Ecography

Supplementary material
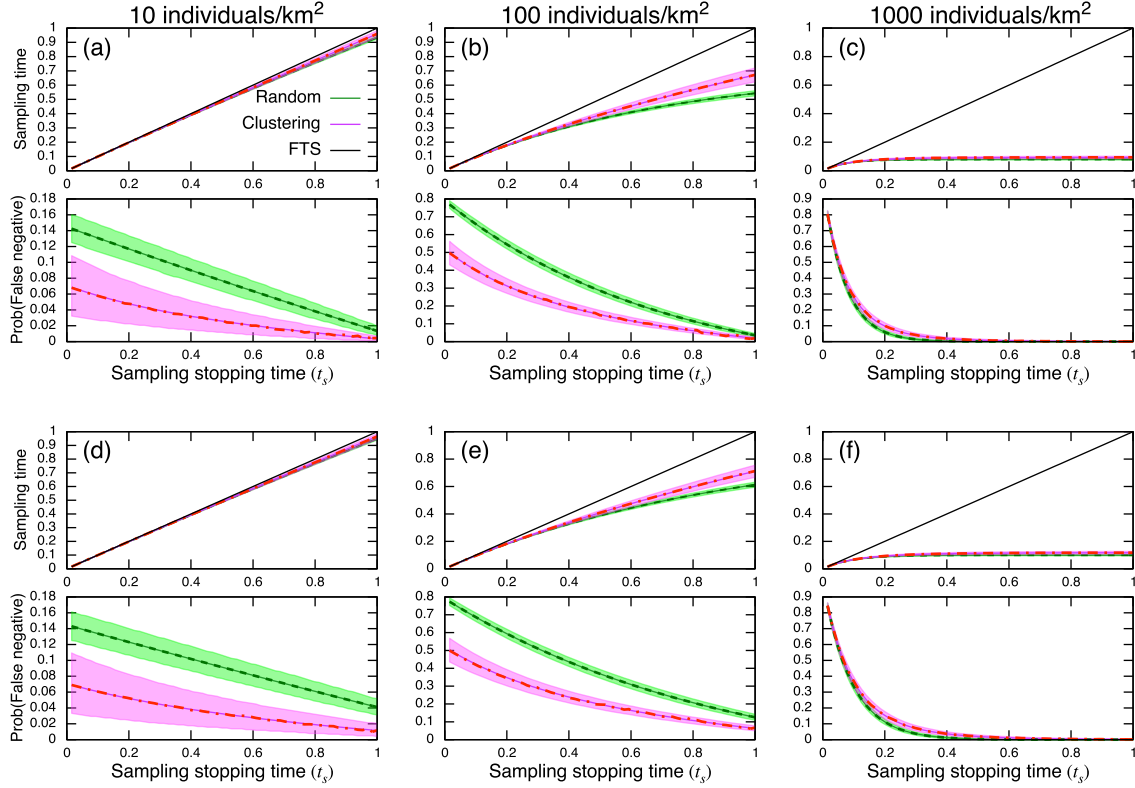
Supporting Information

Figures

Figure S.1: Normalized sampling time and false negative given normalized stopping time t_s with the probability of sampling error $\epsilon = 0.1$ (a-c), and $\epsilon = 0.3$ (d-f). For each individual distribution scenarios, the numerical average (lines) and its theoretical value (dashed) are provided. The shaded area are between 5 and 95 percentiles of a 10^5-time numerical simulation. FTS represents the fixed-time survey. The scales of surveys are are $\nu(W) = 4\text{km} \times 4\text{km}$, $\nu(M) = 2^{-3}\text{km} \times 2^{-3}\text{km}$, $\nu(S) = 2^{-6}\text{km} \times 2^{-6}\text{km}$. The parameters for the points generations are the same as in Fig. A.1.
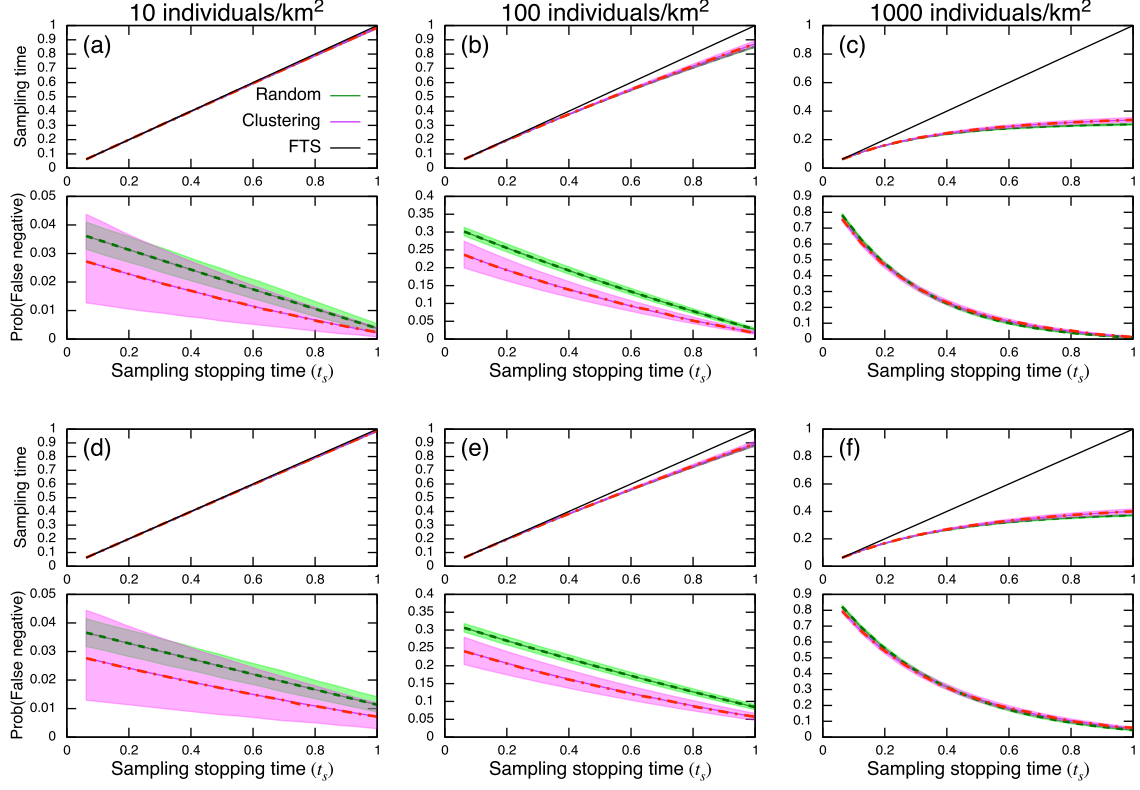
Figure S.2: Normalized sampling time and false negative given normalized stopping time t_s with the probability of sampling error $\epsilon = 0.1$ (a-c), and $\epsilon = 0.3$ (d-f). For each individual distribution scenarios, the numerical average (lines) and its theoretical value (dashed) are provided. The shaded area are between 5 and 95 percentiles of a 10^5-time numerical simulation. FTS represents the fixed-time survey. The scales of surveys are are $\nu(W) = 4\text{km} \times 4\text{km}$, $\nu(M) = 2^{-4}\text{km} \times 2^{-4}\text{km}$, $\nu(S) = 2^{-6}\text{km} \times 2^{-6}\text{km}$. The parameters for the points generations are the same as in Fig. A.1.
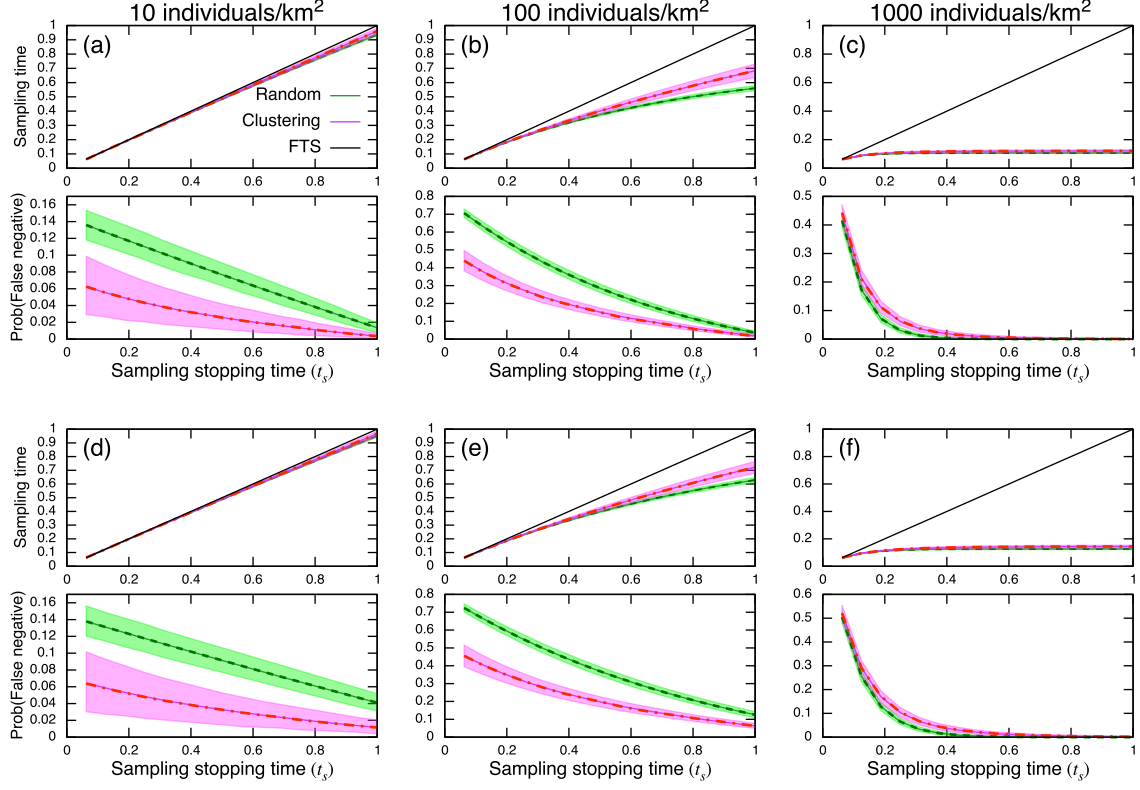
Figure S.3: Normalized sampling time and false negative given normalized stopping time t_s with the probability of sampling error $\epsilon = 0.1$ (a-c), and $\epsilon = 0.3$ (d-f). For each individual distribution scenarios, the numerical average (lines) and its theoretical value (dashed) are provided. The shaded area are between 5 and 95 percentiles of a 10^5-time numerical simulation. FTS represents the fixed-time survey. The scales of surveys are are $\nu(W) = 4\text{km} \times 4\text{km}$, $\nu(M) = 2^{-3}\text{km} \times 2^{-3}\text{km}$, $\nu(S) = 2^{-5}\text{km} \times 2^{-5}\text{km}$. The parameters for the points generations are the same as in Fig. A.1.
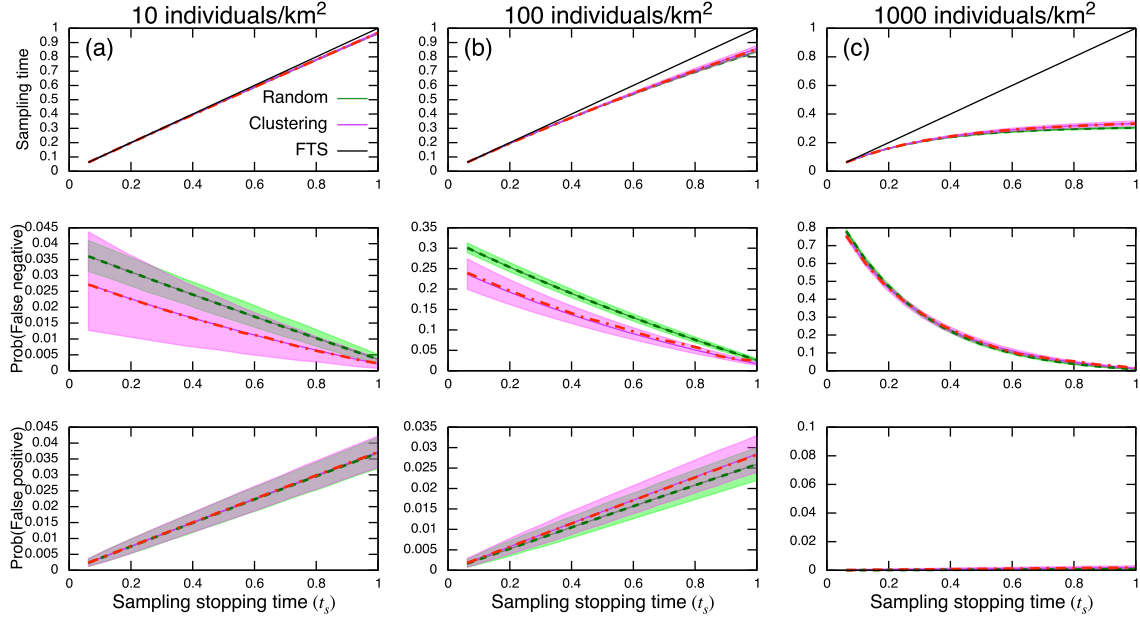
Figure S.4: Normalized sampling time, false negative, and false positive given normalized stopping time t_s with the probability of sampling error $\epsilon = 0.1$. For each individual distribution scenarios, the numerical average (lines) and its theoretical value (dashed) are provided. The shaded area are between 5 and 95 percentiles of a 10^5-time numerical simulation. FTS represents the fixed-time survey. The scales of surveys are are $\nu(W) = 4\text{km} \times 4\text{km}$, $\nu(M) = 2^{-4}\text{km} \times 2^{-4}\text{km}$, $\nu(S) = 2^{-6}\text{km} \times 2^{-6}\text{km}$. The intensity of false positive is $\lambda_{fp} = 10$. The parameters for the points generations are the same as in Fig. A.1.
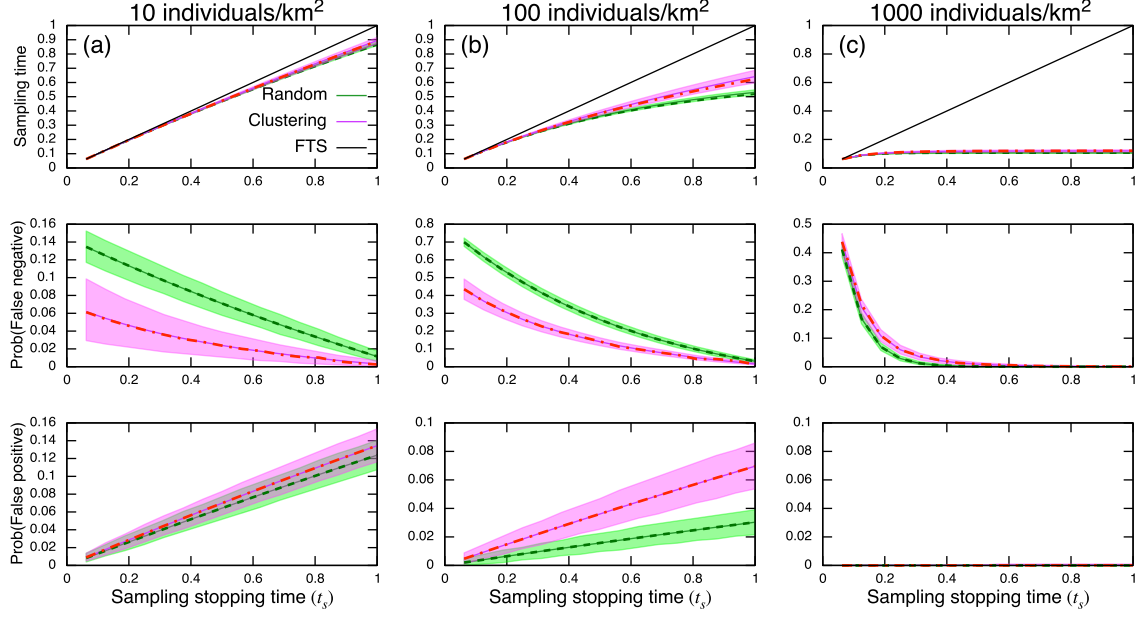
Figure S.5: Normalized sampling time, false negative, and false positive given normalized stopping time t_s with the probability of sampling error $\epsilon = 0.1$. For each individual distribution scenarios, the numerical average (lines) and its theoretical value (dashed) are provided. The shaded area are between 5 and 95 percentiles of a 10^5-time numerical simulation. FTS represents the fixed-time survey. The scales of surveys are are $\nu(W) = 4\text{km} \times 4\text{km}$, $\nu(M) = 2^{-3}\text{km} \times 2^{-3}\text{km}$, $\nu(S) = 2^{-5}\text{km} \times 2^{-5}\text{km}$. The intensity of false positive is $\lambda_{fp} = 10$. The parameters for the points generations are the same as in Fig. A.1.

A Point generations

Individual distributions

The random and clustering individual distributions are generated by applying the theory of spatial point processes that accommodate stochasticity in individual distributions in region W. Specifically, random individual distributions are generated by a homogeneous Poisson process, and a Thomas process, widely applied to characterize intraspecific aggregation pat-

terns (Fig. A.1).

In the homogeneous Poisson process, provided the intensity λ_{po}, the number of individuals X in a given region R with area $\nu(R)$ follows the Poisson distribution with intensity $\lambda_{po}\nu(R)$

$$P(X = k) = \frac{(\lambda_{po}\nu(R))^k}{k!}e^{-\lambda_{po}\nu(R)}. \tag{S.1}$$

The Thomas process is an extension of the homogeneous Poisson process and it is generated by the following three steps:

1. Parents locations are determined according to the homogeneous Poisson process with a parent intensity λ_p.

2. Each parent produces a random number of daughters with an average c that follows the Poisson distribution.

3. The generated daughters are placed around their parents independently with an isotropic bivariate Gaussian distribution with the variance σ_{th}^2, and the parents are removed.

The intensity of the Thomas process is defined [1]

$$\lambda_{th} = \bar{c}\lambda_p, \tag{S.2}$$

In the analysis, we set $\lambda_{th} = \lambda_{po}$ to satisfy the average numbers in a given area are the same under both individual distributions.

The zero probability of the Thomas process in the mapping unit M is [36]

$$P(Y = 0) = \exp\left(-\lambda_p \int_{\mathbf{R}^2} \left(1 - \exp\left(-\bar{c}\int_M \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2}\right)d\mathbf{x}\right)\right)d\mathbf{y}\right). \tag{S.3}$$

For the zero probability $P(X(t_s) = 0)$ where the subregion $S_{t_s} \subset M$ is sampled until time t_s with sampling error ϵ, the power of second exponential of Eq. (S.3) is described as
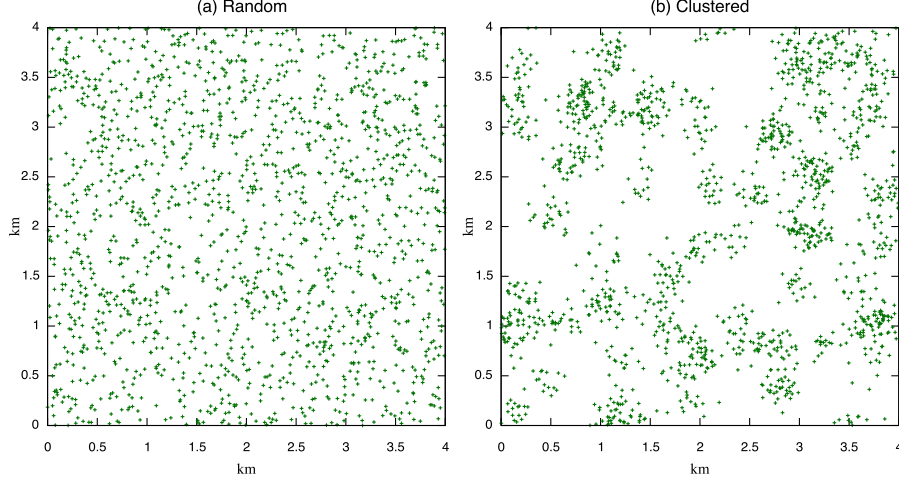
46

Figure A.1: Examples of point patterns based on the (a) homogeneous Poisson process (random); and (b) Thomas (clustered) process. Parameter values are $\lambda_{po} = 100$ for random process and $\lambda_{th} = 5$, $c = 20$, $\sigma_{th} = 0.1$ for clustered process.

$$-\bar{c}(1 - \epsilon) \int_{S_{t_s}} \frac{1}{2\pi\sigma^2} \exp(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2})d\mathbf{x}.$$

References

[1] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, Stochastic Geometry and Its Applications. New York: John Wiley & Sons, 2013.

B   Derivations of sampling probabilities

Here we derive each probability required for Eqs. (1a), (1b). Later, we will show the same manner will immediately follow in the case of clustering individual distributions.

When individuals are distributed randomly (i.e., via homogeneous Poisson distribution), the probability of existing 0 individual in a mapping unit with area $\nu(M)$ is $P(Y = 0) = e^{-\lambda_{po}\nu(M)}$ by Eq. (S.1). The probability of miss-detection conditioned on the existence of individual(s) in this case is directly obtained by calculating (i) the probability of mis-

47

719 detection given individual encounter (*hitting*), and (ii) encounter no species (*non-hitting*)

720 until time t_s given individual existing in a mapping unit as follows

$$
\begin{aligned}
P_{po}(X(t_s) = 0 \mid Y > 0) &= \underbrace{\frac{e^{-\lambda_{po}\nu(S)t_s}}{1 - e^{-\lambda_{po}\nu(M)}}}_{P(hitting)} \underbrace{\left(\epsilon\lambda_{po}\nu(S)t_s + \frac{(\epsilon\lambda_{po}\nu(S)t_s)^2}{2!} + \frac{(\epsilon\lambda_{po}\nu(S)t_s)^3}{3!} + \cdots\right)}_{P(non-detection)} \\
&\quad + \underbrace{1 - \frac{1 - e^{-\lambda_{po}\nu(S)t_s}}{1 - e^{-\lambda_{po}\nu(M)}}}_{P(non-hitting)}, \\
&= \frac{e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)} - e^{-\lambda_{po}\nu(M)}}{1 - e^{-\lambda_{po}\nu(M)}}.
\end{aligned}
\tag{S.4}
$$

Therefore, the probability of false negative can be reduced more efficiently by increasing the sampling stopping time t_s when the intensity λ_{po} and sampling unit $\nu(S)$ are larger and detection error ϵ is smaller. Using Eq. (S.4), we have the followings:

$$
P(X(t_s) > 0 \mid Y > 0) = \frac{1 - e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)}}{1 - e^{-\lambda_{po}\nu(M)}},
\tag{S.5a}
$$

$$
P(X(t_s) = 0, Y > 0) = e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)} - e^{-\lambda_{po}\nu(M)},
\tag{S.5b}
$$

$$
P(X(t_s) > 0, Y > 0) = 1 - e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)},
\tag{S.5c}
$$

721 where Eq. (S.5b) is the false negative probability and Eq. (S.5c) gives the occupancy

722 probability of the presence-absence map. Let $P(t)$ be the probability that a state $X(t) = 0$

723 switches to $X(t) > 0$ at time t, and it is via Eq. (S.5c),

$$
\begin{aligned}
P(t) &= P(X(t) > 0, Y > 0) - P(X(t-1) > 0, Y > 0), \\
&= e^{-\lambda_{po}\nu(S)(t-1)(1-\epsilon)} - e^{-\lambda_{po}\nu(S)t(1-\epsilon)}.
\end{aligned}
\tag{S.6}
$$

724 Then, the average time of this event has the form $\mathrm{E}[t] = \sum_{t=1}^{t_s} tP'(t)$ where, $P'(t)$ is the

725 normalized probability of $P(t)$ with any stopping time t_s obtained by dividing by $\sum_1^{t_s} P(t) =$

$P(X(t_s) > 0, Y > 0)$ so as to satisfy $\sum_0^{t_s} P'(t) = 1$. Thus this is described

$$
\begin{aligned}
\mathrm{E}[t] &= \frac{1}{P(X(t_s) > 0, Y > 0)}\left(\sum_{t=0}^{t_s-1} e^{-\lambda_{po}\nu(S)t(1-\epsilon)} - t_s e^{-\lambda_{po}\nu(S)t_{stop}(1-\epsilon)}\right), \\
&= \frac{1}{1 - e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)}}\left(\frac{1 - e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)}}{1 - e^{-\lambda_{po}\nu(s)(1-\epsilon)}} - t_s e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)}\right).
\end{aligned}
\tag{S.7}
$$

Substituting $P(Y = 0)$, Eqs. (S.5a), (S.5b), and (S.7) into Eq. (16), we have the simpler form:

$$
\begin{aligned}
t^{po}_{samp} &= N_M \sum_{t=0}^{t_s-1} P_{po}(X(t) = 0), \\
&= N_M \frac{1 - e^{-\lambda_{po}\nu(S)t_s(1-\epsilon)}}{1 - e^{-\lambda_{po}\nu(S)(1-\epsilon)}}.
\end{aligned}
\tag{S.8}
$$

This suggests intuitive characteristics of ecological survey: the total sampling time is proportional to the number of mapping unit N_M, the effect of a mapping resolution M. As in the case of Eq. (S.4), the sampling time is reduced more efficiently by increment the sampling stopping time t_s when the factors λ_{po}, $\nu(S)$, and $1 - \epsilon$ are larger.

C Small limit of mapping units

When the mapping unit becomes very small ($\nu(M) \ll 1$; hence the sampling unit S does too) we can discuss the asymptotic behaviors. This is possible when the sampling devices offer highly resolved spatial images. Provided these conditions, it is straightforward, by expanding the exponential terms of corresponding equations above, to show that the sampling time and

probability of false negative asymptotically converge to the same values. These are

$$\lim_{M \to 0} t_{samp}^{dist} = N_M t_s, \tag{S.9a}$$

$$\lim_{M \to 0} P_{dist}(X(t) = 0, Y > 0) = 0. \tag{S.9b}$$

Note Eqs. (S.9) also hold for the limit of individual intensities $\lambda_{po} \to 0$ or $\lambda_{th} \to 0$ (i.e., sparse populations). Intuitively speaking, this limit emerges when the mapping unit becomes sufficiently small that each mapping unit holds at most one individual. In this limit, spatial structure does not matter at the scale of a sampling resolution.

D   Sampling under the possibility of false positive detection

When there is possibility for false positive detection, it is still possible to discuss the sampling performance under our framework. However, the cause of false positive (e.g., random noise, miss classification, etc.) may be much diverse than the false positive detection where its definition is straightforward (i.e., miss-detection). Here we demonstrate the extension and provide some theoretical and numerical results.

D.1   Extended model

To consider a possibility of false positive detection, we introduce the new probability variables X_1 and X_2, where X_1 corresponds to X in the above discussion and X_2 is the indicator of false positive. These are mutually exclusive; hence $X_1 > 0$ and $X_2 > 0$ do not occur

50

749 simultaneously. Then, Eq. (1b) in the main text becomes

$$
\begin{aligned}
T_{samp} \;=\; N_M \Bigg[& t_s \underbrace{\{ P(\mathbf{X}(t_s) = 0, Y = 0) + P(\mathbf{X}(t_s) = 0, Y > 0) \}}_{\text{no detection}} \\
& + \mathrm{E}_1[t] \underbrace{P(X_1(t_s) > 0, X_2(t_s) = 0, Y > 0)}_{\text{detection}} \\
& + \mathrm{E}_2[t] \underbrace{P(X_1(t_s) = 0, X_2(t_s) > 0, Y = 0)}_{\text{false detection within an empty patch}}, \\
& + \mathrm{E}_3[t] \underbrace{P(X_1(t_s) = 0, X_2(t_s) > 0, Y > 0)}_{\text{false detection within a patch with individual(s)}} \Bigg],
\end{aligned}
\tag{S.10}
$$

750 where, $\mathbf{X}(t_s) = 0$ is a concise description of $X_1(t_s) = 0$ and $X_2(t_s) = 0$, $\mathrm{E}_1[t]$, $\mathrm{E}_2[t]$, and
751 $\mathrm{E}_3[t]$ are the average times for a detection, to cause a false positive within an empty patch,
752 to cause a false positive within a patch with individual(s). Note false positive detection is
753 either in an empty patch or a patch with individuals. These influence differently on the
754 total sampling time. Also, our theory discussed in the main text is immediately recovered
755 by turning off the possibility of false positive detection and set (i.e., $P(X_2(t) = 0) = 1$) and
756 replacing the notation X_1 with X.

757 With the new probability variable, Eq. (S.6) is described as:

$$
P_1(t) = P(X_1(t) > 0, X_2(t) = 0, Y > 0) - P(X_1(t-1) > 0, X_2(t) = 0, Y > 0). \tag{S.11}
$$

Similarly, the probability of switching from no false positive to false positive in an empty patch $P_2(t)$ and a patch with individual(s) $P_3(t)$ at time t are respectively described as follows:

$$
P_2(t) = P(X_1(t) = 0, X_2(t) > 0, Y = 0) - P(X_1(t) = 0, X_2(t-1) > 0, Y = 0), \tag{S.12a}
$$

$$
P_3(t) = P(X_1(t) = 0, X_2(t) > 0, Y > 0) - P(X_1(t) = 0, X_2(t-1) > 0, Y > 0). \tag{S.12b}
$$

51

The average time $E_i[t]$ is calculated using these probabilities as:

$$E_i[t] \;=\; \left(\sum_{t=1}^{t_s} P_i(t)\right)^{-1} \sum_{t=1}^{t_s} P_i(t)t, \tag{S.13}$$

where the first summation is the normalization factor.

D.2   Examples

Here, we perform some numerical and theoretical calculations of the extended model. Due to its highly complex nature, we phenomenologically model that probability of false positive via an ordinary assumption in e.g., queueing theory and birth-death process [1]. That is, we assume that the number of false positive detection after sampling a certain region $\nu(S)t$ at time t follows a Poisson distribution with an intensity λ_{fp}, $\text{Po}(\lambda_{fp}\nu(S)t)$. This indicates false positive detection occur independently of individual distributions, and yields a great simplification as we can describe e.g., the probability of detection

$$P(X_1(t) > 0, X_2(t) = 0, Y > 0) = P(X_1(t) > 0, Y > 0)P(X_2(t) = 0). \tag{S.14}$$

As X_1 corresponds to X in the main text, we can use the same probabilities aside from probabilities of X_2. As the number of false positive detection follows the Poisson distribution, we have the followings

$$P(X_2(t) = 0) = \exp(-\lambda_{fp}\nu(S)t), \tag{S.15a}$$

$$P(X_2(t) > 0) = 1 - \exp(-\lambda_{fp}\nu(S)t). \tag{S.15b}$$

Using these probabilities, we calculate theoretically and numerically in three scenarios of mapping and sampling units as in the main text (Figs. ??-S.5). These results demonstrate

52

770 that the theory developed in the main text is easy to extend, and assessment of the false

771 positive probability $P(X_2(t))$ will further improve quality of the sampled data.

References

773 [1] Asmussen, S, Applied probability and queues, Second Edition. New York: Springer-

774 Verlag, 2003