



HAL
open science

On using stratified two-stage sampling for large-scale multispecies surveys

Philippe Aubry, Gwenaël Quaintenne, Jeremy Dupuy, Charlotte Francesiaz,
Matthieu Guillemain, Alain Caizergues

► **To cite this version:**

Philippe Aubry, Gwenaël Quaintenne, Jeremy Dupuy, Charlotte Francesiaz, Matthieu Guillemain, et al.. On using stratified two-stage sampling for large-scale multispecies surveys. *Ecological Informatics*, 2023, 77, pp.102229. 10.1016/j.ecoinf.2023.102229 . hal-04258681

HAL Id: hal-04258681

<https://ofb.hal.science/hal-04258681v1>

Submitted on 25 Oct 2023

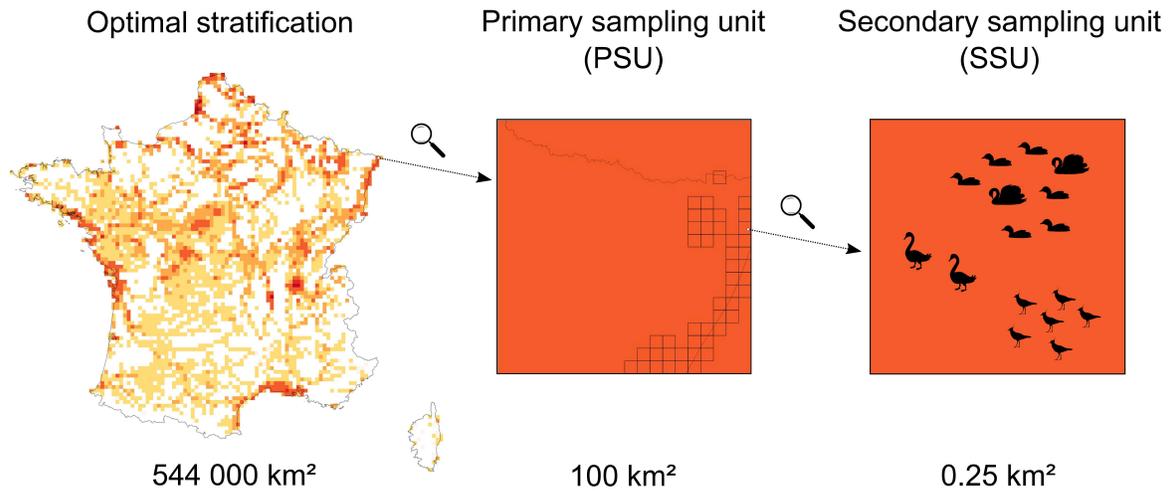
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Graphical Abstract

On using stratified two-stage sampling for large-scale multispecies surveys

Philippe Aubry, Gwenaël Quintenne, Jeremy Dupuy, Charlotte Francesiaz, Matthieu Guillemain, Alain Caizergues



Highlights

On using stratified two-stage sampling for large-scale multispecies surveys

Philippe Aubry, Gwenaël Quaintenne, Jeremy Dupuy, Charlotte Francesiaz, Matthieu Guillemin, Alain Caizergues

- We address the problem of designing a multispecies survey using multistage sampling.
- The multispecies nature of the survey sampling design is based on species richness.
- A distribution atlas documents the species richness for primary sampling units.
- Stratification of the primary sampling units aims to oversample the richest units.
- A stratified two-stage sampling design for a nationwide bird survey is documented.

On using stratified two-stage sampling for large-scale multispecies surveys

Philippe Aubry^{a,*}, Gwenaël Quintenne^b, Jeremy Dupuy^b, Charlotte Francesiaz^c, Matthieu Guillemain^d, Alain Caizergues^e

^aOFB - Office français de la biodiversité - Direction surveillance, évaluation, données - Unité données et appui méthodologique, Saint Benoist, BP 20, 78612 Le Perray-en-Yvelines, France

^bLPO - Ligue pour la protection des oiseaux - Pôle protection de la Nature - Service connaissance, Fonderies royales, CS 90263, 17300 Rochefort Cedex, France

^cOFB - Office français de la biodiversité - Direction de la recherche et de l'appui scientifique - Service conservation et gestion durable des espèces exploitées, Les Portes du Soleil, 147 avenue de Lodève, 34990 Juvignac, France

^dOFB - Office français de la biodiversité - Direction de la recherche et de l'appui scientifique - Service conservation et gestion durable des espèces exploitées, La Tour du Valat, Le Sambuc, 13200 Arles, France

^eOFB - Office français de la biodiversité - Direction de la recherche et de l'appui scientifique - Service conservation et gestion durable des espèces exploitées, 8 Bd Albert Einstein, 44000 Nantes, France

Abstract

Designing a large-scale multispecies survey presents significant challenges due to its multivariate, multiscale, and often multiobjective nature. Because it is impractical to measure or observe all variables of interest at every location in the survey area, spatial sampling is crucial. To ensure a statistically sound survey, free of assumptions, randomization of the selection of spatial sampling units is required.

In a large-scale survey, minimizing travel costs requires multistage sampling in which nearby sampling units are grouped within higher-level sampling units. The design may be supplemented by one or more strata systems — possibly nested — that intervene at specific sampling stages. This article focuses on stratified two-stage sampling that balances statistical efficiency and travel cost minimization.

As presurvey knowledge, we assume that the spatial distribution of the species of interest is available for the sampled spatial domain, either from a spatial database or a published species distribution atlas. The multispecies nature of the survey sampling plan relies on these data. The primary sampling units are defined by the spatial grid resolution of the distribution data, while the presence data are used to stratify the PSUs to ensure that the sampling effort is concentrated where most species occur.

In this article, we formally define the *sampling coverage probability of a species distribution* when the selection of samples is randomized according to a given probability mass function (i.e., a *probability sampling plan* or *sampling design*, in the statistical sense of the term). By maximizing the sampling coverage probability averaged over the species, the stratification and allocation of spatial sampling units concentrate observer effort where most species occur and minimize effort elsewhere.

To guide and illustrate our ideas, we present as an example a multispecies waterbird program at the national scale in European France within the framework of probability sampling from finite populations.

Keywords:

multispecies surveys, spatial sampling, stratification, multistage sampling, species richness, design-based framework

1. Introduction

Large-scale ecological studies face numerous impediments and challenges due to the nature of the features under study. They require the measurement or observation of numerous biotic and abiotic variables (e.g., Legendre et al., 1989), and because of the changing nature of ecological processes with spatial scale, they are plagued by the problem of scale dependence of many parameters of interest (Wiens, 1989; Steele, 1991; Whittaker et al., 2001; Hewitt et al., 2007, 2017; Schneider, 1994, 2009). Measurement issues (Olsen et al., 1999) and the overall cost of collecting field data over large areas are additional hurdles.

*Corresponding author

Email addresses: philippe.aubry@ofb.gouv.fr (Philippe Aubry), gwenael.quintenne@lpo.fr (Gwenaël Quintenne), jeremy.dupuy@lpo.fr (Jeremy Dupuy), charlotte.francesiaz@ofb.gouv.fr (Charlotte Francesiaz), matthieu.guillemain@ofb.gouv.fr (Matthieu Guillemain), alain.caizergues@ofb.gouv.fr (Alain Caizergues)

9 Large-scale studies may consist either of one-shot surveys (i.e., surveys conducted only once) or
10 surveys repeated over time for monitoring purposes (Waite, 2000, Sec. 1.4). The goal may be to
11 assess, for one or more parameters, the status (i.e., the value at a given time point), the change (i.e.,
12 the difference between two time points) or a trend (i.e., a smooth pattern of variation over time). In
13 all cases, spatial sampling is required simply because it is impossible to cover the entire survey area,
14 especially when it is large.

15 Whether or not a survey program is repeated over time, a first limitation consists in investigating
16 biological variables (e.g., presence/absence, abundance) but not abiotic variables. The focus may be
17 on parameters for a single species — for example a threatened one — or the survey may be inherently
18 multispecific when addressing community questions (e.g., Morin, 2011; Sutherland et al., 2013; Keddy
19 and Laughlin, 2022), especially when an abundance-based biodiversity index is the parameter of interest
20 (e.g., Hill, 1973; Patil and Taillie, 1982; Magurran, 2004; Buckland et al., 2011, 2012). In biodiversity
21 assessment, as a pragmatic necessity, only a subset of species can be surveyed. This raises the question
22 of whether this will yield useful results and how these species should be selected (e.g., Lambeck, 1997;
23 Manley et al., 2004 and references cited therein; Wiens et al., 2008; Klibansky et al., 2017).

24 Given the current global biodiversity crisis, documenting and monitoring the conservation status
25 of species has become a critical issue. The shift from single-species to multispecies conservation raises
26 the question of program effectiveness, time and cost savings. In this context, a multispecies survey
27 offers a practical approach to collect data on groups of species simultaneously over large or diverse
28 regions. Such an approach maximizes data collection per field effort (DeWan and Zipkin, 2010). Thus,
29 even if the focus tends to be on independent assessment for each species, it is cost-effective to survey
30 a group of species based on their taxonomy or life history characteristics (e.g., breeding in a common
31 habitat).

32 The multispecies nature of surveys is always challenging. Nevertheless, such surveys can rely on
33 large networks of observers, at least for certain groups of species, mainly animals, and especially diurnal
34 birds, which are probably the most widely documented of all animal groups. It is not possible to cover
35 all types of multispecies surveys for all taxa in all types of environments. Although our focus is not
36 limited to monitoring nor to the case of birds, the examples given below will essentially refer to bird
37 monitoring programs because of their overrepresentation in the ecological and wildlife literature. Of
38 course, what is presented and discussed in this article may be useful in other contexts.

39 Beyond the simple presence/absence, which already provides useful knowledge, abundance —
40 whether absolute or relative — is a variable of interest in most multispecies surveys. However, counting
41 wildlife is a much more difficult task than it appears to be at first glance, even for diurnal conspicuous
42 species such as many bird species. In this article, we will not discuss counting methods and imperfect
43 detection issues (see, among others, Lancia et al., 1996; Lloyd et al., 1998; Pollock et al., 2002; Green-
44 wood and Robinson, 2006a; Buckland et al., 2008; Dénés et al., 2015), as we are concerned only with
45 large-scale spatial sampling issues.

46 Large-scale monitoring programs are crucial from a legal perspective because they play a key role in
47 the classification of species on the IUCN Red List of Threatened Species (IUCN, 2022) and thus drive
48 policy decisions. For birds, monitoring of the conservation status of species is carried out through
49 programs implemented in different parts of the world and on different spatial scales, the oldest of
50 which dates back to the early 20th century. These include the famous *Christmas Bird Counts* (CBC),
51 the *International Waterbird Census* (IWC), and the *Pan-European Common Bird Monitoring Scheme*
52 (PECBMS).

53 Although the tradition of bird counting is quite old and bird counting programs cover different
54 areas, including entire continents, little effort has been devoted to the statistical design of spatial
55 sampling to obtain reliable estimates of parameters of interest (means, changes, trends, etc.). In
56 general, multispecies bird surveys fall into the category of nonprobability sampling.

57 Several forms of nonprobability sampling can be distinguished, in particular (i) *convenience sam-*
58 *pling* (sometimes called *accessibility sampling*, see for example Young and Young, 1998, p. 93 or
59 Barnett, 2002, p. 17); (ii) *preferential sampling*; and (iii) *purposive sampling* (also known as *judgment*
60 *sampling*). In the first case, units are selected for inclusion in the sample for reasons of convenience,
61 generally ease of access and/or safety for the observers. In the second case, sampling units are selected
62 according to partial knowledge or subjective judgments about the variables of interest, for example,
63 by selecting units where high abundances are expected or which are thought to be the most diverse
64 or species-rich. In the third case, units judged to be typical or appropriate for the survey are selected
65 from the population.

66 For example, the *Dutch Breeding Bird Monitoring Program* (BMP) relies on both convenience and
67 preferential sampling since the observers are free to choose their study areas, and in each habitat,
68 they may prefer the most attractive sites, i.e., those that are species-rich and have high bird densities

69 (Van Turnhout et al., 2008). An example of purposive sampling is the *International Waterbird Count*
70 (IWC), a site-based counting scheme for monitoring waterbird numbers organized by Wetlands Inter-
71 national, where sites are defined by the judgment of national coordinators and local observers (Delany,
72 2010, Sec. 4).

73 Convenience, preferential and purposive sampling are nonprobability ways of selecting a sample,
74 that is, with unknown selection probabilities. They are prone to biased inference if the selection process
75 cannot be finely modeled or if external reliable data sources are not used to calibrate the estimators.
76 As pointed out by Olsen et al. (1999, p. 28): “A clear understanding of how sites are selected is
77 necessary but not sufficient to avoid the impact of selection bias when making inferences [...]”.

78 For example, in convenience sampling, the ease of sampling is often accompanied by a lack of
79 awareness of the magnitude and direction of bias, which ultimately leads to unresolved questions
80 about the true reliability of the estimates. In contrast, probability sampling is very strict, both in
81 terms of the definition of the sampling units and in terms of the selection process: (i) the spatial
82 sampling units are defined in advance — and preferably a list of units (*sampling frame*) is set up
83 to allow for a fixed sample size — and (ii) the sample selection process corresponds to a prescribed
84 probability mass function $p(s)$, which defines the probability of selecting a sample s . The distribution
85 $p(s)$ characterizes a *probability sampling plan* (e.g., Royall, 2006, p. 2330; Nguyen, 2006, Sec. 2.1) or
86 *sampling design* in the statistical sense of the term (e.g., Hedayat and Sinha, 1991, p. 3, Definition
87 1.1; Särndal et al., 1992, p. 27). Since the selection probabilities of the samples are known exactly,
88 it follows that the estimates are obtained in a sound and objective statistical framework (*design-based*
89 *estimation*, e.g., Aubry and Francesiaz, 2022).

90 Irrespective of the field considered, in most large-scale surveys, the design-based paradigm has
91 become dominant in the theory and practice of survey sampling, and the reasons for this have been
92 reviewed by Bellhouse (1988) from a historical perspective. Despite regular calls for reliance on proba-
93 bility sampling in fundamental or applied ecology (Albert et al., 2010; Smith et al., 2017; Aubry et al.,
94 2020; Boyd et al., 2023), many programs deviate from such recommendations to a greater or lesser
95 extent. Some programs have some degree of randomization, but not to the extent provided by a (true)
96 probability sampling plan.

97 Reliable estimation involves minimizing both bias and uncertainty, the latter of which is caused
98 by the (spatial) sampling variance. This variance is irrelevant in ecology, except as a nuisance to be
99 minimized (see Link and Nichols, 1994). There are two main approaches to sampling optimization,
100 depending on whether the spatial sampling (i) is guided by a model (e.g., Amorim et al., 2014; Carvalho
101 et al., 2016) or (ii) relies on a probability sampling plan (e.g., Pavlacky et al., 2017; Van Wilgenburg
102 et al., 2020) — with a continuum of possible intermediates that mix the use of both to diverse degrees
103 (e.g., Marta et al., 2019). The present article follows the second approach, that is, it addresses the
104 optimization of a probability sampling plan.

105 Our interest in the problem of large-scale spatial sampling for a multispecies survey was motivated
106 by the example of a scheme aimed at surveying migratory waterbirds during the breeding season on
107 a national scale (European France). The aim of this methodological article is twofold: (i) to propose
108 an approach for optimizing the stratification of sampling units when only a proxy for species richness
109 is available at the design stage and (ii) to present an application of the stratification approach to our
110 illustrative example in the case of a stratified two-stage sampling design, detailing the choices made
111 to face the objectives and constraints.

112 2. Motivating example

113 The Office français de la biodiversité (OFB) — government agency dedicated to the protection and
114 restoration of biodiversity in France — and the Ligue pour la protection des oiseaux (LPO) — the
115 largest French nature conservation NGO — have joined forces to design and implement a nationwide
116 survey (referred to as the LIMAT scheme) to collect data on breeding populations of waterbirds (ducks,
117 geese, swans, coots, waders and grebes) to assess their conservation status in France and in Europe.
118 The main objective of this survey is to estimate the size of breeding populations in European France,
119 which will be used as a regular contribution to the Article 12 report of the EU Birds Directive. This
120 report is produced every sixth year to document assessments such as The State of Nature in the EU
121 (European Environment Agency, 2020), the French and European Red Lists of Birds (IUCN France,
122 MNHN, LPO, SEOF & ONCFS, 2016; International BirdLife, 2021) and African-Eurasian Waterbird
123 Population Estimates (Nagy and Langendoen, 2021). As the species covered by the LIMAT scheme
124 may be greatly affected by hunting activities — and many of these species are in decline — the results of
125 the survey may also contribute to the work on breeding phenologies regularly requested from Member
126 States by the European Commission (Key Concepts of Article 7(4) of Directive 79/409/EEC).

127 The initial implementation of the LIMAT scheme was planned to take place over two field seasons
128 due to the workload involved in such a survey. Among the species covered by the LIMAT scheme,
129 geographically rare species or those with less abundant populations were the subject of an exhaustive
130 approach (census survey), while other more common and widespread species required the implemen-
131 tation of a sample survey (see Table 1). Only the latter situation is considered here.

132 Whenever a descriptive survey cannot be carried out in the form of a census — i.e., whenever it
133 is not possible to list and/or visit all breeding and resting sites for a given species — sampling must
134 be implemented. Allowing for the sampling of observation sites at the discretion of observers leads to
135 serious difficulties in using the data collected (Gregory and Greenwood, 2008, Sec. 3.2.5) because (i)
136 serious sampling biases can be introduced and (ii) statistical processing requires making assumptions
137 that cannot be verified — at least not without further studies that should themselves be free of bias.
138 Therefore, to collect observations for scientific purposes or for compliance with an environmental law,
139 it is advisable to use a (probability) areal sampling approach that avoids the two problems mentioned
140 above. Thus, we decided to rely on the theory and methods of probability sampling from finite
141 populations (see, among others, Cochran, 1977; Särndal et al., 1992; Hankin et al., 2019; Tillé, 2020).
142 Statistical estimation was planned to be performed in a (pure) design-based framework — at least as
143 far as the issue of imperfect detection is not yet accounted for (see Aubry and Francesiaz, 2022, Sec.
144 7.3).

145 In the LIMAT scheme, several objectives and constraints have led us to consider a (relatively)
146 complex sampling design: (a) the fact that we were interested in several species at the same time,
147 whose geographical distributions differ (global scale) and which do not necessarily occupy exactly the
148 same habitats (local scale); (b) the need to optimize sampling effort, i.e., to collect observations for
149 as many species as possible in a limited number of visits; (c) the desire to limit as much as possible
150 the sources of bias of all kinds that can arise in such a survey; and (d) the need to minimize travel
151 between observation sites.

152 The European part of France is divided into 96 departments of various surface areas (see Table
153 2). The *department* is an intermediate administrative subdivision in France and corresponds to the
154 territorial jurisdiction of various stakeholders involved in the LIMAT scheme, so the sample size in
155 each department is important from a logistical point of view. The stakeholders requested that the
156 number of selected ultimate sampling units — i.e., those on which observations were to be made —
157 in each department should not exceed 80 to avoid an excessive survey burden. In what follows, this
158 requirement is referred to as the *logistical constraint*.

159 The sampling design chosen should be able to establish a trade-off between all the aforementioned
160 objectives and constraints. It also necessarily depends on prior knowledge available before the survey
161 is carried out. Since this was the first implementation of a survey based on a sampling design in
162 European France for the species of interest, we did not have reliable quantitative data to determine
163 the expected precision of the estimates. In addition, mapping of potential habitats for each of the
164 species considered was not available. However, to document the distribution of species in European
165 France, range maps of breeding species based on the French Birds Atlas (Issa and Muller, 2015) were
166 available and updated for the postatlas period with citizen-science data collected on the Faune-France
167 web portal (<https://www.faune-france.org/>) in September 2020. These range maps were available
168 at a spatial resolution of $10 \text{ km} \times 10 \text{ km}$. In what follows, the recorded presence of a species in a 10
169 $\text{km} \times 10 \text{ km}$ grid cell is referred to as *occurrence*.

170 An areal sampling frame of $10 \text{ km} \times 10 \text{ km}$ grid cells can be constructed from the bird atlas data.
171 Let U_d be the set of size M_d containing the $10 \text{ km} \times 10 \text{ km}$ grid cells where species d occurs, with
172 $d = 1, \dots, D$ (here, $D = 25$, see Table 1). We call the set U_d the (spatial) *distribution* of species d
173 in European France. The union of the D species distributions leads to a potential common sampling
174 frame U of size $M = 5\,211$:

$$U = \bigcup_{d=1}^D U_d$$

175 Let x be the number of species (or *species richness*) per grid cell. For species richness ranging from 1
176 to $L = 22$, the frequency distribution of x shows a quasimonotone decrease (Fig. 1). The species dis-
177 tributions are pictured in Fig. 2; they show diverse situations from the exclusively coastal distribution
178 of the Kentish plover (*Charadrius alexandrinus*) to the widespread distribution of the mallard (*Anas*
179 *platyrhynchos*) through the distribution of the common sandpiper (*Actitis hypoleucos*) along flowing
180 rivers and streams.

Table 1: List of the $D = 25$ species covered by the spatial sampling in European France, sorted in increasing order of the number of occurrences (M_d) among the $M = 5\,211$ grid cells ($10\text{ km} \times 10\text{ km}$) with at least one species occurrence. % indicates the percentage of the grid cells in which the species occurs.

Rank	Species	Common name	M_d	%
1	<i>Tringa totanus</i>	Common redshank	167	3.2
2	<i>Mergus merganser</i>	Goosander	179	3.4
3	<i>Recurvirostra avosetta</i>	Pied avocet	185	3.6
4	<i>Anser anser</i>	Greylag goose	187	3.6
5	<i>Charadrius alexandrinus</i>	Kentish plover	199	3.8
6	<i>Podiceps nigricollis</i>	Black-necked grebe	213	4.1
7	<i>Anas crecca</i>	Common teal	266	5.1
8	<i>Netta rufina</i>	Red-crested pochard	267	5.1
9	<i>Spatula querquedula</i>	Garganey	280	5.4
10	<i>Alopochen aegyptiaca</i>	Egyptian goose	329	6.3
11	<i>Spatula clypeata</i>	Northern shoveler	382	7.3
12	<i>Himantopus himantopus</i>	Black-winged stilt	392	7.5
13	<i>Mareca strepera</i>	Gadwall	408	7.8
14	<i>Branta canadensis</i>	Canada goose	502	9.6
15	<i>Aythya fuligula</i>	Tufted duck	560	10.7
16	<i>Tadorna tadorna</i>	Common shelduck	562	10.8
17	<i>Actitis hypoleucos</i>	Common sandpiper	594	11.4
18	<i>Aythya ferina</i>	Common pochard	688	13.2
19	<i>Vanellus vanellus</i>	Northern lapwing	1 674	32.1
20	<i>Charadrius dubius</i>	Little ringed plover	1 763	33.8
21	<i>Cygnus olor</i>	Mute swan	1826	35.0
22	<i>Podiceps cristatus</i>	Great crested grebe	2 354	45.2
23	<i>Tachybaptus ruficollis</i>	Little grebe	2 861	54.9
24	<i>Fulica atra</i>	Eurasian coot	3 027	58.1
25	<i>Anas platyrhynchos</i>	Mallard	5 001	96.0

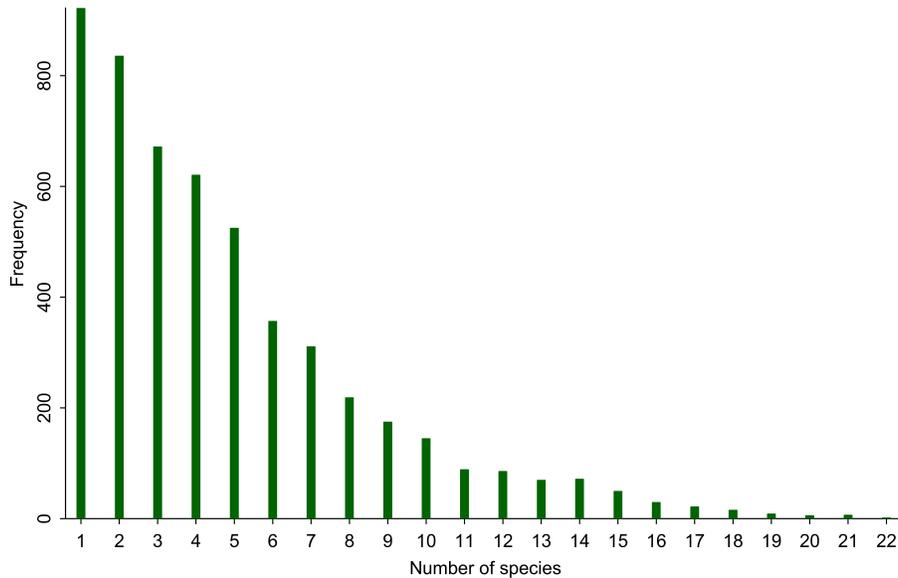


Figure 1: Frequency distribution of the number of species (x) among the $M = 5\,211$ grid cells ($10\text{ km} \times 10\text{ km}$) with at least one species occurrence.

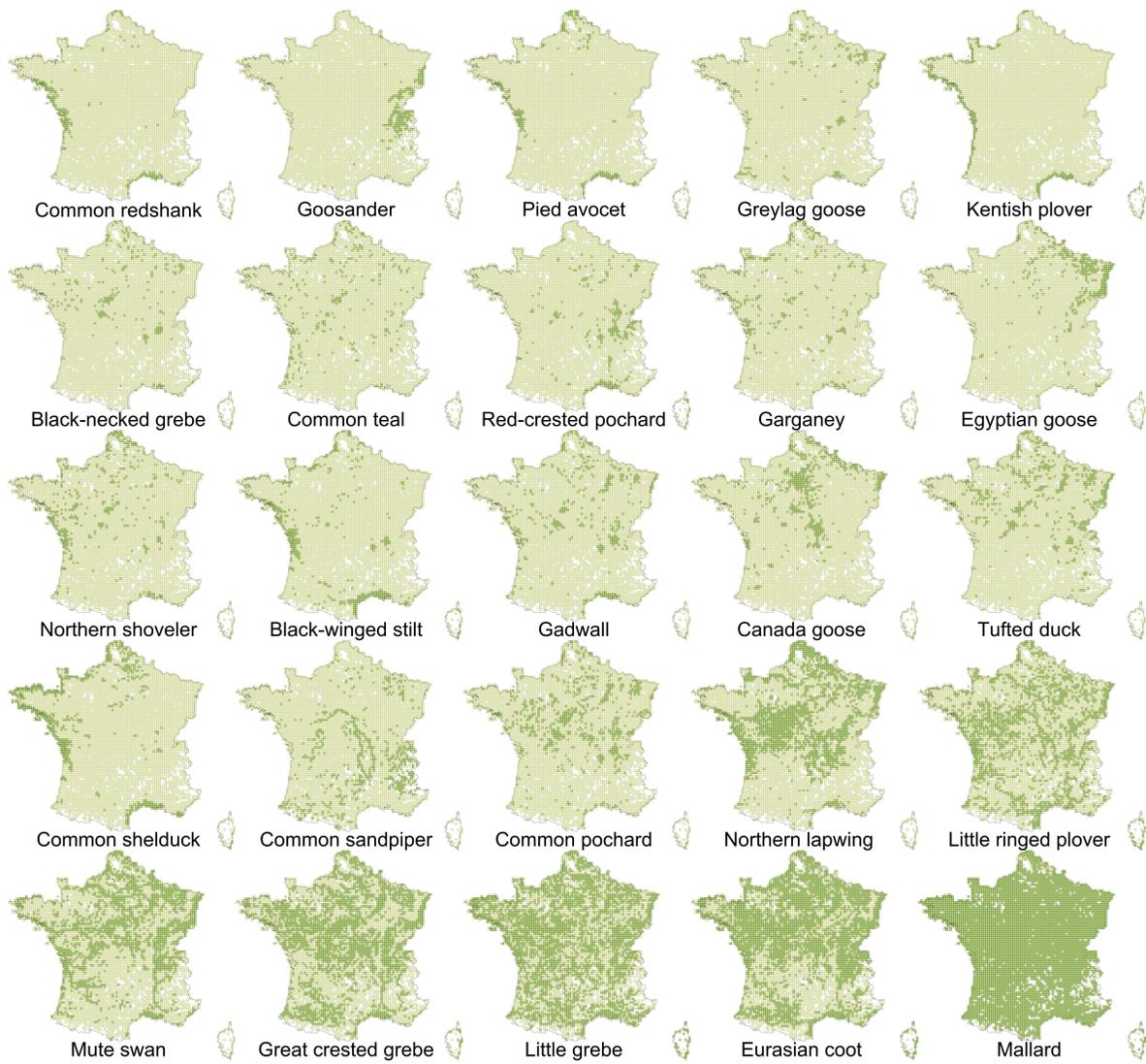


Figure 2: Distributions of the $D = 25$ species covered by the spatial sampling in European France, sorted in increasing order of the number of occurrences M_d — from left to right and from top to bottom — among the $M = 5\,211$ grid cells ($10\text{ km} \times 10\text{ km}$) with at least one species occurrence (light green background).

Table 2: List of the 96 departments that make up the European France. k is the department identifier. A is the cadastral surface area, expressed in km^2 . Note that department $k = 20$ corresponding to the island of Corsica was split in two in 1976, resulting in departments 2A and 2B (so the department identifier 20 is no longer used).

k	Name	A	k	Name	A	k	Name	A
01	Ain	5 762	32	Gers	6 257	64	Pyrénées-Atlantiques	7 645
02	Aisne	7 369	33	Gironde	9 976	65	Hautes-Pyrénées	4 464
03	Allier	7 340	34	Hérault	6 101	66	Pyrénées-Orientales	4 116
04	Alpes-de-Haute-Provence	6 925	35	Ille-et-Vilaine	6 775	67	Bas-Rhin	4 755
05	Hautes-Alpes	5 549	36	Indre	6 791	68	Haut-Rhin	3 525
06	Alpes-Maritimes	4 299	37	Indre-et-Loire	6 127	69	Rhône	3 249
07	Ardèche	5 529	38	Isère	7 432	70	Haute-Saône	5 360
08	Ardennes	5 229	39	Jura	4 999	71	Saône-et-Loire	8 575
09	Ariège	4 890	40	Landes	9 243	72	Sarthe	6 206
10	Aube	6 004	41	Loir-et-Cher	6 343	73	Savoie	6 028
11	Aude	6 139	42	Loire	4 781	74	Haute-Savoie	4 388
12	Aveyron	8 735	43	Haute-Loire	4 977	75	Paris	105
13	Bouches-du-Rhône	5 088	44	Loire-Atlantique	6 809	76	Seine-Maritime	6 278
14	Calvados	5 548	45	Loiret	6 775	77	Seine-et-Marne	5 915
15	Cantal	5 726	46	Lot	5 217	78	Yvelines	2 284
16	Charente	5 956	47	Lot-et-Garonne	5 361	79	Deux-Sèvres	5 999
17	Charente-Maritime	6 864	48	Lozère	5 167	80	Somme	6 170
18	Cher	7 235	49	Maine-et-Loire	7 172	81	Tarn	5 758
19	Corrèze	5 857	50	Manche	5 938	82	Tarn-et-Garonne	3 718
2A	Corse-du-Sud	4 014	51	Marne	8 162	83	Var	5 973
2B	Haute-Corse	4 666	52	Haute-Marne	6 211	84	Vaucluse	3 567
21	Côte-d'Or	8 763	53	Mayenne	5 175	85	Vendée	6 720
22	Côtes-d'Armor	6 878	54	Meurthe-et-Moselle	5 246	86	Vienne	6 990
23	Creuse	5 565	55	Meuse	6 211	87	Haute-Vienne	5 520
24	Dordogne	9 060	56	Morbihan	6 823	88	Vosges	5 874
25	Doubs	5 233	57	Moselle	6 216	89	Yonne	7 427
26	Drôme	6 530	58	Nièvre	6 817	90	Territoire de Belfort	609
27	Eure	6 040	59	Nord	5 743	91	Essonne	1 804
28	Eure-et-Loir	5 880	60	Oise	5 860	92	Hauts-de-Seine	176
29	Finistère	6 733	61	Orne	6 103	93	Seine-St-Denis	236
30	Gard	5 853	62	Pas-de-Calais	6 671	94	Val-de-Marne	245
31	Haute-Garonne	6 309	63	Puy-de-Dôme	7 970	95	Val-d'Oise	1 246

181 3. Methodology

182 To address the various objectives and constraints mentioned above, we propose to rely on stratified
183 two-stage sampling — combining the advantages of stratification and two-stage sampling (for bird
184 surveys, see [Gregory and Greenwood, 2008](#), Sec. 3.2.4, 3.2.9) — with optimized stratification defined
185 from the species distributions.

186 Let $p(s)$ be a sampling design, i.e., a probability mass function corresponding to a well-specified
187 random mechanism for selecting a sample from a population. The expectation operator is defined over
188 all possible samples drawn using the given sampling design and is usually denoted using the subscript
189 p as $\mathbb{E}_p[\cdot]$ to avoid confusion with other possible sources of stochasticity.

190 3.1. Sampling coverage of species distributions

191 Let s be a sample of size m selected by a sampling design $p(s)$; this sample intersects the species
192 distribution U_d , resulting in the subset $s_d \subseteq U_d$, i.e., formally:

$$s_d = s \cap U_d$$

193 Of course, s_d is a random set since it depends on s , which is itself randomly selected according to the
194 sampling design $p(s)$. Therefore, the size of s_d (denoted m_d) is a random integer variable. We can
195 define the sampling coverage probability of U_d by the design $p(s)$ as:

$$\pi_d = \mathbb{E}_p \left[\frac{a_d}{A_d} \right] = \frac{\mathbb{E}_p[a_d]}{A_d} \quad (1)$$

196 where a_d and A_d are the surface areas for s_d and U_d , respectively. Again, a_d depends on s and is
 197 therefore a random variable. If all the sampling units have the same surface area, then the coverage
 198 expressed in terms of the number of sampling units is equal to the coverage expressed in terms of
 199 surface area, and we have:

$$\pi_d = \mathbb{E}_p \left[\frac{m_d}{M_d} \right] = \frac{\mathbb{E}_p [m_d]}{M_d} \quad (2)$$

200 Without loss of generality, in what follows, we assume that the sampling units are all of the same size,
 201 as is usually the case with distribution atlas data (see [Gibbons et al., 2007](#); [Dunn and Weston, 2008](#)).

202 Consider first that the sampling design is *simple random sampling without replacement* (SRSWOR).
 203 For a given species d , the population U is partitioned into U_d and $U - U_d$. Using an SRSWOR, the
 204 random size m_d follows a hypergeometric distribution with parameters M , M_d and m . As a result, its
 205 expected value is $\mathbb{E}_p(m_d) = mM_d/M$, and therefore, we have $\pi_d = m/M$, which does not depend on
 206 species d .

207 Then, consider that the population U is partitioned into strata denoted U_h (of size M_h) for $h =$
 208 $1, \dots, H$, that is, a set of mutually nonoverlapping subpopulations, formally:

$$U = \bigcup_{h=1}^H U_h \quad \text{with} \quad U_h \cap U_{h'} = \emptyset, \quad h \neq h' \quad \text{and} \quad M = \sum_{h=1}^H M_h$$

209 For a given species d , its spatial distribution U_d is crossed with the partition by the strata, resulting
 210 in H classes denoted U_{dh} (of size M_{dh}) for $h = 1, \dots, H$, with possibly $M_{dh} = 0$ for some strata (Fig.
 211 [3](#)). With a stratified sampling design, the sample s itself is partitioned into $s = \cup_{h=1}^H s_h$, and the
 212 sample size is partitioned as $m = \sum_{h=1}^H m_h$. The same applies to the size m_d , which is partitioned as
 213 $m_d = \sum_{h=1}^H m_{dh}$, with possibly $m_{dh} = 0$ for some strata (Fig. [3](#)).

214 If the sampling design is *stratified simple random sampling without replacement* (STR-SRSWOR),
 215 then the coverage probability is written as:

$$\pi_d = \frac{1}{M_d} \mathbb{E}_p \left[\sum_{h=1}^H m_{dh} \right] = \frac{1}{M_d} \sum_{h=1}^H \mathbb{E}_p [m_{dh}] = \frac{1}{M_d} \sum_{h=1}^H m_h \frac{M_{dh}}{M_h} \quad (3)$$

216 In contrast to the SRSWOR case, formula [\(3\)](#) shows that the coverage probability now depends on
 217 species d through M_{dh} and M_d . Thus, stratification provides a way to vary the coverage probability
 218 π_d .

219 For more complex sampling designs than STR-SRSWOR, the derivation of the closed form for
 220 π_d (Eq. [2](#)) may become too complicated or even impossible. In such a situation, by replicating the
 221 sampling design enough times (e.g., at least 10^3 times) and counting the number of times species d is
 222 represented in sample s , one obtains a Monte Carlo approximation of $\mathbb{E}_p [m_d]$, from which that of π_d
 223 is derived.

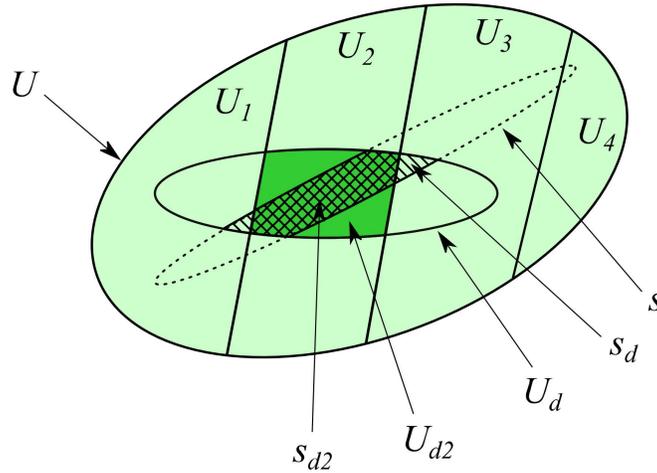


Figure 3: Set relations between the population U , stratified into strata U_1, U_2, U_3 and U_4 (light green background), a species distribution U_d (in solid line) inside the population of sampling units, and a sample s (in dotted line). The intersection between U_d and stratum U_4 is empty. The intersection between U_d and stratum U_2 is U_{d2} (bright green background). The intersection between U_d and s is s_d (hatch and grid patterns). The intersection of U_{d2} and s — or equivalently, between U_2 and s_d — is s_{d2} (grid pattern on bright green background).

224 3.2. Stratified two-stage sampling design

225 A stratified two-stage sampling design combines a two-stage sampling design with a stratified sam-
 226 pling design at the first stage. In other words, a two-stage sampling design is applied independently
 227 within each stratum. This design involves a three-level hierarchy defined from the sampled spatial do-
 228 main (in our motivating example, European France) to the elementary or ultimate sampling units, i.e.,
 229 spatial sampling units that are not further disaggregated and on which measurements or observations
 230 are made. These three levels are as follows (Fig. 4a): (i) a first level composed of *strata*, i.e., groups
 231 of units homogeneous with respect to a criterion to be defined (the strata are not sampled, they are
 232 all represented in the final sample selected); (ii) a second level composed of *primary sampling units* or
 233 PSUs, which group together units of lower level; and (iii) a third level composed of *secondary sampling*
 234 *units* or SSUs, which are the elementary sampling units, included in the PSUs.

235 In this article, we rely on SRSWOR to sample the PSUs and the SSUs. Thus, within each stratum,
 236 a sample of PSUs is selected by SRSWOR; this is the first-stage sampling (Fig. 4b). Within each PSU
 237 selected at the first stage, a sample of SSUs is selected by SRSWOR; this is the second-stage sampling
 238 (Fig. 4c). Sampling in one stratum is independent of sampling in another, and the same is true for
 239 sampling in the PSUs (*independence*, see, for instance, Särndal et al., 1992, pp. 101-102, 134-135 or
 240 Tillé, 2020, pp. 67-68, 149).

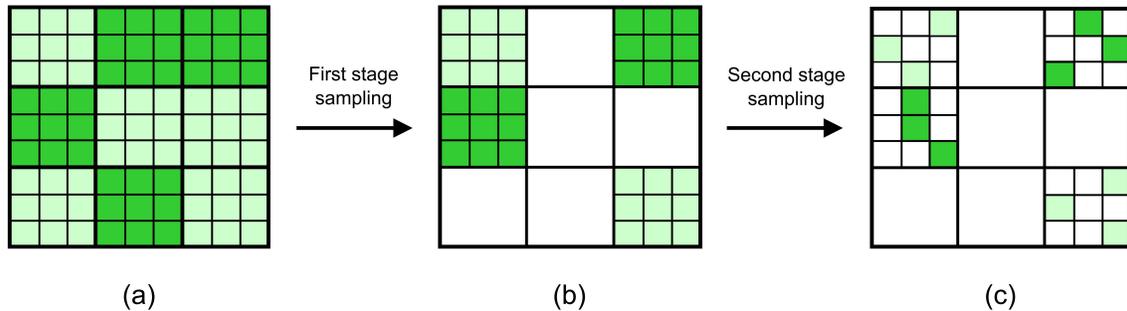


Figure 4: Schematic representation of stratified two-stage sampling. (a) Nine primary sampling units (PSUs) are stratified into two strata (bright and light green backgrounds). Each PSU consists of nine secondary sampling units (SSUs). (b) At the first stage, two PSUs per stratum are selected. (c) At the second stage, three SSUs are selected for each PSU of the first stage sample.

241 *3.3. Stratification*

242 We assume that species distribution data are available on the scale of the PSUs. In this section,
243 we propose an appropriate method for defining PSU stratification based on this assumption.

244 First, we note that the species distributions U_d ($d = 1, \dots, D$) do not define a partition of U since
245 several species can occur in the same PSU. Therefore, a set of species distributions is not a strata
246 system. The strata must be derived from the species distributions by specifying the following: (i)
247 a criterion to be optimized, (ii) a within-strata sample size allocation rule, and (iii) an appropriate
248 optimization method.

249 *3.3.1. Criterion to be optimized*

250 Species distributions can have very different area occupancies and spatial configurations. The
251 fact that we are interested in several species at the same time, along with the need to optimize the
252 sampling effort to collect observations for a maximum number of species in a limited number of visits
253 (see objectives and constraints mentioned above), led us to try to maximize the overall sampling
254 coverage probability for a given number of strata (denoted as H). As a criterion to be maximized, we
255 can choose, for example, the average sampling coverage probability of species distributions, formally:

$$J_1 = E_p \left[\frac{1}{D} \sum_{d=1}^D \frac{m_d}{M_d} \right] = \frac{1}{D} \sum_{d=1}^D \pi_d \quad (4)$$

256 For a given sampling effort (i.e., the overall sampling size m), it is less likely that a sample will broadly
257 cover the distribution of geographically rarer species, for which M_d are the lowest — i.e., of lower
258 prevalence (or weight) $W_d = M_d/M$. To counterbalance this effect of geographical rarity — i.e., to
259 cover the distributions of rare species with a higher probability than by using criterion J_1 — one
260 may still want to maximize the overall sampling coverage probability, but this time focusing on the
261 geographically rare species. For this purpose, we propose a weighted, more general version:

$$J_2 = E_p \left[\left(\sum_{d=1}^D W_d^{-\lambda} \right)^{-1} \sum_{d=1}^D W_d^{-\lambda} \frac{m_d}{M_d} \right] = \left(\sum_{d=1}^D W_d^{-\lambda} \right)^{-1} \sum_{d=1}^D W_d^{-\lambda} \pi_d \quad (5)$$

262 with $\lambda \geq 0$. Criterion J_1 is an instance of criterion J_2 for $\lambda = 0$ (the same weight is used for all
263 species). In this article, we will only use criterion J_1 because, at this stage of our thinking, we have
264 no argument for choosing the value of $\lambda > 0$ other than arbitrarily.

265 *3.3.2. Sample size allocation rule*

266 At the first sampling stage, the overall sample size m must be allocated to the H strata according
267 to a specified rule.

268 Let $x_i \geq 1$ be the number of species whose presence is recorded in the i -th PSU ($i = 1, \dots, M$).
269 The variable x can be termed as a *size variable* since it assigns a size (or importance) to the sampling
270 units. In the context of a multispecies survey, to collect observations for a maximum number of species
271 in a limited number of visits, a PSU i with a value $x_i < x_j$ is less interesting to include in the sample
272 than the PSU j . However, all PSUs in the sampled population must have a nonzero probability of
273 being included in the sample by definition of a sampling design (e.g., Särndal et al., 1992, p. 32, Eq.
274 2.4.6). Taking these elements into account, it follows that the allocation of the sample size can be
275 proportional to the x -variable total (in a classical survey sampling context, see, for example, Särndal
276 et al., 1992, p. 108 or Hidirolou and Lavallée, 2009, Sec. 4.3.2.2):

$$m_h = m \frac{X_h}{X} \quad (6)$$

277 with the partitioning of the overall x -total:

$$X = \sum_{i=1}^M x_i = \sum_{h=1}^H \sum_{i=1}^M Z_{1ih} x_i = \sum_{h=1}^H X_h \quad (7)$$

278 where Z_{1ih} ($i = 1, \dots, M$, $h = 1, \dots, H$) is the membership indicator:

$$Z_{1ih} = \begin{cases} 1 & \text{if PSU } i \text{ belongs to stratum } h \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

279 This allocation of the sampling effort overrepresents strata whose PSUs have a greater number of
 280 species present. Of course, for the variables of interest, this overrepresentation is accounted for in
 281 the definition of the estimators for the population mean or total, so that the estimators are design-
 282 unbiased. To estimate the sampling variance of the estimators, it is necessary to ensure that the
 283 within-stratum variance is estimable. Consequently, a minimum sample size per stratum must be
 284 imposed, i.e., $m_h \geq m_{\min}$. Formally (see, for instance, [Ardilly and Tillé, 2006](#), p. 160 or [Bethlehem,](#)
 285 [2009](#), pp. 116-117), the absolute minimum is $m_{\min} = 2$, but in practice, we have to take a higher value
 286 (i) to anticipate the nonresponse (otherwise, we might end up with $m_h < 2$) and (ii) because reliable
 287 within-stratum sampling variance estimation often requires more than two values.

288 3.3.3. Optimization method

289 Obtaining a stratification that maximizes J_1 (for example) involves solving a combinatorial opti-
 290 mization problem. The first step is to determine whether it is possible to obtain a globally optimal
 291 stratification scheme in a reasonable amount of time. Our problem is simplified by considering a single
 292 size variable (univariate case). For the sake of consistency, PSUs with the same number of species
 293 should belong to the same stratum. Thus, the problem is not expressed in terms of partitioning the
 294 PSUs but in terms of partitioning the (sorted) values of the number of species per PSU; each PSU will
 295 be assigned to the class that contains its number of species. The strata are constructed sequentially in
 296 the order induced by the size variable (here, in decreasing order, so that the first stratum corresponds
 297 to the most species-rich PSUs).

298 Let $\mathcal{L} = (\mathcal{L}_1 > \mathcal{L}_2 > \dots > \mathcal{L}_L)$ be the sequence of values that the integer-valued size variable can
 299 take. Let $P_{L,H}$ be the set of partitions of \mathcal{L} into H contiguous classes. The size of $P_{L,H}$ is written as
 300 (e.g., [Aubry, 2023b](#), Eq. 6):

$$|P_{L,H}| = \prod_{h=1}^{H-1} \frac{1}{h} (L - h) = \binom{L-1}{H-1} \quad (9)$$

301 This size obviously depends on the data through the number of different values of species occurrences
 302 (L) and on the given number of strata (H). The combinatorial enumeration (Eq. 9) allows us to
 303 examine which values of L and H lead to a problem size that can be handled in a reasonable computing
 304 time (see [Aubry, 2023b](#), p. 16). When the set of contiguous partitions to be considered is not too large,
 305 it follows that we can find a globally optimal partition — i.e., a partition that maximizes the criterion
 306 — by exhaustive construction of all distinct contiguous partitions (complete enumeration). For this
 307 purpose, a procedure based on the simulation of nested loops has been described by [Aubry \(2023b,](#)
 308 [Algorithm 10\)](#) and will not be reproduced here. The general algorithm consists of the following steps:

- 309 0. Set $J_1^* \leftarrow 0$;
- 310 1. Construct a new stratification of the L values of the size variable into H strata (if there is no
 311 further new stratification, then the optimization is finished, J_1^* is the global optimum for J_1 and
 312 the saved stratification is optimal);
- 313 2. Compute the strata sizes M_h for $h = 1, \dots, H$;
- 314 3. If a stratum size M_h is inadequate (e.g., too small). Go back to step 1;
- 315 4. Compute the totals X_h for $h = 1, \dots, H$ (Eq. 7);
- 316 5. Compute the sample sizes m_h for $h = 1, \dots, H$ (Eq. 6) and round them while keeping their sum
 317 equal to m ;
- 318 6. Compute criterion J_1 for the current stratification and allocation (Eq. 4, here using Eq. 3);
- 319 7. If $J_1 > J_1^*$, then set $J_1^* \leftarrow J_1$ and save the current stratification. Go back to step 1.

320 4. Application to the motivating example

321 4.1. Sampling units and sampling frames

322 In our motivating example, the primary sampling units (PSUs) are 10 km \times 10 km grid cells, while
 323 the secondary sampling units (SSUs) are 500 m \times 500 m grid cells (Fig. 5).

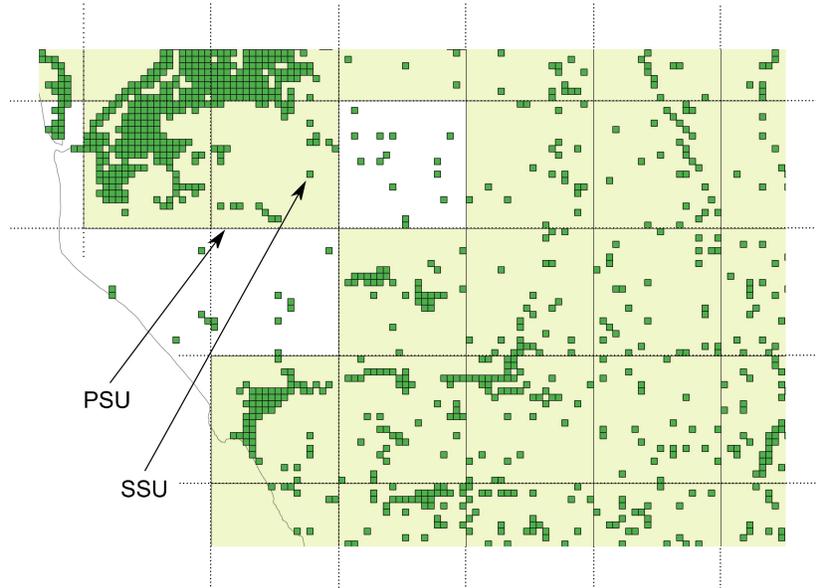


Figure 5: The primary sampling units (PSUs) are $10 \text{ km} \times 10 \text{ km}$ grid cells, and the secondary sampling units (SSUs) are $500 \text{ m} \times 500 \text{ m}$ grid cells. PSUs with fewer than 15 SSUs are removed from the first-stage sampling frame (white background). Details are provided in the text.

324 The SSUs were defined to carry out the field observations. Initially, we considered $250 \text{ m} \times 250$
 325 m grid cells but it appeared that field observations were quite feasible on $500 \text{ m} \times 500 \text{ m}$ grid cells.
 326 There was a trade-off between the size of the sampling frame — through the surface area of the SSUs
 327 — and the practical feasibility of SSU visits in the field.

328 In these SSUs, the fieldwork method was left to the discretion of the observer. However, combining
 329 point counts with observations by transect to cover the maximum of each SSU surface area was strongly
 330 recommended. Most of the data were collected using a smartphone application, called *NaturaList*, from
 331 BioloVision (https://www.faune-france.org/index.php?m_id=20015), which allows the observers
 332 to record their movements within the SSUs, as well as the locations of detected birds. *NaturaList* is
 333 recommended by the LPO and the *European Bird Census Council* (EBCC).

334 The second-stage sampling frame construction (based on elevation and habitat considerations) is
 335 not discussed here, as it is highly dependent on both the species of interest and the context in which
 336 the survey was prepared (available data and time). Furthermore, it is not necessary to be familiar
 337 with the process of second-stage sampling frame construction, as this article is essentially concerned
 338 with the stratification of PSUs. What is important, however, is the number of SSUs per PSU, as this
 339 determines which PSUs could remain in the sampling frame to meet a sample size constraint in each
 340 PSU selected at the first stage (see below).

341 Restricting the European France $10 \text{ km} \times 10 \text{ km}$ grid to land areas only results in a set of 5875
 342 potential PSUs. This number is reduced to 5211 if only PSUs with at least one recorded species
 343 occurrence (at the time of survey design stage, see Section 2) are considered. As noted by Samuel and
 344 Garton (1994), excluding PSUs where no (or relatively few) individuals of the species of interest are
 345 present makes sense and should be cost effective.

346 Let N_i be the number of SSUs included in the i -th PSU. For each PSU i selected at the first stage,
 347 we drew by SRSWOR a sample of constant size $n_i = 10 + 5$ SSUs, i.e., with ten regular SSUs and a
 348 reserve of five SSUs to allow substitution in case of nonresponse due to access difficulties (fenced area,
 349 access denied, etc.). In the reserve, the SSUs were numbered from 1 to 5; if one of the ten planned
 350 SSUs could not be used, the observers in the field could replace it with the first unit from the reserve,
 351 and so on. Thus, the observers could not choose the replacement SSU, but rather the selection was
 352 randomized and planned in advance to avoid selection bias. Due to the sample size mentioned above,
 353 we kept only those PSUs that guaranteed that this second-stage sample size constraint would be met.
 354 We added PSUs with no recorded species presence if they contained at least 15 SSUs (this additional
 355 set included only 48 PSUs). The final sampling frame included 2839 PSUs (Fig. 6). Using an overall
 356 sampling fraction of approximately 10%, the first stage sample contained $m = 285$ PSUs.

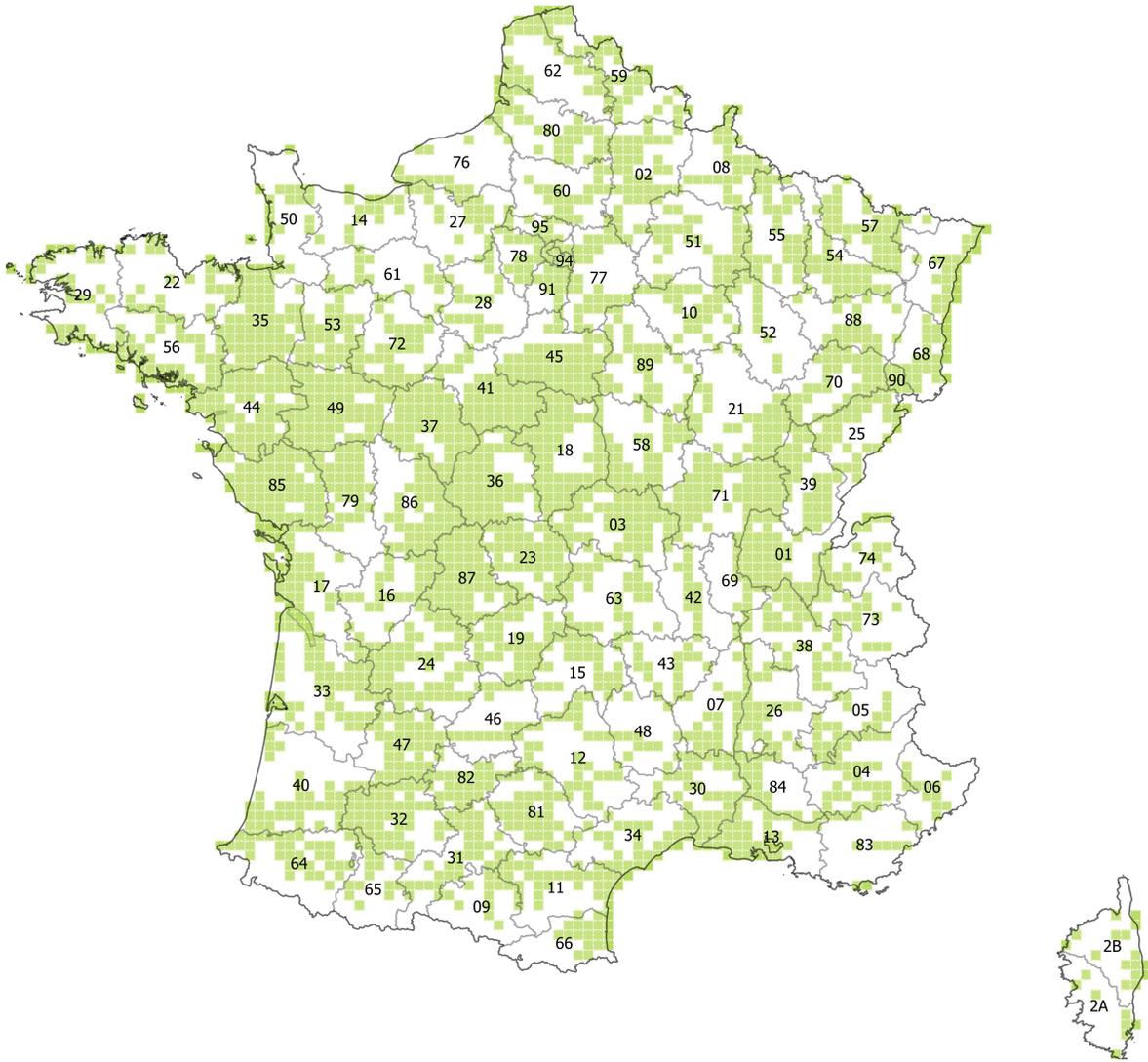


Figure 6: The final sampling frame included $M = 2839$ PSUs with at least 15 SSUs. European France is partitioned into 96 departments — the department is an intermediate administrative subdivision in France — corresponding to the territorial jurisdiction of various stakeholders involved in the LIMAT scheme, for which there is a logistical constraint. Details are provided in the text.

357 *4.2. Stratification*

358 The principle of PSU stratification was adopted to maximize the average sampling coverage prob-
 359 ability of species distributions in the case of STR-SRSWOR. However, as the PSU sampling frame
 360 contained 48 PSUs with no observed species, these PSUs had to be grouped into an extra stratum
 361 that obviously could not participate in the optimal definition of stratification as intended. In this
 362 last stratum, by default, we considered that the status of the PSUs, in terms of species present, was
 363 undetermined, and we applied a sampling fraction of approximately 10%, resulting in the sampling of
 364 five of these PSUs. For the sake of consistency and because it is a reasonable lower bound, we required
 365 a minimum sample size of five PSUs in each stratum. This constraint on the sample size per stratum
 366 imposes the same lower bound on the size of the strata, i.e., we must have $M_h \geq m_{\min}$, bearing in
 367 mind that the sampling fraction in a stratum can be as high as 100% (the so-called *take-all stratum*).

368 Stratification was therefore performed in two steps: (i) optimal stratification into H strata of the
 369 remaining 2791 PSUs with at least one species present and (ii) addition of an $(H + 1)$ -th stratum
 370 comprising the 48 PSUs whose status was considered undetermined. For the optimization, we used a
 371 sample size of $m = 280$ PSUs, as five PSUs were assigned to the $(H + 1)$ -th stratum.

372 For $H = 1$ (i.e., the case of SRSWOR), and for the same sample size, we obtained $\pi_0 \simeq 0.1$
 373 regardless of the species — that is why we no longer use the subscript d to denote this probability in
 374 the case of SRSWOR — so $J_1 = \pi_0 \simeq 0.1$. Increasing the number of strata increases the optimal value
 375 of the criterion J_1 . For example, for $H = 2, 3, 4$, we obtain $J_1^* \simeq 0.149$, $J_1^* \simeq 0.160$ and $J_1^* \simeq 0.170$.
 376 Therefore, the greater the number of strata is, the greater the optimal value of the criterion can be

377 increased. Actually, the maximum number of strata that can be formed is the result of constraints
 378 imposed by the situation at hand. At the first sampling stage, by setting the minimum sample size per
 379 stratum to five, an overall sampling fraction of approximately 10%, and the sampling effort allocation
 380 rule (see Section 3.3.2), the maximum number of strata for the stratification to be optimized could not
 381 be very high. In fact, this maximum number is only $H = 5$.

382 Applying the formula in Eq. 9 — here with $L = 22$ and $H = 5$ — we calculated that only
 383 5985 stratifications could be obtained. Using the algorithm described in Sec. 3.3.3, and imposing a
 384 minimum stratum sample size of five PSUs, the number of distinct stratifications to examine dropped
 385 to 4845. The exhaustive search resulted in saving the best stratification found. The optimal value of
 386 the criterion is $J_1^* \simeq 0.177$. The corresponding values of π_d for the $D = 25$ species are given in Table
 387 3. The partitioning results in a transformation of a discrete variable with support $x = 1, 2, \dots, 22$ into
 388 an ordinal variable with five categories (Fig. 7).

389 The reality of maximizing π_d for a given number of strata $H > 1$ is obvious since only the mallard
 390 (rank 25 in Table 3) had a π_d -value almost equal to π_0 ; this was to be expected, as it is a very common
 391 and widespread species in European France. Goosander (*Mergus merganser*, rank 1 in Table 3) did
 392 not obtain a much higher π_d -value than π_0 because its distribution does not coincide much with that
 393 of the other species. Conversely, for a species with a coastal distribution such as the pied avocet
 394 (*Recurvirostra avosetta*, rank 4 in Table 3), for instance (see Fig. 2), the value of π_d reached up to
 395 $2.5 \times \pi_0$. In summary, the coverage probability π_d was greater than $1.5 \times \pi_0$ for 16 species, and π_d
 396 was even greater than $2 \times \pi_0$ for 9 species. In the case of STR-SRSWOR, these results represented the
 397 best possible compromise in terms of maximizing J_1 , conditional on the overall sample size allowed,
 398 the sample size allocation rule, the minimum within-stratum sample size constraint, and the number
 399 of strata used.

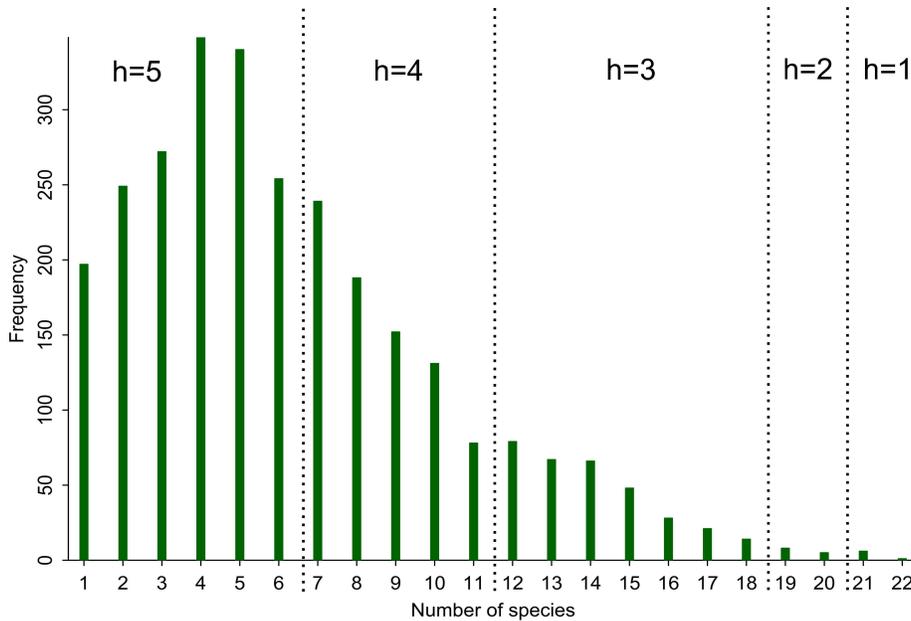


Figure 7: Optimal partitioning into 5 classes for the frequency distribution of the number of species (x) among the $M = 2791$ grid cells ($10 \text{ km} \times 10 \text{ km}$) of the final sampling frame with at least one species occurrence.

Table 3: Values of the sampling coverage probabilities (π_d) for the $D = 25$ species covered by the spatial sampling in European France, sorted in ascending order of the number of occurrences (M_d), restricted to the final PSU sampling frame. The previous rank refers to Table 1.

Rank	Previous rank	Species	Common name	M_d	π_d
1	2	<i>Mergus merganser</i>	Goosander	134	0.114
2	5	<i>Charadrius alexandrinus</i>	Kentish plover	140	0.187
3	1	<i>Tringa totanus</i>	Common redshank	157	0.237
4	3	<i>Recurvirostra avosetta</i>	Pied avocet	164	0.250
5	4	<i>Anser anser</i>	Greylag goose	170	0.232
6	6	<i>Podiceps nigricollis</i>	Black-necked grebe	191	0.245
7	7	<i>Anas crecca</i>	Common teal	221	0.223
8	9	<i>Spatula querquedula</i>	Garganey	247	0.238
9	8	<i>Netta rufina</i>	Red-crested pochard	249	0.193
10	10	<i>Alopochen aegyptiaca</i>	Egyptian goose	262	0.186
11	11	<i>Spatula clypeata</i>	Northern shoveler	330	0.221
12	12	<i>Himantopus himantopus</i>	Black-winged stilt	331	0.204
13	13	<i>Mareca strepera</i>	Gadwall	376	0.218
14	14	<i>Branta canadensis</i>	Canada goose	395	0.152
15	16	<i>Tadorna tadorna</i>	Common shelduck	405	0.180
16	17	<i>Actitis hypoleucos</i>	Common sandpiper	441	0.125
17	15	<i>Aythya fuligula</i>	Tufted duck	485	0.186
18	18	<i>Aythya ferina</i>	Common pochard	599	0.175
19	19	<i>Vanellus vanellus</i>	Northern lapwing	1 137	0.143
20	20	<i>Charadrius dubius</i>	Little ringed plover	1 349	0.132
21	21	<i>Cygnus olor</i>	Mute swan	1 432	0.135
22	22	<i>Podiceps cristatus</i>	Great crested grebe	1 837	0.118
23	23	<i>Tachybaptus ruficollis</i>	Little grebe	1 905	0.118
24	24	<i>Fulica atra</i>	Eurasian coot	2 058	0.115
25	25	<i>Anas platyrhynchos</i>	Mallard	2 745	0.101

400 It is visually verified that the cartographic representation of the size variable (Fig. 8a) is simplified
 401 by the stratification (Fig. 8b).

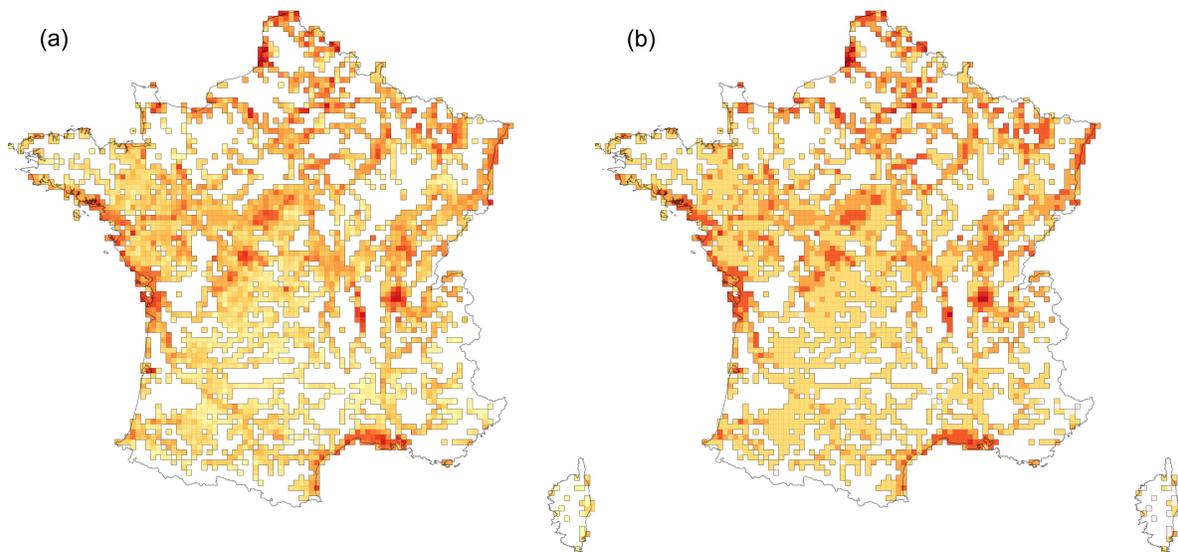


Figure 8: (a) Number of species per PSU. (b) Stratification into six strata. The red color represents the richest PSUs in terms of the number of species present.

402 At the first sampling stage, the allocation of sampling effort between PSU strata is given in Table
 403 4. The stratum containing the richest PSUs in terms of the number of species present (stratum $h = 1$)
 404 is censused (take-all stratum), whereas the poorest PSUs, belonging to the largest stratum ($h = 5$),

405 are sampled with a sampling fraction of less than 6%. Thus, without considering the stratum of PSUs
 406 with undetermined status ($h = 6$), this stratification and allocation overrepresents, as expected, the
 407 most species-rich PSUs in the first stage sample.

Table 4: Allocation of sampling effort between PSU strata. The theoretical sample sizes m_h are deterministically rounded to $[m_h]$ while keeping their sum equal to m . The sampling fraction in the stratum of PSUs with undetermined status is set to the overall value of approximately 10%.

h	M_h	m_h	$[m_h]$	f (%)
1	7	7.00	7	100.0
2	13	8.62	9	69.2
3	323	69.93	70	21.7
4	788	101.43	101	12.8
5	1 660	93.02	93	5.6
6	48	5.00	5	10.4
Total	2 839	285	285	

408 4.3. Sampling under a logistical constraint

409 The logistical organization of the LIMAT scheme ("who surveys where") was carried out according
 410 to the partitioning of the European territory of France into $K = 96$ departments (see Table 2 and Fig.
 411 6).

412 Let N be the total number of SSUs in the second-stage sampling frame. Let s be a sample of
 413 size n drawn from the N SSUs. The number of SSUs in s that belong to a given department k (with
 414 $k = 1, \dots, K$) is denoted by n_k :

$$n_k = \sum_{j=1}^N I_{2j} Z_{2jk} \quad (10)$$

415 with membership indicators I_{2j} ($j = 1, \dots, N$):

$$I_{2j} = \begin{cases} 1 & \text{if SSU } j \text{ belongs to the second-stage sample} \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

416 and Z_{2jk} ($j = 1, \dots, N$, $k = 1, \dots, K$):

$$Z_{2jk} = \begin{cases} 1 & \text{if SSU } j \text{ belongs to department } k \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

417 A maximum number of SSUs (denoted as c) was set to ensure that each department can, logistically,
 418 survey the allocated portion of the sample s . The imposed constraint is therefore written as $n_k \leq c$
 419 (for $k = 1, \dots, K$).

420 The LIMAT scheme was designed to involve OFB agents and volunteer ornithologists from NGOs
 421 under the supervision of the LPO. Under these conditions of workload distribution among observer
 422 teams, we decided to survey a maximum of $c = 80$ SSUs per department.

423 We also decided not to include the logistical constraint at the design stage but to consider it at
 424 the sample selection stage. Indeed, to control the sample size in each department at the design stage
 425 would require departments to define strata of PSUs. However, this posed at least two difficulties:
 426 (i) many PSUs straddle two (or more) departments, which would lead to them being more or less
 427 arbitrarily assigned to a single department (e.g., the one for which the surface area of the intersection
 428 is maximum) and (ii) as we already have another system of PSU strata based on species richness, the
 429 Cartesian product with the departments would have led to a multitude of strata, some of which would
 430 have been too small to be maintained.

431 To account for the logistical constraint at the sample selection stage, the sampling design was
 432 applied many times to obtain a sample that minimized constraint overshoot. Constraint overshoot
 433 was quantified by computing a sample *score*, defined as:

$$f(s) = \sum_{k=1}^K \delta_k \quad (13)$$

434 with

$$\delta_k = \begin{cases} n_k - c & \text{if } n_k > c \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

435 The minimum (min), maximum (max) and average (\bar{n}) numbers of SSUs sampled per department were
 436 approximated by applying the sampling design many times with $m = 285$ and $n = 285 \times 15 = 4275$
 437 (Table 5). The value of \bar{n} varied greatly between departments for three reasons: (i) their surface area
 438 varied (Table 2), (ii) their boundaries were independent of the spatial sampling frame (Fig. 6), and
 439 consequently, (iii) the number of SSUs in departments varied greatly. For example, the lowest value
 440 ($\bar{n} = 1.57$) was obtained for department $k = 75$, which corresponds to Paris, the capital city of France
 441 (see Table 2). As might be expected, some departments had a significantly higher weight than others
 442 in the overall sample, such as Ain ($k = 01$), Aisne ($k = 02$), Somme ($k = 80$), and Vendée ($k = 85$),
 443 because they have many favorable habitats and therefore contain many SSUs and are species-rich.

Table 5: Statistical summaries for 10^6 replications of the sampling design with $m = 285$ and $n = 4275$, 15 SSUs per PSU from the first-stage sample, and $c = 80$. Minimum (min), maximum (max) and average (\bar{n}) numbers of SSUs per department (k). Details are provided in the text.

k	min	max	\bar{n}	k	min	max	\bar{n}	k	min	max	\bar{n}
01	30	274	117.36	32	0	176	38.33	64	0	177	36.60
02	15	263	96.80	33	0	262	78.49	65	0	112	15.49
03	0	258	75.72	34	0	204	51.96	66	0	135	22.84
04	0	162	33.17	35	0	242	66.50	67	0	225	61.30
05	0	125	19.94	36	0	268	83.99	68	0	191	47.51
06	0	103	14.37	37	0	246	75.59	69	0	84	13.42
07	0	125	22.10	38	0	201	50.02	70	0	196	47.34
08	0	181	38.46	39	0	188	49.42	71	0	280	90.65
09	0	118	14.03	40	0	187	47.86	72	0	190	44.55
10	0	163	39.57	41	0	297	94.15	73	0	142	24.17
11	0	208	40.02	42	15	165	48.64	74	0	155	30.22
12	0	165	27.93	43	0	119	18.62	75	0	25	1.57
13	0	234	86.21	44	0	290	103.55	76	0	139	29.60
14	0	175	30.90	45	0	271	82.05	77	0	187	51.06
15	0	138	25.96	46	0	99	12.73	78	0	138	29.15
16	0	144	29.19	47	0	173	36.05	79	0	183	44.20
17	0	281	97.81	48	0	99	13.24	80	30	265	111.10
18	0	226	63.25	49	0	279	92.32	81	0	169	34.12
19	0	165	34.21	50	0	150	35.00	82	0	146	25.98
2A	0	90	10.71	51	0	213	61.03	83	0	105	12.82
2B	0	105	18.85	52	0	108	15.77	84	0	99	14.76
21	0	197	52.75	53	0	197	46.75	85	0	308	108.82
22	0	144	24.61	54	0	196	48.14	86	0	227	68.54
23	0	187	40.04	55	0	239	59.21	87	0	193	47.27
24	0	198	48.27	56	0	236	73.50	88	0	132	23.03
25	0	201	47.77	57	0	266	78.13	89	0	175	45.48
26	0	155	28.14	58	0	187	49.19	90	0	82	12.97
27	0	160	47.65	59	0	253	93.21	91	0	112	17.48
28	0	145	24.74	60	0	163	44.06	92	0	38	3.58
29	0	180	41.33	61	0	104	16.01	93	0	36	3.52
30	0	179	45.59	62	10	212	73.22	94	0	40	4.49
31	0	166	33.93	63	0	163	33.09	95	0	80	10.20

444 Of the 10^6 samples generated by the sampling design, the best one — in terms of meeting the
 445 logistical constraint — yielded the results in Table 6. In practice, it was expected that only 2850 SSUs

446 would be surveyed (the additional $285 \times 5 = 1425$ SSUs are those of the reserve due to possible unit
 447 nonresponse), so the constraint was largely met.

448 For some departments, it was not possible to obtain $\delta = 0$ because of their important weight in the
 449 overall sample, for example, in the case of Ain ($k = 01$), Somme ($k = 80$) or Vendée ($k = 85$) (Table 6).
 450 Conversely, for mountainous — e.g., Ariège ($k = 09$), Corse-du-Sud ($k = 2A$), Haute-Loire ($k = 43$)
 451 — or highly anthropized departments — e.g., Seine-St-Denis ($k = 93$), Val-de-Marne ($k = 94$), both
 452 in the Paris region — the n -value was zero or close to zero.

Table 6: Distribution of the number of SSUs (n) per department (k) for the sample that best satisfies the logistical constraint, that is, minimizes the sample score (sum of the δ -values). Results obtained for 10^6 replications of the sampling design with $m = 285$ and $n = 4275$, 15 SSUs per PSU from the first-stage sample, and $c = 80$. Details are provided in the text.

k	n	δ	k	n	δ	k	n	δ
01	98	18	32	0	0	64	78	0
02	65	0	33	53	0	65	30	0
03	86	6	34	53	0	66	0	0
04	60	0	35	45	0	67	75	0
05	30	0	36	75	0	68	90	10
06	15	0	37	59	0	69	0	0
07	55	0	38	29	0	70	75	0
08	15	0	39	16	0	71	68	0
09	0	0	40	87	7	72	53	0
10	77	0	41	80	0	73	37	0
11	37	0	42	58	0	74	21	0
12	53	0	43	0	0	75	4	0
13	57	0	44	87	7	76	30	0
14	15	0	45	82	2	77	52	0
15	21	0	46	10	0	78	44	0
16	19	0	47	34	0	79	46	0
17	45	0	48	0	0	80	94	14
18	48	0	49	83	3	81	52	0
19	30	0	50	45	0	82	40	0
2A	0	0	51	58	0	83	0	0
2B	45	0	52	15	0	84	20	0
21	59	0	53	66	0	85	100	20
22	30	0	54	35	0	86	59	0
23	45	0	55	75	0	87	57	0
24	37	0	56	45	0	88	40	0
25	46	0	57	75	0	89	46	0
26	21	0	58	37	0	90	44	0
27	45	0	59	67	0	91	23	0
28	60	0	60	41	0	92	8	0
29	90	10	61	15	0	93	0	0
30	43	0	62	53	0	94	0	0
31	75	0	63	45	0	95	44	0

454 In this article, we illustrated the implementation of a stratified two-stage sampling design for
 455 a multispecies survey concerning 25 waterbird species in European France (LIMAT scheme). The
 456 sampling design was based on three spatially nested levels: (i) a level of strata of PSUs defined such
 457 as to maximize the average probability that a sample of PSUs drawn by stratified random sampling
 458 without replacement (STR-SRSWOR) covers the species distributions in European France, (ii) the
 459 units to be sampled at the first stage (PSUs), i.e., 10 km \times 10 km grid cells, and (iii) the units to be
 460 sampled at the second stage (SSUs), defined as 500 m \times 500 m grid cells, to carry out field observations.

461 The knowledge available prior to the start of this survey was the spatial distribution data from
 462 the French Birds Atlas and Faune-France web portal, which documented the recorded presence of
 463 the species in the cells of a 10 km \times 10 km grid. This grid is one of those officially used in Eu-
 464 ropean France for the *National Inventory of Natural Heritage* — in French, *Inventaire National du*
 465 *Patrimoine Naturel* (INPN, see <https://inpn.mnhn.fr/accueil/index?lg=en>). Note that this grid
 466 resolution also exists on the European scale but that the INPN grid does not coincide with that of the
 467 European Environment Agency (EEA reference grid, [https://www.eea.europa.eu/data-and-maps/
 468 data/eea-reference-grids-2](https://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-2)). It was quite natural to choose the grid used by the French Birds
 469 Atlas to define the LIMAT sampling frame of PSUs. Other examples of this practice may be found
 470 in the literature. For instance, in the United Kingdom, where bird counting and monitoring programs
 471 have a long tradition, Donald and Evans (1995) used nested sampling units within larger units defined
 472 by the Atlas of Breeding Birds (Gibbons et al., 1993). The use of the grid referenced by an atlas to
 473 define PSUs for subsequent abundance estimation is a recent development in atlas production (see, for
 474 example, Gibbons et al., 2007; Aizpurua et al., 2015; McCabe et al., 2018). We can therefore see a
 475 link, or even a convergence, between atlas production and monitoring programs.

476 It has been suggested that atlas data “[...] can be used to stratify survey sites and so concentrate
 477 observer effort in areas where most species occurs, and minimize effort outside that range.” (Donald
 478 and Fuller, 1998), but to date, a formalization of this proposal has never been published. Therefore,
 479 the present article fills a methodological gap in the framework of probability sampling applied to
 480 multispecies surveys. The methodological proposal developed in this article (Sec. 3) corresponds to
 481 what Foster et al. (2020, p. 32) call “combined usefulness”, i.e., “the combined utility of each sampling
 482 [unit] to each component of the multivariate observation”, the multivariate nature of the survey here
 483 being due to the multispecies coordination of the sampling effort. This approach can be a source of
 484 inspiration for ecologists or wildlife biologists in charge of multispecies surveys.

485 For bird surveys, the stratified two-stage sampling design that we propose here is unprecedented at
 486 the national level in European France and, to our knowledge, it is also unprecedented at the European
 487 level. In North America, examples of stratified multistage sampling designs exist for bird monitoring
 488 (e.g., Johnson et al., 2009; Pavlacky et al., 2017; Van Wilgenburg et al., 2020). An example of a
 489 recent program is the *Integrated Monitoring in Bird Conservation Regions* (IMBCR) (Pavlacky et al.,
 490 2017). The goal of the IMBCR is “[...] to provide a statistical foundation for reliable knowledge
 491 about bird populations with the ability to address management and conservation objectives at multiple
 492 spatial scales” (Pavlacky et al., 2017). Since 2015, this monitoring program has covered the states of
 493 Colorado, Montana, Wyoming, and parts of Arizona, Idaho, Kansas, Nebraska, New Mexico, North
 494 Dakota, Oklahoma, South Dakota, Texas and Utah. At the first level of stratification, the strata result
 495 from the Cartesian product of the *bird conservation regions* (BCRs) and the states, defining what
 496 Pavlacky et al. (2017) called *BCR-by-state regions*. At the second level, each BCR-by-state region is
 497 stratified according to stakeholder areas of interest, local needs and conservation goals. The PSUs are
 498 defined as 1 km \times 1 km grid cells. Within each PSU, 16 SSUs are defined by points 250 m apart and
 499 125 m from the PSU boundary (see Pavlacky et al., 2017, Fig .2). Another recent example is provided
 500 by Van Wilgenburg et al. (2020) in the boreal region of Canada with the *Boreal Optimal Sampling*
 501 *Strategy* (BOSS). The BOSS was designed to monitor boreal birds for changes in population size and
 502 distribution using a stratified multistage sampling approach that incorporates habitat stratification,
 503 cost constraints, and optimization. In the BOSS, the PSUs are defined as hexagonal grid cells 5 km in
 504 diameter (i.e., a PSU is a hexagon circumscribed by a circle 5 km in diameter), while the SSUs were
 505 defined by a grid of points 300 m apart (see Van Wilgenburg et al., 2020 for more details).

506 It is possible that there is publication bias and that sampling designs (in the statistical sense of
 507 *probability sampling plans*) are rarely described in the ecological literature or are not the subject of
 508 methodological articles. Regardless of whether they are rarely used or rarely published, we found
 509 few reports of large-scale sampling designs for multispecies surveys. Note that the situation is quite
 510 different for environmental monitoring and natural resource surveys (see Gregoire and Valentine, 2008),
 511 especially for forest inventories where complex sampling designs are not uncommon (e.g., Fattorini,

512 2015, and references cited therein; Vallée et al., 2015) and for which very good books are available
513 (e.g., Schreuder et al., 1993; Mandallaz, 2008).

514 5.1. Using stratified two-stage sampling

515 One of the most basic sampling designs is *simple random sampling without replacement* (SRSWOR).
516 The qualifier *simple* does not mean that it is elementary but that the selection probabilities are
517 equal (Tillé, 2020, p. 27). With SRSWOR, the ultimate units are irregularly scattered over the
518 sampled spatial domain. Therefore, SRSWOR is not practical for large-scale surveys due to the cost
519 of transporting observers from one unit to another. One solution is to aggregate the sampled ultimate
520 units into groups that are distributed across the spatial domain of interest. This is usually done by
521 two-stage sampling. Oddly enough — and in agreement with McDonald (2012) — two-stage sampling
522 is still relatively rare in ecology even though Conroy and Smith (1994) consider that “*surveys of wildlife*
523 *populations [...] are inherently multi-stage in design*”. Note that in ecological and wildlife statistics,
524 other terms are sometimes used to refer to multistage sampling, namely, *multilevel sampling* (e.g.,
525 Greenwood and Robinson, 2006b) or *hierarchical nested sampling* (e.g., Skalski, 2012), terms that
526 should be reserved for statistical modeling and variance component estimation.

527 It is widely recognized that stratified multistage sampling is the most common type of design in
528 large-scale surveys (e.g., Sampford, 1962, Preface; Stuart, 1984, p. 78; Singh and Chaudhary, 1986, p.
529 233; Lohr, 2010, p. 282). This type of design is important enough to have its own entry in the famous
530 *Encyclopedia of Statistical Sciences* (see Jain and Hausman, 2006). Moreover, it is one of the most
531 common designs in federal survey programs in the USA (Nusser et al., 1998) and has been used, for
532 instance, for the *National Resources Inventory* (NRI) (Nusser and Goebel, 1997; Nusser et al., 1998).

533 The first reason for using a stratified two-stage sampling design is that it combines the advantages
534 of both stratification and multistage sampling in terms of efficiency and costs. A second reason,
535 specific to stratified sampling, is that it allows estimates to be computed directly at the level of
536 strata (e.g., ecoregions, conservation regions, etc.) without the complications of estimating parameters
537 for unplanned subdomains (e.g., Rao, 2003). A third reason, specific to two-stage sampling, is the
538 possibility of defining the frame of SSUs only for the PSUs selected at the first stage (e.g., Johnson
539 et al., 2009), which can represent a considerable economy in terms of preparatory work.

540 For the LIMAT scheme, the choice of a two-stage stratified sampling design was driven primarily by
541 the need to reduce transportation costs while collecting observations for a maximum number of species
542 — through the sampling allocation rule and the strata optimization — and avoiding major sources of
543 bias in the parameter estimation. In the LIMAT scheme, stratification of the PSUs is central to the
544 multispecies nature of the survey.

545 5.1.1. Optimizing the stratification

546 Beyond the classical applications of sampling designs, there are applications that must be invented
547 to meet the needs of a multispecies survey, depending on the objectives, relevant presurvey data and
548 logistical constraints.

549 In this article, we considered a situation where only the presence of the species in the PSUs was
550 known, but no data were available to anticipate the precision of the estimators. Consequently, the
551 stratification optimization has little to do with what is found in the survey sampling literature, which is
552 essentially concerned with maximizing the precision of the estimators — i.e., minimizing their sampling
553 variance — according to the frequency distribution of the variable of interest. In the case of linear
554 estimators such as the total or mean estimator, maximizing the precision is achieved by minimizing
555 the within-stratum variability for a given number of strata. This defines the objective function to be
556 optimized. Of course, for other finite population parameter estimators, the objective function may be
557 different (Tillé, 2020, Sec. 4.6), for instance, in the case of a quadratic parameter such as the finite
558 population variance. Furthermore, the optimization method to be used depends on both the available
559 knowledge and the number of variables of interest.

560 Here, we propose to optimize a criterion defined as the average sampling coverage probability of
561 species distributions. It is the use of this criterion and the number of species occurring in the PSUs
562 as a size variable that makes the proposed sampling design a multispecies-oriented approach. Other
563 criteria — based, for example, on the variance/covariance or on a diversity measure — are conceivable,
564 but their practical usefulness remains to be evaluated.

565 In the case of the sampling design we used at the first sampling stage (STR-SRSWOR), the criterion
566 to be optimized can be calculated exactly. As mentioned in Section 3.1, it is always possible to use a
567 Monte Carlo approach to estimate the criterion and thus be able to compare different sampling designs
568 used at the first sampling stage, even the most complex ones.

569 The usual methods for optimizing a stratification (see, for instance, Tillé, 2020, Sec. 4.10.2, and
570 references cited therein) are not relevant in the present article since the goal is to maximize the average
571 sampling coverage probability of species distributions, not the between-strata variability. Due to the
572 uniqueness of the size variable considered in this article (univariate case) and the very moderate number
573 of distinct stratifications to be examined for $L = 22$ and $H = 5$, there is no difficulty in finding an
574 optimal stratification by complete enumeration (see Aubry, 2023b).

575 In general, the problem of stratification optimization requires some familiarity with algorithms,
576 whereas most ecologists — even quantitative ones — usually combine already programmed computer
577 procedures (they rarely start completely from scratch). It is therefore necessary to extend the available
578 toolbox. For ease of computer implementation, the necessary basic algorithms can be found in Section
579 3.3.3 of this article and in Aubry (2023b) for the full search algorithm we relied on.

580 Note that it is not impossible that the proposed stratification — in addition to optimizing the
581 coverage probability criterion — could increase the precision of estimation. Indeed, we overrepresented
582 the most species-rich PSUs in the sample, so if there is a nonnegligible correlation between the number
583 of species and the abundance of individuals, at least for some species (as it is reasonable to assume),
584 then we would have increased the precision of the estimation as well. In particular, the use of a take-all
585 stratum for the richest PSUs may be advantageous in that the sample units in that stratum do not
586 contribute to the sampling variance. As far as the LIMAT scheme is concerned, this remains to be
587 verified once the data are compiled, and this will be discussed in a future article.

588 5.1.2. *Optimizing the sample size allocation*

589 In this article, the allocation rule was specified first (Section 3.3.2), and then the stratification was
590 optimized (Section 3.3.3). Conversely, if the strata are given first, then the goal is to optimize the
591 allocation of sampling effort according to a specific rule.

592 The classical literature on survey sampling mostly deals with univariate single objective situations
593 (e.g., Särndal et al., 1992, Sec. 3.7.3; Hankin et al., 2019; Tillé, 2020, pp. 71-75), whereas multispecies
594 surveys deal with multivariate and usually multiobjective situations. Again, the appropriate approach
595 for optimizing the sampling effort allocation depends on the presurvey data available and the objectives
596 of the survey. If available, species richness must be included among the design components considered,
597 and the allocation of sampling effort should be consistent with it. When both species richness and
598 within-stratum variances (or only proxies for these variances, e.g., Van Wilgenburg et al., 2020) are
599 available, they can be used in combination to achieve an appropriate sample size allocation. A different
600 perspective may be to consider the spatial distribution of observers and to allocate the overall sampling
601 effort according to the observer density within strata, as in the *British Breeding Bird Survey* (Gregory
602 and Baillie, 1994; Gregory, 2000; Gregory et al., 2004).

603 Multispecies surveys are not only multivariate and generally multiobjective but are also often
604 multiscale (e.g., Pavlacky et al., 2017; Van Wilgenburg et al., 2020), a feature that can be supported
605 by nesting strata and/or sampling units within each other. When more than one sampling stage is
606 involved in the design, one can also address the allocation of the sampling effort across the different
607 stages. At the present state of our knowledge and skills, we can make four observations: (i) the question
608 of optimizing the sampling effort allocation in a multivariate and/or, multiobjective and/or multiscale
609 situation faces contradictions that can undermine all efforts, i.e., the compromise finally reached is
610 often far from the optimum for the different objectives and design components considered; (ii) the
611 use of sophisticated optimization methods is often not required in the area of probability sampling, as
612 simple approaches generally do not yield distinguishable results from a practical point of view; (iii) the
613 reality of the optimization should be questioned if the data used for this purpose are not themselves
614 sufficiently reliable or recent, and if the phenomenon under study is not stable over time or space
615 (which is precisely what a monitoring program seeks to document); and (iv) allocation optimization
616 leads to a design in which the ultimate sampling units are sampled with unequal overall inclusion
617 probabilities — as is the case in this paper — which complicates the use of the collected data for
618 analytical purposes (e.g., correlative analyses and modeling), since then, one must necessarily account
619 for the sampling weights, which are different (the sampling weights are the inverses of the inclusion
620 probabilities).

621 5.2. *Investigating other possible designs*

622 The stratified two-stage sampling design we are concerned with in this article is a particular instance
623 of a three-stage sampling design with SRSWOR at each stage, where the units at the first stage are
624 sampled at 100% (thus forming strata) (Sukhatme et al., 1984, p. 321; Mandallaz, 2008, p. 34;
625 Gupta and Kabe, 2011, p. 157). This design is referred to as STSI-SI in the taxonomy proposed by
626 Domburg et al. (1997, Table 1) (SI stands for SRSWOR and ST stands for *stratified*, as in Särndal

627 [et al., 1992](#)). According to this taxonomy, other combinations of stratified sampling and two-stage
628 sampling are possible; for instance, considering simple random sampling, stratification may occur at
629 the second stage of sampling, leading to what [Domburg et al. \(1997, Table 1\)](#) abbreviated as the SI-
630 STSI design. Stratification at both stages may also be considered. It is also possible to use more than
631 two sampling stages, but the explicit formulation of the estimators quickly becomes very cumbersome,
632 and an implicit formulation by recurrence is then recommended (see [Aubry, 2021](#)). Furthermore, it
633 is important to understand that each additional stage of sampling is associated with an additional
634 component of sampling variance. Therefore, among the multistage designs, two-stage sampling is both
635 the simplest and the most statistically efficient. It is a reasonable trade-off between precision and
636 travel costs, although not necessarily the most logistically efficient.

637 Given the situation considered in this article — and corresponding to the motivating example
638 — where we have atlas data defined at the PSU level, the key question is whether the species-rich
639 PSUs can be overrepresented in the sample in a way other than the one proposed in this article. One
640 possibility that was explored in the case of the LIMAT scheme, but not finally adopted, was to select
641 the PSUs with probabilities proportional to the size variable (here, the number of species present in the
642 PSUs). This type of sampling design called PPSWOR (an abbreviation for *probability-proportional-*
643 *to-size sampling without replacement*) requires algorithms that can be complex (see [Brewer and Hanif,](#)
644 [1983; Tillé, 2006](#)). This is not a problem as long as their computer implementation is correct (for an
645 example, see [Aubry, 2023a](#)). The main problem with PPSWOR is the stability of the sampling variance
646 estimator since the inclusion probabilities can vary greatly. In the present state of our knowledge and
647 skills, we think that it is wiser to use the stratification and allocation proposed in this article, but the
648 stability of the sampling variance estimators could be studied on fictitious examples, and it would be
649 worthwhile to conduct such a study in depth.

650 While sampling designs (stratified, with unequal probability proportional to size, multistage, mul-
651 tiphase, balanced, etc.) can be combined in a variety of ways, it is important to keep in mind that the
652 choices that are made must be related to the knowledge available and the objectives being pursued
653 and that not all choices are equally relevant.

654 5.3. Dealing with a large-scale multispecies survey

655 As noted by [Yates \(1960, p. 5\)](#) “*The prime requirement of any large-scale sample survey is [...]*
656 *that the organization of the survey should be carried out by a person who has adequate knowledge and*
657 *experience of sampling methods and their application.*”. Indeed, the design of a large-scale survey
658 requires good knowledge of the methods and techniques available. Such skills seem to be relatively
659 rare among (quantitative) ecologists, which is an impediment to the design of such surveys and their
660 widespread use in ecology. Benefiting from the expertise of a survey statistician is usually a rare
661 opportunity in the case of ecological surveys (for counterexamples, see [Fattorini et al., 2004; Baffetta](#)
662 [et al., 2007; Fattorini et al., 2011](#)). Again, note that the involvement of an expert in probability
663 sampling is much more common when financial aspects are at stake, for instance, in forestry. However,
664 statistical expertise is not enough, as a good understanding of thematic issues and logistical constraints
665 is also needed, as [Yates \(1960, p. 6\)](#) pointed out in his time. In fact, a large-scale survey is more an
666 exercise in *statistical engineering* rather than in pure sampling theory ([Mahalanobis, 1944, 1946](#)).
667 Overall, designing a probability sampling plan for surveying multiple species at the same time is
668 a somewhat complicated statistical engineering exercise. For such an exercise, biometricians who
669 specialize in sampling issues can provide valuable assistance.

670 As recently illustrated by [Van Wilgenburg et al. \(2020\)](#) and by the motivating example of the
671 present article, the multiple and often competing constraints force one to make choices to reach the
672 best possible compromise. As far as the LIMAT scheme is concerned, there is room for improvement,
673 especially regarding the second stage sampling frame — the construction of which is not detailed as
674 it is beyond the scope of this article — which does not sufficiently cover the habitats of the various
675 species of interest. Moreover, the logistical constraint is accounted for at the time of sample selection,
676 not at the design stage. If this additional selection process is not accounted for at the estimation stage,
677 this implicitly means that it is considered unrelated to the variables of interest and can therefore be
678 ignored without causing a biased estimation (see [Aubry et al., 2020, Sec. 5.1](#)). This aspect needs to
679 be checked in a timely manner and with an appropriate assessment.

680 A multispecies survey such as the LIMAT scheme faces two main challenges: (i) mapping potential
681 habitats at the scale of the SSUs for each species or group of species and (ii) considering different
682 species simultaneously for spatial and temporal sampling (depending on their phenology). The first
683 point requires a massive amount of work, which is only touched upon in the LIMAT scheme — at least in
684 its current state. It is expected that the mapping of potential habitats will make considerable progress
685 thanks to remote sensing, which is the only way to cover large-scale areas (e.g., [Kerr and Ostrovsky,](#)

686 2003; Fuller et al., 2005). This task requires close collaboration between biologists specialized in the
687 study of the species of interest and geomatics/remote sensing engineers to bridge the gap between
688 ecological expertise and the processing of available environmental data layers (McDermid et al., 2005).
689 Furthermore, the development of automated habitat mapping tools based on deep learning would
690 significantly reduce the financial and human costs of wetland habitat classification, with the advantage
691 that it could be updated annually (e.g., Rezaee et al., 2018; Hosseiny et al., 2021; Jamali et al., 2022a,b).
692 The second point is a thorny one, and the answer given in this article and illustrated by the case of
693 the LIMAT scheme is a solution obtained after eliminating several avenues explored in preliminary
694 studies. The issue of temporal sampling is beyond the scope of this article, which is devoted only to
695 spatial sampling, but it is essential in practice.

696 6. Conclusions and perspectives

697 Given its complexity, the topic of designing a probability sampling plan for multispecies surveys is
698 still in its infancy. The overriding goal is to collect observations for as many species as possible in a
699 limited number of visits. To advance the topic, in this article, we have formalized the proposal that
700 spatial sampling units should be selected to concentrate sampling effort where most species occur and
701 minimize effort elsewhere. To achieve this goal, we proposed an optimal stratification method adapted
702 to the overrepresentation of the most species-rich sampling units in a sample. This method is based
703 on a complete enumeration of admissible stratifications and is therefore exact.

704 We have implemented and illustrated this method in the context of a real-world stratified two-stage
705 sampling design in which primary sampling units (PSUs) are stratified according to species richness,
706 and an allocation rule favors strata in which the PSUs are the richest. The advantage of presenting the
707 method in the context of a real multispecies survey is to highlight both the complexity of the exercise
708 and the practical feasibility of applying a stratified two-stage sampling design, at least for large-scale
709 waterbird monitoring.

710 The methodology described in this article (Sec. 3) is not specific to the motivating example used
711 to illustrate it. It is obviously not specific to fauna and can be used for multispecies surveys of flora
712 or fungi. This methodology should be useful whenever an abundance-based parameter needs to be
713 estimated, whether in a context where each species is considered independently from the others or
714 in the case of estimating an abundance-based biodiversity index. The aim in both cases is to collect
715 observations for as many species as possible in a limited number of visits while avoiding the biases
716 that can arise when spatial sampling is not defined within a sound statistical framework. The question
717 of the usefulness of this methodology for the studies of community patterns — e.g., the analysis of
718 interspecific association, ordination and classification — remains to be examined.

719 To study and document the proposed approach in more detail, for the situation addressed in this
720 article, which is based on the distribution of species among PSUs, future work could simulate cases
721 that allow varying (i) the number of species; (ii) the shape of the frequency distribution for the species
722 richness per PSU; (iii) the species prevalence; (iv) co-occurrence of species in the PSUs; (v) the number
723 of strata; and so on. The aim of such a simulation study would be to identify the breaking points of
724 the proposed approach and to provide guidelines for the design and implementation of multispecies
725 surveys. This assumes first that the technical issues associated with such a simulation study have been
726 adequately addressed.

727 In our view, a great deal of theoretical work remains to be done to examine what other possible
728 paths might be fruitful or, on the contrary, to document why they cannot be used for multispecies
729 surveys. To accomplish this, we could examine different theoretical situations according to what
730 is assumed to be known. Then, in a second step, we could examine to what extent these theoretical
731 situations are close to reality and what possible adaptations should be made to the selected approaches
732 to apply them effectively. This is a methodological research topic in itself, the complexity of which
733 should not be underestimated, but which should be further explored in the future.

734 7. Acknowledgments

735 We would like to dedicate this article to the memory of Thierry Point, who played a crucial role
736 in the development of the LIMAT scheme, especially in developing the sampling frame of secondary
737 sampling units. Unfortunately, Thierry died prematurely in April 2022.

738 We are grateful to Dr Joaquín Hortal and two other anonymous reviewers, as well as to Ecological
739 Informatics' Editor-in-Chief and Associate Editor, for their insightful comments, which helped us to
740 improve the article. We also thank American Journal Experts (AJE) for the final English language
741 editing.

742 **References**

- 743 Aizpurua, O., Paquet, J.Y., Brotons, L., Titeux, N., 2015. Optimising long-term monitoring projects
744 for species distribution modelling: how atlas data may help. *Ecography* 38, 29–40.
- 745 Albert, C.H., Yoccoz, N.G., Edwards, T.C., Graham, C.H., Zimmermann, N.E., Thuiller, W., 2010.
746 Sampling in ecology and evolution - bridging the gap between theory and practice. *Ecography* 33,
747 1028–1037.
- 748 Amorim, F., Carvalho, S.B., Honrado, J., Rebelo, H., 2014. Designing optimized multi-species moni-
749 toring networks to detect range shifts driven by climate change: a case study with bats in the north
750 of Portugal. *PLoS ONE* 9, e87291.
- 751 Ardilly, P., Tillé, Y., 2006. *Sampling methods: exercises and solutions*. Springer, New York, NY, USA.
- 752 Aubry, P., 2021. On the non-recursive implementation of multistage sampling without replacement.
753 *MethodsX* 8, 101553.
- 754 Aubry, P., 2023a. On the correct implementation of the Hanurav-Vijayan selection procedure for
755 unequal probability sampling without replacement. *Commun. Stat. Simul. Comput.* 52, 1849–1877.
- 756 Aubry, P., 2023b. On univariate optimal partitioning by complete enumeration. *MethodsX* 10, 102154.
- 757 Aubry, P., Francesiaz, C., 2022. On comparing design-based estimation versus model-based prediction
758 to assess the abundance of biological populations. *Ecol. Indic.* 144, 109394.
- 759 Aubry, P., Guillemain, M., Sorrenti, M., 2020. Increasing the trust in hunting bag statistics: why
760 random selection of hunters is so important. *Ecol. Indic.* 117, 106522.
- 761 Baffetta, F., Bacaro, G., Fattorini, L., Rocchini, D., Chiarucci, A., 2007. Multi-stage cluster sampling
762 for estimating average species richness at different spatial grains. *Community Ecol.* 8, 119–127.
- 763 Barnett, V., 2002. *Sample survey. Principles & methods*. Third edition. John Wiley & Sons, Chichester,
764 UK.
- 765 Bellhouse, D.R., 1988. A brief history of random sampling methods, in: Krishnaiah, P.R., Rao, C.R.
766 (Eds.), *Handbook of Statistics 6. Sampling*. Elsevier, Amsterdam, The Netherlands, pp. 1–14.
- 767 Bethlehem, J., 2009. *Applied survey methods: a statistical perspective*. John Wiley & Sons, Hoboken,
768 NJ, USA.
- 769 Boyd, R.J., Powney, G.D., Pescott, O.L., 2023. We need to talk about nonprobability samples. *Trends*
770 *Ecol. Evol.* 38, 521–531.
- 771 Brewer, K.R.W., Hanif, M., 1983. *Sampling with unequal probabilities*. Springer, New York, NY,
772 USA.
- 773 Buckland, S.T., Baillie, S.R., Dick, J.M., Elston, D.A., Magurran, A.E., Scott, E.M., Smith, R.I.,
774 Somerfield, P.J., Studeny, P.J., Watt, A., 2012. How should regional biodiversity be monitored?
775 *Environ. Ecol. Stat.* 19, 601–626.
- 776 Buckland, S.T., Marsden, S.J., Green, R.E., 2008. Estimating bird abundance: making methods work.
777 *Bird Conserv. Int.* 18, S91–S108.
- 778 Buckland, S.T., Studeny, A.C., Magurran, A.E., Illian, J.B., Newson, S.E., 2011. The geometric mean
779 of relative abundance indices: a biodiversity measure with a difference. *Ecosphere* 2, Article 100.
- 780 Carvalho, S.B., Gonçalves, J., Guisan, A., Honrado, J.P., 2016. Systematic site selection for multi-
781 species monitoring networks. *J. Appl. Ecol.* 53, 1305–1316.
- 782 Cochran, W.G., 1977. *Sampling techniques*. Third edition. John Wiley & Sons, New York, NY, USA.
- 783 Conroy, M.J., Smith, D.R., 1994. Designing large-scale surveys of wildlife abundance and diversity
784 using statistical sampling principles. *Trans. N. Amer. Wildl. Nat. Resour. Conf.* 59, 159–169.
- 785 Delany, S., 2010. *Guidance on waterbird monitoring methodology: field protocol for waterbird count-*
786 *ing*. Technical Report. Wetlands International. Wageningen, The Netherlands.

- 787 Dénes, F.V., Silveira, L.F., Beissinger, S.R., 2015. Estimating abundance of unmarked animal popu-
788 lations: accounting for imperfect detection and other sources of zero inflation. *Methods Ecol. Evol.*
789 6, 543–556.
- 790 DeWan, A.A., Zipkin, E.F., 2010. An integrated sampling and analysis approach for improved biodi-
791 versity monitoring. *Environ. Manage.* 45, 1223–1230.
- 792 Domburg, P., De Gruijter, J.J., Van Beek, P., 1997. Designing efficient soil survey schemes with a
793 knowledge-based system using dynamic programming. *Geoderma* 75, 183–201.
- 794 Donald, P.F., Evans, A.D., 1995. Habitat selection and population size of Corn Buntings *Miliaria*
795 *calandra* breeding in Britain in 1993. *Bird Study* 42, 190–204.
- 796 Donald, P.F., Fuller, R.J., 1998. Ornithological atlas data: a review of uses and limitations. *Bird*
797 *Study* 45, 129–145.
- 798 Dunn, A.M., Weston, M.A., 2008. A review of terrestrial bird atlases of the world and their application.
799 *Emu* 108, 42–67.
- 800 European Environment Agency, 2020. State of nature in the EU. Results from reporting under the
801 nature directives 2013-2018. Publications Office of the European Union, Luxembourg.
- 802 Fattorini, L., 2015. Design-based methodological advances to support national forest inventories: a
803 review of recent proposals. *iForest* 8, 6–11.
- 804 Fattorini, L., Ferretti, F., Pisani, C., Sforzi, A., 2011. Two-stage estimation of ungulate abundance in
805 Mediterranean areas using pellet group count. *Environ. Ecol. Stat.* 18, 291–314.
- 806 Fattorini, L., Pisani, C., Sforzi, A., 2004. The estimation of wildlife ungulate abundance using sample
807 area surveys: an application to Maremma Regional Park. *Stat. Methods Appl.* 13, 197–212.
- 808 Foster, S.D., Monk, J., Lawrence, E., Hayes, K.R., Hosack, G.R., Langlois, T., Hooper, G., Przeslawski,
809 R., 2020. Statistical considerations for monitoring and sampling, in: Przeslawski, R., Foster, S.
810 (Eds.), *Field manuals for marine sampling to monitor Australian waters. Version 2.* Geoscience
811 Australia and CSIRO, Canberra, Australia, pp. 26–49.
- 812 Fuller, R.M., Devereux, B.J., Gillings, S., Amable, G.S., Hill, R.A., 2005. Indices of bird-habitat
813 preference from field surveys of birds and remote sensing of land cover: a study of south-eastern
814 England with wider implications for conservation and biodiversity assessment. *Glob. Ecol. Biogeogr.*
815 14, 223–239.
- 816 Gibbons, D.W., Donald, P.F., Bauer, H.G., Fornasari, L., Dawson, I.K., 2007. Mapping avian distri-
817 butions: the evolution of bird atlases. *Bird Study* 54, 324–334.
- 818 Gibbons, D.W., Reid, J.B., Chapman, R.A., 1993. *The new atlas of breeding birds in Britain and*
819 *Ireland, 1988-1991.* Poyser, London, UK.
- 820 Greenwood, J.J.D., Robinson, R.A., 2006a. General census methods, in: Sutherland, W.J. (Ed.),
821 *Ecological census techniques. Second edition.* Cambridge University Press, Cambridge, UK, pp. 87–
822 185.
- 823 Greenwood, J.J.D., Robinson, R.A., 2006b. Principles of sampling, in: Sutherland, W.J. (Ed.), *Eco-*
824 *logical census techniques. Second edition.* Cambridge University Press, Cambridge, UK, pp. 11–86.
- 825 Gregoire, T.G., Valentine, H.T., 2008. *Sampling strategies for natural resources and the environment.*
826 Chapman & Hall/CRC, Boca Raton, FL, USA.
- 827 Gregory, R.D., 2000. Development of breeding bird monitoring in the United Kingdom and adopting
828 its principles elsewhere. *The Ring* 22, 35–44.
- 829 Gregory, R.D., Baillie, S.R., 1994. Evaluation of sampling strategies for 1-km squares for inclusion in
830 the Breeding Bird Survey. BTO Research Report No. 139. British Trust for Ornithology, Thetford,
831 UK.
- 832 Gregory, R.D., Baillie, S.R., Bashford, R.I., 2004. Monitoring breeding birds in the united kingdom,
833 in: Anselin, A. (Ed.), *Bird Numbers 1995, Proceedings of the International Conference and 13 th*
834 *Meeting of the European Bird Census Council,* pp. 101–112.

- 835 Gregory, R.D., Greenwood, J.J.D., 2008. Counting common birds, in: Voříšek, P., Klvaňová, A.,
836 Wotton, S., Gregory, R.D. (Eds.), A best practice guide for wild bird monitoring schemes. First
837 edition. CSO/RSPB, Prague, Czech Republic, pp. 21–86.
- 838 Gupta, A.K., Kabe, D.G., 2011. Theory of sample surveys. World Scientific, Singapore.
- 839 Hankin, D.G., Mohr, M.S., Newman, K.B., 2019. Sampling theory for the ecological and natural
840 resource sciences. Oxford University Press, Oxford, UK.
- 841 Hedayat, A.S., Sinha, B.K., 1991. Design and inference in finite population sampling. John Wiley &
842 Sons, New York, NY, USA.
- 843 Hewitt, J.E., Thrush, S.F., Dayton, P.K., Bonsdorf, E., 2007. The effect of spatial and temporal
844 heterogeneity on the design and analysis of empirical studies of scale-dependent systems. *Am. Nat.*
845 *169*, 398–408.
- 846 Hewitt, J.E., Thrush, S.F., Lundquist, C., 2017. Scale-dependence in ecological systems, in: eLS. John
847 Wiley & Sons, Chichester, UK.
- 848 Hidiroglou, M., Lavallée, P., 2009. Sampling and estimation in business surveys, in: Pfeffermann, D.,
849 Rao, C.R. (Eds.), Handbook of statistics 29A. Sample surveys: design, methods and applications.
850 Elsevier, Oxford, UK, pp. 441–470.
- 851 Hill, M.O., 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* *54*,
852 427–432.
- 853 Hosseiny, B., Mahdianpari, M., Brisco, B., Mohammadimanesh, F., Salehi, B., 2021. WetNet: a spatial-
854 temporal ensemble deep learning model for wetland classification using Sentinel-1 and Sentinel-2.
855 *IEEE Trans. Geosci. Remote Sens.* *60*, 4406014.
- 856 International BirdLife, 2021. European Red List of Birds. Publications Office of the European Union,
857 Luxembourg.
- 858 Issa, N., Muller, Y., 2015. Atlas des oiseaux de France métropolitaine. Nidification et présence hiver-
859 nale. Delachaux et Niestlé, Paris, France.
- 860 IUCN, 2022. The IUCN Red List of Threatened Species. URL: <https://www.iucnredlist.org>.
- 861 Jain, A.K., Hausman, R.E., 2006. Stratified multistage sampling, in: Kotz, S., Balakrishnan, N., Read,
862 C.B., Vidakovic, B., Johnson, N.L. (Eds.), Encyclopedia of statistical sciences. Second edition. John
863 Wiley & Sons, Hoboken, NJ, USA. volume 13, pp. 8333–8337.
- 864 Jamali, A., Mahdianpari, M., Brisco, B., Mao, D., Salehi, B., Mohammadimanesh, F., 2022a. 3DUNet-
865 GSFormer: a deep learning pipeline for complex wetland mapping using generative adversarial net-
866 works and Swin transformer. *Ecol. Inform.* *72*, 101904.
- 867 Jamali, A., Mahdianpari, M., Mohammadimanesh, F., Homayouni, S., 2022b. A deep learning frame-
868 work based on generative adversarial networks and vision transformer for complex wetland classifi-
869 cation using limited training samples. *Int. J. Appl. Earth Obs. Geoinf.* *115*, 103095.
- 870 Johnson, D.H., Gibbs, J.P., Herzog, M., Lor, S., Niemuth, N.D., Ribic, C.A., Seamans, M., Shaffer,
871 T.L., Shriver, W.G., Stehman, S.V., Thompson, W.L., 2009. A sampling design framework for
872 monitoring secretive marshbirds. *Waterbirds* *32*, 203–362.
- 873 Keddy, P.A., Laughlin, D.C., 2022. A framework for community ecology. Species pools, filters and
874 traits. Cambridge University Press, Cambridge, UK.
- 875 Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing.
876 *Trends Ecol. Evol.* *18*, 299–305.
- 877 Klibansky, N., Shertzer, K.W., Kellison, G.T., Bachelier, N.M., 2017. Can subsets of species indicate
878 overall patterns in biodiversity? *Ecosphere* *8*, e01842.
- 879 Lambeck, R.J., 1997. Focal species: a multi-species umbrella for nature conservation. *Conserv. Biol.*
880 *11*, 849–856.

- 881 Lancia, R.A., Nichols, J.D., Pollock, K.H., 1996. Estimating the number of animals in wildlife popu-
882 lations, in: Bookhout, T.A. (Ed.), Research and management techniques for wildlife and habitats.
883 Fifth ed., rev.. The Wildlife Society, Bethesda, MD, USA, pp. 215–253.
- 884 Legendre, P., Troussellier, M., Jarry, V., Fortin, M.J., 1989. Design for simultaneous sampling of
885 ecological variables: from concepts to numerical solutions. *Oikos* 55, 30–42.
- 886 Link, W.A., Nichols, J.D., 1994. On the importance of sampling variance to investigations of temporal
887 variation in animal population size. *Oikos* 69, 539–544.
- 888 Lloyd, H., Cahill, A., Jones, M., Marsden, S., 1998. Estimating bird densities using distance sampling,
889 in: Bibby, C., Jones, M., Marsden, S. (Eds.), Expedition field techniques. Bird surveys. Expedition
890 Advisory Centre of the Royal Geographical Society, London, UK, pp. 35–52.
- 891 Lohr, S.L., 2010. Sampling: design and analysis. Second edition. Brooks/Cole, Boston, MA, USA.
- 892 Magurran, A.E., 2004. Measuring biological diversity. Blackwell Science, Oxford, UK.
- 893 Mahalanobis, P.C., 1944. On large-scale sample surveys. *Philos. Trans. R. Soc. Lond. B Biol. Sci* 231,
894 329–451.
- 895 Mahalanobis, P.C., 1946. Recent experiments in statistical sampling in the Indian Statistical Institute
896 (with discussion). *J. R. Stat. Soc. Ser. A Stat. Soc.* 109, 326–378.
- 897 Mandallaz, D., 2008. Sampling techniques for forest inventories. Chapman & Hall/CRC, Boca Raton,
898 Florida, USA.
- 899 Manley, P.N., Zielinski, W.J., Schlesinger, M.D., Mori, S.R., 2004. Evaluation of a multiple-species
900 approach to monitoring species at the ecoregional scale. *Ecol. Appl.* 14, 296–310.
- 901 Marta, S., Lacasella, F., Romano, A., Ficetola, G.F., 2019. Cost-effective spatial sampling designs for
902 field surveys of species distribution. *Biodivers. Conserv.* 28, 2891–2908.
- 903 McCabe, J.D., Anich, N.M., Brady, R.S., Zuckerberg, B., 2018. Raising the bar for the next generation
904 of biological atlases: using existing data to inform the design and implementation of atlas monitoring.
905 *Ibis* 160, 528–541.
- 906 McDermid, G.J., Franklin, S.E., LeDrew, E.F., 2005. Remote sensing for large-area habitat mapping.
907 *Prog. Phys. Geogr.* 29, 449–474.
- 908 McDonald, T., 2012. Spatial sampling designs for long-term ecological monitoring, in: Gitzen, R.A.,
909 Millsaugh, J.J., Cooper, A.B., Licht, D.S. (Eds.), Design and analysis of long-term ecological
910 monitoring studies. Cambridge University Press, New York, NY, USA, pp. 101–125.
- 911 Morin, P.J., 2011. Community ecology. Second edition. John Wiley & Sons, Chichester, UK.
- 912 Nagy, S., Langendoen, T., 2021. Report on the conservation status of migratory waterbirds in the agree-
913 ment area. Eighth edition. Technical Report. Wetlands International. Wageningen, The Netherlands.
- 914 Nguyen, H.T., 2006. An introduction to random sets. Chapman & Hall/CRC, Boca Raton, FL, USA.
- 915 Nusser, S.M., Breidt, F.J., Fuller, W.A., 1998. Design and estimation for investigating the dynamics
916 of natural resources. *Ecol. Appl.* 8, 234–245.
- 917 Nusser, S.M., Goebel, J.J., 1997. The National Resources Inventory: a long-term multi-resource
918 monitoring programme. *Environ. Ecol. Stat.* 4, 181–204.
- 919 Olsen, A.R., Sedransk, J., Edwards, D., Gotway, C., Liggett, W., Rathbun, S., Reckhow, K.H., Young,
920 L.J., 1999. Statistical issues for monitoring ecological and natural resources in the United States.
921 *Environ. Monit. Assess.* 54, 1–45.
- 922 Patil, G., Taillie, C., 1982. Diversity as a concept and its measurement. *J. Am. Stat. Assoc.* 77,
923 548–561.
- 924 Pavlacky, D.C., Lukacs, P.M., Blakesley, J., Skorkowsky, R., Klute, D.S., Hahn, B., Dreitz, V.J.,
925 George, T.L., Hanni, D.J., 2017. A statistically rigorous sampling design to integrate avian moni-
926 toring and management within Bird Conservation Regions. *PLoS ONE* 12, e0185924.

- 927 Pollock, K.H., Nichols, J.D., Simons, T.R., Farnsworth, G.L., Bailey, L.L., Sauer, J.R., 2002. Large
928 scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics* 13,
929 105–119.
- 930 Rao, J.N.K., 2003. *Small area estimation*. John Wiley & Sons, Hoboken, NJ, USA.
- 931 Rezaee, M., Mahdianpari, M., Zhang, Y., Salehi, B., 2018. Deep convolutional neural network for
932 complex wetland classification using optical remote sensing imagery. *IEEE J. Sel. Top. Appl. Earth*
933 *Obs. Remote Sens.* 11, 3030–3039.
- 934 Royall, R.M., 2006. Finite populations, sampling from, in: Kotz, S., Balakrishnan, N., Read, C.B.,
935 Vidakovic, B., Johnson, N.L. (Eds.), *Encyclopedia of statistical sciences*. Second edition. John Wiley
936 & Sons, Hoboken, NJ, USA, pp. 2329–2334.
- 937 Sampford, M.R., 1962. *An introduction to sampling theory: with applications to agriculture*. Oliver
938 & Boyd, Edinburgh, UK.
- 939 Samuel, M.D., Garton, E.O., 1994. Horvitz-Thompson survey sample methods for estimating large-
940 scale animal abundance. *Trans. N. Amer. Wildl. Nat. Resour. Conf.* 59, 170–179.
- 941 Särndal, C.E., Swensson, B., Wretman, J.H., 1992. *Model assisted survey sampling*. Springer, New
942 York, NY, USA.
- 943 Schneider, D.C., 1994. *Quantitative ecology. Spatial and temporal scaling*. Academic Press, San Diego,
944 CA, USA.
- 945 Schneider, D.C., 2009. *Quantitative ecology. Measurement, models, and scaling*. Second edition. Aca-
946 demic Press, London, UK.
- 947 Schreuder, H.T., Gregoire, T.G., Wood, G.B., 1993. *Sampling methods for multiresource forest inven-*
948 *tory*. John Wiley & Sons, New York, NY, USA.
- 949 Singh, D., Chaudhary, F.S., 1986. *Theory and analysis of sample survey designs*. New Age International
950 Publishers, New Delhi, India.
- 951 Skalski, J.R., 2012. Estimating variance components and related parameters when planning long-term
952 monitoring programs, in: Gitzen, R.A., Millsbaugh, J.J., Cooper, A.B., Licht, D.S. (Eds.), *Design*
953 *and analysis of long-term ecological monitoring studies*. Cambridge University Press, New York, NY,
954 USA, pp. 174–199.
- 955 Smith, A.N.H., Anderson, M.J., Pawley, M.D.M., 2017. Could ecologists be more random? Straight-
956 forward alternatives to haphazard spatial sampling. *Ecography* 40, 1251–1255.
- 957 Steele, J.H., 1991. Can ecological theory cross the land-sea boundary? *J. Theor. Biol.* 153, 425–436.
- 958 Stuart, A., 1984. *The ideas of sampling*. Third edition. Charles Griffin, London, UK.
- 959 Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S., Asok, C., 1984. *Sampling theory of surveys with*
960 *applications*. Third edition. Iowa State University Press, Ames, IA, USA.
- 961 Sutherland, W.J., Freckleton, R.P., Godfray, H.C.J., Beissinger, S.R., Benton, T., Cameron, D.D.,
962 Carmel, Y., Coomes, D.A., Coulson, T., Emmerson, M.C., Hails, R.S., Hays, G.C., Hodgson, D.J.,
963 Hutchings, M.J., Johnson, D., Jones, J.P.G., Keeling, M.J., Kokko, H., Kunin, W.E., Lambin,
964 X., Lewis, O.T., Malhi, Y., Mieszkowska, N., Milner-Gulland, E.J., Norris, K., Phillimore, A.B.,
965 Purves, D.W., Reid, J.M., Reuman, D.C., Thompson, K., Travis, J.M.J., Turnbull, L.A., Wardle,
966 D.A., Wiegand, T., 2013. Identification of 100 fundamental ecological questions. *J. Ecol.* 101, 58–67.
- 967 Tillé, Y., 2006. *Sampling algorithms*. Springer, New York, NY, USA.
- 968 Tillé, Y., 2020. *Sampling and estimation from finite populations*. John Wiley & Sons, Hoboken, NY,
969 USA.
- 970 UICN France, MNHN, LPO, SEOF & ONCFS, 2016. *La Liste rouge des espèces menacées en France -*
971 *Chapitre Oiseaux de France métropolitaine*. UICN France, MNHN, LPO, SEOF & ONCFS, Paris,
972 France.
- 973 Vallée, A.-A., Ferland-Raymond, B., Rivest, L.-P., Tillé, Y., 2015. Incorporating spatial and opera-
974 tional constraints in the sampling designs for forest inventories. *Environmetrics* 26, 557–570.

- 975 Van Turnhout, C.A.M., Willems, F., Plate, C., Van Strien, A., Teunissen, W., Van Dijk, A., Foppen,
976 R., 2008. Monitoring common and scarce breeding birds in the Netherlands: applying a post-
977 hoc stratification and weighting procedure to obtain less biased population trends. *Rev. Catalana*
978 *Ornitol.* 24, 15–29.
- 979 Van Wilgenburg, S.L., Mahon, C.L., Campbell, G., McLeod, L., Campbell, M., Evans, D., Easton,
980 W., Francis, C.M., Haché, S., Machtans, C.S., Mader, C., Pankratz, R.F., Russell, R., Smith,
981 A.C., Thomas, P., Toms, J.D., Tremblay, J.A., 2020. A cost efficient spatially balanced hierarchical
982 sampling design for monitoring boreal birds incorporating access costs and habitat stratification.
983 *Plos ONE* 15, e0234494.
- 984 Waite, S., 2000. *Statistical ecology in practice: a guide to analysing environmental and ecological field*
985 *data*. Pearson Education, Harlow, UK.
- 986 Whittaker, R.J., Willis, K.J., Field, R., 2001. Scale and species richness: towards a general, hierarchical
987 theory of species diversity. *J. Biogeogr.* 28, 453–470.
- 988 Wiens, J.A., 1989. Spatial scaling in ecology. *Funct. Ecol.* 3, 385–397.
- 989 Wiens, J.A., Hayward, G.D., Holthausen, R.S., Wisdom, M.J., 2008. Using surrogate species and
990 groups for conservation planning and management. *BioScience* 58, 241–252.
- 991 Yates, F., 1960. *Sampling methods for censuses and surveys*. Third edition. Charles Griffin, London,
992 UK.
- 993 Young, L., Young, J., 1998. *Statistical ecology. A population perspective*. Springer, New York, NY,
994 USA.