

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

BIOLOGICAL SCIENCES

Delivering on a promise: Futureproofing automated insect monitoring methods

Journal:	<i>Philosophical Transactions B</i>
Manuscript ID	RSTB-2023-0105.R1
Article Type:	Opinion piece
Date Submitted by the Author:	n/a
Complete List of Authors:	van Klink, Roel; German Centre for Integrative Biodiversity Research iDiv, Biodiversity Synthesis; Martin-Luther-Universität Halle-Wittenberg
Issue Code (this should have already been entered and appear below the blue box, but please contact the Editorial Office if it is not present):	INSECTS
Subject:	Ecology < BIOLOGY
Keywords:	monitoring, insects, DNA-barcoding, computer vision, bioacoustics

SCHOLARONE™
Manuscripts

1
2
3 **Author-supplied statements**
4

5 Relevant information will appear here if provided.
6

7
8 **Ethics**
9

10 *Does your article include research that required ethical approval or permits?:*

11 This article does not present research with ethical considerations
12

13 *Statement (if applicable):*

14 CUST_IF_YES_ETHICS :No data available.
15

16
17 **Data**
18

19 *It is a condition of publication that data, code and materials supporting your paper are made publicly*
20 *available. Does your paper present new data?:*

21 My paper has no data
22

23 *Statement (if applicable):*

24 CUST_IF_YES_DATA :No data available.
25

26
27 **Conflict of interest**
28

29 I/We declare we have no competing interests
30

31 *Statement (if applicable):*

32 CUST_STATE_CONFLICT :No data available.
33

34
35 **Use of AI**
36

37 *Please provide a statement of any use of AI technology in the preparation of the paper.*
38

39 No, we have not used AI-assisted technologies in creating this article

40 CUST_IF_YES_DECLARATION_OF_AI_USE :No data available.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Delivering on a promise: Futureproofing automated insect monitoring methods

Roel van Klink^{1,2,3*}

¹German Centre for Integrative Biodiversity Research Halle-Jena-Leipzig, Puschstrasse 4, 04103 Leipzig

²Department of Computer Science, Martin-Luther-University Halle-Wittenberg, Germany

³WBBS Foundation, Kanaaldijk 36, 9409 TV, Loon, The Netherlands.

* Corresponding author: roel.klink@idiv.de

ORCID: 0000-0002-8125-1463

Abstract

Due to rapid technological innovations, the automated monitoring of insect assemblages comes within reach. However, this continuous innovation endangers the methodological continuity needed for calculating reliable biodiversity trends in the future.

Maintaining methodological continuity over prolonged periods of time is not trivial, since technology improves, reference libraries grow, and both the hard- and software used now may no longer be available in the future. Moreover, because data on many species are collected at the same time, there will be no simple way of calibrating the outputs of old and new devices.

To ensure that reliable long-term biodiversity trends can be calculated using the collected data, I make four recommendations: (1) Construct devices to last decades, and have a five-year overlap period when devices are replaced. (2) Construct new devices to resemble the old ones, especially when some kind of attractant (e.g. light) is used. Keep extremely detailed metadata on collection, detection and identification methods, including attractants, to enable this. (3) Store the raw data (sounds, images, DNA extracts, radar/lidar detections) for future reprocessing with updated classification systems. (4) Enable forward and backward compatibility of the processed data, for example by in-silico data 'degradation' to match the older data quality.

Key words:

DNA barcoding, bioacoustics, computer vision, radar, lidar, monitoring, insects, arthropods, LTER

Main Text

The development of technological approaches for insect monitoring can allow unprecedented improvements in the spatial, temporal and taxonomic coverage of insect biodiversity assessments [1–4]. To meet the political, societal and industry needs for large-scale biomonitoring [5–7], these technologies can help close an important knowledge gap, since insects and other arthropods are the most species rich group of animals on earth, and perform important ecosystem services (e.g. crop pollination or decomposition) and disservices (e.g. disease transmission or crop damage). Insects are notoriously underrepresented in biodiversity monitoring schemes, since monitoring their diversity by traditional means with morphological identification is extremely time consuming and knowledge intensive. Moreover, some of the largest insect groups, such as flies and parasitoid wasps, are even within insect monitoring programs and ecological assessments rarely assessed. Automated monitoring could thus make large-scale insect biodiversity monitoring possible for a fraction of the effort and costs of traditional monitoring methods, and contribute to solving a number of identified challenges to large scale biomonitoring [6].

However, in order to reliably document changes in species occurrences, population sizes and biodiversity metrics over time, it is important to use the exact same method of monitoring over the whole sampling period. This applies to the collection, detection and identification methods, including any attractants used, as well as the taxonomic precision of the end product provided. This sounds logical, and even trivial, but anyone who has tried to do a sustained monitoring of biodiversity has learned that maintaining methodological continuity is not as easy as it sounds. Even when funding for continuous monitoring is secured (which is challenging even in the richest of countries), traps need to be replaced due to wear, loss or breakdown, workers learn to identify new species, fall ill or make mistakes, and taxonomy changes over time. In addition, there is a constant need for specialists with the right expertise, which is unfeasible in most parts of the world and for most taxa. For this reason, consideration of the methodology and data quality needed, is best done before monitoring commences.

Particularly when using high-tech devices and computer algorithms, the challenges to ensuring methodological continuity compound:

- i. The hardware and software used in these devices are rapidly evolving and improving: camera sensitivity improves, barcoding pipelines change (Iwaszkiewicz-Eggebrecht this issue [8]), energy use becomes more efficient, etc. Although it is almost a moral imperative to use these developments to our advantage, and monitor as many species as possible for the lowest costs, we must also recognize the consequences of these developments for the long-term trends we're trying to calculate.
- ii. Since the devices, which are often custom made for the purpose of insect monitoring, depend on hard- and software produced by third parties, there is no guarantee that these exact components will be available in the future. In fact, it is likely that they will not, because, industrial suppliers have no incentive to produce obsolete products, supply chains change, or new legislation may prevent the continued production or import of specific components.
- iii. Weathering and wear of (parts of) the devices and traps in the field may make repeated use challenging, and parts may need to be exchanged regularly [see for example 9].
- iv. The reference libraries of DNA barcodes, images and sounds used for classification are constantly growing, and will contain more and more species, allowing more accurate classification.

1
2
3 78 v. These devices are designed to collect multivariate data (dozens to thousands of species
4 79 at the same time), and therefore, there will be no simple calibration possible of
5 80 measured variables when monitoring devices are replaced with newer versions,
6 81 especially given the volatility of insect population dynamics and the prevalence of rare
7 82 species [10].
8
9

10 83

11
12 84 In most cases, technological improvements will increase detection and/or identification rates, which,
13 85 when left unaccounted for, will lead to detecting a false increase in diversity over time. But any
14 86 change in detection rates of any species will affect the inferences one can draw from the monitoring
15 87 program in the future. The technologies covered in this Theme Issue (computer vision, DNA
16 88 (meta)barcoding, radar and acoustics) are still in development, and are thus particularly vulnerable
17 89 to the challenges outlined above. Although statistical methods may be able to account for some
18 90 aspects of methodological variability, the reliability of the calculated temporal trends will suffer
19 91 significantly from rapid methodological changes, in comparison to a continuous methodology..
20
21

22 92 I will illustrate the difficulties of ensuring methodological continuity over prolonged periods of time
23 93 by two examples that are orders of magnitude less complex than any of the technologies discussed
24 94 in this Theme Issue: Pitfall trapping of ground beetles (Coleoptera: Carabidae) with morphological
25 95 species identification. In the north of the Netherlands, a program for monitoring ground beetle
26 96 populations by means of standardized, year-round pitfall trapping was started in 1959 by the
27 97 workers of the Willem Beijerink Biological Station, part of what is now Wageningen University. They
28 98 started trapping ground beetles in custom-made square metal cans with an exact perimeter of 1m,
29 99 [11,12]. These traps were replaced in the 1980's and possibly at an earlier time as well, but
30 100 unfortunately this was not well documented. After the biological station was formally dissolved in
31 101 1998, the trapping program was continued by the volunteers of the WBBS foundation using the cans
32 102 constructed in the 1980's. By 2020, the traps were in need of replacement, and we acquired funding
33 103 for the construction of new traps.
34
35
36

37 104 Although we were unable to find back the company that constructed the original traps, this looked
38 105 like a straightforward construction job to us, which any metalworking company could do. However,
39 106 after numerous emails, phone calls and visits to various companies, we found that the technique for
40 107 constructing the rounded edge of the old cans (Fig. 1a), a process called 'edge beading', had fallen
41 108 out of use for this kind of sheet metal, and that a custom-made mold (a 'die') for a bead of exactly
42 109 this size would be excessively expensive (roughly half of our budget for replacing the traps). We
43 110 therefore had to settle for a different edge type for our new traps (Fig. 1b). We hope that, at least
44 111 from a beetle's perspective, there will be no difference between the trap types (Fig. 1c). We have
45 112 replaced the traps in two phases over 2022 and 2023 to test if and how the catch is affected by the
46 113 trap replacement.
47
48
49

50 114 A second example from the same monitoring program is the challenge we have faced regarding the
51 115 transition between data formats. All data collected on a weekly basis from 1959 to 1998 were once
52 116 digitized, and stored on computer tapes. Currently, reading such tapes is close to impossible,
53 117 especially since we don't know which computer brand was used for data entry, or the software
54 118 format the data were stored in. Fortunately, all data are still available on paper sheets, and we are
55 119 currently working on redigitising these, where we ensure compatibility with the upcoming Humboldt
56 120 Extension for ecological inventories to the GBIF Darwin Core. That this is necessary illustrates the
57 121 importance of a timely transition between data formats as hard- and software evolve. In 2009, Borer
58 122 et al. [13] published some excellent advice on data management, and wrote: 'As hard as it is to

1
2
3 123 believe today, we can foresee the day when CD-ROMs might be difficult to read.'. As per 2023, that
4 124 day has come and gone, and it would be well advised to rapidly move all data stored on CD-roms and
5 125 DVD's to the cloud (or better, to make them openly accessible on a FAIR biodiversity data portal like
6 126 GBIF). This trend of soft- and hardware replacement is likely to continue, and it will be important to
7 127 keep up with these developments.

8
9
10 128 Now imagine going through a similar process for replacing a modern camera trap, a radar, a
11 129 sequencer or a barcoding pipeline, or to try to read data 20 years from now. Ideally, we would want
12 130 every single hard- and software component used for detecting and identifying organisms, and for
13 131 data storage to remain constant for as long as the monitoring lasts: several decades. But this is
14 132 exceedingly unlikely, since all technological insect monitoring methods depend on a chain of
15 133 industrial suppliers for the hard- and software used in the devices, as well as for data storage. These
16 134 suppliers have no interest in continuing the production of obsolete products, just as we, as end users
17 135 should use the best products available to monitor as many species as possible. Hence, we will need
18 136 other solutions to ensure methodological continuity.

19
20
21 137 Below, I make four concrete recommendations, from the level of device construction to the
22 138 processed biodiversity data, to ensure the data produced now can be used to calculate reliable
23 139 biodiversity trends in the future. These recommendations are in most cases not only applicable to
24 140 new technologies, but are equally useful for traditional insect monitoring programs:

- 25
26
27 141 a) Build to last. Design devices with the aim of lasting decades, and don't wait for them to
28 142 break down before replacing them. Ideally, aim for an overlap period of 5 years when
29 143 replacing devices, but here it should be considered that two traps set up in close proximity
30 144 may influence each other, especially when an attractant is used. In such cases, a phased
31 145 transition across multiple locations may be a better option.
- 32 146 b) Keep extremely detailed metadata, so that future devices can collect data in the same way,
33 147 even when the sensors improve. This is especially important when an attractant, such as
34 148 light or a colored screen is used, because a change in attractant(s) will inevitably affect
35 149 insect behavior. But also extreme metadata detail is required regarding the sensitivity of the
36 150 sensor(s), as this information can be used to make collected data more comparable.
37 151 Metadata should thus include the exact light spectrum (including parts of the light spectrum
38 152 that are not visible for humans, and luminosity of a light trap, exact screen color and texture
39 153 [see 14 this issue], motion triggers (if used), camera resolution, microphone sensitivity,
40 154 frequency range, and recording bitrate, sequencing depth, biochemical and bioinformatics
41 155 pipelines for (meta)barcoding [see 8 this issue], etc. In addition, all data on the operational
42 156 status of the traps and/or sensors, as well as the exact locations, should be recorded and
43 157 stored. Although a lack of historic metadata may prevent us from precisely redoing historical
44 158 investigations, we can make future resampling campaigns a lot more accurate.
- 45 159 c) Store all raw data (photos, condensed audio recordings, radar/lidar detections, barcoding
46 160 libraries, etc.) in a non-proprietary format for future reprocessing using new algorithms,
47 161 computational facilities and reference libraries. For this, a data infrastructure is needed that
48 162 can handle and process the expected volume of raw data, and that can ensure data
49 163 accessibility in the future. In addition, the energy, and thus environmental, costs of data
50 164 storage and reprocessing should be considered.
- 51 165 d) Ensure forward and/or backward compatibility of the processed data (data with assigned
52 166 taxonomic names), so that the quality of the data collected in the future can be made
53 167 comparable to the data collected now, regarding, for example, the taxonomic depth and the
54 168 sensor sensitivity. This may be done by either bringing currently collected data up to

standards of the future (which will possibly need reprocessing, see previous point), or by in-silico degradation of future data to match the current standards (assuming that future data will be of higher quality than current data). To make this possible, there is a strong need for the automated taxonomic harmonization of species identifications. The GBIF taxonomic backbone, which is based the Catalogue of Life [15], the Barcode Index Numbers from the Barcode of Life project [16], and 103 other taxonomic resources [17], seems the most promising resource for automated harmonization with the most up-to-date taxonomic classification for both traditional and genetic data.

These recommendations do not only apply to the monitoring of insects, but to any type of automated biodiversity monitoring, for example camera trapping of mammals, acoustic monitoring of birds, bats, whales or fish, eDNA, or bird radar.

Conclusions

If the difficulties of securing long-term funding for biodiversity monitoring and the continued training of taxonomic specialists can be overcome, the technological developments of the past decades bring large-scale insect monitoring is closer than ever. But before we start deploying devices whenever an opportunity arises, it will pay off to first consider how we want to use these data now and in the future. What we can learn and infer, and for whom and for what purpose the data will be useful, will crucially depend on the choices we make today. For many purposes, including conservation planning and pest monitoring, accurate species level identifications are of crucial importance. Likewise, for calculating long-term trends, methodological continuity is crucial. If the above recommendations are followed, I am confident that automated insect monitoring will yield us many insights about the changes in insect biodiversity over the coming decades.

Acknowledgements:

This work was funded by DFG Grant FZT 118 to the German Centre for Integrative Biodiversity Research. The new ground beetle traps were funded by the Uyttenboogaart-Eliassen Stichting and the Prins-Bernhard Cultuurfonds. I thank Fons van der Plas and two reviewers for commenting on an earlier version of this manuscript, and Rikjan Vermeulen, Henk de Vries, Alje Woldering and Kees van der Laaken[†] for their dedication to sampling and identifying the beetles. I thank Henk de Vries and Alje Woldering also for the photos used in Fig 1.

References

1. Van Klink R *et al.* 2022 Emerging technologies revolutionise insect ecology and monitoring. *Trends Ecol. Evol.* **37**, 872–885. (doi:10.1016/j.tree.2022.06.001)
2. Gibb R, Browning E, Glover-Kapfer P, Jones KE. 2019 Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods Ecol. Evol.* **10**, 169–185. (doi:10.1111/2041-210X.13101)
3. Høye TT *et al.* 2021 Deep learning and computer vision will transform entomology. *Proc. Natl. Acad. Sci.* **118**, 2002545117. (doi:10.1073/PNAS.2002545117)

- 1
2
3 209 4. Bauer S *et al.* 2017 From agricultural benefits to aviation safety: Realizing the potential of
4 210 continent-wide radar networks. *BioScience* **67**, 912–918. (doi:10.1093/biosci/bix074)
5
6 211 5. Gonzalez A *et al.* 2023 A global biodiversity observing system to unite monitoring and guide
7 212 action. *Nat. Ecol. Evol.* , 1–5. (doi:10.1038/s41559-023-02171-0)
8
9 213 6. Moersberger H *et al.* 2022 Europa Biodiversity Observation Network: User and Policy Needs
10 214 Assessment. , 218.
11
12
13 215 7. Potts SG *et al.* 2020 Proposal for an EU Pollinator Monitoring Scheme.
14
15 216 8. Iwazskiewicz-Eggebrecht E, Zizka VMA, Lynggaard C. In press. placeholder. *Philos. Trans. R. Soc.*
16 217 *B-Biol. Sci.*
17
18 218 9. Ssymank A, Sorg M, Doczkal D, Rulik B, Merkel-Walner G, Vischer-Leopold M. 2018 Praktische
19 219 Hinweise und Empfehlungen zur Anwendung von Malaisefallen für Insekten in der
20 220 Biodiversitätserfassung und im Monitoring. *Ser. Nat.* **1**.
21
22
23 221 10. Fisher ARA, Corbet AS, Williams CB. 1943 The number of animals in a random sample of an
24 222 animal population. *J. Anim. Ecol.* **12**, 42–58.
25
26 223 11. Den Boer PJ. 1990 Density limits and survival of local populations in 64 carabid species with
27 224 different powers of dispersal. *J. Evol. Biol.* **3**, 19–48. (doi:10.1046/j.1420-9101.1990.3010019.x)
28
29 225 12. Den Boer PJ. 1977 *Dispersal power and survival - Carabids in a cultivated countryside.*
30 226 Wageningen: Veenman en zonen B.V. See [http://www.biological-station.com/Publicaties/Piet](http://www.biological-station.com/Publicaties/Piet%20den%20Boer2.html)
31 227 [den Boer2.html](http://www.biological-station.com/Publicaties/Piet%20den%20Boer2.html).
32
33 228 13. Borer ET, Seabloom EW, Jones MB, Schildhauer M. 2009 Some Simple Guidelines for Effective
34 229 Data Management. *Bull. Ecol. Soc. Am.* **90**, 205–214. (doi:10.1890/0012-9623-90.2.205)
35
36 230 14. Roy DB, Hoyer TT. In press. placeholder. *Philos. Trans. R. Soc. B-Biol. Sci.*
37
38
39 231 15. Bánki O *et al.* 2023 *Catalogue of Life Checklist*.
40
41 232 16. Ratnasingham S, Hebert PDN. 2007 bold: The Barcode of Life Data System
42 233 (<http://www.barcodinglife.org>). *Mol. Ecol. Notes* **7**, 355–364. (doi:10.1111/j.1471-
43 234 8286.2007.01678.x)
44
45 235 17. GBIF Secretariat. 2023 GBIF Backbone Taxonomy. Checklist dataset.
46
47 236
48
49 237
50
51
52
53
54
55
56
57
58
59
60

238 FIGURE CAPTIONS

239 **Fig. 1.** *The edges of the old (a) and the new (b) ground beetle traps. Due to technological changes,*
240 *the old, rounded, edge would be excessively hard to reproduce. We have aimed to make the edge as*
241 *similar as possible under field conditions (c). Photo's: Henk de Vries (a), Alje Woldering (b & c).*

242

For Review Only



Fig. 1. The edges of the old (a) and the new (b) ground beetle traps. Due to technological changes, the old, rounded, edge would be excessively hard to reproduce. We have aimed to make the edge as similar as possible under field conditions (c). Photo's: Henk de Vries (a), Alje Woldering (b & c).

682x181mm (300 x 300 DPI)

Delivering on a promise: Futureproofing automated insect monitoring methods

Roel van Klink^{1,2,3*}

¹German Centre for Integrative Biodiversity Research Halle-Jena-Leipzig, Puschstrasse 4, 04103 Leipzig

²Department of Computer Science, Martin-Luther-University Halle-Wittenberg, Germany

³WBBS Foundation, Kanaaldijk 36, 9409 TV, Loon, The Netherlands.

* Corresponding author: roel.klink@idiv.de

ORCID: 0000-0002-8125-1463

Abstract

Due to rapid technological innovations, the automated monitoring of insect assemblages comes within reach. However, this continuous innovation endangers the methodological continuity needed for calculating reliable biodiversity trends in the future.

Maintaining methodological continuity over prolonged periods of time is not trivial, since technology improves, reference libraries grow, and both the hard- and software used now may no longer be available in the future. Moreover, because data on many species are collected at the same time, there will be no simple way of calibrating the outputs of old and new devices.

To ensure that reliable long-term biodiversity trends can be calculated using the collected data, I make four recommendations: (1) Construct devices to last decades, and have a five-year overlap period when devices are replaced. (2) Construct new devices to resemble the old ones, especially when some kind of attractant (e.g. light) is used. Keep extremely detailed metadata on collection, detection and identification methods, including attractants, to enable this. (3) Store the raw data (sounds, images, DNA extracts, radar/lidar detections) for future reprocessing with updated classification systems. (4) Enable forward and backward compatibility of the processed data, for example by in-silico data 'degradation' to match the older data quality.

Key words:

DNA barcoding, bioacoustics, computer vision, radar, lidar, monitoring, insects, arthropods, LTER

33 Main Text

34 The development of technological approaches for insect monitoring can allow unprecedented
35 improvements in the spatial, temporal and taxonomic coverage of insect biodiversity assessments
36 [1–4]. To meet the political, societal and industry needs for large-scale biomonitoring [5–7], these
37 technologies can help ~~close~~eing an important knowledge gap, since insects and other arthropods are
38 the most species rich group of animals on earth, and perform important ecosystem services (e.g.
39 crop pollination or decomposition) and disservices (e.g. disease transmission or crop damage).
40 Insects are notoriously underrepresented in biodiversity monitoring schemes, since monitoring their
41 diversity by traditional means with morphological identification is extremely time consuming and
42 knowledge intensive. Moreover, some of the largest insect groups, such as flies and parasitoid
43 wasps, are even within insect monitoring programs and ecological assessments
44 underrepresented rarely assessed. Automated monitoring could thus make ~~a~~ large-scale insect
45 biodiversity monitoring ~~program~~ possible for a fraction of the effort and costs of traditional
46 monitoring methods, and contribute to solving a number of identified challenges to large scale
47 biomonitoring [6].

48 However, in order to reliably document changes in species occurrences, population sizes and
49 biodiversity metrics over time, it is crucial-important to use the exact same method of monitoring
50 over the whole sampling period. This applies to the collection, detection and identification methods,
51 including any attractants used, as well as the taxonomic precision of the end product provided. This
52 sounds logical, and even trivial, but anyone who has tried to do a sustained monitoring of
53 biodiversity has learned that maintaining methodological continuity is not as easy as it sounds at all
54 trivial. Even when funding for continuous monitoring is secured (which is challenging even in the
55 richest of countries), traps need to be replaced due to wear, loss or breakdown, workers learn to
56 identify new species, fall ill or make mistakes, animals or people vandalize or steal equipment, and
57 taxonomy changes over time. In addition, there is a constant need for specialists with the right
58 expertise, which is unfeasible in most parts of the world and for most taxa. For this reason,
59 consideration of the methodology and data quality needed, is best done before monitoring
60 commences.

61 Particularly when using high-tech devices and computer algorithms, the challenges to ensuring
62 methodological continuity compound:

- 63 i. The hardware and software used in these devices are rapidly evolving and improving:
64 camera ~~s~~ sensitivity improves, barcoding pipelines change (Iwaszkiewicz-Eggebrecht this
65 issue [8]), energy use becomes more efficient, etc. Although it is almost a moral
66 imperative to use these developments to our advantage, and monitor as many species
67 as possible for the lowest costs, we must also recognize the consequences of these
68 developments for the long-term trends we're trying to calculate.
- 69 ii. Since the devices, which are ~~so far~~ often custom made for the purpose of insect
70 monitoring, depend on hard- and software produced by third parties, there is no
71 guarantee that these exact components will be available in the future. In fact, it is likely
72 that they will not, as because, industrial suppliers have no incentive to produce obsolete
73 products, supply chains change, or new legislation may prevent the continued
74 production or import of specific components.
- 75 iii. Weathering and wear of (parts of) the devices and traps in the field may make repeated
76 use challenging, and parts may need to be exchanged regularly [see for example 9].

- 1
2
3 77 iv. The reference libraries of DNA barcodes, images and sounds used for classification are
4 78 constantly growing, and will contain more and more species, allowing more accurate
5 79 classification.
6
7 80 v. These devices are designed to collect multivariate data (dozens to thousands of species
8 81 at the same time), and therefore, there will be no simple calibration possible of
9 82 measured variables when monitoring devices are replaced with newer versions,
10 83 especially given the volatility of insect population dynamics and the prevalence of rare
11 84 species [10].

12 85
13
14
15 86 In most cases, technological improvements will increase detection and/or identification rates, which,
16 87 when left unaccounted for, will lead to detecting a false increase in diversity over time. But any
17 88 change in detection rates of any species will affect the inferences one can draw from the monitoring
18 89 program in the future. The technologies covered in this Theme Issue (computer vision, DNA
19 90 (meta)barcoding, radar and acoustics) are still in development, and are thus particularly vulnerable
20 91 to the challenges outlined above. Although statistical methods may be able to account for some
21 92 aspects of methodological variability, the reliability of the calculated temporal trends will suffer
22 93 significantly from rapid methodological changes, in comparison to a continuous methodology. they
23 94 can only rarely account for methodological changes over time.

24
25
26
27 95 I will illustrate the difficulties of ensuring methodological continuity over prolonged periods of time
28 96 by two examples that are orders of magnitude less complex than any of the technologies discussed
29 97 in this Theme Issue: Pitfall trapping of ground beetles (Coleoptera: Carabidae) with morphological
30 98 species identification. In the north of the Netherlands, a program for monitoring ground beetle
31 99 populations by means of standardized, year-round pitfall trapping was started in 1959 by the
32 100 workers of the Willem Beijerink Biological Station, part of what is now Wageningen University. They
33 101 started trapping ground beetles in custom-made square metal cans with an exact perimeter of 1m,
34 102 left behind by Canadian soldiers after World War II [11,12]. These traps were replaced in the 1980's
35 103 and possibly at an earlier time as well, but unfortunately this was not well documented. In 1967
36 104 these cans were replaced with custom-made traps with the exact same dimensions, which were
37 105 renewed in the 1980's. After the biological station was formally dissolved in 1998, the trapping
38 106 program was continued by the volunteers of the WBBS foundation using the cans constructed in the
39 107 1980's. By 2020, the traps were in need of replacement, and we acquired funding for the
40 108 construction of new traps.

41
42
43
44 109 Although we were unable to trace find back where the company that the constructed the original
45 110 traps were constructed, this looked like a straightforward construction job to us, which any
46 111 metalworking company could do. However, after numerous emails, phone calls and visits to various
47 112 companies, we found that the technique for constructing the rounded edge of the old cans (Fig. 1a),
48 113 a process called 'edge beading', had fallen out of use for this kind of sheet metal, and that a custom-
49 114 made mold (a 'die') for a bead of exactly this size would be excessively expensive (roughly half of our
50 115 budget for replacing the traps). We therefore had to settle for a different edge type for our new
51 116 traps (Fig. 1b). We hope that, at least from a beetle's perspective, there will be no difference
52 117 between the trap types (Fig. 1c). We have replaced the traps in two phases over 2022 and 2023 to
53 118 test if and how the catch is affected by the trap replacement.

54
55
56
57 119 A second example from the same monitoring program is the challenge we have faced regarding the
58 120 transition between data formats. All data collected on a weekly basis from 1959 to 1998 were once
59 121 digitized, and stored on computer tapes. Currently, reading such tapes is close to impossible,

1
2
3 122 especially since we don't know which computer brand was used for data entry, or the software
4 123 format the data were stored in. Fortunately, all data are still available on paper sheets, and we are
5 124 currently working on redigitising these, where we ensure compatibility with the upcoming GBIF
6 125 Humboldt Extension for ecological inventories to the GBIF Darwin Core. This-That this is necessary
7 126 illustrates the importance of a timely transition between data formats as hard- and software evolve.
8 127 In 2009, Borer et al. [13] published some very good/excellent advice on data management, and
9 128 wrote: 'As hard as it is to believe today, we can foresee the day when CD-ROMs might be difficult to
10 129 read.'. As per 2023, that day has come and gone, and it would be well advised to quickly/rapidly
11 130 move all data stored on CD-ROMs and DVD's to the cloud (or better, to make them openly accessible
12 131 on a FAIR biodiversity data portal like GBIF). This trend of soft- and hardware replacement is likely to
13 132 continue, and it will be important to keep up with these developments.

14 133 Now imagine going through a similar process for replacing a modern camera trap, a radar, a
15 134 sequencer or a barcoding pipeline, or to try to read data 20 years from now. Ideally, we would want
16 135 every single hard- and software component used for detecting and identifying organisms, and for
17 136 data storage to remain constant for as long as the monitoring lasts: several decades. But this is
18 137 exceedingly unlikely, since all technological insect monitoring methods depend on a chain of
19 138 industrial suppliers for the hard- and software used in the devices, as well as for data storage. These
20 139 suppliers have no interest in continuing the production of obsolete products, just as we, as end users
21 140 should use the best products available to monitor as many species as possible. Hence, we will need
22 141 other solutions to ensure methodological continuity.

23 142 Below, I make four concrete recommendations, from the level of device construction to the
24 143 processed biodiversity data, to ensure the data produced now can be used to calculate reliable
25 144 biodiversity trends in the future. These recommendations are in most cases not only applicable to
26 145 new technologies, but are equally useful for traditional insect monitoring programs:

- 27 146 a) Build to last. Design devices with the aim of lasting decades, and don't wait for them to
28 147 break down before replacing them. Ideally, aim for an overlap period of 5 years when
29 148 replacing devices, but here it should be considered that two traps set up in close proximity
30 149 may influence each other, especially when an attractant is used. In such cases, a phased
31 150 transition across multiple locations may be a better option.
- 32 151 b) Keep extremely detailed metadata, so that future devices can collect data in the same way,
33 152 even when the sensors improve. This is especially important when an attractant, such as
34 153 light or a colored screen is used, because a change in attractant(s) will inevitably affect
35 154 insect behavior. But also extreme metadata detail is required regarding the sensitivity of the
36 155 sensor(s), as this information can be used to make collected data more comparable.
37 156 Metadata should thus include the exact light spectrum (including parts of the light spectrum
38 157 that are not visible for humans, and luminosity of a light trap, exact screen color and texture
39 158 [see 14 this issue], motion triggers (if used), camera resolution, microphone sensitivity,
40 159 frequency range, and recording bitrate, sequencing depth, biochemical and bioinformatics
41 160 pipelines for (meta)barcoding [see 8 this issue], etc. In addition, all data on the operational
42 161 status of the traps and/or sensors, as well as the exact locations, should be recorded and
43 162 stored. Although a lack of historic metadata may prevent us from precisely redoing historical
44 163 investigations, we can make future resampling campaigns a lot more accurate.
- 45 164 c) Store all raw data (photos, condensed audio recordings, radar/lidar detections, barcoding
46 165 libraries, etc.) in a non-proprietary format for future reprocessing using new algorithms,
47 166 computational facilities and reference libraries. For this, a data infrastructure is needed that
48 167 can handle and process the expected volume of raw data, and that can ensure data

168 accessibility in the future. In addition, ~~one should also take~~ the energy, and thus
 169 environmental, costs of data storage and reprocessing ~~should be considered into account~~.
 170 d) Ensure forward and/or backward compatibility of the processed data (data with assigned
 171 taxonomic names), so that the quality of the data collected in the future can be made
 172 comparable to the data collected now, regarding, for example, the taxonomic depth and the
 173 sensor sensitivity. This may be done by either bringing currently collected data up to
 174 standards of the future (which will possibly need reprocessing, see previous point), or by in-
 175 silico degradation of future data to match the current standards (assuming that future data
 176 will be of higher quality than current data). To make this possible, there is a strong need for
 177 the automated taxonomic harmonization of species identifications. The GBIF taxonomic
 178 backbone, which is based the Catalogue of Life [15], the Barcode Index Numbers from the
 179 Barcode of Life project [16], and 103 other taxonomic resources [17], seems the most
 180 promising resource for automated harmonization with the most up-to-date taxonomic
 181 classification for both traditional and genetic data.

182 These recommendations do not only apply to the monitoring of insects, but to any type of
 183 automated biodiversity monitoring, for example camera trapping of mammals, acoustic monitoring
 184 of birds, bats, whales or fish, eDNA, or bird radar.

185 Conclusions

186 If the difficulties of securing long-term funding for biodiversity monitoring and the continued
 187 training of taxonomic specialists can be overcome, the technological developments of the past
 188 decades bring large-scale, multi-taxon insect monitoring is closer than ever, thanks to the
 189 technological developments of the past decades. But before we start deploying devices whenever
 190 ~~thean~~ opportunity arises, it will pay off to first consider how we want to use these data now and in
 191 the future. What we can learn and infer, and for whom and for what purpose the data will be useful,
 192 will crucially depend on the choices we make today. For many purposes, including conservation
 193 planning and pest monitoring, accurate species level identifications are of crucial importance.
 194 Likewise, for calculating long-term trends, methodological continuity is crucial. If the above
 195 recommendations are followed, I am confident that automated insect monitoring will yield us many
 196 insights about the changes in insect biodiversity over the coming decades.

199 Acknowledgements:

200 This work was funded by DFG Grant FZT 118 to the German Centre for Integrative Biodiversity
 201 Research. The new ground beetle traps were funded by the Uyttenboogaart-Eliassen Stichting and
 202 the Prins-Bernhard Cultuurfonds. I thank Fons van der Plas and two reviewers for commenting on an
 203 earlier version of this manuscript, and Rikjan Vermeulen, Henk de Vries, Alje Woldering and Kees van
 204 der Laaken[†] for their dedication to sampling and identifying the beetles. I thank Henk de Vries and
 205 Alje Woldering also for the photos used in Fig 1.

207 References

- 208 1. Van Klink R *et al.* 2022 Emerging technologies revolutionise insect ecology and monitoring.
 209 *Trends Ecol. Evol.* **37**, 872–885. (doi:10.1016/j.tree.2022.06.001)

- 1
2
3 210 2. Gibb R, Browning E, Glover-Kapfer P, Jones KE. 2019 Emerging opportunities and challenges for
4 211 passive acoustics in ecological assessment and monitoring. *Methods Ecol. Evol.* **10**, 169–185.
5 212 (doi:10.1111/2041-210X.13101)
6
7 213 3. Høye TT *et al.* 2021 Deep learning and computer vision will transform entomology. *Proc. Natl.*
8 214 *Acad. Sci.* **118**, 2002545117. (doi:10.1073/PNAS.2002545117)
9
10 215 4. Bauer S *et al.* 2017 From agricultural benefits to aviation safety: Realizing the potential of
11 216 continent-wide radar networks. *BioScience* **67**, 912–918. (doi:10.1093/biosci/bix074)
12
13 217 5. Gonzalez A *et al.* 2023 A global biodiversity observing system to unite monitoring and guide
14 218 action. *Nat. Ecol. Evol.* , 1–5. (doi:10.1038/s41559-023-02171-0)
15
16 219 6. Moersberger H *et al.* 2022 Europa Biodiversity Observation Network: User and Policy Needs
17 220 Assessment. , 218.
18
19 221 7. Potts SG *et al.* 2020 Proposal for an EU Pollinator Monitoring Scheme.
20
21 222 8. Iwaszkiewicz-Eggebrecht E, Zizka VMA, Lynggaard C. In press. placeholder. *Philos. Trans. R. Soc.*
22 223 *B-Biol. Sci.*
23
24 224 9. Ssymank A, Sorg M, Doczkal D, Rulik B, Merkel-Walner G, Vischer-Leopold M. 2018 Praktische
25 225 Hinweise und Empfehlungen zur Anwendung von Malaisefallen für Insekten in der
26 226 Biodiversitätserfassung und im Monitoring. *Ser. Nat.* **1**.
27
28 227 10. Fisher ARA, Corbet AS, Williams CB. 1943 The number of animals in a random sample of an
29 228 animal population. *J. Anim. Ecol.* **12**, 42–58.
30
31 229 11. Den Boer PJ. 1990 Density limits and survival of local populations in 64 carabid species with
32 230 different powers of dispersal. *J. Evol. Biol.* **3**, 19–48. (doi:10.1046/j.1420-9101.1990.3010019.x)
33
34 231 12. Den Boer PJ. 1977 *Dispersal power and survival - Carabids in a cultivated countryside.*
35 232 Wageningen: Veenman en zonen B.V. See [http://www.biological-station.com/Publicaties Piet](http://www.biological-station.com/Publicaties/Piet%20den%20Boer2.html)
36 233 [den Boer2.html](http://www.biological-station.com/Publicaties/Piet%20den%20Boer2.html).
37
38 234 13. Borer ET, Seabloom EW, Jones MB, Schildhauer M. 2009 Some Simple Guidelines for Effective
39 235 Data Management. *Bull. Ecol. Soc. Am.* **90**, 205–214. (doi:10.1890/0012-9623-90.2.205)
40
41 236 14. Roy DB, Høye TT. In press. placeholder. *Philos. Trans. R. Soc. B-Biol. Sci.*
42
43 237 15. Bánki O *et al.* 2023 *Catalogue of Life Checklist.*
44
45 238 16. Ratnasingham S, Hebert PDN. 2007 bold: The Barcode of Life Data System
46 239 (<http://www.barcodinglife.org>). *Mol. Ecol. Notes* **7**, 355–364. (doi:10.1111/j.1471-
47 240 8286.2007.01678.x)
48
49 241 17. GBIF Secretariat. 2023 GBIF Backbone Taxonomy. Checklist dataset.
50
51
52
53
54
55 242
56
57 243
58
59
60

244 FIGURE CAPTIONS

245 ***Fig. 1.*** *The edges of the old (a) and the new (b) ground beetle traps. Due to technological changes,*
246 *the old, rounded, edge would be excessively hard to reproduce. We have aimed to make the edge as*
247 *similar as possible under field conditions (c). Photo's: Henk de Vries (a), Alje Woldering (b & c).*

248

For Review Only