# The value and purpose of the next generation of climate models and ensembles

## Reto Knutti

*Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland,*

*reto.knutti@env.ethz.ch*

As our understanding improves, more observations become available, and computational capacity increases, climate models continue to increase in complexity to synthesize all that knowledge. The hope is that as more and more processes are considered at greater realism and higher resolution, the models will converge to reality. But are they really, how do we know, and indeed should they? What is the purpose of current global climate models? Are they built to understand processes, to quantify past changes, or to predict the future, and do all of those require the same models?

On the conceptual philosophical level, there is debate whether we are confirming the model itself, or its adequacy for the purpose of making a specific prediction. I argue what we should confirm is a well-understood relevant relationship between the model and the real world. A model is not uncertain, nor is its prediction; what is uncertain is the relevance of the model prediction for the target system. This relevance is inherently specific to the question that is being asked.

On the level of the single climate model, the above implies that the concept of a best model is ill defined without specifying the purpose and metric to judge performance or quality. Yet most institutions continue to build one or two models only, without specifying the purpose explicitly. In many cases, model development is driven by scientific curiosity, and not by a discussion of what matters for the many questions at hand. That question of what the model is relevant for, is formulated only after the model is built, or not at all. In the latter case, model results are used uncritically for potentially important decisions despite obvious limitations. Alternatively, some users dismiss models as unreliable altogether. Both positions are unhelpful, and progress is not only required in the model development but also in guiding users in the interpretation of where models are skillful and where not.

On the level of the model ensemble, climate projections are often summarized as multi model means, assuming that the average of models is better than a single model. Yet averaging models is problematic, because the models are not independent and share biases and code, and the models may not span the full uncertainty range. A model average can significantly underestimate the climate change signal if the changes are spatially heterogeneous. A seemingly obvious step is to select individual models based on how well they simulate the past and present climate. But metrics of model performance and model weighting is a thorny issue. The lack of verification of the actual climate projections means that we do not know, or cannot agree on which metrics are most relevant to identify a good model.

An overview of the performance and limitations of current climate models is given, with a focus on recent coupled model intercomparisons and major challenges in interpreting them. Model agreement with observations continues to improve, but I show that uncertainty in climate projections is difficult to quantify, and has not decreased significantly in the past few years, partly as a result of irreducible climate variability. Progress in model evaluation as well as statistical methods to interpret and combine model projections is urgently needed, in particular as more models of different quality and higher complexity, including perturbed physics ensembles and ensembles with structurally different models become available.