Applied Mathematics and Computation 219 (2013) 9700–9708

Contents lists available at SciVerse ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

An integration-based method for estimating parameters in a system of differential equations

Andrew B. Holder *, Marianito R. Rodrigo

School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, New South Wales, Australia

ARTICLE INFO

Keywords:
Parameter estimation
Systems of ordinary differential equations
Biological models

ABSTRACT

The application of ordinary differential equations to modelling the physical world is extensive and widely studied in many fields including physics, engineering and bioinformatics. Using these models to predict the behaviour of important state variables given particular parameter values has been extensively studied. On the other hand the inverse problem of predicting parameter values that will fit a solution of a differential equation to observed data has traditionally only been considered by using a few methods, many of which approach the problem via a least squares fit method. These methods either can only be applied when the differential equation being studied has a closed form solution or can become very computationally intensive when applying it to a system that can only be solved numerically and hence require optimisation algorithms. We propose an integration-based method that transforms an ordinary differential equation to an algebraic system of equations for which we solve for the unknown parameters in our equation. The method is computationally unintensive, can be extended to systems of differential equations and the number of parameters that can be estimated is not restricted. We demonstrate the method by simulating data, with and without noise, from a number of biological models described by ordinary differential equations and then estimate the parameters via the proposed technique.

1. Introduction

It has been commonplace for many years to use ordinary differential equations (ODEs) to model the physical world. The practise of using these models to predict and determine the behaviour of certain state variables is prolific in mathematical and scientific literature. The inverse problem of predicting the parameter values that appear in an ODE based on observed data has been studied considerably less and traditionally only a few methods have been used, many of which approach the problem via a least squares fit method (see for example [1,3,12,14] and the references therein). The methods are based on minimising the distance between the observed data and values predicted by the model at discrete points in time. If the model can be solved analytically, then the exact solution is used to generate a distance function in which the derivatives are taken with respect to the parameters and the local minima are then found. Since many of these equations do not have just one local minimum, an algorithm must be used to find the global minimum [11]. Alternatively when the model can only be solved numerically, a trial set of parameters is used to generate a solution in which the least squares distance is calculated between the solution and the data. A trajectory is then found by altering the parameters and resolving the system numerically and comparing the least squares distances. This method can be very inefficient and is not guaranteed to return the global minimum, hence much of the focus is on creating stable and efficient algorithms to optimise this approach [11,23].

* Corresponding author.
E-mail addresses: abh920@uowmail.edu.au (A.B. Holder), marianito_rodrigo@uow.edu.au (M.R. Rodrigo).

0096-3003/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2013.03.052

More recently some statistically based methods have been considered by authors including [2,13,22] in which a hierarchical Bayesian approach is used to estimate dynamic parameters in HIV models, as well as a local kernel smoothing-based method to estimate constant parameters considered in [15]. Estimation of dynamic parameters was also considered in [24] using a technique known as principal differential analysis. This technique involves taking discrete data obtained from several sources and fitting a spline to each data set, then the approximations are substituted into an ODE and the dynamic parameters are then found by a least squares procedure. A spline-based smoothing approach has also been considered in [23] to estimate constant parameters. An approach based on integrator theory has been considered in [18] in which a condition is placed on the integral of a particular function within the ODE to ensure convergence to the true parameter values.

This article will develop a new method for estimating parameters in a system of ODEs that solve for some state variable. The method considers an ODE with m unknown parameters and then multiplies the system by a given weight function that contains a controllable parameter β, that we will call an *equation-generating parameter*. The system is then integrated over a finite interval, in what can be thought of as being analogous to a finite Laplace transform, to remove the derivatives contained within the equation. This integration over an interval can also be thought of as taking a weighted average of each term in the system. We will then substitute m values of our controllable parameter β, for example $\beta_1, \ldots, \beta_m$, into our transformed equation to then give us a set of m distinct algebraic equations in the parameters that will have terms which can be approximated by numerical integration using observed data for the values of the state variable. In Section 2 we will outline this procedure in more detail. Section 3 will contain applications of this method to some well-studied problems in which simulated data will be used, with and without noise, to demonstrate the viability of the proposed method. We will make some concluding remarks and discuss potential further analysis for this method in Section 4.

2. Method

To simplify the description of the method we consider an ODE of the form

$$\frac{dx}{dt} = f(x,t;\Theta), \quad x(0) = x_0, \tag{1}$$

where $x : \mathbb{R} \to \mathbb{R}, f : \mathbb{R}^{m+2} \to \mathbb{R}, \Theta \in \mathbb{R}^m, x_0 \in \mathbb{R}$ and we wish to estimate Θ. Assume that the dependent variable x is observed over a time interval I and consider a weight function $\varphi : I \times \mathbb{R} \to \mathbb{R}$ such that φ and $\dot{\varphi}$ are integrable, where $\dot{(\cdot)}$ represents the derivative with respect to time.

Multiplying (1) by $\varphi(t;\beta)$ and integrating over I, we obtain an equation of the form

$$\varphi(t;\beta)x(t)|_I - \int_I \dot{\varphi}(t;\beta)x(t)\,dt = \int_I \varphi(t;\beta)f(x(t),t;\Theta)\,dt. \tag{2}$$

We define $k_\beta := \varphi(t;\beta)x(t)|_I - \int_I \dot{\varphi}(t;\beta)x(t)\,dt$ and $F_\beta(\Theta) := \int_I \varphi(t;\beta)f(x(t),t;\Theta)\,dt$ for some constant k_β and function F_β, so that (2) can be represented as $k_\beta = F_\beta(\Theta)$. Since we have m unknown parameters, we then substitute m distinct values of β, say $\beta_1, \ldots, \beta_m$, into (2) to generate a system of equations given by

$$\begin{bmatrix} k_{\beta_1} \\ \vdots \\ k_{\beta_m} \end{bmatrix} = \begin{bmatrix} F_{\beta_1}(\Theta) \\ \vdots \\ F_{\beta_m}(\Theta) \end{bmatrix}. \tag{3}$$

Thus we have reformulated the problem of parameter estimation as a problem of finding a root of an algebraic system of nonlinear equations.

A special case of (1) is when f is of the form

$$f(x,t;\Theta) = \theta_1 g_1(x,t) + \cdots + \theta_m g_m(x,t) + g_0(x,t), \tag{4}$$

where $g_i : \mathbb{R}^2 \to \mathbb{R}$ and $\Theta = (\theta_1, \ldots, \theta_m)$. This case is where the parameters appear linearly in the equation so that (3) becomes a linear system of algebraic equations. This imposes a restriction on the weight function φ such that the resulting coefficient matrix of Θ in (3) is invertible. Assuming relevant data can be observed, this method can be applied to systems of ODEs by applying it to each equation with unknown parameters that appears within the system.

3. Applications

In order to demonstrate the viability of our method as an alternative technique, we will apply it to some well-known models, for some of which alternative parameter estimation techniques have been applied, for example [15].

3.1. Bernoulli's smallpox model

In his seminal paper [6], Bernoulli proposed a model to determine the prevalence of immune and susceptible individuals, of a certain age, to smallpox and as a result, calculate the gain in life expectancy if this infectious disease were eliminated at

birth as a potential cause of death. A review of this paper has been conducted recently in [10]. Importantly, in Bernoulli's paper an estimate was made about certain parameters in his model based on observed data. It is unclear as to how Bernoulli estimated these values, however this gives us a model in which we can test the effectiveness of our method.

The model proposed by Bernoulli for determining the proportion of susceptible individuals x of a population at age a, was given by

$$\frac{dx}{da} = -\lambda(a)x(a)[1 - c(a)x(a)], \tag{5}$$

where $\lambda(a)$ is the force of infection and determines the rate at which susceptibles are infected and $c(a)$ is the case fatality rate, which determines the proportion of people that die as a result of the infection. This model has an initial condition of $x(0) = 1$ as it is assumed that all newborns will not be immune to the disease.

According to [7,10], when the assumption was made that the parameters in this model were constants, say $\lambda(a) = \lambda$ and $c(a) = c$, Bernoulli was able to calculate parameter values based on data collected by Edmond Halley, which can be found in [7]. Bernoulli estimated that the force of infection was $\lambda = 0.125$ and the case fatality rate was $c = 0.125$.

Under the assumption that the parameters in (5) are constant, we have

$$\frac{dx}{da} = -\lambda x(a) + \lambda c x(a)^2, \tag{6}$$

and by letting $\alpha = \lambda c$ and identifying t with a, we have an equation in the form of (4) and hence we can apply our method to (6). The data in [7] are of observations made of people from ages 0 to 24, therefore we will numerically integrate over the interval $I = [0, 24]$. We choose the weight function to be $\varphi(a;\beta) = e^{-\beta a}$, where a is in years. We chose this type of function since, as mentioned in the introduction, we can think of our method as being analogous to applying a finite Laplace transform to an ODE. However it should be noted that we are not limited to this choice of weight function and theoretically the method will work with any choice of φ as long as system (3) is solvable.

The method yields, with $\beta = \beta_1, \beta_2$, the system

$$\begin{bmatrix} e^{-\beta_1 a}x(a)\big|_0^{24} + \beta_1 \int_0^{24} e^{-\beta_1 a}x(a)\,da \\ e^{-\beta_2 a}x(a)\big|_0^{24} + \beta_2 \int_0^{24} e^{-\beta_2 a}x(a)\,da \end{bmatrix} = \begin{bmatrix} -\int_0^{24} e^{-\beta_1 a}x(a)\,da & \int_0^{24} e^{-\beta_1 a}x(a)^2\,da \\ -\int_0^{24} e^{-\beta_2 a}x(a)\,da & \int_0^{24} e^{-\beta_2 a}x(a)^2\,da \end{bmatrix} \begin{bmatrix} \lambda \\ \alpha \end{bmatrix}. \tag{7}$$

The analysis for the appropriate choice of β_1 and β_2 still needs to be conducted, so values were chosen that do not alter the data set by a large amount as to create a potential bias towards particular sections of the data and not to create any numerical errors as a result of having one equation a large factor greater than the other. Hence we let $\beta_1 = 0$ and $\beta_2 = 0.2$ since the choice of $\beta_1 = 0$ will leave the data unchanged and $\beta_2 = 0.2$ is relatively small. Using MATLAB, we calculate the integrals given in (7) numerically by interpolating the data by a cubic spline method, then integrate and thus calculate the predicted values for λ and α, and as a result c. With these values we obtain that $\lambda = 0.1250$ and $c = 0.1263$ to four significant figures. Hence we have obtained values close to those predicted by Bernoulli in 1766. Note that (6) can be solved analytically, so as a result the least squares method can be used to calculate the parameter values. When we used this method, the parameter values we found are $\lambda = 0.1250$ and $c = 0.1257$ and we can see that the values obtained via the analytical least squares method and the proposed method are very close.

As can been seen in Fig. 1, these parameter values represent an approximation that generates a solution with a strong fit to the data.

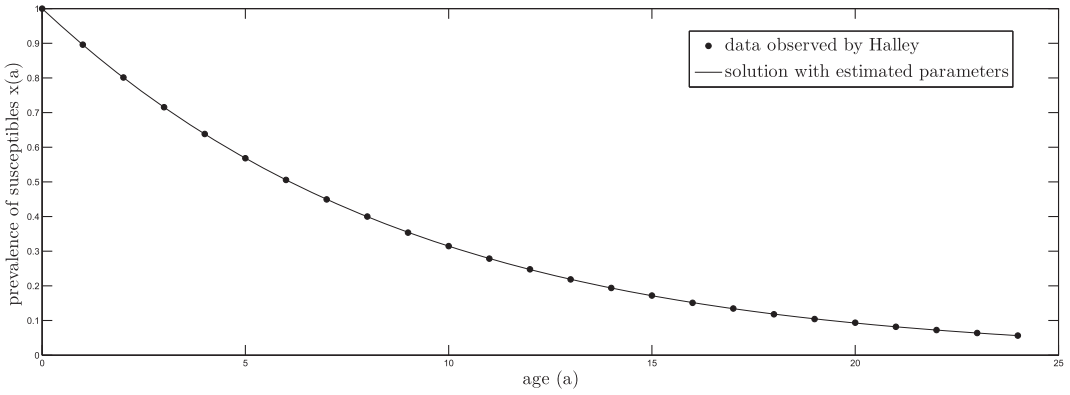

Fig. 1. Comparison of data observed by Halley with solution to Bernoulli's model with estimated parameters $\lambda = 0.1250$ and $c = 0.1263$ for the prevalence of susceptibles from ages 0 to 24.

3.2. HIV dynamic model

Modelling the dynamics of HIV is an important and significant area of study. Models initially considered described the dynamics of the HIV virus and immune cell response. A review of these models has been considered in [17,20,26]. A HIV dynamic model [15,29] in which alternative parameter estimation techniques were applied, is given by

$$\frac{dT_u}{dt} = \lambda - \rho T_u - \eta(t) T_u V, \tag{8}$$

$$\frac{dT_i}{dt} = \eta(t) T_u V - \delta T_i, \tag{9}$$

$$\frac{dV}{dt} = N\delta T_i - cV, \tag{10}$$

where T_u is the concentration of uninfected CD4+ T cells, T_i is the concentration of infected CD4+ T cells, V is the concentration of plasma virus (viral load), λ is the rate at which new CD4+ T cells are generated, ρ is the death rate of uninfected CD4+ T cells, $\eta(t)$ is the infection rate of CD4+ T cells, δ is the death rate of infected cells, c is the clearance rate of free virions and N is the number of virions produced from each infected cell. We note the similarity of this model to the Beretta–Kuang model of viral infection in bacteria [5] and its generalisation [25]. These represent further models for which our parameter estimation technique can be applied.

As is stated in [15], only the total CD4+ T-cell count $T(t) = T_u(t) + T_i(t)$ and the plasma viral load $V(t)$ can be measured. Therefore Eqs. (8)–(10) are transformed such that we have an equation for $T(t)$ and $V(t)$, given by

$$V'(t) + cV(t) = \alpha_0 T'(t) + \alpha_1 T(t) + \alpha_2, \tag{11}$$

where

$$\alpha_0 = \frac{N\delta}{\rho - \delta}, \quad \alpha_1 = \rho\alpha_0, \quad \alpha_2 = -\lambda\alpha_0. \tag{12}$$

See [15] for the full working.

We now have (11) in which the parameters appear linearly and hence we can apply our parameter estimation technique. As is done in [15], we generate data by solving the system of Eqs. (8)–(10) with a particular set of parameters and then apply the technique to (11) to see if we obtain those parameters back, both with and without noise added to the generated data. We also make the same assumption made in [15] that values of δ and c can be obtained from the literature [19,21,27,28]. We generate data by solving (8)–(10), with initial values of $(T_u(0), T_i(0), V(0)) = (600, 30, 10^5)$, parameter values of

$$(\lambda, \rho, N, \delta, c) = (36, 0.108, 10^3, 0.5, 3), \tag{13}$$

and dynamic parameter given by $\eta(t) = 9 \times 10^{-5}[0.9 - \cos(\pi 10^{-3} t)]$, over the interval of time $[0, 20]$ (cf. [15]). Once this data is generated, we add T_u and T_i together to recreate the conditions in a clinical study and then apply the proposed technique to (11). Once again we will let $\varphi(t;\beta) = e^{-\beta t}$ and as we are finding the value of three different parameters, we will need to use three distinct values of β which will be $\beta = 0, 0.2, 0.4$. These values were chosen as they are small and will not skew the observed data by large amounts.

3.2.1. Absence of noise

We solve (8)–(10) with time-steps $0.1, 0.2$ and 0.4, such that we have $201, 101$ and 51 observations respectively over the time period $[0, 20]$. In the absence of noise, when we apply our technique, we obtain parameter values of $(\lambda_{0.1}, \rho_{0.1}, N_{0.1}) = (36.00, 0.1080, 1000)$, $(\lambda_{0.2}, \rho_{0.2}, N_{0.2}) = (35.94, 0.1077, 1000)$ and $(\lambda_{0.4}, \rho_{0.4}, N_{0.4}) = (35.00, 0.1029, 996.8)$ respectively. By comparing these values with (13) we can see that without noise there are small errors in the cases with fewer observations which can be accounted for by a less accurate approximation for our integrals as a result of using larger time-steps.

3.2.2. In the presence of noise

To simulate the effect of applying this technique to real data, we will follow the method used in [15] in which noise is added to the numerically generated solution. The noise will be normally distributed with different variances. We will then run this simulation several times, recording the average and the standard deviation of the parameter values given back. We can then compare our results with those obtained in [15]. We will therefore have data at various points in time t_i of the form

$$\widehat{T}_i = T(t_i) + \varepsilon_{T,i}, \widehat{V}_i = V(t_i) + \varepsilon_{V,i}, \tag{14}$$

where $\varepsilon_{T,i}$ and $\varepsilon_{V,i}$ are normally distributed with mean zero and variances of $\sigma_T^2 = 20, 30, 40$ and $\sigma_V^2 = 100, 150, 200$ respectively. Applying this with data generated by using time-steps of $0.1, 0.2$ and 0.4, we run the simulations 500 times and then calculate the averages and standard deviations of the parameter estimates produced. The results are summarised in Tables 1–3 for the time-steps $0.1, 0.2$ and 0.4 respectively. We can see that the mean of the estimates obtained in the simulations are very close to the actual parameter values (13) and that the standard deviations are relatively small. In the presence of a

Table 1
Means and standard deviations of estimated parameters values from 500 simulations of the HIV model with noise for time-step 0.1.

		Mean			Standard Deviation		
σ_T^2	σ_V^2	λ	ρ	N	λ	ρ	N
20	100	36.01	0.1081	1001.0	0.864	0.00492	16.79
	150	35.94	0.1076	999.4	0.814	0.00474	16.49
	200	36.00	0.1082	1001.8	0.853	0.00476	16.80
30	100	35.95	0.1078	1000.6	1.075	0.00593	20.07
	150	35.97	0.1078	1000.6	1.047	0.00619	21.61
	200	36.05	0.1081	999.8	1.021	0.00588	19.78
40	100	35.95	0.1077	1000.2	1.293	0.00719	24.11
	150	36.01	0.1082	1001.9	1.224	0.00699	23.36
	200	36.04	0.1085	1002.9	1.175	0.00687	24.61

Table 2
Means and standard deviations of estimated parameters values from 500 simulations of the HIV model with noise for time-step 0.2.

		Mean			Standard Deviation		
σ_T^2	σ_V^2	λ	ρ	N	λ	ρ	N
20	100	35.87	0.1072	999.2	1.171	0.00656	18.77
	150	35.89	0.1074	999.8	1.147	0.00637	17.56
	200	35.99	0.1080	1000.9	1.225	0.00670	17.78
30	100	35.97	0.1079	1000.9	1.553	0.00845	22.31
	150	35.90	0.1074	999.7	1.489	0.00827	22.70
	200	35.94	0.1078	1001.4	1.425	0.00783	23.10
40	100	35.96	0.1077	1001.2	1.720	0.00931	26.47
	150	35.97	0.1080	1001.8	1.765	0.00966	26.78
	200	35.88	0.1072	999.4	1.756	0.00963	26.70

Table 3
Means and standard deviations of estimated parameters values from 500 simulations of the HIV model with noise for time-step 0.4.

		Mean			Standard Deviation		
σ_T^2	σ_V^2	λ	ρ	N	λ	ρ	N
20	100	35.02	0.1031	997.6	1.735	0.00929	20.92
	150	34.94	0.1027	997.1	1.724	0.00917	20.21
	200	35.08	0.1034	998.1	1.675	0.00907	20.15
30	100	34.90	0.1027	998.9	2.026	0.01087	24.77
	150	35.03	0.1033	998.8	2.114	0.01142	24.21
	200	35.00	0.1030	998.1	2.266	0.01197	26.03
40	100	34.95	0.1023	995.7	2.437	0.01312	29.28
	150	35.08	0.1033	997.9	2.441	0.01314	29.67
	200	35.01	0.1029	997.6	2.300	0.01273	29.88

small amount of noise we obtain very accurate approximations for the parameter values for the model (8)–(10). When our results are compared to those obtained in [15] by applying the PsLS and SIMEX methods to this model, we can see that this method has obtained more accurate mean approximations for all three parameters and much lower standard deviations for the estimates of λ and ρ and similar standard deviations for the estimates of N. For example, the mean and standard deviation of estimates obtained for the time-step 0.2 and $\sigma_T = 40$ and $\sigma_V = 200$ in [15] were $(\lambda, \rho, N) = (31.1[2.26], 0.090[.22], 935.6[41.9])$ and $(\lambda, \rho, N) = (30.8[9.44], 0.114[.72], 938.1[66.1])$ for the PsLS and SIMEX methods respectively. Here, the values in square brackets represent the standard deviations. When we compare this to the relevant data in Table 2, we can see a much higher level of accuracy has been obtained by the proposed method. If we look at Table 3, in which we have used time-steps of 0.4, we can see that we are obtaining results that are as good if not better than those obtained by the methods used in [15] when the time-step is 0.1.

3.3. Tumour cell and chemotherapy drug interaction model

The study of the growth and treatment of solid tumours is a very important area of analysis as in the developed world about one in five will die from cancer [8]. Many authors have looked at the growth of tumours by modelling them by ODEs [4,9]. A simple model [9] that looks at the dynamics of solid tumour cell and chemotherapy drug interaction is given by

$$\frac{dN}{dt} = kN\left(1 - \frac{N}{\theta}\right) - \mu AN, \tag{15}$$

$$\frac{dA}{dt} = a(t) - \lambda A - \gamma AN, \tag{16}$$

with

$$N(0) = N_0 \quad \text{and} \quad A(0) = A_0. \tag{17}$$

In this, $N(t)$ is the number of tumour cells at time t, $A(t)$ is the average concentration of chemotherapeutic drug within the tumour at time t, k is the cell proliferation rate, θ represents the carrying capacity of the cell population, μ denotes the rate at which the drug kills the tumour cells, λ represents the drug's decay rate, γ is the rate at which the drug becomes ineffective as a result of a cell kill and $a(t)$ is the drug delivery rate to the tumour. The dynamic parameter $a(t)$ can be used to represent various drug infusion protocols. However for simplicity we will let it represent continuous drug infusion, that is $a(t) = a_\infty$ for all t, where a_∞ is a positive constant.

3.3.1. Absence of noise

We can see that (15)–(17) is a system that has functions of the parameters that appear linearly in this equation. Hence we can apply our method to this system to estimate the parameters. Since both the number N of tumour cells and the average concentration A of chemotherapeutic drug can be estimated [16,30], we do not need to alter the state variables as was done for the HIV model (8)–(10). To test the effectiveness of our estimation technique on the model (15)–(17), we simulate data from a numerical solution to the system of ODEs with a particular set of parameter values and we assume that a_∞ is known. To estimate the parameters k, θ and μ, we use our technique on (15) and then to estimate λ and γ, we need to apply the technique to (16). We let

$$(k, \theta, \mu, a_\infty, \lambda, \gamma) = (0.8, 1, 1.1, 0.8, 1.2, 0.6), \tag{18}$$

and use initial values of $N(0) = 1$ and $A(0) = 0.1$. We generate sets of values for N and A over a time period of $[0, 10]$ using time-steps of $0.1, 0.2$ and 0.4 to generate sample sizes of $101, 51$ and 26 respectively. For both equations we let $\varphi(t;\beta) = e^{-\beta t}$ and take $\beta = 0, 0.4, 0.8$ for (15) and $\beta = 0, 0.4$ for (16). These values were chosen as they are small and will not skew the observed data by large amounts. Using the generated data, we obtain estimates for our parameters of $(k_{0.1}, \theta_{0.1}, \mu_{0.1}, \lambda_{0.1}, \gamma_{0.1}) = (0.800, 1.00, 1.10, 1.20, 0.600)$, $(k_{0.2}, \theta_{0.2}, \mu_{0.2}, \lambda_{0.2}, \gamma_{0.2}) = (0.800, 1.00, 1.10, 1.20, 0.600)$ and $(k_{0.4}, \theta_{0.4}, \mu_{0.4}, \lambda_{0.4}, \gamma_{0.4}) = (0.787, 0.993, 1.08, 1.20, 0.602)$ for the data generated with time-steps $0.1, 0.2$ and 0.4 respectively. When we compare this with (18), we can see that the accuracy of the estimation is reduced in the cases with fewer data values as a result of less accurate approximations of the integrals.

3.3.2. In the presence of noise

To test the effectiveness of this model in a clinical situation we will add noise to the data sets generated by solving the system numerically and then estimate our parameters as was done for the HIV model. In the case of the HIV model, noise was added that was normally distributed so that we could compare our results with that of those obtained in [15]. Noise that is normally distributed has the capability of producing large outliers from the data set that would not necessarily be considered for estimation of the parameters. We therefore add noise that is uniformly distributed, say $U(-\sigma_N, \sigma_N)$ and $U(-\sigma_A, \sigma_A)$ for the sets of data produced for N and A respectively. We will use the sets of data generated previously by using time-steps of $0.1, 0.2$ and 0.4 and run 500 simulations with noise added with σ_N and σ_A taking each of the values 0.02 and 0.04, and take the average and standard deviation of the estimates produced for each parameter. The results of this are summarised in Tables 4–6.

We define the error measures mean absolute error (MAE) and mean absolute value (MAV) as

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |x(t_i) - y_i|, \tag{19}$$

Table 4
Means and standard deviations of estimated parameter values from 500 simulations of the tumour cell and chemotherapy drug interaction model with noise for time-step 0.1.

		Mean					Standard Deviation				
σ_N	σ_A	k	θ	μ	λ	γ	k	θ	μ	λ	γ
0.02	0.02	0.801	0.992	1.10	1.20	0.600	0.164	0.072	0.213	0.0144	0.0335
	0.04	0.810	0.995	1.11	1.20	0.601	0.176	0.078	0.232	0.0319	0.0743
0.04	0.02	0.791	0.959	1.09	1.20	0.600	0.310	0.156	0.403	0.0162	0.0381
	0.04	0.816	0.965	1.12	1.20	0.608	0.320	0.163	0.415	0.0307	0.0727

Table 5
Means and standard deviations of estimated parameter values from 500 simulations of the tumour cell and chemotherapy drug interaction model with noise for time-step 0.2.

		Mean					Standard Deviation				
σ_N	σ_A	k	θ	μ	λ	γ	k	θ	μ	λ	γ
0.02	0.02	0.803	0.991	1.10	1.20	0.602	0.182	0.082	0.241	0.0212	0.0515
	0.04	0.817	0.993	1.12	1.20	0.597	0.229	0.103	0.308	0.0426	0.1031
0.04	0.02	0.796	0.943	1.10	1.20	0.600	0.360	0.226	0.477	0.0209	0.0506
	0.04	0.826	0.959	1.14	1.20	0.603	0.375	0.205	0.497	0.0394	0.0947

Table 6
Means and standard deviations of estimated parameter values from 500 simulations of the tumour cell and chemotherapy drug interaction model with noise for time-step 0.4.

		Mean					Standard Deviation				
σ_N	σ_A	k	θ	μ	λ	γ	k	θ	μ	λ	γ
0.02	0.02	0.797	0.981	1.10	1.20	0.596	0.219	0.110	0.295	0.0288	0.0710
	0.04	0.817	0.984	1.12	1.20	0.601	0.276	0.131	0.380	0.0540	0.1351
0.04	0.02	0.815	0.898	1.12	1.20	0.601	0.401	1.127	0.543	0.0296	0.0740
	0.04	0.806	0.917	1.11	1.20	0.610	0.459	0.301	0.622	0.0557	0.1354

and

$$\mathrm{MAV} = \frac{1}{n}\sum_{i=1}^{n}|a_i|, \tag{20}$$

where $x(t_i)$ is the predicted value at t_i, y_i is the observed value at t_i and a_i is the ith element of an arbitrary set of n values. For each simulation we also use the parameters estimated to solve the system of equations given by (15)–(17) and then calculate the MAE between the solution generated with the estimated parameters and the simulated data with noise and we take the average and standard deviation of the MAEs obtained from the 500 simulations. We also calculate the MAV of the noise added to the original solution in each simulation and then take the average and standard deviation of the MAV from the 500 simulations as a basis for comparison. Since the noise is uniformly distributed the expected MAV of the noise generated for N and A is $\sigma_N/2$ and $\sigma_A/2$ respectively. This data is summarised in Tables 7–9.

We can see from the data contained in Tables 7–9 that the averages produced are close to the actual parameter values (18), with the error increasing as the number of observations decreases. However for some of the parameters, the respective standard deviations are large relative to the parameter value, which suggest there is significant probability of obtaining values that can have significant relative errors to the actual parameter values. However if we look at the MAEs of the solutions generated by using the estimated parameters, we can see that these values are close to the MAVs of the noise and this sug-

Table 7
Mean and Standard Deviation of MAE of simulated data and solution obtained with estimated parameters and MAV of noise from simulated data for time-step 0.1.

		MAE		MAV		SD of MAE		SD of MAV	
σ_N	σ_A	N	A	ε_N	ε_A	N	A	ε_N	ε_A
0.02	0.02	0.0102	0.0099	0.0099	0.0099	0.00073	0.00058	0.00057	0.00057
	0.04	0.0102	0.0197	0.0099	0.0199	0.00080	0.00117	0.00058	0.00058
0.04	0.02	0.0202	0.0099	0.0198	0.0099	0.00144	0.00058	0.00111	0.00111
	0.04	0.0203	0.0197	0.0198	0.0198	0.00146	0.00111	0.00113	0.00113

Table 8
Mean and Standard Deviation of MAE of simulated data and solution obtained with estimated parameters and MAV of noise from simulated data for time-step 0.2.

		MAE		MAV		SD of MAE		SD of MAV	
σ_N	σ_A	N	A	ε_N	ε_A	N	A	ε_N	ε_A
0.02	0.02	0.0098	0.0097	0.0097	0.0098	0.00095	0.00079	0.00080	0.00080
	0.04	0.0101	0.0193	0.0098	0.0196	0.00104	0.00167	0.00078	0.00078
0.04	0.02	0.0197	0.0097	0.0196	0.0097	0.00192	0.00085	0.00163	0.00163
	0.04	0.0198	0.0193	0.0197	0.0196	0.00189	0.00168	0.00157	0.00157

Table 9
Mean and Standard Deviation of MAE of simulated data and solution obtained with estimated parameters and MAV of noise from simulated data for time-step 0.4.

		MAE		MAV		SD of MAE		SD of MAV	
σ_N	σ_A	N	A	ε_N	ε_A	N	A	ε_N	ε_A
0.02	0.02	0.0092	0.0092	0.0096	0.0096	0.00129	0.00114	0.00109	0.00109
	0.04	0.0096	0.0184	0.0096	0.0192	0.00144	0.00239	0.00109	0.00109
0.04	0.02	0.0183	0.0093	0.0193	0.0096	0.00259	0.00114	0.00234	0.00234
	0.04	0.0186	0.0186	0.0192	0.0192	0.00250	0.00227	0.00222	0.00222

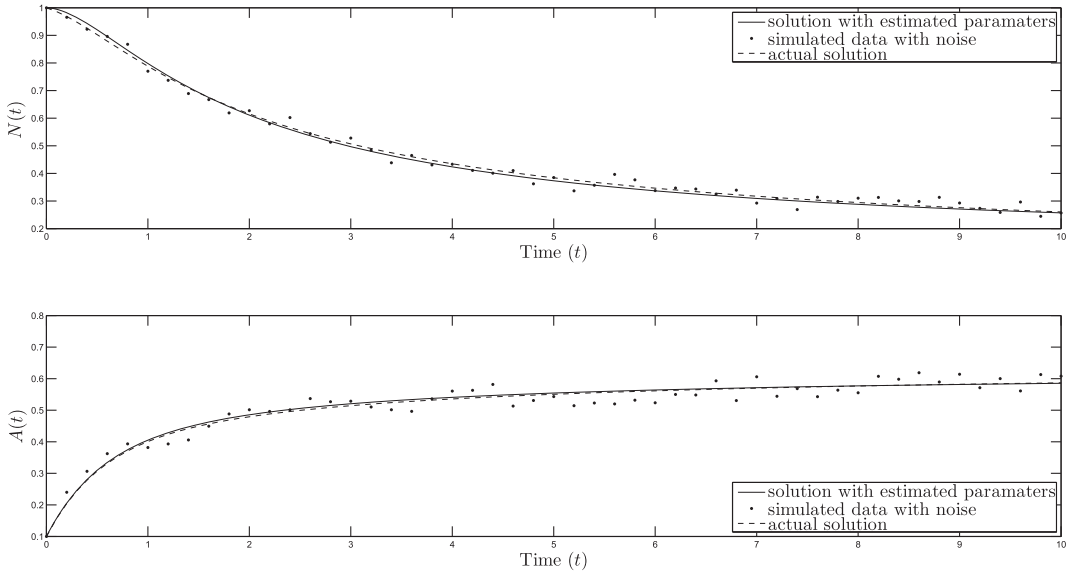

Fig. 2. Comparison of 51 points of simulated data with noise where $\sigma_N, \sigma_A = 0.04$, solution to the model and solution of model with estimated parameters $(k_{0.2}, \theta_{0.2}, \mu_{0.2}, \lambda_{0.2}, \gamma_{0.2}) = (1.12, 1.15, 1.57, 1.23, 0.527)$.

gests that the approximated solutions fit the data on average as well as the solution used to generate the noisy data, which one would assume would be close to the solution of best fit. We can see that if we plot the solution of the ODEs with the estimated parameter values with significant relative errors, we find that the solutions produced still fits the data well. An example of this is shown in Fig. 2, in which we can see there is a 40% error in the value of k and a 43% error in the estimated value of μ, however the solution still represents a good fit to the data. This shows that due to the fact this system has a large number of parameters to be estimated, there is a greater set of values of the parameters that will generate solutions for the ODE that will represent a good fit to the observed data. This suggests that for certain systems our method will not always be appropriate. However our method could potentially be used in conjunction with an alternative method that requires an initial estimate of the parameters values to work, such as the numerical least squares method [12], in order to save computation time by producing a set of values that are close to the values that generate the solution of best fit.

4. Conclusions

We have presented a method for finding unknown parameters, in systems of ODEs, that is integration based. Three model systems have been chosen to demonstrate the validity of this method. We have seen that this method is relatively simple and therefore requires very little computational time. We have noted that this method could be used to obtain an estimate close to the parameter values and then that estimate can be used for the starting point of a numerical least squares procedure that can then be used to find the parameters that generate the least squares best fit to the data. Further work that can be conducted is a review of the class of problems for which this method is appropriate as we have seen that for certain systems, in the presence of noise, this method can become unreliable. A procedure for how to choose appropriate weight functions and equation generating parameters for a particular system is needed. Analysis is needed to determine if certain weight functions are more appropriate for certain sets of data and whether or not certain weight functions can be used to give more weight to sections of the data that is available. An analysis of the errors that are to be expected from this method needs to be conducted

to determine the suitability of this method for particular systems of ODEs and the errors that can be expected as a result of non-uniform data.

References

[1] R.C. Aster, C.H. Thurber, Parameter Estimation and Inverse Problems, Academic Press, Waltham, MA, 2012.
[2] H.T. Banks, S. Grove, S. Hu, Y. Ma, A hierarchical Bayesian approach for parameter estimation in HIV models, Inverse Problems 21 (6) (2005) 1803–1822.
[3] Y. Bard, Nonlinear Parameter Estimation, Academic, London, 1974.
[4] N. Bellomo, J.A. Adam, A Survey of Models for Tumor-Immune System Dynamics, Birkhauser, Boston, 1997.
[5] E. Beretta, Y. Kuang, Modeling and analysis of a marine bacteriophage infection, Mathematical Biosciences 149 (1) (1998) 57–76.
[6] D. Bernoulli, Essai d'une nouvelle analyse de la mortalité caus'ee par la petite vérole. Mém. Math. Phys. Acad. Roy. Sci., Paris, 1766. 1. (Reprinted in: L.P. Bouckaert, B.L. van der Waerden (Eds.), Die Werke von Daniel Bernoulli, Bd. 2 Analysis und Wahrscheinlichkeitsrechnung, Birkhauser, Basel, 1982, p. 235. English translation entitled An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it in: L. Bradley, Smallpox Inoculation: An Eighteenth Century Mathematical Controversy, Adult Education Department, Nottingham, 1971, p. 21. Reprinted in: S. Haberman, T.A. Sibbett (Eds.) History of Actuarial Science, vol. VIII, Multiple Decrement and Multiple State Models, William Pickering, London, 1995, p. 1.).
[7] D. Bernoulli, S. Blower, An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it, Reviews in Medical Virology 14 (5) (2004) 275–288.
[8] N.F. Britton, Essential Mathematical Biology, Springer, New York, 2003.
[9] H.M. Byrne, Modelling avascular tumour growth, in: L. Preziosi (Ed.), Cancer Modelling and Simulation, Chapman and Hall/CRC, 2003, pp. 75–120.
[10] K. Dietz, J.A.P. Heesterbeek, Daniel Bernoulli's epidemiological model revisited, Mathematical Biosciences 180 (2002) 1–21.
[11] W.R. Esposito, C.A. Floudas, Parameter estimation in nonlinear algebraic models via global optimization, Computers and Chemical Engineering 22 (1998) S213–S220.
[12] P.W. Hemker, Numerical methods for differential equations in system simulation and in parameter estimation, in: P.W. Hemker, B. Hess (Eds.), Analysis and Simulation of Biochemical Systems, North-Holland, 1972, pp. 59–80.
[13] Y. Huang, D. Liu, H. Wu, Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system, Biometrics 62 (2) (2006) 413–423.
[14] Z. Li, M. Osbourne, T. Prvan, Parameter estimation of ordinary differential equations, IMA Journal of Numerical Analysis 25 (2005) 264. 235.
[15] H. Liang, H. Wu, Parameter estimation for differential equation models using a framework of measurement error in regression models, Journal of the American Statistical Association 103 (484) (2008) 1570–1583.
[16] F.L. Meyskens Jr, S.P. Thomson, T.E. Moon, Quantitation of the number of cells within tumor colonies in semisolid medium and their growth as oblate spheroids, Cancer Research 44 (1) (1984) 271.
[17] M.A. Nowak, R.M. May, Virus Dynamics: Mathematical Principles of Immunology and Virology, Oxford University Press, Oxford, 2000.
[18] H. Peng, L. Li, Y. Yang, C. Wang, Parameter estimation of nonlinear dynamical systems based on integrator theory, Chaos 19 (3) (2009) 033130.
[19] A.S. Perelson, P. Essunger, Y. Cao, M. Vesanen, A. Hurley, K. Saksela, M. Markowitz, D.D. Ho, Decay characteristics of HIV-1-infected compartments during combination therapy, Nature 387 (6629) (1997) 188–191.
[20] A.S. Perelson, P.W. Nelson, Mathematical analysis of HIV-I: dynamics in vivo, SIAM Review 41 (1) (1999) 3–44.
[21] A.S. Perelson, A.U. Neumann, M. Markowitz, J.M. Leonard, D.D. Ho, HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time, Science 271 (5255) (1996) 1582–1586.
[22] H. Putter, S.H. Heisterkamp, J.M.A. Lange, F. De Wolf, A Bayesian approach to parameter estimation in HIV dynamical models, Statistics in Medicine 21 (15) (2002) 2199–2214.
[23] J.O. Ramsay, G. Hooker, D. Campbell, J. Cao, Parameter estimation for differential equations: a generalized smoothing approach, Journal of the Royal Statistical Society Series B (Statistical Methodology) 69 (5) (2007) 741–796.
[24] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, Springer, New York, 2005.
[25] I. Siekmann, H. Malchow, E. Venturino, An extension of the Beretta–Kuang model of viral diseases, Mathematical Biosciences and Engineering 5 (3) (2008) 549–565.
[26] W.Y. Tan, H. Wu, Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention, World Scientific, Hackensack, NJ, 2005.
[27] H. Wu, A.A. Ding, Population HIV-1 dynamics in vivo: applicable models and inferential tools for virological data from AIDS clinical trials, Biometrics 55 (2) (1999) 410–418.
[28] H. Wu, A.A. Ding, V. De Gruttola, Estimation of HIV dynamic parameters, Statistics in Medicine 17 (21) (1998) 2463–2485.
[29] H. Wu, H. Zhu, H. Miao, A. Perelson, Parameter identifiability and estimation of HIV/AIDS dynamic models, Bulletin of Mathematical Biology 70 (2008) 785–799.
[30] Q. Zhou, P. Guo, G.D. Kruh, P. Vicini, X. Wang, J.M. Gallo, Predicting human tumor drug concentrations from a preclinical pharmacokinetic model of temozolomide brain disposition, Clinical Cancer Research 13 (14) (2007) 4271–4279.