

1 Assessing the extinction risk of plant metapopulations in urban tree
2 bases

3 -

4 Supporting Information

5 Apolline Louvet, Clément Mantoux, Nathalie Machon

6 **A The Paris 12 dataset - Presentation and estimation results**

Name of the street	Abbreviation	Number of tree bases
Rue Baron le Roy	BARO-1	40
	BARO-2	19
Boulevard de Bercy	BERC-1	22
	BERC-2	37
Rue de Charenton	CHAR-1	69
	CHAR-2	70
Rue Daumesnil	DAUM-1	102
	DAUM-2	44
	DAUM-3	39
Rue Joseph Kessel	KESS-1	36
	KESS-2	33
Rue Montgallet	MONT	48
Rue Pommard	POMM-1	17
	POMM-2	22
Quai de la Rapée	RAPE-1	49
	RAPE-2	25
Rue de Bercy	RBER-1	22
	RBER-2	18
Rue de Reuilly	REUI-1	36
	REUI-2	42
Rue Taine	TAIN-1	33
	TAIN-2	29

Table A.1: List of the portions of streets taken into account in this study

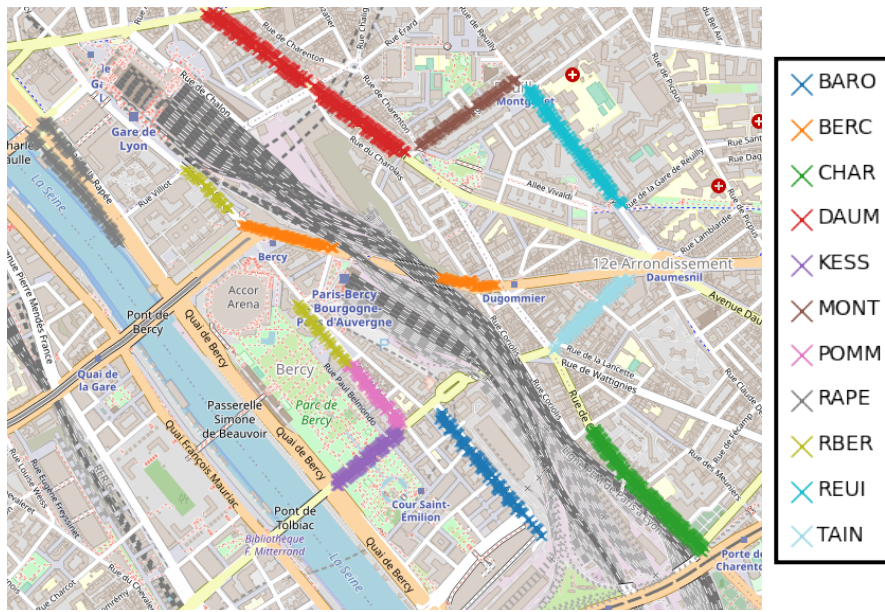


Figure A.1: Map of the study area, adapted from OpenStreetMap. The full street names are listed in Table A.1.

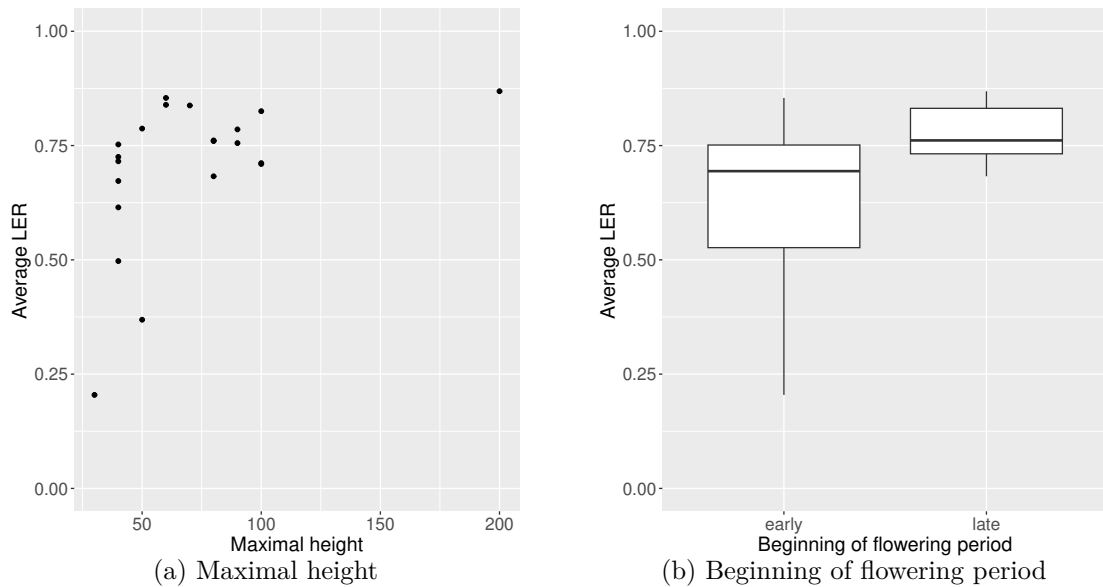


Figure A.2: Illustration of the relation between the local extinction risk (as quantified by the LER metric, and averaged over all streets for each species) and two biological traits for which a borderline significant correlation was identified: (a) the maximal height (p -value = 0.01918), and (b) the beginning of the flowering period (p -value = 0.061). See Table 1 for more details on the biological traits and the correlation tests used.

Name of the plant species	Monitoring in 2013	Number of portions of streets	List of portions of streets
Bromus sterilis	No	2	BERC-1, KESS-2
Capsella bursa-pastoris	Yes	16	BARO-1, BARO-2, BERC-1, BERC-2, CHAR-1, CHAR-2, DAUM-1, DAUM-2, DAUM-3, KESS-1, KESS-2, POMM-1, POMM-2, RBER-1, RBER-2, REUI-1
Chenopodium album	Yes	4	BARO-1, BERC-1, RAPE-1, REUI-1
Conyza sp.	Yes	20	BARO-1, BARO-2, BERC-1, BERC-2, CHAR-1, DAUM-1, DAUM-2, DAUM-3, KESS-1, KESS-2, MONT, POMM-1, POMM-2, RAPE-1, RAPE-2, RBER-1, RBER-2, REUI-1, REUI-2, TAIN-1
Geranium molle	No	3	BARO-1, CHAR-1, CHAR-2
Hordeum murinum	Yes	16	BARO-1, BARO-2, BERC-1, BERC-2, CHAR-1, CHAR-2, DAUM-1, DAUM-2, DAUM-3, KESS-1, KESS-2, RAPE-1, RBER-1, RBER-2, REUI-1, REUI-2
Lactuca serriola	Yes	1	BARO-1
Lolium perenne	No	4	DAUM-1, DAUM-2, DAUM-3, KESS-2
Parietaria judaica	No	2	BARO-1, REUI-2
Plantago lanceolata	Yes	1	RAPE-1
Plantago major	Yes	5	BARO-1, POMM-2, RAPE-1, REUI-1, REUI-2
Poa annua	No	22	BARO-1, BARO-2, BERC-1, BERC-2, CHAR-1, CHAR-2, DAUM-1, DAUM-2, DAUM-3, KESS-1, KESS-2, MONT, POMM-1, POMM-2, RAPE-1, RAPE-2, RBER-1, RBER-2, REUI-1, REUI-2, TAIN-1, TAIN-2
Polygonum aviculare	Yes	11	BARO-1, BARO-2, BERC-1, BERC-2, CHAR-1, CHAR-2, DAUM-1, DAUM-3, RAPE-1, RAPE-2, REUI-1, REUI-2
Senecio inaequidens	Yes	2	BARO-1, BERC-2
Senecio vulgaris	Yes	6	BARO-2, KESS-1, POMM-1, RAPE-1, RBER-2, REUI-1
Sisymbrium irio	Yes	7	BARO-1, BARO-2, BERC-1, BERC-2, DAUM-1, KESS-1, MONT, REUI-1
Sisymbrium officinale	No	2	BARO-1, REUI-1
Sonchus oleraceus	No	13	BARO-1, BARO-2, BERC-1, BERC-2, DAUM-1, KESS-1, KESS-2, MONT, POMM-1, POMM-2, RAPE-1, RAPE-2, RBER-2
Stellaria media	Yes	22	BARO-1, BARO-2, BERC-1, BERC-2, CHAR-1, CHAR-2, DAUM-1, DAUM-2, DAUM-3, KESS-1, KESS-2, MONT, POMM-1, POMM-2, RAPE-1, RAPE-2, RBER-1, RBER-2, REUI-1, REUI-2, TAIN-1, TAIN-2
Taraxacum sp.	Yes	19	BARO-1, BARO-2, BERC-1, BERC-2, CHAR-1, CHAR-2, DAUM-1, DAUM-2, DAUM-3, KESS-1, KESS-2, MONT, POMM-1, RAPE-1, RAPE-2, RBER-1, RBER-2, REUI-1, REUI-2
Veronica persica	Yes	1	RAPE-1

Table A.2: List of the plant species taken into account in this study. The taxonomic reference used is the French Flora Reference TAXREF v8.0 (2014). For each species, the portion of streets which were taken into account in this study were the ones in which the focal species was observed in an average of at least 10% of the patches per year over the monitoring period (2009-2018 for species monitored in 2013, 2014-2018 otherwise).

	Estimate	Standard error	p-value
(Intercept)	0,127	0,161	0,4339
Species	Estimate	Standard error	p-value
Capsella bursa-pastoris	-0,111	0,158	0,4852
Chenopodium album	0,706	0,182	0,0002
Coryza sp.	-0,093	0,157	0,5525
Geranium molle	0,129	0,195	0,5087
Hordeum murinum	-0,118	0,158	0,4554
Lactuca serriola	0,841	0,259	0,0014
Lolium perenne	0,386	0,182	0,0353
Parietaria judaica	0,378	0,214	0,0790
Plantago lanceolata	1,003	0,261	0,0002
Plantago major	0,725	0,179	0,0001
Poa annua	-0,146	0,157	0,3521
Polygonum aviculare	0,587	0,162	0,0004
Senecio inaequidens	0,856	0,213	0,0001
Senecio vulgaris	0,054	0,174	0,7590
Sisymbrium irio	0,067	0,167	0,6893
Sisymbrium officinale	0,797	0,212	0,0003
Sonchus oleraceus	0,606	0,160	0,0002
Stellaria media	-0,146	0,157	0,3522
Taraxacum sp.	-0,145	0,157	0,3574
Veronica persica	0,074	0,261	0,7770
Street	Estimate	Standard error	p-value
BARO-2	0,129	0,086	0,1370
BERC-1	-0,110	0,084	0,1955
BERC-2	0,035	0,085	0,6781
CHAR-1	-0,027	0,092	0,7725
CHAR-2	0,080	0,096	0,4045
DAUM-1	0,002	0,086	0,9791
DAUM-2	0,056	0,097	0,5634
DAUM-3	0,039	0,093	0,6778
KESS-1	0,003	0,089	0,9738
KESS-2	0,038	0,091	0,6794
MONT	0,127	0,101	0,2132
POMM-1	0,060	0,097	0,5340
POMM-2	0,050	0,101	0,6232
RAPE-1	-0,129	0,086	0,1337
RAPE-2	0,053	0,101	0,6042
RBER-1	0,038	0,101	0,7112
RBER-2	-0,006	0,093	0,9509
REUI-1	-0,078	0,081	0,3354
REUI-2	0,049	0,091	0,5911
TAIN-1	0,002	0,132	0,9876
TAIN-2	0,020	0,157	0,9005

Table A.3: Summary of the results of the regression of MaxGER on species and portions of streets. For the species, estimates are expressed relative to *Bromus sterilis*. For the streets, estimates are expressed relative to BARO-1. The species in bold are the ones for which a significantly higher extinction risk was identified.

Biological trait	p-value
Dispersal mechanism	0.6899
Flowering duration	0.002677
Seed mass	0.5479
Heat preference	0.763
Pollination vector	0.5634
Maximal height	0.004763
Beginning of flowering period	0.00915

Table A.4: Summary of the results of the correlation tests of the global extinction risk (as quantified by the MaxGER metric and averaged over all streets for each species) with the species traits listed in Table 1. The biological traits in bold are the ones for which a significant correlation was identified (when accounting for multiple testing using the Holm-Bonferroni method).

	Estimate	Standard error	p-value
(Intercept)	0,745	0,091	0,0000
Species	Estimate	Standard error	p-value
Capsella bursa-pastoris	-0,053	0,090	0,5552
Chenopodium album	0,009	0,103	0,932
Conyza sp.	-0,010	0,089	0,9107
Geranium molle	0,022	0,111	0,8465
<i>Hordeum murinum</i>	-0,254	0,090	0,0054
Lactuca serriola	0,124	0,147	0,3993
Lolium perenne	0,078	0,103	0,4508
Parietaria judaica	0,012	0,121	0,9210
Plantago lanceolata	0,065	0,148	0,6599
Plantago major	0,109	0,102	0,2847
<i>Poa annua</i>	-0,578	0,089	0,0000
Polygonum aviculare	-0,071	0,092	0,4381
Senecio inaequidens	0,062	0,121	0,6096
Senecio vulgaris	0,111	0,099	0,2633
Sisymbrium irio	0,004	0,095	0,9625
Sisymbrium officinale	0,034	0,120	0,7797
Sonchus oleraceus	0,022	0,091	0,8103
Stellaria media	-0,110	0,089	0,2164
<i>Taraxacum sp.</i>	-0,267	0,089	0,0032
Veronica persica	-0,045	0,148	0,7602
Street	Estimate	Standard error	p-value
BARO-2	0,077	0,049	0,1182
BERC-1	-0,045	0,048	0,3518
BERC-2	0,046	0,048	0,3377
CHAR-1	-0,014	0,052	0,7915
CHAR-2	0,076	0,054	0,1635
DAUM-1	0,008	0,049	0,8677
DAUM-2	0,041	0,055	0,4634
DAUM-3	0,033	0,053	0,5351
KESS-1	-0,038	0,051	0,4494
KESS-2	-0,022	0,052	0,6724
MONT	0,082	0,057	0,1544
POMM-1	0,083	0,055	0,1326
POMM-2	0,066	0,057	0,2509
RAPE-1	-0,085	0,049	0,0833
RAPE-2	0,071	0,057	0,2197
RBER-1	0,118	0,057	0,0418
RBER-2	-0,001	0,053	0,9786
REUI-1	-0,046	0,046	0,3121
REUI-2	-0,009	0,052	0,8639
TAIN-1	0,163	0,075	0,0308
TAIN-2	0,229	0,089	0,0109

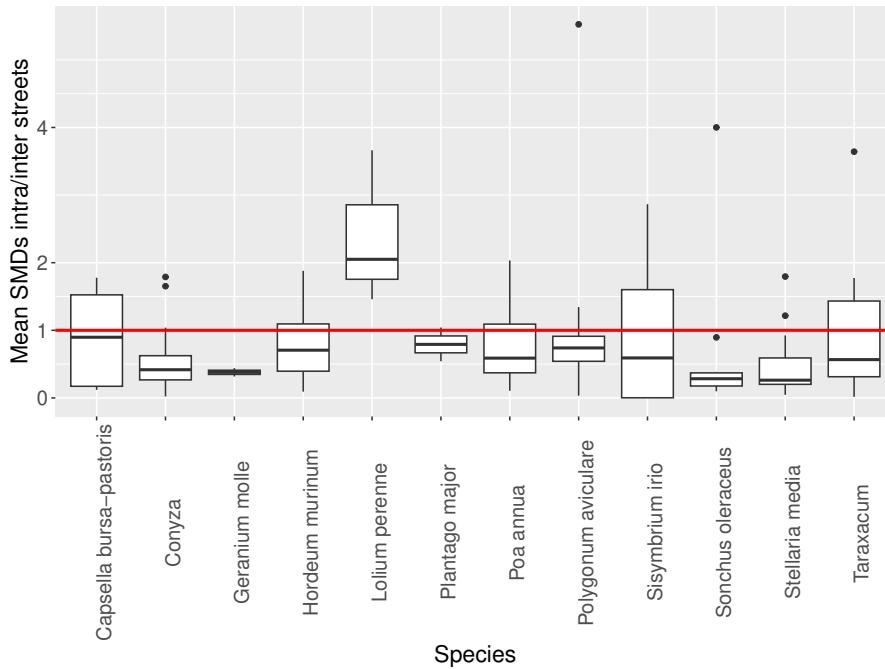
Table A.5: Summary of the results of the regression of LER on species and portions of streets. For the species, estimates are expressed relative to *Bromus sterilis*. For the streets, estimates are expressed relative to BARO-1. The species in italic are the ones for which a significantly lower LER was identified. The portions of streets in bold are the ones for which a significantly higher LER was identified.

Biological trait	p-value
Dispersal mechanism	0.2049
Flowering duration	0.2748
Seed mass	0.3375
Heat preference	0.814
Pollination vector	0.3371
Maximal height	0.01918
Beginning of flowering period	0.061

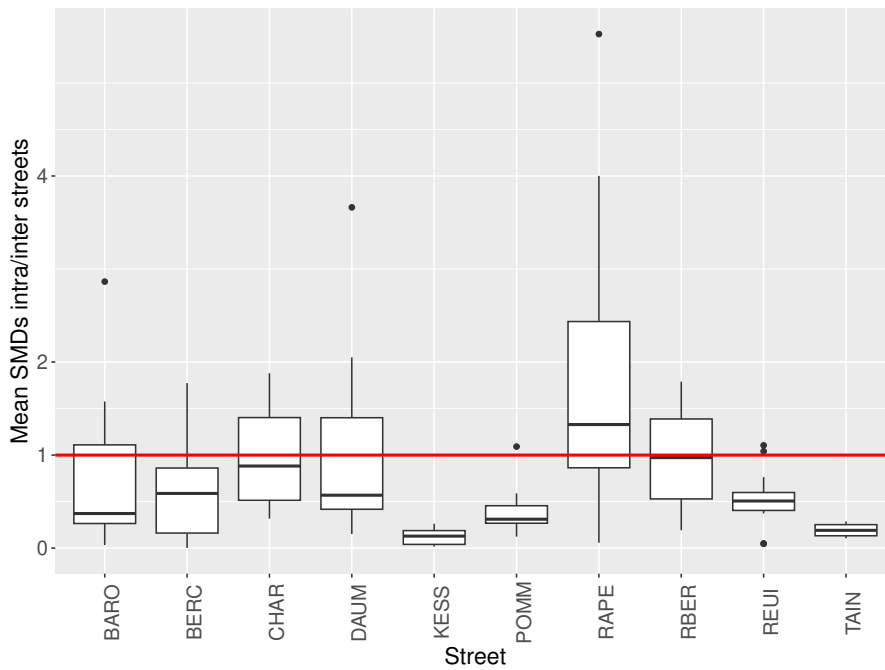
Table A.6: Summary of the results of the correlation tests of the local extinction risk (as quantified by the LER metric and averaged over all streets for each species) with the species traits listed in Table 1. No significant correlation was identified when accounting for multiple testing (using the Holm-Bonferroni method).

Value of H_{inf}	Species
$H_{inf} = 0$	Plantago lanceolata Sisymbrium officinale (*)
$H_{inf} = 1$	Chenopodium album Polygonum aviculare Senecio inaequidens Veronica persica
$H_{inf} = 2$	Parietaria judaica Sonchus oleraceus
$H_{inf} = 3$	Plantago major Poa annua
$H_{inf} = 4$	Lactuca serriola Sisymbrium irio Bromus sterilis Geranium molle Lolium perenne
$H_{inf} = 5$	\
$H_{inf} = 6$	\
$H_{inf} = 7$	Conyza Hordeum murinum Senecio vulgaris Taraxacum
$H_{inf} = 8$	Capsella bursa-pastoris Stellaria media

Table A.7: Value of $H_{inf} = \min\{h \in \llbracket 0, H_{max} \rrbracket : \mathbb{P}(H \leq h | \text{Obs}) \geq 0.05\}$ for each species listed in Table A.2. The posterior distribution of H was obtained by performing parameter inference simultaneously on all portions of streets listed in Table A.2, assuming that only p_{ext} and s differed from one portion of street to another. Species in bold are species for which the absence of a seed bank was identified (that is, for which $\mathbb{P}(H = 0 | \text{Obs}) \geq 0.95$). The asterisk indicates species for which neither the absence (see above) nor the presence (defined as $\mathbb{P}(H \geq 1 | \text{Obs}) \geq 0.95$) of a seed bank was identified.



(a) Distribution of the quotients of the average SMDs inside a street and between streets, grouped by species.



(b) Distribution of the quotients of the average SMDs inside a street and between streets, grouped by street.

Figure A.3: Comparison of the average *Standardised Mean Differences* (SMDs) of the posterior distributions of patch extinction probabilities computed between portions of the same streets or of different streets. The plots correspond to the distribution of the quotients of the mean SMDs inside a street and between streets, grouped by species (a) or by street (b).

We recall that the SMD measures the difference between two probability distributions. Therefore, a quotient smaller than one indicates that posterior distributions of patch extinction probabilities are on average closer between portions of a same street than between portions of different streets, while a quotient larger than one indicates the opposite.

B The BOA Process: Mathematical formulation and estimation procedure

B.1 Formal definition of the BOA process

Formally, the BOA process for a street with N patches observed for T years involves three dependent random variables:

1. The *seed age* $L = (L_{n,t})_{n \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket} \in \mathbb{N}^{N \times T}$.

The random variable $L_{n,t}$ gives the age of the youngest seeds in patch n at the beginning of generation t . These seeds are *viable* (i.e., they can germinate and grow into plants) if and only if their age is below the *maximal dormancy duration* H .

2. The *extinction events* $E = (E_{n,t})_{n \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket} \in \{0, 1\}^{N \times T}$.

We have $E_{n,t} = 1$ if an extinction event occurs in patch n during generation t , and $E_{n,t} = 0$ otherwise.

3. The *observations of standing vegetation* $O = (O_{n,t})_{n \in \llbracket 1, N \rrbracket, t \in \llbracket 1, T \rrbracket} \in \{0, 1\}^{N \times T}$.

We have $O_{n,t} = 1$ if plants are observed in patch n during generation t , after potential patch extinction events and before the seed production step, and we have $O_{n,t} = 0$ otherwise.

At $t = 1$ (initial condition), each patch is randomly chosen to contain viable seeds with probability $s \in [0, 1]$. In that case, the age of the youngest viable seeds is chosen uniformly at random in $\{0, \dots, H\}$. Then, during each generation t , we follow these five steps.

1. GERMINATION STEP: In all patches n such that $L_{n,t} \leq H$ (i.e., in all patches containing viable seeds), some of the viable seeds germinate and grow into plants.
2. EXTINCTION STEP: Each patch containing plants is affected by an extinction event with probability $p_{ext} \in [0, 1]$, independently of other patches. This extinction event kills all the plants in the patch. Formally, the random variables $(E_{n,t})_{1 \leq n \leq N}$ are i.i.d. and follow a Bernoulli distribution with parameter $p_{ext} \in [0, 1]$. The patch n is affected by an extinction event during generation t if, and only if $E_{n,t} = 1$.
3. OBSERVATION STEP: The observer records which patches contain standing vegetation. For each patch n , $O_{n,t}$ is set to 1 if the patch contain plants, and to 0 otherwise.
4. SEED PRODUCTION STEP: For each patch such that $O_{n,t} = 1$, we set

$$L_{n,t+1} = L_{\max(1, n-1), t+1} = L_{\min(N, n+1), t+1} = 0.$$

34 For all remaining non-affected patches n' , we set $L_{n',t+1} = L_{n',t} + 1$.

35 5. All remaining plants die, for instance due to the action of gardeners.

36 The BOA process is thus characterised by three parameters:

- 37 • The initial *proportion of occupied patches* $s \in [0, 1]$,
- 38 • The *maximal dormancy duration* $H \in \mathbb{N}$,
- 39 • The *patch extinction probability* $p_{\text{ext}} \in [0, 1]$.

40 B.2 Bayesian framework used

41 B.2.1 Model likelihood

42 With the above description of the BOA process, we can now write the probability $\mathbb{P}(O, L \mid s, p_{\text{ext}}, H)$
 43 that the observations O and the seed ages L are produced by the process, given that the true parameters
 44 are s, p_{ext} and H . Let us introduce the shortcut notations $O_t = (O_{1,t}, \dots, O_{N,t})$ and $L_t = (L_{1,t}, \dots, L_{N,t})$.

As a first remark, note that the age L_t of the seeds at times $t > 1$ is a deterministic function of the initial age of the seeds L_1 and the previous observations O_1, \dots, O_{t-1} . This function defines the seed production process; it can be written recursively as:

$$L_{n,t+1} = \begin{cases} L_{n,t} + 1 & \text{if } O_{\min(1,n-1),t} = O_{n,t} = O_{\max(N,n+1),t} = 0 \\ 0 & \text{otherwise.} \end{cases}$$

As a consequence, we only need to compute the probability $\mathbb{P}(O, L_1 \mid s, p_{\text{ext}}, H)$, since the remaining values L_2, \dots, L_T are a deterministic function of L_1 and O . This probability can be decomposed as:

$$\mathbb{P}(O, L_1 \mid s, p_{\text{ext}}, H) = \mathbb{P}(L_1 \mid s, p_{\text{ext}}, H) \times \mathbb{P}(O \mid L_1, s, p_{\text{ext}}, H).$$

At time $t = 1$, under the definition of the process introduced in the previous section, the seeds L_1 are either non viable (which we represent with a seed age equal to $H + 1$), or viable, with an age chosen uniformly in $\{0, \dots, H\}$. Therefore, for all $1 \leq n \leq N$, $L_{n,1} \in \{0, \dots, H + 1\}$ and

$$\mathbb{P}(L_{n,1} = \ell \mid s, p_{\text{ext}}, H) = \begin{cases} \frac{s}{H+1} & \text{if } \ell \leq H \\ 1 - s & \text{if } \ell = H + 1. \end{cases}$$

As $L_{1,1}, \dots, L_{N,1}$ are independent, we can then compute $\mathbb{P}(L_1 | s, p_{\text{ext}}, H)$ as

$$\mathbb{P}(L_1 | s, p_{\text{ext}}, H) = \mathbb{P}(L_{1,1} | s, p_{\text{ext}}, H) \times \dots \times \mathbb{P}(L_{N,1} | s, p_{\text{ext}}, H).$$

The probability $\mathbb{P}(O | L_1, s, p_{\text{ext}}, H)$ can be decomposed under the form

$$\mathbb{P}(O | L_1, s, p_{\text{ext}}, H) = \prod_{t=1}^T \mathbb{P}(O_t | L_t, s, p_{\text{ext}}, H).$$

This equation expresses the fact that the random variables O_t only depends on the quantity L_t (which is itself a function of O_1, \dots, O_{t-1} , as described above). Then, each term in this equation can be expressed separately, as follows:

$$\mathbb{P}(O_{n,t} = 1 | L_t, s, p_{\text{ext}}, H) = \begin{cases} 1 - p_{\text{ext}} & \text{if } L_{n,t} \leq H \\ 0 & \text{otherwise.} \end{cases}$$

45 This equation expresses the fact that plants can only be observed in patch n during generation t
 46 if 1) $L_t \leq H$ (i.e., if the youngest seeds present in the patch are viable) and 2) if no extinction
 47 event occurred (which happens with probability $1 - p_{\text{ext}}$). As $O_{n,t}$ is $\{0, 1\}$ -valued, we also obtain
 48 $\mathbb{P}(O_{n,t} = 0 | L_t, s, p_{\text{ext}}, H)$ as $1 - \mathbb{P}(O_{n,t} = 1 | L_t, s, p_{\text{ext}}, H)$.

A consequence of the above observation is that conditionally on L_t , the random variables $O_{1,t}, \dots, O_{N,t}$ are independent. Therefore, we can compute $\mathbb{P}(O_t | L_t, s, p_{\text{ext}}, H)$ as

$$\mathbb{P}(O_t | L_t, s, p_{\text{ext}}, H) = \mathbb{P}(O_{1,t} | L_t, s, p_{\text{ext}}, H) \times \dots \times \mathbb{P}(O_{N,t} | s, p_{\text{ext}}, H).$$

49 We can combine these expressions to obtain the complete likelihood of the model, which is given
 50 by $\mathbb{P}(O, L_1 | s, p_{\text{ext}}, H)$. Note that this quantity is different from the marginal likelihood of the
 51 observations $\mathbb{P}(O | s, p_{\text{ext}}, H)$. The latter would be ideally the one to use in practice (since L_1 is
 52 unknown), but is unfortunately impossible to compute: formally, it writes as sum over all the $(H+2)^N$
 53 possible configurations of seed ages at time $t = 1$, and the number of such configurations grows
 54 exponentially with N . As a consequence, we need to estimate L_1 along with the model parameters
 55 s, p_{ext}, H .

56 B.2.2 Bayesian methodology

57 Since the observed data only spans a period of ten years, our estimation necessarily comprises a certain
 58 amount of uncertainty. In the context of complex hierarchical statistical models (of which the model

59 considered here is an example), Maximum Likelihood estimations only give the most likely value for
60 s , p_{ext} and H , and cannot be used to obtain confidence regions around the estimated values. For this
61 reason, methods like the EM are not suited to our study.

Instead, we propose to work in a Bayesian framework. It consists in determining the posterior distribution $\mathbb{P}(s, p_{\text{ext}}, H \mid O)$ of the parameters s, p_{ext}, H given the observed data O using Bayes' rule:

$$\mathbb{P}(s, p_{\text{ext}}, H \mid O) = \frac{\mathbb{P}(O \mid s, p_{\text{ext}}, H)\mathbb{P}(s, p_{\text{ext}}, H)}{\mathbb{P}(O)}.$$

As we will see in the following sections, this formula can be used to draw samples from the distribution $\mathbb{P}(s, p_{\text{ext}}, H \mid O)$, even though it may not be computed explicitly in practice. These samples are then used to approximate the distribution. For instance, if we have M samples $(s^{(m)}, p_{\text{ext}}^{(m)}, H^{(m)})_{1 \leq m \leq M}$ of the posterior distribution, the expectation of p_{ext} given O can be computed with a Monte-Carlo approximation:

$$\mathbb{E}[p_{\text{ext}} \mid O] \simeq \bar{p}_{\text{ext}} = \frac{1}{M} \sum_{m=1}^M p_{\text{ext}}^{(m)}.$$

Similarly, the uncertainty on p_{ext} can be measured by computing its posterior *variance*

$$\text{Var}(p_{\text{ext}} \mid O) \simeq \frac{1}{M} \sum_{m=1}^M (p_{\text{ext}}^{(m)} - \bar{p}_{\text{ext}})^2.$$

62 **Prior distribution** In a Bayesian framework, the model parameters s, p_{ext} and H are considered
63 as random variables. In order to apply Bayes' rule and obtain the posterior distribution that we are
64 interested in, their distribution *a priori* $\mathbb{P}(s, p_{\text{ext}}, H)$ must be defined in a way that reflects our (prior)
65 knowledge of the model parameters. In this paper, we do not make any initial assumption on the
66 model parameters, and consider the following simple, uninformative prior distributions:

- 67 • The initial proportion of occupied patches s follows a uniform distribution over $[0, 1]$.
- 68 • Similarly, p_{ext} also follows a uniform distribution over $[0, 1]$.
- The maximal dormancy duration H follows a uniform distribution over the set of integers $\{0, \dots, H_{\text{max}}\}$:

$$\mathbb{P}(H) = \begin{cases} 1/(H_{\text{max}} + 1) & \text{if } H \in \{0, \dots, H_{\text{max}}\} \\ 0 & \text{otherwise.} \end{cases}$$

69 The pre-defined upper bound on the maximal dormancy duration is a consequence of the fact that with
70 a limited number of years of observation, identifying long maximal dormancy durations (compared to
71 the length of the observation window) can be an ill-posed problem and lead to identifiability issues.

72 We can then combine the different priors to obtain the complete prior distribution.

73 **B.2.3 Markov Chain Monte-Carlo sampler**

In order to compute the GER and the MaxGER extinction metrics, we need samples from the posterior distribution $\mathbb{P}(p_{\text{ext}}, H \mid O)$. As explained earlier, this distribution takes an intractable form: using the law of total probabilities it writes as

$$\mathbb{P}(p_{\text{ext}}, H \mid O) = \sum_{L_1 \in \{0, \dots, H+1\}^N} \int_0^1 \frac{\mathbb{P}(O, L_1, s, p_{\text{ext}}, H)}{\mathbb{P}(O)} ds,$$

which is a sum over an exponential number of terms. We overcome this hurdle by focusing instead on the posterior distribution of $(L_1, s, p_{\text{ext}}, H)$ given O , which writes as:

$$\mathbb{P}(L_1, s, p_{\text{ext}}, H \mid O) = \frac{\mathbb{P}(O, L_1, s, p_{\text{ext}}, H)}{\mathbb{P}(O)}.$$

74 Since this probability distribution is known up to a normalising constant, the method presented in
 75 the next section can be used to draw samples from it. This method will provide us with samples
 76 $(L_1^{(m)}, s^{(m)}, p_{\text{ext}}^{(m)}, H^{(m)})_{1 \leq m \leq M}$, which then give the marginal distribution samples $(p_{\text{ext}}^{(m)}, H^{(m)})_{1 \leq m \leq M}$
 77 of the posterior distribution $\mathbb{P}(p_{\text{ext}}, H \mid O)$ that we are interested in.

78 In practice, we will use a Markov Chain Monte Carlo (MCMC) sampler to draw approximate
 79 samples from the posterior distribution of $\mathbb{P}(L_1, s, p_{\text{ext}}, H \mid O)$. Such methods rely on simple and fast
 80 iterations which generate a Markov chain whose invariant distribution is the target distribution we wish
 81 to sample from. After a certain amount of iterations, the samples of the Markov chain are very close
 82 from being true samples of the target distribution; they can thus be used in Monte-Carlo estimators
 83 to compute the GER and MaxGER metrics. Note that consecutive MCMC samples are correlated
 84 by definition, which increases the variance of Monte-Carlo estimations compared with independent
 85 samples, but does not affect the expectation of the estimator, which is unbiased as the number of
 86 samples grows large.

87 **B.2.4 Metropolis-Hastings within Gibbs sampler**

88 In this section, we detail the MCMC implemented in practice. For an introduction to classical MCMC
 89 methods, we refer the reader to (Robert and Casella, 2010). In this paper, we use the Metropolis-
 90 Hastings within Gibbs (MHwG) sampling algorithm, which is designed to draw samples from a general
 91 probability distribution $\pi(x_1, \dots, x_d)$ in a high-dimensional space. The general MHwG procedure is
 92 recalled in Algorithm 1. In the algorithm, x_{-k} denotes the vector x where component k has been

93 removed: the term $\pi(x_k | x_{-k})$ thus refers to the conditional distribution of x_k given the remaining
 94 components of x . The algorithm mainly relies on the so-called proposal distributions q_k^m : at each step
 95 m , the distribution q_k^m proposes, given the m -th value of x_{-k} , a random candidate for the next value
 96 of x_k . Then, this candidate is accepted as the next value of x_k with probability α .

Algorithm 1: The Metropolis-Hastings within Gibbs algorithm

input : Target distribution $\pi(x_1, \dots, x_d)$; number of samples M
 Initialize x_1^0, \dots, x_d^0
for $m = 1$ **to** M **do**
 for $k = 1$ **to** d **do**
 Sample a candidate variable y_k^m from a *proposal distribution* q_k^m .
 Define the acceptance ratio $\alpha = \min \left[1, \frac{q_k^m(y_k^m)\pi(x_k^{m-1}|x_{-k}^{m-1})}{q_k^m(x_k^{m-1})\pi(y_k^m|y_{-k}^m)} \right]$
 Sample $B_k^m \sim \text{Bernoulli}(\alpha)$.
 If $B_k^m = 0$, set $x_k^m = x_k^{m-1}$. Otherwise, set $x_k^m = y_k^m$.
 end
end
return $(x^m)_{1 \leq m \leq M}$

97 **MHWG for the BOA process** In this paper, we are interested in the distribution $\mathbb{P}(L_1, s, p_{\text{ext}}, H | O)$.
 98 In the MHWG algorithm, we will thus alternatively be sampling from s, p_{ext}, H and the components of
 99 L_1 . Since L_1 has N components, at each step of the algorithm we only update a single coordinate $L_{n,1}$
 100 for a patch n chosen at random. This allows spending more time on sampling the model parameters
 101 s, p_{ext}, H rather than the seed ages $L_{n,1}$, which in practice only contribute marginally to the likelihood.
 102 In order to obtain a complete algorithmic procedure, we need to specify the proposal distributions q_k^m
 103 used at each step.

- 104 • For H : the new value of H is chosen uniformly at random in $\{0, \dots, H_{\text{max}}\}$.
- 105 • For $L_{n,1}$: the random value is sampled uniformly in $\{0, \dots, H + 1\}$. Values between 0 and H
 106 produce viable seeds, and the value $H + 1$ produces non-viable seeds.
- 107 • For p_{ext} and s : instead of working with p_{ext} and s , we represent them as $p_{\text{ext}} = h(\xi_p)$ and
 108 $s = h(\xi_s)$, with $h(x) = (1 + \exp(-x))^{-1} \in]0, 1[$ a sigmoid function. In the algorithm, the proposal
 109 value for ξ_p at step m is drawn from a Gaussian distribution $\mathcal{N}(\xi_p^{m-1}, \sigma_p^2)$. Similarly, the proposal
 110 value for ξ_s is drawn from a Gaussian distribution $\mathcal{N}(\xi_s^{m-1}, \sigma_s^2)$. The corresponding values for
 111 p_{ext} and s are then obtained by applying h . The sigmoid representation allows sampling values
 112 very close to 0 or 1 as well as more centered values using a common proposal variance σ_p^2 or σ_s^2 .

113 In the form we just described, the algorithm currently has a practical issue: a change in the value
 114 of H can change the initial state of some seeds from viable to non-viable, and conversely. Although

115 this is not a problem in itself, we noticed that it hinders the convergence in practice. Our intuition is
 116 that it is due to the fact that frequent changes in the value of H prevent the distribution of L_1 from
 117 converging, as the viability of the seeds varies not only with the value of $L_{n,1}$, but also with H .

We overcome this issue by defining a duplicate L_1^h of L_1 for each value of $h \in \{0, \dots, H_{\max}\}$. For each h , we impose that $L_1^h \in \{0, h + 1\}$; the distribution of L_1 given $H = h$ corresponds to L_1^h . In other words, by definition:

$$L_1 = \sum_{h=1}^{H_{\max}} \mathbb{1}_{\{H=h\}} L_1^h.$$

118 In practice, at each step of the MHwG algorithm, a random patch n is selected, and the variables
 119 $L_{n,1}^0, \dots, L_{n,1}^{H_{\max}}$ are updated. Theoretically, these variables should be updated using the conditional
 120 distribution $\mathbb{P}(L_{n,1}^h \mid s = s^{(m)}, p_{\text{ext}} = p_{\text{ext}}^{(m)}, H = H^{(m)}, O)$. However, it must be noted that, if $h \neq H^{(m)}$,
 121 the value of L_1^h does not play a role in the distribution of O . In other words, the variables L_1^h and
 122 O are independent conditionally on $H = H^{(m)}$, and the conditional distribution of L_1^h thus simplifies
 123 to a fixed base distribution over $\{0, \dots, h + 1\}$. As a consequence, if $h \neq H^{(m)}$, the variable L_1^h
 124 is randomised to a distribution that does not depend on O . In order to bypass this hurdle, we
 125 change the conditional distribution $\mathbb{P}(L_{n,1}^h \mid s = s^{(m)}, p_{\text{ext}} = p_{\text{ext}}^{(m)}, H = H^{(m)}, O)$ in the MHwG to
 126 $\mathbb{P}(L_{n,1}^h \mid s = s^{(m)}, p_{\text{ext}} = p_{\text{ext}}^{(m)}, H = h, O)$. This modification results in an approximate MHwG
 127 procedure, which provides significantly better performances than the base algorithm.

128 As a last point, the multiple variables $L_1^0, \dots, L_1^{H_{\max}}$ must be modelled by distinct values of s : the
 129 proportion of viable seeds may vary depending on the value of h . In practice, s is thus replaced with
 130 a vector $(s^0, \dots, s^{H_{\max}}) \in [0, 1]^{H_{\max}+1}$, and the probability distribution of L_1^h is defined using s^h .

131 *Remark B.1.* The variances σ_p^2 and σ_s^2 for the proposal distributions of ξ_p and ξ_s are tuned throughout
 132 the convergence of the MCMC in order to obtain an average proportion of accepted samples around
 133 30%. Larger variances lead to larger transitions, which are thus rejected more often; smaller variances
 134 lead to smaller transitions, which are easier to accept.

135 B.3 The noisy BOA process

The noisy BOA process adds a new parameter $\varepsilon \in [0, 1]$, which controls the probability of what can be interpreted as external colonisation (but was introduced to model noise in observations). Compared to the standard BOA process, patches containing non-viable seeds or affected by an extinction event can now contain plants with probability ε . This only changes the probability $\mathbb{P}(O_{n,t} = 1 \mid L_t, s, p_{\text{ext}}, H, \varepsilon)$:

it now expresses as

$$\mathbb{P}(O_{n,t} = 1 \mid L_t, s, p_{\text{ext}}, H, \varepsilon) = \begin{cases} 1 - p_{\text{ext}}(1 - \varepsilon) & \text{if } L_{n,t} \leq H \\ \varepsilon & \text{otherwise.} \end{cases}$$

As with s , p_{ext} and H , the posterior distribution of the parameter ε can be sampled from with the MHwG algorithm. As the parameter ε was introduced to provide a buffer against noise in the dataset rather than to model an actual biological phenomenon, we expect it to take very small values. Therefore, we take a *slack and slab* prior distribution for ε , that is to say, a mixture of two uniform distributions:

$$\varepsilon \sim \frac{1}{2}\mathcal{U}([0, \varepsilon_m]) + \frac{1}{2}\mathcal{U}([0, \varepsilon_M]).$$

136 In practice, we choose $\varepsilon_m \ll \varepsilon_M$, e.g., $\varepsilon_m = 1\%$ and $\varepsilon_M = 5\%$. The idea is to obtain a prior distribution
137 concentrated around relatively small values, with a spike on very small values.

138 We perform Metropolis transitions on the inverse sigmoid of ε , $\xi_\varepsilon = h^{-1}(\varepsilon)$, with a Gaussian
139 proposal with variance a σ_ε^2 tuned adaptively throughout the MCMC convergence.

140 B.4 Handling multiple streets

Finally, the method introduced in the previous sections seamlessly transposes to the case of observations of a given species in multiple streets, denoted O_1, \dots, O_K . We assume that for each species, the maximum dormancy duration H and the noise intensity ε are identical across all streets, but the initial proportion of occupied patches s_k and the extinction probability $(p_{\text{ext}})_k$ depend the street $1 \leq k \leq K$. As in the previous case, we are interested in sampling from the distribution

$$\mathbb{P}((p_{\text{ext}})_1, \dots, (p_{\text{ext}})_K, H, \varepsilon \mid O_1, \dots, O_K),$$

and, as in the previous case, we tackle this problem by instead sampling from the complete posterior distribution

$$\mathbb{P}((L_1, s, p_{\text{ext}})_1, \dots, (L_1, s, p_{\text{ext}})_K, H, \varepsilon \mid O_1, \dots, O_K),$$

which, up to the normalising constant $\mathbb{P}(O_1, \dots, O_K)$, is proportional to the function

$$\mathbb{P}(O_1, \dots, O_K, (L_1, s, p_{\text{ext}})_1, \dots, (L_1, s, p_{\text{ext}})_K, H, \varepsilon).$$

141 The MHwG can then be transposed to the multiple streets setting, by updating H , ε and the param-
142 eters $(L_1, s, p_{\text{ext}})_k$ of each street k at each MHwG step.

143 Using multiple streets at the same time allows reducing the uncertainty on the parameters H
144 and ε , which benefit from the combination of all the observations; this is verified in our numerical
145 experiments (see Supporting Information C.4).

146 **C Assessment of the performances of the estimation procedure**

147 **C.1 Parameter sets**

Parameter	Values
Number of patches N	50, 100
Number of years of observation T	5, 10
Initial proportion of occupied patches s	0.2, 0.8
Patch extinction probability p_{ext}	0.1, 0.2, ..., 0.9
Maximal dormancy duration H	0, 1, 2, 5
Noise intensity ϵ	0.00, 0.01, 0.02, 0.05
<i>Additional parameter sets - $H = 0$</i>	
Patch extinction probability p_{ext}	$p_c(0) + x$, $x = -0.1, -0.08, \dots, 0.08, 0.1$
<i>Additional parameter sets - $H = 1$</i>	
Patch extinction probability p_{ext}	$p_c(1) + x$, $x = -0.1, -0.08, \dots, 0.08, 0.1$
<i>Additional parameter sets - $H = 2$</i>	
Patch extinction probability p_{ext}	$p_c(2) + x$, $x = -0.1, -0.08, \dots, 0.08, 0.1$
<i>Additional parameter sets - $H = 5$</i>	
Patch extinction probability p_{ext}	$p_c(5) + x$, $x = -0.1, -0.08, \dots, 0.08, 0.1$

Table C.1: Parameter sets used to compare the performances of the MaxGER and GER metrics. For each parameter set, we simulated 30 BOA processes, and computed the average MaxGER and GER over the 30 simulated processes.

Parameter	Values
Number of patches N	50
Number of years of observation T	10
Initial proportion of occupied patches s	0.2, 0.8
Patch extinction probability p_{ext}	0.1, 0.35, 0.55, 0.75
Maximal dormancy duration H	0, 1, 2, 5
<i>Additional parameter - False positives</i>	
False positive rate ϵ_{pos}	0.00, 0.01, 0.02, 0.05
<i>Additional parameter - False negatives</i>	
False negative rate ϵ_{neg}	0.00, 0.01, 0.02, 0.05
<i>Additional parameter - External colonization</i>	
External colonization rate ϵ_{col}	0.00, 0.01, 0.02, 0.05

Table C.2: Parameter sets used to compare the performances of the MaxGER and GER metrics on corrupted datasets (containing false positives, false negatives or external colonization). For each parameter set, we simulated 30 BOA processes and computed the average MaxGER and GER over the 30 simulated processes.

Parameter	Values
Number of patches N	30, 50
Number of years of observation T	5, 10
Initial proportion of occupied patches s	0.2, 0.8
Patch extinction probability p_{ext}	0.1, 0.35, 0.55, 0.75
Maximal dormancy duration H	0, 1, 2, 5
Noise intensity ϵ	0, 0.01, 0.02, 0.05
Number of streets M	1, 2, 5, 10

Table C.3: Parameter sets used to assess the effect of performing the estimation simultaneously on multiple streets on the quality of the estimation of H . For each parameter set, we generated 30 simulations and performed parameter inference assuming that H and ϵ were constant across streets.

148 C.2 Comparison of the performances of the GER and MaxGER metrics

149 C.2.1 Performances of the GER metric

150 The following figures show the evolution of the GER metric as a function of the patch extinction
151 probability p_{ext} . For each parameter set listed in Table C.1, we simulated 30 datasets and performed
152 parameter inference under a noisy BOA process. We then computed the average GER across the 30
153 simulations. The black vertical line indicates the critical patch extinction probability $p_c(H)$.

$$(N, T) = (50, 5)$$

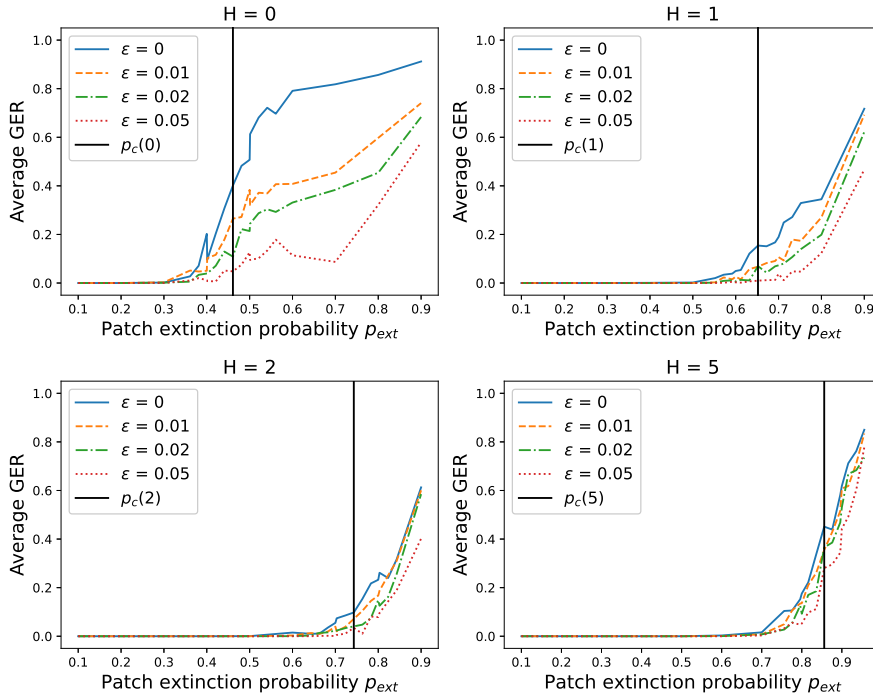


Figure C.1

$$(N, T) = (50, 10)$$

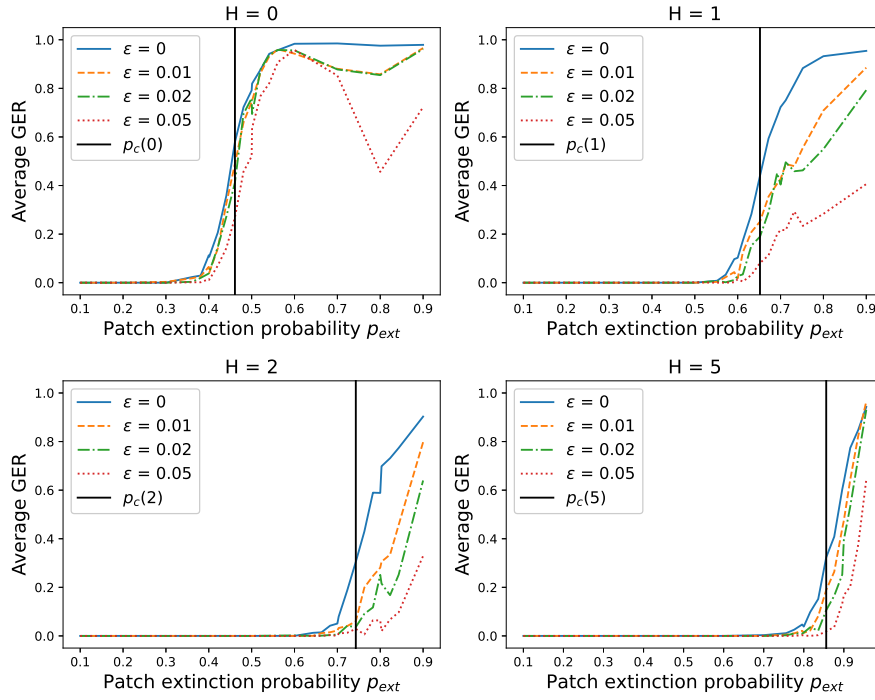


Figure C.2

$$(N, T) = (100, 5)$$

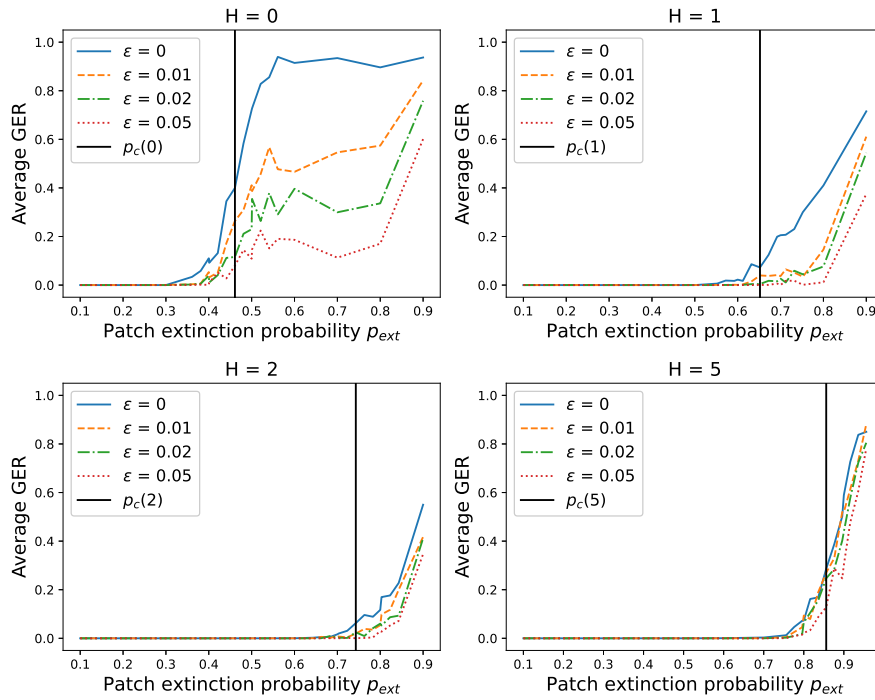


Figure C.3

$(N, T) = (100, 10)$

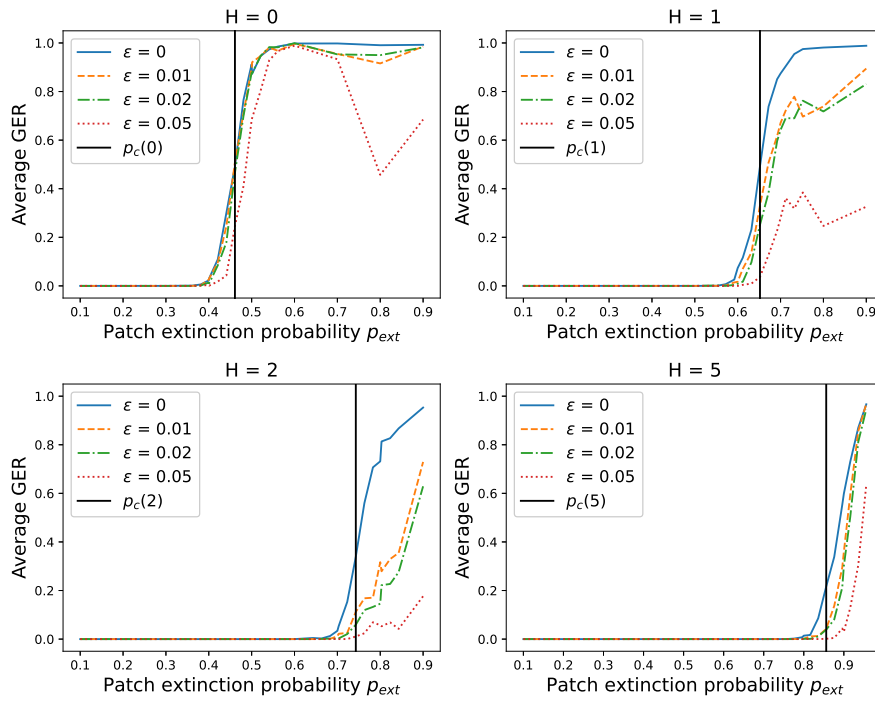


Figure C.4

154 **C.2.2 Performances of the MaxGER metric**

155 The following figures show the evolution of the MaxGER metric as a function of the patch extinction
 156 probability p_{ext} . For each parameter set listed in Table C.1, we simulated 30 datasets and performed
 157 parameter inference under a noisy BOA process. We then computed the average GER across the 30
 158 simulations. The black vertical line indicates the critical patch extinction probability $p_c(H)$.

$$(N, T) = (50, 5)$$

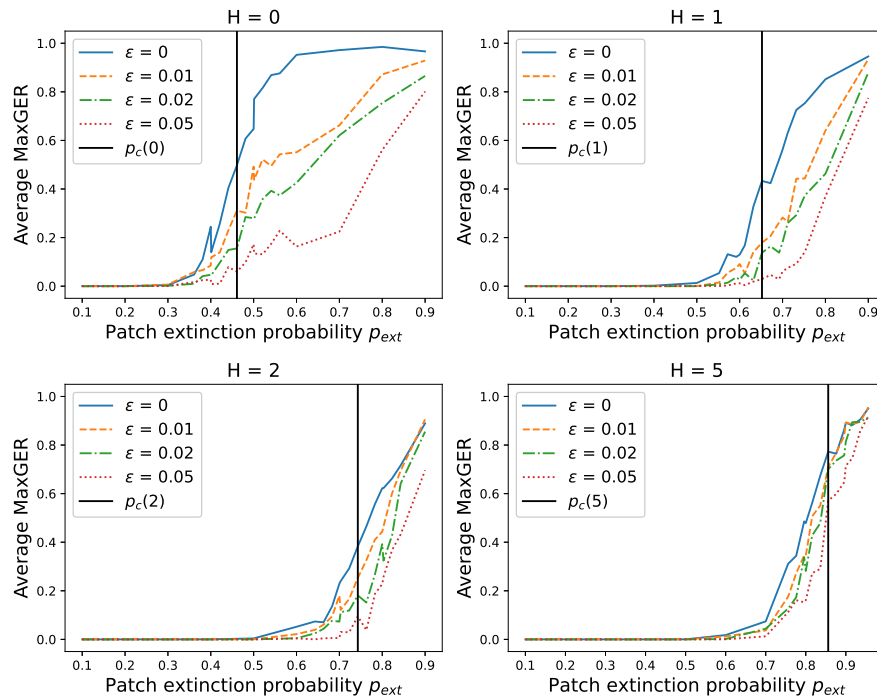


Figure C.5

$(N, T) = (50, 10)$

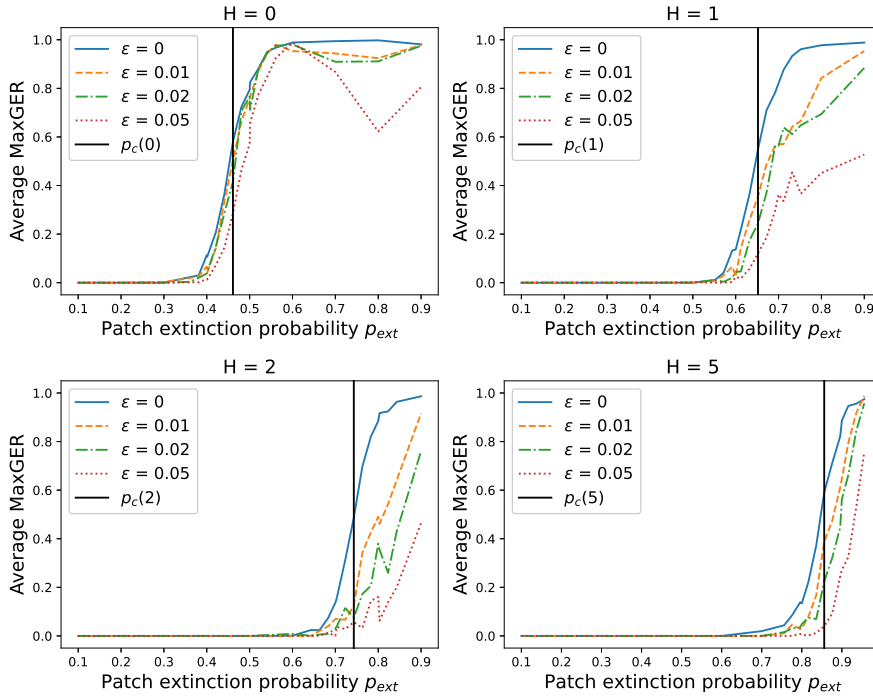


Figure C.6

$(N, T) = (100, 5)$

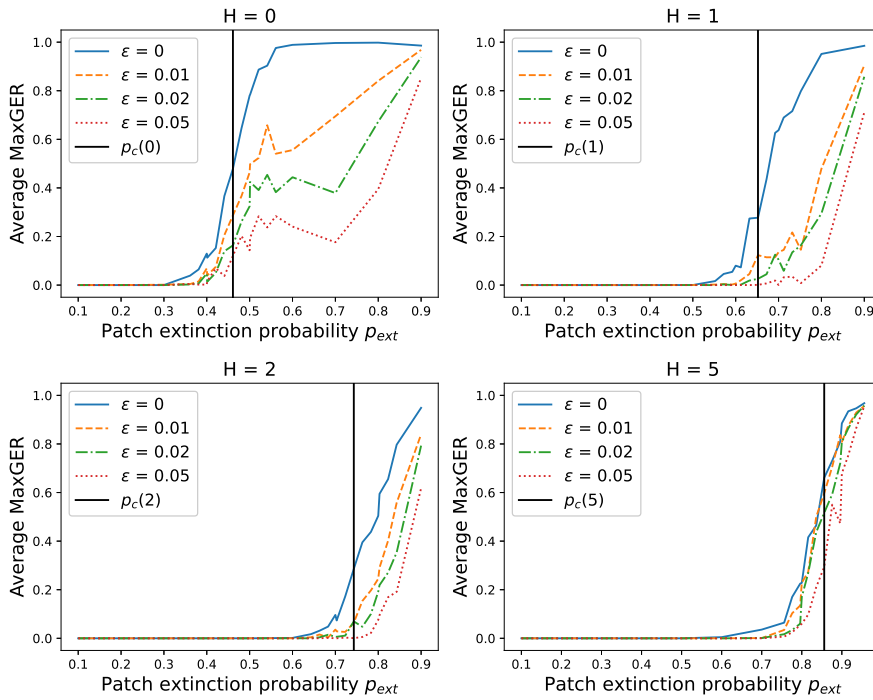


Figure C.7

$(N, T) = (100, 10)$

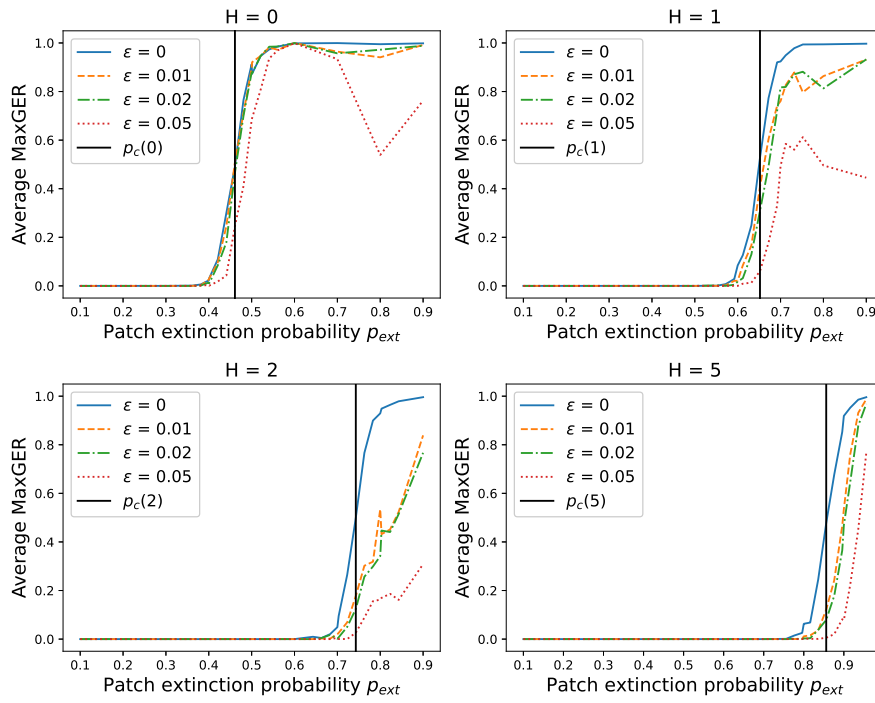


Figure C.8

159 **C.2.3 Performances of the MaxGER metric on corrupted datasets**

160 The effect of the introduction of false negatives on the assessment of the MaxGER metric is mostly vis-
 161 ible when $H = 0$, for high patch extinction probabilities and when parameter estimation is performed
 162 under a BOA process without noise.

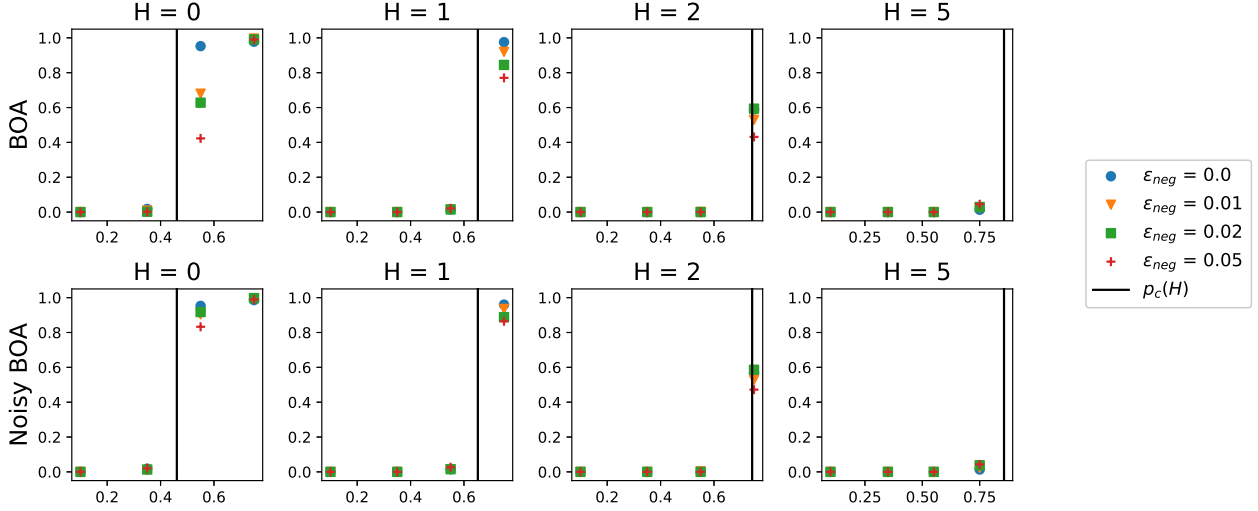


Figure C.9: Effect of the introduction of false negatives on the assessment of the MaxGER metric, when estimation is performed under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. For each parameter set listed in Table C.2, we simulated 30 datasets and performed inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. The black vertical line indicates the critical patch extinction probability $p_c(H)$.

163 The introduction of false positives has a very different outcome depending on whether the estima-
 164 tion is performed under a BOA or noisy BOA process. This is particularly visible when $H = 0$ or 1
 165 and for high patch extinction probabilities.

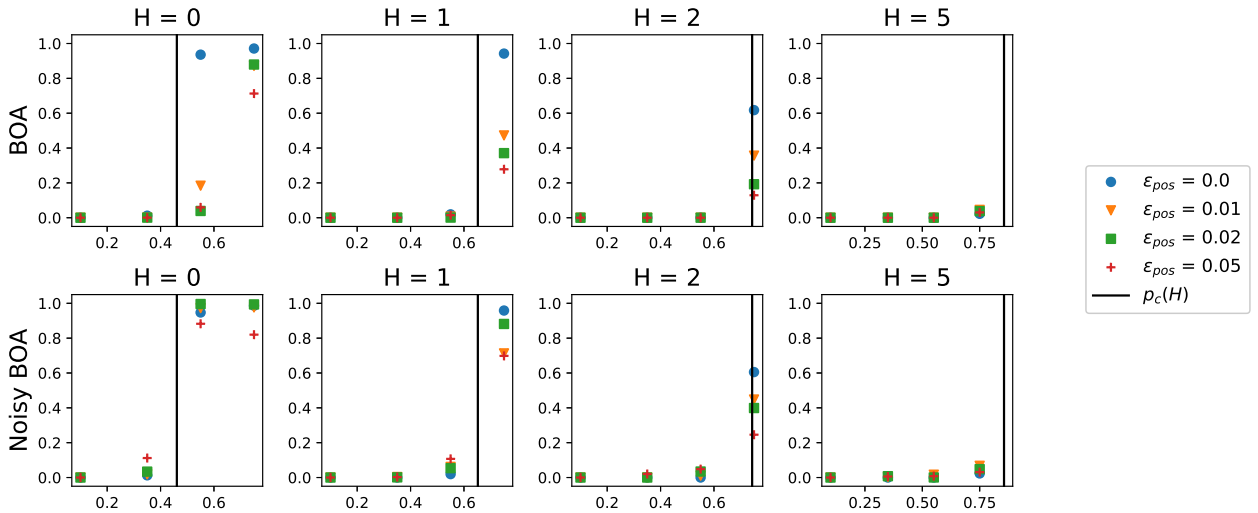


Figure C.10: Effect of the introduction of false positives on the assessment of the MaxGER metric, when estimation is performed under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. For each parameter set listed in Table C.2, we simulated 30 datasets and performed inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. The black vertical line indicates the critical patch extinction probability $p_c(H)$.

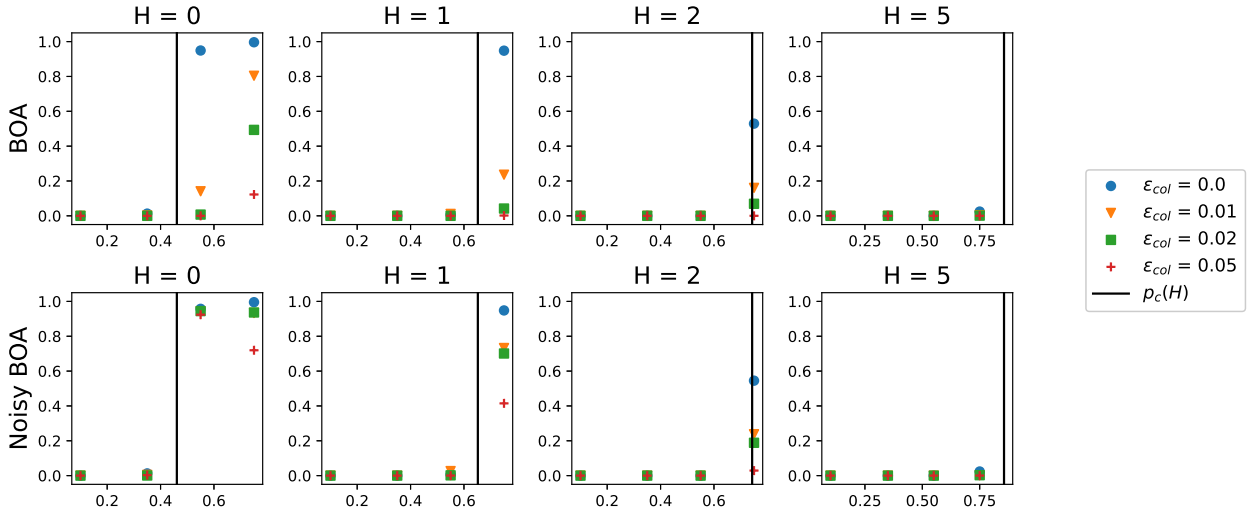


Figure C.11: Effect of the introduction of external colonization on the assessment of the MaxGER metric, when estimation is performed under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. For each parameter set listed in Table C.2, we simulated 30 datasets and performed inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. The black vertical line indicates the critical patch extinction probability $p_c(H)$.

167 **C.3 Performances of the estimation of p_{ext} and H**

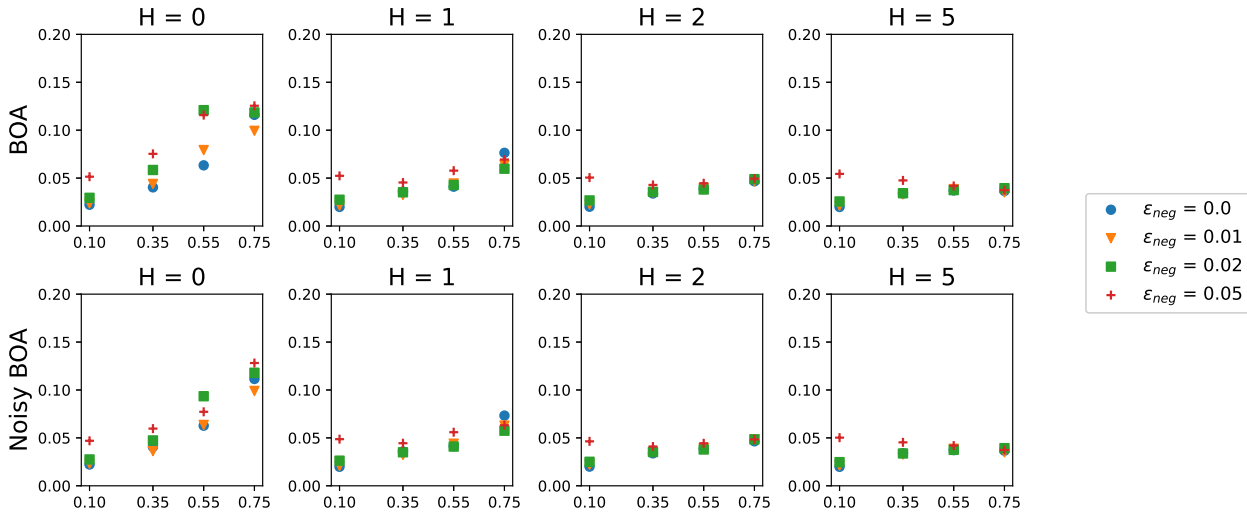


Figure C.12: Root Mean Square Error (RMSE) on the estimation of p_{ext} in the presence of false negatives, when performing parameter inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. For each parameter set listed in Table C.2, we simulated 30 datasets and performed inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$.

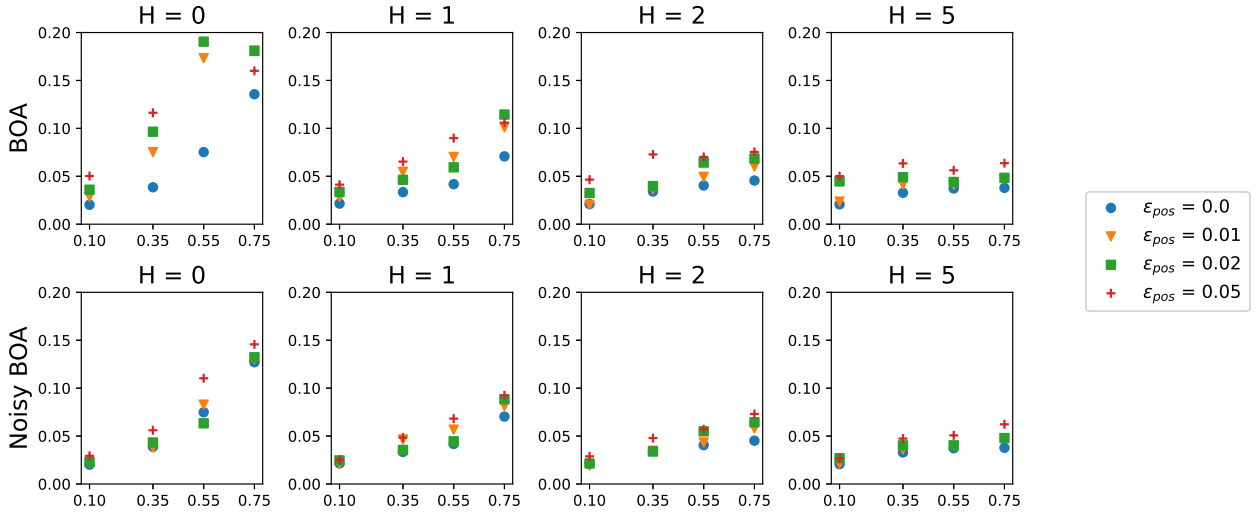


Figure C.13: Root Mean Square Error (RMSE) on the estimation of p_{ext} in the presence of false positives, when performing parameter inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. For each parameter set listed in Table C.2, we simulated 30 datasets and performed inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$.

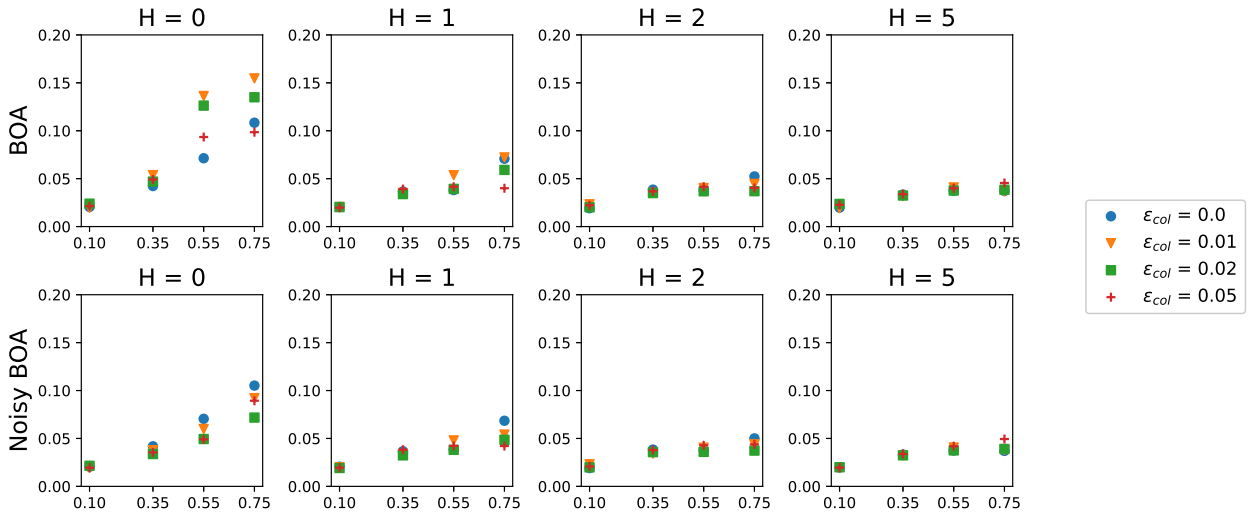


Figure C.14: Root Mean Square Error (RMSE) on the estimation of p_{ext} in the presence of external colonization, when performing parameter inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. For each parameter set listed in Table C.2, we simulated 30 datasets and performed inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$.

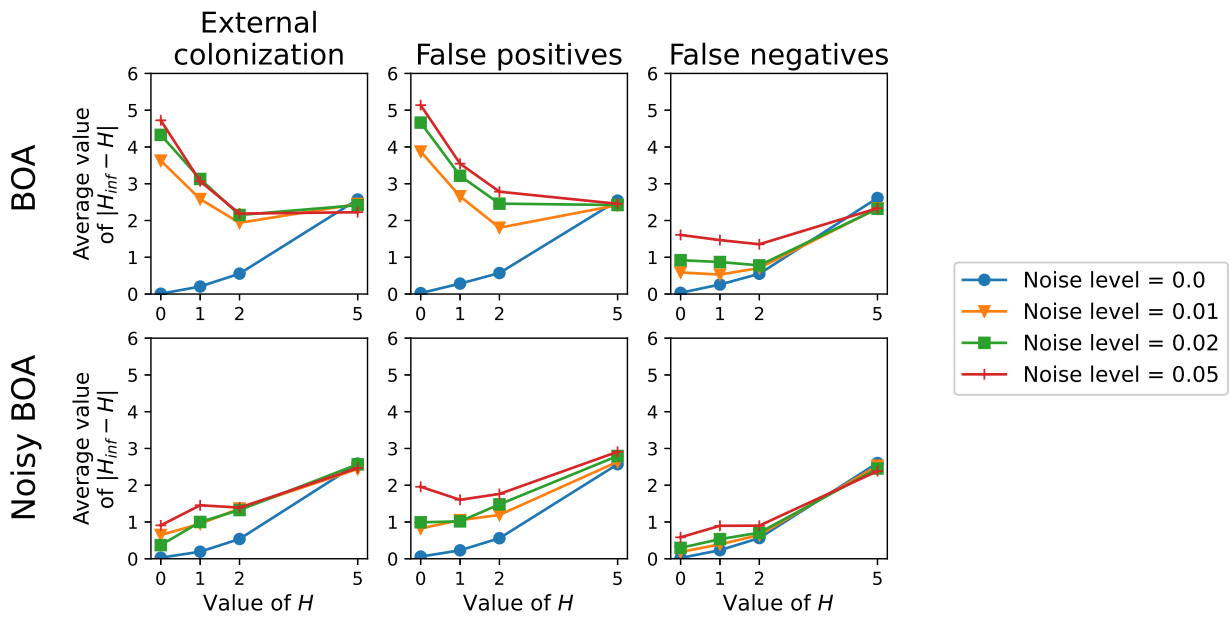


Figure C.15: Average value of $|H_{inf} - H|$ in the presence of false negatives, false positives or external colonization, when performing parameter inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$. For each parameter set listed in Table C.2, we simulated 30 datasets and performed inference under a BOA process or a noisy BOA process with $\epsilon_{max} = 0.1$.

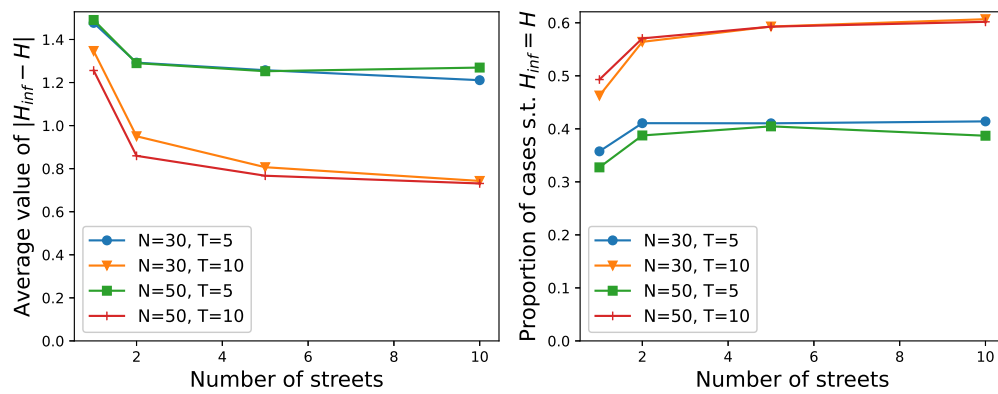


Figure C.16: Effect of performing the estimation simultaneously on multiple streets on the quality of the estimation of H , when the maximal dormancy duration H and the noise intensity ϵ are constant across streets. The parameter sets used are listed in Table C.3.

169 References

- 170 Gargominy, O., Terceirie, S., Régnier, C., Ramage, T., Schoelinck, C., Dupont, P., Vandel, E., Daszkiewicz,
 171 P., & Poncet, L. (2014). Taxref v8. 0, référentiel taxonomique pour la france: Méthodologie,
 172 mise en oeuvre et diffusion. *Rapport SPN*, 42, 2014.
- 173 Robert, C., & Casella, G. (2010). *Monte Carlo Statistical Methods*. Springer.