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

M. Ngungu¹, MY. Shatalov², CR. Kikawa³
Tshwane University of Technology
¹Merciend@yahoo.com
²Shatalovm@tut.ac.za
³Richard.kikawa@gmail.com

Department of Mathematics and Statistics

BIOMATH 2017
OUTLINE

- Introduction
- Objectives
- The Gumel-Mickens Model
- Solution of the problem & numerical results
- Conclusions and applications
- References
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.
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

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

With \( t > t_0 \) and initial conditions:

\[
X(t_0) = X_0, \quad Y_V(t_0) = Y_{V0}, \\
Y_W(t_0) = Y_{W0}, \quad Y_W(t_0) = Y_{W0}, \\
Y_{WV}(t_0) = Y_{WV0}, \quad Y_{WV}(t_0) = Y_{WV0}
\]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X(t) )</td>
<td>Population susceptible to HIV</td>
</tr>
<tr>
<td>( Y_V(t) )</td>
<td>Individuals uninfected by wild type and vaccinated population</td>
</tr>
<tr>
<td>( Y_W(t) )</td>
<td>Individuals infected by wild type and unvaccinated population</td>
</tr>
<tr>
<td>( Y_{WV}(t) )</td>
<td>Individuals infected by wild type and vaccinated population</td>
</tr>
<tr>
<td>( Z = Y_W(t) + Y_{WV}(t) )</td>
<td>Infected by wild type and vaccinated population</td>
</tr>
<tr>
<td>( ZZ = X(t) + Y_V(t) )</td>
<td>Healthy population</td>
</tr>
<tr>
<td>( N(t) = X(t) + Y_V(t) + Y_W(t) + Y_{WV}(t) )</td>
<td>Total (sexually active) population</td>
</tr>
<tr>
<td>( p )</td>
<td>Recruitment rate of susceptible population</td>
</tr>
<tr>
<td>( \mu_1, \mu_2, \mu_3, \mu_4 )</td>
<td>Natural cessation of sexual activity</td>
</tr>
<tr>
<td>( \alpha_w ) and ( \alpha_v )</td>
<td>Rate of transmission of virus and vaccine, respectively</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>Degree of protection against virus</td>
</tr>
</tbody>
</table>

\[ N = X + Y_V + Y_W + Y_{WV} \]

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

\[ Z = Y_W + Y_{WV} = N - X - Y_V \] - also known

But: \( Y_W = Y_W(t) \) and \( Y_{WV} = Y_{WV}(t) \) - 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: \( X = X(t) \) and \( Y_V = Y_V(t) \) - are not known

and coefficients are unknown
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 $\alpha_w$. 

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)

\[
\begin{align*}
\pi_1 & := 2000 \\
p & := 0.8 \\
\xi & := 5.0 \\
\beta_v & := 0.5 \\
\beta_w & := 0.45 \\
\nu & := 0.6 \\
\mu & := 0.031 \\
d_1 & := 0.3 \\
d_2 & := 0.25 \\
d_3 & := 0.2
\end{align*}
\]

- Recalculation of the parameters into new parameters

\[
\begin{align*}
p_1 & := (1-p) \cdot \pi \\
\mu_1 & := \mu \\
\mu_2 & := \mu + d_1 \\
\mu_3 & := \mu + d_2 \\
\mu_4 & := \mu + d_3 \\
p_2 & := p \cdot \pi \\
\alpha_v & := \xi \cdot \beta_v \\
\alpha_w & := \xi \cdot \beta_w \\
g & := (1-\nu) \cdot \alpha_w 
\end{align*}
\]

\( (g \text{ for gamma}) \)
System of ODE with known parameters

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

\[
D(t,u) := \begin{bmatrix}
p_1 - \left[ \alpha \cdot \frac{u_0 \cdot u_1}{(u_0) + (u_1) + (u_2) + (u_3)} \right] - \left[ \alpha \cdot \frac{u_0 \cdot (u_2 + u_3)}{(u_0) + (u_1) + (u_2) + (u_3)} \right] - \mu_1 \cdot u_0 \\
p_2 + \left[ \alpha \cdot \frac{u_0 \cdot u_1}{(u_0) + (u_1) + (u_2) + (u_3)} \right] - \left[ \frac{u_1 \cdot (u_2 + u_3)}{g \cdot (u_0) + (u_1) + (u_2) + (u_3)} \right] - \mu_2 \cdot u_1 \\
\alpha \cdot \frac{u_0 \cdot (u_2 + u_3)}{(u_0) + (u_1) + (u_2) + (u_3)} - \mu_3 \cdot u_2 \\
\frac{u_1 \cdot (u_2 + u_3)}{g \cdot (u_0) + (u_1) + (u_2) + (u_3)} - \mu_4 \cdot u_3
\end{bmatrix}
\]

Initial conditions are also taken from the abovementioned paper:

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

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

\[
\begin{align*}
N_i &= (U^{<1>})_i + (U^{<2>})_i + (U^{<3>})_i + (U^{<4>})_i \\
Z_i &= (U^{<3>})_i + (U^{<4>})_i \\
ZZ_i &= N_i - Z_i \\
Y_W &= U^{<3>} \\
Y_{WV} &= U^{<4>} \\
U^{<1>} &= X \\
U^{<2>} &= Y_V \\
U^{<3>} &= Y_W \\
U^{<4>} &= Y_{WV}
\end{align*}
\]

Interpolation and subsequent differentiation using CAS

\[
\begin{align*}
VYw &= \text{cspline}\left(U^{<0>}, U^{<3>}\right) \\
VYvw &= \text{cspline}\left(U^{<0>}, U^{<4>}\right) \\
Yw(t) &= \text{interp}(VYw, U^{<0>}, U^{<3>}, t) \\
Yvw(t) &= \text{interp}(VYvw, U^{<0>}, U^{<4>}, t)
\end{align*}
\]
NUMERICAL RESULTS

- Numerical solutions

<table>
<thead>
<tr>
<th>N</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9.8·10^4</td>
</tr>
<tr>
<td>1</td>
<td>9.679·10^4</td>
</tr>
<tr>
<td>2</td>
<td>9.52·10^4</td>
</tr>
<tr>
<td>3</td>
<td>9.32·10^4</td>
</tr>
<tr>
<td>4</td>
<td>9.078·10^4</td>
</tr>
<tr>
<td>5</td>
<td>8.798·10^4</td>
</tr>
<tr>
<td>6</td>
<td>8.489·10^4</td>
</tr>
<tr>
<td>7</td>
<td>8.159·10^4</td>
</tr>
<tr>
<td>8</td>
<td>7.82·10^4</td>
</tr>
<tr>
<td>9</td>
<td>7.479·10^4</td>
</tr>
<tr>
<td>10</td>
<td>7.143·10^4</td>
</tr>
<tr>
<td>11</td>
<td>6.816·10^4</td>
</tr>
<tr>
<td>12</td>
<td>6.501·10^4</td>
</tr>
<tr>
<td>13</td>
<td>6.2·10^4</td>
</tr>
<tr>
<td>14</td>
<td>5.913·10^4</td>
</tr>
<tr>
<td>15</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>YW</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8·10^3</td>
</tr>
<tr>
<td>1</td>
<td>1.397·10^4</td>
</tr>
<tr>
<td>2</td>
<td>2.086·10^4</td>
</tr>
<tr>
<td>3</td>
<td>2.823·10^4</td>
</tr>
<tr>
<td>4</td>
<td>3.541·10^4</td>
</tr>
<tr>
<td>5</td>
<td>4.17·10^4</td>
</tr>
<tr>
<td>6</td>
<td>4.655·10^4</td>
</tr>
<tr>
<td>7</td>
<td>4.977·10^4</td>
</tr>
<tr>
<td>8</td>
<td>5.143·10^4</td>
</tr>
<tr>
<td>9</td>
<td>5.181·10^4</td>
</tr>
<tr>
<td>10</td>
<td>5.121·10^4</td>
</tr>
<tr>
<td>11</td>
<td>4.995·10^4</td>
</tr>
<tr>
<td>12</td>
<td>4.826·10^4</td>
</tr>
<tr>
<td>13</td>
<td>4.633·10^4</td>
</tr>
<tr>
<td>14</td>
<td>4.428·10^4</td>
</tr>
<tr>
<td>15</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>YWW</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8·10^3</td>
</tr>
<tr>
<td>1</td>
<td>7.725·10^3</td>
</tr>
<tr>
<td>2</td>
<td>7.544·10^3</td>
</tr>
<tr>
<td>3</td>
<td>7.493·10^3</td>
</tr>
<tr>
<td>4</td>
<td>7.601·10^3</td>
</tr>
<tr>
<td>5</td>
<td>7.869·10^3</td>
</tr>
<tr>
<td>6</td>
<td>8.269·10^3</td>
</tr>
<tr>
<td>7</td>
<td>8.745·10^3</td>
</tr>
<tr>
<td>8</td>
<td>9.236·10^3</td>
</tr>
<tr>
<td>9</td>
<td>9.689·10^3</td>
</tr>
<tr>
<td>10</td>
<td>1.007·10^4</td>
</tr>
<tr>
<td>11</td>
<td>1.037·10^4</td>
</tr>
<tr>
<td>12</td>
<td>1.058·10^4</td>
</tr>
<tr>
<td>13</td>
<td>1.07·10^4</td>
</tr>
<tr>
<td>14</td>
<td>1.076·10^4</td>
</tr>
<tr>
<td>15</td>
<td>...</td>
</tr>
</tbody>
</table>
SOLUTION OF PROBLEM-CONT...

➢ Goal function

\[ G = G(a_1, a_2, a_3, a_4) \]

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

Where;

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

➢ We calculate,

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

\[ Y_{v,k} = \frac{\dot{Y}_{wv,k}(t) + \mu_4 Y_{wv,k}(t)}{\gamma} \frac{N_k}{Z_k} \]
**Figure 1:** Population graph against time with solution of known information
Solution for the known on interpolation and differentiation

**Figure 2:** Population graph against time with known information \((Z = Y_W(t) + Y_{WV}(t))\)
By means of **Least square method** we produce matrix of coefficients and RHS matrices for obtaining of unknown parameters:

\[
\begin{align*}
A &= \begin{pmatrix}
    a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} \\
    a_{0,1} & a_{1,1} & a_{1,2} & a_{1,3} \\
    a_{0,2} & a_{1,2} & a_{2,2} & a_{2,3} \\
    a_{0,3} & a_{1,3} & a_{2,3} & a_{3,3}
\end{pmatrix} \\
|A| &= 9.155 \times 10^{34} \\
B &= b
\end{align*}
\]
Solution of parameters and coefficients

\[ a_{1234} := A^{-1} \cdot B \]

\[ a_{1234} = \begin{pmatrix} 0.444 \\ 0.12 \\ 1.144 \\ 0.274 \end{pmatrix} \]

\[ \alpha w^{-1} = 0.444 \]
\[ \mu_3 \alpha w^{-1} = 0.125 \]
\[ g^{-1} = 1.111 \]
\[ \mu_4 g^{-1} = 0.257 \]

\[ X_i := \left( a_{1234_0} \cdot D Yw_i + a_{1234_1} \cdot Yvw_i \right) \frac{N_i}{Z_i} \]

\[ Yv_i := \left( a_{1234_2} \cdot D Yvw_i + a_{1234_3} \cdot Yvw_i \right) \frac{N_i}{Z_i} \]

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

- Relative error of estimation in %

\[
\left(\frac{c_x w - a_{1234_0}}{c_x w}\right)^{-1} \cdot 100 = 0.036 \quad (\%)
\]

\[
\left(\frac{\mu_3 - \frac{a_{1234_1}}{a_{1234_0}}}{\mu_3}\right) \cdot 100 = 3.551 \quad (\%)
\]

\[
\left(\frac{g - a_{1234_2}}{g}\right)^{-1} \cdot 100 = 2.868 \quad (\%)
\]

\[
\left(\frac{\mu_4 - \frac{a_{1234_3}}{a_{1234_2}}}{\mu_4}\right) \cdot 100 = 3.57 \quad (\%)
\]

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


THANK YOU