A Poisson process model for monitoring and surveillance data from wildlife diseases

Volkmar Liebscher

Ernst-Moritz-Arndt University of Greifswald

Farewell workshop for Prof. Dr. Rupert Lasser
Mathematical Analysis and Applications
2013/09/19
An approach to model monitoring and surveillance data of wildlife diseases—Exemplified by Classical Swine Fever in wild boar

N. Stahnke\textsuperscript{a,*}, V. Liebscher\textsuperscript{a}, C. Staubach\textsuperscript{b}, M. Ziller\textsuperscript{c}

\textsuperscript{a} Department of Mathematics and Computer Science, Ernst Moritz Arndt University Greifswald, Walther-Rathenau-Str. 47, 17487 Greifswald, Germany
\textsuperscript{b} Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Institute of Epidemiology, Seestr. 55, 16868 Wusterhausen, Germany
\textsuperscript{c} Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Biomathematics Working Group, 17493 Greifswald-Insel Riems, Germany
Motivation

Data background

Epidemic Models

Parameter estimation

Simulation Results

CSF data

Conclusion
Why?

Combination of

1. point processes
2. classical biomathematical ODE model
3. data analysis
Why?

Combination of

1. point processes
Why?

Combination of

1. point processes
2. classical biomathematical ODE model
Why?

Combination of

1. point processes
2. classical biomathematical ODE model
3. data analysis
Motivation

- past decades: monitoring and surveillance of wildlife diseases increasingly important with regard to health status of humans and livestock
past decades: monitoring and surveillance of wildlife diseases increasingly important with regard to health status of humans and livestock

implementation of an evaluation model for existing systems
Motivation

- past decades: monitoring and surveillance of wildlife diseases increasingly important with regard to health status of humans and livestock
- implementation of an evaluation model for existing systems
- specify guidelines for future improvement
Motivation

- past decades: monitoring and surveillance of wildlife diseases increasingly important with regard to health status of humans and livestock
- implementation of an evaluation model for existing systems
- specify guidelines for future improvement

New approach: use non-homogeneous Poisson processes the intensities of which are driven by an underlying epidemic model
SIR-models

\[
\begin{align*}
\frac{dS(t)}{dt} &= -\beta S(t) I(t) \\
\frac{dI(t)}{dt} &= \beta S(t) I(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t)
\end{align*}
\]

populations sizes by class:

\(S(t)\) susceptible, \(I(t)\) infectious, \(R(t)\) recovered
SIR-models

\[
\frac{dS(t)}{dt} = -\beta S(t)I(t) + b(I(t) + R(t))
\]

\[
\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t) - bI(t)
\]

\[
\frac{dR(t)}{dt} = \gamma I(t) - bR(t)
\]
SIR-models

\[
\frac{dS(t)}{dt} = -\beta S(t) I(t) + b(I(t) + R(t)) \\
\frac{dI(t)}{dt} = \beta S(t) I(t) - \gamma I(t) - bI(t) \\
\frac{dR(t)}{dt} = \gamma I(t) - bR(t)
\]

fit to data by (weighted) least squares
SIR-models

\[
\frac{dS(t)}{dt} = -\beta S(t)I(t) + b(I(t) + R(t))
\]
\[
\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t) - bI(t)
\]
\[
\frac{dR(t)}{dt} = \gamma I(t) - bR(t)
\]

fit to data by (weighted) least squares
we develop a likelihood based method
Classical swine fever (CSF) in wild boar:

- CSF: viral disease of family suidea
- clinical course depends on age structure, health status, population density, virulence of CSFV and practised hunting methods
- must differentiate between epidemic and endemic phase

Assumptions:

1. direct transmission
2. acute clinical course and (lifelong) immunity after infection
3. births of piglets throughout the entire year
4. routinely contact of male and female boars
Sampling

...is done through
- hunting
...is done through
  ➤ hunting
  ➤ other men-caused deaths, mainly car accidents
...is done through
  ▶ hunting
  ▶ other men-caused deaths, mainly car accidents
  ▶ found dead animals
MOSS Data

... are essentially pairs \((t_n, c_n)_{n=1,...}\) where

- \(t_n \in [0, \infty)\) is the hunting time of the \(n\)th animal
... are essentially pairs \((t_n, c_n)_{n=1,\ldots}\) where

- \(t_n \in [0, \infty)\) is the hunting time of the \(n\)th animal
- \(c_n \in \{s, i, r\}\) is the observed epidemic class of the \(n\)th animal
MOSS Data

... are essentially pairs $(t_n, c_n)_{n=1,...}$ where

- $t_n \in [0, \infty)$ is the hunting time of the $n$th animal
- $c_n \in \{s, i, r\}$ is the observed epidemic class of the $n$th animal

NB: we reduced the time information to weeks
The Rôle of Diagnostics
Diagnostics

- every diagnostic test: own sensitivity and specificity
Diagnostics

- every diagnostic test: own sensitivity and specificity
- classification of samples:
  - s = virological and serological negative
  - i = virological positive
  - r = virological negative and serological positive
transition probabilities during diagnostics:

\[ p = \begin{pmatrix} p_{ss} & p_{is} & p_{rs} \\ p_{si} & p_{ii} & p_{ri} \\ p_{sr} & p_{ir} & p_{rr} \end{pmatrix} \]
transition probabilities during diagnostics:

\[
p = \begin{pmatrix}
  Sp_v \cdot Sp_s & (1 - Se_v) \cdot Sp_s & Sp_v \cdot (1 - Se_s) \\
  1 - Sp_v & Se_v & 1 - Sp_v \\
  Sp_v \cdot (1 - Sp_s) & (1 - Se_v) \cdot (1 - Sp_s) & Sp_v \cdot Se_s
\end{pmatrix}
\]
Paradigm

population: \( S \), \( I \), \( R \)

observation: \( \Pi^S \), \( \Pi^I \), \( \Pi^R \)
population: sampling and diagnostics

observation:
Disease Modelling

- underlying dynamics described by classical SIR model (with extensions)
Disease Modelling

- underlying dynamics described by classical SIR model (with extensions)
- assumption of time-constant contact and birth rate
Disease Modelling

- underlying dynamics described by classical SIR model (with extensions)
- assumption of time-constant contact and birth rate
- estimation of population size $N$ infeasible: normalisation required
underlying dynamics described by classical SIR model (with extensions)

assumption of time-constant contact and birth rate

estimation of population size $N$ infeasible: normalisation required

examination of proportions $s$, $i$ and $r$
Disease Modelling

- underlying dynamics described by classical SIR model (with extensions)
- assumption of time-constant contact and birth rate
- estimation of population size $N$ infeasible: normalisation required
- examination of proportions $s$, $i$ and $r$
- most important parameter: basic reproduction rate $R_0$
Example SIR-Model

\[
\begin{align*}
\frac{ds(t)}{dt} &= -\beta_N s(t)i(t) \\
\frac{di(t)}{dt} &= \beta_N s(t)i(t) - \gamma i(t) \\
\frac{dr(t)}{dt} &= \gamma i(t)
\end{align*}
\]

with \( R_0 = \frac{\beta_N}{\gamma} > 0 \)

**Figure:** model with \( \beta_N = 0.3 \) and \( \gamma = 0.1 \)
Model – Real Proportions

in our SIR-models $N(t) = S(t) + I(t) + R(t) = \text{const.}$

so we set $s(t) = \frac{S(t)}{N(t)}$, $i(t) = \frac{i(t)}{N(t)}$, $r(t) = \frac{R(t)}{N(t)}$

\[
\begin{align*}
\frac{ds(t)}{dt} &= -\beta_N s(t)i(t) \\
\frac{di(t)}{dt} &= \beta_N s(t)i(t) - \gamma i(t) \\
\frac{dr(t)}{dt} &= \gamma i(t)
\end{align*}
\]
in our SIR-models $N(t) = S(t) + I(t) + R(t) = \text{const.}$

so we set $s(t) = \frac{S(t)}{N(t)}$, $i(t) = \frac{i(t)}{N(t)}$, $r(t) = \frac{R(t)}{N(t)}$

\[
\begin{align*}
\frac{ds(t)}{dt} &= -\beta_N s(t) i(t) + b(i(t) + r(t)) \\
\frac{di(t)}{dt} &= \beta_N s(t) i(t) - \gamma i(t) - bi(t) \\
\frac{dr(t)}{dt} &= \gamma i(t) - br(t)
\end{align*}
\]
Data \((t_n, c_n)_{n=1,2...}\) follow three independent Poisson processes: for \(c \in \{s, i, r\}\) set

\[
\Pi^c(t) = \#\{n : t_n \leq t, c_n = c\}.
\]

Then \(\Pi^S, \Pi^I, \Pi^R\) are three independent Poisson processes with intensities

\[
\lambda^c(t) = h(t) \cdot \sum_{c' \in \{s, i, r\}, c \neq c'} p_{c', c} c'(t)
\]
1. SIR-model:
   \[ \beta_N > 0, \gamma > 0, \ b > 0, \ \text{yields} \ R_0 = \frac{\beta_N}{\gamma + b} \]
   \[ s(0) > 0, \ i(0) > 0, \ r(0) > 0 \]

2. Poisson process:
   diagnostic matrix \((p_{c', c})_{c', c \in \{s, i, r\}}\)
   hunting rate \(h(t)\)
Model Justification

2 Processes:

1. Thinning:
   hunting takes out a small proportion of the real life population
   Poisson limit theorem: hunting events follow approximately a
   Poisson point process
2 Processes:

1. Thinning:
   hunting takes out a small proportion of the real life population
   Poisson limit theorem: hunting events follow approximately a Poisson point process

2. Marking:
   hunting does not depend on the class identity
diagnostics is applied in the same way for all individuals independently
What to estimate?

1. infection rate $\beta_N > 0$
2. recovering rate $\gamma > 0$
3. birth=death rate $b > 0$
4. initial values $s(0) > 0$, $i(0) > 0$, $r(0) > 0$
5. diagnostic matrix $(p_{c', c})_{c', c \in \{s, i, r\}}$
6. hunting rate $h(t)$
What to estimate?

1. infection rate $\beta_N > 0 \quad \Rightarrow \text{by fitting}$
2. recovering rate $\gamma > 0 \quad \Rightarrow \text{by fitting}$
3. birth\–death rate $b > 0 \quad \Rightarrow \text{by fitting}$
4. initial values $s(0) > 0, \ i(0) > 0, \ r(0) > 0 \quad \Rightarrow \text{by fitting}$
5. diagnostic matrix $(p_{c',c})_{c',c\in\{s,i,r\}} \quad \Rightarrow \text{from specification}$
6. hunting rate $h(t) \quad \Rightarrow \text{from weekly sample sizes}$
Estimators

1. weighted last squares estimator

\[ WLS(\hat{\theta}) = \sum_{c \in \{s, i, r\}} \sum_{w} \frac{(\Delta C_{cw} - \Delta \hat{C}_{cw})^2}{\Delta \hat{C}_{cw}} \rightarrow \min \hat{\theta} \]

\( \Delta C_{cw} \): observed number of animals of class \( c \) in week \( w \)
Estimators

1. weighted last squares estimator

\[ WLS(\hat{\theta}) = \sum_{c \in \{s, i, r\}} \sum_{w} \frac{(\Delta C_{cw} - \Delta \hat{C}_{cw})^2}{\Delta \hat{C}_{cw}} \rightarrow \min \hat{\theta} \]

2. maximum likelihood estimator

\[ MLE(\hat{\theta}) = \sum_{c \in \{s, i, r\}} \sum_{w} (\Delta C_{cw} \ln[\Delta \hat{C}_{cw}] - \Delta \hat{C}_{cw}) \rightarrow \max \hat{\theta} \]
Estimators

1. weighted last squares estimator

\[ WLS(\hat{\theta}) = \sum_{c \in \{s, i, r\}} \sum_{w} \frac{(\Delta C_{cw} - \Delta \hat{C}_{cw})^2}{\Delta \hat{C}_{cw}} \rightarrow \min_{\hat{\theta}} \]

2. maximum likelihood estimator

\[ MLE(\hat{\theta}) = \sum_{c \in \{s, i, r\}} \sum_{w} (\Delta C_{cw} \ln[\Delta \hat{C}_{cw}] - \Delta \hat{C}_{cw}) \rightarrow \max_{\hat{\theta}} \]

\( \Delta C_{cw} \): observed number of animals of class \( c \) in week \( w \)
1. presetting start values for \{s(0), i(0), r(0)\} and for parameters \(\beta_N\) and \(\gamma\) (for the case of a SIR model without demography) followed by an estimation of the parameters only.
Estimation concept

1. presetting start values for \( \{s(0), i(0), r(0)\} \) and for parameters \( \beta_N \) and \( \gamma \) (for the case of a SIR model without demography) followed by an estimation of the parameters only.

2. estimation of new start values only, based on the parameters from step 1.
1. presetting start values for \( \{s(0), i(0), r(0)\} \) and for parameters \( \beta_N \) and \( \gamma \) (for the case of a SIR model without demography) followed by an estimation of the parameters only.

2. estimation of new start values only, based on the parameters from step 1.

3. new joint estimation of all parameters and start values, estimations from step 1 and 2.
Simulation

1. simulation of a SIR model with given parameters and initial values
Simulation

1. simulation of a SIR model with given parameters and initial values
2. setting a constant hunting rate
Simulation

1. simulation of a SIR model with given parameters and initial values
2. setting a constant hunting rate
3. use of a diagnostic matrix with highest sensitivity and specificity of tests
Simulation

1. simulation of a SIR model with given parameters and initial values
2. setting a constant hunting rate
3. use of a diagnostic matrix with highest sensitivity and specificity of tests
4. simulation of the non-homogeneous Poisson process
Simulation

1. simulation of a SIR model with given parameters and initial values
2. setting a constant hunting rate
3. use of a diagnostic matrix with highest sensitivity and specificity of tests
4. simulation of the non-homogeneous Poisson process
5. estimation of parameters by WLS and MLE with a variation of starting parameters for the optimization process
Simulation

1. simulation of a SIR model with given parameters and initial values
2. setting a constant hunting rate
3. use of a diagnostic matrix with highest sensitivity and specificity of tests
4. simulation of the non-homogeneous Poisson process
5. estimation of parameters by WLS and MLE with a variation of starting parameters for the optimization process
6. evaluation of results by WRMSE
Goodness of fit criteria

- weighted root mean squared error

\[ WRMSE(\hat{\theta}) = \sqrt{\frac{1}{k - df} \sum_{c \in \{s, i, r\}} \sum_{w} \frac{(\Delta C_{cw} - \Delta \hat{C}_{cw})^2}{\Delta \hat{C}_{cw}}} \]
Example for simulation studies

**Table:** Results of an estimation in a simulation of a SIR model with demography.

<table>
<thead>
<tr>
<th>parameter</th>
<th>default</th>
<th>start</th>
<th>WLS</th>
<th>( CI_{WLS} )</th>
<th>MLE</th>
<th>( CI_{MLE} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>0.50</td>
<td>0.10</td>
<td>0.583</td>
<td>[0.547,0.618]</td>
<td>0.503</td>
<td>[0.453,0.553]</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>0.25</td>
<td>0.10</td>
<td>0.315</td>
<td>[0.292,0.338]</td>
<td>0.254</td>
<td>[0.227,0.280]</td>
</tr>
<tr>
<td>( b )</td>
<td>0.01</td>
<td>0.10</td>
<td>0.011</td>
<td>[0.011,0.012]</td>
<td>0.010</td>
<td>[0.009,0.011]</td>
</tr>
<tr>
<td>( s(0) )</td>
<td>0.80</td>
<td>0.90</td>
<td>0.844</td>
<td>[0.798,0.889]</td>
<td>0.767</td>
<td>[0.677,0.856]</td>
</tr>
<tr>
<td>( i(0) )</td>
<td>0.20</td>
<td>0.05</td>
<td>0.094</td>
<td>[0.076,0.117]</td>
<td>0.201</td>
<td>[0.152,0.251]</td>
</tr>
<tr>
<td>( r(0) )</td>
<td>0.00</td>
<td>0.05</td>
<td>0.189</td>
<td>[0.140,0.238]</td>
<td>0.046</td>
<td>[0.000,0.114]</td>
</tr>
</tbody>
</table>

- MLE yields better results \( \rightarrow \) use for data fitting
real data: Staubach et al. (2003)

- come from outbreaks in Rhineland-Palatinate → 5 years of surveillance in regions Eifel and Pfalz
- different measures of control were applied
Region Rhineland–Palatinate 1999–2001

Proportion of $d_C$s

Cumulative number of $C$s

Data and MLE estimates for three regions.
Region Pfalz 2002

Proportion $d_C$

Cumulative number $C$

Data and MLE estimations for Pfalz 2002.
Table: Results of the parameter estimation for the region Rhineland-Palatinate (1999 – 2001), for the region Eifel (2001) and the region Pfalz (2002).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.34</td>
<td>0.17</td>
<td>0.02</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.22</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>$b$</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s(0)$</td>
<td>0.83</td>
<td>0.80</td>
<td>0.67</td>
</tr>
<tr>
<td>$i(0)$</td>
<td>0.06</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>$r(0)$</td>
<td>0.06</td>
<td>0.19</td>
<td>0.30</td>
</tr>
<tr>
<td>$WRMSE$</td>
<td>1.67</td>
<td>1.40</td>
<td>1.63</td>
</tr>
<tr>
<td>$WRMSE ; C_{s}$</td>
<td>1.03</td>
<td>0.63</td>
<td>1.47</td>
</tr>
<tr>
<td>$WRMSE ; C_{i}$</td>
<td>2.25</td>
<td>2.11</td>
<td>1.07</td>
</tr>
<tr>
<td>$WRMSE ; C_{r}$</td>
<td>1.51</td>
<td>1.02</td>
<td>2.17</td>
</tr>
<tr>
<td>$R_0$</td>
<td>1.40</td>
<td>2.13</td>
<td>2.00</td>
</tr>
</tbody>
</table>
1. own class $V$ for vaccinated animals: SIRV model
Vaccination

1. own class $V$ for vaccinated animals: SIRV model
2. vaccination rate $\nu$: constant or periodic (pulse vaccination)
Vaccination

1. own class $V$ for vaccinated animals: SIRV model
2. vaccination rate $\nu$: constant or periodic (pulse vaccination)
   - for present data only pulse vaccination is feasible
Vaccination

1. own class $V$ for vaccinated animals: SIRV model
2. vaccination rate $v$: constant or periodic (pulse vaccination)
   - for present data only pulse vaccination is feasible
   - Likelihood ratio test: no significance for model with vaccination
Vaccination

1. own class $V$ for vaccinated animals: SIRV model

2. vaccination rate $v$: constant or periodic (pulse vaccination)
   - for present data only pulse vaccination is feasible
   - Likelihood ratio test: no significance for model with vaccination
   - DIVA (Differentiating Infected from Vaccinated Animals) vaccines allow use of class $V$
     → possibility for more detailed analysis of vaccination trails
Conclusions

- model is able to interpret real data
Conclusions

- model is able to interpret real data
- estimated parameters enable evaluation of practised surveillance
Conclusions

- model is able to interpret real data
- estimated parameters enable evaluation of practised surveillance
- model can be used to investigate surveillance data from different types of wildlife diseases
Future research

- better choice of start value of optimisation
Future research

- better choice of start value of optimisation
- SIRV-models with extensions
Future research

- better choice of start value of optimisation
- SIRV-models with extensions
- use of additional information (age, sex)
Future research

- better choice of start value of optimisation
- SIRV-models with extensions
- use of additional information (age, sex)
- inclusion of time-limited immunity
Future research

- better choice of start value of optimisation
- SIRV-models with extensions
- use of additional information (age, sex)
- inclusion of time-limited immunity
- spatial models
Future research

- better choice of start value of optimisation
- SIRV-models with extensions
- use of additional information (age, sex)
- inclusion of time-limited immunity
- spatial models
- use on more clean data
All the best,
Rupert Lasser!