Model Hierarchies for Understanding Atmospheric Circulation

Penelope Maher1, Edwin P. Gerber2, Brian Medeiros3, Timothy M. Merlis4, Steven Sherwood5, Aditi Sheshadri6,7, Adam H. Sobel6,8, Geoffrey K. Vallis9, Aiko Voigt8,9, and Pablo Zurita-Gotor10

1Department of Mathematics, University of Exeter, Exeter, UK, 2Courant Institute of Mathematical Sciences, New York University, New York, NY, USA, 3National Center for Atmospheric Research, Boulder, CO, USA, 4Department of Atmospheric and Oceanic Sciences, McGill University, Montreal, Quebec, Canada, 5Climate Change Research Centre, University of New South Wales, Sydney, New South Wales, Australia, 6Department of Applied Physics and Applied Mathematics, Columbia University, USA, 7Department of Earth System Science, Stanford University, Stanford, CA, USA, 8Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA, 9Institute of Meteorology and Climate Research, Department Troposphere Research, Karlsruhe Institute of Technology, Karlsruhe, Germany, 10Instituto de Geociencias, Universidad Complutense de Madrid, Madrid, Spain

Abstract In this review, we highlight the complementary relationship between simple and comprehensive models in addressing key scientific questions to describe Earth’s atmospheric circulation. The systematic representation of models in steps, or hierarchies, connects our understanding from idealized systems to comprehensive models and ultimately the observed atmosphere. We define three interconnected principles that can be used to characterize the model hierarchies of the atmosphere. We explore the rich diversity within the governing equations in the dynamical hierarchy, the ability to isolate and understand atmospheric processes in the process hierarchy, and the importance of the physical domain and resolution in the hierarchy of scale. We center our discussion on the large-scale circulation of the atmosphere and its interaction with clouds and convection, focusing on areas where simple models have had a significant impact. Our confidence in climate model projections of the future is based on our efforts to ground the climate predictions in fundamental physical understanding. This understanding is, in part, possible due to the hierarchies of idealized models that afford the simplicity required for understanding complex systems.

Plain Language Summary Model hierarchies are fundamental to our understanding of the large-scale circulation of Earth’s atmosphere. They have played a critical role in forming and testing our ability to simulate and predict the natural variability of the atmosphere, such as the variations of the extratropical jet streams, and for exploring how the climate will respond to external forcing, such as increased carbon dioxide. In this review we discuss simple models that form the basis of our understanding of the atmosphere and how they connect to the comprehensive models used for climate prediction through the model hierarchies. We describe three principles that help organize the model hierarchies and discuss benchmark models that have been influential in understanding the large-scale circulation in the midlatitudes, middle atmosphere, and tropics.

1. Introduction

In this review, we showcase idealized models which have enabled a deeper understanding of the large-scale circulation of the atmosphere and provide a set of principles for organizing them into hierarchies. We regard a hierarchy to be a sequence that connects our most simple models to our most complex, with the ultimate goal of explaining and predicting the behavior of Earth’s atmosphere. The simplicity, or idealization, of a model is thus defined relative to other members of the hierarchy, where a simpler model seeks to reduce the problem to its most fundamental components at the cost of quantitative accuracy and realism.

We use simple models to ask fundamental science questions, which are ideally validated against observations of the real atmosphere. In practice, simple models are often validated against more complex models in the hierarchy. This is necessary when observations are sparse, such as in the upper stratosphere or Southern
Figure 1. Categorizing atmospheric climate models in terms of (a) complexity as in Bony et al. (2013) and (b) model configurations as in Jeevanjee et al. (2017).
Hemisphere storm tracks, or when data are not available, such as projecting future climates given different emissions scenarios.

We are conscious of the subtle differences between a “theory” and a “model.” Here we consider a model to be a set of equations which seeks to capture the behavior of a system we wish to understand. The model may only capture selected features of the system and may not be broadly applicable. A model may also refer to the implementation of an idea in a more general sense. Theory may be regarded as the assumptions and general principles needed to predict the behavior of some phenomena or system, ideally in a way that is both parsimonious and that has some general applicability. Still, the distinction is blurry. For example, a simple, testable model will have many of the attributes of a theory. Further, the behavior of a complex system may not be directly explainable by a simple theory in the conventional sense, and a model hierarchy itself then becomes a theory, or at least a hypothesis, for the system; some of these issues are discussed further in Vallis (2016).

In this review, we focus on models that may be deliberately simplified and which implement a set of (usually time-dependent) equations, sometimes independently of any specific theory. Having differing degrees of complexity, connected to each other in some way, is the key step in sorting models into a hierarchy.

While the spectrum of available models has increased in the last decade or two, the idea of a “hierarchy of climate models” in itself is not new. Schneider and Dickinson (1974) may have been the first to explicitly discuss the hierarchy in the sense we understand today, commenting that “solid progress in understanding … climate change will require steady development of an almost continuous spectrum or hierarchy of models of increasing physical or mathematical complexity.” A decade later Hoskins (1983) noted the “unhealthy” trend toward building models which are disconnected from one another and the real world, advocating, like Schneider and Dickinson, for a spectrum of connected models to provide a complete and balanced approach. Nof (2008) criticized the trend in climate modeling for higher resolution over increased understanding, and pointed out the danger of regarding comprehensive models as “truth.” Polvani et al. (2017) noted that “Earth system models may be good for simulating the climate system but may not be as valuable for understanding it.”

This gap between our understanding of the atmospheric circulation and the increasing complexity of general circulation models was the focus of Held (2005) and Held (2014). In these essays, Held echoed earlier concerns about relying too much on comprehensive models that we do not fully understand. He argued, however, that we should be equally concerned whether our simpler models are capable of addressing our key scientific questions. He advocated for the community to study a subset of models that are sufficiently complex to capture key elements of the real atmosphere, but still simple enough to provide understanding, that is, a set of “elegant” models.

There is certainly no single unique hierarchy. Instead, a suitable model hierarchy may be constructed based on the key scientific questions of interest. Clearly, no one model is suitable for all purposes. Even for a given scientific problem, individual scientists will make different, perhaps equally defensible, choices. Nevertheless, we can attempt to produce a classification system to describe models as being simple or complex within the spectrum of available models. Bony et al. (2013) intuitively describe the complexity of climate models, see Figure 1a, as a balance between simplicity of the model and complexity of the system that is being modeled. More recently, Jeevanjee et al. (2017) describe the climate model hierarchy, see Figure 1b, in terms of dynamics, boundary layer forcing, and bulk forcing. In section 2 we propose an alternative, but complementary, description based on organizing the model hierarchies in terms of three principles.

In this discussion of the large-scale circulation of the atmosphere, we focus our review on the science questions that have been addressed using key idealized models. We structure our review to start with the most simple models and build up toward the more complicated models used to investigate the large-scale circulation within the midlatitudes, middle atmosphere, and tropics, sections 3–5, respectively. We then discuss the important role moisture plays in setting the atmospheric circulation in section 6 and how the hierarchies have helped improve the representation of, and theory for, the Madden-Julian Oscillation (MJO) in section 7. We then summarize in section 8. To aid the reader we have included the list of the acronyms used in the paper in Table 1.

We do not attempt to review all models. Instead, we describe a subset of simple models, discuss their broad use and then make connections from the simple models through to the coupled Atmosphere-Ocean General
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full name</th>
<th>Section (if relevant)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCM</td>
<td>Single Column Models</td>
<td>6.3</td>
</tr>
<tr>
<td>GCM</td>
<td>General Circulation Model</td>
<td>-</td>
</tr>
<tr>
<td>AGCM</td>
<td>Atmosphere-Only (prescribed SST) GCM</td>
<td>-</td>
</tr>
<tr>
<td>AOGCM</td>
<td>Coupled Atmosphere-Ocean GCM</td>
<td>-</td>
</tr>
<tr>
<td>CRM</td>
<td>Cloud Resolving Model</td>
<td>5.1, 6.1, and 7</td>
</tr>
<tr>
<td>LES</td>
<td>Large-Eddy Simulation Model</td>
<td>6.1</td>
</tr>
<tr>
<td>ESM</td>
<td>Earth System Model</td>
<td>1, Figure 7</td>
</tr>
<tr>
<td>CESM</td>
<td>Community Earth System Model</td>
<td>8, Figure 7</td>
</tr>
<tr>
<td>GFDL</td>
<td>Geophysical Fluid Dynamics Laboratory</td>
<td>Figures 3, 8</td>
</tr>
<tr>
<td>MPI-ESM</td>
<td>The Max Planck Institute for Meteorology Earth System Model</td>
<td>6.2, Figure 5</td>
</tr>
<tr>
<td>ECS</td>
<td>Equilibrium Climate Sensitivity</td>
<td>6.3</td>
</tr>
<tr>
<td>OLR</td>
<td>Outgoing Longwave Radiation</td>
<td>Figure 4</td>
</tr>
<tr>
<td>PV</td>
<td>Potential Vorticity</td>
<td>3.3 and 4.1</td>
</tr>
<tr>
<td>SST</td>
<td>Sea Surface Temperature</td>
<td>5.1, 6.1 and 8</td>
</tr>
<tr>
<td>QG</td>
<td>Quasi-Geostrophic</td>
<td>3.2</td>
</tr>
<tr>
<td>RCE</td>
<td>Radiative Convective Equilibrium</td>
<td>6.1</td>
</tr>
<tr>
<td>WTG</td>
<td>Weak Temperature Gradient</td>
<td>6.1</td>
</tr>
<tr>
<td>ENSO</td>
<td>El Niño–Southern Oscillation</td>
<td>5.1</td>
</tr>
<tr>
<td>ITCZ</td>
<td>Intertropical Convergence Zone</td>
<td>5.2</td>
</tr>
<tr>
<td>MJO</td>
<td>Madden-Julian Oscillation</td>
<td>7</td>
</tr>
<tr>
<td>QBO</td>
<td>Quasi-Biennial Oscillation</td>
<td>4.2</td>
</tr>
<tr>
<td>SSW</td>
<td>Sudden Stratospheric Warming</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Circulation Models (AOGCMs). (The role of GCMs is discussed in more detail in Ghil & Robertson, 2000.) We will not discuss Earth System Models (ESMs), the very complex models that include more processes than typical GCMs (e.g., biogeochemistry), but we do acknowledge that ESMs form an end point (if only by definition) in modeling processes that affect Earth’s climate.

## 2. Three-Principles Guiding Model Hierarchies

Many model hierarchies are possible, depending in part on the science questions to be addressed. Nevertheless, a broad classification of the hierarchies is useful and here we define three principles that can be used to guide the categorization of the model hierarchies. We introduce examples of model hierarchies in Figure 2, which highlights a number of the models to be discussed in this review.

The first principle is a *dynamical* hierarchy of atmospheric fluid flow. Dynamical hierarchies allow us to isolate and explore the importance of different temporal and spatial scales on the governing equations, for example, the quasi-geostrophic equations, appropriate for the synoptic-scale circulation of the midlatitudes, as opposed to the nonhydrostatic equations, necessary for capturing the smaller-scale flow in atmospheric convection.

The second principle is a *process* hierarchy of the atmosphere. Process hierarchies allow for the stepwise integration of important atmospheric processes into the governing equations of the fluid flow. For example, we can systematically advance terms with the thermodynamic equation to form a sequence of models that
Figure 2. The three-principles view of model hierarchies used for understanding the large-scale circulation. A dynamical hierarchy can be constructed by systematically varying the governing equations of the fluid flow. Two sample process hierarchies capture systematic variations in the representation of the thermodynamic processes and the boundary conditions. Convective organization at different scales is used as an example to illustrate the scale hierarchy. For each list, the first element is the simplest or smallest scale and builds down to the most complex or largest scale.

make a “diabatic hierarchy,” as listed in the second group in Figure 2. Another process hierarchy of the general circulation can be constructed by systematically increasing the complexity of the boundary conditions, advancing from an aquaplanet configuration to include the impacts of land and orography.

The third principle is a hierarchy of scale, implicit to both the dynamical and process hierarchies, where the choice of physical domain and numerical resolution allows for the systematic exploration of different dynamical and physical processes across time and spacial scales. There are practical trade-offs between scale and complexity due to the computation expense. Unlike the other two principles, the scale hierarchy is not so much a hierarchy of complexity: a limited area large eddy simulation can require as much computational power as a global model integration that accounts only for the large-scale flow. Nevertheless, the principle of scale helps describe the practical decisions about relevant space and time scales required when building a model.

We appreciate that these three organizing principles are not entirely orthogonal; the choice of a particular process, for example, convection, necessitates a choice of dynamical equations and scales. In addition, almost all theory and modeling efforts can be classified into a hierarchy of some form. An attempt to catalog all the hierarchies would therefore ultimately prove unhelpful.

In the remainder of this paper, we therefore selectively highlight examples of model hierarchies, specifically those that include simple models and that have advanced our understanding of the large-scale circulation of the atmosphere.

We introduce a number of models, highlighting how they can be organized in terms of the three principles of dynamics, process, and scale. Some of these models are shown in Figure 2 as steps within each hierarchy organized from simple (top) to complex (bottom). We will introduce each of these models with the sections that follow, starting with simple dry dynamics models and building up to more complex moist GCMs. We focus on these models not because they necessarily optimally cover the complexity of available models, but rather because they have been extensively studied, thus establishing their value as benchmark models of the atmospheric circulation.
3. The Midlatitude Circulation

The large-scale extratropical circulation provides one of the best success stories for hierarchical climate modeling: some key aspects of the underlying dynamics are now reasonably well understood and part of modern textbooks (e.g., Vallis, 2017). Other aspects are still areas of active research, such as the nonlinear dynamics related to eddy-mean flow interaction. Idealized simulations have played an instrumental role in this progress, providing key insights on the nonlinear behavior of extratropical disturbances. Since the early days of climate modeling, theorists recognized the great power of numerical computing as a means to overcome the stringent limitations of analytical work. Idealized simulations aimed at understanding the atmosphere were performed in parallel with comprehensive simulations. Some of the insights gained with these early simulations constitute the basis of prevalent paradigms on the extratropical circulation.

We begin by highlighting two models that have allowed us to isolate the key elements of the midlatitude circulation. The first model is a class of barotropic vorticity equation models, where collapsing the vertical dimension allows us to focus on feedbacks between the zonal mean flow, Rossby waves, and the spherical geometry of the planet (section 3.1). The second model is the two-layer Quasi-Geostrophic (QG) channel model, which provides perhaps the most simple context for understanding baroclinic instability (section 3.2). We then discuss an idealized approach to combining elements of the baroclinic and barotropic dynamics together in eddy life cycle experiments (section 3.3). We then describe how the model hierarchies in the midlatitudes have improved our understanding of the eddy-driven jet stream (section 3.4).

3.1. Rossby Wave Dynamics: The Barotropic Vorticity Equations on the Sphere

Rossby wave propagation plays a fundamental role in both upper-troposphere synoptic variability and the remote atmospheric response to forcing. The barotropic model provides a simple framework for studying these processes. In addition to providing the first numerical weather simulations (Charney et al., 1950), the barotropic model served as a test bed to understand the influence of topography and localized heating on the general circulation (Grose & Hoskins, 1979; Hoskins & Karoly, 1981). These experiments revealed the important role played by the mean flow structure for Rossby wave refraction in the upper troposphere. The widely used concepts of wave guides and propagation windows are based on these ideas, which are key to our understanding of the extratropical response to the El Niño–Southern Oscillation (ENSO).

So-called “stirred” barotropic models (e.g., Vallis et al., 2004) have seen a resurgence in recent years for understanding upper-troposphere synoptic variability and the dynamics of eddy momentum fluxes and eddy-driven jets without the complexity of baroclinic dynamics. In this model, the impact of baroclinic instability is approximated by a prescribed forcing (the stirring) in the vorticity equation at the synoptic scales. As a result, there are explicitly no feedbacks of the barotropic circulation on eddy generation. The model has been used as a conceptual model of annular mode variability to explain the dependence of zonal index persistence on latitude (Barnes et al., 2010) and to study the interaction between the tropical and subtropical jets (O’Rourke & Vallis, 2013), among other problems.

As a further simplification, when the model is linearized, it is possible to obtain a set of closed solutions (for simple forms of stirring) using stochastic theory (DelSole, 2001). Lorenz (2014) has devised a very sophisticated method to calculate the eddy momentum flux given the full space-time characteristics of the stirring, which can play an important role due to the impact of wave phase speeds on refraction indices and wave propagation (Barnes & Hartmann, 2011). The barotropic model can be a useful tool for exploring “eddy-momentum-flux closures,” that is, the sensitivity of the direction and magnitude of the wave propagation to the mean state and/or model configuration. This remains a challenging open question in general circulation theory.

3.2. Baroclinic Instability: The Two-Layer Quasi-Geostrophic Model

To capture the essence of the eddy generation process, the two-layer quasi-geostrophic model on the β plane stands out as a benchmark, indeed classical model of the extratropical baroclinic circulation (Phillips, 1956). It competes with the Eady model (Eady, 1949) as the simplest model that can produce baroclinic instability in a fashion relevant to the real world. There is only one baroclinic mode and the stratification and radius of deformation are prescribed.

The model also provides a simplified framework for studying the nonlinear extratropical circulation in a forced-dissipative configuration, in which the flow is typically forced by thermal relaxation to a baroclinic jet and damped using Rayleigh friction (e.g., Panetta, 1993; Salmon, 1980; Zurita-Gotor, 2007). The β plane
approximation and constant deformation radius make upper-troposphere dynamics simpler than in the spherical case (the symmetry of the model makes northward and southward propagation equally likely). In this sense, the two-layer QG model complements the spherical barotropic model, for it is devoid of the feedbacks associated with sphericity that are isolated in the latter model. The two-layer model not only reproduces qualitatively the main features of the observed extratropical circulation but also captures more subtle aspects of extratropical dynamics like the clustering of eddies in wave packets (Lee & Held, 1993), the driving of low-frequency baroclinicity variability (Zurita-Gotor et al., 2014), or the character and intensity of the eddy momentum and heat fluxes (Larichev & Held, 1995; Lutsko et al., 2017; Vallis, 1988).

In its forced configuration, the two-layer model provides the lower end of a dynamical hierarchy of forced-dissipative dry models, in which the mean climate is determined by the competition between the eddy fluxes and very idealized forms of forcing. These models can be formulated at different levels of complexity along the dynamical hierarchy depending on the scientific problem of interest (e.g., Jansen & Ferrari, 2013; Lachmy & Harnik, 2014; Zurita-Gotor & Vallis, 2009). At the complex end of this particular dynamical hierarchy, the primitive equation model of Held and Suarez (1994) has been widely used to study various aspects of the extratropical circulation and its sensitivity to climate change (e.g., Butler et al., 2010; Lorenz & DeWeaver, 2007; Yuval & Kaspi, 2016) due to its relatively realistic circulation. This model is set in a spherical domain and is forced by relaxation toward a state approximating Radiative Convective Equilibrium (RCE, see section 6.1), with near moist-neutral stratification in the vertical but strong meridional temperature gradients. Above the tropopause, the atmosphere is simply relaxed toward an isothermal state. A variant of this model better suited for the tropical circulation combines relaxation to pure radiative equilibrium with an idealized convection scheme designed to mimic the stabilizing effect of latent heating by moist convection (Schneider & Walker, 2006).

3.3. Connecting Eddy Growth, Propagation, and Decay: The Eddy Life Cycle Model

Even in the very idealized physical setting described above, the time-dependent evolution of forced-dissipative models is inherently nonlinear and turbulent. As a key simplification to the full nonlinear problem, the series of experiments systematized by Hoskins and collaborators in the 1970s, building on pioneering numerical work by Edelmann (1963) and others, provided insight on the nonlinear evolution of baroclinic modes. The analysis of an eddy life cycle by Simmons and Hoskins (1978) introduced the notions of baroclinic growth and barotropic decay as an idealized conceptual model for the nonlinear evolution of extratropical disturbances. Similar ideas, but in the more general context of a statistical steady state and using quasi-geostrophic theory to interpret the simulations, were introduced independently by Salmon (1980). This simple paradigm has survived to today and plays a fundamental role for our understanding of wave-mean flow interaction and the maintenance of the mean circulation. Additional analysis in Simmons and Hoskins (1980) uncovered the sensitivity of the decay stage in the life cycle to the mean state, identifying two distinct patterns of evolution.

As theoretical advancements clarified the relation between eddy propagation and wave-mean flow interaction (Andrews & McIntyre, 1978; Edmon et al., 1980) and the focus on Potential Vorticity (PV) dynamics highlighted the important role of wave breaking (McIntyre & Palmer, 1983), Thornicroft et al. (1993) proposed a conceptual model for understanding the two idealized life cycles based on the direction of propagation and the typology of wave breaking. Idealized simulations were also useful for demonstrating the relevance of critical layer theory for eddy dissipation and wave-mean flow interaction in eddy life cycles (Feldstein & Held, 1989). The critical layer is a powerful concept for constraining upper-troposphere propagation (Randel & Held, 1991) and plays an important role for extratropical variability and climate sensitivity (Ceppi et al., 2013; Chen & Held, 2007; Lee et al., 2007).

The association between the direction of propagation, the topology of wave breaking, and the sign of the eddy momentum flux uncovered by the idealized studies is central to our understanding of jet shifts and phenomena like the North Atlantic Oscillation (Rivière & Orlanski, 2007). On the sphere, equatorward propagation and poleward momentum fluxes dominate (Balasubramanian & Garner, 1997; Thornicroft et al., 1993) so that we might expect extratropical jets to shift poleward as they strengthen if the stirring does not move. However, idealized studies show that the direction of propagation is affected by many other factors, such as the latitude and scale of the eddies, the barotropic shear and the low-level baroclinicity (Hartmann & Zuercher, 1998; Rivière, 2009; Simmons & Hoskins, 1980), among others. Due to this complexity, we are still far from a complete theory for the eddy momentum flux closure.
We acknowledge that the eddy life cycle sits in the gray area between theory and model. Unlike the barotropic vorticity (section 3.1) or two-layer quasi-geostrophic (section 3.2) models, which involve the integration of a specific set of equations on a well-defined domain, eddy life cycles can be run in a number of configurations with a range of dynamical choices. Thus, the eddy life cycle is not so much a model as an approach to simplifying the full turbulent system. As this approach has provided us with enhanced understanding on the nonlinear evolution of baroclinic instability, the eddy life cycle has also become a paradigm; hence, it is a theory.

The simplifications afforded by the initial-value approach can also be beneficial in other contexts (e.g., Hoskins & Jin, 1991). The main limitation of this approach is the sensitivity to the initial conditions: as
there is an infinitely large space of possible initial conditions that could be explored, the robustness of the results can only be assumed by their consistency with the behavior of the time-dependent system. One reason why the initial-value calculations of Simmons and Hoskins (1978) have become a well-accepted paradigm is their consistency with the energetics of the full turbulent system and the phenomenology of quasi-geostrophic turbulence (Salmon, 1980). Finally, we mention that ensemble “switch-on experiments” (e.g., Nie et al., 2016) may provide a useful approach for bridging the results of initial-value calculations with the full time-dependent system.

### 3.4. Case Study: Eddy Feedbacks and the Variability of the Jet Stream

To illustrate the use of hierarchical modeling in the extratropics, we discuss its application to the analysis of eddy feedbacks in unforced jet variability. We have chosen this example because it lends itself well to the hierarchical approach and because it is a topic of current research.

The leading (and more persistent) mode of extratropical zonal wind variability consists of a meridional shift of the eddy-driven jet concomitant with annular mode variability (Thompson & Wallace, 2000). Lorenz and Hartmann (2001) found a positive correlation between the jet anomalies and their eddy momentum driving in the Southern Hemisphere when the jet leads by a few days, see Figure 3a, which implies that the anomalous eddy momentum fluxes tend to extend the duration of the jet anomalies. They interpreted this positive correlation as depicting the sensitivity of the anomalous eddy momentum flux on the state of the jet or a positive eddy feedback (but see Byrne et al., 2016 for an alternative interpretation).

Climate models are known to be too persistent (Gerber et al., 2008), see Figure 3b, particularly idealized models (Gerber & Vallis, 2007). This is mostly associated with too slow decay of the annular mode autocorrelation function at lags beyond 5 days, see Figure 3c, suggesting an excessive eddy feedback. Two different types of mechanisms have been proposed in the literature for this feedback: barotropic and baroclinic. Barotropic mechanisms rely on changes in upper-troposphere propagation due to changes in refraction in the presence of the anomalous jet, which may involve a number of different mechanisms (Burrows et al., 2017; Lorenz, 2014). In contrast, baroclinic mechanisms attribute the eddy momentum flux changes to changes in the stirring driven by the changes in the barotropic flow (Robinson, 2000).

Idealized models provide a useful framework for studying these two aspects of the problem in isolation. Using the stirred barotropic model, Barnes et al. (2010) investigated the sensitivity of the eddy momentum fluxes to the anomalous jet with fixed stirring. They showed that on the sphere, the eddy momentum flux becomes more asymmetric (equatorward propagation is enhanced) when the jet moves poleward, leading to a positive feedback. This may be understood in terms of changes in the turning latitude/reflecting level (Lorenz, 2014).

In the opposite direction, Zurita-Gotor et al. (2014) analyzed the dynamics of jet variability in idealized two-layer QG simulations and showed that the enhanced persistence in that model was consistent with the baroclinic feedback mechanism of Robinson (2000). They found evidence of baroclinicity driving the barotropic flow and very large coherence between the eddy heat and momentum fluxes at low frequency, with the momentum fluxes leading the variability, see Figure 3e. The covariability between the barotropic and baroclinic components of the wind is also a robust result in observations (Blanco-Fuentes & Zurita-Gotor, 2011) and comprehensive climate models. In Figure 3d the large correlation between the long-lag decay rates of (barotropic) jet anomalies and baroclinicity is shown for a selection of CMIP5 models, so that models with more persistent jet variability also tend to have more persistent baroclinicity.

Stirred barotropic models can capture some aspects of the observed jet variability, like the sensitivity of persistence to latitude (Barnes et al., 2010). On the other hand, the baroclinic mechanism may help explain the excessive persistence bias in comprehensive climate models, which cannot be corrected by eliminating the jet latitude bias (Simpson et al., 2013), or in idealized baroclinic models. Finally, diabatic effects may also play a role for annular mode persistence (Xia & Chang, 2014). The jet persistence problem underscores the importance of making connections across the full model hierarchy, as the mechanisms at work may not be the same in all steps of the hierarchy, in comprehensive climate models and in the real atmosphere.

### 4. The Middle Atmosphere Circulation

Work over the last two decades has established the highly coupled nature of the circulation in the troposphere and stratosphere. Many comprehensive atmospheric models now treat the stratosphere and...
troposphere as one system (e.g., Gerber et al., 2012), recognizing the consequences of underresolving the middle atmosphere for weather and climate prediction (e.g., Manzini et al., 2014; Sigmond et al., 2013). However, historically the middle atmospheric research proceeded on a different track after Charney and Drazin (1961) showed that a detailed representation of the stratosphere was not necessary to capture the basic structure of synoptic variability in the troposphere.

Wave-mean flow theory was developed, in part, to explain and understand the stratospheric circulation. The gross structure of the stratosphere cannot be explained without understanding the essential role of waves in the transport of momentum, mass, and tracers. We highlight three models that capture these interactions, and the more sophisticated steps in the hierarchy they have inspired. First, the Holton-Mass model of wave-mean flow interaction (section 4.1). Second, understanding of transport within the stratosphere (section 4.2). Third, the so-called leaky pipe model that provides a simple context to understand the stratosphere (section 4.3).


Cooling during the polar night generates a strong westerly jet in the winter stratosphere, often referred to as the stratospheric polar vortex, where wind speeds can sometimes reach 100 m/s. In the early 1950s, however, it was observed that the polar vortex in the Northern Hemisphere aperiodically undergoes a rapid breakdown. The reversal of the westerly winds is associated with a dramatic warming (40 K or more in the course of a few days) and hence known as a Sudden Stratospheric Warming (SSW; Scherhag, 1952). SSWs occur on average once every other winter in the Northern Hemisphere, but only one such event (in 2002) has been observed in the Southern Hemisphere.

Baldwin and Dunkerton (2001) showed that SSWs affect the troposphere, shifting the jet stream equatorward with substantial impacts on weather in Europe and Eastern North America. The tropospheric impact persists on the 1- to 2-month time scale that it takes the stratospheric polar vortex to recover back to its climatological state.

Matsuno (1971) proposed a dynamical mechanism for SSWs based on planetary scale wave propagation from the troposphere. Long before this process could be captured in atmospheric GCMs, Holton and Mass (1976) developed a simple, stratosphere-only, model that captures the essence of these abrupt events. They constructed a highly truncated baroclinic quasi-geostrophic model, retaining only wave number 1 and the mean flow. The mean state is forced by Newtonian relaxation toward a specified state of radiative equilibrium, the wave generated by specifying a forcing amplitude on the bottom boundary.

The model reveals an abrupt transition in response to a smooth variation in the amplitude of the wave forcing. For low amplitudes, it exhibits a subcritical state in which westerly winds coexist with a stationary Rossby wave. If the wave amplitude at the lower boundary exceeds a critical threshold, however, the model transitions abruptly to a new equilibrium. The wave grows, weakening the westerlies until they reverse, thereby preventing further wave growth: the essence of an SSW.

Multiple flow equilibria have also been demonstrated in more complex three-dimensional (3-D) stratosphere-only models—again forced by specifying the amplitude of planetary waves at the lower boundary—but permitting arbitrary height and latitude structure in the stratosphere (e.g., Scott & Haynes, 2000; Scott & Polvani, 2006). The highly idealized Holton and Mass (1976) model, however, has continued to inspire research on the role of gravity waves in SSWs (e.g., Albers & Birner, 2014), and the role of the stratosphere on regulating wave activity (e.g., Sjoberg & Birner, 2014).

These models suggest that the near absence of SSWs in the Southern Hemisphere is due to the fact that stationary wave amplitude is weaker, a process explored in full 3-D atmospheric models using a Held and Suarez (1994) forcing, albeit with a modified equilibrium temperature profile in the stratosphere to establish a polar vortex. Transitioning from a Southern Hemispheric state to a Northern Hemispheric state is possible by increasing the amplitude of surface topography (Sheshadri et al., 2015; Taguchi et al., 2001; Taguchi & Yoden, 2002). Held and Suarez (1994) type models have also allowed for exploration of the impact of the vortex strength on the troposphere, both in response to forced changes (Polvani & Kushner, 2002) or SSWs (Gerber & Polvani, 2009).

4.2. The Quasi-Biennial Oscillation: A Physical Model

While high-latitude variability in the stratosphere is dominated by interactions between planetary scale waves and the mean flow, tropical variability is effected by wave-mean flow interactions involving much
smaller scale gravity waves. The Quasi-Biennial Oscillation (QBO) is an oscillation of the zonal mean wind in the tropical stratosphere with a period of ~28 months, associated with the slow downward migration of alternative westerly and easterly jets (see Baldwin et al., 2001, for a comprehensive review). The long time scales of the QBO make it a potential source of predictability in the troposphere. For example, the QBO is associated with changes in the strength and predictability of the MJO (e.g., Yoo & Son, 2016). The QBO also provides another example of the advances that a simplified system can bring about, well ahead of our ability to simulate the phenomenon in comprehensive models.

Pioneering work by Lindzen and Holton (1968) and Holton and Lindzen (1972) proposed that the QBO could be explained as an interaction between gravity waves and the mean flow. Selective absorption or breaking of waves carrying easterly (westerly) momentum on the lower flank of easterly (westerly) jets would generate a momentum tendency that pulls the jets downward, enough to oppose the tendency of the mean tropical upwelling to advect the jet upward. The balance between the two effects could produce the very slow variation in the jets.

These theories of the QBO came long before we had the ability to observe—or simulate—the small-scale gravity waves implicated in the mechanism; even today, gravity waves provide a challenge to observe and model (e.g., Alexander et al., 2010). To establish the mechanism, Plumb and McEwan (1978) developed a novel physical model of the phenomenon. Models of the atmosphere generally refer to numerical models, but we also want to highlight the value of physical, laboratory models in this review.

The Plumb and McEwan (1978) model consists of an annulus of stratified salt water and internal waves forced by mechanically oscillating the lower boundary. The waves generate spontaneous formation of jets (an azimuthal circulation in the annulus), with slow oscillations and reversal of the flow, similar to the QBO of the atmosphere. The model firmly established the validity of the wave-mean flow interaction mechanism, decades before we could fully simulate it with a numerical model.

4.3. Stratospheric Transport: The Leaky Pipe

Transport and chemistry play key roles in the distribution of trace gases throughout the stratosphere, including water vapor, ozone, and the substances that deplete ozone. The meridional overturning circulation of the stratosphere, known as the Brewer-Dobson Circulation, was first inferred from trace gas measurements, without observing the circulation directly (Brewer, 1949; Dobson, 1956). Trace gases are advected by the mean Lagrangian circulation of mass and mixed along isentropic surfaces in the process of wave breaking. The latter mixing process produces no net transport of mass, but will transport a trace gas if there is a horizontal gradient in its concentration.

Efforts to understand stratospheric transport began with limiting cases in the balance between transport of tracers across isentropic surfaces by the mean overturning mass circulation versus the mixing of tracers along isentropic surfaces. Plumb and Ko (1992) consider a circulation where mixing along isentropic surfaces is extremely efficient. In contrast, Plumb (1996) developed the idea of a “tropical pipe,” where upwelling air in the tropics is entirely isolated from the downwelling air in the higher latitudes and transport is set by the mean mass circulation alone. These two limiting cases were combined in a benchmark model in our understanding of transport processes, the “leaky pipe” model of Neu and Plumb (1999).

The leaky pipe divides the stratosphere into two regions, an upwelling “pipe” in the tropics and a downwelling pipe in the extratropics of both hemispheres. Mass is advected up the tropical pipe by the Lagrangian mean circulation, detraining continually out to the extratropics. The boundary between the two regions, the edge of the stratospheric surf zone, is a barrier to transport, but the “leaky” pipe allows for some mixing of mass between the two. A key result of the model is that an increase in the net Lagrangian mass transport will tend to freshen the stratosphere, cycling tracers more quickly through it, while an increase in mixing tends to slow the cycling, as mixing leads to recirculation of air through the stratosphere.

While designed primarily as a conceptual model, the leaky pipe has been applied in a more realistic context to understand the make up of the stratosphere and its response to anthropogenic forcing. Garny et al. (2014) use it to interpret changes in the stratospheric circulation in comprehensive models, separating the roles of mixing from the mean Brewer-Dobson Circulation. Ray et al. (2010) build on the leaky pipe to explain the distribution of trace gases, and Linz et al., 2016 (2016, 2017) use it to quantify the strength of the Brewer-Dobson Circulation from satellite of trace gases, sulfur hexafluoride, and nitrous oxide.
5. The Large-Scale Circulation of the Tropics

Significant progress in understanding the large-scale circulation of the midlatitudes and middle atmosphere was possible in the context of “dry dynamics.” Removing the nonlinearities associated with moist processes simplifies the problem, both conceptually and in terms of the numerical equations, processes, and scales that must be represented or parameterized. Indeed, none of the simple models highlighted in sections 3 and 4 include moist effects. In the tropics, the circulation and moist processes are more intimately coupled. A key scientific challenge for understanding tropical circulation is: How do we deconvolve the tight coupling between circulation, moisture, clouds, and convection?

Nonetheless, there are still dry frameworks for understanding the primary features of the tropical circulation. We will explore the Matsuno-Gill model that captures the equatorial zonal overturning circulation—the Walker circulation—using the dry shallow water equations (section 5.1). We will then focus on the zonal mean tropical overturning circulation—the Hadley circulation—again starting the discussion with a dry atmospheric model and then introducing an idealized GCM that begins to capture moist processes (section 5.2).

5.1. The Walker Circulation: The Matsuno-Gill Model

The Walker circulation describes equatorial atmospheric cells with ascent over the Maritime Continent (equatorial Western Pacific) and descent in the Eastern Pacific or Indian Oceans. The number of equatorial circulation cells and the location of ascending and descending branches are coupled with sea surface temperatures (SSTs) and the phase of ENSO (Julian & Chervin, 1978). Research on the Walker cell spans fundamental questions about the interaction between circulation and convection to very practical concerns, for example, how does ENSO influence the onset of the monsoons, and how will this change with global warming?

Similarly to the midlatitudes, many simple models for the tropical circulation hinge on reducing the dimensions of the atmospheric flow. A key simplification is to vertically truncate the fluid-governing equations. One such model that has been fundamental for understanding the structure of the Walker circulation is the Matsuno-Gill model (Gill, 1980; Matsuno, 1966; Webster, 1972), which uses the dry shallow water equations on an equatorial-β plane with a stationary heating source (e.g., Vallis, 2017, section 8.5). This single-layer model provides an analytic solution for the horizontal structure associated with the first baroclinic mode. This vertical mode captures the circulation driven by heating associated with tropical deep convection and is characterized by opposite signed flow in the upper versus lower troposphere. As the troposphere does not have a rigid upper boundary, it is not a true “mode” as in the ocean, but it often behaves like one.

The model’s solution is generally described as the Matsuno-Gill pattern, in which two steady state circulation cells develop in response to the applied heating, with low-level convergence into and upper-level divergence out of the heating region. This generates an eastward propagating Kelvin wave and a westward propagating Rossby wave. Two off-equatorial low-pressure systems form as the Rossby wave is damped at the equator (see Figure 8.11 of Vallis, 2017). This equatorially symmetric component of the Matsuno-Gill model generally describes the observed structure of the Walker circulation, with analogous tropical convection in the West Pacific and descent over the cold SST in the East Pacific (due to deep water upwelling).

The Matsuno-Gill model has also been used as the atmospheric component of the first successful numerical ENSO prediction model, the Cane-Zebiak model (Cane et al., 1986), a very influential reduced complexity coupled atmosphere-ocean model. In addition, the Matsuno-Gill model captures monsoonal circulations, using off-equatorial heating to mimic the seasonal cycle. Gill (1980) showed that the antisymmetric Matsuno-Gill pattern (see Figure 3 of Gill, 1980), describes the general structure of the monsoon flow (Rodwell & Hoskins, 1996). Furthermore, the Matsuno-Gill model is important for understanding the propagation of the MJO, as detailed in section 7.

While some aspects of the Walker circulation are captured by the Matsuno-Gill model, its primary limitation is that it does not include interactive moisture. As a result, many important moist feedback mechanisms are absent. One approach to studying the moist Walker circulation is to impose a large-scale gradient of SST in a two-dimensional atmospheric model domain, creating a steady state Walker circulation, commonly called the “mock” Walker circulation.
Bretherton et al. (2006) studied the moist Walker circulation using an idealized nonrotating 2-D model following the approach of the quasi-equilibrium tropical circulation model (Neelin & Zeng, 2000). The model used in Bretherton et al. (2006) is vertically truncated (one vertical moisture mode), assumes the Weak Temperature Gradient (WTG) approximation (discussed more in section 6.1) and uses simple precipitation and cloud schemes; see their Figure 4 for the resulting circulation. This is a useful prototype model configuration because it allows explicit Cloud Resolving Model (CRM) and GCM-physics comparisons of a climate relevant problem (Jeevanjee et al., 2017). The beauty of this idealized model is that it includes feedbacks between convection and the large-scale circulation. In comparing to 3-D CRMs, Bretherton et al. (2006) showed many interesting features within the two models: similar precipitation but different humidity distributions, narrowing of the circulation with warming SSTs, and the importance of moist static energy in understanding feedbacks between convection and the large-scale circulation within the Walker circulation.

5.2. The Hadley Circulation: Gray Radiation Aquaplanet Models

The Hadley circulation describes the zonally averaged atmospheric circulation of the tropics, with net ascent near the equator, poleward outflow in the upper troposphere, descent in the subtropics, and an equatorward near-surface return flow. The Hadley circulation separates the moist tropical regions from the dry subtropical climate zones. Key research questions include: what controls its strength and extent (i.e., the tropical edge), the location of the near-equatorial ascending region (i.e., the Intertropical Convergence Zone, ITCZ), and how will these features change with global warming?

The long search for an “ideal” Hadley cell, meaning an idealized, axisymmetric circulation on a planet with no longitudinal asymmetries, was described by Lorenz (1967). No solution was presented in his monograph, for none was available at that time, and it was not clear how far such a circulation would extend meridionally. However, Lorenz (and others before him) did assume that such a circulation existed and that it would be baroclinically unstable.

Theoretical progress in finding the ideal circulation was made by realizing that the poleward moving flow would approximately conserve angular momentum (Schneider, 1977) and would therefore have a finite extent. Held and Hou (1980) then developed approximate analytical solutions that facilitated deeper understanding and led to a number of studies along similar lines that were in fact able to successfully predict many features of the observed Hadley circulation (Fang & Tung, 1996, 1997, 1999; Lindzen & Hou, 1988; Plumb & Hou, 1992). The authors of these studies did not claim that eddies were unimportant to the real Hadley circulation, only that the axisymmetric solutions should be understood as providing the basic state for instability and climate studies, as Lorenz (1967) and Schneider (1977) had argued. (This nuance of interpretation, however, arguably faded somewhat with time, and for a while the axisymmetric models were a paradigm for explaining the actual circulation.)

Regardless of its realism or otherwise, the axisymmetric model may be regarded as forming the base of a hierarchy of models of the Hadley cell, offering plausible explanations of both its finite extent and its seasonal variation. However, baroclinic eddies are important in the real atmosphere, not only in midlatitudes but also in affecting the Hadley circulation, and the next step in the hierarchy is to include models with eddy effects in the simplest possible way. Dry, three-dimensional primitive equation models are the obvious choice and have been particularly important in revealing how large-scale baroclinic eddies can interact with the tropical circulation (Kim & Lee, 2001; Schneider & Bordoni, 2008; Walker & Schneider, 2006). Attempts have also been made to parameterize eddy effects in otherwise axisymmetric models (Sobel & Schneider, 2009; Vallis, 1982), and although such zonally averaged models ultimately run up against the problem of turbulence, they too show how the Hadley cell (and not just the midlatitudes) is affected by eddies.

Although dry models have set the foundations for our understanding of the Hadley cell and reveal the effects of eddies in the cleanest way, moist processes are an important aspect of the tropical circulation. Moisture is critical for determining the width of the ascending branch of the Hadley cell and the circulation’s net energy transport, as Vallis (1982) found in an idealized moist zonally averaged model with simple radiation, condensation, and convection schemes.

Three-dimensional models are more revealing and the idealized moist primitive equation GCM of Frierson (2007) has proven particularly instructive. This model has a semigray radiation scheme (i.e., “gray” in the infrared, with only a single band of radiation that interacts with a specified time-independent optical thickness, with a separate solar scheme) that neglects cloud and water vapor feedbacks, so that dynamic
moisture interactions are decoupled from radiative interactions. As in Vallis (1982), the model uses an idealized large-scale precipitation scheme (condensation upon saturation), and a simple convection scheme that relaxes the atmosphere toward a stable vertical profile. It too shows that the Hadley cell is very sensitive to the presence of condensation, which greatly affects the distribution of the diabatic forcing, as well as the representation of convection, which impacts the stratification, with the moist static energy difference between the upper- and lower-level Hadley circulation playing a key role in the strength of the overturning (Frierson, 2007).

Additional studies have used the gray radiation model to explore the effects of moisture and extratropical forcings on the width and location of the ascending branch of the Hadley cell and its relation to the ITCZ (Byrne & Schneider, 2016; Kang et al., 2009). Bordoni and Schneider (2008) use it to investigate controls on monsoonal circulations, which exist even in an aquaplanet configuration. (This said, a large body of work has emphasized the role of land-sea contrast and orography on the monsoons; e.g., Boos & Kuang, 2010; Chou et al., 2001; Privé & Plumb, 2007.)

The idealized model of Frierson (2007) can be linked to higher levels of the hierarchy by including more processes. The monsoons and annual evolution of the ITCZ can be studied by including the seasonal cycle (Geen et al., 2018; Shaw, 2014; Wei & Bordoni, 2018). A second addition is to include spatial variability in the radiative forcing and its feedbacks, for a more realistic response of the atmospheric energy transport to external forcing (Feldl et al., 2017; Merlis, 2015). A third addition is an idealized ocean heat transport coupled to the surface wind stress of the Hadley cell (Codron, 2012; Held, 2001; Levine & Schneider, 2011) that begins to bridge the gap between full-ocean GCMs and slab-ocean boundary conditions. A further extension is to couple the atmosphere and ocean for more realistic ocean heat uptake and transport, which results in more realistic atmospheric energy transport by the Hadley circulation (Feldl & Bordoni, 2016; Zelinka & Hartmann, 2010). Finally, radiative feedbacks can be introduced to the model by replacing the gray radiation scheme with a more realistic representation or radiative transfer (e.g., Jucker & Gerber, 2017; Merlis et al., 2013; Vallis et al., 2018).

6. Coupling Clouds and Convection to the Large-Scale Circulation

A key simplification of the idealized moist models discussed in section 5.2 is to omit the impact of clouds (and cloud microphysics) on the circulation. Clouds are a visible manifestation of atmospheric convection and play a vital role in the radiative budget both locally, within a single convective system, and globally. Clouds are a key uncertainty in predicting the global temperature response to greenhouse gas forcing (see section 6.3). Individual convective clouds can be isolated and appear as random noise in an otherwise homogeneous environment, such as patchy, fair weather cumulus, but can also interact with nearby convection and the environment to form mesoscale convective systems such as squall lines. Organized convection impacts the radiation budget by changing the distribution of cloudy and clear sky. This is important as the radiative properties of clouds shape the large-scale circulation of the atmosphere (Hunt et al., 1980; Slingo & Slingo, 1988; Randall et al., 1989).

Clouds and convection are also embedded within the large-scale circulation of the atmosphere. The ascending branches of the circulation cells promote deep convection within the ITCZ, while the descending branches suppress convection, creating regions in which clear sky or low-level clouds dominate. This two-way interaction is referred to as cloud-circulation coupling. Understanding cloud-circulation coupling, and representing it in models, is one of the World Climate Research Program’s “Grand Challenges” on clouds, circulation, and climate sensitivity (Bony et al., 2015).

In section 6.1 we focus our discussion on RCE, a conceptual model of the tropical atmosphere that has helped us better understand the organization of convection. In section 6.2 we discuss a “cloud locking” approach that decouples cloud radiative effects from the circulation, forming a bridge from the idealized moist GCMs to full atmospheric models. In section 6.3 we describe how the more complex models within the hierarchy are used to study Earth’s Equilibrium Climate Sensitivity (ECS).

6.1. Convective Organization: Radiative-Convective Equilibrium

In section 5.2 we painted a picture of broad ascent within the equatorial branch of the Hadley circulation. This view of tropical circulation is a reasonable approximation on longer time scales (weeks or more). On shorter time scales (hours to days), however, the tropical atmosphere is highly variable, with both ascent
and descent in most regions. Convection on these shorter time scales is organized on small spatial scales, as within a single convective system, and on large scales, as with the ITCZ and MJO.

Convective organization is not well represented in most atmospheric models (Del Genio, 2012). This deficiency has been partly attributed to convective parameterizations that have a number of shortcomings. For example, convection is generally parameterized in the vertical column without any horizontal interactions, models have limited memory of convection from one time step to the next, and parameterizations generally do not represent interactions with the (unresolved) mesoscale circulation (Mapes & Neale, 2011). A number of persistent model biases have been linked to errors in representing convection (Randall et al., 2016). For example, models (i) exhibit too much light rain, which results in insufficient extreme rainfall, (ii) trigger convection too early, resulting in the wrong diurnal cycle, and (iii) often generate a double ITCZ in the central and eastern Pacific (Dai, 2006; Oueslati & Bellon, 2015; Stephens et al., 2010; Sun et al., 2006).

The need to improve comprehensive atmospheric models motivates the use of a hierarchy of models to understand and (ultimately) address these long-standing model biases. Models can also be used to improve our theoretical understanding of convection and identify how convection interacts with both the local environment and larger scales (e.g., Muller & Bony, 2015).

RCE describes a state in which atmospheric radiative cooling is balanced by convective heating in a domain with no externally imposed horizontal structure, for example, uniform SST and insolation. RCE was first considered in the 1960s by Manabe and Strickler (1964), who originally proposed it to explain the vertical structure of the atmosphere. Since then, it has evolved into a test bed for understanding convection in the absence of large-scale circulation. RCE is an important component of a hierarchical approach connecting physical laws to the complex behavior of the Earth system (Popke et al., 2013). High-resolution models in RCE are a useful starting point for theories of convective organization (Muller & Bony, 2015).

Using a nonrotating CRM in RCE, Bretherton et al. (2005) showed that convection can spontaneously self-organize (see Figure 4), a process known as “self-aggregation.” The integration is initialized from a uniform state, and in the first weeks of integration, seemingly random convection is observed homogeneously across the domain. After ~50 days, however, the system transitions to a single-convecting cluster. Self-aggregation is not solely a spatial reorganization of convection; it dramatically changes the mean cli-
mate in CRMs resulting in a dryer troposphere, more Outgoing Longwave Radiation (OLR), warmer free troposphere and surface. See Wing et al. (2017a) for more details and a full list of references, Mapes (2016) for a broader perspective, and Holloway (2017) for a comparison to observations.

Convection also organizes in RCE simulations using GCMs with parameterized convection, in which large convective clusters form spontaneously (Becker et al., 2017; Coppin & Bony, 2015; Popke et al., 2013; Reed & Chavas, 2015). Once convection begins to organize, a large-scale circulation develops and helps maintain the convection. Based on both RCE and GCM studies, convection is more clustered in simulations without parameterized convection, compared to simulations with convection parameterizations, and more intense rain rates occur (Becker et al., 2017; Maher et al., 2018).

When planetary rotation is included in RCE simulations, self-aggregation transforms into tropical cyclones. Aquaplanet simulations in RCE have been particularly useful for understanding tropical cyclone characteristics (Reed & Chavas, 2015; Satoh et al., 2016; Shi & Bretherton, 2014; Wing et al., 2016) and their response to increasing SSTs (Held & Zhao, 2008; Khairoutdinov & Emanuel, 2013; Merlis et al., 2016). Satoh et al. (2016) used a hierarchy of configurations with a global model to show the multiscale nature of tropical convective systems, and how rotation changes their vertical structure (see Figure 4).

The multiscale structure is also apparent in cloud resolving simulations of RCE. The emergent structures remain similar as the domain size is varied, but not their response to perturbations, such as imposed surface warming (Silvers et al., 2016). Similar experiments with global models are computationally expensive, but one alternative is to test the convergence characteristics of a model’s physics by reducing the planetary radius to increase the effective horizontal resolution; Reed and Medeiros (2016) use this strategy to show how the large-scale convective aggregation seen in GCMs transitions to CRM-like self-aggregation without increased computational cost.

Convective organization more generally is not well understood (Muller & Bony, 2015). For example, it is not clear how important self-aggregation is compared to organization by the mean circulation, or by waves or other mesoscale disturbances, or even the extent to which aggregation may be a model artifact. There are a number of factors that contribute to organization such as cloud radiative feedbacks, SST, and convective-moisture feedbacks, see Sessions et al. (2016) for a full list. Nonetheless, model hierarchies have provided insight into why convection organizes and how it is maintained. While RCE is an idealization for the moist atmosphere, it is still very complicated compared to the models used in the midlatitudes and middle atmosphere.

A further useful idealization, complementary to (and different from) RCE, is the WTG approximation. Under WTG the large-scale circulation, specifically the vertical velocity, is parameterized (Sobel & Bretherton, 2000; Sobel et al., 2001; Raymond & Zeng, 2005). This is done by assuming that horizontal temperature gradients and the local time tendency of temperature are both negligible at synoptic scales in the tropics—an observational fact explained dynamically by Charney (1963)—thus reducing the otherwise prognostic temperature equation to a diagnostic equation that can be solved for the large-scale vertical velocity given the diabatic heating. WTG is a horizontal truncation, as opposed to the vertical truncation in the Matsuno-Gill model described in section 5.1.

WTG has been used to study a range of phenomena, including the Walker and Hadley circulations (Bellon & Sobel, 2010; Bretherton & Sobel, 2002; Burns et al., 2006; Kuang, 2012; Polvani & Sobel, 2002), ENSO teleconnections (Chiang & Sobel, 2002), tropical cyclogenesis (Raymond, 2007), and the MJO (Wang et al., 2013, 2016). Other related parameterizations of large-scale dynamics, solving the same problem in different ways, have been developed (Kuang, 2008; Romps, 2012; Herman & Raymond, 2014), and WTG and the “damped wave” method (Blossey et al., 2009) have been applied to a wide range of models in a recent intercomparison (Daleu et al., 2015, 2016). These parameterizations of large-scale dynamics represent the circulation on scales smaller than the “global” scale at which RCE is relevant—the domain-average vertical motion being parameterized must vanish in RCE by definition. The domain of a WTG single-column or cloud-resolving simulation can be thought of as representing a small fraction of an RCE simulation’s domain.

In such WTG simulations, more than one statistical equilibrium state can occur, depending on the initial humidity, with either dry or persistent deep convection states developing from identical forcing conditions (Sessions et al., 2010; Sobel et al., 2007). These so-called “multiple equilibria” are analogous to self-aggregation in RCE simulations in which a convecting cluster is surrounded by dry subsiding air (Ses-
Figure 5. Extratropical cloud circulation coupling. The impact of clouds on the eddy-driven jet stream response to global warming in a hierarchy of GCMs. The zonal-mean time-mean change in 850-hPa zonal wind (m/s) for each latitude (°) for the ensemble mean (bold line) and individual models (gray) for (a) CMIP5 coupled Earth system models with 4 × CO₂ and (b) aquaplanet CMIP5 models with prescribed SST and 4 K SST warming. For The Max Planck Institute for Meteorology Earth System Model (MPI-ESM) in aquaplanet prescribed SST setup, simulations with the cloud-locking method and imposed global (black) and regional (colors) cloud changes show the cloud radiative contribution to the eddy-driven jet response to warming (c). The global and regional cloud impacts are reproduced in panel (d) using a dry Held-Suarez setup of the MPI-ESM model perturbed with the radiative forcing from cloud changes of panel (c). Because panels (b)–(d) are for aquaplanet simulations, only the Northern Hemisphere is shown. Note the different y scale in panel (d), which reflects the increased jet sensitivity of the Held-Suarez setup. Figure adapted from Voigt and Shaw (2016). GCMs = General Circulation Models; SST = Sea Surface Temperature.

While representing convective organization in ESMs remains problematic, progress is being made through a variety of modeling approaches to develop theories of convective organization and to better represent organization in GCMs. These approaches explore high-resolution large eddy simulations (LES) and CRMs and GCMs utilizing different treatments of convection. Approaches include attempts to resolve convection in global CRMs (Bretherton & Khairoutdinov, 2015; Judt, 2018; Miyamoto et al., 2013; Tomita et al., 2005), contrasting GCMs with and with parameterized convection (Maher et al., 2018; Popke et al., 2013), and the use of superparametrization (Arnold & Randall, 2015).

6.2. Decoupling Clouds and Circulation: Cloud Locking

Two key research questions for understanding and modeling clouds include the following: how do clouds couple to the large-scale circulation (discussed below) and what is the Earth’s ECS (discussed in section 6.3)? A primary challenge in representing clouds and convection in climate models is to adequately capture their interactions (which must be parameterized in global models) with radiation and the resolved circulation.

One opportunity to explore the role of clouds in the climate system is to adapt the diabatic hierarchy to decouple cloud radiative effects from the circulation in which they are embedded. A few different approaches have been developed to achieve this: (i) forcing dry GCMs with atmospheric cloud radiative tendencies simulated by comprehensive AGCMs (Li et al., 2019; Voigt & Shaw, 2016), (ii) the use of simplified physics (e.g., the gray radiation moist GCMs discussed in section 5.2; Kang et al., 2009), (iii) using clouds that are transparent to radiation (Stevens et al., 2012), and (iv) prescribing the cloud fields, often referred to as cloud locking (Zhang et al., 2010). For example, all four modeling approaches have been used to understand how changes in clouds with increased greenhouse gases will impact the position of the eddy-driven jet (Ceppi & Hartmann, 2016; Ceppi & Shepherd, 2017; Voigt & Shaw, 2015, 2016).

The cloud-locking model approach has proven particularly helpful to understand how changes in the radiative properties of clouds impact the circulation response to global warming or hemispheric energy perturbations. Cloud locking removes the coupling between clouds and circulation by prescribing the cloud properties seen by the model’s radiation scheme, generally from an earlier model simulation, which isolates the circulation response to a perturbation as the clouds are invariant (Zhang et al., 2010).

The eddy-driven jet (discussed in section 3.4) is an interesting example, as its equatorward bias in coupled GCMs (Kidston & Gerber, 2010) is associated with Southern Ocean clouds that reflect too little shortwave radiation in models (Ceppi et al., 2012). Coupled GCMs exhibit diverse responses of the eddy-driven jet to global warming, especially in the Southern Hemisphere; see Figure 5a. These broad differences persist...
in aquaplanet simulations (Figure 5b), making aquaplanets a desirable configuration to understand the eddy-driven jet response.

Cloud radiative changes lead to a poleward shift in the eddy-driven jet in cloud-locking simulations for the Max Planck Institute for Meteorology Earth System Model (MPI-ESM) aquaplanet model (Figure 5c). The cloud radiative changes with global warming can be attributed to high-level tropical (orange line) and mid-latitude clouds (blue line). Interestingly, the cloud impact is as large as the differences in jet shifts found in coupled GCMs, which suggests that clouds contribute to uncertainty in future jet shifts. The cloud impact is also reproduced in the dry Held-Suarez simulations perturbed with radiative changes from the cloud-locking simulations (Figure 5d).

A complementary modeling technique to cloud locking is the transparent-cloud approach, that prevents the radiation scheme from “seeing” the clouds and hence sets the radiative heating to cloud-free conditions (Merlis, 2015; Randall et al., 1989). This is easier to implement than cloud locking and is simply achieved by setting the cloud fraction to zero in the radiation scheme. The transparent-cloud approach has helped to demonstrate the importance of cloud radiative effects for the present-day circulation. Such simulations have shown that cloud radiative effects strengthen the Hadley cell and eddy-driven jet stream, reduce tropical-mean precipitation, and narrow the ITCZ (Albern et al., 2018; Harrop & Hartmann, 2016; Li et al., 2015; Popp & Silvers, 2017).

The primary task for understanding the role clouds play in the climate system is to understand their coupling with the circulation and the implications of that coupling for the circulation response to climate change. In this regard, the transparent-cloud approach has proven helpful for understanding the role of clouds in the present-day climate, and the cloud-locking approach for understanding changes in clouds and circulation with global warming. While recent work has clearly shown that a quantitative understanding of the circulation must consider the coupling to clouds, this remains a rather young area of research with many open research questions, including, for example, cloud impacts on the internal variability of the extratropical circulation (Li et al., 2014).

6.3. The Role of Circulation in Earth’s Equilibrium Climate Sensitivity

Understanding how clouds impact the radiative budget, and in turn the circulation, is critical for understanding how increased greenhouse gas emissions will change the climate, that is, how sensitive the climate system is to greenhouse gas emissions. Clouds are at the heart of this question as cloud feedbacks are the largest uncertainty in the ECS—a measure of globally averaged surface temperature change to doubling CO₂. Despite significant improvements in climate models, our estimate of the ECS has not changed appreciably since the Charney report in 1979 suggested a range of 1.5–4.5 K (Stevens et al., 2016).

The representation of clouds in different climate models is diverse, and this results in widely varying cloud responses to the same perturbation (Boucher et al., 2013; Chung & Soden, 2018). Climate models show a relatively robust positive longwave (infrared/greenhouse) cloud feedback (Zelinka & Hartmann, 2010), understood through the fixed anvil temperature hypothesis (Hartmann et al., 2001; Hartmann & Larson, 2002). The shortwave (visible/albedo) cloud feedbacks, however, remains highly uncertain, though most coupled GCMs suggest a weak positive feedback (Ceppi et al., 2017). Answering the open research questions about the ECS comes down to understanding shortwave feedbacks for low-level clouds, which account for much of the model uncertainty in cloud feedbacks. These low-level clouds form below regions of radiative cooling in the descending branches of the Hadley and Walker circulations (Bony & Emanuel, 2005). As such, circulation is key in setting their distribution, as cloud effects also feedback on the circulation.

Single-column models (SCMs) have been used to investigate how parameterized physics impact the climate sensitivity (Dal Gesso et al., 2015). Using SCMs with several configurations, Zhang et al. (2013) showed that the shallow convection and boundary layer turbulence are key differences among models. Care must be taken to meaningfully comparing an SCM to a GCM, however, because of the disconnection of cloud-circulation coupling in SCMs. In addition, physics packages can exhibit different cloud responses in a GCM and SCMs. Progress has been made in understanding cloud feedbacks in the gap between SCMs and GCMs, such as using the WTG approximation to parameterize a circulation in SCMs (Raymond, 2007; Zhu & Sobel, 2012). An alternative approach, again building on the work of Manabe and Strickler (1964), is to run a GCM in RCE to simplify the circulation (Bony et al., 2016; Popke et al., 2013; Wing et al., 2017b).
To capture the impact of circulation on climate sensitivity, efforts have focused on the complex end of the model hierarchy: coupled AOGCMs (Caldwell et al., 2016; Otto et al., 2013; Stevens et al., 2016). Simpler models do not capture all the relevant processes, removing nonlinear behavior that influences the climate sensitivity (Knutti & Rugenstein, 2015). From the perspective of the model hierarchy, AOGCMs are a moving target that evolves in response to both improvements in our understanding of the climate system and to increasing computational resources.

The complexity of modern climate models, however, make it challenging to interpret their results, including the relative role of cloud feedbacks in climate change. The challenges in understanding climate sensitivity in AOGCMs makes a hierarchical approach appealing. The goal then becomes understanding the response of state-of-the-art AOGCMs in a simpler setting to reveal the underlying mechanisms and improve our physical understanding of the system. For example, using a range of boundary conditions and model configurations (ESM, GCM, aquaplanet, and SCM) with the same model parameterizations, Brient and Bony (2013) identified a positive feedback that depends on how moist static energy is transported between the free troposphere and the boundary layer. Additional progress has been made using aquaplanet simulations to identify shallow cumulus clouds as driving the spread in climate sensitivity (Medeiros et al., 2008; Medeiros et al., 2015; Ringer et al., 2014).

Progress is being made using model hierarchies to improve our understanding of how clouds impact the global circulation and the ECS. In section 7 we focus our discussion on the MJO to illustrate how the model hierarchies can be utilized to improve both our understanding of a phenomena and also improve the realism of climate models simulations.

7. Case Study: The Madden-Julian Oscillation

In sections 5 and 6 we described models that have been fundamental for advancing our understanding of tropical circulation and the important role that moisture plays in setting the circulation, specifically how convective organization and clouds impact the radiative structure of the atmosphere. In this section we focus on the MJO, which provides an example where a hierarchical modeling approach aided both our understanding of and ability to simulate an important mode of tropical variability.

The MJO is an organized convective system and the primary source of tropical intraseasonal variability. It is an envelope of organized tropical convection that drifts eastward from the Indian Ocean into the Pacific. It is distinct from most convectively coupled equatorial waves in having a relatively slow speed of propagation (≈4–8 m/s), longer time scales (about 1–2 months), and a relatively large scale (planetary wave numbers 1–3) in comparison to other synoptic disturbances in the tropics.

The MJO continues to challenge our understanding of how circulation couples to clouds, convection, and radiation. Progress is being made in our theoretical understanding of the mechanisms that initiate, propagate and maintain the MJO; however, there is currently no complete theory for the MJO (Ahn et al., 2017). While it has historically been difficult to simulate in global models, in the last decade some dynamical forecasts have become superior to statistical forecasts (e.g., Kang & Kim, 2010), and this new simulation capability allows theoretical ideas to be tested.

Realistic simulations of the MJO require convection to be sensitive to free-tropospheric moisture, that is, a positive moisture convection feedback, where deep convection is favored in regions where the free-tropospheric humidity is larger. CMIP5-class models with the largest moisture sensitivity tend to have the most realistic MJO (Kim et al., 2014). Models that poorly simulate the MJO—generally those with weak to nonexistent MJOs (Ahn et al., 2017)—can be improved by increasing the sensitivity of convection to moisture, such as increasing the entrainment and rain reevaporation. Such tuning to optimize the MJO generally causes biases in mean climate (e.g., Kim et al., 2011), but there is encouraging evidence to suggest a realistic MJO, and mean state can be simultaneously captured with traditional convection schemes (Crueger et al., 2013). There is considerable additional evidence, apart from the MJO, that deep convection in general is quite sensitive to moisture (e.g., Derbyshire et al., 2004), and that typical convective schemes have excessive undilute ascent, as opposed to entraining air about them (e.g., Kuang & Bretherton, 2006; Tokioka et al., 1988).

More recent studies have viewed the MJO through the moist static energy budget, where surface fluxes and radiation are the dominant source terms. (Recall that the moist static energy is conserved under conden-
Figure 6. The connection between simple models of the atmosphere and the comprehensive models used for weather and climate prediction. This figure complements Figure 2, illustrating the connections to atmospheric General Circulation Models (AGCMs) afforded by model hierarchies. Each arm illustrates a different hierarchy: a dynamical hierarchy in terms of the equations, process hierarchies in terms of the boundary conditions (the representation of the ocean) and treatment of diabatic processes, and a hierarchy of scale, focused on convective organization across very different domain sizes and resolutions.

Sensation, while latent heating is the dominant source term in the dry static energy budget in the presence of deep convection. Feedbacks between surface turbulent fluxes and convection were emphasized in early theories (Emanuel, 1987; Neelin et al., 1987) and appear to be important in some GCMs (e.g., Maloney & Sobel, 2004). Other work, however, points to a key role for cloud radiative feedbacks; for example, there is less longwave cooling by high clouds in a moist atmosphere (Andersen & Kuang, 2012; Chikira, 2013). Process-based diagnostics (Kim et al., 2015) and so-called “mechanism denial” experiments (Crueger & Stevens, 2015; Kim et al., 2012; Ma & Kuang, 2016) in which a process is removed in order to test its importance have also lead to progress. This methodology is related to earlier work with more idealized models, for mechanism denial is akin to building a hierarchy from the top down, by systematically eliminating processes from a comprehensive model. Raymond (2001) argued that radiative feedbacks were important to the MJO based on results from a 3-D model of intermediate complexity, while Bony and Emanuel (2005) did so based on 2-D CRM simulations without rotation. In an even simpler context, Hu and Randall (1994) found that radiative feedbacks are critical in a one-dimensional model without large-scale circulation.

The importance of moisture convection feedbacks and cloud radiative feedbacks suggests a view of the MJO as essentially a form of self-aggregation on the equatorial $\beta$ plane, in a domain much larger than CRMs simulations (e.g., Arnold & Randall, 2015). In aquaplanet simulations with superparameterized convection in RCE, Arnold and Randall (2015) found similar energy budgets and radiative feedbacks in nonrotating simulations, where self-aggregation dominates and simulations with rotation where MJO-like variability occurs.

The importance of moisture convection and cloud radiative feedbacks are the core assumptions in a recent set of highly idealized models of the MJO. These models represent the MJO as a moisture mode—a mode that would be absent in a dry atmosphere. In these idealized models, essential information is contained in the moisture field. Truncation to a single vertical mode, as in the Matsuno-Gill model, allows the dry dynamics...
to become shallow water like. The convection schemes depend strongly, and in some cases exclusively, on
the moisture field, building in a strong moisture convection feedback.

Moisture modes emerged in the idealized models of Fuchs and Raymond (Fuchs & Raymond, 2002, 2007;
Raymond & Fuchs, 2007, 2009). The moisture mode was isolated in the simple 1-D linear model of Sobel
and Maloney (2012, 2013) that has a single moisture prognostic variable, assumes WTG in the temperature
equation, and generates winds by assuming a Matsuno-Gill response to quasi-steady heating (approximately
valid as long as the disturbance does not propagate too quickly). In this model it can be shown explicitly
that radiative feedbacks are critical for eastward propagation in a linearly unstable mode (Sobel & Maloney,
2013). While the eastward propagation was initially slower than observations, modifications by Adames and
Kim (2016) increased the propagation speed by accounting for meridional moisture advection. Because the
WTG assumption eliminates the Kelvin waves, the waves that most early theories relied on to explain the
eastward propagation, the propagation of a moisture mode results largely from horizontal moisture advec-
tion, which seems to be supported by a number of observational and modeling studies (e.g., Inoue & Back,
2015a; Kim et al., 2014; Maloney, 2009; Pritchard & Bretherton, 2014).

Moisture mode theory—including the link to self-aggregation in idealized simulations—provides a useful
framework for diagnosing models and observations, although whether moisture mode models correctly cap-
ture the MJO remains a topic of debate. The moisture mode ideas are quite different from those in earlier
MJO theories, most of which excluded both radiative feedbacks and prognostic moisture (e.g., see the review
by Wang, 2005), and also differ from other, more recent models (e.g., Majda & Stechmann, 2009; Yang &
Ingersoll, 2013). Now that some comprehensive models at the top of the model hierarchy can simulate the
MJO with reasonable fidelity, it is a question of linking them to our theories of MJO behavior. A connection to
the moisture mode hypothesis, for example, can be traced through a hierarchical chain from self-aggregation
in idealized simulations to more realistic simulations where moisture convection and radiative feedbacks
are allowed.

8. Synthesis and Outlook

All models are wrong but some are useful. In coining this phrase at a workshop on statistical robustness four
decades ago (Box, 1978), the statistician George Box succinctly made two points. First is the reminder that all
of our models, even the most sophisticated, are inherently simplified—and so in Box’s sense “wrong”—and
thus unable to capture all the potentially relevant processes and scales of the climate system. But second,
we can learn, understand, and make predictions with some models.

In this review, we have identified a number of deliberately simplified models that have proven useful
for understanding and predicting the large-scale circulation of the atmosphere. We have not identified
all possible benchmark models, but have sought to provide a balanced view of the dynamics of the mid-
latitude, middle atmosphere, and tropics. In doing so, we have highlighted processes on scales large, for
example, planetary waves in the Holton-Mass model, to small, for example, convection and clouds within
radiative-convective equilibrium simulations.

In section 2, we proposed three principles to help organize models into hierarchies: dynamics, process,
and scale. These are motivated, in part, by decisions we make in order to create a numerical model of the
atmosphere that captures the large-scale circulation. These decisions include the appropriate governing
equations, the relevant processes that drive the circulation, and the domain and resolution (which deter-
mine the allowable scales). These principles are not independent of one another. Dynamical hierarchies are
designed to isolate particular scales and processes, for example, the quasi-geostrophic equations focus on
Rossby wave processes by filtering out the faster and smaller-scale gravity waves. Likewise, the process hier-
archy influence the choice of dynamics; if we wish to look at nonhydrostatic effects, we must resolve scales
with an order-one aspect ratio, and thus the kilometer scale.

The models featured in sections 3–7 (and introduced schematically in Figure 2) provide several examples
of each of these hierarchies that have emerged organically in the literature. Held (2005) and Held (2014)
emphasize that our simpler models are only valuable in so much as they can be connected with the more
complex models necessary for quantitative prediction. In Figure6, we more explicitly connect these hier-
archies to the state-of-the-art AGCMs used for weather and climate prediction. Their connectedness is
essential; indeed it is what defines a hierarchy. Connectedness does not always need to occur in a sequence
Dynamical hierarchies have played a key role in understanding the midlatitude circulation, where fast rotation and stratification organize the flow. We define an equation hierarchy, see Figure 6, that forms a natural progression of the equation set. The equation hierarchy includes the (i) barotropic vorticity dynamics that capture the evolution of Rossby waves (section 3.1), (ii) quasi-geostrophic flow on two or more layers to capture baroclinic instability (section 3.2), (iii) the dry primitive equation dynamics, for example, as represented in the the Held-Suarez model (sections 3.2, 4.1, and 6.2), (iv) the moist primitive equation dynamics (as in the Frierson model in section 5.2 or a standard AGCM), and ending with (v) the nonhydrostatic equations that includes the vertical momentum equation and are accurate at higher horizontal resolutions, for example, used in CRMs or weather forecast models.

In the tropics, rotation is weak and moist processes are of first-order importance. As such, the dynamical hierarchies generally only include the more complex end of the equation hierarchy. Nonetheless, the primitive equations or nonhydrostatic dynamics can be used with either vertical truncation (the Matsuno-Gill model in section 5.1) or horizontal truncation (the WTG approximation in section 6.1) to simplify the equation set.

The focus on processes is most essential for organizing model hierarchies. The purpose of dynamics and scale hierarchies are to isolate and resolve the processes of interest. An example of a process hierarchy is
the diabatic hierarchy (Figure 6), a term we use to describe a series of GCMs that integrate the primitive equation dynamics on the sphere, with advancing representations of the processes driving the temperature equation and generally with a resolution on the order of 100 km. At the base of the diabatic hierarchy are (i) models utilizing little more than a dry dynamical core, with all diabatic processes replaced by Newtonian temperature relaxation as exemplified by the model of Held and Suarez (1994). The Held-Suarez model has been used to understand jet stream variability (section 3.4), tropical overturning circulation (section 5.2), stratosphere-troposphere coupling (section 4.1), and tracer transport (section 4.3).

The next step in the diabatic hierarchy (ii) is to add moisture, but paired with a gray radiation scheme to decouple the role of water vapor in latent heat transport from its impact on radiation (e.g., Frierson et al., 2006; O’Gorman & Schneider, 2008; section 5.2). At the next level, (iii) water vapor interacts with radiation by including a more complete representation of radiation, but with a prescribed cloud climatology (or no clouds at all), thus removing the key role of cloud microphysics on radiative transfer (e.g., Merlis et al., 2013; Jucker & Gerber, 2017). These models have elucidated the circulation of the tropics and coupling between high and low latitudes (section 5.2).

Moving further up the diabatic hierarchy, (iv) AGCMs account for the importance of cloud and aerosol processes in the diabatic forcing of the circulation (section 6). The most complex end of the diabatic hierarchy is to include (v) the carbon cycle and fully interactive chemistry, which enables a more realistic representation of the processes governing radiative gases, clouds, and aerosols. The complex end of the diabatic hierarchy continues to evolve with time, as more processes are included in Earth System models and computational resources grow.

Another process hierarchy that helps to organize the model hierarchies focuses the lower boundary conditions (Figure 6). Atmospheric models can be created with oceans that have (i) uniform SST, as used in WTG simulations, (ii) prescribed but spatially varying SST, (iii) aquaplanets with a so-called slab ocean which only captures the local thermodynamics of the atmosphere-ocean coupling, and (iv) a slab ocean with q fluxes which include idealized horizontal ocean transport for more realistic atmospheric circulation.

In this review we have not discussed the slab ocean model in much detail, but we refer the reader to Clement et al. (2011) and Clement et al. (2015) for further discussion. The prescribed SST and slab ocean models (without or without q fluxes) can also be configured to have idealized land and topography by changing the heat capacity and boundary layer roughness. The representation of the land surface conditions can be idealized or more realistic, for example, bucket hydrology versus water runoff or a full representation of vegetation. At the most complex end of the hierarchy is an atmospheric model paired with an interactive ocean and land surface to form a coupled atmosphere-ocean model (AOGCM).

Figure 7 illustrates a hierarchy available within the Community Earth System Model (CESM) framework, incorporating elements of both the diabatic hierarchy and varying configurations of the land surface. The SimpleER project (Polvani et al., 2017) makes many of these models an integral part of the CESM structure. The Isca framework (Vallis et al., 2018), using software from Geophysical Fluid Dynamics Laboratory (GFDL) modeling system, both include many of the lower steps of the hierarchy and hooks to add complexity and build models of other planetary atmospheres as needed. One aspect of the process hierarchy that moves beyond these GCMs is to include more comprehensive treatments of microphysical processes that determine the distribution of clouds. An example is the Weather Research and Forecasting (WRF) model, which offers different options for representation of atmospheric processes, such as microphysics, and the treatment of boundary conditions.

Our final principal for organizing the models is scale. One example in which a hierarchy has naturally developed is for studying convective organization (section 6.1). As depicted in Figure 6, the model domain can vary from very high resolution in a small domain (to understand in-cloud properties) to lower resolution on a global scale to understand planetary scale organization such as the MJO. Scale hierarchies are also implicit in dynamical hierarchies. Simplified models have proven useful in problems that intrinsically involve a large spread in scales, such as the QBO, where the evolution of planetary scale jets is driven by small-scale gravity waves and could only recently be captured in AGCMs.

We often think of the model hierarchies from a bottom-up perspective, building up from the most idealized models to the most complex, as we have done in this paper. In section 7 we described how mechanism denial experiments have lead to progress in understanding the MJO. These mechanism denial experiments can be
thought of as a top-down model hierarchy, starting with comprehensive model and removing complexity in order to attribute key processes.

Looking forward, we believe that model hierarchies will continue to improve our understanding of the atmosphere and climate. In particular, the gap in our understanding of the processes coupling clouds, convection, and circulation is mirrored in part with a gap in simple models that isolate the key processes regulating these interactions. There is a large jump between idealized moist models that effectively neglect cloud-aerosol processes (e.g., Frierson et al., 2006; Jucker & Gerber, 2017; Merlis et al., 2013) and comprehensive GCMs that seek to parameterize all the unresolved scales which are critical to clouds and aerosols. This gap partially reflects the difference between what can be done by an individual research group and a full modeling center. Further development of simpler GCMs that capture the essential elements of cloud and aerosol interactions are needed. It requires identifying sufficiently elegant models, in the language of Held (2014), that would merit investment by a modeling center or consortium of research groups to bring in sufficient expertise.

Radiative-convective equilibrium simulations are also aimed at filling this gulf between large-scale dynamics and clouds processes. There is still a fundamental separation between them and the real atmosphere, however, where wind shear plays a vital role in organizing convection. This gap is visually emphasized in Figure 7 by the profound changes in circulation in CESM between simulations in RCE and an aquaplanet (where the large-scale flow is determined by rotation and the temperature). Adding the building blocks of rotation and shear into RCE simulations may help establish these links.

Model hierarchies will continue to play a role in our understanding of climate projections. In fact, we argue that they should play an increasingly important role. Our confidence in global warming projections is not blind faith in a GCMs output; rather, it is fundamentally supported by basic physical laws. However, in their simplest manifestation those laws have little quantitative predictive capability for Earth’s climate. At the other extreme, when comprehensive models are forced into the warmer regimes that may lie in our planet’s future, we do not have the ability to compare parametrizations with observations. A purpose of the model hierarchy is to provide a pathway connecting robust physical laws to our complex reality, via models of varying levels of complexity. Ideally, this enables us to both understand the processes involved and to make useful and trustworthy predictions.

References


