Nonlinear dynamics of the observed HIV-positive population size in Fiji

Abstract: In this paper, the reported numbers of HIV-positive individuals in Fiji between and including the years 1989 and 2002 have been used to study the dynamics of the observed HIV-positive population size. The Fiji data showed a pattern that appears to begin to approximate the "S" shaped growth pattern. It is well-known in the mathematics literature that this nonlinear pattern is a natural growth phenomenon and that there are suitable mathematical equations - from the classical Verhulst equation to the recent Tsoularis equation - that can be used to study it. These equations generate two-dimensional curves, the uses of which are case-dependent. In this paper, the curves used are the Logistic Curve and the Gompertz Curve, which are known to be relatively accurate in predicting the short-term behaviour of the observed HIV-positive population size at the initial stages of growth. This study showed that the curves, based on the 1989-2002 data, gave the scenarios in which new reported cases could stabilize by 2005 or 2013. Based on the 1989-2001 data, both curves gave the worst-case scenario of continuous increase till the year 2020 when the numbers could begin to stabilize. (Pacific Health Dialog 2003, Vol.10 (2); Pg 45-52)

Jito Vanualailai*
Arun Deo**
Bibhya Sharma*
Subhash Chandra*

Introduction

It is only relatively recent that the emphasis was again made that growth curve modeling, which does not include hypothetical kinetics of various aspects of the virus transmission, but which rests solely on the basis of reported numbers, may provide the simplest, yet rigorous, means of studying the dynamics of the HIV-positive population. This suggestion was made in 1997 by Horimoto et al.[2] They argued that the basic epidemiological parameters associated with the spread of the virus were not fully understood. Thus, mathematical models, which were constructed on the basis of the hypothetical kinetics of various aspects of the virus transmission were partially pedagogical and may not include many of the factors, which could and should be included in more realistic models. These pitfalls could be avoided by studying the growth pattern of the HIV-positive population. The work of Horimoto et al.[2] was carried out based on the 1985-1994 data. The unmodified Gompertz curves used in the study projected an approximation that has been remarkably acceptable, at least right up to 2000, when the curves gave an approximation of between 300 and 400 reported cases. The actual figure was 327.[6]

*Department of Mathematics and Computing Science, University of the South Pacific, Suva, FIJI, vanualailai@usp.ac.fj, **School of Public Health and Primary Care, Fiji School of Medicine, Suva, FIJI.

Following the work of Horimoto et al.[2], this paper investigates the behaviour of the HIV-positive population directly from the observed data of the size of the population, rather than from the hypothetical kinetics with the parameters on the population dynamics. This paper also reiterates that growth models of the types considered in this paper, and classified as Type I models by Palloni and Glicklich[4], are appropriate to use in the initial stages of the growth of the HIV-positive population, and hence are particularly well-suited for small South Pacific island nations which are just experiencing the beginning of the AIDS epidemic.

Type I models are useful tools of analysis in situations where an initial slow growth rate, followed by a rapid growth rate, could again exhibit slow growth rate over a time interval due to some inhibiting factors, which, in this case, include safer sexual activities, effective public education on HIV and AIDS, and racial, cultural and religious constraints[1].

Type I models have well-known shortcomings, including, (i) the data may not reflect the true numbers, and (ii) the inability to predict abrupt changes because epidemiological data are not incorporated. However, with respect to point (i), the work of Horimoto et al.[2] showed that if the data reported were true reflections of the size of the HIV-positive population, then that was sufficient (but not necessary) to study the dynamics of the population, given that a reported case was independent of the mode of transmission and/or the behaviour or type of the infected individual. Hence, for example, the 1990 work of Goforth[1] was not based on reported numbers of HIV-positive since no data existed then, but rather on the numbers representing people who would most likely contract the virus, such as male homosexuals and bisexuals and female prostitutes.

In this sense, the work of Goforth[1] was hypothetical, a result of which could be unrealistic estimation. This happened in Goforth's work. With respect to point (ii), Horimoto et al.[2] showed that it was possible to modify Type I models appropriately to capture abrupt features and yet provide estimates that would not significantly differ from those produced by the original models. It is best to stress here that Type I models simply attempt to fit a suitable pre-defined function to observed HIV-positive or AIDS cases - new or accumulative. Assumptions about incubation period, mode of transmission, and epidemiological and biological assumptions, are not incorporated in Type I models. These assumptions are incorporated in more advanced models designated as Type II, Type III and Type IV. More information about these can be found in Palloni and Glicklich[4].

Given that Fiji is at the initial stages of the growth of HIV-positive population and the collection of data started in 1989, it is quite ideal now to attempt Type I modelling for Fiji. This has not been done, hence, this paper. Some advantages of Type I models are simplicity, elegance and the fact they can be used annually to offer new predictions because they depend only on observed data.

Growth Curve Models

In this section, for the sake of the completeness of the paper, we briefly describe the Logistic and Gompertz curves, and discuss why they were used in this study. An excellent and recent review on growth models can be found in Tsoularis.[5]

Both the Logistic and Gompertz curves could be derived from the Generalized Logistic Curve

$$Y = A + \frac{C}{\left[1 + Te^{-B(X-M)}\right]^{1/T}}$$

where Y is the population size, X is the year and A, B, C, M and T are the parameters to be determined, where

(a) A controls the lower asymptote,
(b) C controls the upper asymptote,
(c) M controls the time of maximum growth, with $X = M$ being the X-coordinate of the inflection point (where growth rate, dY/dX, is maximum)
(d) B controls the growth rate at the inflection point, and
(e) T controls where maximum growth occurs - near the lower asymptote if T is small, or near the upper asymptote if T is large.

It is sufficient that the parameters be real-valued, with, $T > 0$, to have the Generalized Logistic Curve continuous at all real X.

Now, when $T = 1$, the Generalized Logistic Curve trivially reduces to the Logistic Curve

$$Y = A + \frac{C}{1 + e^{-B(X-M)}}$$

Given that Fiji is at the initial stages of the HIV-positive population growth (see Figure 1), it can be safely assumed that the growth rate of new HIV-positive cases, dY/dX, is positive, implying that $B > 0$. Now, since M is fixed, a key feature of the Logistic Curve is that $\lim_{X \to \infty} Y = A + C$, which means that the population Y will approach its carrying capacity, $A + C$. Another key feature of the Logistic Curve is that the population size at the inflection point, $Y_{\inf}$, is exactly half the carrying capacity, that is, $Y_{\inf} = (A + C)/2$.

Now, when T is small, $Y_{\inf}$ occurs near the lower asymptote. When $T \to 0$, the Generalized Logistic Curve reduces to the Gompertz Curve $Y = A + Ce^{-e^{-B(X-M)}}$

Note that here the carrying capacity could either be $A + C$ (in which case $e^{-B(X-M)} \to 0$ as $X \to \infty$) or A (in which case $e^{-B(X-M)} \to \infty$ as $X \to \infty$). Thus the Gompertz Curve could be used to model two possible scenarios: since occurs nearer to the lower asymptote,

1. the fastest growth rate has just occurred and the current growth rate, dY/dX, remains positive, but possibly slight, so that the growth pattern will level out later $(B > 0)$ near $Y = A + C$, or,

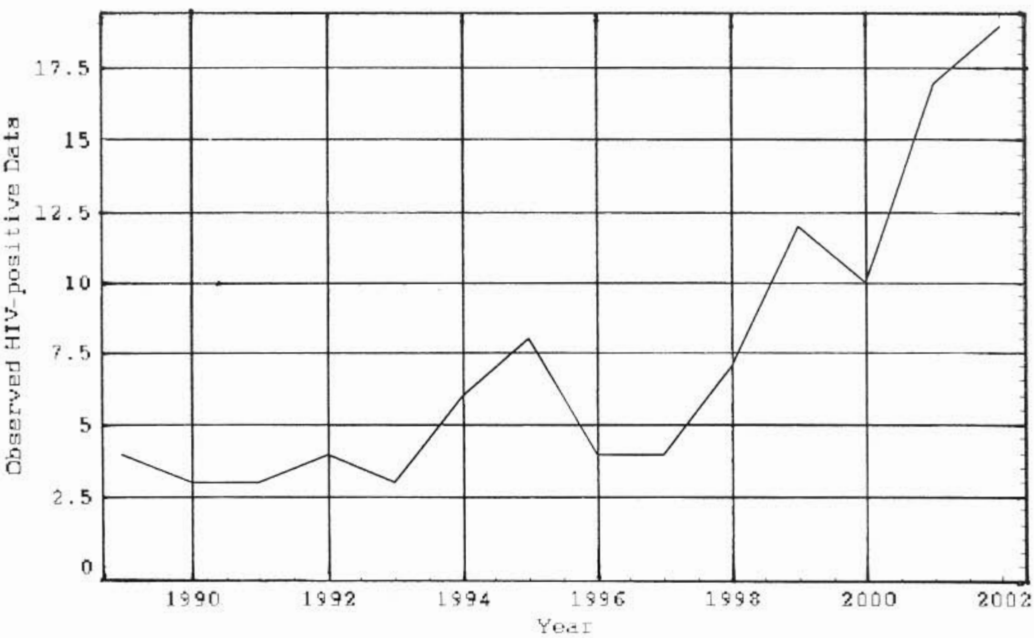
2. the fastest growth rate occurred much earlier and therefore the collected set of data presents a growth pattern that will level out sooner $(B < 0)$ near $Y = A$.

It is now apparent why the Logistic Curve and Gompertz Curve were used in this study. Since Fiji is at the initial stages of the HIV-positive population growth, for best-case scenarios, the Gompertz Curve may produce a more accurate result, the assumption being that the fastest growth rate has just occurred (Y_{inf} is nearer the lower asymptote) due to inhibiting factors, one of which is the on-going government-supported public education on HIV and AIDS. For worst-case scenarios, the fastest growth rate may not have occurred (is further away from the lower asymptote), so the Logistic Curve is ideal.

interval, in which the highest average growth rate was registered in the 2000-2001 interval. The slight increase in average growth rate recorded in 2002 may mask future high growth rates. Hence, for worst-case scenarios afforded by the two models, the numbers leading up to and including 2001 were also used to study the future pattern of the population size.

Growth Pattern Based on the 1989-2001 Data
Letting 1989 be the reference year ($X = 0$), the estimation of the parameters A, B, C and M was carried out using nonlinear regression analysis[2]. For this, the authors have utilized the advanced algorithms and extended precision capability of Mathematica, one of the more powerful Computer Algebra Systems proven independently to have performed reliably on curve-fitting benchmark tests[3]

Figure 1: Fiji is at the initial stages of growth of HIV-positive population, and the trend seems to show the beginning of the "S" shaped growth pattern.



1989-2002 data. Source: Ministry of Health, Fiji Government.

Data Analysis

Table 1: Shows the data on the reported cases of HIV-positive individuals.

Year	89	90	91	92	93	94	95	96	97	98	99	00	01	02
Total	4	3	3	4	3	6	8	4	4	7	12	10	17	19

1989-2002 data: Source: Ministry of Health, Fiji Government

Plotting these numbers as shown in Figure1 gives us a more vivid picture of the average growth rates, $\Delta Y/\Delta X$, in the 1989-2002 time interval.

In general, the 1989 and 2002 period could be divided into two: (i) the slow growth pattern in the 1989-1998 interval and (ii) the high growth pattern in the 1998-2002

In Table 2, the estimates of the parameters, via the steepest descent method, are shown.

Table 1: Estimation of the parameters based on the 1989-2001 data.

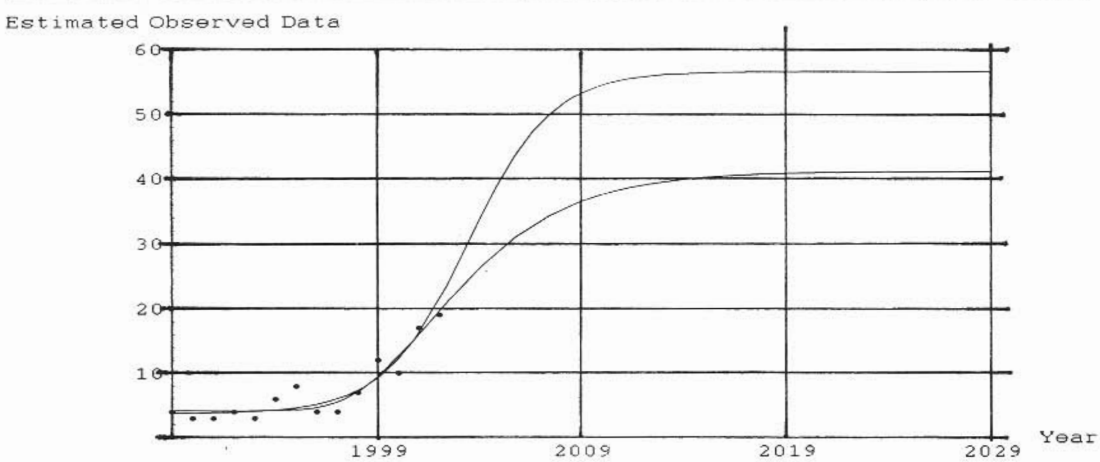
Parameters (Steepest Descent)	Logistic Curve $Y = A + \frac{C}{1+e^{-B(X-M)}}$	Gompertz Curve $Y = A + Ce^{-e^{-B(X-M)}}$
A	3.741	4.182
B	0.482	0.267
C	52.822	36.996
M	14.420	12.475

Among the many diagnostic statistical tests for best fit, the estimated variance and the confidence intervals are two easily understood indicators of the fit. Table 3 gives these. The lower the estimated variance and the smaller the range of the confidence interval in which the parameter is found, the better the fit.

Table 2: Confidence interval (95%) and estimated variance based on the 1989-2001 data.

Parameters (Steepest Descent)	Logistic Curve Confidence Interval	Gompertz Curve Confidence Interval
A	(1.009, 6.474)	(2.335, 6.028)
B	(-0.729, 1.692)	(-0.789, 1.323)
C	(-665.012, 770.656)	(-207.026, 280.958)
M	(-27.685, 56.525)	(-10.427, 35.378)
Estimated Variance	3.906	4.179

Figure 2: The Logistic Curve (top) and the Gompertz Curve based on the 1989-2001 data.



Growth Pattern Based on the 1989-2002 Data

The slight increase in average growth rate in 2002 significantly reduces future estimates. Table 4 gives the parameter estimates, via the Levenberg-Marquardt method, and Figure 3 gives the curves.

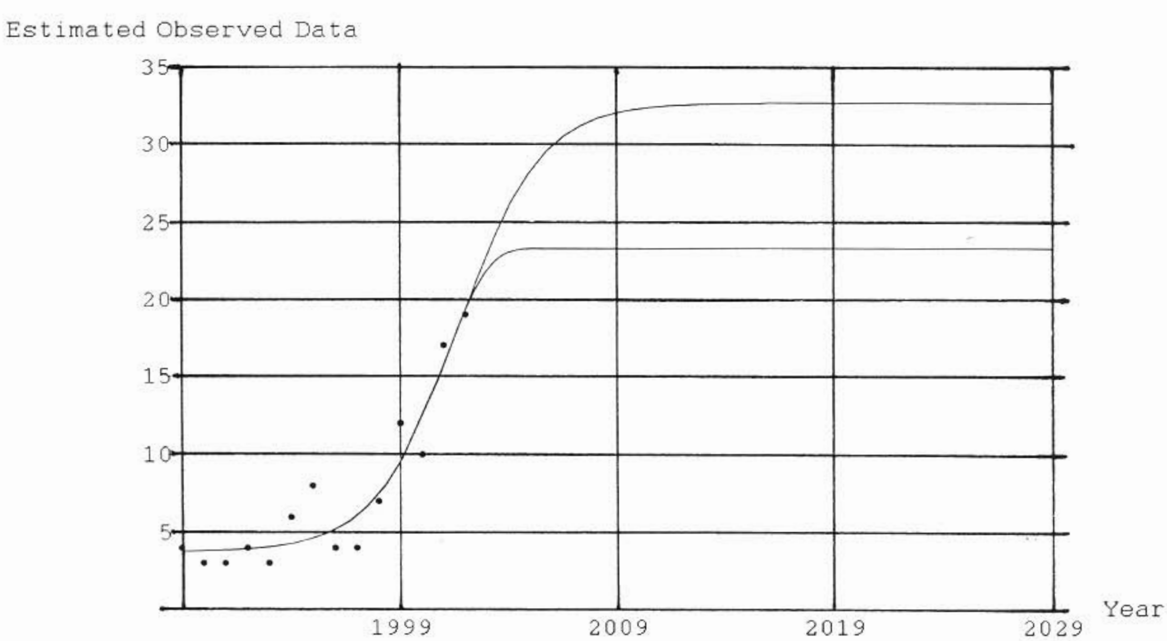
Table 3: Estimation of the parameters based on the 1989-2002 data.

Parameters (Steepest Descent)	Logistic Curve $Y = A + \frac{C}{1+e^{-B(X-M)}}$	Gompertz Curve $Y = A + Ce^{-e^{-B(X-M)}}$
A	3.743	23.336
B	0.516	¯0.505
C	28.915	¯19.588
M	12.687	12.097

Table 4: Confidence interval (95%) and estimated variance based on the 1989-2002 data.

Parameters (Steepest Descent)	Logistic Curve Confidence Interval	Gompertz Curve Confidence Interval
A	(1.477, 6.009)	(-6.175, 52.848)
B	(-0.243, 1.276)	(-1.091, 0.082)
C	(-47.516, 105.346)	(-50.186, 11.010)
M	(2.890, 22.485)	(6.572, 17.622)
Estimated Variance	3.625	3.606

Figure 3: The Logistic Curve (top) and the Gompertz Curve based on the 1989-2002 data.



Discussions

Table 5: Rounded-up estimates compared with reported numbers.

Year	Observed Data	Logistic 1989-2001	Gompertz 1989-2001	Logistic 1989-2002	Gompertz 1989-2002
1989	4	4	4	4	4
1990	3	4	4	4	4
1991	3	4	4	4	4
1992	4	4	4	4	4
1993	3	4	4	4	4
1994	6	4	4	4	4
1995	8	5	5	4	5
1996	4	5	5	5	5
1997	4	6	6	6	6
1998	7	7	7	7	7
1999	12	9	10	10	9
2000	10	12	13	12	12
2001	17	16	16	16	16
2002	19	21	20	19	19
2003		27	23	23	22
2004		34	26	26	23
2005		40	29	28	23
2006		45	32	30	23
2007		49	34	31	23
2008		51	35	32	23
2009		53	37	32	23
2010		54	38	32	23

Some observations can be made:

1. Both models, based on the 1989-2001 data, gave very good estimations of the 2002 observed population size, with the Logistic Curve providing 21 and the Gompertz Curve 20. The reported figure was 19.

2. The models predicted a worst-case scenario of about 27 new reported cases and best-case scenario of about 22 by the end of 2003[3]; 34 and 23 respectively for 2004, 40 and 23 respectively for 2005, etc. At the time of submission of this paper (14 October 2003), the reported cases so far were 25. This is, incidentally, the rounded-up average of the worst-case scenario (27) and the best-case scenario (22).

At the time of submission of this paper (14 October 2003), the reported cases so far were 25. This is, incidentally, the rounded-up average of the worst-case scenario (27) and the best-case scenario (22)

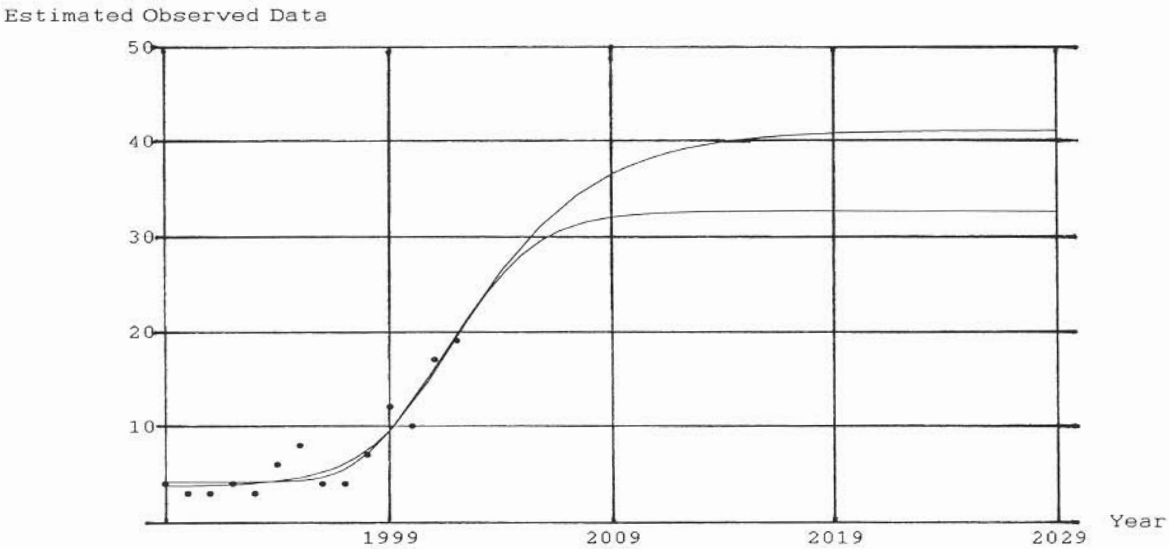
3. Table2 shows that for the Logistic Curve based on the 1989-2001 data, $M = 14.420$, which implies that, in the worst-case scenario, the highest growth rate could be recorded by mid-2003.

4. The curves, based on the 1989-2002 data (Figure 3), gave the scenarios in which new reported cases could stabilize by 2005 (about 23 per year onwards) or 2013 (about 33 per year onwards). Both curves, based on the 1989-2001 data (Figure 2), gave the worst-case scenario of continuous increase till the year 2020, after which the estimated number per year could be about 41 to 57 new reported cases.

5. The Gompertz Curve, based on the 1989-2001 data, and the Logistic Curve, based on the 1989-2002 data, interestingly provided similar trends as shown in Table 6 and in Figure 4.

6. The numbers shown in Table 6 seem more realistic than those published in 1990 by Goforth [1], who gave the HIV-positive accumulated figures, both reported and unreported, of about 60,000 in 2002, 85,000 in 2005 and 95,000 in 2010.

Figure 4: The Gompertz Curve of 1989-2001 (below) and the Logistic curve of 1989-2002 interestingly provide similar trends.



Conclusion

An implication of this study is that it is now possible for South Pacific island nations themselves to quickly carry out, at the end of every year as data becomes available, relatively easy and accurate short-term estimations via the method outlined in this paper. A novelty of the method is the use of both the Logistic and Gompertz curves to generate worst- and best-case scenarios. This was done on the basis of the mathematical properties as outlined in Section 2.

Since the models do not incorporate additional epidemiological data relating to HIV and AIDS, it is difficult to predict abrupt deviations from current trends. A good example of this can be seen in all the three figures in this paper (Figures 2, 3 and 4): the models cannot capture the feature that is the abrupt leap in 1994 and the subsequent fall in 1996 to pre-1994 values. If these abrupt deviations have already occurred, then careful analysis of the data is required to isolate these abrupt deviations and, if data are sufficient, to carry out estimations up to and including these drastic increases to see if reasonable trajectories could be derived from the models. In this case study, it clear that the abrupt leap in the 1994 was an isolated event, since the trend returned to pre-1994 values in 1996 and 1997, before starting on an almost steady growth rate till now.

Finally, having reliable data cannot be emphasized enough. For only then could the results of studies such as this be further strengthened.

Acknowledgement

The authors would like to thank Dr. Graham Roberts, of the Fiji School of Medicine, for his helpful comments and to a reviewer who pointed out a related item published in the Fiji Sunday Times of 14 March 2004 as explained in the Appendix.

Appendix

During the revision of this paper between the original submission date of 14 October 2003 and 14 May 2004, the researchers were alerted by an anonymous reviewer to a HIV/AIDS article published by the Fiji Sunday Times of 14 March 2004.

The figures in Table 1, reproduced below, were provided by the Ministry of Health to the researchers, showing annual figures from 1989 to 2002:

Table 6: 1989-2002 data. Source: Ministry of Health, Fiji Government.

Year	'89	'90	'91	'92	'93	'94	'95	'96	'97	'98	'99	'00	'01	'02
Total	4	3	3	4	3	6	8	4	4	7	12	10	17	19

The Fiji Sunday Times, of 14 March 2004, gave a figure of 26 for the year 2002. The discrepancy between that provided to the researchers (**19 cases**) and that published by the newspaper (**26 cases**) highlights the need for the Ministry of Health to accurately maintain its database given the importance of the matter at hand.

Assuming that the figure 26 for 2002 is correct, then based on the methods in this paper, Table 7 is constructed, showing the worst-case and best-case scenarios generated by the Logistic and Gompertz curves, respectively, on the 1989-2002 data.

Table 7: Likely trends 2003 – 2010 (reported cases only)

Year	2003	2004	2005	2006	2007	2008	2009	2010
Worst-case	36	49	63	75	84	91	95	98
Best-case	31	38	45	52	58	63	68	72

It is noted that the newspaper reported the figure of 31 for the year 2003. If this reported figure is correct, then it is a testimony to the effectiveness of the methods proposed in this paper.

References

1. R. R. Goforth, (1990). Forecasts for HIV infections and AIDS in Fiji, *Fiji Medical Journal*, 16, 4-15.
2. K. Horimoto, I. Hofuku and K. Oshima, (1997). Population dynamics of HIV positive by application of a locally modified Gompertz Curve: a case study for Japan, *Nonlinear Analysis, Theory, Methods and Applications*, 30, 985—993.
3. B. D. McCullough, (2000). The accuracy of Mathematical 4 as a statistical package, *Journal of Computational Statistics*, **15**, 279—299.
4. A. Palloni and M. Glicklich, (1991). Review of approaches to modelling the demographic impact of the AIDS epidemic, *The Aids Epidemic and Its Demographic Consequences*, United Nations and WHO, New York.
5. A. Tsoularis, (2001). Analysis of Logistic Growth Models, *Research Letters in the Information and Mathematical Sciences*, **2**, 23—46.
6. Epidemiological Fact Sheets by Country, UNAIDS/WHO/UNICEF, http://www.who.int/emc-hiv/factsheets/Allcountries.html

The reproduction of mankind is a great marvel and mystery. Had God consulted me in the matter, I should have advised him to continue the generation of the species by fashioning them of clay

Martin Luther (1483 - 1546)