A flexible and precise model for the dynamics of animal diseases with single outbreaks

W.G. Nowak

BOKU Wien – University of Natural Resources and Applied Life Sciences, Vienna, Austria

To Professor Hubert Dürrstein, Rector of BOKU Wien, on his 50th birthday

ABSTRACT: This article provides a general and accurate mathematical model for the epidemics dynamics of a large class of animal diseases. The issue is to offer to the scientist interested in applications a quite concise and practicable explanation how to validitate the model parameters by means of the software tool Microsoft Excel®. Furthermore, predictions concerning hypothetical scenarii can easily be obtained on the basis of this concept. As specific examples, calculations are presented for the BSE epidemics in the British UK and for an outbreak of foot-and-mouth disease in a district of Lower Austria.

Keywords: epidemics; animal diseases; mathematical model; BSE; foot-and-mouth disease

1. Introduction. The search for precise mathematical descriptions of epidemics, both human and animal, has a long and successful history. It started from classic articles like those of Kermack and McKendrick (1927, 1932, 1933) and has been exhibited in monographs of the highest scientific level: See, e.g., Bayley (1975), Murray (1990, ch. 19, 20), Grenfell and Dobson (1995), Schuster (1995), Isham and Medley (1996), and Diekmann and Heesterbeek (1999). These theories involve deep tools from the highest mathematics, including partial differential equations. Usually, little attention is paid to the questions of fitting such models to empirical data and to the accuracy of match between theory and practice. The present article attempts to help fill this gap by describing a model which is relatively simple and, at the same time, sufficiently general for applications to very diverse animal epidemics. The general scope of the discussion will be an animal disease characterized by the following properties:

- The epidemics breaks out, after a while reaches a maximum, and ultimately disappears again, at least for the region and time span under observation.
- The decline is effected by drastic and extensive human actions. There is a certain time interval which elapses until these provisions are brought to full effect.
- There may be a latency period T whose length is not negligible compared to the total duration of the epidemics.
- A sure diagnosis of the disease is always connected with the death of the infested animal.

As two rather divergent examples, BSE and foot-and-mouth disease will be considered. Our issue is to provide for the scientist aimed at applications a very concise explanation how to implement the model on the basis of concrete data, by means of the software package Microsoft Excel®, and how to gain possible predictions as well as computational results on hypothetical scenarii.

2. Basic assumptions of the model. Our considerations apply to one fixed infectious disease involving one kind (species) of animals, restricted to some given geographical region. We define

(continuously differentiable) model functions for the following quantities, all depending on the time variable t:

$g(t)$ = the number of infected[1] animals at time t

$h(t)$ = the number of infectious animals at time t (i.e., those which are likely to transmit the disease to others)

$f_*(t)$ = the number of animals who died/were killed in the time interval $[0, t]$ in an infected state,

$f(t)$ = the number of animals that were slaughtered, resp. culled in the time range $[0, t]$ and were diagnosed as being infectious with the disease

Obeying to the principle that every mathematical model has to compromize between accuracy to reality and relative simplicity, we assume that between these four quantities the following relations hold (c_0, c_1, ... are numerical constants throughout):

(A) The relative proportion of undetected cases is invariant in time, i.e.

$$f_*(t) = c_0 f(t)$$

(B) Let T denote the latency period of the disease. There are $h(t)$ infectious animals at time t, which necessarily must have been infected already at time $t - T$. We suppose that they are a constant proportion of *all* infected animals alive at time $t - T$, i.e.[2]

$$h(t) = c_1 g(t - T)$$

(C) The frequency of slaughtered or culled animals with a positive diagnosis of the disease is always proportional to the number of infectious animals[3]:

$$f'(t) = c_2 h(t)$$

Next we are going to analyze the change rate $g'(t)$ of the number $g(t)$ of infected animals. Obviously $g(t)$ is reduced by kills/deaths whose number at time t is $f_*'(t) = c_0 f'(t)$. On the other hand, it increases by new infections, whose number initially should be proportional to the number of infectious animals, say, $c_3\, h(t)$.

Typically, at a certain time drastic and effective actions are set (we leave the details for special cases to the reader's expertise), which ultimately reduce the factor c_3 to some substantially smaller value c_4. Let[4] $[t_1 - L, t_1 + L]$ denote the time interval which it takes until these provisions against the spread of the disease are fully organized and activated. Summarizing what we said, we obtain

$$g'(t) = -c_0 f'(t) + \begin{cases} c_3 h(t) & \text{for } t \le t_1 - L, \\ c_4 h(t) & \text{for } t \ge t_1 + L \end{cases}$$

Combining this with (C) and (B), we infer that

$$\frac{g'(t)}{g(t-T)} = \begin{cases} c_1 (c_3 - c_0 c_2) & \text{for } t \le t_1 - L, \\ c_1 (c_4 - c_0 c_2) & \text{for } t \ge t_1 + L \end{cases}$$

We avoid an exact analysis of this differential-difference equation; see, e.g., Bellman and Cooke (1963) for a comprehensive classic account on this intrinsic theory. For our present purpose we approximate $g(t - T)$ by its linear Taylor expansion $g(t) - T g'(t)$, to get after a short calculation[5]

$$\frac{g'(t)}{g(t)} = \begin{cases} \dfrac{c_1 (c_3 - c_0 c_2)}{T c_1 (c_3 - c_0 c_2) + 1} =: c_5 & \text{for } t \le t_1 - L, \\ \dfrac{c_1 (c_4 - c_0 c_2)}{T c_1 (c_4 - c_0 c_2) + 1} =: c_6 & \text{for } t \ge t_1 + L \end{cases}$$

After replacing t by $t - T$, one more appeal to (B) and (C) furnishes

$$\frac{f''(t)}{f'(t)} = \begin{cases} c_5 & \text{for } t \le t_0 - L \\ c_6 & \text{for } t \ge t_0 + L \end{cases}$$

[1] $g(t)$ counts the number of *all* infected animals at time t, no matter if they are already infectious or not. Therefore, $g(t) \ge h(t)$ throughout.

[2] If many animals are being infected at a certain time $t - T$, say, then clearly the amount of *infectious* animals will increase drastically as soon as the latency period T has passed. This dependance is modeled in the simplest possible way, namely by a linear equation.

[3] This assumption implies that the relative chances of an infectious animal to be diagnosed remain constant during the outbreak of the disease. I.e., the provisions to test animals and detect the infectious ones are not overrun by the increasing number of cases.

[4] It will turn out convenient to use this notation, instead of, say, $[t_1, t_2]$. This helps to keep the final formulas relatively simple. The same applies to $u + v$, $u - v$ used a bit later.

[5] The deeper mathematical truth may be sketched as follows: The equation $g'(t)/g(t-T) = c$ either possesses an exponential solution; for this, $g'(t)/g(t)$ is a constant, too, and our approximation is essentially correct. Or its solutions are all oscillating functions: This possibility may be ruled out by our general assumption that the epidemics has only one single outbreak.

with the definition $t_0 = t_1 + T$. It is simple and reasonable to assume that the value of the right-hand side decreases linearly from c_5 to c_6, while t ranges from $t_0 - L$ to $t_0 + L$. Writing $c_5 := u + v$, $c_6 := u - v$ for convenience, we thus arrive at the final model equation

$$\frac{f''(t)}{f'(t)} = Z(t) := \begin{cases} u+v & \text{for } t \le t_0 - L \\ -\frac{v}{L}(t-t_0)+u & \text{for } t_0 - L \le t \le t_0 + L \\ u-v & \text{for } t \ge t_0 + L \end{cases} \quad (1)$$

Observe that all the four parameters involved in the definition of (except the variable t) are constants which ultimately will be fitted to concrete data.

3. Solving the differential equation. Integrating eq. (1), the left-hand side becomes $\ln f'(t)$, hence taking exponentials and integrating one more time yields

$$f(t) = \int \exp\left(\int Z(t)\mathrm{d}t\right)\mathrm{d}t$$

Therefore, on each of the intervals $t \le t_0 - L$ and $t \ge t_0 + L$, the general solution reads

$f(t) = A + B\, e^{(u+v)t}$, resp., $f(t) = E + F\, e^{(u-v)t}$

with constants A, B, E, F at our disposal. On the intermediate range the analysis is more subtle. Since here $\int Z(t)dt$ gives a quadratic polynomial, the ultimate result involves the Gauss error function. In fact, let $N(x; s)$ denote the cumulative probability function of a normal distribution with mean 0 and standard deviation s, i.e.,

$$N(x;s) = \frac{1}{s\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-u^2/(2s^2)\right)\mathrm{d}u = \frac{1}{2}\left(1 + \mathrm{erf}\left(x/(s\sqrt{2})\right)\right)$$

Then the general solution of the differential equation (1) reads

$$f(t) = \begin{cases} A + Be^{(u+v)t} & \text{for } t \le t_0 - L \\ C + DN\left(-v(t-t_0) + Lu;\ \sqrt{Lv}\right) & \text{for } t_0 - L \le t \le t_0 + L \\ E + Fe^{(u-v)t} & \text{for } t \ge t_0 + L \end{cases} \quad (2)$$

Here the coefficients $A, B, \ldots, F$ are not completely independent. In fact, writing $f_1(t), f_2(t), f_3(t)$, for the three function terms at the right-hand side of (2), the plausible condition that the graph of $f(t)$ should be connected and smooth, also at $t = t_0 \pm L$, leads to the equations

$$f_1(t_0 - L) = f_2(t_0 - L), \quad f_1'(t_0 - L) = f_2'(t_0 - L)$$

$$f_2(t_0 + L) = f_3(t_0 + L), \quad f_2'(t_0 + L) = f_3'(t_0 + L)$$

From this, a lengthy but straightforward computation yields

$$D = -\sqrt{2\pi}\exp\left(L(u^2 - v^2)/(2v) + t_0(u+v)\right)\left(\frac{u}{v} + 1\right)\sqrt{Lv}\ B$$

$$C = Be^{(t_0 - L)(u+v)} - DN\left(L(u+v);\ \sqrt{Lv}\right) + A$$

$$F = \frac{\sqrt{v}}{\sqrt{2\pi L(v-u)}}\exp\left(L\left(v^2 - u^2\right)/(2v) + t_0(v-u)\right)\ D$$

$$E = -Fe^{(L+t_0)(u-v)} + DN\left((u-v)\sqrt{L};\ \sqrt{v}\right) + C \quad (3)$$

I.e., if the parameters t_0, L, u, v, and A, B, are known, the coefficients C, D, E, F can be computed successively.

4. Example 1: The BSE epidemics in the British UK. According to the website of the British Department for Environment, Food and Rural Affairs of the UK (DEFRA, 2004) in the years 1987 through 2003 the following numbers of confirmed cases of Bovine Spongiforme Encephalopathy in cattle have been reported for the whole United Kingdom:

{446,1641,6958,13038,22931,35276,37011,26096, 15710,8879,4933,3502,2703,1633,1103,1211,670} (4)

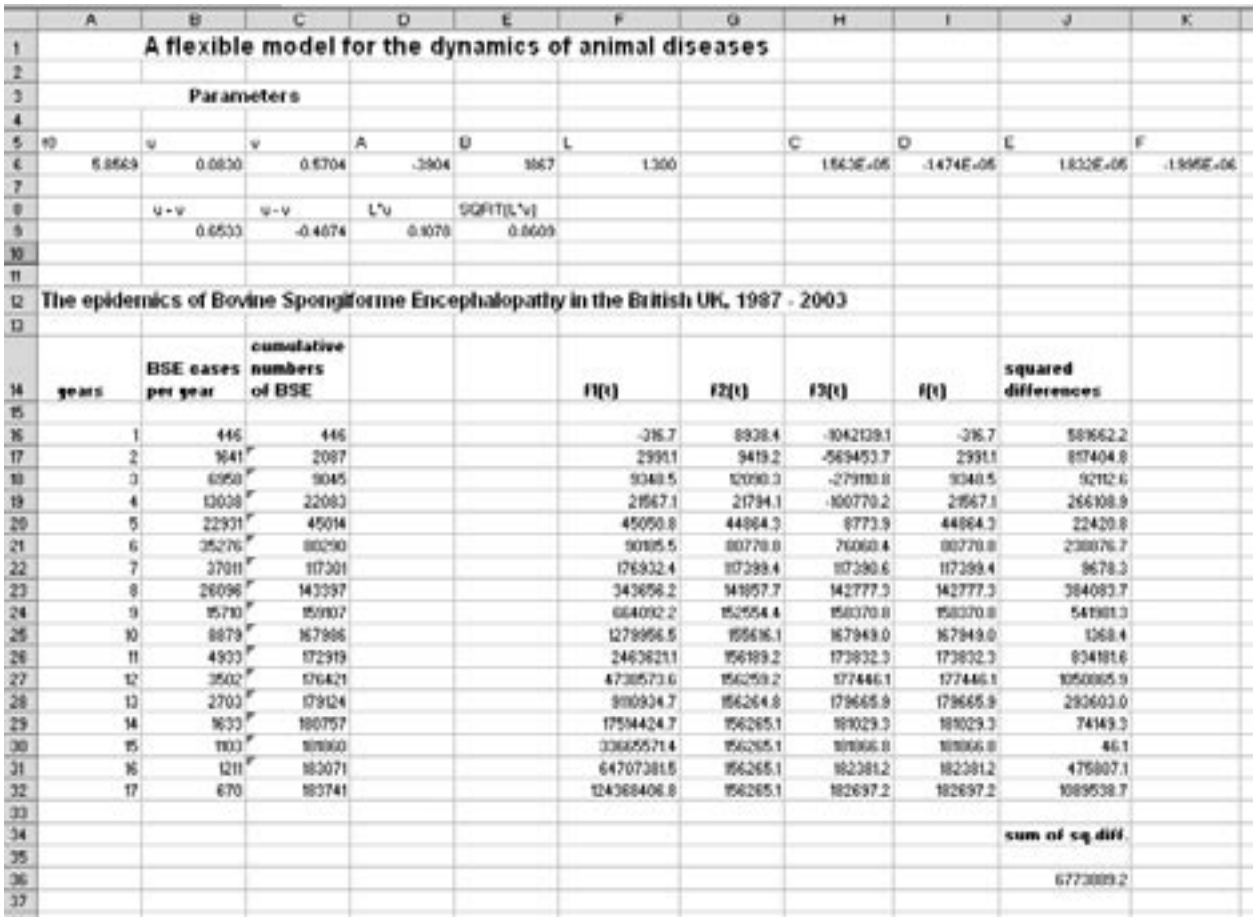
For positive integers $t \le 17$, let $F(t)$ denote the sum of the first t numbers in the list of data (4). Then, by the foregoing theory, the values $F(t)$ and $f(t)$ should approximately coincide for $t = 1, \ldots, 17$. I.e., by the method of least squares, the parameters t_0, L, u, v, A, B are to be determined such that they minimize the sum

$$Q = \sum_{t=1}^{17}\left(F(t) - f(t)\right)^2$$

4.1. Validating the parameters with Excel®. We describe how to carry out this task with the help of the software tool Microsoft Excel® (see Figure 1). To start with the parameters involved, we attribute (by means of the menu options Insert – name – create) to the cells in line 6 the names $t_0, \ldots, F$ written above in line 5. For convenience, we notice below in cells B9 through E9 the values of $u + v$, $u - v$, Lu and $\sqrt{Lv}$ which occur in the model function (2). Into cells H6 through K6 we enter the definitions contained in the formulas (3). E.g., into cell H6 we write

= B*EXP(($t_0 - L$)*($u + v$)) – D*NORMDIST(L*($u + v$); 0;SQRT(L*v);TRUE) + A

and accordingly in I6, J6, and K6.



	A	B	C	D	E	F	G	H	I	J	K
1	A flexible model for the dynamics of animal diseases										
2											
3	Parameters										
4											
5	t0	u	v	A	B	L		C	D	E	F
6	5.8569	0.0830	0.5704	-.3804	.9867	1.300		1.563E-06	-.1474E-06	1.832E-06	-1.995E-06
7											
8		u - v	v - u	L*u	SQRT(L*v)						
9		0.6533	-0.4874	0.1078	0.6609						
10											
11											
12	The epidemics of Bovine Spongiforme Encephalopathy in the British UK, 1987 - 2003										
13											
14	years	BSE cases per year	cumulative numbers of BSE			f1(t)	f2(t)	f3(t)	f(t)	squared differences	
15											
16	1	446	446			-216.7	8938.4	-1042139.1	-216.7	581662.2	
17	2	1641	2087			2991.1	9419.2	-569453.7	2991.1	817404.8	
18	3	6958	9045			9340.5	10090.3	-279180.8	9340.5	92112.6	
19	4	13038	22083			21567.1	21794.1	-100770.2	21567.1	266108.9	
20	5	22931	45014			45050.8	44864.3	8773.9	44864.3	22420.8	
21	6	35276	80290			90805.5	80770.8	76060.4	80770.8	238076.7	
22	7	37011	117301			176932.4	117289.4	117280.6	117289.4	9678.3	
23	8	26096	143397			343656.2	141857.7	142777.3	142777.3	384083.7	
24	9	15710	159107			664092.2	152554.4	158370.8	158370.8	541981.3	
25	10	8879	167986			1279956.5	155616.1	167949.0	167949.0	1368.4	
26	11	4933	172919			2463823.1	156189.2	173832.3	173832.3	834181.6	
27	12	3502	176421			4738573.6	156259.2	177446.1	177446.1	1050865.9	
28	13	2703	179124			9180934.7	156264.8	179665.9	179665.9	293603.0	
29	14	1633	180757			17594424.7	156265.1	181029.3	181029.3	74149.3	
30	15	1103	181860			33655571.4	156265.1	181866.8	181866.8	46.1	
31	16	1211	183071			64707381.5	156265.1	182381.2	182381.2	475807.1	
32	17	670	183741			124368406.8	156265.1	182697.2	182697.2	1089538.7	
33											
34										sum of sq diff.	
35											
36										6773889.2	
37											

Figure 1. Microsoft Excel® sheet for the validation of the model parameters in the UK BSE-epidemics

The next step is to put in the specific data of the British BSE epidemics. After entering the integers 1 to 17 in the A-column, we fill in the numbers of BSE-cases per year from the list of data (4) into the cells B16 through B32. To get in the column C the cumulative BSE-numbers, we write into cell C16:

= SUM(C16: C16)

and copy this down until cell C32. The columns *F*, *G*, *H* are to receive the definitions of the three functions on the right-hand side of (2). Thus we write into F16

$= A + B^*\mathrm{EXP}((u + v)^*\mathrm{A16})$.

Similarly, cell G16 gets the entry

$= C + D^*\mathrm{NORMDIST}(-v^*(\mathrm{A16} - t_0) + L^*u; 0; \mathrm{SQRT}(L^*v); \mathrm{TRUE})$

Furthermore, into H16 we write

$= E + F^*\mathrm{EXP}((u - v)^*\mathrm{A16})$

The distinction of three cases, as required by (2), is effected by entering into I16

$= IF(\mathrm{A16} < t_0 - L; \mathrm{F16}; IF(\mathrm{A16} > t_0 + L; \mathrm{H16}; \mathrm{G16}))$

Column *J* is to contain the squared differences $(F(t) - f(t))^2$, thus we put into J16

$= (\mathrm{C16} - \mathrm{I16})\text{^}2$

Now we mark cells F16 through J16 and copy down till line 32. Finally, the auto-sum device (or, = SUM(J16:J32)) attributes the sum Q to the cell J36.

We are now set to determine the parameters t_0, u, v, A, B, L, with the help of the Solver of Microsoft Excel®. First of all, reasonable initial values are required for our search. A look at a plot of the data points, from columns *A* and *C*, suggests that $t_0 \approx 6$. (see Figure 2.) Further, a little reflection shows that $u + v > 0$ and $u - v < 0$ are in the 10^{-1} range, hence we start with $u \approx 0.2$, $v \approx 0.4$, and further $L \approx 1$. Minimizing cell J36, we first allow only *A*, *B* (i.e., cells D6 and E6) to vary. We apply the Solver once

more, with cells A6 – E6 variable, and a third time with all of A6 – F6 flexible. This yields the (rounded) parameter values[6]

$$t_0 = 5.857,\ u = 0.083,\ v = 0.5704,\ A = -3\ 904,$$
$$B = 1\ 867,\ L = 1.3$$

Automatically we get as well

$$C = 1.563 \times 10^5,\ D = -1.474 \times 10^5,\ E = 1.832 \times 10^5,$$
$$F = -1.995 \times 10^6.$$

The model function for the cumulative numbers of confirmed cases of BSE in cattle in the whole British UK (starting with $t = 0$ at the beginning of 1987) thus reads

$$f(t) = \begin{cases} -3904 + 1867e^{0.6533t} & \text{for } t \le 4.557 \\ 1.563\times10^5 - 1.474\times10^5\, N(-0.5704(t-5.857) + \\ 0.1078; 0.8609) & \text{for } 4.557 \le t \le 7.157 \\ 1.832\times10^5 - 1.995\times10^6\, e^{-0.4874t} & \text{for } t \ge 7.157 \end{cases}$$

4.2. Discussion of the parameter values estimated. Let us recall the practical meaning of some of the constants involved:

- $t_0 = t_1 + T$, where t_1 is the mid-point of the time-interval where actions against the disease are being brought to full effect, and T is the (maximal) latency period. For BSE, $T \approx 4$ years roughly, hence $t_1 \approx 1.85$ ($t = 0$ corresponding to the beginning of 1987), which is a good guess, since the ban of feeding cattle by meat-and-bone meal came into force in the UK in July 1988.
- $2L$ is the length of the time-span until anti-epidemics provisions were fully effective. Again, $2L = 2.6$ years appears quite plausible, recalling how long it took until the threat of BSE was recognized and adequate measures were supported by the majority of people involved.
- For later reference we note that v is a measure for the effectiveness of the provisions taken: Recall that the right-hand side of eq. (1) is decreased from $u + v$ to $u - v$.

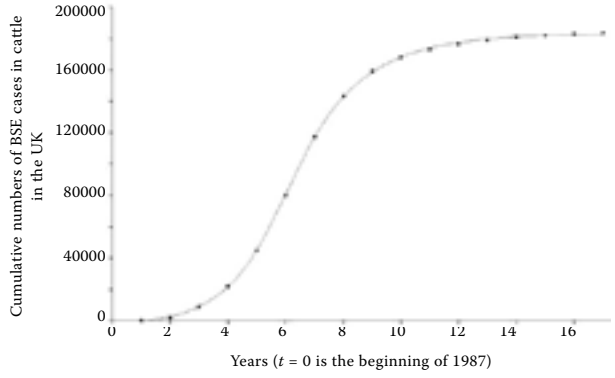


Figure 2. Cases of BSE in cattle in the British UK, 1987–2003, cumulative count. Model function and data points, according to DEFRA (2004)

4.3. Computational investigations of hypothetical scenarii. It should be emphasized that the salient point of the model is not (only) that it matches well empirical data. The important point is that it easily provides the possibility to calculate what would have happened if some conditions (i.e., certain parameters) would have been different. E.g., one can assume the anti-disease actions earlier or later, quicker or slower, more or less efficient than they really were, just by entering different values for t_0, L, v, resp., on the Excel® sheet.

In the present example, we ask how the cumulative loss of cattle by BSE would have developed, (a) if anti-epidemic actions would have been delayed for one year, (b) if they would have taken place half a year earlier. It suffices to put the value of $t_0 = 6.857$ in the first case, and $t_0 = 5.357$ in the second case. The other parameters change automatically, and we obtain from (2) the respective hypothetic model functions

$$f_a(t) = \begin{cases} -3904 + 1867e^{0.6533t} & \text{for } t \le 5.557 \\ 3.039\times10^5 - 2.833\times10^5\, N(-0.5704(t-6.857) + \\ 0.1078; 0.8609) & \text{for } 5.557 \le t \le 8.157 \\ 3.557\times10^5 - 6.242\times10^6\, e^{-0.4874t} & \text{for } t \ge 8.157 \end{cases}$$

and

$$f_b(t) = \begin{cases} -3904 + 1867e^{0.6533t} & \text{for } t \le 4.057 \\ 1.116\times10^5 - 1.063\times10^5\, N(-0.5704(t-5.357) + \\ 0.1078; 0.8609) & \text{for } 4.057 \le t \le 6.657 \\ 1.311\times10^5 - 1.128\times10^6\, e^{-0.4874t} & \text{for } t \ge 6.657 \end{cases}$$

The graphics in Figure 3 displays the drastic effect of earlier or later response to the BSE threat on the total damage caused by the disease. We remark parenthetically that similar calculations work well for smaller parts of the UK, on the basis of the data provided by DEFRA 2004. For the special case of BSE, earlier variants of the model have been exposed in Kühleitner and Nowak (2003), and in Nowak (2004).

[6] The numerical values obtained may depend slightly on the version of Microsoft Excel® employed and on the setting of the Solver options.

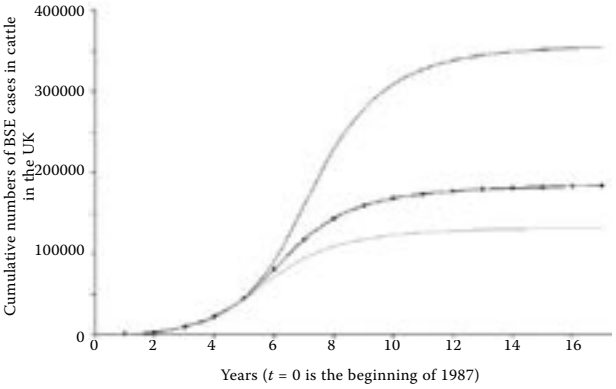


Figure 3. BSE epidemics in the UK, calculations of hypothetical scenarii, assuming counter actions in effect one year later, resp., half a year earlier

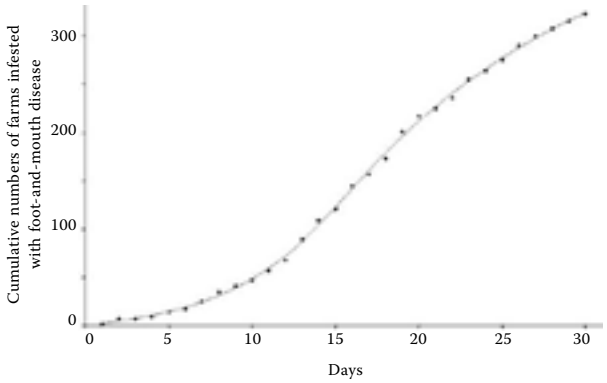


Figure 4. Farms infested by foot-and-mouth disease in a district of Lower Austria, April/May 1973, cumulative count. Model function and data points

5. Example 2: Foot-and-mouth disease in Austria 1973. Our second application of the model concerns a very much different epidemics, namely the outbreak of foot-and-mouth disease in Austria in 1973. By a lucky coincidence, the data of one district of Lower Austria (Mistelbach), for a range of 30 days in April/May 1973, are available in the book of Timischl (1995, p. 6):

{2,5,0,2,5,3,8,10,6,6,10,11,21,20,12,24,12,16,28,
16,8,11,19,9,11,15,10,8,7,8}.

Here the count concerns whole farms reported newly infested with foot-and-mouth disease per day, not individual animals; this, however, does not affect the reasoning in the construction of the model. We may in fact use the same Excel® sheet as before. The only essential change is in column *B* where the new data are to be entered.[7] Further, everything must be copied down until 30 lines altogether. The Excel® Solver readily yields the (rounded) parameter values

$t_0 = 14.48$, $u = 0.0546$, $v = 0.1216$, $A = -9.29$
B = 9.86, L = 3.28, C = 264,54, D = –247.39
E = 441.73, F = –877.77

Using these in (2) furnishes the model function

$$f(t) = \begin{cases} -9.29 + 9.86e^{0.1762t} & \text{for } t \le 11.21 \\ 264.54 - 247.39\, N(-0.1216(t - 14.48) + 0.179; \\ 0.6312) \quad \text{for } 11.21 \le t \le 17.76 \\ 441.73 - 877.77\, e^{-0.067t} & \text{for } t \ge 17.76 \end{cases}$$

Figure 4 shows again an excellent match with the data. We add a few comments on the parameter values obtained. Firstly, $L = 3.28$ suggests that it took about six and a half days until anti-epidemic actions were in full effect. Considering that these included traffic restrictions and cancellations of mass events, apart from the activation of the public alertness to the threat, this might seem plausible. Further, since the (maximal) latency period with FMD is about $T = 7$ days (including one or two days of sometimes delayed diagnosis), we infer that $t_1 = t_0 - T \approx 7.5$. Hence the beginning $t_1 - L$ of counter actions should have been a few days after the initial outbreak, which is again quite reasonable.

We conclude the discussion of this example by two predictions on hypothetical scenarii. This time we ask: What would have happened if the anti-epidemics actions were more or less efficient? To answer this, on our Excel® sheet we

(a) increase v by 0.03 (and decrease u by the same amount to keep $u + v$ constant), and
(b) decrease v by 0.03 (increasing u by 0.03).

Excel® readily yields the corresponding modified parameter values which by (2) give the new model functions

$$f_a(t) = \begin{cases} -9.29 + 9.86e^{0.1762t} & \text{for } t \le 11.21 \\ 223.59 - 203.95\, N(-0.1516(t - 14.48) + 0.0806; \\ 0.7048) \quad \text{for } 11.21 \le t \le 17.76 \\ 282.78 - 1103.92\, e^{-0.127t} & \text{for } t \ge 17.76 \end{cases}$$

and

$$f_b(t) = \begin{cases} -9.29 + 9.86e^{0.1762t} & \text{for } t \le 11.21 \\ 340.94 - 326.94\, N(-0.0916(t - 14.48) + 0.2773; \\ 0.5478) \quad \text{for } 11.21 \le t \le 17.76 \\ 3315.73 - 3544.16\, e^{-0.007t} & \text{for } t \ge 17.76 \end{cases}$$

[7] In this example the per-day data show rather little regularity. This can be visualized by a point plot of the columns *A* and *B* alone. Nevertheless, for the cumulative data our model works well again.

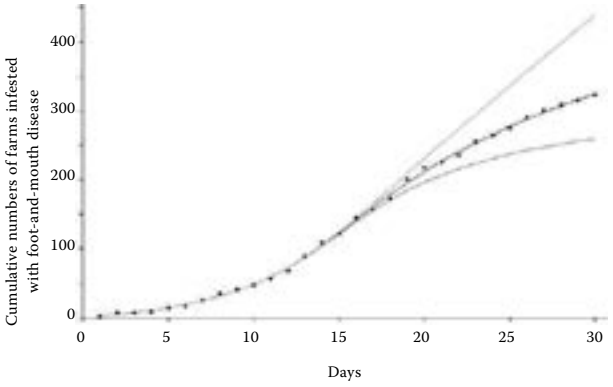


Figure 5. Foot-and-mouth disease in a district of Lower Austria 1973, calculations of hypothetical scenarii, assuming counter actions 25% more, resp., less effective than in reality

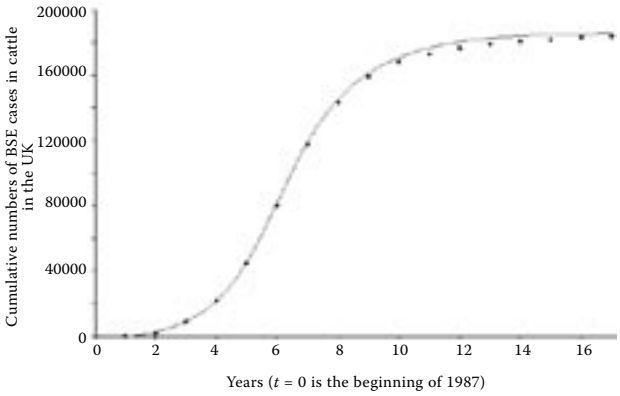


Figure 6. Model function with coefficients $A, B, \ldots,$ recalculated from the first 4 data points only, simulating a prognosis from few initial data to the total development of the outbreak

Figure 5 shows for comparison the graphs of $f(t)$, $f_a(t)$, and $f_b(t)$, giving thus a vivid impression of the influence on the development of the cumulative numbers of afflicted farms of an about 25% more or less effective strategy against the spread of the epidemics.

6. Predictions from few initial data. To stress the applicability of our model, we proceed to show how it admits to predict, with reasonable accuracy, the development of an epidemic outbreak from relatively few data from the initial period. In fact, returning to the example of BSE in the UK, let us suppose we knew only the data of the first 4 years. Assume further that counter actions were as timely and efficient as in reality[8]: I.e., the values of u, v, t_0, and L remain unchanged, while the coefficients A, B are determined anew by a least squares fit involving the first 4 data points only. From these, new values of C, D, E, and F are calculated. This task is carried out most easily, using the Excel® sheet of Figure 1: In cell J36 the definition is changed to = SUM(J16: J19) and the Solver is used again to minimize J36, with only cells D6 and E6 variable now. This yields the new coefficients

$$A = -4\,096,\ B = 1\,898,\ C = 1.588\times10^5,\ D = -1.499\times10^5,\ E = 1.862\times10^5,\ F = -2.029\times10^6,$$

and the "prognosis" model function

$$f_{\text{prog}}(t) = \begin{cases} -4096 + 1898e^{0.6533t} & \text{for } t \le 4.557 \\ 1.558\times10^5 - 1.499\times10^5\, N(-0.5704(t-5.857) + 0.1078;\, 0.8609) & \text{for } 4.557 \le t \le 7.157 \\ 1.862\times10^5 - 2.029\times10^6\, e^{-0.4874t} & \text{for } t \ge 7.157 \end{cases}$$

Figure 6 shows its graph together with all data points and thus confirms the accuracy of the prediction. We can carry out as well analogous calculations for the FMD example, fitting the model function first to the initial 10 data points (out of 30),

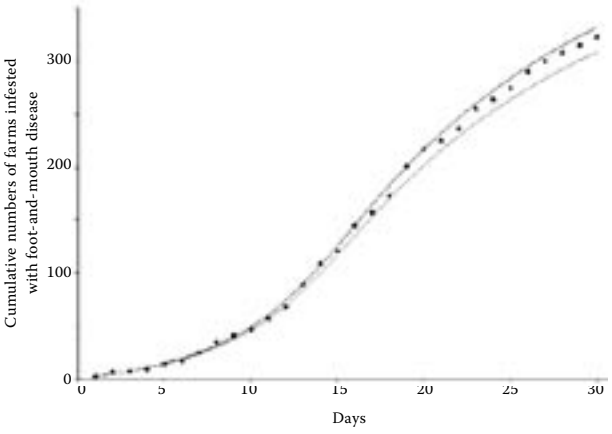


Figure 7. Simulating predictions on a FMD epidemics: Model function with coefficients $A, B, \ldots, F$ calculated from the first 7 (lower curve), resp., 10 data points (upper curve)

[8] It is a priori clear that an assumption of this kind is needed. Initial data alone cannot provide information about counter actions undertaken later. Further, the analysis of other BSE data as given by DEFRA (2004) confirms (at least approximately) that it is reasonable to consider u, v, t_0, and L as constants for BSE outbreaks under similar circumstances. Of course, t_0 may be adapted, e.g., according to the date of the ban of meat-and-bone meal.

then only to the first 7. Allowing again the cells D6 and E6 to vary, we get, using the subscripts 10 and 7 with an obvious meaning,

$$A_{10} = -9.79,\ B_{10} = 10.11,\ C_{10} = 271.02,\ D_{10} = -253.7,$$
$$E_{10} = 452.72,\ F_{10} = -900.14,\ A_7 = -8.61,\ B_7 = 9.38,$$
$$C_7 = 252,\ D_7 = -235.45,\ E_7 = 420.63,\ F_7 = -835.4.$$

This gives the prognosis function

$$f_{\text{prog, }10}(t) = \begin{cases} -9.79 + 10.11e^{0.1762t} & \text{for } t \le 11.21 \\ 271.02 - 253.7\, N(-0.1216(t - 14.48) + \\ 0.179; 0.6312) & \text{for } 11.21 \le t \le 17.76 \\ 452.72 - 900.14e^{-0.067t} & \text{for } t \ge 17.76 \end{cases}$$

and similarly for $f_{\text{prog, }7}(t)$. Their graphs are displayed in Figure 7, showing both a satisfactory match with the data points of the whole outbreak of FMD. Of course, one can also gain predictions for epidemic events based on initial data along with certain hypotheses that the counter actions are undertaken earlier or later, more or less efficiently than in some cases observed before: These assumptions will result in a change of the values u, v, and t_0, as discussed in sections 4.3 and 5. Since there is no immediate way to test the reliability of such predictions, we do not enter into more details.

Acknowledgement

The author wants to express his sincere gratitude to an anonymous referee whose very valuable comments and suggestions helped to improve considerably the exposition of this article.

REFERENCES

Bayley N. (1975): The Mathematical Theory of Infectious Diseases and Its Applications. 2nd ed. Charles Griffin & Co., London.

Bellman R., Cooke K.L. (1963): Differential-difference Equations. Academic Press, New York.

DEFRA (2004), Department for Environment, Food and Rural Affairs of the UK. Web site: http://www.defra.gov.uk/animalh/bse/statistics/bse/res_con.pdf

Diekmann O., Heesterbeek J.A.P. (1999): Mathematical epidemiology of infectious diseases. Model building, analysis and interpretation. Chichester, Wiley.

Grenfell B.T., Dobson A.P. (1995): Ecology of Infectious Diseases in Natural Populations. Univ. Press, Cambridge.

Isham V., Medley G. (1996): Models for infectious human diseases: their structure and relation to data. Publ. Newton Inst., Univ. Press, Cambridge.

Kermack W.O., McKendrick A.G. (1927): Contributions to the mathematical theory of epidemics I. Proceedings of the Royal Society of London, Series A, 115, 700–721.

Kermack W.O., McKendrick A.G. (1932): Contributions to the mathematical theory of epidemics II. Proceedings of the Royal Society of London, Series A, 138, 55–83.

Kermack W.O., McKendrick A.G. (1933): Contributions to the mathematical theory of epidemics III. Proceedings of the Royal Society of London, Series A, 141, 94–122.

Kühleitner M., Nowak W.G. (2003): Ein einfaches mathematisches Modell für die Epidemie-Dynamik bei BSE. Der Mathematische und Naturwissenschaftliche Unterricht, 56, 143–146.

Murray J.D. (1990): Mathematical Biology. Springer Verlag, Berlin.

Nowak W.G. (2004): Modellierung der britischen BSE-Epidemie mit DERIVE. Wissenschaftliche Nachrichten. Herausgegeben vom Österreichischen Bundes-ministerium für Bildung, Wissenschaft und Kultur, 122, 38–41.

Schuster R. (1995): Grundkurs Biomathematik. Teubner Verlag, Stuttgart.

Timischl W. (1995): Biomathematik. 2. Aufl., Springer Verlag, Wien.

Received: 04–12–09
Accepted after corrections: 05–05–02

Corresponding Author

O.Univ.Prof. Dr. Werner Georg Nowak, Institute of Mathematics, Department of Integrative Biology, BOKU Wien – University of Natural Resources and Applied Life Sciences, Vienna, Gregor Mendel-Straße 33, 1180 Vienna, Austria
E-mail: nowak@mail.boku.ac.at, http://www.boku.ac.at/math/