Modelling with real data and technology

Keng-Cheng Ang
kengcheng.ang@nie.edu.sg
National Institute of Education
Nanyang Technological University
Singapore

Abstract: There are several ways in which the teaching and learning of mathematical modelling may be motivated. In this paper, we describe how real data, together with technology, can provide a rich environment for mathematical modelling activities. Examples on modelling tasks for students at different levels are presented and described in detail. These examples not only illustrate the use of real data and technology in mathematical modelling activities, they also underline the importance of keeping such tasks in relevant contexts to provide added motivation for students.

1. Introduction

It is widely accepted that mathematical modelling is one useful way of injecting more real life activities into the mathematics classroom. However, although mathematical modelling has had a reasonably long history, mathematics and mathematics educators alike, cannot seem to agree on a precise definition for this area of work [6].

Some researchers view mathematical modelling as essentially the movement of a physical situation to a mathematical representation [15, 16], while others feel that all applications of mathematics are mathematical models [7]. Galbraith, however, believes that there is a difference between mathematical modelling and applications of mathematics [11]. In addition, Galbraith proposed that the teaching of mathematical modeling may take either an "open" or a "structured' approach. In contrast, Yanagimoto defines mathematical modelling as not just a process of solving a real life problem using mathematics but "applying mathematics which is useful to society" [17]. Still, there are those who believe that mathematical modelling consists of understanding, simplifying and solving a real life problem in mathematical terms [5, 8].

Notwithstanding the differing views, one aspect of mathematical modeling that seems to be commonly regarded as the most important is its connection with real world problems. To further emphasize and focus on this connection, perhaps mathematical modelling activities should involve the use of real data. Real data provides a rich and often relevant platform for developing, designing, learning and applying mathematical models.

The introduction of technological tools such as graphing calculators, computer algebra systems and dynamic geometry software has influenced the approaches to teaching mathematical modelling. Apart from enabling the user to perform computational experiments with the model, technology can help in making it more possible to work with real data [1], which are usually not as "clean" or "sanitized" as textbook or lecture examples.

In this paper, we discuss some examples of modelling tasks motivated by real data and aided or supported by technology. Real data are either collected or obtained through the literature or from some source in the open domain. In addition, the problems in these examples are both real and relevant to students and set in the context with which they are familiar. Hopefully, this will help raise the level of motivation in tackling these tasks.

2. What is mathematical modelling

While there may be different interpretations of mathematical modelling, for the purpose of the current discussion, we shall adopt the definition given in [2]. That is, we define mathematical modelling as the process of representing or describing physical systems or problems in the real world using mathematics so as to gain a more precise understanding of the problem. This process may be depicted as a flow of events as illustrated in Figure 2.1.

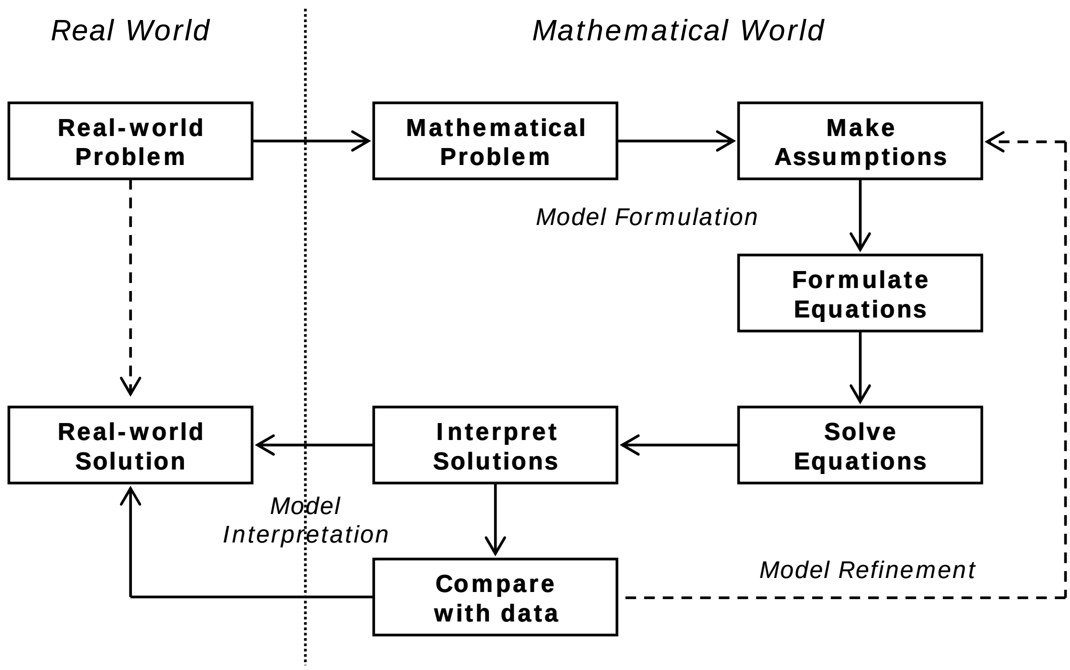


Figure 2.1 Mathematical modelling process

The process begins with a problem in the real world. In the first step, one hopes to describe or represent this problem in some mathematical terms. This may involve stating some variables, and forming some relationships amongst these variables. Usually, some assumptions have to be made in this model formulation phase. With these assumptions and an understanding of the mechanics of the problem, the next step is to formulate some equations to represent the dynamics of the system. The equations (or set of equations) thus form the mathematical model and the next step is to make an attempt to solve them.

The solution of the model equations may be where technology can come in handy. Quite often, the process involves estimation of some unknown parameter. Sometimes, some computational or numerical techniques may be involved, or perhaps some other approximations have to be made. At other times, collected or known data need to be analyzed before a solution method can be proposed or employed to solve the equations. Whatever the case, it is likely that when real data is involved, the use of technological tools will help in solving the problem.

To link the mathematical solution back to the real world problem, one needs to interpret the results or solutions of the model. It would be ideal if one could make some comparison of the results with some real data to validate the model. This represents the end of one cycle of the process because it is often possible to refine the model, improve the methods, or gather more data and repeat the steps in the process.

Although there may be many different versions of the process, most of them are similar and convey the same message essentially. However, the implementation and teaching approaches can vary a great deal, depending on what the teacher wishes to emphasize on. In the next section, we examine three examples, each of which illustrates the use of real data either collected by the students, or obtained from some real source. In every case, some form of technology or technological tool is used to solve or investigate the problem.

3. Examples

In this section, we consider some examples of how the process of mathematical modelling may be introduced in the classroom. In each of the examples, a problem with real data is considered, and a solution approach using a mathematical model and some form of technology is introduced. The examples presented are chosen to illustrate the range of possibilities of introducing modelling tasks to students at different levels of cognitive development.

Example 1: Largest Box Problem

Consider the problem of making an open-top box using a square piece of cardboard (of side S cm) by cutting a square of side x cm from each corner of the cardboard. The resulting piece is folded to form the box, as illustrated in Figure 3.1.

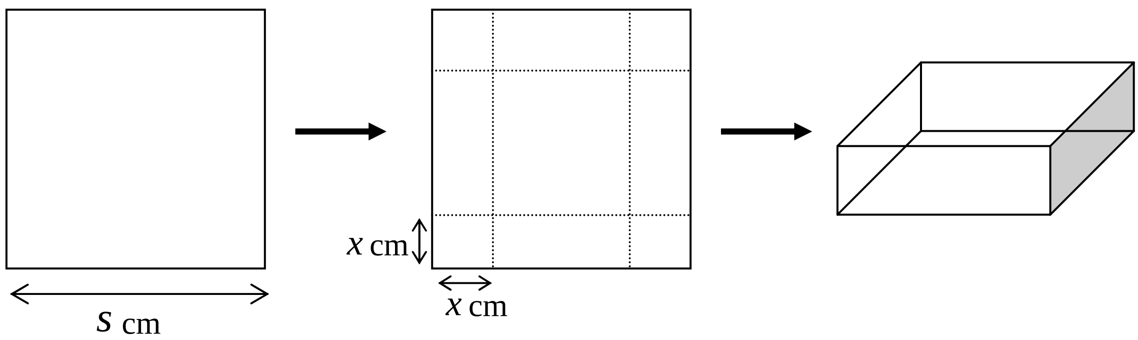
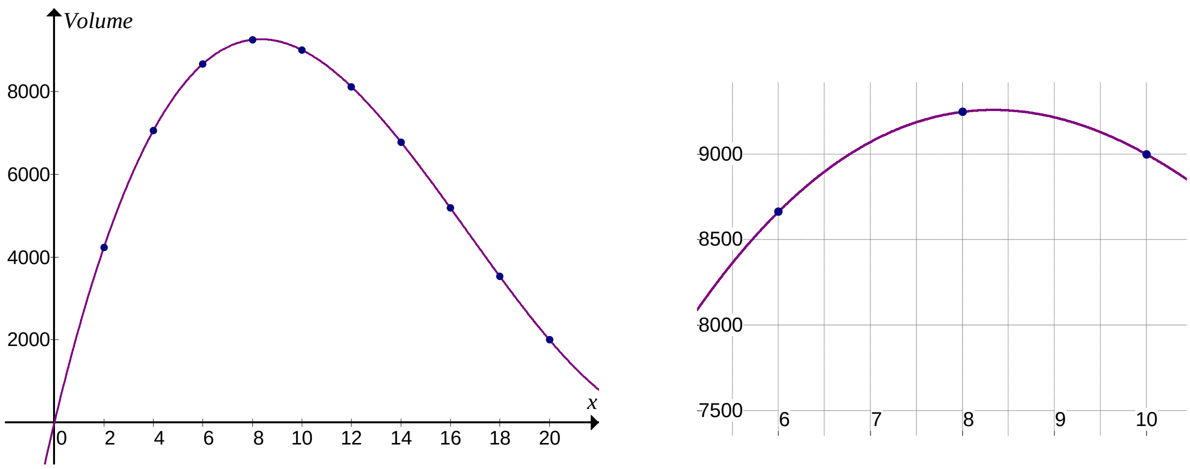


Figure 3.1 Largest Box Problem

The question is: what should x be in order to make the biggest box (in terms of volume)? This could be viewed as an optimization problem for an industry concerned with packing.

This problem was given to some Primary 6 students (11-12 year olds) in a Singapore school. It is important to note that at this level, students have very limited knowledge of algebra and, of course, yet to come across calculus. However, most would have been exposed to simple graphing tools. The students were given sheets of cardboard measuring 50cm by 50cm. In groups of 4 or 5, they made boxes with different values of x and computed the corresponding volumes. By collecting all the measurements, a graph of volume against x similar to that shown in Figure 3.2 may be plotted using a graphing tool such as Graphmatica or a graphics calculator. One could then repeatedly zoom in to estimate the maximum value of the volume and the corresponding value of x.



(a) curve fitting data points   (b) zooming in near maximum

Figure 3.2 Graph of Volume of box against x

Example 2: A logistic model for a disease outbreak

In 2003, some Asian nations including Taiwan, Hong Kong and Singapore were inflicted with an infectious disease known as Severe Acute Respiratory Syndrome, or SARS. In Singapore, there were 206 cases of infection. Among these, 31 lost their lives. The number of SARS cases in Singapore over the period of 70 days has been reported by Heng and Lim [12] and is available in Appendix A.

A modelling task that may be posed to Singapore's Junior College (17-18 year-olds) students is to construct a model (or apply a known one) to represent the SARS outbreak in Singapore. One plausible model based on the so-called "S-I" model for epidemics is described below.

The "S-I" epidemic model consists of two compartments, the susceptible population and the infected population (Figure 3.3). This is called the "S-I" model as it involves susceptible individuals ("S") becoming infected ("I").

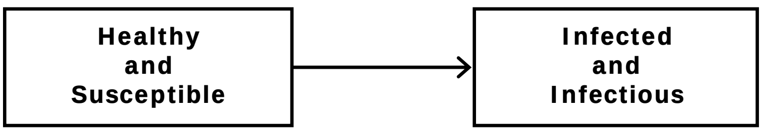


Figure 3.3 A simple epidemic model

Suppose $x(t)$ and $y(t)$ are the number of infected and susceptible individuals at time t (in days) respectively. We further assume that during the course of the epidemic, the total population of the community remains constant. Thus, $x(t)+y(t)=N$, where N is the size of the population. The spread of a highly communicable disease such as SARS may be modelled by the logistic equation given by

$$\frac{dx}{dt}=k\,x\left(1-\frac{x}{N}\right) \tag{3.1}$$

where k is a positive constant representing the transmission rate [3]. Equation (3.1) may be solved using the standard method of separation of variables and integration after performing partial fraction decomposition. Suppose the initial condition is $x(0) = x_0$, then the solution to the equation may be written as

$$x = \frac{N}{1 + \left(N/x_0 - 1\right)e^{-k\,t}}\,. \tag{3.2}$$

The transmission rate k may be estimated from data. For instance, we may use the data for the SARS outbreak in Singapore in 2003 (Appendix A) to find an estimate for k. To do so, we define an "average error",

$$E = \frac{\sqrt{\sum_{i=1}^{n} (\hat{x}_i - x_i)^2}}{n}\,, \tag{3.3}$$

where $\hat{x}_i$ and x_i are data values and model values respectively. A good estimate of k is obtained when E is minimised. One way to do this is to use the "Solver tool" in *Microsoft Excel*. The Solver tool essentially allows the user to minimise (or maximise) the value of a selected cell by varying the values of other cells specified by the user. In the present case, the Solver tool returns a value of $k = 0.1686$ (to four decimal places) with a minimum value of $E = 1.9145$. Figure 3.4 shows the graph of the model, with this value of k, plotted against the real data.

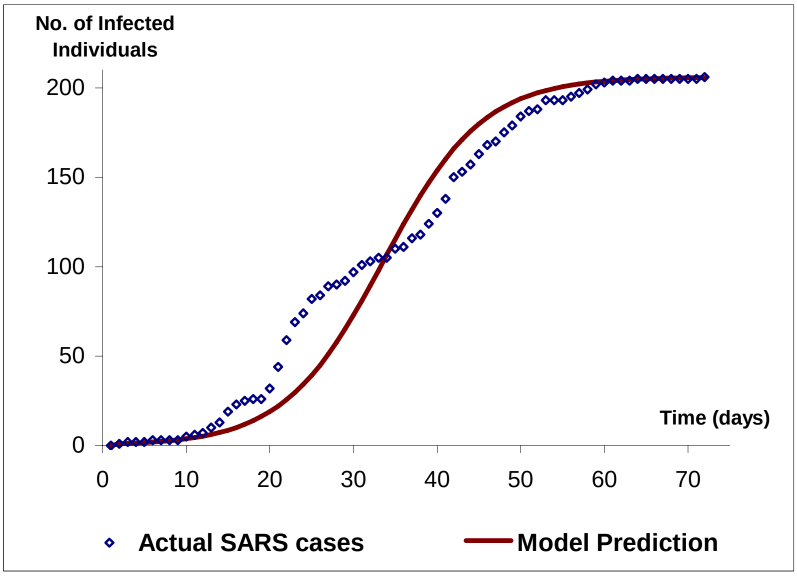


Figure 3.4 Graph of SARS outbreak model and real data

Although the model compares fairly well with the real data, it can be improved and refined, and this is discussed in detail in [3]. The challenge here may be to try and improve or refine the model to provide further insights into the dynamics of the outbreak.

Example 3: Modelling the spread of Dengue

Despite the *Aedes* mosquito control programmes, public health education and law enforcements, there has been a resurgence of dengue in Singapore in recent years. Dengue is transmitted when infected female *Aedes* mosquitoes, notably the *Aedes aegypti* and *Aedes albopictus*, bite human beings. After an incubation period of 5 to 8 days, dengue fever manifests itself with symptoms such as severe headaches, bone or joint and muscular pains, fever and rash. A complete recovery requires 4 to 7 days, and fatality is rare.

In Singapore, the annual incidence rate of the disease has risen progressively from 33 per 100,000 in 1993 to 166.2 per 100,000 in 1998. There are many possible reasons for the resurgence and increase in the number of dengue cases. It is suspected that Singapore's hot and humid weather plays a significant role in the breeding patterns of mosquitoes. As the daily average temperature in this city state remains fairly constant (at around 30°C) throughout the year, it is deemed that changes in the precipitation level (that is, rainfall) are probably the main environmental factor that is linked to the larvae densities of the *Aedes* mosquitoes. Thus, data on the number of dengue cases, larval densities and rainfall in Singapore may prove useful in providing insights to the dynamics of the disease. Such data for the year 1996-1997 are available in [4] (Appendix B).

The challenge, which may be posed to undergraduate students, is to construct a plausible model for the spread of dengue in Singapore based on the available data. One approach is suggested below.

In modelling the spread of dengue, one has to consider the host (that is, the human being) and the vector (that is, the mosquito) and the interactions between them. A possible compartment-model, in which each population type is compartmentalized, is depicted in Figure 3.5.

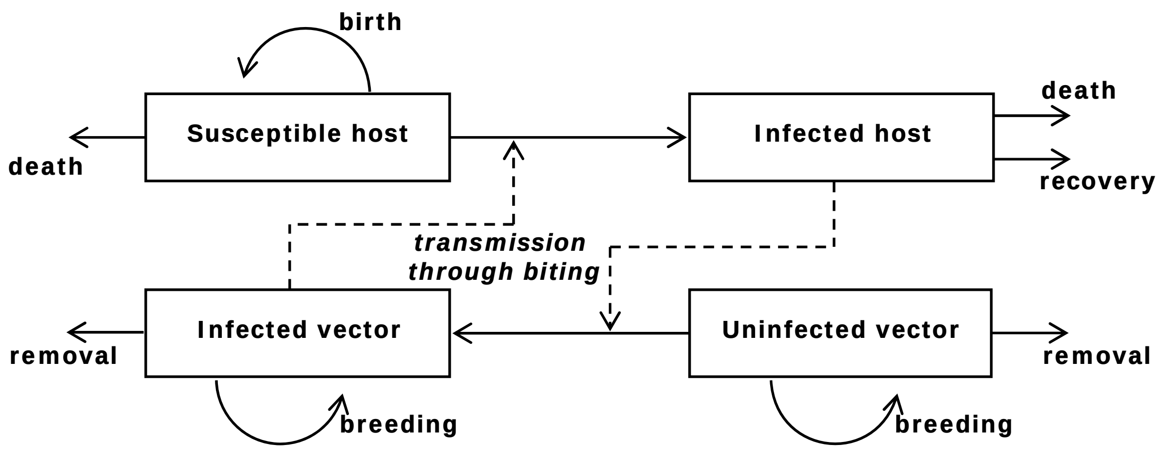


Figure 3.5 A basic model for spread of dengue

Let $S(t)$ and $I(t)$ be the number of susceptible and infected host individuals respectively, and let $N(t)$ and $M(t)$ be the number of uninfected and infected mosquitoes respectively. The model presented in Figure 3.5 may be represented by the following set of differential equations

$$\frac{dS}{dt} = (a - b)S - \upsilon SM \tag{3.4}$$

$$\frac{dI}{dt} = \upsilon SM - (\alpha + \gamma)I \tag{3.5}$$

$$\frac{dN}{dt} = (d - c)N - \omega NI \tag{3.6}$$

$$\frac{dM}{dt} = (d - c)M + \omega NI \tag{3.7}$$

where a and b denote the birth and death (from old age or other illnesses) rates of the host respectively, c denotes the removal rate of the vector by natural death or insecticides, and d denotes the breeding rate of the vector. The transmission efficiency of the host is given by υ while that of the vector is given by ω. The death rate due to the disease, and recovery rate from the disease are represented by α and γ respectively.

While this set of equations may seem complete, the underlying difficulty is the lack of information on the parameter values. Some of these (such as c, d, υ and ω) are non-measurable biological quantities and it is difficult, if not impossible, to obtain accurate values of these quantities. The remaining parameters a, b, α and γ may be estimated from known data.

Although it is possible to solve the set of equations using estimated parameter values, one missing component of this model is the impact or effect of environmental factors on the breeding rate of the vector. Mosquitoes breed rapidly in hot weather and where there are ample breeding sites. It turns out that in Singapore, breeding sites tend to increase during the rainy season, and the hot weather that follows promotes mosquito breeding. It is therefore not unreasonable to assume that the amount of precipitation (or rain, in this case) is linked to larvae densities.

The challenge is to find a way to model the relationship and incorporate it into the set of equations. One of the simplest ways to do so to assume that the population of infected vector is a fraction of the larvae density, and that this fraction depends linearly on the amount of rainfall. Thus, we may write

$$M(t) = P_i L(t) \tag{3.8}$$

$$P_i = \frac{P_{\max}}{R_{\max}} R(t) \tag{3.9}$$

where $L(t)$ and $R(t)$ are the larvae density and rainfall level at time t respectively, P_i is proportion of infected vector, and $P_{\max}$ and $R_{\max}$ are the maximum values of P_i and $R(t)$ respectively. $P_{\max}$ will need to be estimated experimentally.

Based on this assumption, Equations (3.8) and (3.9) can be used to replace Equations (3.6) and (3.7) in the model. Using the data given in Appendix B, the system of equations may be solved numerically using any suitable numerical method such as the Runge-Kutta method of Order 4.

Results from a typical run of the Fortran program written to implement the method is shown graphically in Figure 3.6. In this case, suitable parameter values, as reported in [4], that may be used are $a = 0.012$, $b = 0.007$, $\upsilon = 5\times10^{-7}$, $\gamma = 0.99$ and $\alpha = 3\times10^{-4}$. Figure 3.6 compares actual dengue cases with the model predictions using $P_{\max} = 0.7$ and $P_{\max} = 0.9$.

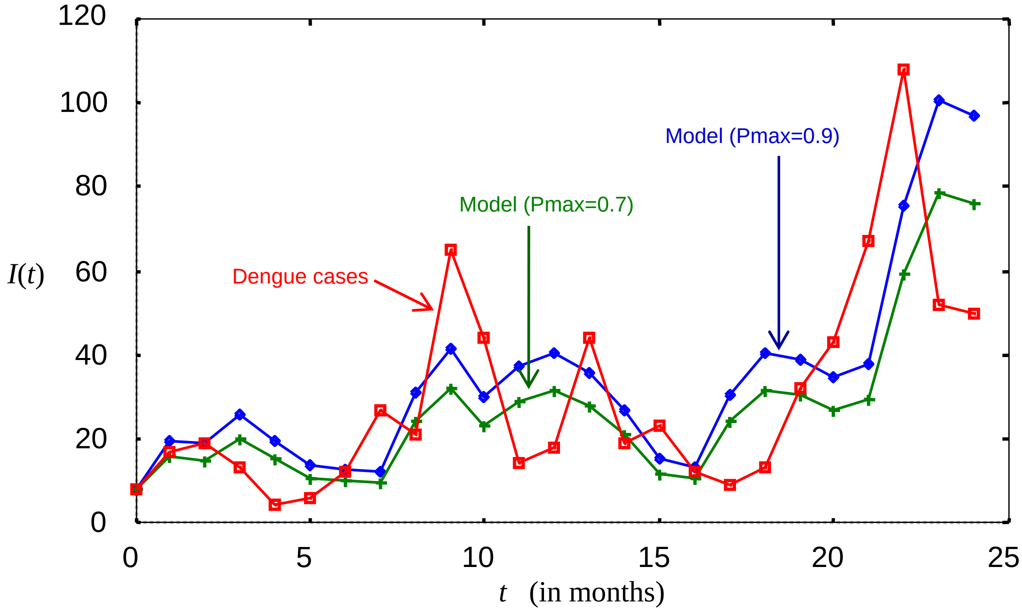


Figure 3.6 Actual dengue cases and model predictions

As can be seen from the graphical output above, the model compares fairly well with recorded data. The two spikes (around Months 9 and 22) in the dengue cases are reasonably well predicted. These spikes represent the dengue outbreaks during the period in question. While it is true that it can certainly be further improved, the present model serves well as a first approximation upon which more accurate models may be based.

4. Discussion

As can be seen, the common thread in all the examples discussed in this paper is the use of real data. In the first example, real data were collected, while in the next two examples, they were obtained from published sources.

In each case, the real data used had provided something concrete for students to work on. In the "biggest box problem", students could construct the boxes and then take measurements. Of course, the problem may be solved easily using calculus. However, the main point in this example is to illustrate how a problem or task may be tackled by modelling with real data. In addition, the use of appropriate technology can help fill some gaps in mathematical skills or knowledge, although one has to bear in mind not to trivialize the mathematical tasks. In this particular instance, technology has served to enrich the learning possibilities as students who do not have the "assumed knowledge" will still be able to construct the model physically and mathematically to tackle the problem.

The next two examples discussed problems related to disease outbreaks and control. In both cases, the context is real and relevant, and of immense public concern to people living in Singapore.

Although the problems are far from being solved, students who embarked on these modelling exercises would have gained a deeper appreciation of real life application of mathematics. The use of real data in modelling a SARS outbreak and dengue transmission has given students a rich mathematical experience in a highly relevant context. Moreover, in using of the real data, the application of technology has helped in reaching the solution of the models in both cases, making the mathematical modelling tasks more accessible.

While it may be possible to look for examples from textbooks or other sources in modelling epidemics or disease outbreaks, the use of data in real life provided learners with realistic opportunities to connect mathematics to important social issues or problems. Even if real data are found or cited in textbook examples, they may not provide as rich an experience as actual real life examples placed in a context that students can identify with. Real problems with real world concerns serve to heighten student interest and motivation [10].

5. Conclusion

Mathematical modelling can be thought of as a form of a scientific inquiry process for mathematics. This paper discusses some practical aspects of this process that may serve to enhance the learning experience in mathematical modelling.

The use of real data in real life problems helps the learner link the mathematical world and the real world. This appears to be an area of weakness in students and has been reported by various researchers [9, 14]. Examples such as those described in this paper will help students build connections between the mathematics they know and the world they live in.

Apart from using real life examples with real data, the timely and appropriate use of technology can help empower students in their modelling tasks. All examples presented in this paper have made use of some form of technology. Notwithstanding the usefulness of technology, it is important to note that the learning of mathematical modelling should not be limited to specific technological tools [13]. In other words, the focus should still be on using mathematics to model or solve a real world problem, and the tools are meant to support the process.

To truly experience the process of mathematical modelling, it would help if one deals with a real problem, handle real data or even test the validity of one’s models through various experiments, including computational or numerical experiments. Performing these tasks by hand without the aid of technology can be tedious, time-consuming and sometimes counter-productive or even impossible. A modelling task set in a real life context and supported by technological tools can provide a very enriching and engaging mathematical modelling experience.

References

[1] Ang, K.C. (2006a). Mathematical Modelling, Technology and H3 Mathematics, *The Mathematics Educator*, Vol. 9, No. 2, pp. 33 – 47.

[2] Ang, K.C. (2006b). *Differential Equations: Models and Methods*, Singapore: McGraw-Hill.

[3] Ang, K.C. (2004). A simple model for a SARS epidemic, *Teaching Mathematics and Its Applications*, Vol. 23, No. 4, pp. 181 – 188.

[4] Ang, K.C. and Pang, F.Y. (2002). *Modelling the Spread of Dengue in Singapore – A Final Report*, Research Report: NTU-ENV Joint Research Scheme, Singapore.

[5] Bassanezi, R.C. (1994). Modelling as a teaching-learning strategy, *For the learning of mathematics*, Vol. 14, No. 2, pp. 31 – 35.

[6] Blum, W. (1993). Mathematical modelling in mathematics education and instruction, in Breiteig, T., Huntley, I. and Daiser-Messmer, G. (eds), *Teaching and Learning Mathematics in Context*, London: Ellis Horwood.

[7] Burghes, D. (1980). Mathematical modelling: A positive direction for the teaching of applications of mathematics at school, *Educational Studies in Mathematics*, Vol. 11, pp. 113 – 131.

[8] Cross, M. and Moscardini, A.O. (1985). Learning the art of mathematical modelling, Chichester, England: Ellis Horwood.

[9] Crouch, R. and Haines, C. (2004). Mathematical modelling: transitions between the real world and the mathematical model, *International Journal of Mathematics Education in Science and Technology*, Vol. 35, No. 2, pp. 197 – 206.

[10] Ferrucci, B.J. and Carter, J.A. (2003). Technology-active mathematical modelling, *International Journal of Mathematics Education in Science and Technology*, Vol. 34, No. 5, pp. 663 – 670.

[11] Galbraith, P. (1999). Important issues in applications and modelling, Paper presented at the AAMT Virtual Conference 1999, Adelaide.

[12] Heng, B.H. and Lim, S.W. (2003). Epidemiology and control of SARS in Singapore, *Epidemiological News Bulletin*, Vol. 29, pp. 42 – 47.

[13] Kadijevich, D., Haapasalo, L. and Hvorecky, J. (2005). Using technology in applications and modelling, *Teaching Mathematics and Its Applications*, Vol. 24, No. 2-3, pp. 114 – 122.

[14] Klymchuk, S. and Zverkova, T. (2001), in Matos, J.F., Blum, W., Houston, S.K. and Carriera S.P. (eds). *Modelling and Mathematics Education*, Chichester: Horwood Publishing.

[15] Mason, J. and Davies, D. (1991). *Modelling with mathematics in primary and secondary schools*, Sydney: Deakin University Press.

[16] Swetz, F. and Hartzler, J.S. (1991). *Mathematical modelling in the secondary school curriculum*, Reston, V.A.: The National Council of Teachers of Mathematics.

[17] Yanagimoto, T. (2005). Teaching modelling as an alternative approach to school mathematics, *Teaching Mathematics and Its Applications*, Vol 24, No. 1, pp. 1 – 13.

Appendix A: Data for the SARS outbreak in Singapore

Day	Number	Day	Number	Day	Number
0	1	24	84	48	184
1	2	25	89	49	187
2	2	26	90	50	188
3	2	27	92	51	193
4	3	28	97	52	193
5	3	29	101	53	193
6	3	30	103	54	195
7	3	31	105	55	197
8	5	32	105	56	199
9	6	33	110	57	202
10	7	34	111	58	203
11	10	35	116	59	204
12	13	36	118	60	204
13	19	37	124	61	204
14	23	38	130	62	205
15	25	39	138	63	205
16	26	40	150	64	205
17	26	41	153	65	205
18	32	42	157	66	205
19	44	43	163	67	205
20	59	44	168	68	205
21	69	45	170	69	205
22	74	46	175	70	206
23	82	47	179		

Appendix B: Dengue Cases, Rainfall and Larvae density

Month	Dengue Cases	Rain (cm)	Larval density	Month	Dengue Cases	Rain (cm)	Larval density
0	8	63	5896	13	44	159.1	5973
1	17	227.46	5070	14	19	65.54	4257
2	19	210.81	4801	15	23	200.4	2959
3	13	261.48	5962	16	12	88.78	4351
4	4	157.6	3813	17	9	111.03	8655
5	6	120.43	4947	18	13	187.04	7100
6	12	129.5	4436	19	32	123.7	7289
7	27	65.48	4878	20	43	82.46	6034
8	21	228.57	9264	21	67	152.82	7964
9	65	117.45	6849	22	108	295.96	12313
10	44	188.48	5281	23	52	88.46	10285
11	14	129.15	8512	24	50	149.77	8700
12	18	103.25	5925				