Linking metapopulation modelling and Information Theory for area-wide pest management

ANTONELLA BODINI
CNR-IMATI

GIANNI GILIOLI
UNIVERSITÀ DI BRESCIA
Non statistical terms ...

- **Metapopulation**: set of local populations within some larger area, where typically migration from one local population to at least some other patches is possible (Hanski and Simberloff, 1997).
Non statistical terms ...

- **Metapopulation**: set of local populations within some larger area, where typically migration from one local population to at least some other patches is possible (Hanski and Simberloff, 1997).

- **Pest**: organisms (rats, insects, etc.) which may cause illness or damage or consume food crops and other materials important to humans. An organism that is considered a nuisance to man, most usually having pathogenic properties (Biology-Online.org dictionary)
**Non statistical terms ...**

- **Metapopulation**: Set of local populations within some larger area, where typically migration from one local population to at least some other patches is possible (Hanski and Simberloff, 1997).

- **Pest**: Organisms (rats, insects, etc.) which may cause illness or damage or consume food crops and other materials important to humans. An organism that is considered a nuisance to man, most usually having pathogenic properties (Biology-Online.org dictionary)

- **Area-wide pest management**: Management of localized populations is the conventional or most widely used strategy, wherein individual producers, other operators and households practice independent pest control. However, since individual producers or households are not capable of adequately meeting the challenge of certain **very mobile and dangerous pests**, the area-wide pest management strategy was developed.
Quantitative evaluation of management strategies

Metapopulation models should be embedded in a decision-making framework to give managers the capability of ranking alternative decisions (Westphal et al., 2003). This means that the objectives of the management should be explicitly and clearly stated in terms of metapopulation model variables (Possingham et al., 2001).
Optimization

Stochastic dynamic programming (SDP) has been recently applied in pest management, coupled with a spatially implicit metapopulation model, e.g. for invasive species control optimization (Bogich and Shea, 2008), or for biological control release strategies optimization (Shea and Possingham, 2000). However, SDP is computationally complex and its applicability limited to small metapopulations (Nicol and Chadès, 2011). Borrowing from epidemiology, a susceptible-infected-susceptible (SIS) model and a finite Markov decision process have been proposed to manage diseases, pest or endangered species in small (<25 nodes) network motifs (Chadès et al., 2011).
The idea

- **Spatially explicit** stochastic processes to predict metapopulation dynamics under the effects of a given strategy

- **Kullback-Leibler divergence** to “compare predictions”
The idea

- **Spatially explicit** stochastic processes to **predict** metapopulation dynamics under the effects of a given strategy

- **Kullback-Leibler divergence** to “compare predictions”

- The best: optimization
The idea

- **Spatially explicit** stochastic processes to **predict** metapopulation dynamics under the effects of a given strategy

- **Kullback-Leibler divergence** to “compare predictions”

- The best: optimization NON LINEAR PROBLEM
The idea

- **Spatially explicit** stochastic processes to predict metapopulation dynamics under the effects of a given strategy

- **Kullback-Leibler divergence** to “compare predictions”

- The best: optimization

- In practice: evaluation of a **finite set** of options
The idea

- The **Incidence Function Model** (Hanski 1994) is the only one spatially explicit metapopulation model in the literature. It has been used to predict metapopulation dynamics in terms of presence/absence of the species.

- The **KL divergence** has been introduced (Gilioli et al. 2008) to evaluate the strategies effects at time $T$ in terms of divergence of the predicted dynamic at time $T$ from the total extinction.
The idea in formula

- The **IFM**.
  Multivariate Markov chain: patches either empty ○ or occupied ●
The idea in formula

- **The IFM.**
  Multivariate Markov chain: patches either empty (○) or occupied (●)
The idea in formula

The **IFM.**

Multivariate Markov chain: patches either empty 🔄 or occupied ⬇️

\[ E_i = P(X_i(t) = 0 | X_i(t - 1) = 1, X_{-i}(t - 1)) = \left( \frac{A_i}{A_0} \right)^x \]
The idea in formula

- **The IFM.**

  Multivariate Markov chain: patches either empty ∅ or occupied ●

\[ E_i = P(X_i(t) = 0|X_i(t-1) = 1, X_{-i}(t-1)) = \left(\frac{A_i}{A_0}\right)^x \]
The idea in formula

- **The IFM.**
  
  Multivariate Markov chain: patches are either empty (circle) or occupied (dot).

\[
E_i = P(X_i(t) = 0 | X_i(t-1) = 1, X_{-i}(t-1)) = \left(\frac{A_i}{A_0}\right)^x
\]

\[
C_i(t) = P(X_i(t) = 1 | X_i(t-1) = 0, X_{-i}(t-1)) = \frac{\Delta_i^2(t-1)}{\Delta_i^2(t-1) + y^2}
\]

\[
\Delta_i(t-1) = \sum_{j=1, j\neq i}^n X_j(t-1) A_j e^{-\alpha d_{ij}} \quad \text{connectivity}
\]

For insect pest, \( t \) means generation.
The idea by graphics

- The KL for strategy evaluation. **Purely spatial:**

CONTROL STRATEGY  dynamic simulation

DATA  KL
The idea by graphics

- The **KL** for strategy evaluation. **Purely spatial:**

  DATA  \[\xrightarrow{\text{CONTROL STRATEGY}}\]  dynamic simulation  \[\xrightarrow{\text{KL}}\]
The idea by graphics

- The KL for strategy evaluation. **Purely spatial:**

  CONTROL STRATEGY  dynamic simulation

DATA  KL
The idea by graphics

- The KL for strategy evaluation. **Purely spatial:**

  CONTROL STRATEGY \[\xrightarrow{\text{dynamic simulation}}\] KL

**IFM**

DATA

1
The idea by graphics

- The KL for strategy evaluation. **Purely spatial:**

  CONTROL STRATEGY → dynamic simulation

  DATA → IFM → KL
The idea by graphics

- The **KL** for strategy evaluation. **Spatio-temporal:**
The idea in formula and simulations

- The **KL of** $P(X_T)$ from Dirac measure on $0_n$: $- \ln P(X_T=0)$
  - The lower the KL the better the strategy for *pest control*
  - The higher the KL the better the strategy for *conservation*

**Simulations** to obtain the distribution of $X_T$:

$$P(X_T=0) = P(X_T=0| X_{T-1}=1) P(X_{T-1}=1)$$

$$+ \sum_{s \neq 1} P(X_T=0| X_{T-1}=s) P(X_{T-1}=s)$$
The idea in formula and simulations

- The **KL of** \( P(X_T) \) from Dirac measure on 0\(_n\): \(- \ln P(X_T=0)\)
  - The lower the KL the better the strategy for *pest control*
  - The higher the KL the better the strategy for *conservation*

- **Simulations** to obtain the distribution of \( X_T \):

\[
P(X_T=0) = P(X_T=0 | X_{T-1}=1) P(X_{T-1}=1)
\]

*explicit:* \( \Pi_i (E_i \text{ or } C_i) \)

\[
+ \sum_{s \neq 1} P(X_T=0 | X_{T-1}=s) P(X_{T-1}=s)
\]


The idea in formula and simulations

- The **KL of** $P(X_T)$ from Dirac measure on $0_n$: $- \ln P(X_T=0)$
  - The lower the KL the better the strategy for *pest control*
  - The higher the KL the better the strategy for *conservation*

- **Simulations** to obtain the distribution of $X_T$:

  $$P(X_T = 0) = P(X_T = 0 | X_{T-1} = 1) P(X_{T-1} = 1)$$

  $$+ \sum_{s \neq 1} P(X_T = 0 | X_{T-1} = s) P(X_{T-1} = s)$$

  *simulated*
The idea in formula and simulations

- The **KL of** \( P(X_T) \) from Dirac measure on \( 0_n \): \( - \ln P(X_T=0) \)
  - The lower the KL the better the strategy for *pest control*
  - The higher the KL the better the strategy for *conservation*

- **Simulations** to obtain the distribution of \( X_T \):

\[
P(X_T = 0) = P(X_T = 0| X_{T-1} = 1) P(X_{T-1} = 1) \text{ explicit} + \text{simulated}
\]

\[
+ \sum_{s \neq 1} P(X_T = 0| X_{T-1} = s) P(X_{T-1} = s)
\]
(Amphibians) **Conservation**

Choice between 2 possible sets of new ponds.

- Only 1 year of data
- Equilibrium assumption:

\[ KL(S) = - \sum_{i=1}^{n} \ln [1 - J_i(S)] \]

measures the divergence of the stationary distribution coming from the strategy \( S, \otimes J_i(S) \), from the total extinction.

Gilioli et al. 2008
Pest control

Gilioli et al. 2013

Pine processionary moth.
Second Application

**Pest control**

SPATIAL ANALYSIS to verify the IFM applicability to this moth.

*Gilioli et al. 2013*
Pest control

SPATIAL STRATEGIES
(a) scattered sites
(b) close sites
(c) “in line” sites.

Three levels of intervention: 15%, 30% and 50% of the total area (low, medium and high intervention level).

SPATIAL ANALYSIS to verify the IFM applicability to this moth.

Pine processionary moth.
Second Application

Pest control

SPATIAL STRATEGIES
(a) scattered sites
(b) close sites
(c) “in line” sites.

Three levels of intervention: 15%, 30% and 50% of the total area (low, medium and high intervention level).

SPATIAL ANALYSIS to verify the IFM applicability to this moth.

SPATIO-TEMPORAL STRATEGIES for high intervention level.

Pine processionary moth.
Pest control

SPATIAL STRATEGIES
(a) scattered sites  
(b) close sites  
(c) “in line” sites.

Three levels of intervention: 15%, 30% and 50% of the total area (low, medium and high intervention level).

SPATIAL ANALYSIS to verify the IFM applicability to this moth.

SPATIO-TEMPORAL STRATEGIES for high intervention level.

Combination of both: distributed effort for high intervention level.

Gilioli et al. 2013

Pine processional moth.
Testing the idea
Testing the idea

Need to

1. understand the performances of the KL in more general (not “linear”!!) situations
   - Different habitat configurations
   - Different types of strategies

2. provide easy interpretation for practitioners
   - Comparison of KL values to non probabilistic indexes
More general habitat configurations

Square of side 50 Km
100 patches

occupied
empty

IFM: $x = 0.15, y = 0.001$
More general habitat configurations

Square of side 50 Km
100 patches

- occupied
- empty

IFM: $x = 0.15, y = 0.001$
More general habitat configurations

Square of side 50 Km
100 patches

occupied
empty

IFM: $x = 0.15$, $y = 0.001$
More general habitat configurations

Square of side 50 Km
100 patches

IFM: $x = 0.15$, $y = 0.001$
More general habitat configurations

Square of side 50 Km
100 patches

- occupied
- empty

IFM: $x = 0.15$, $y = 0.001$
More general habitat configurations & strategies

Square of side 50 Km
100 patches

scattered strategy

- occupied
- empty
- to be treated
More general habitat configurations & strategies

Square of side 50 Km
100 patches

organized strategy

- occupied
- empty
- to be treated
More general habitat configurations & strategies

Square of side 50 Km
100 patches

occupied

empty
to be treated

peripheral-organized strategy
General results: purely spatial

data: 2013, KL 2016

$S_0 = \text{do nothing strategy}$

<table>
<thead>
<tr>
<th>Case</th>
<th>KL</th>
<th>$S_0$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>True param.</td>
<td>141.9</td>
<td>141.0</td>
<td>139.7</td>
<td>140.3</td>
<td>135.4</td>
<td>136.6</td>
<td>133.1</td>
<td></td>
</tr>
<tr>
<td>Estimated param.</td>
<td>146.2</td>
<td>146.1</td>
<td>144.0</td>
<td>141.1</td>
<td>139.6</td>
<td>140.2</td>
<td>137.5</td>
<td></td>
</tr>
<tr>
<td>T.A. (%)</td>
<td>--</td>
<td>23.3</td>
<td>21.9</td>
<td>21.2</td>
<td>22.8</td>
<td>22.6</td>
<td>23.1</td>
<td></td>
</tr>
</tbody>
</table>

---

The table above presents the results of a comparison between the true and estimated parameters for different cases. The values within the table represent the outcomes for each strategy, with $S_0$ indicating the no-action strategy. The graph on the right side visualizes the distribution of data points, with distinct colors representing different cases or strategies.
General results: purely spatial

data: 2013, KL 2016

$$S_0 = \text{do nothing strategy}$$

<table>
<thead>
<tr>
<th>Case</th>
<th>KL</th>
<th>Strategies</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>True param.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S_0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>133.1</td>
</tr>
<tr>
<td>True</td>
<td>141.9</td>
<td>141.0</td>
<td>139.7</td>
<td>140.3</td>
<td>135.4</td>
<td>136.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>param.</td>
<td></td>
<td>146.2</td>
<td>146.1</td>
<td>144.0</td>
<td>141.1</td>
<td>139.6</td>
<td>140.2</td>
<td>137.5</td>
</tr>
<tr>
<td>Estimated</td>
<td></td>
<td></td>
<td>23.3</td>
<td>21.9</td>
<td>21.2</td>
<td>22.8</td>
<td>22.6</td>
<td>23.1</td>
</tr>
<tr>
<td>param.</td>
<td>T.A. (%)</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>True</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connectivity</td>
<td>S_0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>True param.</td>
<td>0.212</td>
<td>0.201</td>
<td>0.145</td>
<td>0.177</td>
<td>0.110</td>
<td>0.117</td>
<td>0.087</td>
<td></td>
</tr>
<tr>
<td>Estimated param.</td>
<td>0.218</td>
<td>0.205</td>
<td>0.150</td>
<td>0.182</td>
<td>0.112</td>
<td>0.119</td>
<td>0.088</td>
<td></td>
</tr>
</tbody>
</table>

![Graph showing data distribution]
### General results, contd.

**KL**

<table>
<thead>
<tr>
<th>Strategies</th>
<th>True param.</th>
<th>Estimated param.</th>
<th>T.A. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_0$</td>
<td>143.7</td>
<td>121.0</td>
<td>--</td>
</tr>
<tr>
<td>1</td>
<td>142.1</td>
<td>119.1</td>
<td>24.8</td>
</tr>
<tr>
<td>2</td>
<td>175.9</td>
<td>152.7</td>
<td>21.3</td>
</tr>
<tr>
<td>3</td>
<td>132.9</td>
<td>115.9</td>
<td>21.0</td>
</tr>
<tr>
<td>4</td>
<td>137.1</td>
<td>113.6</td>
<td>20.6</td>
</tr>
<tr>
<td>5</td>
<td>147.7</td>
<td>124.9</td>
<td>27.0</td>
</tr>
</tbody>
</table>

$S_0 = \text{do nothing strategy}$
General results, contd.

<table>
<thead>
<tr>
<th>KL</th>
<th>Strategies</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S₀</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>True param.</td>
<td>143.7</td>
<td>142.1</td>
<td>175.9</td>
<td>132.9</td>
<td>137.1</td>
<td>147.7</td>
</tr>
<tr>
<td>Estimated</td>
<td>121.0</td>
<td>119.1</td>
<td>152.7</td>
<td>115.9</td>
<td>113.6</td>
<td>124.9</td>
</tr>
<tr>
<td>T.A. (%)</td>
<td>--</td>
<td>24.8</td>
<td>21.3</td>
<td>21.0</td>
<td>20.6</td>
<td>27.0</td>
</tr>
</tbody>
</table>

\[ S₀ = \text{do nothing strategy} \]
### General results, contd.

<table>
<thead>
<tr>
<th>Case</th>
<th>Strategies</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>KL</td>
<td>True param.</td>
<td>143.1</td>
<td>142.3</td>
<td>141.2</td>
<td>156.3</td>
<td>155.7</td>
<td>140.1</td>
<td>140.0</td>
<td>143.6</td>
<td>140.7</td>
<td>138.9</td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>Estimated param.</td>
<td>131.5</td>
<td>132.8</td>
<td>134.1</td>
<td>131.6</td>
<td>159.1</td>
<td>145.6</td>
<td>135.7</td>
<td>137.8</td>
<td>137.6</td>
<td>136.3</td>
<td>129.5</td>
</tr>
<tr>
<td>T.A. (%)</td>
<td>--</td>
<td>22.5</td>
<td>22.1</td>
<td>21.6</td>
<td>16.2</td>
<td>20.8</td>
<td>16.1</td>
<td>23.1</td>
<td>14.0</td>
<td>20.2</td>
<td>22.1</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The values in red indicate the best strategies.
### General results, contd.

<table>
<thead>
<tr>
<th>KL</th>
<th>Strategies</th>
<th>True param.</th>
<th>Estimated param.</th>
<th>T.A. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>So</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>True param.</td>
<td>143.1</td>
<td>142.3</td>
<td>142.3</td>
<td>141.2</td>
</tr>
<tr>
<td>Estimated param.</td>
<td>131.5</td>
<td><strong>132.8</strong></td>
<td>134.1</td>
<td>131.6</td>
</tr>
<tr>
<td>T.A. (%)</td>
<td></td>
<td>--</td>
<td>22.5</td>
<td>22.1</td>
</tr>
</tbody>
</table>

[Diagram showing scatter plot with labeled axes and points]

[Question mark image]
STATISTICAL ISSUES:

- Are 100,000 to few simulations?
  - 200,000 do not change the results: higher order?
STATISTICAL ISSUES:

- Are 100,000 to few simulations?
  - 200,000 do not change the results: higher order?
- Should we use a simulation method different from a paper-pencil method?

\[
P(X_T = 0) = P(X_T = 0 | X_{T-1} = 1) P(X_{T-1} = 1) \\
+ \sum_{s \neq 1} P(X_T = 0 | X_{T-1} = s) P(X_{T-1} = s)
\]
STATISTICAL ISSUES:

- Are 100,000 to few simulations? 200,000 do not change the results: higher order?
- Should we use a simulation method different from a paper-pencil method?
- Is the estimated KL value too sensitive to the representation?

\[
P(X_T = 0) = P(X_T = 0 | X_{T-1} = 1) P(X_{T-1} = 1) + \sum_{s \neq 1} P(X_T = 0 | X_{T-1} = s) P(X_{T-1} = s)
\]
STATISTICAL ISSUES:

- Are 100,000 to few simulations?
  - 200,000 do not change the results: higher order?
- Should we use a simulation method different from a paper-pencil method?
- Is the estimated $KL$ value too sensitive to the representation
  
  $$P(X_T = 0) = P(X_T = 0| X_{T-1} = 1) P(X_{T-1} = 1)$$
  
  $$+ \sum_{s \neq 1} P(X_T = 0| X_{T-1} = s) P(X_{T-1} = s)$$

- Is the simulated model a “biased model”?
  
  $$C_i(t) = \frac{\Delta_i^2(t-1)}{\Delta_i^2(t-1) + y^2} \approx 1 \quad as \quad y^2 \approx 0 \text{ at any time } t$$
How come?

“ECOLOGICAL” ISSUES

- Are strategies really inadequate?
How come?

“ECOLOGICAL” ISSUES
- Are strategies really inadequate?

STATISTICAL & ECOLOGICAL ISSUE
- Is the strategy effect representation adequate?
How come?

“ECOLOGICAL” ISSUES
- Are strategies really inadequate?

STATISTICAL & ECOLOGICAL ISSUE
- Is the strategy effect representation adequate?
- Is the IFM a “good” model?
First possible solution... (English translation...)}
Second possible solution...

- Different representation of strategy effect: colonization reduction

\[ r < 1 \]

IFM

Colonization reduction: \( r < 1 \)

\[ rC_i \]
A few, partial answers.

- **A good case:** extended treatment (59% of total area)

![Graph with data points and labels](image-url)
Increasing treatment effect (i.e., increasing colonization reduction)

- Decreasing $KL$ (i.e. increasing probability extinction)
Conclusions

- After conservation and control we were interested in applying this approach to invasive species, but...
Conclusions

- After conservation and control we were interested in applying this approach to invasive species, but...
- ...we have a lot of matters to understand about the KL and the best way, if any, to use it.
  First of all, we need to understand the true advantage, if any, of using the KL!
Conclusions

- After conservation and control we were interested in applying this approach to **invasive species**, **but**...
- ...we have a lot of matters to understand about the KL and the best way, if any, to use it.
  First of all, we need to understand the **true advantage**, if any, of using the KL!

If our researches will find sound and positive answers, I’ll continue the novel the next SIS (**forewarned is forearmed**...)
After conservation and control we were interested in applying this approach to invasive species, but...

...we have a lot of matters to understand about the KL and the best way, if any, to use it. First of all, we need to understand the true advantage, if any, of using the KL!

If our researches will find sound and positive answers, I’ll continue the novel the next SIS (forewarned is forearmed...)

Thank you very much for your attention and even more for your suggestions!