Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Dynamics of Atmospheres and Oceans
xxx (2006) xxx–xxx
dynamics of atmospheres and oceans
www.elsevier.com/locate/dynatmoce

Deducing the general circulation from basic concepts and a few empirical facts

Richard Grotjahn *

Department of Land, Air, and Water Resources (LAWR), UCD, One Shields Ave., University of California, Davis, CA 95616, USA

Abstract

The general circulation of the Earth's atmosphere is complex, linking a wide range of scales and physical processes. This special issue samples current research on a wide range of properties of the general circulation. This introductory article provides a context in which to place these articles by listing some principal balances such as: conservation of mass and momentum. Some implications and constraints are briefly described. The general circulation may be broadly deduced from those primary physical balances, dynamical constraints, and a few empirical facts (primarily the distribution of precipitation).

Keywords: Atmospheric circulation; Atmospheric physics; Radiation balance; Dynamics; Baroclinic instability; Angular momentum

1. Prologue

What is meant by the "general circulation"? For this special issue of Dynamics of Atmospheres and Oceans, we define the general circulation to be: large scales (everything >1000 km), persisting a season or longer, *and*, all the processes necessary to sufficiently explain (or directly maintain) these large scale, persistent circulations. What follows the '*and*' in that definition greatly widens the definition to include other variables than just wind. Solely using zonal average fields would not work—one needs zonally varying fields to maintain (and thus explain) the zonal mean, as has been demonstrated numerous times.

The question of *why* there should be a general circulation is relatively easy. The atmosphere develops a circulation in an attempt to reach a thermal equilibrium between *net* sources and sinks

* Tel.: +1 530 752 2246; fax: +1 530 752 1793.
E-mail address: grotjahn@ucdavis.edu.

0377-0265/$ – see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.dynatmoce.2006.05.002

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

of radiative heating and cooling. Consequently, the solar radiation absorbed has greater latitudinal variation than does the terrestrial radiation back to space (see Fig. 1a).

Of course neither radiative pattern is completely symmetric with respect to longitude or latitude. (See Fig. 1b and c. While these data have biases, they are easily sufficient to illustrate the qualitative discussion here.) The general pattern is partly explained by surface properties and the geometry of a rotating Earth, tilted with respect to the plane of the ecliptic. As much, if not more of the pattern seems driven by atmospheric circulations that transport heat, cause variable radiative properties (clouds, snow cover, etc.), and absorb and emit radiation. Oceanic circulations and topography further complicate the atmospheric circulation.

Despite the great complexity of the atmospheric and oceanic processes, one can deduce the general properties of the large scale atmospheric circulation by applying some balances augmented by a few empirical facts.

2. Deducing the general circulation

By exploiting the concept of balances, we can deduce much about the general circulation. Certain balances are believed to hold and these give rise to formulas and understanding as to why the observed circulation looks like it does.

One can invoke physical concepts; two examples being: radiative and angular momentum balance. Radiative balance states that when averaged over the surface of the globe, there will be no net heating up (or cooling down) of the Earth if the same amount of energy is emitted as is absorbed. While there appears to be evidence for global warming, that rate of temperature increase is still far less than the seasonal cycle at most places on Earth. Angular momentum balance has several implications for the general circulation. One implication is surface winds need a mixture of easterlies and westerlies so as not to apply a net torque upon the solid Earth. As a trivial example, if the surface winds were westerly everywhere, then the atmosphere would be applying a net torque upon the oceans and solid Earth, as one consequence the rotation of the Earth would speed up.

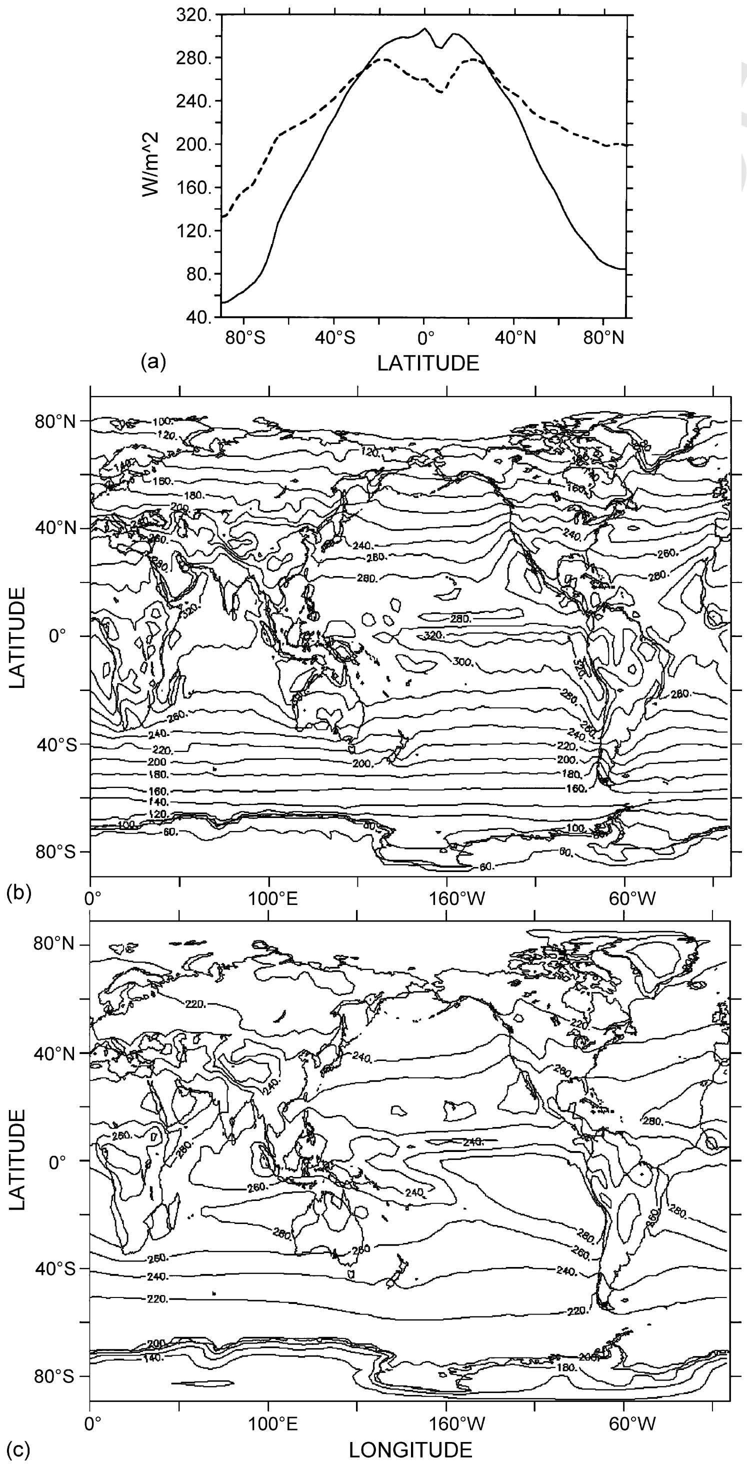
One can use balances, physical reasoning, and the observed precipitation to deduce some basic properties. Various laws, such as ideal gas and conservation of potential temperature for adiabatic motions may be applied. Mathematical simplifications, such as hydrostatic and geostrophic balances are useful. An overarching assumption shall be that the atmosphere can be considered annually stationary: any trends that remain when fields are averaged over a year are small and can be ignored for this discussion.

2.1. Radiation and temperature

Excluding global warming, the Earth does not heat up in the net over a year. So, one may expect a balance between absorption and emission of radiant energy. Absorption is complicated by variations in the reflectance, absorbance, and transmittance of various terrestrial constituents. Despite these complications, the orbital geometry is a primary cause for more absorption averaged

Fig. 1. Global radiative balance for the Earth. (a) Zonal and time average radiation at the top of the atmosphere. Solid line: incoming, solar, absorbed radiation; dashed line: upwelling, longwave radiation from the Earth. (b) Time average incoming, solar, absorbed radiation at the top of the atmosphere. (c) Time average upwelling, longwave radiation from the Earth. ECMWF Reanalysis 1 (ERA-15) monthly mean data from January 1979 through December 1993. Contour interval 20 W/m^2 in b and c.

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002



Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

over the tropics than averaged over polar regions. This fact creates a meridional temperature gradient. The loss of radiation back to space is complicated by several factors; for example, the infrared emission from the Earth is not from a single surface, but includes all levels in the atmosphere as well as the ocean and Earth's surface.

If the rate at which the atmosphere cools down by infrared emission were large compared to the temperature of the Earth divided by the length of a day, then the primary temperature gradient of the atmosphere would be from the subsolar point to the shadow side of the planet. Radiative cooling rates in the troposphere are around 2° per day, which is much less than the roughly 210–310 K temperatures of the troposphere. (This thermal inertia is largely maintained by radiative emission from the Earth's surface.) As a consequence, the diurnal cycle of temperature is small compared to the daily mean temperature. Also, the terrestrial longwave emission is similar to the long term average (Fig. 1c). Similarly, the 20–70 °C range in temperature observed from equator to pole in the troposphere exceeds the temperature change from midday to midnight. (The equator to pole temperature range varies with elevation in the atmosphere.) So, the temperature gradient is primarily meridional rather than zonal. And, longer term (seasonal) average differences in radiation absorbed and emitted therefore dominate the temperature distribution.

Defining net radiation as the local difference of absorption minus emission, net radiation is positive over much of the tropics, and negative in the polar regions (averaged over a year). This pattern of net radiation implies a meridional transport of heat, perhaps by some meridional motions. Such motions could conceivably remove the implied meridional temperature gradient, but if they did, the motions would also remove their driving mechanism.

In the stratosphere the temperature pattern is a bit different. Stratospheric temperatures are dominated by radiative and transport processes, whereas tropospheric temperatures are also influenced by latent heat release, and exchanges with the Earth's surface. The orbital tilt and spherical geometry of the Earth combine to give a rather flat distribution of incoming solar radiation in the high *summer* latitudes. Hence, temperatures have little meridioanl gradient in much of the upper stratosphere in mid-summer. At upper stratospheric levels, temperatures tend to be coldest near the winter pole. In the lower stratosphere, the coldest temperatures are technically in the tropics, owing to the much higher tropopause there than in polar regions. However, in mid and upper stratosphere elevations, temperatures tend to decrease from the summer polar region to the winter polar region during those seasons. In other seasons, the mid and upper level temperatures tend to decrease from the tropics towards both poles.

2.2. Pressure and geopotential

Both water vapor content as well as temperature in the troposphere decrease from tropics to the polar regions. Hence, the virtual temperature decreases in similar fashion. From the hypsometric equation, the meridional virtual temperature gradient thereby creates a meridional geopotential gradient. The meridional temperature gradient applies for each layer (above the boundary layer, say) in the troposphere so the thickness variation in each layer is additive. Thus, the meridional gradient of geopotential increases with elevation. Near the ground, the pressure pattern has comparable longitudinal and latitudinal variations. However, the pressure at elevations above roughly 2 km (and geopotential height of pressure surfaces for $P < 850$ hPa) also decrease from tropics to polar regions.

By similar reasoning, in the stratosphere, the meridional gradient of geopotential decreases with increasing elevation in the summer hemisphere and tends to increase with elevation in the winter hemisphere.

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

Where there is a horizontal gradient of pressure there will be a pressure gradient force (PGF). In the troposphere the PGF is directed towards the poles and provides the impetus to initiate the circulation.

2.3. Winds

The Earth rotates such that observed velocities outside the tropics are subject to a Coriolis force comparable to the poleward PGF. Hence, the meridional gradient of pressure (and geopotential) implies westerly winds from geostrophic balance. Since the gradient increases with height then so must the westerly winds in the troposphere. It was established above that heat must be transported poleward to explain the latitudinal distribution of net radiation. If the heat is transferred in a way that employs meridional motion, then zonal winds are created to conserve angular momentum during such meridional motions. Are the zonal winds implied by the virtual temperature distribution consistent with zonal winds created by large scale angular momentum conserving meridional motions? Presumably a balance can be struck whereby stronger meridional motions more effectively remove the equator to pole temperature gradient thereby lessening the zonal thermal winds, and vice versa.

Conservation of angular momentum by meridional motions can create very large zonal winds. For example, if an air parcel having zero zonal motion at the equator were brought to 30N latitude, an angular momentum conserving zonal wind, U_{amc} would equal 134 m/s. This is twice the time mean speed of the strongest portion of the subtropical jet stream found at a similar latitude. Obviously, one can reduce this acceleration by internal viscosity and vertical mixing of momentum by turbulence or convection. The greater the length of time over which friction and mixing can act, the greater the reduction possible. For realistic estimates of internal friction, *meridional* wind speeds of a few cm/s keep *zonal* winds within observed bounds for motions spanning the tropics and subtropics (Grotjahn, 1993; Section 6.2.3.2) In the tropical upper troposphere observed meridional velocities average 5–10 times larger, suggesting that either the friction is too weak and/or other mechanisms are important. Another mechanism might be convection that mixes momentum vertically. This may be happening within ‘tropical plumes’ created in advance of upper level troughs that penetrate deeply into the tropics; air is accelerated poleward, but actively growing clouds moderate the angular momentum by vertical mixing.

2.4. Heat transport and mass balance

Surface pressure values in polar regions do not have a significant trend. So, a meridional circulation that transports heat poleward would be expected to have compensating equatorward motion, otherwise mass would accumulate in the polar region.

The previous section suggests that meridional velocities are small based on observations and angular momentum considerations. Can such small meridional motion still transport sufficient heat? This question introduces mass and/or temperature field considerations. If the air is moving both poleward and equatorward, then net sensible heat transport depends on the difference in dry static energy (DSE) between the north and southward moving air. (Latent heat processes from water evaporation, transport, and condensation are being ignored for the moment.) Consequently, to have the same net heat flux, a smaller meridional velocity must be balanced by a larger difference in DSE between the north and southward moving air. If the larger DSE difference were due to temperature changes, then that means larger static stability. Alternatively, if the vertical separation were made larger between opposing meridional motions then the DSE difference would be larger and the meridional motions need not be as strong. So in theory such small meridional velocities can transport heat if the DSE change is large enough. Clearly the DSE variation depends on the temperature variation, both horizontally and vertically.

The DSE will increase with height for a statically stable atmosphere. If zonally and vertically integrated mass transport in the meridional is to be zero, then a poleward DSE heat flux occurs when the poleward moving air has higher DSE than the equatorward moving air. Hence, if the upper level flow has larger DSE than the low level flow, the upper level flow would need to be poleward.

Water mass conservation is a different issue because water exists in three different states on Earth. Water could transport heat poleward if it were evaporated in lower latitudes (removing heat there), transported poleward by the winds, and condensed at higher latitudes (releasing the heat there). The mass of atmospheric water in vapor form would not be conserved. Water could conceivably be condensed, fall as precipitation, and run off, returning to the lower latitude in liquid form. While this may seem like a reasonable way to accomplish the heat transport, a major drawback is that this mechanism works against the dry heat transport by a zonally averaged circulation. To see this, one defines a moist static energy (MSE) by adding latent heat content to the DSE. The scale height of water vapor ($\sim$2 km) is much smaller than that of dry air ($\sim$8.5 km). So for high enough moisture content, the moist static energy could conceivably decrease with height and if so, poleward heat transport would be accomplished if the poleward moving air were at *low* levels—opposite to the pattern deduced for DSE. In much of the tropics, MSE *is* observed (e.g. Gustafson and Weare, 2004) to decrease with height in the lower troposphere ($P > 700$). It is conceivable that a circulation with surface poleward motion and mid to lower tropospheric equatorward return flow might transport the necessary heat. Such a circulation might also have a lot of boundary layer friction to control the development of excessive zonal winds from angular momentum conservation. There are potential problems with this idea. First, a mechanism (vertical mixing or solar absorption perhaps) may be needed to overcome a meridional gradient of MSE that may oppose the motion. Second, it may conflict with the observed tropical precipitation pattern.

The observed distribution of precipitation is crucial to deducing the Earth's general circulation. The primary maximum in precipitation is near the equator and relative minima occur in the subtropics; this pattern eliminates the zonal mean meridional 'boundary layer' cell postulated in the previous paragraph. Furthermore, secondary maxima of precipitation in northern and southern middle latitudes must to be explained and that selects for a different midlatitude circulation type than in the tropics.

2.5. Surface torques

When wind blows across the surface of the Earth friction slows down the wind while imparting a torque upon the Earth. If the surface winds were to blow the same direction as the rotation everywhere on a planet with no topographic features, then the winds would be accelerating the planetary rotation. Such reasoning suggests that surface winds on the Earth are unlikely to be westerly (or easterly) everywhere. The simplest conclusion might be that there is a balance between the torque applied in regions of surface easterlies versus the torque applied by regions of surface westerlies.

Surface torques are applied not only by winds blowing across the surface of the Earth but also by surface pressure differences across mountain ranges. Restricting our attention to zonal flows, then westerly winds impart an eastward torque upon the Earth. A surface pressure pattern of low on the windward side and high on the lee side of a topographic feature imparts a "mountain torque" that opposes the torque by the winds.

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

How much of a surface pressure difference across the mountain is needed to balance a given zonal wind? Is such a pressure difference "realistic"? The Earth has several major mountain ranges that may provide such a balance, of particular interest for westerly winds are the north–south oriented ranges such as the Andes and Rockies. To gain a sense of the magnitudes involved, it is instructive to make a crude calculation that balances surface stress τ from the winds against mountain torque. Specifically, for the Andes, the balance (see Peixoto and Oort, 1992; Section 11.1.4) is roughly:

$$\int_0^{2\pi}\int_{-\pi/3}^{0} \tau a^3 \cos^2(\varphi)\,\mathrm{d}\varphi\,\mathrm{d}\lambda = \int_0^{z_{\mathrm{top}}}\int_{-\pi/3}^{0} \{P_{\mathrm{E}} - P_{\mathrm{W}}\} a^2 \cos(\varphi)\,\mathrm{d}\varphi\,\mathrm{d}z \qquad (1)$$

where a is the Earth's radius (~6370 km), φ the latitude, λ the longitude, z_{top} an elevation above the highest point of the Andes, P_{E} the surface pressure on the east side of the Andes, while P_{W} the corresponding pressure on the west side. (Where z exceeds the local height of the Andes, the pressure difference is zero.) Other Southern Hemisphere topography is ignored. Imagine surface winds of the Southern Hemisphere that are a uniform 10 m/s over all longitudes and over latitudes from equator to 60S. Further assume the pressure difference is uniform with elevation and latitude and important for the lowest 3 km. Approximate the surface stress using a drag coefficient of 10^{-3} and average surface density of 1.25 kg m^{-3} such that $\tau \sim 1.25 \times 10^{-1}$ kg m^{-1} s^{-2}. The resultant pressure difference $\mathrm{d}P$ $(= P_{\mathrm{E}} - P_{\mathrm{W}})$ is ~24 hPa. Such a pressure change is within the range of *sea level* pressure observed and therefore is not unrealistically large. Unfortunately, the observed sea level pressure change across most of the Andes has the opposite sign, implying that surface pressure at other Andean elevations might not be able to balance surface westerlies. (Note: such westerlies would exert stress upon the ocean, causing a zonal change in the height of sea level, but the change would largely pass the wind torque along to the solid Earth.) Another problem is westerlies over the Andes would be up, not down the sea level pressure gradient.

To summarize, the discussion of balances and key empirical facts implies the following six properties:

(i) Incoming radiation makes the tropics warmer than the poles, which in turn creates an equatorward pressure gradient (and hence poleward pressure force) above some level in the atmosphere.
(ii) The pressure gradient and Earths rotation lead the T gradient to have associated westerly winds that increase with elevation.
(iii) Net radiation implies that there must be poleward transport of heat.
(iv) Mass balance requires the poleward heat transport to have "equal" mass components north and south.
(v) Angular momentum balance anticipates "roughly equal" areas of easterly and westerly surface winds; and/or high pressure on the lee side and low on the windward side of major north–south oriented mountain ranges.
(vi) Water mass balance and transport are needed to explain precipitation maxima in the midlatitudes and tropics.

3. Thought experiments

To help understand the context of what follows in this special issue, it is useful to have an idea of what circulations one might expect. One can make quite a bit of progress by performing a series of "thought experiments". The discussion that follows is qualitative, but it expands and makes visible the quantitative principles listed above. Some candidate, "imaginary" circulations are proposed until one qualitatively satisfies these six properties discussed above.

3.1. *Idea #1: pure northward motion*

While this motion transports sensible heat poleward, by bringing warm air from the tropics towards the poles, it has already been rejected above due to a lack of mass conservation. Furthermore, angular momentum is not balanced, unless the motion were heavily slowed by friction. There is no mechanism to create the observed precipitation pattern.

3.2. *Idea #2: alternating meridional flows*

In an effort to solve the mass balance problem of idea #1, one might propose meridional circulations that alternate between northward and southward motion (Fig. 2a) through the depth of the troposphere (or perhaps even more of the atmosphere). The arrows shown in Fig. 2a are trajectories. Heat is now mixed both directions, cooling tropical regions directly while warming polar regions. Mass balance is satisfied by matching the mass of poleward and equatorward moving air.

Unfortunately, angular momentum balance might require slow meridional motion to occur mainly within a boundary layer. The motion above the boundary layer would be strongly 'tilted' in the zonal direction as shown in Fig. 2a. In order to avoid a net torque upon the Earth, the strong westerlies (from angular momentum conservation) in high latitudes would be balanced by strong easterlies in low latitudes. If there is vertical mixing of momentum (by convection, if not a large

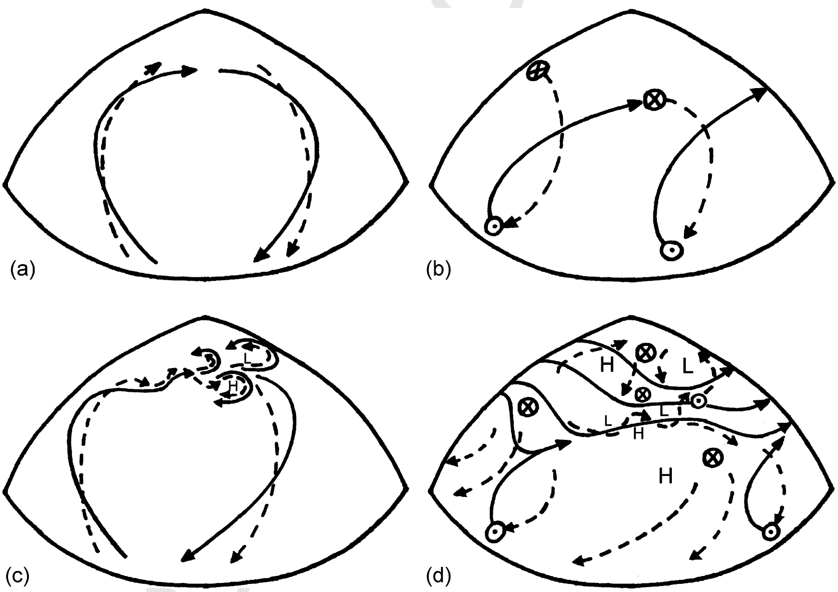


Fig. 2. Schematic diagrams of selected candidate circulations generally showing proposed air parcel trajectories. Solid lines show upper level flow, dashed lines show low level flow. Circles enclosing a dot suggest upward motion; circled × for downward. (a) 'Idea #2' alternating flows; (b) 'idea #3', 'one big Hadley cell'; (c) 'idea #4' barotropic eddies; (d) 'idea #6 hybrid' circulation that combines baroclinic eddies in high latitude with a 'Hadley' circulation in low latitudes.

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

scale overturning) then the winds are somewhat moderated. This candidate circulation would develop large upper level winds similar to the jet streams, however, it has no obvious mechanism by which to limit the jet to subtropical or midlatitude locations.

This candidate would have trouble explaining other important observed features. The midlatitude precipitation maximum might be explainable as follows. The poleward moving air in the boundary layer might be quite moist, as that air migrates poleward it cools (by net radiation, and by surface exchange) to a point where drizzle might occur. Explaining the midlatitude precipitation maxima as drizzle is a bit of a stretch! This scheme also has no obvious explanation for the tropical precipitation maximum. Finally, the predicted strong upper level equatorial easterlies are really not seen; the closest phenomenon being the Indian Ocean monsoon jet.

While this circulation is rejected, it does suggest trying alternating motions in the vertical and in the horizontal.

3.3. Idea #3: "one big Hadley cell"

A circulation having alternating motion in the vertical is the "Hadley cell". The Hadley cell circulation could be similar at all longitudes and form a loop: with rising in the tropics, poleward motion in the upper troposphere, sinking at higher latitudes, and equatorward motion in the lower troposphere. This Hadley cell is depicted in Fig. 2b. The schematic trajectories depict how poleward motion has westerly wind acceleration while equatorward motion in a boundary layer (subject to greater frictional dissipation) has less easterly acceleration. If mountain torques are absent, the wind might have no net torque upon the Earth's rotation by a careful adjustment of where the location lies of zero surface zonal wind. Mass is conserved by matching northward and southward mass fluxes. Heat transport is accomplished if the poleward flow has higher MSE than the equatorward flow. This is possible if the poleward flow is at tropospheric levels high enough so that the gravitational potential in MSE overcomes any high moisture content at low levels. Another reason for having a high troposphere poleward flow is to develop strong westerlies at jet stream level. The rising motion in the equatorial region provides an obvious explanation for the equatorial precipitation maximum (though details are missing like overcoming the vertical gradient of MSE within thunderstorm 'hot towers').

Some observed features are problematic. If this 'big Hadley cell' extends to polar regions, how would the jet streams occur where they do and not much further poleward? The zonal wind speeds reached would be unrealistically large in mid- and higher latitudes. Finally, there is no obvious explanation for the observed midlatitude precipitation maximum.

So while this circulation has promise in the tropics, it cannot explain all the general circulation.

3.4. Idea #4: barotropic eddies

The alternating meridional motions of Fig. 2a lead to very large horizontal wind shears. At some point, large shears may become hydrodynamically unstable. A simple approximate determinate for (barotropic) instability is that the meridional gradient of absolute vorticity Q, change sign in the domain. In spherical geometry (Baines, 1976) Q becomes:

$$Q = \beta - \frac{1}{a}\frac{\partial}{\partial\varphi}\left\{\frac{1}{a\,\cos(\varphi)}\frac{\partial}{\partial\varphi}(U\,\cos(\varphi))\right\} \tag{2}$$

where a is the Earth's radius, φ the latitude, β the meridional gradient of the Coriolis parameter. (Q is the absolute vorticity gradient for a purely zonal flow, the contribution by meridional motion to Q

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

has been ignored.) β decreases while the horizontal shear of U increases for increasing magnitude of latitude. However, since shear increases with latitude, the relative vorticity of such a zonal flow is negative. For the Earth, a zonal flow constructed from angular momentum conservation would have $Q=0$ at all latitudes! This result is a consequence that flow also conserves absolute vorticity. If some internal viscosity is added, then $Q<0$ at all latitudes and the flow is stable to barotropic eddies according to classical reasoning, even if large horizontal shear is still present. But these assumptions are too simplistic.

If topography is added, then waves of a scale similar to the topography may be forced and these would interact and grow, releasing some horizontal shear instability. Also, viscosity may lead to form drag instability (e.g. Frederiksen and Frederiksen, 1990). If moisture is included, perhaps through moist convective heating, then the condition that Q change sign may be "neither a necessary nor a crucial condition for instability for a barotropic flow" (Mak, 1983; p. 2355). In short, it may be that barotropic eddies could form, by several different mechanisms.

This idea (Fig. 2c) might appear to transport heat poleward by motion in both directions (warm air poleward, cold air equatorward). However, the colder air would have smaller thickness (between isobaric surfaces) which in turn creates zonal geopotential gradients which in turn creates geostrophic winds. The consequence of those winds is unclear, perhaps the whole pattern migrates westward, perhaps it breaks down into multiple eddies.

The scheme can satisfy mass balance. A mean horizontal zonal velocity can be added to the flow in such a way as to create compensating amounts of easterly and westerly surface torque. Friction in the boundary layer of barotropic lows implies low level convergence with consequent rising motion that may lead to precipitation, though it is unclear how this gets focused to midlatitudes except by restricting eddies to higher latitudes where some threshold of horizontal shear is crossed.

This idea shows some promise since it recognizes that hydrodynamic instability can play a role. However, there are some critical flaws. First, how the heat transport occurs is unclear. In of themselves, barotropic eddies do not transport heat. However, one might argue that radiative gain (loss) in the tropical (polar) regions leads to a net heat flux if the resident time in each region is long enough for sufficient radiative temperature changes. Second, the energetics of such lows would be a bit strange; while radiation-induced temperature changes drive meridional motion that creates zonal wind shear that is unstable to barotropic eddies, kinetic energy is lost by friction and conversion to eddy potential energy. Third, a mechanism for the tropical precipitation maximum is missing.

3.5. Idea #5: vertically tilted eddies

The barotropic eddies thought experiment invoked instability of horizontally sheared flow. The meridional temperature gradient requires vertical shear if thermal wind balance holds. As for horizontal shear, there is a classical necessary condition for instability of vertically sheared flows. Reasoning by analogy to Cartesian geometry, an appropriate governing equation for stability of a zonal mean flow may be deduced from a potential vorticity conserving governing equation (e.g. Grotjahn and Castello, 2002). Using a streamfunction (Ψ) to represent the zonal mean flow, then

$$\frac{\partial}{\partial \mu}\nabla^2\Psi + 4\mu^2\frac{\partial}{\partial \mu}\left\{\frac{\partial}{\partial p}\left(\frac{p}{\sigma}\frac{\partial \Psi}{\partial p}\right)\right\} + 2 + \text{Boundary Contributions} = Q_y \tag{3}$$

where μ (= sine of latitude), p the vertical coordinate pressure and static stability, and σ can vary in the vertical. Boundary Contributions refers to contributions to Q_y from the bottom and top boundaries. In Cartesian geometry a change in sign of the mean flow potential vorticity gradient

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

is necessary to have unstable waves (Charney and Stern, 1962); one might reasonably extend this classical analysis to the corresponding Q_y in spherical geometry. Simplifying the schematic circulation to be described by flow in two tropospheric layers, Baines and Frederiksen (1978) extend the Phillips (1954, Cartesian geometry) condition for stability to spherical geometry as:

$$U_{\mathrm{U}} - U_{\mathrm{L}} \geq \frac{622.5\ \cos(\varphi)}{a\Omega\ \sin^2(\varphi)}(\Theta_{\mathrm{U}} - \Theta_{\mathrm{L}}) \tag{4}$$

where the subscripts refer to upper and lower layers. U is the zonal mean flow, while Θ is the horizontal mean potential temperature. It is clear that this criterion decreases nonlinearly with latitude. Static stability helps control the latitude at which the instability criterion is exceeded in two ways. The greater the static stability then: (i) more vertical shear is needed for instability but (ii) the vertical difference of DSE is larger. The heat transport is more efficient (smaller meridional motion is needed to transport the same amount of heat). But smaller meridional motion reduces the strength of the U_{amc} (by more time for friction to act).

If zonal mean flow has vertical shear from meridional motion that conserves angular momentum (zonal wind equals U_{amc}) at upper levels, but is strongly damped by friction at low levels, then one might approximate the shear as some fraction of U_{amc}. For illustrative purposes, if the shear is 70% of U_{amc} and the potential temperature difference from 750 to 250 hPa is 30 K (roughly the observed value in the subtropics) then the instability criterion is met for latitudes higher than 32°.

The amplifying waves feed upon the vertical shear and thereby provide a mechanism to control the unrealistically large zonal velocities created by a purely angular momentum conserving flow. The resultant waves will be tilted against the vertical shear. That can occur if the temperature and mass fields are offset, and if so, there will be a net heat flux across the mean wind, i.e. poleward. The needed poleward heat transport is accomplished by mixing cold air equatorward and warm air poleward. The motion around the eddies can satisfy mass balance and it is also possible to balance surface westerly and easterly torques. The poleward heat fluxes are accompanied by vertical motions across isentropes; such that warm air rises to a higher potential temperature surface. Since such warm air is expected to have moisture, the amplifying waves will also have areas where condensation and precipitation are likely to occur, thus providing a new mechanism for the midlatitude precipitation maximum. Surface friction further leads to low level convergence in the troughs of such waves, further amplifying the midlatitude precipitation maximum.

This idea seems to work well in middle and high latitudes. The crudely estimated stability criterion suggests that such baroclinic waves are not likely to be the main factor in the tropical latitudes. That result, coupled with the ability of the Hadley cell to explain tropical features suggests that the tilted eddy and Hadley cell ideas be applied in different latitude bands with some sort of matching in between.

3.6. Idea #6: hybrid circulation

The preceding discussion suggests combining ideas #3 and #5. The challenge is to sort out how the two quite different ideas can be matched. This idea is more difficult to draw than the other candidates since the middle and high latitude baroclinic eddies are traveling and they change their size and other properties as they move. Hence, Fig. 2d is not so much trajectores, or even a snapshot of the motion but is a 'collage' of some aspects of the circulation.

Beginning with the Hadley circulation, sinking is predicted at the poleward end of the Hadley cell. The analysis in idea #5 suggests that the midlatitude eddies become unstable and dominant at latitudes higher than the subtropics (say, 35° latitude). This will define our conceptual boundary

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

between Hadley cell and baroclinic eddy dominated circulations; for compactness the HCBE boundary. It is no coincidence that this HCBE boundary is not far from the observed location of the surface subtropical highs in different seasons. The subtropical highs fit nicely with the Hadley circulation in several ways: winds on the equatorward side of those highs are consistent with the Hadley circulation, from a boundary layer friction argument those highs would have surface divergence that can support the sinking by the Hadley circulation (if some mechanism, such as diabatic cooling allows air parcels to sink to lower potential temperatures).

The same boundary layer friction argument fosters rising motion for surface lows so some mechanism is needed to separate the baroclinic lows from the edge of the Hadley cell. Furthermore, the baroclinic eddies of higher latitudes extract energy from the zonal flow as well as mix it vertically, both processes weaken the upper level flow. This argument suggests that the HCBE boundary is also likely to be a region where the zonal flow is a maximum and that could be a prototype for the subtropical jet stream.

The Hadley cell may extend a bit further poleward in one region due to such things as topographic deflection of the flow, interaction with higher latitude eddies, and land–sea contrasts. Angular momentum conservation implies a stronger jet (and presumably greater vertical shear) in that extended portion of the Hadley cell. Enhanced tropical convection (Sardeshmukh and Hoskins, 1988) in one region (such as a landmass) may also enhance the subtropical jet in a region. It is a bit simplistic, but one might argue that such a region would favor the baroclinic eddy growth compared with a region where the Hadley cell did not extend as far. If one allows this simplistic argument, then a remarkable chain of events unfolds. The eddies take time to grow, and they tend to move with that zonal flow when their amplitude is small. This stage of growth is termed “linear” in the sense that the eddy wind speeds are “small” compared to the mean flow. The linear stage is also the time when the stability criteria mentioned above have any relevance. While the eddy may grow by means that alter its shape (primarily the amount of upstream tilt) during this stage, the orientation and motion of the eddy along the wind flow maximum is sustained. When the eddy grows to such amplitude that its winds are comparable to the mean flow, then nonlinear effects become dominant. Nonlinear effects include such things as: enhanced upper level amplitude, momentum fluxes, and less efficient poleward heat transport (though still large). Of most interest to the problem of matching circulations is that the baroclinic highs and lows tend to migrate to the side of the jet-like flow having the same sign of vorticity. The lows migrate to the poleward side of the jet, while the highs migrate to the equatorward side. The baroclinic highs can thereby reinforce the subtropical highs while the lows carry precipitation to the middle latitudes where precipitation has its secondary maxima.

The primary function of the circulation is to transport heat poleward and that also needs to be matched. Developing baroclinic lows and highs at the poleward edge of the Hadley circulation help to carry further the poleward heat transport by the Hadley cell. Baroclinic highs that migrate equatorward to reinforce the subtropical highs also assist the heat transport. Lagging behind (upstream) the surface baroclinic highs are upper level troughs which also assist in picking up the heat transport by the Hadley cell and carrying it to higher latitudes.

4. Summary and conclusions

The principal balances and fundamental relationships were described for the large scale atmospheric circulation. Some implications of these properties were discussed. For example, a qualitative discussion illustrated how a mountain torque could balance a uniform westerly wind, though the resultant pressure distribution is opposite to that observed in subtropical latitudes.

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002

From this foundation a variety of candidate circulations were proposed and tested against the six principal balances and observed facts.

A hybrid circulation combining a lower latitude "Hadley" cell paired with middle and higher latitude baroclinic eddies can satisfy the criteria specified. This circulation looks a lot like the observed circulation, which is no surprise since our proposed circulations were strongly influenced by foreknowledge of the actual circulation, in particular knowledge of the primary maxima and minima of precipitation.

There are several points to draw from the discussion. First, one is reminded that satisfying known balances places several constraints upon what circulation is workable. Second, that one can avoid any mention of a Ferrel cell circulation, instead placing appropriate emphasis upon baroclinic eddies dominating the midlatitude circulation. Third, that understanding the general circulation requires looking at fields other than the wind field. Finally, there are many details that have been glossed over in this discussion. Such "details" can be subtle and challenging and constitute important subjects of the other articles in this special issue.

Acknowledgements

The author thanks Dr. P. Baines for some comments about instability on a sphere. The author appreciates helpful anonymous reviewer comments. ERA-15 data were prepared by the European Centre for Medium-range Weather Forecasts. The data were obtained from the NOAA-CIRES Climate Diagnostics Center, Boulder CO, USA using their website at http://www.cdc.noaa.gov/. This work was supported by NSF grant ATM0354545.

References

Baines, P., 1976. The stability of planetary waves on a sphere. J. Fluid Mech. 73 (2), 193–213.
Baines, P., Frederiksen, J., 1978. Baroclinic instability on a sphere in two-layer models. Q. J. R. Met. Soc. 104, 45–68.
Charney, J., Stern, M., 1962. On the stability of internal baroclinic jets in a rotating atmosphere. J. Atmos. Sci. 19, 159–172.
Frederiksen, C., Frederiksen, J., 1990. Flow over topography and instability on a sp[illegible]. Atmos. Sci. 48, 2411–2425.
Grotjahn, R., 1993. Global Atmospheric Circulations: Observations and Theory. Oxford, New York, 430 pp.
Grotjahn, R., Castello, C., 2002. Nonmodal growth on a sphere at various horizontal scales. Geophys. Astrophys. Fluid Dyn. 96, 223–238.
Gustafson, W., Weare, B., 2004. MM5 modeling of the Madden-Julian oscillation in the Indian and west Pacific oceans: model description and control run results. J. Climate 17, 1320–1337.
Mak, M., 1983. On moist quasi-geostrophic barotropic instability. J. Atmos. Sci. 40, 2349–2367.
Peixoto, J., Oort, A., 1992. Physics of Climate. Amer. Int. of Physics, New York, 520 pp.
Phillips, N., 1954. Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. Tellus VI, 273–286.
Sardeshmukh, P., Hoskins, B., 1988. The generation of global rotational flow by steady idealized tropical divergence. J. Atmos. Sci. 45, 1228–1251.

Please cite this article in press as: Grotjahn, R., Deducing the general circulation from basic concepts and a few empirical facts, Dyn. Atmos. Oceans (2006), doi:10.1016/j.dynatmoce.2006.05.002