

Considering the non-programming geographer's perspective when designing extracurricular introductory computer programming workshops

Summary: Computer programming is becoming an increasingly important scientific skill, but geographers are not necessarily receiving this training as part of their formal education. While there are efforts to promote and support extracurricular introductory computer programming workshops, there remain questions about how best to deliver these workshops. Therefore, as part of a recent introductory programming extracurricular workshop I organized for non-programming geographers, I tried to understand more about their perceptions of computer programming. I identify that one of the most important aspects for geographers to learn to computer program is to have training that is domain specific to ensure that the training is relevant and achieves a deeper learning outcome.

1 Introduction

Computer programming is becoming increasingly important for science to improve analyses and documentation, but a scientist's formal education may often lack appropriate training [1]. This can result in wasted time and energy, and possibly even retraction of research [5]. This is also true for geographers as with the growth of neogeography, big data, and open geographic information system (GIS) software, more geographers will need to learn how to program a computer [4]. While anecdotally I would say that computer programming is becoming a more common component of geography curricula, the only formal surveys of geography degree curricula I am aware of are from the US where it was found that computer programming is not usually emphasized or required as part of geography degrees [9, 2].

In the absence of formal training at tertiary institutions, there are organizations that have formed to try and address this gap. For example, recognizing this mismatch between computational requirements and formal education, in 2010 the UK research coun-

cils funded the founding of the Software Sustainability Institute, whose mission includes increasing computer programming skills through extracurricular training workshops [3]. The Software Carpentry [10] and Data Carpentry [7] workshops are an excellent example of this type of effort, as these workshops use volunteer effort from the computational science community to provide training in such things as computer programming, version control, data handling, data analysis, and data visualization.

However, as well as telling people what they need to know about computer programming from the perspective of someone with programming skills, I also think it is important to understand how learning computer programming is viewed from the non-programmer's perspective so that the training can be made as effective as possible. Therefore, as part of a recent introductory programming extracurricular workshop I organized for non-programming geographers, I asked questions about their perceptions of computer programming. I have summarized here what I think are some interestingly consistent themes and how these might affect the design of extracurricular introductory computer programming workshops.

2 Pre-workshop survey

Given the limited space available for the workshop, potential participants were asked to complete a survey as part of an application process to ensure that their attendance would be suitable. There were 18 applicants for the workshop, most whom had interests in human geography, so the following does need to be interpreted in that context. However, the applicants did cross a spectrum of educational levels, from BSc/BA to PhD, and careers, with students, researchers, university academics, and professionals all applying.

Perhaps the most important finding was that 83% of applicants reported that they never had the option to undertake any computer programming training as part of their formal geography education. So, there are certainly going to be geographers who need this kind of training, which means that extracurricular workshops clearly do have value. However, the most selected reason for not attending training in the past is an absence of training relevant to geographers (Figure 1).

My perception is that there are lots of online tutorials for teaching programming in a generic sense, but not too much that is directly targeted at geographical applications. This issue of relevance was further underlined by some of the following comments made about what applicants thought might be good ways of getting non-programming geographers to learn programming:

- "... provide examples of programming tasks/techniques related to the discipline of geography, and to relate programming exercises to research tasks likely to be undertaken by geographers on a day-to-day basis."
- "... using different types of analysis but with real data so you can actually try and get the answer out the data. I've tried ... following a handout and it made little sense to me as the numbers that were used didn't mean anything as it was a fake data set."
- "Hands on exercises and clear examples (those that as researchers we normally encounter)."

A desire to learn with relevant examples and real data is something I have encountered before when trying to teach computer programming to geographers within a formal academic setting [4]. The essence of those findings was that by using relevant examples,



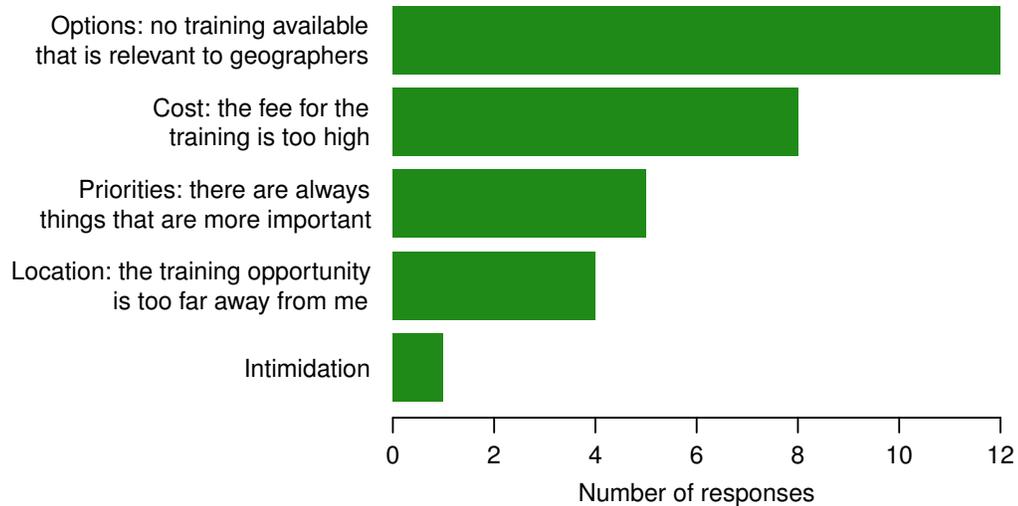


Figure 1: Survey responses to the question: What has stopped you from attending computer programming training in the past?

potentially by opening the learning process such that learners can have some control over what examples they use, facilitates a deeper learning outcome.

Reducing the number of perceived barriers to learning by using relevant examples is also important as the perception of the applicants was that learning to computer program would be difficult (Figure 2a). Fortunately, there was a clear perception that programming would be a very useful skill to develop (Figure 2b). So, while there will be geographers out there who are motivated to learn to program, I do think instructors need to remove as many obstacles to learning as possible, as with a preconceived idea that programming is hard, I think there is the potential for many learners to give up if they perceive or encounter additional difficulties with the learning process. Using relevant geographical data and examples will probably help remove obstacles to understanding context and application. But there are technical obstacles, too, and getting software installed is often harder than using it, especially if participants are using their own computers [11]. Therefore, it is probably prudent to use more reliable software packages as part of an introductory workshop, rather than the most recent packages that may still be developing, to try and avoid technical complications that introduce additional obstacles to learning.

In terms of what non-programming geographers want to learn, there was a clear preference among some applicants for R and Python as desirable programming languages for geographers. In terms of programming tasks, relatively generic tasks such as statistics, plotting graphs, and data handling were popular, but so were other geographer specific tasks such as map making and spatial analyses, which perhaps again underlines the need for workshops to contain some domain-specific content (Figure 3).

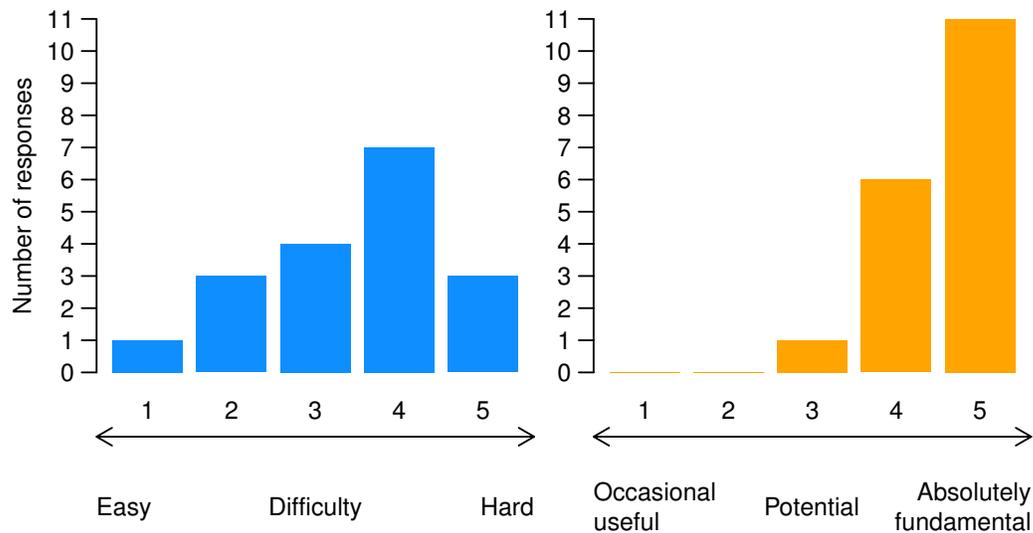


Figure 2: Survey responses to the questions: (a) How do you perceive the difficulty in learning computer programming? (b) How do you perceive the potential benefits to your work of being able to computer program?

3 Workshop design

The workshop was based on the R programming language [6] and was one-day long (Figure 4). Content was designed to introduce programming skills that were highlighted during the pre-workshop survey in as geographically a focused a way as possible. Highly detailed examples and exercises were avoided as: 1) participants clearly could not learn everything they would need in one day, 2) they will all have different specific interests and needs, and 3) people need to be able to continually learn as software changes. Therefore, the workshop objective was simply to try and demystify programming with R by touching briefly on some of the most important topics, to enable the participants to more effectively self-learn what they specifically need in the future. That said, the use of open-ended assessments as a deep learning tool has long been advocated [8]. Therefore, I did build in a “consultancy” session at the end of the day where participants could get some advice and help with their own specific problems and data in the hope this might provide an opportunity to better contextualize the workshop content. Such an open session would require a high instructor to participant ratio, so the workshop was limited to eleven participants and was supported with three instructors. Also, participants were asked to highlight topics of interest during the lunch break to try and find any common interests that could be focused on in smaller groups.

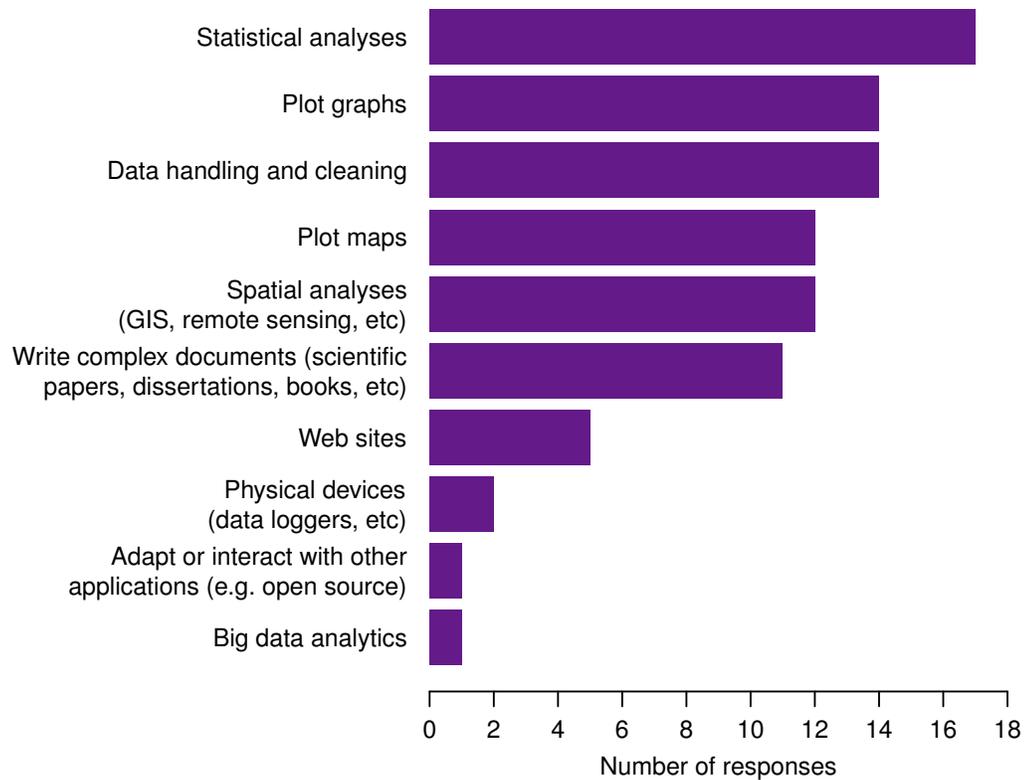


Figure 3: Survey responses to the question: What types of things would you like to be able to do with computer programming?

4 Post-workshop feedback

To try and assess how effective a one-day introductory workshop can be in introducing computer programming I asked for feedback from the participants that reflected on their experience, or lack of experience, with computer programming. From this feedback, I have identified some statements along consistent themes across multiple participants that speak towards some questions I had around the value of such workshops from the perspective of the non-programming geographer.

4.1 Is there a need for introductory programming workshops?

- “The reason I haven’t learned to program before is that I thought it was difficult and that required a considerable amount of time.”
- “With no formal training in computer programming during my studies, I was left needing to acquire programming skills in the workplace, usually as a side venture, and often on my own time. This is a daunting task, requires a lot of time that one

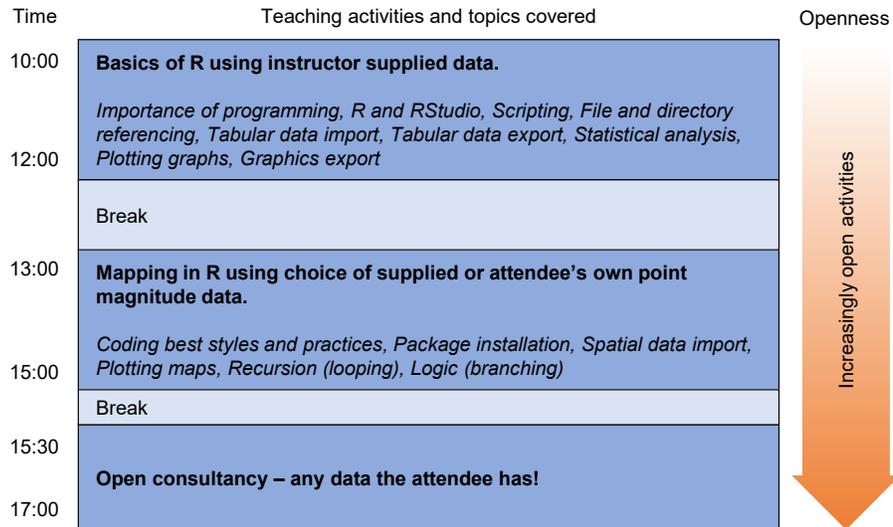


Figure 4: The design of the one-day introductory computer programming workshop for non-programming geographers.

usually does not have and leaves a person more frustrated at the end than when they started.”

- “Learning computer-programming languages has been a major challenge that is dull and painful for me. This understanding started from attending a compulsory computer class on Visual Basic in secondary school. The early frustration deterred me from attending further computer science or programming courses in university.”

The results of the pre-workshop survey indicated that there is at least some section of the geographical community that is not getting appropriate computer programming training as part of their formal education. This may well be changing but these comments indicate there will always be a section of the geographical community that would benefit from this kind of one-day introductory workshop.

4.2 Are in-person workshops necessary?

- “The course run at RGS was ideally designed to provide access to tools, datasets, and knowledgeable instructors removed from the distractions of the typical office environment.”
- “One of the basic mistakes I had is the misspelling of longitude into longtitude...”
- “Having struggled to teach myself languages such as Python in the past with the help of online training courses such as Coursera, I was very grateful for the excellent course structure and the patient and helpful advice given by the three extremely knowledgeable demonstrators.”

There are obviously logistical complications and costs associated with running a workshop in person, so given the proliferation of freely available online material, there really does need to be a strong benefit to organizing a workshop. I think these comments highlight that learning in person helps people to efficiently get over the initial learning obstacles that can otherwise cause so much frustration (e.g., spotting simple errors such as typos) and commits you to focusing on the task at hand.

4.3 Can a one-day workshop really have an impact?

- "...for a one day course I felt I learned enough to begin to read and navigate the software, which for me is the most daunting aspect."
- "...I was able to, with the aid of three very helpful demonstrators write my very first script and perform so rudimental statistical analysis in R. These small but important successes were very pleasing and helped my believe that I was capable of carrying out tasks in computer programming..."
- "The R workshop was a brilliant introduction into programming and has given me some confidence that I will be capable of using R for my data analysis in my thesis"
- "The training has made aware of the importance of acquiring programming skills as a researcher and motivated me to continue learning to program."
- "This course provided me with a starting point for how to set up automated tasks using R, which I am certain will help to improve my productivity over the remaining duration of my studies."

Clearly a one-day workshop is never going to be able to cover all the material that any one of the participants will ever need. So, I was careful to emphasize that the goal of the workshop was to demystify programming enough for people so that they felt more confident and were better equipped to learn themselves after the workshop. It is hard to say whether making this point clear at the beginning of the workshop helped to set expectations for the participants, but certainly these comments would indicate that just one day of concentrated supportive learning can really make a difference to people's confidence and motivation by getting them past the initial uncertainties and apprehensions.

4.4 Does incorporating domain relevance have learning value?

- "I had only once previously spent a short amount of time trying to learn Python but the examples and data that were used were completely unrelated to geography and my work, and therefore I struggled to know if the results coming out were correct and the relationships that were shown in the data were real. By using geographical data, although not related to what my thesis is about, I could understand what was going on, why we were inputting what we did and I could understand the outputs."
- "As a geography student, I particularly enjoyed the aspect of the course which focused on spatial data."
- "Over the years there has been an increase in resources, particularly online, freely available. I have done a couple of these for R but these were often so basic that I was left that one-step too far from the task I needed to accomplish. It is rare to find start-to-finish detail in a specific task in the various online resources available."

Domain relevance was a theme that was highlighted during the pre-workshop survey, so it is perhaps no surprise that designing the workshop to support geographers specifically was appreciated. In doing so, this did mean I had to spend time adapting the workshop material I had developed previously to make it more directly relevant to geographers. Given the comments, I would conclude that this additional effort is worthwhile, and would encourage others to develop workshop materials with this in mind. Much of what needs to be taught in an introductory programming workshop would apply across any domain, so if you develop your materials in a modular fashion, it should not be too much extra work to then adapt the workshop to incorporate examples and data that will be of greater relevance to the target audience.

4.5 Does making the workshop more open have learning value?

- “At the end of the session we had the chance to use our own data but have some help from the team which I found invaluable and it has given me confidence to go forward and learn to use R.”
- “One thing I found helpful was not only working through the examples provided but also then having the time to modify those examples with criteria I typically find in my own work...”
- “The workshop allowed me to not only perform tasks with the supplied data, but also carry out some basic mapping with my own data. To my surprise, I found R to be very enjoyable, and can see many opportunities to use these skills in the future.”
- “I enjoyed that between them [the instructors], they had different ways to do things and it was totally to your preference as to how best achieve the desired outcome and this was a reflection of how R works and its capabilities.”
- “I really appreciated the various approaches of doing some of the basic code writing that the three instructors brought to the day and knowing that there is more than one way to do things.”

There is a definite balance to be struck here. To avoid complete chaos there does need to be some consistency about what people are doing and what data is being used. I tried to build in some openness by stressing that there are multiple ways of achieving the same result, and that there was no single solution to any of the exercises. This was quite helpful as inevitably the instructors proposed different solutions to different people which could be construed as confusing by the participants. I also started the day with everyone using the same data in the first session, then built in a choice of data in the second session. Finally, I ended the day with some time to breakout into smaller groups to discuss and experiment around common topics that were brainstormed over lunch. This seemed to work quite well, and the comments indicate that it was certainly appreciated by some. But with three instructors for 11 participants, I did have the kind of instructor-participant ratio that would be needed to support such an activity.

4.6 How can participants be better supported beyond the workshop?

- “I would be interested if there are future workshops with more extensive content, such as image processing or spatial modelling, such as species distribution models.”

- “It would have been great to have a more extensive course, potentially carried out over a couple of days, of even multiple single day events throughout the year aimed at different levels, allowing attendant to continually improve their skills.”
- “team working can be useful for design and creating programming for larger set of analysis and application development. working in a team can help cross check and verify each other’s common mistakes, and discussing the applicability and effectiveness with the target customers/ researchers. So instead of working on one’s own, it would be valuable to form a team who shares similar levels of knowledge and skills of the subject area, shares the same goal/ vision/ mission, but might varied in advancement of R skill.”
- “I have heard it often suggested that colleagues are a good resource for learning to program code. In my experience this is not always the case. Everyone is busy and while your colleagues may have the knowledge to accomplish a task, it does take more time and effort to teach this in a meaningful way so that you retain that knowledge.”

At the end of a workshop all I could think to do to try and support participants continued learning was to give pointers to some potentially useful online material. But feedback from the participants suggests there is a desire for more detailed and geographically specific workshops. This is perhaps no surprise, but there clearly is a challenge for individuals in trying to develop a whole series of workshops. Efforts such as the Data Carpentry lessons could help with this, by providing a place for lessons that are developed and maintained by a community to be hosted and shared from a central location. I would have hoped that colleagues in home institutions would be able to provide some support, and as one person commented, working in a team is a good way to learn to code. But another comment does highlight the challenges of relying on already busy colleagues. It might be that encouraging organizations to set up code clubs may be a good way to nurture sustainable software skills.

5 Conclusion

In summary, I think the responses to my survey and feedback from the workshop emphasize the need and importance of relevant domain-specific training, and that it is not possible to have a single workshop that will fit the need of every researcher. Therefore, I think the efforts of projects like Data Carpentry that aim to provide domain-specific training are very much needed. As for how this information from the non-programming geographer’s perspective could be used to design better workshops for teaching geographers’ introductory computer programming, I would advise the following:

1. Use real geographical data and examples to maximise relevance,
2. Incorporate uniquely geographical tasks such as map making and spatial analyses,
3. Use R or Python software with reliable and well-supported packages.

While many of the “best practices” for scientific computing [12] will be out of reach of an introductory workshop targeted at non-programming geographers, it should be possible to teach some “good enough practices” [13] that will have an immediate and significant impact on people’s ability to leverage the power of computer programming for geographical research. Therefore, I would encourage others to provide such workshops, but also advise

those that are willing to provide such workshops that there is a wealth of experience that can be drawn upon to try and make the workshops succeed [11].

Thomas R. Etherington
Manaaki Whenua—Landcare Research, New Zealand

Acknowledgments

This work was supported by funding from a Software Sustainability Institute Fellowship, and with logistical support from the Royal Geographical Society (with the Institute of British Geographers).

References

- [1] BAKER, M. Code alert. *Nature* 541 (2017), 563–565. doi:10.1038/nj7638-563a.
- [2] BOWLICK, F. J., GOLDBERG, D. W., AND BEDNARZ, S. W. Computer science and programming courses in geography departments in the United States. *The Professional Geographer* 69, 1 (2017), 138–150. doi:10.1080/00330124.2016.1184984.
- [3] CROUCH, S., CHUE HONG, N., HETRICK, S., JACKSON, M., PAWLIK, A., SUFI, S., CARR, L., DE ROURE, D., GOBLE, C., AND PARSONS, M. The Software Sustainability Institute: changing research software attitudes and practices. *Computing in Science & Engineering* 15, 6 (2013), 74–80. doi:10.1109/MCSE.2013.133.
- [4] ETHERINGTON, T. R. Teaching introductory GIS programming to geographers using an open source Python approach. *Journal of Geography in Higher Education* 40, 1 (2016), 117–130. doi:10.1080/03098265.2015.1086981.
- [5] MERALI, Z. Error . . . why scientific programming does not compute. *Nature* 467 (2010), 775–777. doi:10.1038/467775a.
- [6] R CORE TEAM. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [7] TEAL, T. K., CRANSTON, K. A., LAPP, H., WHITE, E., WILSON, G., RAM, K., AND PAWLIK, A. Data Carpentry: workshops to increase data literacy for researchers. *International Journal of Digital Curation* 10, 1 (2015), 135–143. doi:10.2218/ijdc.v10i1.351.
- [8] UNWIN, D. Make your practicals open ended. *Journal of Geography in Higher Education* 4, 2 (1980), 39–42. doi:10.1080/03098268008708772.
- [9] WIKLE, T. A., AND FAGIN, T. D. GIS course planning: a comparison of syllabi at US college and universities. *Transactions in GIS* 18, 4 (2014), 574–585. doi:10.1111/tgis.12048.

- [10] WILSON, G. Software Carpentry: getting scientists to write better code by making them more productive. *Computing in Science & Engineering* 8, 6 (2006), 66–69. doi:10.1109/MCSE.2006.122.
- [11] WILSON, G. Software Carpentry: lessons learned [version 2; referees: 3 approved]. *F1000Research* 3, 62 (2016). doi:10.12688/f1000research.3-62.v2.
- [12] WILSON, G., ARULIAH, D. A., BROWN, C. T., CHUE HONG, N. P., DAVIS, M., GUY, R. T., HADDOCK, S. H. D., HUFF, K. D., MITCHELL, I. M., PLUMBLEY, M. D., WAUGH, B., WHITE, E. P., AND WILSON, P. Best practices for scientific computing. *PLoS Biology* 12, 1 (2014), e1001745. doi:10.1371/journal.pbio.1001745.
- [13] WILSON, G., BRYAN, J., CRANSTON, K., KITZES, J., NEDERBRAGT, L., AND TEAL, T. K. Good enough practices in scientific computing. *PLoS Computational Biology* 13, 6 (2017), e1005510. doi:10.1371/journal.pcbi.1005510.