System identification and simulation—
A pattern recognition approach*

by W. J. KARPLUS

University of California
Los Angeles, California

INTRODUCTION
Recent years have seen continuing and increasingly-intensive attempts to extend the art of simulation to areas which heretofore were considered too complex and too difficult to lend themselves to conventional modelling and simulation techniques. These include such environment-oriented fields as air-pollution, water conservation, thermal pollution, etc., as well as systems belonging to the biological, the medical, the economic, and sociological areas. For example, in 1970 the Office of Water Resources Research catalogued over 600 on-going research projects concerned with the modelling of water resource systems. The extension of simulation techniques developed in application areas such as control system design, electro-mechanical systems, etc., to these new areas has often been disappointing, if not completely unsuccessful. This is due to the difficulty in constructing a sufficiently-valid mathematical model—a model which can be used for prediction with a reasonable amount of confidence. It is well-known, of course, that even under the best conditions, inverse problems such as system identification problems, do not have unique solutions. That is, inevitably an infinite number of possible models will satisfy a specified set of excitation/response relationships. Where the identification process is further handicapped by uncertainties as to system structure and inadequate experimental data, the pertinent question is often not: “How good is the model?” but rather: “Is there any point to modelling at all?”

It is the purpose of the present paper to suggest an approach to the derivation and utilization of mathematical models. This approach may be considered to be a generalization and formalization of what has broadly been called “gaining insight into the operation of a system,” but is particularly intended to assist in the formulation of relatively-valid mathematical models for subsequent simulation and in the specification of what additional data (in the form of observations and measurements of the system to be modelled) must be provided to permit meaningful modelling. To put the discussion into proper perspective, the conventional approach to modelling and simulation is first summarized, followed by a discussion of the weaknesses of this method when applied to highly-complex systems. The pattern recognition method is then outlined.

CLASSES OF MODELLING PROBLEMS

A variety of techniques is available for the analysis of systems. Where analytical solution and direct experimentation upon a system are impractical, recourse is often made to simulation. Simulation is a numerical technique which takes the following steps:

1. A system existing in the real world (the prototype system) is represented by a model. This model usually characterizes some interesting facets of the prototype system behavior by a set of equations.
2. The model is implemented or programmed on a computer in such a manner that system inputs, system parameters and perhaps system structure can be conveniently varied and the effects of these variations studied.
3. The computer is employed to perform a variety of experiments so as to provide the information that constitutes the basic objectives of the simulation. These experiments usually involve the prediction of the behavior of the prototype system under various conditions.

The necessary first step in any simulation is, therefore, the formulation and validation of the model. The
discussion of system modelling and the development of a comprehensive modelling philosophy is handicapped by the fact that the types of models used, the raw information available to assist in the development of the model, and the objectives of the eventual simulation depend strongly upon the specific application area and upon subfields within specific areas. The following distinctions are particularly important.

In some applications, the modelling is of the so-called “black-box” variety, in which there is virtually no a priori knowledge as to the nature and structure of the equations characterizing dynamic system behavior. In other applications, the problem is one of modelling a “gray-box” of various shades of gray; here one starts from a knowledge of the general nature of the mathematical model and is concerned primarily with the determination of certain system parameters, initial conditions, and structural details. In some identification problems, the excitation/response data used to identify the gray-box are accurately known or specified and are, therefore, dependable starting points for the modelling process; in other situations, the data to be used to identify systems are obtained by measurements on the system to be modelled and may be seriously corrupted by noise, sampling errors, and a variety of shortcomings in the information gathering effort. In that case, the term “estimation” is often used instead of the term “identification.” Finally, a distinction must be made between systems in which the excitation/response observations are made “actively” and those which utilize “passively” obtained data. In active system identifications, experimental data are collected by subjecting the system to be identified to a series of systematic tests involving the application of specified excitations and observing in each case the resulting response. In passive identifications, on the other hand, the analyst is limited to using data which are generated in response to excitations over which the analyst exercises no control and is, therefore, impeded from constructing key experiments to aid in the identification effort. Clearly the above considerations, that is how black the box, the extent to which excitation/response data are free from noise, and the extent to which active data gathering is possible, greatly affect the modelling process and the reliability and validity of the resulting model.

The present discussion is focused on the modelling of a class of systems which are of interest in a variety of environmental studies and similar large-scale system problems. Consider, for example, the problem of modelling an underground water reservoir or aquifer. The water inputs to the underground porous medium, including rainfall and underground streams, are approximately known, and the underground water level has been measured and recorded at a number of wells which have been drilled into the aquifer. It is desired to obtain a model and perform a simulation to assist in the determination of an optimum control policy to specify how much water can be withdrawn from the various wells so as to maintain the water level in the aquifer at a desired level. It is known that the fluid flow in a porous medium obeys the nonlinear parabolic partial differential equation,

$$\frac{\partial}{\partial x} \left( T(x, y, h) \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left( T(x, y, h) \frac{\partial h}{\partial y} \right) = S(x, y, h) \frac{\partial h}{\partial t} - Q(x, y, t)$$  \(1\)

where

- \(h = h(x, y, t)\) = hydraulic potential at any point \((x, y)\) on the water table at time \(t\) and approximately represents the elevation of the water table above a reference plane.
- \(Q = Q(x, y, t)\) = accretion to the water table due to rainfall, lateral flow, wells, etc.
- \(T = T(x, y, h)\) = transmissivity, which is a measure of the fluid conductivity of the aquifer.
- \(S = S(x, y, h)\) = storage coefficient, which is a measure of the fluid capacity of the aquifer.

The nonlinear parameters \(T(x, y, h)\) and \(S(x, y, h)\) are governed by local soil characteristics and are largely unknown. Likewise, the geometry of the field, the location of the boundaries of the aquifer in the \(x-y\) plane, is only very approximately known from geologic explorations. The only information available to permit an inference of these unknown functions are the well histories (the height of the water measured every few months) taken at a number of wells over a period of years. These represent the raw data. It is the objective of the system identification process to derive a mathematical model including particularly a specification of \(T\) and \(S\) and of the boundary configuration.

In terms of the distinctions briefly discussed above, the modelling effort for problems of this type involves the solution of a “gray-box” problem since it is usually more-or-less accepted that the dynamic processes under study are characterized by nonlinear partial differential equations, including two or three space-variables and time as independent variables, in which the field parameters must be determined by the identification procedure. The response data are noisy, subject to
considerable sampling errors, and never sufficiently complete to satisfy the analyst. Usually, these data are passively-obtained, constituting responses to incompletely-known excitations over which the analyst has no control. The ultimate objective of the modelling effort is to generate a computer model which can be utilized, during the simulation phase, to investigate a variety of hypothetical control situations and which can be used to predict the response of the system to these control strategies. There are, of course, many system identification problems which do not have these characteristics. It is conceded, therefore, that the type of modelling discussed in this paper is directly applicable to only one class of a broad spectrum of modelling problems.

THE CONVENTIONAL MODELLING METHOD

The approach most often used in the construction of models of the type discussed in the preceding section involves the iterative refinement of an assumed model, by comparing the response of the model with the response of the prototype system and by modifying the model so as to minimize the difference between the two. This is illustrated in Figure 1 and discussed in considerable detail by Balakrishnan and Bekey. The following are the major steps in the conventional method:

1. **Formulation**: The basic governing equations and all specific physical information applying to the system under study are assembled, together with all available excitation/response data. The basic equations generally have the vector form

   \[ \phi = f(\phi, u, \alpha_T, t) \]  

   where \( \phi \) is the response vector, \( u \) is the excitation, and \( \alpha_T \) is the system parameter vector.

2. **“Starting” Model**: On the basis of all available evidence and insight, an initial hypothesis as to the model is made. This includes an initial specification of the governing equations, the structure or geometry of the system, and the system parameters. The model equations have the general vector form

   \[ \tilde{\psi} = f(\Psi, u, \alpha_M, t) \]  

   where \( \Psi \) is the response of the model to the excitation \( u \), and \( \alpha_M \) is the model parameter vector.

3. **Implementation**: The equations characterizing the “starting” model are programmed on a computer. The computer model is then subjected to excitations similar to those recorded for the prototype system under study, and the response of the model to these excitations is obtained.

4. **Criterion**: A criterion function is specified to serve as a measure of the extent to which the response \( \tilde{\psi} \) of the model conforms to the response \( \phi \) of the prototype system being modelled. Usually this criterion function is defined by an expression of the type

   \[ J(T, \alpha_M) = \int_0^T (\phi - \tilde{\psi})' W (\phi - \tilde{\psi}) \, dt \]  

   where \( W = W(\phi, \psi, t) \) is a suitable weighting function, and \( T \) is the time interval over which the identification takes place. This criterion function \( J(T, \alpha_M) \) is calculated from the system and model responses. That is, the response of the model and the response of the system are compared.

5. **Decision**: The objective of the identification procedure is to seek an optimum set of parameters \( \alpha_M \) which minimize the criterion function such that

   \[ \min_{\alpha_M} J(T, \alpha_M) \]  

   The criterion function calculated in step 4 is, therefore, examined to see if it exceeds a specified minimum \( \epsilon \). If it does not, the identifica-
tion is complete, the model is considered valid and employed for simulation. If the criterion function is not sufficiently small, the model must be modified.

6. Modification: A computational procedure, usually in the form of algorithms, is specified. This routine defines the manner in which the model parameters $\alpha_M$ are to be modified after each iteration, and it may involve gradient methods, random search, relaxation, etc. In any event, it acts to change the parameters of the model hopefully in a manner which results in a smaller $J(T, \alpha_M)$.

The conventional method of modelling is effective for the identification of systems which are "well-behaved." In particular it works well in situations in which the initial guess as to the model is very close to the prototype system, and where the excitation/response data are of very high quality. The method breaks down in many practical applications, however, for two principal reasons: 1. The first hypothetical model is not a sufficiently-close representation of a prototype system, and 2. The excitation/response data available from observations of the prototype system are of such low quality that the attaining of a minimum in the criterion function cannot be taken with confidence as an indication of the validity of the model.

To illustrate the weakness of conventional modelling, consider again the underground water resource modelling problem discussed above. Whereas Equation (1) can be assumed to apply reasonably well throughout the aquifer, the geometry of the field (boundaries of the porous medium), the initial conditions, the excitations, as well as the presence of major inhomogeneities are only incompletely known. The first hypothetical model is, therefore, likely to be substantially different from the actual system. The system response data which are to be used to improve this initial guess are represented by well-logs at haphazardly-spaced points in the field and constitute measurements of the dependent variable sampled at insufficiently-frequent intervals and subjected to serious measurement errors. Nonetheless, the conventional modelling approach requires the iterative refinement of the initial model until a criterion function of the type of Equation (4) is minimized; that is, until the transient response curves of the model are fitted closely enough to the field data. As a result, even if after much laborious and elegant computation, one arrives at a model which provides a tolerable match of the field data, there remains considerable uncertainty as to the meaningfulness of the model and its usefulness in subsequent simulations. This unfortunate consideration applies to models and simulations in a wide variety of important areas of application.

THE PATTERN RECOGNITION APPROACH

The approach to modelling suggested in the present paper is based upon the following premise: the excitation/response data available from experiments or observations of a prototype system contain a large amount of potentially-valuable and useful information which is not adequately utilized in the conventional approach to modelling. In the attempt to employ curve- or data-fitting methods to match the responses of a dubious model to highly error-prone experimental observations, many key features inherent in the experimental data are averaged out, overshadowed, or simply not utilized. A reason for this lies in the application of the criterion function, such as Equation (4), during each iterative cycle, which involves an attempt to compare the "artificial" responses of the model with the "real-world" responses of the system at each stage of the modelling process.

The pattern recognition approach to modelling is similar in some respects to that employed by Duda and others in recognizing and classifying handwritten characters. In that method, the pattern recognition problem is viewed as a sequence of four mappings as shown in Figure 2. The handwritten characters themselves constitute a so-called "object space" ($x$). By means of a video scan of the characters, followed by sampling and digitizing, the object space is mapped into a "representation space" ($y$) consisting of a sequence of binary numbers. Algorithms are then developed to map from the "representation space" into a "feature space" ($z$). This mapping, termed feature extraction, involves ignoring most of the available samples and the focusing of attention on a few key sampled values which are sufficient to distinguish the characters from each other. Finally, there follows a mapping from the "feature space" to the "decision space" ($d$), a classification operation in which the features that have been extracted are used to decide the identity of a character under examination.

![Figure 2-Successive mappings in pattern recognition](From the collection of the Computer History Museum (www.computerhistory.org))
In character recognition, no attempt is made to devise a criterion function of the type of Equation (4) in order to identify characters. Rather the one-dimensional sequence of video signals is subjected to feature extraction, such that a small number of video samples are examined to determine whether they are black or white. The decision as to whether a given character is or is not the letter A, for example, is made on the basis of whether these key characters, sometimes termed the "mask," are of the correct combination of black and white. This mask is developed by postulating a "starting" mask and by working with a "learning set." The learning set is a collection of handwritten characters obtained from representative collections of manuscripts. The "starting mask" and a decision algorithm are then used to examine the learning set, and the success or failure of the character identification is recorded. The mask and the decision algorithm are then modified and applied to the same sequence of characters. This is repeated for many different masks and decision algorithms. The mask and algorithm which manifest the best record of success are adopted as the pattern recognition algorithm, and are then applied to unknown characters as required.

The pattern recognition method of modelling has the same starting point as the conventional approach. Experimental system data (input and output measurements) are assembled, and a first hypothetical model is formulated and implemented on the computer. At this point, the two approaches part company. In the pattern recognition approach, the model implemented on the computer is not regarded primarily as something to be iteratively matched to reality (the system outputs). Rather it is considered as a "learning machine" to develop feature extraction and classification algorithms which will eventually serve to extract pertinent information from the "real world" system data. The primary objective of this first stage of modelling is not to progressively refine the model, but rather to develop a set of computing routines which can subsequently be utilized to analyze the data available from the system to be modelled. The results of this analysis then are used to formulate the "starting" model for conventional parameter identification. As shown in Figure 3, the modelling problem is therefore subdivided into two stages: pattern recognition and parameter identification.

The term "pattern" is used in the present context to connote general or global characteristics of the system being modelled. For any specific modelling problem, these patterns must be known in order to talk meaningfully of parameters and their identification. Accordingly, a list of patterns is prepared, and the nature of these patterns is to be extracted from available system observations (excitations/response data). Where possible, these patterns are formulated in such a manner that their recognition involves the answering of a yes/no question. For example, in the case of a distributed system such as that described by Equation (1), these questions might include:

1. Is a given parameter (for example $S$) present in non-negligible quantities? That is, is it necessary to include that parameter in the model?
2. Is this parameter constant, in the range of dependent and independent variables for which system observations are available?
3. Is this parameter a function of the independent space variables, $x$ and $y$?
4. Is this parameter a function of time?
5. Is this parameter a function of the dependent variable (nonlinear)?
6. Does the magnitude of this parameter everywhere fall within a specified range?
7. Considering the quality of available system observation data (number of measuring stations, sampling interval in time, and measurement...
errors) is it possible to derive a model of a given dimensionality? That is, do available response data permit the meaningful construction of a finite difference grid of a specified truncation interval?

Similar questions can be asked regarding the geometry of the system, that is the location of field boundaries, and even the general structure of the basic equations. Usually in conventional modelling, all patterns of the type listed above are assumed initially, and a basic error in these assumptions invalidates all subsequent modelling efforts. In the pattern recognition method, a set of algorithms is developed with the express purpose of extracting the answers to these questions from available prototype system observations. The computer model is used to develop these computing routines.

Each algorithm is designed to accept, as its input, the response data of the model and eventually response data of the system being modelled. The model is designed to provide data having the same sampling interval, spatial distribution, and measurement noise as the original system. The output of each algorithm is the answer to one question of the type posed above. Each algorithm is, therefore, a separate pattern recognition routine. This routine can conceivably involve transformations or spectral analysis, or it may involve cross-correlations of response data taken at different points in space, but will more often take the form of a “mask.” Instead of processing all the samples obtained from all response functions, attention is focused on a few key sampled values. The yes/no decision is based upon the information contained in these samples. The optimum mask, that is the combination of samples which are processed to determine whether the answer to a question is “yes” or “no,” is determined experimentally using the computer model.

The pattern recognition algorithms used in modelling are developed in a manner basically similar to that used in character recognition. A “starting” algorithm is adopted either from experience or from heuristic considerations. This algorithm is tried out on the model response transients, where these response functions are generated by exciting the computer model with excitations similar to those which excited the prototype system. The algorithm also acts to “perturb” the model, so that the effectiveness of the algorithm over a number of similar yet different model configurations or parameter distributions (patterns) is determined. The algorithm is then modified automatically or by an on-line operator and the process repeated. After a number of such experiments, that algorithm which proved most effective in identifying the desired pattern is selected. The same procedure is followed to obtain successful algorithms for identifying all of the other patterns of interest, so that eventually a library of algorithms is formed—algorithms which are tailor-made for the system being modelled and for the specific excitations and responses which are available from the physical system.

Once this library is complete, attention is turned, for the first time, to the response data of the physical system. These data are now processed by the algorithms that were just developed. That is, the pattern recognition algorithms are employed to determine the patterns of the physical system. This process may demonstrate that the model used for “learning” differs radically from the system being modelled. Accordingly, this model is modified so as to give it the patterns which were found to be contained in the system being modelled. This whole process is repeated until the pattern extracted from the physical system corresponds reasonably closely to those assumed for the system being modelled. At that point, one can conclude that the model is “within the correct ball park,” and the conventional parameter identification method can be employed to determine the fine structure of the model.

The general approach is illustrated in Figure 4 and takes the following steps:

1. **Formulation:** The basic governing equations and all specific information applying to the system under study are assembled, together with all available excitation/response data. The basic equations take the vector form

   \[ \dot{\phi} = f(\phi, u, \alpha, \beta, t) \]  

   where \( \beta \) is a vector of patterns.

2. **“Starting” Model:** On the basis of all available evidence and insight, an initial hypothesis as to the model is made. The model equations have the general form

   \[ \dot{\psi} = f(\psi, u, \alpha_M, \beta_M, t) \]  

   where \( \alpha_M \) is the model parameter vector and \( \beta_M \) is the model pattern vector.

3. **Implementation:** The equations characterizing the “starting” model are programmed on a computer. Provision is made in this implementation for perturbing or modifying the patterns of the model under control of the pattern recognition (P.R.) algorithms. The model is to accept as input data the observations of the excitation, \( u \), of the system being modelled. The model response, \( \psi \), is given as nearly as possible the same characteristics as the system response, \( \phi \). That is, response data are read out from the same...
locations as those at which system response data is available, a similar sampling interval is employed, and if appropriate, noise is artificially added to the model output.

4. "Starting" Pattern Recognition Algorithm: On the basis of previous experience and insight, a separate algorithm is provided for each of the patterns, \( \beta_M \) to be recognized. These algorithms may include masks for selecting key samples for further processing.

5. P.R. Algorithm Implementation: The "starting" algorithms are programmed on the computer. These algorithms may contain loops which act to perturb or modify the patterns of the computer model so as to test the algorithms under a number of different situations. For example, if the purpose of the algorithm is to determine whether a given parameter is constant or not, that parameter is given a number of different constant values as well as caused to vary in a prescribed fashion. The modified patterns imposed by the P.R. algorithm are denoted by \( \beta_M^* \).

6. P.R. of Model Response: The algorithm is employed to process the model response, \( \psi \), and to recognize the model patterns for each of the model perturbations. The patterns recognized by the algorithm are denoted by \( \beta_M^* \).

7. Comparison: For each pattern recognition run, the success or failure of the algorithm is determined by comparing the pattern of the model, \( \beta_M^* \), with that determined by the algorithm, \( \beta_M^* \).

8. Criterion: A figure of merit for each algorithm is determined by totaling the successes and failures of the algorithm over all the experiments conducted with that algorithm.

9. Decision: A decision is made as to whether or not additional modifications of the P.R. algorithms should be attempted.

10. Algorithm Modification: Either automatically or with the aid of an on-line operator, the P.R. algorithm is modified. This modification may involve the re-specification of the mask, a change in the manner in which the samples are processed, or it may involve a more fundamental change in strategy. Evidently, the specific nature of this modification depends upon the patterns to be recognized by the algorithms. In any event, steps 5 through 9 are repeated until no additional modifications are required.

11. Selection: Provided no additional algorithm modifications are required, that algorithm having the best percentage of success is selected and stored.

12. P.R. of System Response: The selected algorithms are now employed to process the system response, \( \phi \), obtained from prototype system observations. That is, the algorithms are employed to recognize the patterns, \( \beta_\phi \), in \( \phi \).

13. Comparison: The patterns \( \beta_\phi \), recognized using the system observations are compared with the patterns \( \beta_M^* \) initially assumed for the model. That is, it is verified whether the model used for algorithm development was "in the correct ballpark."

14. Decision: The results of the comparison of all the members of the pattern vectors \( \beta_M^* \) and \( \beta_\phi \) are analyzed to determine whether the "starting" model was close enough to the system being observed. If agreement between the two is adequate, that is, if the model has most or all of

![Figure 4-The Pattern Recognition (P.R.) modelling method](From the collection of the Computer History Museum (www.computerhistory.org))
the patterns of the physical system, the pattern recognition process is considered complete, and the computer model can be employed as the starting point for conventional parameter identification and eventually for simulation.

15. **Model Modification**: If agreement between the model and the physical system is inadequate, the computer model is modified by giving it the patterns determined in step 12. Steps 5 to 14 are then repeated until adequate agreement is obtained.

The most difficult steps in this method are the selection of the “starting” algorithm for pattern recognition and classification and the specification of the modification strategy of this algorithm. These depend strongly upon the type of patterns to be recognized, upon the computer model, and upon the nature of the response data. It is necessary, therefore, to build up a considerable amount of experience with this method for any specific application area. Occasionally it may turn out that a proven algorithm modification strategy does not lead to adequate convergence for a specific problem. This may then be taken as an indication that the quality of available response data is insufficient to permit meaningful pattern recognition. For example, the time sampling-interval may be too large, or response data may not be available for enough points in the space domain, or the signal-to-noise ratio may be too low. Under these circumstances, the computer model and the pattern recognition method can be employed to determine the approximate extent to which system observation data must be improved to make modelling possible. This can be accomplished by gradually improving the quality of the computer model response (by sampling it more frequently, for example) until the algorithm modification strategy leads to successful convergence. The results of this computer experiment are then used as the basis for better and more complete field measurements.

**CONCLUSIONS**

The pattern recognition method described in this paper is evidently not a panacea. The procedure is useful only for the identification of systems of “a certain shade of gray,” and it leans heavily upon the ingenuity and insight of the analyst. It does, however, constitute a novel utilization of computer models—the development of a “learning set” and the determination as to whether the system response data are of sufficient quality to permit parameter identification. The approach has been used with some success in the modelling of underground water reservoirs of the type characterized by Equation (1) as well as in the study of aquifer pollution problems. The results of these studies will be reported in separate papers.

**REFERENCES**

1. W J KARPLUS, V VEMURI
   *Heuristic optimization and identification in hybrid field simulations*
   Proc Fifth Int Congress of AICA Lausanne Switzerland pp 345-350 1967

2. V VEMURI, W J KARPLUS
   *Identification of nonlinear parameters of ground water basins by hybrid computation*
   Water Resources Research Vol 5 pp 172-185 1969

3. A V BALAKRISHNAN, V PETERKA
   *Identification in automatic control systems*
   Automation Vol 5 pp 817-829 Pergamon Press 1969

4. G A BEKEY
   *System identification—an introduction and a survey*
   Simulation Vol 15 pp 151-166 1970

5. R O DUDA
   *Elements of pattern recognition*