

La publication scientifique est-elle encore à la portée des naturalistes ?

Réflexions sur l'importance accordée aux analyses statistiques informatiques

Par **Michel BARATAUD**
(michel.barataud1@orange.fr)

*Devant la science rationaliste
comme devant l'industrie,
la vie se retire, insaisissable.
L'homme d'action sans sympathie
ne connaît qu'en lui la vie, la
spontanéité, le désir. Rien ne lui
répond et la réalité, passive, est
pour lui une grande prostituée.
[...] Que (la science) prenne un
phénomène naturel, un être
vivant, c'est pour le découper,
idéalement, en pièces inertes.
Force, mouvement, volonté,
conscience, ce qui est la réalité
même, recule d'autant. [...]
Et si pourtant j'avais raison ?
Car lorsqu'on veut douter de soi,
il faut aussi douter de son doute.
Raison contre presque tous,
cela ne présente pas
grande probabilité mais enfin
il n'y a pas que la statistique.*

Robert HAINARD
Nature et mécanisme ; 1946.

*L'histoire naturelle est l'une des
seules sinon la seule matière
dont l'étude tend à affiner la
perception du concret, et surtout
à développer le sentiment, si
nécessaire, de la complexité des
choses. Ces souples contours
du vivant qui narguent toute
réduction à la géométrie, cette
complication indescriptible de la
structure qui fait de chaque être
organisé un petit univers [...], tout
cela fait de l'objet vital une source
irremplaçable d'enseignements,
propres à dissiper certaine illusion
mathématicienne qui est de
croire que les réalités se laissent
rigoureusement définir et qu'on
en épuise le contenu par les outils
de la pure logique.*

Jean ROSTAND
Le droit d'être naturaliste ; 1963.

La publication scientifique est actuellement un exercice très codifié, aux exigences d'autant plus fortes que la revue a un facteur d'impact élevé. Parmi les fondamentaux imposés, le traitement statistique des données concentre une part grandissante de l'attention des comités de relecture. Le succès de cette discipline est justifié car il améliore souvent la compréhension de la causalité des phénomènes ou des limites de leurs interprétations. Bien sûr, comme dans tout sujet, des excès apparaissent, comme pour les biologistes dont la culture les rend plus à l'aise en statistique qu'en recherche de terrain, générant parfois une inversion de logique : les données brutes, matière première d'une étude, sont parfois considérées comme un prétexte à valoriser un type d'analyse statistique innovant. Mais l'intention ici n'est pas de remettre en cause l'utilité des analyses statistiques fréquentielles assistées par ordinateur. Leur puissance de traitement de volumes colossaux de données brutes justifie l'intérêt qui leur est accordé. Le but de cet essai est plutôt d'amener à une réflexion sur la façon dont leur pouvoir s'est imposé sous cette forme exclusive.

Les travaux à caractère scientifique en biologie ou en écologie de type « naturaliste amateur » (couramment accepté comme étant en opposition avec le type « recherche scientifique », majoritairement universitaire), ne peuvent accéder à de nombreux supports de publication lorsque la vérification statistique de leurs résultats selon une méthode *reconnue* est jugée faible ou absente.

¹ Cette étape n'est réalisable sous cette forme que pour les thématiques compatibles avec un processus expérimental ; dans de nombreux domaines de l'écologie ou en biologie de l'évolution par exemple, les méthodes expérimentales sont remplacées par des méthodes déductives.

² Francis Bacon (1561-1626), philosophe et scientifique anglais est l'un des pionniers de la pensée scientifique moderne ; il développe dans « *De dignitate et augmentis scientiarum* » une théorie empiriste de la connaissance, et précise dans *Novum Organum* les règles de la méthode expérimentale.

Ce dogmatisme résulte-t-il d'un processus lent et logique aboutissant à une sélection objective, ou correspond-t-il à un changement assez soudain et récent de paradigme dans notre perception et notre conception du monde ?

Une approche plus tolérante aux capacités d'analyse du cerveau humain pourrait-elle trouver ou retrouver une place dans la ligne éditoriale des revues scientifiques ?

Le raisonnement scientifique : à la recherche de l'objectivité

La difficulté d'accorder au terme « scientifique » une définition consensuelle, tient notamment à trois causes : le caractère diachronique de la pensée scientifique et des méthodes qui lui sont rattachées, le nombre de disciplines qui prétendent inclure un raisonnement scientifique dans leurs pratiques, et la variabilité des domaines dans lesquels ces disciplines ont été classées.

Les sciences naturelles, rangées dans la catégorie des sciences expérimentales, appliquent la méthodologie scientifique classique (observation, interrogation, hypothèse, vérification par l'expérimentation¹, analyse des résultats, publication) déjà décrite dans les grandes lignes par Bacon² au début du XVIIe siècle, qui peut être correctement réalisée quel que soit le cadre des travaux : naturalistes ou universitaires, bénévoles ou professionnels.

Mais le raisonnement scientifique prêle le flanc à un « paradoxe circulaire » : comment isoler une connaissance du monde (en nous et autour de nous), de la subjectivité liée aux sens de l'Humain qui est à la fois auteur du sujet et de la méthode, percepteur des données, acteur de leur collecte et de leur analyse, puis enfin unique censeur des résultats ?

Au cours de l'Histoire, ce paradoxe a été assumé et traité de façons diverses, plus ou moins rationnelles, mais il semble bien que la suprématie des chiffres se soit affirmée pour des raisons d'objectivité.

La puissance des nombres, la fragilité des mots ?

« Au commencement était le verbe » est-il écrit dans la traduction par CRAMPON (1864) du Prologue de l'évangile selon Jean. Cela est vrai sans doute aux débuts de la communication sociale entre individus par langage articulé (VERNIER 2005). Mais la communication entre sociétés par langage écrit a sans doute débuté avec des chiffres, pour des besoins de transactions commerciales (DIAMOND 1997, p. 327). Ainsi le pouvoir des nombres ne date pas d'hier.

Pourtant la communication scientifique en sciences naturelles a longtemps été pratiquée par des personnes lettrées soucieuses de la forme, qui s'exprimaient encore récemment sous forme d'essais narratifs, des récits d'observations et d'expérimentations. Charles Darwin, dans son ouvrage sur l'origine des espèces publié en 1859, a méthodiquement compilé des observations et

des expérimentations durant une bonne partie de sa vie, avant d'avancer sa théorie qui n'a pu être étayée ni par des analyses statistiques informatiques ni par une technologie capable de lui faire appréhender les mécanismes génétiques ; son travail a révolutionné notre vision de la vie alors qu'il n'avait aucune disposition pour les mathématiques (ROSTAND 1963).

Grâce aux nouveautés technologiques (notamment l'informatique et sa puissance de calcul), à un accès facilité aux études supérieures et à une démocratisation des supports de vulgarisation scientifique, ces dernières décennies ont connu une recrudescence du nombre de travaux scientifiques. Publier ses travaux est devenu souvent un indicateur quantitatif vital pour la carrière d'un chercheur. Le nombre de publications augmente de manière exponentielle [passant de 466 419 publications scientifiques toutes disciplines confondues en 1988 selon la National Science Foundation, à 986 099 en 2008 selon le rapport sur la science de l'Institut de statistique de l'UNESCO (2010)]. La plupart des revues impose désormais une codification plus stricte, une norme de plus en plus élitiste.

On retrouve ici une image du schéma reconstitué pour l'évolution historique des petits clans humains vers les grandes sociétés structurées (DIAMOND 1997). Ce changement de paradigme ne serait-il pas lié majoritairement à une question d'échelle quantitative ? Au-delà d'un certain seuil, la quantité entraîne une multiplication des règles standard imposées. Cela s'accompagne d'un accroissement des enjeux de pouvoir ; ce dernier devient de plus en plus centralisé, hiérarchisé et déconnecté de la variabilité des aspirations individuelles. On

observe une standardisation plus stricte des comportements, avec une perte de valorisation des savoirs non codifiés.

Ce processus s'est déroulé en médecine par exemple, avec à l'origine une organisation en clans de petite taille où les chamans (dont la pensée holistique en faisait des précurseurs de l'écologie appliquée) avaient entre autres compétences une maîtrise des soins parfaitement fonctionnelle qui résultait d'un cumul d'expérimentations sur plusieurs générations (la transmission orale était sans doute ici un facteur favorisant car elle permettait de faire passer des subtilités non transcriptibles). L'augmentation des populations, entraînant une structuration forte des sociétés humaines, a imposé un dogme généralisé (en France, principalement religieux pendant l'inquisition³ puis politico industriel en 1941 sous le régime de Vichy⁴) qui a fait disparaître ces savoirs locaux multiples. D'autres acteurs (institutionnels) et d'autres pratiques (industrielles) les ont remplacés, qui se distinguent par leur volonté d'exclusivité en Occident.

On retrouve ici une tendance générale du fonctionnement humain, qui rappelle, à une autre échelle avec d'autres causalités, celle des équilibres ponctués par des extinctions de masse (GOULD 2006) : les innovations majeures ne s'imposent souvent qu'au prix de l'écrasement de nombreux acquis antérieurs.

Serait-ce possible que le raisonnement scientifique, qui par définition tente d'échapper aux pulsions instinctives

en ne retenant que l'analyse objective, n'échappe pas à cette tendance comportementale des sociétés humaines ?

Sciences « dures » et sciences « molles »

Un des exemples fameux de ségrégation en sciences est celui de l'opposition entre sciences exactes et sciences humaines, aussi dénommées sciences dures et sciences molles (on suppose que ce sont les tenants des premières qui ont baptisées ainsi les secondes et non l'inverse : ici les termes semblent significatifs d'un état d'esprit).

Les prémices connues de la pensée scientifique en Occident, dans la Grèce Antique principalement, n'établissent pas de cloisonnement entre les disciplines, la philosophie (en rupture avec les mythologies) se voulant rassembleuse des formes de pensée conceptuelle sur l'esprit et la matière. Une première transition, légère, s'opère entre le VI^e siècle av. J.C. (Héraclite, Démocrite, Pythagore) et le IV^e siècle av. J.C. (Platon, Euclide, Aristote) avec une sécularisation et une rationalisation de la pensée scientifique, qui reste cependant ouverte avec de nombreuses passerelles entre disciplines et un goût pour les échanges (l'art du discours, de la dialectique). Cette perméabilité va continuer à imprégner l'approche des sciences durant toute la période médiévale en Europe avec notamment Augustin d'Hippone (354 - 430) et Thomas d'Aquin (1224 - 1274),

³ Les sorcières brûlées par la Sainte Inquisition n'étaient souvent que de simples herboristes et rebouteuses monnayant leur savoir et leurs aptitudes à guérir auprès de la population..

⁴ La loi du 11 septembre 1941 du gouvernement de Vichy supprima le métier d'herboriste, au profit de la création d'un ordre des pharmaciens à qui furent confiés le monopole de la vente et du conseil en matière de plantes.

même si l'Église catholique impose avec force son rôle ambivalent de mécène et de censeur jusqu'au procès de Galilée en 1633 qui marque la fin progressive de la tradition scolastique, notamment sous l'influence de l'Académie des Lyncéens⁵.

Après 2000 ans de pratique, cette façon globale et perméable d'appréhender globalement les disciplines visant à la connaissance au sens plus large, s'éteint avec la Renaissance. L'importance accordée aux chiffres devient dominante à partir du XVI^e siècle. Galilée (1564 - 1642) écrit : « la nature est écrite dans un langage mathématique » ; René Descartes (1595 - 1650), chantre de la certitude mathématique, combat la rhétorique et confirme l'homme comme « maître et possesseur de la nature » ; Blaise Pascal (1623 - 1662) est à l'origine des probabilités avec son « problème des partis » ; Baruch Spinoza (1632 - 1677) écrit une philosophie « à la manière géométrique » ; Gottfried Wilhelm Leibniz (1646 - 1716) est à l'origine de la machine à calculer et de l'arithmétique binaire ; Isaac Newton (1642 - 1727) fonde la mécanique classique, la théorie de la gravitation universelle et déclare : « les hypothèses n'ont pas leur place dans la philosophie expérimentale ». La plupart des penseurs de cette période condamnent l'imagination et dépouillent la philosophie de sa composante humaine au profit de la véracité des mathématiques et de la physique. Cette influence va imprégner les sociétés occidentales, malgré la réconciliation entre « raison » et « nature » durant le siècle des Lumières, la raison dynamique

de Georg Wilhelm Friedrich Hegel (1770 - 1831) qui lie l'Histoire à son contexte, le positivisme d'Auguste Comte (1798 - 1857) qui tente une classification non normative des sciences, la vision révolutionnaire du monde vivant issue des travaux de Charles Darwin (1809 - 1882) basés sur l'observation et la réflexion, et le naturalisme de Friedrich Nietzsche (1844 - 1900) qui « replonge l'homme dans la nature ». Encore très récemment dans les années 1990, la sociologie par exemple s'est vue contester son statut de science⁶. Pourtant certains essais récents comme celui de Jared DIAMOND (1997) montrent de manière très convaincante qu'une approche méthodologique combinant les sciences humaines et naturelles s'avère très pertinente dans la compréhension de processus diachroniques.

Aux XX^e et XXI^e siècles, le poids des sciences « dures » est tel qu'il imprègne (comme dans la psychanalyse de Jacques Lacan) certains discours de sciences humaines : il semble que pour être crédible à l'heure actuelle, les termes empruntés aux sciences mathématiques et physiques, sans parler des statistiques omniprésentes, soient un parcours obligé. En 1963, Jean Rostand, biologiste et écrivain qui, de son propre aveu, était dans sa jeunesse plus à l'aise avec les mathématiques qu'avec les matières littéraires, s'alarme pourtant de la prise de pouvoir des chiffres sur les mots dans les sciences naturelles et les exigences de réussite scolaire.

⁵ Première académie scientifique d'Europe, fondée à Rome en 1603 ; elle est à l'origine de la tradition des sociétés savantes occidentales.

⁶ Pour plus de détails sur cette synthèse historique, voir notamment BRYON-PORTET 2010 ; BRYSON 2011 et les biographies de wikipedia.

Le pouvoir des chiffres dans la perception actuelle du réel

Il semble donc bien qu'un processus historique ait favorisé la dominance de l'expression mathématique en lien étroit avec notre évolution technologique, pour accompagner la collecte exponentielle de connaissances et faciliter leur organisation. Le langage mathématique a en effet plusieurs avantages par rapport au langage littéraire : il est plus concis (combien de mots pour expliquer une équation ?), plus universel, et il subit moins de dérives (un symbole = un sens). Cette hiérarchisation cartésienne basée sur la prédominance des chiffres, entre disciplines ou entre méthodes et pratiques au sein d'une même discipline, pourrait ainsi être une caractéristique assez récente de la pensée humaine, dont la diffusion s'est accrue de manière spectaculaire ces dernières décennies grâce à la démocratisation de l'usage des ordinateurs.

Les tentatives d'expression de la biologie ou de l'écologie d'une espèce en termes d'équations mathématiques atteignent dans certains articles scientifiques un niveau qui éloigne le lecteur de la réalité observable (HIRZEL *et al.* 2002 ; MACKENZIE *et al.* 2002 ; RODHOUSE *et al.* 2011). Ces études, avec un peu de recul, peuvent être perçues comme une expression artistique, tel un tableau d'art contemporain déformant une source d'inspiration réelle jusqu'à un point tel que chacun y voit ce qu'il veut ou peut - pour les non-initiés un chaos intraduisible en termes concrets. Un nouveau code de langage réservé à une commu-

nauté restreinte... La seule question qui compte est de savoir si nous y gagnons dans notre approche explicative du réel, ou si l'outil mathématique devient peu à peu la véritable finalité intellectuelle ?

Un exemple actuel des préférences de la science pour la comptabilité, est celui de la taxonomie. Bien que de nombreux paléontologues signalent le caractère parfois fragile de la caractérisation de l'espèce dans un continuum d'évolution biologique, la classification des êtres vivants correspond à la fois à une réalité globale et à un moyen fonctionnel pour appréhender le vivant (LECOINTRE & LE GUYADER 2001). Comme le soulignait souvent Stephen Jay GOULD (1997b), l'évolution biologique est souvent confondue avec la notion trompeuse de « progrès », et l'évolution des concepts et des techniques n'échappe pas non plus à cette distorsion. Ainsi les généticiens, lorsqu'ils mettent leur discipline moderne en perspective avec des méthodes descriptives plus anciennes comme la comparaison des phénotypes, affichent souvent une certaine supériorité. Il est pourtant probable que leurs convictions actuelles seront regardées, par le jeu du recul temporel et de l'amélioration constante des connaissances, avec la même distance par leurs futurs collègues dans quelques années ou décennies. Cette façon de bénéficier des acquis antérieurs tout en les jugeant obsolètes semble inscrite dans beaucoup de nos comportements mêlant la connaissance et le pouvoir : ce qui est ancien est jugé archaïque, et doit être dénigré pour mieux asseoir le dogme actuel ; comme si une coexistence avec l'ancien empêchait l'affirmation du nouveau.

La réalité perçue à un instant donné dépend de nos capacités cognitives (KANT 2012) ; elle résulte d'une coévolution entre notre technologie et nos raisonnements intellectuels, donc d'un contexte temporel. Ce mouvement ne correspond pas à celui d'un curseur se déplaçant de l'ignorance vers un point de connaissance absolue : il paraît cohérent que ce point n'existe pas, étant lui-même en mouvement avec l'évolution du vivant et de son environnement sous dépendance de nombreuses règles mais aussi des contingences chères à Stephen Jay Gould. Ce dernier a souvent montré qu'il raisonnait comme un statisticien (voir notamment GOULD 1997b, p. 63), mais il était aussi un bibliophile, un érudit, amoureux des langues, des arts et de l'histoire des sciences, ce qui lui donnait une vision d'ensemble digne d'intérêt. Dans son ouvrage de synthèse sur l'évolution (GOULD 2006, p. 854), il aborde la théorie du gène égoïste (DAWKINS 1990), en considérant que placer le siège de la sélection naturelle au niveau des gènes plutôt qu'au niveau des organismes est une erreur de logique, une confusion entre causalité et comptabilité. Que ce soit Richard Dawkins ou Stephen Jay Gould qui ait raison sur ce sujet précis, cela importe moins que la nature de l'opposition telle que Gould l'interprète : plus globalement, un raisonnement purement comptable n'est peut-être pas le meilleur moyen - en tous cas pas le seul qui vaille - d'accéder à la causalité de faits observés.

L'empirisme cher à Francis Bacon (MILANESE 2016) a perdu ses lettres de noblesse face aux capacités de calcul des ordinateurs jugés sans faille car à l'abri de l'émotionnel, du subjectif ; pourtant les émotions du cerveau humain peuvent aussi améliorer la

qualité du raisonnement, et font partie intégrante de notre compréhension du monde (DAMASIO 1995).

La loi des probabilités : une composante influente de l'évolution biologique ?

Le Merle noir optimise la probabilité de rencontre avec des vers de terre en fouillant autour du point de capture récent (BROSSET 1982) ; le Lynx boréal établit une corrélation négative intuitive entre la distance de sa proie et ses chances de la capturer (RAYDELET 2006) ; l'enfant habitué au trafic citadin évalue la probabilité de traverser une rue sans risque, en fonction de la distance et de la vitesse des nombreuses voitures en mouvements croisés ; les cycles d'émergences des larves de cigales, en utilisant les propriétés des nombres premiers, minimisent la probabilité d'apparition de prédateurs spécialisés (GOULD 1997a) ; l'écologie du Douglas intègre la probabilité que la foudre tombe sur un tronc mort afin de permettre à ses graines de germer (OTTO 1998) ; etc.

Il est très *probable* que beaucoup d'individus dans leur comportement quotidien, et la plupart des espèces dans leurs traits de vie, utilisent les concepts de probabilité, de corrélations positives et négatives, pour améliorer leur survie. Dans tous les cas le phénomène procède par évaluations de plus en plus précises résultant d'un processus de sélection (expérimentale à l'échelle de l'apprentissage chez l'individu ; naturelle – au sens darwinien du terme – à l'échelle de l'évolution d'une espèce), aboutissant à une efficacité fonctionnelle remarquable.

Ainsi les probabilités font partie de nos vies sous forme de « calculs intuitifs », et conditionnent notre survie. Il paraît donc logique que le cerveau humain, doté de capacités cognitives au-dessus de la moyenne, ait quelques chances de ne pas être en reste dans ce domaine.

Notre propension à multiplier nos capacités grâce à la culture et la technologie s'est exprimée de manière particulièrement prégnante dans le domaine des mathématiques et des statistiques assistées par ordinateur, au point de déléguer à cette dernière discipline une part de nos capacités naturelles et qu'elle en vienne à régler nombre de nos comportements. On peut déplorer ses excès (sondages d'opinions qui inondent notre quotidien) ; on peut aussi s'alarmer des effets pervers de notre incapacité à produire par nous-mêmes certains raisonnements probabilistes (la surreprésentation des informations alarmistes dans les médias dérègle notre aptitude à relativiser et nous fait perdre la notion de potentialité des risques dans nos contextes individuels).

Qu'elle soit consciente ou non, la place prise par les probabilités dans nos vies n'a rien de nouveau. Mais elle s'exprime sous la forme de statistiques mathématiques depuis seulement quelques décennies dans nos sociétés technologiques, et ceci d'une manière quasi institutionnalisée qui imprègne nos codes comportementaux. Avons-nous intérêt à développer un esprit critique sur ce

phénomène ? La prise de conscience de son ampleur n'est-elle pas salutaire du point de vue de l'évolution de notre indépendance d'esprit collective – ce fameux libre arbitre individuel garant des libertés sociétales ?

Les statistiques informatiques : un produit humain objectif ou restant sous influence ?

Le traitement statistique informatique est un outil qui permet une lecture plus objective de faits observés, notamment en mettant en évidence leurs liens (influences positives ou négatives) avec des variables connexes. Mais l'on peut craindre qu'une focalisation exclusive sur cette discipline génère une interprétation paradoxalement distanciée de la réalité observable et vérifiée. Il est en effet une tendance récente constatée dans de nombreuses disciplines scientifiques : les chercheurs passent de moins en moins de temps sur le terrain⁷. En sciences naturelles, cette déconnexion sensorielle crée une distance avec l'objet étudié qui devient une entité de plus en plus théorique ; une régression à la fois temporelle et conceptuelle, de la logique d'observation d'Aristote vers le monde des idées de Platon.

⁷ Cette dérive actuelle a au moins un précédent dans l'Histoire : dans la Grèce Antique, la primauté accordée à l'esprit sur les tâches manuelles cantonnait les penseurs patriens à la réflexion aux dépens de l'expérimentation (GOULD 1997a p. 228) ; à ce titre Aristote (384-322 av. J.C.) a été en rupture avec ses prédécesseurs, notamment Socrate et Platon. De nos jours les raisons, si elles tiennent beaucoup à des contraintes triviales (manque de financements pour consacrer du temps au terrain), sont empreintes de la conviction d'une supériorité des modélisations et autres algorithmes informatiques et d'une méfiance vis-à-vis de l'observation directe.

En pratique cela peut être très dommageable à la qualité des données brutes ou à la puissance d'interprétation de l'observateur.

Les exemples sont à manier avec prudence car leur pouvoir de séduction sur notre compréhension risque de masquer la complexité d'un phénomène. C'est averti de cette dérive que je vais tenter d'illustrer mon propos avec précaution, en puisant dans une thématique qui m'est plus familière : le recours aux signaux sonar des chiroptères pour leur identification spécifique, l'étude du comportement des individus et celle de l'écologie des espèces.

Plusieurs études sur l'identification acoustique des chiroptères ont été réalisées en Europe ces vingt-cinq dernières années. Pour la plupart, leur propos est de privilégier une méthode de tri, de mesures et d'analyses automatiques afin de standardiser le processus et le rendre moins sensible aux variations liées à l'intervention humaine (collecte, sélection, mesure et interprétation « manuelles »). Certains de ces travaux expriment explicitement un dénigrement de l'analyse humaine au profit de l'algorithme informatique jugé moins biaisé. Or la gageure première dans ce domaine d'études me semble résider prioritairement dans la nature des données brutes : enregistrer de nombreux individus (pour couvrir la variabilité intraspécifique) dans de nombreuses conditions naturelles de vol (pour appréhender les comportements acoustiques sans biais dus aux méthodes de collecte des sons). Les applications en matière de conservation des espèces et de leurs habitats étant nombreuses, une émulation forte s'est développée dans la recherche universitaire, mais le temps disponible à consacrer au terrain étant réduit, la ma-

tière brute analysée souffre souvent de biais rédhibitoires. Certaines études se basent sur des enregistrements réalisés uniquement ou principalement dans les premiers mètres après le relâcher d'animaux capturés (donc sous effet de stress drastique) ou après leur sortie de gîte, avec dans tous les cas des signaux stéréotypés non représentatifs de l'ensemble du répertoire utilisé au cours de l'activité nocturne (RUSSO & JONES 2002 ; VAUGHAN *et al.* 1997). D'autres, grâce à l'utilisation du travail de naturalistes de terrain, évitent plus ou moins cet écueil mais ne prennent en compte dans un souci d'harmonisation statistique que quelques signaux isolés de leur contexte (MAC LEOD *et al.* 2013 ; WALTERS *et al.* 2012), sous-estimant la variabilité de l'outil sonar et occultant le lien comportemental entre les signaux. Une étude (JENNINGS *et al.* 2008) a même eu pour but de comparer les performances dans l'identification acoustique entre l'analyse automatique (traitements informatiques par réseaux de neurones artificiels) et l'analyse naturaliste (combinaison d'analyses auditive et informatique, cette dernière – sélection, traitement et mesures - étant réalisée manuellement) par des naturalistes plus ou moins expérimentés. Or ce sont les naturalistes qui ont dû se plier aux normes de l'analyse automatique et non l'inverse : ils ont dû travailler sur des courtes séquences de deux à trois signaux isolés de leur contexte, alors que la robustesse de leur méthode est au contraire basée sur la cohérence comportementale d'une série chronologique de plus de 10 signaux (correspondant aux conditions de récolte sur le terrain). Les performances de l'analyse informatique ont ainsi pu apparaître comme légèrement (mais non significativement) supérieures, alors que la méthodologie était biaisée.

Ces travaux ont tous été publiés dans des revues scientifiques à haut facteur d'impact ; l'appartenance des auteurs à des universités renommées et la caution des outils d'analyse statistique informatique doivent sans doute constituer de puissants éléments favorables pour les comités de lecture. Comme le précisent HURLBERT & LOMBARDI (2003), c'est souvent plus facile de publier en utilisant une analyse statistique erronée (38 % des publications examinées en sciences naturelles contenaient des erreurs statistiques conduisant à des conclusions erronées) qu'en n'utilisant aucune analyse statistique.

Autre exemple, celui d'une étude (CELUCH & ZAHN 2008) dont la citation est recommandée par des relecteurs de revues scientifiques. Elle compare, uniquement par le biais de modélisations informatiques, deux techniques d'étude de l'activité nocturne des chiroptères : le radio pistage et la détection acoustique. Démontrant que la seconde surestime l'activité dans des habitats de petite surface, les auteurs concluent que le détecteur d'ultrasons n'est pas un « outil magique ». Au-delà de la critique des résultats eux-mêmes, cette étude pose un problème de confusion dans les objectifs des méthodes comparées. Le radio pistage mesure le temps de présence d'un individu dans un habitat sans savoir ce qu'il y réalise (repos ? chasse ? activité sociale ?) ; la détection acoustique mesure la pression d'utilisation, ventilée par types d'activité, d'un habitat par un nombre inconnu d'individus. Ainsi un chiroptère glaneur chassant ou se reposant pendant 45 minutes dans un arbre sera analysé de la même façon par la technique du radio pistage, alors que la détection ne mettra

en évidence que l'activité de chasse, seule indicatrice de la qualité trophique de l'habitat. Ces seules différences induisent une absence de possibilité correcte de comparaison objective entre les deux techniques, puisqu'elles ne répondent pas aux mêmes questions. Par ailleurs, les auteurs font abstraction des erreurs fréquentes (par réverbération...) de triangulation entraînant une difficulté de localisation précise par radio pistage, problème qui affaiblit considérablement leur démonstration pour les habitats de petite surface. Il semblerait donc que le comité de relecture ayant validé ce travail ait été plus sensible aux aspects rationnel et novateur de la modélisation informatique, qu'aux failles méthodologiques que tout observateur de terrain aguerri ne saurait laisser passer.

Ces égarements montrent à quel point l'influence culturelle - souvent inconsciente - du chercheur scientifique, peut s'exercer à n'importe quel niveau de la conception ou de l'application d'un protocole (DE WAAL 2016, p. 246), même lorsque ce dernier se pare d'une objectivité mathématique.

Appartenir à la communauté scientifique implique d'adopter ses règles fondatrices mais aussi ses convictions du moment (par essence éphémères), qui peuvent résulter d'une volonté de se démarquer d'autres formes de pensée, parfois pour des motifs non rationnels : l'existence de domaines réservés montre l'impact des rapports sociaux humains sur l'objectivité des scientifiques comme des naturalistes (DESPRET 2012, p. 56).

Les études naturalistes : l'approche bayésienne intuitive a-t-elle sa place ?

L'évolution biologique, telle qu'on la conçoit depuis 1859 grâce à Charles DARWIN (1992), déroule un processus que l'on peut résumer de manière imagée comme un cumul d'expériences, chacune étant enrichie du patrimoine de la précédente. Les éléments de l'ensemble, des gènes aux communautés d'organismes, sont parfaitement fonctionnels [avec quelques imperfections de détail, puisque c'est du « bricolage par approches successives » (DAWKINS 1986)] à condition de rester interdépendants. Ce processus explique correctement, selon notre compréhension actuelle du vivant, la réalité observable.

Or nous pouvons malicieusement observer que ce processus cumulatif est assez similaire à la démarche d'un naturaliste de terrain qui tente d'expliquer avec méthode la réalité du vivant par des approches successives, dont chacune bénéficie des précédentes. Avec le temps il se forme progressivement une vision d'ensemble robustement étayée par cette succession d'enseignements confortés ou corrigés.

Une étude particulière, quelle que soit la qualité de son protocole et la richesse de son échantillonnage, n'offre qu'une vue ponctuelle et partielle de la réalité ; ce sont ces quelques données brutes, déconnectées de la continuité temporelle et de la variabilité spatiale, qui sont traitées dans une analyse statistique informatique (la

méthode des méta-analyses, intégrant les données de différentes études, est encore très peu pratiquée). La partie « commentaires » ou « discussion » d'une publication est certes censée, sous forme d'une comparaison narrative avec d'autres travaux, offrir une vision relative ; mais la norme des revues scientifiques sanctionne toute référence à des sources non publiées et non conformes à l'orthodoxie actuelle. D'où un parcours imposé auto-entretenu qui empêche l'intégration de nombreux acquis antérieurs venant améliorer la compréhension du système étudié ponctuellement.

Lorsqu'un chercheur, riche d'une longue expérience de terrain, interprète des résultats ponctuels, il développe – intentionnellement et/ou intuitivement – un processus d'analyse qui intègre l'ensemble de ses observations cumulées dans des contextes comparables. Dans ce processus les influences – positives ou négatives, supposées réelles ou biaisées – des variables explicatives sur les faits observés sont appréhendées à la lumière des nombreuses expériences antérieures, qu'elles relèvent d'études standardisées ou d'observations anecdotiques. Un naturaliste peut ainsi générer dans ses raisonnements une analyse intégrant et croisant une quantité importante d'observations et ne retenant au fil du temps que les hypothèses semblant les plus indépendantes des influences extérieures au système étudié. Lorsqu'il publie des résultats ponctuels sur une étude particulière, la façon dont il écrit son protocole puis interprète ses données est imprégnée de cette analyse globale cumulée avec

le temps ; il pratique ainsi une approche bayésienne⁸ plus ou moins intuitive.

L'expérience d'une personne est une propriété émergente de son équipement cognitif particulier. C'est donc un domaine réservé inaccessible au réductionnisme, dont la subjectivité ne doit pas être considérée comme un handicap mais comme un élément pertinent de la mosaïque que constitue la compréhension d'un phénomène. En écologie nous tentons de comprendre des interactions qui résultent de la nature du biotope et des traits de vie de chaque espèce, mais aussi certainement en partie des comportements individuels ; ces derniers nous sont parfois opaques au sein de notre propre espèce, et notre incompréhension ne peut que grandir vis-à-vis d'organismes plus éloignés comme un Murin de Bechstein, un Alyte accoucheur ou une Rosalie des Alpes (NAGEL 1983). La subjectivité est donc à l'origine même de nos perceptions, qu'elles se réalisent par nos sens ou à travers le filtre rassurant des calculs informatiques ; la sélection aléatoire d'un échantillonnage ou la suppression de variables parasites n'est souvent que le choix d'un nouveau biais pour en estomper d'autres. L'exclusion de telle ou telle forme de perception ou d'analyse me semble correspondre pour beaucoup à des courants historiques de la pensée humaine ; or un naturaliste imprégné en profondeur de son sujet développe un raisonnement qui est un complément utile aux autres formes d'analyses plus cartésiennes.

Des différences peuvent apparaître entre une analyse multivariée informatique (menée sur des résultats ponctuels) et une approche bayésienne intuitive (enrichie de sources multiples et variées) : par exemple une série de valeurs affichant une tendance non validée par les statistiques fréquentielles, mais soutenue comme une hypothèse robuste selon l'expérience de l'auteur. Cette contradiction n'est peut-être problématique qu'en apparence : elle constitue en elle-même une information pour mieux appréhender la globalité des questions posées : chaque source analytique, humaine *versus* informatique, est à la fois soumise à des biais et étayée par des éléments objectifs. Supprimer la première au profit de la seconde sous prétexte d'objectivité « comptable » est une perte d'informations dommageable lorsque l'auteur des travaux dispose d'une expérience suffisamment riche pour exercer ses capacités d'analyse multivariée humaine.

Le comportement actuel des revues scientifiques, se désintéressant de toute étude non étayée par l'analyse statistique informatique, correspond à une émergence culturelle récente : la science des statistiques, par sa puissance de traitement et d'interprétation des données grâce aux ordinateurs, a constitué une pression de sélection efficace sur le monde de la recherche (mais aussi celui de la communication) toujours à la recherche de l'objectivité mathématique. Il s'ensuit un dénigrement des capacités d'intégration et de réflexion humaines

⁸ Thomas Bayes, pasteur britannique, est à l'origine d'une formule simple publiée en 1763 et dénommée depuis le théorème de Bayes. Cette loi des probabilités, oubliée pendant plus de deux siècles, fait l'objet d'une redécouverte récente et fertile dans de nombreuses disciplines qui l'intègrent dans des processus statistiques souvent très complexes. L'approche bayésienne prend en compte les connaissances antérieures (expérimentales ou anecdotiques) sur un système donné, et les associe aux données de terrain qui sont acquises au cours d'une étude ponctuelle ; cette intégration informe souvent mieux et davantage la compréhension du système et peut aussi s'exercer de manière intuitive (GUNNELL 2009).

au profit de la « réflexion robotisée », jugée plus performante (à juste titre pour la précision du calcul, mais elle n'intègre pas le cumul d'expérience) et plus objective (alors qu'elle dépend aussi de composantes humaines : les choix et les capacités de leurs concepteurs et de leurs utilisateurs). Il se pourrait pourtant que la réalité soit mieux appréhendée grâce à plusieurs angles de perception distincts et complémentaires.

Lorsqu'un naturaliste publie des résultats, il présente souvent des tendances jugées pertinentes selon son approche bayésienne. Par exemple pour les chiroptères : la décroissance des effectifs d'une espèce dans un gîte de repos subissant une perturbation, ou l'augmentation de l'activité de chasse d'une espèce en lien avec la complexité structurelle d'un peuplement forestier. Et cela même si certaines sont qualifiées de « non significatives » par une analyse statistique informatique. Cette démarche étant refusée par les revues scientifiques, cela oblige les auteurs - s'ils persistent à publier - à se restreindre aux seuls résultats validés statistiquement. Cette présentation partielle des résultats entraîne une perte d'informations, alors que ces dernières pourraient être exprimées sous forme d'hypothèses à valider ou infirmer par des travaux ultérieurs. En effet, les études écologiques notamment, cumulent plusieurs niveaux de complexité (BLONDEL 2012, p. 76) : celle interne aux organismes étudiés et celle qui concerne les relations multiples au sein d'un système changeant (entre congénères, entre espèces, entre espèces et leur contexte abiotique...). Plus la complexité augmente, plus sa vérification par les mathématiques devient ardue. Richard DAWKINS (1986, p. 17) illustre cette

difficulté par la métaphore suivante : « Le comportement des objets physiques, non biologiques, est si simple qu'il est possible d'utiliser pour le décrire le langage mathématique existant - voilà pourquoi les livres de physique sont remplis de mathématiques. [...] Personne n'a encore inventé les mathématiques qui pourraient décrire en totalité la structure et le comportement d'un objet tel qu'un physicien, ou même une seule de ses cellules. ». Il est en effet particulièrement difficile, pour des études écologiques complexes, de prévoir dans les protocoles une mesure de l'influence de tous les paramètres biologiques, comportementaux, environnementaux... autant de facteurs qui peuvent être progressivement intégrés par un observateur tout au long d'un processus empirique, et dont les interprétations subjectives sont censées diminuer avec l'augmentation dans le temps des qualités d'observation et d'analyse (l'honnêteté intellectuelle étant par ailleurs considérée comme une valeur attendue constante).

Vers une meilleure tolérance envers les différents outils d'analyse ?

Bien sûr il existe quelques rares revues naturalistes qui ont pour but de publier des résultats originaux, et dont les comités de relecture restent à peu près ouverts aux approches empiriques. Mais elles ont tendance avec le temps à suivre deux processus classiques : soit elles manquent de matière de publication (malgré les nombreux et excellents travaux naturalistes réalisés chaque

année, mais beaucoup de leurs auteurs sont découragés par le formatage actuel) et finissent par disparaître, soit elles montent en renommée et versent rapidement dans les exigences dogmatiques pour se rapprocher de la « cour des grandes ».

Serait-il concevable que l'ensemble de la communauté scientifique puisse intégrer dans son concept de description de la réalité observable, le fait qu'un humain est capable dans de nombreux cas d'accéder à un niveau pertinent d'analyses de type statistique, sans pour autant avoir recours aux calculs mathématiques, à partir de ses connaissances acquises grâce à un

processus cumulatif? Ainsi le recours à des analyses informatiques fréquentielles, s'il constitue un atout indéniable (et même indispensable pour les très gros volumes de données), ne constituerait pas pour autant un passage obligé pour publier des résultats. A chaque lecteur de mettre en oeuvre son libre arbitre (élément pivot d'une démocratie, qui nécessite culture et méthode et s'atrophie lorsque la réflexion individuelle est déléguée – notamment – à la technologie) et de juger de la pertinence des travaux présentés, qu'ils aient recours ou non à l'analyse statistique informatique.



Remerciements

Merci à Philippe Grimonprez pour sa relecture avisée du manuscrit et pour son rôle permanent de transmetteur de culture.

Merci à tous ces auteurs d'essais qui excellent dans l'exercice difficile de la vulgarisation des connaissances ; Stephen Jay Gould (1941-2002), référence à la fois scientifique, humaniste et littéraire, est une source permanente d'éveil et d'enrichissement de la pensée.

Bibliographie

BLONDEL, J. 2012. *L'archipel de la vie. Essai sur la diversité biologique et une éthique de sa pratique.* Buchet Chastel, Paris. 256 pp.

BROSSET, A. 1982. La prédation. *Bulletin mensuel de l'Office National de la Chasse* 54 : 14-25.

BRYON-PORTET, C. 2010. Sciences humaines, sciences exactes. Antinomie ou complémentarité ? *Communication* 28(1) : 243-264.

BRYSON, C. 2011. *Une histoire de tout, ou presque...* Payot, Paris. 654 pp.

CELUCH, M. & ZAHN A. 2008. Foraging habitat preferences of bats: new question in interpretation of bat detector data. *Vespertilio* 12: 3-9.

CRAMPON, A. 1864. *La sainte bible.* Traduction d'après les textes originaux par l'abbé A. Crampon, Chanoine d'Amiens. Édition révisée par les Pères de la Cie de Jésus avec la collaboration de Professeurs de S. Sulpice. Société de Saint Jean l'évangéliste, 1923.

DAMASIO, A.R. 1995. *L'erreur de Descartes. La raison des émotions.* Odile Jacob, Paris. 386 pp.

DARWIN, C. 1992. *L'origine des espèces au moyen de la sélection naturelle ou la préservation des races favorisées dans la lutte pour la vie.* Texte établi par Daniel Becquemeont à partir de la traduction de l'anglais d'Edmond Barbier. Flammarion, Paris. 604 pp.

DAWKINS, R. 1986. *L'horloger aveugle.* Robert Laffont, Paris. 380 pp.

DAWKINS, R. 1990. *Le gène égoïste.* Odile Jacob, Paris. 459 pp.

DESPRET, V. 2012. *Que diraient les animaux si... on leur posait les bonnes questions ?* La Découverte, Paris. 325 pp.

DE WAAL, F. 2016. *Sommes-nous trop « bêtes » pour comprendre l'intelligence des animaux ?* Les Liens qui Libèrent, Paris. 408 pp.

DIAMOND, J. 1997. *De l'inégalité parmi les sociétés. Essai sur l'homme et l'environnement dans l'Histoire.* Gallimard, Paris. 695 pp.

GOULD, S.J. 1997a. *Darwin et les grandes énigmes de la vie.* Seuil, Points Sciences, Paris. 308 pp.

GOULD, S.J. 1997b. *L'éventail du vivant. Le mythe du progrès.* Seuil, Points Sciences, Paris. 300 pp.

GOULD, S.J. 2006. *La structure de la théorie de l'évolution.* Gallimard, Paris. 2033 pp.

GUNNELL, Y. 2009. *Ecologie et Société.* Collection U Sciences Humaines et Sociales. Armand Colin, Paris. 432 pp.

HURLBERT, S.H., & C.M. LOMBARDI. 2003. Design and analysis: uncertain intent, uncertain result [review of *Experimental design and data analysis for biologists* by G. Quinn and M. Keough]. *Ecology* 83: 810-812.

HIRZEL, A.H., HAUSER J., HESSEL D. & PERRIN N. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology* 83(7): 2027-2036.

JENNINGS, N., PARSONS S. & POCOCK M.J.O. 2008. Human vs. machine : identification of bat species from their echolocation calls by humans and by artificial neural network.

Canadian Journal of Zoology 86: 371-377.

KANT, E. 2012. *Critique de la raison pure*. Traduction de l'ouvrage publié en 1781. Presses Universitaires de France, Paris. 584 pp.

LECOINTRE, G. & LE GUYADER H. 2001. *Classification phylogénétique du vivant*. 2e édition. Belin, Paris. 543 pp.

MACKENZIE, D.I., NICHOLS J.D., LACHMAN B.B., DROEGE S., ROYLE J.A. & LANGTIMM C.A. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83(8): 2248-2255.

MAC LEOD, N., KRIEGER J. & JONES K. 2013. Geometric morphometric approaches to acoustic signal analysis in Mammalian biology. *Hystrix, the Italian Journal of Mammalogy* 24(1): 110-125.

MILANESE, A. 2016. *Bacon et le gouvernement du savoir. Critique, invention, système: la pensée moderne comme épreuve de l'histoire*. Classiques Garnier. Coll. Constitution de la modernité. 587 pp.

NAGEL, T. 1983. Quel effet cela fait, d'être une chauve-souris ? In : *Questions mortelles*. Presses Universitaires de France, Paris. 247 pp.

OTTO, H. J. 1998. *Ecologie forestière*. Institut pour le Développement Forestier, Paris. 397 pp.

RAYDELET, P. 2006. *Le lynx boréal*. De-lachaux & Niestlé, Paris. 191 pp.

RODHOUSE, T.J., VIERLING K.T. & IRVINE K.M. 2011. A practical sampling design for acoustic surveys of bats. *The Journal of Wildlife Management* 75(5): 1094-1102.

ROSTAND, J. 1963. *Le droit d'être naturaliste*. Stock, Paris. 213 pp.

Russo, D. & JONES G. 2002. Identification of twenty-two bat species (Mamma-

lia: Chiroptera) from Italy by analysis of time-expanded recordings of echolocation calls. *Journal of Zoology of London* 258: 91-103.

UNESCO. 2010. Rapport de l'UNESCO sur la science 2010. Résumé exécutif. L'état actuel de la science dans le monde. Rapport 31 pp. <http://unesdoc.unesco.org/images/0018/001898/189883F.pdf>

VERNIER, P. 2005. Evolution du cerveau et émergence du langage. p. 42-63. In : *Aux origines des langues et du langage*. Sous la direction de Jean-Marie Hombert. Fayard, Paris. 505 pp.

VAUGHAN, N., JONES G. & HARRIS S. 1997. Identification of British bat species by multivariate analysis of echolocation call parameters. *Bioacoustics* 7: 189-207.

WALTERS, C.L., FREEMAN R., COLLEN A., DIETZ C., FENTON M.B., JONES G., OBRIST M.K., PUECHMAILLE S.J., SATTLER T., SIEMERS B.M., PARSONS S. & JONES K.E. 2012. A continental-scale tool for acoustic identification of European bats. *Journal of Applied Ecology* 49: 1064-1074.

Pour citer cet article :

BARATAUD, M. 2017.

La publication scientifique est-elle encore à la portée des naturalistes ? Réflexions sur l'importance accordée aux analyses statistiques informatiques. *Plume de Naturalistes* 1 : 113-128.

ISSN 2607-0510

Pour télécharger tous les articles de Plume de Naturalistes : www.plume-de-naturalistes.fr