GUIDELINES AND TOOLS FOR UNDERSTANDING DYNAMIC MODELS

by

Wil Thissen
Project Globale Dynamica,
Systems and Control Engineering Group,
Eindhoven University of Technology
Eindhoven Netherlands

ABSTRACT

Starting from the aims and difficulties of social systems modeling this paper argues that a good understanding of dynamic mathematical models is indispensible. The author's background, and its relation to System Dynamics is elucidated, and a number of definitions are given of concepts and terms that will be employed. A set of general guidelines, and a list of strategies and tools for understanding follow. Most of the methods presented have been applied successfully in an extensive study of the World Models by Forrester and Meadows et al., and are commonly used in systems and control engineering. The main emphasis is on techniques and points of view that are generally unknown to researchers and practicians in the non--technical disciplines.

CONTENTS

I. INTRODUCTION

One of the main purposes of social systems modeling is to help gain insight into the working of some real system. The investigator's view of the main components of the system and their interactions is described in terms of a set of mathematical equations. While important insights are generated in the phase of model conceptualization, a further step in the direction of insight into the corresponding real system is to study the structure and assumptions of the model in an attempt to understand the causes of its behaviour. Moreover, insight into the working of a preliminary version of a model may be of great help to the modeler: it may attract attention to model parts that need further improvement, or even may lead to a total redesign. Once a model is considered to be completed, a thorough understanding of its inner working is indispensible as a convincing basis for formulating policy starts, and to determine the best approach for possible further study. Moreover, in the uncertain environment of social systems, qualitative insights about behaviour modes, sensitivities, etc. are often more important and robust than quantitative results. Such conclusions, based on a thorough understanding of a model's working can be explained clearly and in simple terms, and offer possibilities for an adequate and convincing communication of the results and conclusions of a modeling project to the clients as well as to the public. Finally, knowledge about the mechanisms governing a model's behaviour can be used for judging the validity of the model.

For these and other reasons, understanding models is an issue important to all modeling studies. In general terms, a model is understood if its results and that of the whole study can be expressed in words and/or simple diagrams*, and be made quite reasonable to anyone. More specifically, one should be able to answer three types of questions about a model's behaviour correctly without performing simulations, namely:
- What will happen if ....(a certain assumption is made)?
- Under what conditions will .... (a specified behaviour take place)?

-----

* It might even be argued that simple patterns or diagrams are superior to verbal description when a model's working has to be explained. Diagrams are able to transfer the pattern of simultaneous interactions in a dynamic model directly to the human mind. Since language is essentially sequential, verbal description cannot do this.

- Why does the model behave the way it does?

The first two questions are closely related, and require a general understanding of the dynamics included in the main components of the system, and of the interactions between these components. The latter question asks for explanation of the general dynamic properties in terms of the assumptions underlying the equations.

Whereas a good understanding of the working of complicated dynamic models is essential, its acquisition remains a difficult and energy-consuming task. At present, model understanding seems to be an underdeveloped part of systems analysis. The lack of effective, systematic, and preferably simple tools that help gain insight is serious. This paper attempts to meet in part the need for such techniques by presenting a list of guidelines that came up and tools that were found to be of use during an extensive analysis of the World2 |3| and World3 |6| models. The paper is a continuation and partial extension of an article by Rademaker |9|, upon which it has drawn freely. For the most part, the guidelines and tools are well-known and used frequently, particularly in control- and systems-engineering circles. In Chapter II, some background information will be given, mainly with respect to systems and control engineering and its relation to System Dynamics, and with respect tot the aims and goals of the 'Global Dynamics' project group. Chapter III presents a list of guidelines. A concise description of a number of tools and techniques that may be of use for understanding dynamic models follows in Chapter IV. Since some of the techniques start from views that are not familiar to all system dynamicists, part of Chapter II is devoted to a description of a few underlying concepts and definitions.

The major part of this paper discusses the application and use of various concepts and tools. Whether an investigator will appreciate a technique as being appropriate and easy to apply is mainly determined by his background and experience, and by his personal preferences. Any judgement about the importance, utility, appropriateness, simpleness, energy costs, etc. of an approach contains a subjective element, particularly if the same results can be obtained along different ways. Clearly, the reader should keep in mind that this also applies to the judgements put forward in this paper.

II. BACKGROUND

II.1 *SYSTEMS AND CONTROL ENGINEERING, AND SYSTEM DYNAMICS*

The quick progression made in the field of aviation, electronics and the area of chemical processes was, at least in part, due to the fast evolution in the field of systems and control engineering in the years before and during World War II. Conversely, the need for practical and feasible solutions to actual problems was stimulating the development of what, nowadays, is called "classical" in control theory. These prior analyses and theories dealt almost exclusively with linear systems, and were heavily based on impiricism and trial and error design, and characterised by the lack of a fundamental theory. However, problems in chemical and aerospace engineering grew more and more complex, and a more fundamental approach was urgently required. "Modern" control theory was born in the late fifties. Rather than centering around frequency-domain techniques and feedback loops, as the classical approach, modern control theory focusses around general time-domain descriptions of dynamic systems. But, developed and presented mainly by applied and pure mathematicians, it is quite inaccessible to most control engineers and others who are interested. Although modern control theory has unmeasurable qualitites, in many practical (and simple) problems the need for the old art has not been eliminated yet.

Many fundamental ideas of System Dynamics as formulated by Forrester in Industrial Dynamics |2| (particularly the emphasis on feedback loops) are based on the classical approach in control theory. But there are many differences, not only paradigmatic, or in the nature of the research objects, but also in the technical field, such as the emphasis on nonlinearity, the time-domain representation, and the abundant use of computer simulation.

Still, System Dynamics and systems and control theory deal with similar problems. Practicians in the two fields attempt to analyse and understand dynamic systems to be able to improve real system behaviour. Also, in both disciplines there is a tendency to tackle more and more complex problems. Increasing complexity was one of the incentives to

adopt a more general and fundamental approach in control theory, whereas - except for the extension of computer-simulation facilities - no further development of tools for solving the complexity problem seems to have been realized in the system-dynamics field. Therefore, System Dynamics may grow up to a higher degree of maturity if it would adopt and/or adapt ideas and techniques developed in modern (and classical) control theory. Certainly, not all concepts will be useful, because System Dynamics focusses on a rather different field of application, but particularly in problem areas of high complexity, where a systematic approach is required, such as parameter estimation, model analysis and understanding, and policy analysis, control theory offers a number of powerful tools which probably can be adapted to System Dynamics without too much trouble. A few examples have already been given by Peterson |7| (parameter estimation) and Sharp |12| (systematic sensitivity analysis). Also in this paper it will - among other things - be tried to demonstrate the utility of certain engineering approaches (state-space concept, total linearisation).

II.2 *THE PROJECT: 'GLOBAL DYNAMICS'*

This paper is one of several outcomes of the project 'Global Dynamics' (see also ref. |8|). The project was started in the course of 1972, after the appearance of the first publications on World models by Forrester |2|, and Meadows et al. |5|. In these publications it was - much more than in most econometric and macro-economic models - focussed on the dynamic properties of systems, including levels and feed-back loops, and causality was emphasized. This might be one of the reasons that the attention of a number of system and control engineers was drawn. Another reason, however, was that there was a strong feeling that only little, if any, of the existing knowledge about and experience with the analysis and control of mathematical models of dynamic systems was used. Therefore, among other reasons, a project was started, the main aims and goals were <u>not</u> to build new models or to criticize the assumptions made by the M.I.T.-groups, but:

- to analyse the models from the control- and systems-science point of view, in an attempt to gain a thorough understanding of their inner working and structure.

- to examine the effects of various kinds of control, particularly (stabilizing) feedback control and optimizing control (dynamic optimization).
- to communicate the information gathered and the insights obtained to an audience as wide as possible, and to develop and compile an array of techniques for the analysis and control of dynamic models.

This paper forms part of an attempt to fulfill the last goal, but it cannot be seen as the result of a separate study. The activities of the group were centered around the study of the World2 and World3 models, and only afterwards the various methods and techniques used were listed, and their general usefulness for understanding was judged. This explains why the application of most of the tools described in Chapter IV will be illustrated using examples of the World models.

II.3 CONCEPTS AND DEFINITIONS

This section will start with a description of the state-variable point of view and of its advantages when looking at a complex dynamic system. The state-variable concept underlies the greater part of modern systems and control theory, and also many of the tools for understanding that are presented in this paper.

The second part of this section will be devoted to the clarification of a number of terms that are used frequently, such as 'time constant', 'model structure', and 'stability'. It is observed that different persons (or the same persons at different times) may adhere different meanings to each of these terms. To avoid misinterpretations, it will be tried to make a clear distinction between the various meanings.

II.3.1 State-space representation

In classical control theory, the feedback loop plays a central part as a tool for influencing (mostly: stabilizing) as well as for explaining a system's behaviour. Similarly, system dynamicists often

see a dynamic model as a conglomeration of interacting feedback loops*. However, in modern control theory the so-called state-space view is prevailing. This means that a dynamic system is seen as a conglomeration of interacting state variables. For ease of discussion a system without time-varying inputs is considered. Its state-space description has the form:

$$\dot{\underline{x}}^{**} = \underline{f}(\underline{x},t),\ \underline{x}(t_o) = x_o,\ \text{or} \quad (1)$$

$$\underline{x}(t) = \underline{x}_o + \int_{t_o}^{t} \underline{f}(\underline{x},\tau)d\tau \quad . \quad (2)$$

$\underline{x}$ is the vector of state variables, t_o the initialization time, and $\underline{f}(\underline{x},t)$ is the vector of algebraic (that means: lag-free) functions relating the rate of change of each state variable to the actual value of all state variables, including itself. State variables are all those variables the value of which is adjusted by integration of the effect of one or more rates. Thus, in system-dynamics models all levels together with the sublevels included in the 'delays'*** constitute the vector of state variables. If the number of state variables is n, $\underline{x}$ is a vector in an n-dimensional space, called the state-space.

The point of view that a dynamic system consists of a restricted number of interacting state variables has a number of advantages:

a. By nature, the state-variables approach focusses attention on those elements that are most important from the dynamical point of view: the levels. The dynamic behaviour of a system finds its origin in the inertia included in the integration process taking place in the levels; without levels, no dynamic behaviour is possible at all!

-----

* In this context, it is amazing to see how little use is made of the feedback loop as an instrument for influencing behaviour in system--dynamics policy analyses.

** $\dot{x}$ means: derivative of x with respect to time.

*** In control theory, the word 'delay' is used for pure or pipeline delays only, while n^{th} order stable systems (such as SMOOTH and 'DELAY' in DYNAMO) are called 'lags'. Terminology in this paper, however, will conform to common system-dynamics practice.

b. Automatically, components that have similar dynamic effects are treated in the same way. Think, e.g., of delays and levels, and of the relations among the state variables: Parallel links that connect two variables in the same direction may easily be lumped together to form one influence.

c. Only two kinds of variables have to be distinguished: state variables and their rates of change. This reflects and exploits the fact that in essence all coupling variables are functions of the state variables only, and that the system is fully specified by the values of its state variables.

d. The state-variable description of a dynamic system is at the same time complete, and irreducible. The number of independent state variables* determines the dimension of the system. Elimination of one or more of the independent state variables essentially excludes part of a system's dynamics. Particularly if large and complex systems are considered, this is a great advantage. The maximum number of links between the state variables is proportional to the square of the system's dimension, whereas the number of possible feedback loops is much higher. One might argue that the number of important feedback loops will not be much larger than the number of levels in a system. However, it is easily seen that in a system in which all links between the state variables are of more or less equal importance, the number of feedback loops that affect system behaviour is much larger than the number of levels, and even than the number of links.

However, like each point of view, the state-variable representation has its shortcomings also: It cannot adequately deal with high-order

-----

* If the set of state variables is dependent, at least one of the state variables can be written as a linear function of (some of) the others, and thus the dimension of the state space is less than the total number of state variables.

delays. An n^{th} order delay would induce n state variables, thus blowing up the dimension of the system. However, most high-order delays (or even pure delays) can be replaced by first-order lags without affecting a system's basic dynamic properties*. Different problems emerge if the system includes elements that display discontinuous behaviour as a function of time (e.g. CLIP-functions), or that contain hysteresis effects. The former can be dealt with by introducing time explicitly in the right hand term of (1), but many state-space-based techniques ignore this possibility.

Finally, many system dynamicists (and others!) will make the objection that the transition to the state-space representation of a dynamic system will lead us too far away from the issue we are interested in, namely understanding the forces and mechanisms acting in the real system. Indeed, there is a real danger of drifting away into mathematical abstractions. But, on the other hand, the analyst should not feel refrained of using any tool that can be helpful. Techniques starting from the state-space principle (such as total linearization, see Chapter IV) systematically and efficiently uncover the main dynamic properties and the most important links in a model. Such information is very helpful to the analyst. Although it is not immediately related to the basic assumptions, it considerably facilitates the further exploration and understanding of a model's behaviour in terms of those underlying assumptions, and also in terms of the feedback-loop structure.

II.3.2 Time constants

The term 'time constant', usually represented by the symbol τ, is used frequently to characterize the time-variability of a dynamic system. Intuitively we feel that the order of magnitude of the time constants included in a system determines the system's inertial properties: large time constants involve slow change, whereas small time constants may give rise to quick variations. However, it appears to be quite difficult to give a precise definition of what a time constant is, covering all meanings in which the concept is used. Upon closer inspection, it appears

-----

* The explanation is that the many levels included in high-order delays are virtually dependent, so that they can be lumped together without inducing major dynamic modifications.

that the term 'time constant' is mostly used to denote one out of three different meanings that will now be defined tentatively.

a. *τ1: a characteristic of a single, lag-free feedback loop.*
The average lifetime of capital, and the adjustment time of a level to its indicated or required value are examples of this kind of use of 'time constant'. In principle, it would be more correct to speak of τ1 as a characteristic of the behaviour of an isolated state variable with only one active lag-free feedback loop, but this was not done to conform with common parlance, and to avoid confusion with the second meaning in which the term is used. If the rate R (affecting the level L) is a function of L itself, a mathematical definition of τ1 is:

$$\tau 1 \triangleq \left(\frac{\partial R}{\partial L}\right)^{-1} . \qquad (3)$$

Thus, if R is a linear function of L, τ1 is the reciprocal of the multiplier that defines the rate in terms of the level. If τ1 is positive, L will display exponential growth, while exponential approach of an equilibrium value will be found if τ1 is negative. However, usually only the absolute value of τ1 is taken into consideration, since negative time has no meaning.

b. *τ2: a characteristic of the behaviour of a state variable.*
More often than not, the rate of change $\dot{x}$ of a state variable x is a function of - among other state variables - the value of x itself. Analogously to the definition (3) of τ1, τ2 is defined as

$$\tau 2 \triangleq \left(\frac{\partial \dot{x}}{\partial x}\right)^{-1} . \qquad (4)$$

τ2 characterises the behaviour of x if, in a model, all other state variables are frozen. It is the reciprocal of the gain of the combination of all lag-free feedback loops around x. If τ2 is negative, it is a measure of the speed at which the state variable x will attain steady state after a small step-wise change, and therefore of the lag implicit in x. If τ2 is positive, the net feedback of all loops around x is

positive, and x is a source of autonomous growth in the system. This interpretation of the term time constant is used frequently in classical control theory. The definition of τ2 fits into the concept of a 'time constant of a first-order delay'.

c. *τ3: a characteristic feature of the overall behaviour of a system*
The speed of the changes that actually take place in a system or model can be characterised by some measure of time, but it is quite difficult to give a simple and precise definition, except for linear systems. It is a well-known result of linear theory, that the behaviour of each system variable can be described by the sum of a limited number (equal to or less than the dimension of the system) of exponential functions of time (including the complex exponentials). The reciprocals of the values of the exponential coefficients(that can be calculated as the eigenvalues of the matrix describing the system in state-space notation) are considered the time constants of the system. The time constants τ3 of a non-linear system may be computed by linearization of the systems's equations, or by analysis of the behaviour of its variables (study of bandwidth characteristics, fitting of behaviour to exponential functions).

In the simple case of one state variable and one loop, the value of the time constant is independent of the choice of definition. Also, it is evident that time constants are really constant in linear systems (or, for τ1, for linear loops) only. In the case of nonlinear relations, the values of the time constants according to all three definitions may change with the state of the system, and thus with time. The time constants τ1 and τ2 may be implicit in table functions, multiplier relationships, etc., but can always be computed directly from the equations by tracing and combining lag-free feedback loops. The relation between τ1 and τ2 on the one hand, and τ3 on the other, is more obscure. Except for a few simple cases (such as first- and second-order linear systems) it cannot be established without solving the model equations.

II.3.3 Structure

The word 'structure' is used frequently with reference to a system or model, but not always with the same meaning. The structure reflects the pattern of relationships between the variables, but to what degree of detail? In fact, the word is used to indicate one out of a whole series of possible meanings. A lot of discussion might be clarified if more precise definitions of 'structure' would be used. As a first step, it is proposed to distinguish between the following three interpretations.

a. An influence diagram

The influence diagram shows the existence or non-existence of relations between the main variables of a model or system. It shows also which variables are endogenous, and which exogenous. In this context, a structural change means a change in model boundary, omission of an existing link, or addition of a new link, but not a change within a link. Figure 1 shows a simple example of such a structure.

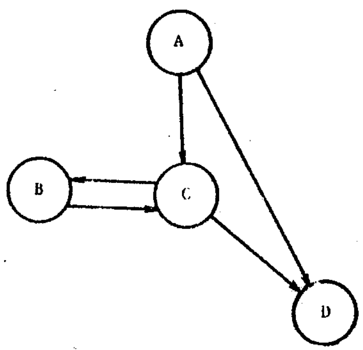


Figure 1: Influence diagram

b. A flow diagram

Often, the word 'structure' can be replaced by 'flow diagram'. It includes all single dynamic relations as well as the nature of the variables (level, rate, intermediate variable). According to this interpretation, a change in structure means a change in flow diagram,

such as the addition or omission of a loop or a delay. A parameter change may have structural implications if its result is to cut or to add a loop or relationship (e.g. if the change results in a zero-value of a multiplier).

c. A picture of the actual working of the system

This meaning of 'structure' will henceforth be denoted as "*operating structure*". It reflects the mechanisms actually governing a system's behaviour, and may involve different degrees of detail. The operating structure of a model can, e.g., be illustrated by a simplified flow diagram containing the most influential loops and variables only, or by a picture showing the major interactions between the subsystems (as shown in Figure 2 in this article).

It is not unusual that a system that is structurally complex according to the first two interpretations of 'structure' (influence diagram, flow diagram), has a very simple operating structure. Whether a model's operating structure is changing with time during a simulation depends on the level of detail considered, and on the degree of non-linearity of the model*. Clearly, changes in parameter values may have implications for the operating structure of a model.

II.3.4 Stability

Often, a system or model is called stable or unstable, but the concept of stability is not always well defined. It is generally agreed that a system is stable if it returns to an equilibrium state upon exogenous perturbations. However, the concept is also used to indicate that the effects of perturbations in a dynamic mode of behaviour will vanish. A third interpretation of stability applies to the general mode of behaviour of a system: If a system displays growth followed by decline, or oscillations, rather than a gradual approach to an equilibrium value

-----

* In systems theory, systems in which nonlinearities play an important part, or which include explicitly time-dependent relations, are often called 'variable-structure systems'.

(like,e.g.,the World models), it may be called unstable.

In this paper, a (sub)system will be called stable for a given state if, all exogenous inputs being constant, it has the inherent tendency to approach an equilibrium starting from the particular state. It will be called unstable if it displays the tendency to grow exponentially under the same conditions. The definition draws upon linear system theory, and is consistent with at least the first two of the interpretations given above. However, the definition implies that it may be difficult to give a single judgement about the stability of a nonlinear system.For a certain state a nonlinear system may display a tendency to grow exponentially (e.g., the World models during the growth phase, i.e. between 1900 en 2010), whereas the same system may possess all properties of stability for a different state (e.g., the decline phase in the same model). Moreover, a system (nonlinear as well as linear) may be composed of stable and unstable parts or subsystems at one and the same time (e.g., the capital subsystem in the world models during the growth phase (unstable), and the persistent pollution subsystem (stable)). Also, the stability properties of a nonlinear system may depend on the exogenous inputs.

The same problems occur in nonlinear systems theory. No simple measure of stability exists for nonlinear systems. Gibson |4| mentions that, for nonlinear systems, more than 28 definitions have been proposed and used by various investigators, such as asymptotic stability (boundedness if time approaches infinity) and monotonic stability (which requires a <u>gradual</u> approach to equilibrium).

III. GUIDELINES AND STRATEGIES

In this chapter a number of guidelines and strategies for understanding dynamic models are presented. They attempt to transmit to the reader part of the empirical knowledge and experience existing in the field of model understanding. Although many are trivial or well-known, they are too important to be ignored.

The first part of this chapter gives a list of guidelines that primarily attempt to draw attention to certain research attitudes that are desirable when the dynamic behaviour of a model has to be understood.

Subsequently, the importance and possible outcomes of a number of research strategies will be discussed. The guidelines and strategies are not cited in a particular order of importance. It should be emphasized that they undoubtedly have been influenced considerably by the author's background, and that the list is far from being exhaustive: there is much room for improvement in this field.

III.1 <u>*GUIDELINES*</u>

1. Be aware of the fact that the only general rule in nonlinear systems is that there are no general rules, except this one. Any approach may be useful in certain conditions, but completely useless or even misleading in other circumstances.

2. Do not limit yourself: Use any idea or approach you can think of. Each has its own merits - and shortcomings.

3. Be not afraid of abstracting from reality. Consider a model as a mathematical structure only, but always return to the original starting point (basic assumptions, reality) afterwards.

4. Always ask: "Why?", and do not rest before you have the correct answer, and you are sure that it is correct.

5. Do not overlook obvious things: they may be of crucial importance.

6. Always be sceptic about obvious explanations. It is too easy and tempting to explain phenomena in a wrong way (apparently strong loops may hardly affect system behaviour, positive loops do not necessarily induce exponential growth).

7. Keep always thinking yourself! No tool, technique or trick will *generate* insight. Their only contribution is to provide information in a meaningful, ordered manner, so that the generation of insight by the analyst is facilitated.

8. Do not expect that any technique will generate unique information. All system properties may be uncovered in different ways, but, in each particular case, certain approaches might be more appropriate than others.

III.2 STRATEGIES

III.2.1 Decomposition

Particularly when the set of equations involved is large, decomposition of a system is fruitful. Dissection facilitates the analysis, because the resulting subsystems are smaller, and therefore often more comprehensible than the overall model, and lend themselves better to further examinations. Decomposition on the basis of partial understanding may enforce the understanding, and facilitate the explanation of the model's working in terms of the original assumptions.

The dissection may be based on several criteria. Large models, for example, can be decomposed into the interacting submodels that have been built more or less separately. Moreover, the flow diagram itself may be revealing: clusters of equations that display many interactions but have only few links with the rest of the model can often be discerned. Other methods of decomposition are based on the actual working of the system. A well-known strategy is to distinguish between active and dormant parts. The so-called "dynamic decomposition" implies making a distinction between subsystems that possess a relatively large amount of inertia (i.e. include relatively large time constants in the sense of τ3), and subsystems containing relatively low time constants*.

Generally speaking, there is no unique way of dividing a system. For a given purpose, one division may be more opportune, whereas, for another purpose, another division may be more appropriate.

III.2.2 Understanding a system on different levels

Understanding a system on different levels is a very simple procedure, and not substantially remote from what we do in everyday life. When people try to understand a complicated system as a whole, they ignore many details and focus only on the interplay of the subsystems that together make up the whole system. In turn, the internal working of each individual subsystem is understood in the same way, and so on. At each level of understanding, it should be observed that the interactions between

-----

* which does not necessarily imply that the subsystem's variables actually display quick variations!

the separate parts are more important than the individual subsystems. In fact, understanding a system on different levels is the only possibility, since the capacity of human mind sets limits to the number of details that can be considered at one and the same time.

III.2.3 Understanding system hierarchy

The interplay of the subsystems causes a system to be more than the sum of its parts. Therefore, understanding the hierarchy of the subsystems is essential to understanding the working of the whole.

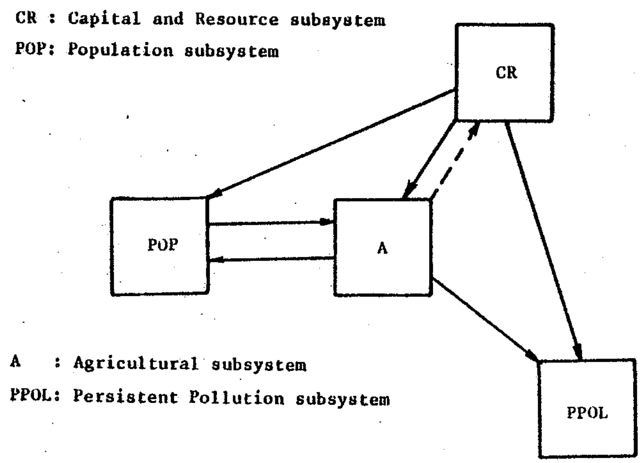


Figure 2: Major interactions among the subsystems of World3 under standard-run conditions, clearly showing the system's hierarchy

The interplay between the main parts of the World3 model under standard-run conditions provides a clear example. Figure 2 shows the hierarchy of the model. Except for the influence of the agricultural sector (quite weak under standard-run conditions), the capital and resource subsystem behaves quite autonomously, and affects the behaviour of all other model sectors.

It is not exaggerated to regard this subsystem as the central power station of the world model. When it goes down too. The growth and decline food production sooner or later have to go down. The growth and decline mode of behaviour is included in the capital and resource sector, and impressed on all other sectors (see |14|). This insight has important implications for policy analyses: when controls are implemented that effectively stabilize the capital and resource subsystem, all other sectors will ultimately tend to equilibrium also! Because the nature of the equilibrum of population is rather undesirable*, a combination of two modifications will have to be introduced to improve the model's behaviour in a fundamental way: one resulting in a stable capital and resource subsystem (e.g. by allocation of part of the industrial output to resource conservation), and another resulting in more desirable equilibrium possibilities for population (e.g. by reducing the number of children per family).

This example clearly illustrates the importance of understanding system hierarchy. It enables the analyst of locating the basic causes of system behaviour, and of formulating policy starts that take advantage of the system's natural properties.

III.2.4 *Modes of behaviour*

For various reasons, it seems wise to start with a thorough investigation of only one mode of behaviour. The techniques and methods that are adopted, and the dissection that is found to be appropriate for understanding one particular mode of behaviour may be of great use to detect the model's properties under largely different conditions. The understanding of a reference behaviour provides an excellent starting point for answering such questions as: "Under what conditions will the conclusions drawn become invalid?", "What will happen if the limit to the validity of certain simplifications is exceeded?", and "Which subsystem

-----

* The natural equilibrum of population in World3 is mainly caused by starvation. A different kind of natural equilibrium may exist at rather high levels of income (about the present U.S. conditions), but it is - under present conditions - extremely unlikely that such high income levels can be sustained for long for the whole world population.

will behave differently under alternative conditions, and which not?".

III.2.5 *Model simplification*

A simple, but expedient strategy is to try to simplify the original set of model equations as much as possible, but without loosing the basic behavioural characteristics. If the set of equations that remains cannot be reduced any further without affecting behaviour, it must consist of the fundamental assumptions leading to the model's overall behaviour, and further research into the causes of this behaviour can be directed to these equations solely. Also, the knowledge that other assumptions are unessential to behaviour may be valuable.

III.2.6 *Irrealistic modifications*

For model-testing purposes, all changes (in equations, or parameters) will have to be limited to the set of more or less realistic possibilities. However, when a model's operating structure is investigated, rigorous modifications and falsifications are permitted or even required, since their effects may be much more pronounced, and hence much easier to interpret*.

III.2.7 *Combinations of tools*

Combination of different tools for gaining information about a dynamic model is not unusual, but it seems desirable to draw attention to this possibility. Particularly, sequential application of various techniques and strategies may be very expedient, e.g., simplification of equations, reformulation** and analytical solution**.

Also, it should be emphasized that almost each technique can be applied for at least two purposes: first, for the detection of information on system properties, and, second, for the verification of hypotheses. If, for instance, omission of a loop does not affect system behaviour, the

-----

* Probably because of their emotional involvement with the product of their own efforts, most modelers are inclined to subject a model to realistic modifications only. Therefore, it might be argued that a modeler should delegate the task of analysis to someone not closely involved in the process of gathering data and building the model.

** See the next chapter.

suggestion is put forward that the particular loop is not influential. The same suggestion, however, can come up as a consequence of another experiment, and then cutting the loop can be used as a technique for verification of this hypothesis.

IV. TOOLS

Various techniques that are helpful in the process of understanding dynamic mathematical models are presented. The list has been divided into groups. Each group consists of a number of tools that are particularly suited for achieving a specific sub-goal (e.g. simplification of equations) or that have a similar character (e.g. model modifications). It must be re-emphasized that, in fact, the distinction is not so clear, and that most techniques can be applied in various ways.

IV.1. *FLOW DIAGRAM MANIPULATION*

A flow diagram may serve many purposes: it may be used as an influence diagram, as an illustration of the elementary assumptions made in a model, as a source of information concerning the form of the equations, as a means for communication of information, as a basis for discussion and, last but not least, as a point of departure and a source of ideas for the model analyst. For any purpose, it is required that a flow diagram is clear and understandable. A few methods of improving DYNAMO flow diagrams for the purpose of model analysis and understanding follow below.

IV.1.1 *Addition of more information*

It is helpful to the analyst if the model equations can be written down directly from the flow diagram. But then, the diagram must contain more information than DYNAMO flow diagrams include. Particularly, adding multiplication, division, addition and subtraction signs, and introducing different symbols for table functions and other algebraic relationships facilitates the direct transition from the diagram to a set of well-defined equations.

IV.1.2 *Reorganization and redesign*

Since diagrams may serve various purposes, they should be redesigned in a different way for each goal. During the insight-generation process, a diagram may be a powerful aid for tracing influences from one variable to another, or for detecting feedback loops. A DYNAMO flow diagram usually is a mixture of an influence diagram and a picture of the flow of physical goods and information. As a consequence, it may be difficult or even impossible to go around loops in the right direction. Out-flow-rates, for example, are connected to the corresponding level by arrows *leaving* the level, whereas their value is actually influencing the level. Therefore, redesign of DYNAMO flow diagrams so that the direction of all arrows corresponds to the direction in which the variables actually influence each other may be useful.

If a diagram has to distinguish as clearly as possible between dynamic elements (state variables) and algebraic ones, it is wise to redesign it in such a way that only two different symbols are used - one for state variables and one for all other variables.

Finally, in all phases of model building and analysis, the modeler should rearrange the diagram time and again in an attempt to bring out the structure (in any sense) of the set of equations as lucidly as possible. This may be achieved by emphasizing similarities in different parts of the model, by separating individual subsystems, by avoiding intersections by the influencing lines as much as possible, and in many other ways.

IV.2 *INVESTIGATION OF THE EFFECTS OF CHANGES AND PERTURBATIONS*

Many techniques for model analysis and testing are based on the introduction of exogenous perturbations or of one or more changes in the set of equations, and comparison and explanation of the differences and similarities between the outcome of the modified or perturbed model and the original one. The introduction of perturbations and modifications can yield information relevant to answering all three questions involved in model understanding (What will happen, if...?, Why does it happen?, and: When will...happen?).

Again, the list of possibilities following below is far from being exhaustive. It will be tried to put emphasisis on those techniques that are simple and expedient, but not so well-known in system-dynamics circles.

IV.2.1 *Modifications in exogenous inputs*

It is started from the point of view that parameters (and constants) are in fact exogenous inputs that do not change in a relevant manner during the simulations. Thus, modification in exogenous inputs includes changes in parameter values as well as in the behaviour of time--varying inputs. The technique is best known as sensitivity analysis. Simple applications consist in changing the value of one or a few parameters at a time. There are more complicated variants also, such as Monte-Carlo tests, the direct calculation of sensitivity functions (see Sharp |12| and Tomovic |15|), and the so-called hill-climbing methods. An advantage of these methods is that they are systematic, and include variations in *all* parameters. However, for the greater part the information they produce is restricted to the possibility of a certain behaviour, or to the sensitivity of the parameters, and therefore their utility is larger for testing than for understanding models.

If an extensive sensitivity analysis has to be performed, the best bet is to first investigate how the parameters occur in the equations. Often, several parameters perform in an exactly analogous way, for instance if only the product or quotient of two coefficients occurs in the equations. These coefficients can be combined into groups of parameters, only one of which must be varied to show the sensitivity of all. In |14|, the use and detection of parameter groups has been illustrated for the World3 capital and resource subsystem.

Sensitivity analysis may yield important information on model hierarchy: usually, each of the subsystems contains sensitive as well as insensitive parameters, but overall behaviour is affected only by modifications in those subsystems that play a leading part in the model hierarchy. Again, an example can be borrowed from the World3 model: Changes that influence the behaviour of the capital and resource subsystem

(such as a change in the value of the industrial capital output ratio (ICOR)* from 3 to 4) cause significant deviations in other parts of the model. Conversely, if modifications are introduced elsewhere, for instance in the persistent pollution sector (such as a change of the assimilation half life in 1970 (AHL70) from 1.5 to 2), the effect on the other subsystems is almost negligible. This result suggests a dominant position of the capital and resource subsystem, and a minor part for the persistent pollution sector in the model hierarchy.

IV.2.2 *Falsification of state variables*

In contrast to sensitivity analysis, the technique of falsification of state variables consists in a perturbation of the values of one or more *endogenous* variables. Because any change in a rate or coupling variable can be explained as the result of a change in a parameter, the attention is focussed on the state variables only. A simple application of the technique is to augment the value of a state variable at the initialisation point or any other point in time during the simulation, and to compare the results with the outcome of the unperturbed simulation. It is an excellent method of isolating the behavioral impacts of the variations that occur in the value of a particular state variable. Also, it may yield information on the time constant ($\tau 2$) associated with a state variable, and on its importance in the model hierarchy.

The foregoing generalization can be illustrated by reference to the World3 model. If the value of industrial capital IC in 1970 is doubled, the overall behaviour of capital and pollution changes considerably, while population is hardly affected, at least in the growth phase (see Figure 3). On the contrary, if the value of persistent pollution in 1970 is doubled, or even multiplied by 10, minor changes can be perceived only during the first 10 years following the perturbation. Even pollution itself returns quickly to its original order of magnitude. These results illustrate that

-----

* A list of letterscripts and their associated meaning is given in an Appendix.

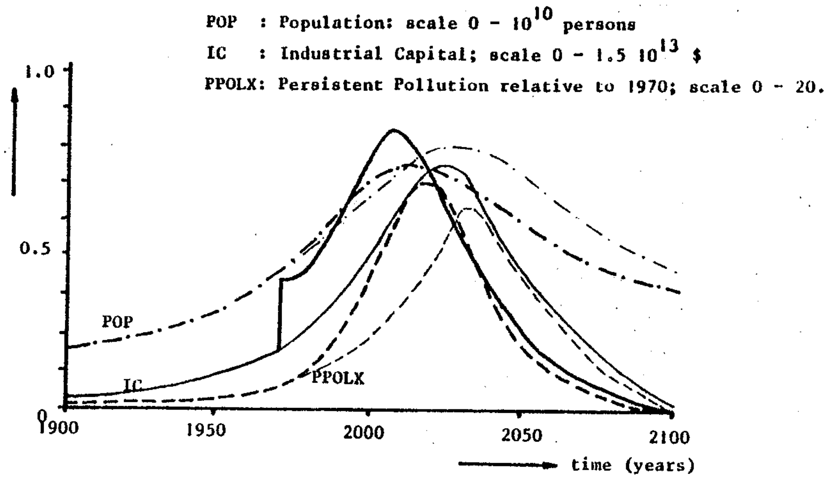


Figure 3: Behaviour of World3 if, during a standard-run simulation, the value of industrial capital IC is doubled in 1970. Thin lines show the unperturbed standard-run results.

the capital and resource subsystem behaves quite autonomously (changes in the values of state variables persist, or even become larger), that its influence on persistent pollution is quite large, that popu-lation is hardly affected by the variations in capital (at least in the growth phase), that the behaviour of pollution is nearly completely determined by the other sectors, and that the time constant 12 of per-sistent pollution is relatively small (a few years or less).

IV.2.3 Cutting links

This method can be applied to investigate the behavioural impact of single relationships or loops, or of combinations. The technique is to freeze the value of one or more coupling variables, table functions or rates from a certain moment onwards. If the model's overall behaviour is

not affected, the contribution of the link that has been frozen may be neglected if only the basic reasons for model behaviour are sought*. In most cases, cutting links implies cutting one or more loops. Therefore, various explanations of the sensitivity to freezing a single link can exist: If the link forms part of only one loop, this loop is probably unimportant to behaviour, but if more than one loop is involved, the working of the different loops might be compensating.

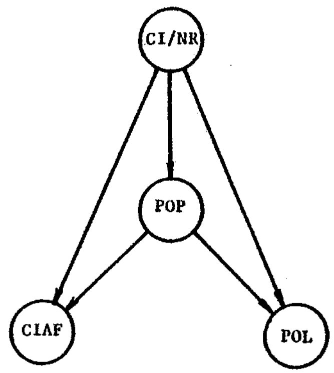
IV.2.4 Freezing state variables

Since the state variables are the basic sources of dynamic behaviour, freezing the value of a state variable means excluding part of the dynamics (or: making part of the dynamics exogenous). In terms of feed-back loops, freezing a state variable means cutting all loops passing through the state variable. The technique is simple but powerful since it may show which part of the dynamics is crucial, and which not, and thus may yield a wealth of information on the model's hierarchy and operating structure.

Let me illustrate the application of the technique using the World2 model as a vehicle. Each of the five state variables included in the model was frozen from 1970 on. Freezing pollution POL or capital-investment--in-agriculture fraction CIAF does not significantly affect the behaviour of any of the other state variables. Freezing population POP from 1970 on affects pollution, leaves CIAF more or less constant, but the behaviour of capital investments CI and natural resources NR is virtually un-changed. If, in turn, CI or NR are frozen from 1970 on, all other variables will be affected as well (as illustrated by Cuypers |1|, freezing CI results in constant POP, CIAF and POL within a few years). The simple diagram of Figure 4, showing the hierarchy in the model's operating structure around 1970 is the result. The diagram points to the central position of capital investments CI and natural resources NR in the model. Moreover, the experiment of freezing CI shows that the other parts of the model are in-herently stable (that is, they will not grow exponentially of their own account).

-----

* The fact that a certain assumption does not affect behaviour is also part of the explanation of that behaviour!



CI : Capital Investments
NR : Natural Resources
POP : Population
CIAF: Capital-Investment-in-Agriculture Fraction
POL : Pollution

Figure 4: Hierarchy of the World2 model around 1970.

IV.3 TOOLS FOR MODEL SIMPLIFICATION

A list of simple techniques for the detection of model areas that may possibly be simplified under certain conditions is given. The background idea is that it facilitates understanding a particular behaviour when all equations and assumptions that do not contribute to that behaviour are omitted, and when the set of equations generating the behaviour of interest is as simple as possible.

IV.3.1 Introduction of changes in the equations

All techniques presented in the previous section can yield information on the importance of single assumptions or equations to model behaviour. Changes that do not significantly affect behaviour point to areas of potential simplification. However, equations or variables should not be omitted before the reasons why, and the conditions under which they do not influence behaviour are well known.

IV.3.2 Observation of the range of behaviour of variables

The general strategy is to look for variables that are more or less constant during the simulations. Because merely small deviations in a variable may not necessarily be insignificant, the investigator must subsequently test whether replacement of the time-varying relation by a

proper constant value changes model behaviour. If no important changes occur, the relationship can be omitted, but not before it is understood why the range of behaviour is so narrow.

The capital utilization fraction CUF in World3 is a clear example of a variable that is virtually constant during a whole simulation. Its value remains equal to 1.0 in the standard run because, during almost the whole simulation, labor force exceeds the total number of jobs. As a consequence, the so-called "job-sector" can be ignored for the analysis of standard-run and similar behaviour.

IV.3.3 Contribution analysis

If a variable is influenced by more than one other variable, changes in its value can be explained as the net result of the different contributions of the influencing variables. Comparison (for example by freezing one link) of the order of magnitude of the different contributions can reveal that some of the influences are more or less negligible compared to the others.

Again, a clear example is found in World3. The land erosion rate and the land removal for urban-industrial use affect the level AL (arable land), but their contribution is only small compared to that of the land development rate, and compared to the value of AL itself. Negligence of these two rates hardly affects the standard-run results of the model, but permits, as is illustrated in |13|, the construction of a considerably less complex-looking diagram of the agricultural sector.

IV.3.4 Combination of functions

This technique is self-evident: if the value of a variable is calculated from one other variable by means of several consecutive functional operations, combination of these functions may simplify the equations, and reflect more clearly the total effect of the influence.

IV.3.5 Combination of parallel links

The underlying philosophy is to combine parallel links between variables in order to simplify the flow diagram and to bring out more clearly the total effect of parallel lag-free influences.

An interesting application is the combination of all parallel, lag-free feedback loops that relate the rate of change of a state variable to its own value. The resulting loop provides information on the time constant τ2 of the state variable, and on its stability properties: when all other state variables are constant, the state variable will grow exponentially if the loop is positive, but, if it is negative, an equilibrium will be approached. Such knowledge may be very helpful for the explanation of model behaviour.

The exponential growth of population in World2, for example, is easily explained if the net feedback population on its own rate of growth is positive. However, it turns out (see |1|) that, all other state variables being equal, an increase in population *reduces* population growth, so that population is inherently stable. Therefore the explanation for the exponential growth of POP must be sought elsewhere in the model (in this case in the net positive feedback of capital investment to its own rate of growth causes capital, and also population to grow exponentially).

IV.3.6 *The use and shape of non-linear (table)functions*

Often, the principal behaviour of a model depends upon only a small part of the model's non-linear functions. If the portion of a function that is behaviourally significant has a distinctive shape, the overall function may be replaced by the simple expression as long as model behaviour remains roughly the same. In particular situations application of the technique may directly give rise to interesting insights.

Let me illustrate the foregoing using an example. In World3, the non-renewable resource usage rate NRUR is calculated as the product of population POP and the per capita resource usage multiplier PCRUM:

$$NRUR = POP * PCRUM. \qquad (5)$$

PCRUM is a table function depending on industrial output per capita IOPC (equal to industrial output IO divided by POP). The shape of the function is shown in Figure 5 (fat line). During the standard-run simulation, IOPC remains below 500$ per head per year, and for this range of values, the relation between PCRUM and IOPC is more or less proportional (see dotted line in Figure 5):

$$PCRUM \simeq a * IOPC \,, \qquad (6)$$

where a is a constant. Substitution of (6) into (5) leads to an intersting result:

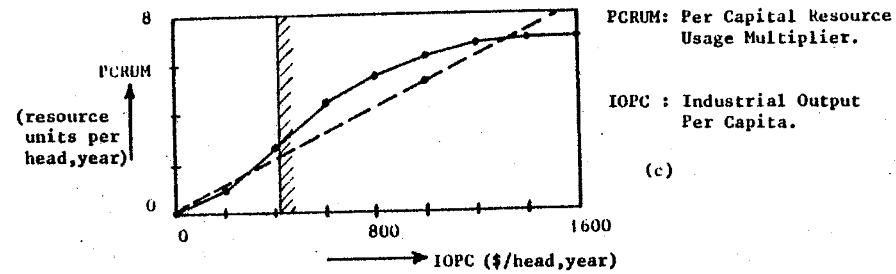


Figure 5: Table function showing PCRUM as a function of IOPC. The dotted line illustrates a linear approximation. Shading shows the part not used during the standard-run simulation

$$NRUR \simeq POP * a * IOPC = a * IO. \qquad (7)$$

This means that in World3 it has implicitly been assumed that, as long as standard-run conditions are held, the usage rate of resources is more or less proportional to industrial output, and virtually independent of the size of population!

Similarly, given the quasi-linearity of two other table functions, it can be shown easily that the direct influence of population on the allocation of industrial output is also virtually nil (see |14|).

IV.3.7 *Elimination of small lags*

The basic idea is that the influences of lags incorporated in stable subsystems depend on the variations in input variables. Under static conditions, the influence of the lags in a stable subsystem is nil, and the subsystem behaves as if it were algebraic since no changes through time can be delayed or accelerated. If the input variations take place only slowly, the effects of relatively small lags might also be neglected. The equations could be simplified by replacing the original differential equation with the proper algebraic relation and thus eliminating a state variable. However, the technique should be applied with great care and, preferably, only by experienced investigators.

Test calculations always have to be performed, because even a small lag may have a great influence, for example, if it helps to calculate the average rate of change of a coupling variable. Moreover, lags that are negligible for a particular behaviour might be very important under different circumstances.

The approach is illustrated using the land-fertility subsystem in World3: Land Fertility LFERT appears as a state variable, calculated according to the equation:

$$\dot{LFERT} = (ILF-LFERT)/LFRT - LFDR * LFERT. \quad (8)$$

Inherent land fertility ILF is a constant, and LFRT (land fertility regeneration time) and LFDR (land fertility degradation rate) are functions of the input variables to the subsystem. In the standard notation for a first-order system, (8) is converted into

$$\dot{LFERT} = - LFERT*(LFDR+1/LFRT) + ILF/LFRT, \quad (9)$$

which means that the two feedback loops from LFERT to $\dot{LFERT}$ have been combined. This equation specifies a system with the time constant $\tau 2$:

$$\tau 2 = -(LFDR+1/LFRT)^{-1} = -LFRT/(1+LFDR*LFRT) \quad (10)$$

and the static function

$$LFERT_{static} = ILF/(1+LFDR*LFRT). \quad (11)$$

The latter expression can be used to compute directly the equilibrium value of LFERT if LFDR and LFRT are constant. From an observation of the standard-run values of LFDR and LFRT, $\tau 2$ apparently has a value of about 5 years or less during the whole simulation. Because the main changes in the input variables take place over a period of about 100 years, the dynamic effect of the LFERT subsystem can therefore probably be neglected. This hypothesis has been fully confirmed by a test calculation in which the land fertility subsystem was replaced by the algebraic relation (11).

IV.4 DETECTION OF FUNDAMENTAL PROPERTIES OF (SUB)SYSTEMS

This section presents techniques that help understanding the basic properties of a (sub)system. They are particularly useful in answering the questions "What will happen, if...?" and "When will...happen?". More often than in the preceding sections, a model will be considered as a set of mathematical relationships only. However, it is usually not difficult to find the elementary assumptions underlying particular system properties once the latter have been isolated. Clearly, most of the techniques presented in the two preceding sections may also be helpful for the detection of fundamental properties.

IV.4.1 *Equilibrium analysis*

The principle of equilibrium analysis is to investigate the conditions under which a (sub)system will be in equilibrium, and the nature of that equilibrium. It is started from the observation that a dynamic system with a set of exogenous inputs $\underline{u}$ can be described by:

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{g}(\underline{u}), \quad (12)$$

where $\underline{f}$ and $\underline{g}$ denote vectors of algebraic functions. The system is in equilibrium if $\dot{\underline{x}} = \underline{0}$, and hence:

$$\underline{f}(\underline{x}) + \underline{g}(\underline{u}) = \underline{0}. \quad (13)$$

Generally, equation (13) cannot be solved since the number of unknown variables (state- and exogenous variables) exceeds the number of equations (equal to the number of state variables). Therefore the equilibria are investigated under specified constant exogenous conditions, which leads to the equilibrium equation for a closed system:

$$\underline{f}'(\underline{x}) = \underline{0} \quad (14)$$

From this equation, possible equilibria, if any, may be computed directly using techniques of numerical analysis.

However, when subsystems with only a few identifiable input variables are under consideration, an analysis using simulation is often simpler, and may yield more information. It consists of freezing the value of all input variables to the investigated subsystem. If the subsystem is stable under the existing conditions, its state variables will gradually tend to equilibrium. By application of various constant input values insight can be gained into the general relation between the equilibrium outputs and the corresponding input values.

Let us take the population sector of World3 as an example: if its input variables IO (industrial output), SO (service output), F (food) and PPOLX (PersistentPollution relative to 1970) are given fixed values, population POP will gradually tend to equilibrium. Investigation of the ultimate equilibrium value of POP for different combinations of values of IO, SO, F and PPOLX shows that the equilibrium value of POP is mainly determined by food (the explanation is obvious: because total food production is kept constant, either population growth continues until starvation causes the death rate to rise considerably, or, when the initial value of POP is *above* its ultimate equilibrium, population declines due to starvation until birth and death rates are equal). This result suggests that population will not continue to grow exponentially unless one or more of the input variables do. It also shows that when a policy succeeds in a stabilization of IO, SO, F and PPOLX, ultimately population will be stable too, but people in the model will have to live under rather miserable conditions.

IV.4.2 *Dynamic tests of subsystems*

This technique is suited to investigate the dynamic properties of stable subsystems with few input variables. It consists in imposing a special test signal (such as a pulse, a step function, a ramp function or a harmonic signal) on one or more of the constant input values. By comparing the reaction of the investigated subsystem with the reaction of simple, well-known systems (for example, a first-order lag) to the same input signal, a fairly good indication of the subsystem's most important dynamic properties can be obtained. Particularly the step-function method is simple and expedient.

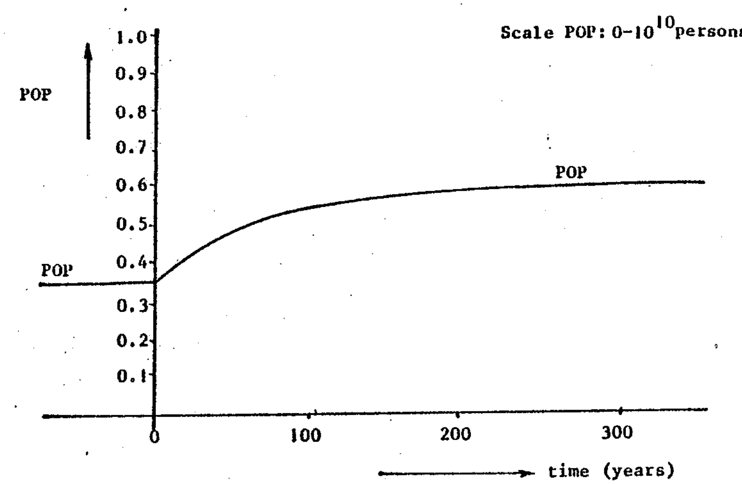


Figure 6: Response of population POP in World3 if industrial output IO, service output SO, food production F and persistent pollution ratio PPOLX are doubled at time=0 (original values: IO=2 10^{11}, SO=3 10^{11}, F=7.5 10^{11} and PPOLX=5)

Figure 6 shows the reaction of population in World3 if, starting from an equilibrium situation (constant inputs and constant population) the value of the inputs is suddenly doubled (step function). Population rises slowly to reach a new equilibrium value. The dynamics of the sector are, apparently, rather slow, and may be characterised by a time constant (13) of about 70 years.

Figure 7 shows the reaction of two output variables of the agricultural sector of World3 if a step function is imposed on population, and all other inputs to the sector are kept constant. Almost immediately, the outputs reach new, fairly constant values, clearly showing that the dynamics of the agricultural subsystem are relatively fast (the time scale is different from that of Figure 6!). Many more examples could be given, but they would lead us beyond the spatial constraints of this paper.

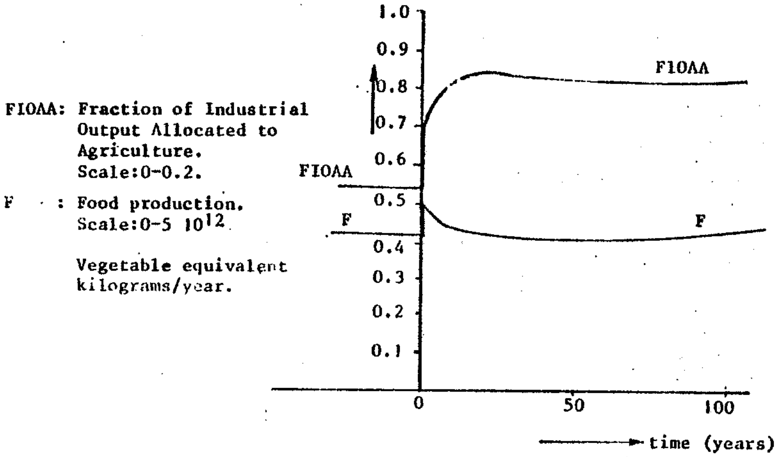


Figure 7: The response of two output variables of the World3 agricultural sector if population is doubled at time=0 (original value: POP=4 10^9).

IV.4.3 *Reformulation of equations*

Equations may be reformulated to bring out their characteristic properties as lucidly as possible, or to uncover implicit assumptions. However, the usefulness of reformulation is closely related to the specific structure and properties of the set of equations to be examined. No general strategy can be formulated. Some experience with the manipulation of mathematical equations is very convenient.

Rademaker (|9| and |10|) discusses a clear example: he starts from the observation that pollution absorption POLA in World2 is a function F' of pollution POL only:

$$POLA = POL/POLAT = F'(POL), \quad (15)$$

because the pollution absorption time POLAT is a nonlinear function of pollution. Figure 8 shows the graph of POLA as a function of POL. The curve shows that pollution absorption will rise when pollution rises from zero to 10 times its 1970 level, that POLA has a fixed value for

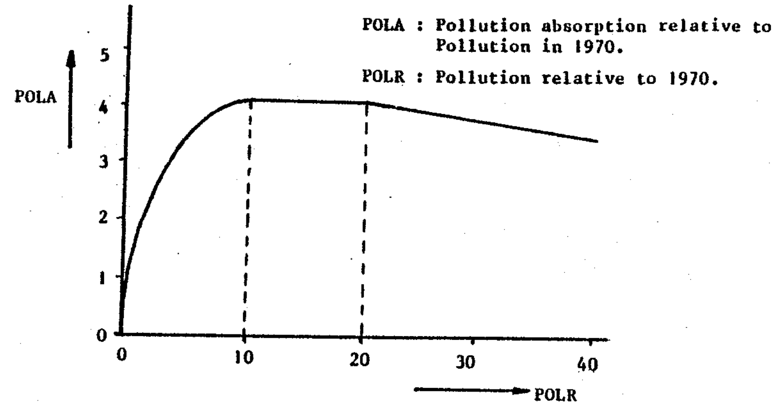


Figure 8: Pollution absorption as a function of pollution

pollution between 10 and 20 times the 1970 value, and that POLA will even decline if pollution rises to higher levels, thus yielding a possibility for a real explosion of pollution at high generation levels. The original formulation shows not nearly as clearly how POLA depends on POL. Starting from this graph, the behaviour of pollution as a function of pollution generation can be easily traced under all circumstances, normal and abnormal, and without any further computation. However, because of spatial constraints, the reader is referred to the original reports |9| and |10| for a further elaboration.

IV.4.4 *Analytical treatment of subsystems*

Analytical solution of simple subsystems is very informative, and it allows direct computation of a system's behaviour at any specified time. However, for the greater part analytical solution is impossible because of high dimensionality and nonlinearity. Yet, more attention should be paid to the possible rewards of studying small subsystems using simple analytical tools, a strategy often overlooked in computer-simulation circles. Analytical treatment, even without explicit solution, can yield far more fundamental and general information on the properties of a (sub)system than numerous simulation runs. However, a fruitful application of the approach calls for some experience and an elementary knowledge of mathematical analysis.

In fact, two of the examples that have already been presented to illustrate various techniques have been handled analytically: the elimination of the lag in the land fertility subsystem of World3 (section IV.3.7), and the linearization of the table function PCRUM (section IV.3.6). Let us continue the latter example. It was concluded that the non-renewable resource usage rate NRUR was, for standard-run conditions, virtually proportional to the industrial output IO. Hence, it follows for $\dot{NR}$, the rate of change of non-renewable resources NR:

$$\dot{NR} \sim -a * IO. \tag{16}$$

The 1900-value of NR is NRI. Integration of (16) yields:

$$NR(t) - NRI = -a \int_{1900}^{t} IO(\tau)d\tau \tag{17}$$

Therefore, since NR(t) cannot be negative:

$$\int_{1900}^{\infty} IO(t)dt \leq NRI/a \quad . \tag{18}$$

In other words, once NRI is given, the total amount of industrial output over time cannot exceed NRI/a. This is the logical consequence of the assumptions that the amount of resources is limited and cannot but decline on the one hand, and, on the other, that each unit of industrial output produced requires a fixed amount of resources. The expression (18) explains why changes in the model that cause a faster growth of industrial output also provoke an earlier system decline. Conversely, if moderated, industrial growth may persist longer. The same characteristic features can also be found by performing several sensitivity simulations, but this simple analytical expression *directly* shows the fundamental reasons of system behaviour.

IV.4.5 *State-space techniques*

This technique is based on the notion that, if the number of independent state variables is n, the state vector $\underline{x}$ corresponds to a vector or point

in the n-dimensioanl state space. Conversely, each point in the space is associated with a state vector, and the rate of change of each vector can be calculated directly from the system's equations (1). Consequently, it can also be derived directly in what direction each point in the state space will move. Thus, starting from different initial positions, the various state-space trajectories $\underline{x}$ may follow through time can be constructed. Therefore, what kind of behaviour will follow from any starting point can be found easily. Also, possible equilibrium points can be isolated, and characteristic features such as limit cycling can be recognized and explained.

Unfortunately, the approach can only be used easily for (sub)systems containing not more than 2 or 3 independent state variables. Typically, some mathematical insight and some energy are required, but the general and fundamental nature of the insights that may be obtained completely counterbalances the disadvantages.

Again, the capital and resource subsystem of World3 offers an elucidating example: In |14| it is shown that, without affecting the basic properties of the subsystem, the interaction between industrial capital IC and non--renewable resources NR for $NR/NRI \leq 0.5$ can be described by:

$$\dot{IC} = b * IC * NR + c * IC \quad \text{and} \tag{19}$$

$$\dot{NR} = d * IC * NR, \tag{20}$$

where b, c and d are constant, with $b>0$, $c<0$, and $d<0$. If the two-dimensional IC-NR state space is considered, some important conclusions can be drawn directly from equations (19) and (20): The subsystem remains at rest for $IC=0$, irrespective of the value of NR. Moreover, IC and NR being positive and d being negative, NR can only decline. The rise or decline of IC depends on the value of NR only: IC will rise if $b*NR + c>0$, and hence $NR>c/b$, but decline if the reverse is true. Therefore, as shown in Figure 9, the IC-NR plane can be divided into two parts (only values for IC and $NR \geq 0$ are considered): If $NR > -c/b$, IC will rise and NR decline; if $NR < -c/b$, both IC and NR will decline; and for $IC=0$, the system will be stable. Finally, for $NR=0$, IC will decline exponentially to zero. The arrows in Figure 9 show the directions in which the state vector will move. As a result, if the values of IC and NR are given, the analyst can immediately deduce the type of behaviour that will follow.

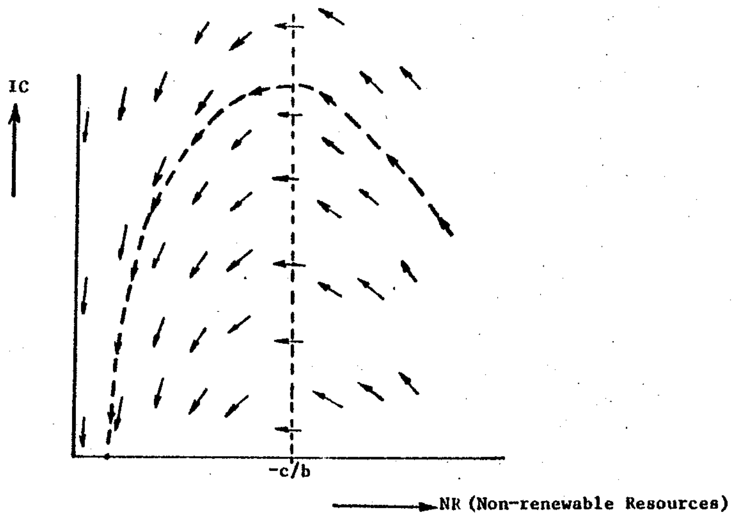


Figure 9: The IC-NR state plane (World3). The fat, dotted curve shows a possible trajectory, whereas the arrows indicate the direction in which the system will move at the specific value-combinations of IC and NR

For values of NR larger than -c/b, IC will rise first, and then decline, while in all other cases decline will set in immediately. Ultimately, IC will always tend to zero.

In this particular case, the equation describing the trajectory in the IC-NR space can be derived analytically from equations (19) and (20):

$$\dot{IC}/\dot{NR} = b/d + c/(d*NR). \quad (21)$$

This is an expression for the slope of the trajectory in the IC-NR plane as a function of NR. The expression for the trajectory can be found as the solution of

$$\frac{dIC}{dNR} = b/d + c/(d*NR). \quad (22)$$

The result is:

$$IC(t)-IC_o = (b/d)*(NR(t)-NR_o)+(c/d)*ln(NR(t)/NR_o), \quad (23)$$

where IC_o and NR_o are the starting values of IC and NR. The curve, sketched in Figure 9, according to equation (23) (fat, dotted line) shows that, unless NR_o=0, the system will always leave resources unused because IC will tend to zero before NR=0. Equation (23) can facilate further investigations, such as the direct calculation of NR(∞), or the calibration of b, c and d if a certain amount of resources has to remain.

IV.5 *ANALYSIS OF LINEAR APPROXIMATIONS*

Linearization of non-linear models, and subsequent analysis of the linear approximation is an extremely powerful technique which is widely utilized in the analysis and control of many dynamic systems. Because the linearization itself, and many of the techniques for study of linear systems can easily be implemented on a computer, application of the technique requires practically no effort, even for the study of models including numerous state variables. Linearization enables the analyst of using all the tools of linear system analysis and design.

Linear system theory provides, among other things, tools for studying the operating structure of linear models (decomposition according to the actual working of the equations), for deriving model hierarchy, for study of the basic stability properties of the whole model as well as of each subsystem, for performing model initialization in a correct way (see also |14|),for systematic order reduction of linear models, and for computation of the time constants (τ2 as well as τ3) included in a system. Moreover, when performed on a computer, linearization is more systematic than many other techniques since it automatically incorporates all relationships of the model.

However, as with all techniques based on approximation or simplification, the validity of the linear model is limited, and all conclusions drawn should be verified in the original model. Moreover, many of the techniques for analysis of linear models that clearly show basic model properties include mathematical operations that obscure the relation between the properties detected and the original model assumptions. Yet, the experience of many investigators tells us that linear models, even of highly nonlinear originals, can reveal many things. In negligible time, much information that may be extremely helpful in studying the original non-linear model can be generated.

Moreover, since computers are available, numerous linearizations and analyses can be easily performed for consecutive points in time, and changes in a model's operating structure caused by non-linearities may clearly come to light.

Let us now consider the technical details more closely. It is started from the state-variable description for a closed system* :

$$\dot{\underline{x}} = \underline{f}(\underline{x}(t)),\ \underline{x}(t_o)=\underline{x}_o. \tag{24}$$

A linear model that reflects the basic properties of the original model around a state $\underline{x}(T)$ can be deduced by a first-order Taylor approximation of $\underline{f}(\underline{x}(t))$ around $\underline{x}(T)$:

$$\dot{\underline{x}}(t) = \dot{\underline{x}}(T) + \underline{A} * (\underline{x}(t) - \underline{x}(T)). \tag{25}$$

where $\underline{A}$ is a matrix the elements A_{ij} of which are equal to

$$A_{ij} = \left[\frac{\partial f_i(\underline{x})}{\partial x_j}\right]_{\underline{x}=\underline{x}(T)} . \tag{26}$$

Thus, each element A_{ij} gives the sensitivity of $\dot{x}_i$ to small changes in x_j. The linear model (25) can be analysed in different ways. The stability properties and time constants τ3 may be derived by calculation of the eigenvalues of $\underline{A}$. The diagonal elements show the strength and sign of the net lag-free feedback of each state variable to its own rate of growth. The actual influence of x_j on $\dot{x}_i$ in the linear model may be evaluated by multiplication of the sensitivity A_{ij} by the changes that actually take place in x_j during a time period ΔT. This can be done for all elements, and then it follows from (25):

-----

* For ease of discussion, it is assumed that all exogenous influences are constant. Variations in exogenous influences, however, can be treated in a similar way.

$$\dot{\underline{x}}_i(T+\Delta T)-\dot{x}_i(T) = \sum_{j=1}^{n} (A_{ij}*\{x_j(T+\Delta T)-x_j(T)\}), \tag{27}$$

where n = the dimension of the system under consideration.
Now, if

$$A_{ij} = A_{ij}*\{x_j(T+\Delta T)-x_j(T)\}, \tag{28}$$

A'_{ij} shows the amount of change in $\dot{x}_i$ in the time period ΔT owing to the variations in x_j during the same time period. Comparison of all elements of each row of $\underline{A}$ teaches which couplings in the linear model are important and which not. Subsequently, $\underline{A}'$ can be simplified by neglecting all weak couplings, and a hierarchical structure of state variables may be derived automatically. Cuypers |1| and Schmidt |11| have demonstrated the utility of the approach in an application to Forrester's model. Recently, the technique has been extended, implemented and tested on a computer using the World3 model as a vehicle. Since it does not affect model properties, all third-order lags were replaced by first-order lags with the appropriate time constant (τ2) to reduce the number of state variables. The matrix $\underline{A}$ was calculated numerically using a difference approximation of (26). Routines have been implemented that rescale $\underline{A}$ according to (28), derive a Boolean matrix indicating the major couplings, reorder it hierarchically, and provide a graphical output of the results. Figure 10 shows the computer output obtained by linearizing World3 around its 1970 state. The matrix was rescaled according to (28) using the changes in each state variable that actually take place in the nonlinear model during the period 1970-2000 (ΔT=30 years)*. The corresponding hierarchical diagram is shown in Figure 11.

-----

* A test calculation has shown that the outcome of the linear model initialised in 1970, hardly differs from the original World3 model for about 30 years.

				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	IC	1	-17.5	*	*	*																		
2	SC	1	-5.5	*	*																			
3	AI	1	-1.0	*		*																		
4	EHSPC	2	-20.0		*		*																	
5	AIØPC	2	-3.0	*				*																
6	DIØPC	2	-20.0	*					*															
7	NR	2	*****	*																				
8	PAL	2	-15.7	*		*					*	*												
9	AL	2	-38.0	*		*					*	*												
10	UIL	2	-10.0	*									*											
11	PPAPR	2	-20.0	*		*								*										
12	FCFPC	3	-20.0	*	*			*							*									
13	PI	3	-7.0	*				*								*								
14	PPØL	3	-2.2											*			*							
15	P2	4	-22.3				*									*		*						
16	LFERT	4	-2.9														*		*					
17	PLE	5	-20.0	*		*	*											*		*				
18	P3	5	-14.0				*											*			*			
19	P4	6	-12.4				*														*	*		
20	PFR	6	-2.0			*						*				*		*			*			
21	LUFD	7	-2.0	*	*							*				*		*			*	*		*

Direction of the influences

Figure 10: Computer output of total linearization program applied to the World3 standard run between 1970 and 2000. The first column gives the row number, the second the letterscript indicating the variable's name (see Appendix for explanation), the third the hierarchical level, and the fourth the time constant τ2 associated with the particular variable, expressed in years. The asteriks show the major couplings in matrix form

The results are in full agreement with the knowledge about the model's working that had been obtained in many other, much more time and energy consuming ways. The capital and part of the agricultural system is virtually autonomous, and influences - directly and/or indirectly - the behaviour of all other variables.

Similarly, linearizations have been performed for other points in time during the World3 standard run. The results clearly show the changes in operating structure after 2020 owing to the nonlinearities in the capital and resource subsystem coming into play: resources rise also to the first level in the hierarchy. However, no further fundamental changes in operating structure are found, which illustrates that the capital and resource subsystem plays a leading part in the decline phase also.

Further elaboration of the technique of linearization, and discussion of conclusions drawn from its application is possible, but it would lead us too far beyond the scope of the present paper.

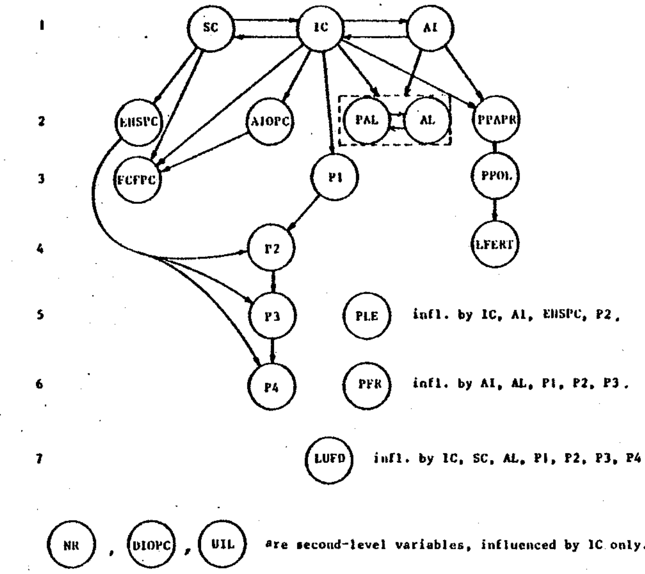


Figure 11: Influence diagram showing the hierarchy of the state variables in the World3 standard run between 1970 and 2000. The letterscripts are explained in the Appendix.

V. CONCLUDING REMARKS

Various methods of acquiring insight into the working of dynamic models have been presented. Yet, the list is far from being complete. Clearly, there is much room left for improvement in the field of understanding. For the time being, experience and intuition will be at least equally important as technical skills. Therefore, rash application of the tools described in this paper is discouraged since many of them may be useless or even dangerous in the hands of the inexperienced.

The author is greatly indebted to Mrs. J. Smulders, who managed to type and re-type the manuscript in almost negligible time.

REFERENCES

|1| Cuypers, J.G.M. and Rademaker, O.
An Analysis of Forrester's World Dynamics Model.
Automatica, Vol.10, Number 2(p.195-201), 1974.

|2| Forrester, J.W.
Industrial Dynamics.
The M.I.T. Press, Cambridge, Massachusetts 1961.

|3| Forrester, J.W.
World Dynamics.
Wright-Allen Press Inc., Cambridge, Massachusetts 1971.

|4| Gibson, J.E.
Nonlinear Automatic Control.
McGraw-Hill Book Company, New York 1963.

|5| Meadows, D.H. et al.
The Limits to Growth.
Universe Books/Potomac Associates, New York 1972.

|6| Meadows, D.L. et al.
Dynamics of Growth in a Finite World.
Wright Allen Press Inc., Massachusetts 1974.

|7| Peterson, D.W.
Statistical Tools for System Dynamics.
Proceedings 1976 International Conference on System Dynamics,
8-15 August 1976, Geilo, Norway.

|8| Rademaker, O. et al.
Project Group Global Dynamics: Progress Reports numbers 1 to 5,
1972-1976.*

|9| Rademaker, O.
On Understanding Complicated Models: Simple Methods.
Presented to the American-Soviet Conference on Methodological
Aspects of Social Systems Simulation, Sukhumi (U.S.S.R.),
October 24-26, 1973.*

|10| Rademaker, O.
The Behaviour of the Pollution in Forrester's World Model.
Proc. IFAC/UNESCO Workshop on Systems Analysis and Modelling
Approaches in Environment Systems,
Zakopane, Poland, Sept. 17-22, 1973.*

|11| Schmidt, G. and Lange, B.
Evolution of the Significant Signal Connections in Forrester's
Model of the World Dynamic (in German).
Regelungstechnik, vol. 23, number 52 (p.145-149), 1975.

------

* Available on request from O. Rademaker, P.O.Box 513, Eindhoven, Netherlands.

|12| Sharp, J.A.
Sensitivity Analysis Methods for System Dynamics Models.
Proceedings 1976 International Conference on System Dynamics,
8-15 August 1976, Geilo, Norway.

|13| Thissen, W. and De Mol, C.
An Analysis of the World3 Agricultural Submodel.
Presented at the Third I.I.A.S.A. Global Modelling Conference,
September 22-25, 1975, Baden, Austria.*

|14| Thissen, W.
Investigations into the World3 Model: The Capital and Resource
Subsystem.
IEEE Transactions on Systems, Man, and Cybernetics,
Vol. SMC-6, number 7 (p.455-466), 1976.

|15| Tomovic, R. and Vukobratovic, R.
General Sensitivity Theory.
Elsevier, New York, 1972.

------

* Available on request from O. Rademaker, P.O.Box 513, Eindhoven, Netherlands.

APPENDIX : List of letterscripts, and their associated meaning.

A	: Agricultural subsystem
AHL70	: Assimilation Half-Life in 1970
AI	: Agricultural Inputs
AIOPC	: Average Industrial Output Per capita
AL	: Arable Land
CI	: Capital Investments
CIAF	: Capital-Investment-in-Agriculture Fraction
CR	: Capital and Resource subsystem
CUF	: Capital Utilization Fraction
DIOPC	: Delayed Industrial Output Per Capita
EHSPC	: Effective Health Services Per Capita
F	: Food
FCFPC	: Fertility Control Facilities Per Capita
FIOAA	: Fraction of Industrial Output Allocated to Agriculture
IC	: Industrial Capital
ICOR	: Industrial Capital-output Ratio
ILF	: Inherent Land Fertility
IOPC	: Industrial Output per Capita
LFDR	: Land Fertility Degradation Ratio
LFERT	: Land Fertility
LFRT	: Land Fertility Regeneration Time
LUFD	: Labor Utilization Fraction Delayed
NR	: Non-renewable Resources (World3); Natural Resources (World2)
NRI	: Non-renewable Resources Initial
NRUR	: Non-renewable Resource Usage Rate
P1	: Population, Ages 0-14
P2	: Population, Ages 14-44
P3	: Population, Ages 45-64
P4	: Population, Ages 65^+
PAL	: Potentially Arable Land
PCRUM	: Per Capital Resource Usage Multiplier
PFR	: Perceived Food Ratio
PLE	: Perceived Life Expectancy
POP	: Population (subsystem)

POL	: Pollution (World2)
POLA	: Pollution Absorption (World2)
POLAT	: Pollution Absorption Time (World2)
POPI	: Population Initial
PPAPR	: Persistent Pollution Appearance Rate
PPOL	: Persistent Pollution (World3)
PPOLX	: Index of Persistent Pollution
SC	: Service Capital
SO	: Service Output
t,T	: Time
$\tau 1$	: Time constant (loop)
$\tau 2$	: Time constant (state variable)
$\tau 3$	: Time constant (overall behaviour)
UIL	: Urban-Industrial Land