Parameter Estimation and Sensitivity Analysis

5

The computer model based on stock–flow diagram discussed in Chap. 4 needs estimation of parameters and sensitivity analysis of the parameters. This chapter presents parameter estimation and sensitivity analysis. Parameter estimation using aggregated and disaggregated data are discussed. Estimation of table function, conversion factor and normal fractional flow rate is highlighted with examples. Sensitivity analysis and size of solution interval have been explained, and some examples of sensitivity analysis have been presented.

5.1 Introduction

System dynamics modelling offers an attractive tool for policy testing and evaluation. Policy alternatives can be simulated by computer. But simulation requires a model, and model building is still an art in many respects. After formulation of dynamic hypothesis, including mapping stock and flow diagram of a model, the parameters of the model should be either assigned numerical values or expressed in terms of equations with parameter values to simulate the model. The estimation of parameters is one of the important steps in model building. The correct estimation of parameters is required to provide plausible behaviours of the system over time. Also sensitivity analysis of the parameters is important to assess how the parameter values affect the system behaviour and how important it is to determine the parameters accurately. This chapter discusses parameter estimation techniques and sensitivity analysis.

5.2 Parameter Estimation Techniques

Parameter estimation techniques can be classified on the basis of assumption and data. The data may be categorised in two categories which are disaggregated and aggregated data.

B.K. Bala et al., *System Dynamics*, Springer Texts in Business and Economics,
DOI 10.1007/978-981-10-2045-2_5

5.3 Estimation Using Disaggregated Data

Many system dynamics studies including the studies of Forrester use parameters determined on the basis of descriptive information obtained from the participants in the systems modelling. Such information is disaggregated data.

For an example of parameters that can be estimated with disaggregated data, consider an equation below which represents death rate of the population model:

$$\text{Death rate}(\text{person}/\text{year}) = \frac{\text{Population}\,(\text{person})}{\text{Average life expectancy (year)}} \tag{5.1}$$

In STELLA

death_rate = population/average_life_expectancy

The model assumes a constant average life expectancy, so that every year, 1/average life expectancy of the population is dead. Average life expectancy can be estimated many ways from disaggregate data, and the equation is not used for computation of average life expectancy. The equation is only needed to define the function in the model. There are many ways to find the average life expectancy. One time-consuming approach is to make a survey to find the number of people dead and their ages at the time of death. A second approach is to examine the ages of the people alive and observe the age at which very few people are alive. The third approach is to ask the oldest people who have observed a lot of deaths, about the ages at which most of the old people died. The fourth approach is to examine the health conditions and extrapolate life time expectancy. The fifth approach is to study the descriptive history of population, food and society, etc., and can have some idea of life expectancy. There may be some other means also to find the life expectancy. The data discussed above come from records, a history book, expert testimony and the modeller’s own day-to-day experience. Thus, disaggregated data are by far the most abundant source of knowledge about real systems.

Another example is organic degradation equation in paddy production model as represented below:

$$\text{Organic degradation (tonne/year)} = \frac{\text{Cumilative organic (tonne)}}{\text{Organic degradation time (year)}} \tag{5.2}$$

The model estimates a constant degradation time. Every year, 1/organic degradation time of cumulative organic is degraded. As discussed above, the same approach also can be used to estimate the degradation time either by survey, experiment, historical data or any other sources of information.

5.3.1 Table Functions

Many of the relationships between the variables in complex systems are non-linear especially in socio-economic systems, and such non-linear relationships are expressed in the form of table functions. This section describes a less straightforward but a common technique to estimate the table function. The table function seems to be a formidable estimation problem, but it can be broken into subproblems, which are:

1. Estimate the value and the slope of the function at (a) one extreme, (b) the normal value, and (c) the other extreme.
2. Connect these known values and slopes with a smooth curve.

Once these subproblems are solved, the table function is known within a narrow range of values.

For example, consider market price and inventory relationship of the commodity production cycle model.

The non-linear relationships between the price and inventory in STELLA are:

```
price = GRAPH(inventory)
(0.00, 100), (1000, 96.5), (2000, 94.0), (3000, 90.0), (4000, 80.0),
(5000, 70.0), (6000, 50.0), (7000, 30.5), (8000, 20.0), (9000, 16.0),
(10000, 10.0)
```

To estimate table function for price, first consider the extreme condition of zero inventory, where the price would be maximum. After that, consider the normal condition when there is some moderate quantity of inventory and the table function has a negative slope. At the other extreme, the inventory is large, and the price is very low. Now draw a smooth curve through the estimated points (Fig. 5.1). Thus, solving the subproblems of extreme conditions, and connecting known points with smooth curves, allows the modeller to estimate a non-linear table function with confidence.

For another example, consider the relationship between soil moisture factor and available soil moisture of the crop irrigation model.

The non-linear relationships between soil moisture factor and available soil moisture in STELLA are:

```
soil_moisture_factor = GRAPH(available_soil_moisture)
(0.00, 0.00), (9.00, 0.03), (18.0, 0.06), (27.0, 0.09), (36.0, 0.12),
(45.0, 0.16), (54.0, 0.19), (63.0, 0.23), (72.0, 0.26), (81.0, 0.32),
(90.0, 0.38), (99.0, 0.45), (108, 0.53), (117, 0.62), (126, 0.72),
(135, 0.86), (144, 0.95), (153, 1.00), (162, 1.00), (171, 1.00), (180, 1.00)
```

To estimate table function for soil moisture factor, first consider the extreme condition of zero available soil moisture, where the soil moisture factor would be zero. Now consider the normal conditions when there are some moderate values

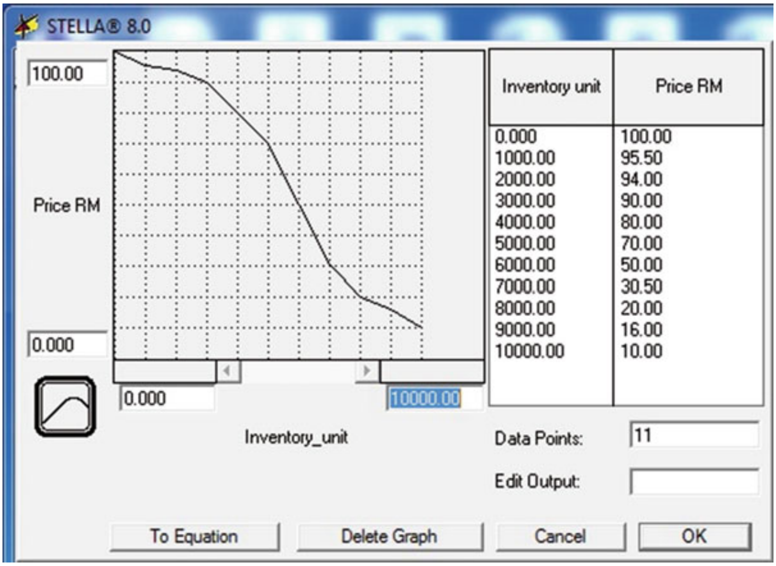

Fig. 5.1 Price versus inventory relationship

and the table function has a positive slope. At the other extreme, the available soil moisture is large and the soil moisture factor is unity. Now draw a smooth curve through the estimated points (Fig. 5.2). Thus, solving the subproblems of extreme conditions, and connecting known points with smooth curves, allows the modeller to estimate a non-linear table function with confidence.

The non-linear relationships between rice per capita consumption and gross domestic product per capita can be expressed in the form of table function. This relationship is illustrated in Fig. 5.3.

```
rice_per_capita_consumption = GRAPH(GDP_per_capita) (4000, 83.7),
(4500, 84.8), (5000, 85.6), (5500, 85.8), (6000, 85.7), (6500, 85.0),
(7000, 84.0), (7500, 82.4), (8000, 80.5), (8500, 78.0), (9000, 75.2)
```

Table functions can also be estimated in normalised form. Consider a cumulative input such as chemical and organic fertilisers in food production. In the long term, the cumulative application of these inputs will affect land fertility, and the fertility will change over time non-linearly. The use of inputs depends on technology and variety of the crop. To develop meaningful relationship between the land fertility effect and cumulative input, it is better to normalise the relationship. In order to normalise the cumulative input, the cumulative input should be divided by its initial value and the land fertility by initial land fertility to determine the fertility effect. After that, the effect of cumulative input on land fertility can be expressed in normalised form as in Fig. 5.4.

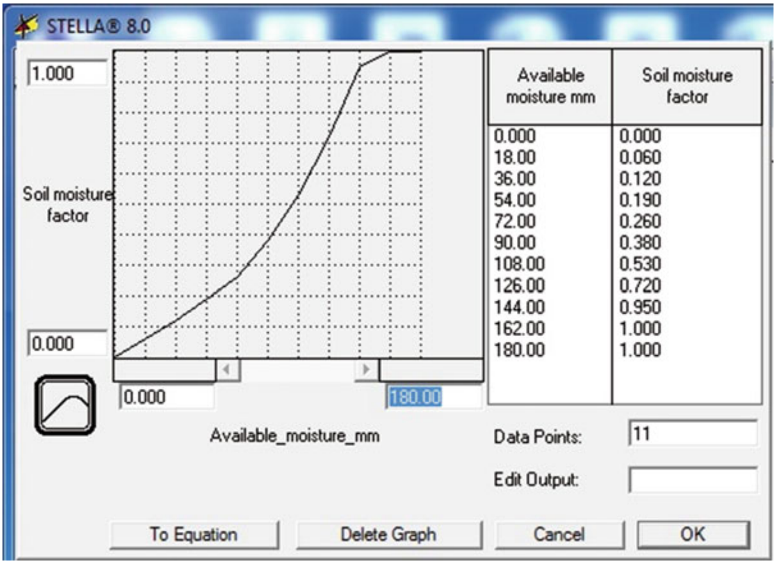

Fig. 5.2 Soil moisture factor versus available soil moisture relationship

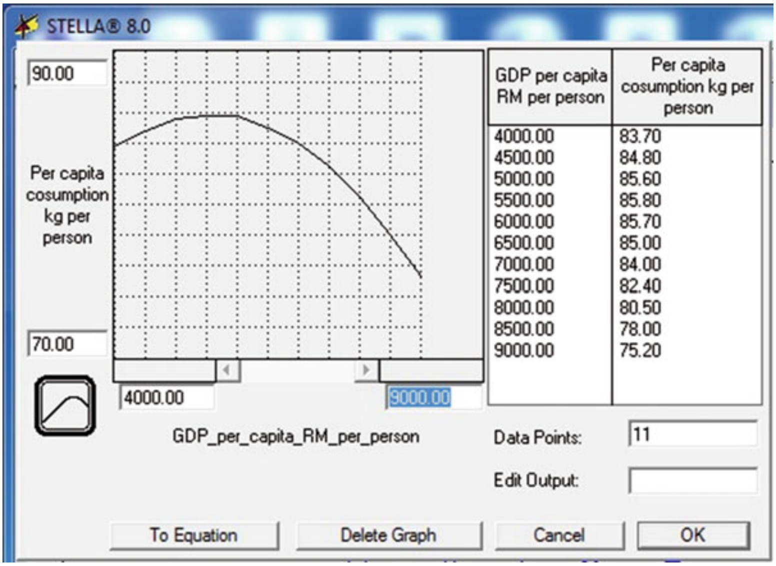

Fig. 5.3 Rice per capita consumption (kg/person) versus GDP per capita (RM/person)

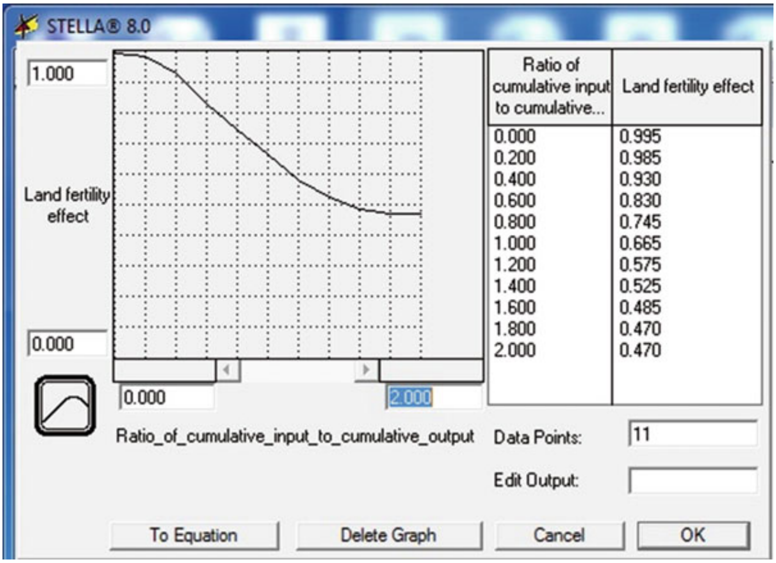

Fig. 5.4 Effect of cumulative input on land fertility

5.4 Estimation Using Aggregated Data

When using aggregated data to estimate parameters, the correctness of model equation(s) is (are) assumed. Two main techniques of parameter estimations using aggregated data are using (a) single equation and (b) multiple equations.

5.4.1 Estimation Using a Model Equation

The model equation is assumed to be correct and the aggregated data that correspond to model variables are used to estimate the parameter value. Estimation using a model equation encompasses all single equation regression techniques.

For an example of estimation using a model equation, consider the population model for calculating the death rate in Eq. 5.1. If the data for death rate and population are available, LE can be estimated by rearranging the formulae in Eq. 5.1. Therefore, the equation for estimated LE is as follows:

$$DR = \frac{P}{LE} \tag{5.3}$$

$$LE = \frac{P}{DR} \tag{5.4}$$

DR Death rate (person/year)
P Population (person)
LE Average life expectancy (year)

Parameters in biological and aquacultural models are essentially estimated using regression techniques. For example, the relation between fasting metabolic rate and body weight of a fish can be expressed as follows:

$$\mathrm{FMR} = \alpha \times \left(\mathrm{WoF}^{\gamma}\right) \tag{5.5}$$

FMR Fasting metabolic rate (Mole ATP/day)
WoF Weight of fish (g)

α and γ are the parameters estimated regressing the experimental data on fasting metabolic rate and body weight of fish. For catfish, these parameters are $\alpha = 0.0012$ and $\gamma = 0.8$.

Estimation using a model equation is less frequent in socio-economic system dynamics studies. However, two forms of model equation estimation—one involving conversion factors and the other fractional rates of flow—have been useful in many studies and are discussed below.

Conversion Factors

Many model parameters are conversion factors, and they convert quantities from one dimension to another. For example, consider the housing sector of the residential community model. Land per house (LPH) converts housing units (H) to an equivalent number of hectares:

$$\mathrm{LFO} = \frac{(H \times \mathrm{LPH})}{A} \tag{5.6}$$

LFO Land fraction occupied (dimensionless)
H Housing unit (unit)
LPH Land per house (ha/unit)
A Area (ha)

Such equation can be manipulated to compute the parameter as function of the real data, in this case, the real LFO, H and A:

$$\mathrm{LPH} = \frac{(\mathrm{LFO} \times A)}{H} \tag{5.7}$$

Consider another example. Here, straw feed per cattle converts cattle population to an equivalent amount of straw feed. The following equation uses straw feed per cattle in the definition of fraction of available straw used as cattle feed (FoS):

$$\text{FoS} = \frac{(\text{SFC} \times CP \times DY)}{\text{ASF}} \tag{5.8}$$

FoS Fraction of Straw (dimensionless)
SFC Straw feed per cattle (kg/day)
CP Cattle population (head)
DY 365 days (day)
ASF Available straw feed (kg/year)

This equation can be used to compute the parameter as a function of the real data in this case FoS, ASF, *CP* and *DY*:

$$\text{SFC} = \frac{(\text{FoS} \times \text{ASF})}{(CP \times DY)} \tag{5.9}$$

Estimating conversion factors offers a straightforward means of ensuring that the absolute magnitudes of model variables are realistic.

Normal Fractional Rates of Flow

Equation below defines the death rate in terms of a level population, a normal fractional rate of death (DRN) and dimensionless multipliers:

$$DR = \text{POP} \times \text{DRN} \times \text{DRFM} \times \text{DRPM} \tag{5.10}$$

DR Death rate
POP Population
DRN Death rate normal
DRFM Death rate multiplier from food
DRPM Death rate multiplier from population

The format for above equation is:

$$\text{Rate} = \text{Level} \times \text{Normal fraction} \times \text{Multipliers} \tag{5.11}$$

This format is widely used and the multiplier can be easily estimated when it is normalised around 1.0. It also facilitates the estimation of the normal fraction, and it can be represented by:

$$\text{Normal fraction} = \frac{\text{Rate}}{(\text{Level} \times \text{Multipliers})} \tag{5.12}$$

Under normal conditions the multipliers assume a value of 1.0, so the fractional flow rate can be obtained by dividing the observed rate by the observed level. For example, suppose 1970 is the normal period for the population model. Then the birth rate normal is obtained from the number of births during 1970 divided by the number of population in 1970. Choosing a normal condition does not bias behaviour of the model. A normal condition is merely a system state about which parameters are defined.

5.5 Estimation Using Multiple Equations

Estimation using several multiple equations consists of using several equations to compute a parameter value. For example, birth rate normal (BRN) can be estimated by finding the value of BRN that causes population growth rate to fit the observed rate of growth. This estimation will use all the equations. The fitting could be performed either with repeated simulations or, if possible, by computation. For an example of such a computation, the population growth rate is 2.5 %. Also from observations of the dead, the life expectancy (LE) is estimated to be 66 years, that is, 1/66 of the population are dead per year. Therefore the BRN should be $1/66 + 0.025 = 0.04$.

5.6 Sensitivity Analysis

Here we consider the parameter sensitivity analysis, and it is conducted to assess how sensitive is the model behaviours to the changes in the values of the parameters. Parameter sensitivity analysis is usually conducted by setting different values such as ±15 % changes in the parameter to study the changes in the model behaviours. Sensitivity analysis provides an opportunity to determine the level of accuracy needed in the estimation of the parameter to make the model valid and useful. If the parameter is insensitive to the model behaviour, it may be used in the policy analysis and design. On the other hand, parameters significantly affecting the model behaviour should be chosen as candidates for additional data collection (Sterman 2000).

The parameters of a system dynamics models are subject to uncertainty. So, sensitivity analysis is an important task for reliability of the simulated results and checking the robustness of the model behaviour for the changes in parameter values. The sensitivity of the important parameters should be estimated. The productivity is one of the most important parameters affecting the model behaviour, i.e. self-sufficiency ratio which is the focus of the food security model (Bala et al. 2014). Figure 5.5 shows the changes in food self-sufficiency ratio for rice productivity for changes of 2.50 t/ha (curve 1), 3.50 t/ha (curve 2) and 4.50 t/ha (curve 3). The food self-sufficiency ratio changes from decrease to increase in values for the changes in the rice productivity from 3.50 t/ha to 2.50 t/ha and 4.50 t/ha, and this corresponds with real-world situation.

Another example of sensitivity analysis is the sensitivity of cocoa production systems in Malaysia to the changes in the subsidy level (Fatimah et al. 2015). Figure 5.6 shows the simulated cocoa plantation area for full subsidy (curve 3),

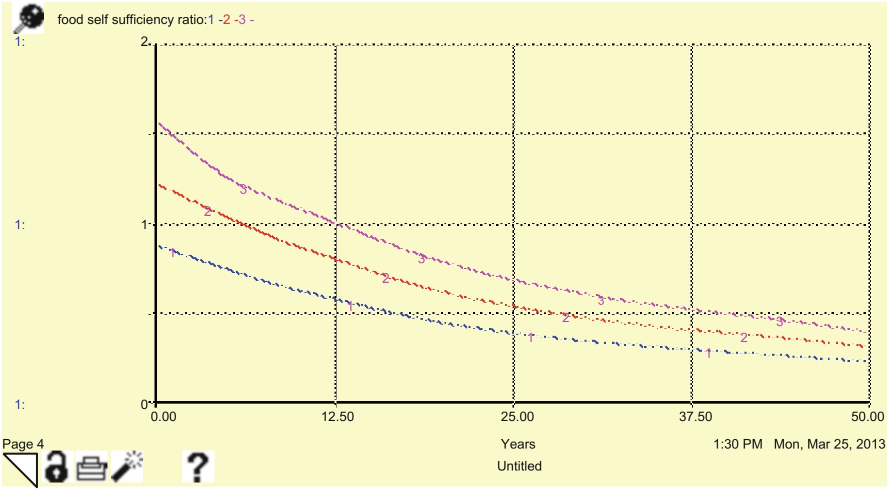

Fig. 5.5 Sensitivity of food self-sufficiency ratio to rice productivity (2.50 t/ha, 3.50 t/ha and 4.50 t/ha)

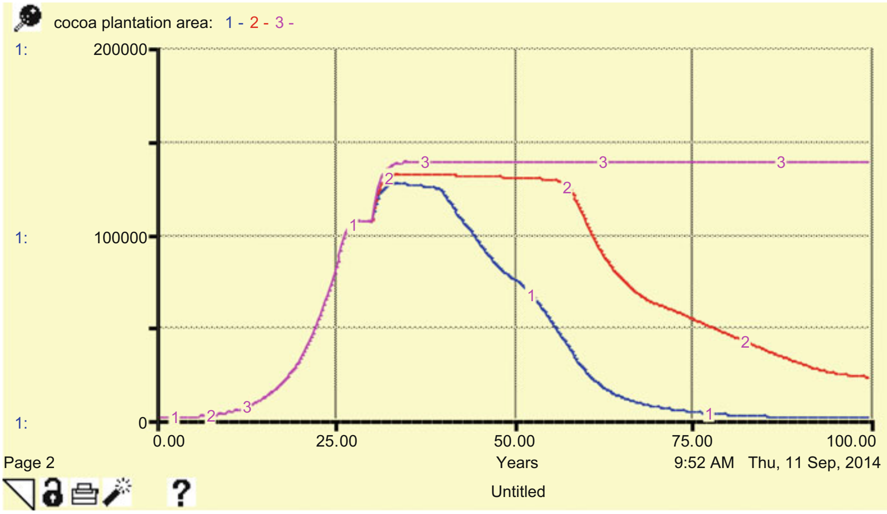

Fig. 5.6 Sensitivity analysis of subsidy for cocoa production systems in Malaysia

80 % subsidy (curve 2) and 60 % subsidy (curve 1) of the cost to cover conservation practices and insect control. The system is sustainable for full subsidy since financial support is provided for joint trade-off of biodiversity and yield along with integrated pest management (IPM) pest control through farmer field schools (FFS), but the sustainability decreases with the degree of reduction of subsidy as lesser opportunities available to maintain biodiversity and insect control in terms of financial support.

5.7 Size of Solution Interval (DT)

It has been mentioned in Chap. 4 that solution interval should be half or less than of the shortest time delay in the model. Let us now consider a simple model of grain storage system as shown in Fig. 5.7 to address the effects of the size of solution interval, and the STELLA equations of the model are given below:

STELLA Equations

```
grain_stock(t) = grain_stock(t - dt) + (receiving_rate) * dt
NIT grain_stock = 50
```

INFLOWS:

```
receiving_rate = (desired_grain_stock-grain_stock)/receeiving_delay
desired_grain_stock = 100
receeiving_delay = 0.5
```

Here in this model, the initial grain stock is 50 tons and desired stock is 100 tons, and receiving time delay is 0.5 months. Let us see what happens to the simulated behaviour of this simple system if solution intervals are 0.20, 0.50 and 1.0. Figures. 5.8, 5.9 and 5.10 show the simulated responses of the system for the solution intervals of 0.20, 0.50 and 1.0, respectively. The simulated response of the system in Fig. 5.8 is the inherent system behaviour for a solution interval of less

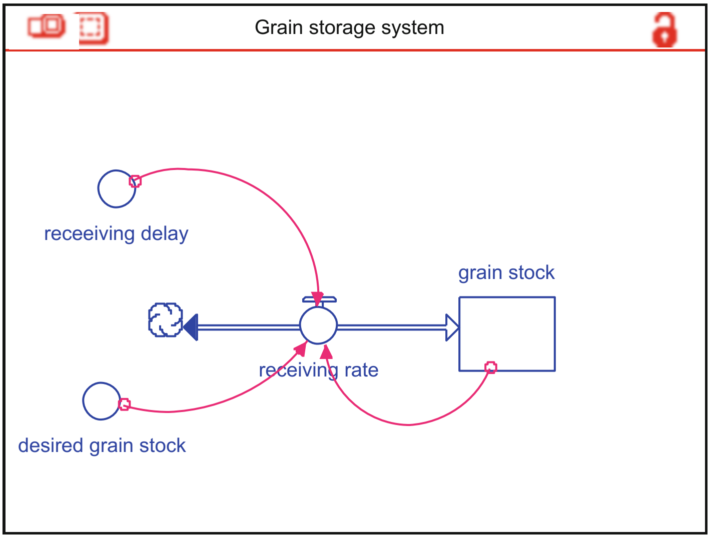

Fig. 5.7 Stock–flow diagram of a simple grain storage system

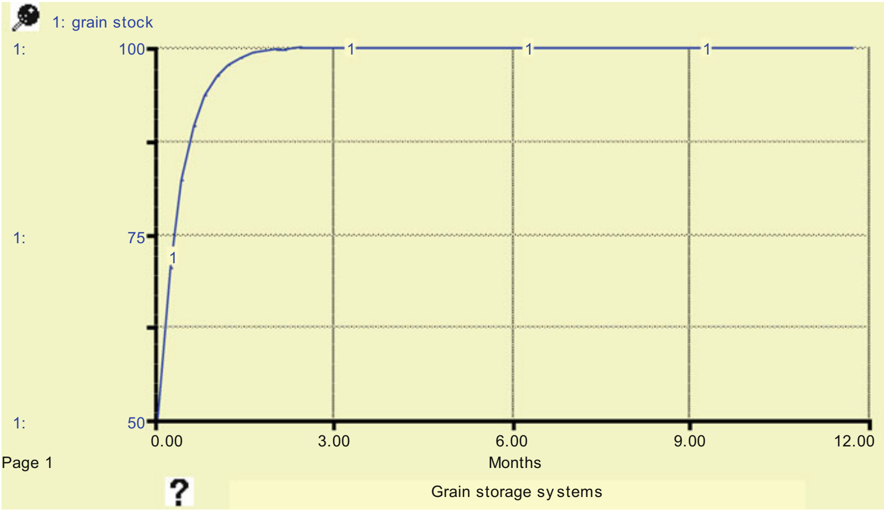

Fig. 5.8 Grain stock response for solution interval of 0.20

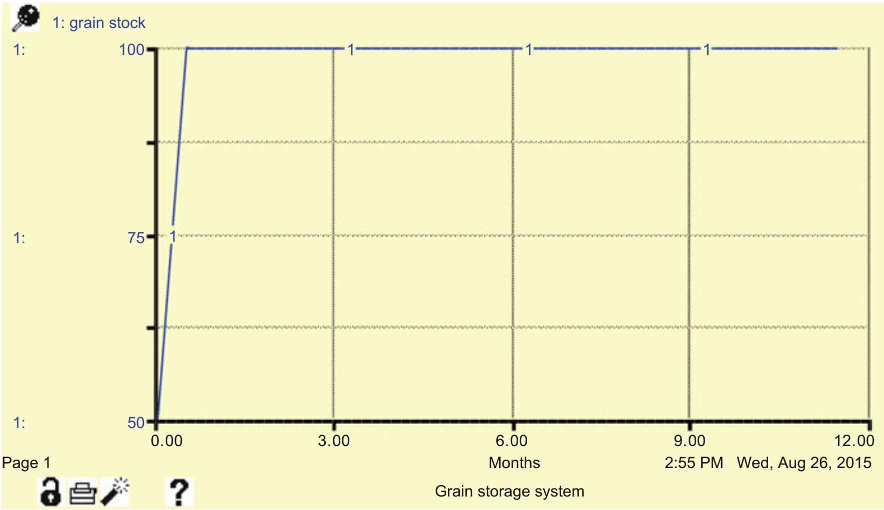

Fig. 5.9 Grain stock response for solution interval of 0.50

than half of the time delay. The simulated response in Fig. 5.9 is poor for a solution interval equal to time delay, and the simulated response in Fig. 5.10 for a solution interval of double of the time delay is the worst, and these responses are not the inherent system behaviour, but these are the results of improper size of solution interval. The solution interval is not part of the real system, but the improper size of it affects the simulated results seriously.

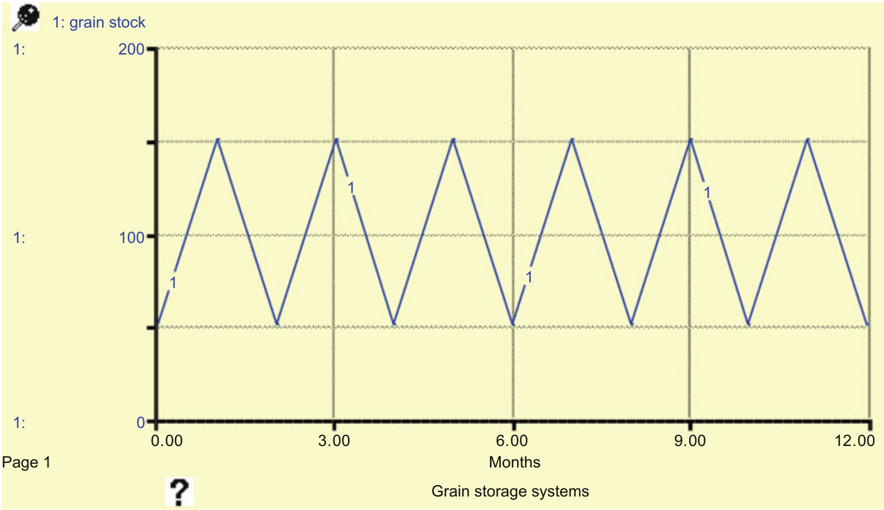

Fig. 5.10 Grain stock response for solution interval of 1.0

Exercises

Exercise 5.1 What is meant by parameter estimation? Describe briefly the techniques of the parameter estimation.

Exercise 5.2 What is table function? Show with examples how table functions are estimated.

Exercise 5.3 What are the main techniques of parameter estimation using aggregated data? Describe how multipliers are estimated?

Exercise 5.4 What is meant by sensitivity analysis? Describe with examples how the sensitivity of a parameter analysed.

Exercise 5.5 What is meant by solution interval? Show that improper use of solution interval can provide explosive behaviour which is not the behaviour of the system.

References

Bala BK, Alias EF, Fatimah MA, Noh KM, Hadi AHA (2014) Modeling of food security in Malaysia. Simul Model Pract Theory 47:154–164

Fatimah MA, Bala BK, Alias EF, Abdulla I (2015) Modeling of boom and bust of cocoa production systems in Malaysia. Ecol Model 309–310:22–32

Sterman JD (2000) Business dynamics: systems thinking and modeling for a complex world. McGraw-Hill Higher Education, Boston

Bibliography

Bala BK (1999) Principles of system dynamics. Agrotech Publishing Academy, Udaipur

Bossel H (1994) Modeling and simulation. K.Peters, Wellesley, 484 pp

Breierova L, Choudhuri M (2001) An introduction to sensitivity analysis. Massachusetts Institute of Technology, Cambridge, MA

Forrester JW (1961) Industrial dynamics. M.I.T. Press, Cambridge, MA

Hekimuglu M, Barlas Y (2015) Sensitivity analysis of system dynamics models by behaviour pattern measures. Department of Industrial Engineering. Bogazici University, Istanbul, Turkey

Maani KE, Cavana RY (2000) Systems thinking and modelling: understanding change and complexity. Prentice Hall, Auckland

Manetsch TJ, Park GL (1982) Systems analysis and simulation with applications to economic and social systems. Department of Electrical Engineering and System Science, Michigan State University, East Lansing

Mohapatra PKJ, Mandal P, Bora MC (1994) Introduction to system dynamics modelling. Universities Press, Hyderabad