Title: Do we need all of this research synthesis in ecology and conservation? A highly subjective criticism of reviews and meta-analyses

Author: Jacopo Cerri, Mammal Research Institute, Polish Academy of Sciences, Stoczek 1D, 17-230, Bialowieza, Poland. email: j.cerri@ibs.bialowieza.pl

Abstract: Research synthesis, which includes systematic reviews and meta-analyses, provides applied ecologists and conservationists with big pictures of ecological processes and the effectiveness of conservation interventions. It can therefore hugely benefit applied ecology and conservation science.

However, while in some areas of ecology and the conservation sciences study design, measurements and statistical analyses have been standardized, in many fast-evolving fields of research they change relentlessly. This raises the question on the extent to which studies, which are inherently different in their design, methods and analyses can be compared. Even worse: is pooling together these studies actually misleading?

I believe four sources of heterogeneity to be particularly troublesome for research synthesis in many areas of ecology and conservation: *i*) the lack of well-agreed measures, *ii*) nonprobability sampling, *iii*) scarce background information about observational units and *iv*) limited control over treatment.

Considering the ongoing increase in research synthesis in ecology and conservation, there is a need to discuss these issues and figure out how to mitigate them. Incentives in favor of publishing research synthesis are too high to hope that scientists will self-regulate. Instead, I believe that research synthesis should: *i*) be published only after some years and a second peer-review, *ii*) consider only studies whose data can be re-analyzed in the future, *iii*) be periodically assessed and retracted, if deemed to be obsolete.

Keywords: evidence-based conservation; applied ecology; wildlife management; biodiversity conservation; researchers degrees of freedom

Disclaimer

This opinion piece is about my personal beliefs about research synthesis, or the practice of gathering and assessing scientific evidence, through systematic reviews and meta-analyses. The more reviews and meta-analyses I read in ecology and conservation, the more I believe them to be potentially more valuable than single studies, yet also potentially more misleading.

In this opinion piece I explicitly mention some reviews and meta-analyses. They were mentioned because I deem them good examples to explain some problems: I did not mean to criticize authors, who wrote them with the best intentions, and who are often better scientists than I will ever be. This opinion piece is also heavily based on my own professional experience, which is limited to the ecology and conservation of mammals and birds, as well as to the conservation social sciences. Finally I also do not mean to throw the baby out with the bathwater: many reviews and meta-analysis are actually methodologically sound and important.

I believe it is too early to submit this piece to a scientific journal. Moreover, I firmly believe that scientists should be wary of publishing opinion pieces, because this practice converts them into pundits. For the moment I will keep it archived as a preprint and readers are encouraged to provide feedback, through the Comments section of EcoEvoRxiv.

The rise of research synthesis and its drivers

Research synthesis is the process through which results from multiple studies are centralized and analyzed, either qualitatively or quantitatively, with the ultimate goal of summarizing broad patterns of interests. Whenever this process is qualitative, research synthesis takes the form of systematic reviews, while a quantitative re-analysis of pre-existing studies leads to a meta-analysis (Gurevitch and Nakagawa, 2015).

By centralizing a high number of pre-existing studies, research synthesis allows ecologists and conservation scientists and practitioners (hereinafter referred to as "conservation scientists") to draw large-scale pictures of socio-economic, biological and environmental processes, as well as about the effectiveness of conservation interventions. Going beyond single case studies is valuable in three main ways. First, it summarizes what has been studied about certain taxa and processes, or which conservation interventions have been implemented so far. This process is crucial to highlight knowledge gaps and guide future research. Moreover, this process allows testing fundamental hypotheses in many fields related to conservation sciences, ultimately improving the knowledge basis sustaining conservation biology. Finally research synthesis allows conservation scientists to quantify the effectiveness of practical interventions (Sutherland et al., 2019), a practice which informs policymakers and stakeholders, and which paves the way for improving the design, implementation and evaluation of future conservation interventions.

Research synthesis has therefore become, as for many other applied sciences, one of the cornerstones of evidence-based conservation and a very important component of its related fields. Nowadays there are journals, such as Biological Review and Mammal Reviews which specifically focus on reviews and meta-analyses. Moreover, most peer-reviewed journals also accept reviews and meta-analyses as a distinct type of article. The number of published reviews and meta-analyses has therefore grown relentlessly in ecology and the conservation sciences (Boice, 2019), with these being more disproportionally cited (Pollo et al., 2025) and highly mentioned (Anderson et al., 2021) in scientific publications. Nowadays it is also increasingly common to include systematic reviews already at the beginning of many PhD curricula, as they are certainly useful to familiarize doctoral students with a specific sub-field of research.

Despite significant advances in methods for literature centralization (Foo et al., 2021; O’Dea et al., 2021) and data analysis (Nakagawa et al., 2023), conducting research synthesis remains problematic for some areas of the conservation science and related fields, particularly those based on observational data and what I have called “post-hoc” science.

“Ad-hoc” versus “post-hoc” science and its impact on standards and re-analysis

In the conventional view of science, exemplified by the PPDAC cycel, data analysis is a process where scientific questions are converted into testable hypotheses, related to specific parameters of a statistical model. Data is then collected on purpose, based on a sampling scheme which is deemed suitable for the estimation of these parameters. This is the case of many randomized experiments: there is a certain scientific question and its corresponding hypotheses, then measurements are collected from observations that are randomly assigned to different controlled conditions, with the explicit scope of testing these hypotheses. This is also the case of surveys, where the question is to estimate a parameter associated with a biological population (number of individuals, density, occupancy etc…) and then a sampling scheme is designed to monitor the study area.

I like to name this way of conducting science “ad-hoc”, because data is collected proactively, with the explicit goal of answering a certain research question, according to the PPDAC cycle. Ad-hoc data collection has the big advantage that people think in advance about what they want to measure, so it encourages standardization in measurements and study design. The use of probabilistic sampling and randomized treatment allocation also have the important advantage that quantities such as p-values become interpretable (Kruschke and Lidell, 2018) and assumptions of hierarchical models are respected (Silk et al., 2020). In some fields, where observing and manipulating observations is relatively easy, this allowed incremental progress and the standardization of data collection, which in turn made studies ready for research synthesis.

In some other fields, the co-evolution of study planning, data collection and data analysis followed a different trajectory. For example, studying the behavior and ecology of many animal species is cumbersome because they are inherently hard to observe, maybe as they move across large areas, are elusive, live far away from humans or are nocturnal. Similarly, studying human opinions about nature and conservation is tricky, because designing good surveys is often incompatible with available resources, or because people increasingly refuse to answer. This has produced two specific trends, in the conservation sciences and related fields.

The first one is the use of technological devices that automatically record the environment and/or the behavior of animals and people, such as satellites, camera traps, acoustic recorders, or biologgers. The second one is the use of “opportunistic data”, or those data that we manage to collect, under the assumption that any data is better than no data. These include citizen science observations, or interviews/questionnaires collected from specific segments of our population.

These two phenomena jointly led to the creation of vast amounts of data (Farley et al., 2018), collected sometimes without having identified all of their future applications in advance, but rather with a certain decoupling between data collection and data analysis.

In turn, this led to the explosion of what I called "post-hoc" science, where research questions are defined once data have already been collected, and scientists test if these questions can be answered with the data at hand. This process often entails compensating for deficiencies in data collection through sophisticated statistical modeling. Importantly, this practice should not be conflated with "HARking", a debatable practice where hypotheses are established only once results are known (Kerr, 1998).

There is nothing intrinsically wrong with "post-hoc" science, which in many scientific fields is the norm and which can produce valuable studies. However I believe its rapid evolution, coupled with incentives from the "publish-or-perish" system, to produce studies that are not immediately re-usable for research synthesis. This is because of four main reasons, which will be explained in the following sections.

Lack of well-agreed measurements

While in "ad-hoc" science, researchers have strong incentives to define valid, reliable and accurate measures before investing time and money in data collection, by decoupling data collection from data analysis, post-hoc science does not provide the same incentives. Although many large collaborative research initiatives are strongly pushing for data standardization, by not knowing in advance how you are planning to analyze your data and being rewarded only for publishing, data users avoid thinking accurately about measurements.

A good example of this is provided by reviews summarizing space use by animals, quantified through satellite or VHF telemetry. In movement ecology, space use has been traditionally associated with the concept of home range, which in its original formulation is defined as "*the area traversed by the individual in its normal activities of food gathering, mating, and caring for young*" (Burt, 1943). Several estimators have been developed to quantify home ranges, however in the last few years a growing emphasis has been placed on the need to distinguish between estimators for the range and the occurrence distribution, deeming the first one to be the most suitable approximation for home range (Alston et al., 2026). Moreover, even when considering only the range distribution, different methods provide different estimates (Noonan et al., 2019). Regardless of this well-known heterogeneity, some reviews and meta-analyses regularly pool together studies using different estimators for the range distribution (Grüner Ofstad et al., 2016), sometimes by using regression models to increase their comparability (Rodríguez-Recio et al., 2022), or studies using estimators for the range and the occurrence distribution (Broekman et al., 2022; Seigle-Ferrand, 2021).

The problem of standardizing measurements is not only conceptual, but also deals with measurement accuracy. This second issue, in my opinion, is exemplified by studies using land cover maps. The development of remote sensing has provided ecologists and conservationists with increasingly accurate maps of vegetation (Pettorelli et al., 2014) and land cover. As for land cover, nowadays, many land cover maps are freely available, with a resolution of approx. 10 m and considerable accuracy with respect to real-world conditions (Venter et al., 2022; Xu et al., 2024). However, until 10-15 years ago, studies in ecology and conservation relied on maps with a much lower resolution and accuracy (Grekousis et al., 2015). Using land cover maps with a different resolution, for example, can produce very different estimates of habitat selection by mammals (Salvatori et al., 2022). It is therefore unclear the extent to which studies based on land cover maps with a different resolution can truly be compared. This problem can affect, for example, reviews summarizing the effect of landscape attributes over the occurrence of wildlife-vehicle collisions (Pagany et al., 2020).

The problem of measurement quality is not exclusive to movement or road ecology. For example, many studies also compared density estimates of wildlife species from multiple studies, but by pooling estimates obtained with different methods, whose comparability even after adjustments is debatable (for ungulates see van Beeck Calkoen et al., 2023 and Melis et al., 2009). In the field of "conservation culturomics" many studies regularly use Internet search volumes to quantify long-term changes in human interest towards nature (Ladle et al., 2016), without any benchmarking against well-established methods such as questionnaire-based surveys.

Nonprobability sampling

Another problem characterizing many reviews based on observational studies lies in the scarce control these usually have over sampling. Sampling is the process through which a subset of units, the sample, is selected from a population, with the goal of better understanding its properties. Statistical inference aims to draw conclusions about the entire population, by generalizing information contained in the sample. This becomes possible when units are collected with a probabilistic sampling, where their probability of being extracted is known. Probabilistic sampling is necessary to create design-based estimators and to draw generalizable conclusions from model-based estimators (Gregoire, 1998; Dumelle et al., 2022; Williams and Brown, 2019).

However, an increasing number of studies are implementing nonprobability sampling (Boyd et al., 2023). In nonprobability sampling, researchers have no control over the process through which they extract units, nor any idea about their probability of being extracted. In most cases, nonprobability samples arise from practical issues connected with data collection (e.g. costs, methods), as well as with difficulties associated with populations without a list, for which nevertheless specific sampling strategies are available (Hankin et al., 2019).

From my experience I can tell that nonprobability sampling is widespread in the conservation social sciences, particularly in studies administering questionnaires. For example, estimating what residents of a certain state think about the presence of wolves (*Canis lupus*) and their management requires well-designed questionnaires and an expensive survey, whose cost often exceeds funding available to research groups or local authorities. Even worse, sometimes researchers might want to quantify what visitors of a certain national park think about wolves, despite not having any indication on how many people visit the park each year, nor about their characteristics. In these situations conservation scientists typically go to the field and administer questionnaires to members of local communities or some specific categories through snowballing. Or simply spend a week on the field, trying to administer as many questionnaires as possible to the people they encounter.

I do not want to criticize nonprobability sampling, which could be valuable to collect preliminary data and which I also adopted extensively in the past. However, whenever this practice becomes the norm, summarizing findings from multiple studies becomes problematic, or even worse, misleading.

For example, Dressel et al. (2015) carried out a meta-analysis to summarise attitudes of European citizens towards bears and wolves, with the goal of understanding temporal patterns and differences between genders and social groups. Yet, half of studies in their sample were based on nonprobability sampling. Findings from these studies cannot be compared quantitatively, not even by using sophisticated statistical models. Because there is no guarantee that these samples are truly comparable: some samples could have had more farmers than others, while other samples could have had more unknown sources of variation. One of the scopes of random sampling lies in its robustness against unobserved sources of heterogeneity between units. In another large-scale review (Barmoen et al., 2024) authors acknowledged that nonprobability sampling of studies could limit the validity of their findings, yet they did not retain only studies based on probabilistic sampling. This issue also probably affected other reviews on this topic, which did not consider probabilistic sampling as an inclusion criteria, nor mentioned potential flaws associated with nonprobability sampling (Franchini et al., 2021; Newsom et al., 2026).

Nonprobability sampling is also widespread in camera trapping studies, as well as in studies using citizen science data about species distribution or wildlife vehicle collisions. Citizen science data are collected mostly by volunteers, often corresponding to biodiversity records that people record on their smartphone or tag on social networks. This data is characterized by persistent issues in spatial coverage, as human observers usually stick to the most accessible areas (e.g. around roads and trails, Dimson et al., 2023). Although advanced modeling allows to researchers to adjust for preferential sampling in the most accessible areas (Sicacha-Parada et al., 2021), many studies using citizen science data do not engage in similar adjustments. As a result, opportunistic data from nonprobability sampling yield very different estimates of species distribution and activity, as well as about their relationship with environmental characteristics, than data obtained from well-designed probabilistic surveys. This point has been emphasized by few studies using opportunistic records of wildlife-vehicle collisions (Cerri et al., 2025), as well as by studies comparing camera trapping data obtained from different designs (Greco et al., 2025; Sollmann et al., 2025). However, it is important to realize that comparisons between estimates from probability and nonprobability samples are almost inexistent in ecology and conservation. As a result, pooling together studies using nonprobability samples might result into biased estimates of ecological parameters (e.g. trapping rates in camera-trapping surveys; e.g., density of wildlife-vehicle collisions) as well as in biased estimates of associations between species and environmental characteristics.

Scarce background information about observational units

In both experimental and observational studies, to assess the effect of a certain treatment over a a sample of units, researchers must be sure that units in the control and the treatment group are either reasonably similar, or that their sources of heterogeneity can be measured and properly accounted for, through interactions or random effects in regression models.

A good example of this is provided by measures to prevent or mitigate large carnivore predation on domestic animals, a potential driver of large carnivore persecution. Assess how well and under which circumstances different measures work is important, especially for lethal control, which is increasingly politicized (e.g. in Europe, Kutal et al., 2025), but whose effectiveness has to be carefully matched against its negative demographic impacts. To address this gap, in the last few years several reviews meta-analyses summarized the effectiveness of measures to prevent or reduce livestock predation by large carnivores. Citing all of them is beyond the scope of this study, however these either analyzed a very low number of studies, due to strict inclusion criteria or sometimes pooled together observational studies where, in my opinion, comparability between units is non-trivial.

A common problem of anti-predatory measures is that in many cases herders implementing them might be systematically different from herders who do not. For example, prevention measures, such as electric fencing or night recoveries might not be paid in advance by local authorities, but rather reimbursed. This could restrict their implementation by herders without enough liquidity to anticipate their cost (Berzi et al., 2021). Therefore, when comparing measures of predation, between herders implementing measures and herders that do not implement them, part of the difference between the two groups might depend upon unobserved differences between the two groups of herders, acting as a confounding factor. For example, poor herders, who struggle to implement anti-predation measures, might also suffer from increased predation rates as they could be located in more marginal agricultural areas, with a higher presence of large carnivores. Alternatively, as they have less revenues from herding, they might need to engage in part-time jobs, therefore allocating less time to guard and/or recover livestock.

From a statistical viewpoint this boils down to one, or more, confounding variables affecting both treatment allocation and differences in the outcome. Confounding can be addressed either by allocating measures at random, or by carefully controlling for differences between groups (e.g. through the “back-door-criterion”, see Pearl and MacKenzie, 2018). In the case or randomized experiments, background information is required to devise a blocked experiment design, while in the case of observational studies, to include relevant confounding variables in statistical models. In both cases, researchers must possess adequate background knowledge about the socio-economic conditions and working habits of herders, which can be collected only through in-depth interviews and the design of questionnaire-based surveys with farmers.

A slightly different problem, related to heterogeneity between units, concerns the differential impact of treatment between different levels of one of their attributes. For example, let’s assume that frightening measures (Lorand et al., 2022) are tested in terms of their effectiveness to reduce wolf presence around sheep farms. It is plausible to hypothesize that the effect of these devices can depend upon the presence of undisposed animal leftovers around farms, which are well known attract wolves (e.g., Nowak et al., 2026). Once scared by the alarm, wolves might decide to avoid farms with little or no leftovers, because they have no incentive for doing so. But at farms with abundant food waste, the impact of scaring devices might be lower, because the benefits deriving from additional food can outweigh perceived risks. In this case, the impact of acoustic scaring devices is highly context-dependent and this significant heterogeneity should be accounted for, by building a statistical model with an interaction term. Ignoring this interactive effect leads to underestimating the effectiveness of acoustic scaring devices and ignoring where their application could be the most effective. The main problem is that accounting for this synergy requires to know which farms leave animal leftovers undisposed, an aspect that again, is not easy to monitor for wildlife managers, as in many countries this practice constitutes a violation of sanitary regulations (for an example about BSE in the Iberian Peninsula, Mateo-Tomás et al., 2022).

Limited control over treatment

In ecology and conservation biology, the importance of manipulative studies is undeniable: by manipulating observations, through a proper experimental design, researchers can identify cause-effect relationships (Pearl and McKenzie, 2018) and mediation mechanisms (Correia et al., 2025). However, while many researchers understand the importance of actively assigning units to one or more treatments, fewer know that assessing the effect of a certain treatment over a quantity of interest requires *i*) no discrepancies between being assigned to a certain treatment and effectively receiving it and *ii*) that treatment does not propagate to observations in the control group.

A first problem lies in the fact that researchers usually have imperfect control over treatment. While some units might be labeled as those which will receive a treatment or not, being therefore "assigned" to the treatment or the control group, there could be imperfect compliance or some confounding factors that produce a discrepancy between being assigned to a certain condition and effectively receiving it. In clinical trials, this is the reason why blinding and placebos are common: some patients in the treatment group might refuse to take a certain drug, why patients in the control group might refuse not to be treated and seek alternative treatements (Chickering and Pearl, 1996).

As in the previous example about comparability between units, similar problems might characterize the use of measures to reduce predation by large carnivores on livestock. Let's hypothesize that, in an area with severe levels of wolf predation on sheep, local authorities decide to run a pilot study to assess the effectiveness of guarding dogs. They therefore assign shepherds to one of two groups: a treatment group, that receives guarding dogs and food for free, and a control group of comparable farmers that does not. However, after a few weeks, guarding dogs turns out to be aggressive towards hikers, therefore, to avoid legal controversies some herders in the treatment group therefore decide to keep dogs in kennels rather than on pastures around sheep, without notifying local authorities. Therefore, the treatment (guarding dogs) is assigned, but it does not manage to truly affect some units (herds). The opposite situation could also happen, for example herders in the control group might decide to implement alternative prevention measures on their own (e.g. by using electric fencing, or by increasing vigilance during grazing).

This creates a mislabeling of units, which affects the outcome of statistical modeling, whenever authorities only compare before-and-after differences in predation rates between the treatment and control group of herders, without accounting for effective levels of compliance. Of course, it is not impossible to quantify the real magnitude of this discrepancy, but this requires ad-hoc surveys which are not be always implemented and in many studies there is not any quality control, explicitly mentioned in the methods.

A second problem, connected to the real level of control attained over treatment, is the fact that its effects should not be transferred to units belonging to the control group. For example, in the hypothetical case mentioned above, it might happen that guarding dogs deployed at a certain sheep farm displace a wolf pack that lived around it. This will decrease predation on livestock at the farm which has received guarding dogs, but also at neighboring farms in the control group.

Again, this creates a situation where observations are mislabeled and where group comparisons are biased. In our hypothetical example, obtaining adequate information on wolves being displaced would require researchers to implement an intensive monitoring scheme to quantify patterns of wolf presence around farms. Absent this detailed information, whenever this process has a clear spatial pattern, like in our example, researchers can test for its occurrence through statistical models that explicitly account for spatial correlation between units (Lamoroux et al., 2025). However, this might not always be the case, as the propagation of treatment effects to units in the control group might follow complex causal mechanisms, which are not necessarily well represented by spatial random effects (Gilbert et al., 2024).

Structural changes to improve the quality of research synthesis

From the previous sections I highlighted four problems that characterize observational studies and can negatively affect research synthesis. It is important to note that these problems are not necessarily solvable through advanced statistical modeling or data standardization, despite the undoubted benefits of these two practices, as they are purely qualitative.

I also firmly believe that these four problems have arisen by a perfect storm: the development of large networks that seamlessly collect data, an unprecedented availability of data scientists and user-friendly software for statistical analysis, as well as the publish-or-perish system, which strongly rewards research synthesis. While an increasing number of manifestos have praised the adoption of voluntary codes of conduct by researchers I believe them to be ultimately ineffective: if we really want to mitigate the rise of misleading research synthesis we need rules acting on the process of scientific publishing, explicitly counteracting the distorsions introduced by the publish-or-perish system. The rules that I believe in, are based on the idea of introducing calibrated friction into the publishing processes (or “sludge” see Sunstein, 2021) and on maintaining a periodic scrutiny on the content of reviews.

The first solution is simple to explain: it should be harder and more time-consuming to publish reviews and meta-analyses. All research synthesis should be based on pre-registration (Nosek et al., 2018), with its methods being carefully evaluated by peer scientists in advance. Then reviews/meta-analyses should be carried out and not be disclosed for 2-3 years. After this embargo, they should be subjected to peer-review: if their methods have not yet become obsolete, these can be considered suitable for publication. This practice will transform publishing research synthesis into a major effort, which will arguably discourage people looking for quick results and not interested in methodological soundness. Given the current proliferation of short-term research positions, this practice will probably discourage many people from engaging in peer-reviews at all. However, I firmly believe that we could easily renounce the vast majority of the reviews and meta-analyses that are currently being produced, if this improves the quality of the remaining ones.

Meta-analyses should also be based solely on studies whose data can be re-analyzed in the future, according to the most up-to-date standards and procedures. Only studies providing completely reproducible raw data will therefore be retained. For example, in the case of movement ecology, it would allow to update pre-existing meta-analyses about space use or habitat use, by shifting from old methods (e.g. MCP, KDE, resource selection functions) to most advanced ones (e.g., aKDE; e.g., step selection analysis). This would therefore transform meta-analyses into an open-process, whose findings can be periodically updated and perfectioned.

This second measure should be coupled with the last practice that I encourage: reviews and meta-analyses should be periodically assessed and retracted, if deemed to be obsolete. So far, article retraction is essentially a voluntary process. Considering the extent to which peer-reviewed studies matter for individual careers, and the stigma associated with retracting a paper, it is not particularly surprising that this practice is seldom implemented, even in high-level journals (Wray and Andersen, 2018). While hoping for a large-scale evaluation and retraction of obsolete papers is unfeasible, in my opinion it is possible to implement this process for reviews and meta-analysis, which are a minority of existing publications. For example, journals explicitly focusing on research synthesis could re-review published reviews and meta-analyses every 5 years, by involving experts in a specific field of research. By doing so, research synthesis available to researchers will comply with the most up-to-date practices of a certain field of studies.

By enforcing these three practices we will have “platinum standard” research synthesis for observational data, with an even greater reputation and impact. A journal which specializes in research synthesis has all the long-term incentives to implement them: by pioneering them it could well position itself among the most important journals in the field of conservation sciences and related disciplines.

Of course, it is possible that these practices will be distorted by dishonest researchers: by subjecting research synthesis to preregistration, dishonest peer-reviewers could prevent their colleagues from realizing a study for a certain amount of time. At the same time, by periodically retracting some reviews and meta-analyses, dishonest researchers could somehow push for retracting specific reviews of competing groups. However, these two adverse effects can be contained through open-peer review (Wolfram et al., 2020) and the evaluation of reviewers themselves.

References

Alston, J. M., *et al*. (2026). Clarifying space use concepts in ecology: range vs. occurrence distributions. Ecology, 107(3), e70300. https://doi.org/10.1002/ecy.70300

Anderson, S. C., *et al*. (2021). Trends in ecology and conservation over eight decades. Frontiers in Ecology and the Environment, 19(5), 274-282. https://doi.org/10.1002/fee.2320

Barmoen, M., *et al*. (2024). Living with wolves: A worldwide systematic review of attitudes. Ambio, 53(10), 1414-1432. https://doi.org/10.1007/s13280-024-02036-1

Berzi, D., *et al*. (2021). Use of European funds and ex post evaluation of prevention measures against wolf attacks (Canis lupus italicus) in the Emilia-Romagna Region (Italy). *Animals*, *11*(6), 1536. https://doi.org/10.3390/ani11061536

Boice, J. (2019). An exploration of systematic review publication trends in conservation biology journals. Issues in Science and Technology Librarianship, (91). https://doi.org/10.29173/istl2

Boyd, R. J., *et al*. (2023). We need to talk about nonprobability samples. Trends in Ecology & Evolution, 38(6), 521-531. https://doi.org/10.1016/j.tree.2023.01.001

Broekman, M. J. E., *et al*. (2023). HomeRange: A global database of mammalian home ranges. Global Ecology and Biogeography, 32(2), 198-205. https://doi.org/10.1111/geb.13625

Cerri, J., *et al*. (2025). Discrepancies between wolf distribution estimates from opportunistic mortality records and probabilistic field-based surveys. https://doi.org/10.32942/X2264T

Chickering, D. M., & Pearl, J. (1996, August). A clinician's tool for analyzing non-compliance. In *Proceedings of the National Conference on Artificial Intelligence* (pp. 1269-1276). https://cdn.aaai.org/AAAI/1996/AAAI96-188.pdf

Correia, H. E., *et al*. (2025). Designing causal mediation analyses to quantify intermediary processes in ecology. *Biological Reviews*, *100*(4), 1512-1533. https://doi.org/10.1111/brv.70011

Dimson, M., & Gillespie, T. W. (2023). Who, where, when: Observer behavior influences spatial and temporal patterns of iNaturalist participation. *Applied Geography*, *153*, 102916. https://doi.org/10.1016/j.apgeog.2023.102916

Dressel, S., *et al*. (2015). A meta-analysis of studies on attitudes toward bears and wolves across Europe 1976–2012. Conservation Biology, 29(2), 565-574. https://doi.org/10.1111/cobi.12420

Dumelle, M., *et al*. (2022). A comparison of design-based and model-based approaches for finite population spatial sampling and inference. Methods in Ecology and Evolution, 13(9), 2018-2029. https://doi.org/10.1111/2041-210X.13919

Farley, S. S., *et al*. (2018). Situating ecology as a big-data science: Current advances, challenges, and solutions. BioScience, 68(8), 563-576. https://doi.org/10.1093/biosci/biy068

Foo, Y. Z., *et al*. (2021). A practical guide to question formation, systematic searching and study screening for literature reviews in ecology and evolution. Methods in Ecology and Evolution, 12(9), 1705-1720. https://doi.org/10.1111/2041-210X.13654

Franchini, M., *et al*. (2021). The return of large carnivores and extensive farming systems: a review of stakeholders' perception at an EU level. Animals, 11(6), 1735. https://doi.org/10.3390/ani11061735

Gilbert, B., *et al*. (2021). A causal inference framework for spatial confounding. *arXiv preprint arXiv:2112.14946*. https://doi.org/10.48550/arXiv.2112.14946

Greco, I., *et al*. (2025). Placement matters: implications of trail-versus random-based camera-trap deployment for monitoring mammal communities. *Ecological Applications*, *35*(5), e70083. https://doi.org/10.1002/eap.70083

Gregoire, T. G. (1998). Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Forest Research, 28(10), 1429-1447. https://doi.org/10.1139/x98-166

Grekousis, G., *et al*. (2015). An overview of 21 global and 43 regional land-cover mapping products. International Journal of Remote Sensing, 36(21), 5309-5335. https://doi.org/10.1080/01431161.2015.1093195

Grüner Ofstad, E., *et al*. (2016). Home ranges, habitat and body mass: simple correlates of home range size in ungulates. Proceedings of the Royal Society B: Biological Sciences, 283(1845), 20161234. https://doi.org/10.1098/rspb.2016.1234

Gurevitch, J., & Nakagawa, S. (2015). Research synthesis methods in ecology. Ecological statistics: Contemporary theory and application, 200-227. https://doi.org/10.1093/acprof:oso/9780199672547.001.0001

Hankin, D., *et al*. (2019). Sampling theory: For the ecological and natural resource sciences. Oxford University Press, USA. https://doi.org/10.1093/oso/9780198815792.001.0001

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. Personality and social psychology review, 2(3), 196-217. https://doi.org/10.1207/s15327957pspr0203_4

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. Psychonomic bulletin & review, 25(1), 178-206. https://doi.org/10.3758/s13423-016-1221-4

Kutal, M., *et al*. (2025). Deeply political and populist decisions on large carnivores in Europe in recent times. Conservation Letters. https://doi.org/10.1111/conl.13125

Ladle, R. J., *et al*. (2016). Conservation culturomics. Frontiers in Ecology and the Environment, 14(5), 269-275. https://doi.org/10.1002/fee.1260

Lamouroux, J., *et al*. (2025). Addressing spatial confounding in geostatistical regression models: An R-INLA approach. *Methods in Ecology and Evolution*, *16*(9), 2082-2097.

Lorand, C., *et al*. (2022). Effectiveness of interventions for managing human-large carnivore conflicts worldwide: Scare them off, don't remove them. *Science of the Total Environment*, *838*, 156195. https://doi.org/10.1016/j.scitotenv.2022.156195

Mateo-Tomás, P., *et al*. (2022). The continued deficiency in environmental law enforcement illustrated by EU sanitary regulations for scavenger conservation. Biological Conservation, 270, 109558. https://doi.org/10.1016/j.biocon.2022.109558

Melis, C., *et al*. (2009). Predation has a greater impact in less productive environments: variation in roe deer, Capreolus capreolus, population density across Europe. Global ecology and biogeography, 18(6), 724-734. https://doi.org/10.1111/j.1466-8238.2009.00480.x

Nakagawa, S., *et al*. (2023). Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. Environmental Evidence, 12(1), 8. https://doi.org/10.1186/s13750-023-00301-6

Newsom, A., *et al*. (2026). Social perceptions of carnivores across the globe–a literature review. Human Dimensions of Wildlife, 31(1), 97-120. https://doi.org/10.1080/10871209.2025.2459733

Noonan, M. J., *et al*. (2019). A comprehensive analysis of autocorrelation and bias in home range estimation. Ecological Monographs, 89(2), e01344. https://doi.org/10.1002/ecm.1344

Nosek, B. A., *et al*. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600-2606. https://doi.org/10.1073/pnas.1708274114

Nowak, S., *et al*. (2026). Discarded carrion of poultry lures wolves to farms. *European Journal of Wildlife Research*, *72*(1), 2. https://doi.org/10.1007/s10344-025-02036-w

O'Dea, R. E., *et al*. (2021). Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. Biological Reviews, 96(5), 1695-1722. https://doi.org/10.1111/brv.12721

Pagany, R. (2020). Wildlife-vehicle collisions-Influencing factors, data collection and research methods. Biological conservation, 251, 108758. https://doi.org/10.1016/j.biocon.2020.108758

Pearl, J., & Mackenzie, D. (2018). The book of why. https://bayes.cs.ucla.edu/WHY/

Pettorelli, N., *et al*. (2014). Satellite remote sensing for applied ecologists: opportunities and challenges. Journal of applied ecology, 51(4), 839-848. https://doi.org/10.1111/1365-2664.12261

Pollo, P., *et al*. (2025). Harnessing meta-analyses' insights in ecology and evolution research. Royal Society Open Science, 12(10), 250759. https://doi.org/10.1098/rsos.250759

Rodríguez-Recio, M., *et al*. (2022). Estimating global determinants of leopard home range size in a changing world. Animal Conservation, 25(6), 748-758. https://doi.org/10.1111/acv.12777

Salvatori, M., *et al*. (2022). Day versus night use of forest by red and roe deer as determined by Corine Land Cover and Copernicus Tree Cover Density: assessing use of geographic layers in movement ecology. Landscape Ecology, 37(5), 1453-1468. https://doi.org/10.1007/s10980-022-01416-w

Seigle-Ferrand, J., *et al*. (2021). A systematic review of within-population variation in the size of home range across ungulates: what do we know after 50 years of telemetry studies?. Frontiers in Ecology and Evolution, 8, 555429. https://doi.org/10.3389/fevo.2020.555429

Sicacha-Parada, J., *et al*. (2021). Accounting for spatial varying sampling effort due to accessibility in Citizen Science data: A case study of moose in Norway. *Spatial Statistics*, *42*, 100446. https://doi.org/10.1016/j.spasta.2020.100446

Silk, M. J., *et al*. (2020). Perils and pitfalls of mixed-effects regression models in biology. PeerJ, 8, e9522. https://doi.org/10.7717/peerj.9522

Sollmann, R., *et al*. (2026). Combining Disparate Camera-Trap Surveys: Impacts of Spatial Bias and Design Variation on Large-Scale Ecological Inference. *Journal of Biogeography*, *53*(8), e70333. https://doi.org/10.1111/jbi.70333

Sunstein, C. R. (2021). Sludge: What stops us from getting things done and what to do about it. mIt Press. https://mitpress.mit.edu/9780262545082/sludge/

Sutherland, W. J., *et al*. (2019). Building a tool to overcome barriers in research-implementation spaces: The Conservation Evidence database. Biological Conservation, 238, 108199. https://doi.org/10.1016/j.biocon.2019.108199

Williams, B. K., & Brown, E. D. (2019). Sampling and analysis frameworks for inference in ecology. Methods in Ecology and Evolution, 10(11), 1832-1842. https://doi.org/10.1111/2041-210X.13279

Wolfram, D., et al. (2020). Open peer review: promoting transparency in open science. Scientometrics, 125(2), 1033-1051. https://doi.org/10.1007/s11192-020-03488-4

Wray, K. B., & Andersen, L. E. (2018). Retractions in science. Scientometrics, 117(3), 2009-2019. https://doi.org/10.1007/s11192-018-2922-4

van Beeck Calkoen, S. T., et al. (2023). Numerical top-down effects on red deer (Cervus elaphus) are mainly shaped by humans rather than large carnivores across Europe. Journal of Applied Ecology, 60(12), 2625-2635. https://doi.org/10.1111/1365-2664.14526

Venter, Z. S.,, et al. (2022). Global 10 m land use land cover datasets: A comparison of dynamic world, world cover and esri land cover. Remote Sensing, 14(16), 4101. https://doi.org/10.3390/rs14164101

Xu, P., *et al*. (2024). Comparative validation of recent 10 m-resolution global land cover maps. Remote Sensing of Environment, 311, 114316. https://doi.org/10.1016/j.rse.2024.114316