G lobal models are used in many applications, ranging from Numerical Weather Prediction (NWP) to seasonal prediction and climate simulation. Despite great computational advances, calculations using these models can still only afford grid spacing for which many of the important processes occurring in the atmosphere—most notably boundary layer processes as well as moist convection and cloud processes—remain unresolved. It is therefore necessary to represent those subgrid-scale processes as a function of the grid-scale variables. The technique to achieve this representation is usually referred to as parameterization (Stensrud 2007).

Despite impressive progress in the development of parameterizations over the last few decades, solutions to some long-standing model problems remain elusive. Examples for such problems are the erroneous representation of the ITCZ and the associated cold sea surface temperature (SST) bias in the central tropical Pacific (e.g., Zhang et al. 2007), the poor representation of the Madden–Julian oscillation and other modes of tropical variability (e.g., Lin et al. 2006), the incorrect representation of the frequency of occurrence of high- and low-intensity rainfall events (e.g., Sun et al. 2006), the poor representation of the diurnal cycle of rainfall (e.g., Yang and Slingo 2001; Betts and Jakob 2002), as well as ongoing difficulties in simulating the El Niño–Southern Oscillation (ENSO) phenomenon (e.g., Neale et al. 2008). It is noteworthy that many of these shortcomings are ascribed to the poor representation of subgrid-scale processes, in
particular that of moist convection. The well-known successes in NWP (e.g., Simmons and Hollingsworth 2002) and the simulation of global mean climate (e.g., Gleckler et al. 2008; Reichler and Kim 2008) demonstrate that these long-standing model errors often do not strongly affect our overall prediction capabilities on large scales. However, they do pose significant problems in many other areas, such as quantitative precipitation forecasting as well as our ability to simulate regional patterns of precipitation in current and future climates, particularly in the tropics (e.g., Sun et al. 2006).

It is worthwhile noting again that significant progress on a number of issues in parameterization development has been made. Particular examples of recent success are the treatment of shallow cumulus convection (e.g., Lock et al. 2000; Neggers 2009) as well as a much-improved treatment of clouds in radiative transfer calculations (Pincus et al. 2003) in global atmospheric models. Although one should acknowledge the possibility that certain processes might not be “parametrizable,” these examples highlight that success in parameterization development is possible and that the community is likely a long way from the “end of the road” on parameterization.

It is evident from the previous discussion that the improvement of models in general is strongly intertwined with the task of improving the parameterizations embedded in them. Why then has the considerable progress in the development of parameterizations not fully translated into alleviating some of the long-standing model problems? It is the aim of this paper to analyze the model development process with the goal of identifying potential reasons for this disconnect and to suggest possible pathways to solutions.

**A PARADIGM SHIFT IN PARAMETERIZATION DEVELOPMENT.** The representation of subgrid-scale processes in models by means of parameterization is almost as old as numerical models of the atmosphere themselves. Although often introduced in early primitive equation models as a “fix”—for instance, to avoid supersaturation when moisture was added as a model variable—modern parameterizations constitute complex conceptual models of the physical processes they are aiming to represent. An example of such a conceptual model is the mass-flux approximation with which convection is represented in many models today (e.g., Arakawa 2004; Tiedtke 1989). The complexity of these conceptual models tends to vary across models and applications. Nevertheless, the overall trend for representing physical processes on subgrid scales in that way has led to a paradigm shift in the development of parameterizations. While early parameterization development used simple relationships and was aimed at finding suitable values for the parameters in those relationships, modern parameterization development is largely concerned with detailed studies of the actual physical processes, either through observations or detailed process model studies, with the aim to build and refine suitable conceptual models. It is not surprising to see this development reflected in international research programs, such as the Global Energy and Water Cycle Experiment’s (GEWEX) modeling activities on clouds, boundary layers, and land surface processes (Randall et al. 2003b; www.gewex.org).

Although the use of conceptual physical models in parameterization has no doubt improved their quality and reputation, the increased complexity leads to a number of difficult issues as well. Complex models—conceptual or otherwise—can easily become an end in themselves, and the connection of the development to the application of the parameterization in a model can easily be lost. Once a parameterization achieves a certain physical realism, improving it further based entirely on findings from process studies (referred to here as both observational and model studies) becomes an appealing approach, one that in the long term will undoubtedly lead to improved models. However, nearly everyone involved in the development of parameterizations will probably attest to the fact that great conceptual improvements based on “first principles” often do not lead to overall model improvements straightaway. Worse still, they will often not address errors found in the application of the model in which the parameterization is embedded.

The opportunities for developing better parameterizations following the conceptual model approach have never been better than today. There is an unprecedented availability of observations based both on ground- and space-based instrumentation. Extensive networks of ground sites equipped with a large suite of remote sensors, such as those operated by the U.S. Department of Energy’s Atmospheric Radiation Measurement Program (ARM; Ackerman and Stokes 2003) or the European Cloudnet project (Illingworth et al. 2007), provide considerable insight into the processes relevant to parameterizations. Combined with the wealth of global information from satellite programs, such as the A-Train (Stephens et al. 2002) or the Tropical Rainfall Measuring Mission (TRMM; Simpson et al. 1988), these observations provide a rich underpinning for model development.
In addition to the wide variety of observations, recent advances in process modeling have also improved the prospects for parameterization development tremendously. Process models have been used in parameterization development for some time (e.g., Browning et al. 1993; Randall et al. 2003b), but advances in both science and computing have allowed for a more effective and reliable use of such models. For example, the development of process models for clouds and convection has recently culminated in the models being used as parameterizations (e.g., Grabowski 2001; Randall et al. 2003a; Khairoutdinov et al. 2005). In addition, the development in computing power has allowed the execution of global models in research mode at grid scales that approach those at which important physical processes, such as deep moist convection, begin to be resolved (e.g., Iga et al. 2007). Such model simulations, although a long way from operational application, do provide great opportunities to illuminate key requirements for parameterizations to successfully eliminate some of the model problems mentioned earlier.

An obvious question arising from the recent trend of the successful application of very high-resolution (~1 km) or “cloud permitting” models in regional and global weather and climate prediction is whether parameterization development will soon be a thing of the past and hence much of the discussion here is of little use. There is no doubt that increased model resolution resolves some significant parameterization issues. Good examples are topographic effects, such as those induced by coastlines and orography. However, parameterization development will remain a key need for model development in the foreseeable future for at least two major reasons. First, even high-resolution models contain parameterizations, such as those of turbulence and microphysical processes. Second, and perhaps more importantly, although the routine application of cloud-permitting models in regional weather prediction is feasible now, global cloud-permitting models for routine use in seasonal and climate prediction as well as in ensemble prediction systems used in NWP are still several decades away. Note that this does not invalidate their great potential as research tools, as recently highlighted by the World Modelling Summit (Shukla et al. 2009).

AN ANALYSIS OF THE MODEL DEVELOPMENT PROCESS. Model development is such a complex process that, for it to be successful, it needs to be informed by a number of considerations that range from knowledge about the overall model performance in its application (e.g., climate simulation) to knowledge about individual and often finescale physical processes (e.g., the nucleation of ice crystals). One of the greatest challenges for model developers today is that both the list of overall model problems ascribed to parameterization shortcomings and the list of shortcomings identified in the parameterizations themselves are very long. If the lists can be prioritized at all (to date they often are not), then mapping them onto each other is very difficult and usually not even attempted. Add to this the relatively small size of the parameterization development community and the often-quoted lack of progress in model development seems understandable. It will be argued here that although techniques and methodologies as well as collaborative research programs for the overall model assessment and the process study approach individually are well established, progress is severely limited by the weak connections between the two. It will further be argued that acceleration in model development can only be achieved by significantly strengthening these weak links through additional research and better coordination across existing programs.

A natural first step in developing the previously mentioned argument is to carry out an analysis of the model development process and to derive recommendations based on the findings of this analysis.

Figure 1 shows a schematic of the model development process. For the sake of illustration, it is assumed that the model in question is an atmospheric general circulation model (AGCM) for application in NWP or climate simulations, but arguments similar to those made here hold for other models and applications. The rectangles depict individual steps in model development, whereas the circles indicate communities involved in each step. It is argued here that model development requires the coordinated efforts of three communities: the data community, the model user/evaluation community, and the model development community. For the purpose of this discussion, those communities are defined as follows:

- **Data community**: Engaged in collecting and analyzing data for enhanced understanding of the system from process to planetary scales; uses models as guidance for observation strategies and dataset generation.
- **Model user/evaluation community**: Engaged in applying models to answer specific questions and to enhance understanding of the system; evaluates models with the aim to understand their suitability and the limits of their applicability.
- **Model development community**: Engaged in developing improved models.
At the core of the iterative development process is the model and its application. The top part of the schematic in Fig. 1 highlights the model application. This application is largely carried out by the model user/evaluation community, but in practice it is very often closely supported by the model development community. Most user communities employ a fairly well-defined approach and standard tools to carry out a general assessment of the model performance. In practice this step is often performed by the model user/evaluation community in conjunction with the data community, as comparisons to observations are at the core of this activity. At the end of this step, overall model weaknesses as they relate to the application are generally well known and some indication of their relation to the model formulation often exist too. Contemporary examples include the very large cold SST biases in coupled models in the central equatorial Pacific, poor monsoon simulations, or the weakness or absence of the Madden–Julian oscillation in atmospheric or coupled models. Here the representation of convection is often implicated as a process that requires improvement in models without being able to say what in current model formulations it is that may need improving, let alone being able to prioritize research. Numerous examples for this overall model assessment process in action can be found in operational NWP centers as well as the recent analysis of the climate simulations in the Coupled Model Intercomparison Project (CMIP) that form the basis of the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report.

The bottom part of Fig. 1 depicts the detailed assessment of the model formulation. This step relies heavily on detailed process studies, which lend themselves naturally to studying and improving the evermore complex and process-oriented modern parameterizations employed in AGCMs. Often the particular process study applied is chosen partly by the need of the model development community (e.g., the need for a “shallow convection case”) and partly by the availability of data. For instance, in the area of convection parameterization development, recent and/or comprehensive field experiments, such as the Tropical Ocean and Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE; Webster and Lukas 1992), are often used to derive process studies. Those are then carried out by all three communities, sometimes jointly but often in isolation. Over the last decade or so, international research programs, such as the GEWEX Cloud System Study (GCSS; Browning et al. 1993; Randall et al. 2003b), as well as national activities, such as the Climate Process Teams in the United States, have promoted this approach and have successfully fostered the collaboration between all three communities in the area of representing clouds in AGCMs. The parameterization community worldwide now relies heavily on taking this approach to development. Techniques and interactions between communities are well established, and just as for general model assessment (as defined earlier), one can argue that the community has the tools to and much knowledge about how to carry out the research required.

The use of process studies for model development in general, and parameterization development in particular, is essential because it ensures that increased knowledge on processes can be incorporated into models in a physically realistic way. At the same time, general model assessment has been highlighted as essential to understand the suitability, applicability, and needs for improvement in the model. A major problem in making model/parameterization develop-
ment more effective lies in the weak links between those two parts of the development process. A too-
strong emphasis on process studies frequently leads to parameterization developments that although are
of scientific interest and are often fully justified at the process level (e.g., “this ice nucleation formula-
tion is too crude”), they do not address major model shortcomings. An example of this very common
phenomenon is the work by Jakob and Klein (2000), who developed a parameterization separating cloudy-
and clear-sky fluxes of precipitation. Although they clearly showed the need for such a parameterization
based on a process study (Jakob and Klein 1999), the implementation of the new parameterization
into the full GCM had very little effect on the model results. On the other hand, ignoring process stud-
ies and aiming at broadly “matching” observations leads to heavy model tuning without improving its
physical basis. Those two situations are indicated as “shortcuts” in the model development loop in Fig. 1.

From the earlier-mentioned argument, it is evident that it must be a critical aim for model development
to balance and strongly connect both strands of model assessment. As indicated in Fig. 1, general
model assessment and process studies intersect in two places: when the process studies are selected and when
model improvements are developed and tested.

CHALLENGES. The selection of process studies

to be used in model development is currently only
loosely connected to the actual model errors. A
number of reasons can be cited for this disconnect.
One reason often given is the limited availability of
data for process studies. Although certainly true a
decade ago, this argument no longer applies. With
the easy availability of evermore complex satellite
data (e.g., Stephens et al. 2002), the availability of
long-term monitoring sites (e.g., Ackerman and
Stokes 2003; Illingworth et al. 2007) and the large
number of field studies performed in recent years, a
large number of process studies can be performed.
Furthermore, improvements in data assimilation

techniques and the growing availability of NWP
techniques to the climate community make it feasible
to perform process studies at every point of the globe
easily by performing short-range weather predictions
(e.g., Boyle et al. 2005). Hence, the more likely reasons
for the disconnect must lie elsewhere.

It is argued here that an improved connection of
overall model assessment to the process studies
necessary to model improvement requires additional
research into diagnostic techniques that address two
specific but connected aims: i) to identify which of
the various model shortcomings needs addressing
with what priority and ii) to identify which aspects
of the model are most likely responsible for the
highest-priority model shortcomings. The level of
difficulty of achieving the first of the two aims is
naturally dependent on the application. For example,
if a model’s main goal is to make fog forecasts two
days from now in a particular region and it consis-
tently fails to produce fog, then this will naturally be
a high-priority area for future development. On the
other end of the scale, prioritizing the list of model
problems in a coupled climate model with the aim to
produce more reliable projections of climate change is
an exceedingly difficult and yet unresolved problem.
Even if the decision on “the most important problem”
can be made, the task of linking this problem to a
particular aspect of the model formulation is often
difficult. Even in the just-given example of the failed
fog predictions, there are numerous possibilities what
might cause the problem, ranging from a misrepre-
sentation of the large-scale conditions to problems
in the model’s microphysics parameterization.
Some promising progress in diagnostic techniques
that address both aspects of this discussion have re-
cently been made. For instance, Bony and Dufresne
(2005) have recently identified differences in the
representation of low clouds over the subtropical
oceans as a possible major source of uncertainty in
climate model sensitivity. Also, early applications of
a regime-oriented approach to model evaluation (e.g.,
Williams and Tselioudis 2007; Hume and Jakob 2007)
have shown much promise. However, more effort in
this area of research is required to establish a solid
connection of process-oriented model development
to overall model errors.

The second connection from process-oriented
development to application is in the implementation
and testing of new ideas in the parameterizations
themselves. This process is currently severely slowed
by the extremely small size of the model development
community worldwide. This naturally means that
model problems are discovered at a much faster rate
than they can be solved. Worse still, implementing
and testing a possible solution to a model problem
often takes as long as—or longer than—finding the
solution itself—sometimes several years. This is be-
cause more often than not a new and physically more
realistic parameterization exposes compensating
model errors, whose diagnosis and alleviation adds
significantly to the original task. It is widely recog-
nized that groups carrying out model development
require a critical mass and need to cover all aspects
of model development to be successful. Despite this,
it has proven difficult for many modeling groups to attract both funding and staff for this important but difficult task. Given the high degree of specialization required, parameterization development is a long-term process and requires attracting and retaining talented staff for long periods of time. Breakthroughs are rare, and the opportunities to publish model development work directly are limited. This often runs counter to funding and promotion models applied in the community today. For a true acceleration of progress in model development, it is vital to increase the size of the model development community. This can be achieved by directly increasing the number of developers at the major modeling centers but also, and maybe more importantly, by improving the engagement of the academic community in the model development process. The latter requires fostering collaborations between academic institutions and modeling centers, both by enhancing the existing collaborations and by creating new opportunities. There have been a number of very good examples on how to achieve a closer collaboration between the academic sector and modeling centers, such as enabling collaboration through infrastructure improvements in the Earth System Modeling Framework (ESMF; www.esmf.ucar.edu) in the United States or the Partnership for Research Infrastructures in Earth System Modelling (PRISM; www.prism.enes.org) in Europe. Similarly, comprehensive collaborations in the area of parameterization development implementing the model development process described here would go a long way toward addressing the issues discussed. Community-based model development efforts already exist [e.g., the very well-known National Center for Atmospheric Research (NCAR) Community Climate System Model (CCSM) effort] and enhancing them through strengthening the missing links highlighted in this short discussion will no doubt increase our success rate in the challenging tasks of model improvement.

Another critical issue that a stronger engagement of academia with modeling centers can address is that of the education of the next generation of model developers. Not only is there a shortage of model developers today, but students and early career researchers are discouraged to take up the challenge of model development. There are a variety of reasons for this, ranging from the perception that model development is a largely engineering-oriented task to the way scientific performance is measured. I can attest from my own experience that model development, if carried out in a strong team environment and with strong focus, can be one of the most intellectually and scientifically rewarding pursuits. It is important to not only convey this message to the next generation of potential model developers but also create an environment in which they can operate successfully. The model development strategy outlined here is meant to constitute a small, but perhaps important, step in opening a debate on how to best achieve this goal.

**SUMMARY.** Better weather and seasonal predictions as well as more reliable climate projections require improved models of the components of the climate systems. It has been shown that the improvement of such models is intricately linked to improving the representation of the physical processes embedded in them. An analysis of the model development process revealed that to accelerate progress in the overall model performance, it is necessary to strengthen the links between model evaluation at the level of the application and the process-oriented refinement of the model formulation, in particular in the area of parameterization. To achieve this requires a closer collaboration of the data, model user, and model development communities on the one hand and the academic and “operational” model development community on the other. It is the responsibility of national and international research programs and the community as a whole to take up the challenge of generating the conditions in which model improvements can be developed on a sound scientific footing at the rate that satisfies society’s needs for improved predictions at all time scales.

**ACKNOWLEDGMENTS.** This essay is the result of discussions with many colleagues and friends to whom I am extremely grateful. My research on model evaluation is supported by the Australian Research Council’s Linkage Project Scheme (LP0883961).

**REFERENCES**


Bony, S., and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback


