



ISSN 0154 - 2109



Essai d'analyse par Distance Sampling de données SHOC en Auvergne (2018-2019-2020).

Par François Guélin, Sylvain Sainnier et Hugo Samain.

LE GRAND-DUC N°88 (ANNEE 2020)



Résumé : Le Suivi Hivernal des Oiseaux Communs – SHOC – est une technique de suivi qui semble pouvoir se prêter à une exploitation par la méthode des distances ou distance-sampling. Les 4 classes de distance notées par les observateurs permettent d'utiliser le logiciel Distance avec l'analyse par groupes (« clusters ») pour les espèces grégaires, afin d'obtenir des estimations de densités. Malgré l'échantillonnage plutôt faible à notre disposition pour le département du Puy-de-Dôme, les estimations obtenues semblent cohérentes, malgré des biais à courte distance. Mais les carrés SHOC suivis en Auvergne sont encore trop peu nombreux, mal répartis, avec des estimations de distance sujettes à des imprécisions.

Mots-clés : Suivi Hivernal des Oiseaux Communs – SHOC - Échantillonnage par la distance – distance sampling – passereaux communs – hivernants - Puy-de-Dôme – 63 – Auvergne Rhône-Alpes

Introduction

Voici maintenant cinq ans qu'un petit groupe de travail s'est créé en Auvergne pour travailler avec la méthode du Distance Sampling. Nous sommes 6 ou 7 à être capables de faire fonctionner le logiciel Distance avec ses subtilités et près de 45 ornithologues, dans l'Allier ou le Puy-de-Dôme, ont pu, ces dernières années, apprendre sur le terrain à pratiquer le Distance Sampling. Les résultats obtenus dans nos enquêtes de terrain sont très intéressants pour estimer les populations d'oiseaux communs. Ces études ont été menées avec des protocoles conçus spécifiquement pour le DS : échantillonnages systématiques et relevés précis des distances.

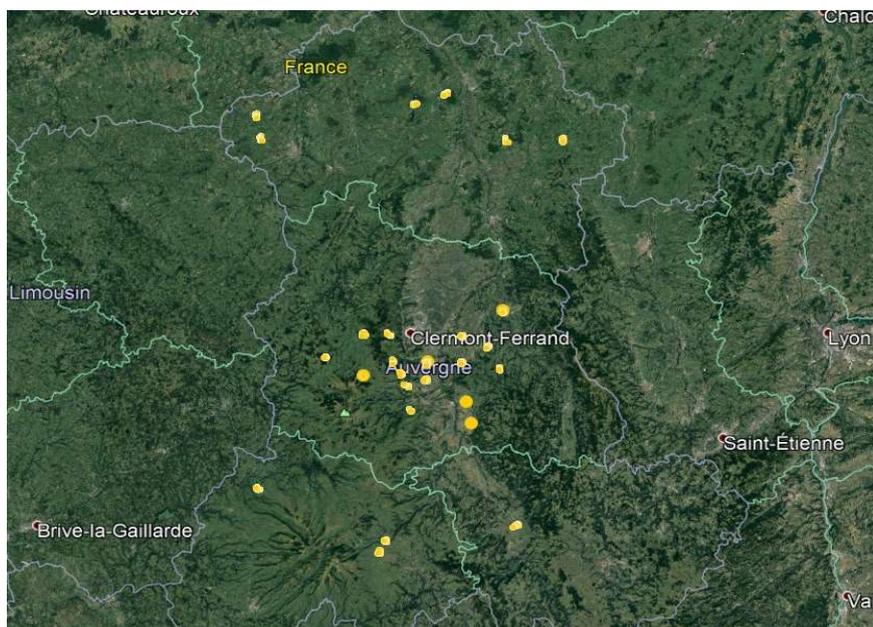
Nous nous sommes demandé si les comptages SHOC, où des valeurs de distance sont spécifiées, pourraient être exploités aussi par la méthode du Distance Sampling. Nous avons donc décidé de regarder de plus près les données SHOC obtenues en Auvergne ces deux dernières années.

Le programme SHOC, organisé par le Muséum National d'Histoire Naturelle, est basé sur des comptages visuels et auditifs d'oiseaux à partir de transects le long desquels l'observateur se déplace. Une paire de jumelles peut être utilisée pour aider à la détermination. Le transect est parcouru à pied, en marchant à environ 4 km/heure (www.vigienature.fr, 2020). Le site de comptage est un carré de 2 km x 2 km tiré au sort dans un rayon de 10 km autour de la commune indiquée par l'observateur. Ceci créé automatiquement un biais, puisqu'en Auvergne, par exemple, la répartition des observateurs est vraiment très hétérogène. Dix transects d'environ 300 mètres (min. 250 et max. 350 mètres, soit 3 kilomètres de transects par carré au total) sont positionnés dans chaque carré, placés bout à bout et traversant les habitats majoritaires du carré. Deux passages sont effectués, un en décembre et un en janvier. Sur le site du Muséum, un petit message est très explicite : « Pourquoi noter les distances ? Ces informations permettent de calculer les probabilités de détection des différentes espèces, indispensables pour que les tendances d'évolution des populations obtenues soient fiables. ». C'est donc bien une exploitation par Distance Sampling qui est prévue.

Matériel & méthode

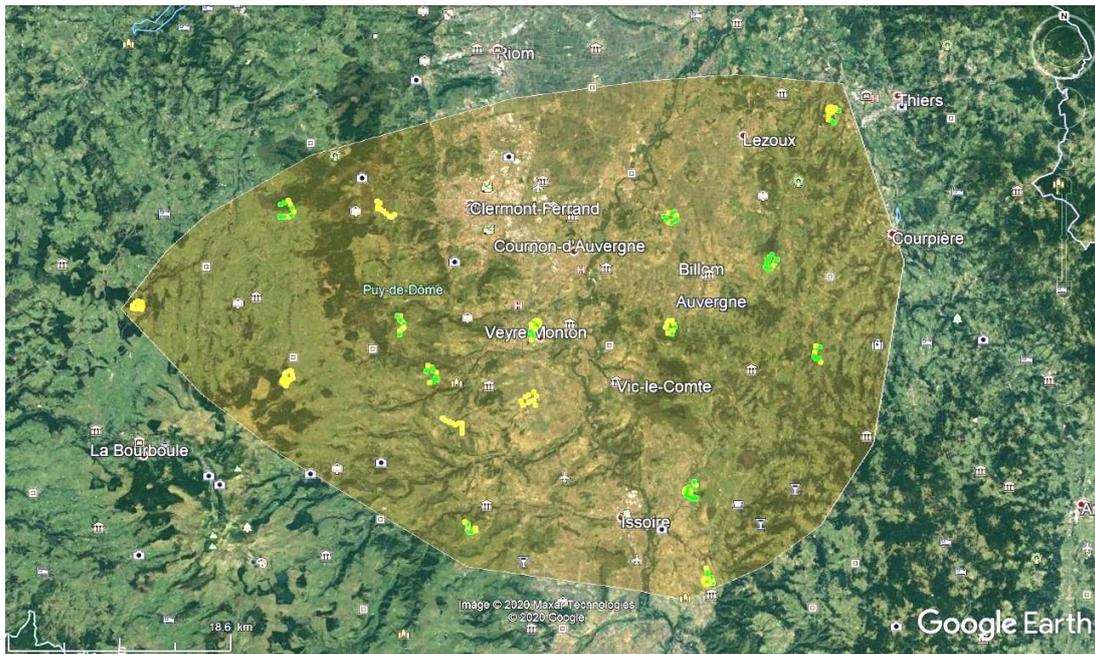
1. Échantillonnage

Nous avons été assez déçus car en prenant les deux derniers hivers 2018-2019 et 2019-2020, seuls 27 carrés échantillons SHOC étaient disponibles sur les quatre départements auvergnats (voir carte du document 1 ci-dessous).



Document 1 : emplacement des 27 carrés SHOC en Auvergne en 2019-2020

Sur ces 27 carrés, nous en avons 6 dans l'Allier, 3 dans le Cantal, 1 seul en Haute-Loire, et 17 dans le Puy-de-Dôme. Cet échantillonnage régional beaucoup trop inégal nous a convaincu qu'il fallait nous restreindre au département du Puy-de-Dôme (17 carrés) pour analyser les données (Document 2).



Document 2 : emplacement des transects sélectionnés pour l'analyse par Distance Sampling dans le département du Puy-de-Dôme. (en vert, hiver 2018-2019, en jaune, hiver 2019-2020) – zone d'extrapolation possible en jaune pâle : 200 000 ha.

Même avec cette limitation, on voit très bien que les carrés échantillonnés sont tous regroupés dans la partie centrale du département où habitent les observateurs. Nous avons considéré que ces 17 carrés permettent de circonscrire une zone d'environ 200 000 hectares, tout en gardant à l'esprit que la Limagne, par exemple, n'est pas couverte par l'enquête (on peut comprendre que peu d'ornithologues veuillent recenser des zones quasi-vides). 17 carrés sur 200 000 ha représentent un échantillonnage vraiment minimal : ils totalisent 174 kilomètres de transects.

2. Méthode

Nous avons choisi d'analyser seulement les espèces avec plus de 100 données individuelles. 19 d'entre elles seulement répondent à ce premier critère (document 3). Un autre critère important pour la méthode du Distance Sampling est de savoir si les individus sont observés un par un, ou bien par groupe ou « cluster » (n individus de la même espèce observés à un temps t et à une distance x). Nous avons donc calculé pour chaque espèce la moyenne d'individus par groupe, pour déterminer quelles seraient les espèces analysées par « cluster » et celles qui le seraient individuellement. Le choix a été fait de considérer comme grégaires, les espèces dont la moyenne d'individus par groupe dépassait 3. Neuf espèces rentrent dans cette catégorie: de la Tourterelle turque (moyenne 3,1 ind/ groupe) à l'Alouette des champs (moyenne 24,8 ind/ groupe).

Par la suite, nous avons analysé les espèces grégaires avec l'option « cluster » du logiciel Distance et éliminé celles qui n'atteignaient pas le nombre minimal de 20 groupes. Le choucas, malgré ses 148 données individuelles est donc éliminé car seuls 10 groupes de cette espèce ont été notés sur les deux hivers, ce qui est très insuffisant. Il ne reste donc que 18 espèces à analyser (document 3).



Document 3 Les espèces fréquentes du SHOC en Centre Puy-de-Dôme

	Nb d'Ind. Avec n>100	Nb de Groupes avec n>20	Nb d'ind. / Nb groupe = Grégaire si n>3	Type d'analyse avec le logiciel Distance
Merle noir <i>Turdus merula</i>	746	371	2,0	Analyse individuelle
Mésange charbonnière <i>Parus major</i>	585	293	2,0	Analyse individuelle
Corneille noire <i>Corvus corone</i>	347	194	1,8	Analyse individuelle
Mésange bleue <i>Cyanistes caeruleus</i>	327	195	1,7	Analyse individuelle
Pie bavarde <i>Pica pica</i>	186	123	1,5	Analyse individuelle
Grive draine <i>Turdus viscivorus</i>	143	97	1,5	Analyse individuelle
Geai des chênes <i>Garrulus glandarius</i>	231	177	1,3	Analyse individuelle
Rougegorge <i>Erithacus rubecula</i>	185	154	1,2	Analyse individuelle
Pic épeiche <i>Dendrocopos major</i>	107	102	1,0	Analyse individuelle
Alouette des champs <i>Alauda arvensis</i>	1015	41	24,8	Analyse par "clusters"
Linotte mélodieuse <i>Carduelis cannabina</i>	453	29	15,6	Analyse par "clusters"
Etourneau sansonnet <i>Sturnus vulgaris</i>	2434	161	15,1	Analyse par "clusters"
Moineau domestique <i>Passer domesticus</i>	1165	121	9,6	Analyse par "clusters"
Pigeon ramier <i>Columba palumbus</i>	746	104	7,2	Analyse par "clusters"
Pinson des arbres <i>Fringilla coelebs</i>	1441	250	5,8	Analyse par "clusters"
Chardonneret élégant <i>Carduelis carduelis</i>	185	33	5,6	Analyse par "clusters"
Mésange à longue queue <i>Aegithalos caudatus</i>	113	25	4,5	Analyse par "clusters"
Tourterelle turque <i>Streptopelia decaocto</i>	236	75	3,1	Analyse par "clusters"
Choucas des tours <i>Corvus monedula</i>	148	10	14,8	Non analysé

Les distances étant notées dans des classes de 0-25, 25-100, 100-200 et plus de 200 m, les valeurs importées dans le logiciel Distance sont fixées par défaut à la médiane de chaque classe. Le logiciel est ensuite paramétré pour travailler en 3 ou 4 classes de distance (selon qu'on conserve ou non la classe > 200 m). La méthode d'analyse consiste à trouver, pour chaque espèce, le meilleur modèle mathématique pour représenter nos données, à partir duquel la population pourra être estimée. Cinq modèles ont été utilisés (Half-normal cosine, Hazard rate cosine et polynomial, Uniform cosine et polynomial). Nous avons réalisé deux tests consécutifs pour chaque espèce : un avec quatre classes de distances, et un second avec trois seulement (en annulant les données obtenues au-delà de 200 m). L'analyse s'est faite soit en mode « classique » individuel (1 oiseau = 1 distance), ou avec l'option « cluster » si le nombre d'individus par groupe est supérieur à 3 (un groupe d'oiseaux = une distance). En mode Cluster le logiciel Distance calcule d'abord la probabilité

de détection des groupes en fonction de la distance (comme il le ferait avec les analyses classiques par individus) puis propose 2 options (nous résumons):

- simplement multiplier l'estimation du nombre de groupes par la moyenne du nombre d'individus par groupe
- ou bien, sachant que la probabilité de détection d'un groupe dépend de la distance mais aussi de la taille de ce groupe, le logiciel propose d'améliorer la qualité de l'estimation en réalisant une régression mathématique entre la probabilité de détection et la taille des groupes. Cela permet - parfois - de réévaluer la moyenne arithmétique de taille des groupes, qui a tendance à être surévaluée, les grands groupes étant plus facilement identifiés à grande distance. Si ce calcul n'est pas statistiquement significatif, il permute automatiquement vers la première méthode. Nous avons donc choisi cette option.

Pour le choix de la meilleure estimation, les critères sont d'abord le plus faible AIC, puis le test du Khi2 (pas toujours possible avec seulement 3 ou 4 classes de distance), et le coefficient de variation de la densité qui doit être le plus bas possible. Les analyses ont toutes été effectuées deux fois par deux personnes différentes.

Résultats

Le tableau ci-dessous (document 4) fait la synthèse des meilleures analyses (AIC, Khi2 p et D CV) retenues pour chaque espèce. Les estimations de densité (min/max à 95 % et estimation principale) sont indiquées, pour 1 kilomètre carré.

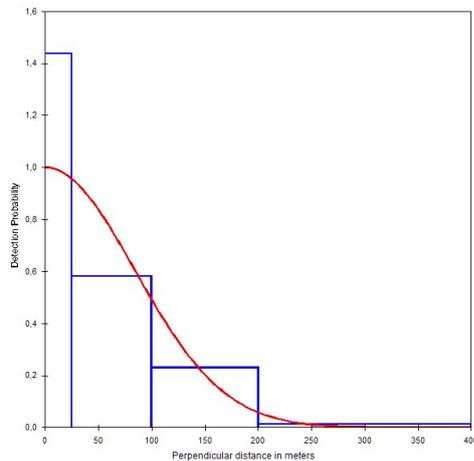
Document 4 : résultats des analyses

	Ind. ou groupes	N	Classes de distance	Khi2 p	D CV %	Dmin Ind/km ²	Densité Ind/km ²	D max Ind/km ²
Alouette des champs (<i>Alauda arvensis</i>)	Groupes	41	4	0,08	52	5,4	14,5	39,1
Chardonneret élégant (<i>Carduelis carduelis</i>)	Groupes	33	4	0,31	49	6,5	16,4	41,6
Corneille noire (<i>Corvus corone</i>)	Ind.	347	4	0,98	18	3,0	4,4	6,4
Etourneau sansonnet (<i>Sturnus vulgaris</i>)	Groupes	161	4	-	28	50,5	86,4	147,9
Geai des chênes (<i>Garrulus glandarius</i>)	Ind.	216	3	-	24	4,7	7,6	12,2
Grive draine (<i>Turdus viscivorus</i>)	Ind.	143	4	0,17	24	1,5	2,5	4,1
Linotte mélodieuse (<i>Carduelis cannabina</i>)	Groupes	29	4	0,39	45	10,9	26,0	62,7
Merle noir (<i>Turdus merula</i>)	Ind.	746	4	0,31	25	35,7	59,5	99,2
Mésange à longue queue (<i>Aegithalos caudatus</i>)	Groupes	25	4	0,53	39	4,6	9,7	20,6
Mésange bleue (<i>Cyanistes caeruleus</i>)	Ind.	327	4	0,87	21	27,6	41,6	62,7
Mésange charbonnière (<i>Parus major</i>)	Ind.	584	3	-	26	53,7	89,1	147,8
Moineau domestique (<i>Passer domesticus</i>)	Groupes	121	3	-	29	70,2	124,4	220,5
Pic épeiche (<i>Dendrocopos major</i>)	Ind.	107	4	0,21	41	2,4	5,2	11,6
Pie bavarde (<i>Pica pica</i>)	Ind.	157	3	0,67	42	1,6	3,8	9,0
Pigeon ramier (<i>Columba palumbus</i>)	Groupes	104	4	-	39	9,3	19,6	41,5
Pinson des arbres (<i>Fringilla coelebs</i>)	Groupes	248	3	-	29	71,6	126,6	223,9
Rougegorge familier (<i>Erithacus rubecula</i>)	Ind.	185	4	0,31	24	14,7	23,4	37,3
Tourterelle turque (<i>Streptopelia decaocto</i>)	Groupes	75	4	-	62	6,5	19,9	61,3

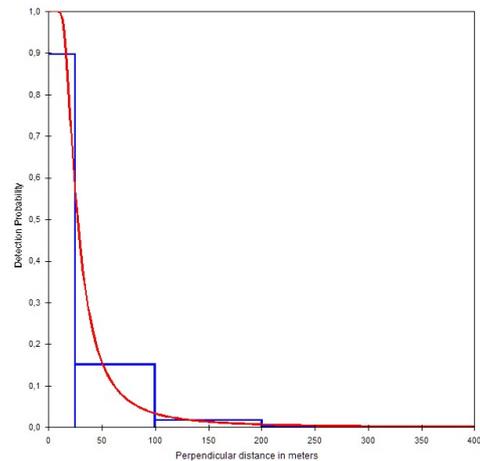
Discussion

En 2019, nous avons testé les résultats du STOC (non publié, mais pour comprendre pourquoi, voir : <http://distancesampling.blogspot.com/p/traiter-les-donnees-stoc-par-distance.html>) : nos conclusions étaient plutôt pessimistes, en constatant l'allure des courbes, souvent incohérentes, probablement à cause d'erreurs de notation des tranches de distances par les observateurs.

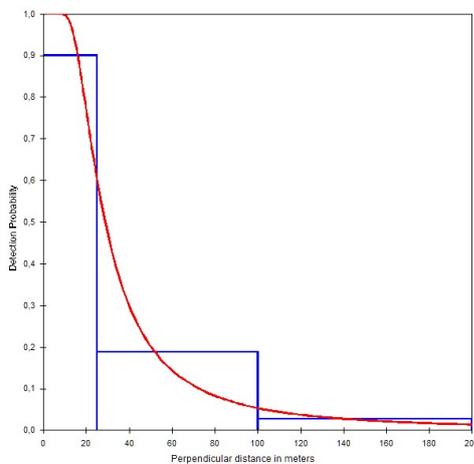
Dans le cas du SHOC, au contraire, nous obtenons souvent des graphes assez cohérents, c'est-à-dire des fonctions décroissantes de probabilité de détection (voir documents 5, 6, 7, 8).



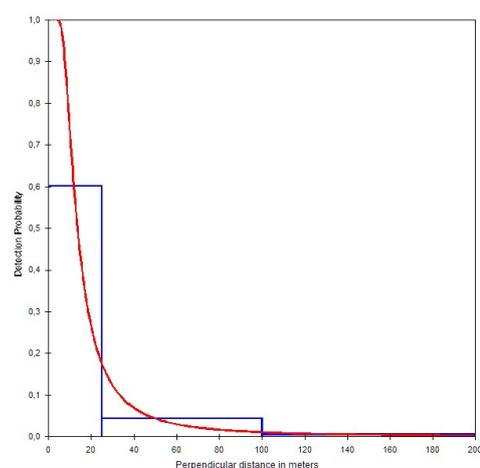
Document 5 : Courbe de probabilité de détection de l'Alouette des champs – *Alauda arvensis*. (Clusters – 4 classes de distance)



Document 6 : Courbe de probabilité de détection du Merle noir – *Turdus merula*. (Analyse individuelle – 4 classes de distance)



Document 7 : Courbe de probabilité de détection du Pinson des arbres – *Fringilla coelebs*. (Clusters – 3 classes de distance)



Document 8 : Courbe de probabilité de détection de la Mésange charbonnière – *Parus major* (Analyse individuelle – 3 classes de distance)

Avec ces 19 espèces, les résultats obtenus sont donc acceptables, mais souvent les fourchettes d'estimation sont extrêmement larges, conséquences de l'échantillonnage faible, des classes de distances assez larges et du petit nombre de classes. Pour les espèces grégaires, les coefficients de variation de la densité sont tous très élevés, ce qui est logique puisque la répartition des individus est par définition en agrégats : l'espèce peut être totalement absente d'un transect et très abondante sur le suivant. Les tests de Khi 2 sont corrects voire très bons pour la moitié des espèces, médiocres pour les autres (<0,30).

Comme pour nos tests sur les points STOC, nous nous sommes posé la question de la précision de notation des distances. Après des centaines d'heures de terrain de travail collectif avec notation de distances sur plan, nous savons que cet exercice est difficile, même pour les ornithologues les plus expérimentés. En milieu découvert, rien n'est plus compliqué que de définir si une alouette est à moins ou plus de 25 m. Il est probable que le taux d'erreur soit important. Les observateurs ont-ils tous le même biais (par exemple sous-estimer les distances?) ou bien la somme de leurs erreurs est-elle compensable ? Nous ne pouvons pas le dire.

En tout cas, il nous semble important que pour les protocoles SHOC (et STOC), l'accent soit mis sur la qualité de notation des distances. L'observateur a tout intérêt à prendre régulièrement des repères sur le terrain pour bien situer les deux premières tranches de distance, cruciales en Distance Sampling.

Il est probable que la méthode SHOC, qui autorise le suivi de chemins existants, entraîne aussi un biais systématique lié aux milieux et à la méthode du transect (ce biais est moindre dans le cas d'un échantillonnage par points) : les chemins en milieu agricole, suivis par les observateurs, sont bordés de haies, et beaucoup d'oiseaux sont concentrés dans ces milieux linéaires, à très courte distance de l'observateur. Cela entraîne donc forcément une surestimation des populations. Les graphes de probabilité de détection sont assez typiques avec une première colonne très haute (exemple du Merle noir, voir doc 6), et une seconde colonne très basse (nous sommes dans les cultures, il n'y a pas de merles !). Il n'y a pas de possibilité de troncature à courte distance par le logiciel à cause du faible nombre d'intervalles : la troncature enlèverait toute la première colonne – la plus importante – du graphe, n'en laissant que 2 ou 3 selon les cas. Il est difficile de lever ce biais sur le terrain : il faudrait soit passer en points d'écoute (mais avec d'autres défauts inhérents à ce type de relevé), soit essayer pour l'observateur de ne pas suivre les chemins systématiquement. Les estimations de populations d'oiseaux hivernants liés aux haies sont-elles surestimées ?

Pour tenter d'évaluer les valeurs de nos estimations de populations, nous aurions voulu les confronter à d'autres études dans la région. Mais il n'existe qu'une seule étude en Auvergne qui fournisse des estimations des populations d'oiseaux hivernants communs sur de grandes surfaces : elle concerne la Réserve Naturelle Nationale du val d'Allier, sur « seulement » 1450 ha avec des milieux très particuliers. Toutes les autres références régionales sont basées sur des comptages relatifs avec des indices d'abondance (IKA) et sont souvent anciennes. Nous n'avons donc pas de base comparative pour valider tout ou partie de nos résultats.

Une autre manière d'évaluer la fiabilité des données obtenues par DS est d'extrapoler les résultats de densité pour examiner leur cohérence générale : si on multiplie les valeurs obtenues sur 1 km² par la surface d'une commune (en moyenne 20 km² en Auvergne), certaines valeurs (Doc 9) paraissent – intuitivement, mais nous pouvons largement nous tromper - trop élevées pour ne pas refléter des erreurs (par exemple Etourneau, Alouette, Linotte) alors que d'autres semblent assez cohérentes : Moineau, Mésanges, Rougegorge, Chardonneret, Merle, Pic épeiche...

	Résultats du SHOC Centre Puy-de-Dôme 2018			Extrapolation à une commune « moyenne » de 20 km ²		
	D min Ind/km ²	Densité Ind/km ²	D max Ind/km ²	Nb minimal d'individus	Estimation Nb d'ind.	Nb maximal d'individus
Alouette des champs (<i>Alauda arvensis</i>)	5,4	14,5	39,1	108	290	782
Chardonneret élégant (<i>Carduelis carduelis</i>)	6,5	16,4	41,6	130	328	832
Corneille noire (<i>Corvus corone</i>)	3	4,4	6,4	60	88	128
Etourneau sansonnet (<i>Sturnus vulgaris</i>)	50,5	86,4	147,9	1 010	1 728	2 958
Geai des chênes (<i>Garrulus glandarius</i>)	4,7	7,6	12,2	94	152	244
Grive draine (<i>Turdus viscivorus</i>)	1,5	2,5	4,1	30	50	82
Linotte mélodieuse (<i>Carduelis cannabina</i>)	10,9	26	62,7	218	520	1 254
Merle noir (<i>Turdus merula</i>)	35,7	59,5	99,2	714	1 190	1 984
Mésange à l. queue (<i>Aegithalos caudatus</i>)	4,6	9,7	20,6	92	194	412
Mésange bleue (<i>Cyanistes caeruleus</i>)	27,6	41,6	62,7	552	832	1 254
Mésange charbonnière (<i>Parus major</i>)	53,7	89,1	147,8	1 074	1 782	2 956
Moineau domestique (<i>Passer domesticus</i>)	70,2	124,4	220,5	1 404	2 488	4 410
Pic épeiche (<i>Dendrocopos major</i>)	2,4	5,2	11,6	48	104	232
Pie bavarde (<i>Pica pica</i>)	1,6	3,8	9	32	76	180
Pigeon ramier (<i>Columba palumbus</i>)	9,3	19,6	41,5	186	392	830
Pinson des arbres (<i>Fringilla coelebs</i>)	71,6	126,6	223,9	1 432	2 532	4 478
Rougegorge familier (<i>Erithacus rubecula</i>)	14,7	23,4	37,3	294	468	746
Tourterelle turque (<i>Streptopelia decaocto</i>)	6,5	19,9	61,3	130	398	1 226

Document 9 : extrapolation des résultats à une surface communale « moyenne » de 20 km²

Conclusion

L'échantillonnage du SHOC, s'il est probablement suffisant pour une analyse nationale (?), ne l'est pas pour une exploitation régionale correcte des données, et encore moins départementale. Outre le faible nombre de carrés, leur répartition pose problème. Il faudrait donc augmenter très considérablement le nombre de carrés, et pour cela convaincre les ornithologues de l'intérêt de ce suivi SHOC. On peut penser que le nombre de données nécessaires pour obtenir des fourchettes de densité fiables doit être d'au moins une centaine. Pour atteindre ce chiffre (et pour un nombre d'espèce suffisant, par exemple les 50 espèces les plus fréquentes), 20 à 30 carrés SHOC nous semblent un bon objectif pour un département (sur deux ans on obtient 2 ans x 2 passages x 3 km x 30 carrés = 360 km d'échantillonnage).

La mauvaise répartition des carrés en Auvergne découle aussi de la répartition très inégale des observateurs : quand on propose de choisir un carré à moins de 10 kilomètres de l'adresse de l'ornithologue, cela réduit considérablement la prospection de certaines zones. Ce critère de choix des carrés ne nous semble pas adapté à un bon échantillonnage régional. La perspective d'évaluations fiables de populations d'oiseaux communs nécessite d'envisager le tirage au sort de carrés témoins sans la contrainte de l'adresse du futur observateur, avec toute l'organisation qui doit accompagner ce type d'enquête. Les observateurs doivent aussi être extrêmement sensibilisés à la nécessité d'obtenir des évaluations de distance les plus fiables possibles, ce qui n'est peut-être pas le cas actuellement.

Remerciements : Merci à Jean-Philippe Meuret pour la mise en forme complexe des données en tableur Excel et fichiers d'imports, à partir de l'extraction d'observations de www.faune-auvergne.org.