https://doi.org/10.5194/gc-2019-22
Preprint. Discussion started: 29 October 2019

Geoscience Communication
Discussions
Open Access

Weather and Climate Science in the Digital Era

Martine G. de Vos[1,2], Wilco Hazeleger[1,3], Driss Bari[4], Jorg Behrens[5], Sofiane Bendoukha[5], Irene Garcia-Marti[6], Ronald van Haren[1], Sue Ellen Haupt[7], Rolf Hut[8], Fredrik Jansson[9], Andreas Mueller[10], Peter Neilley[11], Gijs van den Oord[1], Inti Pelupessy[1], Paolo Ruti[12], Martin G. Schultz[13], and Jeremy Walton[14]

[1]Netherlands eScience center, Amsterdam, the Netherlands
[2]Information and Technology Services, Utrecht University, Utrecht, the Netherlands
[3]Geosciences, Utrecht University, Utrecht, the Netherlands
[4]CNRMSI/SMN, Direction de la Meteorologie Nationale Casablanca, Morocco
[5]German Climate Computing Centre (DKRZ), Hamburg, Germany
[6]Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands
[7]Research Applications Laboratory, National Center for Atmopsheric Research, Boulder, USA
[8]Water Resources Management, Delft University of Technology, Delft, the Netherlands
[9]Centrum Wiskunde & Informatica, Amsterdam, the Netherlands
[10]Numerical methods, European Centre for Medium-Range Weather Forecasts, Reading, UK
[11]The Weather Company/IBM, Boston MA, USA
[12]World Weather Research Division, World Meteorological Organization, Geneva, Switzerland
[13]Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany
[14]Hadley Centre for Climate Science, Met Office, Exeter, UK

Correspondence: Martine G. de Vos (m.g.devos@uu.nl)

Abstract. The need for open science has been recognized by the communities of meteorology and climate science. However, while these domains are mature in terms of applying digital technologies, these are lagging behind where the implementation of open science methodologies is concerned. In a session on “Weather and Climate Science in the Digital Era” at the 14th IEEE International eScience conference domain specialists and data and computer scientists discussed the road towards open weather and climate science.

The studies presented in the conference session showed the added value of shared data, software and platforms through, for instance, combining data sets from disparate sources, increased accuracy and skill of simulations and forecasts at local scales, and improved consistency of data products. We observed that sharing data and code is important, but not sufficient to achieve open weather and climate science and that here are important issues to address.

At the level of technology, the implementation of the FAIR principles to many datasets used in weather and climate science remains a challenge due to their origin, scalability, or legal barriers. Furthermore, the complexity of current software platforms limits collaboration between researchers and optimal use of open science tools and methods.

The main challenges we observed, however, were non-technical and impact the system of science as a whole. There is a need for new roles and responsibilities at the interface of science and digital technology, e.g., data stewards and research software engineers. This requires the personnel portfolio of academic institutions to be more diverse, and in addition, a broader consideration of the impact of academic work, beyond publishing and teaching. Besides, new policies regarding open weather

https://doi.org/10.5194/gc-2019-22
Preprint. Discussion started: 29 October 2019

and climate science should be developed in an inclusive way to engage all stakeholders, including non-academic parties such as meteorological institutions.

We acknowledge that open weather and climate science requires effort to change, but the benefits are large. As can already be observed from the studies presented in the conference it leads to much faster progress in understanding the world.

1 INTRODUCTION

Meteorology and climate sciences are data and compute intensive areas of research by tradition. Being primarily a physical science, empirical data collection has always been important and meteorology was one of the first fields that standardized data collection from the advent of systematic instrumental observations in the mid 1800s (e.g. Maury, 1853; Quetelet, 1874). Also, since the early development of computers, meteorological applications have always been considered. From the first operational weather predictions in the 1950s onwards (Charney et al., 1950) numerical weather prediction has advanced, driven by increasing computing capability and the growing supply of observational data to generate initial conditions and assimilate them into the model state. In many ways, climate research has benefitted from the same developments (see e.g. Lynch, 2008, for an overview). The integration, i.e., assimilation, of observational data into numerical weather prediction (NWP) models has been a turning point for developing high-resolution gridded information of the atmosphere and ocean state (e.g. Kalnay et al., 1996; Dee et al., 2011). The use of this methodology for reanalysis, i.e., generating a comprehensive and physically consistent record of how the weather is changing over time, since ensured a baseline for climate research and triggered the development of downstream climate services. Meteorologists have been using machine learning to post-process model output, blend multiple models, and optimize the weighting of those multiple models for over 20 years (Haupt et al., 2018). Neural nets were also shown to vastly speed the calculation of both incoming shortwave and outgoing longwave radiation in climate models in the 1990s (Krasnopolsky, 2013). Present and future strategies feature an Earth System approach for assimilating environmental data into a more comprehensive coupled system including the atmosphere, ocean, biosphere and sea-ice (Penny and Hamill, 2017).

Recent developments in digital technologies and methods strongly affect meteorology and climate research. The increase of computing power, currently approaching exascale, provides unprecedented opportunities with regard to resolving more scales numerically or coupling more components. At the same time, it poses large software development and data management challenges, as the increasing numerical model resolution impacts the code complexity, the performance profile and the volumes of data that are handled. A parallel development concerns the open availability of data - both standard meteorological data as well as data from many other sources, including citizen science projects and low-cost sensors. Modern data management tools allow for effectively using these data sources. A third development is the increased use of using machine learning methods, in particular deep learning. A plethora of machine learning methods have been and are being applied to problems of weather and climate prediction, from emulating unresolved processes in numerical models to calibrating forecasts produced with numerical models and doing forecasts based on data and machine learning methods only.

Digital technologies enable new research methods, accelerate the growth of knowledge, and spur the creation of new means of communicating that knowledge among researchers and within the wider community. As such, these technologies have reshaped the scientific enterprise and are strongly connected to open science (OECD, 2015; Bourne et al., 2012). Open science methodologies, such as open access publications, FAIR data principles and open source software development, stimulate the reuse of data and software resources and lead to more reproducible research (Wilkinson et al, 2016; Munafò et al., 2017). The need for open research practices has been recognized by the communities of meteorology and climate science. However, while these domains are mature in terms of applying digital technologies, these are lagging behind where the implementation of open science methodologies is concerned.

All these developments spurred organizing a session on “Weather and Climate Science in the Digital Era” at the 14th IEEE International eScience conference. In this session, specialists on the domain of weather and climate science, data and computer scientists came together to discuss the road towards open weather and climate science. This paper describes the main findings of this session and aims to provide input to detail the strategies laid out by institutes and research organizations in the field of weather and climate science.

2 OPEN SCIENCE

In this section we explore the relation between meteorology and climate science and open science developments. Open science refers to open research practices, and includes but is not limited to public access to the academic literature, sharing of data and code (Mckiernan et al., 2016). However, the exact interpretation of the concept of open science is different for different schools of thought (Fecher and Friesike, 2014). In general, open science concerns many different stakeholders: besides scholars, these include institutes, research funders, librarians and archivists, publishers and decision makers (Bourne et al., 2012; OECD, 2015; Fecher and Friesike, 2014).

It has been shown that open research practices bring significant benefits to researchers as these are associated with increases in citations, media attention, potential collaborators, job opportunities, and funding opportunities (Mckiernan et al., 2016). Recently, countries throughout the world have made efforts to adapt legal frameworks and implement policy initiatives to encourage greater openness in scientific research (OECD, 2015; National Science Foundation, 2018). Funders and research institutes have announced policies encouraging, mandating, or specifically financing open research practices (Mckiernan et al., 2016; Wilkinson et al, 2016).

The need for open research practices has been recognized by the communities of meteorology and climate science and even entered into the political arena. For instance, in its report on the “Climatic Research Unit email controversy” in 2009 the Science and Technology Committee of the UK House of Commons stated that climate science is a matter of great importance and the quality of the science should be irreproachable. The committee called for the climate science community to become more transparent by publishing raw data and detailed methodologies (House of Commons, 2010).

Nowadays, there are several examples of open access, open data and open source software in meteorology and climate science. The United States already has a long history of making meteorological observations, model source codes and model output an open public commodity, available to all. A good example of open source software in the meteorological community in Europe is the OpenIFS initiative (Carver, 2019). The European Center for Medium-range Weather Forecasts (ECMWF) provides researchers with a free, and easy-to-use version of the Integrated Forecasting System (IFS), i.e., one of the main global NWP systems. It allows IFS to be used by a much wider community and the academic community contributes to improving the forecast model with new developments. Without being exhaustive, other examples of shared numerical weather and climate model codes are the WRF regional model and the CESM climate model (Skamarock et al., 2019; Hurrell, J.W. et al., 2013).

In addition, co-ordinated coupled model intercomparison projects (CMIP) (Taylor et al., 2012; Eyring et al., 2016) are excellent examples of the climate modeling community working together. The construction of multi-model comparisons and statistics forces research groups to accept common input forcings, provide detailed documentation of the numerical schemes in their model and produce open, standardized output data. As a result, the CMIP initiative enables earth science groups around the world to assess the impact of a changing climate in their domain.

Although open science is growing in popularity and necessity, widespread adoption of these practices has not yet been achieved and this is true for meteorology and climate science as well. Recent studies show that transparency and reproducibility are still a matter of concern to the scientific community as a whole. It requires all the stakeholders of science to work together to create a more open and robust system (Baker, 2016; Munafò et al., 2017; Gil et al., 2016).

3 TOWARDS OPEN WEATHER AND CLIMATE SCIENCE

The IEEE eScience Conference session on weather and climate science included presentations of state-of-the art research at the interface of weather and climate science and digital technologies. Contributions were selected after a peer review on their scientific merit and innovative nature and published in the conference proceedings (Bari; Behrens et al.; Bendoukha; Brangbour et al.; Garcia-Marti et al.; Haupt et al.; Hut et al.; Jansson et al.; Pelupessy et al.; Ramamurthy; Schultz et al.; Stringer et al.; van Haren et al.; van den Oord et al., 2018). In a synthesis session we observed several developments towards open research practices and discussed challenges and opportunities.This section presents the common findings and highlights of the conference session.

3.1 OPEN DATA

Many studies reported in the proceedings of the conference include open data from different sources in their analyses which clearly enrich their research. Enhanced research was shown with the use of open satellite data, geolocated data via Open Street Map and openly available in-situ meteorological observations (Haupt et al., 2018; Garcia-Marti et al., 2018; Bari, 2018; Schultz et al., 2018, and references therein). Also, citizen data like social media posts increasingly leads to new findings (Brangbour et al., 2018) and observations from amateur weather stations can lead to new perspectives on local weather conditions beyond data from traditional meteorological stations (van Haren et al., 2018). All of these studies show that advances in scientific understanding are made with open data and often with combinations of data which are not common in meteorological or

https://doi.org/10.5194/gc-2019-22
Preprint. Discussion started: 29 October 2019

climate research. Besides, a number of studies use standards for file formats and metadata, like NetCDF and CF (van den Oord et al., 2018). The latter formats are increasingly used in climate studies. Such common formats and standard protocols for inter-process communication, like MPI and REST in numerical codes (Behrens et al., 2018; Pelupessy et al., 2018; Schultz et al., 2018), facilitate exchange and use of data.

We recognized that in current weather and climate science the focus is mostly done on making data and software findable and accessible, often via webportals. Although these are necessary first steps towards open data and open science, we acknowledge that these steps are not sufficient. Data and software that is findable and accessible may still be hard to obtain in practice or may be disseminated in a way that it is still difficult to interpret and use. Wilkinson and colleagues (2016) defined guidelines, referred to as the FAIR Principles, to ensure the transparency, reproducibility, and reusability of scientific data. These guidelines state that data - and also the algorithms, tools, and workflows that led to these data- should be Findable, Accessible, Interoperable and Reusable (FAIR). The guidelines put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals.

Regarding open and interoperable weather and climate model data, we consider performance scalability as the foremost technological challenge. Producing FAIR model data via traditional post-processing pipelines is quickly becoming unfeasible for high-resolution climate model data due to the sheer volume and complexity of the model output as noted above. The same is true for most satellite data products. For simulation models, this trend is a consequence of the advance of processor speed compared to storage bandwidth, and can only be countered with (i) increased parallelism in the climate data processing toolchain, or (ii) removing the need for post-processing by incorporating as many steps as possible within the model itself. While many tools exist to support open sharing of geospatial data including comprehensive metadata descriptions, these tools generally do not scale and cannot be employed with the massive amount of weather and climate model data.

In addition to these technological challenges, we observe that some important challenges for open data arise from the political or legal context, and as such require additional efforts beyond the scientific domain. Weather Institutes and commercial entities can see their data as a business advantage and can be reluctant to make it open. Various resolutions by the World Meteorological Organisation (e.g. Resolution 40, 25 and 60) promote open access and exchange of data in order to better manage the risks from weather and climate-related hazards, but leave room for additional conditions. These resolutions have no legal status and national legislation may lead to restricted access to data and charges (Sylla, 2018). Also, policies to promote open data are less mature than those to promote open access to scientific publications (OECD, 2015).

Furthermore, data need to be hosted, and maintained and their quality should be ensured. For large operational data services, such as the European Copernicus program, this is well taken care of, but this is less the case for research data of individual scientists despite the rising attention to data management. Currently, there is no credit or clear policy for data providers to host data and manage good quality, i.e., implement the FAIR principles. The rise of data journals remedies this partially, as it allows for crediting data producers and peer review of (meta)data. Some funding agencies, e.g., the national research funding in the Netherlands NWO, are now requiring that in all projects that they fund software becomes open source and the data are archived and findable unless there are strong reasons not to do so (e.g. privacy). Also, research funded by the European Commission should adhere to FAIR principles and data management plans need to be in place.

3.2 OPEN SOFTWARE

The conference session provided excellent examples where considerable attention is paid to documentation and reuse of tools and methods (Stringer et al., 2018; Behrens et al., 2018; van Haren et al., 2018; Schultz et al., 2018; van den Oord et al., 2018). Moreover, many of these studies present an approach for which open data and software is a prerequisite (Pelupessy et al., 2018; Jansson et al., 2018; Ramamurthy, 2018; Hut et al., 2018; Bendoukha, 2018).

We strongly support open publication of code, even if this code under development, and especially when this code is used in a paper to support research findings. Open code can be inspected and reused by peers, which improves the reproducibility and quality of the corresponding research. This is crucial to science and to climate research in particular, since local and global policies depend on the scientific results. Open publication, however, requires the software code to be documented and tested, which is a time consuming effort. In the current situation this is not standard practice, partially because there is no incentive to do so. There is a need for open science practices where incentives are developed to share scientific information beyond the final result in a scientific paper. We are convinced that these practices will strongly improve scientific practice.

In several of the studies that were presented in the conference machine learning technologies are used for data analysis and prediction (Haupt et al., 2018; Garcia-Marti et al., 2018; Bari, 2018; Schultz et al., 2018). The studies show that use of machine learning methods has added value because models are built with data beyond standard meteorological data. For example, local conditions related to the natural and built environment that cannot be captured easily in simulation models can be taken into account through trained models.

It was observed that in general the use of machine learning approaches in weather and climate science is increasing. These approaches are powerful, for instance, in emulating processes that are not resolved in simulation models, because of computational costs, in calibrating or post-processing simulation results and in building models to describe or forecast meteorological and climatological events. The caveats, on the other hand, are that trained models are not transparent as models based on laws of physics and their results can be hard to interpret. Following the open science principle, machine learning approaches should be understandable and reusable by other researchers. Emerging field like Explainable AI and knowledge based machine learning may provide approaches that help humans experts to understand how machine learning results are produced (Adadi and Berrada, 2018). Data-driven machine learning approaches should be combined with knowledge on physical processes (Dueben and Bauer, 2018; Reichstein et al., 2019) to gain further understanding of Earth system science problems. Moreover, machine learning methods should be accompanied by proper validation and verification.

The use of software as presented above, motivated by open science principles, requires a suitable digital infrastructure. Hardware and software platforms provide tools and services for scientists to perform and disseminate their research and as such facilitates collaboration and reproducibility. Although several of such platforms were presented in the conference session (Ramamurthy, 2018; Hut et al., 2018; Bendoukha, 2018), it was observed that a platform for cross-disciplinary collaboration is still lacking. This may sound contradictory as there are many existing platforms that provide tools and services for weather and climate scientists that aim to facilitate collaboration. The variety of these platforms and the corresponding interfaces, however, pose technical difficulties to researchers and dilute the possibilities of collaboration, especially between different disciplines.

4 DISCUSSION AND CONCLUSION

Based on a session at the 14th IEEE International eScience Conference and the contributions published in the proceedings of the conference, we report on the synthesis of discussions and a further analysis of open science principles in meteorology and climate research. The individual peer reviewed contributions show the value of sharing data, open data, using and developing open source software and using and developing open (software) platforms. Scientific advances are shown, for instance through combining data sets, including non-standard meteorological data such as that of the environment and citizen science sources. The increase in accuracy and skill of forecasts at local scales are shown, improved consistency of data products and improved efficiency and skill of simulations, often crossing different disciplines. The renewed attention of machine learning and increased computational capabilities have facilitated the use of disparate sources of data.

Sharing of data and code offers many opportunities for scientific progress and leads to better reproducible science and it vastly enhances the user base. However, in our conference session we observed that open publication of data and code is not enough to achieve open weather and climate science and that here are important issues to address.

The findability and accessibility of data increasingly gets attention in weather and climate research, and common file and metadata formats increase interoperability. However, for many data sets the implementation of the FAIR principles remains a challenge due to their origin, e.g., citizen data, scalability, e.g., high-resolution climate model data, or legal barriers, e.g., weather forecasts. We also acknowledge that data quality is extremely hard to judge and depends on the actual purpose of the data. This requires a continuous discussion on what aspects of open data can be implemented generically and what aspects are specific.

Technologically, the promise of using modern digital technologies is not always met due to the complexity of software platforms. While this paper hardly addresses hardware, this is true for hardware and software-hardware interaction as well. A further development of platforms should facilitate the ease-of-use and provenance. This also calls for more attention of research software engineering where collaboration and interaction between software engineers and domain researchers can lead to optimal use of open science tools and methods.

As mentioned before, open science science concerns many different stakeholders besides scholars. It is important to acknowledge and define roles, responsibilities and mandates concerning data science, data management, data stewardship and research software engineering. This requires institutional change as the personnel portfolio of academic institutions needs to be more diverse, and in addition, a broader consideration of the impact of academic work, beyond scientific publications and teaching (Akhmerov et al., 2019). In order to remove legal boundaries on sharing data, it is important to also engage non-academic parties such as operational and commercial meteorological institutions in open science. New policies regarding open science should be developed in an inclusive way to engage all stakeholders.

Open science strategies and policies are a means to support better quality science, increased collaboration, and engagement between research and society that can lead to higher social and economic impacts of public research (OECD, 2015). Open science has implications for stakeholders, the institutions and the system of science as a whole. It requires effort to change, but the benefits are large. Sharing data, code, and knowledge openly vastly enhances the user base, which means manifold growth of opportunities for new discoveries. As can already be observed from the studies presented in the 14th IEEE International eScience Conference this leads to much faster progress in understanding the world.

Author contributions. MGdV and WH organized the conference session and were lead writers of the manuscript. All authors contributed to the presentations and discussion in the conference session and to the writing of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would like to acknowledge both the Netherlands eScience Center and the program committee of the Weather & Climate session for their efforts in organizing this conference session. The session created the unique opportunity for specialists on the domain of weather and climate science, data and computer scientist to exchange ideas and knowledge.

References

Adadi, A. and Berrada, M.: Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI), IEEE Access, 6, 52 138–52 160, https://doi.org/10.1109/ACCESS.2018.2870052, 2018.

Akhmerov, A., Cruz, M., Drost, N., Hof, C., Knapen, T., Kuzak, M., Martinez-Ortiz, C., Turkyilmaz-van der Velden, Y., and Van Werkhoven, B.: Raising the Profile of Research Software : Recommendations for Funding Agencies and Research Institutions, Tech. rep., Netherlands eScience Center, Amsterdam, Netherlands, https://doi.org/10.5281/zenodo.3378572, 2019.

Baker, M.: Is there a reproducibility crisis? A Nature survey lifts the lid on how researchers view the 'crisis rocking science and what they think will help., Nature, 533, 452+, 2016.

Bari, D.: Visibility Prediction based on kilometric NWP Model Outputs using Machine-learning Regression, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00048, 2018.

Behrens, J., Biercamp, J., Bockelmann, H., and Neumann, P.: Increasing parallelism in climate models via additional component concurrency, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00044, 2018.

Bendoukha, S.: Towards a new Big Data Analytics Platform for Climate Community, in: IEEE 14th International Conference on e-Science, 2018.

Bourne, P. E., Clark, T., de Ward, D. R., Herman, I., Hovy, E., and Shotton, D.: Force 11 White Paper: Improving the future of research communication and e-scholarship, Tech. rep., Force11, https://doi.org/10.4230/DagMan.1.1.41, https://www.force11.org/about/manifesto, 2012.

Brangbour, E., Bruneau, P., and Marchand-Maillet, S.: Extracting Flood Maps from Social Media for Assimilation, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00045, 2018.

Carver: The ECMWF OpenIFS numerical weather prediction model release cycle 40r1: description and use cases", in preparation to be submitted to GMDD, 2019.

Charney, J. G., FjÖrtoft, R., and Neumann, J. V.: Numerical Integration of the Barotropic Vorticity Equation, Tellus, 2, 237–254, https://doi.org/10.3402/tellusa.v2i4.8607, 1950.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Dueben, P. D. and Bauer, P.: Challenges and design choices for global weather and climate models based on machine learning, Geoscientific Model Development, 11, 3999–4009, https://doi.org/10.5194/gmd-11-3999-2018, 2018.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled MOdel Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization, Geoscientific Model Development, 9, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fecher, B. and Friesike, S.: Open Science: One Term, Five Schools of Thought, in: Opening Science, edited by Bartling, S. and Friesike, S., 1, pp. 1–7, The Author(s), https://doi.org/10.1007/978-3-319-00026-8_2, 2014.

Garcia-Marti, I., Noteboom, J. W., and Diks, P.: Detecting probability of ice formation on overhead lines of the Dutch railway network, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00050, 2018.

Gil, Y., David, C. H., Demir, I., Essawy, B. T., Fulweiler, R. W., Goodall, J. L., Karlstrom, L., Lee, H., Mills, H. J., Oh, J. H., Pierce, S. A., Pope, A., Tzeng, M. W., Villamizar, S. R., and Yu, X.: Toward the Geoscience Paper of the Future: Best practices for documenting and sharing research from data to software to provenance, Earth and Space Science, 3, 388–415, https://doi.org/10.1002/2015EA000136, 2016.

Haupt, S. E., Cowie, J., Linden, S., Mccandless, T., Kosovic, B., and Alessandrini, S.: Machine Learning for Applied Weather Prediction, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00047, 2018.

House of Commons: The disclosure of climate data from the Climatic Research Unit at the University of East Anglia, Science and Technology Committe, https://doi.org/citeulike-article-id:11615640, http://www.publications.parliament.uk/pa/cm200910/cmselect/cmsctech/387/387i.pdf, 2010.

Hurrell, J.W., Holland, M., Gent, P., Ghan, S., Kay, J., Kushner, P., Lamarque, J., Large, W., Lawrence, D., Lindsay, K., and Lipscomb, W.: The Community Earth System Model: A framework for collaborative research, Bulletin of the American meteorological Society, 94, https://doi.org/10.1175/BAMS-D-12-00121.1., 2013.

Hut, R., Drost, N., van Hage, W., and van de Giesen, N.: eWaterCycle II, in: IEEE 14th International Conference on e-Science, 2018.

Jansson, F., van den Oord, G., Siebesma, P., and Crommelin, D.: Resolving clouds in a global atmosphere model - a multiscale approach with nested models, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00043, 2018.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., and Zhu, Y.: The NCEP/NCAR 40-year reanalysis project, Bulletin of the American meteorological Society, 77, 437–472, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Krasnopolsky, V. M.: The application of neural networks in the earth system sciences, in: Neural Networks Emulations for Complex Multidimensional Mappings, Springer, New York, New York, USA, 2013.

Lynch, P.: The origins of computer weather prediction and climate modeling, Journal of Computational Physics, 227, 3431–3444, https://doi.org/10.1016/j.jcp.2007.02.034, 2008.

Maury, M. F.: First International Maritime Conference Held for Devising an Uniform System of Meteorological Observations at Sea, Brussels, 1853.

Mckiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Mcdougall, D., Nosek, B. A., Ram, K., and Soderberg, C. K.: How open science helps researchers succeed, Elife, 5, 1–26, https://doi.org/10.7554/eLife.16800, 2016.

Munafò, M. R., Nosek, B. A., Dorothy V. M. Bishop, K. S. B., Christopher D. Chambers, N. P. d. S., 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.

National Science Foundation: PROPOSAL & AWARD POLICIES AND PROCEDURES GUIDE (PAPPG), Tech. Rep. OMB Control Number 3145-0058, 2018.

OECD: Making open science a reality, OECD Science, Technology and Industry Policy Papers, p. 112, https://doi.org/http://dx.doi.org/10.1787/5jrs2f963zs1-en, 2015.

Pelupessy, I., Werkhoven, B. V., van den Oord, G., Zwart, S. P., van Elteren, A., and Dijkstra, H.: Development of the OMUSE / AMUSE modelling system, in: IEEE 14th International Conference on e-Science, 2018.

Penny, S. G. and Hamill, T. M.: Coupled Data Assimilation for Integrated Earth System Analysis and Prediction, Bulletin of the American Meteorological Society, 98, ES169–ES172, https://doi.org/10.1175/BAMS-D-17-0036.1, http://journals.ametsoc.org/doi/10.1175/BAMS-D-17-0036.1, 2017.

Quetelet, A.: Notice sur Le Capitaine M. F. Maury, in: Associé de l'Académie Royale de Belgique, published by the Academy, Brussels, 1874.

Ramamurthy, M.: Toward a Cloud Ecosystem for Modeling as a Service, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00046, 2018.

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, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Schultz, M. G., Apweiler, S., Vogelsang, J., Kleinert, F., and Mallmann, D.: A web service architecture for objective station classification purposes, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00051, 2018.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Version 4. NCAR Tech. Note NCAR/TN-556+STR, Tech. rep., NCAR, https://doi.org/10.5065/1dfh-6p97, 2019.

Stringer, M., Jones, C., Hill, R., Dalvi, M., Johnson, C., and Walton, J.: A Hybrid-Resolution Earth System Model, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00042, 2018.

Sylla, M. B.: Review of meteorological / climate data sharing policy (WMO Resolution 40) to promote their use to support Climate Information Services uptake in the African continent, in: Expert Group Meeting on data sharing policy in Africa, July, pp. 10–11, Dakar, Senegal, 2018.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

van den Oord, G., Yepes, X., and Acosta, M.: Post-processing strategies for the ECMWF model, in: IEEE 14th International Conference on e-Science, 2018.

van Haren, R., Koopmans, S., Steeneveld, G.-J., Theeuwes, N., Uijlenhoet, R., and Holtslag, A. A. M.: Weather reanalysis on an urban scale using WRF, in: IEEE 14th International Conference on e-Science, https://doi.org/10.1109/eScience.2018.00049, http://www2.mmm.ucar.edu/wrf/users/docs/arw{_}v3.pdf, 2018.

Wilkinson et al, M. D.: Comment: The FAIR Guiding Principles for scientific data management and stewardship, Scientific data, 3, https://doi.org/DOI: 10.1038/sdata.2016.18, 2016.