The role of reproducibility in scientific map making

Eftychia Koukouraki

1 Introduction

The reproducibility of research results is a critical aspect across all scientific disciplines that involve computations, and the domain of Geosciences is no exception. It is widely accepted by the scientific community, that the ability to reproduce results that were published by other working groups enhances the trust and the reliability of the respective research. In an effort to clarify the different existing terminologies for *reproducibility* and *reproducible research*, Barba (2018) defines reproducible research as the case when "authors provide all the necessary data and the computer codes to run the analysis again, re-creating the results". On the other hand, scientific *replication* can be achieved when a study acquires the same results using different methods or different data. This distinction is also known as the Claerbout (Claerbout and Karrenbach, 1992)/Donoho (Donoho et al., 2009)/Peng (Peng, 2011) convention.

In the domain of Geosciences, the relevant work expands to many different directions. In the *Forum on Reproducibility and Replicability in Geography* introduced by Goodchild et al. (2021), the discussed topics range from the theoretical dimension of reproducibility and replicability in geography (Sui and Kedron, 2021), to more tangible matters, such as the review of current technological solutions (Nüst and Pebesma, 2021) and the functionality of software in the context of reproducibility (Wilson et al., 2021). Konkol and Kray (2019) focused on the role of figures as a means of communicating research results and conducted a survey to identify the nature of the incorporated figures (eg. maps, time series, histograms, etc.), as well as the frequency that they are used in geoscientific papers. The survey showed that "maps" were the most frequent among the answers, which seems reasonable for this domain.

Maps cover a wide range of applications, from advertising handouts for tourist navigation to sophisticated cartographic products that analyse complex geographic phenomena. The digital turn in society during the last decades, has enabled the use of maps to become an important part of our daily lives. Nowadays, map making has shifted from being a privilege of trained experts to an opportunity for everyone with access to data and the necessary amount of time (Kraak and Fabrikant, 2017). Moreover, from a user-oriented perspective, maps as a means of transmitting information, can be used by people without prior relevant specialization. Although there are several definitions for what a map is, most of them agree that maps are abstract representations of the real world (Kraak and Fabrikant, 2017), in which certain attributes of the reality are highlighted, while others are completely muffled. The ability to rely on the information that is communicated with maps becomes even more critical nowadays, especially when it comes to analysing geospatial phenomena, such as epidemiological spread or flood risk. Transparency is thus of utmost importance in order to produce trustworthy scientific outcomes and draw reliable conclusions. The capacity to reproduce results builds exactly this trust (Heroux et al., 2018), and this becomes even more important nowadays in the context of climate change.

Results can be purposefully controlled to suit one's interests and maps are a fundamental medium to communicate results. The complexity and significance of the environmental topics analysed and presented, make transparency in map creation imperative. As reproducibility sheds light on how the output was created, it makes only sense to examine its role in the current practices of map making and to suggest ways to facilitate and improve it. In accordance with the agenda for cartographic research that was suggested in Griffin et al. (2017), the proposed work intends to make use of new research methods and technological tools that are emerging, aspiring to highlight the need for map reproduction in the context of environmental studies.

2 Literature Review

The International Cartographic Association (ICA) defines Cartography as the science, art, and technology of map making and map use (Kraak, 2019). In Taylor (1991), cartography in the context of Geographic Information Systems (GIS), is mentioned as "the organization, presentation, communication and utilization of geo-information in graphic, digital or tactile form. It can include all stages from data preparation to end use in the creation of maps and related spatial information products". It is important to note that this definition takes into account the technological aspect in the cartographic process and highlights the importance of it in the creation of spatial information products beyond maps.

The current map making practices in the academia use different software products for different tasks, from spreadsheets to specialized statistical software, resulting in slicing the cartographic process into multiple steps (Giraud and Lambert, 2017), which are not always connected in a straightforward way. These steps include data management, statistical analyses, geoprocessing and graphical display of the results. The delineation of them is necessary, especially in a scientific context, and each one of them needs to be validated for its objectivity (Giraud and Lambert, 2019).

Even though reproducibility is a topic of rising concern for the scientific community, the term "reproducible cartography" is not frequently encountered in the literature. In Giraud and Lambert (2017), reproducible maps are placed within a spectrum, ranging from non-reproducible, when they come as a simple print-out, to fully reproducible, when they are clearly linked with executable code, data and metadata. Although many suggestions have been made regarding the enhancement of map reproducibility in different contexts (terrain representation (Kennelly et al., 2021), knowledge graphs (Mai et al., 2022), etc.), terms as "reproducible cartography" and "map reproducibility" remain officially undefined and rely on the individual intuitive interpretation.

Researchers have tackled the issue of reproducibility in the Geosciences support that open data and methods and open source software are prerequisites for achieving the full potential in transparency (Nüst et al., 2017). However, Ledermann and Gartner (2021) argue that acquiring the source code of an experiment is not enough to reproduce it and rather argue that a well defined and clearly structured programming workflow in an ontological fashion contributes to the transparency, reproducibility and extensibility of scientific experiments. Giraud and Lambert (2017) encourages the use of literate programming reports, e.g. Jupyter notebooks or R Markdown, because such programming solutions provide complete instructions from raw data to the cartographic product. In Giraud and Lambert (2019) an example of reproducible cartographic workflow is implemented, combining different geospatial R packages in a single R Markdown script. On the other other hand, Knoth and Nüst (2017) leverages containers for describing the computational environment of the experiment and for facilitating the reuse of the workflow by others. Kedron et al. (2021) suggests that geospatial data repositories as a means of data sharing should be evaluated regarding their contents, in order to define their usefulness in the reproduction and replication of scientific studies. In more complex visualisation cases, such as terrain representation, Kennelly et al. (2021) commented on the usefulness 3D models that can be used as reference for researchers in order to evaluate and compare reproduced visualizations and stated that in the field of cartography such data models do not exist.

Reproducing and replicating visualizations is not an easy task. Fekete and Freire (2020) impute potential obstacles to the involvement of High Performance Computing and specialized displays, such as Virtual Reality and video walls. Moreover, they consider interactive features as an additional issue for scientific replication, but emphasize on the importance of interactivity in data exploration. A data-centric approach for reproducible visualizations has been proposed by Silva et al. (2007). The authors foreground the role of a well-defined data flow pipeline, stating that despite the plethora of the tools and mechanisms to describe, the purpose of the visualization process is always to gain insights from the data.

3 Aim, Research Questions (RQs) and Objectives

As seen in the respective definition of ICA mentioned in Section 2, cartography as a discipline has also an artistic connotation. For this reason, the present work favours the terms ”map production” and ”map making” in the context of reproducibility.

Based on the definition of Barba (2018) for reproducible research outlined in Section 1 and on the definition of Taylor (1991) for cartography referenced in Section 2, we interpret ”reproducible map production” and ”reproducible map making” as the organization, processing and presentation of all the materials involved in the cartographic process, in such way, that the map as a cartographic product can be recreated in an independent experiment. Although the a *map* can refer to a variety of representational artefacts, for the needs of this doctoral work the focus will be set on maps with an apparent geographical reference.

The aim of the research is to systematically investigate the role of reproducibility in maps that are encountered in scientific publications and to analyse the complexities of map reproduction. Reproducibility of scientific results is not monolithic. By rerunning an experimental workflow, important parameters can change unintentionally. These parameters range from the computational environment to the involved actors and the research goal itself (Fekete and Freire, 2020). When altering these parameters, new insights are revealed, so it is important to observe what information is acquired when reproducing a study (Freire et al., 2016).

Consequently, the assessment of a reproduction should not be binary. In the context of maps, even when two maps (the original and the reproduced) are not identical, they still manage to convey the same message. As reproducibility assessment contains inherently the concept of comparison, we argue that a simple image subtraction may be too simplistic for this purpose. To explore the implications of the aforementioned argument, we seek the answer to the following Research Questions (RQs):

1. What is the current practice of creating maps for scientific publications regarding which tools are used, what data is available and how reproducible the published maps are?

2. In some cases, the reproduced map can differ from the original only in the positioning of the labels. In such cases, how is map reproduction perceived and what is a suitable conceptual model to describe it?

3. How can reproducibility in the context of map making be assessed, so that the particularities of map reproduction are taken into account?

To meet the requirements of the RQs, the subsequent objectives have been defined:

1. Investigate the latest literature of environmental studies regarding the practices (eg. coding, Graphic User Interface) and the tools (eg. libraries, plugins) that are used to produce maps.

2. Identify possible issues or obstacles when reproducing maps of scientific publications.

3. Develop guidelines for facilitating reproducibility in map making.

4. Implement a tool for systematic map comparison.

4 Work Packages (WPs) and Evaluation

4.1 WP 1: Investigate the current practices

This Work Package (WP) covers the Objectives 1 & 2. A reproduction study will be carried out to assess the status quo of the map making process in scientific publications. The study will be based on recent and open access papers, to reduce the risk of unavailable data and/or deprecated software. *Copernicus Publications* meets the open access criterion and offers a good starting point for this search, as most of the journals are of geoscientific interest.

For a paper reproduction, it is necessary that all the materials are available for reuse. Nüst et al. (2018) categorized the availability of materials as following: 1) Input data, 2) Methods and 3) Results. The methods were further divided into 1) pre-processing, 2) method/analysis/processing, and 3) computational environment. After the availability check, the next step is the reproduction attempt. Konkol et al. (2019) developed a coding scheme for categorizing the reproduction issues according to how serious they are.

The aforementioned coding schemes will consist the base for evaluating how effective are the map making practices in scientific publications regarding reproducibility, but are subject to modification to reflect the particularities of map making. A log of the reproducibility attempts will be kept, to reveal the tendencies in map making regarding the preferred coding languages or GUIs, the amount of time that is required to reproduce a map, the nature of the difficulties that arise during the process, etc.

By completing this Work Package, we expect to answer RQ1 and to gain a first impression for RQ2. The findings of this step will be used as input to the other WPs.

4.2 WP 2: Conceptualisation of map reproducibility

This WP intends to fulfill Objective 3. Based on the insights acquired by WP1 regarding the map making practices and impact they have on the reproducibility state, we will run at least one focus group discussion and one round of user interviews (Lazar et al., 2017).

4.2.1 Focus group

With the focus group, we will approach stakeholders with experience in cartography and/or geoscientific reproducible research to receive feedback on the qualitative coding scheme derived in WP1. Furthermore, we will attempt to formalise the concept of reproducibility in the cartographic process and to derive a first draft of guidelines that help to overcome the impediments in map reproduction, that were identified during the elaboration of WP1. Reproducibility assessment is mostly elaborated by comparing two independently created results in a side-to-side fashion, so that it is determined if they are identical or not. Based on preliminary experiments that we have run for the needs of Opening Reproducible Research (O2R) project, we argue that subtraction as a comparison mechanism can be superficial when comparing figures in general, and maps, in particular. A more systematic evaluation and a detailed description regarding the nature of the differences identified during a map comparison will be discussed, to decide the nature of a descriptive framework for quantifying the success level of map reproduction.

4.2.2 User interviews

In the user interviews, the aforementioned conceptual model and the prototype that is built for WP3 will be presented to researchers and professionals from the domain of Geosciences. The participants will be asked for their opinion regarding the conceptualisation of map reproducibility, as well as usefulness of the prototype. In this manner, the prototype validates the practical applicability of the conceptualisation that was derived by the focus group and the user interviews evaluate the prototype.

This Work Package will provide evaluation for WP1 and both feedback and input for WP3. This step will help to answer RQ2 and will provide insights for RQ3.

4.3 WP 3: Development of a prototype

This Work Package addresses Objective 4. For the needs of this WP, a prototype that realises map comparison will be developed. The details of the implementation depend on the outcomes of the other two Work Packages.

It is expected that the majority of the maps encountered in scientific publications will be essentially static figures embedded in a PDF document. Consequently, we will use computer vision algorithms (Szeliski, 2022) for detecting features on maps. Reviewing the related literature and the state-of-the-art, as well as preliminary testing of different algorithms and relevant Python libraries (eg. MapReader (Hosseini et al., 2021)), are also part of this WP.

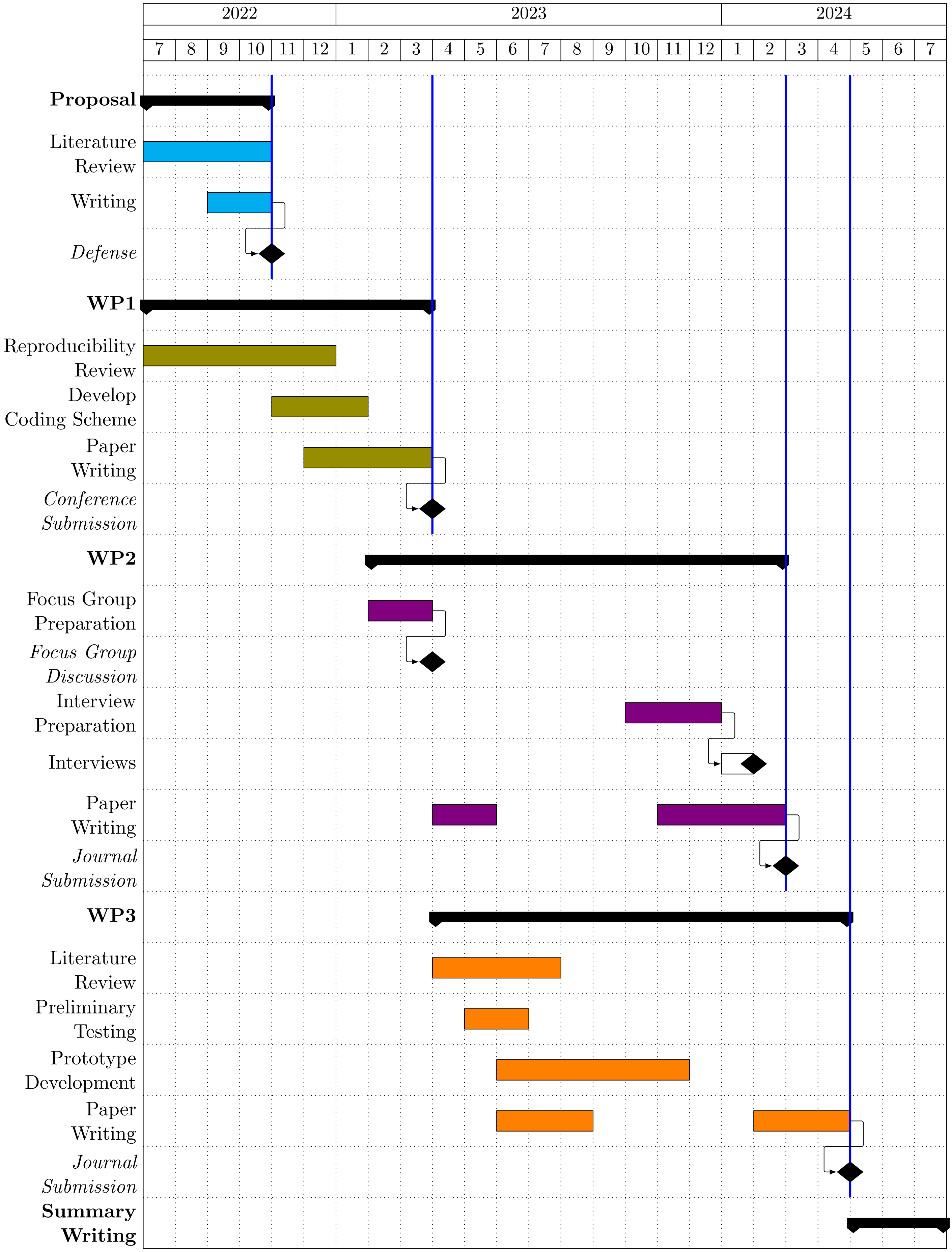
The next steps in the implementation of the prototype are the comparison of the two maps and the integration of the reproducibility evaluation scheme derived in WP2. The computed output of the prototype will be the result of the map comparison, described as outlined in the aforementioned scheme.

The evaluation of the prototype consists of two parts: 1) the efficiency of the feature detecting model will be assessed with the accuracy metrics that will be defined during the phase of literature review for computer vision algorithms, and 2) the usability of the prototype will be evaluated during the round of the user interviews mentioned in WP2.

This Work Package will provide feedback for WP2 and and will help to answer RQ3.

5 Expected Results

With the successful completion of the proposed research, we expect that the tendencies in scientific map making regarding reproducibility will be revealed. A conceptual framework that describes map reproduction and eventually quantifies the differences and similarities in map comparison will be derived. It is expected that this framework will address differences in a geometrical, topological and color-wise level, among others. Based on the estimation that maps encountered in the literature will be mostly static, a tool that facilitates map comparison in a descriptive way will be developed, anticipating that will minimize the need for human intervention in the process. The proposed research tackles an issue that is not of purely scientific or technological nature, but also has a strong artistic dimension. Finally, we expect to contribute to the change of culture in the scientific practices regarding the transparency of the results, by enriching the cartographic process to reflect reproducibility.


2022
2023
2024
7
8
9
10
11
12
1
2
3
4
5
6
7
8
9
10
11
12
1
2
3
4
5
6
7
Proposal
Literature Review
Writing
Defense
WP1
Reproducibility Review
Develop Coding Scheme
Paper Writing
Conference Submission
WP2
Focus Group Preparation
Focus Group Discussion
Interview Preparation
Interviews
Paper Writing
Journal Submission
WP3
Literature Review
Preliminary Testing
Prototype Development
Paper Writing
Journal Submission
Summary Writing

References

Barba, L. A. (2018). Terminologies for Reproducible Research. arXiv:1802.03311 [cs].

Claerbout, J. F. and Karrenbach, M. (1992). Electronic documents give reproducible research a new meaning. In *SEG Technical Program Expanded Abstracts 1992*, SEG Technical Program Expanded Abstracts, pages 601–604. Society of Exploration Geophysicists.

Donoho, D. L., Maleki, A., Rahman, I. U., Shahram, M., and Stodden, V. (2009). Reproducible Research in Computational Harmonic Analysis. *Computing in Science & Engineering*, 11(1):8–18. Conference Name: Computing in Science & Engineering.

Fekete, J.-D. and Freire, J. (2020). Exploring Reproducibility in Visualization. *IEEE Computer Graphics and Applications*, 40(5):108–119.

Freire, J., Fuhr, N., and Rauber, A. (2016). Reproducibility of Data-Oriented Experiments in e-Science (Dagstuhl Seminar 16041). *Dagstuhl Reports*, 6(1):108–159. Place: Dagstuhl, Germany Publisher: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Giraud, T. and Lambert, N. (2017). Reproducible Cartography. In Peterson, M. P., editor, *Advances in Cartography and GIScience*, Lecture Notes in Geoinformation and Cartography, pages 173–183, Cham. Springer International Publishing.

Giraud, T. and Lambert, N. (2019). Reproducible Workflow for Cartography – Migrants Deaths in the Mediterranean. *Proceedings of the ICA*, 2:1–7.

Goodchild, M. F., Fotheringham, A. S., Kedron, P., and Li, W. (2021). Introduction: Forum on Reproducibility and Replicability in Geography. *Annals of the American Association of Geographers*, 111(5):1271–1274. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/24694452.2020.1806030.

Griffin, A. L., Robinson, A. C., and Roth, R. E. (2017). Envisioning the future of cartographic research. *International Journal of Cartography*, 3(sup1):1–8. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23729333.2017.1316466.

Heroux, M. A., Barba, L., Parashar, M., Stodden, V., and Taufer, M. (2018). Toward a Compatible Reproducibility Taxonomy for Computational and Computing Sciences. Technical Report SAND2018-11186, Sandia National Lab. (SNL-NM), Albuquerque, NM (United States).

Hosseini, K., Wilson, D. C. S., Beelen, K., and McDonough, K. (2021). MapReader: A Computer Vision Pipeline for the Semantic Exploration of Maps at Scale. arXiv:2111.15592 [cs].

Kedron, P., Li, W., Fotheringham, S., and Goodchild, M. (2021). Reproducibility and replicability: opportunities and challenges for geospatial research. *International Journal of Geographical Information Science*, 35(3):427–445.

Kennelly, P. J., Patterson, T., Jenny, B., Huffman, D. P., Marston, B. E., Bell, S., and Tait, A. M. (2021). Elevation models for reproducible evaluation of terrain representation. *Cartography and Geographic Information Science*, 48(1):63–77. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15230406.2020.1830856.

Knoth, C. and Nüst, D. (2017). Reproducibility and Practical Adoption of GEOBIA with Open-Source Software in Docker Containers. *Remote Sensing*, 9(3):290. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

Konkol, M. and Kray, C. (2019). In-depth examination of spatiotemporal figures in open reproducible research. *Cartography and Geographic Information Science*, 46(5):412–427. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15230406.2018.1512421.

Konkol, M., Kray, C., and Pfeiffer, M. (2019). Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study. *International Journal of Geographical Information Science*, 33(2):408–429. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/13658816.2018.1508687.

Kraak, M.-J. (2019). Strategic Plan for 2019–2027. Technical report, International Cartographic Association, Enschede.

Kraak, M.-J. and Fabrikant, S. I. (2017). Of maps, cartography and the geography of the International Cartographic Association. *International Journal of Cartography*, 3(sup1):9–31. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23729333.2017.1288535.

Lazar, J., Feng, J. H., and Hochheiser, H. (2017). Chapter 8 - Interviews and focus groups. In Lazar, J., Feng, J. H., and Hochheiser, H., editors, *Research Methods in Human Computer Interaction (Second Edition)*, pages 187–228. Morgan Kaufmann, Boston.

Ledermann, F. and Gartner, G. (2021). Towards Conducting Reproducible Distributed Experiments in the Geosciences. *AGILE: GIScience Series*, 2:1–7.

Mai, G., Huang, W., Cai, L., Zhu, R., and Lao, N. (2022). Narrative Cartography with Knowledge Graphs. *Journal of Geovisualization and Spatial Analysis*, 6(1):4.

Nüst, D., Granell, C., Hofer, B., Konkol, M., Ostermann, F. O., Sileryte, R., and Cerutti, V. (2018). Reproducible research and GIScience: an evaluation using AGILE conference papers. *PeerJ*, 6:e5072. Publisher: PeerJ Inc.

Nüst, D., Konkol, M., Pebesma, E., Kray, C., Schutzeichel, M., Przibytzin, H., and Lorenz, J. (2017). Opening the Publication Process with Executable Research Compendia. *D-Lib Magazine*, 23.

Nüst, D. and Pebesma, E. (2021). Practical Reproducibility in Geography and Geosciences. *Annals of the American Association of Geographers*, 111(5):1300–1310.

Peng, R. D. (2011). Reproducible Research in Computational Science. *Science*, 334(6060):1226–1227. Publisher: American Association for the Advancement of Science.

Silva, C. T., Freire, J., and Callahan, S. P. (2007). Provenance for Visualizations: Reproducibility and Beyond. *Computing in Science & Engineering*, 9(5):82–89.

Sui, D. and Kedron, P. (2021). Reproducibility and Replicability in the Context of the Contested Identities of Geography. *Annals of the American Association of Geographers*, 111(5):1275–1283. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/24694452.2020.1806024.

Szeliski, R. (2022). *Computer Vision: Algorithms and Applications*. Springer Nature. Google-Books-ID: QptXEAAAQBAJ.

Taylor, D. R. F. (1991). CHAPTER 1 - Geographic Information Systems: The Microcomputer and Modern Cartography. In Taylor, F., editor, *Modern Cartography Series*, volume 1 of *Geographic Information Systems*, pages 1–20. Academic Press.

Wilson, J. P., Butler, K., Gao, S., Hu, Y., Li, W., and Wright, D. J. (2021). A Five-Star Guide for Achieving Replicability and Reproducibility When Working with GIS Software and Algorithms. *Annals of the American Association of Geographers*, 111(5):1311–1317. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/24694452.2020.1806026.