BAMS
Meeting Summary

Systematic Errors in Weather and Climate Models

Challenges and Opportunities in Complex Coupled Modeling Systems

Ariane Frassoni, Carolyn Reynolds, Nils Wedi, Zied Ben Bouallègue,
Antonio Caetano Vaz Caltabiano, Barbara Casati, Jonathan A. Christophersen,
Caio A. S. Coelho, Chiara De Falco, James D. Doyle, Laís G. Fernandes,
Richard Forbes, Matthew A. Janiga, Daniel Klocke, Linus Magnusson,
Ron McTaggart-Cowan, Morteza Pakdaman, Stephanie S. Rushley, Anne Verhoef,
Fanglin Yang, and Günther Zängl

6th WGNE Workshop on Systematic Errors in Weather and Climate Models

What: Scientists, ranging from early career to highly experienced, involved in the development of weather and climate models and in the diagnosis of model errors, held an international workshop to discuss the nature, causes, and remedies of systematic errors across time scales and across Earth system modeling components.

When: 31 October–4 November 2022

Where: Reading, United Kingdom

KEYWORDS: Climate models; Coupled models; Model errors; Model evaluation/performance; Numerical weather prediction/forecasting

https://doi.org/10.1175/BAMS-D-23-0102.1

Corresponding author: Ariane Frassoni, ariane.frassoni@inpe.br

In final form 10 May 2023

AFFILIATIONS: Frassoni and Coelho—Brazilian National Institute for Space Research, Center for Weather Forecasting and Climate Studies, Cachoeira Paulista, São Paulo, Brazil; **Reynolds, Christophersen, Doyle, Janiga, and Rushley**—U.S. Naval Research Laboratory, Monterey, California; **Wedi, Bouallègue, Forbes, and Magnusson**—European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom; **Caltabian**o—World Climate Research Programme, Geneva, Switzerland; **Casati and McTaggart-Cowan**—Environment and Climate Change Canada, Dorval, Quebec, Canada; **De Falco**—NORCE Norwegian Research Center, and Bjerknes Center for Climate Research, Bergen, Norway; **Fernandes**—Portland State University, Portland, Oregon; **Klocke**—Max Planck Institute for Meteorology, Hamburg, Germany; **Pakdaman**—Research Institute of Meteorology and Atmospheric Science, Climate Research Institute, Mashhad, Iran; **Verhoef**—Department of Geography and Environmental Science, University of Reading, Reading, United Kingdom; **Yang**—National Centers for Environmental Prediction, College Park, Maryland; **Zängl**—Deutscher Wetterdienst, Offenbach, Germany

The Working Group on Numerical Experimentation (WGNE) organized its 6th Workshop on Systematic Errors (WSE) in Weather and Climate Models, hosted by the European Centre for Medium-Range Weather Forecasts (ECMWF) on 31 October–4 November 2022. The workshop brought together a wide range of experts on simulating the Earth system to advance the understanding of the root causes of systematic model errors across time scales. Here we highlight the progress made in diagnosing and addressing systematic errors since the 5th WGNE WSE (Zadra et al. 2018, hereafter Z18). The hybrid WSE facilitated online and in-person participation with a mix of keynote and contributed oral presentations and dedicated in-person and online poster sessions. The workshop was attended by close to 200 scientists and included 41 oral and 88 poster presentations. An abstract competition for Early Career Scientists (ECS) from selected countries was sponsored by the World Climate Research Programme (WCRP) Earth System Modelling and Observations Core Project. Among the 20 competitors, three winners presented their work during a dedicated oral session. Interactive involvement was encouraged through in-person and online breakout sessions. Early career scientists were invited to serve as session co-chairs and rapporteurs to facilitate interactions across a range of experience levels.

Key topics

The workshop was organized around seven themes: 1) errors in the representation of clouds and precipitation; 2) coupled atmosphere–ocean–land–cryosphere system interactions; 3) (sub-)tropical circulations including errors in the simulation of tropical–extratropical teleconnections; 4) stratosphere–troposphere interactions; 5) novel techniques with particular emphasis on machine learning (ML) and data assimilation (DA) to diagnose, measure, and resolve systematic errors; 6) quantifying uncertainty; and 7) challenges and surprises in simulating the climate system. Breakout group discussions focused on prioritization of the systematic errors that still need to be addressed and recommendations for ways to move forward to reduce errors in coupled systems. Oral sessions and summaries of the breakout discussions were recorded and are accessible on the workshop website (https://events.ecmwf.int/event/241).

Highlights

Clouds and precipitation. While the development of kilometer-scale Earth system models (ESMs) has accelerated in recent years and has corrected some long-standing systematic errors, key issues remain. Since Z18, the community has made progress through advancements in the representation of precipitation-related processes including timing, propagation, statistical characteristics, and the diurnal cycle. Such achievement also benefits subseasonal and climate simulations. In the tropics, coupled global models are now able to represent the seasonal migration of the precipitation belts and the main characteristics of summer monsoon convection using these high-resolution modeling systems. However, accurately simulating oceanic convection and precipitation, and the amplitude of the diurnal cycle of precipitation over land remains a challenge. Systematic errors over the western Pacific and Indian Ocean persist in kilometer-scale coupled models, including the double intertropical convergence zone (ITCZ). The southeast Pacific Ocean stratus cloud deck is still misrepresented in kilometer-scale simulations despite improvements in parameterized and explicit shallow convection.

Convection and precipitation biases are influenced by ocean biases and depend on the coupling methodology, particularly the use of shallow and fast-responding interface layers. Improvements to the treatment of planetary boundary layer (PBL) processes and their interactions across the ocean–land–atmosphere interfaces show potential for improving the representation of shallow clouds and their radiative feedback. Large biases remain in the representation of clouds in the Arctic region, with models suffering from an underestimation of supercooled liquid water in mixed-phase clouds. Specific deficiencies in the parameterization of cloud processes have been linked to model radiation errors through the use of DA diagnostics and short-range forecasts. The increasing use of direct and indirect observations of clouds in DA has the potential to better constrain model cloud properties, such as the amount of condensate, cloud phase, vertical structure, particle properties, and their impact on radiation.

Atmosphere–ocean–land–cryosphere interactions. Increased horizontal and vertical resolution in ocean models was identified by Z18 as a way to reduce systematic errors in sea surface temperature (SST), salinity, Gulf Stream separation, and deep ocean properties. Higher-resolution simulations have recently become more widely available, enabling scientists to study small-scale ocean characteristics. While ocean eddies have a significant impact on the transport of mass, heat, and tracers within the ocean, challenges remain in representing ocean variability because errors arise from multiple sources, including subgrid-scale parameterizations. New approaches and improvements in the parameterization of ocean turbulent flow strongly impact model simulations and reduce systematic errors. There has been progress in the design of parameterizations that reduce excessive dissipation of kinetic energy and that are capable of reducing model biases in SST, sea surface height, salinity, and regional variability. However, challenges remain in applying ocean turbulent flow parameterizations in global ocean forecast and assimilation systems.

Long-standing systematic errors also persist at the ocean–atmosphere interface. Surface flux biases in Coupled Model Intercomparison Project phase 6 models vary throughout the convective life cycle and lead to erroneous ocean feedbacks on convective development. These lead to biases in the Madden–Julian oscillation (MJO) amplitude and propagation, as well as to errors in the forcing of oceanic Kelvin waves, El Niño–Southern Oscillation, and associated teleconnections. Refined bulk flux algorithms improve MJO propagation and reduce the double-ITCZ bias. In coupled models, MJO simulations may be improved by increasing the frequency of the coupling time step. Recent advances using conditional sampling and intercomparison of surface flux diagnostics attempt to further point to the complex sources of systematic errors in MJO simulations.

Short-term simulations of the winter Arctic atmosphere and surface energy budgets were validated against the Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAiC) observations. Only coupled modeling systems accurately simulated Arctic radiation, turbulence, and cloud processes. However, the accurate representation of supercooled liquid clouds, persistent stable PBL, and distinguishing cloudy/clear-sky states remains challenging. Additionally, surface heat fluxes over sea ice leads can affect the modeling of surface energy budgets in the Arctic wintertime. The poor representation of these fluxes in lower-resolution global models causes systematic errors in the region. Despite recent progress in kilometer-scale models to resolve lead-forming processes and consequent surface fluxes over these fractures that expose open water, challenges persist in reducing errors, including those in low-cloud cover, sea ice thickness distributions, and near-surface temperatures. To address these issues, a new parameterization based on the proportional relationship between sensible heat flux and atmospheric stability over sea ice leads has shown promise in improving predictions of these quantities.

Incorrect/incomplete treatment of the land surface often results in systematic errors. The representation of surface processes and parameters varies considerably from model to model, which subsequently results in large variations in atmosphere–land coupling. Therefore, current research in land surface models focuses on advanced representation of vegetation processes – including those in the terrestrial carbon cycle—efforts to improve hydrological processes and flood prediction in ESMs, how to represent anthropogenic activity at the kilometer-scale, and improved soil parameterizations and input data (e.g., better soil maps, soil properties, more soil layers/depth, representation of the effect of spatiotemporally variable soil structure). Offline analysis has helped to identify systematic errors related to atmosphere–land surface coupling. Satellite-based and in situ observations are fundamental to diagnosing errors and improving the representation of land surface processes, in particular in global models.

(Sub-)tropical circulations. Systematic errors in tropical cyclone (TC) intensity and track are sensitive to parameterizations of turbulence, radiation, and moist processes. Storm intensity and the diagnosed pressure–wind relationship are dependent on the surface drag coefficient (including wave model effects) and are affected by numerical dissipation. Air–sea coupling in general reduces overintensification, particularly for slow-moving storms. Very high resolution (1 km) is needed to capture sharp gradients in the inner core and improve the structure of small-scale systems. Rapid intensification of TCs is notoriously difficult to predict, although recent progress has been made using kilometer-scale models. More research is needed on secondary eyewall formation and inner core dynamics, which greatly influence TC intensity and structure.

Despite being a topic of great interest, substantial MJO simulation errors remain (Z18), including biases in frequency, amplitude, speed, growth, decay, and traversing of the Maritime Continent. These errors affect predictions of phenomena that are impacted by the MJO, such as TC genesis. Process-based diagnostics have been used to link MJO intensity and propagation errors to specific model characteristics, such as biases in vertical advection and convection-related moisture adjustment time scales. In-line bias correction methods have been shown to improve MJO simulations by improving the model's basic state, convective parameterization, and representation of near-surface processes. Recommendations for future work include kilometer-scale modeling using integrated parameterizations of PBL and moist convection, perturbed parameter simulations, and comparison of MJO predictions from initial value versus boundary-forced (climate) simulations.

Stratosphere–troposphere interactions. Understanding how increasing the horizontal resolution of global models changes resolved gravity wave forcing is essential because of the control that this process exerts on the general circulation. Global simulations with grid spacings down to 1 km are helpful to understand the representation of resolved gravity waves, to evaluate drag parameterizations and to inspire the development of improved schemes. Even at 3–5-km horizontal grid spacing, gravity waves and their sources are not fully resolved; therefore, we need to parameterize their effects. Model biases can impact subseasonal forecast skill by influencing stratosphere–troposphere coupling, with many models suffering from a similar set of systematic errors, including a global-mean warm bias at the stratopause, a mid-to-lower-stratospheric cold bias in the tropics, a lower-stratospheric cold bias in the Northern Hemispheric summer, and high polar mid-to-upper-stratospheric temperatures in the winter hemisphere. It is now understood that high-top models with sufficient vertical resolution are needed to address stratospheric biases, including simulation errors of the quasi-biennial oscillation. Increasing horizontal resolution with kilometer-scale models can help to resolve more of the spectrum of vertically propagating gravity waves but can also introduce important new stratospheric biases that must be carefully considered.

ML and DA: Novel approaches to diagnose, measure, or reduce systematic errors. Z18 highlighted the need for new observation-based techniques to tune parameterizations. Data-driven approaches have seen a dramatic increase in attention as ML techniques gain popularity. Examples discussed were the relationship between the marine low cloud fraction and meteorological factors that are directly related to model parameterizations, the use of ensemble-based ML algorithms to detect relationships between meteorological parameters in simulations and observations, and the deployment of ML techniques to explore and optimize model parameters. Information from DA has been used to adaptively optimize near-surface parameters (e.g., 2-m temperature) by adjusting uncertain parameters in land surface schemes. Also, hybrid physical model–ML techniques offer a computationally efficient approach to adding ML-based prognostic variables to dynamical model guidance.

Reducing systematic errors through bias correction becomes more challenging as model complexity increases. However, through ML techniques there is renewed interest in the topic. An ocean tendency adjustment technique that accounts for errors associated with model component coupling was proposed at the WSE. The method prognostically applies the climatological increments as a tendency correction term to reduce model errors. It is expected that reductions in ocean model drift will limit drift in the other model components. Analysis increments also have the potential to identify errors before significant feedback occurs. ML techniques were suggested not only to reduce systematic errors but also to diagnose them. For example, causal networks can be used to identify pathways of model biases. Convolutional neural network techniques have also been used to identify causal relationships between the phases of the MJO and warm conveyor belts.

Quantifying uncertainty. Stochastic parameterizations have been employed to represent model uncertainty and to reduce some forms of systematic error (e.g., double-ITCZ biases). Techniques using DA and ML, including genetic algorithms, are alternative ways to move forward with representing and quantifying model uncertainty. Solutions to reduce mean-state errors include both physical model improvements and pragmatic (DA and bias correction) methods. The technique of emergent constraints is also used in climate modeling as a way to reduce uncertainty in the predicted changes of poorly constrained quantities (e.g., precipitation) in a warming climate.

The “Different Models, Same Initial Conditions” project, led by WGNE and presented during the WSE, aims to identify model errors associated with different model formulations. High-quality forecasts can be produced by models when provided with the same high-quality analyses, despite a wide range of model biases. Shared physical parameterizations can lead to similar forecast errors across different models. The multimodel ensemble spread is indicative of important forecast sensitivities to various model formulations, providing evidence of the benefits of diversity in model design for better comprehension of model errors.

Challenges and surprises in simulating the climate system. This session provided a forum to present and discuss the successes and challenges associated with kilometer-scale global modeling. Despite significant progress, increased resolution does not necessarily improve the representation of the large-scale flow and convergence with resolution is not guaranteed. The representation of cloud microphysics is increasingly important as convective motions are better resolved, and the dynamics and radiation are both sensitive to the microphysics formulation and parameters (e.g., fall speeds of hydrometeors, number of moments), that can affect local thermodynamics and cloud–radiation interactions. Advances in microphysics parameterization show the potential to improve predictions of storm dynamics while reducing systematic radiation errors. Simulations of moisture diffusion around deep convection and resultant midlevel moistening can also be improved through refinement of turbulent transport parameterizations. The role of increasing vertical resolution in representing many features, such as the midlatitude circulation, was also highlighted.

The importance of the feedback between transient eddies in the atmosphere and large-scale flow anomalies on seasonal or longer time scales was emphasized. This feedback has been linked to the signal-to-noise paradox, where climate models reflect observed climate variability better than would be expected from their own ensemble statistics. Addressing deficiencies in representing such eddy feedback in models has been linked to improved skill in extratropical regions, suggesting an area for future assessment. Kilometer-scale resolution is also fundamental in reducing ocean model biases, such as the cold SST bias in the central North Atlantic common to many ocean models. At sufficient ocean model resolution, meridional SST gradient biases decrease, mitigating positive biases in low-level baroclinicity and associated errors in atmospheric static stability and diabatic heating. Atmospheric feedback with increased ocean model resolution results in improvements in the representation of European blocking and eddy-driven jet variability.

Conclusions and future plans

With advanced computational technologies, model resolution and complexity have dramatically increased over recent years. Since Z18, efforts in evaluating, testing, and improving models have been rewarded with essential reductions in critical systematic errors. While this has led to considerable improvements in predictive skill of models, some of the biases identified in Z18 remain, while new systematic errors have emerged. Model intercomparison studies can provide insight into how model formulation impacts model biases and will likely continue to guide physical model development and sharing of knowledge on systematic errors. A wide range of observations and field campaigns, including remote sensing, are crucial for verification and informing model development. These data are particularly useful when combined with reduced-order modeling and/or fine-scale simulations to aid understanding. Ensemble and ML-based approaches have shown significant promise for rigorous parameter estimation. Coupled with new stochastic approaches to uncertainty representation, such techniques have the promise to extend the limits of practical predictability in the coming years.

The workshop attendees felt that it was challenging to prioritize systematic errors in terms of importance given the many differences among models and applications, and the need for a better understanding of complex interactions in tightly coupled systems. Using a hierarchy of models or conditional verification approaches can help to isolate and better understand sources of model errors. As a way to move forward to mitigate systematic errors in ESMs we recommend the following based on WSE outcomes:

- Continuously promote model intercomparison activities, especially among kilometer-scale ESMs.
- Employ high-resolution/digital twins of the Earth system for applications such as process studies and coarse-graining.
- Employ hierarchies of models, including single-column models and constrained components.
- Broaden the use of techniques such as ensemble sensitivity, parameter exploration, perturbation experiments, adjoint sensitivity, and relaxation-nudging experiments.
- Carefully consider the mechanism and impact of physics–dynamics and physics–physics cross-component coupling.
- Employ DA methodologies to identify systematic errors and constrain parameters.
- Employ ML to determine and optimize parameters, to identify flow-dependent systematic errors and/or to detect causal connections between seemingly disparate parameters.
- Promote model evaluation using high-resolution, ocean subsurface and process-relevant observations; observations in data-poor regions, particularly those across component interfaces are needed.
- Provide error estimates on reanalyses and observations.
- Weigh the risks and benefits of in-line bias correction versus model improvement.
- Share experience across mesoscale, regional, and global kilometer-scale modeling communities on a regular basis.
- Strengthen connections and communication between the weather and climate modeling communities through seamless prediction experiments and harmonized verification practices; initializing climate predictions and identifying the climatology of weather prediction models.
- Promote cooperation to provide land surface models with suitable and, ideally, dynamic kilometer-scale inputs (soil, vegetation, land use, and land management).
- Unify and standardize field campaign data, model data, and observation network repositories and inventories. Provide the data at various resolutions to account for the increasing size of these datasets.
- Entrain model developers and DA experts when designing field campaigns.
- Promote the career development of ECS and provide opportunities to improve scientific and technical skills in model development; actively involve ECS in shaping the future of Earth system modeling; increase diversity and make efforts to overcome geographic, cultural, and communication barriers.

Acknowledgments. On behalf of WGNE we thank ECMWF for hosting the workshop, and WMO and WCRP for the financial support, especially for sponsoring the in-person participation of ECS. We also thank the scientific committee, keynote speakers, session chairs, and ECS volunteers for their valuable contributions. A final word of gratitude goes to all participants, who made this workshop a memorable event. The authors wish to express their thanks for the financial support to publish this manuscript, which was provided by the Brazilian National Institute for Space Research.

Data availability statement. All materials presented during the WGNE WSE that supported our discussions and conclusions are freely available at https://events.ecmwf.int/event/241.

Reference

Zadra, A., and Coauthors, 2018: Systematic errors in weather and climate models: Nature, origins, and ways forward. *Bull. Amer. Meteor. Soc.*, **99**, ES67–ES70, https://doi.org/10.1175/BAMS-D-17-0287.1.