5

Critical steps in systems simulation

A.A. Dijkhuizen, A.W. Jalvingh & R.B.M. Huirne
Department of Farm Management, Wageningen Agricultural University, Wageningen, the Netherlands

Objectives

From this chapter the reader should gain knowledge of:

- the critical steps to be taken in systems simulation (ie, definition of the system and statement of objectives, analysis of data relevant to the model, model construction, validation of the model, sensitivity analysis, and application of the model)
- how to make a better choice from modelling types and techniques, especially with respect to deterministic and stochastic models

5.1 Introduction

Models are essential tools for understanding animal health economics. Mathematical models are especially useful in this context and generally defined as a set of equations to describe or simulate an interrelated part (system) of the real world (Hillier & Lieberman, 1990). Three broad functions can be distinguished: (1) to provide an objective basis for assessing and assimilating available information about the system, (2) to detect where essential knowledge of the system is lacking or inadequate, indicating needs for further research, and (3) to assist in the management control of the system.

Basically, there are two different modelling approaches to be considered: a positive approach and a normative approach. The **positive approach** can best be indicated as a description of relevant processes and characteristics by statistical/epidemiological data analysis (the so-called empirical modelling). Traditionally, research in livestock production has mainly been conducted in this way. In animal health economics more attention is paid to the **normative approach**, which includes computer simulation techniques (the so-called mechanistic modelling). Computer simulation is a method for analysing a problem by creating a simplified mathematical model of the system under consideration which can then be manipulated by modification of inputs. This method is especially attractive where real-life experimentation would be impossible, costly or disruptive (eg, with highly contagious diseases), and for exploring strategies that have not been applied yet. Special attention has to be paid to the correspondence between model and reality to obtain meaningful results for real-world situations.

In this chapter the critical steps and basic concepts in systems simulation are presented and discussed.

5.2 Systems and systems analysis

The terminology associated with systems and systems analysis is generally a collection of terms that are used in other fields often with different meaning or connotation. In the model ling context, a system is generally described as a - complex - set of related components which exist within some pre-defined boundary and react as a whole to external or internal stimuli (eg, animal, herd, population). Placing of the boundary is considered the key issue

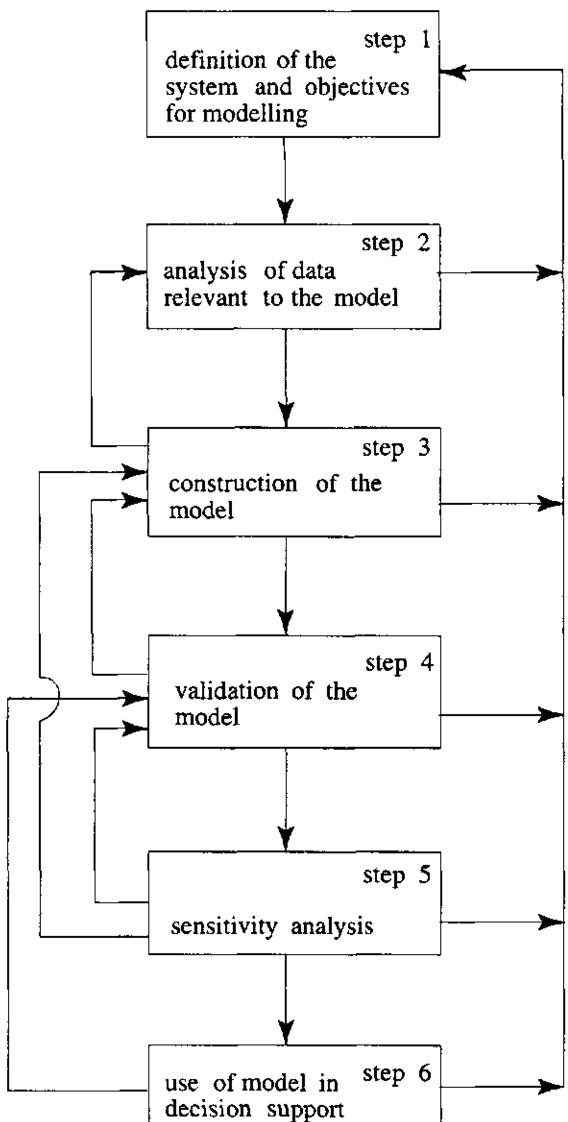

Figure 5.1 The basic steps of systems simulation

in defining and structuring any system, and should depend primarily on the function the model has to fulfil (Dent & Blackie, 1979).
The term systems analysis is generally used to refer to the process of examining complex systems, where all major inputs and outputs are accounted for by the use of mathematical models. Dent & Blackie (1979) consider six critical - and interlinked - steps involved in applying systems analysis, as presented in Figure 5.1. The steps are commented on one by one.

Step 1 Definition of the system and objectives for modelling
A clear description of the system and statement of the reasons why the system simulation work is being carried out is an essential first step. The system under consideration, the nature of the problems to be solved, the relevant data available, and to what degree of detail answers are required highly determine the type of model to be used. Different types of models are available to simulate a system (Law & Kelton, 1991). A first choice that should be made is that between **static** and **dynamic** models. A static model does not contain time as a variable and, therefore, cannot simulate the behaviour of a system over time, as opposed to a dynamic model. A model that makes definite predictions for quantities (such as milk production and live weight) is called **deterministic**. A **stochastic** model, on the other hand, contains probability distributions to deal with uncertainty in the behaviour of a system. These distributions can be used directly or through random sampling. In the latter case, repeated runs of the model are necessary to provide insight into the variation in outcome. A final difference to consider concerns **optimization** versus **simulation**. An optimization model determines the optimum solution given the objective function and restrictions, whereas a simulation model calculates the outcome of pre-defined sets of input variables (scenarios, strategies)

Step 2 Analysis of data relevant to the model
The model design is to a large extent dependent on the data available or on the feasibility of generating data within the time limits set by the research. Complete data availability will seldom, if ever, exist. An obvious shortage occurs when simulation studies are conducted in an early stage of research, eg, to explore new strategies that have not been applied yet. Especially then, however, simulation studies have proved to be beneficial to help structure the problem and set priorities for further (empirical) research. In those cases a close cooperation with experts may help to get the best estimates for the necessary input data and relationships. Once the model is available, the - potential - impact of uncertain estimates can easily be determined through sensitivity analyses (Step 5).

Step 3 Construction of the model
The construction of a mathematical model is usually a multistage procedure. Three functionally different approaches can be distinguished:

- the bottom-up approach, beginning with components of models at the lowest level of organization and combining them without any aggregation;

- the top-down approach, which begins with a simple representation of the entire system and is complete when the resolution of the model is sufficient to satisfy the objectives; and
- the prototyping approach, an iterative compromise between the first two alternatives.

Development of a model with the prototype approach begins with simple modelling of single subsystems. The process of development progresses by formulating more sophisticated representations of the most important subsystems and aggregating, deleting or ignoring subsystems of lesser importance. Because of its flexibility, the prototyping approach is especially favourable for models of large and complex systems, such as livestock production systems. It allows experts (as well as final users) to be included in the modelling process at an early stage. Regular interaction with these people maintains their interest in the simulation study. It can also help to avoid the mistake (often made by novice modellers) to start with too excessive an amount of model detail.

Step 4 Validation of the model

Validation is considered a very important but difficult step in the entire modelling procedure. The key issue here is to judge whether or not the model mimics reality sufficiently well to fulfil the purposes for which it has been developed. If a model is considered 'valid', then the decisions made with the model should be similar to those that would be made by physically experimenting with the system (if possible). If a model is not valid, then any conclusions derived from it will be of doubtful value.

In the literature a distinction is made between internal and external validation (Taylor, 1983). **Internal validation** is a continuous process throughout the development stage of the model, ensuring that all assumptions are in accordance with the theory, experience and relevant general knowledge. Internal validation can thus be described as ensuring that the right answer, decision or recommendation is provided by the correct method, and that each equation or part of the model has a logical basis, uses correct parameters and is correctly written. **External validation** refers to the comparison of the model's performance against the performance of the real system, in which the model is considered a 'black box'. Information should be produced, which enables the user to conclude whether to accept or reject the model's recommendation. This may include a sensitivity analysis (Step 5).

Two fundamental issues relate to the validation of any computer simulation model. First, the fact that a model behaving like reality for one set of inputs and operating rules does not guarantee that it will perform satisfactorily for a different set of conditions. Second, there is no totally objective and accepted approach to validation, because validation necessarily includes: (1) the way(s) in which the model is used, (2) the tests with which to validate the model, (3) the data to serve as a basis for comparison, and (4) the criteria to measure the (in)validity of the model. When a model and the results are accepted by the user as being 'valid', and are used as an aid in making decisions, then such a model is called 'credible' (Law & Kelton, 1991). Although credibility has not been discussed a great deal in literature on systems simulation, it is considered as important as validation in terms of actual implementation of simulation results.

Step 5 Sensitivity analysis

One of the most powerful decision-analytic techniques is **sensitivity analysis**, in which the values of relevant parameters are systematically varied over some range of interest to determine their impact on the results. If each of the assumptions is independent of all other assumptions, then it is reasonable to vary one parameter at a time, assuming all other parameters to be at their baseline or 'best guess' values. On the other hand, if several assumptions are interdependent, or if it seems important to examine the trade-off between specific gains and losses, then it is best to examine several parameters simultaneously.

Good knowledge of sensitive parameters should be available and entered into the model. If this is not available, sensitivity analysis can help to set priorities for further (empirical) research. In this way a valuable interaction between computer simulation and field data analysis is possible. Computer simulation may be used to quantify the significant gaps in (veterinary) knowledge, while knowledge obtained from field data research increases the realness of economic models. This interaction is considered fundamental to the study of disease and disease control.

Step 6 Use of model in decision support

If accepted, the model can be used for providing answers to the questions for which it has been built. This can be done either within a research environment only (ie, providing output for scientific publications) or as an actual tool for decision-support in the field. The latter use is on the increase. A so-called **Integrated Decision Support System** (IDSS) is commonly considered to be an ideal framework for model use under field conditions, and defined as a user-machine system for providing information to support operations, management and decision-making functions in an organization. The system utilizes computer hardware and software, manual procedures, models for analysis, planning, control and decision making, and a database (Davis & Olson, 1985). Marsh (1986), Huirne (1990), Jalvingh (1993) and Houben (1995) developed comprehensive but flexible models for on-farm decision support in the area of animal health and replacement economics in dairy cattle and swine, meant to be included in the model base of such an IDSS.

Different ways can be considered to actually make these models available for use in the field. As a start, the models could be made available for a central computer, which can be accessed by individual users (farmers, advisers). In case of indirect use, the adviser carries out the model calculations and interprets and reports the results to the farmer (intermediary mode). In case of direct use, the farmer and the adviser may use the model on-line (terminal mode), or off-line (clerk mode) by preparing and submitting input to the central computer. The final step is to have the models available on the PC of both the adviser and the farmer.

5.3 Deterministic and stochastic modelling

In the previous section a deterministic model was described as being a model that makes definite predictions for quantities (such as milk production and live weight), whereas a stochastic model contains either probability distributions or random elements to deal with uncertainty in the behaviour of a system. In the literature there is agreement on the distinction between (pure) deterministic models on the one hand and stochastic models containing random elements on the other. This is not the case with respect to stochastic models containing probability distributions. Several authors classify these models as being deterministic, which is not correct and may underestimate their value and potential applications. Therefore, the three different types of models will further be illustrated and discussed, by using a simplified example.

Consider the situation where for a model dealing with sow replacement economics, a young replacement sow is taken from a population where the size of the first litter is normally distributed (see Figure 5.2). Litter size in this example ranges from 6 to 12 pigs and most sows (ie, 30%) fall into the class with 9 pigs per litter. The expected performance of a replacement sow taken from this population depends upon the type of model under consider ation. In case of a **deterministic model**, each replacement sow is considered to produce exactly 9 pigs in her first litter, this being the most likely litter size of the population she comes from. All costs and returns in the calculation then are also derived from such a 9-pig sow. In a **stochastic model with probability distributions**, each replacement sow is expected to produce 0.05 x 6 + 0.10 x 7 + ... + 0.10 x 11 + 0.05 x 12 = 9.0 pigs in the first litter. Now the costs and returns in the calculation are derived (ie, weighed) from the various single litter-size classes under consideration. This is a fundamentally different approach from the deterministic model and will lead to different economic results as well. In case of a **stochastic model with random elements**, each replacement sow is randomly drawn from the specified probability distribution and will have a litter size of 6, 7, ..., or 12 pigs. The costs and returns, therefore, will differ accordingly between the sows. With a sufficiently

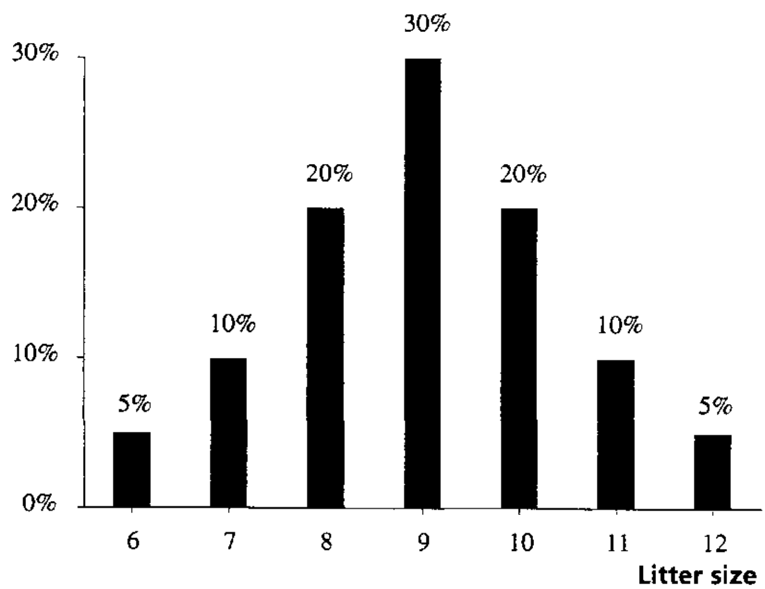

Figure 5.2 First-litter size performance in a sow population

large number of drawings (ie, multiple runs) the average litter size of the replacement sows will approximate 9.0 pigs again, providing the same average economic outcome as the stochastic model using probability distributions.

As illustrated in the example, deterministic models do not take into account uncertainty (ie, variation) associated with future events, resulting in an oversimplification of the conditions under which decisions have to be made in reality. The approach using probability distributions (ie, dynamic probabilistic or (modified) Markov chain simulation) and the one using random numbers (ie, Monte Carlo simulation) have the advantage that, for instance, animals with different performances can be treated differently, eg, a more liberal insemination and replacement policy for high-producing animals. Moreover, future performance can be related to current performance. Therefore, culling of animals with a low performance will influence the performance realized in later production cycles. In the deterministic approach this is not possible; the resulting average performance per production cycle is always equal to the input value.

One advantage of the approach using probability distributions is that there will be observations in all classes, which means that the model will exactly provide the expected value of the results and only **one run** is needed to obtain these results. In fact, the results of a very large herd or population are simulated, with animals available in all possible states. In the model using random numbers, the presence of observations in all classes is not guaranteed. The larger the number of observations, the better the average result will approach the real expected value. Replicated calculations are needed to obtain a reliable estimate of the average results.

One advantage of models with random numbers and multiple runs, on the other hand, is the available information about the expected **standard deviation** in the results, which allows for statistical tests and non-neutral risk analysis. Performing these tests and analyses requires a careful design and analysis of the modelling experiments, in order to obtain reliable estimates of average results and standard deviations. By simply choosing a large number of replications, for instance, a difference between two strategies can always be made significant, due to the fact that the standard error of the mean will then be small.

Applications of stochastic models with random elements often focus on the comparison of average results only, rather than on a trade-off between expected outcome and its variation (ie, non-neutral risk analysis). If so, then the approach using probability distributions had better be applied. Since one run supplies the expected value of the results, various sensitivity analyses can be carried out much easier than is the case with models using random numbers.

An overview of published models and their features in the area of reproduction and replacement economics in dairy cattle and sows was given by Jalvingh (1993).

5.4 Common combinations of modelling type and technique

The choice of the modelling type and technique will depend on a number of factors, including:

- the nature of the problem;
- the resources available, eg, time, money and analytical tools; and

- the availability of data and information about the problem.

Even within specific narrow problem domains such as animal replacement decisions, diffe rent modelling types and techniques are used. Most common combinations in the literature are summarized in Table 5.1.
Model calculations in animal health economics often suffer from a serious lack of accurate data, as discussed before. Since the mid-eighties much effort has been put into designing and implementing integrated veterinary, zootechnical and economic record keeping systems (Morris & Dijkhuizen, 1992). In the future, systematic epidemiological and economic analyses of these databases should be given high priority. A basic question is whether - and, if so, which - standards are available to express the frequency and severity of the various diseases under field conditions. Further research in this field is necessary and can be of great practical value. In this way a valuable interaction between economic research on the one hand and veterinary and zootechnical research on the other is possible.

Table 5.1 Common combinations of modelling type and technique[a]

				Stochastic			
	Static	Dyna-mic	Determi-nistic	Probab. distrib.	Random elements	Optimization	Simulation
Gross Margin Analysis	x		x				x
Partial Budgeting	x		x				x
Cost-Benefit Analysis		x	x				
Decision Analysis	x			x			x
Linear Programming	x		x			x	
Dynamic Programming		x		x		x	
Markov Chain Simulation		x		x			x
Monte Carlo Simulation		x			x		x

[a] The first four (basic) methods of economic analysis in this table were explained in Chapter 3; the other, more advanced ones, will follow in the next chapters.

References

Davis, G.B. & Olson, M.H., 1985. Management information systems. Conceptual foundations, structure and development. McGraw-Hill, New York, 693 pp.

Dent, J.B. & Blackie, M.J., 1979. Systems simulation in agriculture. Applied Science Publisher Ltd., London, 180 pp.

Hillier, F.S. & Lieberman, G.J., 1990. Introduction to operations research. McGraw-Hill, New York, 954 pp.

Houben, E.H.P., 1995. Economic optimization of decisions with respect to dairy cow health management. PhD-Thesis, Department of Farm Management, Wageningen Agricultural University, Wageningen, 146 pp.

Huirne, R.B.M., 1990. Computerized management support for swine breeding farms. PhD-Thesis, Department of Farm Management, Wageningen Agricultural University, Wageningen, 165 pp.

Jalvingh, A.W., 1993. Dynamic livestock modelling for on-farm decision support. PhD-Thesis, Department of Farm Management and Department of Animal Breeding, Wageningen Agricultural University, Wageningen, 164 pp.

Law, A.M. & Kelton, W.D., 1991. Simulation modeling and analysis. McGraw-Hill, New York, 759 pp.

Marsh, W.E., 1986. Economic decision making on health and management in livestock herds: examining complex problems through computer simulation. PhD-Thesis, University of Minnesota, St. Paul, 304 pp.

Morris, R.S. & Dijkhuizen, A.A., 1992. The application of computer software in the management of fertility disorders in pigs. Proceedings 12th Congress of the International Pig Veterinary Society, Amsterdam, I: 45-48.

Taylor, A.J., 1983. The verification of dynamic simulation models. Journal of Operational Research Society 34: 233-242.