

# MEETING SUMMARIES

## SYSTEMATIC ERRORS IN WEATHER AND CLIMATE MODELS

Nature, Origins, and Ways Forward

AYRTON ZADRA, KEITH WILLIAMS, ARIANE FRASSONI, MICHEL RIXEN, ÁNGEL F. ADAMES, JUDITH BERNER, FRANÇOIS BOUYSSSEL, BARBARA CASATI, HANNAH CHRISTENSEN, MICHAEL B. EK, GREG FLATO, YI HUANG, FALKO JUDT, HAI LIN, ERIC MALONEY, WILLIAM MERRYFIELD, ANNELIZE VAN NIEKERK, THOMAS RACKOW, KAZUO SAITO, NILS WEDI, AND PRIYANKA YADAV

The Fifth Workshop on Systematic Errors in Weather and Climate Models (WSE) was hosted by Environment and Climate Change Canada (ECCC) under the auspices of the Working Group on Numerical Experimentation (WGNE), jointly sponsored by the Commission of Atmospheric Sciences of the World Meteorological Organization (WMO) and the World Climate Research Programme (WCRP). This major event welcomed over 200 scientists from the weather and climate communities. The workshop's primary goal was to increase understanding of the nature and cause of systematic errors in numerical models across time scales. Out of 240 abstracts submitted to the workshop, 48 talks and 132 posters were presented.

**KEY TOPICS.** This workshop offered a forum to identify systematic errors and physical processes that are not well represented in current weather and climate models. The program was organized around six themes: the coupled atmosphere–land–ocean–cryosphere system; errors in the representation of clouds and precipitation; resolution issues, including the representation of processes in the so-called gray zones; model errors in ensembles; errors in the simulation of teleconnections between the high/midlatitudes and tropics; and novel metrics and diagnostics.

**HIGHLIGHTS.** Presentations in the clouds and precipitation session focused on the parameterization

### FIFTH WORKSHOP ON SYSTEMATIC ERRORS IN WEATHER AND CLIMATE MODELS

**WHAT:** Hundreds of scientists involved in the development and evaluation of weather and climate models held an international workshop to discuss the nature and causes of systematic model errors across time scales.

**WHEN:** 19–23 June 2017

**WHERE:** Montreal, Quebec, Canada

development bottleneck and the problem of compensating errors. Verification of parameterization schemes helps to attribute biases to specific features of the numerical representation of clouds and precipitation that affect the intensity, distribution, diurnal cycle and timing of maximum precipitation, the transition between convective regimes, and the organization of convective systems. Improved treatments of cloud microphysics and boundary layer processes have been identified as ways to reduce uncertainties in low-cloud radiative feedbacks.

Errors in clouds and precipitation were also linked to biases in column-integrated water vapor, vertical velocity, and sea surface temperature (SST). These biases affect the atmospheric circulation and impact propagation, convective onset, and organization of the Madden–Julian oscillation (MJO), as well as limiting MJO predictability. Discussions also pointed out the limitations of mountain precipitation

forecasts resulting from the representation of subgrid-scale dynamics and deficiencies in the microphysics and convection schemes.

The use of explicit simulations, such as large-eddy simulations (LES) on wide geographical domains, gives rise to new opportunities to improve parameterizations. Such efforts linked with increased international collaboration are a major aim of the Global Atmospheric System Studies Panel.

The rise of convection-permitting climate models was discussed in the session on model resolution issues. These models are crucial for assessing the climate's sensitivity to convective clouds and associated systematic errors and uncertainties. While it is unclear whether high resolution is fundamental for climate projections, there is consensus that convection-resolving simulations can help eliminate systematic errors in weather forecasts and climate predictions, in particular in combination with stochastic perturbations.

However, convection-permitting models still suffer from substantial biases. Tropical convection and rainfall biases require process studies and better observations. Recent progress aims at the unification of boundary layer parameterization, convection in complex terrain, and gray zone/scale-aware parameterizations.

Increased resolution in ocean models is one way to reduce long-standing warm/saline biases and errors in the Gulf Stream separation and in the deep ocean. A promising method is to adapt resolution and discretization to the flow regimes (e.g., by local refinement). Key challenges covered include computational

issues, the difficulties involved with rewriting and optimizing code for emerging high-performance computing architectures, and big data issues of model output and bandwidth limitations. Inexact computing methods and decreased numerical precision might enable increases in ensemble size and/or resolution. Further topics included the relationship between model horizontal and vertical resolution and the impact of high-resolution global modeling on tropical cyclone prediction.

Presentations in the atmosphere–land–ocean–cryosphere interactions session indicated that accurate representation of the coupled system requires a better understanding of the processes involved, supported by diagnostic analysis and the availability of observations. For example, surface heat fluxes are not well observed at the global scale.

Several presentations referred to the Drag Project led by the WGNE ([http://collaboration.cmc.ec.gc.ca/science/rpn/drag\\_project/](http://collaboration.cmc.ec.gc.ca/science/rpn/drag_project/)). The findings demonstrated how different representations of resolved and subgrid orography can affect model performance at both short and long time scales.

More processes are gradually being included in models to represent the complexity of the Earth system. Traditionally, the fully coupled system has been seen as a climate or seasonal time-scale problem, but benefits of coupled processes, such as atmosphere–ocean interactions, were shown to improve weather forecast skill even in the short to medium range.

In the model errors in ensembles session, it was recognized that new metrics are needed to diagnose model errors. For initialized ensembles, reliability

**AFFILIATIONS:** ZADRA, CASATI, AND LIN—Recherche en Prévision Numérique, Environment and Climate Change Canada, Dorval, Quebec, Canada; WILLIAMS AND VAN NIEKERK—Met Office, Exeter, United Kingdom; FRASSONI—Center for Weather Forecasting and Climate Studies, National Institute for Space Research, Cachoeira Paulista, Brazil; RIXEN—World Climate Research Programme, World Meteorological Organization, Geneva, Switzerland; ADAMES—National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, and University Corporation for Atmospheric Research, Boulder, Colorado; BERNER, CHRISTENSEN, AND JUDT—National Center for Atmospheric Research, Boulder, Colorado; BOUYSSSEL—Centre National de Recherches Météorologiques, Météo-France, and Centre National de la Recherche Scientifique, Paris, France; EK—National Oceanic and Atmospheric Administration/National Centers for Environmental Prediction/Environmental Modeling Center, College Park, Maryland; FLATO AND MERRYFIELD—Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, British Columbia, Canada;

HUANG—Monash University, and Centre of Excellence for Climate System Science, Australian Research Council, Melbourne, Victoria, Australia; MALONEY—Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado; RACKOW—Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany; SAITO—Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Ibaraki, Japan; WEDI—European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom; YADAV—Climate Dynamics Program, and Department of Atmospheric, Oceanic, and Earth Sciences, George Mason University, Fairfax, Virginia

**CORRESPONDING AUTHOR:** Ayrton Zadra, [ayrton.zadra@canada.ca](mailto:ayrton.zadra@canada.ca)

DOI:10.1175/BAMS-D-17-0287.1

In final form 26 October 2017

©2018 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

and systematic bias should be evaluated in a state-dependent manner. Emphasis was given to evaluating forecast reliability within a data assimilation framework, which can help in tracing the sources of error in different flow regimes. In climate models, model drift analysis was suggested as a useful technique for understanding model error. Sensitivity experiments using nudging and a hierarchy of models of increasing complexity should also be employed.

Stochastic parameterization was discussed as a useful technique to represent uncertainty in numerical weather prediction (NWP) and seasonal forecasts, and to reduce errors in climate models. The dependency of large-scale errors on small-scale variability was discussed, questioning whether it is necessary to represent it explicitly or statistically. The stochastically perturbed parameterization tendencies (SPPT) scheme appears to reduce variability in certain cases, improving the signal-to-noise ratio of seasonal North Atlantic Oscillation forecasts and reducing the excessive El Niño–Southern Oscillation (ENSO) amplitude, although the detailed impact appears to be model dependent. The wide use of this scheme is related to its simplicity and intuitive appeal, but questions arose regarding the physical consistency of the perturbations. As a result, the community is moving toward targeting model uncertainty in a more process-oriented manner.

The metrics and diagnostics session highlighted the necessity of a continuum of metrics to evaluate models, including holistic metrics for users and decision-makers, and process-oriented metrics for model developers. The fractions skill score (FSS), in conjunction with neighborhood verification methods, allows users to make probabilistic predictions from deterministic forecasts. The common basis function approach was proposed as a means for measuring model skill in simulating extratropical variability. Techniques for isolating compensating biases were considered, allowing for the potential to weight future Coupled Model Intercomparison Project (CMIP) models according to their bias in means and trends.

Easy access to multiple, possibly gridded, observational datasets [e.g., Observations for Model Intercomparisons Project (obs4MIPs)] will enable modelers to use observations for model validations and provide estimates of observation uncertainty. Changes in computing infrastructure are developing toward less data storage and increased simulation repeats from restart files. This requires rethinking the types and frequency of data output and diagnosing some processes inline.

The need to share and apply community-wide diagnostics packages was stressed as a fundamental requirement for an objective validation and verification of models. Many such packages are already developed (e.g., spatial verification software, with object-based approaches) for NWP, and diagnostic packages for CMIP model evaluation.

The teleconnections session covered recent improvements in subseasonal predictions and the influence of the extratropics on tropical variability. The theory of extratropical interactions with the MJO in boreal winter was reviewed, followed by discussions on ways to improve our understanding of teleconnection mechanisms and extratropical predictions.

Links between the low, mid-, and high latitudes were addressed, with a focus on the warming slowdown and Arctic amplification. It was emphasized that errors in external forcings in phase 5 of CMIP (CMIP5) models may partially contribute to the overestimation of recent tropospheric warming on longer time scales.

On subseasonal time scales, the remote Northern Hemisphere response to the fast and slow MJO episodes was discussed. Understanding of the midlatitudes' response to cyclic and pulse heating was stressed. Modest improvements in teleconnection patterns linking the Indo-Pacific intraseasonal variability with the Northern Hemisphere are seen after introducing a higher-resolution ocean component.

The Year of Tropics Midlatitude Interactions and Teleconnections (YTMIT) project was highlighted. Its main objectives are to better understand subseasonal pathways of tropical–extratropical interactions, to identify periods and regions of increased predictability, and to improve subseasonal-to-seasonal forecasts.

## CONCLUSIONS AND WAYS FORWARD.

All model evaluation efforts reveal differences when compared to observations. These differences may reflect observational uncertainty, internal variability, or errors/biases in the representation of physical processes. The following list represents errors that were noted specifically during the meeting:

- convective precipitation—diurnal cycle (timing and intensity); the organization of convective systems; precipitation intensity and distribution; and the relationship with column-integrated water vapor, SST, and vertical velocity;
- cloud microphysics—errors linked to mixed-phase, supercooled liquid cloud, and warm rain;
- precipitation over orography—spatial distribution and intensity errors;

- MJO modeling—propagation, response to mean errors, and teleconnections;
- subtropical boundary layer clouds—still underrepresented and tending to be too bright in models; their variation with large-scale parameters remains uncertain; and their representation may have a coupled component/feedback;
- double intertropical convergence zone/biased ENSO—a complex combination of westward ENSO overextension, cloud–ocean interaction, and representation of tropical instability waves (TIW);
- tropical cyclones—high-resolution forecasts tend to produce cyclones that are too intense, although moderate improvements are seen from ocean coupling; wind–pressure relationship errors are systematic;
- surface drag—biases, variability, and predictability of large-scale dynamics are shown to be sensitive to surface drag; CMIP5 mean circulation errors are consistent with insufficient drag in models;
- systematic errors in the representation of heterogeneity of soil;
- stochastic physics—current schemes, while beneficial, do not necessarily/sufficiently capture all aspects of model uncertainty;
- outstanding errors in the modeling of surface fluxes; errors in the representation of the diurnal cycle of surface temperature;
- errors in variability and trends in historical external forcings;
- challenges in the prediction of midlatitude synoptic regimes and blocking;
- model errors in the representation of teleconnections through inadequate stratosphere–troposphere coupling; and
- model biases in mean state, diabatic heating, SST; errors in meridional wind response and tropospheric jet stream impact simulations of teleconnections.

As a way forward in model development and error reduction, a hierarchy of modeling techniques was recommended. Studies based on constrained systems (e.g., single-column models, nudging techniques, or prescribed surface conditions) and fully coupled systems (possibly with the gradual inclusion

of coupled processes) should be used to improve our understanding of poorly represented processes.

Evaluation and benchmarking tools using long time series with high temporal resolution should also be employed. Optimal uses of observations, LES, and theoretical studies were also recommended to address errors related to clouds and precipitation. New techniques are emerging for the automatic tuning of parameterizations against observations and explicit simulations. An enhanced interaction among model developers, verification groups, and forecasters/users is essential.

WGNE should embrace and foster developments toward kilometer-scale global weather and climate simulations, develop recommendations on best practices, promote the optimal use of future high-performance computing platforms, prioritize systematic errors among the outstanding list, and be a forum to categorize, discuss, and promote solutions for systematic errors in models.

The workshop recommendations include extending the Drag Project and considering momentum processes more generally, launching an intercomparison project focused on surface flux errors, considering a cross weather–climate group looking at initial tendency analyses of common biases, organizing the next workshop within four to five years, and inviting submissions, especially those focused on solutions for systematic errors.

### **EARLY CAREER SCIENTIST OPPORTUNITY AT WSE.**

For the first time, the workshop included a set of dedicated early career scientist (ECS) activities, led by Young Earth System Scientists Community ([www.yess-community.org](http://www.yess-community.org)) members, including the poster competition and best poster award, with 18 participants; ECS participation as rapporteurs in oral sessions; and meet-and-greet sessions with senior scientists. ECSs also acted as coreviewers of the poster competition, together with expert scientists.

**ACKNOWLEDGMENTS.** On behalf of WGNE we thank ECCC for hosting the workshop, and WMO, WCRP, NOAA, and OURANOS for the financial support. 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.