Automated report

Best Practices On Reprogramming A Legacy Global Water Model

Emmanuel Nyenah

March 6, 2025

Report last updated:	March 06, 2025
Compilation date:	March 6, 2025
N total:	217
N completed:	64

This document summarizes a full analysis for detailed results. **It is not a full publication but rather an automated representation of the data presented in this repository.** It will also be distributed along with the data in the course of a journal and data publication. The following analysis only takes into account the fully completed poll from 64 participants.

Contents

List of Figures

1 Aim of this survey

This survey aims to gather user insights on a reprogrammed model's code snippet. It evaluates the code's readability, comprehensibility, modifiability, and documentation quality in implementing the Priestley-Taylor potential evapotranspiration.

2 Results

All data processing and plotting (including building this document) can be executed by running `python run.py`. Plotting details and addtional processing can be found in th script `plot_all.py`.

2.1 Sample Characteristics (demographics)

Who were our participants? Here we present characteristics of the participants in our poll, i.e. their current career stage, their years of research experience, their geo-scientific field and scale as well as their current focus of work. This is purely descriptive statistics. As we welcomed everyone to our poll, we did not form any assumptions regarding sample characteristics. Also, we tried, but might not have reached a representative sample of the population of geo-scientists.
Here, we report basic sample characteristics. Corresponding survey questions:

- DM01 - What phase of your career are you currently in?
- DM02 - How long have you been working in your research field?
- DM03 - To which field within the geosciences does your research mainly belong?
- DM04 - What geographic scale are you working on?
- DM05 - What is the focus of your research?

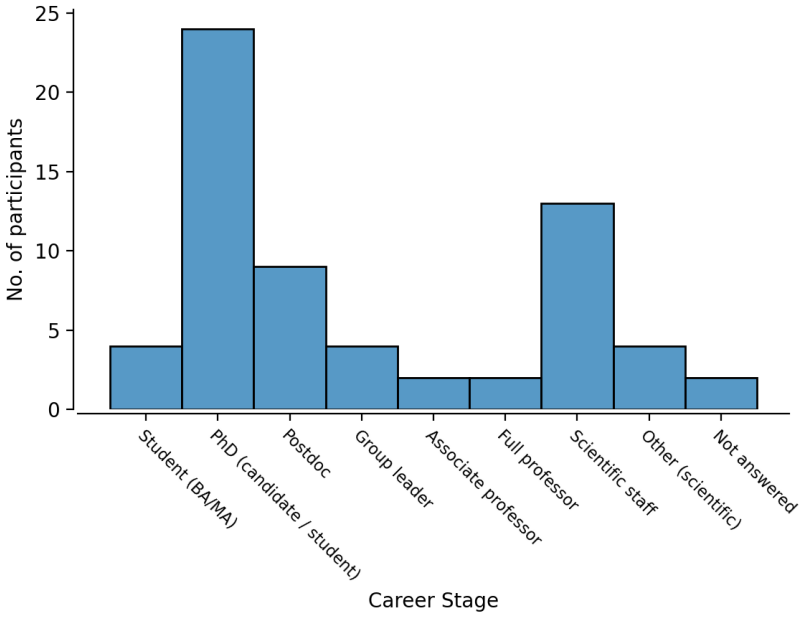

Figure 1: DM01 - Career stage of participants

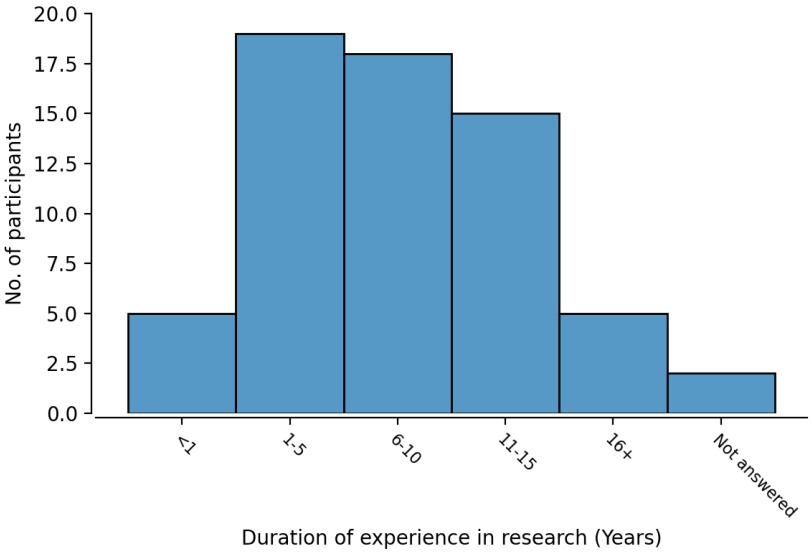

Figure 2: DM02- How long have you been working in your research field?

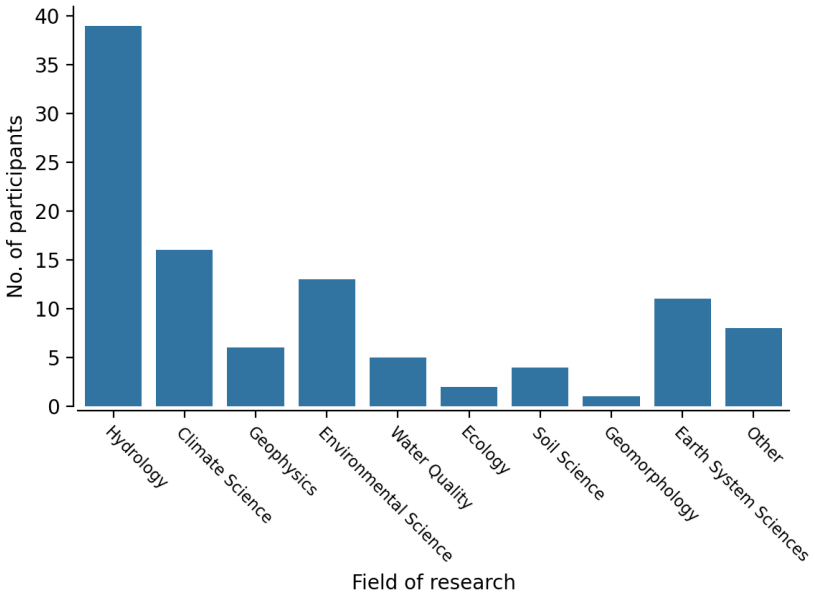

Figure 3: DM03 - To which field within the geosciences does your research mainly belong?

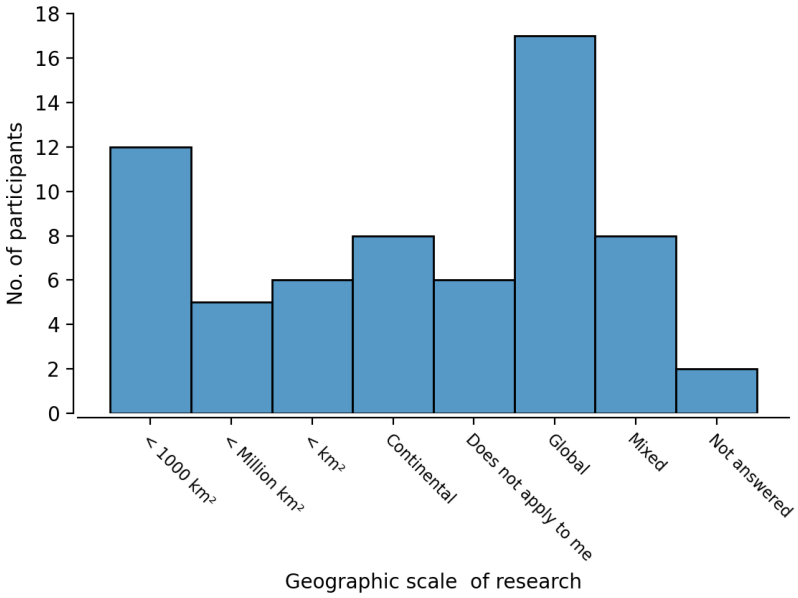

Figure 4: DM04 - What geographic scale are you working on?

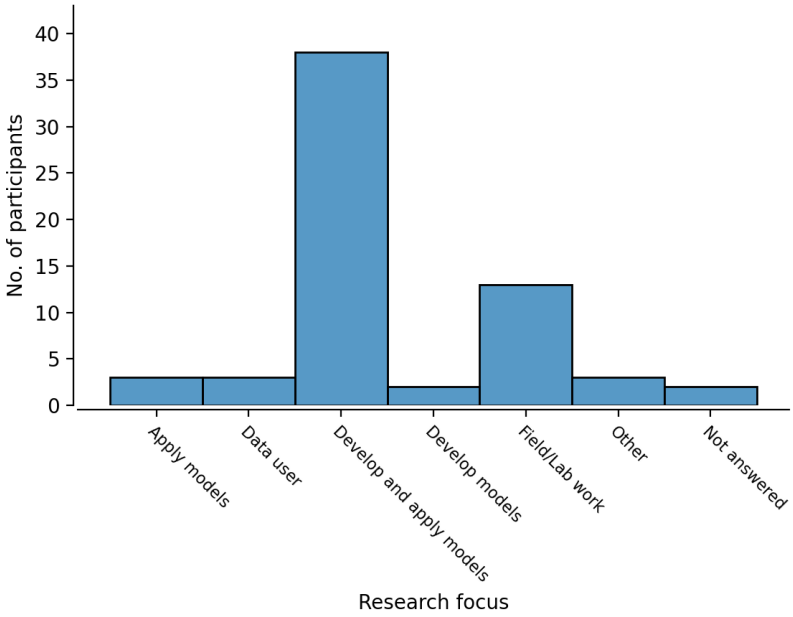

Figure 5: DM05 - What is the focus of your research?

2.2 User Skill

We aim to assess whether participants have experience in programming or using research software, as well as their familiarity with current best practices in software development. For this survey, we define research software as including source code files, algorithms, computational workflows, and executables developed during the research process or for research purposes [1].

1 Barker, M., Chue Hong, N.P., Katz, D.S. et al. Introducing the FAIR Principles for research software. Sci Data 9, 622 (2022). https://doi.org/10.1038/s41597-022-01710-x

Below, we present basic sample characteristics. Corresponding survey questions:

- US01 - How often do you use research software?
- US04 - Do you own a research software?
- US03 - How often do you develop code for research software in your research practice?
- US05- How many years of programming experience do you have?
- US06 - What kind of programming languages do you use?
- US08 - How did you learn to program software?
- US09 - Do you practice any of the following methods / concepts as part of your scientific programming routine?
- US07 - Do you have any background in Hydrology?

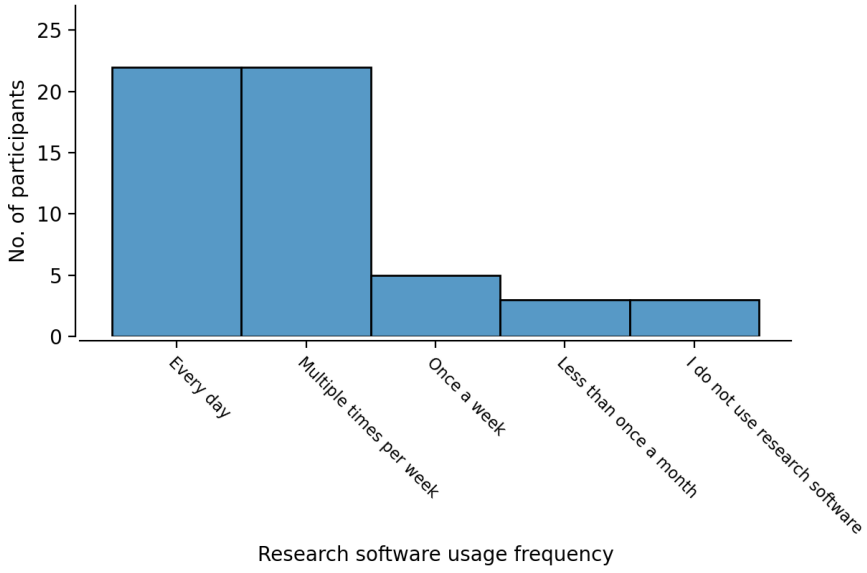

Figure 6: US01 - How often do you use research software?

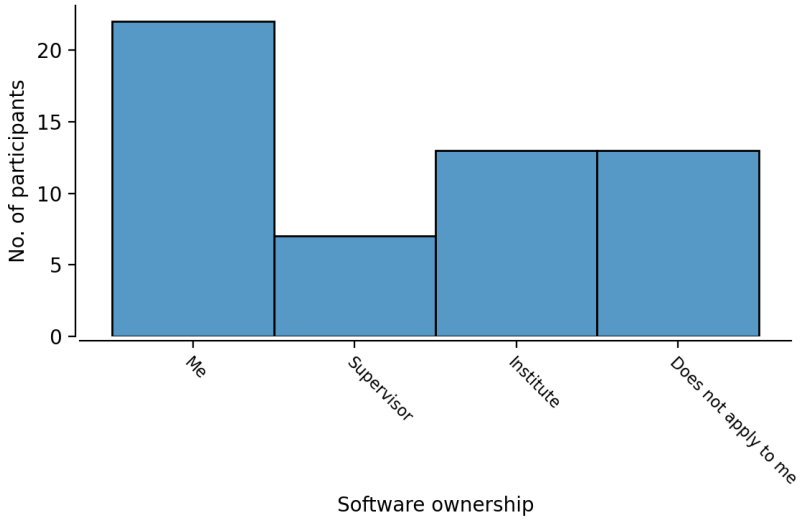

Figure 7: US04- Do you own a research software?

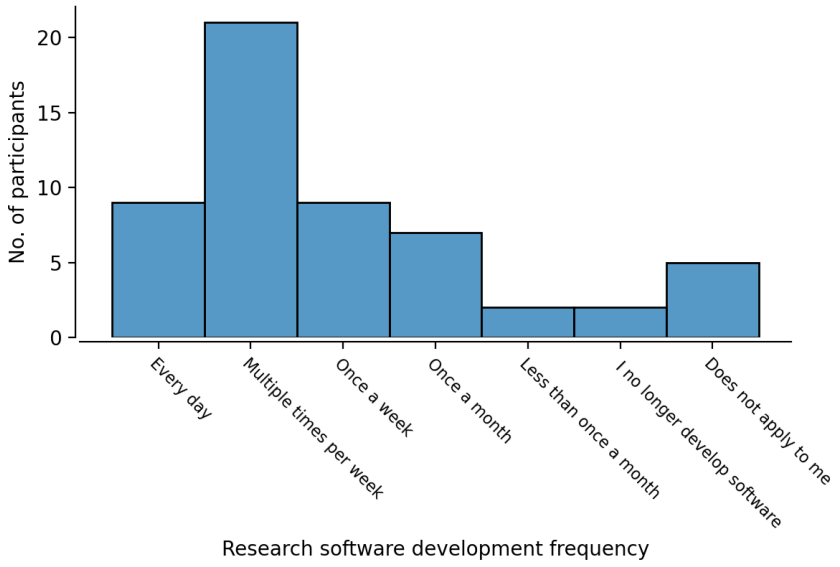

Figure 8: US03 - How often do you develop code for research software in your research practice?

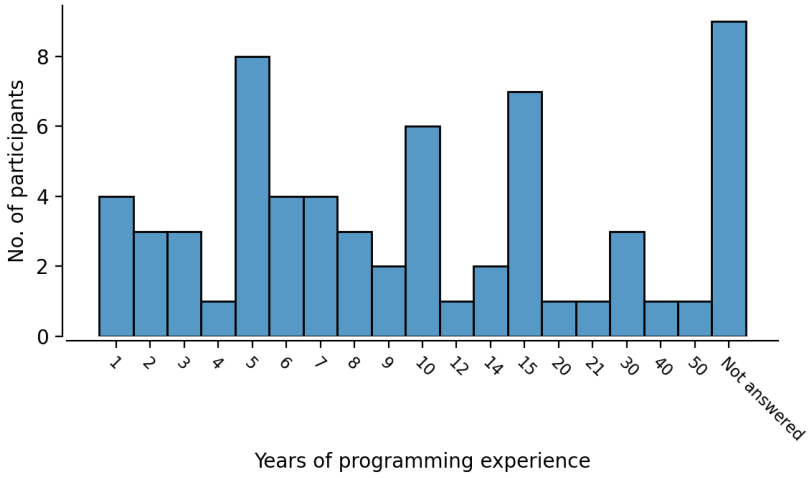

Figure 9: US05 - How many years of programming experience do you have?

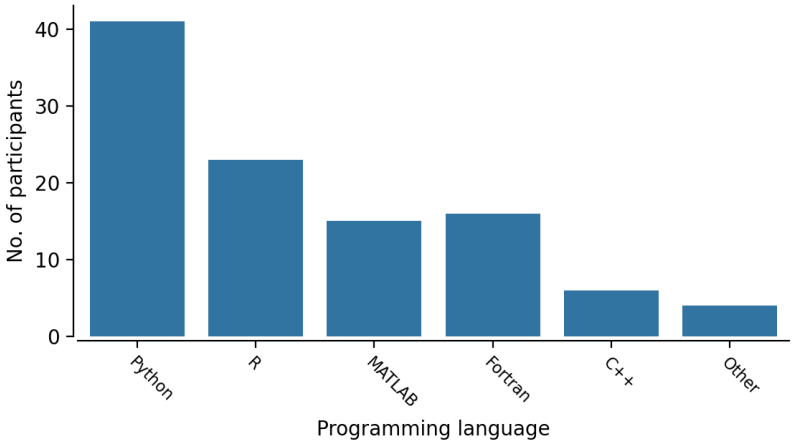

Figure 10: US06 - What kind of programming languages do you use?

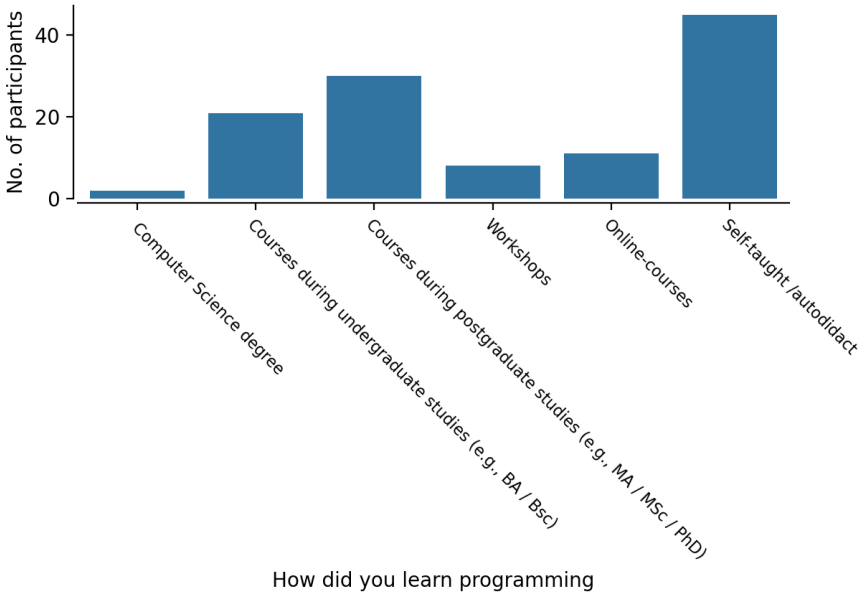

Figure 11: US08 - How did you learn to program software?

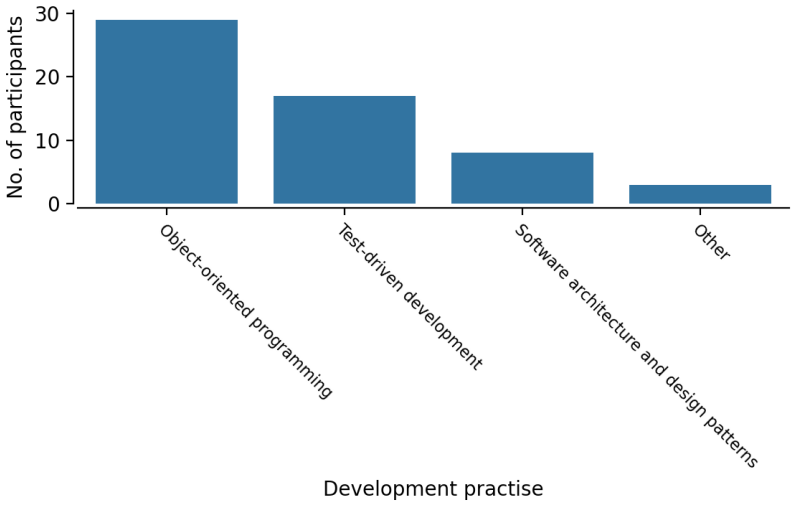

Figure 12: US09 - Do you practice any of the following methods / concepts as part of your scientific programming routine?

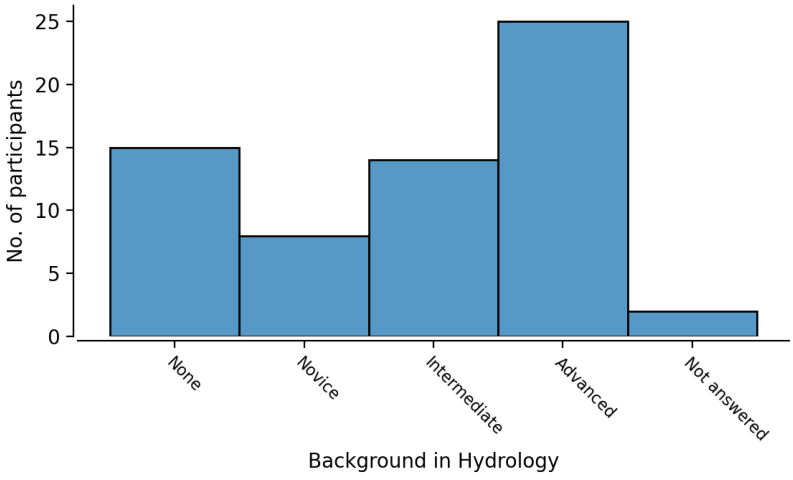

Figure 13: US07 - Do you have any background in Hydrology?

2.3 User Perspective

We asked users to answer the following questions about the Priestley-Taylor potential evapotranspiration code snippet. The goal was to gather feedback to improve the source code's readability, modifiability, and usability, as well as to identify potential contributions to model development.
Below, we present basic sample characteristics.
Corresponding survey questions:

- UP02 - Can you identify what equation is used in the code snippet below?
- UP04 - Is the additional documentation helpful in understanding the function shown previously?
- UP03 - How confident are you to change atmospheric pressure constant from 101.3kPa to 101.325kPa in the equation you identified ?
- UP07 - Which line of code would you modify ?

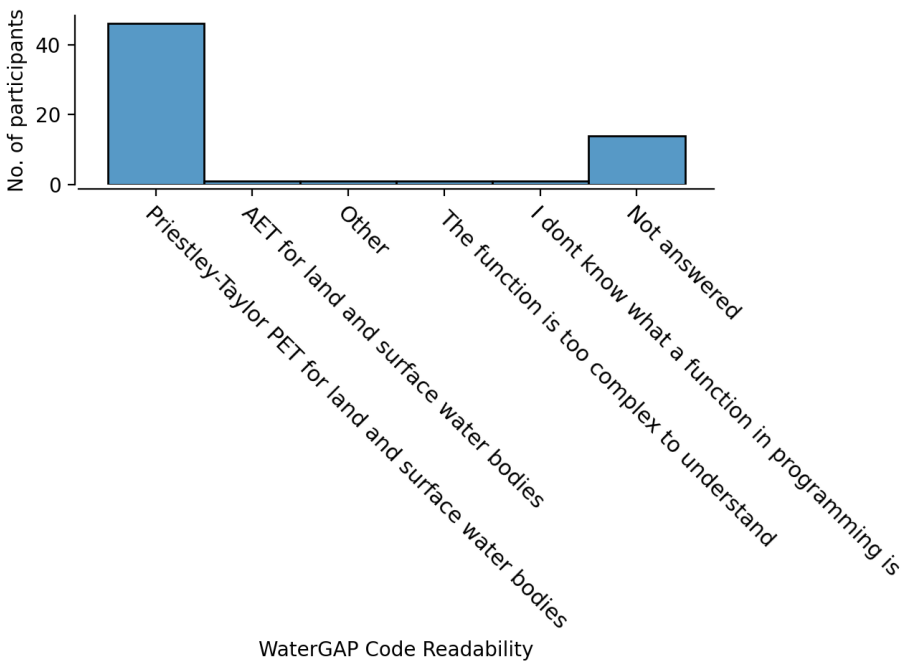

Figure 14: UP02 - Can you identify what equation is used in the code snippet below?

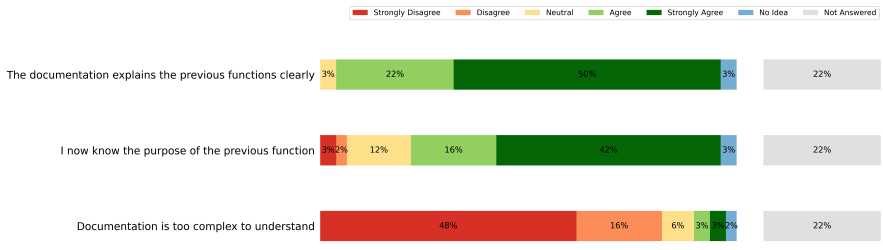

Figure 15: UP04- Is the additional documentation helpful in understanding the function shown previously?

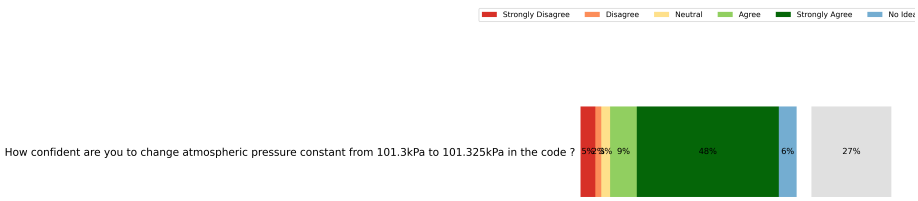

Figure 16: UP03 - How confident are you to change atmospheric pressure constant from 101.3kPa to 101.325kPa in the equation you identified ?

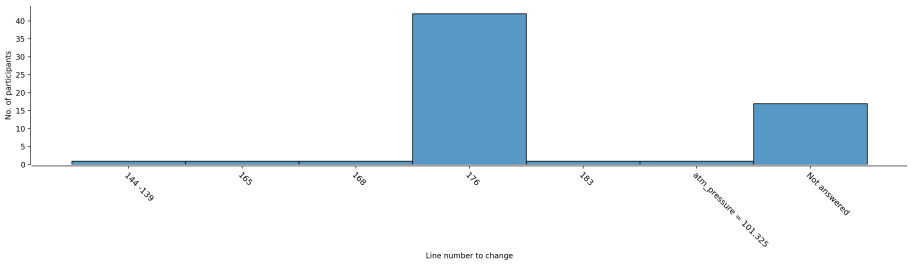

Figure 17: UP07- Which line of code would you modify ?.

2.4 Suggestions for Enhancing Software Sustainability

- No suggestions
- Spread awareness about code style, quality and readability with my peers and friends.
- Nothing
- Annual user workshops or hackathons to foster collaboration
- (1) Take the constants out of the individual equations and store them in a dedicated file. (2) Example workflows are absolutely critical to improve uptake. If those don't exist yet I would strongly recommend making them.
- Modification of packages should not conflict with previous scripts that a script that once was running can no longer run. Community must be created in levels of expertise.
- I think you are doing really great work and bring a huge contribution to the awareness about sustainable research software, at the moment I don't have suggestions, thank you for your work!
- It's a bit a short line for such an elaborate question, the use of tests, code formatters and linters, ongoing maintenance and improvement.
- Make students and PhDs familiar with it. If they are used to it, they will stay with the product and tell others.
- Updating the documents based on the modified scripts ASAP. Less number of classes and try to have a minimum of upper-level classes.
- Offering the video tutorial on utilizing and executing the model would greatly benefit users, particularly those who are new to the field.
- Making it as easily usable as possible, also for beginners / people with limited experience; providing regular courses, workshops, etc. to promote the software; supplying easy-to-follow examples on the documentation web page / on GitHub.
- It looks like very well documented.
- Online classes, good manual. Learning video.
- I did not use this, so sorry.
- Not used.
- Training - using it to solve real-world issues. Case studies. Research that seeks to solve a problem more directly.
- It will depend on each of the program codes.

- Introduction of software in scientific societies like EGU.
- Sample input data.
- A good documentation on how to use the model, tutorials, automatic testing, a good website.
- Active call(s) for interested collaborators.