



Citizen science contributes to SDM validation

[Francisco Lloret](#) based on reviews by Maria Angeles Perez-Navarro and 1 anonymous reviewer

A recommendation of:

Florence Matutini, Jacques Baudry, Guillaume Pain, Morgane Sineau, Josephine Pithon. **How citizen science could improve Species Distribution Models and their independent assessment** (2020), *bioRxiv*, 2020.06.02.129536, ver. 4 peer-reviewed and recommended by Peer Community in Ecology. [10.1101/2020.06.02.129536](https://doi.org/10.1101/2020.06.02.129536)

Submitted: 03 June 2020, Recommended: 30 September 2020

Cite this recommendation as:

Francisco Lloret (2020) Citizen science contributes to SDM validation. *Peer Community in Ecology*, 100059. [10.24072/pci.ecology.100059](https://doi.org/10.24072/pci.ecology.100059)

Open Access

Published: 30 Septembre 2020

Copyright: This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/4.0/>

Citizen science is becoming an important piece for the acquisition of scientific knowledge in the fields of natural sciences, and particularly in the inventory and monitoring of biodiversity (McKinley et al. 2017). The information generated with the collaboration of citizens has an evident importance in conservation, by providing information on the state of populations and habitats, helping in mitigation and restoration actions, and very importantly contributing to involve society in conservation (Brown and Williams 2019). An obvious advantage of these initiatives is the ability to mobilize human resources on a large territorial scale and in the medium term, which would otherwise be difficult to finance. The resulting increasing information then can be processed with advanced computational techniques (Hochachka et al 2012; Kelling et al. 2015), thus improving our interpretation of the distribution of species. Specifically, the ability to obtain information on a large territorial scale can be integrated into studies based on Species Distribution Models SDMs. One of the common problems with SDMs is that they often work from species occurrences that have been opportunistically recorded, either by professionals or amateurs. A great challenge for data obtained from non-professional citizens, however, remains to ensure its standardization and quality (Kosmala et al. 2016). This requires a clear and effective design, solid volunteer training, and a high level of coordination that turns out to be complex (Brown and Williams 2019). Finally, it is essential to perform a quality validation following scientifically recognized standards, since they are often conditioned by errors and biases in obtaining information (Bird et al. 2014). There are two basic approaches to obtain the necessary data for this validation: getting it from an external source (external validation), or allocating a part of the database itself (internal validation or cross-validation) to this function.

Matutini et al. (2020) in his work 'How citizen science could improve Species Distribution

Models and their independent assessment' shows a novel application of the data generated by a citizen science initiative ('Un Dragon dans mon Jardin') by providing an external source for the validation of SDMs, as a tool to construct habitat suitability maps for nine species of amphibians in western France. Importantly, 'Un Dragon dans mon Jardin' contains standardized presence-absence data, the approximation recognized as the most robust (Guisan, et al. 2017). The SDMs to be validated, in turn, were based on opportunistic information obtained by citizens and professionals. The result shows the usefulness of this external data source by minimizing the overestimation of model accuracy that is obtained with cross-validation with the internal evaluation dataset. It also shows the importance of properly filtering the information obtained by citizens by determining the threshold of sampling effort.

The destiny of citizen science is to be integrated into the complex world of science. Supported by the increasing level of the formation of society, it is becoming a fundamental piece in the scientific system dedicated to the study of biodiversity and its conservation. After funding for scientists specialized in the recognition of biodiversity has been cut back, we are seeing a transformation of the activity of these scientists towards the design, coordination, training and verification of programs for the acquisition of field information obtained by citizens. A main goal is that a substantial part of this information will eventually get integrated into the scientific system, and rigorous verification process a fundamental element for such purpose, as shown by Matutini et al. (2020) work.

References

- [1] Bird TJ et al. (2014) Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation* 173: 144-154. doi: [10.1016/j.biocon.2013.07.037](https://doi.org/10.1016/j.biocon.2013.07.037)
- [2] Brown ED and Williams BK (2019) The potential for citizen science to produce reliable and useful information in ecology. *Conservation Biology* 33: 561-569. doi: [10.1111/cobi.13223](https://doi.org/10.1111/cobi.13223)
- [3] Guisan A, Thuiller W and Zimmermann N E (2017) *Habitat Suitability and Distribution Models: With Applications in R*. The University of Chicago Press. doi: [10.1017/9781139028271](https://doi.org/10.1017/9781139028271)
- [4] Hochachka WM, Fink D, Hutchinson RA, Sheldon D, Wong WK and Kelling S (2012) Data-intensive science applied to broad-scale citizen science. *Trends Ecol Evol* 27: 130-137. doi: [10.1016/j.tree.2011.11.006](https://doi.org/10.1016/j.tree.2011.11.006)
- [5] Kelling S, Fink D, La Sorte FA, Johnston A, Bruns NE and Hochachka WM (2015) Taking a 'Big Data' approach to data quality in a citizen science project. *Ambio* 44(Supple. 4):S601-S611. doi: [10.1007/s13280-015-0710-4](https://doi.org/10.1007/s13280-015-0710-4)
- [6] Kosmala M, Wiggins A, Swanson A and Simmons B (2016) Assessing data quality in citizen science. *Front Ecol Environ* 14: 551-560. doi: [10.1002/fee.1436](https://doi.org/10.1002/fee.1436)
- [7] Matutini F, Baudry J, Pain G, Sineau M and Pithon J (2020) How citizen science could improve Species Distribution Models and their independent assessment. *bioRxiv*, 2020.06.02.129536, ver. 4 peer-reviewed and recommended by PCI Ecology. doi: [10.1101/2020.06.02.129536](https://doi.org/10.1101/2020.06.02.129536)
- [8] McKinley DC et al. (2017) Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation* 208:15-28. doi: [10.1016/j.biocon.2016.05.015](https://doi.org/10.1016/j.biocon.2016.05.015)

Revision round #2

2020-09-18

Preprint DOI: <https://www.biorxiv.org/content/10.1101/2020.06.02.129536v3>

Author's reply:

[Download author's reply \(PDF file\)](#)

Revision round #1

2020-07-25

The paper address an interesting topic, which is the feasibility and realibility of data provided by citizen science platforms to furnish information about species distribution models. The topic is extremely novel at a time in which the link between citizens and sciences is becoming strengthened, and natural sciences aim extensive scientific information - for instance for conservation purposes- , while keeping standards of quality. The paper is well structured and written, attaining its objectives. However it still needs some relevant improvements. As pointed by referees, the manuscript needs to reinforce some strategical issues, such as a critical assessment of the use of citizen science in terms of weekenesses, and clarify somewhat its goal, since conservation application of the contributions of the study case is not fully addressed. The revisors are overall positive with the paper, but correctly identify that there are several methodological clarifications that should be addressed: bias treatment (accessibility, attractiveness, sampling effort), particularly when dealing with pseudo-absences, many details on data sources, (access web, program name, institutions,), collection and sampling design, or criteria to set thresholds to establish absence data, among others.

Additional requirements of the managing board:

As indicated in the 'How does it work?' section and in the code of conduct, please make sure that:

-Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.

-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.

-Details on experimental procedures are available to readers in the text or as appendices.

-Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI XXX recommenders."

Preprint DOI: <https://www.biorxiv.org/content/10.1101/2020.06.02.129536v1>

Reviewed by anonymous reviewer, 2020-07-22 13:49

[Download the review \(PDF file\)](#)

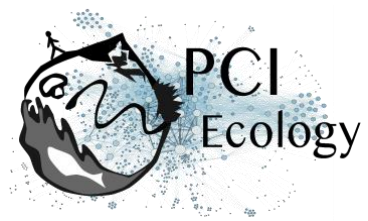
Reviewed by [Maria Angeles Perez-Navarro](#), 2020-07-12 20:01

[Download the review \(PDF file\)](#)

Author's reply:

Dear Editor,

Thank you for your message. We would also like to express our warm thanks to the reviewers for the very relevant evaluation they did on our paper. After a careful reading of the reviewers' comments, we did our best to take into account their comments and suggestions and we hope that the new version of our manuscript has been improved in term of quality. Let us now give a point-by-point answer in the PDF



attached file. We are currently working on additional requirements of the managing board to make some complementary files available to readers (metadata and scripts). It'll be available in few days.

All the best,

Florence Matutini

[Download author's reply \(PDF file\)](#)