The role of criticality on the horizontal and vertical scales of extratropical eddies in a dry GCM

JUNYI CHAI, *

Atmospheric and Oceanic Sciences Program, Princeton University, New Jersey, USA

GEOFFREY K. VALLIS

College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

*Corresponding author address: Junyi Chai, Atmospheric and Oceanic Sciences Program, 300 Forrestal Road, Sayre Hall, Princeton, NJ 08544
E-mail: junyic@princeton.edu
This paper discusses the sensitivity of the horizontal and vertical scales of extratropical eddies when criticality is varied in a dry, primitive equation, general circulation model. Criticality is a measure of extratropical isentropic slope and when defined appropriately its value is often close to one for Earth’s climate. The model is forced by a Newtonian relaxation of temperature to a prescribed temperature profile, and criticality is increased by increasing the thermal relaxation rate on the mean flow. When criticality varies near one, it is shown that there exists a weakly nonlinear regime in which the eddy scale increases with criticality without involving an inverse cascade, while at the same time the Rossby radius may in fact decrease. The quasi-geostrophic instability of Charney problem is revisited. It is demonstrated that both the horizontal and vertical scales of the most unstable wave depend on criticality, and simple estimates for the two scales are obtained. We reconcile the opposite trends of the eddy scale and Rossby radius and obtain an estimate for the eddy scale in terms of the Rossby radius and criticality that is broadly consistent with simulations.
1. Introduction

The theory of the scale of the energy-containing eddies in the extratropics for a given time-mean flow can be regarded as falling between two end-members which we call the linear regime and the turbulent regime. In the linear regime we take the eddy scale to be similar to that of the most unstable baroclinic wave. In Earth’s atmosphere, the observed energy-containing eddy scale (spherical wavenumber $\sim 8$) is only slightly larger than that of the most unstable waves, and the eddy turnover time ($\sim 3-5$ days) is similar to their inverse growth rate (Simmons and Hoskins 1976; Valdes and Hoskins 1988). An estimation for the scale of the most unstable wave originates from Eady’s (1949) classical work in baroclinic instability, in which the basic flow is simplified to be independent of latitude. The Eady model gives simple estimate for the most unstable wavelength $\lambda_m$, that it is only dependent on the Rossby deformation radius $L_R$ according to $\lambda_m \approx 3.9 L_R$. For Earth’s atmosphere, the Eady model gives a good estimate for the eddy scale and eddy turnover time. The Charney model of baroclinic instability does not give nearly as simple an estimate as the Eady problem (a problem we return to later on) but nevertheless the Rossby radius is often used as an estimate eddy scale and indeed it has been shown that for a wide range of mean flows, the Rossby radius multiplied by an empirical constant gives a decent estimate for the scale of the most unstable waves and eddy scale (Merlis and Schneider 2009).

In the other end-member, the eddy scale is considered to be determined by geostrophic turbulence. Now, the baroclinic mode (proportional to temperature in the two-level quasi-geostrophic model) is maintained by differential heating of the atmosphere and stirred by the barotropic flow as if it were a passive tracer. Eddy available potential energy cascades to smaller scales until it arrives at the scale of baroclinic instability, where it is converted into barotropic eddy kinetic energy (Salmon 1980). The kinetic energy is then cascaded to larger scales by nonlinear eddy-eddy interactions much as in two-dimensional turbulence. Unless halted by friction or limited by domain size then, near the Rhines scale, energy becomes preferentially channeled into the zonal flow (Rhines 1979; Vallis and Maltrud 1993). Thus,
in principle, the energy-containing eddy scale can be much larger than the Rossby radius. This conceptual picture generally agrees with the life cycles of nonlinear baroclinic waves (Simmons and Hoskins 1978), and the inverse cascade is seen in many simulations of both the quasi-geostrophic and primitive equations (Vallis and Maltrud 1993; Larichev and Held 1995; Held and Larichev 1996; Zurita-Gotor and Vallis 2009; Jansen and Ferrari 2012).

The regimes in which above two strands of theory are relevant can be separated by a non-dimensional number, the criticality $\xi$. Criticality is a measure of isentropic slope and it can be defined in layered and continuously stratified quasi-geostrophic (QG) and primitive equation (PE) models. Criticality has the cleanest expression in the two-layer QG model on the $\beta$ plane (Phillips 1954). Here, the non-dimensional number $\xi = U/(\beta L_R^2)$ determines the linear instability, where $L_R$ is the deformation radius and $U$ is the mean thermal wind. The system is baroclinic unstable only if $\xi > 1$. However, this particular instability criteria is unique to the two layer model and in continuously stratified model there is generally no minimum requirement on criticality for the system to be unstable. Furthermore, for the full nonlinear model, baroclinic eddies can develop even if it is linearly stable (Farrell and Ioannou 1994; Lee and Held 1991). However, criticality can serve more generally as a measure of instability and as a measure of the extent of the inverse cascade. Generally when criticality $\xi \leq 1$, the inverse cascade is not important, and the energy-containing eddy scale, Rossby radius and Rhines scale are all very similar (Schneider and Walker 2006). When $\xi \gg 1$, there can be substantial inverse cascade. In this case, the eddy scale is much larger than Rossby radius, and an adequate estimate for eddy scale is sometimes found to be the Rhines scale (Jansen and Ferrari 2012). Although the Rhines scale may give a good estimate for the eddy scale in the high criticality case, it is not an a priori or mean field estimate because it assumes knowledge of the root-mean square velocity at the eddy scale or, in the variant proposed by Vallis and Maltrud (1993), one needs to know the energy cascade rate. For baroclinic flow we need a second-order closure to calculate these quantities, such as that provided by Held and Larichev (1996). They took the eddy scale $L_e$ and the Rhines scale
$L_\beta$ to be similar and derived a classical scaling to relate them to the mean field estimates Rossby radius $L_R$ and criticality $\xi$, such that $L_\nu \sim L_\beta \sim \xi L_R$.

In between the linear regime and the turbulent one a theory for eddy scale is less clear. Intuitively, if we increase criticality from below 1 to much much larger than 1, we expect the eddy scale to increase from the deformation radius to the Rhines scale, but it is less clear just how the eddies adjust their scales to an increased criticality and when the inverse cascade becomes developed. However, there are very few studies on the relationship between eddy scale and criticality in a PE model mainly because, in a PE model, the criticality is difficult to change by perturbing model parameters. In a PE model, criticality can be intuitively thought of as the ratio of the potential temperature difference between subtropics and Poles $\Delta h\theta$, to that between the surface and the tropopause $\Delta_v\theta$, written as $\xi = \Delta h\theta / \Delta_v\theta$. In Earth’s atmosphere, this quantity is well-known to stay at about unity throughout the seasons, even though the meridional temperature gradient changes significantly (Stone 1978). Earth’s atmosphere thus appears to be in a low-criticality regime in which nonlinear interaction does not transfer much energy upscale beyond the Rossby radius (Boer and Shepherd 1983; Shepherd 1987; Straus and Ditlevsen 1999). It has been argued that such a marginal critical state for Earth’s atmosphere does not appear by chance; it was first called “baroclinic adjustment” by Stone (1978) and later Schneider (2004) and Schneider and Walker (2006) argued that the constraint $\xi \approx 1$ arises using a turbulent closure for potential vorticity and potential temperature fluxes along with an assumption about the nature of the eddy diffusivity. Unlike a QG model in which the vertical stratification is prescribed, in the PE model the macro-turbulence adjusts its extratropical thermal stratification to a state in which eddy-eddy interactions are weak. Schneider and Walker (2006) concluded that the supercritical state $\xi > 1$ is unobtainable in the PE model.

However, a number of studies have shows that the constraint $\xi \approx 1$ for a PE model can, in some circumstances, be violated, and the criticality can be continuously varied above 1. Zurita-Gotor (2008) first achieved this by changing diabatic heating rate in a PE dry
GCM forced with Newtonian cooling. In order to vary the diabatic heating rate by several orders, Zurita-Gotor only changed the forcing time scale for the zonal mean flow and kept the diabatic damping time scale for the eddies to be the same. He showed that criticality increases when the diabatic heating rate increases and that there is nothing special about the state $\xi \approx 1$. In a later study, Zurita-Gotor and Vallis (2009) applied the same diabatic forcing technique in a two-level PE model on a $\beta$ plane. By using a high diabatic heating rate, along with adjusting $f$ and $\beta$, they obtained the classical $-5/3$ energy spectrum in the barotropic part of the flow, suggesting the existence of an inverse cascade. More recently, Jansen and Ferrari (2012) showed that criticality can be continuously varied by changing the thermal expansion coefficient in a Boussinesq model. When the thermal expansion coefficient decreases from atmosphere-like values to ocean-like values, the mean flow transits from marginally critical ($\xi \sim 1$) to highly supercritical states ($\xi \gg 1$). Jansen and Ferrari (2013a) further suggested that the constraint $\xi \approx 1$ can break down because the Schneider’s (2004) formulation of averaging along isentropes in the surface layer is not consistent with the turbulence diffusive closure. Their subsequent scalings for criticality clearly suggested that criticality can vary with changes in the radiative forcing, and to a greater extent, with changes in the size and rotation rate of the planet (Jansen and Ferrari 2013b). The highly supercritical states are most likely to exist in real geophysical flows where the planetary radius is much larger than the Rossby radius. This may account for the possible existence of inverse cascade in Jupiter’s atmosphere (Choi and Showman 2011) and in Earth’s ocean (Scott and Wang 2005).

These studies provide a basis for studying the relationship between eddy scale and criticality in a PE model. Interestingly, Zurita-Gotor (2008) showed that when criticality is increased by increasing the diabatic heating rate, the eddy scale *increases* while the Rossby radius *decreases*. This opposite trend of eddy scale and Rossby radius is surprising and suggests that the inverse cascade may increase its extent with criticality. However, Zurita-Gotor (2008) noticed that the measure of inverse cascade proposed by Schneider and Walker (2006),...
defined as the ratio between eddy available potential energy and baroclinic eddy kinetic energy, does not clearly support there being a significant inverse cascade in his simulations.

In this study, we aim to quantify the role of the criticality and the inverse cascade in setting the eddy scale in dry primitive equation model. (We do not involve ourselves in the discussion of whether and how states of high criticality can be reached in a natural way; rather, we treat criticality as an external parameter that can be varied using the thermodynamic forcing.) The regime we focus on involves criticality varying near unity, such that the flow transitions from a weakly nonlinear regime to a strongly nonlinear regime. This regime is further complicated by the fact that in linear QG instability theory, the most unstable wavelength has a dependency on criticality when criticality varies near 1. This may provide an additional mechanism for the eddy to adjust its length scale to an increased criticality without involving a well-developed inverse cascade. The classical Charney problem well illustrates this mechanism but fully solving that problem is analytically difficult and uninformative. Instead, by simple rescaling, we derive a simple approximate estimate for the most unstable wavelength’s dependency on criticality in the Charney problem, and this work is presented in section 2. The numerical model we use to test the theory is described in section 3 as are some numerical experiments. We use both a full model and a model that is modified to remove eddy-eddy interactions in order to quantify the role of the inverse cascade. Section 4 further analyzes the modeling results. Our overall goal is to understand the conditions under which, when criticality is varied, the inverse cascade is responsible for the increase in eddy scale, or whether we can understand eddy scale from linear or weakly nonlinear theory; a specific goal would be to estimate the eddy scale as a function of Rossby radius and criticality. Section 5 finishes with some concluding remarks.
2. Notation and linear instability theory

a. Definition of criticality

Following Jansen and Ferrari (2012, 2013), we define criticality as

$$\xi = \left| \frac{f \partial_y \bar{\theta}}{\beta \Delta_V \bar{\theta}} \right|, \tag{1}$$

where $\Delta_V \bar{\theta} = \bar{\theta}_t - \bar{\theta}_s$ is the potential temperature difference between the surface and tropopause, which is often called the bulk stability. The overbars denote temporal and zonal means. In pressure coordinates, we estimate the bulk stability as

$$\Delta_V \approx -\frac{\partial_p \bar{\theta}}{\beta} (\bar{p}_s - \bar{p}_t), \tag{2}$$

where the superscript $s$ means that the evaluation is done near the surface; $\bar{p}_t$ and $\bar{p}_s$ denote the pressure at the tropopause and surface respectively.\(^1\)

In spherical geometry, $f$ and $\beta$ cannot be adjusted independently, but the ratio $f/\beta$ is on the order of the planetary radius. Therefore, $f \partial_y \bar{\theta}/\beta$ approximates the potential temperature difference between the subtropics and the pole, which we denote as $\Delta_h$. The state $\xi \sim 1$ can be interpreted as an isentrope starting at the surface in the subtropics and ending at the tropopause near the pole. Intuitively, criticality can be thought as the isentropic slope if we plot isentropes in a coordinate such that the troposphere has the same meridional and vertical length scales.

\(^1\)In Schneider and Walker's (2006) original paper on criticality, they defined $\Delta_V \approx -2\partial_p \bar{\theta} (\bar{p}_t - \bar{p}_s)$ and derived the constraint $\xi \sim 1$ under this definition. Jansen and Ferrari (2013a) suggested that Schneider and Walker's formulation of averaging along isentropes in the surface layer is not consistent with the diffusive closure and dropped the factor 2 noting there is nothing special about the observed $\xi \sim 1$ as criticality can be varied smoothly. In practice, the exact value of criticality depends on the latitude at which $f$ and $\beta$ are calculated, and also depends on the region one chooses to average over. Therefore, we also drop the factor of 2.
b. Most unstable wave’s dependency on criticality

The most unstable wavelength in the simplest Eady model only depends on Rossby radius. In more realistic models on a β plane, the most unstable wavelength usually has some dependency on criticality. For example in the two-layer Phillips model, the most unstable wavelength increases with criticality, though the dependency is weak. The dependency is much stronger in a continuous model, especially when criticality is much smaller than 1, and becomes weaker when criticality increases above 1. In the following, we show this using the Charney model and derive a simple approximate relationship between the most unstable wave’s scale and criticality.

The Charney problem considers the QG instability of a non-Boussinesq continuously stratified fluid on the β–plane with a rigid bottom boundary condition (Charney 1947; Vallis 2006, Chapter 6; Pedlosky 1979, Chapter 7). The QG equation for the perturbation stream function Ψ, linearized about a zonally and meridionally homogeneous mean flow U(z), is

\[ \left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) q + \frac{\partial \Psi}{\partial x} \frac{\partial Q}{\partial y} = 0, \]  

where q and Q are the potential vorticities of the perturbation and the mean flow respectively, and they are given by

\[ q = \nabla^2 \Psi + \frac{f^2}{\rho} \frac{\partial}{\partial z} \left( \rho \frac{\partial \Psi}{\partial z} \right), \quad \frac{\partial Q}{\partial y} = \beta - \frac{f^2}{\rho} \frac{\partial}{\partial z} \left( \rho \frac{\partial U}{\partial z} \right). \]

The rigid bottom boundary condition is

\[ \left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \frac{\partial \Psi}{\partial z} - \frac{\partial \Psi}{\partial x} \frac{\partial U}{\partial z} = 0, \quad \text{at } z = 0, \]

and the other boundary condition is that the perturbation vanishes at infinity:

\[ \Psi(z = \infty) \rightarrow 0. \]

For the Charney problem, further simplifications for the mean flow and stratification are made. The zonal flow is assumed to have a constant vertical shear, such that \( U(z) = (\partial_z U)z, \)
where $\partial_z U$ is constant. The vertical stratification specified by the Brunt-Väisälä frequency $N$ is assumed to be constant. The density profile is characterized by a density scale height as $\rho = \rho_0 \exp(- z / H_\rho)$. Therefore, we can rewrite (4) as

$$q = \nabla^2 \Psi + \frac{f^2}{N^2} \frac{\partial^2 \Psi}{\partial z^2} - \frac{f^2}{N^2 H_\rho} \frac{\partial \psi}{\partial z}, \quad \frac{\partial Q}{\partial y} = \beta + \frac{f^2 \partial_z U}{N^2 H_\rho}$$

where the boxed terms arise from density’s derivative with respect to height. For a Boussinesq fluid that has no density variation, all the boxed terms above and in the following will disappear. As we will see later, the density variation introduces a significant difference between non-Boussinesq and Boussinesq flow.

As usual, we nondimensionalize the equation using

$$x, \ y = L_R (\hat{x}, \ \hat{y}),$$

$$z = H_\rho \hat{z},$$

$$t = L_R / (H_\rho \partial_z U) \hat{t},$$

where hats denote nondimensional quantities and the horizontal scale is the Rossby radius defined as:

$$L_R = \frac{N H_\rho}{f}.$$  

A nondimensional parameter, which we define to be criticality for the Charney problem, naturally arises after the nondimensionalization:

$$\xi = \frac{f^2 \partial_z U}{\beta N^2 H_\rho}.$$  

All information on the mean flow (shear, stratification, density profile, latitude, and etc.) is absorbed into this single parameter $\xi$ after nondimensionalization, as Eq. (3) becomes

$$\left( \frac{\partial}{\partial \hat{t}} + \hat{z} \frac{\partial}{\partial \hat{x}} \right) \left( \nabla^2 + \frac{\partial^2}{\partial \hat{z}^2} - \frac{\partial}{\partial \hat{z}} \right) \Psi + \left( \frac{1}{\xi} + \Pi \right) \frac{\partial}{\partial \hat{x}} \Psi = 0,$$

and the boundary conditions (5) and (6) become

$$\frac{\partial^2 \Psi}{\partial \hat{t} \partial \hat{z}} - \frac{\partial \Psi}{\partial \hat{x}} = 0, \text{ at } \hat{z} = 0,$$  

9
and

\[ \Psi(\hat{z} = \infty) \rightarrow 0. \]  \hfill (13)

With the thermal wind relation, we can rewrite (10) as

\[ \xi \approx \frac{f}{\beta H_\rho} \left| \frac{\partial \rho \theta}{\partial \hat{z}} \right| \]

and immediately see its relevance to the criticality definition in (1). Compared with the usual definitions for Rossby radius and criticality in a PE system, the density scale height replaces the tropopause height in (9) and (10). For Earth’s atmosphere, the tropopause height and density scale height are of the same order in the extratropics. Also with \( -\partial_\rho \theta^s \sim g^{-1} \partial_\rho \theta^s / \bar{\rho} \sim H_\rho \partial_\rho \theta^s / \bar{\rho}_s \), we can rewrite the bulk stability in (2) as

\[ \Delta V \approx -\partial_\rho \theta^s (\bar{\rho}_t - \bar{\rho}_s) \sim H_\rho \partial_\rho \theta^s \bar{\rho}_s - \bar{\rho}_t. \]

As long as the tropopause pressure is much smaller than surface pressure, \( H_\rho \partial_\rho \theta^s \) can serve as a lowest order estimate for bulk stability. Therefore, we use (9) and (10) as the definitions for Rossby radius and criticality respectively in the Charney problem.

Assuming plane-wave solutions of the form

\[ \Psi = \psi(\hat{z}) \exp[i k (\hat{x} - c \hat{t})] \]  \hfill (14)

to the nondimensionalized equation (11) and its boundary conditions (12-13), the equation describing the vertical structure of the wave is

\[ \frac{d^2 \psi}{d\hat{z}^2} - \left[ \frac{1}{\hat{z}} \right] \partial \psi \hat{z} - \left( k^2 - \frac{1}{\xi} + \frac{1}{\hat{z} - c} \right) \psi = 0, \]  \hfill (15)

with boundary conditions at rigid bottom and at infinity

\[ c \frac{d\psi}{d\hat{z}}(0) + \psi(0) = 0, \quad \psi(\infty) \rightarrow 0. \]  \hfill (16)

For a given pair of criticality \( \xi \) and wavenumber \( k \), the above equation (15) and its boundary conditions (16) form an eigenvalue problem. We can solve it to get an eigenfunction
ψ and eigenvalue c. If the complex part of kc is positive, the wave is unstable. Its growth rate is given by kc, where ci denotes the complex part of c. For a given ξ, there exists a unique km for which km ci reaches its maximum value (kc)i max, where the subscript "max" denotes the maximum value of the bracket for k over (0, ∞) and a definite ξ. The wave solution (14) with the wavenumber km is referred as the most unstable wave. It is often thought that the most unstable wave will dominate over other waves at finite amplitude, so the corresponding most unstable wavelength gives the energy containing eddy scale. Clearly, the most unstable wavenumber km is a function of criticality ξ.

To see how ξ controls most unstable wave, we rescale the variables as

\[ \hat{z} = z^* / \left( \frac{1}{\xi} + 1 \right), \quad c = c^* / \left( \frac{1}{\xi} + 1 \right), \quad k = k^* \left( \frac{1}{\xi} + 1 \right). \]  

(17)

Then the Eqs. (15) and (16) become

\[ \frac{d^2 \psi}{dz^*} - \frac{\xi}{\xi + 1} \frac{d \psi}{dz^*} - \left( k^{*2} - \frac{1}{z^* - c^*} \right) \psi = 0, \]  

\[ c^* \frac{d \psi}{dz^*}(0) + \psi(0) = 0, \quad \psi(\infty) = 0. \]  

(18) (19)

Under this rescaling the growth rate is unchanged as kc = k*c*, so the most unstable wave in the rescaled Eqs. (18) and (19) are also related to that in the Eqs. (15) and (16) through the rescaling relation (17). If the fluid is Boussinesq, the boxed terms are dropped. Therefore Eqs. (18) and (19) have no reference to the mean flow, and the most unstable wavenumber km for the rescaled equations is a constant for any ξ. Therefore for the Boussinesq fluid, rescaling relation (17) gives the familiar results that the most unstable wave has horizontal scale proportional to \( L_m \sim f \partial_z U / (\beta N) \) and vertical scale proportional to the Charney height \( h = f^2 \partial_z U / (\beta N^2) \).

For a compressible fluid, we cannot eliminate the dependency on the mean flow by this simple rescaling. But notice that the dependency on ξ in (18) is weak as ξ/(1 + ξ) only varies from 0 to 1 when ξ varies from 0 to ∞. Without the rescaling, the original Eq. (15) has
a much higher dependency on $\xi$ as $1/\xi$ can vary from 0 to $\infty$. So we may expect that for
the rescaled equation (18) and its boundary condition (19), the most unstable wavenumber
$k_m^*$ is only weakly dependent on $\xi$. If so, we can drop the dependency of $k_m^*$ on $\xi$, such that
$k_m^*(\xi) \approx k_m^*$, where $k_m^*$ represents a characteristic value of $k_m^*$. Use rescaling (17) we get the
dependency of $k_m$ on criticality as $k_m(\xi) \approx (1 + 1/\xi)\overline{k}_m^*$. Finally, we can get the dimensional
most unstable wavenumber for the original equation (3) by adding the length scale described
by nondimensionalization relationship (8). Converting the wavenumber into wavelength, we
get in dimensional variables that the most unstable wavelength is

$$L_m \approx C \frac{\xi}{\xi + 1} L_R,$$

(20)

where $C = 2\pi/k_m^*$ is the approximate constant. Its vertical scale is

$$h_m \approx D \frac{\xi}{\xi + 1} H_\rho,$$

(21)

where $D$ is another approximate constant on the order of 1. To test whether the above
scalings offer a good estimate, we resort to numerical solutions. Before doing this, we note
that the rescaling (17) was used by Lindzen and Farrell (1980a) to obtain an approximate
simple formula for the maximum growth rate. Also, similar forms of the most unstable
wavelength’s estimate as (20) exist in the literature. Branscome (1983) derived the same
estimate (20) through perturbation methods in the short-wave limit ($k \gg 1$) and near
the neutral curve separating Charney and Green modes (the right side boundary of the
growth rate contour shown in Fig. 1), and he determined $C$ to be 12. Wang et al.
(1985) obtained a more complicated estimate by using an integral representation of the
solution and a Frobenious series expansion, and his estimate was later simplified into $L_m \sim
2\pi\xi/(0.48\xi + 1.48)L_R$ (Stone and Yao 1990; Barry, Craig and Thuburn 2002). Above all, the
asymptotic limits for our scaling estimates are well-known (Held 1978; Vallis 2006). When
$\xi \ll 1$ both horizontal and vertical scales of the most unstable wave are linearly proportional
to $\xi$:

$$L_m \sim O(\xi L_R), \ h_m \sim O(\xi H_\rho), \text{ when } \xi \ll 1.$$  

In this limit, the vertical scale of the wave is much smaller than scale height, so we refer to such waves as shallow waves. When $\xi \gg 1$ the sensitivity of $L_m$ and $h_m$ on $\xi$ saturates. The horizontal scale is only given by deformation radius, while the vertical scale is limited by the density scale height, written as

$$L_m \sim O(L_R), \ h_m \sim O(H_\rho), \text{ when } \xi \gg 1.$$  

We refer to these waves as deep waves as their vertical scale is comparable to the scale height. Our scaling estimates (20) and (21) connect the two limits with a smooth transition.

c. Numerical validation

The non-dimensionalized Charney problem described by Eqs. (15) and (16) are solved numerically for various values of criticality $\xi$ and wavenumber $k$. As we are interested in the regime $\xi \sim O(1)$, we vary $\xi$ between 0.1 to 4. Fig. 1 is a contour of the growth rate against inverse wavenumber $1/k$ and $\xi$, together with two curves showing the numerically obtained inverse of the most unstable wavenumber $1/k_m$ and our scaling estimate $1/k_m \approx 1.54\xi/(1+\xi)$, where the constant 1.54 is determined by a least square fit to give the best agreement between the our scaling estimate and numerical result. Only the Charney mode regime is plotted. The other modes have smaller maximum growth rates compared with the Charney mode (Pedlosky 1979). Clearly our scaling estimate gives a good approximation to the numerical results. The difference between the two is within 0.08. To convert wavenumber into wavelength, a factor of $2\pi$ needs to be multiplied. So in dimensional variables, our scaling estimate for most unstable wavelength is

$$L_m \approx 9.7 \frac{\xi}{\xi + 1} L_R.$$  

Fig. 2a shows the vertical structure of the most unstable wave for $\xi = 1$. Three curves are shown, corresponding to amplitude of perturbation $|\psi|$, heat flux $\overline{v^' \partial_z \psi}$, and PV flux.
All of them are normalized by their maximum values. Generally for the most unstable wave, the heat flux decreases rapidly with height. To the lowest order, the vertical profile of heat flux can be approximated by an exponential decay:

\[ v' \partial_z \psi \sim \exp(-z/h_\theta). \]  

(22)

Therefore we use (22) as the definition for the vertical scale of the wave, such that the vertical scale of the wave \( h_\theta \) is the \( e \)-folding height of the heat flux. For the real atmosphere or a primitive equation model, we replace the heat flux in (22) by the potential temperature (PT) flux \( \bar{\psi} \bar{\theta}' \). We can find \( h_\theta \) by fitting the PT flux profile to an exponential function.

To test our estimate for the vertical scale of the most unstable wave in (21), we calculate \( h_\theta \) from the numerical solutions of the most unstable wave for different \( \xi \) and plot their relationship in Fig. 2b. The best estimate is \( h_\theta \sim D\xi/(1 + \xi) \), with the parameter \( D \) determined to be 1.09.\(^2\) Again, our scaling estimate provides a reasonable approximation to the numerical results. Their difference is within 0.11. Because of the analytical difficulty in solving Charney problem, whose solution involves hypergeometric functions, deriving a more accurate estimate for the most unstable wavenumber will presumably make the formula more complicated. Therefore, our scaling estimate is useful in the sense that it provides a simple estimate for both practical use and theoretical understanding.

\(^2\)Held (1978) gave a formally similar estimate for the vertical extent of eddy PV fluxes. His definition of vertical extent is based on the \( e \)-folding scale of \( \bar{\rho} \bar{v}' \bar{q}' \). The density profile multiplied to PV flux limits this vertical extent with an upper bound of density scale height, whereas in our definition the vertical scale \( h_\theta \) can be larger than density scale height. That \( h_\theta \) saturates when \( \xi \) becomes large is not because of being limited by density’s decay with height.
3. Idealized GCM and experiments

a. Model description

We investigate how eddy scale depends on criticality in an idealized GCM. The GCM is a primitive-equation model of a dry ideal gas atmosphere on a sphere of Earth’s radius and rotation rate. The primitive equations are integrated with the spectrum transform method in the horizontal with resolution T42, and centered difference scheme in the vertical, with evenly spaced vertical $\sigma$ levels. There is no bottom topography at the lower boundary. It is forced by idealized Held and Suarez (1994) physics where radiation is represented by a Newtonian relaxation of temperature to a prescribed “radiative equilibrium” profile and dissipation is represented by a Rayleigh damping of the velocities in the “boundary layer” ($\sigma \geq 0.7$). The model setup is very similar to that described in Zurita-Gotor (2008), except that we use a thermal relaxation rate of $1/40 \text{ day}^{-1}$ in the free troposphere, which is the original value used by Held and Suarez (1994), and we use 30 vertical levels.

Criticality is varied by varying thermal forcing. To vary the thermal forcing by several orders without damping out the eddies, Zurita-Gotor (2008) suggests separating the forcing into zonal mean flow and eddy terms, and only vary the forcing time scale on the mean flow. The thermodynamic equation is then

$$\frac{\partial T'}{\partial t} = \ldots - k_T T' - \gamma k_T (\bar{T} - T_{eq}).$$

Overbars denote zonal mean fields and primes denote the deviations from zonal means. The forcing on the zonal mean flow is increased by $\gamma$ times. Presumably, we can have a very high forcing rate on the mean flow so that the time-mean temperature field is close to the prescribed profile, while the eddies are still nearly adiabatic.
b. Reduced GCM

In order to isolate the effects of the inverse energy cascade on the eddy scale, we use the GCM without nonlinear eddy-eddy interactions introduced by O’Gorman and Schneider (2007), which we will refer as the reduced GCM. As a comparison, the original GCM described above is referred as the full GCM. The reduced GCM is constructed from the full GCM by eliminating the nonlinear eddy-eddy interactions but retaining the eddy/mean-flow interactions. Therefore in the reduced GCM, there is no inverse cascade, which is a scale-by-scale energy transfer, but there is still direct energy transfer between eddy and mean flow (Tobias and Marston 2013). We run the reduced GCM in parallel with the full GCM. The reduced GCM has shown to reproduce many atmospheric flow features, including the eddy length scale, the -3 slope in energy spectrum and the jet formation. This is interpreted as the Earth-like regime being weakly nonlinear (O’Gorman and Schneider 2007).

c. Series of simulations

To vary criticality, we vary the rate of thermal forcing on the mean flow by changing \( \gamma \). The baroclinicity is varied by varying \( \delta_y \). For each setting of \( \delta_y = 30, 60, 90, \text{ and } 120 \) K, we consider 9 different values of \( \gamma = (0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 40) \). For \( \gamma = 1 \) run, the forcing rates on zonal mean flow and eddies are the same, and we refer it as the control run. We run each combination of \( \delta_y \) and \( \gamma \) in both the full and reduced GCMs. Each simulation is integrated for 1500 days. The first 1000 days are discarded as spin-up and the last 500 days are used to calculate the flow statistics.

4. Results

A comparison of the basic climatology of the control run with \( \gamma = 1 \) and \( \delta_y = 60 \) K from the full and reduced GCMs is shown in Fig. 3. The full GCM simulation resembles Earth’s
climatology. It has similar magnitude of zonal wind and eddy kinetic energy (EKE = (\overline{u'^2} + \overline{v'^2})/2) as Earth. We determine the tropopause height using Schneider’s (2004) definition, that the tropopause is the boundary layer below which the bulk of the entropy the atmosphere receives by the heating at the surface is redistributed. Schneider’s (2004) definition is based on isentropic diagnostics, and we use Czaja and Marshall’s (2006) method to convert fields from pressure coordinates to isentropic coordinates. This tropopause is lower than the WMO tropopause. We use Schneider’s tropopause for evaluations of criticality and Rossby radius. Repeating the calculations using the WMO tropopause does not change the results much. The reduced GCM can reproduce the upper-level westerly jets, which are driven by the eddy/mean-flow interactions. But the jets are narrower in the reduced GCM, and there is an extra jet in each hemisphere. The EKE in the reduced GCM is about 2 times larger than that in the full GCM. These results are in accordance with O’Gorman and Schneider (2007).

In the following, all simulations are from full GCM if not specified. We will only use reduced GCM in subsection (a).

In our study, the tropopause height varies little because the mean extratropical surface temperature changes little by our model construction (Eq. [17] in Zurita-Gotor 2008; also see Schneider 2007). Therefore the variation of bulk stability \( \Delta_v \) is mostly due to variation of \( |\partial_p \theta| \), which is a measure of vertical stability. \( |\partial_p \theta| \) is set by two competing processes. The first process is the Newtonian relaxation of the temperature field to the prescribed radiative equilibrium profile. As the prescribed profile in extratropics has very small thermal stratification, this process tends to reduce the vertical stability. The second process is the baroclinic eddy activity, which modifies the thermal stratification (Schneider and Walker 2006, 2008), and in our cases it increases the stratification above that of the prescribed profile.

Fig. 4 shows that for the runs with \( \gamma = 1 \) and varying \( \delta_y \), the bulk stability \( \Delta_v = |\partial_p \theta^s (p_t - p_s)\) increases proportionally with the scaled meridional temperature gradient \( \Delta_h = |(f/\beta)\partial_y \theta^s| \). This is because when \( \delta_y \) increases and other parameters are kept the same,
the meridional temperature gradient and baroclinicity increase, and as a result the eddy activity becomes stronger. The stronger eddies transport more heat upward, which increases the vertical stability. The temperature gradients $\partial_y \tilde{\theta}^s$ and $\overline{\partial_y \theta^s}$ are evaluated at 850 mb and averaged over the baroclinic zone, defined as the region where the PT flux at 850 mb $\overline{u' \theta^s} \cos(\phi)$ is within 30% of its maximum. We use superscript $s$ in this section to denote that the evaluation is done at 850 mb. The Coriolis parameters $f$ and $\beta$ are evaluated at the reference latitude that is defined as the centroid of the PT flux at 850 mb in each hemisphere:

$$\phi_{\text{ref}} = \frac{\int_0^{\pi/2} \phi \overline{u' \theta^s} \cos(\phi) d\phi}{\int_0^{\pi/2} \overline{u' \theta^s} \cos(\phi) d\phi}.$$  

Under these conventions, the criticality $\xi = \Delta_h / \Delta_v$ for the control simulations ($\gamma = 1$) is close to one. These results are in accordance with Schneider and Walker’s (2006) that criticality is not sensitive to changes in radiative equilibrium baroclinicity.

Following Zurita-Gotor (2008), we vary the criticality by varying $\gamma$, which controls the thermal forcing rate on the mean flow. When $\gamma$ increases, the vertical stratification transitions from the regime that it is set by eddy activity to the regime that is set by the thermal forcing. For the series of runs with fixed $\delta_y$ but increasing $\gamma$, $\Delta_v$ initially increases nearly proportional to $\Delta_h$ when $\gamma \leq 0.5$ (Fig. 4). This is the regime in which eddy activity dominates over thermal relaxation in setting the thermal stratification. In this regime, the criticality stays close to 1. When $\gamma$ further increases above 1, $\Delta_v$ starts to decrease while $\Delta_h$ saturates. This is because thermal relaxation becomes so strong that it forces the $\Delta_v$ and $\Delta_h$ to approach the values of prescribed profile. For $\delta_y = 90$ K and 120 K, $\Delta_h$ decreases slightly when $\gamma$ is very large because the PT flux moves slightly towards the equator, so the reference latitude and $f/\beta$ decrease slightly.

Our major motivation is to investigate the eddy scale’s relationship to criticality. As criticality is varied by varying $\gamma$, we show the eddy scale $L_e$ vs. $\gamma$ for runs with $\delta_y = 60$ K in Fig. 5a, along with the Rossby radius $L_R$. The eddy scale and Rossby radius are normalized by their control values ($\gamma = 1$). The eddy scale is calculated from the “inverse centroid” of the barotropic EKE spectrum (see Appendix A). The Rossby radius is calculated as
\[ L_R = \frac{N_p (\bar{\rho}_s - \bar{\rho}_t)}{f}, \]  

(23)

where \( N_p^2 = -\left(\bar{\rho}\theta^s\right)^{-1} \frac{\partial}{\partial \rho} \theta^s \) is a vertical stability measure, evaluated at 850 m; \( \bar{\rho}_s \) and \( \bar{\rho}_t \) are the surface pressure and tropopause pressure respectively. \( N_p, \bar{\rho}_s \) and \( \bar{\rho}_t \) are averaged over the baroclinic zone. This Rossby radius multiplied by an empirical constant is shown to be a good estimate for eddy scale for a wide range of climates whose criticality is close to 1 (Merlis and Schneider 2009). However, when \( \gamma \) varies the Rossby radius can no longer offer a good estimate for eddy scale. As shown in Fig. 5a, when \( \gamma \) increases the eddy scale increases monotonically\(^3\); while the Rossby radius initially increases with a similar trend when \( \gamma \leq 0.5 \), and then decreases steady when \( \gamma \) increases above 1. The decrease in Rossby radius is mainly dominated by the decrease in vertical stability \( N_p \) as shown in Fig. 5b. Again, \( N_p \) decreases because the thermal forcing dominates over eddy activity in setting vertical stratification when \( \gamma \) is large, so that \( N_p \) is relaxed towards the prescribed equilibrium profile. The tropopause height \( H_p = \bar{\rho}_s - \bar{\rho}_t \) increases slightly when \( \gamma \) is small and then saturates for large \( \gamma \). If WMO tropopause definition is used, the decreasing trend for Rossby radius is even stronger as the WMO tropopause height slightly decreases for large \( \gamma \). If the density scale height is used to evaluate Rossby radius as that in the Charney problem, the results here and in the following would not change much either. This is because changes in vertical stability dominate changes in Rossby radius in our simulations. For all series of simulations with different \( \delta_y \) as shown in Fig. 6, the criticality and eddy scale increases steady with \( \gamma \), while Rossby radius decreases when \( \gamma \) becomes large enough. For the range of \( \delta_y \) and \( \gamma \) in this study, criticality is approximately a function of \( \gamma \).

Why can the eddy scale and Rossby radius have opposite trends when criticality varies?

\(^3\)In Zurita-Gotor’s (2008), the eddy scale saturates at largest \( \gamma \), while in our simulations eddy scale does not saturate. This is mainly because we calculated eddy scale using the inverse centroid rather than the centroid of the EKE spectrum, which gives more weight to larger scales. Also, the largest thermal forcing rate we used is still 5 times smaller than that used by Zurita-Gotor. This difference does not change our results much as in both cases, the ratio of eddy scale and Rossby radius does not saturate.
One possible explanation comes from the Rhines scale argument. When thermal forcing is accelerated, eddies become more energetic. So it is obvious that the Rhines scale would increase, which gives the correct trend for the eddy scale. Quantitatively, the Rhines scale gives a good estimate for the eddy scale for all the runs as shown in Fig. 7b. On the other hand, Rossby radius can predict the eddy scale only when criticality is nearly constant, as can been seen if we connect the points in Fig. 7a that have the same $\gamma$ and different $\delta_y$. However, the Rossby radius loses its predictive power when criticality varies. The Rhines scale is calculated as

$$L_\beta = \sqrt{\frac{1/2u_{bt}^2}{\beta}}$$

in accordance with Schneider and Walker (2006). The barotropic velocity $u_{bt}$ is calculated by averaging the horizontal velocity (including both zonal and meridional velocities) vertically from the lowest layer to 150 mb. The barotropic EKE $1/2u_{bt}^2$ is averaged over the baroclinic zone.

It may seem that the increase in eddy scale can be interpreted as an increase in the extent of the inverse cascade when criticality increases. However, one needs to be cautious to make this conclusion, as we show in section 2 that the most unstable wavelength increases with criticality, which allows the eddy scale to adjust without involving an inverse cascade. In order to quantify the role of the inverse cascade in determining eddy scale, we compare the results from the full GCM with that from the reduced GCM in the following subsection. The reduced GCM does not have eddy-eddy interaction, so it does not support an inverse cascade. In addition, we calculate the spectral EKE budget for the full GCM runs. Therefore we can clearly see how the inverse cascade contributes to the eddy scale.

a. Role of inverse cascade

The reduced GCM can reproduce many first and second order flow statistics from the full GCM. Eddy statistics are examined using the barotropic EKE spectrum. The spectrum
is defined using spherical wavenumbers (see Appendix A). Fig. 8 compares the barotropic EKE spectra for runs with $\delta_y = 60$ K and varying thermal forcing $\gamma = 0.1, 1$ and 10 from both the full and reduced GCMs. The shape of the energy spectrum in the reduced GCM shares many similarities with that of full GCM. In wavenumber regime $10 \sim 40$, all the full GCM runs exhibits the $n^{-3}$ power-law range. On the other hand, in the reduced GCM, for $\gamma = 0.1$ and 1, the spectrum exhibits the -3 slope. For $\gamma = 10$ the slope is shallower than -3 for wavenumber $20 \sim 40$. This shallower spectrum in the reduced GCM is explained by O’Gorman and Schneider (2007) as the lack of eddy-eddy interactions preventing the build-up of energy at small scales. On the larger scales where wavenumber is smaller than 10, the energy spectra in the reduced GCM are generally lower than that in the full GCM, which suggests that the nonlinear eddy-eddy interaction does transfer some energy upscale.

It is striking that the eddy scales are comparable in the full and reduced GCMs. For $\gamma = 1$, the most energetic wavenumbers are both 8 in full and reduced GCMs, so the eddy scales are very similar in both GCMs. More interestingly, in the reduced GCM eddy scale increases with $\gamma$ similarly to that of the full GCM. Fig. 9a compares the eddy scale for all runs from the full and reduced GCMs. The eddy scales in the full and reduced GCMs roughly changes proportionally to each other with a slope close to 1 for $\gamma \leq 10$. This suggests that in the full GCM the eddy scale can increase with criticality without involving an inverse cascade before it enters the strongly nonlinear regime. In the high criticality end where $\gamma \geq 10$, the eddy scale in the reduced GCM saturates or even decreases while eddy scale in the full GCM continues to increase. This implies that eddy-eddy interactions become important in inversely cascading EKE to larger scales.

Another confirmation is from spectral EKE budget analysis (see Appendix B for the methods). We calculate the average scale at which available potential energy is converted into eddy kinetic energy, which we refer as the EKE generation scale. Fig. 9b compares the EKE generation scale with the eddy scale, both from the full GCM. The energy generation scale generally increases proportional to the eddy scale until $\gamma$ increases above 10, beyond
which the energy generation scale saturates. Therefore, we conclude that when criticality
varies around 1, there exists a weakly nonlinear regime in which the eddy scale can increase
with criticality without involving the inverse cascade, but the EKE generation moves to larger
scales as well. When criticality increases above the weakly nonlinear regime, the nonlinear
eddy-eddy interactions become important, and the inverse cascade plays a role in continuing
to increase the eddy scale. In the weakly nonlinear regime, we need an explanation for why
the eddy scale and Rossby radius can have different trends.

b. Linear instability perspective

Discussions in section 2 suggest that the most unstable wavelength depends on criticality.
When criticality varies, the Rossby radius is not enough to estimate the most unstable
wavelength. This may solve the apparent conundrum that the eddy scale and Rossby radius
can have opposite trends when criticality is varied. In the following, we only use results from
the full GCM.

The Charney problem predicts that when the most unstable wave transitions from being
a shallow wave to a deep wave, it also increases its wavelength. However, if the vertical scale
of the wave is already so large that it is comparable to density scale height, the most unsta-
bile wavelength is no longer sensitive to criticality. Fig. 10 shows the eddy flux of potential
temperature $\bar{\nu} \bar{\theta} \cos(\phi)$ for $\gamma=0.1, 1, 10$ from simulations with $\delta_y=60$ K. The potential tem-
perature flux reaches its maximum near the surface. For very small $\gamma=0.1$, the maximum
occurs above the surface near 850 mb. This is probably because of the Rayleigh damping in
the “surface layer” that reduces eddy amplitude. For larger $\gamma$, the maximum moves towards
the surface. Clearly, the vertical extent of the eddy PT flux gradually increases when $\gamma$
increases. The change in tropopause height also captures the increase of eddies’ vertical
extent to some degree.

To obtain a quantitative estimate for the change in the wave’s vertical scale, we calculate
the vertical scale of the wave $h_\theta$ defined in Eq. (22), where $h_\theta$ is the e-folding scale of the eddy
PT flux. Fig. 11a shows the vertical eddy PT flux profile, which is $v' \theta' \cos(\phi)$ horizontally averaged over the baroclinic zone. Clearly when $\gamma$ increases, PT flux decreases with height slower, which means that the vertical scale of the eddies is larger. Because for very small $\gamma$ the PT flux's maximum is above the surface, and also the PT flux has a second maxima in the upper atmosphere near the tropopause, we only use the PT flux profile between 580 mb and 820 mb to extract the vertical scale of the eddies. This pressure range is marked by dashed lines in Fig. 11a. In practice, the ratio of the eddies’ vertical scale $h_\theta$ to the density scale height $H_\rho$ is naturally obtained by a linear regression between $\ln v' \theta' \cos(\phi)$ and $\ln p$, if we assume that $p \approx p_0 \exp(-z/H_\rho)$. As shown in Fig. 11b, the vertical scale of the eddies $h_\theta$ generally increases monotonically with $\gamma$. For $\gamma \leq 10$, the vertical scale of the eddies generally increases from below the scale height towards the scale height. This suggests that the eddies can also increase their horizontal scales linearly as the most unstable wave. When $\gamma$ increases above 10, the vertical scale of the eddies either saturates (for $\delta_y$=90 and 120 K runs) or increases above the scale height (for $\delta_y$=30 and 60 K runs), which suggests that the eddies enters the deep wave regime and the most unstable wavelength loses its sensitivity on criticality. In the deep wave regime, an inverse cascade is needed to continue increasing the eddy scale.

In the Charney model, the wave’s vertical and horizontal scales increase nearly proportionally to each other. However, for all series of experiments here, the vertical scale of the wave generally increases by about a factor of 3 while eddy scale increases by about a factor of 1.7 when $\gamma$ increases from 0.1 to 40. This discrepancy may result from the vertical inhomogeneity. In the model, the tropopause provides a dynamically relevant boundary layer, which may be viewed as a rigid lid as a first approximation. If we add a rigid lid in the Charney problem, in the limit $\xi \gg 1$, the most unstable wave will approach that in the Eady problem. In this limit, the vertical scale of the most unstable wave in our definition approaches infinity, as the Eady solution is vertically symmetric. Therefore, the vertical inhomogeneity may allow the vertical scale to increase faster than its horizontal scale when
criticality increases.

The estimate for the most unstable wavelength, including its dependency on criticality, is given in Eq. (20) as \( L_m \sim CL_R \xi / (1 + \xi) \). Although \( L_R \) decreases when \( \gamma \) increases, the increase in \( \xi \) can still give an increase in \( L_m \). Therefore, the most unstable wavelength can increase despite the Rossby radius decreasing, giving the correct trend for eddy scale. As the exact value of criticality depends on the conventions we used, and there is a difference between the mean flows in the Charney model and our simulations, we expect to add a degree of freedom to criticality before we use Eq. (20). In the sense that eddy scale scales with the most unstable wavelength, our eddy scale estimate is

\[
L_e \sim L_m \sim CL_R \frac{A\xi}{A\xi + 1},
\]

where \( A \) and \( C \) are two empirical constants. The criticality in Eq. (20) is replaced by an “effective” criticality \( A\xi \). Fig. 12b shows this estimate vs. the eddy scale from the simulations. The empirical constants are determined by a least square fit as \( C=8.7 \) and \( A=0.36 \). We see that our estimate is in good agreement with the simulations. That the “effective” criticality is smaller than criticality may partly result from the vorticity gradient of the jet \( \bar{u}_{yy} \) that adds to \( \beta \) (Zurita-Gotor 2007). Compared with Held and Larichev’s (1996) scaling for eddy scale \( C\xi L_R \), which is also the estimate for the most unstable wavelength in the shallow wave limit, our estimate has one more empirical constant \( A \). However, this is necessary because the dependency of the eddy scale on criticality is not simply linear. This is seen in Fig. 12a, where we show the best fit using \( C\xi L_R \), with \( C \) determined to be 1.8. This estimate increases too fast with criticality when criticality is large.

5. Conclusion and discussion

In this paper we have studied the relationship between criticality and eddy scale using a dry, three-dimensional, primitive equation model. By varying the Newtonian thermal forcing on the mean flow the criticality can be varied smoothly. The main question we have tried
to answer is: what is the role of the inverse cascade on eddy scale when criticality varies?

We showed that when criticality varies near a value of one, there exists a weakly nonlinear regime in which the eddy scale increases with criticality without involving an inverse cascade.

When criticality increases beyond this weakly nonlinear regime, the inverse cascade becomes important in continuing to increase the eddy scale. The Rossby radius is no longer a useful estimate for eddy scale as criticality increases and the Rossby radius and eddy scale can, in fact, have opposite trends.

To understand these results, we began with linear quasi-geostrophic instability theory, as the classical Charney model shows a dependency of the most unstable wavelength on criticality. By using a simple rescaling, we showed that both the horizontal and vertical scales of the most unstable wave increase with criticality and we derived approximate estimates for these scales. In the weakly nonlinear regime, the increase of eddy scale with criticality is found to be associated with the transition of the most unstable wave from a shallow wave into deep wave and the corresponding increase of the most unstable wavelength. The formula for the most unstable wavelength derived in Charney problem is modified to give a semi-empirical estimate for the eddy scale that is consistent with our simulations.

The way in which an eddy can adjust its horizontal and vertical scales when criticality varies has some precedents in the quasi-geostrophic literature although it has not been systematically studied. Thus, Held and O'Brien (1992) showed that in a three-layer QG model the vertical scale of eddies increases with increasing criticality, and that linear instability theory can make an adequate prediction for this vertical scale. Whitaker and Barcilon (1995) showed that in a two-layer QG model the most energetic wave increases its horizontal scale with increasing criticality without involving nonlinear upscale energy transfer. These studies were limited by their low vertical resolutions and QG studies with higher vertical resolutions would be useful.

Although the most unstable wave in the Charney problem shows quantitative similarities with the eddy scale in our simulations, there are still substantial difficulties in bridging linear
instability theory with modeling results. The vertical and horizontal inhomogeneities of the
time-mean flows in modeling prohibit extracting a simple relationship between the most
unstable wave and criticality. Therefore, we need to resort to even simpler models such as
the Charney problem, but still hope to capture the essential physics. A drawback of the
Charney problem is that it does not take into account the tropopause. One step forward
would be to consider the Green problem, which adds a rigid lid to the Charney problem to
represent the tropopause (Green 1960). In the Boussinesq limit, the Green problem can be
formulated with a single non-dimensional parameter, criticality (assuming constant vertical
shear and Brunt-Väisälä frequency, and using tropopause height as the vertical scale). In
this limit, the most unstable wavelength has a similar dependency on criticality as the
Charney problem (see Fig. 3 in Green 1960). This may suggest that for more general time-
mean flows, the most unstable wavelength can increase with criticality when it varies near
1, and saturates when criticality is much larger than 1. Nevertheless, the tropopause has
a significant impact on the vertical structure of the unstable waves. For example, a more
realistic representation of tropopause leads to larger upper-level amplitudes for the unstable
waves (Simmons and Hoskins 1976, 1977).

There are of course alternate theories that may explain the behavior of the horizontal
and vertical scales of eddies in modeling. One alternative is that the horizontal scale is
proportional to the wavelength of the (first) neutral mode in the Charney problem, which also
increases with criticality (see Fig. 1). Exact analytical expressions exist for the wavelength
of neutral modes. For example, the wavelength of the first neutral mode is $4\pi L R \xi / \sqrt{2 \xi + 1}$
(Burger 1962). Another alternative is that the vertical scale of eddies is determined by the
baroclinic adjustment hypothesis (Lindzen and Farrell 1980b; Held 1982), that the vertical
scale is the minimum height above the ground up to which eddies need modify the flow in
order to neutralize it. Applying this hypothesis to the Charney problem leads to a vertical
scale $H_p \ln(1 + \xi)$. If we replace the criticality in these formulas with an “effective” criticality
$A \xi$, it is difficult to check which formula works better.
It is interesting that, although the inverse cascade is not important in most of our simulations, the Rhines scale still serves as a good estimate for the eddy scale in most simulations (Fig. 7), despite that the Rhines scale is usually thought of as a cross-over between homogeneous turbulence and Rossby wave regimes. The result here suggests that the Rhines scale may not crucially depend on the presence of a classical inverse cascade and may have a somewhat more general applicability. The effects of friction are also likely to be important in any real simulation (Sukoriansky et al. 2007).

In our study, criticality varies between about 1 and 3. Consider Jansen and Ferrari’s (2013b) scaling for criticality \( \xi \sim (f \tau_r)^{-1/5}(a/L_R)^{3/5} \), where \( \tau_r \) is the time scale of thermal forcing and is inversely proportional to \( \gamma \) in our experiments, and \( a = f/\beta \) is on the order of the planetary radius. The scaling explains why we have to vary \( \gamma \) by several orders in order to vary criticality, as criticality is only dependent on \( \gamma \) to one-fifth power. Note that we cannot get very high criticality by increasing the thermal forcing rate only. Rather, to get very high criticality, it may be better to increase \( a/L_R \). Indeed, we obtained very high criticality and a substantial inverse cascade in a simulation with planetary radius 10 times Earth’s. Combined with results from this study, it follows that Earth’s atmosphere is within the weakly nonlinear regime because the Rossby radius is on the same order as the planetary radius, and also that the most unstable wave can adjust its horizontal and vertical scale so as to stay weakly nonlinear.

**Acknowledgments.**

We thank Isaac Held, Pablo Zurita-Gotor and Malte Jansen for helpful comments and discussions. We also thank Thomas Flannaghan and Spencer Hill for carefully reading the manuscript. We finally thank the two anonymous reviewers for their helpful comments. The numerical simulations were performed on GFDL’s computer system. This work was funded by the NSF under grant AGS-1144302 and the NOAA under grant NA08OAR4320752. The statements, findings, conclusions, and recommendations are those of the authors and do not
necessarily reflect the views of the NSF, the NOAA, or the U.S. Department of Commerce.
APPENDIX A

Barotropic EKE spectrum and eddy scale

The barotropic EKE spectrum $E_n$ is computed using spherical harmonics as basis functions and the details are described by Boer and Shepherd (1983). Briefly, it is defined as

$$E_n = \frac{1}{4} \sum_{l=-n}^{n} \langle \left| \{u'_{bt}\}_{n,l} \right|^2 + \left| \{v'_{bt}\}_{n,l} \right|^2 \rangle, \quad (A1)$$

where $u'_{bt}$ and $v'_{bt}$ denote the zonal and meridional components of barotropic eddy velocity, which is velocity’s derivation from zonal mean and vertically averaged from 150 mb to 1 bar. The angle brackets denote a time mean. The subscripts $n, l$ denote the spectrum component of the fields with total wavenumber $n$ and zonal wavenumber $l$. The wavenumbers are nondimensional. For the T42 resolution, $n$ ranges from 1 to 42. $E_n$ has unit of m$^2$ s$^{-2}$ and the sum of $E_n$ for all wavenumbers equals the surface area averaged barotropic EKE.

The energy-containing eddy scale, which we usually refer as eddy scale for short, is calculated from the “inverse centroid” introduced by Schneider and Walker (2006, 2008). Compared with the normal centroid, this definition gives a scale closer to the peak in $E_n$. To get the eddy scale, the energy containing wavenumber $n_e$ is first calculated as as

$$n_e(n_e + 1) = \frac{\sum_n E_n}{\sum_n [n(n + 1)]^{-1} E_n}, \quad (A2)$$

and then the eddy scale $L_e$ is given by

$$L_e = \frac{\pi a}{\sqrt{n_e(n_e + 1)}}, \quad (A3)$$

where $a$ is the planetary radius. Using the vertical averaged EKE spectrum defined below in (B2) instead of barotropic EKE spectrum in (B2) gives a slightly smaller eddy scale, which does not change our result much.
Spectral EKE budget

The spectral EKE budget can then be written as

\[ \partial_t \text{EKE}_n = G + T_{EE} + T_{EM} - D + \text{Resi}, \quad \text{where} \]

where \( EKE_n \) denotes the vertically averaged EKE at total wavenumber \( n \), the range between \( p_1 \) and \( p_2 \) is from 150 mb to 1 bar\(^4\), and the angle brackets denote a time mean. By dividing the vertical integral by the pressure range \( \Delta p = p_2 - p_1 \), the units for \( EKE_n \) are \( \text{m}^2 \text{s}^{-2} \). In the right hand side terms above,

\[ G = -\frac{1}{2} \sum_{l=-n}^{n} \left< \int_{p_1}^{p_2} \frac{R}{p} \text{Re}(T'_{n,l} \omega'_{n,l}) dp \right> / \Delta p, \]

denotes the energy conversion from eddy available potential energy to eddy kinetic energy, and we refer it as EKE generation for short,

\[ T_{EE} = \frac{1}{2} \sum_{l=-n}^{n} \left< \int_{p_1}^{p_2} \text{Re}[-u''_{n,l}(\mathbf{u}' \cdot \nabla \mathbf{u}')_{n,l} - v''_{n,l}(\mathbf{u}' \cdot \nabla \mathbf{v}')_{n,l}] dp \right> / \Delta p \]

denotes nonlinear eddy-eddy kinetic energy transfer,

\[ T_{EM} = \frac{1}{2} \sum_{l=-n}^{n} \left< \int_{p_1}^{p_2} \text{Re}[-u''_{n,l}(\mathbf{u}' \cdot \nabla \mathbf{u}')_{n,l} - v''_{n,l}(\mathbf{u}' \cdot \nabla \mathbf{v}')_{n,l} - u''_{n,l}(\overline{\mathbf{u}} \cdot \nabla \mathbf{u}')_{n,l} - v''_{n,l}(\overline{\mathbf{u}} \cdot \nabla \mathbf{v}')_{n,l}] dp \right> / \Delta p \]

\(^4\)In actual calculations, we use the model’s sigma coordinate to approximate the pressure coordinate. This avoids the complications related to the lower boundary condition in pressure coordinate (the planetary surface pierces the near-surface pressure levels). This is a decent approximation as long as the surface pressure variation is small enough.
denotes the eddy/mean-flow kinetic energy transfer, and

\[ D = \frac{1}{2} \sum_{l=-n}^{n} \left( \int_{p_1}^{p_2} k_v(p)(u'_{n,l}u'_{n,l} + v'_{n,l}v'_{n,l}) dp \right) / \Delta p \]

denotes the dissipation by Rayleigh damping of low level winds, where \( k_v(p) \) is the linear damping coefficient as a function of pressure. The contributions from the Coriolis terms, dissipation by hyper-diffusivity, and gravitational potential energy flux through the boundaries are absorbed in the residual term Resi. In a statistical steady state, the sum of the right hand side terms should be zero. Detailed derivations and notations can be found in Lambert (1984) or Koshyk and Hamilton (2001). The corresponding formulation using Fourier series as basis functions in a Cartesian geometry is given in Jansen and Ferrari (2012).

Fig. 13a shows the terms in the spectral EKE budget for the control (\( \gamma = 1 \)) simulation with \( \delta_y = 60 \) K. The EKE generation is predominately transferred into the zonal-mean flow or dissipated by friction at similar scales. The eddy-eddy interactions do transfer some energy upscale, but this does not increase the eddy scale substantially from EKE generation scale. To quantify the role of eddy-eddy interaction, we define EKE generation scale in a similar way as the eddy scale defined in Appendix A. The EKE generation wavenumber is calculated as an inverse centroid of \( G_n \) similar to Eq. (A2), and the EKE generation wavenumber is converted into the EKE generation scale by Eq. (A3). As can be seen in Fig. 13b, the energy generation wavenumber follows closely the peak of \( G_n \), which moves towards larger scales when \( \gamma \) increases from 0.1 to 10.
REFERENCES


1. A contour plot showing the growth rate against the inverse wavenumber $1/k$ and criticality $\xi$ for the Charney model. The solid line marks the most unstable wavenumber determined numerically, and the dashed line shows our scaling estimate $1.54\xi/(1 + \xi)$. The horizontal and vertical length scales are nondimensionalized by Rossby radius $L_R$ and density scale height $H_\rho$ respectively. The contour interval is 0.02, and it is nondimensionalized by $H_\rho \partial_z U/L_R$.

2. (a) Vertical structure of the most unstable Charney wave when $\xi = 1$. Dashed line is the amplitude of the perturbation stream function $|\psi|$, solid line is heat flux $\nabla' \partial_z \psi$, and dotted line is PV flux. All three curves are normalized by their maximum values. (b) Vertical scale of the most unstable wave $h_\theta$ defined in Eq. (22) vs. criticality: solid line is from numerical results and dashed line is our scaling estimate $1.09\xi/(1 + \xi)$.

3. Mean zonal wind (contours) and EKE (colors) for the control run ($\gamma = 1$) with $\delta_y = 60$ K from full GCM (left) and reduced GCM (right). The contour interval for zonal wind is 5 m/s and for EKE is 100 m$^2$/s$^2$. The zero zonal wind contour is thick, and negative contours are dashed. The tropopause calculated from isentropic diagnostics is marked by the thick black line, and the WMO tropopause is marked by the thick gray line.

4. Extratropical bulk stability $\Delta_v = \overline{\theta_p} \theta'(\overline{\rho_t - \rho_\beta})$ vs. scaled meridional temperature gradient $\Delta_h = |(f/\beta)\partial_y \theta'|$ evaluated at 850 mb for different combinations of $\gamma$ and $\delta_y$. The control runs with $\gamma = 1$ but different $\delta_y$ are connected with a black line. Each series of runs with varying $\gamma$ and with the same $\delta_y$ are connected with a blue line, and the corresponding value $\delta_y$ is indicated by the label beside.
Results from simulations with $\delta y = 60$ K and varying $\gamma$. (a) Eddy scale $L_e$ and Rossby radius $L_R$, all normalized by their control run values ($\gamma = 1$). (b) Tropopause height $H_p$ based on isentropic diagnostics, WMO tropopause height $H_p^{\text{WMO}}$ and vertical stability measure $N_p$, all normalized by their control run values.

Criticality $\xi$, Rossby radius $L_R$, and eddy scale $L_e$ as a function of $\gamma$ and $\delta y$. (a) Best Rossby radius estimate $C_R L_R$ and (b) Rhines scale estimate $C_\beta L_\beta$ for eddy scale. The empirical constants $C_R$ and $C_\beta$ are determined to be 2.6 and 3.2 respectively by a least square fit to the modeling results. The legends denote the series of experiments with a certain $\delta y$ and varying $\gamma$.

For the Earth-like settings with $\delta y = 60$ K, $\Omega = \Omega_0$ and $r = a_0$, and varying thermal forcing $\gamma = 0.1, 1, 10$, comparison of EKE spectrum between the full and reduced GCMs. The solid lines denote the full GCM, and the dashed lines denote the reduced GCM. The vertical black lines indicate energy containing wavenumber for the full GCM runs.

(a) Comparison of eddy scale $L_e$ in the full and reduced GCMs. (b) Comparison of eddy scale and eddy kinetic energy generation scale (see Appendix A and B for definitions) in the full GCM. The legends denote the series of experiments with a certain $\delta y$ and varying $\gamma$.

Eddy potential temperature flux $\nu'\theta' \cos(\phi)$ (colors) and mean potential temperature (contours) from simulations with $\delta y = 60$ K, and $\gamma = 0.1$ (top), 1 (middle), and 10 (bottom). The thick black lines denote the tropopause determined from isentropic diagnostics.
11  (a) The vertical structure of eddy potential temperature flux $\overline{v'\theta'}\cos(\phi)$ averaged over the baroclinic zone and normalized by its maximum value from simulations with $\delta_y=60$ K, and $\gamma=0.1, 1, 10$. (b) The vertical scale of the wave $h_\theta$ defined in Eq. (22) for all series of simulations with different $\delta_y$ and varying $\gamma$, normalized by the scale height $H_\rho$. The horizontal dashed lines in (a) denote the pressure range over which we fit potential temperature flux profile to an exponential decay with height to get the vertical scale of the wave $h_\theta$.

12  (a) Best estimate using $C_\xi L_R$ vs observed eddy scale in all runs, with $C=1.8$. (b) Best estimate using the form $C L_R(A_\xi)/(1 + A_\xi)$ vs observed eddy scale in all runs, with $C=8.7$ and $A=0.36$. The legends denote the series of experiments with a certain $\delta_y$ and varying $\gamma$.

13  (a) Spectral EKE budget for the simulation with $\gamma = 1$ and $\delta_y=60$ K: conversion from eddy APE to EKE ($G$, solid red), EKE transfer by eddy-eddy interactions ($T_{EE}$, solid green), energy transfer between zonal-mean flow and eddy ($T_{EM}$, solid blue), explicit dissipation by Rayleigh damping of winds ($-D$, dashed black), and the residual term (Resi, dashed cyan). The residual term is not significant compared to other leading order terms at wavenumbers away from the grid scale. (b) Conversion from eddy available potential energy to EKE for three simulations with $\gamma=0.1, 1, 10$ and $\delta_y=60$ K. The vertical black lines indicate the corresponding EKE generation wavenumber.
**Fig. 1.** A contour plot showing the growth rate against the inverse wavenumber $1/k$ and criticality $\xi$ for the Charney model. The solid line marks the most unstable wavenumber determined numerically, and the dashed line shows our scaling estimate $1.54\xi/(1 + \xi)$. The horizontal and vertical length scales are nondimensionalized by Rossby radius $L_R$ and density scale height $H_\rho$ respectively. The contour interval is 0.02, and it is nondimensionalized by $H_\rho \partial_z U / L_R$. 
Fig. 2. (a) Vertical structure of the most unstable Charney wave when $\xi = 1$. Dashed line is the amplitude of the perturbation stream function $|\psi|$, solid line is heat flux $\bar{v'}\partial_z \psi$, and dotted line is PV flux. All three curves are normalized by their maximum values. (b) Vertical scale of the most unstable wave $h_\theta$ defined in Eq. (22) vs. criticality: solid line is from numerical results and dashed line is our scaling estimate $1.09\xi/(1 + \xi)$.
Fig. 3. Mean zonal wind (contours) and EKE (colors) for the control run ($\gamma = 1$) with $\delta_y = 60$ K from full GCM (left) and reduced GCM (right). The contour interval for zonal wind is 5 m/s and for EKE is 100 m$^2$/s$^2$. The zero zonal wind contour is thick, and negative contours are dashed. The tropopause calculated from isentropic diagnostics is marked by the thick black line, and the WMO tropopause is marked by the thick gray line.
Fig. 4. Extratropical bulk stability $\Delta_v = \frac{\partial \tilde{p}}{\partial \tilde{\theta}^s (\tilde{p}_t - \tilde{p}_s)}$ vs. scaled meridional temperature gradient $\Delta_h = |(f/\beta)\partial_y \tilde{\theta}^s|$ evaluated at 850 mb for different combinations of $\gamma$ and $\delta_y$. The control runs with $\gamma = 1$ but different $\delta_y$ are connected with a black line. Each series of runs with varying $\gamma$ and with the same $\delta_y$ are connected with a blue line, and the corresponding value $\delta_y$ is indicated by the label besides.
Fig. 5. Results from simulations with $\delta_y = 60$ K and varying $\gamma$. (a) Eddy scale $L_e$ and Rossby radius $L_R$, all normalized by their control run values ($\gamma = 1$). (b) Tropopause height $H_p$ based on isentropic diagnostics, WMO tropopause height $H_p^{WMO}$ and vertical stability measure $N_p$, all normalized by their control run values.
Fig. 6. Criticality $\xi$, Rossby radius $L_R$, and eddy scale $L_e$ as a function of $\gamma$ and $\delta_y$. 
Fig. 7. (a) Best Rossby radius estimate $C_R L_R$ and (b) Rhines scale estimate $C_\beta L_\beta$ for eddy scale. The empirical constants $C_R$ and $C_\beta$ are determined to be 2.6 and 3.2 respectively by a least square fit to the modeling results. The legends denote the series of experiments with a certain $\delta_y$ and varying $\gamma$. 
Fig. 8. For the Earth-like settings with $\delta_y = 60 \text{ K}, \Omega = \Omega_0$ and $r = a_0$, and varying thermal forcing $\gamma = 0.1, 1, 10$, comparison of EKE spectrum between the full and reduced GCMs. The solid lines denote the full GCM, and the dashed lines denote the reduced GCM. The vertical black lines indicate energy containing wavenumber for the full GCM runs.
Fig. 9. (a) Comparison of eddy scale $L_e$ in the full and reduced GCMs. (b) Comparison of eddy scale and eddy kinetic energy generation scale (see Appendix A and B for definitions) in the full GCM. The legends denote the series of experiments with a certain $\delta_y$ and varying $\gamma$. 
Fig. 10. Eddy potential temperature flux $\overline{\omega' \theta'} \cos(\phi)$ (colors) and mean potential temperature (contours) from simulations with $\delta_y = 60$ K, and $\gamma = 0.1$ (top), 1 (middle), and 10 (bottom). The thick black lines denote the tropopause determined from isentropic diagnostics.
Fig. 11. (a) The vertical structure of eddy potential temperature flux $\overline{v'\theta' \cos(\phi)}$ averaged over the baroclinic zone and normalized by its maximum value from simulations with $\delta_y = 60$ K, and $\gamma = 0.1, 1, 10$. (b) The vertical scale of the wave $h_\theta$ defined in Eq. (22) for all series of simulations with different $\delta_y$ and varying $\gamma$, normalized by the scale height $H_\rho$. The horizontal dashed lines in (a) denote the pressure range over which we fit potential temperature flux profile to an exponential decay with height to get the vertical scale of the wave $h_\theta$. 

\[
\frac{v'\theta' \cos(\phi)}{\text{max}[v'\theta' \cos(\phi)]}
\]

\[
\gamma = 0.1, 1, 10
\]

\[
\frac{h_\theta}{H_\rho}
\]

\[
\gamma = 30, 60, 90, 120 \text{ K}
\]
Fig. 12. (a) Best estimate using $C\xi L_R$ vs observed eddy scale in all runs, with $C=1.8$. (b) Best estimate using the form $CL_R(A\xi)/(1 + A\xi)$ vs observed eddy scale in all runs, with $C=8.7$ and $A=0.36$. The legends denote the series of experiments with a certain $\delta_y$ and varying $\gamma$. 
Fig. 13. (a) Spectral EKE budget for the simulation with $\gamma = 1$ and $\delta_y = 60$ K: conversion from eddy APE to EKE ($G$, solid red), EKE transfer by eddy-eddy interactions ($T_{EE}$, solid green), energy transfer between zonal-mean flow and eddy ($T_{EM}$, solid blue), explicit dissipation by Rayleigh damping of winds ($-D$, dashed black), and the residual term (Resi, dashed cyan). The residual term is not significant compared to other leading order terms at wavenumbers away from the grid scale. (b) Conversion from eddy available potential energy to EKE for three simulations with $\gamma = 0.1$, 1, 10 and $\delta_y = 60$ K. The vertical black lines indicate the corresponding EKE generation wavenumber.