PERSPECTIVE

The scientific challenge of understanding and estimating climate change

Tim Palmer[a,1] and Bjorn Stevens[b,1]

Edited by Jagadish Shukla, George Mason University, Fairfax, VA, and accepted by Editorial Board Member Robert E. Dickinson October 21, 2019 (received for review May 21, 2019)

Given the slow unfolding of what may become catastrophic changes to Earth's climate, many are understandably distraught by failures of public policy to rise to the magnitude of the challenge. Few in the science community would think to question the scientific response to the unfolding changes. However, is the science community continuing to do its part to the best of its ability? In the domains where we can have the greatest influence, is the scientific community articulating a vision commensurate with the challenges posed by climate change? We think not.

climate change | climate models | high resolution

The idea that the science of climate change is largely "settled," common among policy makers and environmentalists but not among the climate science community, has congealed into the view that the outlines and dimension of anthropogenic climate change are understood and that incremental improvement to and application of the tools used to establish this outline are sufficient to provide society with the scientific basis for dealing with climate change. For certain, some things are settled. We know that greenhouse gases are accumulating in the atmosphere as a result of human activity and that they are largely responsible for warming of surface temperatures globally. We also are confident in our understanding as to why this warming is expected to be amplified over land masses and the Arctic. Likewise, we are confident in our understanding of how the hydrological cycle amplifies the effects of this warming and how warming amplifies the hydrological cycle. For these and other broad brush strokes of the climate change picture, we are also increasingly confident in our ability to usefully bound the magnitude of the effects. From this certainty stems the conviction that additional warming is best avoided by reducing or reversing emissions of long-lived greenhouse gases.

As climate scientists, we are rightfully proud of, and eager to talk about, our contribution to settling important and long-standing scientific questions of great societal relevance. What we find more difficult to talk about is our deep dissatisfaction with the ability of our models to inform society about the pace of warming, how this warming plays out regionally, and what it implies for the likelihood of surprises. In our view, the political situation, whereby some influential people and institutions misrepresent doubt about anything to insinuate doubt about everything, certainly contributes to a reluctance to be too openly critical of our models. Unfortunately, circling the wagons leads to false impressions about the source of our confidence and about our ability to meet the scientific challenges posed by a world that we know is warming globally.

How can we can reconcile our dissatisfaction with the comprehensive models that we use to predict and project global climate with our confidence in the big picture? The answer to this question is actually not so complicated. All one needs to remember is that confidence in the big picture is not primarily derived from the fidelity of comprehensive climate models of the type used to inform national and international assessments of climate change. Rather, it stems from our ability to link observed changes in climate to changes derived from the application of physical reasoning, often as encoded in much simpler models or in the case of the water cycle, through a rather simple application of the laws of thermodynamics. Comprehensive climate models have been effective and essential to address the concern that such a basic understanding could be overly simplistic (i.e., missing something important, such as the existence of a mode of internal variability, which could, if

[a]Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom; and [b]Max Planck Institute for Meteorology, D-20146 Hamburg, Germany

Author contributions: T.P. and B.S. developed the ideas and wrote the paper.

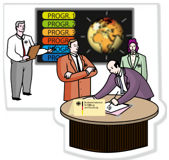
The authors declare no competing interest.

This article is a PNAS Direct Submission. J.S. is a guest editor invited by the Editorial Board.

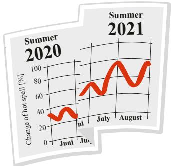
[1]To whom correspondence may be addressed. Email: Tim.Palmer@physics.ox.ac.uk or bjorn.stevens@mpimet.mpg.de.

First published December 2, 2019.

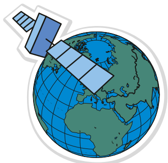
Box 1. Reliable global climate models are vital—for both societal applications and the advancement of science



Decision making: While the basis for decarbonizing rests on simple, unequivocal, physical principles, any policy or strategy that requires knowledge of climate change at the regional level (whether for climate adaptation or for quantifying the impact of solar radiation management) requires models and rests on their fidelity. Crucially, being able to anticipate what would otherwise be surprises in extreme weather and climate variations (possibly outside the range of those experienced to date) requires much more accurate and reliable models than we currently have.



Extended prediction: The predictability of slowly evolving components of the earth system provides a physical basis for useful prediction on timescales beyond those of weather forecasts. Experience with short-term forecasting, where predictability and bias reduction have gone hand in hand, teaches us that realizing this predictability hinges on the fidelity of the models making the predictions. Developing reliable tools for seasonal and decadal prediction will also help society become more resilient to the changing extremes of weather and climate caused by anthropogenic climate change.



Observations: An ability to infer the present or past state of the climate system depends critically on the fidelity of the models used to assimilate these observations. In many cases, weather forecasts fail not because of shortcomings in the observational data but because the models have insufficient resolution to assimilate key information contained in the observational data (22). As the means to measure the world becomes manifest in more and more ways (Internet of Things), improved models will be required to realize the potential of these measurements.

Theory: Theory is vital when making predictions about the future. Robustly simulated changes from models that more directly express physical laws beg explanation. Simulations will stimulate theory by focusing efforts on the salient question of the underpinnings of predicted changes rather than on the elaboration of parameterizations with little physical grounding. Conversely but in the same spirit, what models robustly fail to simulate has proven invaluable in shaping the present discourse on climate change, by discrediting alternate explanations for observed changes.

it were to exist, explain trends in global mean temperature). The enterprise of making models more and more comprehensive through the incorporation of computationally expensive* but poorly understood additional processes has not so much sharpened our ability to anticipate climate change as left the blurry picture established by physical reasoning and much simpler models intact (1). When it comes to global climate change, it is what the present generation of comprehensive climate models do not show—namely, a sensitivity of global changes to either the vagaries of unpredictable regional or global circulations or effects of processes neglected in simpler models—which makes them such a powerful confirmation of inferences from basic physics and simple models.

Now that the blurry outlines of global climate change have been settled, the need to sharpen the picture (Box 1) has become more urgent (2). However, such sharpening is proving to be more challenging than anticipated—something that we attribute to the inadequacy of our models (3–5). Unfortunately, many in the community—notably those in charge of science funding—have no idea how significant and widespread these inadequacies are. This has arisen in part because of a justified desire to communicate, with as much clarity as possible, the aspects of our science that are well settled. While we are certainly not claiming that model inadequacies cast doubt on these well-settled issues, we are claiming that, by deemphasizing what our models fail to do, we inadvertently contribute to complacency with the state of modeling. This leaves the scientific consensus on climate change vulnerable to specious arguments that prey on obvious model deficiencies; gives rise to the impression that defending the fidelity of our most comprehensive models is akin to defending the fidelity of the science; and most importantly, fails to communicate the need and importance of doing better.

Fig. 1, which is taken from the *Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (13), illustrates this situation well. It shows that all of the climate models can adequately reproduce the observed change in temperature—part of what we call the blurry outline of climate change. This is something that the Assessment Report draws attention to in its summary for policy makers. What is not discussed in the summary is what is shown by the thin horizontal lines on the edge of Fig. 1. Even after being tuned to match observed irradiance at the top of the atmosphere,† models differ among themselves in their estimates of surface temperature by an amount that is 2 to 3 times as

*In a comprehensive Earth system model, the computational cost of parametrized representations of poorly known Earth system processes typically exceeds the computational cost of the well-understood fluid mechanics represented by the dynamical core, sometimes by a considerable margin.

†Because parametrizations are only approximate representations of physical processes, model parameters are not well constrained by theory. As a result, they are typically tuned, or optimized, to ensure that there is a good balance between incoming solar and outgoing terrestrial irradiances at the top of the atmosphere and that this balance is achieved at as close to the observed global temperature as possible (6).

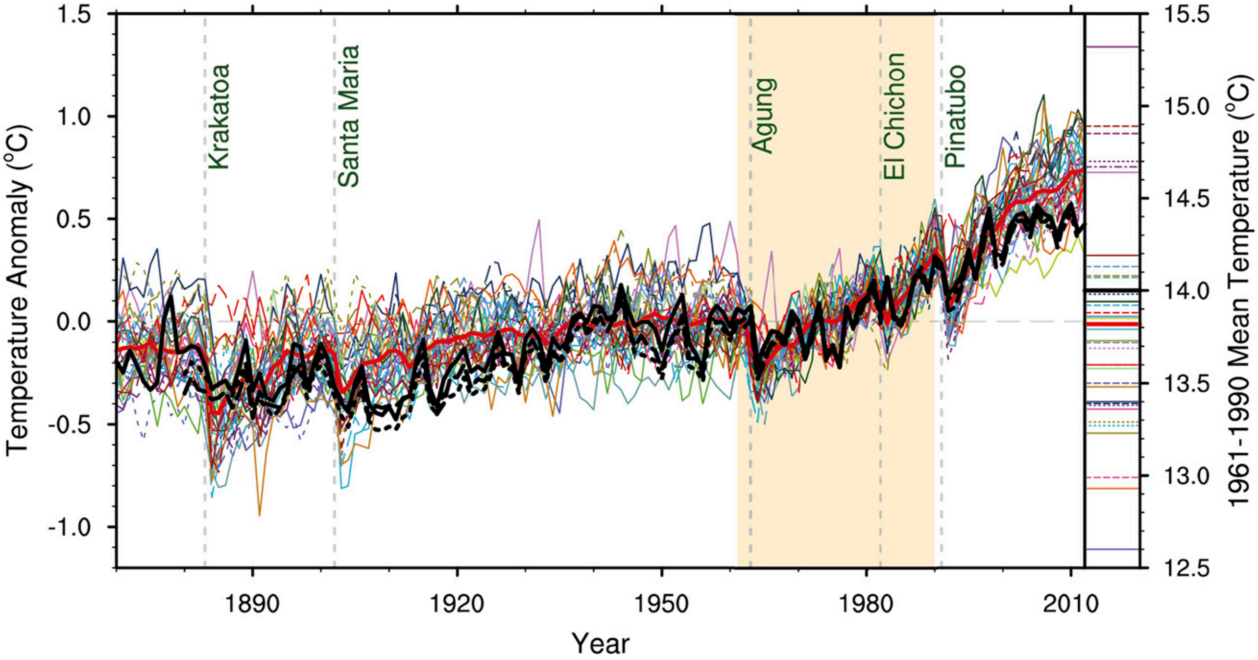


Fig. 1. Estimates of global surface temperature anomaly from model integrations performed in support of the fifth phase of the coupled model intercomparison project (CMIP5). Here, the observed anomaly (black) is estimated relative to the observed climatology, while the model anomalies are estimated relative to each model's estimate of climatology (rather than the multimodel ensemble mean estimate of climatology). This has the implicit effect of removing each model's individual climatological bias against observations. The range of uncorrected model temperatures for the period 1961 to 1990 is shown in a small bar to the right of the figure. While uncertainty in observational estimate of absolute temperatures is about 0.5 °C, uncertainty in the anomalies is much smaller. As such, the spread in model anomalies relative to the multimodel ensemble mean is larger than the observed trend in the anomalies and much larger than the uncertainty in the observed anomalies. Adapted with permission from ref. 13.

large as the observed warming and larger yet than the estimated 0.5 °C uncertainty in the observations. The deemphasis of this type of information, while helpful for focusing the reader on the settled science, contributes to the impression that, while climate models can never be perfect, they are largely fit for purpose. However, for many key applications that require regional climate model output or for assessing large-scale changes from small-scale processes, we believe that the current generation of models is not fit for purpose.

Figs. 2 and 3 develop this point further by showing how, on the regional scale and for important regional quantities (7), these problems are demonstrably more serious still, as model bias (compared with observations) is often many times greater than the signals that the models attempt to predict. In a nonlinear system, particularly one as important as our model of Earth's climate, one cannot be complacent about biases with such magnitudes. Both basic physics and past experience (at least on timescales that observations constrain) teach us that our ability to predict natural fluctuations of the climate system is limited by such biases (8–11). By downplaying the potential significance that model inadequacies have on our ability to provide reliable estimates of climate change, including of course in terms of extremes of weather and climate, we leave policy makers (and indeed, the public in general) ignorant of the extraordinary challenge it is to provide a sharper and more physically well-grounded picture of climate change, essentially depriving them of the choice to do something about it.

What is needed is the urgency of the space race aimed, not at the Moon or Mars, but rather toward harnessing the promise of exascale supercomputing to reliably simulate Earth's regional climate (and associated extremes) globally. This will only be possible if the broader climate science community begins to articulate its dissatisfaction with business as usual—not just among themselves but externally to those who seek to use the models for business, policy, or humanitarian reasons. Failing to do so becomes an ethical issue in that it saddles us with the status quo: a strategy that hopes, against all evidence, to surmount the abyss between scientific capability and societal needs on the back of 2 less-than-overwhelming ideas: 1) that post processing (i.e., empirically correcting or selectively sampling model output) can largely eliminate the model systematic biases that would otherwise make the models unfit for purpose (12) and 2) that incremental changes in model resolution or parametrization can overcome structural deficiencies that otherwise plague the present generation of models (13). Empirical bias corrections can work well for systems that behave linearly, but a large part of the reason for developing comprehensive models is to account for, or help anticipate, the climate system's nonlinearities (14). Were this not reason enough (15) to be dissatisfied with a disproportionate emphasis on idea 1, then surely lessons from numerical weather prediction—where an increase in skill has gone hand in hand with a reduction in systematic error—are (11, 16). As for idea 2, after 6 generations of model development spanning roughly 30 y, it seems safe to conclude that the task that we set for ourselves was, if not impossible, then too difficult to be achieved by incremental improvements alone. Even the most optimistic assessment of model development efforts cannot avoid concluding that progress has been far too slow to justify continuing to prioritize the present path.

This status quo and the complacency that surrounds it give us cause to be deeply dissatisfied with the state of the scientific response to the challenges posed by global warming. Whereas present day climate models were fit for the purpose for which they were initially developed, which was to test the basic tenets of our understanding of global climate change, they are inadequate

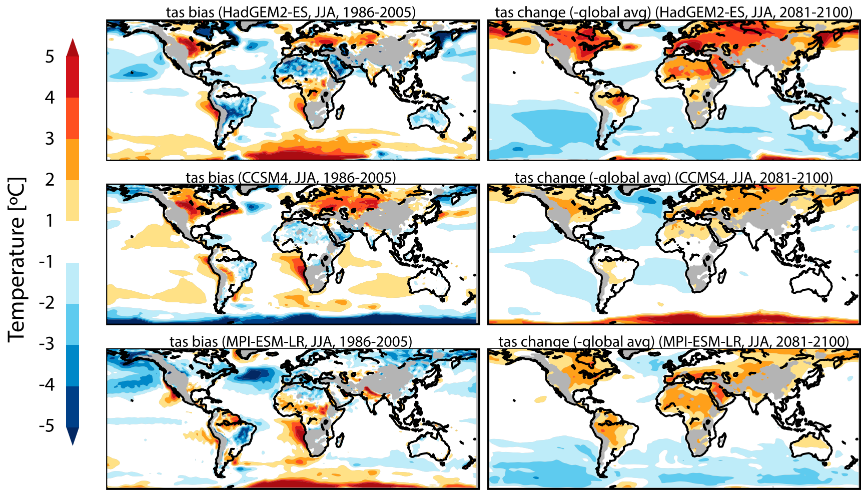


Fig. 2. The left-hand column shows the systematic error of 3 of the leading CMIP5 models for June, July, and August (JJA) mean surface air temperature (tas) computed using ensemble integrations and observations for the period 1986 to 2005. The right-hand column shows the same 3 models' regional temperature response to climate change forcing (computed as the departure from the global mean) based on differences between ensemble integrations for the period 2081 to 2100 following the RCP8.5 scenario and the period 1986 to 2005. The same scale is used for both columns. There arewidespread regions where the magnitude of systematic error is comparable with or exceeds the magnitude of the climate change signal.

for addressing the needs of society struggling to anticipate the impact of pending changes to weather and climate.

If climate science aspires to be relevant to societies dealing with climate change, a new strategy (17) is required (Box 2). A fresh wind, in the form of a step change in the physical content and fidelity of climate simulation systems, must be let loose to fan the flame of basic climate science to challenge our understanding of how global warming becomes manifest in regional climate and its subsystems. The stirrings of such a wind are beginning to be felt, as in different laboratories around the world, experimental efforts aimed at harnessing exascale computing to surmount roadblocks, known to limit the fidelity of existing simulation systems, are taking shape (18–21). However, if these stirrings are to grow to the gale required to give impetus to theory and observations and if society is to fully realize the ensuing benefits, these efforts must be scaled up through bold, sustained, and coordinated multinational initiatives.

As our nonlinear world moves into uncharted territory, we should expect surprises. Some of these may take the form of natural hazards, the scale and nature of which are beyond our present comprehension. The sooner we depart from the present strategy, which overstates an ability to both extract useful information from and incrementally improve a class of models that are structurally ill suited to the challenge, the sooner we will be on the way to anticipating surprises, quantifying risks, and addressing the very real challenge that climate change poses for science. Unless we step up our game, something that begins with critical self-reflection, climate science risks failing to communicate and hence realize its relevance for societies grappling to respond to global warming.

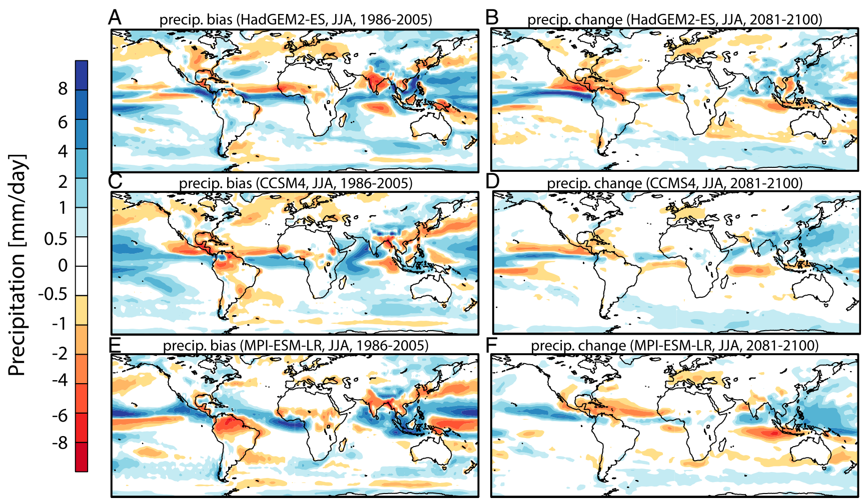
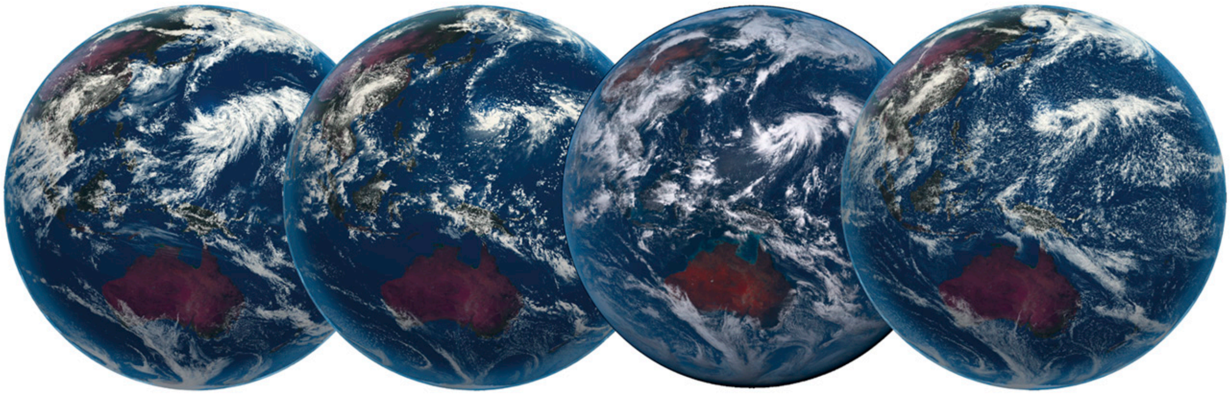


Fig. 3. *A*, *C*, and *E* show the systematic error of the 3 CMIP5 models in Fig. 2 for June, July, and August (JJA) mean precipitation (based on GPCP observational data) computed using ensemble integrations and observations for the same period as in Fig. 2. *B*, *D*, and *F* show the same 3 models' response to climate change forcing based on differences between ensemble integrations for the same period and scenario as in Fig. 2. The same scale is used for *A*–*F*. There are widespread regions where the systematic error exceeds the climate change signal (in some regions by more than a factor of 20).

Box 2. A new approach to global climate modeling

We are suggesting a new approach to climate model development (23). This approach should aim to reduce climate models' dependence on subgrid parameterizations where possible and account for their uncertainty where not. To be successful, this approach must master and motivate technological innovations, particularly in computing, and be given a sense of purpose commensurate to the task at hand.



Global storm and ocean-eddy resolving [O(1 km)] models make it possible to directly simulate deep convection, ocean mesoscale eddies, and important land–atmosphere interactions. Prototypes of such models are already being developed (21), 3 examples of which are compared with a satellite image. By avoiding the need to represent essential processes by semiempirical parameterizations, the simulated climate of such a model is more constrained by the laws of physics. This can be expected to lead to the reduction or even elimination of many systematic biases that plague the present generations of models (24–31).

Commensurate with this focus on high resolution, uncertainties in the parameterization of remaining subgrid (or nonfluid-dynamical) processes should be represented explicitly through some application of stochastic modeling (32). Among other advantages, this will ensure that such parameterizations are not unjustifiably complex—or overfit to past changes—and that the numerical precision of both parameterizations and dynamical cores is commensurate with their information content (33). Data-driven methods could also play an important role in reducing computational costs and in improving the representation of processes that cannot be constrained by first principles (34).

We can expect that such models will have substantially reduced biases against observations and a better characterization of uncertainty. Already, signs of improved prediction skill for tropical intraseasonal variability are emerging from prototype models with convective permitting resolution (35). This will radically enhance the reliability of regional climate projections. However, by how much? There is no overarching theory of the Navier–Stokes equation from which such questions can be answered [indeed, some of the most basic mathematical properties of these equations are still unknown (36)]. Nor can we presently partition uncertainty in predictions among fluid-dynamical vs. nonfluid-dynamical (cloud microphysics, land-surface properties, biogeochemistry) contributions to climate change. This means that the only way to answer such questions is to actually develop and run such models. This is largely the way in which the benefits of increased resolution in numerical weather prediction were quantified. These benefits have been so enormous that no one would seriously contemplate returning to weather forecast model resolutions of say the 1980s and utilize the saved computer resources for other applications (37).

The development of this new generation of models should be sustained, multinational, and coordinated as a flagship application of high-performance computing and information technology. Only as a coordinated technology project will it be possible to meet the computational challenges of running the highest possible resolution models and accessing their full information content. How to structure such an initiative can be debated; indisputable is the necessity to endow it with the same sense of purpose that has made past grand scientific challenges—from weather forecasting to moon landings—so successful.

The importance of developing predictions of climate with reliable regional precision is so important that we simply have to give this our best shot. Failing to do so keeps society in the dark about the possible ways that our climate system might develop in the coming years and decades.

Data Availability

The data used to produce Figs. 2 and 3 were downloaded from the Coupled Model Intercomparison Project 5 (CMIP5) archive and can be accessed from https://esgf-node.llnl.gov/projects/cmip5/. The observational data for precipitation were from the Global Precipitation Climatology Project (GPCP) monthly precipitation dataset (on a 2.5° grid), which can be downloaded from https://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html. The surface temperature observational dataset is from the the interim reanalysis of observational data by the European Centre for Medium Range Weather Forecasts (ERA-interim). It is provided on a 1° grid, which can be downloaded from https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim. The model data were regridded to the corresponding observational grid prior to the analysis.

Acknowledgments

We acknowledge discussions with Peter Bauer in the development of our ideas and acknowledge helpful comments of 3 anonymous referees and the editor. N. Noreiks is thanked for help in drafting Box 1. D. Befort and C. O'Reilly are thanked for help in producing Figs. 2 and 3. T.P. acknowledges the support of the European Research Council Advanced Grant ITHACA Grant 741112.

1 S. Bony *et al*., "Carbon dioxide and climate: Perspectives on a scientific assessment" in *Climate Science for Serving Society*, G. R. Asrar, J. W. Hurrell, Eds. (Springer Netherlands, Dordrecht, the Netherlands, 2013), pp. 391–413.
2 J. Marotzke *et al*., Climate research must sharpen its view. *Nat. Clim. Chang.* **7**, 89–91 (2017).
3 B. Stevens, S. Bony, Climate change. What are climate models missing? *Science* **340**, 1053–1054 (2013).
4 C. Laufkötter *et al*., Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences* **12**, 6955–6984 (2015).
5 T. G. Shepherd, Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.* **7**, 703–708 (2014).
6 T. Mauritsen *et al*., Tuning the climate of a global model. *J. Adv. Model. Earth Syst.* **4**, M00A01 (2012).
7 F. J. Tapiador, R. Roca, A. Del Genio, B. Dewitte, F. Zhang, Is precipitation a good metric for model performance? *Bull. Am. Meteor. Soc.* **100**, 223–233 (2019).
8 E. Hawkins, R. Sutton, The potential to narrow uncertainty in projections of regional precipitation change. *Clim. Dyn.* **37**, 407–418 (2011).
9 J. J. Barsugli, P. D. Sardeshmukh, Global atmospheric sensitivity to tropical SST anomalies throughout the Indo-Pacific basin. *J. Clim.* **15**, 3427–3442 (2002).
10 Y. Dong, C. Proistosescu, K. Armour, D. S. Battisti, Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a green's function approach: The preeminence of the Western Pacific. *J. Clim.* **32**, 5471–5491 (2019).
11 D. M. H. Sexton *et al*., Finding plausible and diverse variants of a climate model. Part 1. Establishing the relationship between errors at weather and climate time scales. *Clim. Dyn.* **53**, 989–1022 (2019).
12 S. A. Klein, A. Hall, Emergent constraints for cloud feedbacks. *Curr. Clim. Change Rep.* **1**, 276–287 (2015).
13 G. Flato *et al*., "2013: Evaluation of climate models" in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al*., Eds. (Cambridge University Press, Cambridge, United Kingdom, 2013), pp. 741–866.
14 W. Steffan *et al*., Trajectories of the Earth system in the anthropocene. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8252–8259 (2018).
15 T. N. Palmer, A. Weisheimer, A simple pedagogical model linking initial-value reliability with trustworthiness in the forced climate response. *Bull. Am. Meteorol. Soc.* **99**, 605–614 (2018).
16 T. J. Phillips *et al*., Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. *Bull. Am. Meteorol. Soc.* **85**, 1903–1916 (2004).
17 T. N. Palmer, A personal perspective on modelling the climate system. *Proc. Math. Phys. Eng. Sci.* **472**, 20150772 (2016).
18 M. Satoh *et al*., Outcomes and challenges of global high-resolution non-hydrostatic atmospheric simulations using the K computer. *Prog. Earth Planet. Sci.* **4**, 13 (2017).
19 O. Fuhrer *et al*., Near-global climate simulation at 1km resolution: Establishing a performance baseline on GPUs with COSMO 5.0. *Geosci. Model Dev.* **11**, 1665–1681 (2018).
20 T. Schneider, S. Lan, A. Stuart, J. Teixeira, Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. *Geophys. Res. Lett.* **44**, 12396–12417 (2017).
21 B. Stevens *et al*., DYAMOND: The DYnamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Prog. Earth Planet. Sci.* **6**, 61 (2019).
22 M. J. Rodwell *et al*., Characteristics of occasional poor medium-range weather forecasts for Europe. *Bull. Am. Meteorol. Soc.* **94**, 1393–1405 (2013).
23 J. Shukla *et al*., Toward a new generation of world climate research and computing facilities. *Bull. Am. Meteorol. Soc.* **91**, 1407–1412 (2010).
24 J. Rougier, D. M. H. Sexton, J. M. Murphy, D. A. Stainforth, Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments. *J. Clim.* **22**, 3540–3557 (2009).
25 M. J. Miller, T. N. Palmer, R. Swinbank, Orographic gravity-wave drag: Its parametrization and influence in general circulation and numerical weather prediction models. *Meteorol. Atmos. Phys.* **40**, 84–109 (1989).
26 D. R. Munday, H. L. Johnson, D. P. Marshall, Eddy saturation of equilibrated circumpolar currents. *J. Phys. Oceanogr.* **43**, 507–532 (2013).
27 K. A. Browning *et al*., The GEWEX cloud system study (GCSS). *Bull. Am. Meteorol. Soc.* **74**, 387–399 (1993).
28 E. W. Doddridge *et al*., Eddy compensation dampens Southern Ocean sea surface temperature response to westerly wind trends. *Geophys. Res. Lett.* **46**, 4365–4377 (2019).
29 J. C. McWilliams, A survey of submesoscale currents. *Geosci. Lett.* **6**, 3 (2019).
30 D. Panosetti, L. Schlemmer, C. Schär, Bulk and structural convergence at convection-resolving scales in real-case simulations of summertime moist convection over land. *Q. J. R. Meteorol. Soc.* **145**, 1427–1443 (2019).
31 M. Satoh *et al*., Global cloud-resolving models. *Curr. Clim. Change Rep.* **5**, 172–184 (2019).
32 T. N. Palmer, Stochastic weather and climate models. *Nat. Rev. Phys.* **1**, 463–471 (2019).
33 M. Chantry, T. Thornes, T. N. Palmer, P. Düben, Scale-selective precision for weather and climate forecasting. *Mon. Weather Rev.* **147**, 645–655 (2019).
34 M. Reichstein *et al*., Deep learning and process understanding for data-driven Earth system science. *Nature* **566**, 195–204 (2019).
35 N. J. Weber, C. F. Mass, Sub-seasonal weather prediction in a global convection-permitting model. *Bull. Am. Meteorol. Soc.* **100**, 1079–1090 (2019).
36 T. N. Palmer, A. Doering, G. Seregin, The real butterfly effect. *Nonlinearity* **27**, R123–R141 (2014).
37 P. Bauer, A. Thorpe, G. Brunet, The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55 (2015).