

PERSPECTIVE

Low availability of code in ecology: A call for urgent action

Antica Culina^{1*}, Ilona van den Berg^{1,2}, Simon Evans^{3,4}, Alfredo Sánchez-Tójar^{5*}

1 Department of Animal Ecology, Netherlands Institute of Ecology, NIOO-KNAW, Wageningen, the Netherlands, **2** Vrije Universiteit Amsterdam, Amsterdam, the Netherlands, **3** Centre for Ecology and Conservation, University of Exeter, Cornwall Campus, Penryn, United Kingdom, **4** Department of Zoology, University of Oxford, Oxford, United Kingdom, **5** Department of Evolutionary Biology, Bielefeld University, Bielefeld, Germany

* A.Culina@nioo.knaw.nl (AC); alfredo.tojar@gmail.com (AST)



Abstract

Access to analytical code is essential for transparent and reproducible research. We review the state of code availability in ecology using a random sample of 346 nonmolecular articles published between 2015 and 2019 under mandatory or encouraged code-sharing policies. Our results call for urgent action to increase code availability: only 27% of eligible articles were accompanied by code. In contrast, data were available for 79% of eligible articles, highlighting that code availability is an important limiting factor for computational reproducibility in ecology. Although the percentage of ecological journals with mandatory or encouraged code-sharing policies has increased considerably, from 15% in 2015 to 75% in 2020, our results show that code-sharing policies are not adhered to by most authors. We hope these results will encourage journals, institutions, funding agencies, and researchers to address this alarming situation.

OPEN ACCESS

Citation: Culina A, van den Berg I, Evans S, Sánchez-Tójar A (2020) Low availability of code in ecology: A call for urgent action. *PLoS Biol* 18(7): e3000763. <https://doi.org/10.1371/journal.pbio.3000763>

Published: July 28, 2020

Copyright: © 2020 Culina et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: AS-T was supported by German Research Foundation (DFG) as part of the SFB TRR 212 (NC³), project numbers 316099922 and 396782608. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: doi, digital object identifier; FAIR, Findable, Accessible, Interoperable, and Reusable; GUI, graphical user interface; TOP Guidelines, Transparency and Openness Promotion Guidelines.

Provenance: Not commissioned; externally peer reviewed.

Introduction

There is a growing appreciation of the need for increased transparency in science. This is informed by the benefits of open and transparent research practices [1–3] and by the alarming lack of both reproducibility (same results obtained using the same data and analytical steps [4]) and replicability (qualitatively similar results obtained using same analytical steps on different datasets [4,5]) of scientific findings [6–8]. Efforts to increase reproducibility have, to date, mostly focused on making research data open and, more recently, FAIR (Findable, Accessible, Interoperable, and Reusable; [9]). Indeed, publishing the data underlying a scientific finding is required by an increasing number of scientific journals and funding agencies. This has led to a considerable increase in the volume, although not necessarily the quality, of available research data [10–12].

Another major component of research, without which reproducibility is difficult to achieve, is the computer code (hereafter code) underlying research findings. Scientists routinely write code for processing raw data, conducting statistical analyses, simulating models, creating figures, and even generating computationally reproducible articles [13,14]. For many ecologists, writing customized code has become an essential part of research, regardless of whether they

conduct laboratory, field-based, or purely theoretical research [13,15,16]. Publicly sharing the code underlying scientific findings helps others to understand the analyses, evaluate the study's conclusions, reuse the code for future analyses, and it increases overall research transparency and reproducibility (e.g., [17]). This might be particularly important in ecology because this discipline commonly requires the use of sophisticated statistical models, yet ecologists are often inadequately trained in quantitative and statistical methods [18].

Across fields, journals are increasingly adopting guidelines or creating policies that require or encourage authors to make the code underlying their findings publicly available [13, 17]. Yet, to what extent researchers follow these guidelines is uncertain, as is the general availability of code in ecology. We address this, specifically focusing on analytical code (i.e., computer code used for statistical analyses and/or simulations). We focus on analytical code because this code is essential for computational reproducibility of quantitative results, and because other types of codes (e.g., data processing) are rarely provided. We randomly sampled 400 articles published between June 2015 and May 2019 in 14 ecological journals that, as of June 2015, had either a mandatory code-sharing policy or explicitly encouraged authors to make their code available upon publication (see [13]). We identified and scrutinized 346 nonmolecular articles conducting statistical analyses and/or simulations to evaluate (i) the extent of code availability in ecology, and whether code availability has increased over time; (ii) the adherence to code publishing practices supporting code findability, accessibility, and reusability; and (iii) the limits to computational reproducibility. Additionally, we reassessed the current percentage of ecological journals with mandatory or encouraged code-sharing policies (Supporting information can be accessed at https://asanchez-tojar.github.io/code_in_ecology/supporting_information.html).

Where are we now?

Amongst 96 ecological journals originally assessed for code-sharing policies by Mislán and colleagues [13], the number of journals with mandatory or encouraged code-sharing policies has increased from 14 in 2015 (15%; [13]) to 72 in 2020 (75%; S1 Table in https://asanchez-tojar.github.io/code_in_ecology/supporting_information.html). This is an encouraging increase that implies that the importance of code-sharing is now widely recognized. However, the existence of code-sharing policies does not necessarily translate into code availability (see below).

Low code availability

Our main objective was to determine what proportion of articles conducting some type of statistical analysis and/or simulations were accompanied by the underlying analytical code. In our dataset, most statistical software used (>ca 90%) were either command line interfaces (e.g., R, Python, SAS, SAS Institute, Cary, NC) or graphical user interfaces (GUIs) that allow users to extract the code or syntax of the analyses (e.g., JMP, SAS Institute, Cary, NC; SPSS, IBM Corp., Armonk, NY). For software for which we were unsure whether code or syntax of the analyses could be extracted, we expanded our definition of code to also include screenshot-based protocols or alike that would allow other researchers to reproduce the quantitative results, and if not provided, we categorized those articles as not providing code.

We found that only 92 of 346 (27%) articles in our random sample of nonmolecular articles published between 2015 and 2019 were accompanied by either seemingly all (75 articles, 22%) or some (17 articles, 5%) of the code underlying the research findings. Code availability has slightly increased over the 5-year period studied here but remains alarmingly low (23% versus 30%, in 2015–2016 and 2018–2019, respectively; Fig 1A). Furthermore, the percentage of code available for each journal ranged between 7% and 53% (median = 22%, mean = 25%; Fig 2, S3

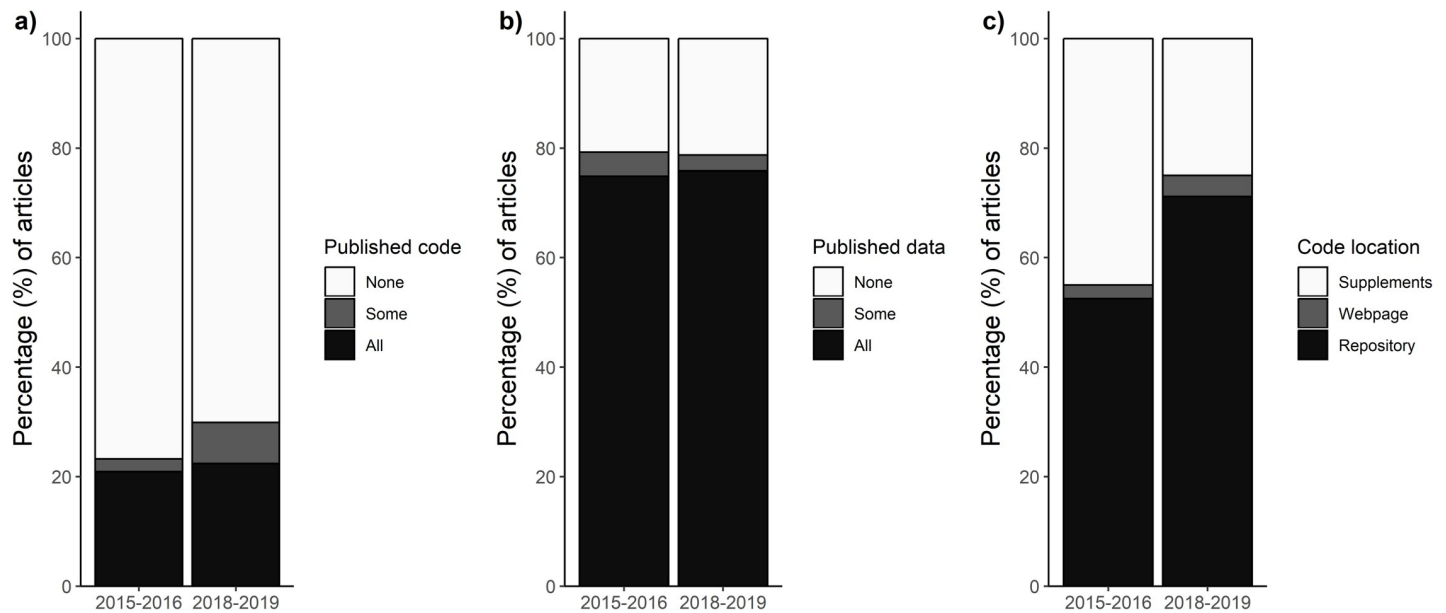


Fig 1. Code-sharing is at its infancy in ecology, whereas data-sharing seems more common. Percentage of articles reviewed that provided code (a) or data (b) for each of the time periods studied (2015–2016: 172 articles, 2018–2019: 174 articles). For studies that provided code (2015–2016: 40 articles, 2018–2019: 52 articles), the percentage of articles reviewed that hosted code in repositories (including nonpermanent platforms such as GitHub; 2015–2016: 3 articles, 2018–2019: 8 articles), web pages, or supplements is shown (c). All percentages calculated in this manuscript and a description of how they were calculated are available in S2 Table in https://asanchez-tojar.github.io/code_in_ecology/supporting_information.html. Data to reproduce this figure are available at Culina and colleagues [19]: script “004_plotting.R”.

<https://doi.org/10.1371/journal.pbio.3000763.g001>

Table in https://asanchez-tojar.github.io/code_in_ecology/supporting_information.html), indicating that low code availability seems to be a general phenomenon and not driven by just a few journals. Given that the number of journals with a code-sharing policy has increased considerably over the last few years (see above), it seems likely that, overall, code availability in ecology has also increased. Nonetheless, the compliancy with the code-sharing policies lags behind (Figs 1A and 2).

Findability, accessibility, and reusability

To be reused, code first needs to be found and accessed [20]. To be found, code availability should ideally be stated within the article or, if not, within the supplementary material of the article. In our dataset, 70 of the 92 articles (76%) that had code available enhanced “findability” by clearly referring to the availability of code within the text, mostly using the keywords “code” and/or “script” (96% of the articles that stated code availability). Code availability was most commonly stated in the “data accessibility” and/or “materials and methods” sections (94% of the articles that stated code availability), and only rarely in the “supplementary material” (4%) or the “discussion” section (1 article). Nonetheless, 24% of the articles that provided code did not explicitly state this code availability within the text. These articles commonly contained a general reference to the supplementary material and/or deposited data within the main text, but did not distinguish between code, primary data, and other supplementary materials, making code availability unclear to the reader.

We consider published code to be easily accessible when it is archived in ways that allow others to access it both now and in the foreseeable future. Our results indicate that there is some room for improvement in code accessibility because only 47 of the 92 articles (51%) that made their code available used online permanent repositories such as the Dryad Digital

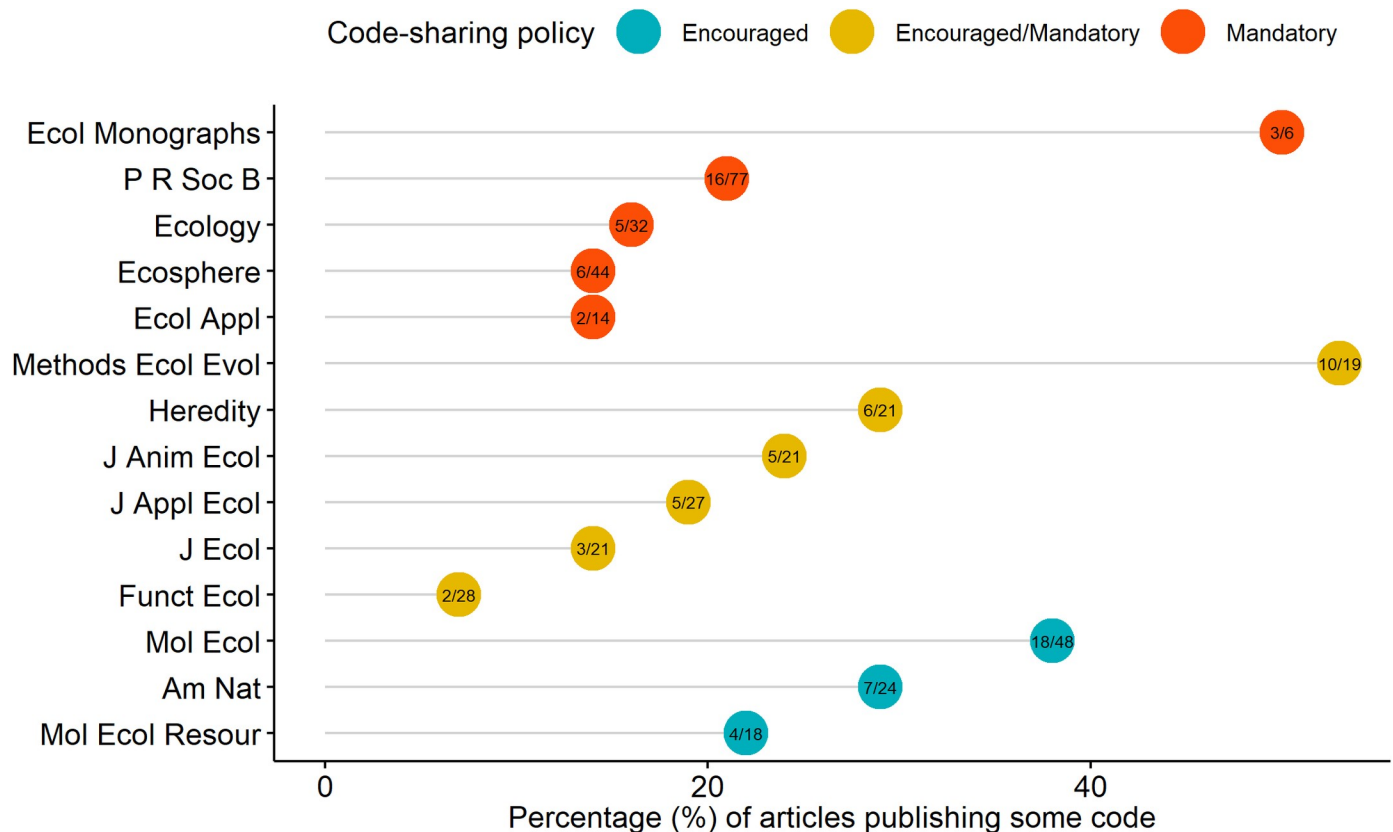


Fig 2. Estimated percentage of code-sharing for each of the 14 ecological journals reviewed. Journals are grouped by the strength of the code-sharing policy, and the ratio of articles with at least some published code to articles reviewed is shown within the circles. Full journal names are shown in S3 Table in https://asanchez-tojar.github.io/code_in_ecology/supporting_information.html. Data to reproduce this figure are available at Culina and colleagues [19]: script “004_plotting.R”. <https://doi.org/10.1371/journal.pbio.3000763.g002>

Repository, and 11 (12%) used exclusively GitHub, which is a nonpermanent, version controlled platform. A relatively large proportion of papers (34%) published their code as a supplementary material of the article. While archiving code as supplementary material does, in principle, support long-term access, not all code will be openly accessible without a journal subscription, which can reduce code accessibility. Furthermore, unlike dedicated repositories, there is no curation service in supplementary materials to ensure that files remain readable into the future and that the archive itself is maintained. Similarly, the use of nonpermanent, version controlled platforms such as GitHub should ideally be combined with publishing the code in a permanent repository (e.g., Zenodo) to protect long-term access. Overall, there has been some improvement in code publishing practices over time, with permanent repositories and GitHub gaining popularity in ecology (52% in 2015/2016 versus 71% in 2018/2019; Fig 1C). Given that many journals provide authors with free archiving of data and code in digital repositories, and given that free-of-charge archiving services such as Zenodo—which can be easily linked to GitHub—exist, we recommend that authors make their code available using these (and alike) services to enhance the long-term accessibility of their code. Furthermore, these services allow assignment of a digital object identifier (doi) to the code’s online location, which in combination with the correct license [21] makes the code more easily citable and reusable.

Finally, we explored code reusability. Code reusability depends on the complexity and documentation of the code, the software used to run the code, and the user’s own experience (e.g.,

[22]). We opted to simply evaluate whether the published code had some form of documentation (e.g., README), either as an accompanying document or embedded at the beginning of the code, and whether the code had inline comments explaining parts of the code. All code in our sample provided either one or both of these, but the level of elaboration ranged from detailed README files and/or comments to very minimal inline comments.

Limits to computational reproducibility

Kitzes and colleagues [23] state that computational reproducibility is achieved when “a second investigator (including the original researcher in the future) can recreate the final reported results of the project, including key quantitative findings, tables, and figures, given only a set of files and written instructions.” Pragmatically, such a definition requires that the data (if any used) and code are available, a requirement met by a worryingly small percentage (21%) of the articles we reviewed (i.e., articles that provided both seemingly all data [if any used] and seemingly all code). Furthermore, the potential for computational reproducibility has not improved over time (20% in 2015/2016 versus 21% in 2018/2019).

Importantly, this figure (21%) is likely an overestimate. First, as highlighted above, we expect the real percentage of sharing to be lowered when also considering journals that do not require or encourage code (and data) sharing. Second, while data availability in our sample was comparatively high (263 out of 333 articles [79%] that used data shared at least some data, including seven embargoes; Fig 1B), this is likely an overestimation of the datasets that are complete and can be fully reused. Indeed, Roche and colleagues [11] recently found that 56% of archived datasets in ecology and evolution were incomplete. Furthermore, recent attempts at estimating true computational reproducibility of published articles within ecology [24] and other fields have shown that the percentage of true computational reproducibility ranged from 18% to 80% [25–28], even when data and code are provided (sometimes via author correspondence). Third, the use of non-free (i.e., proprietary) software can also be an impediment to computational reproducibility [26]. That said, most articles in our sample (74%) used exclusively nonproprietary software, with R being the most popular software (79% of these articles; note that 34 articles [10%] did not state the software used). Last, although we did not quantify this precisely, we noted that many articles failed to provide the version of the software and packages used, which can reduce computational reproducibility substantially [29]. While we did not aim to estimate true computational reproducibility of the research findings (i.e., by running the code on the data), it is alarming that probably fewer than 21% of published research articles in ecology are computationally reproducible, with multiple factors contributing to this low proportion (Fig 3).

Where do we go from here?

The message of our review is clear: code availability in ecology is alarmingly low and, despite the existence of strong policies and guidelines in some journals, represents a major impediment to computational reproducibility of published research (Fig 3). On the one hand, there is no obvious obstacle preventing ecologists embracing the benefits of code-sharing, particularly given that ecology has experienced the shift from reliance on GUI-based statistical tools to a predominant reliance on custom-written scripts, particularly R ([16], this study). As [15] states, “If your code is good enough to do the job, then it is good enough to release—and releasing it will help your research and your field.” Our References section contains a list of several easy-to-follow guidelines on how to write and publish your code [15, 22, 30–34] and some more in-depth guidelines [35–38]. However, it seems clear that the collective benefit of code availability is insufficient encouragement to ensure good practice is widely adopted by

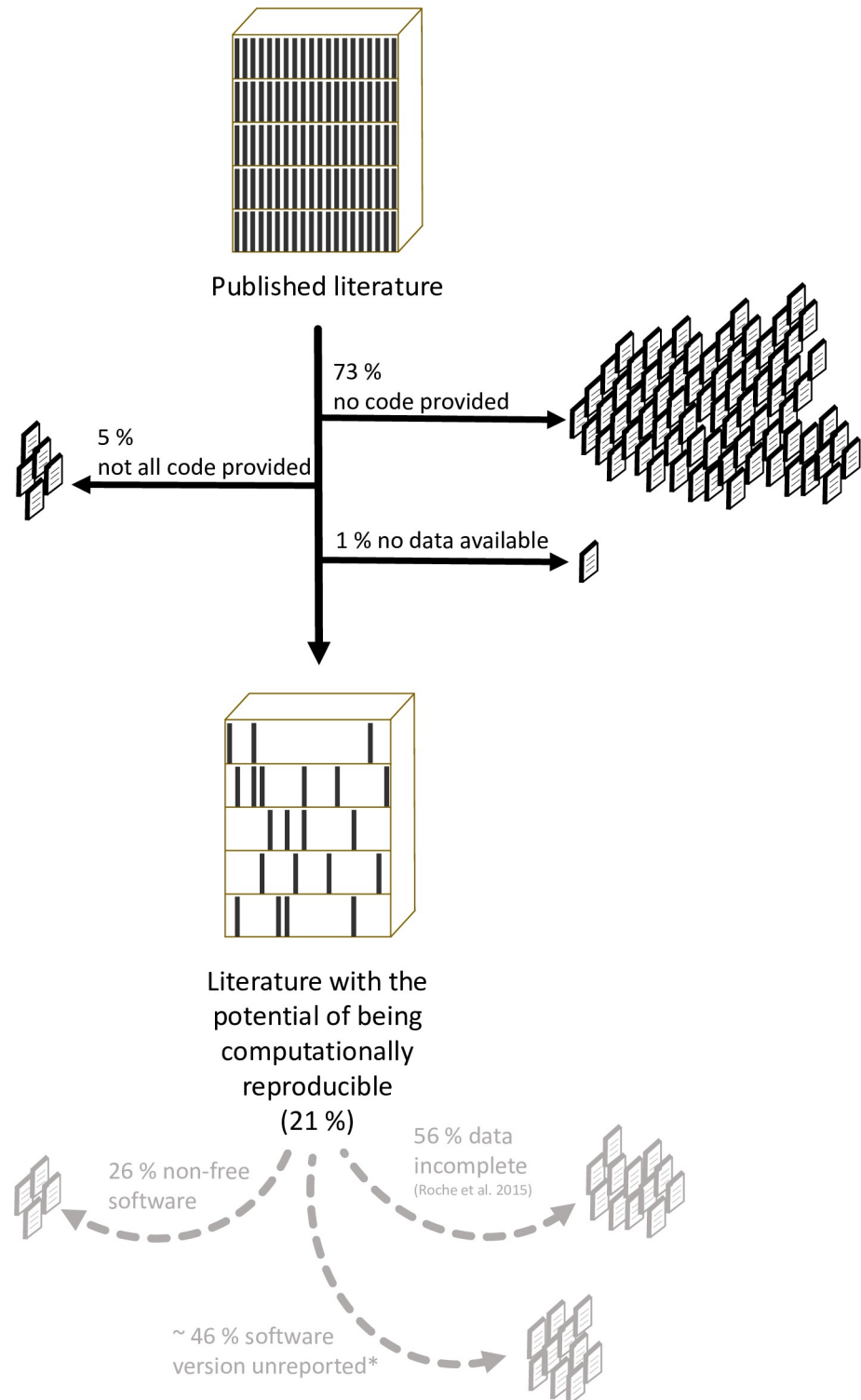


Fig 3. Depletion from published literature to literature with the potential of being computationally reproducible. Percentages estimated from a review of articles ($n = 346$) published in 14 ecological journals with mandatory or encouraged code-sharing policies, and, thus, these percentages likely overestimate the true percentage of computationally reproducible literature. Furthermore, because much code (and data) is published in nonpermanent repositories (see section “Findability, accessibility, and reusability”), long-term computational reproducibility is likely substantially lower. *Rough estimate based on the subset of articles published in 2018–2019. Data to reproduce this figure are available at [19].

<https://doi.org/10.1371/journal.pbio.3000763.g003>

authors, such that code sharing must be mandatory (or strongly encouraged). Journals, institutions, and funding agencies need to instigate and improve explicit code publishing policies (e.g., based on Transparency and Openness Promotion [TOP] Guidelines, <https://www.cos.io/top>) and to ensure these are followed by authors. For example, we noticed that it is often not easy to determine whether a journal mandates or merely encourages code sharing, which might contribute to a general failure on the part of authors to meet conditions of publication. Improving the current situation will require a concerted effort from funders, authors, reviewers, editors, and journals, not only to enforce sharing but also to increase incentives and provide relevant training to enable and encourage code- (and data-) sharing practices (see [15, 39–41]). We believe that access to code should be treated equivalently to access to primary data, i.e., as an essential (and obligatory) part of a publication in ecology.

Acknowledgments

We are grateful to Malika Ihle and Martin Stoffel for their comments on a previous version of this manuscript, and to Nicolas M. Adreani and Antje Girndt for their suggestions on Fig 3. We are, furthermore, grateful to the Library personnel of the NIOO-KNAW for their input on the literature review.

References

1. McKiernan EC, Bourne PE, Brown CT, Buck S, Kenall A, Lin J., McDougall D et al. How Open Science Helps Researchers Succeed. *eLife*. 2016; e16800. <https://doi.org/10.7554/eLife.16800> PMID: [27387362](https://pubmed.ncbi.nlm.nih.gov/27387362/)
2. Allen C, Mehler DMA. Open Science Challenges, Benefits and Tips in Early Career and Beyond. *PLOS Biol*. 2019; 17 (5): e3000246. <https://doi.org/10.1371/journal.pbio.3000246> PMID: [31042704](https://pubmed.ncbi.nlm.nih.gov/31042704/)
3. Gallagher RV, Falster DS, Maitner BS, Salguero-Gómez R, Vandvik V, Pearse WD et al. Open Science principles for accelerating trait-based science across the Tree of Life. *Nat Ecol Evol* 2020; 4, 294–303. <https://doi.org/10.1038/s41559-020-1109-6> PMID: [32066887](https://pubmed.ncbi.nlm.nih.gov/32066887/)
4. The Turing Way Community, Arnold B, Bowler L, Gibson S, Herterich P, Higman R, Krystalli A et al. The Turing Way: A Handbook for Reproducible Data Science (Version v0.0.4); 2019 [cited 2020 Jun 18]. Zenodo. Available from: <https://zenodo.org/record/3233969#.XsUXAcCxVPY>
5. Nosek BA, Errington TM. What is replication? *PLoS Biol*. 2020; 18(3): e3000691. <https://doi.org/10.1371/journal.pbio.3000691> PMID: [32218571](https://pubmed.ncbi.nlm.nih.gov/32218571/)
6. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*. 2015; 349(6251):4716
7. Freedman LP, Cockburn IM, Simcoe TS. The Economics of Reproducibility in Preclinical Research. *PLoS Biol*. 2015; 13(6): e1002165. <https://doi.org/10.1371/journal.pbio.1002165> PMID: [26057340](https://pubmed.ncbi.nlm.nih.gov/26057340/)
8. Nosek BA, Errington TM. Making sense of replications. *eLife*. 2017; 6:e23383 <https://doi.org/10.7554/eLife.23383> PMID: [28100398](https://pubmed.ncbi.nlm.nih.gov/28100398/)
9. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data*. 2016; 3 (1): 1–9. <https://doi.org/10.1038/sdata.2016.18> PMID: [26978244](https://pubmed.ncbi.nlm.nih.gov/26978244/)
10. Vines TH, Andrew RL, Bock DG, Franklin MT, Gilbert KJ, Kane NC, et al. Mandated data archiving greatly improves access to research data. *The FASEB journal*. 2013; 27(4), 1304–1308. <https://doi.org/10.1096/fj.12-218164> PMID: [23288929](https://pubmed.ncbi.nlm.nih.gov/23288929/)
11. Roche DG, Kruuk LEB, Lanfear R, Binning SA. Public Data Archiving in Ecology and Evolution: How Well Are We Doing?. *PLOS Biol*. 2015; 13 (11): e1002295. <https://doi.org/10.1371/journal.pbio.1002295> PMID: [26556502](https://pubmed.ncbi.nlm.nih.gov/26556502/)
12. Evans SR. Gauging the Purported Costs of Public Data Archiving for Long-Term Population Studies. *PLOS Biol*. 2016; 14 (4): e1002432. <https://doi.org/10.1371/journal.pbio.1002432> PMID: [27058254](https://pubmed.ncbi.nlm.nih.gov/27058254/)
13. Mislan KAS, Heer JM, White EP. Elevating the status of code in ecology. *Trends Ecol Evol*. 2016; 31 (1), 4–7. <https://doi.org/10.1016/j.tree.2015.11.006> PMID: [26704455](https://pubmed.ncbi.nlm.nih.gov/26704455/)

14. Maciocci G, Aufreiter M, Bentley N. Introducing eLife's first computationally reproducible article. 2019 [cited 2020 Apr 6]. Available from: <https://elifesciences.org/labs/ad58f08d/introducing-elife-s-first-computationally-reproducible-article>
15. Barnes N. Publish your computer code: it is good enough. *Nature*. 2010; 467(7317), 753. <https://doi.org/10.1038/467753a> PMID: 20944687
16. Lai J., Lortie CJ, Muenchen RA, Yang J, Ma K. Evaluating the popularity of R in ecology. *Ecosphere*. 2019; 10, e02567. <https://doi.org/10.1002/ecs2.2567>
17. Stodden V., Guo P, Ma Z. Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals. *PLOS ONE*. 2013; 8 (6): e67111. <https://doi.org/10.1371/journal.pone.0067111> PMID: 23805293
18. Touchon JC, McCoy MW. The mismatch between current statistical practice and doctoral training in ecology. *Ecosphere*. 2016; 7(8):e01394. <https://doi.org/10.1002/ecs2.1394>
19. Culina A, van den Berg I, Evans S, Sánchez-Tójar A. Code and data needed to reproduce the results shown in: Low availability of code in ecology: a call for urgent action (Version v.1.0.0). Zenodo. 2020 [cited 2020 Apr 6]. Available from: <https://doi.org/10.5281/zenodo.3833928>
20. Lamprecht AL, Garcia L, Kuzak M, Martinez C, Arcila R, Martin E. et al. Towards FAIR Principles For Research Software'. *Data Science*. 2019; 1–23. <https://doi.org/10.3233/DS-190026>
21. Stodden V. Enabling Reproducible Research: Open Licensing for Scientific Innovation. *International Journal of Communications Law and Polic*. 2009 [cited 2020 Apr 6]. Available at SSRN: <https://ssrn.com/abstract=1362040>
22. Stodden V, McNutt M, Bailey DH, Deelman E, Gil Y, Hanson B et al. Enhancing Reproducibility for Computational Methods. *Science*. 2016; 354 (6317): 1240–41. <https://doi.org/10.1126/science.aah6168> PMID: 27940837
23. Kitzes J, Turek D, Deniz F. *The Practice of Reproducible Research: Case Studies and Lessons from the Data-Intensive Sciences*. Univ of California Press. 2017.
24. Archmiller A, Johnson A, Nolan J, Edwards M, Elliott L, Ferguson J. et al. Computational Reproducibility in The Wildlife Society's Flagship Journals. *J Wild Manag*. 2020; <https://doi.org/10.1002/jwmg.21855>
25. Gertler P, Galiani S, Romero M. How to Make Replication the Norm. *Nature*. 2018; 554 (7693): 417–19. <https://doi.org/10.1038/d41586-018-02108-9>
26. Stockemer D, Koehler S, Lentz T. Data Access, Transparency, and Replication: New Insights from the Political Behavior Literature. *Political Sci & Politics*. 2018; 51 (4): 799–803. <https://doi.org/10.1017/S1049096518000926>
27. Stodden V, Seiler J, Ma Z. An Empirical Analysis of Journal Policy Effectiveness for Computational Reproducibility. *PNAS*. 2018; 115 (11): 2584–89. <https://doi.org/10.1073/pnas.1708290115> PMID: 29531050
28. Obels P, Lakens D, Coles NA, Gottfried J, Green SA. Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology. *PsyArXiv*. 2019. <https://doi.org/10.31234/osf.io/fk8vh>
29. Pasquier T, Lau MK, Trisovic A, Boose ER, Couturier B, Crosas M et al. If These Data Could Talk. *Sci Data*. 2017; 4 (1): 1–5. <https://doi.org/10.1038/sdata.2017.114> PMID: 28872630
30. Osborne JM, Bernabeu MO, Bruna M, Calderhead B, Cooper J, Dalchau N, et al. Ten Simple Rules for Effective Computational Research'. *PLOS Comp Biol*. 2014; 10 (3): e1003506. <https://doi.org/10.1371/journal.pcbi.1003506> PMID: 24675742
31. Piccolo SR, Frampton MB. Tools and Techniques for Computational Reproducibility. *GigaScience*. 2016; 5 (1): 30. <https://doi.org/10.1186/s13742-016-0135-4> PMID: 27401684
32. Cooper N, Hsing P-Y. *A Guide to Reproducible Code in Ecology and Evolution*. British Ecological Society; 2017.
33. Eglen S., Marwick B., Halchenko Y., Hanke M, Sufi S, Gleeson P, et al. Toward standard practices for sharing computer code and programs in neuroscience. *Nat Neurosci* 2017; 20, 770–773. <https://doi.org/10.1038/nn.4550> PMID: 28542156
34. Wilson G, Bryan J, Cranston K, Kitzes J, Nederbragt L, Teal TK. Good Enough Practices in Scientific Computing. *PLOS Com Biol*. 2017; 13 (6): e1005510. <https://doi.org/10.1371/journal.pcbi.1005510> PMID: 28640806
35. Wickham H. *testthat: Get Started with Testing*. *The R Journal*. 2011; 3, 5–10 [cited 2020 Apr 6]. Available from: https://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf.
36. Fischetti T. The *assertr* package. 2020 [cited 2020 Apr 6]. Available from: <https://github.com/ropensci/assertr>
37. Hampton SE, Anderson SS, Bagby CS, Gries C, Han X, Hart EM, Jones MB, et al. The Tao of open science for ecology. *Ecosphere*. 2015; 6: 120. <https://doi.org/10.1890/ES14-00402.1>

38. Choose an open source license [cited 2020 Apr 6]. Available from: <https://choosealicense.com/>
39. Gezelter J. D. Open Source and Open Data Should Be Standard Practices *J. Phys. Chem. Lett.* 2015, 6, 1168–1169 <https://doi.org/10.1021/acs.jpcl.5b00285> PMID: 26262967
40. Sholler D, Ram K, Boettiger C, Katz DS. Enforcing Public Data Archiving Policies in Academic Publishing: A Study of Ecology Journals. *Big Data & Soc.* 2019. <https://doi.org/10.1177/2053951719836258>
41. Woolston C. TOP Factor rates journals on transparency, openness. *Nature Index.* 2020 [cited 2020 Apr 6]. Available from: <https://www.natureindex.com/news-blog/top-factor-rates-journals-on-transparency-openness>