



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/127443/>

Version: Accepted Version

Article:

Takashina, N, Beger, M, Kusumoto, B et al. (2018) A theory for ecological survey methods to map individual distributions. *Theoretical Ecology*, 11 (2). pp. 213-223. ISSN: 1874-1738

<https://doi.org/10.1007/s12080-017-0359-7>

© Springer Science+Business Media B.V., part of Springer Nature 2017. This is an author produced version of a paper published in *Theoretical Ecology*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

A theory for ecological survey methods to map individual distributions

Nao Takashina^{1*}, Maria Beger^{2,3}, Buntarou Kusumoto⁴, Suren Rathnayake⁵, Hugh P.

Possingham^{2,6}

¹*Tropical Biosphere Research Center, University of the Ryukyus,*

3422 Sesoko Motobu, Okinawa 905-0227, Japan

²*ARC Centre of Excellence for Environmental Decisions, School of Biological Sciences, The*

University of Queensland, St Lucia, QLD 4072, Australia

³*School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK*

⁴*Center for Strategic Research Project, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa*

903-0213, Japan

⁵*School of Biological Sciences, The University of Queensland, St Lucia, QLD 4072, Australia*

⁶*The Nature Conservancy, 4245 North Fairfax Drive Suite 100 Arlington, VA 22203-1606, USA*

Keywords: Ecological survey, presence/absence map, spatial distribution, spatial point processes

*Corresponding author

Current address: Biodiversity and Biocomplexity Unit, Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa 904-0495, Japan

Email: nao.takashina@gmail.com

1 **Abstract**

2 Spatially-explicit approaches have been widely recommended for various applications
3 of ecosystem management. In practice, the quality of the data involved in the manage-
4 ment decision-making, such as presence/absence or habitat maps, affects the manage-
5 ment actions recommended, and therefore it is a key to management success. However,
6 available data is often biased and incomplete. Although previous studies have advanced
7 ways to effectively resolve data bias and missing data, there still remains a question
8 about how we design the entire ecological survey to develop a dataset through field sur-
9 veys. Ecological survey may inherently have multiple spatial scales to be determined
10 beforehand, such as the spatial extent of the ecosystem under concern (observation
11 window), the resolution to map the individual distributions (mapping unit), and the
12 area of survey within each mapping units (sampling unit). In this paper, we develop a
13 theory to understand ecological survey for mapping individual distributions applying
14 spatially-explicit stochastic models. Firstly, we use spatial point processes to describe
15 individual spatial placements drawn using either random or clustering processes. An
16 ecological survey is then introduced with a set of spatial scales and individual de-
17 tectability. Regardless of the spatial pattern assumed, the choice of mapping unit
18 largely affects presence mapped fraction, and the fraction of the total individuals cov-
19 ered by the presence mapped patches. Tradeoffs between these quantities and the
20 resolution of the map are found, associated with an equivalent asymptotic behaviors
21 for both metrics at sufficiently small and large mapping unit scales. Our approach
22 enables us to directly discuss the effect of multiple spatial scales in the survey, and
23 estimating the survey outcome such as the presence mapped fraction and the number
24 of individuals situated in the presence detected units. The developed theory may sig-
25 nificantly facilitate management decision-making and inform the design of monitoring
26 and data gathering.

1 Introduction

Understanding the spatial characteristics of ecosystems is one of the central challenges in ecology [1]. Such knowledge forms a prerequisite for effective ecosystem management due to an increasing need for spatially explicit approaches in fisheries and wildlife management [2–4] and for the establishment of terrestrial and marine protected areas [5–7].

In ecosystem management, the quality of the data involved in the management decision-making, such as presence/absence or habitat maps, affect the management actions recommended [8–10]. Therefore, creating an ecologically and statistically adequate dataset is key to management success. However, available data is often biased and incomplete [8, 9], due to, for example, different accessibility to sites [8], existence of the favored study sites [8], and imperfect detectability of individuals [11, 12]. These biases hinder the effective implementation of management actions, and may lead to perverse outcomes or wasted management resources. Hence it is important to discuss and benchmark the quality of the spatially explicit data that underlies management decisions.

There is a body of literatures to tackle the challenges of data gathering, including sampling designs for effectively allocating the survey effort under the time and budgetary constraints [13–15], methods for reducing the bias of occurrence data by estimating the detectability of species [12, 16–18], and mathematical theory for ecological sampling [19, 20]. Although these researches have significantly advanced our insight into ecosystem monitoring and ecological survey, there still remains a question about how we actually design the entire ecological survey to systematically develop dataset through a field survey, as the spatial scale issue, such as how to chose the resolution of a map, is often omitted. This is perhaps because many existing studies consider the space to be sampled implicitly. Presence/absence or habitat map is widely used in ecosystem management [16], where at least three different spatial scales may exist; the spatial extension of the ecosystem under concern, resolution to map

52 the individual distributions, and minimum size of survey units. To systematically gather
53 the spatial data, manager should explicitly take into consideration these three spatial scales,
54 because the manner of the sampling and management outcomes depend on the resolution of
55 a map. For example, in fisheries management, finely implemented fishing quota allocations
56 may result in better management outcomes [7, 21], and this can be done with the distri-
57 bution map with a high degree of resolution. However, surveying an area at a fine spatial
58 scale is often impractically expensive in the large scale survey, and the choice of resolution
59 itself faces a budgetary constraint. Hence, quantitatively estimating the performance of a
60 sampling method in advance facilitates survey decision making.

61 In this paper, we develop a theory of ecological survey method for systematically mapping
62 individual distributions by making use of the spatial point processes (SPPs), a spatially
63 explicit stochastic model. The SPPs is widely applied to the study of plant community
64 [22–25], coral community [26], and avian habitat selection [27]. Therefore, they are potential
65 target species of the developed theory. However, the method developed here may be suitable
66 for any organism or the location used by an organism (e.g., nesting site) that is relatively
67 sedentary on a time and spatial scale of the field survey where its spatial distributions can be
68 described by SPPs [28]. In this study, the SPPs describes individual spatial locations by two
69 different processes accounting random or clustering patterns. An ecological survey is then
70 introduced with a set of spatial scales and detectability of individuals. Our spatially-explicit
71 approach is capable of revealing a series of questions important for ecological survey, such
72 as effect of the choice of the spatial scales and spatial distribution patterns of individuals on
73 accuracy of the distribution map. This knowledge enables one to determine the design of an
74 ecological survey beforehand given accuracy of a map required. The developed theory may
75 significantly facilitate management decision making and give solid bases of data gathering.

76 2 Methods

77 2.1 Models of spatial distribution of individuals

78 To develop a theory of ecological survey to map individual distributions, we explicitly model
79 the spatial distribution patterns of individuals. Spatial point processes (SPPs) [22, 25] pro-
80 vide models to describe such patterns with high flexibility and analytical tractability [24].

81 Here, we apply the homogeneous Poisson process and the Thomas process, a family of
82 the Neyman-Scott process (Fig. 1). One of the simplest SPPs is the homogeneous Poisson
83 process where the points (i.e. individuals) are randomly distributed and the number of points
84 of a given region A , $N(A)$, is according to the Poisson distribution with an average μ_A :

$$\text{Prob}(N(A) = k) = \frac{\mu_A^k}{k!} e^{-\mu_A}, \quad (k = 0, 1, \dots) \quad (1)$$

85 where, μ_A is also regarded as the intensity measure [22, 25] described as

$$\mu_A = \lambda \nu(A), \quad (2)$$

86 where, $\lambda = (\text{total points})/(\text{area of concerned region } A)$ is the intensity in the given region,
87 and $\nu(A)$ is the area of A .

88 The Neyman-Scott process [22, 25] provides us more general framework to analyze spa-
89 tial ecological data and characterize the clustering pattern of individuals [22–25]. By the
90 following three steps, the Neyman-Scott process is obtained:

- 91 • Parents are randomly placed according to the homogeneous Poisson process with an
92 intensity λ_p .
- 93 • Each parent produces a random discrete number c of daughters, realized independently
94 and identically for each parent.

95 • Daughters are scattered around their parents independently with an identical spatial
 96 probability density, $f(\mathbf{y})$, and all the parents are removed in the realized point pattern.

97 The intensity of the Neyman-Scott process is [25]

$$\lambda = \bar{c}\lambda_p, \quad (3)$$

98 where, \bar{c} is the average number of daughters per parent. The probability generating functional
 99 (pgfl) of the number of daughters within a given region of the Neyman-Scott process is [22,25]

$$G(v) = \exp \left(-\lambda_p \int_{\mathbf{R}^d} \left[1 - G_n \left(\int_{\mathbf{R}^d} v(\mathbf{x} + \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \right) \right] d\mathbf{x} \right), \quad (4)$$

100 where, $G_n \left(\int_{\mathbf{R}^d} v(\mathbf{x} + \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \right)$ is the probability generating function (pgf) of the random
 101 number c , the number of daughters per parent.

102 The Thomas process is a special case of the Neyman-Scott process, where $f(\mathbf{y})$ is an
 103 isotropic bivariate Gaussian distribution with the variance σ^2 [25]. We also assume that the
 104 number of daughters per parent follows the Poisson distribution with the average number, \bar{c} .
 105 The pgfl of the Thomas process, Eq. (4), within a given region A is obtained by substituting
 106 the pgf of the number of daughters per parent G_n in Eq. (4). It is obtained, by the given
 107 assumptions, as

$$\begin{aligned} G_n \left(\int_{\mathbf{R}^d} v(\mathbf{x} + \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \right) &= \sum_{k=0}^{\infty} \left(\int_{\mathbf{R}^d} v(\mathbf{y}) f(\mathbf{y} - \mathbf{x}) d\mathbf{y} \right)^k \frac{\bar{c}^k}{k!} e^{-\bar{c}}, \\ &= \exp \left[-\bar{c} \left(1 - \int_{\mathbf{R}^d} v(\mathbf{y}) f(\mathbf{y} - \mathbf{x}) d\mathbf{y} \right) \right], \\ &= \exp \left[-\bar{c}(1-t) \left(\int_A f(\mathbf{y} - \mathbf{x}) d\mathbf{y} \right) \right], \end{aligned} \quad (5)$$

108 where, to obtain the last line, $v(\mathbf{y}) = 1 - (1-t)\mathbf{1}_A(\mathbf{y})$ is used, and here $\mathbf{1}_A(\mathbf{y})$ is the indicator
 109 function. Therefore, the pgfl of the number of daughters within the region A of the Thomas

110 process is

$$G(t) = \exp \left(-\lambda_p \int_{\mathbf{R}^2} \left[1 - \exp \left\{ -\bar{c}(1-t) \left(\int_A k(\|\mathbf{x} - \mathbf{y}\|) d\mathbf{y} \right) \right\} \right] d\mathbf{x} \right), \quad (6)$$

111 where, $k(\|\mathbf{x} - \mathbf{y}\|)$ is an isotropic bivariate Gaussian distribution with variance σ^2 ,

$$k(\|\mathbf{x} - \mathbf{y}\|) = \frac{1}{2\pi\sigma^2} \exp \left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2} \right). \quad (7)$$

112 In order to reasonably compare the results of Thomas process with those of the homo-
113 geneous Poisson process, we chose the intensity of the Thomas process so as to have, on
114 average, the same number of points within the concerned region. Namely, the parameters
115 λ_p and \bar{c} satisfy

$$\bar{c}\lambda_p = \lambda, \quad (8)$$

116 where, the left hand side (lhs) is the intensity of the Thomas process and the right hand side
117 (rhs) is the intensity of the homogeneous Poisson.

118 **2.2 Design of ecological survey**

119 **2.2.1 Survey rules and basic properties**

120 Let us consider the situation where an ecological survey takes place for the purpose of
121 creating a presence/absence map of a given region. A presence/absence map is characterized
122 by the three spatial scales: the *observation window* (W), the spatial scale of ecological survey
123 conducted, the spatial scale of the *mapping unit* (M) defining the resolution of the map, and
124 the spatial scale of the *sampling unit* (S) determining the sampling density within each
125 mapping unit (Fig. 2). We assume the following three key sampling rules.

- 126 • The observation window, resolution of map (i.e., scale of the mapping unit), and sam-

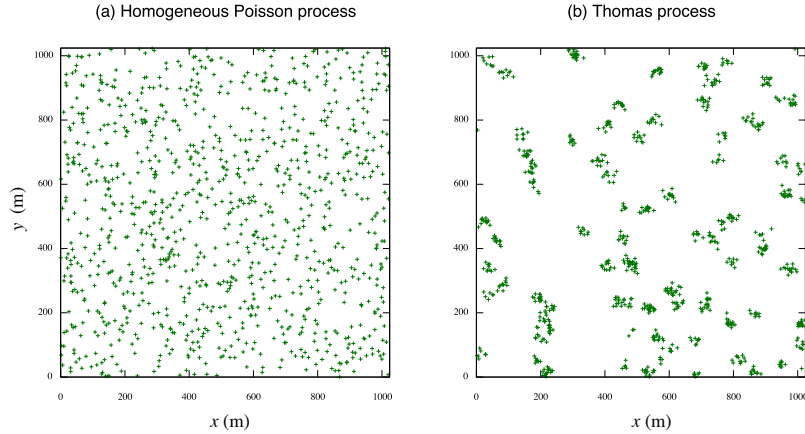


Figure 1: Example of point patterns within a observation window $1024\text{m} \times 1024\text{m}$. (a) Homogeneous point process with the intensity $\lambda = 10^{-3}$; (b) Thomas process with the same intensity value as the homogeneous Poisson process $\lambda_p \bar{c} = \lambda$, where $\lambda_p = 10^{-4}$ and $\bar{c} = 10$. The variance of the bivariate normal distribution $\sigma^2 = 100$. See the text for the interpretations of the parameters.

127 pling unit, are arbitrary determined, but single resolutions are allowed for each of the
 128 spatial scales.

129 • Every mapping unit is assessed by sampling unit, and sampling location is determined
 130 randomly within the mapping unit.

131 • A mapping unit is recorded as presence if at least one individual is detected regardless
 132 of the number of miss detections. If there is no individual or all individuals are miss
 133 detected within the mapping unit, the mapping unit is recorded as absence.

134 Through the second and third assumptions, changing the scale of the mapping unit affects
 135 the obtained presence/absence map (Fig. 3).

136 2.2.2 Modeling the ecological survey

137 Here, we model the ecological survey with the three main assumptions listed above. Let, on
 138 average, N individuals of a species be distributed over a given window W , which is the region

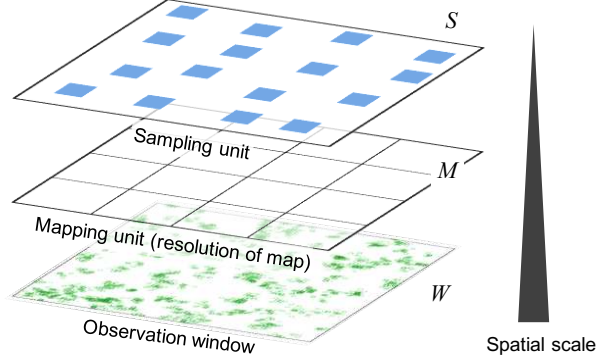


Figure 2: Multiple spatial scales in ecological survey. Each scale is arbitrary determined by managers.

139 under concern (i.e., $N = N(W)$). The manner of individual distribution follows either the
 140 homogeneous Poisson process or the Thomas process. The resolution of the presence/absence
 141 map is defined by the scale of mapping unit M , and every mapping unit is sampled with the
 142 sampling unit S (Fig. 3). The survey is associated by the sampling error for each individual
 143 at a probability $\gamma := 1 - \beta$, which is the probability at which individuals are not detected
 144 despite being present, and where, β is the detectability of an individual. For simplicity, we
 145 assume that the areas of each mapping unit is 1, 2, 4, \dots , or 2^n times smaller than the area
 146 of a given window W . Let $\nu(X)$ be the area of a region X . With the definitions detailed
 147 above, we obtain

$$\nu(M) = \nu(W)/2^m, \quad (m = 0, 1, \dots, n) \quad (9)$$

148 where, the superscript m represents the number of subdivisions of the window W . From Eq.
 149 (9), the number of mapping units within a given window W , is

$$\text{Number of mapping units} = \nu(W)/\nu(M) = 2^m. \quad (10)$$

150 As the record for each mapping unit is based on an survey within the mapping unit, we
 151 obtain

$$\nu(S) = \alpha\nu(M) \leq \nu(M), \quad (0 < \alpha \leq 1) \quad (11)$$

152 where, α is the sampling density within a mapping unit. Combining Eq. (9) and Eq. (11),
 153 we obtain

$$\nu(S) \leq \nu(M) \leq \nu(W). \quad (12)$$

154 Let the intensity of the points within a given window W be [22, 25]

$$\lambda = \frac{N(W)}{\nu(W)}. \quad (13)$$

155 As we noted above, the parameters for the Thomas process are chosen so as to satisfy Eq.
 156 (8).

157 **2.3 Assessing accuracy of presence/absence map**

158 Given the spatial point pattern, sampling density, α , detectability of an individual, β , and
 159 scale of mapping unit, M , we calculate two main quantities of the ecological survey. That
 160 is, the presence mapped fraction (PM fraction), and, the fraction individuals covered by
 161 presence mapped patches (IC fraction):

$$\text{PM fraction} = \frac{\# \text{ presence units mapped}}{\# \text{ presence units exists}}, \quad (14)$$

$$\text{IC fraction} = \frac{\# \text{ individuals in mapped units}}{\# \text{ total individuals}}. \quad (15)$$

162 The presence mapped fraction is the fraction to map presence units correctly and it is
 163 connected with type II error (i.e., $1 - (\text{PM fraction})$ is the probability of type II error in the
 164 obtained map). Instead, the IC fraction, although this measure has not been investigated

165 to our knowledge, connects the PM fraction with population abundance in the observation
 166 window W . For example, let us assume that we find the values of PM and IC fraction are 0.8
 167 and 0.95 respectively given a survey scenario. In that situation, we would expect that 95%
 168 of the total individuals in the observation window are situated within the presence mapped
 169 units. Therefore, the IC fraction also provides useful information for conservation. Examples
 170 of PM and IC fraction values are shown in Fig. 3. It is also expected that the difference
 171 between the average PM fraction and IC fraction increases as the degree of clustering in the
 172 distribution patterns increases, since the individual number is biased to certain (moderate-
 173 sized) mapping units and such sites are more likely to be mapped as presence.

174 The type II error is often a concern in ecological monitoring to estimate how the monitor-
 175 ing is accurate (e.g., [13]). Nevertheless, we apply the PM fraction in the following analysis
 176 to facilitate a comparison of the two measures, since PM and IC fractions have similar curves
 177 as we will see below. As noted above, however, the type II error is easily obtained from the
 178 PM fraction.

179 The presence mapped fraction is obtained by

$$E_{\Lambda,(\beta,S,M)}[\text{PM}] = \frac{1 - p(\text{find 0 individual in } S \mid \beta)}{1 - p(N(M) = 0)}, \quad (16)$$

180 where, Λ indicates the underlying point pattern. On the other hand, the form of the fraction
 181 of total individual situated within presence-mapped units is described as

$$E_{\Lambda,(\beta,S,M)}[\text{IC}] = \frac{2^m}{\mu_W} \sum_{k=1}^{\infty} p(\text{find at least 1 individual in } S \mid N(M) = k, \beta) \quad (17) \\ \times kp(N(M) = k),$$

182 where, 2^m is the number of mapping units as Eq. (10). Since the IC fraction is rather
 183 cumbersome to derive analytically for the Thomas process, we only provide an analytical

184 expression of the IC fraction for the homogeneous Poisson process, and give numerical results
 185 for the Thomas process.

186 2.4 Numerical settings

187 In addition to the IC fraction of Thomas process, we conduct numerical simulations to check
 188 our analytical results by our own C code (available on request). Implementing numerical
 189 simulations is straightforward by taking first two steps (a) and (b) as shown in Fig. 3, and cal-
 190 culate PM and IC fraction values by counting the detected habitats and individuals therein.
 191 We repeat 1000 times this simulation to obtain the 5, 25, 50, 75, and 95 percentile values.
 192 We set a observation window to $1024\text{m} \times 1024\text{m}$, and mapping unit is $2^1, 2^2, \dots, 2^{17}$ times
 193 smaller than the observation window. We also set the sampling density and detectability to
 194 0.5 and 0.9, respectively. The other parameter values are the same as in Fig. 1.

195 3 Results

196 3.1 Ecological survey with individual distributions based on the 197 homogeneous Poisson process

198 Where individuals are distributed in space based on the homogeneous Poisson process, pres-
 199 ence mapped fraction from Eq. (16) is

$$E_{po,(\beta,S,M)}[\text{PM}] = \frac{1 - e^{-\beta\lambda\nu(S)}}{1 - e^{-\lambda\nu(M)}} = \frac{1 - e^{-\alpha\beta\lambda\nu(M)}}{1 - e^{-\lambda\nu(M)}}, \quad (18)$$

200 where, the equality $\nu(S) = \alpha\nu(M)$ is used. Eq. (18) has rather simple form and, thus, we
 201 can easily see the parameter dependence. The intensity of the points λ (Eq. 13) defines the
 202 average number of individuals existing within a given the observation window, W , and since

203 $dE_{po}[\text{PM}]/d\lambda \geq 0$, $E_{po}[\text{PM}]$ increases as the average number of individuals increase, and vice
 204 versa. Especially, when the intensity becomes $\lambda \rightarrow \infty$, $E_{po}[\text{PM}]$ becomes 1 regardless of the
 205 scale of mapping units. Intuitively, as the sampling density α and detectability β increase,
 206 $E_{po}[\text{PM}]$ increases, and vice versa. The asymptotic behavior $M \rightarrow 0$ of Eq. (18) is obtained
 207 by expanding about $\nu(M)$

$$\lim_{M \rightarrow 0} E_{po,(\beta,S,M)}[\text{PM}] \simeq \alpha\beta. \quad (19)$$

208 Since the zero probabilities $p(N(S) = 0 | \beta)$ and $p(N(M) = 0)$ approach to 0 as $M \rightarrow W$
 209 given the observation window, W , is sufficiently large, we obtain

$$\lim_{M \rightarrow W} E_{po,(\beta,S,M)}[\text{PM}] \simeq 1. \quad (20)$$

210 These results show good agreement with the numerical results (Fig. 4a).

211 For the homogeneous Poisson process, we can derive an analytical form of the average
 212 fraction of individuals covered within presence mapped patches (IC) as follows:

$$\begin{aligned} E_{po,(\beta,S,M)}[\text{IC}] &= \frac{2^m}{\mu_W} \sum_{k=1}^{\infty} \left\{ 1 - \left(1 - \beta \frac{\nu(S)}{\nu(M)} \right)^k \right\} k \frac{(\lambda\nu(M))^k}{k!} e^{-\lambda\nu(M)}, \\ &= \frac{2^m}{\mu_W} \sum_{k=1}^{\infty} \{ 1 - (1 - \alpha\beta)^k \} k \frac{(\lambda\nu(M))^k}{k!} e^{-\lambda\nu(M)}, \\ &= \frac{2^m \lambda\nu(M)}{\mu_W} \sum_{k=1}^{\infty} \{ 1 - (1 - \alpha\beta)^k \} \frac{(\lambda\nu(M))^{k-1}}{(k-1)!} e^{-\lambda\nu(M)}, \\ &= 1 - (1 - \alpha\beta) e^{-\lambda\nu(M)} \sum_{k=0}^{\infty} \frac{((1 - \alpha\beta)\lambda\nu(M))^k}{k!}, \\ &= 1 - (1 - \alpha\beta) e^{-\alpha\beta\lambda\nu(M)}, \end{aligned} \quad (21)$$

213 where, on the first line of rhs, 2^m is the number of mapping units within the given window
 214 W , inside of the curly brackets is the probability that none of k points are detected by a
 215 survey given a mapping unit M , and the remaining term is the expected number of points

216 within the mapping unit. The second line is obtained by using the fact $\nu(S) = \alpha\nu(M)$. To
 217 derive the fourth line, we used $\mu_W = 2^m\lambda\nu(M)$, and this equality is easily obtained by Eqs.
 218 (2) and (9). The dependences of the parameters λ , α , and β are qualitatively the same as
 219 those of Eq. (18). In addition, the asymptotic behaviors of Eq. (21) are equivalent to Eqs.
 220 (19) and (20). Fig. (4b) confirms the analytical evaluations of $E_{p_o}[\text{IC}]$.

221 Difference between PM and IC fractions appears with an intermediate mapping unit (see
 222 Fig. A.1 for a direct comparison), but the deviations are relatively small and these curves
 223 have similar forms, suggesting that the degree of clustering is not large.

224 3.2 Ecological survey with individual distributions based on the 225 Thomas process

226 Here we consider the situation where individuals are distributed according to the Thomas
 227 process. By Eq. (16), we calculate the presence mapped fraction for the Thomas process:

$$E_{th,(\beta,S,M)}[\text{PM}] = \frac{1 - p_{th}(N(S) = 0 | \beta)}{1 - p_{th}(N(M) = 0)}, \quad (22)$$

228 where, the probability of each event of the Thomas process is obtained by the pgfl Eq. (4):
 229 $p_{th}(n|A) = 1/n!(d^n G(t)/dt^n)|_{t=0}$. Therefore, $p_{th}(N(A) = 0)$ is

$$p_{th}(N(A) = 0) = \exp\left(-\lambda_p \int_{\hat{A}} \left[1 - \exp\left\{-\bar{c} \left(\int_A \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2}\right) d\mathbf{x}\right)\right\}\right] d\mathbf{y}\right), \quad (23)$$

230 where \hat{A} is the surrounding region of A where parents potentially provide daughters to the
 231 region A . Specifically, the second term inside the square brackets for $p_{th}(N(M) = 0)$ in Eq.
 232 (22) becomes $\exp(-\bar{c} \int_M \frac{1}{2\pi\sigma^2} \exp(-\|\mathbf{x} - \mathbf{y}\|^2/2\sigma^2) d\mathbf{x})$ and that of $p_{th}(0|\beta, S)$ becomes
 233 $\exp(-\alpha\beta\bar{c} \int_M \frac{1}{2\pi\sigma^2} \exp(-\|\mathbf{x} - \mathbf{y}\|^2/2\sigma^2) d\mathbf{x})$, due to the sampling density and the detectabil-
 234 ity. Although Eq. (22) with Eq. (23) is not easy to interpret, we can calculate its asymptotic

235 behaviors by the similar manner to the derivations of Eqs (19) and (20):

$$\lim_{M \rightarrow 0} E_{th,(\beta,S,M)}[\text{PM}] \simeq \alpha\beta, \quad (24)$$

$$\lim_{M \rightarrow W} E_{th,(\beta,S,M)}[\text{PM}] \simeq 1, \quad (25)$$

236 They are equivalent to the asymptotic behaviors of the homogeneous Poisson process Eqs.
237 (19) and (20). Fig. (4a) plots analytical and numerical results, showing the theoretical value
238 has a good agreement with the numerical calculation.

239 To obtain an explicit form for IC fraction of the Thomas process is cumbersome as the
240 pgfl of the Thomas process Eq. (6) is rather complex. Therefore, we only show the numerical
241 value for the IC fraction of the Thomas process (Fig. 4b). The IC for the Thomas process
242 increases faster than Eq. (18) as the mapping scale increases. The asymptotic behavior
243 shows similar trends to the other results.

244 Like in the case of the homogeneous Poisson process, difference between PM and IC
245 fractions appears outside the region where asymptotic behavior occurs (Fig. A.1). However,
246 the deviations are larger in this case, and it occurs with a wider range of the mapping unit
247 size. This is an effect of clustering distributions as discussed above.

248 4 Discussion

249 By explicitly accounting for the spatial distribution patterns of individuals through spatial
250 point processes (SPPs) and multiple spatial scales of field survey, we develop a theory for
251 ecological survey to map individual distributions. The theory quantifies two metrics, the
252 presence mapped fraction (PM fraction) and the fraction of individuals covered by the pres-
253 ence mapped patches (IC fraction), and thus allows us to predict the outcome of an ecological
254 survey under certain survey designs. When both the sampling density α and the detectability

255 β are not equal to 1, we find a tradeoff between the value of the PM and IC fractions and the
256 resolution of the map. The PM and IC fractions show the equivalent asymptotic behaviors
257 for both the homogeneous Poisson process and the Thomas process where $\alpha\beta$ and 1 are the
258 outcomes of the small and large asymptotic limit of mapping units, M , respectively. In fact,
259 these asymptotic limits are the same for any distribution patterns if an observation window
260 holds a sufficiently large number of individuals, which ensures that the probability to miss all
261 the individuals becomes zero. The fine limit of all these asymptotic behaviors are understood
262 as follows: as the mapping unit scale goes to sufficiently small, each mapping unit can hold
263 at most one individual. In such a situation, the probability to detect the single individual is
264 $\alpha\beta$. The asymptotic behavior suggests that there is a certain scale of the mapping unit above
265 or below which the performance of an ecological survey does not change. Thus, in practice,
266 we need to choose a scale of the mapping unit between these limits. The PM fraction of the
267 Thomas process first increases faster than that of the homogeneous Poisson process, because
268 the Thomas process produces mapping units holding clustered individuals which are more
269 likely to be found. However, the PM fraction of the homogeneous Poisson process approaches
270 to its asymptotic limit faster than that of the Thomas process. Because the Thomas process
271 also produces mapping units, due to the clustering pattern, holding a few individuals which
272 is difficult to map as a presence until the mapping unit becomes sufficiently large to hold
273 a sufficient individuals to make a chance of causing a false negative zero. This explanation
274 may be used to any distribution patterns. For example, if individual distributions show
275 highly clustered patterns, the PM fraction becomes steep firstly and becomes gentle as the
276 PM fraction approaches to the asymptotic value 1.

277 Spatial extension of the ecosystem that SPPs accounting individual aggregations de-
278 scribes could be large enough to cover a wide range of spatial scales. For example, Azaele et
279 al. [24] showed that a Thomas model fitted to the distribution map of British rare vascular
280 plant species (see the detailed description of the data set [29]) with three coarse resolutions

281 (40000, 10000, and 2500 km²) can outperform many existing spatially-implicit models in
282 terms of the down-scaling predictions of the species occupancy probability. In addition,
283 Grilli et al. [30] showed that a special case of the Poisson clustering processes, a group of the
284 point processes where parents locations are followed by a Poisson process [25] such as the
285 Neyman-Scott process, recovers the species-area relationship at a local scale to continental
286 scale as predicted by various existing models (e.g., [31]). Hence, even though we used a ob-
287 servation window $\nu(W) = 1024\text{m} \times 1024\text{m}$ as an example, it can be generalized by changing
288 its scale and the sampling intensity. In addition, it is worth noting that albeit individuals of
289 most species are typically aggregated [32,33] the Thomas process could be approximated by
290 the homogeneous Poisson process under a certain condition: when the intensity of individu-
291 als is large, the PM fraction of the Thomas process comes close to that of the homogeneous
292 Poisson process ($\bar{c}\lambda_p = \{10^{-2}, 10^{-1}\}$ in Fig. A.2). This is due to increased parent intensity
293 decreasing spatial heterogeneity over the region concerned, suggesting potential applicability
294 of the simpler model to an abundant ecosystem.

295 For simplicity, we consider a situation where each mapping unit is sampled with the same
296 sampling density, α , and detectability, β , and the location of the sampled unit within a map-
297 ping unit is chosen randomly. These are rather idealized assumptions and may be further
298 generalized. For example, it may be reasonable to assume that the sampling density, α , and
299 the detectability, β , become almost 1 at a certain fine scale of the mapping unit. Although
300 such a fine scale may not be achieved because of budgetary constraints, explicitly taking into
301 account the spatial effect on α and β gives us better understanding about the fine scale of
302 asymptotic behavior. In practice, the location of the sampling unit may be determined by
303 more strategic manner depending on ones purpose. Indeed, previous studies had proposed
304 several sampling strategies which emphasize, for example, a spatially contiguous placement
305 of the sampling units to correctly capture ecological patterns (e.g. [34]), a systematic place-
306 ment to efficiently reflect spatially structured ecological processes [35,36], or a representative

307 design for major environmental gradients to maximize per effort information of organism's
308 distribution [37, 38]. While these strategies have been compared empirically using actual
309 dataset (e.g. [36]), the developed theory in this paper may provide a theoretical base to
310 evaluate the effectiveness and efficiency of such purpose-dependent sampling strategies.

311

312 *Connection to occupancy area and population abundance*

313 Presence/absence map is often used to estimate the occupancy area or population abun-
314 dance [24, 39, 40]. Since our map contains estimated inaccuracy, we need to take into
315 account this effect to estimate these quantities. In our framework, occupancy area is
316 straightforward to obtain using the number of occupancy units. The number of occu-
317 pancy units is calculated from the PM fraction and the presence/absence map from a
318 ecological survey, since we have the relationship: ($\#$ presence mapped units) = $E[PM] \times$
319 ($\#$ total occupancy units). We can also derive the number of occupancy units using the
320 following relationship: ($\#$ total occupancy units) = $2^m(1 - p(N(M) = 0))$, where 2^m is the
321 number of mapping units. Unlike the tradeoffs between mapping resolution and PM or IC
322 fraction, this estimation is improved with a survey with a finer mapping unit since the shape
323 of an occupancy region is better mapped by a finer resolution, but with a certain finer limit.

324 Population abundance is also estimated by using the fact that each mapping unit at most
325 can hold one individual at a sufficiently small mapping scale. In this limit, the estimate
326 number of total occupancy units corresponds to the total population $N(W)$. In fact, we
327 have the following relationship, for example with the homogeneous Poisson process, $N(W) =$
328 $\lim_{M \rightarrow 0} 2^m(1 - p(N(M) = 0)) = \lambda\nu(W)$, by the equality $2^m = \nu(W)/\nu(M)$ and the same
329 expansion as in Eq. (19). $\lambda\nu(W)$ is the unbiased estimator of the total population due to
330 the definition Eq. (13). This estimation is, however, possible only if we have estimated
331 parameter values of the target species.

332

334 For the decision making on field survey designs, mapping resolution must be determined
335 to balance accuracy (i.e., the PM and/or IC fraction) and resolution of the map. Our results
336 show that accuracy of the map is improved with larger mapping resolution. However, it is
337 clear that presence/absence map with too coarse resolution is not practical for many ecolog-
338 ical studies and conservation/management practices. In addition, Takashina and Baskett [7]
339 showed that fisheries management with a coarse management unit inevitably increases in-
340 efficient efforts. Therefore, it may be reasonable to start first determining required map
341 accuracy, and secondly finding the finest possible mapping resolution which an expect PM
342 or IC fraction satisfies the requirement. To see this, let us discuss a rather simple and
343 ideal situation where we have estimations of each parameter value the population abun-
344 dance within a observation window. We assume that all the parameter values are the same
345 as in Fig. 4, and the target species has a clustering distribution pattern, which is described
346 by Thomas process (it corresponds to Fig. 4b. Let us further assume the situation where,
347 through the population viability analysis, we found 55% of the population in the region
348 must be protected to satisfy a 95% chance of persistence next 100 years. Therefore, the
349 minimum requirement of the ecological survey is to obtain the presence/absence map with
350 at least 55% of the total population covered within the presence mapped units. Then, by
351 making use of Fig. 4b, we find that the size of mapping resolution M is required to be about
352 64m^2 or larger to satisfy the requirement, which an expected value of the IC fraction is
353 $E_{th}[\text{IC}] = 0.57$. That is to say, we are expected to get a presence/absence map within which
354 57% of the total population is situated within the presence-mapped units. Of course this
355 example oversimplifies the ecological survey program, since we often do not have parameter
356 values of target species. However, the concept discussed above is rather general and hence
357 applicable to wide variety of ecological surveys. The core of this idea is to clearly set the
358 feasible goal, with time and budgetary constraints, of conservation practice or motivation of

359 ecological study in advance.

360 In practice, the developed theory for ecological survey should be, to an extent, com-
361 plemented by an estimation of the existing number of individuals within the observation
362 window, W since the intensity affects PM and IC fractions (Fig. A.2). An estimation of the
363 population abundance could be done by using historical or surrogate data. Statistical and
364 theoretical methods such as species distribution modeling [41] estimating the occurrence of
365 plant species across scale [24,42] or predicting the population abundance in a coral reef envi-
366 ronment [43] may complement these methods. Conducting a pilot survey is one alternative
367 way to estimate the population abundance with a required estimation accuracy. Takashina *et*
368 *al.* [28] recently developed a framework for the pilot sampling providing a required minimum
369 sampling effort to satisfy the required accuracy. Complemented by these steps, the theory
370 developed here has a potential to significantly improve survey frameworks.

371 Acknowledgements

372 We thank M. Akasaka, B. Stewart-Koster, T. Fung, R.A. Chisholm, L.R. Carrasco and S.
373 Azaele for their thoughtful comments and discussions. NT and BK were funded by the
374 Program for Advancing Strategic International Networks to Accelerate the Circulation of
375 Talented Researchers of the Japan Society for the Promotion of Science. They acknowledge
376 the support for coordinating the research program from Dr Yasuhiro Kubota and Dr James
377 D. Reimer.

378 References

379 [1] S. A. Levin, The problem of pattern and scale in ecology, *Ecology* 73 (6) (1992) 1943–
380 1967.

- 381 [2] E. K. Pikitch, C. Santora, E. a. Babcock, A. Bakun, R. Bonfil, D. O. Conover, P. Dayton,
382 P. Doukakis, D. Fluharty, B. Heneman, E. D. Houde, J. Link, Ecosystem-Based Fishery
383 Management, *Science* 300 (July) (2004) 2003–2003. doi:[10.1126/science.1106929](https://doi.org/10.1126/science.1106929).
- 384 [3] C. White, B. S. Halpern, C. V. Kappel, Ecosystem service tradeoff analysis reveals the
385 value of marine spatial planning for multiple ocean uses., *Proc. Natl. Acad. Sci. U. S.*
386 *A.* 109 (12) (2012) 4696–4701. doi:[10.1073/pnas.1114215109](https://doi.org/10.1073/pnas.1114215109).
- 387 [4] N. Takashina, A. Mougi, Maximum sustainable yields from a spatially-explicit harvest
388 model, *J. Theor. Biol.* 383 (21) (2015) 87–92.
- 389 [5] C. R. Margules, R. L. Pressey, Systematic conservation planning., *Nature* 405 (6783)
390 (2000) 243–53. [arXiv:j.1744-7429.2008.00471.x](https://arxiv.org/abs/j.1744-7429.2008.00471.x), doi:[10.1038/35012251](https://doi.org/10.1038/35012251).
- 391 [6] T. Yamakita, H. Yamamoto, M. Nakaoka, H. Yamano, K. Fujikura, K. Hidaka, Y. Hi-
392 rota, T. Ichikawa, S. Kakehi, T. Kameda, S. Kitajima, K. Kogure, T. Komatsu,
393 N. H. Kumagai, H. Miyamoto, K. Miyashita, H. Morimoto, R. Nakajima, S. Nishida,
394 K. Nishiuchi, S. Sakamoto, M. Sano, K. Sudo, H. Sugisaki, K. Tadokoro, K. Tanaka,
395 Y. Jintsu-Uchifune, K. Watanabe, H. Watanabe, Y. Yara, N. Yotsukura, Y. Shi-
396 rayama, Identification of important marine areas around the Japanese Archipelago:
397 Establishment of a protocol for evaluating a broad area using ecologically and bi-
398 ologically significant areas selection criteria, *Mar. Policy* 51 (2015) 136–147. doi:
399 [10.1016/j.marpol.2014.07.009](https://doi.org/10.1016/j.marpol.2014.07.009).
- 400 [7] N. Takashina, M. L. Baskett, Exploring the effect of the spatial scale of fishery man-
401 agement, *J. Theor. Biol.* 390 (2016) 14–22. doi:[10.1016/j.jtbi.2015.11.005](https://doi.org/10.1016/j.jtbi.2015.11.005).
- 402 [8] H. Possingham, I. Ball, S. Andelman, Mathematical methods for identifying represen-
403 tative reserve networks, in: *Quant. Methods Conserv. Biol.*, Springer New York, 2000,
404 pp. 291–306. doi:[10.1007/0-387-22648-6_17](https://doi.org/10.1007/0-387-22648-6_17).

- 405 [9] K. A. Wilson, M. I. Westphal, H. P. Possingham, J. Elith, Sensitivity of conserva-
406 tion planning to different approaches to using predicted species distribution data, *Biol.*
407 *Conserv.* 122 (1) (2005) 99–112. [doi:10.1016/j.biocon.2004.07.004](https://doi.org/10.1016/j.biocon.2004.07.004).
- 408 [10] P. De Ornellas, E. J. Milner-Gulland, E. Nicholson, The impact of data realities on con-
409 servation planning, *Biol. Conserv.* 144 (7) (2011) 1980–1988. [doi:10.1016/j.biocon.](https://doi.org/10.1016/j.biocon.2011.04.018)
410 [2011.04.018](https://doi.org/10.1016/j.biocon.2011.04.018).
- 411 [11] A. Moilanen, Implications of empirical data quality to metapopulation model parameter
412 estimation and application, *Oikos* 96 (2002) 516–530. [doi:10.1034/j.1600-0706.](https://doi.org/10.1034/j.1600-0706.2002.960313.x)
413 [2002.960313.x](https://doi.org/10.1034/j.1600-0706.2002.960313.x).
- 414 [12] M. A. McCarthy, J. L. Moore, W. K. Morris, K. M. Parris, G. E. Garrard, P. A.
415 Vesk, L. Rumpff, K. M. Giljohann, J. S. Camac, S. S. Bau, T. Friend, B. Harrison,
416 B. Yue, The influence of abundance on detectability, *Oikos* 122 (5) (2013) 717–726.
417 [doi:10.1111/j.1600-0706.2012.20781.x](https://doi.org/10.1111/j.1600-0706.2012.20781.x).
- 418 [13] S. A. Field, A. J. Tyre, H. P. Possingham, Optimizing allocation of monitoring effort
419 under economic and observational constraints, *J. Wildl. Manag.* 69 (2) (2005) 473–482.
420 [doi:10.2193/0022-541X\(2005\)069\[0473:OAOMEU\]2.0.CO;2](https://doi.org/10.2193/0022-541X(2005)069[0473:OAOMEU]2.0.CO;2).
- 421 [14] D. I. Mackenzie, J. A. Royle, Designing occupancy studies: General advice and allocat-
422 ing survey effort (2005). [doi:10.1111/j.1365-2664.2005.01098.x](https://doi.org/10.1111/j.1365-2664.2005.01098.x).
- 423 [15] R. S. Epanchin-Niell, R. G. Haight, L. Berec, J. M. Kean, A. M. Liebhold, Optimal
424 surveillance and eradication of invasive species in heterogeneous landscapes, *Ecol. Lett.*
425 15 (8) (2012) 803–812. [doi:10.1111/j.1461-0248.2012.01800.x](https://doi.org/10.1111/j.1461-0248.2012.01800.x).
- 426 [16] A. J. Tyre, B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, H. P. Possingham,
427 Improving precision and reducing bias in biological surveys: Estimating false-negative
428 error rates (2003). [doi:10.1890/02-5078](https://doi.org/10.1890/02-5078).

- 429 [17] B. A. Wintle, M. A. McCarthy, K. M. Parris, M. A. Burgman, Precision and bias
430 of methods for estimating point survey detection probabilities (2004). [doi:10.1890/
431 02-5166](https://doi.org/10.1890/02-5166).
- 432 [18] G. E. Garrard, S. A. Bekessy, M. A. McCarthy, B. A. Wintle, When have we looked hard
433 enough? A novel method for setting minimum survey effort protocols for flora surveys,
434 *Austral Ecol.* 33 (8) (2008) 986–998. [doi:10.1111/j.1442-9993.2008.01869.x](https://doi.org/10.1111/j.1442-9993.2008.01869.x).
- 435 [19] E. C. Pielou, *An introduction to mathematical ecology*, Wiley-Interscience, New York,
436 1969.
- 437 [20] E. C. Pielou, *Ecological Diversity*, John Wiley and Sons Inc, New York, 1975.
- 438 [21] M. Bode, J. Sanchirico, P. Armsworth, Returns from matching management resolution
439 to ecological variation in a coral reef fishery, *Proc. R. Soc. B.* 283 (1826) (2016) 20152828.
- 440 [22] N. A. C. Cressie, *Statistics for Spatial Data*, John Wiley & Sons, New York, 1993.
- 441 [23] J. B. Plotkin, M. D. Potts, N. Leslie, N. Manokaran, J. Lafrankie, P. S. Ashton, Species-
442 area curves, spatial aggregation, and habitat specialization in tropical forests., *J. Theor.
443 Biol.* 207 (1) (2000) 81–99. [doi:10.1006/jtbi.2000.2158](https://doi.org/10.1006/jtbi.2000.2158).
- 444 [24] S. Azaele, S. J. Cornell, W. E. Kunin, Downscaling species occupancy from coarse
445 spatial scales, *Ecol. Appl.* 22 (3) (2012) 1004–1014. [doi:10.1890/11-0536.1](https://doi.org/10.1890/11-0536.1).
- 446 [25] S. N. Chiu, D. Stoyan, W. S. Kendall, J. Mecke, *Stochastic Geometry and Its Applica-*
447 *tions*, John Wiley & Sons, New York, 2013.
- 448 [26] S. Muko, K. Shimatani, Y. Nozawa, *Spatial analyses for nonoverlapping objects with*
449 *size variations and their application to coral communities* (2014).

- 450 [27] T. S. Bayard, C. S. Elphick, Using Spatial Point-Pattern Assessment to Understand the
451 Social and Environmental Mechanisms that Drive Avian Habitat Selection, *Auk* 127 (3)
452 (2010) 485–494. doi:10.1525/auk.2010.09089.
- 453 [28] N. Takashina, B. Kusumoto, M. Beger, S. Rathnayake, H. P. Possingham, Spatially
454 Explicit Approach To Population Abundance Estimation In Field Surveys, bioRxivdoi:
455 <https://doi.org/10.1101/131037>.
- 456 [29] S. Hartley, W. E. Kunin, J. J. Lennon, M. J. O. Pocock, Coherence and discontinuity in
457 the scaling of species' distribution patterns., *Proc. Biol. Sci.* 271 (1534) (2004) 81–88.
458 doi:10.1098/rspb.2003.2531.
- 459 [30] J. Grilli, S. Azaele, J. R. Banavar, A. Maritan, Spatial aggregation and the species-area
460 relationship across scales, *J. Theor. Biol.* 313 (2012) 87–97. arXiv:arXiv:1209.3591v1,
461 doi:10.1016/j.jtbi.2012.07.030.
- 462 [31] R. H. MacArthur, E. O. Wilson, *The theory of island biogeography*, Vol. 1, Princeton
463 University Press, 1967. doi:10.2307/1796430.
- 464 [32] L. R. Taylor, I. P. Woiwod, J. N. Perry, The Density-Dependence of Spatial Behaviour
465 and the Rarity of Randomness, *J. Anim. Ecol.* 47 (2) (1978) pp. 383–406. doi:10.
466 2307/3790.
- 467 [33] R. Condit, P. S. Ashton, P. Baker, S. Bunyavejchewin, S. Gunatilleke, N. Gunatilleke,
468 S. P. Hubbell, R. B. Foster, A. Itoh, J. V. LaFrankie, H. S. Lee, E. Losos, N. Manokaran,
469 R. Sukumar, T. Yamakura, Spatial patterns in the distribution of tropical tree species,
470 *Science* 288 (1982) (2000) 1414–1418. doi:10.1126/science.288.5470.1414.
- 471 [34] B. Güler, A. Jentsch, I. Apostolova, S. Bartha, J. B. M.G., G. Campetella, R. Canullo,
472 J. HÁzi, J. Kreyling, J. Pottier, G. Szabó, T. Terziyska, E. Uurlu, C. Wellstein, Z. Zim-
473 mermann, J. Dengler, How plot shape and spatial arrangement affect plant species

- 474 richness counts: implications for sampling design and rarefaction analyses, *J. Veg.*
475 [Sci.doi:10.1111/jvs.12411](https://doi.org/10.1111/jvs.12411), 2016.
- 476 [35] D. Bellhouse, *Biometrika*, *Biometrika* 64 (3) (1977) 605–611.
- 477 [36] I. J. Fiske, E. M. Bruna, Alternative spatial sampling in studies of plant demography:
478 Consequences for estimates of population growth rate, *Plant Ecol.* 207 (2) (2010) 213–
479 225. [doi:10.1007/s11258-009-9666-4](https://doi.org/10.1007/s11258-009-9666-4).
- 480 [37] A. N. Gillison, K. R. W. Brewer, The use of gradient directed transects or gradsects in
481 natural resource surveys, *J. Environ. Manage.* 20 (April) (1985) 103–127.
- 482 [38] H. M. Neave, R. B. Cunningham, T. W. Norton, H. A. Nix, Preliminary evaluation
483 of sampling strategies to estimate the species richness of diurnal, terrestrial birds
484 using Monte Carlo simulation, *Ecol. Modell.* 95 (1) (1997) 17–27. [doi:10.1016/
485 S0304-3800\(96\)00016-6](https://doi.org/10.1016/S0304-3800(96)00016-6).
- 486 [39] W. E. Kunin, Extrapolating species abundance across spatial scales, *Science* 281 (5382)
487 (1998) 1513–1515. [doi:10.1126/science.281.5382.1513](https://doi.org/10.1126/science.281.5382.1513).
- 488 [40] W. H. Hwang, F. He, Estimating abundance from presence/absence maps, *Methods*
489 *Ecol. Evol.* 2 (5) (2011) 550–559. [doi:10.1111/j.2041-210X.2011.00105.x](https://doi.org/10.1111/j.2041-210X.2011.00105.x).
- 490 [41] S. Phillips, R. Anderson, R. Schapire, Maximum entropy modeling of species geographic
491 distributions, *Ecol. Modell.* 190 (3-4) (2006) 231–259. [doi:10.1016/j.ecolmodel.
492 2005.03.026](https://doi.org/10.1016/j.ecolmodel.2005.03.026).
- 493 [42] F. He, K. J. Gaston, Estimating Species Abundance from Occurrence, *Am. Nat.* 156 (5)
494 (2000) 553–559. [doi:10.1086/303403](https://doi.org/10.1086/303403).

495 [43] C. Mellin, S. Andrefouet, D. Ponton, Spatial predictability of juvenile fish species rich-
496 ness and abundance in a coral reef environment, *Coral Reefs* 26 (4) (2007) 895–907.
497 [doi:10.1007/s00338-007-0281-3](https://doi.org/10.1007/s00338-007-0281-3).

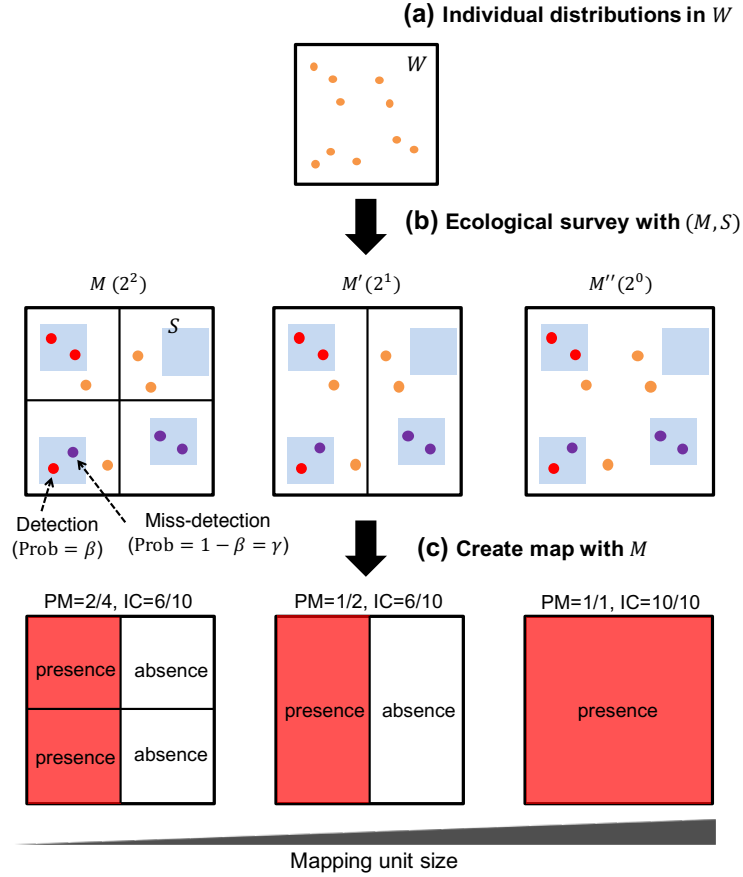


Figure 3: (color online) Ecological survey scheme within the observation window W . (a) Given the individual distributions in the observation window W , (b) ecological survey is conducted with a certain mapping resolution M (Middle left, for example) and sampling unit $S = \alpha M$ (blue regions). Each column represents the result of an ecological survey with a different mapping resolution M (M', M''). 2^2 ($2^1, 2^0$) in the parentheses represents the number of mapping unit within the observation window. (c) With the survey outcome, a presence/absence map is created. If at least one individual is found in a mapping unit (Middle: represented by red point), regardless of miss detecting other individuals situated therein (represented by purple or orange), the unit is mapped as presence, absence otherwise. In this step, PM fraction and IC fraction (see main text for the definitions) are calculated by simply counting the number of presence patches or the number of individuals situated within the (mapped) presence patches. Although the same individual distributions and survey outcome are used, obtained map differs if another mapping resolution is used.

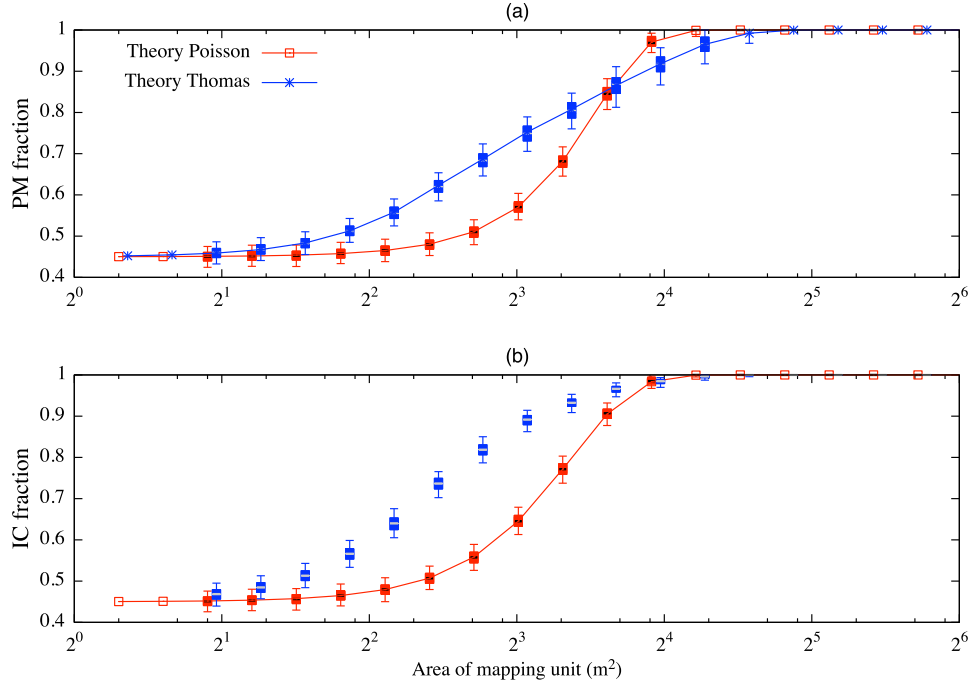


Figure 4: (color online) Analytical and simulated (candlestick) values of (a) the presence mapped fraction (PM fraction); and (b) the fraction of individuals covered within presence mapped patches (IC fraction) across mapping unit scales. x -axis is the area of mapping unit (m^2). Each candlestick shows, from the bottom, 5, 25, 50, 75, and 95 percentile values of 1000 simulation trials. The values of the sampling density and detectability are $\alpha = 0.5$ and $\beta = 0.9$, respectively. The other parameter values are the same as in Fig. 1.

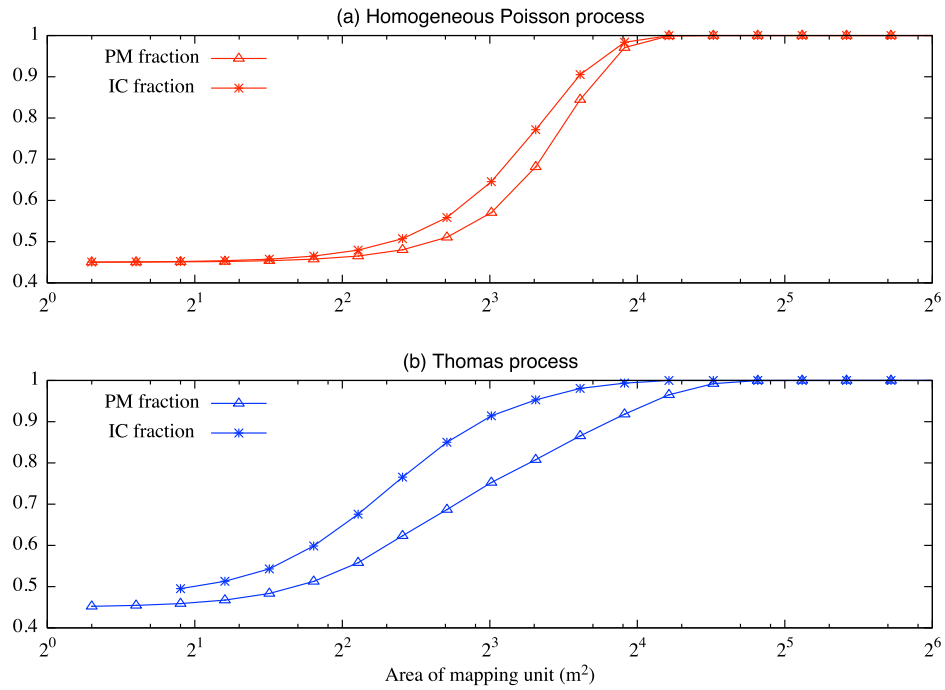


Figure A.1: (color online) Analytical and simulated (only IC fraction of Thomas process) values of (a) the homogeneous Poisson process; and (b) Thomas process. All the values are the same as in Fig. 4 in the main text, but different presentation to facilitate the comparison of PM and IC fractions of each process.

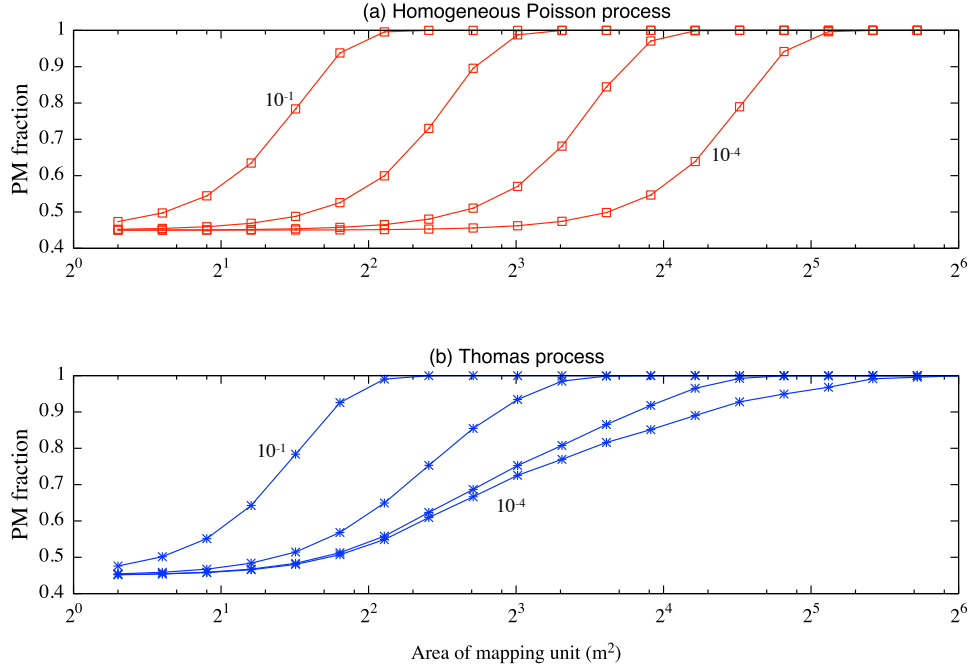


Figure A.2: Effect of the intensity (λ , $\bar{c}\lambda_p$) in the observation window, W , on the theoretical presence mapped (PM) fraction, Eqs. (18), (22). The intensity of the Thomas process is manipulated by changing the parent intensity λ_p . Individual distribution patterns are according to the (a) Homogeneous Poisson process and (b) Thomas process. For the Thomas process, the curves for PM fraction converge as the intensity becomes small, and come close to the corresponding curve of the homogeneous Poisson process as the intensity of the Thomas process increases. This is an effect that the increased parents intensity decreases spatial heterogeneity over the concerned region. For both panels, the order of the intensity monotonically decreases from left to right.