Insect monitoring without pitfalls: seven steps for robust insect sensing systems

Jamie Alison[1✤], Luca Pegoraro[2✤*], Jarrett Blair[1], Yuval Cohen[3,4], Birgen Haest[5], Jacob Idec[6], Jacob Kamminga[7], Jenna Lawson[8], Meng Li[9], Leandro Aparecido Do Nascimento[10], Charlotte L. Outhwaite[11], Benjamin Rutschmann[12], Maximilian Sittinger[13,14], Mariana Abarca[15], Tom August[8], Quentin Geissmann[16], Rob Guralnick[6], Guillaume Mougeot[1], Valentin Ștefan[13,14,17], Kim Bjerge[18], Mikkel Brydegaard[9,19], Toke Thomas Høye[1,20], Eva Knop[21,22], William Kunin[23], Moritz Lürig[6], David Roy[8], Asger Svenning[1], Roel van Klink[14,17]

✤These authors contributed equally to this work
*Corresponding author: luca.pegoraro@wsl.ch

[1]Department of Ecoscience, Aarhus University, Denmark
[2]Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf, Switzerland
[3]Northern Agriculture R&D, MIGAL – Galilee Research Institute, Kiryat Shmona, Israel
[4]Department of Evolutionary and Environmental Biology, Faculty of Natural Sciences, University of Haifa, Haifa, Israel
[5]Bird Migration Unit, Swiss Ornithological Institute, Seerose 1, 6204 Sempach, Switzerland
[6]Florida Museum of Natural History, University of Florida, Gainesville, FL, USA
[7]University of Twente, Enschede, The Netherlands
[8]UKCEH, Wallingford, Oxfordshire, OX10 8BB, UK
[9]Department of Physics, Lund University, Sölvegatan 14C, 22363 Lund, Sweden
[10]Map of Life, New Haven, CT, USA
[11]Institute of Zoology, Zoological Society of London, Outer Circle, Regent's Park, London, NW1 4RY, UK
[12]Agroecology and Environment, Agroscope, Reckenholzstrasse 191, 8046, Zurich, Switzerland
[13]Helmholtz Centre for Environmental Research (UFZ), Permoserstraße 15, 04318 Leipzig, Germany
[14]German Center for Integrative Biodiversity Research Halle-Jena-Leipzig (iDiv), Puschstrasse 4, 04103 Leipzig, Germany
[15]Department of Biological Sciences, Smith College, Northampton, MA, USA
[16]Center for Quantitative Genetics and Genomics, Aarhus University, Denmark
[17]Martin-Luther University Halle-Wittenberg, Universitätsplatz 10, 06108, Halle (Saale), Germany
[18]Department of Electrical and Computer Engineering, Aarhus University, Denmark
[19]Norsk Elektro Optikk, Østensjøveien34, 0667 Oslo, Norway
[20]Arctic Research Centre, Aarhus University, Denmark
[21]Faculty of Biology and Pre-Clinical Medicine, University of Regensburg, Germany
[22]Department of Evolutionary Biology and Environmental Sciences, University of Zurich, Switzerland
[23]School of Biology, University of Leeds, Leeds LS2 9JT, UK

Abstract

Data shortages fuel controversy about an ongoing insect biodiversity crisis. Insects are immensely diverse and functionally critical for ecosystems, yet data on their trends remain patchy and biased. Sensors, ranging from camera-equipped light traps to weather radar stations, are set to transform data collection in entomology. Meanwhile, AI models that extract biological information from sensors are improving at a startling rate. Realising the potential of automated monitoring means progressing from proof-of-concept studies to scalable insect sensing systems. However, stakeholders face severe operational challenges when adopting a growing suite of sensors, models, and protocols for insect surveillance. Deployment of devices is not well co-ordinated, while the risks of relying on AI are overlooked or understated. To achieve monitoring goals, common pitfalls related to sensors and AI need to be exposed and avoided. Here, we trace a seven-step path towards an effective transnational rollout of insect sensing systems. Step (1) reviews strengths, weaknesses and synergies across visual, acoustic, radar and photonic sensors; (2) confronts species determination—a key challenge for sensors and AI—suggesting how to improve identification and make use of uncertain data; (3) promotes creating and sharing standardised, labelled data, offering ecological insights and opportunities to train and evaluate AI; (4) ensures AI is used to complement other resources, both human and digital, given it is not always the best tool for the job; (5) highlights how and why AI models can fail, calling for shrewd approaches to model training and routine evaluations; (6) aims to break barriers to wider uptake of technologies through knowledge sharing, affordable design principles, and equitable computing infrastructures. Finally, (7) emphasizes that any revolution in insect monitoring must be grounded in good sampling design, with established monitoring schemes at its core. We set a trajectory for coordinated development of insect sensing systems, focussing not only on technical performance, but on integration with human expertise, case-based evaluation and harmonisation with historical long-term datasets. We address fundamental challenges of sensors and AI for biodiversity monitoring, producing recommendations that apply to all branches of the tree of life.

Keywords: arthropods, artificial intelligence, computer vision, image classification, invertebrates, machine learning, object detection, pollinators, remote sensing, signal processing

Introduction

Insects form the bulk of macroscopic biodiversity on earth (Stork, 2018). They are extraordinarily diverse, but also functionally indispensable, with roles as pollinators, decomposers, predators, pests, and pest-control agents (Noriega *et al.*, 2018). As such, recent reports of insect decline have raised both economic and conservation concerns (Hallmann *et al.*, 2017; Wagner, 2020). However, the evidence is mixed: a global meta-analysis suggests decreases in terrestrial insects and increases in freshwater insects (van Klink *et al.*, 2020). Furthermore, the magnitude and generality of trends across biomes, regions, and taxa is up for debate (Neff *et al.*, 2022). One thing is clear, though: available data only glimpse the complexity of insect life (Thomas, Hefin Jones & Hartley, 2019), and are heavily biased towards Europe, North America, and certain taxa (e.g. butterflies; Sánchez-Bayo & Wyckhuys, 2019). Meanwhile, international policy instruments such as the Convention on Biological Diversity (United Nations, 1992), UN sustainable development goals (United Nations, 2015), and the EU nature restoration regulation (Directorate-General for Environment, 2022) demand comprehensive monitoring of the diversity and abundance of insects, as well as underlying drivers of change.

Scaling of data collection needs to be cost-effective, as there are limited resources available for insect monitoring (Geijzendorffer *et al.*, 2016). Traditional monitoring methods, including malaise traps, pan traps, pitfall traps, light traps, and visual surveys (Southwood, 1966; McCravy, 2018; Montgomery *et al.*, 2021), have built baseline datasets and fundamental understanding of insect status and trends (Conrad *et al.*, 2006; Brooks *et al.*, 2012; Hallmann *et al.*, 2017). However, traditional monitoring schemes are difficult to expand, especially in tropical regions; manual identification is skilled and laborious work, and local expertise is often a limiting factor (Sánchez Herrera *et al.*, 2024). Furthermore, most traditional methods, such as visual surveys and traps, suffer from detectability and observer biases (e.g. Brereton *et al.*, 2011), while others involve excessive lethal sampling (Lövei & Ferrante, 2024). Methodological continuity is a priority for any long-term monitoring effort (van Klink, 2024). Given their limitations in terms of human workload, scalability, and lethality, traditional methods alone will not be sufficient to address major challenges in entomology (Didham *et al.*, 2020).

Novel technologies offer opportunities for more scalable, standardised and non-lethal insect monitoring (van Klink *et al.*, 2022; Chua *et al.*, 2023; Lövei & Ferrante, 2024; Dyer *et al.*, 2024), and promise datasets that will transform our ecological understanding (Hartig *et al.*, 2024; Gillespie *et al.*, 2025). Molecular methods allow identification of insects to high taxonomic resolution without expert identification (Chua *et al.*, 2023). DNA metabarcoding, in particular, involves extracting, amplifying and sequencing DNA from a sample, followed by taxonomic assignment of sequences according to reference libraries (Chua *et al.*, 2023). This approach has successfully identified insects in bulk samples of specimens (Mata *et al.*, 2021; Buchner *et al.*, 2025), as well as samples of water, soil, air, gut contents, or flowers (Bohmann *et al.*, 2014; Thomsen & Sigsgaard, 2019; Chua *et al.*, 2023). While molecular methods provide a powerful addition to the insect monitoring toolbox, they also face a number of challenges, such as incomplete reference libraries, taxonomic biases, limited capability to infer abundance, high infrastructure costs, and uncertainty about the spatial and temporal origin of detected DNA (Alberdi *et al.*, 2018; Deagle *et al.*, 2019; Chua *et al.*, 2023). Furthermore, because molecular techniques require collecting physical samples, they are limited by the frequency in which data can be collected and analysed; they typically provide discrete snapshots in time rather than continuous observations.

Meanwhile, sensor-based methods show great promise to monitor insects continuously through time. Cameras, microphones, radar, and lidar can offer real-time, non-invasive data that capture the presence, biomass, movements and behaviours of insects (van Klink *et al.*, 2024). Resulting high-frequency time-series offer novel insights into biotic interactions, such as plant-insect interactions (Naqvi *et al.*, 2022; Mehrotra *et al.*, 2024; Serra-Marin *et al.*, 2025). Furthermore, sensors can be affordable and operate autonomously, offering a very low cost per insect observation (Brydegaard *et al.*, 2024). The volume of data produced by sensors is challenging to work with (Høye *et al.*, 2021), but emerging methods for automation, especially artificial intelligence (AI; Box 1), are relieving bottlenecks in data processing. Clearly, insect sensing systems could drive a step-change in the spatial and temporal scale of ecological monitoring worldwide (Besson *et al.*, 2022; Bauer, Tielens & Haest, 2024).

Box 1. What is AI, in the context of insect monitoring?

Artificial intelligence (AI) broadly refers to computational systems that reproduce human skills, such as thinking, acting or interpreting data (Russell & Norvig, 2010). In the context of ecological monitoring, AI usually refers to ***machine learning*** models that learn complex patterns from training data, producing useful outputs, such as locations and identities of organisms in images.

AI-based tools for insect monitoring are diverse, but share common features. Firstly, they are data hungry; as a rule, AI models are trained on large datasets, usually including annotations that describe the biological content of sensor data. Second, they are highly complex; modern AI models tend to include millions of parameters, allowing them to learn rich representations of training data and make accurate predictions about unseen data. Notably, some data processing techniques useful in the context of entomology are not considered AI—including human processing (Russo *et al.*, 2021) and ***algorithmic processing*** (Lürig *et al.*, 2021).

While public perception of AI has been shaped by generative large language models such as ChatGPT (OpenAI *et al.*, 2023) and text-to-image models such as DALL-E (Ramesh *et al.*, 2022), AI models for insect monitoring have so far been specially trained to work for specific taxa and contexts. The role of generalist ***foundation models*** in insect monitoring is expected to grow rapidly in coming years (Box 5).

However, coordination is needed for insect sensing systems to reach their full potential. Hardware designs and deployments are currently fragmented, and species identification remains unfeasible for many taxa and technologies (Hüppop *et al.*, 2019; Hansen *et al.*, 2020; Chiranjeevi *et al.*, 2025). Many datasets are published, but they are not always easy to find, use, or combine (Schneider *et al.*, 2023). AI models are improving rapidly, but are still not used effectively to complement human expertise. Crucially, AI outputs are often viewed with skepticism, so guidance is needed to deal with AI uncertainty transparently and systematically (Cowans *et al.*, 2022). Furthermore, automated methods developed in high-income countries may not be sufficiently affordable and accessible to support insect monitoring schemes at a global scale (Brydegaard *et al.*, 2024).

Here, we provide seven steps for fruitful implementation of insect sensing systems, whether for small scientific projects or global monitoring initiatives (Fig. 1). We provide guidance for choosing the right sensors (Step 1), and set expectations for dealing with taxonomic uncertainty (Step 2). We promote good practices for generating and sharing data (Step 3), and advocate a cautious approach when integrating AI into a data processing workflow (Step 4). We highlight potential pitfalls when training and using AI, and suggest how to avoid them and train more useful AI models (Step 5). Finally, we consider the global accessibility of insect sensing systems (Step 6), which we see as an extension to—not a replacement for—ongoing monitoring initiatives (Step 7).

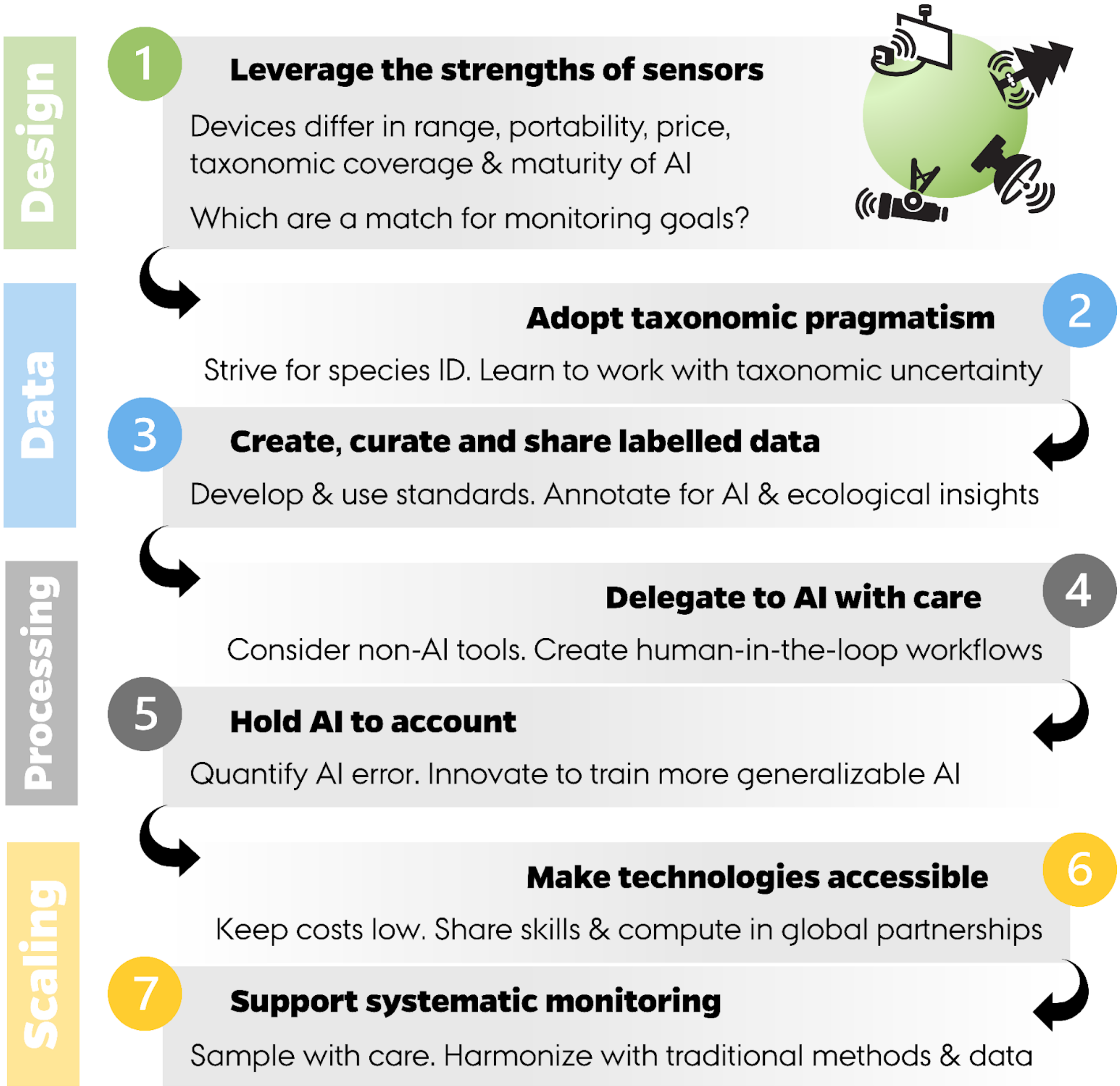

Figure 1. Insect monitoring without pitfalls: Seven steps for robust insect sensing systems.

Scope

This synthetic review focuses on ***insect sensing systems***. We define these as acoustic or electromagnetic wave sensors that continuously record insects *in-situ* (under field conditions), including associated hardware, software and data processing pipelines. We prioritise systems that sense a broad range of insect taxa, while giving some consideration to systems specialised on a few species (e.g. agricultural pests). We do not include systems that tag insects, for example with radio frequency identification (Barlow & O'Neill, 2020). While we centre our review on the topic of insect monitoring, our guidelines (and many of the sensing systems) are applicable for a range of other invertebrates and small animals.

Our scope excludes molecular methods (e.g. DNA metabarcoding), lab-based sensing systems (e.g. imaging museum specimens or dead insects in preservative liquid), and traditional methods (e.g. malaise trapping). However, these methods have their own strengths, provide important context, and can often complement or enable insect sensing systems. For example, molecular methods may provide taxonomic resolution not available to some insect sensing systems. Meanwhile, lab-based sensing can generate vital training data for *in-situ* sensing. Finally, traditional monitoring has massively progressed our understanding of insects and will continue to do so; novel methods must build on this legacy, for example providing a spatio-temporal resolution and coverage that is not reachable with traditional monitoring alone (van Klink *et al.*, 2024).

1. Leverage the strengths of sensors

Sensing systems involving cameras, microphones, radar and lidar offer distinct strengths for automated insect monitoring (Fig. 2). Some sensors excel at detecting and counting large numbers of insects but have low taxonomic resolution, while others provide more detailed information about taxonomic identity or ecological traits. Understanding these differences is crucial to design effective monitoring strategies, and the correct choice of sensor(s) depends greatly on the purpose and goals of the monitoring exercise. Furthermore, the use of attractants can increase the efficiency of data collection by drawing more insects within range of the sensor (Bjerge *et al.*, 2021; Sittinger *et al.*, 2024), however, it can also introduce biases, as taxa often vary in how they sense attractants and move towards them. Below, we examine the relative strengths of four sensor categories (Fig. 2), and consider how diverse technologies can complement each other in insect monitoring.

Sensors excel in terms of consistency and continuity of data collection through time (van Klink *et al.*, 2022), but they also share a set of common challenges. Beyond challenges of species identification and taxonomy (see Step 2), devices face risks of theft, damage or disturbance, particularly in areas densely populated by humans or other animals. These risks scale with the value and visibility of devices, and can be alleviated through strategic placement (Clarin *et al.*, 2014; Meek *et al.*, 2019). Sensors are also operationally complex; off-the shelf systems are not always available or suitable, while more adaptable do-it-yourself (DIY) systems require technical expertise to assemble and use (Bjerge *et al.*, 2021; Droissart *et al.*, 2021; Sittinger *et al.*, 2024). Transport of DIY systems can also be complicated by customs and tax regulations; lithium batteries are often restricted on aeroplanes, while most devices are not compatible with locally available alkaline batteries.

A universal challenge for insect sensing devices is the lack of standardisation *(van Klink, 2024)*. Scalable, commercial solutions for automated insect monitoring are still absent, and most designs rely on custom hardware and have limited user support (Bjerge *et al.*, 2021; e.g. Droissart *et al.*, 2021; Sittinger *et al.*, 2024; Chen *et al.*, 2024). The most established devices for insects include the DIOPSIS (van Klink *et al.*, 2022) and UKCEH AMI (Roy *et al.*, 2024), both of which have a high price tag and require technical skills to maintain. While a range of devices are necessary to monitor different taxa and environments, the fragmentation of system designs hinders standardisation and data integration. Work is ongoing to build an ontology of insect sensing devices, so that their parameters and quirks can be easily and transparently retrieved using data exchange ***standards***.

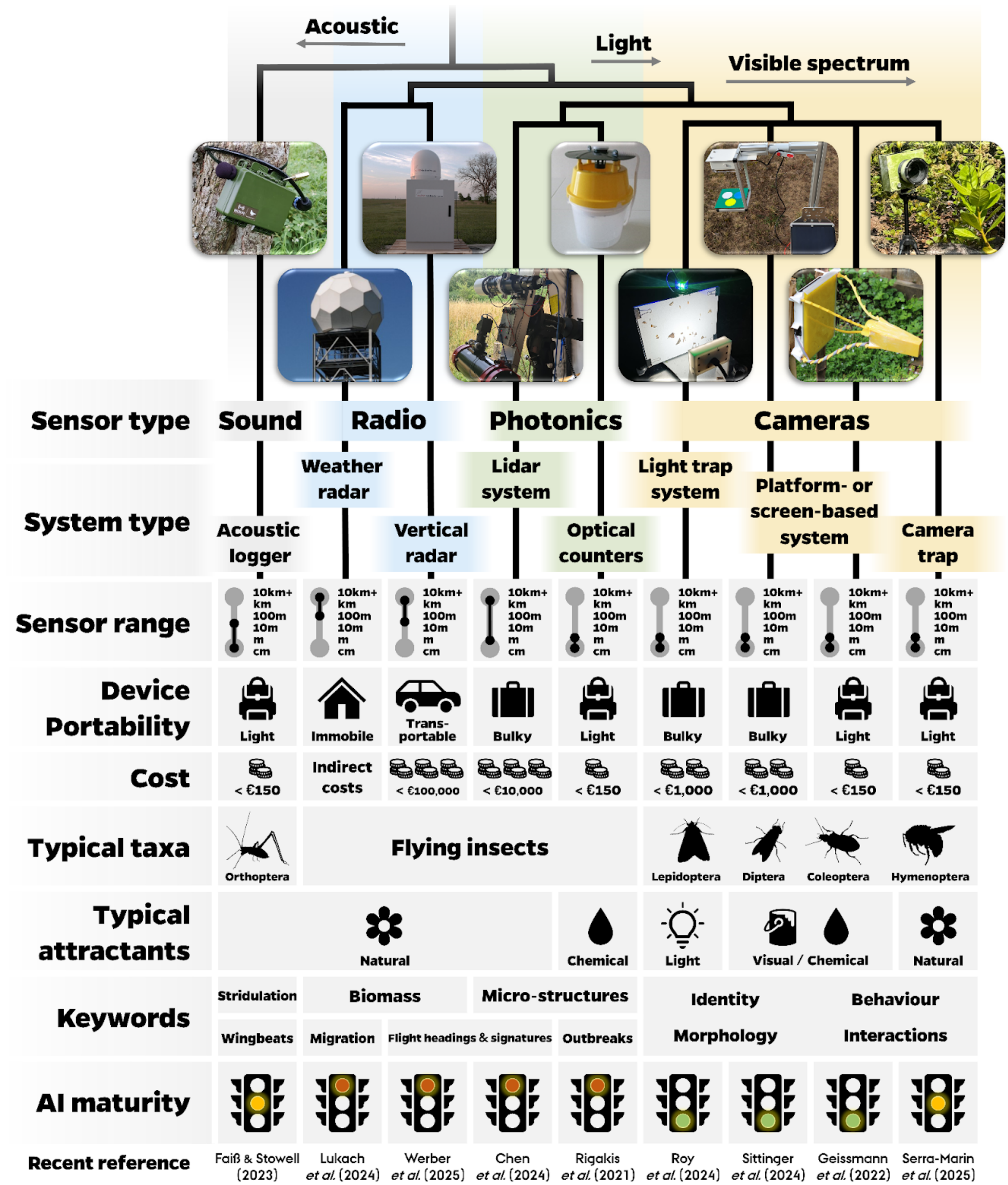

Figure 2. Insect sensing systems vary in their range, portability, cost, taxonomic scope and resolution. Some capture insects in natural contexts, while others lure them with chemical or visual attractants. Data from different sensors provide information on unique aspects of insect life, exemplified here under "Keywords". Some systems can leverage AI more readily than others, and traffic lights represent the maturity of AI systems: **(Red)** Annotated datasets are not widely available, but AI is occasionally used for coarse identification or unsupervised hierarchical clustering. Identification usually relies on ***algorithmic processing***. **(Yellow)** Datasets exist to train AI, but coverage of taxa and contexts is a strong limiting factor. Studies demonstrate reasonable

performance for specialised AI, but models generalise poorly to novel communities and conditions. **(Green)** Large datasets support training of AI models which reliably identify higher insect taxa and distinctive species. Generalisable AI is deployed at scale, and some models are suitable for ***embedded processing***. Image credit from left to right: Jamie Alison (Song Meter Mini 2); Philip Halling (weather radar); Baptiste Schmid (BirdScan MR1 radar); Meng Li (Scheimflug lidar); Rigakis *et al. (e-funnel trap; 2021)*; Jarek Scanferla (UKCEH AMI trap); Max Sittinger (DIY camera trap); Quentin Geissmann (Sticky Pi); Pau Serra-Marin (automated camera system). Insect sensing systems are highly diverse, and these descriptors are indicative, not definitive.

a. Cameras

Modern cameras operate by focusing light through a lens onto an image sensor, converting it into an electrical signal, and storing it as a digital image or video. Image quality depends on lens type, sensor resolution, and onboard image processing, all of which affect the capacity to detect and identify insects or interpret their behaviours (Seimandi-Corda, Hood & Cook, 2024; Bjerge, Karstoft & Høye, 2025).

Camera setups are typically tailored to specific goals, such as observing flowers (Droissart *et al.*, 2021; Alison *et al.*, 2022), nests (Calvus *et al.*, 2025), screens or platforms (Bjerge *et al.*, 2021; Sittinger *et al.*, 2024) or baited locations (Preti, Verheggen & Angeli, 2021). They can also be connected to traditional capture methods, including sticky cards or Malaise traps (Geissmann *et al.*, 2022; Chiavassa *et al.*, 2024). Suitable devices range from trail and compact cameras (Steen, 2017; Naqvi *et al.*, 2022) to action and security cameras (Seimandi-Corda *et al.*, 2024; Varga-Szilay, Szövényi & Pozsgai, 2024), smartphones (Ratnayake, Dyer & Dorin, 2021; Ștefan *et al.*, 2025b), and DIY builds based on single-board computers like Raspberry Pi (Bjerge *et al.*, 2021; Droissart *et al.*, 2021; Geissmann *et al.*, 2022; Sittinger *et al.*, 2024; Wittmann *et al.*, 2024; Serra-Marin *et al.*, 2025; Szczygieł, Dent & Quitmeyer, 2025) or microcontrollers (Darras *et al.*, 2024). The DIY options support add-ons like environmental sensors, light sources, solar panels, and customisable software. Image capture can be scheduled, or triggered by motion or on-device AI-based detection (Steen, 2017; Gardiner, Rowlands & Simmons, 2025b).

Camera traps excel at non-invasive, continuous monitoring of insect presence, activity or abundance, phenology, behaviour and morphology, particularly at times or in locations that are difficult for human observers to reach. They support studies on pollination, predation, and habitat use (Droissart *et al.*, 2021; Bjerge, Mann & Høye, 2022; Ștefan *et al.*, 2025b), and enable fine-scale temporal analyses (Steen, 2017; Geissmann *et al.*, 2022; Sittinger *et al.*, 2024). High-resolution imagery allows species-level classification where distinctive features are visible (Chiranjeevi *et al.*, 2025; Bjerge *et al.*, 2025). It also enables a potential by-catch of ecological and environmental information (Pernat *et al.*, 2024; Alison *et al.*, 2024). However, detection and classification are usually less reliable for smaller insects.

Commercial trail cameras often lack fine control over scheduling and are not optimised to detect small, fast-moving insects (Pegoraro *et al.*, 2020). DIY systems, while adaptable, usually require expertise for assembling components and configuring software, posing a barrier for non-specialists. Open hardware platforms and community-developed projects (Droissart *et al.*, 2021; e.g. Sittinger *et al.*, 2024; Szczygieł *et al.*, 2025) enable flexible, modular designs and can improve reproducibility and uptake in the community, especially in conservation and citizen science efforts (Sheard *et al.*, 2024). Most systems are stationary, capturing imagery of an area smaller than 1m^2.

Video and image data can quickly exceed storage and processing capacity, and scheduled recording may capture long stretches of irrelevant footage (Pegoraro *et al.*, 2020). On the other hand, motion-triggers or real-time detection models can reduce data volume (O'Shea-Wheller *et al.*, 2024; Sittinger *et al.*, 2024; Darras *et al.*, 2024; Bjerge *et al.*, 2025), but may do so with a bias. Detection tends to be easier against simple backgrounds than in complex natural settings (Ștefan *et al.*, 2025a). Detection models are often task-specific, but new generalisable AI tools are emerging (Svenning *et al.*, 2025) as are platforms to increase their accessibility (e.g. Antenna). Classification models, powered by citizen science datasets, are able to identify some large-bodied taxa to species-level (Spiesman *et al.*, 2021).

b. Acoustics

Passive acoustic monitoring (PAM) is the use of autonomous sound recording units to record vibrations in air (Ross *et al.*, 2023), water (Lamont *et al.*, 2022), or solid substrates such as plant tissues (Mehrotra *et al.*, 2024). It is widely used in terrestrial and aquatic ecosystems to detect and analyse animal sounds (Deichmann *et al.*, 2018; Gibb *et al.*, 2019; Darras *et al.*, 2025). Advances in affordable, programmable recording units and software tools have contributed to a surge in PAM studies (Hill *et al.*, 2018; Sugai *et al.*, 2019; Kohlberg, Myers & Figueroa, 2024) although invertebrates remain underrepresented, with only 23 of 460 terrestrial PAM studies targeting insects (see also Mankin *et al.*, 2011; reviewed by Sugai *et al.*, 2019).

PAM devices capture sound across a variety of frequencies: Audible-range units (<20 kHz) can capture crickets, bees, and some aquatic insects, while ultrasonic devices (>20 kHz) capture signals from moths and katydids. Higher sampling rates allow recording of higher frequency sounds, but increase storage demands and drain batteries, which limits deployment duration. Devices range from lightweight, low-cost units like Audiomoth (~80g), to professional-grade recorders like the SM4 (~1.2 kg with batteries). Microphone quality can limit detection of insects, especially in the ultrasonic range, where low-end microphones can miss subtle insect sounds (Riede & Balakrishnan, 2025). Whilst there are several automated hardware options for terrestrial recording, the hydromoth is the principal low-cost option for freshwater environments, which suffers from poor recording quality and low ***signal-to-noise ratio*** (Lamont *et al.*, 2022). Contact microphones and laser vibrometers can record vibrations in a substrate, revealing the presence of insects living on plants (Šturm *et al.*, 2022; Mehrotra *et al.*, 2024), in tree trunks, or underground (Robinson *et al.*, 2024).

PAM excels for loud or "soniferous" taxa, and where visual monitoring is difficult, for example in habitats with dense vegetation, aquatic systems, and forest canopies (Sugai *et al.*, 2019; Desjonquères, Gifford & Linke, 2020). It enables large-scale, long-term surveys with low disturbance and recording effort (Browning *et al.*, 2017; Linke *et al.*, 2018; Gibb *et al.*, 2019; Penar, Magiera & Klocek, 2020). While research has focused on Orthoptera and cicadas (Do Nascimento *et al.*, 2024; Madhusudhana, Klinck & Symes, 2024; Symes *et al.*, 2024; Bennett *et al.*, 2025; Okamoto & Oguma, 2025), there is growing interest in other soniferous taxa, such as aquatic insects (Desjonquères *et al.*, 2024), bees (Bota *et al.*, 2022), and flies (Mukundarajan *et al.*, 2017). Combined with ***machine learning***, PAM can uncover diel and seasonal activity patterns and support behavioural or ecological inference (Lawson, Whitworth & Banks-Leite, 2022; Scarpelli *et al.*, 2023; Symes *et al.*, 2024; Riede & Balakrishnan, 2025).

Data processing is a major bottleneck. The breadth and performance of automated processing tools are improving; AI models are able to distinguish select Orthoptera and cicada species with some success (Madhusudhana *et al.*, 2024; Bennett *et al.*, 2025; Okamoto & Oguma, 2025). However, background noise is a major challenge, and models often perform poorly in new settings (Tang *et al.*, 2022). A limiting factor is the lack of training data for most species, especially in tropical ecosystems (Riede & Balakrishnan, 2025). For manual processing, several software options exist (Audacity, Raven, Whombat, (Rhinehart *et al.*, 2024)), but insect-specific workflows remain scarce. Labelled insect sounds are often not shared (Zefa *et al.*, 2022; Branding *et al.*, 2024; but see Acosta *et al.*, 2024; Rivas *et al.*, 2025), so reference libraries are limited. However, repositories like Xeno-canto facilitate data sharing, while open-source R packages like Rthoptera (Rivas *et al.*, 2025) enable generation of metrics and intuitive visualisations. Solutions that expedite and standardise manual processing would support the development of reference libraries, as well as characterisation of distinct types of bioacoustic signals (sonotypes).

c. Radar

Radar is an active remote sensing technology that emits radio waves and interprets the returning signals reflected by objects. Initially developed for military use, radar has become a powerful tool for studying aerial ecology, including insects (Chapman, Drake & Reynolds, 2011; Shamoun-Baranes *et al.*, 2021; Rhodes *et al.*, 2022; Bauer *et al.*, 2024). Technological and algorithmic advances have enabled radar to deliver continuous data on insect flight patterns, timing, direction, altitude, and spatial distribution, often reaching several kilometres into the atmosphere (Drake & Reynolds, 2012; Chapman, Reynolds & Wilson, 2015; Haest *et al.*, 2024; Tielens & Kelly, 2024). These data support research on insect migration (Bauer *et al.*, 2024; Werber *et al.*, 2025), orientation (Shi *et al.*, 2021), phenology (Haest *et al.*, 2024), behaviour (Gao *et al.*, 2020), biomass dynamics (Hu *et al.*, 2016; Wotton *et al.*, 2019), and environmental drivers (Knop *et al.*, 2023). However, a key limitation is that individuals can rarely be identified to a high taxonomic resolution (Hüppop *et al.*, 2019).

Two radar types are primarily used in insect monitoring: small-scale vertical-looking radars and large-scale weather surveillance radars (Bauer *et al.*, 2024). Vertical-looking radars track individual insects and record flight trajectories, wingbeat frequencies, and body size (Zaugg *et al.*, 2008; Drake, Hao & Wang, 2024). These systems are increasingly portable and commercially available (Bauer *et al.*, 2024). In contrast, weather radars capture biomass (Mungee *et al.*, 2025) and movements (Stepanian *et al.*, 2016; Lukach *et al.*, 2022, 2024) at regional scales but lack the resolution for individual tracking.

Species identification remains challenging: weather radars inherently lack resolution (Drake, 2016; Hao *et al.*, 2020), while vertical-looking radars show promise but require improvements to radar technology, signal interpretation (Gauthreaux & Diehl, 2020; Addison *et al.*, 2022), and more labelled datasets (Haest *et al.*, 2021). Direct species-level identification will remain inherently difficult, but can be achieved under certain conditions—for example using supplementary trait data (e.g. Chapman *et al.*, 2012; Hao *et al.*, 2020). While insect detection with radar may be more reliable for larger bodied insects, size-biases can be mitigated by simultaneously using complementary radar types (Lochmann *et al.*, 2024).

Radar data are generally processed using open-source tools (Dokter *et al.*, 2019; Haest *et al.*, 2024; Kranstauber, Huybrechts & Desmet, 2025), usually through a combination of rule-based and statistical approaches. However, AI methods are being developed, and stored recordings will enable retrospective processing with new models (Sun *et al.*, 2024). Meanwhile, ongoing initiatives, including HiRAD, aim to consolidate radar entomology knowledge, developing open-access tools and datasets for ecological applications (Bauer *et al.*, 2024).

d. Photonics

Photonic sensing of insects involves recording how emitted photons, typically from lasers or LEDs, reflect from insect bodies *in situ* (Brydegaard & Svanberg, 2018). This method leverages principles from optical modulation spectroscopy and laser remote sensing to detect insects based on how they scatter, absorb, or reflect light. While passive methods using scattered sunlight are possible

(Jansson & Brydegaard, 2018), most systems rely on active illumination to ensure reliability across lighting and weather conditions (Brydegaard *et al.*, 2020; Rydhmer *et al.*, 2024).

Photonic sensors can be deployed at various scales (Saha *et al.*, 2023; Li *et al.*, 2023b), ranging from compact e-traps based on simple photo-interruption (Potamitis, Rigakis & Tatlas, 2017; Potamitis *et al.*, 2018; Preti *et al.*, 2021; Rigakis *et al.*, 2021), to short-range monitoring devices (Rydhmer *et al.*, 2024), to lidar systems that survey kilometre-long transects (Brydegaard *et al.*, 2016). The latter enables non-invasive monitoring of insect fluxes (individuals and biomass), flight headings (Li *et al.*, 2020), and dispersal (Månefjord *et al.*, 2024). While generally recording smaller volumes of space than radar, lidar captures fine-scale optical signatures such as wing membrane nanostructures (Müller *et al.*, 2023; Li *et al.*, 2023a) and melanin absorption (Goh *et al.*, 2021). While many such features have primarily been studied under controlled conditions, they could be quantified under field conditions in the future.

While they can be made weather-resistant (Chen *et al.*, 2024), lidar systems are not yet standardised, so comparison across studies is difficult. Comparability can be improved by presenting general metrics like chitin or melanin pathlength (Li *et al.*, 2023a; Månefjord *et al.*, 2024). Lidar systems also require specialised maintenance, and, like radar, can suffer from atmospheric interference. Photonic sensors can generate hundreds of thousands of observations daily but, as for other sensor types, detectability varies with the size and shape of insects (Yamoa *et al.*, 2025).

A lack of training data limits the utility of AI for data processing, though unsupervised methods like hierarchical clustering can help to estimate diversity and reveal ecological patterns (Rydhmer *et al.*, 2024; Yamoa *et al.*, 2025). Data processing generally focuses on extracting oscillation frequencies (Brydegaard *et al.*, 2020; Yamoa *et al.*, 2025), spectral profiles (Li *et al.*, 2022; Månefjord *et al.*, 2024), polarimetric properties (Genoud *et al.*, 2019) and physical parameters of insects (Müller *et al.*, 2023). Species identity can sometimes be derived based on these properties, but is limited by the absence of comprehensive optical property databases. An important step forward is to express sensor-derived traits in SI units, so that they can be matched with future measurements derived manually in the lab.

e. Multimodal sensing

Sensing systems that combine data collection modalities can form more than the sum of their parts. Several studies demonstrate technological synergies for *in situ* biodiversity monitoring, for example between cameras and environmental DNA (Stothut *et al.*, 2024; Tetzlaff *et al.*, 2024). However, studies that combine multiple sensor types are surprisingly rare—especially for insects (Buxton *et al.*, 2018; Wägele *et al.*, 2022; Kline *et al.*, 2025). Multimodal sensing can offer synchronous and asynchronous synergies for biodiversity monitoring. Synchronous synergies involve multiple sensor types recording in the same time and place. For example, recording the same individual insects with multiple sensor types could improve identification using multimodal AI (in line with previous work complementing images and DNA; Badirli *et al.*, 2021), and even create coveted labelled datasets for acoustics, radar and lidar sensing (Giuntini *et al.*, 2024). On the other hand, asynchronous synergies involve complementary insights from sensors that are not perfectly aligned in space and time. For example, cameras and acoustics might capture complementary sets of taxa (Buxton *et al.*, 2018), while radar and lidar data might be complementary in terms of spatial scale. Similarly, aerial environmental DNA (Roger *et al.*, 2022) could provide taxonomic context for entomological radar data, which have high taxonomic uncertainty (Haest *et al.*, 2024).

Beyond light, sound, and radio, other sensing modalities can also be used in specific contexts. For example, changes in capacitance can be used to detect insects walking over a surface (Scherer, Vitzthumecker & Bierl, 2022), crawling through tubes (Campbell, Dahn & Ryan, 2005), or feeding (Itskov *et al.*, 2014). Multimodality generally increases both cost and complexity. However, using one type of sensor to trigger another, for example using an infrared beam to trigger a camera, can offer savings in terms of power or storage (Hobbs & Brehme, 2017).

2. Adopt taxonomic pragmatism

Species identification has always been a major challenge in entomology. There are at least 5.5 million insect species globally, of which 80% remain undescribed (Stork, 2018). Furthermore, species compositions of insect communities are highly uneven, meaning that most species are rare, and thus difficult to record (Preston, 1962). Monitoring of insects is also severely biased towards certain geographic regions and taxa (Rocha-Ortega, Rodriguez & Córdoba-Aguilar, 2021). Small-bodied taxa are rarely monitored, and even meticulous morphological identifications can result in aggregation or “lumping” of cryptic species complexes (Li & Wiens, 2023).

The challenge of identifying insect species is amplified when using only sensor data, which can lack fine-grained morphological features. It can be very difficult for human specialists to distinguish species in visual imagery (Ștefan *et al.*, 2025b), let alone lidar or radar signals (Rydhmer *et al.*, 2022; Drake *et al.*, 2024). Identification is further complicated with the use of AI. AI models inevitably make errors, especially for rare species, owing to bias, missing taxa and class imbalances in training datasets (Buda, Maki & Mazurowski, 2018; Gharaee *et al.*, 2023). Overall, taxonomy is probably the single greatest challenge faced by insect sensing systems.

A key part of the solution is taxonomic pragmatism: specifically, we need to identify species wherever possible, but find ways to work with taxonomically uncertain data. Species-level data are a priority: many applications, from species conservation to surveillance of pests and invasives, fully depend on species identification. Fortunately, some species can be successfully identified through AI processing of images and sounds (Bjerge *et al.*, 2023a; Symes *et al.*, 2024; Jain *et al.*, 2025). Furthermore, in photonics, multi- and hyper-spectral analysis reveal features of insects that are invisible to the human eye, offering new ways to distinguish species and even sexes (Müller *et al.*, 2023; Li *et al.*, 2023a, 2025). However, to maximise the species-level performance of AI models, a priority is to develop and use training datasets verified by morphological or DNA-based identification (Kirkeby *et al.*, 2021; Step 3; Gharaee *et al.*, 2023).

Where species identification is not possible, there are still opportunities to leverage the scale and resolution of sensor data. Higher taxa or functional groups are usually distinguishable with sensors. Cameras can record visits by different pollinator guilds through seasonal or day-night cycles (Alison *et al.*, 2022; Anderson, Rotheray & Mathews, 2023). Acoustic indices or sonotypes help to reveal broad patterns in the activity and diversity of insects (Gomez-Morales & Acevedo-Charry, 2022; Scarpelli *et al.*, 2023). Similarly, radar data capture mass migrations of insects in a range of size classes (Jeffries *et al.*, 2013; Hu *et al.*, 2016; Wotton *et al.*, 2019). Careful aggregation of species can yield indicators that are both inclusive and interpretable (Outhwaite *et al.*, 2019, 2020). Indeed, several of the strongest indicators of insect decline have involved taxonomically coarse data (Brooks *et al.*, 2012; Hallmann *et al.*, 2017), which insect sensing systems can generate in unprecedented quantities (Werber *et al.*, 2025).

Taxonomic pragmatism should not take the form of analytical shortcuts. Working with taxonomic uncertainty means training models that effectively identify insects at both high (Mazen, 2023) and low taxonomic ranks (Bjerge *et al.*, 2022), producing outputs at the highest reasonable taxonomic resolution (Step 5). Species-level records help to build species abundance indicators (Kissling *et al.*, 2018), while higher-level records shed light on wider community biomass, composition and function—for example, by cross-referencing with insect trait databases (Hörren *et al.*, 2022).

Trait-based approaches can capture ecosystem processes using data spanning multiple taxonomic ranks (Moretti *et al.*, 2017). However, such analyses rely heavily on transparent use of taxonomic backbones (Step 3; August *et al.*, 2015).

Above all, taxonomic pragmatism depends on quantification of uncertainty during identification. It is rare for AI models to communicate uncertainty effectively (Gawlikowski *et al.*, 2023), and they sometimes express higher confidence when they are out of their comfort zone (Box 2). More reliable uncertainty measures could reduce the impact of AI identification errors during downstream analyses, but also flag uncertain records for expert review within a ***human-in-the-loop*** workflow (Step 4).

Box 2. AI doesn't know what it doesn't know

AI classifiers often output a "confidence score" between 0 and 1, but this should not be assumed to represent the probability of a correct identification. Confidence scores are a good indicator of AI uncertainty (Chen *et al.*, 2025), but they must be interpreted with caution.

Confidence scores reflect how strongly an AI classifier favours a known class. However, the AI's internal representation of the real world is incomplete: the model hasn't been trained to recognise every possible species in every possible position. As a result, incorrect identifications can be spuriously assigned high confidence values, particularly for unfamiliar species, or familiar species in unfamiliar contexts (Nguyen, Yosinski & Clune, 2015; Hein, Andriushchenko & Bitterwolf, 2019).

How, then, should we quantify AI uncertainty when processing data from novel deployments? One option is to use a small, but representative, labelled dataset from novel deployment to capture model overconfidence, bias, and even calibrate confidence values (Wang, Feng & Zhang, 2021; Gawlikowski *et al.*, 2023; Wood & Kahl, 2024). However, AI models of the future should be trained to flag inputs that fall outside their experience (Gawlikowski *et al.*, 2023), using methods such as Monte Carlo dropout (Gal & Ghahramani, 2015).

3. Create, curate and share labelled data

Labels are essential, representing biological information to sensor data. They can be generated by humans or AI models, and take many forms, for example species IDs attached to bounding boxes, or behavioral descriptors attached to time segments. Human-generated labels, often called ***annotations***, are generally treated as ***ground truth***; they are the gold-standard for training and validating AI models, though sometimes associated with error of their own (Austen *et al.*, 2016; Foody, 2024). However, manual annotation is costly, demanding many human working hours, and forms a severe bottleneck for data processing. Despite advances in AI for machine-generation of labels, the demand for annotations is greater than ever (Box 3). The question is not *whether* to annotate, but *what* to annotate and *how*: the answer depends on what annotations will be used for, as well as which labelled datasets are already available (Blair *et al.*, 2024).

Box 3. Paradoxically, the rise of AI means more work on annotation—not less.

AI is expected to revolutionise entomology by reducing the need for manual annotation of sensor data (Høye *et al.*, 2021). This is true in relative terms, as AI models can automate an increasing share of data processing. However, according to Jevon's paradox, greater efficiency in using a resource can lead to an increase in its consumption (Polimeni & Polimeni, 2006). Similarly, the rise of AI streamlines data processing, but drives a disproportionate increase in demand for labelled data to create more and better AI models (Hellendoorn & Sawant, 2022; Woodruff *et al.*, 2023). This phenomenon is exacerbated by the fact that labelled data drive the development of the next generation of AI models. Demand for annotated data is high because:

1. **Awareness of the potential of sensors and AI is widespread.** Stakeholders are eager to collect and use sensor data, and deploy AI models.
2. **Annotations are required to train new and better AI models.** For example, more training data can improve performance of classifiers for rare species.
3. **Annotations are essential for measuring AI model performance.** We should never blindly trust AI; annotations are key for benchmarking and evaluation.
4. **Where AI fails, annotation is a reliable path to ecological data.** AI is not viable for all sensors and contexts, and annotated data can be analysed biologically.

It is rarely feasible to manually annotate all data generated by a sensor. Instead, it is often necessary to select strategic subsets of the data to achieve annotation goals. Annotations can be used (i) to train AI models, (ii) to evaluate AI outputs, or (iii) as a direct source of biological data. Where the aim is to improve AI model performance, a good strategy is to increase representation of species or conditions where models perform poorly (Koch *et al.*, 2022). If annotations are used to evaluate an AI model in a new context (Cowans *et al.*, 2022; Wood & Kahl, 2024), a good strategy is to subset data evenly across space, time or taxa to fully represent the new context. When annotations are for biological analysis, a good strategy is to target regions and periods of ecological interest, such as open flowers for pollinators (Alison *et al.*, 2022). For non-visual methods it is particularly difficult to generate reliable ***ground truth*** annotations; the process may involve proactively recording known species in the field (Gradišek *et al.*, 2017) or in cages (Rydhmer *et al.*, 2024), or combining multiple recording methods (Van Doren *et al.*, 2023; Giuntini *et al.*, 2024). It is always important to annotate some sensor data where insects are absent, providing true negatives that form a counterfactual for training and evaluation of AI.

Well-documented annotation protocols are vital to ensure quality and reproducibility. Processing tasks include classification, detection, ***segmentation***, measurement or description of insects (Fig. 3), with outputs ranging from a taxonomic ID linked to a recording (Madhusudhana *et al.*, 2024) to a pixel mask highlighting insects in an image ((Svenning *et al.*, 2025); see Fig. 3: Segmentation). Protocols should capture methodological choices that affect reproducibility; especially details related to annotation scope. For example, when is a signal considered too small or weak to annotate? How are the boundaries of an annotation defined? A gold-standard approach is to consult multiple annotators, allowing uncertainty in the human-defined ***ground truth*** to be quantified. This gives context when evaluating an AI model: if two human annotators only agree in 90% of cases, models should not be expected to exceed 90% accuracy. There are a wide variety of viable software tools for annotation, many of which are free and open source (LabelStudio, CVAT, VIA, RoboFlow, Whombat, WildTrax), while platforms like Zooniverse allow crowdsourced ***annotations*** from multiple users (Willi *et al.*, 2019).

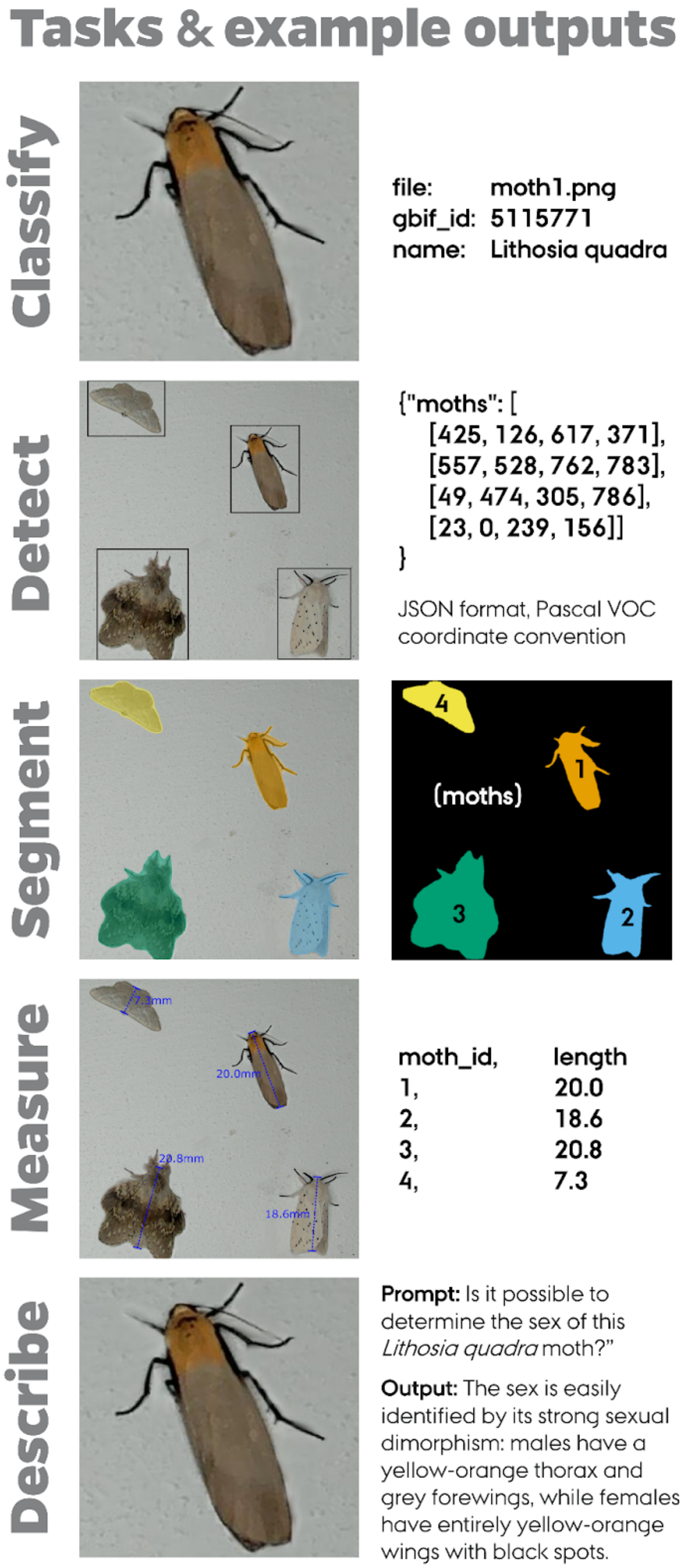

Figure 3. Data processing tasks such as classification, detection, segmentation, measurement and description can be performed by humans, or using automated processing tools such as AI. These tasks output labels in various formats: classification may produce a table with the ID of the insect in each input image; detection may produce an output for each input image, with coordinates of the bounding boxes of each detected object. Outputs can also be more complex: in segmentation, image masks or polygons are generated for each object. For measurement, the objects are assigned a value according to the trait desired. Tasks exemplified here for visual imagery exist in similar forms for acoustic, radar, and lidar sensing. Detections, segmentations and measurements in the Figure were derived using the flatbug model (Svenning *et al.*, 2025), while the description was generated using the Gemini 2.5 Pro model (Comanici *et al.*, 2025).

Large volumes of data and labels are already being generated by insect sensing systems, both in research and industry, yet they are not always shared according to FAIR principles: Findability, Accessibility, Interoperability and Reusability (Wilkinson *et al.*, 2016). While some datasets are published specifically for AI applications (Gharaee *et al.*, 2023), other AI-relevant datasets become available as side effects of research, education and community engagement (August *et al.*, 2015; Unger *et al.*, 2021; Wilson *et al.*, 2023). Datasets accompanying scientific publications are often persistent and findable in general-purpose data repositories, such as Zenodo. Data sharing platforms such as Agouti or Wildlife Insights (Ahumada *et al.*, 2020) have extended reach, catering to a broad ecological audience, but place no emphasis on insects. A catalogue of annotated insect datasets would help to leverage synergies between them, especially to train generalised entomological AI models that perform in a wide variety of contexts (Schneider *et al.*, 2023). For example, Svenning et al. (2025) curated 23 arthropod image datasets, enabling a highly generalised model for arthropod detection and ***segmentation***. Similarly, the TreeOfLife-10M dataset collates the iNat21, BioScan-1M (Gharaee *et al.*, 2023) and Encyclopedia of Life datasets to advance general organism classification, especially at higher taxonomic ranks (Stevens *et al.*, 2023). Still, the accessibility, interoperability and reusability of data ultimately depends on well-defined, shared metadata and ***standards*** (Box 4). Development of, and adherence to, community ***standards*** is absolutely vital for scalability, yet remains an under-resourced frontier.

Box 4. Integrating standards for AI and biodiversity is crucial

While metadata provides context or details about data, standards provide common terminology and structure with which to store and share data. At the heart of a data standard is a set of well-defined terms, such as *scientificName* and *captureMethod*, that link to established ontologies, such as community agreed taxonomic, life-history or anatomical concepts (Yoder *et al.*, 2010). A data standard can also define structures and formats for data: the CamTrapDP standard provides tables for Deployments, Media and Observations, with a JavaScript Object Notation (JSON) format to describe them (Bubnicki *et al.*, 2024). Data standards are diverse and usually tailored to specific applications.

Standards for AI

Most AI tools used in ecology were not originally designed for ecological data. Standards for AI training datasets capture basic details on the location and class of entities in 2-dimensional images—for example the Pascal VOC format (Everingham *et al.*, 2014). Standards for insect sensing systems are more complex, but can incorporate AI data standards—for example to specify the formatting of bounding box coordinates (Fig. 3: Detection). Furthermore, to preserve the provenance of AI predictions, we need to reference models and metadata using unique resource identifiers (leveraging platforms such as kaggle and huggingface).

Standards for biodiversity

Biodiversity data brings unique challenges, especially representing taxonomic concepts and authorities (Sandall *et al.*, 2023). Taxonomic backbones—centralised checklists of biological names, capturing each taxon's lineage, authority and synonyms—are crucial. The GBIF backbone taxonomy collates a huge number of other taxon lists and is widely used in research across many branches of the tree of life (August *et al.*, 2015; GBIF Secretariat, 2023; Roy *et al.*, 2024). The Biodiversity Information Standards community develops, ratifies and promotes standards specifically for biodiversity data. These include the Darwin Core standard, built to enable interoperability of global biodiversity data (Wieczorek *et al.*, 2012).

Standards for insect sensing systems
The Darwin Core standard is designed for primary occurrences of species in nature, especially opportunistic wildlife recording data. However, insect sensing systems generate highly structured data, and require metadata about hardware, media, deployments, AI and identification uncertainty which cannot be expressed in Darwin Core. Some progress has been made with Audiovisual Core (previously the Audubon Core; Morris *et al.*, 2013) which is tailored specifically to biodiversity media, including terms that cover image metadata and regions of images containing organisms. Furthermore, the recently ratified Humboldt extension to Darwin Core provides terms to describe the spatial, temporal and taxonomic scope of sampling events (Guralnick, Walls & Jetz, 2018). Most recently, CamTrap DP is a standard based on Darwin Core which is being adapted for insect camera traps and passive acoustic monitoring (Bubnicki *et al.*, 2024; Reyserhove, Norton & Desmet, 2025).

4. Delegate to AI with care

Given the hype around AI (Pollock *et al.*, 2025), it is easily seen as the solution to every data processing challenge (Placani, 2024). However, AI is not a panacea for insect sensing systems: some sensors are less "AI-enabled" than others (Fig. 2), and integrating AI into monitoring pipelines is far from trivial (Schneider *et al.*, 2023). AI models are costly to develop, energy intensive (Luccioni, Jernite & Strubell, 2023), and not well suited to all situations or tasks. Furthermore, there are many ethical issues around the development and use of AI (Taffel, Bedford & Mann, 2022). Before rushing to assign every task to AI, proper consideration must be given to non-AI or hybrid solutions. We recommend always asking two key questions: (1) Is AI best suited for the task? (2) What is the role of AI in the workflow?

Is AI best suited for the task?

AI models are candidates for many data processing tasks (e.g. Fig. 3), but their availability and performance varies across sensors, tasks and ***domains***. Furthermore, they must be measured against their competition—for example ***algorithmic processing*** (Box 5) and human processing. The suitability of AI for data processing depends on the context; for some tasks, such as insect detection using radar, annotated data are not widely available to train AI models (Gauthreaux & Diehl, 2020; Haest *et al.*, 2021). As such, ***algorithmic processing***, with calibration by experts, remains the standard approach to extract entomological data from radar and lidar (Dokter *et al.*, 2019; Chen *et al.*, 2024). On the other extreme, annotated images of insects are so abundant that ***foundation models***—broad generalist AI—are now extremely useful for image-based monitoring (Box 6).

Box 5. Algorithmic processing can be a viable alternative to AI

Algorithmic processing (also called signal processing; Lürig *et al.*, 2021) is rule-based processing to transform raw sensor data into biologically meaningful information. It differs from AI in that it does not rely on large training datasets, drawing instead on small calibration datasets, expert knowledge and sensor-specific logic. Being simpler than AI, algorithmic processing is usually highly explainable. It is commonly used to process radar or lidar data, including methods for thresholding, filtering and feature extraction (e.g. edges or echoes). It can also be used to preprocess data before using AI.

Algorithmic processing is particularly effective where:

1. **The signal-to-noise ratio is high.** When sensor data are not highly variable or noisy, a simple rule might be sufficient to extract useful information.
2. **Computing resources are limited.** Algorithmic processing requires modest computational resources, and can usually be deployed on low-powered devices.
3. **Training data are scarce.** In the absence of large labelled datasets, AI approaches are generally less viable.

Task complexity also determines the suitability of AI for data processing (Fig. S1). For simple, repetitive tasks with a high ***signal-to-noise ratio***, ***algorithmic processing*** can automate analysis with minimal training data (Lürig *et al.*, 2021). As the complexity and scope of the task increases, and more sensor data are created, annotated and shared, AI gradually becomes a viable solution. For example, it is much harder to detect insects in images of flowers and vegetation (Bjerge *et al.*, 2023a; Ștefan *et al.*, 2025a) than in images of an illuminated screen (Roy *et al.*, 2024); the former task is currently not always achievable with AI, while the latter is simple enough that an algorithmic thresholding method may suffice (Bjerge *et al.*, 2021). When tasks lack sufficient training data for AI and are too complex for ***algorithmic processing***, human processing becomes necessary to expand the training data and generate ecological insights (Fig. S1; Step 3).

Box 6. Foundation AI will shape the insect sensing systems of the future

Foundation models are AI that are adapted to perform a variety of tasks. They are trained on broad, ***multimodal*** datasets, usually comprising vast amounts of text and visual imagery (Bommasani *et al.*, 2021). Self-supervision is often used to improve ***generalisation***, whereby parts of the training data are obscured and the model learns to reconstruct them (He *et al.*, 2021). Foundation models include generalist AI such as GPT-5, but also AI that are broad-specialists in specific ***domains*** such as biomedicine (Zhang *et al.*, 2024) or biodiversity (Stevens *et al.*, 2023). Notably, foundation models use extremely high numbers of parameters, and consume orders of magnitude more processing power than task-specific AI (Luccioni *et al.*, 2023).

It is increasingly feasible to delegate basic insect monitoring tasks to generalist AI. Foundation vision-language models such as GroundingDINO (Liu *et al.*, 2023), CountGD (Amini-Naieni, Han & Zisserman, 2024), and SAM 2 (Ravi *et al.*, 2024) are capable of detecting, counting and segmenting insects in many types of images (Devlin *et al.*, 2025). Similarly, with cautious prompt-engineering, GPT models can describe and classify images (Miao *et al.*, 2025) or help to analyse or interpret data (Potamitis, 2023; Kendall-Bar *et al.*, 2025). However, such generalist models very rarely identify insect species, and demonstrate strong biases about insects (Moser, Krogmann & Wanke, 2025).

Meanwhile, broad-specialist foundation models are beginning to excel at difficult tasks, such as fine-grained classification. BioCLIP 2 is a vision-language model generalizing across the tree of life (Stevens *et al.*, 2023; Gu *et al.*, 2025) which has proven effective for moth species classification, albeit with a small amount of additional training (Gardiner *et al.*, 2025a). While a new foundation model for arthropods may be on the horizon (Nguyen *et al.*, 2024; Truong *et al.*, 2025), a priority is to curate ***benchmark datasets*** to properly test out-of-the-box performance of foundation models (Wu *et al.*, 2019; Schneider *et al.*, 2023).

What is the role of AI in the workflow?

AI models can be characterised based on their task—e.g. detection, classification and measurement of insects (Fig. 3)—but also their role; that is, how they complement other elements, including humans, to achieve monitoring goals (Fig. 4). This affects how, when, and where AI models should be deployed. For example, if AI predictions are used to directly generate a population index for an annual report, models can run long after data collection on a high-performance computer, while training should prioritise accuracy over speed (Fig. 4). On the other hand, AI can assume the role of a "trigger mechanism" to enable reactive monitoring of ecological phenomena, such as pollinator behaviour (Ratnayake *et al.*, 2022). This role demands an AI that runs in real time, perhaps even on-device for ***embedded processing*** *(Sittinger et al., 2024; also known as edge processing; Darras et al., 2024)*. ***Embedded processing*** is trending across animal ecology (Kline *et al.*, 2024) and entomology (Bjerge *et al.*, 2025) but it has many limitations. We advise caution before adopting on-device processing for many applications, and raw data should usually be retained for future verification (Box 7).

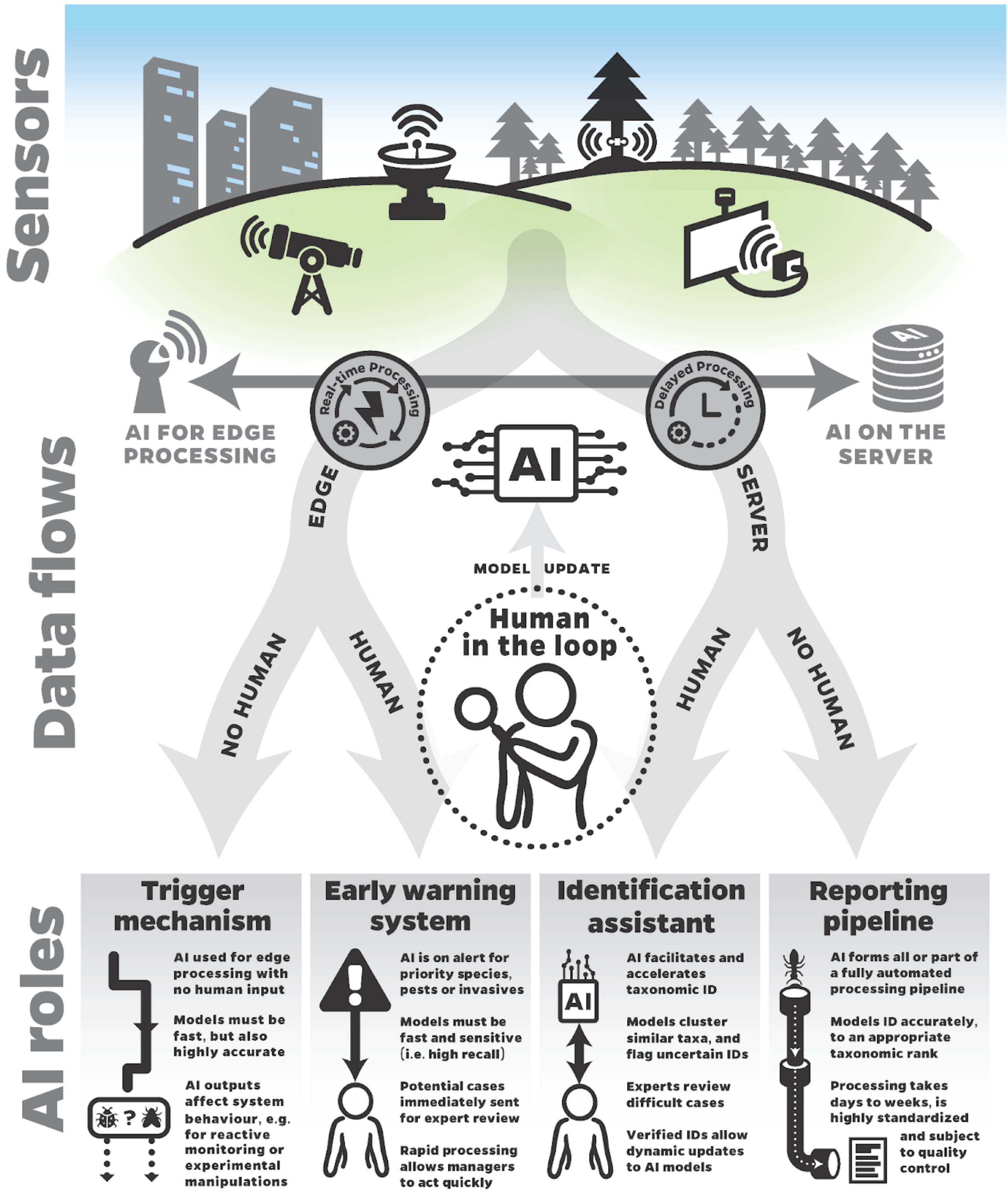

Figure 4. AI can take on a variety of roles in an insect monitoring workflow. In some roles, lightweight models must process data in real-time—perhaps even adjacent to the sensor for ***embedded*** (or edge) ***processing*** (left side). In others, larger models are run some time after data collection on high-performance computing servers (right side). Some roles demand independent working (outer routes); others demand frequent interaction with humans (inner routes). Understanding the role of AI in insect sensing systems is crucial to optimise when, where and how models are deployed

Box 7: An edge case? Embedded AI is powerful, though not always practical

Embedded processing (or edge processing) means processing sensor data locally, for example with AI models running on compact computers attached to the sensor. By processing data in real time, devices can keep track of environmental developments—such as the presence of insects—and react autonomously. Bjerge et al. (2025) demonstrate the potential of embedded AI for real-time detection, classification and tracking of night-flying insects, even identifying moths to species-level. Similar models could be used to generate alerts of high-priority events, such as pest attacks, to enable rapid management interventions (Preti *et al.*, 2021). They could also enable data dashboards to give immediate feedback and help to promote efforts to monitor insects.

A popular use case of embedded AI is for triggered or reactive sampling (Fig. 4). This can maximise the information content of captured media, for example by only recording when insects are detected (Gardiner *et al.*, 2025b). This can be effective where AI models are very accurate (e.g. Sittinger *et al.*, 2024) and data storage or bandwidth is limited. However, AI models cannot reliably detect insects in many contexts (Ștefan *et al.*, 2025a), especially under strict computational constraints. When adopting triggered sampling, we propose maintaining a small stream of representative, unfiltered sensor data (Fig. 5). These data enable us to quantify the false negatives of AI models, and may even include a useful ecological by-catch (Alison *et al.*, 2024).

Whatever the use case, ***embedded processing*** comes at a cost. Compact computers can be affordable, but power consumption can more than double when running embedded AI (Bjerge *et al.*, 2025), necessitating a reliable energy supply. Projects should consider the trade-offs involved; investment in ***embedded processing*** could reduce data storage costs, but also the number of devices that can be deployed within a given budget. Similarly, investment in embedded AI could be better directed toward powerful centralised AI, or improved data management, compression and cold-storage.

AI is often seen as displacing humans in data processing workflows, but humans and AI should perform distinct roles as a part of a complementary team. Humans in particular should take responsibility for critical decisions, and their taxonomic expertise and experience is indispensable (Dyer *et al.*, 2024). Human expertise and AI efficiency may be best exploited using a ***human-in-the-loop*** approach; Frequent interactions between humans and AI models can improve model performance (using active learning feedback loops), but also streamline human processing (Mosqueira-Rey *et al.*, 2023; Kath *et al.*, 2024). Casting AI in a supporting role is especially useful, as even error-prone models can simplify and accelerate human processing. Naturally, different data processing roles call for different qualifications, with implications for AI model selection and training.

One supporting role for AI could be that of an early warning system (Fig. 4), for example to detect outbreaks of pests such as the red palm weevil (*Rhynchophorus ferrugineus*) in palm plantations. AI can be used to detect pests in real-time, with human experts on-call to manually validate detections (Passias *et al.*, 2024). This role may call for specific training targets: a high ***recall*** model would make sure pest outbreaks are not missed, at the cost of a few more false positives. These false positives could be efficiently filtered out by human experts, and used to improve model performance in an active learning feedback loop (Mosqueira-Rey *et al.*, 2023).

A less time-critical role would be that of an identification co-pilot (Fig. 4). Reliable species-level identification with AI is not feasible across all taxa, sensors and ***domains*** (Step 2). However, a general order-level image classifier is within reach (Schneider *et al.*, 2023). As such, AI can take on the role of detecting and sorting insects in sensed media, facilitating review by a suitable taxonomic expert. The process might mirror a taxonomically tiered sorting system, as used in some insect monitoring programs (Karlsson *et al.*, 2020), with sorting at higher tiers being carried out by AI. A desirable model would readily process the most distinctive species, but flag unfamiliar ones for further review (Hogeweg *et al.*, 2024). Of course, experts are not a prerequisite for ***human-in-the-loop*** approaches; AI models can also aid citizen scientists with simpler tasks, such as flagging images or audio files that contain animal signals (Willi *et al.*, 2019; Sheard *et al.*, 2024).

5. Hold AI to account

If applied with care, AI is an increasingly useful processing tool. However, AI models can fail in surprising ways, in spite of—and sometimes owing to—their efficiency and complexity. Some AI pitfalls have roots in statistical autocorrelation; much like statistical models, AI readily exploits non-generalisable correlations in space and time, so we must be vigilant when splitting training and test datasets (Box 8). Despite advances in the field of explainable AI (Mosqueira-Rey *et al.*, 2023; Chiaburu, Haußer & Bießmann, 2024), it is not always clear how AI models function and why they make mistakes. However, flawed AI is still useful if errors can be quantified and understood (Funosas *et al.*, 2026). Robust evaluation improves interpretation of uncertain AI outputs, and helps us train more reliable models for the future.

Box 8. Too clever by half: Overfitting, shortcut learning and data leakage in AI

AI systems can achieve impressive accuracy. However, they can fail following slight changes to input data—even just a few pixels in an image (Su, Vargas & Sakurai, 2019). Such failures arise when models learn patterns in the training dataset too specifically (overfitting), or learn the wrong patterns (shortcut learning & data leakage; Geirhos *et al.*, 2020).

Overfitting
As a rule, an AI model is trained by performing a task hundreds of times using a training dataset. Ideally, it learns general patterns that transfer to new, real-world data. However, the model often learns noise or quirks in that particular dataset, meaning it performs well for training examples but poorly for new ones. This process is called overfitting, and can lead us to overestimate how models will perform in real-world situations. Overfitting happens quickly when models are large, and when training datasets are small or biased. To measure and prevent overfitting, a portion of data is withheld from training and used for validation and testing of the model (Blair *et al.*, 2024).

Shortcut learning and data leakage
Most AI models are trained to find efficient strategies to perform tasks—not necessarily rigorous ones. Shortcut learning occurs when a model uses superficial clues, rather than diagnostic features, to perform a task (e.g. classifying a species based on an indicative background, rather than morphology). Where such clues are accidentally included in the training data, this is called data leakage (Geirhos *et al.*, 2020). A famous example in medicine showed that AI can use hospital-specific symbols in images to “cheat” in diagnosis (Zech *et al.*, 2018). Similar issues could occur in ecology, for example if images of the same insect are included across training, validation and test datasets. While shortcut learning places emphasis on the behavior of the model, data leakage emphasises problems in the training, validation and test data.

As with overfitting, shortcut learning and data leakage produce models that appear to perform well, but fail to generalise to new contexts (Kapoor & Narayanan, 2023). They tend to be most problematic where sensor data have persistent features, for example background conditions, that distract AI models during training and application (Beery, Van Horn & Perona, 2018). Shortcut learning and data leakage can be diagnosed by evaluating models using out-of-distribution data, such as images with novel backgrounds (Bernett *et al.*, 2024).

To properly confront AI uncertainty, we need to be more open about it. A key issue is that models are trained and evaluated using similar data, giving an unrealistic impression of performance in novel contexts. Crucially, there is a lack of accepted ***benchmark datasets*** to fairly evaluate different models (Schneider *et al.*, 2023; Branding *et al.*, 2024 for acoustics; but see Nguyen *et al.*, 2024 for vision). Bjerge *et al.* (2023a) demonstrate issues with siloed evaluation: they trained a model to detect insects in images, achieving an F1 score of 0.932 on their own test dataset. However, when evaluating the model for insect classes not present in training and test data, only ~80% of insects were detected. Similarly, Ștefan *et al.* (2025a) found considerable drops in model performance when exposing models to background conditions and insect sizes that differed from the training data. In general, it is good practice to report evaluation results in detail. Publishing performance for each class, context, or specimen in a ***benchmark dataset*** helps to clarify how often models fail, and for which types of inputs (Burnell *et al.*, 2023).

Crucially, to understand AI uncertainty, we need to encourage evaluation using out-of-distribution (o.o.d.) data. Validation and test datasets are assumed to be independent and identically distributed (i.i.d.) to training data. Ideally, real-world data would also meet this assumption, but this is very rarely the case (Geirhos *et al.*, 2020). Real-world data are often o.o.d.: they differ from the training data based on distinct background conditions, new locations, or even shifts in the rank-abundance of taxa. Evaluation with o.o.d. data helps determine whether an AI model has learned the intended features, so that it can generalise well to new conditions within the target ***domain***. Svenning et al. (2025) evaluate their arthropod detection model across 23 distinct imaging systems, highlighting how o.o.d. evaluation can be carried out and presented during model development.

Importantly, o.o.d. evaluation is not just for model developers; it may be most useful during model deployment, using data from the context in which AI is being applied (see Appendix S1 for a hypothetical example). Annotating a subset of data processed by a model enables quality control of AI outputs (Fig. 5, golden arrows), allowing quantification of false-positives and false-negatives (Cowans *et al.*, 2022) and even calibration of confidence scores ((Wood & Kahl, 2024); Box 2). This evaluation and quality control process generates annotations to train new AI models (see Step 3), but also helps us make sense of unverified AI predictions. Error rates can be accounted for, and it is even possible to explicitly model the AI-detectability of species (Cowans *et al.*, 2022). While o.o.d. evaluation is always insightful, it can be expensive. As a minimum, AI users should consider whether training data differ from the application context, and how this may impact model performance.

Holding AI to account means assessing its reliability, but also improving its performance. A very common approach is fine-tuning—a kind of ***transfer learning*** where an existing model is re-trained on a dataset relevant to a specific task (Kumar *et al.*, 2022). However, there is room for innovation to address more fundamental weaknesses of AI for insect monitoring. For example, models often struggle to perform well against unseen backgrounds. Algorithms that leverage frame-to-frame motion could improve detection in novel contexts (Ratnayake *et al.*, 2021; Bjerge, Frigaard & Karstoft, 2023b). Alternatively, unlabelled data from novel backgrounds could be used to improve model ***generalisation***, through unsupervised domain adaptation (Kay *et al.*, 2025). Data augmentation is also crucial, artificially increasing the variability of the training data through rotation, cropping, color-shifting, or inclusion of synthetic data (Beery *et al.*, 2020; Schneider *et al.*, 2023). In acoustics, augmentation can involve playback re-recordings, or simulation of target and non-target sounds (Madhusudhana *et al.*, 2024; Okamoto & Oguma, 2025).

Many breakthroughs in AI for insect monitoring involve ***multimodal*** approaches, such as supplementing sensor data with taxonomic descriptors (Stevens *et al.*, 2023; Gu *et al.*, 2025) or DNA sequences (Badirli *et al.*, 2021; Gharaee *et al.*, 2023, 2024). By leveraging the hierarchy of the tree of life, ***multimodal*** learning can improve inference using sensor data alone. It can even enable AI to identify taxa not present in the training data, through zero-shot learning (Schneider *et al.*, 2023). ***Multimodal*** approaches also improve uncertainty quantification; AI models are often confronted with species or phenotypes not present in the training data, and hierarchical classification enables models to provide coarse taxonomic labels where appropriate (Bjerge *et al.*, 2023c). Alternatively, models can flag specimens that fall outside of their comfort zone (called novelty or o.o.d. detection; Lee *et al.*, 2018), highlighting where annotation could rapidly improve model performance (Koch *et al.*, 2022). Beyond descriptive or genetic labels, geographic and temporal labels can also improve AI model performance. They can be used, as by human experts, to effectively narrow the list of possible taxa based on distribution and phenology (Terry, Roy & August, 2020).

Finally, an important innovation in AI for insect monitoring is embedding-based modelling. Existing AI, especially ***foundation models*** (Box 6), can be used to generate compressed representations—or embeddings—of sensor data. Then, further models can rapidly explore high-dimensional relationships across those embeddings. This enables effective and computationally-efficient novelty detection, as well as classification of new taxa with very few additional training data (Hogeweg *et al.*, 2024). Furthermore, training small linear classifiers using embeddings from pre-trained models is an efficient form of ***transfer learning*.** So-called "linear probing" trains new models for specific insect monitoring tasks, while retaining some of the generality of the original models (Kumar *et al.*, 2022). BioCLIP 2 embeddings are useful for insect classification (Gardiner *et al.*, 2025a), while BirdNet and Perch embeddings could be useful for acoustic monitoring of insects (Ghani *et al.*, 2023).

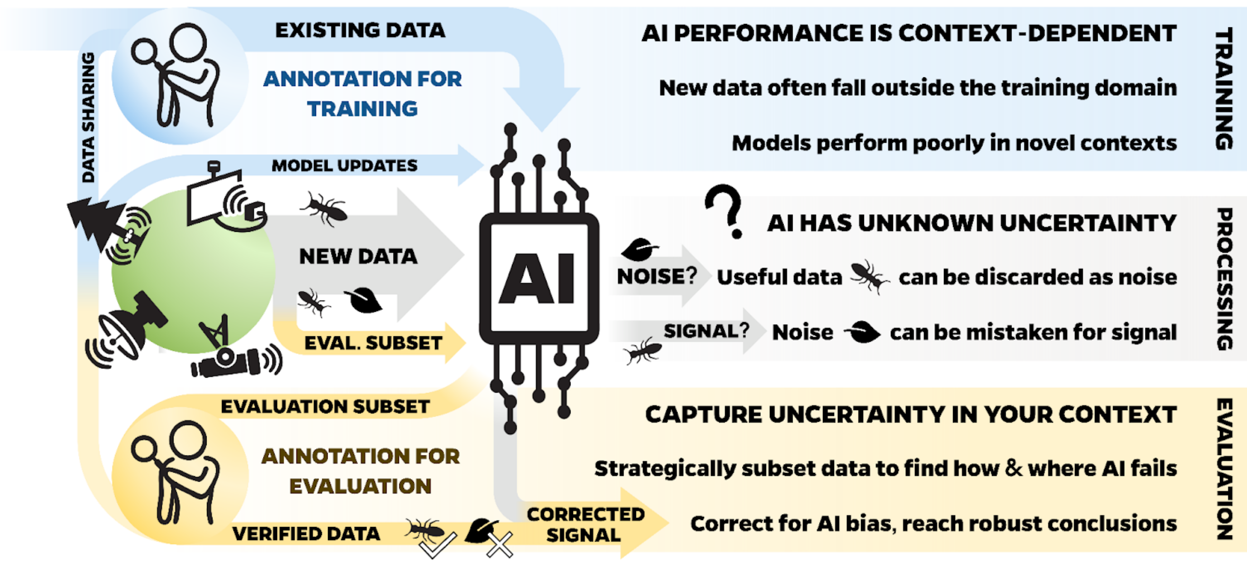

Figure 5. Annotation isn't just for model training—it also helps to capture uncertainty. Strategically annotating a subset of recorded data (golden arrows) can allow bias correction, for example by calibrating AI confidence scores (Wood & Kahl, 2024) or using false-positive models (Cowans *et al.*, 2022). Evaluation permits robust ecological conclusions despite uncertainty in AI predictions

6. Make technologies accessible

In order to be useful, automated monitoring technologies need to be accessible to researchers and practitioners on the ground. However, technology and AI development for biodiversity monitoring is highly concentrated in the global North (Chan *et al.*, 2021). This limits access to technology and training in the global South and for marginalised groups (Speaker *et al.*, 2022), and reduces the availability and relevance of AI models for tropical ecosystems (Pollock *et al.*, 2025). There is a clear risk for existing disparities to be made worse, so that the most biodiverse regions of the world continue to have reduced capacity for biodiversity monitoring (Collen *et al.*, 2008; Ocampo-Ariza *et al.*, 2023).

A variety of technical, financial and socio-economic barriers threaten global adoption of sensors and AI for insect monitoring. First, the complexity of AI, software and hardware is a major barrier to uptake. Many ecologists lack training in not only AI, but also handling of large datasets involved in AI training and data-processing (Kendall-Bar *et al.*, 2025). A lack of training compounds with a fast-evolving landscape of both software and hardware related to sensors for insects (van Klink *et al.*, 2024). Most hardware setups are custom-made rather than off-the-shelf products, so that even collection of data requires considerable expertise. Scaling insect sensing systems means more focus on training and community feedback, and a greater emphasis on ease of use and technical support (including sustainable software development; Durdik *et al.*, 2012). Specifically, overcoming deployment biases depends on equitable Global North-South partnerships that include capacity building, skills sharing, community engagement, and knowledge exchange (Haelewaters, Hofmann & Romero-Olivares, 2021; Sánchez Herrera *et al.*, 2024).

Financial costs related to sensors and AI form another major barrier. For PAM, relatively affordable sensors are available, with commercial devices costing <€150 (Hill *et al.*, 2018). However, other systems are much more expensive; high-end camera systems can cost upwards of €5,000 per device. This is prohibitively expensive in many parts of the world, where there may be added costs for insurance, security and replacement of damaged or stolen equipment; a recent deployment of 20 time-lapse cameras on a mountain in South Africa involved hiring guards for 24-hour security (Alison *et al.*, 2024). Radar and lidar sensors are often even more expensive (Fig. 2), resulting in large gaps in coverage at global scales (Heistermann, Jacobi & Pfaff, 2013). To ensure coverage in low-income countries, system development should use a process of "innovation through simplification", while automated monitoring programmes should converge on the most affordable hardware (Brydegaard *et al.*, 2024).

Even if hardware-related costs are reduced, for example through industrial production of sensors (Hill *et al.*, 2018), costs related to data processing and storage may still be prohibitive. Deploying AI is costly in terms of both computation and energy consumption, especially when using highly generalised ***foundation models*** (Luccioni *et al.*, 2023). Training or fine-tuning AI models is also expensive, and sometimes cost-prohibitive without suitable computing infrastructure (Hellendoorn & Sawant, 2022). Use of smaller, task-specific AI models—including specialised image classifiers—can help to make systems affordable and energy-efficient. However, access to computing infrastructures is a far-reaching issue that is exacerbated by needs to store and access increasing volumes of data. Growing datasets bring rising costs for insect sensing systems, with some already occupying terabytes (Gu *et al.*, 2025). Such barriers could be addressed fundamentally through more equitable international sharing of computational resources.

7. Support systematic monitoring

Systematic monitoring involves repeated, standardised surveys of insect biodiversity, ideally at strategically or randomly selected sites, and is the gold-standard for long-term biodiversity monitoring (Conrad *et al.*, 2006; Brereton *et al.*, 2011). A key example of this is the UK Butterfly Monitoring Scheme, which has monitored butterflies repeatedly at sites across the UK since 1976 (Brereton *et al.*, 2011). Systematic monitoring underpins key indicators used to track the biodiversity crisis (Butchart *et al.*, 2010). Furthermore, international policy structures, such as the Kunming-Montreal Global Biodiversity Framework decision 15/5 (2022) and the European Nature Restoration Law (2022), emphasise the need for systematic monitoring to track progress towards biodiversity targets. Patterns and trends in biodiversity can be derived using opportunistic data (van Strien, van Swaay & Termaat, 2013; Outhwaite *et al.*, 2020), but fixed-effort, systematic observations are more repeatable and information rich (Isaac & Pocock, 2015). Crucially, systematic random sampling is the most direct way to make unbiased inferences about unsampled areas (Boyd, Powney & Pescott, 2023).

How can automated methods support systematic monitoring of insects? First, they can be deployed much more widely, and according to rigorous sampling strategies. Until now, sensors for insects and other wildlife have mostly been placed opportunistically (Burton *et al.*, 2015). While some sensors have practical constraints on their placement, contributing to spatial biases (Bowler *et al.*, 2025), there is clear scope to target deployments of smaller sensors to better achieve monitoring goals. When scaling up deployments, we should strive to delineate spatial sampling units (for example, habitats or features of interest) and deploy sensors so that we represent them evenly and transparently. Automated methods allow for scalable monitoring with lower person-hour costs, increased sampling frequency, and reduced need for lethal sampling (van Klink *et al.*, 2022). Rigorous sampling strategies are the last piece of the puzzle to maximise the scientific contribution of sensors, and fully leverage their potential for standardisation (Høye *et al.*, 2025).

Second, insect sensing systems should build on existing monitoring schemes, according to established site selection protocols. We echo previous assertions that automated methods should improve and expand on, rather than replace, existing insect monitoring initiatives (Dyer *et al.*, 2024). By building on and harmonising with existing datasets, sensor deployments deliver much greater value—both immediately and in the derivation of trends in future. Furthermore, long-term monitoring faces a constant funding crisis, despite calls to expand collection of long-term data to understand insect biodiversity change (Harvey *et al.*, 2020; Didham *et al.*, 2020; Mammola *et al.*, 2020). Automated monitoring and AI are proving highly fundable areas of research, not least due to their potential for biocredit markets (Ford *et al.*, 2024). Well-aligned automated monitoring should scale-up existing monitoring schemes, but also help fund the continuation of traditional monitoring and coveted long-term datasets.

Third, work is needed to integrate data collected by automated methods, which often differ from many traditional approaches. Some of the longest-running time series of insect populations were produced using traditional methods such as light traps (moths; Conrad *et al.*, 2006; Macgregor *et al.*, 2021), suction traps (Bell, Blumgart & Shortall, 2020), and line transects (butterflies; Brereton *et al.*, 2011). Traditional datasets are clearly not replicated by automated systems (but see Chiavassa *et al.*, 2024), so calibrating new datasets with old ones is crucial. Data integration refers to statistical approaches that analyse multiple data sources within a single analytical framework, leveraging all available information to improve ecological inference (Miller *et al.*, 2019; Isaac *et al.*, 2020). It involves integrated modelling approaches that account for differing assumptions, biases, and scales of data collection (Yen *et al.*, 2019). It has been successfully used in species distribution modelling, for example to combine presence-only and abundance data (Pagel *et al.*, 2014; Fithian *et al.*, 2015). However, studies integrating time series from multiple sources are still rare (e.g. Bowler *et al.*, 2019; Hertzog *et al.*, 2021), and we need to move beyond ensemble or synthesis-based approaches toward modelling frameworks that draw on multiple sources of data.

Finally, we need to recognise that sensors are not a panacea in insect monitoring. In most cases, sensors alone do not capture features necessary for species-level identification (see Step 2). Many insect monitoring questions will continue to demand the time of taxonomic specialists, and may gain little from sensor-based approaches (Engel *et al.*, 2021). Here, sensors might support systematic monitoring by focussing on identification and monitoring of common species, freeing up taxonomists for tasks beyond the scope of current AI (Alison & Høye, 2024). Systematic monitoring schemes of the future might build on the strengths of both traditional and automated approaches using a tiered approach: many low-cost, high-frequency, taxonomically-coarse sentinel sensor stations could support a network of core monitoring stations, which combine traditional approaches with molecular methods and high-cost sensors.

Conclusions

1. Automated biodiversity monitoring with sensors promises to massively increase the availability and affordability of data, whether for biodiversity assessments, trend estimation, agricultural monitoring or ecological research. Companies promising to provide automated monitoring services are already emerging worldwide.
2. However, successful automated monitoring of insects requires a coordinated and thoughtful approach; work is needed to confront harsh realities of sensors and AI and bring them up to standard for large-scale deployment. Our attention should be focused on affordable hardware design, agreed protocols, and standardized, annotated datasets.
3. There are limits to what we can expect from AI. We highlight some often neglected caveats, not least ethical concerns. Crucially, AI identification is constrained to morphologically distinct species with large quantities of labelled sensor data. Models must not be blindly trusted, and are most useful following shrewd, case-based evaluations.
4. Human resources are an underappreciated component of insect sensing systems. While sensors and AI can contribute to species monitoring and discovery, human expertise remains a key limiting factor. In short, the role of AI is to complement human workflows—not consume them.
5. Creative use of sensors and AI can improve species detection and identification, whether by leveraging hyperspectral signals, multimodal methods or data augmentation. On the other hand, uncertain identifications can be explored at coarse taxonomic ranks, exploiting the extraordinary spatiotemporal scales and resolutions of sensor data.
6. Deployed in the right places, and with sufficient investment in data management, insect sensing systems can enhance our knowledge of insect biodiversity beyond imagination. However, this vision will only be realised through harmonisation with, and further investment in, existing long-term datasets and systematic monitoring schemes.

Acknowledgements

Paper leads: JA and LP led this work conceiving, writing and curating the core content. Senior author: RvK edited the text and suggested improvements. Section leads: JB, YC, BH, JI, JK, JL, ML, LAdN, CLO, BR and MS took responsibility at the level of leading and drafting a section of the paper. Major contributors: MA, TA, QG, RG, GM, VS contributed major chunks of text or figure content. Contributors: KB, MB, TTH, EK, WK, ML, DR, AS contributed some text, suggestions and comments. All authors approved the current version of the manuscript.

This manuscript is based upon work from COST Action ‘InsectAI - Using Image-based AI for Insect Monitoring & Conservation’, CA22129, supported by COST (European Cooperation in Science and Technology), as well as the ‘Status of Insects’ Research Coordination Network grant from the National Science Foundation (NSF DEB 2225092). LP was supported by the Swiss Data Science Center grants C19-10 and C22-05. RvK was funded by the DFG, grant number FZT-118 to iDiv and the BMFTR through grant number 03LW0654 to the LEPMON project. CLO was supported by funding from Research England and UK Natural Environment Research Council (NERC) grant (NE/V006533/1). TA and TH were supported by the MAMBO project under the European Union's research and innovation programme No.101060639. TA was additionally supported by NERC, through the UKCEH National Capability for UK Challenges Programme NE/Y006208/1. BR was supported through the HiRAD project, funded through the Swiss National Science Foundation (SNF 31BD30_216840) as part of the 2022–2023 Biodiversa+ BiodivMon call for proposals. ML was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement Bug-Flash No. 850463). WEK was supported by the NERC award NE/V006916/1 as part of the DRUID (Drivers and Repercussions of UK Insect Declines) project. The NSF provided support for RG (DBI 2512789) and MA (DBI 2512788). We thank Stacey Jones for bringing Figure 4 to life, and providing assets used in Figures 1 and 5. Generative AI was only used to help condense an early draft of Step 1.

Glossary

Term	Explanation
Annotation, *Label*	Biological information attached to sensor data (e.g.: species presence, identity or behaviour), often used for training, validating, or evaluating AI models. Human-generated labels are often called annotations.
Algorithmic Processing, *Signal Processing*	Extracting meaningful information from raw sensor data using rule-based, mathematical and statistical techniques. One example could be filtering or modification of digital images by means of algorithms and filter matrices (kernels).
Artificial Intelligence (AI)	Computational systems that perform tasks typically requiring human intelligence. Often refers to machine learning models that learn complex patterns from large training datasets to carry out data processing tasks.
Benchmark (dataset)	A curated dataset with associated annotations. Used to compare performance of data processing tools for a specific task, under a set of known constraints and assumptions.
Domain	Entities, contexts and data structures that are in scope for a given data processing task. Can include taxa,

	backgrounds, environmental conditions, recording parameters and file types. An AI model has a source domain, captured by the training dataset, but is tested and applied in target domains. Where source and target domains differ (domain shift) performance may decline.
Embedded processing, Edge processing	Data processing done locally to the sensor device, e.g. on a compact computer. Used to prioritize which data to store or transfer, and for real-time updates or changes to system behaviour.
Foundation models	Large AI models, trained on vast and often multimodal datasets (e.g. text, audio, images), well-suited to general tasks.
Ground truth, Ground reference	Direct observations considered to be true, as opposed to inferred. Can be derived using meticulous calibration or expert annotation.
Generalisation	The ability of a model or monitoring system to perform well when applied to taxa, environments or conditions that differ from those seen during training or calibration.
Insect sensing systems	Acoustic and/or electromagnetic sensors that continuously record insects *in-situ* (under field conditions), including associated hardware, software and data processing pipelines.
Human-in-the-loop	A workflow where human expertise is incorporated to validate, correct or refine outputs of automated data processing. Permits verified records, quality control and feedback loops to train AI.
Machine Learning (ML)	A class of algorithms to train models that learn complex patterns from data to make predictions.
Multimodality	The use or integration of multiple data types or sensing modalities (e.g. imagery, audio, radar, optical signals, DNA, text descriptors) within a single monitoring system or analytical framework.
Segmentation	A data processing task that involves separating or “segmenting” regions of an image based on shared properties, for example separating foreground insects from one another, and from their background. The output is often a set of bounding polygons or pixel masks.
Signal-to-noise ratio	The relative strength of biologically relevant signals compared to irrelevant environmental or sensor noise.
Standards	Sets of terms, structures and formats for storing and sharing data relevant to a specific field or application.
Transfer learning	The process of re-training an AI model for a new task while leveraging features learned for a previous task.

Bibliography

Acosta, R.C., Mendes, D.M.D.E.M., Ramalho, E.E. & Zefa, E. (2024) Oecanthus buxixu sp. nov. (Orthoptera: Grylloidea: Oecanthidae): A new species of tree cricket from Brazilian Amazon rainforest. *Zootaxa* **5492**, 289–296. Magnolia Press.

Addison, F.I., Dally, T., Duncan, E.J., Rouse, J., Evans, W.L., Hassall, C. & Iii, R.R.N. (2022) Simulation Radar Cross Section Noctuid Moth. *Remote Sensing* **14**.

Ahumada, J.A., Fegraus, E., Birch, T., Flores, N., Kays, R., O'Brien, T.G., Palmer, J., Schuttler, S., Zhao, J.Y., Jetz, W., Kinnaird, M., Kulkarni, S., Lyet, A., Thau, D., Duong, M., et al. (2020) Wildlife Insights: A platform to maximize the potential of camera trap and other passive sensor wildlife data for the planet. *Environmental conservation* **47**, 1–6. Cambridge University Press (CUP).

Alberdi, A., Aizpurua, O., Gilbert, M.T.P. & Bohmann, K. (2018) Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in ecology and evolution* **9**, 134–147. Wiley.

Alison, J., Alexander, J.M., Diaz Zeugin, N., Dupont, Y.L., Iseli, E., Mann, H.M.R. & Høye, T.T. (2022) Moths complement bumblebee pollination of red clover: a case for day-and-night insect surveillance. *Biology letters* **18**, 20220187.

Alison, J. & Høye, T.T. (2024) Automated visual systems for insect monitoring and conservation. In *Routledge Handbook of Insect Conservation* pp. 475–4861st Edition. Routledge, London.

Alison, J., Payne, S., Alexander, J.M., Bjorkman, A.D., Clark, V.R., Gwate, O., Huntsaar, M., Iseli, E., Lenoir, J., Mann, H.M.R., Steenhuisen, S.-L. & Høye, T.T. (2024) Deep learning to extract the meteorological by-catch of wildlife cameras. *Global change biology* **30**, e17078. Wiley.

Amini-Naieni, N., Han, T. & Zisserman, A. (2024) CountGD: Multi-Modal Open-World Counting. *arXiv [cs.CV]*. Http://arxiv.org/abs/2407.04619.

Anderson, M., Rotheray, E.L. & Mathews, F. (2023) Marvellous moths! pollen deposition rate of bramble (Rubus futicosus L. agg.) is greater at night than day. *PloS one* **18**, e0281810.

August, T., Harvey, M., Lightfoot, P., Kilbey, D., Papadopoulos, T. & Jepson, P. (2015) Emerging technologies for biological recording. *Biological journal of the Linnean Society. Linnean Society of London* **115**, 731–749. Oxford Academic.

Austen, G.E., Bindemann, M., Griffiths, R.A. & Roberts, D.L. (2016) Species identification by experts and non-experts: comparing images from field guides. *Scientific Reports* **6**, 1–7. Nature Publishing Group.

Badirli, S., Akata, Z., Mohler, G., Picard, C. & Dundar, M. (2021) Fine-grained zero-shot learning with DNA as side information. *arXiv [cs.CV]*. Http://arxiv.org/abs/2109.14133.

Barlow, S.E. & O'Neill, M.A. (2020) Technological advances in field studies of pollinator ecology and the future of e-ecology. *Current opinion in insect science* **38**, 15–25.

Bauer, S., Tielens, E.K. & Haest, B. (2024) Monitoring aerial insect biodiversity: a radar perspective. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230113. The Royal Society.

Beery, S., Liu, Y., Morris, D., Piavis, J., Kapoor, A., Meister, M., Joshi, N. & Perona, P. (2020) Synthetic examples improve generalization for rare classes. In *2020 IEEE Winter Conference* *on Applications of Computer Vision (WACV)*. IEEE.

Beery, S., Van Horn, G. & Perona, P. (2018) Recognition in Terra Incognita. In *Computer Vision – ECCV 2018* pp. 472–489. Springer International Publishing, Cham.

Bell, J.R., Blumgart, D. & Shortall, C.R. (2020) Are insects declining and at what rate? An analysis of standardised, systematic catches of aphid and moth abundances across Great Britain. *Insect Conservation and Diversity* **13**, 115–126. Wiley.

Bennett, D., Nissen, H., Maschke, M.A., Reck, H. & Diekötter, T. (2025) Recent technological developments allow for passive acoustic monitoring of Orthoptera (grasshoppers and crickets) in research and conservation across a broad range of temporal and spatial scales. *Basic and applied ecology* **84**, 147–157. Elsevier BV.

Bernett, J., Blumenthal, D.B., Grimm, D.G., Haselbeck, F., Joeres, R., Kalinina, O.V. & List, M. (2024) Guiding questions to avoid data leakage in biological machine learning applications. *Nature methods* **21**, 1444–1453. Springer Science and Business Media LLC.

Besson, M., Alison, J., Bjerge, K., Gorochowski, T.E., Høye, T.T., Jucker, T., Mann, H.M.R. & Clements, C.F. (2022) Towards the fully automated monitoring of ecological communities. *Ecology letters* **25**, 2753–2775.

Bjerge, K., Alison, J., Dyrmann, M., Frigaard, C.E., Mann, H.M.R. & Høye, T.T. (2023a) Accurate detection and identification of insects from camera trap images with deep learning. *PLOS Sustainability and Transformation* **2**, e0000051. Public Library of Science (PLoS).

Bjerge, K., Frigaard, C.E. & Karstoft, H. (2023b) Object Detection of Small Insects in Time-Lapse Camera Recordings. *Sensors (Basel, Switzerland)* **23**.

Bjerge, K., Geissmann, Q., Alison, J., Mann, H.M.R., Høye, T.T., Dyrmann, M. & Karstoft, H. (2023c) Hierarchical classification of insects with multitask learning and anomaly detection. *Ecological informatics* **77**, 102278. Elsevier BV.

Bjerge, K., Karstoft, H. & Høye, T.T. (2025) Towards edge processing of images from insect camera traps. *Remote sensing in ecology and conservation*. Wiley.

Bjerge, K., Mann, H.M.R. & Høye, T.T. (2022) Real-time insect tracking and monitoring with computer vision and deep learning. *Remote Sensing in Ecology*. Wiley Online Library.

Bjerge, K., Nielsen, J.B., Sepstrup, M.V., Helsing-Nielsen, F. & Høye, T.T. (2021) An automated light trap to monitor moths (Lepidoptera) using computer vision-based tracking and deep learning. *Sensors* **21**, 1–18.

Blair, J.D., Gaynor, K.M., Palmer, M.S. & Marshall, K.E. (2024) A gentle introduction to computer vision-based specimen classification in ecological datasets. *Journal of Animal Ecology* **93**, 147–158. John Wiley & Sons, Ltd.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W. & de Bruyn, M. (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution* **29**, 358–367. Elsevier BV.

Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., et al. (2021) On the opportunities and risks of foundation models. *arXiv [cs.LG]*. Http://arxiv.org/abs/2108.07258.

Bota, G., Traba, J., Sardà-Palomera, F., Giralt, D. & Pérez-Granados, C. (2022) Passive acoustic monitoring for estimating human-wildlife conflicts: The case of bee-eaters and apiculture. *Ecological indicators* **142**, 109158. Elsevier BV.

Bowler, D.E., Boyd, R.J., Callaghan, C.T., Robinson, R.A., Isaac, N.J.B. & Pocock, M.J.O. (2025) Treating gaps and biases in biodiversity data as a missing data problem. *Biological reviews of the Cambridge Philosophical Society* **100**, 50–67. Wiley.

Bowler, D.E., Nilsen, E.B., Bischof, R., O'Hara, R.B., Yu, T.T., Oo, T., Aung, M. & Linnell, J.D.C. (2019) Integrating data from different survey types for population monitoring of an endangered species: the case of the Eld's deer. *Scientific Reports* **9**, 7766. Springer Science and Business Media LLC.

Boyd, R.J., Powney, G.D. & Pescott, O.L. (2023) We need to talk about nonprobability samples. *Trends in ecology & evolution* **38**, 521–531. Elsevier BV.

Branding, J., von Hörsten, D., Böckmann, E., Wegener, J.K. & Hartung, E. (2024) InsectSound1000 An insect sound dataset for deep learning based acoustic insect recognition. *Scientific Data* **11**, 475. Springer Science and Business Media LLC.

Brereton, T.M., Cruickshanks, K.L., Risely, K., Noble, D.G. & Roy, D.B. (2011) Developing and launching a wider countryside butterfly survey across the United Kingdom. *Journal of insect conservation* **15**, 279–290. Springer Science and Business Media LLC.

Brooks, D.R., Bater, J.E., Clark, S.J., Monteith, D.T., Andrews, C., Corbett, S.J., Beaumont, D.A. & Chapman, J.W. (2012) Large carabid beetle declines in a United Kingdom monitoring network increases evidence for a widespread loss in insect biodiversity. *The journal of applied ecology* **49**, 1009–1019. Wiley.

Browning, E., Gibb, R., Glover-Kapfer, P. & Jones, K.E. (2017) Passive acoustic monitoring in ecology and conservation. WWF-UK.

Brydegaard, M., Jansson, S., Malmqvist, E., Mlacha, Y.P., Gebru, A., Okumu, F., Killeen, G.F. & Kirkeby, C. (2020) Lidar reveals activity anomaly of malaria vectors during pan-African eclipse. *Science advances* **6**, eaay5487. American Association for the Advancement of Science (AAAS).

Brydegaard, M., Merdasa, A., Gebru, A., Jayaweera, H. & Svanberg, S. (2016) Realistic instrumentation platform for active and passive optical remote sensing. *Applied Spectroscopy* **70**, 372–385. SAGE Publications.

Brydegaard, M., Pedales, R.D., Feng, V., Yamoa, A.S.-D., Kouakou, B., Månefjord, H., Wührl, L., Pylatiuk, C., Amorim, D. de S. & Meier, R. (2024) Towards global insect biomonitoring with frugal methods. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230103. The Royal Society.

Brydegaard, M. & Svanberg, S. (2018) Photonic monitoring of atmospheric and aquatic fauna. *Laser & photonics reviews* **12**, 1800135. Wiley.

Bubnicki, J.W., Norton, B., Baskauf, S.J., Bruce, T., Cagnacci, F., Casaer, J., Churski, M., Cromsigt, J.P.G.M., Farra, S.D., Fiderer, C., Forrester, T.D., Hendry, H., Heurich, M., Hofmeester, T.R., Jansen, P.A., et al. (2024) Camtrap DP: an open standard for the FAIR exchange and archiving of camera trap data. *Remote sensing in ecology and conservation* **10**, 283–295. Wiley.

Buchner, D., Sinclair, J.S., Ayasse, M., Beermann, A.J., Buse, J., Dziock, F., Enss, J., Frenzel, M., Hörren, T., Li, Y., Monaghan, M.T., Morkel, C., Müller, J., Pauls, S.U., Richter, R., et al. (2025)

Upscaling biodiversity monitoring: Metabarcoding estimates 31,846 insect species from Malaise traps across Germany. *Molecular ecology resources* **25**, e14023. Wiley.

Buda, M., Maki, A. & Mazurowski, M.A. (2018) A systematic study of the class imbalance problem in convolutional neural networks. *Neural networks: the official journal of the International Neural Network Society* **106**, 249–259. Elsevier BV.

Burnell, R., Schellaert, W., Burden, J., Ullman, T.D., Martinez-Plumed, F., Tenenbaum, J.B., Rutar, D., Cheke, L.G., Sohl-Dickstein, J., Mitchell, M., Kiela, D., Shanahan, M., Voorhees, E.M., Cohn, A.G., Leibo, J.Z., et al. (2023) Rethink reporting of evaluation results in AI. *Science (New York, N.Y.)* **380**, 136–138.

Burton, A.C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J.T., Bayne, E. & Boutin, S. (2015) REVIEW: Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. *The journal of applied ecology* **52**, 675–685. Wiley.

Butchart, S.H.M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J.P.W., Almond, R.E.A., Baillie, J.E.M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K.E., Carr, G.M., Chanson, J., Chenery, A.M., Csirke, J., et al. (2010) Global biodiversity: indicators of recent declines. *Science (New York, N.Y.)* **328**, 1164–1168. American Association for the Advancement of Science (AAAS).

Buxton, R.T., Lendrum, P.E., Crooks, K.R. & Wittemyer, G. (2018) Pairing camera traps and acoustic recorders to monitor the ecological impact of human disturbance. *Global Ecology and Conservation* **16**, e00493. Elsevier BV.

Calvus, D., Wueppenhorst, K., Schlösser, R., Klaus, F., Schwanecke, U. & Greil, H. (2025) In-field monitoring of ground-nesting insect aggregations using a scaleable multi-camera system. *Ecological informatics* **86**, 103004. Elsevier BV.

Campbell, J.M., Dahn, D.C. & Ryan, D.A.J. (2005) Capacitance-based sensor for monitoring bees passing through a tunnel. *Measurement science & technology* **16**, 2503–2510. IOP Publishing.

Chan, A., Okolo, C.T., Terner, Z. & Wang, A. (2021) The limits of global inclusion in AI development. *arXiv [cs.CY]*. Http://arxiv.org/abs/2102.01265.

Chapman, J.W., Bell, J.R., Burgin, L.E., Reynolds, D.R., Pettersson, L.B., Hill, J.K., Bonsall, M.B. & Thomas, J.A. (2012) Seasonal migration to high latitudes results in major reproductive benefits in an insect. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 14924–14929. Proceedings of the National Academy of Sciences.

Chapman, J.W., Drake, V.A. & Reynolds, D.R. (2011) Recent insights from radar studies of insect flight. *Annual review of entomology* **56**, 337–356. Annual Reviews.

Chapman, J.W., Reynolds, D.R. & Wilson, K. (2015) Long-range seasonal migration in insects: mechanisms, evolutionary drivers and ecological consequences. *Ecology letters* **18**, 287–302. Wiley.

Chen, H., Li, M., Månefjord, H., Travers, P., Salvador, J., Müller, L., Dreyer, D., Alison, J., Høye, T.T., Gao Hu, Warrant, E. & Brydegaard, M. (2024) Lidar as a potential tool for monitoring migratory insects. *iScience* **27**, 109588. Elsevier BV.

Chen, Y., Lang, N., Schmidt, B.C., Jain, A., Basset, Y., Beery, S., Larrivée, M. & Rolnick, D. (2025) Open-set recognition of novel species in biodiversity monitoring. *arXiv [cs.CV]*. Http://arxiv.org/abs/2503.01691.

Chiaburu, T., Haußer, F. & Bießmann, F. (2024) Uncertainty in XAI: Human perception and modeling approaches. *Machine learning and knowledge extraction* **6**, 1170–1192. MDPI AG.

Chiavassa, J.A., Kraft, M., Noack, P., Walther, S., Kirse, A. & Scherber, C. (2024) The FAIR-Device - a non-lethal and generalist semi-automatic Malaise trap for insect biodiversity monitoring: Proof of concept. *bioRxiv*. Https://www.biorxiv.org/content/biorxiv/early/2024/03/27/2024.03.22.586299 [accessed 18 April 2024].

Chiranjeevi, S., Saadati, M., Deng, Z.K., Koushik, J., Jubery, T.Z., Mueller, D.S., O'Neal, M., Merchant, N., Singh, A., Singh, A.K., Sarkar, S., Singh, A. & Ganapathysubramanian, B. (2025) InsectNet: Real-time identification of insects using an end-to-end machine learning pipeline. *PNAS nexus* **4**, pgae575. Oxford University Press (OUP).

Chua, P.Y.S., Bourlat, S.J., Ferguson, C., Korlevic, P., Zhao, L., Ekrem, T., Meier, R. & Lawniczak, M.K.N. (2023) Future of DNA-based insect monitoring. *Trends in genetics: TIG* **39**, 531–544. Elsevier BV.

Clarin, B.-M., Bitzilekis, E., Siemers, B.M. & Goerlitz, H.R. (2014) Personal messages reduce vandalism and theft of unattended scientific equipment. *Methods in ecology and evolution* **5**, 125–131. Wiley.

Collen, B., Ram, M., Zamin, T. & McRae, L. (2008) The Tropical Biodiversity Data Gap: Addressing Disparity in Global Monitoring. *Tropical Conservation Science*. SAGE PublicationsSage CA: Los Angeles, CA.

Comanici, G., Bieber, E., Schaekermann, M., Pasupat, I., Sachdeva, N., Dhillon, I., Blistein, M., Ram, O., Zhang, D., Rosen, E., Marris, L., Petulla, S., Gaffney, C., Aharoni, A., Lintz, N., et al. (2025) Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *arXiv [cs.CL]*. Http://dx.doi.org/10.48550/arXiv.2507.06261.

Conference of the Parties (2022) COP15: Final text of Kunming-Montreal Global Biodiversity Framework. *Official CBD Press Release*. Https://www.cbd.int/article/cop15-final-text-kunming-montreal-gbf-221222 [accessed 22 January 2026].

Conrad, K.F., Warren, M.S., Fox, R., Parsons, M.S. & Woiwod, I.P. (2006) Rapid declines of common, widespread British moths provide evidence of an insect biodiversity crisis. *Biological conservation* **132**, 279–291. Elsevier BV.

Cowans, Lambin, X., Hare, D. & Sutherland, C. (2022) Improving integration artificial intelligence existing ecological inference workflows. *Methods Ecology Evolution* **13**, 1778–1789.

Darras, K.F.A., Balle, M., Xu, W., Yan, Y., Zakka, V.G., Toledo-Hernández, M., Sheng, D., Lin, W., Zhang, B., Lan, Z., Fupeng, L. & Wanger, T.C. (2024) Eyes on nature: Embedded vision cameras for terrestrial biodiversity monitoring. *Methods in ecology and evolution* **15**, 2262–2275. Wiley.

Darras, K.F.A., Rountree, R.A., Van Wilgenburg, S.L., Cord, A.F., Pitz, F., Chen, Y., Dong, L., Rocquencourt, A., Desjonquères, C., Diaz, P.M., Lin, T.-H., Turco, T., Emmerson, L., Bradfer-Lawrence, T., Gasc, A., et al. (2025) Worldwide Soundscapes: A synthesis of Passive Acoustic Monitoring across realms. *Global ecology and biogeography: a journal of macroecology* **34**. Wiley.

Deagle, B.E., Thomas, A.C., McInnes, J.C., Clarke, L.J., Vesterinen, E.J., Clare, E.L., Kartzinel, T.R. & Eveson, J.P. (2019) Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular ecology* **28**, 391–406. Wiley.

Deichmann, J.L., Acevedo-Charry, O., Barclay, L., Burivalova, Z., Campos-Cerqueira, M., d'Horta, F., Game, E.T., Gottesman, B.L., Hart, P.J., Kalan, A.K., Linke, S., Nascimento, L.D., Pijanowski, B., Staaterman, E. & Mitchell Aide, T. (2018) It's time to listen: there is much to be learned from the sounds of tropical ecosystems. *Biotropica* **50**, 713–718. Wiley.

Desjonquères, C., Gifford, T. & Linke, S. (2020) Passive acoustic monitoring as a potential tool to survey animal and ecosystem processes in freshwater environments. *Freshwater biology* **65**, 7–19. Wiley.

Desjonquères, C., Linke, S., Greenhalgh, J., Rybak, F. & Sueur, J. (2024) The potential of acoustic monitoring of aquatic insects for freshwater assessment. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230109. The Royal Society.

Devlin, J., MacFarlane, F., Karley, A., Manfredini, F. & Williams, D. (2025) BeeSAM2: detecting bees in cherry flowers using timelapse images and foundational models. *bioRxiv*. Http://dx.doi.org/10.1101/2025.08.01.668108.

Didham, R.K., Basset, Y., Collins, C.M., Leather, S.R., Littlewood, N.A., Menz, M.H.M., Müller, J., Packer, L., Saunders, M.E., Schönrogge, K., Stewart, A.J.A., Yanoviak, S.P. & Hassall, C. (2020) Interpreting insect declines: seven challenges and a way forward. *Insect conservation and diversity* **13**, 103–114. Wiley.

Directorate-General for Environment (2022) Nature restoration law. Publications Office of the European Union. Https://data.europa.eu/doi/10.2779/86148 [accessed 23 June 2025].

Dokter, A.M., Desmet, P., Spaaks, J.H., van Hoey, S., Veen, L., Verlinden, L., Nilsson, C., Haase, G., Leijnse, H., Farnsworth, A., Bouten, W. & Shamoun-Baranes, J. (2019) bioRad: biological analysis and visualization of weather radar data. *Ecography* **42**, 852–860. Wiley.

Do Nascimento, L.A., Pérez-Granados, C., Alencar, J.B.R. & Beard, K.H. (2024) Time and habitat structure shape insect acoustic activity in the Amazon. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230112. The Royal Society.

Drake, A. & Reynolds, D. (2012) *Radar entomology: Observing insect flight and migration*. Cabi.

Drake, V.A. (2016) Distinguishing target classes in observations from vertically pointing entomological radars. *International journal of remote sensing* **37**, 3811–3835. Informa UK Limited.

Drake, V.A., Hao, Z. & Wang, H. (2024) Monitoring insect numbers and biodiversity with a vertical-beam entomological radar. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230117. The Royal Society.

Droissart, V., Azandi, L., Onguene, E.R., Savignac, M., Smith, T.B. & Deblauwe, V. (2021) PICT: A low-cost, modular, open-source camera trap system to study plant–insect interactions. *Methods in ecology and evolution* **12**, 1389–1396. Wiley.

Durdik, Z., Klatt, B., Koziolek, H., Krogmann, K., Stammel, J. & Weiss, R. (2012) Sustainability guidelines for long-living software systems. *2012 28th IEEE International Conference on Software Maintenance (ICSM)*, 517–526.

Dyer, L.A., Smilanich, A.M., Gompert, Z. & Forister, M.L. (2024) Insect conservation, technological traps, and the fading arts of natural history and field ecology. *Current opinion in insect science*

66, 101261. Elsevier BV.

Engel, M.S., Ceríaco, L.M.P., Daniel, G.M., Dellapé, P.M., Löbl, I., Marinov, M., Reis, R.E., Young, M.T., Dubois, A., Agarwal, I., Lehmann A., P., Alvarado, M., Alvarez, N., Andreone, F., Araujo-Vieira, K., et al. (2021) The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. *Zoological journal of the Linnean Society* **193**, 381–387. Oxford University Press (OUP).

Everingham, M., Eslami, S.M.A., Van Gool, L., Williams, C.K.I., Winn, J. & Zisserman, A. (2014) The Pascal Visual Object Classes Challenge: A Retrospective. *International Journal of Computer Vision* **111**, 98–136. Springer.

Fithian, W., Elith, J., Hastie, T. & Keith, D.A. (2015) Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in ecology and evolution* **6**, 424–438. Wiley.

Foody, G.M. (2024) Ground truth in classification accuracy assessment: Myth and reality. *Geomatics* **4**, 81–90. MDPI AG.

Ford, H.V., Schrodt, F., Zieritz, A., Exton, D.A., van der Heijden, G., Teague, J., Coles, T. & Field, R. (2024) A technological biodiversity monitoring toolkit for biocredits. *The journal of applied ecology* **61**, 2007–2019. Wiley.

Funosas, D., Sebastián-González, E., Morant, J., Marín Gómez, O.H., Mendoza, I., Mohedano-Muñoz, M.A., Santamaría, E., Bastianelli, G., Márquez-Rodríguez, A., Budka, M., Bota, G., Alonso-Moya, C.D., de la Peña-Rubio, J.M., de la Morena, E.L.G., Santa-Cruz, M., et al. (2026) A global assessment of BirdNET performance: Differences among continents, biomes, and species. *Ecological Indicators* **182**, 114550. Elsevier BV.

Gal, Y. & Ghahramani, Z. (2015) Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. Http://arxiv.org/abs/1506.02142 [accessed 28 October 2025].

Gao, B., Wotton, K.R., Hawkes, W.L.S., Menz, M.H.M., Reynolds, D.R., Zhai, B.-P., Hu, G. & Chapman, J.W. (2020) Adaptive strategies of high-flying migratory hoverflies in response to wind currents. *Proceedings. Biological sciences* **287**, 20200406. The Royal Society.

Gardiner, R.J., Mougeot, G., Rowlands, S., Simmons, B.I., Helsing, F. & Høye, T.T. (2025a) Bridging domain gaps for fine-grained moth classification through expert-informed adaptation and foundation model priors. *arXiv [cs.CV]*. Http://arxiv.org/abs/2508.20089.

Gardiner, R.J., Rowlands, S. & Simmons, B.I. (2025b) Towards scalable insect monitoring: Ultra-lightweight CNNs as on-device triggers for insect camera traps. *Methods in ecology and evolution*. Wiley.

Gauthreaux, S. & Diehl, R. (2020) Discrimination of biological scatterers in polarimetric weather radar data: Opportunities and challenges. *Remote sensing* **12**, 545. MDPI AG.

Gawlikowski, J., Tassi, C.R.N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R. & Zhu, X.X. (2023) A survey of uncertainty in deep neural networks. *Artificial intelligence review* **56**, 1513–1589. Springer Science and Business Media LLC.

GBIF Secretariat (2023) GBIF Backbone Taxonomy. *Checklist dataset*. Https://www.gbif.org/dataset/d7dddbf4-2cf0-4f39-9b2a-bb099caae36c [accessed 25 April 2025].

Geijzendorffer, I.R., Targetti, S., Schneider, M.K., Brus, D.J., Jeanneret, P., Jongman, R.H.G., Knotters, M., Viaggi, D., Angelova, S., Arndorfer, M., Bailey, D., Balázs, K., Báldi, A., Bogers, M.M.B., Bunce, R.G.H., et al. (2016) EDITOR'S CHOICE: How much would it cost to monitor farmland biodiversity in Europe? *The journal of applied ecology* **53**, 140–149. Wiley.

Geirhos, R., Jacobsen, J.H., Michaelis, C., Zemel, R., Brendel, W., Bethge, M. & Wichmann, F.A. (2020) Shortcut learning in deep neural networks. *Nature Machine Intelligence* **2**, 665–673.

Geissmann, Q., Abram, P.K., Wu, D., Haney, C.H. & Carrillo, J. (2022) Sticky Pi is a high-frequency smart trap that enables the study of insect circadian activity under natural conditions. *PLoS biology* **20**, e3001689.

Genoud, P., Gao, Y., Williams, G.M. & Thomas, B.P. (2019) Identification gravid mosquitoes from changes spectral polarimetric backscatter cross-sections. *Journal Biophotonics*.

Ghani, B., Denton, T., Kahl, S. & Klinck, H. (2023) Global birdsong embeddings enable superior transfer learning for bioacoustic classification. *arXiv [eess.AS]*. Http://arxiv.org/abs/2307.06292.

Gharaee, Z., Gong, Z., Pellegrino, N., Zarubiieva, I., Haurum, J.B., Lowe, S.C., McKeown, J.T.A., Ho, C.C.Y., McLeod, J., Wei, Y.-Y.C., Agda, J., Ratnasingham, S., Steinke, D., Chang, A.X., Taylor, G.W., et al. (2023) A step towards worldwide biodiversity assessment: The BIOSCAN-1M Insect Dataset. *arXiv [cs.CV]*. Http://arxiv.org/abs/2307.10455.

Gharaee, Z., Lowe, S.C., Gong, Z., Arias, P.M., Pellegrino, N., Wang, A.T., Haurum, J.B., Zarubiieva, I., Kari, L., Steinke, D., Taylor, G.W., Fieguth, P. & Chang, A.X. (2024) BIOSCAN-5M: A Multimodal Dataset for Insect Biodiversity. *arXiv [cs.LG]*. Http://dx.doi.org/10.48550/ARXIV.2406.12723.

Gibb, R., Browning, E., Glover-Kapfer, P. & Jones, K.E. (2019) Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in ecology and evolution* **10**, 169–185. Wiley.

Gillespie, M.A.K., Bjerge, K., Alison, J., Gerlich, H.S., Helsing-Nielsen, F., Mougeot, G., Svenning, A., Wogram, S.F.A. & Høye, T.T. (2025) Automated insect monitoring with camera traps is transforming ecological understanding. *EcoEvoRxiv*. Https://ecoevorxiv.org/repository/view/10577/ [accessed 5 November 2025].

Giuntini, S., Saari, J., Martinoli, A., Preatoni, D.G., Haest, B., Schmid, B. & Weisshaupt, N. (2024) Quantifying nocturnal thrush migration using sensor data fusion between acoustics and vertical-looking radar. *Remote sensing in ecology and conservation* **10**, 743–754. Wiley.

Goh, B., Ching, K., Soares Magalhães, R.J., Ciocchetta, S., Edstein, M.D., Maciel-de-Freitas, R. & Sikulu-Lord, M.T. (2021) The application of spectroscopy techniques for diagnosis of malaria parasites and arboviruses and surveillance of mosquito vectors: A systematic review and critical appraisal of evidence. *PLoS neglected tropical diseases* **15**, e0009218.

Gomez-Morales, D.A. & Acevedo-Charry, O. (2022) Satellite remote sensing of environmental variables can predict acoustic activity of an orthopteran assemblage. *PeerJ* **10**, e13969. PeerJ.

Gradišek, A., Slapničar, G., Šorn, J., Luštrek, M., Gams, M. & Grad, J. (2017) Predicting species identity of bumblebees through analysis of flight buzzing sounds. *Bioacoustics* **26**, 63–76. Taylor & Francis.

Gu, J., Stevens, S., Campolongo, E.G., Thompson, M.J., Zhang, N., Wu, J., Kopanev, A., Mai, Z., White, A.E., Balhoff, J., Dahdul, W., Rubenstein, D., Lapp, H., Berger-Wolf, T., Chao, W.-L., et al. (2025) BioCLIP 2: Emergent properties from scaling hierarchical contrastive learning. *arXiv [cs.CV]*. Http://dx.doi.org/10.48550/arXiv.2505.23883.

Guralnick, R., Walls, R. & Jetz, W. (2018) Humboldt Core - toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. *Ecography* **41**, 713–725. Wiley.

Haelewaters, D., Hofmann, T.A. & Romero-Olivares, A.L. (2021) Ten simple rules for Global North researchers to stop perpetuating helicopter research in the Global South. *PLoS computational biology* **17**, e1009277.

Haest, B., Liechti, F., Hawkes, W.L., Chapman, J., Åkesson, S., Shamoun-Baranes, J., Nesterova, A.P., Comor, V., Preatoni, D. & Bauer, S. (2024) Continental-scale patterns in diel flight timing of high-altitude migratory insects. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230116. The Royal Society.

Haest, B., Schmid, B., Hertner, F., Aschwanden, J., Früh, D., Kleger, D., Peter, D., Sapir, N., Troupin, D., Werber, Y., Yovel, Y. & Liechti, F. (2021) BirdScan Community Reference Dataset. Zenodo. Http://dx.doi.org/10.5281/ZENODO.5734961.

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D. & de Kroon, H. (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS one* **12**, e0185809. Public Library of Science (PLoS).

Hansen, O.L.P., Svenning, J.-C., Olsen, K., Dupont, S., Garner, B.H., Iosifidis, A., Price, B.W. & Høye, T.T. (2020) Species-level image classification with convolutional neural network enables insect identification from habitus images. *Ecology and evolution* **10**, 737–747. Wiley.

Hao, Z., Drake, V.A., Taylor, J.R. & Warrant, E. (2020) Insect target classes discerned from entomological radar data. *Remote sensing* **12**, 673. MDPI AG.

Hartig, F., Abrego, N., Bush, A., Chase, J.M., Guillera-Arroita, G., Leibold, M.A., Ovaskainen, O., Pellissier, L., Pichler, M., Poggiato, G., Pollock, L., Si-Moussi, S., Thuiller, W., Viana, D.S., Warton, D.I., et al. (2024) Novel community data in ecology-properties and prospects. *Trends in ecology & evolution* **39**, 280–293. Elsevier BV.

Harvey, J.A., Heinen, R., Armbrecht, I., Basset, Y., Baxter-Gilbert, J.H., Bezemer, T.M., Böhm, M., Bommarco, R., Borges, P.A.V., Cardoso, P., Clausnitzer, V., Cornelisse, T., Crone, E.E., Dicke, M., Dijkstra, K.-D.B., et al. (2020) International scientists formulate a roadmap for insect conservation and recovery. *Nature ecology & evolution* **4**, 174–176. Springer Science and Business Media LLC.

Hein, M., Andriushchenko, M. & Bitterwolf, J. (2019) Why ReLU networks yield high-confidence predictions far away from the training data and how to mitigate the problem. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

Heistermann, M., Jacobi, S. & Pfaff, T. (2013) Technical Note: An open source library for processing weather radar data (*wradlib*). *Hydrology and earth system sciences* **17**, 863–871. Copernicus GmbH.

He, K., Chen, X., Xie, S., Li, Y., Dollár, P. & Girshick, R. (2021) Masked Autoencoders Are Scalable Vision Learners. *arXiv [cs.CV]*. Http://arxiv.org/abs/2111.06377.

Hellendoorn, V.J. & Sawant, A.A. (2022) The growing cost of deep learning for source code.

Communications of the ACM **65**, 31–33. Association for Computing Machinery (ACM).

Hertzog, L.R., Frank, C., Klimek, S., Röder, N., Böhner, H.G.S. & Kamp, J. (2021) Model-based integration of citizen science data from disparate sources increases the precision of bird population trends. *Diversity & Distributions* **27**, 1106–1119. Wiley.

Hill, A.P., Prince, P., Covarrubias, E.P., Patrick Doncaster, C., Snaddon, J.L. & Rogers, A. (2018) AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution* **9**, 1199–1211. John Wiley & Sons, Ltd.

Hobbs, M.T. & Brehme, C.S. (2017) An improved camera trap for amphibians, reptiles, small mammals, and large invertebrates. *PloS One* **12**, e0185026. Public Library of Science (PLoS).

Hogeweg, L.E., Gangireddy, R., Brunink, D., Kalkman, V.J., Cornelissen, L. & Kamminga, J.W. (2024) COOD: Combined out-of-distribution detection using multiple measures for anomaly & novel class detection in large-scale hierarchical classification. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* pp. 3971–3980. IEEE.

Hörren, T., Sorg, M., Hallmann, C.A., Zizka, V.M.A., Ssymank, A., Noll, N.W., Schäffler, L. & Scherber, C. (2022) A universal insect trait tool (ITT, v1.0) for statistical analysis and evaluation of biodiversity research data. *bioRxiv*. Http://dx.doi.org/10.1101/2022.01.25.477751.

Høye, T.T., Ärje, J., Bjerge, K., Hansen, O.L.P., Iosifidis, A., Leese, F., Mann, H.M.R., Meissner, K., Melvad, C. & Raitoharju, J. (2021) Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences of the United States of America* **118**, 1–10.

Høye, T.T., Montagna, M., Oteman, B. & Roy, D.B. (2025) Emerging technologies for pollinator monitoring. *Current opinion in insect science* **69**, 101367. Elsevier BV.

Hu, G., Lim, K.S., Horvitz, N., Clark, S.J., Reynolds, D.R., Sapir, N. & Chapman, J.W. (2016) Mass seasonal bioflows of high-flying insect migrants. *Science (New York, N.Y.)* **354**, 1584–1587.

Hüppop, O., Ciach, M., Diehl, R., Reynolds, D.R., Stepanian, P.M. & Menz, M.H.M. (2019) Perspectives and challenges for the use of radar in biological conservation. *Ecography* **42**, 912–930. Wiley.

Isaac, N.J.B., Jarzyna, M.A., Keil, P., Dambly, L.I., Boersch-Supan, P.H., Browning, E., Freeman, S.N., Golding, N., Guillera-Arroita, G., Henrys, P.A., Jarvis, S., Lahoz-Monfort, J., Pagel, J., Pescott, O.L., Schmucki, R., et al. (2020) Data integration for large-scale models of species distributions. *Trends in ecology & evolution* **35**, 56–67. Elsevier BV.

Isaac, N.J.B. & Pocock, M.J.O. (2015) Bias and information in biological records. *Biological journal of the Linnean Society. Linnean Society of London* **115**, 522–531. Oxford University Press (OUP).

Itskov, P.M., Moreira, J.-M., Vinnik, E., Lopes, G., Safarik, S., Dickinson, M.H. & Ribeiro, C. (2014) Automated monitoring and quantitative analysis of feeding behaviour in Drosophila. *Nature communications* **5**, 4560. Springer Science and Business Media LLC.

Jain, A., Cunha, F., Bunsen, M.J., Cañas, J.S., Pasi, L., Pinoy, N., Helsing, F., Russo, J., Botham, M., Sabourin, M., Fréchette, J., Anctil, A., Lopez, Y., Navarro, E., Pimentel, F.P., et al. (2025) Insect identification in the wild: The AMI dataset. In *Lecture Notes in Computer Science* pp. 55–73. Springer Nature Switzerland, Cham.

Jansson, S. & Brydegaard, M. (2018) Passive kHz lidar for the quantification of insect activity and dispersal. *Animal biotelemetry* **6**. Springer Science and Business Media LLC.

Jeffries, D.L., Chapman, J., Roy, H.E., Humphries, S., Harrington, R., Brown, P.M.J. & Handley, L.-J.L. (2013) Characteristics and drivers of high-altitude ladybird flight: insights from vertical-looking entomological radar. *PloS one* **8**, e82278.

Kapoor, S. & Narayanan, A. (2023) Leakage and the reproducibility crisis in machine-learning-based science. *Patterns (New York, N.Y.)* **4**, 100804. Elsevier BV.

Karlsson, D., Hartop, E., Forshage, M., Jaschhof, M. & Ronquist, F. (2020) The Swedish Malaise Trap Project: A 15 year retrospective on a countrywide insect inventory. *Biodiversity data journal* **8**, e47255. pmc.ncbi.nlm.nih.gov.

Kath, H., Serafini, P.P., Campos, I.B., Gouvêa, T.S. & Sonntag, D. (2024) Leveraging transfer learning and active learning for data annotation in passive acoustic monitoring of wildlife. *Ecological informatics* **82**, 102710. Elsevier BV.

Kay, J., Haucke, T., Stathatos, S., Deng, S., Young, E., Perona, P., Beery, S. & Van Horn, G. (2025) Align and Distill: Unifying and improving domain adaptive object detection. *arXiv [cs.CV]*. Http://arxiv.org/abs/2403.12029.

Kendall-Bar, J.M., Payne, A., Czapanskiy, M., Fox, N., Leivesley, J.A., Alksne, M., French, B., Malivert, A., O'Leary, G., Biswas, A., Nealey, I., Parkinson, R., Lowrie, C., Perez, I., Miller, E., et al. (2025) Challenges and solutions for ecologists adopting AI. EcoEvoRxiv.

Kirkeby, C., Rydhmer, K., Cook, S.M., Strand, A., Torrance, M.T., Swain, J.L., Prangsma, J., Johnen, A., Jensen, M., Brydegaard, M. & Græsbøll, K. (2021) Advances in automatic identification of flying insects using optical sensors and machine learning. *Scientific reports* **11**, 1555. Springer Science and Business Media LLC.

Kissling, W.D., Ahumada, J.A., Bowser, A., Fernandez, M., Fernández, N., García, E.A., Guralnick, R.P., Isaac, N.J.B., Kelling, S., Los, W., McRae, L., Mihoub, J.-B., Obst, M., Santamaria, M., Skidmore, A.K., et al. (2018) Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biological reviews of the Cambridge Philosophical Society* **93**, 600–625. Wiley.

Kline, J., O'Quinn, A., Berger-Wolf, T. & Stewart, C. (2024) Characterizing and modeling AI-driven animal ecology studies at the edge. *arXiv [eess.SY]*. Http://dx.doi.org/10.48550/arXiv.2412.01000.

Kline, J., Potlapally, A., Pillai, B., Wani, T., Katole, R., Patil, V., Covey, P., Subramoni, H., Berger-Wolf, T. & Stewart, C. (2025) SmartWilds: Multimodal Wildlife Monitoring Dataset. *arXiv [cs.CV]*. Http://dx.doi.org/10.48550/arXiv.2509.18894.

van Klink, R. (2024) Delivering on a promise: futureproofing automated insect monitoring methods. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230105. The Royal Society.

van Klink, R., August, T., Bas, Y., Bodesheim, P., Bonn, A., Fossøy, F., Høye, T.T., Jongejans, E., Menz, M.H.M., Miraldo, A., Roslin, T., Roy, H.E., Ruczyński, I., Schigel, D., Schäffler, L., et al. (2022) Emerging technologies revolutionise insect ecology and monitoring. *Trends in ecology & evolution* **37**, 872–885. Elsevier BV.

van Klink, R., Bowler, D.E., Gongalsky, K.B., Swengel, A.B., Gentile, A. & Chase, J.M. (2020) Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances.

Science (New York, N.Y.) **368**, 417–420.

van Klink, R., Sheard, J.K., Høye, T.T., Roslin, T., Do Nascimento, L.A. & Bauer, S. (2024) Towards a toolkit for global insect biodiversity monitoring. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230101. The Royal Society.

Knop, E., Grimm, M.L., Korner-Nievergelt, F., Schmid, B. & Liechti, F. (2023) Patterns of high-flying insect abundance are shaped by landscape type and abiotic conditions. *Scientific reports* **13**, 15114. Springer Science and Business Media LLC.

Koch, W., Hogeweg, L., Nilsen, E.B. & Finstad, A.G. (2022) Maximizing citizen scientists' contribution to automated species recognition. *Scientific reports* **12**, 7648.

Kohlberg, A.B., Myers, C.R. & Figueroa, L.L. (2024) From buzzes to bytes: A systematic review of automated bioacoustics models used to detect, classify and monitor insects. *The Journal of Applied Ecology* **61**, 1199–1211. Wiley.

Kranstauber, B., Huybrechts, P. & Desmet, P. (2025) GetRad: Download radar data for biological research. The R Foundation. *CRAN: Contributed Packages*. Http://dx.doi.org/10.32614/cran.package.getrad.

Kumar, A., Raghunathan, A., Jones, R., Ma, T. & Liang, P. (2022) Fine-tuning can distort pretrained features and underperform out-of-distribution. *arXiv [cs.LG]*. Http://arxiv.org/abs/2202.10054.

Lamont, T.A.C., Chapuis, L., Williams, B., Dines, S., Gridley, T., Frainer, G., Fearey, J., Maulana, P.B., Prasetya, M.E., Jompa, J., Smith, D.J. & Simpson, S.D. (2022) HydroMoth: Testing a prototype low-cost acoustic recorder for aquatic environments. *Remote sensing in ecology and conservation* **8**, 362–378. Wiley.

Lawson, J., Whitworth, A. & Banks-Leite, C. (2022) Soundscapes show disruption across the diel cycle in human modified tropical landscapes. *Ecological indicators* **144**, 109413. Elsevier BV.

Lee, K., Lee, K., Min, K., Zhang, Y., Shin, J. & Lee, H. (2018) Hierarchical novelty detection for visual object recognition. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.

Li, M., Månefjord, H., Hernandez, J., Müller, L., Brackmann, C., Merdasa, A., Kirkeby, C., Bulo, M.D., Ignell, R. & Brydegaard, M. (2025) Deadliest Animals with the Thinnest Wings: Near-Infrared Properties of Tropical Mosquitoes. *Applied spectroscopy*, 37028251341317.

Li, M., Runemark, A., Hernandez, J., Rota, J., Bygebjerg, R. & Brydegaard, M. (2023a) Discrimination of hover fly species and sexes by wing interference signals. *Advanced science (Weinheim, Baden-Wurttemberg, Germany)* **10**, e2304657. Wiley.

Li, M., Seinsche, C., Jansson, S., Hernandez, J., Rota, J., Warrant, E. & Brydegaard, M. (2022) Potential for identification of wild night-flying moths by remote infrared microscopy. *Journal of the Royal Society, Interface* **19**, 20220256. The Royal Society.

Linke, S., Gifford, T., Desjonquères, C., Tonolla, D., Aubin, T., Barclay, L., Karaconstantis, C., Kennard, M.J., Rybak, F. & Sueur, J. (2018) Freshwater ecoacoustics as a tool for continuous ecosystem monitoring. *Frontiers in ecology and the environment* **16**, 231–238. Wiley.

Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., Jiang, Q., Li, C., Yang, J., Su, H., Zhu, J. & Zhang, L. (2023) Grounding DINO: Marrying DINO with grounded pre-training for open-set object detection. *arXiv [cs.CV]*. Http://arxiv.org/abs/2303.05499.

Li, X. & Wiens, J.J. (2023) Estimating global biodiversity: The role of cryptic insect species.

Systematic biology **72**, 391–403. Oxford University Press (OUP).

Li, Y., Han, Z., Nessler, R., Yi, Z., Hemmer, P., Brick, R., Sokolov, A.V. & Scully, M.O. (2023b) Optical multiband polarimetric modulation sensing for gender and species identification of flying native solitary pollinators. *iScience* **26**, 108265. Elsevier BV.

Li, Y., Wang, K., Quintero-Torres, R., Brick, R., Sokolov, A.V. & Scully, M.O. (2020) Insect flight velocity measurement with a CW near-IR Scheimpflug lidar system. *Optics express* **28**, 21891–21902. Optica Publishing Group.

Lochmann, M., Kalesse-Los, H., Haest, B., Vogl, T., van Klink, R., Addison, F., Maahn, M., Schimmel, W., Wirth, C. & Quaas, J. (2024) Derivation of aerial insect concentration with a 94GHz FMCW cloud radar. *bioRxiv*. Http://dx.doi.org/10.1101/2024.07.30.605781.

Lövei, G.L. & Ferrante, M. (2024) The use and prospects of nonlethal methods in entomology. *Annual review of entomology* **69**, 183–198.

Luccioni, A.S., Jernite, Y. & Strubell, E. (2023) Power hungry processing: Watts driving the cost of AI deployment? *arXiv [cs.LG]*. Http://arxiv.org/abs/2311.16863.

Lukach, M., Dally, T., Evans, W., Duncan, E.J., Bennett, L., Addison, F.I., Kunin, W.E., Chapman, J.W., Neely, R.R., III & Hassall, C. (2024) Operationalising weather surveillance radar data for use in ecological research. *Ecological informatics* **84**, 102901. Elsevier BV.

Lukach, M., Dally, T., Evans, W., Hassall, C., Duncan, E.J., Bennett, L., Addison, F.I., Kunin, W.E., Chapman, J.W. & Neely, R.R., 3rd (2022) The development of an unsupervised hierarchical clustering analysis of dual-polarization weather surveillance radar observations to assess nocturnal insect abundance and diversity. *Remote sensing in ecology and conservation* **8**, 698–716. Wiley.

Lürig, M.D., Donoughe, S., Svensson, E.I., Porto, A. & Tsuboi, M. (2021) Computer Vision, Machine Learning, and the Promise of Phenomics in Ecology and Evolutionary Biology. *Frontiers in Ecology and Evolution* **9**.

Macgregor, C.J., Williams, J.H., Bell, J.R. & Thomas, C.D. (2021) Author Correction: Moth biomass has fluctuated over 50 years in Britain but lacks a clear trend. *Nature ecology & evolution* **5**, 865–883. Springer Science and Business Media LLC.

Madhusudhana, S., Klinck, H. & Symes, L.B. (2024) Extensive data engineering to the rescue: building a multi-species katydid detector from unbalanced, atypical training datasets. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230444. The Royal Society.

Mammola, S., Riccardi, N., Prié, V., Correia, R., Cardoso, P., Lopes-Lima, M. & Sousa, R. (2020) Towards a taxonomically unbiased European Union biodiversity strategy for 2030. *Proceedings. Biological sciences* **287**, 20202166. The Royal Society.

Månefjord, H., Huzortey, A.A., Boateng, R., Gbogbo, Y.A., Yamoa, A.S.D., Zoueu, J.T., Kwapong, P.K., Anderson, B. & Brydegaard, M. (2024) Hyperspectral lidar for monitoring high-resolution activity patterns of African stingless bee species. *Animal biotelemetry* **12**. Springer Science and Business Media LLC.

Mankin, R.W., Hagstrum, D.W., Smith, M.T., Roda, A.L. & Kairo, M.T.K. (2011) Perspective and promise: A century of insect acoustic detection and monitoring. *American Entomologist* **57**, 30–44. Oxford University Press (OUP).

Mata, V.A., Ferreira, S., Campos, R.M., da Silva, L.P., Veríssimo, J., Corley, M.F.V. & Beja, P. (2021) Efficient assessment of nocturnal flying insect communities by combining automatic light traps and DNA metabarcoding. *Environmental DNA (Hoboken, N.J.)* **3**, 398–408. Wiley.

Mazen, F.M.A. (2023) Arthropod Taxonomy Orders Object Detection in ArTaxOr dataset using YOLOX. *Journal of Engineering and Applied Science* **70**. Springer Science and Business Media LLC.

McCravy, K.W. (2018) A review of sampling and monitoring methods for beneficial arthropods in agroecosystems. *Insects* **9**.

Meek, P.D., Ballard, G.A., Sparkes, J., Robinson, M., Nesbitt, B. & Fleming, P.J.S. (2019) Camera trap theft and vandalism: occurrence, cost, prevention and implications for wildlife research and management. *Remote sensing in ecology and conservation* **5**, 160–168. Wiley.

Mehrotra, D.R., Still, L., Agrawal, V., Gibson, K., Crall, J. & Bick, E.N. (2024) Eavesdropping on Herbivores: Using contact microphones to quantify Plant-Insect Interactions. *bioRxiv*.

Miao, Z., Zhang, Y., Fabian, Z., Hernandez Celis, A., Beery, S., Li, C., Liu, Z., Gupta, A., Nasir, M., Li, W., Holmberg, J., Palmer, M., Gaynor, K., Arbelaez, P., Wang, P., et al. (2025) New frontiers in artificial intelligence for biodiversity research and conservation with multimodal language models. *Methods in Ecology and Evolution*. Wiley.

Miller, D.A.W., Pacifici, K., Sanderlin, J.S. & Reich, B.J. (2019) The recent past and promising future for data integration methods to estimate species' distributions. *Methods in ecology and evolution* **10**, 22–37. Wiley.

Montgomery, G.A., Belitz, M.W., Guralnick, R.P. & Tingley, M.W. (2021) Standards and Best Practices for Monitoring and Benchmarking Insects. *Frontiers in Ecology and Evolution* **8**. frontiersin.org.

Moretti, M., Dias, A.T.C., de Bello, F., Altermatt, F., Chown, S.L., Azcárate, F.M., Bell, J.R., Fournier, B., Hedde, M., Hortal, J., Ibanez, S., Öckinger, E., Sousa, J.P., Ellers, J. & Berg, M.P. (2017) Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology* **31**, 558–567. John Wiley & Sons, Ltd.

Morris, R.A., Barve, V., Carausu, M., Chavan, V., Cuadra, J., Freeland, C., Hagedorn, G., Leary, P., Mozzherin, D., Olson, A., Riccardi, G., Teage, I. & Whitbread, G. (2013) Discovery and publishing of primary biodiversity data associated with multimedia resources: The Audubon Core strategies and approaches. *Biodiversity informatics* **8**. The University of Kansas.

Moser, M., Krogmann, L. & Wanke, D. (2025) Some Insects Are More Equal Than Others: A Comparison of Popular Large Language Model Chatbots' Treatment of Different Insect Groups. *Integrative Conservation* **4**, 254–267.

Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J. & Fernández-Leal, Á. (2023) Human-in-the-loop machine learning: a state of the art. *Artificial intelligence review* **56**, 3005–3054. Springer Science and Business Media LLC.

Mukundarajan, H., Hol, F.J.H., Castillo, E.A., Newby, C. & Prakash, M. (2017) Using mobile phones as acoustic sensors for high-throughput mosquito surveillance. *eLife* **6**.

Müller, L., Li, M., Månefjord, H., Salvador, J., Reistad, N., Hernandez, J., Kirkeby, C., Runemark, A. & Brydegaard, M. (2023) Remote nanoscopy with infrared elastic hyperspectral lidar. *Advanced science (Weinheim, Baden-Wurttemberg, Germany)* **10**, e2207110. Wiley.

Mungee, M., Lukach, M., Shortall, C., Bell, J.R., Duncan, E.J., Addison, F.I., Brown, L.E., Kunin, W.E., Hassall, C. & Neely, R.R., 3rd (2025) Spatio-temporal variation in aerial arthropod abundance revealed by weather radars. *Global change biology* **31**, e70425. Wiley.

Naqvi, Q., Wolff, P.J., Molano-Flores, B. & Sperry, J.H. (2022) Camera traps are an effective tool for monitoring insect-plant interactions. *Ecology and evolution* **12**, e8962. Wiley.

Neff, F., Korner-Nievergelt, F., Rey, E., Albrecht, M., Bollmann, K., Cahenzli, F., Chittaro, Y., Gossner, M.M., Martínez-Núñez, C., Meier, E.S., Monnerat, C., Moretti, M., Roth, T., Herzog, F. & Knop, E. (2022) Different roles of concurring climate and regional land-use changes in past 40 years' insect trends. *Nature Communications* **13**, 7611. Springer Science and Business Media LLC.

Nguyen, A., Yosinski, J. & Clune, J. (2015) Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

Nguyen, H.-Q., Truong, T.-D., Nguyen, X.B., Dowling, A., Li, X. & Luu, K. (2024) Insect-foundation: A foundation model and large-scale 1M dataset for visual insect understanding. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* pp. 21945–21955. IEEE.

Noriega, J.A., Hortal, J., Azcárate, F.M., Berg, M.P., Bonada, N., Briones, M.J.I., Del Toro, I., Goulson, D., Ibanez, S., Landis, D.A., Moretti, M., Potts, S.G., Slade, E.M., Stout, J.C., Ulyshen, M.D., et al. (2018) Research trends in ecosystem services provided by insects. *Basic and Applied Ecology* **26**, 8–23. Urban & Fischer.

Ocampo-Ariza, C., Toledo-Hernández, M., Librán-Embid, F., Armenteras, D., Vansynghel, J., Raveloaritiana, E., Arimond, I., Angulo-Rubiano, A., Tscharntke, T., Ramírez-Castañeda, V., Wurz, A., Marcacci, G., Anders, M., Urbina-Cardona, J.N., de Vos, A., et al. (2023) Global South leadership towards inclusive tropical ecology and conservation. *Perspectives in Ecology and Conservation* **21**, 17–24. Elsevier.

Okamoto, R. & Oguma, H. (2025) A simulation based approach for enhancing automatic detection of cicada songs in challenging chorus conditions. *Ecological informatics* **90**, 103299. Elsevier BV.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., et al. (2023) GPT-4 Technical Report. *arXiv [cs.CL]*. Http://arxiv.org/abs/2303.08774.

O'Shea-Wheller, T.A., Corbett, A., Osborne, J.L., Recker, M. & Kennedy, P.J. (2024) VespAI: a deep learning-based system for the detection of invasive hornets. *Communications biology* **7**, 354.

Outhwaite, C.L., Gregory, R.D., Chandler, R.E., Collen, B. & Isaac, N.J.B. (2020) Complex long-term biodiversity change among invertebrates, bryophytes and lichens. *Nature ecology & evolution* **4**, 384–392. Springer Science and Business Media LLC.

Outhwaite, C.L., Powney, G.D., August, T.A., Chandler, R.E., Rorke, S., Pescott, O.L., Harvey, M., Roy, H.E., Fox, R., Roy, D.B., Alexander, K., Ball, S., Bantock, T., Barber, T., Beckmann, B.C., et al. (2019) Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK, 1970-2015. *Scientific data* **6**, 259. Springer Science and Business Media LLC.

Pagel, J., Anderson, B.J., O'Hara, R.B., Cramer, W., Fox, R., Jeltsch, F., Roy, D.B., Thomas, C.D. & Schurr, F.M. (2014) Quantifying range-wide variation in population trends from local abundance surveys and widespread opportunistic occurrence records. *Methods in ecology and evolution* **5**, 751–760. Wiley.

Passias, A., Tsakalos, K.-A., Rigogiannis, N., Voglitsis, D., Papanikolaou, N., Michalopoulou, M., Broufas, G. & Sirakoulis, G.C. (2024) Insect pest trap development and DL-based pest detection: A comprehensive review. *IEEE Transactions on AgriFood Electronics*, 1–12. Institute of Electrical and Electronics Engineers (IEEE).

Pegoraro, L., Hidalgo, O., Leitch, I.J., Pellicer, J. & Barlow, S.E. (2020) Automated video monitoring of insect pollinators in the field. *Emerging Topics in Life Sciences*, 1–11.

Penar, W., Magiera, A. & Klocek, C. (2020) Applications of bioacoustics in animal ecology. *Ecological complexity* **43**, 100847. Elsevier BV.

Pernat, N., Canavan, S., Golivets, M., Hillaert, J., Itescu, Y., Jarić, I., Mann, H.M.R., Pipek, P., Preda, C., Richardson, D.M., Teixeira, H., Vaz, A.S. & Groom, Q. (2024) Overcoming biodiversity blindness: Secondary data in primary citizen science observations. *Ecological Solutions and Evidence* **5**.

Placani, A. (2024) Anthropomorphism in AI: hype and fallacy. *AI and ethics* **4**, 691–698. Springer Science and Business Media LLC.

Polimeni, J.M. & Polimeni, R.I. (2006) Jevons' Paradox and the myth of technological liberation. *Ecological complexity* **3**, 344–353. Elsevier BV.

Pollock, L.J., Kitzes, J., Beery, S., Gaynor, K.M., Jarzyna, M.A., Mac Aodha, O., Meyer, B., Rolnick, D., Taylor, G.W., Tuia, D. & Berger-Wolf, T. (2025) Harnessing artificial intelligence to fill global shortfalls in biodiversity knowledge. *Nature Reviews Biodiversity* **1**, 166–182. Nature Publishing Group.

Potamitis, I. (2023) ChatGPT in the context of precision agriculture data analytics. *arXiv [cs.AI]*. Http://arxiv.org/abs/2311.06390.

Potamitis, I., Rigakis, I. & Tatlas, N.A. (2017) Automated surveillance of fruit flies. *Sensors* **17**, 1–14.

Potamitis, I., Rigakis, I., Vidakis, N., Petousis, M. & Weber, M. (2018) Affordable bimodal optical sensors to spread the use of automated insect monitoring. *Journal of sensors* **2018**, 1–25. Hindawi Limited.

Preston, F.W. (1962) The canonical distribution of commonness and rarity: Part I. *Ecology* **43**, 185. Wiley.

Preti, M., Verheggen, F. & Angeli, S. (2021) Insect pest monitoring with camera-equipped traps: strengths and limitations. *Journal of pest science* **94**, 203–217.

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C. & Chen, M. (2022) Hierarchical Text-Conditional Image Generation with CLIP Latents. Http://arxiv.org/abs/2204.06125 [accessed 5 November 2025].

Ratnayake, M.N., Amarathunga, D.C., Zaman, A., Dyer, A.G. & Dorin, A. (2022) Spatial Monitoring and Insect Behavioural Analysis Using Computer Vision for Precision Pollination. *arXiv [cs.CV]*. Http://arxiv.org/abs/2205.04675.

Ratnayake, M.N., Dyer, A.G. & Dorin, A. (2021) Tracking individual honeybees among wildflower clusters with computer vision-facilitated pollinator monitoring. *PloS one* **16**, e0239504. Public Library of Science (PLoS).

Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K.V., Carion, N., Wu, C.-Y., et al. (2024) SAM 2: Segment Anything in Images and Videos. *arXiv [cs.CV]*. Http://arxiv.org/abs/2408.00714.

Reyserhove, L., Norton, B. & Desmet, P. (2025) Best Practices for Managing and Publishing Camera Trap Data. Https://doi.org/10.35035/doc-0qzp-2x37 [accessed 28 October 2025].

Rhinehart, T.A., Nicholson, D., Freeland-Haynes, L., Nossan, H., Sochat, V. & Kitzes, J. (2024) Bioacoustics Software Database. Zenodo. Http://dx.doi.org/10.5281/ZENODO.7686202.

Rhodes, M.W., Bennie, J.J., Spalding, A., Ffrench-Constant, R.H. & Maclean, I.M.D. (2022) Recent advances in the remote sensing of insects. *Biological Reviews of the Cambridge Philosophical Society* **97**, 343–360. Wiley.

Riede, K. & Balakrishnan, R. (2025) Acoustic monitoring for tropical insect conservation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **380**, 20240046.

Rigakis, I.I., Varikou, K.N., Nikolakakis, A.E., Skarakis, Z.D., Tatlas, N.A. & Potamitis, I.G. (2021) The e-funnel trap: Automatic monitoring of lepidoptera; a case study of tomato leaf miner. *Computers and electronics in agriculture* **185**, 106154. Elsevier BV.

Rivas, F., Brizio, C., Buzzetti, F.M. & Pijanowski, B. (2025) *Rthoptera*: Standardised insect bioacoustics in R. *Methods in ecology and evolution* **16**, 1084–1094. Wiley.

Robinson, J.M., Annells, A., Cavagnaro, T.R., Liddicoat, C., Rogers, H., Taylor, A. & Breed, M.F. (2024) Monitoring soil fauna with ecoacoustics. *Proceedings. Biological sciences* **291**, 20241595. The Royal Society.

Rocha-Ortega, M., Rodriguez, P. & Córdoba-Aguilar, A. (2021) Geographical, temporal and taxonomic biases in insect GBIF data on biodiversity and extinction. *Ecological entomology* **46**, 718–728. Wiley.

Roger, F., Ghanavi, H.R., Danielsson, N., Wahlberg, N., Löndahl, J., Pettersson, L.B., Andersson, G.K.S., Boke Olén, N. & Clough, Y. (2022) Airborne environmental DNA metabarcoding for the monitoring of terrestrial insects—A proof of concept from the field. *Environmental DNA (Hoboken, N.J.)* **4**, 790–807. Wiley.

Ross, S.R.P.-J., O'Connell, D.P., Deichmann, J.L., Desjonquères, C., Gasc, A., Phillips, J.N., Sethi, S.S., Wood, C.M. & Burivalova, Z. (2023) Passive acoustic monitoring provides a fresh perspective on fundamental ecological questions. *Functional ecology* **37**, 959–975. Wiley.

Roy, D.B., Alison, J., August, T.A., Bélisle, M., Bjerge, K., Bowden, J.J., Bunsen, M.J., Cunha, F., Geissmann, Q., Goldmann, K., Gomez-Segura, A., Jain, A., Huijbers, C., Larrivée, M., Lawson, J.L., et al. (2024) Towards a standardized framework for AI-assisted, image-based monitoring of nocturnal insects. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230108. The Royal Society.

Russell, S. & Norvig, P. (2010) *Artificial Intelligence - A Modern Approach*third. Pearson, Upper Saddle River, NJ.

Russo, S., Besmer, M.D., Blumensaat, F., Bouffard, D., Disch, A., Hammes, F., Hess, A., Lürig, M., Matthews, B., Minaudo, C., Morgenroth, E., Tran-Khac, V. & Villez, K. (2021) The value of human data annotation for machine learning based anomaly detection in environmental systems. *Water research*, 117695. Elsevier BV.

Rydhmer, K., Bick, E., Still, L., Strand, A., Luciano, R., Helmreich, S., Beck, B.D., Grønne, C., Malmros, L., Poulsen, K., Elbæk, F., Brydegaard, M., Lemmich, J. & Nikolajsen, T. (2022) Automating insect monitoring using unsupervised near-infrared sensors. *Scientific reports* **12**, 2603.

Rydhmer, K., Jansson, S., Still, L., Beck, B.D., Chatzaki, V., Olsen, K., Van Hoff, B., Grønne, C., Meier, J.K., Montoro, M., Schmidt, I.K., Kirkeby, C., Smith, H.G. & Brydegaard, M. (2024) Photonic sensors reflect variation in insect abundance and diversity across habitats. *Ecological indicators* **158**, 111483. Elsevier BV.

Saha, T., Genoud, A.P., Williams, G.M. & Thomas, B.P. (2023) Monitoring the abundance of flying insects and atmospheric conditions during a 9-month campaign using an entomological optical sensor. *Scientific reports* **13**, 15606. Springer Science and Business Media LLC.

Sánchez-Bayo, F. & Wyckhuys, K.A.G. (2019) Worldwide decline of the entomofauna: A review of its drivers. *Biological conservation* **232**, 8–27. Elsevier BV.

Sánchez Herrera, M., Forero, D., Calor, A.R., Romero, G.Q., Riyaz, M., Callisto, M., de Oliveira Roque, F., Elme-Tumpay, A., Khan, M.K., Justino de Faria, A.P., Pires, M.M., Silva de Azevêdo, C.A., Juen, L., Zakka, U., Samaila, A.E., et al. (2024) Systematic challenges and opportunities in insect monitoring: a Global South perspective. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230102. The Royal Society.

Sandall, E.L., Maureaud, A.A., Guralnick, R., McGeoch, M.A., Sica, Y.V., Rogan, M.S., Booher, D.B., Edwards, R., Franz, N., Ingenloff, K., Lucas, M., Marsh, C.J., McGowan, J., Pinkert, S., Ranipeta, A., et al. (2023) A globally integrated structure of taxonomy to support biodiversity science and conservation. *Trends in ecology & evolution* **38**, 1143–1153.

Scarpelli, M.D.A., Roe, P., Tucker, D. & Fuller, S. (2023) Soundscape phenology: The effect of environmental and climatic factors on birds and insects in a subtropical woodland. *The Science of the total environment* **878**, 163080. Elsevier BV.

Scherer, L., Vitzthumecker, T. & Bierl, R. (2022) Capacitive PCB sensor for the detection of insects in intelligent traps. *Applied Research Conference 2022 Conference Proceedings: 4th July 2022, University of Applied Sciences Ansbach (Onsite Conference)*. Https://opus4.kobv.de/opus4-oth-regensburg/files/5583/Scherer_Capacitive_PCB_sensor_ARC_SS22-4.pdf [accessed 2 April 2025].

Schneider, S., Taylor, G.W., Kremer, S.C. & Fryxell, J.M. (2023) Getting the bugs out of AI: Advancing ecological research on arthropods through computer vision. *Ecology letters* **26**, 1247–1258.

Seimandi-Corda, G., Hood, T. & Cook, S.M. (2024) Understanding insect predator–prey interactions using camera trapping: A review of current research and perspectives. *Agricultural and forest entomology*. Wiley.

Serra-Marin, P.E., Solé-Ribalta, A., Lana, A., Borge-Holthoefer, J., Hervías-Parejo, S. & Traveset, A. (2025) Comparative assessment of automated and manual monitoring in comprehensive plant–pollinator communities. *Methods in Ecology and Evolution* **16**, 2960–2978. Wiley.

Shamoun-Baranes, J., Bauer, S., Chapman, J.W., Desmet, P., Dokter, A.M., Farnsworth, A., Haest, B., Koistinen, J., Kranstauber, B., Liechti, F., Mason, T.H.E., Nilsson, C., Nussbaumer, R., Schmid, B., Weisshaupt, N., et al. (2021) Weather radars' role in biodiversity monitoring. *Science (New York, N.Y.)* **372**, 248. American Association for the Advancement of Science (AAAS).

Sheard, J.K., Adriaens, T., Bowler, D.E., Büermann, A., Callaghan, C.T., Camprasse, E.C.M., Chowdhury, S., Engel, T., Finch, E.A., von Gönner, J., Hsing, P.-Y., Mikula, P., Rachel Oh, R.Y., Peters, B., Phartyal, S.S., et al. (2024) Emerging technologies in citizen science and potential for insect monitoring. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230106. The Royal Society.

Shi, X., Schmid, B., Tschanz, P., Segelbacher, G. & Liechti, F. (2021) Seasonal trends in movement patterns of birds and insects aloft simultaneously recorded by radar. *Remote sensing* **13**, 1839. MDPI AG.

Sittinger, M., Uhler, J., Pink, M. & Herz, A. (2024) Insect detect: An open-source DIY camera trap for automated insect monitoring. *PloS one* **19**, e0295474.

Southwood, T.R.E. (1966) *Ecological Methods with particular reference to the study of insect populations*. Chapman and Hall, London, England.

Speaker, T., O'Donnell, S., Wittemyer, G., Bruyere, B., Loucks, C., Dancer, A., Carter, M., Fegraus, E., Palmer, J., Warren, E. & Solomon, J. (2022) A global community-sourced assessment of the state of conservation technology. *Conservation Biology* **36**, e13871. John Wiley & Sons, Ltd.

Spiesman, B.J., Gratton, C., Hatfield, R.G., Hsu, W.H., Jepsen, S., McCornack, B., Patel, K. & Wang, G. (2021) Assessing the potential for deep learning and computer vision to identify bumble bee species from images. *Scientific Reports* **11**, 7580. Springer Science and Business Media LLC.

Steen, R. (2017) Diel activity, frequency and visit duration of pollinators in focal plants: in situ automatic camera monitoring and data processing. *Methods in Ecology and Evolution* **8**, 203–213. John Wiley & Sons, Ltd.

Ștefan, V., Stark, T., Wurm, M., Taubenböck, H. & Knight, T.M. (2025a) Successes and limitations of pretrained YOLO detectors applied to unseen time-lapse images for automated pollinator monitoring. *Scientific reports* **15**, 30671.

Ștefan, V., Workman, A., Cobain, J.C., Rakosy, D. & Knight, T.M. (2025b) Utilising affordable smartphones and open-source time-lapse photography for pollinator image collection and annotation. *Journal of pollination ecology* **37**, 1–21. International Commission for Plant Pollinator Relations.

Stepanian, P.M., Horton, K.G., Melnikov, V.M., Zrnić, D.S. & Gauthreaux, S.A., Jr (2016) Dual-polarization radar products for biological applications. *Ecosphere (Washington, D.C)* **7**, e01539. Wiley.

Stevens, S., Wu, J., Thompson, M.J., Campolongo, E.G., Song, C.H., Carlyn, D.E., Dong, L., Dahdul, W.M., Stewart, C., Berger-Wolf, T., Chao, W.-L. & Su, Y. (2023) BioCLIP: A vision foundation model for the tree of life. *arXiv [cs.CV]*. Http://arxiv.org/abs/2311.18803.

Stork, N.E. (2018) How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth? *Annual review of entomology* **63**, 31–45.

Stothut, M., Kühne, D., Ströbele, V., Mahla, L., Künzel, S. & Krehenwinkel, H. (2024) Environmental DNA metabarcoding reliably recovers arthropod interactions which are frequently observed by video recordings of flowers. *Environmental DNA (Hoboken, N.J.)* **6**. Wiley.

van Strien, A.J., van Swaay, C.A.M. & Termaat, T. (2013) Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *The journal of applied ecology* **50**, 1450–1458. Wiley.

Šturm, R., López Díez, J.J., Polajnar, J., Sueur, J. & Virant-Doberlet, M. (2022) Is it time for ecotremology? *Frontiers in ecology and evolution* **10**. Frontiers Media SA.

Sugai, L.S.M., Silva, T.S.F., Ribeiro, J.W., Jr & Llusia, D. (2019) Terrestrial passive acoustic monitoring: Review and perspectives. *Bioscience* **69**, 15–25. Oxford University Press (OUP).

Su, J., Vargas, D.V. & Sakurai, K. (2019) Attacking convolutional neural network using differential evolution. *IPSJ transactions on computer vision and applications* **11**. Springer Science and Business Media LLC.

Sun, Z., Hu, C., Cui, K., Wang, R., Ding, M., Yan, Z. & Wu, D. (2024) Extracting bird and insect migration echoes from single-polarization weather radar data using semi-supervised learning. *IEEE transactions on geoscience and remote sensing: a publication of the IEEE Geoscience and Remote Sensing Society* **62**, 1–12. Institute of Electrical and Electronics Engineers (IEEE).

Svenning, A., Mougeot, G., Alison, J., Chevalier, D., Molina, N.L.C., Ong, S.-Q., Bjerge, K., Carrillo, J., Hoeye, T.T. & Geissmann, Q. (2025) A General Method for Detection and Segmentation of Terrestrial Arthropods in Images. bioRxiv. *Ecology*. Https://www.biorxiv.org/content/10.1101/2025.04.08.647223v1 [accessed 16 June 2025].

Symes, L.B., Madhusudhana, S., Martinson, S.J., Geipel, I. & Ter Hofstede, H.M. (2024) Multi-year soundscape recordings and automated call detection reveals varied impact of moonlight on calling activity of neotropical forest katydids. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230110. The Royal Society.

Szczygieł, H.A., Dent, D.H. & Quitmeyer, A. (2025) Mothbox: inexpensive, lightweight, automated light trap for scalable insect biodiversity monitoring. bioRxiv. *bioRxiv*. Http://dx.doi.org/10.64898/2025.12.03.692171.

Taffel, S., Bedford, L. & Mann, M. (2022) Ecocide isn't ethical: Political ecology and capitalist AI ethics. In *Economies of Virtue: The Circulation of 'Ethics' in AI* (eds T. Phan, J. Goldenfein, D. Kutch & M. Mann), pp. 58–82. Institute of Network Cultures, Amsterdam, The Netherlands.

Tang, T., Long, Y., Li, Y. & Liang, J. (2022) Acoustic domain mismatch compensation in bird audio detection. *International journal of speech technology* **25**, 251–260. Springer Science and Business Media LLC.

Terry, J.C.D., Roy, H.E. & August, T.A. (2020) Thinking like a naturalist: Enhancing computer vision of citizen science images by harnessing contextual data. *Methods in Ecology and Evolution* **11**, 303–315. Wiley.

Tetzlaff, S.J., Katz, A.D., Wolff, P.J. & Kleitch, M.E. (2024) Comparison of soil eDNA to camera traps for assessing mammal and bird community composition and site use. *Ecology and evolution* **14**, e70022. Wiley.

Thomas, C.D., Hefin Jones, T. & Hartley, S.E. (2019) 'Insectageddon': A call for more robust data and rigorous analyses. *Global Change Biology* **25**, 1891–1892. John Wiley & Sons, Ltd.

Thomsen, P.F. & Sigsgaard, E.E. (2019) Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. *Ecology and evolution* **9**, 1665–1679. Wiley.

Tielens, E.K. & Kelly, J. (2024) Temperature, not net primary productivity, drives continental-scale variation in insect flight activity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230114. The Royal Society.

Truong, T.-D., Nguyen, H.-Q., Nguyen, X.-B., Dowling, A., Li, X. & Luu, K. (2025) Insect-Foundation: A Foundation Model and Large Multimodal Dataset for Vision-Language Insect Understanding. Http://arxiv.org/abs/2502.09906 [accessed 29 October 2025].

Unger, S., Rollins, M., Tietz, A. & Dumais, H. (2021) iNaturalist as an engaging tool for identifying organisms in outdoor activities. *Journal of biological education* **55**, 537–547. Informa UK Limited.

United Nations (1992) Convention on Biological Diversity. Https://treaties.un.org/doc/Treaties/1992/06/19920605%2008-44%20PM/Ch_XXVII_08p.pdf.

United Nations (2015) Transforming our world: 2030 Agenda Sustainable Development. Https://digitallibrary.un.org/record/3923923?ln=en&v=pdf.

Van Doren, B.M., Lostanlen, V., Cramer, A., Salamon, J., Dokter, A., Kelling, S., Bello, J.P. & Farnsworth, A. (2023) Automated acoustic monitoring captures timing and intensity of bird migration. *The journal of applied ecology* **60**, 433–444. Wiley.

Varga-Szilay, Z., Szövényi, G. & Pozsgai, G. (2024) Flower visitation through the lens: Exploring the foraging behaviour of Bombus terrestris with a computer vision-based application. *Insects* **15**.

Wägele, J.W., Bodesheim, P., Bourlat, S.J., Denzler, J., Diepenbroek, M., Fonseca, V., Frommolt, K.-H., Geiger, M.F., Gemeinholzer, B., Glöckner, F.O., Haucke, T., Kirse, A., Kölpin, A., Kostadinov, I., Kühl, H.S., et al. (2022) Towards a multisensor station for automated biodiversity monitoring. *Basic and applied ecology* **59**, 105–138.

Wagner, D.L. (2020) Insect declines in the anthropocene. *Annual Review of Entomology* **65**, 457–480. Annual Reviews.

Wang, D.-B., Feng, L. & Zhang, M.-L. (2021) Rethinking calibration of deep neural networks: Do not be afraid of overconfidence. *Neural Information Processing Systems*, 11809–11820.

Werber, Y., Adin, E., Chapman, J.W., Reynolds, D.R. & Sapir, N. (2025) Active navigation and meteorological selectivity drive insect migration patterns through the Levant. *Proceedings. Biological sciences* **292**, 20250587.

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T. & Vieglais, D. (2012) Darwin Core: an evolving community-developed biodiversity data standard. *PloS one* **7**, e29715. Public Library of Science (PLoS).

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, Blomberg, N., Boiten, J.W., Da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, J., Clark, T., Crosas, M. & Dillo, I. (2016) Comment: FAIR Guiding Principles scientific data management stewardship. *Scientific Data* **3**, 1–9.

Willi, M., Pitman, R.T., Cardoso, A.W., Locke, C., Swanson, A., Boyer, A., Veldthuis, M. & Fortson, L. (2019) Identifying animal species in camera trap images using deep learning and citizen science. *Methods in ecology and evolution* **10**, 80–91. Wiley.

Wilson, R.J., de Siqueira, A.F., Brooks, S.J., Price, B.W., Simon, L.M., van der Walt, S.J. & Fenberg, P.B. (2023) Applying computer vision to digitised natural history collections for climate change research: Temperature-size responses in British butterflies. *Methods in* *ecology and evolution* **14**, 372–384. Wiley.

Wittmann, K., Gamal Ibrahim, M., Straw, A.D., Klein, A.-M. & Staab, M. (2024) Monitoring fast-moving animals—Building a customized camera system and evaluation toolset. *Methods in ecology and evolution* **15**, 836–842. Wiley.

Wood, C.M. & Kahl, S. (2024) Guidelines for appropriate use of BirdNET scores and other detector outputs. *Journal of ornithology* **165**, 777–782. Springer Science and Business Media LLC.

Woodruff, J., Schall, D., Oboyle, M.F.P. & Woodruff, C. (2023) *When Does Saving Power Save Planet? Proceedings 2nd Workshop Sustainable Computer Systems*. In pp. 1–6. Association Computing Machinery, New York, NY, USA.

Wotton, K.R., Gao, B., Menz, M.H.M., Morris, R.K.A., Ball, S.G., Lim, K.S., Reynolds, D.R., Hu, G. & Chapman, J.W. (2019) Mass Seasonal Migrations of Hoverflies Provide Extensive Pollination and Crop Protection Services. *Current biology: CB* **29**, 2167–2173.e5. Elsevier Ltd.

Wu, X., Zhan, C., Lai, Y.-K., Cheng, M.-M. & Yang, J. (2019) IP102: A large-scale benchmark dataset for insect pest recognition. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

Yamoa, A.S.-D., Kouakou, B.K., Gbogbo, A.Y., Runemark, A., van Klink, R., Zoueu, J.T. & Brydegaard, M. (2025) Lidar reveals distinct insect daily activity and diversity between habitats. *Scientific Reports* **15**, 43544. Springer Science and Business Media LLC.

Yen, J.D.L., Tonkin, Z., Lyon, J., Koster, W., Kitchingman, A., Stamation, K. & Vesk, P.A. (2019) Integrating multiple data types to connect ecological theory and data among levels. *Frontiers in ecology and evolution* **7**. Frontiers Media SA.

Yoder, M.J., Mikó, I., Seltmann, K.C., Bertone, M.A. & Deans, A.R. (2010) A gross anatomy ontology for hymenoptera. *PloS one* **5**, e15991.

Zaugg, S., Saporta, G., van Loon, E., Schmaljohann, H. & Liechti, F. (2008) Automatic identification of bird targets with radar via patterns produced by wing flapping. *Journal of the Royal Society, Interface* **5**, 1041–1053. The Royal Society.

Zech, J.R., Badgeley, M.A., Liu, M., Costa, A.B., Titano, J.J. & Oermann, E.K. (2018) Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS medicine* **15**, e1002683. Public Library of Science (PLoS).

Zefa, E., DE Pinho Martins, L., Demari, C.P., Acosta, R.C., Centeno, E., Castro-Souza, R.A., DE Oliveira, G.L., Miyoshi, A.R., Fianco, M., Redü, D.R., Timm, V.F., Costa, M.K.M.D.A. & Szinwelski, N. (2022) Singing crickets from Brazil (Orthoptera: Gryllidea), an illustrated checklist with access to the sounds produced. *Zootaxa* **5209**, 211–237. Magnolia Press.

Zhang, K., Zhou, R., Adhikarla, E., Yan, Z., Liu, Y., Yu, J., Liu, Z., Chen, X., Davison, B.D., Ren, H., Huang, J., Chen, C., Zhou, Y., Fu, S., Liu, W., et al. (2024) A generalist vision-language foundation model for diverse biomedical tasks. *Nature medicine* **30**, 3129–3141.