SPECIFICITIES OF CLIMATE MODELING RESEARCH AND THE CHALLENGES IN COMMUNICATING TO USERS

by Ramón de Elía

To better communicate with users, climate modelers need to identify the main barriers to a fruitful exchange and dedicate resources to overcome these obstacles.

After years of being immersed in climate modeling activities, many members of the community have become accustomed to the specificities of our field and less acutely conscious of aspects that may be perplexing to outsiders or newcomers. This situation becomes particularly concrete when interacting with scientists of different backgrounds and users—a frequent event in the era of Earth system models, climate change, and impact and adaptation studies.

Difficulties in communication between groups from different fields of research are certainly commonplace. What situates this community in a particular place of responsibility to overcome communicational barriers is that they are at the core of the climate change community, with a large part of the science being carried out downstream from model simulations.

This essay attempts to sketch some characteristics and practices proper to climate modeling that are both particularly thorny to convey and of relevance for most users. It does not aim to be a recipe to succeed in communication with users, but to simply pinpoint areas that need attention. As can be expected, among the topics that are difficult to transmit there are those that are still somewhat puzzling for the climate community and thus not unreservedly discussed.

In addition to identifying some fundamental aspects of climate modeling that have important communicational repercussions (“Scientific characteristics and practices underlying climate modeling” section), the “Bringing forward sensitive points to better communicate with users” section discusses some possible research avenues, tools, and choices that may help to strengthen our capacity to deliver intelligible and trustworthy climate information to a sophisticated user.

A note on the definition of the term “user”: I refer here to users that have a reasonable level of scientific
background, such as scientists working downstream of climate modeling activities in impact and adaptation studies. Some of these scientists are required to participate in decision-making processes and hence their relation to model-generated information transcends academic purposes. Communications with less sophisticated users or the general public are a very different kind of challenge, as discussed in Somerville and Hassol (2011).

**SCIENTIFIC CHARACTERISTICS AND PRACTICES UNDERLYING CLIMATE MODELING.** This section will be divided in two parts: “Issues related to model complexity and technical aspects” lists those issues emanating from the nature of the climate system or from technical aspects of climate models, and “Issues related to the practice of the science” discusses those issues that relate to the established practice of the science.

**Issues related to model complexity and technical aspects.** Some of the issues discussed below are a consequence of the nature of the climate system while others result simply from technical limitations, but most are a combination of both. There is some overlap between the categories listed below and their content is to some extent arbitrary, but the chosen presentation helps to make salient some fundamental issues. For a more detailed discussion of most of these topics, see Muller and von Storch (2004).

**Slow convergence toward a realistic representation of the climate.** Unlike the textbook cases of classical science that most of us learn at university (e.g., the celestial mechanics revolution in the seventeenth century), no revolutionary step toward a comprehensive description of the climate system is expected. Climate models are built step by step, slowly increasing their complexity toward a more realistic simulation of the climate. This is not only a consequence of the scientific and technical challenges involved: computer power evolves much more slowly than do our ambitions with respect to grid resolution or the integration of new processes and their feedbacks. Usually, actual progress between model version results is underwhelming, and this acts as a counterpoint to the expectations of users. In some cases these expectations are perhaps unwisely raised by members of the climate community; in others, they seem to reflect a classical-science preconception that the climate is a well-defined system just waiting for someone to solve it. Part of the obstacle to a satisfying numerical representation of the climate is the inescapable closure problem in equations that need to be semiempirically parameterized.

Once we accept that climate modeling is an evolutionary process, a central question remains: is an ever-improving simulation of the climate—including future projections—possible? Comparisons among different phases of the Coupled Model Intercomparison Project (CMIP) projections suggest that intermodel convergence in successive generations is not certain (see Knutti and Sedláček 2012, who discuss seven possible reasons behind this). These disappointing results are probably not going to weaken the resolve of the community to keep investing effort in this direction, but they cast an assumption of progress in a new light. As one might suppose, the mere suggestion of intermodel divergence is a bewildering perspective for those users who hold great expectations for model improvement in the coming years.

**Lack of climate model dominance.** In the language of decision theory, model X is “dominant” over model Y when results from the first are better than those of the second for all possible skill scores. Our experience shows that no single climate model is in this sense dominant over the others. There are, however, models that outperform others in the most important criteria (see Gleckler et al. 2008). One of the consequences of this lack of dominance is that, unlike the examples from the history of science in which a new theory supersedes old ones, climate models of different origins and generations may coexist. In general, new models tend to moderately outscore those of the previous generation, improving several aspects and perhaps deteriorating some others. Lack of model dominance—a common trait in many fields the further one gets from physico–mathematical sciences—is a strongly dissonant characteristic of climate modeling for users trained in the “hard sciences.”

**Model skill not necessarily related to good causality.** It is well known that in many situations climate models obtain the right result for the wrong reason. For example, a model may properly simulate the daily surface temperature over a given region, but as a consequence of compensating errors such as those linked to excessive cloudiness that produces cooler days and warmer nights than observed (and hence only a small error in the daily average).

From an empirical or pragmatist point of view, where the primary concern is obtaining good results in a specific aspect, the question of causality may appear to be immaterial. It is of much interest,
however, if we want the model to mimic nature as well as possible (see discussion in Held 2005 and Winsberg 2003). Moreover, in the context of climate change, where good results for today's climate may not guarantee good results for future climate, the pragmatist approach may be inappropriate. Proper causality has received more attention in recent years. With respect to results analysis, there is an increased concentration on model ability to reproduce complex spatiotemporal patterns as well as diurnal and seasonal cycles; with respect to model development, the seamless prediction paradigm is helping to ensure that time averages of the longer time scales are not improved at the expense of shorter time scales. These efforts, however, are not widespread: for example, those requesting future climate local information, where fewer studies regarding proper causality are available, are normally offered conventional validation procedures (i.e., evaluations of averages of a couple of variables), which seems a weak platform on which to build confidence.

**Emergent behavior from the interaction of several components.** The climate system is formed by several components that evolve at vastly different time and spatial scales (i.e., the atmosphere, the hydrosphere, the cryosphere, the lithosphere, and the biosphere). Numerical models represent the climate as a highly hybrid nonlinear dynamical system describing local interactions at successive time steps—at the scale of minutes and of tens of kilometers—but whose emergent behavior may evolve in decadal and continental scales.

There are many sciences involved in climate studies where the notion of "model" differs dramatically, however, from that just expressed (see Muller and von Storch 2004). For example, some models are mainly sophisticated statistical tools that can be adjusted or calibrated to perform a given task. From the climate that emerges from running a climate model, we cannot, for example, deliberately remove a temperature bias by modifying a simple parameter (a common impression among many users without experience with comprehensive climate models). Bias removal is more like a trial and error activity that consists in tuning parameters by means of educated guesses or expensive objective methods—a laborious exercise not free of controversy (see Randall and Wielicki 1997)—after which we frequently fail to remove enough bias, or we succeed at the cost of deteriorating something else. An unyielding model bias may have severe consequences on impact models (e.g., those concerned with the annual cycle of river discharge)—a fact that some users are well aware of and vocal about. Despite users’ useful feedback, not much can be done to correct the situation within a reasonable response time.

**Predictability and natural variability issues.** Even under the hypotheses of a perfect model, the climate system has limited predictability (in the sense that a small error in the initial conditions grows with time, eventually decorrelating predictions among themselves). Lack of predictability does not preclude the ability of anticipating the overall statistical behavior of the system, but it implies a shift from deterministic causes to probabilistic causes—even though we work with deterministic equations. In practice, this means the requirement of sampling the space of solutions with multiple simulations and acknowledging the limited amount of information that can be derived from single realizations (and the observed climate is a single realization). One of the consequences of this is that studies have to be planned within a statistical framework, which raises computing costs enormously, complicates attribution of causes and validation, and dilutes the value of simulated information. Users have to assimilate not only that limited predictability hinders climate prediction for other than long-term statistics (natural variability tends to dwarf short-term trends) but also that it precludes synchronicity between observed past climate and model-simulated past climate (usually a puzzling and unpleasant surprise).

**Very limited verification opportunities.** Observational datasets span a short period in the climatic history of our planet and are sparse and inhomogeneous in space, limiting validation opportunities for climate models. In addition, validation is affected by natural variability, as discussed in the previous subsection. It is also worth mentioning that some observations are used to tune climate models and, consequently, issues of independence with respect to validation attempts may be raised. An additional fundamental obstacle appears in climate model projections for the coming decades, where no data are available for verification (although some data can be used to evaluate the present impact of greenhouse gases). This makes it impossible to estimate the expected model error in future climate projections, and hence we have to rely on partially subjective and ad hoc estimations of uncertainty that raise numerous questions (unlike the situation of weather forecasting modelers, for whom the event of interest such as a two-day forecast is both recurrent and reasonably
A lot of effort has been invested by the climate modeling community to establish ranges of confidence for model future projections through the use of ensembles and postprocessing, but interpreting these outputs is far from trivial. Even the concept of probability—which is comfortably handled in weather forecasting—becomes controversial in this context, and a deeper understanding of the notion of uncertainty is usually needed for a fruitful dialogue with users (see Curry 2011).

**Limited by human error.** Climate models are extremely sophisticated computer programs that may exceed a million coding lines. Although the existent literature suggests that climate models are generally well coded (in the sense of following the recommendations of specialists regarding quality control and code structure; see Easterbrook and Johns 2009), models are, it seems, inexorably burdened by coding bugs. It is not known to what extent model biases or errors can be attributed to coding bugs, but it is perceived that their number and significance may strongly vary among institutions. There is, within a good part of the community, an ambivalent, almost Victorian, relation to their existence: model versions are dismissed on account of known bugs (disregarding many unknown ones) and considered immaculate after detected bugs are corrected. Human error in program coding has technical and scientific consequences, but it especially raises issues of trustworthiness that verge on the ethical.

In the case of a modeling group providing data to users, the moral imperative seems to be to disclose the news when a bug is found, and in certain situations redeliver a corrected dataset. Naturally, from the user standpoint, integrity in disclosing the discovery of bugs is not a replacement for high modeling standards and, in the long-term, modelers’ credibility may suffer.

**Issues related to the practice of the science.** In addition to issues related to climate system itself and the technical demands of modeling, there are social, political, and cultural aspects of climate modeling practice that have an impact on users.

**A science stretched and sheared.** Climate change studies have become a textbook example of what is sometimes called “post-normal science” (e.g., von Storch 2009). This refers to the fact that climate research activities are of great practical significance for policy and decision making by governments and private actors, while the scientific results involve uncertainties difficult or impossible to reduce in the short term. To this situation we can add the growing importance of public opinion and pressure groups, although it is unclear how much—and by which means—these are influencing the science. Feedback among these various stakeholders can potentially obscure the rationale behind certain decisions and provide, as Bray and Krück (2001) put it, “the opportunity for poor science to inform policy and for misconstrued politics to feed back into science.”

To be at the eye of the hurricane has been both a blessing and a curse for climate research. Undoubtedly more funds have been funneled into our field than would have been otherwise, but the blessing comes with many challenges, particularly related to public perception. The challenges are not limited to the complex dialectics of communicating the science with a view, for example, to eventually influence emission policy decisions: climate modelers are in some cases urged to put their tools—where a large part of the funding is going—into effective action by creating practical products.

The community has been asked to stretch research tools into applications and sometimes to tacitly or explicitly offer products with striking names and exciting scientific perspectives that thus far are of questionable practical use. It is only recently that there have been calls demanding that more attention be given to making information providers more responsible, accountable, and credible; and in particular “to temper any undue expectations for the type and time frame for delivery of the needed information” (Asrar et al. 2013).

**A science steered through coordinated experiments.** The complexity and costs associated to issues like those discussed in “Issues related to model complexity and technical aspects”—typical problems of what is sometimes referred to as “big science”—are at the heart of a need for coordinated initiatives such as the CMIP experiments. This approach leads to standardized protocols in climate model practice and makes available model results to a vast community, facilitating wider systematic study and documentation. From the viewpoint of those delivering information to users, such a centralized and organized database has played a key role in raising the standards of the service. Among other things, the exploration of model uncertainty made possibly by coordinate experiments has created a large base of scientific publications very useful for engaging users in discussion.

Historically, the CMIP experiments have had an eye more on researchers than on users. For example,
the public presentation of CMIP3 (Meehl et al. 2007) does not once mention the word “user.” For CMIP5 (Taylor et al. 2012), users seem to have played a larger role in defining output variables. The concentration on research remains clear, however, when in the conclusions they state that “CMIP5 will ultimately be judged on the research it enables. If scientists can successfully [...] use it to address fundamental scientific questions concerning climate and climate change, [...] then CMIP5 should be considered a success.”

An absence of focus on users can also be seen in the fact that CMIP experiments paid limited attention to natural variability. Deser et al. (2012) did an excellent job of illustrating natural variability at certain localities to the point of creating a useful tool for discussion with users. But this study relied on results from forty 60-yr runs, which meant undertaking a large experiment as no such database had been planned within the CMIP umbrella. Had this been the case, the excellent communicational effort made by Deser et al. (2012) could probably have appeared much earlier, avoiding unnecessary misunderstandings from users regarding what we could expect locally in the next decade or so.

Another example is the design of the CMIP5 protocol with respect to modification in the way greenhouse gas (GHG) emissions and concentrations are treated [from the Special Report on Emissions Scenarios (SRES) to representative concentration pathways (RCPs); see Moss et al. 2010]. There are many good reasons for these modifications to occur, but one consequence is the loss of continuity between CMIP3 and CMIP5 for many users. The reconciliation of these two datasets is not simple (see Markovic et al. 2013) and may be de facto forcing users to disregard some past information and to restate their working hypothesis.

A science encumbered by uncoordinated initiatives. The CMIP experiments described above are coordinated in the sense that a set of sophisticated systematic experiments is performed by a large number of global models from different institutions. But the existence and structure of the models themselves is beyond the realm of the CMIP planners. The fact that dozens of climate models exist is due more to the nature and dynamics of individual institutions and governments than to any well-argued scientific reasoning. In addition, models are developed within institutions with varying scientific traditions and resource commitments that influence their quality. The large amount of models available creates on the one hand the possibility of a rich ensemble for analysis and on the other a jungle of data that generally exceed users’ processing capabilities.

Simultaneously with the growing number of global and regional climate models, model complexity is increasingly forcing an intense interchange of expertise to optimize resources, making the effective number of independent models approximately a third of the total for CMIP3 (e.g., Pennell and Reichler 2011). This is not simply an academic matter, as consensus of nonindependent models cannot be used as evidence to build confidence in results with decision-making potential.

Lack of coordination can also be seen between the global and the regional modeling communities. The latter has emerged as the most important player in the transmission of climate information to users [in the first publication of the North American Regional Climate Change Assessment Program (NARCCAP; Mearns et al. 2013) on future projections, the term “impact community” appears in the first sentence of the introduction]; the global modeling community has become mainly responsible for fundamental research and for mitigation policy matters. This split may look like a reasonable division of labor, but in fact it breaks apart an important two-way chain of understanding in the complex endeavor of extracting useful information from climate models.

**BRINGING FORWARD SENSITIVE POINTS TO BETTER COMMUNICATE WITH USERS.** The points outlined thus far offer an indication of the complexities we need to communicate to users and some of the conditions under which this communication takes place.

Within this list are a few issues about which the climate modeling community has undertaken extensive research and has made strong communicational efforts, such as on the topic of model uncertainty and, more lately, natural variability. There are also topics about which the community has been less active.

Regarding these last, I believe we should pay more explicit attention to them in our research activities, primarily because it is difficult to communicate issues that are scarcely discussed in scientific publications. The following are examples of research lines that could be pursued with this aim.

The “Slow convergence toward a realistic representation of the climate,” “Lack of climate model dominance,” and “Emergent behavior from the interaction of several components” sections touch on the evolution of model development. Climate modelers have to deal with relatively slow progress in many aspects,
and for this reason the recent history of the field is very much present. Each Intergovernmental Panel on Climate Change (IPCC) assessment makes this fact clear when a comparison with results from previous reports is carried out. But one may wonder whether a comparison between results from different model generations is the only thing we can do: we have the tools to recreate model development—to run models of evolving complexity at a low cost, thanks to today’s powerful computers—and explore specific questions about the evolution of model skill and model convergence. For example, what is and has been the role of each model component—and at different levels of resolution and complexity—in improving skill and favoring/disfavoring convergence? Lessons from such experiences could help us build more realistic, better-documented expectations about future model developments. A project of this kind would provide a good opportunity to reflect upon and document the way the science has evolved and establish climate modeling credentials more firmly. From a similar perspective, but within the realm of metadata, the new project Common Metadata for Climate Modeling Digital Repositories (METAFOR; Guilyardi et al. 2013) has embraced the historical dimension of the climate modeling undertaking, suggesting that “the important additional benefit of archiving such information is that this ensures that the conditions under which model simulations were performed will be understood well into the future.” A similar concern seems to be valid for the entire enterprise of model development, especially when users’ decision-making accountability is at stake.

The “Model skill not necessarily related to good causality” section suggests that in addition to skill, good causality is fundamental to achieve confidence in model results, and that in the last years efforts have been made to better ensure “good results for the right reasons.” These efforts have concentrated in general on model capabilities to reproduce a limited number of climate phenomena, such as seasonal cycles or blockings. This compromise with good causality has not been, however, extended to ordinary use of simulated data by users.

Those concerned with local climate future impacts are rarely encouraged to enquire into the reasons behind a given model or models having a good local score in the present climate (usually an ad hoc credential of future skill). A line of research could aim to develop criteria and indices that can be used to provide information to users of local, limited datasets (e.g., a couple of variables at a given grid point) related to the “causality health” of the data. An analysis of some other variables at different time and space scales, as well as energy and water budgets, are probably needed to capture the main local climatic processes. Clearly “good causality” will not be assured, but very bad causality could be detected and the guilty datasets put aside.

The technical issue of model bugs developed in the “Limited by human error” section opens questions of model credibility from the user’s standpoint—of a given model and of the community at large. These questions could be better handled if we invest in learning more about model bugs: their life cycle, the expected errors they produce, the expected number of unknown bugs per line of code, the relative importance of model bugs on present biases, the difference in performance among institutions, etc. These are very difficult questions to answer, but the open search for tentative answers will probably give us, paradoxically, more means to strengthen the credibility of climate models. As is the case for model evolution, this issue also has a historical dimension: a systematic, transferable bookkeeping of bug-related issues is a necessity in the search for answers. Published papers tackling these questions could greatly facilitate communication with users.

The conditions discussed in the “A science stretched and sheared” section, regarding different forces that shape climate research and service activities, make it unwise to consider coding bugs as simply an annoyance. The impression of many that coding bugs are not a significant issue in the overall exercise of climate modeling needs to be objectively shown.

All three topics discussed above raise, with greater or lesser severity, issues of trust. Given the exposure to public and interest groups, dealing with information at the heart of decision making without abundant supporting material or guidelines may leave climate service providers in a vulnerable position. Studies that help to clarify these issues will probably be welcome by many. It is interesting to note that the word “actionable,” now in vogue when used with “climate science” to refer to usefulness in decision making, has also another meaning that we should not forget: affording grounds for legal action.

As discussed in the “A science steered through coordinated experiments” section, aspects of the superstructure of the climate modeling community still respond to a purely scientific approach or to mitigation policy issues, and this indirectly affects users interested in impact and adaptation issues (see Asrar et al. 2013). Perhaps a parallel, more modest CMIP-like user-centered approach could be a response, as is already done in regional climate
modeling (e.g., NARCCAP). In many disciplines the separation between research and user-oriented data is strict, and the situation in climate modeling deserves attention. Bottom-up feedback from users, the regional modeling community, and information providers to the global modeling community could be beneficial for the creation of a united climate modeling community including users’ perspectives.

The "A science steered through coordinated experiments" and "A science encumbered by uncoordinated initiatives" sections also touch upon the indirect consequences of model complexity on users. A major problem discussed in Muller and von Storch (2004) is that no single scientist understands the whole climate system, and many understand just a small fraction. One may wonder, then, whether the expertise of those dealing with users is the most appropriate to this activity, and whether capacity building of an equivalent to “family doctors” in medicine should be encouraged. For the moment, it is up to each institution dealing with users to discover the kind of training profile that a desirable team should have. Guidelines on recommended scientific background, even if preliminary, could be of use.

**CONCLUSIONS.** The aim of this essay was twofold. First, to list a core set of characteristics of climate modeling and its associated practices that cause or increase difficulties of communication between practitioners of the field and other scientists and sophisticated users of climate model data. Second, to discuss the need to invest in research and organizational activities that can help illuminate particularly difficult or delicate issues. In a community heavily invested on mutual trust, not only users but also the entire climate modeling community may benefit from a larger investment in detailed scrutiny of controversial points.

The benefits of transforming a scientific experiment such as climate change projections into usable information are self-evident, but the associated risks are much less clear. Some of us have gradually shifted from being climate scientists to becoming climate service providers—a change that has repercussions for the entire community. A few of these repercussions have been discussed here, and hopefully these and others will increasingly be discussed as our field continuous to evolve.

**ACKNOWLEDGMENTS.** I would like to thank the editor Bjorn Stevens and the four reviewers—among them H. von Storch—for their encouragement and fruitful criticism. R. Laprise and J. Tomm have also provided me with valuable feedback. Finally, I would like to thank my colleagues at the Ouranos Consortium and my former students for their stimulating discussions.

**REFERENCES**


