

A community perspective on research software in the geosciences

Robert Reinecke, Tim Trautmann, Thorsten Wagener, Katja Schüler

<http://poll remodel.global>

Is reproducibility an issue in the geosciences?

- Software development has become an integral part of the geosciences¹ as models and data processing get more sophisticated
- Paradoxically, it poses a threat to scientific progress as the pillar of science, reproducibility, is seldomly reached²
- This is especially worrisome as scientific results have potential controversial implications for stakeholders and policymakers and may influence the public opinion for a long time³

Is reproducibility an issue in the geosciences?

- In recent years, progress towards open science has led to more publishers demanding access to data and source code alongside peer-reviewed manuscripts^{4,5}
- Still, recent studies find that results can rarely be reproduced^{6,7}

only a small fraction of articles (1.6% of tested articles) using their available artifacts. We estimated, with 95% confidence, that results might be reproduced for only 0.6% to 6.8% of all 1,989 articles. Unlike prior studies, the survey tool identified key bottlenecks to making work

case. Our results show that there are many limitations in terms of access to the code for these climate models and that the climate modelling community needs to improve its code-sharing practice to comply with best practice in this regard and the most recent editorial publishing policies.

Open Access | Published: 26 February 2019

Assessing data availability and research reproducibility in hydrology and water resources

James H. Stagge, David E. Rosenberg, Adel M. Abdallah, Hadia Akbar, Nour A. Attallah & Ryan James

Current status on the need for improved accessibility to climate models code

Juan A. Añel¹, Michael García-Rodríguez^{1,2}, and Javier Rodeiro²



We would like to find out what the reasons are

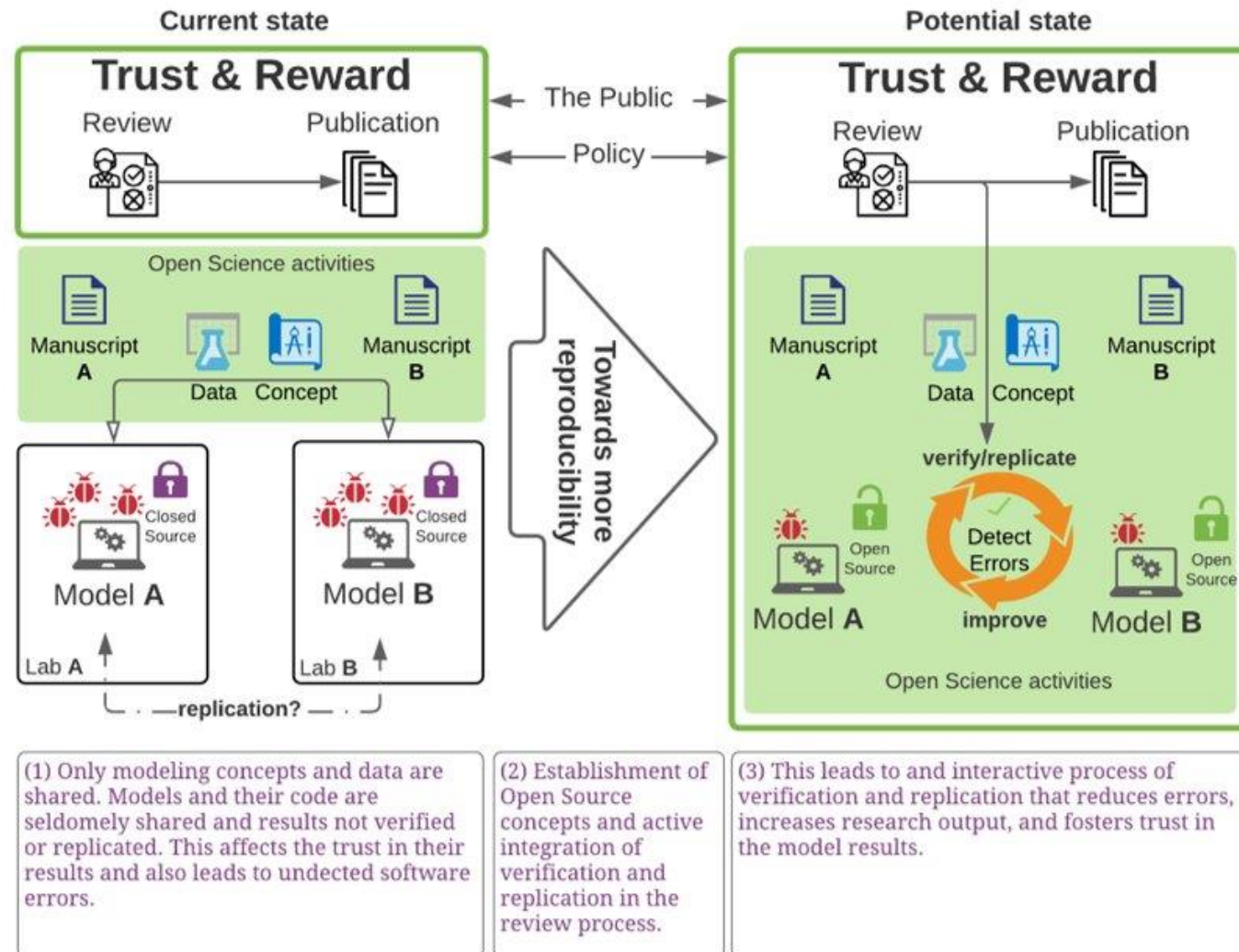
Recent publications suggest that a reproducibility issue exists

- But how does our community reproduce results?
- Are the available publications a mirror of the every-day research reality?
- To what do scientist attribute a lack of reproducible science?

We suspect that some of these reasons might be at the root of the problem

- Increasing software complexity and lack of expertise on how to build such software
- Lack of documentation of that complex software
- Code is rarely shared and a lack of knowledge about Open Source and licensing
- Missing funding and time to improve existing software

Our concept of the current and potential state



What does the community think?

To answer all these questions, we decided to ask, **you**, the geoscience community!

Because the poll is still open (**please share and participate**) and we do not want to introduce a bias we are not able to go into too much detail about current results and the questions that we ask.

-> <http://poll remodel.global>

The concept of our poll

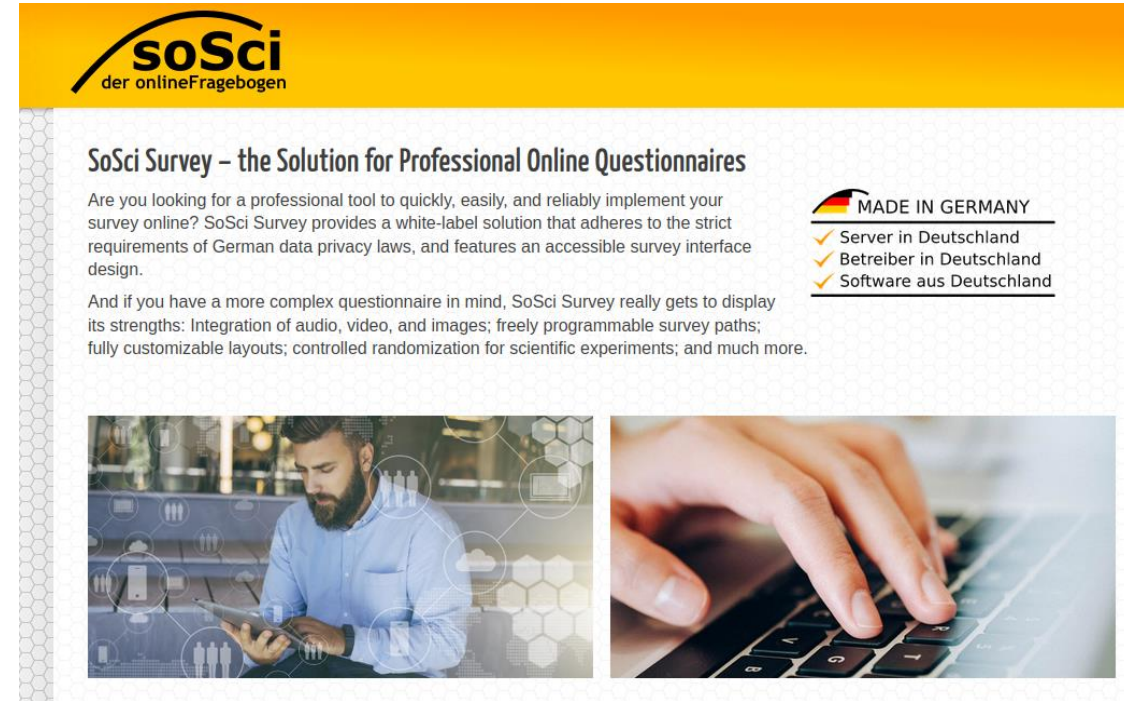
21 questions in total

With 4 Categories:

- Demographic information
- Opinion
- Behavior
- Solutions

Design methodology:

- 1) Internal discussion -> multiple hypothesis
- 2) Development of initial questions based on hypothesis
- 3) Pretest with 10 selected researchers that were not involved in the design of the study
- 4) Refinement of study based on pretest
- 5) -> final poll that is current running



soSci
der onlineFragebogen

SoSci Survey – the Solution for Professional Online Questionnaires

Are you looking for a professional tool to quickly, easily, and reliably implement your survey online? SoSci Survey provides a white-label solution that adheres to the strict requirements of German data privacy laws, and features an accessible survey interface design.

And if you have a more complex questionnaire in mind, SoSci Survey really gets to display its strengths: Integration of audio, video, and images; freely programmable survey paths; fully customizable layouts; controlled randomization for scientific experiments; and much more.

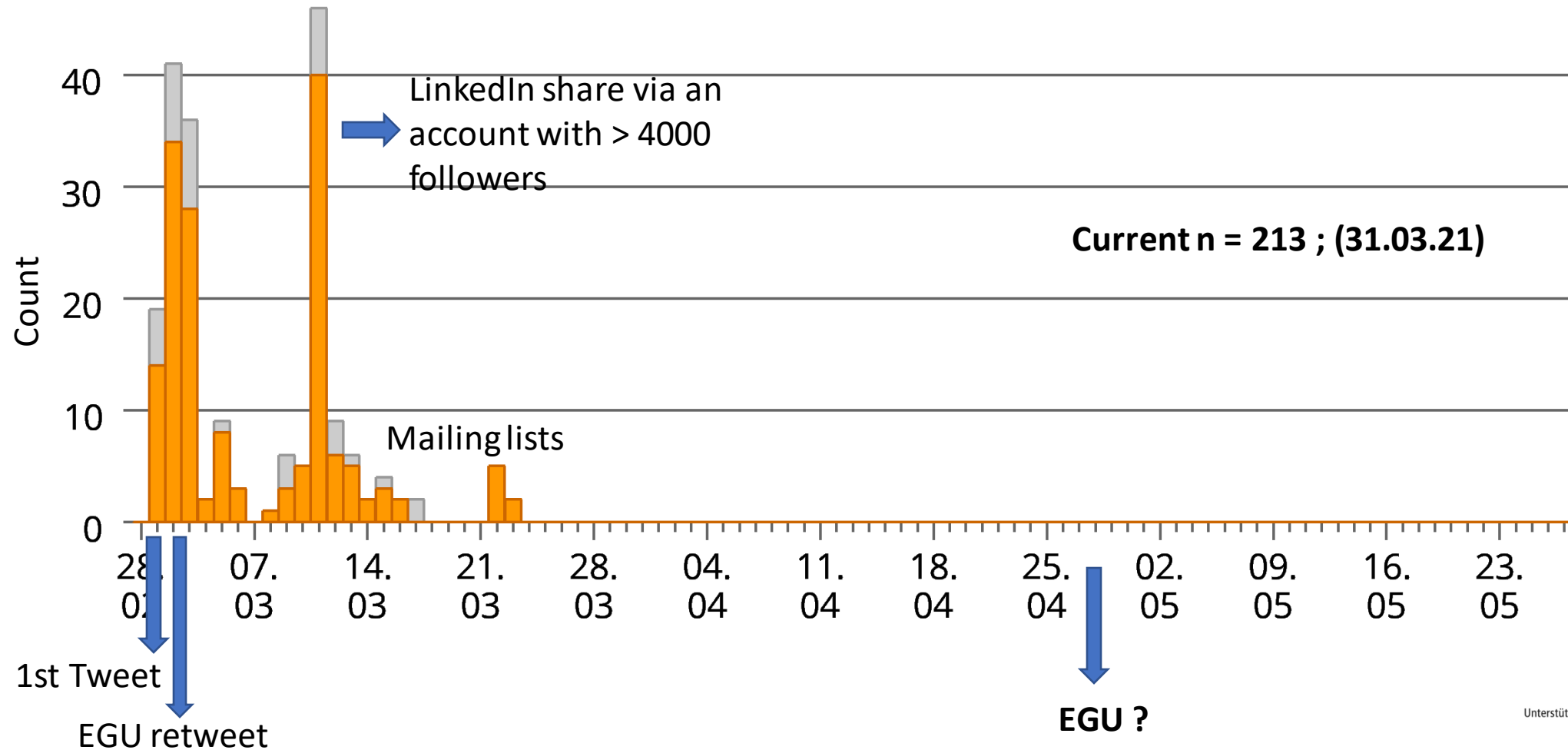
MADE IN GERMANY

- ✓ Server in Deutschland
- ✓ Betreiber in Deutschland
- ✓ Software aus Deutschland

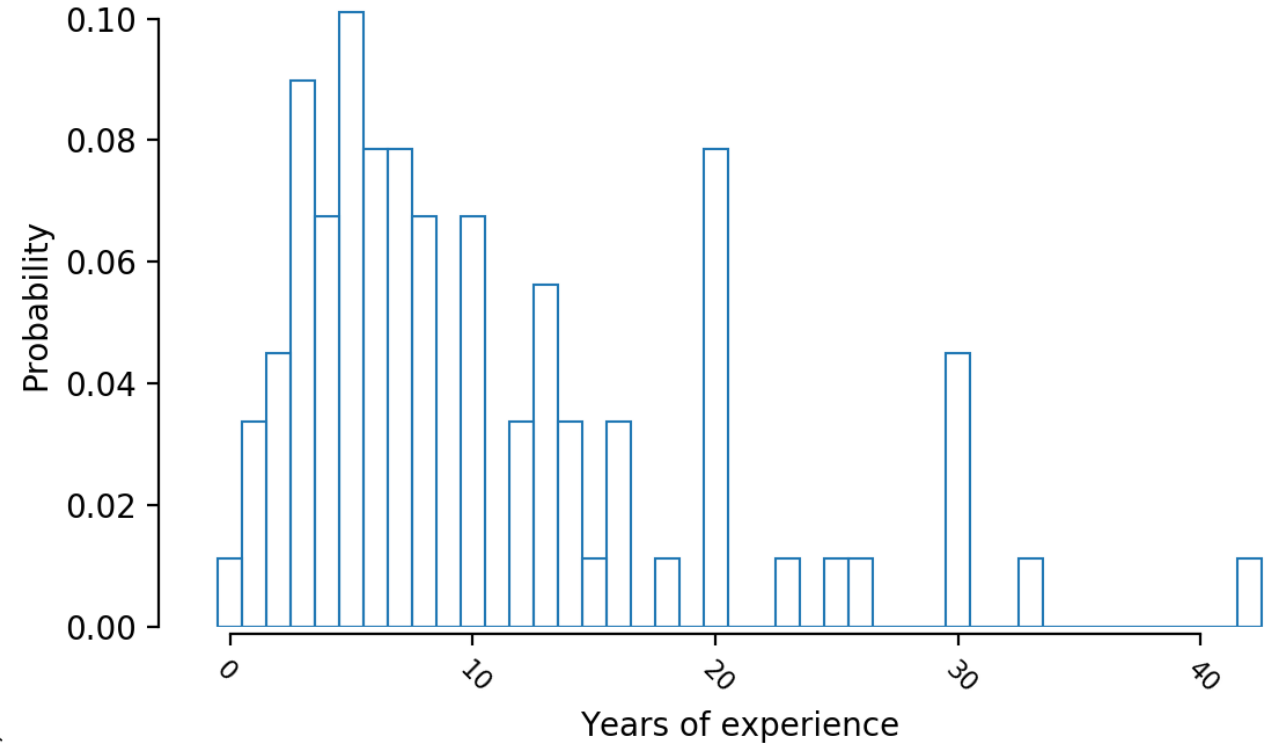
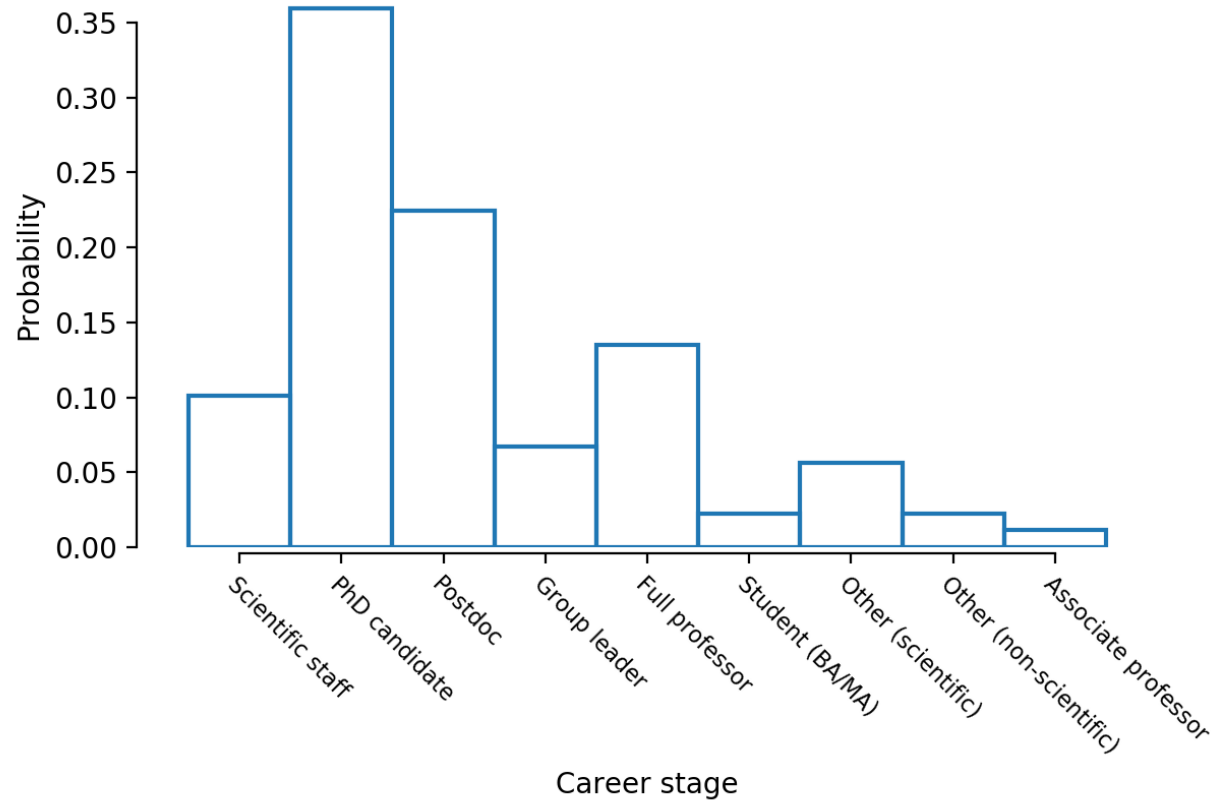


We used SoSci for our poll

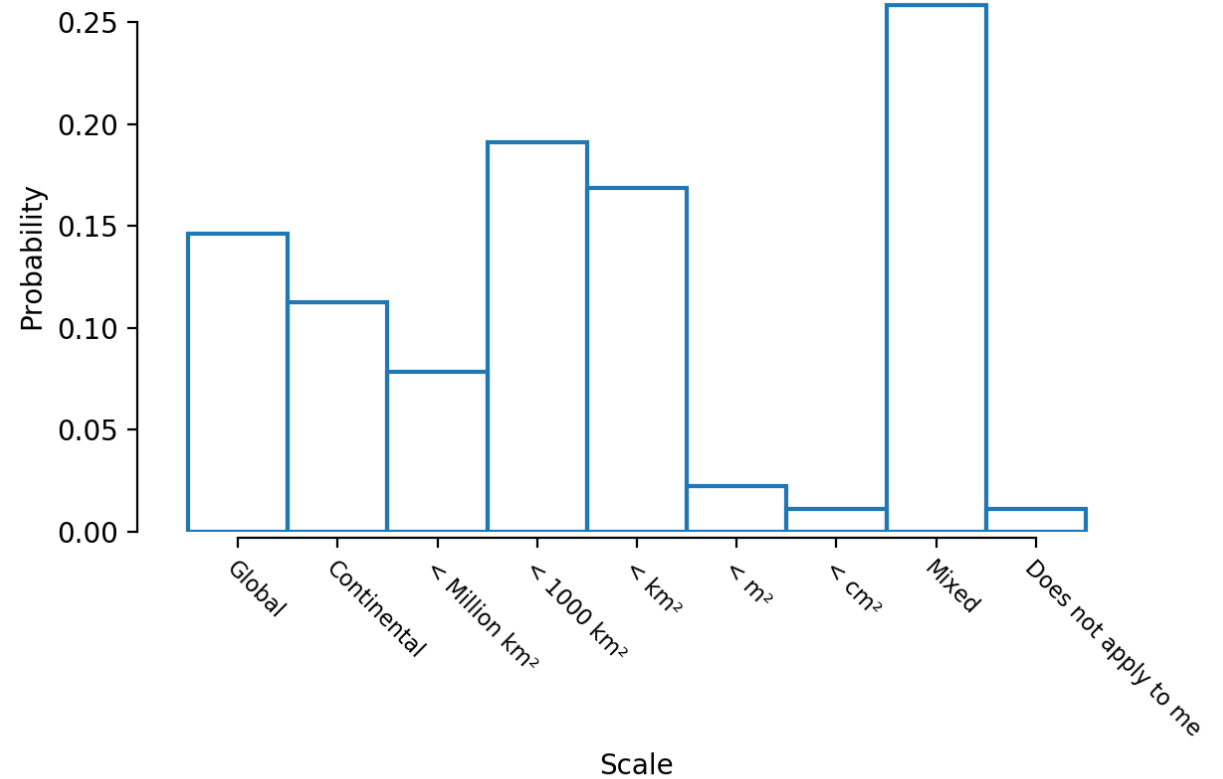
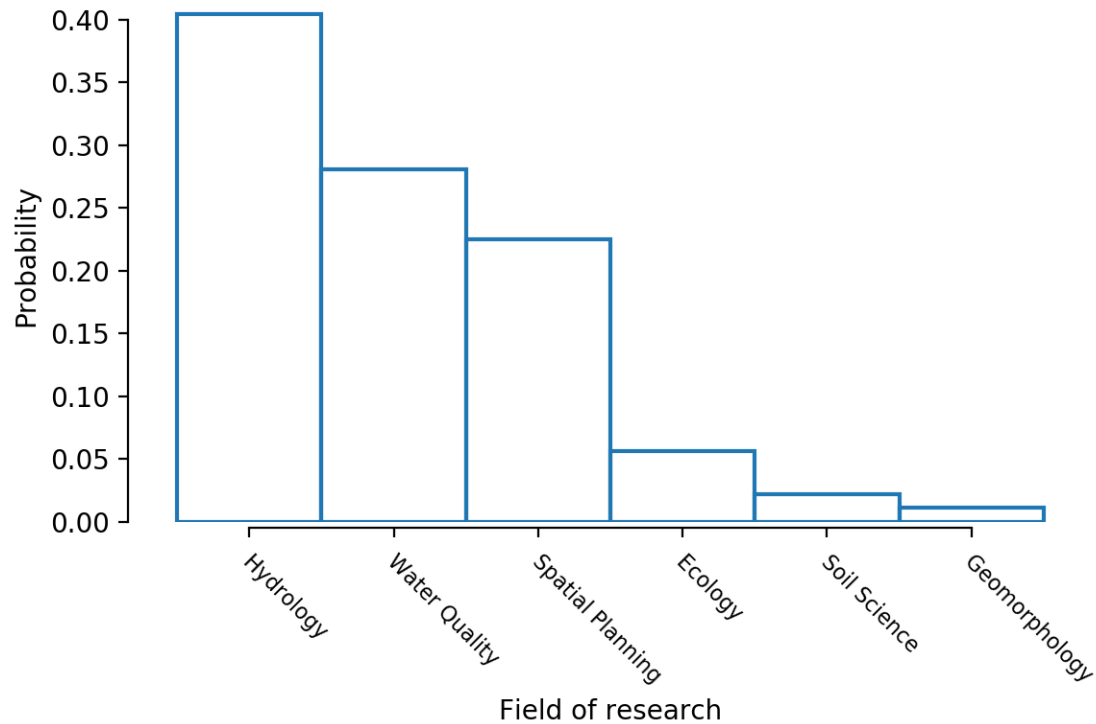
Participation is closely linked to advertisement



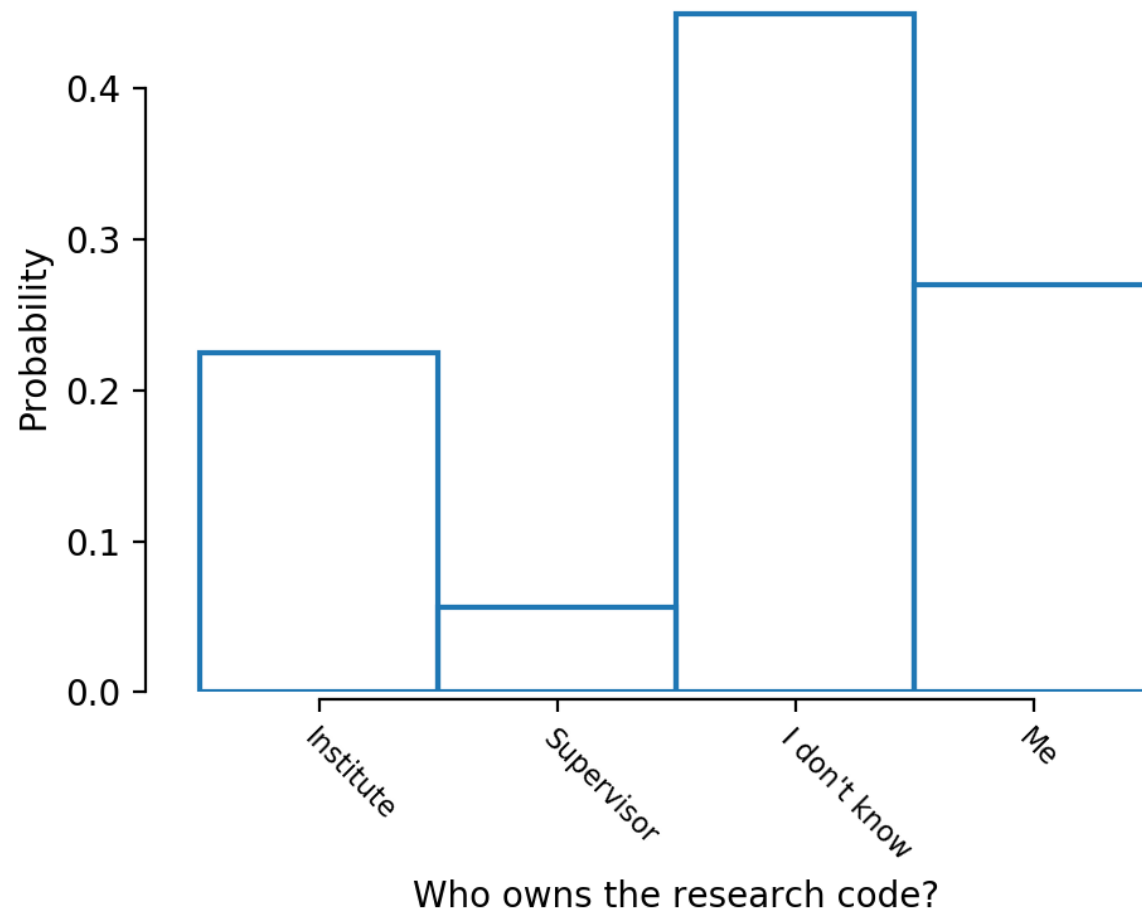
35% are PhD candidates or postdocs with less than 10 years of experience in the field



Most participants are from hydrology and work on various scales



Most scientists don't know who their code belongs to



To prevent introducing a bias we are not able to show more details at this moment :(

Participate! to explore the other questions :)

<http://poll remodel.global>



Data access and closing date

Closing date: **31.05.21**

The data will **openly accessible** and published alongside an article that summarizes the outcomes.

Updates will be available on <http://remodel.global> and on ResearchGate (<https://www.researchgate.net/project/REMODEL-Reproducible-MODELing>)

Questions? Comments?

-> robert.reinecke@uni-potsdam.de



References

- ¹ A comment to "Most Computational Hydrology is not Reproducible, so is it Really Science?" R.W. Hut, N.C. van de Giesen, N. Drost, Water Resources Research, 2017
- ² Hutton, C., Wagener, T., Freer, J., Han, D., Du_y, C., and Arheimer, B., Most computational hydrology is not reproducible, so is it really science? Water Resources Research, 2016
- ³ Munafò, M., Nosek, B., Bishop, D. *et al.*, A manifesto for reproducible science. Nat Hum Behav, 2017
- ⁴ Executive editors, G. Editorial: The publication of geoscientific model developments v1.2. Geoscientific Model Development, 2019
- ⁵ Katz, D. S., Niemeyer, K. E., and Smith, A. M., Publish your software: Introducing the journal of open source software (joss), Computing in Science Engineering, 2018
- ⁶ Stagge, J. H., Rosenberg, D. E., Abdallah, A. M., Akbar, H., Attallah, N. A., and James, R., Assessing data availability and research reproducibility in hydrology and water resources. Scientific data, 2019
- ⁷ Añel, J. A., García-Rodríguez, M., and Rodeiro, J.: Current status on the need for improved accessibility to climate models code, Geosci. Model Dev., 2021.
- ⁸ Stodden, V., The reproducible research standard: Reducing legal barriers to scientific knowledge and innovation. IEEE Computing in Science & Engineering, 2009
- ⁹ <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>
- ¹⁰ Muller, C., Schaphoff, S., von Bloh, W., Thonicke, K., and Gerten, D., Going open-source with a model dinosaur and establishing model evaluation standards. EGU, 2018
- ¹¹ <https://software.rajivprab.com/2019/11/25/the-birth-of-legacy-software-how-change-aversion-feeds-on-itself>