



**Pilkington Library**

---

Author/Filing Title ..... PASSONE .....

.....

Vol. No. .... Class Mark ..... T .....

---

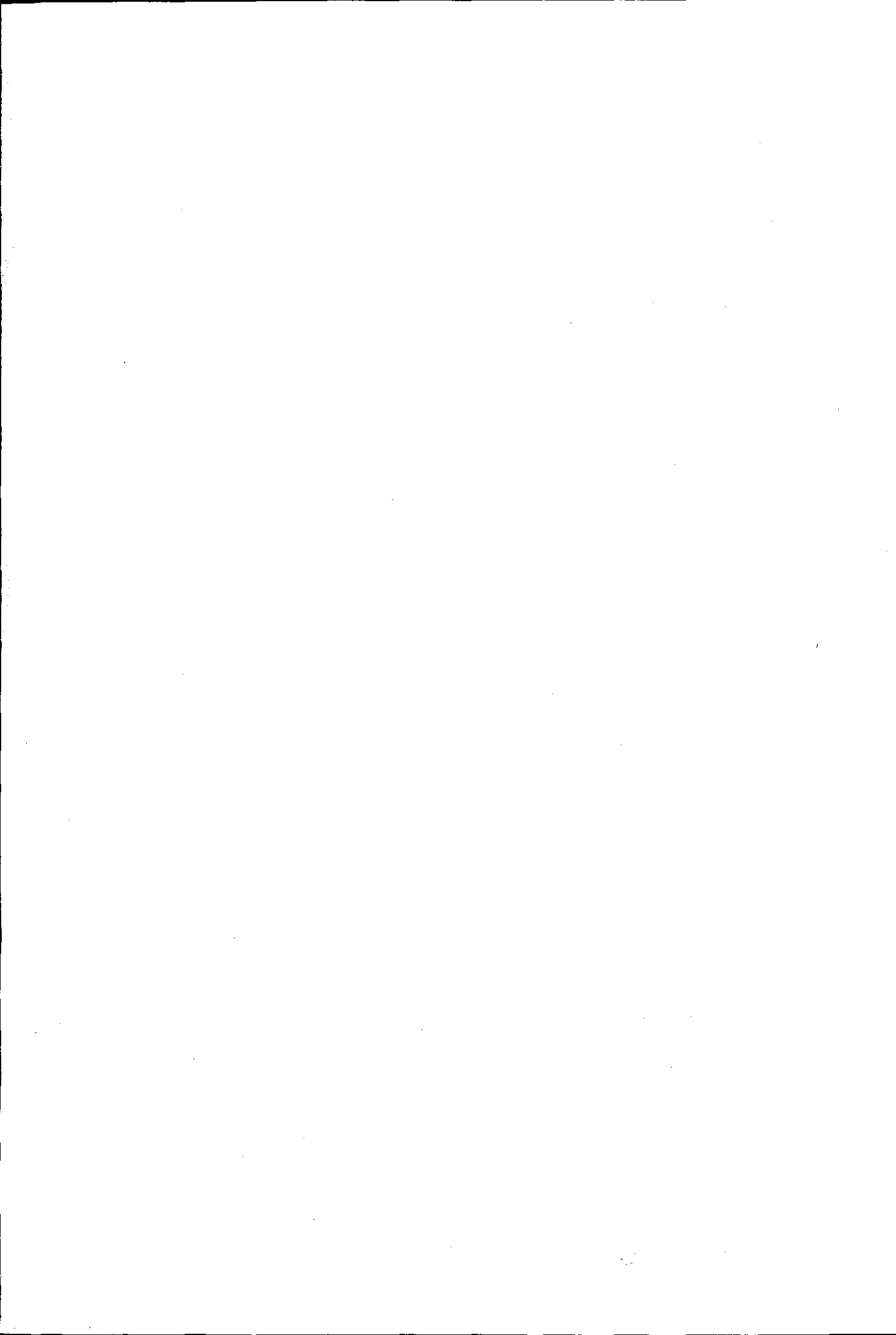
**Please note that fines are charged on ALL  
overdue items.**

--	--	--

FOR REFERENCE ONLY

0402805968





**A HYBRID REASONING SYSTEM FOR SUPPORTING  
THE MODELLING OF ESTUARIES**

By

Sara Passone


A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of

Doctor of Philosophy of Loughborough University

23<sup>rd</sup> November 2002

© by Sara Passone 2002

 <b>Loughborough University</b> Pillington Library
Date <i>Sept 01</i>
Class
Acc No. <i>040.280596</i>

## ABSTRACT

Estuaries are complex natural water systems. Their behaviour depends on many factors, which are possible to analyse only adopting different study approaches. The physical processes within estuaries are generally investigated through computer modelling. However, models are not easily accessible. Their employment is only possible within restricted conditions and assumptions. Furthermore, in depth knowledge is required to interpret the information related to different disciplines and sources for the selection of a correct modelling approach. Therefore, the usability of computational estuarine models appears lower than their actual capability. This thesis describes the application of case-based reasoning methodology to support the design of estuarine models. The system (CBEM - Case-Based reasoning for Estuarine Modelling) aims to provide a general user with the necessary guidance for selecting the model that better matches to his/her goal and the nature of the problem to be solved. The system is based on the co-operative action of three modules: a case-based reasoning scheme and a genetic algorithm and a library of numerical estuarine models. These components are integrated to work as a single tool. The main idea is for the system to learn and adopt solutions from past experience for their application in new problems. With respect to the possible correlation between the features of the estuary and the physical phenomenon to be modelled, the case-based module returns a suitable solution from the system's memory. The selected model is then adapted by the genetic algorithm component, which estimates a valid set of model parameters to suit the particular estuarine environment. Two case studies were carried out to evaluate the system's feasibility. The case studies of Upper Milford Haven and Tay estuaries demonstrate the system's ability to provide the user with a critical and correct estimation of the available model strategies and the robustness of the designed routine for adjustment of the model Manning's friction coefficient. CBEM permits to extend the application of computer-based numerical estuarine models to a broader class of users through an effective organisation and discovery of the knowledge necessary for the design of efficient modelling strategies.

## KEYWORDS

Case-Based Reasoning, Estuaries, Genetic Algorithms, Knowledge Base, Numerical Modelling

To My Parents

*There is a tide in the affairs of men,  
Which, taken at the flood, leads on to fortune;  
Omitted, all the voyage of their life  
Is bound in shallows and in miseries.  
On such a full sea are now afloat,  
And we must take the current when it serves,  
Or lose our ventures.*

*(William Shakespeare: Julius Caesar, Act IV, Scene 3)*

## ACKNOWLEDGEMENTS

I would like to thank:

Professors Paul Chung and Vahid Nassehi, my supervisors, for their encouragement and knowledge. Their guidance was more than about this PhD.

All my friends and colleagues, both in Chemical Engineering Department and outside, for their support in good and rainy days.

To the Engineering and Physical Sciences Research Council and Loughborough University for funding my project.

And finally, to my family and Ágúst and Fosca, for their love and help during the last three years.

## TABLE OF CONTENTS

Abstract .....	i
Acknowledgements .....	iii
List of Figures .....	viii
List of Tables .....	xi
<b>1. Introduction.....</b>	<b>1</b>
1.1. Motivation.....	1
1.2. Overview of the Project.....	1
1.3. Contributions .....	2
1.4. Structure of the Thesis .....	3
<b>2. Estuaries and Estuarine Model Design.....</b>	<b>4</b>
2.1. Introduction .....	4
2.2. Defining an Estuary .....	4
2.2.1. Classifications of Estuaries .....	5
2.2.1.1. Geomorphological Classification.....	6
2.2.1.2. Classification Based on Tidal Range .....	8
2.2.1.3. Classification Based on Circulation.....	8
2.2.1.4. Classification Based on Sedimentation.....	8
2.2.1.5. Quantified Classifications.....	10
2.3. Tidal Modelling.....	14
2.3.1. Hydrodynamical Model Equations .....	16
2.3.1.1. Model Characterisation.....	17
2.3.2. Dispersion Model Equation .....	18
2.3.3. Approximation of the Model Equations.....	18
2.3.4. Boundary Conditions .....	23
2.3.5. Model Parameters.....	24
2.3.6. Numerical Models .....	25
2.4. Conclusions .....	26
<b>3. Hydroinformatics .....</b>	<b>27</b>
3.1. Introduction .....	27
3.2. Old Problems .....	27
3.3. New Problems .....	28
3.4. A New Paradigm .....	29
3.5. Current Developments of Hydroinformatics .....	29
3.6. Conclusions .....	31
<b>4. Case-Based Reasoning .....</b>	<b>32</b>



4.1. Introduction .....	32
4.2. Case-Based Reasoning Definition.....	32
4.3. Motivation.....	33
4.3.1. Cased-Based Reasoning or Expert System? .....	33
4.3.2. The Advantages of Using CBR.....	34
4.3.3. CBR for Estuarine Modelling.....	34
4.4. The CBR Process .....	35
4.5. Case Description .....	36
4.6. Case Retrieval.....	37
4.6.1. Matching and Ranking Procedures .....	38
4.6.2. The Degree of Similarity – Fuzzy Matching.....	40
4.6.2.1. Numerical Values .....	40
4.6.2.2. Symbolic Values .....	41
4.7. Case Adaptation.....	41
4.8. Conclusions .....	43
<b>5. Genetic Algorithms .....</b>	<b>44</b>
5.1. Introduction .....	44
5.2. Motivation.....	44
5.2.1. The Traditional Optimisation Methods .....	46
5.2.2. The Advantage of Using Genetic Algorithms .....	47
5.3. Applications of Genetic Algorithms to Water Resources .....	48
5.4. The Fundamentals of Genetic Algorithms.....	49
5.4.1. Population Representation and Initialisation .....	51
5.4.2. Selection and Fitness Function .....	52
5.4.3. Crossover and Mutation.....	55
5.4.4. The Use of Specific Knowledge .....	57
5.5. Conclusions .....	57
<b>6. Case-Based Reasoning System for Estuarine Modelling.....</b>	<b>59</b>
6.1. Introduction .....	59
6.2. General Architecture .....	59
6.3. Case-Based Module.....	63
6.3.1. Case Description.....	63
6.3.1.1. Estuary Description.....	63
6.3.1.2. Model Description.....	67
6.3.2. Case Retrieval .....	69
6.3.2.1. Model Selection .....	69
6.3.2.2. Similar Estuaries .....	71
6.3.2.3. Similar Models .....	73
6.4. Genetic Algorithm Module .....	74

6.5. Model Library.....	77
6.6. Conclusions .....	77
<b>7. System Design.....</b>	<b>79</b>
7.1. Introduction .....	79
7.2. Case Description .....	80
7.2.1. The Average Width to the Average Depth Ratio.....	80
7.2.2. The Grade of Meandering .....	82
7.2.3. Estimation of the Salinity Stratification.....	83
7.2.4. Coriolis and Wind Forces .....	85
7.2.5. Inlets.....	86
7.3. Case Retrieval.....	86
7.3.1. Similar Estuaries.....	87
7.3.1.1. First Similarity Score .....	87
7.3.2. Similar Models.....	91
7.3.2.1. Criterion of Exclusion .....	91
7.3.2.2. Second Similarity Score .....	92
7.3.2.3. Criterion of Preference.....	92
7.4. Case Adaptation.....	95
7.4.1. Initalisation of the Population.....	96
7.4.1.1. Zonation.....	97
7.4.1.2. Scaling.....	98
7.4.1.3. Use of Case Information.....	98
7.4.2. GA Operators .....	98
7.4.2.1. Selection Mechanism.....	98
7.4.2.2. Crossover of Estuary's Zones .....	99
7.4.2.3. Mutation of Adjacent Genes .....	100
7.4.3. The Advantage of Using an Hybridised Genetic Algorithm.....	100
7.5. Conclusions .....	101
<b>8. Case Studies .....</b>	<b>104</b>
8.1. Introduction .....	104
8.2. Salt Intrusion .....	104
8.2.1. Problem Description.....	104
8.2.2. Search Query Model.....	106
8.3. Numerical Model for Simulating Salt Intrusion .....	106
8.3.1. Model Description.....	107
8.3.2. Input Data.....	108
8.3.3. Model Output.....	109
8.4. Upper Milford Haven Estuary.....	110
8.4.1. Introduction.....	110
8.4.2. Case Description.....	112
8.4.3. Case Retrieval .....	112

8.4.4. Case Adaptation – Genetic Algorithm Optimisation .....	115
8.4.4.1. Application of the Model to the Upper Milford Haven Estuary.....	115
8.4.4.2. Fitness Function Definition.....	116
8.4.4.3. Manning’s Friction Coefficient Calibration .....	117
8.4.4.4. Model Validation – Neap Tide.....	119
8.5. Tay Estuary .....	126
8.5.1. Introduction.....	126
8.5.2. Case Description.....	128
8.5.3. Case Retrieval .....	128
8.5.4. Case Adaptation – Genetic Algorithm Optimisation .....	131
8.5.4.1. Application of the Model to the Tay Estuary.....	131
8.5.4.2. Fitness Function Definition.....	132
8.5.4.3. Manning’s Friction Coefficient Calibration .....	133
8.5.5. Model Validation – Neap Tide.....	134
8.6. Conclusions .....	135
<b>9. Conclusions and Future Work.....</b>	<b>142</b>
9.1. Conclusions .....	142
9.2. Future Work.....	143
<b>References .....</b>	<b>147</b>
<b>Appendices</b>	
A CBEM Cases .....	154
B CB Module – Code .....	157
C GA Module – Code .....	183
D Published and Presented Work .....	195

## LIST OF FIGURES

Figure 2.1. Geomorphological classification from the Estuary Review Classification .....	7
Figure 2.2. Estuary type as function of river flow, tidal force, channel width and depth.....	10
Figure 2.3. Stratification-circulation diagram by Hansen and Rattray .....	12
Figure 2.4. Stratification-circulation diagram showing the contours of the parameters P and $F_m$ .....	13
Figure 5.1. Steps of a genetic algorithm.....	49
Figure 5.2. Chromosome representation using the binary representation.....	50
Figure 5.3. Genetic operators: a) Single-Point Crossover, b) Multi-Point Crossover, c) Uniform Crossover and d) Mutation .....	56
Figure 6.1. Structure of the CBEM system .....	60
Figure 6.2. Case representation as an estuarine problem is given .....	61
Figure 6.3. Estuary description .....	66
Figure 6.4. Inlet description.....	67
Figure 6.5. Model description.....	68
Figure 6.6. Model selection.....	70
Figure 6.7. Similarity rating with respect to the estuary physical and hydrographic features ...	72
Figure 6.8. Similarity rating with respect to the model characteristics.....	73
Figure 6.9. Example of a set of Manning's coefficients and representation of the corresponding chromosome .....	74
Figure 6.10. Structure of the implemented genetic algorithm.....	78
Figure 7.1. The membership functions used to express qualitatively the average width to the average depth ratio .....	81
Figure 7.2. The membership functions used to express qualitatively the degree of meandering. .....	83

Figure 7.3. Simulations of water surface elevations at Easthook, Upper Milford Haven estuary, UK, employing a manual calibration and a classical genetic algorithm optimisation schemes .....	97
Figure 7.4. Example of chromosome initialised using the zonation and scaling options.....	98
Figure 7.5. Chromosomes before and after the crossover operation.....	99
Figure 7.6. Chromosome before and after the mutation operation.....	100
Figure 7.7. a), b) c) and d) Simulations of water surface elevations at Easthook, Upper Milford Haven estuary, UK, obtained for different option combinations (table 7.9) included in the hybridised genetic algorithm optimisation scheme .....	103
Figure 8.1. The Upper Milford Haven Estuary, Wales, UK.....	110
Figure 8.2. Description of the Upper Milford Haven estuary, Wales, UK.....	111
Figure 8.3. Problem statement.....	112
Figure 8.4. Set of cases retrieved from the system's case-base ranked with respect to their similarity to the Upper Milford Haven estuary .....	113
Figure 8.5. Cases retrieved after the second phase of the retrieval process .....	114
Figure 8.6. Finite element discretisation of the Upper Milford Haven estuary (a Manning's coefficient is assigned to each element in the domain) .....	116
Figure 8.7. Simulation of the water surface elevations at a) Port Lion, b) Picton Point and c) Easthook during the spring tide of 25 <sup>th</sup> April 1979.....	121
Figure 8.8. Salinity profiles - a) Dauceddau and Western Cleddau, and b) Dauceddau and Eastern Cleddau during the spring tide of 25 <sup>th</sup> April 1979.....	122
Figure 8.9. Simulation of the water surface elevations at a) Port Lion, b) Picton Point and c) Easthook during the neap tide of 2 <sup>nd</sup> May 1979.....	124
Figure 8.10. Salinity profiles - a) Dauceddau and Western Cleddau, and b) Dauceddau and Eastern Cleddau during the neap tide of 2 <sup>nd</sup> May 1979.....	125
Figure 8.11. The Tay Estuary, Scotland, UK.....	126
Figure 8.12. Description of the Tay estuary, Scotland, UK.....	127
Figure 8.13. Problem statement .....	128

<b>Figure 8.14.</b> Set of cases retrieved from the system's case-base ranked with respect to their similarity to the Tay estuary.....	129
<b>Figure 8.15.</b> Cases retrieved after the second phase of the retrieval process.....	130
<b>Figure 8.16.</b> Discretisation of the Tay estuary (a Manning's coefficient is assigned to each element in the domain).....	132
<b>Figure 8.17.</b> Simulation of water surface elevations at a) Newport and b) Flisk during the spring tide of 12 <sup>th</sup> June 1972.....	136
<b>Figure 8.18.</b> Simulation of water surface elevations at c) Newburgh and d) Inchyra during the spring tide of 12 <sup>th</sup> June 1972.....	137
<b>Figure 8.19.</b> Salinity profile during the spring tide of 12 <sup>th</sup> June 1972.....	138
<b>Figure 8.20.</b> Simulation of water surface elevations at a) Newport and b) Flisk during the neap tide of 20 <sup>th</sup> June 1972.....	139
<b>Figure 8.21.</b> Simulation of water surface elevations at c) Newburgh and d) Inchyra during the neap tide of 20 <sup>th</sup> June 1972.....	140
<b>Figure 8.22.</b> Salinity profile during the neap tide of 12 June 1972.....	141
<b>Figure A.1.</b> Description of the Tees estuary, UK.....	154
<b>Figure A.2.</b> Description of the Fal estuary, UK.....	155
<b>Figure A.3.</b> Description of the Conwy estuary, UK.....	156

## LIST OF TABLES

Table 2.1. Estuary geomorphological types .....	6
Table 2.2. Classification based on tidal range .....	9
Table 2.3. Classification based on salinity .....	9
Table 2.4. Salinity Classification with respect to the ratio of R to V.....	11
Table 5.1. Roulette wheel mechanism.....	54
Table 5.2. Roulette wheel selection .....	54
Table 7.1. Membership grades automatically assigned when the user selects the qualitative ratio of the average width to the average depth.....	82
Table 7.2. Membership grades automatically assigned when the user selects the grade of meandering .....	84
Table 7.3. Degrees of relevance for the features from the estuary description utilised to calculate the first similarity score with respect to the salinity distribution problem .....	88
Table 7.4. Degrees of relevance for the features from the estuary description utilised to calculate the first similarity score with respect to the salt intrusion problem .....	88
Table 7.5. Degrees of match with respect to the index of simulation purpose .....	93
Table 7.6. Degrees of match for the index of simulation time consumed with respect to the simulation purpose of the new investigation.....	93
Table 7.7. Degrees of match for the index of accuracy with respect to the simulation purpose of the new investigation .....	93
Table 7.8. Set of degrees of relevance used to weight the importance of model purpose, simulation time required and accuracy in the computation of the second similarity score.....	93
Table 7.9. The four combinations of GA options utilised to illustrate the advantages given by the hybridised genetic algorithm for estuarine calibration over the classical genetic algorithm scheme .....	101
Table 8.1. Cases currently included within the CBEM's case-base.....	105
Table 8.2. Manning's friction coefficients for the Upper Milford Haven estuary .....	118
Table 8.3. Manning's friction coefficients for the Tay estuary .....	134

## *Chapter 1*

### INTRODUCTION

#### **1.1 Motivation**

The physical behaviour of estuaries is intrinsically complex as they are the places where river and sea waters interact and mix. In addition, high concentration of human activities in these areas has resulted in an ever increasing risk of heavy pollution in estuarine systems. Numerical modelling is considered the most effective tool for studying the complex behaviour of these natural water systems and for predictive simulations of the effects of man-made or natural changes within the estuaries. This is because of the efficiency and increasing ability of numerical modelling systems to produce realistic and comprehensive results. However, modelling hydraulic phenomena involves many different skills. In particular, in order to use these systems correctly it is necessary to have expertise in mathematical representation of the physical processes involved and numerical solution of the derived models. In addition, a sound and exhaustive background in the hydrodynamics of natural water systems is usually needed. The expertise required for the application of numerical schemes in estuary modelling is relatively extensive and has become a limiting factor that restricts the use of these methods. Therefore, modelling of estuaries has, to some extent, remained an art rather than an exact science.

There is a necessity to facilitate the application of estuarine models so that they can be employed by users who have different backgrounds in estuarine science. This can be achieved by guiding the user through model selection. A computer-based environment needs to be developed where models are combined with a library of knowledge to help to identify a suitable modelling strategy with respect to the type of problem to be simulated and the user's requirements. Modelling packages must be part of a systematic structure where solutions are provided through a critical reasoning process. This will facilitate the exploitation of the full potential of estuarine models by reducing time and effort needed in multidisciplinary works of this nature.

#### **1.2 Overview of the Project**

This project aims to investigate means of enhancing the flexibility, cost-effectiveness of



models and making them more accessible to non-specialists. The idea is to provide the user with the necessary guidance and information for the numerical solution of hydro-environmental problems by organising the available knowledge of estuarine modelling into an interactive and dynamic framework.

In order to provide "intelligent support" for model design, in this project a Case-Based reasoning system for Estuarine Modelling (CBEM) has been developed. This is based on the utilisation of the case-based reasoning technique in combination with numerical modelling and genetic algorithms. Case-based reasoning (CBR) is an artificial intelligence methodology in which the solution of a new problem is derived from similar situations encountered in the past. CBEM contains examples of past modelling studies of estuaries. These are codified and organised in the system memory according to a set of common problem descriptors. When a new problem is given to the system, it is compared with past cases contained in the memory. After comparison and evaluation, the case-based search engine returns the most effective modelling scheme available for the solution of the new problem. After finding the best match, the system then optimises, through a genetic algorithm, the physical parameters in the selected model to obtain the most appropriate values for the new application. Therefore, after the parameter adjustment, the user has access to the most suited modelling procedure and physical parameters for a given problem. The information about the new case is also stored in the case-memory so as to extend its knowledge base.

### **1.3 Contributions**

The prime objective of this research is to develop a computer-based system to support the design of numerical models for estuaries. The investigation aims to extend the application of computer-based estuarine models to a broader class of potential users and facilitate the communication of results to policy-makers concerned with water management schemes. The novelty of the approach is in combining numerical modelling with artificial intelligence technologies such as case-based reasoning and genetic algorithm techniques. The knowledge contained is structured in the system in such a way that it enables the potential users, who may not have in depth knowledge of modelling, to define the type of problem and the characteristics of the estuary and obtain meaningful numerical simulations.

A genetic algorithm is specifically implemented for the automatic adjustment of the Manning's

friction coefficient. This coefficient is the most important physical property used for calibration and verification of estuary models.

The application of CBEM in two case studies is then carried out to demonstrate CBEM's feasibility. In particular, the accessibility of computer programs has increased and they have become more cost-effective.

#### **1.4 Structure of the Thesis**

The thesis is organised into 9 chapters including the introduction. In chapter 2 basic definitions of estuaries and estuarine modelling technique are reviewed. This includes discussions on the behaviour of estuaries and the design of tidal models. Chapter 3 defines the newly emerged field of "hydroinformatics". This is the area into which computational hydraulics has evolved and provides the context for the current research.

Chapters 4 and 5 present the principles of the case-based reasoning methodology and genetic algorithm theory, respectively. Chapters 6 and 7 give an overview of the developed CBEM. Chapter 6 illustrates the work-flow and chapter 7 addresses the design and the implementation issues encountered. In chapter 8 the application of CBEM to the cases of the Upper Milford Haven and Tay estuaries is given. Concluding remarks and considerations about possible future work are provided in chapter 9.

## *Chapter 2*

### ESTUARIES AND ESTUARINE MODEL DESIGN

#### **2.1 Introduction**

The objective of this chapter is to describe the basic characteristics of estuaries and estuarine behaviour and give an account of the analysis method used to understand and solve real estuarine problems using numerical modelling techniques. This sets the scene for the research work presented in this thesis. Section §2.2 introduces estuaries and briefly describes their behaviour while section §2.3 is concerned with tidal modelling and solution of estuarine problems using mathematical and numerical techniques.

#### **2.2 Defining an Estuary**

Physical, chemical and biological processes should be modelled in a unique way (A.W. Morris, 1983). Estuaries are under the influence of tides, weather, seasonal river flows and climate. All these aspects control the variation of the water level, salinity, temperature, and sediment load and, consequently the general behaviour of the water-body. Furthermore, the human activities that have grown around almost all estuarine areas have had a great impact on the fragile equilibrium of these water systems. The intensive use of estuaries for transportation, food production, waste-disposal, flood protection, recreation, and other purposes have dramatically modified the morphology and the ecosystem of these water courses (M.J. Kennish, 1986).

Although the individual features of each estuary must be taken into account, there are several recognisable characteristics and behaviours among estuaries which remain similar (R.W. Fairbridge, 1980). Estuaries are studied from different points of view by investigators from many disciplines, and estuarine research is in continuous expansion. Oceanographers, engineers and natural scientists such as biologists and geophysicists, have been studying estuarine systems for many decades and they have defined them in a variety of ways, with over 40 different definitions of estuaries proposed so far (G.M.E. Perillo, 1995).

This mainly depends on the features of interest studied by a group of experts. For example estuarine hydrodynamics may be of prime importance or estuaries may be regarded as reactors where salt water mixes with fresh water inflow from land drainage. The emphasis might also be

on the magnitude and direction of tidal currents. In addition, the definition of estuaries may focus on sedimentation or even on the communities of organisms living in such a complex chemical environment (M.J. Kennish, 1986). However, these widely used definitions only reflect the point of view of a particular discipline and do not have universal applicability. They are limited to a specific characterisation of the estuary and its environment with the exclusion of features that may be of key interest for other disciplines.

Among all the existing definitions, the one given by the Natural Environment Research Council (N.E.R.C., 1975) is preferred in the present context as it proposes a multidisciplinary description of estuaries. By covering all the wide range of features that combine to form these watercourses, an estuary is defined "*...As a partially enclosed body of water, open to saline water from the sea and receiving freshwater from rivers, land run-off and seepage...Estuaries are subject to a usually twice-daily tidal rise and fall and they have mud and sand shoals forming in their shallow basins. Other characteristic features of estuaries include the presence of saltmarshes, shelter from waves, water layering and mixing, temperature and salinity gradients, sediments suspension and transport, high productivity, high levels and rapid exchange of nutrients, the presence of plants and animal particularly adapted to these conditions, and the presence of migrants and seasonally fluctuating populations of animals*".

### 2.2.1 Classification of Estuaries

To predict the characteristics of any estuary a classification scheme is necessary to provide a valid and general framework (K.R. Dyer, 1997). However, a general flexible indexing scheme which can represent all of the fundamental aspects of estuaries and clearly explain the mechanisms of interaction between physical, chemical and biological subsystems, is not available (J.G. Wilson, 1988). Instead, there are several classifications which depending on the criteria used can be considered (K.R. Dyer, 1997). The bases on which these classifications are developed are diverse, focusing on specific aspects of the estuarine environment including geological, physical and chemical factors. The most common classifications are based on geomorphology and physiography, tidal characteristics, hydrography (water circulation, mixing processes and stratification) and sedimentation (availability, transport and distribution of sediments) within estuaries.

### 2.2.1.1 Geomorphological Classification

Estuaries may be classified according to their geomorphology. Although this provides a simple classification, it is useful for understanding the general characteristics of estuaries. Moreover this classification provides an insight about mechanisms, focusing in particular on why and where they occur, what are the common features and how they affect the estuary processes. The geomorphological classification is based on Fairbridge's classification (R.W. Fairbridge, 1980), who extended the topographic classification developed by Pritchard (D.W. Pritchard, 1952). The classification considers factors such as the regional history of the sea level, tectonic and climatic processes, and freshwater and sediment movements, and takes into account long-term processes. According to the geomorphological classification six categories, shown in table 2.1, are identified. The categories are arranged in terms of the relative relief and the degree of blocking of the estuarine system (figure 2.1).

**Table 2.1.** Estuary geomorphological types

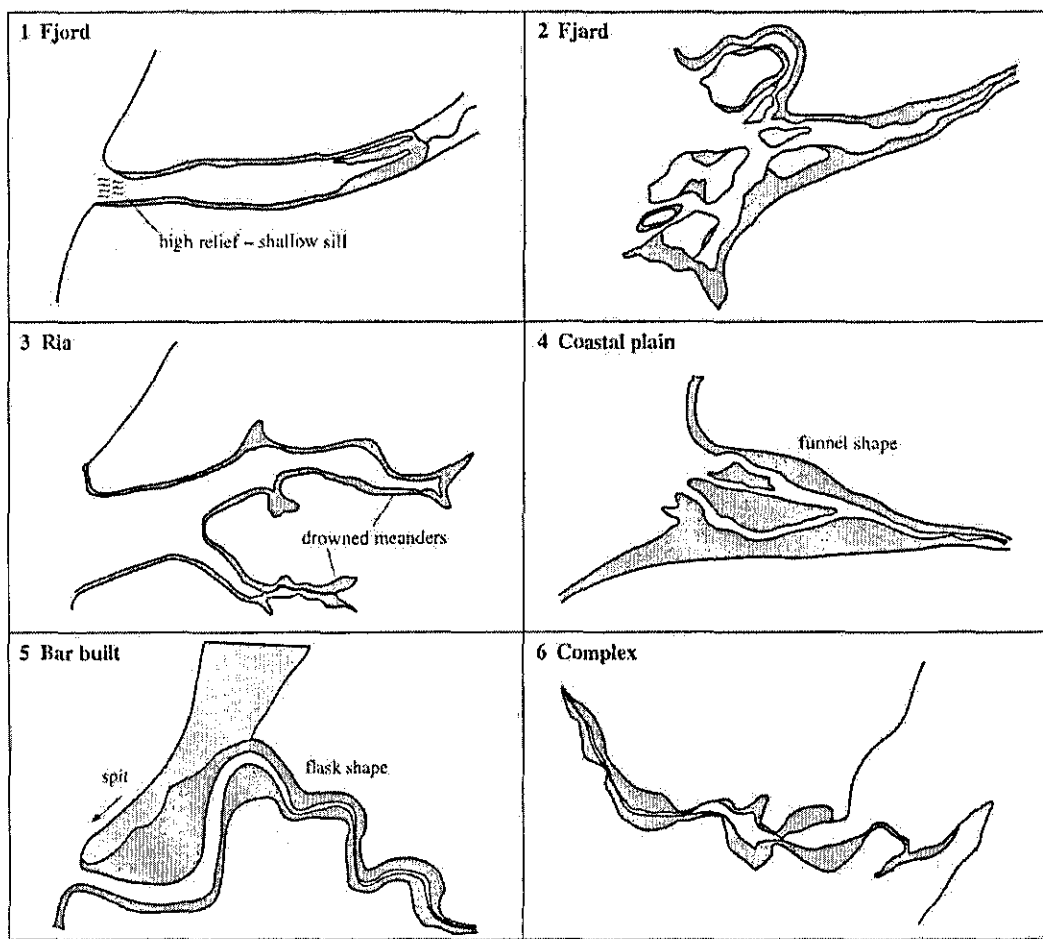
TYPE	CHARACTERISTICS
Fjord	High relief estuary, U shaped valley profile, very low width to depth ratio (i.e. 10:1), very low sinuosity, rectangular cross section.
Fjard	Similar to fjords without the U shape.
Ria	Moderate relief estuary, V cross section, well-defined channel, moderate width/depth ratio.
Coastal plain	Low relief estuary, branching sinuous valley, funnels shaped plan, large width/depth ratio, triangular cross section.
Bar built	Low relief estuary, L shaped plan, large width/depth ratio.
Complex	Estuary with complex origin and physiography such as the large firths in Scotland.

Estuaries can also be considered under other geomorphological aspects. The physical form of estuaries may depend on latitudinal and climatic conditions. Thus, estuaries can also be categorised into three latitudinal-climatic groups (R.W. Fairbridge, 1980), namely:

- polar and subpolar zones
- westerly temperate zones

- tropical and equatorial zones.

Fjords and fjards are typical at polar and subpolar latitude. A feature of many middle and low latitude estuaries is the sandy baymouth bar or barrier, which is absent in high latitude estuaries usually sculpted in hard rock. In some intermediate latitudes (45°-55°N), for instance in Britain, there is a large amount of morainal debris which is moved by wave action to generate bars in glacially scoured estuary. The westerly-temperate zone, located approximately around 55°-30°, is geomorphologically characterised by vegetative cover, apart from those areas where there are human settlements, and large amount of sediments. Between 30° N and 30° S the tropical and equatorial zone is located. Without considering any biological difference, in this



**Figure 2.1.** Geomorphological classification from the Estuary Review Classification (N.C. Davidson et al., 1991)

zone there is a consistent sedimentation, which leads to accumulations (i.e. banks and bars) of considerable dimension.

#### *2.2.1.2 Classification Based on Tidal Range*

This classification scheme is based on the relation between tidal range (the height difference between low tide and high tide) and estuary type (J.L. Davies, 1973; M.O. Hayes, 1975). The tidal range has a considerable effect on the determination of tidal currents, residual current velocities and rate, amounts and patterns of the movement of sediments.

Three different types of tidal ranges, as presented in table 2.2, are identified. To enable comparison between estuaries and minimise the effect of an estuary's individual morphology the range of spring tides are measured at the mouth (N.C. Davidson et al., 1991).

#### *2.2.1.3 Classification Based on Circulation*

This classification, proposed by Pritchard (D.W. Pritchard, 1967), is based on the mixing mechanisms that govern estuarine dynamics (table 2.3). River discharge is mixed with the seawater by the action of tidal motion and wind stress on the surface, and the river discharge flowing towards the sea. The tidal processes and other important estuarine features are affected significantly by salinity (A.W. Morris, 1983), which is an indicator of mixing and pattern of water circulation (K.R. Dyer, 1997). Salinity related factors are the horizontal tidal movement at the estuary mouth, the outflow of fresh water to the estuary from the river and the geometry of the estuary (i.e. estuary width and depth). It should be noted that there are differences in the circulation patterns even for estuaries of the same geomorphological type. Estuarine behaviour is a function of river flow, tidal currents, channel width and depth (figure 2.2). Thus, since all the factors are time-dependent, types A, B, C and D represent stages of a continuous sequence, with estuaries changing in type with time.

#### *2.2.1.4 Classification Based on Sedimentation*

The sedimentation and the relationship between sedimentation and salt intrusion may depend on the Coriolis force and intermittent effects such as wind and the temperature difference between the river water and seawater, which in some zones may generate secondary circulations. Sedimentological criteria have been used to develop a separate classification scheme for the estuaries (M.J. Kennish, 1986). According to this classification, estuaries which are entirely controlled by river transported sediments are called "positive filled". On the other

hand, if the sediments originate from the beaches and the nearshore ocean, they are called "inverse filled". Estuaries, characterised by little or no change in basin volume due to sedimentation are defined as "neutral filled".

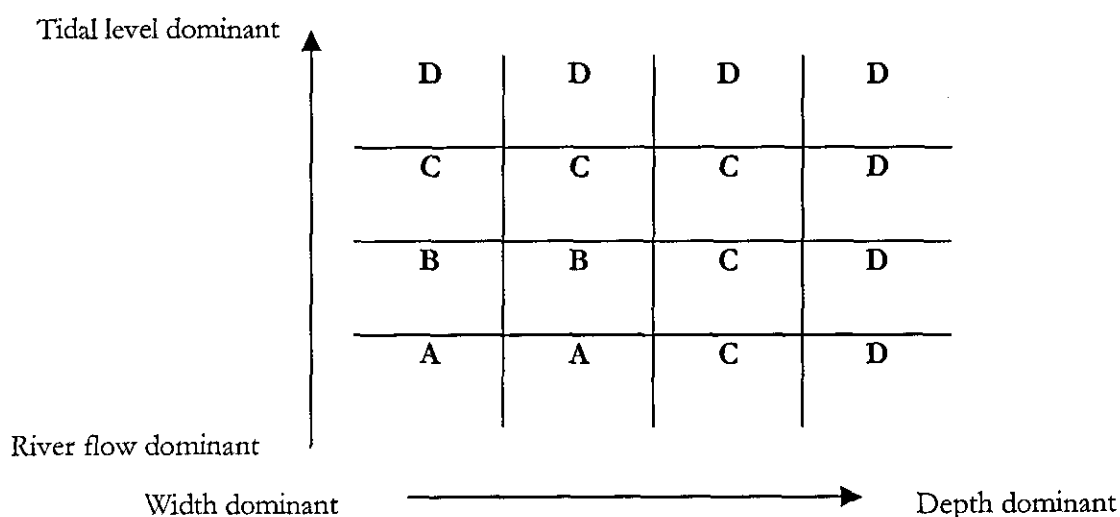
**Table 2.2.** Classification based on tidal range

TYPE	TIDAL RANGE	ESTUARY TYPE	REMARKS
microtidal	<2 m	Dominated by fresh water discharge; Shallow and wide.	Presence of wind-driven waves that produce spits and barrier islands (typical of bar built estuaries). Salt-wedge circulation.
mesotidal	2 - 4 m	Dominated by strong tidal currents; Often short and wide; Meandering tidal channel.	The salt-wedge circulation is excluded.
macrotidal	>4 m	Funnel-shape; Strong tidal and residual currents can extend far inland.	Typical of coastal plain estuaries dominated by the landforms such as flat and salt marshes.

**Table 2.3.** Classification based on salinity

TYPE	CIRCULATION	REMARKS
A	Highly stratified Salt wedge type	Two layered stratified estuaries dominated by high river flows; Salinity gradient (both longitudinal and vertical); Small tidal range and width to the depth ratio; Affected by Coriolis force.
B	Moderately stratified Partially mixed	Controlled by the action of river and tides; Salinity gradient: longitudinal, vertical and lateral; Typically shallow and wide.
C	Vertically Homogenous	Tidal flow greater than river discharge; Significant salinity gradient in longitudinal and lateral directions; High width to depth ratio; Affected by Coriolis and wind forces.
D	Well mixed	Dominated by tidal forces; Salinity gradient: longitudinal; Small the width to the depth ratio.





**Figure 2.2.** Estuary type as function of river flow, tidal force, channel width and depth

#### 2.2.1.5 Quantified Classifications

The estuary classification can also be carried out by considering different dimensionless parameters which characterise them considering river and tidal flows, and the magnitude of the resulting salinity stratification (K.R. Dyer, 1997).

As mentioned before the salinity field is controlled by the balance between the advection of fresh water and the diffusion of the salt due to the turbulence caused by the tidal forces. Thus, representing the fresh water volume entering the estuary from the river over one tidal cycle as  $R$ , and the volume of water brought into the estuary by the flood tide (i.e. tidal volume) as  $V$ , estuaries can be evaluated by considering the ratio of  $V$  to  $R$  over the same tidal period  $T$  (M. Tomczak, 1998) (table 2.4).

Alternatively, this definition can be given using the mean velocity ratio:

$$P = \frac{u_f}{u_t} \tag{2.1}$$

where  $u_f = R/AT$  with  $A$  representing the cross-sectional area of the flow-channel, and  $u_t =$  root mean square tidal velocity. If  $P < 10^{-2}$ , then the estuary is considered well-mixed, else, if  $P < 10^{-1}$ , it is said to be stratified (K.R. Dyer, 1997).

**Table 2.4.** Salinity classification with respect to the ratio of R to V

TYPE	R/V
Salt-wedge	>>1
Highly stratified - Fjords	0.1-1
Slightly stratified	0.005-0.1
Vertically mixed	<<0.005

Another number used to classify estuary in terms of salinity stratification is the 'Estuary Number' which is defined as

$$Ne = \frac{PF_m^2}{TQ_f} \quad (2.2)$$

where  $Q_f$  is the rate of fresh water inflow,  $P$  is the volume of the tidal prism (i.e. volume of water within the estuary between high and low waters), and  $F_m$  is the densimetric Froude number which is expressed as

$$F_m = \frac{u_f}{\sqrt{gh(\Delta\rho/\rho)}} \quad (2.3)$$

with  $h$  representing the water depth and  $\Delta\rho$  the difference between the sea water density  $\rho$  and the fresh water density.

When  $Ne > 0.1$  the estuary is well mixed, otherwise it is stratified.

A fourth number can also be employed to classify estuary. This was originally defined by Ippen and Harleman (A.T. Ippen, 1966) and relates the level of stratification in an estuary to the energy lost by a tidal wave entering the flow channel as:

$$S_t = \frac{G}{J} \quad (2.4)$$

where  $G$  is the rate of turbulent energy dissipation per unit mass of fluid and  $J$  is the rate of potential energy gained per unit of mass over the entire length of the estuary (D.M. McDowell

and B.A. O'Connor, 1977).

G is expressed as:

$$G = \frac{P_{x1} - P_{x2}}{\rho b h (x_1 - x_2)} \quad (2.5)$$

where  $P_{x1} - P_{x2}$  corresponds to the rate of energy loss between the two channel cross-sections  $x_1$  and  $x_2$ .

J is given by

$$J = g \left( \frac{\Delta \rho}{\rho} \right) \frac{h u_f}{L} \quad (2.6)$$

where  $L$  is the distance over which the mean velocity  $u_f$  is estimated.

Another convenient quantified classification is provided by Hansen and Rattray (K.R. Dyer, 1997) who realised the need for more than a single parameter for classifying and comparing estuaries. They proposed a stratification-circulation diagram (figure 2.3) where an estuary is identified by two dimensionless numbers. These are the ratio of mean surface velocity  $u_s$  against fresh water velocity  $u_f$  ( $u_s/u_f$ ), called the circulation parameter, and the ratio of mean salinity difference between surface and bottom  $\Delta S$ , and vertically averaged mean salinity  $S_0$ ,  $\Delta S/S_0$ , called stratification parameter.

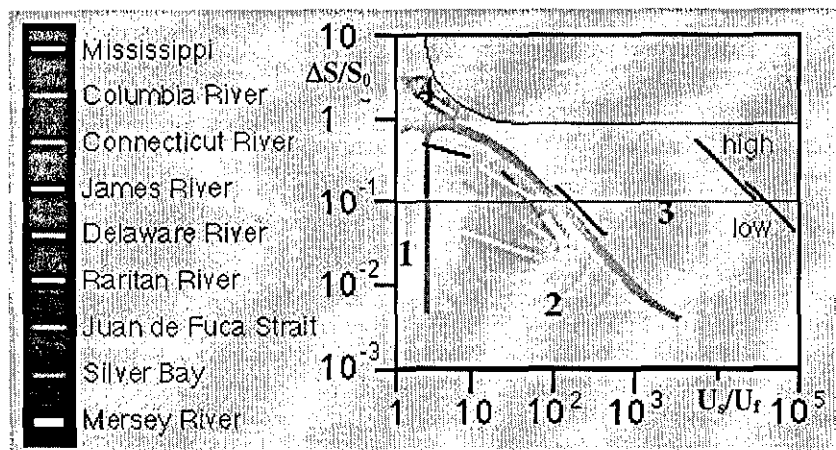
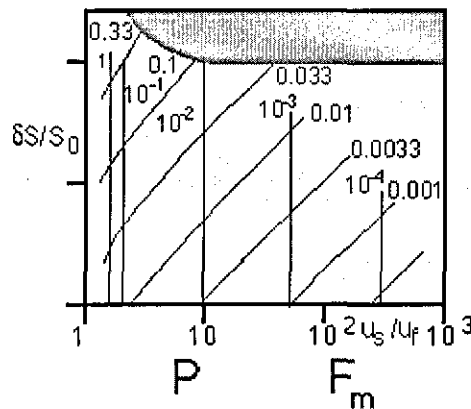


Figure 2.3. Stratification-circulation diagram by Hansen and Rattray (M. Tomczak, 1999) 12

In this classification type 1 represents well-mixed estuaries where the flow at all depths is seaward and diffusion accounts for upstream salt transport. Type 2 indicates a partially mixed estuary which is characterised by a net reverse flow at bottom and an upstream flux of salt. Type 3 corresponds to fjords, which are characterised by a differential flow in the two layers. Type 4 is the salt wedge estuary in which the intense stratification takes place with a very little interaction between layers. Types 1, 2 and 3 are subdivided into high and low categories according to the value of the stratification parameter ( $\Delta S/S_0 > 10^{-1} \rightarrow$ high,  $\Delta S/S_0 < 10^{-1} \rightarrow$ low).

Determination of these parameters requires the measurements of salinity as a function of depth and surface velocity over one tidal cycle, plus the measurements of river input  $R$  over one tidal cycle and the knowledge of the topography for the determination of the cross-sectional area  $A$  (K.R. Dyer, 1997).

The main disadvantage of this scheme is that velocity and salinity are calculated for a specific location along the estuary. Hence, for different points of a particular estuarine basin the state of stratification may change (D.A. Jay and J.D. Smith, 1988). However, it is possible to trace points on the stratification and circulation diagram corresponding to different cross-sections, for the same river discharge (figure 2.3). These points may be joined up to form a line (K.F. Bowden, 1980).



**Figure 2.4.** Stratification-circulation diagram showing the contours of the parameters  $P$  and  $F_m$  (M. Tomczak, 1999)

In order to use the diagram in a predictive way, the quantities  $\Delta S/S_0$  and  $u_s/u_f$  are related to other two parameters which are the flow ratio  $P$  (2.1) and the densimetric Froude Number (2.3). These two parameters are used to relate  $\Delta S/S_0$  and  $u_s/u_f$  to the forcing mechanisms (figure 2.4). While the circulation depends entirely on  $F_m$ , the stratification depends on both parameters.

Other numbers (Fischer, 1972; Scott, 1993) can also be used to classify estuaries according to the Hansen and Rattray scheme (K.R. Dyer, 1997), confirming that this classification provides a rather general mean to represent and compare estuaries. For instance, Fischer (1972) proposed to relate the estuarine circulation to the 'Estuarine Richardson number'. This number, which gives the ratio between the potential energy provided by the fresh water inflow and the mixing force of the tidal currents, is defined as:

$$Ri_e = \frac{\Delta\rho}{\rho} \frac{gQ_f}{bu_i^3} \quad (2.7)$$

where  $b$  is the estuary width.

As the value of  $Ri_e$  increases, so does the stratification of the estuary.

The 'Estuarine Richardson number', related to the densimetric Froude Number and the flow ratio (K.F. Bowden, 1980), is dependent mainly on stratification rather than circulation. Therefore,  $Ri_e$  may be used in combination with  $F_m$  for representing the type of circulation and stratification within an estuary. The two parameters together take into account the influence of river discharge, tidal current and estuary width and depth on the mixing mechanism.

### 2.3 Tidal Modelling

Tidal computations and modelling in estuaries procedure consists of a number of steps as:

- problem observation and assessment
- analysis of physical phenomena and degree of their relevance in the problem to be solved
- definition of the modelling strategy and possible numerical schemes
- estimation of the difference between the real behaviour and the model simulation.

Three general approaches for modelling can be considered: physical hydraulic models, analytical models and numerical models. Although the water movement can be reproduced with some degree of accuracy by physical models, the mixing processes throughout the estuary can not be scaled and simulated by hydraulic prototypes. Both analytical and numerical models are used to simulate the observed behaviour of an estuary, in response to a specific known set of input and particular conditions.

Mathematical models consist of (I.B. Hinwood and I.G. Wallis, 1975)

- a set of equations that describe water movement and related mechanisms such as mixing processes and sediment transport
- boundary conditions and empirical parameters which permit to adjust the model over a particular estuary.

These models are not as expensive as the physical models and they simulate effects that physical models cannot satisfactorily represent. Mathematical models are useful for evaluating the impact of engineering works on tidal motion and the quality of tidal water systems (D.M. McDowell and B.A. O'Connor, 1977). Many different models have been developed to fulfil different needs. The analytical models, developed in pre-computer days, can still be used for very simplified problems. However, research has been progressively concentrating on the study of specific systems by the use of numerical models and analytical schemes.

Mathematical models can be divided into three main categories (D.M. McDowell and B.A. O'Connor, 1977):

- flow models, used to simulate quantitative aspects of the shallow water regimes such as the water surface elevations and the flow velocities;
- mass transport models, based on the mass continuity principles, which include the concentration distribution of substances such as salt and reactive and inert pollutants, and the water quality parameters such as the biochemical oxygen demand (BOD) and its dissolved oxygen content (DO);
- sediment transport models, which are based on the equations of fluid motion and fluid and sediment movement.

Any model of an estuary should be based on a sound mathematical formulation and knowledge of estuarine conditions (Rijkwaterstaat Communications, 1974). Therefore, considering the complexity of the estuaries the construction of tidal models is not a simple task and an adequate model representation depends on many different factors. For instance, the selected model must be defined to include relevant complexity of a given estuary by taking into account the geometry of the estuary plane, its topography, river inflow at its head, exposure to the wind and presence of secondary currents.

The use and purpose of a model often drive the user's preference. The choice of a particular model is affected by not only the characteristics of the problem, but also the accuracy of their application and aim of the investigation. Special needs of some users can influence the degree of the accuracy required and consequently many assumptions and approximations may be made to obtain the desired balance between accuracy and costs. In order to have a realistic and sensible investigation the model selection must be done without an overestimation of the available capability with a careful consideration of the available resources and inevitable consequences deriving from approximations carried out (J.B. Hinwood and I.G. Wallis, 1975). For instance, averaging over one or more spatial dimensions produces a simpler and more economical model. However, this decision may imply many other assumptions and approximations, which must affect the model performance.

### *2.3.1 Hydrodynamical Model Equations*

The dynamic balance within an estuary is mathematically described by the equations of motion. These are also known as momentum and continuity equations and used to mathematically analyse the flow behaviour of estuaries.

The momentum equation concerns with the forces acting on water flowing through the Eulerian control volume. In a three-dimensional co-ordinate system, it is given as

$$\rho \frac{DU}{Dt} = -\nabla P + \mu \nabla^2 U + G \quad (2.8)$$

where  $\rho$ ,  $P$  and  $\mu$  are, respectively, the fluid density, pressure intensity, dynamic viscosity, while  $U$  and  $G$  represent the vectors of velocity and external forces in the x, y and z directions.

The equation of continuity is derived from the law of conservation of mass and is defined in a three-dimensional co-ordinate system as

$$\operatorname{div}(\rho U) = -\frac{\partial \rho}{\partial t} \quad (2.9)$$

### 2.3.1.1 Model Characterisation

In order to be able to describe the flow behaviour of an estuary, the equations of motion need to take into account

- the physical characteristics of the estuary
- the body forces on the water
- water density variations
- turbulence

Estuaries are characterised by the rapid variation of the estuarine cross-section over the length of the channel. Furthermore, estuaries may include geomorphological elements that make the use of equation (2.8) and (2.9) difficult. These can be floodplain, where the presence of water is intermittent (V. Nassehi and A. Kafai, 1999), branches, for which supplementary equations must be necessary (J.H. Bikangaga and V. Nassehi, 1995) and, bends and meanders, which determine energy losses that affect the velocity distribution (J.R. Rossiter and G.W. Lennon, 1965).

In addition to the gravity, other body forces can also be significant in estuaries such as (D.M. McDowell and B.A. O'Connor, 1977)

- wind tangential force
- estuary bed frictional force
- coriolis force

The wind force is responsible for additional shear force on the water surface, especially for estuaries that widen towards the sea. The bed frictional force depends on the type of material the bed channel is made of. It consists of additional shear force that occurs at the bottom of the channel. Finally, the Coriolis force, due to the rotation of the Earth, can have a significant



effect in wide estuary at latitude far from the equator.

Furthermore, density variations and turbulence can affect the flow (J.H. Bikangaga, 1993). Mainly dependent on salinity, density must be mathematically related to the changes in salinity due to the relative strength of river flow and tidal flow. In addition, turbulence can be taken into account by averaging equations (2.8) and (2.9) over a period of time that is long when compared to turbulent fluctuations, but shorter than the tidal period or any time interval required for significant changes in the water course hydrograph. As a result, additional normal and shear stresses is included to embody in the equations turbulence fluctuations.

### 2.3.2 Dispersion Model Equation

The distribution of dissolved matter into estuaries can be investigated using an appropriate continuity of mass equation also known as dispersion equation or convective-diffusion equation. Suitable for studying salt transport and the dispersion of solute pollutants (i.e. conservative and reactive) in estuaries, the equation is defined as

$$\frac{\partial C}{\partial t} + U_x \frac{\partial C}{\partial x} + U_y \frac{\partial C}{\partial y} + U_z \frac{\partial C}{\partial z} - \frac{\partial}{\partial x} \left( K_x \frac{\partial C}{\partial x} \right) - \frac{\partial}{\partial y} \left( K_y \frac{\partial C}{\partial y} \right) - \frac{\partial}{\partial z} \left( K_z \frac{\partial C}{\partial z} \right) = R \quad (2.10)$$

where C represents the concentration and R the sink/source term.  $U_x$ ,  $U_y$  and  $U_z$ , and,  $K_x$ ,  $K_y$ ,  $K_z$  indicate, respectively, the turbulent mean velocities and the coefficients of turbulent diffusion in x, y and z directions. R is null whether (2.10) is used for modelling the estuarine salt balance and conservative pollutants.

### 2.3.3 Approximation of the Model Equations

In modelling it may be necessary to simplify some aspects and ignore others in order to be able to understand some of the processes occurring in an estuary. Therefore, equations (2.8), (2.9) and (2.10) can be simplified by reducing the relative importance of the time or limiting the description of the problem domain to one or two-dimensional situations. Therefore, instead of using the full three-dimensional equations of motion, it may be more practical and realistic to describe the estuarine hydrodynamics through a simplified model.

It can be assumed that, since any vertical force component is small compared to the gravity, the momentum equation in z-direction can be reduced to

$$\frac{1}{\rho} \frac{\partial P}{\partial z} = -g \quad (2.11)$$

For instance, in a well-mixed estuary  $\partial\rho/\partial z$  is negligible (D.M. McDowell and B.A. O'Connor, 1977) and equation (2.11) can be written as

$$\frac{\partial p}{\partial x} = g \left( \rho \frac{\partial h}{\partial x} + h \frac{\partial \rho}{\partial x} \right) \quad (2.12)$$

If the density is also uniform then expression (2.12) can be reduced to

$$\frac{\partial p}{\partial x} = g\rho \frac{\partial h}{\partial x} \quad (2.13)$$

where h is the depth below the estuary surface.

Furthermore, equations (2.8) and (2.9) may be averaged obtaining a two-dimensional representation. For instance, a number of estuaries are wider than deep and a system of two-dimensional equations averaged on the depth may be suitable to describe the estuarine flow behaviour. Thus, the equation of continuity becomes

$$\frac{\partial h}{\partial t} + \frac{\partial(hU_x)}{\partial x} + \frac{\partial(hU_y)}{\partial y} = 0 \quad (2.14)$$

and the momentum equation is given as:

$$\frac{\partial U_x}{\partial t} + U_x \frac{\partial U_x}{\partial x} + U_y \frac{\partial U_x}{\partial y} + g \frac{\partial h}{\partial x} + \frac{\tau_{bx}}{\rho H} - \Omega U_y - W_x F_w = 0 \quad (2.15)$$

$$\frac{\partial U_y}{\partial t} + U_x \frac{\partial U_y}{\partial x} + U_y \frac{\partial U_y}{\partial y} + g \frac{\partial h}{\partial y} + \frac{\tau_{by}}{\rho H} + \Omega U_x - W_y F_w = 0 \quad (2.16)$$

where  $H$  is the river depth over which the average is carried out,  $U_x$  and  $U_y$  are the depth average velocities in the  $x$  and  $y$  directions,  $g$  is the gravitational acceleration, and  $\tau_{bx}$  and  $\tau_{by}$  represent the shear stresses due to the bed friction in the  $x, y$  directions.

The wind force is expressed by  $F_w$  which is defined as

$$F_w = \frac{\tau_{ws}}{\rho H} \sqrt{W_x^2 + W_y^2} \quad (2.17)$$

where  $\tau_{ws}$  indicates the wind stress,  $\rho$  is the water density, and  $W_x$  and  $W_y$  are the components of the wind velocity in  $x, y$  directions, respectively.

The shear stresses  $\tau_{bx}$  and  $\tau_{by}$  are related to the average velocity as follows:

$$\tau_{bx} = \frac{\rho g n^2 U_x}{H^{1/3}} \sqrt{U_x^2 + U_y^2} \quad (2.18)$$

$$\tau_{by} = \frac{\rho g n^2 U_y}{H^{1/3}} \sqrt{U_x^2 + U_y^2} \quad (2.19)$$

where  $n$  is the Manning's friction coefficient.

In equations (2.15) and (2.16)  $\Omega$  is defined as:

$$\Omega = 2\omega \sin(\varphi) \quad (2.20)$$

where  $\omega$  is the angular velocity of the Earth and  $\varphi$  is the angle of the latitude of the Earth.

A one-dimensional representation may also be enough when it is just required to determine the variation of the flow over the length of the estuary. The equations of motions are averaged on the cross-section and may be expressed in form of De St. Venant equations as follows:

$$\frac{\partial Q}{\partial x} + B \frac{\partial h}{\partial t} = q \quad (2.21)$$

$$\frac{\partial h}{\partial x} + \frac{1}{gA} \frac{\partial Q}{\partial t} + \frac{1}{gA} \frac{\partial}{\partial x} \left( \frac{Q^2}{A} \right) + \frac{|Q|Qn^2}{A^2 H^{4/3}} + \frac{H}{2\rho} \frac{\partial \rho}{\partial x} = 0 \quad (2.22)$$

where,  $B$ ,  $A$ ,  $Q$ , and  $q$  are, respectively, breadth of water surface, cross-sectional area, cross-sectional estuary discharge and fresh water inflow at the estuary head. The spatial and temporal lengths are denoted by  $x$  and  $t$ , respectively.

For different estuary types it is also expected that the different terms in equation (2.10) used for representing the salt balance become less significant and, hence, they can be neglected.

The terms can be reduced according to their respective importance regarding

- the degree of stratification present in the flow regime and its change with tide and freshwater conditions;
- the geometrical characteristics of the flow system, and
- the temporal and the spatial characteristics of the engineering problem itself (D.M. McDowell and B.A. O'Connor, 1977).

In particular, (2.10) may be averaged (J.H. Bikangaga, 1993):

- vertically obtaining a two-dimensional depth averaged equation

$$\frac{\partial C}{\partial t} + \frac{\partial(U_x C)}{\partial x} + \frac{\partial(U_y C)}{\partial y} = \frac{1}{H} \frac{\partial}{\partial x} \left( HD_x \frac{\partial C}{\partial x} \right) + \frac{1}{H} \frac{\partial}{\partial y} \left( HD_y \frac{\partial C}{\partial y} \right) + R_h \quad (2.23)$$

where  $D_x$  and  $D_y$  are the dispersion coefficient in  $x$  and  $y$  directions, respectively; and  $R_h$  is the depth averaged source/sink term;

- transversely obtaining a two-dimensional width averaged equation

$$\frac{\partial(b_s C)}{\partial t} + \frac{\partial(b_c U_x C)}{\partial x} + \frac{\partial(b_s U_z C)}{\partial z} = \frac{\partial}{\partial x} \left( b_s D_x \frac{\partial C}{\partial x} \right) + \frac{\partial}{\partial z} \left( b_s D_z \frac{\partial C}{\partial z} \right) + b_s R_b \quad (2.24)$$

where  $b_s$  and  $b_c$  are the storage and current widths, respectively,  $R_x$  is the source/sink term averaged on the width, and  $D_x$  and  $D_z$  are the dispersion coefficients in  $x$  and  $z$  directions, respectively;

- over the cross-section, reducing the equation to a one-dimensional scheme

$$\frac{\partial(AC)}{\partial t} + \frac{\partial(AUC)}{\partial x} - \frac{\partial}{\partial x} \left( AD_x \frac{\partial C}{\partial x} \right) = R_x + q_c \quad (2.25)$$

where  $D_x$ ,  $A$  and  $U$  are, respectively, the longitudinal dispersion coefficient, cross-sectional area and the cross-sectional velocity. The source term,  $q_c$ , specifies a point discharge while  $R_x$  is the cross-sectional averaged source/sink term.

The above approximations can be applied when the magnitude of some terms in equation (2.10) is so small that they can be neglected from the equation. For instance, the averaged over the depth equation (2.23) is suitable for wide and shallow estuaries with concentrations changes likely to be much smaller over the depth than the width of the estuary. It is also used for vertically mixed estuaries for which this equation gives almost as much information on the values of the concentration as the three-dimensional version (2.10) (J.B. Hinwood and I.G. Wallis, 1975).

Similarly, a two-dimensional averaged over the width (2.24) may be prescribed for studying the concentration distribution in narrow and meandering estuaries where transverse concentrations can be assumed uniform across of the width. Instead, if the estuary is narrow with very small vertical density gradient, the variation of the concentration may be satisfactorily described through a one-dimensional model (2.25) (J.B. Hinwood and I.G. Wallis, 1975).

Models can also be time averaged. Therefore, there are steady state models, tidally averaged models, and intratidal models. Tidally averaged models are simpler and more economical but their accuracy is limited. This type of models cannot simulate the concentration distribution throughout a tidal cycle. In case of steady state models, the concentrations are predicted for steady inflows and unvarying tidal conditions, especially when the upstream transport is balanced by downstream advective transport.

#### *2.3.4 Boundary Conditions*

Boundary conditions at the landward and seaward ends of the estuary are central to models. They play a significant part in determining the behaviour of the estuary under particular circumstances. Boundary conditions regard flow, concentrations of salt and, for model estimating the dispersion of pollutants, the concentration of other dissolved matters in the estuary.

At the landward end the flow boundary condition can be determined from river gauging records in case of river flow. If the river flow is null, it can be represented through a complete reflection of the tide. In addition, at the seaward end, tidal elevations need to be specified for the simulated period.

For the salt, while at the estuary head the condition of zero salinity is assumed, the prescription for the seaward salt boundary is more complicated. It is problematic to define realistic sea-ward salinity (J. Woodruff, 2002). This can be estimated by defining empirically the salinity gradient at the estuary mouth for each specific case. Some authors also choose to describe the salt concentration at the mouth as a combination of the salinity of the sea and the salinity of the last parcel of water leaving the estuary. The estuary outflow is assumed to influence the salinity at seaward open boundary for a certain period of time after which the salinity is essentially of the sea (P.A. Gillibrand and P.W. Balls, 1998). Finally, the seaward salt boundary condition can be represented through a constant salinity concentration. Especially, for large macrotidal systems with a wide cross-section at the mouth, it is more likely that the salinity will remain constant under most conditions (V. Nassehi and A. Kafai, 1999).

Similarly, in case of pollutants, if the source of pollutant is located at the estuary mouth the boundary condition may be represented as function of time value at the respective boundary (Dirichlet type). In addition, for modelling the dispersion of pollutants a Neumann boundary condition can also be employed (J.H. Bikangaga, 1993). This permits to describe the situation at the boundary as either as reflection of solid boundary, or a free movement of the matter across the boundary without backflow.

### 2.3.5 Model Parameters

Model parameters are the coefficients representing, respectively, the influence of the bed friction over the velocity distribution and the diffusion/dispersion of a particular substance transported by the estuarine waters. Their physical meaning is influenced by the assumptions and approximations introduced in the model equations.

In the momentum equation (2.8) the shear stress at the bed of the estuary is mainly affected by the bed morphology, and the type and distribution of the sediments observed in the estuary. Particularly significant in  $x$  and  $y$  directions, longitudinal and transversal shear stresses are expressed as a function of the corresponding velocity component multiplied by an empirical factor representing the effects of the estuary bed roughness.

The bed roughness may be described in terms of the Manning's friction coefficient. The values of this parameter has been carefully tabulated for representing the river bed friction to the flow. However, its values in case of estuaries may be slightly different from those adopted for river channels due to the reversal of the tidal flow which may cause different bed forms (D.M. McDowell and B.A. O'Connor, 1977). It must also be noted that with the simplification of the momentum equation, the Manning's coefficient needs also to incorporate those physical effects that, because of the approximation, are no longer considered. Thus, it is expected that the values of Manning's coefficient in (2.22) are higher than those suggested in case of a two-dimensional model (i.e. equations 2.15 and 2.16), as they represent not only the bed friction to the flow but also additional shear stress which counts for losses due to bends of the channel, changes in the estuarine cross-section and secondary circulations (D.M. McDowell and B.A. O'Connor, 1977).

For describing the mixing processes in an estuary, equation (2.10) is also expressed in terms of coefficients of turbulent diffusion. However, in (2.23), (2.24) and (2.25), defined for averaged values of concentration and velocity, the meaning of these coefficients is altered.  $D_x$ ,  $D_y$  and  $D_z$  in equations (2.23), (2.24) and (2.25) are dispersion coefficients which describe the transport of any substance due to both turbulent diffusion and the velocity variations within the estuary section. This also means that, as equation (2.10) is approximated and averaged over one or two dimensions, the resulting dispersion coefficients are uniquely related to the velocity structure and, therefore, must be determined for each flow situation encountered. The values of these

parameters may be obtained using analytical relationships, field data combined with governing equations or semi-empirical formulas (J.R. West and T.W. Broyd, 1979).

### *2.3.6 Numerical Models*

Numerical simulations, among the other types of models, are considered the most effective tool for studying estuaries. They provide very accurate simulations with minimum time consumed (G. Thompson, 1993), revealing important aspects of estuarine environmental dynamics, which may not be evident from field measurements and analytical evaluations (J.R. French and N. J Clifford, 2000). Furthermore, their use for investigating the hydrodynamics, the sediments movement and the mixing processes can be employed to explore different possible conditions.

However, there are different shortcomings in numerical models for estuaries. They are normally developed for particular situations costing a lot of time and money. As other applied sciences, numerical modelling requires substantial mathematical background to formulate and program the computational models, which few estuarine scientists have (P. Dyke, 1996). Furthermore, some of the computational programs are developed by scientists for personal use, making it difficult for other users to access them. Thus, although computational estuarine programs are useful utilities, their usability is usually low (H.J. Van Zuylén, 1994). Finally, the link between estuarine modelling and other aspects of the estuarine science is not widely represented in the model development and in the evaluation of the model results. It is generally required to have a very deep knowledge in order to interpret these complex systems from different points of view and to identify the correct modelling approach for a specific problem.

Numerical methods generally employed for numerically simulating the estuary behaviour are:

- Methods of characteristics
- Finite difference
- Finite element
- Spectral method
- Finite volume
- Diffuse approximation.



The detailed analysis of each of these methods is beyond the scope of this research work and will not be provided here.

#### **2.4 Conclusions**

Estuaries' behaviour depends on many factors, which are possible to analyse only adopting different study approaches. The physical processes within estuaries are generally simulated through computer modelling. In order to provide support for model design, an intelligent modeling environment should be implemented. Such a system has been developed and is called Case-Based Reasoning System for Estuarine Modelling (CBEM). The system is the subject of this thesis.

## HYDROINFORMATICS

This chapter aims to give an overview of hydroinformatics. This is a new paradigm in hydrosience and engineering which aims to bind advanced information technologies and the classic computational hydraulics. Hydroinformatics is also the discipline in whose area the research work presented in this thesis can be included.

### **3.1 Introduction**

CBEM is developed based on the widely accepted idea that the traditional computational hydraulics is no longer sufficient to cope with the complex water related problems that afflict the world's future (M.B. Abbott, 1991). Computational hydraulics developed as an engineering tool with its main focus on the technical and scientific aspects of a problem, to provide answers in relatively short times to water systems related problems (M.B. Abbott et al., 2001). Following this approach, numerical modelling for water engineering and environmental problems has become very sophisticated. Characterised by a great mathematical and computational complexity, modelling systems are now powerful software packages with user-friendly front-ends and a menu of facilities to support pre-processing and post-processing operations (K.W. Chau and W. Chen, 2001). However, although modern model design has resulted in increasing the number of users, there is still a great deal of problems, old and new, that these computer devices cannot solve.

### **3.2 Old Problems**

The increase in accessibility of numerical models has not diminished the necessity of the expertise required to use them. For instance, selecting a suitable model to solve a practical problem remains a very difficult task for which the user needs to have detailed knowledge about the application and possible limitations of the available models. Furthermore, during any phase of the modelling process, the user needs to use his/her expertise in order to achieve the modeling objectives. Hence, thorough exploitation of a model relies on the ability of its user to evaluate and interpret simulation outputs. In case of unrealistic predictions the user must be

able to detect where and why a mistake has been made. Although the current generation of models are very sophisticated and characterised by a great rate of automation, there is still a number of tasks such as the model calibration that cannot be completely automated because it requires the user's experience and judgment. Far from being perfect, these computer tools still lack the ability to evaluate the uncertainty which is usually present in natural water processes. In short, they cannot have and utilise the type of knowledge that only through experiencing a great number of situations is acquired (U. Cortes et al., 2001).

### **3.3 New Problems**

In addition to the problem described in the previous section it is increasingly evident that the scope and scale of water related problems require the involvement of a large number of various experts with different backgrounds. The solution of a problem does not simply rely on the correct estimation of possible scenarios, but needs to be integrated in a wider context so that it will be acceptable from different aspects. Thus, communication during all the phases of the solution process is important. Only through an efficient communication an acceptable final solution can be obtained. However, in order for the communication to be successful, it must be supported by methodologies that allow the parties involved to agree on the available information on which actions and decisions can be based. However, the current models only responds to main modelling requirements of effectiveness, precision and low cost. Therefore, future models need to be implemented with a user-oriented approach, where the modellers are not the final users of the results but, instead, those who need to base their decisions on the simulation are the consumers. Model must not only provide technically sound solutions but it should also support the decision-making process (J.A. Cunge and M. Erlich, 1999).

Numerical models for water resources also need to be considered as a part of a more complex knowledge management system. Modelling software must be integrated with other tools that enable gathering a wide range of information and extending the model's predictive capacity (M.B. Abbott et al., 2001). There is a consensus that an effective solution for problems such as emergency management and risk prevention can only be achieved by having various parts of the system dynamically related with ability to exchange information to find the most suitable solutions.

### **3.4 A New Paradigm**

Thus, within the described context there are two streams of development which can determine the future of the computational hydraulics. The first consists of moving towards the implementation of flexible model systems that will guide and help a broader range of users. These models must be able to assess the uncertainty arising in various phases of the modelling process, and feasibility assessment of the predicted scenarios. They will not only utilise the methods of computational hydraulics but will be able to behave as “intelligent agents” capable of reasoning about the problem and its context (J.A. Cunge and M. Erlich, 1999).

The second stream is that the modelling software must be considered as a component of more extended knowledge management systems, where data are collected, transformed, explained and communicated. Thus, other information tools need to be developed which, integrated as parts of the same units, co-operate in solving problems. The co-ordination of different knowledge facilities, each supervising a particular task, permits sharing expertise, working in parallel and profiting from multi-faceted knowledge bases and different sources of information (U. Cortes, 2001).

The need of combining computational hydraulics and advanced information and communication technologies has given rise to a new discipline called hydroinformatics (M.B. Abbott, 1991). As a way forward for managing, planning and protecting hydro-environment, hydroinformatics aims to bring together many activities in the water engineering research and practice in an innovative way with the help of computer-based technologies.

### **3.5 Current Developments of Hydroinformatics**

In order to build effective decision support systems for integrated water resource management, hydroinformatics research has been directed to four main directions:

- data capture, storage, processing and analysis
- real-time diagnosis and prediction
- uncertainty and risk management.
- open modelling

Raw data, which are today abundant, are regarded to be source of “tacit” knowledge which needs to be exploited and investigated (U. Cortes et al., 2001). Large amounts of data can be treated by using several techniques borrowed from various areas of computational intelligence such as machine learning, data mining and knowledge discovery in databases. These data-driven methods (e.g. cluster analysis, evolutionary programming, artificial neural networks) are employed to reduce the complexity of available data, find new correlations and extract new patterns with the possibility to extend the understanding of the physical processes and the laws that govern the natural phenomena (D.P. Solomatine, 2002).

Real time control is also an important issue. It is fundamental for problems related to emergencies (e.g. flood forecasting, pollutant discharges, management of water network systems) where the reaction needs to be rapid. This means that not only data, models, set of hypotheses and management plans need to be available on the site, but they also need to be dynamically linked so that the entire information system can be immediately updated in response to changes. This permits a greater level of efficiency and awareness throughout emergency situations. However, an effective real time control needs also to include the ability to estimate the uncertainty related to the emergency and mainly sound decisions as it evolves (J.A. Cunge and M. Erlich, 1999).

Uncertainty handling represents one of the fundamental challenging in hydroinformatics. It can be in various forms. It can arise from the variability of natural phenomena, or the type of information which needs to be evaluated. This can be in the form of incomplete and imprecise measurements, approximations (i.e model outputs), or even the particular linguistic patterns usually used by experts (J.W. Hall, 2002). For this purpose, the classic probabilistic theory is combined with alternative methods, generally developed in the field of artificial intelligence, which are more appropriate for those types of uncertain information that cannot be represented in a conventional probabilistic formats.

Another new approach to modelling has now been proposed in order to enable water management systems to understand various impacts of a selected management policy. Research is orientated to create an open modelling framework where any set of models can be integrated and interchanged (P.J.A. Gijsbers et al., 2002; D. Harvey et al., 2002). This is expected to result in a new generation of models characterised by flexibility so that any

decision support system can utilise different types of models in specific problems. In future, models, consisting of software components that can be dynamically integrated or substituted when they are no longer needed, will be developed.

### **3.6 Conclusions**

Hydroinformatics is a relatively new discipline that combines computational hydraulics and communication and information technologies to develop powerful computer-based decision making tools for planning, management and protection of the water environment. It aims to improve communication among those groups of professionals who participate in devising water resources management schemes. This is achieved by creating effective computer based working environments where team work is facilitated and enhanced. Hydroinformatics includes methods of data capture, storage, and processing, advanced modelling, optimisation and uncertainty estimation.

## CASE-BASED REASONING

### **4.1 Introduction**

As shown in chapter 3, the integration of computational hydraulics and information technologies has the potential to create advanced computer systems for assisting in planning, managing and monitoring the water environment. Different AI-hybrids that offer the user support at different operational levels of the decision making process have been proposed. In particular, some attempts to interface and integrate artificial intelligence technologies with sophisticated mathematical tools such as simulation programs have been carried out (K.W. Chau and W. Chen, 2001). The benefits of this combined approach are considered to be enormous.

In this research project the Case-Based Reasoning (CBR) methodology is utilized to provide a flexible computer environment for the design of estuarine models with a significant reduction of the effort needed in a multidisciplinary work of this nature. CBR offers a way of building an intelligent modelling environment where available knowledge about estuaries and estuarine numerical models can be stored and reused. In particular, the system aims to offer a non-expert user in modelling the necessary guidance for selecting a model that matches his goal and the nature of the problem to be solved.

The layout of this chapter is as follows: §4.2 presents the artificial intelligence framework employed in our work, the case-based reasoning method; in §4.3 the choice of this technique for supporting estuarine modelling is discussed; §4.4 describes the mode of operating a case-based reasoning system; §4.5, 4.6 and 4.7 illustrate the sub-units that form the case-based reasoning process: case description, case retrieval and case adaptation.

### **4.2 Case-Based Reasoning Definition**

Case-based reasoning (CBR) is a knowledge-based technique. It provides a problem solving framework in which the user benefits from the knowledge of previous experiences. The

solution of a new problem is derived from similar situations encountered in the past. Imitating the problem solving approach usually adopted by humans, CBR systems organise available experiences in terms of knowledge units (cases), which represent a classification of the past events according to a set of common problem descriptors (J. Kolodner, 1993). As a solution for a similar problem is retrieved from the CBR memory, it is then revised and adapted to the current problem. The new case and its solution then become part of the system and available for future applications. At each iteration, the efficiency and competence of a CBR system improve with the number of cases stored in its memory.

### 4.3 Motivation

To explain why this problem solving method is preferable for supporting the design of estuarine models, it is useful to compare it with the expert system methodology, another technique generally employed in the construction of hydro-environmental decision support systems. It is argued that CBR is a more complete and flexible knowledge based technique because of its versatility. Other artificial intelligence technologies, included rules, can also be implemented in a CBR environment to assist and carry out CBR tasks.

#### 4.3.1 *Case-Based Reasoning or Expert System?*

The use of knowledge-based technologies to provide assistance in model design has a twofold scope (A. Häuslein and B. Page, 1991):

1. to provide the modeller with additional knowledge and
2. to process the available information adequately and cost-effectively.

The first scope can be adequately realised by employing expert system technology. In expert system applications the knowledge, elicited from experts, is encoded in rules, consistent pieces of information organised according to a hierarchical logic (K.W. Chau and W.W. Yang, 1994). On the other hand, the use of rule-based reasoning is not sufficient when flexibility and adaptability are important priorities. Despite the number of rules that can be implemented, the system's answer is limited to the predetermined scenarios included into the system. The addition of any rule may require careful revision of the entire system's logic (A. Holt and G.L. Benwell, 1999). Any change to the expert system structure results in a number of knowledge refinements and distillations, often due to the contentious nature of a subject and non-agreement among experts on a common set of final rules. During this process some rules are changed, others are removed and new ones are proposed. Therefore, it is understandable that



the expert system approach is inadequate and not cost-effective when the complexity of the problem domain increases and involves different areas of expertise. CBR overcomes some of the problems encountered in rule-based systems, such as knowledge analysis and organization, system maintenance, and the ability for a system to respond in a real and practical context (I. Watson and F. Marir, 1994).

#### *4.3.2 The Advantages of Using CBR*

The implementation of a CBR system consists of identifying an essential indexing scheme by which the cases are described. When what is to be stored of an experienced event is established, the knowledge extraction consists of describing the case features of past problems according to the pre-defined scheme. This permits the description of cases to be partially filled. Therefore, by assuming the past as precedence, CBR is able to provide a plausible explanation for the solution of a new situation even when the problem is not fully understood.

In the CBR process the reasoning and learning phases are completely integrated. When a problem has been solved, the new information is incorporated in the system's memory with an automatic and incremental improvement of its knowledge content. By accumulating new experience into the memory, the size of the case-base increases and the CBR system becomes more useful.

CBR can also be used in combination with other technologies such as expert system, genetic algorithms, neural networks and statistics to exploit the full potential of the CBR methodology. For instance, these techniques have been successfully applied to establish the similarity of a new situation with past experience (F. Azuaje et al., 2000) and to adapt the retrieved solution to the need of the current problem (D.P. Finn and P. Cunningham, 1998).

#### *4.3.3 CBR for Estuarine Modelling*

The logic of the CBR solving procedure is particularly appropriate to a complicated discipline such as estuarine modelling. This is mainly due to the variety of expertises required. The study of any estuarine phenomenon via numerical modelling relies on an example-by-example based knowledge. Purely theoretical approach in estuarine modelling is not possible and the effort should often be supported by practical experience.

In order to provide "intelligent support" for model design, the complexity of estuaries and the actual interaction of many problem factors must be estimated through the codification of previous studies and correct assessment of numerous assumptions for different cases and problems. In general, the modelling process starts from determining the possible correlation between the physical features of an estuary and the physical phenomena to be simulated. Based on a set of attributes representing the estuary and models' characteristics, the modeller formulates the necessary hypotheses, which are used as bases of the simulation strategy. Therefore, the combination of CBR technology and simulation software programs can result in a very powerful system that will reduce the time required for the necessary field measurements and estimating studies.

#### **4.4 The CBR process**

A case is defined as a "problem situation" (A. Aamodt and E. Plaza, 1994). Therefore, cases represent unique knowledge related to specific situations. A case may be previously experienced and thus stored in the CBR system's library to be used for solving future problems; or it may be new and not yet classified and needs to be included in the CBR system. It can be kept as a concrete experience or clustered with others with which it shares similar features. Cases can also be represented as single knowledge units or the information incorporated in each of them can be split and distributed in the different subunits of the system's knowledge structure.

A case can be divided into three major parts (J. Kolodner, 1993):

- the state of the problem
- the proposed solution
- the outcomes from the application of the solution.

A past problem and its solution can provide useful information to solve new cases. In addition, the outcomes, if included in the case representation, can be used to evaluate the performance of the proposed solution when applied to the current situation.

The CBR process is cyclical and includes the following four steps (B.C. Jeng and T.P. Liang, 1995):

- definition of the problem state and the domain configuration
- selection and retrieval of past cases similar to a new considered event
- revision of the proposed solution if necessary
- storage of the new solution as a part of a new case.

In order to create a system that is capable of searching for similar cases, it is necessary to identify the key features that provide a standard scheme for storing all possible cases in the system memory. This step of CBR is called case description. The chosen case indices are then used to expedite the search process during the phase called case retrieval for calculating case similarities and finding the best match. The search engine is an intelligent component, which allows complete or partial matches according to established prescriptions. However, because every case is likely to be unique, it is necessary to adjust the old solution retrieved by the system in order to apply it to the new situation. The CBR term for employing the retrieved experience is case adaptation. At the end of the problem solving process the current problem and the adopted solution are stored as a new case.

#### 4.5 Case Description

A very important component of the CBR process is case description. Cases are described, then retained in the system's library, and recalled when a new problem needs to be solved. The description of a case consists of finding a suitable and manageable structure to represent the case content. The memory of a CBR system must be organised according to an indexing scheme that permits the retrieval process to be effective and time efficient. A case must be defined in terms of those attributes that are relevant to the domain. Indices should be chosen by following the concepts of functionality and availability of the information to use for representing a case (J. Kolodner, 1993).

Therefore, by using an appropriate domain-specific vocabulary, the indices need to be:

- inherent to the specific goal that needs to be achieved by the reasoner
- abstract enough to make the case generally applicable
- concrete to be recognisable and directly referable to a practical context
- sufficiently predictive to anticipate the solution and outcomes of the case.

For instance, as seen in chapter 2, in case of estuarine modelling the problem representation can count on several definitions and classifications, especially for describing the physical behaviour of estuaries. However, the information is often duplicated and the possible relationships among many of the estuarine features are not clearly expressed. Furthermore, some of this information can be obtained only after specific and often expensive investigations which include field measurements and laboratory experiments. The indexing scheme must be derived from the available sets of definitions and classifications to give adequate description of an estuarine problem that needs to be modelled. Indices can be expressed numerically or qualitatively. Qualitative features can be either discrete concepts expressed as class hierarchies (object symbols) or adjectives representing qualitative values (P.W.H. Chung and R. Inder, 1992).

Based on the set of properties (specificity, abstractness, concreteness, predictability) previously mentioned, the indices should be drawn from the case features and customized in a way that is understandable and applicable. The level of detail in the case description should only include those features that enable the reasoning process and the access to the cases whenever it is necessary. Possible relationships between indices should also be established. Two features may be related (e.g. the average width and depth of an estuary) and their composition may offer more precision and accuracy in the description of a case and in the retrieval process than the single indices (J. Kolodner, 1993). The capturing of a relationship between indices requires the knowledge to link a descriptor to another in the form of functions.

Indexing can be done automatically or manually (I. Watson and F. Marir, 1994). However, despite the possible use of automated indexing programs, it is believed (J. Kolodner, 1993) that for practical CBR applications it is more useful to design the description scheme manually.

#### **4.6 Case Retrieval**

Case retrieval is the step of CBR in which the system identifies and returns the cases stored in the library that can help to achieve the goal or to perform the task associated with the new problem situation. Since it is generally not possible to have a perfect match between the new case and the old ones, the retrieval process is expected to find in the library cases that are the most similar ones. Therefore, it is more appropriate to define the retrieval task as a fuzzy or similarity matching process.

The retrieval task consists of comparing each of the cases contained in the CBR library with the new situation along the indices provided in the case description and calculating the corresponding degree of match. With respect to the degrees of match for all the specified dimensions a set of past cases is selected (*matching*). At this point a more comprehensive evaluation of the similarity with the new case is performed by taking into account the importance of each descriptor. Based on these considerations, the selected cases are ranked in order (*ranking*).

To determine the similarity between cases it is essential to take into account the context of the match. Among the indices available in the case description only a set of dimensions must be chosen that bear upon the problem proposed in the input case and the solution to be found. Degrees of relevance are also given to the selected indices, as the importance of each dimension on the similarity assessment is different with respect to the type of investigation undertaken and the purpose for which the retrieved cases will be used.

#### *4.6.1 Matching and Ranking Procedures*

There are three common methods for carrying out the matching and ranking procedures: the heuristic method, the nearest neighbour matching and the mixed approach that combines the first two (J. Kolodner, 1993).

The heuristic method uses rules to search the case-memory. In a heuristic scheme each difference between an input and a stored case is associated with one or more rules (evidence rules). The old cases that pass the test of the evidence rules are returned to the user. The heuristic method helps to work out the degree of similarity when there is not a perfect correspondence between the features used for representing the new case, and those used for describing the stored cases. For example, the heuristic method is used when two cases need to be compared but they are included in different contexts or the level of representation is more abstract in one case or more detailed in the other. The heuristic rules are also very important when the value of a feature is missing either in the stored case or in the new one and the direct comparison along that index is not possible. The heuristic approach is also used to express preference. Some cases may be preferred over others if certain features along which the degree of similarity is calculated, assume particular values (*salient-feature preference*) or if they address

in a particularly successful way the scope of the reasoning goal that guides the retrieval process (goal-directed preference).

In the nearest neighbour matching for each chosen dimension in the input case the corresponding feature in each past case is found with respect to a numeric evaluation function. The two values are compared to each other and the degree of similarity is calculated. The resulting value is then multiplied by a coefficient representing the grade of relevance of that feature to the new problem. When the weighted similarity for each dimension is determined, the scores are summed to provide an overall similarity measure for the past case and the new one. The case with highest score is retrieved as the best match.

The aggregate match score of a past case with respect to the new problem situation is expressed as:

$$\text{match aggregate score} = \frac{\sum_{K=1}^N W_K (\text{SIM}(F_K^I, F_K^R))}{\sum_{K=1}^N W_K} \quad (4.1)$$

where  $W_K$  represents the importance of the index  $K$  and  $\text{SIM}$  is the degree of similarity between the values of  $K$  for the input case ( $F_K^I$ ) and the retrieved case ( $F_K^R$ ), respectively.

The matching between cases can also be carried out through a combined approach that mixes numeric and heuristic criteria. For instance, cases can be selected on the basis of exclusion (heuristic criterion), before applying the nearest neighbour matching scheme. When the exclusion is adopted, cases are not retrieved if there are unacceptable mismatches between cases along some important case features. Furthermore, it also mixes the matching procedure in which the nearest neighbour matching employs different sets of importance values for different reasoning goals. Therefore, the numeric matching criterion is more sensitive to not only the problem context but also the specific type of problem represented in the new case.

The nearest neighbour matching is a very simple and straightforward procedure. However, it is not characterised by the articulated reasoning process that supports the heuristic approach. It also true that the heuristic method does not provide a definite procedure to establish the best

match among the retrieved case. It is believed (J. Kolodner, 1993; L.D. Xu, 1994) that for the implementation of a computer CBR system a mixed approach is preferred.

#### 4.6.2 The Degree of Similarity – Fuzzy Matching

Once it is established which feature corresponds to each other and along which indices the cases should be compared then the degree of similarity should be calculated. However, since different kinds of data (numeric, classes of objects and qualitative attributes) can commonly be found in case descriptions, it is essential to have a common framework for computing the degree of similarity. Fuzzy matching (B.C. Jeng and T.P. Liang, 1995) based on fuzzy set theory and fuzzy logic, permits the calculation for each index, relevant to the retrieval process, of the degree of similarity  $SIM$  between the corresponding values  $F^I$  for the input case and  $F^R$  for an old case. The use of fuzzy matching gives the possibility to manipulate all types of data (numeric parameters and qualitative attributes) and computing similarity scores on a normalized scale between 0 (minimum similarity) and 1 (maximum similarity).

##### 4.6.2.1 Numerical Values

In case of numerical data the degree of similarity  $SIM(F^I, F^R)$  is calculated using evaluation functions based on the possible expression used for the characteristic functions in fuzzy set theory (P.W.H. Chung and R. Inder, 1992).

For instance, a linear approximation could be used for computing  $SIM$ :

$$SIM(F^I, F^R) = \max\left(0, 1 - \frac{|F^I - F^R|}{K}\right) \quad (4.2)$$

A more complex function could be applied to avoid the sharp cut-off point and to concentrate the emphasis on the near misses.

$$SIM(F^I, F^R) = \left[1 + \left(\frac{|F^I - F^R|}{\sqrt{F^R}}\right)^2\right]^{-1} \quad (4.3)$$

Both functions are characterized by the fact that the value of  $SIM$  decreases as the distance between  $F^I$  and  $F^R$  becomes greater.

There is not criterion that can help to choose between these two functions or any other proposed. Only an expert can decide which function is most appropriate for calculating the degree of similarity along each feature and, perhaps, modify existing functions or propose new ones.

#### 4.6.2.2 *Symbolic Values*

There are two types of symbols: qualitative descriptors, which are adjectives describing qualitative values (e.g.: low, average, high) and object symbols, which are labels for classes of objects (e.g.: microtidal, mesotidal, macrotidal).

In fuzzy matching the degree of similarity between two qualitative values is measured by their positions in a discrete scale. For example, the scale may be: very low, low, moderate, high, very high. The distance between adjacent categories is 1, e.g. low and very low and the total distance between very low and very high is 4.

The similarity between two adjectives defining the values of a feature for the input case and the old case, respectively, can then be expressed as:

$$SIM(F^I, F^R) = 1 - \frac{\text{distance}(F^I, F^R)}{\text{Total distance}} \quad (4.4)$$

It is more difficult to establish the appropriate similarity function for object symbols. It is up to the expert, or in general the system's designer, to define the possible symbols to associate with a specific index and organize them in a classification scheme. The similarity function must be, then, defined by giving each pair of symbols a weighted distance that takes into account the possible relationship between the two classes, and their characterisation. General methods based on the distance between two nodes exist but they are not always applicable.

### 4.7 Case Adaptation

Once the best match has been retrieved from the system's library, its solution needs to be made applicable to the new problem. The CBR term for the flexible use of the retrieved experience is case adaptation. The case adaptation process is responsible for modifying and amending the old case accordingly to the new situation requirements. The adaptation may



involve the whole case structure or just parts of it without the need to change the entire scheme. Some parts of the solution may be substituted, others deleted or transformed.

Substitution consists of replacing some components of the old case with others that better suit the new problem. These components may be objects or parameter values that have to be modified. By comparing the contexts of the new and old cases and identifying the differences between them, it is possible to establish which components of the old solution need to be changed and where to search for the necessary information. For instance, it may be necessary to substitute the objects of the old solution with others that have the same functional role. It may also be required to modify the values of the solution parameters to guarantee correct and realistic output for the new case. However, sometimes the information is not directly available from the new problem context, and the necessary new data need to be found by using specialised adjustment heuristics, as in the case of parameters, or querying the system's memory or other libraries that may contain useful alternatives.

Transformation methods use heuristics to replace, delete or add components to an old case. Some items may have to be transformed as they violate the constraints and limits imposed by the new case and no alternatives are available. The heuristic rules resolve the possible conflicts guided by the weighted importance of the different components in the old case. Rules can be of general purpose (commonsense transformation) or based on the causal explanation of the role played by each component in the old solution scheme (model-guided repair).

If the adaptation heuristics is domain-specific then the adaptation technique is said to be special-purpose. The adaptation is addressed by the particular information gathered from the old case and its possible performance in case of its application to the new case. This technique is important when particular structure modifications, such as the insertion of a new element in the scheme, are required. The special-purpose adaptation is applied when the old solution, directly tested on the new case, has failed and consequently needs to be repaired. Based on the type of failure registered specific rules are activated to appropriately change the old solution and eliminate the cause of failure.

Because many parts of a case might need to be changed, the case adaptation would result in a combination of different techniques for substitution and transformation. Therefore, for large

and complex cases, the problem must be decomposed and each part treated and adapted separately. However, this does not guarantee that once put together again, the modified parts will work properly as there may be unresolved conflicts between the parts.

*After the selected adaptation strategy has been applied and the old case has been appropriately updated to meet the new problem requirements, the modified retrieved case is ready to be used.*

At the end of the problem solving process the knowledge contained in the library is refined and incremented. The CBR system becomes more useful (case retention).

#### **4.8 Conclusions**

This chapter discusses the benefits of using the Case-Based Reasoning methodology to create decision support systems for complex and dynamic tasks involving multidisciplinary approaches such as the design of numerical models for estuaries. CBR systems utilize the knowledge from solving past problems to solve similar cases. In contrast with the expert system technology, used for similar kind of applications, the CBR technology does not only automate the reasoning and implement domain-specific knowledge, but also permits the system to increase and advance its knowledge base with use.

## GENETIC ALGORITHMS

### **5.1 Introduction**

Case adaptation is a key phase in the CBR process. Once it has been established which parts of the selected case need to be changed, an appropriate adaptation strategy must be identified. Its choice depends on the application domain and, thus, its implementation needs to reflect the specific knowledge involved. The selection of a suitable adaptation method is guided by the kinds of items to be substituted within the old solution and the roles played by them in the problem definition.

For numerical models, case adaptation is essentially based on adjusting parameters where the values in the old solution are changed with respect to differences between specifications of the new and old cases. The input and model parameters are updated to ensure that the model accurately simulates the real behaviour of the new case's estuary. In particular, while the new input data are directly derived from the new case's specification, the values of the model parameters need to be identified and substituted into the model datafile.

In this project genetic algorithm theory is applied to the adaptation of the model parameters. Based on the mechanics of natural selection and genetics, genetic algorithm technology can be used to find the optimal parameter values to suit the particular environmental conditions of the new estuarine domain to be studied. In this chapter the basis of genetic algorithm theory and its application to case adaptation are described.

The chapter is organised as follows: §5.2 discusses the motivation behind the use of the genetic algorithm technique for case adaptation; §5.3 presents previous work in which the genetic algorithms have been used to evaluate parameters, in particular for hydrologic systems; in §5.4 the genetic algorithm methodology is illustrated.

### **5.2 Motivation**

For adapting a numerical model to new situations, it requires:

- 1) discretising the new problem domain into a computational grid

2) updating the input data (e.g. hydrographic data and boundary and initial conditions)

3) modifying the model parameters (e.g. bed friction and dispersion coefficients).

The domain discretisation can be obtained by using appropriate pre-processor software programs that subdivide the domain into a mesh according to the requirements of the selected numerical method. There are many commercially available software (e.g. GEOSTAR<sup>1</sup>) which after the insertion of the geographical boundaries of the domains generate the required grid.

The input data are derived from existing field measurements and navigation charts that provide all the information related to the characteristics of the specific estuary and the particular problem that requires solution (D.M. McDowell and B.A. O'Connor, 1977).

Due to inevitable uncertainties in the measurement of complex hydrodynamical parameters, these are usually found by specific optimisation techniques. Unlike the hydrographic data, the identification of appropriate model parameter values is a matter of continuous debate. This is because such coefficients serve in effect to match the model prediction with the existing system's behaviour (M.B. Beck, 1991). Therefore, the model parameters must be optimised to reproduce the essential features of a watercourse. The selection of their values is affected by physical phenomena characterising the specific water system considered, and scale-effects due to approximations introduced during the model development. Although it is important to take into account the physical meaning of model parameters, their calibration in exact agreement with the nature is not possible (M.B. Abbott et al., 2001). Thus, because their best values cannot be obtained by direct field measurements, special techniques, based on the minimisation of the difference between simulated values and the observed data, need to be employed. This process is commonly called model calibration. A calibrated model is verified in a further step in which the optimised parameters are used without alteration in conjunction with a new set of boundary conditions, and the model and expected results are compared.

In the present CBR system a genetic algorithm is used to identify appropriate values for the Manning's coefficient. As mentioned in chapter 2, this parameter is utilised to represent the bed resistance to the flow of water in the hydraulic equation of motion. For numerical models where the problem domain is discretised into elements (up to several hundreds), the resistance to the flow is expressed by associating to each section a specific value of the Manning's

---

<sup>1</sup> Geostar is a COSMOS product that offers parametric and automatic meshing of complex 2D/3D domains ([www.cosmosm.com](http://www.cosmosm.com)).

coefficient. Thus, calibrating the Manning's coefficient in a numerical hydrodynamic model of an estuary means to find the set of Manning's coefficient values that gives realistic simulations. It must be noted that because of the interdependency of the Manning's coefficient values on each other, just changing a Manning's coefficient value for one section of the domain may result in the alteration of the entire model performance and the quality of the output.

#### *5.2.1 The Traditional Optimisation Methods*

Model calibrations have been traditionally carried out either manually or using numerical optimisation programs. Both manual and computer-based parameter adjustments require an experienced modeller to conduct the task.

The manual parameter assessment essentially relies on the modeller's experience and his/her personal judgment. The modeller must estimate how the error is distributed between the simulated predictions and the corresponding field measurements for different sections of the computational grid. He/she must be able to identify the locations along the estuarine channel that, due to the specific physical features of those sections (e.g. a junction) and the level of accuracy required in the simulation, need to be carefully checked, during the calibration process. It is hence not surprising that in large problem domains calibration based on the trial and error method is usually a tedious process which requires the modeller to have considerable experience. The manual optimisation is very subjective. Different modellers will probably produce different sets of values (R.H. McCuen and W.M. Snyder, 1986).

The computer-based optimisation programs use mathematical procedures that minimise the difference between the simulated results and the observed data. The most common criteria are based on least square procedures which depend on the maximum absolute error or bias (A.M. Wasantha Lal, 1995). The mathematical formulation of such minimising expressions is called objective function. The methodologies used for estimating the parameter values may be different. Mathematical algorithms based on the Gauss-Newton, the minimax or the Kalman filter techniques are commonly used to find optimum parameters (A.M. Lal Wasantha, 1995). Non linear problems are generally optimised by using non-linear programming techniques such as sequential quadratic programming (R. Ramesh et al., 2000).

The scope of these automatic procedures is to speed up the search, with minimum involvement of a modeller, for a set of parameters that bring the numerical results as close as

possible to the experimental data. However, these calibration techniques are of limited applicability because of the

- strong dependency on the initial guess
- practical impossibility of searching the entire parameter space
- tendency to converge toward local optimal solutions.

Furthermore, these calibration procedures are not completely automatic. Their rate of convergence closely depends on a modeller's decision and the selection of the specific optimisation factors. Information such as the initial set of values for the parameters, the step size used in exploring the parameter space and the value of the gradient for non-linear optimisation methods must be all selected with extreme care.

### *5.2.2 The Advantage of Using Genetic Algorithms*

Previously described calibration methods are insufficiently robust and lack efficiency. Many practical parameter spaces, such as the domain of possible sets of Manning's coefficients in hydrodynamic models, are too large to be investigated either manually or even using computer based numerical algorithms (D.E. Goldberg, 1989). These methods, which tend to give local and subjective solutions, have a limited capacity to provide the real best solution for an optimisation problem. In contrast to these optimisation methods genetic algorithms have the characteristics of being simple, quick and highly efficient in locating the global optimum values in the calibration process.

The genetic algorithm technique is a stochastic global search which is based on the analogy with biological evolution.

An initial population of solutions, randomly distributed in the search space, is defined as the starting point of a search. The population evolves in a number of generations by eliminating the unfit individuals and generating new solutions by combining those that have provided a good performance. Occasionally the individuals in the population are randomly modified in an attempt to produce new improvements by chance.

The main advantage of using a genetic algorithm is due to the way it searches the solution space. First, the search is simultaneously conducted for a population of points equally probable, rather than point by point. Second, the evolution of succeeding generations is randomly guided by probabilistic transition rules, instead of using deterministic mathematical

methods that do not allow a free movement within the search space. Third, despite a random search, a genetic algorithm uses the information related to the individuals in the previous generation to obtain new search points with possible improvement of the algorithm's performance (L.D. Davis, 1991). Fourth, genetic algorithms do not need additional information (e.g. step size, a gradient for the step progress and initial guessed values) to work properly. They are blind. Useful information about search space is simply utilised to narrow down the search space. Furthermore, genetic algorithms can also be easily implemented in parallel mode with a great advantage in terms of precision and speed especially when a genetic algorithm is used in conjunction with numerical models characterised by high computational time (B.H.V. Topping et al., 1998).

These characteristics make genetic algorithms more robust and flexible than other search procedures. Many applications, especially in the area of water resource modelling, have shown the usefulness of this approach even in situations involving highly complex multiple parameter domains.

### **5.3 Applications of Genetic Algorithms to Water Resources**

The genetic algorithm approach, developed by Holland and co-workers at the University of Michigan in the seventies (J.H. Holland, 1973), has been successfully used in many engineering applications. However, only recently it has been applied to model hydro-environmental problems, demonstrating that it is superior to the classical optimisation methods (P. Reed et al., 2001).

These applications are mainly focused on the calibration of the parameters related to rainfall-runoff models in forecasting of the river flow (Q.J. Wang, 1991; Q.J. Wang, 1997). The genetic algorithms limit the uncertainty related to these parameters and their mutual relationships. Several groundwater applications are also reported. In particular, genetic algorithms have been implemented for groundwater management problems to reduce the uncertainty related to the hydrodynamics of aquifer systems (D.C. McKinney and M.D. Lin, 1994; K. Lakshmi Prasad and A.K. Rastogi, 2001; M.M. Aral et al., 2001). In these applications a genetic algorithm is coupled with the working equations that, transformed using finite element or finite difference methods, permit the objective function calculation (i.e. the difference between the predicted values of the unknown and the corresponding observed data). Optimisation by genetic algorithms has also been demonstrated to be straightforward and to provide reliable parameter

estimation even when the field measurements are sparse and approximate (K. Lakshmi Prasad and A.K. Rastogi, 2001). It has also been demonstrated that the optimisation by genetic algorithms can be used for a single goal or multiple goals, obtaining more than one compatible solution for a given problem. Multiple conflicting objectives problems have been solved by genetic algorithms, e.g. treatment of contaminant waters (B.J. Ritzel et al., 1994) and groundwater monitoring (V.M. Johnson and L.L. Rogers, 1995).

#### 5.4 The Fundamentals of Genetic Algorithms

Genetic algorithms perform the search process in four stages (K. Shin and I. Han, 1999) (figure 5.1):

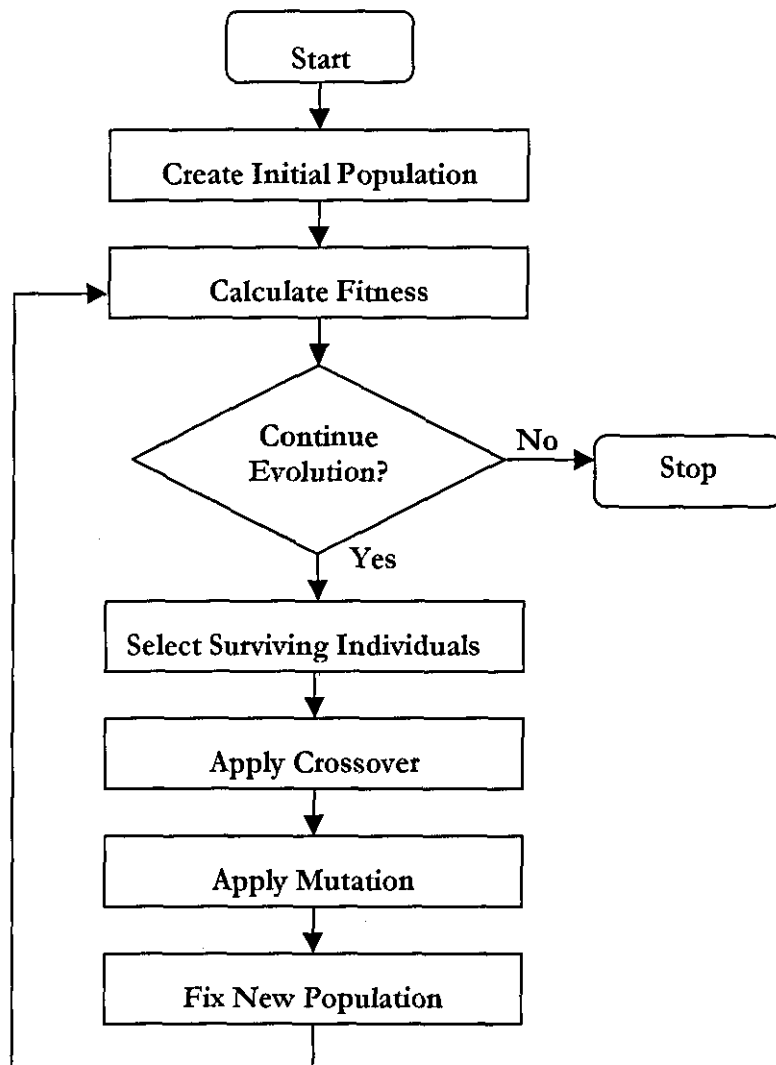
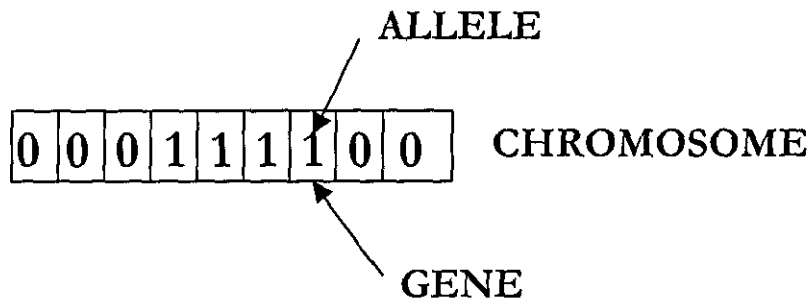


Figure 5.1. Steps of a genetic algorithm



- Initialisation
- Selection
- Crossover
- Mutation.

In the initialisation stage a population of possible solutions is chosen randomly in the search space. The solutions might be related to a single parameter value or a combination of variables that need to be simultaneously optimised. To have a degree of correspondence with biological evolution, the solutions are represented in a predefined mathematical expression, which corresponds in biology to the transcription of biological information in the form of genes. Borrowing the terminology from biology, the representation is called *genotype* and the corresponding parameter variable is called *phenotype*. The genetic structures corresponding to the symbolic representation of the true parameter values are called *chromosomes*. Chromosomes are made up of unites, *genes*. In case of binary encoding, the genes correspond to the bits of the strings. The possible values of the units are called alleles: in the binary representation the only permitted values are “0” and “1” (figure 5.2).



**Figure 5.2.** Chromosome representation using the binary representation

Starting from the initial population, each chromosome is evaluated according to a user-defined objective function (also known as fitness function). The goal of the fitness function is to estimate the performance of the chromosome as a possible solution. Based on their fitness, chromosomes are selected through a stochastic process. The chromosomes with high fitness are chosen to propagate and be replicated in the next generation.

Imitating the classical Darwinian evolution theory, genetic operators such as crossover and mutation are used to form new chromosomes. Through a probabilistic mechanism, individuals

in the population are chosen based on their performance to create new offspring individuals. Some chromosomes are created by swapping corresponding parts of their symbolic representation (crossover). The new solutions are supposed to inherit the good features of their parents and extend the search in areas of the solution space not yet explored. After the crossover operation the mutation operation is often performed. Inspired by what happens in nature when the genetic material is transferred between generations, the mutation operator slightly modifies the genetic structure at the level of the symbolic representation. A parent chromosome is randomly chosen, mutated and replaced by the new version. The genetic algorithm guarantees that all chromosomes have a chance to be picked for mutation.

The selection, crossover and mutation steps are repeated for each generation until the termination criterion is satisfied. Deciding on the termination criterion of a GA run is non-trivial. This is because it is rather difficult to decide on the required performance. Different criteria can be adopted for this purpose. A time-control criterion can be chosen by imposing a limit on the run time. The genetic algorithm can also end when a specified number of generations is reached or a threshold distance between the simulated results and measured values is achieved.

#### *5.4.1 Population Representation and Initialisation*

The genetic algorithm population is represented by a number of individuals, which can be anything between 30 and 100. Smaller sizes (i.e. 10 individuals) can be employed with limited selection and reproduction capacity if the execution time is a practical constraint to the success of the search. The binary representation is the most commonly used genotype, encoding the individuals in bit strings (figure 5.2). In the standard binary coding, each candidate solution is a string of binary digits of chosen length. In particular, for solutions that represent a combination of parameters, each variable is transformed into a binary string and these strings are then concatenated to form a single string. The length of each string depends on the required precision. For instance, if the domain of the parameter is 5 digits and the required precision is 2 places after the decimal point, the range of the domain should be divided into  $5 \times 10^2$  equal size ranges. This means that each chromosome is 9 bits long:

$$256 = 2^8 < 5 \times 10^2 < 2^9 = 512 \quad (5.1)$$

Alternative encoding procedures are also employed: for instance, integer and real-valued representations. Integer encoding is claimed to provide a convenient and natural way of

mapping the phenotype into the genotype, especially for the subset selection problems (P.J. Fleming and A.M.S. Zalzal, 1997). In this case the chromosome is represented by a number of integers each of which refers to a specific variable. In real-valued encoding the genotype and phenotype are the same and less memory are required. Furthermore, the precision of this representation is claimed to be superior to that of binary code, which relies on the chosen length of the chromosome (P.J. Fleming and A.M.S. Zalzal, 1997).

After having decided the genotype form, the population of the genetic algorithm must be initialised. The initialisation of the population is carried out by randomly creating each chromosome. For instance, in the case of binary representation with a population of  $N$  individuals, each having  $L$  bits,  $N \times L$  bits each with a value 0 or 1, are randomly computed and distributed in groups of  $L$  bits to each chromosome.

#### 5.4.2 Selection and Fitness Function

The fitness function is used to estimate the performance of the chromosomes. Based on this calculation the chromosomes are then selected to be included in the next generation. The definition of the fitness function must be based on a correct understanding of the type of field measurements available. A fitness function is dependent on the problem and it must satisfy a set of chosen constraints associated with the optimisation (i.e. economic, technical conditions). More generally a fitness function is either a minimisation or maximisation measurement (i.e. maximum water demand, minimum cost, minimum discrepancy between observed and simulated points).

Each chromosome is decoded into the corresponding phenotypical value to calculate the fitness function. For instance, in order to evaluate the performance of a model parameter's value generated by the genetic algorithm, the fitness function can be calculated by measuring the deviation between the results obtained from the model run with that specific parameter's value and the corresponding observed data.

One obvious possibility is to express this difference as (V. Babovic et al., 1994):

$$\varphi(x_i) = \left( \sum_{j=1}^N |F_j - P_j|^2 \right)^{1/2} \quad i = 1, M \quad (5.2)$$

denoting by  $\varphi(x_i)$ ,  $N$ ,  $F_j$  and  $P_j$  the fitness value of the chromosome  $i$ , the number of sample points, the simulated and observed value of a chosen model variable (i.e. velocity, water surface

elevation) for the sample point  $j$ , respectively. In this expression, weights can also be included to give emphasis to the fact that the deviation is not well distributed among the  $N$  samples.

$$\varphi(x_i) = \left( \sum_{j=1}^N w_j |F_j - P_j|^2 \right)^{1/2} \quad i = 1, M \quad (5.3)$$

As it is evident from the above mathematical expressions, a fitness function is defined to be non-negative. There is usually the preference to express the fitness on a normalised scale and to maximise the fitness (V. Babovic and M.B. Abbott, 1997):

$$\varphi^*(x_i) = 1 - \frac{\varphi(x_i)}{\text{Max}_{i=1, M}[\varphi(x_i)]} \quad (5.4)$$

$$\varphi^*(x_i) = \frac{1}{\varphi(x_i)} \quad (5.5)$$

According to their fitness the chromosomes are chosen with respect to a predefined selection mechanism to form a new population. Many selection techniques are based on the roulette wheel mechanism that consists of selecting the chromosomes in a probabilistic way based on the measure of their fitness (P.J. Fleming and A.M.S. Zalzal, 1997). The sum of the fitness values over all the individuals in the current population is calculated.

$$F_{TOT} = \sum_{i=1}^M \varphi(x_i) \quad (5.6)$$

A random number  $n$  in the range  $[0, 1]$  is then generated.

For each chromosome the selection probability and cumulative probability are calculated respectively as:

$$p_i = \frac{\varphi(x_i)}{F_{TOT}} \quad (5.7)$$

$$q_k = \sum_{i=1}^k p_i \quad \text{with } k = 1, M \quad (5.8)$$

If  $n \leq q_1$ , then the first chromosome is selected, otherwise the chromosome  $i$  in the population for which  $q_{i-1} < n \leq q_i$ . The mechanism is operated  $M$  times, i.e. the total number of chromosomes (Tables 5.1 and 5.2).

**Table 5.1.** Roulette wheel mechanism

chromosome	$\varphi(x_i)$	$P_i$	$q_i$
1	10	0.106383	0.106383
2	18	0.191489	0.297872
3	4	0.042553	0.340425
4	11	0.117021	0.457446
5	17	0.180851	0.638297
6	5	0.053192	0.691489
7	8	0.085106	0.776595
8	13	0.138298	0.914893
9	2	0.021277	0.936170
10	6	0.063830	1

**Table 5.2.** Roulette wheel selection

n	Chromosome chosen
0.601431	5
0.322071	3
0.876501	8
0.184235	2
0.661015	6
0.285423	2
0.965284	10
0.543265	5
0.087523	1
0.485613	5

To make sure that the fittest candidates are selected, this classical sampling technique is coupled with a number of possible extensions such as the elitist selection, expected-value and crowding approaches (D.E. Goldberg, 1989). The elitist approach chooses one or more fittest individuals and directly include them in the next generation. In the expected-value approach, the probability of each chromosome of being selected is proportional to its fitness. In the crowding approach, old individuals, albeit still considered fit to be selected, are substituted with offspring chromosomes that most closely resembled the older elements in the genotypical representation. This modification ensures a certain grade of diversity to the population through the generations, avoiding the dominance of some chromosomes over other individuals. Noticed that while the first two approaches improve the genetic algorithm's performance for single-objective functions, the last one is preferably applied to multi-objective optimisations (D.E. Goldberg, 1989).

### 5.4.3 Crossover and Mutation

The crossover operator combines segments from different individuals. Crossover produces new individuals which are made up of parts from both parents' genetic material. The simple way to determine the crossover point is to choose a random cut-point. An integer  $k$  is randomly chosen from the range  $[1, L-1]$ , where  $L$  is the chromosome length. The offspring chromosomes are generated by merging the first  $k$  bits segment of one parent with the  $k+1$  to  $L$  bits from another parent (figure 5.3a). This is the simplest crossover scheme called single-point crossover.

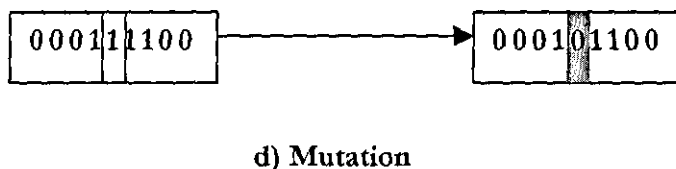
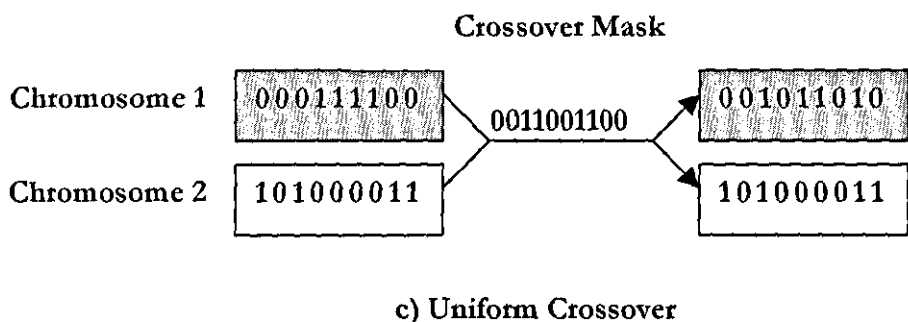
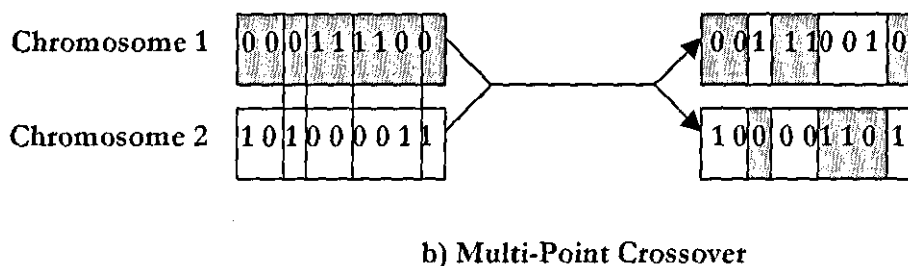
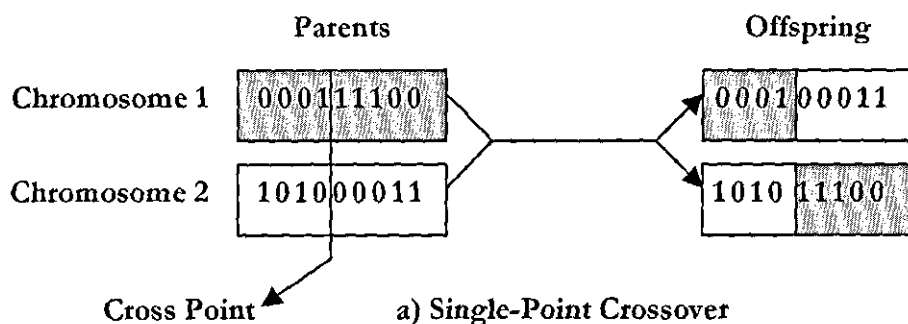
Other possible schemes are multi-point crossover and uniform crossover. For multi-point crossover,  $m$  crossover positions  $k_i = \{1, 2, \dots, L-1\}$  are chosen at random with no duplicates and sorted in ascending order. The bits included between two adjacent crossover points are exchanged between the parents to generate two new chromosomes (figure 5.3b). The multi-point crossover responds to the fact that in some cases the good performance of a chromosome is not due to the information contained in adjacent substrings, but may be distributed in different parts of the chromosome.

Uniform crossover generalises the multi-point crossover operation by making every gene a potential crossover point. It combines bits sampled uniformly from the two parents as illustrated in figure 5.3c. A crossover mask, of the same length as a chromosome, is created at random. The offspring chromosomes are produced by taking from one parent the bits that corresponds to the digits of the mask equal to 1, and from the other the bits that correspond to the digits of the mask equal to 0.

The crossover operator is only applied to a group of chromosomes. It is applied with probability  $P_c$  based on which parents are randomly selected from the population for breeding.

After the application of crossover, the mutation of chromosome is performed. The mutation operation introduces new qualities in the population and enables a genetic algorithm to consider unexplored areas of the search space. With a probability  $P_m$ , some individuals are selected from the population and small random changes are made to the chromosomes. For instance, with binary encoding, mutation causes a randomly chosen bit of a chromosome to change its value either from 1 to 0 or from 0 to 1 (figure 5.3d). With non-binary encoding, mutation may slightly vary the value of a gene or substitute it with a new value randomly picked from an allowed range. The mutation probability  $P_m$  is generally between 0.001 and 0.1

in the case of binary encoding. This probability might be higher in the case of real-coded chromosomes, increasing the possibility of exploring the search space more effectively without affecting the level of convergence (Z. Michalewicz, 1994).



**Figure 5.3.** Genetic operators: a) Single-Point Crossover, b) Multi-Point Crossover, c) Uniform Crossover and d) Mutation

#### *5.4.4 The Use of Specific Knowledge*

One of the most important characteristics of genetic algorithms is their ability to perform a blind search. The information about the specific problem is generally limited to the construction of fitness function. However, when additional sources of problem-specific knowledge are available, it is appropriate to incorporate them to form a hybridised genetic algorithm structure. The combination of problem-specific information and genetic algorithms increases the accuracy of the search and the capacity to identify the best solution (J.J. Grefenstette, 1987). Problem-specific knowledge can be implemented in any phase of the genetic algorithm scheme. For instance, heuristic rules can be utilised to initialise the population. Several populations can be generated and the best performing chromosomes can be chosen among them to represent the actual initial population (D. E. Goldberg, 1989). The initial population can also be seeded with particular chromosomes that speed up the search toward the best solution, with particular care in avoiding any premature convergence (S.J. Louis and J. Johnson, 1997). These specific chromosomes from previous searches, which yield useful information about the search space can be injected into the initial population reducing the time taken to find the appropriate solution. For this purpose genetic algorithms may be combined with case-based reasoning principles, based on the idea that similar problems have similar solutions (C.L. Ramsey and J.J. Grefenstette, 1993)

For many specific problems it may also be feasible to define the genetic operators (i.e. crossover and mutation) based on the characteristics of the considered search space. The encoding scheme for hybridised genetic algorithms generally uses real or integer representation. The familiar genetic operators must be adapted to the problem domain. The crossover and mutation operators developed employing problem-specific knowledge can be more useful than the original versions, which can create non realistic offspring chromosomes that are meaningless in terms of the problem domain (D.E. Goldberg, 1989).

#### **5.5 Conclusions**

Case adaptation is the most difficult phase in the case-based reasoning methodology. An adaptation strategy must be chosen with respect to the specific problem domain, the knowledge represented in the system and the reasoning task to be supported. In CBEM case adaptation is based on the genetic algorithm approach, since the adaptation is focused on model parameter adjustment. A genetic algorithm is employed to find an appropriate set of



Manning's coefficients for a new case. Superior to traditional optimisation techniques, the genetic algorithm method permits the finding of the global optimum quickly and efficiently. It also allows the incorporation of domain-specific knowledge to enhance its performance.

## CASE-BASED REASONING SYSTEM FOR ESTUARINE MODELLING

### 6.1 Introduction

This chapter outlines CBEM, a Case-Based Reasoning System for Estuarine Modelling (CBEM). CBEM is an intelligent computational environment for supporting estuarine model design. By organising the available knowledge on existing numerical models for estuaries and their applications the user is provided with necessary guidance to solve numerically hydro-environmental problems for estuaries. The system is developed using the case-based reasoning methodology, employed in conjunction with numerical modelling theory and genetic algorithms.

The system has three main software components: a case-based module where the case description and the case retrieval take place, a genetic algorithm module, responsible for the case adaptation and a library of numerical simulation models, which contains the computer codes of the numerical models described in the case-based module and called by the genetic algorithm during the adaptation phase. Based on the features of the estuary under investigation and the physical phenomenon to be modelled, the case-based module retrieves a suitable solution among the past cases contained in the system's memory. The selected model is then returned from the model library and specifically adapted by the genetic algorithm unit, which estimates a valid set of model parameters to suit the new estuarine physical conditions.

The layout of this chapter is as follows: in §6.2 the general architecture is illustrated; §6.3, 6.4 and 6.5 outline the system's main components: the case-based module, the genetic algorithm module and the model library.

### 6.2 General Architecture

The Case-Based Reasoning system for Estuarine Modelling (CBEM) operates through the co-operative action of its three main components: a case-based module, a genetic algorithm module and a library of numerical models (figure 6.1). These modules are activated to perform

specific tasks of the case-based problem solving process. The case-based (CB) module allows the user to describe new and past cases (case description). It is also responsible for the retrieval process (case retrieval).

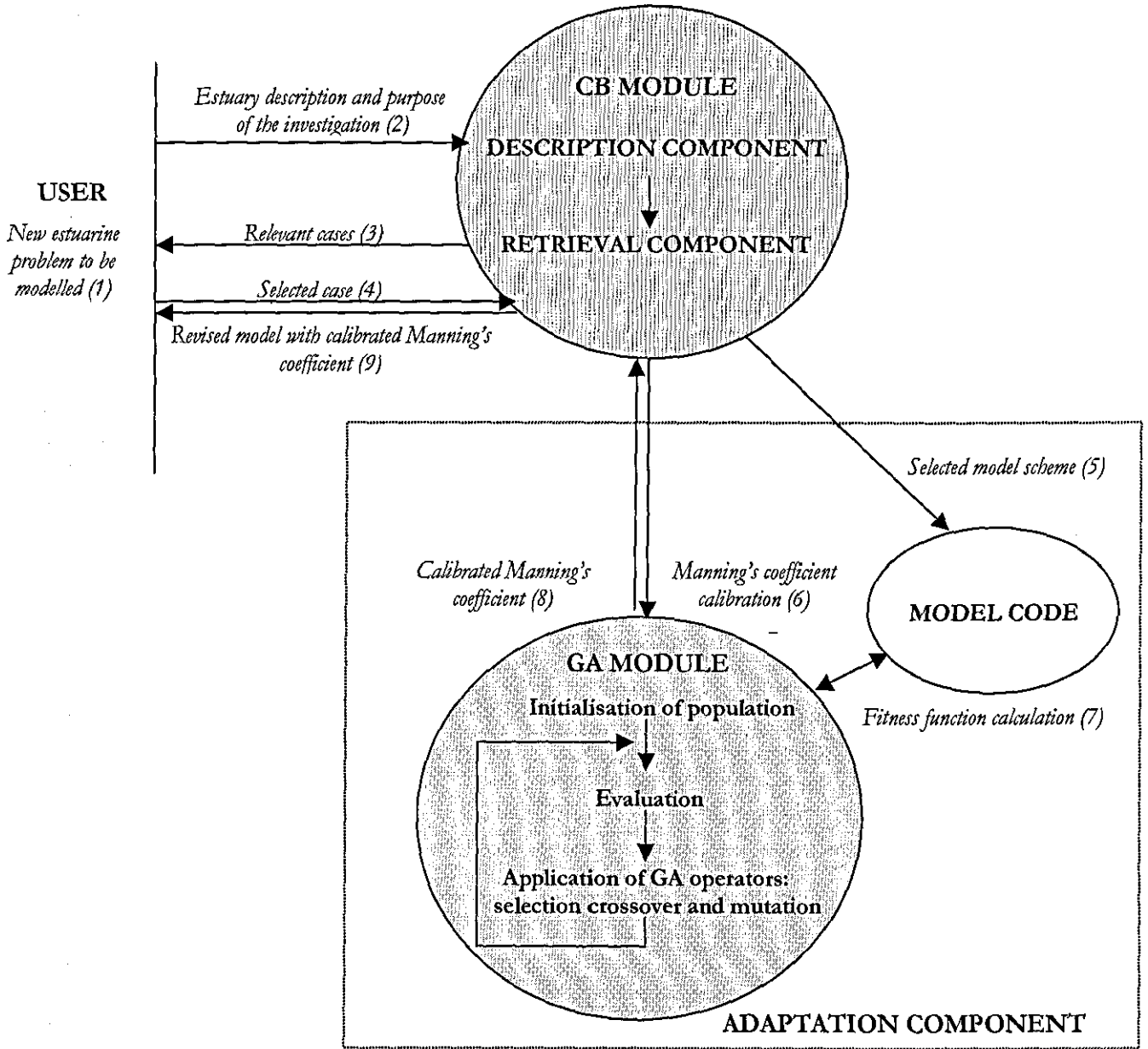


Figure 6.1. Structure of the CBEM system

**NEW PROBLEM:** i.e. salinity distribution to be investigated for the estuary X

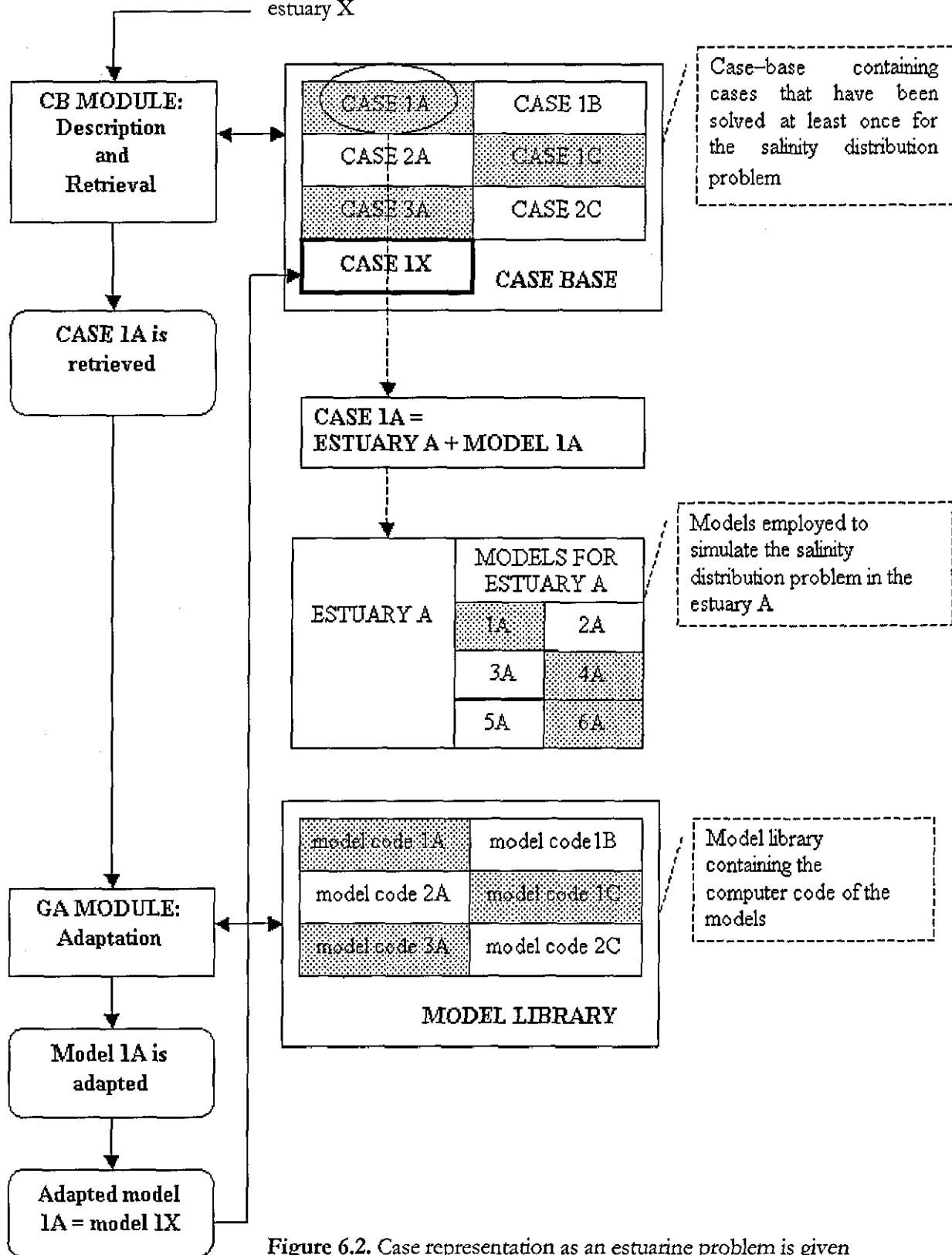


Figure 6.2. Case representation as an estuarine problem is given

In CBEM a case is divided into two parts: the estuary, which is the object of the investigation, and the related models, each of which is employed to simulate a specific physical phenomenon for that estuary. The estuary description contains indices representing the features of the estuary domain, while the model description includes information about the model characteristics and the estuarine problem simulated. This distinction is due to several practical reasons. The same estuary may have been studied and modelled for different purposes, or a specific estuarine process may have been repeatedly simulated for the same estuary but using different model strategies to satisfy different quality requirements of the results and the specific purpose of the simulation (figure 6.2). The separated descriptions for estuaries and models permit the identification of the different aspects that characterise a model strategy, which is constructed taking into account the assumptions on the physical and hydrographic behaviour of an estuary, and specific conditions on the problem definition (e.g. cost-effectiveness and accuracy). This distinction also makes the case representation more accessible and readable to the user, who has to supply the necessary information. Furthermore, it facilitates and speeds up the retrieval process. Dividing the cases into parts enables the search mechanism to identify easily similarity between parts of the cases and make rapid conclusions about the cases themselves (J. Kolodner, 1993). The retrieval mechanism computes the similarity with respect to both the estuary and model descriptions.

As the user decides to investigate a problem using CBEM, he/she enters the feature values of the estuary to be modelled in the estuary description scheme. He/she then defines the type of problem and the purpose of the investigation. At this point the retrieval process begins. Initially the search engine selects from the system's case-base (figure 6.2) only those cases for which the current problem has been previously modelled. A similarity rating for each of the past cases is established based on the values of the indices contained in the estuary's description.

In addition, the past cases are ranked with respect to their performance expressed in terms of accuracy, simulation time consumed and purpose for which they were employed. At the end of the retrieval process, the user is presented with a list of past cases graded with respect to the estuary description as well as the problem definition. The user is responsible for the final selection based on personal judgment.

Once the model scheme is chosen, it is selected from the model library, which contains the computer codes of the models described in the case-base unit, and it is given to the genetic algorithm module (GA) to be adapted, i.e. the case adaptation phase of the case-based reasoning process. The genetic algorithm (GA) module is activated to optimise the parameters of the selected model to suit the new application. The required parameter optimisation is carried out by combining the classical evolutionary theory with problem-specific information. The GA module incorporates knowledge from the practice of estuarine model calibration by using modified genetic operators. Furthermore, the present scheme benefits from the co-operation with the CB module by including in the initial population the parameter values from the most similar cases. As it will be illustrated later, the use of knowledge augmented operators and case-based initialisation improves the search performance. It finds a better set of parameters for the new problem and requires less time than the classical genetic approach. The CBEM procedure terminates and returns to the user a model scheme retrieved among the past cases as the best match, and a new set of parameters that provides satisfactory performance. The model can now be applied to simulate different scenarios.

The description and the retrieval phases are implemented using ACCESS (R. Smith and D. Sussman, 1997), while the genetic algorithm in the GA module is programmed in FORTRAN. ACCESS provides a form-like interface with each case described by a set of fields and their values. A database is built where cases are stored. Different forms are provided for viewing and searching for the estuary and model descriptions, and the results from the retrieval process.

The use of FORTRAN for the genetic algorithm permits to link directly the genetic algorithm program and the code of the numerical models, which is also written in FORTRAN. Thus, the simulated results are made automatically available to the GA module for calculating the fitness of the chromosomes at each generation without any interruption.

## **6.3 Case-Based Module**

### *6.3.1 Case Description*

In the following sections the schemes for describing estuaries and models are illustrated.

#### *6.3.1.1 Estuary Description*

In CBEM the organisation and representation of estuaries is assessed on the basis of the

existing estuarine classifications previously presented in chapter 2. However, the information contained in these classifications cannot be used directly but needs to be carefully combined. This is because each of these classifications is not complete. They focus on specific aspects of the estuarine environment and they do not take into account the possible interdependence between different aspects of an estuary. They also overlap in terms of information and estuary characterisation. Therefore, the present estuary description is designed to systematically organise the information scattered through these classifications into formal and meaningful indices, suitable for the case-based reasoning process. For this reason, only the classifications for estuarine geomorphology, physiography, tidal characteristics and hydrography (water circulation, mixing process and stratification) have been considered. Instead, canonical parameters for estuary classification (e.g. stratification parameter, circulation parameter, densimetric Froude number and estuarine Richardson number) have not been included in the estuary description as this type of data is not always available and inexperienced users in estuarine hydrodynamics may misuse them (K.R. Dyer, 1997).

Following the case-based reasoning approach, illustrated in chapter 4, the present estuary description scheme does not intend to provide a detailed review of the complex physical, chemical and biological processes that take place in estuaries. Instead, it aims to represent the estuary domain through indices that are generally understandable and applicable, making existing relationships between features explicit. For this purpose, the inventories of British and American estuaries (A.L. Buck and N.C. Davidson, 1997; NOAA report, 1985) have been useful inspiring information sources in determining a well-defined description scheme. These inventories offer a reasoned estuary representation, which includes only key estuarine features that are of interest and significant from the perspective of estuary management.

Some of the indices in the estuary description are defined using qualitative values in order to facilitate the indexing and the retrieval process, while other are numeric data. In addition, there are other features that are object symbols defined according to the description provided by the estuary's classification schemes.

After initially describing in which country the estuary is located and its latitude, the description is then organised in terms of physical and hydrographic characteristics. These represent the

major dimensions of an estuarine system, which play an important role in determining the dynamics of an estuary and its behaviour (figure 6.3).

The physical characteristics are

- geomorphological type – Derived from the geomorphological classification drawn by Fairbridge (R.W. Fairbridge, 1980). The proposed classification consists of six types: *fjord, fjard, ria, complex, coastal plain* and *bar built*.
- tidal range - According to the classification by tides, the tidal range is a choice between *microtidal, microtidal-mesotidal, mesotidal-macrotidal* and *macrotidal*.
- estuarine total area
- intertidal area
- maximum channel bank area
- maximum, minimum and average widths
- average depth
- average width to average depth ratio – Three classes are used to qualitatively define the ratio of the average width to average depth. These are *low, moderate, high* and *very high*.
- channel length
- valley length
- grade of estuary meandering – The sinuosity of the estuary channel is classified as *relatively straight, meandering* and *extremely meandering*.
- bed shape - The bed shape can be either *flat* or *sloping bed*.

The hydrographic characteristics are:

- freshwater inflow
- tidal volume
- salinity stratification – Based on the salinity distribution and flow characteristics within an estuary, the estuary can be defined as *salt wedge, partially mixed, vertically mixed* and *well mixed*.



- limit of salt intrusion
- average longitudinal velocity
- tidal period

The Coriolis and wind forces are also included in the description of the estuary physics. Although they are external factors, they can significantly affect the mechanism of mixing within estuaries especially under particular physical conditions (P. Dyke, 1996). Therefore, their effects are represented in the model equations with specific mathematical terms. Through a dimensional analysis approach, the corresponding magnitude of the wind and Coriolis forces for each case is established with respect to the dominant friction stress. Based on the

The screenshot shows a Microsoft Access database form titled "ESTUARY". The form is divided into several sections with various input fields and dropdown menus. The fields are as follows:

- name:** X
- ID:** 7
- country:** ENGLAND
- latitude:** 54
- Physical characteristics:**
  - geomorphological type:** coastal plain
  - tidal range:** macrotidal
  - total area (ha):** 1347
  - intertidal area (ha):** 500
  - max channel bank area (ha):** 30
  - channel length (km):** 33
  - valley length (km):** 25
  - meandering/sinuosity:** relatively straight
  - min. width (m):** 10
  - max width (m):** 800
  - bed:** flat bed
  - avr. width (m):** 600
  - avr. depth (m):** 10
  - avr. width to depth ratio:** low
- Hydrographic characteristics:**
  - tot. freshwater inflow (m3/s):** 1
  - tidal volume (m3):** 3000
  - salinity:** well mixed
  - salinity limit (km):** 30
  - averaged velocity (m/s):** 3
  - tidal period (hr):** 12.5
  - inlets
  - How many inlets? 1
  - coriolis effect
  - wind
- comments:** prevalence of mud flats and salt marshes

At the bottom of the form, there are navigation buttons: "Apply Model", "First", "Previous", "Next", "Last", and "New". The status bar at the bottom indicates "Record: 7 of 8" and "Form View".

Figure 6.3. Estuary description

dimensional analysis's outcome, the system suggests either to take into account or discard their effects in the modelling procedure. The user can then decide to follow the system's advice or to proceed against its suggestion.

In the estuary description special consideration is given to the presence of inlets. These, which are evaluated separately from the main estuary description (figure. 6.4), may influence the dynamics and mixing process within an estuary. Their contribution to the estuarine motion of water and mass continuity must be represented differently in the model equations if classified as either branches or tributaries (J. Bikangaga and V. Nassehi, 1995). Additional mathematical relationships may be necessary to represent the flow conditions at the junction of branching estuaries.

The description unit evaluates the nature of any inlet by comparing it with the estuary's main channel over the average depth, width and velocity dimensions. At this point the user is still able to ignore the response from the system's dimensional analysis and change the computer's recommendation.

The screenshot shows a software window titled "INLETS OF THE ESTUARY". The window contains several input fields and a dropdown menu:

- "number of the inlet" with a text box containing "1".
- "Estuary" with a text box containing "X".
- "inlet name" with a text box containing "Y".
- "position from the mouth(km)" with a text box containing "20".
- "average velocity(m/s)" with a text box containing "2".
- "average depth (m)" with a text box containing "10".
- "the inlet must be considered" with a dropdown menu showing "tributary".

At the bottom of the window, there is a record navigation bar that reads "Record: 1 of 1 (Filtered)".

Figure 6.4. Inlet description

### 6.3.1.2 Model Description

The proposed scheme for model description (figure 6.5) was designed based on the 'Classification of the models of tidal waters' by Hinwood and Wallis (J.B. Hinwood and I.G. Wallis, 1975) and 'Guidelines of the use of computational models in coastal and estuarial studies' by Lawson and Gunn (J. Lawson and I.J. Gunn, 1996).

The screenshot shows a Microsoft Access window titled "Microsoft Access - [frmMaintModels]". The form is titled "MODELS" and contains the following fields and controls:

- estuary:** A text box containing "X".
- problem:** A dropdown menu with "flood routing" selected.
- dimension:** A dropdown menu with "2-D avg on depth" selected.
- numerical method:** A dropdown menu with "finite element" selected.
- purpose:** A dropdown menu with "management tool" selected.
- accuracy:** A dropdown menu with "high" selected.
- manning's coefficient:** A dropdown menu with "varying manning" selected.
- dispersion coefficient:** A dropdown menu with "salt interpolation" selected.
- Governing Assumptions:** Four checkboxes: "wind", "coriolis force", "branching", and "slope". All are currently unchecked.
- limitation of use:** A text box that is empty.
- time consumed:** A dropdown menu with "moderate" selected.

At the bottom of the form, there is a record navigation bar showing "Record: 1 of 1 (Filtered)" and a "Form View" button.

**Figure 6.5.** Model Description

The model is described as:

- dimension - The model dimension can be classified as *1-D*, *1-D network* (i.e for branching estuaries), *2-D averaged on depth*, *2-D averaged on width* and *3-D*.
- numerical technique - The listed numerical techniques are *characteristics*, *finite difference*, *finite element*, *finite volume*, *spectral method* and *diffuse approximation method*.
- model assumptions - The model assumptions are the presence of wind and Coriolis forces, the bed slope and branching effect.
- dispersion coefficient - The dispersion coefficient is classified as either Thatcher/Halerman, if it is calculated using the Thatcher-Halerman formula (Thatcher and Halerman, 1972), or salinity interpolation, if it is determined from the model equations using the salinity measurements obtained through field surveys as known variables.

- Manning's coefficient - The Manning's coefficient can be *frictionless, constant, linear* or *varying*.
- problem - The problem is represented as: *salinity distribution, salt intrusion, sediments, flood-routing* or *pollutant distribution*. If the pollutant distribution option is chosen then the user is offered the possibility to specify if the pollutant is conservative or reactive, as the model equation for the concentration is expressed differently in these two circumstances (see chapter 2).

Considerations related to the model strategy are also included. Thus, some qualitative attributes are included in the model description to measure the model's performance. These features are essential for the retrieval purposes as they are used to estimate the appropriateness of a model when the user's requirements for efficiency and accuracy are taken into account. For instance, a model may provide a sufficiently correct simulation procedure but may be inappropriate according to the aim of the investigation as far as the time required or the accuracy are concerned. These indices are:

- accuracy - The accuracy is a choice between *overall, moderate* and *high*.
- time required for the simulation - The time necessary for running the model code can be defined as *low, moderate* and *high*.
- simulation purpose - The simulation purpose is also included, which expresses the aim of the investigation as *management tool, water quality* or *research*.

### 6.3.2 Case Retrieval

The CB module is also designed to carry out the retrieval process and select from the system's library the model that is most likely to fulfil the user's needs.

#### 6.3.2.1 Model Selection

After entering the new case according to the estuary description scheme, the user clicks on the button APPLY MODEL (figure 6.3) to activate the retrieval process. At this point a new screen, the model selection screen, is opened, where the user can specify the context of the new case. With the purpose of increasing the retrieval efficiency the type of problem to be simulated and the purpose of the undergoing investigation must be given to restrict the search to those portions of the case-base that are likely to provide useful cases. The indices for the

purpose and the problem definition are defined using the same classes employed for the homonymous indices within the model description.

The model selection screen also contains the model approximations. These are selected conforming to what was previously declared in the estuary description about the presence of inlets, the bed shape and the inclusion of wind and Coriolis forces. In the model selection screen the user can change the previous choice and define different governing assumptions.

Furthermore, a lower-bound limit for the degree of similarity (i.e. match aggregate score) must

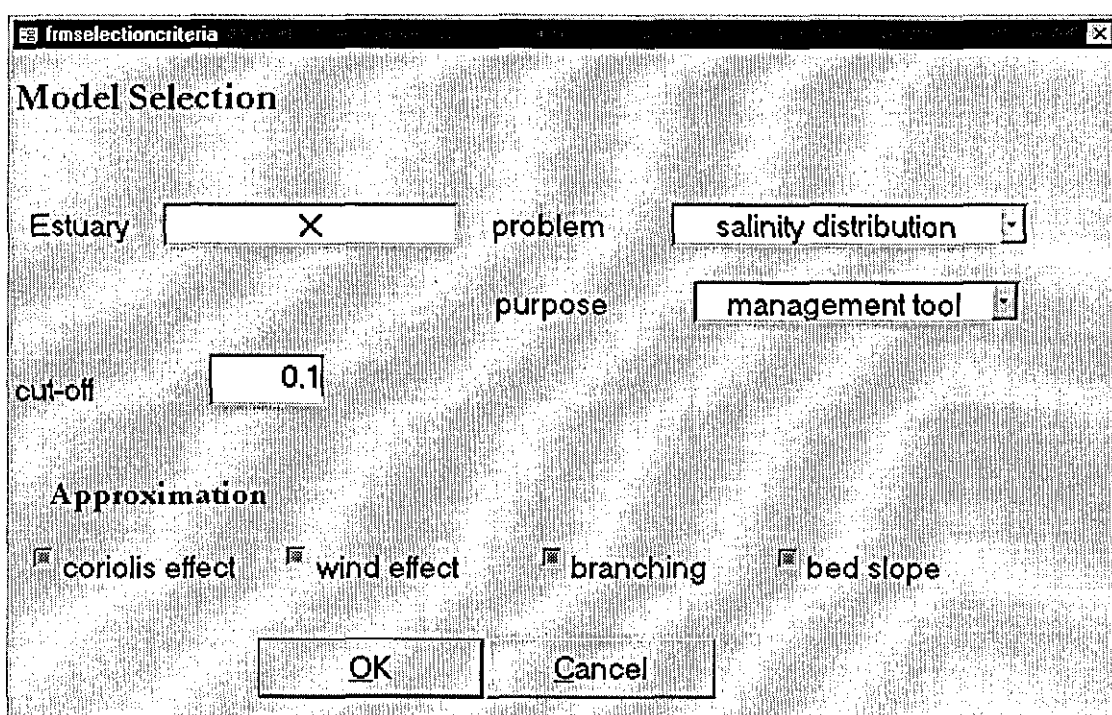


Figure 6.6. Model selection

also be indicated. In the model selection screen this is called cut-off. Only cases with similarity greater than or equal to the cut-off value will be retrieved. The value of the cut-off must be between 0 and 1 as the grade of similarity is calculated on a normalised scale (see chapter 4).

Once the indices of the model selection screen have been chosen, the retrieval process can be activated. The degree of similarity is computed using the nearest neighbour matching procedure, which evaluates the similarity in two stages, based on both the estuary description

and the problem characteristics. The similarity rating between a new and a past case is initially defined based on the values of the physical and hydrographic characteristics contained in their estuary descriptions. A second score is then calculated with respect to the type of investigation to be conducted, the accuracy and simulation time required.

The similarity is measured using a fuzzy approach if the attributes are described on a quantitative scale (i.e. ratio of the total area to the intertidal area). Alternatively, if the descriptors are expressed qualitatively (i.e. the degree of meandering, the model purpose), the matching criterion consists of computing the distance between the two symbols. A more detailed account of the matching process is provided in chapter 7.

The nearest neighbour matching is also combined with the heuristic criteria of exclusion and preference expressed with respect to the model dimension. It must be noted that the model dimension, which depends on the type of problem as well as the estuary's physics, needs to be chosen so that the physical phenomenon under investigation is well-represented without underestimating or oversizing the problem domain. Cases are not retrieved from the case-base if the dimension of the related models is inappropriate for the type of problem considered (criterion of exclusion). Some cases are also preferred over others if their model dimension as evaluated by the system is more suited to simulate the current problem (criterion of preference).

Furthermore, the retrieval procedure implemented in the present scheme uses different sets of matching and importance criteria according to the type of estuarine phenomenon to be modelled. The CBEM employs different sets of importance values  $W_k$  for problems such as salinity distribution and salt intrusion.

To illustrate the retrieval process implemented in CBEM the specific problem of salinity distribution investigated for management purpose is here considered.

#### *6.3.2.2 Similar Estuaries*

In the retrieval process the similarity based on the values of the features of the estuary description is computed first. The similarity of the new case's estuary is estimated as follows:

1. select a set of 5 features  $F$  from the estuary description: the ratio of the average width to the average depth (a), the geomorphological estuary type (b), the tidal range (c), the meandering rate (d), the ratio of the total area to the intertidal area (e), the ratio of maximum bank channel area to the intertidal area (f) and the ratio of the channel length to the average depth (g).
2. assign the degree of relevance  $W= 1$  to (a), 0.75 to (b) and (e), 0.5 to (c), (d) (f) and (g).
3. determine the similarity values  $S = \text{sim}(F_{\delta}^I, F_{\delta}^R)$ , with  $\text{sim}$  as similarity function and,  $I$  and  $R$  referring to the input and retrieved cases, respectively.

ESTUARY	SIMILARITY RATING	NUMBER OF MODELS
TEES	0.848	1
FAL	0.762	2
CONWY	0.773	1

Figure 6.7. Similarity rating with respect to the estuary physical and hydrographic features

4. normalize the match aggregate score  $= \frac{\sum_{k=1}^7 W_k S_k}{\sum_{k=1}^7 W_k}$ .

For each retrieved estuary (figure 6.7) CBEM gives the value of the similarity rating and the number of model previously employed for the specific estuary to simulate the considered estuarine problem (i.e. salinity distribution). The user is then required to select, based on the similarity rating, the retrieved estuaries for further examination in the second phase of the

retrieval process. This computes the similarity of the case with respect to the context of the problem and the model's characteristics.

### 6.3.2.3 Similar Models

The retrieval process continues with the computation of the adequacy of each retrieved model. This is evaluated through a pre-determined set of match values that rank the accuracy, simulation time consumed and purpose of the model based on the investigation aim of the new problem (i.e. management purpose). The procedure for calculating the likelihood of each model to fulfil the user's requirements is described below.

ESTUARY	DIMENSION	ACCURACY	ESTUARY SIM	MODEL SIM
TEES	1-D	overall	0.848	0.545

Record: 1 of 1 (Filtered)

**Figure 6.8.** Similarity rating with respect to the model characteristics

1. exclude case-models with respect to the model dimension.
2. select the following features from the model description: purpose (h), accuracy (i), simulation time (j).
3. apply a set of pre-determined rules to establish the functional role  $M_k$  of each feature with respect to the purpose of the current investigation.
4. assign a grade of relevance  $P=0.75$  to (h) and 0.5 to (i) and (j).
5. apply the criterion of preference with respect to the model dimension. If a model is "preferred", assign the value 1 to the match value  $M_k$  and to the grade of relevance  $P$ .



6. normalise the match aggregate score =  $\frac{\sum_{k=1}^N P_k M_k}{\sum_{k=1}^N P_k}$  with N equal to 4 if the criterion of

preference is valid, otherwise N is equal to 3.

CBEM returns to the user the chosen models, described by the value of their similarity, accuracy, dimension and the similarity of the related estuary (figure 6.8). Based on the information provided, the user, aware of the degree of similarity estimated for both the estuary and the model is responsible for the final selection of the model.

#### 6.4 Genetic Algorithm Module

CBEM's adaptation phase focuses on generating a new set of Manning's coefficients. The input values of these model parameters are adapted to suit the physics of the estuary to be simulated and give a realistic simulation of the estuarine hydrodynamics. This module is developed by combining the classical evolutionary with problem-specific information in the form of heuristic rules and case-based reasoning principles. The classical genetic operations of initialisation, selection, crossover and mutation are modified to incorporate practical information about the estuarine model calibration. This implementation narrows down the search toward areas of the search space that are more likely to contain the best set of parameters. A considerable reduction of the necessary computational time is then obtained.

Since the detailed description of the genetic algorithm will be given in chapter 7, here the design of the genetic algorithm and the logic behind it are presented.

MANNING'S COEFFICIENT SERIES					
Element	n1	n2	n3	n4	n5
Manning's number	0.020	0.020	0.021	0.034	0.021

CHROMOSOME REPRESENTATION					
Genes	n1	n2	n3	n4	n5
Allele	20	20	21	34	21

Figure 6.9. Example of a set of Manning's coefficients and representation of the corresponding chromosome

In the genetic algorithm an initial population of sets of Manning's coefficients is generated. Their genotypical representation is given using the decimal base (figure 6.9). This chromosome representation provides the necessary correspondence between the sets of Manning's coefficients and the chromosomes. Therefore, as the Manning's coefficients differ from one value to another only in the last two digits, the chromosomes are expressed as integers corresponding to the second and third decimal places of Manning's numbers. This representation is more practical since, to preserve the accuracy of modelling, the number of elements in a discretised domain is usually high. Therefore, with a high number of Manning's coefficients the use of integers for the genes significantly facilitates the passage to and from the phenotypical representation and the transformation by genetic operators. The transformation of a chromosome into the corresponding set of Manning's coefficients is required anytime that the fitness function needs to be calculated.

The fitness function, as mentioned in § 5.4.2, can be computed by estimating the discrepancy  $\rho$  between the water surface elevations ( $H_m$ ) measured at different locations within the estuary, and their corresponding simulated values ( $H_s$ ). Each sampling station  $j$  is characterised by a set of experimental data corresponding to the water surface elevations observed at different time levels, indicated by  $n$ . Denoting the total number of sampling stations by  $J$  and the total number of samples, collected at each station during a tidal period, by  $N$ , the series of all measured water surface elevations can be represented as  $H_m = \{(h_m)_j^n, j=1, \dots, J; n=1, \dots, N\}$  and the set of all simulated values as  $H_s = \{(h_s)_j^n, j=1, \dots, J; n=1, \dots, N\}$ .

Hence, the discrepancy between  $H_m$  and  $H_s$  is given as:

$$\rho(H_s, H_m) = \left[ \sum_{j=1}^J \sum_{n=1}^N \left( (h_s)_j^n - (h_m)_j^n \right)^2 \right]^{1/2} \quad (6.1)$$

The fitness of each chromosome is calculated as the reciprocal of  $\rho$ :

$$\varphi_I = \frac{1}{\rho} \quad (6.2)$$

In order to find which chromosome gives a minimum for equation (6.1), water surface elevations for all chromosomes in each generation must be simulated. Hence,  $h_s$  at each station  $j$  for the time levels  $n$  is calculated. For this purpose each set of Manning's coefficients, corresponding to a specific chromosome, needs to be included in the model input data required. The model is run for a number of tidal cycles until the convergence is achieved. The simulated results are then written to an output datafile which is given to the genetic algorithm for calculating the fitness of the chromosome (figure 6.10).

Therefore, the genetic algorithm consists of several steps.

Step 1. Read the datafile containing the values of the parameters for the genetic algorithm:  $M$  for the size of the population,  $R$  for the number of generations,  $P_C$  and  $P_M$  for the probability of crossover and mutation,  $H_m$  for the set of observed water surface elevations.

Step 2. Generate  $M$  chromosomes employing the integer representation as discussed

Step 3. Transform the chromosomes into the corresponding sets of Manning's coefficients (genotype to phenotype transformation).

Step 4. Include the set  $I$  of Manning's coefficients into the model input datafile

Step 5. Call and run the model program

Step 6. Compute the simulated water surface elevations  $H_s$

Step 7. Calculate the fitness function  $\phi_f$

Step 8. Transform the sets of Manning's coefficients into the corresponding chromosomes (phenotype to genotype transformation).

Step 9. Select surviving individuals

Step 10. Apply crossover

Step 11. Apply mutation

The steps are repeated for the number of generations  $R$ , while the steps 4-7 are also repeated for each chromosome in the population.

### **6.5 Model Library**

The model library contains the FORTRAN codes of the models described in the CB module. The model library is the form of subroutines so that the appropriate model can be called automatically by the GA module and the simulated results used for calculating the fitness functions.

### **6.6 Conclusions**

The system presented consists of three modules: a case-based module, a genetic algorithm module and a library of numerical model codes. Through the case-based module a model strategy is defined by considering the similarity of the new situation to the past cases with respect to the physical behaviour of the estuary and the specific problem to be modelled. The CB module also handles and evaluates the available knowledge on existing models and their applications accordingly to the principles of the case description and retrieval. Once an appropriate model procedure is selected, the genetic algorithm module adjusts the model Manning's coefficients to suit the new case. The genetic algorithm is directly linked to the computer code of the selected model contained in the model library. By testing different sets of parameters according to the genetic algorithm theory, the genetic algorithm provides a robust and efficient methodology for optimising the Manning's coefficients with a considerable reduction in computational time.

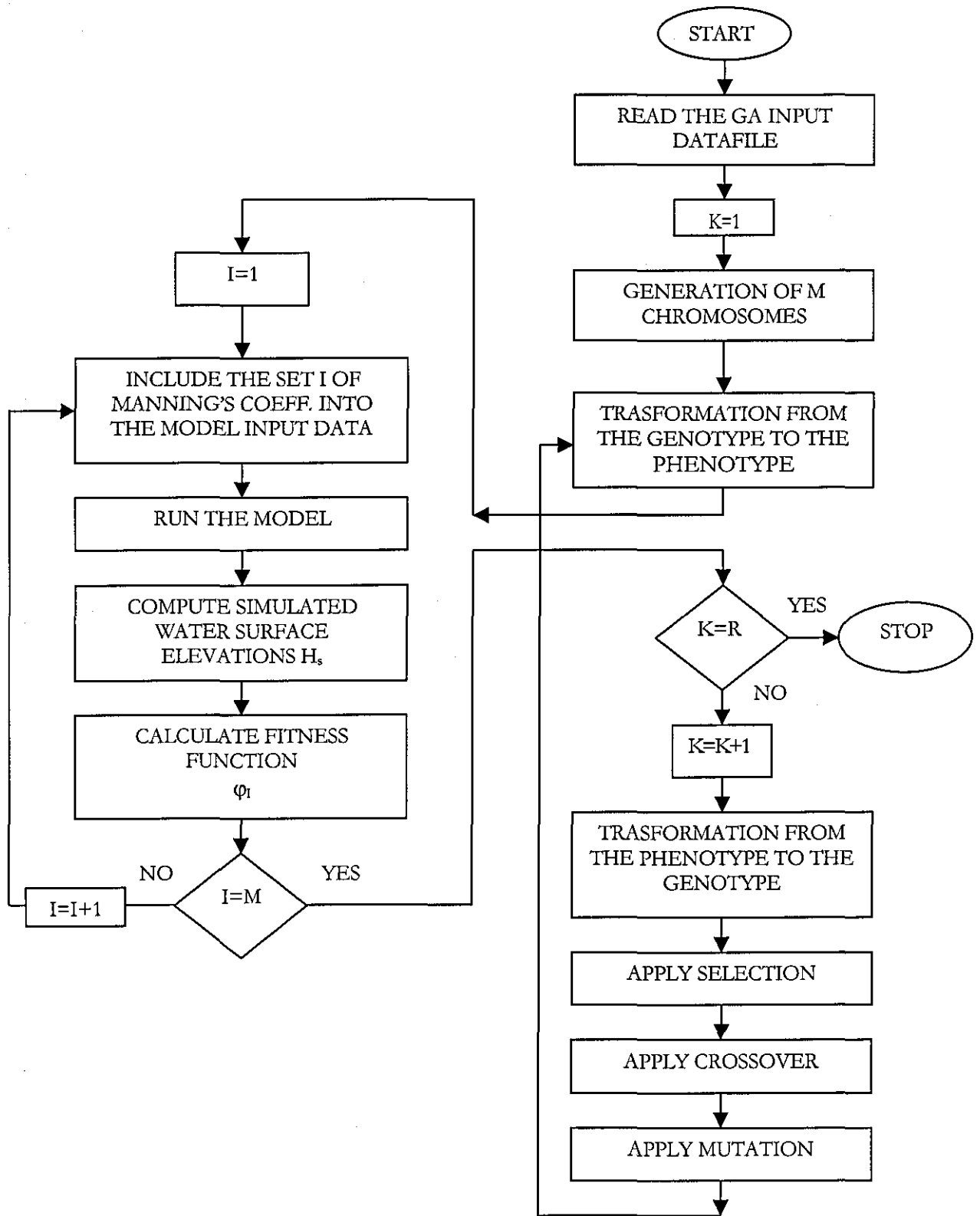


Figure 6.10. Structure of the implemented genetic algorithm

## SYSTEM DESIGN

### 7.1 Introduction

This chapter highlights the design issues encountered during the development of CBEM. It describes how the knowledge related to estuaries and models is specifically structured within the system defining the case description, retrieval and adaptation.

The knowledge base implemented derives from the direct experience of various experts (e.g. experts in estuarine modelling, geomorphology and oceanography). An extensive literature search was also made. Literature accounts, maps and papers describing the physics of estuaries, modelling procedures of estuarine processes with particular reference to estuaries in Britain (A.L. Buck and N.C Davidson, 1997), North America (NOAA, 1985), South Africa (J.A.G. Cooper, 2001) and Malaysia (Z.Z. Ibrahim et al., 1997), were considered. The available literature provided the necessary base for discussion with the experts. Interviewing the experts helped extending the information gathered from the literature search and including 'best practice' guidelines, which may not be published yet. Also assumptions and hypotheses, on which existing models are based, were studied and integrated in the knowledge system.

The knowledge is organized in the CB module in the form of "IF... THEN... ELSE" rules to estimate feature values and quantify the similarity between cases. In the GA module domain-specific knowledge from the practice of estuarine model calibration is used to modify the classical genetic operators into a new set of operators that behave differently.

The layout of this chapter is as follows: §7.2 looks at the characterisation of the description component; §7.3 covers the retrieval mechanisms; and in §7.4 the design of the genetic algorithm for the adjustment of the set of Manning's coefficients employed in a model is described.

## 7.2 Case Description

In the case description some indices have been constructed in such a way that their values is suggested by the system as the estuary description gradually progresses. These indices are: the average width to average depth ratio, the degree of meandering, the salinity stratification, the Coriolis and wind forces (figure 6.3) and the nature of the inlets (figure 6.4). The characterisation of these indices is given below.

### 7.2.1 The Average Width to the Average Depth Ratio

The average width and average depth are indicated in the estuary description as numbers. However, their ratio is expressed qualitatively. A qualitative measurement for this index makes it easier to relate the average width to the average depth ratio to other dimensions of the case description with a consequent increase of the retrieval process flexibility.

Using fuzzy set theory (E. Cox, 1994), four membership functions are constructed to convert the numerical value of the average width to the average depth ratio into a categorical scale expressed as low, moderate, high and very high. The four membership functions were implemented by trial and error, using a number of British estuaries to adjust the input parameters of each function.

The four membership functions are given as follows:

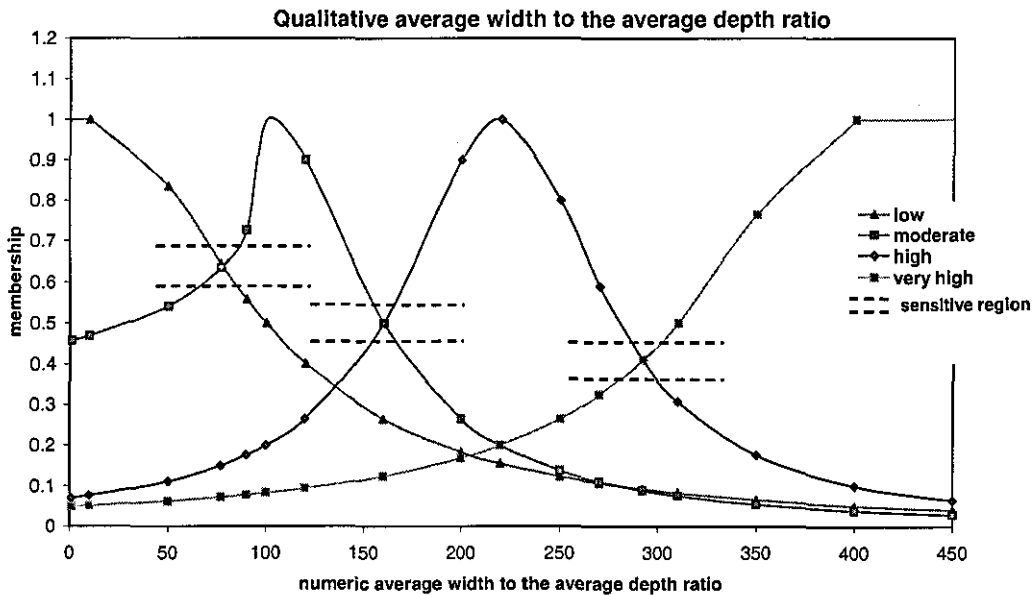
$$\text{if ratio} \leq 10 \text{ THEN } m_{low} = 1 \text{ ELSE } m_{low} = \left[ 1 + \left( \frac{|\text{ratio} - 10|}{90} \right)^2 \right]^{-1}$$

$$\text{if ratio} \leq 100 \text{ THEN } m_{moderate} = \left[ 1 + \left( \frac{|\text{ratio} - 100|}{70} \right)^{0.5} \right]^{-1} \text{ ELSE } m_{moderate} = \left[ 1 + \left( \frac{|\text{ratio} - 100|}{60} \right)^2 \right]^{-1}$$

$$m_{high} = \left[ 1 + \left( \frac{|\text{ratio} - 220|}{60} \right)^2 \right]^{-1}$$

$$m_{very\ high} = \left[ 1 + \left( \frac{|\text{ratio} - 400|}{100} \right)^2 \right]^{-1} \text{ , if ratio} \geq 400 \text{ THEN } m_{very\ high} = 1$$

where *ratio* indicates the numerical value of ratio of the average width to the average depth.



**Figure 7.1.** The membership functions used to express qualitatively the average width to the average depth ratio

Through this indexing the average width to the average depth ratio is classified over the classes of low, moderate, high and very high with different membership grades. The class with the highest membership value is displayed on the case description screen. For instance, if the estuary is represented by (low/0.44, moderate/0.97, high/0.23, very high/0.21) the ratio is defined in the case description as moderate. However, if two classes are assigned with the same maximum value (e.g. low/0.64, moderate/0.64, high/0.15, very high/0.09), or the absolute difference between the first and the second highest values is less or equal to 0.1 (e.g. low/0.65, moderate/0.60, high/0.15, very high/0.09) then the index is defined as low/moderate, moderate/high or high/very high, depending on the two classes that share the same or similar maximum values. In figure 7.1, the “sensitive regions” indicate those areas of the graph where the difference between the values of two neighbour functions is less or equal to 0.1.

The user can also select the qualitative ratio of the average width to the average depth. With the selection of one of these classes, the user automatically determines the value of all four membership functions. For instance, by choosing “low” he/she assigns the membership value 1 to the class “low”, and gives 0 to the other three classes. The memberships grades associated to each class if the user selects the qualitative value of the ratio are shown in table 7.1.



**Table 7.1.** Membership grades automatically assigned when the user selects the qualitative ratio of the average width to the average depth

USER SELECTION	MEMBERSHIP VALUE			
	LOW	MODERATE	HIGH	VERY HIGH
LOW	1	0	0	0
LOW/MODERATE	0.5	0.5	0	0
MODERATE	1	0	0	0
MODERATE/HIGH	0	0.5	0.5	0
HIGH	0	0	1	0
HIGH/VERY HIGH	0	0	0.5	0.5
VERY HIGH	0	0	0	1

The user's decision always overrides the calculated values based on the membership functions.

### 7.2.2 The Grade of Meandering

The grade of meandering of an estuary is definable in a similar way to rivers (D. Rosgen, 1996), as the ratio of the channel length to the length of the valley within which the water stream is located.

Again, in this system, the degree of meandering is expressed qualitatively transforming the ratio of the channel length to the valley length.

The conversion to a categorical scale is provided by the following membership functions:

$$\text{if } prop \leq 1.3 \text{ THEN } m_{\text{relatively straight}} = 1 \text{ ELSE } m_{\text{relatively straight}} = \left[ 1 + \left( \frac{|prop - 1.3|}{0.1} \right)^2 \right]^{-1}$$

$$m_{\text{meandering}} = \left[ 1 + \left( \frac{|prop - 1.6|}{0.1} \right)^2 \right]^{-0.5}$$

$$m_{\text{extremely meandering}} = \left[ 1 + \left( \frac{|prop - 1.9|}{0.1} \right)^2 \right]^{-2}, \text{ if } prop \geq 1.9 \text{ THEN } m_{\text{extremely meandering}} = 1$$

where *prop* is the numerical value of the ratio of the channel length to the valley length.

As for the ratio of the average width to the average depth, the membership functions for the grade of meandering were determined by trial and error, using a set of British estuaries to calibrate the functions.

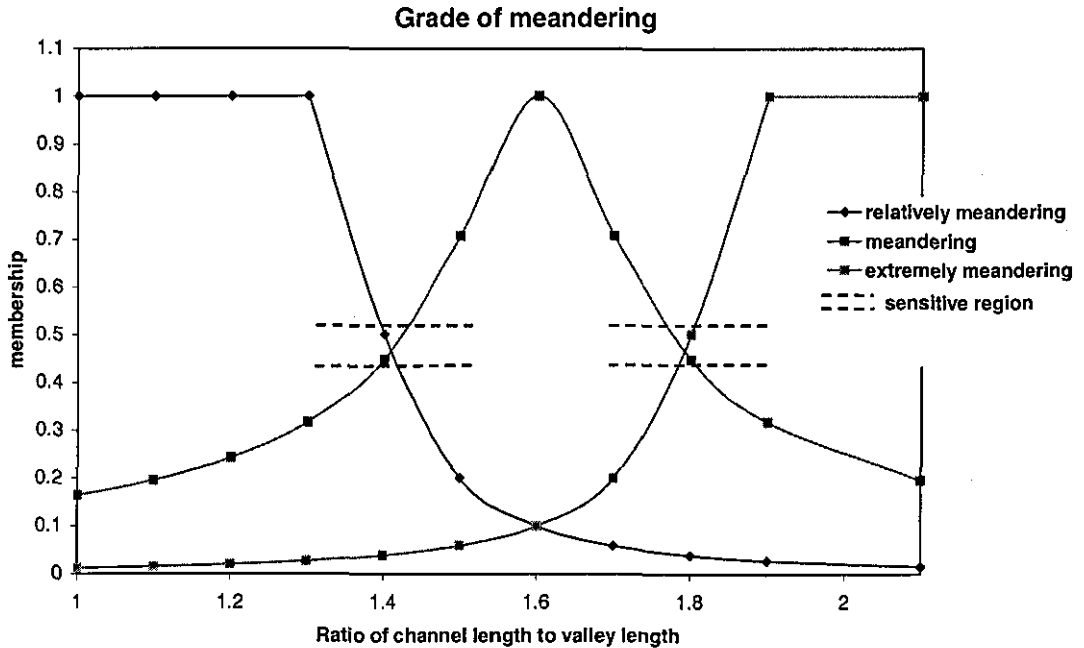


Figure 7.2. The membership functions used to express qualitatively the degree of meandering

The category with the highest membership value is assigned to the grade of meandering in the case description. For instance, if for an estuary the grade of meandering is classified as (relatively straight/0.96, meandering/0.34, extremely meandering/0.0008), this is represented in the case description as “relatively straight”. It is also possible that either there are two membership functions having the same maximum value (e.g. relatively straight/0.46, meandering/0.46, extremely meandering/0.002) or the absolute difference between the first and the second highest maximum is less or equal to 0.1 (e.g. relatively straight/0.5, meandering/0.42, extremely meandering/0.002). In these two cases, the grade of meandering in the case description is defined as relatively straight/meandering or meandering/extremely meandering. In the case description the qualitative value of the grade of meandering can also be selected by the user. By doing that, the value of all three membership functions is automatically given as shown in table 7.2.

### 7.2.3 Estimation of the Salinity Stratification

By using the estuary classifications, given in chapter 2, it is possible to relate the salinity

**Table 7.2.** Membership grades automatically assigned when the user selects the grade of meandering

USER SELECTION	MEMBERSHIP VALUE		
	RELATIVELY STRAIGHT	MEANDERING	EXTREMELY MEANDERING
RELATIVELY STRIGHT	1	0	0
REALITIVELY STRAIGHT/MEANDERING	0.5	0.5	0
MEANDERING	1	0	0
MEANDERING/EXTREMELY MEANDERING	0	0.5	0.5
EXTREMELY MEANDERING	0	0	1

stratification to the tidal range and the average width to the average depth ratio. Thus, the following rules are included in the system's knowledge base:

1. **IF** *the estuary is macrotidal* **THEN**  
     **IF** *the width-depth ratio is "low or moderate or low/moderate"* **THEN** *salinity is "well mixed"* **ELSE** *salinity is "vertically mixed"*
2. **IF** *the estuary is mesotidal/ macrotidal* **THEN**  
     **IF** *the width-depth ratio is "low or moderate or low/moderate"* **THEN** *salinity is "partially mixed"* **ELSE** *salinity is "vertically mixed"*
3. **IF** *the estuary is microtidal/ mesotidal* **THEN**  
     **IF** *the width/ depth ratio is "low or moderate or low/moderate"* **THEN** *salinity is "partially mixed"* **ELSE** *salinity is "vertically mixed"*
4. **IF** *the estuary is microtidal* **THEN**  
     **IF** *the width/ depth ratio is "low or moderate or low/moderate"* **THEN** *salinity is "salt wedge"* **ELSE** *salinity is "partially mixed"*

#### 7.2.4 Coriolis and Wind Forces

Although the dimensional analysis can only be taken as a general guide, it forms an essential element in eliminating the unnecessary terms in the equation of motion used in the estuary model. In the case description module, specific conditions based on the dimensional analysis are implemented to provide an estimate of the magnitude of the wind and Coriolis forces.

The Coriolis effect is expressed as:

$$F_c = \frac{2U\omega\sin(\phi)}{g} \quad (7.1)$$

where  $U$  is the average estuarine velocity,  $g$  is the gravity acceleration,  $\omega$  is the angular velocity of the Earth and  $\phi$  is the angle of the latitude of the Earth. The angular velocity of the Earth is approximately  $7.29 \times 10^{-5}$  rads per second.

The wind force is represented as a function of the wind stress:

$$\frac{\tau_{ws}}{\rho g H} \quad (7.2)$$

where  $\tau_{ws}$  indicates the wind stress,  $\rho$  is the water density and  $H$  is the average estuarine depth.

The wind stress  $\tau_{ws}$  is expressed as:

$$\tau_{ws} = c_D \rho_a U_w^2 \quad (7.3)$$

where  $c_D$  is a dimensionless drag coefficient equal to 0.00015,  $\rho_a$  is the density of the air equal to  $1.2 \text{ kg/m}^3$ , and  $U_w$  is the wind velocity.

The friction term is estimated as:

$$F_f = \frac{n^2 U^2}{H^{4/3}} \quad (7.4)$$

where  $n$  is the Manning's coefficient (see chapter 2), assumed for the purpose of the dimensional analysis, equal to  $0.02 \text{ sm}^{-1/3}$ .

The system suggests to the user to include the effects of the Coriolis and the wind forces only if the terms calculated using expressions (7.1) and (7.2), respectively, give values that are more than 10% of the friction term evaluated according to (7.4).

In addition to the dimensional analysis, the following rules are also included to evaluate the importance of the Coriolis force:

**IF** *the average width to the average depth ratio is "low"* **THEN** *the Coriolis force must be excluded.*

**IF** *the average width to the average depth ratio is "high or very high or high/very high" AND the latitude is in the range between 50 and 60 degree AND the tidal range is not "macrotidal"* **THEN** *the Coriolis force must be included.*

The rules override the dimensional analysis.

#### 7.2.5 Inlets

The system also suggests to the user if the described inlets into the main channel are classifiable as tributaries or branches. The nature of a inlet is decided by the system according to the following rule:

**IF** *the ratio of the inlet's average velocity to the average estuarine velocity is greater or equal to 0.1 AND the ratio of the average inlet's depth to the estuarine average depth is greater or equal to 0.1* **THEN** *the inlet is a branch* **ELSE** *it is a tributary.*

### 7.3 Case Retrieval

The retrieval process is structured to provide the user with two similarity scores that quantify the similarity between cases with respect to the features contained in the estuary descriptions and characteristics of the models previously used. Thus, the case retrieval consists of two different frameworks, each of which estimates the corresponding similarity score through a specific data retrieval scheme.

### *7.3.1 Similar Estuaries*

The first similarity score is calculated by comparing cases along a certain number of indices used in the estuary description. The system employs different sets of indices according to the type of problem considered (i.e. sediments, salinity distribution and etc.).

In CBEM, which is currently designed to advise the user over the problems of salinity distribution and salt intrusion, only two sets of features from the estuary description are considered. In case of salinity distribution the features used for calculating the estuary similarity score are: the ratio of the average width to the average depth, geomorphological type, tidal range, ratio of the intertidal area to the total area, ratio of the maximum channel bank area to the intertidal area, grade of meandering and ratio of the channel length to the average depth. For the salt intrusion problem, the first similarity rating is calculated using indices such as the average width to the average depth ratio, geomorphological type, tidal range and the ratio of the channel length to the average depth.

According to the procedure for calculating the first similarity score illustrated in chapter 6, each feature within both sets is associated with a specific importance value that defines the feature's relevance for the assessment of the similarity between cases (see chapter 4). The degree of relevance of each feature mainly depends on the type of the problem. For instance, the average width to the depth ratio, which is used to compute the similarity in case of both salinity distribution and salt intrusion, has a different degree of relevance with respect to these two problems.

The degrees of relevance associated with the salinity distribution and salt intrusion problems are given in table 7.3 and 7.4, respectively.

#### *7.3.1.1 First Similarity Score*

The indices employed for computing the first similarity score are of different kinds. The ratios of the intertidal area to the total area, the maximum bank channel to the intertidal area and the channel length to the average depth are numerical. Instead, the geomorphological type and

**Table 7.3.** Degrees of relevance for the features from the estuary description utilised to calculate the first similarity score with respect to the salinity distribution problem

<b>SALINITY DISTRIBUTION: FEATURES</b>	<b>DG. OF RELEVANCE</b>
ratio of average width to average depth	1
geomorphological type	0.75
tidal range	0.75
grade of meandering	0.5
ratio of intertidal area to the total area	0.5
ratio of max. bank channel area to intertidal area	0.5
ratio channel length to average depth	0.5

**Table 7.4.** Degrees of relevance for the features from the estuary description utilised to calculate the first similarity score with respect to the salt intrusion problem

<b>SALT INTRUSION: FEATURES</b>	<b>DG. OF RELEVANCE</b>
ratio of avr. width to avr. depth	1
geomorphological type	1
tidal range	1
ratio channel length to average depth	0.5

tidal range are defined as classes of objects and, the degree of meandering and the average width to the average depth ratio are represented on a qualitative scale. The degree of similarity along each of these indices is differently calculated according to their kinds (chapter 4).

For those indices that are represented by numerical values, the similarity is measured using the similarity function (4.3). For instance, the similarity with respect to the ratio of the intertidal area to the total area is given as:

$$SIM(\text{itratio}^I, \text{itratio}^R) = \left[ 1 + \left( \frac{|\text{itratio}^I - \text{itratio}^R|}{\sqrt{\text{itratio}^R}} \right)^2 \right]^{-1} \quad (7.5)$$

with *itratio* indicating the ratio of the intertidal area to the total area and I and R referring to the input and the retrieved case, respectively.

Equivalent functions are also defined for the ratio of the maximum bank channel to the intertidal area and the ratio of the channel length to the average depth.

For the qualitative indices such as the average width to the average depth ratio and grade of meandering, the similarity measure is based on "the geometric model distance" (T.W. Liao et al., 1998), which computes the degree of similarity as function of the multiple class memberships previously defined using fuzzy sets.

For instance, if the multiple membership values of the average width to the average depth ratio for a new case I and a past case R are represented as  $I(\text{low}/a_1, \text{moderate}/a_2, \text{high}/a_3, \text{very high}/a_4)$  and  $R(\text{low}/b_1, \text{moderate}/b_2, \text{high}/b_3, \text{very high}/b_4)$ , the degree of similarity is estimated as follows:

$$SIM(F^I, F^R) = 1 - \frac{\sum_{k=1}^4 |a_k - b_k|}{4} \quad (7.6)$$

Similarly, if the multiple membership values for the degree of meandering is represented for a new case I and a past case R as  $I(\text{relatively straight}/a_1, \text{meandering}/a_2, \text{extremely meandering}/a_3)$  and  $R(\text{relatively straight}/b_1, \text{meandering}/b_2, \text{extremely meandering}/b_3)$ , the similarity function for the grade of meandering is given as:

$$SIM(F^I, F^R) = 1 - \frac{\sum_{k=1}^3 |a_k - b_k|}{3} \quad (7.7)$$



Differently from the previous two types of indices considered, the degree of similarity along object symbols such as geomorphological type and tidal range cannot be computed employing any standard function. As mentioned in chapter 4, it is up to the system's designer to establish the similarity between classes of objects. Except for the event when two cases are represented by the same value and the associated similarity is clearly 1, in any other situation it is necessary to assign to each pair of values a pre-determined similarity value.

In case of the geomorphological type, the degree of similarity between its values is obtained by grouping the geomorphological classes in three categories:

1. Fjord, Fjard, Ria
2. Coastal Plain, Complex
3. Bar Built

Given such a classification, the similarity between two geomorphological types is defined as follows:

**IF** *two cases are represented by the same geomorphological type* **THEN**  $\text{Sim}(F^I, F^R) = 1$

**IF** *two cases are represented by geomorphological types in the same category (i.e. ria and fjard)* **THEN**  
 $\text{Sim}(F^I, F^R) = 0.8$

**IF** *two cases are represented by geomorphological types in adjacent categories (i.e. ria and costal plain; complex and bar built)* **THEN**  $\text{Sim}(F^I, F^R) = 0.4$

**IF** *two cases are represented by geomorphological types in not-adjacent categories (i.e. ria and bar built)* **THEN**  $\text{Sim}(F^I, F^R) = 0.2$

In case of the tidal range, the degree of similarity between its values is determined by arranging them in the following order (microtidal, microtidal-mesotidal, mesotidal-macrotidal, macrotidal), and establishing the following rules:

**IF** *two cases are represented by the same tidal range value* **THEN**  $\text{Sim}(F^I, F^R) = 1$

**IF** *two cases are represented by adjacent tidal range values* **THEN**  $\text{Sim}(F^I, F^R) = 0.7$

**IF** *two cases are represented by tidal range values separated by one value* **THEN**  $\text{Sim}(F^I, F^R) = 0.3$

**IF** *two cases are represented by tidal range values separated by two values* **THEN**  $\text{Sim}(F^I, F^R) = 0.1$

### 7.3.2 Similar Models

The cases retrieved from the system's memory with respect to the estuary similarity score are further ranked with respect to their model's characteristics. A second similarity score is calculated based on the values of indices such as simulation purpose, time required for the computation and accuracy. During this stage of the retrieval process the cases are also filtered with respect to the criteria of exclusion and preference, formulated with respect to the model dimension.

#### 7.3.2.1 Criterion of Exclusion

Exclusion is applied to whittle down from the retrieved cases whose model dimension is considered inappropriate for modelling the new case. It is implemented through different sets of rules based on the type of problem considered. The rules for exclusion in case of salinity distribution and salt intrusion are given below.

- **IF** *the estuary has inlets* **AND** *at least one of them is a "branch"* **THEN** *eliminate 1-D models*  
**ELSE** *eliminate 1-D network models*
  
- **IF** *the problem to investigate is salinity distribution* **THEN**
  1. **IF** *the estuary is a "ford, "ria" OR "fjord"* **THEN**  
*eliminate all 3-D AND 2-D moving boundary models*  
**IF** *(the width/depth ratio is "high" OR "very high" OR "high/very high")* **THEN** *eliminate 2-D average on width models*
  2. **IF** *the estuary is a "coastal plain" OR "complex"* **THEN**  
**IF** *(the salinity is "well mixed" OR "vertically mixed") OR (salinity is "partially mixed" AND width/depth ratio is "very high")* **THEN** *eliminate all 3-D models*

**IF** (the width/depth ratio is "high" OR "very high" OR "high/very high") **THEN** eliminate all 2-D average on width models

**IF** (salinity is "partially mixed" OR "salt wedge") **AND** (the width/depth ratio is "low" OR "moderate" OR "low/moderate") **THEN** eliminate all 2-D average on depth models

3. **IF** the estuary is bar built **THEN**

**IF** (the salinity is "salt wedge" OR "partially mixed") **AND** (the width/depth ratio is "very high") **THEN** eliminate from the selection 3-D models

• **IF** the problem to investigate is salt intrusion **THEN** eliminate all 3-D and 2-D moving boundary models

#### 7.3.2.2 Second Similarity Score

The application of the criterion of exclusion is followed by the computation of the second similarity score formulated with respect to the indices of accuracy, simulation time consumed and model purpose contained in the model description. A degree of match is assigned to each value of these indices, defined with respect to the purpose of the new investigation specified in model selection screen (chapter 6). These match values estimate the capacity of each retrieved model to address the goal of the new investigation in terms of model purpose, accuracy and simulation time consumed. The degrees of match are presented below in three tables: the first one (table 7.5) is constructed for calculating the similarity between the values of the model purpose associated with the new and past case, while table 7.6 and 7.7 quantify the correspondence between the model purpose of the new case and the dimensions of simulation time consumed and accuracy, respectively.

A degree of relevance is also assigned to each of these three matches to take into account their importance (table 7.8). Finally, it must be noted that the second similarity score is calculated without making any difference with respect to the type of problem considered (i.e. salinity distribution, salt intrusion).

#### 7.3.2.3 Criterion of Preference

As mentioned in chapter 6, during the computation of the second similarity score the criterion of preference is also employed. Preference is expressed for those cases whose model dimension is more suitable for the new case. The criterion of preference is treated as an extra

**Table 7.5.** Degrees of match with respect to the index of simulation purpose

SIMULATION PURPOSE (NEW CASE)	SIMULATION PURPOSE (PAST CASE)			
		MANAGEMENT TOOL	WATER QUALITY	RESEARCH
	MANAGEMENT TOOL	1	0.8	0.7
	WATER QUALITY	0.5	1	0.8
	RESEARCH	0.1	0.5	1

**Table 7.6.** Degrees of match for the index of simulation time consumed with respect to the simulation purpose of the new investigation

SIMULATION PURPOSE (NEW CASE)	SIMULATION TIME CONSUMED (PAST CASE)			
		LOW	MODERATE	HIGH
	MANAGEMENT TOOL	1	0.5	0.2
	WATER QUALITY	1	0.7	0.4
	RESEARCH	1	0.8	0.6

**Table 7.7.** Degrees of match for the index of accuracy with respect to the simulation purpose of the new investigation

SIMULATION PURPOSE (NEW CASE)	ACCURACY (PAST CASE)			
		OVERALL	MODERATE	HIGH
	MANAGEMENT TOOL	0.5	0.8	1
	WATER QUALITY	0.4	0.7	1
	RESEARCH	0.2	0.6	1

**Table 7.8.** Set of degrees of relevance used to weight the importance of model purpose, simulation time required and accuracy in the computation of the second similarity score

FEATURES	DEGREE OF RELEVANCE
Model purpose	0.75
Simulation time consumed	0.5
Accuracy	0.5

feature. For those cases for which preference is satisfied, the second similarity score is computed with respect to four indices instead of three (chapter 6). Preference is used in the numerical function (4.1) with a match value and a degree of relevance both equal to 1.

In case of salinity distribution the following rules are employed to filter cases with respect to the criterion of preference:

1. **IF** (*the estuary has inlets and at least one is a "branch"*) **IF** (*salinity is "vertically mixed" OR "well mixed"*) **AND** (*the width/depth ratio is "low" OR "moderate" OR "low/moderate"*) **THEN** 1-D network models are preferred
2. **IF** (*the estuary has inlets and at least one is a "branch"*) **AND** (*the salinity limit is beyond the junction*) **AND** (*the investigative purpose is "research" OR "water quality"*) **THEN** 2-D moving boundary models are preferred
3. **IF** *the estuary is a "fjord"* **THEN**  
**IF** *the salinity is "partially mixed" AND (the width/depth ratio is "low" OR "moderate" OR "low/moderate")* **THEN** 2-D average on width models are preferred  
**IF** *the salinity is "partially mixed" AND (the width/depth ratio is "high" OR "very high" OR "high/very high")* **THEN** 2-D average on depth models are preferred  
**IF** *the salinity is "well mixed"* **THEN** 1-D models and 1-D network models are preferred
4. **IF** *the estuary is a "ria" OR "fjord"* **THEN**  
**IF** *the salinity is "partially mixed" IF (the width/depth ratio is "low" OR "moderate" OR "low/moderate")* **THEN** 2-D average on width models are preferred **ELSE IF** (*the width/depth ratio is "high" OR "very high" OR "high/very high"*) **THEN** 2-D average on depth models are preferred  
**IF** *the salinity is "well mixed"* **THEN** 1-D models and 1-D network models are preferred  
**IF** *the salinity is "vertically mixed" AND (the width/depth ratio is "low" OR "moderate" OR "low/moderate")* **THEN** 1-D models and 1-D network models are preferred
5. **IF** *the estuary is a "coastal plain" OR "complex"* **THEN**  
**IF** (*the salinity is "partially mixed" OR "salt wedge"*) **AND** (*the width/depth ratio is "high" OR "very high" OR "high/very high"*) **THEN** 2-D average on depth models are preferred

IF (the salinity is "partially mixed" OR "salt wedge") AND (the width/depth ratio is "low" OR "moderate" OR "low/moderate") THEN 2-D average on width models are preferred

IF (the salinity is "partially mixed" OR "salt wedge") AND (the width/depth ratio is "high" OR "very high" OR "high/very high") AND (the intertidal area/total area is greater than 0.6) THEN 2-D moving boundary models are preferred

IF the salinity is "well mixed" THEN 1-D models and 1-D network models are preferred

IF the salinity is "vertically mixed" AND (the width/depth ratio is "low" OR "moderate" OR "low/moderate") THEN 1-D and 1-D network models are preferred ELSE 2-D average on depth models are preferred

6. IF the estuary is bar built THEN

IF (the salinity is "salt wedge" OR "partially mixed") THEN

IF the width/depth ratio is "very high" THEN 2-D average on depth models are preferred  
ELSE 3-D models are preferred.

In case of the salt intrusion problem preference is expressed according to the following set of rules:

7. IF (the estuary is a "fjord", "fjard" OR "ria") THEN 1-D models and 1-D network models are preferred

8. IF (the estuary is "complex" OR "coastal plain") THEN

IF (the width/depth ratio is "low" OR "moderate" OR "low/moderate") THEN 1-D models and 1-D network are preferred ELSE 2-D average models on depth are preferred

9. IF the estuary is "bar built" THEN 2-D average on depth models are preferred

#### 7.4 Case Adaptation

The GA module performs the case adaptation by optimising the set of Manning's coefficients for the models through a genetic algorithm, which is hybridised with knowledge related to the estuarine model calibration. The proposed GA has been demonstrated to be more efficient than the classical genetic algorithm formulation. In the tests conducted using the classical genetic algorithm, realistic results were produced for a limited range of the Manning's coefficient (0.015-0.030). Figure 7.3 shows simulated water surface elevations for the location

of Easthook on the Upper Milford Haven estuary, South-Wales<sup>1</sup>. These are generated carrying out the model optimisation by either the classical genetic algorithm or manual calibration. As it is evident, the simulated values, obtained using a classical genetic algorithm, are less accurate than those obtained through a manual calibration

The hybridisation of the classical genetic algorithm formulation with specific knowledge of estuarine model calibration offers a genetic based optimisation method able to provide accurate and realistic simulation results with a consistent reduction of the discrepancy between the observed and the model values. The hybridisation consists of modifying versions of the classical genetic operations of initialisation, selection, crossover and mutation. These are designed to incorporate specific knowledge related to estuaries and the calibration of estuarine models, addressing the search towards zones of the search space that more likely contained suitable solutions.

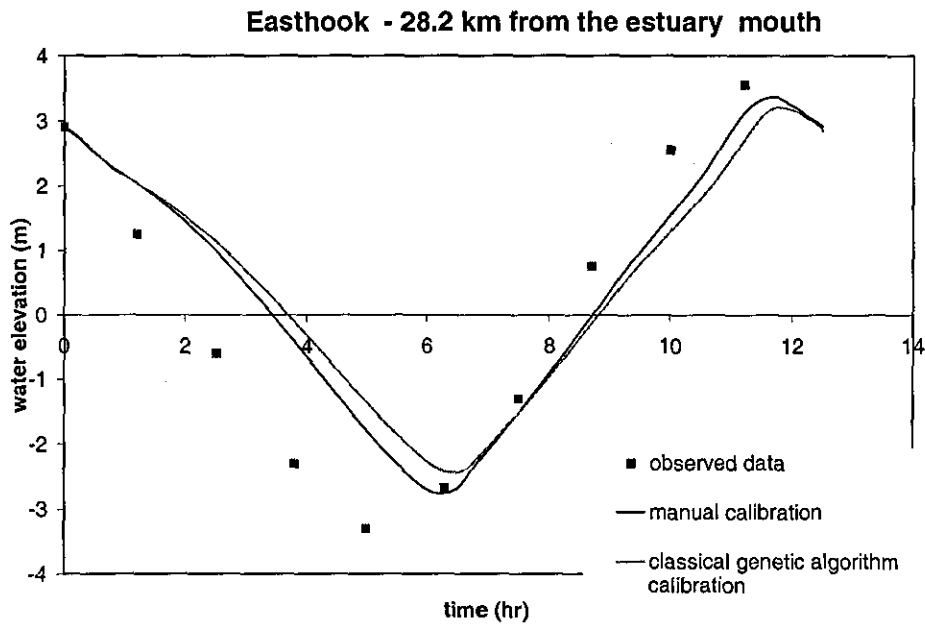
#### *7.4.1 Initialisation of the Population*

Although the classical genetic algorithm generates the chromosomes randomly, this is not a feasible choice as the Manning's coefficient remains the same, or varies very little, for adjacent elements in a discretised domain. The reason is that the bed resistance depends on the physical conditions of the estuarine locations and, reaches with similar physical characteristics are expected to have similar values for the Manning's coefficient. Therefore, using randomly generated coefficients there is the danger of obtaining unrealistic simulations. In order to have chromosomes that represent the existing estuarine flow resistance more closely, the present GA scheme uses three different modes for initialising the population of chromosomes: the division of the chromosome in segments each of which has the same allele value (*zonation option*); the arrangement of the values of the genes in descending order (*scaling option*); and the inclusion, in the initial population, of chromosomes, of sets of Manning's coefficients inherited from similar cases (*use of case information*).

Thus, the initial population consists of 90% of chromosomes generated using the zonation and scaling options and 10% of individuals obtained from the selected cases. The use of case

---

<sup>1</sup> The estuary will be described in detail in Chapter 8 as it is one of the two study cases employed for testing the system and demonstrating the system's feasibility. The water elevations were observed during the spring tide on the 25<sup>th</sup> of April 1979. The simulation results showed in this chapter are obtained by using a one-dimensional Taylor-Galerkin finite element model.



**Figure 7.3.** Simulations of water surface elevations at Easthook, Upper Milford Haven estuary, UK, employing a manual calibration and a classical genetic algorithm optimisation schemes

information is limited to 10% to avoid premature convergence and ensure the necessary population diversity.

#### 7.4.1.1 Zonation

The initial population is based on the estuary divided into zones of influence each having a specific physical behaviour. This is due to the fact that the resistance to the flow changes with respect to the variation of the estuarine physical characteristics. Therefore, it is expected adjacent elements to have similar values for the Manning's coefficient. The division of the estuary is reflected in each chromosome by partitioning it into a number of segments corresponding to the estuarine zones. Based on this chromosome's structure the genes in the same segment are initially assigned with the same value, which is randomly generated.

Figure 7.4 shows a chromosome corresponding to a set of Manning's coefficients for an estuary domain discretised into 28 elements. The zonation option is applied through the division in 3 zones imposed over the element-based domain discretisation. The zonation is carried out from the estuary's head (zone A) towards the estuary mouth (zone C).



#### 7.4.1.2 Scaling

The observation that the flow resistance generally decreases towards the estuary mouth is also taken into account during the initialisation. Based on this evidence, the GA program sorts the alleles of chromosomes in descending order, with lower values for genes that correspond to elements of the domain allocated towards the estuary mouth (scaling option). In the example provided in figure 7.4, the values of the genes gradually decrease from zone A towards zone C, which include the estuary's head and mouth, respectively.

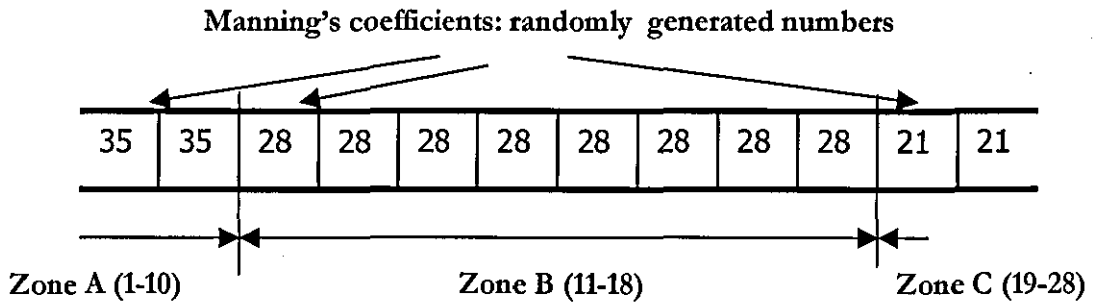


Figure 7.4. Example of chromosome initialised using the zonation and scaling options

#### 7.4.1.3 Use of Case Information

The last feature implemented for generating the initial population consists of seeding the cluster with appropriate Manning's coefficient series selected from the system's case-base. Based on the principle that similar problems should have similar solutions (S.J. Louis and J. Johnson, 1997), estuaries that do not significantly differ from one another should have similar sets of Manning's coefficients. The sets are preventively adapted to suit the discretisation scheme employed for the estuary under investigation.

#### 7.4.2 GA Operators

The GA operators of selection, crossover and mutation are also designed to incorporate concepts from the theory of estuarine calibration for the purpose of Manning's coefficient optimisation.

##### 7.4.2.1 Selection Mechanism

Starting from the initial population the subsequent generations are formed by selecting the chromosomes according to their fitness. The fitness of the chromosomes is calculated with

respect to the function (6.1), which consists of quantifying the discrepancy between the simulated results obtained using a specific sequence and the corresponding field measurements.

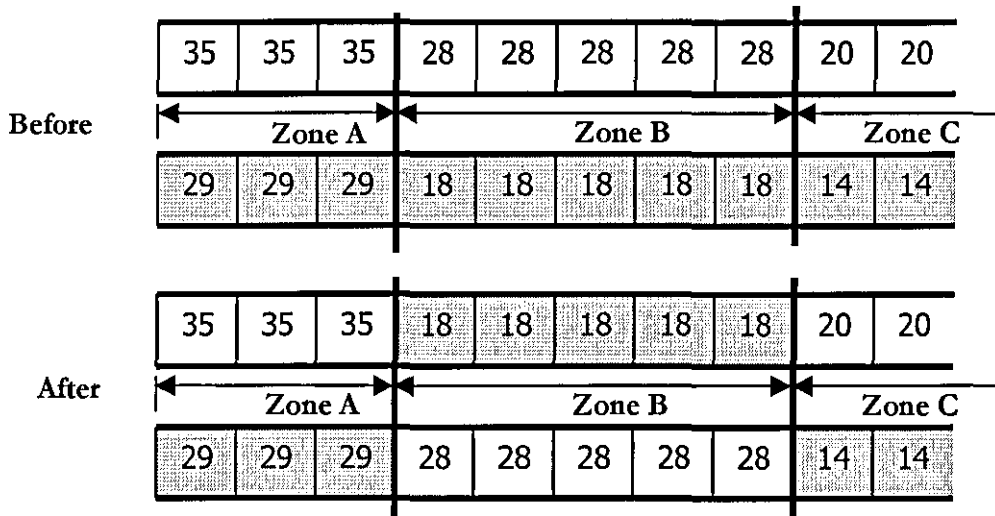
The selection procedure implemented consists of

- keeping 10% of the best chromosomes (i.e. with highest fitness values) in the next generation (i.e. elitist approach)
- and having the other 90% of the next generation randomly reproduced according to their fitness values (i.e. roulette wheel) and then transformed by crossover and mutation in order to introduce diversity in the population.

The application of the elitist logic stops the search to converge too quickly.

#### 7.4.2.2 Crossover of Estuary's Zones

The present scheme also contains different forms of the more common random mutation and crossover. The crossover and the mutation operators are devised to guide the search towards



**Figure 7.5.** Chromosomes before and after the crossover operation

chromosomes with a real physical meaning for estuarine calibration. Therefore, the traditional genetic operators are modified according to the previously made observations of adjacent genes representing adjacent elements and chromosome's segments corresponding to specific estuary's zones. The crossover operator swaps between chromosome segments which

correspond to specific estuary's zones. The number of cut-points in a chromosome is randomly chosen each time the crossover operator is applied (figure 7.5).

#### 7.4.2.3 Mutation of Adjacent Genes

The mutation operator implemented here is based on the concept that close elements are generally characterised by similar Manning's coefficients. Thus, the chromosomes are mutated by changing the value of a randomly chosen gene and its closest neighbours (figure 7.6).

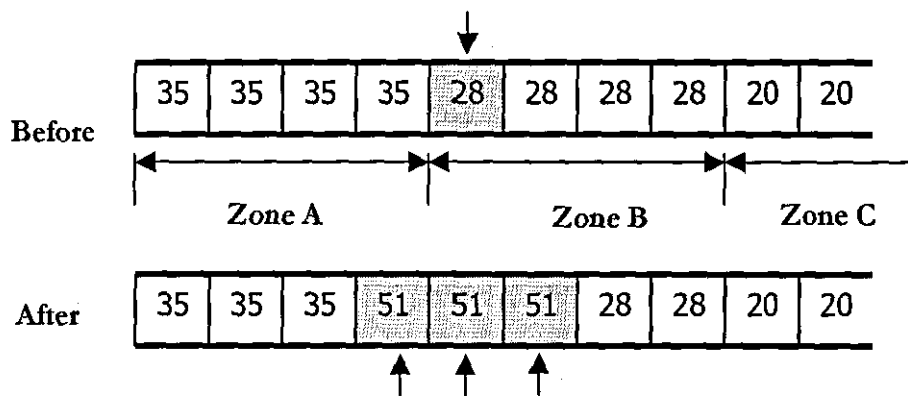


Figure 7.6. Chromosome before and after the mutation operation

#### 7.4.3 The Advantage of Using a Hybridised Genetic Algorithm

The water elevation profiles at Easthook station on the Upper Milford Haven estuary (figure 7.9) obtained through the calibration of the model by the hybridised genetic algorithm are here presented to demonstrate the validity of the designed scheme. It is evident that the simulation improves when the modified versions of initialisation, crossover and mutation are considered.

The different combinations of options are given in table 7.9 and the results of the different alternatives are shown in figure 7.7.

The use of zonation and scaling options and the mutation of adjacent genes generates sets of Manning's coefficients that gradually reduce the discrepancy between the simulated results and the observed data (figures 7.7a, 7.7b and 7.7c). Furthermore, the employment of the crossover of estuary's zones (figure 7.7d) reduces the computational time required for the generation of an optimum set of Manning's coefficients. Figure 7.7d proposes the simulated water surface elevations with and without the modified crossover operator (i.e. crossover of estuary's zones).

**Table 7.9.** The four combinations of GA options utilised to illustrate the advantages given by the hybridised genetic algorithm for estuarine calibration over the classical genetic algorithm scheme

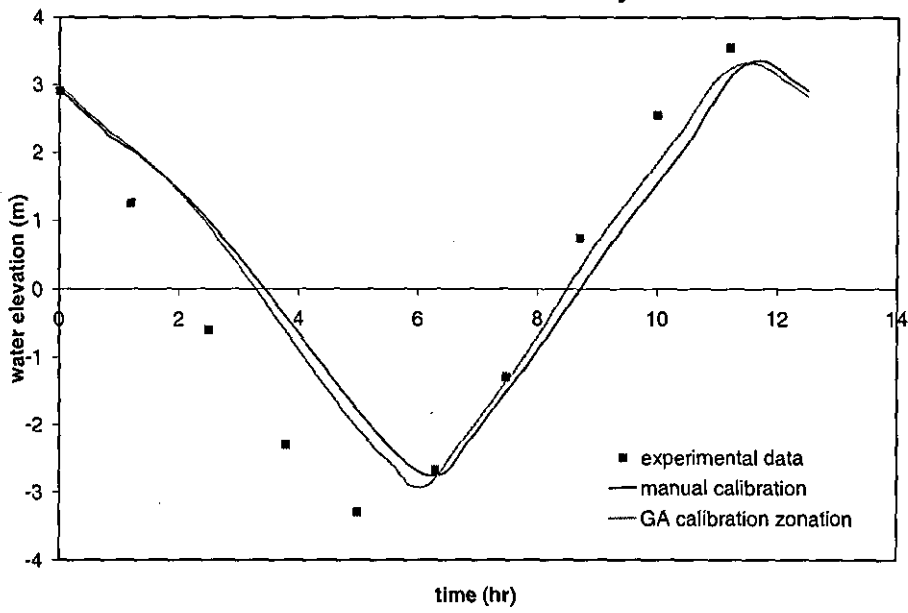
GRAPH	ZONATION	SCALING	CROSSOVER	MUTATION
a	✓		classical	classical
b	✓	✓	classical	classical
c	✓	✓	classical	mutation of adjacent genes
d	✓	✓	crossover of zones	mutation of adjacent genes

The two profiles overlap, as the set of Manning's coefficient adopted in both simulations is identical. However, the combination of options d) employs 1 less run than c) to generate the expected set of Manning's coefficients with a considerable saving on the computational time (~approximately one hour).

### 7.5 Conclusions

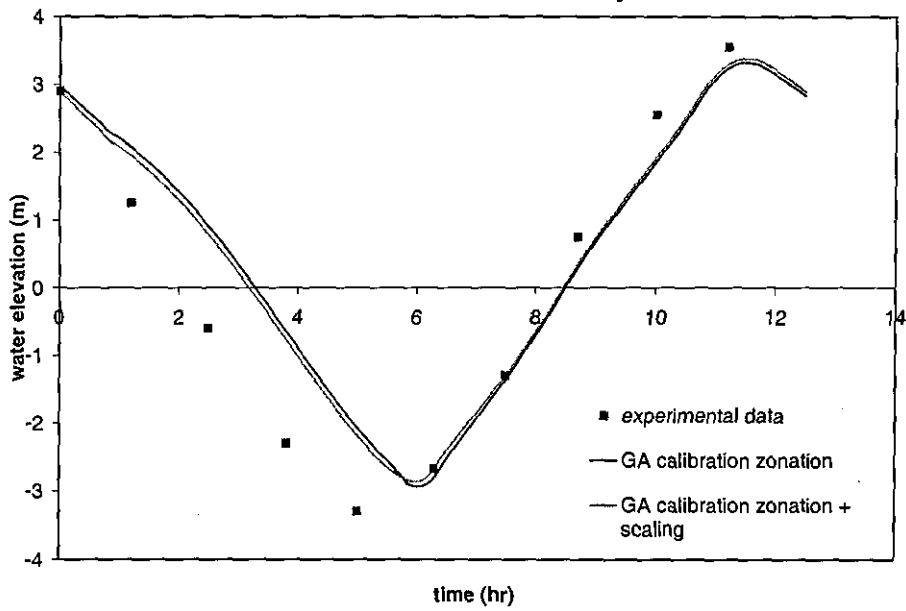
This chapter describes the inside of CBEM and in particular how the case description, retrieval and adaptation are structurally implemented. The knowledge is organised to provide a specialised system able to carefully evaluate the information given by the user. Different sets of rules are employed by the CB and GA modules, constructed with respect to the principles of the case-based reasoning process and the genetic algorithm theory.

Easthook - 28.2 km from the estuary mouth

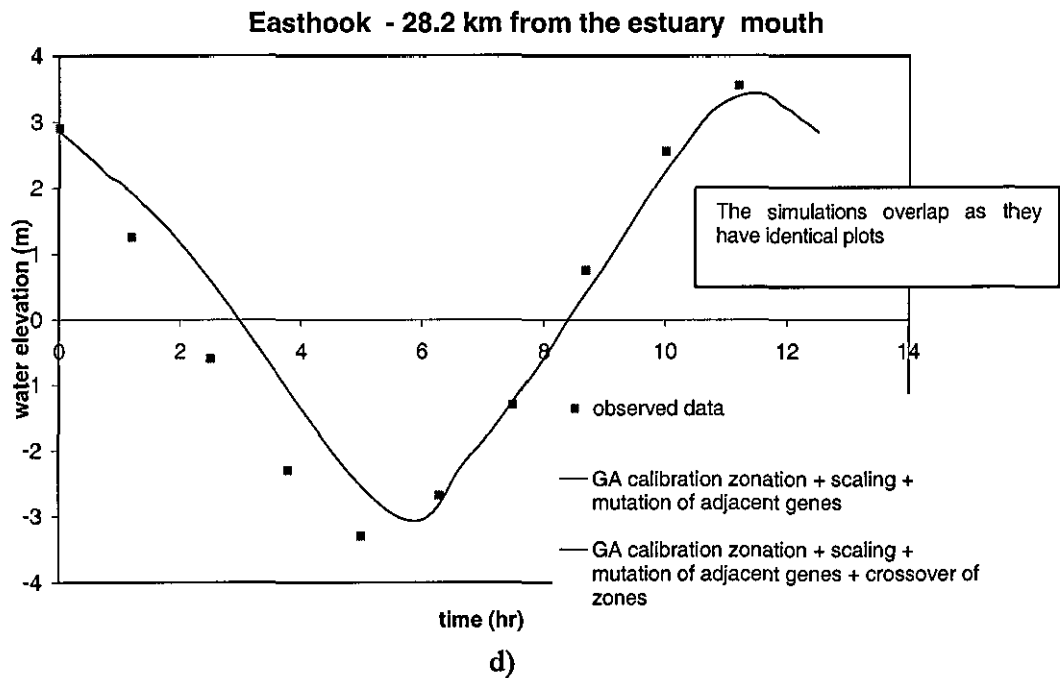
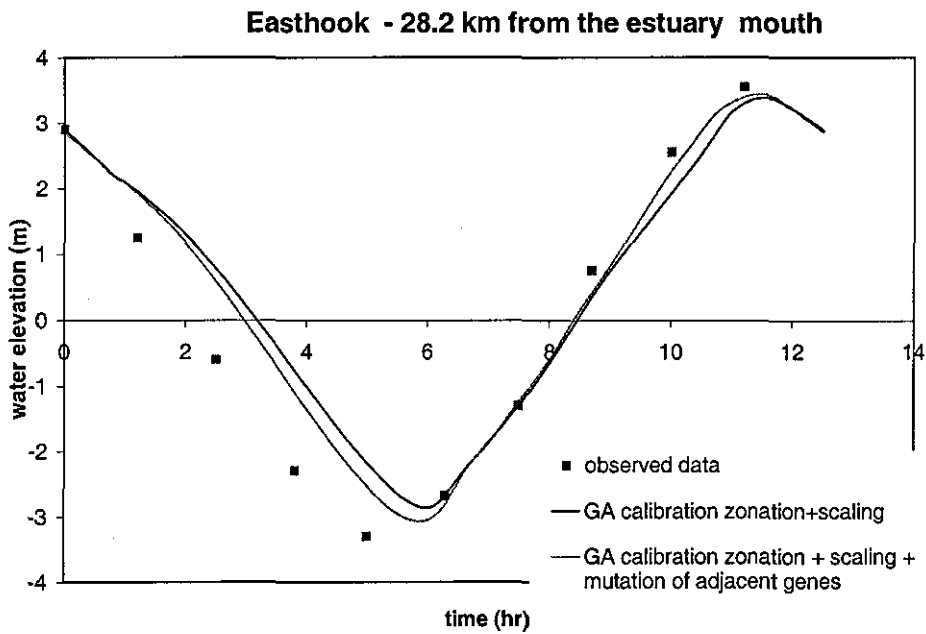


a)

Easthook - 28.2 km from the estuary mouth



b)



**Figure 7.7.** a), b) c) and d) Simulations of water surface elevations at Easthook, Upper Milford Haven estuary, UK, obtained for different option combinations (table 7.9) included in the hybridised genetic algorithm optimisation scheme

## CASE STUDIES

### 8.1 Introduction

This chapter aims to run through 2 case studies to show CBEM at work. The system is employed for finding an appropriate model to simulate the salt intrusion in the Upper Milford Haven estuary, in Wales, and the Tay estuary, in Scotland. The following discussion has a twofold scope: proving the system's ability to provide the user with a critical and correct estimation of the available model strategies and the robustness of the designed genetic algorithm.

The case-base of CBEM currently contains 12 studies about five estuaries: Conwy, Fal, Upper Milford Haven, Tay and Tees (table 8.1). Each estuary is associated with a number of models for simulating different problems (i.e. salt intrusion, salinity distribution, pollutant distribution and flood routing) for different investigative purposes.

The layout of this chapter is as follows: the specific problem is described in detail in §8.2 while the model employed is illustrated in §8.3. The two cases are given in sections §8.4 and §8.5 respectively, presenting the system's output for phases of case description, retrieval and adaptation.

### 8.2 Salt Intrusion

CBEM is employed to search for an appropriate model to investigate the axial salinity distribution (i.e. salt intrusion) of the Upper Milford Haven and Tay estuaries.

#### *8.2.1 Problem Description*

Salinity is probably the most important variable that characterises the estuarine environment. Studying salinity and its transport along an estuary channel helps to understand the mechanisms governing the estuarine water circulation and the mixing processes (K.R. Dyer, 1997). Not only salinity affects the hydrodynamics and the sediment movement within an estuary, but also it is responsible for the life of the estuarine organisms (i.e. fauna and flora).

Any variation in salt concentration can cause substantial changes of the ecosystem. Finally, as the physical behaviours of salt and most polluting solutes are expected to be similar, modelling salt intrusion provides an insight into the transport of pollutants within the estuary from the sea or the near coast (J.H. Bikangaga, 1993).

The penetration of salt into an estuary is due to the mixing of the fresh water from river discharge and sea water by the tidal action. The overall flow regime is mainly affected by the topography and geometry of the estuary and the magnitude of parameters such as river flow, tidal amplitude and wind force.

**Table 8.1.** Cases currently included within the CBEM's case-base

n.	estuary	problem	dimension	numerical method	time	purpose	accuracy
A	FAL	salt intrusion	1-d network	finite element	low	water quality	moderate
B	CONWY	salt intrusion	1-d	finite difference	low	management tool	moderate
C	TEES	salt intrusion	2-d avg on width	finite element	moderate	water quality	high
D	TEES	salt intrusion	1-d	finite difference	low	management tool	high
E	TAY	salt intrusion	2-d moving	finite element	high	research	high
F	TAY	salt intrusion	2-d avg on depth	finite element	high	management tool	moderate
G	CONWY	salinity distribution	2-d avg on depth	finite element	moderate	water quality	high
H	TEES	salinity distribution	1-d	finite element	low	management tool	overall
I	FAL	salinity distribution	2-d avg on depth	finite element	high	research	high
J	FAL	salinity distribution	2-d avg on depth	finite difference	low	management tool	moderate
K	UP.MILFORD HAVEN	pollutant distribution	2-d avg on depth	spectral method	high	water quality	high
L	UP.MILFORD HAVEN	flood routing	2-d avg on depth	finite element	high	research	high

Salt intrusion means the determination of the axial salinity profile within the estuary at some specific moments during the tidal period. For instance, salt intrusion is generally determined during high and low water at the estuary mouth. As previously mentioned (chapter 2), the salt concentrations along any direction in an estuary are dependent on the estuary's



hydrodynamical behaviour. The techniques generally employed to estimate saline penetration can be analytical, numerical or based on laboratory experiments.

Numerical models for simulating estuarine salt intrusion consist of solving, numerically, the equations for continuity and momentum of water and mass continuity of salt. This solution gives the variation of salt concentration between the mouth and the head of an estuary under different scenarios. Based on their specifications, numerical models provide predictions with different degree of accuracy and cost-effectiveness. The selection of a specific solution scheme must be based on the complexity of the estuarine hydrodynamics and geometry and the precision required from the model results.

### *8.2.2 Search Query Model*

The search query is identical for the two estuaries used in this project. Addressed using the model selection screen (chapter 6), it consists of selecting "salt intrusion" as the problem and "management tool" as the purpose of the investigation. The cut-off value is 0.5 for both applications.

### **8.3 Numerical Model for Simulating Salt Intrusion**

The same model is selected from those retrieved by CBEM for simulating the salt intrusion in the Upper Milford Haven and Tay estuaries. This helps to show the ability of the developed system to estimate correctly which model is more suitable for simulating in a given estuary a specific estuarine phenomenon. Moreover, it permits to demonstrate the effectiveness of the genetic algorithm routine implemented to support the parameter calibration in the retrieved models.

Developed by J.H. Bikangaga (J.H. Bikangaga, 1993) the model is a one-dimensional Taylor-Galerkin finite element numerical scheme for simulating estuarine salt intrusion. The model was employed by its author to simulate the hydrodynamics of estuaries such as Upper Milford Haven, Tay, Tees and Fal estuary.

### 8.3.1 Model Description

The hydrodynamics and the salinity distribution are represented in this model by one-dimensional shallow water and convection-dispersion equations, respectively. These are expressed as the following equations of continuity (8.1) and motion of water (8.2),

$$\frac{\partial Q}{\partial x} + B \frac{\partial h}{\partial t} = q \quad (8.1)$$

$$\frac{\partial h}{\partial x} + \frac{1}{gA} \frac{\partial Q}{\partial t} + \frac{1}{gA} \frac{\partial}{\partial x} \left( \frac{Q^2}{A} \right) + \frac{|Q|Qn^2}{A^2 H^{4/3}} + \frac{0.75H}{2000+1.5S} \frac{\partial S}{\partial x} = 0 \quad (8.2)$$

where, B, A, h, Q, H, S and q are, respectively, breadth of water surface, cross-sectional area, water surface elevation with respect to a datum, cross-sectional estuary discharge, hydraulic depth, salinity concentration and fresh water inflow at the estuary head. The spatial and temporal lengths are denoted by x and t, respectively, whereas, g and n represent the gravitational constant and the Manning's friction coefficient, respectively.

The convection-dispersion equation for salt transport is expressed as,

$$\frac{\partial(AS)}{\partial t} + \frac{\partial}{\partial x}(AUS) - \frac{\partial}{\partial x} \left( AD_x \frac{\partial S}{\partial x} \right) = 0 \quad (8.3)$$

where, U and  $D_x$  are depth averaged velocity and dispersion coefficient, respectively.

In case of a branching system, the conservation of mass and momentum at a junction cannot be represented by equations (8.1) to (8.3) and, instead, the following equations should be used.

For continuity of water,

$$\sum_{i=1}^M Q_i = A_j \frac{\partial h_j}{\partial t} \quad (8.4)$$

for water surface elevation,

$$h_I + \frac{\alpha_I \left(\frac{Q}{A}\right)^2}{2g} = h_{II} + \frac{\alpha_{II} \left(\frac{Q}{A}\right)^2}{2g} = \dots = h_M + \frac{\alpha_M \left(\frac{Q}{A}\right)^2}{2g} \quad (8.5)$$

where  $\alpha$  represents the energy coefficient related to the surface area of the junction,

and for conservation of mass, assuming complete mixing at the junction

$$\sum_{I=1}^M (QS)_I = A_J \frac{\partial(h_J S_J)}{\partial t} \quad (8.6)$$

where,  $A_J$ ,  $h_J$ ,  $S_J$  are the junction area and elevation and concentration at the junction, respectively.

When the junction surface area is small, as assumed by Bikangaga and Nassehi (1995), the above equations are simplified for the momentum transfer and the mass continuity to, respectively,

$$h_I = h_{II} = \dots = h_M \quad (8.7)$$

$$S_I = S_{II} = \dots = S_M \quad (8.8)$$

where,  $M$  represents the number of branches at the junction  $J$ .

The described equations are discretised using the Taylor-Galerkin finite element method. This scheme has an inherent ability to cope with complex geomorphology and boundary conditions in natural water systems. The detailed derivation of the working equations of this model has been presented elsewhere (J.H. Bikangaga and V. Nassehi, 1995) and is not repeated here.

### 8.3.2 Input Data

The model is defined by giving a set of input data such as elemental and hydrographic data, boundary and initial conditions, Manning's and dispersion coefficients and the time step. The elemental data consists of the elements of the discretised domain and the nodal coordinates of the elements' vertices. The hydrographic data required are the time-dependent values of the

cross-sectional area  $A$ , hydraulic depth  $H$  and water surface breadth  $B$ , at each node along the estuary domain.

The fresh water flow upstream and the water surface elevations at the estuary mouth need to be specified as boundary conditions. The salinity variation at the mouth and zero salinities at tidal limits must also be supplied. Set of arbitrary values must also be given for  $h$ ,  $Q$  and  $S$  as initial conditions to start the solution algorithm.

In the simulation of the salinity distribution in both estuaries, while the sets of Manning's friction coefficients are obtained through the designed genetic algorithm scheme, the dispersion coefficient is assumed to be already known. In case of the Upper Milford Haven estuary the dispersion coefficient is estimated according to the empirical expression proposed by Thatcher and Harleman (1972) as,

$$D_x = 64K_1n\left(\frac{Q}{A}\right)H^{5/6} \quad (8.9)$$

where  $K_1$  is an empirical coefficient that depends on the estuary type.

The semi-empirical expression (8.9) is found to be suitable for the Upper Milford Haven. However, the results provided by using it for the Tay estuary are very inaccurate. Therefore, for the Tay estuary, the variation of the dispersion coefficient along the estuary channel is computed through the interpolation of the measured salt concentration values collected during a spring and neap tide (V. Nassehi and D.J.A. Williams, 1986).

### 8.3.3 Model Output

The output data consists of the tidal variations at specific nodal points of

- Water surface elevations
- Discharges
- Water surface breadths
- Axial salinity profiles at different times

## 8.4 Upper Milford Haven Estuary

### 8.4.1 Introduction

Milford Haven is a “ria” type estuary system (A.L. Buck and A.C. Davidson, 1997) located in the south-western region of Wales, U.K. It consists of the Daucleddau river which branches at Picton Point into the Western and the Eastern Cleddau rivers. In this study the upper part of the estuary northwards Carr Jetty, the Upper Milford Haven estuary (figure 8.1), is examined.

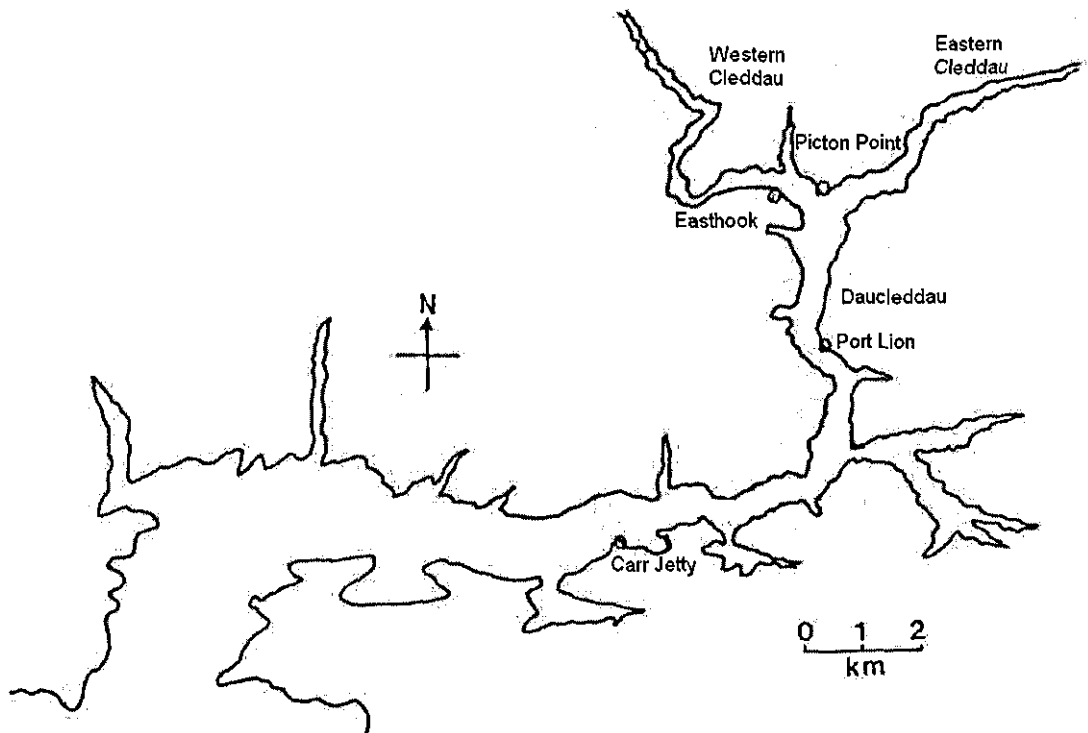


Figure 8.1. The Upper Milford Haven Estuary, Wales, UK

The estuary has been presented in many studies focusing on its physical behaviour and the risk assessment of a potential contamination of its waters (A. Nelson-Smith, 1965; V. Nassehi and D.J.A. Williams, 1986; P. McLaren and D.I. Little 1987; J.H. Bikangaga and V. Nassehi, 1995; S. Passone et al. 2002).

The estuary is a narrow branched tidal waterway and the ratio of the width of the estuary to its depth is moderate and almost constant, contributing with 110 km of shoreline. Furthermore, as pointed out by Nassehi and Williams (1986), the difference between the salinity levels at the surface and the bottom is negligible ( $\sim 0.3\%$ ). Hence, the estuary is considered to be a well

**frmMainEstuary**

**ESTUARY**

name:  ID:  county:  latitude:

Physical characteristics

geomorphological type:  tidal range:

total area (ha):  intertidal area (ha):  max channel bank area (ha):

channel length (km):  valley length (km):  meandering/sinuosity:

min width (m):  max width (m):  bed:

avr. width (m):  avr. depth (m):  avr. width to depth ratio:

Hydrographic characteristics

tot. freshwater inflow (m<sup>3</sup>/s):  tidal volume (m<sup>3</sup>):  salinity:

salinity limit (km):  averaged velocity (m/s):  tidal period (hr):

inlets: How many inlets?   coriolis effect  wind

comments:

Record:  of 8

**Figure 8.2.** Description of the Upper Milford Haven estuary, Wales, UK

mixed system as far as the salinity is concerned. The Upper Milford Haven estuary is also macrotidal and the tidal dynamics dominates the estuarine flow. Any secondary effect induced by fresh water discharge and, the wind and Coriolis forces can be neglected. These conditions support the notion that the cross-sectional salinity distribution within the Upper Milford Haven is uniform. This also allows the representation of the longitudinal mixing by one-dimensional computational scheme.

The hydrodynamics of the Upper Milford Haven is significantly influenced by branching of the main channel at Picton Point to the Eastern and Western Cleddau. For example, the geomorphology of the estuary affects the phase difference between tidal waves propagating in the Eastern and Western Cleddau rivers. Phase differences between the main channel and the branches are also observed. However, as the Upper Milford Haven is a typically short British

branching estuary, the phase lag between the Western and Eastern Cleddau and the Dauceddau channel is limited to approximately one hour at most for a low neap tide. Nevertheless, these phase differences may significantly influence the longitudinal water density distribution and consequently the salt transport along the branches.

#### 8.4.2 Case Description

In order to identify a model within the CBEM's case-base that is suitable to simulate the salt intrusion in the Upper Milford Haven estuary, the user must provide CBEM with the necessary information to guide the retrieval process. The user specifies the physical and hydrological characteristics of the Upper Milford Haven estuary through the case description form (figure 8.2). Once satisfied with the definition of the new case, the user clicks the button APPLY MODEL that opens the model selection screen.

The screenshot shows a window titled "frmselectioncriteria" with a close button in the top right corner. The main title is "Model Selection". Below the title, there are several input fields and checkboxes:

- Estuary:** A text box containing "JP.MILFORD HAVEN".
- problem:** A dropdown menu showing "salt intrusion".
- purpose:** A dropdown menu showing "management tool".
- cut-off:** A text box containing "0.5".
- Approximation:** A section with four checkboxes:
  - coriolis effect
  - wind effect
  - branching
  - bed slope
- Buttons:** "OK" and "Cancel" buttons at the bottom.

Figure 8.3. Problem statement

#### 8.4.3 Case Retrieval

As the problem to be investigated (i.e. salt intrusion) for the Upper Milford Haven estuary is specified through the model selection screen (figure 8.3), CBEM is ready to start the retrieval process.

In the first half of the retrieval process the estuary description of each past case is compared to the description of the Upper Milford Haven estuary given by the user. CBEM then returns a list of 4 estuaries that have initial similarity scores greater than 0.5 (figure 8.4).

ESTUARY	SIMILARITY RATING	NUMBER OF MODELS	
TEES	0.690	2	<input checked="" type="checkbox"/>
FAL	0.886	1	<input checked="" type="checkbox"/>
TAY	0.534	2	<input checked="" type="checkbox"/>
CONWY	0.722	1	<input checked="" type="checkbox"/>

Record: 14 of 4 (Filtered)

**Figure 8.4.** Set of cases retrieved from the system's case-base ranked with respect to their similarity to the Upper Milford Haven estuary

At this point the user is asked to select those cases that he/she would like to see through the second phase of the retrieval process. For the purpose of the present discussion, all four retrieved estuaries are admitted. The second half of the process consists in the computation of the second similarity score that quantifies the appropriateness of each retrieved model with respect to the specified purpose of the investigation (i.e. "management tool"). Furthermore, the retrieval process uses the criteria of exclusion and preference to eliminate or prefer the cases whose model dimension is considered inadequate or very adequate, respectively, for the simulation of the salt intrusion problem.

At the end of the retrieval process the selected case are classified with respect to both the first and second similarity scores (figure 8.5). However, only 3 cases of the initially retrieved 6 are in



this list. Cases B), D) and E) (table 8.1) are withdrawn from the set. In particular, these cases are eliminated because of the principle of exclusion, which tells,

1. "If the estuary has inlets and at least one of the inlets is a "branch" then eliminate all 1-D models".
2. "In case the problem is salt intrusion all 3-D and 2-D moving boundary models must be not considered".

Furthermore, preference is expressed for the 1-D network model.

ESTUARY	DIMENSION	ACCURACY	ESTUARY SIM	MODEL SIM
TAY	2-D avg on depth	moderate	0.534	0.455
TEES	2-D avg on width	high	0.690	0.482
FAL	1-D network	moderate	0.886	0.882

Record: 14 of 3 (Filtered)

Figure 8.5. Cases retrieved after the second phase of the retrieval process

As evident from figure 8.5, case A) (i.e. "FAL, 1-D network, estuary sim. = 0.886, model sim. = 0.882") has the highest first and second similarity score. The Fal estuary is, among those estuaries in the system's memory, the most similar to the Upper Milford Haven in terms of physical and hydrological characteristics (more information about the Fal estuary can be found in Appendix A, which contains the description of the cases included in CBEM case-base). The model dimension, 1-D network matches quite well the geometry and the geomorphology of Upper Milford Haven. Furthermore, as evident from table 8.1, its model characteristics (i.e. "low" simulation time, "moderate" accuracy and "water quality" investigation purpose) make case A) to be selected for providing the model that more appropriately simulates the longitudinal salt transport into the Upper Milford Haven estuary.

The selected model is the one-dimensional Taylor-Galerkin finite element scheme for branching estuaries previously described in section §8.3.

#### *8.4.4 Case Adaptation – Genetic Algorithm Optimisation*

The adaptation of the selected model is provided by the GA module, which finds an appropriate set of Manning's coefficients to represent the frictional resistance of the Upper Milford Haven channel to the flow. The calibration of the model is carried out by using the field measurements collected during the spring tide of the 25<sup>th</sup> April 1979. Finally, the calibration provided by the GA module is validated by employing the model to simulate the salt intrusion during the neap tide of the 2<sup>nd</sup> May 1979, while the set of Manning's coefficients utilised is the one previously selected by the genetic algorithm routine.

The effectiveness of the developed procedure is illustrated through the comparison of the water surface elevations and salinity profiles simulated for these spring and neap tides using the genetic algorithm based calibration, and the results obtained under the same circumstances by J.H. Bikangaga (J.H. Bikangaga, 1993), who manually calibrated the model.

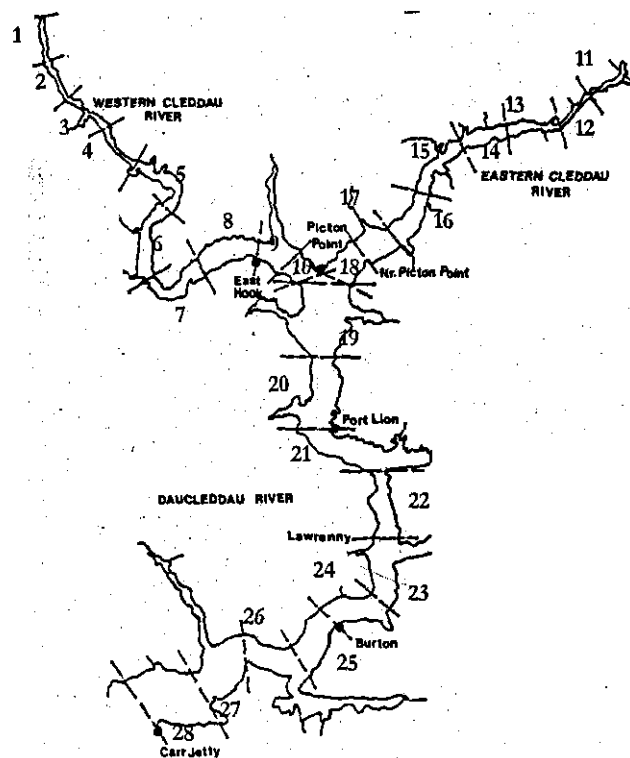
##### *8.4.4.1 Application of the Model to the Upper Milford Haven Estuary*

The estuary domain is discretised into 28 linear one-dimensional finite elements (figure 8.6). The nodes coincide with the sampling stations considered in the 1977 survey of the estuary by Nassehi (V. Nassehi, 1981).

Finite element locations and sizes are the same as those utilised by J.H. Bikangaga in his modelling study of the estuary (J.H. Bikangaga, 1993).

The necessary hydrographic data are obtained from the mentioned survey. These are tabulated for all stations, at interval of 0.2 m from low water with respect to a datum O.D. Newlyn. The fresh water elevations and measured salinities at Carr Jetty and the water discharge and zero salinities at the tidal limits observed during the spring tide of 25<sup>th</sup> April 1979 are used as boundary conditions. In equation (8.9) used for calculating the dispersion coefficient  $K_1$  is assumed to be equal to 700.

Fresh water inflow on the 25<sup>th</sup> April 1979 was 5.75 m<sup>3</sup>/s and 5.15 m<sup>3</sup>/s for Western Cleddau and Eastern Cleddau rivers, respectively.



**Figure 8.6.** Finite element discretisation of the Upper Milford Haven estuary (a Manning's coefficient is assigned to each element in the domain)

It must be noticed that in order to have an acceptable comparison between the model performances obtained using the manual and GA based calibrations, both model applications should be specified under identical conditions. This means that except for the set of Manning's friction coefficients, the input data employed in the model calibrated by GA routine, is the same as those utilised by J.H. Bikangaga for modelling the salt intrusion in the Upper Milford Haven estuary (J.H. Bikangaga, 1993).

#### 8.4.4.2 Fitness Function Definition

The calculation of the fitness function is based on the minimisation of the discrepancy between the water surface elevations simulated by the model and observed at Easthook and Picton Point stations during the tide of 25<sup>th</sup> April 1979 (figure 8.6). These locations are selected because the model is very sensitive to the hydrodynamic conditions at these two stations. Thus, the fitness function (6.1) is then expressed for the Upper Milford Haven case as

$$\rho(H_s, H_m) = \left[ \sum_{j=1}^2 \sum_{n=1}^{10} \left( (h_s)_j^n - (h_m)_j^n \right)^2 \right]^{1/2} \quad (8.10)$$

where the total number of samples, collected at each station during the tidal period, is equal to 10.

#### 8.4.4.3 Manning's Friction Coefficient Calibration

The model calibration operated by the designed genetic algorithm scheme (§7.4) is executed with the population set to 30 individuals and the rate of crossover and mutation equal to 0.5 and 0.01, respectively. The GA based calibration is carried out for 15 generations. The estuary is divided into three main zones corresponding to the two branches (i.e. Western Cleddau and Eastern Cleddau rivers) and the main channel (the Decleddau reach) (figure 8.6). Based on this partition of the estuary, the chromosome population is initialised using the zonation and the scaling options. These chromosomes are then transformed by the modified mutation and crossover operators. Only one set of Manning's coefficients from the case-library is included in the initial population. This is the set of parameters employed for this particular model to simulate the salt intrusion in the Fal estuary.

The set of Manning's coefficients selected by the genetic algorithm scheme after 15 generations is presented in table 8.2. which also contains the set of values for the parameter obtained by Bikangaga (J.H. Bikangaga, 1993) through the manual calibration of the model. The values of both sets are well within the expected range for an estuary modelled by using a one-dimensional scheme.

The simulated water surface elevations, generated using the set of Manning's coefficients selected by the GA module, are presented for the stations of Port Lion, Picton Point and Easthook in figure 8.7. The simulations at these locations show that the tidal wave has a regular shape within the main channel becoming gradually distorted as it propagates upstream from the junction into the branches. This is consistent with the geomorphological characteristics of the Upper Milford Haven estuary. The observed and simulated water surface elevations, obtained by the manual optimisation of the model, at the described locations are also shown in figure 8.7. The comparison between the observed data and the simulated water

Table 8.2. Manning's friction coefficients for the Upper Milford Haven estuary

element	GA Calibration	Manual Calibration
1	0.04	0.02
2	0.022	0.02
3	0.022	0.02
4	0.017	0.02
5	0.017	0.02
6	0.017	0.02
7	0.017	0.02
8	0.013	0.02
9	0.012	0.02
10	0.013	0.02
11	0.013	0.02
12	0.013	0.02
13	0.013	0.02
14	0.013	0.02
15	0.013	0.02
16	0.013	0.02
17	0.013	0.02
18	0.013	0.02
19	0.013	0.02
20	0.013	0.02
21	0.013	0.018
22	0.013	0.018
23	0.013	0.018
24	0.013	0.018
25	0.013	0.018
26	0.013	0.018
27	0.013	0.018
28	0.013	0.018

surface elevations for the manual and the GA based calibrations shows that the model using the GA module yields a better performance. The superiority of the GA based calibration over the trial and error optimisation is also demonstrated by considering the time necessary to carry out these two processes. In general, manual calibration of a model requires two weeks to one month (working time) while the GA based calibration takes 10 hours of CPU time on a shared SUN workstation (i.e.sun-cc211).

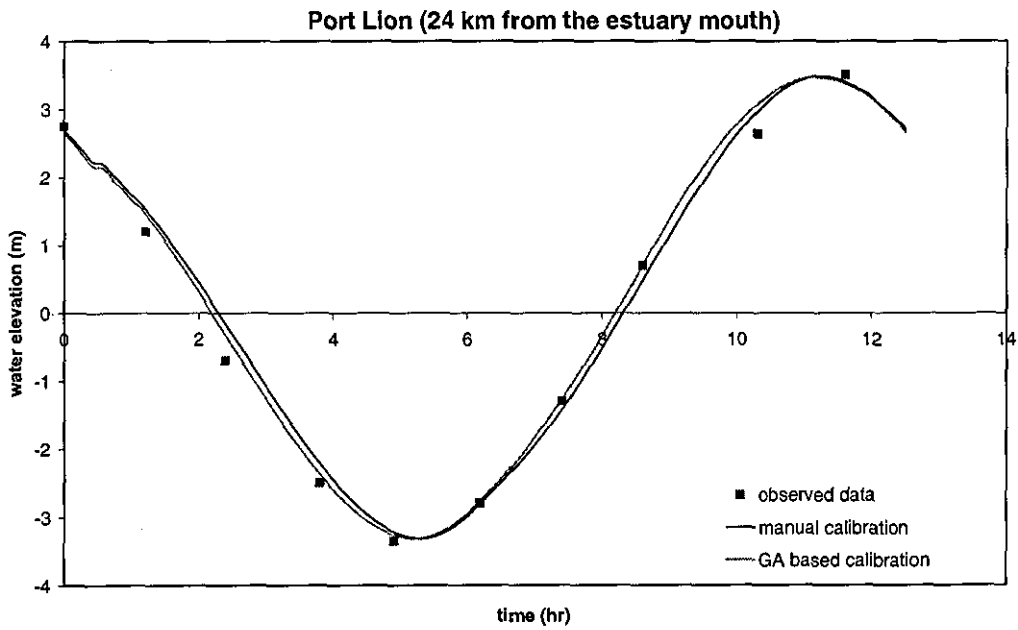
Furthermore, results for the salt intrusion (figure 8.8) also confirm the improvement for the model accuracy obtained through the GA based calibration. Although the dispersion coefficients are identical for both applications, the salinity profile modelled using the set of Manning's coefficient selected by the GA module is in better agreement with the measured values than the axial salt distribution given by the model manually calibrated.

However, figure 8.8 also shows that the longitudinal salinity distribution has an unexpected peak in correspondence with the junction. This anomaly may be due to the fact that there is a time lag of approximately one hour between the observed and simulated low water at the junction (figure 8.7), which can explain the sudden variation of the predicted salt concentration between the main channel and the Western Cleddau branch.

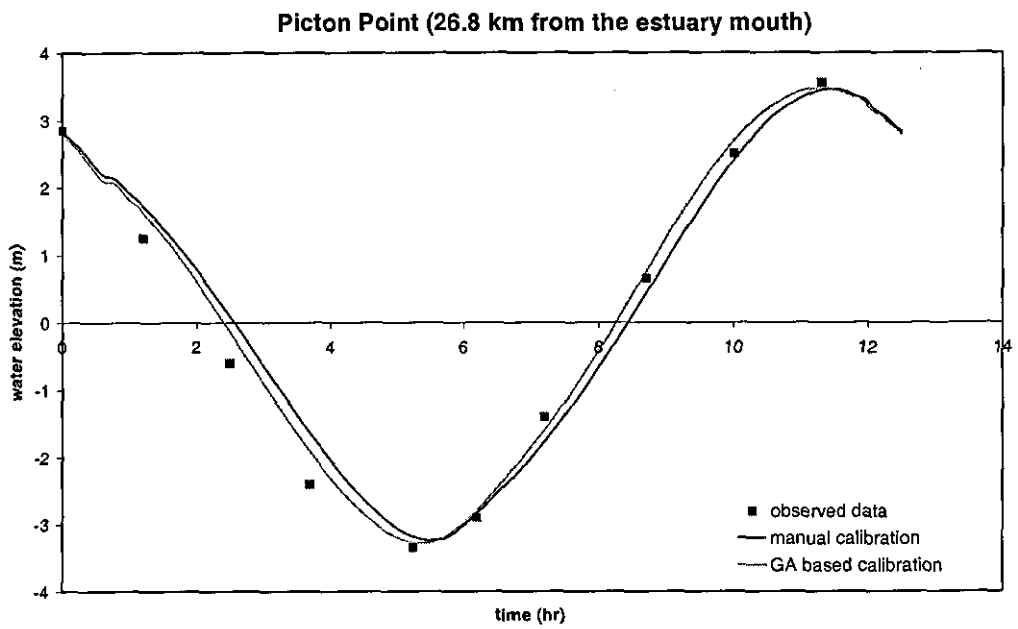
#### *8.4.4.4 Model Validation - Neap Tide*

The appropriateness of the set of parameters provided by the GA routine is verified using the same set to simulate the salt intrusion in the Upper Milford Haven estuary under different conditions. For this purpose, the validation is conducted for the neap tide of 2<sup>nd</sup> May 1979. The results obtained for the water surface elevations (figure 8.9) and the longitudinal salinity profile (figure 8.10) during this tide confirm the suitability of the set of values given by the genetic algorithm and, consequently, the validity of the developed GA based optimisation procedure.

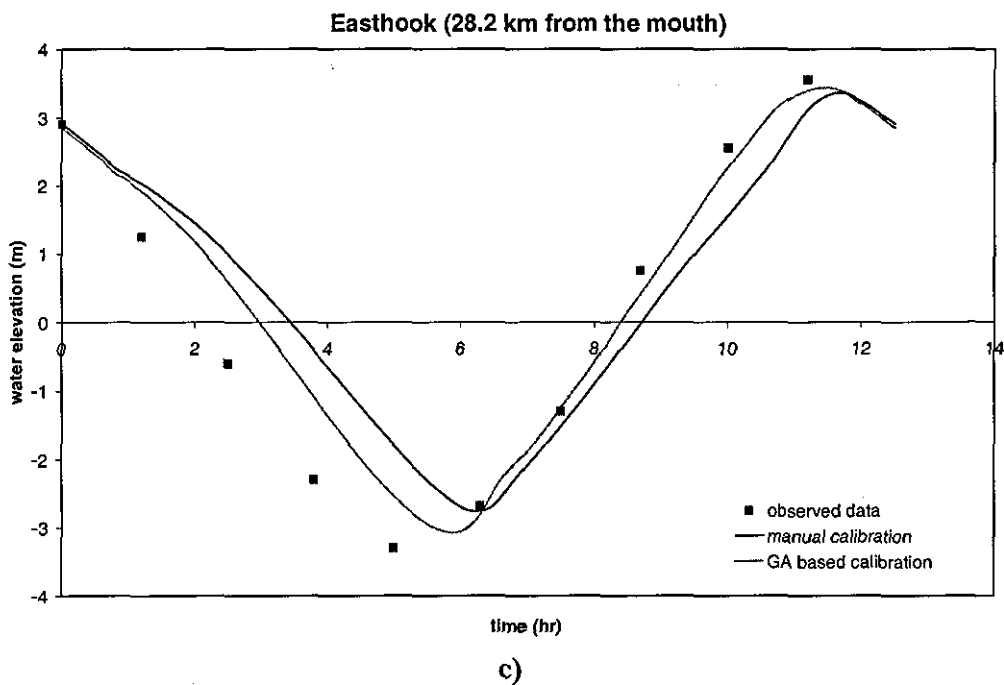
Furthermore, it must be noted that as for the simulation of the salt intrusion during the spring tide, the model outputs for the axial salinity distribution during the neap tide are also affected by the dispersion coefficient, which is identical for both the manually calibrated model and the model optimised through the genetic algorithm scheme.



a)

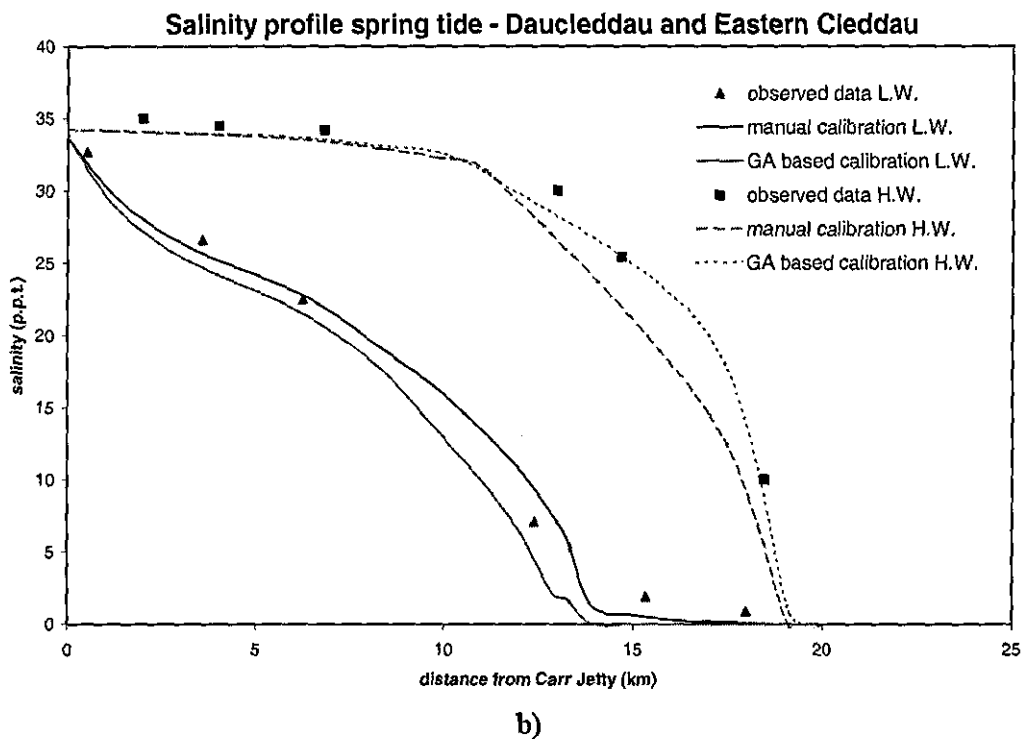
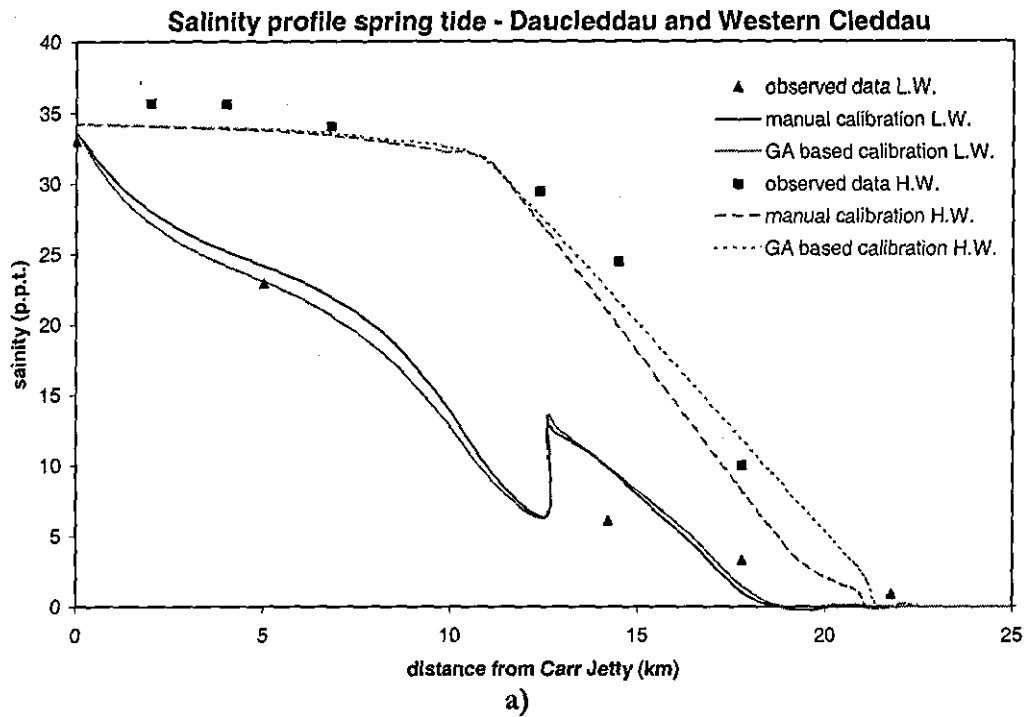


b)

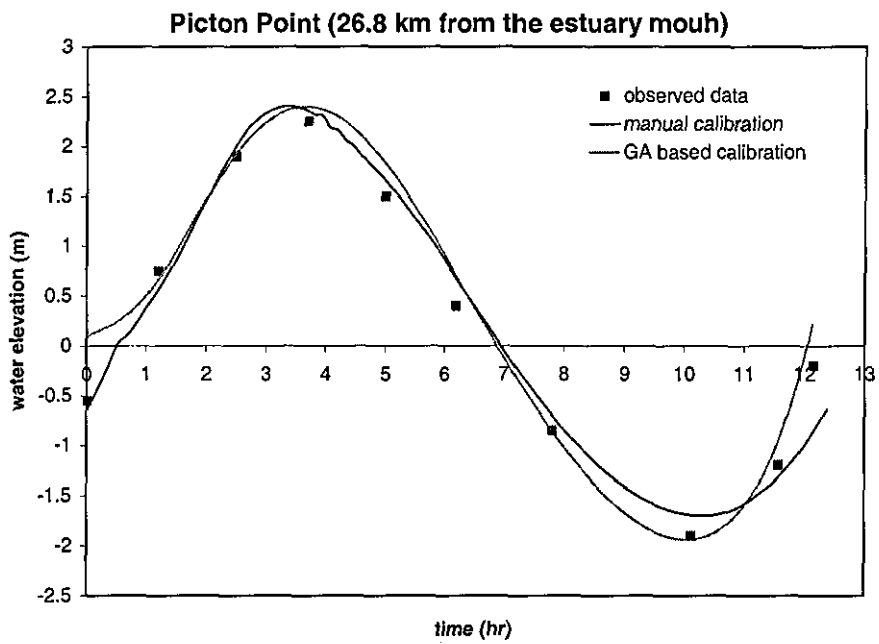
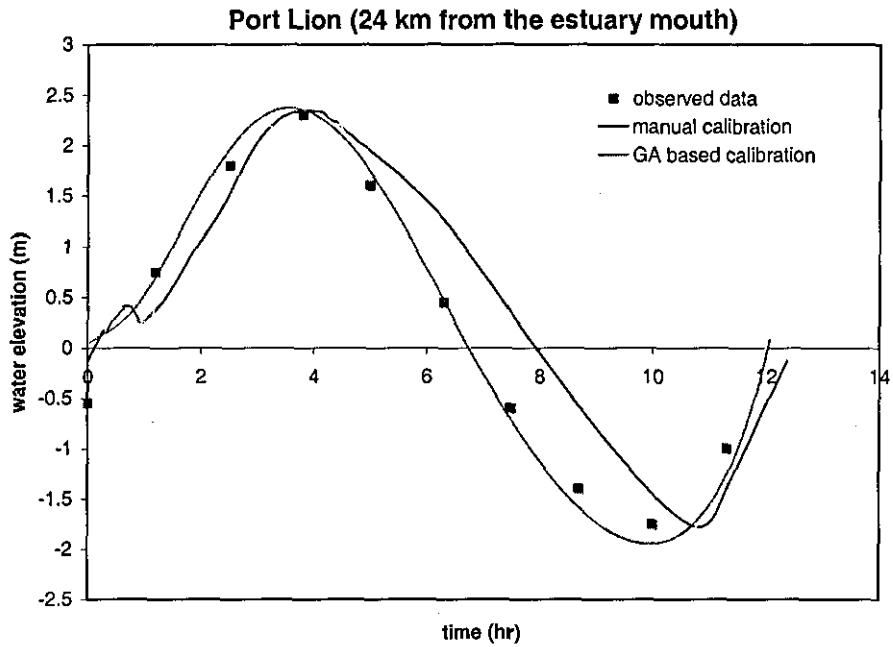


**Figure 8.7.** Simulation of the water surface elevations at a) Port Lion, b) Picton Point and c) Easthook during the spring tide of 25<sup>th</sup> April 1979

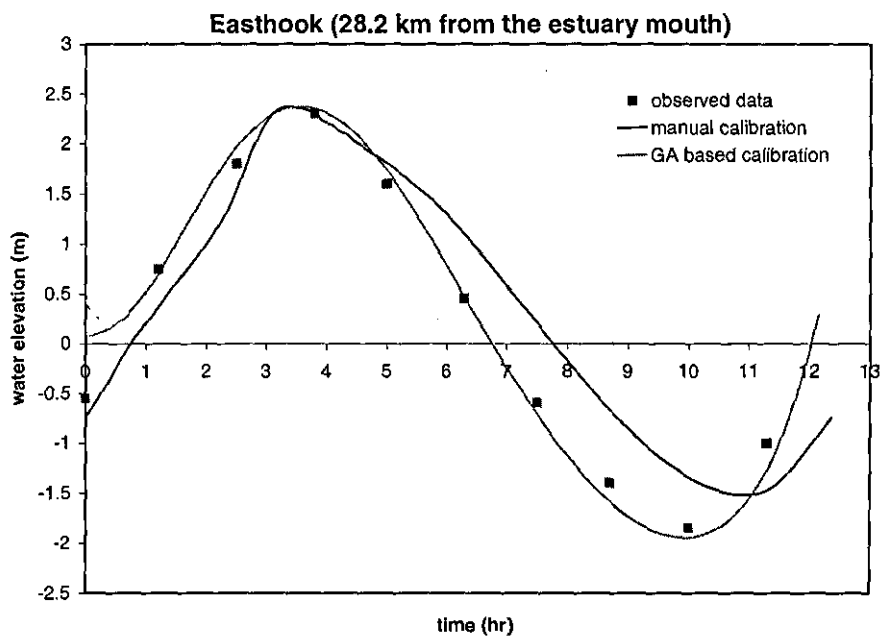




**Figure 8.8.** Salinity profiles - a) Dauceddau and Western Cleddau, and b) Dauceddau and Eastern Cleddau during the spring tide of 25<sup>th</sup> April 1979

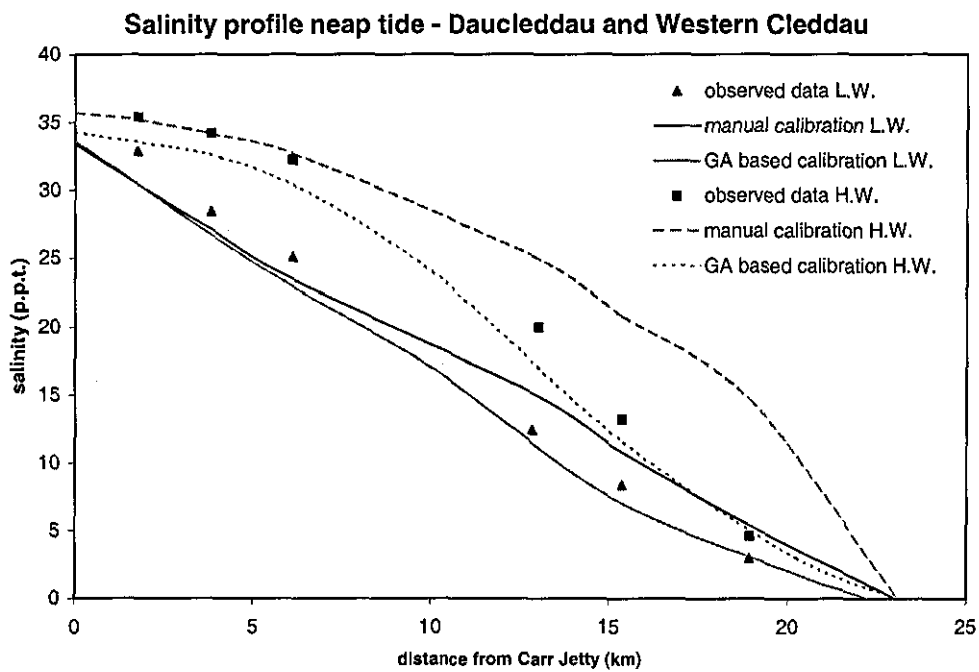


b)

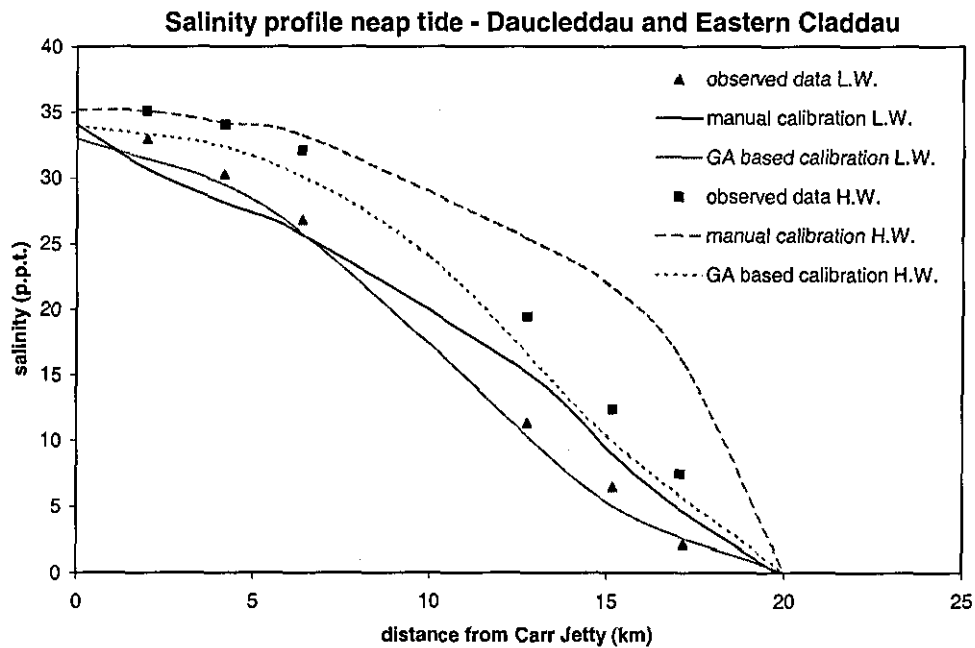


c)

**Figure 8.9.** Simulation of the water surface elevations at a) Port Lion, b) Picton Point and c) Easthook during the neap tide of 2<sup>nd</sup> May 1979



a)



b)

**Figure 8.10.** Salinity profiles - a) Daucleddau and Western Cleddau, and b) Daucleddau and Eastern Cleddau during the neap tide of 2<sup>nd</sup> May 1979

## 8.5 Tay Estuary

The second case study is about the Tay estuary in Scotland, UK.

### 8.5.1 Introduction

The Tay estuary in Scotland, UK, consists of the rivers Tay and Earn. It has the highest freshwater discharge of any British estuary (figure 8.11), which consequently affects the mixing between fresh and salt water over a significant length. The estuary is 54 km long and the front of salt intrusion extends for 45 km while the limit of tidal motion is near Perth, 50 km from the estuary mouth.

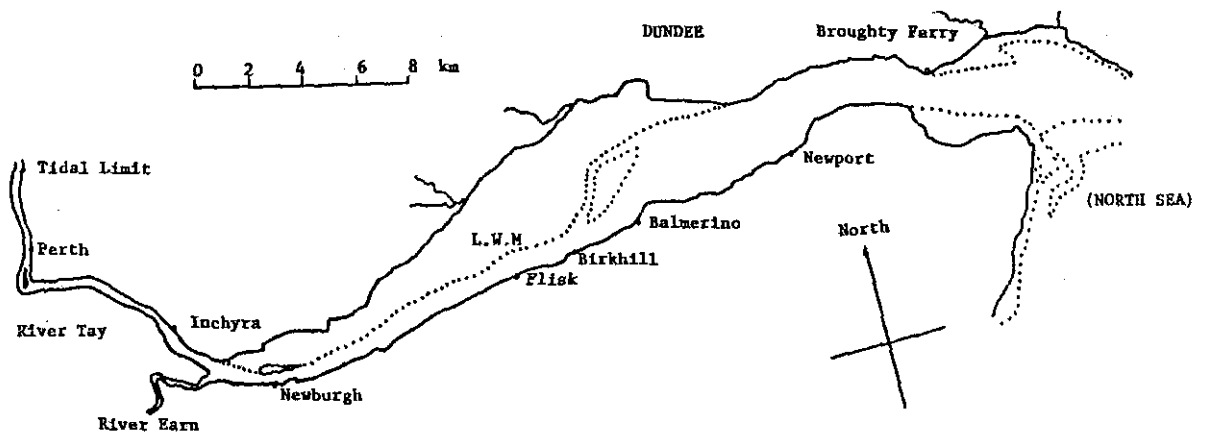


Figure 8.11. The Tay Estuary, Scotland, UK

The estuary is mesotidal-macrotidal with variation of the tidal range between 2.8 and 4.7 m. Hydrodynamically influenced by its geomorphology, the Tay estuary is classified as "complex" (A.L. Buck and A.C. Davidson, 1997). The flow channel geometry in Tay has resulted from the combined effects of glaciations, river erosion and sea-level changes. Hence, the Tay Estuary is quite meandering and characterised by 165 km of shoreline. The dominant flow channel on its southern side is deep and narrow having several sand point-bars. The floodplain on the northern side is wide, shallow and includes a number of sand over-banks. The width and depth of the estuary are not regular. From the entrance of the estuary at Buddon Ness, which is 1.9 km wide at low water, the width of the estuary increases exponentially inland to a maximum of 3 km at Dundee. At high water, the width of the estuary reaches a maximum of 5 km at

Birkhill (22.5 km from the estuary mouth) with an exponential decrease from this point towards the estuary head.

Because of its typical flow behaviour and morphology, the Tay estuary has been the subject of many hydrological and statistical analyses in the last four decades (e.g., J. McManus, 1968; A.T. Buller et al. 1972; D.J.A. Williams and J.R. West, 1973; J.A. Charlton and W. McNicoll, 1975; R.A. Pontin and J.A. Reid, 1975; V. Nassehi and D.J.A. Williams, 1987; V. Nassehi and A. Kafai, 1999; S. Passone et al., 2001).

The screenshot shows a Microsoft Access database form titled "ESTUARY". The form is divided into several sections:

- Basic Information:** name: TAY, ID: 4, country: Scotland, latitude: 55.
- Physical characteristics:**
  - geomorphological type: complex
  - tidal range: mesotidal/macrotidal
  - total area (ha): 12128
  - intertidal area (ha): 5583
  - max channel bank area (ha): 2600
  - channel length (km): 54
  - valley length (km): 33
  - meandering/sinuosity: meandering
  - min. width (m): 2000
  - max. width (m): 5000
  - bed: flat bed
  - avr. width (m): 3000
  - avr. depth (m): 10
  - avr. width to depth ratio: very high
- Hydrographic characteristics:**
  - tot. freshwater inflow (m<sup>3</sup>/s): 260
  - tidal volume (m<sup>3</sup>): 63500000
  - salinity: vertically mixed
  - salinity limit (km): 45
  - averaged velocity (m/s): 2
  - tidal period (hr): 12.5
  - inlets: How many inlets? 2
  - coriolis effect
  - wind
  - comments: generally meandering channel

At the bottom of the form, there are navigation buttons: "Apply Model...", "First", "Previous", "Next", "Last", and "New". The status bar at the bottom indicates "Record: 4 of 8" and "Form View".

Figure 8.12. Description of the Tay estuary, Scotland, UK

Considerable variation in the quantity and the velocity of water have been observed in the flow channel along the estuary. During the propagation of the tidal wave the phase lag can reach a maximum of 3 and 1.5 hours at low and high water, respectively. The propagation of the tidal wave is accompanied by the deformation in its shape, which increases with the phase lag. The

irregular geometry of the estuary also affects the mixing and salt transport along the channel. While the differences in surface and bottom salinity throughout the estuary are relatively small (3g/kg), salinity distributions, measured at different locations, show significant asymmetries (V. Nassehi and A. Kafai, 1999).

### 8.5.2 Case Description

The user describes the Tay estuary through the case description form (figure 8.12) and clicks, once the necessary information are given, on the button APPLY MODEL which automatically opens the model selection screen.

### 8.5.3 Case Retrieval

Through the model selection screen (figure 8.13) the user specifies the problem (i.e. salt intrusion) he/she wants to examine, defining the investigation purpose as "management tool". CBEM is now ready to launch the retrieval process. Initially, this provides the user with a list of 4 estuaries which have a first similarity score greater than the given cut-off value of 0.5 (figure

The screenshot shows a window titled "frmselectioncriteria" with the following content:

- Model Selection**
- Estuary: TAY
- problem: salt intrusion
- purpose: management tool
- cut-off: 0.5
- Approximation**
- coriolis effect
- wind effect
- branching
- bed slope
- OK
- Cancel

Figure 8.13. Problem statement

8.14). Only 3 estuaries are selected, excluding from the second part of the retrieval the Tay estuary.

Microsoft Access - [malchestuary]

Similar Models for the estuary **TAY**

in case of **salt intrusion**

ESTUARY	SIMILARITY RATING	NUMBER OF MODELS	
TEES	0.661	2	<input checked="" type="checkbox"/>
FAL	0.545	1	<input checked="" type="checkbox"/>
TAY	1	2	<input type="checkbox"/>
CONWY	0.752	1	<input checked="" type="checkbox"/>

Record: 1 of 4 (Filtered)

Figure 8.14. Set of cases retrieved from the system's case-base ranked with respect to their similarity to the Tay estuary

During the second phase of the retrieval scheme each selected case is ranked with respect to the appropriateness of its corresponding model using the second similarity score. The criteria of exclusion and preference, formulated with respect to the model dimension, are also employed.

After the second retrieval phase, only 2 of the selected 4 cases are returned to the user (figure 8.15). Cases B) and D) (table 8.1.) are withdrawn from the set by the principle of exclusion that tells "If the estuary has inlets and at least one of the inlets is a "branch" then eliminate all 1-D models".

Figure 8.15 shows that the Fal estuary is not considered by the CBEM to be quite similar to the Tay estuary by assigning to it a first similarity score that is just above the cut-off value of



ESTUARY	DIMENSION	ACCURACY	ESTUARY SIM.	MODEL SIM.
FAL	1-D network	moderate	0.546	0.518
TEES	2-D avrg on width	high	0.661	0.482

Figure 8.15. Cases retrieved after the second phase of the retrieval process

0.5. This is also evident from the second similarity score that is not as high as the score for the Upper Milford Haven estuary given in section §8.4.

The suggestion given through the retrieval process can be explained by considering that the Fal and Tay estuaries are generally regarded to be very different from each other in terms of physical and hydrological characteristics (see Appendix A). In particular, because of the width and irregular cross-section of the Tay estuary the use of a depth average 2-D model may be preferable to describe the salt transport into this watercourse (D.M. McDowell and B.A. O'Connor, 1977).

Nevertheless, the 1-D model network for the Fal estuary fulfils the minimum requirement needed for investigating the salt intrusion with the purpose as “management tool”. The aim is to have a simulation characterised by the lowest possible computational time together with an acceptable level of accuracy. Thus, the 1-D model network can be considered sufficient for giving a rough simulation of the longitudinal salt distribution into a branching watercourse such as the Tay estuary.

Therefore, this model is selected and given to the adaptation component to be adjusted by the GA module.

#### *8.5.4 Case Adaptation – Genetic Algorithm Optimisation*

The GA module is activated to determine the set of Manning's friction coefficients that are appropriate regarding the particular hydrodynamics and geomorphological characteristics of the Tay estuary. The calibration of the model is carried out using the field measurements obtained during the spring tide of 12<sup>th</sup> June 1972. Once calibrated, the model is validated. While the set of parameters given by the GA module remained invariant, the model is utilised to simulate the salt intrusion during the neap tide of 20<sup>th</sup> June 1972.

The water surface elevation and salt concentrations observed during these spring and neap tides, are compared to those which are either simulated using the genetic algorithm based calibration or obtained by J.H. Bikangaga (J.H. Bikangaga, 1993) through the manual calibration of the model. The comparison confirms again that the GA routine can very effectively calibrate estuarine models under realistic situations with a significant reduction in the time normally required for optimising a numerical model.

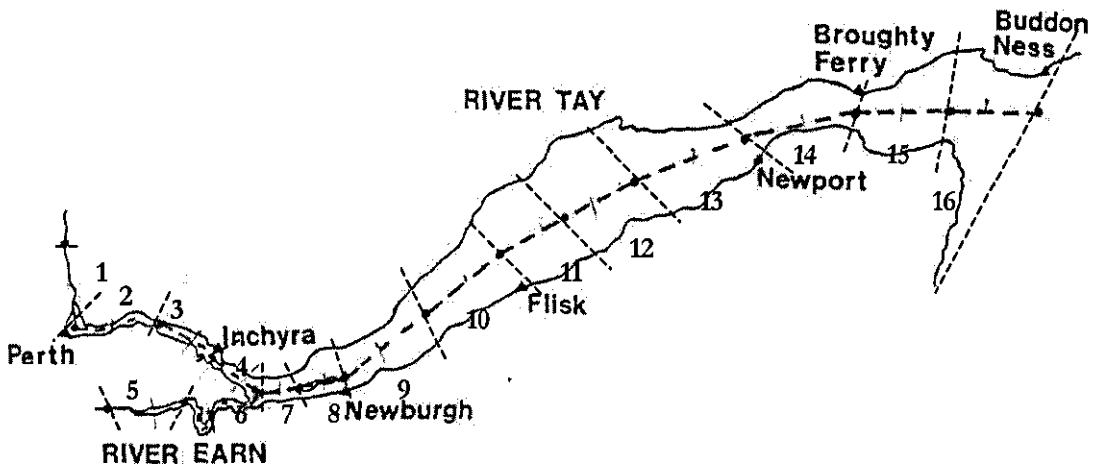
##### *8.5.4.1 Application of the Model to the Tay Estuary*

The application of the 1-D network model to the Tay estuary is based on considering the estuary as a network of three one-dimensional channels with their junction located at the confluence of the rivers Tay and Earn. The estuary is discretised into 16 linear elements (figure 8.16) with nodes coinciding with the sampling stations on the channel for the 1972 survey (A.T. Buller et al., 1972). These stations are located at Broughty Ferry, Newport, Flisk, Newburgh and Inchyra. Hydrographic data are obtained from the mentioned survey, which are tabulated at intervals of 0.15 m from the low water with respect to a datum at Dundee Harbour. Data for nodes not coinciding with the sampling stations are obtained through interpolation.

The model boundary conditions are chosen to be water surface elevations at Buddon Ness, where the estuary mouth is assumed to be located, and fresh water flows of 229 m<sup>3</sup>/s and 31 m<sup>3</sup>/s for the rivers Tay and Earn, respectively.

It must be noted that in order to have an acceptable comparison between the model performances obtained using the manual and GA based calibrations, both model applications

should be specified under identical conditions. This means that except for the set of Manning's friction coefficients, the input data employed in the model calibrated by GA routine, is the same as that utilised by J.H. Bikangaga for modelling the salt intrusion in the Tay estuary (J.H. Bikangaga, 1993).



**Figure 8.16.** Discretisation of the Tay estuary (a Manning's coefficient is assigned to each element in the domain)

#### 8.5.4.2 Fitness Function Definition

The calculation of the fitness function is based on the minimisation of the discrepancy between the water surface elevations simulated by the model and observed at Inchyra and Newburgh during the tide of 12<sup>th</sup> June 1972 (figure 8.16). These locations are selected because the propagation of a tidal wave towards the estuary head is accompanied by significant deformation of its shape at these two stations.

The fitness function (6.1) is defined for the Tay estuary as follows:

$$\rho(H_s, H_m) = \left[ \sum_{j=1}^2 \sum_{n=1}^{13} \left( (h_s)_j^n - (h_m)_j^n \right)^2 \right]^{1/2} \quad (8.11)$$

where the total number of samples, collected at each station during the tidal period is equal to 13.

#### *8.5.4.3 Manning's Friction Coefficient Calibration*

The model calibration by the GA routine is executed with a population of 30 individual and the rate of crossover and mutation equal to 0.5 and 0.01, respectively. The genetic algorithm has been run for 15 generations. The estuary is divided in two zones. The first zone corresponds to the area of the estuary between elements 1 and 10, while the second zone corresponds to the part of the channel from element 11 towards the estuary mouth. Based on this partition of the estuary the chromosome population is initialised using the zonation and scaling options. These chromosomes are then transformed by the modified mutation and crossover operators. Only one set of Manning's coefficients from the case-library is included in the initial population, which is the set of parameters employed in the simulation of salt intrusion into the Fal estuary using the 1-D model network.

The set of Manning's coefficients selected by the genetic algorithm after 15 generations is presented in table 8.3, which also contains the set of parameters manually obtained by Bikangaga (J.H. Bikangaga, 1993). It must be noticed that both sets are within expected range for an estuary modelled using a one-dimensional scheme.

The simulated water surface elevations, generated using the set of Manning's coefficients selected by the GA routine, are presented for the stations of Newport, Flisk, Newburgh and Inchyra in figures 8.17 and 8.18.

The simulation shows that the tidal wave propagates with a progressive deformation of its shape towards the estuary's head (i.e. Inchyra). The comparison between the observed data and the simulated water surface elevations for the manual and the GA based calibrations shows that the model optimised by the GA routine yields a better performance. In particular, while the two sets generate similar water elevations at Newport, the GA based calibration is able to cope with the deformation of the tidal wave, which takes place between the extensive floodplain and the estuary head. There are only small discrepancies between the observed and simulated water surface elevations, obtained using the GA based calibration at Flisk, Newburgh and Inchyra.

**Table 8.3.** Manning's friction coefficients for the Tay estuary

element	GA calibration	manual calibration
1	0.014	0.02
2	0.014	0.02
3	0.013	0.018
4	0.013	0.018
5	0.013	0.016
6	0.013	0.016
7	0.013	0.022
8	0.013	0.022
9	0.016	0.021
10	0.013	0.023
11	0.013	0.023
12	0.013	0.023
13	0.013	0.023
14	0.022	0.023
15	0.022	0.021
16	0.022	0.021

The simulation of the longitudinal salt transport into the Tay estuary for the spring tide of 12<sup>th</sup> June 1972 does not provide any evidence for the improvement of results obtained by the GA optimisation. Both simulations obtained using the GA based and the manual calibrations, yield very similar values for the salt concentration (figure 8.19). This can be explained by the fact that salt intrusion beyond a certain point is so small that it cannot be detected. Furthermore, for locations towards the estuary mouth such as Newport, where significant salt concentrations are observed, the simulated water surface elevations obtained from the GA based calibration and the manual optimisation are very similar. Thus, no comparison between the model performances given by two sets of Manning's coefficients is possible over the simulation of the salt intrusion during the spring tide of 12<sup>th</sup> June 1972.

#### *8.5.4.4 Model Validation – Neap Tide*

Once calibrated, the model needs to be validated under a different condition. For this purpose, the set of Manning's coefficients generated by the GA routine is utilised to simulate the water surface elevations and salt intrusion in the Tay estuary during the neap tide of the 20<sup>th</sup> June 1972.

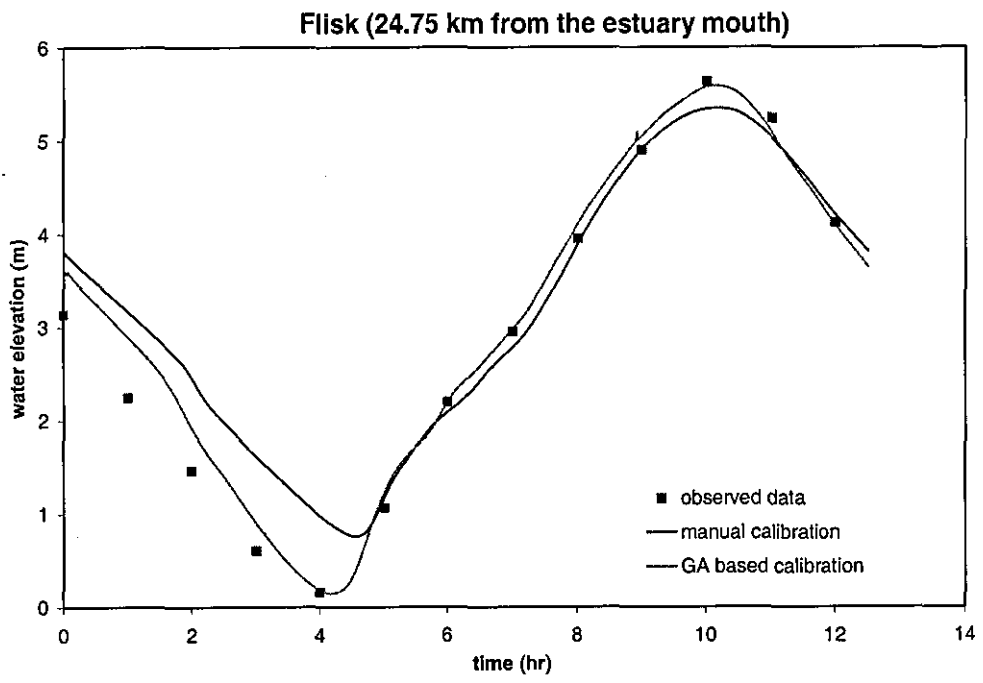
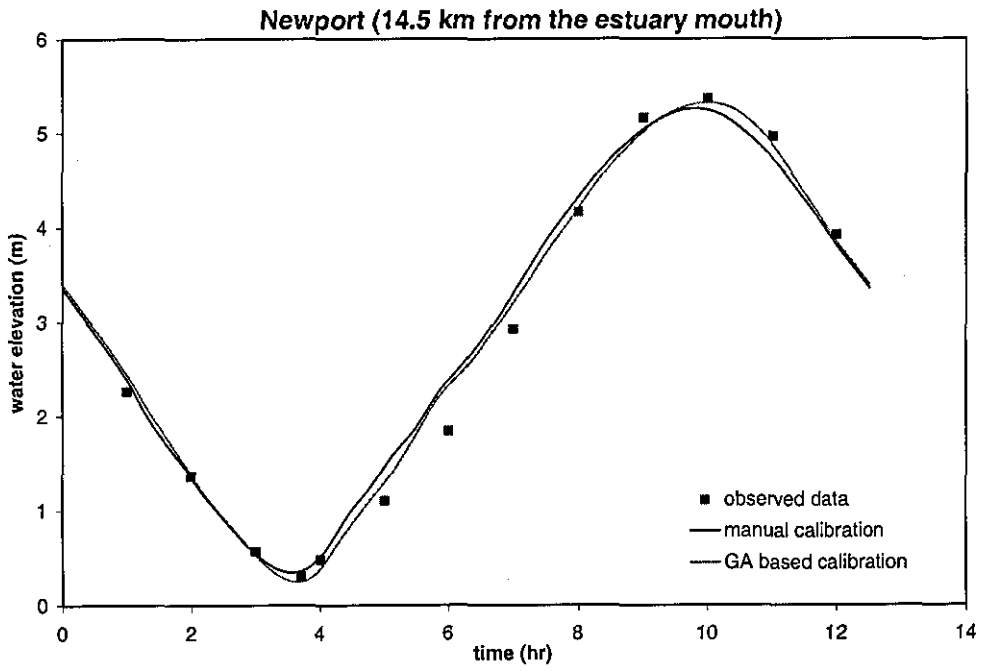
The simulated water surface elevations for this neap tide (figures 8.20 and 8.21) confirm the effectiveness and validity of the GA based calibration.

However, any comparison between the model performances given by the two sets of Manning's coefficients is not possible for the simulation of the salt intrusion during this neap tide (figure 8.22). As already discussed for the spring tide, significant salt concentrations are observed only in that part of the estuary where water surface elevations have the same values regardless the calibration technique used.

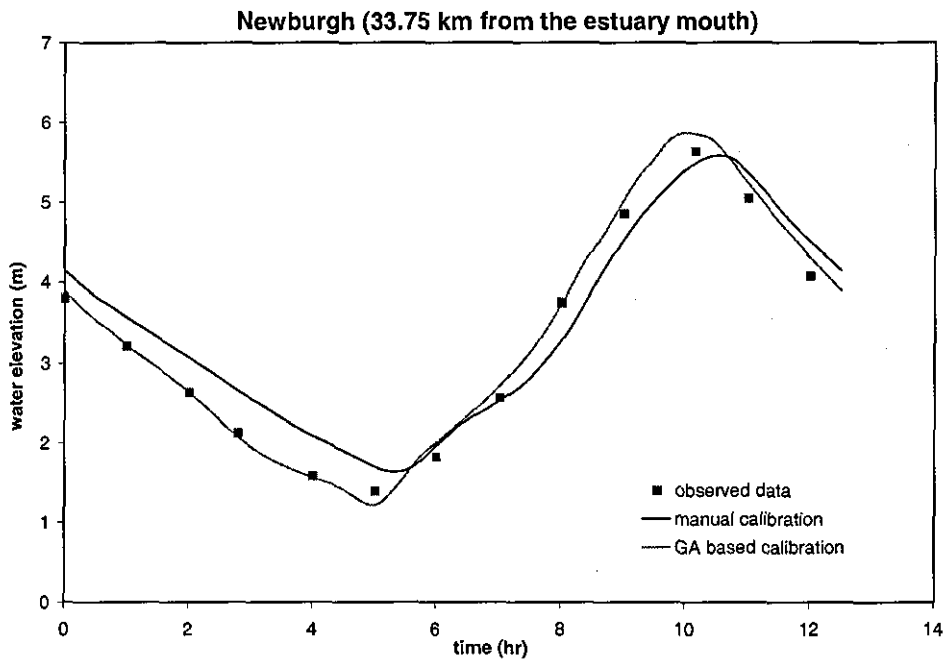
## 8.6 Conclusions

Practical applications of the CBEM have been illustrated in this chapter. The cases of the Upper Milford Haven and Tay estuaries are used to demonstrate the details of the mechanism implemented in the system. They show that the system is able to guide the user towards the most suitable model for a new case with a correct estimation of the model adequacy for the type of problem considered and characteristics of the estuary to be investigated.

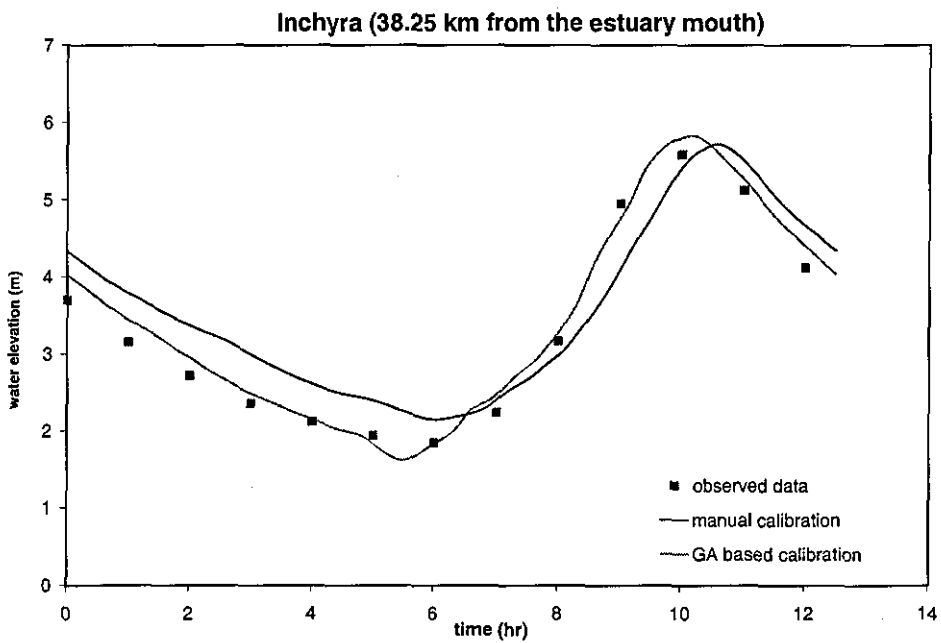
The effectiveness of the implemented adaptation is also demonstrated. The results obtained for the two case studies indicate the feasibility and the effectiveness of the GA based calibration under realistic conditions. The comparison with the simulations given by a manually calibrated model shows that the hybridisation of the classical genetic algorithm theory with problem specific information can provide a superior optimisation process with a significant reduction in time normally required for the implementation of a numerical modelling scheme.



**Figure 8.17.** Simulation of water surface elevations at a) Newport and b) Flisk during the spring tide of 12<sup>th</sup> June 1972



c)



d)

**Figure 8.18.** Simulation of water surface elevations at c) Newburgh and d) Inchyra during the spring tide of 12<sup>th</sup> June 1972



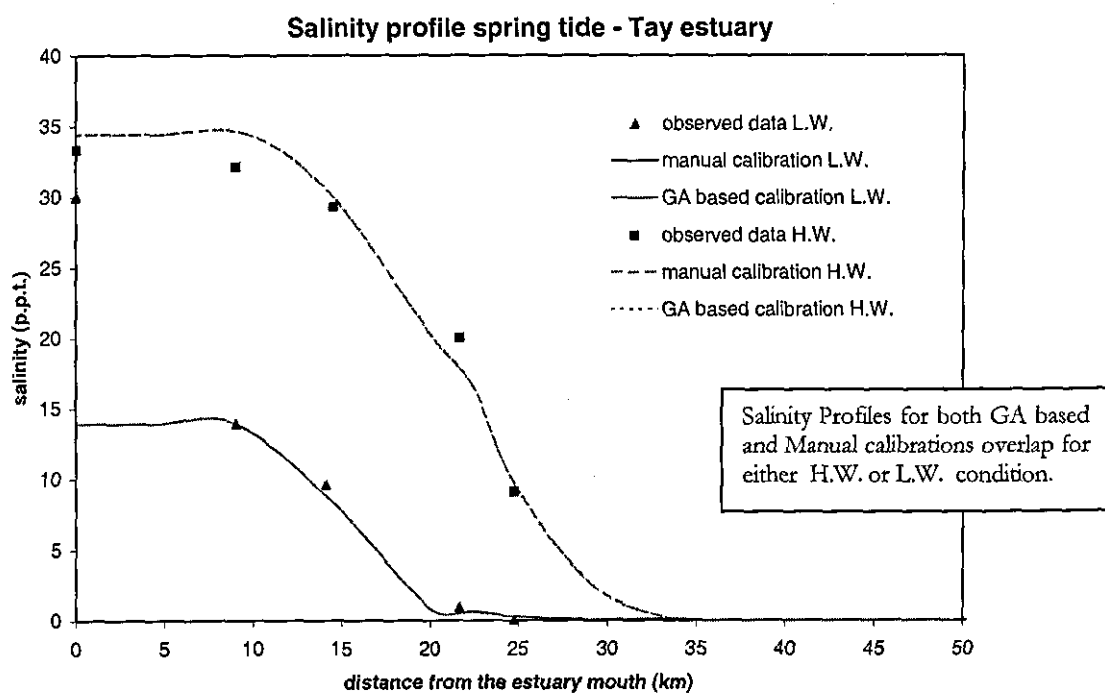
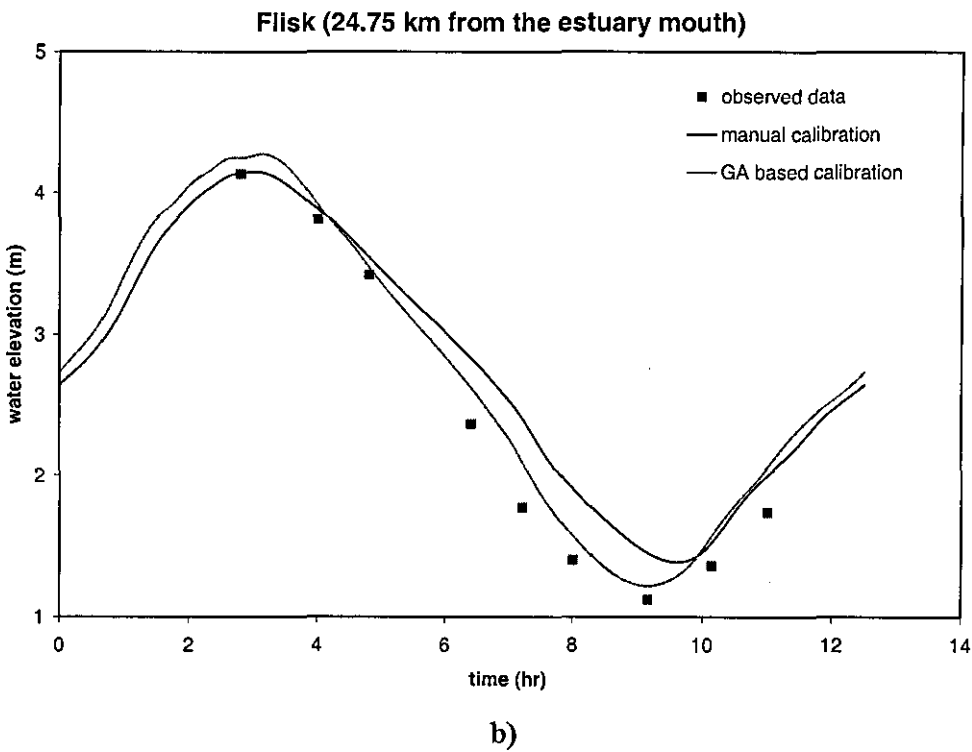
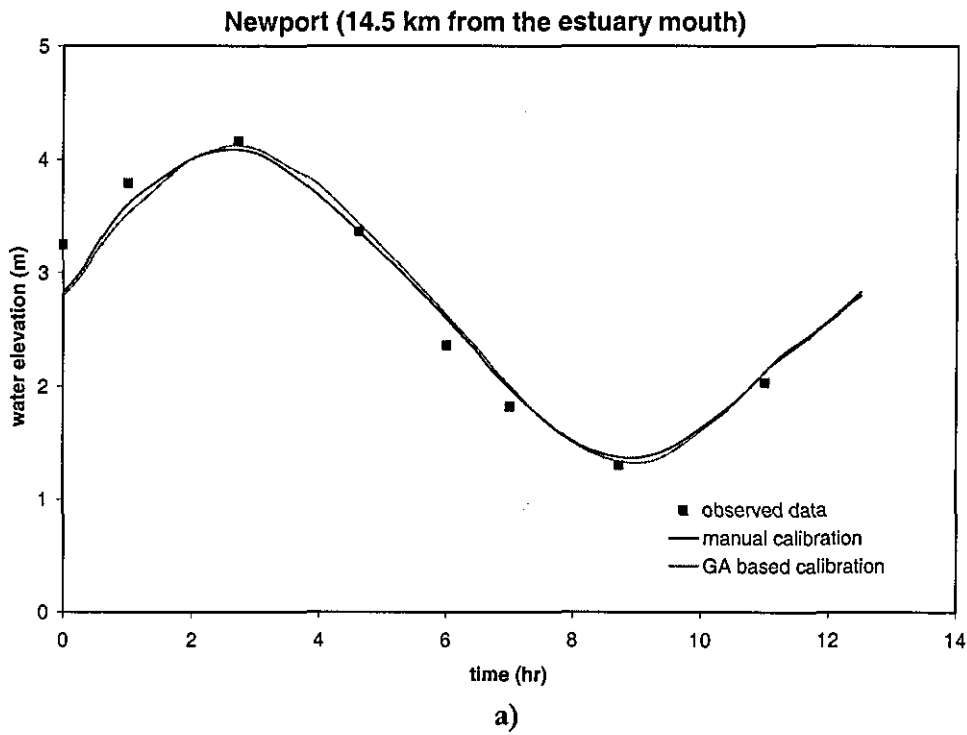
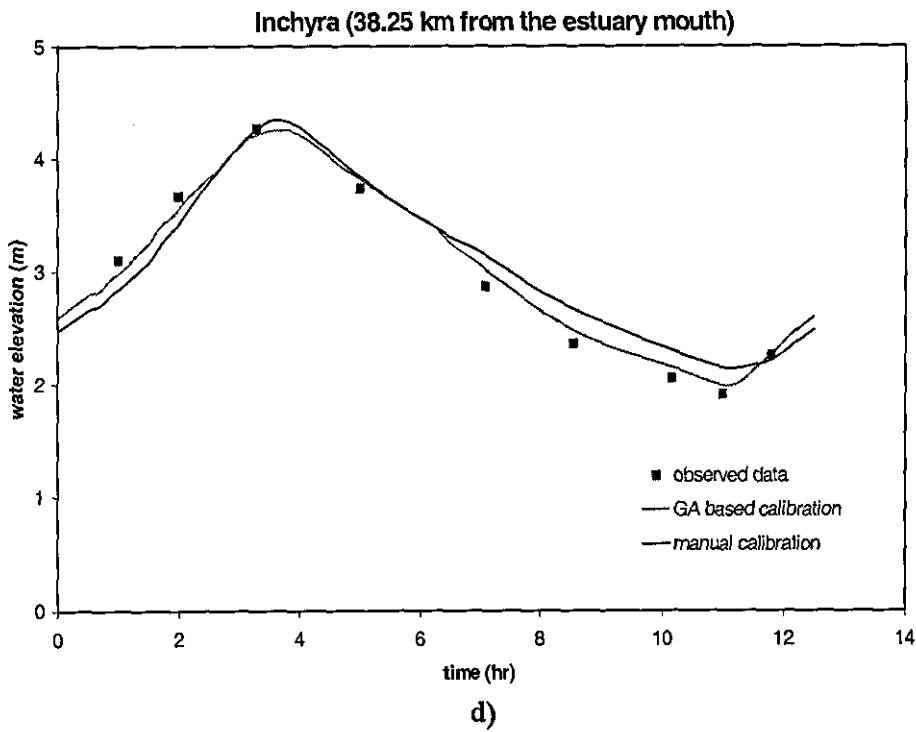
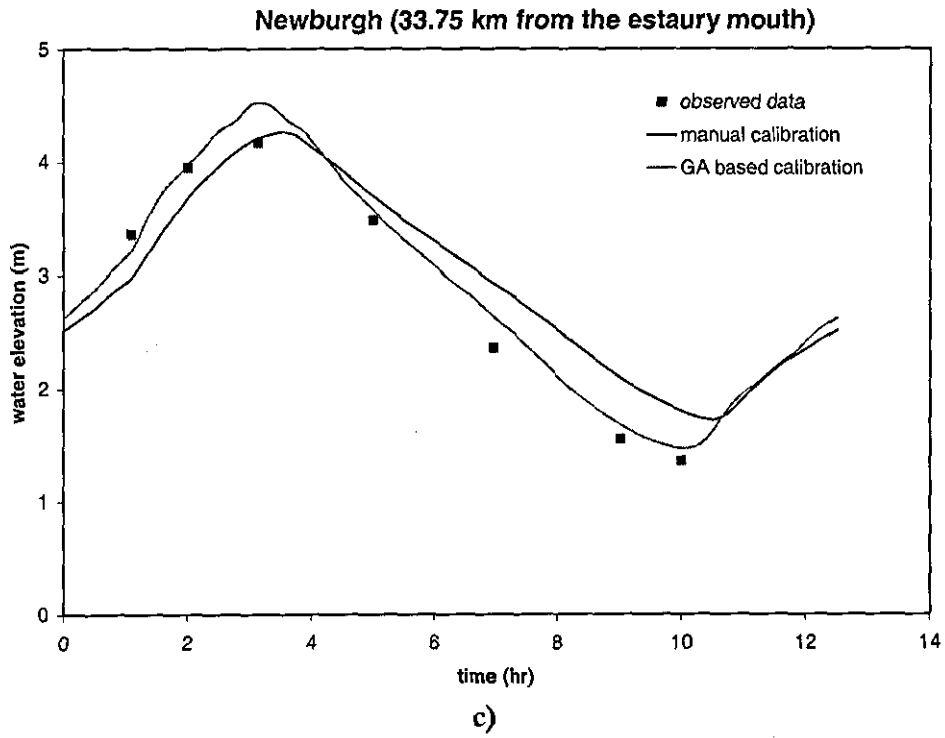


Figure 8.19. Salinity profile during the spring tide of 12<sup>th</sup> June 1972



**Figure 8.20.** Simulation of water surface elevations at a) Newport and b) Flisk during the neap tide of 20<sup>th</sup> June 1972



**Figure 8.21.** Simulation of water surface elevations at c) Newburgh and d) Inchyra during the neap tide of 20<sup>th</sup> June 1972

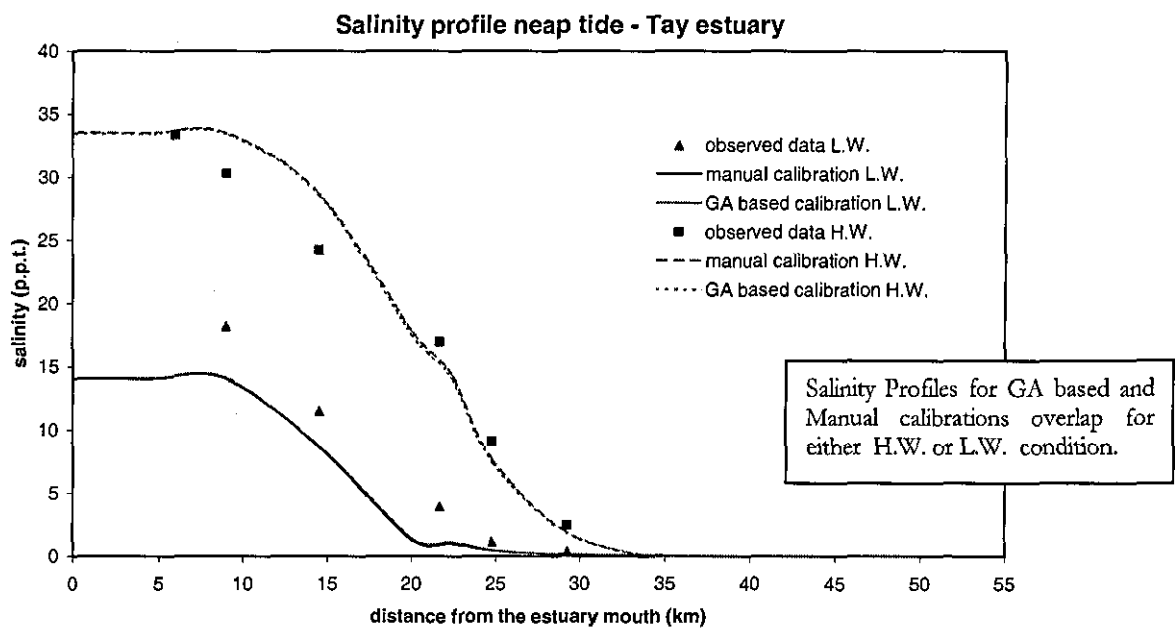


Figure 8.22. Salinity profile during the neap tide of 20<sup>th</sup> June 1972

## CONCLUSIONS AND FUTURE WORK

This chapter outlines the conclusions that can be drawn from the thesis and prospective work that may be undertaken to advance the developed system.

### 9.1 Conclusions

In this research a case-based reasoning system for estuarine modelling (CBEM) has been developed. The system consists of three modules: a case-based module, a genetic algorithm module and a library of numerical codes. These modules are activated to perform specific tasks of the case-based reasoning methodology. These components are integrated to work as a single tool. The main idea is for the system to accumulate past experience, change and adapt it to solve new problems. The current prototype is equipped with examples of past solutions recorded according to the features of specific estuaries and the models employed. The system's ability extends beyond simple retrieval of similar cases. It provides the user with direct information regarding the modification of previous schemes according to a given new problem. The implemented genetic algorithm, hybridised with specific knowledge of estuarine model calibration, is used for adjusting the values of the Manning's friction coefficient for the selected model to suit particular estuarine environments.

Two case studies have been carried out. These are the simulation of tidal dynamics in the Upper Milford Haven and Tay estuaries located in Wales and Scotland, respectively. In particular, the specific problem of salt intrusion for a management tool purpose has been studied.

This work shows that CBEM provides a convenient and effective approach to support the modelling of estuaries. The system is able to provide the user with a critical and accurate evaluation of the suitability of the model strategies contained in its memory for a new case. Although the indexing scheme for the estuary description is limited to the physical and hydrological features of estuaries, nevertheless it has been demonstrated to provide a satisfactory set of indices for estimating the similarity between estuaries. The system is also able to distinguish between the constraints and requirements related to the estuary's physical characteristics as well as those regarding the specific problem to be simulated. Thus, through

the comparison with past cases the user is able to understand how the selection of a modelling strategy is affected by the specific estuary's physical and hydrological characteristics, and the model requirements for accuracy, investigation purpose and simulation time consumed.

Furthermore, the system's inference engine implements different types of knowledge to drive the model selection. For instance, during case description the user is helped in assessing the relevance of the wind and the Coriolis forces and the nature of the estuarine inlets. He/she is also guided during the phase of the estuary description during which the system suggests possible values for other features as the description progresses. The reasoning mechanism of the retrieval component also combines the computation of the similarity based on the model adequacy with the heuristic criteria of exclusion and preference expressed with respect to the model dimension.

The various knowledge related to the practice of estuarine modelling implemented in the description and retrieval components, is organised in CBEM using the rule-based approach and the fuzzy set theory. CBR demonstrates its ability to provide a good framework for integrating different reasoning techniques to overcome the intrinsic complexity of defining a modelling procedure.

CBR is also combined with genetic algorithm theory used in case adaptation for carrying out the automatic adjustment of the Manning's friction coefficient. The implemented GA optimisation, characterised by the incorporation of domain knowledge to guide the search, shows to be a potentially useful technique for adaptation. Not only the GA can very effectively calibrate estuarine models under realistic situations, but it also allows a significant reduction in the time normally required for the validation of a numerical model by other means.

## **9.2 Future Work**

Although CBEM as currently implemented is able to support estuarine model design for realistic problems such as salinity distribution and salt intrusion, the system needs to be further developed. Its performance is limited by the number of cases contained in the system's library. It is also programmed to deal with only two types of problems (i.e. salinity distribution and salt intrusion). No interface currently exists between the CB and the GA modules. Furthermore,

the inference engine is not provided with any form of reasoning mechanism to guide the user during all phases of the modelling.

In order to make CBEM a complete working system, the inclusion of more cases should be considered for ensuring the necessary variability within the case-base. Furthermore, CB and GA modules should be interfaced to permit the automatic communication between the modules.

Other improvements should be carried out to refine and extend the system's reasoning mechanism.

1. The adaptation component needs to be improved so that it is able to adjust multi-model parameters. In particular, through the implementation of a multi-objective genetic algorithm parameter adjustment should provide the calibration of both Manning's friction and dispersion coefficients.
2. A better description of the existing models should also be given. For this purpose, it is suggested that some aspects of the model characterisation (e.g. numerical method, time step, accuracy) should be described in more detail. This will provide the user with a more precise estimation of the overall model's adequacy. Thus, in the logic of model composition, this will allow the retrieval mechanism to identify among the past cases those containing important information (e.g. set of specific assumptions, routines and data) that should be used to form a model strategy for the new case.
3. It would be useful to extend CBEM to support the interpretation of simulation results and, in particular, risk and reliability of the predicted future scenarios using the case-based reasoning approach. The system could be provided with a critical background and uncertainty analysis to help the user to design specific policies for estuarine management.
4. New procedures should also be implemented to support the modelling of other types of problems such as pollutant dispersion, flood routing and sediments movements within estuaries.

Another potential useful avenue for research is to extend the approach implemented in CBEM by creating a new computer tool that focuses on the management of the estuarine water quality. It should be able to

- control resources and process information in real time;
- analyse data, discover trends in a pre-modelling phase and automatically formulate assumptions;
- handle uncertainty;
- facilitate the communication among various levels of the decision-making process.

The idea consists of developing a real-time control system that is able to monitor the water quality and plan appropriate actions to take. The system could be based on a temporal reasoning framework, which helps to integrate the existing computer tools (model programs, GIS, data elaborations), identify the interrelationships between various states of the estuarine water-body and generate, through a realistic time representation, possible management plans. The temporal representation of various processes and events intends to establish the level of interaction on the time scale between the observed phenomena and the action to take for improving the estuarine water quality. This will allow the capture of knowledge in the form of trend discovery, identification of occurred changes and event characterisation. The evaluation of temporal knowledge (observed data, previous experience, events, tasks to be performed, time constraints) is also inextricably linked to other kinds of knowledge, such as knowledge about causality and process decomposition. By identifying processes and functions and their causal relationships, it is possible to represent the estuarine water quality problems in detail with a global view that takes into account the work flow from the origin of the problem to the selection of the final solution.

The adoption of such approach requires unifying the temporal reasoning framework with the ability of the system to deal with uncertainty related to past and future events, in terms of incompleteness, fuzziness of the available information and the limitations of the computer tools. This means that throughout the entire process different kinds of uncertainty, risk and reliability of various phases (data, imposed conditions, missing information, model approximation, predicted scenarios) are evaluated and quantified. The real-time knowledge



based system for estuarine water quality should be able to characterise the environmental conditions and allocate the technical resources for tailoring the response according to the existing requirements and constraints. By making the communication among the managers and other professionals during the decision making process continuous and transparent, the impact of several decision strategies will be estimated. They can then be refined and, eventually, turn into a complete solution.

## REFERENCES

- Aamodt, A. and Plaza, E. (1994) *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches*, In AICom, IOS Press, 7, 39-59
- Abbott, M.B., Babovic, V. and Cunge, J.A. (2001) *Towards the Hydraulics of the Hydroinformatics Era*, Journal of Hydraulic Research, 39(4), 339-349
- Abbott, M.B. (1991) *Hydroinformatics - Information Technology and the Aquatic Environment*, Avebury Technical, Aldershot
- Aral, M.M., Guan, J.B. and Maslia, M.L. (2001) *Identification of Contaminant Source Location and Release History in Aquifers*, Journal of Hydrologic Engineering, 6(3), 225-234
- Azuaje, F., Dubitzky, W., Black, N. and Adamson, K. (2000) *Discovering relevance knowledge in data: A growing cell structures approach*, IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics, 30(3), 448-460
- Babovic, V. and Abbott, M.B. (1997) *The Evolution of Equations from Hydraulic Data, 1. Theory*, Journal of Hydraulic Research 35(3), 397-410
- Babovic, V. Wu, Z. and Larsen, L.C. (1994) *Calibrating Hydrodynamics Models by Means of Simulated Evolution*, Proceedings of the First International Conference in Hydroinformatics, A.A. Balkema, 193-200
- Beck, M.B. (1991) *Principles of Modelling*, Water Science and Technology, 24(6), 1-8
- Bikangaga, J.H. and Nassehi, V. (1995) *Application of Computer Modelling Techniques to The Determination of Optimum Effluent Discharge Policies in Tidal Water Systems*, Water Resources, 29(10), 2367-2375
- Bikangaga, J.H. (1993) *Mathematical Modelling of the Hydrodynamics and Transport of Soluble Reactive Pollutants in Narrow Tidal Rivers*, PhD Thesis
- Bowden K. F. (1980) *Physical Factors: Salinity, Temperature, Circulation, and Mixing Processes*, in Chemistry and Biochemistry of Estuaries, Edited by E. Olausson and I. Cato, 37-69, John Wiley & Sons, New York
- Buck, A.L. and Davidson, A.C. (1997) *An Inventory of UK Estuaries*, Joint Natural Conservation Committee, Peterborough
- Buller, A. T., Charlton, J.A and McManus, J. (1972) *Data from Physical and Chemical Measurements in the Tay Estuary for Neap and Spring Tides*, Tay Estuary Research Report, No.: 2, University of Dundee, Dundee.
- Charlton, J. A. and McNicoll, W. (1975) *Tidal and Freshwater Induced Circulation in the Tay Estuary*, Proceedings Royal Society Edinburgh B, 75(2), 11-27

- Chau, K.W. and Chen, W. (2001) '*A Fifth Generation Numerical Modelling System in Coastal Zone*', Applied mathematical Modelling, **25**, 887-900
- Chau, K.W. and Yang, W.W. (1994) '*Structuring and Evaluation of VP-Expert Based Knowledge Bases*', Engineering Applications of Artificial Intelligence, **7**(4), 447-454
- Chung, P.W.H. and Inder, R. (1992) '*Handling Uncertainty in Accessing Petroleum Exploration Data*', Revue de L'Institut Francais du Petrole, **47**(3), 305-314
- Cooper, J.A.G. (2001) '*Geomorphological Variability Among Microtidal Estuaries from the Wave-Dominated South African Coast*', Geomorphology, **40**(1-2), 99-122
- Cortés U., Sánchez-Marrè, M., Sangüesa, R., Comas, J., Roda, J.R., Poch, M. and Riaño, D. (2001) '*Knowledge Management in Environmental Decision Support System*', Ai Communication, **14**, 3-12
- Cox, E. (1994) The Fuzzy System Handbook, AP Professional, Chestnut Hill, MA
- Cunge, J.A. and Erlich, M. (1999) '*Hydroinformatics in 1999: What is to be done?*', Journal of Hydroinformatics, **1**(1), 21-31
- Davidson, N.C., d'A Laffoley, D., Doody, J.P., Way, L.S., Gordon, J., Key, R., Pienkowski, M.W., Mitchell, R. and Duff, K.L. (1991) Nature Conservation and Estuaries in Great Britain, Peterborough, Nature Conservancy Council
- Davies, J.L. (1973) Geographical Variation in Coastal Development, Longman, London
- Davis, L. (1991) Handbook of Genetic Algorithms, Van Nostrand Reinhold, New York, U.S.A.
- Dyke, P. (1996) Modelling Marine Processes, Prentice-Hall, London
- Dyer, K.R. (1997) Estuaries: a Physical Introduction, John Wiley & Sons, New York
- Fairbridge R.W. (1980) '*The Estuary: Its Definition and Geodynamic Cycle*', in Chemistry and Biochemistry of Estuaries, Edited by E. Olausson and I. Cato, 1-35, John Wiley & Sons, New York
- Finn, P.D., Cunningham, P. (1998) '*Solution Reuse for Model Generation in Numerical Simulation*', Artificial Intelligence in Engineering, **12**, 297-314
- Fischer, H.B. (1972) '*Mass Transport Mechanisms in Partially Stratified Estuaries*', Journal of Fluid Mechanics, **53**, 672-687
- Fleming, P.J. and Zalzalá, A.M.S. '*Special Issue on Genetic Algorithms in Engineering Systems*', Proceedings of The Institution Of Mechanical Engineers Part I-Journal Of Systems And Control Engineering, **211**(5), I-II

- French, J.R. and Clifford, N.J. (2000) *Hydrodynamic Modelling as a Basis for Explaining Estuarine Environmental Dynamics: Some Computational and Methodological Issues*, *Hydrological Process* **14**, 2089-2108
- Gillibrand, P.A. and Balls, P.W. (1998) *Modelling Salt Intrusion and Nitrate Concentrations in the Ythan Estuary*, *Estuarine Coastal and Shelf Science*, **47**(6), 695-706
- Gijsbers, P.J.A., Moore, R.V. and Tindall, C.I. (2002) *HarmonIT: towards an Open Modelling Interface and Environment to Harmonise European Developments in Water Related Simulation Software*, Proceedings of the 5<sup>th</sup> International Conference on Hydroinformatics, edited by I.D. Cluckie, D. Han, J.P. Davis and S. Heslop, 1268-1275
- Goldberg, D.E. (1989) *Genetic Algorithm in Search, Optimisation and Machine Learning*, Addison-Wesley Publishing Company, Massachusetts, U.S.A.
- Grefenstette, J. J. (1987) *Incorporating Problem Specific Knowledge into Genetic Algorithms*, In *Genetic Algorithms and Simulated Annealing*, L. D. Davis (Ed.), 42-61, London: Pitman
- Hall, J.W. (2002) *Random Set Uncertainty Handling in Hydroinformatics*, Proceedings of the 5<sup>th</sup> International Conference on Hydroinformatics, edited by I.D. Cluckie, D. Han, J.P. Davis and S. Heslop, 1388-1393
- Harvey, D., Han, D. and Cluckie, I. (2002) *A Blueprint for the Next Generation Modelling Software*, Proceedings of the 5<sup>th</sup> International Conference on Hydroinformatics, edited by I.D. Cluckie, D. Han, J.P. Davis and S. Heslop, 1276-1281
- Häuslein, A and Page, B. (1991) *Knowledge-Based Approaches to Modeling and Simulation Support*, *Systems Analysis Modelling Simulation*, **8**(4-5), 257-272.
- Hayes, M.O. (1975) *Morphology of Sand Accumulation in Estuaries*, in *Estuarine Research, Volume 2: Geology and Engineering*, Edited by L.E. Cronin, 3-22, Academic Press, New York
- Hinwood, J.B. and Wallis, I.G. (1975) *Classification of Models of Tidal Waters*, *Journal of Hydraulics Division, Proc. of A.S.C.E.*, **101**, HY10 1315-1351
- Holland J.H. (1973), *Genetic Algorithms and the Optimal Allocations of Trials*, *SIAM Journal of Computing*, **2**(2), 88-105
- Holt, A. and Benwell, G.L. (1999) *Applying Case-Based Reasoning Techniques in GIS*, *The International Journal of Geographical Information Science*, **13**(1), 9-25
- Ibrahim, Z.Z., Moi, L.S., Abdullah, R. and Arshad, A. (1996) *Classification of Malaysian Estuaries for Development Planning*, *Aquatic Conservation-Marine and Freshwater Ecosystems*, **6**(4), 195-203
- Ippen, A.T. (1966) *Estuary and Coastline Hydrodynamics*, McGraw-Hill, New York

- Jay D. A and Smith J. D. (1988) *Residual Circulation in and Classification of Shallow, Stratified Estuaries*, in Physical Processes in Estuaries, Edited by J. Dronkers and W. Van Leussen, Springer-Verlag
- Jeng, B.C. and Liang, T.P. (1995) *Fuzzy indexing and retrieval in case-based systems*, Expert Systems With Applications, 8(1), 135-142
- Johnson, V.M. and Rogers L.L. (1995) *Location Analysis in Groundwater Remediation Using Neural Networks*, Groundwater 33(5), 749-758
- Kennish, M.J. (1986) Ecology of Estuaries, Vol.1: Physical and Chemical Aspects, Boca Raton (Fla.): CRC Press
- Kolodner, J. (1993) Case-Based Reasoning, Morgan Kaufmann, San Mateo, California
- Lal Wasantha, A.M. (1995) *Calibration of Riverbed Roughness*, Journal of Hydraulic Engineering, ASCE 121(9), 664-671
- Lakshmi Prasad, K. and Rastogi, A.K. (2001) *Estimating net aquifer recharge and zonal hydraulic conductivity values for Mahi Right Bank Canal project area, India by genetic algorithm*, Journal of Hydrology, 243, 149-161
- Lawson, J. and Gunn, I.J. (1996) Guidelines for the Use of Computational Model in Coastal and Estuarial Studies, Report SR 450, HR Wallingford, Oxon, UK
- Liao, W.T., Zhang, Z. and Mount, C.R. (1998) *Similarity Measures for Retrieval in Case-Based Reasoning Systems*, Applied Artificial Intelligence, 12, 267-288
- Louis, S.J. and Johnson, J. (1997) *Solving Similar Problems Using Genetic Algorithms and Case-Base Memory*, Proceedings of the Seventh International Conference on Genetic Algorithms, 101-127, Morgan Kauffman, San Mateo, CA
- McCuen, R.H. and Snyder, W.M. (1986) Hydrologic Modelling, Statistical Methods and Applications, Englewood Cliffs, N.J. London, Prentice-Hall
- McDowell, D.M. and O'Connor, B.A. (1977) Hydraulic Behaviour of Estuaries, Macmillan Press LTD, London
- McKinney, D.C. and Lin, M.D. (1994) *Genetic Algorithm Solution of Groundwater-Management Models*, Water Resources Research, 30(6) 1897-1906
- McLaren, P. and Little, D.I. (1987) *The Effect of Sediment Transport on Contaminant Dispersal: an Example from Milford Haven*, Marine Pollution Bulletin, 18, 586-594
- McManus, J. (1968) The Hydrology of the Tay basin in Dundee and District, (Ed. Jones, SD). British Association for the Advancement of Science, Dundee, 107-124.
- Michalewicz, Z. (1994) Genetic Algorithms, 2<sup>nd</sup> edition, Springer-Verlag, Berlin, New-York

- Morris, A.W. (1983): Practical Procedures For Estuarine Studies, A Handbook Prepared by the Estuarine Ecology Group of the Institute For Marine Environmental Research, Edited by A.W. Morris, Natural Environment Research Council, London
- Natural Environment Research Council (1975) Estuaries Research, Natural Environment Research Council Publications Series 'B', No. 9
- Nassehi, V. and Kafai, A. (1999) '*A Moving Boundary Computer Model for Hydrodynamics and Salt Transport in the Tay Estuary*', *Advances in Environmental Research*, **3**(3), 293-308
- Nassehi, V. and Williams, D.J. (1987) '*A Mathematical Model for Salt Intrusion in the Tay Estuary*', *Proceedings Royal Society Edinburgh B*, **92**, 285-295.
- Nassehi, V. and Williams, D.J.A. (1986) '*Mathematical Model of Upper Milford Haven – A Branching Estuary*', *Estuarine, Coastal and Shelf Science*, **23**, 403-418
- Nelson-Smith, A. (1965) *Marine Biology of Milford Haven: the Physical Environment*, *Field studies*, **2**, 155-188
- NOAA (1985) Strategic Environmental Branch National Estuarine Inventory: Data Atlas, Volume I: Physical and Hydrologic Characteristics. Rockville, MD:NOAA
- Passone, S., Das, D.B., Nassehi, V. '*Design of Discharge Policies for Multiple Effluent Sources and Returning Pollutants Scenarios in a Branching Estuarine*', *Coastal and Shelf Science* (under publication).
- Passone, S., Das, D.B. and Nassehi, V. (2001) '*Development of an Effluent Discharge Policy for the Tay Estuary Based on a Finite Element Model*', *Water Science and Technology*, **43**(7), 247-256
- Perillo, G.M.E. (1995) *Geomorphology and Sedimentology of Estuaries*, Elsevier
- Pontin, R.A. and Reid, J.A. (1975) '*The Freshwater Input to the Tay Estuary*', *Proceedings Royal Society Edinburgh B*, **75**, 1-9
- Pritchard D.W. (1967) '*Observation of Circulation in Coastal Plain Estuaries*', *Estuaries* Edited G. H. Lauff, American Association for the Advancement of Science
- Pritchard, D.W. (1952) '*Salinity Distribution and Circulation in the Chesapeake Bay Estuaries System*', *Journal of Marine Resources*, **11**, 106-123
- Ramesh, R., Datta, B., Murty Bhallamudi, S. and Narayana, A. (2000) '*Optimal Estimation of Roughness in Open-Channel Flows*', *Journal of Hydraulic Engineering*, **126**(4), 299-302
- Ramsey, C.L. and Grefenstette, J.J. (1993) '*Case-based Initialization of Genetic Algorithms*', *Proceedings Fifth International Conference on Genetic Algorithms (ICGA93)*, San Mateo: Morgan Kaufmann, 84-91
- Ritzel, B.J., Eheart, J.W. and Ranjithan S. (1994) '*Using Genetic Algorithms to solve a Multiple Objective Groundwater Pollution Containment Problem*', *Water Resources Research*, **30**(5), 1589-1603

- Rijkswaterstaat Communications (1976) Salt Distribution in Estuaries, Proceedings of Seminar Held in 1974
- Rosgen, D. (1996) Applied River Morphology, Pagosa Springs, Colo., Wildland Hydrology
- Rossiter, J.R. and Lennon, G.W. (1965) 'Computation of Tidal Conditions in the Thames Estuary by the Initial value Method', Proceedings Institution of Civil Engineers, **31**, 25-56
- Scott, C. (1993) 'Canonical Parameter for Estuarine Classification', Estuarine, Coastal Shelf Science, **13**, 529-540
- Smith, R. and Sussman, D. (1997) Beginning Access 97 VBA programming, Birmingham : Wrox Press
- Shin, K. And Han, I. (1999) 'Case-Based Reasoning Supported by Genetic Algorithms for Corporate Bond Rating', Expert Systems with Applications, **16**, 85-95
- Solamatine, D.P. (2002) *Data-driven Modelling – Paradigm, Methods*, Experiences, Proceedings of the 5<sup>th</sup> International Conference on Hydroinformatics, edited by R.A. Falconer, B. Lin, E.L. Harris and C.A.M.E. Wilson, 757-763
- Thatcher, M.L. and Harleman, D.R.F. (1972) A mathematical model for the prediction of unsteady salinity intrusion in estuaries. Ref. 144 Department of Civil Engineering, MIT, Cambridge, Mass. U.S.A.
- Thompson, G. (1993) *Mathematical-Models and Engineering Design*, Journal of the Institution of Water and Environmental Management, **7**, 18-23
- Tomczak Matthias (1999) Shelf and Coastal Oceanography, <http://gaea.es.flinders.edu.au/~mattom/ShelfCoast/>
- Topping, B.H.V., Sziveri, J., Bahreinejad, A., Leite, J.P.B. and Cheng, B. (1998) *Parallel processing, Neural Network and Genetic Algorithm*, Advances in Engineering Software, **29**(10), 763-786
- Van Zuylen, H.J. (1994) 'Knowledge Based Support of Users of Numerical Programs', Mathematics and Computers in Simulation, **36**, 327-336
- Wang, Q.J. (1997) 'Using Genetic Algorithms to Optimise Model Parameters', Environmental Modelling & Software, **12**(1), 27-34
- Wang, Q.J. (1991) *The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models*, Water Resources Research, **27**(9), 2467-2471
- Watson, I. and Marir, F. (1994) 'Case-Based Reasoning - A Review', Knowledge Engineering Review, **9**(4), 327-354
- West, J.R. and Broyd, T.W. (1981) *Dispersion Coefficients in Estuaries*, Proceedings Institution of Civil Engineers, **71**, 721-737

- Williams, D.J.A. and West, J.R. (1973) '*A One-Dimensional Representation of Mixing in the Tay Estuary*', Technical Paper in Water Pollution Resources, **13**, 117-125
- Wilson, J.G. (1988): *The Biology of Estuarine Management*, Croom Helm
- Woodruff, J. (2002) Personal Communication
- XU, L.D. (1994) '*Developing a Case-Based Knowledge System for Aids-Prevention*', Expert Systems, **11**(4), 237-244



# APPENDIX A

## CBEM CASES

The appendix contains the estuary description of the past cases in CBEM case-base used for the case studies presented in chapter 8.

The estuaries are: Tees (figure A.1), Fal (figure A.2), and Conwy (figure A.3).

The screenshot shows a software window titled "frmMaintestuary" with a sub-header "ESTUARY". The interface is divided into several sections for data entry:

- General Information:** name: TEES, ID: 1, country: England, latitude: 54.
- Physical characteristics:**
  - geomorphological type: coastal plain, tidal range: macrotidal
  - total area (ha): 1347, intertidal area (ha): 471, max channel bank area (ha): 0
  - channel length (km): 33, valley length (km): 18, meandering/sinuosity: extremely meandering
  - min. width (m): 10, max. width (m): 1000, bed: flat bed
  - avr. width (m): 300, avr. depth (m): 10, avr. width to depth ratio: high
- Hydrographic characteristics:**
  - tot. freshwater inflow (m<sup>3</sup>/s): 1, tidal volume (m<sup>3</sup>): 3000, salinity: vertically mixed
  - salinity limit (km): 30, averaged velocity (m/s): 0.5, tidal period (hr): 12.5
  - checkboxes for inlets, coriolis effect, and wind.
- Comments:** A text box containing "prevalence of mud flats and salt marshes".
- Buttons:** "Apply Model", "inlets", "models", "Next", "Last", "New".
- Status Bar:** "Record: 14 / 1 of 9".

Figure A.1. Description of the Tees estuary, UK

frmMaintestuary

### ESTUARY

name  ID  country  latitude

Physical characteristics

geomorphological type  tidal range

total area (ha)  intertidal area (ha)  max channel bank area (ha)

channel length (km)  valley length (km)  meandering/sinuosity

min width (m)  max width (m)  bed

avr. width (m)  avr. depth (m)  avr. width to depth ratio

Hydrographic characteristics

tot. freshwater inflow (m<sup>3</sup>/s)  tidal volume (m<sup>3</sup>)  salinity

salinity limit (km)  averaged velocity (m/s)  tidal period (hr)

inlets How many inlets?   coriolis effect  wind

comments

Record: 14.4 | 3 | 11/20 of 8

Figure A.2. Description of the Fal estuary, UK

**frmMaintestuary**

**ESTUARY**

name:  ID:  country:  latitude:

**Physical characteristics**

geomorphological type:  tidal range:

total area (ha):  intertidal area (ha):  max. channel bank area (ha):

channel length (km):  valley length (km):  meandering/sinuosity:

min. width (m):  max. width (m):  bed:

avr. width (m):  avr. depth (m):  avr. width to depth ratio:

**Hydrographic characteristics**

tot. freshwater inflow (m<sup>3</sup>/s):  tidal volume (m<sup>3</sup>):  salinity:

salinity limit (km):  averaged velocity (m/s):  tidal period (hr):

inlets  coriolis effect  wind

comments:

Record:  of

Figure A.3. Description of the Conwy estuary, UK

*APPENDIX B*

CB MODULE - CODE

## Description Component

```
Option Compare Database
Option Explicit
```

```
Private Sub cbomeandering_AfterUpdate()
```

```
    If Me!cbomeandering = "relatively straight" Then
        Me!Text227.value = 1
        Me!Text231.value = 0
        Me!Text233.value = 0
    End If
    If Me!cbomeandering = "meandering" Then
        Me!Text227.value = 0
        Me!Text231.value = 1
        Me!Text233.value = 0
    End If
    If Me!cbomeandering = "extremely meandering" Then
        Me!Text227.value = 0
        Me!Text231.value = 0
        Me!Text233.value = 1
    End If
    If Me!cbomeandering = "relatively straight/meandering" Then
        Me!Text227.value = 0.5
        Me!Text231.value = 0.5
        Me!Text233.value = 0
    End If
    If Me!cbomeandering = "meandering/extremely meandering" Then
        Me!Text227.value = 0
        Me!Text231.value = 0.5
        Me!Text233.value = 0.5
    End If
End Sub
```

```
Private Sub cboqualitative_AfterUpdate()
```

```
    Call qualitative
    If Me!cboqualitative = "low" Then
        Me!Text219.value = 1
        Me!Text221.value = 0
        Me!Text223.value = 0
        Me!Text225.value = 0
    End If
```

```
End If
    If Me!cboqualitative = "moderate" Then
        Me!Text219.value = 0
        Me!Text221.value = 1
        Me!Text223.value = 0
        Me!Text225.value = 0
    End If
    If Me!cboqualitative = "high" Then
        Me!Text219.value = 0
        Me!Text221.value = 0
        Me!Text223.value = 1
        Me!Text225.value = 0
    End If
    If Me!cboqualitative = "very high" Then
        Me!Text219.value = 0
        Me!Text221.value = 0
        Me!Text223.value = 0
        Me!Text225.value = 1
    End If
    If Me!cboqualitative = "low/moderate" Then
        Me!Text219.value = 0.5
        Me!Text221.value = 0.5
        Me!Text223.value = 0
        Me!Text225.value = 0
    End If
    If Me!cboqualitative = "moderate/high" Then
        Me!Text219.value = 0
        Me!Text221.value = 0.5
        Me!Text223.value = 0.5
        Me!Text225.value = 0
    End If
    If Me!cboqualitative = "high/very high" Then
        Me!Text219.value = 0
        Me!Text221.value = 0
        Me!Text223.value = 0.5
        Me!Text225.value = 0.5
    End If
```

```
End Sub
```

```
Private Sub cbotidal_range_Change()
```

```

Dim msg As String
Dim flow As String
Dim strmsg As String
Dim title As String
flow = Me.cbottidal_range

Select Case flow
Case "macrotidal"
Me!txtcomments.value = "prevalence of mud flats and salt marshes"
strmsg = "Prevalence of mud flats and salt marshes typical of
coastal plain estuaries." _
    & "Check the estuary geomorphological type."

title = "suggestion"
MsgBox strmsg, vbOKOnly, title
Case "mesotidal/macrotidal"
Me!txtcomments.value = "generally meandering channel"
strmsg = "Generally meandering channel. Check the meandeirng value."
title = "suggestion"
MsgBox strmsg, vbOKOnly, title
Case "microtidal/mesotidal"
Me!txtcomments.value = "generally meandering channel"
strmsg = "Generally meandering channel. Check the meandeirng value."
title = "suggestion"
MsgBox strmsg, vbOKOnly, title
Case "microtidal"
Me!txtcomments.value = "presence of spits and barrier islands"
msg = Me!txtcomments.value
strmsg = "Presence of spits and barrier islands. Control the wind
force and the width and depth value"
title = "suggestion"
MsgBox strmsg, vbOKOnly, title
msg = msg And "control the wind drive force and the dimensions"
End Select
End Sub

Private Sub chkcoriolis_Click()
'calculus of coriolis force magnitudo with respect to
'the averaged velocity and the averaged depth

Dim strinput As Variant, strmsg As String
Dim vardepth As Variant, varcoriolis As Variant, sngvelocity As
Single
Dim intlatitude As Integer
Dim sngstressbottom As Single, n As Single
Dim dischargevelocity As Single
Dim msg As String
Const omega = 0.000073
Const density = 1025
Const pigreco = 3.1415
Const gravityacceleration = 10
Const chezycoefficient = 50

```

```

If (Me!chkcoriolis = True And Me!txtvelocity.value = 0) Then
msg = "the velocity must be inserted "
MsgBox msg
Me!txtvelocity.SetFocus
Me!chkcoriolis = False
End If
If (Me!chkcoriolis = True And Me!txtvelocity.value <> 0) Then
sngvelocity = Me!txtvelocity
intlatitude = Me!txtlatitude
varcoriolis = sngvelocity * strinput * 2 * omega *
Sin(intlatitude * pigreco / 180)
dischargevelocity = Me!txtvelocity
sngstressbottom = (density * gravityacceleration) /
chezycoefficient ^ 2
sngstressbottom = sngstressbottom * dischargevelocity ^ 2

If ((varcoriolis / sngstressbottom) <= 0.1) Then
strmsg = "the coriolis effect can be neglected"
MsgBox strmsg
txtvelocity.SetFocus
Me!chkcoriolis = False
Else
txtvelocity.SetFocus
Me!chkcoriolis = True
End If
End If
End Sub

Private Sub chkinlets_Click()
If Me!chkinlets = True Then Me!Txtnumberinlets.Visible = True
If Me!chkinlets = False Then Me!Txtnumberinlets.Visible = False
End Sub

Private Sub chkmodel_BeforeUpdate(Cancel As Integer)
chkmodel.value = False
End Sub

Private Sub chkwind_Click()
'calculus of wind velocity W magnitudo with respect to the averaged
longitudinal velocity
'the value is calculated with respect to
'surface shear stress due to wind =(air density*friction
coefficient)*(wind velocity)^2
'and bottom shear stress= (water density*gravity acceleration)(water
velocity)^2

If Me!chkwind = False Then
Me!txtwindvelocity.Visible = False
Exit Sub

```

```

End If

Dim windvelocity As Single
Dim dischargevelocity As Single
Dim n As Single
Dim m As Single
Dim msg As String
Const densityair = 1.25
Const densitywater = 1025
Const frictioncoefficient = 0.001
Const gravityacceleration = 10
Const chezycoefficient = 50
Dim strinput As Variant, strmsg As String
If (Me!chkwind = True And Me!txtvelocity.value = 0) Then
    msg = "the velocity must be inserted "
    MsgBox msg
    Me!txtvelocity.SetFocus
    Me!chkwind = False
End If
If (Me!chkwind = True And Me!txtvelocity.value <> 0) Then
    strmsg = "insert the value of wind velocity"
    strinput = InputBox(prompt:=strmsg, title:="information")
    dischargevelocity = Me!txtvelocity
    m = (((densitywater * gravityacceleration) / (densityair *
frictioncoefficient)) ^ 0.5) / chezycoefficient
    m = m * dischargevelocity

    If ((strinput / m) <= 0.1) Then
        msg = "the wind effect can be neglected"
        MsgBox msg
        txtvelocity.SetFocus
        Me!txtwindvelocity.Visible = False
        Me!chkwind = False
    Else
        msg = "the wind effect must be included"
        MsgBox msg
        Me!txtwindvelocity.Visible = True
        Me!txtwindvelocity.value = strinput
    End If
End If

End If
End Sub

Private Sub cmdFirst_Click()
On Error GoTo Err_cmdFirst_Click

    DoCmd.GoToRecord , , acFirst

Exit_cmdFirst_Click:

```

```

Exit Sub

Err_cmdFirst_Click:
    MsgBox Err.Description
    Resume Exit_cmdFirst_Click

End Sub

Private Sub cmdPrev_Click()
On Error GoTo Err_cmdPrev_Click

    DoCmd.GoToRecord , , acPrevious

Exit_cmdPrev_Click:
    Exit Sub

Err_cmdPrev_Click:
    MsgBox Err.Description
    Resume Exit_cmdPrev_Click

End Sub
Private Sub cmdNext_Click()
On Error GoTo Err_cmdNext_Click

    DoCmd.GoToRecord , , acNext

Exit_cmdNext_Click:
    Exit Sub

Err_cmdNext_Click:
    MsgBox Err.Description
    Resume Exit_cmdNext_Click

End Sub
Private Sub cmdLast_Click()
On Error GoTo Err_cmdLast_Click

    DoCmd.GoToRecord , , acLast

Exit_cmdLast_Click:
    Exit Sub

Err_cmdLast_Click:
    MsgBox Err.Description
    Resume Exit_cmdLast_Click

End Sub

```

```

Private Sub Form_Close()
Dim db As Database
Dim rec As Recordset

Set db = CurrentDb()
Set rec = db.OpenRecordset("estuary", dbOpenDynaset)
If rec.Restartable = True Then

    rec.Requery
Else
    MsgBox "NO"
End If

End Sub

Private Sub Form_Current()
Dim recClone As Recordset
Dim intnewrecord As Integer, n As Integer
Dim msg As String
Set recClone = Me.RecordsetClone()
intnewrecord = IsNull(Me.id)
If intnewrecord Then
    cmdFirst.Enabled = True
    cmdNext.Enabled = False
    cmdPrev.Enabled = True
    cmdLast.Enabled = True
    Cmdnew.Enabled = False
    Me!chkinlets = False
    Me!Txtnumberinlets.Visible = False
    Me!chkwind = False
    Me!txtwindvelocity.Visible = False
    n = recClone.RecordCount
    Me!IDL.value = n + 1
Exit Sub
End If
Cmdnew.Enabled = True

If recClone.RecordCount = 0 Then
    cmdFirst.Enabled = False
    cmdNext.Enabled = False
    cmdPrev.Enabled = False
    cmdLast.Enabled = False
Else
    recClone.Bookmark = Me.Bookmark
    recClone.MovePrevious
    cmdFirst.Enabled = Not (recClone.BOF)
    cmdPrev.Enabled = Not (recClone.BOF)
    recClone.MoveNext

```

```

recClone.MoveNext
cmdLast.Enabled = Not (recClone.EOF)
cmdNext.Enabled = Not (recClone.EOF)
recClone.MovePrevious
End If
recClone.Close

If Me!chkinlets = False Then Me!Txtnumberinlets.Visible = False
If Me!chkinlets = True Then Me!Txtnumberinlets.Visible = True
If (Me!Txtnumberinlets.value = 0) Then
    Me!Txtnumberinlets.Visible = False
    Me!chkinlets = False
Else
    Me!Txtnumberinlets.Visible = True
End If
If Me!chkwind = False Then Me!txtwindvelocity.Visible = False
If Me!chkwind = True Then Me!txtwindvelocity.Visible = True
If (Me!txtwindvelocity.value = 0) Then
    Me!txtwindvelocity.Visible = False
    Me!chkwind = False
Else
    Me!txtwindvelocity.Visible = True
End If

End Sub

Private Sub Form_GotFocus()

Dim reciestuary As Recordset
Dim db As Database

Set db = CurrentDb()
Set reciestuary = db.OpenRecordset("estuary")

reciestuary.MoveFirst
Do
    reciestuary.Edit
    reciestuary!chkmodel = "0"
    reciestuary.Update
    reciestuary.MoveNext
Loop Until reciestuary.EOF

End Sub

Private Sub Form_Load()

```



```

Dim reciestuary As Recordset
Dim db As Database

Set db = CurrentDb()
Set reciestuary = db.OpenRecordset("estuary")

reciestuary.MoveFirst
Do
    reciestuary.Edit
    reciestuary!chkmodel = "0"
    reciestuary.Update
    reciestuary.MoveNext
Loop Until reciestuary.EOF

End Sub

Private Sub Form_Open(Cancel As Integer)
Dim db As Database
Dim rec As Recordset

Set db = CurrentDb()
Set rec = db.OpenRecordset("estuary", dbOpenDynaset)
If rec.Restartable = True Then
    rec.Requery
Else
    MsgBox "NO"
End If
Dim reciestuary As Recordset

Set db = CurrentDb()
Set reciestuary = db.OpenRecordset("estuary")

reciestuary.MoveFirst
Do
    reciestuary.Edit
    reciestuary!chkmodel = "0"
    reciestuary.Update
    reciestuary.MoveNext
Loop Until reciestuary.EOF

End Sub

```

```

Private Sub Form_Resize()
Dim reciestuary As Recordset
Dim db As Database

Set db = CurrentDb()
Set reciestuary = db.OpenRecordset("estuary")

reciestuary.MoveFirst
Do
    reciestuary.Edit
    reciestuary!chkmodel = "0"
    reciestuary.Update
    reciestuary.MoveNext
Loop Until reciestuary.EOF

End Sub

Private Sub Form_Unload(Cancel As Integer)
If Not IsNull(Me.OpenArgs) Then Forms(Me.OpenArgs).Visible = True
End Sub

Private Sub txtaveragedepth_AfterUpdate()
Call depth
Call qualitative
End Sub

Private Sub txtaveragedepth_Change()
Call depth
Call qualitative
End Sub

Private Sub txtaveragewidth_Change()
Call depth
End Sub

Private Sub txtfreshwater_AfterUpdate()
Call freshtidal
End Sub

```

```

Private Sub Txtnumberinlets_AfterUpdate()
Dim stDocName As String
Dim stLinkCriteria As String

stDocName = "frmMaintinlets"
stLinkCriteria = "[estuary]=" & "" & Me![txtname] & ""

DoCmd.OpenForm stDocName, , , stLinkCriteria

End Sub
Private Sub models_Click()
On Error GoTo Err_models_Click

Dim stDocName As String
Dim stLinkCriteria As String

stDocName = "frmMaintModels"

stLinkCriteria = "[estuary]=" & "" & Me![txtname] & ""
DoCmd.OpenForm stDocName, , , stLinkCriteria

Exit_models_Click:
Exit Sub

Err_models_Click:
MsgBox Err.Description
Resume Exit_models_Click

End Sub

Private Sub Cmdnew_Click()
On Error GoTo Err_Cmdnew_Click

DoCmd.GoToRecord , , acNewRec

Exit_Cmdnew_Click:
Exit Sub

Err_Cmdnew_Click:
MsgBox Err.Description
Resume Exit_Cmdnew_Click

End Sub

Private Sub cmdinlets_Click()
On Error GoTo Err_cmdinlets_Click

```

```

Dim stDocName As String
Dim stLinkCriteria As String

stDocName = "frmMaintinlets"

stLinkCriteria = "[estuary]=" & "" & Me![txtname] & ""
DoCmd.OpenForm stDocName, , , stLinkCriteria

Exit_cmdinlets_Click:
Exit Sub

Err_cmdinlets_Click:
MsgBox Err.Description
Resume Exit_cmdinlets_Click

End Sub

Private Sub txttidalperiod_AfterUpdate()
Call freshtidal
End Sub

Private Sub txttidalvolume_AfterUpdate()
Call freshtidal
End Sub

Private Sub txtvalley_length_AfterUpdate()
Call meandering
End Sub

Private Sub txtvalley_length_Change()
Call meandering
End Sub

Private Sub txtvalley_length_Exit(Cancel As Integer)
Call meandering
End Sub

Private Sub txtvelocity_AfterUpdate()
Dim msg As String
If (Me!txtvelocity.value = 0) Then
msg = "impossible to calculate the wind effect without the
velocity value"
MsgBox msg
txtvelocity.SetFocus
Else
Me!chkwind.value = True

```

```

    Call windii
End If

End Sub

Private Sub txtwindvelocity_Change()
'if you change the velocity the system is able to recalculated the
value of the
'effect of the wind and if it has to be considered
Call windii
End Sub

Public Sub windii()
Dim windvelocity As Single
Dim dischargevelocity As Single
Dim n As Single
Dim m As Single
Dim msg As String
Const densityair = 1.25
Const densitywater = 1025
Const frictioncoefficient = 0.001
Const gravityacceleration = 10
Const chezycoefficient = 50
Dim strinput As Variant, strmsg As String

    strmsg = "insert the value of wind velocity"
    strinput = InputBox(prompt:=strmsg, title:="information")
    dischargevelocity = Me!txtvelocity
    m = (((densitywater * gravityacceleration) / (densityair *
frictioncoefficient)) ^ 0.5) / chezycoefficient
    m = m * dischargevelocity

    If ((strinput / m) <= 0.1) Then
        msg = "the wind effect can be neglected"
        MsgBox msg
        txtvelocity.SetFocus
        Me!txtwindvelocity.Visible = False
        Me!chkwind = False
    Else
        msg = "the wind effect must be included"
        MsgBox msg
        Me!txtwindvelocity.Visible = True
        Me!txtwindvelocity.value = strinput
    End If
End Sub

```

```

End Sub

Private Sub togFilter_Click()
Dim frm As String
Dim strname As Form

inlets = Me!Txtnumberinlets
If togFilter Then
DoCmd.Minimize
DoCmd.OpenForm "frmselectionCriteria", , , , , acDialog,
(Me.name)
    If Me.FilterOn = True Then
        togFilter.Caption = "Model Applied"
    Else
        togFilter = False
    End If
Else
    Me.FilterOn = False
    togFilter.Caption = "Apply Model..."
End If

End Sub

Public Sub qualitative()
Dim wide As String
Dim flow As String
Dim salinity As String
Dim latitude As Integer

flow = Me.cbotidal_range
Select Case flow
Case "macrotidal"
If (Me!cboqualitative.value = "low") Or (Me!cboqualitative.value =
"moderate") Then
    Me!cbosalinity.value = "well mixed"
Else
    Me!cbosalinity.value = "vertically mixed"
End If
Case "mesotidal/macrotidal"
    If (Me!cboqualitative.value = "low") Or (Me!cboqualitative.value =
"moderate") Then
        Me!cbosalinity.value = "partially mixed"
    Else
        Me!cbosalinity.value = "vertically mixed"
    End If
Case "microtidal/mesotidal"
    If (cboqualitative.value = "low") Or (Me!cboqualitative.value =
"moderate") Then
        Me!cbosalinity.value = "partially mixed"
    End If

```

```

Else
    Me!cbosalinity.value = "vertically mixed"
End If
Case "microtidal"
    If (cboqualitative.value = "low") Then Me!cbosalinity.value =
    "salt wedge"
    If (Me!cboqualitative.value = "moderate") Then
        Me!cbosalinity.value = "salt wedge"
    Else
        Me!cbosalinity.value = "partially mixed"
    End If
End If

Me.Requery

End Select
cbosalinity.SetFocus
'determination of the coriolis force
'through the value of the width /depth ratio and the latitude

wide = Me!cboqualitative
latitude = Me!txtlatitude
If ((wide = "high" Or wide = "very high") And (latitude < 60 And
latitude > 50) And (flow <> "macrotidal")) Then
    Me.chkcoriolis.value = True
End If
End Sub

Public Sub depth()
Dim width As Variant
Dim depth As Variant
Dim max As Variant
Dim i As Integer
Dim j As Integer
Dim ratio(1 To 5) As Variant
Dim def(1 To 7) As String
Dim rat As Variant

width = Me!txtaveragewidth
depth = Me!txtaveragedepth
def(1) = "low"
def(2) = "moderate"
def(3) = "high"
def(4) = "very high"
def(5) = "low/moderate"
def(6) = "moderate/high"
def(7) = "high/very high"

rat = width / depth
ratio(1) = 1 / (1 + (Abs(rat - 10) / 90) ^ 2)

```

```

ratio(2) = 1 / (1 + (Abs(rat - 100) / 70) ^ 0.5)
Else
ratio(2) = 1 / (1 + (Abs(rat - 100) / 60) ^ 2)
End If

ratio(3) = 1 / (1 + (Abs(rat - 220) / 60) ^ 2)
ratio(4) = 1 / (1 + (Abs(rat - 400) / 100) ^ 2)
Me!Text219.value = ratio(1)
Me!Text221.value = ratio(2)

Me!Text223.value = ratio(3)
Me!Text225.value = ratio(4)
If ((rat < 10) Or (rat = 10)) Then
    Me!cboqualitative.value = "low"
    Me!Text219.value = 1
ElseIf (rat > 400) Then
    Me!cboqualitative.value = "very high"
    Me!Text225.value = 1
Else
    max = ratio(1)
    Me!cboqualitative.value = def(1)
    MsgBox Me!cboqualitative

    For j = 2 To 4
        If ratio(j) > max Then
            max = ratio(j)
            MsgBox max
            Me!cboqualitative.value = def(j)
            MsgBox Me!cboqualitative
            ElseIf ratio(j) = max Then
                If j = 2 Then Me!cboqualitative.value =
                If j = 3 Then Me!cboqualitative.value =
                If j = 4 Then Me!cboqualitative.value =
            MsgBox Me!cboqualitative
        End If
    Next

End If
End Sub

Public Sub meandering()
Dim channel As Variant
Dim valley As Variant
Dim i As Integer
Dim j As Integer
Dim prop(1 To 3) As Variant
Dim defi(1 To 6) As String
Dim propr As Variant

```

```

Dim max As Variant

channel = Me!txtChannel
valley = Me!txtvalley_length
defi(1) = "relatively straight"
defi(2) = "meandering"
defi(3) = "extremely meandering"
defi(4) = "relatively straight/meandering"
defi(5) = "meandering/extremely meandering"

propr = channel / valley
prop(1) = 1 / (1 + ((propr - 1.3) / 0.1) ^ 2)
prop(2) = 1 / (1 + ((propr - 1.6) / 0.1) ^ 2) ^ 0.5
prop(3) = 1 / (1 + ((propr - 1.9) / 0.1) ^ 2) ^ 2
Me!Text227.value = prop(1)
Me!Text231.value = prop(2)
Me!Text233.value = prop(3)
max = prop(1)
cbomeandering.value = defi(1)

    For j = 2 To 3
        If prop(j) > max Then
            Me!cbomeandering.value = defi(j)
            ElseIf prop(j) = max Then
                If j = 2 Then Me!cboqualitative.value =
defi(4)
                If j = 3 Then Me!cboqualitative.value =
defi(5)
            End If
        Next

    If propr < 1.3 Then
        cbomeandering.value = defi(1)
        Me!Text227.value = 1
        End If
    If propr > 1.9 Then
        cbomeandering.value = defi(3)
        Me!Text233.value = 1
        End If

End Sub

Public Sub freshtidal()
Dim fresh As Variant
Dim tidal As Variant
Dim rv As Variant
Dim strmsg As String
Dim title As String
Dim period As Variant

```

```

fresh = Me!txtfreshwater
tidal = Me!txttidalvolume
period = Me!txttidalperiod
If ((tidal <> 0) And (fresh <> 0) And (period <> 0)) Then
rv = (fresh * period) / tidal
If (rv > 1) Then Me!cbosalinity.value = "salt wedge"
If ((rv < 1) Or (rv = 1)) And (rv > 0.1)) Then
Me!cbosalinity.value = "partially mixed"
strmsg = "fresh water discharge and tidal voume typical of a fjord
estuary.Check the geomorphological type"
title = "suggestion"
MsgBox strmsg, vbOKOnly, title
End If
If ((rv <= 0.1) And (rv > 0.005)) Then Me!cbosalinity.value =
"partially mixed"
If (rv < 0.005) Then
If (Me!cboqualitative = "high" Or Me!cboqualitative = "very high")
Then
Me!cbosalinity.value = "vertically mixed"
End If
End If
End If
End Sub

```

Option Compare Database

Option Explicit

Private Sub average\_velocity\_BeforeUpdate(Cancel As Integer)

```

Dim strFormName As Form
Dim strvelocity As Variant
Dim n As Variant

```

```

Set strFormName = Forms!frmmainestuary
strvelocity = strFormName.txtvelocity
n = Me.velocity
n = strvelocity / n
If (n >= 5) Then
Me!cbonatureinlet = "branch"
Else
Me!cbonatureinlet = "tributary"
End If

```

End Sub

```

Private Sub Form_Load()

Dim db          As Database
Dim rec         As Recordset
Dim reci       As Recordset
Dim rex        As Recordset
Dim strFormName As Form
Dim counter    As Integer
Dim strname    As Variant
Dim straname   As String
Dim strmainame As String
Dim strmsg     As String
Dim n          As Integer
Dim m          As Integer
Set strFormName = Forms!frmmaintestuary
n = strFormName.Txtnumberinlets
strname = Forms!frmmaintestuary!txtname
Set db = CurrentDb()
Set rec = db.OpenRecordset("Inlets")
Set rex = db.OpenRecordset("Inlets", dbOpenDynaset)
m = 0
Set reci = db.OpenRecordset("Inlets", dbOpenSnapshot)
straname = "[Estuary]= '" & strname & "'"
If Not IsNull(n) Then
    reci.FindFirst straname
    Do While reci.NoMatch = False
        m = m + 1
        reci.FindNext straname
    Loop
    If (n < m) Then
        m = m - n

        For counter = 1 To m
            If rex.NoMatch = False Then
                rex.FindLast straname
                rex.Delete
                rex.MovePrevious

            End If
            rex.Requery
        Next

    Else
        If IsEmpty(straname) = True Then
            counter = 0
            rec.Edit
            Do While counter < n
                counter = counter + 1
                rec.AddNew

```

```

                rec("ID1") = counter
                rec("estuary") = straname
                rec.Update
            Loop
            rec.Bookmark = rec.LastModified

        Else
            reci.FindFirst straname
            Do While reci.NoMatch = False
                counter = counter + 1
                reci.FindNext straname
            Loop
            Do While counter < n
                counter = counter + 1
                rec.AddNew
                rec("ID1") = counter
                rec("estuary") = straname
                rec.Update
            Loop

        End If
    End If
End If
Me.Requery
End Sub

Private Sub average_depth_Change()
Dim strFormName As Form
Dim strvelocity As Variant
Dim n            As Variant

Set strFormName = Forms!frmmaintestuary
strvelocity = strFormName.txtvelocity
n = Me.velocity
n = strvelocity / n
If (n >= 10) Then
Me!cbonatureinlet = "branch"
Else
Me!cbonatureinlet = "tributary"
End If

End Sub

```

## Retrieval Component

Estuary Similarity

Option Compare Database  
Option Explicit

```
Private Sub cboproblem_Change()
Dim strwhere As String
strwhere = Me!cboproblem
Select Case strwhere
    Case "pollutant distribution"
        Me!cbopollutantnature.Visible = True
End Select
End Sub

Public Sub selectestuary(intrelqual() As String, relquali() As
Variant, simivalue As String, value As Integer, strsql As String)
Dim db As Database
Dim rec As Recordset
Dim rex As Recordset
Dim recproblem As String
Dim recpollutant As String
Dim recestuary As String
Dim interrecords As Integer
Dim strproblem As String
Dim counter As Integer
Dim method As String
Dim dimestuary As String
Dim strpurpose As String
Dim strSQLA As String
Dim inrel() As Variant
Dim n As Integer
Dim cut As Variant
If Not IsNull(Me.cboproblem) Then
    strproblem = Me.cboproblem
    txtproblem = strproblem
    strpurpose = Me.cbopurpose
    txtpurpose = strpurpose
    estu = Me.estuary
End If
Set db = CurrentDb()
```

```
Set rec = db.OpenRecordset("models")
interrecords = rec.RecordCount
numbrec = interrecords

rec.MoveFirst
ReDim intrelqual(1 To interrecords)
ReDim relquali(1 To interrecords)
ReDim inrel(1 To interrecords)
'comparison among models which simulate the undergoing problem
value = 0
'evaluation of different problems and cases for which the stored
models have been applied
Select Case strproblem
    Case "salinity distribution", "pollutant distribution"
        'the pollution distribution in case of conservative is
        considered a similar case to salinity distribution
        For counter = 1 To interrecords
            inrel(counter) = 0
            recproblem = rec("problem")
            If recproblem = "pollutant distribution" Then
                recpollutant = rec("pollutant")
            If (recproblem = "pollutant distribution" And
                recpollutant = "conservative") Or (recproblem = "salinity
                distribution") Then
                    recestuary = rec("estuary")
                    value = value + 1
                    intrelqual(value) = recestuary
            End If
            rec.MoveNext
        Next

    Case "salt intrusion"
        For counter = 1 To interrecords
            recproblem = rec("problem")
            If (recproblem = "salt intrusion") Then
                recestuary = rec("estuary")
                value = value + 1
                intrelqual(value) = recestuary
            End If
            rec.MoveNext
        Next
End Select
```

```
totvalue = value
rec.Close
'comparison among the estuaries associated to the selected models
```

```
selectestuarytype intrqual, relquali, simivalue, value
cut = Me.cut_off
strsql = "ratio > " & cut
```

```
ctestuary_Exit:
Exit Sub
```

```
selectestuary_Err:
MsgBox Err.Description
```

```
End Sub
```

```
Public Sub selectestuarytype(intrqual() As String, relquali() As
Variant, simivalue As String, value As Integer)
'the routine is used to evaluated the similarity between the estuary
underconsideration and the estuaries stored
' in the system's library
```

```
Dim db As Database
Dim reci As Recordset
Dim rex As Recordset
Dim rec As Recordset
Dim widthdepth As Variant
Dim id As Integer
Dim recid As Integer
Dim counter As Integer
Dim intrecords As Integer
Dim relquale() As Variant
Dim relqual() As Variant
Dim strform As Form
Dim strname As String
Dim recname As String
Dim strqualitative As String
Dim recwidthdepth As String
Dim rectotalarea As Variant
Dim strwhere As String
Dim recintertidalarea As Variant
Dim ratarea As Single
Dim recratarea As Single
Dim meandering As String
Dim recmeandering As String
Dim tidalrange As String
Dim rectidalrange As String
Dim tid As Integer
Dim channel As Single
Dim avrdepth As Single
Dim recavrdepth As Single
Dim recchannel As Single
```

```
Dim sha As Single
Dim recsha As Single
Dim criterion As String
Dim rectid As Integer
Dim estutype As String
Dim rectype As String
Dim n As Variant
Dim m As Integer
Dim inrel() As Integer
Dim ma As Integer
Dim mea As Integer
Dim recmea As Variant
Dim st As Integer
Dim sd As Integer
Dim max As Single
Dim le As Variant
Dim mx As Variant
Dim mn As Variant
Dim strcriteri As String
Dim rexratio As Variant
Dim strsql As String
```

```
Dim totalarea As Variant
Dim intertidalarea As Variant
Dim i As Integer
Dim j As Integer
Dim ratio(1 To 6) As Variant
Dim rat As Variant
Dim qualitative As String
Dim maxbank As Variant
Dim ratmaxbank As Variant
Dim recratmakbank As Variant
Dim recnome As String
Dim com As Variant
Dim total As Variant
Dim rati(1 To 4) As Variant
Dim recrati(1 To 4) As Variant
Dim mearati(1 To 3) As Variant
Dim recmearati(1 To 3) As Variant
```

```
Dim sumrat As Variant
Set db = CurrentDb()
Set reci = db.OpenRecordset("estuary")
Set rec = db.OpenRecordset("estuary")
Set strform = Forms!frmmaintestuary
ReDim inrel(1 To value)
```

```
'qualification of the estuary under consideration
widthdepth = strform!cboqualitative
totalarea = strform!txttotal
intertidalarea = strform!txtintertidal
ratarea = intertidalarea / totalarea
```



```

channel = strform!txtChannel
avrdepth = strform!txtaveragedepth
sha = (channel) / avrdepth
meandering = strform!cbomeandering
maxbank = strform!txtbankarea
rati(1) = strform!Text219
rati(2) = strform!Text221
rati(3) = strform!Text223
rati(4) = strform!Text225
mearati(1) = strform!Text227
mearati(2) = strform!Text231
mearati(3) = strform!Text233
Dim recmaxbank As Variant
Dim recratmaxbank As Variant
total = 0
Select Case widthdepth
    Case "low"
        sd = 1
    Case "moderate"
        sd = 2
    Case "high"
        sd = 3
    Case "very high"
        sd = 5
End Select

estutype = strform!cbogeomorphological
Select Case estutype
    Case "fjord", "fjard", "ria"
        ma = 1
    Case "coastal plain", "complex"
        ma = 2
    Case "bar built"
        ma = 3
End Select

Select Case meandering
    Case "relatively straight"
        mea = 1
    Case "meandering"
        mea = 2
    Case "extremely meandering"
        mea = 3
End Select

tidalrange = strform!cbotidal_range
Select Case tidalrange
    Case "microtidal"
        tid = 1
    Case "microtidal/mesotidal"
        tid = 2
    Case "mesotidal/macrotidal"
        tid = 3

```

```

        Case "macrotidal"
            tid = 4
    End Select
    id = strform!ID1
    intrecords = reci.RecordCount
    ReDim relquali(1 To value)
    reci.MoveFirst
    ReDim relquale(1 To value)
    ReDim relqual(1 To value)
    For i = 1 To value
        relquale(i) = intrrelqual(i)
        relqual(i) = intrrelqual(i)
    Next

    For i = 1 To value
        If (relquale(i) <> "0") Then
            inrel(i) = 1
            n = i + 1
            If (n > value) Then
                Exit For
            Else
                For j = n To value
                    If relquale(j) = relquale(i) Then
                        relquale(j) = "0"
                        inrel(i) = inrel(i) + 1
                    End If
                Next
            End If
        Next
    End If

    Next
End If

Select Case txtproblem
    Case "salinity distribution"
    For n = 1 To value
        total = 0
        If relquale(n) = "0" Then
            relquali(n) = 0
        Else

            reci.Index = "name"
            reci.Seek "=", relquale(n)

            recnome = reci("name")
            MsgBox recnome
            recwidthdepth = reci("widthdepratio")
            rectype = reci("geomorphological type")
            recchannel = reci("channel length")
            recavrdepth = reci("average depth")
            recmeandering = reci("meandering")
            rectotalarea = reci("total area")
            recintertidalarea = reci("intertidal area")

```

```

rectidalrange = reci("tidal range")
recratarea = recintertidalarea / rectotalarea
recsha = (recchannel) / recavrdepth
recmaxbank = reci("max channel bank area")
recrati(1) = reci("ratio1")
recrati(2) = reci("ratio2")
recrati(3) = reci("ratio3")
recrati(4) = reci("ratio4")
recmearati(1) = reci("mearat1")
recmearati(2) = reci("mearat2")
recmearati(3) = reci("mearat3")
sumrat = (1 - ((Abs(recrati(1) - rati(1)) + Abs(recrati(2)
- rati(2)) + Abs(recrati(3) - rati(3)) + Abs(recrati(4) - rati(4)))
/ 4))
ratio(1) = sumrat
MsgBox sumrat

total = total + 1
relquali(n) = ratio(1) * 1
If (estutype = rectype) Then
    ratio(2) = 1
Else
    Select Case rectype
        Case "fjord", "fjard", "ria"
            m = 1
        Case "coastal plain", "complex"
            m = 2
        Case "bar built"
            m = 3
    End Select
    If (Abs(ma - m) = 0) Then ratio(2) = 0.8
    If (Abs(ma - m) = 1) Then ratio(2) = 0.4
    If (Abs(ma - m) = 2) Then ratio(2) = 0.1

End If
total = total + 0.75
relquali(n) = relquali(n) + ratio(2) * 0.75
ratio(3) = (1 + (Abs(ratarea - recratarea) / (ratarea) ^ 0.5)
^ 2) ^ -1
relquali(n) = relquali(n) + ratio(3) * 0.5
total = total + 0.5
If (maxbank <> 0) Then
    ratmaxbank = maxbank / intertidalarea
    If (recmaxbank <> 0) Then
        recratmaxbank = recmaxbank / recintertidalarea
        ratio(3) = (1 + (Abs(ratmaxbank - recratmaxbank) /
(ratmaxbank) ^ 0.5) ^ 2) ^ -1
        relquali(n) = relquali(n) + ratio(3) * 0.5
        total = total + 0.5
    End If

```

```

End If

sumrat = (1 - ((Abs(recmearati(1) - mearati(1)) +
Abs(recmearati(2) - mearati(2)) + Abs(recmearati(3) - mearati(3))) /
3))
ratio(4) = sumrat
total = total + 0.5
relquali(n) = relquali(n) + ratio(4) * 0.5
Select Case rectidalrange
    Case "microtidal"
        rectid = 1
    Case "microtidal/mesotidal"
        rectid = 2
    Case "mesotidal/macrotidal"
        rectid = 3
    Case "macrotidal"
        rectid = 4
End Select
rat = 1 - (Abs(tid - rectid) / 3)
If rat = 1 Then com = 1
If rat = 2 / 3 Then com = 0.666
If rat = 1 / 3 Then com = 0.33
If rat = 0 Then com = 0
ratio(5) = com
total = total + 0.75
relquali(n) = relquali(n) + ratio(5) * 0.75
ratio(6) = (1 + (Abs(sha - recsha) / (sha) ^ 0.5) ^ 2) ^ -1
total = total + 0.5
relquali(n) = relquali(n) + ratio(6) * 0.5
relquali(n) = relquali(n) / total
End If

Next
reci.Close
Case "salt intrusion"
n = 0
For n = 1 To value
    relquali(n) = 0
Next
For n = 1 To value
    If relquale(n) = "0" Then
        relquali(n) = 0
    Else
        Dim recmaxmin As Variant
        Dim recmax As Variant
        Dim recmin As Variant
        Dim reclength As Variant
        Dim maria As Variant
        rec.Index = "name"
        rec.Seek "=", relquale(n)
        recname = rec("name")
    End If

```

```

MsgBox recname
recwidthdepth = rec("widthdepratio")
rectype = rec("geomorphological type")
rectidalrange = rec("tidal range")
reclength = rec("channel length")
recavrdepth = rec("average depth")
recsha = reclength / recavrdepth
recrati(1) = rec("ratio1")
recrati(2) = rec("ratio2")
recrati(3) = rec("ratio3")
recrati(4) = rec("ratio4")

sumrat = (1 - ((Abs(recrati(1) - rati(1)) +
Abs(recrati(2) - rati(2)) + Abs(recrati(3) - rati(3)) +
Abs(recrati(4) - rati(4))) / 4))
ratio(1) = sumrat
MsgBox sumrat
relquali(n) = ratio(1)

If (estutype = rectype) Then
    ratio(2) = 1
Else
    Select Case rectype
        Case "fjord", "fjard", "ria"
            m = 1
        Case "coastal plain", "complex"
            m = 2
        Case "bar built"
            m = 3
    End Select
    rat = Abs(ma - m)
    If ma = m Then ratio(2) = 0.8
    If (rat = 1) Then ratio(2) = 0.4
    If (rat = 2) Then ratio(2) = 0.2
End If
ratio(2) = ratio(2) * 1
relquali(n) = relquali(n) + ratio(2)
Select Case rectidalrange
    Case "microtidal"
        rectid = 1
    Case "microtidal/mesotidal"
        rectid = 2
    Case "mesotidal/macrotidal"
        rectid = 3
    Case "macrotidal"
        rectid = 4
End Select
rat = 1 - (Abs(tid - rectid) / 3)
If rat = 1 Then ratio(3) = 1
If rat = 2 / 3 Then ratio(3) = 0.666
If rat = 1 / 3 Then ratio(3) = 0.33
If rat = 0 Then ratio(3) = 0

```

```

relquali(n) = relquali(n) + ratio(3) * 1
ratio(4) = (1 + (Abs(sha - recsha) / (sha) ^ 0.5) ^
2) ^ -1
relquali(n) = relquali(n) + ratio(4) * 0.5
relquali(n) = relquali(n) / 3.5
End If

Next
rec.Close
End Select

Set rex = db.OpenRecordset("estuary", dbOpenDynaset)

If rex.Updatable = True Then
For counter = 1 To value
    If ((relquali(counter) > 0) And (inrel(counter) > 0)) Then
        rex.FindFirst "name = " & "" & relquale(counter) & ""
        Do While rex.NoMatch = False
            n = rex("ratio")
            rex.Edit

            rex("ratio") = relquali(counter)
            rex("modnumber") = inrel(counter)
            rex.Update
            rex.FindNext "name = " & "" & intrelqual(counter) & ""

            Loop
        End If
    Next
    rex.Close
    Set db = Nothing

End If

n = 0
max = 0
For m = 1 To value

    If relquali(m) > max Then
        max = relquali(m)

        n = m
    End If

Next

simivalue = relquale(n)

```

```

selectestuarytype_Exit:
    Exit Sub

selectestuarytype_Err:
    MsgBox Err.Description
    Resume selectestuarytype_Exit

End Sub
Private Sub cmdCancel_Click()
Forms(Me.OpenArgs).Visible = True
DoCmd.Close acForm, Me.name
End Sub

Private Sub cmdOK_Click()

Dim strSQL As String
Dim simivalue As String
Dim sensor As Integer
Dim message As String
Dim dimestuary As String
Dim intrelqual() As String
Dim relquali() As Variant
Dim value As Integer
Dim qdf As QueryDef
Dim lngRecords As Long
Dim strcriteria As String
Dim strMessage As String
Dim intResponse As Integer
Dim frm As Form
Dim frmA As Form
If IsNull(cboproblem) Then
    MsgBox "you must enter the problem descriptor", vbExclamation
    Exit Sub
End If
If IsNull(cut_off) Then
    MsgBox "you must enter a value for the cut off parameter",
vbExclamation
    Exit Sub
End If

selectestuary intrelqual, relquali, simivalue, value, strSQL

If makeatable1(vRecordcount:=lngRecords, vmethod:="DAO") Then

Select Case lngRecords
Case 0
    MsgBox "no records matched the criteria you specified", _
vbExclamation, "no matches found"

Case Else

```

```

        strMessage = lngRecords & " record(s)matched the criteria "
        selection intrelqual, relquali, simivalue, value
        DoCmd.OpenForm "matchestuary", , , , ,
(Me.OpenArgs)
        DoCmd.Maximize

        Set frm = Forms!matchestuary

        frm.Filter = strSQL
        frm.FilterOn = True
        DoCmd.Close acForm, Me.name

End Select
End If

End Sub

```

```

Private Sub Form_Current()

Dim strform As Form
Dim strname As String
Dim strbranch As String
Dim inlet As Integer
Dim strwind As Variant
Dim db As Database
Dim rec As Recordset
Dim strcriteria As String
Dim counter As Integer

Set db = CurrentDb()

Set strform = Forms!frmmaintestuary
strname = strform!txtname
Set rec = db.OpenRecordset("inlets", dbOpenDynaset)
rec.MoveFirst
counter = 0
inlet = strform!txtnumberinlets
If inlet > 0 Then
    strcriteria = "[Estuary] = '" & strname & "'"
    rec.FindFirst strcriteria
    Do While rec.NoMatch = False
        strbranch = rec!natureinlet

        If strbranch = "branch" Then
            counter = 1
            Me!branching = True
        End If
        rec.FindNext strcriteria
    Loop

```

```

End If

If strform!chkcoriolis = True Then Me!coriolis = True
If strform!chkwind = True Then Me!wind = True
If strform!cbobed = "sloping bed" Then Me!slope = True

If IsNull(Me.OpenArgs) Then
    MsgBox "This form must be opened from within another form",
vbCritical, "Cannot Open Form"
    DoCmd.Close acForm, (Me.name)
End If
End Sub

Public Function makeatable1(Optional vTableName = "similarity", _
    Optional vRecordcount, _
    Optional vtimetaken, _
    Optional vmethod = "DAO") _
As Boolean

Dim strSQLA As String
Dim lngTime As Long
Dim intFromPos As Integer
Dim qdf As QueryDef
Dim db As Database
Dim fld As Field
Dim fld1 As Field
Dim tbl As TableDef
Dim idx As Index

Select Case vmethod
    Case "DAO"
        Set db = CurrentDb()

        CurrentDb.TableDefs.Delete vTableName
        Set tbl = db.CreateTableDef(vTableName)

        Set fld = tbl.CreateField("name", dbText)
        fld.OrdinalPosition = 1
        fld.Size = 50
        fld.Required = True
        tbl.Fields.Append fld
        db.TableDefs.Append tbl
        RefreshDatabaseWindow

```

```

'Now add the data
strSQLA = "INSERT INTO [" & vTableName & "]SELECT
name FROM [estuary]"

'And run the insert query to add the data
Set qdf = CurrentDb.CreateQueryDef("")
qdf.SQL = strSQLA
qdf.Execute

'Return the number of records in the table if needed
If Not IsMissing(vRecordcount) Then _
    vRecordcount = qdf.RecordsAffected

'Close the querydef
qdf.Close
RefreshDatabaseWindow
'indicate the creation of the table was successful
If Not IsMissing(vtimetaken) Then vtimetaken = Now - lngTime
makeatable1 = True

MsgBox "The " & tbl.name & " table was successfully created"
End Select
makeatable_Exit:
Exit Function

makeatable_Err:
Select Case Err.Number
    Case 3265 'Occurs if the table we are deleting
doesn't exist
        Resume Next
    Case Else
        MsgBox Err.Description
        Resume makeatable_Exit
        Resume
End Select

End Function

Public Sub selection(intrelqual() As String, relquali() As Variant,
simivalue As String, value As Integer)
Dim rec As Recordset
Dim db As Database
Dim counter As Integer

Set db = CurrentDb()
Set rec = db.OpenRecordset("similarity", dbOpenDynaset)

Do

```

```
rec.Edit
rec.Delete
rec.MoveNext
Loop Until rec.EOF
```

```
Set db = Nothing
```

```
If rec.Updatable = True Then
```

```
For counter = 1 To value
```

```
  If (relqual(counter) <> 0) Then
```

```
    rec.AddNew
```

```
      rec("name") = intrelqual(counter)
```

```
      rec.Update
```

```
  End If
```

```
Next
```

```
rec.Close
```

```
Set db = Nothing
```

```
End If
```

```
End Sub
```

```
Private Sub Form_Unload(Cancel As Integer)
```

```
If Me.branching = True Then
```

```
  txtbranching = 1
```

```
Else
```

```
  txtbranching = 0
```

```
End If
```

```
MsgBox txtbranching
```

```
End Sub
```

MODEL SIMILARITY

Option Compare Database  
Option Explicit

```
Private Sub cmdSelect_Click()
Dim pref As Variant
Dim rec As Recordset
Dim resim As Variant
Dim db As Database
Set db = CurrentDb()
Set rec = db.OpenRecordset("models")
rec.MoveFirst
Do
rec.Edit
resim = rec("simodel")

rec.Update
rec.MoveNext
Loop Until rec.EOF
```

```
pref = Me!txtpreference
pref = "estuary = " & "'" & Me!txtestuary & "'" & " AND problem = " & "'" & txtproblem & "'" & " AND dimension = " & "'" & Me!txtdimension & "'" & " AND accuracy = " & "'" & Me!txtaccuracy & "'"
MsgBox pref
DoCmd.OpenForm "frmmaintmodels", acNormal, , pref, , acDialog
```

End Sub

```
Private Sub Form_Load()
```

```
DoCmd.Maximize
End Sub
```

```
Private Sub Form_Open(Cancel As Integer)
```

```
Dim rec As Recordset
Dim db As Database
Dim recaccuracy As String
```

```
Dim rectime As String
Dim recpurpose As String
Dim recsim As Single
Dim recproblem As String
Dim recestuary As String
Dim sim() As Variant
Dim ratio(3) As Variant
Dim n As Integer
Dim total As Single
Dim risim As Variant
Dim criteria As String
Dim io As Integer
ReDim sim(1 To totvalue)
Set db = CurrentDb()
Set rec = db.OpenRecordset("models", dbOpenDynaset)
n = 0
```

```
rec.MoveFirst
For io = 1 To numbrec
recproblem = rec("problem")
If (recproblem = txtproblem) Then
n = n + 1
sim(n) = 0
total = 0
recestuary = rec("estuary")
MsgBox recestuary
```

```
recaccuracy = rec("accuracy")
rectime = rec("time")
recpurpose = rec("purpose")
If (txtpurpose = "management tool") Then
If recpurpose = "management tool" Then ratio(1) = 1
If recpurpose = "water quality" Then ratio(1) = 0.7
If recpurpose = "research" Then ratio(1) = 0.5
End If
If (txtpurpose = "water quality") Then
If recpurpose = "management tool" Then ratio(1) = 0.5
If recpurpose = "water quality" Then ratio(1) = 1
If recpurpose = "research" Then ratio(1) = 0.8
```

```
End If
If (txtpurpose = "research") Then
If recpurpose = "management tool" Then ratio(1) = 0.1
If recpurpose = "water quality" Then ratio(1) = 0.5
If recpurpose = "research" Then ratio(1) = 1
End If
sim(n) = ratio(1) * 0.75
MsgBox sim(n)
total = total + 0.75
```

```
If (txtpurpose = "management tool") Then
If recaccuracy = "high" Then ratio(2) = 1
```

```

        If recaccuracy = "moderate" Then ratio(2) = 0.8
        If recaccuracy = "overall" Then ratio(2) = 0.5
    End If
    If (txtpurpose = "water quality") Then
        If recaccuracy = "high" Then ratio(2) = 1
        If recaccuracy = "moderate" Then ratio(2) = 0.7
        If recaccuracy = "overall" Then ratio(2) = 0.4
    End If
    If (txtpurpose = "research") Then
        If recaccuracy = "high" Then ratio(2) = 1
        If recaccuracy = "moderate" Then ratio(2) = 0.6
        If recaccuracy = "overall" Then ratio(2) = 0.2
    End If
    sim(n) = sim(n) + ratio(2) * 0.5

total = total + 0.5
    If (txtpurpose = "management tool") Then
        If rectime = "low" Then ratio(3) = 1
        If rectime = "moderate" Then ratio(3) = 0.6
        If rectime = "high" Then ratio(3) = 0.2
    End If
    If (txtpurpose = "water quality") Then
        If rectime = "low" Then ratio(3) = 1
        If rectime = "moderate" Then ratio(3) = 0.7
        If rectime = "high" Then ratio(3) = 0.4
    End If
    If (txtpurpose = "research") Then
        If rectime = "low" Then ratio(3) = 1
        If rectime = "moderate" Then ratio(3) = 0.8
        If rectime = "high" Then ratio(3) = 0.6
    End If
    sim(n) = sim(n) + ratio(3) * 0.5

    total = total + 0.5
    rec.Edit
    sim(n) = sim(n)
    rec("simodel") = sim(n)
    rec.Update
End If
rec.MoveNext
Next
rec.Requery

Dim reci As Recordset
Dim branch As Integer
Dim branchnumber As Integer
Dim interecords As Integer
Dim i As Integer
Dim resim As Single

```

177

```

Dim strcriteria As String
Dim strcriter As String
Dim inletve As Variant
Dim inletnature As String
Dim strci As String
Dim salimit As Variant
Dim inlet As Recordset
Set db = CurrentDb()
Dim estutype As String
    Dim salinity As String
    Dim strform As Form
    Dim tidalrange As String
    Dim widthdepth As String
    Dim totalarea As Single
    Dim intertidalarea As Single
    Dim meandering As String
    Dim purpose As String
    Dim modsim As Single
    Dim intotal As Variant
    Set strform = Forms!frmmainestuary
    Set inlet = db.OpenRecordset("inlets", dbOpenDynaset)
    estutype = strform!cbogeomorphological
    salinity = strform!cboosalinity
    tidalrange = strform!cbotidal_range
    widthdepth = strform!cboqualitative
    totalarea = strform!txttotal
    intertidalarea = strform!txtintertidal
    meandering = strform!meandering
    salimit = strform!txtsalinitylimit

intotal = 0
Set reci = db.OpenRecordset("retrievedmodels", dbOpenDynaset)
If txtbranching <> 0 Then
    strcriteria = "[estuary] = " & "'" & estu & "'"
    MsgBox strcriteria
    inlet.FindFirst strcriteria
    Do While inlet.NoMatch = False
        inletnature = inlet("natureinlet")
        If (inletnature = "branch") Then intotal =
intotal + 1
        inlet.FindNext strcriteria
    Loop

    If intotal >= 1 Then
        strcriteria = "[dimension] = '1-D'"
        reci.FindFirst strcriteria
        Do While reci.NoMatch = False
            reci.Edit
            reci.Delete
            reci.FindNext strcriteria
        Loop
    End If
End If

```



```

End If
Else
  strcriteria = "[dimension] = '1-D network'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit
    reci.Delete
    reci.FindNext strcriteria
  Loop
End If

Select Case txtproblem
Case "salinity distribution"
  If (txtpurpose = "research" Or txtpurpose = "waterquality") Then
  If ((txtbranching <> 0) And (salimit > 0)) Then
    strcriteria = "[estuary]=" & "" & estu & ""
    reci.FindFirst strcriteria
    inlet.FindFirst strcriteria
    Do While inlet.NoMatch = False
      inletve = inlet("distance")
      inletnature = inlet("natureinlet")
      If (salimit > inletve) And (inletnature = "branch") Then
        strcriter = "[dimension] = '2-D moving'"
        reci.FindFirst strcriter
        Do While rec.NoMatch = False
          reci.Edit
          resim = reci("simodel")
          resim = (resim + 1)
          reci("simodel") = resim
          total = total + 1
          reci.Update
          reci.FindNext strcriter
        Loop
      End If
      inlet.FindNext strcriteria
    Loop
  End If
End If
Select Case estutype
Case "fjord"
  strcriteria = "[dimension] = '3-D'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit
    reci.Delete
    reci.FindNext strcriteria
  Loop
  strcriteria = "[dimension] = '2-D moving'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit

```

```

    reci.Delete
    reci.FindNext strcriteria
  Loop
  If (salinity = "partially mixed") Or (widthdepth = "low" Or
widthdepth = "moderate") Then
    strcriter = "[dimension] = '2-D avrg on width'"
    reci.FindFirst strcriter
    Do While reci.NoMatch = False
      reci.Edit
      resim = reci("simodel")
      resim = (resim + 1)
      reci("simodel") = resim
      total = total + 1
      reci.Update
      reci.FindNext strcriter
    Loop
  End If
  If (salinity = "partially mixed") Or (widthdepth = "high"
Or widthdepth = "very high") Then
    strcriter = "[dimension] = '2-D avrg on depth'"
    reci.FindFirst strcriter
    Do While reci.NoMatch = False
      reci.Edit
      resim = reci("simodel")
      resim = (resim + 1)
      reci("simodel") = resim
      total = total + 1
      reci.Update
      reci.FindNext strcriter
    Loop
  End If
  If ((salinity = "well mixed") And intotal = 0) Then
    strcriter = "[dimension] = '1-D'"
    reci.FindFirst strcriter
    Do While reci.NoMatch = False
      reci.Edit
      resim = reci("simodel")
      resim = (resim + 1)
      reci("simodel") = resim
      total = total + 1
      reci.FindNext strcriter
    Loop
    ElseIf (((tidalrange = "macrotidal") Or (salinity =
"vertically mixed" Or salinity = "well mixed")) And (intotal >= 1))
Then
      strcriteria = "[dimension] = '1-D network'"
      reci.FindFirst strcriteria
      Do While reci.NoMatch = False
        reci.Edit

```

```

        resim = reci("simodel")
        resim = (resim + 1)
        reci("simodel") = resim
        total = total + 1
        reci.Update
        reci.FindNext strcriteria
    Loop
End If

Case "ria", "fjard"
    strcriteria = "[dimension] = '2-D moving'"
    reci.FindFirst strcriteria
    Do While reci.NoMatch = False
        reci.Edit
        reci.Delete

        reci.FindNext strcriteria
    Loop
    strcriteria = "[dimension] = '3-D'"
    reci.FindFirst strcriteria
    Do While reci.NoMatch = False
        reci.Edit
        reci.Delete

        reci.FindNext strcriteria
    Loop
    If (salinity = "partially mixed") And (widthdepth = "low" Or
widthdepth = "moderate") Then
        strcriter = "[dimension] = '2-D avrg on width'"
        reci.FindFirst strcriter
        Do While reci.NoMatch = False
            reci.Edit
            resim = reci("simodel")
            resim = (resim + 1)
            reci("simodel") = resim
            total = total + 1
            reci.Update
            reci.FindNext strcriter
        Loop
    End If

    If (salinity = "vertically mixed" Or salinity = "well mixed") And
(widthdepth = "low" Or widthdepth = "moderate") Then
        strcriter = "[dimension] = '2-D avrg on depth'"
        reci.FindFirst strcriter
        Do While reci.NoMatch = False
            reci.Edit
            resim = reci("simodel")
            resim = (resim + 1)
            total = total + 1
            reci("simodel") = resim
            total = total + 1

```

```

        reci.Update
        reci.FindNext strcriter
    Loop
End If

If (widthdepth = "high" Or widthdepth = "very high") Then
    strcriteria = "[dimension] = '2-D avrg on width'"
    reci.FindFirst strcriteria
    reci.FindFirst strcriteria
    Do While reci.NoMatch = False
        reci.Edit
        reci.Delete
        reci.FindNext strcriteria
    Loop
    End If
    If (salinity = "vertically mixed" Or salinity = "well
mixed") And (widthdepth = "low" Or widthdepth = "moderate") Then
        If intotal >= 1 Then
            strcriter = "[dimension] = '1-D network'"
            reci.FindFirst strcriter
            Do While reci.NoMatch = False
                reci.Edit
                resim = reci("simodel")
                resim = (resim + 1)
                total = total + 1
                reci("simodel") = resim
                total = total + 1
                reci.Update
                reci.FindNext strcriter
            Loop
        Else
            strcriter = "[dimension] = '1-D'"
            reci.FindFirst strcriter
            Do While reci.NoMatch = False
                reci.Edit
                resim = reci("simodel")
                resim = (resim + 1)
                reci("simodel") = resim
                total = total + 1
                reci.Update
                reci.FindNext strcriter
            Loop
        End If
    End If

Case "coastal plain", "complex"
    If (salinity = "well mixed" Or salinity = "vertically mixed") Or
(salinity = "partially mixed" And widthdepth = "very high") Then
        strcriteria = "[dimension] = '3-D'"
        reci.FindFirst strcriteria

```

```

Do While reci.NoMatch = False
  reci.Edit
  reci.Delete
  reci.FindNext strcriteria
Loop
End If
Then
  If (widthdepth = "high" Or widthdepth = "very high")
    strcriteria = "[dimension] = '2-D avrg on width'"
    reci.FindFirst strcriteria
    Do While reci.NoMatch = False
      reci.Edit
      reci.Delete
      reci.FindNext strcriteria
    Loop
    strcriter = "[dimension] = '2-D avrg on depth'"
    reci.FindFirst strcriter
    Do While reci.NoMatch = False
      reci.Edit
      resim = reci("simodel")
      resim = (resim + 1)
      reci("simodel") = resim

      reci.Update
      reci.FindNext strcriter
    Loop
  End If
  If (widthdepth = "low" Or widthdepth = "moderate") Then
    strcriteria = "[dimension] = '2-D avrg on depth'"
    reci.FindFirst strcriteria
    Do While reci.NoMatch = False
      reci.Edit
      reci.Delete
      reci.FindNext strcriteria
    Loop
    strcriter = "[dimension] = '2-D avrg on width'"
    reci.FindFirst strcriter
    Do While reci.NoMatch = False
      reci.Edit
      resim = reci("simodel")
      reci("simodel") = (resim + 1)

      reci.Update
      reci.FindNext strcriter
    Loop
  End If
  End If
  If ((intertidalarea / totalarea) > 0.6) Then
    strcriter = "[dimension] = '2-D moving'"

```

```

    reci.FindFirst strcriter
    Do While reci.NoMatch = False
      reci.Edit
      resim = reci("simodel")
      resim = (resim + 1)
      reci("simodel") = resim
      reci.Update
      reci.FindNext strcriter
    Loop
  End If
  If (intotal = 0) Then
    If (salinity = "vertically mixed" Or salinity = "well
mixed") Then
      If widthdepth = "low" Then
        strcriteria = "[dimension] = '1-D'"
        reci.FindFirst strcriteria
        Do While reci.NoMatch = False
          reci.Edit
          resim = reci("simodel")
          resim = (resim + 1)
          reci("simodel") = resim
          reci.Update
          reci.FindNext strcriteria
        Loop
      Else
        strcriteria = "[dimension] = '2-D avrg on depth'"
        reci.FindFirst strcriteria
        Do While reci.NoMatch = False
          reci.Edit
          resim = reci("simodel")
          resim = (resim + 1)
          reci("simodel") = resim
          reci.Update
          reci.FindNext strcriteria
        Loop
      End If
    End If
  End If
  If (intotal >= 1) Then
    If (salinity = "vertically mixed" Or salinity = "well mixed") And
(widthdepth = "low" Or widthdepth = "moderate") Then
      strcriter = "[dimension] = '1-D network'"
      reci.FindFirst strcriter
      Do While reci.NoMatch = False
        reci.Edit
        resim = reci("simodel")
        resim = (resim + 1)
        total = total + 1
        reci("simodel") = resim

        reci.Update
        reci.FindNext strcriter

```

```

Loop
  End If
  End If
Case "bar built"
  If (salinity = "salt wedge" Or salinity =
"partially mixed") Then
    If (widthdepth = "high" Or
widthdepth = "very high") Then
      strcriter = "[dimension] = '2-D
avrg on depth'"
      reci.FindFirst strcriter
      Do While reci.NoMatch =
False
        reci.Edit
        resim =
        resim = (resim +
1)
        reci("simodel") =
resim
      strcriter
    Loop
  Else
    strcriter = "[dimension] = '3-
D'"
    reci.FindFirst strcriter
    Do While reci.NoMatch =
False
      reci.Edit
      resim =
      resim = (resim + 1)
      reci("simodel") =
resim
    Loop
  End If
End Select
  strcriter = "[problem]='salinity distribution'"
  reci.FindFirst strcriter
  Do While reci.NoMatch =
False
    reci.Edit
    resim =
    reci("simodel")
  Loop

```

```

2.75
Round(resim, 3)
resim
strcriter
Loop
Case "salt intrusion"
  strcriteria = "[dimension] = '2-D moving'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit
    reci.Delete
    reci.FindNext strcriteria
  Loop
  Select Case estutype
    Case "fjord", "fjard", "ria"
      If (intotal = 0) Then
        strcriteria = "[dimension] = '1-D'"
        reci.FindFirst strcriteria
        Do While reci.NoMatch = False
          reci.Edit
          resim = reci("simodel")
          resim = (resim + 1)
          total = total + 1
          reci("simodel") = resim
          reci.Update
          reci.FindNext strcriteria
        Loop
      Else
        strcriter = "[dimension] = '1-D network'"
        reci.FindFirst strcriter
        Do While reci.NoMatch = False
          reci.Edit
          resim = reci("simodel")
          resim = (resim + 1)
          total = total + 1
          reci("simodel") = resim
          reci.Update
          reci.FindNext strcriter
        Loop
      End If
    Case "coastal plain", "complex"
      If (widthdepth = "low" Or widthdepth = "moderate")
Then
        resim = resim /
        resim =
        reci("simodel") =
        reci.Update
        reci.FindNext
      Loop

```

```

If (intotal = 0) Then
  strcriteria = "[dimension] = '1-D'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit
    resim = reci("simodel")
    total = total + 1
    resim = (resim + 1)
    reci("simodel") = resim
    reci.Update
    reci.FindNext strcriteria
  Loop
Else
  strcriteria = "[dimension] = '1-D network'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit
    resim = reci("simodel")
    resim = (resim + 1)
    total = total + 1
    reci("simodel") = resim
    reci.Update
    reci.FindNext strcriteria
  Loop
End If
End If
If (widthdepth = "high" Or widthdepth = "very high")
Then
  strcriteria = "[dimension] = '2-D avrg on depth'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit
    resim = reci("simodel")
    resim = (resim + 1)
    total = total + 1
    reci("simodel") = resim
    reci.Update
    reci.FindNext strcriteria
  Loop
End If

Case "bar built"
  strcriteria = "[dimension] = '2-D avrg on depth'"
  reci.FindFirst strcriteria
  Do While reci.NoMatch = False
    reci.Edit
    resim = reci("simodel")
    resim = (resim + 1)
    total = total + 1
    reci("simodel") = resim
    reci.Update

```

```

reci.FindNext strcriteria
Loop
End Select

strcriter = "[problem]='salt intrusion'"
reci.FindFirst strcriter
Do While reci.NoMatch =
  reci.Edit
  resim =
  resim = resim /
  reci("simodel") =
  reci.Update
  reci.FindNext
Loop

End Select

rec.Requery
rec.Close
reci.MoveFirst
Do
  reci.Edit
  reci.Update
  resim = reci("simodel")

  reci.MoveNext
Loop Until reci.EOF
End Sub

```

APPENDIX C

GA MODULE - CODE

```
program data
```

```
c  
c
```

```
integer ista,ibrach,isava,isavl,juncr,ibcup,  
1 ibcdn,iflux,ibrahd,izone  
double precision tidper,tcon,aK2,testv,deltat  
real Qflow,Qflowb  
character*60 dfnam1,hydfnm,dfnam9,dfnam7,dfnam2,namedata  
character*60 dfnam8  
parameter (mlchr=1000,mpsiz=500,msize=5000,msample=3000)  
integer lchrom, popsize,irun,ttime(msample)  
real random(mpsiz,mlchr),randx  
integer seed,ncycle  
integer nstat, npoint(msample),stat(msample), run  
integer numele(500),crosstri,inods(msample),isca,bran(10)  
real randx,tstep,cycle,nstep,pcross,pmut  
real low, high
```

```
c
```

```
character*60 dname(30)*20,title,sname  
character*60 fname  
character*60 name(30),dfnam3,dfnam4,dfnam5
```

```
c
```

```
open(unit=45,file='data',form='formatted',status='new')  
print*,'is the river/estuary branching? 0...no, 1...yes'  
read*,ibrach  
write(45,*)ibrach  
print*  
print*,'is simulation starting from'  
print*,'low water...enter 0,other...enter 1'  
read*,ista  
write(45,*)ista  
print*  
print*,'is the geometrical data available for all branches?'  
print*,'0... no, 1...yes'  
read*,ibrahd  
write(45,*)ibrahd  
print*,'enter name of input element datafile'  
print*  
read*,dfnam1  
write(45,*)dfnam1  
print*  
print*,'enter name of hydrografic datafile'  
read*,hydfnm  
write(45,*)hydfnm  
print*  
print*,'enter constant landward discharge boundary data'  
print*,'enter main channel discharge(cu.m/s)'  
print*  
read*,Qflow  
write(45,*)Qflow  
print*
```

```
if (ibrach.eq.1) then  
    print*,'enter branch channel discharge (cu.m/s)'  
    print*  
    read*,Qflowb  
    write(45,*)Qflowb  
    print*  
end if  
print*,'enter duration of the tidal cycle(secs)'  
print*  
read*,tidper  
write(45,*)tidper  
print*  
print*,' enter the name of the elevation boundary datafile'  
read*,dfnam9  
write(45,*)dfnam9  
print*  
print*,'are salinity changes significant along the river?'  
print*,'enter 0 for no,1 for yes'  
print*  
read*,isava  
write(45,*)isava  
print*  
print*,'enter salinity boundary datafile'  
print*  
read*,dfnam7  
write(45,*)dfnam7  
print*  
print*,'is salinity data available...0 o simulated...1?'  
read*,isavl  
write(45,*)isavl  
if (isavl.eq.0) then  
    print*,'enter the salinity datafile name'  
    read*,dfnam8  
    aK2=0  
    tcon=0  
    juncr=0  
    ibcup=5  
    iflux=5  
    ibcdn=5  
elseif (isavl.eq.1) then  
    dfnam8='mamma'  
    if (ibrach.eq.1) then  
        print*,'enter the junction relationship for the salt'  
        print*,'simulation'  
        print*,'enter:'  
        print*,'1... unsteady state'  
        print*,'2... steady state'  
        print*,'3... common node'  
        print*,'4... three equal nodes'  
        print*  
        read*,juncr  
    endif  
endif
```

```

print*
print*, 'enter if upstream boundary is specified as:'
print*
print*, '0... Neumann condition, 1... Dirichlet condition'
read*, ibcup
print*
ibcdn = 1
iflux = 5
if(ibcup.eq.0) then
  ielem=0
  if(ielem.eq.0) then
    print*, 'flux at the boundary is it:'
    print*, '0...enter 0, constant...1'
    read*, iflux
    print*
  endif
endif
print*, 'enter value of coefficient in dispersion relation:'
print*, ' D=64.K1.n(Q/A)**(5/6)+K2(ds/dx)'
print*, 'K1 and K2, respectively'
print*
read*, tcon, aK2
endif
write(45,*) dfnam8
write(45,*) juncr
write(45,*) ibcup
write(45,*) ibcdn
write(45,*) iflux
write(45,*) tcon
write(45,*) aK2
print*
print*, 'enter the name of your output convergence datafile'
read*, dfnam2
write(45,*) dfnam2
print*

print*, 'enter time step'
print*
read*, deltat
write(45,*) deltat
print*
print*, 'enter no. of tidal cycles'
print*
read*, ncycle
write(45,*) ncycle
print*
print*, 'enter tolerance level'
read*, testv
write(45,*) testv
print*

```

```

into'
print*, 'enter the number of elements the domain is divided
read*, nelelem
write(45,*) nelelem
print*
c
if(ibrach.eq.1) then
  print*, 'enter the number of element for each branch'
  print*, ' starting from left clockwise'
  do 10 i=1,3
    read*, bran(i)
    write(45,*) bran(i)
10  continue
  else
    do 20 i=1,3
      bran(i)=0
      write(45,*) bran(i)
20  continue
endif
print*
print*, 'enter the population size'
read*, popsize
write(45,*) popsize
print*
print*, 'enter the number of GA runs'
read*, irun
write(45,*) irun
c
print*
print*, 'enter a positive integer seed number'
read*, seed
write(45,*) seed
print*
print*, 'enter the probability of crossover'
read*, pcross
write(45,*) pcross
print*
print*, 'enter the probability of mutation'
read*, pmut
write(45,*) pmut
print*
print*, 'do you want to mutate adjacent chromosomes'
print*, 'enter 1...yes, 0...no'
read*, mutad
write(45,*) mutad
print*
print*, 'do you want to cross all the estuary branches'
print*, 'enter 1...yes, 0...no'
read*, crosstri
write(45,*) crosstri
c
c
dfnam4 is the name file used for storing the possible

```



```

c combination due to the probability of crossover
c
print*
print*,'enter the name of the datafile for storing the'
print*,'chromosomes selected as new population'
read*, dfnam4
write(45,*) dfnam4
c
print*
print*,'enter the lower limit and upper limit for'
print*,'the Manning*s coefficient'
read*,low
read*,high
low=1000*low
high=1000*high
write(45,*) low
write(45,*) high
c
print*
print*,'would you like to generate the Manning*s series'
print*,'randomly or with the zonation...enter 1 or 0'
read*,ima
write(45,*) ima
c
if (ima.eq.0) then
print*
print*,'enter the number of zones in which you would'
print*,'like to divide the estuary (zonation)'
read*,izone
write(45,*) izone
print*,'enter the node/upperlimit of each zone'
print*,'starting from the estuary source'
do 150 ij=1,izone
read*, inods(ij)
write(45,*) inods(ij)
150 continue
print*
print*,'would you like to scale the coefficient'
print*,'yes 1, no 0'
read*, isca
write(45,*) isca
c
else
izone=0
inods(1)=0
isca=5
endif
if(ima.eq.1) isca=9
c dfnam3 is the datafile containing the manning*s
c coefficient series from estuaries evaluated as similar
c
print*

```

```

print*,'do you want to include previously used mannings'
print*,' coefficient series...yes 1,no 0'
read*, imani
write(45,*) imani
print*
if (imani.eq.1)then
print*,'enter the name of the datafile containing mannings'
print*,'coefficient series that you would like to use'
read*,dfnam3
write(45,*) dfnam3
endif
c
print*
print*,'enter the name of the fitness function datafile'
read*,fname
write(45,*) fname
c
print*
print*,'enter the name of the file for storing the'
print*,'Manning*s series'
read*, dfnam5
write(45,*) dfnam5
c
print*
print*,'enter the name of the experimental datafile'
read*, title
write(45,*) title
stop
end
c
c -----
c program initialise
c -----
c this program reads the necessary input information for
starting the program

```

```

c   and initialise the program by generating the necessary
chromosomes
c   according to the input information. Some of the chromosomes
are generated
c   from the cases retrieved by the CBEM system.
c
c
c   GA variables
c
c   popsize= population size
c   lchrom= chromosome length
c   irun= number of runs
c   maxgen= maximum number of generation
c   pcross= crossover probability
c   pmut= mutation probability
c   rando= array of number generated through a random process
c   oldpop=matrices of the population before the mutation and
crossover
c   newpop=matrices of the population after the mutation and
crossover
c
c   -----
c   simulation variables to be used in the GA routines
c   -----
c
c   tstep=time step used in the simulation
c   cycle=duration of the tidal cycle
c   nstep=number of time step in which the tidal cycle is divided
c   hsim= array of the simulated water elevations
c   mea(k,2)= water elevation field data
c   mea(k,1)=timeinstants when the water elevations were sampled
c   nstat= number of stations at which the field data were sampled
c   sname= name of the station
c   npoint(i)= number of sample for each station i
c
c
c   parameter (mlchr=1000,mpsiz=500,msize=5000,msample=3000)
integer lchrom, popsize,irun,ttime(msample),lzone
real random(mpsiz,mlchr),randx
integer seed, irandx, counter, ista,ibrach,isava,ncycle
integer nstat, npoint(msample),stat(msample), run
integer numele(500),timeh,time,crosstri,inods(msample)
integer ibrahd,juncr,ibcup,ibcdn,iflux,isavl
real randx,tstep,cycle,nstep,pcross,pmut
real high,low,coeff,coeff2,coeff3, coef1
c
c   double precision mea(msample,2),ob(mpsiz),obi(mpsiz)
c   double precision sumfit,objfunl,mano(mlchr)
c   double precision sim(3000,100),mannsi(mlchr,mpsiz)
c   double precision hmea(msample),mann(mlchr,mpsiz)
c   double precision imann(mlchr),ahsim(3000,100)
c   double precision hsim(2000,30,50,50),newmann(mlchr,mpsiz)
c   double precision tidper,deltat,tcon,aK2, testv

```

```

c
c   character*60 dname(30)*20,title,dfnam1,sname
c   character*60 dfnam2,dfnam9,dfnam7,hydfnm, fname
c   character*60 name(30),dfnam3,dfnam4,dfnam5,dfnam8
c
c
c
c   open(unit=145,file='dataestuary7',form='formatted',status='old')
c   rewind(145)
c   read(145,*) ,ibrach
c   read(145,*) ,ista
c   read(145,*) ,ibrahd
c   read(145,*) ,dfnam1
c   read(145,*) ,hydfnm
c   read(145,*) ,Qflow
c   read(145,*) ,Qflowb
c   read(145,*) ,tidper
c   read(145,*) ,dfnam9
c   read(145,*) ,isava
c   read(145,*) ,dfnam7
c   read(145,*) ,isavl
c   read(145,*) ,dfnam8
c   read(145,*) ,juncr
c   read(145,*) ,ibcup
c   read(145,*) ,ibcdn
c   read(145,*) ,iflux
c   read(145,*) ,tcon
c   read(145,*) ,aK2
c   read(145,*) ,dfnam2
c   read(145,*) ,deltat
c   read(145,*) ,ncycle
c   read(145,*) ,testv
c   read(145,*) ,nelem
c
c   read(145,*) ,popsize
c   read(145,*) ,irun
c   lchrom=nelem
c
c
c   read(145,*) , seed
c   read(145,*) , pcross
c   read(145,*) , pmut
c   read(145,*) ,mutad
c   read(145,*) ,crosstri
c
c
c   dfnam4 is the name file used for storing the possible
c   combination due to the probability of crossover
c
c   read(145,*) , dfnam4
c
c
c   read(145,*) ,low
c   read(145,*) ,high

```

```

    read(145,*) , izone
    do 1008 ij=1, izone
        read(145,*) , inods(ij)
1008 continue
    read(145,*) , ima
    read(145,*) , imani
c
    if (imani.eq.1) then
        read(145,*) , dfnam3
        open(unit=119, file=dfnam3, form='formatted', status='old')
        read(119,*) , numes
        do 1011 i=1, numes
            read(119,*) , numele(i)
            do 1012 j=1, numele(i)
                read(119,*) , mannsi(j,i)
1012 continue
1011 continue
            do 1013 i=1, numes
                if (numele(i).eq.lchrom) then
                    do 1014 j=1, lchrom
                        mann(j,i)=mannsi(j,i)
1014 continue
                    l=0
                else if (numele(i).lt.lchrom) then
                    coef1=lchrom/numele(i)
                    ncoeff=int(lchrom/numele(i))
                    coeff3=0
                    do 1015 k=1, numele(i)
                        coeff=coeff*k-coeff3
1016 l=l+1
                        if (l.le.coef1) then
                            mann(l,i)=mannsi(k,i)
                            coeff2=coeff-1
                            goto 1016
                        else
                            mann(l,i)=mannsi(k,i)*coeff2+mannsi(k+1,i)*(1-coeff2)
                            coeff3=1-coeff2
                            goto 1015
                        endif
                        if (l.eq.lchrom) goto 1013
1015 continue
                else if (numele(i).gt.lchrom) then
                    coef1=numele(i)/lchrom
                    ncoeff=int(lchrom/numele(i))
                    coeff2=coeff-ncoeff
                    rest=0
                    summa=0
                    do 1017 k=1, lchrom
                        coeff=coef1*k
                        mann(k,i)=rest
1018 l=l+1
                        if (l.le.coef1) then

```

```

                            mann(k,i)= mann(k,i)+mannsi(l,i)
                            summa=summa+1
                            goto 1018
                        else
                            mann(k,i)=mann(k,i)+coeff2*mannsi(l,i)
                            summa=summa+coeff2
                            rest=(1-coeff2)*mannsi(l,i)
                        endif
                        mann(k,i)=mann(k,i)/summa
1017 continue
                    endif
1013 continue
                end if
c
                if (numes.ge.1) then
                    numsta=numes+1
                else
                    numsta=1
                endif
c
c dfnam2 is the convergence output file that is open in
c order to be treated as an old object in the simulation
c open(unit=20, file=dfnam2, form='formatted', status='new')
c write(220,*) 'start'
c close(unit=220)
c
c fname is the file used for storing the values of the
c fitness function calculated for all elements in the GA a
c population for different runs
c
c read(145,*) , fname
c open(unit=141, file=fname, form='formatted', status='new')
c rewind(141)
c
c read(145,*) , dfnam5
c open(unit=162, file=dfnam5, form='formatted', status='new')
c rewind(162)
c
c read experimental data for a datafile containing sampling
c elevations from different stations
c
c read(145,*) , title
c
c call readdata(title, deltat, name, ttime,
1 hmea, nstat, npoint, msample, stat)
c
c if (ima.eq.1) then
    do 1000 counter=numsta, popsize
        do 1010 j=1, lchrom
            call manning(seed, randx, low, high, raso)
            mann(j, counter)=raso
1010 continue

```

```

1000  continue
      endif
      if (ima.eq.0) then
        do 1020 counter=numsta,popsize
          call mannbranch(lchrom,seed,randx,mano,low,high,
1         izone,inods,msample)
          do 1030 i=1,lchrom
            mann(i,counter)=mano(i)
1030    continue
1020    continue
      endif
      open (unit=160,file=dfnam4,form='formatted',status='new')
      write(162,*) 'run=1'
      do 1031 i=1,lchrom
        write(162,5000) (mann(i,counter),counter=1,popsize)
1031  continue
      endfile(162)
c
      do 1033 counter=1,popsize
        do 1034 i=1,lchrom
          imann(i)=mann(i,counter)
1034  continue
          do 1035 jk=1,nstat
            do 1036 jl=1,ntimes
              sim(jl,jk)=0.0
1036  continue
1035  continue
          call simulation (imann,sim,deltat,tidper,ntimes,nstat,
1         stat,ista,ibrahd,dfnam1,hydfnm,ibrach,
2         Qflow,Qflowb,dfnam9,isava,dfnam7,
3         isavl,dfnam8,juncr,ibcup,ibcdn,iflux,
4         tcon,aK2,dfnam2,ncycle,testv)
c
          do 1040 i=1,nstat
            do 1050 k=1,ntimes
              hsim(k,i,counter,1)=sim(k,i)
1050  continue
1040  continue
1033  continue
c
      run=1
      sumfit=0.0
      do 1060 counter=1,popsize
        call object1(objfun1,nstat,npoint,hsim,counter,
1         hmea,ttime,run,msample)
        ob(counter)=objfun1
        obi(counter)=objfun1
        write(141,*) obi(counter)
        sumfit=sumfit+objfun1
1060  continue
      endfile (unit=141)
c

```

```

open(unit=161,name='roulette',form='formatted',status='new')
do 1090 i=2,irun
  backspace 141
  run=i
c
c   the first two chromosomes in the new population will be the
c   chromosomes with the best two values of the object function
c
  call firsttwo (ob,lchrom,popsize,mann,newmann)
c
c   start creating a new generation thorough the subroutine
c   selection, crossover and mutation. We assume that the
c   population size is constant.
c
  call generation (pcross,pmut,seed,sumfit,obi,mann,
1         lchrom,popsize,newmann,mutad,crosstri,
2         high,low)
c
  do 1100 j=1,popsize
    do 1110 k=1,lchrom
      mann(k,j)=newmann(k,j)
1110  continue
1100  continue
      backspace (162)
      write(162,*) 'run=',i
      do 1115 ima=1,lchrom
        write(162,5000) (mann(ima,k),k=1,popsize)
1115  continue
      endfile(162)
c
  do 1120 counter=1,popsize
    do 1125 ik=1,lchrom
      imann(ik)=mann(ik,counter)
1125  continue
      do 1126 jk=1,nstat
        do 1127 jl=1,ntimes
          sim(jl,jk)=0.0
1127  continue
1126  continue
      call simulation(imann,sim,deltat,tidper,ntimes,nstat,
1         stat,ista,ibrahd,dfnam1,hydfnm,ibrach,
2         Qflow,Qflowb,dfnam9,isava,dfnam7,isavl,
3         dfnam8,juncr,ibcup,ibcdn,iflux,tcon,
4         aK2,dfnam2,ncycle,testv)
      do 1130 j=1,nstat
        do 1140 k=1,ntimes
          hsim(k,j,counter,i)=sim(k,j)
1140  continue
1130  continue
1120  continue
c
      sumfit=0.0

```

```

do 1150 counter=1,popsize
  call object1 (objfun1,nstat,npoint,hsim,counter,
1      hmea,ttime,run,msample)
  obi(counter)=objfun1
  write(141,*) obi(counter)
  ob(counter)=objfun1
  sumfit=sumfit+objfun1
1150 continue
c
  endfile (unit=141)
1090 continue
  close(unit=141)
  do 1160 k=1,nstat
  open(unit=151+k,file=name(k),form='formatted',status='new')
  do 1170 i=1,irun
    timeh=0
    do 1180j=1,ntimes
      time=idint(deltat)
      timeh=time*j

write(151+k,6000)timeh,(hsim(j,k,counter,i),counter=1,popsize)
1180 continue
1170 continue
  close (unit=151+k)
  close (unit=160)
  close (unit=161)
  close (unit=141)
  close (unit=162)
1160 continue
5000 format(500f8.4)
6000 format(I6,500f8.4)
  close(unit=145)
  stop
  end
c
c -----
c
c subroutine randomize (seed,x)
c -----
c
integer seed
real x
seed= 2045*seed+1
seed= seed - ((seed/1048576)*1048576)
x= real(seed+1)/1048577.0
return
end
c
c -----
c
c subroutine manning (seed,randx,low,high,raso)
c -----
c
c c c c
c This subroutine imposes to the Manning's coefficient to
c be between a lower and an upper limits chosen by the user

```

```

c
real randx,high,low,raso,man
integer iman,seed
call randomize (seed,randx)
man=(randx*(high-low+1))+low
if (man.gt.high) man=high
iman=int (man)
raso= 0.001*iman
return
end
c
c -----
c
c subroutine manbranch (lchrom,seed,randx,mano,low,
1      high,izone,inods,msample)
c -----
c
c This subroutine divides the estuary in n zones according to
the
c zonation made by the expert and imposes the same randomly
generated
c Manning's coefficient for each zone.
c
integer lchrom,seed,irandx,nm(msample),izone
integer inods(msample)
real randx,riso,high,low
double precision mano(1000),mana(msample),max
n=1
do 1000 i=1,izone
  call manning (seed,randx,low,high,riso)
  mana(i)=riso
1000 continue
  do 1005 i=1,izone
  do 1006 j=i+1,izone
    if (mana(j).gt.mana(i)) then
      max=mana(i)
      mana(i)=mana(j)
      mana(j)=max
    endif
  1006 continue
1005 continue
  do 1008 i=1,izone
  do 1010 j=n,inods(i)
    mano(j)= mana(i)
  1010 continue
  n=1+inods(i)
1008 continue
  return
  end
c
c -----
c
c subroutine readdata(title,deltat,name,ttime,hmea,
1      nstat,npoint,msample,stat)
c -----

```

```

c   this subroutine reads the experimental data necessary
c   for calculating the fitness function.

double precision hmea(msample),time(msample),deltat
integer nstat,npoint(msample),stat(msample)
integer ttime(msample),av3,av4
double precision bv1,bv2,av1,av2
character*60 sname,title,name(30)
n=0
m=1
open(121,file=title,form='formatted',status='old')
read(121,*) nstat
do 2000 i=1,nstat
  read(121,8020) sname
  name(i)=sname
  read(121,*) stat(i)
  read(121,*) npoint(i)
  n=n+npoint(i)
  do 2015 k=m,n
    read(121,*) time(k)
2015   continue
  do 2017 k=m,n
    read(121,*) hmea(k)
2017   continue
  do 2020 l=m,n
    av1=dint(time(l)/deltat)
    av2=av1+1
    av1=av1*deltat
    av2=av2*deltat
    bv1=dabs(av1-time(l))
    bv2=dabs(av2-time(l))
    if(bv1.le.bv2) then
      av3=int(av1)
      ttime(l)=av3
    else
      av4=int(av2)
      ttime(l)=av4
    endif
2020   continue
  do 2030 l=m,n
2030   continue
  m=1+n
2000 continue
8020 format(a60)
close(unit=121)
return
end

-----
subroutine object1(objfun1,nstat,npoint,hsim,counter,
1      hmea,ttime,run,msample)

```

```

c   -----
c   this subroutine calculates the objective function based
c   on the estimation of the error between the simulated
c   results and the observed value. The present objective
c   function consists of calculating the standard deviation.
c
double precision objfun1,sum1,sum2
double precision hmea(msample)
double precision hsim(2000,30,50,50)
integer nstat,npoint(msample),run,counter
integer ttime(msample),sel,mel
n=0
sum1=0.0
sum2=0.0
objfun1=0.0
m=1
do 1010 i=1,nstat
  n=n+npoint(i)
  do 1020 j=m,n
    sel=ttime(j)
    mel=sel/30
    sum1=sum1+(hsim(mel,i,counter,run)-hmea(j))**2
1020   continue
  m=n+1
  sum2=sum2+sum1
1010 continue
objfun1= sum2**0.5
objfun1=1/objfun1
return
end

c   -----
c   subroutine firstwo (ob,lchrom,popsiz,mann,newmann)
c   -----
c   This subroutine determines the first two chromosomes in the
c   population with the best values for the object function.
c
parameter (mlchr=1000,mpsiz=500)
integer lchrom, popsiz
double precision ob(500),max1,max2,mann(mlchr,mpsiz)
double precision newmann(mlchr,mpsiz)

c
max1=ob(1)
n=1
do 1000 i=2,popsiz
  if(ob(i).gt.max1) then
    max1=ob(i)
    n=i
  endif
1000 continue
max2=0.0

```

```

do 1010 i=1,popsize
  if((ob(i).gt.max2).AND.(ob(i).ne.max1)) then
    max2=ob(i)
    m=i
  endif
1010 continue
do 1020 i=1,lchrom
  newmann(i,1)=mann(i,n)
  newmann(i,2)=mann(i,m)
1020 continue
return
end

c
c -----
c subroutine selection (popsize,seed,sumfit,obi,senum)
c -----
c the subroutine selects a single individual via roulette wheel
c mechanism.
c
double precision random,choic,sumfit
double precision obi(500)
integer senum,seed
real rand

c
backspace 161
call randomize (seed,rand)
random=rand*sumfit
write(161,*)'random number=', rand
write(161,*) 'the product between random number and sumfit is'
write(161,*) random
choic=0.0
i=0
1000   i=i+1
      choic=choic+obi(i)
      if ((choic.ge.random).OR.(i.eq.popsize))then
        write (161,*) 'the partial sum of the values of'
        write (161,*) 'object functions is', choic
        senum=i
      else
        goto 1000
      endif
endfile (161)
return
end

c
c -----
c subroutine crossover (pcross,pmut,child1,child2,parent1,
1 parent2,seed,npos,ncross,lchrom,mutad,
2 high,low)
c -----
c this subroutine calculates the crossover for those selected
c chromosomes according to the corresponding value of the

```

```

c objective function.
c
parameter(mlchr=1000)
double precision child1(mlchr),child2(mlchr),random
double precision parent1(mlchr),parent2(mlchr)
double precision chil,chi2
integer lchrom,ncross,npos,seed,nmut,mutad
real pcross,pmut,randa,rando,high,low

c
nmut=0
call randomize (seed,randa)
if (randa.le.pcross) then
  write(160,*)'crossover operated'
  l=lchrom-1
  call randomize(seed,rando)
  npos=int(rando*(l-1+1)+1)
  if (npos.gt.l) then
    npos=l
    ncross=ncross+1
  endif
else
  npos=lchrom
endif
do 1000 i=1,npos
  child1(i)=parent1(i)
  chil=child1(i)
  call mutation (pmut,seed,chil,nmut,high,low)
  if (nmut.ne.0) write(160,*) 'nmut=',nmut
  child1(i)=chil
  if (mutad.eq.1) then
    child1(i-1)=chil
    child1(i+1)=chil
  endif
  child2(i)=parent2(i)
  chi2=child2(i)
  call mutation (pmut,seed,chi2,nmut,high,low)
  child2(i)=chi2
  if (mutad.eq.1) then
    child2(i-1)=chi2
    child2(i+1)=chi2
  endif
  if (nmut.ne.0) write(160,*) 'nmut=',nmut
1000 continue
if (npos.ne.lchrom) then
  do 1010 j=npos+1,lchrom
    child1(j)=parent2(j)
    chil=child1(j)
    call mutation (pmut,seed,chil,nmut,high,low)
    if (nmut.ne.0) write(160,*) 'nmut=',nmut
    child1(j)=chil
    if (mutad.eq.1) then
      child1(j-1)=chil

```

```

        child1(j+1)=chil
        endif
        child2(j)=parent1(j)
        chi2=child2(j)
        call mutation (pmut, seed, chi2, nmut, high, low)
        if (nmut.ne.0) write(160,*) 'nmut=', nmut
        child2(j)=chi2
        if (mutad.eq.1) then
            child2(j-1)=chi2
            child2(j+1)=chi2
        endif
1010    continue
    else
        goto 2000
    endif
2000    return
end
c
c -----
subroutine crssex (pcross, pmut, child1, child2, parent1,
1          parent2, seed, lchrom, mutad, high, low,
2          ize, inods, msample)
c
c -----
parameter (mlchr=1000)
double precision child1(mlchr), child2(mlchr), random
double precision parent1(mlchr), parent2(mlchr)
double precision chil, chi2
integer lchrom, ncross, npos(3), seed, nmut, mutad, ncount
integer numb1, numb2, numb3, ize, inods(msample)
real pcross, pmut, randa, rando, high, low, hi, ho, umb

c
nmut=0
numb3=0
call randomize (seed, randa)
if (randa.le.pcross) then
do 1000 ij=1, ize
    npos(ij)=inods(ij)
1000    continue
    mj=ize
else
    npos(1)=lchrom
    mj=1
endif
m=1
call randomize(seed, randa)
hi=real(ize)
ho=1.0
umb=(randa*(hi-ho+1))+ho
if (umb.gt.hi) umb=hi
numb1=int(umb)
do 1010 i=1, numb1
1020    call randomize (seed, randa)

```

```

        hi=real(ize)
        ho=1.0
        umb=(randa*(hi-ho+1))+ho
        if (umb.gt.hi) umb=hi
        numb2=int(umb)
        if (numb2.eq.numb3) goto 1020
        nm=npox(numb2)
        if (numb2.eq.1) then
            nm=0
        else
            nm=npox(numb2-1)
        endif
        nm=m+nm
        do 1030 j=nm, nm
            child1(j)=parent2(j)
            child2(j)=parent1(j)
1030    continue
        numb3=numb2
        if (nm.eq.lchrom) goto 1040
1010    continue
1040    do 1050 j=1, lchrom
            child1(j)=parent2(j)
            chil=child1(j)
            call mutation (pmut, seed, chil, nmut, high, low)
            if (nmut.ne.0) write(160,*) 'nmut=', nmut
            child1(j)=chil
            if (mutad.eq.1) then
                child1(j-1)=chil
                child1(j+1)=chil
            endif
            child2(j)=parent1(j)
            chi2=child2(j)
            call mutation (pmut, seed, chi2, nmut, high, low)
            if (nmut.ne.0) write(160,*) 'nmut=', nmut
            child2(j)=chi2
            if (mutad.eq.1) then
                child2(j-1)=chi2
                child2(j+1)=chi2
            endif
1050    continue
        return
end
c
c -----
subroutine mutation (pmut, seed, random, nmut, high, low)
c
c -----
c This subroutine carries out the mutation of the chromosomes'
c bits in the present population.
c
double precision random, low, high
real pmut, rana, rano
integer nmut, seed, irandx

```



```

integer seed
c
call randomize (seed,rana)
if (rana.le.pmut) then
  call manning(seed,randx,low,high,rano)
  nmut=nmut+1
  random=rano
  write (160,*) 'mutation operated'
endif
return
end
c
c
-----
subroutine generation (pcross,pmut,seed,sumfit,obi,mann,
1      lchrom,popsize,newmann,mutad,crosstri,
2      high,low)
c
c
This subroutine creates a new generation of the population.
It activates the subroutines selection, crossover
and mutation.
The size of the population is constant.
c
parameter (mlchr=1000,mpsiz=500)
double precision parent1(mlchr),parent2(mlchr)
double precision child1(mlchr),child2(mlchr)
double precision mann(mlchr,mpsiz),obi(500)
double precision sumfit,newmann(mlchr,mpsiz)
real pcross,pmut,randx
integer senum,seed,irandx,ncross,nmut,mutad
integer mate1,mate2,popsize,lchrom,npos
c
ncross=0
backspace 160
i=0
1000 i=i+1
  call selection (popsize,seed,sumfit,obi,senum)
  mate1=senum
  call selection (popsize,seed,sumfit,obi,senum)
  mate2=senum
  write(160,*) mate1, mate2
  do 1010 j=1,lchrom
    parent1(j)=mann(j,mate1)
    parent2(j)=mann(j,mate2)
1010 continue
  if (crosstri.eq.1) then
    call crsser (pcross,pmut,child1,child2,parent1,
1      parent2,seed,lchrom,mutad,high,low)
  else
    call crossover (pcross,pmut,child1,child2,parent1,
1      parent2,seed,npos,ncross,lchrom,
2      mutad,high,low)
  endif

```

```

  if (ncross.ne.0) then
    write(160,*) 'ncross=',ncross
  endif
  n=i+2
  do 1020 j=1,lchrom
    newmann(j,n)=child1(j)
1020 continue
  n=i+3
  do 1030 j=1,lchrom
    newmann(j,n)=child2(j)
1030 continue
  if (n.le.popsize) then
    i=i+1
    goto 1000
  else
    goto 1040
  endif
1040 endfile(160)
return
end

```

## APPENDIX D

### PUBLISHED AND PRESENTED WORK

The following work, which has been part of the research, has been published or presented.

#### PUBLICATIONS

Passone, S., Das, D.B. and Nassehi, V. (2001). *Development of a discharge policy for the Tay Estuary based on a finite element model*. Water Science and Technology, 43(7), p. 247-256.

**Abstract.** The tidal hydrodynamics and effluent distribution in estuaries involve a complicated range of solute transport phenomena modelled by partial differential equations. Therefore, the quantitative estimation of the risks of water and soil contamination of coastal areas as a result of polluted estuary flows, or effects of the effluent input on the chemical loads, involves the solution of these equations. Generally, the pollutants load in an estuary is determined by the nature of land use which by altering the watershed hydrology or chemical detention/release in the river banks affect the water quality of the estuaries. The present modelling work aims to investigate the solute transport behaviour in the Tay Estuary in Scotland. Based on this study, an attempt to devise an estuary specific discharge strategy for the Tay has been made. The numerical calculations are based on using 2D Galerkin finite element discretisation of the governing equations in an Eulerian co-ordinate system. The flexibility of the formulation allows it to be extended to moving boundary situations encountered in most tidal water systems.

Passone, S., Chung, P.W.H. and Nassehi, V (2001). *Development of a decision support system for estuarine modelling using the case-based reasoning technology*. Water Resources Management. Edited by C.A. Brebbia, P. Anagnosopoulos, K. Katsifarakis and A.H.-D. Cheng, WIT Press, Southampton, UK.

**Abstract.** Computer modelling is regarded as an important investigative tool in the formulation of correct guidelines and effective environmental policies for estuaries. Its ability in revealing important aspects of the estuarine dynamics under different conditions has provided strong motivation for the use and success of this approach. However, computer models are normally developed for specific cases, which limits their applicability. Integration of numerical modelling and artificial intelligence techniques has the potential for creating advanced computational environments to assist in design of models for coastal hydrodynamic systems. Case based reasoning is a knowledge-based technique, which helps to capture and structure past experience in the form of cases studied. Through the comparison among similar cases CBR systems provide a support for the user in understanding novel situations and solving new problems. CBR for estuarine modelling (CBEM) imparts a way of building an intelligent computational environment where the available expertise about estuaries and the described models can be stored and reused. Through an intelligent interface, CBEM allows users to define the characteristics required for the numerical simulation. Based on the user's input, the modelling process starts from determining possible correlation between a new phenomenon to be simulated and the previous studies contained in the system. In this paper, the main emphasis is focused on the estuary and model description and, the retrieval of the models to be used in new simulations. Special attention is given to the organisation of categorisation indices adopted from different classifications of estuaries and models and, to the formal criteria established for retrieving "similar" computational programs for their use in a given problem.

Passone, S., Das, D.B., Nassehi, V. and Bikangaga, J.H. (2002). *Design of discharge policies for multiple effluent sources and returning pollutants scenarios in a branching estuary*. Estuarine, Coastal and Shelf Science (accepted).

**Abstract.** For realistic simulation of contaminants' transport in an estuary, the presence of multiple sources and circulation of pollutants returning upstream of the discharge point should be considered. These factors contribute to a complex mechanism for effluent transport in estuaries, which can be quantitatively analysed using computational methodologies. In this work, a finite element model has been used to study the transport of solute pollutants in a branching estuary with controlled discharges from multiple sources. The scheme has been applied to test five discharge policies for the Upper Milford Haven estuary in Wales, UK, a typical branching estuary. The outcomes of each discharge schedule in relation to the minimisation of the pollutants' concentration in the estuary, has been investigated for both spring and neap tides. Geomorphological features such as meanders and the bed slope of the estuary were included in the present study. To determine the transport and dilution of effluents in the estuary, other factors such as tidal dynamics and discharge point spacing and flow reversal have also been incorporated in the model. The developed scheme is used to quantify the interactions between multiple effluent discharges for different tidal conditions in conjunction with the effects of returning pollutants. Reaction mechanisms are not considered as the conservative behaviour for the effluents is assumed. The results of the present analysis can be generalised to networks of narrow branching estuaries characterised to predict the consequences of various discharge programmes.

#### CONFERENCE PAPERS

Passone, S., Mokhtarzadeh, M.R and Nassehi, V *Numerical Solution of 4th order Partial Differential Equations by Diffuse Approximation Method*. Accepted for presentation in **International Conference On Spectral and High Order Methods June 11-15, 2001, at Uppsala University, Sweden**

**Abstract.** For important categories of industrial problems mathematical models are generally formulated in terms of 4th order partial differential (bi-harmonic) equations. An excellent example of similar applications is the plate bending analysis. Accurate solution of these equations is, hence, of significant relevance when plates made from metals or polymer composites are used during the design of body parts for vehicles, household goods and different types of casings. Traditionally, high order finite element schemes based on Hermite or spectral elements are utilised to solve the plate bending problems. These schemes, however, depend on a computationally costly pre-processing stage to generate a suitable mesh for the problem domain. In this paper a novel mesh-free technique based on diffuse approximation method is proposed for solving homogeneous bi-harmonic equations. The present scheme mainly consists of formulating a functional, which represents the least square of error between the true solution and a selected trial function approximations of field unknowns. This functional is weighted, using locally variable test functions. High order approximations of the unknowns are, hence, generated by minimisation of the functional. The approximating functions are finally used via the standard Galerkin method to develop a suitable weighted residual statement for the governing equations. The comparison between the numerical solution and the analytical results for a test case is, also, presented.

Passone, S., Chung, P.W.H. and Nassehi, V *The use of the Case-Based Reasoning method for designing estuarine models*. Accepted for presentation **13th European Simulation Symposium Marseille, France, October 18-20, 2001**

**Abstract.** Estuaries' behaviour relies on many factors, which are possible to analyse only adopting incomplete study approaches. The physical processes within estuaries, such as floods and pollutant dispersion, are generally investigated through the use of computer modelling. However, computational models are usually limited to selective instances and oriented towards a narrow view of estuarine phenomena. The application of an artificial

intelligence technique such as case based reasoning (CBR) technology provides a flexible computer environment to support estuarine modelling. In a CBR system precision and realism of numerical modelling predictions can be improved by including multiplex domain aspects and extending the models' applications to a larger class of problems. In this paper it is shown how the CBR technology can help in the definition of suitable numerical models to simulate the physical behaviour and provide qualitative predictions for estuary systems. The description and retrieval phases are mainly described.

Passone, S., Chung, P.W.H. and Nassehi, V *A case based reasoning system for estuarine modelling support*. Accepted for presentation in the **5th International Conference on Hydroinformatics, July 1-5, 2002, Cardiff, Wales, UK.**

**Abstract.** Communication between modellers and policy makers is crucial in devising water resources management schemes. However, the differences in expertise and approach adopted by various groups of professionals often complicate useful and productive co-operation in selecting a right strategy for realistic situations. The aim of this paper is to present a computing system for estuarine modelling that can be utilised by users who have different backgrounds in estuarine science. This system facilitates understanding of complex concepts and communication between the users so that time and effort needed in a multidisciplinary work of this nature is significantly reduced. The system has three software components: a case-based reasoning scheme, a genetic algorithm and a library of numerical simulation models. These components are integrated to work as a single tool. The main idea is for the system to learn and adopt solutions from past experience for applications in new problems. The current prototype is equipped with examples of past solutions which correspond to specific features of real estuaries and employed models. The system's ability extends beyond simple retrieval of similar cases. The genetic algorithm component is responsible for fine-tuning of the parameters in the selected models to suit a particular estuarine environment being studied. This paper also describes an example of these applications to demonstrate the utility of the system in designing estuarine models.

Passone, S., Chung, P.W.H. and Nassehi, V (2002). *Use of a genetic algorithm in the calibration of estuary models*. Accepted for presentation in the **15th European Conference on Artificial Intelligence (ECAI), July 21-26, 2002, Lyon, France.**

**Abstract.** This paper describes an artificial intelligence (AI) system for estuarine model design. It is created by the combination of case-based reasoning and genetic algorithm techniques. This application aims to make the utilisation of complicated and expensive hydrodynamic models flexible, cost-effective and accessible to non-specialists. By organising the available knowledge of estuarine modelling into an interactive and dynamic framework, the AI system provides the user with the necessary guidance and information for numerically solving hydro-environmental problems related to estuaries. As soon as a new problem is given to the system, the case-based module for estuarine modelling (CBEM) is activated. This module accesses information about estuarine models and estuaries to which numerical solutions have been previously applied. After comparison and evaluation, the case-based search engine returns from its memory the most effective modelling scheme available for the solution of the new problem. The system then calls the genetic algorithm (GA) module which optimises the physical parameters of the selected modelling procedure to suit the new application. The main focus of this paper is on the description of the GA module. This module is developed by combining the classical evolutionary approach with problem-specific information to carry out the required parameter optimisation. The effectiveness of this procedure is illustrated using a one-dimensional hydrodynamic model for the Upper Milford Haven estuary in UK. The comparison between manual and genetic algorithm based calibrations for this specific case suggests that the GA routine can very effectively calibrate estuarine models under realistic situations. This means a significant reduction in the time normally necessary for the implementation of a numerical modelling scheme.

Passone, S., Chung, P.W.H. and Nassehi, V. *Case-based reasoning for estuarine model design*. Accepted for presentation in the **6th European Conference on Case-Based Reasoning (ECCBR)**, September 4-7, 2002, Aberdeen, Scotland.

**Abstract.** Estuaries are complex natural water systems. Their behaviour depend on many factors, which are possible to analyse only by adopting different study approaches. The physical processes within estuaries, such as floods and pollutant dispersion, are generally investigated through computer modelling. In this paper the application of case-based reasoning technology to support the design of estuarine models is described. The system aims to provide a non-expert user in modelling with the necessary guidance for selecting a model that matches his goal and the nature of the problem to be solved. The system is based on three components: a case-based reasoning scheme, a genetic algorithm and a library of numerical estuarine models. An example based on the Upper Milford Haven estuary (UK) is used to demonstrate the efficacy of the system's structure for supporting estuarine model design.

## POSTER

Passone, S., Mokhtarzadeh, M.R and Nassehi, V *Computation of eigen functions used in the spectral expansion of underground flow models*. Accepted for presentation in the **3rd International Conference on Future Groundwater at Risk**, June 25-27, 2001, Lisbon, Portugal.

**Abstract.** Using spectral expansions computationally efficient numerical models for the simulation of underground flow processes can be developed. To obtain such expansions the corresponding eigen values and eigen functions of the flow model need to be found. These can be calculated by the solution of appropriate singular second order Sturm-Liouville equations. In this paper we describe a novel method for the solution of these equations and the calculation of eigen functions required for the expansion of the governing equations of a typical underground flow model. This method is based on the application of a generalised moving weighted least-square scheme to obtain highly accurate smooth approximations for the field variables in terms of well defined weight and basis functions. The substitution of these approximations into the Sturm-Liouville equations generates residuals which are in turn weighted and solved using the standard Galerkin method. Accuracy and performance of the scheme in the context of groundwater flow models is discussed.

Das, D.B., Passone, S. and Nassehi, V. (2001). *Deterministic modelling of the total environment*. Proceedings of the **International conference on 'Detecting Environmental change: Science and Society**, July 17-20, 2001, London, UK.

**Abstract.** Though the natural world is far from being deterministic where one physical or chemical condition leads uniquely to another, the deterministic models have been very useful in engineering description of the environmental transport processes. The random (statistical) nature of the environmental parameter can be said to result from variety of factors. For example, pollutant concentration in air can vary due to changing weather conditions or other factors such as humidity, ambient temperature etc. Similarly, in the surface and sub-surface transport processes, the water quality may fluctuate within a range due to microbial growth and reactions. The deterministic models cannot predict such variability, as it is most suitable for calculating mean values. Nevertheless, they have been particularly useful in identifying the risks and hazards owing to degraded environmental quality (what-if analysis). Deterministic models of the environmental flow processes are based on the physical laws of conservation of mass, momentum and energy. They are represented in terms of non-linear partial differential equations, which usually require a numerical scheme for solution. These models allows the transport processes to be expressed with great accuracy but also demand a great deal of understanding about the phenomena. This realisation comes from the fact that the primary source of erroneous results is external logic (model assumptions) rather than the internal arguments in the mathematical formulations. Furthermore, to solve the model equations, special techniques need to be employed based on the desired tasks and model assumptions. The present paper has been motivated by mainly two broad aims. Firstly, to demonstrate the utility of deterministic methods for modelling complex environmental transport processes. For this purpose, modelling

cases of flow phenomena in all three components of the total environment, i.e., air, surface water and sub-surface are illustrated. The first example studies dispersion of a heavy gas in air. The dense chemical is released in air instantaneously at the ground level due to a catastrophic rupture of a storage tank. The second example studies solute convection and dispersion phenomenon in an estuary as a typical case of flow in surface water. The third example is given for the pollutant mobility in land and groundwater zone. The second objective of the paper is to demonstrate the utility of specific solution techniques in specific cases. For the heavy gas dispersion modelling, analytical methods have been adopted, while, for the other two cases, numerical techniques such as, finite element and finite volume method, respectively, are used. The use of a particular method for each case is justified in the paper.

## AWARDS

Travel grant awarded by ECCAI (European Coordinating Committee for Artificial Intelligence) for attending the **15th European Conference on Artificial Intelligence 2002, July 21-26, 2002, at Lyon, France.**

Travel grant awarded by ICOSAHOM-01 organising committee and Brown University for attending the **International Conference on Spectral and High Order Methods, June 11-15, 2001, at Uppsala University, Sweden.**

