



Overcoming barriers to reproducibility in geoscientific data analysis: Challenges and practical implementation strategies

Matthias Schlögl^{1,2}, Laura Waltersdorfer^{3,4}, Peter Regner⁵, Andrea Siposova⁶, and Alexander Brenning⁷

¹Department of Landscape, Water and Infrastructure, BOKU University, Peter-Jordan Straße 82, 1190 Vienna, Austria

²Department for Climate Impact Research, GeoSphere Austria, Hohe Warte 38, 1190 Vienna, Austria

³Institute of Information Systems Engineering, Technical University of Vienna, Karlsplatz 13, 1040 Vienna, Austria

⁴Institute for Data, Process and Knowledge Management, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria

⁵Institute of Sustainable Economic Development, BOKU University, Feistmantelstraße 4, 1180 Vienna, Austria

⁶Research Group Machine Learning and Data Management, SBA Research gGmbH, Floragasse 7, 1040 Vienna, Austria

⁷Department of Geography, Friedrich Schiller University Jena, Löbdergraben 32, 07743 Jena, Germany

Correspondence: Matthias Schlögl (matthias.schloegl@boku.ac.at)

Abstract. Reproducibility is a cornerstone of the scientific method, yet it remains elusive in many domains of contemporary research, including the geosciences. In this Perspective, we examine reproducibility in the context of computational workflows for geospatial analyses. Building on established frameworks, we disentangle key dimensions—methodological, results, and inferential reproducibility—and identify critical barriers, including irreproducibility arising from obscurity, obfuscation, and uncontrollable conditions. We argue that enhancing reproducibility in geoscientific research requires both cultural transformation and practical, domain-specific interventions. Our focus lies on methodological and computational reproducibility, with particular attention to challenges posed by spatial data structures, diverse data sources and infrastructures, and the integration of statistical and machine-learning methods. We outline actionable guidance across the research workflow, including data governance, analysis design and documentation, code development, and long-term accessibility. Emphasis is placed on the use of open-source software, script-based automation, version control, and the adoption of FAIR principles to support findability, accessibility, interoperability, and reusability of data and code. Our aim is to provide a structured synthesis that supports reproducible and transparent geospatial research. By implementing even incremental improvements, researchers can strengthen the robustness, transparency, and reuse potential of their scientific contributions.

1 Introduction

The reproducibility crisis in science (Ioannidis, 2005) has sparked extensive debate on the reliability of research findings, leading to what some describe as a credibility revolution (Korbmacher et al., 2023; Angrist and Pischke, 2010). Reproducibility—the ability to independently verify research findings using the same data and methods—is a fundamental principle of the scientific method and a cornerstone of empirical inquiry (Nichols et al., 2021). However, scientific practice does not always adhere to this ideal (Ioannidis, 2012; Simonsohn et al., 2014; Ivie and Thain, 2018; Marshall-Cook and Farley, 2024). Researchers frequently struggle to reproduce not only the work of others but also their own findings (Baker, 2016; Schwab et al.,



2000), raising concerns that a substantial fraction of published results may be unreliable or merely reflect prevailing biases (Ioannidis, 2005).

Scientific reproducibility is often framed as a technical problem requiring standardized methodologies, transparent workflows, and open data. However, sociology of science perspectives emphasize deeper social and institutional challenges, including structural factors such as incentive systems that prioritize novelty over verification, lack of dedicated funding and infrastructure, tacit knowledge that is difficult to formalize, and barriers to data accessibility (Joelving, 2024; Van Noorden, 2023; Nichols et al., 2021; Wasserstein et al., 2019; Fanelli, 2019; Goodman et al., 2016; Gewin, 2016; Ioannidis, 2014; Nosek et al., 2012). These challenges are particularly pronounced in data-intensive fields, where vast, heterogeneous datasets and complex models further complicate reproducibility efforts.

The notion of ‘research reproducibility’ is a multifaceted concept with inconsistent understanding across disciplines, shaped by vantage point, context, and idiosyncrasies of each scientific domain (Korbmacher et al., 2023; Peng and Hicks, 2021; National Academies of Sciences, 2019; Konkol et al., 2018; Benureau and Rougier, 2018; Plessner, 2018; Barba, 2018; Stark, 2018; Nature, 2016; Goodman et al., 2016; Leek and Peng, 2015). Goodman et al. (2016) advocated moving beyond debates over technical terminology towards adopting concise descriptors that capture the essence of the underlying constructs. They introduced definitions for the core dimensions of ‘methodological reproducibility’, ‘results reproducibility’ and ‘inferential reproducibility’ (Tab. 1).

Table 1. Conceptions of reproducibility based on descriptors of the underlying constructs after Goodman et al. (2016)

Type	Definition
Methodological reproducibility	Enabling the exact repetition of study procedures by providing sufficient detail about study methods, code and data. This notably comprises ‘computational reproducibility’.
Results reproducibility	Obtaining corroborating results from an independent study, whose procedures closely match the original experiment, but which uses its own data. This is often referred to as ‘replicability’ (National Academies of Sciences, 2019).
Inferential reproducibility	Drawing qualitatively similar conclusions from either an independent replication of a study or a reanalysis of the original study.

Across these perspectives, reproducibility can be regarded as a spectrum (Peng, 2011), ranging from essentially irreproducible research to the gold standard of full replicability. In the context of computational reproducibility this spectrum comprises access to raw data, code and a defined execution environment as core components.

With growing awareness of open science (Nüst et al., 2018) and advances in scientific computing (Peng and Hicks, 2021), reproducible research has gained prominence. As early as 1992, Claerbout and Karrenbach (1992) argued that publications merely advertise research, while the true essence lies in the underlying computer programs, input data, parameter values, and other related components. Since then, extensive guidelines have emerged to enhance computational reproducibility: simple rules (Sandve et al., 2013; Goodman et al., 2014; Osborne et al., 2014; List et al., 2017; Hunter-Zinck et al., 2021), best practices



45 (Wilson et al., 2014) and ‘good enough’ practices (Wilson et al., 2017) for scientific computing provided recommendations on techniques and tools to support computing workflows in a scientific context. Principles for systematic and reproducible workflows in academic data analysis have been outlined (Stoudt et al., 2021), experiences on open sharing and reuse examined (Chen et al., 2019), and guidelines for data stewardship (Wilkinson et al., 2016) and open-source software proposed (Hunter-Zinck et al., 2021).

50 While reproducibility has gained broad attention in the life sciences (McNutt, 2014) and psychology (Open Science Collaboration, 2015), its uptake in geography and the geosciences has been more recent (Steventon et al., 2022; Nüst and Pebesma, 2020; Kedron et al., 2020). Specifically, Konkol et al. (2018) assessed computational reproducibility in geographic research, identifying technical issues and proposing solutions. Expanding on this, Nüst and Pebesma (2020) addressed practical reproducibility, emphasizing scholarly communication, literate programming (Knuth, 1984), and research compendia (Gentleman
55 and Temple Lang, 2007). In the qualitative geographic information systems (GIS) domain, Muenchow et al. (2019) reviewed 380 papers, revealing a lack of transparency regarding software use and advocating for open-source GIS. Wilson et al. (2020) proposed a five-level framework for data and code sharing in geospatial research, echoing the ‘spectrum of reproducibility’ coined by Peng (2011). Building on broader literature, Kedron et al. (2020) reviewed the adaptation of reproducibility practices to geospatial analyses, highlighting the need for interdisciplinary collaboration to address open issues.

60 Geospatial data science, which increasingly relies on machine learning (ML) and artificial intelligence (AI) methods, faces unique reproducibility challenges. As the distributed spatial data infrastructures that are often used for geospatial data analysis workflows evolve, ensuring that geospatial analyses remain replicable over time requires careful documentation and planning. Moreover, unique geospatial data structures, standards, and computing environments require special attention. The increasing adoption of GeoAI, including large language models and spatially aware AI frameworks, introduces new reproducibility
65 challenges, particularly due to biases in training data, model uncertainty, and the inherent spatial heterogeneity of geospatial processes, as highlighted by Li et al. (2024).

While reproducibility is often discussed in academic research, it is equally critical in industry and government contexts, where geospatial analyses inform high-stakes decisions in areas such as urban planning, environmental management, disaster response, and public policy. Reproducible workflows enhance accountability, legal defensibility, and transparency, ensuring
70 that spatial analyses can be audited, verified, and trusted by stakeholders.

Building on previous work, this article advocates for improving reproducibility in geospatial analyses within the geosciences and related fields. We focus on methodological and results reproducibility, emphasizing the distinctive characteristics of geospatial data: its dual nature as geometric and attribute information, its association with spatial features (e.g., points, lines, polygons, and rasters), its dependence on location references (Pebesma and Bivand, 2023), and its frequent origination from
75 data collection process that differ from controlled experimental designs. Our goal is to synthesize existing deliberations, address emerging challenges, and provide structured guidance to make computational reproducibility in geospatial workflows more accessible to researchers with varying levels of experience in scientific programming. In doing so, we aim to raise awareness of the critical importance of reproducibility in geospatial data science, provide actionable advice to researchers and practitioners, and propose a practical checklist to facilitate reproducible geospatial workflows. The checklist, included as a supplement to



80 this article, is designed to serve as a resource for ensuring transparency, consistency, and replicability in geospatial analyses across diverse applications.

2 Adversaries of practical reproducibility

Reproducibility in research is inherently challenging and requires careful navigation of several intricate aspects along the research workflow. Albeit reproducible research is not immune to errors (Leek and Peng, 2015; Ioannidis, 2005), thorough
85 documentation and control enhance transparency and enable comprehensive evaluation. Despite ongoing endeavors to promote reproducibility, practical adherence to quality standards continues to face obstacles.

2.1 Research in general

Most of the reproducibility literature addresses fields reliant on empirical experiments, such as medicine (Pusztai et al., 2013; Landis et al., 2012), biomedical research (Munafò et al., 2017), biology (Freedman et al., 2015), or psychology (Open Science
90 Collaboration, 2015; Wicherts et al., 2006). Common challenges include cognitive biases (e.g. confirmation and hindsight bias; Xiong et al., 2020; Munafò et al., 2017; Nuzzo, 2015; Birch and Bloom, 2007; Nickerson, 1998), incomplete documentation of study design (Landis et al., 2012), statistical malpractices (underpowered statistics, p-hacking, HARKing; Bishop, 2019; Munafò et al., 2017) and systemic pressures to publish (van Wesel, 2015). These factors distort data interpretation, foster overconfidence in results, and undermine scientific rigor.

95 Dissemination biases, including publication bias, citation bias, gray literature bias, early extreme bias, media attention bias, the decline effect, and the small-study effect (Wasserstein et al., 2019; Fanelli et al., 2017; Munafò et al., 2017; Song et al., 2010), further skew the scientific record toward positive results, distorting meta-analyses and overestimating effects (Kicinski et al., 2015; Fanelli et al., 2017; Thornton, 2000; Fanelli, 2010). Additional challenges arise from limited programming expertise and lack of standardized practices (National Academies of Sciences, 2019; Mesirov, 2010). Together, these issues
100 contribute to what Feynman (1974) critically termed ‘Cargo Cult Science’—the superficial adherence to scientific rigor without its substance.

Counteracting these cognitive and structural drivers of irreproducibility requires systemic attentiveness throughout the research workflow, including study design, reporting, data sharing, and incentive structures (Munafò et al., 2017). Additionally, integrating established software engineering practices into research workflows can enhance the reliability and transparency of
105 computational methods across disciplines (Connolly et al., 2023; Arvanitou et al., 2021).

2.2 Geosciences

Despite increasing efforts to improve the availability of data and code in geoscientific research (Easterbrook, 2014), accessibility remains fragmented (Ireland et al., 2023). A survey by Baker (2016) showed that over 60 % of researchers in the Earth and Environment domain were unable to reproduce other studies, while more than 40 % even failed to reproduce their own work.



110 Reproducibility challenges in the geosciences are multifaceted (Konkol et al., 2018) and can be categorized into three key dimensions:

- **Culture.** The interdisciplinary and broad nature of geosciences leads to inconsistent standards and varying adoption of reproducible practices across domains. Technical and sociological barriers often hinder open sharing of results (Gewin, 2016; Easterbrook, 2014; Reichman et al., 2011; Barnes, 2010). Furthermore, reproducibility is often underemphasized
 115 in education and training (Tenopir et al., 2018).
- **Technology.** Geoscientific research relies heavily on geospatial data and involves a diverse community with varying levels of expertise in software development and computational methods (Nüst and Pebesma, 2020; Yan et al., 2020; Muenchow et al., 2019). Many researchers and practitioners depend on desktop GIS software, which, while effective for certain tasks, complicates documenting processing steps. Although some tasks can be transitioned to script-based
 120 workflows, others, such as manual mapping, remain inherently difficult to automate.
- **Data complexity.** Geospatial data present unique challenges due to their diversity in formats, geometry types, and collection methods (Rolf et al., 2024). They originate from a wide range of modalities and sources, including satellite imagery, aerial vehicles, GNSS, sensor networks, field surveys, digital field work, model outputs, administrative data sources, historical archives, public participation GIS and citizen science contributions (Thompson, 2016).

125 Common difficulties pertain to data quality, heterogeneous data sources, and managing large, multidimensional datasets. In contrast to controlled experiments in other fields, ensuring traditional data quality metrics such as accuracy, consistency, completeness or timeliness (Ehrlinger and Wöß, 2022) is often challenging in real-world geospatial applications. For instance, studies in hazard research showcased that inconsistencies and inaccuracies in landslide inventory data negatively impact data quality (Milledge et al., 2022; Pokharel et al., 2021). Variability in spatial resolution, scales, coordinate systems and map
 130 projections often complicate data alignment and analysis, while temporal changes and high-dimensional datasets, such as spatiotemporal arrays or data cubes, add further complexity.

These cultural, technical and data-related challenges make reproducibility in geosciences fundamentally different from controlled laboratory sciences. Overcoming these obstacles requires specialized technical skills and tools for data handling, processing, modeling and visualization, as well as community efforts to improve training to build on and adopt suitable open
 135 science principles.

2.3 Aspects of reproducibility in practice

Practical reproducibility of scientific research is shaped by several interconnected dimensions that collectively influence transparency, reliability, and validity of the underlying research. Building upon previous work, we postulate that the key dimensions affecting practical reproducibility are:

- 140 – **Data governance,** concerning the management and publication of data in terms of, e.g., availability, data formats, licenses and quality.



- **Methodological approach**, defining the selection and correct application of appropriate methodologies for solving the research questions under consideration.
- **Technical or computational artifacts**, consisting of software, code or models.
- 145 – **Documentation**, describing non- technical artifacts providing information about relevant steps, parameters and configurations.
- **Organisational or cultural factors and policies**, influencing types of collaborations, choice of technology or publishing of results.

2.4 Types of irreproducibility

150 Recognizing that the adversaries represent root causes of reproducibility barriers, we now investigate three major types of irreproducibility across the aforementioned dimensions:

Irreproducibility by obscurity (unintentional)

- **Data governance**. Limited findability, accessibility, interoperability and reusability of data, compounded by a general lack of open data, hinder various aspects of reproducibility.
- 155 – **Methodological shortcomings**. Flawed experimental designs and inadequate research methodologies impede inferential reproducibility and independent validation, even when procedures are published. Re-executing code primarily verifies portability rather than results validity.
- **Code quality**. Poorly written, undocumented, or unmaintainable code creates barriers to understanding and building upon existing implementations. Bugs can be difficult to identify and hamper inferential reproducibility.
- 160 – **Computing environment**. Variations in software versions, libraries, and computational resources complicate workflow replication.
- **Documentation**. Inadequate reporting standards. particularly the lack of detailed methodology documentation, obstruct understanding and reuse of published results.

Irreproducibility by obfuscation (intentional)

- 165 – **Proprietary data**. High-quality data often requires substantial investments in data collection and processing (e.g., field work, instrumentation, labeling). Due to these costs, researchers or institutions may hesitate to share such data, fearing a loss of competitive or financial advantage (Köster et al., 2021). This reluctance can result in incomplete disclosure of data collection methods or preprocessing steps, limiting the understanding of the research context.



- **Opaque methods.** Computational or statistical models with no (black box) or limited (grey box) insight into their internal workings obscure how inputs are transformed into outputs. This opacity hampers validation, error diagnosis, understanding of limitations, and further development.
- **Selective reporting.** Beyond selective reporting bias (Marks-Anglin and Chen, 2020), this includes the omission of code snippets or undocumented manual processing steps, whether deliberate or accidental.

Irreproducibility due to uncontrollable conditions

- **Computing hardware demands.** Specialized or expensive hardware (e.g., supercomputers, proprietary configurations) makes reproduction costly or inaccessible. Examples include climate projections, numerical weather predictions (Bauer et al., 2015), satellite embeddings (Brown et al., 2025) and non-deterministic model outputs from generative AI and Monte-Carlo simulation models.
- **Specialized sensing / measurement equipment.** Data collection may rely on expensive and specialized tools (e.g. earth observation satellites, ground-penetrating radar, superconducting gravimeters, aeromagnetic sensors or airborne laserscanning systems).
- **Technological obsolescence.** Hardware and software used may become outdated, difficult to access, or even unavailable over time.
- **Environmental context.** In geosciences, measurements are commonly taken under variable environmental conditions that cannot be fully controlled or replicated (Ryan, 2011). Challenges include spatial inhomogeneities (including effects of administrative boundaries in data collection), temporal inconsistencies (e.g., breakpoints caused by relocating measurement stations or changing equipment), and logistical difficulties of fieldwork in demanding environments (e.g., glaciological monitoring).
- **Transient events.** Singular natural phenomena (e.g. volcanic eruptions, gravitational natural hazards), occur at specific points in time. Studies rely on observations collected around the time of event occurrence, which cannot be independently replicated if missed.
- **Limited control over data collection.** Researchers occasionally rely on datasets collected by third parties for different purposes or even without scientific intent. Consequently, the available data may not align well with specific requirements and constrain the research process.
- **Confidentiality concerns.** Privacy or security restrictions (e.g., property locations and attributes) prevent open sharing, as they may pose risks to individuals or communities if disclosed.



3 Guiding principles and implementation strategies for geoscientific workflows

Promoting reproducibility often requires navigating challenges of varying complexity.

1. **Low-hanging fruits.** Simple, broadly applicable solutions can resolve many reproducibility issues. For instance, adopting open-source formats, programming languages, and libraries ensures accessibility and facilitates rerunning published code.
2. **Moderate issues.** Some well-known practices remain underutilized in practice. These include using open licenses for code and data, preparing thorough documentation, and compiling relevant metadata (c.f. FAIR data principles). While these steps reduce future hurdles for both the original investigators and others, their adoption is often inhibited by insufficient awareness and inadequate incentives (c.f. Easterbrook, 2014; Osborne et al., 2014).
3. **Wicked challenges.** Reproducibility in data analysis presents inherent complexities. Tools tailored to software engineering (e.g. version control, continuous integration) are powerful but face limitations when dealing with large datasets, long computation times and experimental coding. In contrast to software projects, data analysis typically prioritizes answering immediate research questions over long-term maintainability. Reproducibility exists on a spectrum, requiring practical implementations to strike a balance between ease of implementation and potential benefits for the intended audience and use cases. While automating computational workflows is valuable, it does not inherently validate the correctness of models or their alignment with research objectives. Researchers should identify whether their primary contribution lies in sharing the methodology, results, or both, and tailor their technical artifacts accordingly. Common use cases include reusing the code as a library, leveraging computation results, or understanding experiment design and implementation details.

These aspects inform the broader goals of reproducibility and reusability. With these priorities in mind, we now outline guidelines for enhancing reproducibility in geospatial workflows, organized by task and increasing complexity.

3.1 Overarching principles & planning

The following fundamental principles hold as general recommendations throughout the entire research workflow.

Use a good folder structure and naming conventions

A clear folder structure and meaningful filenames are essential to efficiently manage and collaborate on computational projects (Noble, 2009; Wilson et al., 2017; Van Lissa et al., 2021). By keeping data, code and results systematically organized, locating and managing different project components is greatly facilitated (Fig. 1). This streamlines workflows by providing a clear layout from the outset, thus not only saving time for setup and maintenance of the project structure but also reducing errors and confusion, especially in collaborative settings. Moreover, a scalable structure accommodates project growth, allowing new data, scripts, and analyses to be added without clutter.

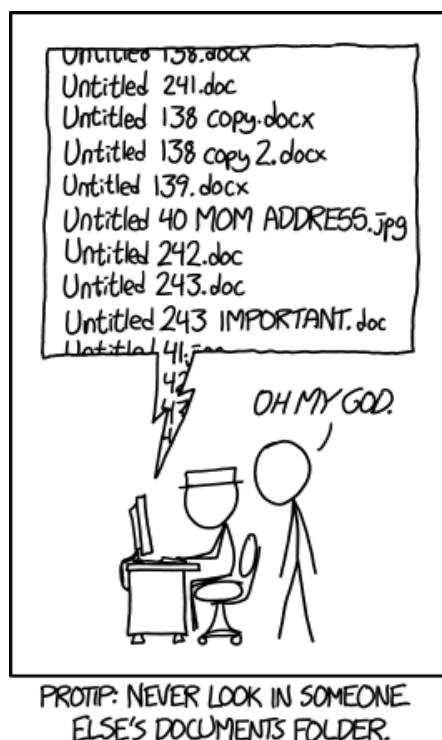


Figure 1. xkcd #1459: Documents. Source: <https://xkcd.com/1459/>. [Licensed under CC BY-NC 2.5].

To get started, create a dedicated folder for each project, using a clear and descriptive title as the folder name. Tools like cookiecutter¹ can be used to create an initial folder hierarchy and commonly needed files from templates, which can also serve as inspiration for structuring your project (e.g., Cookiecutter Data Science²).

- 230 Good naming practices involve concise, easy-to-type terms that balance specificity and generality. Create appropriate terms or abbreviations if necessary. Use lowercase letters and hyphens (or underscores if hyphens are not allowed, e.g., in package names), avoiding spaces and special characters to ensure compatibility across systems. Reserved filenames such as `README.md` or `LICENSE`, are typically in uppercase. For more on coding conventions, see Adhere to a consistent coding style.

Adopt open standards and technologies

- 235 Using open standards and technologies enhances transparency and reproducibility throughout the research workflow. This includes open data formats (c.f. Use common open data formats), free and open-source software (c.f. Use free and open-source software), and open-access publishing (c.f. Publish under a permissive or copyleft license). Open standards enable sharing and verification of data and methods, while open-source software enables others to inspect, modify, and build upon the published output. Open-access publishing greatly facilitates access to scientific findings, broadening their reach and impact. Together,

¹<https://github.com/cookiecutter/cookiecutter>

²<https://cookiecutter-data-science.drivendata.org>



240 these practices improve knowledge dissemination, simplify collaboration, reduce entry barriers, and support the validation and reuse of research outputs.

Develop (and implement) an actionable data management plan (DMP)

Effective implementation of best practices in reproducibility requires strategic planning and resource allocation. A DMP helps structuring this process by identifying key milestones where data and code should be archived, shared, or released, either
245 internally or publicly. It summarizes key aspects such as data types, storage, security, sharing and quality. Many funding agencies now mandate DMPs to ensure compliance with ethical, legal, and institutional frameworks (Miksa et al., 2019), yet their actual implementation is rarely monitored. Institutional research data management units can provide valuable support in developing DMPs and training research staff, ensuring that data stewardship is not only planned but also effectively executed.

Document both high-level and low-level provenance

250 Provenance refers to information about the origins, processes, and contributors involved in producing data or results, which helps to assess their quality, reliability or trustworthiness (Belhajjame et al., 2013). This includes method parameters, input and output data, software versions, and the computational environment (Khan et al., 2019).

To support reproducibility, both high-level provenance (i.e., ‘What steps were followed in this analysis?’) and low-level provenance (i.e., ‘How was this specific result computed?’) should be documented. This information is valuable for debugging,
255 monitoring, and reproducing results.

Isolate and document irreproducible workflow components

Some workflow components, such as expert decisions, external web services, non-deterministic AI inferences, or continuously updated data streams, are inherently irreproducible (c.f. section 2). To mitigate their impact, strategically isolate these components and document them thoroughly. One approach is to save intermediate results of the irreproducible steps to ensure
260 conditional downstream reproducibility. For workflows dependent on external APIs or databases, caching responses or storing queried datasets locally prevents inconsistencies caused by data updates or service outages. Submitted queries or prompts should be stored along with the responses especially in the case of AI models and dynamic web contents. When expert decisions are involved, maintaining detailed records—such as meeting protocols, internal memos, issue comments, or email correspondence—enhances transparency and enables future researchers to assess their influence.

265 Automate repetitive tasks

Automation ensures consistency, reduces human error and saves time (c.f. Use script-based workflows). By automating repetitive tasks, workflow steps are implicitly documented, fostering a more efficient and reliable research environment. Steps that cannot be automated may require more thorough documentation, as there is no code for reference (c.f. Isolate and document irreproducible workflow components).



270 Use tidy data

Data cleaning and preparation are often tedious, especially when data do not adhere to a defined structure (c.f. database normalization). The concept of tidy data, introduced by Wickham (2014), describes a standardized way of organizing (tabular) data to simplify analysis and visualization by mapping the meaning of a dataset to its structure. Tidy data adheres to three principles: (1) each observation forms a row, each row represents an observation, (2) each variable forms a column, each
275 column contains a variable, and (3) each type of observational unit forms a table, with each value being a cell and each cell being a single value.

In geospatial research, tidy data principles can easily be applied to vector datasets, which are essentially dataframes with a geometry column. Packages such as `sf` in R and `geopandas` in Python integrate well within their respective ecosystems and thus facilitate preparing data in a tidy format.

280 Having established these overarching principles, the following sections delve into specific aspects of geospatial research workflows.

3.2 Data collection & governance

Geospatial analyses, whether using big data or small samples, require high-quality, well documented, and interoperable data to achieve reliable and reproducible results. The objective is to ensure that data can be interpreted and reused reliably by both the
285 original researcher over time and by external stakeholders.

3.2.1 Description & assessment of data

Identify and describe your data

Begin each research endeavor by identifying and evaluating available data sources to ensure they are sufficient and suitable for addressing your research question.

290 For geospatial data, use appropriate file formats with clearly defined coordinate reference systems (CRS), specified using standardized formats like EPSG-codes³, OGC WKT strings⁴ or PROJ-strings⁵. Pay attention to practical challenges, such as custom definitions of well-defined standard CRS, discrepancies between documented CRS and those stored in file metadata, or handling uncommon projections (e.g., the equidistant rotated pole projection used in EURO-CORDEX). For spatiotemporal data, document the meaning of timestamps (including interval labeling on the left, center or right), time zones, calendar types
295 (e.g., leap years or 360-day years), and aggregation functions.

³<https://epsg.org>

⁴<https://www.ogc.org/standard/wkt-crs>

⁵<https://proj.org>



Consider the spatial / spatiotemporal nature of data

Inferential reproducibility in geospatial data science requires modeling strategies that explicitly address the unique characteristics of spatial data, specifically spatial and spatiotemporal autocorrelation (Wikle, 2014). Such approaches include spatial interpolation using Kriging (Li and Heap, 2011), geographically and temporally weighted regression (Fotheringham et al., 2015) or spatiotemporal clustering (Birant and Kut, 2007). Whether or not such explicitly spatial models are used, it is crucial that performance estimation takes into account the predictive setting in which the models will be applied, thus requiring the consideration of concepts such as spatial or temporal prediction distance and autocorrelation through the use of suitable cross-validation techniques (Brenning, 2023; Schratz et al., 2024). In this context there is a risk of information leakage from validation data into the training process unless hyperparameters are tuned in a nested cross-validation (Schratz et al., 2019). Model performance metrics should focus on the user's perspective rather than the producer's, as standard metrics like overall classification accuracy fail to account for the varying costs of different types of misclassification. Additionally, extrapolation in geographic space should be flagged by identifying a model's area of applicability (Meyer and Pebesma, 2021), and trade-offs between statistical model performance versus process-based plausibility should be considered (Reichstein et al., 2019; Steger et al., 2016).

Assess the quality of your data

Data quality is highly contextual, depending on the domain and task. In geospatial research, standard quality metrics often require adaptation to address domain-specific needs. For example, in geospatial ML models, the intended use cases and the scope of training data are crucial for evaluating their suitability for future applications. Established methodologies (Batini et al., 2009) and metrics such as accuracy, consistency, completeness and timeliness (Ehrlinger and Wöß, 2022) guide assessments and improvements.

High-quality metadata is essential for geospatial data. This includes information on dataset provenance, completeness, logical consistency and accuracy across positional, temporal and attribute dimensions (Lush et al., 2012). However, collecting such metadata is resource-intensive and requires expertise and effort beyond standard data collection processes. Additionally, information on licensing, the reputation of the data provider, and 'soft knowledge' about uncertainties when quantitative estimates are unavailable enhances data reusability.

3.2.2 Standards & common formats

Use common open data formats

Using open, standardized data formats enhances accessibility and reusability (Fig. 2). The choice of format should balance accessibility and performance based on the requirements of the use case. Human-readable formats like CSV and (Geo)JSON are easier to inspect and arguably more stable over time, while binary formats (like GeoParquet, Arrow/Feather, FlatGeobuf)



offer better performance in terms of I/O speed, size and resource efficiency, particularly for large-scale data processing. Use a common character encoding, preferably UTF-8, to ensure compatibility across systems.

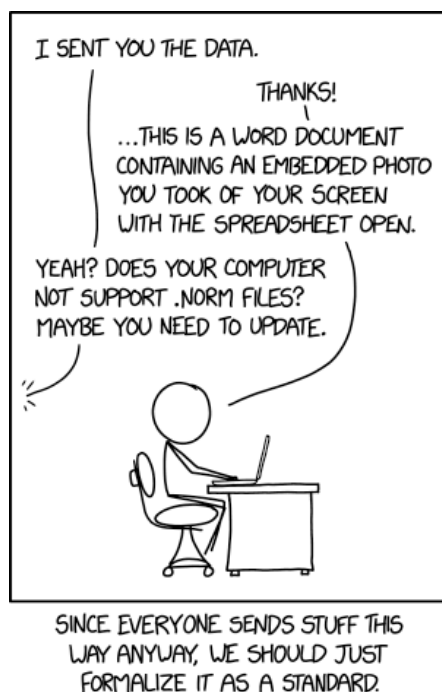


Figure 2. xkcd #1459: .NORM Normal File Format. Source: <https://xkcd.com/2116/>. [Licensed under CC BY-NC 2.5].

Use geodata formats supported by GDAL/OGR raster and vector drivers (GDAL/OGR contributors, 2024; Warmerdam, 2008) and OGC standards⁶. For vector data, GeoPackage is preferred over the widely used shapefile format, which has notable technical and practical limitations. For raster data, GeoTIFF is a feasible standard format, with COG (for efficient streaming) and NetCDF, HDF5 or Zarr (for multi-dimensional arrays) being also viable for specific needs. Note that advanced features like compression or chunking require some familiarity with the underlying data engineering concepts.

Adopt taxonomies, controlled vocabularies and ontologies

Taxonomies and controlled vocabularies provide knowledge management tools for organizing structured information for both data and metadata. While geoscience-specific resources exist (e.g., landslide type classifications, (Hung et al., 2014)), they are often not formalized in machine-readable formats. Ontologies provide such a formal representation format and support both knowledge sharing and interoperability. They are well established in certain scientific domains such as medicine (Ivanović and Budimac, 2014) or computer science (Salatino et al., 2018). Large open knowledge graphs, for instance, based on Wikipedia

⁶<https://www.ogc.org/standards/>



data, have been created collaboratively and can be used for various use cases (e.g. DBpedia⁷, Wikidata⁸). However, adoption
340 of ontologies in geosciences faces challenges: a steep learning curve of the underlying technologies, the need for extensions in
complex modeling (e.g., spatiotemporal data), and limited adaptation due to knowledge transfer and integration issues from do-
main experts (Claramunt, 2020). Suitable ontologies for reuse in geospatial applications include W3C Basic Geo Vocabulary⁹,
Dublin Core¹⁰ (including the spatial extension), SSN/SOSA¹¹ for sensor and observation data and GeoSPARQL¹² for query-
ing purposes. The KnowWhere Graph integrates geospatial data on various natural hazards, climate variables, topographical
345 information and other relevant information for reuse (Shimizu et al., 2025).

Use established location reference systems

Relying on standardized spatial reference systems is crucial for long-term reproducibility in geospatial analyses (c.f. ISO
19111). Specify coordinate reference systems by their EPSG code¹³ to ensure unambiguous interpretation. The unprojected
WGS84 (EPSG:4326) is generally preferred for global applications due to its widespread adoption. For analyses requiring
350 equal-area projections, selecting an appropriate standardized CRS (e.g., EPSG:6933 for global equal-area mapping) is crucial
to avoid distortions in spatial measurements. For multi-scale or hierarchical data, standardized Discrete Global Grid Systems
(DGGS) or equal-area hexagonal grids (e.g., ISEA3H) provide robust spatial partitioning. While newer systems such as the H3
grid have gained popularity, they lack standardization and may still evolve, potentially introducing reproducibility challenges.

When working with toponyms (place names), following established semantic ontologies ensures interoperability. The GeoN-
ames gazetteer¹⁴ provides a well-structured ontology linking place names to Uniform Resource Identifiers (URIs) in the Se-
355 mantic Web, facilitating precise spatial referencing. Advances in AI and natural language processing (e.g. named entity recog-
nition), enable the automated extraction of location references from text (Hu et al., 2023). Incorporating such tools can enhance
reproducibility by establishing explicit, machine-readable geographic references within the GeoNames framework.

Use templates and checklists

360 A shared understanding of data and models is crucial for reproducibility. Clear documentation of their purpose, performance,
and limitations improves transparency and accountability. Using templates, checklists and adhering to standards facilitates the
process of documenting important information on models and datasets. For example, model cards (Mitchell et al., 2019) and
datasheets (Gebu et al., 2021) offer structured frameworks to capture metadata, including intended use and potential biases,
and are widely adopted on platforms like Huggingface¹⁵. In the geospatial domain, standards like ISO 19115, which addresses

⁷<https://www.dbpedia.org>

⁸<https://www.wikidata.org>

⁹<https://www.w3.org/2003/01/geo/>

¹⁰e.g. <http://purl.org/dc/terms/>

¹¹<https://www.w3.org/TR/vocab-ssn/>

¹²<https://www.ogc.org/publications/standard/geosparql/>

¹³<https://epsg.io>

¹⁴<https://www.geonames.org/>

¹⁵c.f. <https://huggingface.co/docs/hub/en/model-cards>



365 geographic information metadata (e.g., update frequency, spatial resolution and reference system), provide a foundation which could be adapted to include domain-specific parameters and ethical considerations, such as those relevant to satellite imagery (Rolf et al., 2024).

3.3 Analysis design & automation

Major data manipulation activities (e.g., cleaning, pre-processing, modeling, and analysis), along with their inputs and out-
370 puts, should be described and tracked to ensure reproducibility and traceability. The goal is to make these steps transparent, repeatable and automated.

3.3.1 Provenance

Track model development

Machine Learning Operations (MLOps) tools help streamline the development, improvement, deployment, and maintenance
375 of ML models (Kreuzberger et al., 2023). Experiment tracking and model registry tools, such as MLflow (Chen et al., 2020; Zaharia et al., 2018), ClearML (ClearML, 2024) or Weights & Biases (Biewald, 2020), support the documentation of the entire system lifecycle for ML models, notably including automated recording of hyperparameters and performance metrics (c.f. Explore dataflow programming and pipeline tools). While open-source tools (including MLflow and ClearML) can be self-hosted, proprietary platforms may require licensing for larger teams. Geospatial projects come with specific requirements in
380 terms of tasks and data size, calling also for the development of domain-specific support tools and higher scalability.

Document provenance and workflow rationale

Provenance tracking ensures transparency in data processing and workflow design. Tools like noWorkflow for Python (Murta et al., 2015), E2ETools (Lerner et al., 2023, 2018) and dtrackr (Challen, 2022) for R, and more comprehensive workflow engines such as VisTrails (Callahan et al., 2006), Taverna (Hull et al., 2006) or Kepler (Altintas et al., 2006) support (script-
385 based) collection of provenance traces (Pérez et al., 2018; Pimentel et al., 2019a). For custom workflows, generic provenance data models (Prov-DM; Lebo et al., 2013) and P-Plan Ontology (Garijo and Gil, 2012) can be adapted and extended.

Capturing domain-specific annotations is particularly valuable for encoding relevant design steps (Khan et al., 2019; Garijo et al., 2014). General examples include data preparation steps, e.g. format transformation or data grouping. Relevant specialized motifs in geospatial data processing include reprojection of input data sources into a common coordinate reference system,
390 spatial alignment of data via coregistration, homogenization of observational datasets, atmospheric correction in remote sensing data, specification of spatial relationships (c.f. DE-9IM, Clementini et al., 1994) when combining/joining datasets, or numerical simulation parameters.



3.3.2 Scripted & automated workflows

Use script-based workflows

395 We advocate prioritizing script-based workflows over GUI-based or manual processes wherever possible. Scripts automate
processing steps, ensuring consistent and repeatable results while minimizing the risk of human error. They also serve as self-
documenting records, enabling reconstruction and verification of analytical processes. When combined with version control
systems like Git (c.f. Use a version control system (VCS)), scripts provide precise and transparent tracking of modifications
over time. They also facilitate error-handling to detect exceptions and ensure robustness and reliability across different datasets,
400 environments, and scales of operation.

Additionally, script-based workflows offer modularity for structured development, support collaborative research environ-
ments, and enable scalable and parallelizable analyses of big data in distributed and high-performance environments. This
improves efficiency and allows for optimization of runtime and memory requirements.

Explore tools for literate programming

405 Literate programming, introduced by Knuth (1984), integrates code and documentation into a cohesive, executable narrative,
producing dynamic reports where source code is embedded within explanatory text (Stodden et al., 2018). Tools like Quarto,
knitr (Xie, 2015), R Markdown (Xie et al., 2018), and Jupyter Notebooks (Pimentel et al., 2019b) enhance clarity, accessibility,
and practical reproducibility by combining code with narrative (Vassilev et al., 2016). This approach is used in software de-
velopment, such as R package ‘vignettes’ for long-form documentation, and in practical geospatial analyses, where executable
410 notebooks facilitate the presentation of workflows and results. Literate programming also underpins frameworks like workflow
in R (Blischak et al., 2019), which support the creation and sharing reproducible research code.

Explore dataflow programming and pipeline tools

Workflow management systems and pipeline tools enhance reproducibility, scalability, and modularity by structuring processes
as Directed Acyclic Graphs (DAGs), which represent data flow and task dependencies. These tools formalize workflows,
415 automate task execution, and ensure traceability of data transformations.

For low to moderate complexity workflows, tools such as make (Feldman, 1979), Snakemake¹⁶ (Koster and Rahmann,
2012; Mölder et al., 2021), doit¹⁷, pytask¹⁸ or targets¹⁹ (Landau, 2021) provide lightweight solutions for task automation. ML

¹⁶<https://snakemake.readthedocs.io/en/stable/>

¹⁷<https://pydoit.org/>

¹⁸<https://pytask-dev.readthedocs.io/en/stable/>

¹⁹<https://books.ropensci.org/targets/>



pipeline frameworks such as scikit-learn²⁰, Pythonflow²¹, mlr3pipelines²² or recipes²³ help streamline preprocessing, modeling and evaluation.

420 For complex workflows requiring scheduling, monitoring and deployment, platforms like Apache Airflow²⁴, Luigi²⁵, MLflow²⁶, KubeFlow²⁷ (Bisong, 2019) or Dagster²⁸ offer robust orchestration capabilities. CI/CD tools further support executing pipelines and deploying models in suitable environments (c.f. Perform automatic testing).

When transitioning research code into operational services, early consideration of task allocation—separating code development (scientists) from execution and deployment (DevOps)—is essential for scalability and reproducibility (c.f. Consider code
425 life cycle and maintenance). Infrastructure as Code complements workflow management systems by enabling the automated provisioning and configuration of computational environments. This approach ensures consistency across development, testing, and production stages, reduces manual errors, and enhances reproducibility by version-controlling infrastructure alongside code and workflows.

3.4 Code & scientific computing

430 This section covers a broad range of aspects commonly encountered during the main development phase of geospatial data science projects. The aim is to obtain maintainable, shareable and reproducible computational code.

3.4.1 Tool & environment selection

Use free and open-source software

Using common free and open-source software increases the accessibility and practical reproducibility of results (Fortunato and
435 Galassi, 2021; Pebesma et al., 2012). The growing popularity of open-source programming languages and GIS reflects a shift from proprietary tools, lowering barriers to entry and promote transparency in geoscientific workflows (Lovelace, 2021).

In geocomputation, languages like Python (Dorman et al., 2024), R (Lovelace et al., 2019; Pebesma and Bivand, 2023) and Julia (Hoffmann, 2023), along with command line utilities like `gdal`²⁹, `cdo`³⁰ and `nco`³¹, enable efficient geodata processing. Open-source GIS software such as QGIS, GRASS GIS and SAGA provides viable alternatives to proprietary desktop GIS tools
440 like ArcGIS, and integrates seamlessly into script-based workflows.

²⁰<https://scikit-learn.org/stable/modules/generated/sklearn.pipeline.Pipeline.html>

²¹<https://pythonflow.readthedocs.io/en/latest/>

²²<https://mlr3pipelines.mlr-org.com/>

²³<https://recipes.tidymodels.org>

²⁴<https://airflow.apache.org/docs/>

²⁵<https://luigi.readthedocs.io/en/stable>

²⁶<https://mlflow.org/docs/latest>

²⁷<https://www.kubeflow.org/docs/>

²⁸<https://docs.dagster.io/>

²⁹<https://gdal.org/>

³⁰<https://code.mpimet.mpg.de/projects/cdo>

³¹<https://nco.sourceforge.net/>



Use defined environments

Reproducibility depends not only on code but also on the computing environment, which influences results and should be documented for transparency. The hardware layer includes processing units (CPU, GPU), storage (RAM and disk), and network access. While exact requirements may be difficult to determine, specifying the system used helps others replicate analyses.

445 The logical or software layer includes the operating system, dependencies (applications and libraries), and project-specific code. These components and their versions can affect computations, making isolated environments or containers essential for reproducibility. Dependency versions should be exported to a text file and shared with the code.

Language-specific environments (e.g., `venv`³², `poetry`³³, `uv`³⁴ or `renv`³⁵) manage dependencies within a single programming language, while cross-language environments ((`mini`)`conda`³⁶, (`micro`)`mamba`³⁷, `pixi`³⁸) provide broader dependency manage-
450 ment and better isolation. Containers like `docker`³⁹ or `podman`⁴⁰ encapsulate operating system-level dependencies but may add unnecessary complexity for most data analysis applications (c.f. Moreau et al., 2023; Choi et al., 2023).

Leverage collaborative development platforms

Platforms like GitLab or GitHub provide valuable support for managing scientific projects beyond software development. Features such as issue tracking and merge requests facilitate collaboration, improve code quality through peer reviews, and
455 document discussions and decisions. Issue boards can serve as Kanban-style project management tools, aiding both organizational and technical workflows. While designed for teams, these platforms are equally beneficial for individual researchers, helping structure tasks and maintain clear documentation. Following an issue → branch → merge request workflow is best practice for improving organization, documentation and reproducibility, even in solo projects.

3.4.2 Coding & development practices

460 Adhere to a consistent coding style

Clean, consistent code facilitates effective scientific communication (Filazzola and Lortie, 2022), and improves accessibility. Using meaningful filenames without special characters, and prepending script filenames with numbers to indicate execution order helps to navigate and locate files and grasp their meaning. Adhering to coding style guides⁴¹ ensures uniformity, facili-

³²<https://docs.python.org/3/library/venv.html>

³³<https://python-poetry.org>

³⁴<https://docs.astral.sh/uv>

³⁵<https://rstudio.github.io/renv>

³⁶<https://docs.anaconda.com>

³⁷<https://mamba.readthedocs.io>

³⁸<https://pixi.sh/dev/>

³⁹<https://docs.docker.com>

⁴⁰<https://podman.io>

⁴¹e.g., <https://peps.python.org/pep-0008> for Python or <https://style.tidyverse.org/index.html> for R



tating readability, collaboration, and maintenance. Tools like linters and autoformatters⁴² can automate style enforcement (c.f.
465 Perform automatic testing).

Don't repeat yourself (DRY)

Minimizing code duplication is a best practice, as repeated code increases maintenance efforts and the risk of inconsistencies. Modularizing code with functions and packages improves reusability, maintainability, and coding efficiency. However, deeply nested code increases complexity and complicates debugging, especially with complex geospatial data. For ill-defined prob-
470 lems, a flat, interactive workflow is often better suited for exploration and rapid prototyping. While deviating from the DRY principle can be pragmatic during early development, transitioning prototypes into maintainable software requires substantial effort, even without adding new features. Recognizing this challenge early helps balance exploratory coding with long-term maintainability.

Consider code life cycle and maintenance

475 The purpose and lifespan of scientific code influence how it should be written and maintained (Adorf et al., 2019). For one-time analyses, prioritizing human readability is key, as future users may need to understand the methodology rather than re-run or modify the code. Conversely, code intended for long-term operational use requires careful planning for maintainability. Research prototypes often evolve into operational services; but without dedicated resources for refactoring, optimization, automation, and testing, transitioning them into sustainable systems can be difficult. Early planning, ideally in collaboration with
480 software developers, helps address funding and technical challenges.

Provide logging information

Logging⁴³ captures key information during script execution, including system details, parameters, configuration, and progress (List et al., 2017). Log messages, including warnings, errors, and executed steps, are typically printed on-screen or saved with timestamps in a file or database. This facilitates process monitoring, error tracing, and provenance documentation. To enhance
485 transparency, we recommend publishing log files of the final computation alongside the results to document when and how the analysis was performed.

3.4.3 Validation & versioning

Perform automatic testing

Testing is crucial for ensuring code correctness. Different techniques serve specific purposes: *unit tests* validate individual
490 components or functions with various inputs, helping to isolate errors efficiently. *Functional tests* or *integration tests* verify that all components work together as expected. *Plausibility tests* check whether results fall within reasonable bounds (e.g., non-

⁴²e.g., Black (Łukasz and contributors to Black, 2024) for Python or styler (Müller and Walthert, 2024) for R

⁴³e.g. <https://docs.python.org/3/howto/logging.html> for Python or logger (Daróczy and Wickham, 2024) for R



negative values). *Regression tests* ensure code changes do not unintentionally alter results (c.f. Use a version control system (VCS)). *Performance tests* monitor runtime and memory usage help prevent efficiency regressions.

495 In the geosciences, where data often originate from heterogeneous sources, *input validation* is essential to assess data correctness and meaningfulness by combining technical and plausibility checks. Automating tests with code snippets (*test cases*) ensures consistent execution after changes, helping detect bugs early, saving development time and increasing confidence in results.

Code linters⁴⁴ can automatically identify syntax errors, flag potential issues, and enforce coding standards (c.f. Adhere to a consistent coding style).

500 For further automation, Git pre-commit hooks and continuous integration (CI)⁴⁵ can enforce successful test runs after each commit or pull/merge request.

The impact of undetected errors is often underestimated. While testing adds complexity, it reduces time lost to debugging. Automated tests are a valuable investment but should complement, not replace, expert judgment in addressing unexpected errors or implausible results.

505 Use a version control system (VCS)

VCS such as Git, Mercurial or Subversion facilitate collaboration, tracking, and management of code and data projects. They provide key features such as lineage (tracking changes and provenance over time), time travel (navigating or restoring previous states) and auditability (providing a detailed change log). Features like branching (creating separate development lines) and merging (integrating changes) enable parallel development and efficient maintenance.

510 Git stands out as a free, open-source, and widely supported tool, with integration into most editors and IDEs. Its web-based platforms for collaborative development and code management further enhance accessibility.

Although VCS tools have a learning curve (Fig. 3), mastering basic commands already provides long-term benefits beyond collaboration. Even for individual researchers, regular snapshots not only improve reproducibility and safeguard progress and allow safe experimentation without the risk of losing previous work.

515 Version-control data as well

While version control for source code is well established, data versioning—particularly for large binary files—remains challenging. Data often requires additional capabilities, such as deleting old versions to free storage space or partially cloning repositories. Standard VCS operations like merging or cherry-picking are typically unsuitable for binary data. Additionally, the exploratory nature of data analysis does not always align with traditional branching workflows.

⁴⁴e.g. <https://flake8.pycqa.org/> or <https://docs.astral.sh/ruff/> for Python or <https://lintr.r-lib.org/> for R

⁴⁵e.g., GitLab CI/CD, GitHub Actions or Jenkins

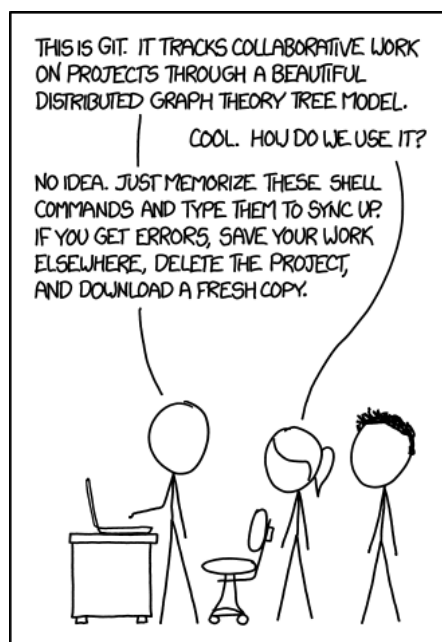


Figure 3. xkcd #1597: Git. Source: <https://xkcd.com/1597/>. [Licensed under CC BY-NC 2.5].

520 Despite these challenges, automating the addition of generated data and results to a VCS repository enhances traceability by linking output files to the code and parameters that created them. Tools like Git LFS⁴⁶, DVC⁴⁷, lakeFS⁴⁸, Dolt⁴⁹, Nessie⁵⁰, and Pachyderm⁵¹ address data versioning but can be complex and may not fit all research workflows. A simpler approach is to store analysis outputs in a dedicated folder and associate them with the corresponding Git commit hash, ideally through automation.

3.5 Publishing and reuse of results

525 Code and data availability statements are commonly required in renowned scientific journals, including *Nature*⁵², *Science*⁵³, as well as the journals of the American Geophysical Union⁵⁴ and the European Geophysical Union⁵⁵. While open-access publishing is increasing (Piwowar et al., 2018; Munroe, 2013), data is not always provided, even when required by journals (Alsheikh-Ali et al., 2011; Savage and Vickers, 2009). While willingness to share research data has been linked to the strength

⁴⁶<https://git-lfs.com>

⁴⁷<https://dvc.org>

⁴⁸<https://lakefs.io>

⁴⁹<https://www.dolthub.com>

⁵⁰<https://projectnessie.org>

⁵¹<https://www.pachyderm.com>

⁵²<https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards>

⁵³<https://www.science.org/content/page/science-journals-editorial-policies#data-and-materials-after-publication>

⁵⁴<https://www.agu.org/publish-with-agu/publish/author-resources/data-and-software-for-authors>

⁵⁵https://publications.copernicus.org/services/data_policy.html



of evidence and quality of statistical reporting (Wicherts et al., 2011), concerns about data misuse and lack of proper acknowl-
530 edgment remain barriers to sharing in the geoscience community (Tenopir et al., 2018).

Best practices are increasingly promoted through specific conference tracks, for instance NeurIPs for datasets and bench-
marks⁵⁶ or semantic web resource tracks⁵⁷, requiring well-documented and reusable artifacts, such as datasets or code.

Our recommendations draw inspiration from Goodman et al.'s short guide on the care and feeding of scientific data (Good-
man et al., 2014).

535 3.5.1 Sharing of artifacts & results

Publish all data, code, and results

All methodological and technical artifacts used to derive scientific conclusions should be made available and properly refer-
enced (Chen et al., 2019; Stodden et al., 2016, 2018). This includes input data, computational workflows (code and environ-
ment), output data and metadata (Konkol et al., 2020; Barba, 2012). Computational models should also be published as digital
540 research objects alongside other results.

Providing full access to data, methods, and results supports precise replication, fostering transparency, accountability, and
longevity of research contributions. When data cannot be shared (e.g. due to confidentiality, legal or proprietary constraints),
synthetic datasets can be used as alternatives (Endres et al., 2022; Wimmer and Finger, 2022). However, care must be taken to
avoid biases or misrepresentations of real-world characteristics (Chen et al., 2021).

545 **Publish under a permissive or copyleft license**

Choosing an appropriate open license is essential to enable reuse, modification, and redistribution of scientific outputs. Permis-
sive licenses (e.g., MIT, Apache, BSD) impose minimal restrictions, allowing for broad adoption and integration. In contrast,
copyleft licenses (e.g. GNU GPL, CC BY-SA) require derivative works to maintain the same licensing, preventing proprietiza-
tion.

550 Ideally, all relevant output (code, data, publications) should be published under explicit open licenses, as default copyright
laws restrict reuse. For software (c.f. Morin et al., 2012), permissive licenses are generally preferred for interoperability, as
they allow integration with open and proprietary projects.

For data, the Creative Commons licenses like CC0 (public domain dedication) or CC BY (attribution required) are the
most compatible, whereas CC BY-SA (share-alike) enforces copyleft principles that may limit integration with datasets under
555 different licenses. Managing and tracking licenses throughout a project is thus critical.

Open-access publishing remains a priority, despite ongoing debates over publishing models (Frank et al., 2023), costs
(Van Noorden, 2013) and predatory publishers (Beall, 2012). Many publishers and journals permit preprint sharing (c.f. Bourne
et al., 2017), although policies vary and may be unclear (Klebel et al., 2020; Massey et al., 2020).

⁵⁶<https://neurips.cc/Conferences/2024/CallForDatasetsBenchmarks>

⁵⁷<https://2024.eswc-conferences.org/call-for-contributions-eswc-2024/>



Publish in a machine-readable, well-documented, common and stable format

560 Choosing widely supported, stable, and future-proof formats is important for reliable data storage (c.f. Use common open data formats). Machine-readable formats (i.e., structured data) are crucial to ensure that data can be readily processed by computers (Fig. 2). Caution is advised with certain file types, including (i) proprietary formats that require specific software (which may become obsolete or no longer supported), (ii) niche binary formats with limited support, and (iii) compressed formats, which may risk corruption during decompression. For instance, while the Feather format (version 2) is now stable and suitable for
565 interoperable, language-agnostic data storage (c.f. Apache Arrow IPC), its legacy version 1 was explicitly not designed for long-term use.

3.5.2 Long-term access

Publish in long-preserving manner with a persistent identifier

Data should be published in open repositories with strong institutional support for long preservation⁵⁸, such as Zenodo⁵⁹ or
570 PANGAEA⁶⁰. These provide persistent digital object identifiers (DOIs) to ensure long-term accessibility, unlike generic URLs. Computational code and workflows can be shared via platforms like Code Ocean⁶¹, which not only assign DOIs but also ensure reproducibility by preserving the execution environment (Nature Physics Editorial, 2025).

Adhere to data stewardship & management best practices

Guiding standards like the FAIR (Findability, Accessibility, Interoperability, Reusability) principles (Wilkinson et al., 2016)
575 and 5-Star Open Data⁶² provide best practices for (meta)data management and sharing. While these principles outline ideal conditions, limited funding and resources in research projects often necessitate prioritization frameworks, such as the FAIR maturity model (Bahim et al., 2020).

⁵⁸<https://www.re3data.org/>

⁵⁹<https://zenodo.org>

⁶⁰<https://www.pangaea.de/>

⁶¹<https://codeocean.com>

⁶²<https://5stardata.info/en/>



4 Conclusions

Reproducibility is a cornerstone of the scientific method, yet upholding it in geoscientific research practice remains challenging. Geoscientific data analysis thrives on a dynamic balance between creativity and rigor, embodied in the interplay between divergent and convergent thinking. Divergent thinking drives exploratory phases, enabling researchers to uncover patterns and generate innovative insights from complex datasets. Convergent thinking ensures that these insights are distilled into reproducible workflows and well-documented outputs. Together, these approaches bridge the gap between exploration and reproducibility, fostering transparency, reliability, and long-term impact.

This work identified adversaries of practical reproducibility and outlined a set of implementation strategies for geoscientific workflows to address them. While certain challenges can be tackled directly (e.g., using open-source software and adhering to guidelines), other aspects require broader community efforts (e.g. domain-specific support and higher scalability for geoscientific data). By following these guidelines, even partially, researchers can strengthen the integrity, reliability, robustness of geoscientific research.

Code and data availability. The code and data used for supplementary analyses are publicly available on GitLab at <https://gitlab.com/Rexthor/reproducibility-in-geosciences>. Additionally, the supplementary material is hosted via GitLab Pages and can be accessed at <https://rexthor.gitlab.io/reproducibility-in-geosciences/>.

Author contributions. MS: Conceptualization, Software, Formal analysis, Data curation, Visualization, Writing - Original Draft, Writing - Review & Editing; LW: Conceptualization, Writing - Original Draft; PR: Conceptualization, Writing - Original Draft; AS: Conceptualization, Writing - Original Draft; AB: Writing - Original Draft, Writing - Review & Editing

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. Dedicated to the memory of Fritz Leisch (1968–2024).



Appendix A: Guiding principles and implementation strategies for reproducible geoscientific data analysis

Available at <https://rexthor.gitlab.io/reproducibility-in-geosciences/checklist.html>.

600 This checklist provides a structured framework to support reproducibility in geoscientific data analysis throughout the research workflow. It offers actionable guidance for scientists at all career stages.

To improve accessibility and usability, action items are prioritized using the MoSCoW method. This suggested prioritization offers a practical starting point to help researchers identify and focus on critical tasks. However, it is important to note that these priorities are flexible and may need to be adjusted based on the specific requirements and goals of individual projects.

605 Designed specifically for projects in the earth and environmental sciences that involve geoscientific data analysis, this checklist aims to serve as a practical reference to promote transparency, ensure reliability, and support the long-term usability of geoscientific research outputs.

Appendix B: Exploratory analysis of data and code availability statements in Copernicus journals

Available at https://rexthor.gitlab.io/reproducibility-in-geosciences/copernicus_report.html.

610 This supplement examines reproducibility practices in geosciences, with a focus on data and code availability in scientific publications. It provides a short exploratory analysis of data and code availability statements from 25 European Geosciences Union journals, covering the period from 2016 to 2024.



References

- Adorf, C. S., Ramasubramani, V., Anderson, J. A., and Glotzer, S. C.: How to Professionally Develop Reusable Scientific Software—And
 615 When Not To, *Computing in Science & Engineering*, 21, 66–79, <https://doi.org/10.1109/mcse.2018.2882355>, 2019.
- Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H., and Ioannidis, J. P. A.: Public Availability of Published Research Data in High-Impact
 Journals, *PLoS ONE*, 6, e24357, <https://doi.org/10.1371/journal.pone.0024357>, 2011.
- Altintas, I., Barney, O., and Jaeger-Frank, E.: Provenance Collection Support in the Kepler Scientific Workflow System, in: *Provenance and
 Annotation of Data*, edited by Moreau, L. and Foster, I., International Provenance and Annotation Workshop (IPAW), p. 118–132, Springer
 620 Berlin Heidelberg, ISBN 9783540463030, ISSN 1611-3349, https://doi.org/10.1007/11890850_14, 2006.
- Angrist, J. D. and Pischke, J.-S.: The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of
 Econometrics, *Journal of Economic Perspectives*, 24, 3–30, <https://doi.org/10.1257/jep.24.2.3>, 2010.
- Arvanitou, E.-M., Ampatzoglou, A., Chatzigeorgiou, A., and Carver, J. C.: Software engineering practices for scientific software develop-
 ment: A systematic mapping study, *Journal of Systems and Software*, 172, 110848, <https://doi.org/10.1016/j.jss.2020.110848>, 2021.
- 625 Bahim, C., Casorrán-Amilburu, C., Dekkers, M., Herczog, E., Loozen, N., Repanas, K., Russell, K., and Stall, S.: The FAIR Data Maturity
 Model: An Approach to Harmonise FAIR Assessments, *Data Science Journal*, 19, <https://doi.org/10.5334/dsj-2020-041>, 2020.
- Baker, M.: 1,500 scientists lift the lid on reproducibility, *Nature*, 533, 452–454, <https://doi.org/10.1038/533452a>, 2016.
- Barba, L. A.: Reproducibility PI manifesto, https://figshare.com/articles/reproducibility_pi_manifesto/104539, slides for lightning talk at the
 ICERM workshop on "Reproducibility in Computational and Experimental Mathematics", 2012.
- 630 Barba, L. A.: Terminologies for Reproducible Research, <https://doi.org/10.48550/arxiv.1802.03311>, 2018.
- Barnes, N.: Publish your computer code: it is good enough, *Nature*, 467, 753–753, <https://doi.org/10.1038/467753a>, 2010.
- Batini, C., Cappiello, C., Francalanci, C., and Maurino, A.: Methodologies for data quality assessment and improvement, *ACM Computing
 Surveys*, 41, 1–52, <https://doi.org/10.1145/1541880.1541883>, 2009.
- Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, *Nature*, 525, 47–55,
 635 <https://doi.org/10.1038/nature14956>, 2015.
- Beall, J.: Predatory publishers are corrupting open access, *Nature*, 489, 179–179, <https://doi.org/10.1038/489179a>, 2012.
- Belhajjame, K., B'Far, R., Cheney, J., Coppens, S., Cresswell, S., Gil, Y., Groth, P., Klyne, G., Lebo, T., McCusker, J., et al.: Prov-dm: The
 prov data model, *W3C Recommendation*, 14, 15–16, 2013.
- Benureau, F. C. Y. and Rougier, N. P.: Re-run, Repeat, Reproduce, Reuse, Replicate: Transforming Code into Scientific Contributions,
 640 *Frontiers in Neuroinformatics*, 11, <https://doi.org/10.3389/fninf.2017.00069>, 2018.
- Biewald, L.: Experiment Tracking with Weights and Biases, <https://www.wandb.com/>, software available from wandb.com, 2020.
- Birant, D. and Kut, A.: ST-DBSCAN: An algorithm for clustering spatial-temporal data, *Data & Knowledge Engineering*, 60, 208–221,
<https://doi.org/10.1016/j.datak.2006.01.013>, 2007.
- Birch, S. A. and Bloom, P.: The Curse of Knowledge in Reasoning About False Beliefs, *Psychological Science*, 18, 382–386,
 645 <https://doi.org/10.1111/j.1467-9280.2007.01909.x>, 2007.
- Bishop, D.: Rein in the four horsemen of irreproducibility, *Nature*, 568, 435–435, <https://doi.org/10.1038/d41586-019-01307-2>, 2019.
- Bisong, E.: Kubeflow and Kubeflow Pipelines, p. 671–685, Apress, Berkeley, CA, ISBN 9781484244708, https://doi.org/10.1007/978-1-4842-4470-8_46, 2019.



- Blischak, J. D., Carbonetto, P., and Stephens, M.: Creating and sharing reproducible research code the workflowr way, *F1000Research*, 8, 1749, <https://doi.org/10.12688/f1000research.20843.1>, 2019.
- Bourne, P. E., Polka, J. K., Vale, R. D., and Kiley, R.: Ten simple rules to consider regarding preprint submission, *PLOS Computational Biology*, 13, e1005473, <https://doi.org/10.1371/journal.pcbi.1005473>, 2017.
- Brenning, A.: Spatial machine-learning model diagnostics: a model-agnostic distance-based approach, *International Journal of Geographical Information Science*, 37, 584–606, <https://doi.org/10.1080/13658816.2022.2131789>, 2023.
- Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S., Gorelick, N., Zhang, L. L., Alj, S., Schechter, E., Askay, S., Guinan, O., Moore, R., Boukouvalas, A., and Kohli, P.: AlphaEarth Foundations: An embedding field model for accurate and efficient global mapping from sparse label data, <https://doi.org/10.48550/ARXIV.2507.22291>, 2025.
- Callahan, S. P., Freire, J., Santos, E., Scheidegger, C. E., Silva, C. T., and Vo, H. T.: VisTrails: visualization meets data management, in: *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*, edited by SIGMOD, SIGMOD '06, p. 745–747, Association for Computing Machinery, New York, NY, USA, <https://doi.org/10.1145/1142473.1142574>, 2006.
- Challen, R.: dtrackr: An R package for tracking the provenance of data, *Journal of Open Source Software*, 7, 4707, <https://doi.org/10.21105/joss.04707>, 2022.
- Chen, A., Chow, A., Davidson, A., DCunha, A., Ghodsi, A., Hong, S. A., Konwinski, A., Mewald, C., Murching, S., Nykodym, T., Ogilvie, P., Parkhe, M., Singh, A., Xie, F., Zaharia, M., Zang, R., Zheng, J., and Zumar, C.: Developments in MLflow: A System to Accelerate the Machine Learning Lifecycle, in: *Proceedings of the Fourth International Workshop on Data Management for End-to-End Machine Learning*, edited by DEEM, DEEM '20, pp. 1–4, Association for Computing Machinery, New York, NY, USA, <https://doi.org/10.1145/3399579.3399867>, 2020.
- Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. K., and Mahmood, F.: Synthetic data in machine learning for medicine and healthcare, *Nature Biomedical Engineering*, 5, 493–497, <https://doi.org/10.1038/s41551-021-00751-8>, 2021.
- Chen, X., Dallmeier-Tiessen, S., Dasler, R., Feger, S., Fokianos, P., Gonzalez, J. B., Hirvonsalo, H., Kousidis, D., Lavasa, A., Mele, S., Rodriguez, D. R., Šimko, T., Smith, T., Trisovic, A., Trzcinska, A., Tsanaktisidis, I., Zimmermann, M., Cranmer, K., Heinrich, L., Watts, G., Hildreth, M., Lloret Iglesias, L., Lassila-Perini, K., and Neubert, S.: Open is not enough, *Nature Physics*, 15, 113–119, <https://doi.org/10.1038/s41567-018-0342-2>, 2019.
- Choi, Y.-D., Roy, B., Nguyen, J., Ahmad, R., Maghami, I., Nassar, A., Li, Z., Castronova, A. M., Malik, T., Wang, S., and Goodall, J. L.: Comparing containerization-based approaches for reproducible computational modeling of environmental systems, *Environmental Modelling & Software*, 167, 105760, <https://doi.org/10.1016/j.envsoft.2023.105760>, 2023.
- Claerbout, J. F. and Karrenbach, M.: Electronic documents give reproducible research a new meaning, in: *SEG Technical Program Expanded Abstracts 1992*, edited by SEG, pp. 601–604, Society of Exploration Geophysicists, <https://doi.org/10.1190/1.1822162>, 1992.
- Claramunt, C.: Ontologies for geospatial information: Progress and challenges ahead, *Journal of Spatial Information Science*, 20, <https://doi.org/10.5311/josis.2020.20.666>, 2020.
- ClearML: ClearML - Your entire MLOps stack in one open-source tool, <https://clear.ml/>, software available from <http://github.com/allegroai/clearml>, 2024.
- Clementini, E., Sharma, J., and Egenhofer, M. J.: Modelling topological spatial relations: Strategies for query processing, *Computers & Graphics*, 18, 815–822, [https://doi.org/10.1016/0097-8493\(94\)90007-8](https://doi.org/10.1016/0097-8493(94)90007-8), 1994.



- Connolly, A., Hellerstein, J., Alterman, N., Beck, D., Fatland, R., Lazowska, E., Mandava, V., and Stone, S.: Software Engineering Practices in Academia: Promoting the 3Rs—Readability, Resilience, and Reuse, *Harvard Data Science Review*, 5, <https://doi.org/10.1162/99608f92.018bf012>, 2023.
- Daróczy, G. and Wickham, H.: logger: A Lightweight, Modern and Flexible Logging Utility, <https://doi.org/10.32614/cran.package.logger>, 2024.
- Dorman, M., Graser, A., Nowosad, J., and Lovelace, R.: Geocomputation with Python, *geocompx*, <https://py.geocompx.org>, 2024.
- Easterbrook, S. M.: Open code for open science?, *Nature Geoscience*, 7, 779–781, <https://doi.org/10.1038/ngeo2283>, 2014.
- Ehrlinger, L. and Wöß, W.: A Survey of Data Quality Measurement and Monitoring Tools, *Frontiers in Big Data*, 5, <https://doi.org/10.3389/fdata.2022.850611>, 2022.
- Endres, M., Mannarapotta Venugopal, A., and Tran, T. S.: Synthetic Data Generation: A Comparative Study, in: *International Database Engineered Applications Symposium*, edited by IDEAS, IDEAS '22, pp. 94–102, Association for Computing Machinery, New York, NY, USA, <https://doi.org/10.1145/3548785.3548793>, 2022.
- Fanelli, D.: “Positive” Results Increase Down the Hierarchy of the Sciences, *PLoS ONE*, 5, e10068, <https://doi.org/10.1371/journal.pone.0010068>, 2010.
- Fanelli, D.: A theory and methodology to quantify knowledge, *Royal Society Open Science*, 6, 181 055, <https://doi.org/10.1098/rsos.181055>, 2019.
- Fanelli, D., Costas, R., and Ioannidis, J. P. A.: Meta-assessment of bias in science, *Proceedings of the National Academy of Sciences*, 114, 3714–3719, <https://doi.org/10.1073/pnas.1618569114>, 2017.
- Feldman, S. I.: Make — a program for maintaining computer programs, *Software: Practice and Experience*, 9, 255–265, <https://doi.org/10.1002/spe.4380090402>, 1979.
- Feynman, R. P.: Cargo Cult Science, *Engineering and Science*, 37, 10–13, <https://resolver.caltech.edu/CaltechES:37.7.CargoCult>, 1974.
- Filazzola, A. and Lortie, C.: A call for clean code to effectively communicate science, *Methods in Ecology and Evolution*, 13, 2119–2128, <https://doi.org/10.1111/2041-210x.13961>, 2022.
- Fortunato, L. and Galassi, M.: The case for free and open source software in research and scholarship, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379, <https://doi.org/10.1098/rsta.2020.0079>, 2021.
- Fotheringham, A. S., Crespo, R., and Yao, J.: Geographical and Temporal Weighted Regression (GTWR): Geographical and Temporal Weighted Regression, *Geographical Analysis*, 47, 431–452, <https://doi.org/10.1111/gean.12071>, 2015.
- Frank, J., Foster, R., and Pagliari, C.: Open access publishing – noble intention, flawed reality, *Social Science & Medicine*, 317, 115 592, <https://doi.org/10.1016/j.socscimed.2022.115592>, 2023.
- Freedman, L. P., Cockburn, I. M., and Simcoe, T. S.: The Economics of Reproducibility in Preclinical Research, *PLOS Biology*, 13, e1002 165, <https://doi.org/10.1371/journal.pbio.1002165>, 2015.
- Garijo, D. and Gil, Y.: Augmenting PROV with plans in P-PLAN: scientific processes as linked data, in: *Proceedings of the Second International Workshop on Linked Science 2012 - Tackling Big Data*, Boston, MA, USA, November 12, 2012, edited by Kauppinen, T., Pouchard, L. C., and Keßler, C., vol. 951 of *CEUR Workshop Proceedings*, pp. 1–4, CEUR-WS.org, 2012.
- Garijo, D., Alper, P., Belhajjame, K., Corcho, O., Gil, Y., and Goble, C.: Common motifs in scientific workflows: An empirical analysis, *Future Generation Computer Systems*, 36, 338–351, <https://doi.org/10.1016/j.future.2013.09.018>, 2014.
- GDAL/OGR contributors: GDAL/OGR Geospatial Data Abstraction software Library, Open Source Geospatial Foundation, <https://doi.org/10.5281/zenodo.5884351>, 2024.



- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., III, H. D., and Crawford, K.: Datasheets for datasets, *Communications of the ACM*, 64, 86–92, <https://doi.org/10.1145/3458723>, 2021.
- Gentleman, R. and Temple Lang, D.: Statistical Analyses and Reproducible Research, *Journal of Computational and Graphical Statistics*, 16, 1–23, <https://doi.org/10.1198/106186007x178663>, 2007.
- Gewin, V.: Data sharing: An open mind on open data, *Nature*, 529, 117–119, <https://doi.org/10.1038/nj7584-117a>, 2016.
- Goodman, A., Pepe, A., Blocker, A. W., Borgman, C. L., Cranmer, K., Crosas, M., Di Stefano, R., Gil, Y., Groth, P., Hedstrom, M., Hogg, D. W., Kashyap, V., Mahabal, A., Siemiginowska, A., and Slavkovic, A.: Ten Simple Rules for the Care and Feeding of Scientific Data, *PLoS Computational Biology*, 10, e1003542, <https://doi.org/10.1371/journal.pcbi.1003542>, 2014.
- Goodman, S. N., Fanelli, D., and Ioannidis, J. P. A.: What does research reproducibility mean?, *Science Translational Medicine*, 8, <https://doi.org/10.1126/scitranslmed.aaf5027>, 2016.
- Hoffmann, J.: Geospatial Data Science with Julia, <https://doi.org/10.5281/zenodo.10150870>, 2023.
- Hu, X., Zhou, Z., Li, H., Hu, Y., Gu, F., Kersten, J., Fan, H., and Klan, F.: Location Reference Recognition from Texts: A Survey and Comparison, *ACM Computing Surveys*, 56, <https://doi.org/10.1145/3625819>, 2023.
- Hull, D., Wolstencroft, K., Stevens, R., Goble, C., Pocock, M. R., Li, P., and Oinn, T.: Taverna: a tool for building and running workflows of services, *Nucleic acids research*, 34, W729–W732, 2006.
- Hungr, O., Leroueil, S., and Picarelli, L.: The Varnes classification of landslide types, an update, *Landslides*, 11, 167–194, 2014.
- Hunter-Zinck, H., de Siqueira, A. F., Vásquez, V. N., Barnes, R., and Martinez, C. C.: Ten simple rules on writing clean and reliable open-source scientific software, *PLOS Computational Biology*, 17, e1009481, <https://doi.org/10.1371/journal.pcbi.1009481>, 2021.
- Ioannidis, J. P. A.: Why Most Published Research Findings Are False, *PLoS Medicine*, 2, e124, <https://doi.org/10.1371/journal.pmed.0020124>, 2005.
- Ioannidis, J. P. A.: Why Science Is Not Necessarily Self-Correcting, *Perspectives on Psychological Science*, 7, 645–654, <https://doi.org/10.1177/1745691612464056>, 2012.
- Ioannidis, J. P. A.: How to Make More Published Research True, *PLoS Medicine*, 11, e1001747, <https://doi.org/10.1371/journal.pmed.1001747>, 2014.
- Ireland, M., Algarabel, G., Steventon, M., and Munafò, M.: How reproducible and reliable is geophysical research?: A review of the availability and accessibility of data and software for research published in journals, *Seismica*, 2, <https://doi.org/10.26443/seismica.v2i1.278>, 2023.
- Ivanović, M. and Budimac, Z.: An overview of ontologies and data resources in medical domains, *Expert Systems with Applications*, 41, 5158–5166, <https://doi.org/10.1016/j.eswa.2014.02.045>, 2014.
- Ivie, P. and Thain, D.: Reproducibility in Scientific Computing, *ACM Computing Surveys*, 51, 1–36, <https://doi.org/10.1145/3186266>, 2018.
- Joelving, F.: Paper trail, *Science*, 383, 252–255, <https://doi.org/10.1126/science.ado0309>, 2024.
- Kedron, P., Li, W., Fotheringham, S., and Goodchild, M.: Reproducibility and replicability: opportunities and challenges for geospatial research, *International Journal of Geographical Information Science*, 35, 427–445, <https://doi.org/10.1080/13658816.2020.1802032>, 2020.
- Khan, F. Z., Soiland-Reyes, S., Sinnott, R. O., Lonie, A., Goble, C., and Crusoe, M. R.: Sharing interoperable workflow provenance: A review of best practices and their practical application in CWLProv, *GigaScience*, 8, <https://doi.org/10.1093/gigascience/giz095>, 2019.
- Kicinski, M., Springate, D. A., and Kontopantelis, E.: Publication bias in meta-analyses from the Cochrane Database of Systematic Reviews, *Statistics in Medicine*, 34, 2781–2793, <https://doi.org/10.1002/sim.6525>, 2015.



- Klebel, T., Reichmann, S., Polka, J., McDowell, G., Penfold, N., Hindle, S., and Ross-Hellauer, T.: Peer review and preprint policies are unclear at most major journals, *PLOS ONE*, 15, e0239518, <https://doi.org/10.1371/journal.pone.0239518>, 2020.
- Knuth, D. E.: Literate Programming, *The Computer Journal*, 27, 97–111, <https://doi.org/10.1093/comjnl/27.2.97>, 1984.
- Konkol, M., Kray, C., and Pfeiffer, M.: Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study, *International Journal of Geographical Information Science*, 33, 408–429, <https://doi.org/10.1080/13658816.2018.1508687>, 2018.
- Konkol, M., Nüst, D., and Goulier, L.: Publishing computational research - a review of infrastructures for reproducible and transparent scholarly communication, *Research Integrity and Peer Review*, 5, <https://doi.org/10.1186/s41073-020-00095-y>, 2020.
- Korbmacher, M., Azevedo, F., Pennington, C. R., Hartmann, H., Pownall, M., Schmidt, K., Elsherif, M., Breznau, N., Robertson, O., Kalandadze, T., Yu, S., Baker, B. J., O'Mahony, A., Olsnes, J. Ø. S., Shaw, J. J., Gjoneska, B., Yamada, Y., Röer, J. P., Murphy, J., Alzahawi, S., Grinschgl, S., Oliveira, C. M., Wingen, T., Yeung, S. K., Liu, M., König, L. M., Albayrak-Aydemir, N., Lecuona, O., Micheli, L., and Evans, T.: The replication crisis has led to positive structural, procedural, and community changes, *Communications Psychology*, 1, <https://doi.org/10.1038/s44271-023-00003-2>, 2023.
- Koster, J. and Rahmann, S.: Snakemake – a scalable bioinformatics workflow engine, *Bioinformatics*, 28, 2520–2522, <https://doi.org/10.1093/bioinformatics/bts480>, 2012.
- Köster, M., Moors, A., De Houwer, J., Ross-Hellauer, T., Van Nieuwerburgh, I., and Verbruggen, F.: Behavioral Reluctance in Adopting Open Access Publishing: Insights From a Goal-Directed Perspective, *Frontiers in Psychology*, 12, <https://doi.org/10.3389/fpsyg.2021.649915>, 2021.
- Kreuzberger, D., Kühl, N., and Hirschl, S.: Machine Learning Operations (MLOps): Overview, Definition, and Architecture, *IEEE Access*, 11, 31 866–31 879, <https://doi.org/10.1109/access.2023.3262138>, 2023.
- Landau, W.: The targets R package: a dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing, *Journal of Open Source Software*, 6, 2959, <https://doi.org/10.21105/joss.02959>, 2021.
- Landis, S. C., Amara, S. G., Asadullah, K., Austin, C. P., Blumenstein, R., Bradley, E. W., Crystal, R. G., Darnell, R. B., Ferrante, R. J., Fillit, H., Finkelstein, R., Fisher, M., Gendelman, H. E., Golub, R. M., Goudreau, J. L., Gross, R. A., Gubit, A. K., Hesterlee, S. E., Howells, D. W., Huguenard, J., Kelner, K., Koroshetz, W., Krainc, D., Lazic, S. E., Levine, M. S., Macleod, M. R., McCall, J. M., Moxley III, R. T., Narasimhan, K., Noble, L. J., Perrin, S., Porter, J. D., Steward, O., Unger, E., Utz, U., and Silberberg, S. D.: A call for transparent reporting to optimize the predictive value of preclinical research, *Nature*, 490, 187–191, <https://doi.org/10.1038/nature11556>, 2012.
- Lebo, T., Sahoo, S., McGuinness, D., Belhajjame, K., Cheney, J., Corsar, D., Garijo, D., Soiland-Reyes, S., Zednik, S., and Zhao, J.: Prov-o: The prov ontology, <https://www.w3.org/TR/prov-o/>, 2013.
- Leek, J. T. and Peng, R. D.: Reproducible research can still be wrong: Adopting a prevention approach, *Proceedings of the National Academy of Sciences*, 112, 1645–1646, <https://doi.org/10.1073/pnas.1421412111>, 2015.
- Lerner, B., Boose, E., and Perez, L.: Using Introspection to Collect Provenance in R, *Informatics*, 5, 12, <https://doi.org/10.3390/informatics5010012>, 2018.
- Lerner, B., Boose, E., Brand, O., Ellison, A. M., Fong, E., Lau, M., Ngo, K., Pasquier, T., Perez, L. A., Seltzer, M., Sheehan, R., and Wonsil, J.: Making Provenance Work for You, *The R Journal*, 14, 141–159, <https://rjournal.github.io/>, 2023.
- Li, J. and Heap, A. D.: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors, *Ecological Informatics*, 6, 228–241, <https://doi.org/10.1016/j.ecoinf.2010.12.003>, 2011.



- Li, W., Hsu, C.-Y., Wang, S., and Kedron, P.: GeoAI Reproducibility and Replicability: A Computational and Spatial Perspective, *Annals of the American Association of Geographers*, 114, 2085–2103, <https://doi.org/10.1080/24694452.2024.2373787>, 2024.
- 800 List, M., Ebert, P., and Albrecht, F.: Ten Simple Rules for Developing Usable Software in Computational Biology, *PLOS Computational Biology*, 13, e1005265, <https://doi.org/10.1371/journal.pcbi.1005265>, 2017.
- Lovelace, R.: Open source tools for geographic analysis in transport planning, *Journal of Geographical Systems*, 23, 547–578, <https://doi.org/10.1007/s10109-020-00342-2>, 2021.
- Lovelace, R., Nowosad, J., and Muenchow, J.: *Geocomputation with R*, CRC Press, ISBN 9781138304512, <https://r.geocompx.org>, 2019.
- 805 Łukasz, L. and contributors to Black: Black: The uncompromising Python code formatter, <https://black.readthedocs.io/en/stable>, 2024.
- Lush, V., Bastin, L., and Lumsden, J.: Geospatial Data Quality Indicators, in: *Proceedings of the 10th international symposium on spatial accuracy assessment in natural resources and environmental sciences*, edited by Vieira, C., Bogorny, V., and Aquino, A., pp. 121–126, International Spatial Accuracy Research Association, 2012.
- Marks-Anglin, A. and Chen, Y.: A historical review of publication bias, *Research Synthesis Methods*, 11, 725–742, <https://doi.org/10.1002/jrsm.1452>, 2020.
- 810 Marshall-Cook, J. and Farley, M.: The hidden sustainability cost of the reproducibility crisis, *Nature Reviews Physics*, 6, 4–5, <https://doi.org/10.1038/s42254-023-00674-0>, 2024.
- Massey, D. S., Opere, M. A., Wallach, J. D., Ross, J. S., and Krumholz, H. M.: Assessment of Preprint Policies of Top-Ranked Clinical Journals, *JAMA Network Open*, 3, e2011127, <https://doi.org/10.1001/jamanetworkopen.2020.11127>, 2020.
- 815 McNutt, M.: Journals unite for reproducibility, *Science*, 346, 679–679, <https://doi.org/10.1126/science.aaa1724>, 2014.
- Mesirov, J. P.: Accessible Reproducible Research, *Science*, 327, 415–416, <https://doi.org/10.1126/science.1179653>, 2010.
- Meyer, H. and Pebesma, E.: Predicting into unknown space? Estimating the area of applicability of spatial prediction models, *Methods in Ecology and Evolution*, 12, 1620–1633, <https://doi.org/10.1111/2041-210x.13650>, 2021.
- Miksa, T., Simms, S., Mietchen, D., and Jones, S.: Ten principles for machine-actionable data management plans, *PLOS Computational Biology*, 15, e1006750, <https://doi.org/10.1371/journal.pcbi.1006750>, 2019.
- 820 Milledge, D. G., Bellugi, D. G., Watt, J., and Densmore, A. L.: Automated determination of landslide locations after large trigger events: advantages and disadvantages compared to manual mapping, *Natural Hazards and Earth System Sciences*, 22, 481–508, <https://doi.org/10.5194/nhess-22-481-2022>, 2022.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., and Gebru, T.: Model Cards for Model Reporting, in: *Proceedings of the Conference on Fairness, Accountability, and Transparency*, edited by FAT, FAT* '19, pp. 220–229, Association for Computing Machinery, New York, NY, USA, <https://doi.org/10.1145/3287560.3287596>, 2019.
- 825 Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., Forster, J., Lee, S., Twardziok, S. O., Kanitz, A., Wilm, A., Holtgrewe, M., Rahmann, S., Nahnsen, S., and Köster, J.: Sustainable data analysis with Snakemake, *F1000Research*, 10, 33, <https://doi.org/10.12688/f1000research.29032.1>, 2021.
- 830 Moreau, D., Wiebels, K., and Boettiger, C.: Containers for computational reproducibility, *Nature Reviews Methods Primers*, 3, <https://doi.org/10.1038/s43586-023-00236-9>, 2023.
- Morin, A., Urban, J., and Sliz, P.: A Quick Guide to Software Licensing for the Scientist-Programmer, *PLoS Computational Biology*, 8, e1002598, <https://doi.org/10.1371/journal.pcbi.1002598>, 2012.
- Muenchow, J., Schäfer, S., and Krüger, E.: Reviewing qualitative GIS research—Toward a wider usage of open-source GIS and reproducible research practices, *Geography Compass*, 13, <https://doi.org/10.1111/gec3.12441>, 2019.
- 835



- Müller, K. and Walthert, L.: styler: Non-Invasive Pretty Printing of R Code, <https://doi.org/10.32614/cran.package.styler>, 2024.
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., and Ioannidis, J. P. A.: A manifesto for reproducible science, *Nature Human Behaviour*, 1, <https://doi.org/10.1038/s41562-016-0021-2017>.
- 840 Munroe, R.: The Rise of Open Access, *Science*, 342, 58–59, <https://doi.org/10.1126/science.342.6154.58>, 2013.
- Murta, L., Braganholo, V., Chirigati, F., Koop, D., and Freire, J.: noWorkflow: Capturing and Analyzing Provenance of Scripts, in: *Provenance and Annotation of Data and Processes*, edited by Ludäscher, B. and Plale, B., International Provenance and Annotation Workshop (IPAW), p. 71–83, Springer International Publishing, Cham, ISBN 9783319164625, ISSN 1611-3349, https://doi.org/10.1007/978-3-319-16462-5_6, 2015.
- 845 National Academies of Sciences: Reproducibility and Replicability in Science, National Academies Press, ISBN 9780309486163, <https://doi.org/10.17226/25303>, 2019.
- Nature: Reality check on reproducibility, *Nature*, 533, 437–437, <https://doi.org/10.1038/533437a>, 2016.
- Nature Physics Editorial: An ocean of code, *Nature Physics*, 21, 177–177, <https://doi.org/10.1038/s41567-025-02808-w>, 2025.
- Nichols, J. D., Oli, M. K., Kendall, W. L., and Boomer, G. S.: A better approach for dealing with reproducibility and replicability in science, *Proceedings of the National Academy of Sciences*, 118, <https://doi.org/10.1073/pnas.2100769118>, 2021.
- 850 Nickerson, R. S.: Confirmation Bias: A Ubiquitous Phenomenon in Many Guises, *Review of General Psychology*, 2, 175–220, <https://doi.org/10.1037/1089-2680.2.2.175>, 1998.
- Noble, W. S.: A Quick Guide to Organizing Computational Biology Projects, *PLOS Computational Biology*, 5, <https://doi.org/10.1371/journal.pcbi.1000424>, 2009.
- 855 Nosek, B. A., Spies, J. R., and Motyl, M.: Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability, *Perspectives on Psychological Science*, 7, 615–631, <https://doi.org/10.1177/1745691612459058>, 2012.
- Nüst, D. and Pebesma, E.: Practical Reproducibility in Geography and Geosciences, *Annals of the American Association of Geographers*, 111, 1300–1310, <https://doi.org/10.1080/24694452.2020.1806028>, 2020.
- Nüst, D., Granell, C., Hofer, B., Konkol, M., Ostermann, F. O., Sileryte, R., and Cerutti, V.: Reproducible research and GIScience: an evaluation using AGILE conference papers, *PeerJ*, 6, e5072, <https://doi.org/10.7717/peerj.5072>, 2018.
- 860 Nuzzo, R.: How scientists fool themselves – and how they can stop, *Nature*, 526, 182–185, <https://doi.org/10.1038/526182a>, 2015.
- Open Science Collaboration: Estimating the reproducibility of psychological science, *Science*, 349, <https://doi.org/10.1126/science.aac4716>, 2015.
- Osborne, J. M., Bernabeu, M. O., Bruna, M., Calderhead, B., Cooper, J., Dalchau, N., Dunn, S.-J., Fletcher, A. G., Freeman, R., Groen, D., Knapp, B., McInerney, G. J., Mirams, G. R., Pitt-Francis, J., Sengupta, B., Wright, D. W., Yates, C. A., Gavaghan, D. J., Emmott, S., and Deane, C.: Ten Simple Rules for Effective Computational Research, *PLoS Computational Biology*, 10, e1003506, <https://doi.org/10.1371/journal.pcbi.1003506>, 2014.
- 865 Pebesma, E. and Bivand, R.: *Spatial Data Science: With Applications in R*, Chapman and Hall/CRC, ISBN 9780429459016, <https://doi.org/10.1201/9780429459016>, 2023.
- 870 Pebesma, E., Nüst, D., and Bivand, R.: The R software environment in reproducible geoscientific research, *Eos, Transactions American Geophysical Union*, 93, 163–163, <https://doi.org/10.1029/2012eo160003>, 2012.
- Peng, R. D.: Reproducible Research in Computational Science, *Science*, 334, 1226–1227, <https://doi.org/10.1126/science.1213847>, 2011.



- Peng, R. D. and Hicks, S. C.: Reproducible Research: A Retrospective, *Annual Review of Public Health*, 42, 79–93, <https://doi.org/10.1146/annurev-publhealth-012420-105110>, 2021.
- 875 Pimentel, J. F., Freire, J., Murta, L., and Braganholo, V.: A Survey on Collecting, Managing, and Analyzing Provenance from Scripts, *ACM Computing Surveys*, 52, 1–38, <https://doi.org/10.1145/3311955>, 2019a.
- Pimentel, J. F., Murta, L., Braganholo, V., and Freire, J.: A Large-Scale Study About Quality and Reproducibility of Jupyter Notebooks, in: *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)*, edited by MSR, pp. 507–517, IEEE, <https://doi.org/10.1109/msr.2019.00077>, 2019b.
- 880 Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., and Haustein, S.: The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles, *PeerJ*, 6, e4375, <https://doi.org/10.7717/peerj.4375>, 2018.
- Plesser, H. E.: Reproducibility vs. Replicability: A Brief History of a Confused Terminology, *Frontiers in Neuroinformatics*, 11, <https://doi.org/10.3389/fninf.2017.00076>, 2018.
- Pokharel, B., Alvioli, M., and Lim, S.: Assessment of earthquake-induced landslide inventories and susceptibility maps using slope unit-based logistic regression and geospatial statistics, *Scientific Reports*, 11, <https://doi.org/10.1038/s41598-021-00780-y>, 2021.
- 885 Pusztai, L., Hatzis, C., and Andre, F.: Reproducibility of research and preclinical validation: problems and solutions, *Nature Reviews Clinical Oncology*, 10, 720–724, <https://doi.org/10.1038/nrclinonc.2013.171>, 2013.
- Pérez, B., Rubio, J., and Sáenz-Adán, C.: A systematic review of provenance systems, *Knowledge and Information Systems*, 57, 495–543, <https://doi.org/10.1007/s10115-018-1164-3>, 2018.
- 890 Reichman, O. J., Jones, M. B., and Schildhauer, M. P.: Challenges and Opportunities of Open Data in Ecology, *Science*, 331, 703–705, <https://doi.org/10.1126/science.1197962>, 2011.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, *Nature*, 566, 195–204, <https://doi.org/10.1038/s41586-019-0912-1>, 2019.
- Rolf, E., Klemmer, K., Robinson, C., and Kerner, H.: Mission Critical – Satellite Data is a Distinct Modality in Machine Learning, <https://doi.org/10.48550/arxiv.2402.01444>, 2024.
- 895 Ryan, M. J.: Replication in Field Biology: The Case of the Frog-Eating Bat, *Science*, 334, 1229–1230, <https://doi.org/10.1126/science.1214532>, 2011.
- Salatino, A. A., Thanapalasingam, T., Mannocci, A., Osborne, F., and Motta, E.: The Computer Science Ontology: A Large-Scale Taxonomy of Research Areas, in: *The Semantic Web – ISWC 2018*, edited by Vrandečić, D., Bontcheva, K., Suárez-Figueroa, M. C., Presutti, V., Celino, I., Sabou, M., Kaffee, L.-A., and Simperl, E., p. 187–205, Springer International Publishing, Cham, ISBN 9783030006686, ISSN 1611-3349, https://doi.org/10.1007/978-3-030-00668-6_12, 2018.
- 900 Sandve, G. K., Nekrutenko, A., Taylor, J., and Hovig, E.: Ten Simple Rules for Reproducible Computational Research, *PLoS Computational Biology*, 9, e1003285, <https://doi.org/10.1371/journal.pcbi.1003285>, 2013.
- Savage, C. J. and Vickers, A. J.: Empirical Study of Data Sharing by Authors Publishing in PLoS Journals, *PLoS ONE*, 4, e7078, <https://doi.org/10.1371/journal.pone.0007078>, 2009.
- 905 Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., and Brenning, A.: Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data, *Ecological Modelling*, 406, 109–120, <https://doi.org/10.1016/j.ecolmodel.2019.06.002>, 2019.
- Schratz, P., Becker, M., Lang, M., and Brenning, A.: mlr3spatiotempcv: Spatiotemporal Resampling Methods for Machine Learning in R, *Journal of Statistical Software*, 111, 1–36, <https://doi.org/10.18637/jss.v111.i07>, 2024.
- 910



- Schwab, M., Karrenbach, N., and Claerbout, J.: Making scientific computations reproducible, *Computing in Science & Engineering*, 2, 61–67, <https://doi.org/10.1109/5992.881708>, 2000.
- Shimizu, C., Stephen, S., Barua, A., Cai, L., Christou, A., Currier, K., Dalal, A., Fisher, C. K., Hitzler, P., Janowicz, K., et al.: The KnowWhereGraph ontology, *Journal of Web Semantics*, 84, 100 842, 2025.
- 915 Simonsohn, U., Nelson, L. D., and Simmons, J. P.: P-curve: A key to the file-drawer., *Journal of Experimental Psychology: General*, 143, 534–547, <https://doi.org/10.1037/a0033242>, 2014.
- Song, F., Parekh, S., Hooper, L., Loke, Y., Ryder, J., Sutton, A., Hing, C., Kwok, C., Pang, C., and Harvey, I.: Dissemination and publication of research findings: an updated review of related biases, *Health Technology Assessment*, 14, <https://doi.org/10.3310/hta14080>, 2010.
- Stark, P. B.: Before reproducibility must come preproducibility, *Nature*, 557, 613–613, <https://doi.org/10.1038/d41586-018-05256-0>, 2018.
- 920 Steger, S., Brenning, A., Bell, R., Petschko, H., and Glade, T.: Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps, *Geomorphology*, 262, 8–23, <https://doi.org/10.1016/j.geomorph.2016.03.015>, 2016.
- Stevenson, M. J., Jackson, C. A.-L., Hall, M., Ireland, M. T., Munafo, M., and Roberts, K. J.: Reproducibility in Subsurface Geoscience, *Earth Science, Systems and Society*, 2, <https://doi.org/10.3389/esss.2022.10051>, 2022.
- 925 Stodden, V., McNutt, M., Bailey, D. H., Deelman, E., Gil, Y., Hanson, B., Heroux, M. A., Ioannidis, J. P., and Taufer, M.: Enhancing reproducibility for computational methods, *Science*, 354, 1240–1241, <https://doi.org/10.1126/science.aah6168>, 2016.
- Stodden, V., Leisch, F., and Peng, R. D.: Implementing Reproducible Research, Chapman and Hall/CRC, ISBN 9781315373461, <https://doi.org/10.1201/9781315373461>, 2018.
- Stoudt, S., Vásquez, V. N., and Martinez, C. C.: Principles for data analysis workflows, *PLOS Computational Biology*, 17, e1008 770, <https://doi.org/10.1371/journal.pcbi.1008770>, 2021.
- 930 Tenopir, C., Christian, L., Allard, S., and Borycz, J.: Research Data Sharing: Practices and Attitudes of Geophysicists, *Earth and Space Science*, 5, 891–902, <https://doi.org/10.1029/2018ea000461>, 2018.
- Thompson, M. M.: Upside-Down GIS: The Future of Citizen Science and Community Participation, *The Cartographic Journal*, 53, 326–334, <https://doi.org/10.1080/00087041.2016.1243863>, 2016.
- 935 Thornton, A.: Publication bias in meta-analysis its causes and consequences, *Journal of Clinical Epidemiology*, 53, 207–216, [https://doi.org/10.1016/s0895-4356\(99\)00161-4](https://doi.org/10.1016/s0895-4356(99)00161-4), 2000.
- Van Lissa, C. J., Brandmaier, A. M., Brinkman, L., Lamprecht, A.-L., Peikert, A., Struiksmá, M. E., and Vreede, B. M.: WORCS: A workflow for open reproducible code in science, *Data Science*, 4, 29–49, <https://doi.org/10.3233/ds-210031>, 2021.
- Van Noorden, R.: Open access: The true cost of science publishing, *Nature*, 495, 426–429, <https://doi.org/10.1038/495426a>, 2013.
- 940 Van Noorden, R.: How big is science’s fake-paper problem?, *Nature*, 623, 466–467, <https://doi.org/10.1038/d41586-023-03464-x>, 2023.
- van Wesel, M.: Evaluation by Citation: Trends in Publication Behavior, Evaluation Criteria, and the Strive for High Impact Publications, *Science and Engineering Ethics*, 22, 199–225, <https://doi.org/10.1007/s11948-015-9638-0>, 2015.
- Vassilev, B., Louhimo, R., Ikonen, E., and Hautaniemi, S.: Language-Agnostic Reproducible Data Analysis Using Literate Programming, *PLOS ONE*, 11, e0164 023, <https://doi.org/10.1371/journal.pone.0164023>, 2016.
- 945 Warmerdam, F.: The Geospatial Data Abstraction Library, p. 87–104, Springer Berlin Heidelberg, ISBN 9783540748311, https://doi.org/10.1007/978-3-540-74831-1_5, 2008.
- Wasserstein, R. L., Schirm, A. L., and Lazar, N. A.: Moving to a World Beyond “ $p < 0.05$ ”, *The American Statistician*, 73, 1–19, <https://doi.org/10.1080/00031305.2019.1583913>, 2019.



- Wicherts, J. M., Borsboom, D., Kats, J., and Molenaar, D.: The poor availability of psychological research data for reanalysis., *American Psychologist*, 61, 726–728, <https://doi.org/10.1037/0003-066x.61.7.726>, 2006.
- Wicherts, J. M., Bakker, M., and Molenaar, D.: Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results, *PLoS ONE*, 6, e26 828, <https://doi.org/10.1371/journal.pone.0026828>, 2011.
- Wickham, H.: Tidy Data, *Journal of Statistical Software*, 59, <https://doi.org/10.18637/jss.v059.i10>, 2014.
- Wikle, C. K.: Modern perspectives on statistics for spatio-temporal data, *WIREs Computational Statistics*, 7, 86–98, <https://doi.org/10.1002/wics.1341>, 2014.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data*, 3, <https://doi.org/10.1038/sdata.2016.18>, 2016.
- 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, e1001 745, <https://doi.org/10.1371/journal.pbio.1001745>, 2014.
- Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., and Teal, T. K.: Good enough practices in scientific computing, *PLOS Computational Biology*, 13, e1005 510, <https://doi.org/10.1371/journal.pcbi.1005510>, 2017.
- Wilson, J. P., Butler, K., Gao, S., Hu, Y., Li, W., and Wright, D. J.: A Five-Star Guide for Achieving Replicability and Reproducibility When Working with GIS Software and Algorithms, *Annals of the American Association of Geographers*, 111, 1311–1317, <https://doi.org/10.1080/24694452.2020.1806026>, 2020.
- Wimmer, S. and Finger, R.: A note on synthetic data for replication purposes in agricultural economics, *Journal of Agricultural Economics*, 74, 316–323, <https://doi.org/10.1111/1477-9552.12505>, 2022.
- Xie, Y.: Dynamic Documents with R and knitr, Chapman and Hall/CRC, <https://doi.org/10.1201/9781315382487>, 2015.
- Xie, Y., Allaire, J. J., and Golemund, G.: R Markdown: The Definitive Guide, Chapman and Hall/CRC, ISBN 9781138359444, <https://doi.org/10.1201/9781138359444>, 2018.
- Xiong, C., Van Weelden, L., and Franconeri, S.: The Curse of Knowledge in Visual Data Communication, *IEEE Transactions on Visualization and Computer Graphics*, 26, 3051–3062, <https://doi.org/10.1109/tvcg.2019.2917689>, 2020.
- Yan, A., Huang, C., Lee, J., and Palmer, C. L.: Cross-disciplinary data practices in earth system science: Aligning services with reuse and reproducibility priorities, *Proceedings of the Association for Information Science and Technology*, 57, <https://doi.org/10.1002/pra2.218>, 2020.
- Zaharia, M. A., Chen, A., Davidson, A., Ghodsi, A., Hong, S. A., Konwinski, A., Murching, S., Nykodym, T., Ogilvie, P., Parkhe, M., Xie, F., and Zumar, C.: Accelerating the Machine Learning Lifecycle with MLflow, *IEEE Data Eng. Bull.*, 41, 39–45, 2018.