IDENTIFICATION OF GUMEL-MICKENS HIV MODEL WITH INCOMPLETE DATA ON A POPULATION

M. Ngungu[1], MY. Shatalov[2], CR. Kikawa[3]

Tshwane University of Technology

[1]Merciend@yahoo.com

[2]Shatalovm@tut.ac.za

[3]Richard.kikawa@gmail.com

Department of Mathematics and Statistics

BIOMATH 2017

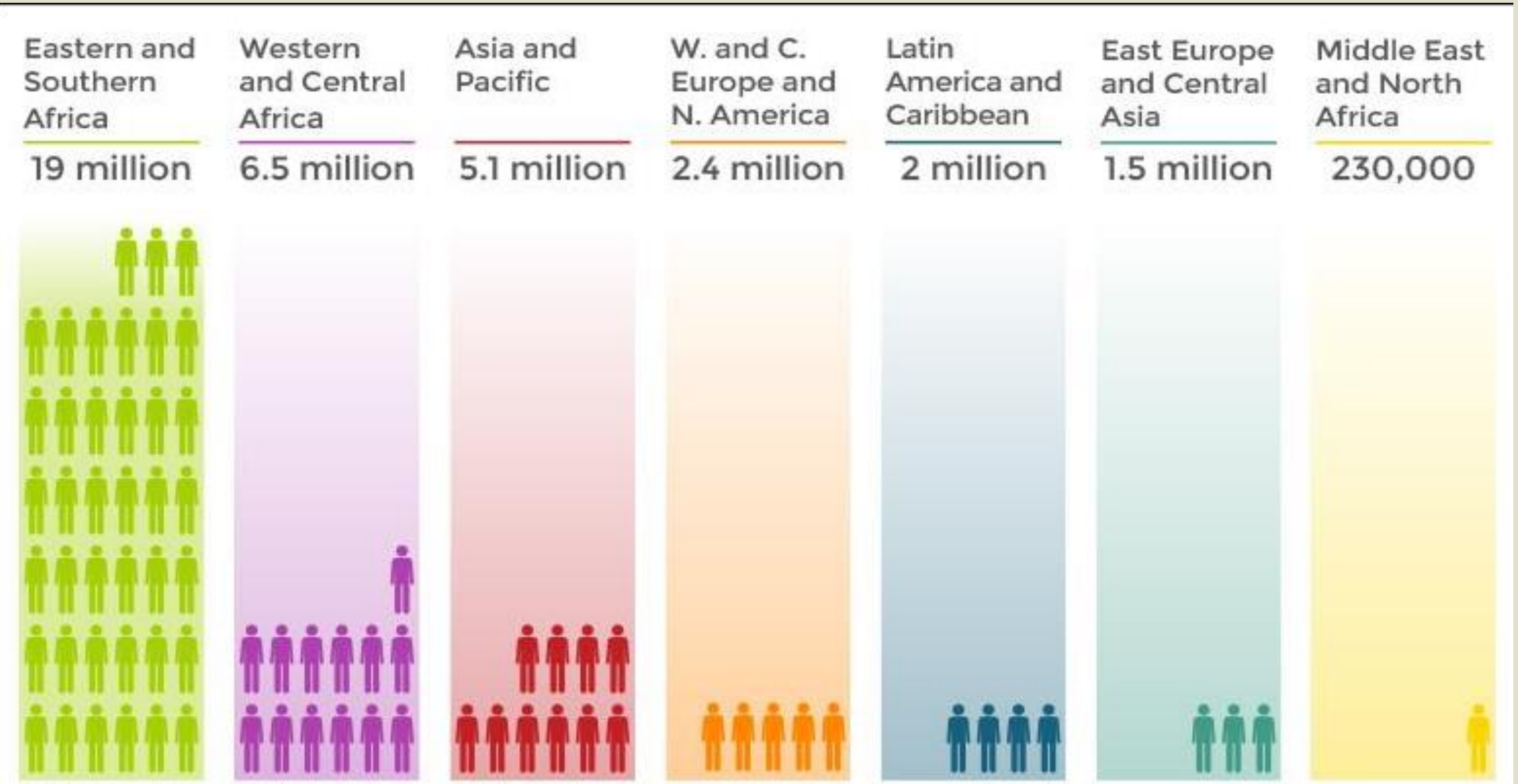
OUTLINE

INTRODUCTION

- Challenges/problems of incomplete information or unavailable data in modelling of disease dynamics
- Previously, (Fedotov & Shatalov, 2007) problems on estimation of dynamic system parameters from experimental data in linear systems with respect to their unknown coefficients have been discussed
- This talk presents a method of identifying mathematical models that can be used to model HIV/AIDs transmission dynamics in the presence of incomplete data.

INTRODUCTION-CONT…

HIV/AIDS population statistics by end 2015



Geographical region	World	Sub-Saharan Africa	South Africa
Living with HIV/AIDS	36.7 million	25.6 million	6.9 million
New Infections	2.1 million	0.97 million	0.38 million
Deaths	1.1 million	0.47 million	0.18 million

HIV/AIDS related challenges

- Affects health of a population,
- consequently, impacts development and social economic growth of households, communities and nations at large.
- Owing to these statistics there is a need to study, design, analyse and solve mathematical models for determining indicators that could give insight into eradicating HIV/AIDs.

OBJECTIVES

- In *"On identification of dynamical system parameters from experimental data"* the authors, (Mathye, Fedotov and Shatalov, 2004) the Gumel-Mickens HIV transmission model was considered with the aim of estimating the unknown parameters and restoring the information
- Analogous, with different parameters and coefficients unknown;
  - Fully identify the Gumel-Mickens HIV transmission model,
  - find all coefficients of this mathematical model and
  - restore information on
    - HIV- susceptible population
    - and uninfected-vaccinated population
- Ultimately, with the key objective of giving insights in eradicating HIV from the community.

THE GUMEL-MICKENS MODEL

- HIV transmission model that monitors the dynamics of HIV in four sexually active populations in the presence of vaccines as described by Gumel Moghadas and Mickens

$$D(T) = \begin{cases} \dot{X} = P_1 - \mu_1 X - \alpha_v \dfrac{XY_V}{N} - \alpha_w \dfrac{X(Y_W + Y_{WV})}{N} \\ \dot{Y}_V = P_2 - \mu_2 Y_V + \alpha_v \dfrac{XY_V}{N} - \gamma \dfrac{Y_V(Y_W + Y_{WV})}{N} \\ \dot{Y}_W = -\mu_3 Y_W + \alpha_w \dfrac{X(Y_W + Y_{WV})}{N} \\ \dot{Y}_{WV} = -\mu_4 Y_{VW} + \gamma \dfrac{Y_V(Y_W + Y_{WV})}{N} \end{cases}$$

With $t > t_0$ and initial conditions:

$X(t_0) = X_0,$ $\quad Y_W(t_0) = Y_{W0},$

$Y_V(t_0) = Y_{V0},$ $\quad Y_W(t_0) = Y_{W0},$

Parameter	Interpretation
$X(t)$	Population susceptible to HIV
$Y_V(t)$	Individuals uninfected by wild type and vaccinated population
$Y_W(t)$	Individuals infected by wild type and unvaccinated population
$Y_{WV}(t)$	Individuals infected by wild type and vaccinated population
$Z = Y_W(t) + Y_{WV}(t)$	Infected by wild type and vaccinated population
$ZZ = X(t) + Y_V(t)$	Healthy population
$N(t) = X(t) + Y_V(t) + Y_W(t) + Y_{WV}(t)$	Total (sexually active) population
P	Recruitment rate of susceptible population
$\mu_1, \mu_2, \mu_3, \mu_4$	Natural cessation of sexual activity
α_w and α_v	Rate of transmission of virus and vaccine, respectively
γ	Degree of protection against virus

Above parameters are taken from "Effect of a preventive vaccine on dynamics of HIV transmission" by A.B. Gumel, S.M. Moghadas and R.E. Mickens //*Communications in nonlinear Science and Numerical Simulations* , 9 2004, pp.649-659.

THE GUMEL-MICKENS MODEL-CONT…

$$N = X + Y_V + Y_W + Y_{WV}$$

P. Mathye : $N,\ X,\ Y_V$ - are known

Z= $Y_W + Y_{WV} = N - X - Y_V$ - also known

But: $Y_W = Y_W(\text{t})$ and $Y_{WV} = Y_{WV}(\text{t})$ – are not known

In this case: N, Y_W, Y_{WV} - are known

ZZ= $N - (Y_W + Y_{WV}) = X + Y_V$ - also known

But: $X = X(\text{t})$ and $Y_V = Y_V(\text{t})$ – are not known

and coefficients are unknown

THE GUMEL-MICKENS MODEL-CONT…

Assumptions:

- all dependent variables of G-M model and parameters are non-negative,
- all new individuals recruited into the society (either via birth or immigration) at a rate P per year are considered to be susceptible analogous to Gumel *et al.* ,
- population is reduced by the natural cessation of sexual activity at a constant rate l, and by infection with the virus which may be acquired from each new sexual partner at a rate α_w .

SOLUTION OF PROBLEM

- **Stage 1:**
  - Use direct method of initial problem solving
  - assume that all parameters and initial conditions of the four populations are known as below and solve the system numerically using Computer Algebra System (CAS)

$pi := 2000$	$p := 0.8$	$c := 5.0$	$\beta v := 0.5$	$\beta w := 0.45$	$\psi := 0.6$
$\mu := 0.031$	$d1 := 0.3$	$d2 := 0.25$	$d3 := 0.2$		

  - Recalculation of the parameters into new parameters

$p1 := (1 - p) \cdot pi$	$\mu 1 := \mu$	$\mu 2 := \mu + d1$	$\mu 3 := \mu + d2$	$\mu 4 := \mu + d3$
$p2 := p \cdot pi$	$\alpha v := c \cdot \beta v$	$\alpha w := c \cdot \beta w$	$g := (1 - \psi) \cdot \alpha w$	(g for gamma)

SOLUTION OF PROBLEM-CONT…

- System of ODE with known parameters

Right hand side of the Cauchy system $(u_0 = X,\ u_1 = Y_v, u_2 = Y_w\ u_0 = Y_{vw})$:

$$D(t,u) := \begin{bmatrix} p1 - \left[\alpha v \cdot \frac{u_0 \cdot u_1}{(u_0)+(u_1)+(u_2)+(u_3)}\right] - \left[\alpha w \cdot \frac{u_0 \cdot (u_2+u_3)}{(u_0)+(u_1)+(u_2)+(u_3)}\right] - \mu 1 \cdot u_0 \\ p2 + \left[\alpha v \cdot \frac{u_0 \cdot u_1}{(u_0)+(u_1)+(u_2)+(u_3)}\right] - \left[g \cdot \frac{u_1 \cdot (u_2+u_3)}{(u_0)+(u_1)+(u_2)+(u_3)}\right] - \mu 2 \cdot u_1 \\ \left[\alpha w \cdot \frac{u_0 \cdot (u_2+u_3)}{(u_0)+(u_1)+(u_2)+(u_3)}\right] - \mu 3 \cdot u_2 \\ \left[g \cdot \frac{u_1 \cdot (u_2+u_3)}{(u_0)+(u_1)+(u_2)+(u_3)}\right] - \mu 4 \cdot u_3 \end{bmatrix}$$

Initial conditions are also taken from the abovementioned paper:

$$u := \begin{pmatrix} 80 \cdot 10^3 & 2 \cdot 10^3 & 8 \cdot 10^3 & 8 \cdot 10^3 \end{pmatrix}^T$$

SOLUTION OF PROBLEM-CONT…

Time interval (T) is different (shorter than in the previously cited paper). **NN** is number of interval onto given time interval.

$T := 5 \quad NN := 25 \quad i := 0..NN$

Solution of system of equation by the specific method:

$U := \text{Adams}(u, 0, T, NN, D) \qquad TOL = 1 \times 10^{-3}$

Time interval (T) is different (shorter than in the previously cited paper). **NN** is number of interval onto given time interval.

$T := 5 \quad NN := 25 \quad i := 0..NN$

Solution of system of equation by the specific method:

$U := \text{Adams}(u, 0, T, NN, D) \qquad TOL = 1 \times 10^{-3}$

SOLUTION OF PROBLEM-CONT…

- **Stage 2:**
  - Forget the values of the parameter and assume that we know only N - total population and Y_W – population infected, $Y_{WV}(t)$ -population infected-vaccinated
  - It is also known that the sum N -Y_W -Y_{WV}= N - Z= $X + Y_V$, that is the sum of healthy population

$$N_i := \left(U^{\langle 1 \rangle}\right)_i + \left(U^{\langle 2 \rangle}\right)_i + \left(U^{\langle 3 \rangle}\right)_i + \left(U^{\langle 4 \rangle}\right)_i \qquad Z_i := \left(U^{\langle 3 \rangle}\right)_i + \left(U^{\langle 4 \rangle}\right)_i$$

$$ZZ_i := N_i - Z_i \qquad YW := U^{\langle 3 \rangle} \qquad YVW := U^{\langle 4 \rangle}$$

$$U^{\langle 1 \rangle} = X \quad U^{\langle 2 \rangle} = Y_v \quad U^{\langle 3 \rangle} = Y_w \quad U^{\langle 4 \rangle} = Y_{vw}$$

- Interpolation and subsequent differentiation using CAS

$$VYw := \mathrm{cspline}\left[\left(U^{\langle 0 \rangle}\right), U^{\langle 3 \rangle}\right] \qquad VYvw := \mathrm{cspline}\left(U^{\langle 0 \rangle}, U^{\langle 4 \rangle}\right)$$

$$Yw(t) := \mathrm{interp}\left(VYw, U^{\langle 0 \rangle}, U^{\langle 3 \rangle}, t\right) \qquad Yvw(t) := \mathrm{interp}\left(VYvw, U^{\langle 0 \rangle}, U^{\langle 4 \rangle}, t\right)$$

NUMERICAL RESULTS

- Numerical solutions

N =

	0
0	$9.8 \cdot 10^4$
1	$9.679 \cdot 10^4$
2	$9.52 \cdot 10^4$
3	$9.32 \cdot 10^4$
4	$9.078 \cdot 10^4$
5	$8.798 \cdot 10^4$
6	$8.489 \cdot 10^4$
7	$8.159 \cdot 10^4$
8	$7.82 \cdot 10^4$
9	$7.479 \cdot 10^4$
10	$7.143 \cdot 10^4$
11	$6.816 \cdot 10^4$
12	$6.501 \cdot 10^4$
13	$6.2 \cdot 10^4$
14	$5.913 \cdot 10^4$
15	...

YW =

	0
0	$8 \cdot 10^3$
1	$1.397 \cdot 10^4$
2	$2.086 \cdot 10^4$
3	$2.823 \cdot 10^4$
4	$3.541 \cdot 10^4$
5	$4.17 \cdot 10^4$
6	$4.655 \cdot 10^4$
7	$4.977 \cdot 10^4$
8	$5.143 \cdot 10^4$
9	$5.181 \cdot 10^4$
10	$5.121 \cdot 10^4$
11	$4.995 \cdot 10^4$
12	$4.826 \cdot 10^4$
13	$4.633 \cdot 10^4$
14	$4.428 \cdot 10^4$
15	...

YVW =

	0
0	$8 \cdot 10^3$
1	$7.725 \cdot 10^3$
2	$7.544 \cdot 10^3$
3	$7.493 \cdot 10^3$
4	$7.601 \cdot 10^3$
5	$7.869 \cdot 10^3$
6	$8.269 \cdot 10^3$
7	$8.745 \cdot 10^3$
8	$9.236 \cdot 10^3$
9	$9.689 \cdot 10^3$
10	$1.007 \cdot 10^4$
11	$1.037 \cdot 10^4$
12	$1.058 \cdot 10^4$
13	$1.07 \cdot 10^4$
14	$1.076 \cdot 10^4$
15	...

SOLUTION OF PROBLEM-CONT…

- Goal function

$$G = G(a_1, a_2, a_3, a_4)$$

$$G = \frac{1}{2}\sum_{k=0}^{n}\{ a_1\dot{Y}_{w,k} + a_2 Y_{w,k} + a_3\dot{Y}_{wv,k} + a_4 Y_{wv,k} - \frac{ZZ_k - Z_k}{N_k}\}^2$$

Where;

$a_1 = \alpha_w^{-1}$, $a_2 = \mu_3\alpha_w^{-1}$, $a_3 = \gamma^{-1}$, $a_4 = \mu_4\gamma^{-1}$

- We calculate,

$$X_k = \frac{\dot{Y}_{w,k}(t) + \mu_3 Y_{w,k}(t)}{\alpha_w}\frac{N_k}{Z_k}$$

$$Y_{v,k} = \frac{\dot{Y}_{wv,k}(t) + \mu_4 Y_{wv,k}(t)}{\gamma}\frac{N_k}{Z_k}$$

NUMERICAL RESULTS-CONT…

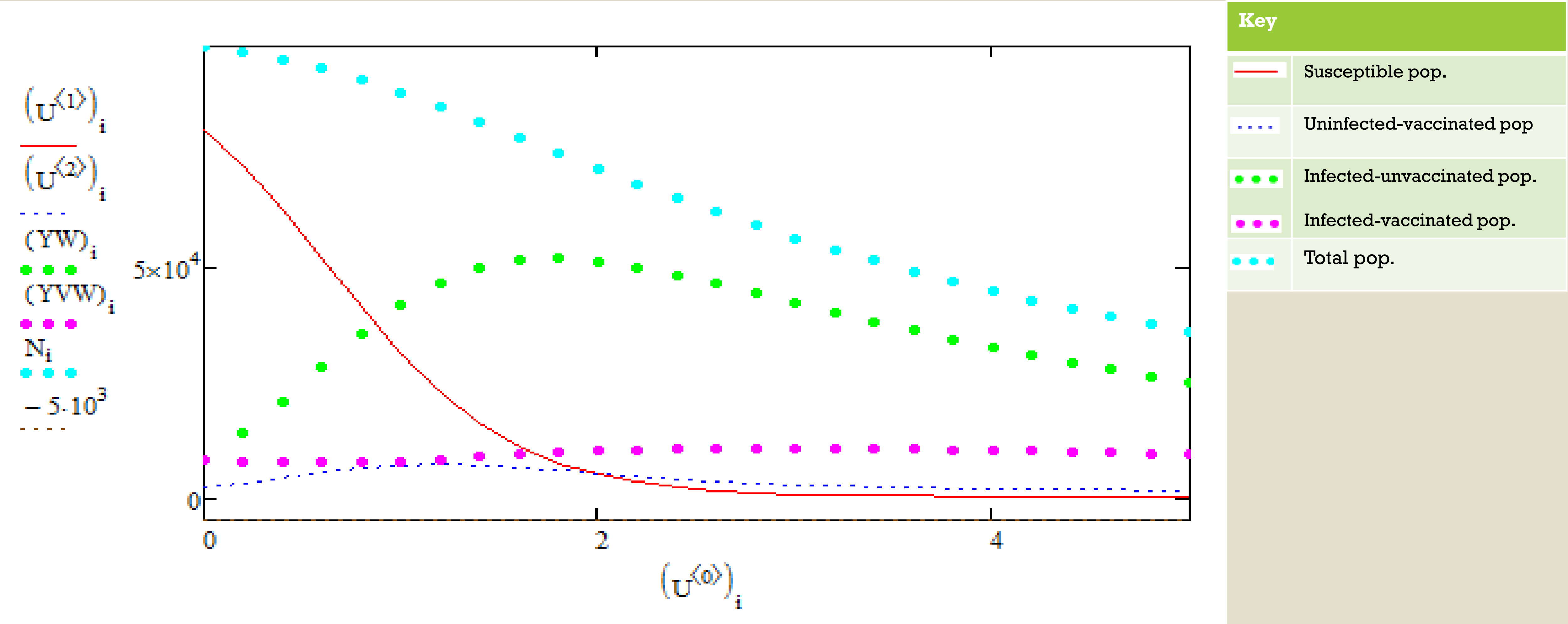


Figure 1: Population graph against time with solution of known information

NUMERICAL RESULTS-CONT…

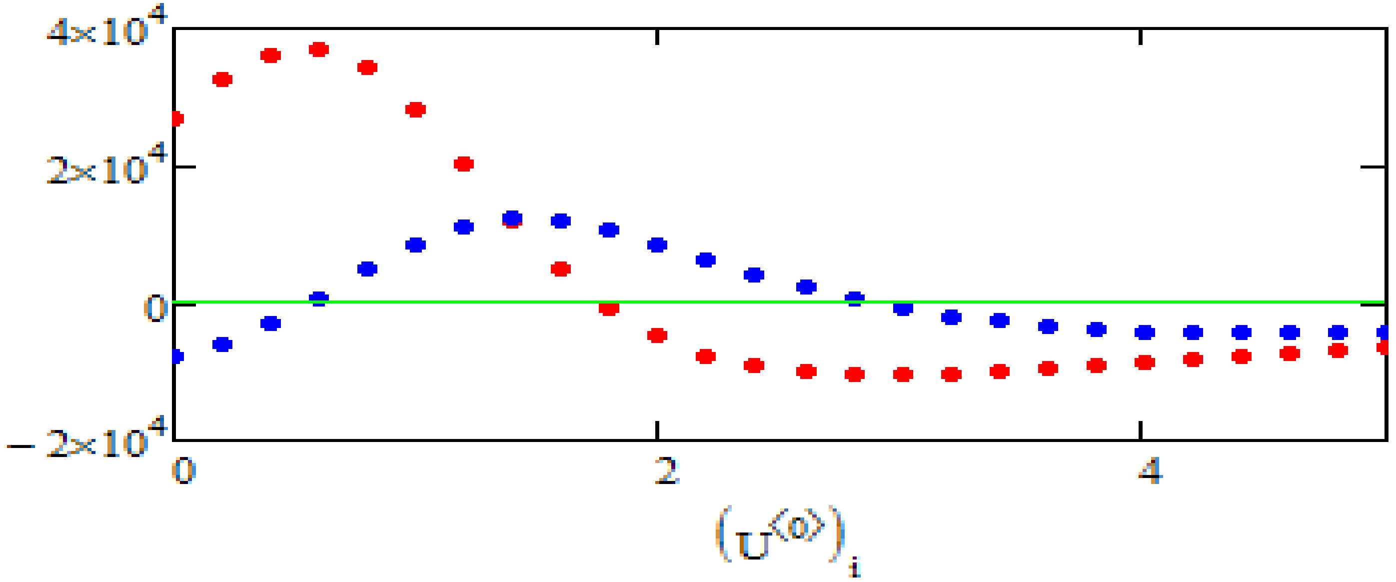
- Solution for the known on interpolation and differentiation

$$dYw(t) := \frac{d}{dt}Yw(t) \qquad dYvw(t) := \frac{d}{dt}Yvw(t) \qquad t := U^{\langle 0 \rangle}$$

$dYw\left[\left(U^{\langle 0 \rangle}\right)_i\right]$

$dYvw\left[\left(U^{\langle 0 \rangle}\right)_i\right]\cdot 5$

0



$$DYw_i := dYw\left[\left(U^{\langle 0 \rangle}\right)_i\right] \qquad DYvw_i := dYvw\left[\left(U^{\langle 0 \rangle}\right)_i\right] \qquad F_i := \frac{(N_i - Z_i)\cdot Z_i}{N_i}$$

Key	
red dots	Infected-unvaccinated pop
blue dots	Infected-vaccinated pop.

Figure 2: Population graph against time with known information ($Z = Y_W(t) + Y_{WV}(t)$)

SOLUTION OF PROBLEM-CONT…

- By means of Least square method we produce matrix of coefficients and RHS matrices for obtaining of unknown parameters:

$$a_{0,0} := \sum_i \left[(DYw_i)^2\right] \quad a_{0,1} := \sum_i \left[(DYw_i)\cdot(Yw_i)\right] \quad a_{0,2} := \sum_i \left[(DYw_i)\cdot(DYvw_i)\right] \quad a_{0,3} := \sum_i \left[(DYw_i)\cdot(Yvw_i)\right] \quad b_0 := \sum_i \left[(DYw_i)\cdot(F_i)\right]$$

$$a_{1,1} := \sum_i \left[(Yw_i)^2\right] \quad a_{1,2} := \sum_i \left[(Yw_i)\cdot(DYvw_i)\right] \quad a_{1,3} := \sum_i \left[(Yw_i)\cdot(Yvw_i)\right] \quad b_1 := \sum_i \left[(Yw_i)\cdot(F_i)\right]$$

$$a_{2,2} := \sum_i \left[(DYvw_i)^2\right] \quad a_{2,3} := \sum_i \left[(DYvw_i)\cdot(Yvw_i)\right] \quad b_2 := \sum_i \left[(DYvw_i)\cdot(F_i)\right]$$

$$a_{3,3} := \sum_i \left[(Yvw_i)^2\right] \quad b_3 := \sum_i \left[(Yvw_i)\cdot(F_i)\right]$$

$$A := \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} \\ a_{0,1} & a_{1,1} & a_{1,2} & a_{1,3} \\ a_{0,2} & a_{1,2} & a_{2,2} & a_{2,3} \\ a_{0,3} & a_{1,3} & a_{2,3} & a_{3,3} \end{pmatrix} \qquad |A| = 9.155 \times 10^{34} \qquad B := b$$

NUMERICAL RESULTS-CONT…

- Solution of parameters and coefficients

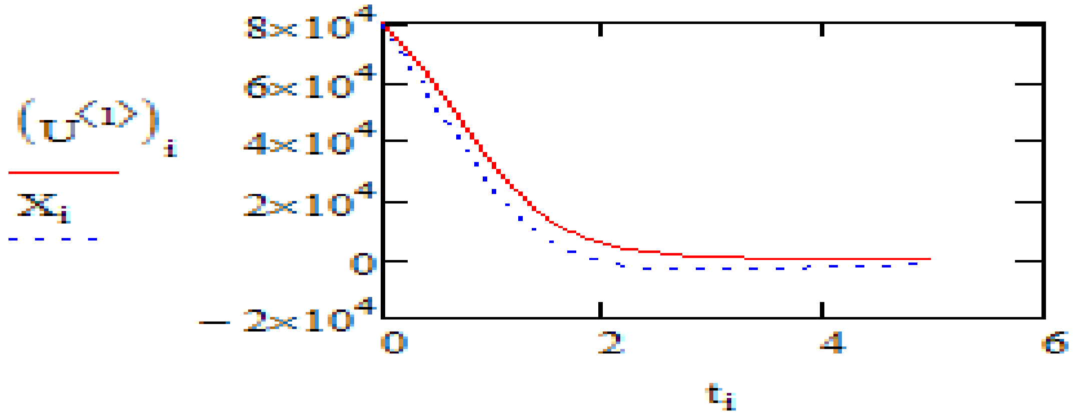
$$a1234 := A^{-1} \cdot B$$

$$a1234 = \begin{pmatrix} 0.444 \\ 0.12 \\ 1.144 \\ 0.274 \end{pmatrix}$$

$$\alpha w^{-1} = 0.444$$
$$\mu 3 \cdot \alpha w^{-1} = 0.125$$
$$g^{-1} = 1.111$$
$$\mu 4 \cdot g^{-1} = 0.257$$

$$X_i := \left(a1234_0 \cdot DYw_i + a1234_1 \cdot Yvw_i\right) \cdot \frac{N_i}{Z_i}$$

$$Yv_i := \left(a1234_2 \cdot DYvw_i + a1234_3 \cdot Yvw_i\right) \cdot \frac{N_i}{Z_i}$$



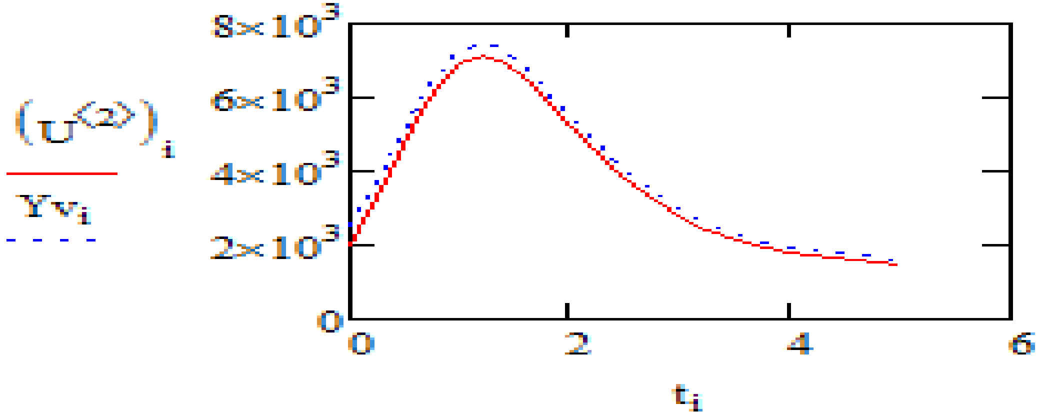


Figure 3: Population graph against time comparing the estimated and actual results

Results

- **Relative error of estimation in %**

$$\left|\frac{\left[\alpha w - \left(a1234_0\right)^{-1}\right]}{\alpha w}\right| \cdot 100 = 0.036 \quad (\%)$$

$$\left|\frac{\left(\mu 3 - \frac{a1234_1}{a1234_0}\right)}{\mu 3}\right| \cdot 100 = 3.551 \quad (\%)$$

$$\left|\frac{\left[g - \left(a1234_2\right)^{-1}\right]}{g}\right| \cdot 100 = 2.868 \quad (\%)$$

$$\left|\frac{\left(\mu 4 - \frac{a1234_3}{a1234_2}\right)}{\mu 4}\right| \cdot 100 = 3.57 \quad (\%)$$

Note:

- relative error can be further minimised by increasing the NN points
- All other parameters can be also estimated

DISCUSSIONS AND CONCLUSIONS

- The estimated parameters are calculated accurately
- The obtained $X(t)$ and $Y_v(t)$ values are in good qualitative and quantitative correspondence and possible improvement of accuracy could be achieved by proposing other estimation methods
- Applications,
  - by researchers in the medical or epidemiology filed for estimation
  - big data solutions aimed at restoring incomplete data or information
  - predictive analytics

REFERENCES

1. FEDOTOV I. & SHATALOV M. 2007. On identification of dynamic system parameters from experimental data. *RGMIA,* Victoria University, 10(1), 2: 106 – 116.
2. WHO, Global HIV/AIDs Report 2016. Available online from: http: //www.who.int/tb/publications/global_report/en [Accessed on 28th February 2017].
3. Mathye P. , FEDOTOV I. & SHATALOV M. 2015. On identification of dynamical system parameters from experimental data . BIOMATHS, 4: 1-7.
4. GUMEL A. B., MOGHADAS S. M., MICKENS R. E. 2004. Effect of a preventive vaccine on the dynamics of HIV transmission. *Communications in Nonlinear Science and Numerical Simulations*, 9: 649-659.


THANK YOU
ASANTE
GRAZIE
KIA ORA
MULTUMESC
OBRIGADO
SALAMAT
MAMANA
MATONDO
SPASIBO
MERCI
KIITOS
MATUR NUWUN
MOCHCHAKKERAM
ARIGATO
DANKON
DANK JE
CAM ON BAN
MAAKE
JUSPAXAR
VINAKA
WELALIN