Demonstrating the potential of GIS technology in hydrosalinity modelling through interfacing the DISA model and a GIS

by

Brendon Wolff-Piggott

Institute for Geographical Analysis
University of Stellenbosch

in association with

Ninham Shand Incorporated.

Report on a Project Funded Jointly by the Water Research Commission
and the Department of Water Affairs and Forestry

Project Leaders: Prof. H. L. Zietsman
Dr. A. H. M. Görgens

WRC Report No. 588/1/95
ISBN No. 1 86845 178 X

EXECUTIVE SUMMARY

Background

The middle Breede River valley in the Western Cape is an important area for the growing of vines and fruit under irrigation. Water is supplied from the Greater Brandvlei Dam, near Worcester. The bulk of the water is released into the Breede river, from which it is abstracted by the farmers. The river also serves as a drain for saline irrigation return-flows. This results in severe water quality problems in the summer. The expansion of irrigated lands will only add to this problem. Strategies which could be used to relieve this problem include the installation of drains to catch saline return-flow before it reaches the river, the upgrading of the canal water delivery system and the use of freshening releases from the Brandvlei Dam to lower salinity levels. These options either require significant capital investment, or the unproductive release of high-quality water.

The Daily Irrigation and Salinity Analysis (DISA) model is a physically-based hydrological simulation tool for the management of water supply to irrigation schemes. It was developed using the extensive research that had been conducted on salinization in the Breede River valley. The model requires the manual entry of large data sets, so that the configuration of the model for different scenarios or different catchments is a tedious procedure. This project stemmed from a perceived need by the Department of Water Affairs and Forestry to assist in the planning process by developing an interface between a geographical information system (GIS) and DISA.

Objectives

The aims of the project were as follows:

*The integration of a Geographical Information System with the DISA hydrosalinity model in order to demonstrate the application potential of GIS in hydrological modelling.

The following secondary aims are identified:

(a) To undertake a comprehensive review of the current status of efforts to integrate GIS and distributed hydrological models.

(b) To develop an application methodology that will enable the use of a GIS to collect, manipulate and provide the spatial data requirements for the Daily Irrigation and Salinity Analysis Model (DISA).

(c) To integrate the DISA model with Arc/Info.

(d) To verify that the methodology developed is functionally correct.

(e) To depict results of DISA model simulations by using and demonstrating GIS capabilities with maps, graphs, tables etc."

Summary of Findings

Literature Review

A preliminary study of the literature indicated that while the coupling of hydrological models and GIS has been proceeding for some time, little has been written about the principles and fundamental concepts involved in this activity. One approach has been to describe both hydrological models and GIS in terms of subsystems eg. subsystems for the user interface, data capture, data manipulation and analysis as well as data management and display. The coupling of models and GIS commonly occurs by linking their data management subsystems (see Figure A).

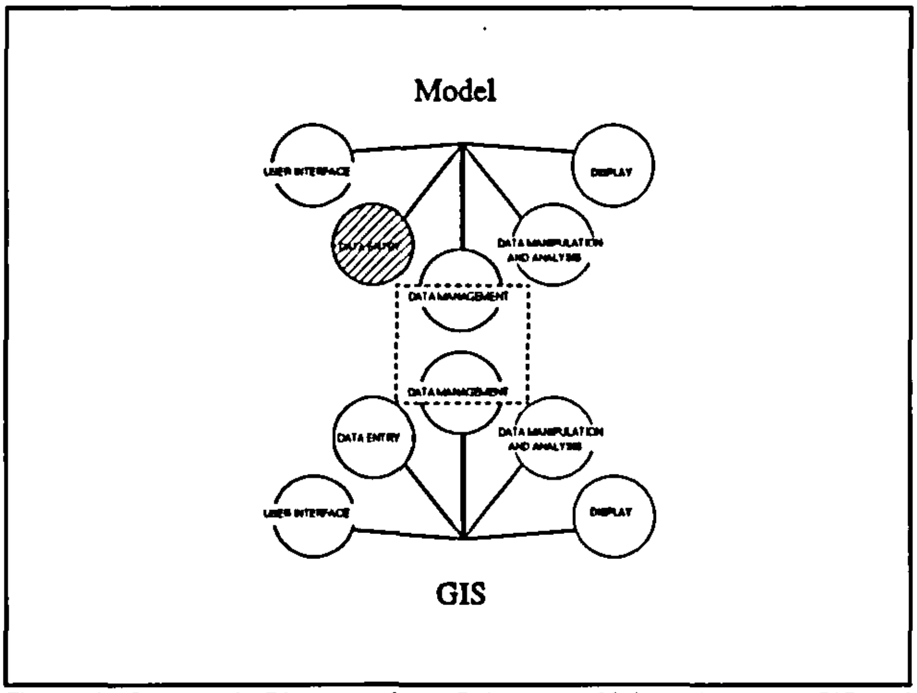


Figure A Schematic Diagram of the Subsystem Linkage between GIS and Models.

The different ways in which models and GIS can be coupled may then then be conceptualized in terms of the type of linkage established between the data management subsystems. The simplest form of linkage is the transfer of files between the model and the GIS, while an integrated system would combine the two sets of functionality in a single package. When data is exchanged between systems it is important that the data be represented in compatible ways eg. a model dealing only with features such as river reaches should be coupled with a GIS that can model linear networks rather than with a raster-based GIS.

Other perspectives on the coupling of models and GIS use the method of modelling and the user interface as organizing concepts. The modelling may be *internal* to the GIS, or it may be performed by software outside of the GIS environment, when it would be termed *external*. Coupled systems can be differentiated on the basis of whether a single user interface is presented or whether two different user interfaces have to be utilized to apply the software. These two views are closely related to the subsystems linkage classification, as an integrated system would also use internal modelling and a single user interface, while systems linked by simple file transfer would use external modelling and would probably have two different user interfaces. Figure B shows how the three perspectives on interfacing GIS and models may be integrated conceptually.

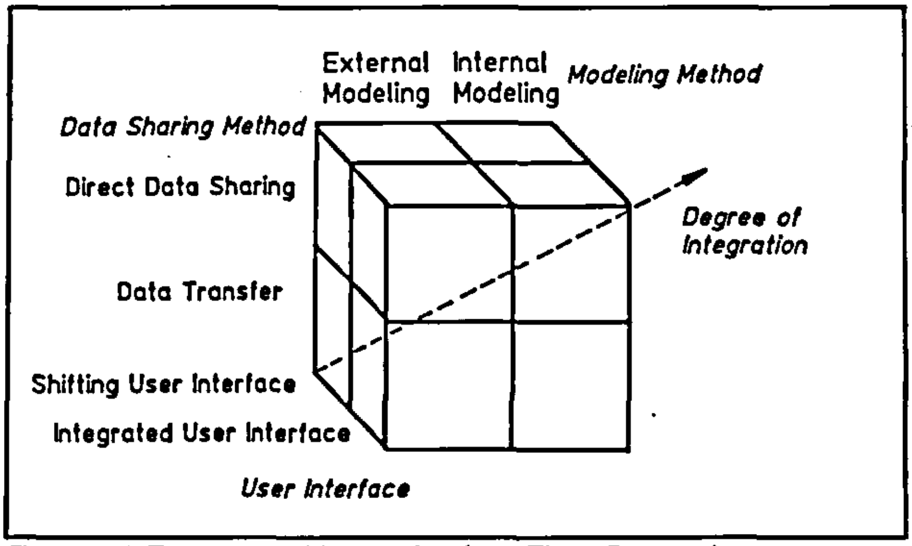


Figure B A Taxonomy of Integration from Three Perspectives

While the mechanics of establishing a simple coupling between a model and a GIS may be relatively straightforward, there are questions about the quality of results which may be obtained using this methodology. The management of error in GIS and the effect on GIS products is an area of ongoing research, but it is important to have an estimate of error in inputs to hydrological models. This is further complicated by the computational complexity and data intensive nature of distributed hydrological models. It has been suggested that the development of user-friendly GIS interfaces to hydrological models may lead to an abuse of modelling by individuals that do not understand the limitations of GIS data and the inherent complexity of hydrological models.

A number of authors have suggested ways to address the underlying problems in an automated way, but these are research proposals which are unlikely to provide practical solutions in the near future. Effective use of coupled GIS and hydrological models requires a sound understanding of both GIS and the model involved, however easy the system may be to operate. One closely related field of research is spatial decision support systems (SDSS). SDSS share a number of characteristics with decision support systems (DSS), namely the ability to address problems that are not well defined by using analytical and modelling techniques together with database systems, powerful presentation functions and ease of operation and modification. SDSS add to this the capability to capture, represent and analyse spatial data.

The literature review also examined the available references on applications of GIS with hydrological modelling, and noted that limited attention was given to the details of the coupling in the hydrological literature. The design and development of the spatial database was discussed in some detail, but little was mentioned on systems linkage and the user interface. Few references addressed the important topic of the effectiveness of coupled systems, i.e. the degree of confidence that can be attached to the results of such modelling. It was noted that most applications involved the coupling of pre-existing hydrological modelling software and GIS, and that a simple file transfer system was frequently used. Although the coupling of GIS and hydrological models had been widely demonstrated, it was not possible to determine whether or not this is a cost-effective strategy. It is difficult to evaluate the cost-effectiveness of information technology such as GIS, as its importance lies in the way that it frees skilled human resources to become more productive.

The Interfacing of DISA and Arc/Info

This project involved the use of an existing model together with Arc/Info. A simple file transfer linkage was therefore established between the two systems. As the particular strengths of GIS lie in its data capture and integration capabilities, and DISA already has a well-developed user interface, it was decided that Arc/Info could contribute most as a pre-processor of spatial information for DISA . The DISA user-interface would still be used to enter non-spatial information (see Figure C).

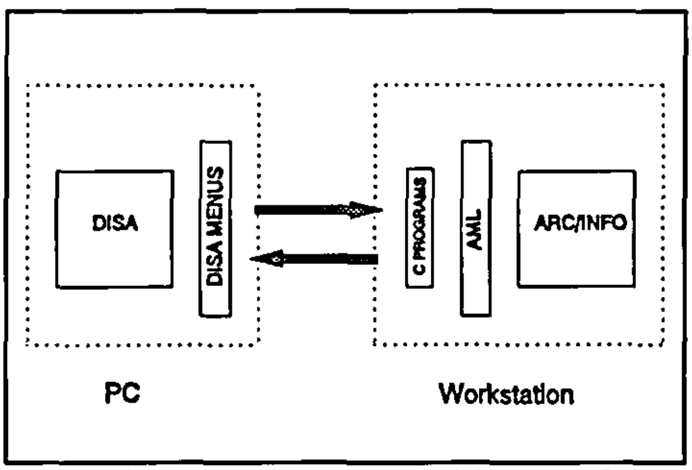


Figure C Schematic Diagram of the Linkage of DISA and Arc/Info

A GIS database was designed to receive all the spatial information required by DISA. The available spatial data was captured, and the derived data transferred into the database. It was found that the visual display of spatial data was very helpful in detecting errors in the data capture process. The results of the DISA model for the original configuration were then compared against the results obtained using the GIS-derived information. It was found that the results of the model were very similar at the point where the Breede river leaves the study area, and maximum salinities are expected (Zanddrift weir). Substantial differences were observed at return-flow cells with widely differing areas, however.

A configuration file was generated for the DISA model containing the information held in the GIS database, using the Arc/Info's own macro language as well as the C programming language. It was successfully read into the DISA configuration editor, demonstrating the success of the interface. An existing configuration file for DISA was read into the Arc/Info database, exported to DISA format again and visually compared against the original to confirm that the interface process did not introduce errors.

It was concluded that the use of GIS offers a number of significant advantages in the application of the DISA model, including faster spatial data capture, powerful visual tools for error detection and automated spatial data analysis. Although the presentation of the results of DISA were limited to the use of the existing model post-processor, all of the other objectives of the project have been met in full. The literature review synthesized both the theoretical and applications literature in a field that has been noted for its fragmentation, and provided a sound basis for the development of the interface between the DISA model and the Arc/Info GIS.

Research Needs

- The use of GIS offers much greater advantages when data capture from maps can be reduced or eliminated. A review of digital data sources available in Southern Africa, and their suitability for use in hydrological modelling is recommended. This would help reduce the large data capture costs associated with GIS.

- The identification of suitable standards for data storage in hydrological modelling is recommended, so that interfaces between models and GIS can be established more easily. The WDM format that has been established by the USGS is an example of one such standard for time series data.

- The representation of variation over time is limited in GIS. This means that the presentation of time series data from hydrological modelling within a GIS environment requires a great deal of programming expertise and effort. The Institute for Water Quality Studies has developed a sophisticated system (AQCES) to perform such a task for water quality time series. It is recommended that a system such as this should be adapted for use with models rather than independently developing GIS post-processors for models, wherever possible.

- It is recommended that the presentation of DISA output within GIS be addressed by interfacing

AQCES and DISA, as it was not possible to develop this aspect within the scope of the project.

- The interface that has been developed in the course of this project is suitable for operational use. It is recommended that the software system be applied in the configuration of DISA for a new catchment in order to demonstrate the benefit of GIS for hydrological modelling.

Technical Issues

Technical issues have been addressed in a series of appendices to the main text. These include details of the GIS database design, a discussion of the coding of the interface, the methodology developed for the use of Arc/Info with the DISA model and a detailed listing of the results obtained from GIS data capture in comparison with the original model configuration. The software that was developed in the course of this project is being archived by Ninham Shand Inc. (Cape Town), from whom it may be obtained.

ACKNOWLEDGEMENTS

The research reported on in this document emanated from a project jointly funded by the Water Research Commission (WRC) and the Department of Water Affairs and Forestry (DWAF). The Steering Committee responsible for this project consisted of the following persons:

Mr H M du Plessis	WRC (Chairman)
Mr F P Marais	WRC (Secretary)
Mr H Maaren	WRC
Prof I J van der Merwe	University of Stellenbosch
Mr L A Visagie	DWAF
Mr D R McPherson	DWAF
Dr A H M Görgens	Ninham Shand Inc.

The financing of the project by the Water Research Commission and the Department of Water Affairs and Forestry, and the contribution of the members of the Steering Committee is gratefully acknowledged.

I also wish to record my sincere thanks to Mr L A Visagie, of the DWAF, Mr S Forster, formerly of the DWAF, and to Prof J H Moolman of the University of Stellenbosch, for their assistance in obtaining reference materials, including the draft standards for the DWAF data dictionary. Mr Andre Greyling of Ninham Shand provided invaluable advice on the analysis and operation of the DISA model. Especial thanks are due to the project leaders, Prof H L Zietsman of the University of Stellenbosch and Dr A H M Görgens of Ninham Shand for their valuable comments and criticism in the course of this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 - INTRODUCTION

1.1 Background

The middle Breede River valley in the Western Cape is an important area for the growing of vines and fruit under irrigation. Water is supplied from the Greater Brandvlei Dam, near Worcester. The bulk of the water is released into the Breede river, from which it is abstracted by the farmers. The river also serves as a drain for saline irrigation return-flows. This results in severe water quality problems in the summer (Forster *et al*, 1991). The expansion of irrigated lands will only add to this problem. Strategies which could be used to relieve this problem include the installation of drains to catch saline return-flow before it reaches the river, the upgrading of the canal water delivery system and the use of freshening releases from the Brandvlei Dam to lower salinity levels. These options either require significant capital investment, or the unproductive release of high-quality water (Görgens *et al*, 1992).

An urgent need therefore exists for planning tools which can assist decision-makers to identify the costs and benefits of these different development scenarios. Hydrological simulation modelling can be used to explore these issues (Abbott *et al*, 1986). This requires the assembly and analysis of large amounts of detailed information about the state of the catchment. This is a lengthy and tedious process, and the development of an adequate database is recognized as a major issue in the modelling process (Beven, 1989).

Geographical information systems (GIS) are computerized systems which can capture, store and manipulate geographical information. It has been demonstrated that they have powerful capabilities for creating and maintaining an inventory of natural resources. Their ability to perform sophisticated analytical tasks is limited, however (Goodchild, 1992b).

The Daily Irrigation and Salinity Analysis (DISA) model is a semi-distributed hydrological simulation tool for the management of water supply to irrigation schemes. DISA has already been used to assist in planning in the Breede River valley (Greyling, Beuster and Görgens, 1993). However, the model requires the manual entry of large data sets, so the configuration of the model for different scenarios or different catchments is a tedious procedure. This project stemmed from a perceived need by the Department of Water Affairs and Forestry to assist in the planning process by developing an interface between a GIS and DISA.

1.2 Scope of the Project

"The integration of a Geographical Information System with the DISA hydrosalinity model in order to demonstrate the application potential of GIS in hydrological modelling.

The following secondary aims are identified:

(a) To undertake a comprehensive review of the current status of efforts to integrate GIS and distributed hydrological models.

(b) To develop an application methodology that will enable the use of a GIS to collect, manipulate and provide the spatial data requirements for the Daily Irrigation and Salinity Analysis Model (DISA).

(c) To integrate the DISA model with Arc/Info.

(d) To verify that the methodology developed is functionally correct.

(e) To depict results of DISA model simulations by using and demonstrating GIS capabilities with maps, graphs, tables etc."

The project was jointly funded by the Water Research Commission (WRC) and the Department of Water Affairs and Forestry (DWAF).

1.3 Report Structure

The literature review is presented in Chapter 2, and the results of the review are summarised to identify the principles on which to base the project. The DISA model and the Arc/Info GIS are discussed in Chapter 3, in order to introduce the systems and to provide insight on aspects of particular importance to the interface development. The DISA model is considered in terms of the modelling approach it embodies, its structuring of the catchment information, the data required to operate the model and the feasibility of data interchange. Similarly, Arc/Info is analysed in terms of its data model, software structure and interfacing capabilities.

The GIS database design is dealt with in Chapter 4. First, the data requirements of DISA are considered in the light of data availability, and some implications for the database design are drawn. A conceptual design for the interface between DISA and Arc/Info is then proposed, and the implementation of this interface is then described.

The verification of this interface is discussed in Chapter 5. An internal verification is performed, checking the consistency of the interface by matching the file format produced from Arc/Info against the original file. The GIS-derived data is compared to the original configuration file and model results using both data sets are compared for the external verification. Finally, the achievements of the project are critically assessed in the concluding chapter.

Technical information is presented in a number of appendices to the main text. This includes a detailed description of the GIS database design, an overview of the coding undertaken for the interfacing and a guide to setting up and using the interfaced systems. The guide is intended to enable readers familiar with both UNIX Arc/Info and DISA to apply the interface described in this report in a different study area.

CHAPTER 2 - LITERATURE REVIEW

The literature review has taken a broad approach to the issue of interfacing GIS and hydrological models. First of all, the GIS literature was investigated for information on the topic. This revealed a set of general principles, which could only be applied after some knowledge of hydrological modelling systems had been gained. The subsequent section is therefore devoted to examining relevant issues in hydrological modelling. Next, examples of interfaced GIS and hydrological modelling systems were reviewed, and the related fields of spatial analysis and spatial decision support systems were discussed. Finally the results of the literature review were synthesized in a concluding section.

2.1 Interfacing GIS and modelling systems

There is agreement that GIS systems generally lack the capability to perform complex modelling tasks (Goodchild, 1987; Rhind, 1988), but there has been strong interest in GIS as a technology to assist in the parameterisation of models and in the presentation of their results (Goodchild, 1993). The interfacing of GIS and modelling systems has been proposed as a way to overcome the limitations of GIS analysis (Burrough, 1990; Goodchild, 1992a). Although there has been interest in this topic for a number of years, little attention has been given to defining how this should be carried out (Nyerges, 1992). In this section the interfacing of GIS and modelling systems is treated first from a structural perspective, in terms of subsystem linkage, and then the prerequisites for effective use of interfaced systems are discussed.

2.1.1 Software Subsystems

Nyerges (1992) proposed that both GIS and modelling systems can be described in terms of subsystems. GIS can be considered as consisting of subsystems for the user interface, data entry, data manipulation and analysis, data management and display. Modelling software often consists of the same subsystems, except for data entry, but has an emphasis on analysis and display. Nyerges (1992) stated that coupling can be implemented for various combinations of the GIS and model subsystems, but that the most common linkage occurs through the data management subsystems. This view is implicit in the work of a number of other authors, including Burrough (1989), Chou (1991), Chou and Ding (1992) and Goodchild *et al* (1992).

There is agreement that the selection of a coupling strategy involves a number of trade-offs, principally between the performance required from the system and the resources available for the development effort (Chou and Ding, 1992; Goodchild *et al*, 1992; Nyerges, 1992). Nyerges described strategies listed by a number of authors to develop a linkage, and synthesized a framework to illustrate the various trade-offs that may be considered, on the basis of the closeness of the coupling achieved. He identified four different *application coupling environments*, which he termed *isolated*, *loose*, *tight* and *integrated*, which are described in more detail in Table I below.

Isolated and loosely coupled interfaces rely on human operators to supervise the communication between the systems, based on file transfer. They are relatively easy to develop, especially if a suitable model and GIS are already available. The looseness of the coupling can also lead to a clumsy and slow exchange, however.

Tight coupling also relies on file transfer between the systems, but the transfer is performed automatically by the software as required. They are therefore much easier to use, but require much more skill and effort to develop. In an integrated system there is no file transfer, as the 'GIS' and model components have been developed as a single software system, from the bottom up. The development and modification of such systems is a major undertaking.

Nyerges concluded that while an isolated environment is inexpensive, it may not perform adequately. The integrated environment should meet client's needs almost exactly, but involves a high cost of development, or modification. He suggested that standard interface services will be developed further in order to support the tight coupling of different systems, but that more effort is required to specify the requirements of such interfaces.

Chou (1991) discussed various strategies for integrating spatial analysis into GIS based on the method of data sharing and the modelling approach used, which included the possibility that the analytical functions could be performed within the GIS itself. Chou and Ding (1992) incorporated these proposals and suggested that there are three important perspectives on the issue, which should be viewed in conjunction. These are (a) the way in which data is shared between the model and the GIS, (b) the way in which the modelling functions have been implemented and (c) the user interface of the coupled system. This is illustrated in Figure 1 below. Schematic diagrams of examples of loose, tight and integrated coupling are given below in Figure 2, Figure 3 and Figure 4, respectively. These show how data-sharing, user-interface and modelling development strategies in coupled systems affect each other.

Chou and Ding also proposed a methodology for the coupling of modelling systems and GIS. The following steps were recommended:

(a) identify integration purpose and performance requirement;
(b) develop a conceptual model;
(c) map the conceptual model to a data model;
(d) evaluate different integration alternatives; and
(e) implement the design.

Table 1 Applications coupling environments (Nyerges, 1992)

APPLICATIONS COUPLING ENVIRONMENT :				
Issue	Isolated	Loose	Tight	Integrated
coupling strategy	•data transfer off-line	•reference/cross index on-line	•interoperable store services	•common store services
data model similarity (internal system)	•different or same	•different or same	•different or same	•same
data model construct resolution (data transfer service)	•manual resolution of data construct differences; •stored off-line	•manual resolution of data construct differences; stored •on-line	•automatic resolution of data construct differences; stored •on-line	•no need for data construct resolution;
data communications	•not needed	•high speed phone or ethernet	•data bus or ethernet	•data bus
user interface	•independent process supervisor	•independent process supervisor	•independent or common process supervisor	•common process supervisor
frequency of interaction	•one-off	•few	•often	•all the time
speed of interaction	•slow	•tolerable	•reasonable	•fast
amount of software development	•none-low	•low-some	•some-lots	•lots
hardware platform	•different or same	•different or same	•different or same	•same
personnel involved	•different	•different or same	•different or same	•same
example :	•off-line transfer GIS data to model	•on-line transfer between model and GIS	•seamless GIS & model integration through AIF	•custom decision support system

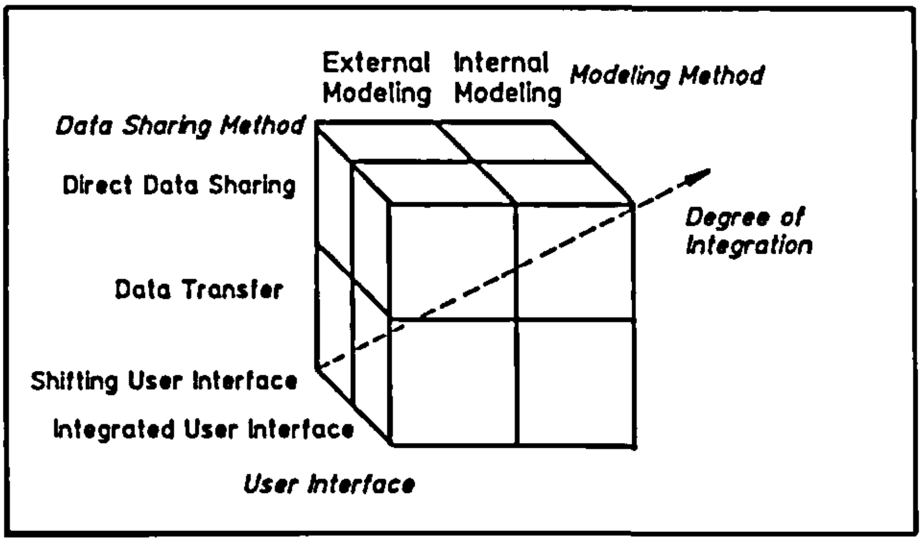


Figure 1 A Taxonomy of Integration from Three Perspectives (from Chou and Ding, 1992)

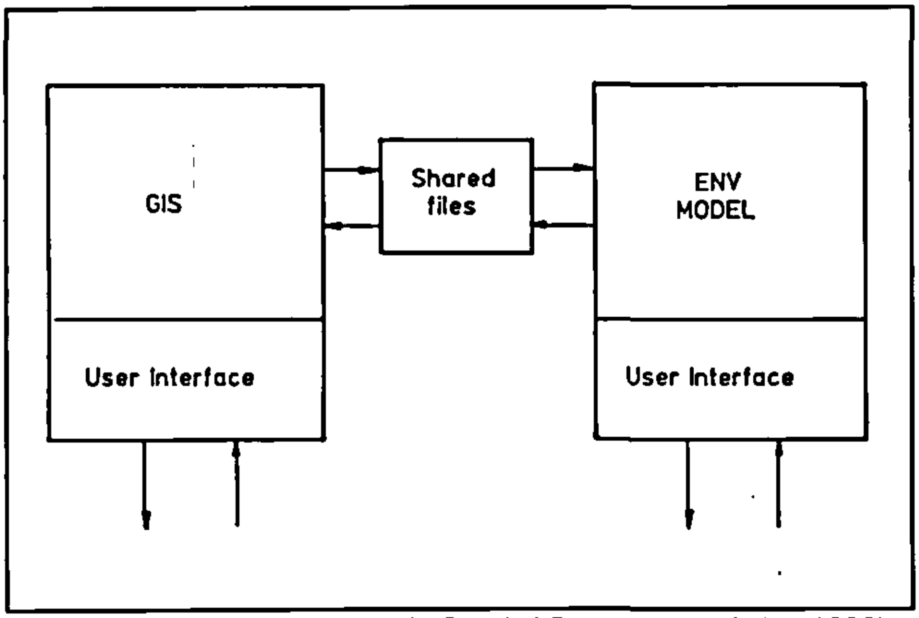


Figure 2 A Structure for a Loosely Coupled System (from Fedra, 1993)

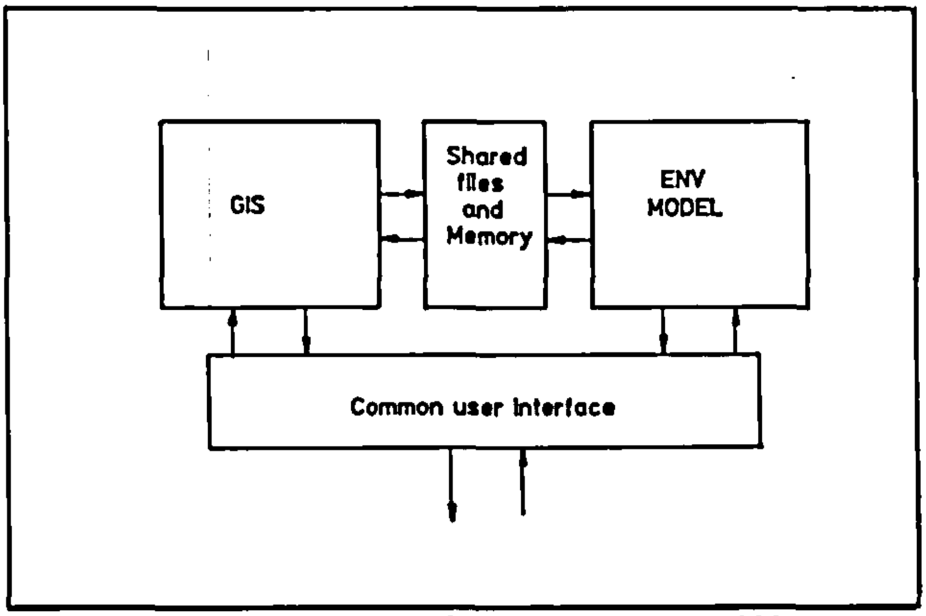


Figure 3 An Architecture for Tight Coupling (from Fedra, 1993)

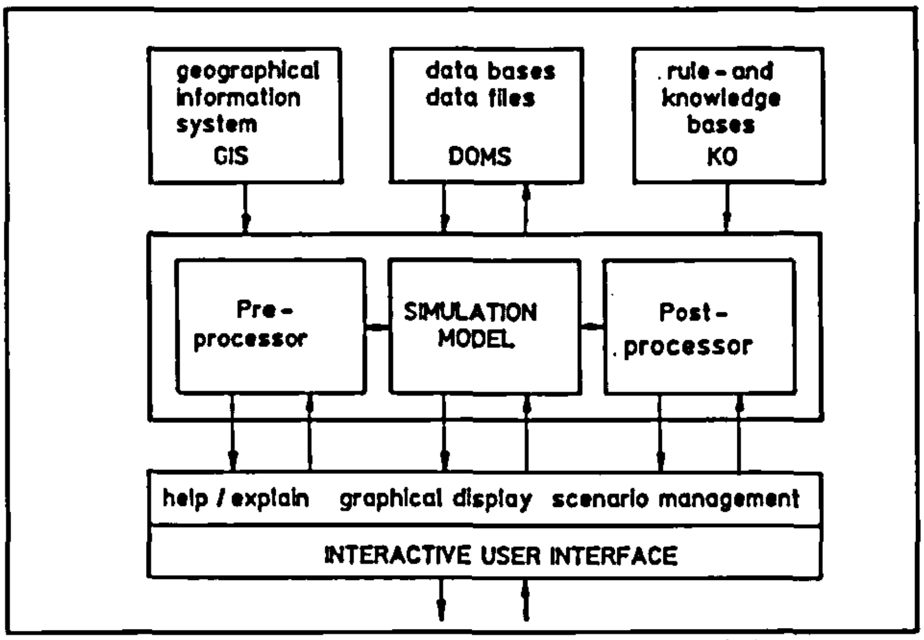


Figure 4 An Architecture for an Integrated System (from Fedra, 1993)

2.1.2 Effective Interfacing

Nyerges (1992) stated that the effectiveness of coupled systems should be evaluated before considering a coupling strategy, using the criterion that it should provide "practically adequate information for decision-making (p. 536)". As coupling generally proceeds through the data management subsystems, as discussed earlier, particular attention needs to be given to the role of data in models.

Data is necessary in order to calibrate a model, i.e. to quantify model parameters and/or coefficients so that the model results agree with observed values. Empirical models tend to make use of regression techniques for calibration, while process models can be calibrated using field measurements. The model should also be validated, that is, the results of the model should agree with observed values at points not used for calibration (Beven, 1989).

It is also important to know how errors in the input information affect the model results, and, in particular, whether the errors exceed acceptable levels. In simpler models this may be done analytically, but in more complex models this cannot be done, and the errors in the model results are estimated by running a number of simulations where inputs are varied, and the changes in results observed. It is important to note which variables are most sensitive to error, i.e. those variables which produce the most significant model output error, when subjected to realistic variation. This process is known as sensitivity analysis (Jørgenson, 1990).

In order to take account of spatial variation in the area being simulated, model users might discretize the study area into units that can be considered homogenous. Although the use of smaller cells may be considered a more realistic structure, independent data are often not available for each cell, and values have to be interpolated from sampled points. Burrough (1989) has termed this "the parameter crisis (p. 7)". The precision of the spatial resolution of the model exceeds the accuracy of the information available to calibrate it. While there are a variety of geostatistical methods available to assist in such problems, they should be used judiciously (Burrough, 1986). These issues are discussed with specific reference to the hydrological literature in section 2.2.

Veregin (1989) stated that decision-makers should, as a basic minimum, be provided with a means of assessing the accuracy of the information upon which their decisions are based in the context of GIS. In the light of the preceding discussion, this remark appears to be equally applicable to the development of coupled models and GIS. The uncertainty in the GIS products passed to the model must be known, so that error propagation or sensitivity analysis can be performed, to estimate the reliability of the model results.

Lanter (1990) investigated various mechanisms for automatically tracing the processes to which spatial data is subjected in a GIS, termed *lineage*, and concluded that none of the existing mechanisms were adequate. The US National Committee for Digital Cartographic Data Standards has proposed that a lineage statement accompany data to be exchanged, in a formal quality report. Other authors have also stressed the importance of formalising procedures to ensure the quality of GIS products, whether suggesting the use of artificial intelligence approaches (Couclelis, 1992; Lanter, 1992) or extensions to GIS data models (Burrough, 1992a; Goodchild, 1992b). The link between data quality in GIS and the effective use of GIS with modelling has been commented on by Openshaw, Cross and Charlton (1990), Burrough (1992b) and Stoms, Davis and Cogan (1992). Only Burrough has discussed the broader implications of this issue for the coupling of GIS and models, however.

2.2 Relevant issues in distributed hydrological modelling

2.2.1 Classification

The purpose of this section is to provide a framework for the discussion of hydrological models. The schema of mathematical methods in hydrology developed by Fleming (1975) (see Figure 5 below) is used as a guide. Fleming first of all makes a division between *optimization* and *non-optimization* methods. Optimization methods provide explicit solutions to clearly-posed problems, while "nonoptimizing methods are generally associated with the assessment of hydrologic data and are used to quantify physical processes" (p.27).

At the second tier, a distinction is drawn between *statistical* and *physical/deterministic* hydrology. Since this report is concerned with the group of models that attempt to represent physical processes, they fall under the heading physical/deterministic. As discussed previously, *empirical* models, are concerned more with achieving appropriate output than with faithfully representing physical processes. Process models are termed *conceptual* by Fleming. Digital models are listed as *indirect simulation* models. These models are further subdivided using a number of criteria, the most important one for the purpose of this project being the spatial structure, described here as *lumped* or *distributed.*

Lumped catchment models need to be calibrated using time series of inputs and outputs. Distributed models may in principle be parameterised using field measurements (Beven, 1989). Beven (1985) lists the most significant application areas as:

1) forecasting the effects of land-use change;

2) forecasting the effects of spatially varying inputs and outputs;

3) forecasting the movement of pollutants and sediments and

4) forecasting the hydrological response of ungauged catchments where no information is available

to support the use of a lumped model.

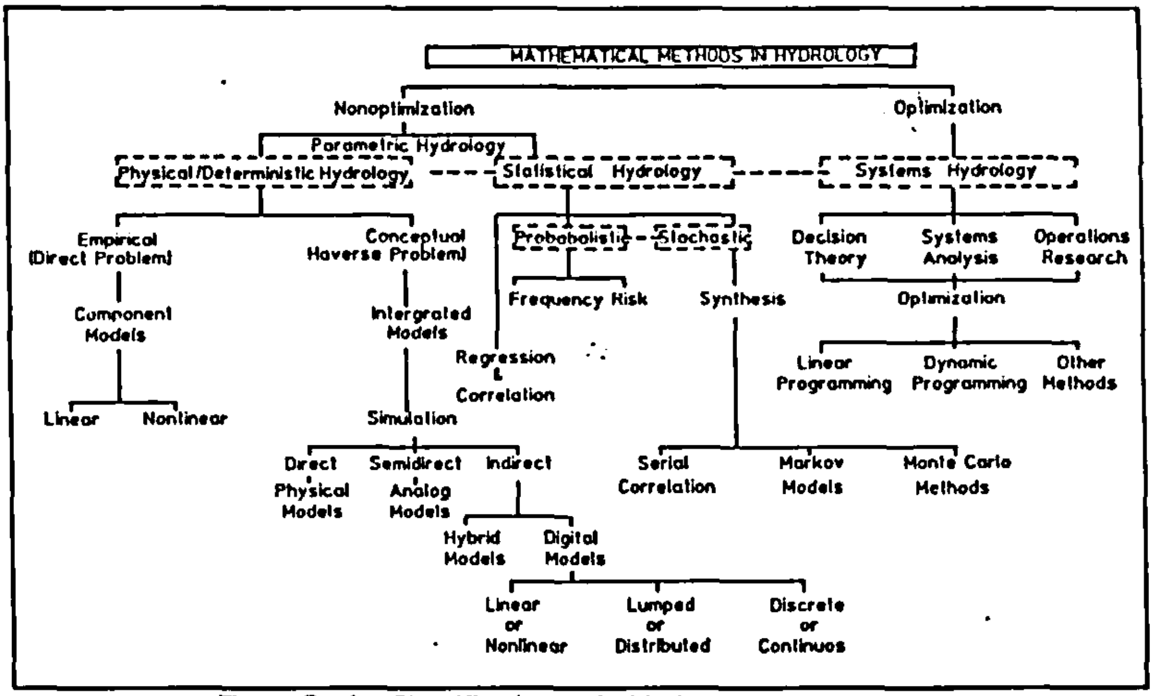


Figure 5 A Classification of Mathematical Methods in Hydrology (from Fleming, 1975)

There are a number of obstacles to the development and use of distributed models, a process which has been described as 'slow and faltering' by Beven (1985, p. 406). Issues surrounding resources needed for applying distributed models are discussed in section 2.2.3.

2.2.2 Model Structure and Spatial Structure

Lumped catchment models generally draw on ordinary differential equations or a series of linked ordinary differential equations (Karnieli, 1991). Detailed physical processes cannot be simulated in lumped catchment models because of the coarseness of their spatial structure. Only a model which uses spatially varying parameters and input can attempt this. Two distinct representations of space are generally integrated: a surface from which runoff is modelled, and secondly a channel network for the routing of flow (Kemp, 1992).

One way to do this is to treat the catchment as a series of cells which are effectively internally homogenous. Output from each cell is then routed between cells until they leave the catchment. Since the processes occurring in each cell are treated separately, calculations remain relatively simple, although the routing system may be quite sophisticated. This approach is known variously as cell-based (Schulze *et al*, 1989), semi-distributed (Karnieli, 1991) or representative elementary area (Moore *et al*, 1993). The term semi-distributed will be used in this report.

A more rigorous approach to modelling treats the study area as a continuum as far as possible. This requires the use of complex partial differential equations, which can only be solved approximately (Anderson and Rogers, 1987). An appropriate discretisation of space and time must be chosen for the study, in order to achieve a solution using numerical methods. Although the actual processes being simulated occur in three dimensions, it is usual to represent the system in terms of interacting one and two-dimensional components because of the computational requirements of the former (Beven, 1985). This type of approach has been called distributed or physically-based (Anderson and Rogers, 1987; Beven, 1985).

One prominent example of the distributed approach is the SHE model (Abbott *et al*, 1986). The SHE model uses a regular grid in its standard form (Abbott *et al*, 1986; Anderson and Rogers, 1987); Rohdenburg, Diekkrüger and Bork (1986) described a modified version of SHE using a variable grid. The Institute of Hydrology Distributed Model (IHDM) uses a series of planes (Beven, 1985; Rogers *et al*, 1985). Both models represent river reaches as one-dimensional features.

It should be noted that the term 'distributed' seems to be used loosely, with different meanings in different contexts. One example is the ANSWERS model. In describing the model Beasley, Huggins and Monke (1980) state that a distributed parameter concept is applied in the model. They also state that processes in each element in the model are treated independently. It would therefore be described as a semi-distributed model according to the terminology developed above. However, Jõao and Walsh (1992) describe the model as distributed in their account of the linkage of ANSWERS with a GIS.

ANSWERS is an example of a model which includes an explicit description of the terrain, like the SHE and IHDM, but it is not as strict in its application of physical equations. This type of model falls between the semi-distributed and distributed classes as described above, and will be referred to as terrain-based (Moore, O'Loughlin and Burch, 1988; Band, 1989).

2.2.3 Constraints to Hydrological Modelling

Anderson and Burt (1985) state that while there are two major constraints in modelling efforts, i.e. computing facilities and the available data base, it is the second factor that is truly limiting. As mentioned in section 2.1, data is required for the parameterization, calibration and verification of models.

Although distributed models can in principle be parameterized using field measurements, the large number of parameters and spatial elements typical of such models make this impractical and extensive interpolation is necessary (Beven, 1985). For these reasons, distributed models are usually calibrated (Anderson and Rogers, 1987). There are significant problems associated with the calibration of distributed models because of the large number of parameters they use, and the possibility of complex interactions between them.

These observations indicate that it is important to gauge the uncertainty inherent in model results (Beven, 1989; Binley and Beven, 1991). Beven (1985) argued that it is neither possible to eliminate input error nor even to assess it directly; instead he suggested that techniques should be applied to evaluate model output as a stochastic function of uncertain parameterisation. Abbott *et al* (1986) describe an application procedure for the SHE model using several different parameter sets to estimate the variation in model results for a given scenario. The use of formal mathematical techniques is discussed by Moore and Rowland (1990), Binley and Beven (1991) and Beven and Binley (1992). Some of these methodologies can be very demanding of computer time (Binley and Beven, 1991).

When model performance is expressed in a stochastic form, validation becomes much more complex, and Anderson and Rogers (1987) state that this is a field which still requires much research. The question of adequate data availability surfaces again in this regard (Anderson, 1989). Although the literature discussed here is focused on distributed models, similar problems apply to models that are less precise in their characterisation of spatial variability (Beven, 1989).

Remote sensing has been cited as a possible source of information for hydrological modelling, and is considered suitable for integrating spatial information at the catchment level (Anderson and Burt, 1985). Suggestions that it may even be used at a model grid scale (Abbott *et al*, 1986; Drayton, Wilde and Harris, 1992) have been questioned, however (Beven, 1989). Bathurst and O'Connell (1992) contend that remote sensing may be able to fulfil such a role in the future, but concur with Beven (1989) that field studies are of critical importance in supporting model development.

2.3 Applications of GIS in hydrological modelling

Moore *et al* (1993) have distinguished several levels of modelling with GIS. These are hydrological assessment, hydrological parameter derivation and linking hydrological models with GIS. The use of GIS for data management in water resources is briefly discussed, together with hydrological assessment. Applications of remote sensing and digital terrain modelling are considered with reference to GIS, before turning to the literature on the linkage of hydrological models and GIS.

2.3.1 GIS for Water Data Management and Hydrological Assessment

The use of GIS for data management in hydrology has been discussed by a number of authors, including Goulter and Forrest (1987), Schoolmaster and Marr (1992) and Leipnik, Kemp and Loagicia (1993). Samuels *et al* (1991) described the Environmental Display Manager, a GIS-like system developed at the US Environmental Protection Agency, for the integration of their national database, especially on water quality. Schoolmaster and Marr discussed the use of GIS for analysing water use data, together with SAS, an example of the linking of GIS and analysis software. Leipnik, Kemp and Loagicia (1993) looked at the implementation of GIS for water-resources management, addressing management level issues rather than scientific/technical questions. Some weaknesses of GIS were identified, specifically the difficulty that present GIS have in handling data that are not precisely defined in space or that change over time. Seasonally variable dams and braided river courses are examples of such data.

Shea *et al* (1993) state that custom software was used to handle gridded data for hydrological modelling as early as the 1970's, of which Davis (1978) is a good example. DeVantier and Feldman (1993) cite the work of the US Army Corps of Engineers' Hydrologic Engineering Centre in developing GIS-type raster software for use with hydrological models. This work was extended to receive input from satellite imagery, and an interface was developed to a digital elevation model (DEM). Goulter and Forrest (1987) developed a conceptual model for the use of GIS in water resources management, including how GIS could be articulated with mathematical models to form a decision support system. Their paper addresses GIS on an institutional level, which is beyond the scope of this report, but it will be discussed in more detail under related research.

Maidment (1993) used the term hydrological assessment to describe the use of GIS to investigate a problem by combining the relevant information in a map presentation, usually for the purpose of risk assessment. Walsh (1987) discusses the assessment of non-point source pollution, while Lynch (1993) describes the application of a weighted map overlay procedure to map groundwater pollution potential. This kind of approach presents a static snapshot of a situation, in contrast to the dynamic character of a process model.

2.3.2 Satellite Remote Sensing, Digital Terrain Modelling and Hydrological Modelling

Remote sensing, digital terrain modelling and GIS are synergistic tools for hydrological modelling (Moore, 1991; Drayton, Wilde and Harris, 1992; DeVantier and Feldman, 1993; Moore *et al*, 1993). The literature on remote sensing and hydrological modelling and digital elevation modelling and hydrological modelling is extensive, and the purpose of this section is to show their relevance to the project topic rather than to review them.

Remote sensing is a source of grid-structured information that has been explored as an input for hydrological models (Rango *et al*, 1983; Rango, 1985) before the advent of GIS. Although remote sensing has been used to provide land-cover (Rango, 1985), snow-melt (Engman and Gurney, 1991) and rainfall (Rott, 1986) information for hydrological modelling, Engman (1993) stated that the only truly successful applications have involved parameterizing hydrological models which draw on land-cover information. Engman also suggests that remote sensing has great potential for the determination of soil moisture, although this has yet to be properly investigated.

GIS is increasingly being seen as an appropriate way to manage remotely sensed information (Davis and Simonett, 1991; Davis *et al*, 1991; Estes, 1992), and a number of papers discussing the use of remote sensing for hydrological modelling integrate both of these approaches (Drayton, Wilde and Harris, 1992; Kouwen *et al*, 1993; Stuebe and Johnston, 1990; White, 1988). These papers discuss the use of the US Soil Conservation Service (SCS) runoff curve method.

Digital elevation modelling is another approach to data management for hydrological modelling which has close links to GIS. Moore *et al* (1991) reviewed applications of digital elevation models (DEMs) in a variety of disciplines, including hydrology, while Moore *et al* (1993) examined digital elevation modelling applications in the context of GIS. Studies such as those by Bork and Rohdenburg (1986), Jenson (1991) and Moore *et al* (1991) indicate that important parameters for hydrological modelling may be obtained from the analysis of digital elevation models, which may be used to parameterise distributed models (Band, 1989). Table 2 below lists important topographic parameters for hydrological modelling.

Table 2 Important Topographic Attributes for Hydrology (adapted from Moore *et al*, 1991)

Attribute	Definition	Hydrologic Significance
Altitude	Elevation	Climate, vegetation type
Upslope height	Mean height of upslope area	Potential energy
Aspect	Slope azimuth	Solar irradiation
Slope	Gradient	Overland and subsurface flow
Upslope slope	Mean slope of upslope area	Runoff velocity
Dispersal area	Area downslope from a short length of contour	Runoff volume, steady-state runoff rate
Catchment length[1]	Distance from highest point to outlet	Overland flow attenuation

There are a number of ways in which elevation information may be represented, the main one being a) the regular grid, b) the triangular irregular network (TIN) and c) the contour-based network (Moore, Grayson and Ladson, 1991). These various spatial representation are illustrated in Figure 6.

Grids consist of regularly spaced points on a triangular, rectangular or square mesh. The fundamental unit in a grid is the cell, enclosed by three or four grid points for triangular and rectangular meshes,

[1] All attributes except this one are defined at points within the catchments

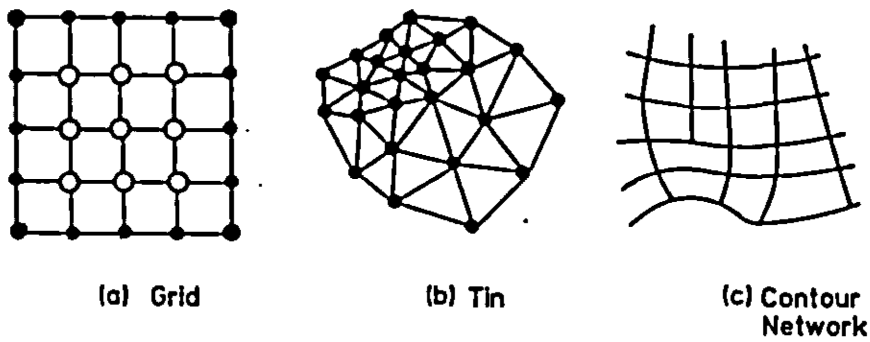


Figure 6 Commonly-Used Methods for Representing Topography (adapted from Moore, Grayson and Ladson, 1991)

respectively. The TIN method only requires that elevation data be captured at topographically significant points in the area of interest. An irregular network of points is stored, the plane joining three adjacent points being known as a facet. Contour networks use the cartographic convention of representing height using lines of equal altitude.

Grid systems are widely used, because of the ease with which computations can be carried out on the data. The TIN structure has the advantage that the spatial resolution of the model can be varied, and it is therefore a more efficient storage structure than the grid. It is more difficult to calculate derived characteristics such as slope and aspect from a TIN than from a grid, however. Contour approaches require significantly more storage than the other methods, and are no easier to handle computationally. Moore, Grayson and Ladson (1991) state that a structure chosen to handle digital elevation information for hydrological modelling should take the process being modelled into account, rather than being dictated by data storage considerations. Contour-based DEMs have significant advantages over the other representations used, because their basic structure corresponds with the way water flows overland (Moore, O'Loughlin and Burch, 1988).

All three types of DEM have been used in hydrological modelling, but the grid store approach has been used most commonly. The SHE model (Abbot *et al*, 1986), ANSWERS (Beasley, Huggins and Monke, 1980) and AGNPS (Panuska, Moore and Kramer, 1991) are three examples of hydrological models incorporating terrain information using a regular grid. Jett, Weeks and Grayman (1982) and Vieux (1991) demonstrate the use of TINs.

Moore, Grayson and Ladson (1991) state that the development of DEM-based models is still in a very early stage, and call for further research into the relative effectiveness of these structures.

The analysis of digital elevation models may also be used to automatically extract channel networks, delineate watershed boundaries and flow paths (Jenson and Domingue, 1988; Tarboton, Bras and Rodriguez-Iturbe, 1991). Digital elevation models have been incorporated into the structure of a number of hydrological models (Moore *et al*, 1991), as was indicated in section 2.2. One function of such DEMs has been to route the flow of water in both terrain-based and fully distributed models (Gandanoy-Bernasconi and Palacios-Velez, 1990; Quinn *et al*, 1991).

Much research remains to be done on the use of DEMs in hydrological models (Moore *et al*, 1991). Stuebe and Johnston (1990) note problems with basin delineation in flat terrain, while Jenson (1991) indicated that the cell resolution chosen in the analysis of slope from a DEM had a significant effect.

2.3.3 Interfaced Systems

DeVantier and Feldman (1993) and Zhang, Haan and Nofziger (1990) have reviewed the literature on the use of GIS in hydrological modelling. Both reviews indicated that the major application has been in the pre-processing of data for the model. This consists of the calculation of the spatial distribution of hydrological parameters, usually using automated map overlay techniques. Horn and Granger (1993) describe the use of a spatial database of river reaches for pure network modelling, which does not involve any overlay process. Presentation of the results of the modelling was discussed in less detail.

A large number of papers which discuss applications involving the interfacing of GIS and hydrological models focus on the modelling, and discuss only a few of the many issues that have been raised in this literature review. References have therefore been organised into three tables, each of which uses a different set of criteria to organise the literature. The first table (Table 3) is based on the aspects of linking GIS and models discussed in section 2.1, and is divided into spatial database development, the system interface, the user interface and the effectiveness of the linkage. Table 4 looks at the model type, as discussed in section 2.2. Finally, applications involving remote sensing and digital terrain modelling with GIS are listed in Table 5.

The classification of papers by model type is meant only as a guide, as many of the papers only discuss the model briefly. Even so, it is obvious that the bulk of applications have been in the use of semi-distributed and terrain based models with GIS. The complexity of distributed models probably accounts for the low number of applications with GIS. The SHE model has recently been coupled with GIS in a number of applications (Danish Hydraulic Institute, 1993; Gao, Soroosh and Goodrich, 1993).

The topic of database development emerges clearly as the most common element in the references. The systems- and user-interfaces receive much less attention, as might be expected in a literature that is primarily concerned with applications rather than technology development. Most of the references noted here deal with the coupling of existing hydrological models and GIS, rather than the development of integrated systems. The effectiveness of the coupled systems is not evaluated in most cases. Heuvelink, Burrough and Stein (1989) and De Roo, Hazelhoff and Heuvelink (1992) are important exceptions. Grayson *et al* (1993), Maidment *et al* (1993) and Moore *et al* (1993) discuss the place of GIS in hydrological modelling critically, and conclude that more research and technology development are necessary in this area. These papers draw on both the GIS literature, discussed in section 2.1, and issues arising from the hydrological literature mentioned in section 2.2.

The papers that include remote sensing and digital elevation modelling tended to be more research oriented, and often involved the use of custom developed GIS-like systems. In remote sensing these were generally grid data management systems (Drayton, Wilde and Harris, 1992; Kouwen *et al*, 1993), while digital elevation model software focused on the analysis of the data rather than data management (Moore, O'Loughlin and Burch, 1988; Band, 1989; Vieux, 1991). Mainstream GIS systems are now starting to include DEM analysis capabilities for hydrology. The GRASS GIS has had a basin delineation module incorporated into it, as a result of ongoing research using this system (Stuebe and Johnston, 1990). Arc/Info and IDRISI are two other systems that incorporate tools for the hydrologic analysis of DEMs (ESRI, 1991c; Eastman, 1992).

GIS has been used successfully to parameterise a range of hydrological models, particularly those designed to handle sub-catchment variability explicitly. Dramatic time-savings have been reported (Wolfe, 1992), but the conclusion of review papers has been that the cost-effectiveness of linking GIS and hydrological models has yet to be proved (Zhang, Haan and Nofziger, 1990; DeVantier and Feldman, 1993). The capture and manipulation of data using GIS is expensive, and it may be more difficult to work with inadequate data using GIS than if an experienced individual tackled the problem manually (Hastings and Moll, 1986). DeVantier and Feldman (1993) believe that a broad based study needs to be undertaken to resolve these issues, but suggest that this is not likely to be done before this approach becomes more widely used. It is agreed that GIS is best used when there will be repeated use of the spatial database. Maidments' call (1993) for more remote sensing and a national soils database suitable for use in a GIS can be seen as an attempt to address the matter.

Maidment (1993) and Moore *et al* (1993) discuss the conceptual basis for linking hydrological models and GIS, as well as the opportunities and obstacles in this field. Maidment suggests that GIS can be used to move beyond the capture of data and display of model output to address more substantive issues, such as the modelling of partial area flow and the interaction of surface-water and groundwater. He also proposed that areas that require development in GIS include data structures to handle variation with time and three dimensional data structures for groundwater modelling. The essentially static nature of present GIS hinders the use of models to identify spatial patterns as a function of process, as suggested by Moore *et al* (1993).

Grayson *et al* (1993) suggest that GIS can be dangerous to hydrological modelling, because it is not appreciated that the generation of datasets at a finer resolution than the source data does not lead to an increase in hydrological information. They also state that user-friendly GIS interfaces may lead to an abuse of hydrological models if unqualified individuals are allowed to operate them, and because they hide the inherent complexity of hydrological modelling. These concerns are shared by Moore *et al* (1993), as well as other researchers who wish to see GIS which can indicate the sources of derived data sets, specify the uncertainty in GIS products and interact intelligently with models (Burrough, 1989; Lanter, 1989; Burrough, 1992a; Burrough, 1992b; Couclelis, 1992; Lanter, 1992).

The use of expert systems (Couclelis, 1992; Grayson *et al*, 1993) and object-oriented approaches (Burrough, 1992a) have been suggested as techniques which may assist in these matters. Djokic and Maidment (1991) discuss the use of an expert system to assist in the GIS parameterisation of a hydrological model, while McKim, Cassell and LaPotin (1993) note the potential of GIS to supplement the integration of remote sensing and object-oriented simulation.

Table 3 References of Applications Classified according to the Aspects of System Interfacing Discussed.

DATABASE DEVELOPMENT
Anderson, 1991; Bhaskar, James and Devulapalli, 1992; Bork and Rohdenburg, 1986; Cline, Molinas and Julien, 1989; Davis 1978; De Roo, Hazelhoff and Burrough, 1989; Jõao and Walsh, 1992; Karnieli, 1991; Loveland and Johnson, 1983; Meyer, Salem and Labadie, 1993; Moore, O'Loughlin and Burch, 1988; Rango *et al*, 1983; Ross and Tara, 1993; Sasowsky and Gardner, 1991; Tarboton, 1992; Vieux, 1991.
SYSTEMS INTERFACING
Cline, Molinas and Julien, 1989; Jõao and Walsh, 1992; Kilborn, Rifai and Bedient, 1990; Meyer, Salem and Labadie, 1993; Rango *et al*, 1983; Shea *et al*, 1993; Tarboton, 1992.
USER INTERFACE
Bitters, Restrepo and Jourdan, 1992; Cline, Molinas and Julien, 1989; Tarboton, 1992.
EFFECTIVENESS
De Roo, Hazelhoff and Burrough, 1989; Grayson *et al*, 1993; Maidment *et al*, 1993; Meyer, Salem and Labadie, 1993; Moore, O'Loughlin and Burch, 1988; Moore *et al*, 1993; Stuebe and Johnston, 1990.

Table 4 References of Applications classified according to the Type of Model Used.

LUMPED MODEL
Rango *et al*, 1983.
SEMI-DISTRIBUTED MODEL
Purwanto and Donker, 1991; Karnieli, 1991; Kouwen *et al*, 1993; Sasowsky and Gardner, 1991; Stuebe and Johnston, 1990; Tao and Kouwen, 1989; Tarboton, 1992.
TERRAIN-BASED MODEL
De Roo, Hazelhoff and Burrough, 1989; De Roo, Hazelhoff and Heuvelink, 1992; Drayton, Wilde and Harris, 1992; Jõao and Walsh, 1992; Moore, O'Loughlin and Burch, 1988; Vieux, 1991.
DISTRIBUTED MODEL
Gao, Sorooshian and Goodrich, 1993; Bathurst and O'Connell, 1992; Wolfe, 1992.

Table 5 References of Applications which discuss Satellite Remote Sensing and Digital Terrain Modelling.

SATELLITE REMOTE SENSING
Drayton, Wilde and Harris, 1992; Engman, 1993; Kouwen *et al*, 1993; McKim, Cassell and LaPotin, 1993; Rango, 1985; Rango *et al*, 1983; Sasowsky and Gardner, 1991; Tao and Kouwen, 1989.
DIGITAL TERRAIN MODELLING
Band, 1989; Bhaskar, James and Devulapalli, 1992; De Roo, Hazelhoff and Burrough, 1989; Donker, 1992; Drayton, Wilde and Harris, 1992; Jett, Weeks and Grayman, 1982; Moore, Grayson and Ladson, 1991; Moore, O'Loughlin and Burch, 1988; Silfer, Hassett and Kinn, 1987; Stuebe and Johnston, 1990; Vieux, 1991.

2.4 Spatial decision support systems

The impulse for the coupling of GIS and modelling systems has come from a number of different groupings, with rather different interests. The spatial analytic research community have been interested in performing spatial analysis while using GIS (Goodchild, 1987; Goodchild *et al*, 1992). The emphasis in spatial analysis is on the application of specialised statistical tests to data held in a GIS, which requires mathematical analysis generally not practical to perform using the native capabilities of a GIS. This has lead to the interfacing of GIS and specialized spatial analysis software, as reported by Openshaw, Cross and Charlton (1990) and Ding and Fotheringham (1992), while Haslett, Wills and Unwin (1990) describe an integrated system. Goodchild *et al* (1992) discuss various ways in which this interfacing could be achieved, and developed a schema similar to the one discussed in section 2.1.

Spatial analysis has been used with a GIS approach in several hydrological applications. Ive, Walker and Cocks (1992) used spatial statistics to predict dryland salinization potential on a grid-cell basis. See, Naftz and Qualls (1992) used regression analysis with GIS to identify sources of selenium contamination in an investigation of the impact of irrigation return-flow on water quality. Both of these papers focused on the application, without describing the GIS interface. Rhodes and Myers (1993) discussed the use of kriging using a statistical package in conjunction with the GRASS GIS. Spatial analysis is directed to applying a test to a data set and returning a result, it does not have the dynamic element of process modelling, and is generally a more straightforward process. The passing of static datasets by file transfer is therefore perfectly adequate in most cases. The structures used in spatial analysis are also quite similar to those commonly used in GIS (Armstrong and Densham, 1990).

Spatial decision support systems (SDSS) is another related area of research. The developers of SDSS also want to move beyond the limited capabilities of GIS and integrate flexible modelling capabilities with their geoprocessing systems (Densham, 1991). Nyerges (1992) has noted that there is some debate about the distinction between a GIS and an SDSS. Densham states that in many cases GIS is defined as a tool to support decision-making; Cowen (1988) in fact defines GIS as SDSS. Armstrong and Densham (1990) and Densham (1991) take a contrary position, which is discussed in more detail below.

Densham (1991) states that the field of SDSS has developed in parallel with business Decision Support Systems (DSS), but with a lag of 10-15 years. According to Sprague and Carlson (1982), DSS has been differentiated from management information systems (MIS) and electronic data processing (EDP) in practice by its focus on high-level decision-making.

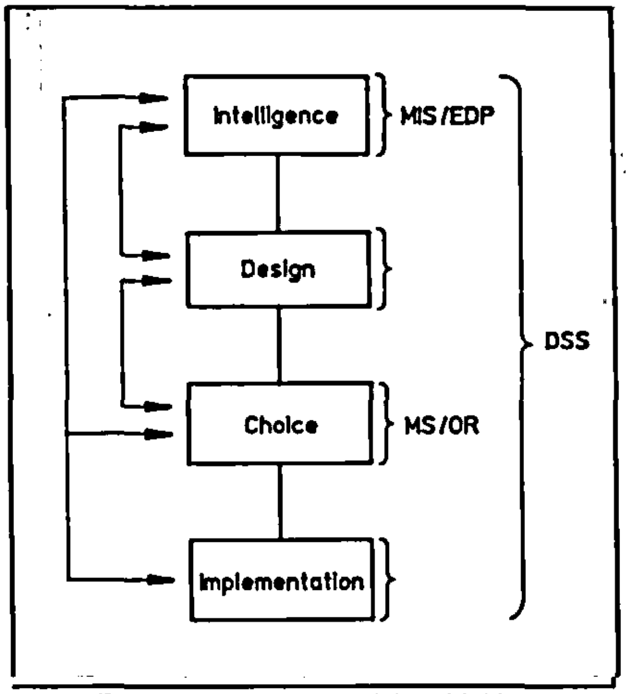


Figure 7 The Phases of Decision Making (from Sprague and Carlson, 1982)

A widely accepted model of the decision-making process in this context simply consists of three steps, called intelligence, design and choice (Sprague and Carlson, 1982). The first step is a search for issues requiring decisions, involving the collection and examination of data. The second step, design, consists of conceptualizing and analysing various strategies that might be adopted. Choice is the selection of a particular course of action, and its implementation. The scope of decision support systems can be contrasted with other approaches using this model of decision making (see Figure 7).

Drawing on Sprague and Carlson (1982), Guariso and Werthner (1989) and Armstrong and Densham (1990), some of the essential features of a DSS are:

1) they are designed to solve ill-structured problems by flexibly combining statistical analysis, models and data;

2) they have intuitive user interfaces and powerful graphical presentation capabilities which make them easy to use;

3) they are easily modified to suit the changing needs of the user.

Densham (1991) states that SDSS require the following characteristics in addition to those specified above:

1) the ability to capture spatial data;

2) the ability to represent the complex relations and structures of spatial data;

3) they must include analytical procedures "that are unique to both spatial and geographical analysis (p.406)";

4) they must produce output of spatial data in a variety of forms, not only as maps.

In order to achieve this variety of objectives, in particular the requirement of flexibility, a modular architecture is often proposed for DSS/SDSS (Sprague and Carlson, 1982; Guariso and Werthner, 1989; Armstrong and Densham, 1990; Densham, 1991). An example of such an architecture is given in Figure 8.

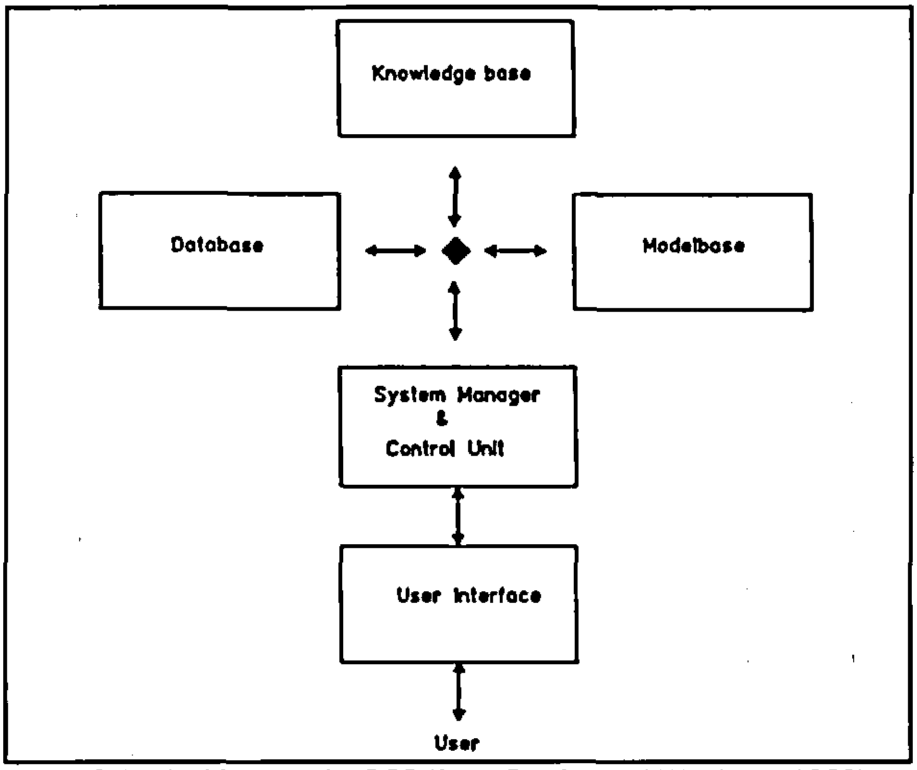


Figure 8 An Architecture for DSS (from Guariso and Werthner, 1989)

According to this proposal, a DSS will integrate several different types of functionality. The database management system (DBMS) is the heart of an SDSS (Armstrong and Densham, 1990). The inclusion of modelling capabilities can be done in a number of ways: using the query language of the DBMS, external subroutine libraries or a modelbase management system (Densham, 1991). Modelbase management systems (MBMS) are a new research field. They manage the procedural steps which are the atomic components of models. When an MBMS is required to apply a numerical technique, it retrieves the appropriate steps and executes them in the required order. The knowledge base contains higher order information such as the rules for constructing models, as well as information specific to the problem. The system manager controls the way in which the system components should interact (Guariso and Werthner, 1989).

An SDSS is clearly a highly sophisticated system, according to this description. Some authors have proposed that systems which do not fulfil all the criteria mentioned above can be used as DSS under certain conditions. Abbott *et al* (1986) stated that the SHE model can be considered to function as a DSS when it is used to investigate the impacts of different development scenarios. Bathurst and O'Connell (1992) suggested that SHE can be run "within decision support systems" (p.276), if the technically complex stage of setting up the model is separated from its application. One of the main requirements on the models to be integrated in a decision support system is that they must be easy to understand and use (Guariso and Werthner, 1989). It is not clear whether the two-stage solution proposed by Bathurst and O'Connell would be able to meet this criterion.

Fedra (1993) takes a position close to that of Bathurst and O'Connell, stating that "Calibration is a .... difficult task, but it can be separated from the interactive decision support use of a model" (p.41), while discussing the coupling of GIS and environmental models. Thus it appears that a hydrological model interfaced with a GIS might be considered an SDSS.

This discussion of SDSS has indicated that while this field is concerned with system architectures in a way that can inform the interfacing of modelling systems with GIS, it is focused on the development of integrated applications. This type of development is seldom attempted, and is usually of research interest only.

2.5 Conclusion

This review has indicated that while the conceptual basis for the coupling of hydrological models and GIS is relatively undeveloped, the practical use of GIS as a pre- and post-processor of data is well established. The present use of GIS with hydrological modelling is limited however, and there are significant obstacles to overcome in order for the full potential of this approach to be realised.
The advanced status of hydrological modelling (Kemp, 1992) has meant that sophisticated models are in common use. The complexity of these models makes them difficult to implement within GIS (Zhang, Haan and Nofziger, 1990), and the coupling of GIS and hydrological models therefore commonly involves a loose linkage. In this type of scenario, the spatial database is transferred from the GIS environment to a dedicated modelling system for processing, and the GIS is limited to a role as a pre- and post-processor of data. Another consequence is that user-interface development is generally limited, as the two systems operate quite independently of each other apart from at the time of data transfer.

Several areas have been identified where the present approach to interfacing GIS and hydrological models falls short. GIS data and products often have significant levels of error in them, but current GIS are not designed to handle this kind of information (Burrough, 1992b; Couclelis, 1992). A number of authors have stated that it is important to indicate the uncertainty in the results of hydrological modelling (Beven, 1989; Binley and Beven, 1991), which requires some estimate of input error. Moore *et al* (1993) have contended that lineage information, i.e. a description of the original sources and transformations applied to the data, should also be available in GIS for hydrological modelling. This would allow users of the database to assess the suitability of the data for different applications (Clouclelis, 1992).

Loosely interfaced GIS and modelling systems rely on the skills and judgement of their operators to ensure that the quality of the data supplied to the model is adequate. If tools were to become available to track lineage and data quality in GIS, a much more complex data set would have to be transferred to the modelling system in order to communicate this additional information. Burrough (1992b) suggested the development of a formal description of data quality, GIS analysis functions and the data requirements, sensitivity and error propagation characteristics of models. He envisaged this metadata being incorporated in a knowledge base with GIS in order to guide preparation of data for *inter alia*, modelling.

This idea is closely related to the concept of DSS discussed earlier, with an architecture including database and modelbase management systems, as well as a knowledge base of higher-order information about the functioning of the system. The type of integration envisaged by Burrough (1992) would involve the development of an SDSS with a highly sophisticated knowledge base. Although models of error propagation in GIS have been developed by researchers, they are only beginning to be implemented in commercial GIS (Goodchild, 1992a). Formal descriptions of data quality and model characteristics proposed by Burrough simply do not exist. In addition, the development of such integrated systems involves substantially larger investments than the loose coupling which is prevalent today. The cost-effectiveness of even the latter approach has yet to be demonstrated, and it seems likely that skilled human supervision will continue to be an essential in the operation of coupled GIS and hydrological models.

CHAPTER 3 - CHARACTERISTICS OF THE DISA MODEL AND THE ARC/INFO GIS

The features of the Daily Irrigation and Salinity Analysis (DISA) model and the Arc/Info GIS of importance to the interface development are discussed in this chapter, in order to be able to assess how they might be coupled.

3.1 DISA

The basic concepts underlying the DISA model are considered in this section, with special reference to the model structure and the spatial data requirements of the model. This section draws heavily on the DISA Model Development Report (Beuster, Görgens and Greyling, 1990a) and DISA User Guide (Beuster, Görgens and Greyling, 1990b).

3.1.1 Background and Concept

The Department of Water Affairs and Forestry (DWAF) commissioned the development of the DISA model in 1988, specifying that it should be able to:

(1) predict the effect of expanding irrigation and changing irrigation practices on the water draining into the Breede River and

(2) indicate the impact of a variety of planning and operational methods on river salinity.

Given these requirements, the simulation of physical processes was emphasised and a choice was made to avoid a calibration approach. While the model had to realistically simulate the flow of water and salts in the system, the data demands of the model also had to be kept from becoming prohibitive.

The Breede river catchment is in a winter rainfall region, and as salinisation is only a problem in the summer the model was designed without surface runoff processes. In accordance with the objective of avoiding calibration, the model was run from the middle of the winter season to allow equilibrium conditions to be established within the model by the start of the irrigation season.

3.1.2 Model Structure

As a physical process model, DISA had to explicitly describe spatial variation within the study area. The catchment was divided into a number of relatively homogenous physiographic units, with similar agricultural and hydrological features. These physiographic regions are shown in Figure 9 below. Subdivisions were made at gauging stations and wherever future control points had been identified by

DWA planning staff. Thirty-three physiographic regions were defined in this way. These regions were further divided into sixty-six return-flow cells, distinguishing between the river alluvium and the higher-lying terrace areas.

Beuster, Görgens and Greyling (1990a) state that "the water supply system acts as the main driving force in the model (p.13)". The return-flow cells and water supply system therefore had to be integrated by the model. This was done using the concept of a network of modelling nodes. Five different types of nodes were used:

a) Abstraction nodes, points where water is abstracted from the river or canals.
b) Inflow nodes, which manage the input of water and salts from areas outside the model boundary.
c) Farm dam nodes, which combine the storage of multiple farm dams on each return flow cell, and accept input from and give output to a variety of other model elements.
d) Canal nodes, which model canal reaches, and perform daily mass-balance calculations.
e) River channel routing nodes, which simulate river channel processes.

Another type of node was also described, the reservoir release control node (Greyling, Beuster and Görgens, 1993). This differs somewhat from the other nodes, in that it regulates the behaviour of the main reservoir (the Greater Brandvlei Dam) in response to the behaviour of the simulated system, using a user-defined set of rules.

The Breede river system and associated water distribution system was modelled by assembling elements to logically represent the physical layout. Part of the configuration of the model for the existing Breede River system is given in Figure 10, Figure 36 and Figure 11.

This report focuses on the spatial aspects of the model structure, but it should be noted that the DISA model simulates the physical processes in irrigated model elements in detail. DISA is able to model the behaviour of irrigated areas with a number of different crop types, using several different water sources and irrigation methods. The movement of water and salts is treated in such a way that each responds with different time lags. The river channel routing system responds sensitively to changes in the water supply regime because it uses a sub-daily time step.

The central processes simulated in DISA and the flows between the parts of the system are shown in Figure 12. The irrigated and non-irrigated areas are modelled separately (b), and that a distinction is drawn between the active and inactive soil within the irrigated areas (c). This figure also shows that the various areas noted above are lumped for the purposes of modelling, which indicates that DISA is a semi-distributed model in terms of the terminology introduced in Chapter 2.

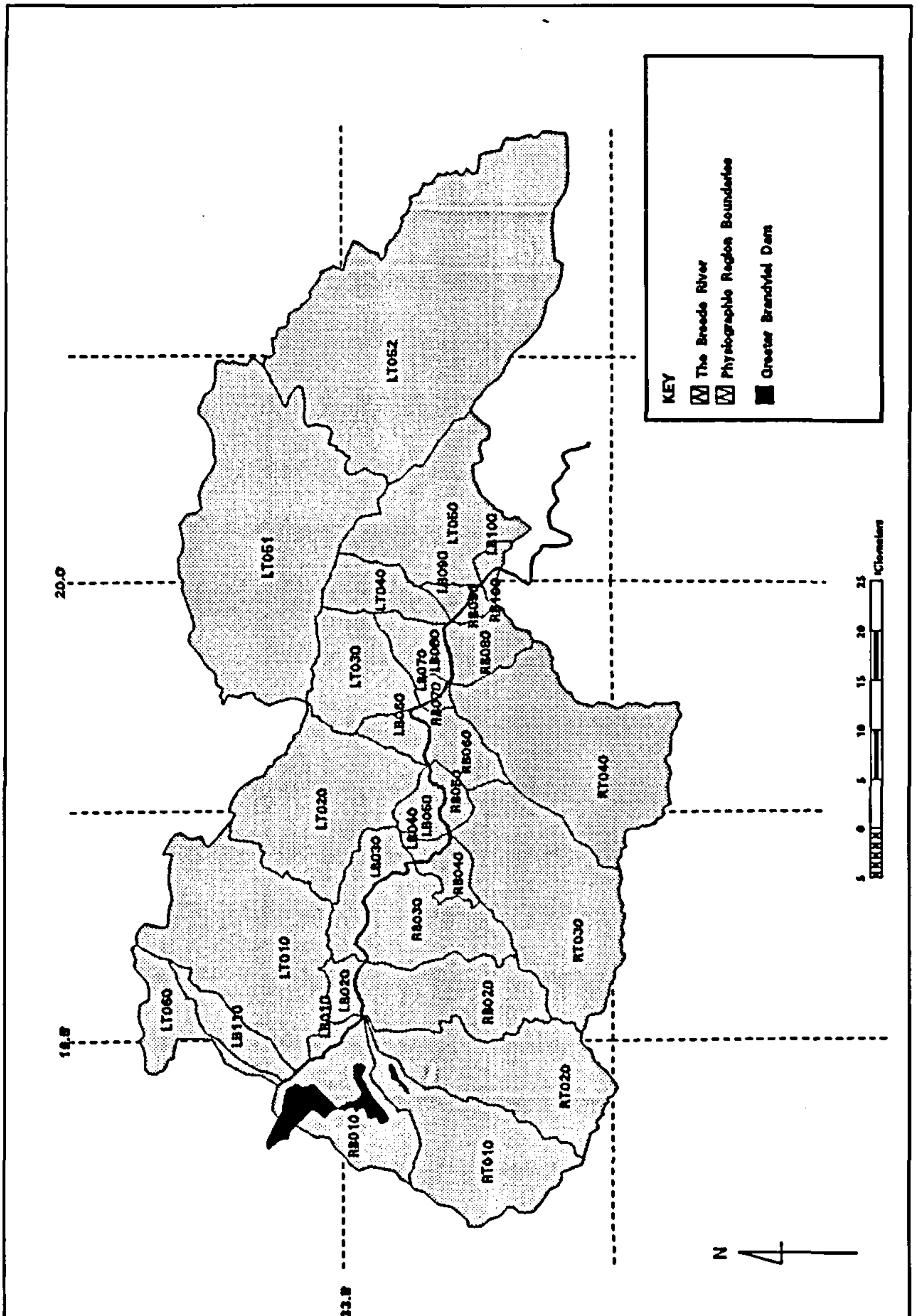


Figure 9 The Breede River Valley and Physiographic Units.

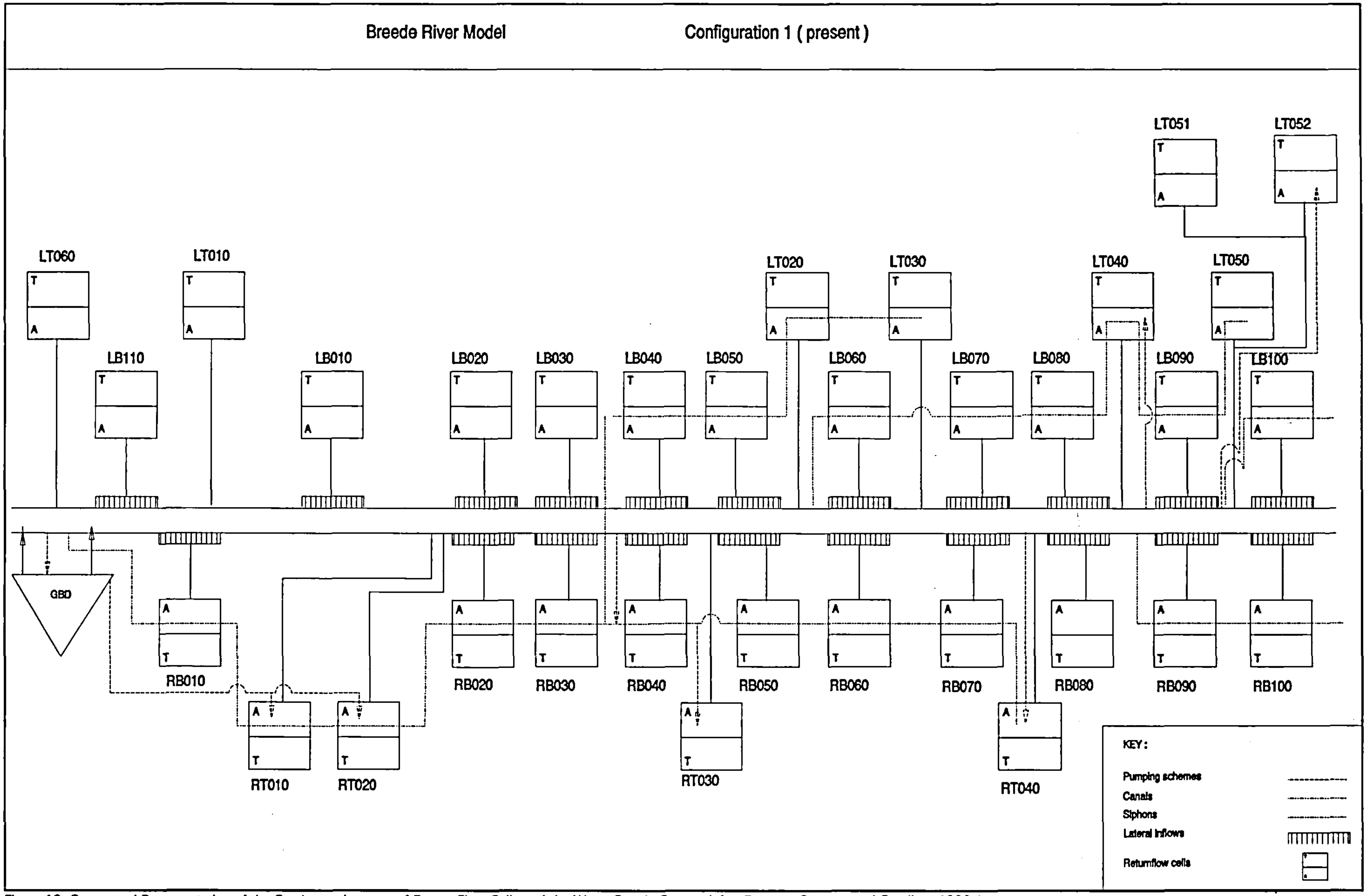


Figure 10 Conceptual Representation of the Catchment in terms of Return-Flow Cells and the Water Supply System (after Beuster, Görgens and Greyling, 1990a)

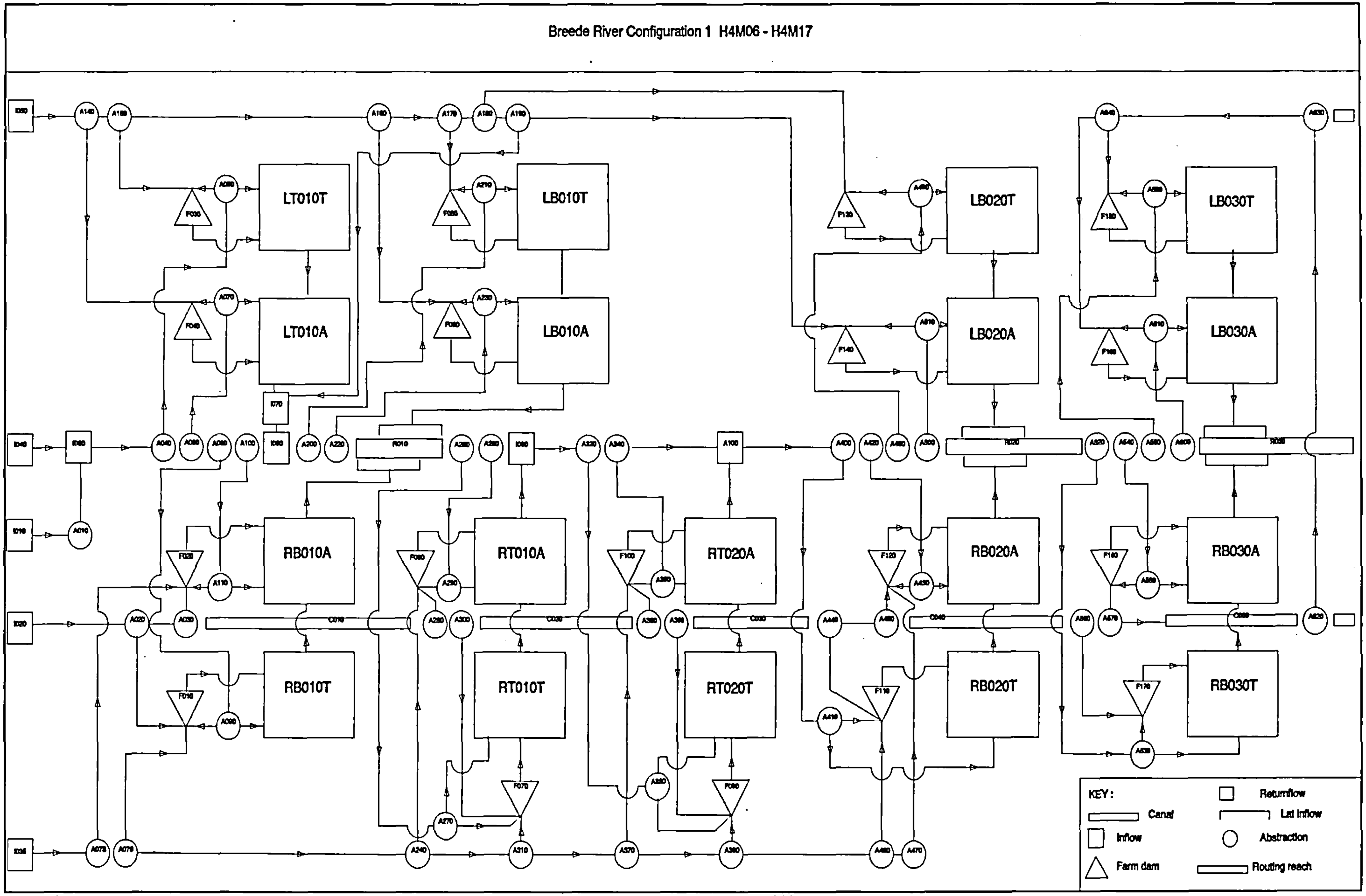


Figure 10 Detailed Configuration Diagram for the Upper Part of the Breede River Valley (after Beuster, Görgens and Greyling, 1990a)

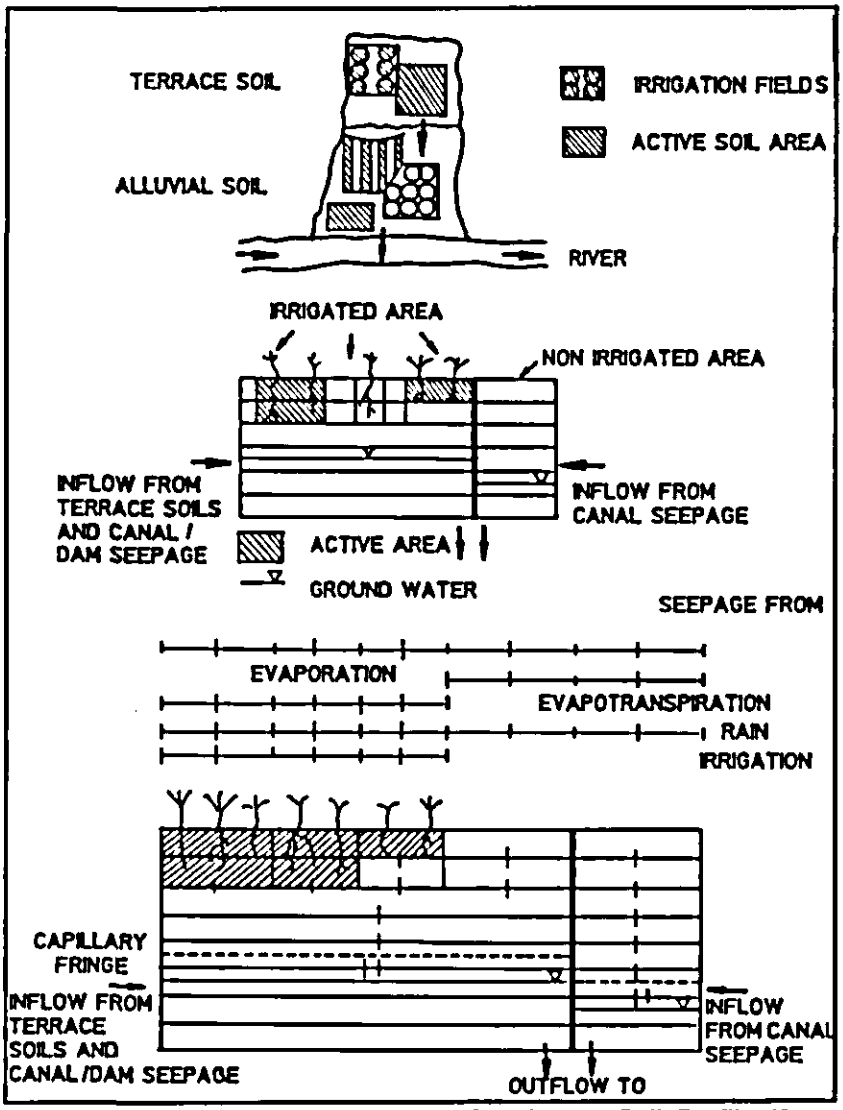


Figure 12 Simulation of Water and Salt in the Soil Profile (from Beuster, Görgens and Greyling, 1990a)

3.1.3 Data Requirements

The data required by the model consists of two distinct sets: a set of time series of evaporation, rainfall, flow and salinity that represent inputs from beyond the boundary of the catchment being simulated, and a second set describing the modelling elements. Only the second set of data has an explicit spatial component, and is considered in this project.

The data describing the modelling elements may be summarised as follows:

a) global parameters, which are constant for all elements in the system, including crop factors, soil characteristics, root distributions and active soil layers;

b) information specific to each modelling node must also be supplied, to indicate the flows between the various nodes, crop types and extent of planting and irrigation methods used;
c) the nature and location for which output is required.

The spatial data required by the DISA model for the configuration of the system is summarised in Table 6 below. Thematic attributes are included where appropriate, in brackets.

Table 6 Summary of Spatial Data Requirements of the DISA Model

Modelling Element Type	Associated Spatial Information
Abstraction Node	None
Inflow Node	None
Farm Dam Node	Maximum volume.
	Volume, area and depth relationship for area-capacity curve (seepage rate).
Canal Node	Canal section length and width.
Routing Node	River reach length.
	Section length and slope.
	Cross-section geometry.
Return-flow Cell	Cell type, soil type, irrigated area and aquifer width (drained outflow).
	Irrigated units, areas and crop types (water sources and type of irrigation systems used on each unit).

3.1.4 Software Structure

The software was developed as a modular system, consisting of:

a) the *database manager*, which controls and manages the time series data required for the simulation;
b) the *system configurator*, which sets process parameters, specifies the water distribution

network and the points for which simulation output is required;

c) the *system model*, which is composed of a number of process sub-modules simulating the physical processes involved in the flow of water and salts, and draws on the time-series database and the system description in each run of the model;

d) the *output manager*, which presents the results of model runs in either tabular or graphical form. Time series, percentile curves and longitudinal salinity profiles can be displayed.

The system configurator stores data in a configuration file, in ASCII format. This file is the key to the development of an interface between DISA and Arc/Info.

Figure 13 below shows a schematic representation of the DISA model suite.

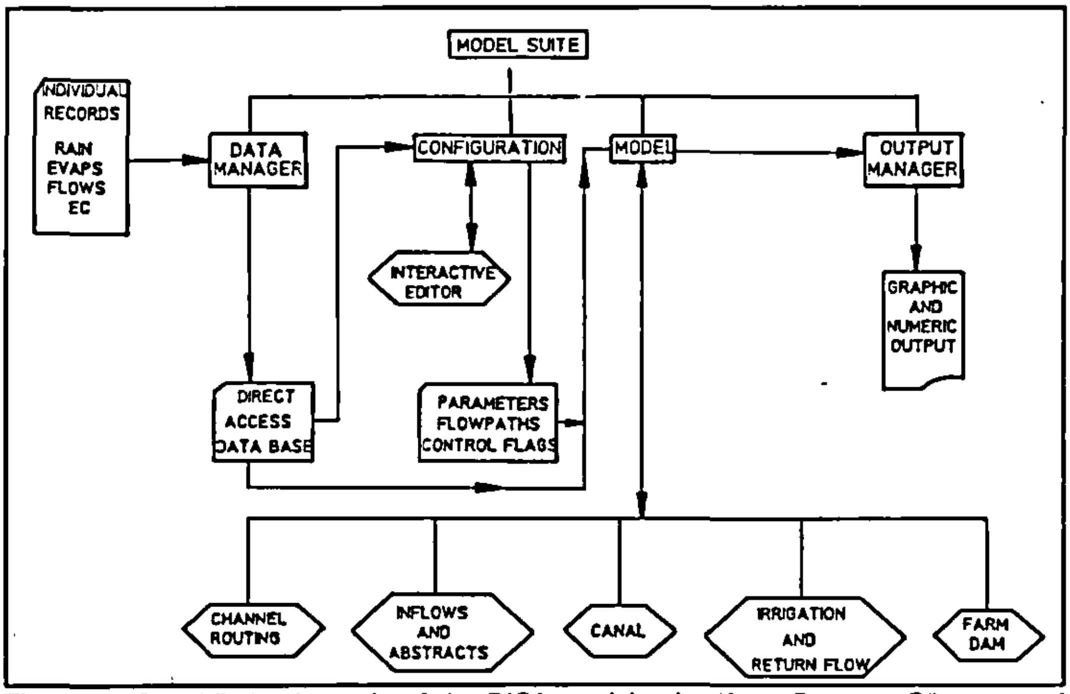


Figure 13 Simplified schematic of the DISA model suite (from Beuster, Görgens and Greyling, 1990a).

The output generated by the simulation run from the system model may be viewed from the output manager. Tabular output may consist of monthly mass balances of flow volumes and salt tonnage, or monthly deficits for abstraction and farm dam nodes only. Graphical output consists of daily water and salt time series, or groundwater levels and salinities in return flow cells. This output may be viewed on the screen or plotted on a pen plotter.

DISA was developed with a fully menu-driven user interface, which integrates the various software modules. The operation of the model and display of results can therefore be performed with ease. The initial design of the model network and the configuration of the model are still formidable tasks,

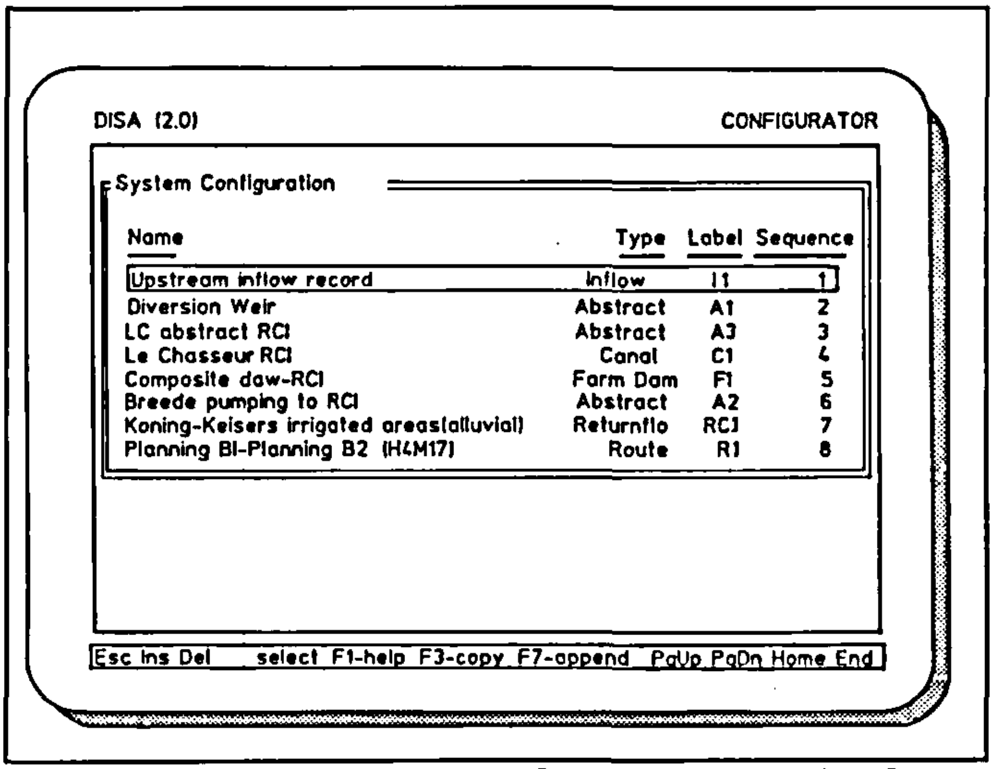


Figure 14 DISA Model Elements Menu for System Configuration (from Beuster, Görgens and Greyling, 1990b).

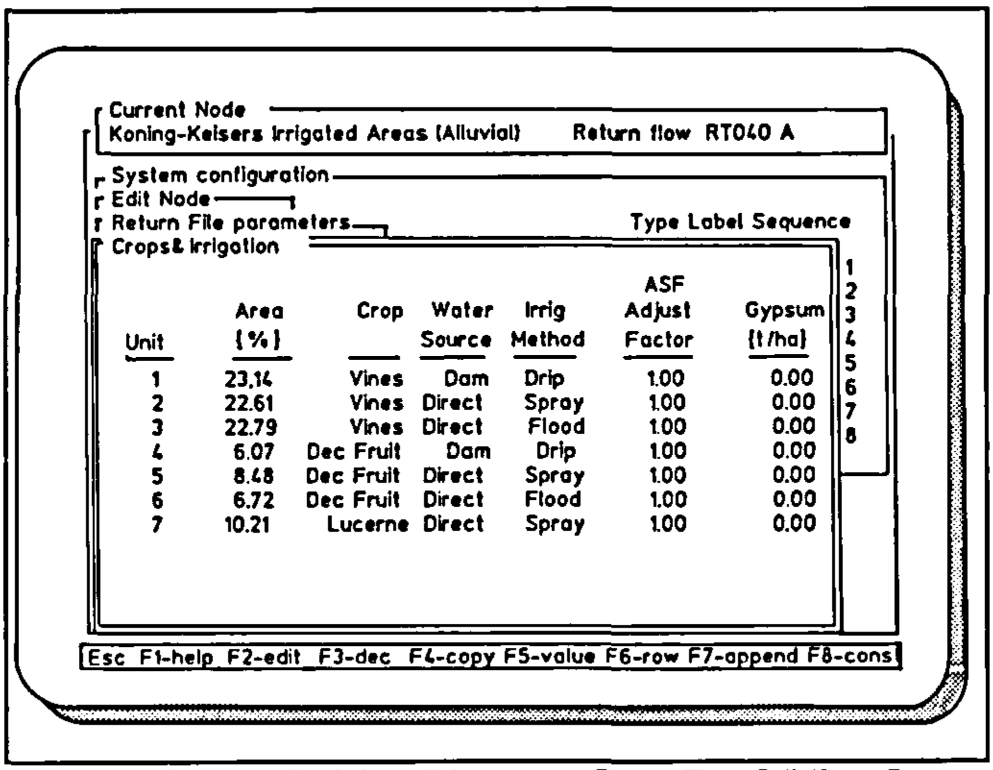


Figure 15 DISA menu for Irrigated Areas on a Return-Flow Cell (from Beuster, Görgens and Greyling, 1990b)

although the menu system helps structure these processes. The menu system for the configuration of the model is organized around a list of the model elements as shown in Figure 14, which corresponds to the configuration diagram shown in Figure 11. Further information for a model element can be obtained after it has been selected from the main list (see Figure 15)

3.2 The Arc/Info GIS

This section discusses the Arc/Info GIS with an emphasis on the features of relevance to using it with a hydrological model. These are the software structure, data models, analytical capabilities and automation and interfacing facilities.

3.2.1 Overview

The Arc/Info geographic information system is one that is widely used, and has been employed in a range of natural resources management applications. The overall architecture of this GIS is known as the hybrid or georelational model, because a standard relational database management system (INFO) is used to handle thematic attribute data, while separately developed software handles the spatial data (Morehouse, 1992). Arc/Info draws on the relational data model to link spatial and attribute data using a common key, the user identity number, just as INFO can link separate data tables using a common key item. This is discussed at some length in the following section.

The software structure of Arc/Info is essentially modular, with four different environments. Spatial data capture and editing is performed in the ArcEdit module, batch processing (including topological overlays) occurs in the Arc environment, ArcPlot caters for display and query of the database and the INFO database is used for complex manipulation of the tabular data. Arc/Info can also be connected to a number of industry standard database management systems.

3.2.2 The Arc/Info Data Model

The Arc/Info data model makes provision for handling a variety of types of spatial data, including raster, vector, surface, image and CAD information. The vector representation uses points, lines and polygons to represent features on the earth's surface projected onto a two-dimensional surface. Arc/Info uses a cartographic data model (Tomlin, 1990), organizing different data sets into separate 'layers', known as *coverages* in Arc/Info (see Figure 16).

Figure 17 illustrates how point information is handled using Arc/Info's georelational data model. Arc/Info generates a data table for this coverage, with one record in the table for every point. This Point Attribute Table contains the user-defined identifying numbers assigned to the points during

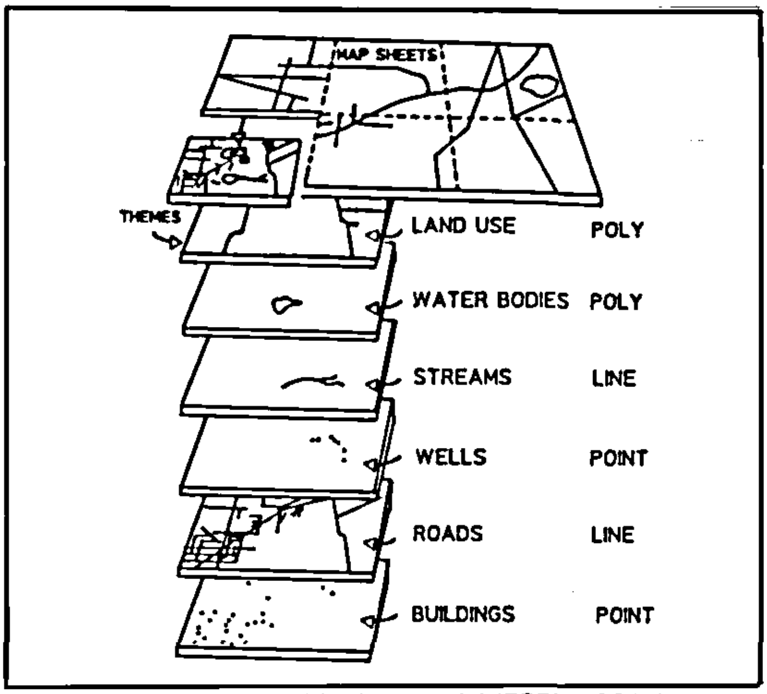


Figure 16 The cartographic data model (ESRI, 1991a)

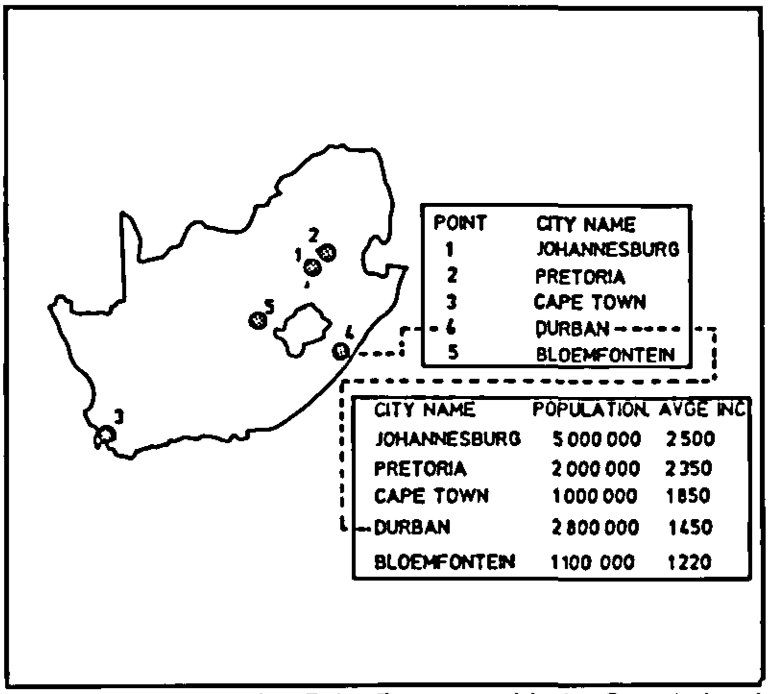
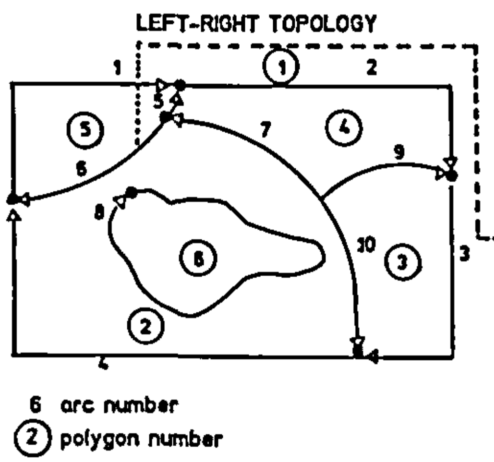


Figure 17 Representing Point Features with the Georelational Data Model

data capture, as well as system-defined entries. The point coordinates are not displayed in this file, instead this information is maintained separately by the software. Additional fields may be added to the attribute table by the user to hold additional thematic information, and it may also be linked to other tables in the database using a relational join.

Similarly, line and area features have corresponding Arc Attribute and Polygon Attribute Tables (AATs and PATs), which are illustrated in Figure 18. These tables contain additional information about these features, such as length for lines and area and perimeter for areas. Arc/Info also stores information about the spatial relationship between linear and area features explicitly, drawing on the theory of directed graphs. In graph theory, a line is conceptualized as consisting of a set of vertices, called an arc, and its endpoints, which are termed nodes (Frank, 1992). Lines may only join at nodes. This is called arc-node topology, and provides the basis for systematically describing how arcs connect to each other in the database. Arc/Info also assigns a direction to each arc, by identifying a from- and to-node in the AAT.



Arc#	From node	To node	LPoly	RPoly	Length
1	-	-	1	5	-
2	-	-	1	4	-
3	-	-	1	3	-
4	-	-	1	2	-
5	-	-	5	4	-
6	-	-	2	5	-
7	-	-	2	4	-
8	-	-	2	6	-
9	-	-	4	3	-
10	-	-	3	2	-

Poly#	Area	Perimeter
1	-	-
2		
3	-	-
4	-	-
5		

Figure 18 The Storage of Topological Relationships between Arcs and Polygons in Arc/Info's Attribute Tables (adapted from ESRI, 1991a)

Area features, called polygons in Arc/Info, are represented using a list of the arcs that bound them. This is called polygon-arc topology, and it is stored in the AAT as a list of right- and left-neighbour polygons for each arc. This formally defines which polygons are adjacent to each other. This data model has a number of advantages. First, by 'building up' features from less complex ones rather than specifying each one independently there is a reduction of redundant data being stored. Secondly and more importantly, the topological information stored in the attribute tables can be used to perform analytical operations, which are discussed in the following section.

Arc/Info offers a variety of extensions to this data model. Attribute tables may be generated for nodes. Multiple arcs can be assigned to a single *route*, for the purposes of network analysis. Changing conditions along the course of a single arc can be represented using a feature called dynamic segmentation. The events along the arc are assigned relative positions in the database, the system calculates the absolute position at run-time. This is commonly applied to traffic networks, eg. to identify different speed zones along a road, but it also has hydrological applications such as representing the characteristics of a river reach.

Raster data can also be handled in Arc/Info, the raster grid having an associated data table called a Value Attribute Table containing a record of cell values and the total number of cells with the same value in the same data set. Unlike other attribute tables, there is no user-id. Satellite images may be imported as a raster, and displayed together with the vector data sets.

Arc/Info can model three-dimensional surfaces using either raster or vector representations. Each cell in a grid will represent a discrete elevation value, giving a blocky effect. The more sophisticated TIN model can use both spot heights and contours to develop a DEM, and breaklines can be incorporated to indicate discontinuities in the landscape. The TIN model does not have a fixed resolution, unlike the Arc/Info raster grid, so more data can be added to the model in areas of variable terrain without having to sample plateaux more intensively. The advantages of the TIN model are being brought into question by recent research, however (Goodchild, 1992). Both the grid and TIN models in Arc/Info can only handle a single z-value at a given x,y position, and do not have true three-dimensional topology. Raper and Kelk have called this type of model a "two-and-a-half dimensional representation (p.308)", because the z-value is stored as an attribute of a two-dimensional geometric data structure.

3.2.3 Analytical Capabilities

GIS are generally considered limited in their analytical capabilities, but Arc/Info offers a wide range of tools in comparison to most other systems. It has a raster modelling language based on the cartographic modelling system proposed by Tomlin (1990). This can be used to analyse grid DEMs and extract terrain parameters as dicussed in the literature review, which may then be passed to a hydrological model (see Kienzle and Lorentz, 1993). Heuristic models using cellular automatons can also be applied, eg. a model of fire spread is included as a demonstration with Arc/Info 6.1. This technique has also been applied to model drainage (Eli, 1990 cited in DeVantier and Feldman, 1993).

The vector analysis operations are based on Arc/Info's topological data model, and are much more limited. Points and lines can be buffered to produce polygons which include all areas within a specified distance. Thiessen polygon generation produces contiguous polygons from a set of points. Location/allocation modelling may be performed on a linear network, using the Route and Allocate submodules. Although these capabilities can be applied to the routing of water flow, the algorithm used is designed for traffic analysis, and cannot be used to route water when the flow is not steady (Maidment, 1993).

Topological overlays can also be performed between a polygon coverage and a point, line or polygon coverage. A topological overlay creates a new coverage which contains information on where the features being input coincide. Figure 18 shows the result of the topological overlay of two polygon coverages. TIN models may be used to perform visibility analysis, as well as the derivation of surface areas, volumes and profiles.

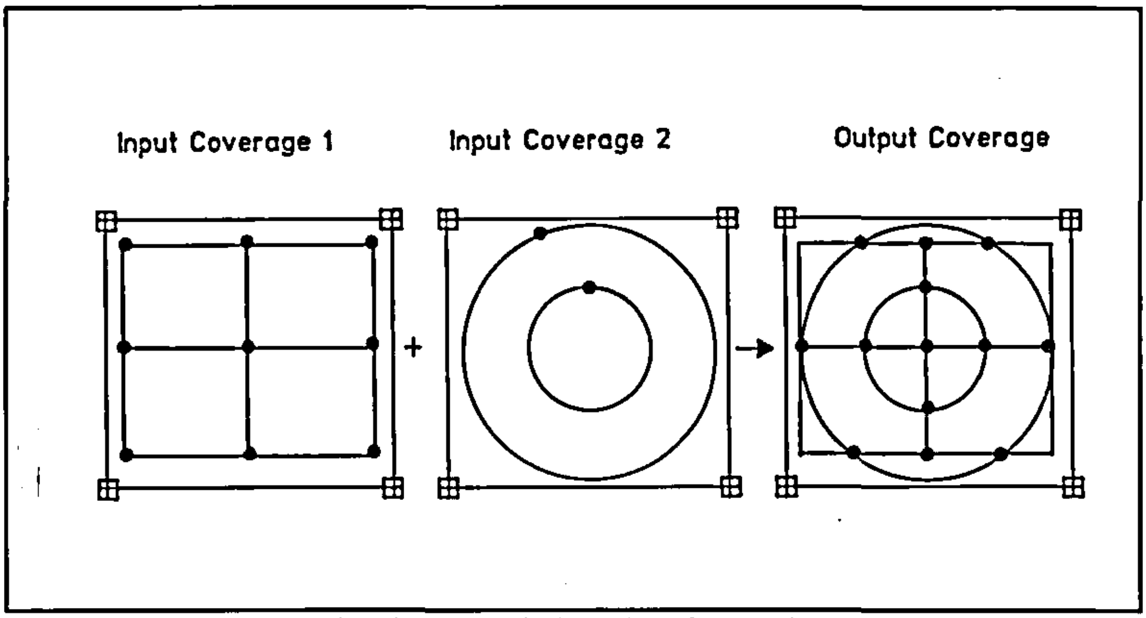


Figure 19 Schematic showing the topological overlay of two polygon coverages (from ESRI, 1991b)

INFO database programming may be used to relate and manipulate attribute tables, in order to carry out spatial analysis using the topological information that they contain. This can be a powerful means to analyse linear networks.

3.2.4 Customisation and Interfacing

The Arc Macro Language (AML) enables the automation of procedures in Arc/Info. AML is also a programming language, with standard flow-of-control and looping constructs, although it only provides scalar variables. AML can provide information about Arc/Info datasets in the form of standard variables, making it easy to interrogate a database.

Graphical user interfaces can be created using AML, in a variety of different styles. The form menu systems are the most complex, but the scrolling lists, check boxes and buttons that can be made available with form menus make it possible to design user-friendly systems. The ArcTools system demonstrates the use of an object-oriented approach to AML programming which uses form menus to present an attractive user interface.

AML has the ability to call an application programmed outside Arc/Info and take the value returned by the application. This enables an AML interface to be integrated with external software to provide functionality beyond Arc/Info's, while the user does not have to leave the Arc/Info environment. Arc/Info can import and export both spatial and tabular data in a variety of standard formats, making communication with external systems a straightforward matter.

3.3 Conclusion

DISA is a semi-distributed model, which requires detailed spatial information on the water distribution system and the irrigated lands for its configuration. It could therefore benefit considerably from linkage with GIS, which can capture and analyse spatial data. The DISA configurator file holds all the spatial data required by the model, and can easily be transferred to the GIS because it is in ASCII format.

Arc/Info can handle a variety of types of spatial data and can draw on a variety of different input sources, making it a flexible data capture tool. Arc/Info also has a range of data analysis functions which can be used to manipulate data to extract the information required by a model. The Arc Macro language provides a means to automate the interface system, and can be used to develop sophisticated user interfaces.

DISA and Arc/Info provide two complementary sets of functionality, which need to be brought together in order to assist in the application of DISA. Chapter Four discusses the process of designing an interface to accomplish this task.

CHAPTER 4 - INTERFACE DESIGN

The process of designing the interface for Arc/Info and DISA will be discussed by reviewing the data required, developing a conceptual view of the interface and then describing the practical implementation of the interface. A further distinction has been drawn between the system interface, which enables the two software systems to communicate, and the user interface which structures the way in which the system operator interacts with the overall system.

4.1 Data Analysis

Burrough *et al* (1988) recommended that an analysis of the data requirements of a model be performed before starting to design an interface between it and a GIS. The importance of data for effective distributed modelling identified in Chapter 2 also suggests that this is an essential step.

4.1.1 Available Data

The data requirements of the DISA model relevant to GIS were summarised in Chapter 3. This consisted of a list of data that would be suitable for capture by GIS in principle. GIS data capture usually proceeds from the digitising or scanning of maps. Not all of the spatial data that has been identified exists in map form, however.

The spatial information base for this project consists of the 1: 50 000 topographical map sheets of the study area, and overlays prepared for the entry of data for the model development project reported by Beuster, Görgens and Greyling (1990). Neither map sheets nor overlays were available for certain information however, and it was necessary to use the information from the DISA configurator file used in the verification study. The maps and overlays available for the project and the information to be derived from them for the configuration of DISA are summarised in Table 7 below.

Table 7 Available Spatial Data and Derived Information

Maps	Information derived for DISA
Model Cells	*Return-flow cell area*
Soils	" " "
River reach	Routing node length
Canals	Canal node length
Aquifers, geology, homogenous climate zones and points of interest	None

4.1.2 Original Sources

Only a very limited subset of the spatial data requirements of DISA were available in a form suitable for capture using GIS in this project. Several data sets were originally derived in a way that could have made use of GIS capabilities, however. It is very important that the location of irrigated areas be accurately identified for successful modelling with DISA (Beuster, Görgens and Greyling, 1990a), and the process used to derive this information illustrates the important role GIS could play.

An up-to-date survey of irrigated and potentially irrigable areas in the Greater Brandvlei Dam Government Water Scheme was available when DISA was being developed, but the results were presented on the basis of the original farm boundaries. DISA requires this data for each return-flow cell, and it was therefore necessary to obtain a more precise indication of the location of irrigated lands. A classified Landsat image was used together with digitised return-flow cell boundaries to obtain a first estimate of irrigation per return-flow cell. This information was then reconciled with the results of the farm survey, using digitised farm boundaries to reassign the image data on a common basis. The digitisation of the various boundaries was achieved using specially developed software, as was its integration with the remotely-sensed image. This capability is now available within the Arc/Info software.

The river reach cross-section information was obtained from analysis of 1: 6 000 scale aerial photography. Again, GIS can now be used to capture this information, along with coarser data such as the reach and section lengths. The spatial data used by DISA to characterise farm dams could also be derived from the analysis of aerial photography, using a GIS approach. Individual dams are lumped together in the DISA configuration, a common modelling practice (Mallory, 1993).

The soils information was derived from the maps prepared by the Soils and Irrigation Research Institute (now known as the Institute for Soil, Climate and Water). These maps were examined by a group of experts, and appropriate delineations of soil groups were made onto overlays by the DISA development team. These maps are now being made available in digital form, and the reclassification of soil categories is a straightforward procedure in GIS.

4.1.3 Lineage

The availability of data of adequate coverage and quality is essential for the successful interfacing of a GIS and a model (Burrough, 1989). A wide range of data sources have been integrated for the configuration of the DISA model. The present configuration was set up by a skilled team. The development of a GIS interface is intended to ensure that the DISA model can be more widely used, without such intensive involvement by specialist personnel. Grayson *et al* (1993) have warned that the development of GIS interfaces for models can mislead users into believing that the application of a complex hydrological model is straightforward.

As has been noted, the DISA model is very data demanding, and it has been necessary to synthesize various data sources and extrapolate from observed data even in a situation where current ground survey information was available. Thus the error component in the data used to configure DISA is due to error in the original inputs as well as the procedures used to integrate these sources. Even if error estimates were available for the original sources, it would not be possible to quantify the error in the derived information. A quantitative indication of data quality is therefore impossible.

Various authors have stressed the importance of *lineage* information in GIS (Chrisman, 1984; Lanter, 1990; Lanter, 1992). Lanter (1992) has reported that lineage has been defined by the US National Committee for Digital Cartographic Data Standards as "... information describing source materials and transformations used to derive final digital cartographic data files.... (p.2)". This information is intended to enable a user of data to determine whether it is fit for its intended use . Moore *et al* (1993) have stated that this is an issue of particular importance in the interfacing of hydrological models and GIS.

4.2 Conceptual Design

4.2.1 Data Flow

It was decided that the coupling of the DISA model and Arc/Info would be focused on assisting in the configuration of the model, with the GIS serving primarily as a spatial data capture and management system. The use of the interfaced system would start with the development of a system network for the DISA model. Having defined the features required by the model, the GIS would then be used to capture the spatial component of the data. The partial data set built up in this way would then be passed to DISA, where the rest of the information would be added using the existing model interface. Figure 20 overleaf illustrates this concept.The GIS would also be used to capture lineage information for the spatial database, to assist users to use the information appropriately.

The discussion in section 4.1 has indicated that only a small part of the potential spatial data sets were available for this project. As a result, a large part of the GIS database had to be read from the original configurator file. Data capture for the project therefore proceeded along two fronts: the capture and processing of spatial data sets to derive parameters for DISA, and the development of an import procedure for the configurator file.

This project has concentrated on the development of a GIS database and the automation of the exchange of data between the Arc/Info and DISA systems, and the user interface development has been focused on these areas. Data capture will proceed using the basic functionality of the GIS, although the development of the application methodology (see Appendix E) has included some automation of the analysis procedures.

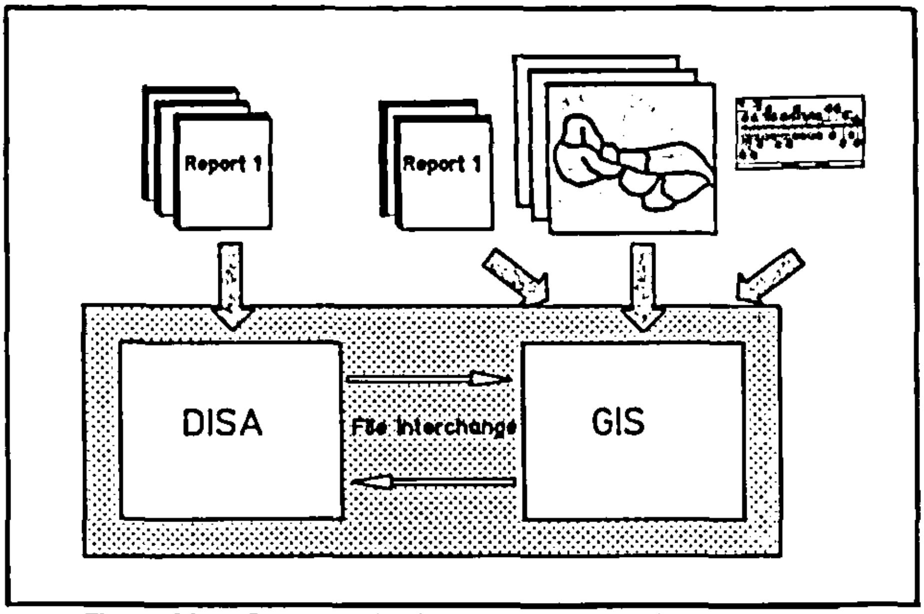


Figure 20 Conceptual diagram of the DISA - Arc/Info interface.

4.2.2 The Systems Level

Methodologies for the integration of GIS and modelling systems were discussed in the literature review. *The "taxonomy of integration" proposed by Chou and Ding (1992) is used to discuss the strategy decided* on for the interfacing of DISA and Arc/Info. The three criteria cited by Chou and Ding are the modelling method, the user interface employed and the data sharing method. The PC-based DISA model had been fully implemented and tested before this project commenced. The complexity of DISA meant that it was not feasible to consider implementing it within Arc/Info, and it was decided that the system would be used as it stood with Arc/Info providing only data capture, management and display services. This is known as *external modelling*, as opposed to internal modelling performed by the use of the GIS facilities.

The Arc/Info GIS was available on both PC and UNIX workstation for this project. The use of the PC offered the possibility of establishing a very close linkage between DISA and Arc/Info. A decision was made to use workstation Arc/Info, for a number of reasons. The workstation version of Arc/Info offers a number of facilities not found on the PC, including the extended data model discussed earlier, more powerful visualisation facilities and a much more sophisticated macro language and interface system. It is also the standard Arc/Info version, and applications developed for the workstation can easily be used with different combinations of hardware and software.

It was not possible to develop a single user interface for the DISA and Arc/Info systems in this project, because of the decision to use different platforms for the two systems. Instead, a *shifting interface* approach was adopted: the user interface already established for DISA was used to run it, while a totally separate user interface was developed to manage the data in Arc/Info. The awkward nature of communications between IBM PCs and UNIX workstations meant that *file transfer* was selected as the means of data sharing between the DISA model and Arc/Info, as opposed to directly sharing the same files.

The strategy that was decided upon makes full use of the established capabilities of the DISA model and required a minimum of additional programming. It was not optimal in terms of user-friendliness and the speed of the system interface, but the overall advantages associated with the use of workstation Arc/Info more than compensate for this.

4.2.3 The User View

The decision to develop a separate user interface for Arc/Info meant that users would be faced with learning two menu systems. It was therefore resolved that the Arc/Info interface should present the user with a view of the database which closely resembles that given by the DISA software itself. The DISA menu system for the configurator module presents a summary of the configured network, displayed in the order in which calculation takes place. Each type of model element has at least one menu dedicated to entering the data for it.

The system diagram which was shown in Chapter 3 (Figure 11) is central to guiding the configuration of the model. GIS offers the possibility of integrating this information into the database. Although a demonstration of this concept was developed, it was not possible to realise this goal in the operational user interface due to time constraints.

4.3 Implementation

4.3.1 Systems Coupling

Having decided on the use of separate computers, and therefore the use of linkage by file transfer, the details of the implementation had to be determined. DISA stores its configuration information in an ASCII file, which was simply transferred across an ethernet linkage to the workstation using the standard UNIX File Transfer Protocol (FTP). A 'layered' approach to the interface implementation was taken. The more complex manipulation of data was undertaken by programming in the C-language (see Appendix C), in conjunction with UNIX utilities and shell programs. Arc/Info macro language (AML) was used to drive the overall process and load the database after the translation of files was completed (see Appendix B). Some menu programming was also done using AML. This is illustrated in the schematic diagram below.

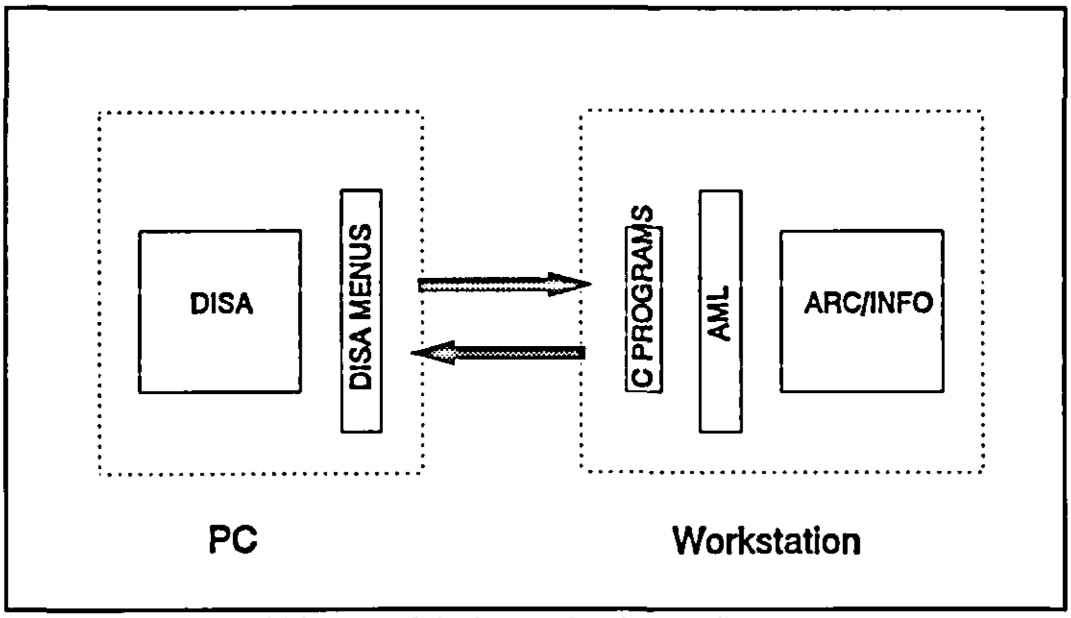


Figure 21 Conceptual Diagram of the Integration Approach

4.3.2 GIS Database Design

The DISA modelling element concept was used to guide the design of the GIS database. Arc/Info is built around a relational database system. The central task in the GIS database design was therefore to define the structure of the tables to be established in the database, as well as the relationships between them. Each type of modelling element was assigned a fixed set of associated data tables, based on the data requirements of that element and roughly corresponding to the DISA menus. The details of the full database design are given in Appendix A. The data tables were designed with a common key, the model element label, so that all the relevant records can be related to each other (see Figure 21). The model element labels are indicated on the system diagram (see Figure 11).

It was originally intended that a relational join would be used to access the data in the various tables, as this is the type of access generally supported in Arc/Info. Although this approach is adequate when there is a one-to-one correspondence between records, there are several situations when there are multiple records associated with a particular model element, requiring a one-to-many relationship i.e return-flow cells have multiple irrigated units on them. Arc/Info can only display one related record at a time, which is not adequate to emulate the DISA menu system (see Figure 15). This approach also required that both the file name and the name of the related item be specified each time data was called up.

A data dictionary was developed in order to overcome this limitation. A file, called the cross-reference dictionary was set up to store the names of data files and menus associated with each type of model element. This was then accessed at run-time to select the associated data files after a specific model element had been chosen (see Figure 22, Figure 23). The model element key was still used to make sure that only the selected model element records were displayed from the data file chosen. The effect of this

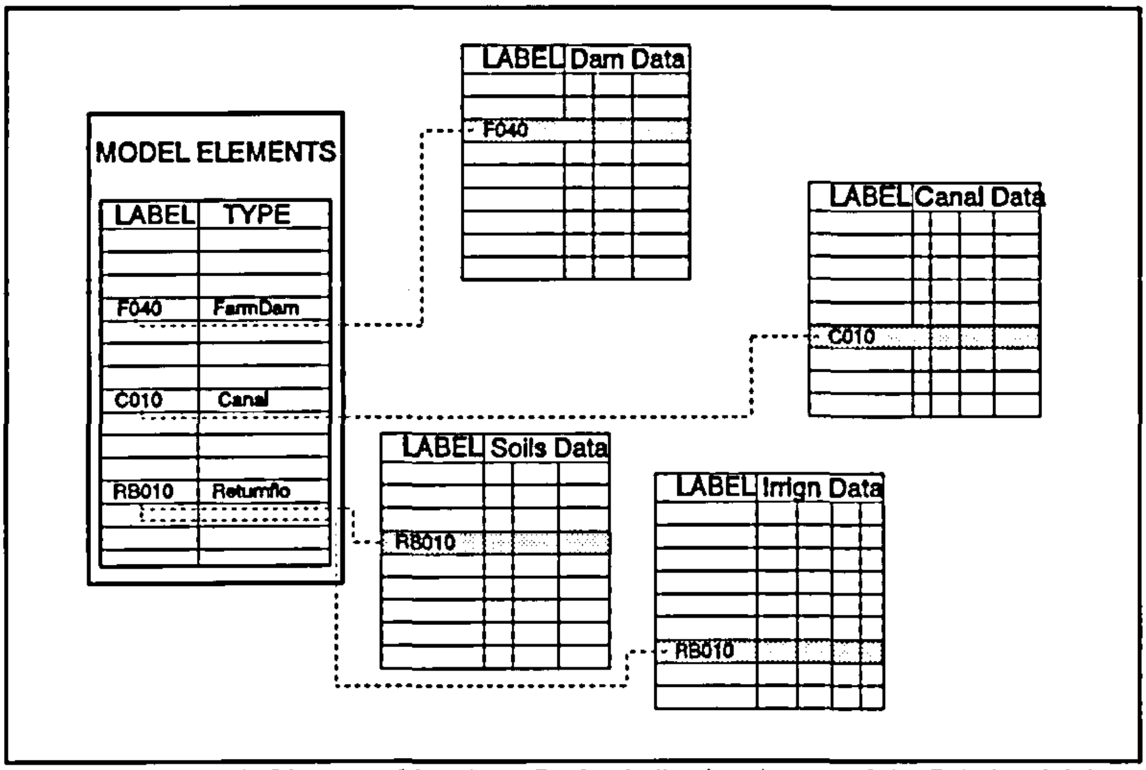


Figure 22 Schematic Diagram of Database Design indicating the use of the Relational Join to access Model Element Data

is to cause Arc/Info to behave as if the underlying data model had been extended, as the system now incorporates the 'intelligence' of the associations between model elements and data files i.e. as if each model element type had been defined in an object-oriented DBMS.

The discussion of lineage in section 4.1.1 emphasises the importance of giving users a means by which they can judge the utility of the information in the database. The present version of Arc/Info does not provide error propagation functions, and it was not possible to determine the level of error in the input data *in any case (see section 4.1.2). It was therefore decided that lineage information should be incorporated* in the GIS database. This data was stored in additional data dictionary files, and accessed using the system described in the previous paragraph.

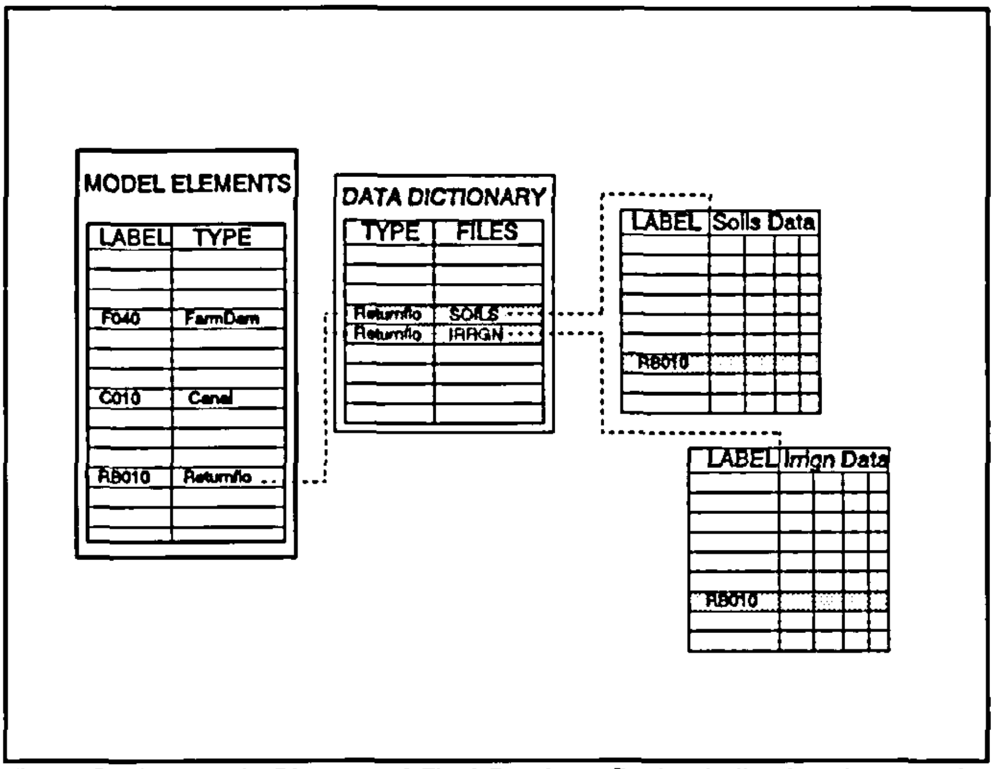


Figure 23 Schematic Diagram of Final Database Design indicating the use of a Data Dictionary to access Model Element Data.

4.3.3 The User Interface

The design of the GIS user interface based on the model element structure not only provided a consistent feel to the interface, it also made it possible to extend the database without having to make major modifications to the underlying code. This is one of the advantages cited in object-oriented programming (Meyer, 1991). The same cannot be said of the C-language file formatting programs, however.

The user interface draws together two distinct sets of operations: batch processes for defining the database structure and interchanging data, and interactive listing and editing of the database contents. The use of model elements has also enabled the development of an interface where database queries may be made either by drawing on a list of the model elements, or by 'picking' an element from the conceptual diagram. The data associated with the model element becomes available via a series of submenus. Metadata is also available in this menu system.

4.4 Conclusion

A loose coupling was considered most appropriate for the linkage of DISA and Arc/Info, because of the need to use separate hardware platforms. The user interface for Arc/Info was designed to imitate the DISA menu interface, in order to provide a familiar feel to the system. A data dictionary was set up to support the user interface, as Arc/Info's relational join was unable to display multiple related records once a model element had been chosen. Lineage information was also included in the database.

CHAPTER 5 - VERIFICATION AND EVALUATION

This chapter discusses the internal and external verification of the DISA - Arc/Info interface, and then critically evaluates the coupling.

5.1 Verification

The verification of the interface consisted of two main stages. First, a check on the internal consistency of the systems communications was carried out, to ensure that the data exported from Arc/Info to DISA was of the correct format. Then the parameters derived for DISA using GIS analysis were compared to the figures in the original configuration file for DISA, and the results for model runs using the two different configurations were analysed.

5.1.1 *Internal Verification*

The internal verification consisted of reading in a configurator file from DISA, immediately exporting it using the interface system developed, and then comparing the two files. As the Arc/Info output is an initialised file an automated comparison was not practical, and the relevant sections were compared visually. A sample of ten percent of the elements was selected for the comparison. The file was then successfully read into the DISA model, demonstrating that the existing model menu system could be used to complete the model configuration.

5.1.2 External Verification

The data derived from the GIS database was compared to the data from the DISA configurator file. The percentage difference between each pair of GIS- and original DISA figures was expressed using the formula:

$$\textit{Percentage Difference} = ((N_{Arc/Info} - N_{DISA}) / N_{DISA}) \times 100$$

Figure 24, Figure 25 and Figure 26 show box-and-whisker plots of the percentage difference between the DISA configurator and Arc/Info values for return-flow cell area, river reach length and canal node length respectively. This data is presented in tabular form in Appendix F. The box-and-whisker plot is explained in Table 7.

The return-flow cell data (Figure 24) shows the most variation. Half of the data values show less than ten percent difference, and most show less than twenty percent difference. There are several outliers with a much greater value, one of approximately 160%, which lies beyond the right-hand margin of the plot. The middle two quartiles of the canal figures show a similar distribution.

Table 8 The Box-and-Whisker Plot

The Box-and-Whisker plot is a technique of exploratory data analysis. It presents a visual summary of the spread of data values, and is useful for detecting skewness in distributions. The central box indicates the distribution of the middle 50% of the data values, while the whiskers extend to show the extremes. Whisker length is limited to one and a half times the size of the box. Points which do not fall within this distance are shown individually (STSC, 1989).

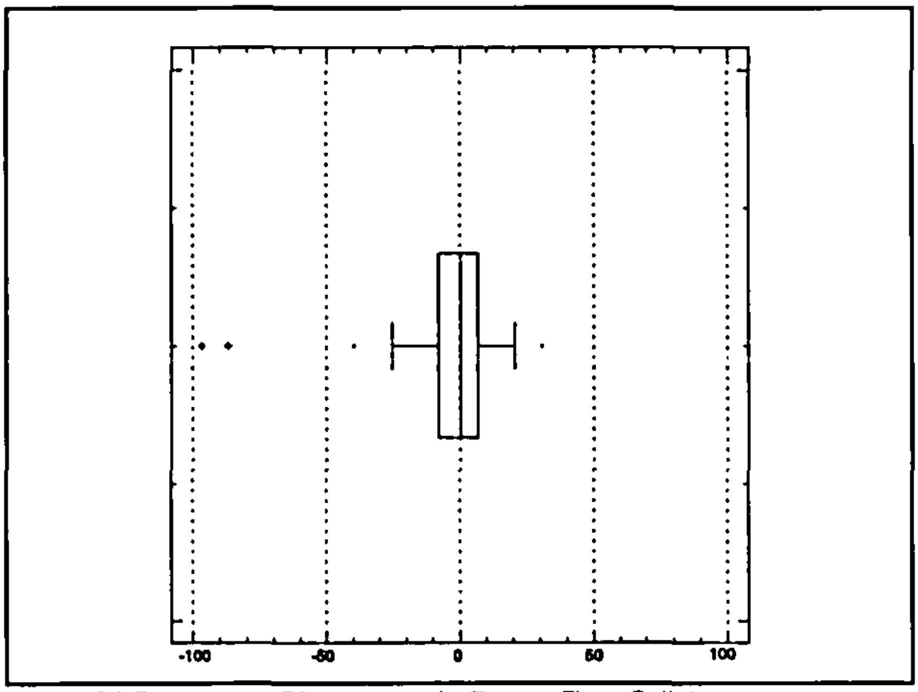


Figure 24 Percentage Discrepancy in Return-Flow Cell Areas.

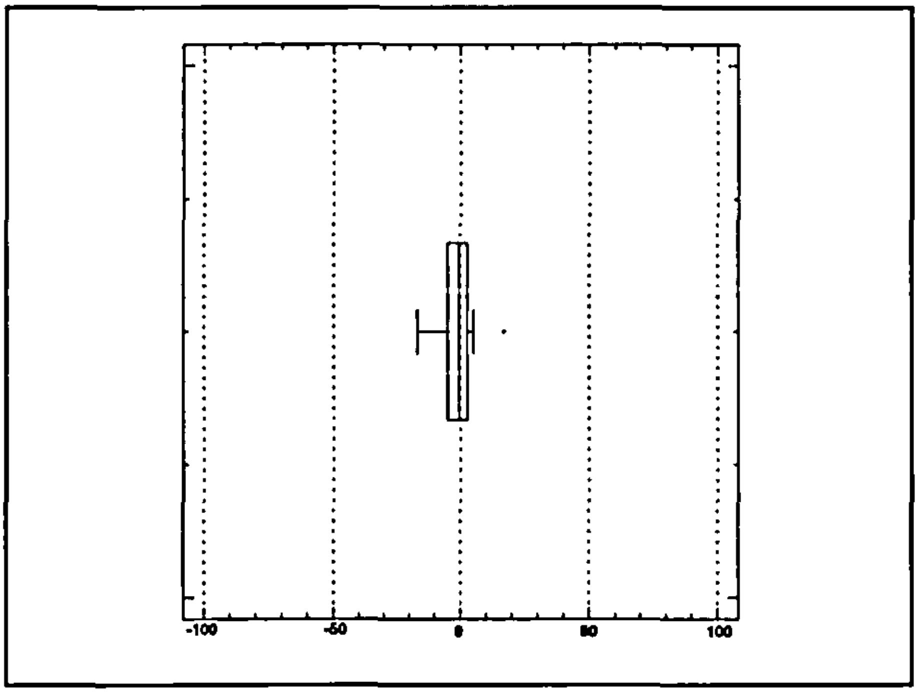


Figure 25 Percentage Discrepancy in River Reach Lengths.

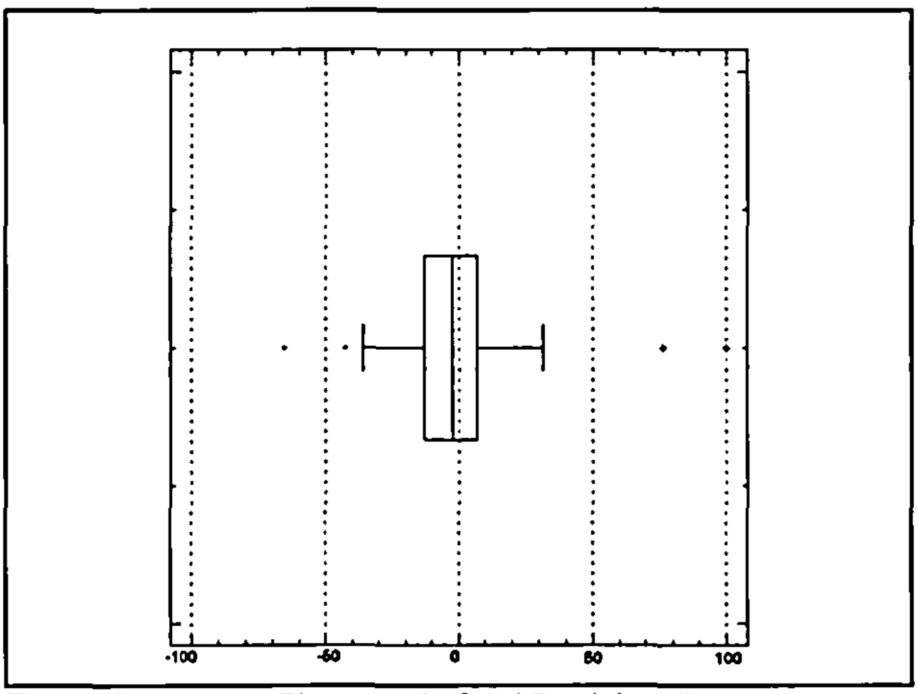


Figure 26 Percentage Disrepancy in Canal Reach Lengths.

(Figure 26) to the return-flow cell data, but the outliers are slightly less extreme. By contrast, the river reach differences are no greater than seventeen percent (Figure 25), and the box shows that the central 50% of figures show a variation of approximately five percent. The skewness in the route figure distributions is not statistically significant, because there were only ten river reaches.

There are a number of possible explanations for these discrepancies:

1) digitising error;
2) undocumented departures from the configuration expressed in the map overlays and
3) error in the original data capture for the DISA configuration.

The areas where major discrepancies appeared were inspected for digitising error, but this could not account for the size of the deviations. Two cases were detected where return-flow cells that were demarcated separately on the overlays had been merged for the configuration of DISA. These cells were then combined using the GIS, and the two sets of figures were compared again to yield Figure 24. It seems likely that the deviations reported here were a result of the combination of factors 2) and 3) above. It was not possible to investigate this issue further given the time constraints of the project and the amount of time that has elapsed since the original configuration of DISA.

The values derived using Arc/Info for the DISA configurator file were manually entered into the existing file for the 1985/86 verification of DISA. The model was then run using the original file and the modified file, and the results were compared. As can be seen from Figure 27 and Figure 28, the results for both flow and salinity concentrations on the Breede river at the Zandvliet weir could not be distinguished.

Return-flow cells with the biggest change in area between the DISA and Arc/Info values were then selected (see Table 9 below), and the results from the different cells for the same configuration were plotted together (see Figure 29, Figure 30, Figure 31 and Figure 32). These show that the water level and salinity time series were similar, except for cell RB040A.

Table 9 Selected Return-Flow Cell Areas

Record	Label (1985)	Area - DISA (ha^2)	Label (1994)	Area - Arc/Info (ha^2)	Percentage Difference
1	RB040A	195.00	XB040A	255.34	31
2	RB070A	32.00	XB070A	83.36	161
3	RT040T	5164.00	XT040T	3856.30	-25

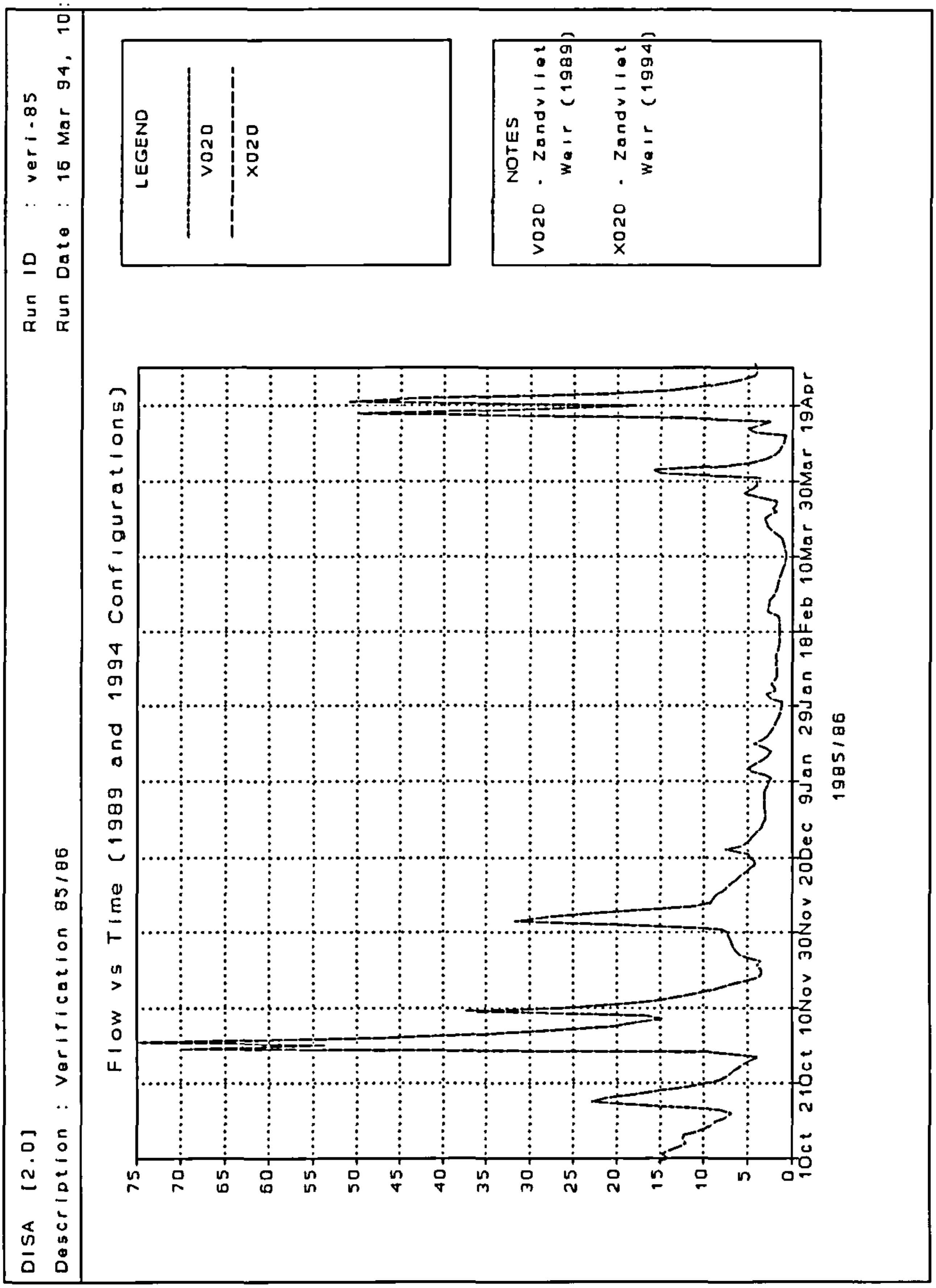


Figure 27 Time-series of Flow at the Zandvliet Weir.

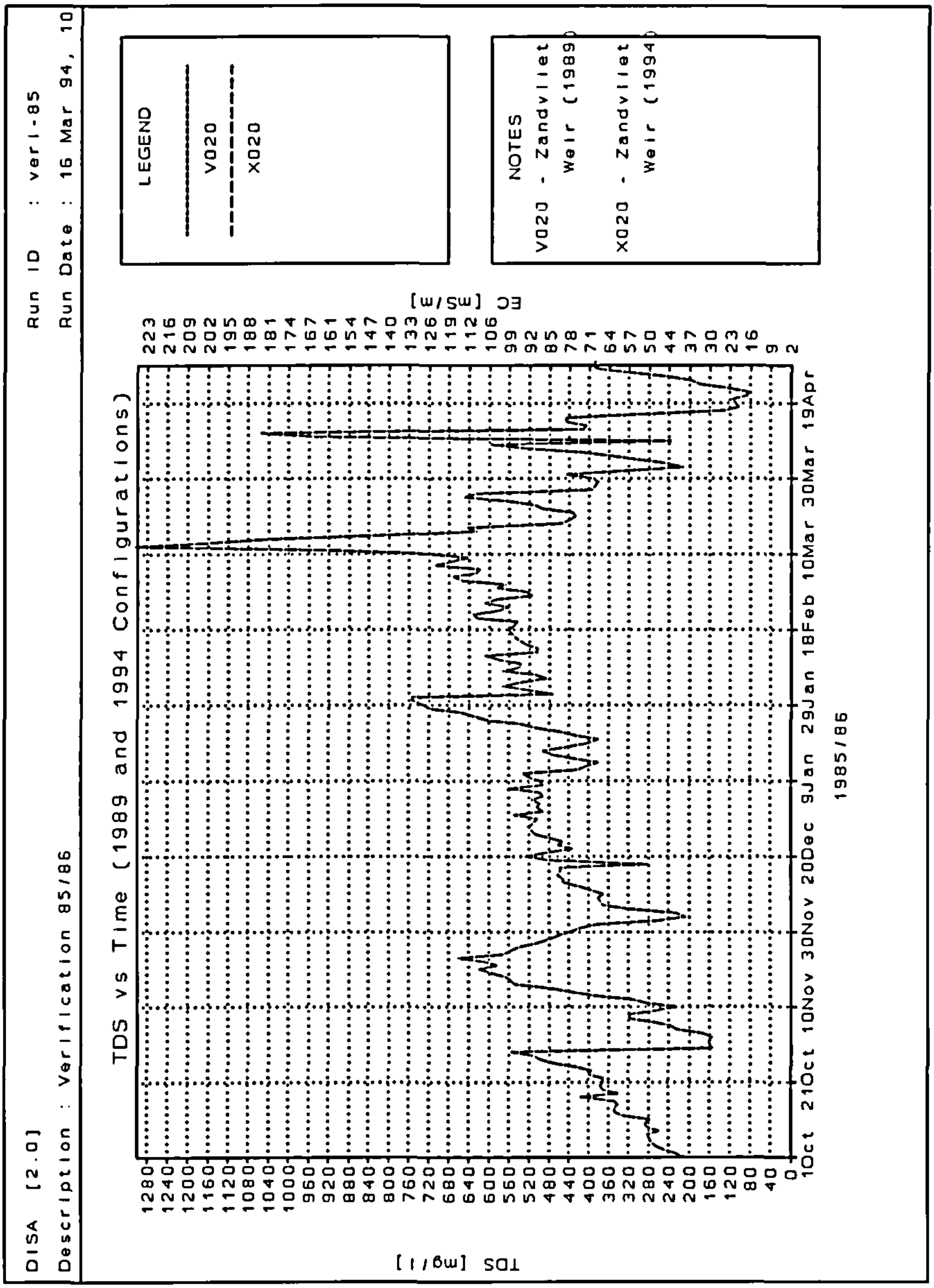


Figure 28 Time-series of Salt Concentration at the Zandvliet Weir

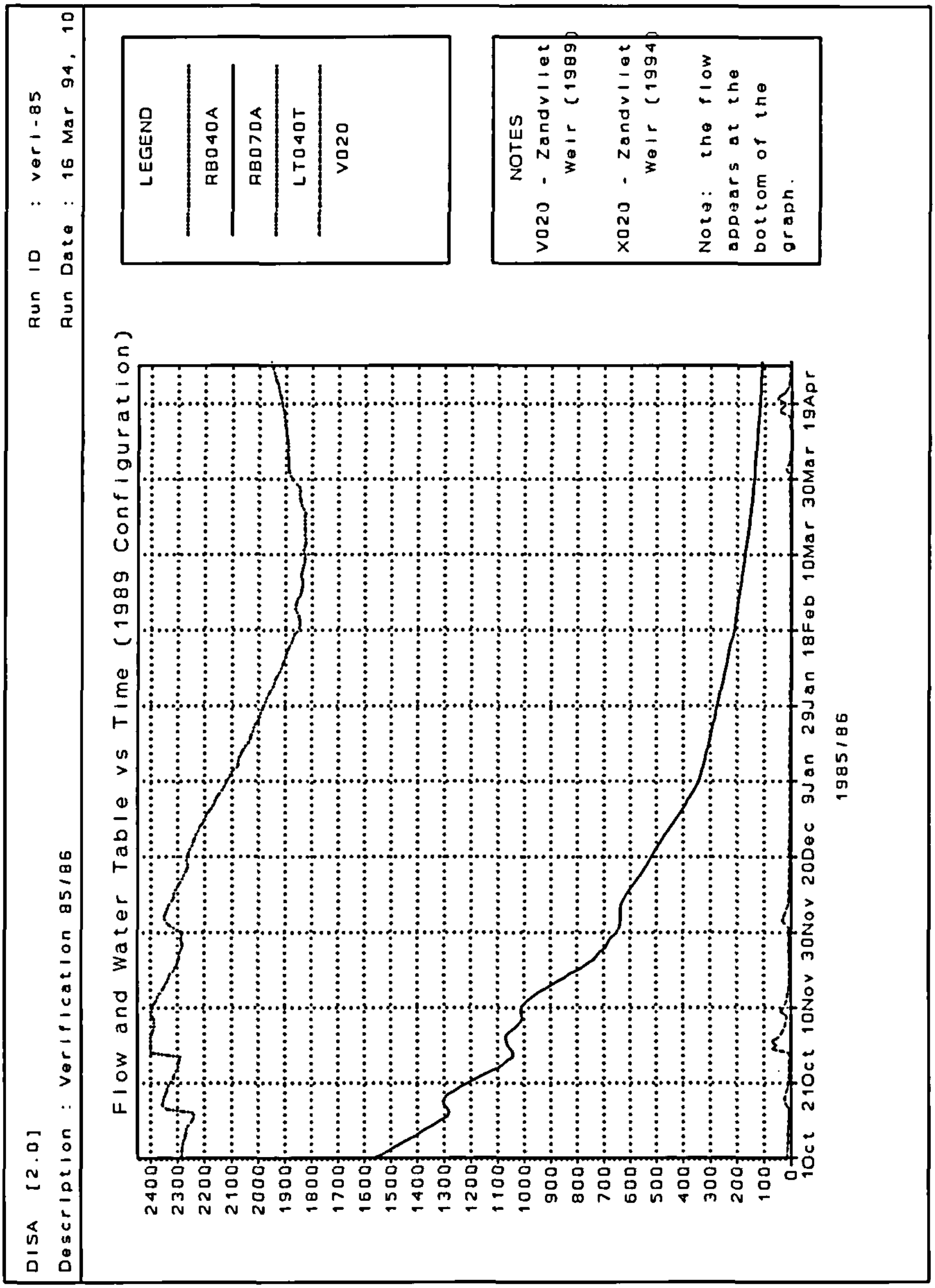


Figure 29 Flow and Water Level at the Key Points using the 1985 configuration file.

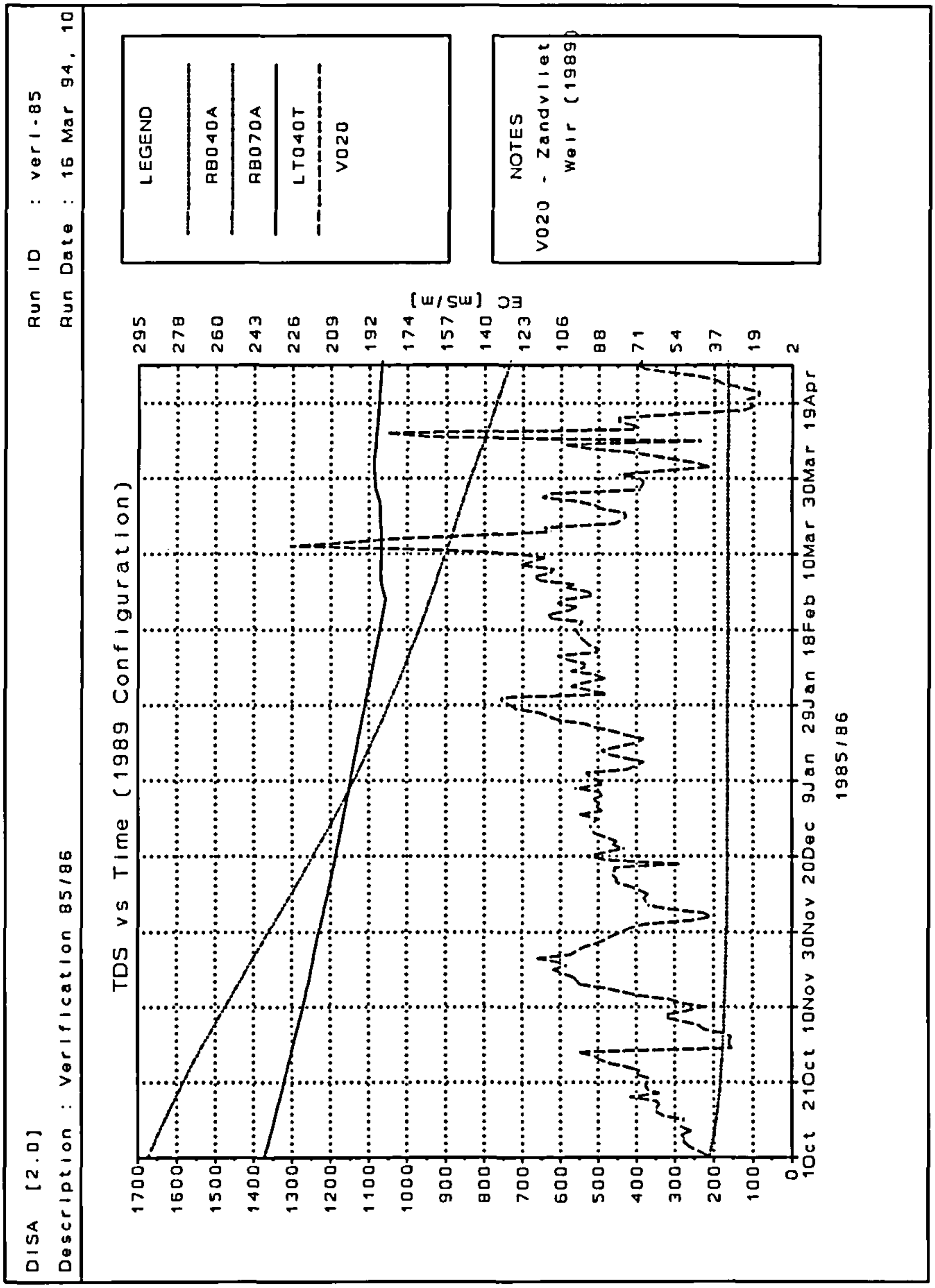


Figure 30 Salt Concentration at the Key Points using the 1985 configuration file.

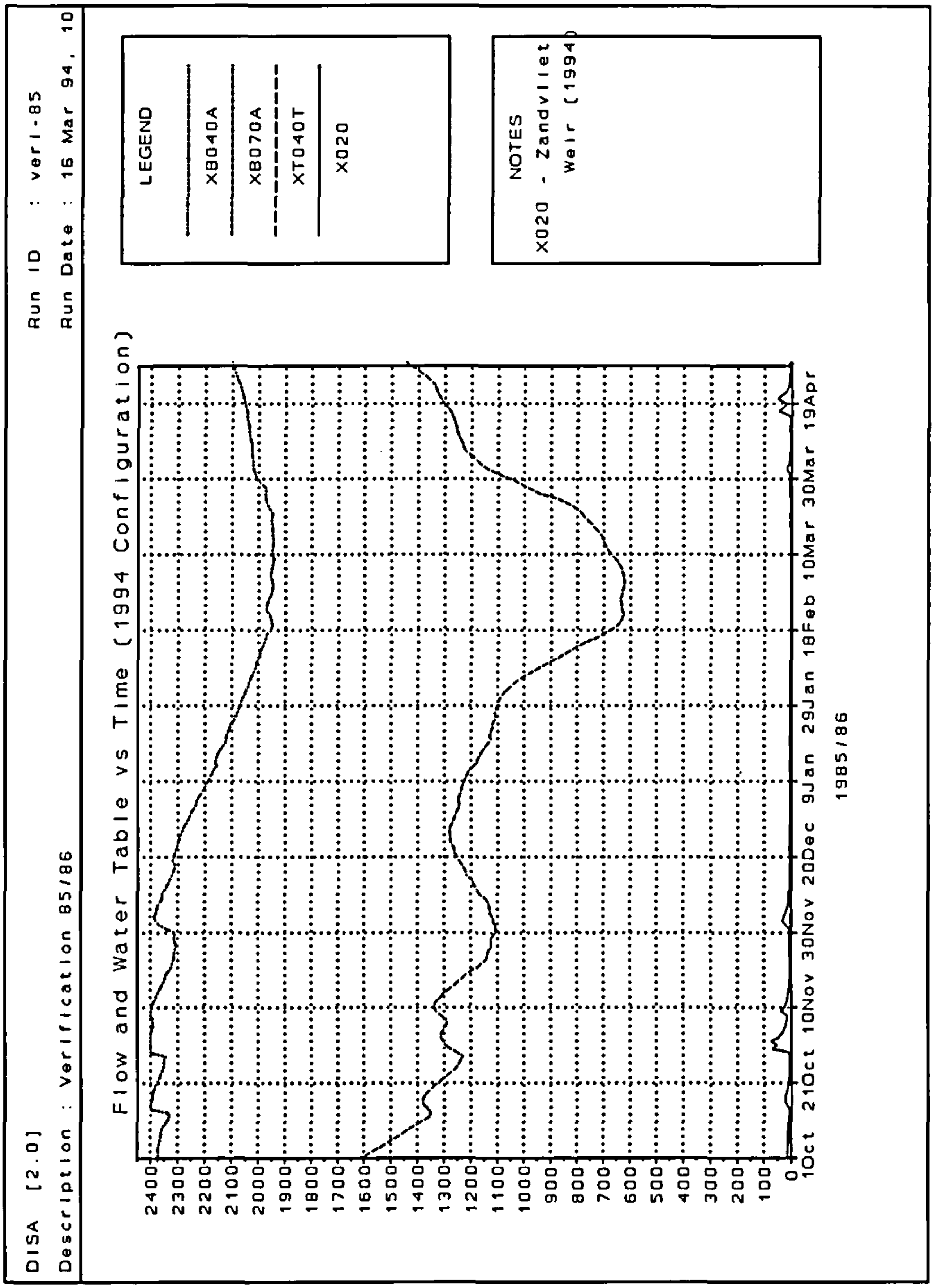


Figure 31 Flow and Water Level at the Key Points using the 1994 configuration file.

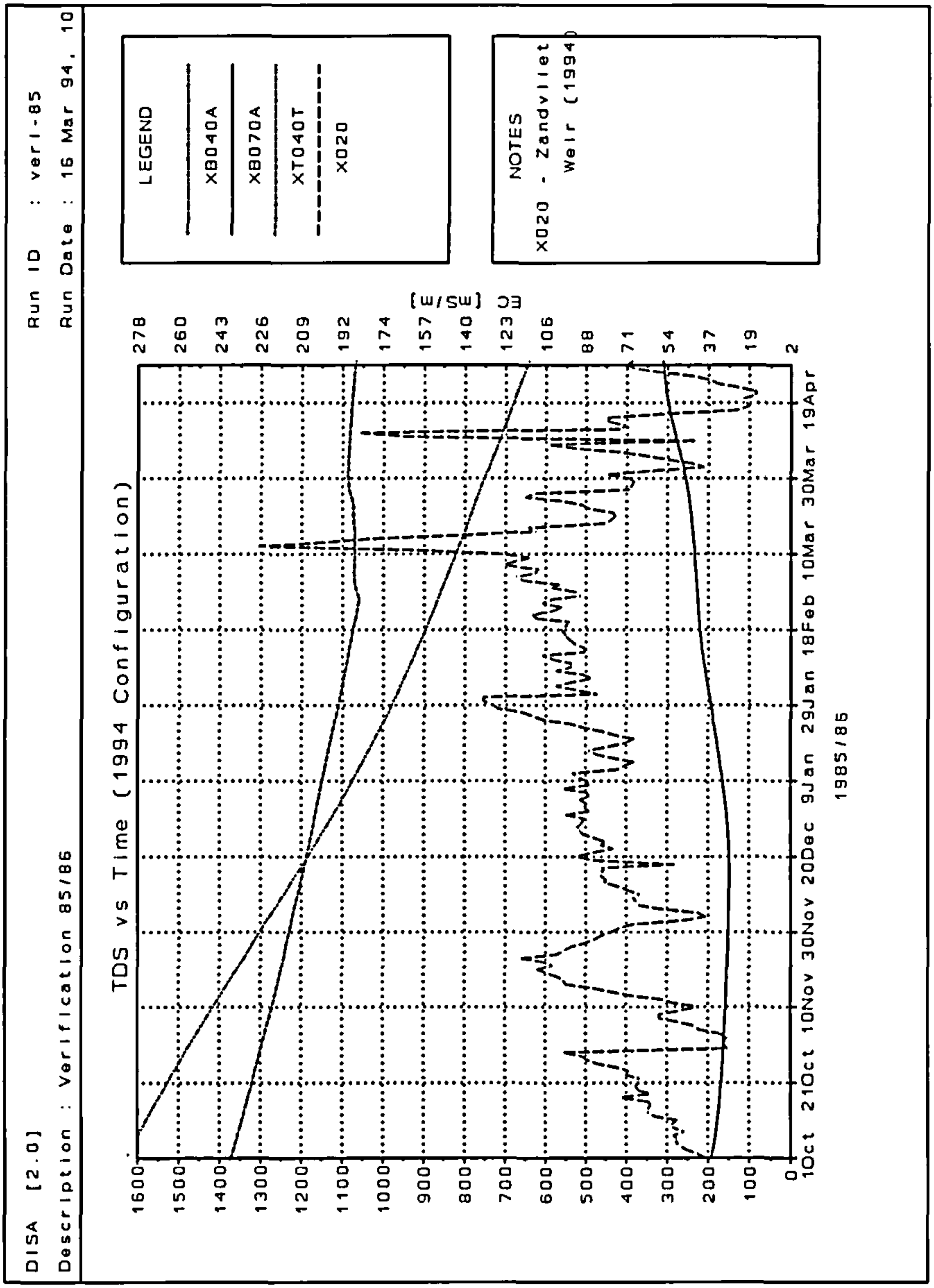


Figure 32 Salt Concentration at the Key Points using the 1994 configuration file.

5.2 Evaluation

The evaluation of the interfacing is further elaborated with a discussion of the different aspects identified from the literature review: the establishment of the spatial database, the systems linkage and the user interface. The overall effectiveness of the coupled system is dicussed in the conclusion.

5.2.1 The Spatial Database

The literature on the use of GIS as a spatial database for hydrological modelling has often referred to the advantage of increased speed of data capture under GIS versus manual data capture. This project has also demonstrated this feature, especially when taking into consideration the fact that GIS-like custom software was developed for the original configuration of DISA. The original soil groups delineated for the modelling effort were automatically conflated using the GIS in order to yield the soil groups used for the modelling effort (see Figure 33 and Figure 34). However, GIS also demands that the data to be captured be clearly and consistently defined, eg. all polygons must close and neighbouring polygons must have different types. The effective use of GIS requires a clear understanding of the particular needs of this type of approach. The availability of overlays from the original DISA project substantially assisted the data capture for GIS.

A spatial database offers another important advantage over conventional means of data capture for hydrological modelling: it preserves the input data from which the model configuration parameters are derived. It can be used again for the same model, or others if the input data captured was of primary data. This is an important point to be considered in the development of a GIS database for modelling purposes. The storage of the input data as well as derived information in a GIS also helps keep a record of the processes used to obtain the configuration, and GIS provides a means to check the interpretation of the data if the derivation is ever questioned . This contrasts with the situation encountered in this project, where it is not possible to confirm that modification of the delineated data may account for the discrepancies with the original configuration.

The graphic display capabilities of GIS can also give an overall view of the spatial database that is not achievable when working with a number of map sheets. This in turn assists in the detection of errors which might go unnoticed until a much later stage, improving the efficiency of the modelling effort. Figure 35 shows the information used in the DISA configuration (together with the topographic information) to derive the physiographic regions. This draws together the aquifer system, the regional geology and the designated points of interest in the catchment.

5.2.2 The Systems Linkage

The systems linkage developed in this project exports the Arc/Info database and generates a configurator file for DISA in less than three minutes. The ethernet transfer of the file from the workstation to the PC takes a matter of seconds, from which the non-spatial configuration of the model can be completed. This is an insignificant delay in comparison to the time required to prepare the DISA configuration.

The Arc/Info system does not incorporate rigorous checks on the completeness and integrity of the database. What has been done is that the number of unique elements in each database file is determined during the export procedure and compared to the number given in the header file. The export process is aborted if these figures do not agree. All of these files are required to be sorted in the same order - this is checked by the C-program that integrates the different ASCII files exported by Arc/Info. The process is aborted if a mismatch between the element labels in the different files is encountered. It should be noted that while the interface provides default values for sections of the configuration file not being parameterised from Arc/Info, it does not assign default values where entries have not been made in Arc/Info. The database has been designed so that the database definitions correspond exactly to the format requirements of DISA.

The systems linkage provides an efficient data exchange service which incorporates error checking. The default data values for any particular application have to be set in the DISA models configuration file, and entered manually into Arc/Info.

5.2.3 The User Interface

The user interface automates the definition of database files, as well as transferring data to and from the DISA file format. It also provides a view on the database similar to that given by the PC menu system, to assist the user to assess the completeness of the database. The edit mode of the interface using form menus allows only one record to be accessed at a time, which detracts from the intended consistency of the user interface. Only metadata for the DISA attributes is presented in the user interface. Only the metadata for the data exported to DISA is incorporated in the user interface.

The user interface drives procedures to automate fundamental tasks, as well as providing a familiar interface to the data for model users.

5.3 Conclusion

The use of the DISA model involves much more time and effort in the configuration of the system than in the actual running of the model. This project took loose linkage of DISA and Arc/Info as its guiding principle, and concentrated on the development of a spatial database to support the configuration of DISA because the configuration of the model is the single most significant task in the application of the model.

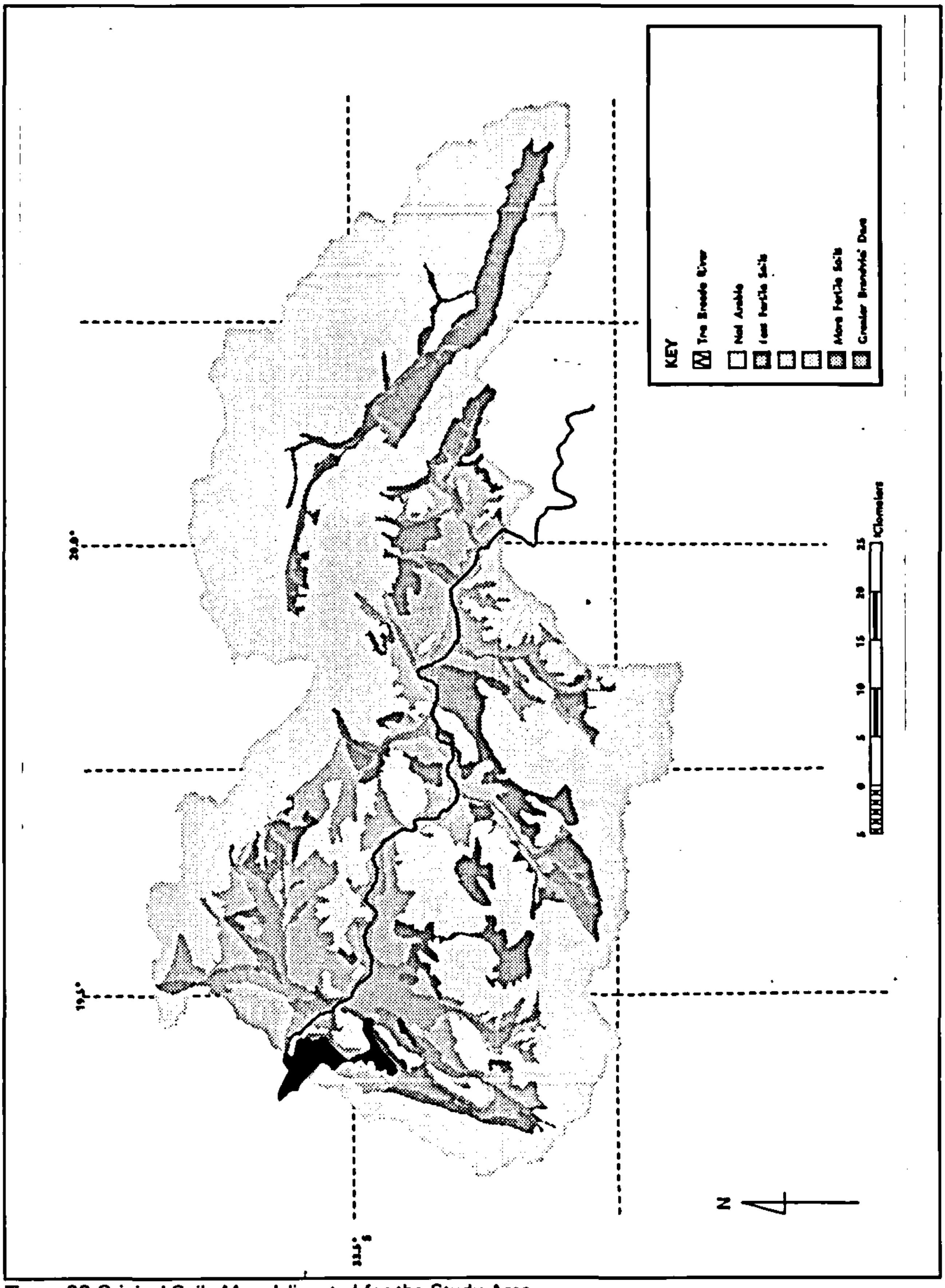


Figure 33 Original Soils Map delineated for the Study Area

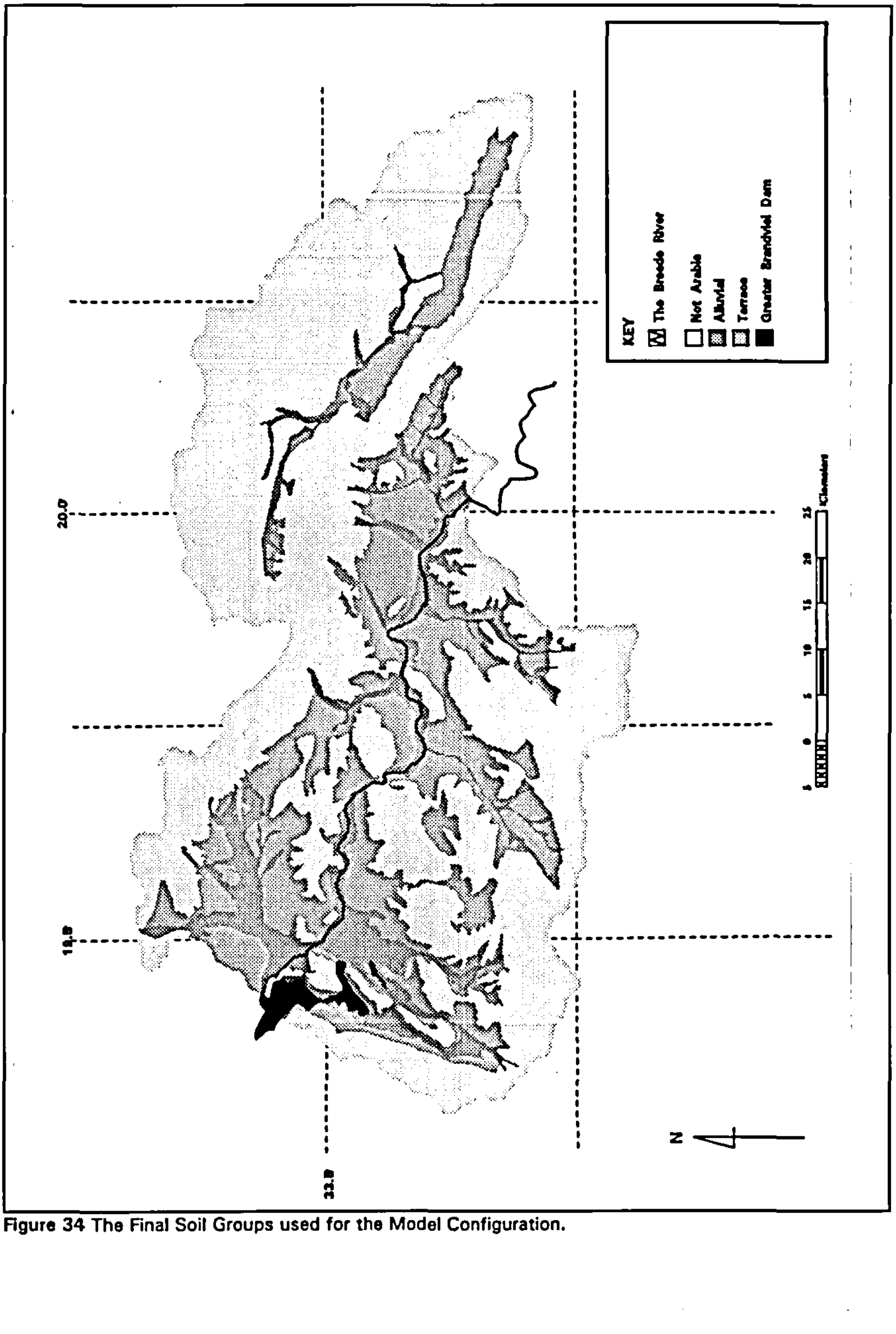


Figure 34 The Final Soil Groups used for the Model Configuration.

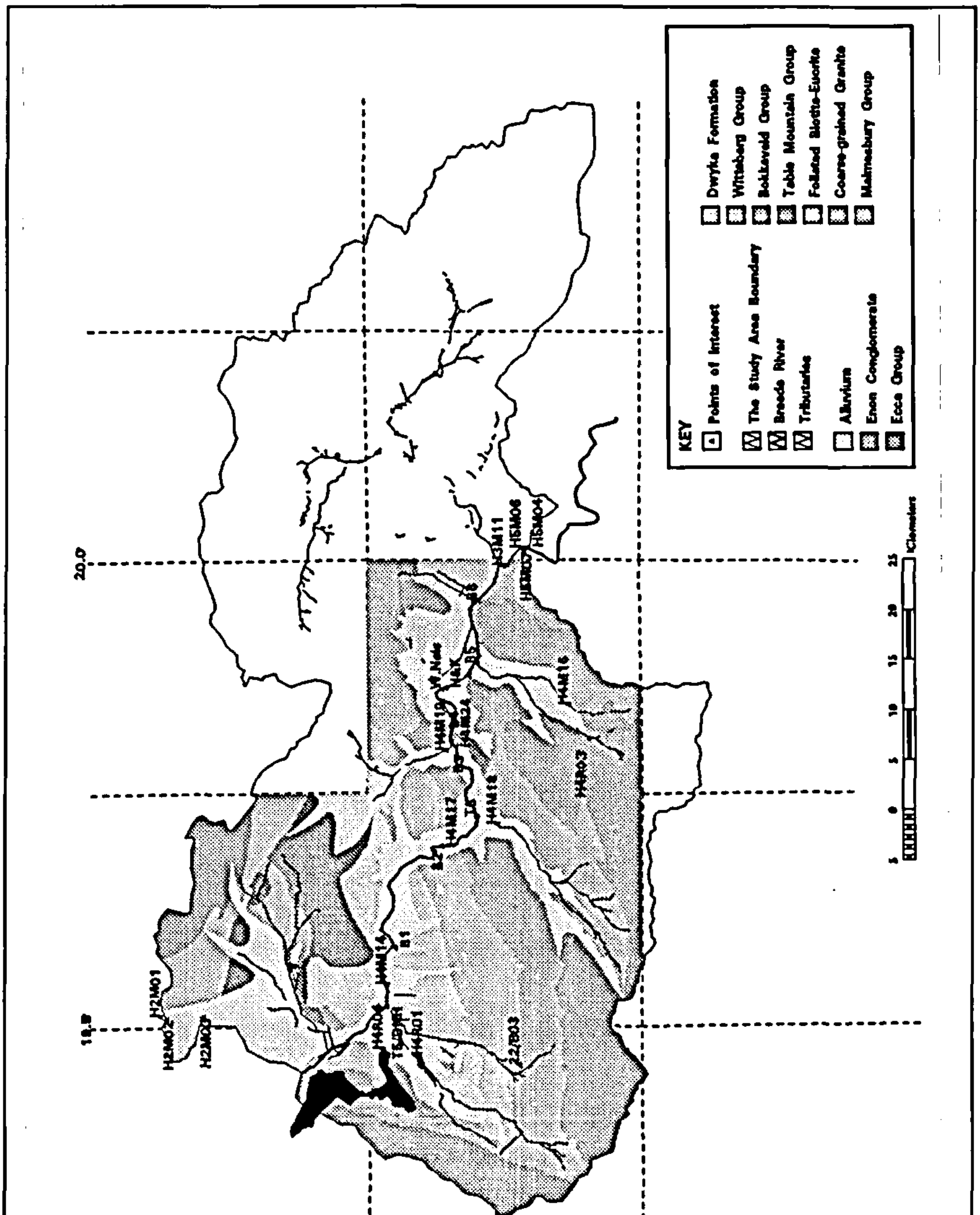


Figure 35 General Map showing the Information used to Assist in the Delineation of Physiographic Regions.

The use of GIS reduces the time required for data capture because GIS offers established spatial data capture capabilities, as well as tools for analysis of the information by overlay. The ability to integrate a number of map sheets of various themes into the database gives an overview of the information captured for the configuration which cannot be achieved with conventional map sheets, and assists in the detection of errors in the database.

The interface developed provides a verified and efficient means of transferring the data from the Arc/Info database into a format which can be read and completed using the DISA model. Although differences were observed between the original DISA data and the GIS-derived information, it was not possible to determine the reasons for all of thse discrepancies. It seems possible that undocumented deviations from the map overlays drafted for the project may account for a substantial number of these problems.

The development of a spatial database with GIS means that the data capture effort for one project can be reused for other purposes, as the primary data can be reworked. This also provides a partial record of how the model was configured, which, when combined with the use of metadata tools such as a data dictionary, can assist different users with different needs to assess the database for their own purposes.

CHAPTER 6 - CONCLUSION

6.1 General Discussion

The coupling of GIS and hydrological models has a considerable history, which has its origins in research into spatial data management systems for modelling in the 1970s (Shea *et al*, 1993). Remote sensing and digital elevation modelling were identified as important technologies in spatial database development in this period (DeVantier and Feldman, 1993). The maturing of GIS in the 1980s has made it more generally accessible, and several recent reviews have noted the great interest that has been shown in hydrological modelling with GIS (Zhang, Haan and Nofziger, 1990; DeVantier and Feldman, 1993; Moore *et al*, 1993).

The discussion on linking GIS and models has often considered this to be a matter of establishing communications between the different data management subsystems of the two software systems. Although it is possible to carry out modelling within GIS, the developed nature of hydrological modelling has meant that it is much more common for GIS to serve as a pre- and post-processor for existing hydrological models. Weaknesses have been identified in such a use of GIS, particularly in its lack of support for time-varying data structures, lineage and product quality information. These are areas of active research in GIS.

The approach taken in this project has been to focus on the development of the Arc/Info GIS as a spatial database to support the configuration of the DISA model. The configuration of a hydrological model is a technically demanding and tedious process, which takes much more effort than the actual running of model simulations.

Although only a limited set of spatial data was available for this project, the original sources for the DISA model development included the use of aerial photography and satellite imagery. Arc/Info offers the ability to integrate this information together with conventional map data, to form an enduring spatial database. Such a database can be used repeatedly, and the original input data can be preserved, enabling other users to check the derivation of configuration information if errors are suspected or to analyse the data for use with other models. The visual presentation capabilities of GIS also assist in the detection of configuration errors.

A loose linkage was established between DISA and the GIS, in order to transfer a partially completed configuration file after GIS capabilities had been applied to capture and analyze the spatial data for the model. Two important aspects of this linkage are the user interface established for the GIS database and the metadatabase of sources drawn on in the development of the spatial database. The application procedure for the coupled system requires that the DISA menus be used to complete the configuration of the model. The GIS user interface for the database was therefore developed to resemble the model menu system in order to establish a similar 'look and feel' between the GIS and DISA, so that when the operators move from the one system to the other there is a minimal learning curve.

The coupling established between the Arc/Info GIS and the DISA model will considerably assist in the mechanical aspects of the DISA configuration. It does not safeguard against the use of inappropriate input data or GIS analyses. A system that could provide these features would have to involve a tight coupling of the two systems, and as presently available GIS does not incorporate product quality checking it would involve developing the GIS component from scratch.

The approach that has been taken in this project has been to make the best use of the available resources in order to address this problem, without claiming to have completely overcome it. An effective use of coupled GIS and hydrological models will continue to require a firm grasp of both GIS and hydrological modelling, until it becomes possible to tightly couple these systems and propagate estimates of uncertainty from the GIS to the modelling results with reasonable ease. This will require the development of communications standards between software systems, and formal descriptions of both GIS product quality and model characteristics. These are areas that urgently require research, but it is unlikely that practical progress will be made for some time.

6.2 Research Needs

The interfacing of hydrological models and GIS can provide a valuable decision support tool. The research proposed here will be of direct benefit to developers of such tools. Wolff-Piggott (1994) discusses the fundamental research topics associated with this field.

- The use of GIS offers much greater advantages when data capture from maps can be reduced or eliminated. A review of digital data sources available in Southern Africa, and their suitability for use in hydrological modelling is recommended. This would help reduce the large data capture costs associated with GIS.

- The identification of suitable standards for data storage in hydrological modelling is recommended, so that interfaces between models and GIS can be established more easily. The WDM format that has been established by the USGS is an example of one such standard for time series data.

- The representation of variation over time is limited in GIS. This means that the presentation of time series data from hydrological modelling within a GIS environment requires a great deal of programming expertise and effort. The Institute for Water Quality Studies has developed a sophisticated system (AQCES) to perform such a task for water quality time series. It is recommended that a system such as this should be adapted for use with models rather than independently developing GIS post-processors for models, wherever possible.

- It was not possible to develop a presentation of DISA output in GIS within the scope of this project. It is recommended that this issue be addressed by interfacing AQCES and DISA.

- The interface that has been developed in the course of this project is suitable for operational use. It is recommended that the software system be applied in the configuration of DISA for a new catchment in order to demonstrate the benefit of GIS for hydrological modelling.

BIBLIOGRAPHY

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E. and Rasmussen, J. 1986. An Introduction to the European Hydrological System _ Systeme Hydrologique Europeen, "SHE", 1: history and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*. 87, 9: 45-59.

Anderson, K. L. 1991. Spatial changes in the hydrology of portions of the twin cities, Minnesota: a simulation study. *Physical Geography*. 12, 2: 147-166.

Armstrong, M. P. and Densham, P. J. 1990. Database organizational strategies for spatial decision support systems. *International Journal of Geographical Information Systems*. 4, 1: 3-20.

Band, L. E. 1989. A terrain-based watershed information system. *Hydrological Processes*. 3, 2: 151-162.

Beasley, D. B., Huggins, L. F. and Monke, E. J. 1980. ANSWERS: a model for watershed planning. *Transactions of the American Society of Agricultural Engineers*. 123, 4: 938-944.

Beuster, J., Görgens, A. H. M. and Greyling A. J. 1990a. *Development of a Daily Hydrosalinity Model (DISA)*. Pretoria. Department of Water Affairs and Forestry. Directorate of Project Planning, Report No. H 000/00/0790.

Beuster, J., Görgens, A. H. M. and Greyling A. J. 1990b. *DISA User's Guide*. Pretoria. Department of Water Affairs and Forestry. Directorate of Project Planning, Report No. H 000/00/0690.

Beven, K. 1989. Changing ideas in hydrology - the case of physically-based models. *Journal of Hydrology* 105, 1/2:157-172.

Bhaskar, N. R., James, W. P. and Devulapalli, R. S. 1992. Hydrologic parameter estimation using geographic information system. *Journal of Water Resources Planning and Management*. 118, 5: 492-512.

Bitters B., Restrepo, P. J. and Jourdan, M. R. 1991. Using geographic information systems to predict effects of flooding on the Han river, South Korea. In *Proceedings of the ACSM-ASPRS Annual Convention*.

Bork, H. -R. and Rohdenburg, H. 1986. Transferable parameterization methods for distributed hydrological and agroecological catchment models. *Catena*. 13: 99-117.

Burrough, P. A. 1986. *Principles of geographical information systems for land resources assessment*. Clarendon. Oxford.

Burrough, P. A. 1989. Matching spatial databases and quantitative models in land resource assessment. *Soil Use and Management*. 5, 1: 3-8.

Burrough, P. A. 1992a. Are GIS data structures too simple minded? *Computers and Geosciences*. 18, 4: 395-400.

Burrough, P. A. 1992b. Development of intelligent geographical information systems. *International Journal of Geographical Information Systems*. 6, 1: 1-11.

Chou, H. C. 1991. Design of a language for spatial analysis. *ACSM-ASPRS Technical Papers*. 2: 56-64. Baltimore.

Chou, H. C. and Ding, Y. 1992. Methodology of integrating spatial analysis/modelling and GIS. *Proceedings of the Fifth International Symposium on Handling Spatial Data*. 2: 514-523. IGU Commission on GIS.

Cline, T. J., Molinas, A. and Julien, P. Y. 1989. An Auto-Cad-based watershed information system for the hydrologic model HEC-1. *Water Resources Bulletin*. 25, 3: 641-652.

Couclelis, H. 1992. Geographic knowledge production through GIS: towards a model for quality monitoring. In Couclelis, H. and Beard, K. and Mackaness, W. *Two Perspectives on Data Quality*. Technical Report 92-12. National Center for Geographic Information and Analysis. Santa Barbara.

Danish Hydraulic Institute. 1993. *DHI Software News*, 2. Danish Hydraulic Institute.

Date, C. J. 1985. *An introduction to database systems volume II*, Addison-Wesley Systems Programming Series. Addison-Wesley. Reading.

Davis, D. W. 1978. Comprehensive flood plain studies using spatial data management techniques. *Water Resources Bulletin*. 14, 3: 587-604.

Davis, F. W. and Simonett, D. S. 1991. GIS and remote sensing. In *Geographic Information Systems Volume 1: Principles*. eds. D. J. Maguire, M. F. Goodchild and D. W. Rhind, pp. 191-213. Longman Scientific and Technical. London.

Davis, F. W., Quattrochi, D. A., Ridd, M. K., Lam, N. S-N., Walsh, S. J., Michaelsen, J. C., Franklin, J., Stow, D. A., Johannsen, C. J. and Johnston, C. A. 1991. Environmental analysis using integrated GIS and remote sensed data: some research needs and priorities. *Photogrammetric Engineering and Remote Sensing*. 57, 6: 689-697.

Densham, P. J. 1991. Spatial decision support systems. In *Geographic Information Systems Volume 1: Principles*. eds. D. J. Maguire, M. F. Goodchild and D. W. Rhind, pp. 403-412. Longman Scientific and Technical. London.

De Roo, A. P. J., Hazelhoff, L. and Burrough, P. A. 1989. Soil erosion modelling using ANSWERS and geographical information systems. *Earth Surface Processes and Landforms*. 14, 6-7: 517-532.

DeVantier, B. D. and Feldman, A. D. 1993. Review of GIS applications in hydrologic modeling. *Journal of Water Resources Planning and Management*. 119, 2: 246-261.

Ding, Y. and Fotheringham, A. S. 1992. The integration of spatial analysis and GIS. *Computers, Environment and Urban Systems*. 16, 1: 3-19.

Djokic, D. and Maidment, D. R. 1991. Terrain analysis for urban stormwater modelling. *Hydrological Processes*. 5, 1: 115-124.

Donker, N. H. W. 1992. Automatic extraction of catchment hydrologic properties from digital elevation data. *ITC Journal*. 3: 257-265.

Drayton, R.S., Wilde, B. M. and Harris, J. H. K. 1992. Geographical information system approach to distributed modelling. *Hydrological Processes*. 6, 3: 361-368.

Eastman, J. R. 1992. *IDRISI user's guide*, version 4.0. Clark University. Worcester.

Eli, R. N. 1990. A cellular automation drainage processor for large scale spatially distributed hydrologic models. *EOS. Transactions, American Geophysical Union*. 71, 43: 1348. cited in DeVantier, B. D. and Feldman, A. D. 1993. Review of GIS applications in hydrologic modeling. *Journal of Water Resources Planning and Management*. 119, 2: 246-261.

Engman, E. T. and Gurney, R. J., eds. 1991. *Remote sensing in hydrology*. Chapman and Hall. London.

Engman, T. 1993. Remote sensing: a developing technology for hydrology (preface). *Hydrological Processes*. 7, 2: 119-120.

ESRI. 1991a. *Arc/Info data model, concepts and key terms*. Arc/Info user guide, version 6.0. ESRI. Redlands.

ESRI. 1991b. *Arc command references, version 6.0*. ESRI. Redlands.

ESRI. 1991c. *Arc/Info user's guide: What's new in Arc/Info rev 6.0.* ESRI. Redlands.

Estes, J. E. 1992. Remote sensing and GIS integration: research needs, status and trends. *ITC Journal.* 1: 2-10.

Fedra, K. 1993. GIS and environmental modeling. In *Environmental modeling with GIS.* eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert, pp. 35-50. Oxford University Press. Oxford.

Frank, A. U. 1992. Spatial concepts, geometric data models and geometric data structures. *Computers and Geosciences.* 18, 4: 409-417.

Gandanoy-Bernasconi, W. and Palacios-Velez, O. 1990. Automatic cascade numbering of unit elements in distributed hydrological models. *Journal of Hydrology.* 112, 3/4: 375-393.

Goodchild, M. 1987. A spatial analytic perspective on GIS. *International Journal of Geographic Information Systems.* 1, 4: 327-334.

Goodchild, M. 1992a. *Closing report on the research initiative on error in spatial databases.* National Center for Geographic Information and Analysis. Santa Barbara.

Goodchild, M. 1992b. Geographical information science. *International Journal of Geographic Information Systems.* 6, 1: 31-45.

Goodchild, M. 1993. The state of GIS for environmental modelling. In *Environmental modeling with GIS.* eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert, pp. 94-104. Oxford University Press. Oxford.

Goodchild, M. Haining, R. Wise, S. and twelve others. 1992. Integrating GIS and spatial data analysis: problems and possibilities. *International Journal of Geographical Information Systems.* 6, 5: 407-423.

Grayson, R. B., Blöschl, G., Barling, R. D. and Moore, I. D. 1993. Process, scale and constraints to hydrological modelling in GIS. In *HydroGIS 93: Application of Geographic Information Systems in Hydrology and Water Resources.* IAHS Publ. no. 211. 83-92. IAHS.

Greyling, A., Beuster, H. and Görgens, A. H. M. 1993. *DISA salinity predictions for various planning scenarios.* Department of Water Affairs and Forestry. Report H 000/00/0890. Pretoria.

Guariso, G. and Werthner, H. 1989. *Environmental decision support systems.* Ellis Horwood Limited. Chichester.

Haslett, J., Wills, G. and Unwin, A. 1990. SPIDER - an interactive statistical tool for the analysis of spatially distributed data. *International Journal of Geographic Information Systems.* 4, 3: 285-296.

Hastings, D. A. and Moll, S. H. 1986. On geographic information systems: the new explores, a one act play. *Geotimes.* 31, 7: 8-10.

Jenson, S. K. and Domingue, J. O. 1988 Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing.* 54, 11: 1593-1600.

Jenson, K. 1991. Applications of hydrological information automatically extracted from digital elevation models. *Hydrological Processes.* 5, 1: 31-44.

Jett, C., Weeks, A. D. and Grayman, W. M. 1982. Geographic information systems in hydrological modelling. *Hydrologic Transport Modeling Symposium, December 1979.* pp. 127-137. The American Society of Agricultural Engineers. St. Joseph.

Jõao, E. M. and Walsh, S. J. 1992. GIS implications for hydrologic modeling: simulation of non-point pollution generated as a consequence of watershed development scenarios. *Computers, Environment and Urban Systems.* 16, 1: 43-64.

Karnieli, A. 1991. Stepwise overlay approach for utilizing a GIS with a soil moisture accounting model. *ITC Journal.* 1: 11-18.

Kienzle, S. W. and Lorentz, S. A. 1993. Production of a soil erodibility map for the Henley Dam catchment, Natal, using a GIS approach. In *Proceedings of the sixth South African national hydrological symposium.* Volume 2. pp. 565-574.

Kilborn, K., Rifai, H. S. and Bedient, P. B. 1990. The integration of ground water models and GIS. In *Technical Papers of the ASPRS/ACSM Annual Convention.*

Kouwen, N., Soulis, E. D., Pietroniro, A., Donald, J. and Harrington, R. A. 1993. Grouped response units for distributed hydrologic modeling. *Journal of Water Resources Planning and Management.* 119, 3: 289-305.

Lanter, D. 1990. *Lineage in GIS: the problem and a solution.* Technical Report 90-6. National Center for Geographic Information and Analysis. Santa Barbara.

Lanter, D. 1992. *Intelligent assistants for filling critical gaps in GIS.* Technical Report 92-4. National Center for Geographic Information and Analysis. Santa Barbara.

Laurini, R. and Thompson, D. 1992. *Fundamentals of spatial information systems.* Academic Press. London.

Loveland, T. R. and Johnson, G. E. 1983. The role of remotely sensed and other spatial data for predictive modeling: the Umatilla, Oregon, example. *Photogrammetric Engineering and Remote Sensing.* 49, 8: 1183-1192.

Lynch, S. D, Reynders, A. G. and Schulze, R. E. 1993. Determining input data for a national scale groundwater vulnerability map of Southern Africa. In *Proceedings of the Sixth South African National Hydrological Symposium.* 1, 747-758.

Maidment, D. R. 1993. GIS and hydrologic modeling. In *Environmental modeling with GIS.* eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert, pp. 147-167. Oxford University Press. Oxford.

Mallory, S. J. L. 1993. Hydrological modelling of farm dams. *Sixth South African Hydrological Symposium.* 2: 243-250. Pietermaritzburg.

McKim, H. L., Cassell, E. A. and LaPotin, P. J. 1993. Water resource modelling using remote sensing and object-oriented simulation. *Hydrological Processes.* 7, 2: 153-165.

Meyer, B. 1991. *Object-oriented software construction.* Prentice-Hall International Series in Computer Science. Prentice-Hall. 1988.

Meyer, S. P., Salem, T. H. and Labadie, J. W. 1993. Geographic information systems in urban storm-water management. *Journal of Water Resources Planning and Management.* 119, 2: 206-228.

Moore, I. 1991. Editorial: digital terrain modelling in hydrology. *Hydrological Processes.* 5, 1: 1.

Moore, I. D., O'Loughlin, E. M. and Burch, G. J. 1988. A contour-based topographic model for hydrological and ecological application. *Earth Surface Processes and Landforms.* 13, 4: 305-320.

Moore, I. D., Grayson, R. B. and Ladson, A. R. 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes.* 5, 1: 3-30.

Moore, I. D., Turner, A. K., Wilson, J. P., Jenson, S. K. and Band, L. E. 1993. GIS and land-surface--subsurface process modeling. In *Environmental modeling with GIS.* eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert, pp. 196-230. Oxford University Press. Oxford.

Morehouse, S. 1992. The Arc/Info geographic information system. *Computers and Geosciences*. 18, 4: 435-441.

Nyerges, T. 1992. Coupling GIS and spatial analytic models. *Proceedings of the Fifth International Symposium on Handling Spatial Data*. 2: 534-543. IGU Commission on GIS.

Openshaw, S. Cross, A. and Charlton, M. 1990. Developing a prototype geographical correlates exploration machine. *International Journal of Geographical Information Systems*. 4, 3: 297-311.

Purwanto, E. and Donker, N. H. W. 1991. Semi-distributed hydrologic modelling of the humid tropical upper Cimandri catchment (west Java) using HEC-1 model. *ITC Journal*. 4: 241-253.

Quinn, P., Beven, K., Chevallier, P. and Planchon, O. 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*. 5, 1: 59-79.

Rango, A. 1985. Assessment of remote sensing input to hydrological models. *Water Resources Bulletin*. 21, 3: 423-432.

Rango, A., Feldman, A., George, T. S. and Ragan, R. M. 1983. Effective use of Landsat data in hydrologic models. *Water Resources Bulletin*. 19, 2: 165-174.

Raper, J. F. and Kelk, B. 1991. Three-dimensional GIS. In *Geographical Information Systems: Principles and Applications*. eds. D. J. Maguire, M. F. Goodchild and D. W. Rhind. Longman Scientific and Technical. Burnt Mill.

Rohdenburg, H., Diekkrüger, B. and Bork, H. -R. 1986. Deterministic hydrological site and catchment models for the analysis of agroecosystems. *Catena*. 13: 119-137.

Ross, M. A. and Tara, P. D. 1993. Integrated hydrologic modeling with geographic information systems. *Journal of Water Resources Planning and Management*. 119, 2: 129-140.

Rott, H. 1986. Estimation of daily runoff based on Meteosat data. In *Hydrologic applications of space technology*. IAHS publication no. 160. ed. A. I. Johnson, pp. 321-330. IAHS. Wallingford.

STSC. 1989. *Statgraphics Version 4.0*. STSC. Rockville.

Sasowsky, K. C. and Gardner, T. W. 1991. Watershed configuration and geographic information system parameterization for SPUR model hydrologic simulations. *Water Resources Bulletin*. 27, 1: 7-18.

Shea, C., Grayman, W., Darden, D., Males, R. M., and Sushinsky, P. 1993. Integrated GIS and hydrologic modeling for countywide drainage study. *Journal of Water Resources Planning and Management*. 119, 2: 112-127.

Smith, M. B. 1993. A GIS-based distributed parameter hydrologic model for urban areas. *Hydrological Processes*. 7, 1: 45-61.

Stuebe, M. M. and Johnston, D. M. 1990. Runoff volume estimation using GIS techniques. *Water Resources Bulletin*. 26, 4: 611-620. Tao, T. and Kouwen, N. 1989. Remote sensing and fully distributed modeling for flood forecasting. *Journal of Water Resources Planning and Management*. 115, 6: 809-823.

Tarboton, D. G., Bras, R. L. and Rodriguez-Iturbe, I. 1991. On the extraction of channel networks from digital elevation data. *Hydrological Processes*. 5, 1: 81-100.

Tarboton, K. C. 1992. Interfacing GIS and hydrological modelling: Mgeni case study. *Water SA*. 18, 4: 273-278.

Tomlin, C. D. 1990. *Geographic information systems and cartographic modeling*. Prentice-Hall. Englewood Cliffs.

Veregin, H. 1989. Error modeling for the map overlay operation. In *Accuracy of Spatial Databases*. eds. M. Goodchild and S. Gopal, pp. 107-114. Taylor and Francis. London.

Vieux, B. E. 1991. Geographic information systems and non-point source water quality and quantity modelling. *Hydrological Processes*. 5, 1: 101-113.

Weibel, R. and Heller, M. 1991. Digital terrain modelling. In *Geographic Information Systems Volume 1: Principles*. eds. D. J. Maguire, M. F. Goodchild and D. W. Rhind, pp. 269-297. Longman Scientific and Technical. London.

White, D. 1988. Grid-based application of runoff curve numbers. *Journal of Water Resources Planning and Management*. 114, 6: 601-613.

Wolfe, M. L. 1992. GIS-assisted input data set development for the finite element storm hydrograph model (FESHM). *Applied Engineering in Agriculture*. 8, 2: 221-228.

Wolff-Piggott, B. B. 1994. MSc Thesis. *Coupling geographical information systems and catchment hydrological models*. University of Stellenbosch.

Zhang, H., Haan. C. T. and Nofziger, D. L. 1990. Hydrologic modelling with GIS: an overview. *Applied Engineering in Agriculture*. 6, 4: 453-458.

APPENDICES

Appendix A Detailed Database Design

1 Data Files

1.1 General Data

1.1.1 Run-time Data: RUN.DAT

ITEM_NAME	WDTH	OPUT	TYP
RUN-ID	15	15	C
PARAM_FIL	12	12	C
PARAM_FIL_DATE	20	20	C
DATABASE	12	12	C
NO_CELLS	4	4	C
START_DATE	12	12	C
END_DATE	8	8	C
OUTPUT_FLAGS	12	12	C
DESCRIPTION	40	40	C

1.1.2 The model element list: ELEMENTS.DAT

ITEM_NAME	WDTH	OPUT	TYP
SEQ	4	4	C
LABEL	8	8	C
TYPE	12	12	C
NAME	45	45	C

The redefined item DESCRIPTION is composed of all of the above items.

1.1.3 Precipitation and Evaporation: PPN_EVAPN.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
MAP	4	10	F	3
MAE	4	10	F	3

1.2 Return-flow Cell Data

1.2.1 Soils and related information: SOIL.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
CELL_TYPE	12	12	C	
SOIL_TYPE	12	12	C	
AREA	4	10	F	2
IRG_AREA	4	10	F	2
AQUIFER_WIDTH	4	10	F	2
DRAINED_OUTFLOW	4	10	F	2

1.2.2 Irrigation: IRRIGATION.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
UNIT	2	2	B	
PCT_AREA	4	10	F	3
CROP	12	12	C	
WATER_SOURCE	12	12	C	
IRRIG_METHOD	12	12	C	

1.3 DAMS

1.3.1 Primary dam information: DAM1.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
MAX_VOL	4	12	F	3

1.3.2 Area-Capacity Information: DAM2.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
VOLUME	4	12	F	3
AREA	4	12	F	3
DEPTH	4	12	F	3
SEEPAGE	4	12	F	3

1.4 Canal node: CANAL.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
LENGTH	4	12	F	3
WIDTH	4	12	F	3

1.5 Routing node

1.5.1 Primary Reach Information: ROUTE1.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
LENGTH	4	8	F	2
NO_SECTIONS	4	8	B	

1.5.2 Section Information: ROUTE2.DAT

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
SECTION	4	8	B	
LENGTH	4	8	F	1
SLOPE	4	8	F	3

1.5.3 Cross-section information: point event database

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
LABEL	8	8	C	
LOCATION	4	12	F	3
SECTION	4	8	B	
X	4	7	F	1
Y	4	7	F	1
N	4	7	F	3

2 Meta-data Files

2.1 Project dictionary: PROJ.DIC

ITEM_NAME	WDTH	OPUT	TYP
PROJID	8	8	C
PROJNME	20	20	C
PROJLOC	25	25	C
PROJDESC	50	50	C
PROJLDR	25	25	C
PROJSDTE	8	8	D
PROJCDTE	8	8	D
PROJSTAT	8	8	C

2.2 Model element dictionary: ELEM.DIC

ITEM_NAME	WDTH	OPUT	TYP
ELEMTYPE	10	10	C
ELEMDESC	50	50	C

2.3 Feature dictionary: FEAT.DIC

ITEM_NAME	WDTH	OPUT	TYP	N_DEC
FEATID	4	6	I	
FEATNME	10	10	C	
FEATTYPE	10	10	C	
FEATDESC	10	10	C	
FED	4	6	F	2
FSD	4	6	F	2
FWD	4	6	F	2
FSYMTYP	5	5	C	
FSYMNO	4	6	I	

2.4 Attribute dictionary: ATTRIB.DIC

ITEM_NAME	WDTH	OPUT	TYP
ATTRID	4	6	I
ATTRDESC	50	50	C
ATTRNME	16	16	C
TABLE	32	32	C
INDEXED	3	3	C

2.5 Relate dictionary: REL.DIC

ITEM_NAME	WDTH	OPUT	TYP
RELATION	8	8	C
TABLE-ID	128	128	C
DATABASE	8	8	C
ITEM	16	16	C
COLUMN	32	32	C
TYPE	16	16	C
ACCESS	4	4	C

2.6 Element - table correspondence and menu dictionary: CROSSREF.DIC

This dictionary is explicitly for the use of the system. It would automate the range of tables offered by the user interface after a model elements has been selected, and specify the form menu to be used in editing a data table. This dictionary would enable Arc/Info to act as if it had a degree of object-orientation.

ITEM_NAME	WDTH	OPUT	TYP
ELEMTYPE	10	10	C
TABLE	32	32	C
TABLEDESC	20	20	C
MENU	14	14	C

A redefined item DESCRIPTION is also defined, consisting of the last two items.

2.7 Feature lineage dictionary: FLIN.DIC

ITEM_NAME	WDTH	OPUT	TYP
FEATID	4	6	I
CAPDTE	8	8	D
CAPSCL	15	15	C
SRCMAT	15	15	C
CAPMETH	12	12	C
CAPDEV	4	4	I
OPERATOR	25	25	C
CAPORG	30	30	C
MAINTORG	30	30	C
OWNORG	30	30	C
PUBDTE	8	8	D
CAPROJ	12	12	C

2.8 Attribute lineage dictionary: ALIN.DIC

ITEM_NAME	WDTH	OPUT	TYP
ATTRIBID	4	6	I
CAPDTE	8	8	D
SOURCE1	50	50	C
SOURCE2	50	50	C
SOURCE3	50	50	C
CAPMETH	50	50	C
CAPORG	30	30	C
CAPROJ	50	50	C

Appendix B   Arc macro language code

Almost all of the macros that were developed were integrated using a menu system. This appendix gives a general description of the code to assist people who wish to modify it. Those wishing only to use the existing system should refer to Appendix E (Setup and Use of the Interface) for directions.

The standards suggested for utility tools as specified in the ArcTools Coding Standards were followed in the AML coding for the project, so that the code could be easily integrated with a graphical user interface. Headers were used to document each program, and the dependencies between them.

The AML code was designed to work in a structured directory system (see Appendix E), but global variables have been used for directory and file names so that the software can easily be customised i.e. the file names and directory structure may be modified from those used in this project. All of the global variables are set in the file *envt.aml*. There are three main AMLs: *db_mngr.aml* which manages the setup and maintenance of the data files, *df_mngr.aml* which manages the transfer of configuration data between Arc/Info to DISA format and *ui_mngr.aml* which provides the user interface for editing and listing the database contents.

The interface system is launched by *disa.aml*, in the ArcPlot environment. This calls *envt.aml*, which sets the general information about the application including the directory system, file names and file suffix conventions using a series of global variables. The menu interface system is then invoked, which runs the three major macros mentioned above.

The critical variables set in envt.aml have been given upper-case names. These variable values may be changed, but the file names and suffixes are set separately in the C header file *disa.h* again, which must also be altered for the interface to work properly. The field delimiter for Arc/Info export files is also set in both sets of code. The file names are also entered explicitly in the cross-reference data dictionary (see Appendix A, 2.6), which is used only by the macro ui_mngr.aml.

A number of utility macros are called by df_mngr.aml and db_mngr.aml. Both of the main AMLs call *del.aml* to delete data files. Del.aml can be run in a 'silent' mode which makes no prompts or messages, or it can be directed to request user input. When used in interactive mode it creates a global variable *.action* in order to communicate the choice made to the calling AML. The calling AML deletes this variable as soon as it has been read.

Df_mngr.aml also calls *exp.aml* to export data files from INFO. Exp.aml creates an ASCII file delimited using a user-defined character (set in envt.aml), without explicitly specifying the data fields for each file. It is therefore very flexible. *Freq.aml* counts the number of records per model element for the data files that have a variable number of records. The data files produced by this macro and the ASCII files exported are erased as soon as the information has been translated. Df_mngr.aml calls *imp.aml* to import ASCII text files generated by C code to INFO.

Additional Notes: DB_MNGR.AML

The definition of the data file structures was recorded using a spreadsheet originally, and is shown in Appendix A. This information was converted to comma-delimited ASCII format and transferred to the workstation. The UNIX utility *awk* and the C-shell program *indef.csh* were used to reformat the ASCII files into a layout readable by the INFO DEFINE statement.

This AML automates the setup of the database structure, and also makes it easy to modify the database structure reported on here. File names may be changed by altering the variables set in envt.aml, while a different file structure may be achieved by modifying the ASCII files stored in the IMPORT directory (see Appendix E). If fields are added, deleted or their order is changed in the file it will be necesary to alter the C code described in Appendix C.

Db_mngr.aml can also be used to purge and delete the database, and to copy it to another workspace.

2. UI_MNGR.AML

This uses a look-up table approach to ensure that the data set is displayed in the correct format after a model element has been selected. The model element type is used to look up the associated data tables in the cross-reference dictionary (CROSSREF.DIC - see Appendix A). The user may then select the desired data table from the list. Each data table contains information for all of the relevant model elements eg. the soils data table contains soils information for every return-flow cell in the database. A selection is performed on the table on the model element label before listing it's contents, ensuring that only the records for the model element of interest are displayed.

Although the model element label is present in every table of the configuration database, making it possible to join them using the relational capabilities of INFO, only one related record can be dislayed at a time in Arc/Info 6.1. It is necessary for multiple related records to be displayed, if the user interface is to be faithful to the original PC menu system. This is why the "look-up" approach outlined above has been taken.

Appendix C System Program Code

The AML programs which export the contents of the GIS database workspace (see appendices A and B) create temporary files in the directory containing the C program. The C program then assembles all these temporary files into one file compatible with the DISA configurator in the export directory, containing dummy values where they are not set by the GIS.

The C program also uses three files which are generated by the AML program which contain statistical information rather than data values for DISA. These files are:

irrigation.stat - contains the number of irrigation units per return-flow cell

dam2.stat - contains the number of records in the area-capacity curve per farm dam element

route3.stat - contains the number of points per section describing the cross-sectional characteristics of the element. The number of sections per element is stored explicitly in *route1.txt*.

The purpose of these files is explained in more detail below.

The C code which converts the DISA configurator file to and from flat ASCII files is stored in several files: the main files (*import_cfg.c* and *export_cfg.c*), a shared header file (*disa.h*) and a shared file of low-level functions (*disa.c*). The import and export programs are compiled separately.

1. DISA.H

The header file defines a number of symbolic constants which are used throughout the code. They are divided into a number of different functional groups:

(a) *Program constants describing the configuration which are set as global variables in the Pascal* code for DISA.

(b) The data and statistical file names, as well as the field delimiter expected by the program.

N.B. these may be changed, but the definitions in the AML file ENVT.AML *must also be altered* so that they agree exactly.

(c) The placeholder values which will be assigned in the DISA configurator file when the GIS database has no data.

(d) Constants internal to the working of the C program, including one called DEBUG which will cause the program to report additional information on it's operation according to the setting made.

The header file also indicates which system function libraries are to be included for use by the program, contains the function declarations for the code in file *disa.c* and specifies that this file should be included in the compilation process.

2. DISA.C

The functions in this file are called repeatedly from *export_cfg.c*. They are stored separately in order to make the main program more readable.

3. EXPORT_CFG.C

The main function is very simple, and calling a number of subfunctions to process each the various major sections of the DISA configurator file in sequence. These sections are the header, the element summary, the cell input data, the cell data and the global data.

The information for the first two sections is contained in two INFO files, and is simply rearranged and written out in the correct format. Function write_header deals with the header, and write_summary follows it. The summary function also loads this information into a C data structure that can pass it to the other major functions. The summary information is used to partially write the cell input section, and the cell data proper is then read from the GIS data files and formatted in the body of the interface.

All the data files are opened, and the program steps through the element summary, reading the appropriate data files line-by-line. Both the summary and the data files list the element label, and these are cross-checked each time. The program aborts if a mismatch is encountered, and provides information to help in tracing the cause of the problem.

It is ESSENTIAL that the sequence of elements is the same in all the files, and it is assumed that the routing elements will always be labelled in alphabetically ascending order, according to their processing order.

The entire program is built on a process of reading files one line at a time, and then extracting character strings from the line and checking that they are of the expected type and length before converting them to integers or floating point numbers where appropriate.

Appendix D Map Projection Information

The source data for the GIS was mostly at 1: 50 000 scale, and in the Gauss projection. The study area fell into both the lo 19 and the lo 21 zones, and so all data was reprojected into the Albers Equal Area projection before it was integrated and analysed.

Appendix E Setting up and Using the System

1. Hardware and Software Specifications

This project was carried out using a SUN Sparc-2, with 32 MB of RAM. The operating system was SunOS 4.1.3, and Arc/Info 6.1 was used. SunOS 4.1.3 includes a C-compiler, and this was used for system programming. Some basic C-shell programming was used to develop the interface.

The C code developed as part of this project was compiled successfully on a PC using Borland Turbo-C 2. The only modification required was a renaming of one of the included *.h* files from *strings.h* to *string.h.*

2. System Set-Up

As indicated in the appendices dealing with the coding, the interface system was designed to be flexible. The description of the setup here is for the project. Appendices A, B and C should be consulted if you would like to customise your system.

The project directory structure is central to the running of the system. Following ESRI guidelines separate directories were used for data for programs, raw data, import, export, working use and for the cleaned database. The directory structure is used in the project as illustrated in Figure 36 below:

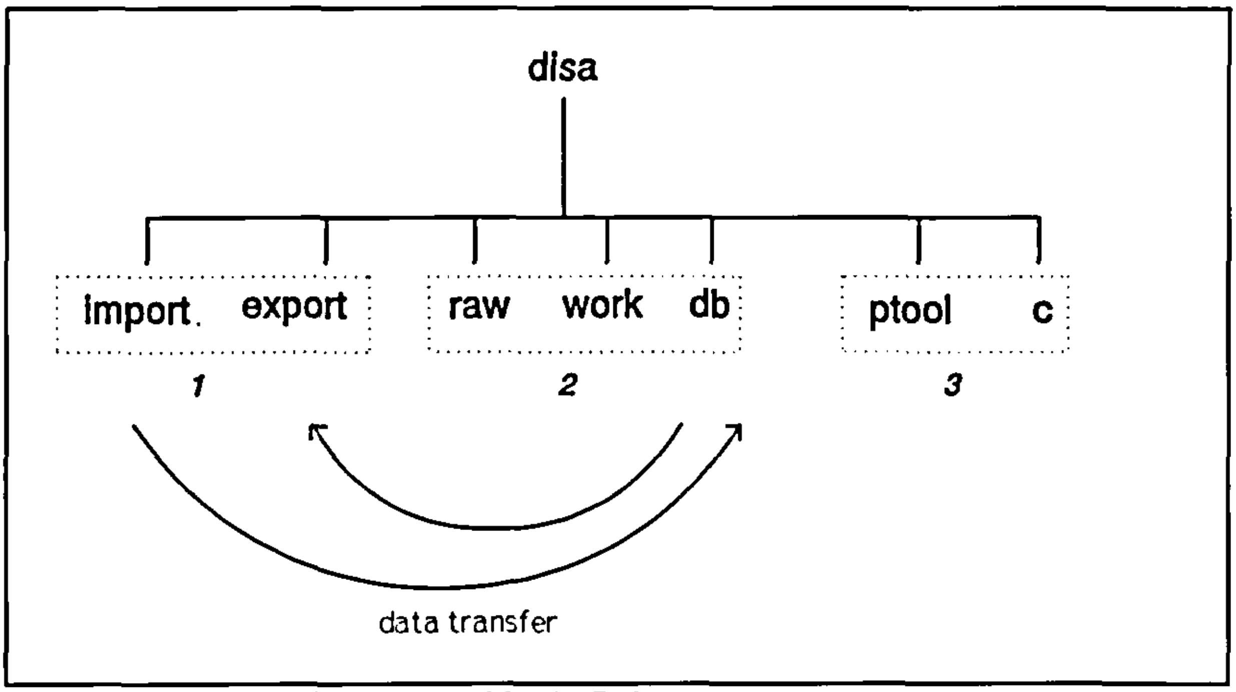


Figure 36 The Directory Structure used for the Project

The directories have been grouped according to their purpose. Group one directories are used to hold data being imported to or exported from the database via the interface system. These would be ASCII files such as the data file definition tables (see Appendix B, Additional Notes) or the DISA configurator file.

Group two directories are used for storing data at different stages of processing. The *raw* directory is used to hold data that has to be processed before importation to Arc/Info. The *work* directory is to hold data that is being captured and cleaned before entry to the database. This directory may have a number of subdirectories. The *db* directory is used to hold the final database. The interface system generates temporary files in the Arc/Info workspace where the database is established.

The group three directories hold the code which drives the interface. AML code and menus are stored under *ptool*, while C-code and C-shell scripts are held in the directory named *c*. It should be noted that the awk script *indef.awk* is also stored in this directory, but it needs to be relocated to the directory where the database structure is to be created.

3. Application Procedure

In this project only canal, river reach, soils and model cell information was available in a format suitable for capture in a GIS database. The application procedure will be discussed with particular reference to these data sets, but the methodology can be extended to cover the other data as well.

3.1 Preparing to Configure DISA

The DISA models structure consists of a network of modelling elements, which is graphically represented in the system diagram (). The has to be developed before the model can be configured. This requires that a physical overview of the catchment be acquired, in order to be able to decide how it should be discretized.

GIS provides a powerful tool to integrate and present the base information for an entire catchment, and this should be the first step in preparing to use DISA. The first use for this catchment information is in guiding the delineation of the physiographic regions, using the geology, morphology and points of interest in the study area (see Figure 34). Then the alluvial and terrace soil groups have to be generated and overlaid with the physiographic regions to form the return-flow cells.

The position of the return-flow cells relative to the water-supply elements suh as farm dams, canals and the hydrologically defined river reaches can then be used to develop a to represent the system.

3.2 GIS Procedures

In this project only canal, river reaches, soils and physiographic units were available in a format suitable for capture using GIS, and so this discussion will focus on these datasets. The procedure could be extended considerably following the methodology outlined here. The soils and physiographic regions are polygon coverages, which need to be overlaid to derive the return-flow cells, while the river reach and canals need to be discretized according to hydrological criteria and the return-flow cells they overlay respectively.

The four data sets were digitized and cleaned, and user identity numbers were assigned to differentiate the different features. The soils had to be reclassified into alluvial and terrace soil groups before the overlay could proceed. Another item was added to the PAT and each polygon was assigned the text "ALLUVIAL" or "TERRACE" using the INFO database. The DISSOLVE command was then used to merge adjoining soil polygons of the same type. The model cell names were also added to the coverage PAT.

A polygon-polygon overlay was then carried out in order to merge the soil groups and model cell coverages. The IDENTITY command was used in order to minimise slivers on the boundary. The ratio of perimeter to area of the polygons in this coverage was used as a criterion to help identify slivers, which were noted using WRITESELECT in ArcPlot, followed by ELIMINATE. A macro utility called *writelab.aml* was then run to create labels for the return-flow cells which had been delineated, following the convention established for the DISA model eg. the overlay of a terrace soil and model cell RB100 resulted in a model cell label of RB100T, while the overlay of the same model cell with an alluvial cell was labelled RB100A.The STATISTICS command was then used to summarize the areas of all return-flow polygons having the same label. This data can then be tranferred directly to the main database file SOIL.DAT.

The canals were discretized by overlaying them with the physiographic regions coverage, and entering the appropriate label in the AAT for each arc. It was found that the return-flow cells coverage was too detailed to use for this overlay. The river reaches should be derived from the central dividing line between the physiographic regions, rather than being independently digitised. It may be necessary to unsplit some of the arcs here, as a reach may adjoin a number of physiographic regions. A river reach should not be defined at a point other than a physiographic region boundary however.

Dynamic segementation may be used to attach the river cross-section information to the arcs. This was not achieved in this study, however, as it appears to require interactive specification of the location of cross-sectional points. It would be feasible if GIS was used to capture the data from the original source materials.

3.3 Using the Menu System

The interface menu system offers three main headings, which deal with the definition of the database structure, the editing and listing of the database and the interchange of data (Figure 37). The definition menu allows the automated definition of a database structure as described in Appendix A. It can also be used to purge the database, delete the files or copy the entire database to a new workspace. By default, the database templates will be created in the db directory.

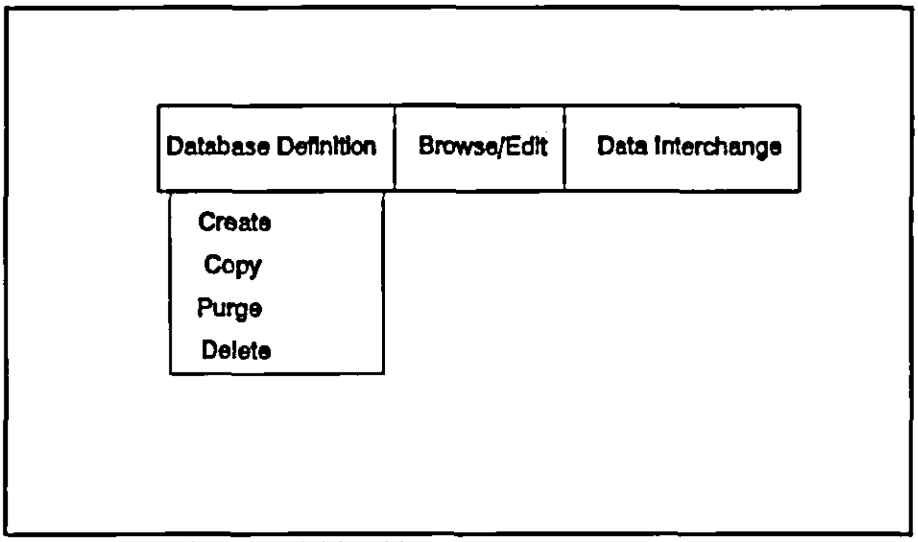


Figure 37 Database Definition Menu

The Browse/Edit submenus provide a DISA-like view of the database using the model elements as the entry-point to the system. This viewing system can be used to determine whether all the spatial information for each model elements has been entered into the system. The edit option provides form menus to aid data entry;

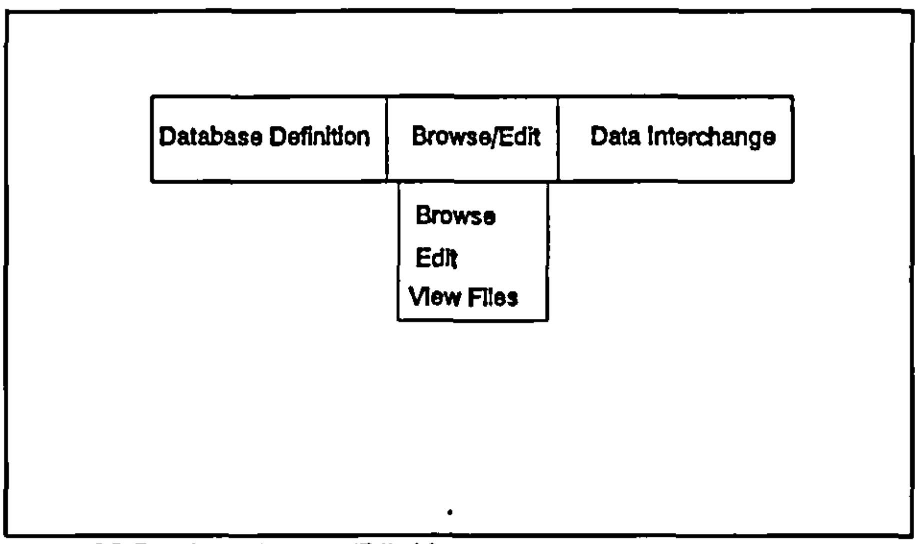


Figure 38 Database Browse/Edit Menu

only one data record can be viewed at one time using form menus however. The file view option lists individual database files, including the metadata files and attribute tables. Data entry to these last two sets of files cannot be done from the menu system.

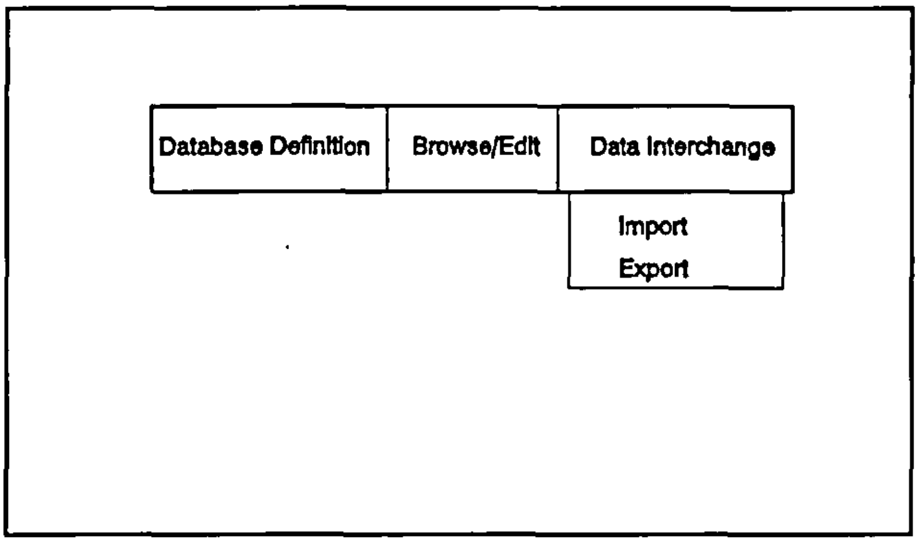


Figure 39 Data Interchange Menu

Data interchange has a very simple menu system, which imports or exports data in batch mode. It uses the import and export subdirectories to locate the DISA configurator file, and the db directory for the database by default. It should be noted that the configurator file generated from Arc/Info data is incomplete, and a number of dummy values are entered into the file so that the DISA model will read it without error. These dummy values have to be modified to observed values and considerable amounts of non-spatial data also have to be entered before running the model.

Appendix F External Verification Results

1 DISA and Arc/Info input data

The results for the three spatial parameters derived from the GIS procedure are tabulated below.

Table 10 Percentage Difference in Return-Flow Cell Areas

Record	Label	Area - DISA (ha^2)	Area - Arc/Info (ha^2)	Percentage Difference
1	RB010T	3008.00	2495.07	-17
2	RB010A	2717.00	2941.22	8
3	LT010T	10248.00	9805.57	-4
4	LT010A	6936.00	6978.56	1
5	LB010T	428.00	325.86	-24
6	LB010A	463.00	474.40	3
7	RT010T	4417.00	4051.18	-8
8	RT010A	3840.00	3819.81	-1
9	RT020T	3097.00	2989.43	-4
10	RT020A	2598.00	2719.46	5
11	RB020T	5526.00	5547.56	0
12	RB020A	813.00	942.51	16
13	LB020T	1002.00	954.58	-5
14	LB020A	438.00	435.21	-1
15	LB030T	2141.00	1849.32	-14
16	LB030A	518.00	556.68	8
17	RB030T	3223.00	3158.40	-2
18	RB030A	817.00	744.71	-9
19	LB040T	328.00	348.13	6
20	LB040A	265.00	246.57	-7
21	RB040T	1307.00	1421.42	9
22	RB040A	195.00	256.34	31
23	LB050T	1489.00	1538.37	3
24	LB050A	514.00	525.34	2
25	RT030T	8950.00	1187.48	-87
26	RT030A	1720.00	58.26	-97
27	RB050T	833.00	740.31	-11

28	RB050A	292.00	321.84	10
29	LT020T	6768.00	6686.73	-1
30	LT020A	952.00	977.33	3
31	LB060T	1010.00	956.38	-5
32	LB060A	286.00	303.09	6
33	RB060T	2237.00	2342.25	5
34	RB060A	463.00	488.86	6
35	LT030T	2619.00	2402.38	-8
36	LT030A	364.00	325.95	-11
37	RB070T	282.00	340.45	21
38	RB070A	32.00	83.36	161
39	LB070T	238.00	210.29	-12
40	LB070A	382.00	397.01	4
41	LB080T	2822.00	2834.12	0
42	LB080A	418.00	452.89	8
43	RT040T	5184.00	3858.30	-25
44	RT040A	1223.00	737.31	-40
45	RB080T	1341.00	1554.23	16
46	RB080A	230.00	230.77	0
47	LT040T	1096.00	1438.02	31
48	LT040A	691.00	778.70	12
49	LB090T	698.00	672.69	-4
50	LB090A	326.00	358.17	10
51	RB090T	202.00	200.87	-1
52	RB090A	96.00	91.30	-5
53	RB100T	68.00	69.97	3
54	RB100A	25.00	29.08	16
55	LB100T	460.00	385.29	-16
56	LB100A	84.00	86.26	3

Table 11 Percentage Difference in Canal Section Lengths

Record	Label	Length - DISA (km)	Length - Arc/Info (km)	Percentage Difference
1	C010	11200.00	11299.97	1
2	C020	700.00	829.87	19
3	C030	1100.00	379.72	-66
4	C040	7000.00	7092.45	1
5	C050	20300.00	13018.38	-36
6	C060	6400.00	6576.50	3
7	C070	5700.00	6636.77	16
8	C080	1450.00	2558.13	76
9	C090	7300.00	7342.10	1
10	C100	7500.00	7397.52	-1
11	C110	385.00	0.00	100
12	C120	7450.00	7454.70	0
13	C130	8450.00	8171.71	-3
14	C140	3400.00	2203.49	-35
15	C150	4250.00	3982.50	-6
16	C160	3300.00	3035.39	-8
17	C170	7000.00	6653.36	-5
18	C180	1500.00	1241.35	-17
19	C190	3000.00	3953.58	32
20	C200	4900.00	2804.81	-43
21	C210	3300.00	2179.79	-34
22	C220	5100.00	5676.39	11
23	C230	2700.00	2506.43	-7
24	C240	3000.00	2836.19	-6

Table 12 Percentage Difference in River Reach Lengths

Record	Label	Length - DISA (m)	Length - Arc/Info (m)	Percentage Difference
1	R010	9552.00	9392.83	-2
2	R020	6306.00	5976.21	-5
3	R030	14738.00	14804.75	1
4	R040	5370.00	5646.82	5
5	R050	8316.00	7609.73	-9
6	R060	8064.00	8284.34	3
7	R070	4206.00	4915.01	17
8	R080	7056.00	6804.22	-4
9	R090	4790.00	4880.84	2
10	R100	3756.00	3123.87	-17