

Ensuring reproducibility in computational catalysis

Reproducibility is a cornerstone of science. It is imperative that everyone involved in the generation of scientific knowledge holds themselves to the highest standard to ensure reproducibility.

The way scientific work is conducted has changed drastically in the recent decades, which has had an inevitable impact on the way we communicate the corresponding results. Research articles no longer hinge on one single experiment, but rather a collection of techniques are used to prove not only an ultimate goal but also each of the intermediate conclusions along the path of reaching such goal. As a result, large amounts of data are generated in every single project. This permeates all areas in catalysis, and is particularly relevant for studies involving computational approaches where advancement in technology now allows for massive calculations, which in turn may challenge the standards for reproducibility.

Reproducibility has often been referred to as a cornerstone of science. Every single person involved in the production and dissemination of scientific outcome should be committed to ensure reproducibility. Accessibility of data allows the results to be reproduced by the community. We, as editors, have to make sure that every article we publish complies with the highest degree of reproducibility, meaning that all relevant information garnered during the course of the study should be documented and made available — ideally within the publication or at an external accessible platform.

At the forefront of reproducibility lies computational catalysis. Not only because it has been increasingly growing, but also due to its intrinsic readily accessible nature. This was very nicely addressed by François-Xavier Coudert in the Editorial of *Chemistry of Materials* of April last year¹. All computational data related to an article, from its generation (inputs) to its results (outputs), contain crucial information that is surely valuable to everyone working on the subject. At *Nature Catalysis* we advocate for reproducibility. We require authors to always include a data availability statement and strongly encourage the use of repositories, such as those highlighted by our sister journal *Scientific Data* (<https://www.nature.com/sdata/policies/repositories>).

On top of data availability, our commitment to reproducibility also

embraces computer code. Along with the data availability statement, those articles with specifically developed code must also include a code availability statement. Together with all *Nature Research* journals, we have recently adopted a new policy regarding custom computer code to ensure transparency and replicability. On a case-by-case basis, we might ask reviewers to check the code if we consider it central to the main claims of the work, as introduced in a recent editorial in *Nature*². In addition, we find that transparency and accessibility of custom computer code, especially when deemed pivotal to the conclusions of the study, is of utmost importance, and thus it should be accessible via the Supplementary Information or an external repository.

Exemplifying our advocacy for reproducibility, in this issue of *Nature Catalysis* you will find three research articles, in addition to a Comment, that — completely, or for their most part — rely on computational approaches. All three research articles, which are summarized hereafter, adhere to our data and code availability guidelines.

At *Nature Catalysis* we advocate for reproducibility. We require authors to always include a data availability statement and strongly encourage the use of repositories.

William Schneider, Jason Hicks, David Go and co-workers use a microkinetic model to demonstrate that ammonia synthesis at near-ambient conditions can reach similar rates as those of the Haber–Bosch process when coupled to a plasma source. The code of the plasma-assisted microkinetic model has been uploaded in a repository whereas the python scripts to reproduce all of the figures and experimental data are included as Supplementary Data files. Furthermore, the authors went one step further and all density functional theory energies used in the kinetic model are extracted from the CatApp database³,

demonstrating that the benefits of data availability extend beyond reproducibility.

Andrew Peterson and colleagues show that scaling relations can be circumvented by applying strain to the catalytic surface and that, in conjunction with the intrinsic strain induced by the adsorbate, a specific response can be engineered. The cartesian coordinates of all intermediate species can be found in the Supplementary Information as a Supplementary Data file.

Peijun Hu, Hai-Feng Wang and co-workers found — by means of molecular dynamics simulations and microkinetic analyses — that the rate of photocatalytic oxygen evolution on titanium dioxide is limited by the concentration of photoholes reaching the surface. The code of the microkinetic model can be found at an external repository and the cartesian coordinates of all intermediate species are included in the Supplementary Information as a Supplementary Data file.

As expressed by John Kitchin in his Comment article on machine learning, the format of the datasets and custom code in the Supplementary Information is also an important subject that we should pay attention to. These files should facilitate the extraction and subsequent manipulation of the data^{4,5}.

The data-driven nature of computational catalysis puts it directly under the spotlight of reproducibility. Needless to say, reproducibility is the main pillar of good scientific practice that extends beyond computational catalysis. At *Nature Catalysis*, we will continue to work towards ensuring accessibility and reproducibility in all areas. □

Published online: 16 April 2018
<https://doi.org/10.1038/s41929-018-0068-7>

References

1. Coudert, F.-X. *Chem. Mater.* **29**, 2615–2617 (2017).
2. *Nature* **555**, 142 (2018).
3. Hummelshøj, J. S., Abild-Pedersen, F., Studt, F., Bligaard, T. & Nørskov, J. K. *Angew. Chem. Int. Ed.* **124**, 278–280 (2011).
4. Rudsteyn, B., Acharya, A. & Batista, V. S. J. *Phys. Chem. A* **121**, 9680–9681 (2017).
5. Kitchin, J. R. *ACS Catal.* **5**, 3894–3899 (2015).