Chapter 48
Modelling of Infectious Disease with Biomathematics: Implications for Teaching and Research

Sergiy Klymchuk, Ajit Narayanan, Norbert Gruenwald, Gabriele Sauerbier, and Tatyana Zverkova

Abstract The chapter compares a variety of models from biomathematics and bioinformatics of the spread of severe acute respiratory syndrome (SARS) that hit dozens of countries worldwide in 2003. It also investigates students’ and lecturers’ opinions regarding differences in predictions from three different models. All models were based on the real data for Hong Kong published by the World Health Organization (WHO). Although the models were based on the same data, they gave very different predictions of the spread of the disease. The models were discussed with two groups of people: undergraduate students majoring either in engineering or applied mathematics and university lecturers who teach mathematics or mathematical modelling courses. In this chapter we present, analyze, and compare responses to the same questionnaire given to the two groups.

1 Introduction and Framework

Mathematical modelling is a complex process consisting of a number of interrelated steps. Many researchers and practitioners consider skills in mathematical modelling to be different from skills in mathematics. George (1988) states that “model building is an activity which students often find difficult and sometimes rather puzzling. The process of model building requires skills other than simply knowing the appropriate mathematics” (George 1988). Modelling “can be learnt but not taught in a usual way”
(Neunzert and Siddiqi 2000). Relationships between mathematical competencies of students and their skills in modelling were investigated in Galbraith and Haines (1998). Caron and Belair in their exploratory study (2007) examined the phases of the mathematical modelling process that received greater attention from undergraduate students and the competences that were displayed in each phase. They suggested that “some modelling heuristics should be explicitly taught.” They also suggested that “more time should be spent discussing the purpose of a model; this would help students clarify the expected outcome and benefits of each stage” (Caron and Belair 2007).

In many cases a major purpose for doing mathematical modelling of a phenomenon is to make predictions. Taking into account uncertainty and a variety of possible models and a number of assumptions in each model, the task of prediction cannot have the ‘correct’ answer. This fact alone can confuse many students. This chapter investigates students’ opinions regarding differences in predictions from three different models based on the same real data. The task given to the students might look very simple. They neither needed to build a model nor solve the given models. All they needed to do was to read the given real life problem, look at the predictions from three different models and give their reasons for the differences in the predictions. We tested one of the modelling competences described by Kaiser in (2007): “Relating back to the real situation and interpreting the solution in a real-world context”. We also gave the same task to university lecturers who teach mathematics or mathematical modelling courses. Our idea was to compare the responses of the students and lecturers. The main research question was to investigate possible patterns within each group and also similarities and differences between the two groups when they do the same modelling task. In particular, to which extent the two groups use their intuition, common sense and past experience explaining the differences in predictions from three familiar models.

The theoretical framework of this study was based on the works of Haines and Crouch (2001, 2004). A measure of attainment for stages of modelling has been developed in (Haines and Crouch 2001). The authors expanded their study in (Crouch and Haines 2004) where they compared undergraduates (novices) and engineering research students (experts). They suggested a three-level classification of the developmental processes which the learner passes in moving from novice behaviour to that of an expert. One of the conclusions of that research was that “students are weak in linking mathematical world and the real world, thus supporting a view that students need much stronger experiences in building real world mathematical world connections” (Crouch and Haines 2004). This was consistent with the findings from the study by Klymchuk and Zverkova (2001) on possible practical, not cognitive reasons for students’ difficulties linking mathematics and real world. Referring to that study Crouch and Haines wrote: “…students across nine countries all tended to feel that they found moving from the real world to the mathematical world difficult because they lacked such practice in application tasks” (Crouch and Haines 2004).

We believe that doing even simple mathematical modelling activity can be beneficial for students. We agree with Kadjevich who pointed out that “although through solving such … [simple modelling] … tasks students will not realize the examined nature of modelling, it is certain that mathematical knowledge will become alive for them and that they will begin to perceive mathematics as a human enterprise, which improves our lives” (Kadijevich 1999).
2 The Study

2.1 The Models

Infectious diseases hit mankind from time to time on a large scale. In the mid-fourteenth century, the Black Death plague epidemic killed one-third of the population of Europe – 34 million people. In 1563, up to half of the population of London died from the Bubonic plague. In 1918 the Spanish Flu pandemic killed, by different estimations, from 50 to 100 million people worldwide (Patterson and Pyle 1991). An estimated 500 million people were infected. In 1957 the Asian Flu killed up to 4 million people. In 2003 a new highly infectious disease Severe Acute Respiratory Syndrome (SARS) spread rapidly around the world. In March 2003 the WHO, for the first time in history, issued a global warning about the disease. Globally 8,422 people were infected and 916 died. Hong Kong was one of the countries that was hit most by the disease: 1,755 people were infected and 299 died. “Predicting the trend of an epidemic from limited data during early stages of the epidemic is often futile and sometimes misleading. Nevertheless, early prediction of the magnitude of an epidemic outbreak is immeasurably more important than retrospective studies” (Hsieh et al. 2004). Two common types of epidemic models were used to analyze the spread of SARS in Hong Kong. The first type was a long-established susceptible-infected-recovered (SIR) model and its modifications. The model was developed by Kermack and McKerdrick (1927). The fixed population \( N \) is divided into three distinct groups: Susceptible (\( S \)) (those at risk of the disease), Infected (\( I \)) (those that have it), and Removed (\( R \)) (those that are quarantined, dead or have acquired immunity). That is: \( N = S + I + R \). The model is represented by the following system of differential equations:

\[
\begin{align*}
\frac{dS}{dt} &= -rSI \\
\frac{dI}{dt} &= rSI - aI \\
\frac{dR}{dt} &= aI
\end{align*}
\]

where \( r > 0 \) is the infection rate and \( 0 < a < 1 \) is the removal rate.

A typical prediction from this model based on the first 30 days since WHO started publishing daily reports for Hong Kong was given in Shi and Small (2003). The predicted number of infected people was 1,700 versus 1,755 in reality. So the deviation was only 3.1%.

The second model was a relatively new Small-World (SW) network model. The concept of the SW model was imported from the study of social networks into the natural sciences by Watts and Strogatz (1998). As a full epidemic network in Hong Kong was not available, numerical simulations were applied to construct an epidemic chain based on social contacts. The model was established on a grid network weaved by \( m \) parallel and \( m \) vertical lines. Every node in the network represented a
person. The value of $m$ was 2700 for the population $N = m^2 = 7.29 \times 10^6$. The model predicted 1,830 cases, so the deviation was 4.3% (Shi and Small 2003).

Public measures were very important in controlling the epidemic. If the epidemic was allowed ‘to run its natural course,’ in other words, to die down by itself, up to several million people would fall victim to SARS in Hong Kong alone. An epidemic will die down only when the basic reproductive number (number of people infected by a patient) is less than 1. This can be achieved only in two ways: when herd immunity is high enough (natural course of events), or when effective public health measures limit the spread of the epidemic.

Apart from complicated mathematical models, three easier models – linear, exponential and logistic – were used for the analysis of the epidemic. These models were offered as a student project in calculus in Hughes-Hallett et al. (2005). Although the models were based on the same data reported in March 2003, they gave very different predictions of the spread of the disease for June 12, 2003 when the last case was reported in Hong Kong. We decided to ask two groups of people – students and lecturers – about their opinions on the differences in predictions from these three familiar models in an unfamiliar (for students) context. Below is the questionnaire given to the participants of the study.

### 2.2 The Questionnaire

The questionnaire took the following form:

Please read the case below and answer the questions. You don’t need to solve anything.

**Models of the Spread of SARS**

In 2003 a highly infectious disease SARS spread rapidly around the world. Predicting the course of the disease – how many people would be infected, how long it would last – was important to officials trying to minimize the impact of the disease. A number of mathematical models of the spread of SARS were developed to make the predictions. Below are three simple models of the spread of SARS in Hong Kong. We measure time $t$, in days since March 17, the date the World Health Organization (WHO) started to publish daily SARS reports. Let $P(t)$ be the total number of cases reported in Hong Kong by day $t$. On March 17, Hong Kong reported 95 cases. We compare predictions for June 12, the last day a new case was reported in Hong Kong (87 days since March 17). The constants in the differential equations were determined using WHO data from 17 to 31 March (15 days).

(continued)
A Linear Model \( \frac{dP}{dt} = 30.2, \ P(0)=95. \) The prediction for June 12 was 2,722 cases.

An Exponential Model \( \frac{dP}{dt} = 0.12P, \ P(0)=95. \) The prediction for June 12 was 3,249,000 cases.

A Logistic Model \( \frac{dP}{dt} = P(0.19 - 0.0002P), \ P(0)=95. \) The prediction for June 12 was 950 cases.

The actual number of cases on June 12 was 1,755.

Questions:

1. What were possible reasons for the differences in the predictions from the three models above?
2. On what were your reasons from question 1 based (e.g. your experience in modelling, common sense, etc.)?
3. What could make the predictions more accurate?
4. Are you interested in learning more about epidemic modelling and possibly doing research projects in this area (e.g. modelling the spread of swine flu)? Why?

3 The Responses

3.1 Students

The students’ group consisted of first-year undergraduate students majoring in engineering from a German university and second and third year students majoring in applied mathematics from a New Zealand university. Ninety questionnaires were distributed over 2009 and 2010. Forty-eight responses were received so the response rate was 53%. It was a self-selected sample. We systematized and grouped students’ answers into different categories according to the nature of their responses. We used either the key words or exact quotes to name the categories. Some students gave multiple responses to some of the questions and some students did not answer all the questions. The students’ categorized responses are presented below.

1. What were possible reasons for the differences in the predictions from the three models above?
   Different models (16), lack of biological factors (10), different ideas of the speed of spread (8), isolation of infected people (8), population density (6), different assumptions of cases per day, report of cases is not correct (3), different infection rates (3), counter actions, for example pharmaceuticals, different side conditions
(1), different assumptions for each model (1), probability of onset (1), people developed immunity (1), the predictions are theories, which are different from the reality (1), not enough data (1).

2. On what were your reasons from question 1 based (e.g. your experience in modelling, common sense, etc.)?

Common sense (19), mathematical knowledge and experience in modelling (7), both modelling experience and common sense (3), the given information (1), idea of spread of disease (1), I have never seen such problems in [a] mathematical context before, so I don’t know exactly, how to solve it (1), reality, never a constant number of persons will be sick (1), my knowledge about curves of elementary functions (1).

3. What could make the predictions more accurate?

Use experiences from studies of other epidemics, in other regions (14), use more data (7), more knowledge of the virus (3), look for preventive steps, compulsory registration (2), improve data collection (1), average value of cases from 7 days (1), a constant showing the rate of infections (1), side effects like number of travellers to and from Hong Kong (1), information of medical doctors or scientists for the course of disease (1), a study of people behaviour and their health state (1), more facts (1), evaluation of the models (1), the logistic model looks more realistic and it could be improved by using more variables (1), set up a limit of resources (1), adjust the models results to the reality all the time (1), compare the first 2–3 days to find the initial condition (1).

4. Are you interested in learning more about epidemic modelling and possibly doing research projects in this area (e.g. modelling of the spread of swine flu)? Why?

Yes – 4. An interesting topic (4), important for the science on viruses (2).

No – 39. Not my area of interest (18), lack of time (8), this is only making panic (3), don’t see the point to play with numbers or equations which are not correct (2), it is going to have too many factors and the predictions may not be that reliable (1), not enough knowledge in mathematics (1).

3.2 Lecturers

The lecturers’ group consisted of university lecturers from different countries who teach mathematics or mathematical modelling courses. Some of them were involved in research on teaching mathematical modelling and applications. Some of the lecturers were from the same universities as the students participated in the study. Thirty-eight questionnaires were distributed over 2009 and 2010. Twenty-three responses were received so the response rate was 63%. It was a self-selected sample. We systematized and categorized the lecturers’ answers in the same way as the students’ answers. The lecturers’ categorized responses are presented below.

1. What were possible reasons for the differences in the predictions from the three models above?
The models (19), different ideas of the spread of the disease, certain factors were not considered (2), the models were developed for other epidemics, SARS does not fit (1), the assumptions are not the same in all three models (1), did not consider the spread style of the disease (1), infinite number of predictions exist (1).

2. On what were your reasons from question 1 based (e.g. your experience in modelling, common sense, etc.)?
   Experience in modelling (13), common sense (5), both modelling experience and common sense (3).

3. What could make the predictions more accurate?
   More data (6), a better model (3), better parameter estimation (3), knowledge about infection mechanism and other factors e.g. travelling routes, social patterns (2), more accurate analysis of influencing factors (2), a deeper understanding of how infectious disease spreads (1), the parameters in all the models must be the same (1), distribute the observing time in intervals and use different models in different intervals (1), use learning methods (1).

4. Are you interested in learning more about epidemic modelling and possibly doing research projects in this area (e.g. modelling of the spread of swine flu)? Why?
   Yes – 14. Because I work in a similar field, overlapping in research (2), interested in modelling real-life situations to get a grip on it (2), it is important to learn about this modelling since nowadays the disease is increasing (1), interesting subject for my students MS theses (1), modelling of real things is fascinating (1), I am always thrilled how a model helps us in understanding a process and forecasting future data (1), important and relevant area to see maths applied (1), most important in considering today’s swine flu, it is a fascinating subject to consider modelling, teachers and students would be motivated by both the mathematics and the consequences for cities and countries to consider (1).
   No – 8. Not my subject (3), no time (2).

3.3 Analysis of the Responses

After consultations with professional mathematicians specialising in epidemic modelling we estimated percentages of appropriate answers to questions 1 and 3 in both groups. The results are presented in the table below. ‘CS’ means ‘common sense’ and ‘Exp’ means ‘experience’ (Table 48.1).

|        | Question 1 | Question 2 | Question 3 | Question 4 |
|--------|------------|------------|------------|------------|
|        | N | Appropriate | CS | Exp | Both | Other | Appropriate | Yes | No |
| Students | 48 | 73% | 56% | 20% | 9% | 15% | 74% | 9% | 81% |
| Lecturers | 23 | 92% | 24% | 62% | 14% | 0% | 90% | 64% | 36% |
The majority of the students had no or very little experience in mathematical modelling. The closest activity to real mathematical modelling for them was solving application problems. To our surprise the students did quite well in both modelling questions 1 and 3. They were not much behind the lecturers, giving 73% appropriate reasons for the differences in the predictions from the models versus 92% given by the lecturers. They were not much behind the lecturers giving 74% appropriate ways to improve the accuracy of the predictions in the models versus 90% given by the experts. This is consistent with the findings of Haines and Crouch (2001, 2004) where the researchers found that sometimes novices exhibited aspects of expert behaviour although they were not consistent in doing so. In particular, in their study on self-assessment and tutor assessment they found that students were almost as good as tutors in assessing group (project) presentations on modelling and so they could recognize modelling behaviour in others. It is the consistency that demonstrates expert behaviour that perhaps the differentiates lecturers.

In question 2 the reverse polarity on the answers by the students and the lecturers was anticipated: the students relied more on common sense (56%) rather than on experience (20%) compared to the lecturers (24% on common sense and 62% on experience). Apart from lack of modelling experience by the students, one of possible reasons for this reverse polarity might be elements of the lecturers’ behaviour where they were reluctant to attribute their responses to common sense, preferring to classify them as experience. After all they have invested a great deal of time in mathematics/modelling.

Based on the participants’ comments in the questionnaire and follow-up interviews with some of them, we attempted a comparison of the processes used by the students and the lecturers in terms of links between the mathematical world and the real world in a similar way to that done by Crouch and Haines (2004). We took the first “level (a) where there was clear evidence that the participants took into account the relationship between the mathematical world and the real world” (Crouch and Haines 2004). The students referred explicitly to that relationship in 65% of cases (though not always in a correct way) whereas the lecturers in 20% of cases. The lecturers tended to concentrate more on the mathematical aspects of the models probably implicitly assuming that relationship. One of the possible reasons might be that the lecturers used their experience in modelling and knowledge in mathematics much more than their common sense whereas the students relied more on their common sense and life experiences, lacking the experience in mathematical modelling.

In question 4 very few students (9%) reported that they were interested in doing a research project in epidemic modelling. This was understandable taking into account that the majority were majoring in engineering. The lecturers were more enthusiastic in doing research in epidemic modelling (64%), mostly because of the importance of the topic and/or relevance to their current research. The lecturers also indicated that the topic was very useful for teaching purposes because it was timely and could increase students’ motivation.
4 Conclusions

This study indicates that in spite of lack of experience in real mathematical modelling, students can effectively use their common sense and general knowledge of mathematics to evaluate some modelling issues dealing with prediction. The responses at a more general level indicated that both students and lecturers would have preferred to include more parameters in the model to make the modelling more realistic and intuitive, that is, to have a theoretical basis for the modelling that included hypothetical rates of spread, infection mechanisms, etc. It is possible that engineering students, in particular, would have engaged more with the modelling exercise if they saw it as a parameter optimisation problem so that the model was both explanatory and predictive.

We are very aware of the limitations of the study. It was intended as a pilot study to check our assumptions and share the findings with the mathematics education community. Future work should explore students’ and lecturers’ (i.e., novices and experts according to Haines and Crouch 2004) responses to more sophisticated mathematical models that allow for the adjustment of parameters to optimize the output from the model.

Acknowledgements We would like to express our gratitude to Professor Chris Haines from City University, UK for his help with analyzing participants’ responses.

References

Caron, F., & Belair, J. (2007). Exploring university students’ competences in modelling. In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), Mathematical modelling (ICTMA 12): education, engineering and economics (pp. 120–129). Chichester: Horwood.

Crouch, R., & Haines, C. (2004). Mathematical modelling: Transitions between the real world and the mathematical world. International Journal on Mathematics Education in Science and Technology, 35(2), 197–206.

Galbraith, P., & Haines, C. (1998). Some mathematical characteristics of students entering applied mathematics courses. In J. F. Matos et al. (Eds.), Teaching and learning mathematical modelling (pp. 77–92). Chichester: Albion.

George, D. A. R. (1988). Mathematical modelling for economists. London: Macmillan.

Haines, C., & Crouch, R. (2001). Recognizing constructs within mathematical modelling. Teaching Mathematics and Its Applications, 20(3), 129–138.

Hsieh, Y., Lee, J., & Chang, H. (2004). SARS epidemiology modeling. Emerging Infectious Diseases, 10(6), 1165–1167.

Hughes-Hallett, D., Gleason, A., McCallum, W., et al. (2005). Calculus: Single and multivariable (4th ed.). Hoboken: Wiley.

Kadijevich, D. (1999). What may be neglected by an application-centred approach to mathematics education? Nordisk Matematikkdidatikk, 1, 29–39.

Kaiser, G. (2007). Modelling and modelling competences in school. In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), Mathematical modelling (ICTMA 12): Education, engineering and economics (pp. 110–119). Chichester: Horwood.
Kermack, W. O., & McKerdrick, A. G. (1927). Contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, 115*, 700–721.

Klymchuk, S., & Zverkova, T. (2001). Role of mathematical modelling and applications in university service courses: An across countries study. In J. F. Matos, W. Blum, S. K. Houston, & S. P. Carreia (Eds.), *Modelling and mathematics education: ICTMA-9: Applications in science and technology* (pp. 227–235). Chichester: Horwood.

Neunzert, H., & Siddiqi, A. H. (2000). *Topics in industrial mathematics*. New York: Springer.

Patterson, K. D., & Pyle, G. F. (1991). The geography and mortality of the 1918 influenza pandemic. *Bulletin of the History of Medicine, 65*(1), 4–21.

Shi, P. and Small, M. (2003). Modelling of SARS for Hong Kong. *Populations and Evolution*. http://arxiv.org/abs/q-bio/0312016. Accessed 13 Sep 2009.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature, 393*, 440–442.