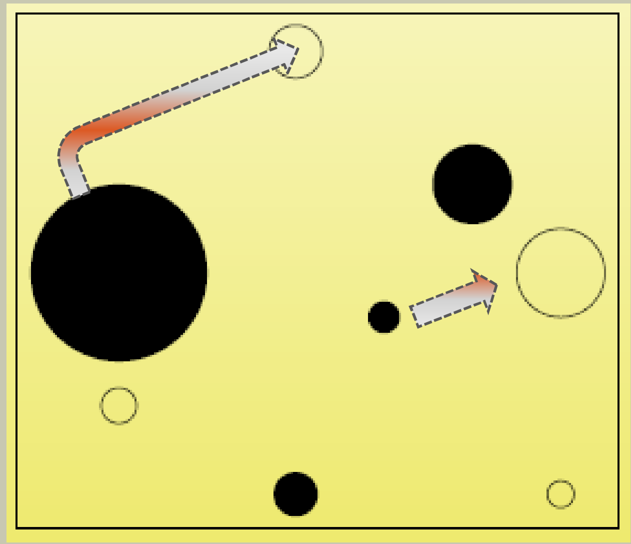
Linking metapopulation modelling and Information Theory for area-wide pest management

ANTONELLA BODINI
CNR-IMATI

GIANNI GILIOLI
UNIVERSITÀ DI BRESCIA

Non statistical terms …

- **Metapopulation** :[*s*]et 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** :[*s*]et 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** :[*s*]et 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.

… & statistical problems

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).

… & statistical problems

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 a given strategy

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

The idea

- **Spatially explicit** stochastic processes to **predict** metapopulation dynamics under the effects 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 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 a given strategy
- **Kullback-Leibler divergence** to "compare predictions"
- The best: optimization
- In practice: evaluation of a **finite set** of options

The idea

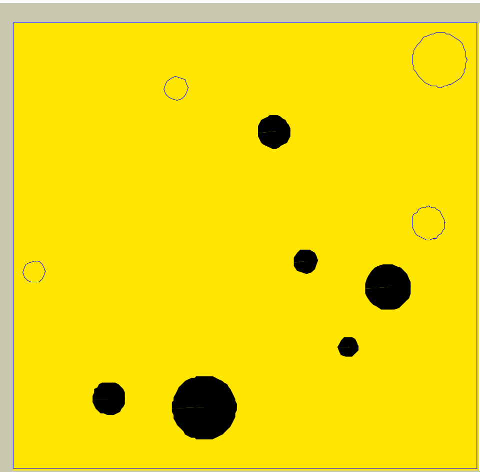
- The **Incidence Function Model** (Hanski 1994) is the only one spatially explicit metapopulation model in the literature. It has been used to predict metapopulation dynamics in terms of *presence/absence* of the species.

- 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.**

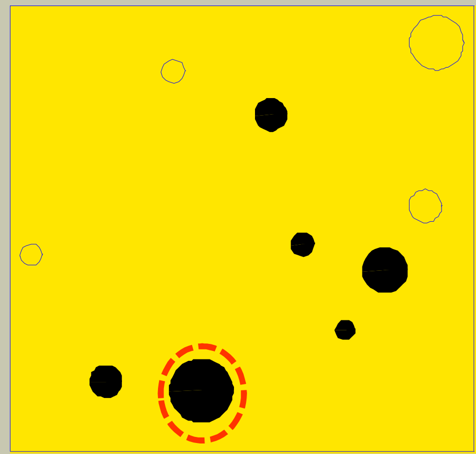
  Multivariate Markov chain: patches either empty or occupied



The idea in formula

- The **IFM.**

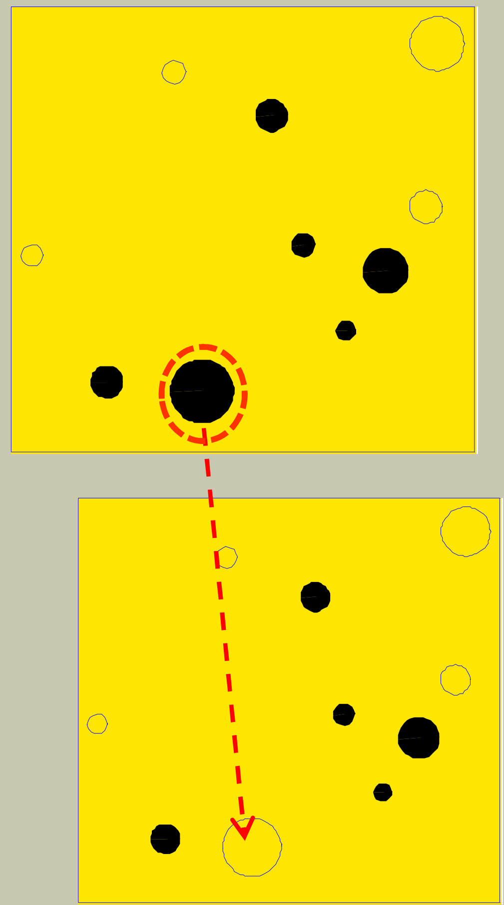
  Multivariate Markov chain: patches either empty or occupied



The idea in formula

- The **IFM.**

  Multivariate Markov chain: patches either empty or occupied



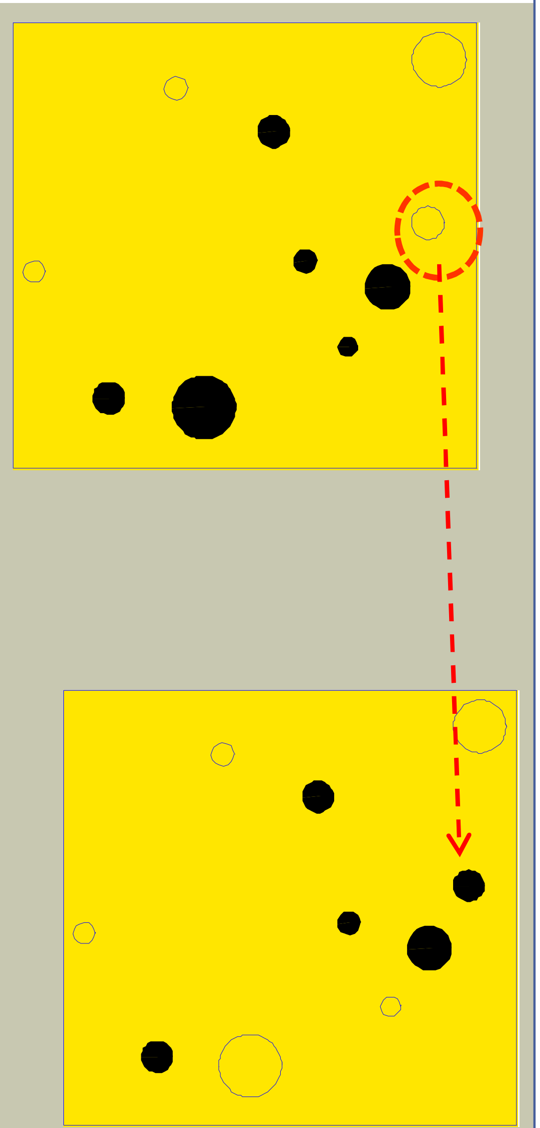
$$E_i = P(X_i(t) = 0 \mid X_i(t-1) = 1, \boldsymbol{X}_{-i}(t-1)) = \left(\frac{A_i}{A_0}\right)^x$$

The idea in formula

- The **IFM.**

  Multivariate Markov chain: patches either empty ○ or occupied ●

$$E_i = P(X_i(t) = 0 \mid X_i(t-1) = 1, \boldsymbol{X}_{-i}(t-1)) = \left(\frac{A_i}{A_0}\right)^x$$



The idea in formula

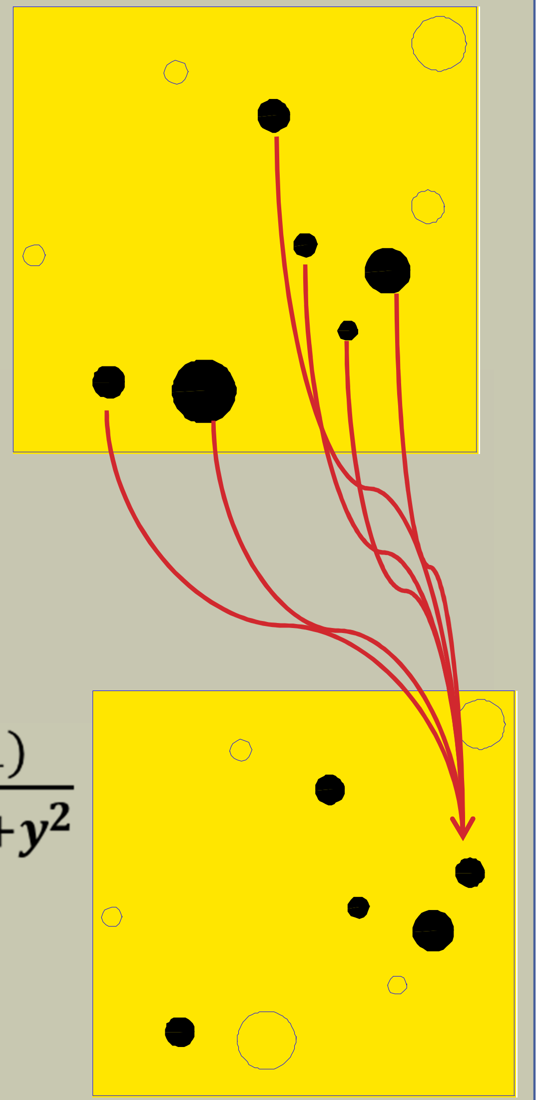
- The **IFM.**

  Multivariate Markov chain: patches are either empty ◯ or occupied ●

$$E_i = P(X_i(t) = 0 \mid X_i(t\text{-}1) = 1, \mathbf{X}_{-i}(t\text{-}1)) = \left(\frac{A_i}{A_0}\right)^x$$

$$C_i(t) = P(X_i(t) = 1 \mid X_i(t\text{-}1) = 0, \mathbf{X}_{-i}(t\text{-}1)) = \frac{\Delta_i^{\,2}(t-1)}{\Delta_i^{\,2}(t-1) + y^2}$$

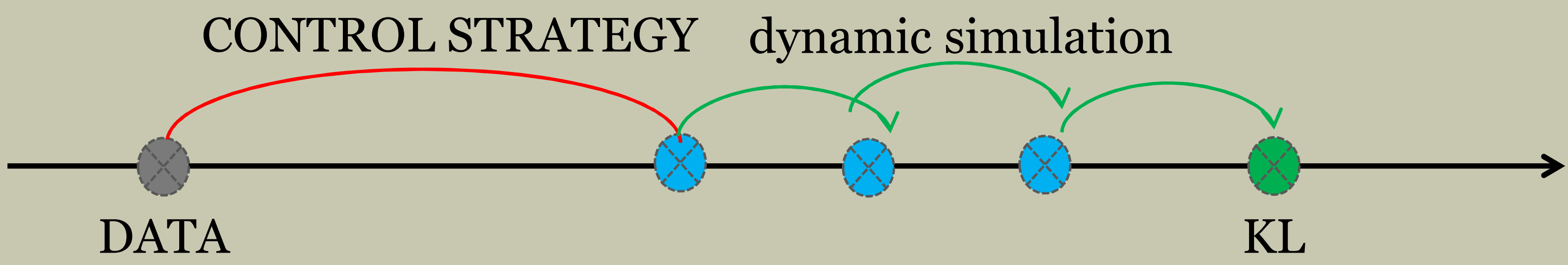
$$\Delta_i(t\text{-}1) = \sum_{j=1, j \neq i}^{n} X_j(t\text{-}1)\, A_j e^{-\alpha d_{ij}}$$ ***connectivity***

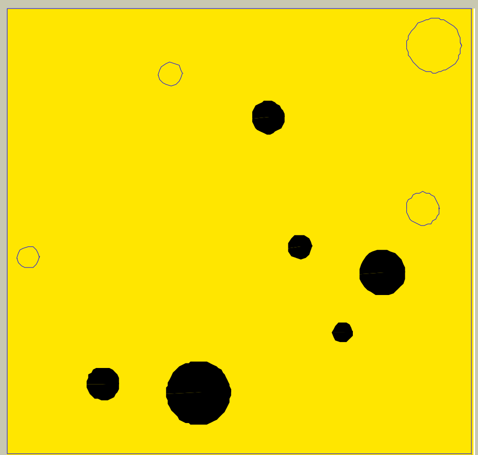


For insect pest, t means *generation*.

The idea by graphics

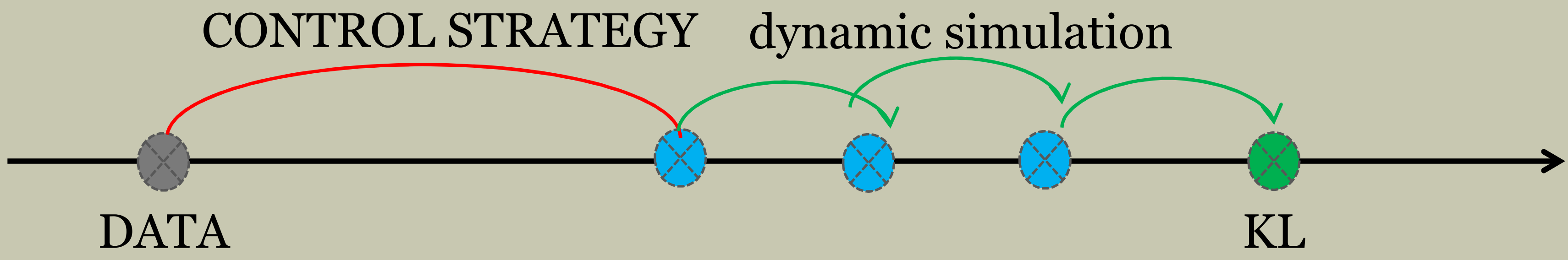
- The **KL** for strategy evaluation. **Purely spatial:**

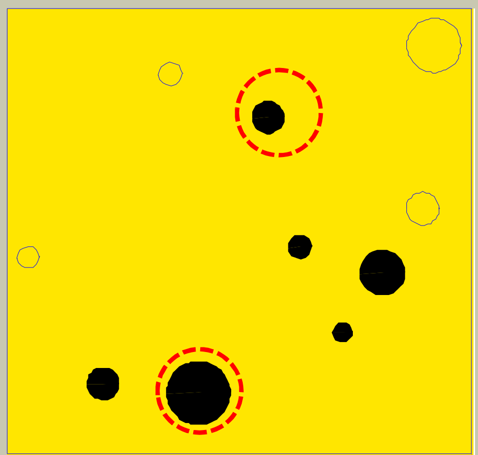




The idea by graphics

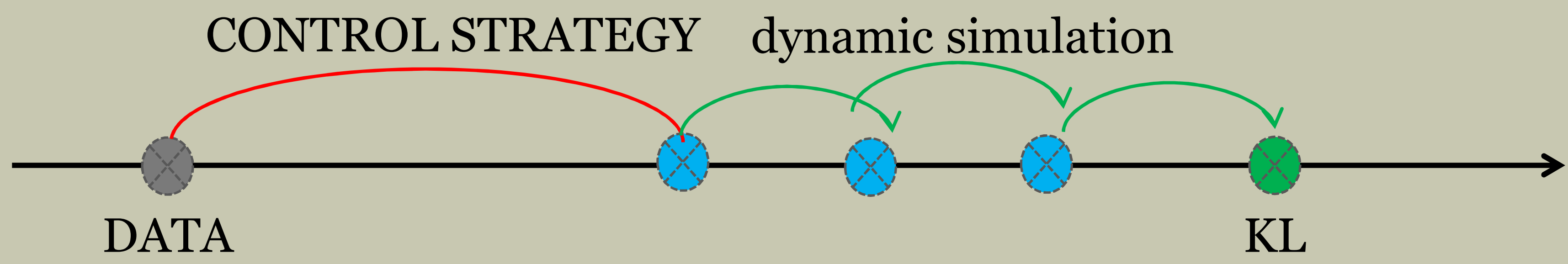
- The **KL** for strategy evaluation. **Purely spatial:**

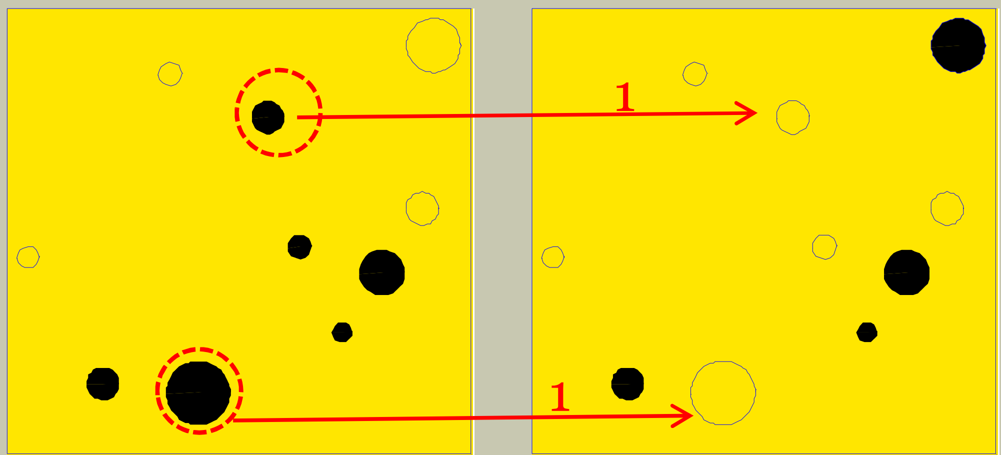




The idea by graphics

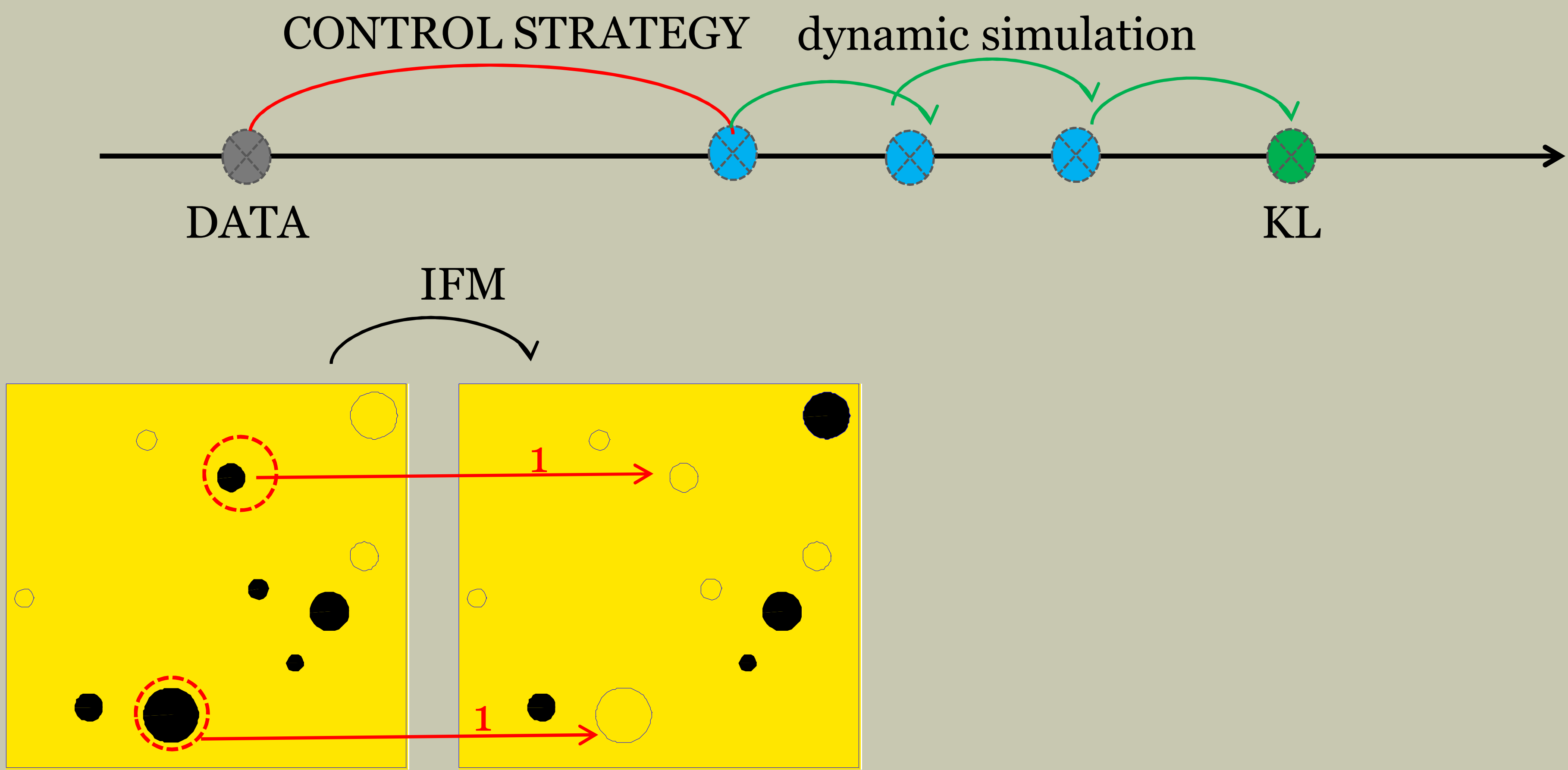
- The **KL** for strategy evaluation. **Purely spatial:**





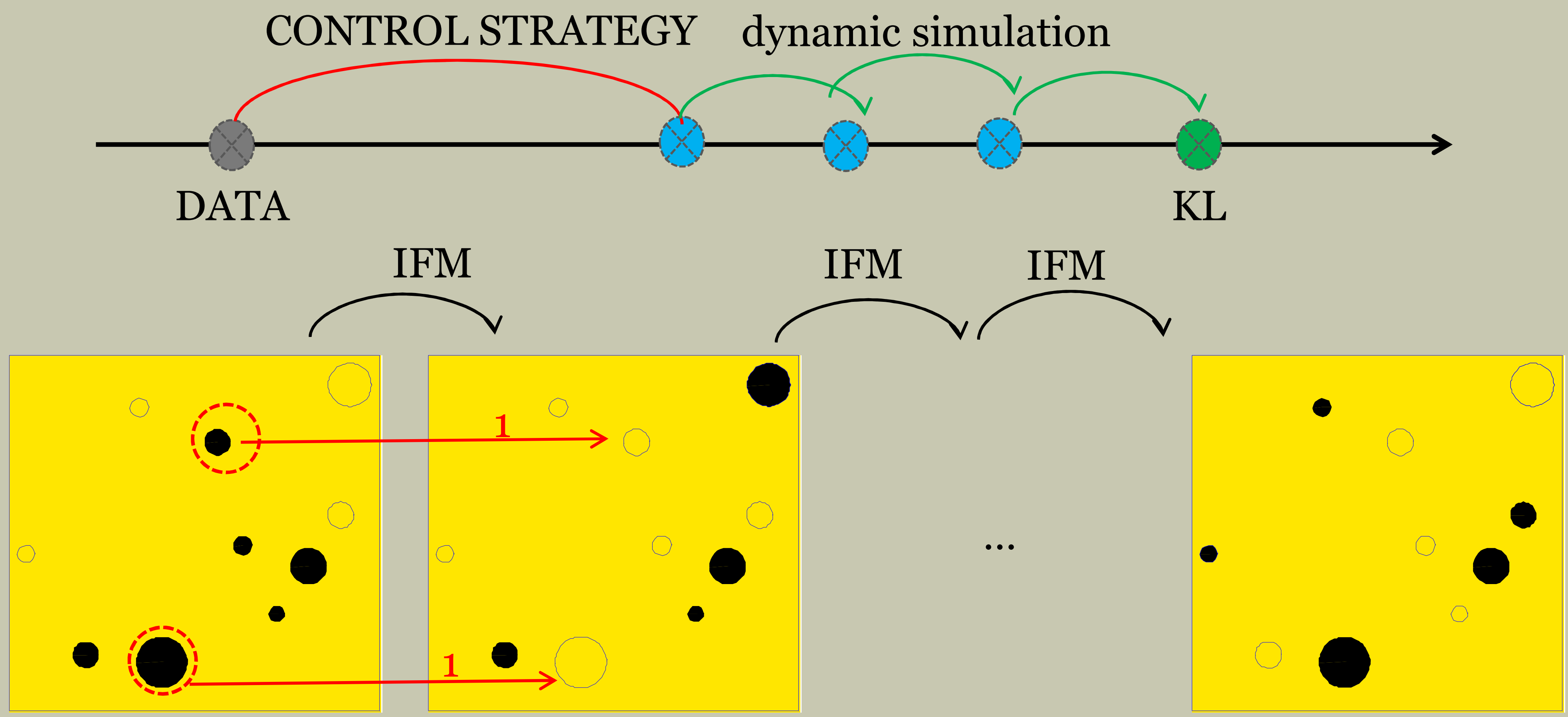
The idea by graphics

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



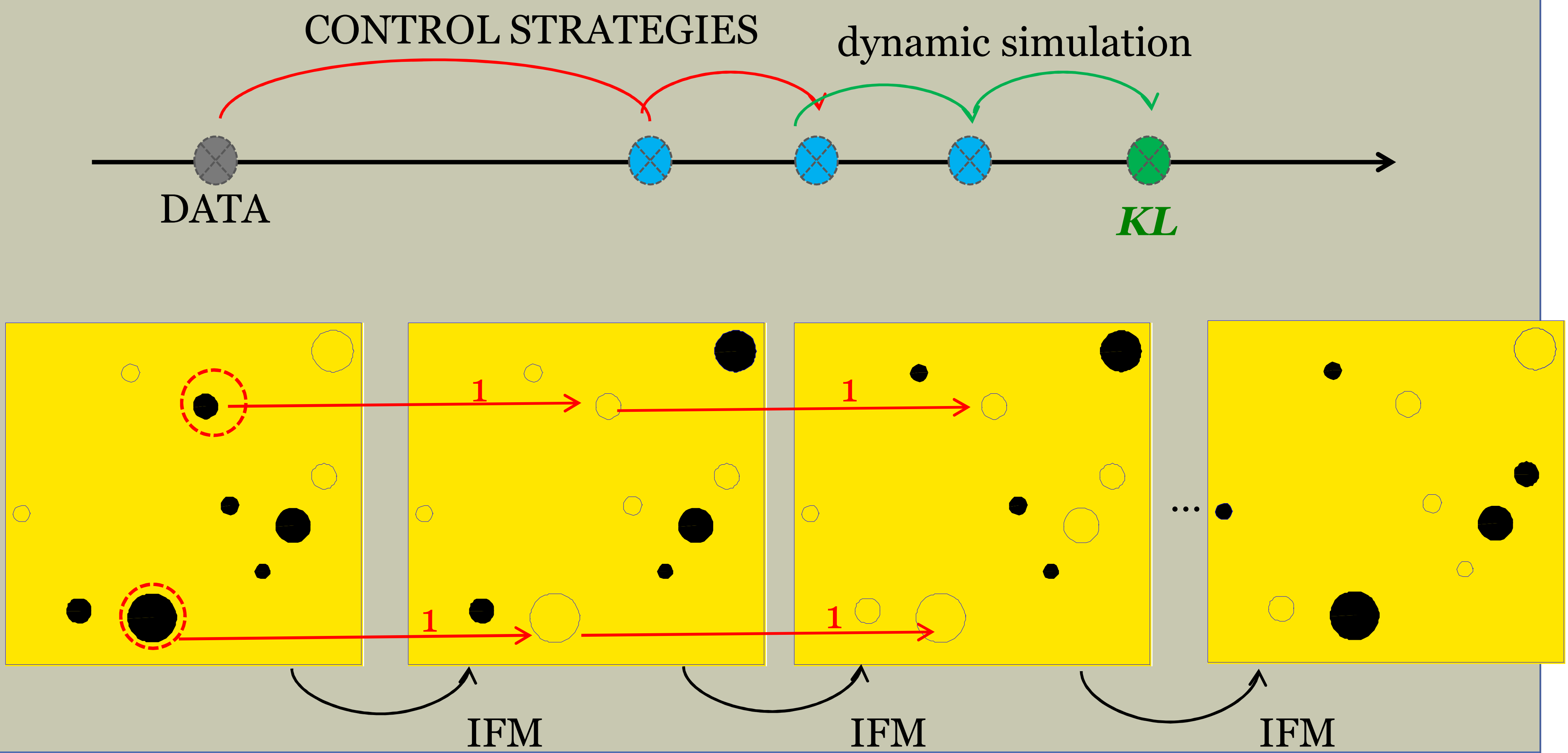
The idea by graphics

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



The idea by graphics

- The **KL** for strategy evaluation. **Spatio-temporal:**



The idea in formula and simulations

- The **KL of** $\boldsymbol{P}(\boldsymbol{X_T})$ from Dirac measure on $0_{\boldsymbol{n}}$: $\boldsymbol{-\ln P(X_T=0)}$
  - 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 $\boldsymbol{X_T}$:

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

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

The idea in formula and simulations

- The **KL of $\boldsymbol{P}(\boldsymbol{X_T})$** from Dirac measure on $0_{\boldsymbol{n}}$: $\boldsymbol{-\ln P(X_T{=}0)}$
  - 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 $\boldsymbol{X_T}$:

$$P(X_T = 0) = P(X_T = 0 \mid X_{T-1} = 1)\, P(X_{T-1} = 1)$$

explicit: $\prod_i (E_i \text{ or } C_i)$

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

The idea in formula and simulations

- The **KL of** $\boldsymbol{P(X_T)}$ from Dirac measure on $0_{\boldsymbol{n}}$: $\boldsymbol{-\ln P(X_T{=}0)}$

  - 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 $\boldsymbol{X_T}$:

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

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

simulated

The idea in formula and simulations

- The **KL of $\boldsymbol{P}(\boldsymbol{X_T})$** from Dirac measure on $0_{\boldsymbol{n}}$: $-\ln \boldsymbol{P}(\boldsymbol{X_T}=0)$

  - 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 $\boldsymbol{X_T}$:

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

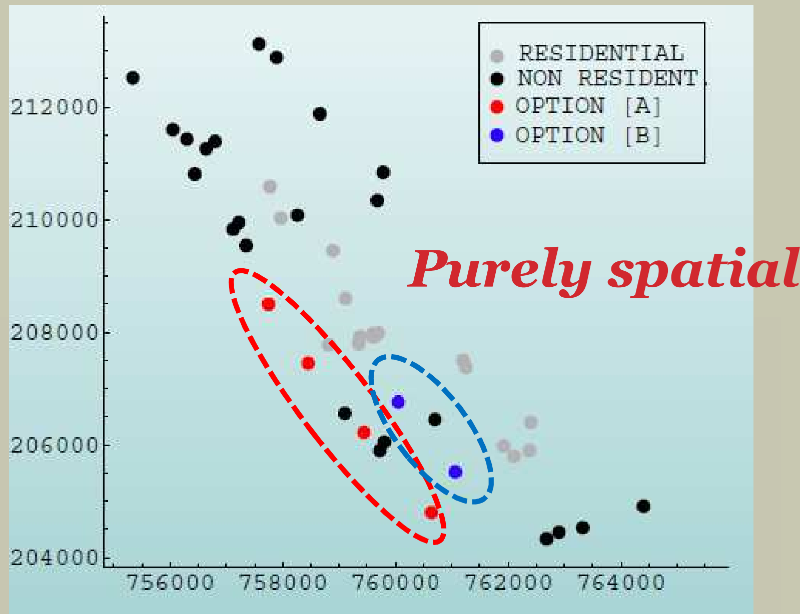
explicit + simulated

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

First Application

- (Amphibians) **Conservation**

Choice between 2 possible sets of new ponds.



- Only 1 year of data
- Equilibrium assumption:

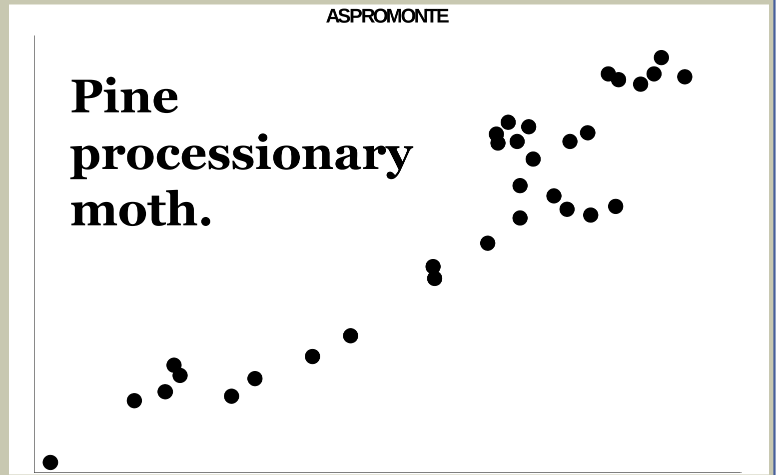
$KL(S) = -\sum_{i=1}^{n} \ln\left[1 - J_i(S)\right]$ measures the divergence of the stationary distribution coming from the strategy S, $\otimes J_i(S)$, from the total extinction.

Gilioli et al. 2008

Second Application 

Gilioli et al. 2013

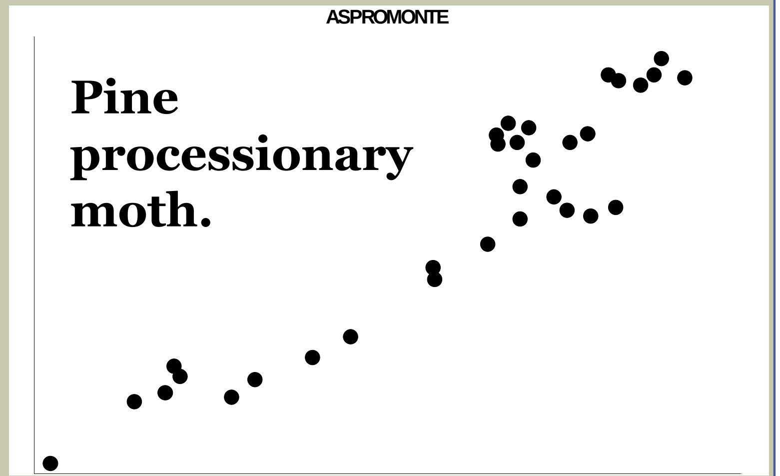
Pest control



Second Application

Gilioli et al. 2013

<u>Pest control</u>

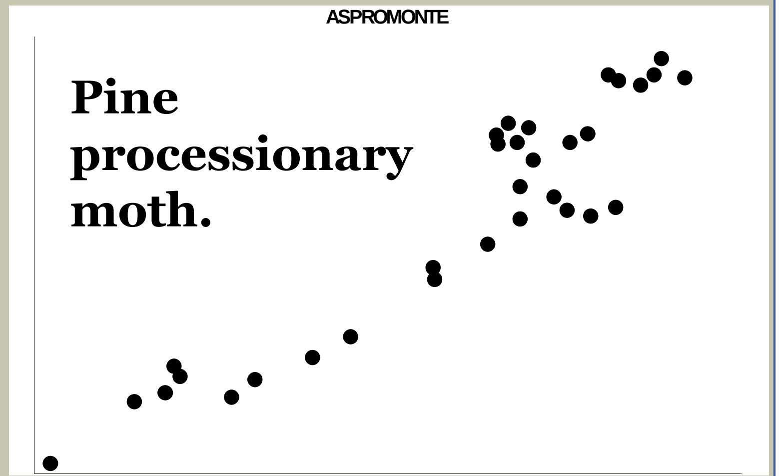


SPATIAL ANALYSIS to verify the IFM applicability to this moth.

Second Application

Gilioli et al. 2013

Pest control



SPATIAL ANALYSIS to verify the IFM applicability to this moth.

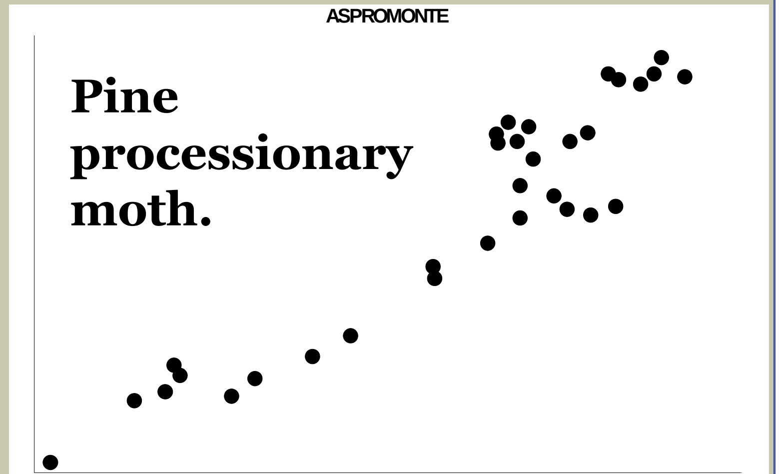
SPATIAL STRATEGIES
(a) scattered sites
(b) close sites
(c) "in line" sites.

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

Second Application

Gilioli et al. 2013

Pest control



SPATIAL ANALYSIS to verify the IFM applicability to this moth.

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

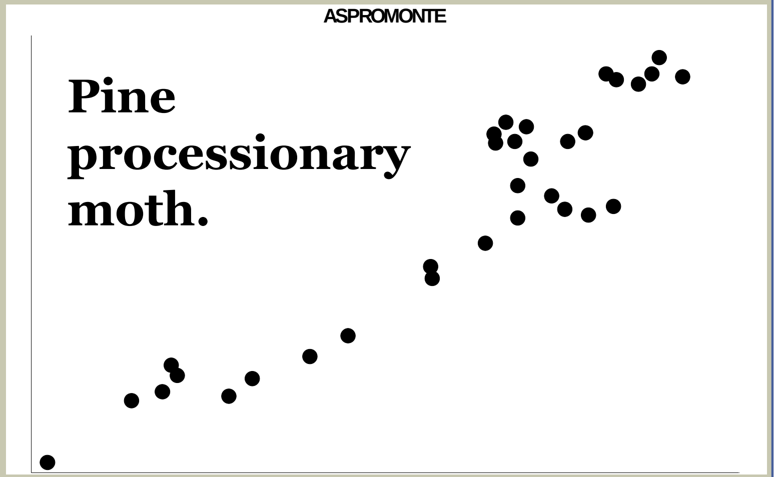
SPATIO-TEMPORAL STRATEGIES for high intervention level.

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

Second Application

Gilioli et al. 2013

<u>Pest control</u>



SPATIAL ANALYSIS to verify the IFM applicability to this moth.

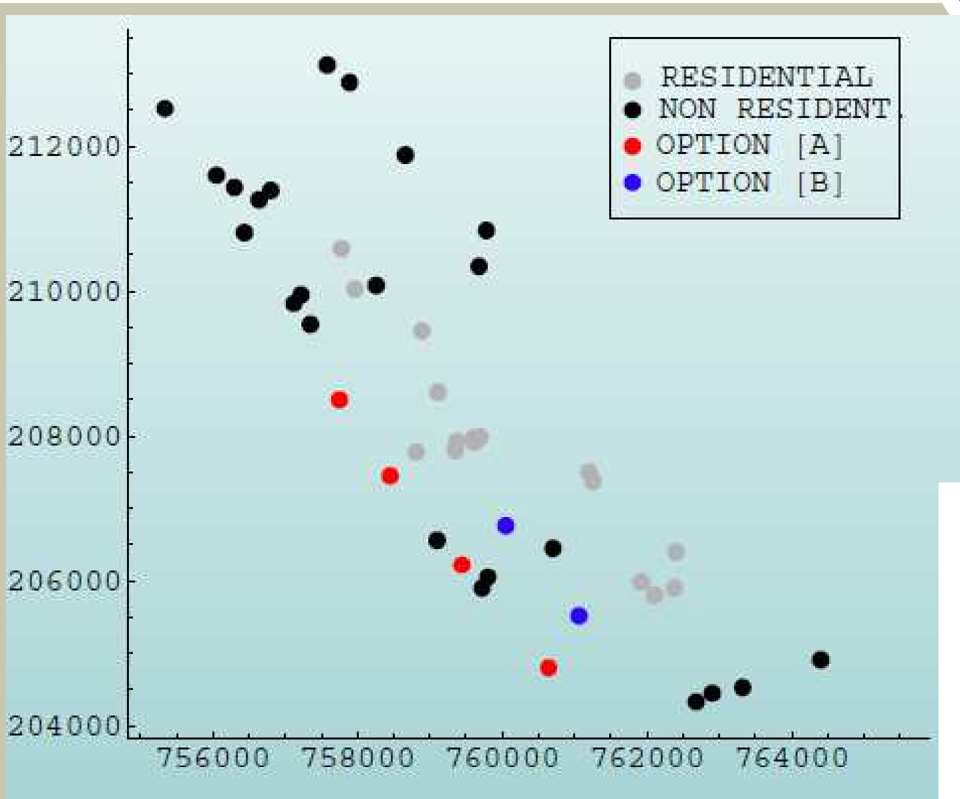
SPATIAL STRATEGIES
(a) scattered sites
(b) close sites
(c) "in line" sites.

SPATIO-TEMPORAL STRATEGIES for high intervention level.

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

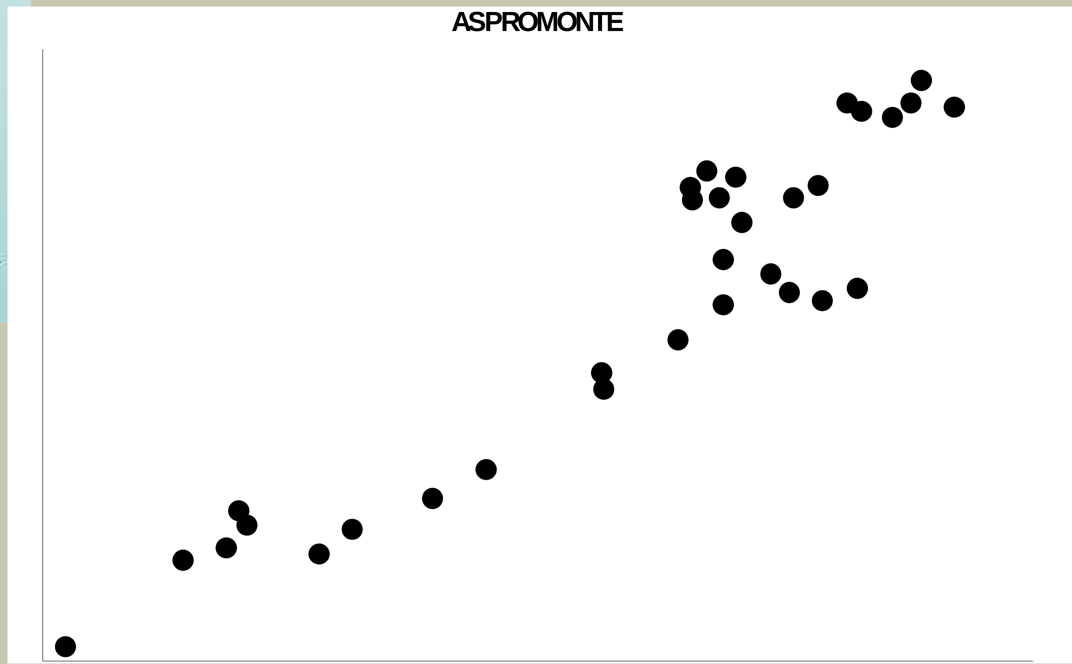
Combination of both: distributed effort for high intervention level.

Testing the idea



CONSERVATION

PEST CONTROL



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

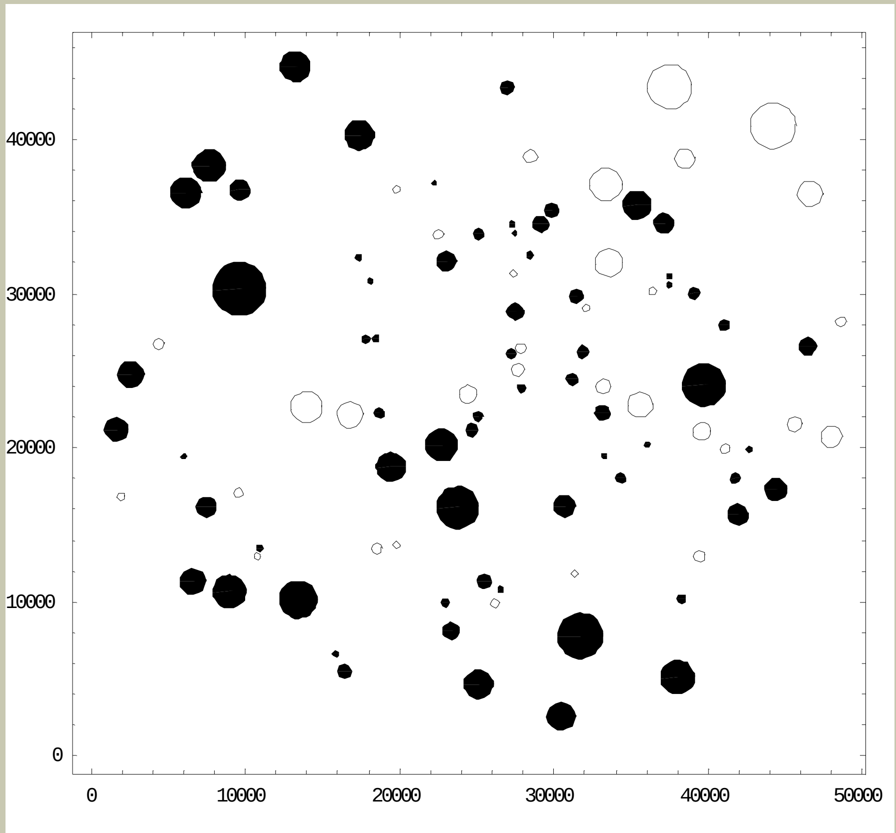
100 patches

● occupied

○ empty

IFM: $x = 0.15$, $y = 0.001$

More general habitat configurations



Square of side 50 Km

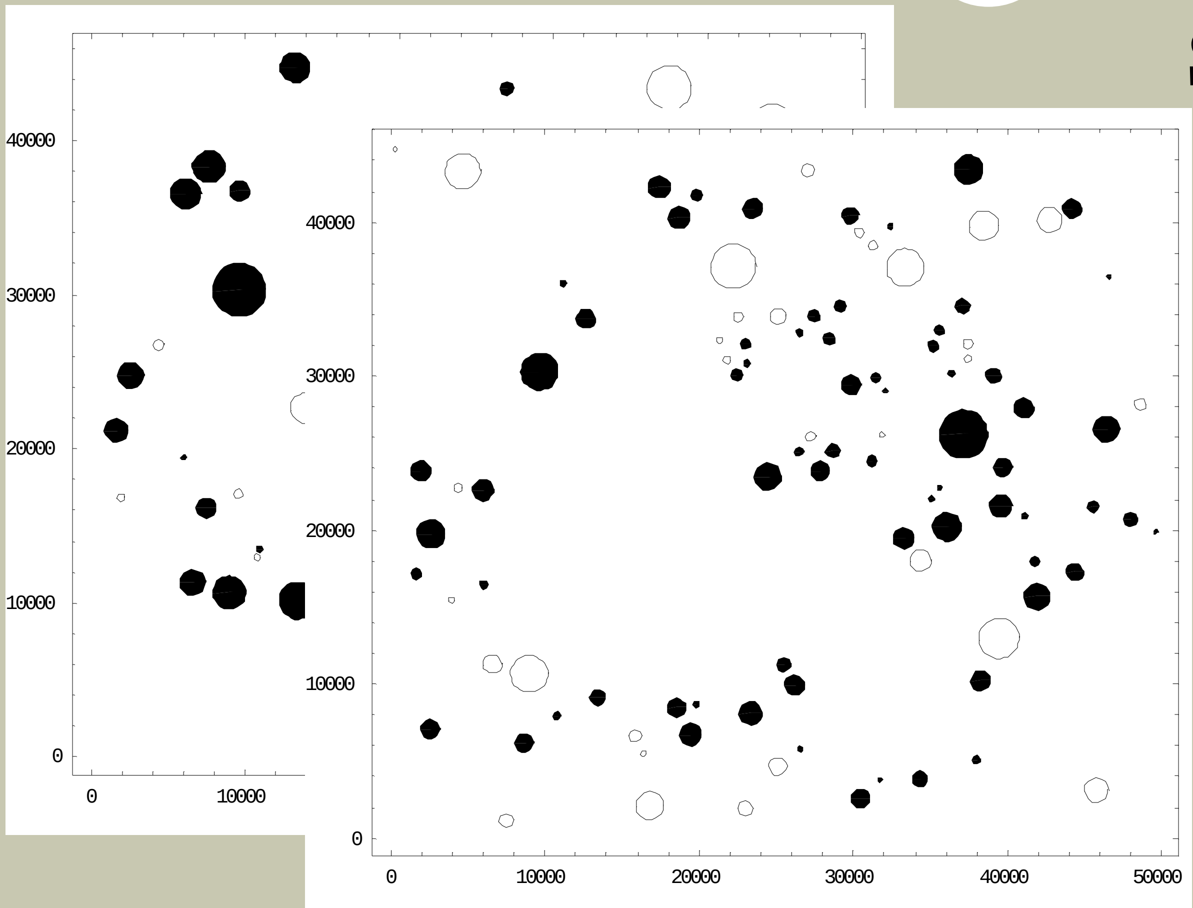
100 patches

occupied

empty

IFM: $x = 0.15$, $y = 0.001$

More general habitat configurations



Square of side 50 Km

100 patches

occupied

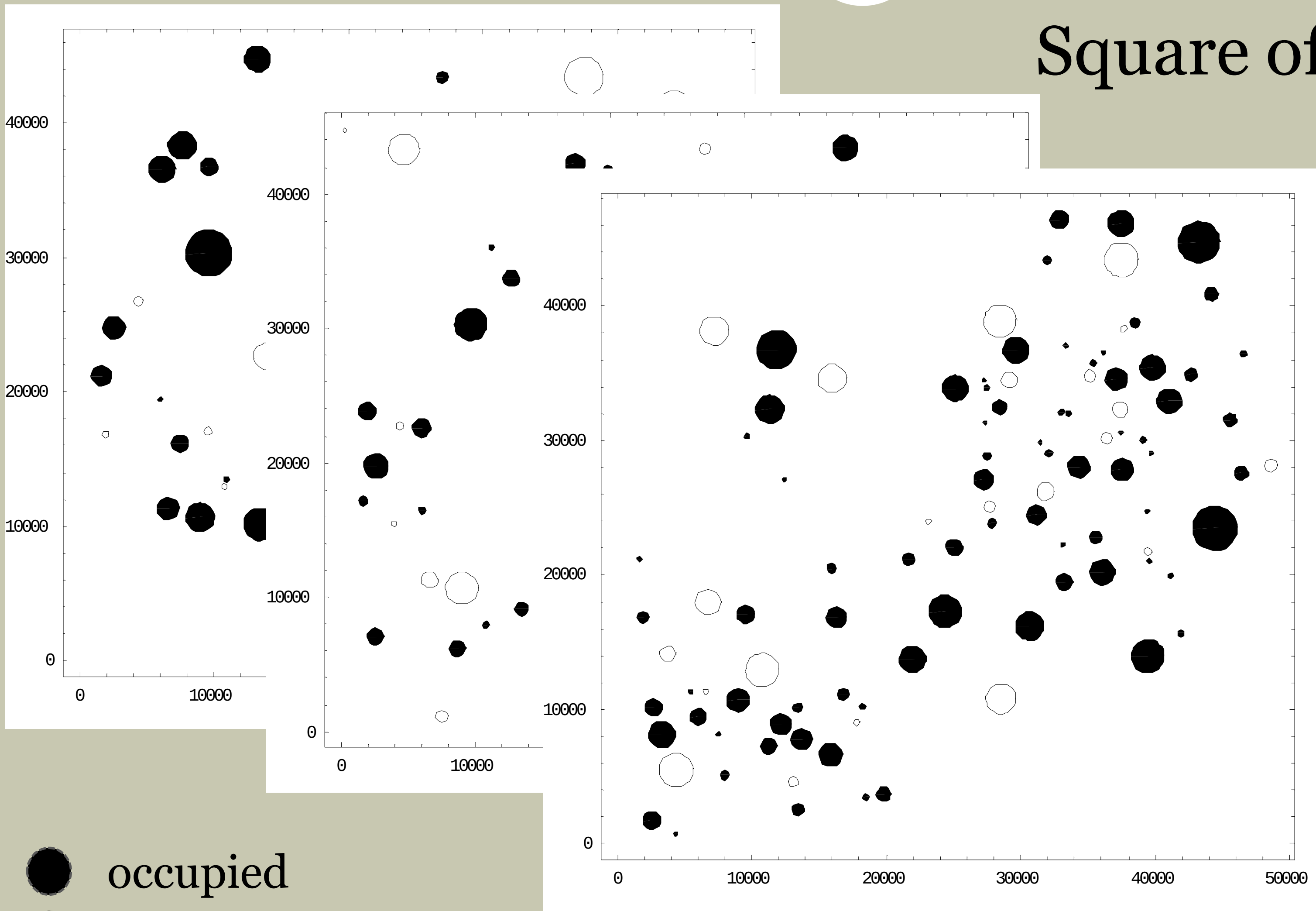
empty

IFM: $x = 0.15$, $y = 0.001$

More general habitat configurations

Square of side 50 Km

100 patches



occupied

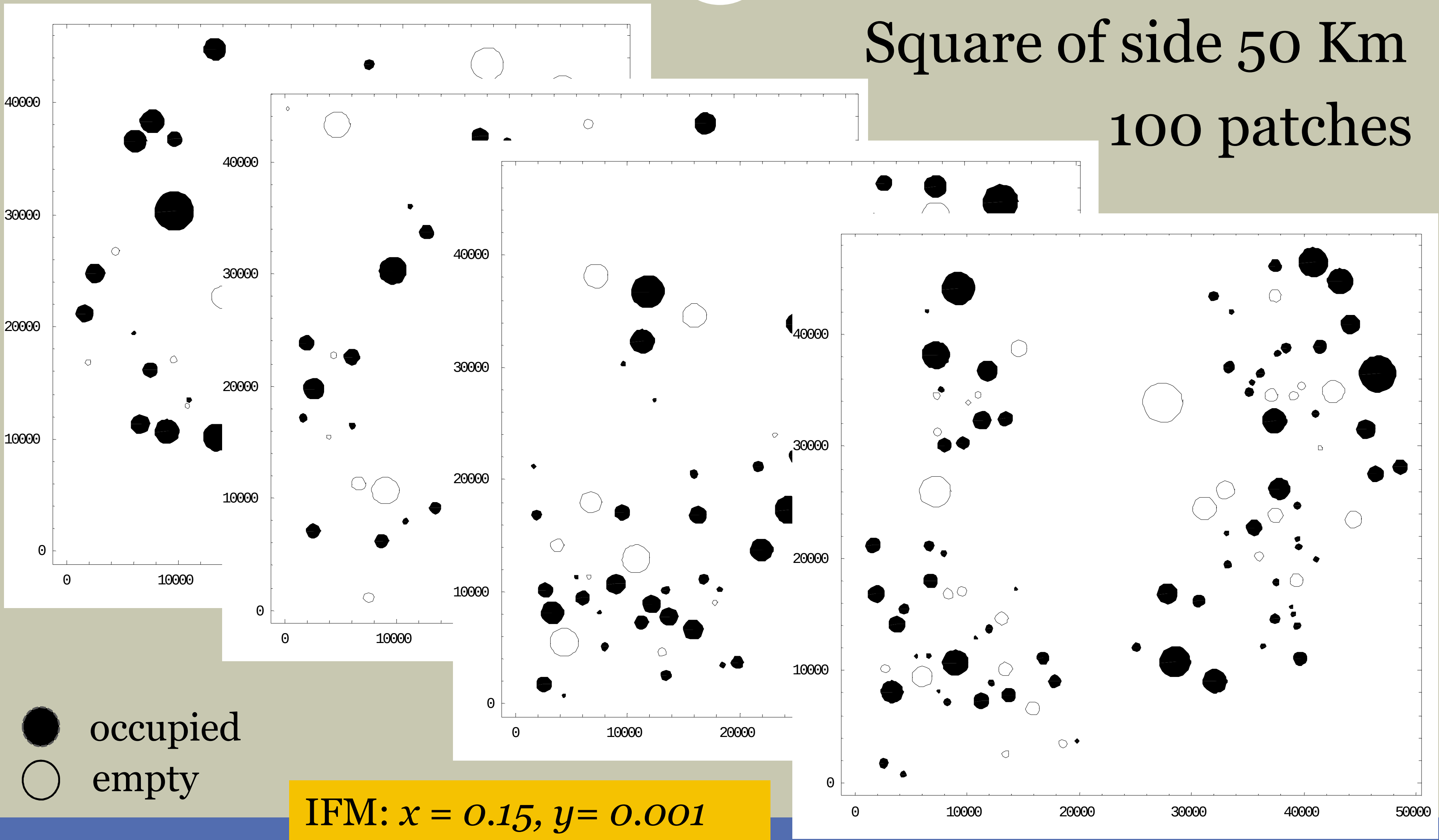
empty

IFM: $x = 0.15$, $y = 0.001$

More general habitat configurations

Square of side 50 Km

100 patches

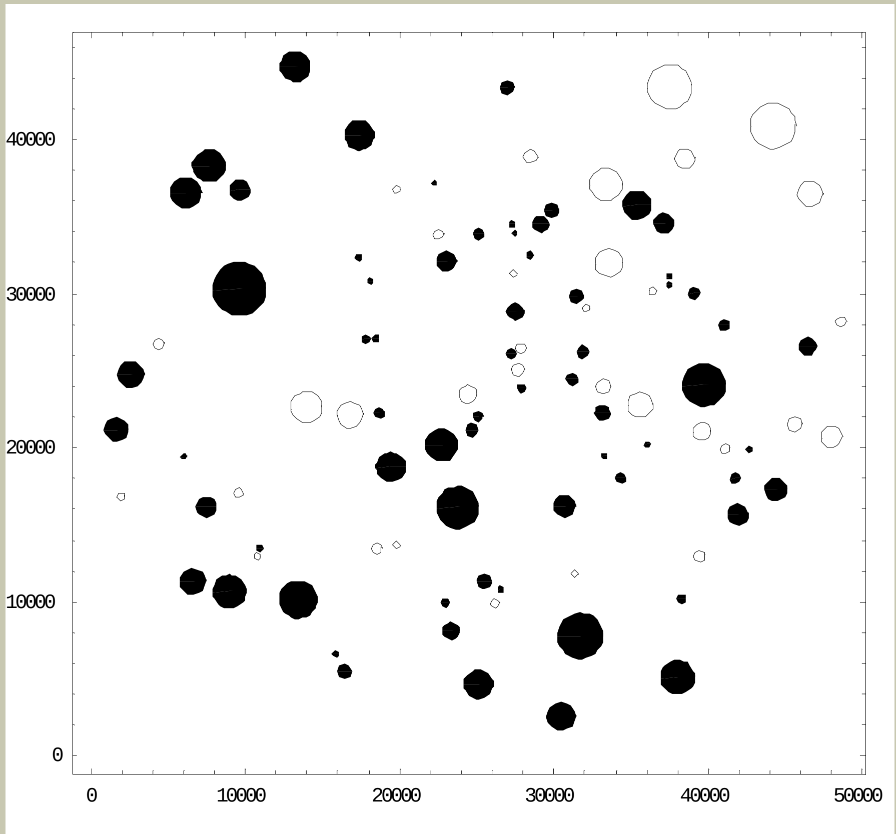


● occupied

○ empty

IFM: $x = 0.15, y = 0.001$

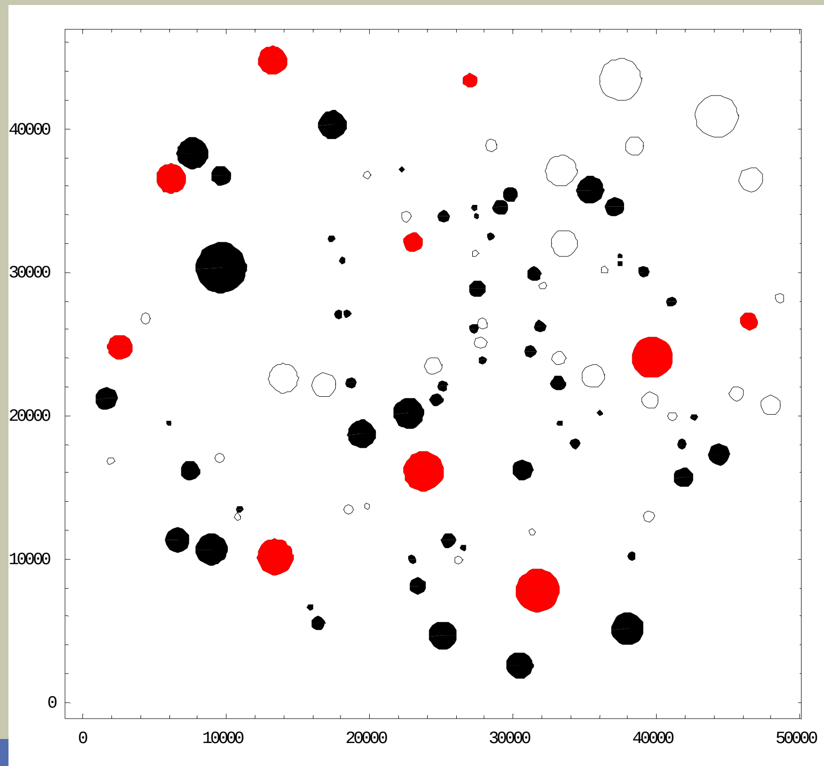
More general habitat configurations & strategies



Square of side 50 Km

100 patches

scattered strategy

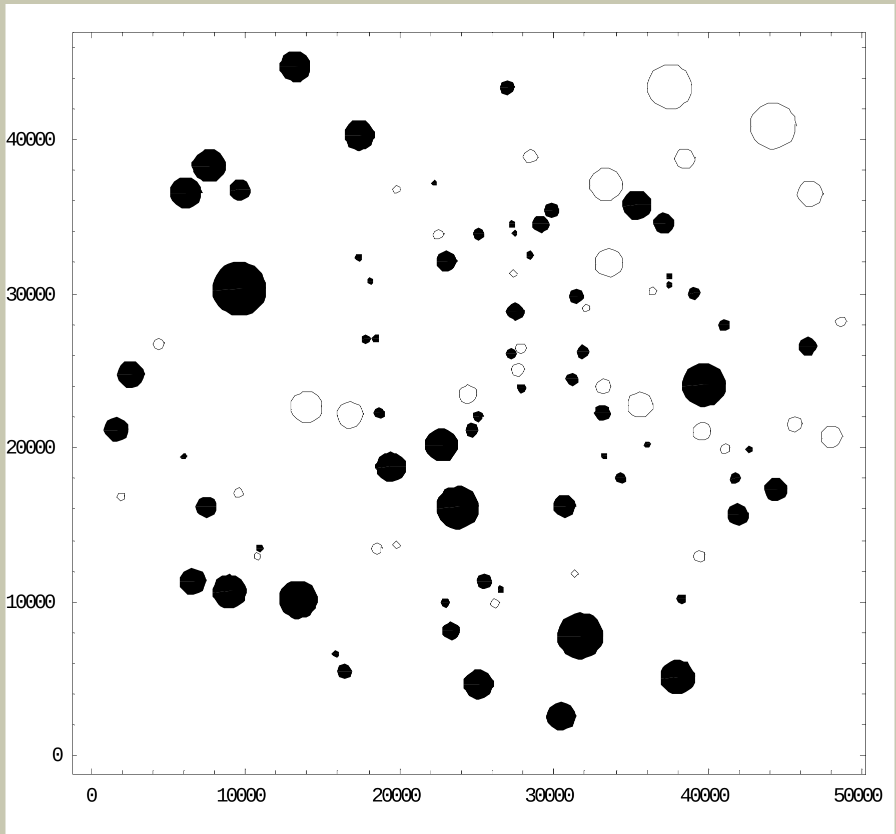


occupied

empty

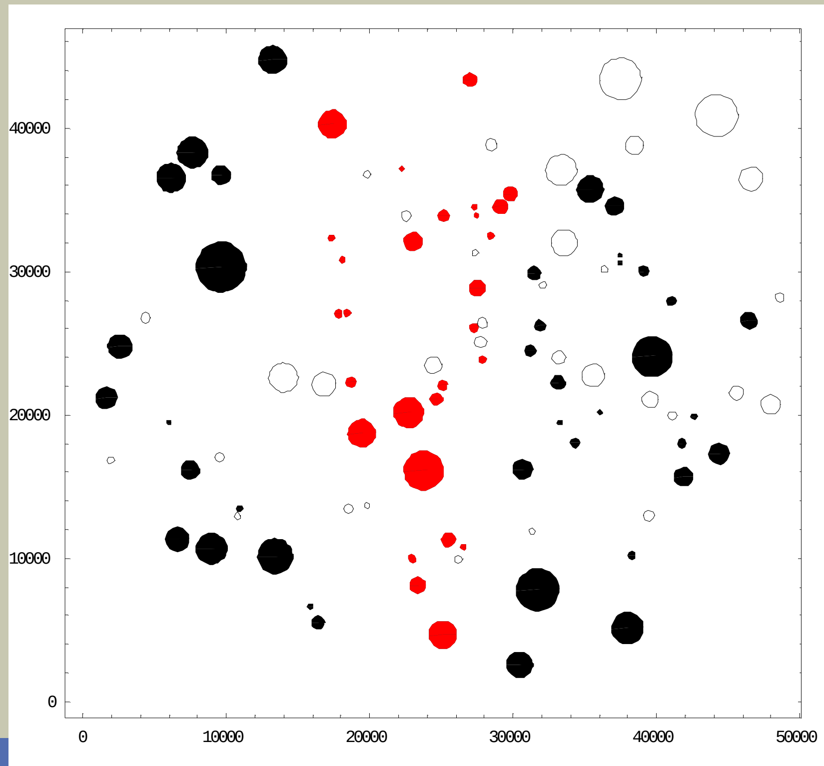
to be treated

More general habitat configurations & strategies



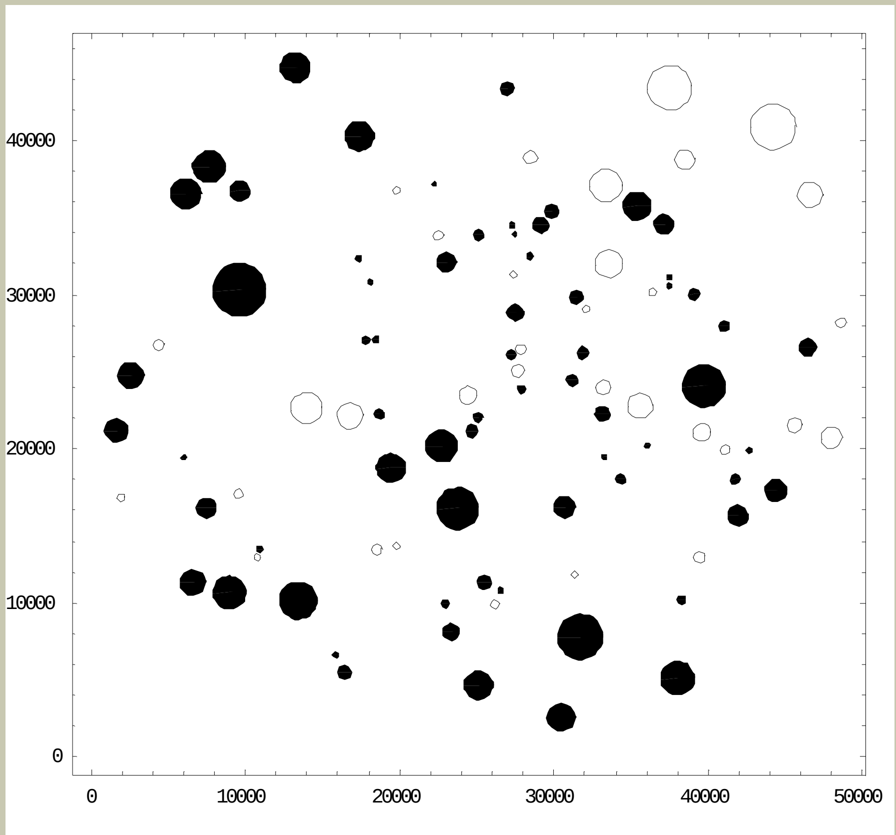
Square of side 50 Km
100 patches

organized strategy



occupied
empty
to be treated

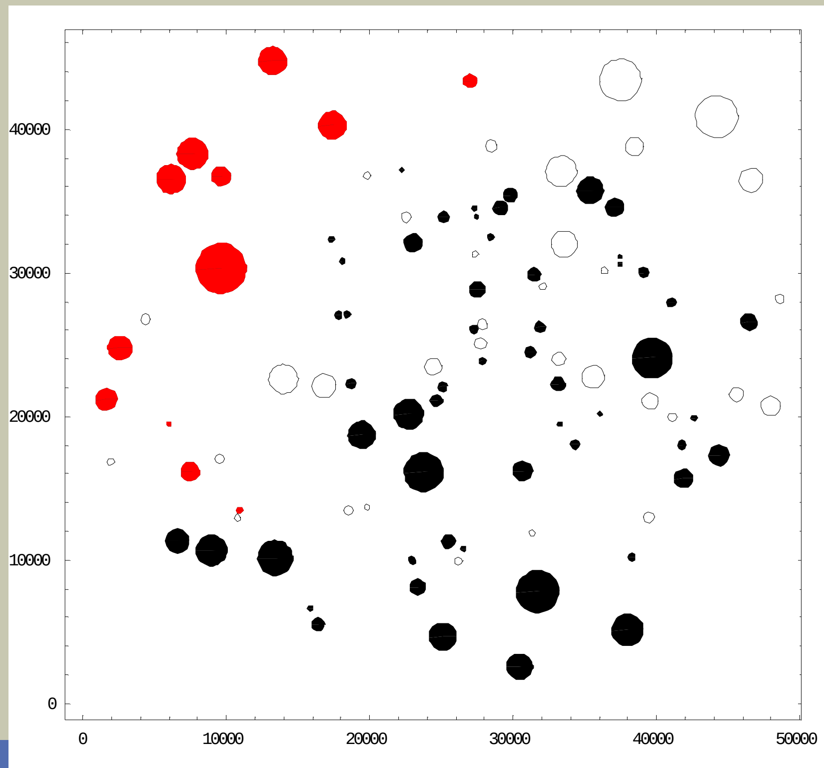
More general habitat configurations & strategies



Square of side 50 Km

100 patches

peripheral-organized strategy



occupied

empty

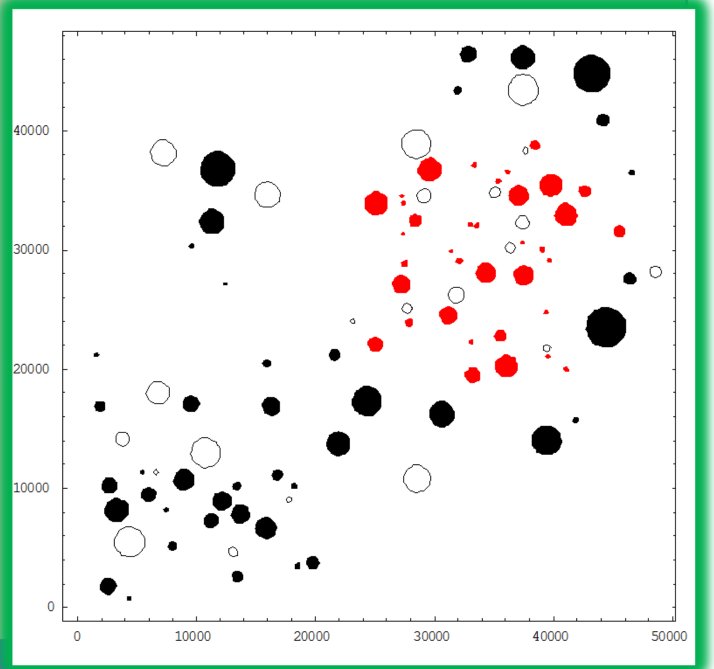
to be treated

General results: purely spatial

data:2013, KL 2016

S_0 = do nothing strategy

KL		Strategies						
		S_0	1	2	3	4	5	6
Case	True param.	**141.9**	141.0	139.7	140.3	135.4	136.6	133.1
	Estimated param.	**146.2**	146.1	144.0	141.1	139.6	140.2	137.5
T.A. (%)		**--**	**23.3**	**21.9**	**21.2**	**22.8**	**22.6**	**23.1**



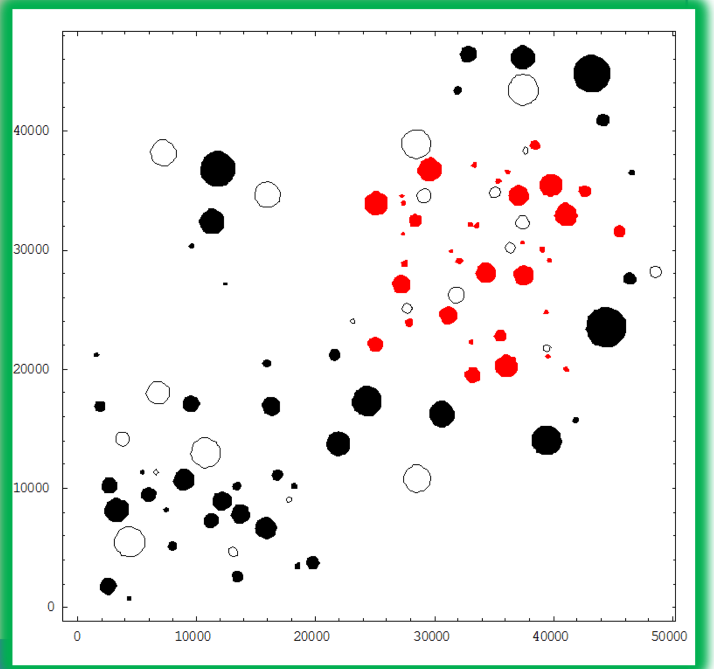
General results: purely spatial

data:2013, KL 2016

S_0 = do nothing strategy

KL		Strategies						
		S_0	1	2	3	4	5	6
Case	True param.	**141.9**	141.0	139.7	140.3	135.4	136.6	133.1
	Estimated param.	**146.2**	146.1	144.0	141.1	139.6	140.2	137.5
T.A. (%)		**--**	**23.3**	**21.9**	**21.2**	**22.8**	**22.6**	**23.1**

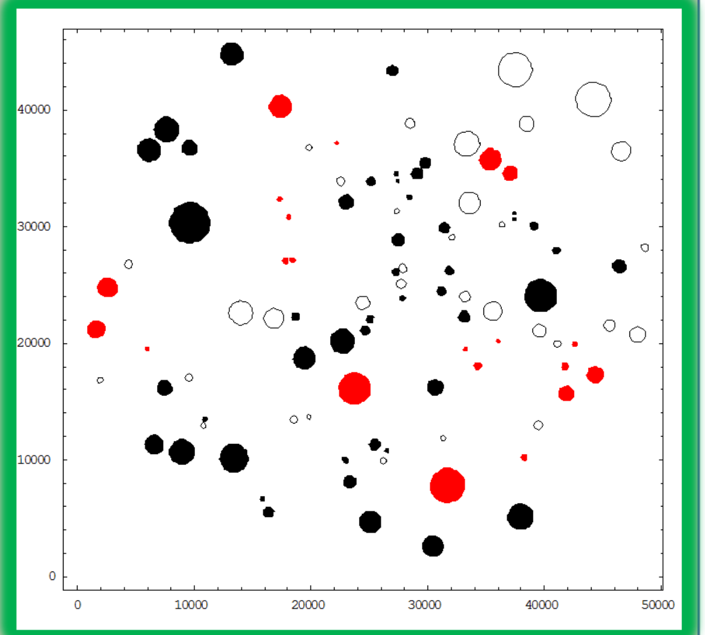
Connectivity		Strategies						
		S_0	1	2	3	4	5	**6**
Case	True param.	**0.212**	0.201	0.145	0.177	0.110	0.117	**0.087**
	Estimated param.	**0.218**	0.205	0.150	0.182	0.112	0.119	**0.088**



General results, contd.

KL		Strategies					
		S_0	1	2	3	4	5
Case	True param.	**143.7**	142.1	175.9	132.9	137.1	147.7
	Estimated param.	**121.0**	119.1	152.7	115.9	113.6	124.9
T.A. (%)		**--**	**24.8**	**21.3**	**21.0**	**20.6**	**27.0**

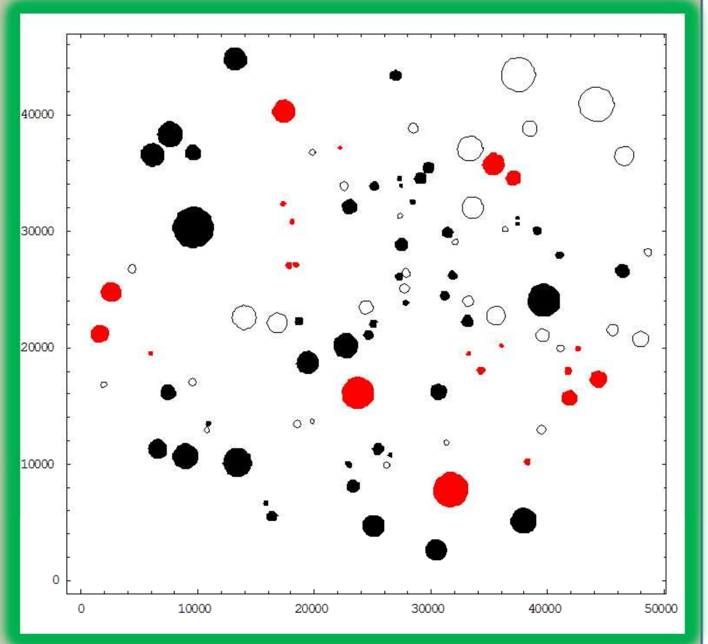
S_0 = do nothing strategy

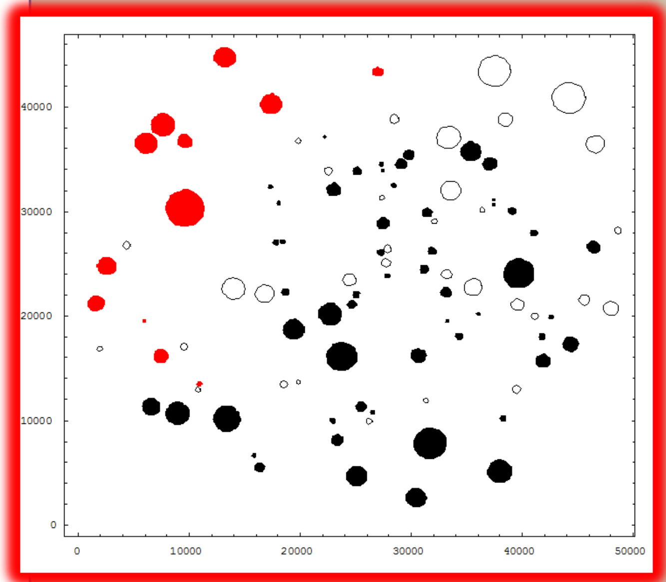


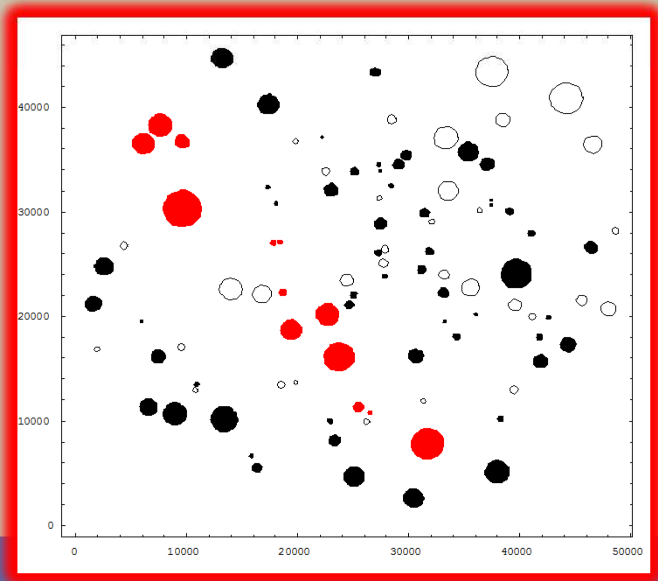
General results, contd.

KL		Strategies					
		S_0	1	2	3	4	5
Case	True param.	**143.7**	142.1	175.9	132.9	137.1	147.7
	Estimated param.	**121.0**	119.1	152.7	115.9	113.6	124.9
T.A. (%)		**--**	**24.8**	**21.3**	**21.0**	**20.6**	**27.0**

S_0 = do nothing strategy

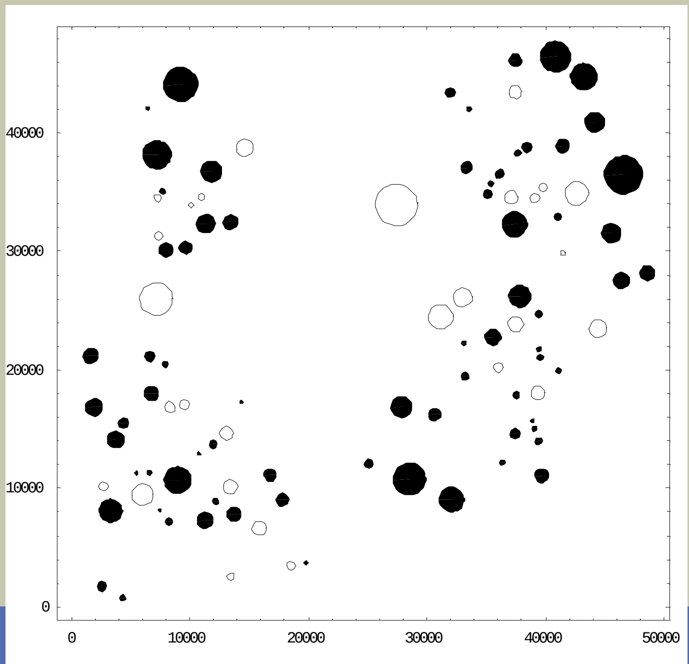






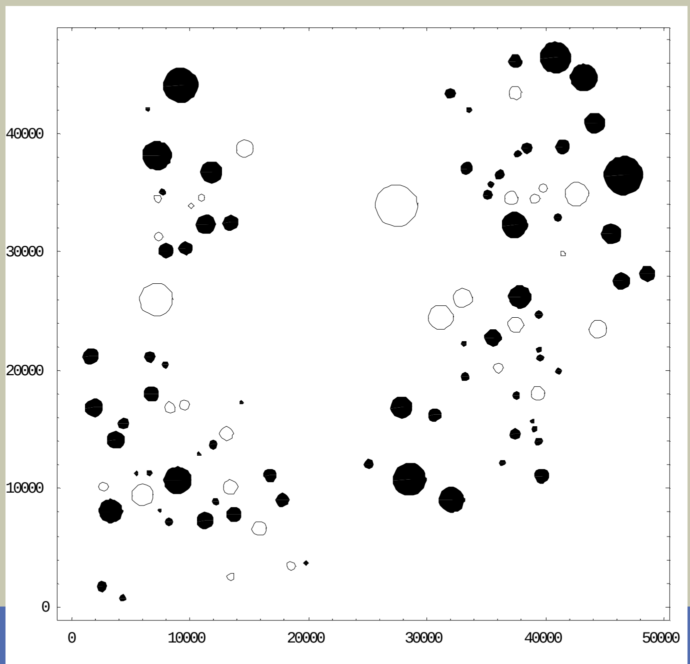
General results, contd.

KL		Strategies										
		S_o	1	2	3	4	5	6	7	8	9	10
Case	True param.	**143.1**	142.3	142.3	141.2	156.3	155.7	140.1	140.0	143.6	140.7	138.9
Case	Estimated param.	**131.5**	132.8	134.1	131.6	159.1	145.6	135.7	137.8	137.6	136.3	129.5
T.A. (%)		--	22.5	22.1	21.6	16.2	20.8	16.1	23.1	14.0	20.2	22.1



General results, contd.

KL	Strategies										
	S_0	1	2	3	4	5	6	7	8	9	10
Case: True param.	**143.1**	142.3	142.3	141.2	156.3	155.7	140.1	140.0	143.6	140.7	138.9
Case: Estimated param.	**131.5**	132.8	134.1	131.6	159.1	145.6	135.7	137.8	137.6	136.3	129.5
T.A. (%)	--	22.5	22.1	21.6	16.2	20.8	16.1	23.1	14.0	20.2	22.1



How come?

STATISTICAL ISSUES:

- Are 100,000 to few simulations?

  - 200,000 do not change the results: higher order?

How come?

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?

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

How come?

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

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

How come?

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

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

- Is the simulated model a "biased model"?

$$C_i(t) = \frac{\Delta_i^{\,2}(t-1)}{\Delta_i^{\,2}(t-1) + y^2} \approx 1 \qquad as\ \ y^2 \approx 0 \text{ at any time } t$$

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

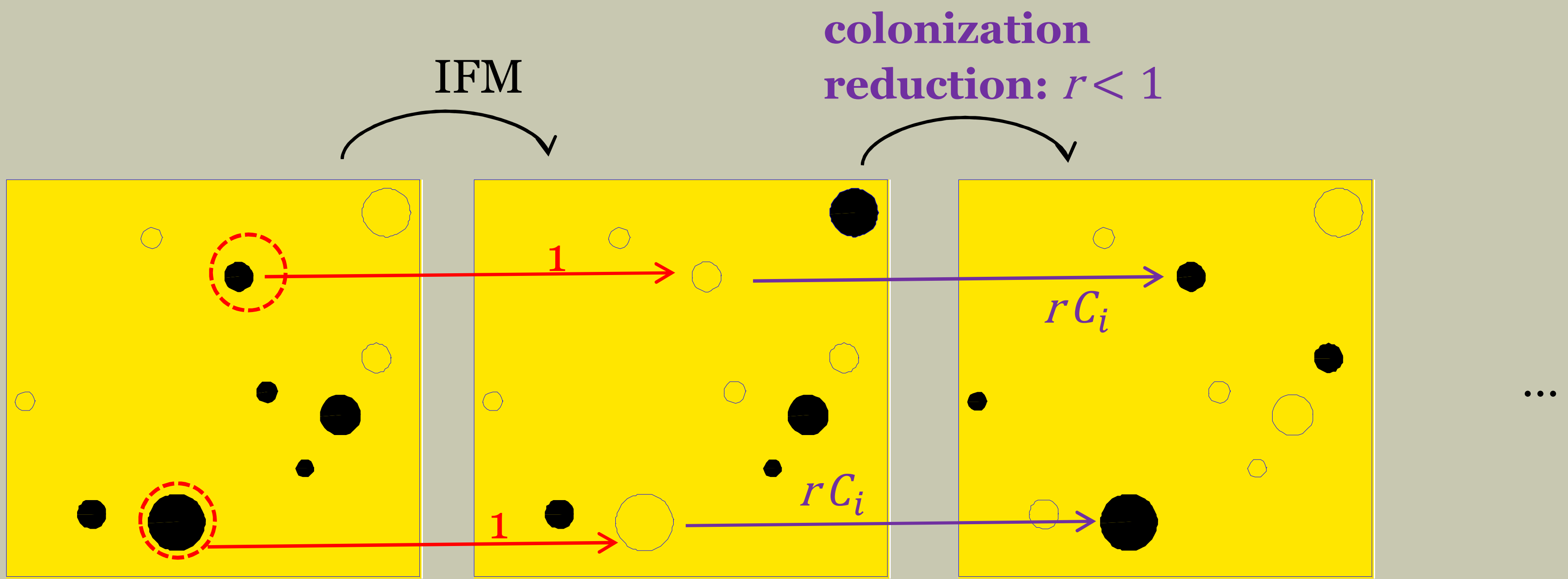
- Is the strategy effect representation adequate?
- Is the IFM a “good” model?

First possible solution…

(English translation...)

Second possible solution…

- Different representation of strategy effect: colonization reduction



A few, partial answers.

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

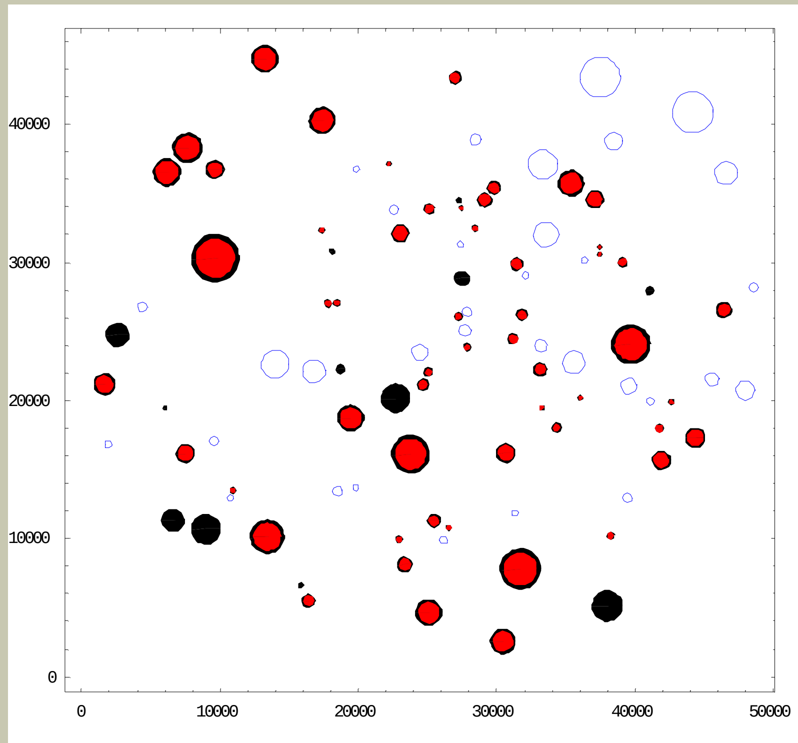
data | | | ***KL***

2013 → 2014 → 2015 → 2016

strat.

col. reduct.
r=0.5
r=0.1
r=0.1

col. reduct.
r=0.5
r=0.5
r=0.1



A few partial answers, contd.

	Do nothing	no reduction	(0.5, 0.5)	(0.1, 0.5)	(0.1,0.1)
KL	143.7	140.5	123.4	99.1	77.0

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*…)

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…*)

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