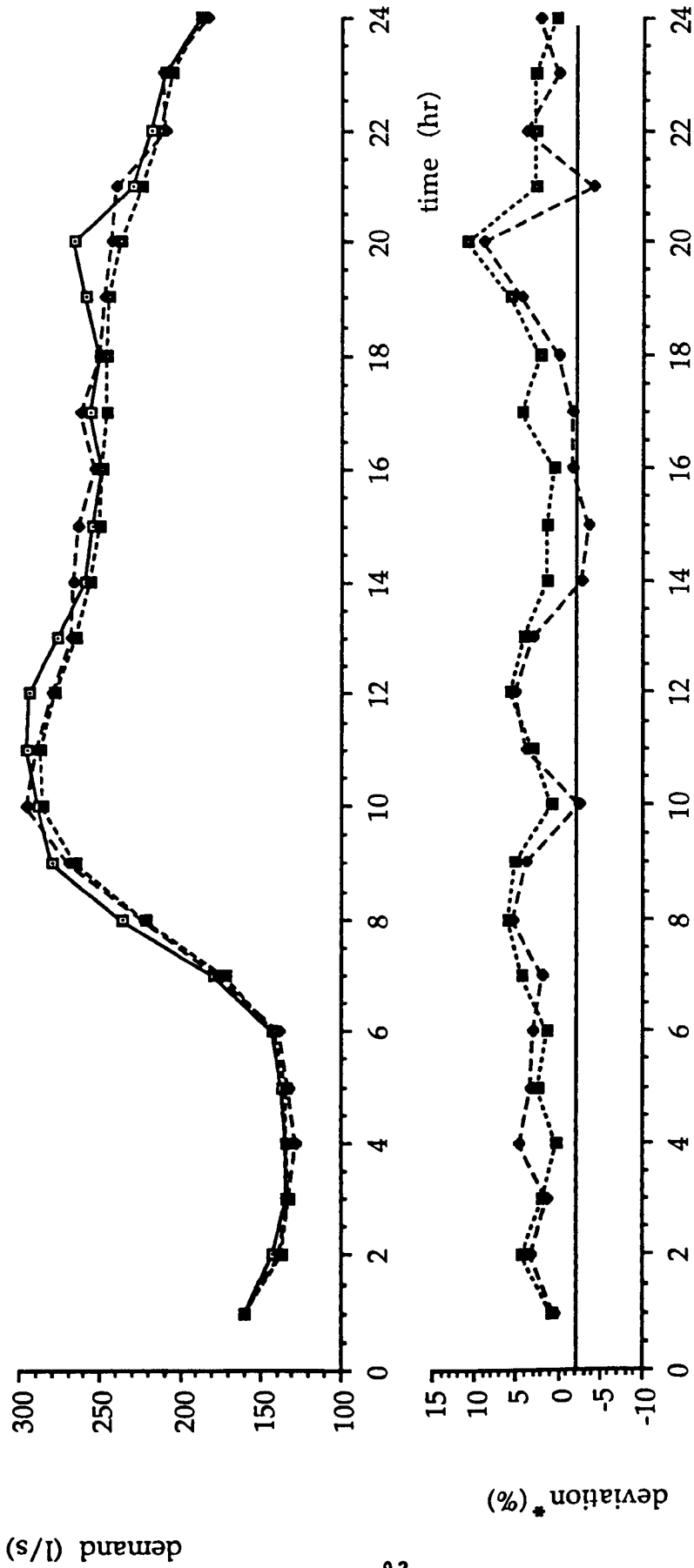
While comparing the performance of the two models, it should be remembered that the time series models are forecasting up to 24 steps ahead. The exponential model used in GIDAP, on the other hand, is performing 24 one-step ahead forecasts. Further the time series method employs only one complex model, while GIDAP requires 24 separate but simple models; and these latter do not take into account interaction between successive hourly demands. In conclusion, it might be profitable to incorporate the time series methodology within GIDAP.

TABLE 3.10 A COMPARISON OF FORECASTS FROM A TIME SERIES MODEL
AN EXPONENTIAL SMOOTHING MODEL FOR 3 JUNE 1988

TIME	ACTUAL	FORECASTS BY ARIMA	DEVIATION FROM ACTUAL (%)*	FORECASTS BY GIDAP	DEVIATION FROM ACTUAL (%)*
1	160.38	159.74	0.40	159.39	0.62
2	142.75	138.11	3.25	136.77	4.19
3	134.00	132.14	1.39	131.39	1.95
4	134.25	128.29	4.44	133.91	0.25
5	136.75	132.47	3.13	133.54	2.35
6	142.50	138.23	3.00	140.62	1.32
7	179.50	176.26	1.80	171.82	4.28
8	235.75	223.33	5.27	221.49	6.05
9	279.50	269.40	3.61	265.01	5.18
10	288.50	295.60	-2.46	286.12	0.82
11	296.00	289.85	3.77	287.34	2.93
12	294.50	279.36	5.14	278.01	5.60
13	276.50	268.40	2.93	265.54	3.96
14	259.00	266.48	2.89	255.78	1.24
15	254.25	263.28	-3.55	250.99	1.28
16	249.50	253.90	-1.76	248.42	0.43
17	256.88	261.50	-1.80	246.14	4.18
18	250.63	251.03	-0.16	245.60	2.01
19	259.75	246.95	4.39	244.72	5.79
20	266.50	242.66	8.95	237.29	10.96
21	230.00	240.02	-4.36	223.99	2.61
22	218.63	210.20	3.86	213.11	2.52
23	210.00	210.43	-0.21	204.69	2.53
24	187.63	183.05	2.11	187.44	0.10
			ave. 3.10 ⁺		ave. 3.00 ⁺

* see the footnote of Table 3.6.

+ see the footnote of Table 3.6.



* see the footnote of FIG 3.7

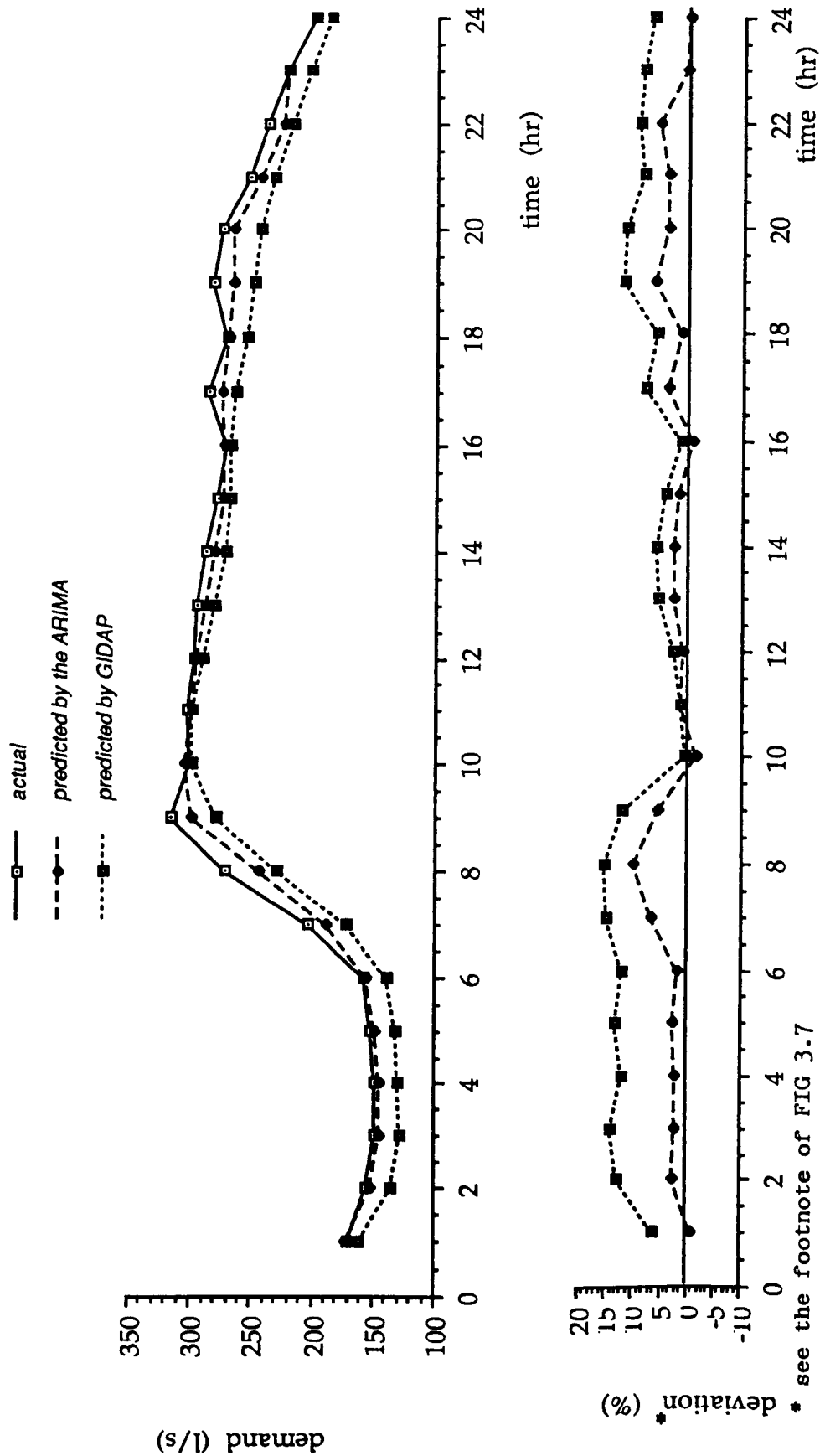
FIG 3.9 A COMPARISON OF FORECASTS FOR 3 JUNE 1988

TABLE 3.11 A COMPARISON OF FORECASTS FROM A TIME SERIES MODEL
AN EXPONENTIAL SMOOTHING MODEL FOR 10 JUNE 1988

TIME	ACTUAL	FORECASTS BY ARIMA	DEVIATION FROM ACTUAL (%)*	FORECASTS BY GIDAP	DEVIATION FROM ACTUAL (%)*
1	170.04	171.81	1.04	159.85	5.99
2	154.88	151.38	2.26	135.52	12.50
3	148.00	144.95	2.06	127.37	13.94
4	147.25	144.25	2.03	130.20	11.58
5	150.63	147.21	2.27	131.18	13.21
6	157.25	154.68	1.64	138.69	11.93
7	202.25	189.00	6.55	172.55	14.68
8	270.25	243.81	9.78	229.39	15.12
9	315.50	299.19	5.17	278.34	11.78
10	299.50	304.46	-1.66	298.79	0.24
11	301.50	298.73	0.92	297.98	1.17
12	297.25	295.66	0.53	290.14	2.39
13	295.00	287.78	2.45	279.98	5.10
14	286.75	280.40	2.21	270.53	5.66
15	278.00	273.60	1.58	267.08	3.92
16	270.67	272.76	-0.77	267.46	1.18
17	285.25	274.89	3.63	263.58	7.59
18	270.75	268.21	0.94	255.24	5.73
19	282.25	264.73	6.21	248.76	11.87
20	274.38	264.82	3.49	243.13	11.39
21	252.00	242.71	3.69	232.15	7.88
22	238.25	225.70	5.27	217.41	8.74
23	222.00	221.27	0.33	203.67	8.26
24	199.00	199.29	-0.14	186.16	6.45
			ave. 2.80 ⁺		ave. 8.30 ⁺

* see the footnote of Table 3.6.

+ see the footnote of Table 3.6.



* see the footnote of FIG 3.7

FIG 3.10 A COMPARISON OF FORECASTS FOR 10 JUNE 1988

3.6 CONCLUSION

This chapter has described in detail how time series analysis is applied to water demand forecasting. Among the general class of multiplicative ARIMA models, AR (Autoregressive) models have to their advantage, the fact that model parameters are estimated using simple linear least-squares. In contrast, the parameters of a general ARIMA model are estimated using more complex, non-linear least-squares. Further, the model identification for AR models is also quite simple. The forecasting results show that the parallel forecasting method is better than the series forecasting method under normal demand conditions but the latter is more suitable under abnormal conditions. Limited studies demonstrate that forecasting precision is raised by combining the forecasting results from the two methods, although the problem of how to best combine them is left open and needs further study. The application results have shown that AR models are quite suitable for the demand studied. However, the applicability of AR models could be limited to the demand patterns which are similar to the demand pattern studied, since AR models are only a class of general ARIMA models.

Due to the availability of various routines for time series analysis in the commercial NAG Fortran Library, the general Box-Jenkins approach to multiplicative ARIMA models has been employed. Comprehensive computer programs, which cover model identification, parameter estimation, diagnostic checking and demand forecasting, have been developed and applied to demand forecasting for the Woverhampton water system. Numerical results are presented in detail.

Typical forecasting results show that the general Box-Jenkins approach to general ARIMA models is satisfactory for water demand forecasting (average deviation 3.1%, refer to Table 3.6 and Fig. 3.5). It has also been verified through comparative tests that weekdays and weekends should be placed in different categories and forecasts should be made within each category.

In terms of theory and performance, comparisons have been made between a time series model and the exponential model employed in the demand analysis and prediction program, GIDAP. For the former the corresponding data requirement is great and model building is somewhat tentative and could vary considerably from demand to demand. With regards to the latter the data requirement for parameter estimation is lower. The prediction model is easier to build and the process straightforward. However, exponential smoothing is only strictly correct for the process that is described equivalently by an IMA(1,1) time series model and therefore, the applicability and effectiveness of such a method to water demand prediction could be limited. The comparative tests indicate that the forecasting precision of time series analysis and exponential smoothing are comparable when the demand series is relatively stationary and regular, the average deviation of the former being 3.1% and of the latter being 3.0% (refer to Table 3.10 and Fig.3.9). However, for non-stationary and irregular demands, the time series analysis model is able to follow fluctuations in demand more closely. This is confirmed by an average deviation of 2.8% for time series compared to 8.3% for exponential smoothing (refer to Table 3.11 and Fig.3.10). Consider this in the light of the fact that the time series model is forecasting up to 24 steps ahead. The

conclusion is that the time series analysis method could usefully be incorporated into GIDAP.

It should be evident that the various methods described could be easily adapted to the requirements for long-term demand forecasting as well as other short-term demand forecasting.

The integration of the various demand forecasting methods will result in a comprehensive computer program which provides facilities to enable the user to switch to different methods. For the demand forecasting of a particular system, it is worth first employing AR models, since the simplicity of parameter estimation and convenience of model-order selection make this method practical and friendly for non-expert users. If the forecasting results from the derived AR model are not satisfactory in terms of practical requirements, then the Box-Jenkins approach to general ARIMA models will have to be employed, where further work will be required in order for the computer program to choose the most appropriate model automatically. After a well-defined model has been applied for a long time, satisfactory forecasting results may no longer be obtainable. This could indicate that recent demand data will be required as input to the computer program either to update the parameters of the model or to build a new model, which will be determined by the computer program. The latter may happen when the demand pattern has changed significantly.

4.1 INTRODUCTION

Generally, a water supply and distribution network is composed of hundreds or even thousands of pipes and other elements and thus is a very complicated system. Moreover, time-varying consumers demands occur at different points throughout the network, few of which are measured hourly or daily. As discussed in (CHEN, 1986; DEMOYER and HORWITZ, 1975a, 1975b), it is impractical to evaluate optimal controls using a detailed model of the actual network. Therefore, certain reasonable and effective simplifications are essential.

As discussed in Chapter 2, most conventional distribution system models use what is called a microscopic approach by incorporating in the model all components down to the smallest pipe diameter included in a skeletonized network. Although some sort of aggregation work, involving the cancellation of pipes of small diameters, replacement of parallel or serial pipes by 'equivalents' and so on, can be done, there are still a large number of pipes and nodes in the skeletonized network and a single network 'balance' requires the iterative calculation of all pipe flows and nodal pressures. This is very time consuming and even infeasible for optimal control purpose (CHEN, 1985). Owing to this difficulty, some complicated networks have been

simplified by some researchers by eliminating a large part of the network to leave only a few pipes. In this case the characteristics of the original network might not be properly reflected and customers service requirements as well as some important components in the network cannot be taken into account directly.

Owing to the practical requirements, this chapter presents two independent methods of network modelling and simplification. The first of these methods extends the macroscopic model, derived by DeMoyer et al (DEMOYER and HORWITZ, 1975a, 1975b), to cater for more general water supply and distribution systems. The second method has been developed by the author based on conventional theory of water network analysis.

4.2 SUMMARY OF THE MACROSCOPIC MODEL

DeMoyer et al developed what is called a macroscopic model in that it deals only with major heads and flows associated with pumps and tanks (or reservoirs). In such a model, variables are related by empirical equations derived from statistical analysis of operating data (DEMOYER and HORWITZ, 1975a, 1975b).

4.2.1 Pumping Station Relationships

$$H_p(i) = C_p(i,k,1) + C_p(i,k,2) \times Q(i)^{1.85} \quad (4.1)$$

where:

$H_p(i)$ = head increase across station i , pump combination k .
(e.g., m)

$Q(i)$ = pump flow at station i . (e.g., m³/h)

C_p = pump constant array.

For accuracy, the constants C_p are determined from multiple regression or stepwise regression analysis of operational data (DEMOYER and HORWITZ, 1975a, 1975b; KENNEDY, 1986; OSTLE, 1974).

$$H_s(i) = C_s(i,1) + C_s(i,2) \times Q_d^{1.85} + C_s(i,3) \times Q(i)^{1.85} \quad (4.2)$$

where:

$H_s(i)$ = suction head at station i . (e.g., m)

Q_d = total demand flow. (e.g., m^3/h)

C_s = suction head constants.

The constants C_s can also be determined from multiple regression or stepwise regression analysis of operating data. Then the pumping station discharge head, $H_d(i)$, is a combination of eqn(4.1) and eqn(4.2):

$$H_d(i) = H_s(i) + H_p(i) \quad (4.3)$$

4.2.2 Internal Network Nodal Pressure Relationship

$$H(h) = C_n(h,1) + C_n(h,2) \times Q_d^{1.85} + \sum_{i=1}^I C_n(h,i+2) \times H_d(i) + \sum_{j=1}^J C_n(h,j+I+2) \times H_t(j) \quad (4.4)$$

where:

I = number of pumping stations in the system.

J = number of tanks in the system.

$H(h)$ = internal network pressure at node h.(e.g., m)

$H_t(j)$ =head of tank j.(e.g., m)

C_n = network pressure constant array determined by
regression.

For other related empirical relationships and more details see
(DEMOYER and HORWITZ, 1975a, 1975b).

The macroscopic model can be, for the purpose of control, the equivalent of a full network model. It is convenient and feasible for on-line implementation, and is also effective for those systems having only partial information available, because of the stochastic nature of water consumption and the constant change in hydraulic network characteristics.

It should be noted that such macroscopic models are based on the assumption that the actual distributed demands throughout the network are proportional to total demand (i.e., proportional loading); the accuracy of the models will decrease in a district having large non-proportional loads. Unfortunately, a lot of urban water distribution networks do not satisfy this assumption. In many water supply and distribution systems, it is difficult to distinguish residential districts from industrial districts thus the conditions for proportional loading may not be satisfied. Direct applications of the macroscopic models to such systems might not be satisfactory. However, it was found by examining operating data that the load patterns of many systems do not vary much over short time horizons, and within these horizons the patterns can be considered to

approximate the assumption of proportional loading. Therefore, the author has developed a set of piecewise macroscopic models, i.e., the constants in the models are determined from operating data over several periods during one day (CHEN, 1988b), which are discussed in the following section.

4.3 PIECEWISE MACROSCOPIC MODEL

From investigations of practical water supply and distribution systems in China and using the macroscopic model as a basis, the piecewise macroscopic model was developed as follows:

4.3.1 Pumping Station Relationship

$$\begin{aligned}
 H_d^{k(i)} = & C_d^{k(i,1)} + C_d^{k(i,2)} \times Q_d^\alpha \\
 & + \sum_{j=1}^n C_d^{k(i,j+2)} \times Q(j)^\alpha \\
 & + \sum_{\substack{j=1 \\ j \neq i}}^n C_d^{k(i,j+2+n)} \times Q(j) \times Q(i), \quad (4.5)
 \end{aligned}$$

where:

n = number of pumping stations.

$H_d^{k(i)}$ = discharge head of station i in period k . (e.g., m)

Q_d = total demand flow. (e.g., m^3/h)

$Q(i), Q(j)$ = discharge flow at station i, j , respectively. (e.g., m)

α = 1.85 to 2.0.

C_d^k = discharge head constant array in period k ,
 which is to be determined from regression analysis of
 operating data in period k .

The third item of the R.H.S. of eqn(4.5), which is not present in the corresponding relationship of the original macroscopic model, was found out to be useful in achieving better results for the practical system studied.

4.3.2 Internal Network Nodal Pressure Relationship

$$H^k(j) = B^k(j,1) + \sum_{i=1}^n B^k(j,i+1) \times Q^a(i) + B^k(j,n+2) \times Q_d^a \quad (4.6)$$

(when there is no reservoir or no significant storage capacity in the system)

where:

$H^k(j)$ = internal network pressure at node j in period k . (e.g., m)

B^k = network pressure constant array in period k .

4.3.3 Model Validations

The model validations were carried out using data from the Pudong district in Shanghai. The water supply system in this district contains three water purification plants (each plant with a pumping station, which discharges the purified water into the network) and five internal pressure monitoring nodes, at the locations shown in Fig 4.1. There are no water storage reservoirs in the network.

For example, for station 3, in period 1 (from 0.00 hour to 6.00 hours

every day). By multiple regression, eqn(4.5) is:

$$\begin{aligned} H_d^1(3) = & 32.7055 + 3.3245 \times Q_d^2 - 7.4109 \times Q^2(1) \\ & -74.9065 \times Q^2(2) - 148.5306 \times Q^2(3) \\ & + 20.3285 \times Q(1) \times Q(3) \\ & + 142.0468 \times Q(2) \times Q(3) \end{aligned} \quad (4.7)$$

where: correlation coefficient $\rho = 0.802$, and
standard deviation $s_y = 1.737$

and by stepwise regression, eqn(4.5) becomes:

$$H_d^1(3) = 31.1768 - 21.6010 \times Q^2(3) \quad (4.8)$$

where: correlation coefficient $\rho = 0.754$, and
standard deviation $s_y = 1.010$

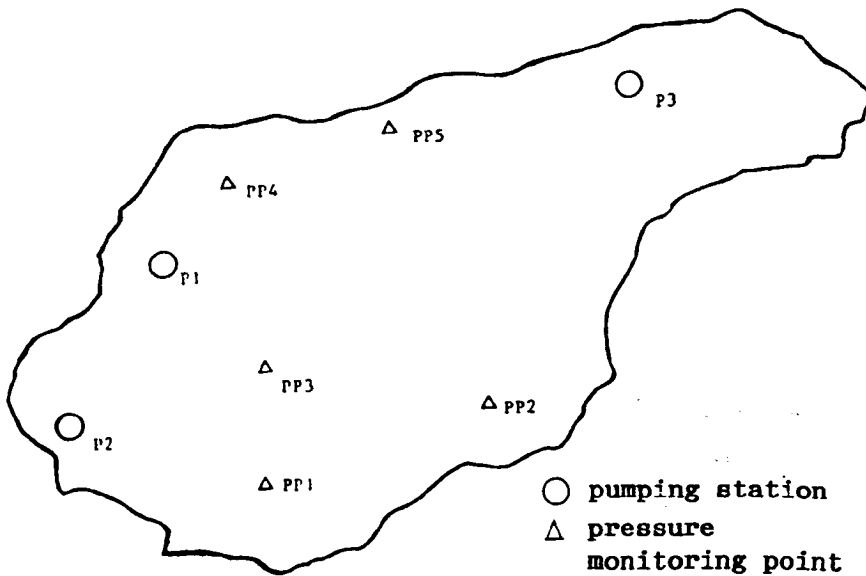


Fig 4.1 Layout of the Pudong System, in Shanghai

Parts of the simulation results are listed in Table 4.1, which shows that both eqn(4.7) and eqn(4.8) are acceptable. This is true for the simulation of other pumping stations.

As for the simulation of internal network nodal pressures. By multiple regression, the equation for point 1 is:

$$H^1 = 20.8327 - 1.2304 \times Q_d^2 + 2.1097 \times Q^2(1) + 10.9899 \times Q^2(2) + 2.0821 \times Q^2(3) \quad (4.9)$$

where: correlation coefficient $\rho = 0.697$, and
standard deviation $s_y = 1.835$.

and by stepwise regression, it is:

$$H^1(1) = 23.3292 - 0.9906 \times Q_d^2 + 1.8435 \times Q^2(1) + 9.6945 \times Q^2(2) \quad (4.10)$$

where: correlation coefficient $\rho = 0.603$, and
standard deviation $s_y = 1.432$.

also, the simulation results of both are quite accurate.

Table 4.1 SIMULATION RESULTS OF PIECEWISE MACROSCOPIC MODEL

Q_d ($10^3 \text{xm}^3/\text{hr}$)	Q(1)	Q(2)	Q(3)	$H_d(3)$ actual (m)	Multiple Regression		Stepwise Regression	
					$H_d(3)$ pre- dicted (m)	ERR.* (m)	$H_d(3)$ pre- dicted (m)	ERR.* (m)
1	3.9820	2.5200	0.6100	31.00	32.7875	-1.7875	32.0580	-1.0580
2	3.9920	2.5200	0.6200	32.00	32.9481	-0.9481	31.9386	0.0614
3	2.9989	1.5580	0.5889	30.00	28.6511	1.3489	29.0073	0.9927
4	3.0539	1.7500	0.6179	28.00	27.1947	0.8653	27.1769	0.8231
5	3.2300	1.7500	0.8849	29.00	29.4064	-0.4064	29.5710	-0.5710
6	4.8210	3.2480	0.8969	34.00	34.4129	-0.4129	34.0450	-0.0450
7	3.0300	1.7360	0.6179	28.00	27.6081	0.3919	28.1146	-0.1146
8	3.2500	1.7360	0.8969	28.00	29.0612	-1.0612	29.1682	-1.1682
9	3.2900	1.7360	0.6200	31.00	27.9915	3.0085	29.1183	1.8817
10	4.5220	3.0480	0.8849	34.00	32.1778	1.8222	34.0993	-0.0993
						ave 1.2053 ⁺		ave 0.6815 ⁺

* is calculated from (actual head - predicted head).
⁺ is calculated from $\sum |ERR| / 10$.

4.3.4 Division of a Day into Proportional Loading Periods

The periods when the assumption of proportional loading is approximately satisfied may vary with different water supply and distribution systems. From the author's studies, the division into periods of proportional loading depends mainly on the total demand. In other words, a period in which the total demand does not change significantly can be considered to be a proportional loading period. This is further demonstrated through the following numerical results.

For the period from 0.00 hr to 6.00 hrs, the multiple regression model of station 2 in the Pudong system is,

$$\begin{aligned} H_d(2) = & 20.8929 + 0.7209 \times Q_d^2 - 1.9072 \times Q^2(1) \\ & - 77.0007 \times Q^2(2) - 121.7429 \times Q^2(3) \\ & + 4.9584 \times Q(1) \times Q(2) + 196.3082 \times Q(2) \times Q(3) \end{aligned} \quad (4.10a)$$

where: correlation coefficient $\rho = 0.87$, and
standard deviation $s_y = 1.443$.

Whereas for the period from 0.00 hr to 12.00 hrs, the multiple regression model of the same station is,

$$\begin{aligned} H_d(2) = & 28.4666 + 15.0469 \times Q_d^2 - 12.6832 \times Q^2(1) \\ & + 41.8231 \times Q^2(2) - 45.4219 \times Q^2(3) \\ & - 63.3089 \times Q(1) \times Q(2) - 63.7255 \times Q(2) \times Q(3) \end{aligned} \quad (4.10b)$$

where: correlation coefficient $\rho = 0.59$, and

standard deviation $s_y = 3.259$.

It shows that eqn(4.10a) is more suitable than eqn(4.10b) in terms of the correlation coefficient ρ and the standard deviation s_y . More accurate simulation results are also achieved from eqn(4.10a) than those from eqn(4.10b). This is mainly because the total demand within the sub-period from 1.00 hr to 6.00 hrs is relatively stable, and the total demand within the sub-period from 6.00 hrs to 12.00 hrs is also relatively stable; but the total demand changes significantly from the first sub-period to the second sub-period. Therefore the period from 0.00 hr to 12.00 hrs should be divided into two different proportional loading sub-periods and regression models should be built separately within each sub-period.

4.4 EQUIVALENT NETWORK MODEL

4.4.1 Theoretical Analysis and Deduction

From the above discussions, it will be clear that both the macroscopic model and the piecewise macroscopic model were developed based on empirical relationships obtained from statistical analysis of operating data. Consequently the derived parameters have no clear and definite physical meaning (LIU and DUAN, 1986). This does not sound strong theoretically, and large errors may occur, especially in those water distribution systems that do not strictly satisfy the condition of proportional loading.

To avoid these disadvantages, the concept of an equivalent network model is developed by the author hereafter.

The essence of equivalent network modelling is to deal only with major components in a network (such as reservoir nodes, pressure monitoring nodes and nodes which link control elements) and to inter-connect them using fictitious pipes in order to construct a simplified equivalent network from the original detailed network. The constructed equivalent network should be able to replace the main aspects of the original network for simulation or optimization purpose.

To further explain the concepts of the equivalent network modelling procedure, take a simple example as shown in Fig 4.2 (CHEN, 1988b). For the original network, drawn with solid lines, from Chapter 2, the well-known nodal equation will be:

$$\sum_{j \in N(i)} q_{ij} = u_j - y_i \quad i \in N \quad (4.11)$$

where:

N = total number of nodes in the network.

$N(i)$ = subset of j connections to node i .

q_{ij} = pipe flow from node i to j (positive flow).(e.g., l/s)

u_i = pump flow into node i .(e.g., l/s)

y_i = consumption flow out of node i .(e.g., l/s)

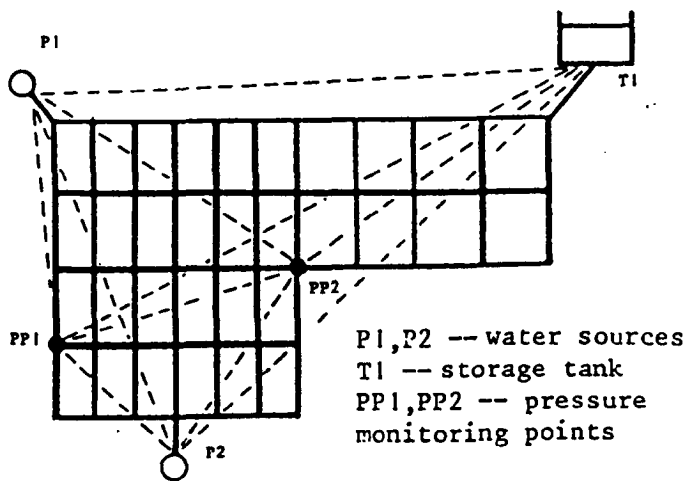


Fig 4.2 An Example Water Supply Network

In Fig 4.2 all the dashed lines which link the main components of the water supply distribution system will constitute its initial equivalent network. The equivalent network will be determined by the statistical analysis of correlation, ie, if two variables of the system are correlated, a dashed line (a fictitious pipe) will connect

these two variables, otherwise there will be no dashed line between them.

The procedure of correlation analysis is stated as follows:

Since a water supply distribution system operates somewhat cyclically, nodal pressures will fluctuate around their average values. In the long term, nodal pressures can be assumed to follow a Gaussian distribution. Therefore, nodal pressures can be selected as statistical variables of a Gaussian distribution. If any pair of the two statistical variables are not correlated, this means they are independent from each other as well. From the above discussion the correlation analysis can be performed as follows.

Suppose r is a sample estimate of the population parameter ρ , then:

$$\rho = \frac{\sigma_{uv}}{\sigma_u \sigma_v} \quad (4.12)$$

$$r = \frac{\hat{\sigma}_{uv}}{\hat{\sigma}_u \hat{\sigma}_v} \quad (4.13)$$

where:

$$\hat{\sigma}_{uv} = 1/(L-1) \sum_{i=1}^L (u_i - \bar{u})(v_i - \bar{v}), \text{ sample covariance.}$$

$$\hat{\sigma}_u^2 = 1/(L-1) \sum_{i=1}^L (u_i - \bar{u})^2, \text{ sample variance.}$$

$$\hat{\sigma}_v^2 = 1/(L-1) \sum_{i=1}^L (v_i - \bar{v})^2, \text{ sample variance}$$

$$\bar{u} = 1/L \sum_{i=1}^L u_i, \text{ sample mean.}$$

$$\bar{v} = 1/L \sum_{i=1}^L v_i, \text{ sample mean.}$$

L = length of sample.

u_i = pressure value of node u (i = 1, 2, ...L).

v_i = pressure value of node v (i = 1, 2, ...L).

To test hypothesis H: $\rho = 0$ versus the alternative A: $\rho \neq 0$, it is necessary to calculate:

$$t = (r-0)/S_r = r(L-2)^{0.5}/((1-r^2)^{0.5}) \quad (4.14)$$

and reject H if $t \geq t_{(1 - \alpha/2)(L-2)}$

or if $t \leq - t_{(1 - \alpha/2)(L-2)}$

$t_{(1 - \alpha/2)(L-2)}$ can be found in the t - distribution table; where 100 α is the significance level.

This correlation analysis will be performed between each pair of the nodes. At the same time, practical engineering understanding of the particular network should be helpful in the final construction of the equivalent network.

After the correlation analysis, assume the final equivalent network is as shown in Fig 4.3

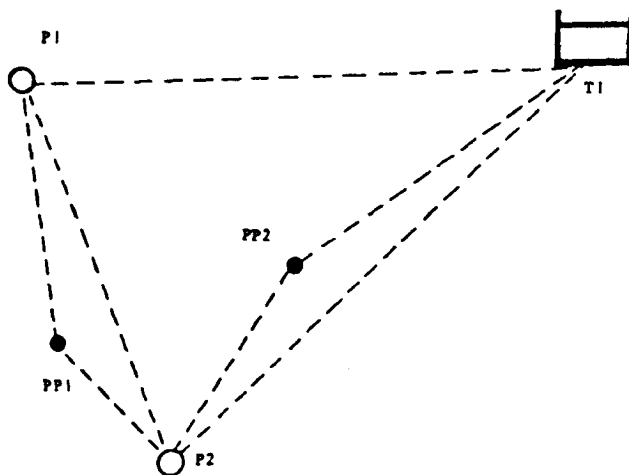


Fig 4.3 Equivalent Network of Fig. 4.2

Using the operating data or the data obtained from dynamic simulation of a given water supply distribution system as discussed in Chapter 2, it is possible to obtain the nodal consumptions and pipe resistances of its equivalent network so that the equivalent network can be dealt with as a real network. This will be demonstrated through the following deduction.

For the equivalent network, the nodal equations will still be:

$$\sum_{j \in N(i)} Q_{ij} = u_i - y_i \quad i \in N \quad (4.15)$$

But where:

N = total number of nodes in the equivalent network.

$N(i)$ = subset of node j connections to node i .

Q_{ij} = the "flow" through an equivalent pipe from node i to node j (positive flow). (e.g., l/s)

u_i = pump flow into node i . (e.g., l/s)

y_i = equivalent consumption flow out of node i . (e.g., l/s)

$$y_i = \begin{cases} y_i, & \text{if } i = 1, \dots, N - N_s \\ 0, & \text{if } i = N - N_s + 1, \dots, N. \end{cases} \quad (\text{here it is required that source nodes are numbered last}).$$

N_s = number of source nodes.

Using Manning's formula as given in Chapter 2:

$$\Delta H_{ij} = H_i - H_j = R_{ij} Q_{ij}^2$$

then

$$Q_{ij} = (H_i - H_j) / (R_{ij}^{0.5} |H_i - H_j|^{0.5}) \quad (4.16)$$

where

ΔH_{ij} = head drop from node i to node j . (e.g., m)

H_i, H_j = head at node i and node j , respectively. (e.g., m)

R_{ij} = resistance coefficient of an equivalent pipe from node i to node j .

The N nodal equations in eqn.(4.15) can be combined into a vector equation as follows:

$$\underline{A} \underline{Q} = \underline{U} - \underline{Y} \quad (4.17)$$

where:

\underline{A} = connection matrix of dimension $N \times N_p$ (whose elements are 0 or 1).

\underline{Q} = vector of dimension N_p with elements Q_{ij}

\underline{U} = vector of dimension N with elements u_i .

\underline{Y} = vector of dimension N with elements y_i .

N_p = number of "pipes" in the equivalent network.

For M sets of measurements, there are M vector equations:

$$\begin{aligned} \underline{AQ}^{(1)} &= \underline{U}^{(1)} - \underline{Y}^{(1)} \\ \dots & \quad \dots \quad \dots \\ \underline{AQ}^{(k)} &= \underline{U}^{(k)} - \underline{Y}^{(k)} \\ \dots & \quad \dots \quad \dots \\ \underline{AQ}^{(M)} &= \underline{U}^{(M)} - \underline{Y}^{(M)} \end{aligned} \quad (4.18)$$

If R_{ij} changes slowly, the following approximations may be valid:

$$R_{ij}^{(1)} = R_{ij}^{(2)} = \dots = R_{ij}^{(M)}$$

then

$$\begin{aligned} \Delta^{H_{ij}}(k) / \Delta^{H_{ij}}(1) &= \\ [Q_{ij}^{(k)} \mid Q_{ij}^{(k)} \mid R_{ij}^{(k)}] / [Q_{ij}^{(1)} \mid Q_{ij}^{(1)} \mid R_{ij}^{(1)}] \\ &= [Q_{ij}^{(k)} \mid Q_{ij}^{(k)}] / [Q_{ij}^{(1)} \mid Q_{ij}^{(1)}] \end{aligned}$$

$$k = 2, 3, \dots, M$$

or

$$Q_{ij}^{(k)} = \gamma_{ij}^{(k)} Q_{ij}^{(1)} \quad (4.19)$$

in which

$$\gamma_{ij}^{(k)} = \text{sign}[\Delta H_{ij}^{(k)} / \Delta H_{ij}^{(1)}] \times |\Delta H_{ij}^{(k)} / \Delta H_{ij}^{(1)}|^{0.5}$$

By substituting eqn.(4.19) into eqn.(4.18), eqn(4.18) can be rewritten as:

$$\begin{aligned} \underline{A}^{(1)} \underline{Q}^{(1)} &= \underline{U}^{(1)} - \underline{Y}^{(1)} \\ \dots \quad \dots \quad \dots & \\ \underline{A}^{(k)} \underline{Q}^{(1)} &= \underline{U}^{(k)} - \underline{Y}^{(k)} \\ \dots \quad \dots \quad \dots & \\ \underline{A}^{(M)} \underline{Q}^{(1)} &= \underline{U}^{(M)} - \underline{Y}^{(M)} \end{aligned} \quad (4.20)$$

where:

$$\underline{A}^{(1)} = \underline{A}; \quad \underline{A}^{(k)} = \underline{A} \times [\gamma_{ij}^{(k)}], \quad k = 2, 3, \dots, M.$$

$[\gamma_{ij}^{(k)}]$ = a diagonal matrix of dimension $N_p \times N_p$ with elements $\gamma_{ij}^{(k)}$.

Define a_i as the fractional consumption rate at node i , such that:

$$a_i = y_i / Y_D \quad \text{and} \quad \sum_{i=1}^{N-N_s} a_i = 1 \quad (4.21)$$

or

$$\sum_{i=1}^{N-N_s} (a_i \times Y_D) = Y_D$$

where

Y_D = total demand flow.

Combining eqn.(4.21) with eqn.(4.20) leads to the partitioned matrix-vector equations as follows:

$$\begin{bmatrix} \underline{A}^{(k)} & \underline{Y}_D^{(k)} \\ \underline{0} & \underline{E} \end{bmatrix} \times \begin{bmatrix} \underline{Q}^{(1)} \\ \underline{a} \end{bmatrix} = \begin{bmatrix} \underline{U}^{(k)} \\ 1 \end{bmatrix}$$

$k = 1, 2, \dots, M$ (4.22)

where

$\underline{E} = [1 \ 1 \ \dots \ 1]$ is a row vector of dimension $N-N_s$

$\underline{a} = [a_1 \ a_2 \ \dots \ a_{N-N_s}]^T$

$$\underline{Y}_D^{(k)} = \begin{bmatrix} Y_D^{(k)} & & & 0 \\ Y_D^{(k)} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & Y_D^{(k)} \\ 0 & \dots & 0 & \dots & 0 \\ \dots & & \dots & & \dots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} \quad N \times (N-N_s)$$

Finally, the parameters to be estimated are:

$$\underline{\theta} = \left[\frac{Q^{(1)}}{a} \right]$$

$\underline{\theta}$ can easily be estimated by using the method of least-squares (HSIA, 1977):

Let

$$\underline{X}_k = \left[\begin{array}{c|c} \underline{A}^{(k)} & \underline{Y}_D^{(k)} \\ \hline \underline{0} & \underline{E} \end{array} \right] \\ k = 1, 2, \dots, M$$

then

$$\underline{X}^T = [\underline{X}_1^T, \underline{X}_2^T, \dots, \underline{X}_M^T]$$

and let

$$\underline{Z}_k = \left[\begin{array}{c} \underline{U}^{(k)} \\ 1 \end{array} \right] \\ k = 1, 2, \dots, M$$

then

$$\underline{Z}^T = [\underline{Z}_1^T \ \underline{Z}_2^T \ \dots \ \underline{Z}_M^T]$$

Eqn.(4.22) becomes

$$\left[\begin{array}{c} \underline{X}_1 \\ \underline{X}_2 \\ \vdots \\ \underline{X}_M \end{array} \right] \cdot \underline{\theta} = \left[\begin{array}{c} \underline{Z}_1 \\ \underline{Z}_2 \\ \vdots \\ \underline{Z}_M \end{array} \right]$$

or:

$$\underline{X} \underline{\theta} = \underline{Z} \tag{4.23}$$

Define an error vector $\underline{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_m)^T$ and let

$$\underline{\epsilon} = \underline{z} - \underline{x}\underline{\theta}$$

The estimate $\hat{\underline{\theta}}$ of $\underline{\theta}$ will be chosen in such a way that the criterion, J, given by:

$$\begin{aligned} J &= \sum_{i=1}^M \epsilon_i^2 = \underline{\epsilon}^T \underline{\epsilon} = (\underline{z} - \underline{x}\underline{\theta})^T (\underline{z} - \underline{x}\underline{\theta}) \\ &= \underline{z}^T \underline{z} - \underline{\theta}^T \underline{x}^T \underline{z} - \underline{z}^T \underline{x}\underline{\theta} + \underline{\theta}^T \underline{x}^T \underline{x}\underline{\theta} \end{aligned} \quad (4.24)$$

is minimized.

Differentiate J with respect to $\underline{\theta}$ and equate the result to zero to determine the conditions on the estimate $\underline{\theta}$ that minimizes J. Thus

$$\left. \frac{\partial J}{\partial \underline{\theta}} \right|_{\underline{\theta} = \hat{\underline{\theta}}} = -2 \underline{x}^T \underline{z} + 2 \underline{x}^T \underline{x} \hat{\underline{\theta}} = 0$$

This yields

$$\underline{x}^T \underline{x} \hat{\underline{\theta}} = \underline{x}^T \underline{z}$$

from which $\hat{\underline{\theta}}$ can be solved for as

$$\hat{\underline{\theta}} = (\underline{x}^T \underline{x})^{-1} \underline{x}^T \underline{z} \quad (4.25)$$

A program coding EQUINET was produced in Fortran 77 to implement this algorithm.

However, in some cases of the study, it was found that some "flows" can occur from lower pressure nodes to higher pressure nodes without the existence of pumping. To avoid these problems a constrained

least-squares formulation needs to be solved. The problem is:

$$\begin{aligned}
 \min_{\underline{\theta}} J &= \underline{\epsilon}^T \underline{\epsilon} \\
 &= (\underline{Z} - \underline{X} \underline{\theta})^T (\underline{Z} - \underline{X} \underline{\theta}) \\
 &= \underline{Z}^T \underline{Z} - \underline{\theta}^T \underline{X}^T \underline{Z} - \underline{Z}^T \underline{X} \underline{\theta} + \underline{\theta}^T \underline{X}^T \underline{X} \underline{\theta} \\
 \text{s.t. } \theta_i &\geq 0 \quad i = 1, 2, \dots, N_p + N - N_s
 \end{aligned} \tag{4.26}$$

Note that in problem (4.26), $\underline{\theta}^T \underline{X}^T \underline{Z}$ is a scalar, then $\underline{\theta}^T \underline{X}^T \underline{Z} = (\underline{\theta}^T \underline{X}^T \underline{Z})^T = \underline{Z}^T \underline{X} \underline{\theta}$ and $\underline{Z}^T \underline{Z}$ is a constant. Therefore, problem (4.26) is equivalent to

$$\begin{aligned}
 \min_{\underline{\theta}} J_1 &= \underline{C}^T \underline{\theta} + (1/2) \underline{\theta}^T \underline{H} \underline{\theta} \\
 \text{s.t. } \theta_i &\geq 0, \quad i = 1, 2, \dots, N_p + N - N_s
 \end{aligned} \tag{4.27}$$

where:

$$\begin{aligned}
 \underline{C}^T &= -\underline{Z}^T \underline{X}, \\
 \underline{H} &= \underline{X}^T \underline{X}
 \end{aligned}$$

In fact, problem (4.27) is a Linear-Quadratic Programming problem, which can be solved with the routine E04NAF in the NAG FORTRAN LIBRARY (NAG, 1983). Therefore, a program coding ENCQP was produced in Fortran 77, which employed E04NAF, to cater for this algorithm.

With the estimate $\hat{\underline{\theta}}$, substituting the equivalent pipe flow vector $\underline{Q}^{(1)}$ into eqn.(4.16), the "resistance" R_{ij} of the equivalent network can be computed; substituting the fractional consumption rate vector \underline{a} into eqn.(4.21), the distributed consumption y_i can also be computed.

So far, the equivalent network has been thus constructed. Note that the parameters in the model have a clear and definite physical

meaning; and are consistent with the conventional theory of water network analysis. In particular the method does not rely on the assumption of proportional loading, as compared with the macroscopic model (DEMOYER and HORWITZ, 1975a, 1975b) or the piece-wise macroscopic model, which cannot be guaranteed to apply to most water supply systems.

4.4.2 Model Validations

The model validations were carried out on the Goldthorn zone of the Wolverhampton water supply and distribution system as shown in Fig.4.4 (COULBECK and ORR, 1986). A dynamic hydraulic simulation was performed using GINAS (COULBECK, 1985) for one day's operation of the Goldthorn zone, where half-hourly simulation time steps were utilized to obtain 48 half-hourly sets of data. The data simulated were then used for the parameter estimation of the equivalent network.

The results of the nodal correlation analysis are listed in Table 4.2 and the initial equivalent network is shown in Fig. 4.5.

Using the program coding EQUINET for the unconstrained least-squares procedure, the computed results are as summarized in Table 4.3.

From Table 4.3, it may be noticed that the flow of pipe 8 and the flow of pipe 10 are negative, which mean that the two flows are from node 5 to node 7 and from node 1 to node 3, respectively. However, the pressures at node 7 and node 3 are always higher than those at node 5 and node 1, respectively. This means that the flows are from lower

pressure nodes to higher pressure nodes, which are therefore not acceptable.

Consequently, the program coding ENCQP for the constrained least-squares procedure has been employed.

The computed results from program ENCQP are summarized in Table 4.4 and Table 4.5, respectively and their related equivalent networks are shown in Fig 4.6 and Fig.4.7, respectively. The C-values used in Hazen-Williams relationships in Fig.4.6 and Fig.4.7 are computed by assuming all pipe lengths to be equal to 1 m, all pipe diameters to be equal to 100mm for convenience and by use of the Hazen-Williams relationship in Chapter 2.

The results in Table 4.4 are obtained by using the simulation data from 0.00hr to 8.00hrs, when all pumps are in use. Whereas the results in Table 4.5 are obtained by using the simulation data from 8.30hrs to 23.00hrs when one pump is off (the lower pump between nodes 122 and 6 in Fig 4.5).

In order to validate the equivalent network model, GINAS has been employed to perform dynamic simulations over the equivalent network under the same operational conditions as were performed over the original network. Both of these results are listed in Table 4.6 to give a comparison; this shows that the accuracy of the equivalent network model is satisfactory.

Validation of the suitability of the equivalent network for different operating conditions is investigated by changing the initial head of

reservoir 4 from 182.00m to 183.00m and scaling the total demand by a factor of 1.3. The simulation results, performed for the original network and the equivalent network, respectively, are listed in Table 4.7. The largest difference in Table 4.7 is 2.1m, the relative error being 1.17%, which indicates that the equivalent network is accurate for simulation and control purposes.

Further, using the parameters for the equivalent network obtained for the period of 8.30hr to 23.00hrs to simulate over the period of 0.00hr to 8.00hrs, gives the results as tabulated in Table 4.8. The largest difference is 2.89m, the relative error being 1.63% which is still small enough. It seems, at least for the network studied, that the equivalent network is not too dependent on the variation of pump combinations when used for simulations. In other words, it is possible to use one set of parameters to represent the equivalent network for varied operational conditions, regardless of the variations of pump combinations.

There are only 10 nodes and 13 pipes in the equivalent network. Whereas there are 36 nodes and 47 pipes in the original network. The reduction is about 72%, which is very effective in such a type of non-linear system. It follows that the reduction in system complexity would be much more significant in more complicated systems.

TABLE 4.2 RESULTS OF CORRELATION

ANALYSIS BETWEEN NODES ($\alpha=0.01$)

NODE PAIRS	$t_1 - \alpha/2$	Computed t	Correlation Result
Node1/Node2	2.6868	36.5346	YES
Node1/Node3	2.6868	3.4001	YES
Node1/Node4	2.6868	-1.4525	NO
Node1/Node5	2.6868	3.1621	YES
Node1/Node6	2.6868	4.8032	YES
Node1/Node7	2.6868	-1.4525	NO
Node2/Node3	2.6868	4.2317	YES
Node2/Node4	2.6868	-0.8713	NO
Node2/Node5	2.6868	3.9889	YES
Node2/Node6	2.6868	4.6607	YES
Node2/Node7	2.6868	-0.8717	NO
Node3/Node4	2.6868	-3.6431	YES
Node3/Node5	2.6868	-0.5052	NO
Node3/Node6	2.6868	7.9354	YES
Node3/Node7	2.6868	-3.6456	YES
Node4/Node5	2.6868	8.7352	YES
Node4/Node6	2.6868	-6.5673	YES
Node4/Node7	2.6868	3186.6274	YES
Node5/Node6	2.6868	-1.7295	NO
Node5/Node7	2.6868	8.7337	YES
Node6/Node7	2.6868	-6.5712	YES

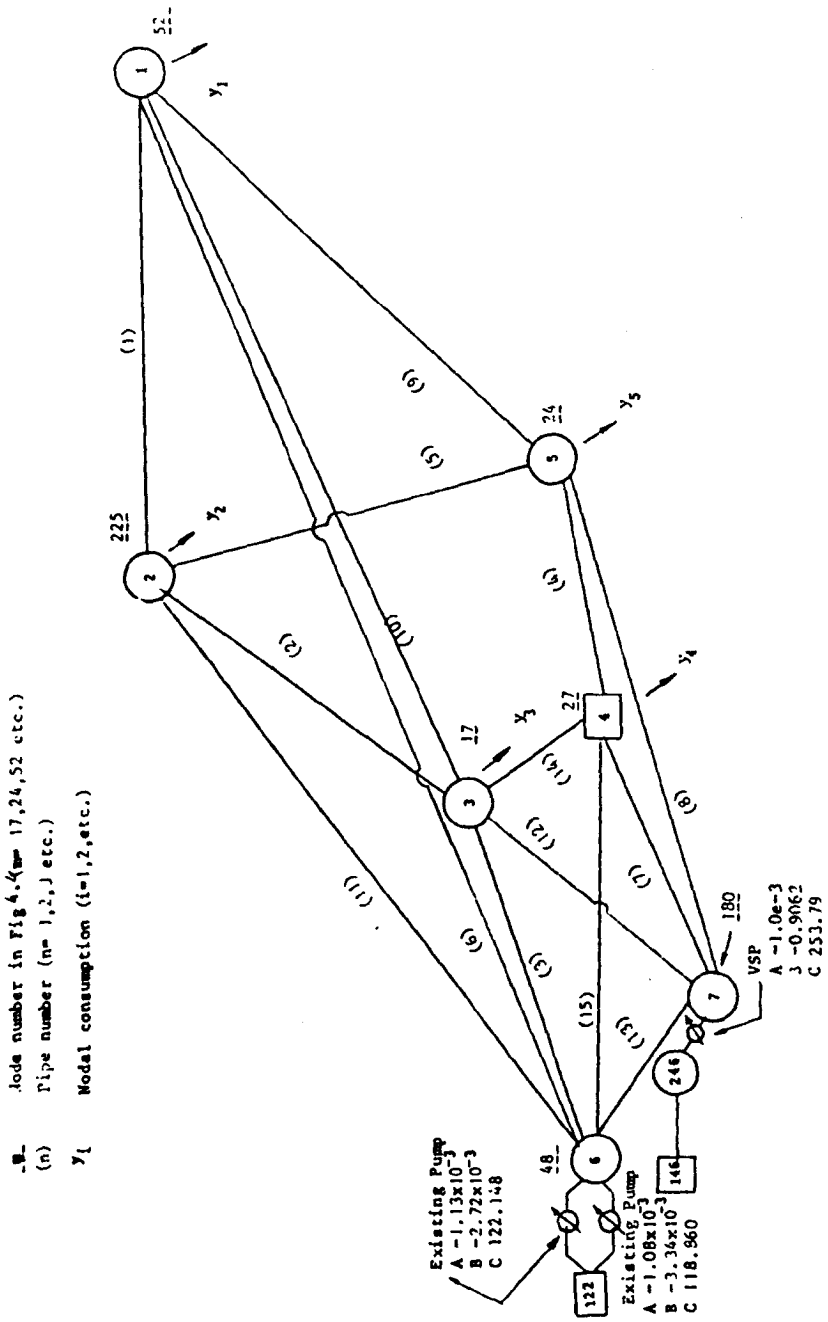


Fig 4.5 Equivalent Network of Fig 4.4 (to be further simplified)

TABLE 4.3 RESULTS FOR EQUIVALENT NETWORK MODELLING
OBTAINED FROM PROGRAM EQUINET

Equivalent Pipe Flow Qp(i) 1/s									
Qp(1)	Qp(2)	Qp(3)	Qp(4)	Qp(5)	Qp(6)	Qp(7)	Qp(8)	Qp(9)	Qp(10)
1.061	0.428	18.531	16.323	2.563	8.083	117.60	-4.782	3.587	-1.805

TABLE 4.3 (Continued)

Equivalent Pipe Flow Qp(i) 1/s					Fractional Consumption Rate a(i)				
Qp(11)	Qp(12)	Qp(13)	Qp(14)	Qp(15)	a(1)	a(2)	a(3)	a(4)	a(5)
0.103	1.577	3.916	2.257	186.441	0.05600	0.08063	0.09825	0.79862	0.02902

**TABLE 4.4 RESULTS FOR EQUIVALENT NETWORK MODELLING
FROM PROGRAM ENCQP (0.00 HR. TO 8.00 HRS)**

Equivalent Pipe Flow Qp(i) 1/s									
Qp(1)	Qp(2)	Qp(3)	Qp(4)	Qp(5)	Qp(6)	Qp(7)	Qp(8)	Qp(9)	Qp(10)
0.0	0.415	48.983	2.494	0.0	0.0	115.966	0.0	0.0	0.705
Equivalent Pipe Resistance RE(i) m/(1/s)²									
RE(1)	RE(2)	RE(3)	RE(4)	RE(5)	RE(6)	RE(7)	RE(8)	RE(9)	RE(10)
∞	3.2487	0.0091	0.0275	∞	∞	0.0007	∞	∞	2.3209

TABLE 4.4 (Continued)

Equivalent Pipe Flow Qp(i) 1/s				Fractional Consumption Rate a(i)					
Qp(11)	Qp(12)	Qp(13)	Qp(14)	Qp(15)	a(1)	a(2)	a(3)	a(4)	a(5)
0.0	0.0	0.0	3.757	161.366	0.00379	0.00237	0.24869	0.72037	0.01321
Equivalent Pipe Resistance RE(i) m/(1/s)²									
RE(11)	RE(12)	RE(13)	RE(14)	RE(15)					
∞	∞	∞	0.0196	0.0008					

**TABLE 4.5 RESULTS FOR EQUIVALENT NETWORK MODELLING
FROM PROGRAM ENCQP (8.30 HRS TO 23.00 HRS)**

Equivalent Pipe Flow Qp(i) l/s									
Qp(1)	Qp(2)	Qp(3)	Qp(4)	Qp(5)	Qp(6)	Qp(7)	Qp(8)	Qp(9)	Qp(10)
7.811	2.099	38.755	29.103	24.680	0.0	115.783	0.0	0.0	5.017
Equivalent Pipe Resistance RE(i) m/(l/s) ²									
RE(1)	RE(2)	RE(3)	RE(4)	RE(5)	RE(6)	RE(7)	RE(8)	RE(9)	RE(10)
0.0234	1.7843	0.0124	0.0004	0.0029	∞	0.0007	∞	∞	0.2556

TABLE 4.5 (Continued)

Equivalent Pipe Flow Qp(i) l/s					Fractional Consumption Rate a(i)				
Qp(11)	Qp(12)	Qp(13)	Qp(14)	Qp(15)	a(1)	a(2)	a(3)	a(4)	a(5)
0.0	0.0	0.0	53.609	110.228	0.03778	0.05587	0.22549	0.66376	0.01537
Equivalent Pipe Resistance RE(i) m/(l/s) ²									
RE(11)	RE(12)	RE(13)	RE(14)	RE(15)					
∞	∞	∞	0.0035	0.0007					

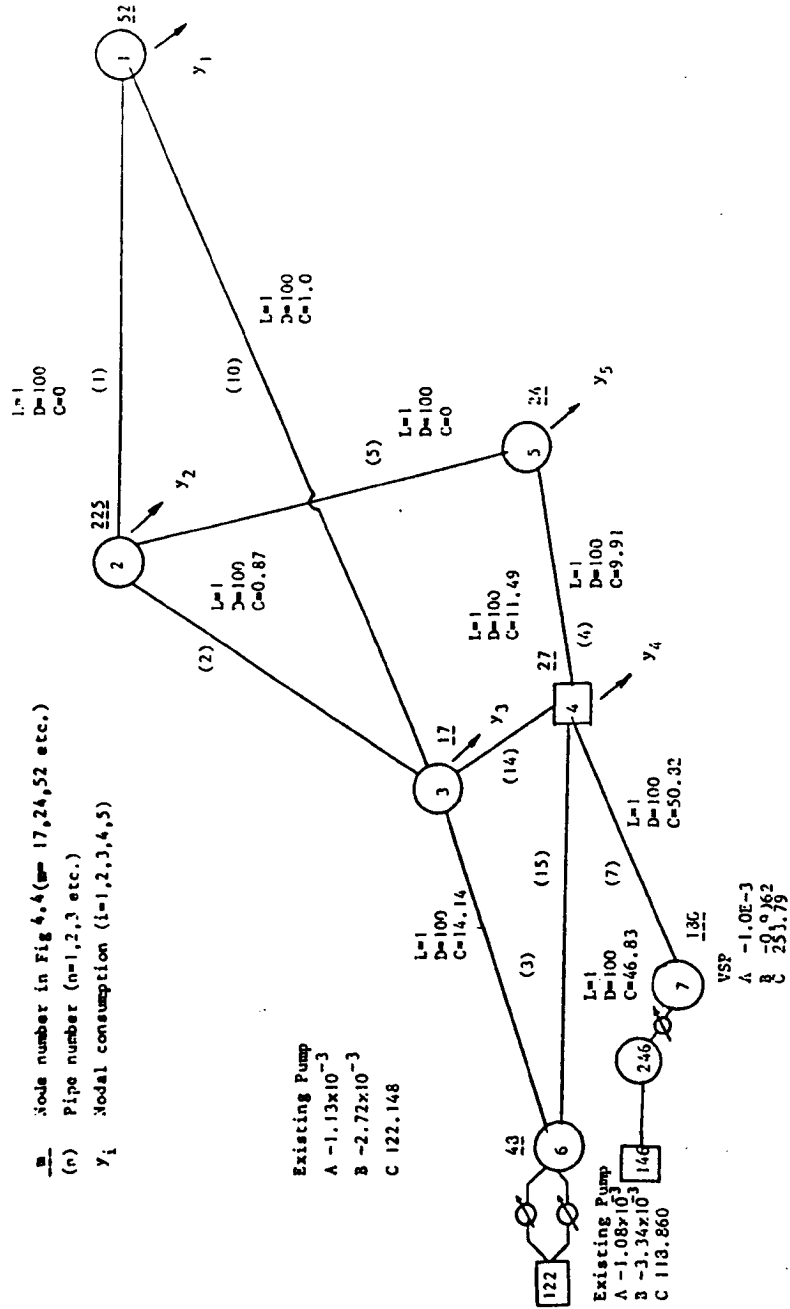


Fig 4.6 Equivalent Network of Fig 4.4 (0.0 hr to 8.0 hrs)

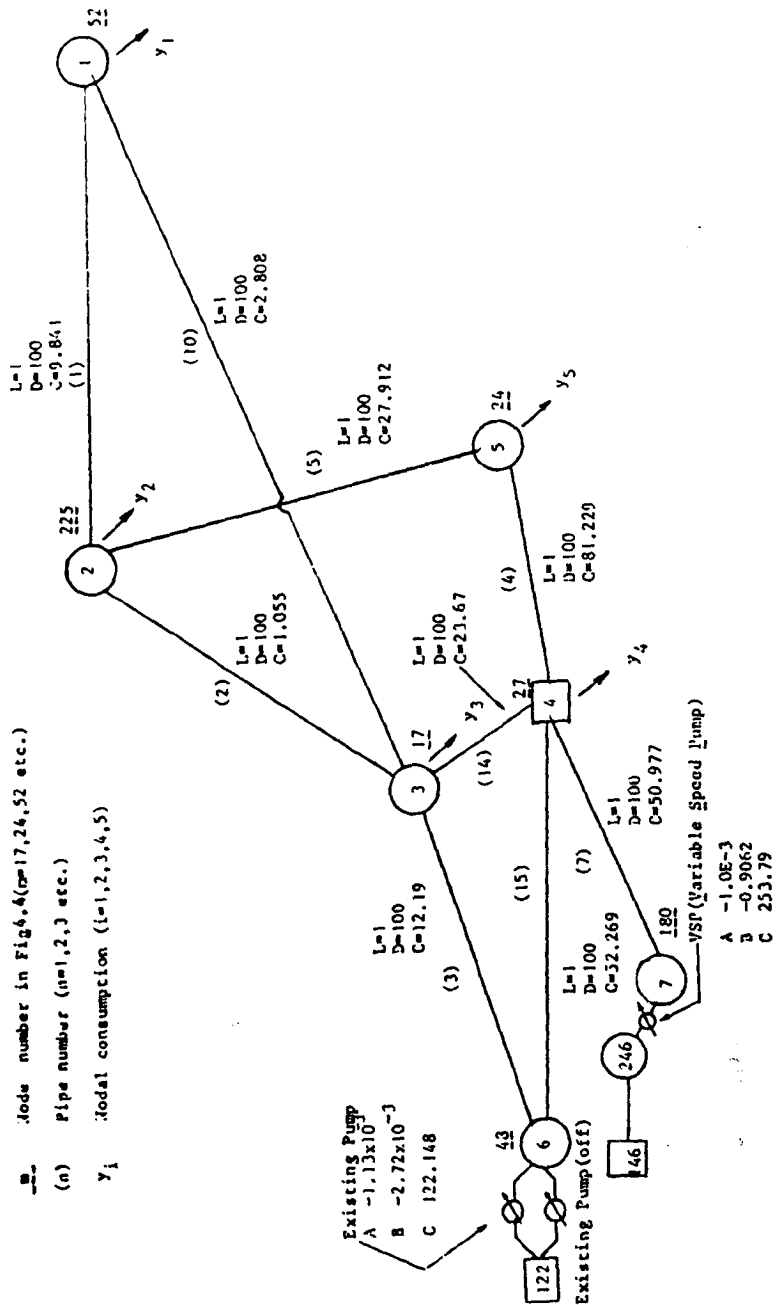


Fig 4.7 Equivalent Network of Fig 4.4 (9.0 hrs to 23.00 hrs)

TABLE 4.6
VALIDATION OF EQUIVALENT NETWORK MODELLING

		NODAL HEADS (M)																							
		node 1		node 2		node 3		node 4		node 5		node 6		node 7		node 246		node 146		node 122					
VSON*	VSEN+	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN		
0	180.57	180.57	181.16	181.16	181.72	181.89	182.00	182.00	181.83	181.81	203.56	203.96	192.06	192.13	56.87	56.88	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
1	180.88	181.19	181.38	181.69	182.85	182.27	182.27	182.27	181.82	181.92	203.80	204.01	192.13	192.19	56.88	56.89	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
2	181.37	181.74	181.69	182.05	184.76	184.40	182.17	182.17	182.04	182.08	204.25	204.23	192.21	192.28	56.89	56.90	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
3	181.80	181.90	181.99	181.90	186.50	186.94	182.28	182.28	182.19	182.23	204.72	204.56	192.31	192.38	56.90	56.91	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
4	181.98	182.37	182.15	182.75	186.91	188.44	182.41	182.41	182.32	182.37	204.86	204.68	192.42	192.50	56.91	56.92	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
5	182.04	182.83	182.24	182.79	186.70	187.77	182.53	182.53	182.44	182.48	204.50	204.65	192.53	192.60	56.92	56.92	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
6	182.03	182.50	182.28	182.81	186.10	186.47	182.65	182.65	182.54	182.59	204.66	204.56	192.63	192.71	56.93	56.93	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
7	181.50	182.02	182.20	182.57	184.50	183.85	182.76	182.76	182.62	182.64	204.28	204.35	192.73	192.80	56.93	56.93	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
8	180.93	180.17	181.66	181.10	179.05	179.01	182.82	182.82	182.61	182.50	203.27	203.96	192.86	192.77	56.94	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
9	179.28	178.75	180.71	179.90	172.85	173.45	182.78	182.78	182.45	182.18	191.47	192.18	192.75	192.77	56.93	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
10	176.20	175.85	179.66	179.17	175.50	176.14	182.73	182.73	182.26	182.18	191.79	192.22	192.71	192.73	56.93	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
11	172.70	173.17	178.88	179.05	179.24	180.12	182.70	182.70	182.13	182.29	192.45	192.43	192.68	192.70	56.93	56.93	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
12	173.12	173.79	179.19	180.60	180.65	179.37	182.69	182.69	182.17	182.17	192.81	192.58	192.67	192.69	56.93	56.93	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
13	174.52	174.52	179.73	180.03	181.44	180.99	182.70	182.70	182.26	182.44	193.05	192.70	192.68	192.70	56.93	56.93	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
14	174.00	174.62	179.60	180.50	181.54	181.29	182.71	182.71	182.25	182.45	193.09	192.70	192.69	192.71	56.93	56.93	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
15	174.19	174.53	179.63	180.49	181.37	180.69	182.72	182.72	182.26	182.45	193.04	192.69	192.69	192.72	56.93	56.93	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
16	177.70	178.12	180.65	181.30	181.48	180.84	182.73	182.73	182.40	182.51	193.08	192.78	192.71	192.73	56.93	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
17	180.22	180.57	181.46	181.55	181.46	181.33	182.76	182.76	182.52	182.57	193.08	192.86	192.73	192.75	56.93	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
18	179.98	180.05	181.32	181.29	180.81	180.75	182.78	182.78	182.52	182.55	192.88	192.80	192.76	192.77	56.93	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
19	179.44	179.48	181.09	181.00	180.13	180.12	182.79	182.79	182.49	182.51	192.67	193.73	192.76	192.78	56.94	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
20	180.45	179.98	181.53	181.26	179.96	180.67	182.80	182.80	182.57	182.56	192.68	192.80	192.77	192.79	56.94	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
21	181.36	180.67	181.96	181.64	180.13	180.73	182.83	182.83	182.65	182.64	192.78	192.92	192.79	192.81	56.94	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
22	181.64	181.12	182.13	181.89	180.85	181.29	182.86	182.83	182.70	182.68	192.98	192.98	192.82	192.81	56.94	56.94	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	
23	181.77	181.50	182.24	182.13	181.56	182.26	182.91	182.91	182.76	182.79	193.20	193.08	192.88	192.89	56.95	56.95	69.33	69.33	69.33	69.33	69.33	69.33	69.33	69.33	

* VSON - Value Simulated over Original Network

+ VSEN - Value Simulated over the Equivalent Network

TABLE 4.7
VALIDATION OF EQUIVALENT NETWORK MODELLING
FOR DIFFERENT OPERATIONAL CONDITIONS

NODAL HEADS (M)																				
node 1	node 2		node 3		node 4		node 5		node 6		node 7		node 146	node 122						
VSON*	VSEN+	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN					
0	180.73	179.37	181.70	180.35	179.78	183.00	183.00	182.78	182.69	203.44	204.05	192.94	193.02	56.96	56.96	69.33	69.33	96.18	96.18	
2	181.85	182.22	182.39	182.70	183.96	183.27	183.11	182.95	182.96	204.22	204.44	193.04	193.11	56.97	56.97	69.33	69.33	96.18	96.18	
4	182.66	183.63	182.94	184.83	186.86	187.07	183.32	183.32	183.21	183.26	204.94	204.81	193.23	193.30	56.98	56.99	69.33	69.33	96.18	96.18
6	182.57	184.20	182.97	184.52	185.63	184.93	183.52	183.52	183.39	183.41	204.66	204.67	193.41	193.43	57.00	57.00	69.33	69.33	96.18	96.18
8	180.61	179.23	181.82	180.89	175.56	173.86	183.65	183.65	183.35	183.12	203.01	203.64	193.52	193.59	57.01	57.01	69.33	69.33	96.18	96.18
10	172.82	171.25	178.45	177.34	171.30	171.25	183.44	183.44	182.68	182.50	191.64	192.15	193.34	193.36	56.99	56.99	69.33	69.33	96.18	96.18
12	167.77	169.33	177.62	179.20	179.00	177.59	183.31	183.31	182.46	182.75	192.71	192.63	193.22	193.24	56.98	56.98	69.33	69.33	96.18	96.18
14	169.08	171.28	178.19	180.33	180.17	178.08	183.25	183.25	182.50	182.80	192.95	192.76	193.17	193.19	56.98	56.98	69.33	69.33	96.18	96.18
16	175.00	176.47	179.81	180.70	180.00	178.95	183.19	183.19	182.65	182.80	192.89	192.83	193.11	193.13	56.97	56.97	69.33	69.33	96.18	96.18
18	178.64	178.23	180.81	180.57	178.98	178.69	183.16	183.17	182.76	182.77	192.61	192.78	193.10	193.12	56.97	56.97	69.33	69.33	96.18	96.18
20	179.32	178.02	181.07	180.43	177.66	178.46	186.11	183.11	182.77	182.69	192.34	192.72	193.05	193.06	56.97	56.97	69.33	69.33	96.18	96.18
22	181.17	179.97	181.98	181.40	178.83	180.62	183.11	183.10	182.90	182.84	192.62	192.98	193.04	193.05	56.97	56.96	69.33	69.33	96.18	96.18

* VSON - Value Simulated over Original Network

+ VSEN - Value Simulated over the Equivalent Network

TABLE 4.8
VALIDATION OF EQUIVALENT NETWORK MODELLING
FOR DIFFERENT OPERATIONAL CONDITIONS
 (the influence of pump combinations is ignored)

NODAL HEADS (M)																				
node 1	node 2		node 3		node 4		node 5		node 6		node 7		node 146		node 122					
VSON*	VSEN+	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN	VSON	VSEN			
0	180.73	179.96	181.70	181.32	179.78	181.48	183.00	183.00	182.78	182.73	203.44	203.39	192.94	192.97	56.96	56.96	69.33	69.33	96.18	96.18
2	181.85	181.84	182.39	182.38	183.96	183.09	183.11	183.11	182.95	182.99	204.22	203.55	193.04	193.06	56.97	56.97	69.33	69.33	96.18	96.18
4	182.66	182.92	182.94	183.07	186.86	183.97	183.32	183.32	183.21	183.28	204.94	203.68	193.23	193.25	56.98	56.98	69.33	69.33	96.18	96.18
6	182.57	182.67	182.97	183.03	185.63	183.64	183.52	183.52	183.39	183.44	204.66	203.72	193.41	193.43	57.00	57.00	69.33	69.33	96.18	96.18
8	180.61	177.80	181.82	180.47	175.46	179.14	183.65	183.65	183.35	183.15	203.01	203.44	193.52	193.54	57.01	57.01	69.33	69.33	96.18	96.18

* VSON - Value Simulated over Original Network

+ VSEN - Value Simulated over the Equivalent Network

4.5 CONCLUSIONS

This chapter has presented in detail several methods of network modelling and simplification in order to make the on-line optimal control of system operations feasible, particularly to speed up the solution time of the optimization problems. With a set of explicit nonlinear regressional equations, a macroscopic model, originated from DeMoyer et al's work, can represent the major aspects of the original much more detailed network. However, its applicabilities are restricted by the requirements of proportional loading, which are difficult to satisfy in many realistic water supply and distribution systems. The extended work on the macroscopic model by the author leads to the piecewise macroscopic model which can be successfully applied to those systems in which the loading pattern is not too far from proportional loading assumptions. For the validation of the piecewise macroscopic model, both sets of results, from multiple regression and stepwise regression, respectively, are presented and are shown to be satisfactory (the average errors of typical results are 1.21m and 0.68m, respectively, refer to Table 4.1). Furthermore, the stepwise regressional equations, in terms of covariance contributions in regression, only take important factors into account and delete those unimportant factors. This might increase the accuracy of the predicted results (OSTLE, 1983). Consequently, as can be seen from the numerical results, the stepwise regressional equations are much simpler than their corresponding multiple regressional equations. Also, the standard deviations, and thus the accuracies, of the stepwise regressional equations are much improved, although their correlation coefficients are insignificantly lower,

than those of their corresponding multiple regression equations. In conclusion, it is better to adopt the stepwise regression method to build a piecewise macroscopic model for a system. Studies show that the division of a day into proportional loading periods mainly depends on the total demand. For a system where there are periods within which the total demand is relatively stable, there should be corresponding piecewise macroscopic models for each period. However, in a system where the total demand varies significantly and irregularly, the work of building-up the sets of piecewise macroscopic models (regression computations) would become cumbersome, and even impractical. In this sense, piecewise macroscopic models are no longer applicable to that system.

Based on the well-known nodal equations, the equivalent network model is developed by the author by introducing the concept of fictitious pipes. By employing matrix algebra, the detailed mathematical deduction leads to a least-squared estimation problem which in essence minimizes the discrepancies between the original detailed network model and the simplified equivalent network model. For practical applications, two algorithms which cater for unconstrained and constrained least-squared estimations, have been derived. Once the least-squared estimation has been performed, the pipe resistances and nodal consumption rates will become known and then the equivalent network can be treated as an ordinary network. The derivation of the equivalent network model has a clear and definite physical interpretation and is consistent with the conventional theory of water network analysis. In particular, the method, as compared with the macroscopic model or the piecewise macroscopic model, does not rely on

the assumptions of proportional loading.

The practical validation of the equivalent network model using the Wolverhampton system confirms that the accuracy of the model is satisfactory. The simulation results derived by GINAS performed both over the original network and over its equivalent network show that the discrepancies are within 1.53m (relative error, 0.8%, refer to Table 4.6); validation of the suitability of the equivalent network for varied operation conditions shows that the maximum discrepancy is 2.89m (relative error, 1.63%, refer to Table 4.8). The equivalent network model is very effective in the reduction of system complexity. For the system currently studied, this reduction is about 70% in terms of the numbers of network components. This reduction could be much more significant for more complicated systems. Solving time is directly related to the numbers of network components and it follows that this reduction will result in significant improvements in network simulation times -- essential for control purposes.

The macroscopic model and the piecewise macroscopic model, which include pumps and controls and incorporate reservoir dynamics directly, can allow an explicit solution. This is particularly beneficial for formulating a very useful optimal control model, which will be further discussed in Chapter 5. In contrast, the equivalent network model does still require an iterative solution. Further, at the present, the simplification by the equivalent model is restricted to the distribution part of a system. Pumps and their controls have not been taken into account directly. In other words, pump controls could influence the parameter values and even the configuration of the

equivalent network.

From the above discussion, the applicabilities and accuracies of the various modelling methods presented should become very obvious. This should guide the selection among these methods for a particular system. However, the selection of which model to use for a particular system will not only depend on the characteristics of the system itself (e.g., load pattern), but will also primarily depend on what kind of algorithm for optimal operations of the system is to be employed. In Chapter 5, two algorithms to cater for the optimal operations of different classes of water supply and distribution systems are developed based on the piecewise macroscopic model and the equivalent network model, respectively. These algorithms will be fully discussed in Chapter 5.

5.1 INTRODUCTION

As stated in Chapter 1, in order to save energy and raise economic efficiency for water supply and distribution systems, it is considered inevitable to have to adopt computerized optimal system operations.

Mathematically, the objective function of the optimal operations of a certain class of systems, can be stated generally as follows:

$$J = \sum_e \int_{t_1}^{t_f} [c_e Q_e(t) + b_e(t) H_e(t) Q_e(t) / \eta(Q_e(t))] dt + \sum_r f[X_r(t_f)] \quad (5.1)$$

where:

t_1, t_f = initial and final time of the operational (control) periods, respectively. (e.g., h)

c_e = the unit cost of water production associated with pump e. (e.g. £/m³)

b_e = the unit price of energy for pump e at time t (usually a piecewise constant function). (e.g., £/kWh)

$Q_e(t)$ = flow of pump e at time t. (e.g., m³/h)

$H_e(t)$ = head increase through pump e at time t. (e.g., m)

$\eta(Q_e(t))$ = is the efficiency as a function of flow, including
a conversion factor for electrical power relating
water quantities to electrical energies. (%)

$f[X_r(t_f)]$ = reservoir final level penalty function. (e.g., £)

In its above form the optimal operation problem so described presents several serious difficulties: in eqn(5.1), $Q_e(t)$ and $H_e(t)$ are related by non-linear functions, $\eta_e(Q_e(t))$ is also a non-linear function about $Q_e(t)$, whereas $b_e(t)$ is a two-valued or a three-valued piecewise constant function; pump flow $Q_e(t)$ could be a discrete variable for fixed speed pumps or a continuous variable for variable speed or variable throttle pumps, and $f[X_r(t_f)]$ is usually expressed in a kind of non-linear function form. Furthermore the constraint set should include state equations describing the dynamics of the system which in general will be non-linear and time-varying. The constraint set should also include system equations describing the hydraulics of the system which are constituted of a set of simultaneous non-linear equations and which must be solved iteratively as stated in Chapter 2. In addition there should be included bounds on the capacities of pumps and reservoirs, etc.

The formulation still neglects certain costs or implied costs which may be significant, e.g., pump switching costs, electricity maximum charge, and costs reflecting pressure effects on leakage, etc.

In summary, this is a large-scale non-linear dynamic optimization problem with discrete and continuous variables, which represents one

of the most difficult problems to solve and there is almost no solution method in the field of optimization that can cope with all these difficulties simultaneously.

For more than ten years, many researchers have tried to solve this problem and various methods have been employed. Of the major advances, dynamic programming, in particular discrete dynamic programming, can theoretically solve this problem, but it is effective and practical for only single reservoir systems because of the problem of dimensionality and time of computation (COULBECK, 1977; COULBECK and ORR, 1983; SABET, 1983). Hierarchical optimization methods can be applied to multi-source, multi-reservoir systems when a linearized dynamic model of the system can be derived and the objective function can be approximated by a linear or quadratic form. In this case, the accuracy of the linearized model, convergence of the algorithm, as well as the design of coordinators, etc. are problematic and left open (COULBECK, 1977; COULBECK et al, 1985; FALLSIDE, 1975; JOALLAND and COHEN, 1980). Consequently, it has been recognized that research has not yet reached the stage of a universal approach which can cope with all the difficulties described above and which is applicable to any water supply distribution system. It is only by exploring the particular features of certain classes of systems, that appropriate and simplified optimization methods can be developed. These particular methods can then be very successfully applied to those classes of systems. (BRDYS et al, 1988; BRDYS, 1988; CHEN, 1988b; COULBECK, 1988).

In this chapter, two different algorithms for the optimal operations

of different classes of water supply and distribution systems, based on the author's several years of work both in China and in the U.K., are presented.

The first algorithm is applicable to a class of multi-source water supply and distribution system without significant storage within the system. This represents quite a few small water supply and distribution systems and sub-systems in reality. Under this condition, there are no dynamics in the system and hence the optimization problem for each time interval can be solved separately. Thus the number of decision variables is small yielding a very much simpler problem than eqn(5.1). Only pumping costs and treatment costs are included in the objective function, this represents the major part of the true cost and it is still highly non-linear. The pressure requirements of some pressure monitoring nodes are taken into account directly in the set of constraints, by employing the piecewise macroscopic models developed in Chapter 4, which can be solved explicitly, rather than the network simultaneous hydraulic equations system, which has to be solved iteratively and is thus very time-consuming. The restrictions on pumping capacities and so on are included in the set of constraints. The final formulation of the problem of optimal operation of the system by this algorithm results in a constrained non-linear programming problem, which can be solved using standard optimization methods.

The second algorithm caters for a class of multi-source, multi-reservoir systems, in which pumping flows and especially pumping consumed powers are more or less constant during the whole control period. This condition can be satisfied in quite a few systems with

the presence of significant storage, since the storage of reservoirs can compensate for the variations in water demand to a certain degree and keep the pumping flows (and thus consumed powers) of certain pump combinations from changing significantly. Based on the equivalent network model developed in Chapter 4, and by choosing times of pumping (WRc, 1985), rather than the usual pumping flows as decision variables, this algorithm forms a large-scale linear programming problem for which global optima can be guaranteed. This algorithm is able to deal with the difficulties of the mixture of fixed speed pumps (discrete variables), and variable speed and/or variable throttle pumps (continuous variables), by discretizing the speed range or throttle factor respectively. The discretization scheme is systematically constructed through a post-optimality analysis.

5.2 ALGORITHM FOR OPTIMAL OPERATIONS OF MULTI-SOURCE SYSTEMS CONTAINING INSIGNIFICANT WATER STORAGE

5.2.1 Formulation of Optimization Problem

For water supply and distribution systems with multiple water sources, there exists an optimal proportion of water delivered by each of the sources and pumping stations so that the total operating cost will be minimized, under the conditions that customers requirements, for water quantities and service pressures, will be met. This section consider the above case without significant water storage in the network.

Let e_0 denote the electrical power cost (e.g., Yuan*/KWh), $e_1(i)$ denote the cost of water treatment at plant i (Yuan/m³), including the cost of chemical dosage and the consumption of electrical energy in the process of water abstraction and purification, then:

$$\begin{aligned} e_2(i) &= e_0 \times 1000 / (3600 \times 102 \times \eta_m(i) / 100) \\ &= e_0 / [3.673 / \eta_m(i)] \end{aligned}$$

where:

$e_2(i)$ = cost of electricity at pumping station i for lifting one cubic metre of water to one metre height in a water column (Yuan/m⁴)

$\eta_m(i)$ = mean efficiency of station i (%).

The total hourly cost of water supply, $F_w(Q)$, can be expressed as follows:

* Yuan is the Chinese monetary unit.

$$F_w(Q) = \sum_{i=1}^n Q(i) \times [H_d(i) \times e_2(i) + e_1(i)] \quad (5.2)$$

where:

n = number of water sources.

$Q(i)$ = flow of pumping station i (m^3/hr).

$H_d(i)$ = delivery head of pumping station i (m).

In eqn(5.2), both $Q(i)$ ($i=1, \dots, n$) and $H_d(i)$ ($i=1, \dots, n$) are decision variables. However, $H_d(i)$ ($i=1, \dots, n$) are dependent on $Q(i)$ ($i=1, \dots, n$). Furthermore, each $H_d(i)$ is not only dependent on its corresponding pumping station flow but also, generally, on the flows of other pumping stations. To determine this dependency involves solving the network simultaneous hydraulic equations system iteratively as stated in Chapter 2. This will be very time-consuming and is considered infeasible for on-line control purposes (CHEN, 1985; DEMOYER and HORWITZ, 1975a, 1975b)

Fortunately, this problem can be avoided if the piece-wise macroscopic model developed in Chapter 4 is employed. Furthermore, the pressure requirements of some critical nodes in the system can be considered directly in this algorithm, because of the adoption of the piece-wise macroscopic model. These will be demonstrated as follows.

Substituting eqn(4.5) into eqn(5.2) yields:

$$F_w(Q) = \sum_{i=1}^n Q(i) \times \{ [C_d^{k(i,1)} + C_d^{k(i,2)} \times Q_d^\alpha + \sum_{i=1}^n C_d^{k(i,j+2)} \times Q^\alpha(j) + \sum_{\substack{i=1 \\ j \neq i}}^n C_d^{k(i,j+2+n)} \times Q(j) \times Q(i)] \times e_2(i) + e_1(i) \} \quad (5.3)$$

After this substitution, eqn(5.3) is then a nonlinear function only about $\{Q(i), i=1, \dots, n\}$. The $Q(i)$ in eqn(5.3) will usually be subject to the following constraints:

$$\sum_{i=1}^n Q(i) = Q_d \quad (5.4)$$

$$Q_{\min}(i) \leq Q(i) \leq Q_{\max}(i), \quad i = 1, 2, \dots, n \quad (5.5)$$

where:

Q_d = total demand flow (m^3/hr), which can be predicted by using the demand forecasting program developed in Chapter 3.

$Q_{\min}(i), Q_{\max}(i)$ = minimum and maximum discharge flow for pumping station i , respectively (m^3/hr).

In order to meet the customer requirements of service pressures, the pressures at the pressure monitoring nodes in the network should be higher than the lower limit of pressure at the node, i.e.:

$$H^k(j) \geq H_{\min}(j), \quad j=1, 2, \dots, n_h \quad (5.6)$$

where:

$H_{\min}(j)$ = the lower limit of pressure at node j in the network (m).

n_h = total number of pressure monitoring nodes in the network.

Substituting eqn(4.6) into eqn(5.6), gives:

$$\begin{aligned}
 H^k(j) &= B^k(j,1) + \sum_{i=1}^n B^k(j,i+1) \times Q^\alpha(i) + B^k(j,n+2) \times Q_d^\alpha \\
 &\geq H_{\min}(j)
 \end{aligned}
 \tag{5.7}$$

Eqn(5.3) to eqn(5.7) constitute a nonlinear programming problem about $\{ Q(i), i=1, \dots, n \}$:

$$\begin{aligned}
 \min_Q F_w(Q) &= \sum_{i=1}^n Q(i) \times \{ [C_d^k(i,1) + C_d^k(i,2) \times Q_d^\alpha + \sum_{i=1}^n C_d^k(i,j+2) \times \\
 &\quad Q^\alpha(j) + \sum_{\substack{j=1 \\ j \neq i}}^n C_d^k(i,j+2+n) \times Q(j) \times Q(i)] \times e_2(i) + e_1(i) \}
 \end{aligned}$$

s. t.

$$\sum_{i=1}^n Q(i) - Q_d = 0
 \tag{5.8}$$

$$Q_{\min}(i) \leq Q(i) \leq Q_{\max}(i), \quad i=1, 2, \dots, n$$

$$B^k(j,1) + \sum_{i=1}^n B^k(j,i+1) \times Q(i)^\alpha + B^k(j,n+2) \times Q_d^\alpha - H_{\min}(j) \geq 0$$

$$j=1, 2, \dots, n_h$$

By solving this optimization problem, the optimal apportioning of water among the pumping stations and water sources will be obtained.

Since it is assumed that there is no significant storage in the distribution system, there are no dynamics and, therefore, the optimization problem can be solved independently for each time interval. For the whole control period, if it is twenty four hours,

there will be twenty four such small optimization problems rather than a single large-scale optimization problem, which is much more complicated in non-linear programming problems.

5.2.2 Summary of the Optimization Technique

For the above nonlinear programming problem, the Sequential Unconstrained Minimization Technique (SUMT) is employed, which is very effective in solving many practical problems. Its principle is summarized as follows:

For a generalized nonlinear programming problem of :

$$\begin{aligned}
 & \min f(X) \\
 & \quad X \\
 & \text{s.t.} \\
 & g_i(X) \geq 0 \quad i = 1, 2, \dots, m \\
 & h_j(X) = 0 \quad j = 1, 2, \dots, p
 \end{aligned} \tag{5.9}$$

An inside penalty function is constituted in inequality constraints $g_i(X)$ ($i = 1, \dots, m$) and an outside penalty function is constituted in equality constraints $h_j(X)$ ($j = 1, \dots, p$). Adding these two functions to the objective function $f(X)$ gives:

$$P(X, r) = f(X) + r \sum_{i=1}^m 1/g_i(X) + (1/ r^{0.5}) \times \sum_{j=1}^p [h_j(X)]^2 \tag{5.10}$$

where r is a penalty factor, which is an infinite decreasing series of positive numbers, and can be formed from $r^{(k)} = r^{(k-1)} c$, where $0 < c < 1$ (often 0.01 to 0.20).

For $r = r^{(k)}$, minimizing penalty function $P(X, r^{(k)})$, the corresponding extreme point $X(r^{(k)})$ is found. For a series of penalty

factors $\{ r^{(k)} \}$, when $k \rightarrow \infty$, then $r^{(k)} \rightarrow 0$. Under certain conditions, the corresponding extreme point $X(r^{(k)})$ makes:

$$\lim_{k \rightarrow \infty} \left\{ r^{(k)} \sum_{i=1}^m 1/g_i(X) \right\} = 0 \quad (5.11)$$

$$\lim_{k \rightarrow \infty} \left\{ 1/(r^{(k)})^{0.5} \sum_{j=1}^p [h_j(X)]^2 \right\} = 0 \quad (5.12)$$

therefore

$$\lim_{k \rightarrow \infty} P[X(r^{(k)})] = f(X) \quad (5.13)$$

In practice, when k is large enough, $X(r^{(k)})$ can be considered as the approximate solution to eqn(5.9).

From the above statement, we can see that the constrained nonlinear programming problem of $f(X)$ is changed into the unconstrained nonlinear programming problem of $P(X,r)$, which is solved by using Powell's minimization method (one of the conjugate direction methods, FLETCHER, 1987; WAN, 1983).

In order to avoid the dependency among the n successively generated search directions S_i ($i=1,2,\dots,n$), when obtaining the conjugate direction S_{n+1} each time, the following inequalities must hold:

$$f_3 < f_1 \quad (5.14)$$

and

$$(f_1 - 2f_2 + f_3)(f_1 - f_2 - \Delta_m)^2 < 0.5 \Delta_m (f_1 - f_3)^2 \quad (5.15)$$

where:

f_1 = value at initial point X_0 .

f_2 = value at X_n along search direction S_n .

f_3 = value at X_{n+1} , which is the reflection point along the conjugate direction S_{n+1} , i.e., $X_{n+1} = 2X_n - X_0$.

$$\Delta_m = \max_{1 < i < n} \{ \Delta_i \} = \max_{1 < i < n} \{ P(X_i) - P(X_{i-1}) \}.$$

S_m = the direction corresponding to Δ_m .

Then S_{n+1} will take place of S_m . Otherwise, the original search directions set will still be used in the iterative search for a solution.

The one-dimensional search along each direction is conducted first by extrapolation to determine the search interval for optimal step length and then by quadratic interpolation to decide the optimal step length a^* .

The rule for convergence of the unconstrained minimization is:

$$\| X^{(k)} - X^{(k-1)} \| \leq \epsilon_1 \quad (5.16)$$

If $X^{(k)}$ is still not the constrained optimal point, then a penalty function $P(X, r^{(k+1)})$ will be constructed by a new penalty factor $r^{(k+1)}$, and will be manipulated for the search of its unconstrained extreme point. By repeatedly doing this, a series of points will be generated:

$$X(r^{(1)}), X(r^{(2)}), \dots, X(r^{(k-1)}), X(r^{(k)})$$

And when the penalty function satisfies:

$$|[P(X(r^{(k)})) - P(X(r^{(k-1)}))]/P(X(r^{(k-1)}))| \leq \epsilon_2, \quad (5.17)$$

$X(r^{(k)})$ is considered as the constrained optimal solution to the optimization problem (5.9).

Using the SUMT, the optimal operation problem (5.8) can be solved. Some of the relevant subroutines in (WAN, 1983) have been employed in designing a program OPWAP (OPTimal Water Apportioning) to cater for this algorithm.

5.2.3 Application Results

This algorithm has been applied to the Pudong water supply and distribution system, Shanghai, China. As mentioned in Chapter 4, it contains three water purification plants (each plant with a pumping station discharging the purified water into the distribution system). There are no reservoirs in the network.

The computed optimal apportioning of water among the different pumping stations for the 12-hour period of one day is listed in Table 5.1; the actual proportions of water are also listed in Table 5.1 so as to allow a comparison.

From Table 5.1, for the 12 hour period, 108.57 Yuan of cost can be saved. In other words, the cost of water supply can be reduced by 11.65% by using this optimization algorithm. This is proved, more or

less, to be true in other related studies.

The optimal pump schedules are produced based on the following principle:

Once the optimal pumping station flows $\{Q^*(i), i=1, \dots, n\}$ are found at each time interval, the corresponding delivery head $\{H_d^*(i), i=1, \dots, n\}$ can be computed from eqn(4.5). In pumping station i , there could be several single pumps or pump combinations of which the actual delivery head is equal to or higher than $H_d^*(i)$ when supplying flow $Q^*(i)$. They are first considered as feasible pumps (or combinations). Then the single pump (or combinations) will be chosen in operation among those feasible pumps (or combinations) which has the least cost. The procedure will be conducted in each station at each time interval to form the optimal pump schedules of all pumping stations for the whole control period.

It should be clear now that in this algorithm the optimal operations of a water supply and distribution system are carried out in two levels as illustrated in Fig 5.1.

Table 5.1 OPTIMIZATION RESULTS OF PUDONG WATER SUPPLY SYSTEM

	station 1 discharge ($10^3 \text{xm}^3/\text{hr}$)		station 2 discharge ($10^3 \text{xm}^3/\text{hr}$)		station 3 discharge ($10^3 \text{xm}^3/\text{hr}$)		total cost of water supply (Yuan)	
	actual	minimized	actual	minimized	actual	minimized	actual	minimized
0-1 hr	3.2480	2.6741	0.8970	0.5255	0.6760	1.6210	64.21	51.37
1-2 hrs	1.7360	1.6765	0.6180	0.4384	0.6760	0.9147	39.83	33.92
2-3 hrs	1.7360	1.6765	0.6180	0.4384	0.6760	0.9147	39.83	33.92
3-4 hrs	1.7360	1.6765	0.6180	0.4384	0.6760	0.9147	39.83	33.92
4-5 hrs	1.7360	1.5241	0.8970	0.7817	0.6170	0.9443	47.52	44.49
5-6 hrs	3.2080	2.8140	0.8970	0.6403	1.6910	1.3149	67.24	60.29
6-7 hrs	3.6340	3.0542	1.5580	1.4247	1.0110	1.7243	112.28	91.77
7-8 hrs	4.4360	4.1612	1.6300	1.5701	1.0760	1.4101	116.64	105.09
8-9 hrs	4.2880	4.6004	1.6300	0.9688	1.1190	1.4740	101.00	95.08
9-10 hrs	4.2140	5.3591	1.6450	0.8926	1.1190	0.7262	100.64	88.49
10-11 hrs	4.3620	5.3131	1.6300	0.9543	1.1190	0.8846	101.62	91.68
11-12 hrs	4.3620	5.1130	1.6300	1.0690	1.1190	0.9121	101.33	93.36
total	38.6960	39.6430	14.2410	10.1422	10.5750	13.7145	931.97	823.38

note: Both actual and minimized cost of water supply are computed from eqn(5.2)

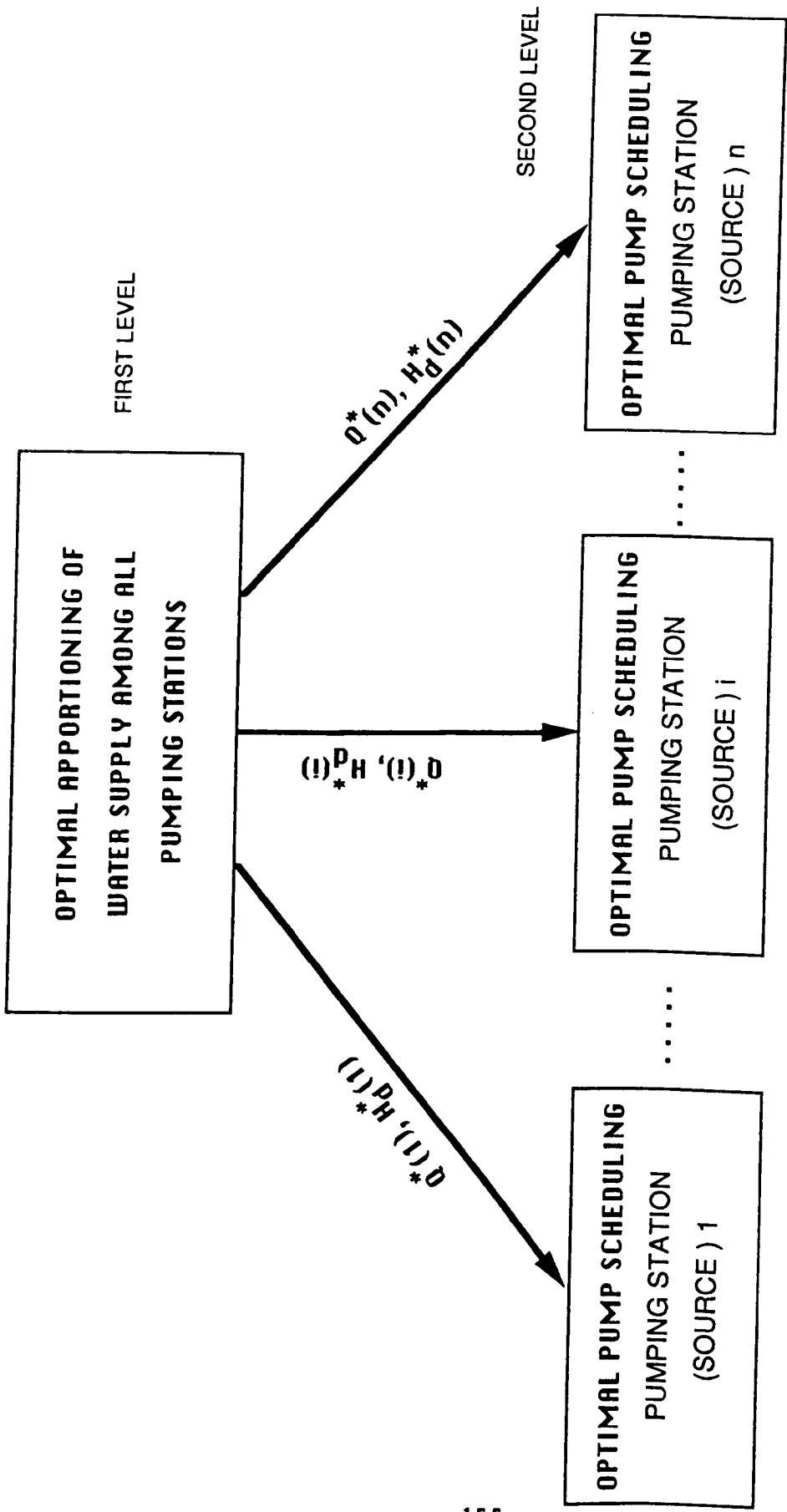


Fig. 5.1 Structure for Optimal Operations

5.3 ALGORITHM FOR OPTIMAL OPERATIONS OF A CLASS OF MULTI-SOURCE, MULTI-RESERVOIR SYSTEMS

From eqn(5.8), we may see that the objective function (cost) of the optimal operation problem is a nonlinear function about pumping flow. From eqn(5.1), we may see that the objective function is even a non-linear function about pumping flow and pumping head. These are generally true and are the main difficulty of optimization. Since some of the constraints are also non-linear, for this kind of nonlinear programming problem, finding global optima cannot be guaranteed, usually only local optima can be found (LASDON, 1970; PIERRE, 1969).

However, there is a class of multi-source, multi-reservoir systems in reality, in which pumping flows and hence pumping head do not vary much during the whole control period, even though demands vary. In these systems, the influence of demand variations can be compensated for by variations of reservoir storages (in these systems, reservoir storages must be significant).

The pumping cost of a pumping station can be expressed as:

$$\begin{aligned} \text{Cost} &= \sum_e P \times UC \times T \\ &= \sum_e (gHQ / \eta) \times UC \times T \quad (\text{e.g., } \pounds) \end{aligned} \quad (5.18)$$

where:

e = pump combination.

P = pump consumed power.(e.g., kWh)

UC = unit charge (cost) of electricity, which is usually a

piecewise constant function.(e.g., £/kWh)

T = time of pumping. (e.g., h)

H = pumping head increase. (e.g., m)

Q = pumping flow.(e.g., l/s)

η = pumping efficiency.(%)

g = unit conversion factor for electrical power relating water quantities to electrical energies.(e.g., 0.98 kWh/m³)

From eqn(5.18), for certain single pumps or pump combinations (note that, a single pump is considered to be a type of combination hereafter), if the flow is a constant, and since H and η can be calculated from the characteristic equations (as described in Chapter 2). Then (gHQ/η) is also a constant, although it is generally a nonlinear function about Q and H. In this way, the pumping cost in eqn(5.18) is a linear function about time of pumping for each pump combination.

Unlike the algorithm presented in section 5.2 , and other optimization approaches, where the time of pumping for a certain pump combination is the length of the time interval at each stage, here the time interval is made up by the time of pumping for several pump combinations and is itself a decision variable.

5.3.1 Formulation of Optimization Problem

(1). Objective Function

The total cost of pumping for the whole control period in a water supply distribution system can be expressed as:

$$TC = \sum_{l=1}^{L_p} \sum_{k=1}^{K_T} \sum_{j=1}^{j(1)} C_p(j,l,k) \times UC(l,k) \times T(j,l,k) \quad (5.19)$$

where:

L_p = number of pumping stations.

K_T = number of time stages in the whole control period.

$j(1)$ = number of pump combinations in pumping station 1.

$C_p(j,l,k)$ = pumping power consumption for pump combination 1 in pumping station j at time stage k . (e.g., kW)

$UC(l,k)$ = unit electricity charge (cost) for pumping station 1 at time stage k . (e.g., £/kWh)

$T(j,l,k)$ = time of pumping at stage k for pump combination j in pumping station 1. (e.g., h)

From the above discussion, it is clear that TC is a linear function about time of pumping $T(j,l,k)$.

(2). Constraints

a) on time of pumping

At each time stage for each pumping station, the times of pumping

added together should be equal to the length of that time stage, i.e.:

$$\sum_{j=1}^{j(1)} T(j,1,k) = TT(k) \quad (5.20)$$

$$l = 1, 2, \dots, Lp$$

$$k = 1, 2, \dots, KT$$

where:

$$TT(k) = \text{length of time stage } k \text{ (e.g., h)}$$

b) on reservoir capacities

$$XMIN(i) \leq X(i,k) \leq XMAX(i) \quad (5.21a)$$

$$i = 1, \dots, IR$$

$$k = 1, \dots, KT-1$$

where:

$$X(i,k) = \text{storage of reservoir } i \text{ at time stage } k. (\text{e.g., } m^3)$$

$$XMIN(i) = \text{lower operational bound of reservoir } i. (\text{e.g., } m^3)$$

$$XMAX(i) = \text{upper operational bound of reservoir } i. (\text{e.g., } m^3)$$

$$IR = \text{number of reservoirs in the system.}$$

and

$$XKT(i) - ALPHA1(i) \leq X(i,KT) \leq XKT(i) + ALPHA2(i), \quad (5.21b)$$

$$i = 1, 2, \dots, IR$$

where:

$$XKT(i) = \text{final storage of reservoir } i, \text{ which is usually prescribed or set to the initial storage since water supply and distribution systems are operated somewhat in a cyclic manner. (e.g., } m^3)$$

ALPHA1(i) = lower permissible deviation from reservoir final storage XKT(i). (e.g., m³)

ALPHA2(i) = upper permissible deviation from reservoir final storage XKT(i), for strict condition, ALPHA2(i) and ALPHA1(i) are set to zero, then eqn(5.21b) becomes a set of equality constraints. (e.g., m³)

From the principle of reservoir mass balance stated in Chapter 2, the following relationship should hold, (which is also referred to as a set of state equations):

$$X(i,k) = X(i,k-1) + \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)Q(j,l,k)T(j,l,k) - d(i,k) + \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k)TT(k) \quad (5.22)$$

and $X(i,0) = X0(i)$.

where:

$d(i,k)$ = demand related to reservoir i at time stage k. (e.g., m³)

$QC(ic,k)$ = ic-th constant flow (non-decision variable) at time stage k. (e.g., m³/h)

IQC = number of constant flows in the system.

$X0(i)$ = initial storage of reservoir i, which is usually given or known in advance. (e.g., m³)

$$E(i,l) = \begin{cases} 1, & \text{pumping station } l \text{ supplies water to reservoir } i; \\ -1, & \text{pumping station } l \text{ abstracts water from reservoir } i; \\ 0, & \text{otherwise.} \end{cases}$$

$$F(i,ic) = \begin{cases} 1, & \text{constant flow } ic \text{ supplies water to reservoir } i; \\ -1, & \text{constant flow } ic \text{ abstracts water from reservoir } i; \\ 0, & \text{otherwise.} \end{cases}$$

Further, since reservoir initial storage $X_0(i)$ is given or known in advance, then by expanding eqn(5.22), we have:

$$\begin{aligned}
 X(i,1) = X_0(i) + \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)Q(j,l,1)T(j,l,1) \\
 - d(i,1) + \sum_{ic=1}^{IQC} F(i,ic)QC(ic,1)TT(1) \quad (5.23)
 \end{aligned}$$

$$\begin{aligned}
 X(i,2) = X_0(i) + \sum_{k1=1}^2 \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)Q(j,l,k1)T(j,l,k1) \\
 - \sum_{k1=1}^2 d(i,k1) + \sum_{k1=1}^2 \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \quad (5.24)
 \end{aligned}$$

... ..

$$\begin{aligned}
 X(i,KT) = X_0(i) + \sum_{k1=1}^{KT} \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)Q(j,l,k1)T(j,l,k1) \\
 - \sum_{k1=1}^{KT} d(i,k1) + \sum_{k1=1}^{KT} \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \quad (5.25)
 \end{aligned}$$

or in general:

$$\begin{aligned}
 X(i,k) = X_0(i) + \sum_{k1=1}^k \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)q(i,l,k1)T(j,l,k1) \\
 - \sum_{k1=1}^k d(i,k1) + \sum_{k1=1}^k \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \\
 i=1,2,\dots,IR \\
 k=1,2,\dots,KT \quad (5.26)
 \end{aligned}$$

From the above expansions, the storage of reservoir i at time stage k , $X(i,k)$, is now expressed in terms of decision variables and other known quantities but excludes the previous reservoir storage $X(i,k-1)$. These expansions make the state equations more convenient for mathematical manipulation.

Substituting eqn(5.26) into eqn(5.21a) gives:

$$\begin{aligned}
& X0(i) + \sum_{k1=1}^k \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)Q(j,l,k1)T(j,l,k1) \\
& - \sum_{k1=1}^k d(i,k1) + \sum_{k1=1}^k \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \\
& \geq XMIN(i) \\
& i = 1,2,\dots,IR \\
& k = 1,2,\dots,KT-1
\end{aligned} \tag{5.27}$$

or:

$$\begin{aligned}
& \sum_{k1=1}^k \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)Q(j,l,k1)T(j,l,k1) \geq XMIN(i) - X0(i) \\
& + \sum_{k1=1}^k d(i,k1) - \sum_{k1=1}^k \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \\
& i = 1,2,\dots,IR \\
& k = 1,2,\dots,KT-1
\end{aligned} \tag{5.28}$$

and

$$\begin{aligned}
& \sum_{k1=1}^k \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,l)Q(j,l,k1)T(j,l,k1) \leq XMAX(i) - X0(i) \\
& + \sum_{k1=1}^k d(i,k1) - \sum_{k1=1}^k \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \\
& i = 1,2,\dots,IR \\
& k = 1,2,\dots,KT-1
\end{aligned} \tag{5.29}$$

Substituting eqn(5.26) into eqn(5.21b) results in:

$$\begin{aligned}
& \sum_{k1=1}^{KT} \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,1)Q(j,1,k1)T(j,1,k1) \geq XKT(i) - ALPHA1(i) \\
& - X0(i) + \sum_{k1=1}^{KT} d(i,k1) - \sum_{k1=1}^{KT} \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \\
& i = 1,2,\dots,IR
\end{aligned} \tag{5.30}$$

and

$$\begin{aligned}
& \sum_{k1=1}^{KT} \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,1)Q(j,1,k1)T(j,1,k1) \leq XKT(i) + ALPHA2(i) \\
& - X0(i) + \sum_{k1=1}^{KT} d(i,k1) - \sum_{k1=1}^{KT} \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1) \\
& i = 1,2,\dots,IR
\end{aligned} \tag{5.31}$$

Eqn(5.28) to eqn(5.31) are all linear functions about $T(j,1,k)$.

c) on source capacities

At each stage, for source pumping stations the amount of water abstracted should be less than the maximum capacities of that source, i.e.:

$$\sum_{j=1}^{j(1)} Q(j,1,k)T(j,1,k) \leq V_{\max}(1,k)$$

$$l = 1,2,\dots,Lp$$

$$k = 1,2,\dots,KT$$

(5.32)

where:

$V_{\max}(1,k)$ = maximum allowable amount of water to be abstracted by source pumping station 1 at time stage k.(e.g., m^3)

Obviously these are also linear functions about $T(j,1,k)$.

In conclusion, eqn(5.19), eqn(5.20), eqn(5.28), eqn(5.29), eqn(5.30), eqn(5.31) and eqn(5.32) constitute a large-scale linear programming problem summarized as follows:

$$\min_{T(j,1,k)} TC = \sum_{l=1}^{Lp} \sum_{k=1}^{KT} \sum_{j=1}^{j(1)} C_p(j,1,k) UC(1,k) T(j,1,k)$$

subject to

$$\sum_{j=1}^{j(1)} T(j,1,k) = TT(k)$$

$$l = 1, 2, \dots, Lp$$

$$k = 1, 2, \dots, KT$$

$$\sum_{k1=1}^k \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,1) Q(j,1,k1) T(j,1,k1) \geq XMIN(i) - X0(i)$$

$$+ \sum_{k1=1}^k d(i,k1) - \sum_{k1=1}^k \sum_{ic=1}^{IQC} F(i,ic) QC(ic,k1) TT(k1)$$

$$i = 1, 2, \dots, IR$$

(5.33)

$$k = 1, 2, \dots, KT-1$$

$$\sum_{k1=1}^{KT} \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,1) Q(j,1,k1) T(j,1,k1) \geq XKT(i) - ALPHA1(i)$$

$$- X0(i) + \sum_{k1=1}^{KT} d(i,k1) - \sum_{k1=1}^{KT} \sum_{ic=1}^{IQC} F(i,ic) QC(ic,k1) TT(k1)$$

$$i = 1, 2, \dots, IR$$

$$\sum_{k1=1}^k \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,1) Q(j,1,k1) T(j,1,k1) \leq XMAX(i) - X0(i)$$

$$+ \sum_{k1=1}^k d(i,k1) - \sum_{k1=1}^k \sum_{ic=1}^{IQC} F(i,ic) QC(ic,k1) TT(k1)$$

$$i = 1, 2, \dots, IR$$

$$k = 1, 2, \dots, KT-1$$

$$\sum_{k1=1}^{KT} \sum_{l=1}^{Lp} \sum_{j=1}^{j(1)} E(i,1)Q(j,1,k1)T(j,1,k1) \leq XKKT(i) + ALPHA2(i)$$

$$- X0(i) + \sum_{k1=1}^{KT} d(i,k1) - \sum_{k1=1}^{KT} \sum_{ic=1}^{IQC} F(i,ic)QC(ic,k1)TT(k1)$$

$$i = 1, 2, \dots, IR$$

$$\sum_{j=1}^{j(1)} Q(j,1,k)T(j,1,k) \leq V_{\max}(1,k)$$

$$l = 1, 2, \dots, Lp$$

$$k = 1, 2, \dots, KT$$

This is a large-scale linear programming problem with $\sum_{l=1}^{Lp} KT$ decision variables and $2KT(Lp+IR)$ constraints, for which the Revised Simplex Method is employed.

In practice, it worth noting that it is not easy to determine the reservoir related demand, $d(i,k)$, in eqn(5.22) from an original detailed network; it is also difficult to derive the proportions a pumping station supplies to different reservoirs, particularly for a strongly interactive or coupled network. These will hinder the application of this algorithm.

However, these difficulties will be overcome if the equivalent network modelling technique developed in Chapter 4 is employed. This is because once the equivalent network model for an original detailed network is obtained, the fractional consumption rate to the total demand at reservoir nodes, hence the reservoir related demand, is known. Also, with the much simplified equivalent network, the proportion a pumping station supplies to different reservoirs can be

easily and clearly derived. Therefore it is concluded that this algorithm should be applied based on the equivalent network model. This will be further illustrated in the application results of this algorithm.

5.3.2 Summary of the Revised Simplex Method

The Revised Simplex Method was developed by Dantzig and Orchard-Hays in 1953. In many aspects it employs the ideas underlying the Simplex Method but it has the advantage that it only calculates those quantities that are actually needed. It computes the simplex multipliers and the inverse of the basis directly, whereas these were only apparent indirectly from the Simplex Method in solving linear programming problems (LASDON, 1970; MURTAGH, 1981).

A general Linear Programming (L.P.) problem can be put in the form:

$$\begin{aligned}
 &\text{minimize } Z = C_1X_1 + C_2X_2 + \dots + C_nX_n \\
 &\text{subject to } a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n = b_1 \\
 &\qquad\qquad a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n = b_2 \\
 &\qquad\qquad \cdot \qquad\qquad\qquad \cdot \\
 &\qquad\qquad \cdot \qquad\qquad\qquad \cdot \\
 &\qquad\qquad \cdot \qquad\qquad\qquad \cdot \\
 &\qquad\qquad \cdot \qquad\qquad\qquad \cdot \\
 &\qquad\qquad a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n = b_m \\
 &\text{and } X_1 \geq 0, X_2 \geq 0, \dots, X_n \geq 0
 \end{aligned} \tag{5.34}$$

where:

b_i 's, c_i 's, a_{ij} 's = fixed real constants.

X_i 's = real variables to be determined.

Any other forms of L.P. problems can be transformed to the above

standard form, by introducing additional variables (BUNDAY, 1984).

In more compact vector notation, this standard problem becomes:

$$\begin{aligned}
 &\text{minimize } Z = \underline{C} \underline{X} \\
 &\text{subject to } \underline{A} \underline{X} = \underline{b} \\
 &\qquad \qquad \underline{X} \geq \underline{0}
 \end{aligned} \tag{5.35}$$

where:

$$\begin{aligned}
 \underline{C} &= (C_1, C_2, \dots, C_n) \\
 \underline{X} &= (X_1, X_2, \dots, X_n)^T \\
 \underline{A} &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \\
 \underline{b} &= (b_1, b_2, \dots, b_n)^T \\
 \underline{0} &= (0, 0, \dots, 0)^T
 \end{aligned}$$

Let \underline{B} denote the submatrix of the original \underline{A} matrix consisting of the m columns of \underline{A} corresponding to the basic variable, \underline{B} is referred to as the basis matrix.

Without loss of generality, assume that \underline{B} consists of the first m columns of \underline{A} . Then by partitioning \underline{A} , \underline{X} and \underline{C} as:

$$\begin{aligned}
 \underline{A} &= [\underline{B}, \underline{D}] \\
 \underline{X} &= [\underline{X}_B^T, \underline{X}_D^T]^T, \quad \underline{C} = [\underline{C}_B, \underline{C}_D]
 \end{aligned}$$

the standard linear programming problem becomes:

$$\begin{aligned}
 &\text{minimize} && \underline{C}_B \underline{X}_B + \underline{C}_D \underline{X}_D \\
 &\text{subject to} && \underline{B} \underline{X}_B + \underline{D} \underline{X}_D = \underline{b} \\
 &&& \underline{X}_B \geq \underline{0}, \underline{X}_D \geq \underline{0}
 \end{aligned} \tag{5.36}$$

The basic solution which is also assumed feasible, corresponding to the basis \underline{B} is $\underline{X} = [\underline{X}_B^T, \underline{0}^T]^T$ where $\underline{X}_B = \underline{B}^{-1}\underline{b}$. The basic solution results from setting $\underline{X}_D = \underline{0}$. However, for any value of \underline{X}_D the necessary value of \underline{X}_B can be computed from (5.36) as:

$$\underline{X}_B = \underline{B}^{-1} \underline{b} - \underline{B}^{-1} \underline{D} \underline{X}_D \tag{5.37}$$

and this general expression when substituted in the objective function (or cost function) yields:

$$\begin{aligned}
 z &= \underline{C}_B (\underline{B}^{-1} \underline{b} - \underline{B}^{-1} \underline{D} \underline{X}_D) + \underline{C}_D \underline{X}_D \\
 &= \underline{C}_B \underline{B}^{-1} \underline{b} + (\underline{C}_D - \underline{C}_B \underline{B}^{-1} \underline{D}) \underline{X}_D
 \end{aligned} \tag{5.38}$$

which expresses the cost of any solution to (5.36) in terms of \underline{X}_D .

Thus

$$\underline{r} = \underline{C}_D - \underline{C}_B \underline{B}^{-1} \underline{D} \tag{5.39}$$

is the relative cost vector. It is the components of this vector that are used to determine which vector to bring into the basis.

Given the inverse \underline{B}^{-1} of a current basis, and the current solution \underline{X}_B

= $\underline{Y}_0 = \underline{B}^{-1} \underline{b}$, the conceptual procedure of the Revised Simplex Method is this:

Step 1 Calculate the current relative cost coefficients $\underline{r} = \underline{C}_D - \underline{C}_B \times \underline{B}^{-1} \underline{D}$. This can best be done by first calculating $\underline{C}_B \underline{B}^{-1}$ and then the relative cost vector $\underline{C}_D - \underline{C}_B \underline{B}^{-1} \underline{D}$, if $\underline{r} \geq \underline{0}$, the current solution is optimal.

Step 2 Determine which vector \underline{a}_j is to enter the basis by selecting the most negative cost coefficient; and calculate $\underline{Y}_j = \underline{B}^{-1} \underline{a}_j$ which gives the vector \underline{a}_j expressed in terms of the current basis.

Step 3 Calculate the ratios Y_{j0}/Y_{ij} to determine which vector is to leave the basis.

Step 4 Update \underline{B}^{-1} and the current solution $\underline{B}^{-1} \underline{b}$. Return to Step 1.

A program OPPUS (OPTimal PUmP Scheduling of multi-source and multi-reservoir Water Supply System), which employs the Revised Simplex Method, is designed for this algorithm. Two versions, which are based on the Revised Simplex Method from (BUNDAY, 1984) and the routine E02MBF for solving linear programming problem in (NAG, 1983), respectively, are available.

5.3.3 Application Results

This algorithm was applied to the NURTON/BUSHBURY zone of Woverhampton water supply and distribution system as shown in Fig 5.2.

The basic mode of operation of the Nurton/Bushbury combined system is to transfer water from the sources located in the sparsely populated western fringes of Wolverhampton to the main consumer areas which are principally fed from Bushbury reservoir (at node 500 in Fig 5.2a), whilst meeting the service requirement in the Nurton zone. Nurton reservoir (at node 600 in Fig 5.2a) and the in-line booster pumps at Tettenhall provide the necessary additional storage capacity and delivery pressure required to maintain a satisfactory level of service.

The five borehole sources in Nurton zone (namely Neachley, Cosford, Copley, Stableford and Hilton) are capable of supplying up to 40 megalitres (Ml) of water per day. The total average daily consumer demand for Nurton zone is about 15 Ml. This includes about 5 Ml/d which is currently being fed to Bridgenorth zone from Hilton pump station. The total capacity of Nurton reservoir is 22.9 Ml. The booster pumps at Tettenhall draw water from Nurton reservoir and, together with the direct supply from Neachley, supplies Bushbury reservoir with about 25 Ml/d in order to meet consumer demands. The total capacity of Bushbury reservoir is 36.8 Ml. The details of the two reservoirs are summarized in Table 5.3.

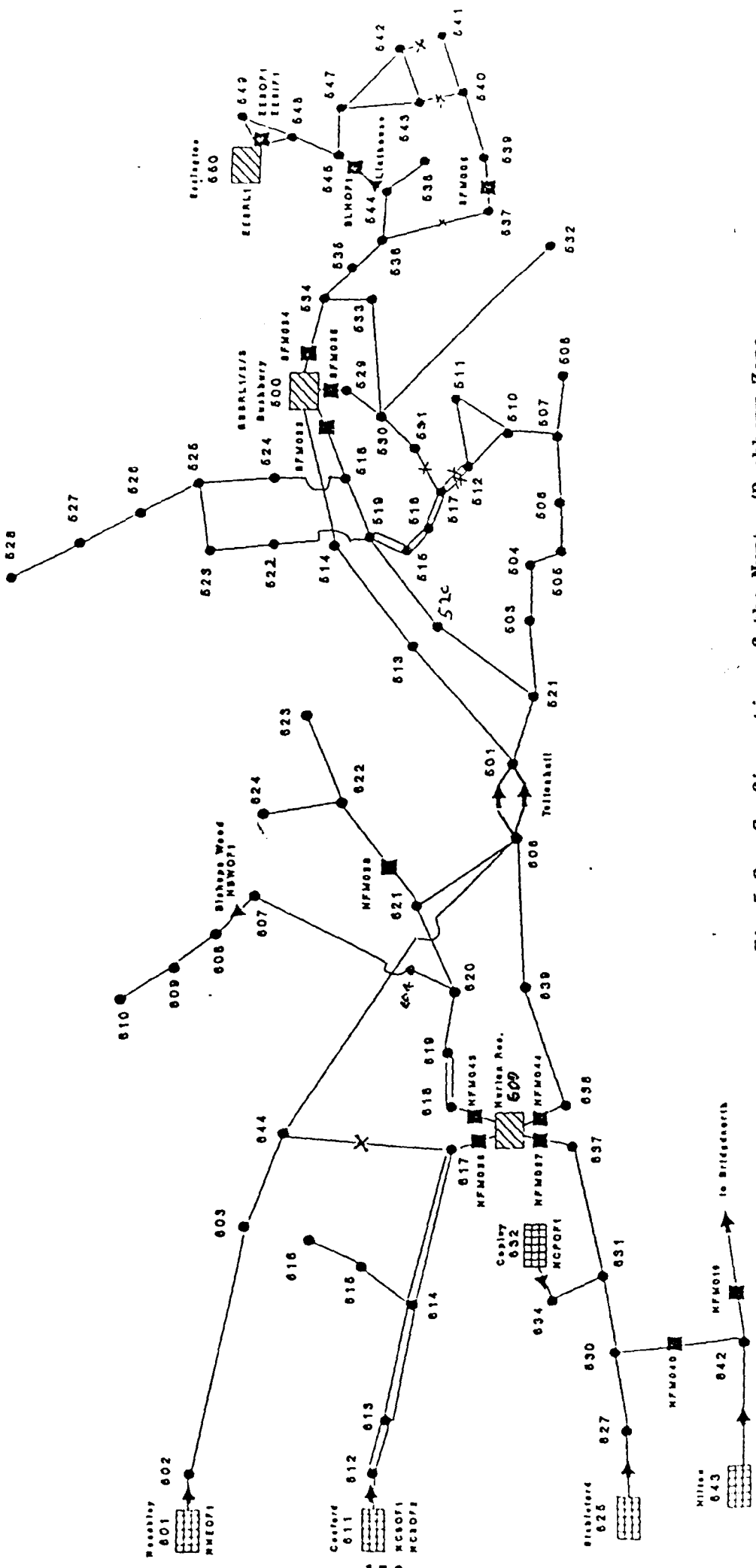


Fig 5.2a Configuration of the Nurton/Bushbury Zone

TABLE 5.3 DATA FOR RESERVOIRS IN FIGURE 5.2a

	NURTON RESERVOIR	BUSHBURY RESERVOIR
LOCATION	NODE 600	NODE 500
STORAGE (m ³)	22866.0	36822.0
CROSS-SECTIONAL AREA (m ²)	3811.0	4845.0
FLOOR ELEVATION (m)	163.0	175.3
UPPER OPERATIONAL BOUND:		
storage (m ³)	21722.7	34980.0
head (m)	168.7	182.5
LOWER OPERATIONAL BOUND:		
storage (m ³)	7622.0	13566.0
head (m)	165.0	178.1
INITIAL CONDITIONS:		
storage (m ³)	17149.5	27616.5
head (m)	167.5	181.0
FINAL CONDITIONS:		
storage (m ³)	17149.5	27616.5
head (m)	167.5	181.0
upper permissible deviation:		
storage (m ³)	1086.1	1749.1
percentage	5.0%	5.0%
lower permissible deviation:		
storage (m ³)	1086.1	1749.1
percentage	5.0%	5.0%

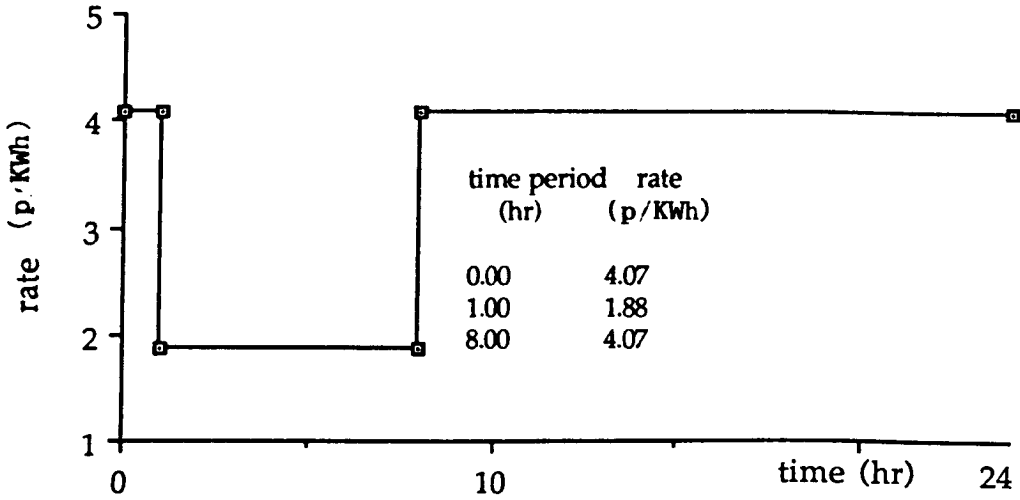
The one fixed speed pump at Cosford is capable of supplying 10 Ml/d and must be on at all times in order to meet consumer demands between Cosford station and Nurton reservoir. Cosford supply represents a significant proportion of the total water requirement for Nurton and Bushbury. The sources at Copley and Stableford each provide relatively smaller, but nevertheless significant, proportions of the total water requirements for Nurton and Bushbury. Owing to their locality within the Nurton sub-system, either of these sources can be switchable (on or off). The present status at Hilton pump station (which is currently undergoing extensive modernization) is that the fixed speed operation must remain on at all times, since part of the supply from Hilton is used to feed Bridgenorth zone. The single fixed speed pump at Neachley supplies, on average, 7.7 Ml of water daily. Current operating policy requires that Neachley pump station remains on all day in order to maintain the supply to consumers between Neachley and Tettenhall. The two variable speed booster pumps at Tettenhall are capable of pumping 230 l/s at 40 metres delivery head.

The details of all the pumps are tabulated in Table 5.2. Typical electricity tariff data are shown in Fig 5.2. Their applicabilities are indicated in Table 5.2.

TABLE 5.2 DATA FOR PUMPS IN FIGURE 5.2a

PUMP NAME	NEACHLEY	COSFORD	STABLEFORD	COPLEY	HILTON	TETTENHALL	
PUMP TYPE	FSP	FSP	FSP	FSP	FSP	VSP1	VSP2
PUMP MODEL							
A:	-0.00071	-0.00228	-0.0264	-0.01170	-0.00271	-0.00020	-0.00020
B:	-0.04250	-0.30555	-0.2583	-0.02680	-0.17220	-0.01514	-0.01514
C:	104.85	126.51	210.98	193.06	280.93	63.86	63.86
PEAK EFFICIENCY (%)	60.0	73.0	60.0	60.0	75.0	75.0	75.0
PEAK EFFICIENCY FLOW (l/s)						220.0	220.0
NOMINAL SPEED (rpm)						1550.0	1550.0
OPERATIONAL RANGE OF SPEED (rpm):							
lower						1100.0	1100.0
upper						1600.0	1600.0
TARIFF TYPE	A1	A2	A1	A1	A2	A2	A2

1) TYPE A1



2) TYPE A2

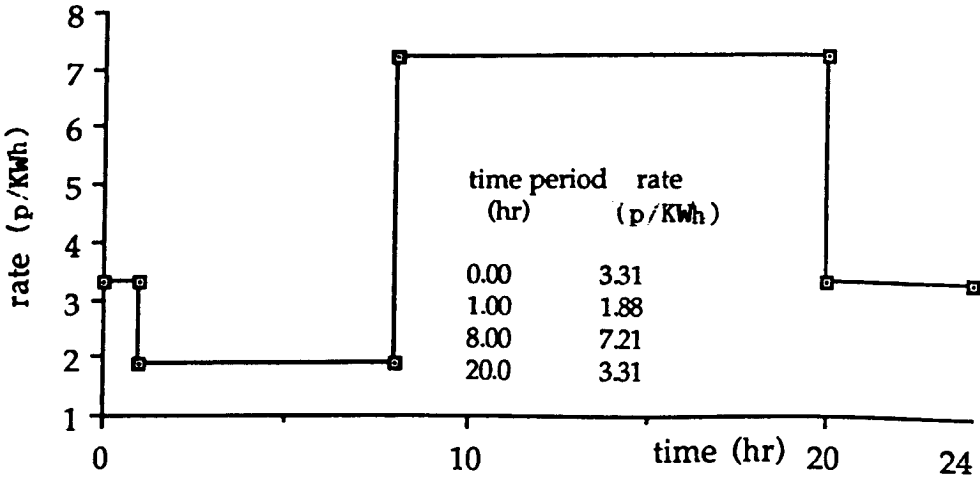


Fig 5.2 Tariff Data (unit charge)

A fully detailed model for the system, which includes the details of pumps, reservoirs, pipes, demands, etc., is available and can be used to simulate the operations of the system using GINAS5 under different operational conditions (such as FSP's on and off and VSP's operating at different speeds, etc.). Table 5.4 lists the dynamic simulation results for the system when all pumps are on, and the two identical VSP's are operating at their nominal speed of 1550 rpm. Table 5.5 and Table 5.6 list the dynamic simulation results for the system under the same condition, except that pumps at Stableford and Copley are off. From an engineering point of view, the flows and powers for all pumps in Table 5.4 could be considered as being practically constant. Furthermore, by comparing Table 5.4, Table 5.5 and Table 5.6 with each other, it can be seen that the average values of flow and power for each pump under different operational conditions are similar. Therefore, for a certain pump in the system when in operation, its flow (and thus power) is more or less constant and can be determined in advance. This suits the applicability of the presented algorithm.

TABLE 5.4 SIMULATION RESULTS FOR NURTON/BUSHBURY ZONE (ALL PUMPS ON)

TIME (hr)	PUMP FLOW AND POWER AT LOCATION												RESERVOIR ASSOCIATED DEMAND (1/s) (L/s)			
	NEAGLEY ON		COSFORD ON		HILTON ON		STABLEFORD ON		COPLY ON		TETTENHALL(1) SPEED=1550rpm			TETTENHALL(2) SPEED=1550rpm		
	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)		FLOW (1/s)	POWER (KW)	
0	87.7	196.0	113.8	95.4	136.1	378.9	38.2	136.0	44.9	139.6	186.3	134.9	186.3	134.9	105.9	122.5
2	86.7	195.7	113.9	95.3	136.1	379.0	38.3	136.0	45.0	139.6	186.1	134.9	186.1	134.9	93.5	97.0
4	86.8	195.7	114.0	95.3	136.2	379.2	38.3	136.0	45.0	139.7	185.2	134.6	185.2	134.6	92.5	96.0
6	88.8	196.3	114.4	95.1	136.7	380.3	38.4	135.9	45.3	139.8	182.8	133.9	182.8	133.9	135.5	220.1
8	91.6	199.0	115.1	94.7	140.1	388.0	39.6	135.0	46.8	140.8	174.4	131.4	174.4	131.4	295.6	438.5
10	95.5	198.3	115.4	94.5	140.1	388.1	39.6	135.0	46.9	140.9	176.7	132.0	176.1	132.0	248.1	407.9
12	95.1	198.2	115.7	94.4	139.3	386.2	39.4	135.2	46.7	140.8	177.1	132.2	177.1	132.2	226.4	382.8
14	95.3	198.3	116.0	94.2	139.5	386.7	39.5	135.1	46.9	140.9	177.1	132.2	177.1	132.2	220.6	377.1
16	95.3	198.3	116.2	94.0	139.7	387.3	39.6	135.0	47.2	141.0	177.2	132.2	177.2	132.2	217.1	396.2
18	96.7	199.0	116.7	93.8	140.0	388.1	39.7	134.9	47.4	141.2	175.0	131.6	175.0	131.6	248.5	392.1
20	95.5	198.3	116.5	93.9	139.0	385.5	39.3	135.2	47.0	140.9	176.9	132.1	176.9	132.1	180.3	232.6
22	94.8	198.1	116.6	93.8	139.1	385.7	39.4	135.2	47.1	141.0	177.1	132.2	177.1	132.2	164.1	215.3
24	92.3	197.4	116.7	93.8	138.0	383.3	39.0	135.4	46.7	140.7	178.7	132.6	178.7	132.6	107.4	119.5
AVERAGE	92.5	197.6	115.5	94.5	138.5	384.3	39.1	135.4	46.4	140.5	179.3	132.8	179.3	132.8	179.6	269.0
MAXIMUM DEVIATION* (%)	6.2	1.0	1.5	1.0	1.7	1.4	2.3	0.4	3.2	0.6	3.9	1.6	3.9	1.6	3.9	1.6

* IS CALCULATED FROM $[\text{MAX}(|\text{MINIMUM VALUE} - \text{AVERAGE}|, |\text{MAXIMUM VALUE} - \text{AVERAGE}|)] / \text{AVERAGE} \times 100\%$.

TABLE 5.5 SIMULATION RESULTS FOR NURTON/BUSHBURY ZONE (PUMP AT LOCATION NODE 625 TO NODE 627 OFF)

TIME (hr)	PUMP FLOW AND POWER AT LOCATION												RESERVOIR ASSOCIATED			
	NEACHLEY ON		COSFORD ON		HILTON ON		STABLEFORD OFF		COOPLY ON		TETTENHALL(1) SPEED=1550rpm			TETTENHALL(2) SPEED=1550rpm		
	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)		FLOW (1/s)	POWER (KW)	
0	87.7	196.0	113.8	95.4	141.3	390.9	0.0	0.0	47.6	141.3	186.3	134.9	186.3	134.9	105.9	122.5
2	87.0	195.7	144.0	95.3	141.2	391.2	0.0	0.0	47.7	141.4	186.0	134.9	186.0	134.9	93.5	97.0
4	87.0	195.7	114.1	95.2	141.5	391.5	0.0	0.0	47.8	141.4	185.0	134.5	185.0	134.5	92.5	96.0
6	89.2	196.4	114.6	95.0	142.0	392.7	0.0	0.0	48.1	141.6	182.5	133.8	182.5	133.8	135.5	220.1
8	98.0	199.1	115.4	94.5	145.1	400.5	0.0	0.0	49.5	142.5	174.0	131.2	174.0	131.2	295.5	438.5
10	96.0	198.5	115.7	94.3	145.3	400.8	0.0	0.0	49.7	142.6	176.1	131.9	176.1	131.9	248.1	407.9
12	95.7	198.4	116.1	94.2	144.6	399.2	0.0	0.0	49.6	142.5	176.5	132.0	176.5	132.0	226.4	382.8
14	96.0	198.5	116.4	94.0	144.9	399.8	0.0	0.0	49.8	142.7	176.3	131.9	176.3	131.9	220.6	377.1
16	96.1	198.5	116.8	93.7	145.2	400.6	0.0	0.0	50.1	142.8	176.4	132.0	176.4	132.0	217.1	396.2
18	98.0	199.3	117.3	93.5	145.6	400.6	0.0	0.0	50.3	143.0	174.1	131.3	174.1	131.3	248.5	392.1
20	96.5	198.7	117.2	93.5	144.7	399.3	0.0	0.0	50.1	142.8	175.9	131.8	175.9	131.8	180.3	232.6
22	95.9	198.5	117.4	93.4	144.8	399.7	0.0	0.0	50.2	142.9	176.0	131.8	176.0	131.8	164.1	215.3
24	93.5	197.8	117.4	93.4	143.9	399.4	0.0	0.0	50.2	142.7	177.5	132.3	177.5	132.3	107.4	119.5
AVERAGE	93.6	197.8	115.9	94.3	143.9	397.4	0.0	0.0	49.3	142.3	178.7	132.6	178.7	132.6	179.6	269.0
MAXIMUM DEVIATION* (%)	7.1	1.1	1.8	1.2	1.9	1.6	0.0	0.0	3.4	0.7	4.3	1.7	4.3	1.7	4.3	1.7

* IS CALCULATED FROM $[\text{MAX}(|\text{MINIMUM VALUE} - \text{AVERAGE}|), |\text{MAXIMUM VALUE} - \text{AVERAGE}|] / \text{AVERAGE} \times 100\%$.

TABLE 5.6 SIMULATION RESULTS FOR NURTON/BUSHBURY ZONE (PUMP AT LOCATION NODE 632 TO NODE 634 OFF)

TIME (hr)	PUMP FLOW AND POWER AT LOCATION												RESERVOIR ASSOCIATED DEMAND (1/s) (1/s)			
	NEACHLEY ON		COSFORD ON		HILTON ON		STABLEFORD ON		COPLEY OFF		TETTENHALL(1) SPEED=1550rpm			TETTENHALL(2) SPEED=1550rpm		
	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)	FLOW (1/s)	POWER (KW)		FLOW (1/s)	POWER (KW)	
0	87.7	196.0	113.8	95.4	139.8	387.3	39.7	134.9	0.0	0.0	186.3	134.9	186.3	134.9	105.9	122.5
2	87.0	195.7	113.9	95.3	139.9	387.6	39.8	134.8	0.0	0.0	186.0	134.8	186.0	134.8	93.5	97.0
4	87.1	195.8	114.1	95.2	140.0	387.9	39.8	134.8	0.0	0.0	184.9	134.5	184.9	134.5	92.5	96.0
6	89.3	196.4	114.6	95.0	140.5	389.2	40.0	134.6	0.0	0.0	182.4	133.7	182.4	133.7	135.5	220.1
8	97.3	199.1	115.5	94.5	143.8	397.2	41.1	133.7	0.0	0.0	173.8	131.2	173.8	131.2	215.5	438.5
10	96.2	198.6	115.9	94.3	144.0	397.6	41.2	133.6	0.0	0.0	176.0	131.8	176.0	131.8	248.1	407.9
12	95.9	198.5	116.2	94.1	143.3	395.9	41.0	133.8	0.0	0.0	176.3	131.9	176.3	131.9	226.4	382.8
14	96.2	198.6	116.6	93.9	143.6	396.6	41.1	133.7	0.0	0.0	176.1	131.9	176.1	131.9	220.6	377.1
16	96.4	198.6	117.0	93.7	143.9	397.4	41.3	133.6	0.0	0.0	177.1	131.9	177.1	131.9	217.1	396.2
18	97.4	199.4	117.5	93.4	144.3	398.4	41.4	133.4	0.0	0.0	175.3	131.2	175.3	131.2	248.5	392.1
20	96.8	198.8	117.4	93.4	143.3	396.0	41.1	133.7	0.0	0.0	174.8	131.7	174.8	131.7	180.3	232.6
22	96.3	198.6	117.7	93.3	143.4	396.4	41.2	133.6	0.0	0.0	175.0	131.7	175.0	131.7	164.1	215.3
24	94.0	197.9	117.7	93.3	142.6	394.1	41.0	133.8	0.0	0.0	177.1	132.2	177.1	132.2	107.4	119.5
AVERAGE	93.7	197.8	116.0	94.2	142.5	393.9	40.7	134.0	0.0	0.0	178.5	132.6	178.5	132.6	179.6	269.0
MAXIMUM * DEVIATION (%)	7.1	1.1	1.9	1.3	1.9	1.7	2.5	0.7	0.0	0.0	4.4	1.7	4.4	1.7		

* IS CALCULATED FROM [MAX(|MINIMUM VALUE - AVERAGE|, |MAXIMUM VALUE - AVERAGE|) / AVERAGE] X 100%.

Further, as discussed in section 5.3.2, for the application of this algorithm, it is necessary to obtain the equivalent network of this system. By using the program, ENCQP, for equivalent network modelling developed in Chapter 4, the equivalent network of the system can be obtained as shown in Fig 5.3. The numerical results are summarized in Table 5.7. Thus the reservoir related demand, and the proportion a pump supplies or abstract water to or from a reservoir (eqn(5.22)) can be easily determined.

As stated above, the pumps at Neachley, Cosford, and Hilton must be on for all of the day due to operational considerations. Therefore they are simulated by constant flows in this algorithm. The decision variables are the times of pumping for all other pumps.

Practically, pumps at Stableford, and Copley can be amalgamated into one bigger pumping station (named as pumping station 1), which consists of three possible combinations as follows:

- 1) pump at Stableford on.
- 2) pump at Copley on.
- 3) pumps at Stableford and Copley both on.

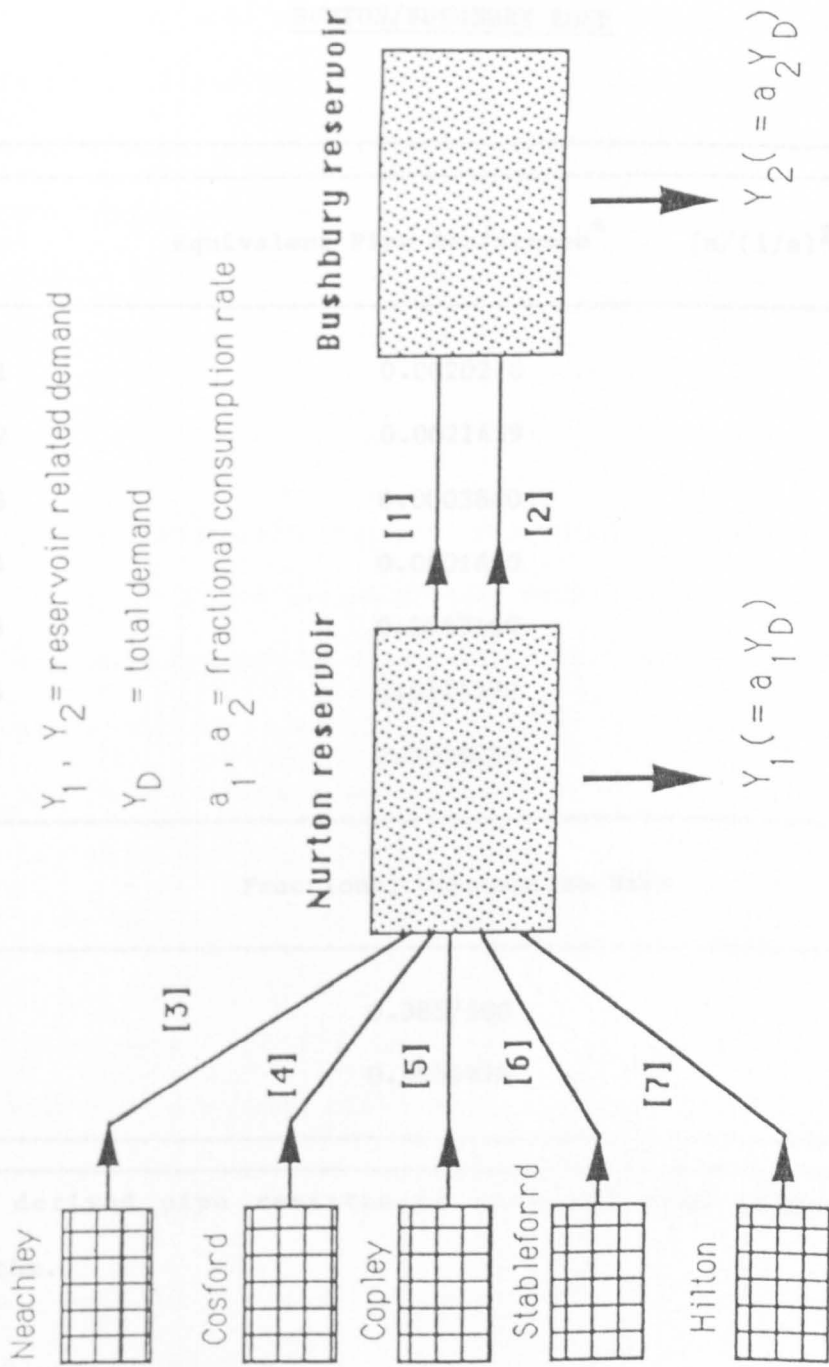


FIG. 5.3 EQUIVALENT NETWORK OF NUR/BUSHBURY ZONE

TABLE 5.7 DATA FOR THE EQUIVALENT NETWORK OF
NURTON/BUSHBURY ZONE

	Equivalent Pipe Resistance*	[m/(1/s) ²]
PIPE 1	0.0020270	
PIPE 2	0.0021459	
PIPE 3	0.0003880	
PIPE 4	0.0001850	
PIPE 5	0.0053160	
PIPE 6	0.0111570	
PIPE 7	0.0008430	
Fractional Consumption Rate		
a ₁	0.3857500	
a ₂	0.5934970	

* the derived pipe resistances were not used directly in this algorithm.

The two identical variable speed pumps at Tettenhall are in one pumping station (named as pumping station 2) and their speed range is from 1100 rpm to 1600 rpm, as shown in Table 5.2. One of the difficulties, in solving the problem of optimal control, arises from the mixture of fixed speed pumps, the controls of which are discrete, and variable speed pumps, the control of which are continuous within the speed range, (see the discussion in section 5.1). This problem has proved to be one of the main obstacles for the application of some algorithms (BRDYS' et al, 1988; COULBECK, 1977; JOALLAND and COHEN, 1980)

However, this obstacle can be overcome in this algorithm by discretizing the speed range of the variable speed pumps. Thus a variable speed pump can be discretized into a number of fixed speed pumps corresponding to a number of speed steps. In this way, this system can be transformed to a system containing discrete variables only. Variable throttle pumps (VTP's) could be dealt with in a similar way, by discretizing the range of throttle factors.

As an illustration of the treatment, the two VSP's are firstly discretized into six fixed speed pumps corresponding to speed steps of 1100 rpm, 1550 rpm, and 1600 rpm for each individual VSP.

For the above six fixed speed pumps, there could be various pump combinations. However, when the two VSP's are operating together in parallel, the least cost combinations will obviously be those combinations for which the two VSP's are operating at the same speed. This is because it is most economical when a pump is operating around

its peak efficiency, this can only be achieved for both two identical VSP's operating in parallel when they are operating at the same speed.

Table 5.8 summarizes the average values of flow and power of all economical pump combinations for the complete system. These are derived from various dynamic simulations of the system.

With the availability of the necessary data required by this algorithm, the program OPPUS has been run to solve this particular linear programming problem with 108 variables and 72 constraints (not including slack variables for mathematical manipulation of linear programming as described in 5.3.2). This requires about 1 minute of elapsed time to solve on a Prime 2550 computer.

The optimized pump operations and relevant power consumptions and costs are given in Table 5.9. The complete pump schedules are drawn in Fig 5.4. Finally, the reservoir trajectories are plotted in Fig 5.5. Unfortunately, the actual site pump schedules and costs, etc., are not available for comparison.

It is worth noting that in Fig 5.5, Bushbury reservoir is topped up overnight to take advantage of the cheap night electricity tarriff rate (refer to Fig 5.2, Type A2). However, Nurton reservoir is not topped up overnight. This could be due to the limited pump combinations and their cost factors in Station 1, and/or the tarriff rate difference between peak and off pead hours (refer to Fig 5.2, Type A1) and demand conditions in Nurton zone so that Nurton reservoir cannot be or is not necessary to be topped up overnight in this particular case.

TABLE 5.8 PUMP AVERAGE VALUES OF FLOW AND POWER

	FLOW (1/s)	POWER (KW)
<u>PUMPING STATION 1 (Stableford + Copley)</u>		
COMBINATIONS:		
1) PUMP AT STABLEFORD	39.1	135.3
2) PUMP AT COPLEY	45.7	140.5
3) PUMPS AT STABLEFORD AND AT COPLEY	84.8	275.8
<u>PUMPING STATION 2 (Tettenhall)</u>		
COMBINATIONS:		
1) ONE VSP AT SPEED 1100 rpm	181.2	57.6
2) ONE VSP AT SPEED 1550 rpm	295.3	183.6
3) ONE VSP AT SPEED 1600 rpm	307.7	204.0
4) BOTH VSP'S AT SPEED 1100 rpm	216.8	89.6
5) BOTH VSP'S AT SPEED 1550 rpm	357.5	265.8
6) BOTH VSP'S AT SPEED 1600 rpm	434.2	313.9
<u>PUMPING STATIONS WITH CONSTANT FLOWS</u>		
1) PUMP AT NEACHLEY	90.2	196.0
2) PUMP AT COSFORD	115.0	94.8
3) PUMP AT HILTON	139.1	385.9

TABLE 5.9 RESULTS FOR THE OPTIMAL OPERATIONS
OF NURTON/BUSHBURY ZONE

OPERATIONS OF PUMPING STATION 1 (Stableford + Copley)

TIME (hr.min)	0.00	18.00	18.00	18.30	18.30	24.00
PUMP COMBINATION IN OPERATION (NO.)	3	3	2	2	3	3
FLOW (1/s)	84.8	84.8	45.7	45.7	84.8	84.8

OPERATIONS OF PUMPING STATION 2 (Tettenhall)

TIME (hr.min)	0.00	1.00	1.00	8.00	8.00	20.00	20.00	22.50	22.50	24.00
PUMP COMBIN- ATION IN OPER- ATION (NO.)	1	1	6	6	1	1	4	4	1	1
FLOW (1/s)	181	181	434	434	181	181	216	216	181	181
VSP 1 SPEED (rpm)	1100	1100	1600	1600	1100	1100	1100	1100	1100	1100
VSP2 SPEED (rpm)	0	0	1600	1600	0	0	1100	1100	0	0

ENERGY CONSUMPTIONS AND MINIMIZED COSTS

PUMP NAME	ENERGY CONSUMPTION (KWH)	COST (£)
PUMP AT STABLEFORD	3179.6	108.67
PUMP AT COPLEY	3372.0	115.70
PUMPS AT TETTENHALL	3267.1	103.68
PUMP AT NEACHLEY	4704.0	161.41
PUMP AT COSFORD	2275.2	110.19
PUMP AT HILTON	9261.6	448.53
TOTAL	26059.5	1048.18

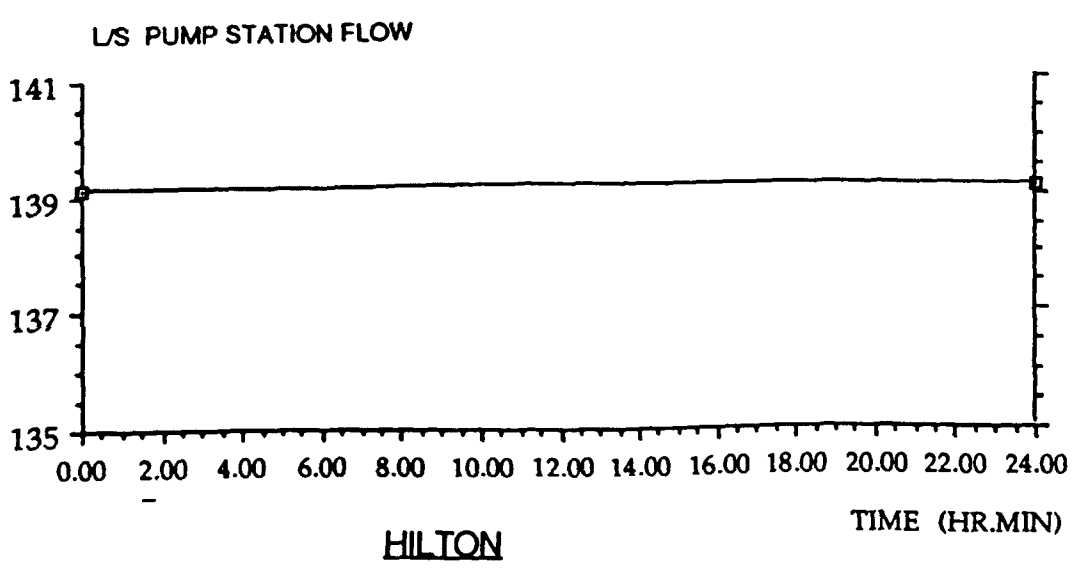
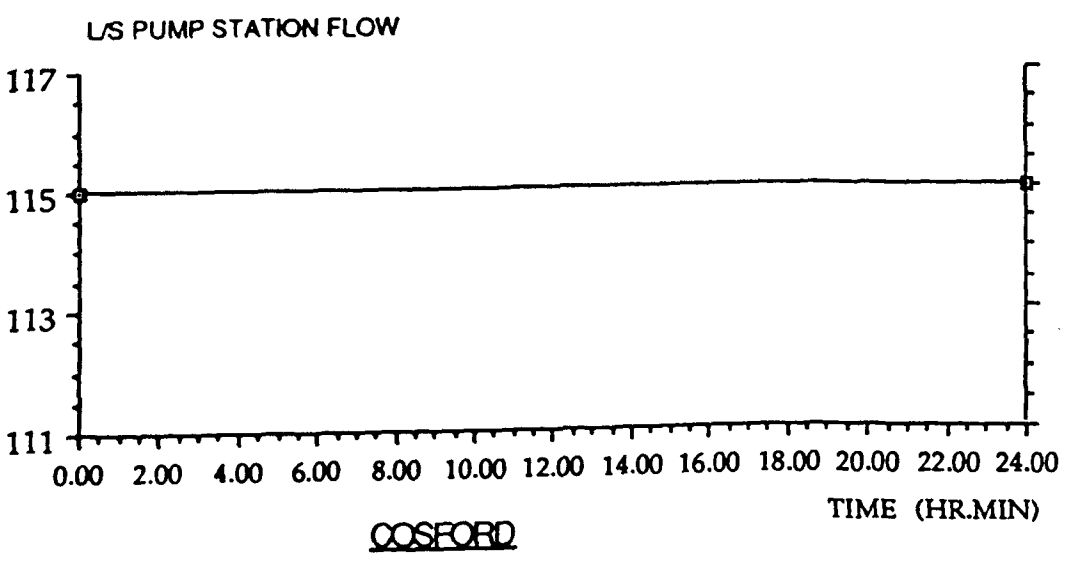
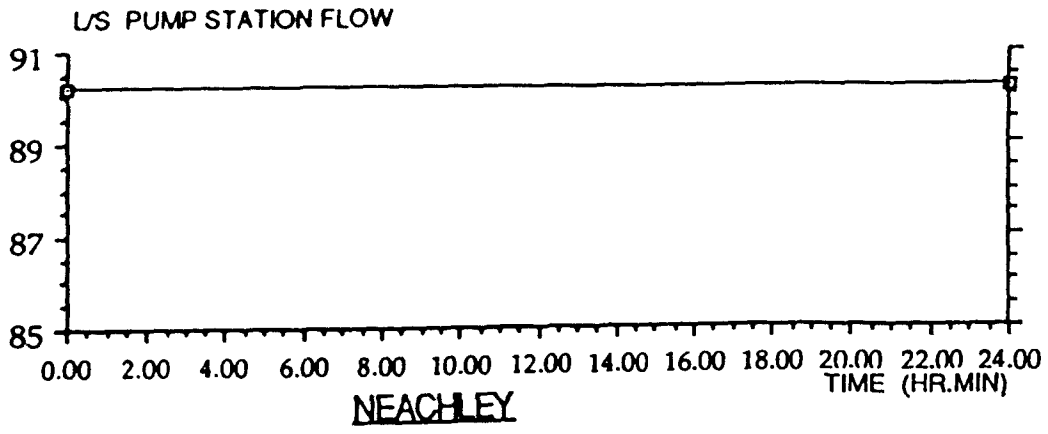
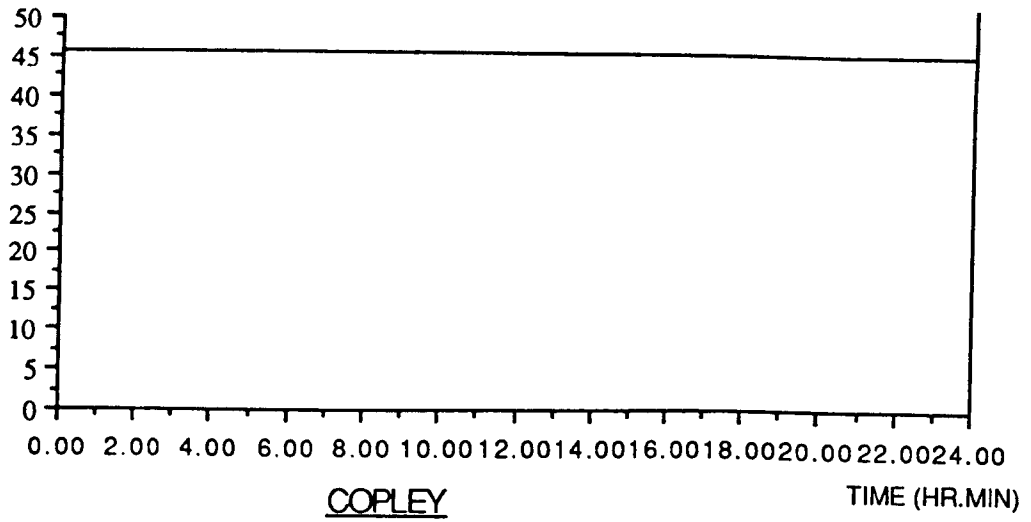
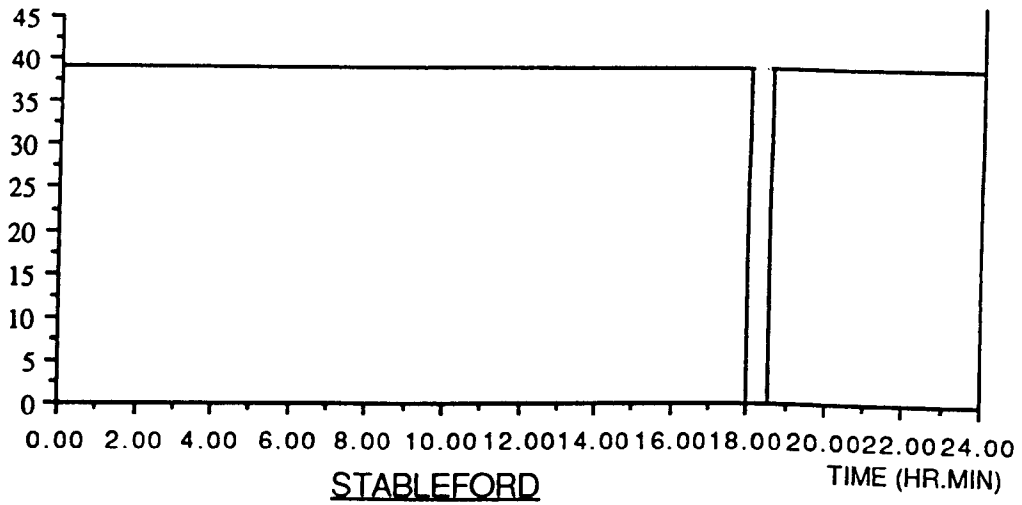


FIG 5.4 OPTIMIZED PUMP SCHEDULES

L/S PUMP STATION FLOW



L/S PUMP STATION FLOW



L/S PUMP STATION FLOW

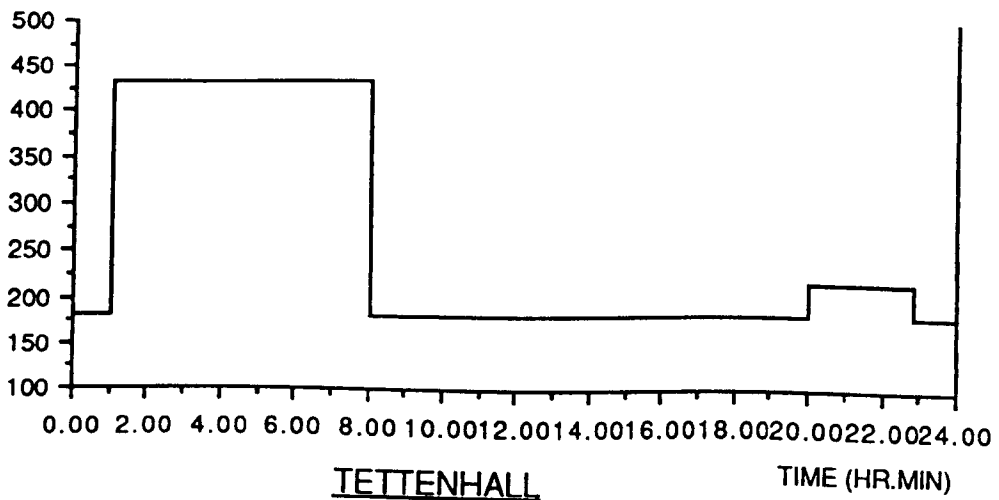


FIG 5.4 (continued)

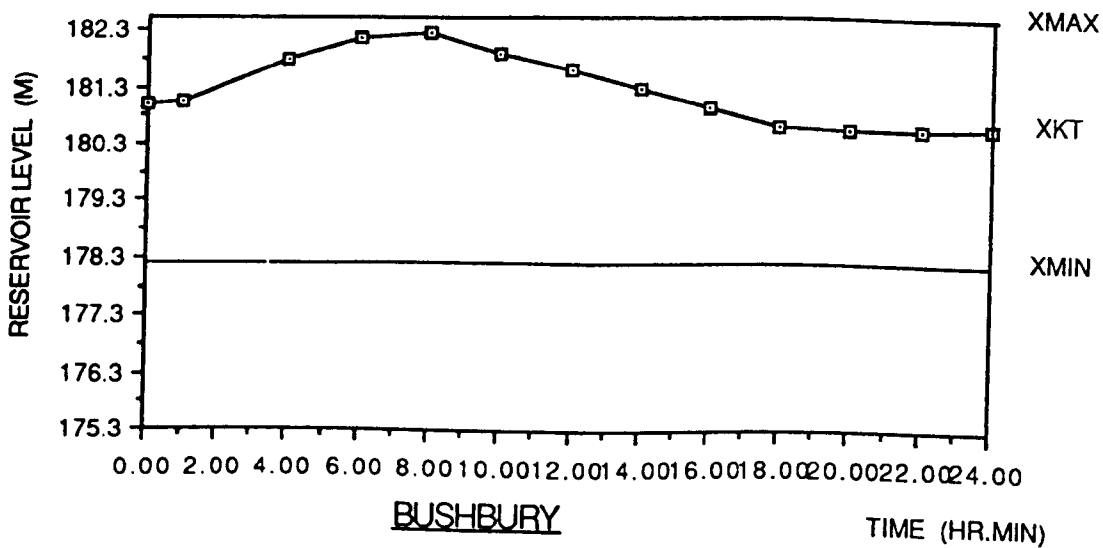
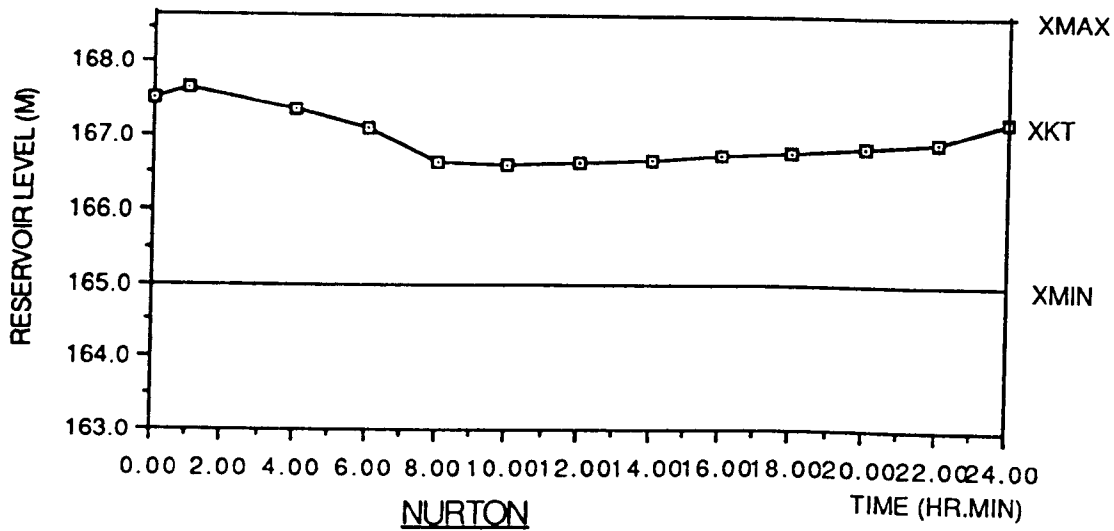


FIG 5.5 RESERVOIR TRAJECTORIES

5.3.4 Studies on the Decomposition of Variable Speed Pumps

From the above studies, it may be seen that variable speed pumps can be dealt with in this algorithm by discretizing their speed ranges. In this way, one VSP is transformed into a number of FSP's corresponding to different speed steps. This overcomes the difficulty of mixing discrete and continuous variables, which is a major difficulty in applying many algorithms. However, it is not clear how to properly discretize the speed range of a variable speed pump. Over-discretization may unnecessarily increase the numbers of decision variables (or dimensions) and increase the computation time of solving the L.P. problems dramatically. Under-discretization may result in a loss of optimality. A study of the discretization of VSP speed ranges was initially conducted through a sensitivity analysis as discussed in the following.

Generally, if some additional variables are to be taken into account in a previously solved linear programming problem, one very crude way is to change the mathematical problem to take account of the additional variables and solve the new problem 'from scratch'. However, this procedure may be very inefficient and does not take account of the useful work that has already been done in solving the problem (BUNDAY, 1984; MURTAGH, 1981). Alternatively, sensitivity analysis or post-optimality analysis is a technique for avoiding this inefficiency.

Multiplying the constraints in eqn(5.34) by numbers $\pi_1, \pi_2, \dots, \pi_m$ and adding to the objective function Z gives:

$$\begin{aligned}
& X_1(C_1 + \sum_{i=1}^m a_{i1} \pi_i) + X_2(C_2 + \sum_{i=1}^m a_{i2} \pi_i) + \dots \\
& + X_n(C_n + \sum_{i=1}^m a_{in} \pi_i) = Z + \sum_{i=1}^m b_i \pi_i \qquad (5.40)
\end{aligned}$$

We can choose the π_i so that the coefficients of the basic variables in eqn(5.40) are zero. The π_i are known as the Simplex Multipliers. If X_1, X_2, \dots, X_m are basic (there is no loss of generality here) the π_i are determined from:

$$\begin{aligned}
a_{11} \pi_1 + a_{21} \pi_2 + \dots + a_{m1} \pi_m &= -C_1 \\
a_{12} \pi_1 + a_{22} \pi_2 + \dots + a_{m2} \pi_m &= -C_2 \\
\dots & \qquad \qquad \dots \qquad \qquad \dots \\
a_{1m} \pi_1 + a_{2m} \pi_2 + \dots + a_{mm} \pi_m &= -C_m
\end{aligned} \qquad (5.41a)$$

i.e.

$$\underline{B}^T \underline{\pi} = - \underline{C}_B \qquad (5.41b)$$

where:

\underline{B} = is the matrix of coefficients of the basic variables.

$\underline{C}_B = (C_1, \dots, C_m)^T$, the coefficients of the basic variables in the first form for Z.

$$\underline{\pi} = (\pi_1, \pi_2, \dots, \pi_m)^T$$

thus

$$\underline{\pi} = -(\underline{B}^{-1})^T \underline{C}_B \qquad (5.42)$$

The value of $\underline{\pi}$ is already available upon the completion of the computation of the Revised Simplex Method.

Suppose we have solved the original problem. Suppose for the optimal

basis, its matrix of coefficients in \underline{A} is \underline{B} with inverse \underline{B}^{-1} . The values of the basic variables in the original problem will be given by

$$\underline{X}_B = \underline{B}^{-1} \underline{b} = \underline{b}' \geq 0 \quad (5.43)$$

The value of the objective function will be given by

$$z^{opt} = - \sum b_i \pi_i \quad (5.44)$$

and in eqn(5.40), for $j=1,2,\dots,n$, the following holds;

$$C_j + \sum_{i=1}^m a_{ij} \pi_i \geq 0 \quad (5.45)$$

of which the coefficients of the basic variables are 0, and the coefficients of the non-basic variables ≥ 0 .

Now we introduce an additional variable X_{n+1} , then its corresponding item in eqn(5.40) will be

$$C_{n+1} + \sum_{i=1}^m a_{i,n+1} \pi_i \quad (5.46)$$

If

$$C_{n+1} + \sum_{i=1}^m a_{i,n+1} \pi_i \geq 0, \quad \text{or} \quad C_{n+1} \geq - \sum_{i=1}^m a_{i,n+1} \pi_i,$$

X_{n+1} remains non-basic and the original solution remains optimal. If, however,

$$C_{n+1} + \sum_{i=1}^m a_{i,n+1} \pi_i < 0, \text{ or } C_{n+1} < - \sum_{i=1}^m a_{i,n+1} \pi_i$$

then we should make X_{n+1} basic to enter the basis and continue the computation.

The above summarizes the principle of inclusion of additional variables in a sensitivity analysis. This principle has been applied to study the discretization of VSP's, which is illustrated below.

If for the two VSP's, we introduce an additional speed step 1300 rpm, there will be the additional feasible pump combinations in Table 5.9a.

TABLE 5.9a ADDITIONAL PUMP COMBINATIONS AT TETTENHALL

COMBINATION	VSP1		VSP2	
	ON/OFF	SPEED (rpm)	ON/OFF	SPEED (rpm)
7	ON	1300	OFF	--
8	ON	1300	ON	1300

Their average values of flow and power obtained from dynamic simulations are listed in Table 5.9b.

TABLE 5.9b AVERAGE VALUES OF FLOW AND POWER OF
ADDITIONAL PUMP COMBINATIONS (ADDITIONAL SPEED STEP 1300 RPM)

ADDITIONAL COMBINATIONS	FLOW (l/s)	POWER (KW)
7)	272.5	122.7
8)	293.6	155.1

For each of those additional pump combinations, there will be KT additional variables to enter the linear programming problem, where KT is the number of time stages (here $KT=12$).

Based on the original solution, it can be checked by eqn(5.46) whether any of the additional variables remain non-basic or not. Table 5.10 summarizes the results.

Table 5.10 shows that all of the combinations no longer remain non-basic at some stages, which means that the additional speed step 1300 rpm is necessary and beneficial in obtaining the optimized pump schedules. Table 5.11, Fig 5.6 and Fig 5.7 show the solution for the linear programming problem including the additional variables, where there are 132 decision variables and 72 constraints. It takes about 3 minutes computation time to solve this problem, which is 3 times longer than that of the original problem with 108 variables and 72 constraints. From Table 5.11 and Fig 5.6, it can be seen that pump combination no.7 in Tettenhall station (refer to Table 5.9a) has been introduced in operation, which confirms the results of sensitivity analysis in Table 5.10.

The computed schedule for Stableford pump in Fig 5.6 (continued) requires the pump to be off for 30 minutes at 20.00hours, which could incur some difficulties and switching costs for practical application. If it is not switched off at that time, the extra cost will only be £2.75 (0.26% of system total cost) and the reservoir level of Nurton is still within operational constraints. In conclusion, for practical application of the computed schedule, the pump at Stableford can remain on all the day without increasing system total cost significantly. The schedule for Stableford pump in Fig 5.4 can be modified in a similar way.

TABLE 5.10 SENSITIVITY ANALYSIS FOR ADDITIONAL SPEED

OF 1300 RPM OF THE VSP'S

ADDITIONAL COMBINATION IP=7	C_{n+1}	m $-\sum_{I=1} a_{I,n+1}$	π_I	CONCLUSION
Time Stage				
1	6.769	7.705		BASIC
2	3.845	1.817		NON-BASIC
3	3.845	1.817		NON-BASIC
4	3.845	1.817		NON-BASIC
5	14.744	11.449		NON-BASIC
6	14.744	11.449		NON-BASIC
7	14.744	11.449		NON-BASIC
8	14.744	11.449		NON-BASIC
9	14.744	11.449		NON-BASIC
10	14.744	11.449		NON-BASIC
11	6.769	7.705		BASIC
12	6.769	7.705		BASIC
ADDITIONAL COMBINATION IP=8				
Time Stage				
1	8.556	8.751		BASIC
2	4.486	2.863		NON-BASIC
3	4.486	2.863		NON-BASIC
4	4.486	2.863		NON-BASIC
5	18.638	12.495		NON-BASIC
6	18.638	12.495		NON-BASIC
7	18.638	12.495		NON-BASIC
8	18.638	12.495		NON-BASIC
9	18.638	12.495		NON-BASIC
10	18.638	12.495		NON-BASIC
11	8.556	8.751		BASIC
12	8.556	8.751		BASIC

TABLE 5.11 RESULTS FOR THE OPTIMAL OPERATIONS
OF NURTON/BUSHBURY ZONE

OPERATIONS OF PUMPING STATION 1 (Stableford + Copley)

TIME (hr.min)	0.00	20.00	20.00	20.30	20.30	24.00
PUMP COMBINATION IN OPERATION (NO.)	3	3	2	2	3	3
FLOW (l/s)	84.8	84.8	45.7	45.7	84.8	84.8

OPERATIONS OF PUMPING STATION 2 (Tettenhall)

TIME (hr.min)	0.00	1.00	1.00	8.00	8.00	22.50	22.50	24.00
PUMP COMBIN- ATION IN OPER- ATION (NO.)	1	1	6	6	1	1	7	7
FLOW (l/s)	181	181	434	434	181	181	273	273
VSP 1 SPEED (rpm)	1100	1100	1600	1600	1100	1100	1300	1300
VSP2 SPEED (rpm)	0	0	1600	1600	0	0	0	0

ENERGY CONSUMPTIONS AND MINIMIZED COSTS

PUMP NAME	ENERGY CONSUMPTION (KWH)	COST (£)
PUMP AT STABLEFORD	3179.6	108.67
PUMP AT COPLEY	3372.0	115.70
PUMPS AT TETTENHALL	3252.7	103.19
PUMP AT NEACHLEY	4704.0	161.41
PUMP AT COSFORD	2275.2	110.19
PUMP AT HILTON	9261.6	448.53
TOTAL	26045.1	1047.69

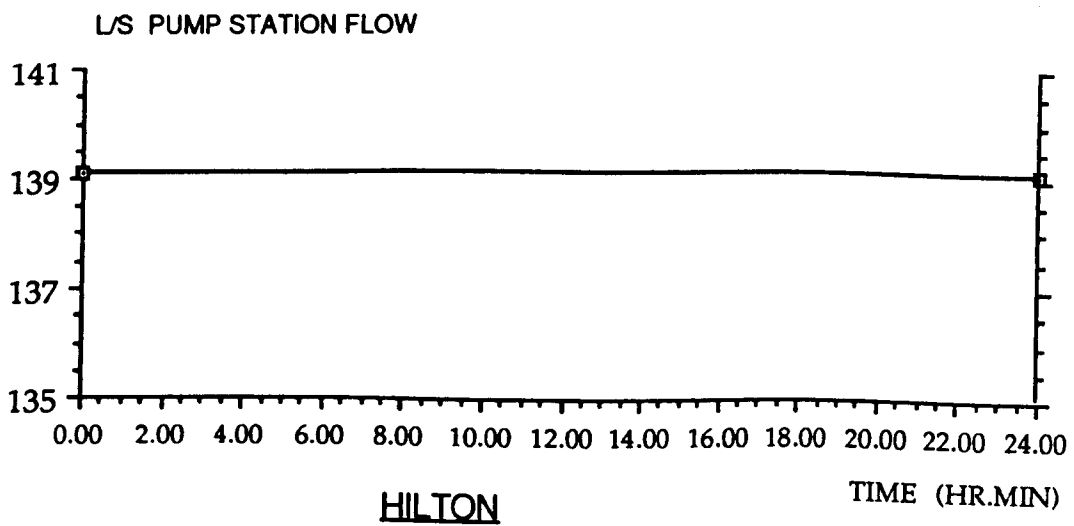
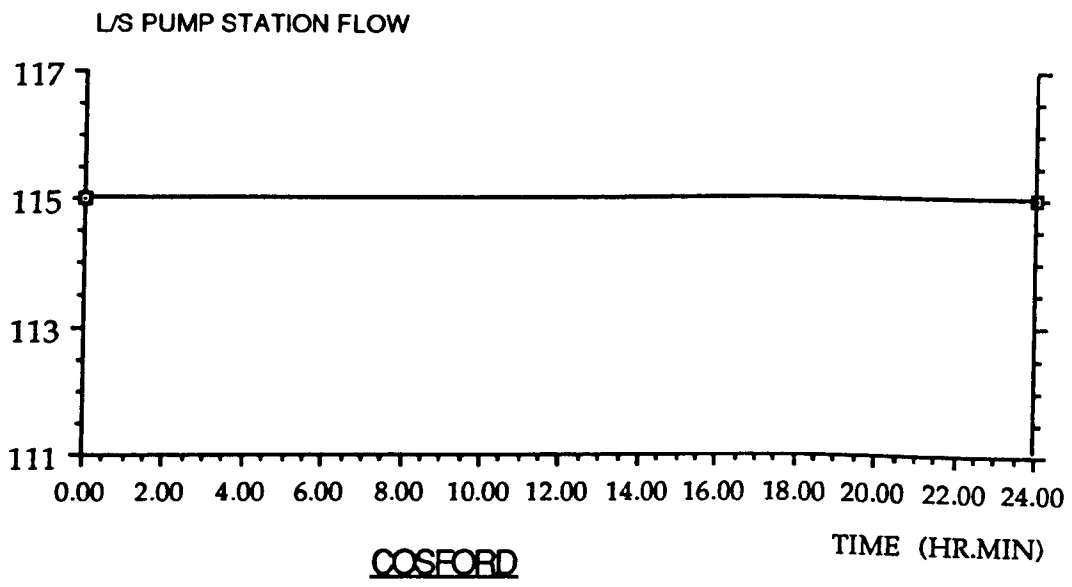
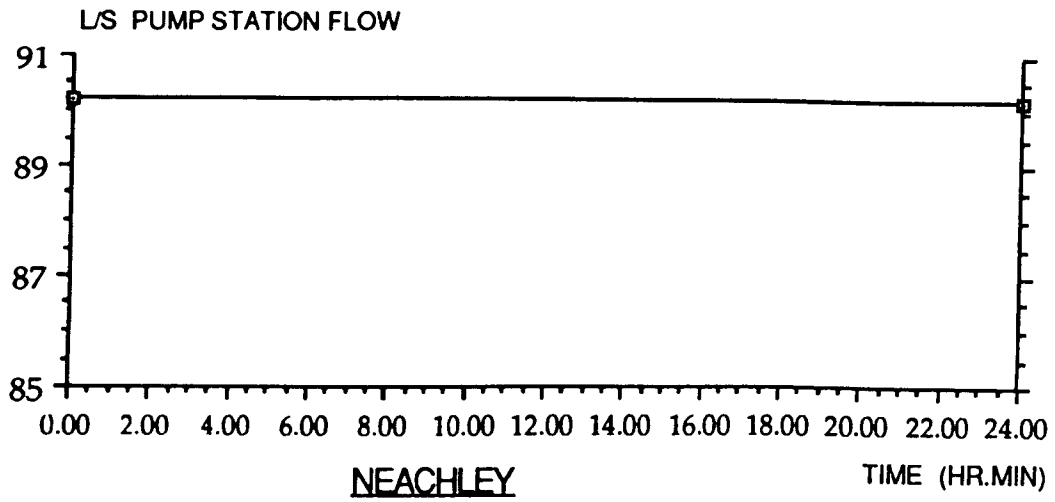


FIG 5.6 OPTIMIZED PUMP SCHEDULES

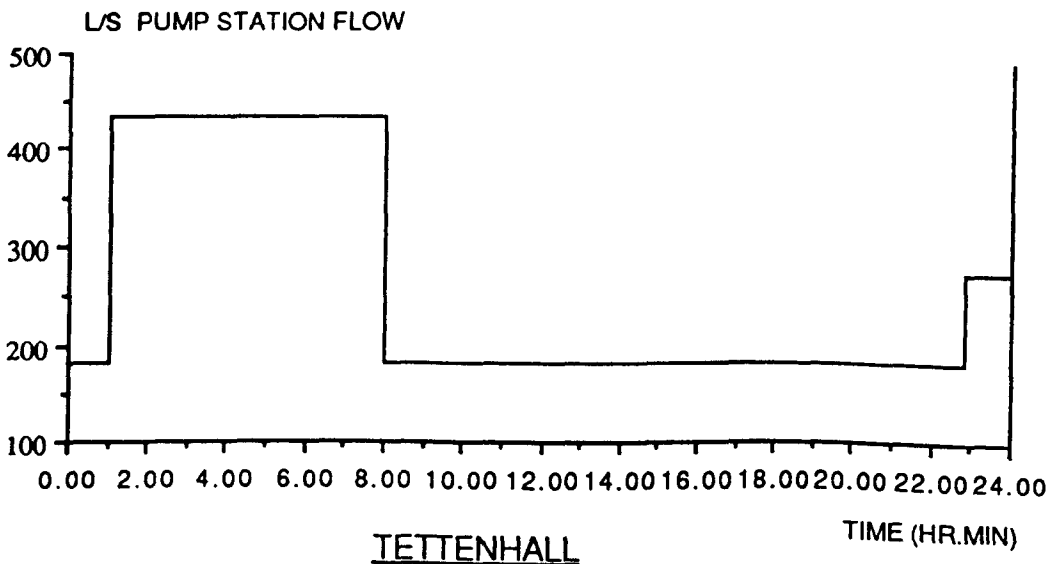
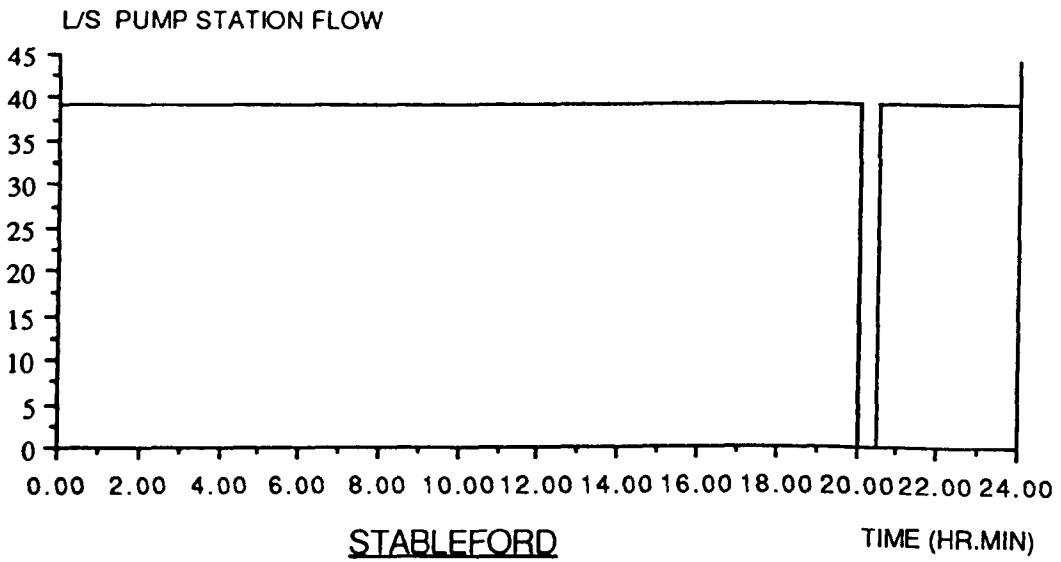
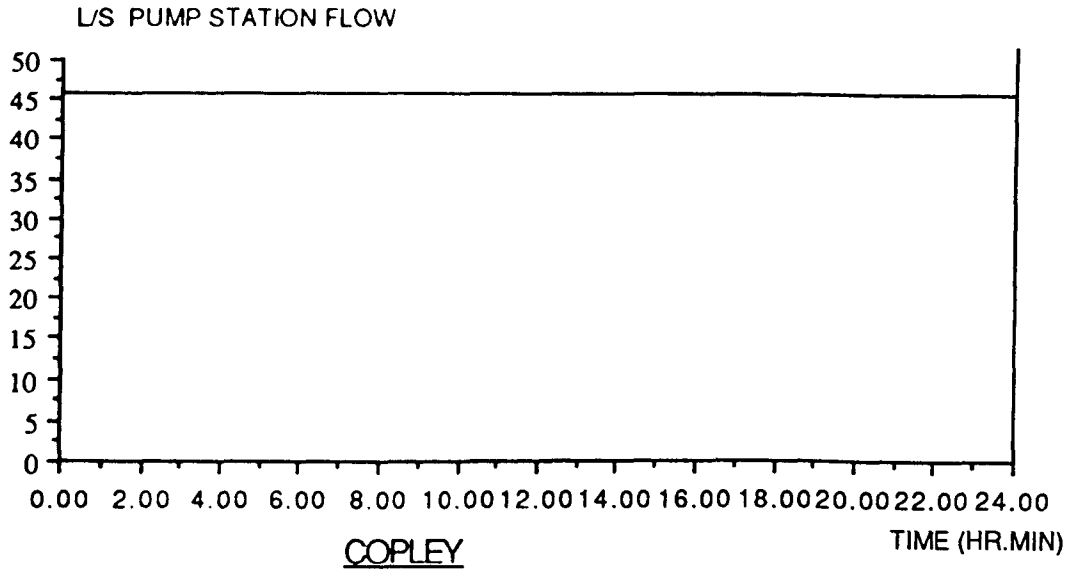


FIG 5.6 (continued)

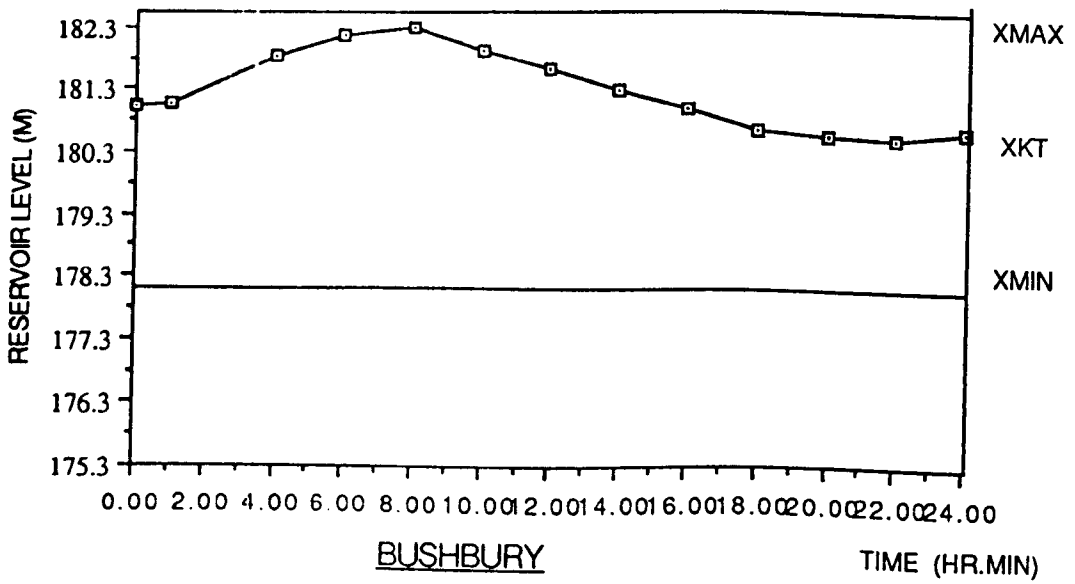
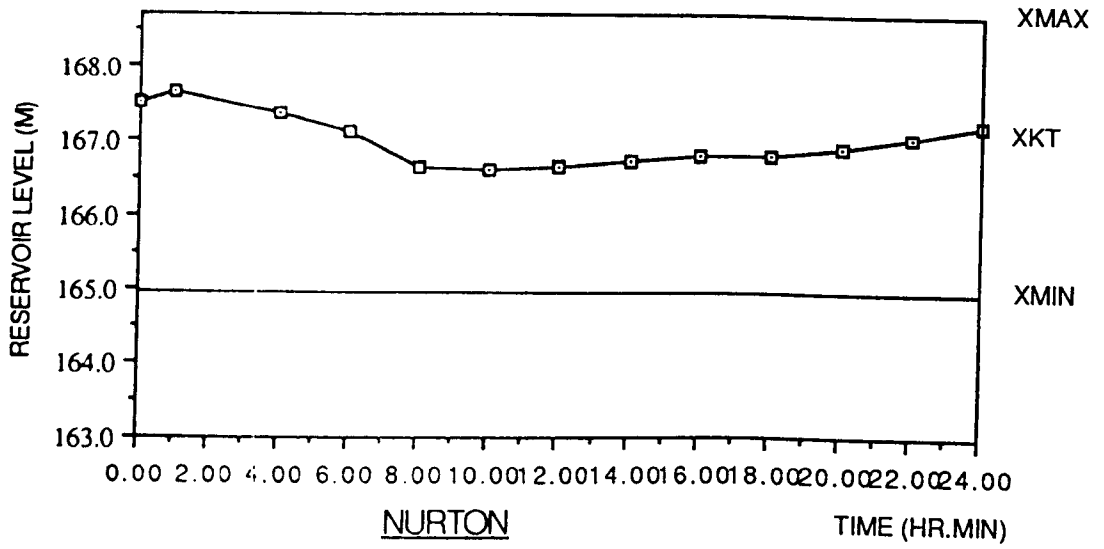


FIG 5.7 RESERVOIR TRAJECTORIES

Similarly, if a speed of 1525 rpm is introduced as an additional discretized point, which is only a 25 rpm decrement from the existing point of 1550 rpm, the corresponding additional feasible pump combinations are listed in Table 5.12. The sensitivity analysis results for these additional variable are given in Table 5.12a.

TABLE 5.12 AVERAGE VALUES OF FLOW AND POWER OF ADDITIONAL PUMP COMBINATIONS (ADDITIONAL SPEED STEP 1525 RPM)

ADDITIONAL COMBINATIONS	FLOW (l/s)	POWER (KW)
9) ONE VSP AT SPEED 1525 rpm, ANOTHER OFF	284.6	171.9
10) BOTH VSP's AT SPEED 1525 rpm,	354.4	253.4

The sensitivity analysis results shows that all the relevant additional variables remain non-basic, which means such a small step change from 1525 rpm to 1550 rpm produces no benefit. Also, this small step may be difficult to implement practically.

The discretization scheme of VSP's included in this algorithm has been illustrated through the above examples. In this way, we can first solve a relatively low dimensional problem with a few speed steps, then, through sensitivity analyses, we can incorporate any necessary additional speed steps. This approach is much more efficient than simply and arbitrarily discretizing the speed range of VSP's and solving every new problem 'from scratch'. It can be inferred that

this discretization scheme based on the sensitivity analysis will also be useful in studying the problems of pumping station expansions, e.g., choosing lower cost pumps, while satisfying service and other technical requirements. Sensitivity Analysis or Post-optimality Analysis (FLETCHER, 1986) is a unique advantage of linear programming based algorithms.

TABLE 5.12a SENSITIVITY ANALYSIS FOR ADDITIONAL SPEED
1525 RPM OF THE VSP'S

ADDITIONAL COMBINATION IP=9	C_{n+1}	m $-\sum_{i=1} a_{i,n+1} \pi_i$	CONCLUSION
Time Stage			
1	9.483	8.394	NON-BASIC
2	5.386	2.506	NON-BASIC
3	5.386	2.506	NON-BASIC
4	5.386	2.506	NON-BASIC
5	5.386	2.506	NON-BASIC
6	20.657	12.138	NON-BASIC
7	20.657	12.138	NON-BASIC
8	20.657	12.138	NON-BASIC
9	20.657	12.138	NON-BASIC
10	20.657	12.138	NON-BASIC
11	9.483	8.394	NON-BASIC
12	9.483	8.394	NON-BASIC

ADDITIONAL
COMBINATION IP=10

Time Stage			
1	13.979	11.766	NON-BASIC
2	7.940	5.878	NON-BASIC
3	7.940	5.870	NON-BASIC
4	7.940	5.870	NON-BASIC
5	30.450	15.510	NON-BASIC
6	30.450	15.510	NON-BASIC
7	30.450	15.510	NON-BASIC
8	30.450	15.510	NON-BASIC
9	30.450	15.510	NON-BASIC
10	30.450	15.510	NON-BASIC
11	13.979	11.766	NON-BASIC
12	13.979	11.766	NON-BASIC

5.3.5 Comparisons with the Modified GIPOS Application

The Nurton/Bushbury optimization problem has also been approached using a GIPOS based algorithm (PARKAR, 1989), which is introduced briefly as follows:

Under normal operational conditions, the global optimization problem is considered as two, weakly coupled, sub-system optimization problems. In the first instance, the Bushbury sub-system optimization problem includes Nurton reservoir as a source reservoir with the Tettenhall booster pumps providing the required delivery pressure to transfer water into Bushbury reservoir. In addition, the supply from Neachley, which is boosted through Tettenhall, is also considered as part of the Bushbury sub-system (Fig 5.8). The second sub-system comprises the sources at Cosford, Copley, Stableford and Hilton together with Nurton reservoir (Fig 5.9). The hydraulic coupling of the two sub-systems is achieved through the outflow from Nurton reservoir. This outflow is initially obtained by optimization of the Bushbury sub-system which consequently determines the required demand flow from Nurton reservoir.

The optimization of the Bushbury sub-system was approximated by the program GIPOS.

GIPOS (Graphical Interactive Pump Optimization and Scheduling) is a program to perform optimized scheduling for groups of parallel fixed speed and variable speed pumps. Both interactive and graphical display features are incorporated within the program to provide users with simple operations and easy interpretation of results.

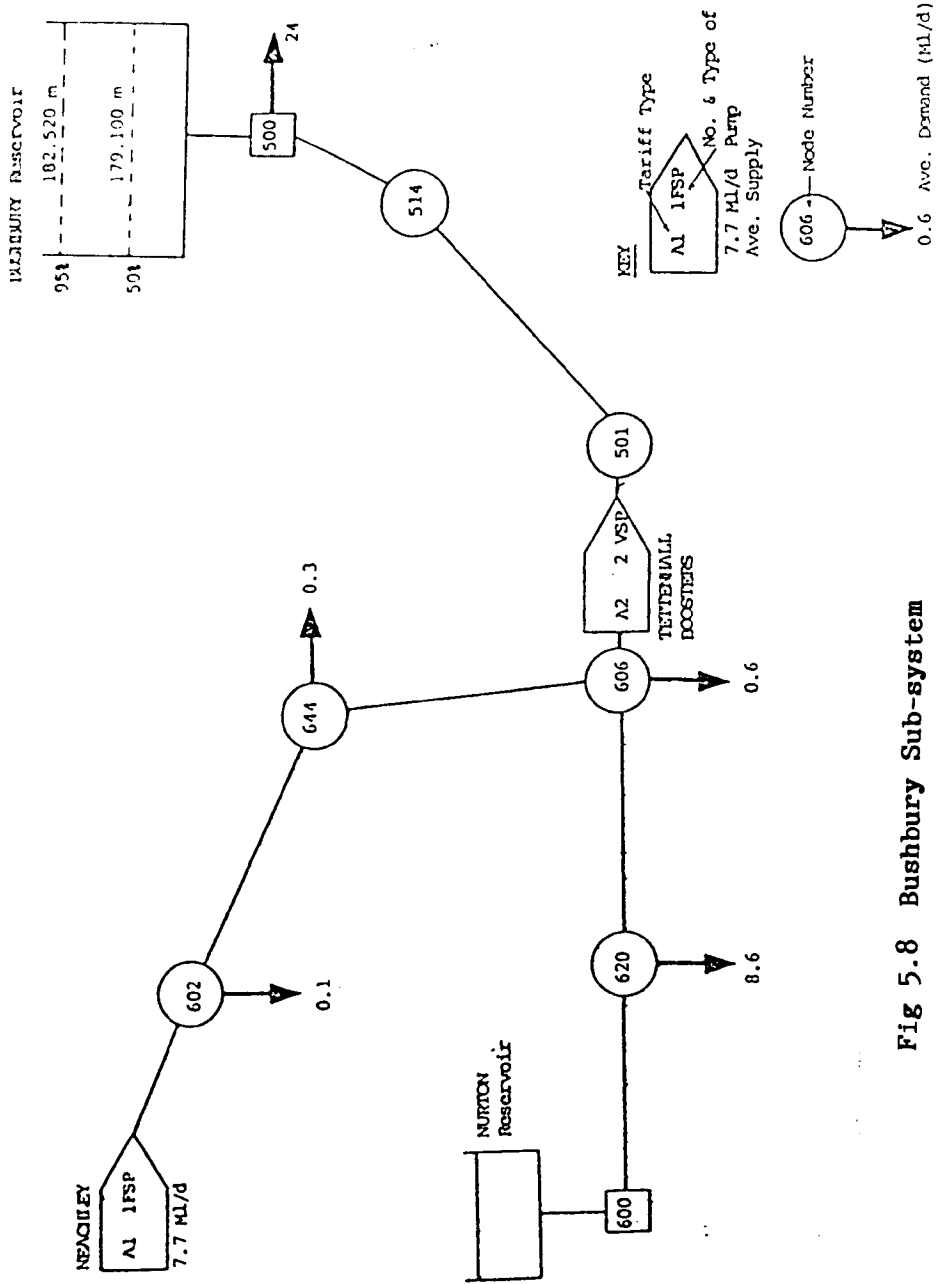


Fig 5.8 Bushbury Sub-system

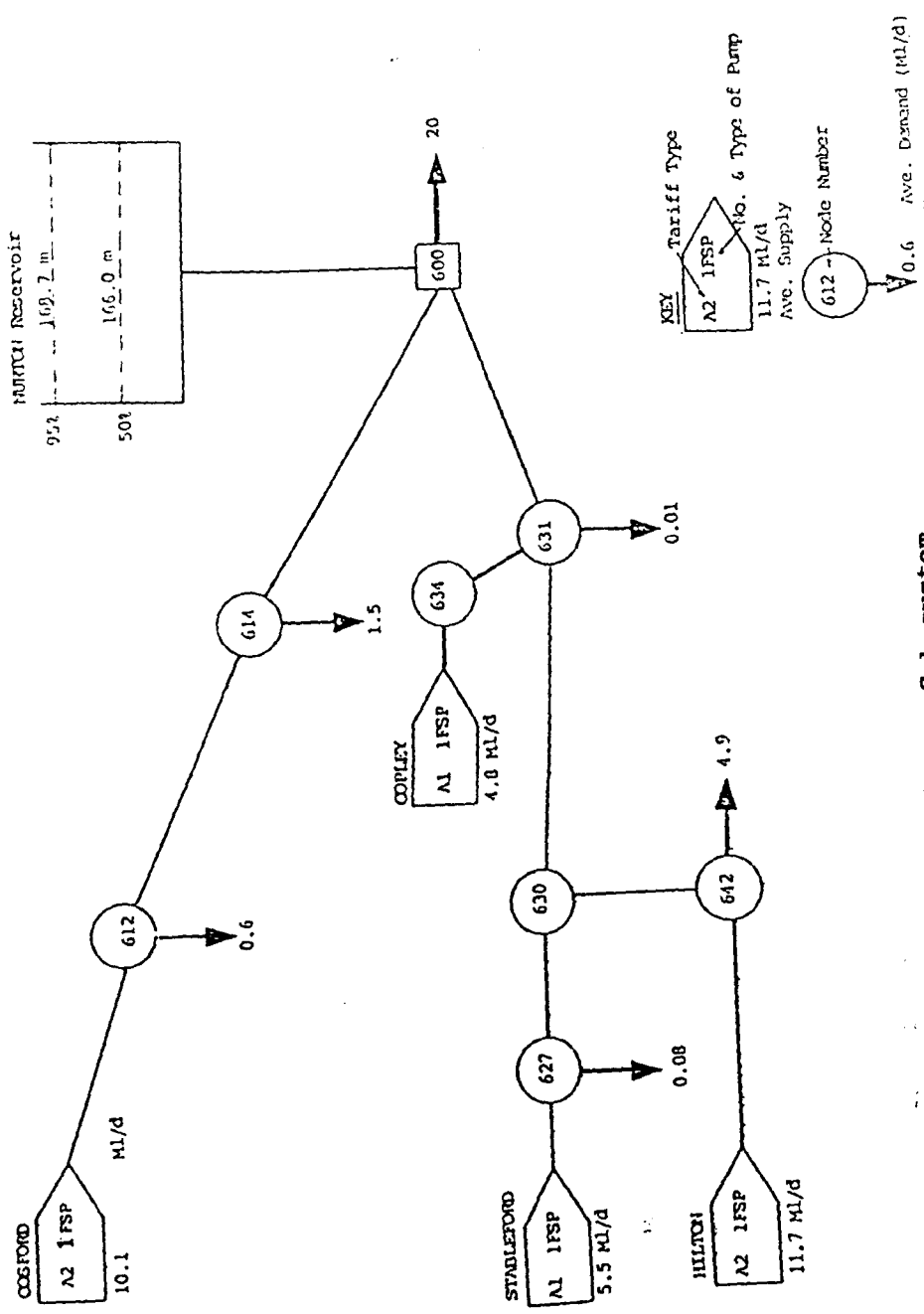


Fig 5.9 Nurton Sub-system

The algorithm employed within the program is based on the Forward Dynamic Programming Technique. The program takes into account all pump combinations together with pump speed for all variable speed pumps. Calculations are performed, under a set of prescribed operational conditions, to obtain schedules of the number of parallel pumps in use and their speeds. These schedules give the cheapest operational cost over the time duration studied. However, this program is only applicable to sub-systems with a single source reservoir supplying water to a single controlled reservoir via one pump station and an equivalent direct pipe-line, with intermediate and final take-off demands.

When GIPOS is applied to the Bushbury sub-system, Bushbury reservoir is taken as the controlled reservoir and Nurton reservoir as the source reservoir. The optimized schedules are derived for the Two VSP's at Tettenhall. In order to take into account the supply from Neachley and the demand requirement between the source reservoir (Nurton) and pump station (Tettenhall), some modifications have been made to GIPOS. Upon the completion of the GIPOS computations, the optimized pump schedules at Tettenhall and the required demand flow from Nurton reservoir are determined.

For Nurton sub-system, the objective of optimization is to minimize the total pumping costs from the four fixed speed pump stations at Cosford, Copley, Stableford and Hilton (practically, Cosford and Hilton pumps must be on all the day). In doing so, the operational constraints regarding reservoir levels at Nurton reservoir, demand flows from Bushbury sub-system optimization and source flows, are all taken into consideration. The optimized pump schedules of the fixed

speed pumps are obtained by using a direct search method to calculate the costs of all possible fixed speed pump combinations, within the operational constraints for each pump station, and then choosing the least cost combination.

From the above statement, it is worth noting several points: Firstly, the Dynamic Programming based algorithm is very effective and accurate for a single controlled reservoir system. For a system containing more than one controlled reservoir, the computation time, and for some computers, the storage requirement will be critical, especially for on-line control purposes. In this approach, Bushbury sub-system is initially optimized regardless of Nurton sub-system, Nurton sub-system is subsequently optimized bound by the results from Bushbury sub-system optimization. Clearly, in doing so, the optimality of the global problem might be lost to some degree. In other words, only sub-optimal results for the whole system might be obtained without including optimal coordination of the two independent results. Whereas with the Linear Programming based algorithm, developed by taking the system as a whole, global optimal results can always be found. Secondly, the direct search method applied to the Nurton sub-system is computationally time-consuming and inefficient. Consequently, this approach may be impracticable when more pumps, particularly variable speed pumps, are introduced in the Nurton sub-system. Such pumps are currently being commissioned (PARKAR et al, 1989). Thirdly, this approach utilizes constant pumping times which are equal to the length of prescribed intervals. Currently the 24 hour control period is divided into 4 intervals, namely, 0.00 hr to 1.00 hr, 1.00 hr to 8.00 hrs, 8.00 hrs to 20.00 hrs and 20.00 hrs to

24.00 hrs (in the author' approach, the time intervals are 0.00 hr to 1.00 hr, 1.00 hr to 4.00hrs, then every two hours). Consequently, the prescribed final reservoir level of Nurton, which is important for long-term optimal operation purposes, is not always reachable using the limited fixed speed pump combinations. In order to achieve the prescribed final reservoir level using the limited fixed speed pump combinations, it is necessary to divide the whole control period into many smaller intervals. This, however, will dramatically increase the direct search and dynamic programming computing times.

Numerical comparisons between the Linear Programming based program OPPUS and the Dynamic Programming based program GIPOS (modified, including the direct search method) have been made. In order to make the results from both programs comparable, the average values of pump flows and powers used by OPPUS have been adjusted to the same values used by the modified GIPOS, which are summarized in Table 5.13. Other operating conditions are also kept identical in deriving the results from the both algorithms. Firstly, comparisons are made in Tables 5.14 and 5.15 and Figures 5.10 and 5.11 with the same final reservoir level requirement. From Fig 5.11, it can be seen that the final level of Bushbury reservoir (prescribed as the same as the initial reservoir level of 179.9m) was attained by GIPOS (the deviation is about 0.4%). The final reservoir level derived from OPPUS is also very close to the prescribed level. However, the final level of Nurton reservoir derived from the direct search method is much lower than the prescribed level (the deviation is about 11.1%). While the final level derived from OPPUS is the same as the initial level.

Secondly, by adjusting the final reservoir levels of Nurton and Bushbury for OPPUS to be the same as those derived from the modified GIPOS, gives the comparison results in Table 5.16, Figures 5.12 and 5.13. From these numerical comparisons, it can be seen that the optimization results derived from both OPPUS and the modified GIPOS are very close to each other. The computation time required to solve this problem by OPPUS is about 1 minute. In contrast, the computation time required to solve the same problem by the modified GIPOS is about 2.5 minutes (PARKAR, 1989).

TABLE 5.13 AVERAGE VALUES OF PUMP FLOW AND POWER
USED BY BOTH OF THE APPROACHES

	FLOW (1/s)	POWER (kW)
PUMPING STATION 1		
COMBINATIONS:		
1) PUMP AT STABLEFORD	72.4	63.9
2) PUMP AT COPLEY	63.9	150.6
3) PUMPS AT STABLEFORD AND AT COPLEY	136.3	214.5
CONSTANT FLOWS		
1) PUMP AT NEACHLEY	93.1	197.3
2) PUMP AT COSFORD	112.8	95.8
3) PUMP AT HILTON	128.8	361.3

TABLE 5.14 RESULTS FROM OPPUS FOR THE OPTIMAL

OPERATION OF NURTON/BUSHBURY ZONE
OPERATIONS OF PUMPING STATION 1 (Stableford + Copley)

TIME (hr.min)	0.00	20.00	20.00	21.05	21.05	24.00
PUMP COMBINATION IN OPERATION (NO.)	3	3	1	1	3	3
FLOW (l/s)	136.3	136.3	72.4	72.4	136.3	136.3

OPERATIONS OF PUMPING STATION 2 (Tettenhall)

TIME (hr.min)	0.00	1.00	1.00	8.00	8.00	20.00	20.00	21.50	21.50	24.00
PUMP COMBIN- ATION IN OPERA- TION (NO.)	4	4	6	6	1	1	6	6	4	4
FLOW (l/s)	216.8	216.8	434.2	434.2	181.2	181.2	434.2	434.2	216.8	216.8
VSP 1 SPEED (rpm)	1100	1100	1600	1600	1100	1100	1600	1600	1100	1100
VSP 2 SPEED (rpm)	1100	1100	1600	1600	0	0	1600	1600	1100	1100

WATER PUMPED, ENERGY CONSUMED AND COSTS

PUMP NAME	WATER PUMPED (M ³ /D)	ENERGY CONSUMED (KWH)	COST (£)
STABLEFORD	6255.4	1533.6	52.62
COPLEY	5271.8	3451.3	117.38
TETTENHALL	24106.7	3747.6	119.58
NEACHLEY	8043.8	4735.2	162.48
COSFORD	9745.9	2299.2	111.35
HILTON	11128.3	8671.2	419.94
TOTAL	64551.9	24438.1	983.35

TABLE 5.15 RESULTS FROM MODIFIED GIPOS FOR THE OPTIMAL

OPERATION OF NURTON/BUSHBURY ZONE

NURTON SUB-SYSTEM

TIME (hr.min)	0.00	1.00	1.00	8.00	8.00	20.00	20.00	24.00
COSFORD								
PUMP FLOW (l/s)	112.8	112.8	112.8	112.8	112.8	112.8	112.8	112.8
COPLEY								
PUMP FLOW (l/s)	0.0	0.0	63.9	63.9	63.9	63.9	0.0	0.0
STABLEFORD								
PUMP FLOW (l/s)	0.0	0.0	72.4	72.4	72.4	72.4	72.4	72.4
HILTON								
PUMP FLOW (l/s)	128.8	128.8	128.8	128.8	128.8	128.8	128.8	128.8

BUSHBURY SUB-SYSTEM

TIME (hr.min)	0.0	1.00	1.00	8.00	8.00	20.00	20.00	24.00
NEACHLEY								
PUMP FLOW (l/s)	98.9	98.9	107	107	82.5	82.5	99.3	99.3
TETTENHALL								
PUMP FLOW	304	304	399	399	189	189	339	339
VSP1 SPEED (rpm)	1335	1335	1583	1583	1051	1051	1417	1417
VSP2 SPEED (rpm)	1335	1335	1583	1583	1051	1051	1417	1417

WATER PUMPED, ENERGY CONSUMED AND COSTS

PUMP NAME	WATER PUMPED (M ³ /D)	ENERGY CONSUMED (KWH)	COST (£)
STABLEFORD	5994.7	1469.7	50.02
COPLEY	4370.8	2861.4	93.37
TETTENHALL	24195.6	3955.6	137.10
NEACHLEY	8046.4	4736.2	161.86
COSFORD	9745.9	2299.2	111.35
HILTON	11128.3	8671.2	419.94
TOTAL	63481.7	23993.3	973.64

— results from GIPOS
 -□- results from OPPUS

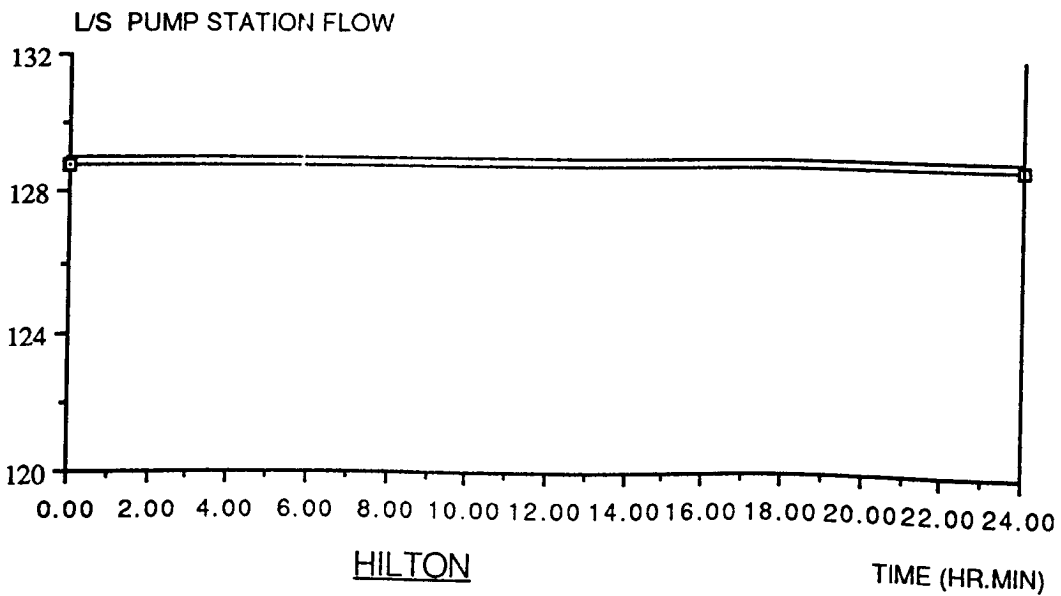
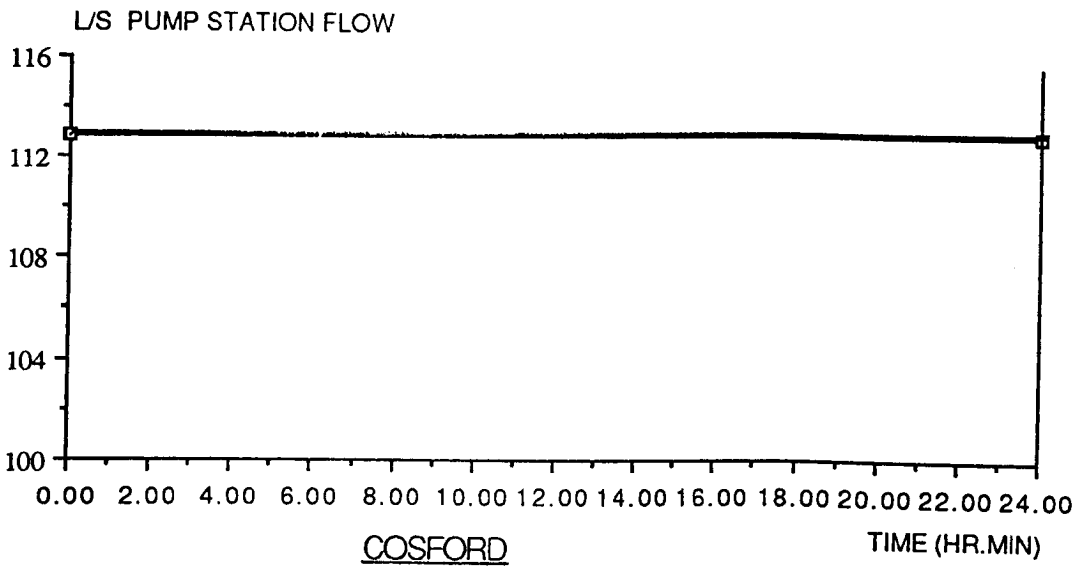
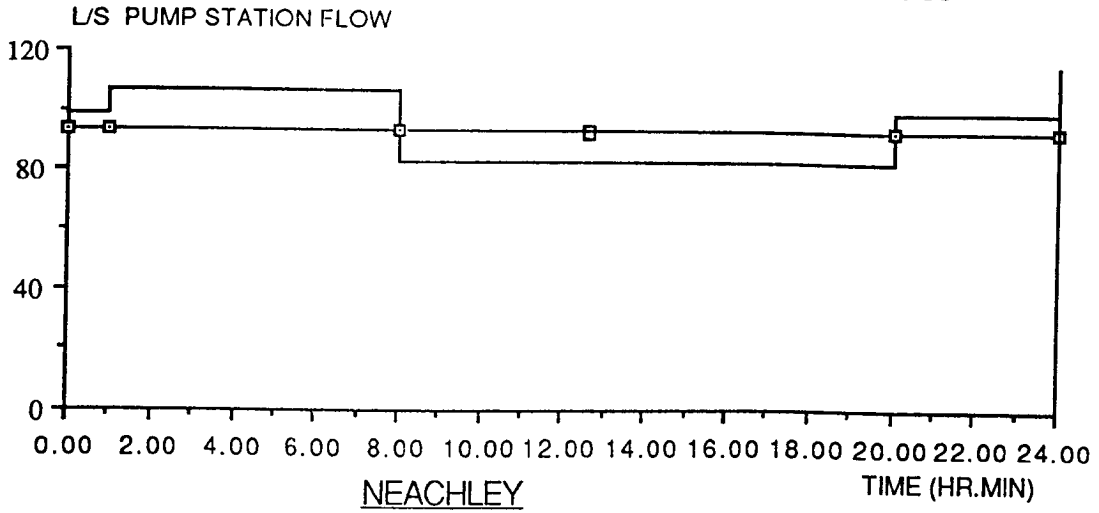


FIG 5.10 COMPARISON OF PUMP SCHEDULES

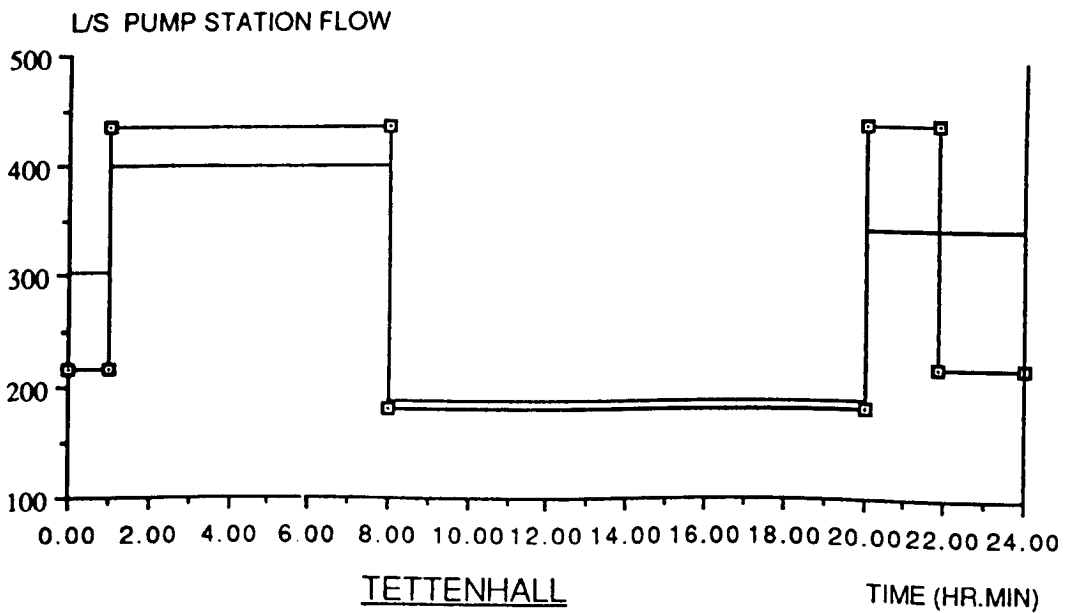
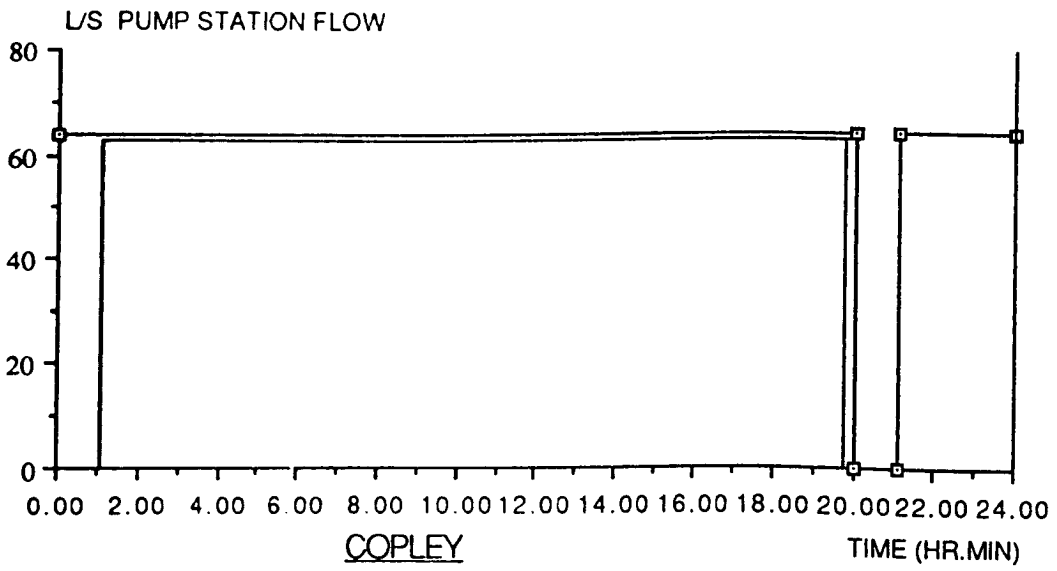
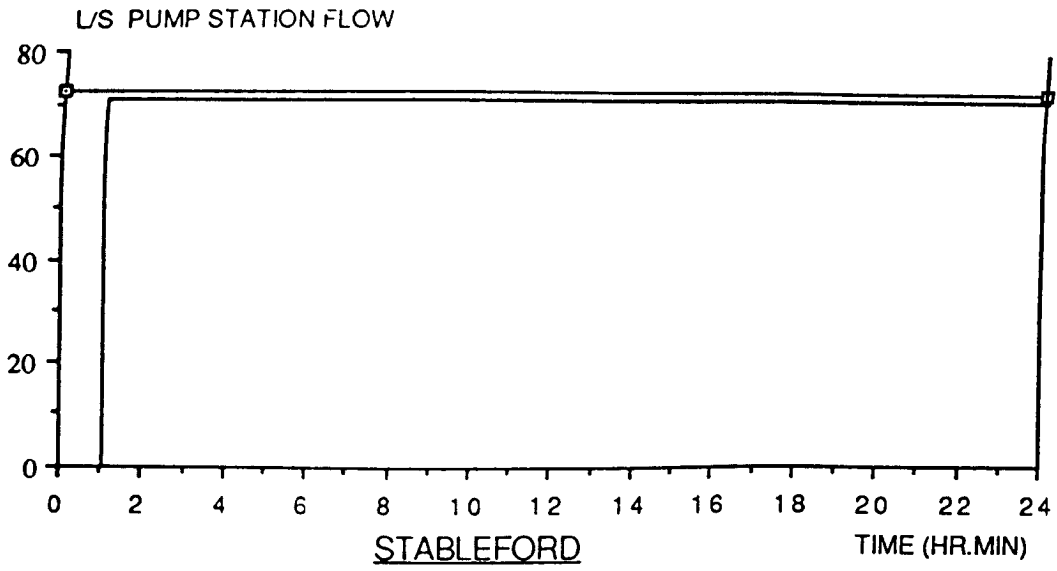


FIG 5.10 (continued)

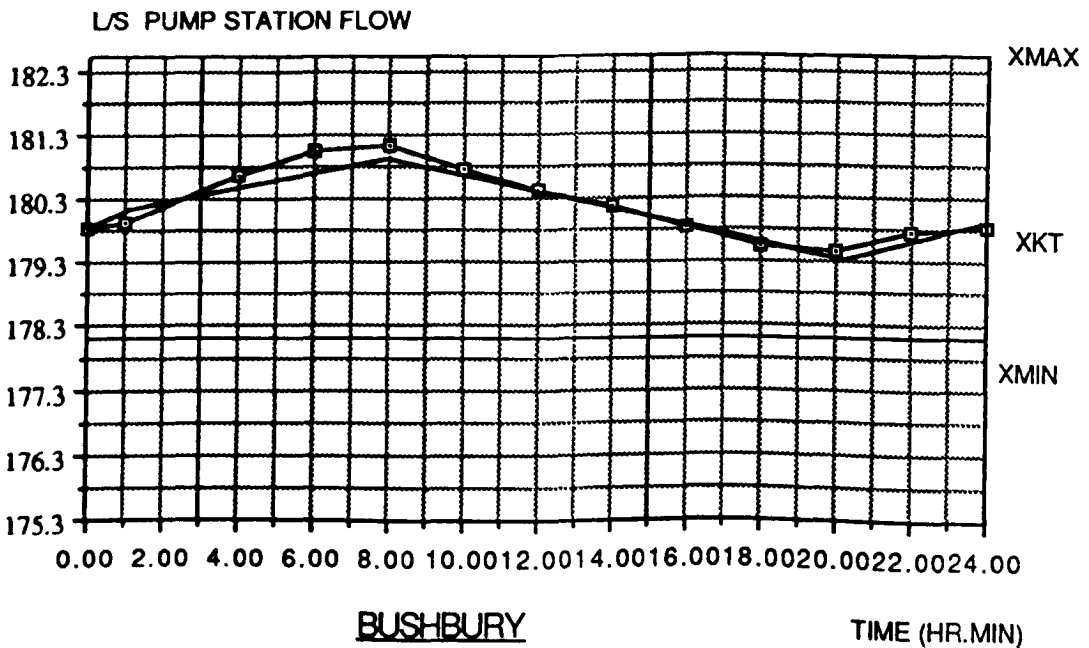
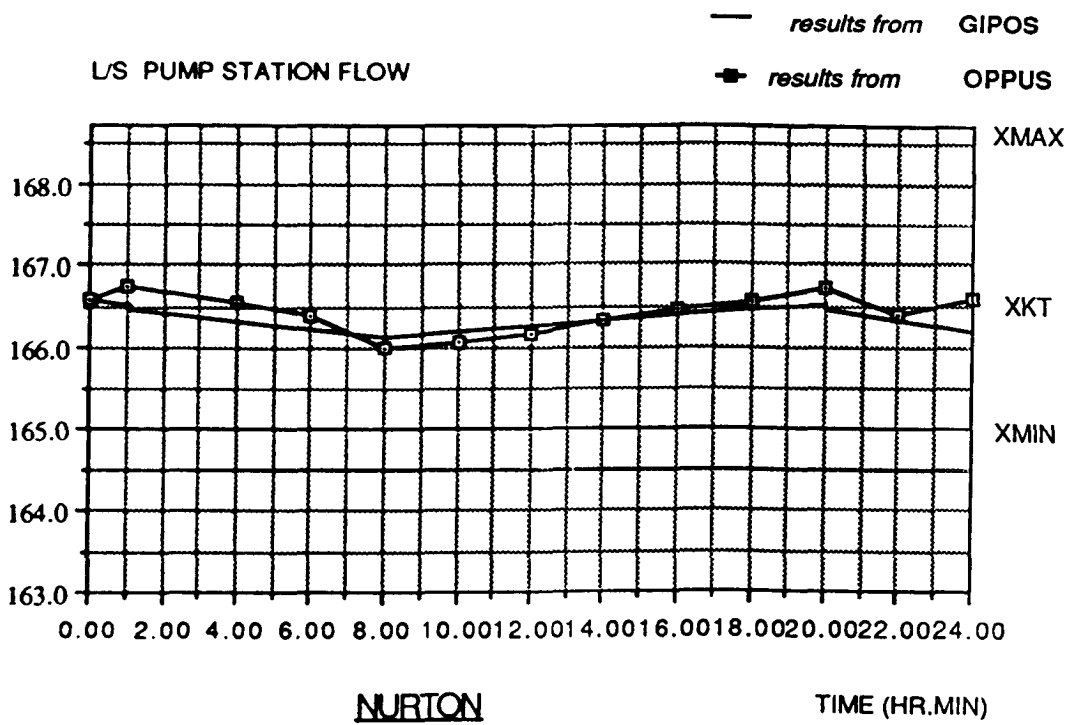


FIG 5.11 COMPARISON OF RESERVOIR TRAJECTORIES

TABLE 5.16 RESULTS FROM OPUS FOR THE OPTIMAL

OPERATION OF NURTON/BUSHBURY ZONE

OPERATIONS OF PUMPING STATION 1 (Stableford + Copley)

TIME (hr.min)	0.00	10.00	10.00	16.10	16.10	24.00
PUMP COMBINATION IN OPERATION (NO.)	3	3	1	1	3	3
FLOW (l/s)	136.3	136.3	72.4	72.4	136.3	136.3

OPERATIONS OF PUMPING STATION 2 (Tettenhall)

TIME (hr.min)	0.00	1.00	1.00	8.00	8.00	20.00	20.00	21.57	21.57	24.00
PUMP COMBIN- ATION IN OPERA- TION (NO.)	4	4	6	6	1	1	6	6	4	4
FLOW (l/s)	216.8	216.8	434.2	434.2	181.2	181.2	434.2	434.2	216.8	216.8
VSP 1 SPEED (rpm)	1100	1100	1600	1600	1100	1100	1600	1600	1100	1100
VSP 2 SPEED (rpm)	1100	1100	1600	1600	0	0	1600	1600	1100	1100

WATER PUMPED, ENERGY CONSUMED AND COSTS

PUMP NAME	WATER PUMPED (M ³ /D)	ENERGY CONSUMED (KWH)	COST (£)
STABLEFORD	6255.4	1533.6	52.62
COPLEY	4102.4	2685.7	86.22
TETTENHALL	24198.2	3773.9	120.45
NEACHLEY	8043.8	4735.2	162.48
COSFORD	9745.9	2299.2	111.35
HILTON	11128.3	8671.2	419.94
TOTAL	63474.0	23698.8	953.06

— results from GIPOS
 -□- results from OPPUS

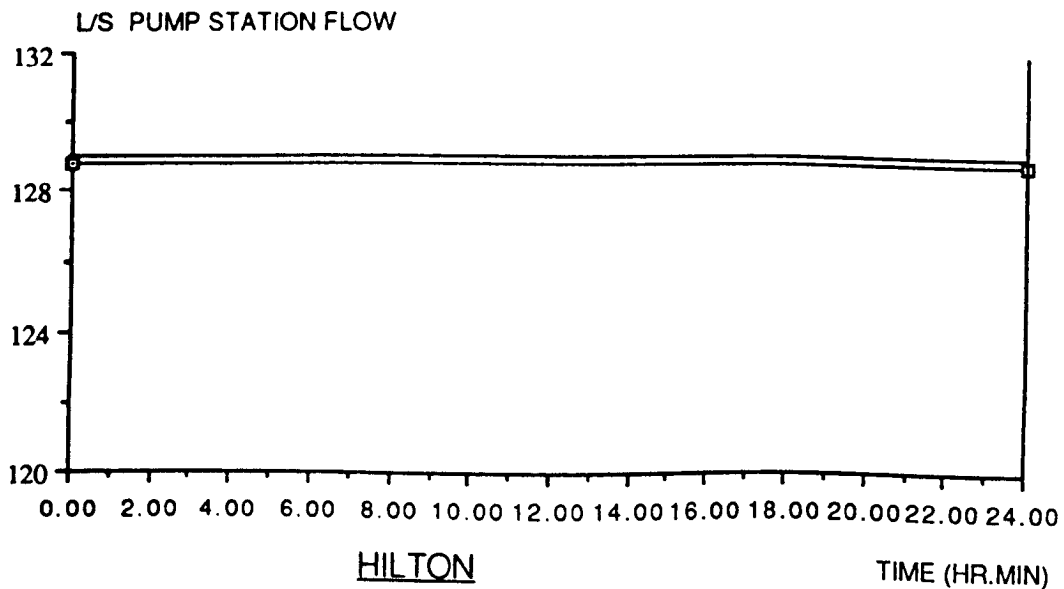
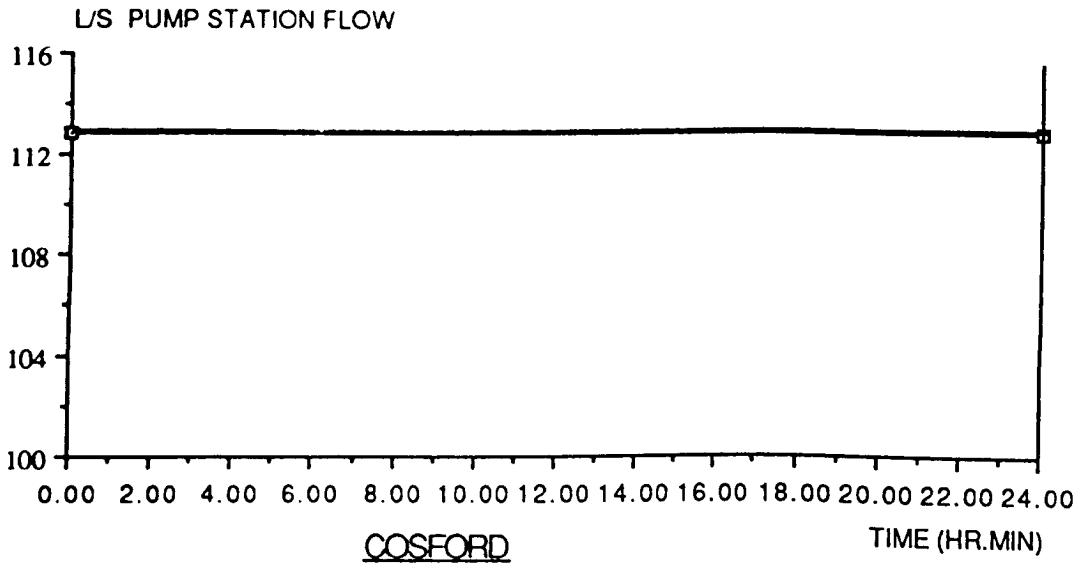
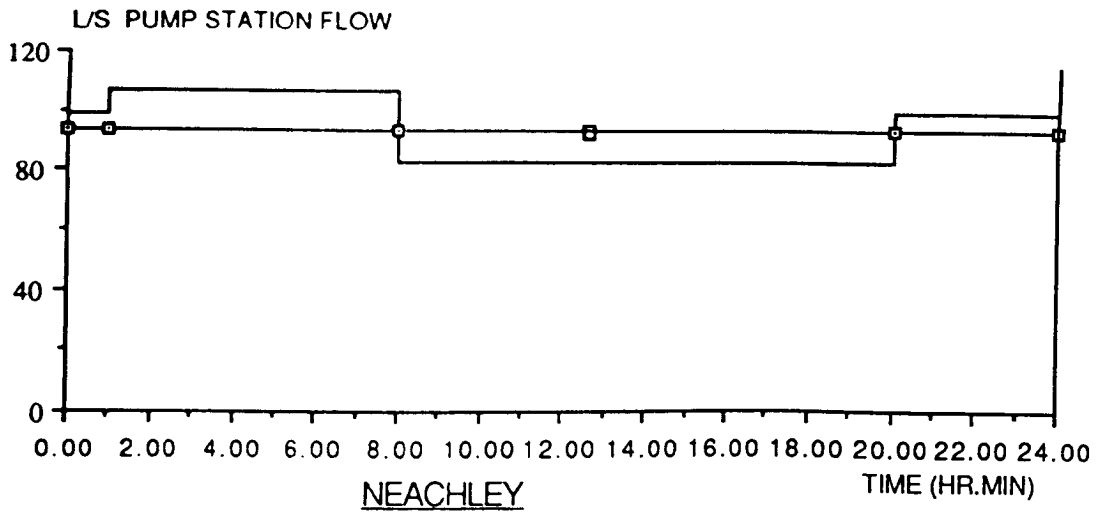


FIG 5.12 COMPARISON OF PUMP SCHEDULES

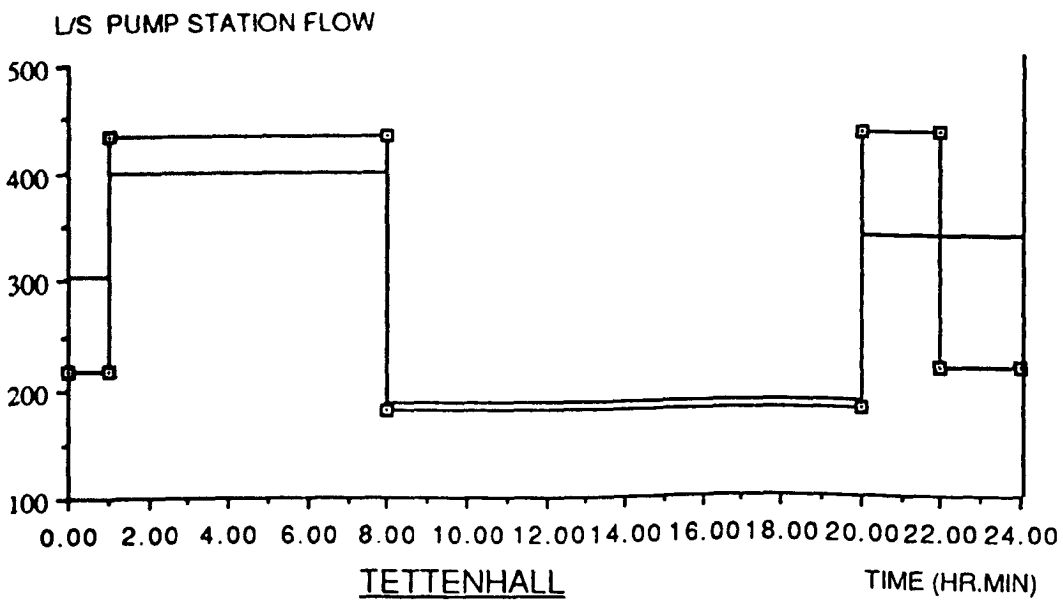
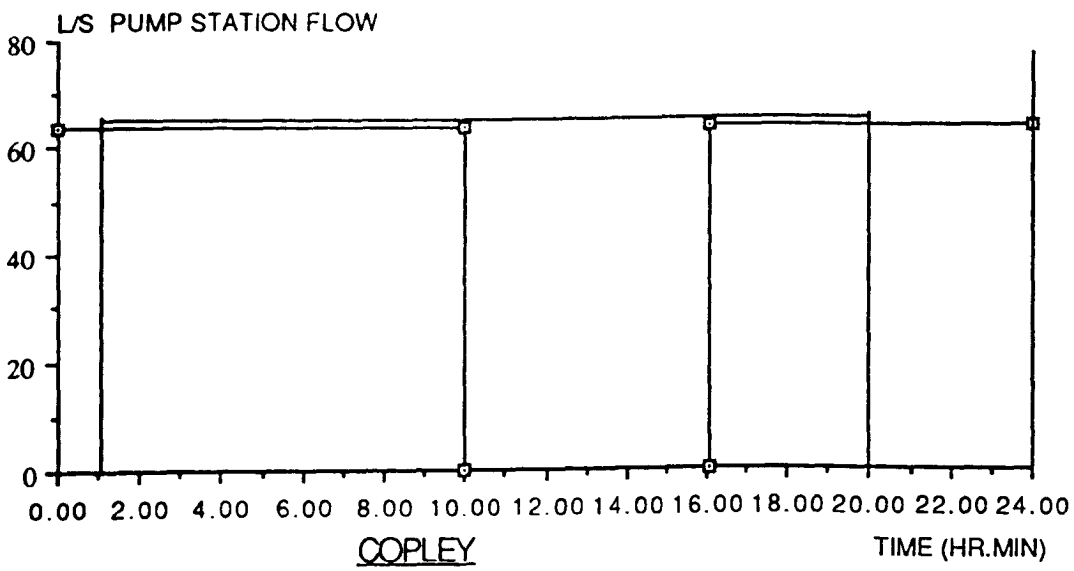
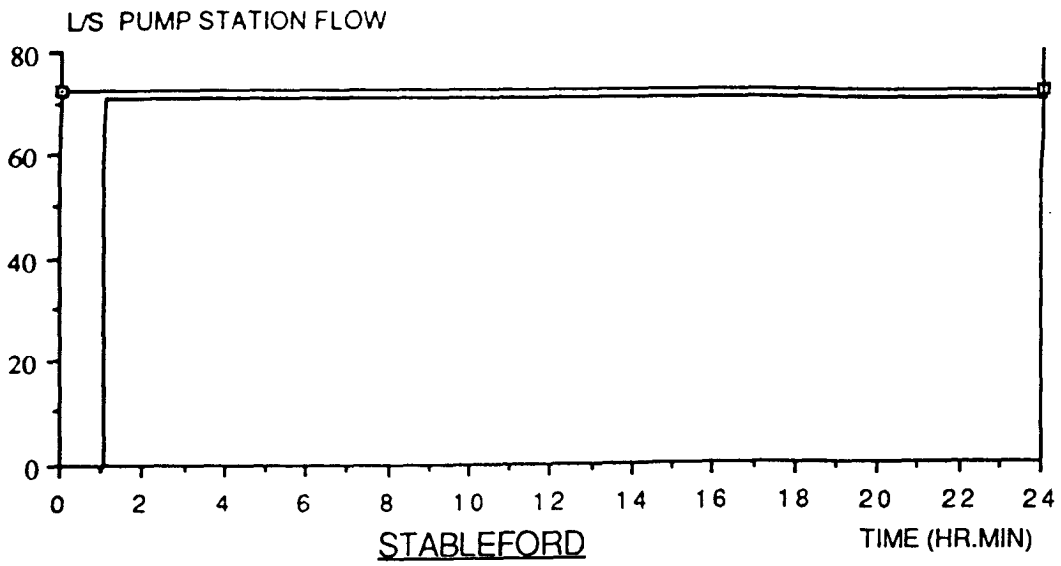


FIG5.12 (continued)

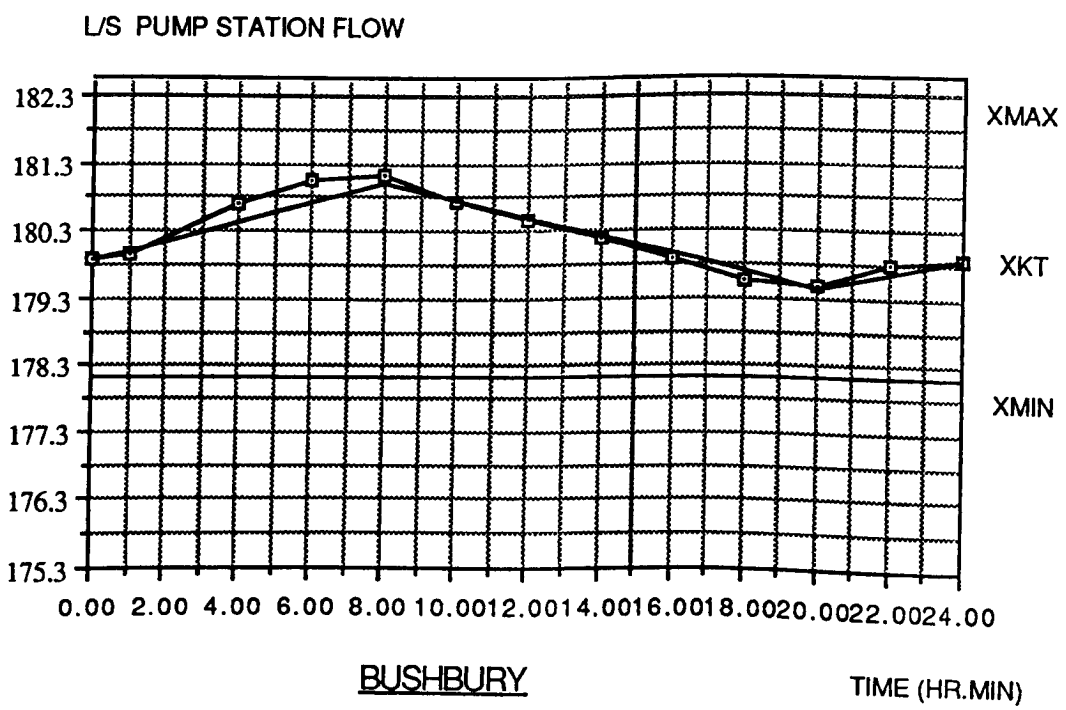
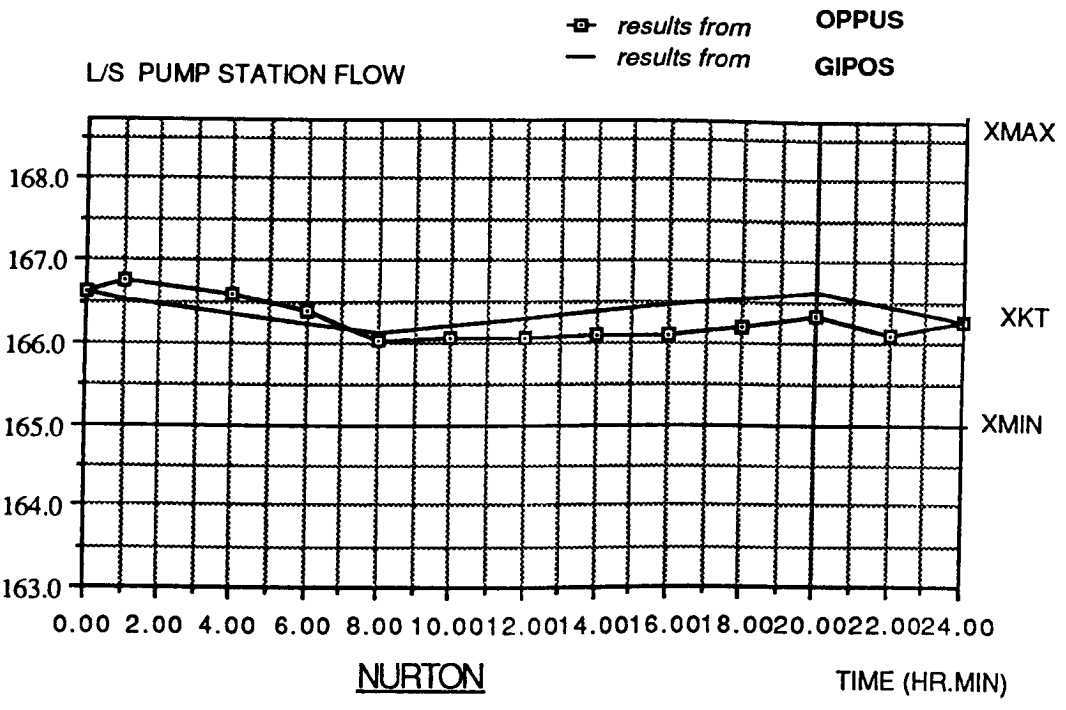


FIG 5.13 COMPARISON OF RESERVOIR TRAJECTORIES

5.4 CONCLUSIONS

In this chapter, two different optimization algorithms catering for two different classes of water supply and distribution systems were presented. The two classes considered included systems with and without significant water storage.

The first algorithm is directly applicable to multi-source systems without significant water storage in the distribution part of the system. This class represents many small water supply and distribution systems or sub-systems in reality. In this case, there are no dynamics in the system. Therefore, the optimal operation problem for each time interval of the complete control period can be solved separately, and thus the number of decision variables is small. In this algorithm, only the pumping costs and treatment costs are used as the objective function. By adopting the piecewise macroscopic model developed in Chapter 4, instead of solving simultaneous nonlinear network equations iteratively, the objective function becomes a function only about station flows. This model also incorporates the pressure requirement of selected nodes in the system directly. The adoption of the piecewise macroscopic model in this algorithm has speeded up the solution time significantly, in order to make this algorithm feasible for on-line control purposes. The operational constraints on pumping capacities and other relevant system variables are also incorporated into the constraint set. The formulation of the optimal operations by this algorithm results in a constrained nonlinear programming problem. The chosen solution method, SUMT (Sequential Unconstrained Minimization Technique), has been discussed and summarized. The application results show that

approximately 11% of cost savings are achievable when applying this algorithm to a real system.

This algorithm can easily be extended to deal with systems having significant storages, by incorporating macroscopic relationships for reservoirs (DEMOYER and HORWITZ, 1975a, 1975b) into the constraint set. However, this will increase the dimension of the problem, and much greater difficulties will be encountered in the computation of the nonlinear programming algorithm. The main disadvantage of this algorithm is that a global optimum cannot be guaranteed. In practice, we can start the computation from different initial points and compare the solutions to see whether the solution is unique otherwise we choose the least cost solution as the final one.

The second algorithm can be applied to multi-source, multi-reservoir systems, in which pump flows and powers are more or less constant during the whole control period. This condition can be approximately satisfied in many systems where significant storage somewhat decouples direct pump flows from major consumer demands. Based on the equivalent network model developed in Chapter 4, and by using time of pumping instead of pump flow as a decision variable, the formulation of the optimal operation problem for a complete control period results in a large-scale dynamic linear programming problem for which a global optimum can always be guaranteed. The chosen solution method, the Revised Simplex Method, has been introduced in principle. The incorporated discretization scheme can overcome the difficulties of the mixture of fixed speed pumps (discrete variables) with variable speed and/or throttle pumps (continuous variables). A post-optimality

analysis method plays an important role in the discretization scheme.

This algorithm has been fully tested through the application to the Nurton/Bushbury zone in the city of Wolverhampton system. The application results confirm that least cost pump schedules and reservoir trajectories are obtainable from the algorithm for various operational conditions. The solution is usually available within 3 minutes, which indicates that this algorithm is practical for on-line control purposes. Note that computational solving times were derived by using a PRIME 2550 multi-user computer. Faster solving times could be achieved by using other readily available computing devices such as dedicated workstations or PC's (e.g., Sun 3/60 which is at least twice as fast as the PRIME computer or the Sparc workstation which is at least 10 times faster).

The application of the modified GIPOS (dynamic programming based plus direct search method) to the same system has been introduced in brief and compared with this algorithm. The disadvantages in comparison with the author's algorithm are that, firstly, only sub-optimal results for the whole system are obtainable; secondly, the direct search method is computationally time-consuming and inefficient and even inapplicable for future operations of the system; thirdly, the prescribed final level of Nurton reservoir is not always reachable. The numerical comparisons show that the optimization results between the modified GIPOS and the author's algorithm(OPPUS) are not very different from each other.

This algorithm can be extended to deal with those multi-source, multi-reservoir systems in which the pump flows are varying significantly.

The can be done by taking average values of the flows and powers. The derived solution from this algorithm can provide a good reference for their optimal operations.

The major shortcoming of this algorithm is that for some systems, the linear programming problem would be so large that the computation time required by the Revised Simplex method would be critical. If this happens, some decomposition method for the linear programming problem, such as prime-dual decomposition and/or compact storage technique, such as LU Factorization, could be employed to modify this algorithm.

Since the first algorithm is a non-linear programming problem, a global optimum cannot be guaranteed theoretically. In the author's experience, by changing the initial guess to different values, the derived solutions, in most cases, are very close to each other and are satisfactory. However, one or two solutions are not very satisfactory. Therefore, special care is needed to choose the best available solution for on-line implementation of this algorithm.

For the second algorithm, which is based on a linear programming method, a global optimum can be guaranteed. The computer program has the options of either calling the Revised Simplex subroutine in (BUNDAY, 1984) or the subroutine EO4MBF in (NAG, 1983), and the solutions from these two subroutines agree with each other accurately for all applications.

6.1 OVERALL COMMENTS

In this thesis, two methods, which cover the various aspects of the research topic from water demand forecasting, network modelling and simplification, to optimization of system operation, have been systematically developed in parallel. These two methods, catering for two different classes of water supply and distribution systems, can, together, solve the optimization problems of fairly general systems.

It has been shown that the overall control of water supply and distribution systems is very complex. Consequently, most research and development have concentrated on a restricted class of systems, due mainly to the limitations of computing power and available mathematical techniques. On the other hand, it has been shown that successful implementations of suitable models are possible, providing that certain properties of a system are exploited correctly. This fact provided the motivation for the current research project.

This thesis represents a significant contribution to the above research field. It presents the development of new, practical methods for modelling and optimization, with extended solution capabilities. The refinement of existing techniques is also described. The integration of these techniques has led to the development of two

algorithms which cater for the overall optimal control of a wider range of water supply and distribution system than was previously possible.

The theoretical aspects of optimal system operation have even thoroughly investigated. Theoretical results have indicated the benefits of applying the theory in practice. It remains to actually implement these methods in a real system. Fig 6.1 is an illustration of an integrated scheme for the dynamic optimal control of a water system. The scheme uses an on-line computer employing the mathematical models and software algorithms developed in this thesis. The scheme would realize the control strategy envisaged in Fig 2.5 (Chapter 2).

When utilizing this scheme for a particular system, one should examine the characteristics of the system and adopt the most appropriate of the two methods described here. For a large scale system, however, both methods could be applied to different sub-systems under the control of an upper level coordinator. The latter would coordinate the local optimal results in order to achieve the global optimum.

An optimal control strategy can be thus derived which will facilitate control actions over the control period (in terms of pump scheduling and so on). The control strategy would be based upon predicted demands and previously derived network models.

The on-line monitoring of significant parameters such as demand and reservoir level enable discrepancies between actual and predicted

values to be evaluated. Significant discrepancies would necessitate the evaluation of a revised optimal control schedule, thus ensuring optimality over the control period.

The individual components of the scheme were developed as stand-alone units, able to be applied independently to a variety of analyses. The techniques of demand forecasting, piecewise macroscopic modelling, equivalent network modelling, linear programming optimization, and nonlinear programming optimization, are considered to represent significant developments in their own right. The programs presented in this thesis can be applied independently or in conjunction with other existing methods to provide a comprehensive analysis and control tool.

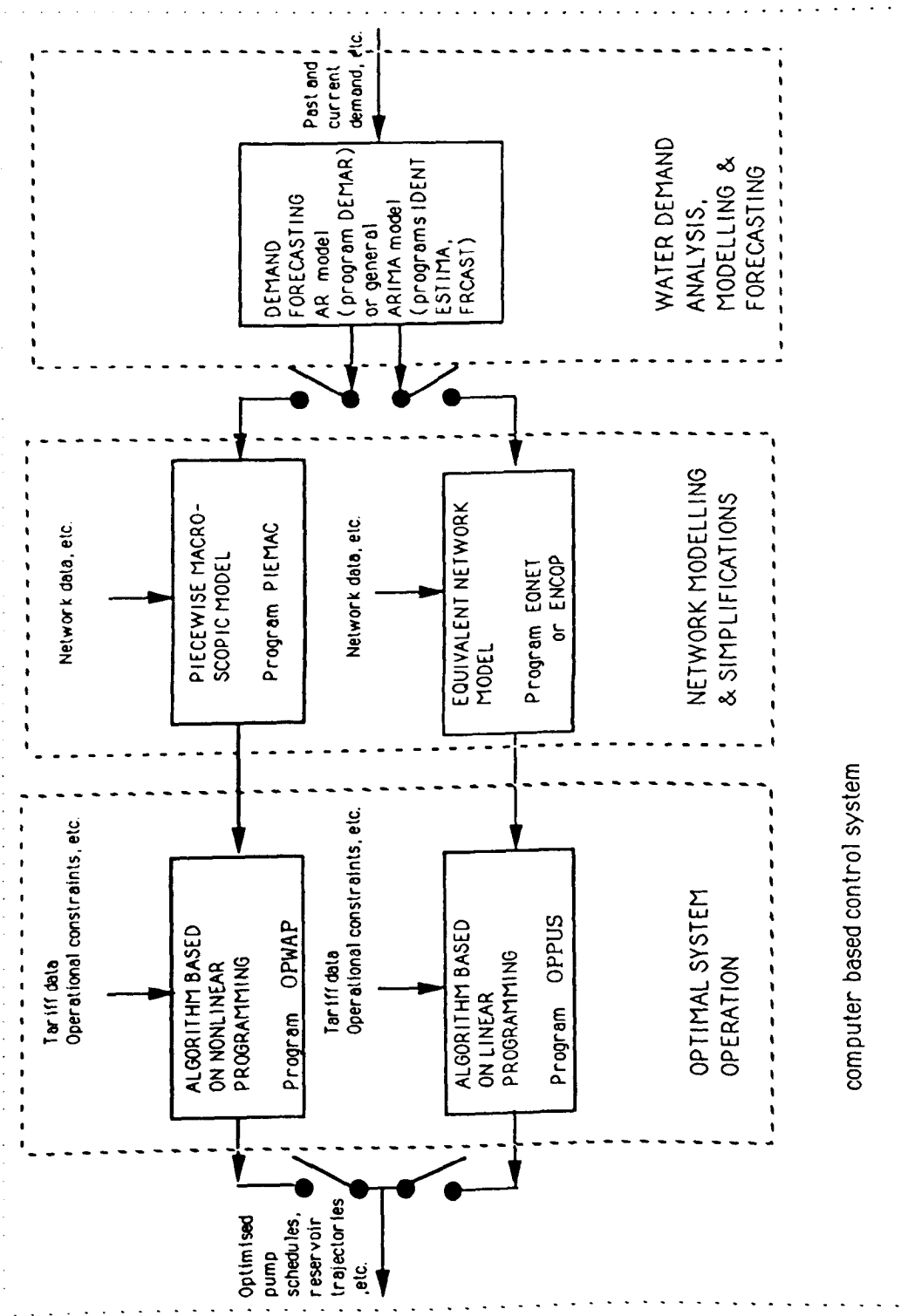


FIGURE 6.1 AN INTEGRATED CONTROL SCHEME FOR A WATER SUPPLY & DISTRIBUTION SYSTEM

6.2 GENERAL CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

This thesis describes a comprehensive, and complete, methodology for the overall optimization of water systems operations. However, further work could be done to improve and refine the methods based upon implementation results. In this section of the thesis, general conclusions are drawn and further work is suggested.

Chapter 3 Water Demand Analysis and Forecasting

Chapter 3 showed in detail how time series analysis is applied to short-term water demand forecasting. Among the general class of multiplicative ARIMA models, pure AR models offer distinct advantages over other models. Namely, parameter estimations are relatively simple and the determination of the model is likewise simple. Such factors lend themselves to less complex formulation and easier implementation. Further investigation into the combination of different forecasting models would be helpful in further raising forecasting precision.

AR models are, however, only applicable to a class of water demand patterns. Studies have shown that, by employing the Box-Jenkins approach to general ARIMA models, a more extensive range of demands could be modelled in a more systematic manner. The model building procedure is however rather complex and tentative, which would present an obstacle to the non-expert users and would complicate implementation procedures.

The models developed so far would benefit from further investigation

in the following areas:

- a) the study of the deterministic components of demand data in order to incorporate a deterministic function into the demand model (QUEVEDO et al, 1988).
- b) how to take account of the effect of external factors such as weather conditions. This could lead to a multivariate time series model or a mixed time-series/regressional model.
- c) the introduction of a data preprocessing capability similar to the data screening process described by Coulbeck (COULBECK et al, 1985). Such a process would remove noise and enhance the on-line performance of the prediction models.
- d) the design of a methodology for tuning of the time series parameters to enhance capability for on-line implementation.
- e) further work on the computer program in order to make it possible that the computer program can deal with the identification, parameter estimation, diagnostic checking, and forecasting of general ARIMA models automatically. This is particularly important for non-expert users.

Chapter 4 Network Modelling and Simplification for System Operation

In this chapter, several methods were presented for the modelling and simplification of water networks for purposes of on-line optimal

control. During development, particular emphasis was placed upon the need to reduce solution times in order to make such control feasible.

The macroscopic model developed by Demoyer (DEMOYER and HORWITZ, 1975a, 1975b) used a set of explicit nonlinear regression equations, to represent the major features of a detailed network. However, the applicability of such a model is restricted by the assumptions of proportional loading. Such assumptions are not always realistic for many water supply and distribution networks. The author's extensions to this approach has culminated in the piecewise macroscopic model. Such a model is further applicable to those systems in which the load patterns are not too inconsistent with the assumptions of proportional loading. The model has provided a good foundation for the development of an optimization method, described in Chapter 5. The idea of dividing a day into several proportional loading periods has been studied but additional research is required to derive more precise and comprehensive rules.

The equivalent network model developed by the author is based on the well-known nodal equations. The modelling methodology was developed around the concept of fictitious pipes. The model derivation has a clear and definite physical interpretation and is consistent with the conventional theory of water network analysis. In particular, and in contrast to the macroscopic and piecewise macroscopic models, the method does not rely on the assumption of proportional loading. The method was applied to the Wolverhampton water supply and distribution system with satisfactory results. The derivation of the model has facilitated the application of the linear programming optimization method described in Chapter 5. It should be worthwhile to investigate

its application to more varied systems and to enhance the model accordingly. For example, its performance when used in conjunction with different optimization algorithms could be studied. It should also be worthwhile to conduct research on the incorporation of pump controls into the model directly.

Finally, it will be worth making numerical comparisons between the equivalent network model and the macroscopic model or the piecewise macroscopic model in order to further learn their applicabilities.

Chapter 5 Optimal System Operations

Two algorithms were presented in Chapter 5. One catered for water systems without an storage. The other catered for those with such storage. The former is directly applicable to those systems without significant distribution storage. These include some sub-systems and small supply and distribution systems. The algorithm treats the pumping and treatment costs as the objective function. By adopting the piecewise macroscopic model of Chapter 4, as opposed to the iterative solution of nonlinear simultaneous equations, the objective function is solely in terms of station flows. The pressure requirements of certain system nodes can also be modelled. The formulation of the optimal operation problem using such an approach takes the form of a constrained nonlinear programming problem. Test results indicate that cost savings of up to 11% can be obtained. The extension of this algorithm to incorporate the piecewise macroscopic relationship for reservoirs to lead to a dynamic problem is expected to be straightforward. This should extend the applicability of the

algorithm to multi-source and multi-reservoir systems, although the computation of optimal solutions by the nonlinear programming technique will present greater problems.

The second algorithm can be applied to multi-source, multi-reservoir systems in which pump flows are more or less constant throughout the control period. The formulation of the optimal operation problem, when based on the equivalent network model (Chapter 4), results in a dynamic linear programming problem, where the decision variables are pump times rather than pump flows. The discretization scheme incorporated into the algorithm overcomes the difficulties of using fixed (discrete variables) and variable speed and/or variable throttle (continuous variables) pumps together. The algorithm has been fully tested on a real system and confirms the hypothesis that least-cost pump schedules and reservoir trajectories are indeed obtainable. The algorithm should also cope with those systems with fluctuating pump flows. This can be done by taking the average pump flows and pump powers, whereupon the solutions thus derived will provide a good reference for optimal operation. Further research is required on the inclusion of electrical maximum demand charges. This would be essential for winter operations.

The algorithm could be extended to include treatment costs without affecting the linearity of the formulation. Finally, in an attempt to speed up solution times for very large systems having many VSP's, it is worth considering the possibility of employing decomposition techniques for linear programming, and/or compact storage techniques.

REFERENCES

- [1] ABRAMOV, N.: Methods of Reducing Power Consumption in Pumping Water in Water Supply Systems, Including the Use of Booster Stations. Moscow Construction Engineering Institute, 1975
- [2] AN, H.Z.: Time Series Analysis and Its Applications. Science Press, China, 1983.
- [3] ASHCROFT, J.A., ELDRIDGE, R.H., PAULSON, R.W. and WILSON, G.A.: Programming With Fortran 77. Collins Professional and Technical Books, William Collins Sons & Co. Ltd, U.K., 1986.
- [4] BLANCHARD, B.S., FABRYCKY, W.J.: Systems Engineering and Analysis. Englewood Cliffs, N.J., London, U.K., 1981.
- [5] BOX, G.E.P. and JENKINS, G.M.: Time Series Analysis, Forecasting and Control. (Revised Edition), Holden-Day, San Francisco, U.S.A., 1976.
- [6] BRDYS, M.: Hierarchical Methods and Application to Water Supply Systems. Proceedings of the 9th IFAC Congress, Budapest, Hungary, 1984
- [7] BRDYS, M., COULBECK, B. and ORR, C.H.: A Method for Scheduling of Multi-Source, Multi-Reservoir Water Supply Systems Containing

Only Fixed Speed Pumps. Proc. IEE International Conference, Control 88, Oxford, U.K., April 1988.

- [8] BRDYS, M.: An Algorithm for Optimal Scheduling of Water Supply Systems Containing Only Fixed Speed Pumps. (submitted for publication), 1988.
- [9] BROGAN, W.L.: Modern Control Theory. Quantum Publishers, New York, U.S., 1974.
- [10] BUCKLEY, R.V.: Control Engineering, Theory, Worked Examples and Problems. Macmillan, London, U.K., 1976.
- [11] BUNDAY, B.D.: Basic Linear Programming. Edward Arnold (Publishers)Ltd, U.K. 1984
- [12] CARLSON, C.E.C.: The Denver System of Water Works Controls. Journal of AWWA, U.S., Aug. 1971
- [13] CEMBRANO, G., BRDYS, M., QUEVEDO, J., COULBECK, B. and ORR, C.H.: Optimization of a Multi-reservoir Water Network Using a Conjugate Gradient Technique -- A Case Study. Proceedings, 8th International Conference, Analysis and Optimization of Systems, Antibes, France, 1988
- [14] CHEN, Y.C.: Optimal Dispatch of Water Supply Systems (in Chinese). MSc thesis, Tongji University, Shanghai, China, 1985
- [15] CHEN, Y.C.: Optimal Control of Water Distribution Systems with

a Microcomputer (in Chinese). China Water & Wastewater, Vol.2, 1986

- [16] CHEN, Y.C.: Optimization of Water Supply Systems (in Chinese). Proceedings of the First Conference of China Environmental Scientists, Beijing, China, 1987a.

- [17] CHEN, Y.C.: Mathematical Modeling of Water Supply Networks (in Chinese). Proceedings of Symposium on Mathematics and Its Applications, Guangzhou, China, 1987b.

- [18] CHEN, Y.C.: Application of Time Series Analysis to Water Demand Predictions. In COULBECK, B. AND ORR, C.H.(eds): Computer applications in Water Supply, Vol.1, Systems Analysis and Simulations, Research Studies Press, U.K. and John Wiley & Sons Inc., U.S., 1988a

- [19] CHEN, Y.C.: Simplification of Water Supply and Distribution Systems for Optimal Operations. In COULBECK, B. AND ORR, C.H.(eds): Computer Applications in Water Supply, Vol.2, Systems Optimization and Control, Research Studies Press, U.K. and John Wiley & Sons Inc, U.S. 1988b.

- [20] CHEN, Y.C., COULBECK, B. and ORR, C.H.: Computer Control of Water Supply-- Network Simplification by Equivalent Modelling. Research Report No.64, Leicester Polytechnic, U.K. 1988a.

- [21] CHEN, Y.C., COULBECK, B. and ORR, C.H.: Computer Control of

Water Supply-- Studies on Water Demand Forecasting. Research Report No.72 , Leicester Polytechnic, U.K. 1988b.

- [22] COHEN, G.: Optimal Control of Water Supply Networks. In TZAFESTAS, S.G.(ed): Optimisation and Control of Dynamic Operational Research Models. North-Holland Publishing Company, 1982.
- [23] COHEN, S.S. and STANLEY, S.: Operational Research. Edward Arnold, London, U.K., 1985.
- [24] COLLINS, M.A., COOPER, L. and KENNINGTON, J.L.: Multiple Operating Points in Complex Pump Networks. Proc. ASCE, HY3, March, 1979.
- [25] COULBECK, B.: Optimization & Modelling Techniques in Dynamic Control of Water Distribution Systems, PhD thesis, University of Sheffield, U.K. 1977.
- [26] COULBECK, B. and STERLING, M.J.H.: Modelling Techniques in Dynamic Control of Water Distribution Systems. Measurement and Control, No.11, 1978a.
- [27] COULBECK, B and ORR, C.H.: Optimized Pumping in Water Supply Systems -2. Research Report No.33, Leicester Polytechnic, U.K., Dec., 1983.
- [28] COULBECK, B. and STERLING, M.J.H.: Optimised Control of Water Distribution Systems. Proc. IEE, Vol.125, No.9, Oct. 1978b.

- [29] COULBECK, B., ORR, C.H., TENNANT, S.T. and RANCE, J.P.: Computer Control of Water Supply-- The Development of a Demand Prediction Program for Use in Optimal Control of Water Supply, Research Report No.38, Leicester Polytechnic, U.K. 1985.
- [30] COULBECK, B.: Computer Control of Water Supply-- Applications Documentation for GINAS (A Graphical Interactive Network Analysis and Simulation Program). Research Report No.16, Leicester Polytechnic, U.K. 1985.
- [31] COULBECK, B. and ORR, C.H.: Computer Control of Water Supply-- GINAS Applications Workshop Training Guide, Research Report No.40, Leicester Polytechnic, U.K., 1986.
- [32] COULBECK, B.: A Review of Methodologies for Modelling and Control of Water Supply. In COULBECK, B. and ORR, C.H.(eds): Computer Applications in Water Supply, Vol.2, Systems Optimization and Control, Research Studies Press, Ltd, U.K. and John Wiley & Sons Inc. U.S. 1988.
- [33] COULBECK, B., ORR, C.H., PARKAR, M.A. and RANCE, J.P.: GINAS5 Applications Workshop Training Guide. Research Report 62, Leicester Polytechnic, U.K., 1988.
- [34] CREASEY, J.D.: Pump Scheduling in Water Supply: More Than a Mathematical Problem. In COULBECK, B. and ORR, C.H.(eds): Computer Applications in Water Supply, Vol.2, Systems

- Optimization and Control, Research Studies Press, Ltd, U.K., and John Wiley & Sons Inc, U.S., 1988.
- [35] CUMMINGS, W.J.: The Operation of the Control Systems of the Yorkshire Derwent Scheme. Progress of Water Technology, Vol.9, Nos. 5/6, Pergamon Press, U.K. 1977.
- [36] DEKAY, C.F.: The Evolution of Water Demand Forecasting, Journal of AWWA, U.S. Oct. 1985.
- [37] DEMOYER, R. JR., and HORWITZ, L.B.: Macroscopic Distribution-System Modelling, Journal of AWWA, July, U.S. 1975a.
- [38] DEMOYER, R. JR. and HORWITZ, L.B.: A System Approach to Water Distribution Modelling and Control, Lexington Books, D.C. Heath & Co. Lexington, Mass., U.S. 1975b.
- [39] DILLINGHAM, J.H.: Automating Distribution System Operation. Journal of AWWA, U.S, March 1979.
- [40] FAIR, G.M., GEYER, J.C. and OKUN, D.A.: Water and Wastewater Engineering, Vol.1, Water Supply and Wastewater Removal, John Wiley and Sons, Inc. U.S. and Toppan Company, Ltd, Japan, 1966.
- [41] FALLSIDE, F. and PERRY, P.F.: Hierarchical Optimisation of A Water-Supply Network. Proc. IEE, Vol.122, No.2, U.K. Feb. 1975a.
- [42] FALLSIDE, F. and PERRY, P.F.: Hierarchical Model for Water-

- Supply-System Control. Proc. IEE, Vol.122, No.4, U.K. April 1975b.
- [43] FALLSIDE, F. and PERRY, P.F.: On-line Prediction of Consumption for Water Supply Network Control. presented at the IFAC World Congress, Boston, Mass., U.S. 1975c.
- [44] FALLSIDE, F. and BURCH, R.H.: On-line Control of a Water Supply System. Progress of Water Technology, Vol.9, Nos.5/6, Pergamon Press, U.K. 1977.
- [45] FLETCHER, R.: Practical Methods of Optimization. Second Edition, Wiley, Chichester, U.K., 1987.
- [46] GILL, P.E., and MURRAY, W. and WRIGHT, M.H.: Practical Optimization. Academic Press, London, U.K., 1981.
- [47] GRACIE, G.: Analysis of Distribution Demand Variations. Journal of AWWA, U.S. Aug. 1971.
- [48] GRENDER, V. and ROSENBLATT, M.: Statistical Analysis of Stationary Time Series. Wiley, New York, U.S., 1957.
- [49] HARPER, W.H. and LIM, H.C.: Operational Research. Macdonald & Evans, Plymouth, U.K., 1982.
- [50] HILSDON, C.W.: Using Pump Curves and Valving Techniques for Efficient Pumping. Journal of AWWA, U.S. March, 1982.

- [51] HSIA, T.C.: System Identification, Least-Squares Method. Lexington Books, U.S. 1977.
- [52] JEPPSON, R.W.: Analysis of Flow in Pipe Networks. Ann Arbor Science Publishers, Inc. U.S. 1979.
- [53] JEPPSON, R.W.: Equivalent Hydraulic Pipe for Parallel Pipes. Proc. ASCE HY1, U.S. Jan. 1982.
- [54] JOALLAND, G. and COHEN, G.: Optimal Control of a Water Distribution Network by Two Multi-Level Methods. Automatica, 16, 1980.
- [55] JOWITT, P.W., GARRETT, R.T., COOK, S.C. and MONTOLIU, J.M.: Real-Time Forecasting and Control for Water Distribution. In COULBECK, B. and ORR, C.H.(eds): Computer Application in Water Supply. Vol.2, Systems Optimization and Control, Research Studies Press, Ltd, U.K. and John Wiley and Sons, Inc. U.S. 1988.
- [56] KANBAYASHI, T.: Computer Control and Operation Information System for the Large Water Supply Systems. Progress of Water Technology, Vol.9, Nos.5/6, Pergamon Press, U.K. 1977
- [57] KENNEDY, W., Jr. and GENTLE, J.E.: Statistical Computing. M. Dekker, New York, U.S., 1986.
- [58] LASDON, L.: Optimization theory for Large Systems. The

Macmillan Company, London, 1970.

- [59] LEWIS, F.L.: Optimal Control. John Wiley and Sons, Inc. New York, U.S., 1986.

- [60] LIU, S.R. and DUAN, W.Z.: A Study of Modelling Techniques on Water Supply Systems (in Chinese). Proceedings of the Conference for Applications of Automation Technology, China, 1986.

- [61] MARQUARDT, D.W.: An Algorithm for Least-squares Estimation of Non-linear Parameters. Jour. Soc. Ind. Appl. Math., 11, U.S.A., 1963.

- [62] MATSUMOTO, K.: Control Methods for Computer Operation of the Large Water Supply System. preprints IAWPR, Japan, MAY, 1977.

- [63] MATSUMOTO, K. and MIYAOKA, S.: Development of a Hierarchical Operation Scheduling Model and its Application to Large Scale Supply Systems. IFAC 8th Triennial World Congress, Kyoto, Japan, 1981.

- [64] MCPHERSON, M.B.: Generalized Distribution Network Head Loss Characteristics. Trans. ASCE, Part I, U.S. 1960.

- [65] MCPHERSON, M.B.: Ground Storage Booster Pumping Hydraulics, Journal of AWWA, U.S. 1966.

- [66] MOSS, S.M.: On-line Optimal Control of Water Supply Networks.

PhD thesis, Cambridge University, U.K. 1979.

- [67] MURTAGH, B.A.: Advanced Linear Programming. McGraw-Hill Inc., U.K., 1981.
- [68] NAG Fortran Library Routine Document, U.K. 1983.
- [69] OHMURO, T. and TACHIBANA, Y.: An Efficient Control Method for Operation of Multiple Pumps. IFAC 8th World Congress, Kyoto, Japan, 1981.
- [70] ORR, C.H., COULBECK, B., RANCE, J.P. and KORZENIEWSKI, R.E.: Computer-Aided Pump Scheduling for Efficient Operation of Water Supply Systems Using the Program GIPOS3. Research Report No.42, Leicester Polytechnic, U.K., March 1986.
- [71] ORR, C.H., COULBECK, B., BRDYS, M. and PARKAR, M.A.: Computer Control of Water Supply and Distribution Systems: Structure, Algorithms and Management. In COULBECK, B. and ORR, C.H.(eds): Computer Applications in Water Supply, Vol.2, Systems Optimization and Control, Research Studies Press, Ltd, U.K. and John Wiley & Sons, Inc., U.S.A. 1988a.
- [72] ORR, C.H., COULBECK, B. and RANCE, J.P.: Computer Aided Pump Analysis, Design and Simulation Using the Program GIPADS. Research Report No.43, Leicester Polytechnic, U.K. 1988b.
- [73] OSTLE, B.: Statistics in Research, Second Edition. The Iowa State University Press, U.S.A. 1974.

- [74] PARKAR, M.A., ORR, C.H., COULBECK, B. and TENNANT, S.T.: Description and Solution Methodology for Optimal Pump Scheduling of the Nurton/Bushbury System. Technical Report WCU/STWA-9, U.K., Feb., 1989.
- [75] PARZEN, E.: Some Recent Advances in Time Series Modelling. IEEE Transactions on Automatic Control, Vol.AC-19, No.6, U.S., Dec. 1974.
- [76] PERRY, P.F.: An On-line Optimal Control Scheme for Water Supply Networks. PhD thesis, Cambridge University, U.K., July, 1975.
- [77] PERRY, P.F.: Decentralized Optimum Control Methods for Water Distribution Systems Optimization. Proc. of the 1977 Joint Automatic Control Conference, Vol.2, U.S. 1977.
- [78] PERRY, P.F.: Demand Forecasting in Water Supply Networks. ASCE HY9, U.S. Sept. 1981.
- [79] PIERRE, D.A.: Optimization Theory with Applications, John Wiley & Sons, Inc., U.S.A., 1969.
- [80] QUEVEDO, J. and CEMBRANO, G.: Water Demand Forecasting through Time Series Analysis. Proc. Water and Data Processing, Presse Ponts et Chaussees, Paris, France, 1986.
- [81] RADER, J.E.V.: Telemetric Control of Water Distribution Systems.

- [82] RAO, H.S., BREE, D.W. and BENZVI, R.: Extended Period Simulation of Water Distribution Networks. Final Report by Systems Control Inc. Document No. PB-230 148, Palo Alto, California, U.S. Feb. 1974.
- [83] RAO, H.S. and SEITLE, R.A.: Computer Applications in Urban Water Distribution System Control. Journal of Dynamic Systems, Measurement and Control, U.S.A., 1975.
- [84] RAO, H.S. and BREE, D.W., Jr.: Extended Period Simulation of Water Systems -- Part A. Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol.103, No.HY2, U.S., Feb. 1977a.
- [85] RAO, H.S., MARKEL, L.C. and BREE, D.W., Jr: Extended Period Simulation of Water Systems -- Part B, Journal of Hydraulics Division, Proceedings of American Society of Civil Engineers, Vol.103, No.HY3, U.S. March, 1977b.
- [86] RATKOWSKY, D.A.: Nonlinear Regression Modelling, A Unified Practical Approach, M. Dekker, New York, U.S., 1983.
- [87] SABET, M.H.: Increasing Energy Efficiency of Ground Water Supply Systems. PhD thesis, Dept. of Civil Engineering, University of California, Davis, U.S., 1982.
- [88] SABET, M.H. and HELWEG, O.J.: Cost Effective Operation of Urban

- Water Supply System Using Dynamic Programming. Water Resources Bulletin, American Water Resources Association, Vol.21, No.21, U.S., Feb. 1985.
- [89] SAGE, A.P.: Systems Engineering, Methodology & Applications. IEEE Press, New York, U.S., 1977.
- [90] SHAMIR, U. and HOWARD, C.D.D.: Water Distribution System Analysis. Proc. ASCE HY1, U.S., Jan. 1968.
- [91] SHAMIR, U.: Optimal Design and Operation of Water Distribution Systems. Water Resources Research, U.S., Feb. 1974.
- [92] SHAMIR, U. and HOWARD, C.D.D.: Engineering Analysis of Water-Distribution Systems. Journal of AWWA, U.S., Sept. 1977.
- [93] SHAMIR, U.: Computer Applications for Real-time Operation of Water Distribution Systems. in Computer Applications in Water Resources, ed by TORNO, H., ASCE, U.S., 1985.
- [94] SINGH M.G. and TITLI, A.: Systems: Decomposition, Optimization and Control, Pergamon Press, U.K., 1978.
- [95] STERLING, M.J.H. and COULBECK, B.: A Dynamic Programming Solution to Optimization of Pumping Costs. Proc. Institute of Civil Engineers, Part 2, U.K., Dec. 1975a.
- [96] STERLING, M.J.H. and COULBECK, B.: Optimization of Water Pumping

- Costs by Hierarchical Methods. Proc. Institute of Civil Engineers, Part 2, U.K., Dec. 1975b.
- [97] STERLING, M.J.H. and COULBECK, B.: Linear Dynamic Models in Water Systems. Proc. Institute of Civil Engineers, U.K., Part 2, Sept. 1976.
- [98] TAHA, H.A.: Integer Programming, Theory, Applications and Computation. Academic Press, New York, U.S., 1975.
- [99] TAMURA, H.: Dual Algorithm for Solving Large-Scale Scheduling Systems Control Problem with Hierarchical Structure. Proc. of the IFAC 5th World Conference, 1972.
- [100] TENNANT, S.T., COULBECK, B. and ORR, C.H.: Computer Control of Water Supply, A System Description of GIDAP, Research Report No.46, Leicester Polytechnic, U.K., 1986.
- [101] TENNANT, S.T.: Short Term Demand Analysis and Prediction for Control of Water Supply. MPhil thesis, Leicester Polytechnic (CNA), U.K., 1987.
- [102] TSUKIYAMA, M. and FUKUDA, T.: The Control of Water Supply Network Based on Fuzzy Information. Proc. of the 1981 Conference of Japan Water Works Association, Japan, 1981a.
- [103] TSUKIYAMA, M. and FUKUDA, T.: A Heuristic Methodology of Estimation and Control for Large Scale Water Supply Network. IFAC 8th World Congress, Kyoto, Japan, 1981b.

- [104] VANDAELE, W.: Applied Time Series and Box-Jenkins Methods. Academic Press, Inc., New York, U.S.A., 1983.
- [105] WALKER, R.: Water Supply, Treatment and Distribution. Englewood Cliffs, Prentice-Hall, London, U.K., 1978.
- [106] WAN, Y.C.: Programming Collections for Optimization Techniques (in Chinese). Workers Press, China, 1983.
- [107] WRc: Pump Scheduling in Water Supply. Water Research Centre, TR232, U.K., 1985.
- [108] YANG, Q.: Design and Computation of Water Distribution Networks (in Chinese). Constructional Engineering Press, China, 1956.
- [109] YANG, Q.: On Computer Applications to Hydraulic Analysis of Water Distribution Networks (in Chinese). Journal of Tongji University, Shanghai, China, 1979.
- [110] ZARGHAMEE, M.S.: Mathematical Model for Water Distribution Systems. Proc. ASCE, HY1, U.S., Jan. 1971.

PAGES NOT SCANNED AT THE
REQUEST OF THE UNIVERSITY

SEE ORIGINAL COPY OF THE THESIS
FOR THIS MATERIAL