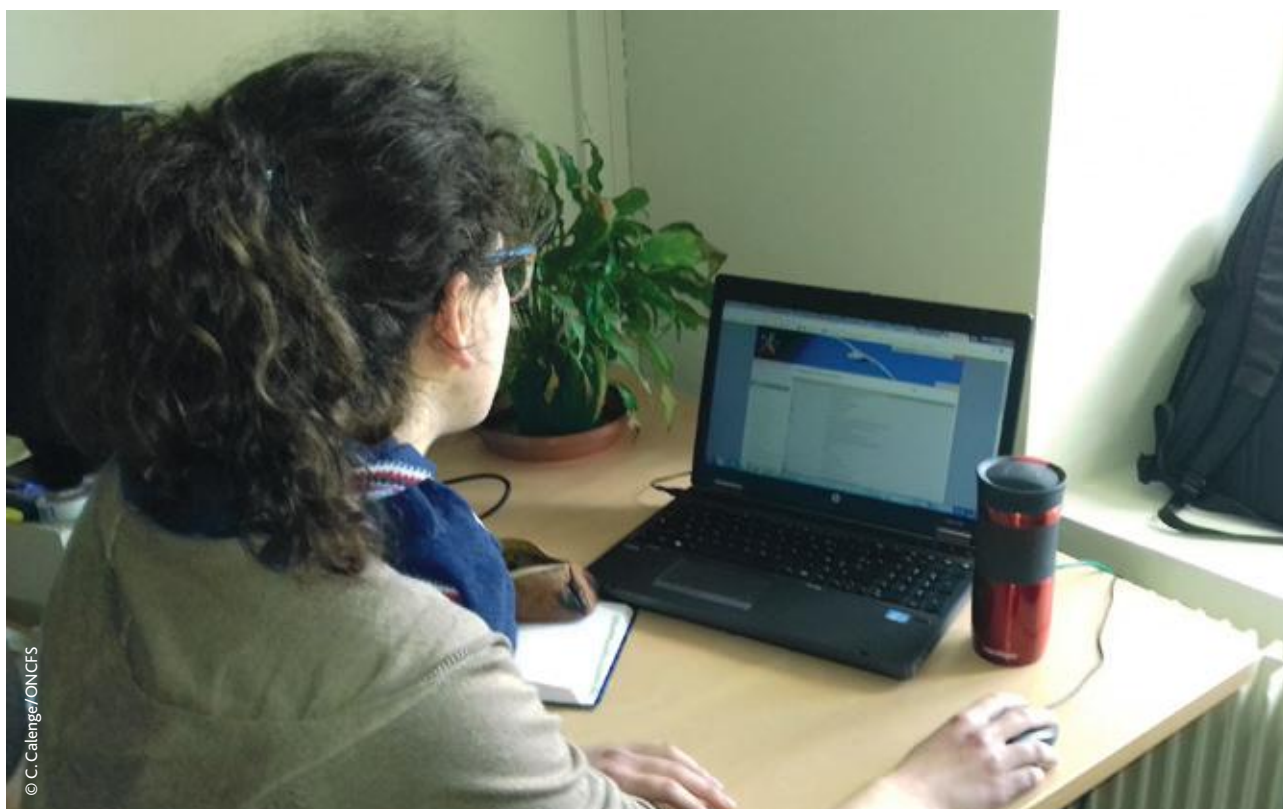




Sciences participatives, sciences citoyennes : quelles méthodes de modélisation pour la connaissance de la faune sauvage ?



© C. Calenge/ONCFS

▲ L'outil internet a permis le développement des sciences citoyennes.

CLÉMENT CALENGE¹, CAMILLE CORON²,
CHRISTOPHE GIRAUD²,
ROMAIN JULLIARD³,
NIRMALA SÉON-MASSIN⁴,

SANDRINE RUETTE⁵, CYRIL ERAUD⁶

¹ ONCFS, Direction de la recherche et de l'expertise, Cellule d'appui méthodologique – Le Perray-en-Yvelines.

² Laboratoire de mathématiques d'Orsay, UMR 8628, Université Paris-Sud – Orsay.

³ CESCO, UMR 7204, MNHN-CNRS-UPMC – Paris.

⁴ ONCFS, Direction de la recherche et de l'expertise – Le-Perray-en-Yvelines.

⁵ ONCFS, Direction de la recherche et de l'expertise, Unité Prédateurs-animaux déprédateurs – Birieux.

⁶ ONCFS, Direction de la recherche et de l'expertise, Unité Avifaune migratrice – Chizé.

Contact : clement.calenge@oncfs.gouv.fr

Le développement récent des programmes de sciences citoyennes en écologie, reposant sur le suivi de la faune par des citoyens, pose de nombreuses interrogations quant à la fiabilité des données collectées. De nouvelles méthodes statistiques permettent aujourd'hui de proposer un traitement de ces données pour corriger certains biais importants inhérents à ces suivis. En combinant l'information apportée par ces programmes avec celle issue des suivis standardisés plus classiques, il devient possible d'améliorer considérablement la précision des estimations d'abondance des espèces et, dans certaines conditions, d'apporter des informations sur les effectifs, les tendances et la répartition des espèces non suivies par les programmes classiques.

Le suivi des espèces animales : des programmes coûteux

Le suivi des populations animales est indispensable pour pouvoir tirer des conclusions argumentées sur leur état de conservation et définir des politiques de gestion appropriées. La quantité et la qualité

des données collectées dans le cadre de ces suivis ont cependant une très grande influence sur la confiance que l'on peut accorder aux conclusions tirées de leur analyse ou de leur modélisation. On comprend aisément qu'il est délicat de conclure lorsque les données sont peu abondantes ou collectées uniquement dans

des zones peu représentatives de la situation pour laquelle l'inférence est recherchée (par exemple, uniquement dans les zones où les effectifs de l'espèce décroissent). La collecte de données sur des surfaces parfois importantes (par exemple, échelle nationale), en général assez coûteuse en temps et en personnel, constitue à cet égard une source d'information de premier plan, et les responsables des programmes de suivis cherchent autant que possible à construire des programmes permettant de collecter d'abondantes données de qualité pour le coût le plus réduit possible.

Travailler en partenariat avec des acteurs locaux compétents est alors une solution commune. En effet, établir des collaborations avec des partenaires locaux professionnels de la faune sauvage – par exemple des gestionnaires d'espaces naturels, des forestiers, des techniciens de fédération départementale des chasseurs, des agriculteurs, etc. – peut être un moyen de partager les coûts et les bénéfices de ces suivis. On parle alors de *sciences participatives* (**encadré 1**). Les réseaux de suivis de la faune sauvage chassable française animés par l'ONCFS, dont certains sont en partenariat avec les fédérations nationale et départementales des chasseurs, mobilisent aussi bien des agents de l'État et de ses établissements publics que des collectivités, des associations, des fédérations, et jusqu'à des particuliers bénévoles pour certains (cf. Body, 2016 pour plus d'information).

Les sciences citoyennes : une alternative très attirante...

Le développement récent d'internet a permis le développement parallèle d'une nouvelle stratégie de suivi de la faune sauvage : les *sciences citoyennes* (**encadré 1**). Ces programmes reposent sur la participation de citoyens amateurs bénévoles, éclairés ou non, à la collecte et éventuellement à l'analyse des données. Le plus souvent, ces programmes fédèrent un grand nombre de citoyens volontaires autour d'un site internet, à partir duquel les utilisateurs renseignent leurs observations sur la faune sauvage. Ces programmes sont souvent pensés comme des réseaux sociaux, permettant aux observateurs bénévoles d'interagir entre eux, ce qui les motive à partager beaucoup de leurs observations avec la communauté. Ces observations sont ensuite utilisées pour tirer des conclusions sur les effectifs, tendance et répartition des espèces suivies. Il s'agit par exemple de données de présence ou d'abondance concernant des espèces rattachées à une grande variété de taxons (oiseaux, mammifères, insectes, batraciens...). Le nombre d'observateurs participant à ce type de programme étant en général très



▲ Les opérations de recensement des perdrix grises reposent sur la participation de nombreux chasseurs.

important, il devient possible de détecter des espèces rares plus facilement qu'avec des suivis classiques. Le coût de tels programmes est en outre modéré du fait de leur caractère bénévole. Enfin, ils permettent de diffuser des connaissances au grand public, de le sensibiliser aux thématiques de recherche et à la conservation de certaines

espèces. Il est donc peu surprenant que ces programmes soient très développés dans le milieu associatif (cf. <http://www.naturefrance.fr/sciences-participatives/le-collectif-national-sciences-participatives-biodiversite> pour une liste de programmes de ce type).



▲ Le Centre commun de recherche de l'UE a développé l'application mobile « Invasive Alien Species Europe », qui permet aux citoyens européens de recueillir et partager des informations dans le cadre de la lutte contre les espèces envahissantes.

► Encadré 1 • Sciences citoyennes, sciences participatives : de quoi s'agit-il ?

Un récent rapport commandé à François Houllier, alors président d'AllEnvi (Alliance nationale de recherche pour l'environnement), par les ministres de l'Éducation nationale et de la Recherche, propose la définition suivante (Houllier *et al.*, 2016) : « Les sciences participatives sont définies comme les formes de production de connaissances scientifiques auxquelles des acteurs non-scientifiques-professionnels, qu'il s'agisse d'individus ou de groupes, participent de façon active et délibérée. »

La notion de sciences participatives recouvre donc un large panel d'approches. Tout d'abord, il s'agit de produire des connaissances scientifiques ; que ce soit pour la recherche fondamentale (par exemple pour faire progresser une théorie scientifique), la recherche appliquée (par exemple pour le développement de nouveaux outils de gestion), ou encore les suivis de populations... Les acteurs mobilisés peuvent également être très divers : du citoyen bénévole et amateur jusqu'au professionnel (non-scientifique donc) qui participe dans le cadre de ses fonctions. Ces acteurs peuvent éventuellement être formés pour assurer la qualité des données recueillies. Enfin, la notion de sciences participatives ne préjuge pas du mode de recueil et de validation des données, tant que celles-ci contribuent à la production de connaissances scientifiques. Ainsi, même si la collecte de données opportunistes (c'est-à-dire sans dispositif d'échantillonnage prédéterminé) est sans doute la plus connue du grand public, le terme de sciences participatives est aussi utilisé dans des études caractérisées par un dispositif très complexe et mis en œuvre de façon stricte. En se basant sur les caractéristiques d'un échantillon considérable de programmes de sciences participatives, Houllier *et al.* (2016) ont défini trois grandes familles de programmes :

- **recherches s'appuyant sur une communauté (Community based research)** : ce type de programme, principalement mis en œuvre dans le domaine médical, correspond aux formes de collaboration entre chercheurs et groupes concernés, pour résoudre les problèmes qui les affectent (on pense par exemple aux collaborations entre chercheurs en médecine et patients atteints d'une maladie, ayant pour objectifs d'améliorer les conditions d'existence de ces derniers) ;

- **recherches participatives proprement dites** : ce type de programme décrit des collaborations entre chercheurs et groupes de citoyens ou de professionnels (chasseurs, agriculteurs, forestiers, associations de protection de la nature...) pour résoudre des problèmes scientifiques ou sociaux. Ainsi, dans le domaine de l'écologie, ces programmes peuvent avoir pour objectifs de définir des règles de gestion d'une population animale, de limiter les dégâts de la faune sur les cultures, de permettre le suivi de populations, etc. Dans ce cadre, certains programmes demandent une formation spécifique quand d'autres peuvent mobiliser largement les bonnes volontés. Ainsi, la capture ou même le comptage de certaines espèces (chevreuil, perdrix grise...) reposent sur la mise en œuvre d'opérations de recensements qui nécessitent des dizaines voire centaines de personnes pouvant venir de tous horizons : chasseurs locaux, scolaires ou tout simplement passionnés. D'autres opérations mobilisent des compétences plus spécifiques : identification des espèces (comptages au chant dans le cadre du réseau « Oiseaux de passage » ONCFS/FNC/FDC), capture et baguage (réseaux « Bécasse » et « Bécassines » ONCFS/FNC/FDC), ou même éviscération du gibier (suivi de la reproduction des cerfs). Une formation spécifique, assurée par les responsables du programme scientifique, est alors indispensable pour participer.

- **Sciences citoyennes** : ce type de programme repose sur une contribution des citoyens à la collecte et éventuellement à l'analyse de données. Avec le développement récent des interfaces Web et en particulier des applications mobiles (Pl@ntnet, Invasive Alien Species Europe...), il devient de plus en plus facile et économique de mettre en œuvre des programmes de ce type à grande échelle et de s'appuyer sur une grande diversité de citoyens pour collecter les données. Ces derniers jouent alors un rôle dans la recherche et l'acquisition de connaissances, et constituent des communautés intéressées par un sujet donné qui vont interagir avec les scientifiques, permettant également de mieux faire connaître les approches scientifiques et donc de sensibiliser la société aux enjeux tant de la recherche que de la protection de l'environnement.



▲ L'ONCFS dispense notamment aux chasseurs volontaires une formation à l'éviscération des grands ongulés leur permettant de participer au suivi de la reproduction.

... mais non sans défauts !

Les programmes de sciences citoyennes ne sont toutefois pas dénués de défauts. Certains sont évidents : le fait que les participants ne soient pas des professionnels soulève naturellement des questions quant à leur compétence pour identifier sans erreur une espèce donnée. La marouette ponctuée reportée à tel endroit n'était-elle pas en réalité un râle d'eau ? Dès lors, si la confusion entre espèces est fréquente, quelle confiance accorder aux données ?

Conscients de ces défauts, la plupart des programmes de sciences citoyennes ont mis en place des procédures pour tenter de les réduire au minimum. Ainsi, beaucoup travaillent à limiter les erreurs d'identification d'espèces en incitant aux échanges entre observateurs, à la prise de photos et aux descriptions détaillées accompagnant les observations. Les responsables de programmes peuvent identifier des observateurs clés et fiables, et valider ainsi une partie des observations *a posteriori*.

Mais malgré ces efforts, les données collectées par de tels programmes ont longtemps été quelque peu dénigrées par les analystes. En effet, les procédures développées dans le domaine de la statistique pour produire des estimations fiables requièrent en règle générale d'importants pré-requis, tels que la nécessité de récolter les données selon un dispositif d'échantillonnage respectant certaines règles, et l'application rigoureuse d'un protocole défini au préalable – le dispositif d'échantillonnage déterminant les lieux et

dates auxquels les observations doivent être faites ; le protocole réglant les conditions de réalisation de ces observations (nombre d'observateurs, durée, etc.).

Certaines de ces contraintes demeurent peu intuitives pour l'observateur non averti. Par exemple, le fait qu'il est important d'aller compter les individus d'une espèce dans des milieux peu favorables, surtout si ces derniers représentent une composante importante du paysage. Ainsi, pour une espèce donnée, une entité paysagère d'une superficie de 10 000 hectares recelant une densité de 0,1 individu/km² seulement, hébergera autant d'individus qu'une zone plus intéressante de 10 hectares où la densité est de 100 individus/km². Pourtant, lorsque le choix de la zone à prospecter est laissé à l'appréciation de l'observateur, les risques sont grands de le voir sélectionner une zone où la densité est élevée et donc les chances importantes de contacter un grand nombre d'animaux. Un comptage dans une zone peu favorable conduira probablement à très peu d'individus vus, voire aucun, et il n'est jamais agréable de clore une session sans avoir fait d'observation ! Les observateurs d'un programme de sciences citoyennes montreront donc naturellement une préférence envers certains sites riches en espèces ou certains types d'habitats plus favorables, préférence difficilement quantifiable à partir des seules données collectées. Ajoutons à cela que la distribution des observateurs à grande échelle dépend également de la densité de population humaine : les observations sont en général plus nombreuses à proximité des villes.

Ces problèmes centraux des sciences citoyennes se traduisent par une variabilité spatiale de l'effort de recherche des animaux. Si l'on était capable de quantifier cet effort (par exemple, en comptant le nombre d'heures passées par les observateurs à recenser des oiseaux en chaque point de la zone d'étude), il serait possible d'utiliser directement les comptages pour estimer des effectifs et des tendances d'évolution des populations, par exemple, en estimant des indices d'abondance du type « nombre d'animaux vus par heure d'observation ». Or cet effort de recherche n'est le plus souvent pas connu, ni même quantifiable. Il est en effet difficile d'évaluer la durée de l'observation lorsqu'un citoyen reporte sa rencontre fortuite avec une espèce rare au cours d'un déplacement depuis son domicile ou son lieu de travail. Ceci pose un problème fondamental pour l'estimation : si l'on observe deux fois plus d'individus dans une zone que dans une autre, c'est peut-être parce que les observateurs y ont passé deux fois plus de temps, ou parce que la densité y est deux fois plus forte. Et il est impossible de distinguer la part de la densité de l'espèce de celle de l'effort de recherche dans le résultat du comptage sans plus d'information. Les données de sciences citoyennes peuvent permettre d'assurer que des espèces protégées sont présentes sur certains sites, justifiant leur statut de protection. Mais il semble difficile de les utiliser pour des estimations d'effectifs, de répartition ou de tendance d'évolution.

▼ Les opérations de captures organisées par l'ONCFS et ses partenaires sont un exemple de recherche participative engageant une collaboration entre professionnels et particuliers bénévoles.





© P. Massit/ONCFS

▲ Les comptages menés par les agents de l'ONCFS peuvent permettre de corriger les biais de collecte des programmes citoyens.

Combiner les programmes citoyens et les programmes classiques

Plusieurs approches ont été proposées dans la littérature scientifique pour corriger ce problème de distribution inconnue et inégale de l'effort de recherche dans les programmes citoyens (e.g. Warton *et al.*, 2013; Isaac *et al.*, 2014; Giraud *et al.*, 2015). Celle que nous jugeons la plus prometteuse consiste à corriger ces défauts en s'appuyant sur les données collectées en parallèle par un programme de suivi plus classique, pour lequel l'effort d'échantillonnage est connu voire, dans l'idéal, contrôlé. Il est intéressant de noter que le principe de cette approche a été simultanément et indépendamment proposé par plusieurs équipes de recherche dont la nôtre (cf. Giraud *et al.*, 2015).

Cette approche consiste à s'appuyer sur la connaissance partielle des populations animales étudiées, que l'on acquiert avec un programme de suivi classique, pour corriger les biais de collecte associés au programme de sciences citoyennes et compléter ces connaissances. En d'autres termes, cela revient à calibrer le programme de sciences citoyennes à l'aide du programme classique pour en corriger certains biais. Le principe de cette approche, extrêmement simplifié ici, est illustré dans l'**encadré 2** (cf. Giraud *et al.* (2015) pour le principe mathématique général; une méthode pratique basée sur ce principe est en cours de développement – Coron *et al.*, en prép.).

Considérons le cas où l'on dispose de deux jeux de données, collectés respectivement par un programme de sciences citoyennes et par un programme de suivi plus

► Encadré 2 • Combiner des données de suivi professionnel avec des données issues des sciences citoyennes

Imaginons la situation suivante : un programme de sciences citoyennes conduit à une détection de 20 alouettes des champs dans une commune A, et de 60 alouettes dans une commune B. Un programme de suivi professionnel, avec effort d'échantillonnage contrôlé et connu, mené sur les mêmes communes, produit une estimation de 20 alouettes dans la commune A, et 10 alouettes dans la commune B. Ainsi, selon le programme de sciences citoyennes, les alouettes seraient trois fois plus abondantes dans la commune B que sur la commune A. À l'inverse, selon le programme de suivi professionnel, l'espèce serait deux fois moins abondante dans la commune B que dans la A.

Comment expliquer ce désaccord entre les deux programmes ? L'explication la plus probable est que l'effort de recherche du programme de sciences citoyennes était disproportionné dans la commune B par rapport à la commune A. Ainsi, tout se passe comme si le programme de sciences citoyennes avait conduit six opérations de recensement dans la commune B et avait sommé les résultats de ces six opérations. Dans la mesure où nous disposons d'un programme de suivi professionnel, nous sommes alors en mesure d'estimer l'effort de recherche caractérisant le programme de sciences citoyennes dans chaque commune (ici : effort identique pour les deux programmes dans la commune A, et six fois plus important pour la commune B), et donc de corriger l'estimation du programme de sciences citoyennes.

À présent, supposons que le même programme de sciences citoyennes poursuive également l'objectif de suivre les populations de perdrix grises, mais que ce ne soit pas le cas du programme professionnel. Imaginons que ce programme de sciences citoyennes conduise de la même façon à un dénombrement de 40 perdrix grises dans la commune A et de 120 perdrix grises dans la commune B. Si l'on suppose que l'effort de recherche du programme de sciences citoyennes est le même pour toutes les espèces suivies, alors nous savons que l'abondance réelle de la perdrix grise est correctement estimée pour la commune A, mais qu'elle est six fois trop importante pour la commune B en comparaison de la réalité. Une meilleure estimation ici serait donc de 20 perdrix grises pour la commune B. On voit alors qu'il est possible d'obtenir une estimation d'abondance pour la perdrix grise pour toutes les communes, et ce, sans que cette espèce ne soit suivie par un programme professionnel. Cette pratique consistant à corriger les biais des programmes de sciences citoyennes, en « calant » les estimations sur celles obtenues par un suivi professionnel avec effort contrôlé, est rapportée de plus en plus souvent dans la littérature (cf. Giraud *et al.*, 2015).



▲ *Le biais lié à la variabilité de l'effort de recherche dans les programmes citoyens peut être corrigé en s'appuyant sur les données recueillies dans le cadre des programmes de suivi plus classiques.*

professionnel. La zone géographique d'intérêt est découpée en « unités spatiales » (par exemple des communes, des quadrats, des départements, des pays, etc.). Chaque jeu de données contient un nombre d'individus de différentes espèces détectés dans les différentes unités. Toutes les espèces ne sont pas forcément suivies par les deux types de programmes : certaines d'entre elles peuvent n'être suivies que par le programme citoyen. Nous disposons d'observations pour les deux jeux de données dans toutes les unités étudiées (voir application dans l'**encadré 3**). Nous connaissons par ailleurs la composition environnementale de chaque unité (par exemple, proportion de forêt, de prairie, etc.) pertinente biologiquement pour l'espèce étudiée. La méthode développée permet d'estimer :

- **l'indice d'abondance de chaque espèce**, aussi appelé « abondance relative », dans chaque unité : il s'agit de la densité réelle de l'espèce dans l'unité, multipliée par un coefficient de valeur inconnue et propre à chaque espèce. Ainsi, notre méthode ne nous permet pas de dire si l'abondance d'une espèce est importante ou faible dans l'absolu, mais elle nous permet par exemple de dire que cette abondance est trois fois plus importante dans telle unité que dans telle autre. Notons que les estimations de ces indices d'abondance, obtenues pour les espèces suivies par les deux programmes, sont beaucoup plus précises que si l'on avait utilisé uniquement les données du seul programme professionnel ; l'abondance de données des programmes citoyens permet d'augmenter cette précision. L'intérêt de cette approche est également de permettre d'estimer l'abondance relative d'espèces suivies uniquement par le programme citoyen. Enfin, pour les espèces très rares, la possibilité de s'appuyer sur la forte pression

d'observation de programmes citoyens, afin de dégager des tendances populationnelles, est aussi ce qui rend cette approche très intéressante ;

- **l'effort de recherche** des animaux dans une unité réalisé par les observateurs du programme de sciences citoyennes : comme indiqué plus haut, le nombre variable d'observateurs dans les différentes unités, et leurs préférences pour tel site d'observation emblématique ou tel type d'habitat plus favorable, vont résulter en un temps inégal passé dans les différentes unités qui composent la région d'intérêt. En revanche, cet effort est connu pour le programme professionnel. Notre approche permet d'estimer l'effort de recherche à une constante près pour le programme citoyen, sous l'hypothèse que cet effort est identique pour toutes les espèces et ne varie qu'en fonction de l'unité ;

- **un paramètre combinant la probabilité de détection et la probabilité de reporter l'observation d'une espèce dans chaque programme** : on suppose que les différentes espèces ne sont pas détectables de façon identique (selon leur discrétion). De même, la détectabilité des espèces peut varier entre les deux programmes (par exemple, des observateurs professionnels pourront détecter plus facilement certaines espèces que des amateurs). Par ailleurs, la motivation des citoyens à reporter leurs observations peut varier en fonction de l'espèce. Un observateur amateur détectant un oiseau rare sera motivé à reporter son observation à la communauté d'observateurs bénévoles qui met en œuvre le programme, mais ne prendra pas toujours la peine de reporter l'observation d'une espèce très commune. Notre modèle permet de prendre en compte ces probabilités de détection et de reports inégales entre espèces et entre programmes

lors de l'estimation des indices d'abondance. En revanche, nous supposons que la probabilité de détection et la probabilité de report ne varient pas d'une unité à l'autre ;

- **La préférence des observateurs du programme de sciences citoyennes pour un type d'habitat** : notre modèle permet de déterminer quels sont les habitats préférentiellement prospectés par les observateurs du programme citoyen, et de prendre en compte ces préférences pour tendre vers une estimation sans biais des indices d'abondance ;

- **La préférence d'une espèce pour un type d'habitat** : les espèces étudiées n'utilisent pas les différents types d'habitats avec la même fréquence ; notre approche permet de déterminer et prendre en compte leurs préférences sur l'ensemble de la zone d'étude.

Les hypothèses du modèle

Notre approche repose sur un certain nombre d'hypothèses listées ci-après, qui doivent être impérativement respectées.

- Les espèces doivent être identifiées sans erreur. Comme nous l'avons évoqué précédemment, la présence d'erreurs d'identification par les observateurs peut conduire à des biais dans les estimations d'abondance des espèces. Notons que si le risque de biais sera vraisemblablement mineur pour les espèces les plus fréquentes, il sera plus important pour les espèces rares. Par exemple, si une proportion même infime d'une espèce fréquente (le râle d'eau) est confondue avec une espèce rare (la marouette ponctuée), cela risque de conduire à une surestimation considérable de l'abondance de cette dernière. Ce point doit être pris très au sérieux car l'un des principaux attraits des sciences citoyennes est justement de permettre le suivi des espèces rares.

- Il ne doit pas y avoir d'effet observateur sur la probabilité de détection et de report dans le programme de suivi « professionnel ». Le principe de la méthode que nous proposons repose sur la connaissance de l'effort de recherche des animaux pour chaque unité dans le programme de suivi « professionnel » : cette connaissance permet de corriger les biais du programme citoyen en « calant » les estimations du programme citoyen sur le programme professionnel. Or, l'hétérogénéité individuelle, liée entre autres à la motivation, à l'expérience et aux capacités de perception des observateurs dans le programme professionnel, peut conduire à une hétérogénéité de la probabilité de détection d'un animal entre observateurs difficile à quantifier, mais qui peut jouer sur cette connaissance de l'effort de recherche. Ainsi, on comprendra aisément que le temps passé est une mesure inappropriée de l'effort

► Encadré 3 • Combiner les données de « Faune Aquitaine » et du protocole ACT

L'article de Giraud *et al.* (2015) présente un exemple d'application de ce cadre théorique, que nous reproduisons en partie ici. Il repose sur les deux jeux de données ci-après.

- **Le jeu de données standardisé « ACT »** collecté par le réseau participatif « Oiseaux de passage » (ONCFS/FNC/FDC) : ce jeu de données est recueilli selon un dispositif et un protocole d'échantillonnage standardisés. L'acronyme ACT signifie « Alaudidés-Columbidés-Turdidés » (alouettes, pigeons et tourterelles, grives et merles), ces groupes d'espèces contenant les principaux oiseaux suivis par ce programme (qui suit également plusieurs espèces de corvidés et de galliformes). Nous nous sommes concentrés ici sur 13 espèces d'oiseaux (voir Giraud *et al.* (2015) pour plus de détails). La région Aquitaine a été découpée en 64 quadrats, et dans chacun une route de 4 km a été placée aléatoirement dans l'habitat non urbain. Chaque route comportait cinq points de comptages séparés de 1 km et était parcourue deux fois entre avril et mi-juin. Sur chaque point, le comptage durait dix minutes et pour chaque espèce d'intérêt, le plus élevé des deux comptages a été retenu comme nombre détecté.

- **Le jeu de données citoyen « Faune Aquitaine »** collecté par la LPO Aquitaine : le programme « Faune Aquitaine » est organisé autour du site internet www.fauneaquitaine.org, et permet à toute personne de reporter ses observations d'oiseaux (toutes espèces présentes) dans la région éponyme. Nous nous sommes concentrés sur les 34 espèces d'oiseaux qui y sont le plus communément observées. Notons qu'elles incluent les 13 espèces suivies par le programme ACT. Au total, quelque 115 000 observations d'oiseaux ont pu être effectuées dans le cadre du programme « Faune Aquitaine » dans les quadrats définis par le programme ACT. Nous ne disposons pour aucun endroit de la région d'une mesure de l'effort de recherche, et il y a de fortes chances que cet effort soit distribué très inégalement dans l'espace. Par conséquent, il est impossible d'utiliser directement le résultat du dénombrement d'une espèce dans un quadrat donné comme un indice d'abondance de l'espèce à cet endroit. Il est donc nécessaire d'utiliser notre approche de modélisation pour calibrer ces données citoyennes à l'aide des données standardisées, estimer ainsi l'effort de recherche et par suite l'abondance relative des 34 espèces (c'est-à-dire les 13 espèces suivies par les deux programmes et les 21 espèces suivies uniquement par « Faune Aquitaine »).

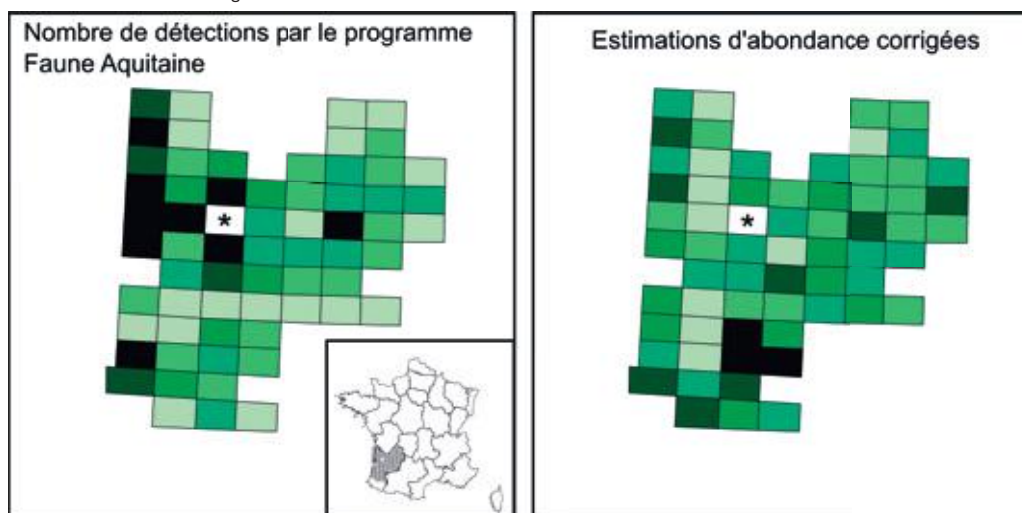
Nous avons combiné ces deux jeux de données à l'aide d'une version simplifiée de notre approche, corrigeant la distribution inégale de l'effort de recherche et la détectabilité des espèces, mais ne prenant pas en compte les préférences variables des observateurs et des espèces pour différents habitats (un modèle plus complexe intégrant ces préférences est en cours de développement). Nous avons pu estimer des indices d'abondance pour toutes les espèces d'intérêt, y compris celles suivies uniquement par le programme citoyen. La *figure* illustre ce dernier point pour la mésange à longue queue, suivie uniquement par « Faune Aquitaine ».

Afin de pouvoir juger de la qualité de nos estimations, nous avons utilisé un jeu de données indépendant, collecté dans la région Aquitaine et concernant les mêmes espèces d'oiseaux, dans le cadre du programme STOC-EPS (*Suivi temporel des oiseaux communs*) du Muséum national d'histoire naturelle. Ce programme repose sur un échantillonnage aléatoire stratifié et offre l'avantage, comme le programme ACT, d'être standardisé. Nous avons alors calculé le coefficient de corrélation entre les estimations de densités des 34 espèces à l'aide de notre approche et les comptages réalisés par le programme STOC, afin d'évaluer la qualité des prédictions de notre modèle. Il apparaît très clairement que :

- pour les espèces suivies par les programmes professionnels et de sciences citoyennes, on prédit plus précisément les comptages des espèces dans une unité lorsqu'on combine données citoyennes et données standardisées professionnelles, que lorsqu'on s'appuie uniquement sur les données standardisées professionnelles. Les abondantes données citoyennes permettent d'améliorer fortement la précision des estimations et peuvent apporter une réelle plus-value aux données standardisées professionnelles ;
- cette combinaison permet également d'estimer correctement les abondances des espèces non suivies par les programmes professionnels. Les suivis standardisés professionnels permettent de calibrer les programmes citoyens en corrigeant les biais de collecte, et donc permettent de tirer parti de l'importante pression d'observation mise en œuvre par les programmes citoyens pour suivre les espèces les plus rares.

Figure Exemple de correction des biais dans l'estimation d'abondance de la mésange à longue queue dans différentes unités spatiales à l'aide de notre approche.

Les mailles les plus sombres correspondent aux valeurs les plus élevées. Cette espèce n'était pas suivie par le programme ACT, mais ce dernier nous a malgré tout permis de corriger les biais de collecte du programme citoyen. À noter que nous avons éliminé le quadrat contenant Bordeaux car le milieu y était très différent du reste de la région.



de recherche, si pendant un même laps de temps un observateur A voit moitié moins d'animaux qu'un observateur B au même endroit. Autrement dit, la présence de cette hétérogénéité individuelle attachée à la probabilité de détection signifie que l'effort de recherche par observateur n'est en réalité pas connu exactement dans le programme professionnel, et qu'il est donc difficile de s'en servir pour corriger les biais du programme citoyen. Cependant, la violation de cette hypothèse n'induit pas de biais très important si, globalement et en moyenne, l'hétérogénéité de la probabilité de détection est la même d'une unité à l'autre (c'est-à-dire, si les observateurs expérimentés et performants sont distribués uniformément dans l'espace, tout comme les observateurs inexpérimentés et/ou moins performants).

- Nous supposons que la préférence des observateurs du programme citoyen pour certains types d'habitats est constante dans l'espace. Cette hypothèse peut se révéler assez contraignante. En effet, la sélection des types d'habitats par l'observateur peut dépendre de la région géographique dans laquelle il se trouve. Par exemple, si le programme se concentre sur le suivi d'oiseaux, un observateur dans la Dombes (Ain) pourra vouloir observer l'avifaune inféodée aux zones humides (plus caractéristiques de la région) et donc préférer ce type d'habitat ; tandis qu'un observateur dans les Cévennes pourra vouloir observer préférentiellement des rapaces... De telles variations de préférence risquent de générer des biais dans les estimations d'abondance ; les espèces des milieux préférés dans une région se trouvant alors surestimées, contrairement aux espèces des milieux délaissés par les observateurs.

- Notons que nous supposons, pour les mêmes raisons, que la préférence des espèces pour différents types d'habitats est constante dans l'espace ; autrement dit, une espèce ne devra pas préférer tel type de milieu dans une région et l'éviter dans une autre (une hypothèse souvent raisonnable, mais à vérifier).

- Notre approche suppose que la probabilité de détection et de report d'un animal d'une espèce donnée par un observateur présent au même endroit est constante d'une unité à l'autre. Plusieurs sources de violation de cette hypothèse peuvent entraîner des biais dans les estimations. Par exemple, la détectabilité des animaux peut être plus faible dans certaines unités à cause de l'importance de la végétation. Les enjeux de conservation peuvent également changer d'une unité à l'autre, entraînant une attention variable des observateurs envers différentes espèces. Cette question doit être considérée sérieusement lors de la mise en œuvre de la méthode.

Les sciences citoyennes comme complément, non comme substitut

L'approche que nous proposons permettrait de réaliser de nouvelles formes de suivi de la faune. L'exemple donné dans l'**encadré 3** illustre les bénéfices potentiels de cette approche collaborative : les données citoyennes permettent d'améliorer la précision des estimations d'abondance pour les espèces suivies par des programmes professionnels (qui manquent souvent de précision pour les espèces les plus rares), et les données de suivi professionnel permettent d'éliminer certains des biais inhérents aux programmes de sciences citoyennes, autorisant ainsi la cartographie des abondances ou l'estimation de tendances d'évolution de populations d'espèces non suivies par les programmes professionnels. Notons qu'il serait intéressant de travailler en amont lors de la conception des programmes citoyens et professionnels, de façon à définir les zones et périodes d'études, les résolutions spatiales et les conditions de collectes des données qui optimiseraient les conditions d'application de notre approche. De plus amples recherches sont encore nécessaires, afin de déterminer dans quelles conditions cette combinaison de données de différentes natures se traduirait par les estimations les plus précises.

Remerciements

Nous remercions très sincèrement Denis Roux (ONCFS) et tous les observateurs du réseau ACT. Nous tenons également à exprimer notre gratitude aux responsables du programme « Faune d'Aquitaine », ainsi qu'à tous les observateurs bénévoles ayant participé à ce programme. Nous remercions enfin les responsables du programme STOC-EPS et tous les observateurs de ce réseau pour leur contribution. ●

Bibliographie

- Body, G. 2016. La faune sauvage à la trace. Le suivi des espèces à l'ONCFS. *Brochure de l'Office national de la chasse et de la faune sauvage*, coll. « Carnets ». 96 p.
- Coron, C., Giraud, C., Calenge, C. & Julliard, R. En prép. Combining citizen science and standardized data for the monitoring of biodiversity.
- Giraud, C., Calenge, C., Coron, C. & Julliard, R. 2015. Capitalizing on opportunistic data for monitoring relative abundances of species. *Biometrics* 72: 649-658.
- Houllier, F. et al. 2016. Les sciences participatives en France – Synthèse. *Rapport élaboré à la demande des ministres en charge de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche*. 8 p.
- Isaac, N.J., Strien, A.J., August, T.A., Zeeuw, M.P. & Roy, D.B. 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* 5: 1052-1060.
- Warton, D.I., Renner, I.W. & Ramp, D. 2013. Model-Based Control of Observer Bias for the Analysis of Presence-Only Data in Ecology. *PLoS one*, 8, e79168.

▼ Les données citoyennes permettent d'affiner les estimations produites par les programmes de suivi professionnel ; tandis que ces derniers permettent d'éliminer certains biais liés aux programmes de sciences citoyennes, comme l'illustre le cas de la mésange à longue queue (cf. encadré 3).

