

Atmospheric (Climate) Dynamics

Fred Kucharski

This course consists of 15 lectures.

- Each lecture is 90 minutes long.
- At the end of each lesson, exercises will be given as homework and discussed in the beginning of the following lesson.
- Recommended textbooks:
 - James R. Holton*: Dynamic Meteorology, Third edition, Academic Press.
 - Joseph Pedlosky*: Geophysical Fluid Dynamics, Springer-Verlag.Others are suggested in the individual lectures. Many others are good as well, so choose!
- Lecture notes will be available at http://users.ictp.it/~kucharsk/lecture_notes_AD_section1.pdf, etc.
- If you find mistakes, corrections are highly appreciated!

Topics in the course

- Vorticity equation for synoptic-scale motion; potential vorticity conservation (barotropic and general) [1.5 h]
- Quasi-geostrophic motion; Thermo-Hydrodynamic equations in pressure coordinates [1.5 h]
- Rossby waves; free Rossby waves; forced Rossby waves; turning latitude [1.5 h]

- Baroclinic instability; two-layer model [1.5 h]
- Equatorial waves; Rossby-gravity waves; Kelvin waves [1.5 h]
- ENSO atmosphere and ocean feedback mechanisms; Gill model; Reduced Gravity Model [1.5h]
- Boundary Layer Processes; turbulent fluxes; Ekman pumping [1.5 h]
- Rainfall responses to heating; Ekman pumping effect; upper-level divergence [1.5h]
- The General Circulation; Hadley Cell; Ferrell Cell [1.5 h]
- Tropical zonal and meridional circulations; Walker circulation; Sverdrup balance [1.5h]
- Energetics of the General Circulation; Lorenz' energy cycle [1.5 h]
- Vertically integrated moisture balance; precipitation; evaporation; moisture flux convergence; application to Sahel drought; Charney feedback mechanism [1.5 h]
- Analysis of climate Variability; EOF analysis, PCA analysis [3 h]
- Modes of variability in the climate system: ENSO, PDO, NAO, AMO [1.5]
- Predictability, Lorenz' Model; measures of predictability [3 h]

1 Vorticity equation for synoptic-scale motion

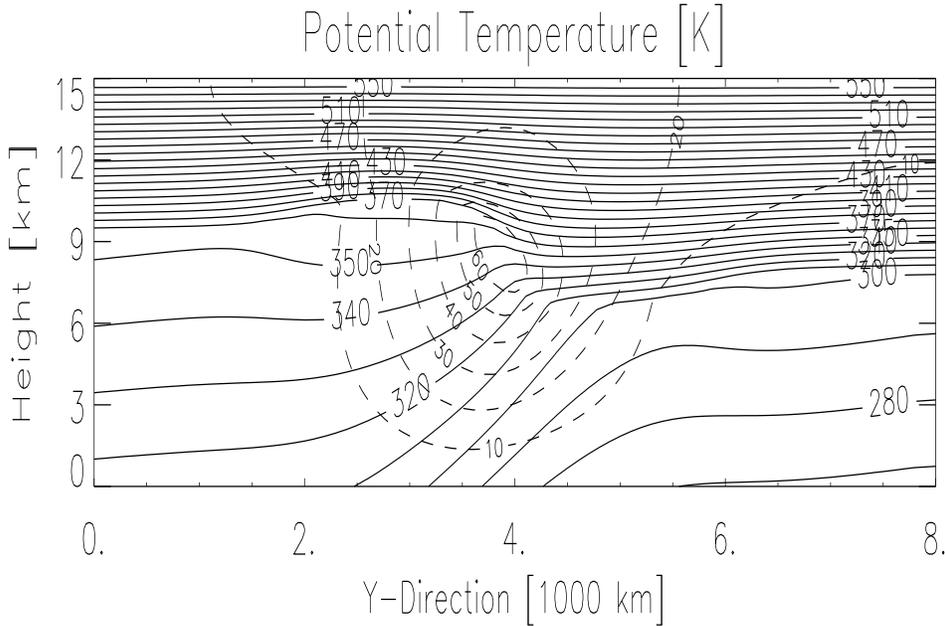


Figure 1: Idealised situation (meridional-vertical section) of the extratropical mean state. Potential temperature (solid lines, K) and zonal wind (dashed, m/s). As we will show later in this Climate Dynamics course the wind approximately fulfill the thermal wind equation $\partial u_g / \partial z \approx -g / (fT) \partial T / \partial y$.

Vorticity is an important concept for the analysis of all kind of atmospheric motions, but in particular for large-scale atmospheric motions. We use the approximate horizontal equations of motion (in the vertical the equation of motion degenerates to the hydrostatic equation) on a sphere, but neglecting all metric terms that occur in the total derivative. Furthermore, we use the abbreviations $dx = r \cos \phi d\lambda, dy = r d\phi, dz = dr$). Also recall the definition of the Coriolis parameter $f \equiv 2\Omega \sin \phi$, and note that we have already neglected the small term proportional to the vertical velocity Coriolis term.

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - fv = -\frac{1}{\rho} \frac{\partial p}{\partial x} \quad (1)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + fu = -\frac{1}{\rho} \frac{\partial p}{\partial y} \quad (2)$$

If we apply $\partial/\partial y$ to Eq. 1 and $\partial/\partial x$ to Eq. 2 and subtract the first from the second, we obtain using the definition $\xi = \partial v/\partial x - \partial u/\partial y$ (Exercise!)

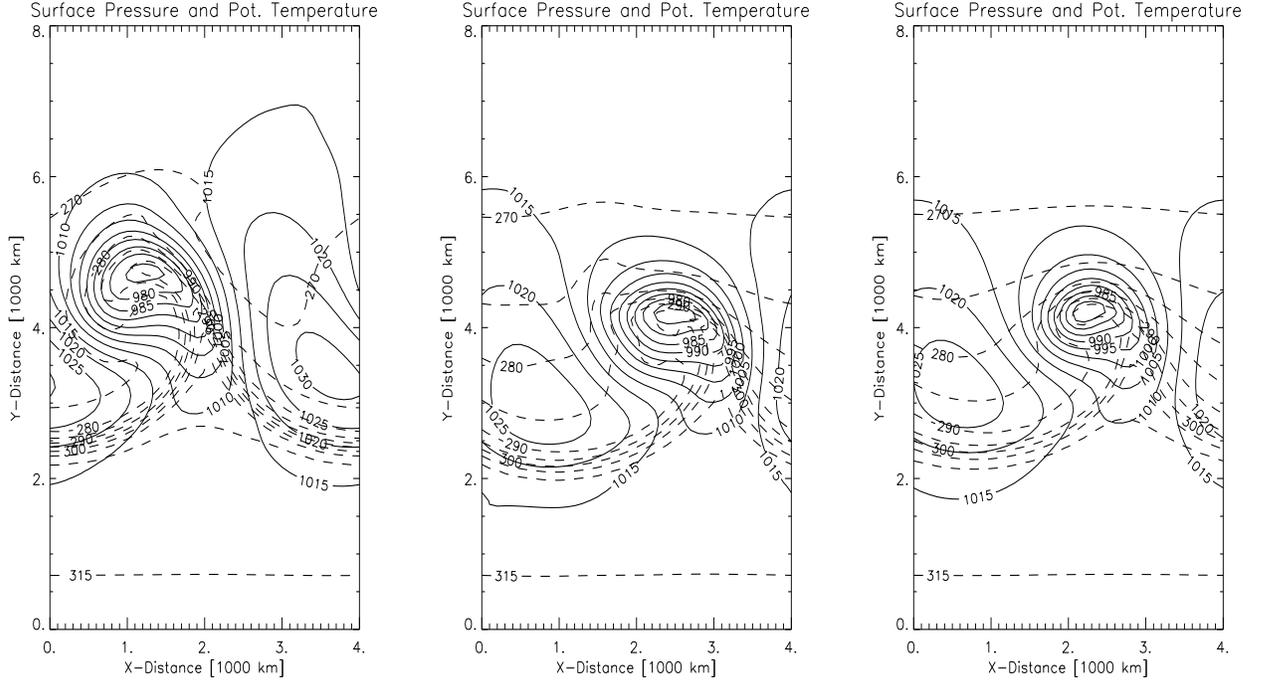


Figure 2: Typical surface pressure [hPa] and potential temperature [K] distributions in extratropical cyclones.

$$\begin{aligned} \frac{\partial \xi}{\partial t} + u \frac{\partial \xi}{\partial x} + v \frac{\partial \xi}{\partial y} + w \frac{\partial \xi}{\partial z} + (\xi + f) \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \\ + \left(\frac{\partial w}{\partial x} \frac{\partial v}{\partial z} - \frac{\partial w}{\partial y} \frac{\partial u}{\partial z} \right) + v \frac{df}{dy} = \frac{1}{\rho^2} \left(\frac{\partial \rho}{\partial x} \frac{\partial p}{\partial y} - \frac{\partial \rho}{\partial y} \frac{\partial p}{\partial x} \right) \end{aligned} \quad (3)$$

The coriolis parameter only depends on y , so we may write:

$$\begin{aligned} \frac{d}{dt}(\xi + f) = - (\xi + f) \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \\ - \left(\frac{\partial w}{\partial x} \frac{\partial v}{\partial z} - \frac{\partial w}{\partial y} \frac{\partial u}{\partial z} \right) + \frac{1}{\rho^2} \left(\frac{\partial \rho}{\partial x} \frac{\partial p}{\partial y} - \frac{\partial \rho}{\partial y} \frac{\partial p}{\partial x} \right). \end{aligned} \quad (4)$$

This is the vorticity equation usually used to analyse synoptic-scale motions (the direction is perpendicular to the earth's surface). It states that the rate of change of absolute vorticity following the motion is given by the sum of three terms, called the divergence term, the tilting or twisting term, and the solenoidal term, respectively.

The first term on the right-hand side may be interpreted as an expression of angular momentum conservation. Imagine an ice skater who rotates and while rotating moves his arms closer to his body: His rotation accelerates (see Fig. 3). But since we are dealing with large scales, in Eq. 4 the absolute vorticity, $\eta = \xi + f$, has to be considered. The interpretation of the second term is that vertical vorticity may be generated by the tilting of horizontal vorticity components by a non-uniform vertical motion field. The meaning of the third term is the solenoidal term. It can be expressed as (exercise!):

$$\frac{1}{\rho^2} \left(\frac{\partial \rho}{\partial x} \frac{\partial p}{\partial y} - \frac{\partial \rho}{\partial y} \frac{\partial p}{\partial x} \right) = \frac{1}{\rho^2} (\nabla \rho \times \nabla p) \cdot \mathbf{k}, \quad (5)$$

where \mathbf{k} is the unit vector in vertical direction. In order to create vorticity through the solenoidal term, lines of constant density have to intersect with lines of constant pressure. A land-sea breeze is a typical example where vorticity is created in such a way.

1.1 Scale analysis of the vorticity equation

In order to understand which terms and therefore mechanisms are the dominant ones in the vorticity equation (4) in this section a 'scale-analysis' will be performed. The scale analysis uses the dimensions of the synoptic scales we are interested in, but also observed magnitudes of flow velocities and other quantities. It is not a rigorous procedure (you use part of the answer as input), but it helps to identify dominant mechanisms.

The scales are given in the following table:

Table 1: Scale parameters for synoptic-scale flow.

$U \sim 10 \text{ m s}^{-1}$	horizontal velocity scale
$W \sim 1 \text{ cm s}^{-1}$	vertical velocity scale
$L \sim 10^6 \text{ m}$	length scale
$H \sim 10^4 \text{ m}$	vertical scale
$\delta p \sim 10 \text{ hPa}$	horizontal pressure scale
$\rho \sim 1 \text{ kg m}^{-3}$	mean density
$\delta\rho/\rho \sim 10^{-2}$	fractional density fluctuation
$L/U \sim 10^5 \text{ s}$	time scale
$f_0 \sim 10^{-4} \text{ s}^{-1}$	Coriolis parameter
$\beta = df/dy \sim 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$	'beta' parameter

This gives

$$\xi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \sim \frac{U}{L} \sim 10^{-5} \text{ s}^{-1}, \quad (6)$$

and

$$\xi/f_0 \sim U/(f_0 L) \equiv Ro \sim 10^{-1}, \quad (7)$$

the ratio of relative to planetary vorticity is equal to the *Rossby number*, which is small for synoptic flow. Therefore, ξ may be neglected compared to f in the divergence term in the vorticity equation

$$(\xi + f) \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \approx f \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \quad (8)$$

The magnitudes of the various terms in equation 4 can now be estimated as follows:

$$\frac{\partial \xi}{\partial t}, u \frac{\partial \xi}{\partial x}, v \frac{\partial \xi}{\partial y} \sim \frac{U^2}{L^2} \sim 10^{-10} \text{ s}^{-2}$$

$$\begin{aligned}
w \frac{\partial \xi}{\partial z} &\sim \frac{WU}{HL} \sim 10^{-11} s^{-2} \\
v \frac{df}{dy} &\sim U\beta \sim 10^{-10} s^{-2} \\
f \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) &\sim \frac{f_0 U}{L} \sim 10^{-9} s^{-2} \\
\left(\frac{\partial w}{\partial x} \frac{\partial v}{\partial z} - \frac{\partial w}{\partial y} \frac{\partial u}{\partial z} \right) &\sim \frac{WU}{HL} \sim 10^{-11} s^{-2} \\
\frac{1}{\rho^2} \left(\frac{\partial \rho}{\partial x} \frac{\partial p}{\partial y} - \frac{\partial \rho}{\partial y} \frac{\partial p}{\partial x} \right) &\sim \frac{\delta \rho \delta p}{\rho^2 L^2} \sim 10^{-11} s^{-2}
\end{aligned}$$

The estimation of the divergence is an overestimation. Indeed it will be shown later in this course the the divergent part of the flow (which is also the non-geostrophic part) is an order of magnitude smaller that the rotational part (geostrophic). We have therefore

$$\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \sim 10^{-6} s^{-1} \quad , \quad (9)$$

which means that the divergence is typically one order of magnitude smaller than the vorticity of synoptic-scale motion. Therefore,

$$f \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \sim 10^{-10} s^{-2} \quad . \quad (10)$$

Therefore in the vorticity equation (4), we have the first order balance

$$\frac{d_h(\xi + f)}{dt} = -f \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \quad , \quad (11)$$

where

$$\frac{d_h}{dt} \equiv \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \quad . \quad (12)$$

The synoptic-scale vorticity equation 11 states that the rate of change of absolute vorticity following the horizontal motion is approximately given by the generation (destruction) of vorticity owing to horizontal convergence (divergence; see sketch 3). Indeed, this is considered to be the main mechanism of cyclone (and anticyclone) developments, connecting cyclonic motion to low pressure and anticyclonic motion to high pressure (only for large-scale motions!).

1.2 The Barotropic (Rossby) Potential Vorticity Equation

As barotropic model of the Atmosphere we assume that there is incompressibility and the flow may be confined by the height of two given boundaries, $h(x, y, z, t) = H_t - H_b$ (see also lecture on equatorial waves). The incompressibility condition may be expressed as

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = - \frac{\partial w}{\partial z} \quad . \quad (13)$$

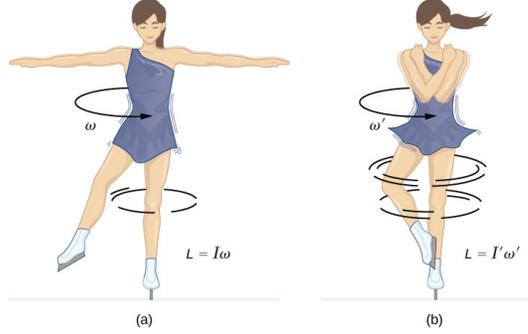


Figure 3: Sketch of ice scater closing in arms. From: courses.lumenlearning.com

We also assume that the horizontal velocities are independent of height. With this we can integrate the vorticity equation (11) vertically to obtain

$$h \frac{d_h(\xi + f)}{dt} = (f + \xi)[w(H_t) - w(H_b)] \quad . \quad (14)$$

Note that in equation (14), the relative vorticity ξ may be replaced, to a first approximation, by the geostrophic relative vorticity

$$\xi \approx \xi_g \equiv \nabla^2 gh / f_0 = \nabla^2 \Phi / f_0 \quad , \quad (15)$$

where we have assumed that the meridional scale, L , is small compared to the radius of the earth so that the geostrophic wind may be defined using a constant reference latitude of the Coriolis parameter $f \approx f_0 \equiv 2\Omega \sin \phi_0$. Also, in the operator (12) the horizontal velocities may be approximated by the geostrophic ones

$$\mathbf{v} \approx \mathbf{v}_g \equiv f_0^{-1} \mathbf{k} \times \nabla gh = f_0^{-1} \mathbf{k} \times \nabla \Phi \quad , \quad (16)$$

Equations 15 and 16 can be derived from the geostrophic equations and integration of the hydrostatic equation for an incompressible fluid (exercise!). Note that, for beauty, we have re-introduced the small ξ effect on the rhs of Eq. (14). Since $w(H_t) = dH_t/dt, w(H_b) = dH_b/dt$ we have,

$$\frac{1}{\xi + f} \frac{d_h(\xi + f)}{dt} = \frac{1}{h} \frac{d_h h}{dt} \quad . \quad (17)$$

Integrating left and right side leads to

$$\frac{d_h}{dt} [\ln(\xi + f)] = \frac{d_h}{dt} [\ln h] \quad , \quad (18)$$

which implies that

$$\frac{d_h}{dt} \left(\frac{\xi + f}{h} \right) = 0 \quad , \quad (19)$$

which is the potential vorticity conservation theorem for a barotropic fluid, first obtained by C. G. Rossby. The quantity conserved following the horizontal motion is

the barotropic potential vorticity. It explains nicely some features of the observed stationary waves, e.g. induced by the Rocky mountains. If the flow is purely horizontal, i.e. rigid lid and lower boundary, then we obtain the *barotropic vorticity equation*

$$\frac{d_h(\xi + f)}{dt} = 0 \quad , \quad (20)$$

which states that the absolute vorticity is conserved following the horizontal motion. The flow in the mid-troposphere approximately fulfils this condition and equation (20) may be used to explain the movement of air particles in Rossby waves!

Note that using the approximations (15) and (16) the barotropic vorticity equation (20) can be re-written in terms of the streamfunction $\psi \equiv \Phi/f_0 = gh/f_0$ or equivalently also in terms of geopotential $\Phi = gh$ (exercise!)

$$\frac{d_h}{dt} \nabla^2 \psi + \beta \frac{\partial \psi}{\partial x} = 0 \quad . \quad (21)$$

Also the operator d_h/dt can be expressed in terms of streamfunction

$$\frac{d_h}{dt} = \frac{\partial}{\partial t} - \frac{\partial \psi}{\partial y} \frac{\partial}{\partial x} + \frac{\partial \psi}{\partial x} \frac{\partial}{\partial y} \quad (22)$$

Equations (21) and (22) can be used conveniently to compute Rossby wave solutions numerically (exercise in section 3!)

1.3 The exact potential vorticity conservation law; Ertel's potential vorticity

The barotropic potential vorticity conservation law 19 for is a very instructive special case (incompressible, barotropic fluid) of a much more general conservation law. Ertel (1942, Meteorologische Zeitung, **59**, 271-281) was the first to derive the law in the most general form. In order to derive it, we start from a general form of the equations of motion

$$\rho \frac{d\mathbf{v}}{dt} = -\nabla p - 2\rho\boldsymbol{\Omega} \times \mathbf{v} - \rho\nabla\phi - \nabla \cdot \mathbf{F} \quad , \quad (23)$$

where \mathbf{F} is the frictional tensor, $\boldsymbol{\Omega}$ is the (constant) rotation vector of the earth. From Eq. 23 we may derive (exercise!) the full 3-dimensional vorticity equation, which is a generalization of Eq. 4

$$\frac{d\mathbf{vort}_{\mathbf{a}}}{dt} = \mathbf{vort}_{\mathbf{a}} \cdot \nabla \mathbf{v} - \mathbf{vort}_{\mathbf{a}} \nabla \cdot \mathbf{v} + \frac{\nabla \rho \times \nabla p}{\rho^2} - \nabla \times \frac{\nabla \cdot \mathbf{F}}{\rho} \quad , \quad (24)$$

where $\mathbf{vort}_{\mathbf{a}} = 2\boldsymbol{\Omega} + \xi$ is the 3-dimensional absolute vorticity vector. Note, that whereas $\boldsymbol{\Omega}$ is constant, its components following the earth surface are not! Using the continuity equation, this may be re-written as (exercise!)

$$\frac{d}{dt} \frac{\mathbf{vort}_{\mathbf{a}}}{\rho} = \frac{\mathbf{vort}_{\mathbf{a}}}{\rho} \cdot \nabla \mathbf{v} + \frac{\nabla \rho \times \nabla p}{\rho^3} - \frac{1}{\rho} \nabla \times \frac{\nabla \cdot \mathbf{F}}{\rho} \quad . \quad (25)$$

Let us consider a scalar quantity, λ , which is a function of just ρ and p . For a one component system, any thermodynamic quantity could be used, but in the end it turns out that quantities that obey a conservation law are particularly useful. Therefore, the entropy is a very good candidate (or potential temperature in the atmosphere or potential density in the ocean). The equation may then look like

$$\frac{d\lambda}{dt} = \sigma_\lambda \quad . \quad (26)$$

Let us perform the following simple calculation

$$\mathbf{vort}_a \cdot \frac{d}{dt} \nabla \lambda = \mathbf{vort}_a \cdot \nabla \sigma_\lambda - [\mathbf{vort}_a \cdot \nabla \mathbf{v}] \cdot \nabla \lambda \quad . \quad (27)$$

On the other hand, if we multiply as scalar product Eq. 25 by $\nabla \lambda$, we get

$$\nabla \lambda \cdot \frac{d}{dt} \frac{\mathbf{vort}_a}{\rho} = \left[\frac{\mathbf{vort}_a}{\rho} \cdot \nabla \mathbf{v} \right] \cdot \nabla \lambda + \nabla \lambda \cdot \frac{\nabla \rho \times \nabla p}{\rho^3} - \frac{1}{\rho} \nabla \lambda \cdot \nabla \times \frac{\nabla \cdot \mathbf{F}}{\rho} \quad . \quad (28)$$

Combining Eqs. 27 and 28 leads to

$$\frac{d}{dt} \left[\frac{\nabla \lambda \cdot \mathbf{vort}_a}{\rho} \right] = \frac{1}{\rho^3} \nabla \lambda \cdot (\nabla \rho \times \nabla p) + \frac{\mathbf{vort}_a}{\rho} \cdot \nabla \sigma_\lambda - \frac{1}{\rho} \nabla \lambda \cdot \nabla \times \frac{\nabla \cdot \mathbf{F}}{\rho} \quad . \quad (29)$$

The quantity on the lhs of Eq. 29 is conserved following the motion if (why?, Exercise):

- The fluid is barotropic ($\nabla \rho \times \nabla p = 0$) or λ is a thermodynamic function of p and ρ , i.e. $\lambda = \lambda(p, \rho)$,
- the quantity λ is itself conserved following the motion, i.e. $\sigma_\lambda = 0$,
- the flow is frictionless ($\mathbf{F} = 0$).

In case the above criterion are fulfilled, we may call the quantity

$$q = \left[\frac{\nabla \lambda \cdot \mathbf{vort}_a}{\rho} \right] \quad (30)$$

potential vorticity. So far, apart from the momentum budget, no physical constrain is on the variable λ , which may be chosen as suitable. A good choice is the entropy s , which for adiabatic-reversible processes is a constant following the motion. For atmospheric purposes it is convenient to choose alternatively the potential temperature θ . Furthermore, for large-scale atmospheric (and ocean) dynamics it is generally a good approximation that the vorticity is dominated by its vertical component (note that this approximation may break down if there are strong horizontal potential temperature gradients). In this case the potential vorticity can be expressed as

$$q = \left[\frac{\eta}{\rho} \frac{\partial \theta}{\partial z} \right] \quad , \quad (31)$$

where $\eta = \xi + f$ is the vertical component of the absolute vorticity. The form 31 is most conveniently explored in the pressure coordinate system, which will be introduced in the next section. Note that a convenient and exact version of the potential vorticity can be derived using a coordinate system with the potential temperature as vertical coordinate.

The barotropic form of the potential vorticity Eq. 19 may be derived by considering the quantity

$$\lambda = \frac{z - H_b}{h} \quad , \quad (32)$$

which measures the relative height of a parcel with respect to the total height of the fluid and turns out to be conserved for barotropic flow with constant density.

Exercises

1. Show that applying $\partial/\partial y$ to Eq. 1 and $\partial/\partial x$ to Eq. 2 leads to Eq. 3.
2. Show the validity of Eq. 5.
3. Show that the geostrophic formulations 15 and 16 can be derived from the usual geostrophic wind

$$\mathbf{v}_g \equiv \frac{1}{\rho f_0} \mathbf{k} \times \nabla p$$

and that with this the barotropic vorticity equation 20 can be written as Eq. 21 with 22.

4. An air parcel at 30 N moves northward conserving absolute vorticity. If its initial relative vorticity is $5 \times 10^{-5} \text{ s}^{-1}$, what is its relative vorticity upon reaching 90 N?
5. An air column at 60 N with $\xi = 0$ initially stretches from the surface to a fixed tropopause at 10 km. If the air column moves until it is over a mountain barrier 2.5 km high at 45 N, what are its absolute vorticity and relative vorticity as it passes the mountain top, assuming that the flow satisfies the barotropic potential vorticity equation?
6. Derive Eq. 24 from 23 and Eq. 25 from 24.
7. Discuss under which conditions 30 is conserved following the motion.

2 Quasi-geostrophic motion

2.1 The basic equations in isobaric Coordinates

The basic governing equations are (see Eqs. 1 and 2)

The horizontal momentum equations

$$\frac{d\mathbf{v}}{dt} + f\mathbf{k} \times \mathbf{v} = -\frac{1}{\rho}\nabla p \quad , \quad (33)$$

where $\mathbf{v} = \mathbf{i}u + \mathbf{j}v$ and the nabla operator has just the horizontal components.

The vertical equation of motion degenerates for all large-scale motion (e.g. scales more than 100 km) into the hydrostatic equation (discuss how good this approximation is):

$$\frac{\partial p}{\partial z} = -\rho g \quad . \quad (34)$$

Equation (34) states that there is a monotonic relation between pressure p and height z , which leads to the possibility of using p as a vertical coordinate. The basic equation for deriving all transformation from the height to the pressure coordinate system is: $\psi(x, y, p, t) = \psi(x, y, z, t)$, which leads, for example to

$$\frac{\partial \psi}{\partial x} \Big|_p = \frac{\partial \psi}{\partial x} \Big|_z + \frac{\partial \psi}{\partial z} \frac{\partial z}{\partial x} \Big|_p \quad . \quad (35)$$

Inserting $\psi = p$ and applying Eq. (35) also to the derivative in y direction gives the transformation for the horizontal pressure gradient force $\nabla_z p = \rho g \nabla_p z = \rho \nabla_p \Phi$. Thus the horizontal momentum equation reads

$$\frac{d\mathbf{v}}{dt} + f\mathbf{k} \times \mathbf{v} = -\nabla_p \Phi \quad . \quad (36)$$

This looks a little like the horizontal momentum equation of the shallow water model, but it is not! Applying $\psi(x, y, p, t) = \psi(x, y, z, t)$ to a vertical derivative and letting $\psi = p$ gives the hydrostatic equation in pressure coordinates

$$\frac{\partial \Phi}{\partial p} = -\frac{1}{\rho} = -\alpha = -\frac{RT}{p} \quad . \quad (37)$$

The total derivative d/dt is invariant and can be expressed as (as follows directly from $\psi = \psi(x, y, p, t)$)

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \frac{dx}{dt} + \frac{\partial}{\partial y} \frac{dy}{dt} + \frac{\partial}{\partial p} \frac{dp}{dt} \quad (38)$$

$$= \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + \omega \frac{\partial}{\partial p} \quad (39)$$

$$= \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_p + \omega \frac{\partial}{\partial p} \quad . \quad (40)$$

$\omega = dp/dt$ (called the 'omega' vertical velocity) is the pressure change following the motion. Note that when w is positive ω is typically negative.

The Continuity Equation

The easiest way to derive the continuity equation is through the principle of mass conservation. For an infinitesimal mass element we may write:

$$\delta m = \rho \delta V = \rho \delta x \delta y \delta z = -\delta x \delta y \frac{1}{g} \delta p \quad . \quad (41)$$

Note that the first part of equation (41) is just the definition of the density. In the second part the hydrostatic equation (34) has been used to replace the vertical perturbation by a pressure perturbation. Let's calculate the derivative of (41) following the motion (conservation of mass)

$$\frac{1}{\delta m} \frac{d}{dt} \delta m = \frac{g}{\delta x \delta y \delta p} \frac{d}{dt} \frac{\delta x \delta y \delta p}{g} = 0. \quad (42)$$

After applying the product rule of differentiation, and changing the order of differentiation we obtain

$$\frac{1}{\delta x} \delta \frac{d}{dt} x + \frac{1}{\delta y} \delta \frac{d}{dt} y + \frac{1}{\delta p} \delta \frac{d}{dt} p = \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} + \frac{\delta \omega}{\delta p} = 0 \quad . \quad (43)$$

Letting $\delta x, \delta y, \delta z \rightarrow 0$, it follows the continuity equation in pressure coordinates:

$$\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)_p + \frac{\partial \omega}{\partial p} = 0 \quad . \quad (44)$$

In pressure coordinates the full continuity equation takes the form of that of an incompressible fluid, i.e. the time derivative of density does not occur anymore explicitly.

The Thermodynamic Energy Equation

Recall Eqs. 108 or 110 for the Enthalpy, and after multiplying with T from our Earth System Thermodynamics course (for $ds = 0$), which was approximately valid for the atmosphere in which phase transitions from water vapour to liquid water are allowed (do you remember what the symbols L_{lv} and m_v stand for?):

$$c_p \frac{dT}{dt} - \frac{RT}{p} \frac{dp}{dt} = -L_{lv} \frac{dm_v}{dt} \quad . \quad (45)$$

If we further allow diabatic processes to occur (e.g. radiation), then we can simply add $T ds/dt$ on the rhs and abbreviate those terms as Q .

$$c_p \frac{dT}{dt} - \alpha \frac{dp}{dt} = Q \quad . \quad (46)$$

Q is thus the heat added by diabatic processes (i.e. condensation, radiation).

Using the total derivative in pressure coordinates and the definition of ω we have

$$\left(\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} \right)_p - S_p \omega = \frac{Q}{c_p} \quad , \quad (47)$$

where the stability factor

$$S_p = \frac{RT}{c_p p} - \frac{\partial T}{\partial p} = -\frac{T}{\theta} \frac{\partial \theta}{\partial p} \quad (48)$$

has been introduced. In Eq. (48) we have used the definition of the potential temperature (Exercise!)

$$\theta = T \left(\frac{p_0}{p} \right)^{\frac{R}{c_p}} \quad . \quad (49)$$

p_0 is a constant reference pressure here. Using the dry adiabatic lapse rate $\Gamma_d = g/c_p$, we have also (exercise!)

$$S_p = (\Gamma_d - \Gamma)/(\rho g) \quad , \quad (50)$$

where the definition of the lapse rate $-dT/dz = \Gamma$ has been used.

The set of equations (36), (37), (44) and (47) is the basis for our analysis of synoptic-scale motion, but also the basis for many numerical models of the atmospheric circulation.

It is also useful to derive the approximate version of the potential vorticity 31 in pressure coordinates, because it takes a more convenient form. We can write (e.g. using $\psi(x, y, p, t) = \psi(x, y, z, t)$) to evaluate a vertical derivative

$$\frac{\partial \theta}{\partial z} = \frac{\partial \theta}{\partial p} \frac{\partial p}{\partial z} = -\rho g \frac{\partial \theta}{\partial p} \quad . \quad (51)$$

With this 31 becomes

$$q_p = \left[\eta_p \frac{\partial \theta}{\partial p} \right] \quad , \quad (52)$$

where the constant factor $-g$ has been excluded from the definition (this does not matter, why?). Note, that the absolute vorticity $\eta_p = \mathbf{k} \cdot \nabla_p \times \mathbf{v} + f$ is the vertical component of the absolute vorticity and has to be evaluated on pressure levels. The physical interpretation of the approximate potential vorticity in pressure coordinates, q_p , is that a fluid element within an isentropic flow may be considered as confined between two potential temperature values $\Delta\theta$. The thickness of the fluid element, Δp , however, may change. If this occurs, then the absolute vorticity has to adjust in order to maintain potential vorticity conservation.

2.2 Some observed features of the extratropical mean flow

A primary goal of dynamic meteorology is to interpret the observed structure of large-scale atmospheric motions in terms of physical laws governing the motions. These laws, which express the conservation of momentum, mass, and energy completely determine the relationships among the pressure, temperature, and velocity fields. However, the pure laws provide an enormously complicated picture of the motions. For extratropical synoptic-scale motions, however, the horizontal velocities are approximately geostrophic. Such motions, which are often referred to as *quasi-geostrophic*, are simpler to analyze than, for example, tropical disturbances. They are also the major systems of interest in traditional short-range weather forecasting and are thus a reasonable starting point for dynamical analysis. In this section we show that for extratropical synoptic-scale systems the twin requirement of hydrostatic and geostrophic balance constrain the baroclinic motions so that to a good approximation the structure and evolution of the three-dimensional velocity field are determined by the distribution of geopotential height on isobaric surfaces. The equations that express these relationships constitute the quasi-geostrophic system. Before developing this system of equations it is useful to summarize briefly the observed structure of mid-latitude synoptic systems and the mean circulations in which they are embedded.

Zonally averaged cross sections do provide some useful information on the gross structure of the planetary-scale circulation, in which synoptic-scale eddies are embedded. Fig. 4 and 5 show the zonal mean meridional-vertical sections of temperature (left) anomaly from zonal mean and zonal velocity (right) for northern (December-to-February; DJF) and southern winter (June-to-August; JJA), respectively. The vertical direction is measured in pressure (hPa). The average pole-to-equator temperature gradient in the Northern Hemisphere troposphere is much larger in winter than in summer. In the southern hemisphere the difference between summer and winter temperature distributions is smaller, owing mainly to the large thermal inertia of oceans together with the greater fraction of the surface that is covered by oceans in the Southern Hemisphere. The zonal flow and the meridional temperature gradients satisfy to a large degree the thermal wind relation (Exercise!), the largest zonal wind speeds are located in upper levels in regions with the largest meridional temperature gradients

$$\frac{\partial u_g}{\partial p} = \frac{R}{f p} \left(\frac{\partial T}{\partial y} \right)_p . \quad (53)$$

The core of maximum zonal wind speed (called *jet stream axis*) is located just below the *tropopause* (the boundary between troposphere and stratosphere). In both hemispheres the location is about 30°-35° during winter and 40°-45° during summer.

However, there are some important deviations from the zonal mean picture. Fig. 6 shows the zonal wind distribution at the 200 hPa level. As can be seen the largest wind speeds are concentrated just off the coast of Asia and North America, where also the largest meridional temperature gradients occur. Also, whereas the

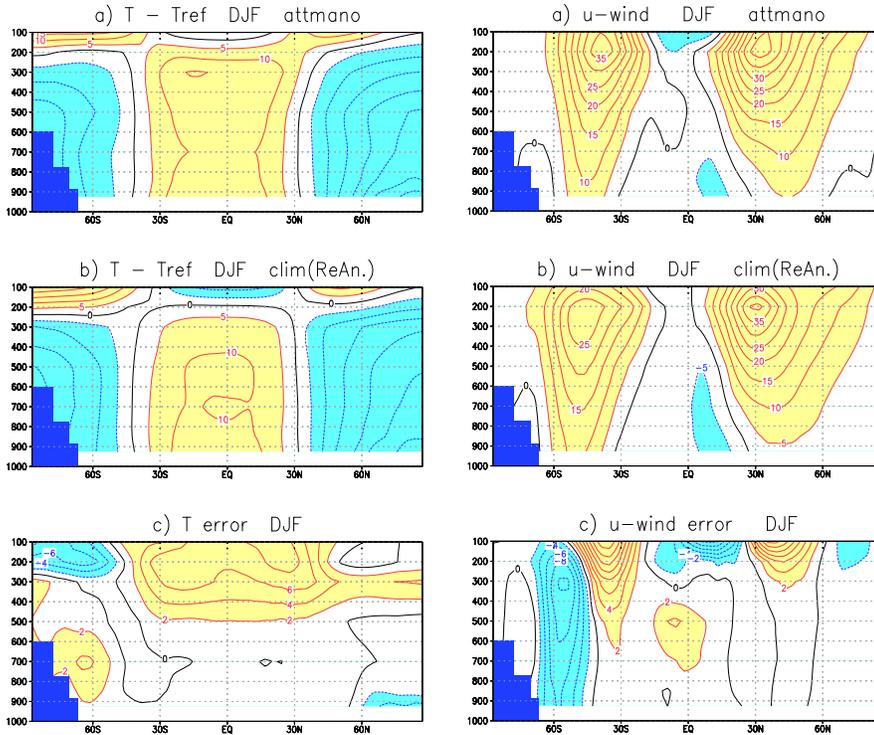


Figure 4: Northern winter meridional-height sections of temperature deviations from zonal mean (left) and zonal wind (right). Units are K for temperature and m/s for wind.

pacific jet is quite zonal, the Atlantic one is clearly tilted from the south-west to the north-east. It is in these regions where most extratropical cyclones and anticyclones develop. I will be shown in section 4 that the mechanisms where these systems draw energy from is the meridional temperature gradient due to an instability called *baroclinic instability*. The systems propagate downstream along the *storm tracks* that approximately follow the jet axis.

The large departure of the northern winter climatological jet stream from zonal symmetry can also be inferred from examination of Fig. 7, which shows the DJF mean 500 hPa geopotential contours (the z from $\Phi = gz$ in Eq. 36). Even after averaging the geopotential height contours for one season, very striking departures from zonal symmetry remain. These are clearly linked to the distribution of continents and Oceans (for example orographic *Rossby waves* due to approximate barotropic potential vorticity conservation [see section 1.3]).

The most prominent asymmetries are the throughs to the east of the American and Asian continents. Referring back to Fig. 6, we see that the intense jet at 35° N and 140° E is a result of the semi-permanent trough in that region (that is the isolines of height show strong gradient in that region).

Thus, it is apparent that the mean flow in which synoptic systems are embedded

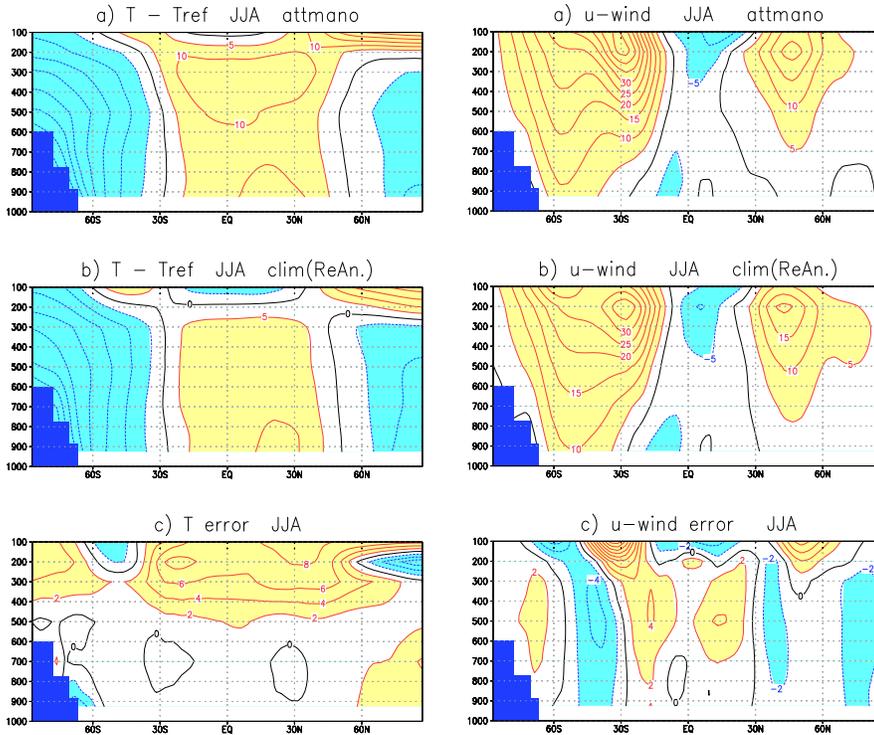


Figure 5: Southern winter meridional-height sections of temperature deviations from zonal mean (left) and zonal wind (right). Units are K for temperature and m/s for wind.

should really be regarded as a longitude-dependent time-averaged flow. In addition to its longitudinal dependence, the planetary-scale flow also varies from day to day owing to its interactions with transient synoptic-scale disturbances.

It is a common observation in fluid dynamics that jets in which strong velocity shears occur may be unstable with respect to small perturbations. By this is meant that any small disturbance introduced into the jets will tend to amplify, drawing energy from the jet as it grows. Most synoptic-scale systems in mid-latitude appear to develop as the result of an instability of the jet-stream flow. This instability, called *baroclinic instability*, depends on the meridional temperature gradient, particularly at the surface. Hence, through the thermal wind relationship, baroclinic instability depends on vertical wind shear.

2.3 The Quasi-geostrophic approximation

The main goal of this chapter is to show how the observed structure of midlatitude systems can be interpreted in terms of the constraints imposed on synoptic-scale motions by the dynamical equations. Specifically we show that for equations that are hydrostatic and nearly geostrophic the three-dimensional flow is determined

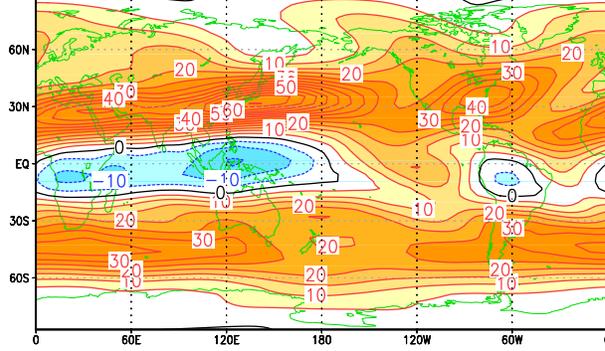


Figure 6: Northern winter (DJF) 200 hPa zonal wind. Units are m/s.

approximately by the isobaric distribution of geopotential height $[\Phi(x, y, p, t)]$. For this analysis, it is convenient to use the isobaric coordinate system both because meteorological measurements are generally referred to constant-pressure surfaces and because the dynamical equations are somewhat simpler in isobaric coordinates than in height coordinates. Thus, use of the isobaric coordinate system simplifies the development of approximate prognostic and diagnostic equations.

2.3.1 Scale Analysis in Isobaric Coordinates

We consider the set of equations (36), (37), (44) and (47). In the following we will drop the notation $(\)_p$ to indicate derivatives at constant pressure, which is valid in this section for all horizontal and time derivatives. The stability parameter is positive [$S_p \approx 5 \times 10^{-4} \text{ K Pa}^{-1}$ in the mid-troposphere]. This set of equations, although simplified by use of the hydrostatic approximation and by neglect of some small terms that appear in the complete spherical coordinate form, still contains terms that are of secondary significance for mid-latitude synoptic-scale systems. They can be further simplified by the observation that the horizontal flow is nearly geostrophic and that the ratio of the magnitudes of vertical to horizontal velocities is of the order of 10^{-3} .

We first separate the horizontal velocity into geostrophic and ageostrophic parts by letting

$$\mathbf{v} = \mathbf{v}_g + \mathbf{v}_a \quad , \quad (54)$$

where the geostrophic wind is defined as

$$\mathbf{v}_g \equiv f_0^{-1} \mathbf{k} \times \nabla \Phi \quad , \quad (55)$$

and \mathbf{v}_a is just the difference between the total horizontal wind and the geostrophic wind. Here we have assumed that the meridional scale, L , is small compared to the radius of the earth so that the geostrophic wind may be defined using a constant reference latitude of the Coriolis parameter ($f \approx f_0$ as in equation 15). Note that the definition (55) implies that the geostrophic wind is non-divergent.

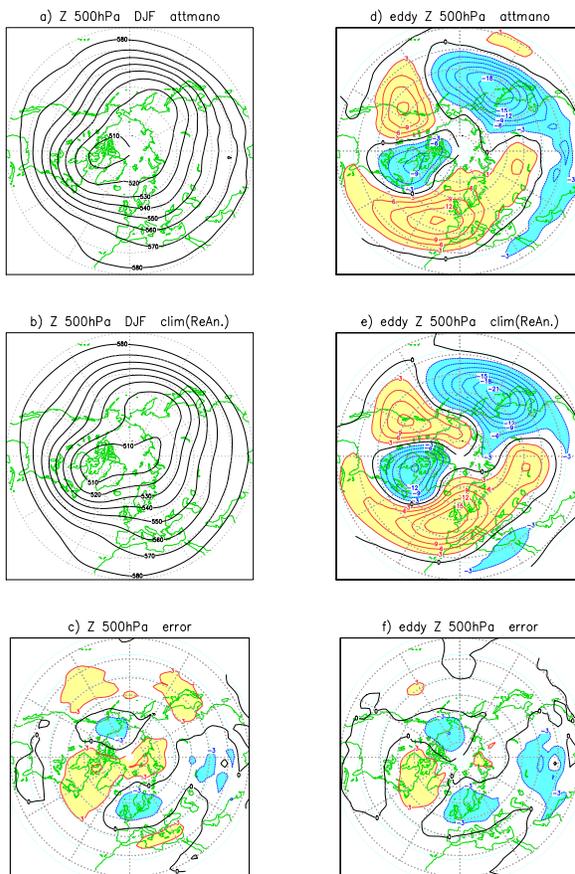


Figure 7: Northern winter (DJF) 500 hPa zonal geopotential height. Units are decametre.

For the systems of interest $|\mathbf{v}_g| \gg |\mathbf{v}_a|$. More precisely,

$$\frac{|\mathbf{v}_a|}{|\mathbf{v}_g|} \sim O(Ro) \approx 10^{-1} . \quad (56)$$

The *Rossby number* Ro has been introduced in Eq. (7).

The momentum can then be approximated to $O(Ro)$ by its geostrophic value, and the rate of change of momentum or temperature following the horizontal motion can be approximated to the same order by the rate of change following the geostrophic wind. Thus, in the total derivative (40), \mathbf{v} can be replaced by \mathbf{v}_g and the vertical advection, which arises from ageostrophic flow, can be neglected. The rate of change of momentum following the total motion is then approximately equal to the rate of change of the geostrophic momentum following the geostrophic wind:

$$\frac{d\mathbf{v}}{dt} \approx \frac{d_g \mathbf{v}_g}{dt} , \quad (57)$$

where

$$\frac{d_g}{dt} = \frac{\partial}{\partial t} + \mathbf{v}_g \cdot \nabla = \frac{\partial}{\partial t} + u_g \frac{\partial}{\partial x} + v_g \frac{\partial}{\partial y} \quad . \quad (58)$$

Note, however, that the vertical advection in the thermodynamic equation, 47, has been combined already with the adiabatic expansion term to provide the stability term $S_p \omega$.

Although a constant f_0 can be used in defining \mathbf{v}_g , it is still necessary to retain the dynamical effect of the variation of the Coriolis parameter with latitude in the Coriolis force term in the momentum equation. This variation can be approximated by expanding the latitudinal dependence of f in a Taylor series about a reference latitude ϕ_0 and retaining only the first two terms to yield

$$f = 2|\Omega| \sin \phi \approx f_0 + \beta y \quad , \quad (59)$$

that is the $\sin \phi$ -dependence is approximated linearly for a given latitude ϕ_0 by a Taylor series expansion (therefore $\beta = 2|\Omega| \cos \phi_0 / a$; a being the mean radius of the earth). This approximation is usually referred to as *mid-latitude beta-plane* approximation. For synoptic-scale motions the ratio of the first two terms in the expression of f has the order of magnitude

$$\frac{\beta L}{f_0} \approx \frac{\cos \phi_0 L}{\sin \phi_0 a} \sim O(Ro) \ll 1 \quad . \quad (60)$$

This justifies letting the coriolis parameter have a constant value f_0 in the geostrophic approximation and approximating its variation in the coriolis force term by (59).

From Eq. (36) the acceleration following the motion is equal to the difference between the Coriolis force and the pressure gradient force. This difference depends on the departure of the actual wind from the geostrophic wind. We can write, using (54), (59) and (55)

$$\begin{aligned} f \mathbf{k} \times \mathbf{v} + \nabla \Phi &= (f_0 + \beta y) \mathbf{k} \times (\mathbf{v}_g + \mathbf{v}_a) - f_0 \mathbf{k} \times \mathbf{v}_g \\ &\approx f_0 \mathbf{k} \times \mathbf{v}_a + \beta y \mathbf{k} \times \mathbf{v}_g \quad . \end{aligned} \quad (61)$$

The approximate horizontal momentum equation thus has the form

$$\frac{d_g \mathbf{v}_g}{dt} = -f_0 \mathbf{k} \times \mathbf{v}_a - \beta y \mathbf{k} \times \mathbf{v}_g \quad . \quad (62)$$

Since the geostrophic wind (55) is non-divergent, the continuity equation (44) may be written as

$$\nabla \cdot \mathbf{v}_a + \frac{\partial \omega}{\partial p} = 0 \quad , \quad (63)$$

which shows that ω is only defined by the ageostrophic part of the wind field (i.e. it is the ageostrophic wind that drives vertical motions that are relevant for energy conversions!!!).

In the thermodynamic energy equation (47) the horizontal advection can be approximated by its geostrophic value. However, as mentioned above, the vertical advection is not neglected, but forms part of the adiabatic heating and cooling term. This term must be retained because the static stability is usually large enough on the synoptic scale that the adiabatic heating or cooling owing to vertical motion is of the same order as the horizontal temperature advection despite the smallness of the vertical velocity. It can be somewhat simplified, though, by dividing the total temperature field T_{tot} , into a basic state (standard atmosphere) portion that depends only on pressure, $T_0(p)$, plus a deviation from the basic state, $T(x, y, p, t)$

$$T_{tot} = T_0(p) + T(x, y, p, t) \quad . \quad (64)$$

Since $|dT_0/dp| \gg |\partial T/\partial p|$ only the basic state portion of the temperature field need to be included in the static stability term and the quasi-geostrophic thermodynamic energy equation may be expressed in the form

$$\frac{\partial T}{\partial t} + \mathbf{v}_g \cdot \nabla T - \left(\frac{\sigma p}{R} \right) \omega = \frac{Q}{c_p} \quad , \quad (65)$$

where $\sigma \equiv -RT_0 p^{-1} d \ln \theta_0 / dp$ and θ_0 is the potential temperature corresponding to a basic state temperature T_0 ($\sigma \approx 2 \times 10^{-6} \text{ m}^2 \text{ Pa}^{-2} \text{ s}^{-2}$ in the midtroposphere).

Equations (62), (55), (37), (63) and (65) constitute the quasi-geostrophic equations. If Q is known these form a complete set in the dependent variables $\Phi, T, \mathbf{v}_g, \mathbf{v}_a$ and ω .

2.4 The Quasi-Geostrophic Vorticity Equation

Just as the horizontal momentum can be approximated to $O(Ro)$ by its geostrophic value, the vertical component of the vorticity can also be approximated geostrophically. Using Eq. (55) the geostrophic vorticity $\xi_g = \mathbf{k} \cdot \nabla \times \mathbf{v}_g$ can be expressed in terms of the Laplacian of the geopotential

$$\xi_g = \frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} = \frac{1}{f_0} \nabla^2 \Phi \quad . \quad (66)$$

Equation (66) can be used to determine $\xi_g(x, y)$ at any given time from a known field $\Phi(x, y)$. Alternatively, (66) can be solved by inverting the Laplacian operator to determine Φ from a known distribution of ξ provided that suitable conditions on Φ are specified on the boundaries of the region in question. This *invertibility* is one reason why vorticity is such a useful forecast diagnostic; if the evolution of vorticity can be predicted, then inversion of Eq. (66) yields the evolution of the geopotential field, from which it is possible to determine the geostrophic wind. Since the Laplacian of a field tends to be a maximum where the function itself is a minimum, positive vorticity implies low values of geopotential and vice versa (see Fig. 7). We will use the invertibility to solve a problem numerically in section 3.

The quasi-geostrophic vorticity equation can be obtained from the x and y components of the quasi-geostrophic momentum equation (62) and yields (exercise!)

$$\frac{d_g \xi_g}{dt} = -f_0 \left(\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y} \right) - \beta v_g \quad , \quad (67)$$

which should be compared with Eq. (11). Thus the quasi-geostrophic vorticity equation takes the form of the barotropic vorticity equation! Using (63), Equation (67) can be re-written as

$$\frac{\partial \xi_g}{\partial t} = -\mathbf{v}_g \cdot \nabla (\xi_g + f) + f_0 \frac{\partial \omega}{\partial p} \quad , \quad (68)$$

which states that the local rate of change of geostrophic vorticity is given by the sum of the advection of the absolute vorticity by the geostrophic wind plus the concentration or dilution of vorticity by stretching or shrinking of fluid columns (the divergence effect). The vorticity tendency owing to vorticity advection [the first term on the right in Eq. (68)] may be rewritten as

$$-\mathbf{v}_g \cdot \nabla (\xi_g + f) = -\mathbf{v}_g \cdot \nabla \xi_g - \beta v_g \quad . \quad (69)$$

The two terms on the right represent the geostrophic advections of relative vorticity and the planetary vorticity, respectively. For disturbances in the westerlies, these two effects tend to have opposite signs. In the upstream of a 500 hPa trough, the geostrophic wind is directed from the negative vorticity maximum at the ridge toward the positive vorticity maximum at the trough so that $-\mathbf{v}_g \cdot \nabla \xi_g < 0$. But at the same time, since $v_g < 0$ in that region, the geostrophic wind has its y component directed down the gradient of planetary vorticity so that $-\beta v_g > 0$. Hence, in this region the advection of relative vorticity tends to decrease the local relative vorticity, whereas the advection of planetary vorticity tends to increase the local relative vorticity. Similar arguments (but with reversed signs) apply to a region downstream a trough. Therefore, advection of relative vorticity tends to move the vorticity and trough (and ridge) pattern eastward (downstream). But advection of planetary vorticity tends to move the troughs and ridges westward against the advecting wind field.

The net effect of advection on the evolution of the vorticity pattern depends upon which type of advection dominates. Given a geopotential height wavy field, the vorticity increases with the square of the wave number, so that the first term on the right of Eq. (69) is larger for large wave numbers (i.e. short waves; typically $L_x < 3000$ km), while for long waves ($L_x > 10000$ km) the planetary vorticity advection tends to dominate. Therefore, as a general rule, short wavelength synoptic-scale systems should move eastward with the advecting zonal flow while long planetary waves should tend to be stationary or move against the zonal advection. This will be discussed in more details when we derive the dispersion relation for Rossby waves.

Vorticity advection does not alone determine the evolution of meteorological systems. The orographic effects, for example seems to have vanished from Eq. (68). But they are still present, because orography will lead to vertical motions that make the second term on the right important.

Exercises

1. Show that

$$\frac{RT}{c_p p} - \frac{\partial T}{\partial p} = -\frac{T}{\theta} \frac{\partial \theta}{\partial p} = (\Gamma_d - \Gamma)/(\rho g)$$

using the definition of potential temperature, and dry adiabatic and actual lapse rates.

2. Show that from (62) follows the quasi-geostrophic vorticity equation (67). [Hint: apply $\partial/\partial x$ to the second component of Eq. (62) and subtract $\partial/\partial y$ applied to the first component of Eq. (62)].
3. Derive the thermal wind equation for u-component of the zonal wind (Eq. 53) and also for the v-component in pressure coordinates using the geostrophic relation 55.
4. Suppose that on the 500 hPa surface the relative vorticity at a certain location at 45° N latitude is increasing at a rate of $3 \times 10^{-6} s^{-1}$ per 3 h. The wind is from southwest at 20 m/s. and the relative vorticity decreases toward the northeast at a rate of $4 \times 10^{-6} s^{-1}$ per 100 km. Use the quasi-geostrophic vorticity equation to estimate the horizontal divergence at this location on a β plane.
5. Given the following expression for the geopotential field:

$$\Phi(x, y, p, t) = \Phi_0(p) + f_0[-Uy + k^{-1}V \cos(\pi p/p_0) \sin k(x - ct)] \quad , \quad (70)$$

where U, V, c, k, p_0 are constants, use the quasi-geostrophic vorticity equation (68) to obtain an estimate for ω . Assume that $df/dy = \beta$ is a constant (not zero) and that ω vanishes for $p = p_0$.

3 Rossby Waves

Suggested Literature:

1. Hoskins, B.J. and Karoly D.J., 1981: 'The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing', *J. Climate*, **38**, 1179-1196

3.1 Free Barotropic Rossby Waves

The dispersion relation for free barotropic Rossby waves can be derived by linearizing the barotropic vorticity equation in the form (21). This equation states that the absolute (geostrophic) vorticity is conserved following the horizontal (geostrophic) motion. As usual, we assume that the fields can be expressed as small perturbations from a basic state $\psi = \bar{\psi} + \psi'$. We linearize using a basic state that has only flow in zonal direction $\bar{\psi} = -\bar{u}y + \text{const}$. This mean state fulfills Eq. (21). With this mean state $\nabla^2\psi = \nabla^2\psi'$. Thus, by linearizing, in the first term the total derivative operator can be replaced by the mean operator and it follows

$$\left(\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x}\right)\nabla^2\psi' + \beta\frac{\partial\psi'}{\partial x} = 0 \quad . \quad (71)$$

As usual, we seek for solutions of the type

$$\psi' = Ae^{i(kx+ly-\nu t)} \quad . \quad (72)$$

Inserting (72) into (71) yields the dispersion relation

$$(-\nu + k\bar{u})(-k^2 - l^2) + k\beta = 0 \quad , \quad (73)$$

which we can solve immediately for ν

$$\nu = \bar{u}k - \beta k/K^2 \quad , \quad (74)$$

where $K^2 \equiv k^2 + l^2$ is the total horizontal wave number squared. Recalling that $c_x = \nu/k$, we find that the zonal phase speed relative to the mean wind is

$$c_x - \bar{u} = -\beta/K^2 \quad . \quad (75)$$

Thus, the Rossby wave zonal phase propagation is always westward relative to the mean zonal flow. Furthermore, the Rossby wave phase speed depends inversely on the square of the horizontal wave number. Therefore, Rossby waves are dispersive waves whose phase speeds increase rapidly with increasing wavelength. This result is consistent with the discussion in section 2.4, in which we showed that the advection of planetary vorticity, which tends to make the disturbances *retrogress*, increasingly dominates over relative vorticity advection as the wavelength of a disturbance increases. Equation (75) provides a quantitative measure of this effect in cases where the disturbance is small enough in amplitude.

From Eq. (75) we may calculate the stationary free Rossby wave wavelength

$$K^2 = \beta/\bar{u} \equiv K_s^2 \quad . \quad (76)$$

This means that stationary free Rossby waves only exist if there is a positive mean flow \bar{u} . This condition is important for Rossby waves that may be generated by tropical convection.

The group velocity of Rossby waves may be calculated as (exercise!):

$$c_{gx} \equiv \frac{\partial \nu}{\partial k} = \bar{u} + \beta \frac{k^2 - l^2}{K^4} \quad (77)$$

$$c_{gy} \equiv \frac{\partial \nu}{\partial l} = 2 \frac{\beta k l}{K^4} \quad . \quad (78)$$

Therefore, the energy propagation of stationary Rossby waves is always eastward (Fig. 8; exercise!).

These waves can also be derived from the original, compressible equations, but the analysis is much more complicated. There are some minor modifications in the phase velocities if the full equations are considered, but the main results remain valid.

3.2 Forced Topographic Rossby waves

Forced stationary Rossby waves are of primary importance for understanding the planetary-scale circulation pattern. Such modes may be forced by longitudinal dependent latent heating, or by flow over topography. Of particular importance for the Northern Hemisphere extratropical circulation are stationary Rossby modes forced by flow over the Rockies and the Himalayas.

As the simplest possible dynamical model of topographic Rossby waves we use the barotropic vorticity equation for a homogeneous fluid of variable depth (e.g. Eqs. 14 or 17). We assume that the upper boundary is at fixed height H and the lower boundary is at the variable height $h_T(x, y)$. We also use the quasi-geostrophic scaling $|\xi| \ll f_0$. Then, from 14 and 17 we have

$$H \frac{d_h(\xi + f)}{dt} = -f_0 \frac{dh_T}{dt} \quad , \quad (79)$$

where it has been also assumed that $h \equiv H - h_t \approx H$ on the left side (i.e. the mountain height is much smaller than the troposphere height). After linearizing (as we did to derive Eq. 71)

$$\left(\frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} \right) \nabla^2 \psi' + \beta \frac{\partial \psi'}{\partial x} = -\frac{f_0}{H} \bar{u} \frac{\partial h_T}{\partial x} \quad . \quad (80)$$

Lets consider the solutions of Eq. (80) for the special case of a sinusoidal lower boundary. We specify the topography to have the form

$$h_T(x, y) = h_0 \sin(kx + \phi) \cos ly \quad , \quad (81)$$

where ϕ is an arbitrary phase (therefore equivalent to $A \cos kx + B \sin kx$). If we insert the streamfunction perturbation

$$\psi' = \psi_0 \sin(kx + \phi) \cos ly \quad , \quad (82)$$

then Eq. (80) has the steady-state solution (i.e. dropping the partial time derivative) [exercise!]

$$\psi_0 = f_0 h_0 / [H(K^2 - K_s^2)] \quad . \quad (83)$$

The streamfunction is either exactly in phase (ridges over the mountains) or exactly out of phase (troughs over the mountains) with the topography depending on the sign $K^2 - K_s^2$. For long waves, ($K < K_s$), the topographic vorticity source in Eq. (80) is primarily balanced by meridional advection of planetary vorticity (the β effect). For short waves ($K > K_s$) the source is balanced primarily by the zonal advection of relative vorticity.

The topographic wave solution (83) has the unrealistic characteristic that when the wave number exactly equals the critical wave number K_s the amplitude goes to infinity. This is the resonant response case when the wave number reaches the stationary wave number of free Rossby waves.

Fig. 8 gives another example for a stationary Rossby wave, caused by ENSO forcing (discuss Eq. 11).

3.3 Turning Latitude

In reality the theory applied here with a constant β and u is a little too over-simplified, and a more correct treatment would make use of the dynamics in spherical coordinates (as in e.g. Hoskins and Karoly, 1981). However, we can derive some properties for an initially north-eastward propagating stationary Rossby wave here knowing that β slowly decreases in the meridional direction. Let us consider the stationary Rossby wave 85

$$k^2 + l^2 = \beta / \bar{u} \quad . \quad (84)$$

Let's assume a wave generated by ENSO in the tropics moves north-eastward, and that its zonal wave number is a constant. If we further take into account that β decreases to the north with the cosine of the latitude, then the meridional wave number l must decrease until it becomes 0. From this point the wave turns southward again. The latitude in which this occurs is called *turning latitude*, and it is an important property of stationary Rossby waves generated in the tropics. Try to identify the turning latitude in Fig. 8.

We can go a step further, and let also the mean wind u depend on latitude, in which case Eq. 84 has an additional term:

$$k^2 + l^2 = \frac{\left(\beta - \frac{d^2 u}{dy^2}\right)}{\bar{u}} \quad . \quad (85)$$

There are some metric terms missing in this equation, but this expression gives a hint why a strong jet can modify stationary Rossby waves (the full correct expression can be found in Hoskins and Karoly, 1981). Strong jets are therefore also able to modify the turning latitude and other properties of stationary Rossby waves. Fig. 9 shows two examples of stationary wave number distribution in the meridional direction versus the zonal wavenumber for regions with a strong jet (South Asian/Western Pacific region; solid line) and one with a weaker jet (Eastern Pacific region). North of a strong jet the turning latitude is reduced, and we get an effect called *waveguide*, e.g. wave numbers 5 and 6 are essentially trapped in the region between 25° and 35°N . How do you determine the turning latitude in this graph?

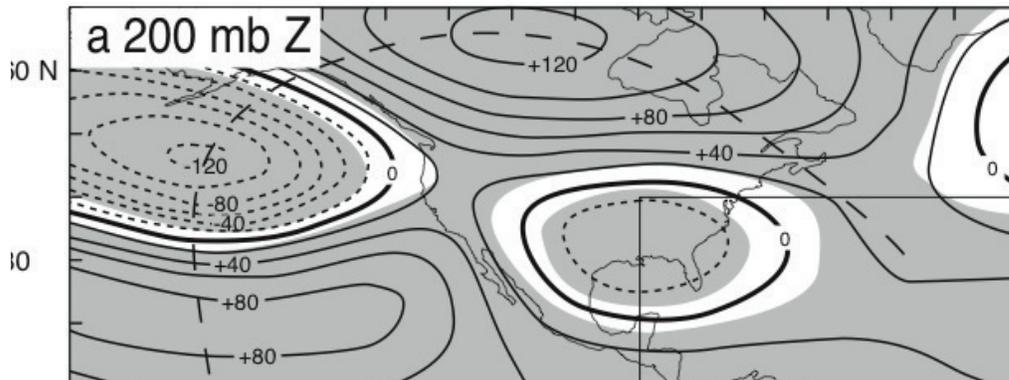


Figure 8: Stationary Rossby wave induced by ENSO.

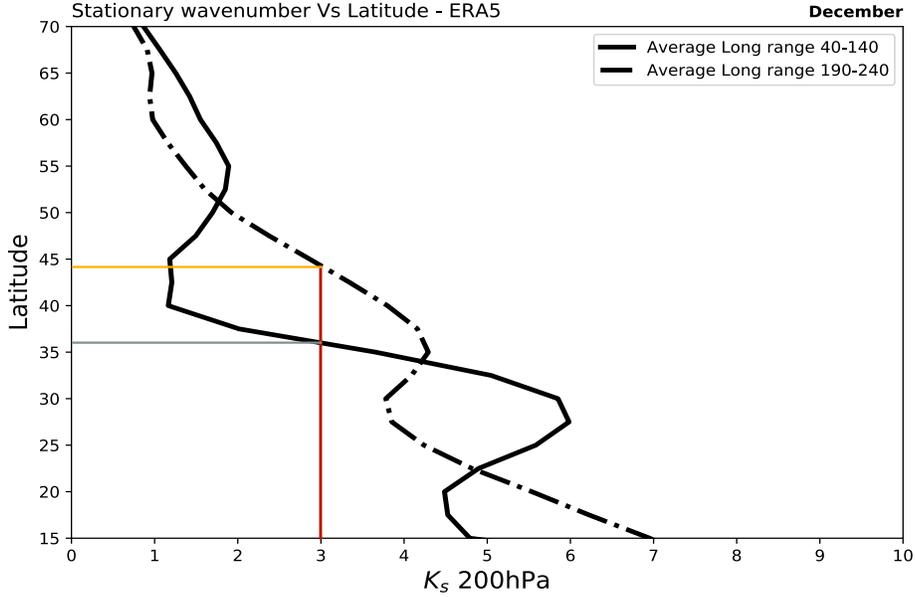


Figure 9: Meridional profile of stationary wave number (K_s). From Master thesis of Alessandro Raganato.

Exercises

1. Derive the group velocities for Rossby waves (77) and (78) and show that for stationary Rossby waves fulfilling Eq. (85), the c_{gx} component is always positive.
2. Show that (83) is the solution of (80) with (81).
3. Using the linearized form of the potential vorticity equation (11) and the β -plane approximation, derive the Rossby wave speed for a homogeneous incompressible ocean of depth h . Assume a motionless basic state and small perturbations that depend only on x and t ,

$$u = u'(x, t), \quad v = v'(x, t), \quad h = H + h'(x, t) \quad , \quad (86)$$

where H is the mean depth of the ocean. With the aid of the continuity equation for a homogeneous layer

$$\frac{\partial h'}{\partial t} + H \frac{\partial u'}{\partial x} = 0 \quad (87)$$

and the geostrophic wind relationship $v' = g f_0^{-1} \partial h' / \partial x$. Show that the perturbation vorticity equation can be written in the form

$$\frac{\partial}{\partial t} \left(\frac{\partial^2}{\partial x^2} - \frac{f_0^2}{gH} \right) h' + \beta \frac{\partial h'}{\partial x} = 0 \quad (88)$$

and that $h' = h_0 e^{ik(x-ct)}$ is a solution provided that

$$c = -\beta(k^2 + f_0^2/gH)^{-1} \quad . \quad (89)$$

If the ocean is 4 km deep, what is the Rossby wave speed at latitude 45° N for a wave of 10000 km zonal wavelength?

4. Rossby-type waves can be generated in a rotating cylindrical vessel if the depth of the fluid is dependent on the radial coordinate. To determine the Rossby wave speed formula for this equivalent β effect, we assume that the flow is confined between rigid lids in an annular region whose distance from the axis of rotation is large enough so that the curvature terms in the equations can be neglected. We then can refer the motion to cartesian coordinates with x directed azimuthally and y directed toward the axis of rotation. If the system is rotating at angular velocity Ω and the depth is linearly dependent on y ,

$$H(y) = H_0 - \gamma y \quad , \quad (90)$$

show that the perturbation (shallow water) continuity equation ($dH/dt = -H\nabla \cdot \mathbf{v}$) can be written as

$$H_0 \left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} \right) - \gamma v' = 0 \quad (91)$$

and that the perturbation quasi-geostrophic vorticity equation is thus

$$\frac{\partial}{\partial t} \nabla^2 \psi' + \beta \frac{\partial \psi'}{\partial x} = 0 \quad , \quad (92)$$

where ψ' is the perturbation geostrophic streamfunction and $\beta = 2\Omega\gamma/H_0$. What is the Rossby wave speed in this situation for waves of wavelength 100 cm in both the x and y directions if $\Omega = 1s^{-1}$, $H_0 = 20$ cm, and $\gamma = 0.05$? (Hint: Assume that the velocity field is geostrophic except in the divergence term.)

5. Solve the nonlinear potential vorticity conservation equation (19) using (22) and including a Ekman pumping term $r_e \xi$, $r_e = 1/day$ numerically for a channel with the centre at 45°N of $L_y = 3 \cdot 10^6$ m meridional width and a zonally periodic domains of a length of $L_x = 2 \cdot 10^7$ m using a spatial discretization of $\Delta x = \Delta y = d = 1 \cdot 10^5$ m and a Δt of 1 h. Assume that the top is fixed at a height $H = 1.2 \cdot 10^4$ m, so that the total height of the fluid is given by $h = H - h_t$. Let a sinusoidal mountain be if the shape

$$h_t(x, y) = h_0 \sin(N2\pi x/L_x) \sin(\pi y/L_y) \quad ,$$

where $h_0 = 1 \cdot 10^3$ m, and let N (the number of the mountain waves in the channel) be a) N=2 and b) N=8. The initial condition zonal flow, which, expressed in streamfunction means

$$\psi(x, y, 0) = -10(y - L_y) \quad .$$

Compare the solutions for a) and b) (after 7 days) and especially compare the position of the eddy streamfunction crests relative to the mountain crests for both cases and interpret the results. Hint: Equation (19), with the condition of a fixed upper height H and given lower topography, including Ekman pumping terms can be written as

$$\frac{\partial}{\partial t}\xi = F(x, y, t) - r_e\xi = - \left[\frac{\partial}{\partial x}(u\xi) + \frac{\partial}{\partial y}(v\xi) \right] - \beta v + \frac{f_0}{h} \left[\frac{\partial}{\partial x}(uh) + \frac{\partial}{\partial y}(vh) \right] - r_e\xi \quad . \quad (93)$$

and

$$u = - \frac{\partial\psi}{\partial y} \quad (94)$$

$$v = \frac{\partial\psi}{\partial x} \quad (95)$$

$$\xi = \nabla^2\psi \quad , \quad (96)$$

given that the quasi-geostrophic approximations (15) and (16) are valid. Note that Eq. (93) can be written in its formulation because the geostrophic wind (Eq. 94) is divergence free and the height does not depend explicitly on time.

Discretize the terms in Eq. (93) as (using an implicit discretization of the Ekman damping)

$$\xi(t + \Delta t) = \xi(t) + \Delta t F_{i,j}(t) - \Delta t r_e \xi(t + \Delta t) \quad (97)$$

with

$$\begin{aligned} F_{i,j}(t) &= - \frac{1}{2d} [(u_{i+1,j}\xi_{i+1,j} - u_{i-1,j}\xi_{i-1,j}) + (v_{i,j+1}\xi_{i,j+1} - v_{i,j-1}\xi_{i,j-1})] - \beta v_{i,j} \\ &+ \frac{f_0}{h_{i+1,j}} \frac{1}{2d} \\ &[(u_{i+1,j}h_{i+1,j} - u_{i-1,j}h_{i-1,j}) + (v_{i,j+1}h_{i,j+1} - v_{i,j-1}h_{i,j-1})] \quad (98) \end{aligned}$$

where the right side is evaluated at the time t , where the fields are already known. This leads to

$$\xi(t + \Delta t) = (\xi(t) + \Delta t F_{i,j}) / (1 + \Delta t r_e) \quad (99)$$

Knowing the vorticity, the streamfunction can be determined by Eq. (96), which can be discretized as (see MMG lectures)

$$(\psi_{i+1,j} + \psi_{i-1,j} + \psi_{i,j+1} + \psi_{i,j-1} - 4\psi_{i,j}) / d^2 = \xi_{i,j} \quad . \quad (100)$$

If we can solve this equation for $\psi_{i,j}$, we can derive the velocity fields by using Eqs. (94), (95) in discretized form

$$u_{i,j} = -(\psi_{i,j+1} - \psi_{i,j-1}) / (2d) \quad (101)$$

$$v_{i,j} = (\psi_{i+1,j} - \psi_{i-1,j}) / (2d) \quad (102)$$

$$(103)$$

A scheme how to solve the initial value problem is:

- (a) Given the initial condition $\psi_{i,j}$, (100) can be solved to determine $\xi_{i,j}$, (101) and (102) can be used to determine the velocities
- (b) Evaluate $F_{i,j}$ using Eq. (98).
- (c) $\xi_{t+\Delta t}$ can be determined by integrating (99).
- (d) Knowing $\xi_{t+\Delta t}$, (100) can be inverted to calculate $\psi_{t+\Delta t}$
- (e) Go back to (a).

Integrate this scheme for 7 days to reach a steady-state solution and plot the eddy streamfunction and topography.

4 Baroclinic Instability

4.1 A two-layer Model

Even for a highly idealized mean flow profile the mathematical treatment of baroclinic instability in a continuously stratified atmosphere is rather complicated. Thus, we focus on the simplest model that can incorporate baroclinic processes. The atmosphere is represented by two discrete layers bounded by surfaces numbered 0, 2, and 4 (generally taken to be the 0-, 500-, and 1000-hPa surfaces, respectively). The quasi-geostrophic vorticity equation for the midlatitude β plane is applied at levels denoted 1 and 3 and the thermodynamic energy equation is applied at level 2. Before writing the specific equations of the two-layer model, it is convenient to define a *geostrophic streamfunction*, $\psi \equiv \Phi/f_0$ (see definitions leading to Eq. 21). Then the geostrophic wind (Eq. 55) and the geostrophic vorticity (Eq. 66) can be expressed as

$$\mathbf{v} = \mathbf{k} \times \nabla\psi, \quad \xi = \nabla^2\psi \quad (104)$$

The quasi-hydrostatic vorticity equation (68) and the hydrostatic thermodynamic equation (65) can be written with help of (37) in terms of ψ and ω as (assuming no diabatic processes)

$$\frac{\partial}{\partial t} \nabla^2\psi + \mathbf{v} \cdot \nabla(\nabla^2\psi) + \beta \frac{\partial\psi}{\partial x} = f_0 \frac{\partial\omega}{\partial p} \quad (105)$$

$$\frac{\partial}{\partial t} \left(\frac{\partial\psi}{\partial p} \right) = -\mathbf{v} \cdot \nabla \left(\frac{\partial\psi}{\partial p} \right) - \frac{\sigma}{f_0} \omega \quad . \quad (106)$$

We now apply the vorticity equation (105) at the two levels designated as 1 and 3, which are in the middle of the two layers. To do this we must estimate the divergence term $\partial\omega/\partial p$ at these levels using finite difference approximations to the vertical derivatives

$$\left(\frac{\partial\omega}{\partial p} \right)_1 \approx \frac{\omega_2 - \omega_0}{\delta p}, \quad \left(\frac{\partial\omega}{\partial p} \right)_3 \approx \frac{\omega_4 - \omega_2}{\delta p}, \quad (107)$$

where δp is the pressure interval between levels 0-2 and 2-4 and subscript notation is used to designate the vertical level for each dependent variable. The resulting vorticity equations are

$$\frac{\partial}{\partial t} \nabla^2\psi_1 + \mathbf{v}_1 \cdot \nabla(\nabla^2\psi_1) + \beta \frac{\partial\psi_1}{\partial x} = f_0 \frac{\omega_2}{\delta p} \quad (108)$$

$$\frac{\partial}{\partial t} \nabla^2\psi_3 + \mathbf{v}_3 \cdot \nabla(\nabla^2\psi_3) + \beta \frac{\partial\psi_3}{\partial x} = -f_0 \frac{\omega_2}{\delta p} \quad , \quad (109)$$

where we have used the fact that $\omega_0 = 0$ and assumed that $\omega_4 = 0$, which is approximately true for a level lower boundary surface. We next write the thermodynamic energy equation (106) at level 2. Here we must evaluate $\partial\psi/\partial p$ using the difference formula

$$(\partial\psi/\partial p) \approx (\psi_3 - \psi_1)/\delta p \quad .$$

The result is

$$\frac{\partial}{\partial t}(\psi_1 - \psi_3) = -\mathbf{v}_2 \cdot \nabla(\psi_1 - \psi_3) + \frac{\sigma \delta p}{f_0} \omega_2 \quad . \quad (110)$$

The first term on the right-hand side in Eq. (110) is the advection of the 250-750 hPa thickness by the wind at 500 hPa. However, ψ_2 , the 500 hPa streamfunction, is not a predicted field in this model. Therefore, ψ_2 must be obtained by linearly interpolating between the 250- and 750-hPa levels

$$\psi_2 = (\psi_1 + \psi_3)/2 \quad . \quad (111)$$

If this interpolation formula is used, (108)-(110) become a closed set of prediction equations in the variables ψ_1, ψ_3 , and ω_2 .

4.2 Linear Perturbation Analysis

To keep the analysis as simple as possible we assume that the streamfunctions ψ_1 and ψ_3 consist of basic state parts that depend linearly on y alone, plus perturbations that depend only on x and t (similar to section 3). Thus, we let

$$\begin{aligned} \psi_1 &= -U_1 y + \psi'_1(x, t) \\ \psi_3 &= -U_3 y + \psi'_3(x, t) \\ \omega_2 &= \omega'_2(x, t) \quad . \end{aligned} \quad (112)$$

The zonal velocities at levels 1 and 3 are then constants with the values U_1 and U_3 , respectively. Hence, the perturbation field has meridional and vertical velocity components only. Inserting (112) into (108)-(110) and linearizing yields the perturbation equations (see section 3)

$$\left(\frac{\partial}{\partial t} + U_1 \frac{\partial}{\partial x} \right) \frac{\partial^2 \psi'_1}{\partial x^2} + \beta \frac{\partial \psi'_1}{\partial x} = f_0 \frac{\omega'_2}{\delta p} \quad (113)$$

$$\left(\frac{\partial}{\partial t} + U_3 \frac{\partial}{\partial x} \right) \frac{\partial^2 \psi'_3}{\partial x^2} + \beta \frac{\partial \psi'_3}{\partial x} = -f_0 \frac{\omega'_2}{\delta p} \quad (114)$$

$$\left(\frac{\partial}{\partial t} + U_m \frac{\partial}{\partial x} \right) (\psi'_1 - \psi'_3) - U_T \frac{\partial}{\partial x} (\psi'_1 + \psi'_3) = \frac{\sigma \delta p}{f_0} \omega'_2 \quad , \quad (115)$$

where we have linearly interpolated to express \mathbf{v}_2 in terms of ψ_1 and ψ_3 and have defined

$$U_m \equiv (U_1 + U_3)/2, \quad U_T \equiv (U_1 - U_3)/2 \quad .$$

Thus, U_m and U_T are, respectively, the vertically averaged mean zonal wind and the mean thermal wind for the interval $\delta p/2$. The dynamical properties of this system are more clearly expressed if (113)-(115) are combined to eliminate ω'_2 . We first note that (113) and (114) can be rewritten as

$$\left(\frac{\partial}{\partial t} + (U_m + U_T) \frac{\partial}{\partial x} \right) \frac{\partial^2 \psi'_1}{\partial x^2} + \beta \frac{\partial \psi'_1}{\partial x} = f_0 \frac{\omega'_2}{\delta p} \quad (116)$$

$$\left(\frac{\partial}{\partial t} + (U_m - U_T) \frac{\partial}{\partial x} \right) \frac{\partial^2 \psi'_3}{\partial x^2} + \beta \frac{\partial \psi'_3}{\partial x} = -f_0 \frac{\omega'_2}{\delta p} \quad . \quad (117)$$

We now define the barotropic and baroclinic perturbations as

$$\psi_m \equiv (\psi'_1 + \psi'_3)/2, \quad \psi_T \equiv (\psi'_1 - \psi'_3)/2 \quad (118)$$

Adding (116) and (117) and using the definitions in (118) yields

$$\left[\frac{\partial}{\partial t} + U_m \frac{\partial}{\partial x} \right] \frac{\partial^2 \psi_m}{\partial x^2} + \beta \frac{\partial \psi_m}{\partial x} + U_T \frac{\partial}{\partial x} \left(\frac{\partial^2 \psi_T}{\partial x^2} \right) = 0 \quad , \quad (119)$$

while subtracting (117) from (116) and combining with (115) to eliminate ω'_2 yields

$$\left[\frac{\partial}{\partial t} + U_m \frac{\partial}{\partial x} \right] \left(\frac{\partial^2 \psi_T}{\partial x^2} - 2\lambda^2 \psi_T \right) + \beta \frac{\partial \psi_T}{\partial x} + U_T \frac{\partial}{\partial x} \left(\frac{\partial^2 \psi_m}{\partial x^2} + 2\lambda^2 \psi_m \right) = 0 \quad , \quad (120)$$

where $\lambda^2 \equiv f_0^2/[\sigma(\delta p)^2]$. Equations (119) and (120) govern the evolution of the barotropic (vertically averaged) and baroclinic (thermal) perturbation vorticities, respectively. As usual we assume that wavelike solutions exist of the form

$$\psi_m = A e^{ik(x-ct)}, \quad \psi_T = B e^{ik(x-ct)} \quad . \quad (121)$$

Substituting these assumed solutions into (119) and (120) and dividing through by the common exponential factor, we obtain a pair of simultaneous linear algebraic equations for the coefficients of A, B

$$ik[(c - U_m)k^2 + \beta]A - ik^3 U_T B = 0 \quad (122)$$

$$ik[(c - U_m)(k^2 + 2\lambda^2) + \beta]B - ik U_T (k^2 - 2\lambda^2)A = 0 \quad . \quad (123)$$

From the Mathematical Methods course we know that a homogeneous set of equations has only nontrivial solutions if the determinant of the coefficients for A and B is zero. Thus the phase speed c must satisfy the condition

$$\begin{vmatrix} (c - U_m)k^2 + \beta & -k^2 U_T \\ -U_T(k^2 - 2\lambda^2) & (c - U_m)(k^2 + 2\lambda^2) + \beta \end{vmatrix} = 0 \quad , \quad (124)$$

which gives a quadratic dispersion equation in c

$$(c - U_m)^2 k^2 (k^2 + 2\lambda^2) + 2(c - U_m)\beta(k^2 + \lambda^2) + [\beta^2 + U_T^2 k^2 (2\lambda^2 - k^2)] = 0 \quad , \quad (125)$$

The solution for c is

$$c = U_m - \frac{\beta(k^2 + \lambda^2)}{k^2(k^2 + 2\lambda^2)} \pm \delta^{1/2} \quad , \quad (126)$$

where

$$\delta \equiv \frac{\beta^2 \lambda^4}{k^4 (k^2 + 2\lambda^2)^2} - \frac{U_T^2 (2\lambda^2 - k^2)}{(k^2 + 2\lambda^2)} \quad . \quad (127)$$

We have shown that (121) is a solution for the system (119) and (120) only if the phase speed satisfies (126). Although (126) appears to be rather complicated, it is immediately apparent that if $\delta < 0$ the phase speed will have an imaginary part and

the perturbations will amplify exponentially. Before discussing the general physical conditions required for exponential growth it is useful to consider two special cases.

As the first special case we let $U_T = 0$ so that the basic state thermal wind vanishes and the mean flow is barotropic. There can be no instability if the thermal wind vanishes (i. e. without horizontal mean-state temperature gradients). The *available potential energy* stored in the mean state temperature gradients is responsible for baroclinic growth! The phase speeds in this case are

$$c_1 = U_m - \beta k^{-2} \quad (128)$$

and

$$c_2 = U_m - \beta(k^2 + 2\lambda^2)^{-1} \quad (129)$$

These are real quantities that correspond to the free (normal mode) oscillations for the two-level model with a barotropic basic state current. The phase speed c_1 is simply the dispersion relationship for a barotropic Rossby wave with no y dependence (see Eq. [75]). Substituting the expression (128) in place of c in (122) and (123) we see that in this case $B = 0$, so that the perturbation is barotropic in structure. The expression (129), on the other hand, may be interpreted as the phase speed of an internal baroclinic Rossby wave. Note that c_2 is a dispersion relationship analogous to the Rossby wave speed for a homogeneous ocean with a free surface, which was given in problem 3 of section 3. But, in the two-level model, the factor $2\lambda^2$ appears in the denominator in place of the f_0/gH for the oceanic case. In each of these cases there is vertical motion associated with the Rossby wave so that static stability modified the phase speed.

Comparing (128) and (129) we see that the phase speed of the baroclinic mode is generally much less than that of the barotropic mode, since for average midlatitude tropospheric conditions $\lambda^2 \approx 2 \times 10^{-12} \text{ m}^{-2}$, which is comparable in magnitude to k for zonal wavelength of $\sim 4500 \text{ km}$.

Returning to the general case where all terms are retained in (126), the stability criterion is most easily understood by computing the *neutral curve*, which connects all values of U_T and k for which $\delta = 0$ so that the flow is *marginally stable*. From Eq. (126), the condition $\delta = 0$ implies that

$$\frac{\beta^2 \lambda^4}{k^4(2\lambda^2 + k^2)} = U_T^2(2\lambda^2 - k^2) \quad (130)$$

or

$$k^4/(4\lambda^4) = 1/2\{1 \pm [1 - \beta^2/(4\lambda^4 U_T^2)]^{1/2}\} \quad (131)$$

Fig. 4.2 shows nondimensional quantity $k^2/2\lambda^2$, which is a measure of the zonal wavelength, plotted against the nondimensional parameter $2\lambda^2 U_T/\beta$, which is proportional to the thermal wind, according to Eq. (131).

As indicated in the figure, the neutral curve separates the unstable region of the U_T, k plane from the stable region. It is clear that the inclusion of the β effect serves to stabilize the flow, for unstable roots exist only for $|U_T| > \beta/(2\lambda^2)$. In addition to a

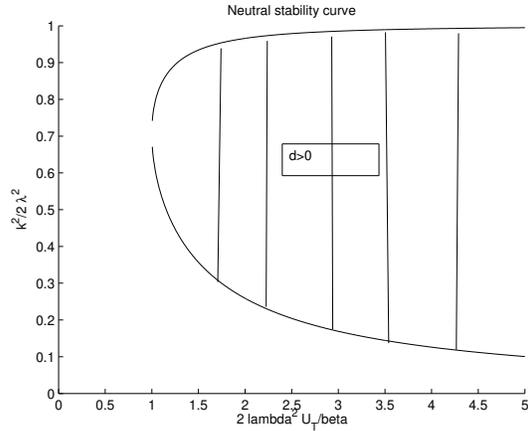


Figure 10: Neutral stability curve for the two-level baroclinic model

minimum value of U_T required for unstable growth depends strongly on k . Thus, the β effect strongly stabilizes the long-wave end of the wave spectrum ($k \rightarrow 0$). Again, the flow is always stable for waves shorter than the critical wavelength $L_c = \sqrt{2}\pi/\lambda$ (why?). The long-wave stabilization associated with the β effect is caused by the rapid westward propagation of long waves, which occurs only when the β effect is included in the model.

Differentiating Eq. (130) with respect to k and setting $dU_T/dk = 0$, we find the minimum value of U_T for which unstable waves exist occurs when $k^2 = \sqrt{2}\lambda^2$. This wave number corresponds to the wave of maximum instability. Wave numbers for observed disturbances should be close to the wave number of maximum instability, if U_T were gradually raised from zero the flow would first become unstable for perturbations of wave number $k = 2^{1/4}\lambda$. Those perturbations would then amplify and in the process remove energy from the mean thermal wind, thereby decreasing U_T and stabilizing the flow. Under normal conditions of static stability the wavelength of maximum instability is about 4000 km, which is close to the average wavelength for midlatitude synoptic systems.

Exercises

1. Suppose that a baroclinic fluid is confined between two rigid horizontal lids in a rotating tank in which $\beta = 0$ but friction is presented in the form of linear drag proportional to the velocity (i.e., $\mathbf{Fr} = -\mu\mathbf{v}$). Show that the two-level model perturbation vorticity equations in cartesian coordinates can be written as

$$\left(\frac{\partial}{\partial t} + U_1 \frac{\partial}{\partial x} + \mu \right) \frac{\partial^2 \psi'_1}{\partial x^2} - f_0 \frac{\omega'_2}{\delta p} = 0$$

$$\left(\frac{\partial}{\partial t} + U_3 \frac{\partial}{\partial x} + \mu \right) \frac{\partial^2 \psi'_3}{\partial x^2} + f_0 \frac{\omega'_2}{\delta p} = 0 \quad ,$$

where perturbations are assumed in the form given in Eq. (112). The thermodynamic equation remains (115). Assuming solutions of the form (121), show that the phase speed satisfies a relationship similar to (126), with β replaced everywhere by $i\mu k$ and that as a result the condition for baroclinic instability becomes

$$U_T > \mu(2\lambda^2 - k^2)^{-1/2} \quad .$$

5 Equatorial Wave Theory

Lecture based on 'An Introduction to Dynamic Meteorology' by J. R. Holton, *Academic Press, INC.*, 3rd edition, 511 pp.

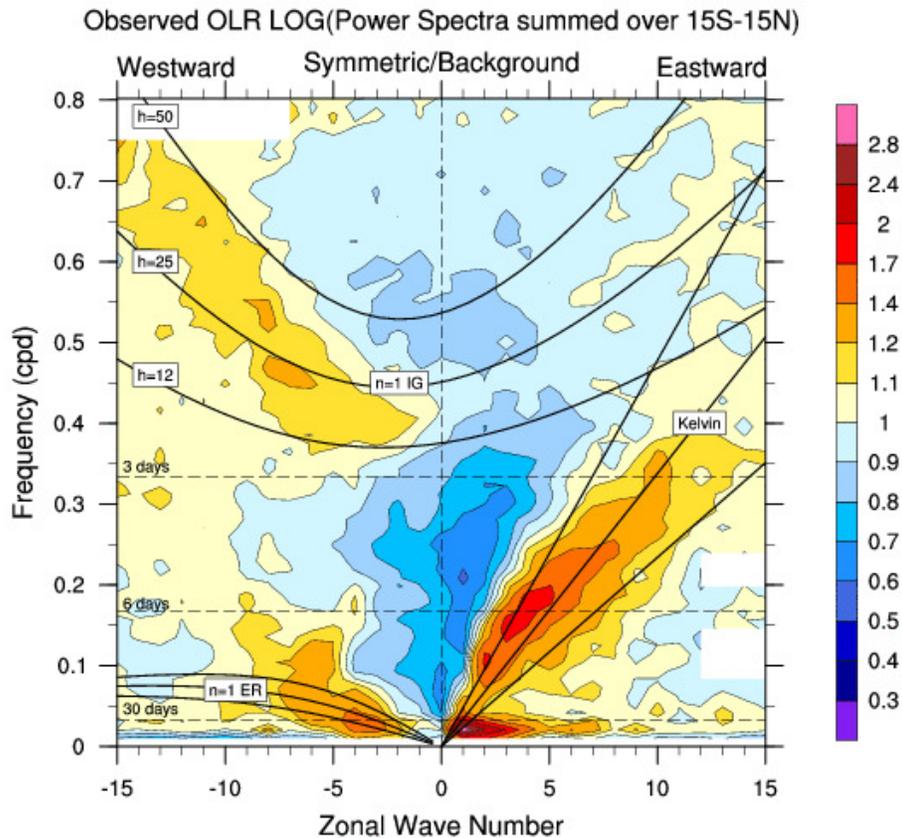


Figure 11: Dispersion diagram for tropical Outgoing longwave radiation. Source: www.cgd.ucar.edu, NOAA.

Equatorial waves are an important class of eastward and westward propagating disturbances that are trapped about the equator (that is, they decay away from the equatorial region). In a dispersion diagram for observed equatorial quantities, these wave can be identified as regions of increased energy density (Fig. 11).

Diabatic heating by organized tropical convection may excite equatorial wave motions (see Fig. 12). Through such waves the dynamical effects of convective storms can be communicated over large longitudinal distances in the tropics. Such waves can produce remote responses to localized heat sources. Furthermore, by influencing the pattern of low-level moisture convergence they can partly control the spatial and temporal distribution of convective heating. In order to introduce equatorial waves in the simplest possible context, we here use a shallow-water model

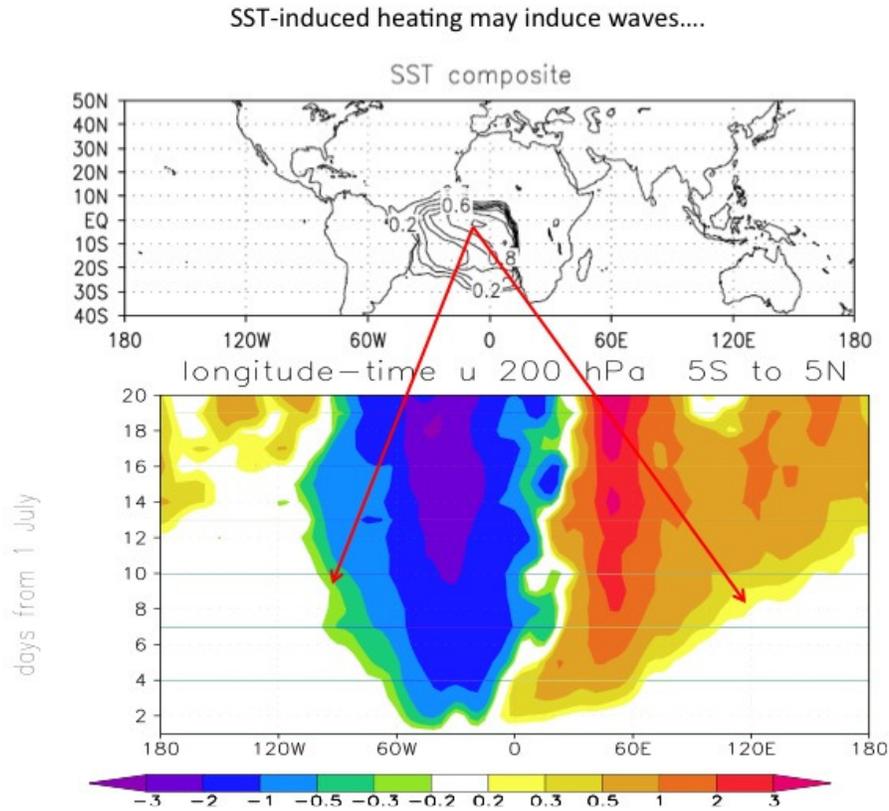


Figure 12: Equatorial Kelvin and Rossby waves triggered by an SST-induced heating. Source: Kucharski et al. 2008: A Gill-Matsuno-type mechanism explains the tropical Atlantic influence on African and Indian monsoon rainfall. *Q. J. R. Meteorol. Soc.* (2009), 135, 569-579, DOI: 10.1002/qj.406

and concentrate on the *horizontal* structure.

5.1 The shallow water equations

The shallow water equations are a drastic simplification to the real atmospheric flow. However, despite its simplicity it gives often a good insight into many atmospheric wave phenomena. The basic assumptions in the shallow water model are

- (i) The flow is incompressible $\rho = \text{const.}$
- (ii) The flow is shallow enough so that the horizontal velocity components are independent of height.
- (iii) The flow is hydrostatic. Accelerations in the vertical direction may be neglected.

Let us consider the horizontal momentum equations 1 and the hydrostatic equation 34

$$\frac{\partial u}{\partial t} + (\mathbf{v} \cdot \nabla)u = -\frac{1}{\rho} \frac{\partial p}{\partial x} + fv \quad (132)$$

$$\frac{\partial v}{\partial t} + (\mathbf{v} \cdot \nabla)v = -\frac{1}{\rho} \frac{\partial p}{\partial y} - fu \quad (133)$$

$$\frac{1}{\rho} \frac{\partial p}{\partial z} = -g, \quad (134)$$

Further consider the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (135)$$

Integrating the hydrostatic equation from a height z to the top of the fluid leads to (assuming the pressure is vanishing there)

$$\int_z^{h(x,y,t)} \frac{\partial p}{\partial z} dz = - \int_z^{h(x,y,t)} \rho g dz, \quad \text{or} \quad (136)$$

$$-p(x, y, z, t) = -\rho g [h(x, y, t) - z]. \quad (137)$$

Thus the horizontal pressure gradient force in the equations of motion 132, 133 may be expressed as

$$-\frac{1}{\rho} \frac{\partial p}{\partial x} = -g \frac{\partial h}{\partial x} = -\frac{\partial \Phi}{\partial x} \quad (138)$$

$$-\frac{1}{\rho} \frac{\partial p}{\partial y} = -g \frac{\partial h}{\partial y} = -\frac{\partial \Phi}{\partial y}, \quad (139)$$

where we have defined $\Phi(x, y, t) = gh(x, y, t)$. Thus, keeping in mind that there the horizontal velocities do not depend on the vertical direction and ignoring the coriolis term proportional to the vertical velocity, the horizontal equations of motion may be written as

$$\frac{\partial u}{\partial t} + (\mathbf{v} \cdot \nabla)u = -\frac{\partial \Phi}{\partial x} + fv \quad (140)$$

$$\frac{\partial v}{\partial t} + (\mathbf{v} \cdot \nabla)v = -\frac{\partial \Phi}{\partial y} - fu, \quad (141)$$

The number of dependent variables in Eqs. 140 and 141 is reduced to 3, (u, v, Φ) . Thus, if we have another equation only containing (u, v, Φ) , then the system may be complete. This is achieved by simplification of the continuity equation 135 and vertical integration. First, we note that because of $\rho = \text{const}$, Eq. 135 reduces to

$$\frac{\partial w}{\partial z} = -\nabla \cdot \mathbf{v}. \quad (142)$$

If we integrate this equation vertically from 0 to $h(x, y, t)$ we have

$$\int_0^h \frac{\partial w}{\partial z} dz = - \int_0^h \nabla \cdot \mathbf{v} dz \quad (143)$$

$$w(h) := \frac{dh}{dt} = \frac{\partial h}{\partial t} + \mathbf{v} \cdot \nabla h = -(\nabla \cdot \mathbf{v})h \quad (144)$$

Eq. 144 may as well be written as

$$\frac{\partial \Phi}{\partial t} + \mathbf{v} \cdot \nabla \Phi = -\Phi \nabla \cdot \mathbf{v} . \quad (145)$$

Eqs. 140, 141 and 145 build a complete set of differential equations for (u, v, Φ) , and are called the *shallow water equations*.

5.2 Linearization for an Equatorial β -plane

Now we linearize the set of equations 140, 141 and 145 about a motionless mean state with height h_e on an equatorial β -plane. Generally speaking, the β -plane assumption states that $f = 2|\omega| \sin \phi \approx f_0 + \beta y$, that is the $\sin \phi$ -dependence is approximated linearly for a given latitude ϕ_0 by a Taylor series expansion (therefore $\beta = 2|\omega| \cos \phi_0 / a$; a being the mean radius of the earth). If we set the base point at the equator we have $f_0 = 0$, therefore $f \approx \beta y$.

$$\frac{\partial u'}{\partial t} = -\frac{\partial \Phi'}{\partial x} + \beta y v' \quad (146)$$

$$\frac{\partial v'}{\partial t} = -\frac{\partial \Phi'}{\partial y} - \beta y u' \quad (147)$$

$$\frac{\partial \Phi'}{\partial t} = -gh_e \left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} \right) , \quad (148)$$

where the primed variables denote the perturbations from the basic state. This is our basic set of linearized equations (with variable coefficients!) to study equatorial wave dynamics. By adjusting the scale height h_e as well the ocean case may be included.

Discuss Inertia-Gravity waves for extratropical situation and approximation $f = f_0 = \text{const}$, and assume $u'(x, t), v'(x, t), \Phi'(x, t)$.

5.2.1 Equatorial Rossby and Rossby-Gravity Modes

In order to find solutions to the linearized system 146, 147 and 148, we assume that the y -dependence can be separated

$$\begin{pmatrix} u' \\ v' \\ \Phi' \end{pmatrix} = \begin{bmatrix} \hat{u}(y) \\ \hat{v}(y) \\ \hat{\Phi}(y) \end{bmatrix} e^{i(kx - \nu t)} . \quad (149)$$

Substitution of Eq. 149 into 146-148 then yields a set of ordinary differential equations in y for the meridional structure functions $\hat{u}, \hat{v}, \hat{\Phi}$:

$$-i\nu\hat{u} = -ik\hat{\Phi} + \beta y\hat{v} \quad (150)$$

$$-i\nu\hat{v} = -\frac{\partial\hat{\Phi}}{\partial y} - \beta y\hat{u} \quad (151)$$

$$-i\nu\hat{\Phi} = -gh_e \left(ik\hat{u} + \frac{\partial\hat{v}}{\partial y} \right). \quad (152)$$

If Eq. 150 is solved for $\hat{u} = k/\nu\hat{\Phi} + i\beta y\hat{v}/\nu$ and inserted into Eq. 151 and 152 we obtain

$$(\beta^2 y^2 - \nu^2)\hat{v} = ik\beta y\hat{\Phi} + i\nu\frac{\partial\hat{\Phi}}{\partial y} \quad (153)$$

$$(\nu^2 - gh_e k^2)\hat{\Phi} + i\nu gh_e \left(\frac{\partial\hat{v}}{\partial y} - \frac{k}{\nu}\beta y\hat{v} \right) = 0. \quad (154)$$

Finally, Eq. 154 is inserted into Eq. 153 to eliminate $\hat{\Phi}$, yielding a second-order differential equation in the single unknown, \hat{v}

$$\frac{\partial^2\hat{v}}{\partial y^2} + \left[\left(\frac{\nu^2}{gh_e} - k^2 - \frac{k}{\nu}\beta \right) - \frac{\beta^2 y^2}{gh_e} \right] \hat{v} = 0. \quad (155)$$

We seek solutions of this equation for the meridional distribution of \hat{v} , subject to the boundary condition that the disturbance fields vanish for $|y| \rightarrow \infty$. This boundary condition is necessary since the approximation $f \approx \beta y$ is not valid for latitudes much beyond $\pm 30^\circ$, so that solutions must be equatorially trapped if they are to be good approximations to the exact solutions on the sphere. Equation 155 differs from the classic equation for a harmonic oscillator in y because the coefficient in square brackets is not a constant but is a function of y . For sufficiently small y this coefficient is positive and solutions oscillate in y , while for large y , solutions either grow or decay in y . Only the decaying solutions, however, can satisfy the boundary conditions.

It turns out that solutions to Eq. 155 which satisfy the condition of decay far from the equator exist only when the constant part of the coefficient in square brackets satisfies the relationship (which is as well the dispersion relation!)

$$\frac{\sqrt{gh_e}}{\beta} \left(-\frac{k}{\nu}\beta - k^2 + \frac{\nu^2}{gh_e} \right) = 2n + 1; \quad n = 0, 1, 2, \dots \quad (156)$$

which is a cubic dispersion equation determining the frequencies of permitted equatorially trapped free oscillations for zonal wave number k and meridional mode number n . These solutions can be expressed most conveniently if y is replaced by the nondimensional meridional coordinate

$$\xi = \left(\frac{\beta}{\sqrt{gh_e}} \right)^{1/2} y. \quad (157)$$

With the Eqs. 156 and 157, Eq. 155 becomes

$$\frac{\partial^2 \hat{v}}{\partial \xi^2} + (2n + 1 - \xi^2) \hat{v} = 0 . \quad (158)$$

This is the differential equation for a quantum mechanical, simple harmonic oscillator. The solution has the form

$$\hat{v}(\xi) = H_n(\xi) e^{-\xi^2/2} , \quad (159)$$

where $H_n(\xi)$ designates the n th *Hermite polynomial*. The first of these polynomials have the values

$$H_0 = 1, \quad H_1(\xi) = 2\xi, \quad H_2(\xi) = 4\xi^2 - 2 . \quad (160)$$

Thus, the index n corresponds to the number of nodes in the meridional velocity profile in the domain $|y| < \infty$. Inserting the solution 159 into Eq. 158 leads to one of the defining differential equations for Hermite polynomials. In general, the three solutions of Eq. 156 can be interpreted as eastward- and westward-moving equatorially trapped gravity waves and westward-moving equatorial Rossby waves. The case $n = 0$ (for which the meridional velocity perturbation has a gaussian distribution centered at the equator) must be treated separately. In this case the *dispersion relationship* 156 (which is something like a characteristic equation that gives the $\nu(k)$ -dependence from which we may derive the phase velocities) factors as

$$\left(\frac{\nu}{\sqrt{gh_e}} - \frac{\beta}{\nu} - k \right) \left(\frac{\nu}{\sqrt{gh_e}} + k \right) = 0 . \quad (161)$$

The root $\nu/k = -\sqrt{gh_e}$, corresponding to a westward-propagating gravity wave, is not permitted since the second term in parentheses in Eq. 161 was explicitly assumed not to vanish when Eqs. 153 and 154 were combined to eliminate Φ . The roots given by the first term in parentheses in Eq. 161 are

$$\nu = k\sqrt{gh_e} \left[\frac{1}{2} \pm \frac{1}{2} \left(1 + \frac{4\beta}{k^2\sqrt{gh_e}} \right)^{1/2} \right] . \quad (162)$$

The positive root corresponds to an eastward-propagating equatorial inertio-gravity wave, while the negative root corresponds to a westward-propagating wave, which resembles an inertio-gravity wave for long zonal scale $k \rightarrow 0$ and resembles a Rossby wave for zonal scales characteristic of synoptic-scale disturbances. This mode is generally referred to as a Rossby-gravity wave.

5.2.2 Equatorial Kelvin Waves

In addition to the modes discussed in the previous section, there is another equatorial wave that is of great practical importance. For this mode, which is called the equatorial *Kelvin wave*, the meridional velocity perturbation vanishes and Eqs. 150 to 152 are reduced to the simpler set

$$-i\nu\hat{u} = -ik\hat{\Phi} \quad (163)$$

$$\beta y\hat{u} = -\frac{\partial\hat{\Phi}}{\partial y} \quad (164)$$

$$-i\nu\hat{\Phi} = -gh_e(ik\hat{u}) . \quad (165)$$

Eliminating Φ between Eq. 163 and Eq. 165, we see that the Kelvin wave dispersion equation is that of the shallow-water gravity wave

$$c^2 = \left(\frac{\nu}{k}\right)^2 = gh_e . \quad (166)$$

According to Eq. 166 the phase speed c can be either positive or negative. But, if Eq. 163 and Eq. 164 are combined to eliminate Φ we obtain a first-order equation for determining the meridional structure

$$\beta y\hat{u} = -c\frac{\partial\hat{u}}{\partial y} , \quad (167)$$

which may be integrated immediately to yield

$$\hat{u} = u_0 e^{-\beta y^2/(2c)} , \quad (168)$$

where u_0 is the amplitude of the perturbation zonal velocity at the equator. Equation 168 shows that if solutions decaying away from the equator are to exist, the phase speed must be positive ($c > 0$). Thus Kelvin waves are eastward propagating and have zonal velocity and geopotential perturbations that vary in latitude as Gaussian functions centered on the equator. The e-folding decay width is given by

$$Y_K = |2c/\beta|^{1/2} , \quad (169)$$

which for a phase speed $c = 30 \text{ m s}^{-1}$ gives $Y_K = 1600 \text{ km}$. The meridional force balance for the Kelvin mode is an exact geostrophic balance between the zonal velocity and the meridional pressure gradient. It is the change in sign of the Coriolis parameter at the equator that permits this special type of equatorial mode to exist.

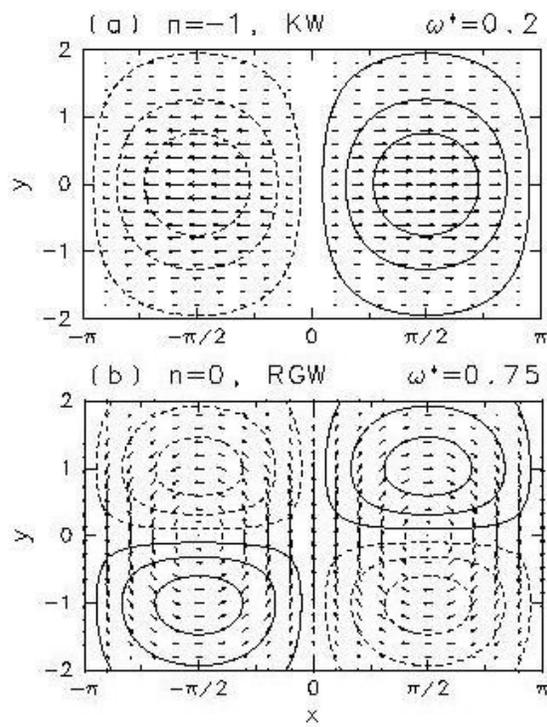


Fig.1 Horizontal structure of (a) Kelvin wave and (b) mixed-Rossby gravity wave. Contours show geopotential height component and arrows show horizontal wind components.

Figure 13: Illustration of Kelvin (upper panel) and Rossby-gravity (lower panel) waves.

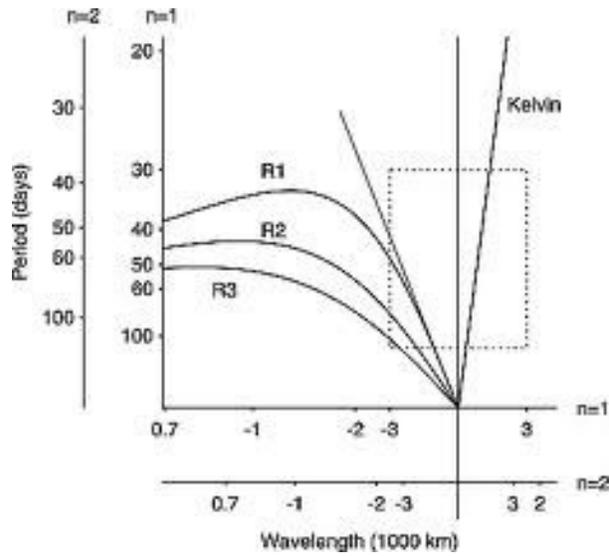


Figure 14: Dispersion diagram for equatorial Rossby-gravity and Kelvin waves.

6 ENSO atmosphere and ocean feedback mechanisms

Suggested Literature:

- a) Gill, A. E., 1980: 'Some simple solutions for heat-induced tropical circulation', *Quart. J. R. Met. Soc.*, **106**, 447-462.

- b) Reference for reduced gravity model in oceans:
'Ocean Circulation' by R. X. Huang, *University Press, Cambridge*, 791 pp.

- c) Online textbooks:
oceanworld.tamu.edu/resources/ocng_textbook/contents.html
oceanworld.tamu.edu/resources/oceanography-book/contents.htm

- d) 'El Nino, La Nina, and the Southern Oscillation' by S. G. Philander, *Academic Press, INC.*, 293 pp.

- e) For the ocean simulations presented:
Chang, P., 1994: 'A study of the seasonal cycle of sea surface temperature in the tropical Pacific Ocean using reduced gravity models', *J. Geophys. Res.*, **99**, C4, 7725-7741

6.1 Some general observations

Lets consider the sea surface temperature anomaly on 26th November, 2015 (Fig. 15). It's a manifestation of an El Nino development in the year 2015. Let's try to understand what is going on in the Pacific region.

The typical equatorial Pacific background state is shown in Fig. 16, in a vertical-zonal section. As can be seen it gets colder as we go from the surface downward (why?). And it also gets colder if we go from west to east. Looking at a map of sea surface temperature (SST), we see this equatorial asymmetry clearly (Fig. 17).

A schematic what is the situation in normal conditions (or La Nina) is given in 18. Trade winds are blowing near the equator from east to west (the reason theses trade winds will be discussed later in this course, but briefly it is due to the maximum convective heating that occurs in the mean around the equator). These trade winds push the warm surface waters to the west. What is happening then in the east?

However, there is an additional *Bjerknes* feedback mechanism working at the equator, which even further strengthens the equatorial easterly winds as shown in Fig. 19.

In El Nino conditions (Fig. 21) this normal situation breaks down and we get to a situation where also the eastern Pacific is flushed with warm waters (intuitively one would think that this is the 'normal' situation).

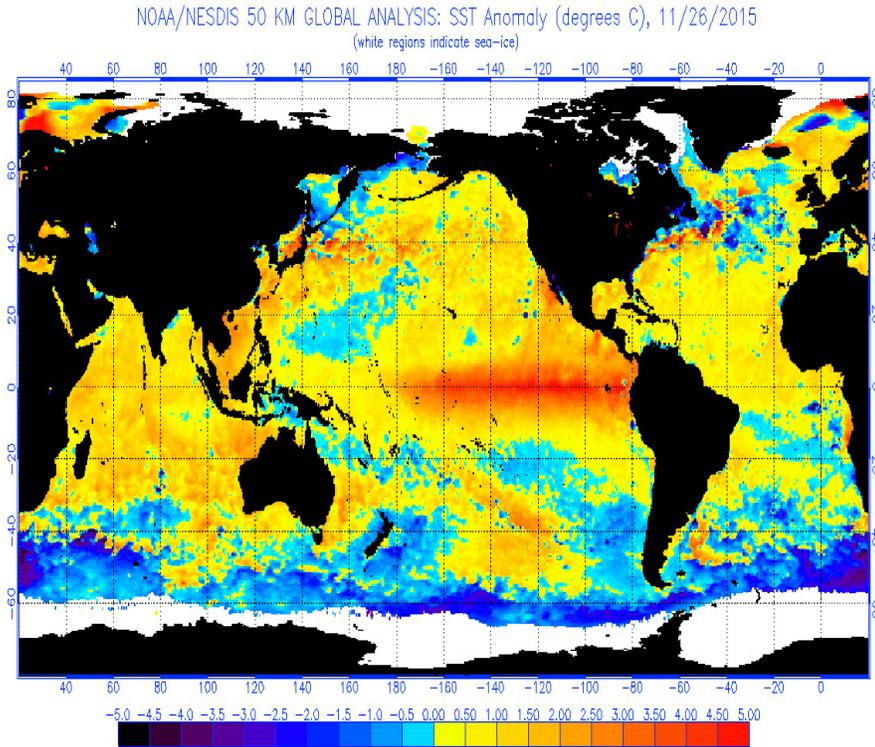


Figure 15: SST anomaly on 26. Nov 2015. Source: <http://www.ospo.noaa.gov/data/sst/anomaly/2015/anomnight.11.26.2015.gif>.

6.2 Atmospheric (ENSO) Teleconnections

In winter 2015/2016 one of the largest El Nino event in recent history happened (Fig. 22).

Such an ENSO event is considered to be the major source of seasonal predictability due to ENSO teleconnections Fig. 23.

In this section the physical basis for atmospheric teleconnections induced by tropical SST anomalies will be briefly discussed. How can we understand the origin of these teleconnections? Assuming we have an area warm SST anomalies. How do we assess what are the atmospheric adjustment processes. Surely a warm area of SST's will modify the stability of a parcel (see Fig. 24), but how exactly?

Note that the parcels that are considered here are much smaller in scale (1km) compared to the large-scale flow adjustments that we want to understand eventually. From the Atmospheric Physics course in the last term you know that for a moist, saturated parcel to be unstable the following criterion has to be fulfilled:

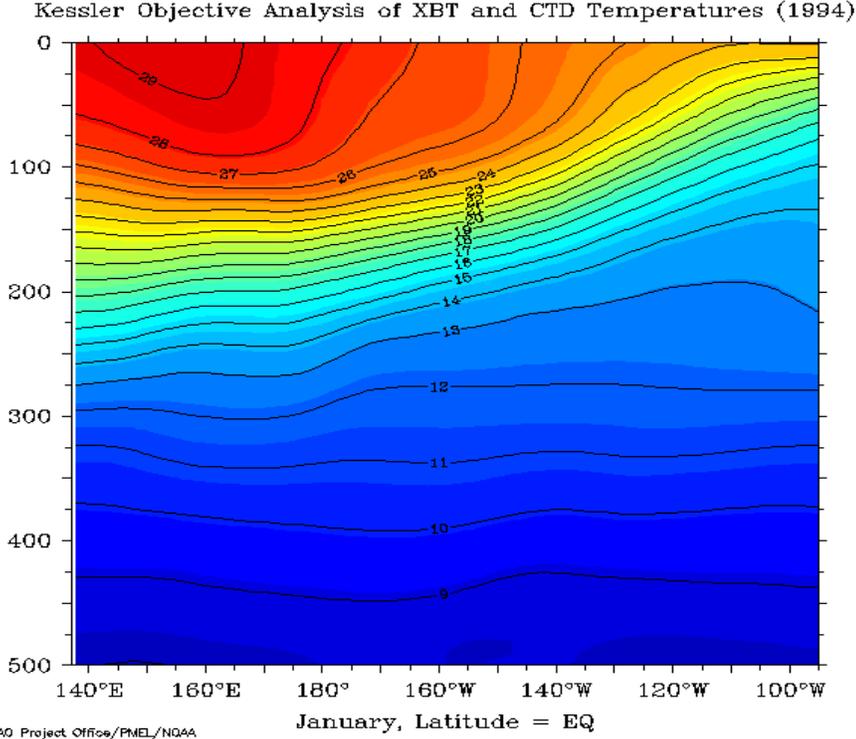


Figure 16: Average January temperature as a function of depths and longitude along the equator. Map from Pacific Marine Environmental Laboratory of NOAA, Seattle.

$$\frac{\partial \bar{T}}{\partial z} \leq -\frac{g}{c_p} \left(1 + \frac{L_{lv} m_{vs}}{RT} \right), \quad (170)$$

where quantities refer to the large-scale environment. As long as the parcel is not saturated the stability condition governed by the dry adiabat by setting m_{vs} to zero. This means that a warm surface would favour unstable conditions (Fig. 25).

For the moist saturated adiabatically rising parcel we have

$$\frac{dT}{dt} = -w \frac{g}{c_p} \left(1 + \frac{L_{lv} m_{vs}}{RT} \right). \quad (171)$$

Again, we can recover the unsaturated case by setting m_{vs} to zero, and this leads to the dry adiabatic lapse rate. This means the parcel always cools! It can be shown (look at your Atmospheric Physics course) that due to the decreasing temperature the Relative Humidity (RH) of a rising, unsaturated parcel always increases until saturation is reached

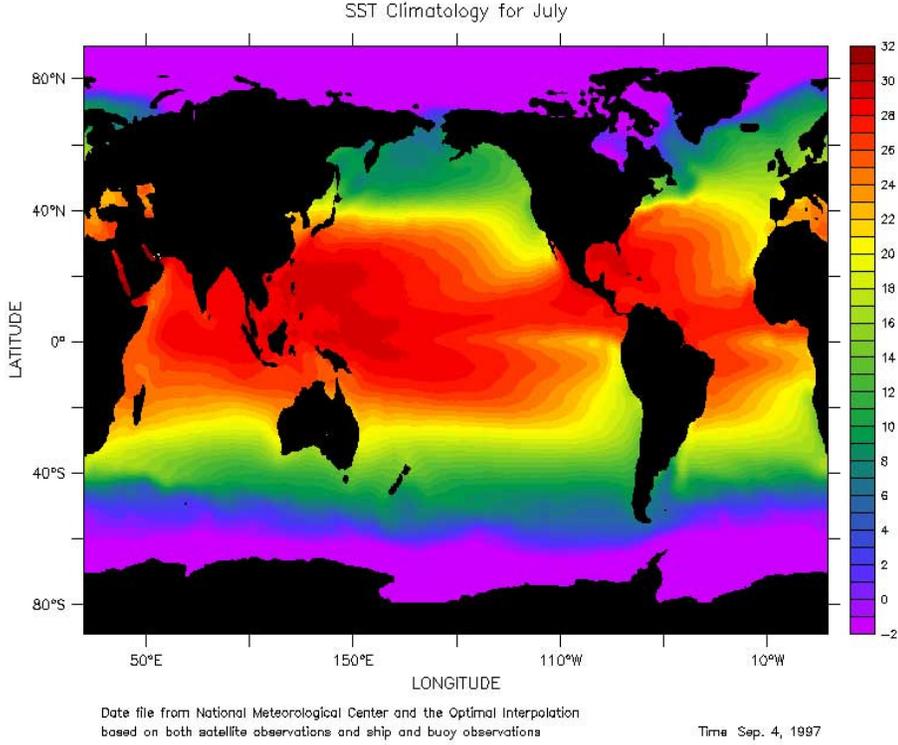


Figure 17: Mean sea-surface temperature calculated from the optimal interpolation technique (Reynolds and Smith, 1995) using ship reports and AVHRR measurements of temperature.

$$\frac{dRH}{dt} = -\frac{m_v}{m_{vs}^2} \frac{dm_{vs}}{dt} \geq 0 \quad , \quad (172)$$

If the parcel is rising further then condensation will occur. If the parcel raises far enough then nearly all its moisture may condensate and fall out as rain. This leads to diabatic heating of the atmosphere. In an unstable environment this leads to a huge asymmetry between rising and sinking parcels, because the sinking parcel may just evaporate the little moisture content in it. Note that if no rainfall is occurring, then the process of rising (condensation) and sinking (evaporation) would be reversible (no Entropy generation). It is the net heating related to the falling rain that makes the convection process irreversible. See also Foehn effect mechanism (Fig. 26).

As a net effect there is a large-scale tropospheric convective heating due to the SST-induced unstable conditions. From the thermodynamic equation (e.g. 47), we know that this will lead to further adjustments. In particular, for tropical regions we may assume that horizontal and time derivatives are negligible. We get the

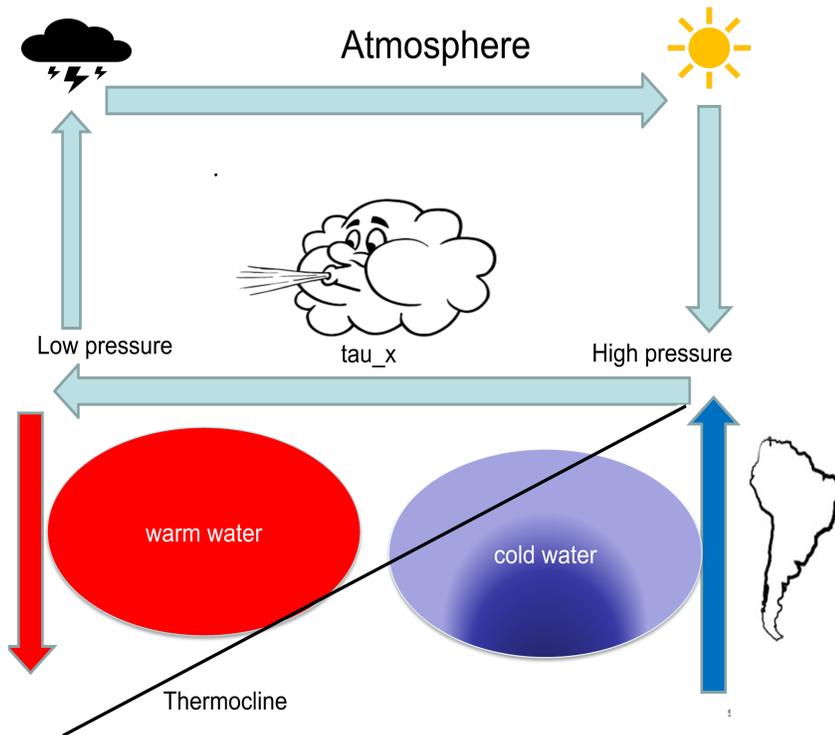


Figure 18: Normal conditions in the Pacific with strong trade winds pushing surface water toward the west and heavy rain in the west driving the atmospheric circulation.

approximate relations

$$\omega \approx -\rho g w \approx -\frac{Q}{c_p} \frac{1}{S_p} \quad (173)$$

or

$$w \approx \frac{Q}{c_p} \frac{1}{S_p \rho g} \quad (174)$$

This relation may also be applied to perturbations of Q and w .

$$\Delta w \approx \frac{\Delta Q}{c_p} \frac{1}{S_p \rho g} \quad (175)$$

if we assume that perturbations of S_p and ρ are small. It follows that increased heating in the region of warm SST anomalies has to be compensated by rising motion. Further adjustments may be explained by continuity as illustrated in Fig. 27.

For ENSO this could look like in Fig. 28. The signal in the upper troposphere may be spread through equatorial Rossby and Kelvin waves, and the resulting upper-level divergence may act as a Rossby-wave source in the vorticity equation 11.

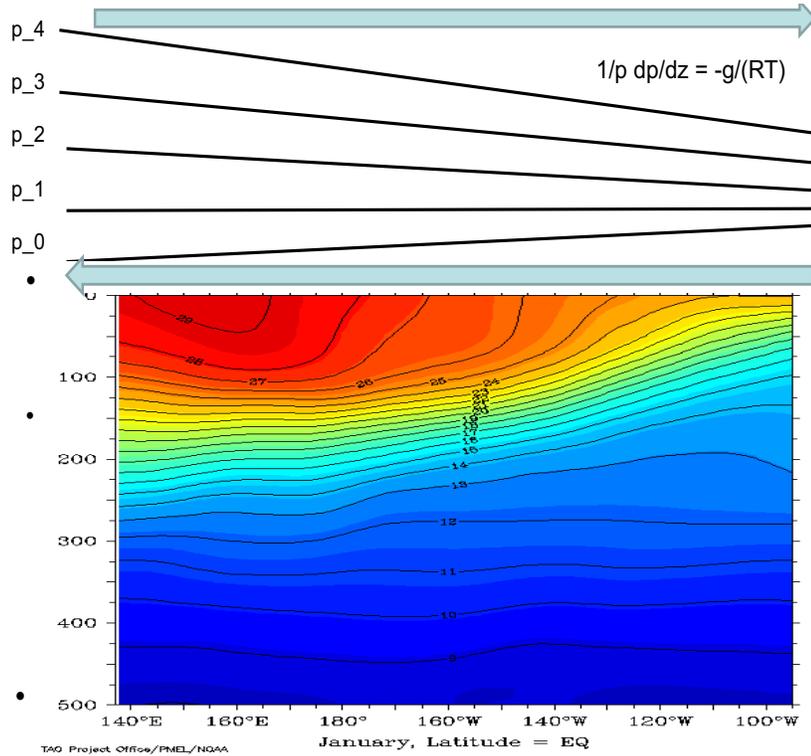


Figure 19: Cold conditions in east and warm conditions in west increase trade winds further due to the hydrostatic equation.

The typical sea surface temperature anomaly for an El Nino event is shown in Fig. 29. A warming of typically more than 1K is occurring in the eastern Pacific and surrounded by a cooling (so-called *Horse-shoe pattern*).

Fig. 30 shows the typical (composite) response of the atmosphere (rainfall [or heating!], and low-level winds) to the typical (composite) El Nino SST anomaly of Fig. 29.

The response in the central equatorial Pacific is a *weakening* of the trade winds, which is the positive atmospheric feedback, because a initial warm anomaly in the eastern Pacific will cause a response that is strengthening the original SST anomaly (why?). What is the typical period of ENSO? We will try to understand this local atmospheric response and the subsequent ocean response in the following subsections from a more theoretical point of view.

6.3 Atmospheric response to SST or heating anomaly

Kelvin and equatorial Rossby-gravity waves are also relevant for shaping the stationary response to an equatorial heating, so-called Gill response (Gill, 1980). In

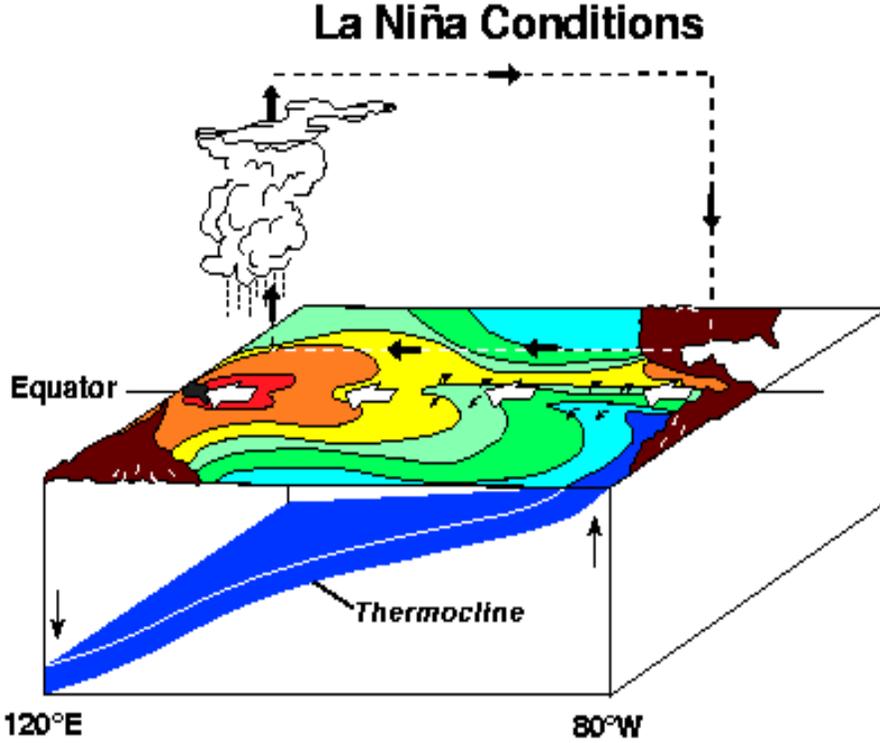


Figure 20: La Niña conditions in the Pacific with strong trade winds (black arrows) pushing surface water toward the west (white arrows) and heavy rain in the west driving the atmospheric circulation (black arrows). Colors give temperature of the ocean surface, red is hottest, blue is coldest. From: NOAA Pacific Marine Environmental Laboratory.

the Gill model a simple parameterization of the effect of heating on divergence, Q , is added to the continuity equation (in the Gill's paper and references therein proper justifications for this approach are presented)

$$\frac{\partial u'}{\partial t} = -\frac{\partial \Phi'}{\partial x} + \beta y v' \quad (176)$$

$$\frac{\partial v'}{\partial t} = -\frac{\partial \Phi'}{\partial y} - \beta y u' \quad (177)$$

$$\frac{\partial \Phi'}{\partial t} = -g h_e \left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} \right) - Q, \quad (178)$$

Finally, rayleigh friction and Newtonian cooling are added to the equations by replacing the time derivatives $\partial/\partial t$ by $\partial/\partial t + \epsilon$, and the equations are solved for a stationary state. The result of the (complicated) computations are illustrated in Fig. 31. To the west of the heating, we find the Rossby-gravity wave-type response,

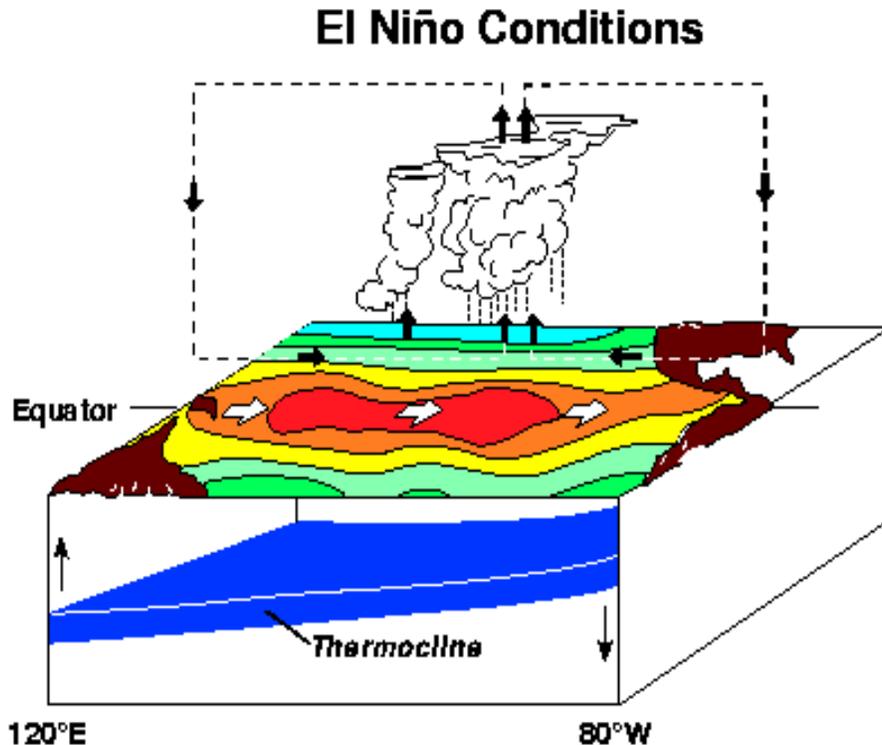


Figure 21: El Niño conditions in the Pacific with weak or reversed trade winds in the west (black arrows) allowing surface water to surge eastward (white arrows) and with heavy rain in the central equatorial Pacific driving the atmospheric circulation (black arrows). Colors give temperature of the ocean surface, red is hottest, blue is coldest. From: NOAA Pacific Marine Environmental Laboratory.

to the east we find the Kelvin wave-type response. Note that the Gill model is used, for example, to explain the atmospheric part of the positive (Bjerknes) feedback that leads to the ENSO phenomenon. Also note that a positive 'heating' Q explains the low-level convergent motions, but not the upper level ones (one would have to change sign to explain the upper-level response; upper-level divergence in case of a positive heating! Discuss!).

6.4 Response in the equatorial Ocean

6.4.1 Reduced gravity model

In this subsection we want to derive the equatorial ocean response to an atmospheric forcing of the Gill-type. If we consider the mean stratification of the equatorial Pacific Ocean of Fig. 16, one sees that warm water is residing on top of colder waters, divided by the thermocline. The simplest model of the upper Pacific are assuming therefore that there are 2 layers, divided by a density jump. Let layer 2 be the lower layer of density ρ_2 and height h_2 , and the upper layer be of density ρ_1 and

height h_1 . The aim here is to derive the pressure gradient force in the upper layer in this situation. We derive the pressure gradient force by integrating the hydrostatic equation 134 first in layer 2 from an arbitrary vertical position to the boundary h_2

$$\int_z^{h_2(x,y,t)} \frac{\partial p_2}{\partial z} dz = - \int_z^{h_2(x,y,t)} \rho_2 g dz , \quad \text{or} \quad (179)$$

$$p_2(h_2) - p_2(x, y, z, t) = -\rho_2 g [h_2(x, y, t) - z] . \quad (180)$$

Then we apply the horizontal gradient

$$\nabla p_2(h_2) = \nabla p_2(x, y, t) - \rho_2 g \nabla h_2(x, y, t) , \quad (181)$$

noting that the horizontal pressure gradient is independent on the vertical position.

Next we integrate further in layer 1

$$\int_{h_2(x,y,t)}^z \frac{\partial p_1}{\partial z} dz = - \int_{h_2(x,y,t)}^z \rho_1 g dz , \quad \text{or} \quad (182)$$

$$p_1(x, y, z, t) - p_1(h_2) = -\rho_1 g [z - h_2(x, y, t)] . \quad (183)$$

We again apply the horizontal gradient

$$\nabla p_1(x, y, t) = \nabla p_1(h_2) + \rho_1 g \nabla h_2(x, y, t) , \quad (184)$$

Continuity demands that $p_1(h_2) = p_2(h_2)$, therefore inserting 184 into 181 leads to:

$$\nabla p_1(x, y, t) = \nabla p_2(x, y, t) - (\rho_2 - \rho_1) g \nabla h_2(x, y, t) . \quad (185)$$

Assuming the lower layer motionless and without pressure gradient $\nabla p_2(x, y, t) = 0$ and $H = h_1 + h_2 = \text{const}$ (rigid lid; an approximation here), and therefore $\nabla h_1 = -\nabla h_2$ we get

$$\nabla p_1(x, y, t) = (\rho_2 - \rho_1) g \nabla h_1(x, y, t) , \quad (186)$$

or

$$\frac{1}{\rho_1} \nabla p_1(x, y, t) = \frac{\rho_2 - \rho_1}{\rho_1} g \nabla h_1(x, y, t) . \quad (187)$$

The pressure gradient in the upper layer can be, to a first approximation, expressed in terms of the change in density between lower and upper layer and the gradient of the *thermocline depth* h_1 . Since with respect the standard one-layer case the factor g is replaced by $(\rho_2 - \rho_1)/\rho_1 g = g'$, this model is called *reduced gravity model*. The density change $(\rho_2 - \rho_1)/\rho_1$ is typically about 1%. The *reduced gravity model* is identical to the shallow water equations, but with the pressure gradient force 187, because if the slight change of density between the 2 layers. We add a wind-stress forcing on the rhs of the following equations to mimic the ocean forced case. These equations are similar to the atmospheric Gill model, but with forcing in the momentum equations instead of in the continuity equation.

$$\frac{\partial u}{\partial t} + (\mathbf{u} \cdot \nabla)u = -g' \frac{\partial h}{\partial x} + fv + \frac{1}{\rho h} \tau_x \quad (188)$$

$$\frac{\partial v}{\partial t} + (\mathbf{u} \cdot \nabla)v = -g' \frac{\partial h}{\partial y} - fu + \frac{1}{\rho h} \tau_y \quad (189)$$

$$\frac{\partial h}{\partial t} + \mathbf{u} \cdot \nabla h = -h \nabla \cdot \mathbf{u} \quad (190)$$

We have used h instead of h_1 for simplicity as the *thermocline depth*, and ρ instead of ρ_1 as density in the upper layer. It is quite instructive to consider the motionless stationary state of such a model (even in a more complex situation this may be a good approximation):

$$\frac{\partial h}{\partial x} = \frac{1}{\rho h g'} \tau_x \quad (191)$$

$$\frac{\partial h}{\partial y} = \frac{1}{\rho h g'} \tau_y \quad , \quad (192)$$

assuming a typical situation in the equatorial Pacific with purely easterly trade winds ($\tau_x < 0$, $\tau_y = 0$) it becomes clear why the equatorial Pacific thermocline is tilted from west to east!!!!

It is interesting to note that the (forcing) free Eq. (188) - (190) are formally identical to the shallow water equations 140, 141 and 145 that we used to derive the equatorial Rossby and Kelvin waves, if we replace g by g' , the reduced gravity. This means that the reduced gravity equations support the same solutions as the shallow water equations near the equator, if we also in all phase velocities replace g by g' . We will see these kind of waves in the following example.

In the following we present and discuss the ocean adjustment to a constant atmospheric wind-stress forcing of the ocean derived from a model similar, but slightly more complicated than Eq. 188 - 190. The addition is basically an equation for the surface temperature that is not present in Eq. 188 - 190. In a strict sense, the surface temperature would never be influenced by the thermocline depth as long as this is positive. However, in the real world, there is still a gradient present in the upper-layer, meaning that if the thermocline is nearer to the surface, the temperature is lower there, just as seen in Fig. 16. Furthermore, if temperature is allowed to vary horizontally then there will also be horizontal advection of temperature.

6.5 Ocean response to a zonal wind stress anomaly

Fig. 32 shows the wind-stress forcing applied with maximum at equator at 180 E, and a Gaussian shape of with 10 degrees in east-west and 5 degrees in north-south direction. The forcing is of magnitude 0.015 N/m^2 , which is a typical response of the atmosphere to a typical ENSO anomaly. This forcing is mimicking the atmospheric Gill-response on the equator to the ENSO-induced heating (SST) anomaly. We are considering the time evolution of the thermocline response to the constant forcing.

Fig. 33a shows the response after the first month of forcing, and we see interestingly that the response resembles the Gill response that we discussed for the atmosphere to a diabatic heating anomaly, with an equatorial Kelvin wave to the east of the wind stress forcing and 2 Rossby waves straddling the equator to the west. The Kelvin wave signal moves to the east as time evolves and the Rossby wave signal moves to the west (as they should). The Kelvin waves reach the eastern boundary after about 4 month (Fig. 33b) and appear to transform into coastal Kelvin waves that move north- and southward from the equator. The maximum response is reached after about 6 month, and is then decaying somewhat. The stationary response is seen in Fig. 33f. The thermocline tilt in the steady-state solution seems to balance the wind-stress forcing, showing that Eqs. 191 are a good approximation.

Fig. 34 shows the corresponding response in sea surface temperatures (using an additional equation not present in Eq. (188) - (190)). Remember that the atmospheric forcing was mimicking the Gill-type response to a warm SST anomaly. From Fig. 34 we clearly identify that the ocean response to the atmospheric forcing provides a positive feedback: The SST in the eastern equatorial Pacific is further increased. The SST response evolves slower than the thermocline response (because the response starts in the central/western Pacific where the thermocline is relatively deep in the mean state and its fluctuations are therefore less coupled to the surface. The maximum positive feedback seems to occur after about 6 to 12 months and it decays to reach an equilibrium. This decay could be interpreted as the delayed negative feedback provided by the ocean to the ENSO phenomenon.

Exercises

1. Assume a mean wind stress distribution along the equator:

$$\begin{aligned}
 \tau_x &= 0 & \text{for } \text{lon} \leq 170 \text{ E} \\
 \tau_x &= -0.06 \text{ N/m}^2 & \text{for } 170 \text{ E} \leq \text{lon} \leq 240 \text{ E} \\
 \tau_x &= 0 & \text{for } \text{lon} \geq 240 \text{ E}
 \end{aligned}
 \tag{193}$$

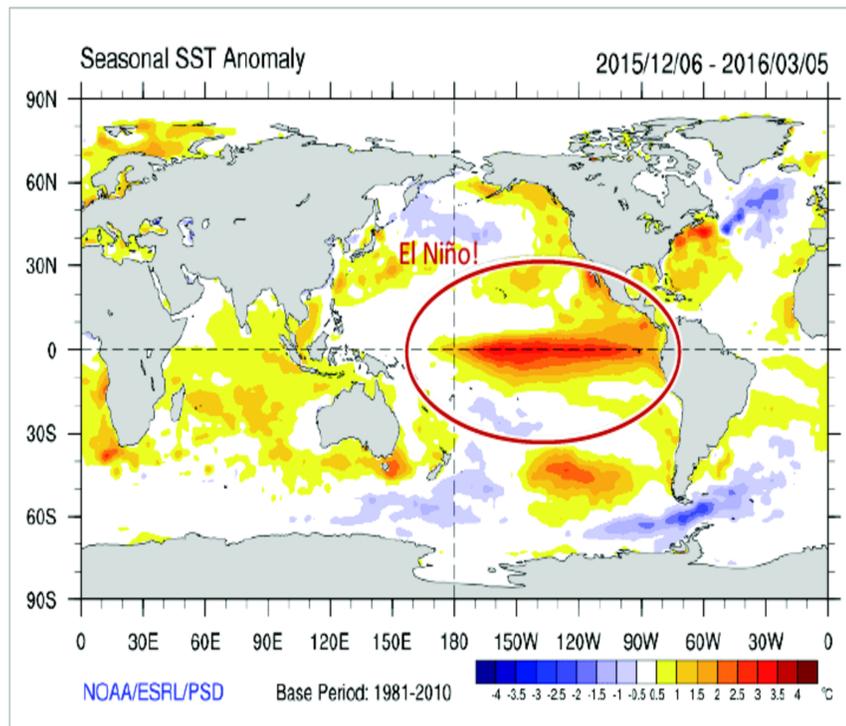
Using the approximation 191, calculate the thermocline distribution along the equator, assuming that the thermocline depth at the western edge is 150 m. What is the total change in height between 170 E and 240 E?

2. Calculate the oceanic Kelvin wave speed of (using the reduced gravity approximation) of an ocean with a mean thermocline depth of 100 m.

EL NIÑO 2015/2016

Perhaps the strongest event in the recent history?

Dec 2015 – Mar 2016 SST anomalies



7

This slide, as many in the following from Bianca Mezzina's Master thesis defence

Figure 22: Sea Surface Temperature Anomalies in winter 2015/2016.

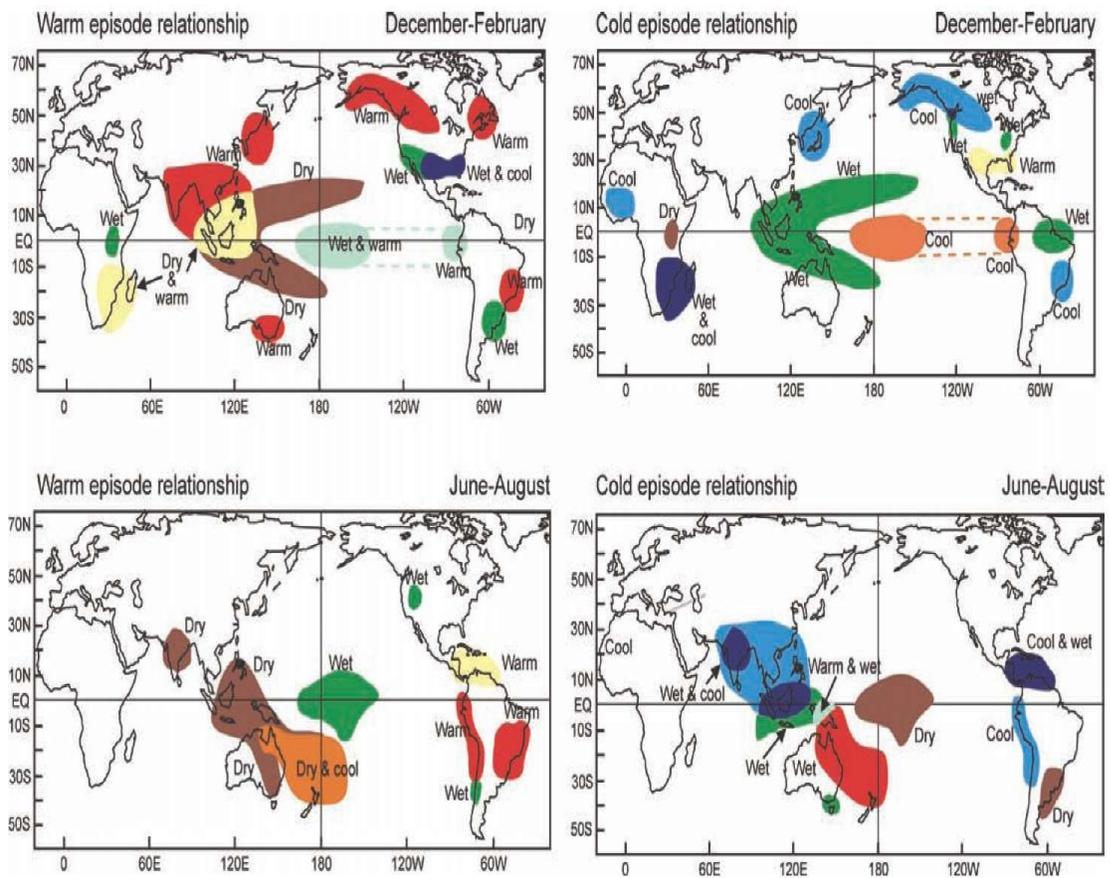
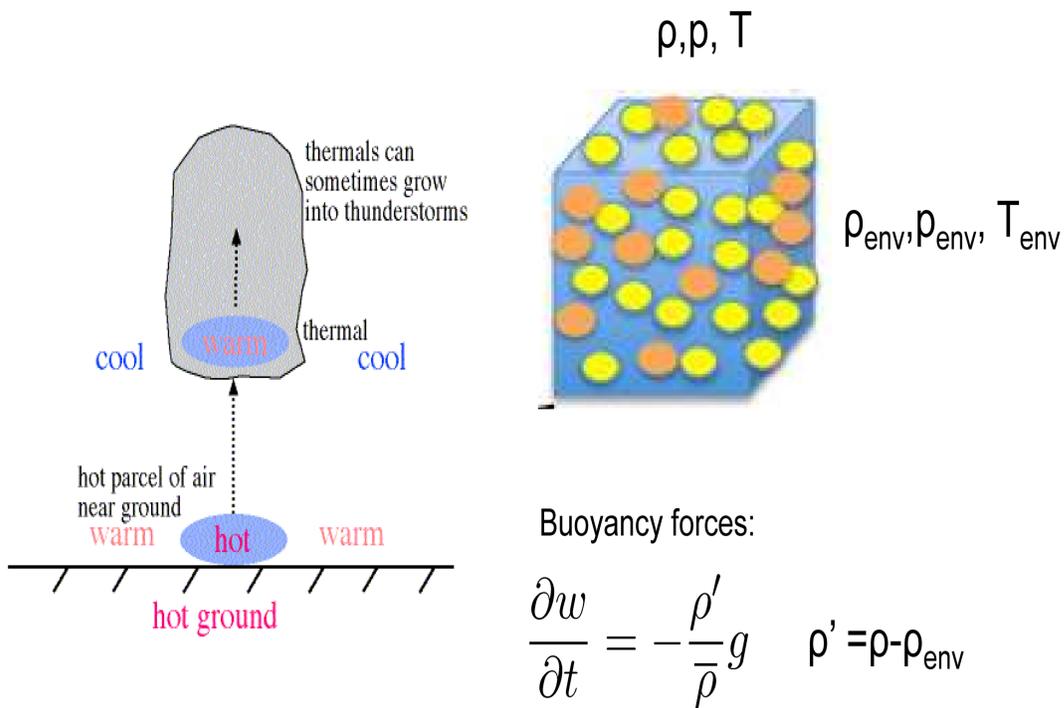


Figure 23: Some known ENSO teleconnections. Note that not all teleconnections are shown, only the most robust ones. Source: NOAA

Evaluate parcel properties with respect to its environment



Source: appollo.lsc.vsc.edu



Figure 24: Sketch of a parcel in an environment that is warmed from below.

Unstable condition – negative static stability

$$\Gamma_d < \Gamma_{env}$$

Γ_d , DALR (parcel T follows this profile)

z

$\Gamma_{env} = -dT_{env}/dz$
Atmosphere,
Environment

T_{env} T_{parcel} T



Figure 25: Sketch how the environment is favorable or unfavorable for convection and may be influenced by surface conditions.

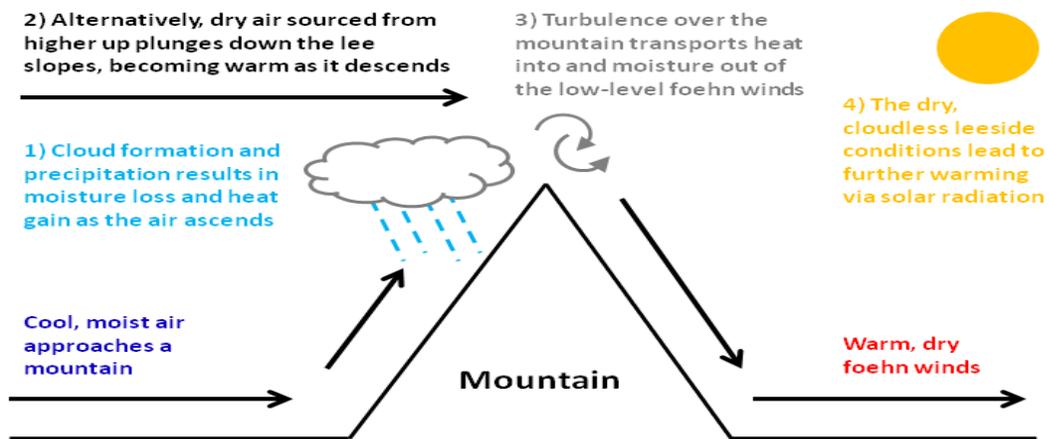


Figure 26: Foehn effect mechanism.

In Stratosphere: $\Delta w \approx \frac{\Delta Q}{c_p S_p \rho g}$ small, because $S_p = (\Gamma_d - \Gamma)/(\rho g)$ Large!
 $-dT/dz = \Gamma = 0$
 Tropopause ~ 10km

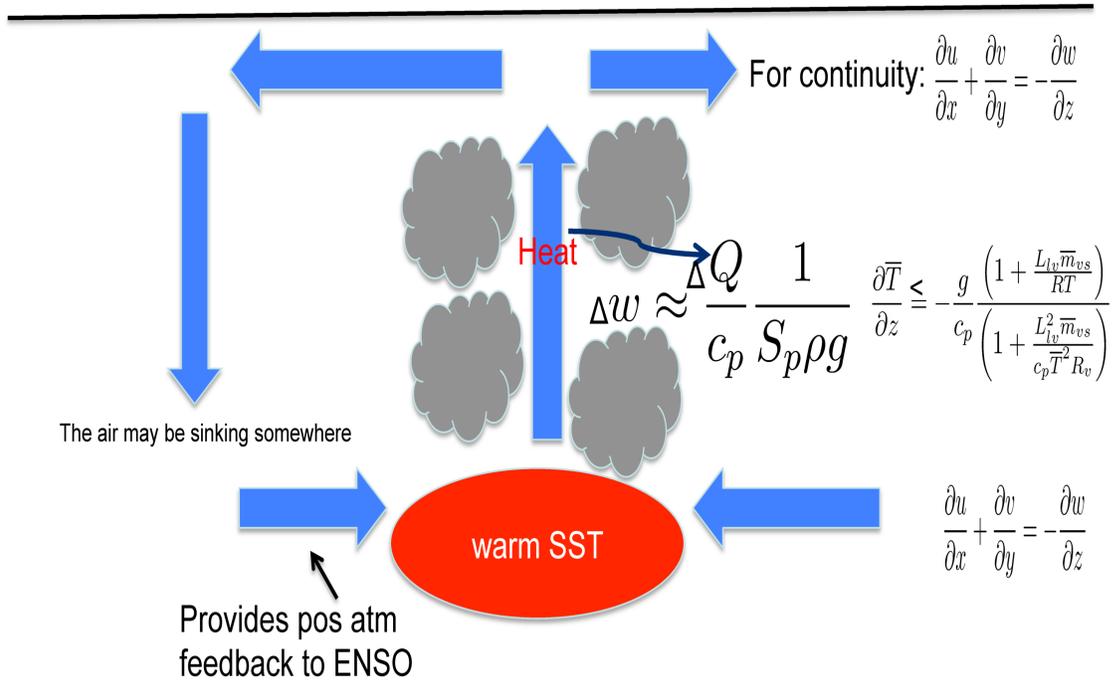
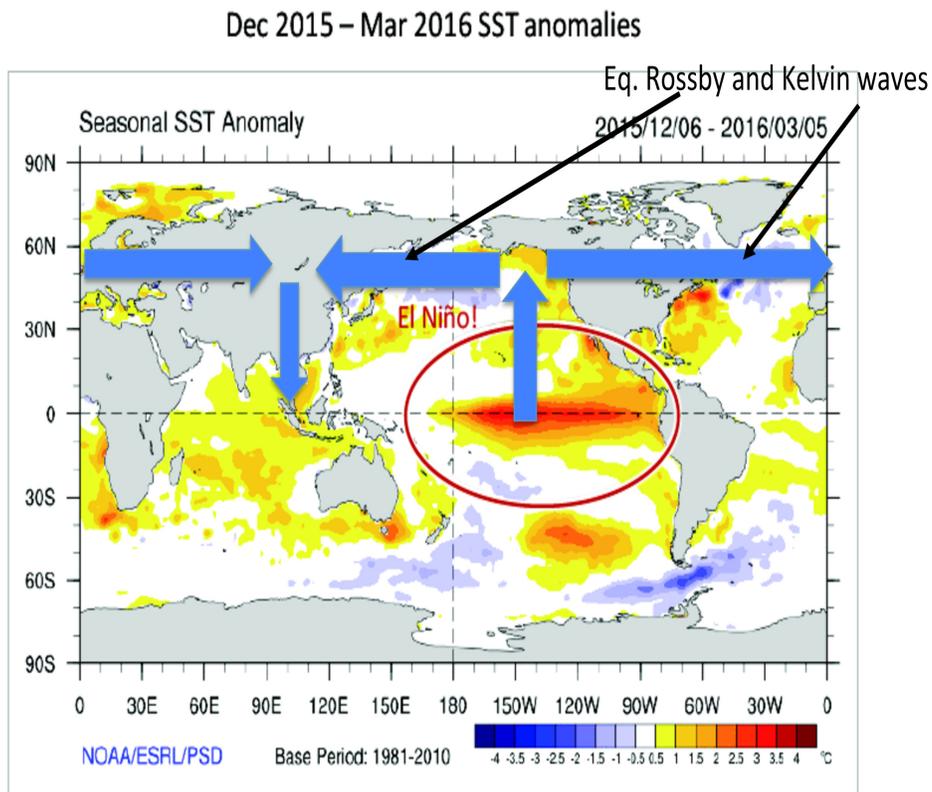


Figure 27: Large-scale adjustments to diabatic convective heating.

EL NIÑO 2015/2016



7

Figure 28: Large-scale adjustments to diabatic convective heating induced by ENSO.

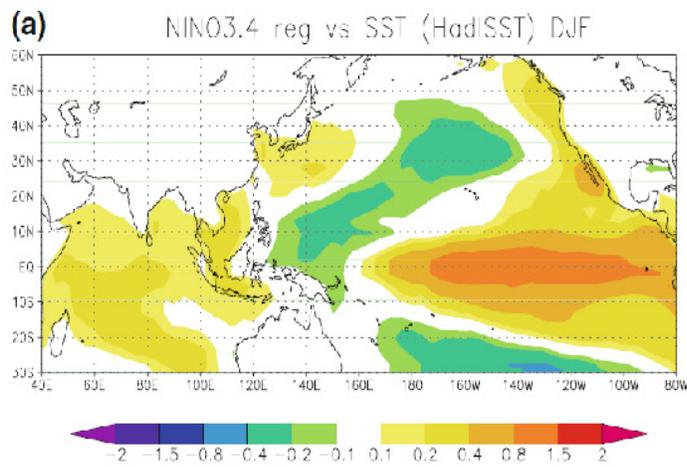


Figure 29: Composite sea surface temperature anomaly for an El Niño condition. Units are K.

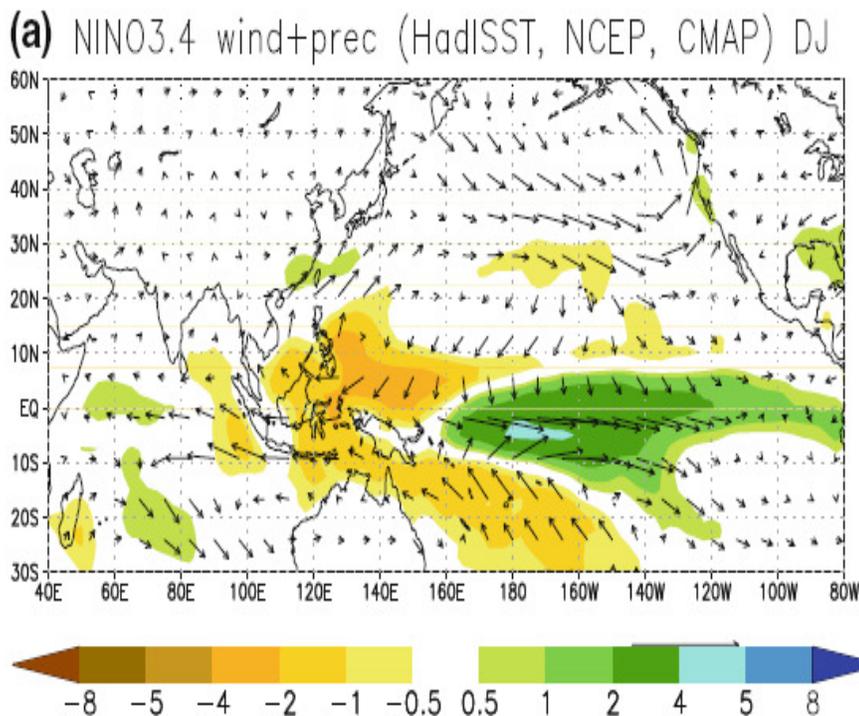


Figure 30: Composite of response to an El Niño forcing. Shading Precipitation in mm/day, vectors 925 hPa wind.

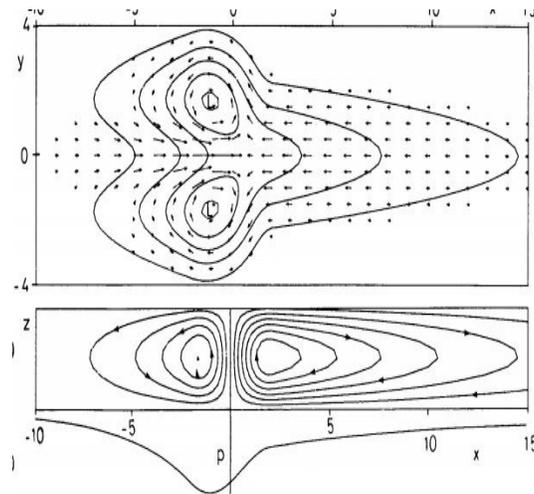


Figure 31: Illustration response to equatorial heating according to the gill model. Equatorial rossby waves shape the response to the west and Kelvin waves shape the response to the east

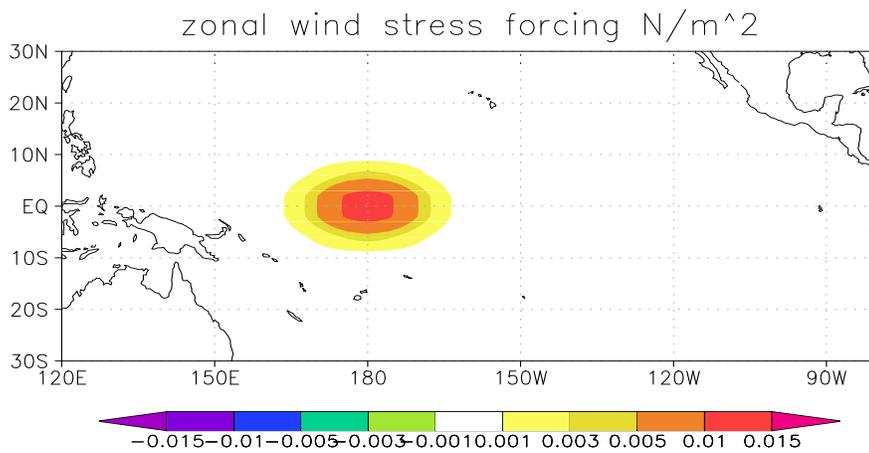


Figure 32: Zonal wind stress forcing. Units are N/m^2 .

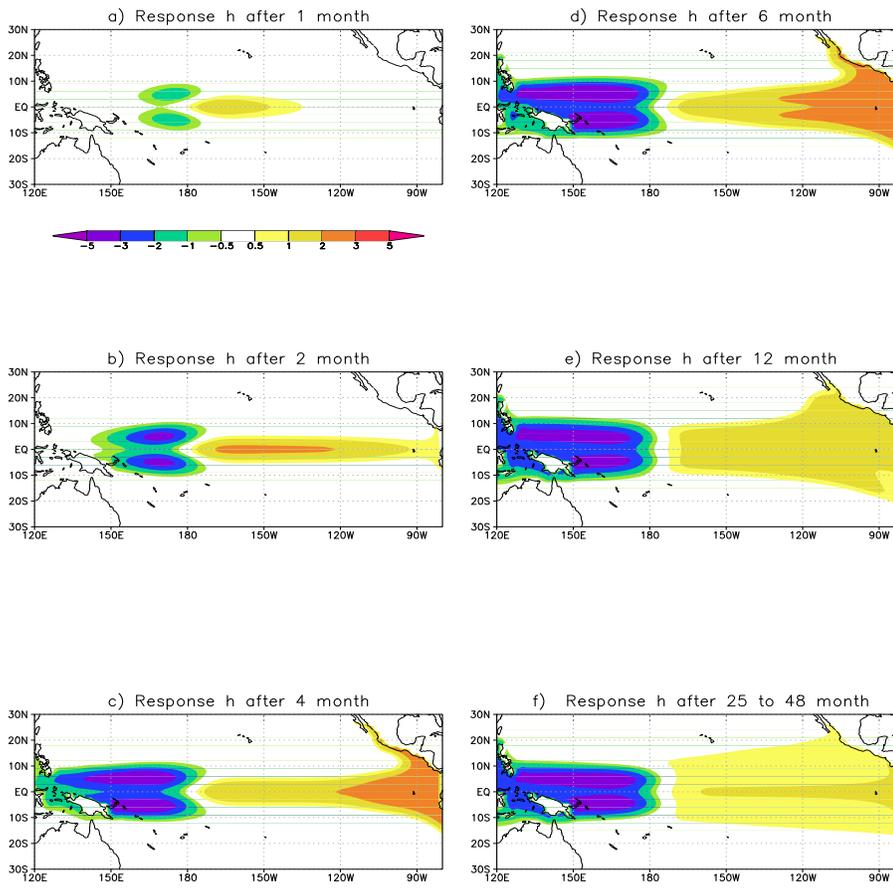


Figure 33: Thermocline response to anomalous zonal wind stress forcing of Fig. 32. a) after 1 month, b) after 2 months, c) after 4 months, d) after 6 months, e) after 12 months. f) stationary response. Units are m.

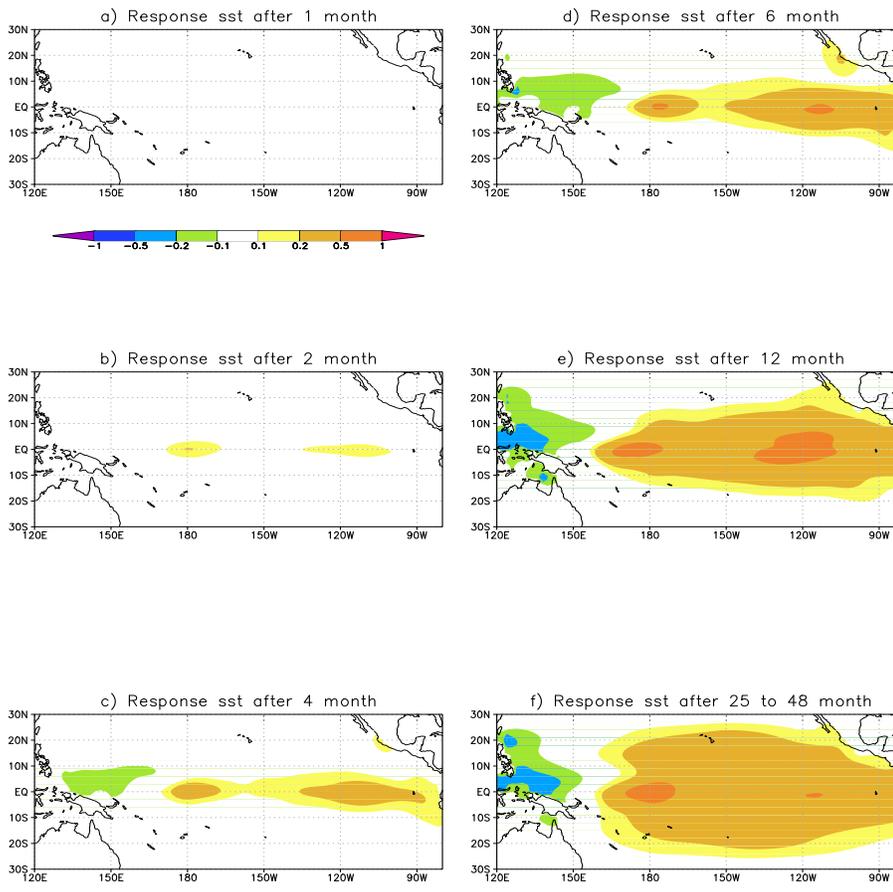


Figure 34: SST response to anomalous zonal wind stress forcing of Fig. 32. a) after 1 month, b) after 2 months, c) after 4 months, d) after 6 months, e) after 12 months. f) stationary response. Units are K.

7 (Atmospheric) Planetary Boundary Layer Processes

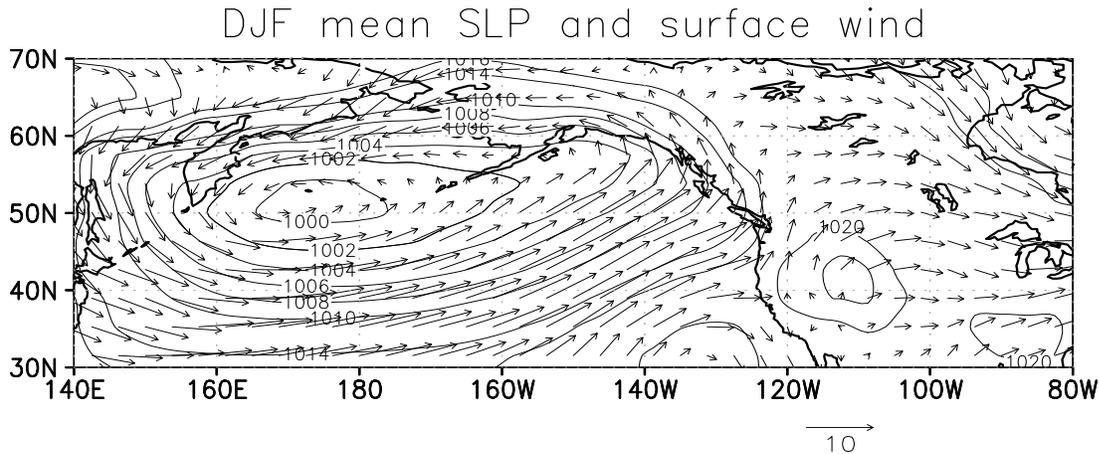


Figure 35: Mean DJF Sea Level Pressure (SLP) and surface winds in the North Pacific from observations (NCEP-NCAR re-analysis). Units are m/s for wind and hpa for SLP.

Figure 35 shows the winter mean (DJF) climatology of Sea Level Pressure (SLP) and surface winds in the North Pacific. As we expect for geostrophy, winds are mainly parallel to the isobars (lines of constant pressure). However, there seems to be a systematic tendency for a component of the winds towards the low pressure. In this section we will try to understand this systematic departure from geostrophy. Do you have a guess why this departure exists?

So far we have ignored the effect of friction on the flows. However, in order to understand climate dynamics it is important to consider the effects of friction, particularly the one provided by the *planetary boundary layer*, which covers roughly the lowest kilometer(s) of the atmosphere. The boundary layer frictional processes are ultimately induced by the molecular viscosity. However, this effect is only relevant in the few millimeters closest to the surface. In the largest part of the planetary boundary layer turbulent eddies take over the role of molecular friction. This is part of the *energy cascade*, meaning that in the large-scale flow ever smaller eddies are embedded that carry energy to ever smaller scales until finally molecular viscosity takes over. As for all sections, a whole lecture course could be devoted to this topic. Therefore we will concentrate on the features that are most relevant and essential to understand climate dynamics.

7.1 The Boussinesq Approximation

The density in the lowest part of the Atmosphere varies little (about 10 % of its mean value). The flow may be considered as essentially incompressible if only the

momentum equation is considered. Therefore, for simplicity we assume the density to be constant. Let us consider therefore horizontal momentum equations 132 and 133, in which ρ is considered to be a constant

$$\frac{du}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + fv \quad (194)$$

$$\frac{dv}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial y} - fu . \quad (195)$$

and the continuity equation for incompressible flows 142

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 . \quad (196)$$

This set of equations is equivalent to the barotropic shallow water equations introduced and used several times already. However, the *Boussinesq approximation* goes beyond this, because it also involves an approximation to the vertical momentum equation that is different from the hydrostatic equation and allows for buoyancy effects there (as has been used in the EST course of the last term). Here we do not need to consider this equation.

7.2 Reynolds Averaging

In order to simulate the effect of the smaller scale eddies on the larger scale (“resolved”) flow, is it useful to apply an averaging operator to the equations. The idea is that the total flow can be divided into a slow evolving large scale field and into small-scale eddy fluctuations

$$u = \bar{u} + u', \quad v = \bar{v} + v' . \quad (197)$$

Formally, the operator could be a temporal and/or spatial average. For the total derivative of a quantity A

$$\frac{dA}{dt} = \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \right) A + A \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) , \quad (198)$$

where we have added a zero according to the incompressibility condition 196. Therefore we may write the total derivative (for incompressible flow!) as

$$\frac{dA}{dt} = \left(\frac{\partial A}{\partial t} + \frac{\partial Au}{\partial x} + \frac{\partial Av}{\partial y} + \frac{\partial Aw}{\partial z} \right) , \quad (199)$$

Application of the averaging operator yields

$$\frac{d\bar{A}}{dt} = \left(\frac{\partial \bar{A}}{\partial t} + \frac{\partial(\bar{A}\bar{u} + \overline{A'u'})}{\partial x} + \frac{\partial(\bar{A}\bar{v} + \overline{A'v'})}{\partial y} + \frac{\partial(\bar{A}\bar{w} + \overline{A'w'})}{\partial z} \right) , \quad (200)$$

because

$$\overline{ab} = \overline{(\bar{a} + a')(\bar{b} + b')} = \overline{\bar{a}\bar{b}} + \overline{\bar{a}b'} + \overline{a'\bar{b}} + \overline{a'b'} = \overline{\bar{a}\bar{b}} + \overline{a'b'} ,$$

and $\overline{a'} = \overline{b'} = 0$. Therefore application of the averaging operator to Eqs. 194 and 195 yields

$$\frac{d\bar{u}}{dt} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x} + f\bar{v} - \frac{\partial \overline{u'u'}}{\partial x} - \frac{\partial \overline{u'v'}}{\partial y} - \frac{\partial \overline{u'w'}}{\partial z} \quad (201)$$

$$\frac{d\bar{v}}{dt} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial y} - f\bar{u} - \frac{\partial \overline{v'u'}}{\partial x} - \frac{\partial \overline{v'v'}}{\partial y} - \frac{\partial \overline{v'w'}}{\partial z} , \quad (202)$$

where

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} + \bar{v} \frac{\partial}{\partial y} + \bar{w} \frac{\partial}{\partial z} \quad (203)$$

is the rate of change following the large-scale (or resolved) flow. Applying the zonal average to the continuity equation leads to

$$\frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y} + \frac{\partial \bar{w}}{\partial z} = 0 \quad . \quad (204)$$

If we compare Eqs. 201 and 202 with 195 and 195 we see that extra terms emerge if we follow a particle with the average large-scale flow. These can be interpreted as the effects of the small-scale eddies on the large-scale flow, and are called *convergence of eddy momentum fluxes*. Obviously, in order to solve Eqs. 201 and 202 for the large-scale flow, the additional terms have to be *parameterized* in terms of mean flow properties. This is big topic in fluid dynamics, and is called *closure problem*. Note that a very similar problem occurs when we write down the Navier-Stokes equations for numerical models which intrinsically have a grid (and time) spacing. Turbulence theories give some clues how such parameterizations should look like. One of the simplest one is the *flux-gradient theory*, which states that the effect of small-scale eddies on the large-scale flow is similar to the effect of molecular viscosity on smaller scale flow. The effect of viscosity on the small scale flow is to bring the flow into equilibrium, that is to reduce contrasts or gradient. Also note that the geometry (horizontal surface) means that changes in the vertical direction are much larger than in the horizontal direction (horizontal homogeneity). We have parameterizations of the type:

$$\overline{u'w'} = -K_m \frac{\partial \bar{u}}{\partial z} \quad (205)$$

$$\overline{v'w'} = -K_m \frac{\partial \bar{v}}{\partial z} , \quad (206)$$

where K_m can be a function of the vertical coordinate z . All other momentum fluxes can be approximated to be close to zero in Eqs. 201 and 202.

7.3 The Ekman Layer

The Ekman layer is the layer that connects a layer very close to the surface to the free atmosphere where we have near geostrophic equilibrium. Using the geostrophic wind (note that we can consider for the current analysis $f = f_0 = \text{const}$)

$$u_g = -\frac{1}{f\rho} \frac{\partial p}{\partial y}, v_g = \frac{1}{f\rho} \frac{\partial p}{\partial x},$$

with this and Eqs. 205-206, the stationary (equilibrium) approximation to Eqs. 201 and 202 are

$$K_m \frac{\partial^2 u}{\partial z^2} + f(v - v_g) = 0 \quad (207)$$

$$K_m \frac{\partial^2 v}{\partial z^2} - f(u - u_g) = 0, \quad (208)$$

where we have dropped the overbar for average quantities for convenience (only average quantities appear). Note that also the mean vertical advection term has been dropped because of smallness compared to the other terms. The horizontal advection terms are dropped because of the horizontal homogeneity condition. If we assume $u_g = \text{const.}$, and $v_g = \text{const.}$ with height, then we can substitute $u_* = u - u_g$ and $v_* = v - v_g$. to get a system of the type

$$K_m \frac{\partial^2}{\partial z^2} \begin{pmatrix} u_* \\ v_* \end{pmatrix} + \begin{pmatrix} 0 & f \\ -f & 0 \end{pmatrix} \begin{pmatrix} u_* \\ v_* \end{pmatrix} = 0 \quad . \quad (209)$$

We assume a solution of the type $u_* = Ae^{imz}$, $v_* = Be^{imz}$, then it follows

$$\begin{pmatrix} -K_m m^2 & f \\ -f & -K_m m^2 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = 0 \quad . \quad (210)$$

Following basic algebra, non-trivial solutions of such a linear are found by setting the determinant of the 2x2 matrix to zero

$$K_m^2 m^4 + f^2 = 0. \quad (211)$$

The four solutions are for positive f (northern hemisphere; otherwise we have to use negative f for southern hemisphere)

$$m_1 = \sqrt{i} \sqrt{\frac{f}{k_m}}, m_2 = -\sqrt{i} \sqrt{\frac{f}{k_m}}, m_3 = \sqrt{-i} \sqrt{\frac{f}{k_m}}, m_4 = -\sqrt{-i} \sqrt{\frac{f}{k_m}}. \quad (212)$$

With $\sqrt{i} = (1+i)/\sqrt{2}$, we have

$$m_1 = (1+i) \sqrt{\frac{f}{2k_m}}, m_2 = -(1+i) \sqrt{\frac{f}{2k_m}}, m_3 = (i-1) \sqrt{\frac{f}{2k_m}}, m_4 = (1-i) \sqrt{\frac{f}{2k_m}}. \quad (213)$$

The boundary conditions are geostrophy ($u = u_g, v = v_g$) as z goes to infinity, therefore $u_* = v_* = 0$ and $u = v = 0$ or $v_* = -v_g, u_* = -u_g$ at $z=0$. The boundary condition as z goes to infinity excludes solutions that grow, therefore the solutions with negative i (m_2 and m_4) are excluded. If we insert the solutions m_1 and m_3 back into the original system 210, we can determine the two eigenvectors, which are

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

and

$$\mathbf{x}_3 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

The two eigenvectors are complex conjugate and therefore not independent. Therefore the solution simply given by m_1 ,

$$u_* = ae^{im_1z} = a[\cos(\gamma z) + i \sin(\gamma z)]e^{-\gamma z} \quad (214)$$

$$v_* = ia e^{im_1z} = ia[\cos(\gamma z) + i \sin(\gamma z)]e^{-\gamma z}, \quad (215)$$

where we have used $\gamma = \sqrt{f/(2K_m)}$. Let $a = b + ic$, then the real part of the solution is

$$u_* = b \cos(\gamma z)e^{-\gamma z} - c \sin(\gamma z)e^{-\gamma z} \quad (216)$$

$$v_* = -b \sin(\gamma z)e^{-\gamma z} - c \cos(\gamma z)e^{-\gamma z} \quad (217)$$

and with $z = 0$: $u_*(z = 0) = -u_g = b, v_*(z = 0) = -v_g = -c$ or

$$u = u_g - [u_g \cos(\gamma z) + v_g \sin(\gamma z)]e^{-\gamma z} \quad (218)$$

$$v = v_g + [u_g \sin(\gamma z) - v_g \cos(\gamma z)]e^{-\gamma z}, \quad (219)$$

The height of the boundary layer may be defined where the wind is for the first time parallel to the geostrophic wind, which is at $De = \pi/\gamma = \pi\sqrt{2K_m/f}$. We can use this formula to estimate the value of the eddy viscosity K_m . Observations of the mean boundary layer height in mid-latitudes give $De \approx 1$ km, therefore $K_m = 1/2f(De/\pi)^2 \approx 5 \text{ m}^2 \text{ s}^{-1}$. An important application of the Ekman solutions 218 and 219 is that we can calculate the vertical velocity at the top of the Ekman Layer induced by the action of turbulent eddies.

Let us calculate the divergence of the winds in the Ekman Layer

$$\begin{aligned} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= \left(\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} \right) (1 - \cos(\gamma z)e^{-\gamma z}) \\ &- \left(\frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} \right) \sin(\gamma z)e^{-\gamma z} \\ &= -\xi_g \sin(\gamma z)e^{-\gamma z}. \end{aligned} \quad (220)$$

This equation states that the divergence in the Ekman layer is proportional to the negative geostrophic vorticity, a very important effect of the boundary layer. Positive

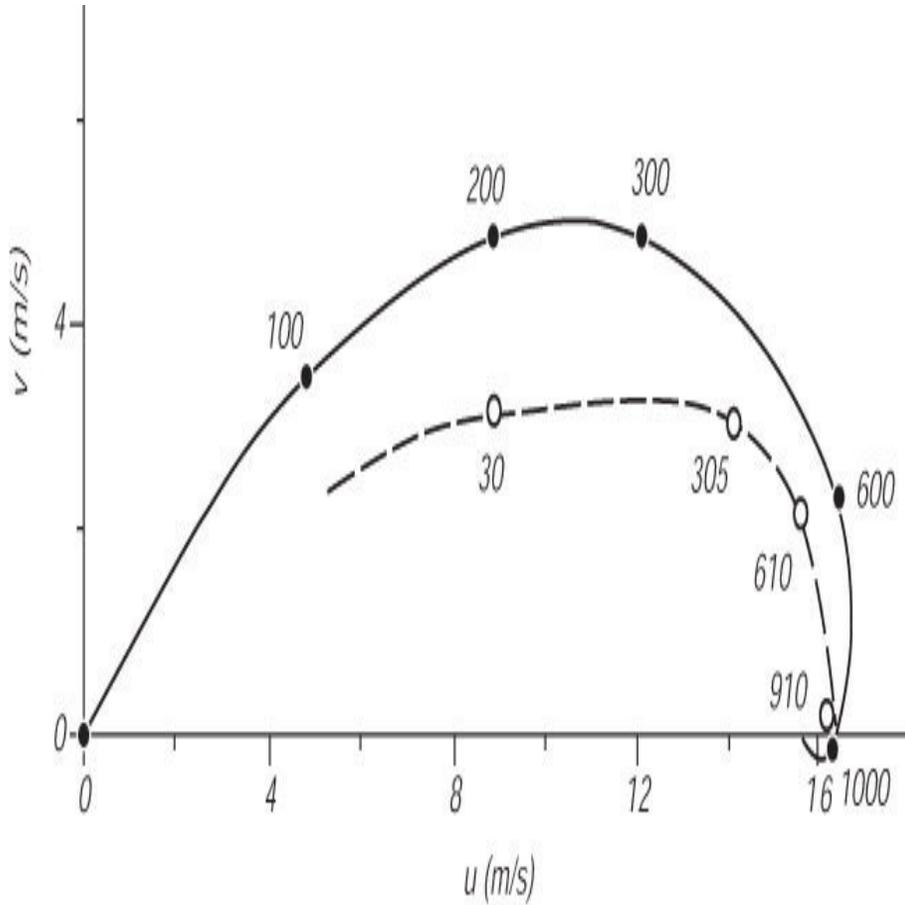


Figure 36: Idealized and observed Ekman Layer velocities Source: http://oceanworld.tamu.edu/resources/ocng_textbook/chapter09/chapter09_02.htm.

(cyclonic vorticity) leads to convergence! According to the continuity equation 196, this will lead to vertical motion, which is on top of the Ekman Layer

$$\begin{aligned}
 w(De) &= - \int_0^{De} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dz & (221) \\
 &= \xi_g \frac{e^{-\gamma z}}{2\gamma} [\sin(\gamma z) - \cos(\gamma z)] \Big|_0^{De} \\
 &= \frac{\xi_g}{2\gamma} (1 + e^{-\pi}) \\
 &\approx \frac{\xi_g}{2\gamma} = \xi_g \sqrt{\frac{K_m}{2f}},
 \end{aligned}$$

where we have assumed that the geostrophic wind is independent of height within the Ekman Layer. This is again an important result, a positive vorticity leads to upward

motion through Ekman effects on top of the boundary layer. This is called *boundary layer pumping* or *Ekman pumping*. It may be used to explain vertical motions and therefore rainfall anomalies induced by the Gill responses in tropical regions as derived in Section 6.3. It states that whenever we calculate a flow response that has (geostrophic) vorticity, this will lead to vertical motion and therefore a rainfall response. Given that geostrophy is valid from approximately 10 degrees away from the equator, this rule can be used for many flow responses. Remember that we have shown in chapter 5 that even the zonal winds in the close equatorial Kelvin waves are in exact geostrophic equilibrium. We can estimate the typical magnitude of the vertical velocity 221 by inserting $\xi_g = 10^{-5}\text{s}^{-1}$, $De = 1\text{km}$ or $\gamma = 3 \times 10^{-3} \text{ m}^{-1}$ to be $w(De) \approx 10^{-5}/(2 \times 3 \times 10^{-3}) \text{ m s}^{-1}$ or $2 \times 10^{-3} \text{ m s}^{-1}$. This a substantial vertical velocity, comparable to the one induced by a heating anomaly of about Q/c_p 1 k/day in the tropical regions, if we use equation 174 and $S_p \approx 5 \times 10^{-4} \text{ K Pa}^{-1}$ to estimate the vertical velocity:

$$w \approx \frac{Q}{c_p} \frac{1}{S_p \rho g} .$$

Also in the Ocean Ekman Layers exist (have you discussed them?). Clearly at the bottom of the ocean very similar processes take place as discussed here. Even at the top of the oceans we have an Ekman Layer (have you discussed this?). However, the main change is the boundary condition at the surface, which is given by the atmospheric winds that drive the ocean, in the interior the boundary condition can be assumed to be geostrophic again. Otherwise we can use the above derived methodology also to derive the ocean surface Ekman Layer.

Exercises

1. Verify that the Ekman solution 218 and 219 is indeed a solution of the original system of equations 207 and 208.
2. Calculate the scalar product between the pressure gradient and the wind within the Ekman layer given by Eqs. 218 and 219. Is the wind directed into or out of a low pressure system?
3. Write a fortran code that uses the Eqs. 218 and 219 and plot the solution as as phase space diagram (u,v) as in Fig 36. Also, solve the original equations 207 and 208 numerically by keeping the local time derivative in the Ekman equations:

$$\frac{\partial u}{\partial t} = K_m \frac{\partial^2 u}{\partial z^2} + f(v - v_g) \quad (222)$$

$$\frac{\partial v}{\partial t} = K_m \frac{\partial^2 v}{\partial z^2} - f(u - u_g) . \quad (223)$$

For both analytical and numerical solutions use $K_m = 5 \text{ m}^2 \text{ s}^{-1}$, the coriolis parameter at $45^\circ N$, $u_g = 10 \text{ m s}^{-1}$, $v_g = 0$. The vertical domain should be [0

m , 3000 m]. Use as initial condition $u = u_g, v = 0$. Compare the numerical stationary with the analytical solution. How long does it take for the solution to become approximately stationary?

8 Mechanisms for tropical rainfall responses to equatorial heating

More reading:

1. Hamouda, M. and Kucharski, F. (2019) Ekman pumping Mechanism driving Precipitation anomalies in Response to Equatorial Heating, *Clim Dyn*, DOI: 10.1007/s00382-018-4169-4

In section 6.2 we have discussed the basic large-scale circulation adjustment mechanisms to a localized SST anomaly. In this section mechanisms for regional rainfall responses to a localized equatorial sea surface temperature (SST)-induced heating anomaly will be discussed (e.g. ENSO-induced as shown in Fig. 23). It is mainly based on the paper Hamouda and Kucharski (2018) and references therein. This is an important topic as rainfall modifications due to, for example, ENSO can have substantial positive and negative effects and the understanding of the physical mechanisms for such responses is important to assess seasonal predictability. In the literature, several mechanisms have been proposed for rainfall responses to ENSO and other tropical SST-induced heating anomalies (e.g. see Fig. 23). These include:

- a) Destabilization in the region with strong SST anomalies, stabilization of the atmosphere in the surrounding regions. This leads to increased convective rainfall in the regions with substantial SST anomalies, and to reduced convective rainfall outside. The stabilization is induced by the upper-level tropical wave propagations (equatorial Rossby and Kelvin waves) discussed in section 5, which spread the signal in the tropical belt.
- b) Upper-level convergence in the region surrounding the heating, which compensates for the upper-level divergence in the heating region. This may, according to the continuity equation in pressure coordinates, 44, induce vertical velocities driven from the upper-levels.
- c) Flow induced by the heating in remote regions which may be forced to rise because of orography.
- d) Ekman pumping induced by remote atmospheric responses (e.g. Gill response discussed in section 6.3) to the heating.
- e) Changes of the tropical meridional temperature gradients induced by land masses that drive monsoons. These changes are also communicated from region with the SST perturbation to other tropical regions by equatorial wave propagation.

There are more mechanisms which have been proposed in the scientific literature for rainfall responses to an SST anomaly, but the above list contain the most commonly used hypothesis. Here we will use an idealized *aquaplanet* (explain!)

modelling approach that eliminates the hypothesis c) and e), and leaves the hypothesis a), b) and d). Such an aquaplanet setting has the advantage that some of the complexities in the real world (e.g. land-surface interactions) that make understanding of the responses very difficult, are removed, while the basic dynamical feedbacks are retained. Fig. 37 shows the ICTPAGCM (SPEEDY) models climate in such an aquaplanet setting. All fields shown are annual means.

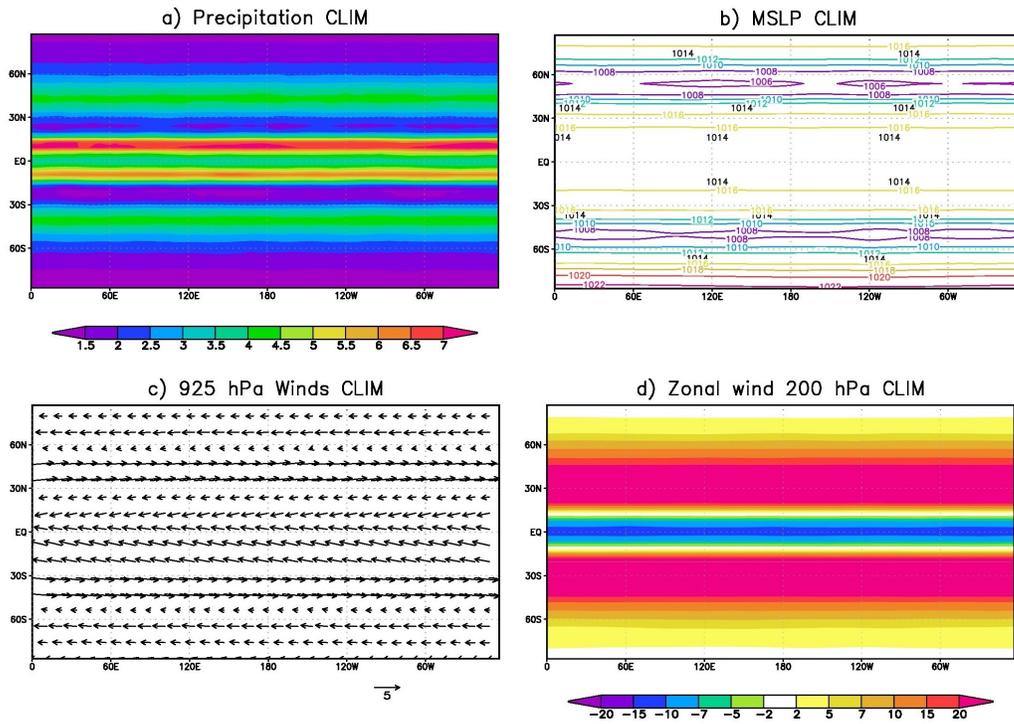


Figure 37: Aquaplanet annual mean climatologies of a) precipitation [mm/day], b) mean sea level pressure [hPa], c) surface winds [m/s] and d) 200 hPa zonal winds [m/s]

Now, the response to a gaussian Gill-type SST anomaly is analysed. For this purpose an additional experiment is performed in which such an anomaly is added to the aquaplanet SST field (Fig. 38). The responses of mean surface pressure, zonal wind, 200 hPa velocity potential, 200 hPa eddy streamfunction, precipitation and 850 hPa pressure vertical velocity are shown in Fig. 39. This response may be interpreted as the Gill-type response reproduced by an Atmospheric General Circulation Model (AGCM). Indeed, in the surface pressure we can identify off-equatorial Rossby-gyres to the west and high pressure at the equator resembling the Kelvin-wave type response. The precipitation and 850 hPa pressure vertical velocity response match each other quite closely, which may not be surprising because on one hand upward motion will lead to condensation (remember lecture course on EST). On the other hand, as discussed in section 6.2, in the tropics there is an

approximate equilibrium between the diabatic heating (cooling) and the adiabatic cooling (warming) term (see equation 173)

$$-S_p\omega \approx \frac{Q}{c_p} \quad . \quad (224)$$

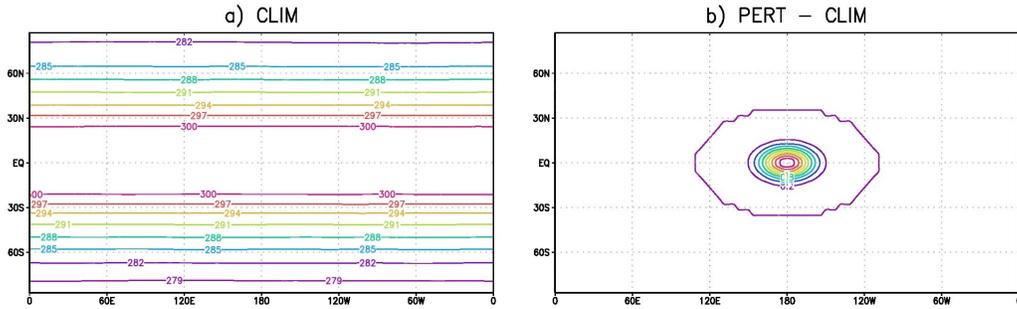


Figure 38: Aquaplanet annual mean SST distribution and Gill-type SST perturbation. Units are K.

Also this leads to a good agreement between precipitation (e.g. vertical integrated heating) and the vertical velocity field. Note, however, that in the Gill solution there is no precipitation, and the vertical velocity field is very different from the one shown in Fig. 39 (there we have sinking motion everywhere outside the heating region). These differences are a weakness of the simplified equations used in the Gill solution (essentially modified shallow water equations) rather than a weakness of the AGCM solution, which essentially solves the full Navier-Stokes equations. Another important difference with respect to the traditional Gill solution is that there is an infinite zonal domain is assumed, whereas in reality it is periodic. This leads to differences in the atmospheric adjustments to the heating. For example, to the east of the heating, we do find off-equatorial anticyclonic Rossby gyres in the surface pressure, which resemble the cyclonic ones to the west. We may interpret these as response to the upper-level convergence (maximum in velocity potential) at around 60W to 30W. However, we do note that in the regions with the imposed SST anomaly (around 180E), we get increased rainfall (heating) due to the destabilization of the atmosphere. The rainfall structure appears to be related to the meridional zonal wind gradient (main component of vorticity), which is cyclonic in the off-equatorial regions to the west, and anticyclonic in the off-equatorial regions to the east of the heating (explain!), which should, according to the section 7, induce rising and sinking motion respectively (e.g. from Eq. 221). However, since we are using a numerical model to simulate the effects of Ekman pumping, it is best to use the model's boundary-layer parameterization, which is slightly different. Using the definition

$$|v_0| = \sqrt{u^2 + v^2 + V_{gust}^2} \quad (225)$$

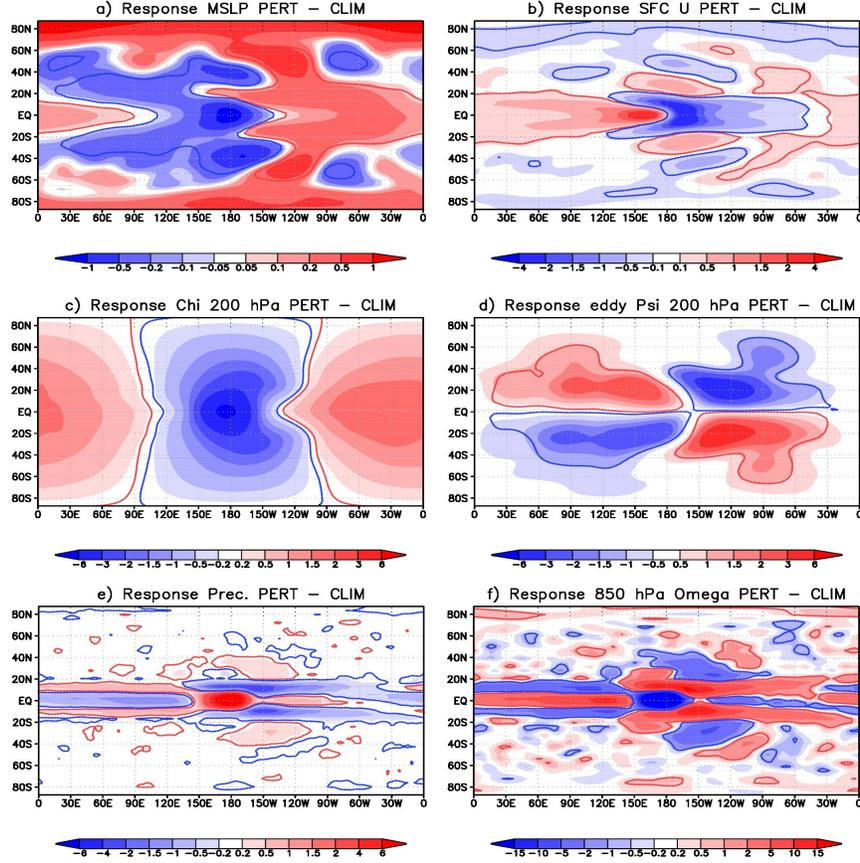


Figure 39: Response to Gill-type SST perturbation of a) mean sea level pressure [hPa], b) surface zonal wind [m/s], c) 200 hPa velocity potential [10^6 m 2 /s 2], d) 200 hPa eddy streamfunction [10^6 m 2 /s 2], e) precipitation [mm/day] and d) omega at 850 hPa [10^{-3} Pa/s].

where u, v are near surface winds (representative for the mean wind in the whole depth of the boundary-layer of the model), $V_{gust} = 5ms^{-1}$ is representing the effect of sub-grid scale gusts, the equations equivalent to Eqs. 207 become for the AGCM:

$$-Cu\sqrt{u^2 + v^2 + V_{gust}^2} + f(v - v_g) = 0 \quad (226)$$

$$-Cv\sqrt{u^2 + v^2 + V_{gust}^2} - f(u - u_g) = 0 \quad (227)$$

where $C = \frac{C_D}{h}$, $C_D = 1.8 \times 10^{-3}$ is the drag coefficient over sea, $h = 1000m$ is the depth of the boundary layer, and the geostrophic wind is u_g, v_g . Also, the AGCM is posed in the pressure coordinate system, and since we want to assess solutions including the equator it is convenient to use the geostrophic forcing in terms of

geopotential height gradients using the geostrophic relation:

$$\mathbf{v}_g = f^{-1} \mathbf{k} \times \nabla \Phi \quad (228)$$

in pressure coordinates. Since Eqs. 226 and 227 are nonlinear, they are solved numerically for the near surface winds, given the geopotential at 850 hPa (by re-introducing the time derivative as in the exercise 3) of section 7). Note that in Eqs. 226 and 227 the near surface winds are non-zero. Once the solutions are found, the divergence of the near surface wind is calculated and the continuity equation in pressure coordinates, 44, is vertically integrated (assuming the near surface winds are constant) to find the vertical velocity, ω_{ek} , induced by Ekman pumping in the model (1000 m corresponds according to the hydrostatic equation approximately to a pressure change of 100 hPa). The resulting Ekman vertical velocity is shown in Fig. 40, and shows very good agreement with the vertical velocity field at 850 hPa shown in Fig. 39 outside the region where the SST perturbation is present. This indicates that Ekman pumping is indeed a very powerful mechanism to induce vertical motion outside the heating region. Also, the upper-level convergence field is calculated for comparison (Fig. 41). There is also some correspondence of this field with 850 hPa vertical velocity and thus rainfall, but it should be noted that the upper-level convergence field and the ω_{ek} field are not independent (e.g. Ekman pumping could induce vertical velocity, this induces convection and thus upper-level divergence). The thermodynamic mechanism a) turns out to be relevant only in the region with SST perturbation and seems to be largely irrelevant outside that region.

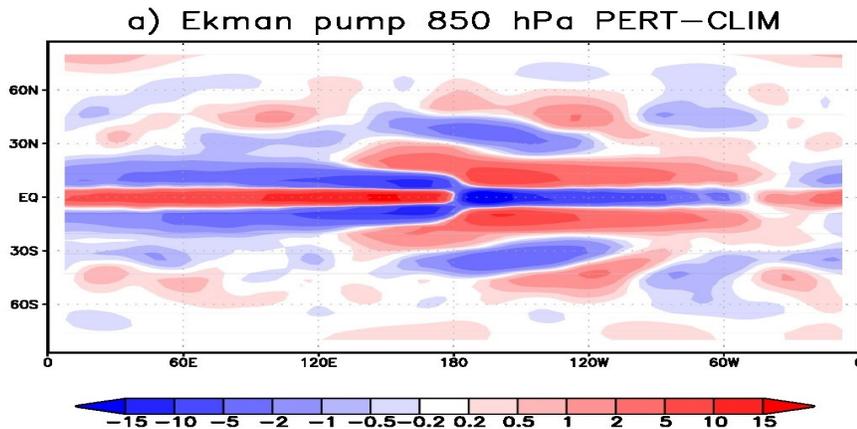


Figure 40: Ekman pumping induced ω_{ek} [10^{-3} Pa/s].

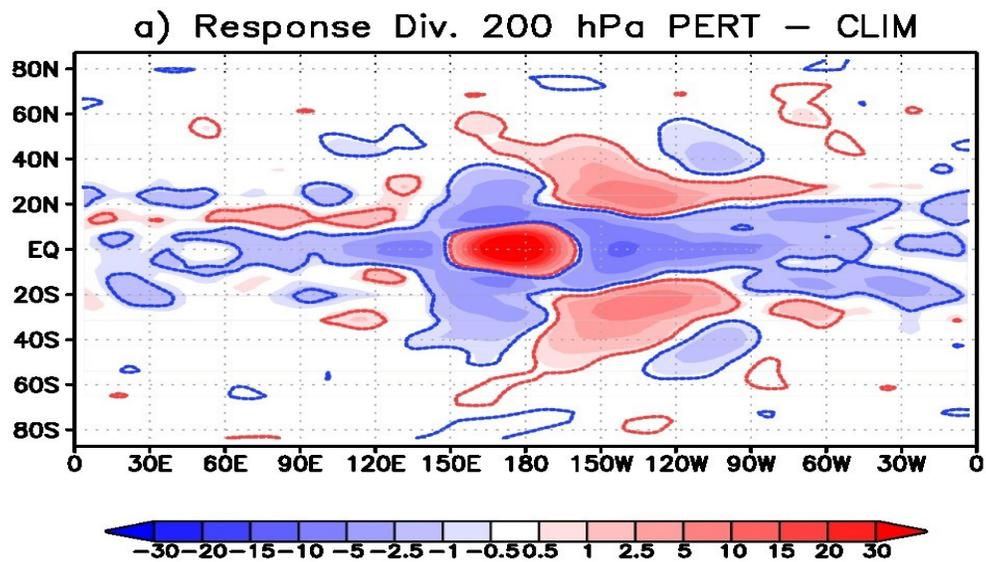


Figure 41: Response in 200 hPa divergence [1/s].

9 The general circulation

The *general circulation* of the atmosphere is usually considered to include the totality of motions that characterise the global-scale atmospheric flow. Climate dynamics is one of the main topics of the study of the general circulation. Here we are interested in the temporally (i.e. monthly) averaged fields of wind, temperature, humidity, precipitation, and other meteorological variables and their long-term variations (also called *low-frequency variability*). For example, monsoon systems are a very important feature of the general circulation. For example, on the web-page <http://users.ictp.it/~kucharsk/speedy8.clim.html> we find some features relevant to the general circulation.

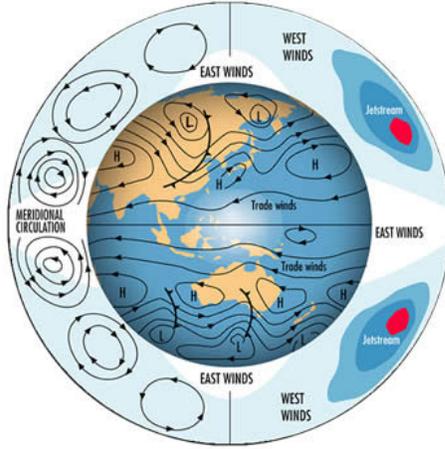


Figure 42: Schematic of some features of the general circulation.

9.1 Zonally averaged circulation

The aim of this section is to analyse the zonal mean circulation. The basis for the following analysis are the thermo-hydrodynamic equations in pressure coordinates Eqs. (36, 37, 44 and 47)

We apply in the following an averaging operator to these equations

$$\bar{A} \equiv \frac{1}{2\pi r \cos\phi} \int_0^{2\pi} A r \cos\phi d\lambda \quad . \quad (229)$$

All quantities are then expressed as the zonal mean plus a deviation from the zonal mean $A = \bar{A} + A'$ (see Figure 7).

For the total derivative of a quantity A

$$\frac{dA}{dt} = \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + \omega \frac{\partial}{\partial p} \right) A + A \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p} \right) , \quad (230)$$

where we added zero on the rhs according to the continuity equation. Therefore we may write the total derivative (in pressure coordinates!) as

$$\frac{dA}{dt} = \left(\frac{\partial A}{\partial t} + \frac{\partial Au}{\partial x} + \frac{\partial Av}{\partial y} + \frac{\partial A\omega}{\partial p} \right), \quad (231)$$

Application of the zonal operator (229) yields (as in section 7)

$$\frac{\overline{dA}}{dt} = \left(\frac{\partial \overline{A}}{\partial t} + \frac{\partial(\overline{A}\overline{v} + \overline{A}'v')}{\partial y} + \frac{\partial(\overline{A}\overline{\omega} + \overline{A}'\omega')}{\partial p} \right), \quad (232)$$

because $\overline{\partial(\cdot)\partial x} = 0$ and

$$\overline{ab} = \overline{(\overline{a} + a')(\overline{b} + b')} = \overline{a}\overline{b} + \overline{a'b'} + \overline{a'\overline{b}} + \overline{a'\overline{b}'} = \overline{a}\overline{b} + \overline{a'b'},$$

because the quantities $\overline{(\cdot)}$ are independent of x and $\overline{a'} = \overline{b'} = 0$. Applying the zonal average to the continuity equation leads to

$$\frac{\partial \overline{v}}{\partial y} + \frac{\partial \overline{\omega}}{\partial p} = 0 \quad . \quad (233)$$

Note that with Eq. 233 we can define a streamfunction:

$$\Psi = \int_0^p \overline{v} dp \quad , \quad (234)$$

so that

$$\overline{v} = \frac{\partial \Psi}{\partial p} \quad ; \quad \overline{\omega} = -\frac{\partial \Psi}{\partial y} \quad . \quad (235)$$

To show 235 will be an Exercise! From Eq. (232) we can also derive

$$\frac{\overline{dA}}{dt} = \frac{\overline{d}}{dt} \overline{A} + \frac{\partial \overline{A}'v'}{\partial y} + \frac{\partial \overline{A}'\omega'}{\partial p}, \quad (236)$$

where

$$\frac{\overline{d}}{dt} = \frac{\partial}{\partial t} + \overline{v} \frac{\partial}{\partial y} + \overline{\omega} \frac{\partial}{\partial p} \quad (237)$$

is the rate of change following the mean motion. Averaging the zonal component of the momentum equation 36 and the thermodynamic equation 47 leads to

$$\frac{\partial \overline{u}}{\partial t} - f_0 \overline{v} = -\frac{\partial \overline{u}'v'}}{\partial y} \quad (238)$$

$$\frac{\partial \overline{T}}{\partial t} - S_p \overline{\omega} = -\frac{\partial \overline{v}'T'}}{\partial y} + \frac{\overline{Q}}{c_p} \quad (239)$$

Here several further approximations have been introduced which are all consistent with quasi-geostrophic scaling. A similar scaling shows that the meridional momentum equation is in quasi-geostrophic balance. For the zonal averaged meridional momentum equation, the first order geostrophic approximated balance is

$$f_0 \bar{u} = -\frac{\partial \bar{\Phi}}{\partial y} \quad . \quad (240)$$

Together with the zonal average of the hydrostatic equation 37 this leads to the thermal wind equation for zonal averaged motion

$$\frac{\partial \bar{u}}{\partial p} = \frac{R}{f_0 p} \frac{\partial \bar{T}}{\partial y} \quad . \quad (241)$$

This equation is similar to equation (53), but for zonal averages. Equations (238) and (239) tell us that in order to get a steady-state meridional, vertical circulation cell ($\bar{v}, \bar{\omega}$) we must have the balances

Coriolis force $f_0 \bar{v} \approx$ divergence of eddy momentum fluxes

Adiabatic cooling \approx diabatic heating plus convergence of eddy heat fluxes

Also note that any \bar{v} and $\bar{\omega}$ separately induces the other due to continuity 233.

Analysis of observations shows that outside the tropics these balances appear to be approximately true above the boundary layer. Close to the equator we have that the heating is mainly balanced by mean vertical motion, driving the *Hadley Cell*, whereas in the extratropics the meridional, vertical circulations are mainly driven by the convergence of eddy momentum and eddy heat fluxes. These cells are called *Ferrell Cells*. Discuss that the (angular) momentum fluxes should be toward the extratropics because of the absolute (angular) momentum loss of the atmosphere in the extratropics and gain in the tropics (Fig. 44, upper panel). Also discuss effect of tilt of waves, and the fact that Rossby waves radiate energy away from the jet (baroclinic zone), means at the same time that they carry momentum towards the jet (see Fig. 45) One can use the radiation condition considering the meridional energy propagation Eq. 78, suggesting $kl > 0$ (and $l > 0$) north and $kl < 0$ (and $l < 0$) south of jet, implying $k > 0$ in both cases)! Also discuss how the shear induced by the jet as well as $\beta(y)$ may modify the tilt of the phases of waves, with the latter one responsible for the dominance of poleward eddy momentum transport south of the jet.

Exercises

1. Show that with the streamfunction definition Eq. 234, the relationships 235 are fulfilled.

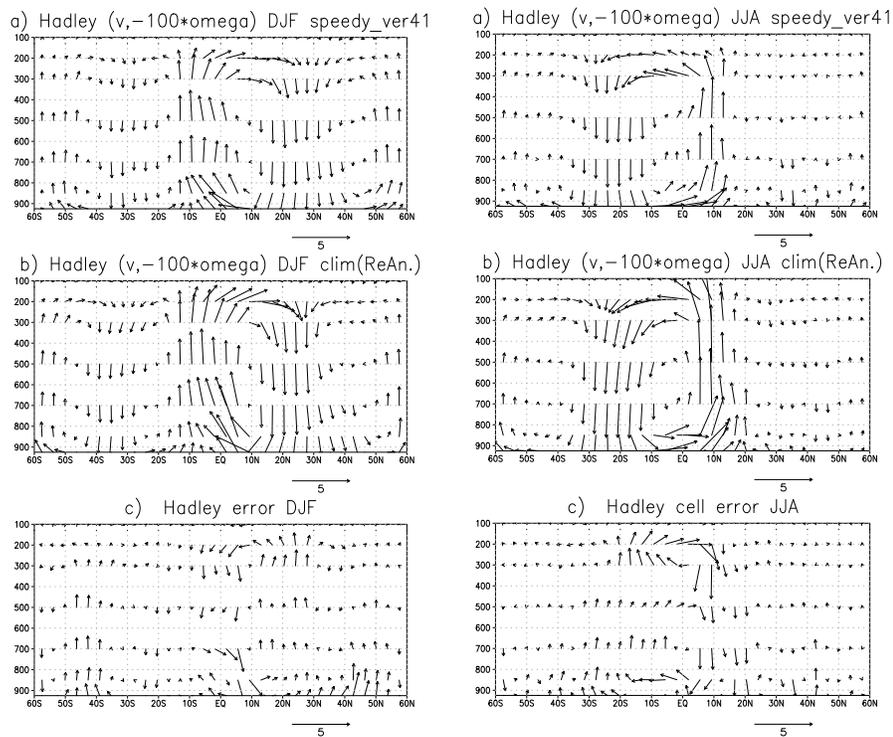


Figure 43: Illustration of the Hadley cell by a $(\bar{v}, -\omega)$ vector plot. left panel: Boreal winter, right panel: boreal summer.

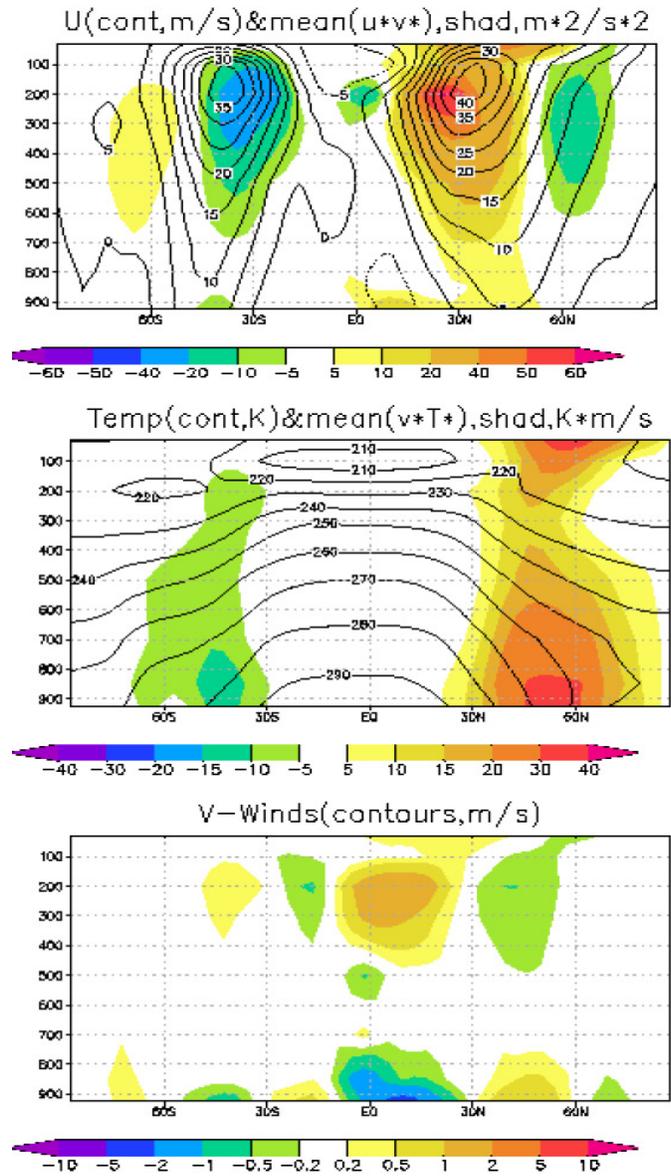


Figure 44: Upper panel: \bar{u} (contours), and $\overline{u'v'}$, middle panel: \bar{T} (contour) and $\overline{v'T'}$, lower panel \bar{v} .

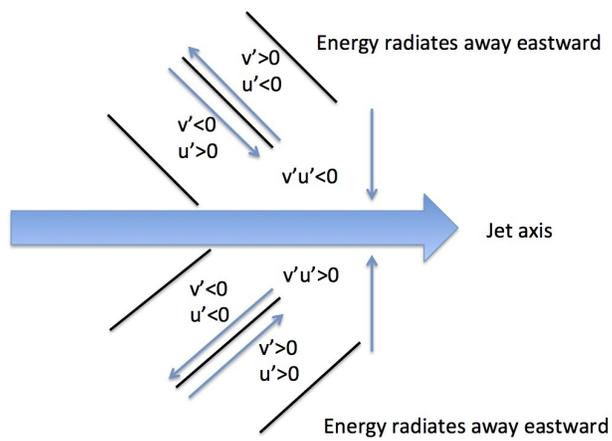


Figure 45: Sketch of why the eddy momentum flux is always towards the Jet.

10 Tropical zonal and meridional circulations

More reading:

1. Rodwell MJ, Hoskins BJ (2001) Subtropical Anticyclones and Summer Monsoons. *J Clim*, **14**, 3192–3211
2. Chao WC, Chen B (2001) The origin of Monsoons. *J Atmos Sci*, **58**, 3497–350
3. Kucharski, F, Bracco, A, Barimalala, R, Yoo, J-H (2011) Contribution of the east–west thermal heating contrast to the South Asian Monsoon and consequences for its variability, *Clim Dyn*, **37**, 721–735, DOI 10.1007/s00382-010-0858-3

Let’s have a look at the global June-to-September rainfall distribution in Fig. 46. We can clearly identify the Intertropical Convergence Zone (ITCZ), identified by the rainfall maximum north of the equator that can reach 30°N in some land region. We also see that this strip of large rainfall is not zonally homogeneous, but is stronger in some locations than in others. Some of the features may be explained by the distributions of sea surface temperatures (SSTs) in Fig. 47. Note that the rainfall is the column integrated $-L_{lv} \frac{dm_v}{dt}$ and therefore related to the column integrated diabatic heating due to condensation according to Eqs. 171 and 46.

The western Pacific rainfall maximum is related to high SSTs in that region, as we have already discussed several times. Also, the land-sea contrast are likely important due to different heat capacities. As we have already discussed regarding the ENSO phenomenon, the distribution of SSTs, rainfall and atmospheric circulations in the tropical Pacific provide positive feedbacks between them, so that it is difficult to say what is cause and what is effect (chicken-egg problem). The mean zonal circulation in the tropical regions is called *Walker circulation* (see also Fig. 20). Note that this circulation is not a strict closed circulation cell as we could derive for the zonal mean circulation (Hadley Cell). A good measure of this zonal tropical circulation is the upper-level *velocity potential*, χ , for which we have the relation to the divergent wind

$$\mathbf{v}_\chi = \nabla \chi \quad (242)$$

The distribution of the 200 hpa velocity potential χ is shown in Fig. 48a (which height is this, approximately). According to the definition 242, a minimum means divergent wind. The centre of upper-level divergence (rising motion, why?) is in the western Pacific region, and the centers of upper-level convergence (sinking motion, why?) are located in the eastern Pacific (that is the classical Walker circulation) and in the tropical South Atlantic region.

Fig. 48 b) shows the *streamfunction*, ψ , which is related to the rotational flow in the following way (see, e.g. Eq. 16)

$$\mathbf{v}_\psi = \mathbf{k} \times \nabla \psi \quad . \quad (243)$$

It is clear that there is a systematic relationship between upper-level velocity potential and upper-level streamfunction, they are in 'quadrature', that is the extreme values of one lie in the gradient regions of the other. Can we propose an explanation for this behaviour? The solution is an approximate version of the vorticity equation 67 for the tropics

$$\frac{d_g \xi_g}{dt} = -f_0 \left(\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y} \right) - \beta v_g \quad . \quad (244)$$

It turns out that in the tropical regions relative vorticity changes and advections are relatively small, mainly because of approximate geostrophy and small pressure gradients. We may estimate it to be one order of magnitude smaller than in the extratropics. Therefore, the left hand side is Eq. 244 may be set to zero. This leaves us with the approximation

$$\beta v_\psi = -f_0 \left(\frac{\partial u_\chi}{\partial x} + \frac{\partial v_\chi}{\partial y} \right) \quad , \quad (245)$$

where we have used that the geostrophic wind is the rotational wind defined in Eq. 243 (note that according to the quasigeostrophic vorticity definition 66, $\psi = \Phi/f$ is the streamfunction as long as we are not too close to the equator), and the ageostrophic wind is the divergent wind defined in Eq. 242. Eq. 245 is called *Sverdrup balance*. If we take a divergence field as given (e.g. the field that corresponds to Fig. 48a; we can imagine it has been caused by the dominance of the west Pacific heating), then according to the Sverdrup balance, this will cause rotational winds. An upper level divergence maximum will lead to southward rotational motion in the northern hemisphere, an upper-level convergence maximum will lead to northward rotational motion. This is consistent with the streamfunction distribution in Fig. 48b, and explains why velocity potential and streamfunction are in quadrature (exercise!). Note that the interpretation of the Asian monsoon high to be partially forced by Sverdrup balance from the heating differences between the western Pacific and the Atlantic Ocean is a relatively new one. The conventional point of view is that the Asian monsoon high (*Tibetan high*) is forced exclusively by the land sea contrast between the Asian land mass (including importance of Himalayas) and the Indian Ocean. We also note that all the structures that we have considered for upper levels should be reversed for low levels, thus a part of the low-level monsoon trough can be attributed to the western Pacific/Atlantic heating differences. Keep in mind that secondary vertical motions can be induced then by surface friction, because the streamfunction centres that we can derive are fields with vorticity, which according to equation 221 induces vertical motion (see section 8). Positive (negative) vertical motion, in turn, leads to increased (decreased) rainfall (why?). These effects have been analysed with idealized numerical experiments in the paper cited in the beginning of the section (see also Figs. 49, 50, 51).

Exercises

1. Let the velocity potential distribution in zonal direction at latitude 30°N be

$$\chi = A\cos(\lambda) ,$$

where λ is longitude. Calculate using the Sverdrup balance 245 to derive the zonal distribution of the streamfunction at the same latitude.

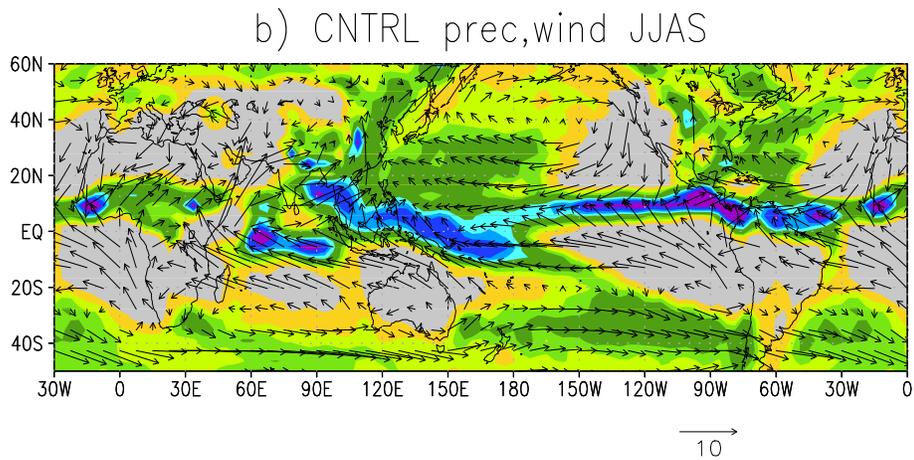
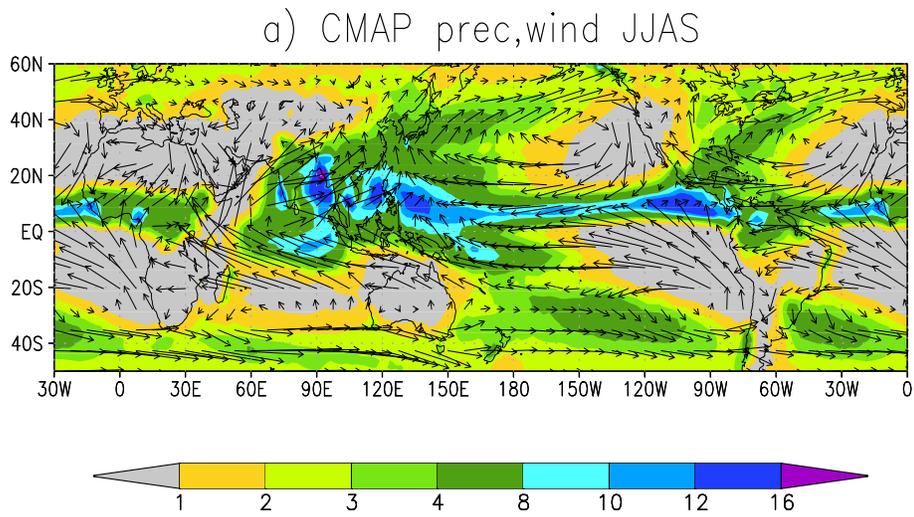


Figure 46: Mean JJAS rainfall and surface winds a) from observations (CMAP,NCEP-NCAR re-analysis), b) from the ICTPAGCM. Units are mm/day for rainfall and m/s for wind.

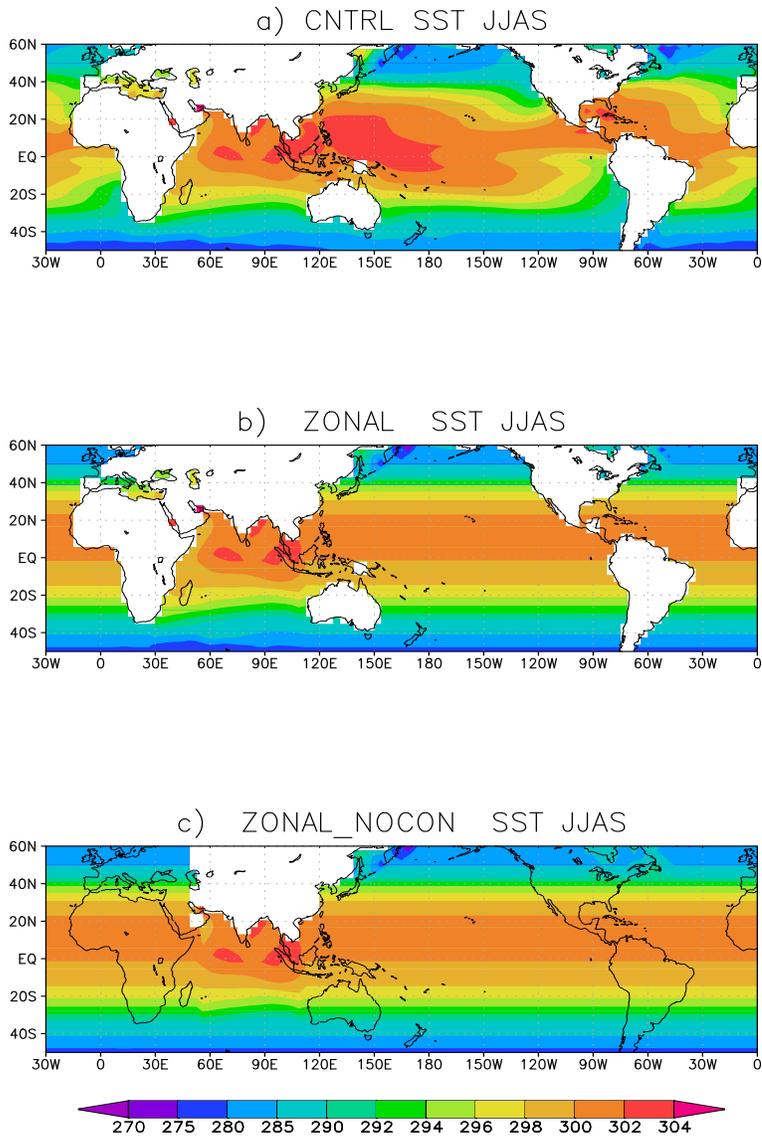


Figure 47: Mean JJAS sea surface temperature distribution a) observed, b) and c) idealized distribution for numerical experimentation.

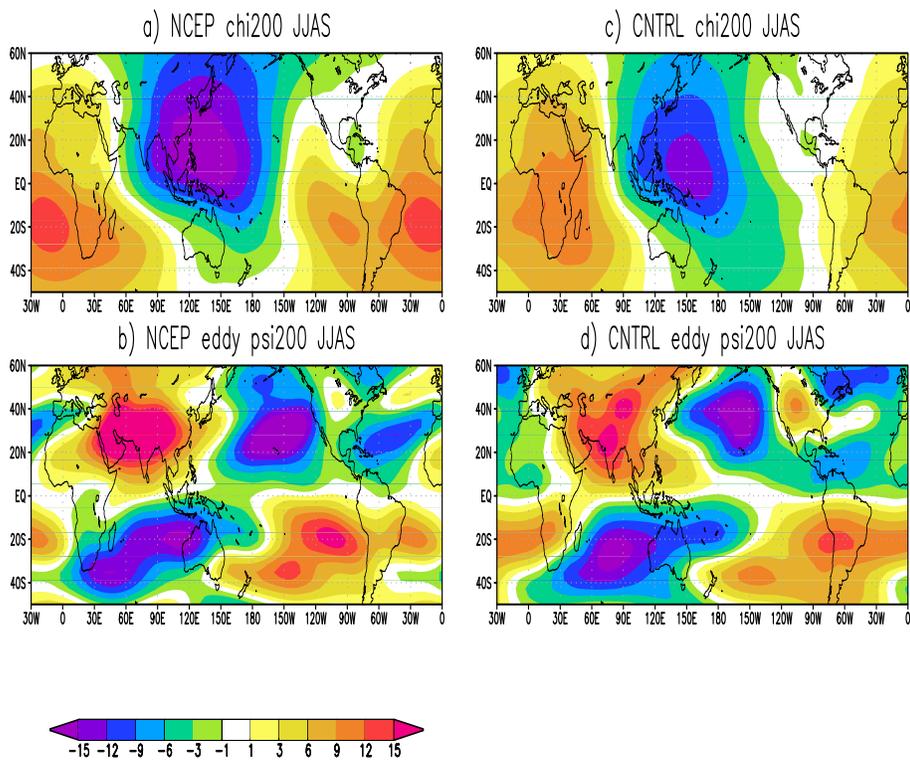


Figure 48: Mean JJAS distributions at 200 hPa of a) observed velocity potential χ , c) modeled observed velocity potential χ , b) observed streamfunction ψ , d) modeled streamfunction ψ . Units are $10^6 \text{ m}^2 \text{ s}^{-2}$.

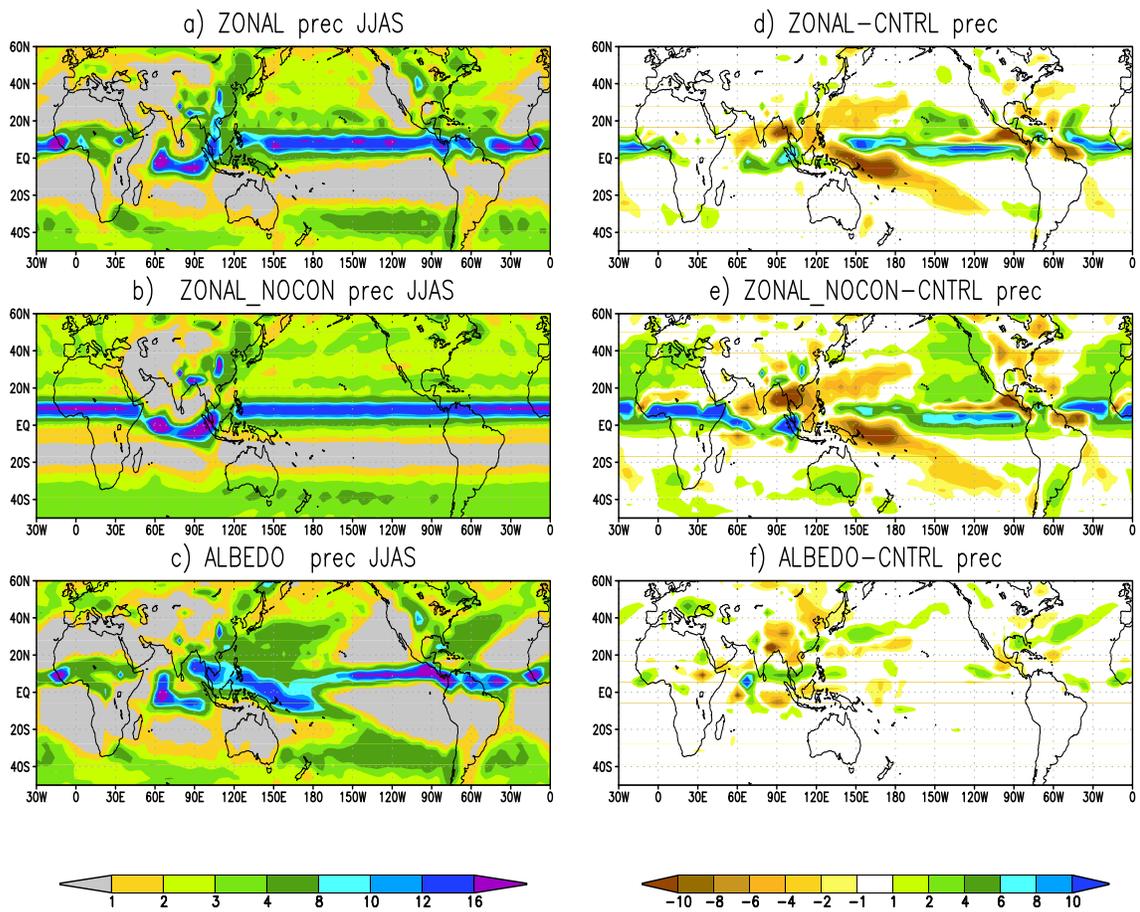


Figure 49: Response in rainfall to zonal (a, d)) and zonal with removed African and American continents (b, e). Units are mm/day.

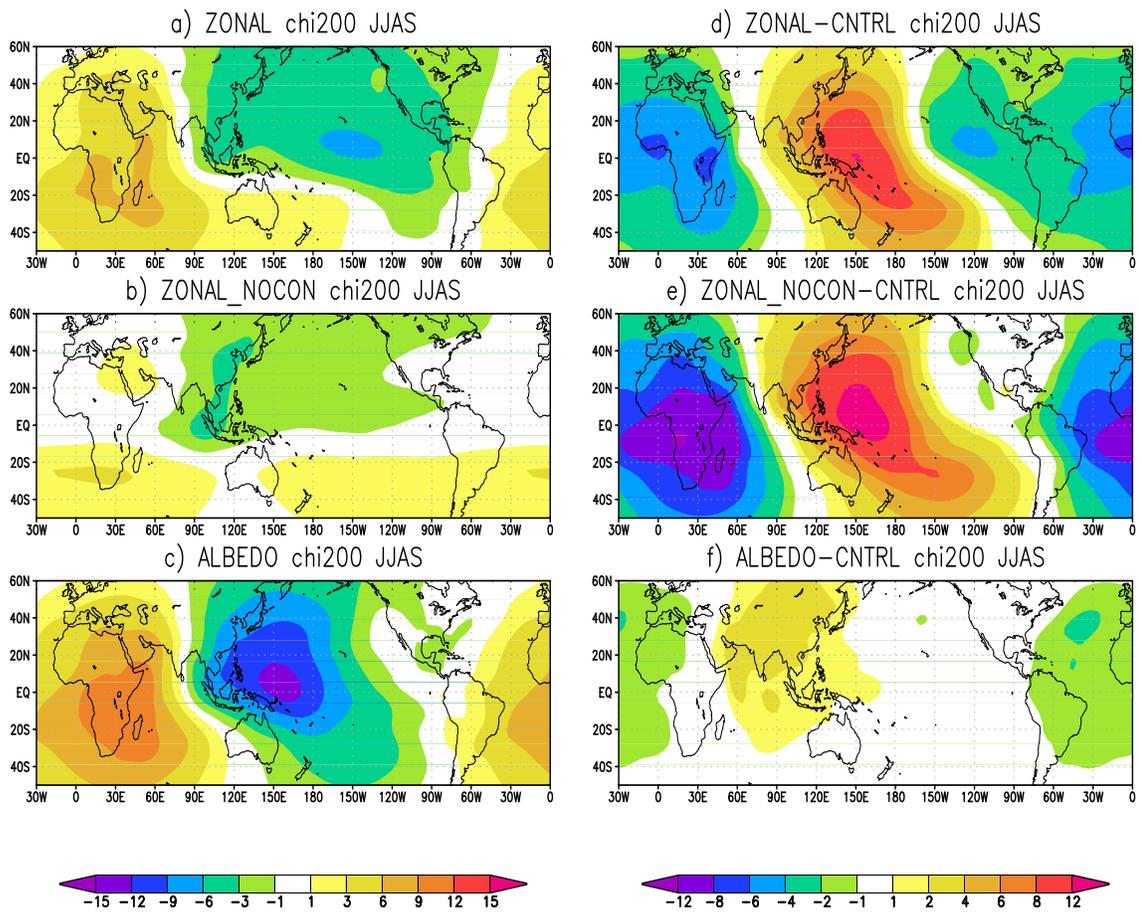


Figure 50: Response in velocity potential χ to zonal (a, d)) and zonal with removed African and American continents (b, e). Units are $10^6 \text{ m}^2 \text{ s}^{-2}$.

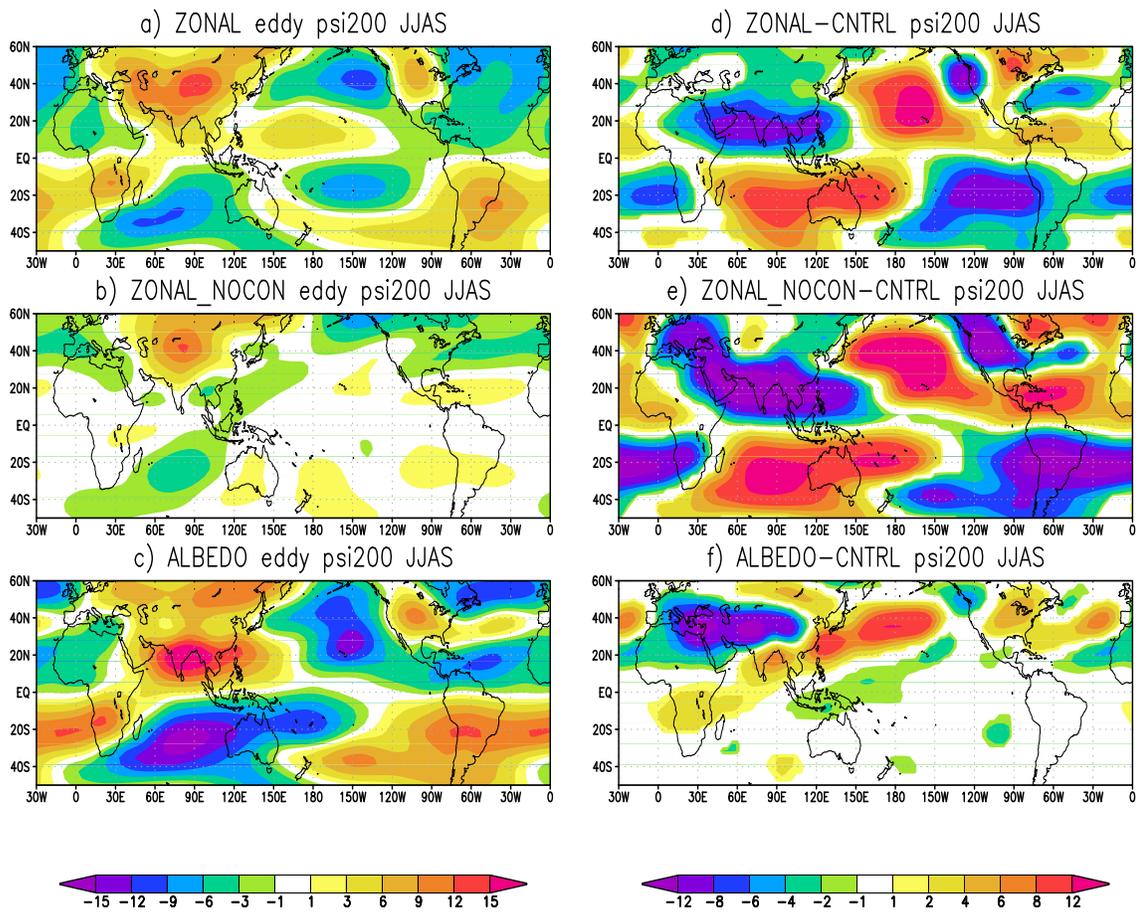


Figure 51: Response in eddy streamfunction ψ to zonal (a, d)) and zonal with removed African and Americal continents (b, e). Units are $10^6 \text{ m}^2 \text{ s}^{-2}$.

11 Energetics of the general circulation

Suggested textbooks, reading:

- a) Lorenz, E. N.; 1955: Available Potential Energy and the maintenance of the General Circulation. *Tellus*, **7**, 157-167
- b) Tailleux, R.; 2013: Available Energy and Exergy in Stratified Fluids. *Annual Review of Fluid Mechanics*, **45**, 35-58, doi: 10.1146/annurev-fluid-011212-14620
- c) Marquet, M.; 1991: On the concept of Exergy and available enthalpy: application to atmospheric energetics. *Q. J. R. Meteorolog. Soc.*, **117**, 449-475
- d) Kucharski, F.; 1997: On the concept of exergy and available potential energy. *Q. J. R. Meteorolog. Soc.*, **123**, 2141-2156

We go back to the full (unapproximated) momentum equation 23

$$\rho \frac{d\mathbf{v}}{dt} = -\nabla p - 2\rho\boldsymbol{\Omega} \times \mathbf{v} - \rho\nabla\phi - \nabla \cdot \mathbf{F} \quad , \quad (246)$$

where \mathbf{F} is the frictional tensor. Multiplying with the velocity gives

$$\rho \frac{dk}{dt} = -\nabla \cdot (p\mathbf{v} + \mathbf{F} \cdot \mathbf{v}) + p\nabla \cdot \mathbf{v} - \rho\mathbf{v} \cdot \nabla\phi + \mathbf{F} : \nabla\mathbf{v} \quad , \quad (247)$$

where $k = 1/2(\mathbf{v} \cdot \mathbf{v})$ is the kinetic energy. We can reformulate equation 247 as equation

$$\rho \frac{d(k + \phi)}{dt} = -\nabla \cdot (p\mathbf{v} + \mathbf{F} \cdot \mathbf{v}) + p\nabla \cdot \mathbf{v} + \mathbf{F} : \nabla\mathbf{v} \quad . \quad (248)$$

The equation for internal energy is

$$\rho \frac{du}{dt} = -\nabla \cdot \mathbf{J} - p\nabla \cdot \mathbf{v} - \mathbf{F} : \nabla\mathbf{v} \quad . \quad (249)$$

\mathbf{J} can contain the radiative flux vector and the diffusive (molecular) heat flux and these only heating terms for a one-component system in which no phase transitions are possible. The addition of condensational heating is an approximation to a one-component system and strictly possible only if more components are considered (dry air, water vapour and liquid water). The pressure work term $p\nabla \cdot \mathbf{v}$ may therefore be interpreted as reversible conversion term between kinetic plus potential and internal energy. The dissipational heating, $\mathbf{F} : \nabla\mathbf{v}$, is just transferring energy into the internal energy reservoir. The other terms are energy fluxes into the climate system. If we for the time being assume a closed system, then these terms vanish after an integration over the this system. In this case, we may assume that reservoirs of

internal, potential and kinetic energy can only change by exchanging energy between them.

Imagine an initial situation with a given temperature distribution, but without motion (this is, for example, the typical initial condition of ICTP AGCM). How much kinetic energy could the climate 'gain' in such a situation? A naive estimation would be $u = c_v T \approx 2 \cdot 10^5$ J/kg as typical local value. However, experience tells us that typical values of specific kinetic energy are about 10^2 J/kg, so that there is an overestimation of a factor of about 1000 in this simple (an naive) approach. Clearly, only a small portion of the large internal energy reservoir can be released into kinetic energy. It is obvious that we have to consider *differences* with respect to some reference state in order to define the proper amount of internal energy that is available for conversion into kinetic energy. An approach like $\Delta u = c_v \Delta T$, where $\Delta T = T - T_r$, with T_r some reference temperature, leads still to overestimations, but has the even more severe problem of not being positively definite. Note that the above arguments also apply to the potential energy, because of the proportionality of their total amounts in the atmosphere ($P = R/c_v U$; Exercise!). The problem of identifying the energy available for conversion into kinetic energy is a classical one. For the atmosphere, Lorenz (1955) has developed a new energy concept, called *Available Potential Energy*. We can re-write the equations in a more suitable form:

$$\rho \frac{dk}{dt} = -\nabla \cdot (\mathbf{F} \cdot \mathbf{v}) - \mathbf{v} \cdot (\nabla p + \rho \nabla \phi) + \mathbf{F} : \nabla \mathbf{v} \quad . \quad (250)$$

$$\rho \frac{d(u + \phi)}{dt} = -\nabla \cdot (p\mathbf{v}) + \mathbf{v} \cdot (\nabla p + \rho \nabla \phi) + T\rho \frac{ds}{dt} \quad . \quad (251)$$

s is the specific entropy, governing all irreversible processes (for an ideal gas we have $s = c_p \ln \theta$). From these equations we see that in a hydrostatic atmosphere ($\nabla p = -\rho \nabla \phi$), there is no reversible conversion from the internal plus potential to kinetic energy.

11.1 On Dry and Moist Static Energy

Before we move on to discuss the problem of the available potential energy, it is useful to derive from these exact equations the (approximate) conservation law for Dry Static Energy $= h + \phi$ (DSE; Eq. 132 from the lecture course on Earth System Thermodynamics). Please note that

$$-\nabla \cdot (p\mathbf{v}) = -\rho \frac{d}{dt}(p/\rho) + \frac{\partial p}{\partial t} \quad . \quad (252)$$

With this adding Eq. (250) and Eq. (251) we get

$$\rho \frac{d(k + u + p/\rho + \phi)}{dt} = \rho \frac{d(k + h + \phi)}{dt} = \frac{\partial p}{\partial t} - \nabla \cdot (\mathbf{F} \cdot \mathbf{v}) + \mathbf{F} : \nabla \mathbf{v} + T\rho \frac{ds}{dt} \quad . \quad (253)$$

This demonstrates that the DSE is approximately conserved for adiabatic-reversible flow, for which kinetic energy changes are small compared to changes of enthalpy and

potential energy, and for which the local pressure tendency is also sufficiently small. These last approximations go along with the hydrostatic approximation which has been made when deriving the dry static energy conservation (Eq. 132 of EST lecture course).

If we consider a system with moisture, we may approximate it's effect on the thermodynamic energy equation 249 according to 171 by adding the term $-L_{lv} \frac{dm_v}{dt}$ on the rhs of Eq. 253. This, using the good approximation of $L_{lv} = \text{const.}$ leads to the moist version of Eq. 253, *Moist Static Energy* = $h + \phi + L_{lv}m_v$ (MSE) equation:

$$\rho \frac{d(k + h + \phi + L_{lv}m_v)}{dt} = \frac{\partial p}{\partial t} - \nabla \cdot (\mathbf{F} \cdot \mathbf{v}) + \mathbf{F} : \nabla \mathbf{v} + T\rho \frac{ds}{dt} \quad . \quad (254)$$

The conservation of MSE requires the same approximate conditions as the conservation of the DSE.

Let's go back to derive the available potential energy. If in Eq. 251 we define a function that just depends on entropy, $T_0(s)$, then it is possible to remove from the internal plus potential energy a part that merely depends on entropy:

$$\rho \frac{d(u + \phi - q_0(s))}{dt} = -\nabla \cdot (p\mathbf{v}) + \mathbf{v} \cdot (\nabla p + \rho \nabla \phi) + (T - T_0(s))\rho \frac{ds}{dt} \quad , \quad (255)$$

where $q_0(s) = \int_{s_B}^s T_0(s') ds'$. Note that for isothermal processes, $T_0(s) = T$, then $q_0(s) = Ts + \text{const.}$ and $u + \phi - q_0(s) = u - Ts + \phi$, and $(u - Ts)$ is the *Free Energy* (a famous quantity for available energy for isothermal processes). The interpretation is that we can remove from the internal plus potential energy amounts a portion that is not available for conversion into kinetic energy. However, the problem of the positive definiteness of the *useful* energy $u + \phi - q_0(s)$ remains. We have to use a transformation of thermodynamic variables that guarantees positive definiteness in the end.

11.2 Exergy transformation

The exergy transformation is usually applied to the internal energy u to provide a positive quantity with properties of a thermodynamic potential (see Figure 52).

Application of this transformation to the energy $u + \phi - q_0$ results in

$$e_{ape} = \Delta u - \Delta q_0 + p_R \Delta \alpha \quad , \quad (256)$$

where $\Delta \psi = \psi - \psi_R$, where the subscript R indicates the reference state that depends only on z . The quantity e_{ape} is the available potential energy first introduced by Lorenz (1955) in a globally integrated form. It is a positive quantity the for a closed volume and the only *reversible* production/destruction term is the conversion to/from kinetic energy. This quantity therefore has all the required properties. The reference state is arbitrary at this point, but following our idea of e_{ape} being the available energy, it should be derived from a variational principle, minimizing the volume integral of e_{ape} . However, for practical purposes, determination of the

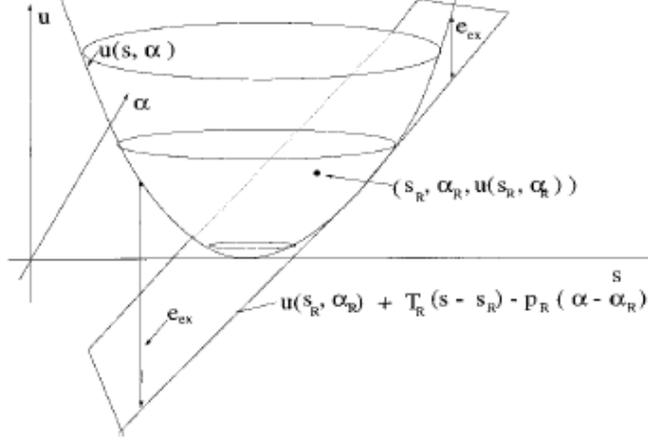


Figure 52: The convex internal energy, u , as a function of specific entropy, s , and specific volume, α , in phase space. The exergy transformation is just the (always positive) difference between u the tangente at the reference point.

reference state by, for example, horizontal averaging of temperature is sufficient. All other reference state variables can be determined by vertically integrating the hydrostatic equation for an ideal gas. Note that e_{ape} according to Eq. 256 becomes the classical exergy in case of an isothermal reference temperature $T_0(s) = const. = T_R$. Finally, for an ideal gas, the lowest order approximation of e_{ape} is

$$e_{ape} \approx \frac{1}{2} \left[\frac{g^2}{N_R^2} \left(\frac{\Delta\theta}{\theta_R} \right)^2 + RT_R \frac{c_v}{c_p} \left(\frac{\Delta p}{p_R} \right)^2 \right], \quad (257)$$

where

$$N_R^2 := \frac{g}{c_p} \frac{ds_R}{dz} = \frac{g^2}{c_p^2} \left[\frac{T_R}{c_p} - \frac{dT_R}{dz} \left(\frac{ds_R}{dz} \right)^{-1} \right]^{-1}. \quad (258)$$

From Eq. 258 follows that e_{ape} is positive if the stratification of the reference state is statically stable ($N_R^2 > 0$). For application of this concept to the general circulation of the atmosphere this condition is certainly fulfilled (just look at a potential temperature cross section in meridional and height direction, e.g. Fig. 3). However, to a system where the stratification is unstable even in the horizontal average, the available potential energy concept cannot be applied, and we have to return to other methods of identifying useful energy. The balance equation for e_{ape} can be derived by noting that

$$\frac{de_{ape}}{dt} = (T - T_0(s)) \frac{ds}{dt} - \Delta p \frac{d\alpha}{dt} + \Delta\alpha \frac{dp_R}{dt} \quad (259)$$

The energy equations 250 and 251 can be reformulated in the following form

$$\rho \frac{dk}{dt} = -\nabla \cdot (\Delta p \mathbf{v} + \mathbf{F} \cdot \mathbf{v}) + \Delta p \nabla \cdot \mathbf{v} + \rho \frac{\Delta\alpha}{\alpha_R} \mathbf{v} \cdot \nabla \phi + \mathbf{F} : \nabla \mathbf{v} \quad (260)$$

$$\rho \frac{de_{ape}}{dt} = -\Delta p \nabla \cdot \mathbf{v} - \rho \frac{\Delta \alpha}{\alpha_R} \mathbf{v} \cdot \nabla \phi + \rho \frac{T - T_o(s)}{T} q \quad , \quad (261)$$

where the abbreviation for diabatic processes $Tds/dt = q$ has been used. The factor $(T - T_o(s))/T = \eta$ may be interpreted as *Carnot factor* controlling the efficiency of the energy gain for a given heating. It can be approximated to the first order as (exercise!)

$$\eta \approx \left(\frac{g^2}{c_p N_R^2} \frac{\Delta \theta}{\theta_R} + T_R \frac{R}{c_p} \frac{\Delta p}{p_R} \right) \frac{1}{T} \quad . \quad (262)$$

The interpretation is then that the volume integrated generation of e_{ape} is dependent of the correlation of η and q . It's not the heating per se that is important, but the differential heating has to be correlated with potential temperature perturbations (dominant part in Eq. 262). On the other hand, frictional dissipation is the main process that destroys kinetic energy. A few remarks are appropriate here. If we forget for the time being about the approximative expressions given for an ideal gas, e_{APE} , and its evolution equation are really very general for a stably stratified one-component system. For example consider a simple solid body. We know from experience that even if a differential heating is applied, the only process that may result is heat diffusion (diabatic process), but certainly no kinetic energy can be gained. Indeed, in case of a solid body the function $q_0(s)$ can be chosen as the internal energy u plus the constant potential energy, because $T = T(s)$, therefore we can define $T_0(s) = T$. In case of a simple incompressible fluid, we can also identify $T_0(s) = T$ and $q_0(s)$ can be chosen to equal to the internal energy, but the potential energy may not be constant, because of fluctuations of surface height, and could therefore provide the available potential energy.

To derive a slightly different formulation that has been derived by Lorenz (1955), equation has to be integrated over the systems volume

$$\int_{\tau} \frac{de_{ape}}{dt} d\tau = \int_{\tau} \left\{ -p \frac{d\alpha}{dt} + p_R \alpha \nabla \cdot \mathbf{v} - \frac{\alpha}{\alpha_R} \mathbf{v} \cdot \nabla \phi + \frac{d\phi}{dt} + \frac{T - T_o(s)}{T} q \right\} d\tau \quad , \quad (263)$$

where $d\tau = \rho dx dy dz$ is a mass element. Using

$$p \frac{d\alpha}{dt} = \frac{d}{dt}(p\alpha) - \alpha \frac{dp}{dt}$$

where $\alpha = 1/\rho$. With $P = R/c_v U$ (see exercise 1!), it follows

$$\int_{\tau} \frac{de_{ape}}{dt} d\tau = \int_{\tau} \left\{ \frac{RT}{p} \omega + \frac{T - T_o(s)}{T} q \right\} d\tau \quad , \quad (264)$$

where the orographic term from the exercise has been ignored. This is the integral available potential energy balance for a hydrostatic atmosphere that was first derived by Lorenz (1955). Apart from the diabatic production term, which can be reformulated for an idealized gas to get a formulation identical to Lorenz, this form

is highlighting the conversion term to kinetic energy $-\omega RT/p$, which states that energy is converted into kinetic energy by *rising of warm air and sinking of cold air*. This makes the conversion term positive on average. This process is lowering the *centre of mass* of the atmosphere, thus releasing kinetic energy. In baroclinic waves, warm air is typically moving northward, and cold air southward. This together with the tendency for motions to be adiabatic, thus following lines of constant potential temperature (see Fig. 3), means that the warm air moving to the north tends to rise and the cold air moving to the south tends to sink, thus providing the conditions for conversion of available potential energy into kinetic energy.

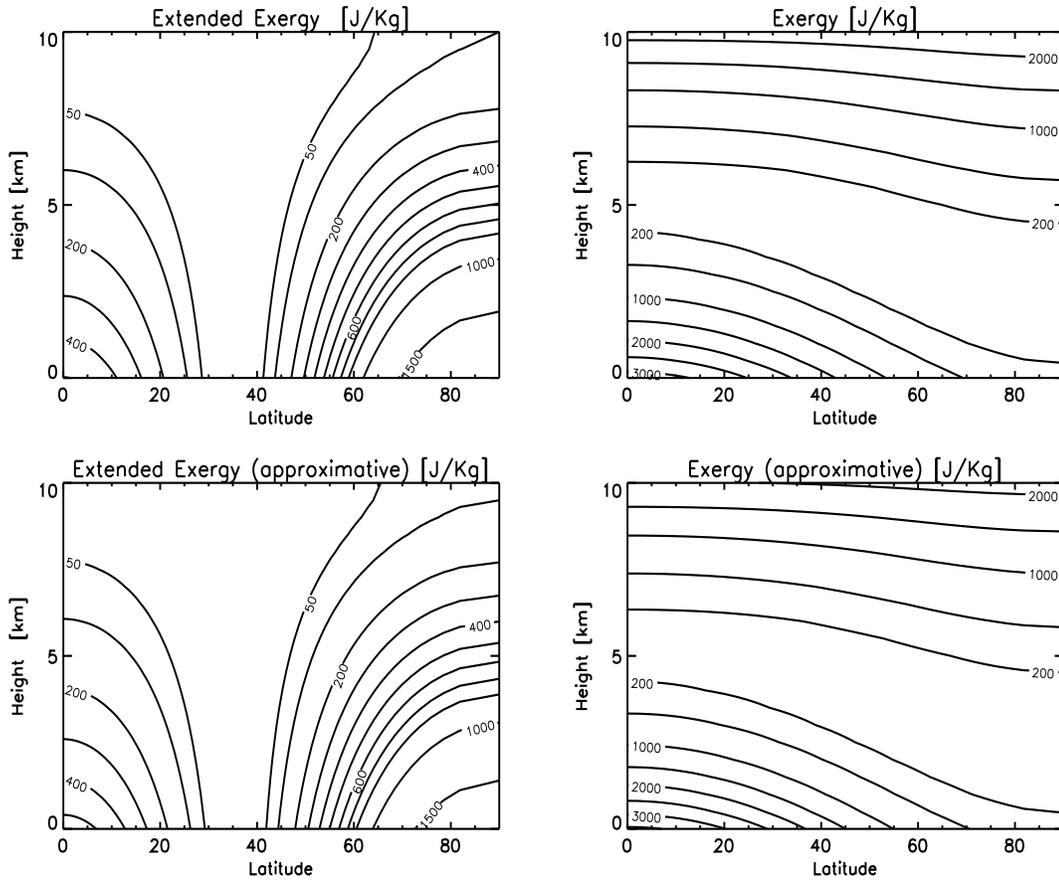


Figure 53: Distribution of classical Exergy and extended Exergy (local available potential energy). Units are J/kg.

It is further possible to decompose the kinetic and available potential energy in their zonal mean and eddy components (see Section 9.1), and to derive their evolution equations. If we identify the global volume averages of k and e_{ape} and their mean and eddy components as E_K, E'_K and E_P, E'_P then the resulting global

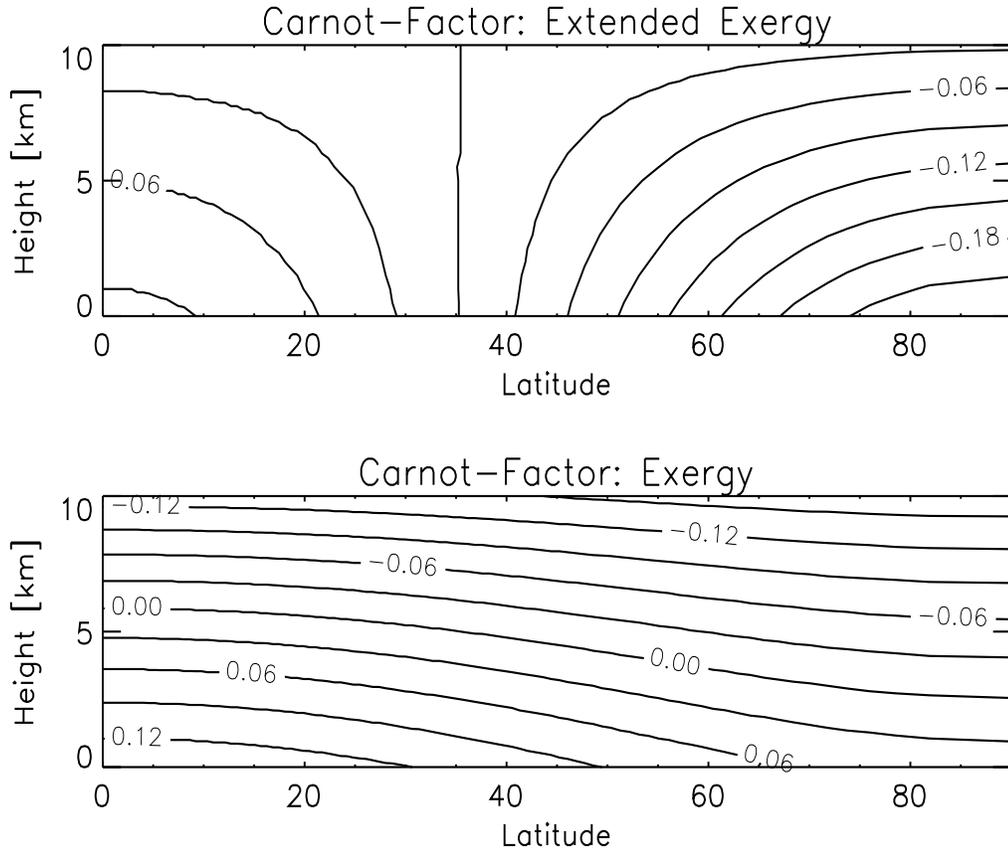


Figure 54: Distribution of Carnot factors of Exergy and extended Exergy (local available potential energy).

energy cycle according to Peixoto and Oort (1983) is shown in Fig. 55. As can be seen most production going into E_p , but also a considerable part into E'_p . These are related to heating in equatorial regions and cooling in polar regions for E_p , and diabatic heat release in cyclones in case of E'_p . The energy is then transformed into eddy kinetic energy, E'_K , from where most of the dissipation occurs. A smaller part is transferred into mean kinetic energy E_K , and eventually dissipated and also transformed back into E_p .

For the ocean we have a much simpler picture, mainly driver by atmospheric winds and friction:

Exercises

1. Show that for a hydrostatic atmosphere the volume integral of potential energy is proportional to the volume integral of internal energy, i.e. $P = R/c_v U$.

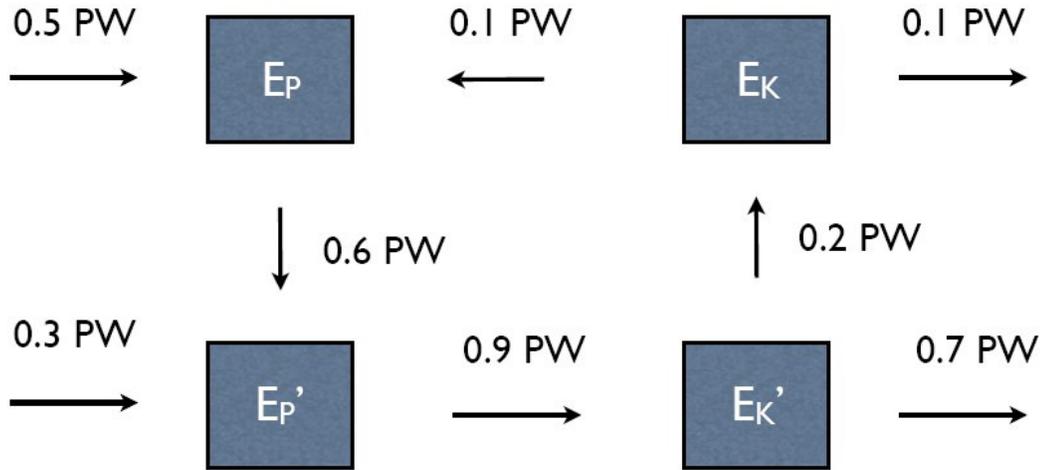


Figure 55: The global atmospheric energy cycle for the global integrals of mean and eddy available potential and kinetic energies (E_P , E_P' , E_K , E_K'), respectively.

(Hint: Use a partial integration!)

$$P = \int \int_A \int_{Z_s}^{\infty} \rho g z dz dx dy = \frac{R}{g} \int \int_A \int_0^{p_s} T dp dx dy + \int \int_A \frac{\Phi_s p_s}{g} dx dy$$

2. Show that

$$T - T_0(s) \approx \left(\frac{g^2}{c_p N_R^2} \frac{\Delta \theta}{\theta_R} + T_R \frac{R}{c_p} \frac{\Delta p}{p_R} \right)$$

in first order by developing $T(s, p)$ and $T_0(s)$ in Taylor series around a reference state s_R, α_R . Hint: Note that $\partial T / \partial s = T / c_p$, $\partial T / \partial p = \alpha / c_p$ and

$$\frac{dT_0}{ds}(s_R) = \frac{dT_R}{dz} \left(\frac{ds_R}{dz} \right)^{-1}$$

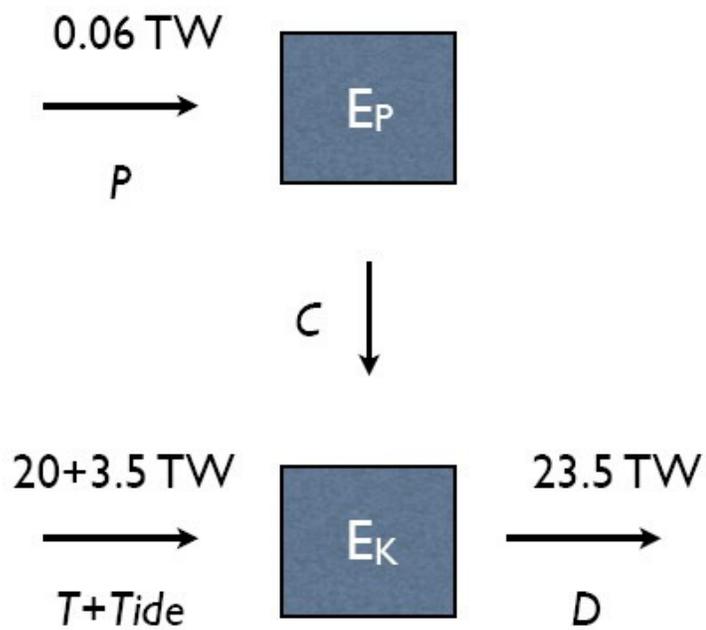


Figure 56: The global oceanic energy cycle therefore the global integrals of available potential and kinetic energies.

12 Moisture budget equation and application to Sahel Drought

- a) de Groot and Mazur, Non-Equilibrium Thermodynamics, Dover Publications, Inc., New York, 501 pp.
- b) Charney, J.G., 1975: Dynamics of deserts and drought in Sahel. *Q J Roy Meteorol Soc*, **101**, 193–202
- c) Zeng N., Neelin J.D., Lau K.-M., Tucker C.J., 1999: 'Enhancement of inter-decadal climate variability in the Sahel by vegetation interaction. *Science*, **286**, 1537-1540
- d) Xue, Y., et al., 2016: 'West African monsoon decadal variability and surface-related forcings: second West African Monsoon Modeling and Evaluation Project Experiment (WAMME II).' *Climate Dynamics*, **47**, 3517-3545.
- e) Kucharski, F., Zeng, N., Kalnay E., 2012: A further assessment of vegetation feedback on decadal Sahel rainfall variability. *Climate Dyn.*, DOI:10.1007/s00382-012-1397-x

Let us consider the general equation for conservation of the water mass in an infinitesimal volume of air (de Groot and Mazur)

$$\rho \frac{dm_v}{dt} = -\nabla \cdot \mathbf{F}_{m_v} - S_{m_v} , \quad (265)$$

where m_v is the mass fraction of water vapour $m_v = \rho_v/\rho$ (or specific humidity), \mathbf{F}_{m_v} is the diffusive water vapour flux, and S_{m_v} are the local sinks or sources of water vapour (e.g. local precipitation, or evaporation in the air). A typical near-surface specific humidity distribution is shown in Fig. 57. Discuss the distribution! Why do we see maximum values in some specific regions?

Using the continuity equation 135, Eq. 265 can be reformulated as (Exercise!)

$$\frac{\partial \rho_v}{\partial t} = -\nabla \cdot (\rho_v \mathbf{v} + \mathbf{F}_{m_v}) - S_{m_v} . \quad (266)$$

If we vertically integrate this equation, we get

$$\frac{\partial \int_0^\infty \rho_v dz}{\partial t} = -|\rho_v w|_0^\infty - |F_{m_v}^z|_0^\infty - \int_0^\infty \nabla_h \cdot (\rho_v \mathbf{v}_h + \mathbf{F}_{m_v}) dz - \int_0^\infty S_{m_v} dz . \quad (267)$$

The first term on the rhs vanishes, because $w(z=0) = 0$ and $\rho_v(z=\infty) = 0$. Furthermore, assuming $F_{m_v}^z(z=\infty) = 0$, and $F_{m_v}^z(z=0) = E$ (surface evaporation flux), $\int_0^\infty S_{m_v} dz = P$ (precipitation). Assuming further a stationary state (or rate of change in time to be small), as well as the horizontal specific humidity diffusion to be negligible, the vertically integrated moisture budget becomes

$$P = E - \int_0^\infty \nabla_h \cdot (\rho_v \mathbf{v}_h) dz . \quad (268)$$

This is an important equation, which is often used in the analysis of the origin of precipitation. It tells us that any local precipitation must come from either local evaporation or the vertical integrated moisture flux (horizontal) convergence, and obviously we have $P = E$ for global integrals of it. Equation 268 can be further split for a more detailed interpretation

$$P = E - \int_0^\infty (\rho_v \nabla_h \cdot \mathbf{v}_h + \mathbf{v}_h \cdot \nabla_h \rho_v) dz \quad . \quad (269)$$

The first term of the integral may be interpreted in the way that any horizontal convergence (or almost equivalently positive vertical motion, why?) in the presence of moisture will lead to a precipitation increase. This is interesting, because it combines nicely with the fact that a parcel that rises high enough will necessarily reach supersaturation and therefore water vapour will condense. The second term in the integral may be interpreted as moisture advection from wetter regions to dryer regions. This second term is less important than the first one, but plays sometimes also an important role. Note also that often Eq. 268 or 269 are evaluated in pressure coordinates using $\int \rho \psi dz = - \int \psi dp/g$, leading to

$$P = E - \int_0^{ps} \nabla_h \cdot (m_v \mathbf{v}_h) dp/g \quad . \quad (270)$$

Fig. 58 shows the annual mean distribution of evaporation and precipitation. Discuss the important features of these distributions. Where is evaporation maximal and where is precipitation maximal?

12.1 An application to Sahel drought

The Sahel Drought is one of the most important and most studied climatic events in the 20th century. Fig. 59 shows the world's desert regions and clearly identifies the Sahara as the largest and brightest desert regions.

The Sahel region is the region at the southern edge of the Sahara Desert, and has the characteristics of modest, but highly variable rainfall as indicated by mean rainfall, decadal standard deviation and Coefficient of Variation (Figs 60 and 61a,c).

If we concentrate now on Sahel rainfall (averaged rainfall in the region 15° W to 37.5° E, 12.5° to 17.5° N), then we see striking decadal rainfall variations compared to interannual variations (Fig. 62).

As can be seen from the time series in the 1970s to the 1990s prolonged drought periods have been observed, but a much wetter period in the 1950s and 1960s. An intriguing feature for scientists (but problematic for the Sahel population) is the long duration of the dry periods, with sometimes nearly 10 years of continuous negative or positive anomalies (note that no filtering has been applied to this time series, only in the red smooth curve). As said above, many scientific studies have been performed to address the physical mechanisms for the Sahel drought from the 1970s to the 1990s. A general consensus is now that it is induced by sea surface temperature variations, but enhanced by land-surface feedbacks. Also human-induced land-use

changes and aerosols may have played a role (e.g. Xue et al, 2016 for a review). In order to highlight these decadal changes further, we show in Fig. 63 the rainfall difference 1980 to 1994 minus 1950 to 1964, roughly representing the core periods for dry and wet conditions, respectively. It is seen that the drying is a very large-scale phenomenon, but it can also be seen that there are slightly wet conditions to the south of the Sahel.

Here we will focus on the role that interactive vegetation may have played. In order to assess this we will show some results from the analysis of Kucharski et al. (2012), where the vegetation impact on Sahel rainfall variability has been investigated using an AGCM coupled to a dynamic interactive vegetation model (see Fig. 64).

An equation for the dynamic vegetation may be formulated as (Zeng et al., 1999)

$$\frac{dV}{dt} = a\beta_{veg}(W)(1 - e^{-\kappa L}) - \frac{V}{\tau} , \quad (271)$$

where V is the vegetation cover, a is a coefficient, $\beta_{veg}(W)$ is the soil moisture W dependence of vegetation growth (which could be linear), κ is the extinction coefficient of photosynthesis, L is the leaf area index, τ is the vegetation time scales (set to 1 year; Discuss the solution for small L and $L = 0$). There is assumed also a linear relation between L and V , i.e. $L = L_{max}V$. Obviously, there is also an equation for W , which will tell us essentially that soil moisture will grow when precipitation is large, and so further. For us its important to note the main impact for the atmosphere will be the relation between albedo and vegetation cover (or L), which is parameterized empirically as

$$A = 0.38 - 0.3(1 - e^{-\kappa L}) \quad (272)$$

Between what values varies the albedo (reflectivity) according to this equation? Obviously, with this the vegetation cover modified strongly the atmospheric radiative balance.

The first question that has to be addressed is whether the model can reproduce the mean climate in the regions in terms of rainfall and vegetation. Fig. 60, lower panel, shows the modeled rainfall and LAI climatologies. The model represents at least the gross regional characteristics.

Also the model decadal standard deviation and COV (Fig. 61, b,d) show overall spatial characteristics similar to observations. Does the model reproduce the Sahel drought?

Fig. 65 compares the models response for the Sahel drought rainfall difference when vegetation is interactively coupled and prescribed to be constant. Clearly the interactive vegetation version of the model gives a response that is much closer in magnitude compared the the non-interactive version, even though the overall drought signal is present in both simulations.

Let's investigate the feedback mechnism that the interactive vegetation model can provide for drought/wet conditions. Fig. 66 shows several components of the

positive feedback mechanism involved when dynamic vegetation coupling is present in the response. Panel (a) shows the vegetation cover response, and indicates that there is a reduction in vegetation when the precipitation is reduced (see Eq.271). This is leading to an increase in albedo (surface reflectivity) as shown in panel (b) (see Eq. 272). This means that more solar radiation at the surface is reflected (c). The feedback loop is closed by the impact of the reduced net surface radiation on the circulation. Charney (1975) suggested that the reduced surface net radiation will lead to less surface warming, thus inducing high pressure, Ekman divergence, and sinking motion (e.g. moisture flux divergence according to Eq. 269) and thus reduced rainfall. On the other hand, the model without interactive vegetation cannot produce this feedback (panel d). Fig. 67 summarises the *Charney feedback mechanism*.

We will see however, that the feedback mechanism is slightly more complex than initially suggested by Charney, and involves more components, also non-local feedbacks. Fig. 68a and b shows the moisture flux convergence and the evaporation terms of Eq. 268. As can be seen, in the Sahel region local evaporation plays a substantial role in the moisture budget. However, the moisture flux convergence term plays the dominant role, supporting Charney's hypothesis. The reduced evaporation means, however, that at least part of the reduced surface heating is compensated by reduced surface latent heatflux. Panels b) and d) show the moisture flux convergence and evaporation for the model with fixed vegetation. Both are substantially weaker.

Fig. 69a shows the corresponding surface pressure and low-level wind responses for the Sahel. As hypothesised by Charney there is high pressure and low-level divergent response. However, in particular the high pressure response is shifted to the north, so that the advection of dry air from the Sahara will also contribute to the drying (e.g. second term in the moisture flux convergence of Eq. 269). It can also be noted that the dynamic feedback is much stronger in the case with interactive vegetation compared with the non-interactive vegetation (Fig. 69b).

Finally, we compare the model Sahel rainfall time series with the observed (Fig. 70). Whereas the time series from the interactive vegetation run has characteristics similar to the observed rainfall (but slightly weaker magnitude; Fig. 70b), the model without interactive vegetation (Fig. 70c) has much more power at interannual time scales and less persistent anomalies as well as smaller magnitudes.

Exercises

1. Show with help of the full continuity equation that Eq. 266 follows from Eq. 265.

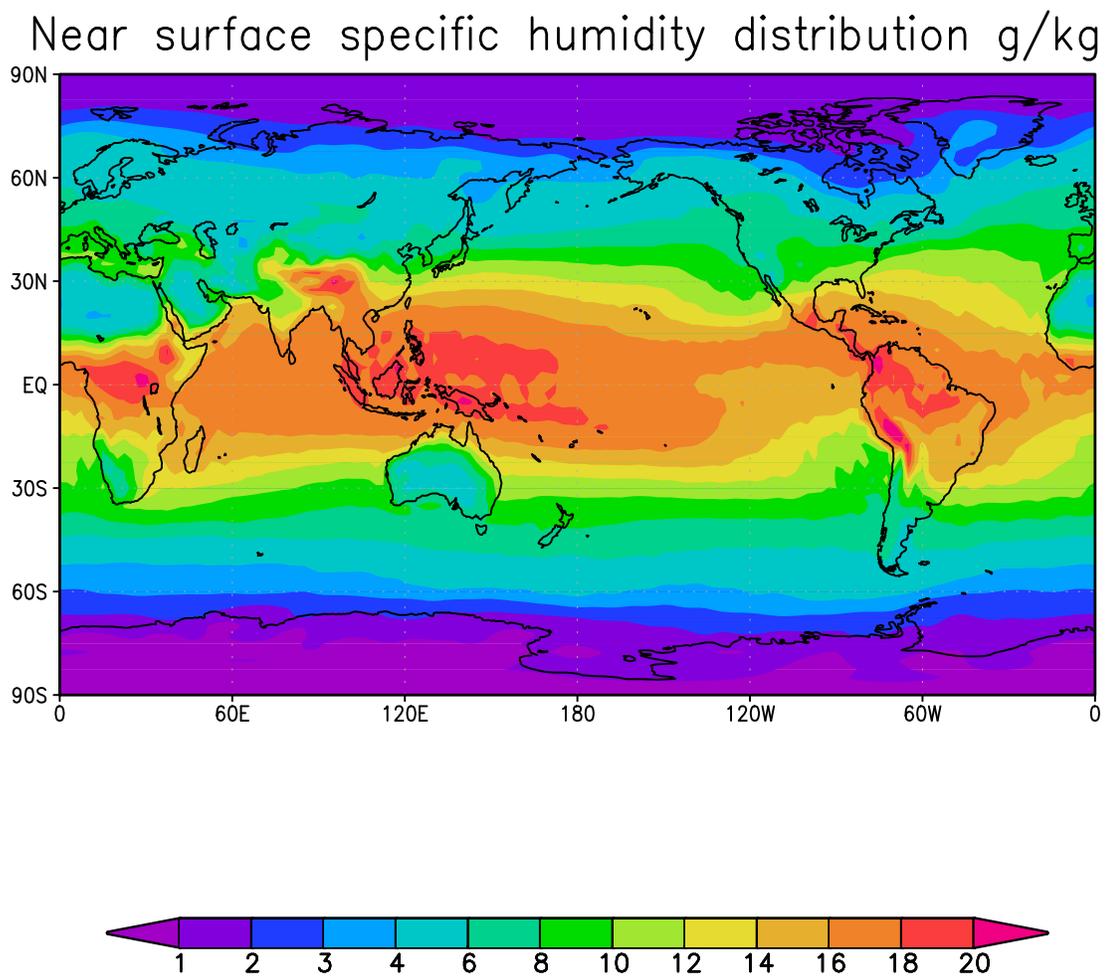
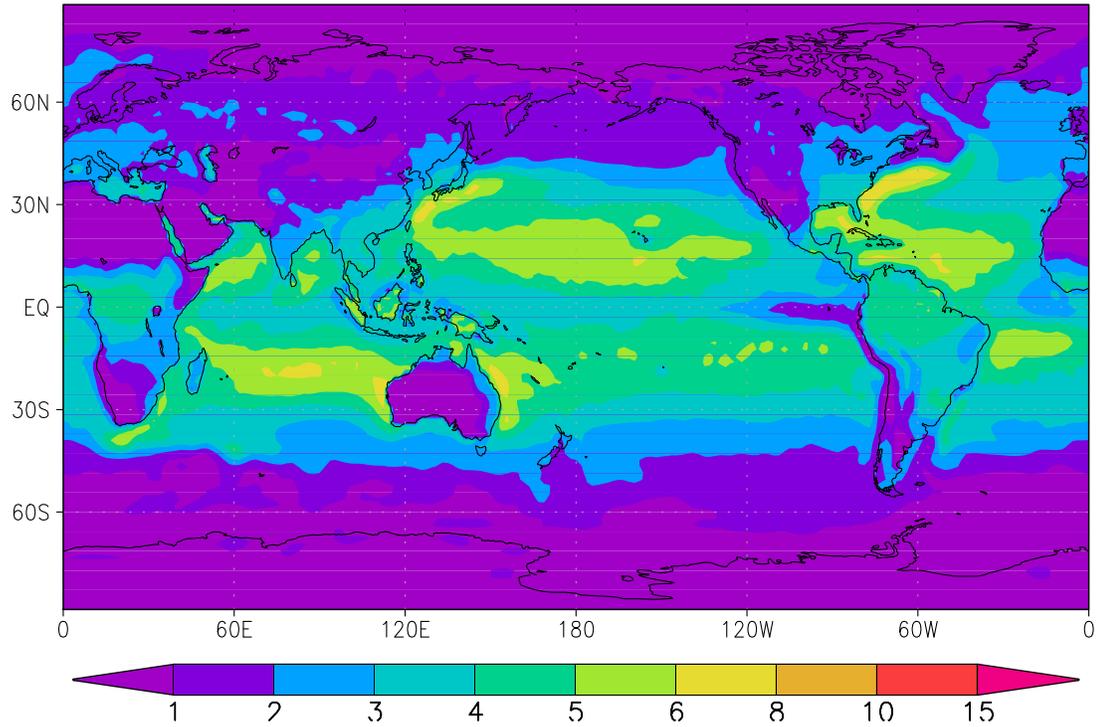


Figure 57: Near surface distribution of specific humidity (vapour mass fraction).
Units are g/Kg.

Annual mean evaporation mm/day



Annual mean precipitation mm/day

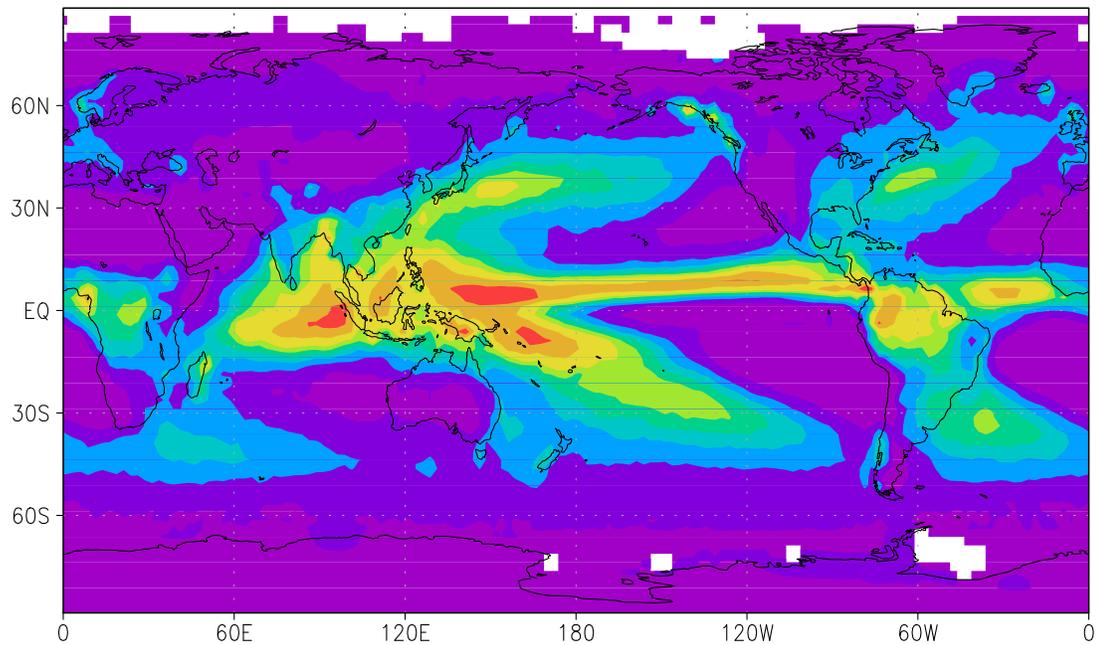


Figure 58: Annual mean observed evaporation (top) and precipitation (bottom).



Figure 59: Vegetations and brightness of the Earth.

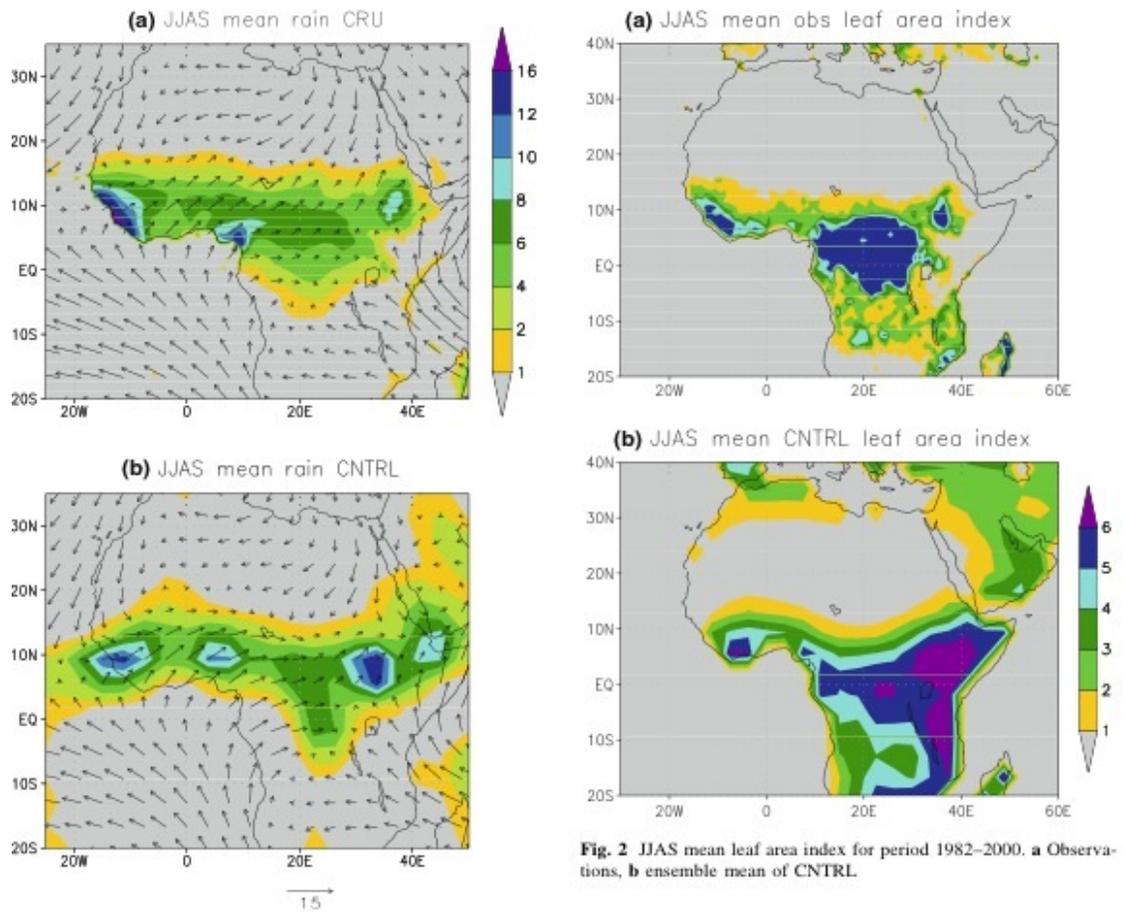


Figure 60: Precipitation in mm/day and Leaf Area Index (LAI; leaf area/ground area in m^2/m^2) of observations (upper panel) and model (lower panel).

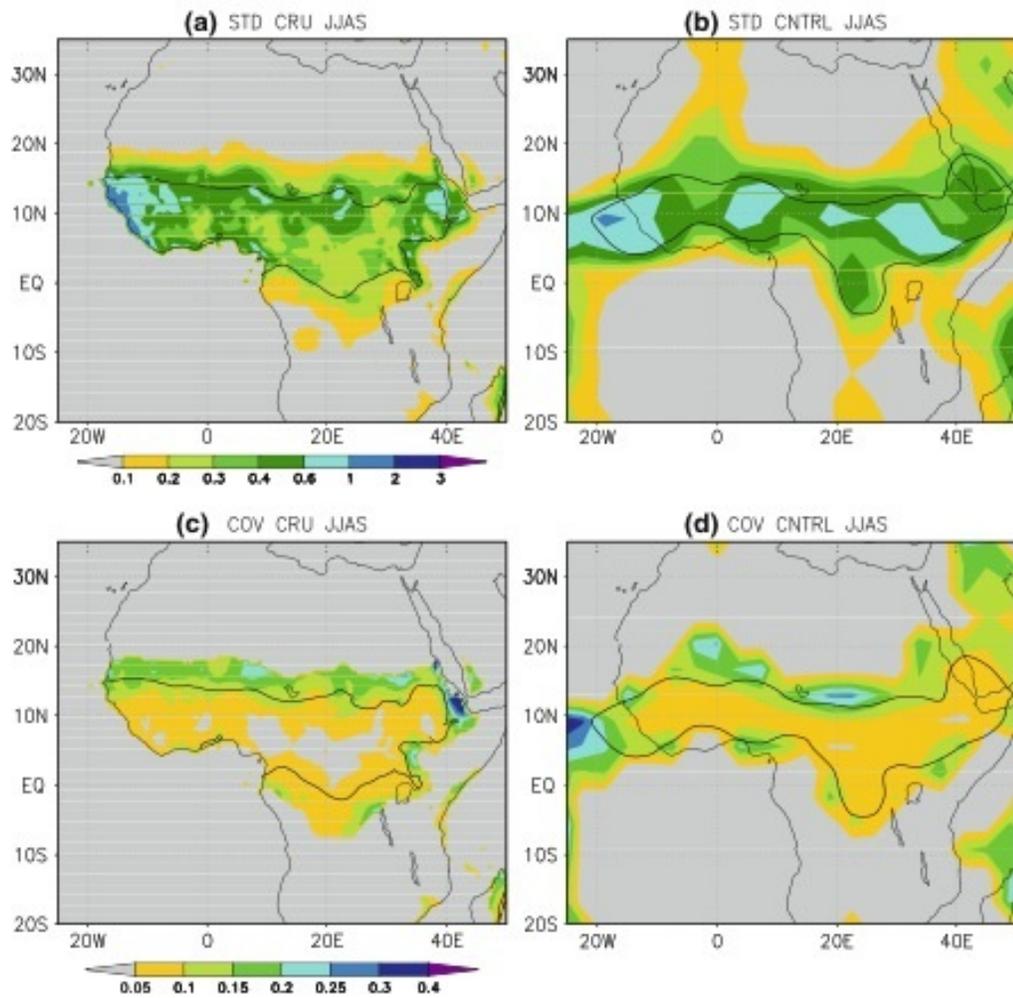


Figure 61: Observed (a, c) and model (b, d) precipitation standard deviations in mm/day and Coefficient of Variation (COV), defined as rainfall standard deviation divided by mean rainfall.

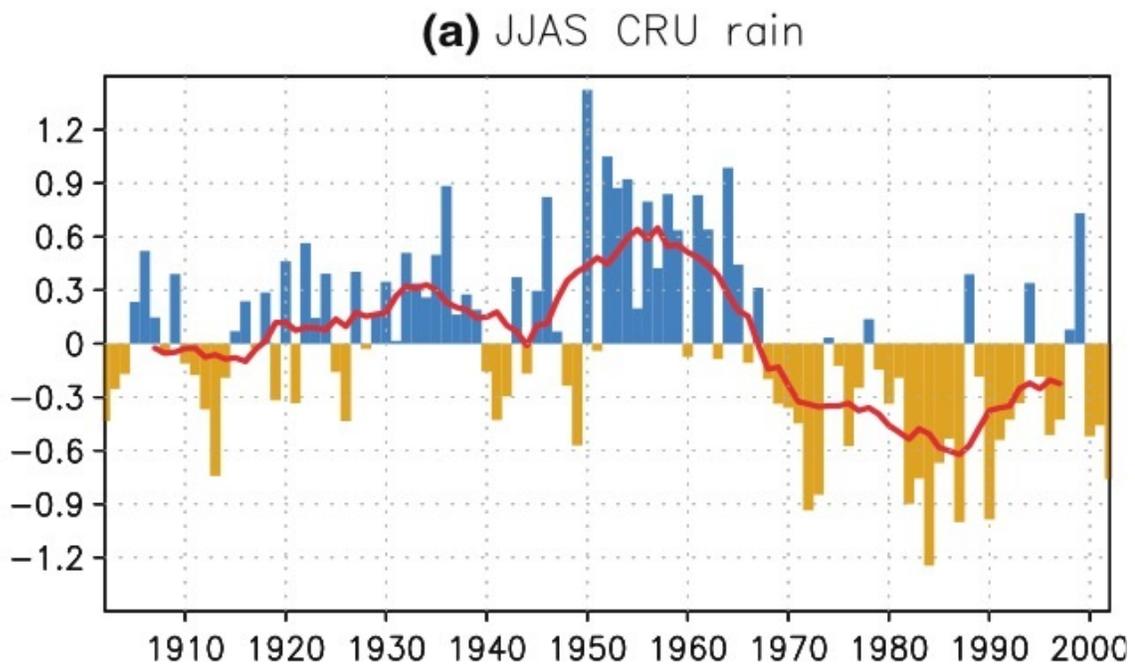


Figure 62: Sahel rainfall anomalies in mm/day.

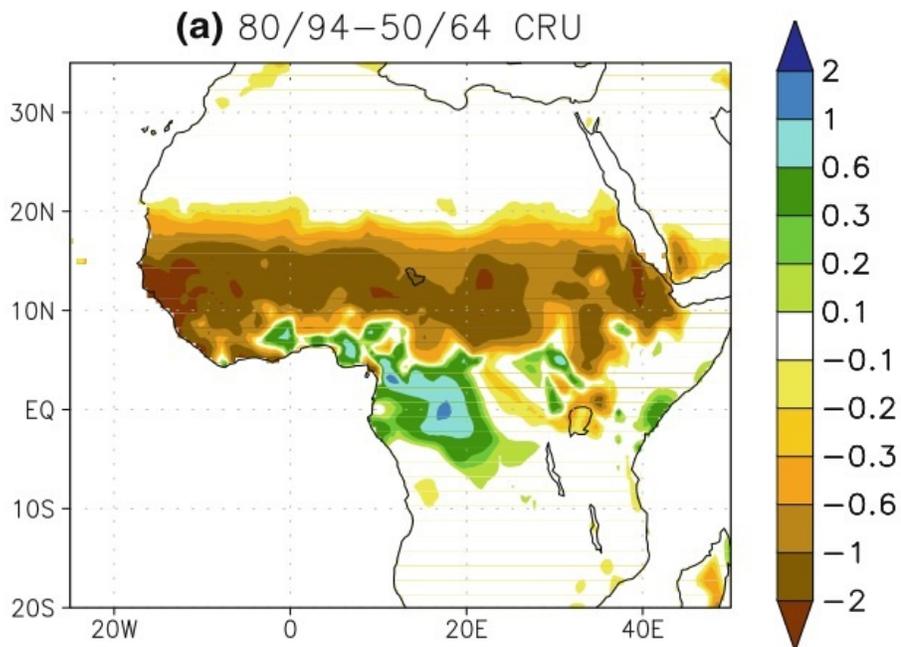


Figure 63: Rainfall difference 1980 to 1994 minus 1950 to 1964 in mm/day.

The UMD-ICTP earth system model (SPEEDY-VEGAS)

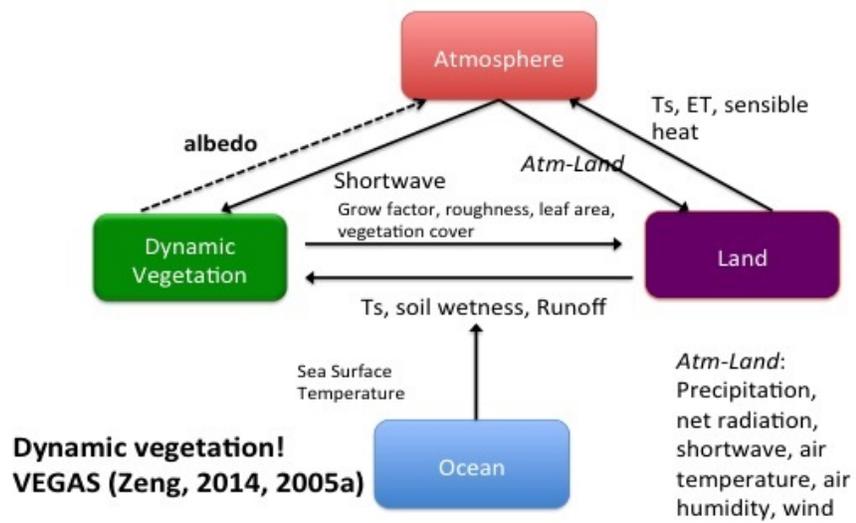


Figure 64: Model set-up.

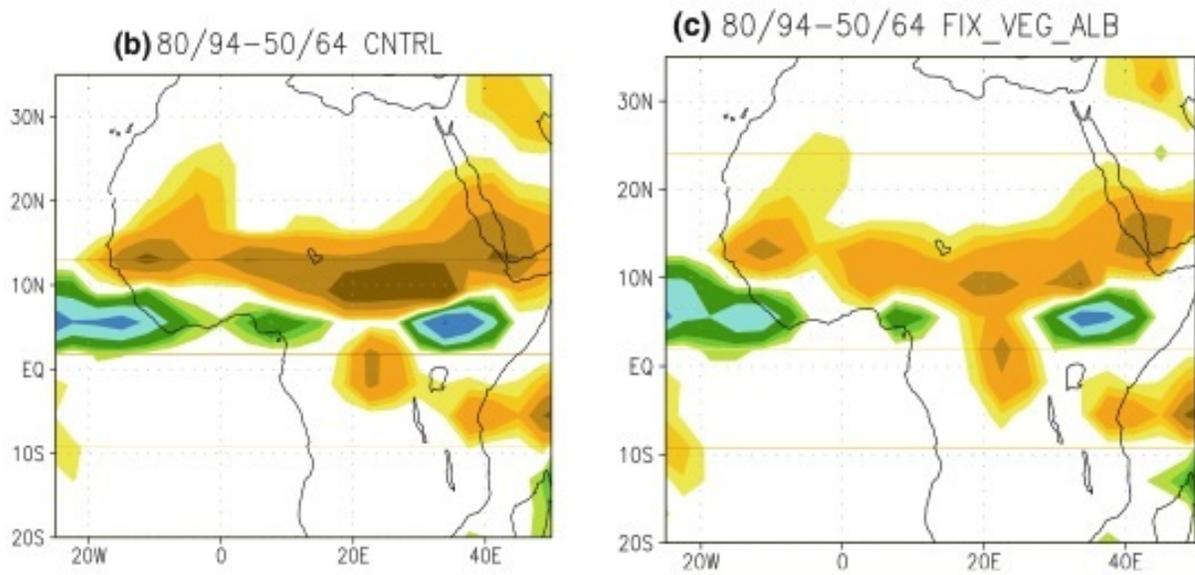


Figure 65: Modeled rainfall difference 1980 to 1994 minus 1950 to 1964 in mm/day. Left: with interactive vegetation, right: without interactive vegetation). of observations and model.

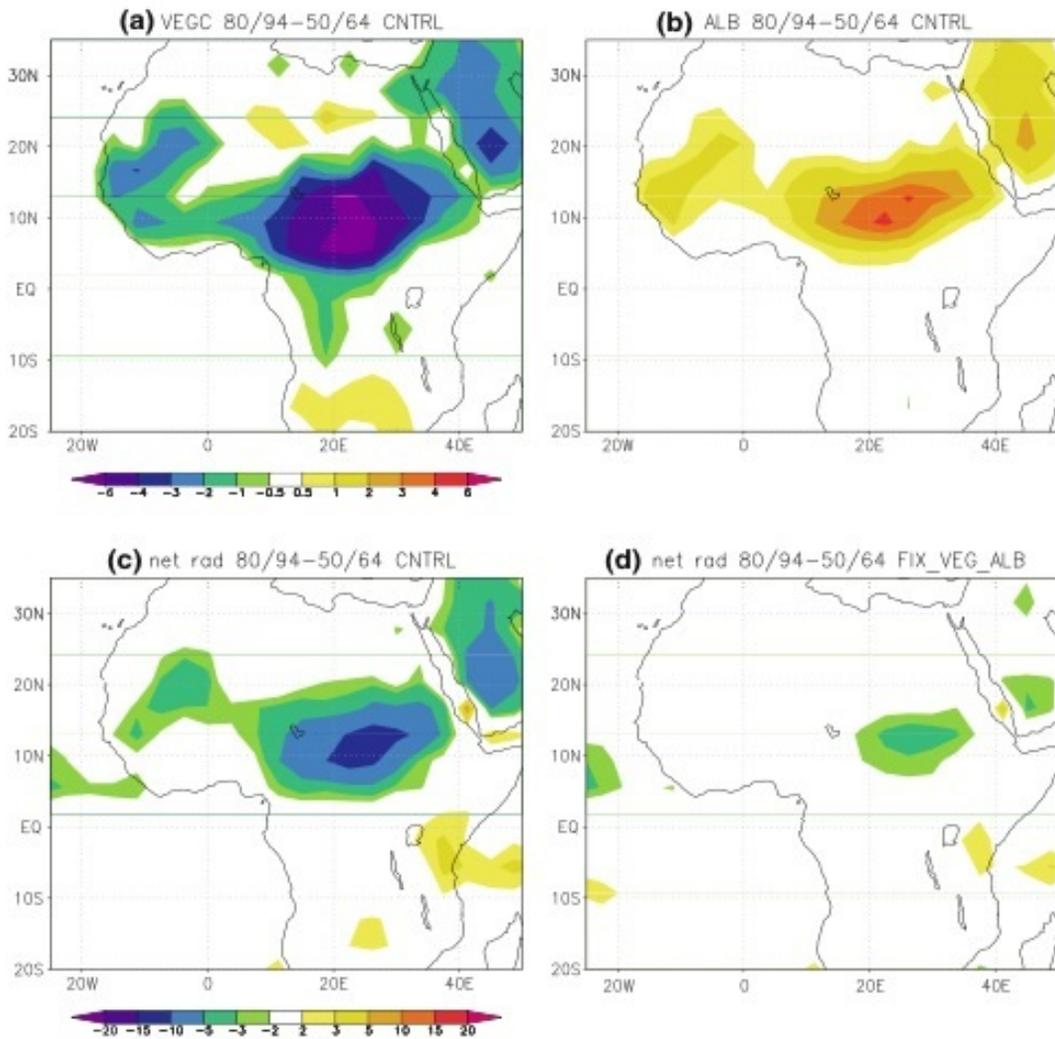


Figure 66: Features of the Charney Feedback mechanism responsible for the responses 1980 to 1994 minus 1950 to 1964 in the model with dynamics vegetation: a) Vegetation cover change [%], b) Albedo change [%], c) net surface radiation change [W/m^2], d) as c) but without vegetation feedback [W/m^2].

Charney and Sud Feedback Mechanisms for the Sahel Drought associated with **overgrazing**

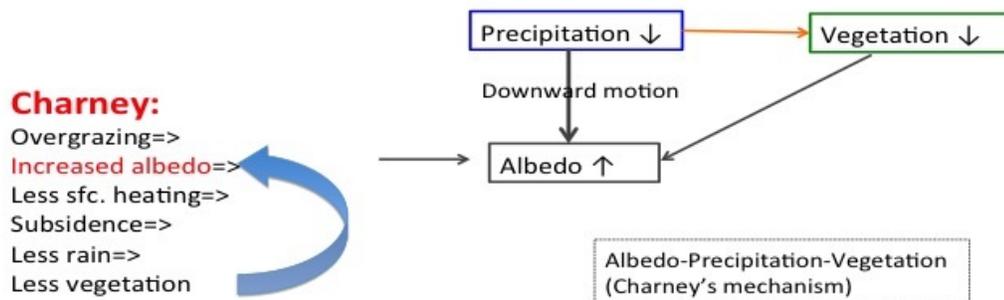


Figure 67: Schematic of the Charney feedback mechanism.

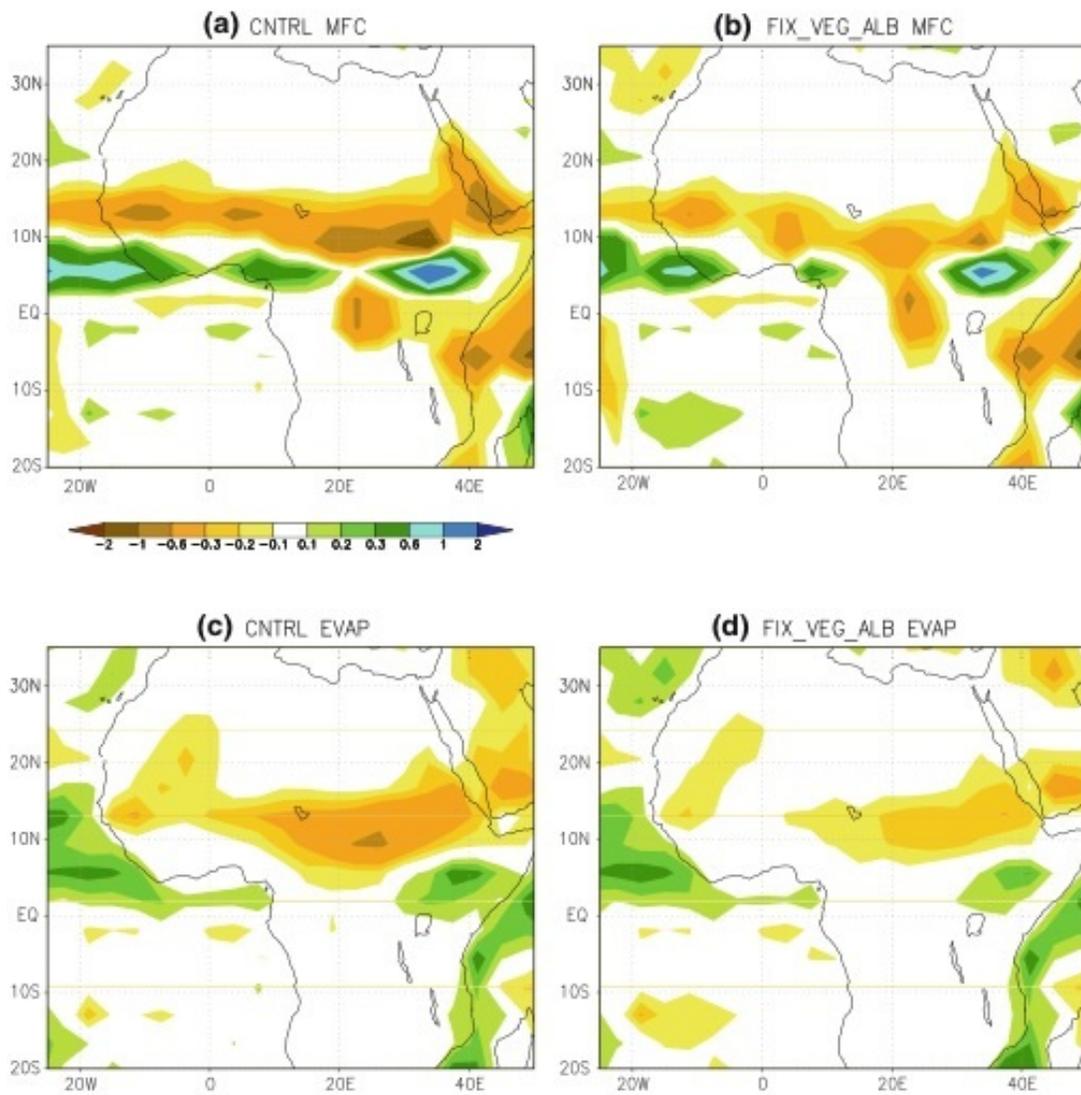


Figure 68: Moisture budget terms in mm/day. a) Moisture flux convergence term of the interactive vegetation, b) moisture flux convergence term of the fixed vegetation model, c) evaporation on the interactive vegetation, d) evaporation of the fixed vegetation

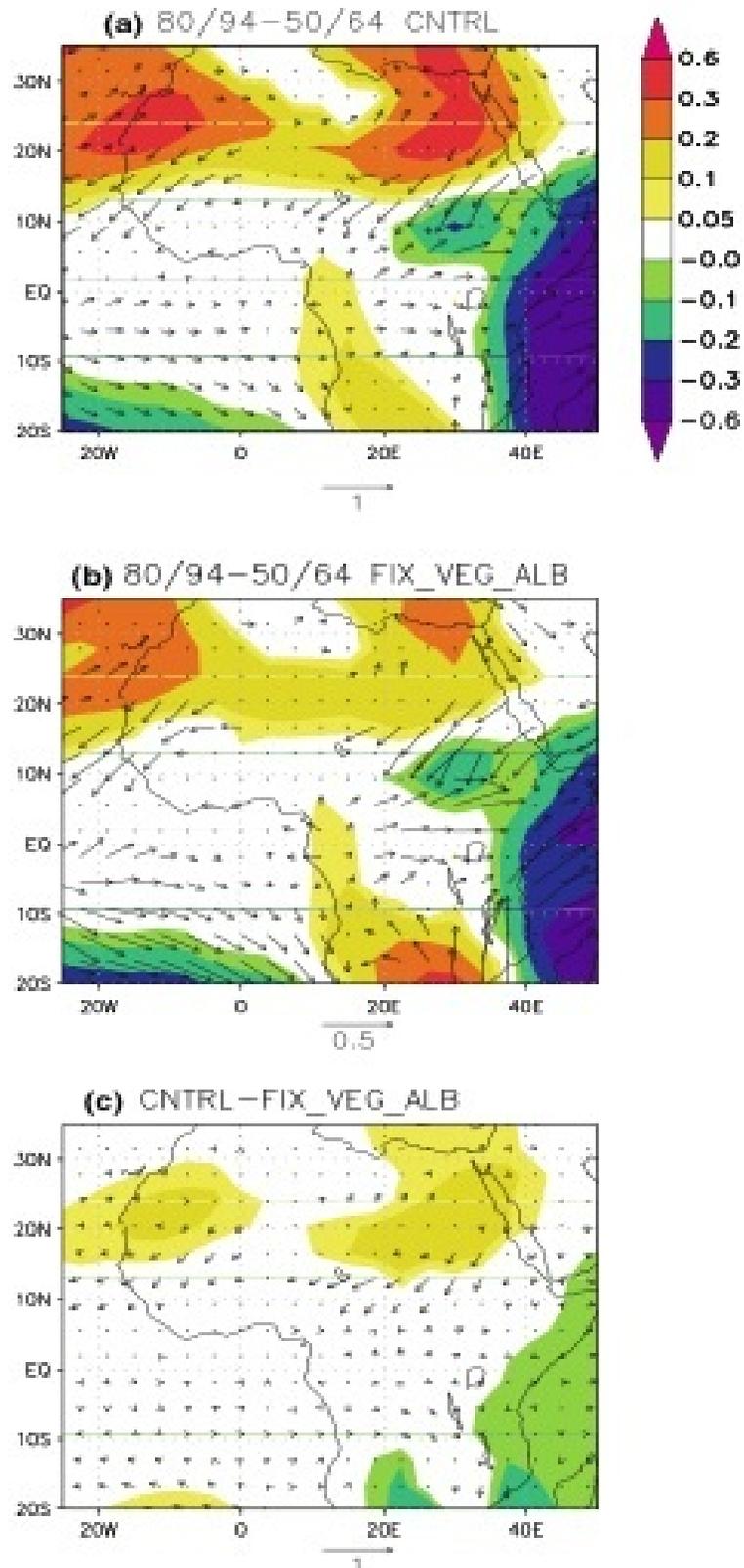


Figure 69: Response of surface pressure [hPa] and low-level winds [m/s] for a) interactive vegetation model, b) non-interactive vegetation model and c) difference between (a) and (b)

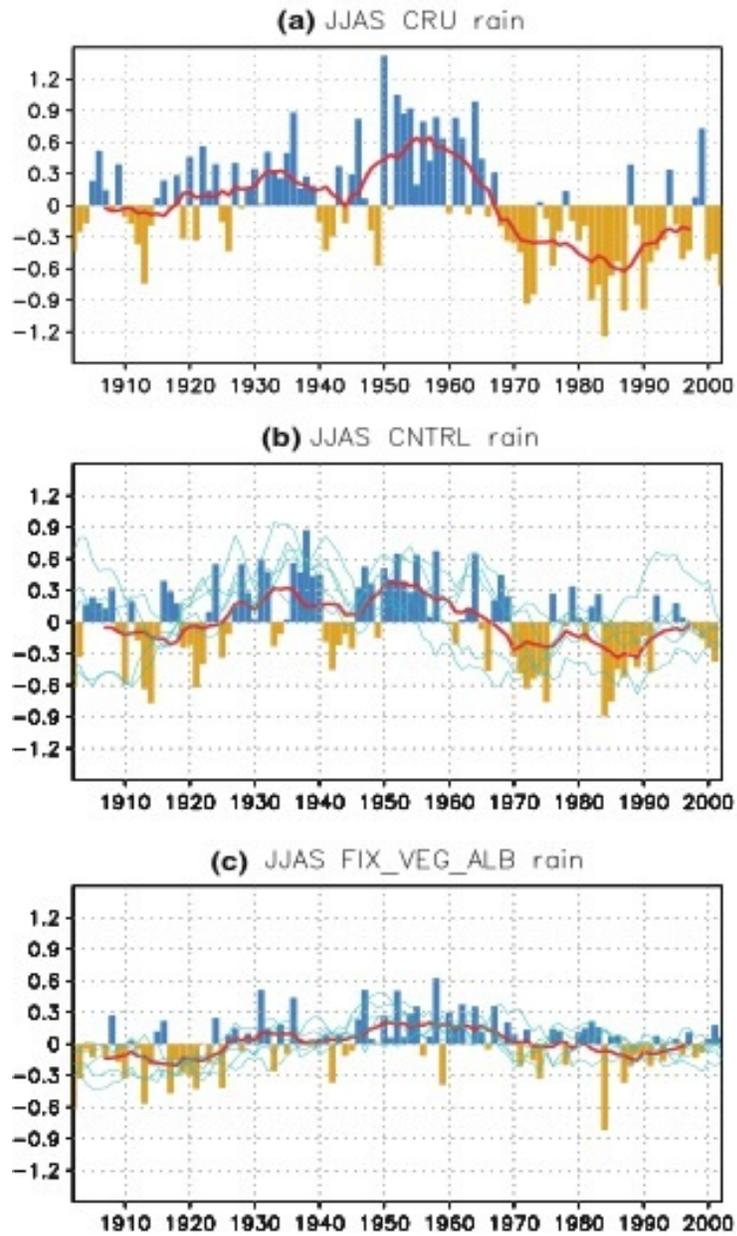


Figure 70: Sahel rainfall anomalies in mm/day. a) Observations, b) model with interactive vegetation, c) model without interactive vegetation.

13 Analysis of Climate Variability: EOF/PCA Analysis

Suggested textbooks:

- a) Statistical Methods in the Atmospheric Sciences. D. S. Wilks, Second Edition, International Geophysics Series, Academic Press, 2006
- b) Statistical Analysis in Climate Research. H. von Storch and F. W. Zwiers, Cambridge University Press, 1999
- c) Analysis of Climate Variability. H. von Storch, A. Navarra (Eds.), Springer, 1995.
- d) Or simply look things up on Wikipedia.....

13.1 Motivation

The problem and necessity of the *analysis of climate variability* becomes clear if we consider the series of 500 hPa winter mean anomaly fields shown in Fig. 71. Lacking a precise theory of what we are seeing (apart from the fact that we know that what we see are solutions of the complex Navier-Stokes equations), how can we try to find some order in the *chaos* that we are confronted with? One way to tackle this problem is the *Empirical Orthogonal Function (EOF)* analysis (guess who introduced this in climate analysis?) or *Principle Component Analysis (PCA)*.

13.2 What does the EOF analysis do?

The EOF analysis solves our problem (how, see below) by finding orthogonal functions (EOFs) to represent a time series of horizontal fields in the following way:

$$Z'(x, y, t) = \sum_{l=1}^L PC_l(t) EOF_l(x, y) \quad . \quad (273)$$

$Z'(x, y, t)$ is the original (anomaly) time series as a function of time (t) and (horizontal) space (x,y), for example the fields that are displayed in Fig. 71. $EOF_l(y, x)$ show the spatial structures of the major factors that can account for the temporal variations of Z' . $PC_l(t)$ are the principal components that tell you how the amplitude of each EOF varies with time. In practice, time and space dimensions are discretized (as in the Numerical Methods Course!). Therefore, dealing with $Z'(x, y, t)$ and $EOF_l(y, x)$ means to deal with matrices.

13.3 Some useful specific definitions and notations

In the following, matrices will be denoted by capital boldface roman letters (**A**, **B**, **Y**, etc.). Vectors will be denoted by a lowercase boldface letters. Let's consider the

data matrix:

$$\mathbf{Z} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1k} \\ z_{21} & z_{22} & \dots & z_{2k} \\ \vdots & \vdots & & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nk} \end{bmatrix} \quad (274)$$

In the following we assume that time and space are discretized and time is represented by the columns of this matrix, whereas space is represented by the rows (space (x,y) is just discretized a one vector, i.e. order $f(i, j)$ as one long vector $f(i, 1), f(i, 2), \dots, f(i, M), i = 1, N$, with $N \times M = k$). EOF analysis is based on anomalies, therefore anomaly data has to be defined. In order to define anomalies, a mean has to be defined. This is done in time, meaning a k-dimensional vector of means can be defined by averaging along the columns of the matrix of Eq. 274 (i.e. the time mean at every grid point). This mean has to be subtracted at every time and gridpoint in order to define the anomaly matrix. The mean subtracted is in general different at different gridpoints, but must be the same at a fixed gridpoint. An elegant way to write this is:

$$\mathbf{Z}' = \mathbf{Z} - \frac{1}{n} \mathbf{1} \mathbf{Z} \quad , \quad (275)$$

where $\mathbf{1}$ is a $n \times n$ matrix that contains 1 everywhere which is multiplied with \mathbf{Z} (to confirm, simply try this procedure with a 2x2 matrix!).

With these notations Equation 273 may be re-written in (discretized) matrix notation as

$$\mathbf{Z}' = \sum_{l=1}^k \mathbf{p} \mathbf{c}_l \mathbf{e}_l^T \quad , \quad (276)$$

where $\mathbf{p} \mathbf{c}_l$ is a $n \times 1$ vector and \mathbf{e}_l is a $k \times 1$ vector, therefore the transpose \mathbf{e}_l^T is a $1 \times k$ vector. Note that the product of an arbitrary $n \times 1$ vector \mathbf{a} and a $1 \times k$ vector \mathbf{b}^T results in

$$\mathbf{a} \mathbf{b}^T \equiv \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \dots & b_k \end{bmatrix} \equiv \begin{bmatrix} a_1 b_1 & a_1 b_2 & \dots & a_1 b_k \\ a_2 b_1 & a_2 b_2 & \dots & a_2 b_k \\ \vdots & \vdots & & \vdots \\ a_n b_1 & a_n b_2 & \dots & a_n b_k \end{bmatrix} \quad (277)$$

If we demand the vectors \mathbf{e}_l to be orthogonal unit vectors, such that $\mathbf{e}_i^T \mathbf{e}_j = 0$ for $i \neq j$, and $\mathbf{e}_l^T \mathbf{e}_l = 1$, then we have

$$\mathbf{Z}' \mathbf{e}_m = \sum_{l=1}^k \mathbf{p} \mathbf{c}_l \mathbf{e}_l^T \mathbf{e}_m = \sum_{l=1}^k \mathbf{p} \mathbf{c}_l \delta_{lm} = \mathbf{p} \mathbf{c}_m \quad , \quad (278)$$

where $\delta_{lm} = 1$ if $l=m$ and zero otherwise. We call $\mathbf{Z}' \mathbf{e}_m$ the *projection* (or in climate analysis sometimes called *regression*) of the data matrix onto the subspace defined by the EOF \mathbf{e}_m . Thus the principle components corresponding the m th EOF can simply be derived by projection of the data matrix \mathbf{Z}' onto the m th EOF. The vector $\mathbf{p} \mathbf{c}_m$ has therefore n components.

13.4 Minimum criterium leading to EOF definition

EOF analysis can be interpreted as a recursive process, we start to determine the first EOF (\mathbf{e}_1), then the second, and so on. The criterion to determine the first EOF is the minimization of the residual

$$\epsilon_1 = \| \mathbf{Z}' - \mathbf{Z}'\mathbf{e}_1\mathbf{e}_1^T \|^2 , \quad (279)$$

with respect to the k dimensional vector \mathbf{e}_1 designing the first EOF in our notation. Here, if \mathbf{Y} is any matrix,

$$\| \mathbf{Y} \|^2 = \frac{1}{(nk)} \mathbf{Y}^T : \mathbf{Y} \equiv \text{tr} \left(\frac{1}{(nk)} \mathbf{Y}^T \mathbf{Y} \right) = \frac{1}{(nk)} \sum_{i=1}^n \sum_{j=1}^k y_{ij}^2 . \quad (280)$$

This means first the matrix product of \mathbf{Y}^T and \mathbf{Y} , then the trace of the resulting matrix by summing up the diagonal elements and this is the total variance of \mathbf{Y} . The normalization by (nk) is arbitrary, but represents the *natural* definition of the total variance. In some cases you may find that the normalization is just done by n , meaning in time. The final results is however independent of this. The meaning of Eq. 279 is that we are searching for a k - dimensional subspace \mathbf{e}_1 to represent the data such that the residual (279) is minimal.

Note that $\mathbf{Z}'\mathbf{e}_1$ is a n -dimensional vector to be matrix multiplied by the k -dimensional vector \mathbf{e}_1^T to give a $k \times n$ matrix according to Eq. 277. Also note that $\mathbf{Z}'\mathbf{e}_1$ is just the definition of the vector of (discretized) Principle Components corresponding to the first EOF in Eq. 276. Some further manipulation leads to:

$$\epsilon_1 = \| \mathbf{Z}' \|^2 - \| \mathbf{Z}'\mathbf{e}_1 \|^2 , \quad (281)$$

which means that minimizing ϵ_1 according to Eq. 279 with respect to \mathbf{e}_1 is equivalent to maximizing the principle component projections

$$\epsilon_{proj} = \| \mathbf{Z}'\mathbf{e}_1 \|^2 \quad (282)$$

with respect to \mathbf{e}_1 (see, e.g. Wikipedia). This leads to the often used 2-dimensional example of the *geometrical* interpretation of EOFs shown in Fig. 73, where samples of 2-dimensional data vectors are considered and we search for the unit vector (EOF) that maximizes the variance of the projection of the data on this vector (straight line).

The minimization (a lot of matrix calculus) leads to the eigenvalue problem

$$\mathbf{S}\mathbf{e}_1 = \lambda\mathbf{e}_1 , \quad (283)$$

where λ is the largest eigenvalue and $\mathbf{S} = \frac{1}{nk} \mathbf{Z}'^T \mathbf{Z}'$ is the $k \times k$ variance-covariance matrix of the anomalies. Therefore the first EOF \mathbf{e}_1 becomes the eigenvector of the matrix \mathbf{S} corresponding to the largest eigenvalue. The other EOFs are found by simply iteratively minimizing the reduced residual

$$\epsilon_2 = \| \mathbf{Z}' - \mathbf{Z}'\mathbf{e}_1\mathbf{e}_1^T - \mathbf{Z}'\mathbf{e}_2\mathbf{e}_2^T \|^2 ' \quad (284)$$

and

$$\epsilon_l = \| \mathbf{Z}' - \mathbf{Z}'\mathbf{e}_1\mathbf{e}_1^T - \mathbf{Z}'\mathbf{e}_2\mathbf{e}_2^T - \dots - \mathbf{e}_l\mathbf{e}_l^T \|^2 , \quad (285)$$

and the results is that \mathbf{e}_2 is the eigenvector of \mathbf{S} that corresponds to the second largest eigenvalue, and \mathbf{e}_l is the eigenvector of \mathbf{S} that corresponds to the l th largest eigenvalue. Since \mathbf{S} has k eigenvectors we can continue this until $l=k$.

13.5 Some further properties

Note that also the principal components are orthogonal, that is $\mathbf{pc}_i \cdot \mathbf{pc}_j^T = 0$ for $j \neq i$. For practice purposes, we hope that a good approximation for the data matrix is given by

$$\mathbf{Z}' \approx \sum_{l=1}^N \mathbf{pc}_l \mathbf{e}_l^T , \quad (286)$$

with $N \ll k$.

A further property is

$$\sum_{l=1}^k \lambda_l = \frac{1}{(nk)} \mathbf{Z}'^T : \mathbf{Z}' = \frac{1}{(nk)} \sum_{i=1}^n \sum_{j=1}^k z_{ij}^2 , \quad (287)$$

which means that the sum of all eigenvalues gives the trace of the variance-covariance matrix \mathbf{S} which is the total variance of \mathbf{Z}' . To evaluate the importance of EOFs it is useful to consider the portion of variance explained by it:

$$\text{expl var of } \lambda_i = \frac{\lambda_i}{\sum_{l=1}^k \lambda_l} \quad (288)$$

A further property of eigenvalues of a matrix is of importance for the practical implementation of the EOF analysis, and is indeed used in the fortran program that you will use in the exercises of this section: If λ is an eigenvalue of the variance-covariance matrix $k \times k$ $\mathbf{Z}'^T \mathbf{Z}'$ (we drop the scaling $1/(nk)$ for here because it is just a factor), then it is also an eigenvalue of the $n \times n$ matrix $\mathbf{Z}' \mathbf{Z}'^T$. In this case the variance-covariance matrix is defined as by the spatial variances and covariances. Thus if $n \ll k$, then we may prefer to find the eigenvalues of $\mathbf{Z}' \mathbf{Z}'^T$. If there are m independent eigenvectors ($\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_s$) of $\mathbf{Z}'^T \mathbf{Z}'$ the eigenvectors of $\mathbf{Z}' \mathbf{Z}'^T$ are $\mathbf{Z}'\mathbf{e}_1, \mathbf{Z}'\mathbf{e}_2, \dots, \mathbf{Z}'\mathbf{e}_s$, which are the projections of the data matrix on the EOFs \mathbf{e}_s which are therefore the (normalized) principal components of the original problem. This means that EOFs and principal components are exchangeable. Instead of calculating the eigenvectors of $\mathbf{Z}'^T \mathbf{Z}'$, we may calculate the eigenvectors of $\mathbf{Z}' \mathbf{Z}'^T$, interpret the eigenvectors as the principal components and calculate the EOFs as projections of the transpose data matrix \mathbf{Z}'^T onto the eigenvectors: $\mathbf{e}_1 = \mathbf{Z}'^T \mathbf{pc}_1$. In this case the principal components are normalized (that is standard deviation = 1), whereas the EOFs are not. In the approximation 286 it does not matter if the principle component or the EOF is normalized, because they are multiplied with each other.

As stated above EOFs are found by determining the eigenvalues and eigenvectors of the variance-covariance matrix. Do you remember how to find these? You have to demand that the *determinant* of the variance-covariance matrix vanishes, this leads to an equation, the *characteristic equation* that contains k th-order polynomials and has at most k roots. There are standard techniques to find eigenvalues and eigenvectors, you may have learned some in your Numerical Methods course?

13.6 Geometric interpretation of PCs and EOFs

The geometric interpretation of the principle components mentioned before is as follows: The eigenvectors empirical orthogonal function (EOF) define a new coordinate system in which to view the data. This coordinate system is oriented such that each new axis is aligned along the direction of the maximum joint variability of the data, consistent with that axis being orthogonal to the preceding one.

The goal is to account for the variation in a sample in as few variables as possible. In the example here, the data is essentially 1-dimensional in the new coordinate system defined by the EOFs.

13.7 Interpretation of EOFs

As we have learned by now, EOFs may be useful to compress the information contained in complex data sets and to structure the data (according to the largest variances). As for the physical interpretation of EOFs, it is tempting to try to give physical explanations to the first few EOFs of a complex data set. Indeed, we expect that if the variability of our fields are governed by a strong low-dimensional physical mechanism (e.g. ENSO in the Pacific region), then one of the first EOFs will reflect this mechanism (indeed in case of EOFs of the interannual variability in the tropical Pacific, we find that the first EOF reflects the canonical ENSO pattern). Unfortunately, the opposite is not true: Not every first (or second or third, ...) is related to a simple and unique physical mechanism! Furthermore it is often even misleading to try to provide a physical mechanism for higher EOFs (e.g. EOF4, EOF5, etc.), because of the orthogonality of the EOFs. This constraint may make higher EOFs less 'physical' than the first or second EOF! The EOF analysis applied to the fields in Fig. 71 gives as first 2 EOFs the maps displayed in Fig. 74. Do you have ideas about possible 'physical' explanations of these EOFs? They are at least well know patterns, do you know their names?

13.8 Related Methods of Climate Analysis

The EOF analysis is probably the most basic of all analysis methods of climatic fields. For example a different question could arise considering 500 hPa geopotential height fields and sea surface temperature fields together. We may ask the question are the 500 hPa fields and the sea surface temperature (SST) fields we see related? This could be due to the fact that one is *forcing* the other. We may get some idea performing an EOF analysis on both fields separately and then try to connect the

emerging EOFs by a physical interpretation (e.g. similar to what we will do in the exercise in this section). We could go one step further and compare (e.g. correlate) the principle components (pcs) of the first EOFs, etc. If we are lucky and the pcs are highly correlated, then there is likely some physical connection between the two first EOFs. However, it could also be that the first pc in geopotential height is a little correlated with the first pc in SSTs and also a little with the second, and so on. This means our interpretation of the connections between 500 hPa geopotential height and SST fields are not much easier after the EOF analysis. There are methods to address this question systematically. For example, the Canonical Correlation Analysis (CCA) or Maximum Covariance Analysis (MCA) provide tools to address the question stated above in a systematic way.

Exercises

1. Using the fortran programme provided, calculate the (winter-mean: DJF) EOFs of a) surface temperature and b) 200 hPa geopotential height in the tropical Pacific. Display the covariance of the resulting principal components related to the first EOF with the global surface temperature and 200 hPa geopotential height fields and interpret the results. How much variance does the first EOF explain in each case? Are the first EOFs of surface temperature and 200 hPa geopotential height related? If yes, what could be the physical mechanism?

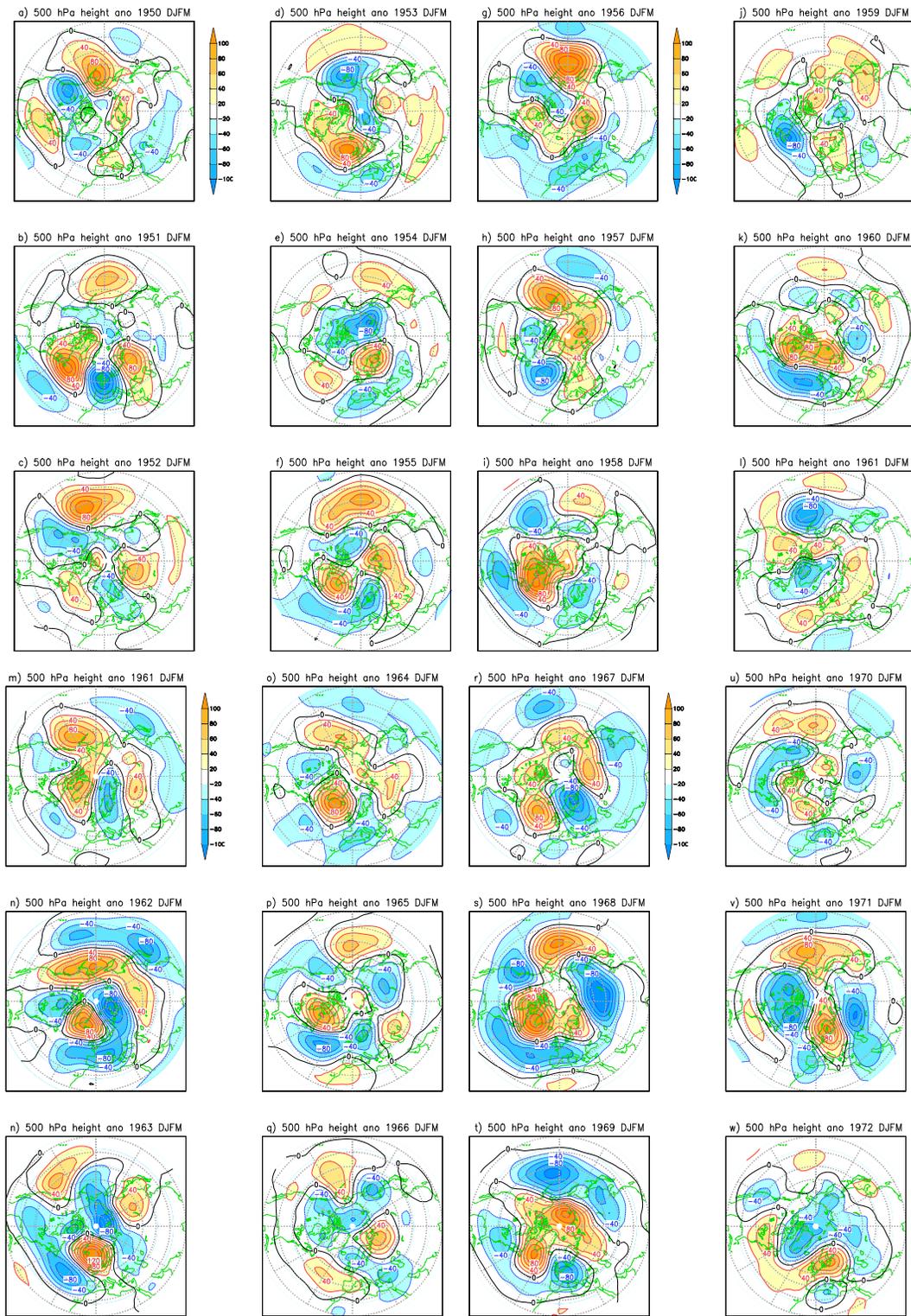


Figure 71: Anomalies of winter 500 hPa height fields for several years. Units are m.

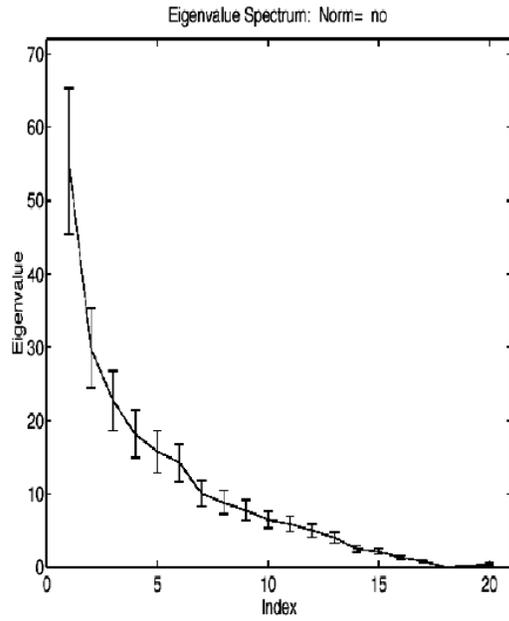


Figure 72: A typical example of the distribution of eigenvalues.

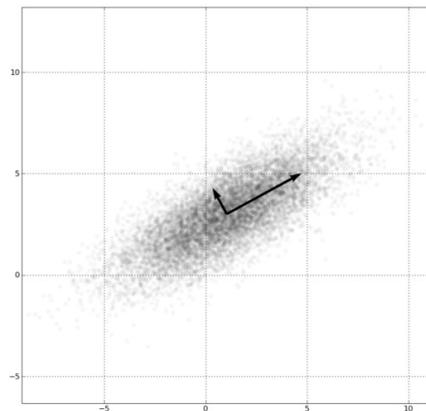


Figure 73: A sample of n observations in the 2-D space $x = (x_1, x_2)$.

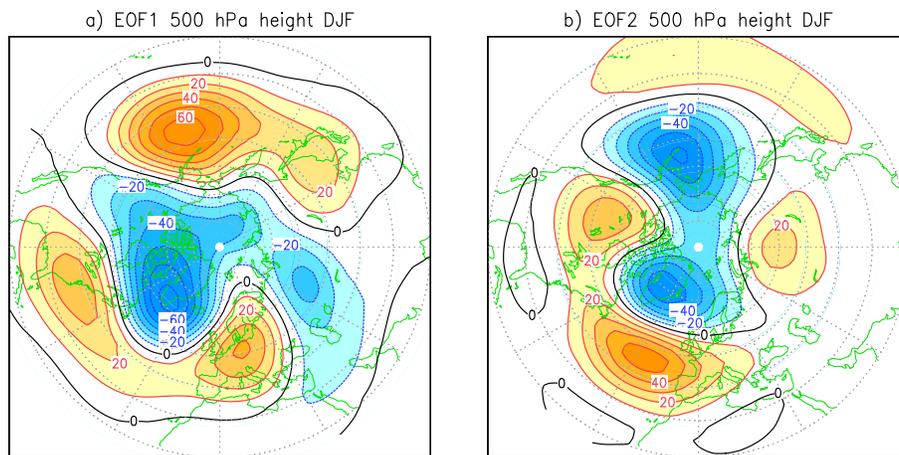


Figure 74: EOFs of the 500 hPa fields presented in Fig . 71

14 Modes of Variability in the Climate System

Suggested Literature:

- a) Hasselmann, 1976: ' Stochastic Climate Models'. Tellus, 6, 473-485
- b) Deser, Clara, Trenberth, Kevin and National Center for Atmospheric Research Staff (Eds). Last modified 2022-09-09 "The Climate Data Guide: Pacific Decadal Oscillation (PDO): Definition and Indices." Retrieved from <https://climatedataguide.ucar.edu/climate-data/pacific-decadal-oscillation-pdo-definition-and-indices> on 2023-09-21.
- c) Hurrell et al. (2003) in The North Atlantic Oscillation: Climate Significance and Environmental Impact, 2003. J.W. Hurrell, Y. Kushnir, G. Ottersen, and M. Visbeck, Eds. Geophysical Monograph Series, 134, 279pp.
- d) Clement, J, et al., (2015) The Atlantic Multidecadal Oscillation without a role for ocean circulation. Science 350, 320-324. DOI:10.1126/science.aab3980
- e) Capotondi, A., McGregor, S., McPhaden, M.J. et al. (2023) Mechanisms of tropical Pacific decadal variability. Nat Rev Earth Environ (2023). <https://doi.org/10.1038/s43017-023-00486-x>

14.1 Mechanisms for low-frequency variability: The Hasselmann model

Hasselmann (1976) suggested a simple model (derived from Damped Brownian Motion) of how atmospheric white noise forcing could generate low-frequency variability through interaction with a static Ocean layer:

$$\rho c H \frac{dT}{dt} = N - \lambda T \quad , \quad (289)$$

where N is the *weather noise*. $\rho = 1000 \text{kg/m}^3$ is the density of water, $c = 4186 \text{J/kg}$ the specific heat capacity of water and $H = 50 \text{m}$ the thickness of the layer considered, $\lambda = 15 \text{W/(m}^2 \text{K)}$ is a damping coefficient, and also represents damping through longwave radiation from the ocean layer. T are temperature anomalies with respect to a reference value (e.g. 290K). Solutions of this simple system are already quite complicated, since N is an arbitrary function of time in principle, and therefore Eq. 289 is non autonomous. If we assume $N = N_\nu e^{i\nu t}$, and $T = T_\nu e^{i\nu t}$ then

$$T_\nu = \frac{N_\nu}{\lambda + i\nu\rho c H}, \quad (290)$$

which leads to the power spectrum

$$|T_\nu|^2 = \frac{N_\nu^2}{\lambda^2 + (\nu\rho c H)^2} \quad . \quad (291)$$

This means that for high frequency noise the resulting amplitude is small, but for low frequency noise the temperature amplitude is larger.

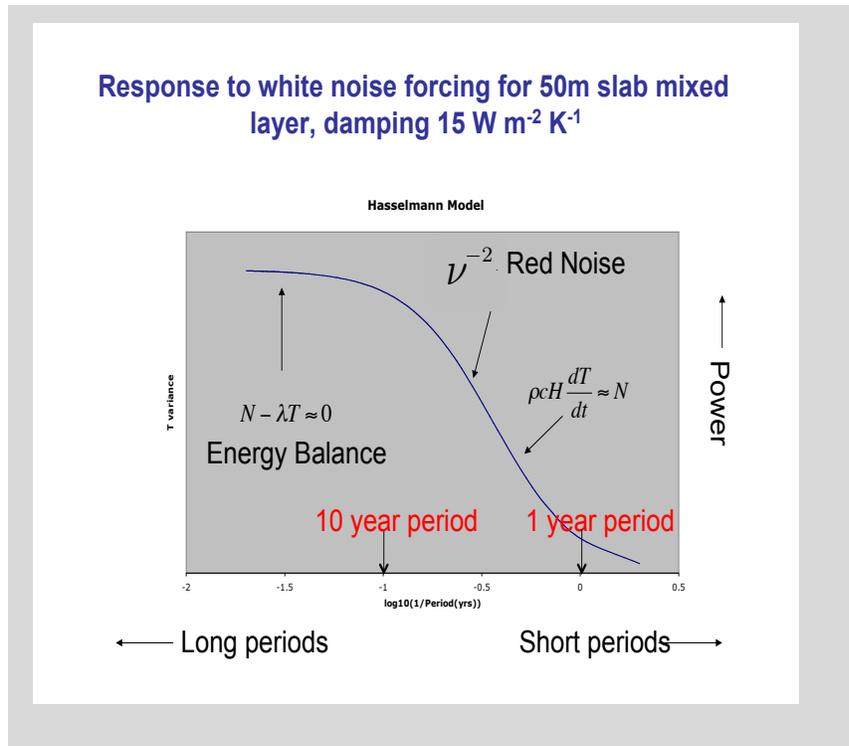


Figure 75: Power spectrum 291 as a function of frequency. Taken from presentation of Ed Schneider at the ICTP workshop on Hierarchical modelling.

14.2 ENSO

The ENSO phenomenon and its mechanisms have been extensively discussed in section 6. This section provides some further information, and in particular its link with other climate modes of variability. Fig. 76 shows the SST anomalies of the ongoing El Nino event. Fig. 77 shows the regions often used to define ENSO time series. Different regions may capture different *flavors* of the ENSO phenomenon. For example, the maximum SST anomaly can be in different locations. Maximum anomalies are typically located in the Nino3 region, but ENSO impacts are typically larger in the maximum anomalies are located in the Nino3.4 region (discuss why). The time series for the Nino3.4 index is presented in Fig. 78. All major El Nino events in the recent history

can be clearly identified, namely 1983, 1998, and 2016. As we can see, also 2023 is on its way to become a major El Nino event, whereas in the last 3 years, La Nina conditions were prevailing. Typical pattern for El Nino and La Nina events are shown in Fig. 79. It can be seen that El Nino events are typically with larger magnitude and further to the east compared to La Nina events (discuss!). An important aspect of ENSO is its relation with global mean surface temperatures. Overall, an El Nino events releases heat to the atmosphere, and this leads to a slow overall warming of other ocean basins as well as land surfaces. In contrast, less than normal heat is released to the atmosphere during La Nina events. As a consequence, the global mean surface temperature shows a clear relation with ENSO, particularly what the global temperature is lagging by about one year (Fig. 80). This relation has let to a lot of discussion in the scientific community, e.g. concerning the slight global cooling in the early 2000s. What is the period of the ENSO phenomenon?

Did you note the SST dipole anomaly (Indian Ocean Dipole or IOD) in the Indian Ocean? It has been recently recognised that a wave train emanating from the Indian Ocean in early winter can generate an important response in the European region. Figure 81 shows a sketch of the response to ENSO and IOD in early and late winter. Whereas in late winter we see a wave train directly induced by ENSO heating anomalies, in early winter the direct ENSO wave train is still weak, but we do see a strong signal coming from the Indian Ocean. This is demonstrated in the composite for IOD events in Fig. 82, left panel, where the stationary wave from the Indian Ocean in December impacts the European region with corresponding precipitation and temperature anomalies (Fig. 83). The ENSO response in February instead is shown in Fig. 82, right panel.

14.3 Pacific Decadal Oscillation/Pacific Decadal Variability (PDO/PDV)

The Pacific Decadal Oscillation (PDO) or its very close relative, the Interdecadal Pacific Oscillation, are basin-wide modes of variability in the Indo-Pacific region. They may be defined through statistical methods, like Principle Component Analysis (PCA; discuss) or indices applied to SSTs (sometimes low-pass filtered). A classical definition for the PDO would be the Pacific SSTs North of 20°N . Fig. 84, top panel shows the pattern and time series of the PDO (derived from an PCA). The lower panel shows the ENSO pattern and time series (again derived from a PCA). As can be seen, pattern and time series are rather similar, with the PDO being lower frequency than ENSO, and PDO pattern more pronounced in the North Pacific. This points to the possible 'null' hypothesis: Let's say ENSO is a given and well explained phenomenon, with its well established teleconnections from the tropical to extra-tropical regions. Could it be that it provides the noise in Eq. 289 (with periods from 4-7 years) and the ocean provides a reddening of the spectrum? This is one hypothesis,

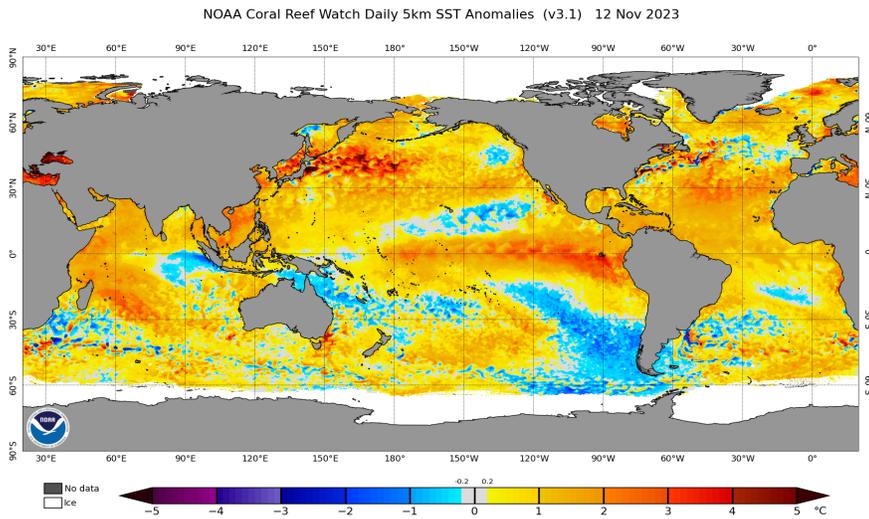


Figure 76: SST anomaly on 12 Nov 2023. Source: <https://www.ospo.noaa.gov/data/cb/ssta/ssta.daily.current.png>.

but it is likely that more than just one process is relevant for the PDO, and also North Pacific Ocean dynamics may play a role to generate a stronger response there, and also in the equatorial parts. For example, it has been suggested that decadal SST variability in the equatorial Pacific could be generated through atmospheric teleconnections to the extratropical Pacific, which modify the Pacific Ocean gyre circulation, thus providing a negative feedback to the initial equatorial Pacific SST signal and leading to a reversal of these anomalies, and thus initiating an oscillation (see Capotondi et al., 2023). For example, Fig. 85 shows how strongly ocean flow convergence and SSTs are related, suggesting an important role of Ocean dynamics in the PDO. Also other Ocean basins (e.g. Indian and Atlantic Oceans) may impact the Pacific and lead to PDO-type variations. A very similar phenomenon compared to the PDO is the Interdecadal Pacific Oscillation (IPO). The IPO is also usually defined through a PCA applied to low-pass filtered (e.g. 11-years), detrended SST in the whole Indo-Pacific region (e.g. 11-years). The PDO/IPO have widespread climatic impacts, particularly in the North American continent, but also for Asian monsoons.

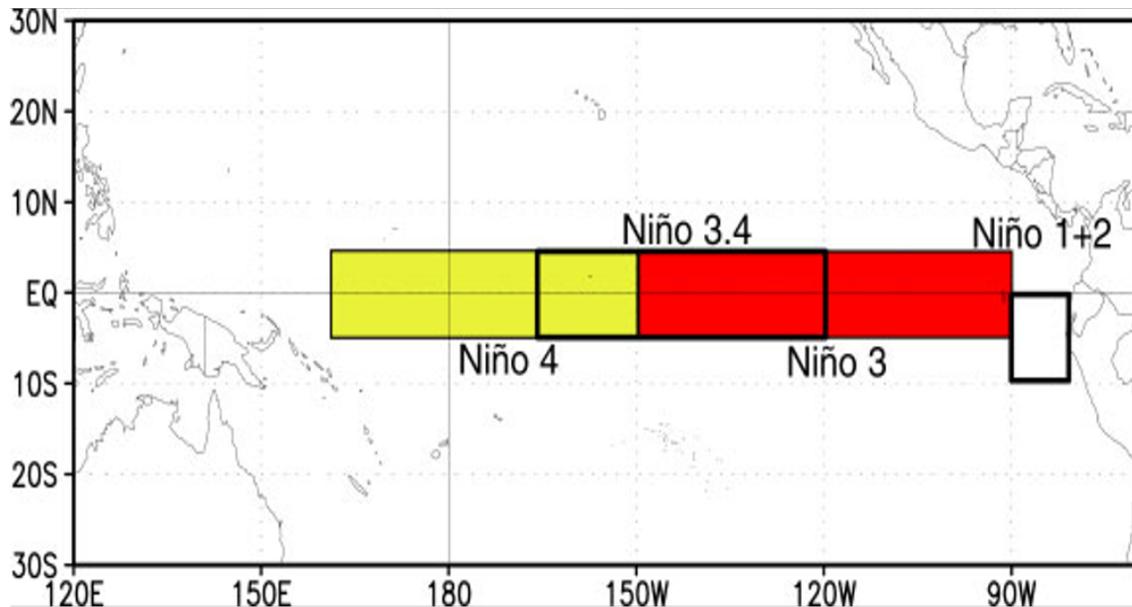


Figure 77: Definition of often used indices to define ENSO time series. Source: http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/lanina/enso_evolution-status-fcsts-web.pdf

14.4 North Atlantic Oscillation (NAO)

The North Atlantic oscillation (NAO) is a dominant pattern of atmospheric variability in the North Atlantic region. It is a pressure see-saw between Iceland and the Azores. It can be defined either in terms of indices or as strongest variability pattern resulting from an PCA. Fig. 86 shows the NAO pattern in Sea Level Pressure and the associated time series. The NAO has important impacts particularly on European winter climate (see Fig. 87 for the surface temperature anomalies associated with the positive NAO phase). The last really cold winters in Europe coincide with strong and persistent negative NAO phases around 2010.

The mechanism for the NAO is rather complicated. It is mainly an internal atmospheric variability mode, but some feedbacks from local and also remote SST forcing may exist. The most prominent theory for the NAO is an eddy-mean state interaction. Rossby wave breaking (e.g. the classical case of breaking of shallow water waves is shown in Fig. 88 and is related that the phase speed of these waves is a function of the height of the fluid) is thought to be responsible for generating positive and negative NAO phases. In case of Rossby waves, the breaking occurs because of meridional variations of the zonal wind (jet). Then there is a positive feedback between the wavebreak-

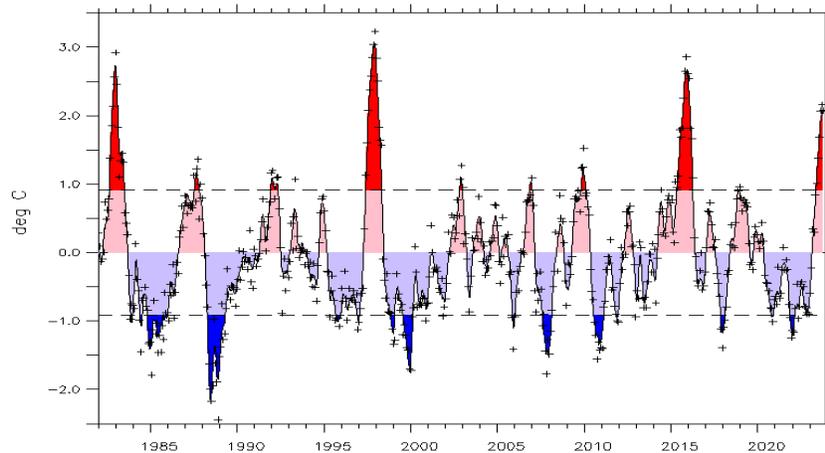


Figure 78: Nino3 SST anomaly timeseries. <https://stateoftheocean.osmc.noaa.gov/sur/pac/nino3.php>

ing and the NAO, because the NAO changes the jet, which causes more wave breaking that favour the same NAO phase, and so on.

14.5 Atlantic Multidecadal Oscillation/Atlantic Multidecadal Variability (AMO/AMV)

The AMO/AMV is the leading mode of variability at decadal time scales in the North Atlantic region. As PDO/PDV it can be defined through PCA, but the more commonly used method is through an SST index covering the whole North Atlantic. The SST field is low-pass filtered and detrended before the analysis, or, alternatively, the global mean SST are subtracted from the time series. Fig. 89a shows the AMO pattern, and Fig. 90 the associated time series. The AMO impacts climate around the Atlantic region and beyond on decadal time scale. Several mechanisms have been proposed for the AMO variability, and there is a heated scientific debate on which is the most relevant. The first one, the 'null' hypothesis, is again an application of the Hasselmann model 289. The (negative) NAO SST imprint (Figs. 89b and 87 show characteristic pattern in SST (also referred to as SST tripole). This pattern does indeed show similarities to the AMO pattern Figs. 89a). However, in the AMO pattern the cooling in the centre is missing. Nevertheless, several studies have argued that the atmospheric forcing from the NAO, plus reddening of the spectrum through the Hasselmann model 289 can largely explain the AMO (e.g. Clement et al., 2015). The second hypothesis is also related to NAO

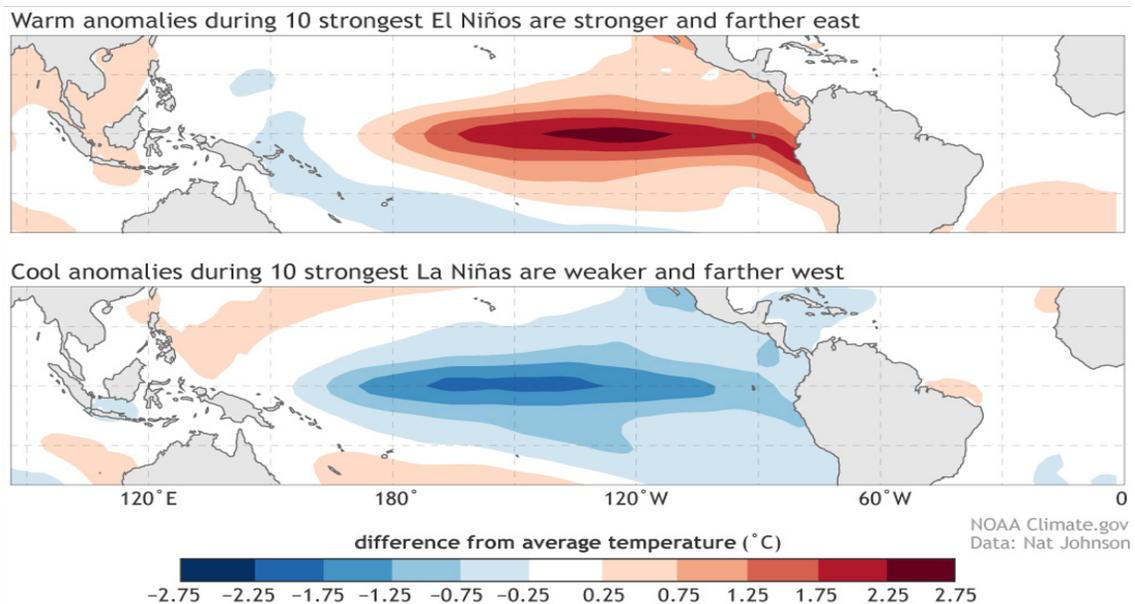


Figure 79: Composites of 10 strongest El Niño and La Niña events.

forcing, but involves dynamically more complicated processes. A positive NAO can lead to an enhanced Atlantic Meridional Overturning Circulation (AMOC; see Riccardo's lectures) though enhances ocean convection close to Labrador Sea. An increased AMOC leads to a warming in the North Atlantic subpolar gyre region. This warming forces in turn the negative phase of the NAO, which then weakens the AMOC, and so on, leading to an oscillation. All the above processes involve delays of several years, and therefore multidecadal oscillations, as shown in Fig. 90 can be explained. The third hypothesis proposed is 'external' forcing, mainly aerosols. This explanation has been proposed because, quite surprisingly, the multimodel ensemble mean of the recent models used for the IPCC reports (e.g. AR5 and AR6) show AMO variability similar to observations, but with much reduced amplitude. The explanation must be external forcing, because a mean of many models would remove internally atmosphere-ocean generated variability, and only the forcing common in each simulation should be visible. It is possible that all 3 process are relevant, and dominate in different time intervals. The observed record of atmospheric and particular Ocean data is not long enough to really decide which is the most important mechanism. Note, that the first 2 mechanisms involve interactions with the NAO, and lead-lag relations between NAO and the AMO can be used to distinguish the two. For the first, it is expected that there is a negative correlation between the NAO and AMO, with the NAO leading by several years. For hypothesis 2 it is expected that the correlations between NAO and AMO are positive when NAO is leading and negative when

Global temperatures anomalies during El Niño, La Niña, and ENSO neutral months

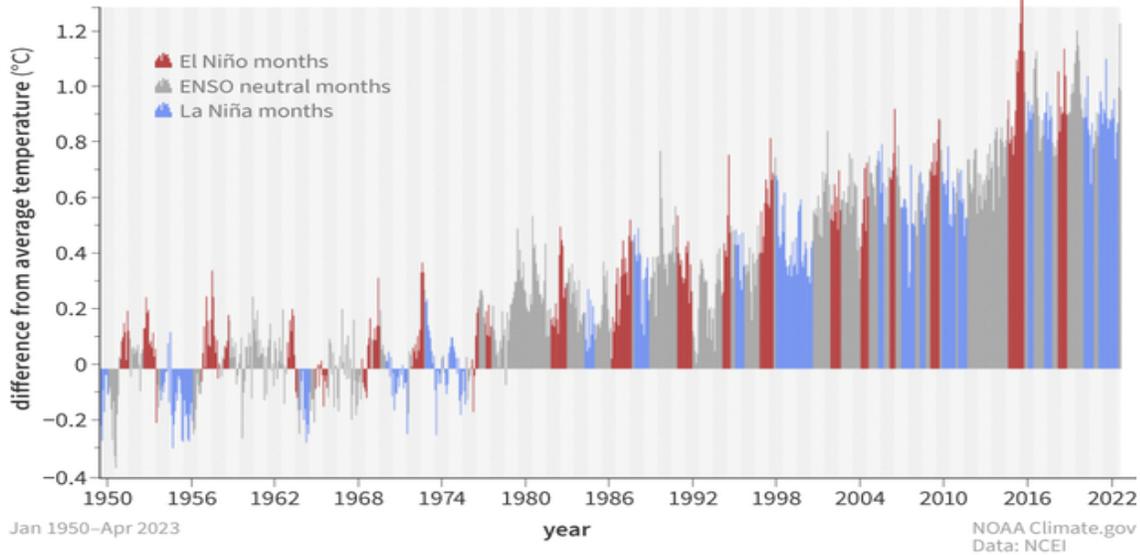
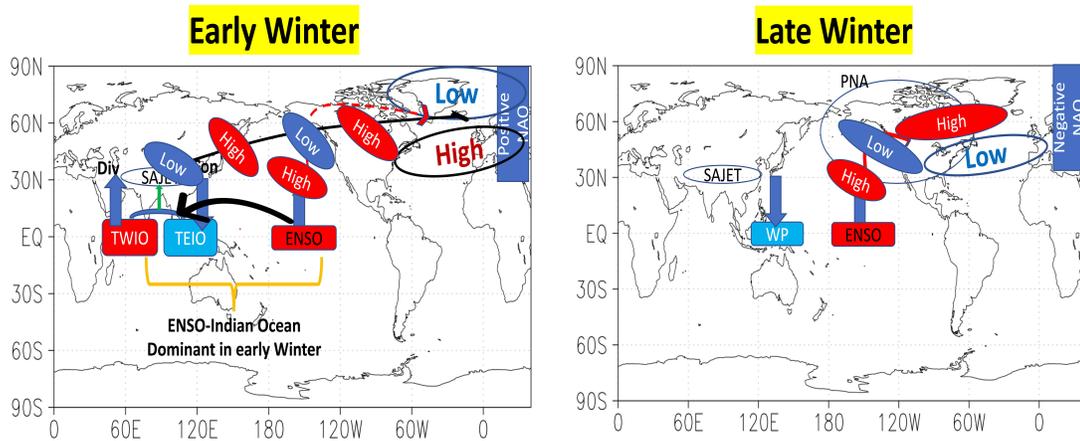


Figure 80: Monthly global surface temperature for January 1950–May 2023 compared to the 20th-century average, colored by monthly ENSO values. NOAA NCEI image. <https://www.climate.gov/news-features/blogs/enso/how-do-noaa-scientists-predict-annual-global-temperature-ranking-ahead>

the AMO is leading. Using a proper definition of indices, there is a clear evidence that hypothesis 2 is strongly supported by observations (see Fig. 91).

Exercises

- Show the validity of 291.
- Given the values in Eq. 289, what is the resulting damping time scale associated with the values of λ .
- Integrate Eq. 289 for $N = N_\nu(\sin \nu_1 t + \sin \nu_2 t)$, where $\nu_1 = 2\pi/180\text{days}^{-1}$ days, $\nu_2 = 2\pi/20\text{days}^{-1}$ and $N_\nu = 50\text{W}/\text{m}^2$. You may use $T = 0\text{K}$ as initial condition. Use a time step of $86400\text{ s} = 1\text{ day}$. Plot the solution and the forcing for 2 years.



Abid et al, 2021, J. Climate, DOI: 10.1175/JCLI-D-20-0075.1

Figure 81: Sketch of early (left) and late winter (right) ENSO and IOD-induced teleconnections.

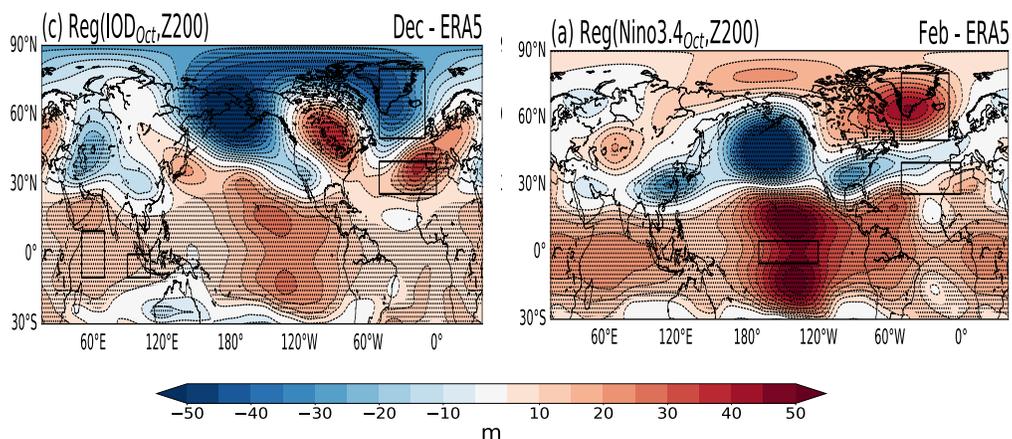


Figure 82: Composite response of 200 hPa geopotential height to October IOD perturbations for December (left) and to ENSO perturbations for February (right). Units are m.

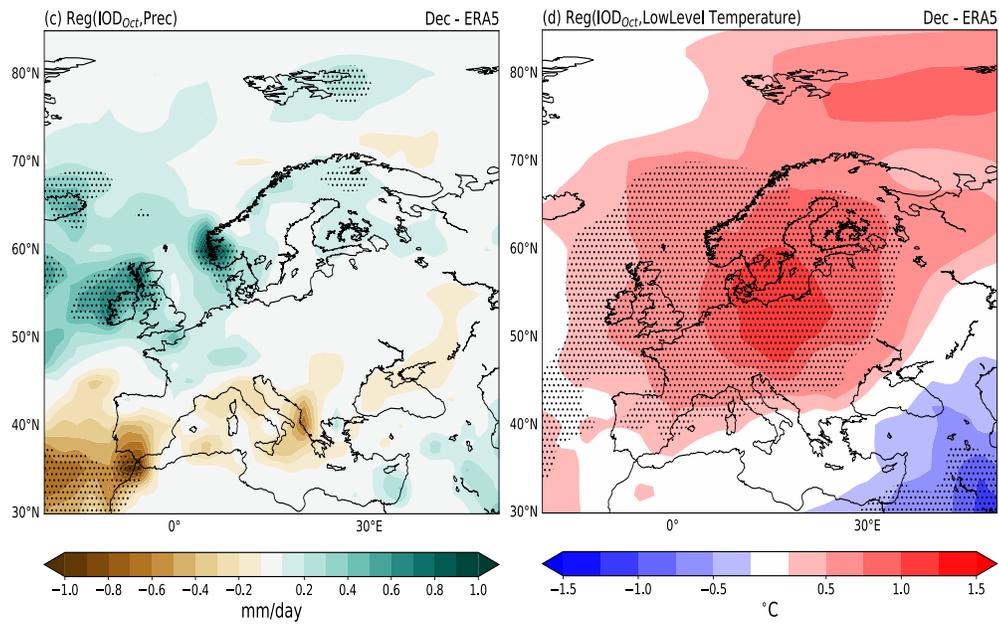


Figure 83: Composite response European precipitation (left) and surface temperature (right) to IOD perturbations. Units are mm/day and K.

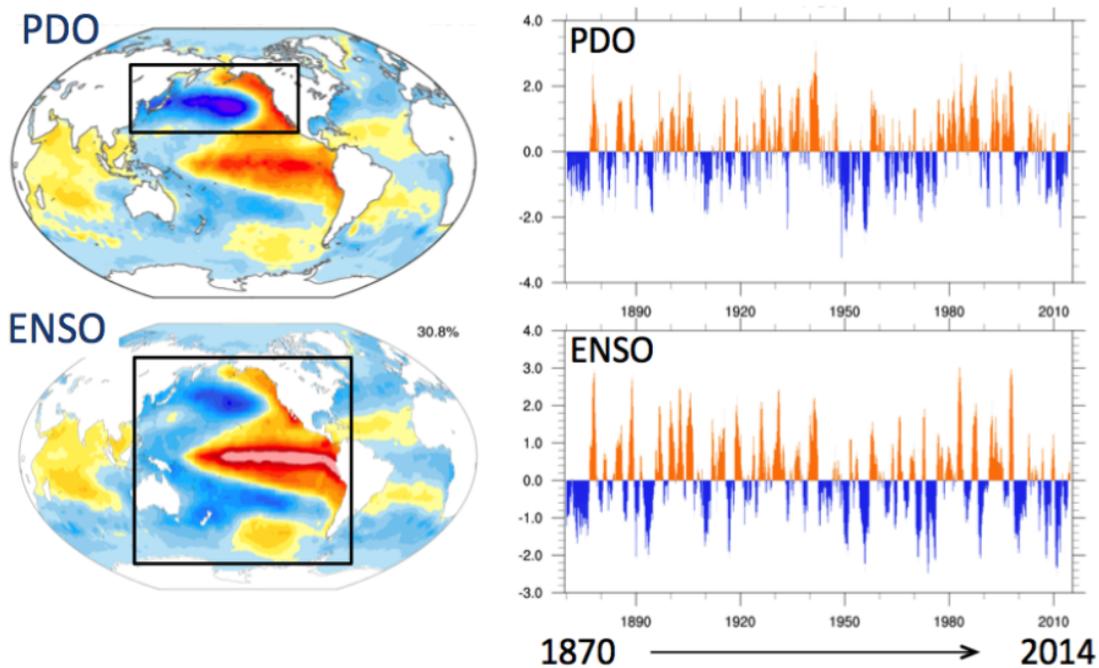


Figure 84: Global expression of the PDO (top left) and ENSO (bottom left), obtained by linearly regressing monthly SST anomalies at each grid box upon the leading Principal Component time series based on the domains outlined in the black boxes. The PDO and ENSO time series, defined by their corresponding PCs, are shown in the top right and bottom right, respectively. Based on the HadISST data set (Rayner et al., 2003) for the period 1870-2014. Adapted from Deser et al. (2010). (Contributed by C. Deser) <https://climatedataguide.ucar.edu/climate-data/pacific-decadal-oscillation-pdo-definition-and-indices>.

d Meridional transport convergence (low-pass filtered)

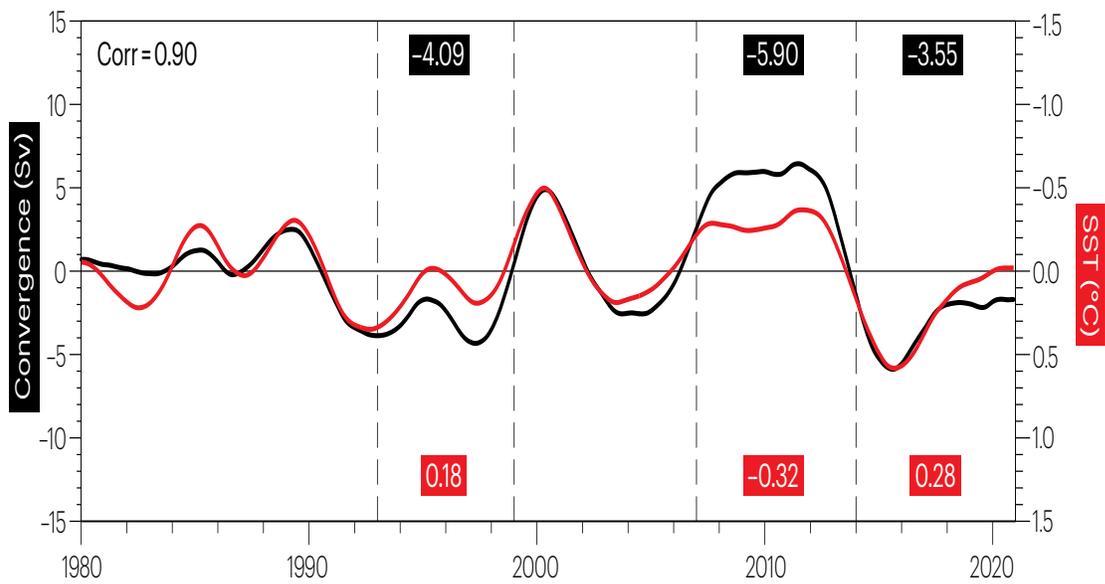


Figure 85: Reanalysis meridional transport convergence anomalies (seasonal cycle removed) across 9.5 N and 9.5 S in the Pacific (black), and SST anomalies averaged over 9.5 N–9.5S, 90 W–180 W (red). Meridional velocity anomalies used to compute transports, and SST anomalies are linearly detrended. Figure from Capotondi et al, 2023

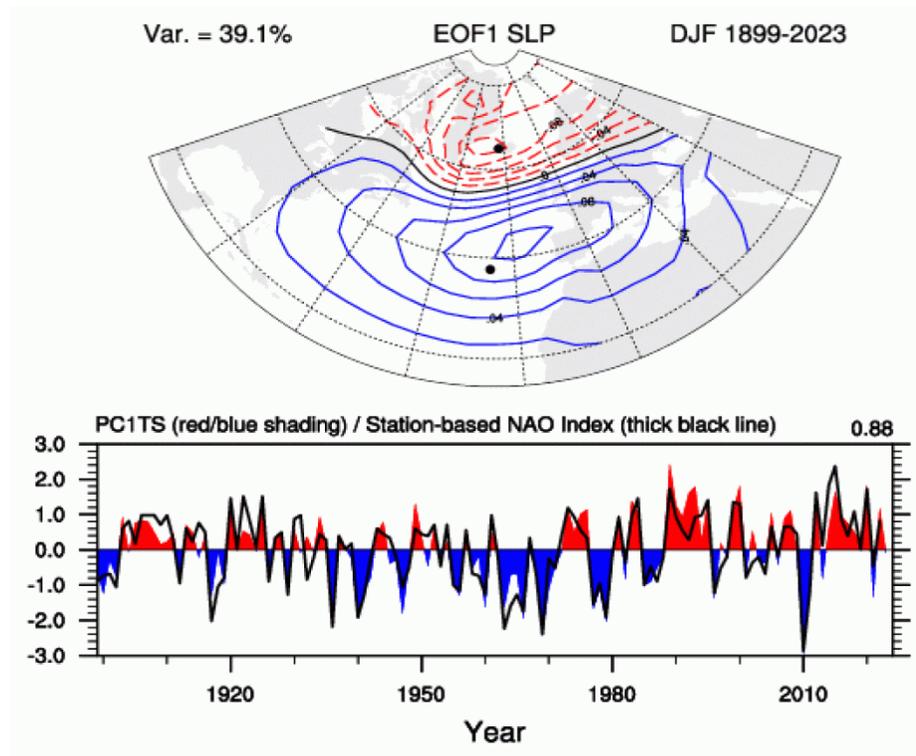
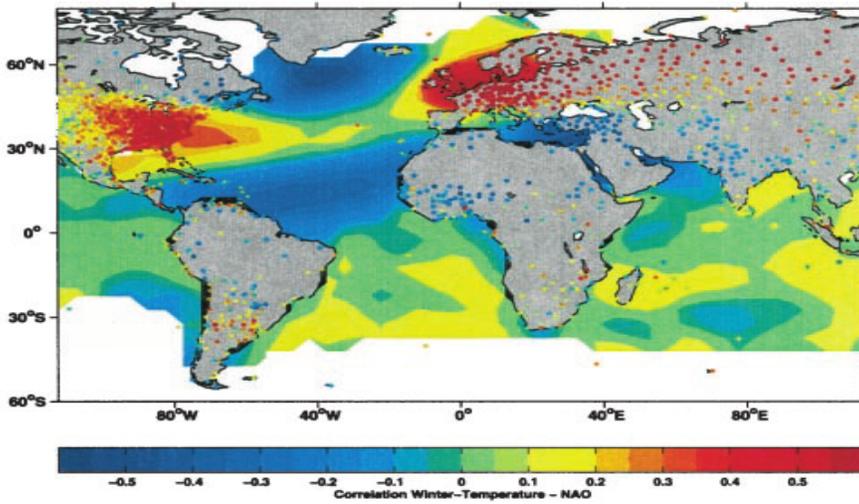


Figure 86: The principal component (PC) time series of the leading EOF of DJF SLP anomalies over the Atlantic sector (20-80N, 90W-40E) serves as an alternative index. The DJF PC timeseries is shown below in color, and the station based index is given by the thick black line. The correlation between the two is 0.88 over the period 1899-2022. The black dots on the EOF panel show the location of the stations used in the DJF station-based index. (Climate Data Guide; A. Phillips).

Winter (DJFM) SST and Land Temperature correlated with NAO index



SLP difference Island – Portugal (NAO index)

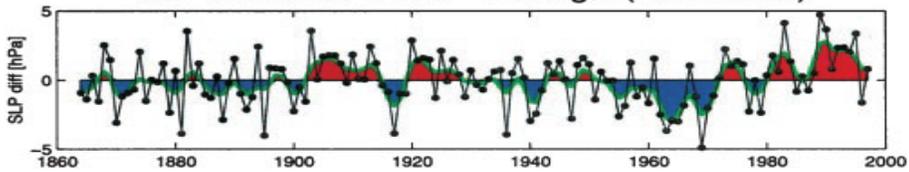


Figure 87: The surface temperature and other impacts associated with the positive phase of the NAO. Figure from Visbek et al., 2001

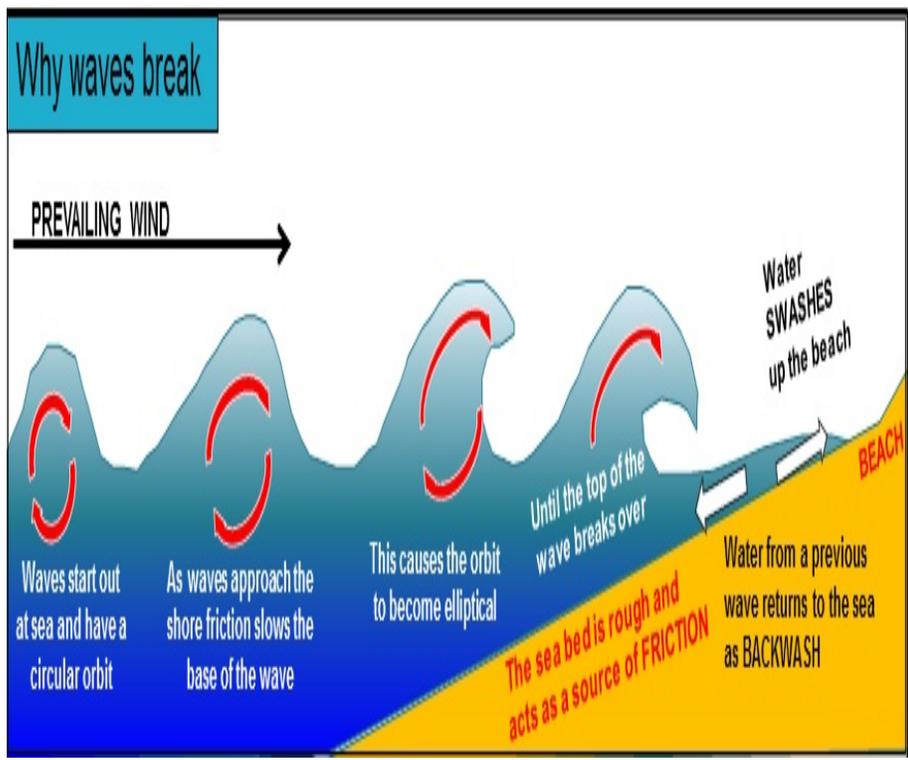


Figure 88: Breaking of shallow water waves. Source: <https://www.coolgeography.co.uk/GCSE/AQA/Coastal%20Zone/Processes/waves.htm>

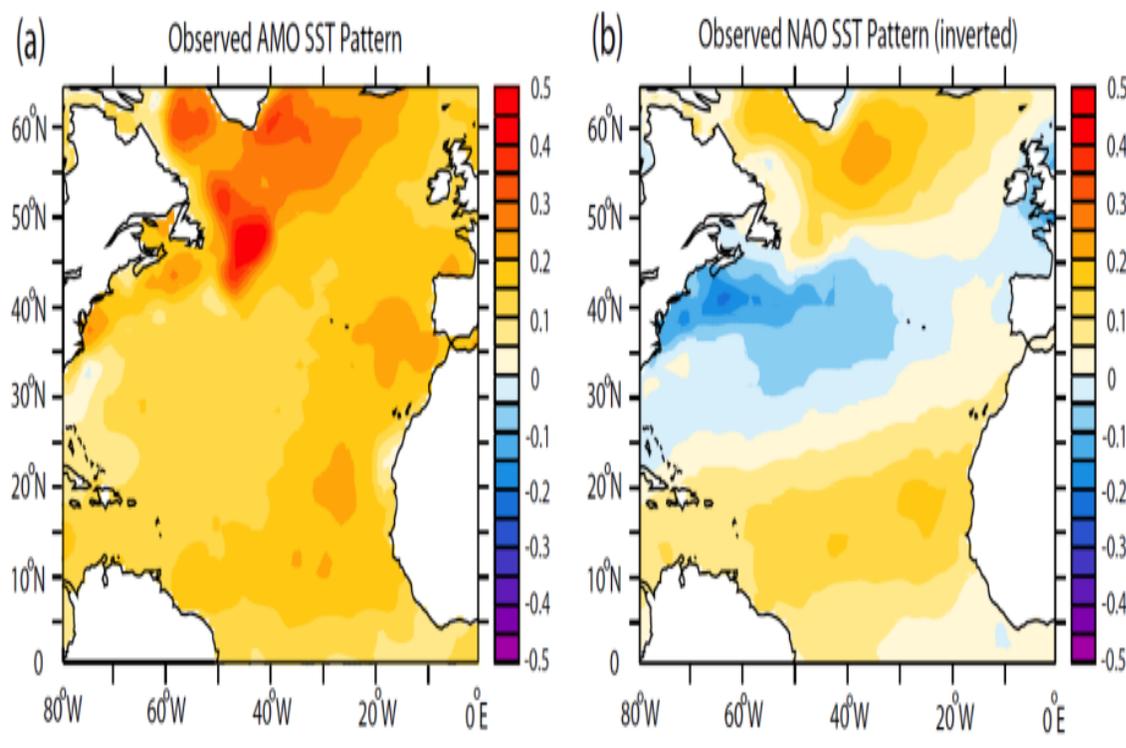


Figure 89: (a) Observed AMO SST pattern, derived by regressing detrended North Atlantic annual mean SST anomalies on the observed AMO index (Figure 1), using HadISST dataset (Rayner et al. 2013) for the period 1870-2015. The regression corresponds to 1 standard deviation of the observed AMO index. (b) Observed NAO SST pattern (inverted), derived by regressing detrended North Atlantic annual mean SST anomalies (HadISST dataset) on the inverted detrended observed winter NAO index (Hurrell Station-Based DJFM NAO Index) for the period 1870-2015. The regression corresponds to 1 standard deviation of the inverted observed NAO index. (Created by Dr Rhong Zhang for the Climate Data Guide)

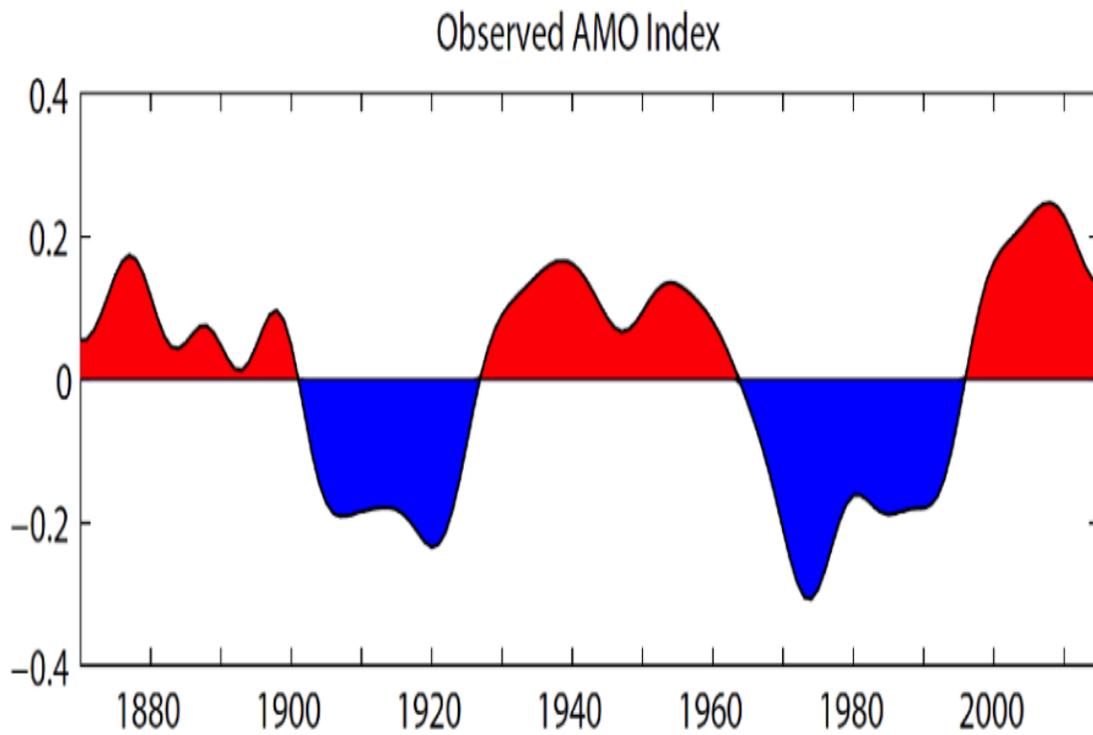


Figure 90: Observed AMO index, defined as detrended 10-year low-pass filtered annual mean area-averaged SST anomalies over the North Atlantic basin (0N-65N, 80W-0E), using HadISST dataset (Rayner et al. 2003) for the period 1870-2015. (Created by Dennis Shea and Dr Rhong Zhang for the Climate Data Guide)

Lead-lag relationship with NAO in observations inconsistent with slab-ocean mechanism

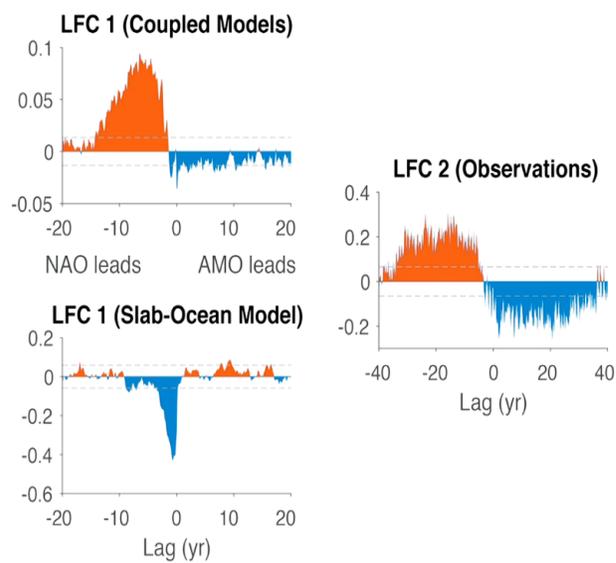


Figure 91: Lead-Lag correlations between the NAO and AMO indices for the 2 first hypothesis. Figure taken from Rob Will's talk at the 4th summer school on Climate dynamic, in Trieste

15 Predictability Basics

15.1 A Model of Convecting Fluids: The Lorenz Model

Essentially from the textbook: Chaos and Nonlinear Dynamics. Robert C. Hilborn, Oxford University Press, 1994.

The set of nonlinear equations (discuss linear oscillator, and exponential decay!) derived in this section is a highly simplified model of a convecting fluid. The model was introduced in 1963 by MIT meteorologist Edward Lorenz, who was interested in modelling convection in the atmosphere. What Lorenz set out to demonstrate was that even a very simple set of equations may have solutions whose behaviour is essentially unpredictable. Unfortunately for the development of the science of chaos, Lorenz published his results in the respectable, but little read Journal of the Atmospheric Sciences, where they languished essentially unnoticed by mathematicians and scientists in other fields until the 1970s (about 10 citations until 1972, then 3000 in one year.....). Now that chaos is more widely appreciated, a minor industry studying the Lorenz model equations has developed.

Here, the Lorenz equations will not be derived, we will just say enough to give a feeling for what the equations tell us. In simple physical terms, the Lorenz model treats the fluid system (say the atmosphere) as a fluid layer that is heated at the bottom (due to the sun's heating the earth's surface, for example) and cooled at the top. The bottom of the fluid is maintained at a temperature T_w (the 'warm' temperature), which is higher than the temperature T_c (the 'cold' temperature) at the top. We will assume that the temperature difference $T_w - T_c$ is held fixed. (This type of system was studied experimentally by Benard in 1900. Lord Rayleigh provided a theoretical understanding of some basic features in 1916. Hence, this configuration is now called *Rayleigh-Benard* cell.) Fig. 92 shows the principle set-up.

If the temperature difference $\delta T = T_w - T_c$ is not too large, the fluid will remain stationary. Heat is transferred from bottom to top by means of thermal conduction. The tendency of warm (less dense) fluid to rise is counterbalanced by a loss of heat from the warm fluid 'packet' to the surrounding medium. The damping due to the fluid viscosity prevents the packet from rising more rapidly than the time required for it to come to the same temperature as its neighbours. Under these conditions the temperature drops *linearly* with the vertical position from T_w at the bottom of the layer to T_c at the top. However, if the temperature difference becomes large enough, the buoyant forces eventually become strong enough to overcome viscosity and steady circulation currents develop. In this situation heat is transferred from the bottom to the top by the process of convection, the actual mass motion of the fluid. In simple terms, when the warm packet of fluid reaches the top of the layer, it loses heat to the cool region and then sinks to the bottom, where its temperature goes up

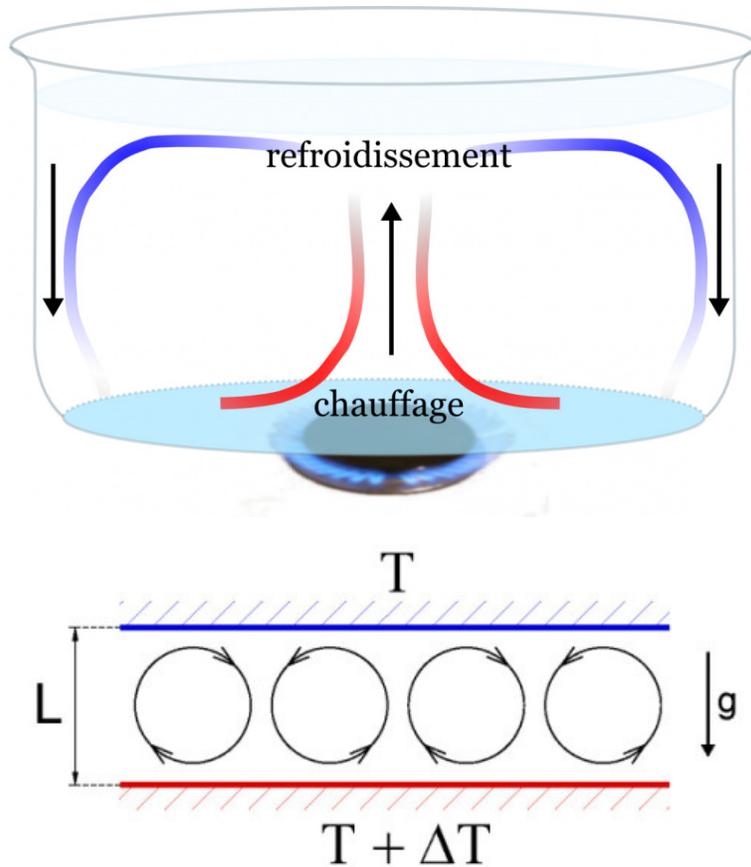


Figure 92: Setting of Rayleigh-Benard convection.

again. The net result is a circulation pattern that is stable in time.

With a further increase in temperature difference δT , the circulation currents and the resulting temperature differences within the fluid start to vary in time. This never occurs for a linear system with frictional forces included. If a linear system is subject to steady forces, (after an initial transition period) will be steady in time.

15.2 The Lorenz Equations

The Lorenz model is based on a gross simplification of the fundamental Navier-Stokes equations for fluids (explain the approach to insert fixed spatial dependencies of solutions into the Navier-Stokes equations, then look for solutions of the time-dependent amplitudes). The fluid motion and resulting temperature differences can be expressed in terms of these three variables, conventionally called $X(t)$, $Y(t)$ and $Z(t)$. These are *not* spatial variables. X is related to

the time-dependence of the so-called fluid stream function. The variables Y and Z are related to the time dependence of the temperature deviations away from the linear temperature drop from bottom to top, which one obtains for the nonconvective steady-state situation. In particular, Y is proportional to the temperature difference between the rising and falling parts of the fluid at a given height, while Z is proportional to the deviation from temperature linearity as a function of vertical position.

Using these variables, we may write the Lorenz model equations as three coupled differential equation

$$\begin{aligned}\dot{X} &= p(Y - X) \\ \dot{Y} &= -XZ + rX - Y \\ \dot{Z} &= XY - bZ .\end{aligned}\tag{292}$$

p, r, b are adjustable parameters: p is the so-called Prandtl number, which is defined to be the ratio of kinetic viscosity of the fluid to its thermal diffusion coefficient. r is proportional to the *Rayleigh number*, which is a dimensionless measure of the temperature difference between the bottom and top of the fluid. As the temperature difference increases, the Rayleigh number increases. The final parameter b is related to the ratio of the vertical height h of the fluid layer to the horizontal size of the convection rolls. It turns out that for $b = 8/3$, the convection begins for the smallest value of the Rayleigh number, that is for the smallest value of the temperature difference δT . This value is usually chosen to study the Lorenz model. p is then chosen for the particular fluid under study. Lorenz (LOR63) used the value $p = 10$ (which corresponds roughly to cold water), a value that had been used in a previous study of Rayleigh-Benard convection by Saltzman (SAL62). We let r , the Rayleigh number, be the adjustable control parameter. The Lorenz model, although based on what appears to be a very simple set of differential equations, exhibits very complex behaviour. The equations look so simple that one is led to guess that it would be easy to write down their solutions. In fact, it is now believed that it is in principle impossible to give the solutions in analytical form. Thus, we must solve the equations numerically (explain possible discretization!). Here we will discuss a few results of those integrations.

15.3 Behaviour of Solutions to the Lorenz Equations

For small values of the parameter r , that is, for small temperature differences, δT , the model predicts that the stationary, nonconvecting state is the stable condition. In terms of the variables X, Y, Z , this state is described by the values $X_{s1} = 0, Y_{s1} = 0, Z_{s1} = 0$. For values of r greater than 1, steady convection sets in (it is actually quite easy to perform a linear stability analysis of this stationary point at zero, and it is left to you as an exercise, if you wish.....).

There are two possible convective states: one corresponding to clockwise rotation, the other to counterclockwise rotation (discuss calculation of stationary points: $X_{s2,3} = \pm\sqrt{b(r-1)}$, $Y_{s2,3} = \pm\sqrt{b(r-1)}$, $Z_{s2,3} = r-1$, solutions 2,3 only exist for $r > 1$). Some initial conditions lead to one state, other initial conditions to the other state. For $p > b+1$, this steady convection is *unstable* for large enough r and gives way to more complex behaviour. As r increases, the behaviour has regions of chaotic behaviour intermixed with regions of periodic behaviour and regions of *intermittency*, which cycle back and forth, apparently random, between chaotic and periodic behaviour. Solutions are shown in Fig. 93 for different values of r (discuss dependence of solutions on initial conditions for all cases).

Note that even though the Lorenz equations strictly only apply to the Rayleigh-Benard convection experiment (and even here they are crude approximations to the full equations), these equations are applied to study the behaviour of many complex systems. For example, there is a vast literature on application to Indian monsoon intraseasonal variability, extratropical flow regimes, etc.

15.4 Kinds of predictability

It is the dependence of the solutions on small variations (uncertainties) in the initial conditions of our nonlinear system (Navier-Stokes equations), that leads us to the introduction of the concepts of predictability. Adrian Tompkins will in his lecture on Numerical Weather Prediction (NWP) discuss what the sources of uncertainties in the initial conditions are (obviously related to measurement errors). There are other sources of uncertainties related to model imperfections that may also be included in the treatment, but are excluded here at the moment.

From Lorenz system we can understand the two principle *kinds of predictabilities*. The first kind is the (atmospheric) initial condition predictability, which can be illustrated looking at the development of an initial condition ensemble (this means many different initial conditions that have to be considered because of our uncertainties in the initial conditions). Fig. 94 illustrates the time development of such an initial condition ensemble in the Lorenz system. Predictability that results from the atmospheric initial conditions is the subject of NWP, and the limit turns out to be a few days. The NWP problem will be discussed by Adrian Tompkins in depth.

The more general case also including model uncertainty is illustrated in Fig. 95, which again illustrates the principle of growth in time of the differences between the different simulations, which could be interpreted as measure of uncertainty.

The predictability is usually measured using the *ensemble* method, meaning that several realizations of a prediction are performed from only slightly dif-

ferent initial conditions.

Fig. 96 shows the mean (root-mean-square) difference of the near surface air temperature at one gridpoint in Europe (upper panel) and equatorial Africa (lower panel) from the ECMWF seasonal hindcast ensemble (15 members) for slightly different initial conditions (we will discuss later what exactly this system is). As can be seen the error grows quickly in the first few days (even faster for equatorial Africa), then saturates more or less around 15-20 days for Europe, but already after 2-3 days in equatorial Africa, and at a much smaller saturation value. This feature has important implications for seasonal predictability in extratropical and tropical regions!

Usually, initial condition predictability of the atmosphere (particularly in the extratropics) is limited by a typical *error growth* time scale of a few days, which effectively means that no reasonable prediction can be made after a week or so (see Fig. 96). This initial condition predictability is also referred to as *predictability of the first kind*. The *predictability of the second kind* results from the fact that the solutions of nonlinear systems (like e.g. the Lorenz system) stay in any case relatively close to the unstable stationary points, or *Attractors*. If any *external forcing* changes these stationary points, the solutions will change systematically with the changing stationary points. We can imagine that such a situation occurs, for example, in the atmosphere in the case of a large *El Nino* event. The atmospheric weather or *noise* will be chaotic and unpredictable on a seasonal time-scale (see Fig. 96), but the (imaginary) stationary state around which the atmosphere evolves may be shifting far enough that we can predict the average climate in this situation. This is what is referred to as predictability of the second kind (for the atmosphere). This is the usual situation we consider in *seasonal forecasts*. It is interesting for this to consider the differences in the saturation errors seen in Fig. 96, which is much larger in extratropical regions compared with tropical regions. This indicates potentially more predictability of the second kind in the tropics, even if predictability of the first kind is less. We will see in the following sections that this is indeed generally the case.

Note, however, that the separation between predictability of first and second kind is a little arbitrary and we need to specify exactly for which system. In all above reference to predicability of first and second kind I have added *atmosphere*. Indeed, seasonal forecasts are performed with coupled ocean-atmosphere models. For the ocean-atmospheric system, we are looking again at the initial condition predictability, or predictability of the first kind even for seasonal forecasts. The predictability relies then on the fact that some parts of the oceans evolve on much longer time-scales compared to the atmosphere (e.g. ENSO has a period of 2-7 years!). The same argument may apply to other 'external' forcing, for example, carbon dioxide may be considered as external forcing in many standard models, but if our model includes a carbon cycle, then it is part of the dynamical system itself. As a note, some researchers have

also argued that an external forcing may not necessarily shift the attractor, but may express itself as a change of probability of the system to be close to one or the other stationary state (see Fig 97 from paper by Corti, Molteni and Palmer), which may make it difficult to identify, e.g. Climate change from observations.

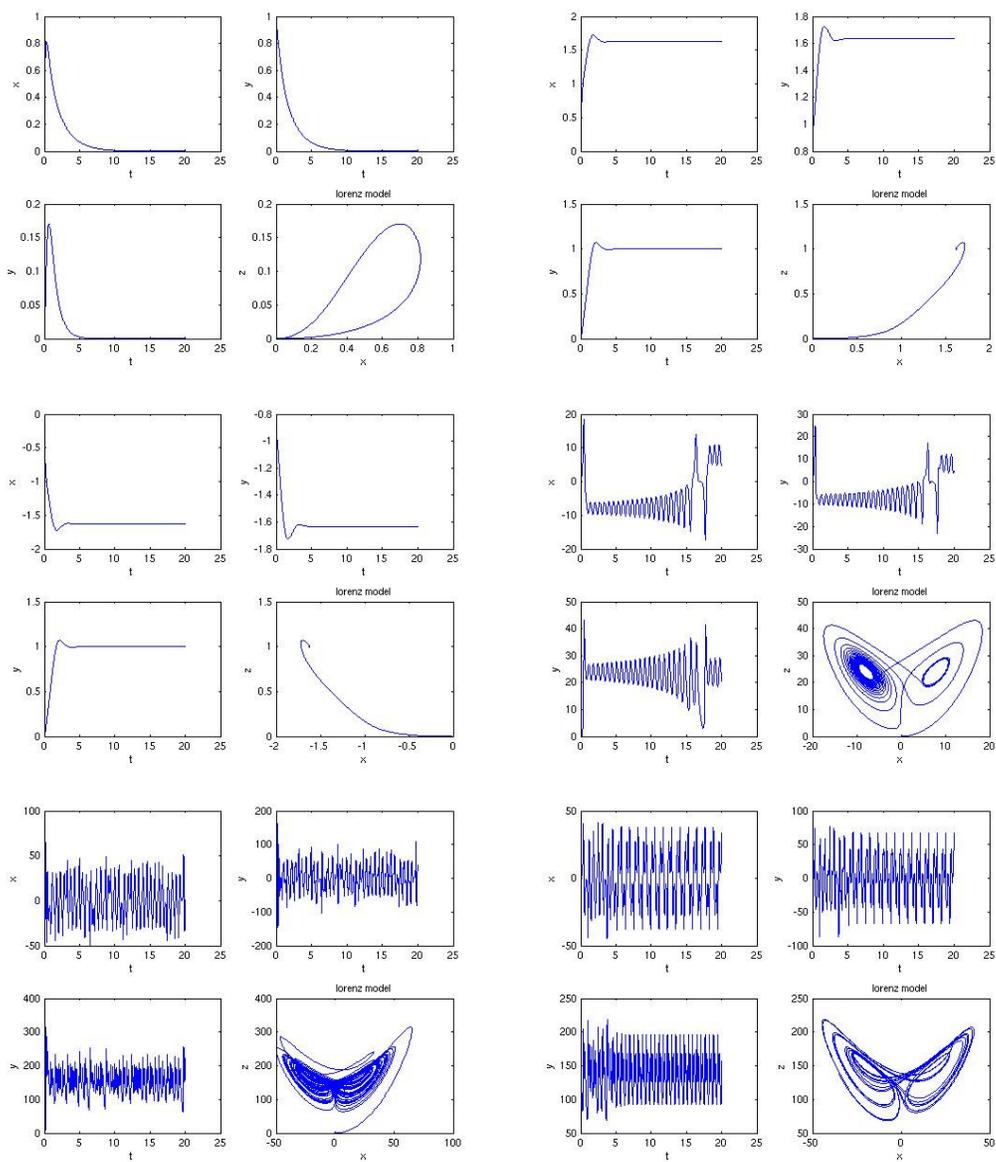


Figure 93: Solutions of the Lorenz equations for different parameters r .

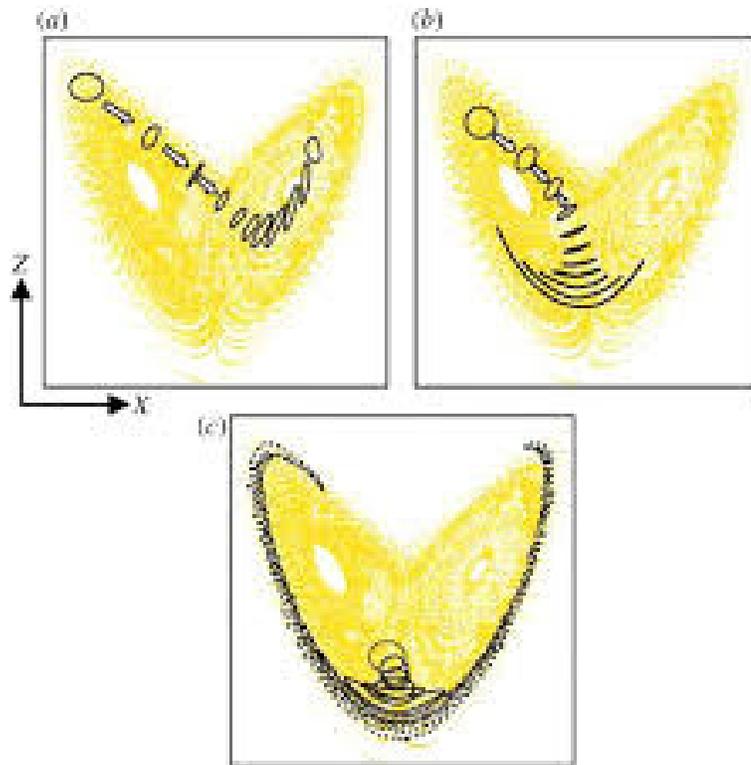


Figure 94: Illustration of time evolution of an initial condition ensemble in the Lorenz system.

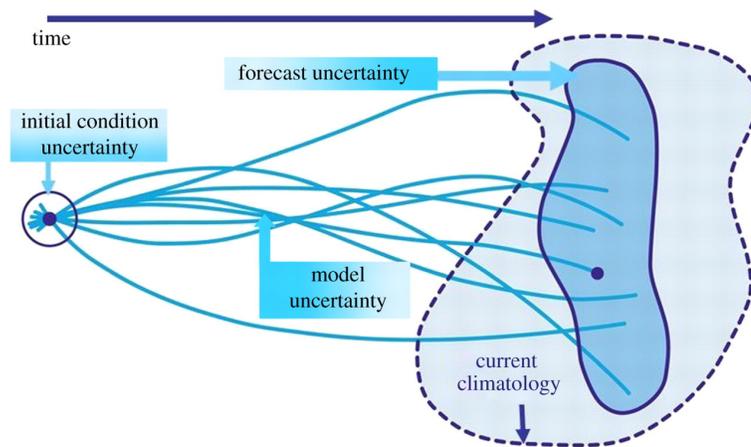
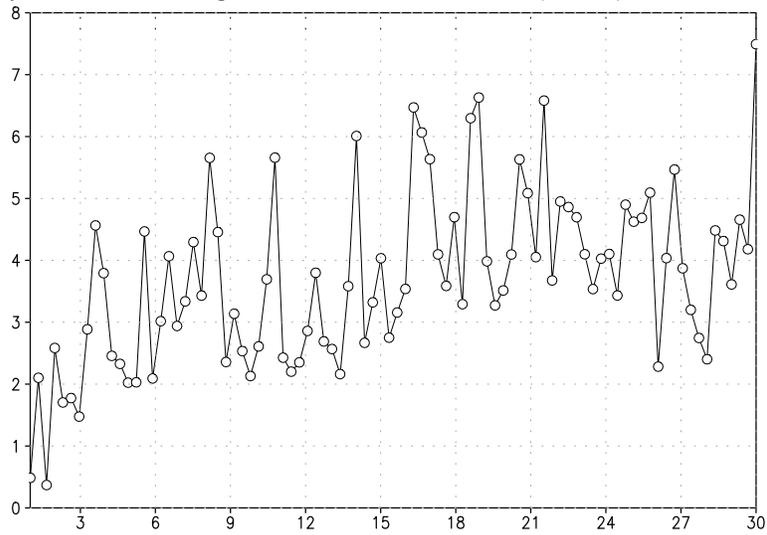


Figure 95: Illustration of the predictability problem.

sys4 error growth 2mt Europe (15E,50N)



sys4 error growth 2mt Eq Africa (20E,0N)

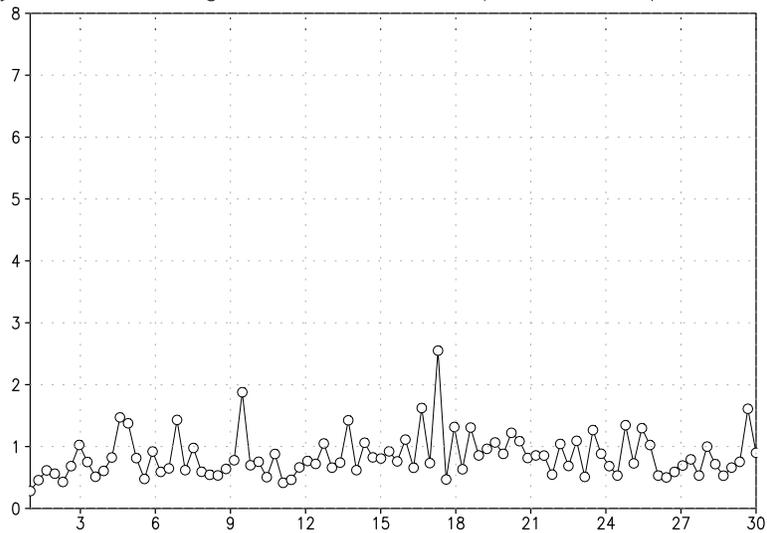


Figure 96: The mean (root-mean-square) difference of the near surface air temperature at one gridpoint in Europe (15E, 50N) and equatorial Africa (20E, 0N) from the ECMWF seasonal hindcast ensemble (15 members) for slightly different initial conditions. Units are K.

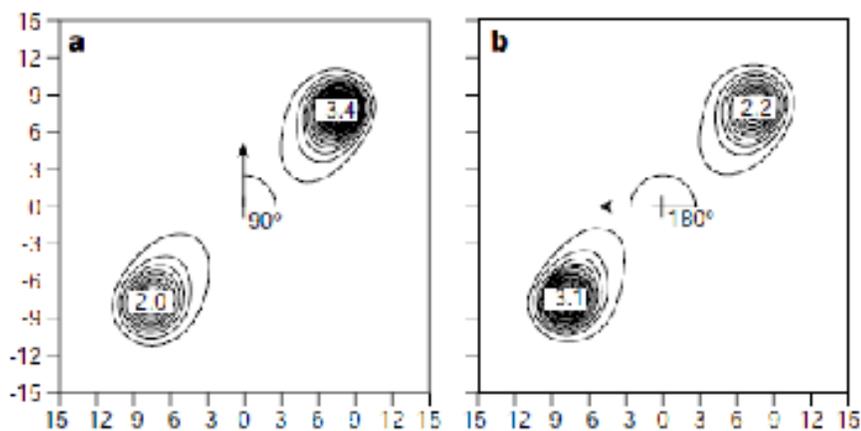


Figure 1 Response of a nonlinear chaotic model to imposed forcing. Illustrated is the state vector PDF (invariant measure) of a forced Lorenz² attractor, with governing equations $\dot{X} = -10X + 10Y + 2.5 \cos \theta$, $\dot{Y} = -XZ + 28X - Y + 2.5 \sin \theta$, $\dot{Z} = XY - (8/3)Z$, in X, Y phase space. The arrow shown is the forcing vector $(2.5 \cos \theta, 2.5 \sin \theta)$ in this two-dimensional phase space. A short running time-average has been applied to the state vector to emphasise the regime character of the Lorenz attractor. **a**, $\theta = 90^\circ$; **b**, $\theta = 180^\circ$. The numbers shown correspond to maxima of the PDF at the regime centroids. The quantities plotted on the horizontal and vertical axes are values of X and Y respectively. See ref. 3 for details.

Figure 97: Illustration of frequency of occurrence changes in a Lorenz system from paper Corti, S., F. Molteni and T. N. Palmer, 1999: Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature*, **398**, 799-802

16 Predictability Measures

16.1 Measures of model-derived predictability

Before we formally introduce the concepts of signal and noise in the seasonal prediction context, let us first consider an example of a seasonal prediction in Fig. 98.

Let x_{ij} be a model variable (e.g. near-surface temperature or rainfall) at a certain gridpoint at a discretized time $i = 1, \dots, N$ for the ensemble member $j = 1, \dots, M$. The *noise* in weather and climate is then usually defined as the variance of the deviations from the ensemble mean

$$NO = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (x_{ij} - [x]_i)^2 \quad , \quad (293)$$

where $[x]_i$ is the ensemble mean

$$[x]_i = \frac{1}{M} \sum_{j=1}^M x_{ij} \quad . \quad (294)$$

We can obviously define these values at every gridpoint of the model and thus define maps of e.g. NO . The meaning of this definition becomes clear if we imagine that the results of an ensemble of 2 (or more) simulations are identical. In this case the noise NO is zero. The ensemble mean of an ensemble is supposed to define the signal, thus

$$SI = \frac{1}{N} \sum_{i=1}^N \left([x]_i - \overline{[x]} \right)^2 \quad , \quad (295)$$

where $\overline{[x]}$ is the time mean of the ensemble mean

$$\overline{[x]} = \frac{1}{N} \sum_{i=1}^N [x]_i \quad . \quad (296)$$

It should be noted that the ensemble mean usually contains a noise residual, particularly if the ensemble size is small, and provides therefore a *biased* estimate of the signal. Imagine we have just 2 simulations, then it is clear that the mean of these 2 simulations will not be effective to identify the signal (unless they are identical, and therefore the noise is zero). There are ways to correct/improve this, but we will not deal with this here. Discuss this using ensemble Nino3.4 forecasts! With this the *signal-to-noise* ratio $S - N$ is simply

$$S - N = \frac{SI}{NO} \quad . \quad (297)$$

In principle, $S - N$ can become infinity, if NO is zero. In practise this is usually not the case, although values can become quite large (in which case we are lucky!!!). A useful threshold to be considered could be $S - N = 1$, for which the signal has the same variance as the noise, indicating some predictability. In practical applications to important variables such as surface temperatures, precipitation or geopotential height in seasonal forecasts, $S - N$ turns out to be typically small for precipitation and other variables outside the tropical Pacific regions (see Fig. 99).

A useful transformation of the $S - N$ is called *theoretical limit of predictability*

$$R_{limit} = \sqrt{\frac{SI/NO}{SI/NO + 1}} = \sqrt{\frac{SI}{SI + NO}} = \sqrt{\frac{SI}{T}} \quad , \quad (298)$$

where $T = SI + NO$ is the *total* variance. By definition R_{limit} lies between 0 and 1. It may be interpreted as maximum expected correlation skill if we were to correlate the ensemble mean with the observations (correlation skill will be introduced in the next section). Zero means there is no predictability, 1 means there is perfect predictability. A value of $S - N$ of 1 translates into a value of R_{limit} of about 0.7. There are many more indicators of predictability, that are related to information theory (e.g. relative entropy), but we will restrict ourselves here to just the basic ones. Fig. 100 shows the R_{limit} for seasonal mean (September-to-November) precipitation over land points. As we can see, unfortunately the seasonal mean theoretical limit of Predictability is typically low over land points. Another global assessment of R_{limit} for surface temperature, mean sea level pressure and precipitation for the DJF season is shown in Fig. 101

Another, purely model derived predictability measure is the *potential correlation skill*. The idea is to calculate some kind of mean correlation derived from the model that may be compared with the correlation of an ensemble mean with an *observation*. We may treat every single ensemble member of an ensemble of realizations as observation (it contains internal and forced variability components). We can correlate each ensemble member with an ensemble mean of the *remaining* simulations. For example, we use x_{i1} and correlate this with the ensemble mean

$$[x]_i^{no1} = \frac{1}{M-1} \sum_{j=2}^M x_{ij} \quad . \quad (299)$$

This is because if x_{i1} were included in the ensemble mean calculation then we get trivial correlations due to this. Then we can calculate M correlation coefficients

$$\rho_j = \frac{\frac{1}{N} \sum_{i=1}^N \left([x]_i^{no_j} - \overline{[x]}^{no_j} \right) (x_{ij} - \bar{x}_j)}{\sigma_{[x]^{no_j}} \sigma_{x_j}} \quad (300)$$

where the standard deviations in time are defined as

$$\sigma_{x_j} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{ij} - \bar{x}_j)^2} \quad (301)$$

and

$$\sigma_{\bar{x}^{no_j}} = \sqrt{\frac{1}{N} \sum_{i=1}^N ([x]_i^{no_j} - \overline{[x]_i^{no_j}})^2} \quad (302)$$

Now one should average the correlations ρ_j . However, we should not just average these correlations since they are limited by [-1,1]. Instead, we may average them after applying a *FischerZ*-transformation

$$r_j = \frac{1}{2} \ln \left(\frac{1 + \rho_j}{1 - \rho_j} \right) \quad , \quad (303)$$

to get

$$[r] = \frac{1}{M} \sum_{j=1}^M r_j \quad . \quad (304)$$

After this, in order to get the average correlation, we have to transform back using the inverse transformation

$$[\rho] = \frac{e^{2[r]} - 1}{e^{2[r]} + 1} \quad . \quad (305)$$

Note that, in practise, this potential correlation skill $[\rho]$ is very similar numerically to the theoretical limit of predictability, R_{limit} . The reason for this is that the square of a correlation is the *explained variance*, which applied to our case is the explained variance fraction by the ensemble mean or by the signal. An example if a Potential Correlation Skill calculation is shown in Fig. 102.

16.2 Predictability from comparison with observations

Of course, in order to investigate the 'goodness' or *skill* of a model simulation, we should compare the model output with observations. There are a number of quality measures of a model simulation. The most basic one, perhaps, is the *bias* of the model. If the model has several realizations, that is, ensemble members, the (time-mean) bias is best evaluated by comparing the ensemble mean of a field, $[x]_i$, with the corresponding observation, ox_i

$$bias = \frac{1}{N} \sum_{i=1}^N ([x]_i - ox_i) \quad . \quad (306)$$

If an ensemble is not available, but just a single simulation, the single simulation can be used to assess the bias. One of the most commonly used measures of error that takes variability into account is the *Root-mean-square error (RMSE)*

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N ([x]_i - ox_i)^2} \quad . \quad (307)$$

Again, a single realization may be also used to evaluate the RMSE of a model. The problem with the *RMSE* is that we need to know what is an acceptable value of it, which can be quite tricky. In a forecast for a week, is a *RMSE* of 1 K acceptable? A more generically comparable predictability measure is the *Correlation Skill*

$$\rho = \frac{\frac{1}{N} \sum_{i=1}^N (ox_i - \overline{ox})([x]_i - \overline{[x]})}{\sigma_{ox}\sigma_{[x]}} \quad . \quad (308)$$

Such a measure may be somewhat simpler to evaluate as it has values between -1 and 1 and we may simply pick a generic threshold of, e.g. 0.5, for any variable. On the other hand, physically we may prefer the *RMSE* error measure (for example if your model has variations that are a factor of 1000 smaller than the observations, the *Correlation Skill* may still be 1, but the *RMSE* would be large). In order to better evaluate the *RMSE* we should compare it with some kind of trivial forecast without real skill. Such a forecast could be a climatological forecast or a persistence forecast. The *Brier Skill Score* uses this idea to define skillful forecasts

$$BS = 1 - \frac{RMSE^2}{RMSE_{cl}^2} \quad , \quad (309)$$

where $RMSE_{cl}$ is the base-line *RMSE* of, for example a climatological forecast. A forecast is skillful (compared to the base-line forecast) if the *BS* is positive.

An advantage of the *Correlation Skill* measure, ρ , is that we can directly compare it with the model-derived potential correlation skill $[\rho]$ (Eq. 305) or the theoretical limit of predictability, R_{limit} (Eq. 298). Fig. 103 shows an example of real seasonal prediction skill from a multimodel ensemble of seasonal *hindcasts*

16.3 Some other useful and simple techniques

A very useful technique in climate research is the *regression* analysis. Assume we want (in a model or in observations) to investigate what is the influence of ENSO (or any other phenomenon) on rainfall. Assuming that we can characterize ENSO by a single time-series (i.e. the Nino3.4 SST index), then the

influence of ENSO on rainfall at a certain location may be determined by a linear regression of rainfall at time i , r_i onto the Nino3.4 index (I_i)

$$r_i = a + bI_i \quad . \quad (310)$$

In the linear regression we are looking now for the coefficients a and b that minimize the sum of the squared differences between the linear model, Eq. (310), and the observational (or numerical model) counterpart ro_i

$$\epsilon = \frac{1}{N} \sum_{i=1}^N (ro_i - r_i)^2 \quad (311)$$

The theory of the linear regression tells us now how to determine (exercise!) the coefficients a and b (see Fig. 104). The mostly used coefficient b can be expressed as

$$b = \frac{\frac{1}{N} \sum_{i=1}^N (ro_i - \bar{ro})(I_i - \bar{I})}{\sigma_I^2} \quad . \quad (312)$$

Thus b can be interpreted as covariance between the index I_i and ro_i , divided by the variance of I_i .

Note that this formulae can be evaluated at every gridpoint, leading to a map of regression coefficients. The map of the coefficients b would tell us what is the typical linear *response* to a 1 K Nino3.4 SST anomaly in global rainfall (the dimension is mm/day per K). Also note that sometimes a different scaling is used

$$b^* = \frac{\frac{1}{N} \sum_{i=1}^N (ro_i - \bar{ro})(I_i - \bar{I})}{\sigma_I} \quad , \quad (313)$$

so that the dimension of b^* is simply mm/day. This would be the result if we had taken from the beginning a normalized index $I_i^* = (I_i - \bar{I})/\sigma_I$ which has standard deviation 1 by definition. Thus b^* is simply the covariance between the normalized index I_i and ro_i . This may be interpreted as the response to a *normalized* index I_i or the response to one standard deviation of the regression coefficient. One may interpret this also as a *composite* map based on linear regression. If we have an ensemble of simulations we could, for example, compare the ENSO regression map of the ensemble mean onto global rainfall with the observed regression map and try to identify if there are errors in the ENSO *teleconnections*.

A comparison between observations and an AGCM for the b^* regression coefficients for the Nino3.4 index regression onto winter (December-to-February) mean precipitation is shown in Fig. (105).

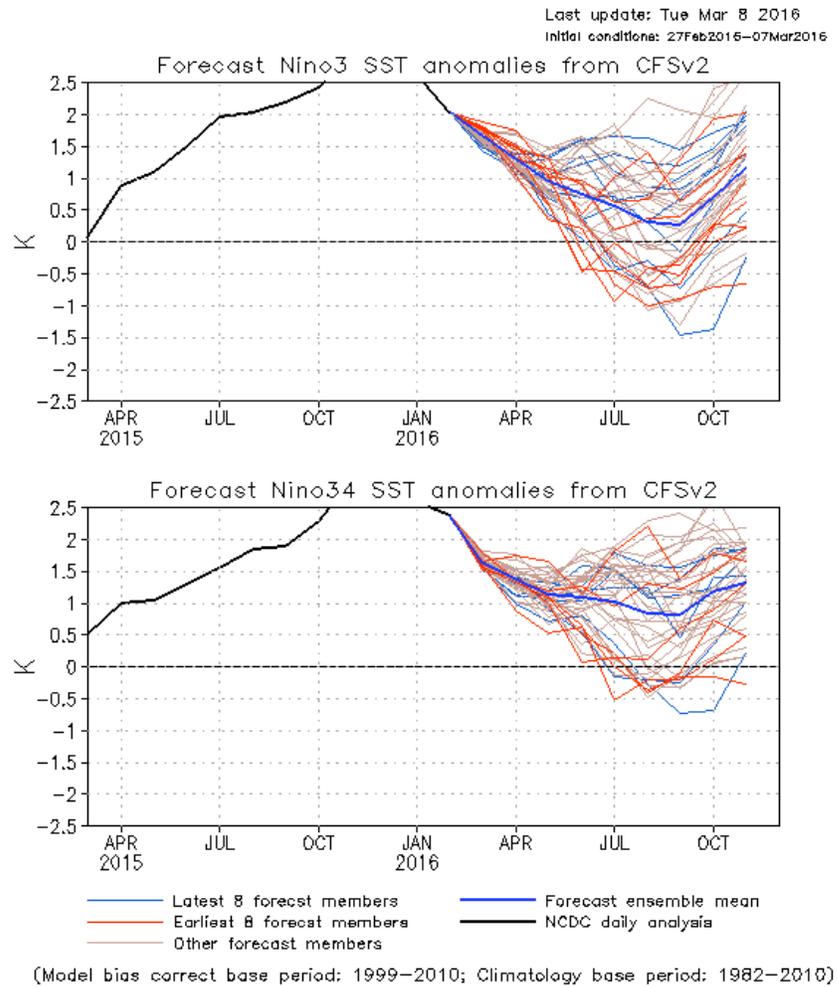


Figure 98: An example of Nino3 and Nino3.4 SST index prediction to illustrate the Signal-to-Noise ratio concept.

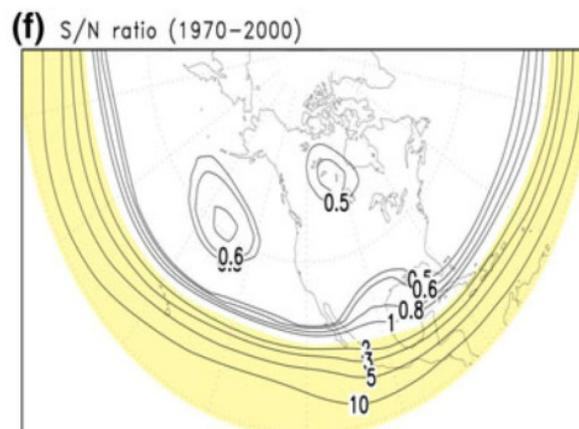
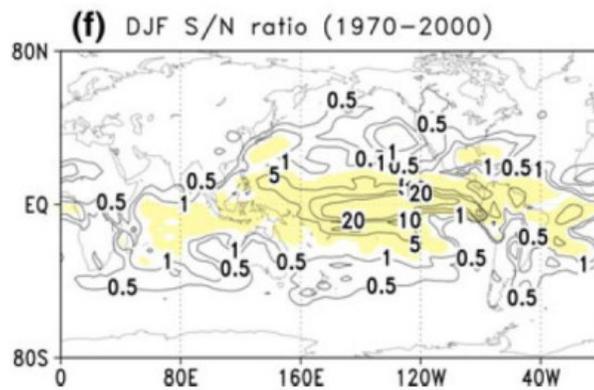


Figure 99: Signal-to-Noise ratio for winter (December-to-February) mean derived from an AGCM ensemble. Upper panel: Precipitation, lower pane: 200 hPa geopotential height. From paper: Ehsan et al., 2013: A quantitative assessment of changes in seasonal potential predictability for the 20th century. *Clim Dyn*, **41**, 2697-2709, doi: 10.1007/s00382-013-1874-x.

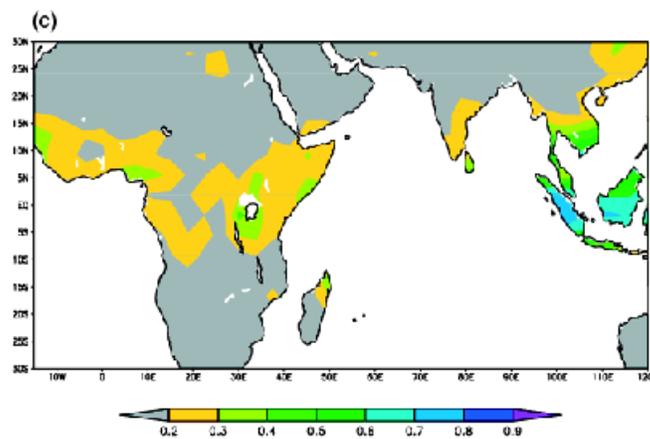
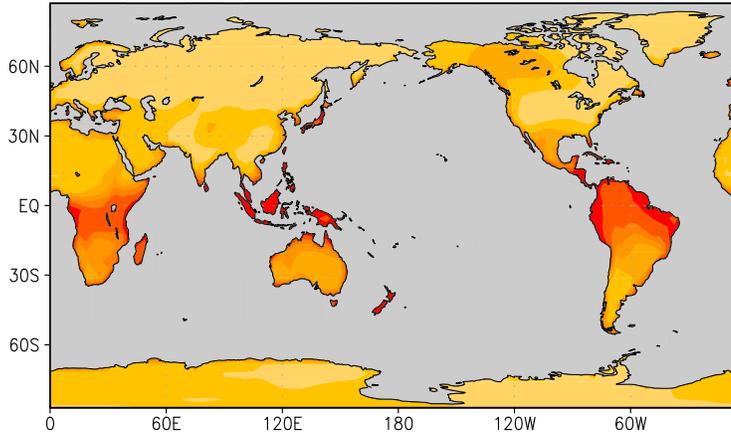
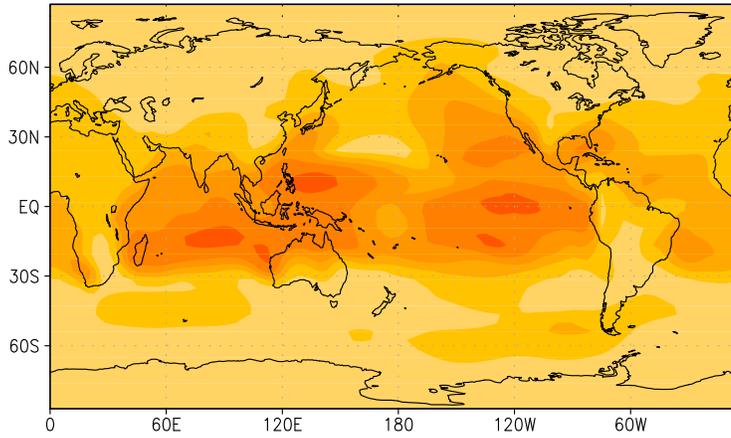


Figure 100: Theoretical limit of predictability, R_{limit} for September-to-November mean precipitation derived from an AGCM ensemble. From paper: Bahaga et al., 2015: Potential predictability of the sea-surface forced Equatorial East African short rains interannual variability in the 20th century. *Q. J. R. Meteorol. Soc.*, **141**, 16-26, doi: 10.1002/qj2338.

(a) DJF R_{lim} ST



(b) DJF R_{lim} MSLP



(c) DJF R_{lim} PREC

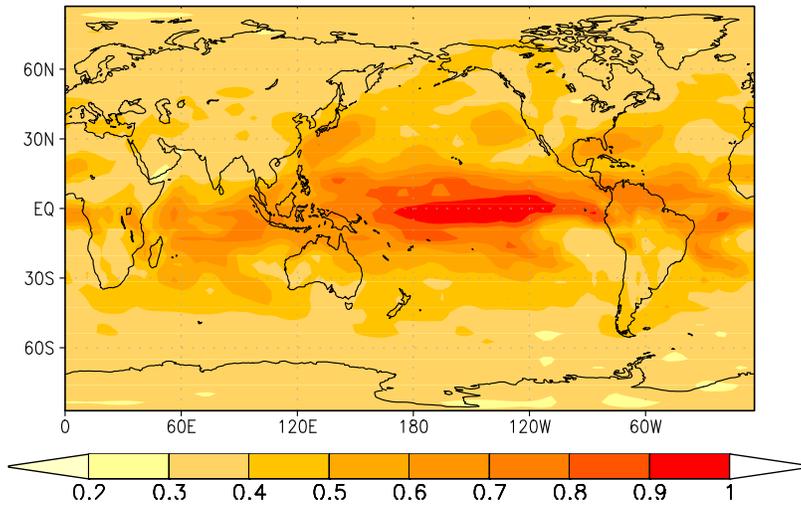


Figure 101: Theoretical limit of predictability, R_{limit} for December-to-February mean a) surface temperature, b) mean sea level pressure and c) precipitation for the DJF season derived from an AGCM ensemble. From Thesis: Bianca Mezzina, 2016: Seasonal influences of SST variability on European climate. University of Trieste Master thesis. 168

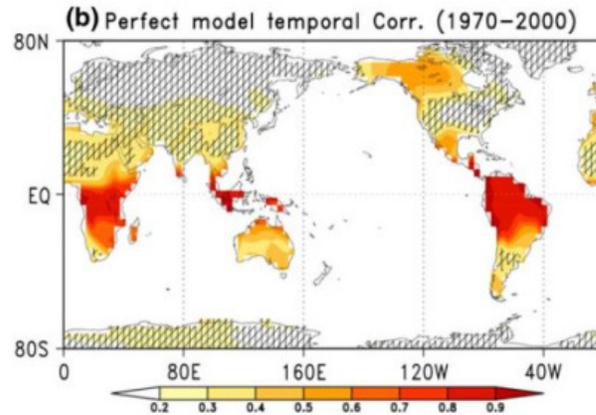


Figure 102: Potential Correlation Skill of near-surface temperature for winter (December-to-February) mean derived from an AGCM ensemble. From paper: Ehsan et al., 2013: A quantitative assessment of changes in seasonal potential predictability for the 20th century. *Clim Dyn*, 41, 2697-2709, doi: 10.1007/s00382-013-1874-x.

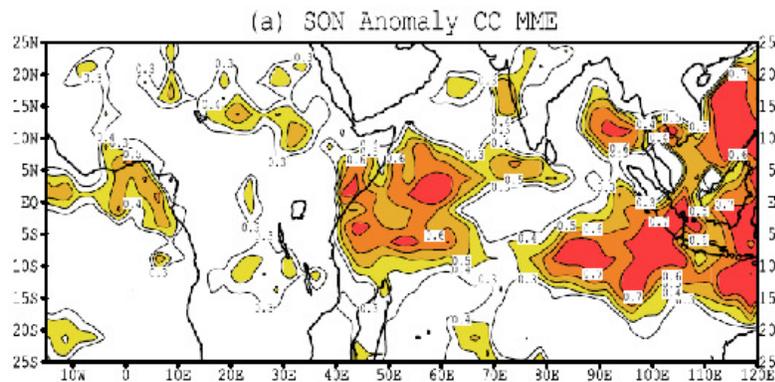


Figure 103: Correlation Skill of seasonal mean (September-to-November) precipitation from a multimodel seasonal forecast ensemble. From paper: Bahaga et al., 2015: Assesment of prediction and predictability od short rains over equatorial East Africa using a multi-model ensemble. *Theor. Appl. Climatol.* , doi: 10.1007/s00704-014-1370-1.

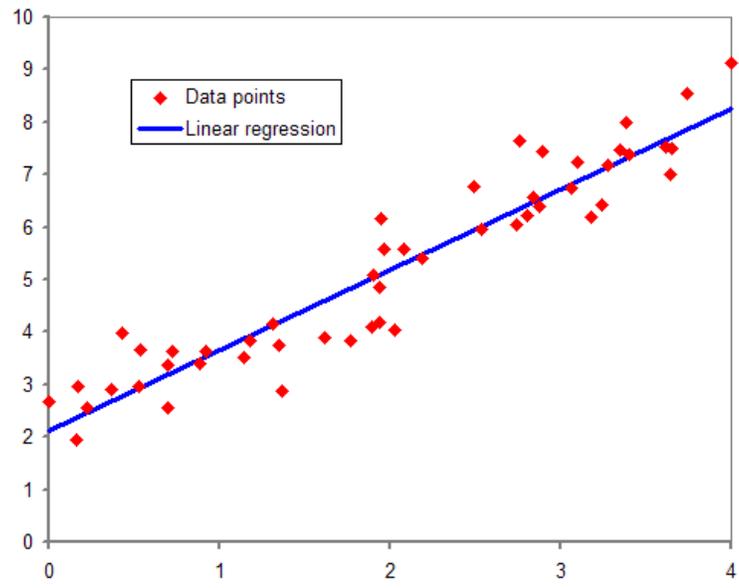


Figure 104: Linear regression method.

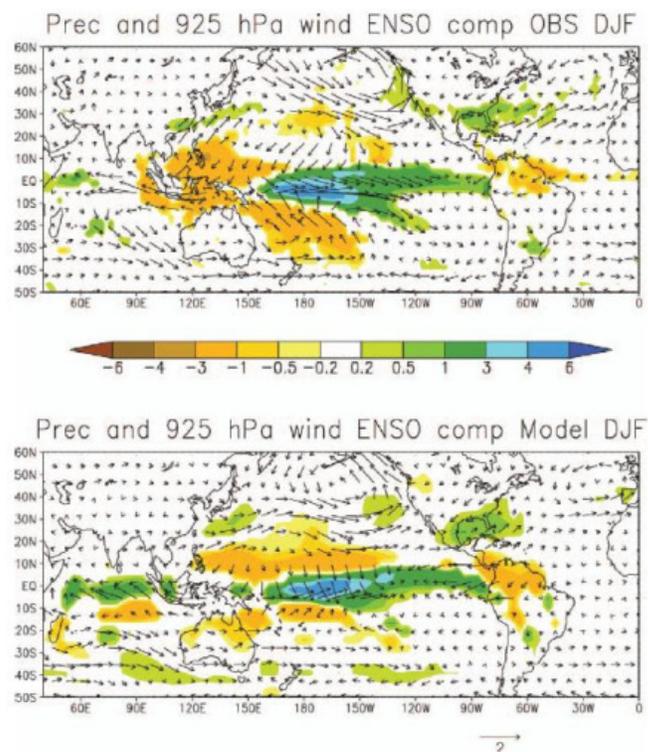


Figure 105: Regression coefficients b^* of the Nino3.4 index onto winter (December-to-February) mean precipitation. a) Observations, b) AGCM. Units are mm/day. From paper: Kucharski et al., 2015: On the need of intermediate complexity General Circulation Models: A “SPEEDY” Example. *BAMS*, **94**, 25-30, DOI: 10.1175/BAMS-D-11-00238.1.