International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Modeling Support System for Systems Analytic Research

Nakamori, Y. and Sawaragi, Y.

IIASA Working Paper

WP-87-024

March 1987

MODELING SUPPORT SYSTEM FOR SYSTEMS ANALYTIC RESEARCH

Y. Nakamori* and
Y. Sawaragi**

March 1987
WP-87-24

*Department of Applied Mathematics, Konan University, Kobe, Japan

**Department of Computer Science, Kyoto Sangyo University, Kyoto, Japan

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

Recently the complexity of systems being investigated in Systems Analysis increased essentially. Consequently, the task of model building becomes more difficult and in many important practical cases the existing model building methodologies are not sufficient. This applies especially to such cases, when the system under study contains ill-defined parts, like social and ecological subsystems.

Therefore, in such cases, it is necessary to consider during the model building process the expert knowledge about the systems. One of the possible ways of utilizing such a knowledge is to apply interactive procedures for model building and verification. The paper presents such a methodology, named IMSS (Interactive Modeling Support System) together with a detailed description of the computer implementation. One of the advantages of the presented system is the facility for structuring both mental and mathematical models, that facilitates the model understanding and confidence. According to the IMSS methodology, the model building process consists of several stages performed interactively; during the process the model builder has to specify a'priori information about causality relations between model variables and has to supervise the statistical procedures during the parameter estimation phase. The interactive approach is also being used for model simplification and elaboration. The system presented in the paper is an example of new methodology of computers usage, based on new mathematical techniques and algorithms for real-time man-machine interaction.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

MODELING SUPPORT SYSTEM FOR SYSTEMS ANALYTIC RESEARCH

Y. Nakamori

Department of Applied Mathematics, Konan University

Y. Sawaragi

Department of Computer Science, Kyoto Sangyo University

Keywords: Man-machine interfaces and dialogue systems, mental models, training, human decision making, decision support systems, computer-assisted modeling.

Abstract: A computer-assisted mathematical modeling system based on an advanced interactive methodology is presented. The system aims at the extraction of a trade-off between human mental models and regression-type models based on the use of numerical data, by using a flexible combination of statistical methods and graph-theoretical analysis. One of the main advantages of using this system is the facilities for the structuring of both mental and mathematical models.

1. INTRODUCTION

The primary aim of a system analytic study is to find a synthesis of formal and informal, mathematical and non-mathematical methods, procedures, approaches, etc., and to design a computer-based system as a qualitatively new tool for the analysis of concrete problems. This process of design should be able to incorporate different types of available knowledge and information about the underlying system. The non-quantifiable knowledge of people, who from their experience know many important properties of the system, is often of a high value. Therefore, having efficient channels of communicating this type of knowledge into the process of design is very desirable. This type of a communication channel is one of the characteristic features of the modeling procedure described in this paper. The use of the highly interactive computer system based on this procedure is an interactive process in the course of which the subjective expert knowledge can be utilized to a great advantage.

The advanced interactive methodology presented in this paper consists of a combined modeling techniques of statistical and graph-theoretical approaches, and related multi-stage man-machine dialogues. The corresponding computer system developed is called the Interactive Modeling Support System (IMSS). One of the main advantages of using this system is the facility for the structuring of both mental and mathematical models, that facilitates the model understanding and confidence. It is based on the interpretive structural modeling (Warfield, 1976) and contains mathematically new techniques and algorithms for real-time man-machine interactions. The other characteristic features are a flexible treatment of the data set, self-organization methods including the stepwise regression procedures and the group method of data handling (Ivakhnenko, 1968), and the graphic-display facilities for information exchange. One of the ideal stages in developing the system IMSS is that a graph works as if an advanced computer language for communication between the rigid, logical computer and the vague, incomplete human minds. The computer and human interaction and communication can be conceptualized as in Figure 1.

2. INTERACTIVE MODELING METHODOLOGY

The modeling procedure to be presented in this paper requires the following three types of information:

- A set of variables to describe the system under study:

$$S = \{ x_1, x_2, \cdots, x_n \}.$$

- The corresponding measurement data table:

$$X = (x_{ij}), \quad i=1,2,\ldots,n, \quad j=1,2,\ldots,N,$$

where x_{ij} represents the j-th measurement of the variable x_i.

- A relation B, on the product set $S \times S$, defined by

(x_i, x_j) is in B if and only if x_i influences x_j,

or, equivalently, an adjacency matrix $A=(a_{ij})$ defined by

$$a_{ij} = \begin{cases} 1 & \text{if } (x_i, x_j) \text{ is in B} \\ 0 & \text{otherwise.} \end{cases}$$

The objective of the modeling is to obtain a set of linear equations:

$$x_i = c_{i0} + \sum_{x_j \varepsilon S_i} c_{ij} x_j , \quad i=1,2,\cdots,n,$$

where $S_i = S - \{ x_i \}$, $i=1,2,\ldots,n$, that describes the underlying system and is capable of predicting the behavior of the system.

The modeling process consists of three different but interdependent stages of dialogues as shown in Figure 2. The first stage dialogue is required for preparation of the modeling, including input of measurement data and the initial version of cause-effect relation on the set of variables, transformation of variables, data screening, and refinement of the cause-effect relation. The second stage dialogue is devoted to finding a trade-off between the measurement data and the modeler's knowledge about dependencies between variables. Based on the measurement data and the initial version of the cause-effect relation, using the option of regression method, the computer finds a linear model and the corresponding digraph model. The modeler can modify the new relation referring to these computer models and his or her knowledge. The process continues repeatedly until no change occurs or the modeler is satisfied with the modified relation. The third stage dialogue is related to model simplification and elaboration. Model simplification is based on the use of equivalence relation, and model elaboration is an application of regression analysis including the hypothesis testing on estimated coefficients, and examination of the explanatory and predictive powers of the model.

The first craft required is the selection of descriptive variables. The variable set S can include nonlinear reexpression or time-delayed variables of the initial variables. Following the traditional usage, we use the term "linear model" to describe a set of equations whose structural parameters are embedded linearly. Reexpression and time-shifting enable us to analyze nonlinear relationships and multiple autoregressive processes, respectively. At the second and third stages, the corresponding data set is required to be complete in the sense that it is screened enough to avoid multicollinearity or the influence of outliers. This does not imply that the data should be measured absolutely correctly. Soft observation is allowed to compensate for lacking or extraordinary data. Hereafter, we use the term "observation" instead of "measurement", meaning that observation includes data estimated or modified by the modeler.

Because both the complexity and ambiguity of an object depend on the interests and capabilities of the individual, the introduction of a cause-effect relation is also a craft work. But in-depth considerations are not required initially; the remaining ambiguities are resolved after some iterative modeling sessions. In the second stage, the self-organization methods are used to obtain linear equations and graph-theoretic techniques are used for man-machine interactions. The required human input is knowledge of the structural image of the system. This stage includes part of the model verification, because the modeler should judge whether the model behaves in general as he or her intends. Even the experts can hardly tell whether the obtained model is appropriate or not because the coefficients of a linear model do not necessarily have practical meanings. Therefore, the structure of the model is extracted and shown in the form of digraphs to help the understanding and modification of the computer model. The third stage is concerned with judgments about the validity of the model in terms of its explanatory and predictive powers.

3. INTERACTIVE MODELING SUPPORT SYSTEM

The Interactive Modeling Support System (IMSS) is a highly user-friendly software providing for an interactive person-computer dialogue facilitated by the use of advanced techniques to communicate directly graphic information to the computer and receive graphic output. The original system is implemented on the NEC's microcomputer, and the revised version is developed for the use of IBM-PC and VAX 11/780. The following main advantages of its use are emphasized:

- The data-screening features provide a powerful tool for debugging the data-set.
- The structural modeling features are helpful for organizing one's thinking with respect to the system under study.
- It enables rapid access to the set of relationships comprise the statistical model.
- It makes possible rapid validation and easy refinement of the statistical model.

The modeling information (S,X,A) is fed into the computer at the first step. IMSS has several facililies to read and preprocess the data set. The option of data transformation makes it possible to analyze time-lag effects or functional relationships. Data transformation is also needed to make distributions of variables symmetric because, according to Hartwig and Dearing (1979), non-symmetric distributions and nonlinear relationships often exist together. If every distribution of variables is roughly symmetric, then we will have a high chance to obtain a linear model. IMSS prepares histograms of the original and transformed variables to assist the modeler's judgment.

If at some step the modeler want to check distributions or outliers of the data for some variables, IMSS assists him by showing the list of candidates of outliers, histograms or scattergrams. The modeler can designate the case numbers which he does not want to use in modeling. IMSS also checks and displays pairs of variables which have high correlation coefficients. To avoid the problem of multicollinearity and also to simplify the model, it is recommended that one of the pair is set aside when they are supposed to be linearly dependent. The modeler can be referred to the condensed basic statistics and scatter plots.

The manner of filling the adjacency matrix A should be negative. Here negative means that the modeler should enter the computer a part of his knowledge, putting 0's at the right places. The rest of entries will be filled with 1's by the computer. The underlying idea is that the modeler should inquire into strength of relationship between every pair of variables except those which are definitely irrelevant. An extension of binary relation is allowed in filling the matrix $A = (a_{ij})$:

$$a_{ij} = \begin{cases} 2 & \text{if } x_i \text{ certainly influences } x_j \\ 0 & \text{if } x_i \text{ never influences } x_j \\ 1 & \text{otherwise.} \end{cases}$$

There is no difference between 1 and 2 in structural modeling, but they are treated differently in statistical modeling, i.e., the variables indicated by 2 are regarded as the core variables and those indicated by 1 the optional variables.

IMSS has another option of filling the matrix A. The relation considered is the cause and effect that is not necessarily transitive. But it may be quite feasible to employ the assumption of transitivity to develop a linear model. The modeler can choose the option of a transitive embedding method that is a modified version of that in Warfield (1976). The advantage of this method is that it can reduce the number of pairwise comparisons remarkably. One caution in using this method is that the modeler should consider indirect cause-effect relationships as well as direct ones. IMSS provides the digraph of hierarchy based on the adjacency matrix A, taking its transitive closure and extracting the skeleton. Figure 3 shows the process of the first stage dialogue.

After the first stage dialogue, the set of variables S and the data matrix X are fixed and will be used in the subsequent stages as they are. The adjacency matrix A is alone open for further modification. The purpose of the second stage is to elaborate the cause-effect relations which are summarized in A. First, the modeler must choose one of the options of regression methods with self-selection of explanatory variables, which will be used in the next step. The options of these include:

- the forward selection procedure,
- the backward elimination procedure,
- the all possible selection procedure, and
- the group method of data handling.

The last one is recommended to use when the number of data points is very small. It is a modified or simplified version of the original one (Ivakhnenko, 1968), i.e., the partial descriptions is written in a linear form with respect to variables.

Let us demote by $\bar{C} = (c_0, C)$, where c_0 is an n-column vector and C an $n \times n$ matrix, the coefficient matrix of a set of linear equations which the computer will search from now on:

$$x = c_0 + Cx \text{ with } c_{ii} = 0, \text{ all } i,$$

where x denote the n-column vector whose components correspond to the variables $x_1, x_2, \cdots, x_n$. By the selected automatic modeling method, the computer will estimate the row vectors of $\bar{C}$ one by one referring to the matrix A that can have been converted into a reachability matrix.

For example, suppose now a turn is the i-th row vector of $\bar{C}$. Two subsets of S are defined by referring to $A = (a_{ij})$:

$$S_i^2 = \{ x_j \ ; \ a_{ji} = 2, \ j \neq i \}$$
$$S_i^1 = \{ x_j \ ; \ a_{ji} = 1, \ j \neq i \}.$$

The variables x_j in S_i^2 are always chosen as explanatory variables for x_i, and those in S_i^1 are examined whether they should be explanatory variables or not for x_i. Thus the computer finds a linear model:

$$M_c = (S, \bar{C}) \text{ or } x = c_0 + Cx.$$

Then the adjacency matrix A will be modified as follows. Let us denote the new (i,j) entry of A by a_{ij}^n and the old (i,j) entry by a_{ij}^o. The new (i,j) entry is obtained by

$$a_{ij}^n = \begin{cases} 0 \text{ if } c_{ji} = 0 \text{ and } x_i \in S_j^1 \\ a_{ij}^o \text{ otherwise.} \end{cases}$$

Thus some of 1's in A will turn into 0's. The corresponding relation B is then modified by an apparent manner.

Let us introduce a digraph D defined by

$$D = (S, B),$$

where the elements of S are identified as vertices and those of B arcs. The vertices are represented by points and there is a directed line heading from x_i to x_j if and only if (x_i, x_j) is in B.

Let $\bar{B}$ denote the transitive closure of B, and suppose the variable set S can be divided into m cycle sets $e_1, e_2, \cdots, e_m$; here e_p is defined by

$$x_i, x_j \in e_p \text{ if and only if } (x_i, x_j), (x_j, x_i) \in \bar{B}.$$

Then we can define new sets:

$$\bar{S} = \{ e_1, e_2, \cdots, e_m \}$$

$$B^* = \{ (e_p, e_q) ; \text{ some } (x_i, x_j) \in \bar{B}, x_i \in e_p, x_j \in e_q \}$$

and the corresponding digraph is called the condensation digraph.

Finally we introduce the skeleton digraph $\bar{D}$ which is a minimum arc subdigraph of D, for which removal of any arc would destroy reachability presented in the relation. Actually the above process is carried out by some matrix operations in the computer. After all, the computer will have found the digraph model:

$$M_D = (\bar{S}, \bar{D}),$$

that is a visual version of the linear model M_C.

If the modeler is satisfied with the model structure, the modeling process will proceed to the third stage. Otherwise, the second stage will be repeated again after amendments of the digraph model M_D. The modification facility of the digraph is one of the most fascinating parts of IMSS; but the detail description is omitted here. The reader is referred to Nakamori et. al. (1985). Even an expert can hardly tell whether the obtained linear model is appropriate or not because of the difficulties of checking validity of the hypothesis testing and giving meaning to regression coefficients. The most emphasized point of using IMSS that is the refinement of the statistical model can be done by the modification of the structural model. Figure 4 and 5 show the man-machine interactions and the matrix exchanging process in the second stage, respectively.

The third stage dialogue consists of two modes: model simplification and model elaboration. Model simplification is based on the use of equivalence relation and model elaboration is related to applications of regression analysis. IMSS prepares most of the classical procedures in regression analysis including the hypothesis testing on the estimated coefficients, and the examination of explanatory and predictive powers of the model with the aid of graphic facilities of the computer. Figure 6 schematizes the third stage dialogue.

4. USAGE OF THE SYSTEM IMSS

The modeling process of using IMSS can be drawn as in Figure 7. The program consists of 50 submodules as shown in Table 1, and the data files to be produced are shown in Table 2. The relations between the upper level menus are shown in Figure 8.

Figure 9 is the openning menu and Figure 10 is the problem selection menu that appears when the first menu is selected in the openning menu. Figure 11 is the master menu that can call the first, second and third stage modeling menus. But the second and third stage menus are not accessible until the first stage modeling has been finished. The menu 4 in the master menu leads the menu to show the mathematical model developed up to then.

Figure 12 shows the main manu for first stage dialogue, in which the menus form an hierarchical structure as shown in Figure 13. The menu 1: INITIALIZATION should be selected at the beginning of the modeling, which initializes the commonly used parameters such as the number of variables, and prepares several data files. If the menu 2: MEASUREMENT DATA is chosen, the data input menu will appear as shown in Figure 14, in which the works of data input, up-dating and correction are carried out.

The menu 3: CAUSAL RELATION in Figure 12 leads to Figure 15 that is the input menu for causal relationships. Two types of input procedures are prepared: one is the sequential input and the other is the transitive embedding method, as shown in Figure 16. If the latter one is selected, the list of variables and a matrix will appear as Figure 17. After some dialogues with the computer, a hierarchical digraph will be drawn as shown in Figure 18. The modification of digraph can be carried out if the menu 2 in Figure 15 is selected. The adjacency, reachability and skeleton matrices and the corresponding digraphs can be seen from the menu 5: SYSTEM STRUCTURE in Figure 12.

The data transformation menu, Figure 19, is led by choosing the menu 4: TRANSFORMATION in Figure 12. Figure 20 shows the prepared list of transformations. After a data transformation is carried out, the causal relationships are inferred by the computer, but the menu 8: BASIC CALCULATION should be called to calculate basic statistics of the new variable. From the menu 6: LINEAR RELATION, the causal relationships can be modified by referring to the correlation coefficients.

Figure 22 is the menu for data screening led by the menu 7: DATA SCREENING in Figure 12. Figure 23 shows the list of candidates of outliers. The basic statistics such as means, variances and correlation coefficients are calulated and memorized as in Figure 24 which is called by the menu 8: BASIC STATISTICS in Figure 12.

Figure 25 is the menu to show the modeling information which is followed by the menu I: INFORMATION in Figure 12, in which we can see the initial version of the cause-effect relation, the original and standardized data, the sample means and variances, the correlation coefficients and the scatter plats. Figure 26 is the menu for the scatter plats. Figure 27 is a dialogue to see the distributions of data and Figure 28 is an example of two dimensional scatter plots.

The opening menu of the second stage dialogue includes the menu for the selection of regression methods as shown in Figure 29. In Figure 30, the explanatory variables for X1 are being selected by the forward selection procedure. After this screen, the computer obtains a linear model and the corresponding digraphs, and finally leads Figure 31 which is the main menu of the second stage dialogue.

Figure 32 will appear if the menu 1 is chosen in Figure 31, which is the menu for the modification of model structure. Figure 33 is an example of dialogue concerning to the modification of adjacency relationships in a cycle set. The direct modification of the adjacency matrix can be done as shown in Figure 34.

The main menu for the third stage dialogue is shown in Figure 35. The first menu in Figure 35 is related to the model simplification, and Figure 36 is a part of dialogues to simplify the model. In model simplification, the reachability of the model can be checked and the statistics of the revised model are shown.

Figure 37 is the menu for regression analysis, which is led by selecting the second option in Figure 35. Here, the facilities for the traditional regression analysis, residual plots, multicollinearity checking and extrapolation are prepared. Figures 38 and 39 are related to the regression analysis. The model can be elaborated by referring to statistical information at the third stage.

After returning to the master menu, Figure 11, and selecting the menu 4, we can see the linear model obtained until then. Figure 40 is the menu to see the current model and the statistics for hypothesis testing, and Figure 41 shows an equation in the model.

5. CONCLUSIONS

The interactive modeling support system (IMSS) is a tool for enlightening both the modeler and the computer about the underlying system. The main point is how effectively extract reality from human mental models and also from measurement data with computer assistance. One important thing involved in developing any kinds of models is the learning and training experience. The system IMSS aimes at not only model building but also learning exercise. An application of using this system to the model simplification can be found in Walsum and Nakamori (1985) which reports how the modeler obtained a deep insight into the underlying system and problems.

REFERENCES

Harwig, F., and B. E. Dearing (1979). Exploratory Data Analysis, Sage, London.

Ivakhnenko, A. G. (1968). The group method of data handling, a rival of the method of stochastic approximation. Soviet Automatic Control, 13, 3, 43-55.

Nakamori, Y., M. Ryobu, H. Fukawa, and Y. Sawaragi (1985). An Interactive Modeling Support System. IIASA Working Paper, WP-85-77. IIASA, Laxenburg, Austria.

Van Walsum, P. E. V., and Y. Nakamori (1985). Simplification of a comprehensive hydrological model for scenario analysis. IIASA Working Paper, WP-85-92. IIASA, Laxenburg, Austria.

Warfield, I. N. (1976). Societal Systems: Planning, Policy and Complexity. John Wiley, New York.

Table 1: The program package of the system IMSS.

code	stage	level	function
100	-	1	opening message, data-file control
200	-	1	MASTER MENU
300	1	2	MENU for the first stage dialogue
310	1	3	initialization of modeling parameters
320	1	3	MENU for input of numerical data
321	1	4	filing of numerical data
322	1	4	appending data of new variables
323	1	4	appending up-dated data
324	1	4	correction of the data file
330	1	3	MENU for input of causal information
331	1	4	sequential input of causal information
332	1	4	the transitive embedding method
333	1	4	calculation of transitive closures
334	1	4	modification of causal information
340	1	3	data transformation
350	1	3	graphic display of causal information
360	1	3	check of linear relationships
370	1	3	check of extraordinary data
380	1	3	calculation of basic statistics
400	2	2,3	MENU for the second stage dialogue
411	2	4	the forward selection procedure
412	2	4	the backward elimination procedure
413	2	4	the all possible selection procedure
414	2	4	the group method of data handling
420	2	3	memorization of model equations
430	2	3	extraction of causal relationships
440	2	3	graphic display of causal relationships
450	2	3	modification of causal relationships
500	3	2	MENU for the third stage dialogue
510	3	3	model simplification
520	3	3	MENU for model elaboration
521	3	4	hypothesis testing by regression methods
522	3	4	residual plots of the model
523	3	4	multicollinearity checking
524	3	4	estimation by the model
525	3	4	display of regression results
530	3	3	prediction by new data
540	3	3	extraction of causal relationships
600	1	3	MENU for modeling information
610	1	4	the initial version of causal relation
620	1	4	display of the original numerical data
630	1	4	display of the standardized data
640	1	4	display of sample means and variances
650	1	4	display of correlation coefficients
660	1	4	MENU for hisrograms and scattergrams
661	1	5	one-dimensional scatter plots
662	1	5	two-dimensional scatter plots
663	1	5	three-dimensional scatter plots
700	-	2	display of the current model
800	-	1	ending message

Table 2: Data files to be produced in using IMSS.

code	modeling stage	contents
00	-	the problem lists, the current problem
01	1,2,3,4	commonly used modeling parameters
02	1,2,3,4	the list of system variables
03	1,2,3,4	the list of outliers
04	1,2,3,4	the table of original numerical data
05	1,2,3	the table of sample means
06	1,2,3	the table of sample variances
07	1,2,3	the table of standardized data
08	1,2,3	the table of correlation coefficients
09	1	the initial version of adjacency matrix
10	1	the initial version of reachability matrix
11	1	the initial version of skeleton matrix
12	2,3,4	the current version of adjacency matrix
13	2,3	the current version of reachability matrix
14	2,3	the current version of skeleton matrix
15	3	the results of hypothesis testing
16	4	the list of model equations
17	3	the data dable for prediction

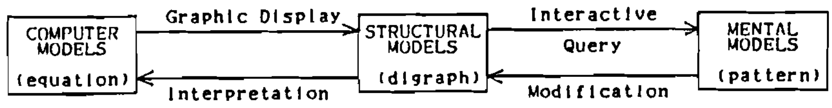


Figure 1: Concept of human-model-computer interaction and communication.

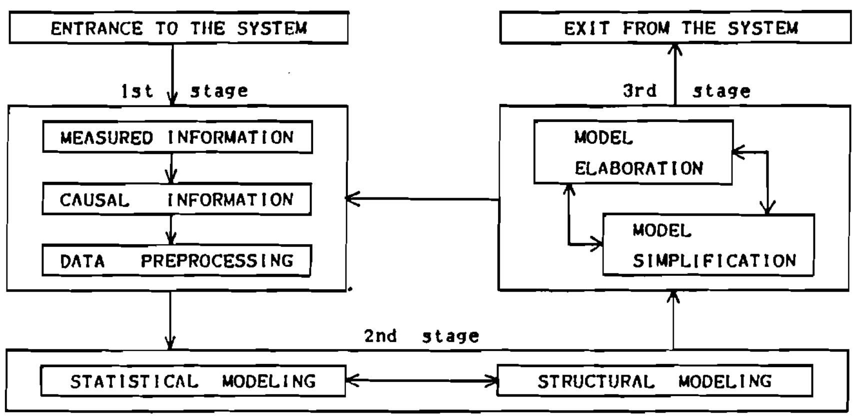


Figure 2: The structure of the modeling process of using IMSS.

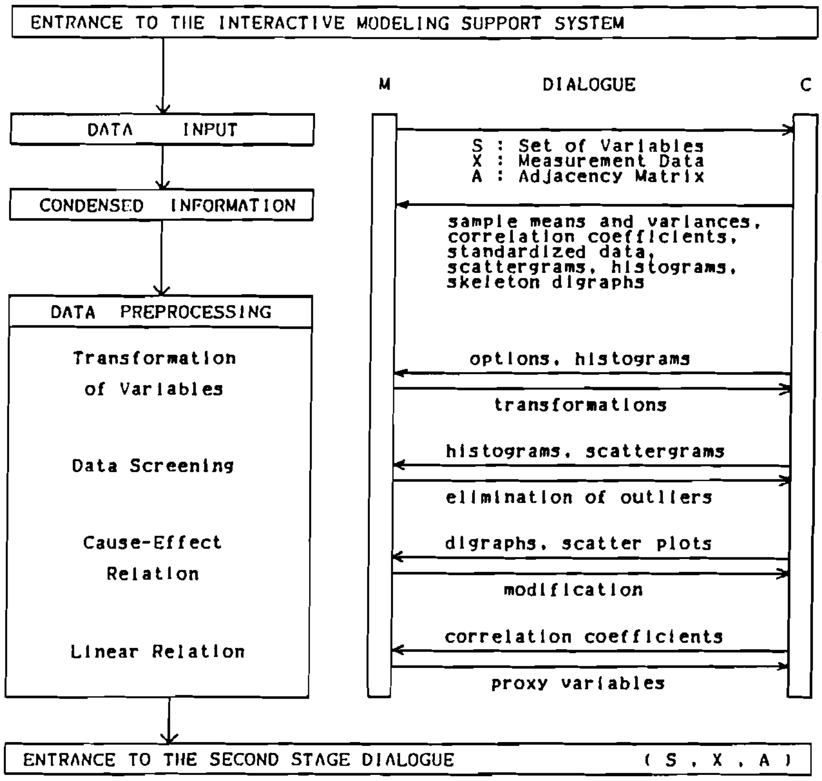


Figure 3: The process of the first stage dialogue of using IMSS.

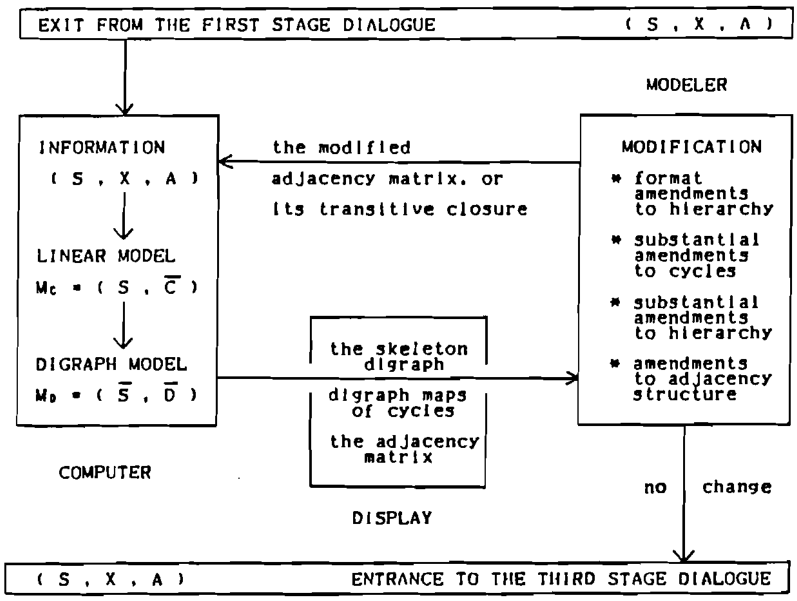


Figure 4: The man-machine interactions in the second stage.

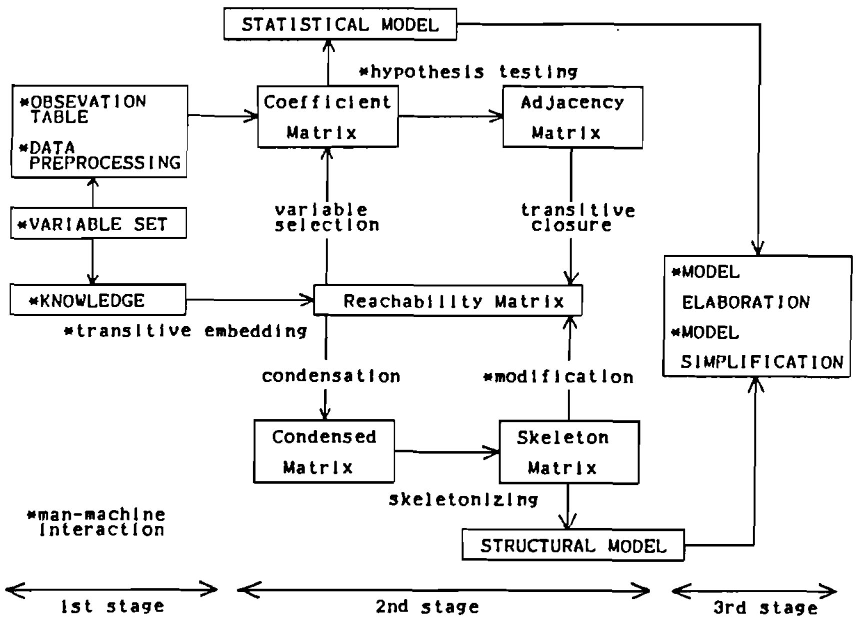


Figure 5: The matrix exchanging process in the second stage.

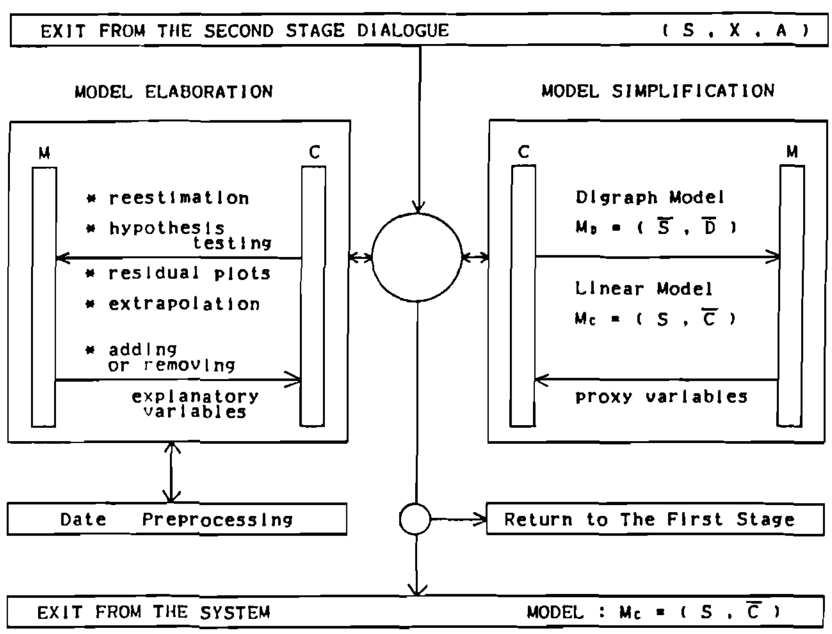


Figure 6: Model elaboration and simplification in the third stage.

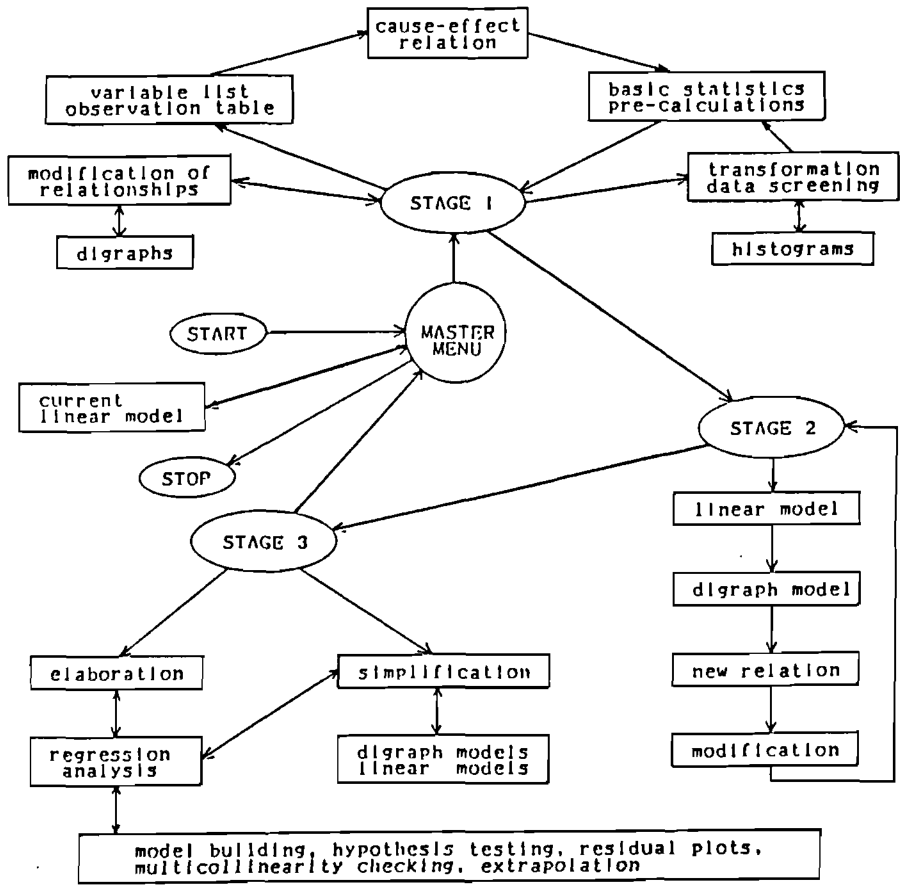


Figure 7: The total modeling process of using IMSS.

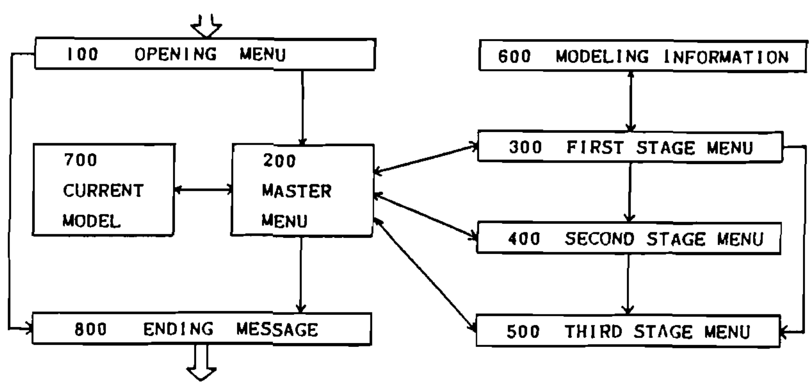


Figure 8: The interrelation between the high level menus.

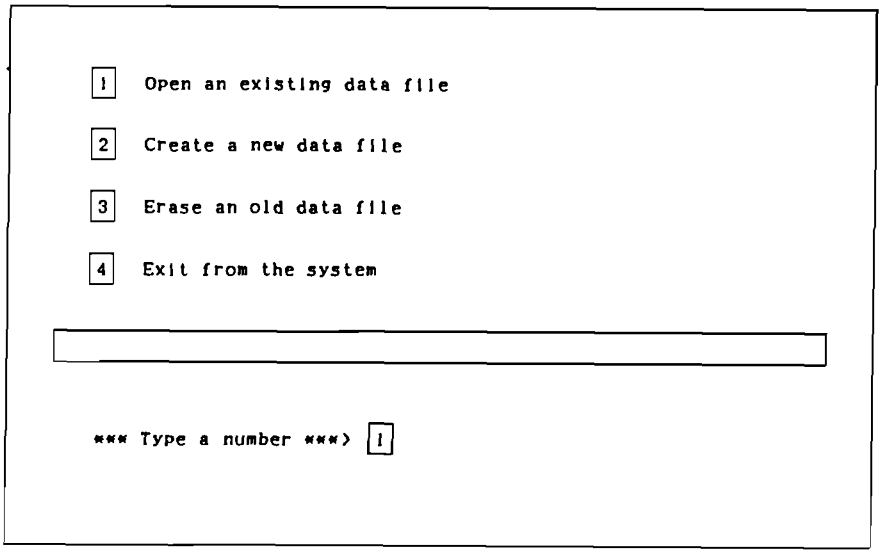


Figure 9: The opening menu for file control.

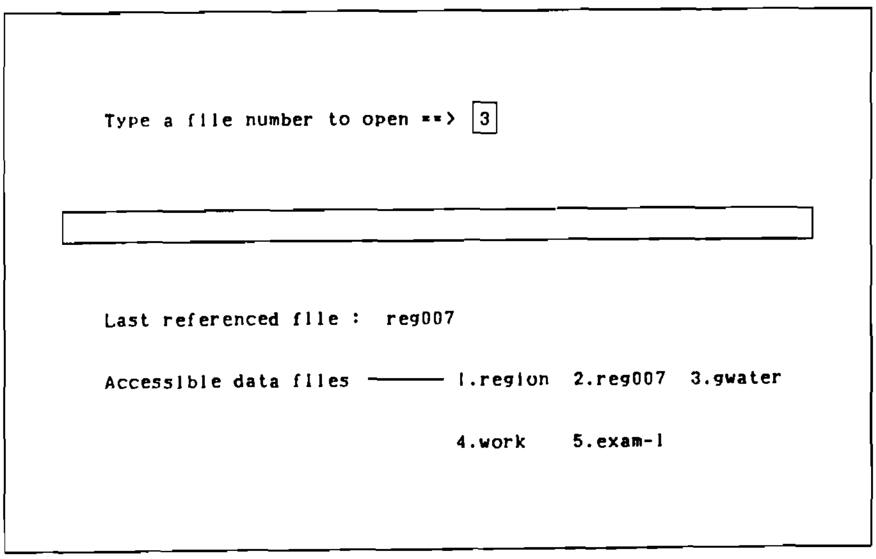


Figure 10: The selection mode to open an existing data file.

AN INTERACTIVE MODELING SUPPORT SYSTEM * data=gwater *

1 FIRST STAGE 2 SECOND STAGE 3 THIRD STAGE

4 CURRENT MODEL 5 TERMINATION NO. ==> [1]

Figure 11: The master menu of the system IMSS.

F I R S T S T A G E D I A L O G U E

1. INITIALIZATION
2. MEASUREMENT DATA
3. CAUSAL RELATION
4. TRANSFORMATION
5. SYSTEM STRUCTURE
6. LINEAR RELATION
7. DATA SCREENING
8. BASIC CALCULATION
I. ▽ INFORMATION ▽
S. ▽ SECOND STAGE ▽

Select a work number ==> [2]

Figure 12: The main menu for the first stage dialogue.

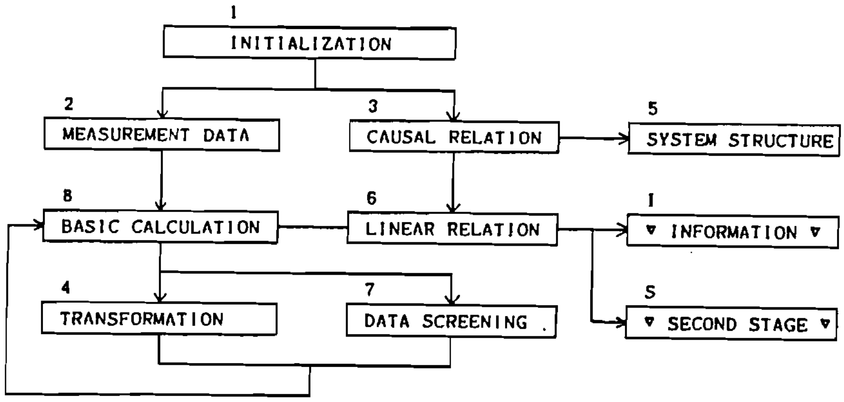


Figure 13: The hierarchical structure of the first stage menus.

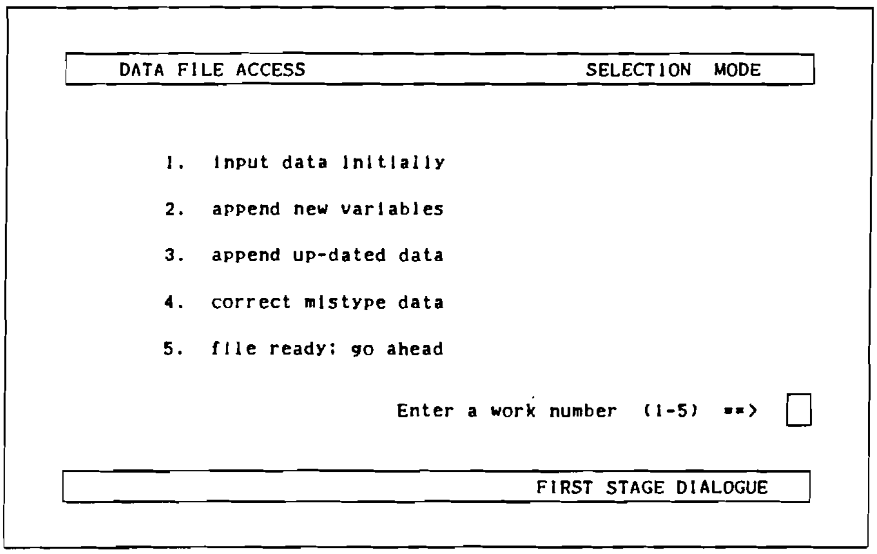


Figure 14: The menu for numerical data input, append or correction.

```
CAUSE-EFFECT RELATION                                   SELECTION MODE

        1. initial input of relation

        2. modification  of relation

        3. cause-effect file ready;  go ahead

                          Enter a work number  (1-3)  ==>  [ ]

                                                FIRST STAGE DIALOGUE
```

Figure 15: The menu for the input of causal information.

```
CAUSE-EFFECT RELATION                                   SELECTION MODE

        1.  Sequential Input

        2.  Transitive Embedding

                          Enter a work number (1-2)  ==>  [ ]

                                                FIRST STAGE DIALOGUE
```

Figure 16: The selection menu for input procedures of causal information.

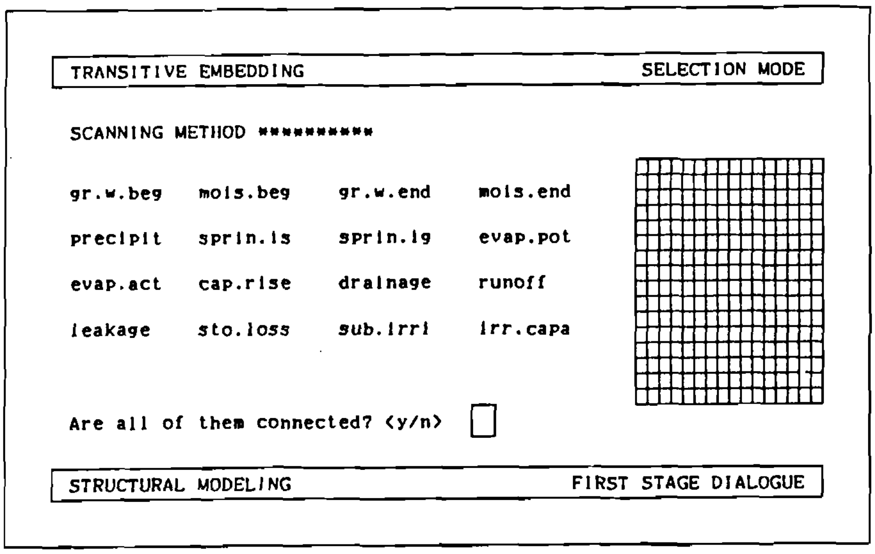


Figure 17: The beginning of the transitive embedding method.

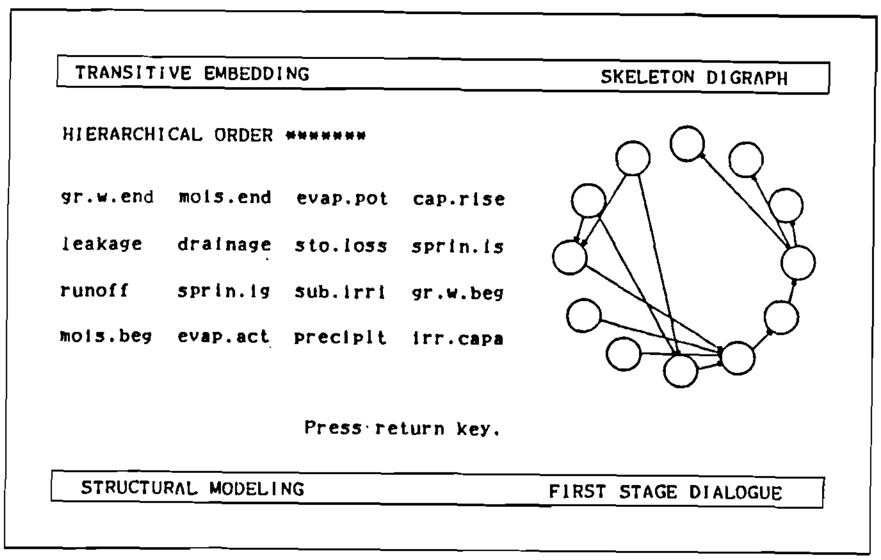


Figure 18: The initial version of the hierarchical relation.

```
TRANSFORMATION OF VARIABLES                    SELECTION MODE

     1. Trials of  Transformations

     2. Actual     Transformations

     3. Recalculation of Hierarchy

     4. Return to  Antecedent Menu

          O  Enter a work number  ==> [ ]

                                 FIRST STAGE DIALOGUE
```

Figure 19: The menu for data transformation.

```
TRANSFORMATION OF VARIABLES                    SELECTION MODE

          O Available Transformation List O

     1.  y=a*log(b*x+c)   2.  y=a*exp(b*x+c)

     3.  y=a*b^x+c        4.  y=log(x/(1-x))

     5.  y=a*x^b+c        6.  y(k)=x(k-1)

     7.  y=a*x+b*z+c      8.  y=a*x*b*z+c

     O   Enter a number (end=0)   ==>  [ ]

TRIAL MODE                       FIRST STAGE DIALOGUE
```

Figure 20: The list of prepared data transformations.

```
▼▼▼ CAUSAL RELATIONSHIP ▼▼▼

X certainly affects Y ==> 2

X   might   affect  Y ==> 1

X   never   affects Y ==> 0

                              correlation ( 1 , 15 ) = 0.9553

                              causality from  1  to  15  ==> 1

                              causality from  15  to  1  ==> 0

HIGH CORRELATION CHECK        Modify [no change : hit return]

                              causality from  1  to  15  ==> □
```

Figure 21: Looking again the cause-effect relation by correlations.

```
OUTLIER CHECKING                              SELECTION MODE

1. Current Status of Outlier List

2. Checking by Standardized Data

3. Histograms or Scattergrams

4. Data Elimination or Restoration

5. Return to Antecedent Menu

          O  Enter a work number  ==> □

                                         FIRST STAGE DIALOGUE
```

Figure 22: The menu for data screening.

Case No.	Variable No.	Standardized Data
9	3	-2.3415
27	3	-2.3415
9	4	2.6571
27	4	2.6571
4	6	2.1681
25	6	2.0781
4	7	2.1668
25	7	2.0811
9	9	2.0009
12	9	2.1441

hit return key

FIRST STAGE DIALOGUE

Figure 23: The list of the candidates of outliers.

STATISTICS WORKING

mean values ************************** completed

variances ************************** completed

standard data ************************** completed

correlations *****************

FIRST STAGE DIALOGUE

Figure 24: Calculation of the basic statistics.

No. (return=7) ==>

INFORMATION SERVICE

ORIGINAL	DATA	DATA
1	2	3
RELATIONSHIPS	ORIGINAL	STANDARDIZED
MEANS	CORELATION	SCATTER
4	5	6
VARIANCES	COEFFICIENTS	DIAGRAMS

Figure 25: The menu for modeling information.

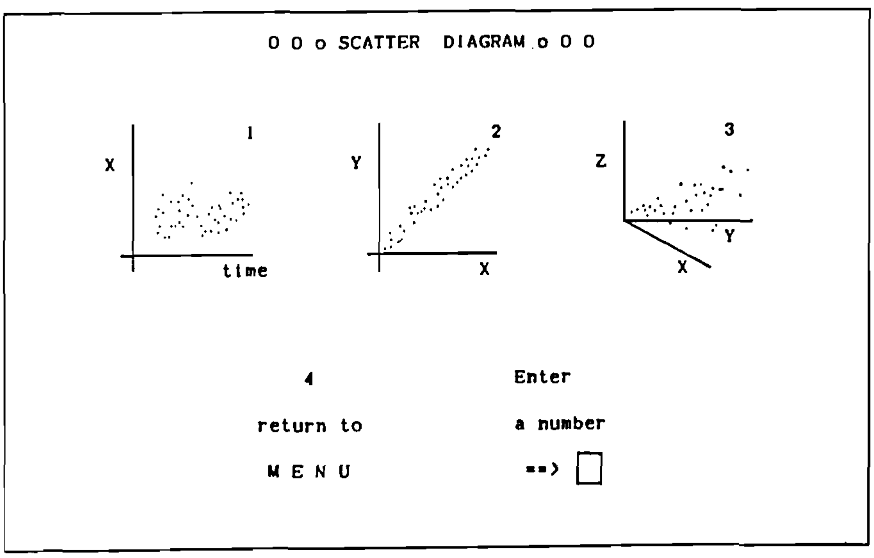


Figure 26: The menu for visualization of numerical data.

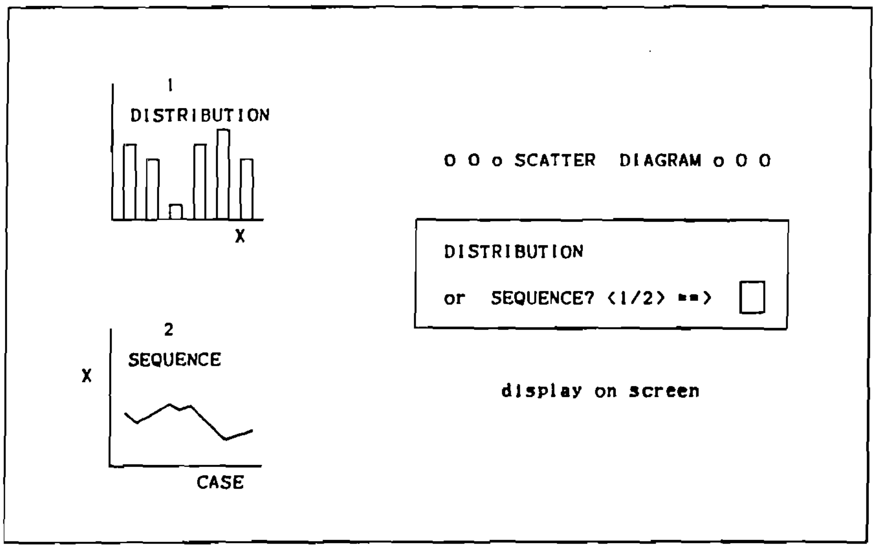


Figure 27: The menu for histograms or scattergrams.

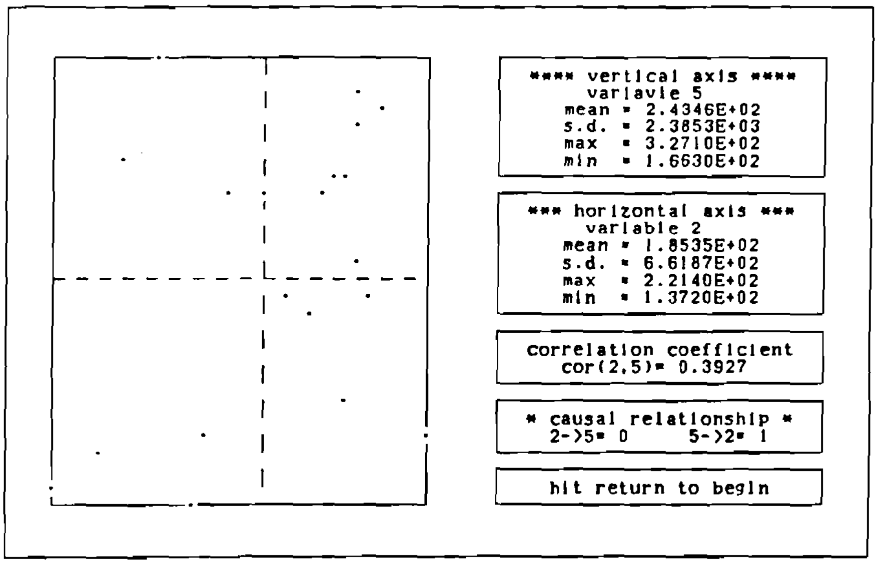


Figure 28: An example of a two-dimensional scatter plot.

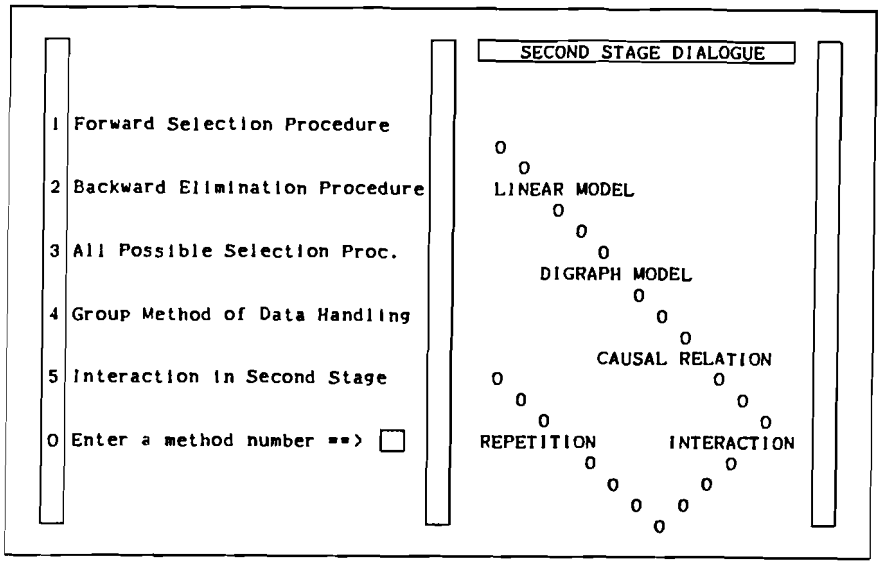


Figure 29: The opening menu of the second stage dialogue.

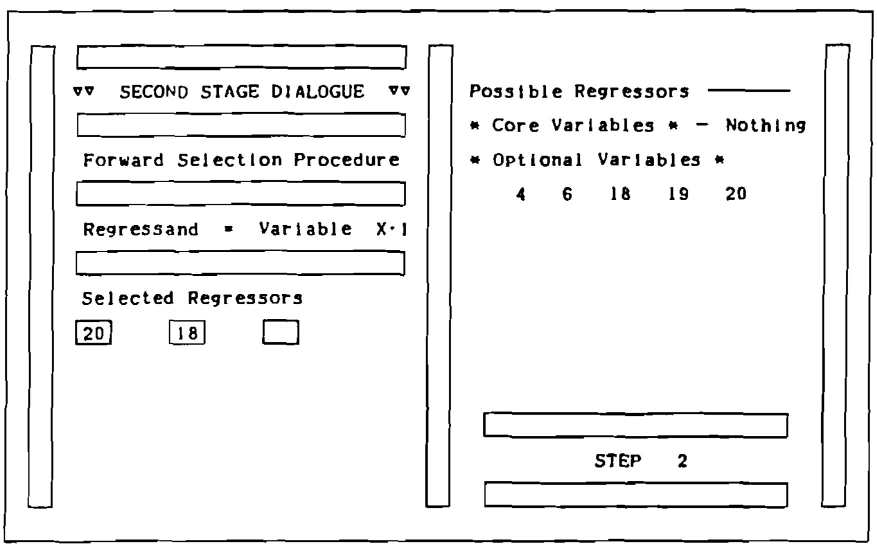


Figure 30: Model building by the forward selection procedure.

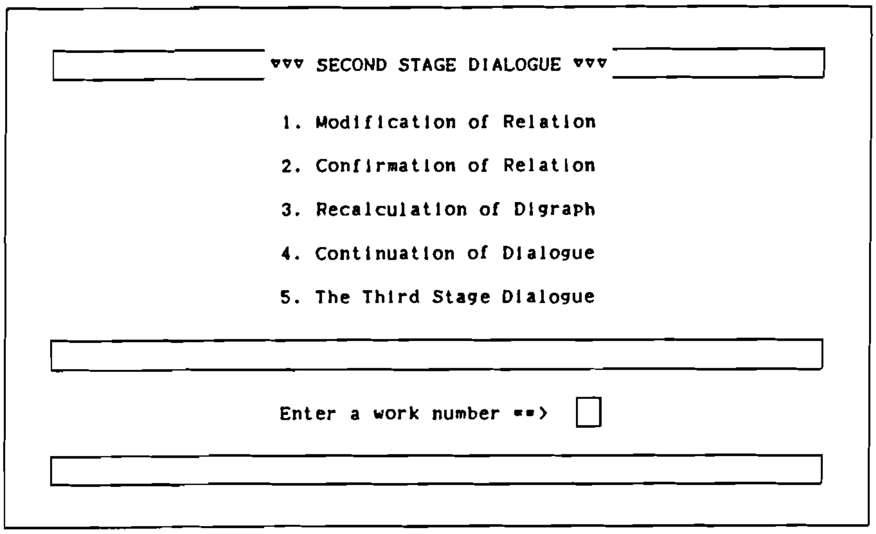


Figure 31: The main menu for the second stage dialogue.

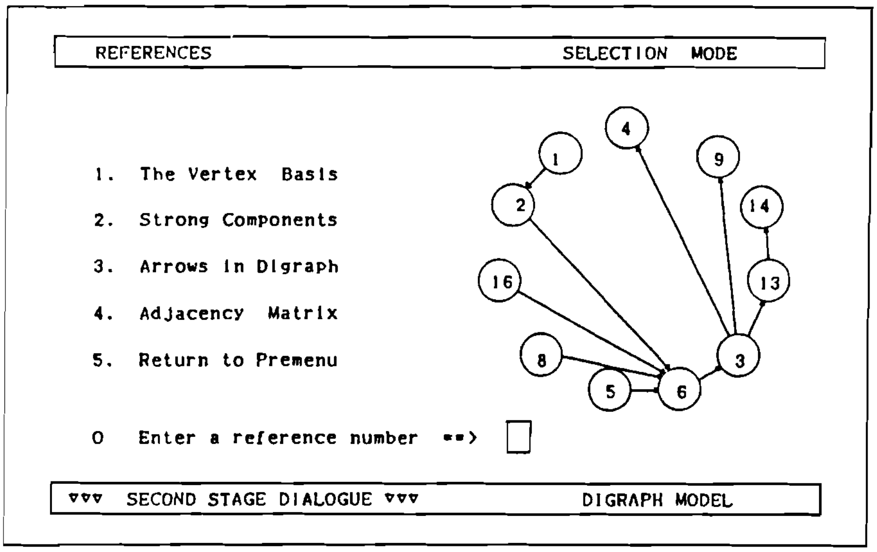


Figure 32: The menu for modification of cause-effect relation.

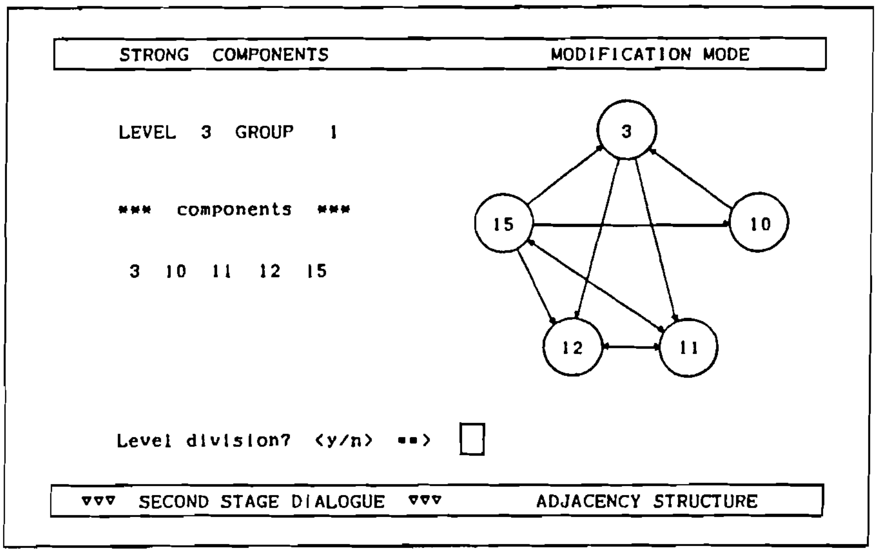
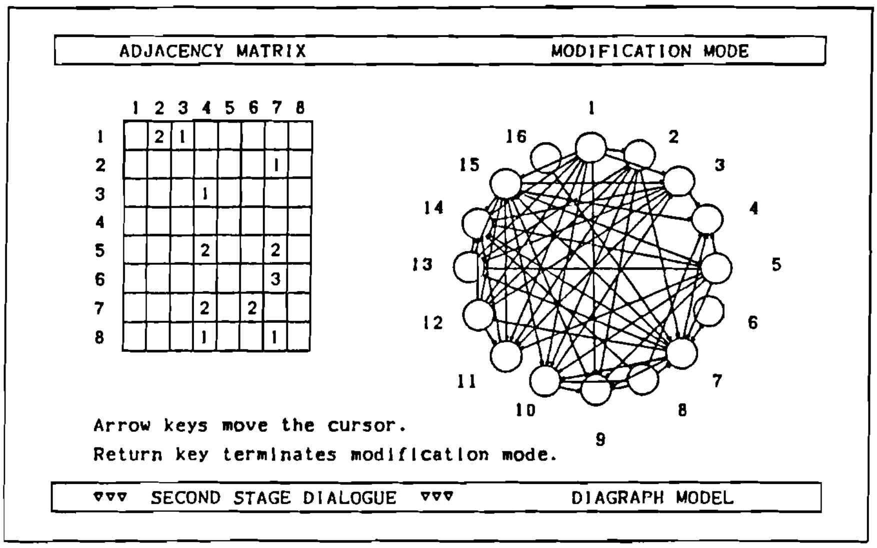


Figure 33: Level division of a cycle set.



	1	2	3	4	5	6	7	8
1		2	1					
2							1	
3				1				
4								
5				2			2	
6							3	
7				2		2		
8				1			1	

Figure 34: Modification of the adjacency matrix.

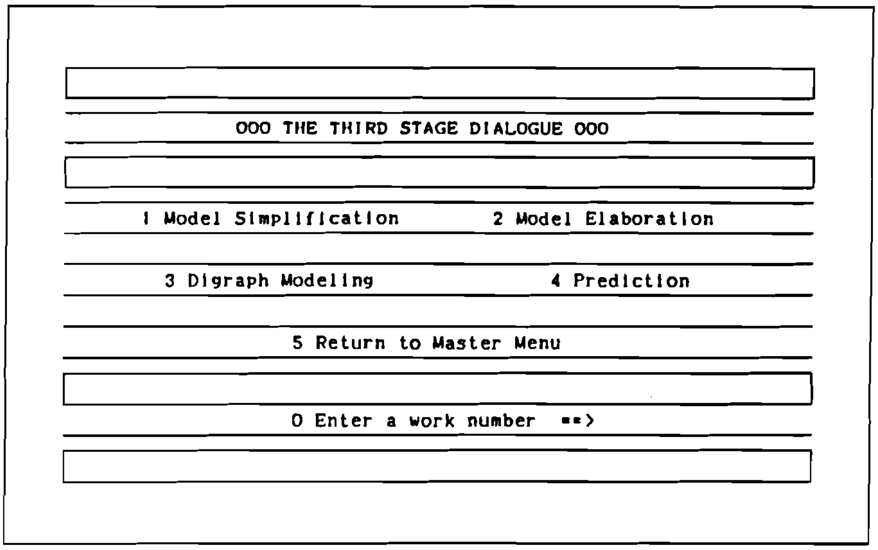


Figure 35: The main menu for the third stage dialogue.

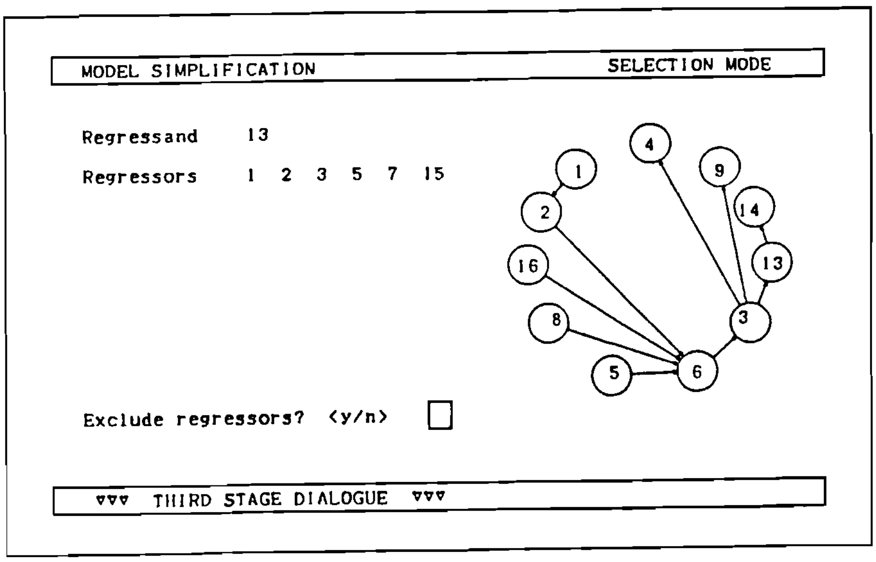


Figure 36: A dialogue for model simplication.

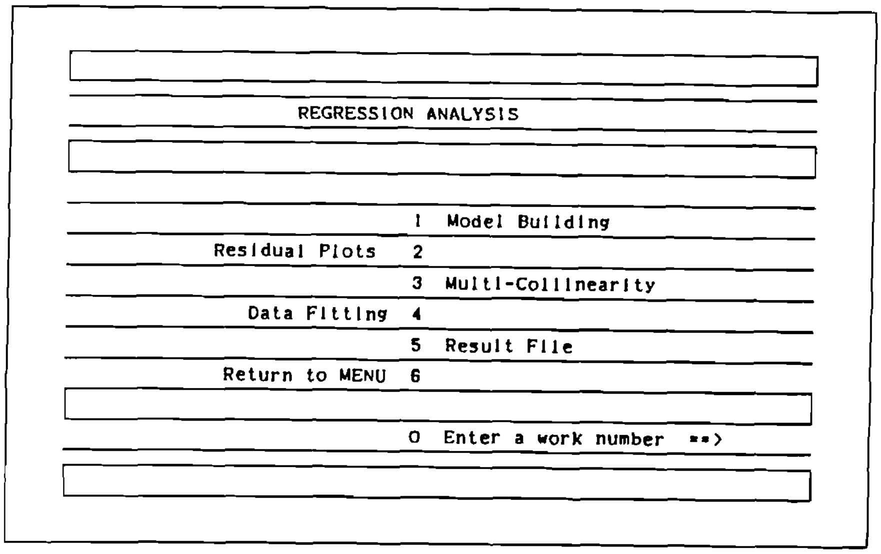


Figure 37: The menu for classical regression analysis.

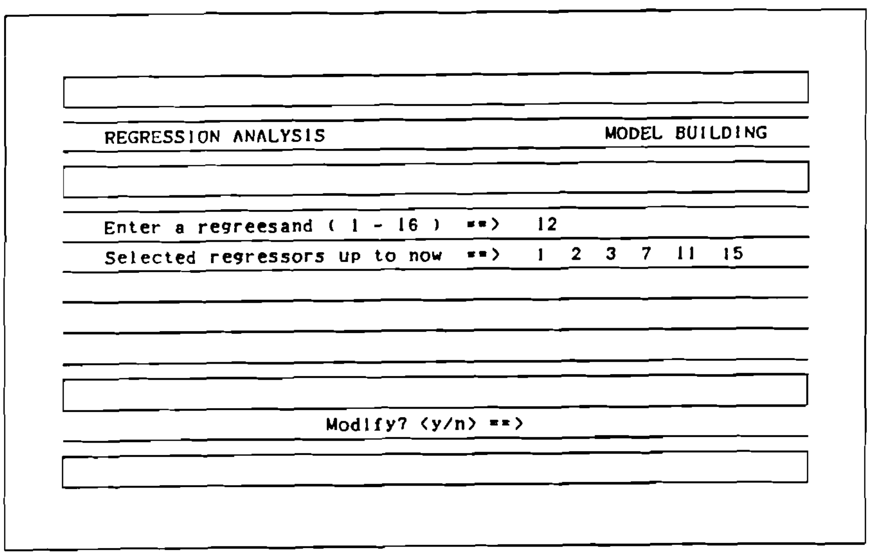


Figure 38: An interaction for regression analysis.

RESULT 2	Regressand ==> variable X12		Ranking 1 / 1	
variable	coefficient	standard error	t-ratio	correlation
X1	-.2194D+01	0.8382D+00	-.2617D+01	0.5314
X2	-.2051D-01	0.6428D-02	-.3191D+01	-.4480
X3	-.1583D+01	0.6150D+00	-.2573D+01	0.4364
X7	0.1009D-01	0.4673D-02	0.2159D+01	0.3787
X11	0.3967D-01	0.5077D-02	0.7815D+01	0.6658
X15	-.6587D-01	0.1732D-01	-.3804D+01	0.4786
constant	0.8956D+01		[hit return]	
Degrees of Freedom = 21		Adjusted R-Square = 0.7893		
S.D. of Residual = 0.2444D+00			F-Ratio = 0.1785d+02	
T(21 , 0.05) = 2.0796		F(6 , 21 , 0.05) = 2.5727		

Figure 39: A result of linear modeling.

```
*** CURRENT LINEAR MODEL *** gwater ***

1  display memorized results

    --- output both on screen and printer ---

2  current model with statistics

    --- output only on printer ---

3  return to master menu                 Type 1, 2 or 3 ==> 
```

Figure 40: The menu for desplaying the current linear model.

```
*** CURRENT LINEAR MODEL *** gwater ***

--- equation for variable  mois.end  ---

mois.end = 2.6751D+02  -1.2739D+02  gr.w.end  -1.4655-02  precipit

                        3.4723D-02  sprin.ig  -6.7754D-02 evap.pot

                        5.8402D-01  sub.irri

                                                    hit return
```

Figure 41: An example of the model equation.