Wildlife population assessment methods

Takis Besbeas

CAPAM 2015, La Jolla, CA
October 2015
Outline

- IPM in ecology

- IPM aspects
  - Modelling data reflecting population size
  - Goodness of fit (with B.J.T. Morgan)
  - New methods of model-fitting (with R. Langrock)

- Conclusions
Motivation

- In ecology, aspects of the demography of wild animals are estimated through different studies, resulting in multiple data sets.

- Main types of study:
  - individual studies (MRR data)
  - population studies (abundance data)
Individual studies (MRR data)

- These provide information on survival from studying previously marked animals.
- Animals may be observed again alive (recaptures) or dead (recoveries):

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>A</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963</td>
<td>1147</td>
<td>14</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1964</td>
<td>1285</td>
<td>20</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>1106</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>1615</td>
<td>9</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>1618</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>
Population studies (abundance data)

- These also arise in a variety of different forms:

  - **Heron**
  - **Lapwing**
  - **Kaka**
  - **Comma**
Illustrating example I: British herons (Ardea cinerea)

- Mark-recovery and census data

<table>
<thead>
<tr>
<th>Year Ringed</th>
<th>Number Ringed</th>
<th>Year Recovered</th>
<th>'56</th>
<th>'57</th>
<th>'58</th>
<th>...</th>
<th>'97</th>
<th>'98</th>
</tr>
</thead>
<tbody>
<tr>
<td>1955</td>
<td>348</td>
<td></td>
<td>33</td>
<td>5</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1956</td>
<td>141</td>
<td></td>
<td>17</td>
<td>6</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1957</td>
<td>265</td>
<td></td>
<td>31</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>4488</td>
<td></td>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>1997</td>
<td>4339</td>
<td></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

![Graph showing # frost days over years](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAABAA0bl+AAAABwAExcnd2AAAACXBIWXMAAAsAAAABUFmIwDAALiJ2AAGDAAAAACnQ6AAADuB4+PAAAAAHRSTlMAAAAlwEyAACAgEwJxAAAAHh7JwAAAIx1Q8AAAACeZwQQAAAHFh7BotAAAAABJRU5ErkJggg==)
Methods and notation

- Use of capture-recapture methods or state-space models to form MRR likelihood(s) (Lebreton et al, 1992; Gimenez et al, 2007).

- Use of state-space models to form “census” likelihood:

\[ N_{t+1} = T_t N_t + \eta_t \]
\[ y_t = z_t N_t + \epsilon_t \]

- Use of integrated population modelling to estimate parameters.

- \( \phi \) denotes annual survival probability,
  \( \lambda \) denotes reporting probability,
  \( p \) denotes productivity rate.
Relevant state-space models

\[
\begin{pmatrix}
N_1 \\
N_2 \\
\vdots \\
N_{a-1} \\
N_{a+}
\end{pmatrix}_{t+1} =
\begin{pmatrix}
0 & p\phi_1 & \cdots & p\phi_1 & p\phi_1 \\
\phi_2 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & \phi_a & \phi_a
\end{pmatrix}
\begin{pmatrix}
N_1 \\
N_2 \\
\vdots \\
N_{a-1} \\
N_{a+}
\end{pmatrix}_t
+ 
\begin{pmatrix}
\eta_1 \\
\eta_2 \\
\vdots \\
\eta_{a-1} \\
\eta_{a+}
\end{pmatrix}_t
\]

\[y_t = (0 \ 1 \ \cdots \ 1) \begin{pmatrix} N_{1,t} & N_{2,t} & \cdots & N_{a+,t} \end{pmatrix} + \epsilon_t\]

based on Poisson for recruitment and binomial for survival assumptions.
## Model fitting method

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est</td>
<td>SE</td>
<td>Mean</td>
</tr>
<tr>
<td>$\phi_1$ intercept</td>
<td>0.540</td>
<td>0.068</td>
<td>0.543</td>
</tr>
<tr>
<td>$\phi_1$ slope</td>
<td>-0.196</td>
<td>0.059</td>
<td>-0.197</td>
</tr>
<tr>
<td>$\phi_a$ intercept</td>
<td>1.542</td>
<td>0.070</td>
<td>1.55</td>
</tr>
<tr>
<td>$\phi_a$ slope</td>
<td>-0.241</td>
<td>0.039</td>
<td>-0.243</td>
</tr>
<tr>
<td>$\lambda$ intercept</td>
<td>-3.945</td>
<td>0.082</td>
<td>-3.910</td>
</tr>
<tr>
<td>$\lambda$ slope</td>
<td>-0.034</td>
<td>0.004</td>
<td>-0.034</td>
</tr>
<tr>
<td>$p$ intercept</td>
<td>-0.693</td>
<td>0.090</td>
<td>-0.668</td>
</tr>
<tr>
<td>$p$ slope</td>
<td>-0.026</td>
<td>0.004</td>
<td>-0.027</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>159.47</td>
<td>21.87</td>
<td>169.11</td>
</tr>
</tbody>
</table>

- Minor differences but methods differ in ease of application
## Model fitting method

**Computational demand**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est</td>
<td>SE</td>
<td>Mean</td>
</tr>
<tr>
<td>$\phi_1$ intercept</td>
<td>0.540</td>
<td>0.068</td>
<td>0.543</td>
</tr>
<tr>
<td>$\phi_1$ slope</td>
<td>-0.196</td>
<td>0.059</td>
<td>-0.197</td>
</tr>
<tr>
<td>$\phi_a$ intercept</td>
<td>1.542</td>
<td>0.070</td>
<td>1.550</td>
</tr>
<tr>
<td>$\phi_a$ slope</td>
<td>-0.241</td>
<td>0.039</td>
<td>-0.243</td>
</tr>
<tr>
<td>$\lambda$ intercept</td>
<td>-3.945</td>
<td>0.082</td>
<td>-3.910</td>
</tr>
<tr>
<td>$\lambda$ slope</td>
<td>-0.034</td>
<td>0.004</td>
<td>-0.034</td>
</tr>
<tr>
<td>$p$ intercept</td>
<td>-0.693</td>
<td>0.090</td>
<td>-0.668</td>
</tr>
<tr>
<td>$p$ slope</td>
<td>-0.026</td>
<td>0.004</td>
<td>-0.027</td>
</tr>
<tr>
<td>$\sigma$</td>
<td><strong>159.47</strong></td>
<td>21.87</td>
<td><strong>169.11</strong></td>
</tr>
</tbody>
</table>

- Minor differences but methods differ in ease of application
Advantages of integration

- Simultaneous description of all the data
- Estimation of parameters not estimable from separate analyses
- Generally more precise parameter estimators
- Reduction in correlation between estimators
IPM aspects

- Assumption of independence (Besbeas et al, 2009; Abadi et al 2010; Chandler and Clark, 2014)
- Model selection (Besbeas et al, 2015)
- **Modelling data reflecting population size**
  - Pseudo replication
  - Penalised likelihood
- Goodness-of-fit
- New methods of model-fitting
Modelling data reflecting population size

- In many cases, the only type of data available are a time-series of population abundances $y_t$.
  - See, eg, the Global Population Dynamics Database

- A frequently used model for log-abundance is the Gompertz SSM

\[
X_t = a + cX_{t-1} + \eta_t, \quad \eta_t \sim N(0, \tau^2) \\
Y_t = X_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2)
\]
Modelling data reflecting population size

- However practical performance eg wrt partition variance is less than ideal:

Simulation $a=1$, $c=0.3$, $\sigma=\tau = 0.2$, $n =25$
IPM protection

- Recall, eg for two age classes, $\mathbf{N}_t = (N_{1,t}, N_{2+,t})$
  
  \[
  N_{1,t+1} | \mathbf{N}_t \sim \text{Po}(p \phi_1 N_{2+,t})
  \]
  
  \[
  N_{2+,t+1} | \mathbf{N}_t \sim \text{bin}(N_{1,t} + N_{2+,t}, \phi_a)
  \]
  
  \[
  y_t | \mathbf{N}_t \sim \mathcal{N}(N_{2+,t}, \sigma^2)
  \]

- Thus process variances $\tau$ are fully specified in IPM.
IPM breakdowns: fully time-dependent parameters

- Model notation: $\phi_1, \phi_2, \phi_3, \phi_a/\lambda_t/p = \{c,t\}$

  Best $K^{th}$-order time-dependent model:

<table>
<thead>
<tr>
<th>Model</th>
<th>$-\ell$</th>
<th>$\hat{\sigma}$</th>
<th>$\Delta$AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>K=0</td>
<td>cccc/\lambda_t/c</td>
<td>8942.41</td>
<td>651.97</td>
</tr>
<tr>
<td>1</td>
<td>cccc/\lambda_t/t</td>
<td>8781.12</td>
<td>16.95</td>
</tr>
<tr>
<td>2</td>
<td>ccct/\lambda_t/t</td>
<td>8652.30</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>tcct/\lambda_t/t</td>
<td>8600.97</td>
<td>0.00</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>tttt/\lambda_t/t</td>
<td>8539.98</td>
<td>0.00</td>
</tr>
<tr>
<td>vvvv/\lambda_t/t</td>
<td>8695.92</td>
<td>0.00</td>
<td>0</td>
</tr>
</tbody>
</table>

- Several models, incl. best AIC, result in unrealistic findings, eg $\hat{\sigma} = 0$
IPM breakdowns: individual heterogeneity

- Several authors (King et al, 2008; Besbeas et al, 2009) have considered the use of more general process error structures as might result for example from individual heterogeneity.

- We replace the binomial number of survivors

\[ N_{t+1} \sim \text{bin}(n, \phi) \]

with beta-binomial

\[ N_{t+1} \sim \text{betabin}(\text{mean} = n\mu, \text{var} = n\mu(1 - \mu)(1 + \frac{n - 1}{1 + \theta}\theta)) \]

where \( \theta \geq 0 \) is an unknown overdispersion parameter.

- Similarly for recruitment (negative binomial)

\[ N_0 \sim \text{NB}(\text{mean} = pn, \text{variance} = pn(1 + \theta pn)) \]

- For illustration we concentrate on adult survival \( \phi_a \)
Profile log-likelihoods from selected model
Two solutions: a) Pseudo replication

- In non-IPM, various authors (Dennis et al, 2010; Knape et al, 2012) have proposed the use of replicated sampling as a means to improve estimation:

\[ y_{t,k} = X_t + \epsilon_{t,k}, \quad k = 1, \cdots, K \]

- Data from Robert and Casella (2010):
Replicated sampling

- Practical performance is vastly improved, even when $K=2$.

\[ y_{t,k} = X_t + \epsilon_{t,k}, \quad k = 1, \cdots, K \]
Four pseudo replicates

- In IPM, Besbeas and Morgan (2015) propose pseudo replication if no replicates are available:

\[ \tilde{y}_t = \text{random}(y_{t-1}, y_t, y_{t+1}) \]

- Several variants, incl. suggestions by referee.
### Parameter estimate comparison

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No replication</th>
<th>Replication with 3 values</th>
<th>Replication with 5 values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>se</td>
<td>Av. Estimate</td>
</tr>
<tr>
<td>1</td>
<td>-0.171</td>
<td>0.049</td>
<td>-0.174</td>
</tr>
<tr>
<td>2</td>
<td>-0.021</td>
<td>0.004</td>
<td>-0.023</td>
</tr>
<tr>
<td>3</td>
<td>0.411</td>
<td>0.074</td>
<td>0.410</td>
</tr>
<tr>
<td>4</td>
<td>-0.022</td>
<td>0.006</td>
<td>-0.018</td>
</tr>
<tr>
<td>5</td>
<td>0.931</td>
<td>0.105</td>
<td>0.924</td>
</tr>
<tr>
<td>6</td>
<td>-0.023</td>
<td>0.009</td>
<td>-0.019</td>
</tr>
<tr>
<td>7</td>
<td>1.325</td>
<td>0.100</td>
<td>1.341</td>
</tr>
<tr>
<td>8</td>
<td>-0.017</td>
<td>0.005</td>
<td>-0.013</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.053</td>
<td>0.009</td>
<td>0.043</td>
</tr>
<tr>
<td>10</td>
<td>-2.019</td>
<td>0.026</td>
<td>-2.021</td>
</tr>
<tr>
<td>11</td>
<td>-0.831</td>
<td>0.046</td>
<td>-0.833</td>
</tr>
<tr>
<td>12</td>
<td>-0.094</td>
<td>0.092</td>
<td>-0.110</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.000</td>
<td>–</td>
<td>213.22</td>
</tr>
</tbody>
</table>
Two solutions: b) Penalised likelihood

- An alternative approach is to adopt a penalised likelihood approach (Wang & Lindsay, 2005) to bound $\sigma$ away from 0.

- Thus
  \[ \log L_j^\gamma = \log L_j - \gamma h(\sigma) \]
  where $\gamma > 0$ and $h(\sigma)$ becomes large when $\sigma$ is small.

- We adopt $h(\sigma) = -\log \sigma$ and we observe performance by application to the heron data and simulation.
Profile log-likelihoods

- Model $\frac{\nu \nu \nu v}{\lambda_t / t}$

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>Log-Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>-8730</td>
</tr>
<tr>
<td>0.10</td>
<td>-8720</td>
</tr>
<tr>
<td>1.00</td>
<td>-8710</td>
</tr>
<tr>
<td>10.00</td>
<td>-8700</td>
</tr>
<tr>
<td>100.00</td>
<td>-8698</td>
</tr>
</tbody>
</table>

- Model $\frac{\nu \nu \nu v}{\lambda_t / t}$

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>Log-Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>-8725</td>
</tr>
<tr>
<td>0.10</td>
<td>-8715</td>
</tr>
<tr>
<td>1.00</td>
<td>-8705</td>
</tr>
<tr>
<td>10.00</td>
<td>-8700</td>
</tr>
<tr>
<td>100.00</td>
<td>-8695</td>
</tr>
</tbody>
</table>
(Penalised) MLEs

$\gamma = 0.1 \text{ vs } \gamma = 0$

$\gamma = 1 \text{ vs } \gamma = 0$

$\gamma = 10 \text{ vs } \gamma = 0$
Simulation

- Three models (simple to complex)
- Four levels of heterogeneity (S/M/L/XL)
- Observed data sample size
- 500 replications
RMSE ratios: $\text{RMSE}(\gamma=i) / \text{RMSE}(\gamma=0)$

- **Complex model**

  Heterogeneity=S

  Heterogeneity=L

  - $\gamma=0.1$
  - $\gamma=1$
Data weighting and parameter stability

- Recall that observation std $\sigma$ has multiple roles:
  1) Partitions variance in SSM
  2) Weights influence of different data sets
  3) Controls model-selection (anonymous referee)

- Penalising $\sigma$ away from zero offers a natural way to investigate influence of population data and explore sensitivity of model results to different data weightings.
95% CIs for parameters as population data are downweighted
IPM aspects

- Assumption of independence
- Model selection
- Modelling data reflecting population size
  - Pseudo replicate
  - Penalised likelihood
- **Goodness-of-fit**
- New methods of model-fitting
Motivation

- State-space methods are receiving increasing attention in ecology.

- "Constructing diagnostics for judging the goodness of fit of state-space models to data is complex" (Newman et al, 2014, p117).
IPM Goodness-of-fit

- Component-wise but informal:
  - eg comparing observed vs expected m-arrays for MRR components.
  - comparing observed vs fitted trends for census components.

- Bootstrap gof options available but not widely employed.
A new goodness-of-fit procedure

Besbeas and Morgan (2014) propose the use of calibrated simulation to assess component gof:

- Simulate $\theta_i \sim N(\hat{\theta}, \Sigma), i = 1, \cdots, s$
- For each $\theta_i$
  - Work out fitted values $\hat{x}_i$ and simulate new data $x_i \sim f(\cdot | \theta_i)$ for each component model
  - Plot $D(x; \hat{x}_i)$ vs $D(x_i; \hat{x}_i)$ for some measure of discrepancy $D$.
- If the model fits the data well, one would expect approximately half of the points to lie above the diagonal (Johnson, 2004; Fletcher, 2009).

Note:
- Model is fitted only once.
- The approach is akin to Bayesian p-values.
Choice of discrepancy measure

- There is complete freedom in the choice of distance measure
  - cf Millar and Meyer, 2000: \( p \)-value=0.69, 0.27, 0.50 and 0.42.

- MRR / GLM
  - Freeman-Tukey:  
    \[
    D(\mathbf{x}; \mathbf{e}) = \sum_i (\sqrt{x_i} - \sqrt{e_i})^2
    \]

- Time series \( \mathbf{x}_t \)
  - MAPE  
    \[
    D(\mathbf{x}; \hat{\mathbf{x}}) = \frac{100}{n} \sum_t |(x_t - \hat{x}_t)/x_t|
    \]
  - MPE  
    \[
    D(\mathbf{x}; \hat{\mathbf{x}}) = 100 \max\{(x_t - \hat{x}_t)/x_t\} \]
Illustrating example II: lapwings (Vanellus vanellus)

- Mark-recovery and index of abundance data

<table>
<thead>
<tr>
<th>Year Ringed</th>
<th>Number Ringed</th>
<th>Year Recovered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>'56 '57 '58 '97 '98</td>
</tr>
<tr>
<td>1955</td>
<td>348</td>
<td>33 5 0 ... 0 0</td>
</tr>
<tr>
<td>1956</td>
<td>141</td>
<td>17 6 ... 0 0</td>
</tr>
<tr>
<td>1957</td>
<td>265</td>
<td>31 ... 0 0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>... ... ...</td>
</tr>
<tr>
<td>1996</td>
<td>4488</td>
<td>... 16 0</td>
</tr>
<tr>
<td>1997</td>
<td>4339</td>
<td>18</td>
</tr>
</tbody>
</table>

![Graph showing index and weather over years]
Motivating example: IPM

Model $\phi_1(w_t), \phi_a(w_t)/\lambda(year)/p(year)$, following model-selection:

$p_c = 0.39$

$\text{Freeman-Tukey}$

$p_c = 0.72$

$\text{MAPE}$

$p_c = 0.52$
Simulation comparison with bootstrap: MRR component

- Calibrated simulation vs bootstrap p-values when true model (vv/v) is fitted by:

  - vv/v
  - vc/v
  - vv/c
  - cc/c
Calibration of test statistic

- 100 simulations from selected model, fitted 3-ways
Modelling site by year survey data

- In practice, census data are rare.
- Often we collect spatio-temporal observations, $c_{it}$ (eg BBS, CBC).
- We typically model the $c_{it}$ using a Poisson-based linear model:

  \[ c_{it} \sim Po(\mu_{it}) \]
  \[ \log(\mu_{it}) = s_i + u_t \]
Modelling site by year survey data

- In practice, census data are rare.
- Often we collect spatio-temporal observations, $c_{it}$ (eg BBS, CBC).
- We typically model the $c_{it}$ using a Poisson-based linear model:

  \[ c_{it} \sim \text{Po}(\mu_{it}) \]
  \[ \log(\mu_{it}) = s_i + u_t \]

- “Census” likelihood can be formed in two or one stages
  - Based on a derived index of abundance, $y_t \propto e^{u_t}$
  - Building a dynamic model directly on the $u_t$
Pros and Cons of the two approaches

- Details in
  - Mauder (2001) — fisheries
  - Besbeas and Freeman (2006) — ecology

- Similar performance trading off speed vs efficiency

- However one-stage method offers greater flexibility not just for modelling counts but for incorporating data of an entirely different nature.
  - For example presence/absence data (Freeman and Besbeas, 2012).
Illustrating example II revisited

- Simulated calibration gof for Poisson distribution of $C_{it}$
Site by year survey data: model selection

<table>
<thead>
<tr>
<th>Model</th>
<th>$-\ell$</th>
<th>$np$</th>
<th>$\Delta AIC$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson($\mu_{it}$)</td>
<td>3629.8</td>
<td>266</td>
<td>246.8</td>
</tr>
<tr>
<td>ZIP($\mu_{it}, \pi$)</td>
<td>3629.8</td>
<td>267</td>
<td>248.8</td>
</tr>
<tr>
<td>Negbin($\mu_{it}, \theta$)</td>
<td>3594.0</td>
<td>267</td>
<td>177.2</td>
</tr>
<tr>
<td>Poisson-Tweedie($\mu_{it}, D, a$)</td>
<td>3589.3</td>
<td>257</td>
<td>167.8</td>
</tr>
<tr>
<td>ZIP($\mu_{it}, \pi_i$)</td>
<td>3520.4</td>
<td>497</td>
<td>490.0</td>
</tr>
<tr>
<td>Hurdle Po($\mu_{it}, \pi_i$)</td>
<td>3478.4</td>
<td>497</td>
<td>406.0</td>
</tr>
<tr>
<td>Negbin($\mu_{it}, \theta_i$)</td>
<td>3426.3</td>
<td>497</td>
<td>301.8</td>
</tr>
<tr>
<td>Negbin($\mu_{it}, \theta(CV)$)</td>
<td>3504.4</td>
<td>268</td>
<td>0</td>
</tr>
</tbody>
</table>
Individual site $\theta_t$ vs $\theta(CV)$

- We can then fit the (integrated) model exactly as we do for Poisson data.
From deterministic to stochastic transitions

- Recall

\[ c_{it} \sim Po(\mu_{it}) \]

\[ \log(\mu_{it}) = s_i + u_t \]

\[ e^{u_t} = (\phi_a e^{u_{t-1}} + p\phi_1 \phi_a e^{u_{t-2}}) \]

- We introduce lognormal “process error”

\[ e^{u_t} = (\phi_a e^{u_{t-1}} + p\phi_1 \phi_a e^{u_{t-2}}) \varepsilon \]

and adopt a Bayesian approach.
Deterministic vs stochastic transitions

- Comparison of performance between “best” and “worst” (constant) models from analysis of index:

«Every time I think I know what’s going on, suddenly there’s another layer of complications. I just want this damn thing solved.»

John Scalzi
IPM aspects

- Assumption of independence
- Model selection
- Modelling data reflecting population size
  - Single replicate
  - Multiple replicate
- Goodness-of-fit
- New methods of model-fitting
Hidden Markov Models - overview

- Flexible and mathematically tractable time series model.
- Two discrete-time stochastic processes: observed and hidden.
- Hidden process is an N-state Markov chain.
- Observed process determined by underlying hidden state.

Applications, inter alia, in speech recognition, biology, medicine, psychology, economics and finance.
SSM estimation via hidden Markov models

- HMMs have the same two-process structure as SSMs.
- In SSMs: $S_t$ continuous-valued.
- Discretizing $S_t$ yields approximation by HMM.
- Benefit: HMM methodology becomes applicable.
Illustration of an HMM-based population model

- $N_t$: true (unknown) number of individuals at time $t$ (the states).
- $N_t = N_{1t} + N_{at}$, where $N_{1t} \sim Po(pN_t)$ and $N_{at} \sim bin(N_t, \phi)$.
- Specifying some upper bound for $N_t$, this is a Markov chain (with a lot of states, but only two parameters).
- Observations, conditional on states: $X_t|N_t = j \sim bin(j, r_t)$
- HMM likelihood structure as usual, using a forward algorithm.
- If the upper bound for $N_t$ is in the thousands, computer memory can be an issue
  - consider coarser state-space, e.g. $[0, 10), [10, 20), ..., [4990, 5000)$ instead of $0, 4999$.
- Integrated Population model can be formulated and fitted as usual.
Reference example: lapwings

- Model $\phi_1, \phi_a/\lambda/p$ (Homogeneous Markov chain)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>HMM</th>
<th>KF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w=100</td>
<td>w=40</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.427</td>
<td>0.419</td>
</tr>
<tr>
<td>$\phi_a$</td>
<td>1.448</td>
<td>1.401</td>
</tr>
<tr>
<td>$\lambda$ intercept</td>
<td>-4.678</td>
<td>-4.680</td>
</tr>
<tr>
<td>$p$ intercept</td>
<td>-0.978</td>
<td>-0.885</td>
</tr>
<tr>
<td>$\log \sigma^2$</td>
<td>10.977</td>
<td>11.693</td>
</tr>
</tbody>
</table>
Reference example: lapwings

- Model \( \phi_1, \phi_a, \lambda, p \) (Homogeneous Markov chain)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>HMM</th>
<th>KF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w=100</td>
<td>w=40</td>
</tr>
<tr>
<td>( \phi_1 )</td>
<td>0.427</td>
<td>0.419</td>
</tr>
<tr>
<td>( \phi_a )</td>
<td>1.448</td>
<td>1.401</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>-4.678</td>
<td>-4.680</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.978</td>
<td>-0.885</td>
</tr>
<tr>
<td>( \log \sigma^2 )</td>
<td>10.977</td>
<td>11.693</td>
</tr>
</tbody>
</table>

Max % RAM | 0.3 | 0.5 | 3.7 | 34.1 | 61.2 |
Conclusions

- **Modelling data reflecting population size**
  - Replication has advantages and permits more sophisticated models to be fitted.
  - Penalising $\sigma$ natural way to investigate sensitivity of model results to different data weightings.

- **Goodness-of-fit**
  - For integrated models, gof preferably component-wise
  - Different distance measures highlight different aspects of fit.
  - Calibrated simulation appears to have reasonable potential.

- **Model-fitting**
  - HMMs provide a promising alternative.
  - Computer memory can be an issue but more development to follow.
References

Thank you
10% MRR
Simulation comparison with bootstrap: GLM
Sampling distributions of parameter estimates