CONVECTIVE QUASI–EQUILIBRIUM

J.–I. Yano,[1] and R. S. Plant[2]

R. S. Plant, Department of Meteorology, University of Reading, UK. (r.s.plant@reading.ac.uk)

J.–I. Yano, GAME/CNRM, Météo-France and CNRS, 31057 Toulouse Cedex, France. (jun-ichi.yano@zmaw.de)

[1]GAME/CNRM, Météo-France and CNRS, 31057 Toulouse Cedex, France.
[2]Department of Meteorology, University of Reading, UK.

The concept of convective quasi–equilibrium (CQE) is a key ingredient in order to understand the role of deep moist convection in the atmosphere. It has been used as a guiding principle to develop almost all convective parameterizations and provides a basic theoretical framework for large–scale tropical dynamics. The CQE concept as originally proposed by *Arakawa and Schubert* [1974] is systematically reviewed from wider perspectives. Various interpretations and extensions of Arakawa and Schubert's CQE are considered in terms of both a thermodynamic analogy and as a dynamical balance. The thermodynamic interpretations can be more emphatically embraced as a *homeostasis*. The dynamic balance interpretations can be best understood by analogy with the slow manifold. Various criticisms of CQE can be avoided by taking the dynamic balance interpretation. Possible limits of CQE are also discussed, including the importance of triggering in many convective situations, as well as the possible self–organized criticality of tropical convection. However, the most intriguing aspect of the CQE concept is that, in spite of many observational tests supporting and interpreting it in many different senses, it has never been established in a robust manner based on a systematic analysis of the cloud–work function budget by observations as was originally defined.

1. INTRODUCTION

The concept of "equilibrium" is used in various contexts in atmospheric sciences. It may be needless to emphasize at the outset a difficulty in strictly applying the concept in its thermodynamic sense to the climate system: the climate system is always evolving and so is never perfectly in equilibrium. Note that a system is in perfect thermodynamic equilibrium only if it is steady and motionless (see section 2.3 for more). Thus, the concept of equilibrium has only limited applicability in the climate system. In other words, the concept is strictly applicable only under certain approximations or idealized settings. This limitation makes the concept of equilibrium somewhat subtle in the atmospheric sciences.

In order to expand its applicability, the concept of a "quasi–equilibrium" is often introduced. Probably the best-known example is the *convective quasi–equilibrium* (CQE) originally introduced by *Arakawa and Schubert* [1974]. The concept is especially important as a key principle for "closing" convection parameterization, as originally proposed by *Arakawa and Schubert* [1974]. A series of review papers by *Emanuel et al.* [1994], *Emanuel* [2000, 2007] furthermore emphasizes its importance for constructing a theory of tropical large–scale circulations. However, the interpretation and range of validity of CQE remains controversial in convection dynamics as well as in tropical meteorology. Thus the present contribution focuses on a critical analysis of the concept itself.

Loosely speaking, the notion behind *Arakawa and Schubert's* [1974] CQE is that convection is almost in "equilibrium" with the large–scale, non-convective processes. The main goal of the present review is to re–examine the use of this terminology in the light of the original ideas of equilibrium in both thermodynamics and dynamics. For this reason, we perform a historical review of the concept of equilibrium and related issues. We believe such a historical approach is crucial in order to settle various current confusions concerning CQE. Especially, interpretations of CQE have multiplied over the years, and new interpretations continue to be propounded. Unfortunately, authors are not always careful about stating precisely what interpretation is intended when invoking CQE in a particular study. A broad historical context provides the most powerful framework in order to settle such a situation, allowing us to organize and compare the various interpretations.

The present review also addresses the extent to which the atmosphere can be treated as being at CQE. This is both a practical and a conceptual issue. The latter aspect is emphasized here, because without clarifying exactly what the question is, we cannot talk about its practical implications. Various related scientific issues as well as methodologies potentially applicable for addressing CQE are considered. The review attempts to be general and abstract, rather than being specific and practical, because our goal is to suggest directions for investigations on CQE for years to come.

We begin by examining the concept of thermodynamic equilibrium. In particular, in section 2 the concept is contrasted with that of *balance* in mechanics. The two concepts are closely related but not the same. Thus, a natural question to ask is: which of these do we really invoke in the context of CQE? Section 3 closely examines the concept of CQE in the form originally introduced by *Arakawa and Schubert* [1974]. Section 4 discusses various interpretations of CQE as well as some related concepts. Section 5 examines some of the criticisms of CQE and section 6 discusses various alternative paradigms. The review is concluded in section 7.

2. WHAT IS EQUILIBRIUM?

2.1. Etymology

The etymology for "equilibrium" is found in Latin *aequilibrium*, which consists of the two stems, *aequi–* and *libera*, meaning weight and balance, respectively. Thus, etymologically speaking, "equilibrium" suggests *weighted balance*. In this respect, the concept of "equilibrium" is very close to "balance", but the former implies something more than the latter. However, it is curious to note that French, for example, has only one word *équilibre* for referring to both "equilibrium" and "balance".

2.2. Thermodynamics and Mechanics

Before the advent of quantum mechanics and relativity, classical physics could be divided into two distinct disciplines: thermodynamics and mechanics (dynamics). Classical thermodynamics is inherently interested with equilibrium. For this reason, in order to understand what *equilibrium* is, we have first to understand something about thermodynamics.

Arguably the most successful application of classical thermodynamics is to the heat engine problem. In this problem a heat engine cycle is defined as progressing through a sequence of equilibrium states, with each transition being necessitated as a response to an external condition that is imposed on the engine. For example, the heat engine may be subjected to an externally-specified temperature which is changed sequentially in discrete steps. On each occasion that the external condition is changed, a new equilibrium state of the system is calculated, the state of interest being that which would be reached after waiting for an indefinitely long time. The sequencing of external conditions is continued in such a way as to progress through a closed cycle of such states. Then the obtained "useful" energy (called "work") is evaluated, this quantity being the main interest of the heat engine problem.

In performing a heat engine calculation, the question of how long it takes to complete the cycle is not considered. How fast one can turn the engine is clearly a question of enormous practical importance, but a classical heat engine problem (such as the Carnot cycle) does not pose the question. In other words, in classical thermodynamics the concept of *time* remains implicit.

By contrast, the main interest of classical mechanics is the time evolution of a given system. Perhaps the greatest success of classical mechanics is in explaining the movements of planets and moons in the solar system. The degree of success in such an application is judged by the predictability: for example, precise timings of solar and lunar eclipses.

Arguably, the concept of thermodynamic equilibrium as just described is inherently foreign to mechanics, because motion is inherent to mechanics. Notions of equilibrium, however, began to play an important role in the development of mechanics in the 18th century [*Hepburn*, 2007] through a French school, perhaps most notably by d'Alembert, which led to the discovery of a variational principle by Lagrange (*cf.*, section 2.4 below).

It seems fair to say that modern meteorology, originating from the Bergen school led by Vihelm Bjerkness in the late 19th century, developed as an outgrowth of classical mechanics with its main interest being the prediction of weather. For this reason, one might even argue that the notion of equilibrium is inherently foreign to modern meteorology. However, before we consider that argument, we must first discuss further the thermodynamic notion of equilibrium.

2.3. Equilibrium as a Thermodynamic Concept

As an example of a process leading to thermodynamic equilibrium, take two boxes (with the same heat capacity, for simplicity) with two different temperatures, T_1 and T_2. We bring these two boxes into contact (but with no interactions with the outside world: *i.e.*, we assume that interactions happen only between the boxes). Heat is exchanged between the boxes in affecting a transition towards an equilibrium. More specifically, heat is transferred from the box with higher temperature to the box with lower temperature and the process continues until the temperatures of the two boxes are equal. In this case, the final temperature is the mean, $(T_1 + T_2)/2$, of the original temperatures and the final state is called an equilibrium.

The example just given, albeit an extremely simple one, is a typical problem considered in thermodynamics: define an equilibrium thermodynamic state of a system in contact with an external system, under given constraints (*e.g.*, the volume remains constant, and the initial temperatures are given). In the example, one of the boxes is the system of interest, and the other box is an external system.

In many cases the external system is considered to be much larger than the system of interest (such as in the heat engine problem). It is sometimes called a *thermal bath* when the temperature is fixed, and in that case the final equilibrium temperature of the system would be the same as the external temperature.

Carnot's Reflections [1824] are widely recognized as a key foundation of thermodynamics, although they had to be more carefully formulated and rescued from obscurity by *Clapeyron* [1834]. Interestingly, *Clapeyron's* article [1834] does not speak of equilibrium although *Carnot* [1824] uses the concept frequently without ever defining quite what he means by it. The closest Clapeyron comes is in the description "an equilibrium considered as destroyed by any cause whatever, by chemical action such as combustion, or by any other." This, and the broader sense of his article, is certainly consistent with the interpretation given above.

Importantly, the thermodynamic equilibrium is always a motionless state, except for an obvious situation such as a uniform translation. From a thermodynamic point of view, dissipation plays a key role when dealing with motion. For example, if a coffee cup is stirred then a rotating flow arises, and persists until dissipated away by friction so that the system recovers a quiescent state of equilibrium. This is distinct from mechanics, which prefers to take a dissipationless limit for describing many motions (more on this in section 2.8). We may even perpetualize motion, for example by imposing a temperature gradient on a given system, either horizontally or vertically. However, from a thermodynamic point of view, this simply forces the system perpetually away from thermodynamic equilibrium.

As the discussion just above suggests, application of a thermodynamic equilibrium to any real physical system is necessarily approximate. Therefore, while time invariance of macroscopic properties may be a useful practical description of a system in a state of statistical equilibrium it does not provide an appropriate definition of that state, since the macroscopic properties will undergo fluctuations for any system of finite extent. Rather, in order to define the statistical equilibrium state for a finite system, one may consider that "a system is in a condition of equilibrium when the information one has about it has reached a time-independent minimum" [*Andrews*, 1963, p24]. In such a condition one can derive probability distributions from a variational principle (*cf.*, section 2.4) such that the extent of the fluctuations becomes a known aspect of the equilibrium state. In this way, a time-independence remains inherent to the concept of equilibrium.

It is often desirable to try to apply the equilibrium concept without necessarily waiting a very long time for a complete adjustment to a state of time-independence. As discussed by *Landau and Lifshitz* [1980] for instance, complete adjustment can be a particularly awkward restriction of the concept in practice because the time–scale for such adjustment will increase with the system size. It is therefore often necessary to invoke a partial equilibrium in which the probability distributions of observables take functional forms derived from the limit of a complete statistical equilibrium but containing parameters that are considered to change slowly in space and time. The notion of partial equilibrium can also be applied in a local sense, considering that thermodynamic equilibrium is established at any individual macroscopic point, although the system as a whole (such as the climate) never reaches a thermodynamic equilibrium because it is perpetually perturbed by external forces. The need for such a local equilibrium concept to derive the standard equation set for geophysical flow is made explicit (for example) in Ch. 1 of *Salmon* [1998].

2.4. Variational Principle

The main goal of thermodynamics is to describe equilibrium states in as general a manner as possible. For this purpose, a variational principle is often invoked, in which the thermodynamic equilibrium is defined by a minimization (or maximization) of a certain thermodynamic quantity (potential). For the example of joining two boxes in the previous subsection, the equilibrium state can be deduced by maximizing the entropy. The textbook by *Kondepudi and Prigogine* [1998] emphasizes the variational principle. A more concise description of the principle can be found, for example, in section 7.4 of *Adkins* [1983], and in chapter 2 of *Chandler* [1987].

The strength of the variational principle is that it enables one to determine a thermodynamic equilibrium state in many situations with a few general rules. Due to its generality, this principle can also be applied to many other physical problems. In other words, the concept of equilibrium can be generalized by invoking the variational principle. Indeed, it was a consideration of equilibrium in its broadest sense that led Maupertius, Euler, Lagrange and others to introduce the variational principle into classical mechanics [*Hepburn*, 2007], contemperaneously with Johann Bernoulli's characterization of static equilibrium through an optimality principle [*Hildebrant and Tromba*, 1996]. Ch. 5 of *Salmon* [1998] reviews applications of the variational principle to geophysical flows.

Properties of the convective equilibrium state may also be determined using a variational principle. Although a different approach was taken in their original article, the Boltzmann-like distribution of mass flux per cloud derived by *Craig and Cohen* [2006] can easily be re-derived from a maximum entropy condition, as shown in the Appendix A. We believe that the variational principle can potentially be used also for determining other aspects of atmospheric convective systems, including the mean equilibrium state (*e.g.*, total convective mass flux). Nevertheless, much investigation is still awaited.

However, a major limitation of the variational principle in thermodynamics is that it restricts attention to the most likely state of a given system. In other words, the problem is formulated in such a way that a state of maximum probability is sought, often equated with a state of maximum entropy. In many cases, applications of a variational principle may require an assumption that a large number of ensemble members can be associated with the system of interest. Clearly the maximum entropy principle makes this assumption: only if the ensemble size is large enough can we assume that the most likely state will correspond to actual realizations (*cf.*, section 2.7).

2.5. Why Equilibrium is a Useful Concept

The great strength of classical thermodynamics resides in excluding the need for consideration of the time evolution of a system by focussing on the final, equilibrium state of a given system under given constraints. Within this framework, even a full initial condition of the system is not important so long as the constraints are well specified. In the above example with the two boxes, it suffices merely to specify the initial temperatures of the boxes, and there is no need to ask further questions, such as the initial positions of the boxes, their volumes, etc.

This approach greatly facilitates the computations by focussing only on the final state, and removing the need to consider a complicated initial value problem for the complete evolution. In many practical applications, it is indeed only the final state that is of interest, especially if the given system approaches an equilibrium rapidly. The heat engine problem is a good example for making this point: in many practical engineering applications, a system can be considered to reach equilibrium almost instantaneously whenever an operation is performed. The transient adjustment process is not of interest.

The geostrophic adjustment problem originally developed by *Rossby* [1938] can be considered as a similar approach based on this concept of equilibrium. The problem asks for the final state of a flow under a geostrophic constraint, starting from an initial non-geostrophic state (assumed to be a state of rest in Rossby's original problem) with a height anomaly (recall that Rossby's original problem considers a shallow–water system). By invoking a principle of potential–vorticity conservation, the final state under geostrophy can be determined without solving a complicated initial value problem.

Conceptually speaking, a synoptic-scale weather system can be considered to evolve under a continuous sequence of geostrophic adjustments. The process of geostrophic adjustment is fast (say, a few hours) relative to a typical time–scale (say, a day and beyond) of interest in traditional synoptic weather forecasts. Hence there may be no need to consider explicitly the details of the adjustment process. This reasoning essentially leads to the idea of quasi–geostrophy. To some extent, the concept of convective quasi–equilibrium can be interpreted in a similar manner, as discussed later in section 4.6.

Nevertheless, a limitation of the "geostrophic adjustment" concept should also be emphasized: it is not necessarily consistent with the cascade point of view based on geostrophic turbulence [Maarten Ambaum, personal communication, 2010]. Certain issues beyond this metaphor of reality are discussed in section 4.7.

2.6. Macroscopic and Microscopic Views

The macroscopic point of view of thermodynamics may be contrasted with the microscopic view from statistical mechanics. A goal of statistical mechanics is, presumably, that of deriving and even of proving the principles of thermodynamics, originally derived in a completely empirical phenomenological manner, from more fundamental principles of atomic theory. Note that we have added the qualifier "presumably" in order to be cautious about this point of view.

In contemporary standard physics education, atomic theory is often considered to be more fundamental than macroscopic phenomenologies. Convective quasi–equilibrium can also be interpreted from a perspective of statistical mechanics. For this reason, a review by *Emanuel* [2000] emphasizes the importance of "statistical equilibrium thinking" in order to understand properly convective quasi–equilibrium.

However, this perspective has not always been accepted as common wisdom in the history of science. Just recall the philosophical debates between Ernest Mach, Stephan Boltzmann and others a little more than hundred years ago. Mach strongly argued that macroscopic phenomenology is robust enough for establishing the physics without invoking a then–speculative atomic theory. Mach's position almost carried the debate, greatly distressing Boltzmann and possibly contributing to his suicide. The positions taken by Planck are nicely illustrative of the debates. In the preface to his *Treatise on Thermodynamics* published in 1897 he described "the most fruitful" treatment of the subject as that which "starts direct from a few very general empirical facts" and contrasted it with "essential difficulties ... in the mechanical interpretation of the fundamental principles of Thermodynamics" [*Planck*, 1922, p. viii]. By the time of the second edition in 1905, following the success of his 1901 quantization postulate for black-body radiation, he had somewhat reluctantly adopted a more statistical perspective, arguing that "the full content of the second law can only be understood if we look for its foundations in the known laws of the theory of probability" [*Planck*, 1922, p. x].

This historical anecdote suggests that, against contemporary urgings, it may not necessarily be a good idea to turn to more "elementary" theories in order to re–establish something already phenomenologically established. A major strength of classical thermodynamics is in providing a robust theoretical framework without having to rely on more elementary atomic theory.

In the context of atmospheric convection, the use of cloud–resolving models (CRM) is nowadays extremely popular for studying convective systems. However, it is always worth exercising some caution with this major trend, reminding ourselves that while CRM studies might be elementary they are not necessarily fundamental. CRMs themselves contain many uncertainties, especially in the cloud physics. In this respect, CRM studies are not necessarily more robust than a macroscopic phenomenological approach.

2.7. Law of large numbers

As just discussed in the last subsection, the advent of statistical mechanics from the 1870s led to a link between microscopic dynamics and thermodynamics, and this gave rise to other conceptions of equilibrium. For example, in a standard undergraduate textbook on thermodynamics [*Adkins*, 1983] we find the statement (p7) "equilibrium is itself a macroscopic concept. We may only apply the idea of equilibrium to large bodies, to systems of many particles". The view of equilibrium as a large-scale statistical concept implies that strictly it holds only in the limit of infinite system size.

This argument can be grossly summarized by invoking the "law of large numbers": statistical expectation values (mean, variance, etc) become increasingly more reliable as the sample size increases. More precisely, the so–called "law of large numbers" in probability theory [*e.g.*, Ch. 10 of *Feller*, 1968] guarantees convergence of the mean for a sequence of mutually independent events to the probabilistic expectation. Furthermore, the central limit theorem guarantees that the fluctuation of this mean value around the expectation value decreases as $\sim N^{-1/2}$ as the sample number N increases. These two mathematical theorems provide an explanation of why macroscopic thermodynamic quantities, such temperature and pressure, can be measured in a stable and reliable manner.

Sometimes the same argument is invoked in order to justify convective quasi–equilibrium. A review by *Emanuel* [2000], for example, stresses the separation of convective and large scales. Convective processes are of much smaller scales (*ca.*, 1 km) than a typical synoptic scale (*ca.*, 10^3 km). As a result, the number of convective elements (or convective towers more precisely) contained within a typical synoptic-scale disturbance may be substantial, if the elements are not too widely spaced. Thus it is tempting to appeal to the law of large numbers and the central limit theorem.

The limitation of this argument must be kept in mind. The number of gas molecules contained in, say, 1 mm^3 volume is more than 10^{16} for standard atmospheric parameters (10^3 hPa and 300 K). On the other hand, just for the sake of an estimate, let us assume that a convective tower is found every 10 km in two horizontal directions. Even under this relatively dense situation, we find only 100 convective towers within an area of size (100 km)^2. The convection within such an area may be under equilibrium in a certain sense, but clearly with far more statistical uncertainties than the case of gas in a 1 mm^3 volume. Following this argument, *Plant and Craig* [2008] proposed that statistical fluctuations about equilibrium be explicitly taken into account in a stochastic convective parameterization, since such fluctuations may have larger-scale impacts and can also be of interest in their own right [*Ball and Plant*, 2008; *Groenemeijer and Craig*, 2012].

It is especially important to keep in mind that the central limit theorem guarantees only a relatively slow convergence towards the statistical expectation with increasing sample size N. A good example for making this point is the population dynamics of a two-species predator–prey system. Phenomenologically, many such systems have an oscillatory evolution, whereas a conventional description based on the number density of each species (Lotka–Volterra equation) gives only a steady state after an initial transient dependency. This apparent dilemma can be resolved by explicitly calculating the evolution of the number of each species as a particular realization of a stochastic process: oscillatory behavior is then revealed [*McKane and Newman*, 2005]. The reason for the discrepancy is traced to a discrepancy between an expectation value (as calculated by the Lotka–Volterra system) and a single actual realization of a stochastic process. The two solutions can be qualitatively different, even if an apparently large system of $\sim 10^3$ individuals is considered, because purely internal noise can induce resonance effects.

This study with a biological system emphasizes the need for caution and careful interpretation when applying statistical ideas of equilibrium that are based on large samples to real systems, such as a system of convective clouds, which have relatively small populations. We refer to *Solé and Bascompte* [2006] for biological descriptions of population dynamics. In particular, their book includes extensive discussions on *self–organization*, an issue to be discussed later in section 6.3. *van Kampen* [2007] provides more general discussions on probabilistic descriptions of physical (as well as biological) systems of finite extent. Some possibilities for and perspectives on applying explicitly stochastic descriptions to convective systems can be found in *Majda and Khouider* [2002], *Plant and Craig* [2008], *Khouider et al.* [2010] and *Plant* [2012].

2.8. Balance as a dynamical concept

The closest analogy to thermodynamic equilibrium in classical mechanics may be the concept of stationary or steady solutions. A steady solution occurs if there exists an inertial frame in which all of the particles of a given system remain at rest. The notion can, furthermore, be generalized to a case in which a "mode" of movement of the system does not change with time. For example, a planet rotating around a star along a fixed orbit with a fixed period, can also be considered to be in a steady state.

Steady solutions can be said to be subject to a "balance" condition. Under a perfect steady state with no motion, all the forces acting on a given particle sum to zero,

$$\sum_j \mathbf{F}_j = 0, \tag{2.1}$$

where $\mathbf{F}_j$ designates an individual force acting on the particle. In the example of the orbiting planet with steady motions, all of the forces acting on the planet remain perpendicular to the direction of movement and so are balanced along a given curved coordinate.

These *balanced* states are clearly similar to the thermodynamic concept of equilibrium in the sense that both can be characterized by a certain time invariance. However, the two concepts carry quite different implications, as illustrated in Fig. 1. Balance does not necessarily imply stability, although equilibrium implies stability. Balance is simply the identification a state which may be of interest, with no consideration made of whether and how that state may actually arise. On the other hand, equilibrium implies that the given system will arrive at the state of interest (assuming that the external conditions so allow) after certain transient adjustments.

A good example to illustrate the distinction is a standing egg: we can make an egg stand upright after careful adjustments of its position. However, this state is hardly stable, and an egg would never reach that position spontaneously. Thus, the state is under balance, but not in equilibrium.

Although convective quasi-equilibrium is most often considered in the context of a thermodynamic or statistical mechanics interpretation of equilibrium, some discussions based on the concept of dynamical balance can also found in the literature. An example is *Emanuel's* [2000] discussion on the entropy budget in his section II. However, in the context of convection, we should emphasize the role of dissipation as a major difference between the "balance" concept of classical mechanics and the equilibrium concept of thermodynamics (section 2.3). Clearly convection is subject to dissipation, although its importance may depend on the situation.

The concept of dynamical balance can be generalized under the Hamiltonian formulation into various other conservative systems. However, the distinction between conservative and non–conservative systems, or alternatively between non–dissipative and dissipative systems is important in this framework as well: the Hamiltonian formulation is essentially developed for non–dissipative systems.

Dissipative behavior requires a more general geometric description such as a metriplectic approach reviewed by *Guha* [2007]. Here, recall that a Hamiltonian system can be described in terms of the Poisson bracket [*cf.*, *Goldstein et al.*, 2002]. The basic idea of the metriplectic approach is to include both a Poisson bracket and a symmetric (metric) bracket. Whether such a mathematical structure can also be used to study CQE is an intriguing possibility. An example application relevant to atmospheric science can be found in *Bihlo* [2010].

3. ARAKAWA AND SCHUBERT'S CONVECTIVE QUASI–EQUILIBRIUM

In the previous section we discussed various concepts of equilibrium, all of which have been instrumental in various interpretations of convective quasi–equilibrium (CQE). Before we discuss those interpretations in section 4, we first present the original specification of CQE, as it appeared in *Arakawa and Schubert* [1974].

3.1. The Original Purpose: Closure Problem

As in other parameterization problems, a closure is required for convection parameterization. *Arakawa and Schubert* [1974, AS hereinafter] originally proposed CQE as a closure hypothesis for their development of a convection parameterization based on the mass–flux formulation.

The basic idea of the mass flux formulation is to represent convective clouds as an ensemble of plumes, with a one-dimensional steady plume model used to describe the vertical structure of each cloud type. The plume model essentially considers parcel ascent with additional hypotheses on the interactions between in-plume air and the surrounding environment. These interactions take the form of lateral exchanges of air known as entrainment and detrainment. For a given environment, the plume equations can be integrated vertically to obtain a vertical profile for any physical variable associated with a given cloud type, provided that values of those variables are known at the bottom of the plume. Most variables are assumed equal to the environmental state at the bottom of the plume, which is usually near to the top of the boundary layer. Small "triggering" perturbations may also be applied, to temperature [*Gregory and Rowntree*, 1990] or to vertical velocity [*Kain and Fritsch*, 1990, 1992] for example.

Thus, the plume model provides the vertical profile for a convective cloud type within a given environment, but it provides no information about the number of plumes that are present and the intensity of each. The problem of determining those values (*i.e.*, an overall magnitude for the convection) is called the "closure". In its original form, CQE is simply a diagnostic relationship proposed by AS in order to solve this closure problem. The relationship is a balance condition between large-scale forcing for convection and the convective activity itself, allowing for determination of the magnitude of parameterized convection. This diagnostic relationship is more precisely described next.

3.2. Formal Definition

The CQE hypothesis is based on a consideration of the energy cycle for the ensemble of convective plumes.

Convection is, by definition, driven by buoyancy, and thus the rate of generation of convective kinetic energy is primarily controlled by two factors: the strength of the convective vertical velocity w_c (or more precisely, the vertical momentum ρw_c) and the buoyancy, b. In order to calculate the generation rate over a synoptic area (or over a model grid box), we also have to pay attention to the fact that convection is not everywhere, but occupies only a fraction, σ_c, of the total area under consideration. Thus, the generation rate of convective kinetic energy defined per unit area is given by

$$\int_{z_B}^{z_T} \sigma_c \rho w_c b dz,$$

where the vertical integral extends from the convection base (z_B) to the convection top (z_T).

By starting from this expression, and by following AS, a potential efficiency of the kinetic–energy generation rate is estimated as:

$$A = \int_{z_B}^{z_T} \hat{w} b dz. \tag{3.1}$$

Here, we have introduced a normalized convective vertical momentum (plume profile) given by

$$\hat{w} \equiv \frac{\sigma_c \rho w_c}{\hat{M}} \tag{3.2}$$

with a normalization factor $\hat{M} = \sigma_{c,B}\rho_B w_{c,B}$ which is the rate of vertical mass-transport (the mass flux). A subscript B has been used to denote the normalization variables, because these are traditionally defined at the convection base z_B. Note that as a result, the generation rate of convective kinetic energy is simply given by $\hat{M}A$.

Arakawa and Schubert name the quantity A the cloud–work function. In their formulation, the buoyancy b and the plume profile, $\hat{w}$, are defined in terms of an entraining plume model, but this is not essential at all in order to understand the cloud–work function, as emphasized by *Yano et al.* [2005]. The cloud–work function can be evaluated once the profiles b and $\hat{w}$ are defined from any convective cloud model, even one without any explicit notion of entrainment or detrainment. In order to distinguish the generalized quantity in Eq. (3.1) from the cloud–work function defined by *Arakawa and Schubert* [1974] in a more narrow sense, we call the generalization the potential energy convertibility (PEC) as suggested by its relationship with the energy cycle. It is also convenient to introduce at this point the quantity CAPE (convective available potential energy) which is frequently used in the literature and is a special case of Eq. (3.1) for a "forced" ascent that maintains a constant lifting rate, $\hat{w} = 1$, without any mass exchange with the environment.

The discussion so far has considered a single cloud type only, but by following the original formulation of *Arakawa and Schubert* [1974], we now generalize to an ensemble of convective clouds (*e.g.*, entraining plumes with different entrainment rates) with λ labelling a cloud type. Proceeding directly from its definition in Eq. (3.1) the time evolution of PEC for each type can be computed given the cloud model (as in Appendix B of AS for example) and the result can be expressed in the following general form

$$\frac{d}{dt}A_\lambda = F_{L,\lambda} - D_{c,\lambda} \tag{3.3}$$

where $F_{L,\lambda}$ is the rate at which A_λ is generated by large–scale processes, and $D_{c,\lambda}$ is the rate at which A_λ is consumed by convective processes.

We shall discuss the concept of a scale separation in section 5.2, but it may be worth stressing at this point how the distinction arises between large-scale and convective processes in Eq. (3.3). Terms appearing on the right-hand side of the equation can be grouped into those which are independent of the cloud-base convective mass-flux (*i.e.*, $F_{L,\lambda}$) and those which are proportional to it (*i.e.*, $D_{c,\lambda}$: *Plant* [2010]). The scale on which various physical processes might act is therefore not of immediate relevance to the derivation and the structural form of this particular equation, which we emphasize is derived directly from the definition of A_λ and from any cloud model. Rather it is in the interpretation and application of the equation that notions of scale arise.

In practice the two most important forcing processes in $F_{L,\lambda}$ are radiative cooling and the adiabatic cooling of the environment associated with large–scale ascent, although surface fluxes will also contribute. The most important convective process contributing to $D_{c,\lambda}$ is latent heating associated with condensation of water. Phenomenologically speaking, the condensation process is to a good extent balanced by the adiabatic cooling associated with convective vertical motions, and thus the PEC consumption rate $D_{c,\lambda}$ due to a cloud type λ becomes proportional to the mass flux for the given cloud type. Noting also that all of the cloud types are able to reduce the PEC for other cloud types in a similar manner, we find that the consumption rate takes the form

$$D_{c,\lambda} = \sum_{\lambda'} \mathcal{K}_{\lambda\lambda'} \hat{M}_{\lambda'} \tag{3.4}$$

where $\mathcal{K}_{\lambda\lambda'}$ is the rate at which each unit of cloud-base mass flux for the cloud type λ' contributes to the reduction of A_λ. The idea is schematically summarized in Fig. 2.

Arakawa and Schubert's CQE hypothesis is formally obtained by assuming a stationary solution to the above tendency equation (3.3): *i.e.*,

$$F_{L,\lambda} - D_{c,\lambda} = 0. \tag{3.5}$$

3.3. What is the use of CQE?

In order to understand the purpose of imposing the condition (3.5), we have to understand slightly better the structure of Arakawa and Schubert's closure problem. A key variable required by the parameterization is the convective mass flux $M_\lambda = \hat{M}_\lambda \hat{w}_\lambda$ for each cloud type. Once these mass fluxes are known, all of the convective feedbacks to the parent model which must be determined by the parameterization can be evaluated in a more or less straightforward manner. However, we should add the caveat that a satisfactory treatment of cloud microphysical processes at convective scales is far from straightforward [*cf.*, *Donner*, 1993].

The mass flux is factorized into two parts as defined by Eq. (3.2): its normalized vertical profile $\hat{w}_\lambda$, and its amplitude $\hat{M}_\lambda$. As discussed in section 3.1, the vertical profile $\hat{w}_\lambda$ can be calculated with any preferred cloud model. This procedure is indeed relatively straightforward in all existing mass–flux parameterizations, because it is assumed that each cloud type is in a steady state for a given environment. We call this assumption the steady–plume hypothesis, since the clouds are typically modeled as plumes. Here, however, we emphasize that the whole formulation is independent of how a set for the vertical profiles $\hat{w}_\lambda$ is defined. Thus, the amplitude, $\hat{M}_\lambda$, becomes the major undetermined factor of the convection parameterization.

The purpose of Arakawa and Schubert in introducing the condition (3.5: convective quasi–equilibrium hypothesis) is to determine $\hat{M}_\lambda$. Substitution of (3.4) into (3.5) leads to

$$\sum_{\lambda'} \mathcal{K}_{\lambda\lambda'} \hat{M}_{\lambda'} = F_{L,\lambda}. \tag{3.6}$$

The large–scale forcing $F_{L,\lambda}$ is known from the parent model in which the parameterization operates, while the matrix $\mathcal{K}_{\lambda\lambda'}$ is known from the cloud model adopted, specifically the vertical profiles of convective plumes and the thermodynamic state of the detrained air. Note again that otherwise the details of the cloud model are not at issue in this formulation. Thus, the convective–cloud strength $\hat{M}_\lambda$ is obtained by inverting the matrix $\mathcal{K}_{\lambda\lambda'}$.

In this manner, Arakawa and Schubert's original CQE (Eq. 3.6) is defined in terms of the response of convection (left–hand side) to the large–scale forcing (right–hand side): its analogy with the dynamical balance (Eq. 2.1) is hard to miss. For the sake of promoting a conceptual understanding of Eq. (3.6), we may further consider that the matrix, $\mathcal{K}_{\lambda,\lambda'}$, is climatologically given. Under this approximation, Eq. (3.6) literally becomes a linear problem for determining the convective–cloud strength, $\hat{M}_\lambda$, from a given large–scale forcing. However, the formal matrix inversion is hardly trivial, because each convective type, λ', projects onto all convective types, λ, in response. For example, deep large–scale forcing is not necessarily responded to by a deep mode alone but rather by a spectrum of deep and shallow modes. Issues associated with inversion are further discussed next.

3.4. Problems associated with the inversion method based on Eq. (3.6)

Some potential problems can be foreseen in attempting to invert the $\mathcal{K}_{\lambda\lambda'}$ matrix using Eq. (3.6).

i) *Does the inversion solution exist?* An issue here is the numerical stability of the inversion as well as the question of invertibility when a continuous spectrum of convective cloud modes is considered under a discretization into types. However, the most serious issue is the constraint that the normalized mass flux, $\hat{M}_\lambda$, must be positive (because it is expected to represent a convective *updraft*), but Eq. (3.6) does not guarantee it. According to *Lord* [1982] this is not merely an issue of principle but one that poses real practical difficulties in identifying a solution. However no systematic investigation of this basic question has been performed since then to the best of our knowledge.

ii) *Does the system actually evolve slowly?* Although slowness of the evolution of A_λ, as observationally known [*cf.*, Fig. 3: reproduced from Fig. 13 of *Arakawa and Schubert*, 1974], suggests the validity of Eq. (3.5) for estimating $\hat{M}_\lambda$, the inverse is not necessarily true.

Note that by replacing a prognostic equation (3.3) by a diagnostic relation (3.5), we lose the ability to predict PEC, A_λ, directly. Instead its evolution must be diagnosed, based on the predicted evolution of the thermodynamic fields. The situation is analogous to the use of hydrostatic balance. Recall that hydrostatic balance is obtained by setting $Dw/Dt = 0$ in the vertical momentum equation. The vertical velocity w can then be calculated diagnostically from mass continuity. This does not automatically mean that the evolution of the vertical velocity so-computed is actually slow. Indeed, the resulting evolution of vertical velocity could be much more rapid than an evolution that would support the hydrostatic balance assumption.

iii) *Does the system evolve at all under the steady constraint (3.6)?* In contrast to the previous question, we now ask whether the imposition of Eq. (3.6) might over-constrain the evolution of temperature and moisture. Consider the case that the large–scale forcing, $F_{L,\lambda}$, is slowly varying in time. The quasi–equilibrium constraint (3.6) implies that the

PEC, A_λ, is stationary with time, and the convective strength, $\hat{M}_\lambda$ is obtained by assuming stationarity of a set of PECs. The PEC for each cloud type is defined by a vertical integral, Eq. (3.1), of a function of temperature and moisture, the function being different for each type. The stationarity of PEC would therefore suggest that both temperature and moisture are stationary with time, if enough cloud types with enough functional forms of the integrand are considered to obey the constraint, and if those functional forms are a complete basis set. Thus, it is not clear that we still see temporal evolution of the thermodynamic variables. This is rather a serious issue, because it suggests that the convective quasi–equilibrium may destroy the predictability of thermodynamic fields (*cf.*, section 6.1 for a related issue).

It should be appreciated that the question is not trivial, again through analogy with the hydrostatic balance. In that case, although the steadiness of the vertical velocity field is implied by the balance assumption, the vertical velocity field nonetheless evolves with time as estimated by mass continuity. Unfortunately, it is not immediately clear whether an analogous situation occurs for the convective quasi–equilibrium hypothesis or not.

iv) *Is the inverted solution stable against a perturbation?* The obtained balanced solution may turn out to be unstable against any linear perturbation. In that case, in practice, the system would never stay at convective quasi–equilibrium. This possibility was recently investigated by *Wagner* [2010, but see also *Plant and Yano*, 2011]. In his case study, it was found that such an instability occurred about 30% of the time. Note that this issue has a link to the concept of self–organized criticality, which will be discussed in section 6.3.

v) *A choice of spectrum?* The characteristics of the $\mathcal{K}_{\lambda\lambda'}$ matrix depend on the choice of the spectrum, $\{\hat{w}_\lambda\}$, for convective plumes. *Arakawa and Schubert* [1974] took the entraining plumes as such a set with λ standing for the fractional entrainment rate. However, the CQE hypothesis (3.5) itself does not depend on the entraining plume hypothesis, but could be applied to any choice of set, $\{\hat{w}_\lambda\}$, of convective plumes. This generality is important especially considering the fact that accumulating evidences [*cf.*, *de Rooy et al.*, 2012] suggest that a more complex entrainment–detrainment formulation is required in order realistically to represent atmospheric convection. The point that Arakawa and Schubert's original CQE can be interpreted independent of their specific choice of entrainment–detrainment hypothesis is hardly overemphasized. Thus, a major challenge is to identify an appropriate framework for defining a set of convective plumes under a more general entrainment–detrainment formulation.

3.5. Relaxation and Prognostic Approaches

In the face of both practical and conceptual difficulties in inverting Eq. (3.6), various relaxation approaches have been proposed, using either an iterative minimization procedure for dA_λ/dt or else a prognostic calculation incorporating Eq. (3.3) [*e.g.*, *Moorthi and Suarez*, 1992; *Randall and Pan*, 1993; *Randall et al.*, 1997; *Pan and Randall*, 1998]. Here, we discuss the physical issues in treating the convective energy cycle prognostically, as considered by *Randall and Pan* [1993], *Randall et al.* [1997], *Pan and Randall* [1998].

Recall that PEC, A_λ, as defined by Eq. (3.1), is a measure of the efficiency with which convective kinetic energy, K_λ, is generated from potential energy. The actual energy generation rate is $\hat{M}_\lambda A_\lambda$ as already noted. Denoting by $D_{K,\lambda}$ the dissipation rate of convective kinetic energy, a prognostic equation for K_λ is given by

$$\frac{d}{dt}K_\lambda = \hat{M}_\lambda A_\lambda - D_{K,\lambda}. \tag{3.8}$$

The set of equations (3.3) and (3.8) provides a description for an energy cycle of a convective system. According to Eq. (3.8), PEC generates convective kinetic energy. Increased kinetic energy would be expected to be associated with an increase of convective mass flux, $\hat{M}_\lambda$, which modifies A_λ by Eq. (3.3). The system can be closed by assuming a functional relationship between K_λ and $\hat{M}_\lambda$, as well as an expression for the dissipation rate $D_{K,\lambda}$.

Details of those assumptions may be important at a conceptual as well as a practical level. This can most easily be seen from a comparison of the systems proposed by *Randall and Pan* [1993] and *Yano and Plant* [2012a, b] which have different stability characteristics for departures from the equilibrium state (recall the question iv) above).

However, leaving such details aside here, an important aspect indicated by Eq. (3.8) is the possibility for another type of convective quasi–equilibrium by assuming a steady state for this equation, *i.e.*,

$$\hat{M}_\lambda A_\lambda - D_{K,\lambda} = 0 \tag{3.9}$$

This condition is called kinetic–energy quasi–equilibrium by *Lord and Arakawa* [1980] in contrast to the cloud–work function quasi–equilibrium defined by Eq. (3.5). Physical interpretations for Eq. (3.9) can be developed in a similar manner as for the original quasi–equilibrium hypothesis.

3.6. Operational implementation of CAPE closure

The majority of current convection parameterizations used in operational models take a bulk approach, in which only a single convective mode is considered, rather than a spectral approach as introduced by *Arakawa and Schubert* [1974], and as also discussed so far. *Plant* [2010] discusses the assumptions required in order to reduce a spectral formulation to a bulk one. Although various approaches to the formulation of CQE could be taken in making the reduction, in practice most bulk parameterizations consider a formulation based on replacing PEC (cloud work function) with CAPE.

We noted in section 3.2 that CAPE is a particular limiting case of PEC, and as such its evolution equation can immediately be written analogously to Eq. (3.3):

$$\frac{d}{dt}\mathrm{C}APE = \left.\frac{d\mathrm{C}APE}{dt}\right|_{\mathrm{conv}} + \left.\frac{d\mathrm{C}APE}{dt}\right|_{\mathrm{L}} \tag{3.9}$$

Thus, the total CAPE tendency can be divided into the contributions from convective (subscript conv) and large–scale processes (subscript L). A statement of CQE for the CAPE budget is therefore

$$\left.\frac{d\mathrm{C}APE}{dt}\right|_{\mathrm{conv}} + \left.\frac{d\mathrm{C}APE}{dt}\right|_{\mathrm{L}} = 0. \tag{3.10}$$

From a parallel argument as for deriving Eq. (3.4) above, the convective contributions to the CAPE tendency may be written as

$$\left.\frac{d\mathrm{C}APE}{dt}\right|_{\mathrm{conv}} = -\mathcal{K}\hat{M} \tag{3.11}$$

where the matrix $\mathcal{K}_{\lambda\lambda'}$ has been reduced to a single coefficient $\mathcal{K}$ which is defined by both the environmental profile and the normalized vertical profile of the bulk plume. Substitution of Eq. (3.11) into Eq. (3.10) leads to a closure condition:

$$\hat{M} = \frac{1}{\mathcal{K}} \left.\frac{d\mathrm{C}APE}{dt}\right|_{\mathrm{L}}. \tag{3.12}$$

The above condition is analogous to the Arakawa and Schubert CQE. However, the actual CAPE-based closure used in many operational models is defined differently. In *Zhang and McFarlane* [1995] and *Gregory et al.* [2000], for example, it is assumed that convection consumes CAPE at a rate that is determined by a characteristic closure time–

scale τ_c. Hence,

$$\left.\frac{dCAPE}{dt}\right|_{\rm conv} = -\frac{CAPE}{\tau_c}. \tag{3.13}$$

Substitution of Eq. (3.13) into Eq. (3.11) leads to

$$\hat{M} = \frac{\rm CAPE}{\tau_c \mathcal{K}} \tag{3.14}$$

The formulation (3.13) is inspired by an iterative CAPE–based closure originally introduced by *Fritsch and Chappell* [1980], and adopted by *Kain and Fritsch* [1990] and *Bechtold et al.* [2001]. It expresses the action of convection as a relaxation process for CAPE. We discuss this and other relaxation ideas more fully in section 4.5.

Notice that in order for the closure (3.14) to be equivalent to the CQE closure condition (3.10), the large–scale forcing must satisfy the condition

$$\left.\frac{dCAPE}{dt}\right|_{\rm L} = \frac{CAPE}{\tau_c} \tag{3.15}$$

or else the relaxation time–scale, τ_c, would have to be defined in such a manner that Eq. (3.15) is satisfied at every time step.

3.7. Concluding Remarks

In this section, we have examined *Arakawa and Schubert's* [1974] original formulation for convective quasi–equilibrium (CQE) in some detail. This is because, apart from its own importance, in our opinion their original formulation still provides the most solid basis for examining the CQE concept even today. Various other forms and interpretations for CQE have been proposed, as will be discussed in the next section, but frequently lack the simplicity and clarity of *Arakawa and Schubert's* [1974] original formulation in terms of the cloud work function budget. Further studies of the limitations of CQE and on alternative perspectives, as will be discussed in remainder of the paper, would also be best conducted using their framework.

Another reason for this emphasis is because their original CQE proposal is surprisingly little examined in its original form. The cloud work function based on the entraining plume hypothesis can relatively easily be evaluated from any sounding data. However, little such data analysis is reported in the literature. Many of the observational CQE analyses to be discussed in the following sections are based instead on CAPE. Furthermore, various other proxies for convective instabilities can be proposed. For example, a simple quantity intermediate between CAPE and the cloud work function is the dilute CAPE (or entraining CAPE) [e.g. *Zhang et al.*, 2009] which is given by Eq. (3.1) with the buoyancy computed for an entraining plume but the plume profile $\hat{w}$ set to unity.

Another possibility arises from *Holloway and Neelin*'s [2009] demonstration that the parcel–lifted buoyancy profile (for various different entrainment hypotheses) can be well stratified using column–integrated water vapor (CWV, *cf.*, section 6.3). CWV is an attractive quantity to use as it allows much valuable satellite data to be brought to bear, and it is known to exhibit some similar behaviours to the dilute CAPE [*Neelin et al.*, 2008; Shanay *et al.*, 2012] but it must nonetheless be recognized that the correspondence of any of the alternative buoyancy-related measures is hardly one–to–one with the cloud work function itself.

The usefulness of the undilute parcel hypothesis adopted in the CAPE calculation is also open to question, given that the existence of undilute air in convective plumes has long been debated [see *e.g.*, *de Rooy et al.*, 2012 for a recent review on entrainment issues]. Although recent CRM studies have been performed in order to address this question

[*Fierro et al.*, 2009, 2012; *Romps and Kuang*, 2010], it has not been widely recognized that an inversion of the CQE equation (3.6) would provide a more direct estimate of the contribution of undilute (or almost undilute) convective plumes to the whole spectrum.

4. INTERPRETATIONS OF CQE

In the previous section, Arakawa and Schubert's convective quasi–equilibrium hypothesis (3.5) was introduced in terms of the PEC budget equation (3.3). The goal of the present section is to review the subsequent development of interpretations of the hypothesis. The original argument for its physical justification can be found in section 7 of *Arakawa and Schubert* [1974]. The discussion of the present section can be considered as an expanded version of that original, incorporating various later comments and re–interpretations.

The concept of CQE can perhaps be most simply understood, and is certainly most often interpreted, by analogy with thermodynamic equilibrium (section 4.2). The idealization of a radiative–convective equilibrium can be considered as a starting point (section 4.1) for this perspective. A perspective from statistical mechanics (section 4.3) may provide a more robust basis. In order to justify quasi–equilibrium from the perspective of statistical mechanics, both the law of large numbers and the principle of scale separation become important issues (section 4.4). The scale-separation principle further suggests two other possible interpretations of quasi–equilibrium: as a relaxation process (section 4.5) and as a slow manifold (sections. 4.6 and 4.7). A more specific interpretation can be developed by a lifting parcel argument (section 4.8). Finally, we point out that CQE may be interpreted as part of a wide category of processes called *homeostasis* (section 4.9). The following discussions, as a whole, trace tendencies for both generalizations and re–interpretations of the CQE concept.

4.1. Radiative–convective equilibrium

The notion of radiative–convective equilibrium originates from the context of a radiative transfer calculation within a static one-dimensional atmosphere. Under such conditions, a stationary solution for the vertical profile of temperature $T(z)$ can be sought, such that

$$Q_R = 0,$$

where Q_R is the radiative heating rate of the atmosphere.

The obtained radiative equilibrium state is typically unrealistic in the sense that the lapse rate in the troposphere is too steep and even convectively unstable, contradicting observations. For example, such a super–adiabatic state was found in one of the original full radiative equilibrium calculations for the atmosphere by *Möller and Manabe* [1961]. Phenomenologically, convection would arise in a situation, and act to adjust the vertical profile to a stable configuration. Hence, for the problem of radiative equilibrium to be of practical interest it must be at least minimally modified into the problem of radiative–convective equilibrium, by adding the effects of moist convection to the above equation:

$$Q_R + Q_c = 0.$$

Here, Q_c refers not only to diabatic heating (due to latent heating, both by condensation and freezing of water) but also to vertical transfers of heat associated with convective motions.

Manabe and Strickler [1964] introduced a critical lapse rate into their calculations, adjusting the computed lapse rate to the critical value wherever the former exceeds the latter. They proposed to call this procedure *convective adjustment*, and to call the profile so-obtained the *radiative–convective equilibrium* (RCE). Note that under this procedure, the "convective" heating, Q_c, is completely implicit, and the term represents all dynamical and thermodynamical processes that were neglected in the radiative transfer calculations.

Today, the concept of RCE has been rather taken away from this original context, and often refers to a final state obtained by integrating a CRM for an extensive period, without imposing any large–scale vertical motion [*e.g.*, *Grabowski*, 2003; *Cohen and Craig*, 2006; *Stephens et al.*, 2008; *Parodi and Emanuel*, 2009]. In such applications, the radiative heating rate may be computed with a complex radiation code that is fully responsive to the simulated cloud field, but alternatively, it may sometimes be treated in extremely simplified manner: for example, by imposing a prescribed, fixed tropospheric cooling rate. In the latter case, the cooling rate is not necessarily typical of the rates in the tropical atmosphere. Indeed, the purpose of an investigation may be to assess the scaling of convective properties with the strength of the cooling [*e.g.*, *Shutts and Gray*, 1999; *Parodi and Emanuel*, 2009].

Finally, the problem can be further generalized to the situation in which the given atmospheric column is no longer static, but is subject to vertical motion. Vertical heat transport from the column-averaged vertical motion, w, must then be added to the above equation:

$$-w\frac{\partial\theta}{\partial z} + Q_R + Q_c = 0. \tag{4.1}$$

Here, θ is the potential temperature, and then the terms Q_R and Q_c must be re-interpreted in terms of the potential temperature. Such a generalization may include simulations where w is imposed from the outset [*e.g.*, *Sui et al.*, 1994]. Alternatively, w can be diagnosed based on the assumption that Eq. (4.1) is satisfied at every time step [*e.g.*, *Sessions et al.*, 2010; also see section 6.1].

Note that there is no ambiguity about the above generalizations of the concept of RCE so long as the system evolves to a configuration that remains perfectly steady with time, even with a non–vanishing vertical velocity, $w \neq 0$. However, as soon as the system becomes time evolving, the concept of RCE suddenly becomes subtle, even elusive. The degree to which RCE is satisfied under time–evolving situations is not an easy question to answer. However, the basic assertion of CQE is, essentially, that Eq. (4.1) remains a good approximation even when the system is evolving with time. We return to this issue in section 6.1.

4.2. Thermodynamic Analogy

If a convective system can be treated as analogous to a thermodynamical system then both Eqs. (3.6) and (4.1) can be justified. For this purpose, we suppose that the convective system constitutes only a small part of the whole "atmospheric system". More precisely, we suppose that the given large–scale atmospheric state can be considered as a fixed external environment to the convective system similarly to a "thermal bath" for a thermodynamical system. In this manner, the convective system is then slaved to the given large–scale environment.

As a result, if the large–scale state changes (through some suitably slow process), as described by $F_{L,\lambda}$ in Eq. (3.6) or $F_L \equiv -w(\partial\theta/\partial z) + Q_R$ in Eq. (4.1), then the convective state is considered to adjust almost immediately (as a fast process) in order to produce a $D_{L,\lambda}$ or Q_c which satisfies Eq. (3.6) or Eq. (4.1), respectively. This is the same concept as a heat engine adjusting itself to a new equilibrium as an external condition (its environment) is modified (*cf.*, section 2.1).

It is important to note that cause and effect are clearly distinguished under this thermodynamic analogy: the large scale is regarded as the cause, and the convective system responds to any changes in the large scale as an effect. This logic is also consistent with the original purpose of the CQE hypothesis as discussed in section 3: to determine the magnitude of convection, $\hat{M}_\lambda$ or Q_c, given the large–scale forcing.

In general, CQE under a thermodynamic analogy as defined by Eq. (4.1) should apply to all vertical levels. However, versions of the quasi–equilibrium hypothesis focussing on particular vertical levels have been proposed. In particular, *Raymond* [1995] proposed that the magnitude of convection over the tropical oceans is governed by quasi–equilibrium of the boundary layer. Note that when the quasi–equilibrium principle given by Eq. (4.1) is applied to the boundary layer, previously–neglected eddy heat transfer effects in the equation become important. Thus, the boundary–layer quasi–equilibrium is essentially established as a balance between heating and moistening by surface fluxes and cooling and drying by convective downdrafts.

CAPE budget analyses of sounding data over both the tropics [*Zhang*, 2003] and mid–latitudes [*Zhang*, 2002] have shown that the free troposphere is closer to an equilibrium than the boundary layer. The analysis of *Donner and Phillips* [2003] supported this conclusion, whilst also demonstrating that a boundary–layer quasi–equilibrium is reasonable on time–scales of around one day or longer. These authors proposed the term parcel–environment quasi–equilibrium to refer to the equilibrium concept as applied to the free troposphere only. In essence, this concept recovers the free tropospheric balance (4.1).

4.3. Statistical Cumulus Dynamics

A rigorous justification of the analogy of convective quasi–equilibrium with thermodynamic equilibrium discussed in the last subsection would be "an eventual goal of statistical cumulus dynamics" as stated in *Arakawa and Schubert* [1974]. As foreseen by *Arakawa and Schubert* themselves, improvements in cloud–resolving modeling are expected to lead to the development of statistical theories describing ensembles of cumulus clouds. Such theories should reduce to quasi–equilibrium in suitable limits. However, in spite of rapid modeling improvements in recent decades, rigorous theories of statistical cumulus dynamics remain in their infancy [*cf.*, *Cohen and Craig*, 2006; *Craig and Cohen*, 2006; *Plant*, 2009, 2012; *Davoudi et al.*, 2010; *Khouider et al.*, 2010].

Nonetheless, some important ingredients for developing a statistical theory that would support the concept of CQE can easily be pointed out. In particular, we should be able to identify a large enough number of convective "elements" [*i.e.*, convective plumes in the original formulation of *Arakawa and Schubert*, 1974] within a single large–scale domain in order to ensure that the ensemble statistics which might be predicted by the statistical theory would be representative of the modeled domain-mean statistics. In other words, a large–scale, macroscopic state must itself be well defined and must to a good approximation be actually realized in practice. In much of the parameterization literature, the large–scale domain is equated with a grid–box of the parent model, but in practice the size of the large–scale domain must be much larger than a single grid–box in a typical GCM (global circulation model) in order to guarantee a smooth description of large–scale processes as argued by *Lander and Hoskins* [1997 , see also *Xu et al.*, 1992; *Cohen and Craig*, 2006; *Shutts and Palmer*, 2007; *Jones and Randall*, 2011]. The law of large numbers discussed in section 2.6 can then support a statistical cumulus dynamics approach. A simple corollary is the need for a clear separation of scales between convective and large–scale processes.

4.4. Scale Separation Principle

Partially for the reason discussed in the previous subsection, the notion of scale separation is often invoked in order to justify convective quasi–equilibrium. As discussed in *Gombosi* [1994] or Ch. 1 of *Salmon* [1998] for example, the vast difference in scales between those of interest in typical fluid–mechanics problems and the mean free path between molecular collisions is important in order to justify the statistical thinking that leads to the Navier-Stokes equations. *Emanuel* [2000] draws the analogy with convective scales that are much smaller than those scales of interest in synoptic meteorology, and thus a similar statistical thinking might be justified, although obviously such an approximation must be very much more tentative in the convective case.

It may be worth restating in the present context that although scale separations are implied both in space and time, the relationship to quasi–equilibrium in the sense described in section 4.2 emphasizes the fundamental importance of a separation in the time–scales. Time–scale separation is not necessarily equivalent to space–scale separation [*cf.*, *Yano*, 1999]. Here, the convective time–scale is expected to be much shorter than the time–scales of interest for the evolution of the "large–scale" atmosphere.

Davies et al. [2009] explicitly demonstrate the importance of time–scale separation in order to establish CQE. They consider a dynamical system consisting of a fixed periodic large–scale forcing and a convective relaxation time–scale. Convection is no longer *slaved* to the large–scale forcing when the forcing period is reduced such that it approaches the relaxation time–scale. The same point has also been demonstrated in analogous CRM experiments [*Davies et al.*, 2012].

The notion of a time–scale separation, while recognizing a non-zero convective time–scale, leads to the idea of considering convective processes as adjustments towards an equilibrium state. Due to the time–scale difference, the adjustment is accomplished relatively rapidly. Thus, CQE can be re–interpreted as a fast adjustment process, as proposed by *Neelin and Yu* [1994]. We further discuss this concept in the next subsection.

Overall, one may argue, the system remains on a "slow manifold" due to such rapid adjustment processes, in the same sense in which geostrophic adjustment maintains mid-latitude dynamics on a slow manifold. *Emanuel* [2000] concludes his essay by invoking this notion, which we consider further in section 4.6.

Thus, from the point of view of considering quasi–equilibrium as a consequence of time–scale separation, we are led to two alternative views: moist–convective adjustment and slow manifold. More generally, there is an extensive literature in applied mathematics for systems consisting of two distinguishable and well-separated time–scales. It may be that some of these approaches are under-explored for convective problems, and we refer to the review by *van Kampen* [1985] for various possibilities.

4.5. Moist Convective Adjustment

As suggested in the previous subsection, CQE may be re–interpreted as a "moist adjustment" process, the point of view adopted by *Neelin and Yu* [1994] and *Yu and Neelin* [1994] for example, in their treatment of tropical convective–dynamics interactions [see also *Neelin*, 1997 as a review]. Several convection parameterizations have been developed based on the idea of moist convective adjustment. *Manabe et al.* [1965] used a hard adjustment in which any convectively unstable atmospheric profile is instantaneously reset to a moist adiabat, in the same spirit as *convective adjustment* in *Manabe and Stricker's* [1964] RCE. The *Kuo* [1974] scheme can be regarded as imposing a soft adjustment over a finite time–scale [*Arakawa*, 2004], but perhaps the most familiar such parameterization to the modern reader is that of *Betts* [1986]. Based on this scheme, full global atmosphere models of intermediate complexity have been developed, known as quasi-equilibrium tropical circulation models [QTCMs: *Neelin and Zeng*, 2000].

However, differences between the convective adjustment interpretation of CQE and Arakawa and Schubert's original definition should be noted. First of all, moist–convective adjustment clearly generalizes the notion of CQE. The rapid adjustment towards a target state can be defined with the target specified in terms of any convection–related variables, whereas in Arakawa and Schubert's original formulation the equilibrium is based very specifically upon the budget of the cloud work function.

More importantly, however, one should realize that this re–interpretation leads to a qualitatively different formulation to CQE. Arakawa and Schubert's original formulation attempts to define an equilibrium directly without explicitly considering a transitional phase. On the other hand, in the re–interpreted version, the assumed–short, but nonetheless non-zero time–scale for adjustment is explicitly recognized.

In this manner, unlike the original Arakawa and Schubert's CQE, convection is no longer slaved to the large–scale state, but two–way interactions between convection and the large–scale state are established. More precisely, in the formulation of *Neelin and Yu* [1994] and *Yu and Neelin* [1994], the large-scale state is adjusted towards a reference profile by convection, with the reference profile itself being dependent on the large–scale state. Note also that the reference moisture profile is defined in terms of the background temperature profile. Convection naturally furthermore modifies the background thermodynamic state. Such a modulation effect is considered by *Kuang* [2011] by coupling a cloud–resolving model with self–contained dynamics for large–scale gravity waves, as introduced in section 5 of *Yano et al.* [1998]. Here, a slight modification is to replace the temporal tendency of the large–scale vertical velocity with a Reynolds–like damping tendency.

Such a two–way interaction approach allows the specification of a vertical profile for tropical flows, as formulated in *Neelin and Yu* [1994] and numerically presented by Fig. 2 of *Yu and Neelin* [1994]. The main assumption is that a perturbation thermodynamic profile satisfies a moist adiabat, while the vertical profile for the basic state can remain unspecified. *Holloway and Neelin* [2007] present supporting evidence for this assumption above the boundary layer up to almost the top of convection. Thus, a reference perturbation temperature can be determined, which is then used in *Neelin and Yu*'s [1994] parameterization.

However, we must recognize that the "adjustment" interpretation of CQE is different from the interpretation of CQE through a thermodynamic analogy as discussed in section 4.2. Although this type of Newtonian relaxation description is very convenient for many purposes (such as a Newtonian cooling form for radiation: *e.g.*, *Gill*, [1980]), it is clearly a qualitative description with no obvious correspondence to any approximation of a more physically–based formulation. Indeed the introduction of *Betts* [1986] is careful to make no such claim, stressing instead the operational utility (which is undeniable even today) and that the relaxation "sidesteps all the details of how the subgrid-scale cloud and mesoscale processes maintain the quasi-equilibrium structure we observe".

An indication that the simplest Newtonian relaxation is not in fact an accurate description of the adjustment process is the recent study of *Raymond and Herman* [2011]. These authors demonstrate that the relaxation time–scale may be height dependent, and moreover that there may not be a well-defined or unique target profile. Another challenge arises from observations of power-law behaviour (to be discussed in section 5.2 and 6.3) which suggests that a single convective relaxation timescale cannot be well defined as the rate of decay to equilibrium will depend on the initial departure from the equilibrium state [*e.g.*, *Yano et al.*, 2001; section 6 of *Neelin et al.*, 2009; *cf.*, *Yano and Plant*, 2012a].

In the Betts–Miller scheme [*Betts*, 1986], the relaxation is performed towards a target profile of temperature and moisture. This leads to a further re–interpretation of CQE as being a means of maintaining the atmosphere close to a reference profile. Over the tropics, intuitively the most likely reference profile would be a moist adiabat, as emphasized by *Emanuel* [2007]. This then leads further to an anticipation that CQE maintains the tropical atmosphere close to a state of convective neutrality, whenever convection is a dominant process. An early application of the reference–state idea can be found in *Lord et al.* [1982] (their Eqs. 8 and 9) where convective forcing is evaluated in terms of departures from a set of time–averaged reference values for the cloud work function. *Xu and Emanuel* [1989] argue that the convectively–neutral reference state has zero CAPE, for a definition of CAPE based on reversible ascent (*i.e.*, assuming that all condensed water is retained by the lifting air parcel). The analysis of sounding data by *Roff and Yano* [2002] shows that such a state of zero reversible CAPE is indeed realized as a time mean over the tropical Western Pacific, but also emphasizes that deviations from the mean state are substantial over synoptic time–scales.

In summary, the moist adjustment re–interpretation of CQE leads to a different perspective than the original definition by Arakawa and Schubert. Whilst Arakawa and Schubert's original formulation (Eq. 3.5) defines CQE as a response to forcing, the adjustment re–interpretation considers CQE as a function of a large–scale state (typically a thermodynamic vertical profile). The latter obviously has the advantage of intuitive appeal, especially as further developed into a focus on the concept of "transition to strong convection" (*cf.*, section 6.6).

4.6. Balance Condition and Slow Manifold

An alternative way of interpreting CQE is, as already suggested in section 2.7, merely as a dynamical balance condition. We believe that the concept of the *slow manifold* best elaborates this point of view. The slow manifold may be considered a generalization of quasi–geostrophic flows. However, we should be careful with a subtle difference between quasi–geostrophy and the slow manifold. Quasi–geostrophy implies an approximate solution to an exact system (*i.e.*, geostrophy), whereas the slow manifold refers to where in phase space an exact solution actually resides, albeit after some filtering may have been applied.

The concept of slow manifold was originally proposed by *Leith* [1980], and revisited by *Lorenz* [1986; see also *Lorenz*, 1992]. An analogy between the ideas of the slow manifold and quasi-equilibrium was made by *Schubert* [2000]. In his defence, Schubert does not use the terminology *slow manifold*, but it is easy to infer this concept behind his essay.

Schubert [2000], more specifically, draws attention to the analogy between Arakawa's original ideas of quasi–equilibrium and quasi–geostrophic theory: the CQE condition is considered to filter out the transient adjustment of a convective cloud ensemble in the same sense that quasi–geostrophic theory filters out transient inertia–gravity waves. By extending this analogy, the state of CQE may be considered as analogous to the slow manifold.

To illustrate his point, *Schubert* [2000] first considers a linear one–dimensional shallow–water system with an exponentially–decaying mass source of the form (after appropriate nondimensionalization):

$$\alpha^2 t \exp(-\alpha t).$$

The cases of $\alpha \ll 1$ and $\alpha \gg 1$ correspond respectively to slow and fast forcing, and the time–integrated mass source is normalized to unity by definition. *Schubert* [2000] shows that provided the forcing time–scale is slow, the inertia–gravity wave mode appears only at an initial stage of the evolution starting from a stationary initial condition. At later times, only the geostrophic mode remains. One way to filter out the transient inertia–gravity waves *a priori* is to introduce a balance condition (his Eq. 15) so that the wave modes never arise.

By the same token, we might consider a balance condition being applied to a convection parameterization in such a way that fast *convective adjustments* are filtered out from the model evolution. *Schubert* [2000] argues that this is the original idea behind the CQE hypothesis. He makes this point more explicitly by analyzing Arakawa's earlier version of a cumulus parameterization for a four–layer GCM [*Arakawa*, 1969]. Thus, *Schubert's*

[2000] perspective invites an interpretation of convective quasi–equilibrium as a type of balance condition which holds convective dynamics on a slow manifold.

In their linear stability analysis of convectively–coupled waves, *Neelin and Yu* [1994], and *Yu and Neelin* [1994] also emphasize the importance of the distinction between fast and slow modes. In their formulation with convective adjustment, the fast modes always damp with the fast convective time–scale, thereby ensuring the maintenance of the equilibrium defined as a basic state of the model. On the other hand, the slow modes may be interepreted as explaining aspects of observed convectively–coupled equatorial waves.

These perspectives lead to further implications by analogy with the issues encountered in the original slow–manifold problem as developed by Lorenz and others.

4.7. Slow Manifold and Lighthill's Theorem

The main issue encountered with the concept of the slow manifold is whether it is actually possible to construct a system consisting solely of the slow time–scale processes. If that is the case, then we can develop a fully self-consistent description of geophysical flow on a slow manifold. The same question can be asked of the concept of CQE, because it implicitly assumes that a self–contained description of large–scale flows is possible while leaving implicit the fast convective–scale processes [*cf.*, *Ooyama*, 1982].

A more specific way of addressing this question is to consider a full system initialized only with slow modes (*i.e.*, the system initially resides on the slow manifold, or alternatively is in a state of convective quasi–equilibrium), and to ask whether the system evolves in such a way as to remain on the slow manifold (or alternatively to ask whether the CQE condition remains satisfied). As discussed in the previous subsection through an example from *Schubert* [2000], the case with a linear system is relatively obvious: so long as the slow and the fast modes are orthogonal in phase space (as is the case for geostrophic and inertia–gravity modes, *cf.*, *Greenspan* [1968]), the system remains on the slow manifold. However, once the system becomes nonlinear, the question is far from trivial. Lighthill's theorem casts light on this question.

In general, *Lighthill's* [1952, 1954] theorem as re–interpreted by the McIntyre school [*Ford*, 1994; *Ford et al.*, 2000; and summarized in *McIntyre*, 2000] says that a system initialized under geostrophic balance (or other balanced condition) will spontaneously generate gravity waves. The theorem suggests that a slow manifold can exist only in a limited sense, and that it is not possible to construct a system purely consisting of slow modes (*i.e.*, geostrophic modes, or Rossby modes).

Recent work by *Ring* [2009] illustrates this issue more concretely. Consider an expansion for some variable A,

$$A = A_0 + \epsilon A_1 + \epsilon^2 A_2 + \cdots$$

in terms of the expansion parameter ϵ which is a Rossby number. Thus, the expansion describes systematic departures from geostrophy. Note that *Neelin and Yu* [1994] perform a similar expansion in their linear stability analysis, taking the convective adjustment time–scale to be their expansion parameter. In the specific case studied by *Ring* [2009], the expansion is performed for a shallow–water system on an f-plane. It is shown that even when a (fast) intertia–gravity mode is absent from the initial state, $A_2(t = 0) = 0$, nonetheless A_2 grows quickly so that geostrophy breaks down due to the nonlinearity of the system. At a final stage, the contribution of the inertia–gravity waves reaches 10 % of the total non–zonal energy of the system. This level of contribution from the inertia–

gravity waves is also comparable with an estimate from global data analysis [*Zagar et al.*, 2009].

A simple extrapolation of the above result into the context of CQE suggests that even a system initialized without explicit convective modes may rapidly and continually develop fast convective modes. In the parameterization context, this implication might seem rather pessimistic because it suggests that it is fundamentally not possible to keep the fast convective processes implicit and completely exclude them from the parameterization.

It appears that the absence of the slow manifold *in the strict sense* is already well established in the literature. For this reason, a proposal has been made to replace the original concept of the slow manifold by the *slow quasimanifold*, a system consisting primarily of the slow modes, but allowing for finite departures within a stochastic layer [*Ford et al.*, 2000]. The notion of the quasimanifold is based on an anticipation that the departure is small enough that the system remains within a "fuzzy" zone close to the original manifold. A complementary view of this situation is that there are infinitely many, slightly–different slow manifolds that could be constructed, and that it is their non-uniqueness that leads to the notion of a quasimanifold [e.g., *Cox and Roberts*, 1994].

Applying the analogy to convection suggests that it may be important to take into account a "fuzzy" zone arising from convective–scale fluctuations that interact more directly with the large–scale processes in order to formulate CQE in a more robust manner.

In this respect, *Neelin and Zeng's* [2000] QTCM, mentioned in subsection 4.6, may arguably be considered as a slow manifold formulation for tropical large–scale circulations, at least in a conceptual sense. In their case, there is no specially–designed filtering procedure or initialization performed in order to maintain the system on slow manifold, but rather the fast modes are *effectively* eliminated by the "damping" provided by an adjustment form of the convection parameterization. This construction is almost guaranteed to avoid a problem with Lighthill's theorem, although gradual leaking from fast modes into slow modes by nonlinear interactions could still be an issue.

For a more general examination of the issue, we suggest that the convective energy–cycle system discussed in section 3.5 might fruitfully be considered under a coupling with large–scale dynamics. An important question to be addressed is iv) in section 3.4: if all linear perturbations around a CQE solution are damping, as is the case with QTCM, the system is almost guaranteed to stay on the slow manifold. However, if some perturbations turn out to be exponentially growing, the slow manifold is no longer a well–defined concept, but at least a "fuzzy" zone away from the strict slow manifold must be considered. The concept of self–organized criticality, to be discussed in section 6.3, would appear to provide a mechanism for establishing a "fuzzy" zone that is relatively narrow.

4.8. Equilibrium Control

Mapes [1997] has proposed an interpretation of CQE that is based on a lifting parcel argument. The interpretation is introduced by means of a contrast with an alternative principle, "activation control", which will be discussed in section 6.2.

Consider an air parcel which is being lifted upwards within an atmosphere that is conditionally unstable. The parcel first experiences negative buoyancy due to the adiabatic cooling associated with lifting, until it reaches saturation. Latent heating effects may then be sufficient to overcome adiabatic cooling, such that the parcel begins to feel positive buoyancy, and is accelerated upwards.

The initial energy barrier can be measured by the convective inhibition (CIN), a vertical integral of the buoyancy over the negative buoyancy zone. Further parcel ascent converts potential energy into kinetic energy, by following a downslope of the potential energy. A relevant question to ask is whether the generation rate of convective kinetic energy within the atmosphere is controlled primarily by "changes in amount of the downhill plunge" (equilibrium control) or by "the rate at which parcels are lifted over the activation energy barrier (*i.e.*, CIN) by intense small scale lifting processes" (activation control).

Mapes [1997] points out that "equilibrium control" lies behind many conceptions for deep convection. His section 4.1 lists six historical or observational points that he believes have led to the general adoption of the equilibrium control assumption, and he then argues that the evidence for equilibrium control is susceptible to other interpretations. One such interpretation is discussed in section 5.1.

4.9. Homeostasis

We close this section by introducing a very general concept from a very different discipline as a well–marked reference point for CQE. *Homeostasis* can be considered a biological counterpart for quasi–equilibrium. It refers to self–regulating processes in biological systems that maintain the constancy of properties such as acidity, salinity, and other compositional aspects of the blood, as well as body temperature, against changing external conditions. The concept especially refers to the ability of biological systems to maintain their stability against external perturbations such as abnormal food in–take or a change of external temperature. Etymologically, *homeostasis* consists of the two stems: prefix *homeo* means "similar" or "like" in Latin, whereas *stasis* comes from Greek meaning "standstill". Thus, as a whole, the word can be translated as *quasi–equilibrium.*

The concept was originally introduced by a physiologist, *Cannon* [1929, 1932]. In his own words, "The constant conditions which are maintained in the body might be termed *equilibra*. That word, however, has come to have fairly exact meaning as applied to relatively simple physico–chemical states, in closed systems, where known forces are balanced. The coordinated physiological processes which maintain most of the steady states in the organism are so complex and so peculiar to living beings — involving, as they may, the brain and nerves, the heart, lungs, kidneys and spleen, all working cooperatively — that I have suggested a special designation for these states, *homeostasis*."

By same token, although atmospheric convection may not be as complex as biological systems, it is far more complex than "relatively simple physico–chemical states". Moreover atmospheric convection is an open system like biological systems. In these respects, it may be more relevant to call it *convective homeostasis* rather than *convective quasi–equilibrium*. It may furthermore be worthwhile to recall that the Gaia hypothesis argues that the atmosphere is in homeostasis with Earth's biosphere [*Lovelock and Margulis*, 1974]. One might even speculate that convective homeostasis *may* contribute to homeostasis of the whole climate system [*cf.*, *Yano et al.*, 2012b].

At a very conceptual level, probably the thermodynamic analogy interpretation of CQE is the most well aligned with the concept of *homeostasis*, in the sense that it suggests stability of the system. This concept furthermore makes a very good counterpoint to the concept of self–organized criticality, which will be discussed in section 6.3.

However, a qualitative difference between *convective homeostasis* and *biological homeostasis* must be emphasized. In biological systems, homeostasis maintains an internal state (*e.g.*, constant body temperature) regardless of the external conditions (*e.g.*, how cold or warm the outside is). On the other hand, convective homeostasis must be environment dependent: the term may be a useful one in referring to a state that is uniquely defined by its environment and in which the stability of that state is maintained by its own self–regulation.

5. CRITICISMS OF CQE

The two most serious criticisms raised against CQE (or more accurately, against some of the interpretations of CQE) are the causality arguments of *Mapes* [1997] and the lack of evidence for a clear scale separation. In this section, we will consider those two criticisms in turn.

5.1. Cause and Effect?

As discussed above, various interpretations of CQE are based on the thermodynamic analogy developed in section 4.2, according to which the convection acts in direct response to an (assumed–known) large–scale forcing in order to maintain a state of quasi–equilibrium. Although it must be accepted that convective activity does feedback on the forcing (for example, through cloud-radiative interactions) nonetheless the forcing is essentially treated as an external constraint imposed on the convective system.

However, the observed smallness of the tendency of CAPE, or the cloud–work function, in comparison with the strength of the large-scale forcing, $dA_\lambda/dt \ll F_{L,\lambda}$ (Fig. 3), does not in itself say anything about the causality. *Mapes* [1997] points out that the same result would be entirely consistent with a complete reversal of the assumed causality. That is, instead of convection responding to the large scale, it could rather be that the large scale is responding to convective processes. The alternative point of view may be called "heating–response control".

In order to illustrate the apparent plausibility of this alternative interpretation, Mapes presented a Gedankenexperiment using a linear shallow–water model with a localized white–noise forcing. The shallow–water system has a long history as an analogue model for the tropical atmosphere [*e.g.*, *Gill*, 1980], while the white noise is designed to mimic random convective heating. Under this analogue model, CAPE is measured by fluctuations in the height of the shallow water. It was shown that the fluctuation of this analogue of CAPE becomes much smaller than the imposed forcing strength after times equivalent to a few hours. This happens because thermal anomalies generated by convective heating are rapidly smoothed out by gravity waves in the tropical atmosphere. As a result, a smooth temperature field is left behind once the gravity waves have re–adjusted the atmospheric thermodynamic structure in response to convective heating.

Although this alternative interpretation is fully consistent with Fig. 3, of course it does not follow that the equilibrium–control picture is necessarily wrong. The actual causality may be dominated by equilibrium–control, or heating–response control, or more plausibly, through a genuinely two–way interaction between convection and its environment (*cf.*, section 4.5). Indeed, as *Mapes* [1997] discussed, a suitable picture may depend on both the scale and on the phenomena of interest: the equilibrium–control picture has proved fruitful in seeking to understand aspects of global scale behavior, while at the other extreme, both observational and model-based case studies of individual convective storms are interested in exactly where and when a storm occurs and naturally adopt an heating–response control picture. Perhaps neither picture is truly satisfactory on its own to address other questions and other phenomena. These arguments seem to imply that a more general framework is needed: one which reduces to heating–response or equilibrium control in appropriate limits.

A good first step would be to look for data in which the outcome of very many individual heating–response controlled events can be shown to produce an equilibrium-control situation on a much broader scale. It could be fruitful to explore some recent large-domain, cloud-system resolving model data [e.g., *Shutts and Palmer*, 2007; *Liu et al.*, 2009; *Holloway et al.*, 2012] from this perspective.

An important issue with the experiment of *Mapes* [1997, 1998] is that the interpretation of the tropical atmosphere as being driven by white–noise convective forcing is not consistent with observations. A simple application of asymptotic analysis in *Yano et al.* [2000] shows that in that case the CAPE power spectrum must be proportional to the square of the frequency. That is not what is observed.

Yano et al. [2000] examine various alternatives. Among those, they show that when the large–scale forcing, F_L, is prescribed as in typical CRM simulations, convection actively responds in order to maintain the system close to the equilibrium as defined by Eq. (3.6). Slight deviations from this equilibrium behave as white noise, and as a result CAPE evolves as a Brownian motion. However, this is not what is observed either. Instead, the frequency spectrum of CAPE has the form of $1/f$–noise. This leads to the notion of self–organized criticality, to be discussed in section 6.3.

More importantly, we should recognize that *Mapes's* [1997] criticisms raise legitimate concerns about the validity (certainly about the range of validity) of the thermodynamic–analogy based interpretation of CQE, but they do not discredit the whole idea of CQE, especially if it is interpreted as a balance condition (*cf.*, section 4.6). The point of view of "heating–response control", *i.e.*, the large–scale dynamics responding to convective forcing, will be revisited in the context of the "free ride" principle in the discussions of section 6.1.

5.2. Absence of Scale Separation

A more fundamental obstacle for accepting CQE is the ostensible absence of a clear scale separation between convection and the large scales. *Yano et al.* [2000] strongly argue that if CQE is interpreted as convection responding very rapidly to slow large–scale processes, then we should be able to identify a fast convective adjustment time–scale. It follows that the CAPE timeseries should qualitatively behave like red noise with a damping time–scale characterized by a convective scale. However, as noted in the previous subsection, the observed CAPE timeseries does not exhibit such a damping time–scale, but rather it has a power–law spectrum.

Although one should be mindful that power–law behavior has undoubtedly been claimed too readily and too strongly in the scientific literature [*Clauset et al.*, 2009], nonetheless the ubiquitous presence of power laws and scaling behavior in the atmosphere does raise a major challenge to interpretations of CQE based on the scale separation principle. Observations of cloud fractality [*e.g.*, *Lovejoy*, 1982; *Yano and Takeuchi*, 1987] and analyses of precipitation timeseries [*e.g.*, *Peters et al.*, 2010 and references therein] may be taken as illustrative of the evidence for power laws in the atmosphere. Extensive recent reviews on observational evidence for atmospheric fractality have been given by *Tuck* [2008] and *Lovejoy and Schertzer* [2010].

Very recently, a careful analysis of a characteristic convective time–scale has been conducted by *Zimmer et al.* [2010]. Their time–scale is measured by

$$\tau \sim \mathrm{CAPE}/P$$

assuming that the precipitation rate P provides a measure of the rate of change of CAPE. The precipitation rate is an average value within a 50 km radius and 3 hr window of the radiosonde ascent that is used to calculate the CAPE. We have omitted a normalization factor from the above definition, which is required to obtain the correct dimensions, but obscures the main point. The frequency of occurrence of this time–scale shows a slow alegbraic decay, approximately following a slope of $\tau^{-1.3}$, over scales of 10^{-1} to 10^{4} hr: another example of scaling behavior.

However, we must distinguish between the *elusiveness* of the scale separation principle in the face of observed scale-free behavior for various aspects of the atmosphere, and the very clear usefulness of the scale separation principle. The principle may be *elusive*, but it may well be *useful*. Understanding the behavior of an idealized system in a suitable limit may of course provide valuable insights into the much more complicated behavior of a real system.

In order to illustrate these points, quasi–geostrophic theory again provides a good example. Arguably this theory is also based on a scale separation principle, by singling out the scale of the Rossby deformation radius as a characteristic scale for large–scale flow. However, observations do not single out this scale as having any particular importance in the face of atmospheric scaling behavior. Nonetheless most dynamicists would consider that the absence of clear, simple observational support in this sense hardly diminishes the usefulness of quasi–geostrophic theory. Singling out a particular scale is clearly "useful" in deriving the quasi–geostrophic system. Quasi–geostrophic theory is even capable of explaining the observed scaling behavior of kinetic energy [*cf.*, *Charney*, 1971].

By the same token, despite the lack of observational evidence for a single, simple characteristic convective time–scale, that does not exclude the usefulness of the concept for deriving a theoretical principle. The usefulness of the principle must then be judged *a posterori* from its applications, such as the performance of parameterizations. Again, it should be emphasized that the interpretation of CQE as a balance condition stands without invoking a scale separation principle.

The existence of meoscale organization provides a similar objection to the scale separation principle from phenomenology. A more basic issue with such organized structures is less the use of a quasi–equilibrium hypothesis, but rather the absence of an explicit representation of the structures in convection parameterizations. To the best of our knowledge, the only mass–flux parameterization to have included a mesoscale downdraft component is that of *Donner* [1993]. *Moncrieff's* [1981, 1992] archetype model could be a more formal answer to this challenge. However, no operational centers seem to have taken this proposal seriously to this date.

6. ALTERNATIVE THEORIES

6.1. A Link to the Notion of "Free ride"

Despite the caution expressed towards the end of section 4.1, it turns out that radiative convective equilibrium (RCE) as defined by Eq. (4.1) is generally a good approximation for large–scale tropical atmospheric processes. Almost any tropical sounding (but especially over the oceans) can demonstrate this point, as shown in Fig. 4 for example [reproduced from Fig. 1 of *Yano*, 2001]. It summarizes the relationship for very many soundings from the TOGA-COARE (Tropical Ocean Global Atmosphere–Coupled Ocean-Atmosphere Response Experiment) campaign. *Charney's* [1963] adiabatic scale analysis essentially demonstrates this point [*cf.*, *Yano et al.*, 2009; *Delayen and Yano*, 2009]. It is therefore tempting to apply the constraint as a dynamical balance condition for studying large–scale tropical circulations by analogy with geostrophic balance for mid–latitudes [*cf.*, *Yano and Bonazzola*, 2009]. The central importance of this balance is emphasized by *Neelin and Held* [1987]. *Fraedrich and McBride* [1989] propose to refer to the balance as a "free ride", whilst *Sobel et al.* [2001] in a different context introduce the name "weak temperature gradient" (WTG) approximation. The earlier terminology seems more intuitive for the purposes of the present discussion and so will be adopted here.

However, the application of the "free ride" balance as a large–scale dynamical constraint may not be appropriate if it is simultaneously to be used along with CQE as a parameterization closure. In order to demonstrate this point more explicitly, let us re–derive the RCE balance statement of Eq. (4.1) from a more formal statement of convective quasi–equilibrium, Eq. (3.6). As noted in section 3.2, the large–scale forcing term, $F_{L,\lambda}$, in Eq. (3.6) is dominated by the adiabatic cooling associated with large–scale uplift and radiative cooling ($-Q_R$). Thus,

$$F_{L,\lambda} \simeq -\int_0^H \frac{\rho g \hat{w}}{\theta} \left(w \frac{\partial \theta}{\partial z} - Q_R \right) dz \tag{6.1}$$

Here, the vertical integral is taken from the surface, for simplicity, to the top of the convection, denoted as the height H. The dominant contribution to the convective damping term, $D_{c,\lambda}$, is convective heating Q_c that stabilizes the atmosphere. Thus,

$$D_{c,\lambda} \simeq \int_0^H \frac{\rho g \hat{w}}{\theta} Q_c dz \tag{6.2}$$

Substitution of Eqs. (6.1) and (6.2) into Eq. (3.6) leads to

$$\int_0^H \frac{\rho g \hat{w}}{\theta} \left(-w \frac{\partial \theta}{\partial z} + Q_R + Q_c \right) dz = 0.$$

This condition must be satisfied for every vertical profile of convection, $\hat{w} = \hat{w}(z, \lambda)$. The possible profiles are scarcely an arbitrary set of test functions, but it is nonetheless reasonable to suppose that there are a sufficient number of sufficiently distinct profiles for the above integral constraint to be satisfied by a vanishing integrand at each vertical level. Thus,

$$-w \frac{\partial \theta}{\partial z} + Q_R + Q_c = 0 \tag{6.3}$$

so that the RCE balance of Eq. (4.1) is recovered, a balance that may be considered as an approximate simplification of CQE.

The basic idea of a "free ride" is to diagnose the vertical velocity w, given the total diabatic heating rate, $Q_R + Q_c$. On the other hand, the idea of applying CQE as a parameterization closure is to determine the convective heating rate, Q_c, given the two remaining terms in Eq. (6.3). Clearly one cannot use these two ideas at the same time. An alternative, and slightly more general way of looking at this issue is therefore to regard Eq. (6.3) simply as a balance condition relating the vertical velocity and the convective heating rate. The problem is clearly degenerate, just as with geostrophic balance.

6.2. Activation Control

"Activation–control" proposed by *Mapes* [1997] was introduced in section 4.8 as an alternative principle for the control of large–scale variations of deep convection. This principle emphasizes, in contrast to CQE ("equilibrium control"), the importance of overcoming an energy barrier of CIN. We now discuss the viability of this alternative principle.

We note first of all that the “activation control” idea focuses attention on shorter time–scale processes than those considered by *Arakawa and Schubert* [1974]. Recall from section 3.3 that their CQE hypothesis is formalized by assuming steady–state plumes from the outset. Under the steady–plume hypothesis, the vertical structure of the plume is assumed to be in equilibrium. By contrast, activation control focuses on an initial transient stage of convective growth, as a boundary-layer eddy breaks through the local energy barrier. Whether that short time–scale process has a key importance that needs to be recognized in the large–scale evolution of convective systems is an open question. The implications from Lighthill’s theorem discussed in section 4.7 remind us that the possibility cannot easily be excluded.

It is also interesting to note that the concept of activation control implicitly adopts the perspective of atmospheric convection as consisting of a series of ascending “bubbles”, rather than as an ensemble of steady plumes as assumed by *Arakawa and Schubert* [1974]. Historically speaking, bubble theory [*e.g.*, *Ludlam and Scorer*, 1953; *Scorer and Ludlam*, 1953; *Levine*, 1959] was seen a strong alternative theory for describing atmospheric convection. The idea was largely abandoned during the 1970s, although strong echoes persist to this day, most notably in the ongoing debates on entrainment [*e.g.*, *Blyth et al.*, 1988; *Heus and Jonker*, 2008] and in the usual textbook and lecture–course introductions of CAPE. *Mapes’s* [1997] arguments on activation control urge us to reconsider seriously the alternative possibility of bubble theory.

Unfortunately, the distinction is not always clear in *Mapes’s* [1997] article between two time–scales, that associated with individual convective plumes and that associated with the *ensemble* of convective plumes. Recall that the latter aspect is prognostically described by Eqs. (3.3) and (3.8). Issues raised by *Mapes* [1997] in his section 6 are not associated with lifting parcels, and could be equally well be interpreted in terms of the evolution of an ensemble of convective plumes as described by the prognostic equations [*cf.*, *Yano and Plant*, 2012a, b]. Note that in this description, CIN enters the problem only as a part of the vertical integral defining A_λ, and only if the vertical integral includes the boundary layer (*cf.*, Eq. 3.1). Recall that the lower limit, z_B, of the integral is usually taken at the cloud base, which is close to the top of the boundary layer [*cf.*, *Romps and Kuang*, 2011]. The issue raised in *Mapes's* [1997] section 6 is that variations in deep convection occur due to processes that simultaneously increase CAPE and reduce CIN so that it is ambiguous which controls the variation. An answer is suggested by the prognostic ensemble equations, Eqs. (3.1), (3.3) and (3.8): the combination that matters is dictated by a weighting function provided by the vertical plume profile.

Finally, a limit of thinking solely in terms of the CIN barrier should be emphasized. The barrier is typically expressed in terms of a neutrally–buoyant air parcel artificially lifted from the middle of a well–mixed boundary layer. The ascending air within a well–mixed boundary layer is likely to be positively buoyant. Thus when the buoyancy variable to be integrated over is defined as a weighted average using the vertical velocity of actual local air parcels, then we no longer see any negatively–buoyant barrier zone, except perhaps for an inversion layer at the top of the well–mixed layer. The result of such an analysis from CRM data is shown in Fig. 2 of *Yano* [2003], and Fig. 1 of *Yano* [2011]. Thus, the role of inhibition control is not as strong as it appears from a simple parcel analysis. This might also suggest that "activation control" is less important than it at first appears, although clearly it does not discredit the whole argument. The idea of "activation control" is also discussed in section 11.2 of *Emanuel* [1994] as a concept of "triggered convection". Some dimensional analyses are presented there.

Most importantly, if activation control were to be accepted as a guiding principle, then a rather drastic modification of the formulation of convection parameterization would be required. *Mapes* [1997, 1998] does not address such formulation issues. Unfortunately, subsequent attempts to implement the ideas [*Mapes*, 2000; *Kuang and Bretherton*, 2006; *Fletcher and Bretherton*, 2010; *Hohenegger and Bretherton*, 2011] have all been made within a traditional framework that assumes CQE. Thus, they lack in self-consistency, an issue that is further discussed in *Yano* [2011]. The concept of a trigger function as described by *Kain and Fritsch* [1992] for example can also be understood as arising from an activation control perspective [*cf.*, section 6.5].

6.3. Self–organized criticality

The concept of self–organized criticality (SOC) was originally proposed by *Bak et al.* [1987] to explain $1/f$–noise behavior in the sand–pile system. The concept refers to a state of a macroscopic system which is analogous to the critical state of a thermodynamic system at a phase transition [*cf.*, *Stanley*, 1972; *Yeomans*, 1992]. However, the major difference from the thermodynamic phase transition is that the system remains at, or close to, the critical state due to its self–maintaining tendency.

An SOC system remains at a type of *equilibrium* state called "criticality" but this is due more to the system's own critical behavior rather than externally–imposed conditions. In other words, it is not the environment that defines the equilibrium, as the conventional thermodynamic analogy suggests, but rather the internal dynamics. *Bak et al.* [1987] demonstrate their idea by introducing a simple multi–variable dynamical system that perpetually remains at a marginally unstable state. Common features of the textbook systems exhibiting SOC [*e.g.*, *Jensen*, 1998] are that the system is slowly–driven by some external forcing, with threshold behavior of the individual degrees of freedom, and that there are interactions between those degrees of freedom. Internal interactions drive the system towards criticality, developing large variability and structures on many scales without any need for external tuning.

Clearly SOC has a very different emphasis on interactions than the conventional interpretation of CQE, in which the convective plumes do not interact in any direct sense, only via their influence on the environment. Nonetheless, the system of equations given by (3.3) and (3.8), if suitably extended to incorporate a spatial aspect and hence a localized interaction, could be considered as a good starting point for theoretically considering a similar behavior for convection.

An important ingredient of SOC from a dynamical–systems point of view is its linear instability around the critical state. A numerical time–integration of a simplified version of the pair equation, (3.3) and (3.8), by *Wagner* [2010, but see also *Plant and Yano*, 2011] indeed suggests that this can be the case for a convective system (*cf.*, section 3.4). From the perspective of the slow manifold and Lighthill's theorem discussed in section 4.7, the convective system remains within a "fuzzy" interface zone due to the combination of its own instability and self–maintaining tendency. In both interpretations, an important contribution from fast convective processes is suggested unlike the conventional interpretation of CQE.

The state of SOC is often associated with $1/f$–noise behavior of the frequency spectrum (*i.e.*, the power spectrum is a power law with an exponent of -1), and thus findings of

$1/f$–noise behavior [*Yano et al.*, 2000, 2001, 2004] in various tropical time series, including CAPE, are suggestive that tropical convection is also at SOC.

Stronger evidence has more recently been found by *Peters and Neelin* [2006, see also *Peters et al.*, 2002; *Neelin et al.*, 2008]. The behavior of a system close to a state of criticality can be characterized by a power–law relationship

$$P \sim (I - I_c)^{\alpha}, \tag{6.4}$$

where I is a control variable, I_c the critical point defined in terms of this control variable, P a variable that represents self–organized behavior and α is a positive critical exponent less than one [*cf.*, Ch. 11, *Stanley*, 1972]. By identifying the column–integrated water vapor (CWV) and precipitation rate from satellite retrievals with I and P, respectively, they show that tropical convection exhibits such behavior with $\alpha = 0.215$.

However, some subsequent observational analyses and modeling simulations have had some difficulties in recovering the same results. *Raymond et al.* [2007, 2009] have pointed to qualitatively different behavior characterized by a relation

$$P \sim (I_c - I)^{-1}. \tag{6.5}$$

Other recent analyses include *Holloway and Neelin* [2009, 2010], who investigated ARM (Atmospheric Radiation Measurement) data but concluded that it was "impossible to test the power–law relationship at high total column water". A large–domain CRM experiment by *Posselt et al.* [2012] produced a scatter plot for precipitation that splits into two directions suggestive of the two possibilities in Eqs. (6.4) and (6.5). Global model data analysis by *Bechtold* [2009] does show a flattening tendency of the precipitation rate as a function of column integrated water for the largest values of column water (his Fig. 17), as would be expected from Eq. (6.4). However, at lower values, this flattening shape is preceded by a curve that is well fit by an exponential for precipitation rates varying over two orders of magnitude.

Note that the singular relationship (6.5) indicates a tendency for stabilization of the system by a negative feedback: as the column–integrated water (CWV) approaches I_c the precipitation rate dramatically increases. Such a tendency not only prevents the system from reaching the critical point, but it tends to stabilize the system by rejecting a highly moist state. This could be considered a good example of the *convective homeostasis* discussed in section 4.9, maintaining the stability of the system by self–regulation.

On the other hand, the relationship (6.4) indicates a critical behavior in the system, with a slower increase of the precipitation rate with increasing column–integrated water above I_c. As a result, above the critical point, the system will tend to accumulate more and more moisture into a given atmospheric column under sufficiently strong forcing. The accumulated column water is lost only gradually by precipitation. Such behavior is a reflection of the inherent instability of a system under SOC. It may furthermore be remarked that SOC is potentially important for understanding convective organization [cf., *Peters et al.*, 2009; *Yano et al.*, 2012a].

As discussed above, an SOC system is inherently unstable at the critical point and will therefore tend to evolve further away from that point, relative to the departures from the equilibrium point that might be expected for a thermodynamic system. In the latter case, one would expect to find a system that stays at its equilibrium point stably for substantial periods of time. To distinguish the two cases, it is important to study the frequency of occurrence of CWV. Such an analysis [*Neelin et al.*, 2008, 2009; *Lintner et al.*, 2011] reveals a Gaussian core but with tails that are much longer than would be expected from a Gaussian, suggesting some occasional, substantial deviations away from the critical point. Thus the identified critical point cannot be straightforwardly interpreted only in terms of a standard thermodynamic equilibrium state.

6.4. Activation and SOC: Complementary or contradictory with CQE?

Conceptually both activation control and SOC propose very different principles in comparison with CQE as interpreted through a thermodynamic analogy in section 4.2. Both of these alternative principles emphasize that convective processes are *not* passively defined as an equilibrium dictated by the given large–scale environment, but rather that they represent their own *autonomous* actions.

The activation–control principle emphasizes the importance of the local threshold: *i.e.*, the individual air parcel or boundary–layer eddy that triggers a convective element. However, it emphasizes less *how* an individual convective element modifies its environment, how such modifications may then affect subsequent triggering and consequently also how a system comprised of multiple convective elements behaves collectively. *Mapes* [1997] states that "clearly this situation is hopeless in detail" and so advocates an empirical approach.

On the other hand, SOC does emphasize the collective behavior of many convective elements. Individual convective elements are considered more as fluctuations. Details of their triggering and of their local environmental modifications are not considered to be important but rather the focus is on the collective behavior that emerges from the general character of their interactions. An ensemble average of the fluctuations provides crucial feedbacks to large–scale behavior due to nonlinear interactions between the convective elements. This is a qualitative difference of a critical phenomenon [*cf.*, *Wilson*, 1983] from a normal thermodynamic equilibrium. In the latter case, microscopic (convection) fluctuations may be simply averaged out at the macroscopic scale (large–scale).

Clearly activation control and SOC can be compatible: the former focussing on triggering while neglecting details of how collective behavior arises, while the latter ignores details of triggering and focuses on collective behavior. However, even if activation–control and SOC are relevant, then CQE considered as a balance condition, as discussed in sections 4.6 and 4.7, may nonetheless remain valid. In both of the alternative paradigms, however, the most serious implication is that fast convective processes have crucial impacts on the evolution of large–scale processes. Indeed SOC suggests that the equilibrium solution given by Eq. (3.6) would be unstable under linear perturbations. To what extent do we need to consider explicitly the fast, fluctuating processes? At the time of writing, it is not immediately clear whether we can still retain the quasi–equilibrium description as given by Eq. (3.6), or whether we have to move to a more prognostic or stochastic formulation.

Neelin et al. [2008] interprets that SOC can be regarded as an extension of an adjustment interpretation of CQE. In this respect, an important ingredient to be added to CQE in order to accomodate SOC is the transition to strong precipitating convection [*Neelin et al.*, 2009] above a critical threshold. This perspective is further discussed in section 6.6.

6.5. Phenomenological Limitations

Arguably the concept of CQE has been developed with the tropical atmosphere in mind, and mainly for maritime situations. In such situations, both temperature and moisture are relatively–speaking horizontally homogeneous, leading to the "free ride" principle and an equivalent balance for the moisture (*cf.*, section 6.1). These situations are consistent with CQE in observational diagnoses [*e.g.*, *Betts*, 1986, *Donner and Phillips*, 2003, *Holloway and Neelin*, 2007, *Zhang*, 2009]. However, the situations over land as well as in midlatitudes are likely to be very different. Both temperature and moisture are much more horizontally heterogeneous and it is less obvious how and when CQE is supportable from observational diagnoses [*e.g.*, *Zhang*, 2003, *Zimmer and Craig*, 2011].

Consider the situation over land in summer in the midlatitudes. The US Great Plains is perhaps the best studied example of this situation. It is phenomenologically known that (assuming fine weather) the surface heats up strongly during the day, so that CAPE becomes large around noon indicating a conditionally highly unstable atmosphere. Thus, a simple application of CQE would predict the onset of convection well before noon [*cf.*, *Guichard et al.*, 2004]. However, in practice convection is typically triggered in the late afternoon: thus, an external "triggering" (by either a synoptic or a boundary–layer process) seems to be required to realize the conditional instability.

A strict CQE hypothesis does not work in this type of situation, and many operational schemes introduce a trigger condition for just this reason. *Kain and Fritsch* [1992] demonstrate the sensitivity of mesocale simulations to the formulation of triggering in some circumstances. *Sud et al.* [1991] investigate the use in a GCM of critical onset values of the cloud work function in an adapted form of the *Arakawa and Schubert* [1974] parameterization. *Rogers and Fritsch* [1996] propose a general framework for trigger functions.

An alternative possibility for taking into account these phenomenological limitations is to adopt a prognostic energy–cycle description, as an extension of CQE, as discussed in section 3.5. This description can provide, at least, a partial answer to a phenomenologically–observed delayed onset of convection [*Yano and Plant*, 2012a]. The concept of transition to strong convection, to be discussed in the next section, may also be considered as an alternative possibility for overcoming phenomenological limitations of CQE.

6.6. Transition to Strong Convection

From a purely phenomenological point of view, probably the most important aspect revealed by a series of observational analyses initiated by *Peters and Neelin* [2006], which further led to a SOC interpretation as already discussed in section 6.3, is the fact that there is a well–defined onset of convection at $I = I_c$, beyond which the major proportion of tropical precipitation occurs. The concept of such an onset is something missing, or at least implicit, in the original sense of CQE.

Neelin et al. [2008] use the phrase "transition to strong convection" to describe the onset and further analyses examining its characteristics are presented by *Holloway and Neelin* [2009, 2010], *Neelin et al.* [2009] and *Sahany et al.* [2012]. By focussing on identifying the onset and its dependencies, these studies attempt to develop a phenomenological theory relatively independent of implications in terms of SOC. Note that the (I, P) space description characterizes convection as a function of a state (*i.e.*, column–integrated water vapor, or CWV) rather than of forcing (*cf.*, section 3.3).

Neelin et al. [2009] note that the critical behavior represented by Eq. (6.4) for $I > I_c$ can be interpreted as a nonlinear extension of a linear relaxation convection scheme originally developed by *Betts* [1986]. The scheme by *Betts* [1986] can be recovered as a special case of Eq. (6.4) with $\alpha = 1$ and so the transition from strong convection back to the onset state can be considered as being a natural extension of the adjustment re–interpretation discussed in section 4.5.

Neelin et al. [2009] and *Sahany et al.* [2012] furthermore show that variations in the onset value, I_c, can be defined to a good approximation as a function of the column–integrated tropospheric temperature, denoted $\widehat{T}$. Thus, the onset is characterized as a critical thermodynamic state in terms of both the column–integrated water vapor and temperature. Intriguingly, their analyses suggest that the onset is independent of the sea surface temperature (SST), which instead appears to be manifest as a stronger drive towards onset from below resulting in a frequency distribution of CWV that is shifted towards the onset boundary.

Holloway and Neelin [2009] examine the evolution of the vertical structure of the atmosphere associated with transition to convection by constructing various composites, emphasizing the importance of water vapor in the lower free troposphere. *Holloway and Neelin* [2010] focus more on lag–lead relationships between CWV and precipitation and so argue that high values of CWV occur primarily as a result of external forcing mechansisms rather than as a response to strong convection. *Holloway and Neelin* [2009] and *Sahany et al.* [2012] furthermore demonstrate that that the onset boundary in $(\mathrm{CWV}, \widehat{T})$ space can be approximately reproduced by some relatively simple bulk plume models, suggesting a link to conditional instability. The main requirement for the plume model is that it should have sufficiently strong entrainment in the lower free troposphere, so that the environmental water vapor then plays a sufficiently important role in the calculated plume buoyancy.

The emphasis on the onset of convection by these authors is, to some extent, reminiscent of points emphasized by the activation control principle [*Mapes*, 1997] discussed in section 6.2, and related to the trigger function [*Kain and Fritsch*, 1992] as discussed in the last section. On the other hand, there is an importance difference in these recent analyses. *Neelin et al.* (2008) emphasize that the concept of transition to strong convection need not be distinct from CQE but rather that is closely related to it [*cf.*, section 4.5].

7. CONCLUSIONS

The concept of convective quasi–equilibrium (CQE) originally proposed by *Arakawa and Schubert* [1974] has multiplied in its interpretations over the years. The purpose of the present review has been to provide a coherent picture of the various, sometimes competing, interpretations. For this purpose, a possibly–unusual historical–philosophical perspective has been taken.

It seems fair to say that the concept of thermodynamic equilibrium was developed from a tradition of French authors such as Carnot and Clausius. It is interesting then to note that the French word *équilibre* essentially corresponds to both equilibrium and balance in English. It may even be something of a historical accident that "équilibre thermodynamique" was translated into English as "thermodynamic equilibrium". According to the Oxford English dictionary, "equilibrium" dates back to 1608 in English, and its earliest textual references from around the 1660s onwards are clearly scientific. On the other hand, the word "balance" is known in English since the 13th century.

The word equilibrium has a more mystifying power in English than in French, and so perhaps does the notion of "convective quasi-equilibrium" (CQE). For this reason, we have extensively examined the connotations behind the concept of thermodynamic equilibrium first. We have also tried to suggest how these connotations have influenced thoughts on the concept of CQE.

In current English scientific language, the word "balance" is used for dynamical rather than thermodynamical concepts. We have also reviewed dynamical balance as a counterpart to thermodynamic equilibrium. The word "balance" is associated with fewer additional connotations than "equilibrium". Thus, if CQE had originally been simply coined as "convective quasi-balance" instead, its interpretation may have been less controversial.

With the given literature, there are two possible ways for interpreting Akio Arakawa's original philosophical argument for justifying the CQE hypothesis. First, an unbiased reading of section 7 of *Arakawa and Schubert* [1974] suggests that they have a thermodynamic analogy with the convective system in mind, an argument developed here in section 4.2. On the other hand, *Schubert* [2000] re–tells the history of the development of the CQE concept by Akio Arakawa by analogy with quasi–geostrophic theory. We have expanded this argument by invoking the concept of slow manifold in section 4.6. Akira Kasahara [1997, personal communication] also supports the latter view.

CQE is a key concept in order to understand the role of deep moist convection in the atmosphere. The concept serves a wide range of purposes. It has been used as a guiding principle to develop almost all convective parameterizations (especially for their closure) that are used in weather forecasting and climate modeling. More fundamentally, it also provides a basic theoretical framework in order to understand the role of convection in large–scale tropical dynamics. Although the concept is frequently invoked, there are different interpretations that may be relevant for different purposes. Unfortunately, rather few authors take care to explain their own interpretation before applying the concept, and as a result there now appears to be some confusion in some quarters.

The present review has attempted to consider the various interpretations of CQE as systematically as possible, as summarized in Fig. 5. However, because so much has already been said about CQE, it is fair to acknowledge that only selective materials have been examined. Our focus has been to examine the existing interpretations under the two basic interpretations identified above: as a thermodynamic analogy and as a dynamical balance. We have also remarked that a biological counterpart to quasi–equilibrium, *homeostasis*, may help our understanding of CQE.

Probably the main issue for CQE interpretations based on the thermodynamic analogy has been best expressed by *Mapes's* [1997, 1998] criticisms of the assumed causality. However, if CQE is interpreted as a dynamical balance condition, then no particular form of the causality has then to be assumed, and such criticisms immediately become irrelevant. Indeed, our review as a whole suggests that it would be more fruitful to consider CQE as being primarily a balance condition, with the thermodynamic analogy being a more specific view that can be useful in particular, more limited situations. The concept of a slow manifold provides a robust, but as–yet–to–be fully exploited, theoretical basis for developing CQE theories from a dynamical balance interpretation. *Neelin and Zeng's* [2000] QTCM may be considered as taking such a first step.

However, various obstacles for further developing CQE theories have also been identified. The most basic of these is the relationship between CQE and the "free ride" principle that has been used to constrain large–scale tropical dynamics. By referring to essentially the same balance as CQE, simultaneous use of these two principles may lead to a tautological situation in which neither the large–scale circulation nor the convective heating is predictable.

A potentially serious issue for CQE is the possibility that the atmospheric convective system is at self–organized criticality (SOC). In contrast with the thermodynamic analogy for CQE, this would suggest the need for explicit consideration of contributions from fast convective processes to the large–scale processes. A similar issue may be anticipated by analogy with the slow manifold, as a consequence of Lighthill's theorem. The latter theorem might help us to tame the issues arising from the possibility of convective SOC.

In developing this review, it comes as something of a surprise how little observational verification of CQE has been performed based on the original definition (Eq. 3.6) introduced by *Arakawa and Schubert* [1974]. CQE is based on the idea of posing a balance condition within an energy cycle description of the convective system. Within this energy cycle, the cloud work function plays a key role. However, strangely speaking, the cloud–work function budget equation (3.3) has never been comprehensively investigated, even for cloud types represented by simple entraining plumes as originally considered by *Arakawa and Schubert* [1974]. The cloud–work function has rarely been evaluated from observational data, the main exceptions being *Arakawa and Schubert* [1974], and *Lord and Arakawa* [1980]. Most of the observational analysis is instead for CAPE. Some recent work, such as that by *Zhang* [2009], has used a diluted parcel buoyancy and so comes closer to a cloud work function analysis, but dilute CAPE assumes that $\hat{w} = 1$ in Eq. (3.1). Other recent analyses based on CWV or other proxies may also have a closer link to the cloud work function, but systematic work function analyses are still awaited. Thus, we still lack an observationally robust basis to discuss the extent and range of validity of the originally–formulated quasi–equilibrium hypothesis.

This situation is unfortunate because the range of validity of CQE could in fact be wider than is generally supposed, and also because alternative paradigms such as activation–control might be able to be incorporated with a modest generalization of the energy cycle framework, as suggested in section 6.2. As an intermediate step, PEC (potential energy convertibility) was proposed by *Yano et al.* [2005] to provide a bulk estimate for the cloud work function from CRM experiments. Even for the diurnal cycle of US Great Plains, the cloud work function (as estimated by PEC) may provide a good measure for predicting the onset of afternoon convective precipitation, as suggested by a good positive correlation of PEC with the precipitation rate. As remarked in section 6.5, the midlatitude continental summer situation is often considered to be far from CQE, but to the best of our knowledge a careful budget analysis in terms of PEC or cloud work function is still to be performed.

There still remain many investigations to be performed at a theoretical level. For example, it may be revealing to formulate and study possible statements of a variational principle for the thermodynamic analogy to CQE. The structure and the invertibility of the $\mathcal{K}$ matrix, and the stability of the equilibrium state would also be valuable subjects of study. Moreover, although a re–interpretation of CQE as a slow manifold suggests the applicability of rich resources from dynamical systems thinking, as well as Hamiltonian dynamics, this simple point is yet to receive full attention.

Studies of atmospheric convection over many years have no doubt greatly enriched both our understandings and interpretations of convective quasi–equilibrium. Rich satellite data that has become available over the last decade for convection studies may especially be highlighted. Not least, this has led to the recent series of papers discussed in section 6.6 through which the concept of the "transition to strong convection" has emerged as promising direction.

However, our conceptual understanding of CQE is hardly converged. Our own point of view is that *Arakawa and Schubert's* [1974] equations defining the CQE hypothesis provide the basic, and far from exhausted, statement of this fundamental issue. We furthermore propose to make it customary to define the meaning of convective quasi–equilibrium in a given context whenever the phrase is invoked, because its meaning has so multiplied over the years that it can be hard to judge otherwise exactly what some authors mean by CQE. We hope that the present review serves as a baseline for clarifying the inter-related but diverged meanings used for CQE.

APPENDIX A: CANONICAL ENSEMBLE STATISTICS FOR CONVECTIVE MASS FLUXES

The purpose of this appendix is to provide a simple example to show how a variational principle can be applied to an ensemble of convective clouds. Specifically, we use this approach to re–derive an expression from *Craig and Cohen* [2006] (their Eq. (7)) for the distribution of convective mass fluxes in a equilibrium convective system.

For simplicity, suppose for the moment that each cloud has a discrete value of convective mass flux taken from the set $\{m_1, m_2, \cdots,\}$ where i is a whole number and $m_i = i\Delta m$. The continuous limit of $\Delta m \to 0+$ will be taken later to establish the final result.

The goal is to determine the most likely number n_i of convective clouds in each state i. Taking an analogy with the canonical ensemble in statistical mechanics, we assume that both the total number, N, of convective clouds and the total mass flux, M, are known in the system so that

$$\sum_i n_i = N, \tag{A.1a}$$

$$\sum_i n_i m_i = M. \tag{A.1b}$$

The number of ways of arranging a distribution $\mathbf{n} = \{n_i\}$ is

$$W(\mathbf{n}) = \frac{N!}{n_1!n_2!n_3!\cdots} \tag{A.2}$$

and the most likely state is obtained by invoking a variational principle: we maximize W subject to the constraints of (A.1a, b).

In practice, it is more convenient to consider variations of the entropy–like variable $\ln W$ rather than W itself. The constraints (A.1a, b) are taken into account using Lagrange's method of undetermined multipliers. Thus, the variational principle may be stated as

$$\delta\{\ln W + (\alpha - 1)\sum_i n_i - \beta \sum_i n_i m_i\} = 0$$

with $\alpha - 1$ and β being the multipliers. (The -1 included in the first multiplier is purely for algebraic convenience.) Since the above condition must be satisfied for all the possible changes of n_i, it is equivalent to

$$\frac{\partial}{\partial n_i}\{\ln W + (\alpha - 1)\sum_i n_i - \beta \sum_i n_i m_i\} = 0$$

By invoking Stirling's approximation

$$\ln n! \simeq n(\ln n - 1),$$

the above condition reduces to

$$-\ln n_i - \alpha - \beta m_i = 0$$

Thus, the most likely number in the state i is

$$n_i = \exp(-\alpha - \beta m_i)$$

The multipliers α and β are determined by the constraints (A.1a, b), and the final expression is made simpler by taking the continuous limit to produce

$$n_i = N\beta \exp(-\beta m)\Delta m$$

with $\beta = N/M$.

Note that this simple example does not answer the question of defining the large-scale equilibrium state for grid–box averaged variables. A variational approach could be developed for that aim, but additional constraints would have to be imposed that would account for the interactions between the clouds and the environmental state.

ACKNOWLEDGMENTS. We would like to acknowledge discussions with Laura Davies of Monash University on the range of interpretations of CQE, with Mick Short of the University of Lancaster on etymology, and also with Paul Williams and Maarten Ambaum of the University of Reading on slow manifolds and geostrophic adjustment. We also thank David Neelin and two anonymous reviewers for reviewing the manuscript. The present work has been supported by a UK–French collaboration sponsored by the Royal Society and CNRS, and is performed under a framework of COST Action ES0905.

REFERENCES

Adkins, C. J. (1983), *Equilibrium Thermodynamics*, 3rd Ed., Cambridge University Press.

Andrews, F. C. (1963), *Equilibrium Statistical Mechanics*, John Wiley & Sons.

Arakawa, A. (1969), Parameterization of cumulus convection. In: Proceedings of the WMO/IUGG Symposium on Numerical Weather Prediction, Tokyo, 26 Nov.–4 Dec., 1968, Japan Meteor. Agency, IV, 8, 1–6.

Arakawa., A. (2004), The cumulus parameterization problem: Past, present, and future. *J. Climate, 17*, 2493–2525.

Arakawa, A., and W. H. Schubert (1974), Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci., 31*, 674–701.

Bak, P., C. Tang, and K. Wiesenfeld (1987), Self-organized criticality: An explanation of the $1/f$ noise. *Phys. Rev. Lett., 59*, 381–384.

Ball, M. A., and R. S. Plant (2008), Comparison of stochastic parameterization approaches in a single-column model. *Phil. Trans. Roy. Soc. A, 366*, 2605-2623.

Bechtold, P., E. Bazile, F. Guichard, P. Mascart, and E. Richard (2001), A mass-flux convection scheme for regional and global models. *Quart. J. Roy. Meteor. Soc., 127*, 869 –889.

Bechtold, P. (2009), Convection parameterization. In: Proceedings of the ECMWF Seminar on Parameterization of Subgrid Physical Processes, 1–4 September 2008, Reading, UK
http://www.ecmwf.int/publications/library/ecpublications/_pdf/seminar/2008/Bechtold.pdf

Betts, A. K. (1986), A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quart. J. Roy. Meteor. Soc., 112*, 677–691.

Bihlo, A. (2010), Symmetry methods in the atmospheric sciences. PhD thesis, University of Vienna.

Blyth, A. M., W. A. Cooper, and J. B. Jensen (1988), A study of the source of entrained air in Montana cumuli. *J. Atmos. Sci., 45*, 3944–3964.

Cannon, W. B. (1929), Organization for physiological homeostasis. *Physiol. Rev., 9*, 399–431.

Cannon, W. B. (1932), *The Wisdom of the Body*, W. W. Norton, Inc., New York.

Carnot, S. (1824), Reflections on the motive power of fire, and on machines fitted to develop that power. Reprinted in Mendoza, E. (ed.): *Reflections on the Motive Power of Fire*, Dover Publications, New York (1960), 3–59.

Chandler, D. (1987), *Introduction to Modern Statistical Mechanics*, Oxford University Press.

Charney, J. G. (1963), A note on large–scale motions in the tropics. *J. Atmos. Sci., 20*, 607–609.

Charney, J. G. (1971), Geostrophic turbulence, *J. Atmos. Sci., 28*, 1087–1095.

Clapeyron, E. (1834), Memoir on the motive power of heat. Reprinted in Mendoza, E. (ed.): *Reflections on the Motive Power of Fire*, Dover Publications, New York (1960), 73–105.

Clauset, A., C. R. Shalizi and M. E. J. Newman (2009), Power-law distributions in empirical data. *SIAM Review 51*, 661–703.

Cohen, B. G., and G. C. Craig (2006), Fluctuations in an equilibrium convective ensemble. Part II: Numerical experiments. *J. Atmos. Sci., 63*, 2005–2015.

Cox, S. M., and Roberts, A. J. (1994), Initialization and the quasi-geostrophic slow manifold. *Preprint*, arXiv:nlin/0303011v1.

Craig, G. C., and B. G. Cohen (2006), Fluctuations in an equilibrium convective ensemble. Part I), Theoretical formulation. *J. Atmos. Sci., 63*, 1996–2004.

Davoudi, J., N. A. McFarlane, T. Birner (2010), Fluctuation of mass flux in a cloud-resolving simulation with interactive radiation. *J. Atmos. Sci., 67*, 400–418.

Davies, L., R. S. Plant, and S. H. Debyshire (2009), A simple model of convection with memory. *J. Geophys. Res., 114* (D17202), doc:10.1029/2008JD011653.

Davies, L., R. S. Plant, and S. H. Derbyshire (2012), Departures from convective equilibrium with a rapidly-varying surface forcing. Submitted to: *Quart. J. Roy. Met. Soc.*

Delayen, K., and J.-I. Yano (2009), Is asymptotic non–divergence of the large–scale tropical atmosphere consistent with equatorial wave theories? *Tellus 61A*, 491–497.

de Rooy, W. C. de, P. Bechtold, K. Froehlich, C. Hohenegger, H. J. J. Jonker, D. Mironov, A. P. Siebesma, J. Teixeira and J.-I. Yano (2012), Entrainment and detrainment in cumulus convection: an overview *Quart. J. Roy. Meteor. Soc.,* in press, DOI: 10.1002/qj.1959.

Donner, L. J. (1993), A cumulus parameterization including mass fluxes, vertical momentum dynamics, and mesoscale effects. *J. Atmos. Sci., 60*, 137–151.

Donner, L. J., and V. T. Phillips (2003), Boundary layer control on convective available potential energy: Implications for cumulus parameterization. *J. Geophys. Res., 108* D22, doi:10.1029/2003JD03773.

Emanuel, K. A. (1994), *Atmospheric Convection*, Oxford University Press.

Emanuel, K. (2000), Quasi–equilibrium thinking. In: *General Circulation Model Development* (D. A. Randall, Ed.), Elsevier, Ch. 8, 225–255.

Emanuel, K. A. (2007). Quasi–equilibrium dynamics of the tropical atmosphere. In: *The Global Circulation of the Atmosphere* (T. Schneider and A. H. Sobel, Eds.), Princeton University Press, Ch. 7, 186–218.

Emanuel, K. A., J. D. Neelin, and C. S. Bretherton (1994), On large–scale circulation in convective atmospheres. *Quart. J. Roy. Meteor. Soc., 120*, 1111–1143.

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd Ed., John Wiley and Sons.

Fierro, A. O., R. F. Rogers, F. D. Marks, D. S. Nolan (2009), The impact of horizontal grid spacing on the microphysical and kinematic structures of strong tropical cyclones simulated with the WRF-ARW model. *Mon. Wea. Rev., 137*, 3717–3743.

Fierro, A. O., E. J. Zipser, M. A. LeMone, J. M. Straka, J. (Malkus) Simpson (2012), Tropical oceanic hot towers: Need they be undilute to transport energy from the boundary layer to the upper troposphere effectively? An answer based on trajectory analysis of a simulation of a TOGA COARE convective system. *J. Atmos. Sci., 69*, 195–213.

Fletcher, J. K., and C. S. Bretherton (2010), Evaluating boundary layer–based mass flux closures using cloud–resolving model simulations of deep convection. *J. Atmos. Sci., 67*, 2212–2225.

Ford, R. (1994), Gravity wave radiation from vortex trains in rotating shallow water. *J. Fluid Mech., 281*, 81–118.

Ford, R., M. E. McIntyre, and W. A. Norton (2000), Balance and the slow quasimanifold: Some explicit results. *J. Atmos. Sci., 57*, 1236–1254.

Fraedrich, K., and J. L. McBride (1989), The physical mechanism of CISK and the free-ride balance. *J. Atmos. Sci., 46*, 2642–2648.

Fritsch, J. M., and C. F. Chappell (1980), Numerical prediction of convectively driven mesoscale pressure systems. Part i: convective parameterization. *J. Atmos. Sci., 37*, 1722–1733.

Gill, A. E. (1980), Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc., 449*, 447–462.

Goldstein, H., C. Poole, and J. Safko (2002), *Classical Mechanics*, 3rd Ed. Addison Wesley, San Francisco.

Gombosi, T. I. (1994), *Gaskinetic theory.* Cambridge University Press.

Grabowski, W. W. (2003), Impact of cloud microphysics on convective–radiative quasi–equilibrium revealed by cloud-resolving convection parameterization. *J. Climate, 16*, 3463–3475.

Gregory, D., and P. R. Rowntree (1990), A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Wea. Rev., 118*, 1483–1506.

Gregory, D., J.–J. Morcrette, C. Jacob, A. C. M. Beljaars, and T. Stockdale (2000), Revision of convection, radiation and cloud schemes in the ECMWF Integrated Forecast System. *Quart. J. Roy. Meteor. Soc., 126*, 1685–1710.

Greenspan, H. P. (1968), *The Theory of Rotating Fluids.* Cambridge University Press, New York.

Groenemeijer, A. P., and G. C. Craig (2012), Ensemble forecasting with a stochastic convective parametrization based on equilibrium statistics. *Atmos. Chem. Phys., 12*, 4555-4565.

Guichard F., J. C. Petch, J. L. Redelsperger, P. Bechtold, J.–P. Chaboureau, S. Cheinet, W. Grabowski, H. Grenier, C. G. Jones, M. Kohler, J.–M. Piriou, R. Tailleux, M. Tomasini (2004), Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quart. J. Roy. Meteor. Soc., 604*, 3139–3172.

Guha, P. (2007). Metriplectic structure, Leibniz dynamics and dissipative systems. *J. Math. Anal. Appl.,*, **326**, 121-136.

Heus, T., and H. J. J. Jonker (2008), Subsiding shells around shallow cumulus clouds. *J. Atmos. Sci., 65*, 1003-1018.

Hepburn, B. S. (2007), Equilibrium and explanation in 18th century mechanics. PhD thesis, University of Pittsburgh.

Hildebrandt, S. and Tromba, A. (1996), *The Parsimonious Universe: Shape and Form in the Natural World*. Springer-Verlag, New York.

Hohenegger, C., and C. S. Bretherton (2011), Simulating deep convection with a shallow convection scheme. *Atmos. Chem. Phys., 11*, 10389–10406.

Holloway, C. E., and J. D. Neelin (2007), The convective cold top and quasi–equilibrium. *J. Atmos. Sci., 64*, 1467–1487.

Holloway, C. E., and J. D. Neelin (2009), Moisture vertical structure, column water vapor, and tropical deep convection. *J. Atmos. Sci., 66*, 1665–1683.

Holloway, C. E., and J. D. Neelin (2010), Temporal relations of column water vapor and tropical precipitation. *J. Atmos. Sci., 67*, 1091–1105.

Holloway, C. E., S. J. Woolnough and G. M. S. Lister (2012), Precipitation distributions for explicit versus parameterized convection in a large-domain high-resolution tropical case study. To appear in: *Quart. J. Roy. Meteor. Soc.*, doi: 10.1002/qj.1903

Jensen, H. J. (1998), *Self-organized criticality.* Cambridge University Press.

Jones, T. R., and Randall, D. A. (2011), Quantifying the limits of convective parameterizations. *J. Geophys. Res., 116* (D08210), doi:10.102910/2010JD014913.

Kain, J. S., and J. M. Fritsch (1990), A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci., 47*, 2784–2802.

Kain, J. S., and J. M. Fritsch (1992), The role of the convective “trigger function” in numerical forecasts of mesoscale convective systems. *Meteorol. Atmos. Phys., 49*, 93–106.

Khouider, B., Biello, J. A. and Majda, A. J. (2010), A stochastic multicloud model for tropical convection. *Comm. Math. Sci., 8*, 187–216.

Kondepudi, D., and I. Prigogine (1998), *Modern Thermodynamics: From Heat Engines to Dissipative Structures*, John Wiley & Sons.

Kuang, Z. (2011), The wavelength dependence of the gross moist stability and the scale selection in the instability of column-integrated moist static energy. *J. Atmos. Sci., 68*, 61–74.

Kuang, Z., and C. S. Bretherton (2006), A mass–flux scheme view of a high–resolution simulation of a transition from shallow to deep cumulus convection. *J. Atmos. Sci., 63*, 3421–3436.

Kuo, H. L. (1974), Further studies of the parameterization of the influence of cumulus clouds on the large-scale flow. *J. Atmos. Sci., 31*, 1232-1240.

Lander, J. L., and B. J. Hoskins (1997), Believable scales and parameterizations in a spectral transform model. *Mon. Wea. Rev., 125*, 292–303.

Landau, L. D., and Lifshitz, E. M. (1980), *Statistical Physics Part 1*, 3rd Ed., Course of Theoretical Physics Volume 5, Elsevier.

Leith, C. E. (1980), Nonlinear normal mode initialization and quasi–geostrophic theory. *J. Atmos. Sci., 37*, 958–968.

Levine, J. (1959), Spherical vortex theory of bubble–like motion in cumulus clouds. *J. Meteor., 16*, 653–662.

Lighthill, M. J. (1952), On sound generated aerodynamically. I. General theory. *Proc. Roy. Soc. Lond., A211*, 564–587.

Lighthill, M. J. (1954), On sound generated aerodynamically. II. Turbulence as a source of sound. *Proc. Roy. Soc. Lond., A222*, 1–32.

Lintner, B. R., C. E. Holloway, and J. D. Neelin (2011), Column water vapor statistics and their relationship to deep convection, vertical and horizontal circulation, and the moisture structure of Nauru. *J. Climate, 24*, 5454–5466.

Liu, P., M. Satoh, B. Wang, H. Fudeyasu, T. Nasuno, T. Li, H. Miura, H. Taniguchi, H. Masunaga, X. Fu and H, Annamalai (2009), An MJO simulated by the NICAM at 14- and 7-km resolutions. *Mon. Wea. Rev., 137*, 3254–3268.

Lord, S. J. (1982), Interaction of a cumulus cloud ensemble with the large-scale environment. Part III: Semi-prognostic test of the Arakawa-Schubert cumulus parameterization. *J. Atmos. Sci., 39*, 88–103.

Lord, S. J., and A. Arakawa (1980), Interaction of a cumulus cloud ensemble with the large-scale environment, Part II. *J. Atmos. Sci., 37*, 2677–2692.

Lord, S. J., W. C. Chao and A. Arakawa (1982), Interaction of a cumulus cloud ensemble with the large-scale environment, Part IV. *J. Atmos. Sci., 39*, 104–113.

Lorenz, E. N. (1986), On the existence of a slow manifold. *J. Atmos. Sci., 43*, 1547–1557.

Lorenz, E. N. (1992), The slow manifold — What is it? *J. Atmos. Sci., 49*, 2449–2451.

Lovejoy, S. (1982), Area–perimeter relation for rain and cloud areas. *Science*, *216*, 185–187.

Lovejoy, S., and D. Schertzer (2010), Towards a new synthesis for atmospheric dynamics: Space–time cascades. *Atmos. Res.,96*, 1–52.

Lovelock, J. E., and L. Margulis (1974), Atmospheric homeostasis by and for biosphere - Gaia hypothesis. *Tellus, 26*, 2–10.

Ludlam, F. H., and R. S. Scorer (1953), Convection in the atmosphere. *Quart. J. Roy. Meteor. Soc., 79*, 317–341.

Majda, A. J. and B. Khouider (2002), Stochastic and mesoscopic models for tropical convection. *Proc. Nat. Acad. Sci.,99*, 1123–1128.

Manabe, S., and R. F. Strickler (1964), Thermal equilibrium of the atmosphere with a convective adjustment. *J. Atmos. Sci., 21*, 361–385.

Manabe, S., J. S. Smagorinsky, and R. F. Strickler (1965), Simulated climatology of a general circulation model with a hydrological cycle. *Mon. Wea. Rev., 93*, 769–798.

Mapes, B. E. (1997), Equilibrium vs. activation controls on large–scale variations of tropical deep convection. In: *the Physics and Parameterization of Moist Atmospheric Convection*, (R. K. Smith, Ed.), NATO ASI, Kloster Seeon, Kluwer Academic Publishers, Dordrecht, 321–358.

Mapes, B. E. (1998), The large–scale part of tropical mesoscale convective system circulations: A linear vertical spectrum band model. *J. Met. Soc. Japan 76*, 29–55.

Mapes, B. E. (2000), Convective inhibition, subgrid-scale triggering energy, and stratiform instability in a toy tropical wave model. *J. Atmos. Sci., 57*, 1515–1535

McKane, A. J., and T. J. Newman (2005), Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett., 94*, 218102.

McIntyre, M. (2001), Balance, potential-vorticity inversion, Lighthill radiation and the slow quasimanifold. In: Proceedings of the IUTAM/IUGG/Royal Irish Academy Symposium on Advances in Mathematical Modelling of Atmosphere and Ocean Dynamics, Limerick, Ireland, (P. F. Hodnett, Ed.), 45–68. Kluwer Academic Publishers.

Möller, F., and S. Manabe (1961), Über das Strahlungsgleichgewicht in der Atmosphäre. *Z. Meteor., 15*, 3–8.

Moncrieff, M. W. (1981), A theory of organized steady convection and its transport properties. *Quart. J. Roy. Meteor. Soc., 107*, 29–50.

Moncrieff, M. W. (1992), Organized convective systems: Archetypal dynamical models, mass and momentum flux theory, and parametrization. *Quart. J. Roy. Meteor. Soc., 118*, 819–850.

Moorthi, S., and M. J. Suarez (1992), Relaxed Arakawa-Schubert. A parameterization of moist convection for general circulation models. *Mon. Wea. Rev., 120*, 978–1002.

Neelin, J. D. (1997), Implications of convective quasi–equilibrium for the large–scale flow. In: *the Physics and Parameterization of Moist Atmospheric Convection*, (R. K. Smith, Ed.), NATO ASI, Kloster Seeon, Kluwer Academic Publishers, Dordrecht, 413–446.

Neelin, J. D., and I. M. Held (1987), Modeling tropical convergence based on the moist static energy budget. *Mon. Wea. Rev., 115* 3–12

Neelin, J. D., and J.-Y. Yu (1994), Modes of tropical variability under convective adjustment and the Madden–Julian oscillation. Part I: Analytical theory. *J. Atmos. Sci., 51*, 1876–1894.

Neelin, J. D., and N. Zeng (2000), The first quasi-equilibrium tropical circulation model–formulation. *J. Atmos. Sci 57*, 1741–1766.

Neelin, J. D., O. Peters, J. W.–B. Lin, K. Hales, and C. E. Holloway (2008), Rethinking convective quasi–equilibrium: Observational constraints for stochastic convective schemes in climate models. *Phil. Trans. Roy. Soc., A366*, 2581–2604.

Neelin, J. D., O. Peters, and K. Hales (2009), The transition to strong convection. *J. Atmos. Sci., 66*, 2367–2384.

Ooyama, K. V. (1982), Conceptual evolution of the theory and modeling of the tropical cyclone, *J. Met. Soc. Japan 60*, 369–380.

Pan, D.-M., and D. A. Randall (1998), A cumulus parameterization with prognostic closure, *Quart. J. Roy. Meteor. Soc., 124*, 949–981.

Parodi, A., and K. Emanuel (2009), A theory for buoyancy and velocity scales in deep moist convection. *J. Atmos. Sci., 66*, 3449–3463.

Peters, O., and D. Neelin (2006), Critical phenomena in atmospheric precipitation. *Nature Phys., 2*, 393–396.

Peters, O., C. Hertlein, and K. Christensen (2002), A complexity view of rainfall. *Phys. Rev. Lett., 88*, 018701, doi:10.1103/PhysRevLett.88.018701

Peters, O., J. D. Neelin, and S. W. Nesbitt (2009), Mesoscale convective systems and critical clusters. *J. Atmos. Sci., 66*, 2913–2924.

Peters, O., A. Deluca, A. Corral, J. D. Neelin, and C E Holloway (2010), Universality of rain event size distributions. *J. Stat. Mech. Theor. Exp., P11030*. doi:10.1088/1742-5468/2010/11/P11030

Planck, M. (1922), *Treatise on Thermodynamics.* 7th edition. Translation by Ogg, A. (1926), 3rd edition translated from the seventh German edition. Dover Publications, New York.

Plant, R. S. and G. C. Craig (2008), A stochastic parameterization for deep convection based on equilibrium statistics. *J. Atmos. Sci., 65*, 87–105.

Plant, R. S. (2009), Statistical properties of cloud lifecycles in cloud-resolving models. *Atmos. Chem. Phys., 9*, 2195–2205.

Plant, R. S. (2010), A review of the theoretical basis for bulk mass flux convective parameterization. *Atmos. Chem. Phys., 10*, 3529–3544.

Plant R. S. (2012), A new modelling framework for statistical cumulus dynamics. *Phil. Trans. Roy. Soc., A370*, 1041–1060.

Plant, R. S., and J.-I. Yano (2011), Comment on "An ensemble cumulus convection parameterisation with explicit cloud treatment" by T. M. Wagner and H.-F. Graf. *J. Atmos. Sci., 68*, 1541–1544.

Posselt, D. J., S. van den Heever, G. Stephens, and M. R. Igel, (2012), Changes in the interaction between tropical convection, radiation, and the large-scale circulation in a warming environment. *J. Climate, 25*, 557–571.

Randall, D. A., and D.-M. Pan (1993), Implementation of the Arakawa-Schubert cumulus parameterization with a prognostic closure. In: *The Representation of Cumulus Convection in Numerical Models*, (Emanuel, K. A., and D. J. Raymond, Eds.), Meteorological Monographs No. 46, Amer. Meteor. Soc., 137–144.

Randall, D. A., D.–M. Pan, P. Ding, and D. G. Gripe (1997), Quasi-equilibrium. In: *the Physics and Parameterization of Moist Atmospheric Convection*, (R. K. Smith, Ed.), NATO ASI, Kloster Seeon, Kluwer Academic Publishers, Dordrecht, 359–385.

Raymond, D. J. (1995), Regulation of moist convection over the warm tropical oceans. *J. Atmos. Sci., 52*, 3945–3959.

Raymond, D. J., and M. J. Herman (2011), Convective quasi-equilibrium reconsidered. *J. Adv. Model. Earth Syst., 3*, Art. 2011MS000079, 14 pp, doi:10.1029/2011MS000079.

Raymond, D. J., S. L. Sessions, and Z. Fuchs (2007), A theory for the spinup of tropical depressions. *Quart. J. Roy. Meteor. Soc., 133*, 1743–1754.

Raymond, D. J., S. Sessions, A. Sobel and Z. Fuchs (2009), The mechanics of gross moist stability. *J. Adv. Model. Earth Syst., 1*, Art. #9, 20 pp., doi:10.3894/JAMES.2009.1.9

Ring, D. (2009), Non–linear wave interactions in rotating stratified fluid flows. PhD Dissertation, John Hopkins University.

Roff, G. L., and J.-I. Yano (2002), Tropical convective variability in the CAPE phase space. *Quart. J. Roy. Meteor. Soc., 128*, 2317–2333.

Rogers, R.F., and J. M. Fritsch (1996), A general framework for convective triggering functions. *Mon. Wea. Rev.,124*, 2438–2452.

Romps, D. M., and Z. Kuang (2010), Do undiluted convective plumes exist in the upper tropical troposphere? *J. Atmos. Sci., 67*, 468–484.

Romps, D. M., and Z. Kuang (2011), A transilient matrix for moist convection. *J. Atmos. Sci., 68*, 2009–2025.

Rossby, G.-G. (1938), On the mutual adjustment of pressure and velocity distribution in certain simple current systems. II. *J. Mar. Res., 1*, 239–263.

Sahany, S., J. D. Neelin, K. Hales, and R. Neale (2012), Temperature–moisture dependence of the deep convective transition as a constraint on entrainment in climate models. *J. Atmos. Sci., 69*, 1340–1358.

Salmon, R. (1998), *Lectures on Geophysical Fluid Dynamics*. Oxford University Press.

Scorer, R. S., and F. H. Ludlam (1953), Bubble theory of penetrative convection. *Quart. J. Roy. Meteor. Soc., 79*, 94–103.

Schubert, W. H. (2000), A retrospective view of Arakawa's ideas on cumulus parameterization. In: *General Circulation Model Development* (D. Randall, Ed.), Elsevier, Ch. 6, 181–198.

Sessions, S. L., Sugaya, S., Raymond, D. J., and Sobel, A. H. (2010), Multiple equilibria in a cloud-resolving model using the weak temperature gradient approximation. *J. Geophys. Res., 115*(D12110), doi.10.1029/2009JD013376.

Shutts, G. J. and Gray, M. E. B. (1999), Numerical simulations of convective equilibrium under prescribed forcing. *Quart. J. Roy. Meteor. Soc., 125* 2767–2787.

Shutts, G. J. and Palmer, T. N. (2007), Convective forcing fluctuations in a cloud-resolving model: Relevance to the stochastic parameterization problem. *J. Climate, 20*, 187–202.

Sobel, A. H., J. Nilsson, and L. M. Polvani (2001), The weak temperature gradient approximation and balanced tropical moisture waves. *J. Atmos. Sci., 58*, 3650–3665.

Solé, R. V., and J. Bascompte (2006), *Self–Organization in Complex Ecosystems*. Princeton University Press.

Stanley, H. E. (1972), *Introduction to Phase Transition and Critical Phenomena*. Oxford University Press.

Stephens, G. L., S. van den Heever, and L. Pakula (2008), Radiative–convective feedbacks in idealized states of radiative–convective equilibrium. *J. Atmos. Sci., 65*, 3899–3916.

Sud, Y. C., W. C. Chao, and G. K. Walker, (1991), Contribution to the implementation of the Arakawa–Schubert cumulus parameterization in the GLA GCM. *J. Atmos. Sci.,48*, 1573–1586.

Sui, C. H., K. M. Lau, W. K. Tao, and J. Simpson (1994), The tropical water and energy cycles in a cumulus ensemble model. Part I: Equilibrium climate. *J. Atmos. Sci., 51*, 711–728.

Tuck, A. *Atmospheric turbulence: A molecular dynamics perspective*, Oxford University Press.

van Kampen, N. G. (1985), Elimination of fast variables. *Phys. Rep., 124*, 71–160.

van Kampen, N. G. (2007), *Stochastic Processes in Physics and Chemistry*, 3rd Ed., Elsevier.

Wagner, T. M. (2010), A dynamical convective cloud field model and the effects of aerosols. PhD thesis, University of Cambridge.

Wilson, G. K. (1983), The renormalization group and critical phenomena. *Rev. Mod. Phys., 55*, 583–600.

Xu, K.-M., A. Arakawa, and S. K. Krueger (1992), The macroscopic behavior of cumulus ensembles simulated by a cumulus ensemble model. *J. Atmos. Sci., 49*, 2402–2420.

Xu, K.-M., and K. A. Emanuel (1989), Is the tropical atmosphere conditionally unstable? *Mon. Wea. Rev., 117*, 1471–1479.

Yanai, M. S., S. Esbensen and J. H. Chu (1973), Determination of bulk properties of tropical cloud clusters from large–scale heat and moisture budgets. *J. Atmos. Sci., 30*, 611–627.

Yano, J.-I. (1999), Scale–separation and quasi–equilibrium principles in Arakawa and Schubert's cumulus parameterization. *J. Atmos. Sci., 56*, 3821–3823.

Yano, J.-I. (2001), Residual cumulus parameterization. *Quart. J. Roy. Meteor. Soc., 127*, 1261–1276.

Yano, J.-I. (2003), The cumulus parameterization problem in the context of MJO simulations. In: Proceedings for the MJO workshop, ECMWF, 2–5 November 2003, 115–127.

http://www.ecmwf.int/publications/library/do/references/list/17124

Yano, J.–I. (2011), Interactive comment on "Simulating deep convection with a shallow convection scheme" by C. Hohenegger and C. S. Bretherton. Atmos. Chem. Phys. Discuss.,11, C2411–C2425.

Yano, J. I., and Y. Takeuchi (1987), The self-similarity of horizontal cloud pattern in the intertropical convergence zone, *J. Met. Soc. Japan 65*, 661-667.

Yano, J.-I., M. W. Moncrieff, and J. C. McWilliams (1998), Linear stability and single–column analyses of several cumulus parameterization categories in a shallow–water model. *Quart. J. Roy. Meteor. Soc., 124* 983–1005.

Yano, J.-I., W. W. Grabowski, G. L. Roff, and B. E. Mapes (2000), Asymptotic approaches to convective quasi–equilibrium. *Quart. J. Roy. Meteor. Soc., 126*, 1861–1887.

Yano, J.-I., K. Fraedrich, and R. Blender (2001), Tropical convective variability as $1/f$–noise. *J. Climate, 14*, 3608–3616.

Yano, J.-I., R. Blender, C. Zhang, and K. Fraedrich (2004), $1/f$ noise and pulse–like events in the tropical atmospheric surface variabilities. *Quart. J. Roy. Meteor. Soc., 300*, 1697–1721.

Yano, J.-I., and M. Bonazzola (2009), Scale analysis for the large-scale tropical atmospheric dynamics. *J. Atmos. Sci., 66*, 159–172.

Yano, J.-I., J.–P. Chaboureau, and F. Guichard (2005), A generalization of CAPE into potential–energy convertibility. *Quart. J. Roy. Meteor. Soc., 131*, 861–875.

Yano, J.-I., S. Mulet, and M. Bonazzola (2009), Tropical large–scale circulations: Asymptotically nondivergent? *Tellus, 61A*, 417–427.

Yano, J.-I., and R. S. Plant (2012a), Finite departure from convective quasi-equilibrium: Periodic cycle and discharge-recharge mechanism. *Quart. J. Roy. Meteor. Soc., 128*, 626–637.

Yano, J.-I., and R. S. Plant (2012b), Interactions between shallow and deep convection under a finite departure from convective quasi–equilibrium. *J. Atmos. Sci.*, accepted.

Yano, J.–I., H.-F. Graf, and F. Spineanu (2012a), Meeting report: Theoretical and operational implications of atmospheric convective organization. *Bull. Amer. Meteor. Soc.*,, *93*, ES39-ES41, doi: http://dx.doi.org/10.1175/BAMS-D-11-00178.1

Yano, J.–I., C. Liu, and M. W. Moncrieff (2012b), Atmospheric convective organization: homeostasis or criticality? accepted to *J. Atmos. Sci.*,.

Yeomans, J. (1992), *Statistical Mechanics of Phase Transitions*, Oxford University Press.

Yu, J-Y., and J. D. Neelin (1994), Modes of tropical variability under convective adjustment and the Madden–Julian oscillation. Part II: Numerical results. *J. Atmos. Sci., 51*, 1895–1914.

Zagar, N., J. Tribbia, J. L. Anderson, and K. Raeder (2009), Uncertainties of estimates of inertia–gravity energy in the atmosphere. Part I: Intercomparison of four analysis systems. *Mon. Wea. Rev., 137*, 3837–3857.

Zhang, G. (2002), Convective quasi–equilibrium in midlatitude continental environment and its effect on convective parameterization. *J. Geophys. Res., 107*(D14), doi:10.1029/2001JD001005.

Zhang, G. (2003), Convective quasi–equilibrium in the tropical western Pacific. Comparison with midlatitude continental environment. *J. Geophys. Res., 108* (D19), doc.10.1029/2003/JD003520.

Zhang, G. J. (2009), Effects of entrainment on convective available potential energy and closure assumptions in convection parameterization. *J. Geophys. Res., 114*, D07109, doc.10.1029/2008/JD010976.

Zhang, G. J., and N. A. McFarlane (1995), Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre General Circulation Model. *Atmos.–Ocean, 33*, 407–446.

Zimmer, M., Craig, G. C., Wernli, H. and Keil, C. (2011), Classification of precipitation events with a convective response timescale. *Geophys. Res. Lett., 38*, L05802, doi:10.1029/2010GL046199.

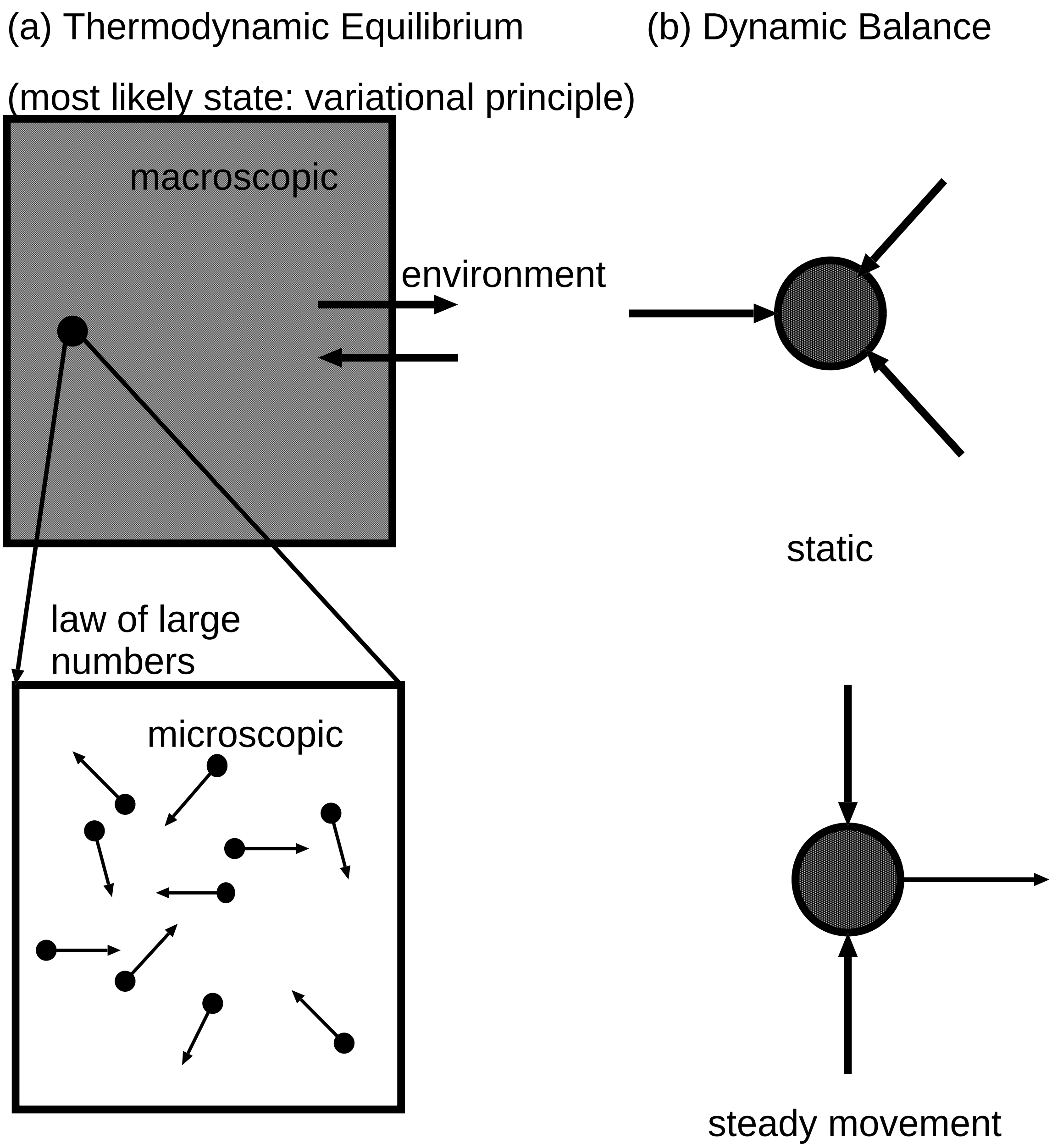

Figure 1. Schematics to illustrate the difference between (a) thermodynamic equilibrium and (b) dynamic balance. Richer implications of thermodynamic equilibrium are seen (a relationship with the environment, between macroscopy and microscopy, and the role of the law of large numbers).

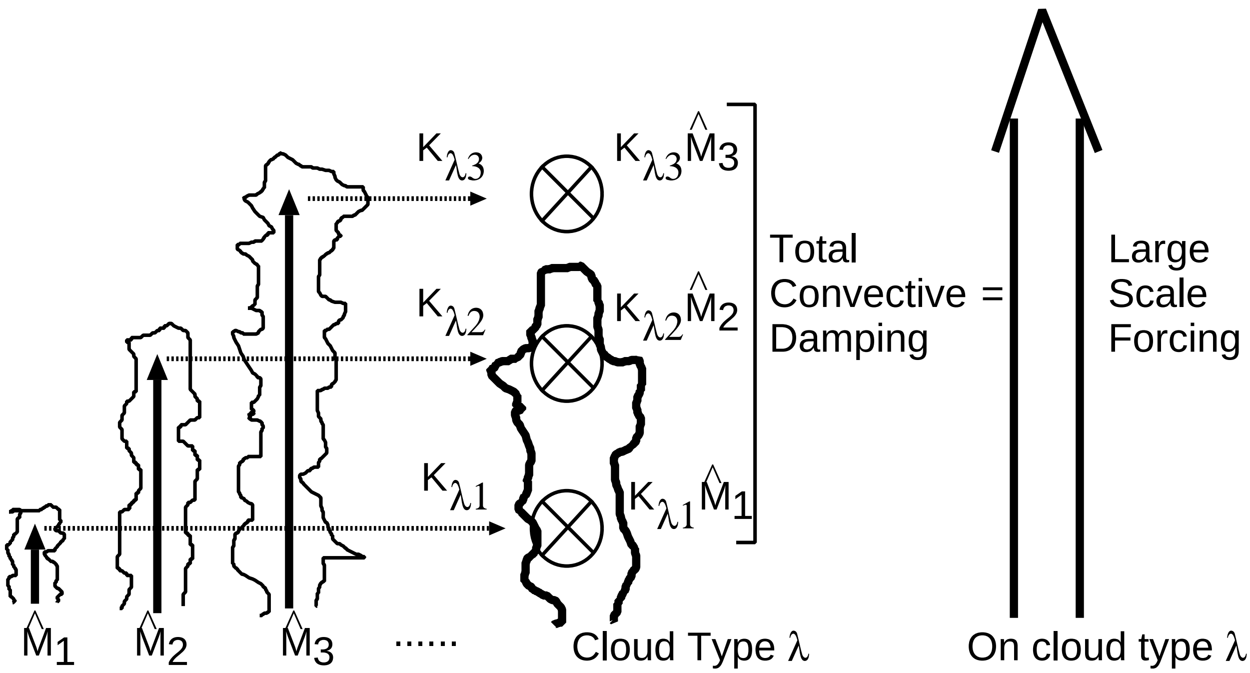

Figure 2. A schematic to illustrate the original concept of convective quasi–equilibrium (CQE) by *Arakawa and Schubert* [1974]. Each convective cloud is characterized by its cloud-base mass flux, $\hat{M}_1, \hat{M}_2, \hat{M}_3, \ldots$ (left) and damps the PEC for cloud type λ (center) with a rate $\mathcal{K}_{\lambda 1}\hat{M}_1, \mathcal{K}_{\lambda 2}\hat{M}_2, \mathcal{K}_{\lambda 3}\hat{M}_3, \ldots$. CQE is the assumption that the sum (*i.e.*, the total damping rate) balances with the large–scale forcing for each cloud type (right).

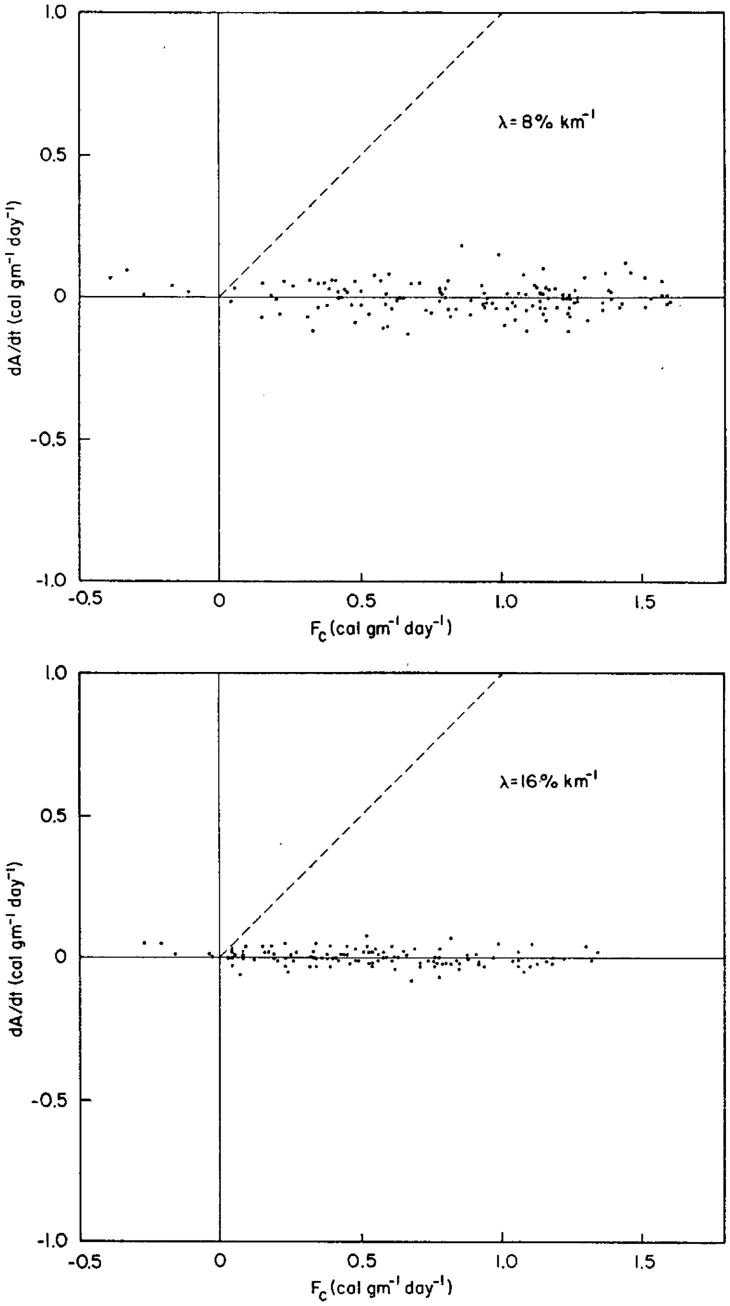

Figure 3. An observational demonstration of convective quasi–equilibrium: the horizontal axis is the large–scale forcing, $F_{L,\lambda}$, and the vertical axis is the rate of change of cloud work function, dA_λ/dt. The top and bottom panels are for entraining plumes with constant fractional entrainment rates of 0.08 km^{-1} and 0.16 km^{-1} respectively. Marshall Islands data provided by *Yanai et al.* [1973] was used. The dashed line corresponds to $dA_\lambda/dt = F_{L,\lambda}$. [Taken from Fig. 13 of *Arakawa and Schubert*, 1974].

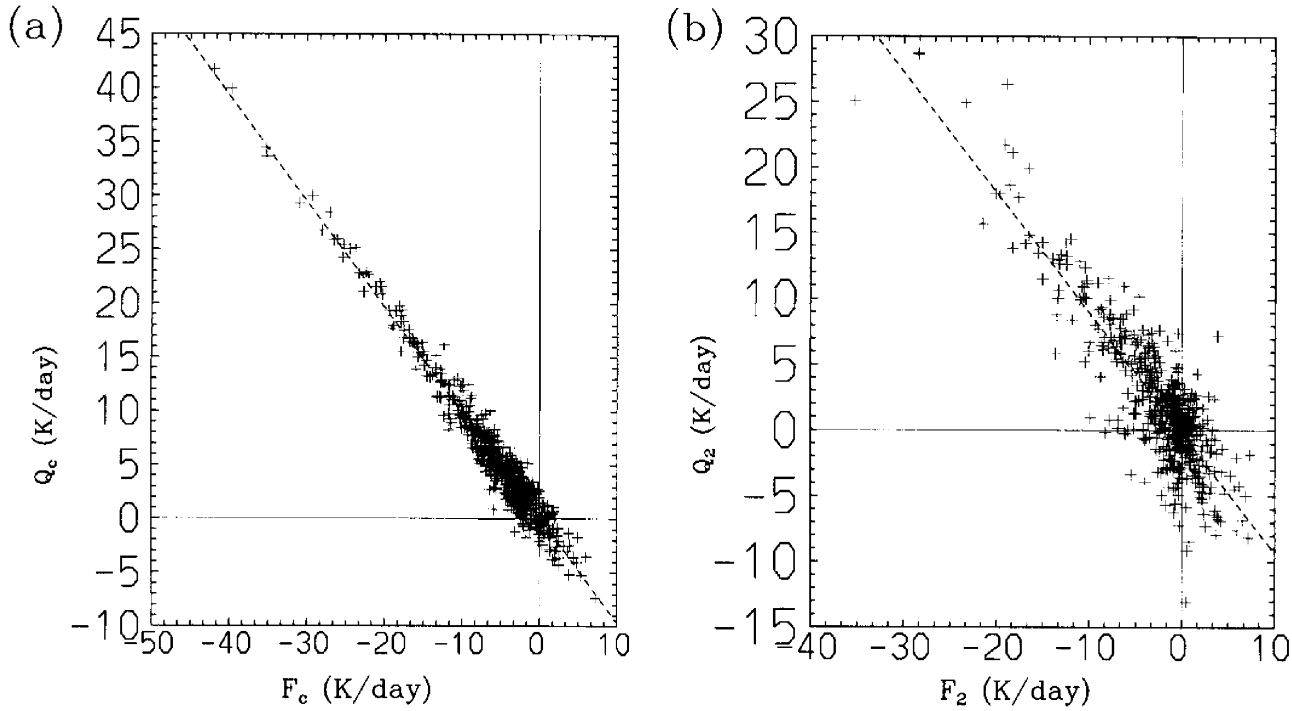

Figure 4. An observational demonstration of the "free ride" principle: for (a) the thermodynamic and (b) the moisture equations. The horizontal axis is the large–scale forcing, and the vertical axis is convective forcing. The values at 500 hPa are shown from the TOGA–COARE IFA (Intensive Flux Array). [Taken from Fig. 1 of *Yano*, 2001].

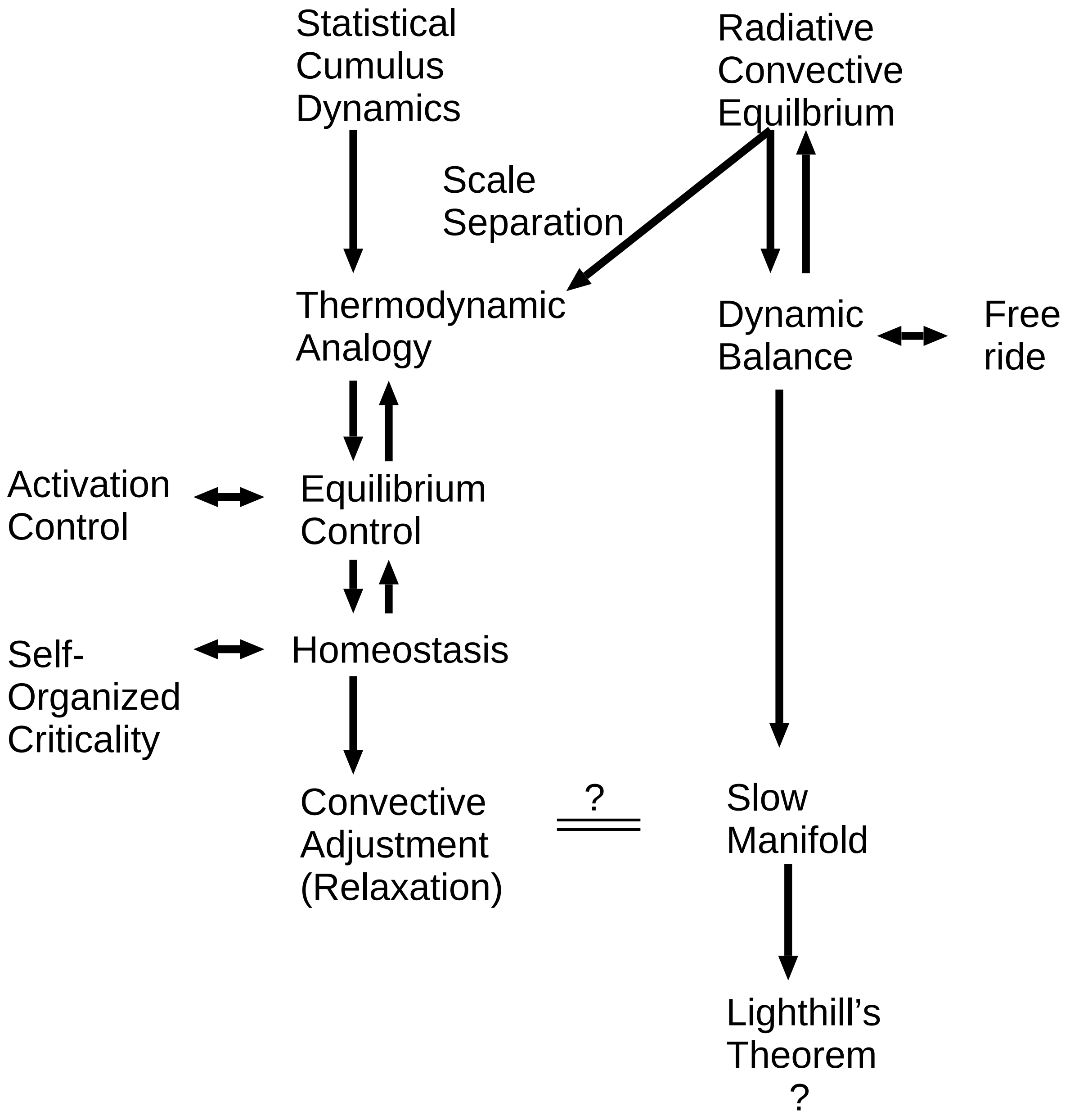

Figure 5. A flowchart for summarizing the links between various concepts discussed in the review. The arrows indicate directions of evolution of the concepts. Where linking arrows are shown in both directions, it suggests that the two concepts are *almost* equivalent. On the other hand, the double arrow suggests two conflicting concepts. The concepts of convective adjustment and the slow manifold are linked by an equal sign with a question mark (=?), because they are closely related but clearly not equivalent.