1

Introduction to Mathematical Modelling

The definitions of mathematical modelling of systems, along with analysis, synthesis and compensation of a system are given in this chapter. The basic steps in the development of model equations are given. The classification of model equations and the types of mathematical equations encountered are given. The need for the black box modelling and grey box modelling and their limitations are given.

1.1 MATHEMATICAL MODEL

Description of a system by mathematical equations is called mathematical modelling. It is important to capture the essential features of a system to describe (design), forecast (predict), optimize the operating conditions and to design a suitable controller. Some of the application areas of model equations are in process design, process control, process safety, operating training simulators, and environmental impact assessment, etc. Each application area may require different form of mathematical model equations. Mathematical modelling method involves analysis, numerical simulation, followed by experimental tests. Development of mathematical model needs general laws and constitutive relations. The general laws are the conservation or balance equations of mass, momentum and energy. Some of the terms or variables in the general conservation equation are to be obtained by the constitutive equation(s).

The constitutive relations are experimental in nature and strongly depend on the phenomenon under consideration. Some examples include Fourier's laws of heat conduction, Fick's law of diffusion of a substance, reaction rates, equilibrium relations and equation of state, Newton's law of viscosity, isotherms for adsorption, hold up of catalyst, etc. This is also known as phenomenological law, i.e., a law which is arrived at by observing natural phenomena and not derived from the basic principles. The mathematical equations are algebraic equations, ordinary differential equations, partial differential equations, etc. A detailed physical insight has to be gained before the mathematical description can be formulated.

In general, the development of a completely exact model is rather difficult. Sometimes the theory of the phenomenon is not completely known; sometimes the experimental facts are not available; sometimes there is no need to include all details—often we simply look for the trends. Oftentimes, the analytical solution of differential equations is rather difficult or impossible because the equations are nonlinear or several coupled equations constitute the model. In such cases, the use of the numerical solutions using computers, or specifically designed software package (as standard subroutines) is desirable to get the solution.

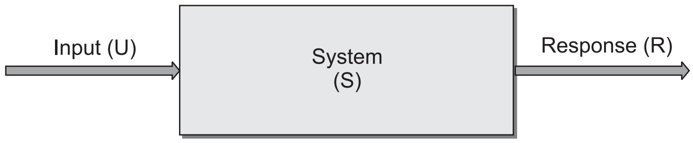

Fig. 1.1 Input, system and response

The three general problems we face with respect to any system (Fig. 1.1) are:

Synthesis: given the input U and desired response (R) find S (design the system to get the desired output for the given input). The synthesis problem is known as the *design problem.*

Analysis: Given U and S, find R (given the system and the input find the output). The analysis problem is known as the *prediction* problem.

Compensation: given S and R, find U (given the system and the desired output, find the input). This problem is called *control problem*.

If we are good at the analysis problem, then the other two problems (synthesis and compensator) can be solved by reformulating the problem as an analysis problem. For example, in the synthesis problem, we assume the system and for the given input, by the analysis we can find the output. If this output is matching with the given output, then the assumed system is correct. Otherwise, we have to modify the system and repeat the problem. Similarly, for the compensator problem, we assume the input and by analysis problem calculate the output. If this output matches well with the given output, then the assumed input value is correct. Otherwise, the input value is to be changed, and the procedure is repeated. How to change the input depends on how good we are at the analysis problem.

In many cases, theoretical treatments are available for models for which there is no perfect physical realization is available. In this case, the only possible test for an appropriate theoretical solution is to compare with the 'data' generated from a computer simulation. An important advantage of the simulations is that different physical effects, which are present simultaneously in real systems, may be isolated and, through separate consideration by simulation, may provide a much better understanding. Simulations of the simplified models can 'switch off' or 'switch on' these effects and thus determine the particular consequences of each contributing factor. We wish to emphasize that the aim of simulations is not to provide better 'curve fitting' to experimental data than does by the analytic theory.

The goal is to create an understanding of the physical properties and processes which is as complete as possible, making use of the perfect control of 'experimental' conditions in the 'computer experiment' and of the possibility to examine every aspect of system configuration in detail. The desired result is then the elucidation of the physical mechanisms that are responsible for the observed phenomena. The relationship among the theory, experiment and simulation is similar to those of vertices of a triangle, each is distinct, but each is strongly connected to the other two.

Empirical models: The method is easy to develop. The idea is to fit a curve through a set of data and use this curve in order to predict the outcome. We may not be confident that the method is applicable outside the range (extrapolation) .

Stochastic model: ('stochastic' in Greek means to guess). We try to estimate the probability of certain outcomes based on the available data.

1.2 DEVELOPMENT OF MATHEMATICAL MODEL

$$\text{Knowledge of the system} + \text{Mathematics} \rightarrow \text{Modelling} \quad (1.1)$$

To develop a mathematical model of a system, we must know both the subject knowledge and mathematics. Subject knowledge must both precede and follow the mathematical modelling to ensure that the essential features of the system are taken into account and to check the mathematical model solution gives a meaningful solution.

Mathematical modelling is an art. It requires experience, insight and understanding. Teaching this art is also another art. Criteria for a successful model include a good agreement between the prediction and the observations, of drawing valid conclusions and simplicity of the model . Modelling forces us to think clearly about the system. We may need different models for explaining different aspects of the same situation or even for different ranges of the variables. The search for a unified model continues. Sticking to one model may prevent the insight. Comparison of the prediction with the observations reveals the need for new experiments to collect the needed data. Mathematical models can also lead to the development of new concepts. Every model contains some parameters, which are to be estimated. The model must itself suggest experiments. The parameter, if possible, is to be calculated independently. Sometimes, new mathematical methods have to be developed. If the model is oversimplified, it may not represent the reality. If the model is overambitious, then the mathematical model equations may need very complicated methods of solutions. Mathematical models are constantly updated (or improved), to make them more realistic. Thus, we need to develop hierarchy of models, each of which is applicable to increased complex situation.

The modelling process starts (Fig. 1.2) with the analysis of the problem for idealization and for various assumptions. Once the idealized model is formulated, it can be then translated into the corresponding mathematical model in terms of differential equations. The modelling step involves idealization and simplification of the real world problem, to develop a simple model to extract the essential features of the system. From the physical model, we can use the first principles such as the conservation of mass, energy and momentum to translate in to mathematical equations. Then, the model equations will be solved to find the prediction under appropriate conditions. This step will involve in solving the model equations analytically or numerically by appropriate techniques. The process is known as the simulation. The predictions will be validated against the existing models, or well established bench-mark and the experimental data.

If the results are satisfactory (rare at first attempt), then the mathematical model can be accepted. If not, both the assumptions and idealization of the physical model and mathematical modelling will be modified based on the feedback and then the new simulation and prediction will be carried out.

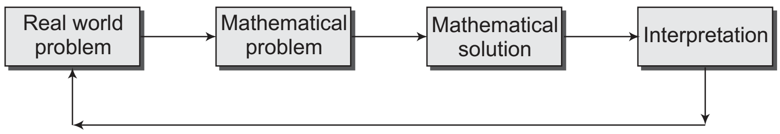

Fig. 1.2 Principle of mathematical modelling (Bender, 1978)

The mathematical model equations must be verified. This is done using a prototype and an experimental verification of the model behaviour. It is also desirable to consider special limiting cases to ensure that the model has the proper behaviour. Every term in the mathematical model must have the same units as in every other term. The algebraic sign of the term makes sense in the overall context of the equation. The resources available consist of available man-power, time, the availability of computing resources, and the software programs. It is valuable to start with a simple model having a limited purpose, and then to improve upon it as required. Table 1.1 gives the classification of the model and Table 1.2 gives the forms of model equations encountered.

Table 1.1 *Model classification (Hangos and Cameron, 2001)*

Group of models	Classification	Criterion of classification
I	Mechanistic	Based on mechanism/underlying phenomena
	Empirical	Based on input-output, trials or experiments
II	Stochastic	Contains model elements that are probabilistic nature
	Deterministic	Based on cause effect analysis
III	Lumped parameters	Dependent variables not function of special position
	Distributed parameter	Dependent variables are a function of special position
IV	Linear	Super position principle applies
	Nonlinear	Super position principle does not apply (occurrence of product of dependent variables and/or its derivatives)
V	Continuous	Dependent variables defined over continuous space time
	Discrete	Only defined for discrete values of time/ or space
	Hybrid	Containing continuous and discrete behaviour

Table 1.2 *Forms of model equations (Hangos and Cameron, 2001)*

Type of model	Equation Types	
	Steady state problem	**Dynamic problem**
Deterministic	Nonlinear algebraic	ODEs/PDEs
Stochastic	Algebraic/difference equation	
Lumped parameter	Algebraic equation	ODEs
Distributed parameter	Elliptic PDEs	Parabolic PDEs
Linear	Linear algebraic equations	Linear ODEs
Nonlinear	Nonlinear algebraic equations	Nonlinear ODEs
Continuous	Algebraic equations	ODEs
Discrete	Difference equations	Difference equations

ODEs: Ordinary Differential Equations; PDEs: Partial Differential Equations

1.3 SIMULATION

Simulation is the process of analyzing a whole process or a part of it, using the model equations. The purpose of simulation is to find the optimal operating conditions/parameters, analyzing the effect of the input variables on the performance of the system and design of controllers, Simulation is defined as experimentation with the mathematical equations/models. Solving the mathematical model equations to understand the behaviour of the system is called simulation.

1.3.1 Nonlinear differential equations

Whenever there is a product of dependent and/or its derivative terms present in the equation, then the system is called nonlinear. Examples are:

$$[dy/dx]^2 + y = 0 \tag{1.2}$$

$$y[d^2 y/dx^2] + dy/dx + y = 0 \tag{1.3}$$

$$[d^2y/dx^2] + dy/dx + \exp(y) = 0 \tag{1.4}$$

The boundary conditions are usually linear.

The nonlinearity is with respect to the dependent variable (and its derivative). If any term present in the right side of equation [as a function of independent variable], then it is called a non-homogeneous equation. If the coefficient is nonlinear in the independent variable then the equation is called as a variable coefficient equation:

$$(d^2y/dx^2) + (x^2+x)\,(dy/dx) + y = 0 \tag{1.5}$$

We have considered here the ordinary differential equations. Similarly, we can write the partial differential equations.

Autonomous system:

$$dy_i/dt = f_i\,(y_{1,}\, y_2,\, y_3),\ i = 1, 2, 3 \tag{1.6}$$

The above model equation, without the appearance of independent variable in the function f, is called an autonomous system (also known as a time invariant system).

1.4 CONSERVATION OF MASS/ENERGY/MOMENTUM

Input – output = accumulation (transient condition) (1.7)

Input – output = 0 (steady state condition) (1.8)

The input terms represent the mass entering the system and the mass generated by the reaction or mass transferred from other phase to the phase under consideration. Similarly, the output term consists of mass leaving out, mass consumed by the reaction, mass transported from this phase to another phase. Total mass balance and the component mass balance in each phase are to be written. The conservation equation for the mass of the i-th chemical species at steady state condition is given by

Rate of mass of the i-th species in – rate of mass of the i-th species out + rate of generation of mass of the i-th species in the system = 0 (1.9)

The mass of i-th species may enter or leave the system by two means: (1) by inlet or outlet stream; and (2) by exchange of mass between the system and its surroundings through the boundaries of the system, i.e., inter-phase mass transfer. If it is a reacting system, we need to know the reaction kinetics for the reactions.

The conservation statement for total energy under steady state condition takes the form

Rate of energy in – rate of energy out + rate of energy generated in the system = 0 (1.10)

Energy must enter or leave the system by two means: (i) by inlet or outlet steams; (ii) by exchange of energy between the system and its surroundings through the boundaries of the system in the form of heat and work.

Depending upon the situation, the modelling may require only the mass balance and/or energy balance and /or momentum balance.

Rate of momentum in – rate of momentum out + rate of momentum generation = 0 (1.11)

Rate of momentum generation is equal to the summation of forcing on the system. The basic steps in the development of the balance equations are given by:

Define the system: if possible, draw the sketch; list the suitable assumptions made; write down the mass balance and /or energy balance and/or momentum balance equations.

Use the appropriate correlation to calculate the mass transfer/energy transfer coefficients relevant to the problem at hand: write the initial conditions and solve the ordinary differential equations.

Let us discuss on the classification of models given in Table 1.1

(1) First principle modelling versus empirical modelling

The first principles modelling (or white box or mechanistic modelling) is based on the conservation laws (of mass/energy/momentum). For the transport coefficient and the reaction kinetics, we use the empirical relationships. The mechanistic models have an advantage of preserving the physical insight into the process as the model variables coincide with the process variables. A major drawback of the physical modelling is that it is time demanding approach.

In empirical modelling (or black box modelling), an input–output relationships is constructed using the experimental data of the process. Whenever there is a lack of knowledge of the process and /or time limitation to develop a model, the empirical modelling is desirable. A disadvantage of this approach is that it may require extensive experimentation to obtain the input–output data of sufficient quality for a newly developed process or a process under development. Extrapolation to outside the experimental data condition is not recommended. Mathematical representation include time series models such as Auto Regressive Moving Average (ARMA), Auto Regressive with eXogenous input (ARX), Output Error Model (OE) models; artificial neural network (ANN) models; Fuzzy models and partial least square (PLS) models, etc. The model parameters have no direct relationship to that of the first principles.

2) Lumped versus distributed parameters models

The model equations should be as simple as possible. Hence, the lumped parameter models are preferred over the distributed parameter models (described by partial differential equations) as they are simple and computationally less demanding. However if the variable (say, temperature) is varying along the spatial coordination, it may be desirable to develop a DPS (Distributed Parameters System) models. For controller design purposes, the DPS model should be converted in to a set of LPM and hence the number of equations is larger for the design of the control systems.

3) Linear versus nonlinear models

For analysis of the processes and optimization of the systems, appropriate nonlinear model is required. For purpose of designing controllers, a linear model may be adequate. Sometimes, switching over linear models, for tuning the controllers, is recommended. For a highly nonlinear process such as pH process, a nonlinear model based control law is required. The nonlinear models may be difficult to solve analytically.

4) Steady-state versus dynamic models

For continuous time processes, to design a process or optimization of the operating variables we require steady state models. For analyzing the oscillatory behaviour, we may need the unsteady state models. For design of the controller also we need dynamic models. For batch processes, the dynamic model is required.

5) Continuous versus discrete-time models

Usually the continuous time modelling is required to analyze the behaviour of the process. For design of controllers, particularly adaptive control or optimal control, etc., we may need discrete time models. Some numerical solution of ODE models may require discrete models.

6) Deterministic models versus stochastic models

It is traditional to model the nature in the deterministic terms. However, neither the nature nor the engineered systems behave in a precisely predictable fashion. Systems are almost always inherently 'noisy'. Therefore, in order to model a system realistically, a degree of randomness is to be considered in the model. Even though one cannot precisely predict a next event, one can predict how next events will be distributed. Unlike the traditional data analysis, where the statistics (such as mean, standard deviation, and the like) are calculated from the given data, we can generate here a set of data having pre specified statistics. The actual input signal will never be replaced, but its statistics will most likely still be stationary.

1.5 RANDOM PROCESSES

In simulating dynamic processes, it is important to model not only the system itself but also the input data that drives the system. Input signals are rarely deterministic. In such cases, we call them random processes because the input signals have an inherent fluctuation built around the underlying unpredictable processes.

In most stochastic models, the system *per se* is fixed and deterministic. Even though they might vary with time, systems are usually well behaved and not subjected to erratic fluctuations. In contrast, the signals that drive the system often appear somewhat random and noisy. Signals are in essence time series of statistically related processes, while systems are simply devices that transform the processes. The stochastic models are very similar to deterministic models except that a random process is the input to the system and the effects are analyzed statistically.

Deterministic models have finite sets of input–output vectors. Each signal is simulated by an explicit formula expressed as a function of time. In contrast, stochastic signals do not have explicit formulas. Instead, the stochastic signals are characterized by defining statistical time series descriptions such as the auto correlation or the spectral density. Another major difference between deterministic and stochastic systems lies in their long-term behaviour. Deterministic systems tend to produce output signals with regular steady state qualities such as constant, periodic, or chaotic states. Stochastic systems are usually random, even in the long run. Although the randomness still has no explicit definition, the output can still be characterized by the auto correlation and spectral density functions.

Table 1.3 *Comparisons of deterministic and stochastic signals (Severance, 2005)*

Signal	Deterministic	Stochastic
Input	Single input vector	Ensemble of input vectors
Output	Single output vector	Ensemble of output vectors
Analysis	Transient Phase	Initial non-stationery phase
	Steady state phase	Stationary phase
Defining input	Impulse	White noise
Signal descriptor	Explicit formula	Autocorrelation
		Spectral density

Table 1.3 shows several of the characteristic difference between deterministic and stochastic systems. In the deterministic case, the output can be a constant steady state or periodic. One of these two cases will always occur in a stable linear system. In a nonlinear system, its behaviour may be more exotic and perhaps even chaotic. A deterministic system is described using the formulas that are explicit functions of time. On the other hand, stochastic systems seem to have underlying deterministic behaviour with the randomness superimposed. In the case of linear systems, the output signal will have steady states analogous to those of the deterministic systems. If the steady

state appears essentially constant with a fixed amount of randomness superimposed, the signal has achieved the stationary condition, since all statistics are the same for each time slice. However, if there is no stationary, differing time slices have different averages.

1.6 BLACK BOX MODEL AND GREY BOX MODEL

The modelling of the system by the first principles model (physical, mechanistic) is known as a white box modelling. The requirement for a good understanding of the physical back ground of the problem at hand proves to be a severe limiting factor in practice when complex and poorly understood systems are considered. Difficulties encountered white box modelling can arise, for instance, from the poor underlying phenomena, inaccurate values of various parameters, or from the complexity of the resulting model.

In such cases, the process is described by a general ‘black box’ structure used as a function approximation. In black box modelling, the structure of the model is hardly related to the structure of the real system. The identification problem consists of estimating the parameters in the model. If representative process data is available, the black box model can be developed quite easily, without requiring process specific knowledge. Time series models, polynomial models, neural network models fall under the category. A severe drawback of this approach is that the structure and parameters of these models usually do not have any physical significance. Such models (i) cannot be used for analyzing the system behaviour otherwise than by the numerical simulation, (ii) cannot be scaled up or scaled down moving from one process scale to another. Therefore, the black box model is less useful for industrial practice. The black box model is valid only in a limited range.

Neural network (NN) model is an example of black box modelling of a system. The neural network models have gained much attention because of their capability of nonlinear function approximation. Theoretically, neural networks with three layers can approximate any nonlinear function with arbitrary accuracy. This feature makes the neural networks a promising nonlinear modelling tool. The development of NN model involves three basic steps. These include the generation of (or compilation of available) data required for training of NN, and the evaluation and validation of the optimal configuration of the NN model. Training the NN models with such a large number of parameters can lead to a poor prediction.

There is a range of modelling methods that attempt to combine the advantages of white box and black box approaches, such that known parts of the system are modelled using the physical knowledge and the unknown or less certain parts are approximated in a black box manner, using process data and black box modelling structures with suitable approximation properties. These methods are often denoted by hybrid, semi-mechanistic or grey box modelling. Attention can be given on the semi-parametric approach, which combines a neural network with a fixed form of a parametric model, either in series or in parallel. For example, a first principle model can be used to describe a particulate solid drying process, whereas, the unknown parameters in the first principle model can be approximated by a black box model. Figure 1.3 shows the architecture of a serial grey box model structure. The serial grey box modelling technique with black box network offers substantial advantages over a black-box (such as neural networks) modelling approach. The training of the model is easier and the model can be updated and it is more reliable for prediction purposes.

Sometimes, we have a situation, where a complete description of the entire process is not available. For example, in an exothermic reactor, the heating/cooling system is well studied but the kinetics of the reaction system, may not be easily modelled. The serial approach cannot be applied here, but an alternate (parallel) approach can be used (refer to Fig. 1.4)

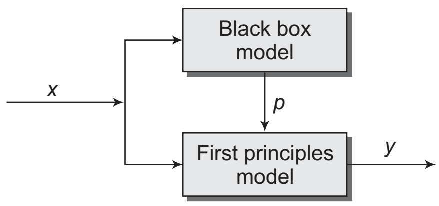

Fig. 1.3 Serial structure of a grey-box model with black box model (Xiong and Jutan, 2002)

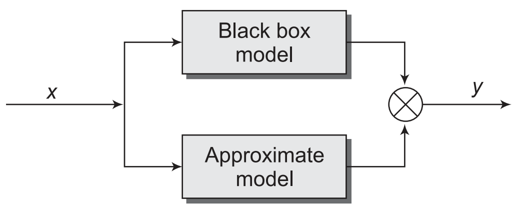

Fig. 1.4 Parallel structure of a grey-box model with black box model (Xiong and Jutan, 2002)

Review Problems

1. What is meant by mathematical modelling of systems?
2. Discuss briefly analysis, synthesis and compensation of a system.
3. Describe briefly the development of mathematical models.
4. With relevant equations, explain what is meant by nonlinear differential equations.
5. What are the classifications of mathematical models?
6. Explain briefly black box modelling and grey box modelling.
7. Discuss briefly random processes.