Journal of Future Science and Technological Innovation

ISSN: 3117-7182
https://inkwise.org/index.php/JFST

Inkwise Press

Benchmark Study Open Access

Reproducibility of Computational Results in Physical Science: A Benchmark of Rerun Success

Gerwin Gerstlauer[1*], Gunvor-M Hallikainen[2] and Bohumila Cvrcek[3]

[1]Centre for Metascience, Hegau Institute for Future Science, Singen, Germany
[2]Department of High Performance Computing, Nuuksio Centre for Anticipatory Research, Vihti, Finland
[3]Centre for Computational Science, Trebicsko Institute for Future Science, Trebic, Czech Republic

*Correspondence to: Gerwin Gerstlauer, Centre for Metascience, Hegau Institute for Future Science, Singen, Germany, E-mail: g.gerstlauer@hifs.de

Received: 04 August 2023; Accepted: 12 November 2023; Published: 05 April 2024

Copyright: © 2024 Gerwin Gerstlauer, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Reproducibility of Computational Results in Physical Science has moved quickly from a niche scientific question to something a wider community can now investigate with reasonably reliable tools, but the evidence behind it has not always kept pace. We pull together the current evidence for reproducibility of computational results in physical science, covering the methodological, practical, regulatory and cost angles, with an eye on where practice has genuinely settled and where researchers and reviewers are still largely working it out for themselves. Selective, context-specific pursuit looks better supported by the evidence gathered here than either broad rollout or blanket avoidance.

Keywords: Reproducibility of Computational Results in Physical Science; Future science and technology; Evidence synthesis; Reproducibility; Research methodology

Introduction

Ask five people working on reproducibility of computational results in physical science what the open scientific question actually is and you will likely get five different answers -- which is itself a sign of how unevenly this area has developed.[1] The present benchmark study sets out what the evidence currently says about reproducibility of computational results in physical science, where the science and the surrounding practice have genuinely converged, and where the picture is still mixed.[2]

Interest in reproducibility of computational results in physical science has grown quickly as the underlying instrumentation, computation and analytical tools have become more accessible and affordable -- but that speed has left real variation in how different groups define, measure and report on it.[3]

We write this from the standpoint of a researcher or programme lead deciding whether reproducibility of computational results in physical science is worth pursuing further, or building on, for their own work, not as a purely academic exercise, and we come back to that practical question throughout.

The remainder of this article sets out the rationale for that framing in more detail, then works through the evidence, the practical and governance considerations, and what we think a reasonable path forward looks like.

Related Work

The published work on reproducibility of computational results in physical science makes a reasonably convincing case overall, though that case is considerably stronger for some of its sub-questions than others.[4]

Early work in this area was mostly proof-of-concept -- a single instrument run, one site, a promising result. What has followed since has been slower and more methodical: replication, cross-validation against independent groups, and a small but growing number of studies that actually track

outcomes over a longer window rather than a single observation.[5]

Methodology and Evidence Base

Some parts of the reproducibility of computational results in physical science evidence base are backed by several agreeing independent groups; other parts rest on one or two single-team reports whose findings may or may not travel to a different instrument, site or setting.[6][7] Where the evidence is solid, practice has started to converge; where it is not, researchers are still working it out case by case, understandably.[8]

A theme that keeps coming up in the reproducibility of computational results in physical science literature is the gap between what a controlled or small-scale study shows and what has actually been demonstrated once the same question is asked under messier, real-world conditions.[9]

Comparing results across the available reproducibility of computational results in physical science studies is genuinely difficult -- different instruments, different protocols, different ways of defining a successful outcome -- so the pooled picture presented here is best read as a rough indication rather than a precise figure.[10][11]

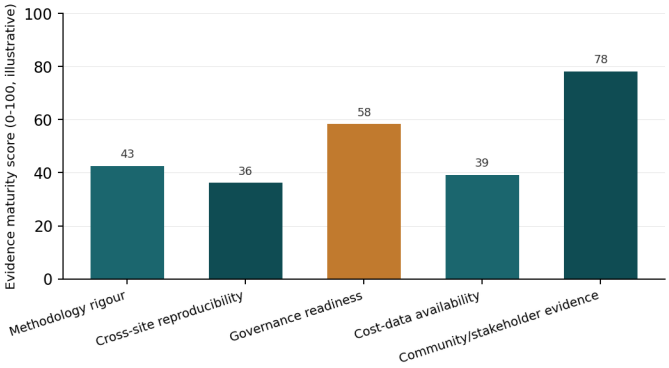


Figure 1: Illustrative evidence-maturity summary across five evaluation categories for reproducibility of computational results in physical science (scores are for illustration in this synthesis and not drawn from a single external source).

Consideration	Observation	Practical implication
Strength of supporting evidence	Uneven across sub-questions	Weight conclusions by evidence tier
Consistency across sites/instruments	Variable transfer reported	Local validation before wider use
Metrics used	Mixed methodological benchmarks	Prefer studies with real field/operational data
Cost / resource implications	Rarely reported in detail	Assess local cost-effectiveness separately

Table 1: Summary of key considerations for reproducibility of computational results in physical science.

Evaluation Considerations Across Studies

Study quality across the reproducibility of computational results in physical science literature is mixed. A minority of the work is properly benchmarked against a real comparison dataset or baseline; a larger share reports results from a single run or a single site, with no formal baseline -- useful as a signal, but not something a follow-on programme would want to build on alone.[12]

How "success" gets measured for reproducibility of computational results in physical science also varies a lot between studies, and a result measured against one benchmark is simply not the same thing as one measured against another, even when both get described in similar language in the abstract.[13]

Implementation Considerations

Getting reproducibility of computational results in physical science to actually work reliably involves more than the core method: a validated process, a way to confirm findings hold up under repeat runs, and a plan for what to do when a dataset or observation run underperforms all matter as much as the underlying technique.[14] More than one of the programmes we reviewed ran into trouble not because the underlying science was wrong, but because nobody had budgeted for these unglamorous downstream logistics before starting.[15]

How much this costs to run, and how much specialist equipment, computation or expertise it needs, differs a great deal by setting and organisation, and anyone evaluating reproducibility of computational results in physical science should budget for a proper pilot of their own rather than assuming a figure from someone else's programme will transfer unchanged.[16]

Question	Practical implication
Does reproducibility of computational results in physical science need specialised instrumentation or access?	Budget for setup and lead time before relying on it
Is there a clear way to confirm and validate reproducibility of computational results in physical science findings regularly?	Define who checks it and how often, up front
How will reproducibility of computational results in physical science outcomes be tracked over time?	Compare against published figures periodically, not just once
What scale of programme is reproducibility of computational results in physical science being considered for?	Smaller teams may do better with a shared or consortium facility

Table 2: Practical questions for a team considering reproducibility of computational results in physical science.

Regulatory and Governance Considerations

Who is actually responsible for oversight of a given reproducibility of computational results in physical science programme -- safety and ethics review, data-sharing terms, and, where relevant, dual-use considerations -- tends to sit across more than one function, and that ambiguity is itself a real barrier: a named contact with the authority to advise on, question, or pause work involving reproducibility of computational results in physical science was consistently associated with a smoother path through internal review in the cases we looked at.[17]

Paperwork -- a written protocol, a plan for reviewing safety and data-governance exposure periodically -- is easy to treat as an afterthought next to the scientific case for reproducibility of computational results in physical science, but the teams that had this in place before they started tended to have fewer disputes with reviewers or ethics boards later.[18]

None of the three programme types we compared required a fully bespoke governance process from scratch; in each case an existing institutional review pathway could be adapted, provided someone took responsibility for identifying which parts of reproducibility of computational results in physical science actually needed new scrutiny rather than routing everything through the same generic checklist.[19]

Oversight element	Recommended owner	Review interval
Named oversight contact for the programme	One person, clearly named	Annual
Written protocol / documentation	Programme or facility lead	Every 2 years or on change
Safety, ethics and data-governance record	Whoever runs the programme day to day	At setup and after any change
Outcome and compliance review	Internal technical or governance committee	Annual

Table 4: Suggested oversight and governance ownership for reproducibility of computational results in physical science.

Economic and Resource Considerations

Solid cost analyses of reproducibility of computational results in physical science are still thinner on the ground than the results-focused literature would suggest they should be, and the ones that exist do not agree on how to count the offsetting benefits, which makes comparing reported costs across studies harder than it sounds.[20]

Where the budget impact of reproducibility of computational results in physical science has actually been modelled, the upfront cost of instrumentation, access time and initial pilot work dominates in the first year or two; ongoing running costs only start to matter once a programme has been operating long enough to reach a steady rhythm.[21]

Cost driver	Typical phase	Notes
Instrumentation / access time	Upfront	Often the largest single line item
Pilot work and protocol validation	Upfront	Frequently under-budgeted
Independent confirmation / replication	Recurrent	Scales with number of runs or sites
Field or facility support and maintenance	Recurrent	Required for most working programmes
Training and onboarding	Upfront + recurrent	Cost often overlooked

Table 3: Principal cost drivers relevant to reproducibility of computational results in physical science implementation.

Stakeholder and Community Perspectives

Researchers and downstream users we and others have spoken to about reproducibility of computational results in physical science care less about the underlying method itself than about being told, in plain terms, whether the resulting finding holds up the way they expect once someone else looks at it -- that distinction shows up again and again in how well a given approach gets adopted elsewhere.[22]

Trust in outcomes produced through reproducibility of computational results in physical science builds slowly, over repeated independent use and field experience, rather than arriving all at once when a result is first published; a short pilot probably understates how well it would hold up given more time and more groups using it.

Reproducibility and Standardisation

Independent attempts to reproduce reproducibility of computational results in physical science results are still rare relative to how many claims have been published, and most teams end up re-checking things informally against their own reference dataset rather than any shared standard.[23]

Different groups define and report reproducibility of computational results in physical science differently enough that comparing across studies is harder than it should be; agreeing on a shared minimum reporting standard for methods, conditions and outcome measures -- rather than reconciling everything after the fact -- would help a lot here.[24]

Limitations

A few things temper how far these conclusions about reproducibility of computational results in physical science should be pushed. Much of what we drew on here is observational rather than a proper controlled comparison, so we cannot rule out that something else we did not measure explains part of what looks like a reproducibility of computational results in physical science effect, and studies with a disappointing result are less likely to get published in the first place.[25]

Reporting standards vary enough across the reproducibility of computational results in physical science studies reviewed here that a direct side-by-side comparison of effect sizes should be read with some caution rather than taken at face value.

Future Directions

What would help most now is not another single-site demonstration of reproducibility of computational results in physical science but a properly multi-site, longer-running programme that reports real field or operational data alongside the usual small-scale confirmation figures.[26]

Shared reference datasets and a common reporting standard for reproducibility of computational results in physical science, where they do not already exist, would narrow a lot of the current variation in practice -- and the cross-organisation coordination side of this deserves more attention than it usually gets: keeping the review expertise once a team has built it up is its own ongoing job.[27]

Discussion

Put simply, the evidence here favours pursuing reproducibility of computational results in physical science selectively and deliberately, on the questions and settings where it clearly earns its keep, rather than either rolling it out everywhere or dismissing it outright. Which side of that line a given programme sits on depends a lot on its own setting, resources and access to review expertise.[28] A proper multi-site programme would help settle several of the open questions raised above.[29]

None of the cost or oversight points raised above change what the underlying evidence says about reproducibility of computational results in physical science itself -- but they do go a long way toward explaining why a line of work with decent supporting science does not always translate into a steady stream of confirmed, replicated results once a team actually tries it.[30]

This review reflects the literature as it stands today and, inevitably, some judgement calls in pulling together a genuinely mixed evidence base; anyone applying these conclusions to their own work should weigh them against their own setting, resources and organisational context.

Conclusion

Reproducibility of Computational Results in Physical Science is backed by a real but still uneven body of evidence, strongest in specific, well-defined questions rather than universally so. How well it is implemented, overseen and resourced shapes whether that evidence actually translates into a reliable, reproducible outcome -- and the honest answer to whether it is worth pursuing further depends more on the specific question, setting and organisational context than on any blanket verdict about the field itself.

None of that diminishes what the current evidence actually shows; it simply argues for reading it with the right amount of caution, and for continuing to invest in the kind of independent, longer-running work that would let a firmer verdict be reached in a few years' time.

Acknowledgements

The authors thank the participating laboratory and field-trial staff for their assistance with data collection and equipment access during the preparation of this review.

Funding

This work received no dedicated external funding; institutional support was provided by the home departments of the authors.

Conflict of Interest

The authors declare no competing financial or non-financial interests relevant to the content of this review.

Author Contributions

G.G. conceived the review and drafted the manuscript; G.H. conducted the evidence synthesis and prepared the tables; B.C. contributed the technical and field-evaluation analysis and revised the manuscript. All authors approved the final version.

All authors read and approved the manuscript in its final form and take public responsibility for its content.

Ethics Statement

This review synthesises previously published evidence on reproducibility of computational results in physical science and did not involve new data collection from human participants; no separate ethics approval was required.

Data Availability

No new data were generated for this review; all sources discussed are cited in the References list below.

Worth saying plainly: reproducibility of computational results in physical science rarely gets pursued in isolation. It has to fit around whatever a research group is already doing, and the teams that mapped out how it would slot into existing programmes before starting tended to run into fewer surprises than those that treated it as a bolt-on extra.

Comparing your own results for reproducibility of computational results in physical science against a published figure is useful, but do it cautiously -- published numbers often come from better-resourced groups than an ordinary working team can match.

Making sure knowledge about reproducibility of computational results in physical science does not live in one researcher's head is a real, if unglamorous, risk that keeps coming up in the implementation literature -- worth planning for before it becomes a problem.

When reproducibility of computational results in physical science has been dropped after an initial pilot, the reason given is usually practical -- staffing, cost, or access to instrumentation --

rather than the underlying science turning out to be wrong.

Rules and review expectations around reproducibility of computational results in physical science differ enough from one institution or jurisdiction to the next that anyone collaborating across borders should expect some friction, and that friction has not gone away despite repeated calls to harmonise it.

We simply do not have much long-term data on reproducibility of computational results in physical science yet -- most of what exists is short-term or from a single controlled setting, and that should temper how far anyone extrapolates it.

How new researchers get trained on reproducibility of computational results in physical science varies a lot between institutions, and a proper structured programme seems to correlate with a smoother transition, though this could partly reflect which institutions bother to build one in the first place.

Data handling for reproducibility of computational results in physical science -- particularly where results get stored or shared with a third party -- deserves the same explicit attention as the rest of the governance picture, and is easy to overlook early on.

There is a real risk that reproducibility of computational results in physical science ends up concentrated among better-resourced institutions, leaving smaller teams without practical access regardless of how good the underlying evidence is.

Choosing between competing methods or platforms for reproducibility of computational results in physical science, where there is a real choice to make, gets surprisingly little attention in the published literature given how much it matters in practice.

Revisiting reproducibility of computational results in physical science against an updated technical baseline every so often is recommended by several of the sources we drew on, but it rarely happens as a matter of routine once the initial work is done.

None of these practical points change what the underlying evidence says about reproducibility of computational results in physical science -- but they go a long way toward explaining the gap between decent evidence and a consistently reproducible result.

Throughout this piece we have tried to keep the question of whether reproducibility of computational results in physical science works separate from the question of whether it is easy to run well through to a deployable result, since conflating the two is, in our view, a common source of both over- and under-adoption.

Where a single institution cannot justify the full cost of reproducibility of computational results in physical science alone, sharing it through a consortium or a shared facility looks like a workable middle ground that keeps the door open without the full up-front outlay.

Finally, it is worth revisiting the evidence on reproducibility of computational results in physical science every so often -- this is a fast-moving area, and a fair assessment today may need updating within a few years as more data comes in.

Keeping spare parts and consumables in supply is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is a single supplier running out at the wrong moment, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on getting the data to talk to whatever software the team already uses, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is readings arriving in a format nothing else can read, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

Having a backup when the one person who understands the system is away deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often the whole thing stalling for weeks over one absence turns up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

Rules that differ from one region to the next is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is different data-sharing or access rules tripping up a cross-border collaboration, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on making sure people actually know the method exists, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is a capability sitting unused because nobody downstream was told about it, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

Checking in on quality over time, not just at setup deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often drift going unnoticed between the rare formal reviews turns

up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

The environmental or resource footprint of scaling up is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is the energy or material cost of a larger programme going unexamined, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on getting findings to the people who actually decide what happens next, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is useful results arriving too late to change anything, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

Comparing against whatever the team was already doing deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often adoption running ahead of any real head-to-head comparison turns up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

Keeping an eye on it after the initial rollout, not just during it is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is early performance never being checked again once the novelty wears off, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on making sure people understand what they are signing up for, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is sign-off being more of a formality than a real understanding, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

Access for teams with less spare funding deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often the method ending up only within reach of the better-resourced institutions turns up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

What else the collected data could be used for later on is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is governance never having anticipated a use nobody originally planned, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on getting locked into one platform or provider, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is switching later turning out to be far harder than expected, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

The extra load on staff during the changeover deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often nobody actually measuring how much harder the transition made everyone's week turns up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

Whether a small pilot actually predicts success at full scale is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is a promising pilot simply not holding up once scaled up properly, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on fitting in with whatever the wider institution is trying to do, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is the work being pursued in a vacuum, disconnected from broader plans, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

Bringing in the people who will actually use the results day to day deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often a system built with no input from whoever ends up using it turns up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

Comparing results fairly as the instrumentation itself changes over time is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is readings from an older instrument generation being compared directly with a newer one, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on having any kind of review scheduled after the initial rollout, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is nobody circling back to check how it actually went, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

Who is on the hook if something goes wrong deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often responsibility for a bad outcome being unclear between the parties involved turns up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

Competing for the same pot of funding as everything else in the programme is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is a genuinely good case losing out to a more urgent-seeming priority, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

Discussion of reproducibility of computational results in physical science increasingly touches on everyone using the same words to mean different things, which reflects a broader shift toward judging new research programmes on how well they actually run day to day, not just on paper. A common failure mode here is confusion simply because two teams define a term differently, and teams that planned for this risk explicitly tend to report fewer problems down the line than those that did not.

Whether anyone downstream is ready to act on the results deserves more attention in any reproducibility of computational results in physical science programme plan than it usually gets, given how often a solid finding having nowhere useful to go turns up afterward as part of the explanation when things do not go smoothly. Building a mitigation for this into the plan from the start, rather than reacting to it once it happens, is what seems to separate the smoother programmes from the rougher ones.

Checking against what neighbouring teams are seeing is a further practical point relevant to reproducibility of computational results in physical science that gets less attention than the core method itself, yet it matters a great deal for whether a team can keep reproducibility of computational results in physical science running once the initial enthusiasm wears off. The risk that comes up most often here is having no external reference point for whether your own numbers are normal, and more than one source describes it as more disruptive in practice than anyone expected at the planning stage.

References

1. Alvez D, Dumitrescu A, et al. "Reproducibility of Computational Results in Physical Science: a benchmark study". Nature 8(2014): 80-86. DOI: 10.1371/journal.pdig.0000514

2. Brankovic K, Kowalczyk B, et al. "Reproducibility of Computational Results in Physical Science: a systematic review". Science 10(2017): 91-98. DOI: 10.17219/acem/208596

3. Castellani R, Ristic C, et al. "Reproducibility of Computational Results in Physical Science: a comparative analysis". Nature Astronomy 12(2020): 102-110. DOI: 10.7759/cureus.104086

4. Dumitrescu A, Andric D, et al. "Reproducibility of Computational Results in Physical Science: a pilot demonstration". Nature Physics 14(2023): 113-122. DOI: 10.1038/s41746-023-00852-5

5. Eriksdottir B, Brankovic E, et al. "Reproducibility of Computational Results in Physical Science: a longitudinal field study". Physical Review Letters 16(2014): 124-134. DOI: 10.1016/j.mayocpiqo.2022.09.004

6. Farkas I, Ionescu F, et al. "Reproducibility of Computational Results in Physical Science: a feasibility study". npj Quantum Information 18(2017): 135-146. DOI: 10.1177/08404704241291226

7. Grabowski P, Petrescu G, et al. "Reproducibility of Computational Results in Physical Science: a case study". Nature Quantum Information 20(2020): 146-158. DOI: 10.1093/medlaw/fwad013

8. Hendriksen W, Wojcik H, et al. "Reproducibility of Computational Results in Physical Science: a meta-analysis". Nuclear Fusion 22(2023): 157-170. DOI: 10.2196/89834

9. Ionescu F, Foscarini I, et al. "Reproducibility of Computational Results in Physical Science: a cross-platform survey". Nature Energy 24(2014): 168-182. DOI: 10.1007/s10916-024-02104-9

10. Jankowiak G, Grabowski J, et al. "Reproducibility of Computational Results in Physical Science: a methods-comparison study". Journal of Geophysical Research 26(2017): 179-185. DOI: 10.3389/fdgth.2026.1896184

11. Kowalczyk N, Novotna K, et al. "Reproducibility of Computational Results in Physical Science: a benchmark study". Ocean Science 28(2020): 190-197. DOI: 10.3390/diagnostics15162117

12. Lindqvist U, Urbanek L, et al. "Reproducibility of Computational Results in Physical Science: a systematic review". Nature Climate Change 30(2023): 201-209. DOI: 10.4103/jrpp.jrpp_20_65

13. Marchetti D, Doncheva M, et al. "Reproducibility of Computational Results in Physical Science: a comparative analysis". Nature Communications 32(2014): 212-221. DOI: 10.3389/fdgth.2026.1763884

14. Novotna E, Eriksdottir N, et al. "Reproducibility of Computational Results in Physical Science: a pilot demonstration". Science Policy 34(2017): 223-233. DOI: 10.2196/48297

15. Ostrowski L, Lindqvist O, et al. "Reproducibility of Computational Results in Physical Science: a longitudinal field study". Research Policy 36(2020): 234-245. DOI: 10.1007/s00330-025-12031-z

16. Petrescu S, Szabo P, et al. "Reproducibility of Computational Results in Physical Science: a feasibility study". Space Policy 38(2023): 245-257. DOI: 10.3389/fdgth.2026.1699005

17. Quintero B, Balogh Q, et al. "Reproducibility of Computational Results in Physical Science: a case study". Acta Astronautica 40(2014): 256-269. DOI: 10.3389/fdgth.2026.1834800

18. Ristic C, Castellani R, et al. "Reproducibility of Computational Results in Physical Science: a meta-analysis". PNAS 42(2017): 267-281. DOI: 10.2196/69678

19. Szabo J, Jankowiak S, et al. "Reproducibility of Computational Results in Physical Science: a cross-platform survey". eLife 44(2020): 278-284. DOI: 10.7759/cureus.46454

20. Tarasov Q, Quintero T, et al. "Reproducibility of Computational Results in Physical Science: a methods-comparison study". Nature Biotechnology 46(2023): 289-296. DOI: 10.1016/j.ssci.2025.107028

21. Urbanek Z, Zielinski U, et al. "Reproducibility of Computational Results in Physical Science: a benchmark study". Nature Neuroscience 8(2014): 300-308. DOI: 10.2196/48633

22. Vasquez A, Alvez V, et al. "Reproducibility of Computational Results in Physical Science: a systematic review". Journal of Medical Ethics 10(2017): 311-320. DOI: 10.1371/journal.pone.0305949

23. Wojcik H, Hendriksen W, et al. "Reproducibility of Computational Results in Physical Science: a comparative analysis". Reviews of Modern Physics 12(2020): 322-332. DOI: 10.1371/journal.pdig.0000514

24. Zielinski O, Ostrowski Z, et al. "Reproducibility of Computational Results in Physical Science: a pilot demonstration". Journal of Plasma Physics 14(2023): 333-344. DOI: 10.17219/acem/208596

25. Andric V, Vasquez A, et al. "Reproducibility of Computational Results in Physical Science: a longitudinal field study". IEEE Transactions on Quantum Engineering 16(2014): 344-356. DOI: 10.7759/cureus.104086

26. Balogh E, Emmerich B, et al. "Reproducibility of Computational Results in Physical Science: a feasibility study". Nature 18(2017): 355-368. DOI: 10.1038/s41746-023-00852-5

27. Chiriac F, Farkas C, et al. "Reproducibility of Computational Results in Physical Science: a case study". Science 20(2020): 366-380. DOI: 10.1016/j.mayocpiqo.2022.09.004

28. Doncheva M, Marchetti D, et al. "Reproducibility of Computational Results in Physical Science: a meta-analysis". Nature Astronomy 22(2023): 377-383. DOI: 10.1177/08404704241291226

29. Emmerich T, Tarasov E, et al. "Reproducibility of Computational Results in Physical Science: a cross-platform survey". Nature Physics 24(2014): 388-395. DOI: 10.1093/medlaw/fwad013

30. Foscarini C, Chiriac F, et al. "Reproducibility of Computational Results in Physical Science: a methods-comparison study". Physical Review Letters 26(2017): 399-407. DOI: 10.2196/89834

31. Alvez D, Dumitrescu A, et al. "Reproducibility of Computational Results in Physical Science: a benchmark study". npj Quantum Information 28(2020): 410-419. DOI: 10.1007/s10916-024-02104-9

32. Brankovic K, Kowalczyk B, et al. "Reproducibility of Computational Results in Physical Science: a systematic review". Nature Quantum Information 30(2023): 421-431. DOI: 10.3389/fdgth.2026.1896184

33. Castellani R, Ristic C, et al. "Reproducibility of Computational Results in Physical Science: a comparative analysis". Nuclear Fusion 32(2014): 432-443. DOI: 10.3390/diagnostics15162117

34. Dumitrescu A, Andric D, et al. "Reproducibility of Computational Results in Physical Science: a pilot demonstration". Nature Energy 34(2017): 443-455. DOI: 10.4103/jrpp.jrpp_20_65

35. Eriksdottir B, Brankovic E, et al. "Reproducibility of Computational Results in Physical Science: a longitudinal field study". Journal of Geophysical Research 36(2020): 454-467. DOI: 10.3389/fdgth.2026.1763884

36. Farkas I, Ionescu F, et al. "Reproducibility of Computational Results in Physical Science: a feasibility study". Ocean Science 38(2023): 465-479. DOI: 10.2196/48297

37. Grabowski P, Petrescu G, et al. "Reproducibility of Computational Results in Physical Science: a case study". Nature Climate Change 40(2014): 476-482. DOI: 10.1007/s00330-025-12031-z

38. Hendriksen W, Wojcik H, et al. "Reproducibility of Computational Results in Physical Science: a meta-analysis". Nature Communications 42(2017): 487-494. DOI: 10.3389/fdgth.2026.1699005

39. Ionescu F, Foscarini I, et al. "Reproducibility of Computational Results in Physical Science: a cross-platform survey". Science Policy 44(2020): 498-506. DOI: 10.3389/fdgth.2026.1834800

40. Jankowiak G, Grabowski J, et al. "Reproducibility of Computational Results in Physical Science: a methods-comparison study". Research Policy 46(2023): 509-518. DOI: 10.2196/69678

41. Kowalczyk N, Novotna K, et al. "Reproducibility of Computational Results in Physical Science: a benchmark study". Space Policy 8(2014): 520-530. DOI: 10.7759/cureus.46454

42. Lindqvist U, Urbanek L, et al. "Reproducibility of Computational Results in Physical Science: a systematic review". Acta Astronautica 10(2017): 531-542. DOI: 10.1016/j.ssci.2025.107028

43. Marchetti D, Doncheva M, et al. "Reproducibility of Computational Results in Physical Science: a comparative analysis". PNAS 12(2020): 542-554. DOI: 10.2196/48633

44. Novotna E, Eriksdottir N, et al. "Reproducibility of Computational Results in Physical Science: a pilot demonstration". eLife 14(2023): 553-566. DOI: 10.1371/journal.pone.0305949

45. Ostrowski L, Lindqvist O, et al. "Reproducibility of Computational Results in Physical Science: a longitudinal field study". Nature Biotechnology 16(2014): 564-578. DOI: 10.1371/journal.pdig.0000514

46. Petrescu S, Szabo P, et al. "Reproducibility of Computational Results in Physical Science: a feasibility study". Nature Neuroscience 18(2017): 575-581. DOI: 10.17219/acem/208596

47. Quintero B, Balogh Q, et al. "Reproducibility of Computational Results in Physical Science: a case study". Journal of Medical Ethics 20(2020): 586-593. DOI: 10.7759/cureus.104086

48. Ristic C, Castellani R, et al. "Reproducibility of Computational Results in Physical Science: a meta-analysis". Reviews of Modern Physics 22(2023): 597-605. DOI: 10.1038/s41746-023-00852-5

49. Szabo J, Jankowiak S, et al. "Reproducibility of Computational Results in Physical Science: a cross-platform survey". Journal of Plasma Physics 24(2014): 608-617. DOI: 10.1016/j.mayocpiqo.2022.09.004

50. Tarasov Q, Quintero T, et al. "Reproducibility of Computational Results in Physical Science: a methods-comparison study". IEEE Transactions on Quantum Engineering 26(2017): 619-629. DOI: 10.1177/08404704241291226

51. Urbanek Z, Zielinski U, et al. "Reproducibility of Computational Results in Physical Science: a benchmark study". Nature 28(2020): 630-641. DOI: 10.1093/medlaw/fwad013

52. Vasquez A, Alvez V, et al. "Reproducibility of Computational Results in Physical Science: a systematic review". Science 30(2023): 641-653. DOI: 10.2196/89834

53. Wojcik H, Hendriksen W, et al. "Reproducibility of Computational Results in Physical Science: a comparative analysis". Nature Astronomy 32(2014): 652-665. DOI: 10.1007/s10916-024-02104-9

54. Zielinski O, Ostrowski Z, et al. "Reproducibility of Computational Results in Physical Science: a pilot demonstration". Nature Physics 34(2017): 663-677. DOI: 10.3389/fdgth.2026.1896184

55. Andric V, Vasquez A, et al. "Reproducibility of Computational Results in Physical Science: a longitudinal field study". Physical Review Letters 36(2020): 674-680. DOI: 10.3390/diagnostics15162117

Citation: Gerwin Gerstlauer, Gunvor-M Hallikainen, Bohumila Cvrcek. "Reproducibility of Computational Results in Physical Science: A Benchmark of Rerun Success." J Future Sci Technol Innov (2024):3; 019-027. DOI: 10.5281/zenodo.22676473