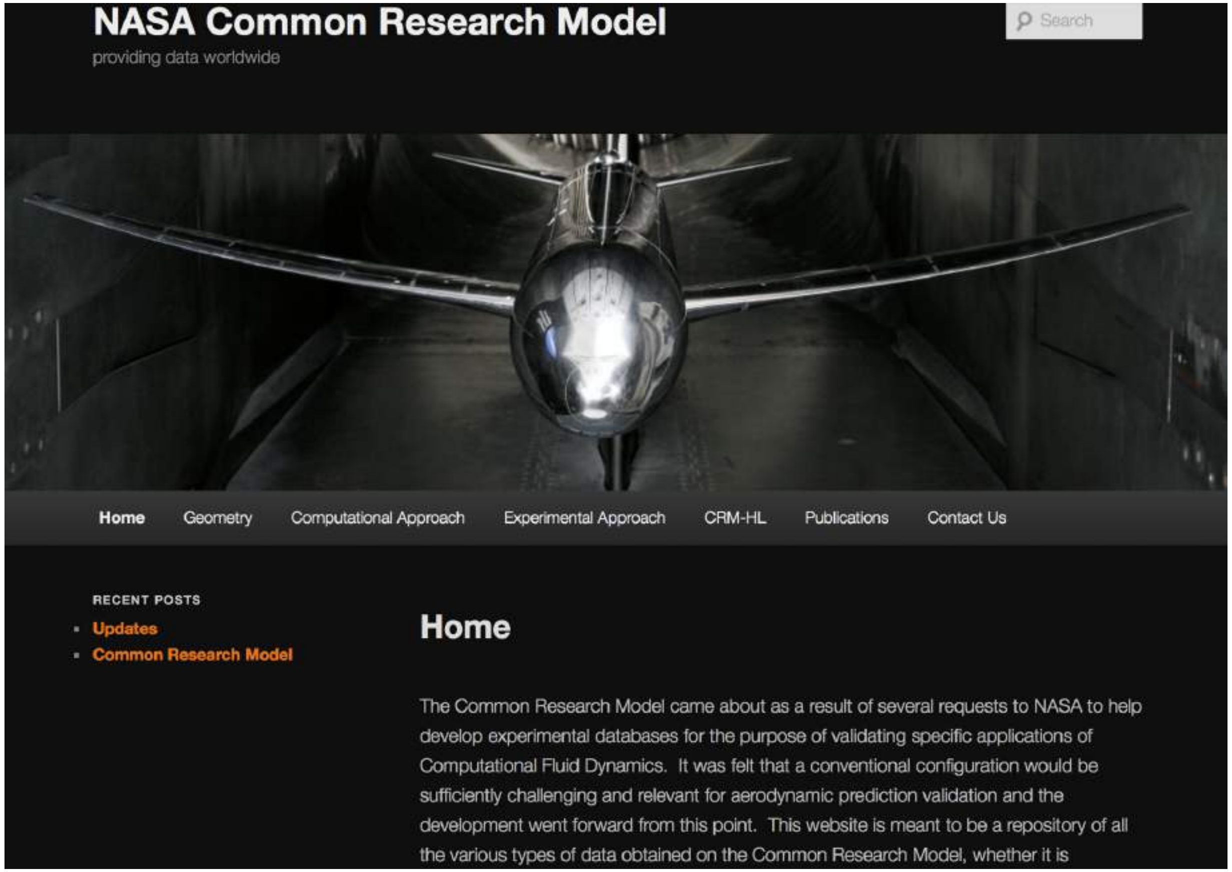
Reproducible Papers
Prof. J. Morlier

Slide Sources:
Victoria Stodden CompareML: Structuring Machine Learning Research in Data Driven Science
Charles Sutton, You and Your Code, Univ. of Edinburgh
Arnaud Legrand Reproducible Research: Where to Begin With?

A (Very) Good Example from … NASA



- Round Robin Test (Test inter Laboratoire en sciences experimentales)
- http://sem.org/dic-challenge/
- https://www.lanl.gov/projects/national-security-education-center/engineering/software/shm-data-sets-and-software.php
- http://www.garteur.org

MDOlab popularity increases through online a

Home > Publications > AIAA Journal > Volume 53, Issue 4 > Aerodynamic Shape Optimization Investigations of the Common Research Model Wing Benchmark

Previous Article | VOLUME 53, ISSUE 4 (APRIL) | Next Article

Add to Favorites | Email | Download to Citation Manager | Track Citations

Full-text | PDF | PDF Plus (3,157 KB) | Supplemental Material | Cited By

Zhoujie Lyu, Gaetan K. W. Kenway, and Joaquim R. R. A. Martins. "Aerodynamic Shape Optimization Investigations of the Common Research Model Wing Benchmark", AIAA Journal, Vol. 53, No. 4 (2015), pp. 968–985.
https://doi.org/10.2514/1.J053318

Check for updates

SUBSCRIBE

Online
Member: $140.00
Non-member: $2,425.00

Buy/Renew

Aerodynamic Shape Optimization Investigations of the Common Research Model Wing Benchmark

Zhoujie Lyu*, Gaetan K. W. Kenway†, and Joaquim R. R. A. Martins‡
Department of Aerospace Engineering, University of Michigan, Ann Arbor, Michigan 48109

*Ph.D. Candidate, Department of Aerospace Engineering. Student Member AIAA.

†Postdoctoral Research Fellow, Department of Aerospace Engineering. Member AIAA.

‡Associate Professor, Department of Aerospace Engineering. Associate Fellow AIAA.

RELATED PUBLICATIONS

Google Scholar

Search for other articles
By author
Zhoujie Lyu

Acknowledgements The authors are grateful for support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1256260 and from the AFOSR MURI on multi-information sources of multi-physics systems under Award Number FA9550-15-1-0038, program manager Jean-Luc Cambier. The authors would like to thank Shamsheer Chauhan for contributing his figures from the MDO course project, as well as Joseph Morlier and Nathalie Bartoli for their support in the ISAE-SUPAERO course.

Structural and Multidisciplinary Optimization
https://doi.org/10.1007/s00158-018-1912-8

EDUCATIONAL ARTICLE

CrossMark

Open-source coupled aerostructural optimization using Python

John P. Jasa[1] · John T. Hwang[2] · Joaquim R. R. A. Martins[1]

Received: 2 August 2017 / Revised: 28 November 2017 / Accepted: 15 January 2018

Abstract
To teach multidisciplinary design optimization (MDO) to students effectively, it is useful to have accessible software that runs quickly, allowing hands-on exploration of coupled systems and optimization methods. Open-source software exists for low-fidelity aerodynamic or structural analysis, but there is no existing software for fast tightly coupled aerostructural analysis and design optimization. To address this need, we present OpenAeroStruct, an open-source low-fidelity aerostructural analysis and optimization tool developed in NASA's OpenMDAO framework. It uses the coupled adjoint method to compute the derivatives required for efficient gradient-based optimization. OpenAeroStruct combines a vortex lattice method and 1-D finite-element analysis to model lifting surfaces, such as aircraft wings and tails, and uses the coupled-adjoint method to compute the aerostructural derivatives. We use the Breguet range equation to compute the fuel burn as a function of structural weight and aerodynamic performance. OpenAeroStruct has proved effective both as an educational tool and as a benchmark for researching new MDO methods. There is much more potential to be exploited as the research community continues to develop and use this tool.

Keywords Aerostructural design optimization · Wing design · Multidisciplinary design optimization · Project-based learning · Python

1 Summary

In this paper, we discuss OpenAeroStruct,[1] an open-source coupled aerostructural analysis and design optimization tool. OpenAeroStruct couples the vortex-lattice method (VLM) and finite-element analysis (FEA) using six degree-of-freedom (DOF) spatial beam elements with axial, bending, and torsional stiffness. It is mostly implemented in Python, but some of the more intensive computations use Fortran.

OpenAeroStruct is developed within the OpenMDAO framework (Heath and Gray 2012), a NASA-developed open-source software framework for multidisciplinary design optimization (MDO). OpenMDAO facilitates derivative computation for gradient-based optimization using the modular analysis and unified derivatives (MAUD) architecture (Hwang and Martins 2018), which unifies the adjoint method with the chain rule and all other methods for computing discrete derivatives (Martins and Hwang 2013). OpenAeroStruct computes derivatives for the aerostructural system using the coupled adjoint method (Martins et al. 2005; Kenway et al. 2014). The aerodynamic forces and structural displacements are transferred between disciplines in a consistent and conservative manner. This process is simplified because the aerodynamic and structural meshes have the same spanwise discretization, so no interpolation is necessary to transfer the loads or displacements. A variety of solvers can be used to converge the coupled aerostructural system, including block Gauss–Seidel, GMRES, or LU decomposition for the lin-

[1] https://github.com/mdolab/openaerostruct

✉ John P. Jasa
johnjasa@umich.edu

John T. Hwang
john.hwang@eperless.com

Joaquim R. R. A. Martins
jrram@umich.edu

Another nice example

Standard papers (no online supplementary materials)

1 Have you ever tried to reproduce some research results ?

2 Have you ever failed ?

what we can do with standard papers:

read the formulas

believe the results

$ check results

$ reproduce the results

$ see the pictures in detail

$ see the graphs in detail

FAQ:
How numerical integral is implemented in this paper?
How are estimated the optimization hyperparameters ?
…
What are the postprocessing (stress field) tricks?
…

For PhDs…

Don't really understand an algorithm unless you can code it.

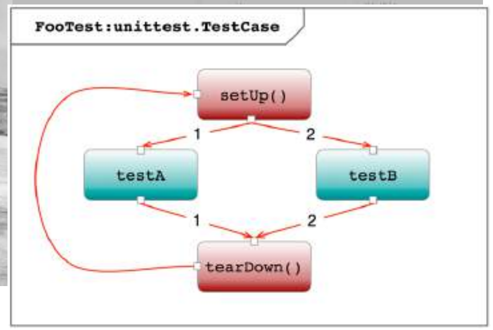
Remember: Software is not your product. → Your product is knowledge:

- ★ New algorithms (often… improvement of existing algorithms)
- ★ New theorems (not often…)
- ★ New models
- ★ New design, optimum, trade-off etc…

Reproducible research:

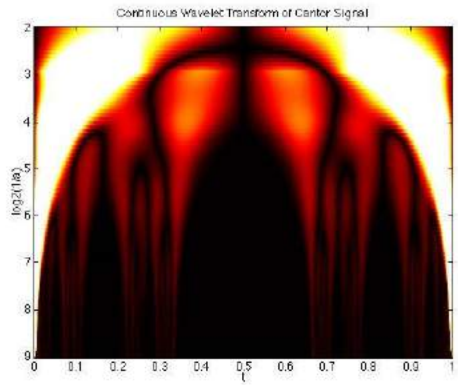
Authors provide all the necessary data and the computer codes to run the analysis again, re-creating the results.

https://hackernoon.com/barba-group-reproducibility-syllabus-e3757ee635cf



Lockheed P-80A airplane (1946). Credit: NASA Commons. — A reminder to test your code.

WAVELAB



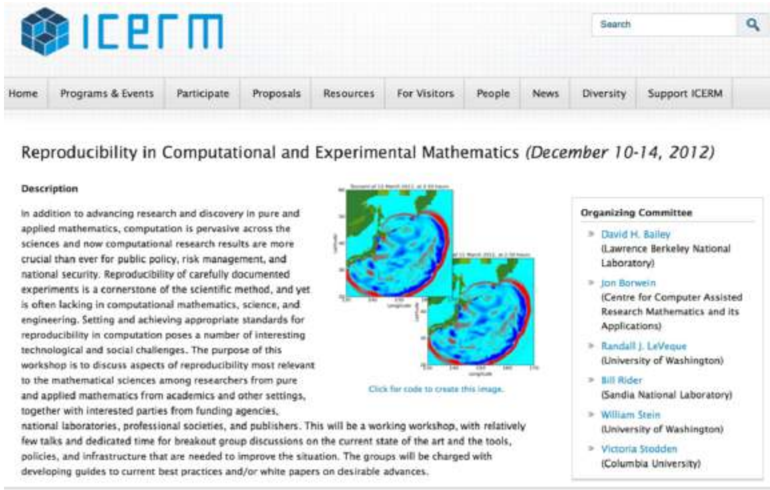
Jump to Donoho et al. (2009). This could be the first group to explicitly associate reproducible research with open code and data:

Reproducible computational research, in which all details of computations — code and data — are made conveniently available to others, is a necessary response to [the credibility] crisis.

My favorite quote from Donoho et al. (2009) is: ***"… if everyone on a research team knows that everything they do is going to someday be published for reproducibility, they'll behave differently from day one."***

… si tous les chercheurs d'une équipe savaient qu'ils devaienit publier avec des contraintes de reproductibilité, ils se comporteraient différemment dès le premier jour

BUT



In fact, David Donoho, 1998
paraphrased a Stanford Prof

"Really Reproducible Research" pioneered by Jon Claerbout:

"The idea is: An article about computational science in a scientific publication is ***not*** the scholarship itself, it is merely ***advertising*** of the scholarship. The actual scholarship is the complete … set of instructions [and data] which generated the figures."

L'idée principale est: Un article sur la simulation dans une publication scientifique n'est pas le projet (l'étudiant venant avec sa bourse) en lui-même, c'est simplement la publicité autour de ce projet. Ce projet, lui est l'ensemble … des instructions [et des données] qui ont généré les figures/tableaux

An interesting example…

Running the same code twice with identical input will produce the same output.

If the computation is done in serial, this assumption is good; **but** with parallel computing, it is not always the case.

Diethelm (2012) ran an experiment using an application of finite-element analysis in computational mechanics. Executing the same simulation (same code, same input data) with varying number of processors gave different results!

The answer

A vector dot-product, computed in parallel over several partial sums. On each execution, individual processors may complete their portion of the sum in different order. In finite precision, addition is not associative and the final sum depends on the order of the partial sum

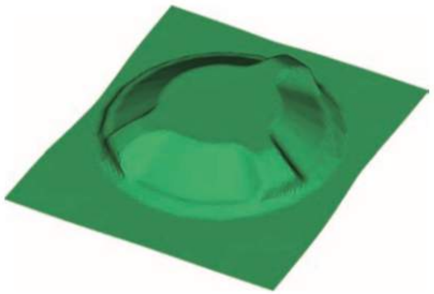


Figure 2. The part to be manufactured. Manufactured from a 0.6-mm thick sheet of steel, the part is about 350-mm long and 400-mm wide. The finite elements used for the simulation have an edge length of 4 mm.

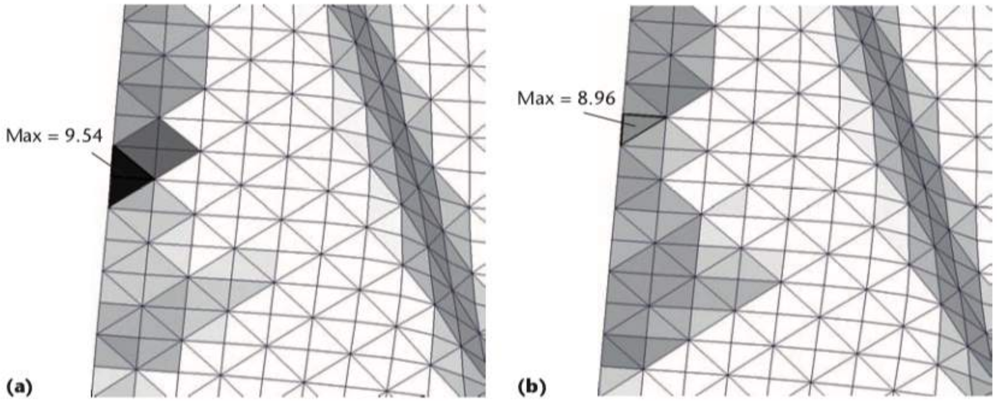


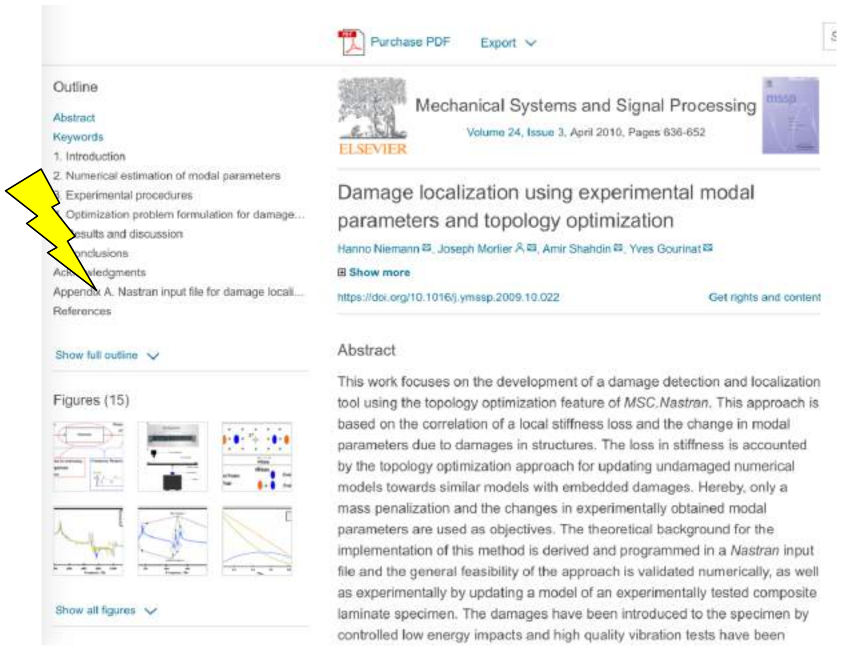
Figure 4. Location of the computed maxima of the sheet thickness change. (a) The simulation with one processor. (b) The second run of the simulation with four processors. The darker the element is colored, the larger the corresponding sheet-thickness change. Elements colored in white have a sheet thickness change of less than 8.5 percent.

My definition

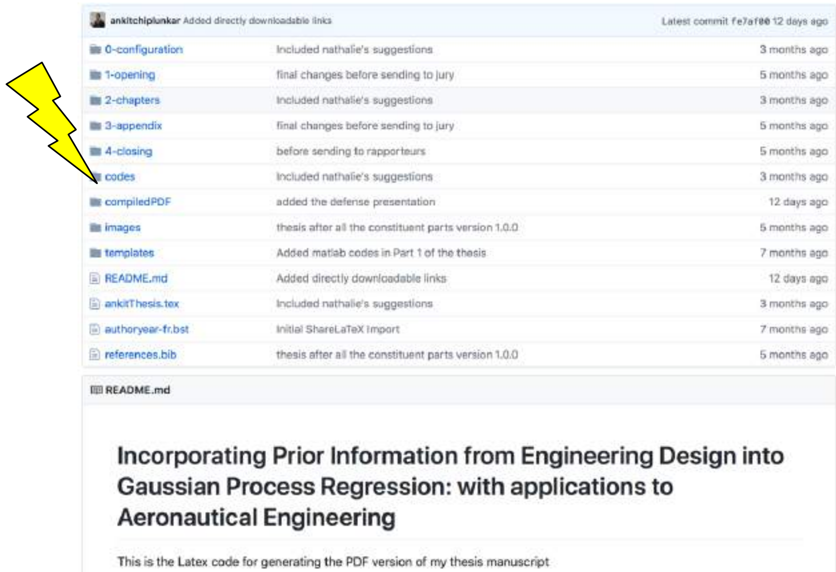
sustainability of the research works
pérennisation des travaux de recherche

"reproducible research" means submitting at minimum:

1. the paper
2. all code & data to reproduce results under open source licenses
3. README files describing code & data



https://github.com/ankitchiplunkar/thesis_isae



Code is not a paper but…

- https://distill.pub
- https://www.software.ac.uk/resources/guides/which-journals-should-i-publish-my-software
- https://www.journals.elsevier.com/softwarex/

PROs and CONs

- From: Survey of the Machine Learning Community, NIPS (Stodden 2010)

Code		Data
91%	Encourage scientific advancement	81%
90%	Encourage sharing in others	79%
86%	Be a good community member	79%
82%	Set a standard for the field	76%
85%	Improve the calibre of research	74%
81%	Get others to work on the problem	79%
85%	Increase in publicity	73%
78%	Opportunity for feedback	71%
71%	Finding collaborators	71%

Survey of the Machine Learning Community, NIPS (Stodden 2010)

Code		Data
77%	Time to document and clean up	54%
52%	Dealing with questions from users	34%
44%	Not receiving attribution	42%
40%	Possibility of patents	-
34%	Legal Barriers (ie. copyright)	41%
-	Time to verify release with admin	38%
30%	Potential loss of future publications	35%
30%	Competitors may get an advantage	33%
20%	Web/disk space limitations	29%

Survey of the Machine Learning Community, NIPS (Stodden 2010)

Response from Within the Sciences

The ***Reproducible Research Standard*** (***RRS***) (Stodden, 2009)
A suite of license recommendations for computational science:

Release media components (text, figures) under CC BY,

https://web.stanford.edu/~vcs/talks/VictoriaStoddenCommuniaJune2009-2.pdf

Benefit for Scientists

- Openness means increased citation.
- Working reproducibly engenders better science.
- Easier for the scientists to build on his or her own work.
- Showcase of skillset for potential collaborators/funders/employers

The pledge:

http://lorenabarba.com/gallery/reproducibility-pi-manifesto/

I will teach my graduate students about reproducibility.

All our research code (and writing) is under version control.

We will always carry out verification and validation (V&V reports are posted to figshare)

For main results in a paper, we will share data, plotting script & figure under CC-BY

We will upload the preprint to arXiv at the time of submission of a paper.

We will release code at the time of submission of a paper.

We will add a "Reproducibility" declaration at the end of each paper.

I will keep an up-to-date web presence.

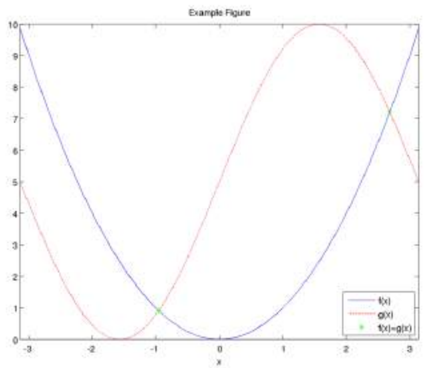
→ Higher confidence in our (students) work can create a competitive advantage

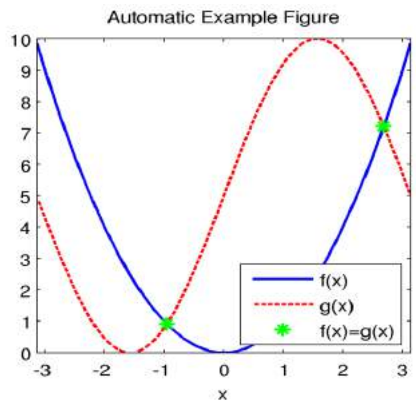
→ Our research will become deeper more visible & reusable (even for us)

Finally…

Reach HQ standards

- **Manuscript**
- **Presentation**





What is the best way to present your results?

En France, depuis 2013

http://www.runmycode.org/home/?/CompanionSite/

https://www.lemonde.fr/sciences/article/2013/07/15/pour-une-recherche-reproductible-publiez-vos-codes-et-donnees_3447825_1650684.html

https://www.fun-mooc.fr/courses/course-v1:inria+41016+session01bis/about

SCIENCES

TRIBUNE

Antoine Blanchard et Elifsu Sabuncu
Fondateurs de Deuxième Labo

Yvan Stroppa
Ingénieur CNRS, cofondateur et "chief computing officer" de RunMyCode

Pour une recherche reproductible, publiez vos codes et données

Selon Antoine Blanchard et Elifsu Sabuncu, fondateurs du Deuxième Labo, et Yvan Stroppa, ingénieur CNRS, la lutte contre la fraude scientifique passe par une meilleure transparence des résultats des chercheurs.

Publié le 15 juillet 2013 à 14h42 - Mis à jour le 10 octobre 2013 à 14h35 | Lecture 5 min.

En avril, un pavé a été jeté dans la mare de la recherche en économie : l'un des articles économiques les plus cités ces dernières années par les partisans de l'austérité, selon lequel le taux de croissance d'un pays devient négatif quand sa dette dépasse 90 % du produit intérieur brut (PIB), serait entaché d'erreurs de calcul et de manipulations de données. Ce sont trois économistes de l'université du Massachusetts qui ont découvert la fraude après avoir

Conclusion

- Changes in funding agency requirements
- Changes in journal/conferences publication requirements (SMO?)
- Cultural changes in our relation to publication (more work)

- Reproducible papers are more cited? No Proof at this time

The article is only the top of the iceberg, we need a way to dive and unveil what's behind every graphics and number...

BUT

- It's Definitely more efficient (not only in the long run and for the community)
- It's simply more satisfying…
- Train our researchers and students to use better tools, better research methodology,

→ https://github.com/alegrand/RR_webinars

REFERENCES

1. M., Karrenbach, N., Claerbout, J. (2000) Making scientific computations reproducible, ***Comp. Sci. Eng***. 2(6):61–67, doi: 10.1109/5992.881708
2. Donoho, D. et al. (2009), Reproducible research in computational harmonic analysis, ***Comp. Sci. Eng.*** 11(1):8–18, doi: 10.1109/MCSE.2009.15
3. Reproducible Research, by the Yale Law School Roundtable on Data and Code Sharing, ***Comp. Sci. Eng.*** 12(5): 8–13 (Sept.-Oct. 2010), doi:10.1109/mcse.2010.113
4. Peng, R. D. (2011), Reproducible research in computational science, ***Science*** 334(6060): 1226–1227, doi: 10.1126/science.1213847
5. Diethelm, Kai (2012) The limits of reproducibility in numerical simulation, ***Comp. Sci. Eng.*** 14(1): 64–72, doi: 10.1109/MCSE.2011.21
6. Setting the default to reproducible (2013), ICERM report of the Workshop on Reproducibility in Computational and Experimental Mathematics (Providence, Dec. 10–14, 2012), Stodden et al. (eds.), https://icerm.brown.edu/tw12-5-rcem/ // report PDF
7. Sandve, G. K. et al. (2013), Ten simple rules for reproducible computational research, ***PLOS Comp. Bio.*** (editorial), Vol. 9(10):1–4, doi: 10.1371/journal.pcbi.1003285
8. Leek, J. and Peng, R (2015), Opinion: Reproducible research can still be wrong: Adopting a prevention approach, ***PNAS*** 112(6):1645–1646, doi: 10.1073/pnas.1421412111
9. M. Liberman, "Replicability vs. reproducibility — or is it the other way around?," Oct. 2015, http://languagelog.ldc.upenn.edu/nll/?p=21956
10. Goodman, S. N., Fanelli, D., & Ioannidis, J. P. (2016). What does research reproducibility mean? ***Science Translational Medicine*** 8(341), 341ps12–341ps12, doi: 10.1126/scitranslmed.aaf5027