Smart Solutions, Big Returns: Closing Biodiversity Knowledge Gaps with Digital Agriculture

Authors
Ruben Remelgado[1,*]
Moudrý Vítězslav[2]
Elisa Padulosi[3]
Michela Perrone[2]
Petteri Vihervaara[4]
Christopher Marrs[5]
Anette Eltner[5]
Duccio Rocchini[2,3]
Anna F. Cord[1]

Affiliations
[1]Chair of Computational Landscape Ecology, Dresden University of Technology, Helmholtzstr. 10, 10169 Dresden, Germany.
[2]Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic.
[3]BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126, Bologna, Italy.
[4]Finnish Environment Institute (Syke), Nature Solutions, Latokartanonkaari 11, FI-00790, Helsinki, Finland
[5]Institute of Photogrammetry and Remote Sensing, Dresden University of Technology, Helmholtzstr. 10, 10169
* Corresponding author (ruben.remelgado@gmail.com)

SUMMARY

The global expansion and intensification of food production threaten biodiversity, vital for ecosystem services and food security. The Kunming-Montreal Global Biodiversity Framework (GBF) advocates drastic changes in agricultural management, yet translating recommendations into local action is challenging. Biodiversity-friendly practices carry highly uncertain benefits, dissuading their adoption. Reducing uncertainties demands systematic data on biodiversity-yield interactions. Yet, many biodiversity studies lack such detailed data, and food production systems remain underrepresented in global biodiversity datasets. Here, we illustrate how Digital Agriculture can address these issues. It uses technologies also applied in biodiversity monitoring, but is currently treated separately, leading to duplication of effort and costs. Digital Agriculture provides a low-cost, low-effort solution for monitoring biodiversity in food production systems, linking it directly to land management practices, and benefiting multiple stakeholders without creating additional monitoring requirements. This integration has the potential to increase the effectiveness of the GBF in promoting sustainable agricultural practices.

KEYWORDS
GBF, monitoring, agroecology, GBIF, uncertainty, integration

INTRODUCTION

The global expansion and intensification of food production systems has led to drastic losses of habitat and biodiversity[1]. As the human population continues to grow, the increasing demand for food[2] and the associated expansion of farmland are threatening thousands of species with extinction[1]. However, many of these species actually provide ecosystem services that benefit food production. For instance, as reported in 2015, 5-8% of food production worldwide was directly dependent on pollinators, valued at an estimated US$235-577 billion[6]. Soil is home to more than 50% of the Earth's species[4] and enables the growth of over 140 million metric tons of food annually[5], and vertebrate diversity is important to halt the spread of pests[6] that can otherwise cause up to 40% of global yield losses[7]. Conserving biodiversity is thus essential to ensuring food security[3] and resilience[8].

Literature suggests that changes in food production practices are urgently needed[9], and that reducing the agricultural footprint is critical for ecosystem regeneration[10]. This is also acknowledged in the Kunming-Montreal Global Biodiversity Framework (GBF)[11], which was recently adopted at the 15th meeting of the Conference of the Parties to the Convention on Biological Diversity (CBD). In addition to reducing direct threats to biodiversity, the GBF advocates more sustainable land use that helps conserve biodiversity and nature's contribution to people. In particular, Target 10 promotes sustainable intensification[15] or agroecological practices[13] that help maintain biodiversity-provided services in food production systems. There is evidence that such practices can promote biodiversity gains without compromising food production requirements[14]. However, translating GBF recommendations into local action is not straightforward. Despite a large body of literature on the general ecosystem benefits of more sustainable management practices, these benefits may be gradual and uncertain[15], may take several years (as with pollination[16]) or decades (as with soil services[17]) to manifest, or may not occur at all[18]. In turn, it has been argued that biodiversity-enhancing management practices will lead to transition periods of lower productivity[19], which may put millions of people at risk of hunger by 2050[20].

Systematic monitoring capabilities are necessary to provide reliable and scalable recommendations on when, where, and which agricultural management practices should be implemented to promote biodiversity[21]. Over the past decade, the biodiversity monitoring community has largely reached consensus on key variables for measuring and monitoring biodiversity, referred to as Essential Biodiversity Variables[22] (EBVs). Recently, a similar set of Essential Ecosystem Service Variables[23] (EESVs) was proposed. Nonetheless, effective management of changes in food production systems requires reference data on crop conditions and management practices at the time the species (or trait) was observed. Combining this information with field and farm-level yield measurements will be a critical step in understanding complex biodiversity yield trade-offs and in guiding the translation of changes in EBVs and EESVs in agroecosystems into confident policy recommendations.

Yet, biodiversity studies often lack such detailed reference data on agricultural management, crop condition and yield. Instead, state-of-the-art literature often relies on coarser agricultural statistics (e.g., at sub-national scales[24]) as proxies of yield. Limited data on farmland biodiversity also constrain choices of methods and are one reason for thematic focuses of research (e.g. on single crops[25] or selected taxonomic groups such as birds[25] and butterflies[26]) during the analysis of drivers of biodiversity change. Species distribution models, the most common class of models in ecology, evolution, and conservation[27], have been used to study how land use drives biodiversity patterns (e.g. ref[28]) – though, as we will show later, the species observations informing these models tend to originate from outside food production systems. All of this prevents drawing comparable causal links between incremental changes in biodiversity and concurrent changes in management practices[29].

One way to address these data limitations is involving farmers to improve the collection of biodiversity data in food production systems[30]. Farmers are responsible for implementing conservation policies (e.g. as acknowledged in the EU Common Agricultural Policy[31]) and control access to lands where data is to be collected. Following this premise, various participatory strategies have been proposed to improve biodiversity monitoring in food production systems. These suggestions range from involving farmers in the design of conservation measures[32] or as citizen scientists[33] to proposing a networked design of stakeholders, data, tools, and biodiversity monitoring programs up to the global scale[34,35].

However, we argue that current participatory strategies are insufficient and rely on financial incentives to motivate farmers[31] to participate in what is ultimately an additional and challenging task. Involving farmers in the governance, organisation and execution of biodiversity monitoring also poses some challenges such as limited representativeness of sampled farmland due to varying willingness to participate in such programs[30]. Issues of varying data quality[36] have also been reported. Finally, farmers may view biodiversity as pests[37] and thus not feel the urgency to contribute with data[38].

To tackle these issues and assure systematic data acquisitions, we must ensure the participation of farmers without imposing additional challenges on them (**Fig. 1**). We propose this can be achieved through technologies used to optimise food production (hereafter 'digital agriculture'). Whereas digital agriculture helps farmers optimise food production, they may also provide highly valuable, but currently overlooked, biodiversity data streams (**Fig. 2**). Here, 'biodiversity' refers specifically to species observations and, potentially, species traits. The use of digital agriculture is essential to tackle land system biases in biodiversity monitoring. Our global analysis of existing biodiversity data indicates that current monitoring efforts inadequately capture global land use patterns (**Fig. 3**) and do not reflect the global distribution of biodiversity within them (**Fig. 4**). We then discuss how data biases relate to political factors and land privatisation (**Fig. 5**), and provide recommendations for improving biodiversity monitoring through digital agriculture. We aim to stimulate technical advances that reduce redundancy and costs in environmental monitoring, while accelerating benefits for nature and people.

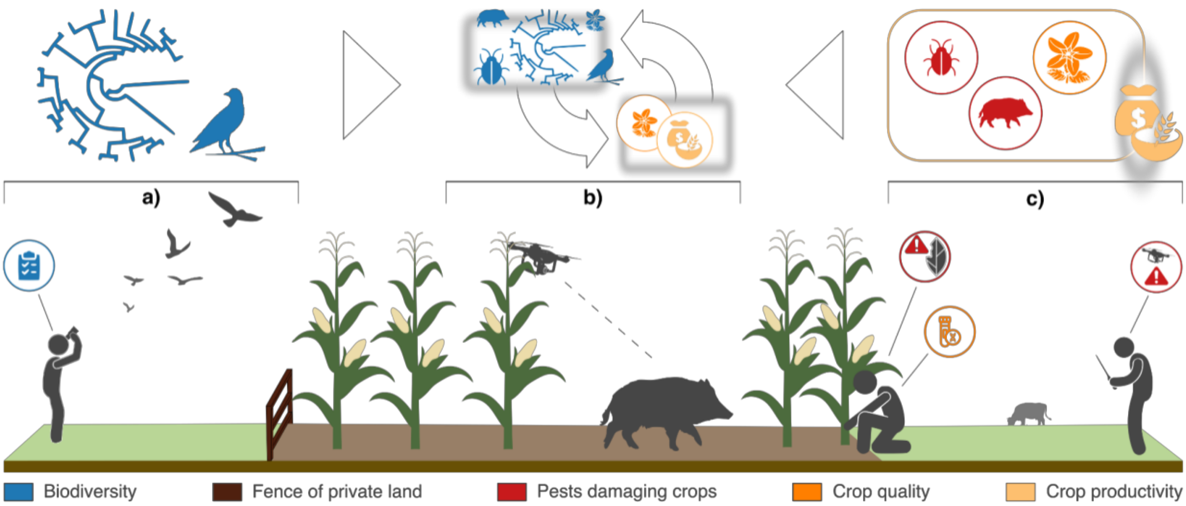

Figure 1. Unaccounted biodiversity in food production systems. a) A surveyor of biodiversity is located outside private farmland, but can record overpassing birds or other species detectable from the distance. Such information can enter public or private databases and then contributes to knowledge and monitoring of local biodiversity (shown in blue). The observation of areas within the farmland is, however, restricted (e.g. due to fencing or dense crop cultivation). **b)** Within the maize plantation, a wild pig is spotted by a drone. The drone was deployed by a land manager (**c,** on the right**)** with the intention of surveying crop conditions and detecting potential pests (in red). Simultaneously, another land manager collects data on pests in the cultivated crops (e.g. bugs, in red) and on crop conditions (in orange) to make management decisions. As shown at the top of panel **c**), information on pests and crop conditions informs on crop productivity (in yellow). Combining the biodiversity information and data given in **a)** with the pest and crop condition data given in **c)** provides a more complete picture of biodiversity and how biodiversity responds to (and affects) crop productivity (shown above **b**).

Parallels of Digital Agriculture and automated biodiversity monitoring

Digital Agriculture and automated biodiversity monitoring share many similar technologies. Drones, for instance, are employed to detect, locate, and count pests (e.g., insects[39], rodents[40], wild pigs[41]), to detect plant diseases[42] and to monitor cattle in large pasture areas[43] in support of improving food production and quality. Similarly, drones are being used in biodiversity surveys to detect wildlife more efficiently than human observers[44], including rare[45] and elusive[46] species. Passive acoustic sensors at ground level can both measure crop height[47] and be used to detect soniferous species (e.g. birds[53]). On the other hand, active acoustic sensors can provide information on crop health[49] and physiological traits that distinguish non-crop plant species[50]. More recently, robots equipped with artificial intelligence are enabling the extraction of environmental DNA (eDNA) to detect organisms harmful to crops[56] (e.g. insects), information that would otherwise be used to distinguish taxa[57] (e.g. insect, microbial species). Robots are also increasingly used to assist in farmland management (e.g. to remove weeds[53]) or to enable biodiversity surveys of inaccessible habitats[54] (e.g. large farmlands). All of these technologies can be integrated with satellite remote sensing to monitor biodiversity change[55] or long-term trends in food production[56].

Digital Agriculture: a hidden source of biodiversity data

Although digital agriculture and biodiversity monitoring have obvious parallels, they are treated as separate branches of environmental monitoring in research, university education and practice, resulting in duplication of efforts and costs. In turn, combining these branches can yield critical and novel insights (**Fig. 2**). Species traits and occurrences can be directly linked to concurrent biophysical measurements of crop conditions to obtain data-driven knowledge on species-specific patterns of resource and habitat selection. This will significantly advance our understanding of habitat vs. matrix[57] in agricultural landscapes. For instance, there is evidence that some species can adapt to man-made habitats[58], and that even species thought to have been displaced by cropland expansion can return to those lands[59]. Conversely, if a species is not recorded on farmland, despite being detected outside of it, this can provide data on true species absences. Knowing whether a species is present or absent, and its relation to specific crop conditions, can support systematic causal analysis of which management practices enhance or diminish biodiversity[60] and thus ecosystems functions and ecosystem services[61]. Ultimately, the concurrent monitoring of biodiversity and food production enables thorough and reproducible landscape-level experiments to fully comprehend cross-scale biodiversity-yield relationships. In contrast, maintaining biodiversity monitoring as a specialised effort creates persistent land-use biases in biodiversity data streams, as demonstrated in the following sections.

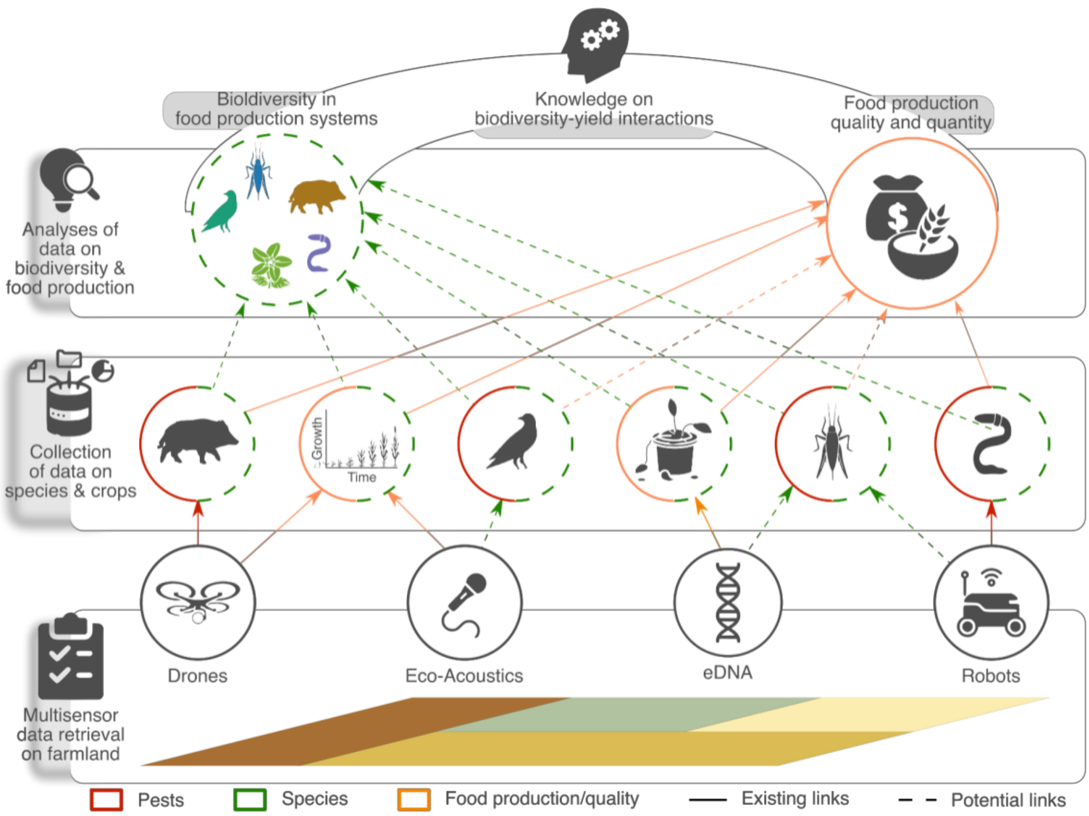

Figure 2. Sensing biodiversity through Digital Agriculture. A cropland area (where colours distinguish fields with different crops) is observed using different sensing technologies (indicated in the circles above the fields).

These technologies provide several pieces of information, such as on pests, plant growth and conditions. All of the existing information links (shown with full lines) feed into food production and quality information systems (in orange). We here propose new information links (shown with dashed lines) that can feed monitoring systems for both biodiversity (in green) and food production and quality. Here, 'biodiversity' refers specifically to species observations and, potentially, species traits. For instance, whereas drones and robots inform on the presence of pests so that farmers can make management decisions, this information can additionally be used to distinguish different species for biodiversity information. Similarly, ecoacoustics and eDNA used to monitor crop growth and health can simultaneously be used to acquire information on roaming species not captured directly through image recognition. All of this information can be fed into biodiversity monitoring workflows that can distinguish and catalogue species occurrences and assess species traits. In addition, information on biodiversity can be combined with that on food production and quality to acquire new knowledge on species habitat preferences. This can help us establish causal links between biodiversity and food production and quality that inform on the provision of ecosystem services by particular species, and which can then feed the mapping of these services and subsequent policymaking, monitoring, and conservation.

Food production system are underrepresented in global biodiversity datasets

The Global Biodiversity Information Facility (GBIF) provides access to Big Data on species occurrences that directly inform the CBD and, by extension, the GBF. Yet, global differences in data-sharing cultures and geographical mobility[62], preferential sampling of certain taxonomic groups (as seen for some insect taxa[63]), and the disproportionate sampling of populated areas[64], lead to spatial and taxonomic biases in GBIF that can distort biodiversity assessments[65]. Here, we focus on previously unreported biases in observations in cropland (see the **SI** for details about our methodology).

GBIF provides access to over half a billion species observations between 2015 and 2019 (reference period with land cover data, see **SI**). Of these, 71.9% were collected in 0.25° cells with <30% cropland cover (**Fig. 3a**). Conversely, 22.2% of global cropland was missing species observations. This includes large parts of countries where the pressure of food production on biodiversity is high and may even further increase. For instance, 33.6% of the cropland cover in China, which produced 20.7% of the world's cereals in 2019[66], lacked any species observations during the 2015-2019 period. Similarly, more than one third of cropland in Angola, where 73.5% of the population faced moderate to high food insecurity in 2019[70], lacked species observations.

Species observations were concentrated in relatively stable cells with 53.4% occurring in cells with a change in cropland cover < 1% (**Fig. 3b**). In turn, species occurrence in nearly 32.2% of the cropland area undergoing change was not recorded (**Fig. 3c**). This includes areas of persistent land abandonment (e.g. in Kazakhstan[67], which lost >2 million ha of agricultural land between 2015 and 2019[66]) and cropland expansion (e.g. along the Sahel belt, which gained >4 million ha between 2015 and 2019[66], associated with losses of biodiversity-rich shrubland ecosystems[68]). Our results hence suggest biases in the selection of sites for biodiversity studies. Areas with limited cropland cover, or where cropland cover is stable or rapidly changing, appear to be favoured. Indeed, we found that the cumulative global distribution of cropland-covered cells differed significantly from the distribution of the subset of cells with species observations (Kolmogorov-Smirnov test with a $p < 0.005$, **Fig. 3d**). We found similar differences in the distributions of cropland-cover changes ($p < 0.005$, **Fig. 3e**).

Land system biases in species observations also reflect biases in the monitoring of global species richness. Focusing on cells with cropland cover, we found that the cumulative distribution of species richness over 0.25° cells differed significantly from the cumulative distribution drawn only from cells with species observations ($p < 0.005$, **Fig. 4a)**. This includes biodiversity hotspots such as the Amazonian forest and the Congo Basin, where international food demands and investments threaten biodiversity through cropland expansion[69,70] (**Fig. 4b**).

A species-focused analysis of these data indicates that tackling land system biases in species observations may challenge our assumptions about how different species interact with cropland. For instance, whereas the IUCN Red List reports only 8,466 species inhabiting cropland-related ecosystems, 60,544 species were observed in cropland between 2015 and 2019 alone in our analysis (based on 100-m resolution data, see **SI**). Of these species, 36,980 were not reported to occur in cropland, and more than one fifth of these are considered threatened (**Fig. 4c**). This includes species that have been reported to forage in cropland (e.g., Eastern Gorilla, *Gorilla beringei*[71]) and data deficient species assumed to be dependent on forests (e.g. Alcathoe Bat, *Myotis alcathoe*[72]). While assuming that classification errors in the land cover data play a role in our assessment, our results create a reasonable demand for additional scrutiny.

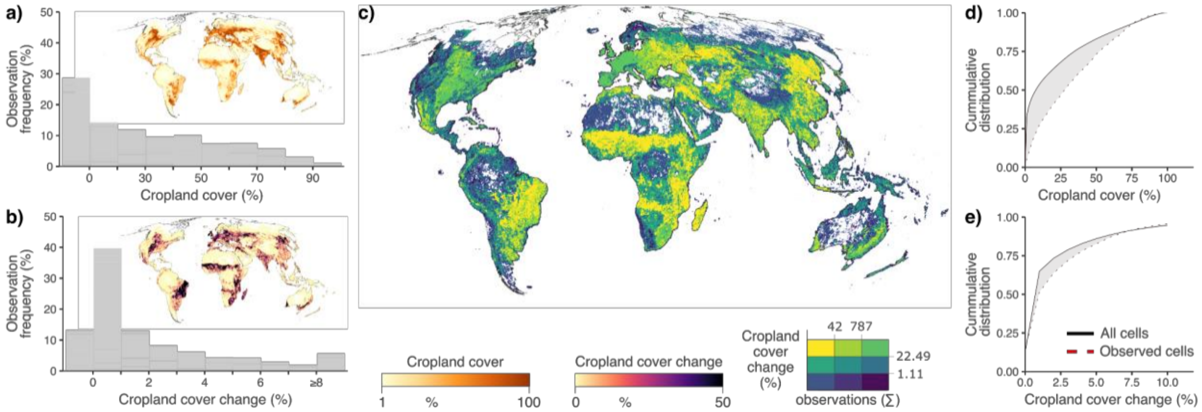

Figure 3 - Sampling gaps and biases in food production systems. a) Global distribution of cropland cover. The map shows per-cell cropland cover percentages, and the histogram shows the distribution of species observations per percentage of cropland cover. **b)** Similar to **a),** but instead characterising the global distribution of cropland-cover change. **c)** Per-cell characterisation of cropland-cover changes and the number of species observations at a 0.25° resolution. For instance, yellow indicates cells with a low number of species distributions and a high cropland-cover change, whereas dark blue shows a low cropland cover and a high number of species observations. Here, 'low' values are below the 33rd percentile of the global distribution of the corresponding variable. In turn, 'high' values are above the 66th percentile. **d)** Comparison of the cumulative distribution of cells with cropland cover (**d**, black line) and the cumulative distribution drawn by a subset of cells containing species observations (dashed red line). The pink polygon indicates the distance between distributions. **e)** Similar to **d)**, but comparing distributions of cropland-cover change values.

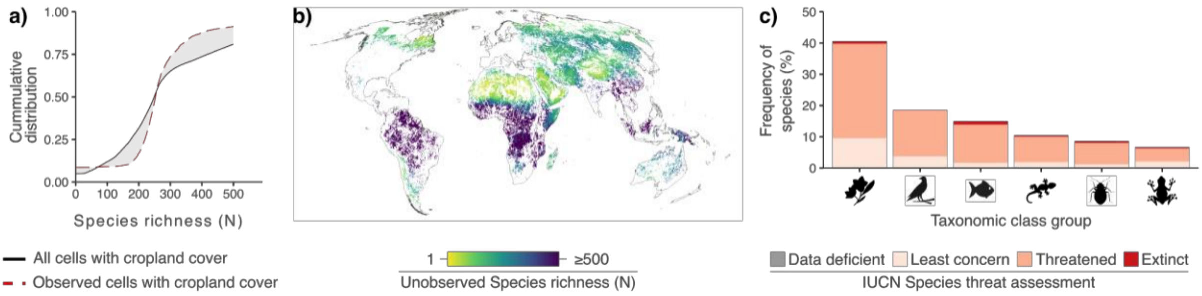

Figure 4. Biodiversity knowledge gains and biases in food production systems. a) Comparison between the cumulative distribution of richness in pixels with cropland cover anytime between 2015 and 2019 (black line) against the cumulative distribution of the subset of pixels with species observations made during the same period (red line). **b)** Global map of species richness of mammals, birds, reptiles, and amphibians in pixels that experienced cropland gains between 2015 and 2019, but where no species observation was made during the same period. The richness map was obtained from the IUCN Red List of species. **c)** Proportion of species (y-axis) per taxonomic group (x-axis) that are not reported in the IUCN Red List as inhabitants of cropland-related ecosystems, but for which GBIF provides species observations found in cropland pixels between 2015 and 2019. The plot shows the leading taxonomic group that together account for 80% of all species not reported as inhabitants of cropland-related ecosystems, but with observations in 100-m cells with cropland cover. The colour of the bars indicates the proportions of species associated to different threat categories.

Political (dis)incentives and land privatisation hinder biodiversity monitoring

The biases revealed here are likely driven by biodiversity-related policies. For instance, article 7 of the CBD[73], approved in 1992, promotes efforts to identify and monitor species in biodiversity-rich and wilderness areas. In contrast, food production systems are only mentioned in relation to domesticated or cultivated species. The lack of established biodiversity monitoring programs in many countries[34] makes data collection heavily reliant on short-term projects[74]. In line with CBD guidelines, these projects (unless their focus is on agroecology) tend to prioritise areas of high biodiversity[79] or unique and pristine ecosystems. Comparatively lower species detection rates in food production systems[76] have likely discouraged systematic scientific investments.

We see such biases, for instance, in Brazil's legal Amazonia (**Fig. 5a-c**), where expanding food production systems pose a major threat to biodiversity[77]. Improved access to remote areas and reduced travel time occurs with road construction facilitating commodity transportation, but potentially also the establishment of new biodiversity monitoring sites. (**Fig. 5a**). Between 2000 and 2019, the majority of species observations in food production systems of legal Amazonia (combining cropland and pastures) occurred in expansion areas (64.5%, compared to 9.4% in areas of abandonment). Yet, these observations totaled only 47,622 out of all 219,529 observation sites (i.e. unique year-coordinate combinations) recorded in the region . Of these, 51.3% were recorded in areas where travel times did not improve substantially (i.e. < 1h) despite the concurrent expansion of food production systems (**Fig. 5b**). Moreover, in the expansion areas, most observation sites were visited only once (**Fig. 5c**). This makes it difficult to systematically assess biodiversity change due to agricultural activities in general or specific management practices in particular.

However, even if systematic assessments were sustained, patterns of land access and privatisation can be a source of spatial biases in species observation data. In California, for instance, we recorded 2,936 species observations made within food production systems in 2019 (0.6% of all observations in the state). Based on these data, we found a correlation between the distance of observation points to the centre of the nearest land parcel and the approximate radius of that parcel (Spearman's rho=0.61, $p < 0.005$, **Fig. 5d**). This indicates observations concentrated along the boundaries of land parcels (e.g., **Fig. 5e**), limiting insight into differences between field core and edge. When this occurs, the effort to observe the entirety of a land parcel increases (**Fig. 5f**), and the combined effect of the Earth's curvature and topography may make ground-dwelling species imperceptible (**Fig. 5g**). Land parcels larger than 100 ha composed 85.7% of the world's food production systems by 2020, and this trend is likely to continue as improvements in technologies enable managing increasingly larger areas[78]. We therefore argue that the combined effects of globally varying field sizes[79] and the associated limitation of access to lands require more attention in biodiversity monitoring, where an emphasis is currently placed on travel distances[80].

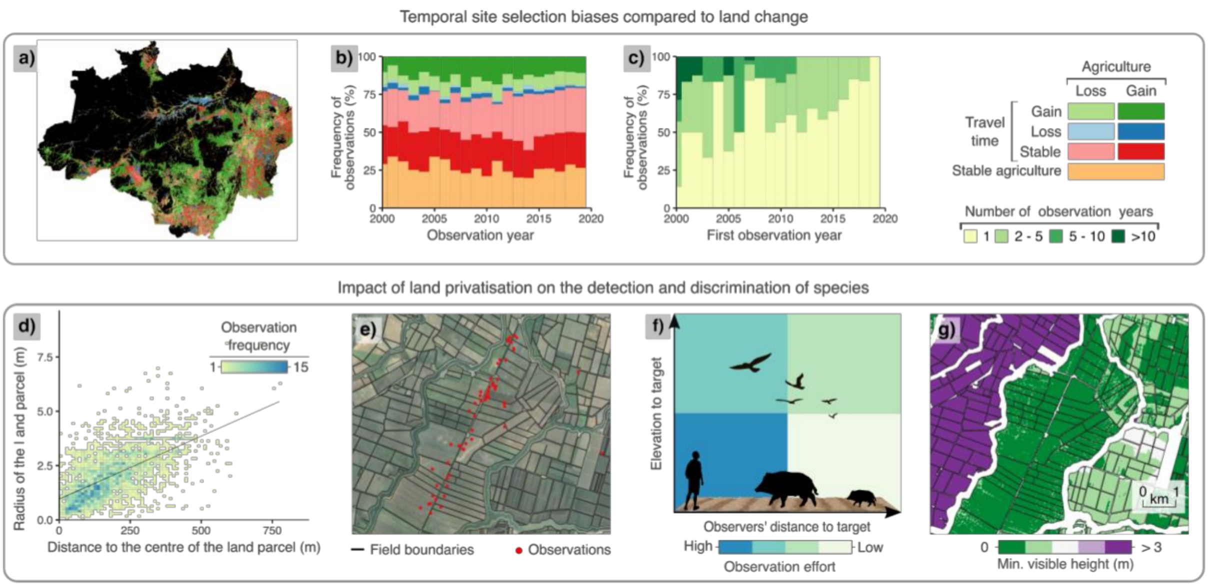

Figure 5. Sources of land system biases in species observation data. a) Map of the legal Amazonia, Brazil. Each 1-km cell classifies the expansion and abandonment of agriculture, and changes in travel time (by more than 1h) from a given cell to the nearest city with at least 50,000 habitats. In addition, the map distinguishes stable agriculture. **b)** Relative distribution of species observations per each class in **a)**, and for each year between 2000 and 2019. **c)** For each year, the cells first observed in that year are coloured based on the number of subsequent years of observation. In each year, the number of cells per year-frequency category is normalised by the number of cells observed in that year. **d)** Correlation between the radii of land parcels – estimated as □□□□(□□□□□ □□□□/□) – and the distance between field centroids and species observations. Colours indicate frequency of observation. **e)** Example of cropland area in California where black lines indicate field boundaries, and red dots locate species observations collected along a road. **f)** Effort to observe and discriminate ground-dwelling and airborne species as a function of the distance and elevation measured between the observer and the species. **g)** Minimum elevation at which a species can be perceived. This is estimated based on the species observation locations mapped in **e),** and based on a digital elevation model with a resolution of 1-m.

Closing biodiversity knowledge gaps with Digital Agriculture

The limitations of biodiversity data we exemplified do not diminish the huge and extremely valuable efforts of both professional and citizen science biodiversity observers[33]. Furthermore, the expertise of taxonomic specialists remains indispensable[81]. Yet, with urban populations projected to increase by 13% by 2050 at the cost of those in rural areas[82], citizen science at least is likely to be displaced away from food production systems. And, as alluded to, factors such as restricted access to land parcels and funding trends contribute increasing data gaps. It is vital to ensure that policy recommendations, such as those on biodiversity-friendly agricultural management practices (GBF, target 10), are not skewed by sampling biases but based on systematic and global biodiversity monitoring capabilities[37]. Such capabilities would support rapid detections of biodiversity changes and the subsequent attribution of their causes[35], enabling more confident policy recommendations[29].

However, large-scale biodiversity monitoring programs are still lacking in most countries[34]. Their implementation would cost millions of US dollars annually[83], clashing with global inequalities in economic power[76]. In contrast, investments promoting innovation in agriculture are increasing rapidly. Globally, the agriculture market reached US$6.2 trillion in value in 2021 after an exponential growth[85], which is more than three times the GDP of Sub-Saharan Africa. These investments are accompanied by those in Digital Agriculture to increase yields, improve efficiency, reduce waste, cost and environmental impact, and sustain food security[86]. For instance, the UN-led *50 by 2030* initiative will invest US$500 million to digitise food production in 50 countries in Africa, Asia, the Middle East and Latin America by 2030[87].

Digital Agriculture technologies are similar to those used to survey biodiversity (see *Parallels of Digital Agriculture and biodiversity monitoring*). Therefore, data originally intended to monitor food production can also help detect and describe non-crop biodiversity in food production systems. It is important to note, however, that Digital Agriculture is not a replacement for traditional biodiversity monitoring. Exploratory research is still required (e.g., subterranean biodiversity is largely unknown[88]), and biodiverse ecosystems will require continuous and dedicated monitoring (e.g. such as in tropical moist forests, which are reported to be home to more than half of the world's vertebrates[89]). In fact, focusing on Digital Agriculture alone may even shift current spatial and thematic biases in biodiversity data. Nonetheless, Digital Agriculture offers an immediate and cost-effective solution to address apparent current knowledge gaps on biodiversity in food production systems. It also provides a platform for systematic assessments of biodiversity-yield interactions that can improve recommendations for sustainable agricultural management practices.

To enable synergies in the short-term, we must ensure that the data generated by Digital Agriculture becomes Findable, Accessible, Interoperable, and Reusable (FAIR). This would empower biodiversity experts to apply their own methods to translate digital agriculture data into biodiversity data (e.g., by applying computer vision methods to drone imagery originally intended to provide information on crop conditions). Combined with knowledge of concurrent management practices and crop conditions, biodiversity experts can generate knowledge and models on biodiversity-yield interactions. To support data sharing, initiatives are already underway to promote FAIR agricultural data principles[90], and platforms are being designed to provide such data[91]. In addition, the UN-funded Consultative Group on International Agricultural Research (CGIAR) is advancing generalised principles and tools to enable the sharing and distribution of big agricultural data[92].

In the long term, sustaining the benefits of Digital Agriculture for biodiversity monitoring demands collaborative workflows between farmers, biodiversity experts and decision-makers. However, we need to go beyond current participatory strategies. Rather than involving farmers in biodiversity monitoring, which adds to the challenges of current farming, collaborations could, for instance, coordinate smart solutions to deploy sensing technologies in ways that maximise returns for all stakeholders involved. For example, drones used to assess crop growth could also be employed to monitor green infrastructure in agricultural land without compromising the farmers' needs. Similarly, night-time and automated deployments of acoustic sensors would enable the detection of pests (from a farmers perspective), while also allowing for the recording of nocturnal species (e.g. bats, insects). This could also help bridge the divide that many farmers feel between society's expectation to conserve nature and the desire to appear productive to their peers[93]. Further, studies have shown that knowledge of the ecological effectiveness of agricultural practices can help increase the likelihood of their adoption by farmers[94]. Cooperation could be extended to other disciplines. For instance, the involvement of engineers will support the adoption of new sensing technologies, such as robots (e.g. ref[95]). Experts from other disciplines, such as agronomy and computer vision, could help design strategies for calibrating sensing routines to increase knowledge benefits across disciplines (and farmers) while reducing monitoring costs.

Despite this potential, we should acknowledge that there are risks associated with relying on rapidly advancing technologies. This may further concentrate data collection efforts in countries with the financial resources to invest in acquiring, maintaining, and distributing these technologies. Furthermore, disparities in technical and scientific development may affect the uptake of new technologies[96]. To avoid this, we must ensure that data contributions are not restricted to only advanced technologies. Where financial capacity is lower, even data obtained through manual sampling and visual assessments of crops would be immensely valuable given the persistence of spatial and taxonomic biases in biodiversity data[62–64]. However, we also need to ensure that different tiers of data contributions are accompanied by uncertainty metrics. To this end, metadata standards and quality control measures have been proposed[91] that can support the development and calculation of quality metrics (e.g. on data completeness, clarity). Such metrics can then be used as part of models and analytical frameworks fed with data from Digital Agriculture, such as through 'weight of evidence' approaches[97].

Our recommendations have global relevance. In the global north, where monitoring capabilities are most advanced[62], Digital Agriculture can help tackle current land system biases in biodiversity data as outlined above. For instance, the European Biodiversity Observation Network (Europa BON) reported biases in biodiversity data that prevent systematic assessments of land-use threats to biodiversity[98]. To achieve this, financial incentives are required to enable systematic biodiversity monitoring capabilities in Europe, which are currently not explicitly supported in the European Common Agricultural Policy for the period 2023-2027[31]. We see particular opportunities to tackle general taxonomic biases. Insects encounter large taxonomic gaps in existing biodiversity databases (e.g. as shown in GBIF[63]). Agriculture has the potential to enhance insect diversity under certain conditions, such as through crop heterogeneity[99]. Considering insects' significant role as pests, it becomes plausible that extensive datasets on insect species occurrences can be efficiently derived through the application of Digital Agriculture.

In the global south, our recommendations may also help address gaps in monitoring capacity. In countries where biodiversity monitoring is, by necessity, considered a lower priority compared to other development issues (e.g. food security), Digital Agriculture offers a cost-effective solution to address multiple development challenges without large additional financial or organisational burdens. In fact, research shows that most of the investments in agriculture across the global south are already aimed at innovation[100]. The agricultural market could likely bear the financing costs if policies were in place to motivate the sharing of acquired biodiversity data. In turn, as new data enables more concrete and confident valuations of biodiversity and its contributions to people, thereby transforming nature from a mere by-product to a quantifiable asset, private investments in the agricultural industry can additionally increase data returns.

CONCLUSION

The GBF has set an ambitious agenda to prevent further biodiversity losses. It aims to restore 30% of global ecosystem extents by 2030 (Target 2) and emphasises the importance of changes in agricultural management (Target 10) to achieve this mammoth task. However, the advancement and implementation of sustainable agricultural practices faces significant challenges due to uncertainties surrounding the optimal management of trade-offs between biodiversity and yield. These complexities pose major obstacles to addressing the critical issues related to global food security. . Reducing uncertainties therefore demands tackling biases in biodiversity data and pairing them with data on food production. Here, we propose that Digital Agriculture offers a cost-effective solution. It employs some of the same technologies used to monitor biodiversity, which means it can in principle provide systematic biodiversity data directly linked to information on food production. This can address current land system biases in biodiversity data without imposing new burdens on farmers, who can instead benefit immediately from the same data. This would enable co-designing nature-based solutions sensitive to the needs and challenges of both farming and biodiversity monitoring. Nonetheless, we emphasise that Digital Agriculture should not be regarded as a direct substitute for traditional biodiversity monitoring, and the expertise of taxonomic specialists remains indispensable. In fact, if not properly designed, monitoring through Digital Agriculture may even shift current biases in biodiversity data. Nonetheless, the integration of Digital Agriculture into biodiversity monitoring systems could substantially improve our understanding of the interactions between biodiversity and yield, how these interactions generate ecosystem services, and how species use the agricultural matrix for movement and foraging. Ultimately, this integration would enable smart solutions to optimise data extraction, leading to big knowledge returns.

DATA AND CODE AVAILABILITY

The code data generated in this paper are provided as supplementary material.

ACKNOWLEDGEMENT

This work was supported by the Horizon Europe project EarthBridge (Grant No. 101079310) as well as the ECO²SCAPE project funded by the German Federal Ministry of Education and Research (BMBF, Grant No. 16LW0079K). We would also like to thank all of our collaborators in EarthBridge who provided us with a valuable scientific platform to discuss and clarify our ideas.

REFERENCE

1. IPBES. *Global Assessment Report on Biodiversity and Ecosystem Services of the*

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (IPBES secretariat, 2019).
2. van Dijk, M., Morley, T., Rau, M. L. & Saghai, Y. A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nat. Food* **2**, 494–501 (2021).
3. IPBES. *Summary for policymakers of the assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on pollinators, pollination and food production.* (2016).
4. Anthony, M. A., Bender, S. F. & Heijden, M. G. A. van der. Enumerating soil biodiversity. *Proc. Natl. Acad. Sci.* **120**, e2304663120 (2023).
5. Fonte, S. J., Hsieh, M. & Mueller, N. D. Earthworms contribute significantly to global food production. *Nat. Commun.* **14**, 5713 (2023).
6. Lindell, C., Eaton, R. A., Howard, P. H., Roels, S. M. & Shave, M. E. Enhancing agricultural landscapes to increase crop pest reduction by vertebrates. *Agric. Ecosyst. Environ.* **257**, 1–11 (2018).
7. IPCC. *Scientific review of the impact of climate change on plant pests – A global challenge to prevent and mitigate plant pest risks in agriculture, forestry and ecosystems.* https://doi.org/10.4060/cb4769en (2021).
8. Isbell, F. *et al.* Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature* **526**, 574–577 (2015).
9. Delabre, I. *et al.* Actions on sustainable food production and consumption for the post-2020 global biodiversity framework. *Sci. Adv.* **7**, eabc8259 (2021).
10. Beyer, R. M., Hua, F., Martin, P. A., Manica, A. & Rademacher, T. Relocating croplands could drastically reduce the environmental impacts of global food production. *Commun. Earth Environ.* **3**, 49 (2022).
11. Bommarco, R., Kleijn, D. & Potts, S. G. Ecological intensification: harnessing ecosystem services for food security. *Trends Ecol. Evol.* **28**, 230–238 (2013).
12. Inger Andersen. *Adoption of the "Kunming-Montreal Global Biodiversity Framework" (GBF)*. https://policycommons.net/artifacts/3348807/adoption-of-the-kunming-montreal-global-biodiversity-framework-gbf/ (2022).
13. McGreevy, S. R. *et al.* Sustainable agrifood systems for a post-growth world. *Nat. Sustain.* **5**, 1011–1017 (2022).
14. MacLaren, C. *et al.* Long-term evidence for ecological intensification as a pathway to sustainable agriculture. *Nat. Sustain.* **5**, 770–779 (2022).
15. Garibaldi, L. A. *et al.* Farming Approaches for Greater Biodiversity, Livelihoods, and Food Security. *Trends Ecol. Evol.* **32**, 68–80 (2017).
16. Grab, H., Poveda, K., Danforth, B. & Loeb, G. Landscape context shifts the balance of costs and benefits from wildflower borders on multiple ecosystem services. *Proc. R. Soc. B* **285**, 20181102 (2018).
17. Brady, M. *Report on Ecological-Economic Models and Evaluation of Effects of Ecological Intensification on Farm Income*. (2016).
18. Elek, Z. *et al.* Mixed effects of ecological intensification on natural pest control providers: a short-term study for biotic homogenization in winter wheat fields. *PeerJ* **8**, e8746 (2020).
19. Kovács-Hostyánszki, A. *et al.* Ecological intensification to mitigate impacts of conventional intensive land use on pollinators and pollination. *Ecol. Lett.* **20**, 673–689 (2017).
20. Fujimori, S. *et al.* Land-based climate change mitigation measures can affect agricultural markets and food security. *Nat. Food* **3**, 110–121 (2022).
21. Toivonen, M., Herzon, I. & Kuussaari, M. Differing effects of fallow type and landscape structure on the occurrence of plants, pollinators and birds on environmental fallows in

Finland. *Biol. Conserv.* **181**, 36–43 (2015).
22. Pereira, H. M. *et al.* Essential Biodiversity Variables. *Science* **339**, 277–278 (2013).
23. Balvanera, P. *et al.* Essential ecosystem service variables for monitoring progress towards sustainability. *Curr. Opin. Environ. Sustain.* **54**, 101152 (2022).
24. Beckmann, M. *et al.* Conventional land-use intensification reduces species richness and increases production: A global meta-analysis. *Glob. Change Biol.* **25**, 1941–1956 (2019).
25. Scholefield, P. *et al.* Modelling the European Farmland Bird Indicator in response to forecast land-use change in Europe. *Ecol. Indic.* **11**, 46–51 (2011).
26. Swaay, C., Nowicki, P., Settele, J. & Van Strien, A. Butterfly monitoring in Europe: Methods, applications and perspectives. *Biodivers. Conserv.* **17**, 3455–3469 (2008).
27. Zurell, D. *et al.* A standard protocol for reporting species distribution models. *Ecography* **43**, 1261–1277 (2020).
28. Martínez-Núñez, C., Martínez-Prentice, R. & García-Navas, V. Land-use diversity predicts regional bird taxonomic and functional richness worldwide. *Nat. Commun.* **14**, 1320 (2023).
29. Gonzalez, A., Chase, J. M. & O'Connor, M. I. A framework for the detection and attribution of biodiversity change. *Philos. Trans. R. Soc. B Biol. Sci.* **378**, 20220182 (2023).
30. Alblas, E. & Zeben, J. van. Public participation for a greener Europe: The potential of farmers in biodiversity monitoring. *Land Use Policy* **127**, 106577 (2023).
31. Pe'er, G. *et al.* How can the European Common Agricultural Policy help halt biodiversity loss? Recommendations by over 300 experts. *Conserv. Lett.* **15**, e12901 (2022).
32. Hölting, L. *et al.* Co-design: Working with farmers in Europe to halt the loss of biological diversity. *Ecol. Solut. Evid.* **3**, e12169 (2022).
33. Billaud, O., Vermeersch, R.-L. & Porcher, E. Citizen science involving farmers as a means to document temporal trends in farmland biodiversity and relate them to agricultural practices. *J. Appl. Ecol.* **58**, 261–273 (2021).
34. Kühl, H. S. *et al.* Effective Biodiversity Monitoring Needs a Culture of Integration. *One Earth* **3**, 462–474 (2020).
35. Gonzalez, A. *et al.* A global biodiversity observing system to unite monitoring and guide action. *Nat. Ecol. Evol.* (2023) doi:10.1038/s41559-023-02171-0.
36. Steinke, J., van Etten, J. & Zelan, P. M. The accuracy of farmer-generated data in an agricultural citizen science methodology. *Agron. Sustain. Dev.* **37**, 32 (2017).
37. de la Torre, J. A. *et al.* There will be conflict – agricultural landscapes are prime, rather than marginal, habitats for Asian elephants. *Anim. Conserv.* **24**, 720–732 (2021).
38. Maas, B., Fabian, Y., Kross, S. M. & Richter, A. Divergent farmer and scientist perceptions of agricultural biodiversity, ecosystem services and decision-making. *Biol. Conserv.* **256**, 109065 (2021).
39. Stumph, B. *et al.* Detecting Invasive Insects with Unmanned Aerial Vehicles. in *2019 International Conference on Robotics and Automation (ICRA)* 648–654 (2019). doi:10.1109/ICRA.2019.8794116.
40. Keshet, D., Brook, A., Malkinson, D., Izhaki, I. & Charter, M. The Use of Drones to Determine Rodent Location and Damage in Agricultural Crops. *Drones* **6**, (2022).
41. Friesenhahn, B. A. *et al.* Using drones to detect and quantify wild pig damage and yield loss in corn fields throughout plant growth stages. *Wildl. Soc. Bull.* **47**, e1437 (2023).
42. Abbas, A. *et al.* Drones in Plant Disease Assessment, Efficient Monitoring, and Detection: A Way Forward to Smart Agriculture. *Agronomy* **13**, (2023).
43. Soares, V. H. A., Ponti, M. A., Gonçalves, R. A. & Campello, R. J. G. B. Cattle counting in the wild with geolocated aerial images in large pasture areas. *Comput. Electron. Agric.* **189**, 106354 (2021).
44. Hodgson, J. C. *et al.* Drones count wildlife more accurately and precisely than humans.

Methods Ecol. Evol. **9**, 1160–1167 (2018).
45. Zhang, H. *et al.* Commercial drones can provide accurate and effective monitoring of the world's rarest primate. *Remote Sens. Ecol. Conserv.* **n/a**, (2023).
46. Stokes, H. J., Mortimer, J. A., Laloë, J.-O., Hays, G. C. & Esteban, N. Synergistic use of UAV surveys, satellite tracking data, and mark-recapture to estimate abundance of elusive species. *Ecosphere* **14**, e4444 (2023).
47. Sharma, L. K., Bu, H., Franzen, D. W. & Denton, A. Use of corn height measured with an acoustic sensor improves yield estimation with ground based active optical sensors. *Comput. Electron. Agric.* **124**, 254–262 (2016).
48. Fischer, S., Edwards, A. C., Garnett, S. T., Whiteside, T. G. & Weber, P. Drones and sound recorders increase the number of bird species identified: A combined surveys approach. *Ecol. Inform.* **74**, 101988 (2023).
49. Colaço, A. F., Molin, J. P., Rosell-Polo, J. R. & Escolà, A. Application of light detection and ranging and ultrasonic sensors to high-throughput phenotyping and precision horticulture: current status and challenges. *Hortic. Res.* **5**, 35 (2018).
50. Rostami, B. & Nansen, C. Application of active acoustic transducers in monitoring and assessment of terrestrial ecosystem health—A review. *Methods Ecol. Evol.* **13**, 2682–2691 (2022).
51. Kestel, J. H. *et al.* Applications of environmental DNA (eDNA) in agricultural systems: Current uses, limitations and future prospects. *Sci. Total Environ.* **847**, 157556 (2022).
52. Aucone, E. *et al.* Drone-assisted collection of environmental DNA from tree branches for biodiversity monitoring. *Sci. Robot.* **8**, eadd5762 (2023).
53. Rai, N. *et al.* Applications of deep learning in precision weed management: A review. *Comput. Electron. Agric.* **206**, 107698 (2023).
54. Angelini, F. *et al.* Robotic Monitoring of Habitats: The Natural Intelligence Approach. *IEEE Access* **11**, 72575–72591 (2023).
55. Vihervaara, P. *et al.* How Essential Biodiversity Variables and remote sensing can help national biodiversity monitoring. *Glob. Ecol. Conserv.* **10**, 43–59 (2017).
56. Basso, B. & Antle, J. Digital agriculture to design sustainable agricultural systems. *Nat. Sustain.* **3**, 254–256 (2020).
57. Fahrig, L. How much habitat is enough? *Biol. Conserv.* **100**, 65–74 (2001).
58. Gruber, T. *et al.* Cultural change in animals: a flexible behavioural adaptation to human disturbance. *Palgrave Commun.* **5**, 64 (2019).
59. Jess, T. Farmer Unknowingly Captures Animal Extinct in Region for Over a Century. *News Week 90* (2023).
60. Basile, M., Asbeck, T., Cordeiro Pereira, J. M., Mikusiński, G. & Storch, I. Species co-occurrence and management intensity modulate habitat preferences of forest birds. *BMC Biol.* **19**, 210 (2021).
61. Magnano, A. L. *et al.* Relationships between land-use intensity, woody species diversity, and carbon storage in an arid woodland ecosystem. *For. Ecol. Manag.* **529**, 120747 (2023).
62. Meyer, C., Kreft, H., Guralnick, R. & Jetz, W. Global priorities for an effective information basis of biodiversity distributions. *Nat. Commun.* **6**, 8221 (2015).
63. Rocha-Ortega, M., Rodriguez, P. & Córdoba-Aguilar, A. Geographical, temporal and taxonomic biases in insect GBIF data on biodiversity and extinction. *Ecol. Entomol.* **46**, 718–728 (2021).
64. Mentges, A., Blowes, S. A., Hodapp, D., Hillebrand, H. & Chase, J. M. Effects of site-selection bias on estimates of biodiversity change. *Conserv. Biol.* **35**, 688–698 (2021).
65. Bowler, D. E. *et al.* Temporal trends in the spatial bias of species occurrence records. *Ecography* **2022**, e06219 (2022).
66. World Bank. World Development Indicators. *World Bank database*

https://data.worldbank.org/ (2023).
67. Dara, A. *et al.* Mapping the timing of cropland abandonment and recultivation in northern Kazakhstan using annual Landsat time series. *Remote Sens. Environ.* **213**, 49–60 (2018).
68. Nzabarinda, V. *et al.* Impact of cropland development intensity and expansion on natural vegetation in different African countries. *Ecol. Inform.* **64**, 101359 (2021).
69. Putzel, L. *et al. Chinese trade and investment and the forests of the Congo Basin: Synthesis of scoping studies in Cameroon, Democratic Republic of Congo and Gabon*. (2011).
70. Mahlich, L., Jung, C. & Schaldach, R. The Biodiversity Footprint of German Soy-Imports in Brazil. *Sustainability* **14**, (2022).
71. Seiler, N. & Robbins, M. M. Factors Influencing Ranging on Community Land and Crop Raiding by Mountain Gorillas. *Anim. Conserv.* **19**, 176–188 (2016).
72. Hutson, A. M. & Paunovic, M. *Myotis alcathoe*. https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T136680A518740.en. Accessed on 20 December 2023. (2016).
73. CBD. *Convention on Biological Diversity : text and annexes*. (2011).
74. Vihervaara, P., Mathieu, B., Mandon, C., Suni, T. & Lipsanen, A. *Report on the mapping of ministries, agencies and organisations that fund and steer national/ sub-national biodiversity monitoring schemes,*. (2023).
75. Tydecks, L., Jeschke, J. M., Wolf, M., Singer, G. & Tockner, K. Spatial and topical imbalances in biodiversity research. *PloS One* **13**, e0199327 (2018).
76. Morelli, F. *et al.* Detection Rate of Bird Species and What It Depends on: Tips for Field Surveys. *Front. Ecol. Evol.* **9**, (2022).
77. WWF. *Derstroying native vegetation to produce beef and soy is the main driver affecting biodiversity in the Cerrado and Amazon*. (2021).
78. FAOSTAT. *Agricultural production statistics: 2000-2020*.
79. Lesiv, M. *et al.* Estimating the global distribution of field size using crowdsourcing. *Glob. Change Biol.* **25**, 174–186 (2019).
80. Monsarrat, S., Boshoff, A. F. & Kerley, G. I. H. Accessibility maps as a tool to predict sampling bias in historical biodiversity occurrence records. *Ecography* **42**, 125–136 (2019).
81. European Commission *et al. European Red List of insect taxonomists*. (Publications Office of the European Union, 2022). doi:10.2779/364246.
82. Heilig, G. K. *World urbanization prospects: the 2018 revision*. (2019).
83. Juffe-Bignoli, D. *et al.* Assessing the Cost of Global Biodiversity and Conservation Knowledge. *PLOS ONE* **11**, 1–22 (2016).
84. World Bank. *Poverty and Shared Prosperity 2022: Correcting Course.* (2022).
85. FAO. *Agricultural investments and capital stock: 2011–2021 Global and Regional trends*. (2023).
86. Gebbers, R. & Adamchuk, V. I. Precision Agriculture and Food Security. *Science* **327**, 828–831 (2010).
87. Zezza, A., Gourlay, S. & Molini, V. Closing the data gap in agriculture through sustainable investment in the data value chain: Realizing the vision of the 50x2030 Initiative. *Stat. J. IAOS* **38**, 1–6 (2022).
88. Titus, T. N. *et al.* A roadmap for planetary caves science and exploration. *Nat. Astron.* **5**, 524–525 (2021).
89. Pillay, R. *et al.* Tropical forests are home to over half of the world's vertebrate species. *Front. Ecol. Environ.* **20**, 10–15 (2022).
90. Top, J., Janssen, S., Boogaard, H., Knapen, R. & Şimşek-Şenel, G. Cultivating FAIR principles for agri-food data. *Comput. Electron. Agric.* **196**, 106909 (2022).
91. Ewert, F. *et al.* FAIRagro - A FAIR Data Infrastructure for Agrosystems (proposal). (2023)

doi:10.5281/zenodo.8366884.

92. Basel, A. M., Nguyen, K. T., Arnaud, E. & Craparo, A. C. W. The foundations of big data sharing: A CGIAR international research organization perspective. *Front. Environ. Sci.* **11**, (2023).
93. Home, R., Balmer, O., Jahrl, I., Stolze, M. & Pfiffner, L. Motivations for implementation of ecological compensation areas on Swiss lowland farms. *J. Rural Stud.* **34**, 26–36 (2014).
94. Bartkowski, B. *et al.* Adoption and potential of agri-environmental schemes in Europe: Cross-regional evidence from interviews with farmers. *People Nat.* **5**, 1610–1621 (2023).
95. Pringle, S. *et al. Robotics and Autonomous Systems for Environmental Sustainability*. (2023).
96. UNCTAD. *Technology and innovation report 2023 : opening green windows : technological opportunities for a low-carbon world*. (2023).
97. Fan, D. *et al.* Weight of Evidence Method and Its Applications and Development. *Procedia Environ. Sci.* **11**, 1412–1418 (2011).
98. Moersberger, H. *et al.* Europa Biodiversity Observation Network: User and Policy Needs Assessment. *ARPHA Prepr.* **3**, ARPHA Preprints (2022).
99. Priyadarshana, T. S., Lee, M.-B., Ascher, J. S., Qiu, L. & Goodale, E. Crop heterogeneity is positively associated with beneficial insect diversity in subtropical farmlands. *J. Appl. Ecol.* **58**, 2747–2759 (2021).
100. Prasad, P. V. V. *et al.* Patterns of investment in agricultural research and innovation for the Global South, with a focus on sustainable agricultural intensification. *Front. Sustain. Food Syst.* **7**, (2023).

SUPPLEMENTARY INFORMATION

METHODS

Analysis of gaps and biases in species observations

Species observations

We mobilised all species observation data collected between 2015 and 2019 from the Global Biodiversity Information Facility (GBIF) to analyse potential biases in biodiversity observations in food production systems. We focused on the period 2015-2019 to align our analysis of species observation data with concurrent global land cover data (see next section). As a basis for the detection of spatial biases, we derived global layers of per-cell species observation counts at a resolution of 0.25°. This resolution offers sufficient detail to detect both regional and global biases and data gaps, as well as general land system patterns. It also facilitates the visualisation of spatially sparse species observations. To derive these layers, we used the function *occ_count()* from the R package *rgbif* to count species observations within each cell. Using this approach, we first derived annual layers(2015 to 2019), to compare annual sampling effort per cell with corresponding cropland cover percentages.

Measurements of land system biases

To detect land system biases in species observations, we assessed how the presence and frequency of observations interacted with cropland cover and cropland-cover change. While the first variable provided information on site selection biases, the second one informed on how these biases affected the representation of land system changes in species observations. Areas experiencing such changes provide ideal opportunities for monitoring immediate and long-term biodiversity-yield interactions under different land management practices.

We used the Copernicus Land Cover dataset[1] due to its global extent, high spatial and thematic resolution and annual coverage (which enabled generalised statements on the persistence of site selection biases). It also has the major advantage that it directly reports per-cell proportions of land cover (which allows inferring cell-area proportions to account for regional variations in cell sizes). Consistent with our analysis of species observations, we aggregated the 100-m resolution annual cropland layers provided by the CLC dataset to a 0.25° resolution through averaging (hereafter 'cropland cover'), to match the resolution of the species observation layers. We then calculated the mean cropland cover between 2015 and 2019, and the change in cropland cover between the beginning and end of this period (hereafter 'cropland-cover change').

We analysed these data in several ways. Firstly, we compared the distribution of species observations with the corresponding cropland cover (**Fig. 3a** in the main manuscript), and against the cropland-cover change measured between 2015 and 2019 (**Fig. 3b** in the main manuscript). This enabled us to detect whether the site selection biases persisted from year to year. Secondly, we focused on cells covered by cropland at any time between 2015 and 2019. For each aggregated layer (i.e., mean cropland cover, cropland-cover change), we compared the cumulative distributions of all cells to those of cells with species observations using Kolmogorov-Smirnov[2], a nonparametric test for equality of continuous and one-dimensional probability distributions (**Fig. 3d-e** in the main manuscript). We used a weighted version of this test, as implemented in the function *ks_test()* of the R package *Ecume*[3]. For the global distributions, we used a static weight of 1 for every cell. For cells with species observations, the weight in each cell was the corresponding number of observations between 2015 and 2019.

We rejected the null hypothesis that both cell samples came from the same population if the p-value was below 0.005, following recommendations on significance testing[4].

Implications of biases in biodiversity monitoring

We compared the observed biases and gaps in species observations with data on the combined species richness of birds, amphibians, mammals, and reptiles provided by the IUCN Red List of species[5] as a global grid with a resolution of 5km. Richness in this dataset informs on the number of species potentially occupying a given cell and is estimated by summing the range maps of individual species on a cell-by-cell basis. In line with our analysis of land system biases, we aggregated these data to a resolution of 0.25° through averaging. First, focusing on cells with cropland cover, we analysed differences between the global cumulative distribution of richness across cells with cropland cover and the cumulative distribution drawn from the subset of cells with species observations. Again, we used the weighted Kolmogorov-Smirnov test (**Fig. 4a** in the main manuscript). As described previously, the weights corresponded to the number of species observations per cell, and unobserved cells were given a constant weight of 1. Second, we mapped richness for cells without species observations and with changes in cropland cover to describe global patterns of undersampling (**Fig. 4b** in the main manuscript).

Third, we compared species observations with knowledge of species habitat preferences to determine how observation gaps and biases may have led to misconceptions of species ecologies (**Fig. 4c** in the main manuscript). To do so, we retrieved species-specific assessments of habitat preferences from the IUCN Red List of Threatened Species. From the list of species with threat assessments, we distinguished those that were not associated with cropland-related habitats as described in the IUCN Habitat class scheme[6] (classes 14.1, 14.4, 15.7, and 15.8).

For each species in this group that was observed between 2015 and 2019, we collected all corresponding GBIF records without coordinate issues (e.g., coordinates associated with country centroids or scientific institutions). For each observation, we then extracted the proportion of cropland based on the 100-m resolution Copernicus Land Cover data[1]. During the extraction process, we considered all cells within the radius defined by the coordinate uncertainty. In case this was not directly provided along with the species observation, we inferred it from the number of decimal places in the coordinates. If there were ≥5 decimal places, the error was assumed to be ≤1 m. As the number of decimal places decreases, the error increases by a factor of 10 for each decimal place. If there were no decimal places, the error was assumed to be 1 degree. Based on the extracted data, we quantified the percentage of observations per species overlapping cells with cropland-cover.

Assessing sources of bias

We used data from two regions to provide examples of potential reasons for site selection biases. Specifically, we analysed data from Brazil to evaluate revisits to species observation sites relative to agricultural expansion (to inform on site selection biases motivated by an interest in drivers of biodiversity loss). We further analysed data from California to evaluate the geographic locations of species observations within agricultural land parcels (to inform on the impacts of privatisation and access to land).

Site selection biases

We compared changes in agriculture with the spatial and temporal distribution of species observations to: *i)* assess whether species observations occur in areas of agricultural expansion or abandonment (**Fig. 5a-b** in the main manuscript); *ii)* detect revisits to these sites, the absence of which implies the lack of a systematic biodiversity assessment (**Fig. 5c** in the main manuscript). We focused this assessment on the legal Amazonia, anticipating that the active deforestation frontiers driven by agricultural expansion would motivate biodiversity surveys.

We mapped changes in agriculture using a national land-cover dataset for Brazil[7]. These data is provided at a 30-m resolution, but we aggregated it to a 1 km resolution expressing per-cell proportions of 30-m cells classified as 'agriculture'. In the reference land cover dataset, 'agriculture' included cells classified as 'pastures', 'sugar cane', 'palm oil', 'soybean', 'rice', 'other temporary crops', 'coffee', 'citrus', 'other perennial crops', or 'cotton'. Although the land cover dataset is available annually between 1985 and 2022, we focused on the period 2000-2019. The start of this period coincides with massive improvements in Landsat data frequency, allowing for more confident land cover mapping[8]. The end precedes the start of the global COVID-19 pandemic, which negatively impacted the frequency and quality of species observations[9].

In our comparison of species observations with changes in agriculture, we additionally considered changes in travel time. We assumed that the expansion of agriculture would be accompanied by the development of the infrastructure needed for the transport of the produced commodities. As new roads connect previously inaccessible land, we expected new species observations in these lands. To measure changes in travel time, we used global data with a 1-km resolution, mapping the travel time to the nearest city with at least 50,000 inhabitants from any given cell[10]. These data are available for the years 2000 and 2015, and we subtracted them to derive a layer on long-term changes in travel time requirements.

To assess the effect of agricultural land change and travel time on species observations, we mobilised all records collected across the legal Amazonia between 2000 and 2019[11]. We then translated these data into 'observation sites', corresponding to unique 1 km cells with one or more species observations in a given year. This aggregation step is more meaningful than the original observation coordinates. While the coordinates of specific observations may change from year to year, sites of interest for biodiversity monitoring are likely to be visited in multiple years. Furthermore, this aggregation allowed us to focus on the geographical location rather than on the frequency of species observations, which may vary widely.

Impacts of privatisation and access to land

We demonstrated the effect of privatisation of agricultural land parcels on species observations (**Fig. 5d-f** in the main manuscript). To do so, we compared GBIF data for 2019[12] with land cadastre data for California, USA, for the same year[13]. We chose this region because of the availability of annual, high-resolution, and authoritative land cover data, the relatively high frequency of species observations, and the openness of cadastre data.

To limit our analysis to observations made in agricultural areas, we removed GBIF data not overlapping land covers described as 'cultivated crops' or 'hay/pastures' in the authoritative national land cover data for 2019[14]. We used these data for filtering rather than the land cadastre data because species observations might occur at field boundaries. In turn, the classification of each cell is sensitive to the dominant composition of the land surface, meaning that observations made just outside of land parcels might still be classified as agriculture. In addition, land cover data provided us with confirmation that a land parcel was actively managed in the observation year, which would make it more difficult to access than fallow land. Additionally, we excluded species observations made outside of a 100-m radius around agricultural land parcels to account for species observations made on cropland within urban areas. This is because land parcels used for agriculture are not distinguished in the cadastral data if located within urban areas.

For each species observation, we calculated the minimum distance to the centroid of the nearest land parcel as well as the size of that parcel. We then correlated these data to assess whether distances were related to parcel sizes (**Fig. 5d** in the main manuscript). A high correlation would hence indicate that species observations are persistently collected along or near the edges of land parcels (as seen in **Fig. 5e** in the main manuscript). Because land parcels can have various shapes, prior to our correlation analysis, we translated parcel sizes into their approximate radius, estimated as □□□□(□/□), where a is the area of the parcel in ha.

As a follow-up analysis, we evaluated how the position of observers may have limited their ability to detect species within land parcels (**Fig. 5f** in the main manuscript). We used the function *viewshed()* from the R package *terra* that estimates how high above the ground a species must be to be perceived by the observer. This estimate is based on the difference in elevation between the observers' location and the surrounding terrain. Here, we assumed an observer height of 1,80 m and a target height of 1 m (e.g., a wild pig, a species which is invasive to California). The data on the elevation of the terrain originates from a digital elevation model with a 1-m resolution provided by the United States Geological Survey[15].

REFERENCES

1. Buchhorn, M. *et al.* Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual. (2021) doi:10.5281/zenodo.4723921.
2. Massey, F. J. The Kolmogorov-Smirnov Test for Goodness of Fit. *J. Am. Stat. Assoc.* **46**, 68–78 (1951).
3. Bezieux, H. R. de. Ecume: Equality of 2 (or k) Continuous Univariate and Multivariate Distributions. (2021).
4. Ioannidis, J. P. A. The Proposal to Lower P Value Thresholds to .005. *JAMA* **319**, 1429–1430 (2018).
5. Lumbierres, M. *et al.* Area of Habitat maps for the world's terrestrial birds and mammals. *Sci. Data* **9**, 749 (2022).
6. IUCN. Habitats Classification Scheme Version 3.1. (2020).
7. MapBiomas. MapBiomas Alert Project V4.0 - Validation and Refinement System for Deforestation Alerts with High Resolution Images. (2021).
8. Remelgado, R., Conrad, C. & Meyer, C. Limitations in historical satellite archives bias SDG monitoring. *preprint* (2023) doi:https://doi.org/10.31223/X5QH37.
9. Hochachka, W. M., Alonso, H., Gutiérrez-Expósito, C., Miller, E. & Johnston, A. Regional variation in the impacts of the COVID-19 pandemic on the quantity and quality of data collected by the project eBird. *Biol. Conserv.* **254**, 108974 (2021).
10. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333 (2018).
11. GBIF. Occurrence Download. (2023) doi:https://doi.org/10.15468/dl.yvjz93.
12. GBIF. Occurrence Download. (2023) doi:https://doi.org/10.15468/dl.9cq346.
13. CNRA. Statewide Crop Mapping. (2019).
14. Yang, L. *et al.* A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS J. Photogramm. Remote Sens.* **146**, 108–123 (2018).
15. USGS. 1 meter Digital Elevation Models (DEMs) - USGS National Map 3DEP Downloadable Data Collection. (2023).