
ESTIMATING THE PARAMETERS OF AN AIDS SPREAD MODEL USING
OPTIMISATION SOFTWARE: RESULTS FOR TWO COUNTRIES COMPARED

Carole Roberts and Brian Dangerfield
Centre for O.R. and Applied Statistics
University of Salford
SALFORD M5 4WT
U.K.

A B S T R A C T

Few real life case study examples exist concerning optimisation in system dynamics models. This study reports an attempt to estimate relevant parameters of an AIDS spread model in order to check whether the chosen model structure can be separately parameterised and thereby explain the course of the epidemic for more than one country. The UK and USA are the two countries selected and the parameter values derived are reported for each. The values obtained are not inconsistent with emerging knowledge about the epidemic and the subsequent optimised projections reveal that the peak of the homosexual epidemic has been or is about to be reached in both countries.

Introduction

A system dynamics model of the spread of AIDS in a susceptible population of male homosexuals has been developed and refined over a number of years (Dangerfield and Roberts, 1989; Roberts and Dangerfield, 1990a, 1990b). This model handles complexities such as heterogeneity in sexual activity (different groups of susceptibles engaging in different frequencies of sexual activity) and a temporal variation in infectiousness over the long and variable incubation period.

The values used for the parameters in this model have been taken from the published literature reporting on cohort studies, statistical analysis of small sample data or clinical case histories. However, time series data on the number of AIDS cases is available for many countries now and this data is commonly disaggregated by risk group. For the homosexual group



specifically, data of this type could be used in order to offer estimates of model parameters which can then be compared with those estimates reported in the medical literature.

System dynamics and time series have had, historically, a rather uneasy relationship. There are those within the system dynamics community who take the view that past time series data offer no utility in the formulation or even testing of a system dynamics model. This view concerns the use of models for evaluating various policy choices for the future; the theme centres on just that and not on trying to explain the past. Others consider time series data can be employed during model testing but only in a judgemental sense, for comparison of turning points and phase relationships between the model and the real-world.

However, to employ time series data for the identification of parameter values is heresy in system dynamics rendering the methodology almost indistinguishable from econometrics. The rationale behind this study though is to allow a comparison between the parameter values derived and those obtained through direct clinical information and statistical surveys. As will be revealed below, it is possible that a set of parameters arising from a sub-optimal fit is the one which provides a scenario for the epidemic which is not obviously implausible. In addition, the results obtained for two countries' parameters, after fitting quarterly data spanning the same period in each, can be compared and similarities or differences highlighted. Finally, with a fitted (and plausible) model it is then possible to make an estimate of the number of HIV seropositives in each country.

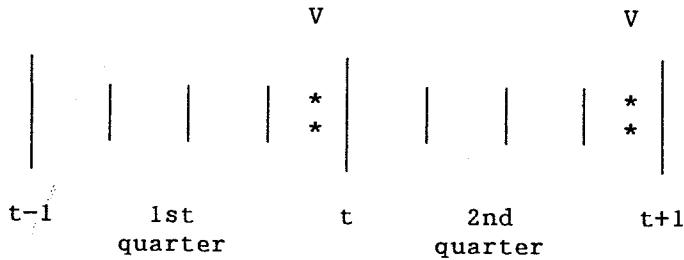
An immediate problem is deciding how to utilise the time series data in a model of the transmission of AIDS. Up until now such data has been restricted to use in statistical curve fitting models or those based on the so-called 'back projection' method of calculation, neither of which is able to reflect the behavioural and biological processes ongoing. However, within system dynamics, special parameter optimisation software, DYSMOD, (Luostarinen, 1982) can be employed to synthesise the information obtainable from time series data on the epidemic with the basic structure provided by a transmission model. (The DYSMOD software has now been ported to a PC environment at the University of Salford and retitled as DYSMOD/386.)

The ideas of parameter optimisation embodied in DYSMOD are due to Keloharju (1983). The principles and practice of this approach are described elsewhere (Keloharju and Wolstenholme, 1988; 1989). Essentially the software varies the values of selected parameters (constants or table functions) in a controlled fashion and within prescribed ranges such that an objective function (one of the equations in the model) is optimised.



Statistical and Computational Considerations

A number of issues have to be resolved when fitting model variables to reported data, if the data is reported on a more coarse time interval than that under which the model is running. In our case we are working with the reported incidence of AIDS on a quarterly basis, yet DT in our model is only 0.0625 years. If, say, 100 new cases of AIDS are reported in one quarter then there are four time slices to choose to allocate these cases. We assigned each such value to the final DT in each quarter (see figure 1) and compared it with the instantaneous incidence computed by the model for that DT.



V = values compared in this time step

Figure 1 Comparison of reported quarterly incidence data with model generated data

The equations below show the detail of the approach adopted. In addition to the table function for the reported statistics, equations required to compute the metric chosen for the fitting process are also listed.



```

A REPDAT.K=TABHL(REPDATT,TIME.K,1982.25-DT,1991.5-DT,0.25)
A TRIGGER.K=0+PULSE(1,STIME+0.25-DT,0.25)
A DEVC.K=OCLIP(0,REPDAT.K-EXPDAT.K,TIME.K,1991.5,
X 1982.25-DT-DT,TIME.K)
A CHI.K=RATIO(DEVC.K*DEVC.K,EXPDAT.K)
A LOWEXP.K=CLIP(0,CHI.K,EXPDAT.K,1)
L SUMCHISQ.K=SUMCHISQ.J+DT/DT*(CHI.J-LOWEXP.J)*TRIGGER.J
N SUMCHISQ=0
A OBJ.K=SUMCHISQ.K

```

Since we have adopted the principle of maximum likelihood as the basis of estimation, a chi-square statistic is computed in the objective function. For a population of size N with a predicted probability (p) of an individual contracting AIDS during a particular quarter, the number of new infections follows a binomial distribution with parameters (N,p) . For large N the maximum likelihood estimate of p reduces to a chi-square, which should be calculated from frequency data rather than cumulative data. Hence the need for consideration of the details of how this might be accomplished.

Strictly speaking, the model's analogue to the quarterly reported incidence of AIDS is the difference between cumulative AIDS cases at quarter t and quarter $t-1$, for all t , assuming $DT = 0.25$ years. However, with system dynamics software, a level is computed at the beginning of each DT and, therefore, even though it is easy to hold the 'old' value of the level at time $t-1$ through the next time step, the new value of the level at time t is not known until the start of the interval $(t,t+1)$. Thus, to carry out the differencing between reported and expected cases of AIDS would require that the reported statistic be pushed forward one reporting interval in order to marry up with the appropriate expected value derived as suggested. This is cumbersome and, furthermore, means that the first reporting period in the simulation is devoid of any "observed-expected" value whatsoever. The inadequacies of this approach led us to favour the one based on incidence data as explained above.

The TRIGGER variable is employed to ensure that the model accumulates in SUMCHISQ only the CHI values from the final instantaneous expected incidences in each quarter. This mechanism is similar to the PICK macro developed by Sterman (1984) to handle the same problem. The CHI values are obtained by computing DEVC (DEVIations of Cases) squaring this value and dividing by the expected number of cases. DEVC computes the



deviations between reported and expected values and restricts this to only those values arising at or between the limits of the data employed, namely 1982 (Q1) and 1991 (Q2). The OCLIP function used in the equation for DEVC effects the necessary restriction. The "-DT-DT" term is required because the selected time step for the comparison is one DT before the end of the quarter and the operational test between the fifth and sixth (and third and fourth) arguments of an OCLIP function is "greater than or equal to". We do not want the differencing of reported and expected cases to be carried out until time reaches one DT before the first quarter of 1982.

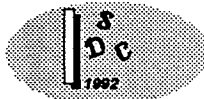
Computationally there can be problems with adoption of chi-square since the denominator in the formula (the expected number of cases) may be zero, or close to zero, at the beginning of the epidemic. The RATIO function is used to circumvent this problem. In addition a variable LOWEXP is employed to deal with the consequences of any low expected values of cases. Low is defined as less than 1.0. When this happens, LOWEXP exactly offsets the computed value of CHI and hence no change occurs when chi-square is being cumulated. The TRIGGER variable ensures this takes place only at every quarterly interval where real data exists and the "DT/DT" term permits the cumulation of the exact chi-square values and not 1/DT th of them. Finally, the equation for OBJ is necessary because the DYSMOD software requires that the objective function be an auxiliary.

Data and Results

The data employed in the study were quarterly from 1982 (Q1) to 1991 (Q2) inclusive. USA data were available monthly but were used in quarterly form. The time series in each case was derived by a program which analysed the entire list of case reports provided for each country to 30 June 1991. For the UK this was N= 4758 and for the USA N= 182834. The program extracted valid records for the homosexual risk group only, ignoring cases where an individual was classified into multiple risk groups. This produced a series of N= 3621 (UK) and N= 101868 (USA) homosexual AIDS case reports by quarter.

As part of our earlier work (Dangerfield and Roberts, 1989) an optimisation was conducted and an estimate of the HIV seropositive population of homosexuals was made for the UK at the end of 1987. This, however, suffered from a number of disadvantages which are overcome in the current study:

- (1) The model was fitted to cumulative, not incidence data.
- (2) Minimisation of sums of squares was adopted as the fitting metric, not chi-square.



(3) A time series of AIDS case reports, for the homosexual risk group only, was not isolated from the original data set; it was used as it stood.

(4) No attempt was made to assess changing sexual behaviour by homosexuals; this study does that.

(5) The incidence of AIDS diagnoses was equated with the incidence of report. In the current study a reporting delay is estimated on an ex-ante basis using the data on diagnosis and report dates for cases in each country. The reporting lags were fitted to a negative exponential distribution and the mean value of the best fit distribution was employed in the model as a constant in the SMOOTH function used to handle this feature. The values obtained were 0.6225 years for the UK and 0.581 years for the USA.

The parameters which were estimated via the fitting process on the full model numbered nine as follows:

ETPnAR Estimated Total Population At Risk, where $n=1, 2$ and 3 for each of the three strata of sexual activity which make up the heterogeneous model.

PIPPSn Probability of Infection Per Partner where $n=1, 2$ and 3 representing the three different stages of a U-shaped infectivity profile over the course of the incubation period. This was fixed at 10 years (1, 8 and 1 for each phase of infectiousness) from the outset. This approximates a third-order Erlang distribution of incubation time.

MNDPn Mean Number of Different Partners (per year) where $n=1, 2$ and 3 to equate with the number of partners taken by each of the three sexual activity groups. For $n=2$ and $n=3$ a table function (against time) was employed to capture the effect of changing behaviour. Three year increments were chosen for convenience, making five in all over the 15 year run of the model. (For each country the model was initialised half way through 1976 by one infected individual being introduced into the sexual activity class with the highest rate of partner change.)

The two data series are illustrated in figure 2 and, apart from the scale difference together with the slightly earlier take-off in the USA, the epidemics can be seen to possess a degree of similarity.



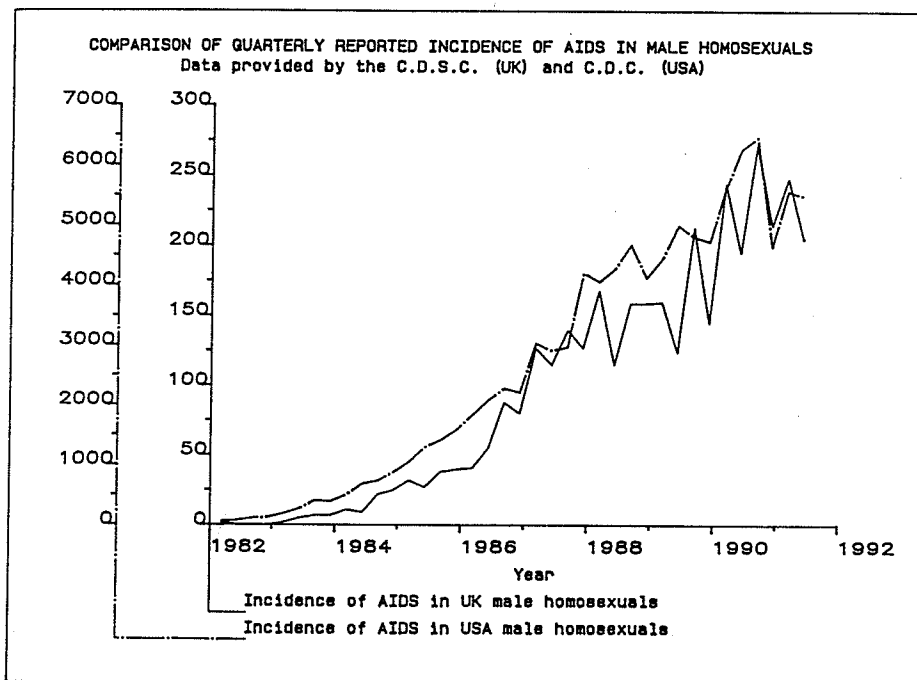


Figure 2 Comparison of the UK and USA quarterly data series on new AIDS cases in homosexuals

The results for the optimised parameters are given in the table below, together with the ranges imposed on them for the purposes of the search algorithm in DYSMOD and the value attained by the objective function.

UNITED KINGDOM

| <u>Range</u> | <u>Parameter</u> | <u>Optimised Value</u> |
|-----------------|------------------|------------------------|
| (100E3 - 500E3) | ETP1AR | 214133 |
| (300E3 - 800E3) | ETP2AR | 340650 |
| (20E3 - 150E3) | ETP3AR | 51148 |
| | | (Total= 605931) |
| (0.05 - 0.20) | PIPPS1 | 0.067 |
| (0.01 - 0.08) | PIPPS2 | 0.019 |
| (0.10 - 0.20) | PIPPS3 | 0.162 |
| (0.5 - 1.5) | MNDP1 | 0.53 |
| () | MNDP2(1) | 10.6 |
| () | MNDP2(2) | 8.0 |
| (2 - 15) | MNDP2(3) | 6.0 |
| () | MNDP2(4) | 9.0 |
| () | MNDP2(5) | 3.0 |
| () | MNDP2(6) | 2.2 |
| () | MNDP3(1) | 21.7 |
| () | MNDP3(2) | 36.7 |
| (2 - 40) | MNDP3(3) | 30.0 |
| () | MNDP3(4) | 18.9 |
| () | MNDP3(5) | 5.2 |
| () | MNDP3(6) | 2.6 |

Objective function = 158.9



UNITED STATES

| <u>Range</u> | <u>Parameter</u> | <u>Optimised Value</u> |
|---------------|------------------|------------------------|
| (2E6 - 15E6) | ETP1AR | 7185811 |
| (2E6 - 15E6) | ETP2AR | 2004290 |
| (400E3 - 2E6) | ETP3AR | 716627 |
| | | (Total= 9,906,728) |
| (0.05 - 0.20) | PIPPS1 | 0.0588 |
| (0.01 - 0.08) | PIPPS2 | 0.0199 |
| (0.10 - 0.20) | PIPPS3 | 0.1166 |
| (0.5 - 1.5) | MNDP1 | 0.58 |
| () | MNDP2(1) | 2.9 |
| () | MNDP2(2) | 7.3 |
| (2 - 25) | MNDP2(3) | 20.0 |
| () | MNDP2(4) | 10.7 |
| () | MNDP2(5) | 4.2 |
| () | MNDP2(6) | 2.0 |
| () | MNDP3(1) | 49.1 |
| () | MNDP3(2) | 46.6 |
| (2 - 80) | MNDP3(3) | 29.4 |
| () | MNDP3(4) | 11.3 |
| () | MNDP3(5) | 2.3 |
| () | MNDP3(6) | 2.0 |

Objective function = 914.6

The results for both countries support the view of there being a dip in infectiousness during the course of the incubation period, with a strong rise towards the time of onset of clinical AIDS. The numbers of different partners taken by the most sexually active and moderately sexually active strata (numbers 3 and 2 respectively) have declined considerably from the mid-1980's which again supports the view of widespread adoption of reduced frequency of partner change by homosexuals cited in many sexual surveys. Although the average number of different partners actually climbed for the moderate sexual activity stratum in the USA, this may well have occurred given that the recognition of the seriousness of the situation (and the health promotion campaign) was not in place until the mid-1980's. The ranges for the number of different partners were made wider for the USA because of the a priori belief, derived from surveys, that a greater degree of heterogeneity in sexual activity exists there.



Plots of the quarterly reported incidence of AIDS in homosexuals, together with the expected incidence derived from the model, are shown below for the two countries. Inspection of the printed results reveals that by the end of the second quarter of 1991, 14322 homosexuals had become seropositive since the start of the epidemic in the UK with 9976 of these still not progressed to the clinical definition of AIDS. For the USA the figures are 338492 and 223081 respectively.

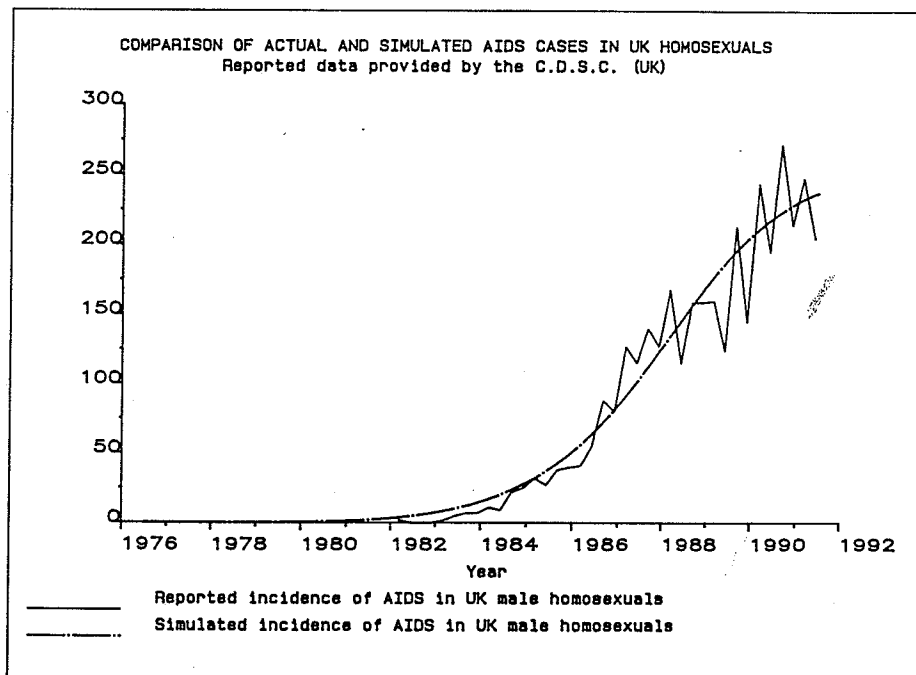
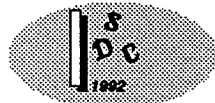


Figure 3 Quarterly reported incidence of AIDS in UK homosexuals with fitted model trajectory

For the United Kingdom analysis, the fitted parameters reported above were not those associated with the lowest chi-square value. Possibly because of the greater variability in the quarterly incidence data for the UK, a run which produced a value for chi-square of 138 also produced a table for the mean number of different partners that was much more varied and, in particular, showed a higher figure as the final (sixth) value. Although this gave a 'best fit', the resultant behaviour of the epidemic following directly after the period of the fit was just not plausible; not surprisingly the incidence of new AIDS cases exhibited a second take-off. While this could actually happen, especially if the homosexual risk group become less circumspect about their rate of partner change, it does seem unlikely and the result serves as a warning to those who prefer to determine all parameter values in a model by time-series data alone.



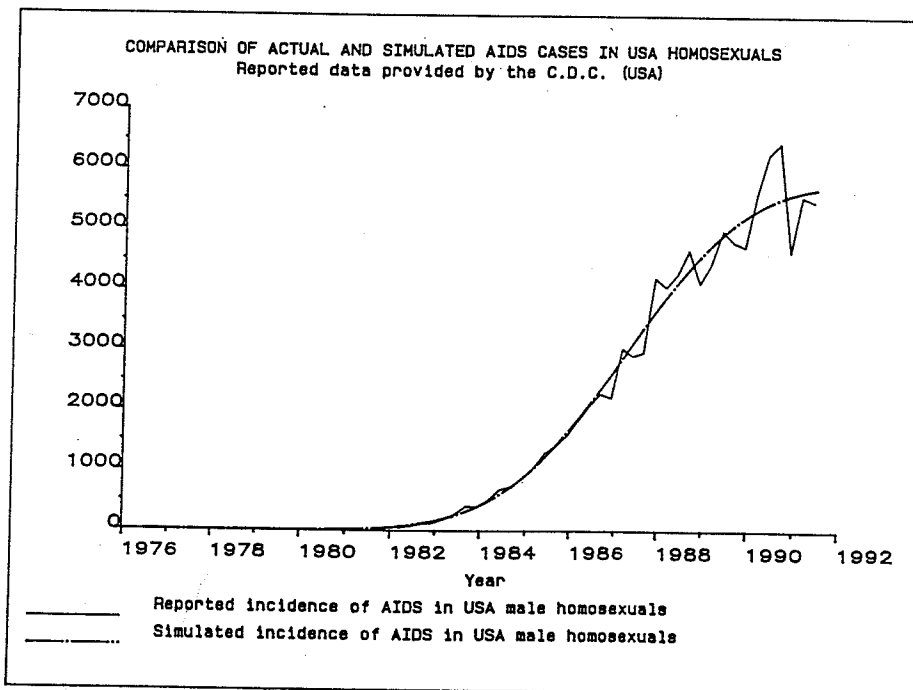


Figure 4 Quarterly reported incidence of AIDS in USA homosexuals with fitted model trajectory

Finally, a co-plot is given below of the two countries' homosexual AIDS epidemics projected over 50 years by the best fit models in each case. The overall character of the epidemic is remarkably similar in each country. It reveals that peak incidence has almost been reached or indeed has been passed in the case of the United States which exhibits the leading curve as it does the reported data. However, it cannot be stressed enough that this graph is not a forecast, but merely a projection based on the assumptions incorporated into our model. Further release of data sets might cause the parameter specification to change but there is strong evidence from this study that we are close to the peak incidence of AIDS in homosexuals.

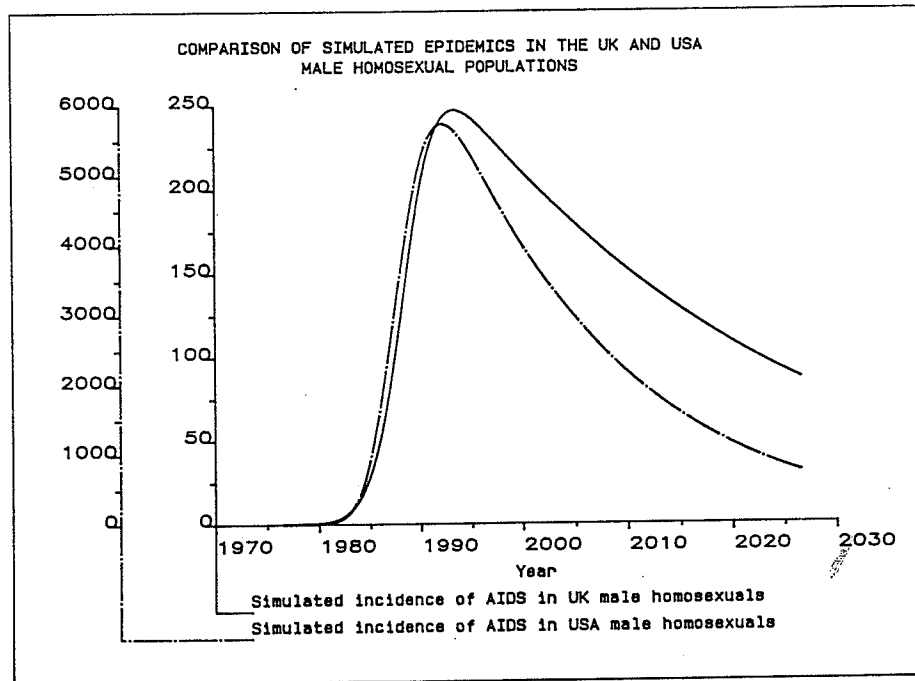


Figure 5 A projection of AIDS incidence in the UK and USA derived from a model parameterised by optimising the fit to reported data

Conclusion

This study has demonstrated that optimisation of system dynamics models using special purpose software such as DYSMOD is a powerful research methodology capable of offering additional insight into complex systems. Few applications of this methodology exist other than on textbook examples. While system dynamics models should never be specified entirely by such a method, the work reported above stands testimony to its use as a viable adjunct to conventional system dynamics modelling.

REFERENCES

- Dangerfield B C. and C.A. Roberts. 1989. A Role for System Dynamics in Modelling the Spread of AIDS. *Transactions of the Institute of Measurement and Control* 11: 187-195.
- Keloharju R. 1983. *Relativity Dynamics*. Helsinki School of Economics, Helsinki.



-
- Keloharju R. and E.F. Wolstenholme. 1988. The Basic Concepts of System Dynamics Optimisation. *Systems Practice* 1: 65-86.
- Keloharju R. and E.F. Wolstenholme. 1989. A Case Study in System Dynamics Optimisation. *Journal of the Operational Research Society* 40: 221-230.
- Luostarinen A. 1982. *DYSMOD User's Manual*. Helsinki School of Economics, Helsinki.
- Roberts C. A. and B.C. Dangerfield. 1990a. Modelling the Epidemiological Consequences of HIV Infection and AIDS: a contribution from Operational Research. *Journal of the Operational Research Society* 41: 273-289.
- Roberts C. A. and B.C. Dangerfield. 1990b. A System Dynamics Framework for Understanding the Epidemiology of HIV/AIDS. In *O.R. Work in HIV/AIDS*. Operational Research Society, Birmingham.
- Sterman J. 1984. Appropriate Summary Statistics for Evaluating the Historical Fit of System Dynamics Models. *Dynamica* 10: 51-66.

