Multi-year Earth system variability, predictability and prediction

Topical Leads: *Benjamin Kirtman, Gerald Meehl, Christina Patricola*

Individual weather events are generally not predictable more than a couple of weeks in advance. This is because the atmosphere is chaotic, so that errors in the initial conditions grow over a few days into large-scale disturbances. However, the atmosphere can also be influenced by external factors. This is clearly illustrated by the annual cycle. For example, we expect more mid-latitude storms during winter than summer, although the precise time and location of each winter storm cannot be predicted more than a couple of weeks in advance.

Predicting the annual cycle is, of course, trivial. Furthermore, there are other natural factors that can cause particular seasons, or even years or decades, to be abnormal. The combination of natural variability along with forced changes will be felt most acutely through changes in the frequency and intensity of extreme events, including droughts, floods, storms, fires and heat waves. There is therefore an increasing need to know how likely climate events with large societal impact will be in the coming seasons to decades. This section of the white paper *very* briefly reviews the current scientific understanding and challenges, and some of the near-term and longer term research goals of multi-year Earth system variability, predictability and prediction.

Grand Challenge Question: *What is the interplay between internal variability (from extreme weather to low frequency modes) and external forcing (both anthropogenic and natural) that affects predictability and prediction of near-term regional climate in decadal climate predictions and what critical processes limit our ability to improve predictability and predictions of multi-year Earth system variability?*

Addressing this Grand Challenge question will require understanding limit of predictability, and implicitly mechanisms, of phenomena operating at interannual to multi-decadal time scales such as El Niño-Southern Oscillation (ENSO), Atlantic Multidecadal Variability (AMV), Pacific Decadal Variability (PDV). A current overview of these topics and references related to initialized prediction is given in Meehl et al. (2020).

Description of Challenges and Current Research in RGMA

a) Multi-Year ENSO

The largest source of seasonal forecast skill is ENSO, which is a coupled mode of variability in the tropical Pacific. ENSO grows through positive feedbacks between sea surface temperature (SST) and winds: a weakening of the easterly trade winds produces a positive SST anomaly in the eastern tropical Pacific which in turn alters the atmospheric zonal (Walker) circulation to further reduce the easterly winds. ENSO influences seasonal climate almost everywhere either by directly altering the tropical Walker circulation, or through Rossby wave trains that propagate to mid and high latitudes. The strongest impacts occur in Indonesia, North and South America, east and South Africa, India, Australia and East Asia. There is also a notable influence on the NAO, especially in late winter. ENSO also modulates the vertical wind shear and stability in the tropical Atlantic atmosphere, leading to fewer (more) hurricanes during El Niño (La Niña) years. ENSO's heat discharge can fuel intense hurricanes in the East North Pacific and these hurricanes may impact the south west US and Hawaii islands.

Observed El Niño and La Niña events are mostly not opposites of one another. For example, most El Niño events last a few seasons and frequently transition quickly into La Niña. From time to time El Niño events tend to reach extreme intensities. In contrast, one out of two observed La Niña events lasts 2 years or longer. Moreover, very few La Niña events transition directly into El Niño as expected from oscillatory behavior. Instead, the great majority of La Niña events slowly decay, oftentimes taking several years of near-neutral conditions until the next El Niño event is triggered. These observational findings suggest that there is multi-year predictability associated with La Nina events and frequent El Niño to La Niña transitions also implies potential predictability. Not all these observed characteristics of ENSO are simulated by climate models, and improving the simulation and thus prediction remains a challenge. For example, the diversity in ENSO's spatial pattern is largely missing in all climate models. The asymmetry between El Niño and La Niña amplitude and duration are not well captured by many models.

b*) Atlantic and Pacific Multi-Year Mid-Latitude Variability*

Atlantic Multidecadal Variability (AMV) is likely to be a major source of decadal predictability. Observations and climate models indicate that north Atlantic SSTs fluctuate with a period of about 30–80 yrs., linked to variations of the Atlantic Meridional Overturning Circulation (AMOC). Climate models suggest that the AMOC and AMV can vary naturally or through external influences including volcanoes and greenhouse gases. Idealized model experiments suggest that natural fluctuations of the AMOC and AMV are potentially predictable at least a few years ahead. If skillful AMV predictions can be achieved, observations and modelling studies suggest that important climate impacts, including rainfall over the African Sahel, India and Brazil, Atlantic hurricanes and summer climate (e.g., heat waves, droughts) over Europe and America, might also be predictable.

Pacific decadal variability (PDV) is also associated with potentially important climate impacts, including rainfall over America, Asia, Africa and Australia. The combination of PDV, AMV, and trends appears to explain nearly all of the multi-decadal US drought frequency, including key events like the American dust bowl of the 1930s.

c) Prediction and Predictability

In terms of the state-of-the-science in multi-year system prediction, Kirtman et al. (2013) provides a comprehensive review largely based on studies that analyze the experiments made as part of the Coupled Model Intercomparison Project phase 5 (CMIP5). The upshot of the Kirtman et al. (2013) review was there is considerable and more recently, mounting evidence for skillful predictions of temporally averaged temperatures up to ten years in advance particularly in the North Atlantic, Indian, and western subtropical Pacific Oceans. The skill in predicting terrestrial surface temperatures was considerably more modest. All of the models participating in the CMIP5 decadal prediction experiments used eddy-parameterized ocean components. More recently there have been numerous follow on multi-model studies and detailed analysis based on a specific models and experimental designs, and, in particular, ocean eddy-permitting models. A comprehensive review is beyond the scope of this white paper.

Finally, we note that initialization of the current state of the climate is essential for seasonal to decadal forecasts. Here we note an important distinction between initialization and assimilation. There are a number of data assimilation methodologies (e.g., Kalman filter techniques, optimal interpolation techniques, coupled vs. uncoupled) and these primarily focus on how to combine observational estimates with models whereas, initialization emphasizes how the results of these various data assimilation methodologies (or other approaches) are used to actually initialize predictions. In this white paper we focus on the initialization issue. There are a number of possible initialization strategies but in the broadest terms they fall into one of four categories, and all have various strengths and weaknesses:

1) Unified initialization. This can include separate assimilation in the various components of the climate system or fully coupled earth system models. The assimilation is viewed as unified in the sense that the assimilation system uses the component models from the prediction system for the data assimilation.
2) Disparate data assimilation and prediction systems. In this category, assimilation products from a specific data assimilation system involving a particular model is used to initialize predictions made with a completely different model.
3) Anomaly assimilation whereby ocean observations are introduced into the model by relaxing the fully coupled model towards gridded ocean analysis.
4) Ocean only simulation where the initial state comes from a forced ocean - sea-ice hindcast simulation driven by, for example, the Coordinated Ocean-Ice Reference Experiments (CORE) atmospheric data sets. One challenge is the typically-used atmospheric data suitable for forcing coarse-resolution ocean simulations has been demonstrated to introduce substantial errors in high-resolution ocean simulations.

Gaps in Current Research

Improving fundamental understanding should remain the main driver for research within RGMA. That said, given the current research in RGMA and the overarching research goals described above, there are some specific research gaps that could be addressed through leveraging existing expertise and capabilities within RGMA. These research gaps fall within three main categories: 1) gaps in our quantitative understanding of predictability sources of multi-year variability, 2) gaps in models' ability to simulate and predict the multi-year variability, and 3) gaps in methods for skill assessment.

1) *Sources of multi-year predictability*

- The time between El Niño events is typically about 2–7 years, but the mechanisms controlling the initiation of the warm phase and the reversal to the opposite cold phase are not understood completely.
- Mechanisms underlying AMV and to a greater extent for PDV have yet to explain what are the key underlying processes that control their prolonged persistence and their phase transitions.
- AMOC is believed to be driven to some extent by the NAO which does not have a long memory of its own from the underlying atmospheric dynamics alone. Improved predictions of the evolution of the AMOC and associated climate will therefore likely require low frequency variations of the NAO which to be predicted – the predictability in this regard has not been fully assessed.
- There is clear evidence that seasonal forecasts are more skillful when ENSO is active. Not only is ENSO itself more predictable once established, but climate in teleconnected regions is more strongly constrained, and therefore more predictable, when ENSO is active. Idealized experiments also suggest that the predictability of AMV depends on the initial state. Regime dependence of skill could therefore be exploited further to increase confidence in predictions under certain circumstances. These windows of opportunity during which very skillful predictions could be achieved could therefore be used to give forecasts with higher (but conditional) skill. This could arise, for example, if the effects of several different sources of skill align to produce a particularly strong signal. Finally, the importance of multiple constructive and compensating sources of variability in driving the statistics of extreme events highlights the need to understand future changes in joint probabilities of leading modes of variability.
- There has been substantial research on climate variability and on climate change, however, the interplay between variability and a changing background state has received less attention. Understanding the two jointly is important for two reasons. First, modes of variability shape the statistics of extreme events, and second, the response of extremes to patterns of variability depends on the background state, sometimes non-linearly and with

threshold behavior. Variations in background state can arise not only from mean climate change (e.g., changes in interhemispheric, interbasin, and intrabasin SST gradients), but also from changes in spatial patterns of variability and the seasonal cycle. For example, the importance of considering ENSO from such a perspective is highlighted by the non-linear relationship between SST and deep convection, together with the strong background SST gradients associated with the West Pacific warm pool and East Pacific cold Tongue. An additional important consideration is how variability in a changing background state can influence the interactions between thermodynamic and dynamic factors important for extremes.

2) *Model errors*

- Model biases remain one of the most serious limitations in the delivery of more reliable and skillful predictions. Biases in the mean tropical-subtropical SST, land-atmosphere coupling, and precipitation are among the most persistent throughout generations of climate models and problematic in terms of impacting simulated extreme events. Many of these deficiencies stem from a poor representation of sub-grid processes, particularly in the atmosphere - namely those associated with the representation of clouds, convection, and precipitation. The current practice of model mean state bias correction is unphysical and neglects entirely the non-linear relationship between the climate mean state and modes of weather and climate variability. In addition, the AMIP simulations often used to circumvent SST biases in coupled models, while informative, lack important coupled processes and feedbacks. Reducing model bias is arguably the most fundamental requirement going forward. A key activity must be the evaluation of model performance with a greater focus on processes and phenomena that are fundamental to reducing model bias and for delivering improved confidence in the predictions.

- In addition to mean state biases, there are other systematic errors related to feedback processes. For example, in some instances coupled models can simulate ENSO events with the correct amplitude for the wrong reason, owing to strong cancellation of coupled thermodynamic and dynamic feedbacks. What is the cause of errors in feedbacks and how these errors affect models' ability to simulate ENSO asymmetry, extreme events, and ENSO predictability are open questions.

- Likewise, the potential predictability in the climate system for monthly to decadal timescales is probably underestimated because of model shortcomings. For example, ENSO's spatial pattern diversity is largely missing in all climate models thus far. The asymmetry in amplitude and duration between El Niño and La Niña events are not well captured in many models. Improving the simulation and thus prediction of ENSO remains a challenge. There is therefore considerable scope for improved skill through model development aimed at reducing biases and improving the simulation of teleconnections. This will be achieved both by increased resolution as computers become more powerful, and improved parameterization of unresolved processes. Progress in model development may accelerate by studying the development of errors in seamless seasonal to decadal predictions. This is particularly relevant for ocean model errors (e.g. tropical SST errors) that develop with time scales longer than those in the atmosphere, for which it is possible to determine if erroneous atmospheric forcing given realistic SSTs is causing the SST errors.

- One key focus of model improvement efforts should be the representation of clouds, convection, precipitation, and radiation - processes which do not always improve over the range of resolutions in the global models used for multi-year predictions. While it is well-known that errors in these processes contribute to mean-state biases, there is increasing evidence that these errors in these processes also contribute to errors in the simulated variability. In particular, recent studies have demonstrated how low cloud processes contribute to errors in the amplitude of AMV, PDV and ENSO. While low clouds are particularly important for variability because their impact on solar radiation reaching the surface systematically varies with SST itself (i.e. the cloud feedback), the latent heat release in deeper clouds and convection drives atmospheric circulation anomalies that can affect surface wind stress and hence impact climate variability. Spatial gradients in the response of radiative fluxes to changes in composition have similar impacts on circulations and climate variability; errors in the representation of radiation and especially its sensitivity to composition changes (the "radiative forcing") may hamper predictability.
- Ocean and atmosphere model resolution may play an important role in simulating a stronger forcing of the atmosphere from the ocean than is achieved in most climate models, and indeed, there is some evidence that the atmosphere is more strongly coupled to the ocean in higher resolution models.

3) *Evaluation and skill assessment methods*

- The process of forecast calibration and skill assessment using hindcasts presents some serious challenges, however, when the lead time of the predictions extends beyond days to months, seasons and decades. That is because to have a high enough number of cases in the hindcast set means testing the system over many realizations, which can extend to many decades in the case of decadal prediction. Given the limited observational record, forecast calibration and skill assessment continues to require careful consideration.

Future Directions

Based on these gaps identified in our current knowledge, we formulate 3-5 year and 10 year research goals to address these gaps in the context of the Grand Challenge.

Short Term (3- 5 years) Research Goals

- Use the CMIP6 DCPP simulations to better quantify regional prediction skill on different timescales especially from the known predictability sources (MJO, ENSO, land surface, sea ice, stratosphere).

- Perform and analyze process experiments to determine interaction among ocean basins on decadal timescales, as well as to elucidate processes and mechanisms that could be producing decadal climate variability and define the role of the effects of volcanic eruptions in decadal climate prediction.

- Perform large ensemble hindcast simulations (30 ensemble members of 10 years each for each start date) with 1 degree class earth system models, and a moderately large ensemble (10 members of 10 years each for each start date) with ¼ degree class earth system models to determine the relative contributions of natural and anthropogenic forcing agents in the initialized hindcasts.

- Explore of the benefits and limitations of high resolution on synoptic weather events that produce extreme precipitation in a set of initialized hindcasts at 1 degree and ¼ degree and examine processes connected with synoptic systems that produce precipitation extremes ranging from daily to seasonal to interannual in the initialized hindcasts of the DCPP CMIP6 simulations.

- Study how earth system processes will be affected by short-lived climate forcers to better understand how storms could change in initialized hindcasts due to changes in forcing.

- Study predictability of extremes in a changing climate and establish a benchmark of present-day predictability of extremes to establish reliability bounds for initialized predictions.

- Understand how climate model biases influence the predictability of extreme events using coarse and fine resolution and atmosphere-only and coupled atmosphere-ocean simulations from HighResMIP.

- Quantify how decadal-timescale base state changes affect interannual phenomena such as MJO, ENSO, and monsoons.

- Use the CMIP6 RFMIP and CFMIP simulations to understand how the total response to composition changes is partitioned between variability in forcing and variability in sensitivity.

- Analyze CMIP6 simulations to identify the role of specific atmospheric and oceanic processes in Earth system variability. Advanced multi-model diagnostic techniques such as emergent constraints should be used. The goal is to provide information how these processes affect the character of model simulated variability. Example processes to be considered include clouds, convection, precipitation and fresh water forcing of the ocean, boundary layer mixing (in both the atmosphere and ocean), and oceanic upwelling.

- Use the latest in new observations (both satellite and ground-based from programs such as the DOE Atmospheric Radiation Measurement (ARM) program) to evaluate the model representation of processes (e.g., clouds, precipitation, surface wind stress, soil moisture, sea-surface height) that are critical to simulation of the mean state and multi-year variability of the Earth System. A key goal is the determination of which critical processes are not improving with increases in model resolution, in order that model developers may be made aware of where their focus should be placed.

- Utilize a hierarchy of models (e.g. aqua-planet, non-linear oscillators, models with specific dynamics or diabatic processes) to illustrate and understand the multi-scale nature of Earth-System variability. A key goal is to the determination of the impact of small-scale processes (e.g. boundary layer mixing, clouds) on Earth-system variability and its predictability as well as the feedback of variability back onto the small-scale processes.

Long Term (10 years) Research Goals

- Elucidate mechanisms that produce decadal climate variability in concert with external forcing, and apply these directly to decadal climate predictions using models from low resolution to high resolution.
- Improved understanding of processes and better quantification of impact that can provide a springboard for developing practically useful decadal prediction system beyond the upper ocean heat content.
- Formulate probabilistic climate information from initialized predictions, and quantify the reliability of such predictions, for regional and local space scales and from seasonal to decadal timescales. This will heavily rely on observations to help characterize model processes and mechanisms that produce regional and local decadal climate variability.
- Analyze high resolution global cloud-permitting model simulations that will be suitable for studies of near-term regional climate predictability without deficiencies of cumulus parameterization.
- Analyze earth system models of the ¼ degree class for a suite of hindcasts with more frequent start dates larger ensembles for each start date to compile better statistical distributions of storm characteristics involving midlatitude and tropical systems.
- Study initialized hindcasts and predictions of regional/local sea level rise and storm surge with ¼ degree class to ⅛ degree class earth system models with consequent improved representations of midlatitude and tropical cyclones in the models, and the input from ice sheet melt from the next generation coupled ice sheet models.
- Quantify the characteristics of near-term earth system predictability associated with MJO, ENSO and decadal variability phenomena in high resolution initialized hindcasts with changes in short-term climate forcers, GHGs, volcanoes, and solar variability.
- Provide strong constraints on the past temporal evolution of radiative forcing to which the Earth itself has been subject to sharpen interpretation of this historical record.