École Doctorale des Sciences de l'Environnement d'Île-de-France
Année Universitaire 2024-2025

Modélisation Numérique de l'Écoulement Atmosphérique et Assimilation de Données

Olivier Talagrand
Cours 1

2 Avril 2025

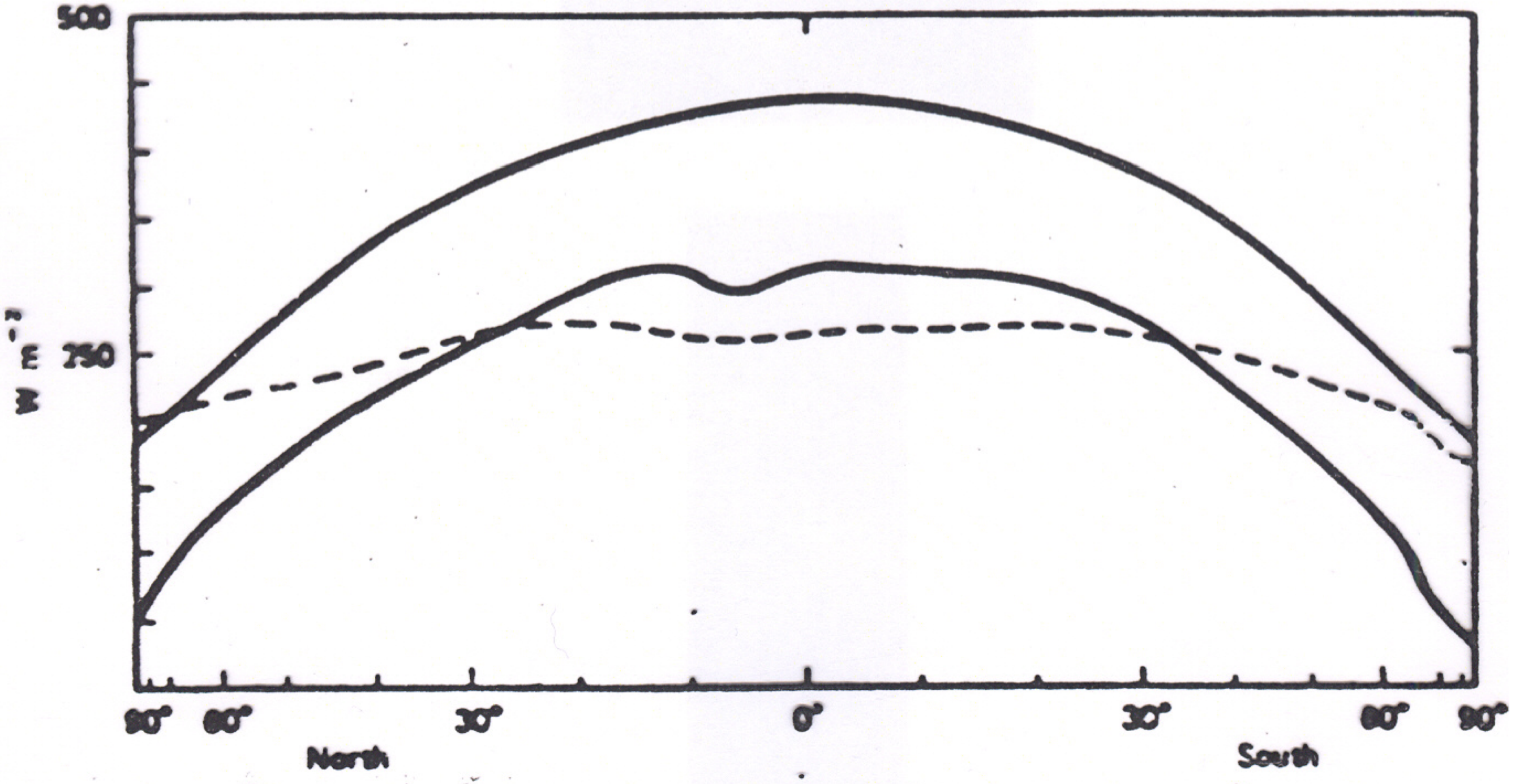
Programme of the course

1. Numerical modelling of the atmospheric flow. The *primitive* equations. Discretization methods. Numerical Weather Prediction. Present performance. Prevision by machine learning.
2. The meteorological observation system. The problem of 'assimilation’. Bayesian estimation. Random variables and random functions. Meteorological examples.
3. ‘Optimal Interpolation'. Basic properties. Meteorological applications. The theory of *Best Linear Unbiased Estimator*.
4. Advanced assimilation methods.

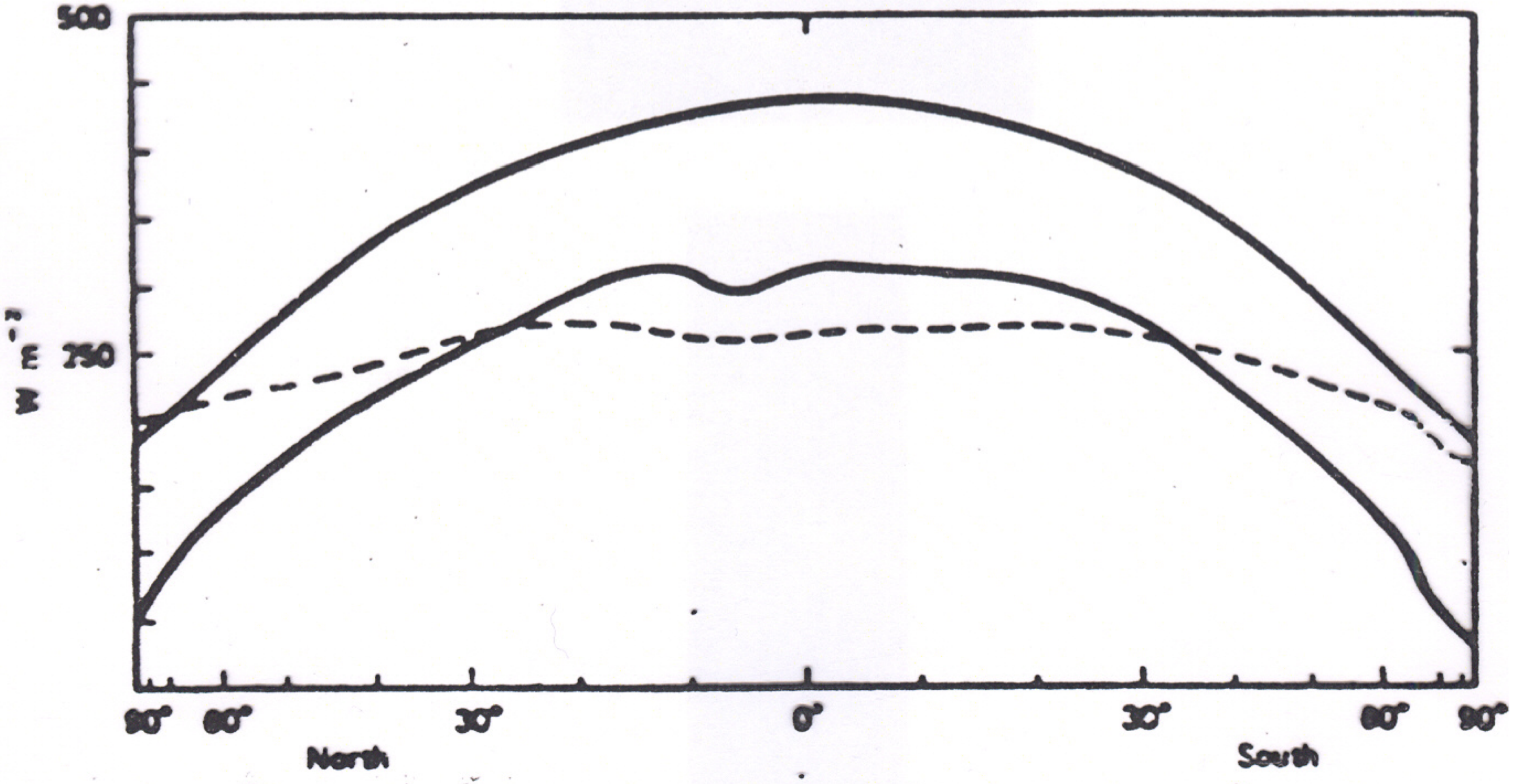
   - Kalman Filter. Ensemble Kalman Filter. Present performance and perspectives.

   - Variational Assimilation. Adjoint Equations. Present performance and perspectives.
5. Advanced assimilation methods (continuation).

   - Bayesian Filters. Theory, present performance and perspectives.
6. Assimilation and Machine Learning. General conclusions.



500
250
W m⁻²
90°
60°
30°
0°
30°
60°
90°
North
South



Earth Radiative Budget, yearly average

- Short description of the general circulation of the atmosphere and the ocean

- The physical laws which govern the circulation

- The *primitive equations*

- Spatial and temporal discretization

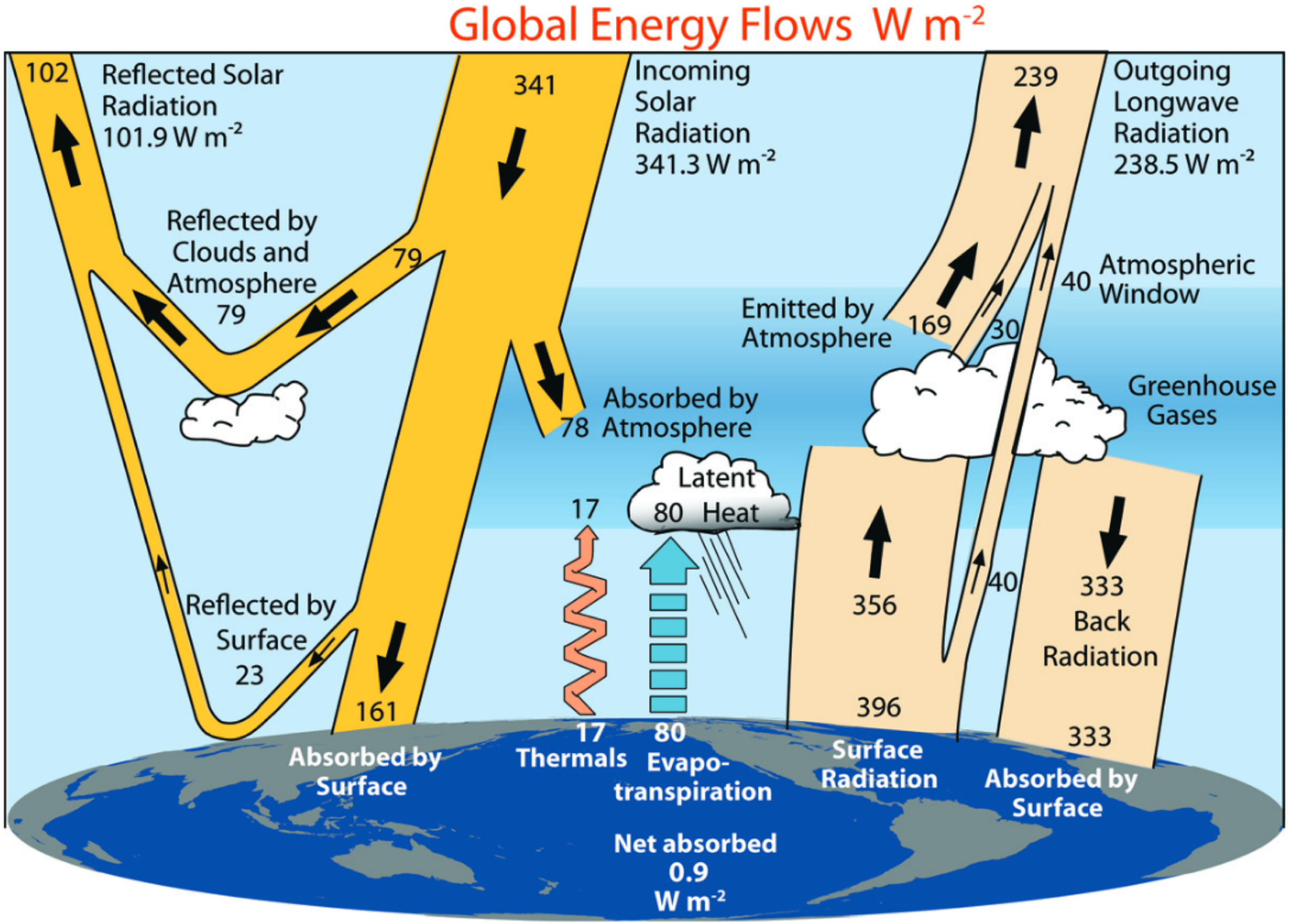
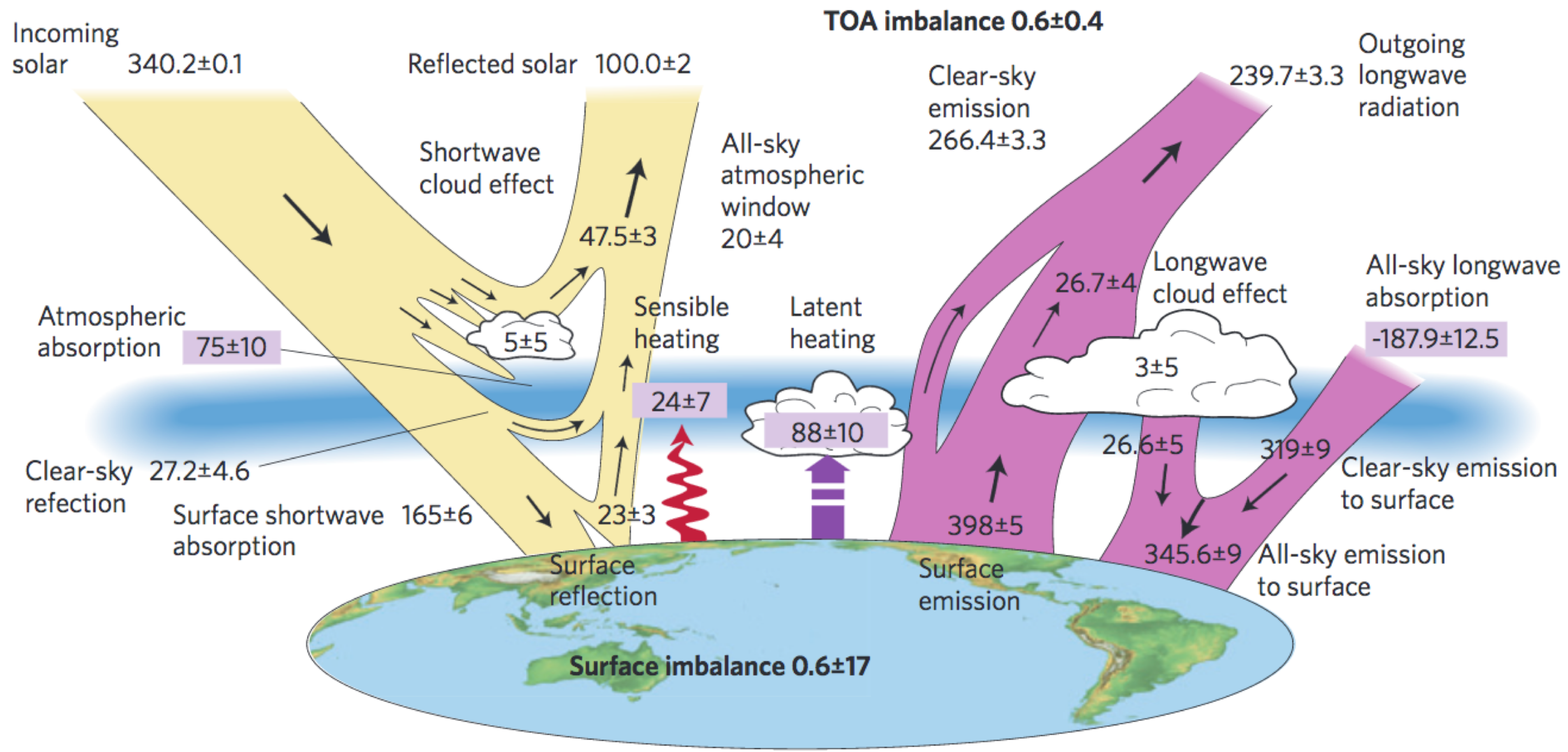


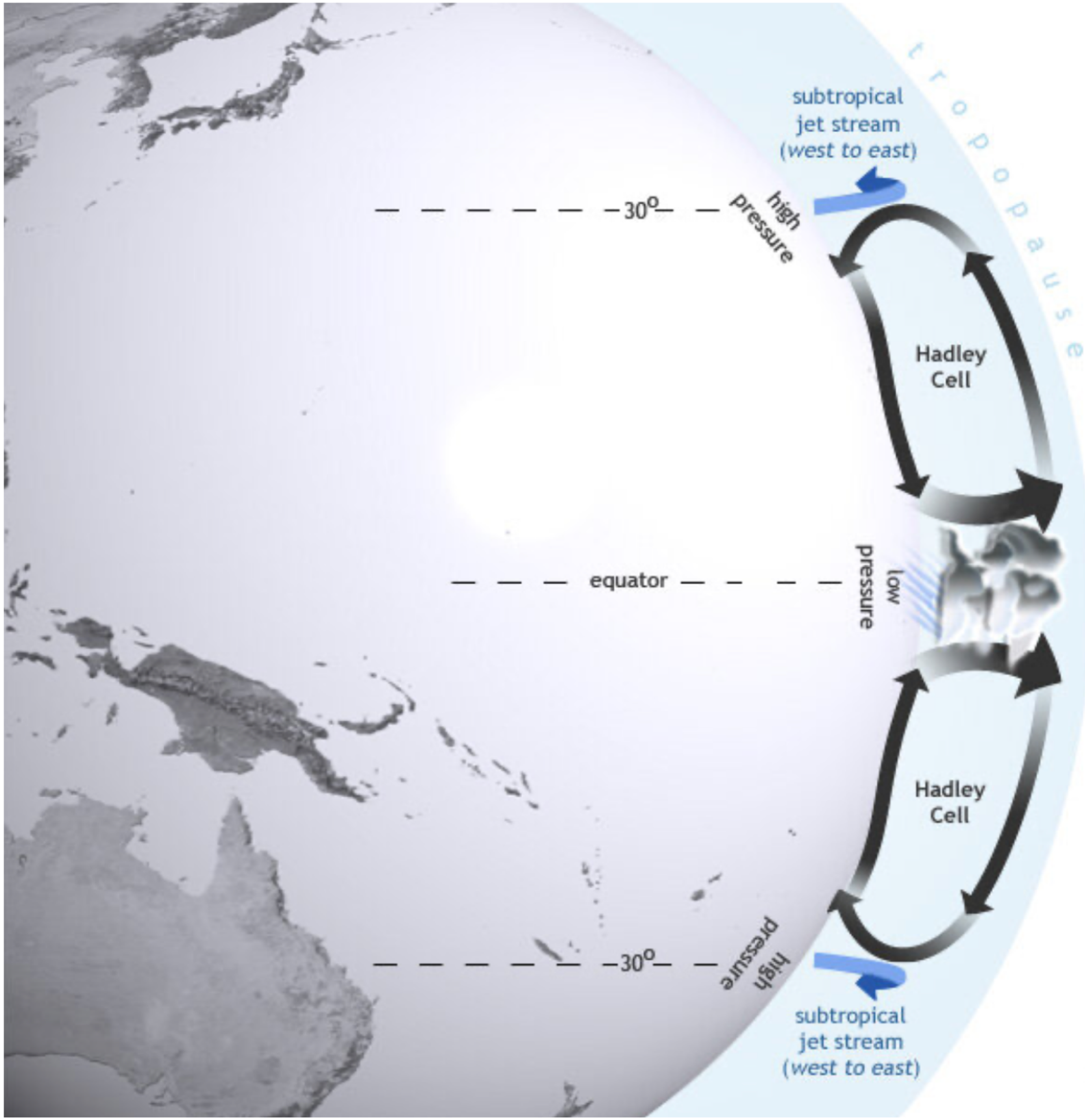
FIG. 1. The global annual mean Earth's energy budget for the Mar 2000 to May 2004 period (W m^{-2}). The broad arrows indicate the schematic flow of energy in proportion to their importance.

Trenberth *et al.*, *BAMS*, 2009

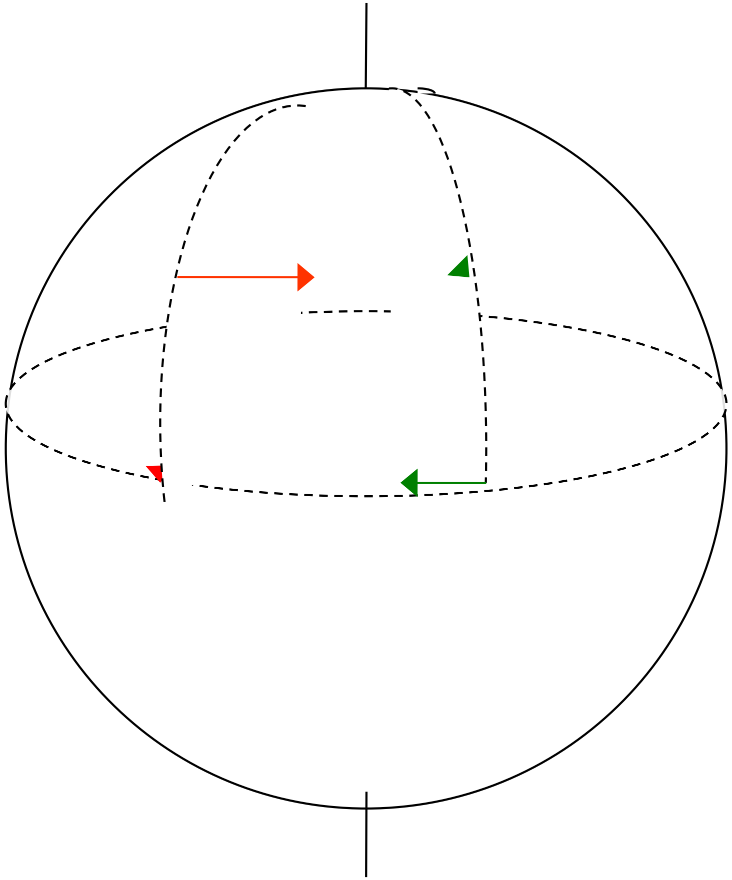


Annual mean energy budget of atmosphere for period 2000–2010. Unit W.m^{-2}

Stephens *et al*., *Nature Geoscience,* 2012



tropopause
subtropical
jet stream
(*west to east*)
30°
high
pressure
Hadley
Cell
equator
low
pressure
Hadley
Cell
30°
high
pressure
subtropical
jet stream
(*west to east*)



Particle moves on sphere with radius R
under the action of a force lying
in meridian plane of the particle

$\rightarrow$ Angular momentum wrt axis of rotation

$$(u + \Omega R \cos\varphi)\, R \cos\varphi = \mathrm{Cst}$$

On Earth, $\Omega \approx 2\pi\, 10^{-5}\, s^{-1}$, $R \approx 6.4\ 10^{6}\ m$.

If $u = 0$ at equator, $u = 329\ ms^{-1}$ at latitude $\varphi = 45^\circ$. If $u = 0$ at 45°, $u = -232\ ms^{-1}$ at equator.

Hadley, G., 1735, Concerning the cause of the general trade winds, *Philosophical Transactions of the Royal Society*

Fig. 10.1. Streak photographs illustrating the dependence of the flow type on rotation rate Ω for a laboratory 'dishpan' experiment. The values of Ω in rad s^{-1} are (a) 0.41; (b) 1.07; (c) 1.21; (d) 3.22; (e) 3.91; (f) 6.4. Working fluid was a water–glycerol solution of mean density 1.037 g cm^{-3} and kinematic viscosity 1.56×10^{-2} cm^2 s^{-1}. The streak photographs show the flow at a depth of 0.5 cm below the free upper surface (see also problem 10.1.) (From Hide & Mason, 1975)

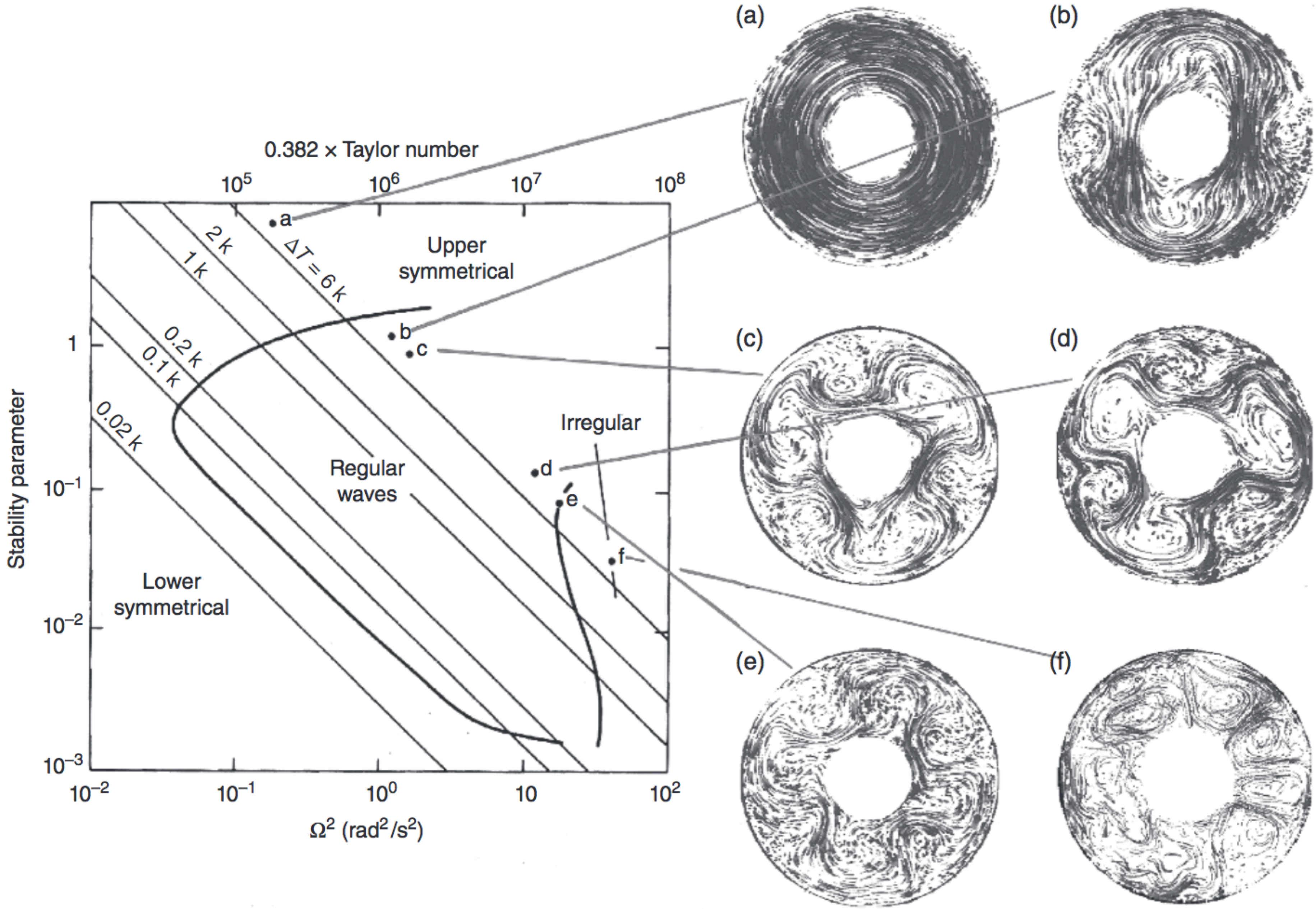


Figure 1.3. Schematic regime diagram for the thermally driven rotating annulus in relation to the thermal Rossby number Θ (or stability parameter, $\propto \Omega^{-2}$) and Taylor number $\mathcal{T} \propto \Omega^2$, showing some typical horizontal flow patterns at the top surface, visualized as streak images at upper levels of the experiment.

Read *et al*., 2015

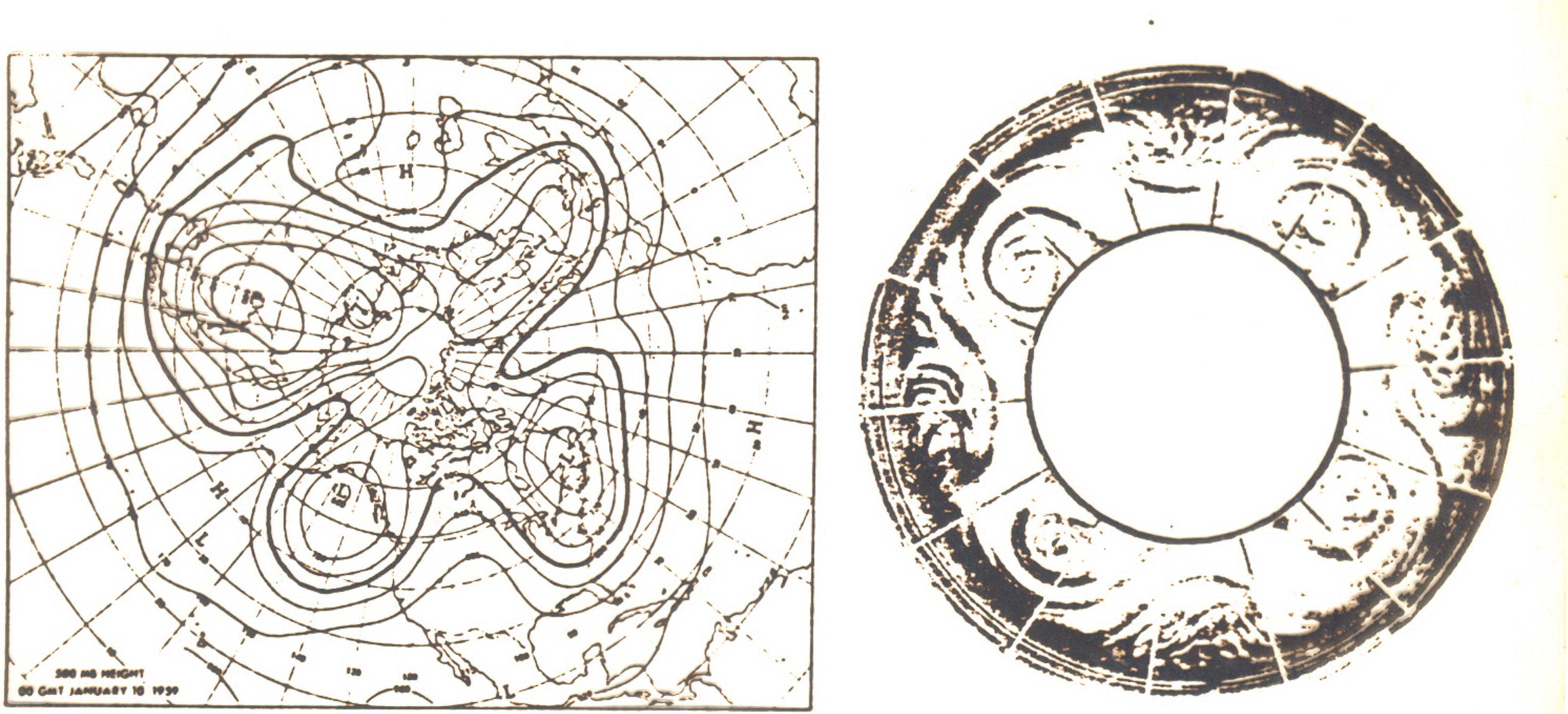
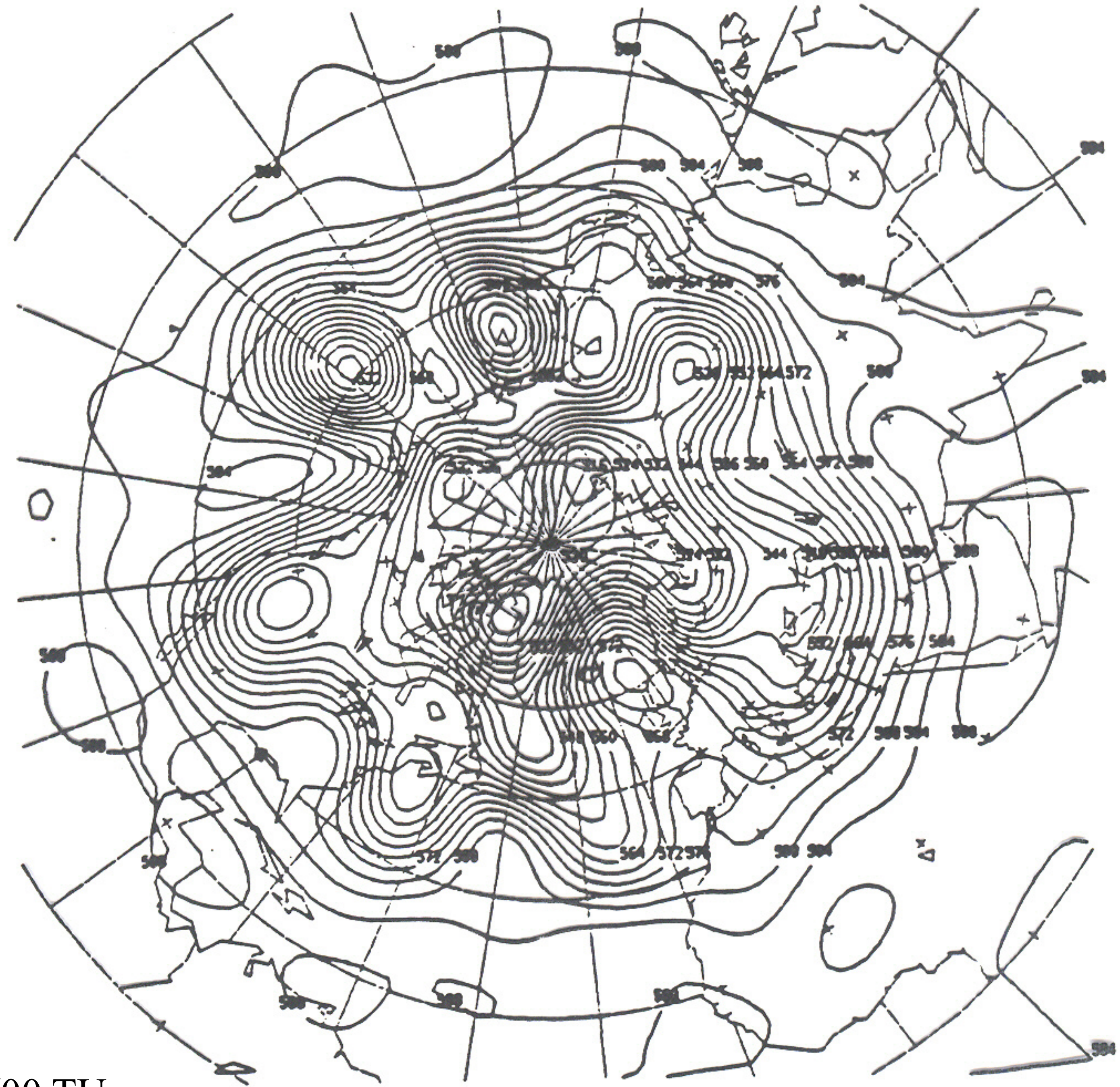
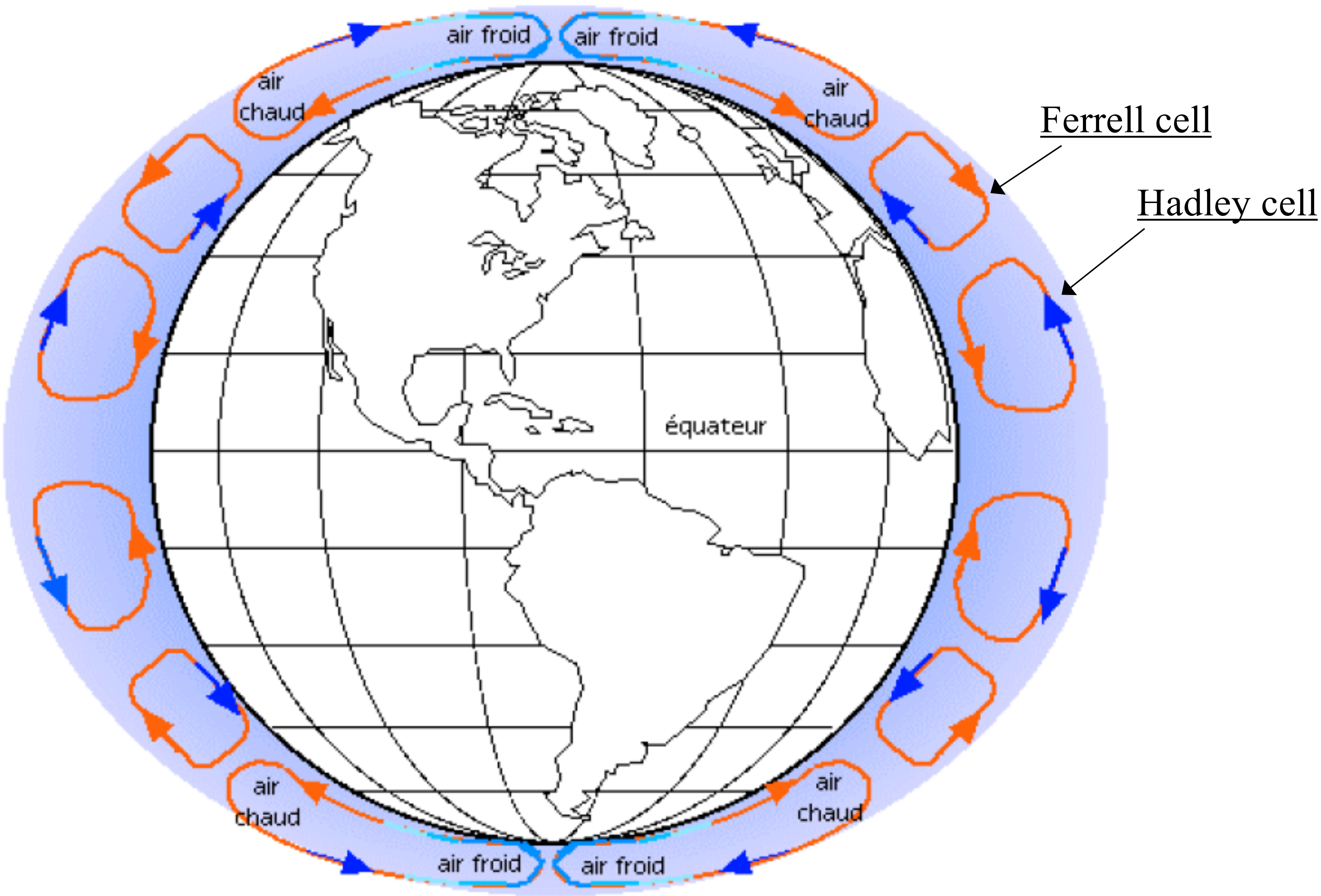
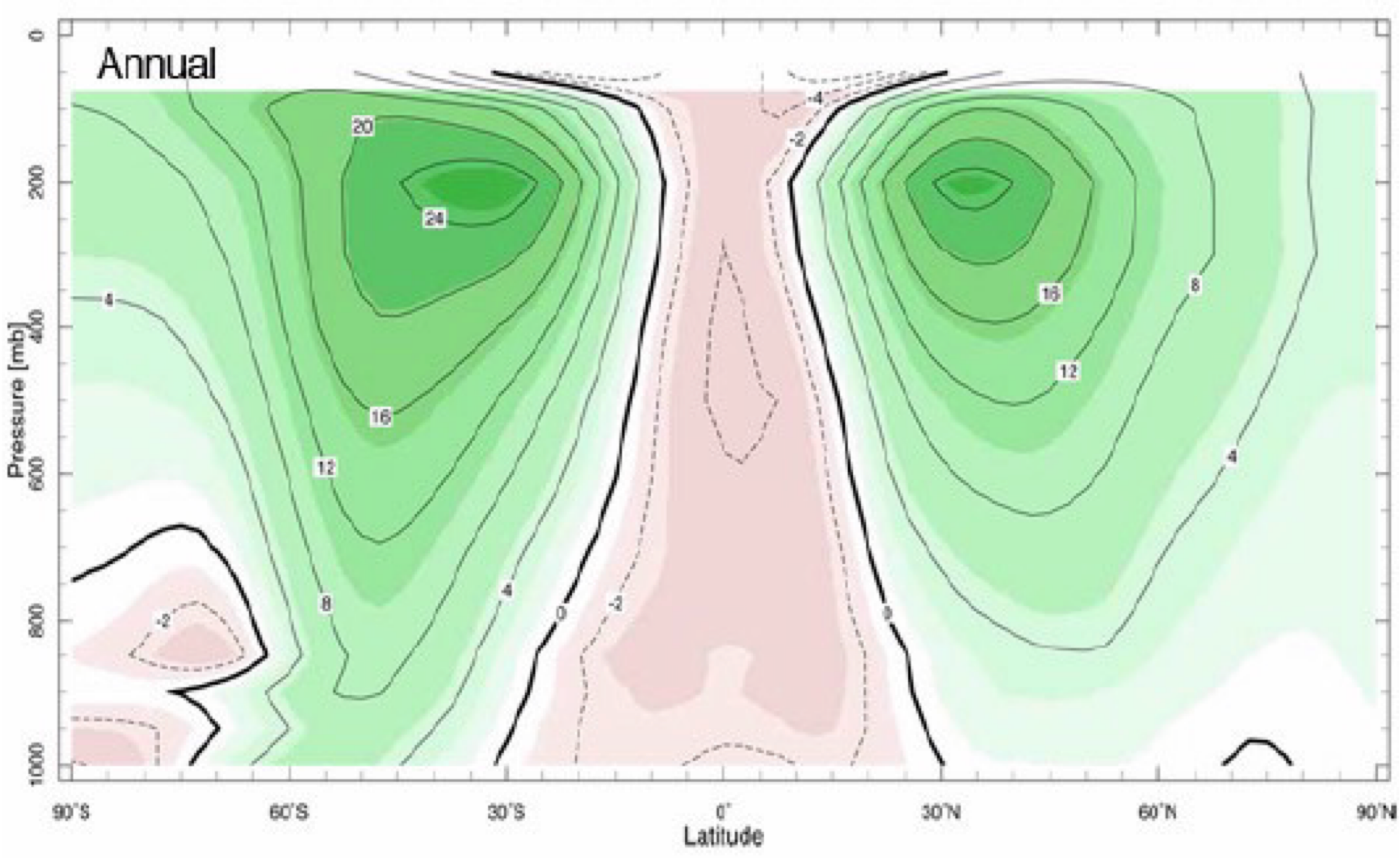


Figure 2. Comparison shows similarities between the global 500 mb pressure pattern in the upper atmosphere of the Northern Hemisphere and a four-wave pattern in the laboratory. (Laboratory flow conditions were similar to those in Fig. 1, except $\Omega = 1.95$ radians per sec.) In the atmosphere the flow is approximately parallel to the isobars (the flow is to the right, from high to low pressure), with speed inversely proportional to the spacing. Changes in the wave pattern have a significant effect on large-scale weather and climate.

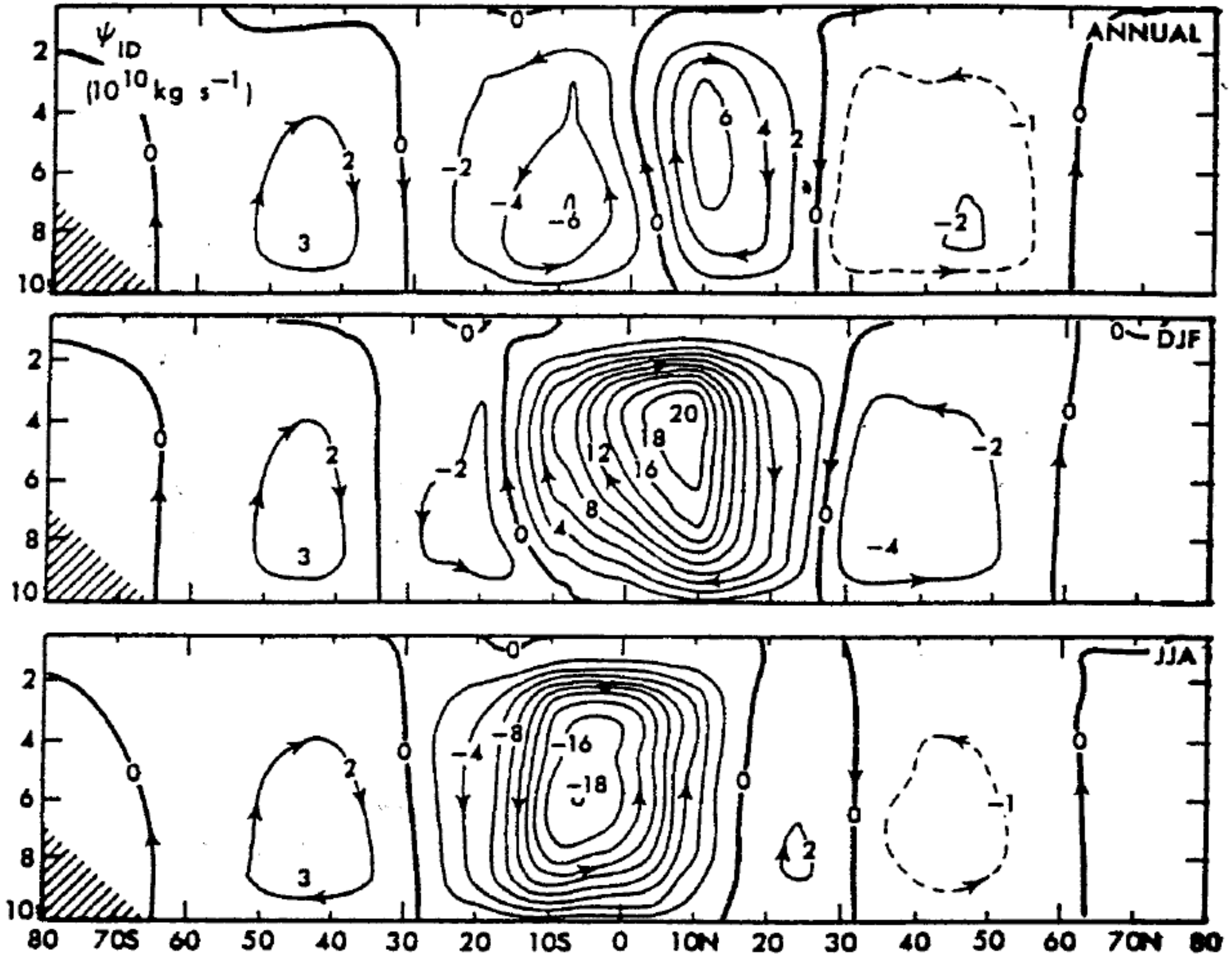


26/04/1984, 00/00 TU


air froid
air froid
air chaud
air chaud
Ferrell cell
Hadley cell
équateur
air chaud
air chaud
air froid
air froid

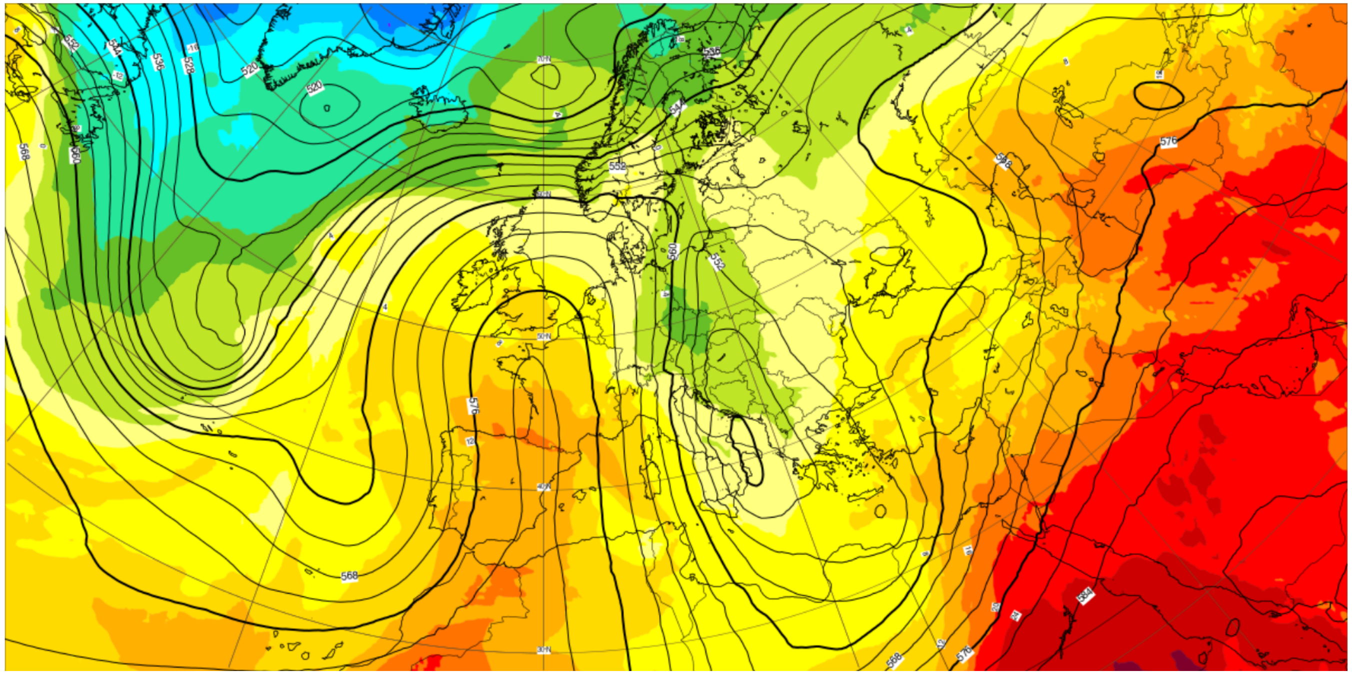


Zonal wind; annual longitudinal average (m.s^{-1})
http://paoc.mit.edu/labweb/notes/chap5.pdf,
Atmosphere, Ocean and Climate Dynamics, by J. Marshall and R. A. Plumb,
International Geophysics, Elsevier)

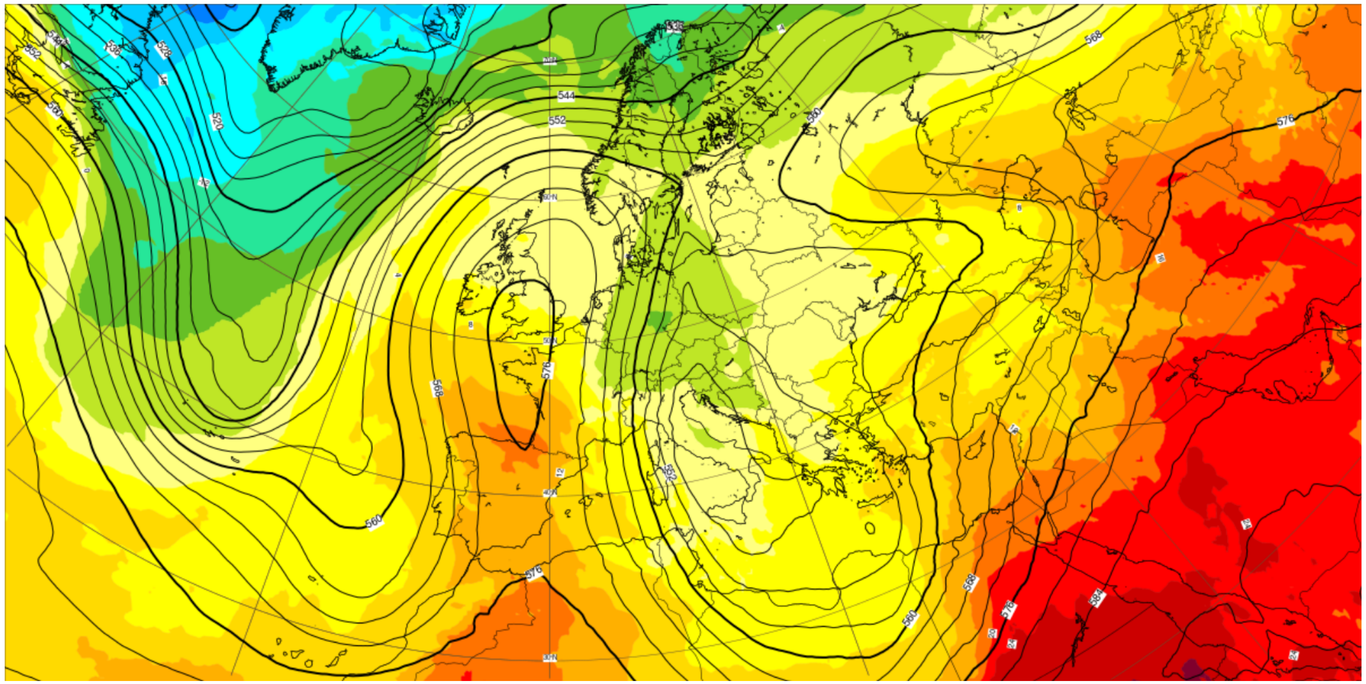


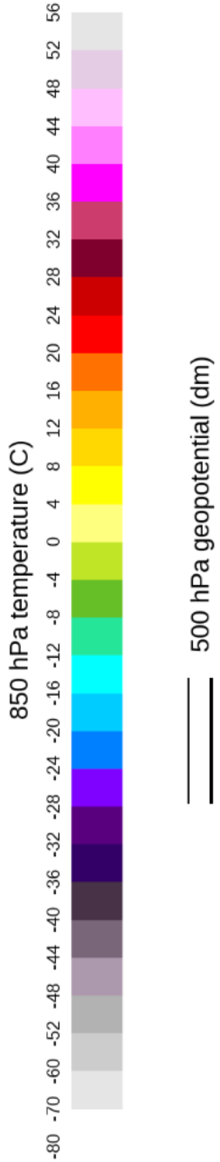
Peixoto and Oort, 1992, *The Physics of Climate,* Springer-Verlag

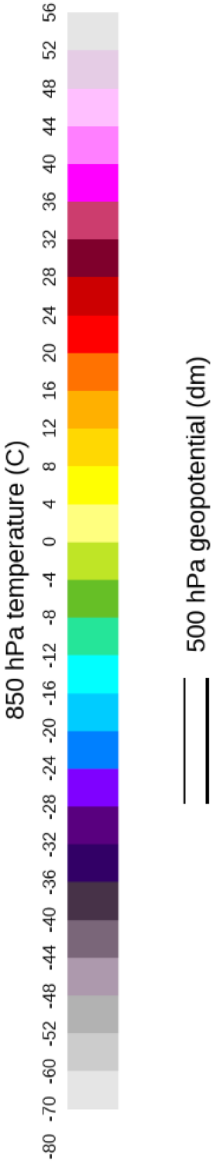
Base time: Thu 27 Mar 2025 00 UTC Valid time: Tue 01 Apr 2025 00 UTC (+120h) Area : Europe



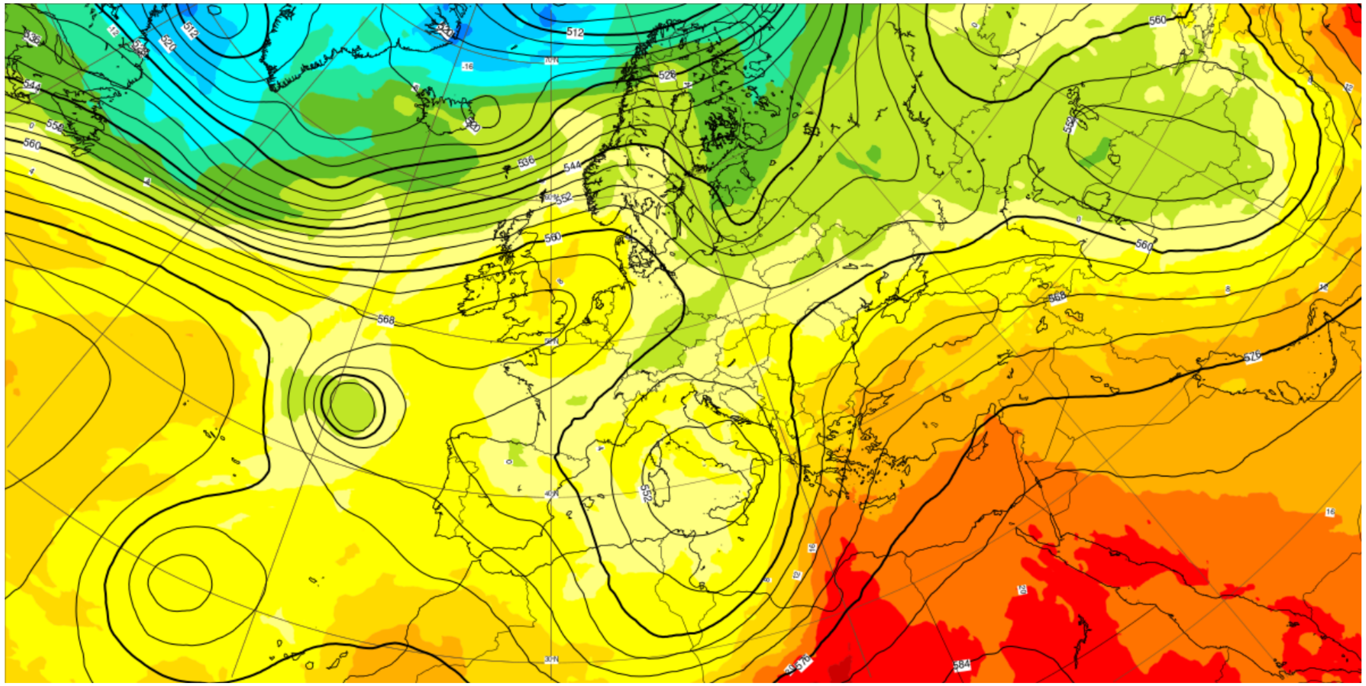
Base time: Tue 01 Apr 2025 00 UTC Valid time: Tue 01 Apr 2025 00 UTC (+0h) Area : Europe



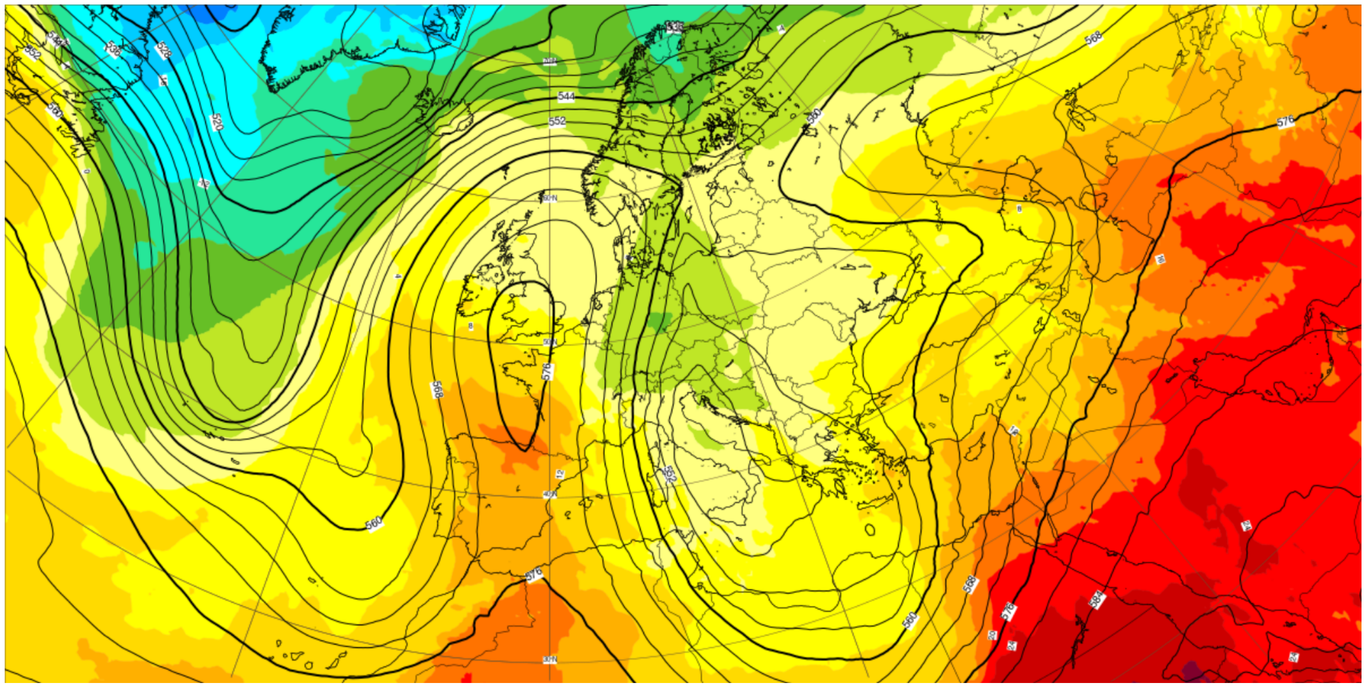




Base time: Thu 27 Mar 2025 00 UTC Valid time: Thu 27 Mar 2025 00 UTC (+0h) Area : Europe



Base time: Tue 01 Apr 2025 00 UTC Valid time: Tue 01 Apr 2025 00 UTC (+0h) Area : Europe



Base time: Tue 01 Apr 2025 00 UTC Valid time: Thu 10 Apr 2025 00 UTC (+216h) Area : Europe

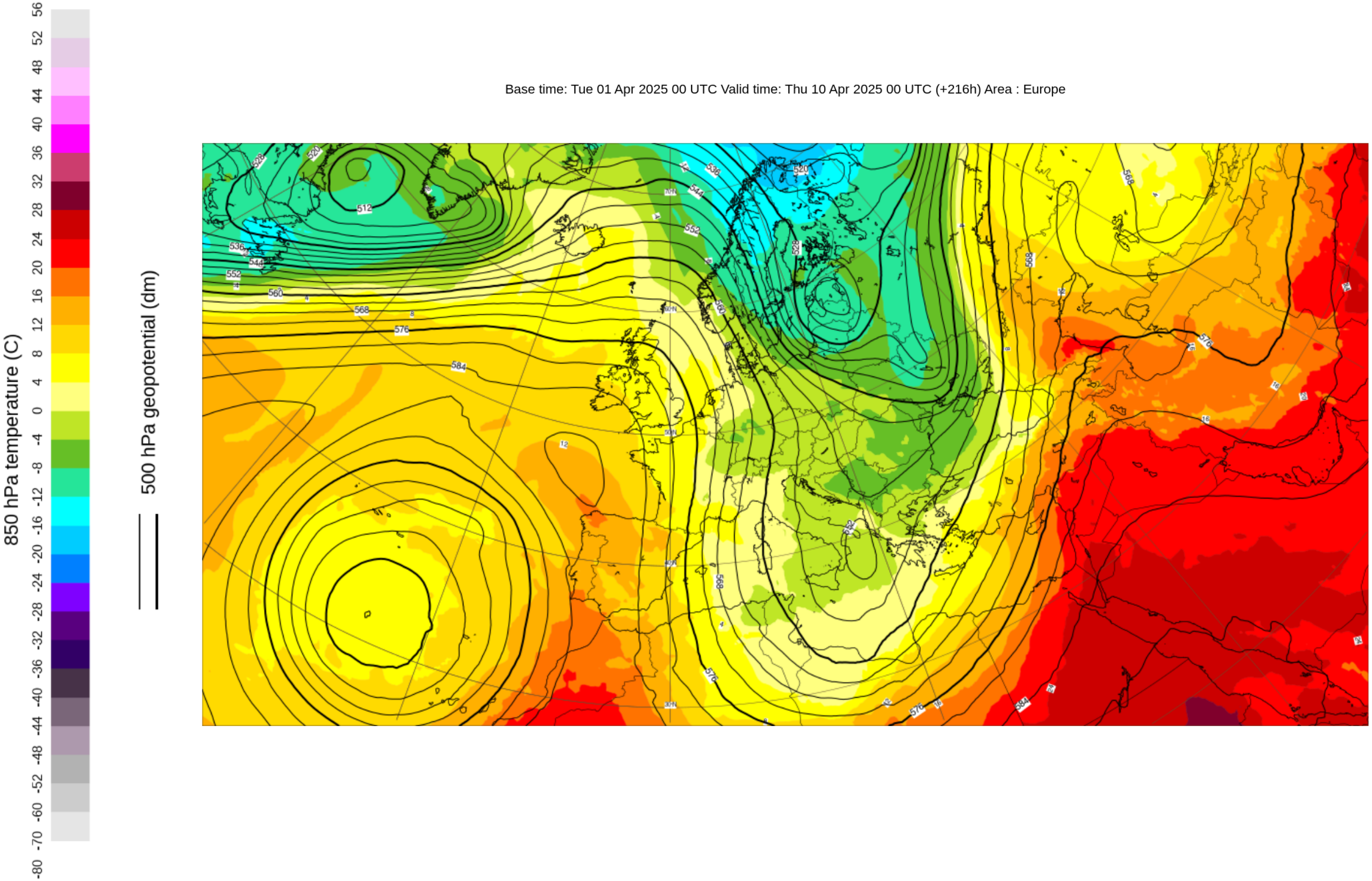


Fig. 1: Members of day 7 forecast of 500 hPa geopotential height for the ensemble originated from 25 January 1993.

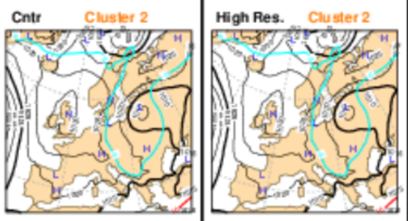
Figure 6 Hurricane Katrina mean-sea-level-pressure (MSLP) analysis for 12 UTC of 29 August 2005 and t+84h high-resolution and EPS forecasts started at 00 UTC of 26 August:

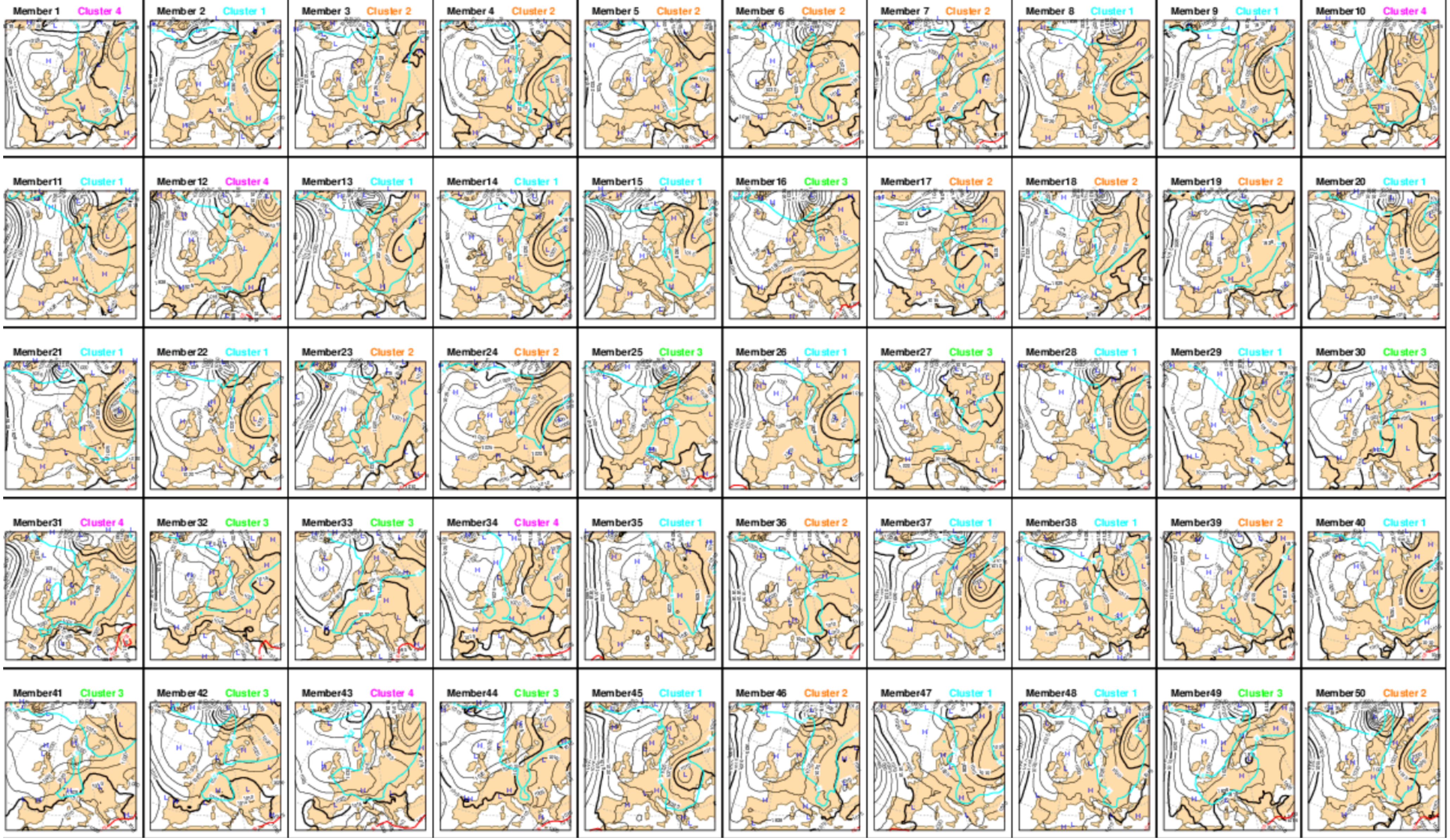
1st row: 1st panel: MSLP analysis for 12 UTC of 29 Aug
2nd panel: MSLP t+84h $T_L511L60$ forecast started at 00 UTC of 26 Aug
3rd panel: MSLP t+84h EPS-control $T_L255L40$ forecast started at 00 UTC of 26 Aug
Other rows: 50 EPS-perturbed $T_L255L40$ forecast started at 00 UTC of 26 Aug.

The contour interval is 5 hPa, with shading patters for MSLP values lower than 990 hPa.

ECMWF ENSEMBLE FORECASTS

Tuesday 01 April 2025 0000 UTC ECMWF forecast t+168 VT:Tuesday 08 April 2025 0000 UTC
MSLP (contour every 5hPa) Temperature at 850hPa (only -6 and 16 isolines are plotted)





Pourquoi les météorologistes ont-ils tant de peine à prédire le temps avec quelque certitude ? Pourquoi les chutes de pluie, les tempêtes elles-mêmes nous semblent-elles arriver au hasard, de sorte que bien des gens trouvent tout naturel de prier pour avoir la pluie ou le beau temps, alors qu'ils jugeraient ridicule de demander une éclipse par une prière ? Nous voyons que les grandes perturbations se produisent généralement dans les régions où l'atmosphère est en équilibre instable. Les météorologistes voient bien que cet équilibre est instable, qu'un cyclone va naître quelque part ; mais où, ils sont hors d'état de le dire ; un dixième de degré en plus ou en moins en un point quelconque, le cyclone éclate ici et non pas là, et il étend ses ravages sur des contrées qu'il aurait épargnées. Si on avait connu ce dixième de degré, on aurait pu le savoir d'avance, mais les observations n'étaient ni assez serrées, ni assez précises, et c'est pour cela que tout semble dû à l'intervention du hasard.

H. Poincaré, *Science et Méthode,* Paris, 1908

Why have meteorologists such difficulty in predicting the weather with any certainty? Why is it that showers and even storms seem to come by chance, so that many people think it quite natural to pray for rain or fine weather, though they would consider it ridiculous to ask for an eclipse by prayer? We see that great disturbances are generally produced in regions where the atmosphere is in unstable equilibrium. The meteorologists see very well that the equilibrium is unstable, that a cyclone will be formed somewhere, but exactly where they are not in a position to say; a tenth of a degree more or less at any given point, and the cyclone will burst here and not there, and extend its ravages over districts it would otherwise have spared. If they had been aware of this tenth of a degree they could have known it beforehand, but the observations were neither sufficiently comprehensive nor sufficiently precise, and that is the reason why it all seems due to the intervention of chance.

H. Poincaré, *Science et Méthode*, Paris, 1908

(English transl. by F. Maitland, *Science and Method*,

T. Nelson and Sons, London, 1914)

The case of the ocean

Same basic physics.

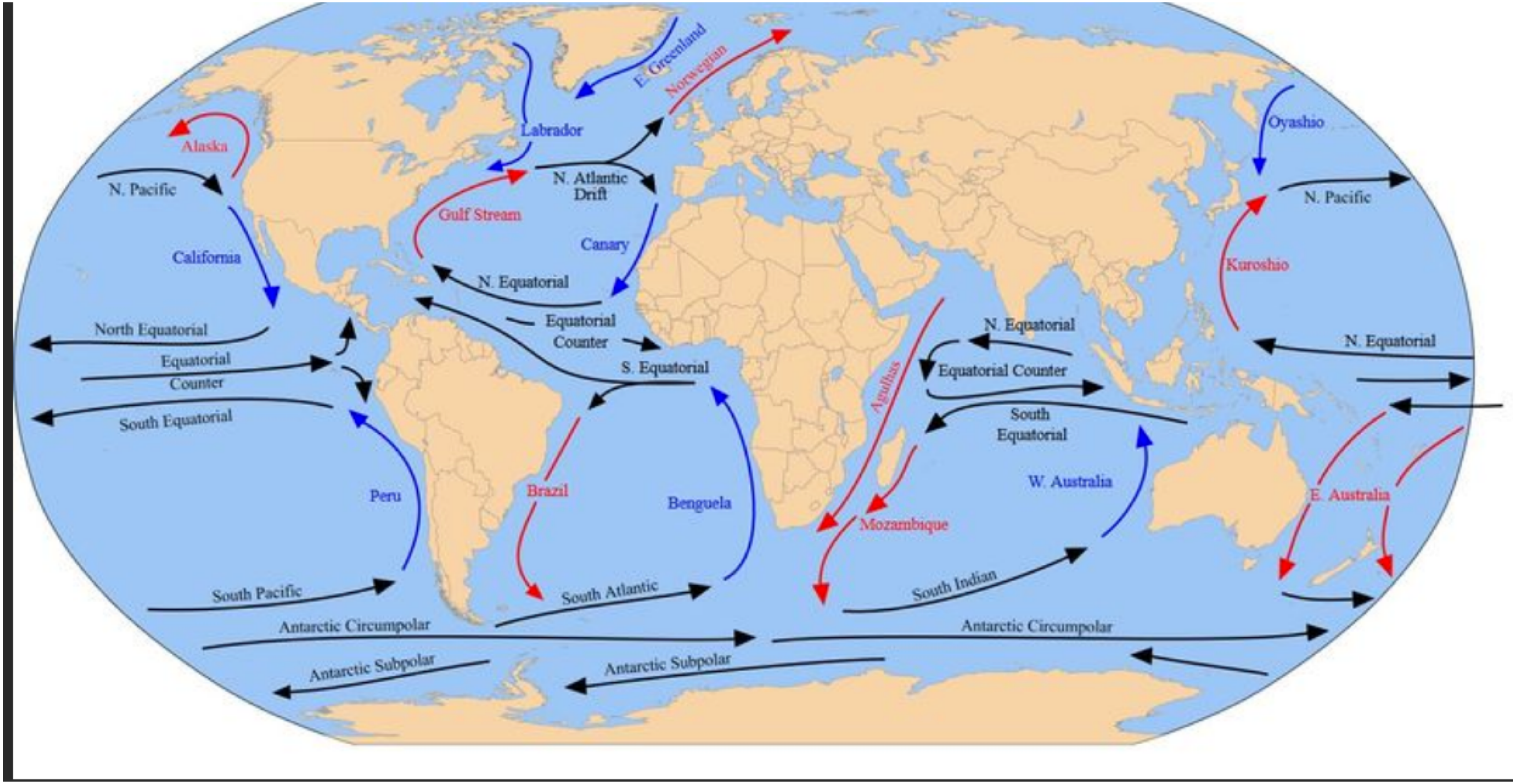
Absorption of solar energy limited to upper layers ocean heated from above). Circulation largely due to surface friction by wind.

But

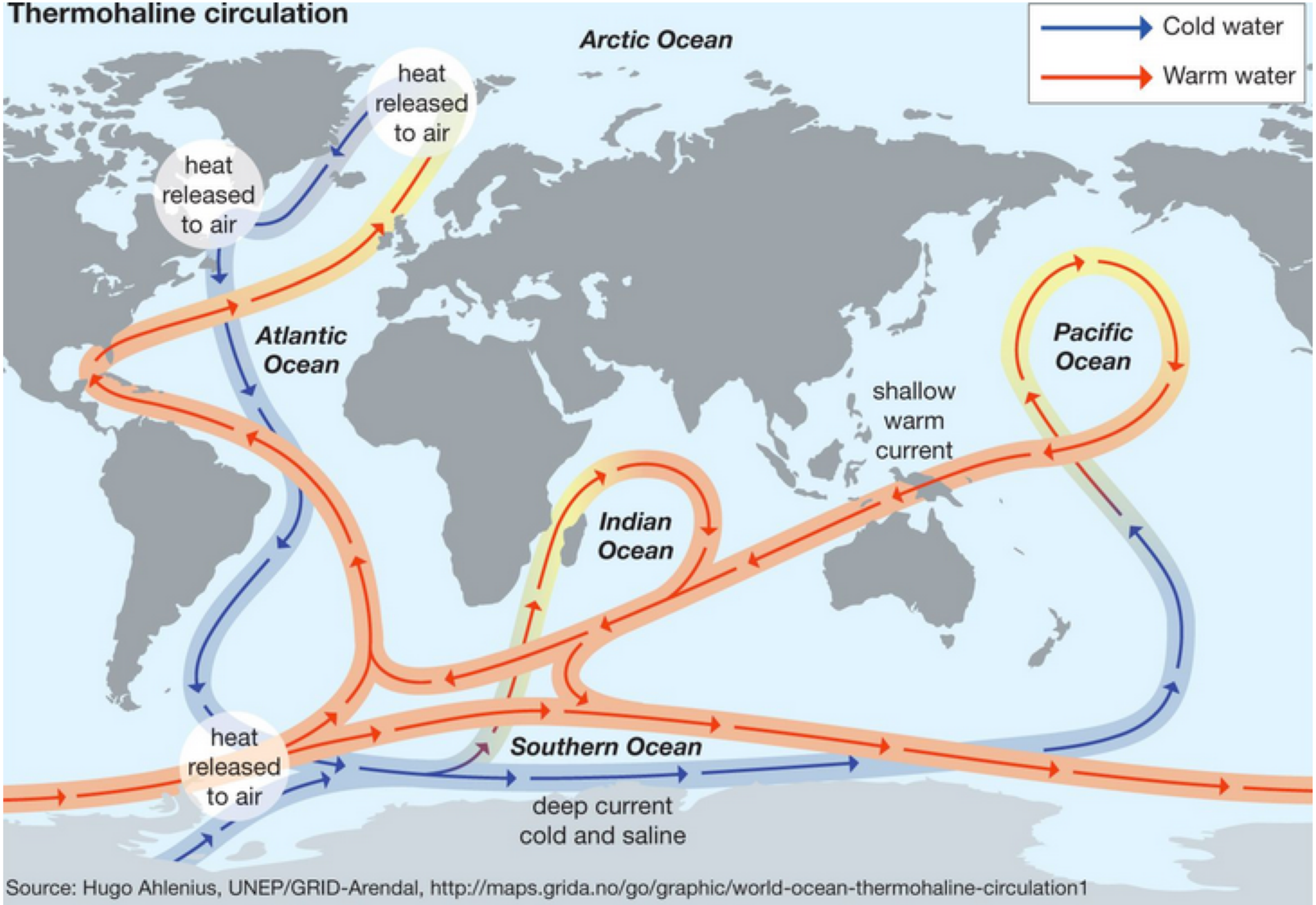
- Closed oceanic basins

- Importance of density, which strongly depends on salinity (*thermohaline circulation*)

Latitudinal temperature gradient, rotation, poleward energy transport



Surface oceanic circulation



Conveyor belt

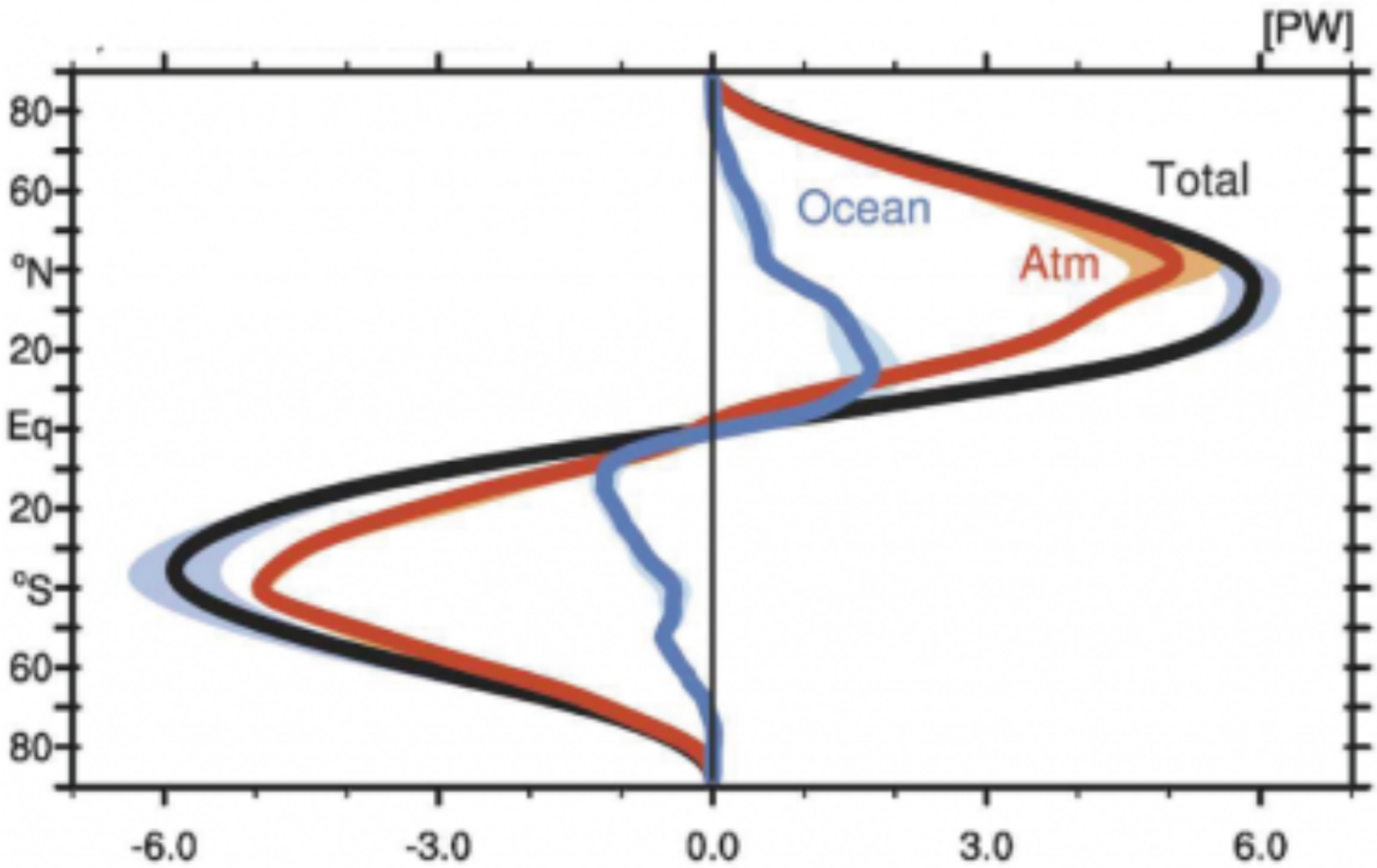


Fig. 2: The oceanic and atmospheric meridional heat transport in Petawatt per latitude

1 PW = 10^{15} W

Authors: Benjamin Meier and Julia Wouters

Physical laws governing the flow

- Conservation of mass

  $D\rho/Dt + \rho\,\mathrm{div}\underline{U} = 0$

- Conservation of energy

  $De/Dt - (p/\rho^2)\,D\rho/Dt = Q$

- Conservation of momentum

  $D\underline{U}/Dt + (1/\rho)\,\underline{\mathrm{grad}p} - \underline{g} + 2\,\underline{\Omega}\wedge\underline{U} = \underline{F}$

- Equation of state

  $f(p, \rho, e) = 0$ $(p/\rho = rT,\ e = C_v T)$

- Conservation of mass of secondary components (water, chemical species, …)

  $Dq/Dt + q\,\mathrm{div}\underline{U} = S$

These physical laws must be expressed in practice in discretized (and necessarily imperfect) form, both in space and time

$$\underbrace{Da/Dt}_{\substack{\textit{Particular}\\ \textit{(Lagrangian)}\\ \textit{derivative}}} = \underbrace{\partial a/\partial t}_{\substack{\textit{Eulerian}\\ \textit{derivative}}} + \underbrace{\underline{U}.\underline{\mathrm{grad}}a}_{\substack{\textit{Advection}\\ \textit{(due to motion)}}}$$

The case of the ocean

- Same basic equations

But

- Different equation of state

- Major secondary component is now salt (convective instability associated with variation of density of the fluid)

Physical laws must in practice be discretized in both space and time

$\Rightarrow$ *numerical models*, which are necessarily imperfect.

Models that are used for large scale weather prediction and for climatological simulation cover the whole volume of the atmosphere. These models are based, at least so far, on the *hydrostatic* hypothesis

in the vertical direction :

$\partial p/\partial z + \rho g = 0$

Eliminates momentum equation for vertical direction. In addition, flow is incompressible in coordinates (x, y, p) $\Rightarrow$ number of equations decreased by two units.

Hydrostatic approximation valid, to accuracy $\approx 10^{-4}$, for horizontal scales > 20-30 km

More costly nonhydrostatic models are used for small scale meteorology.

Using pressure p as independent vertical coordinate

- Flow is incompressible

- Pressure gradient term $(1/\rho)\ \mathrm{grad}_z\ p$ becomes $\mathrm{grad}_p \Phi$, where $\Phi \equiv gz$ is geopotential

Spatial Discretization

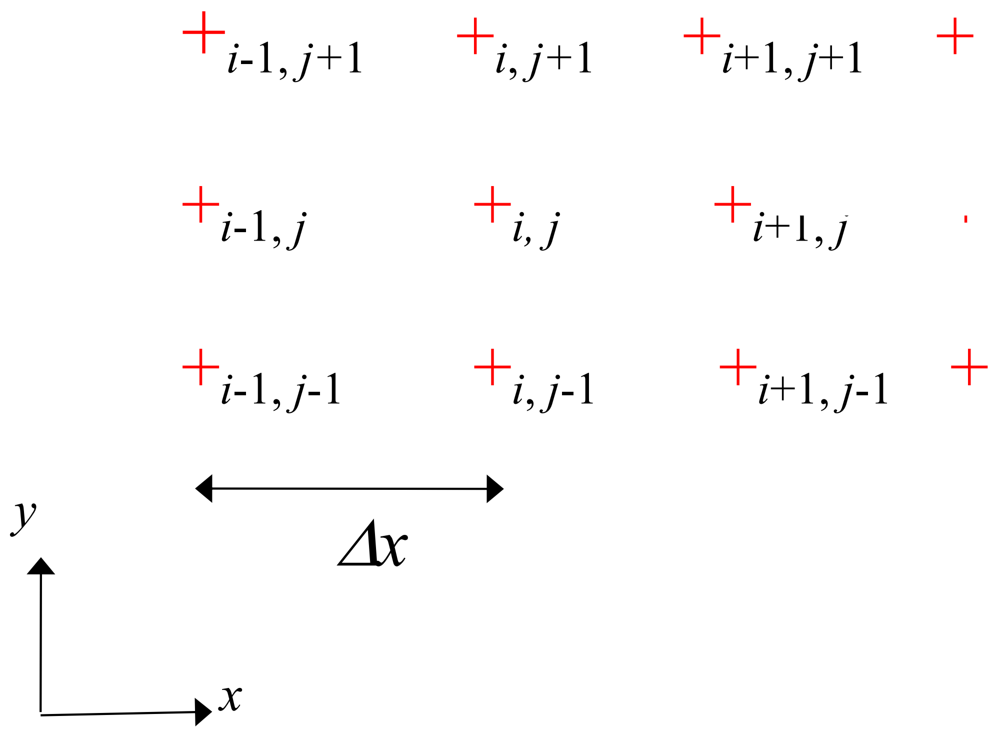
There exist at present two forms of spatial discretization

- Gridpoint discretization

- (Semi-)spectral discretization (mostly for global models, and most often only in the horizontal direction)

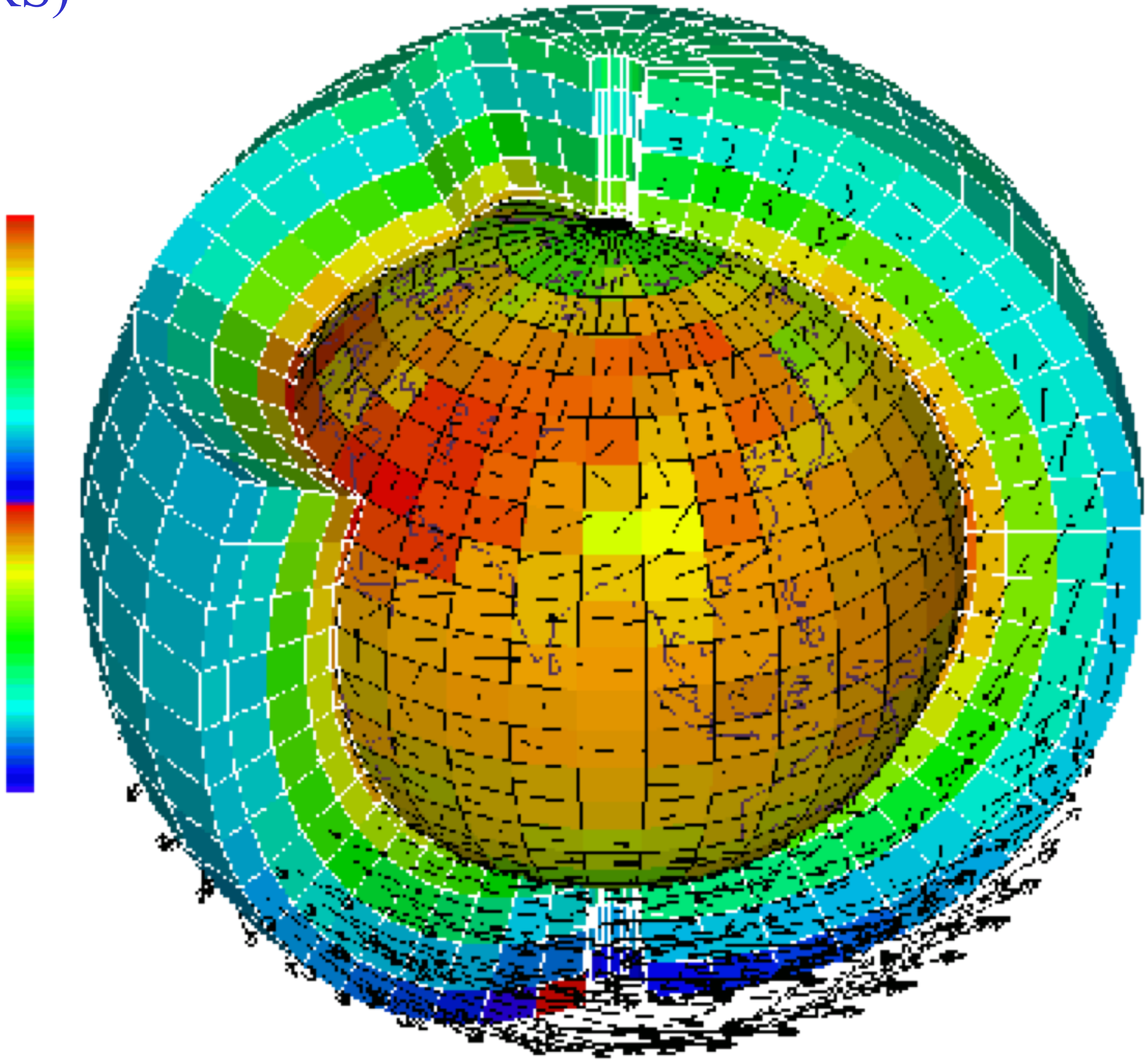
Finite element discretization, which is very common in many forms of numerical modelling, is rarely used for modelling of the atmosphere, except for discretization in the vertical direction. It is more frequently used for oceanic modelling, where it allows to take into account the complicated geometry of coast-lines.

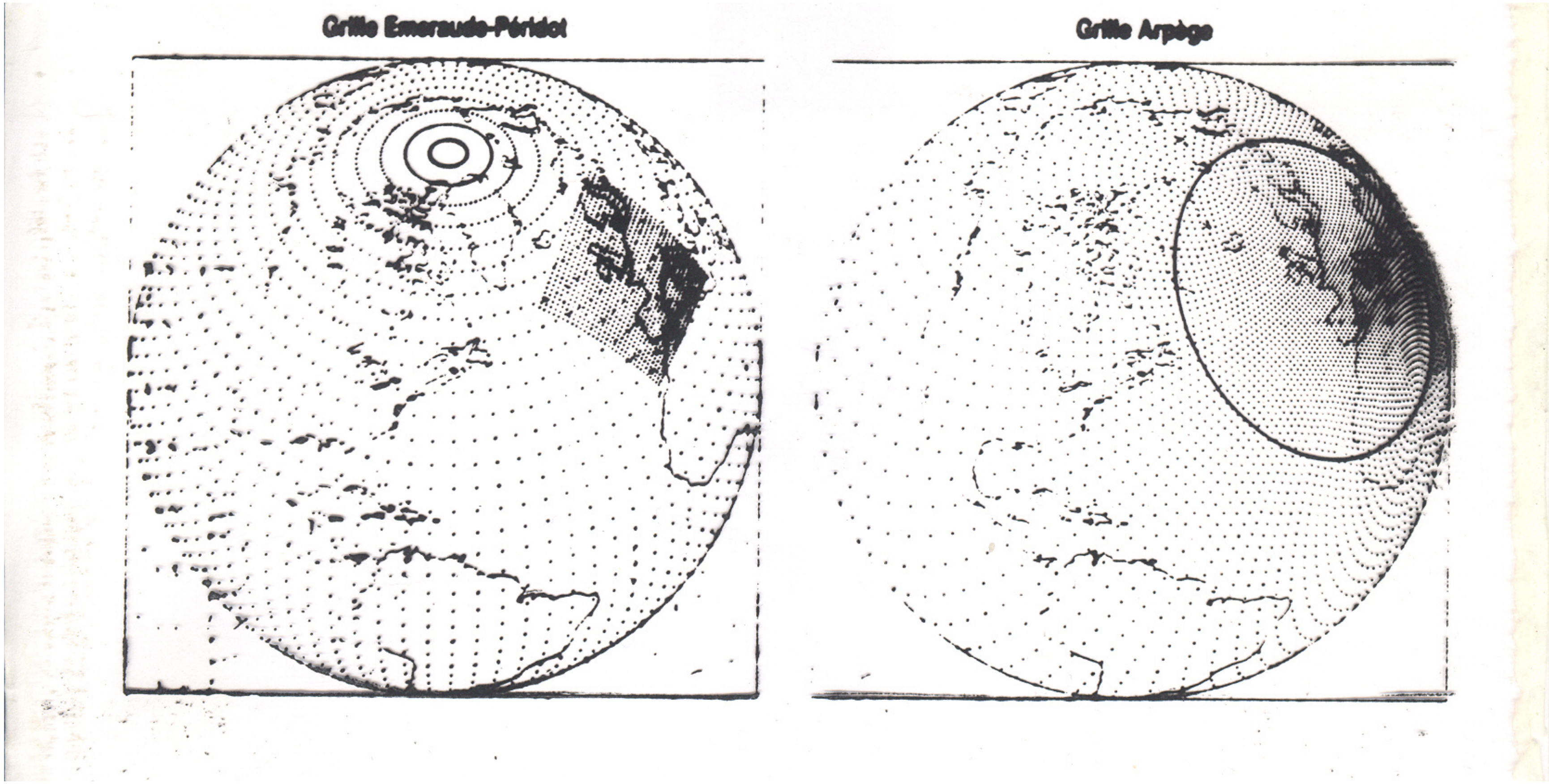
In gridpoint models, meteorological fields are defined by values at the nodes of a grid covering the physical domain under consideration. Spatial derivatives are expressed by finite differences.



$$\left(\frac{\partial \Phi}{\partial x}\right)_{i,j} \approx \frac{\Phi_{i+1,j} - \Phi_{i-1,j}}{2\,\Delta x}$$

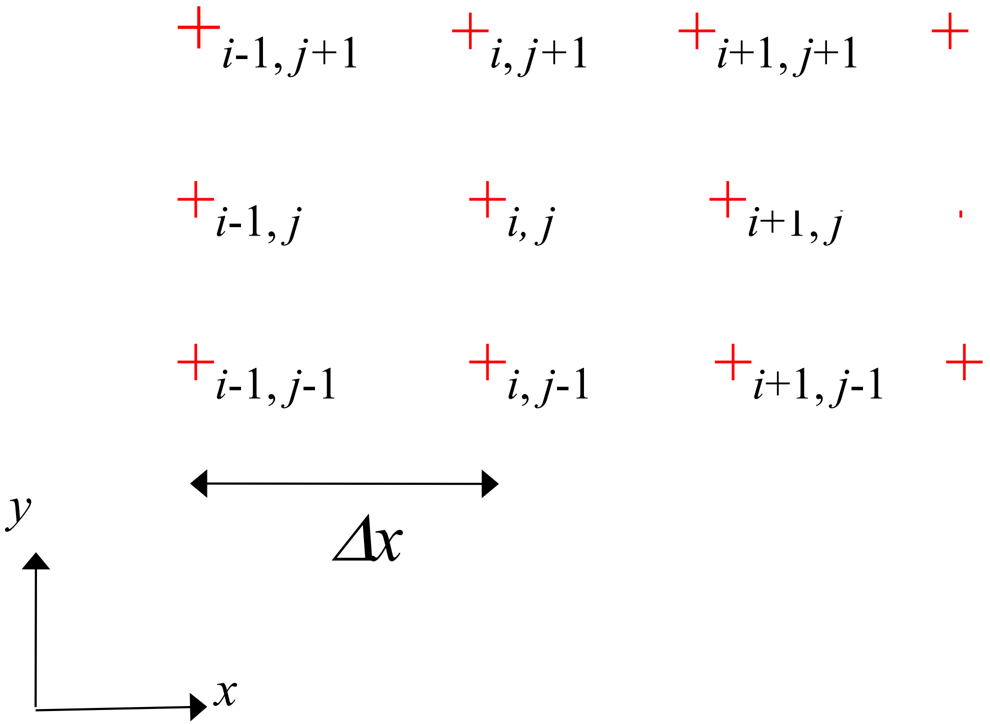
A schematic of an Atmospheric General Circulation Model (L. Fairhead /LMD-CNRS)





Grilles de modèles de Météo-France *(La Météorologie)*

Gridpoint discretization. Much care is to be taken in the definition of the discretizing schemes, to ensure for instance underlying conservation laws (mass, energy, momentum, …), both locally and globally



$$\left(\frac{\partial \Phi}{\partial x}\right)_{i,j} \approx \frac{\Phi_{i+1,j} - \Phi_{i-1,j}}{2\,\Delta x}$$

In *spectral models*, fields are defined by the coefficients of their expansion along a prescribed set of basic functions. This is similar to Fourier expansion in a periodic domain in R^n, which uses imaginary exponential functions (sines and cosines) as basis functions

$$F(x, y) = \Sigma_{k,\, m} \mathcal{F}(k, m) \exp\,[2i\pi\,(kx/L_x + my/L_y)]$$

where the function F has periods L_x and L_y in the directions x and y respectively.

In the discretized case, the integer indices k and l are limited to a set of finite values

$$k = -K, \ldots, 0, \ldots, K \quad ; \quad m = -M, \ldots, 0, \ldots, M$$

Advantage : spatial differentiation is obvious and exact

In the case of global meteorological models, which cover the whole atmosphere, the basis functions are the *spherical harmonics*

$$T(\mu=\sin(\text{latitude}),\ \lambda=\text{longitude}) = \sum_{\substack{0\leq n<\infty \\ -n\leq m\leq n}} T_n^m Y_n^m(\mu,\lambda)$$

où les $Y_n^m(\mu,\lambda)$ sont les *harmoniques sphériques*

$$Y_n^m(\mu,\lambda) \propto P_n^m(\mu)\exp(im\lambda)$$

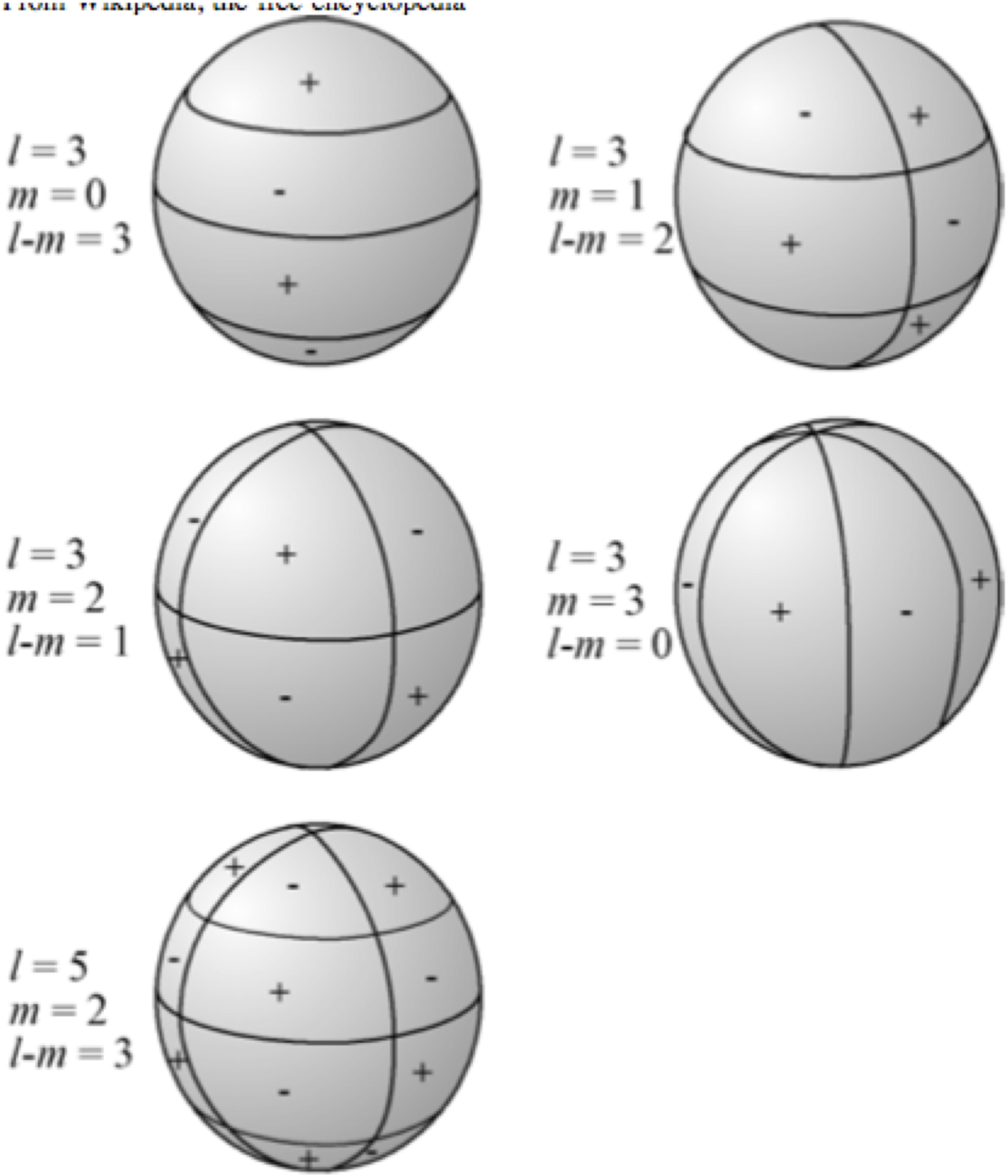
$P_n^m(\mu)$ est la *fonction de Legendre* de deuxième espèce

$$P_n^m(\mu) \propto (1-\mu^2)^{\frac{m}{2}} \frac{d^{n+m}}{d\mu^{n+m}}(\mu^2-1)^n$$

n et m sont respectivement le *degré* et l'*ordre* de l'harmonique $Y_n^m(\mu,\lambda)$

$n = 0, 1, \ldots \qquad -n \leq m \leq n$

From Wikipedia, the free encyclopedia



Spherical harmonics, such as Fourier harmonics in R^n, are eigenfunctions of Laplacian Δ on the surface of the sphere

$$\Delta Y_n^m = -n(n+1)Y_n^m$$

'Triangular' truncation *TN* ($n \le N$, $-n \le m \le n$) is independent of choice of polar axis. Corresponding representation is perfectly homogeneous on the surface of the sphere

Linear computations, such as spatial differentiation, are performed in spectral space. Nonlinear computations are performed in physical space, on a grid that is appropriate for avoiding *aliasing* (often a latitude-longitude grid, called 'gaussian'). The required changes of representation can be performed at a non-prohibitive cost in the longitudinal direction through the use of *Fast Fourier Transforms*. There also exist a fast version of Legendre transforms relative to the variable μ.

Owing to those repeated transformations between physical and spectral spaces, the corresponding models are often called *semi-spectral*.

In addition to hydrostatic approximation, the following approximations are (almost) systematically made in global modeling :

- Atmospheric fluid is contained in a spherical shell with negligible thickness. This does not forbid the existence within the shell of a vertical coordinate which, in view of the hydrostatic equation, can be chosen as the pressure p.

- The horizontal component of the Coriolis acceleration due to the vertical motion is neglected (this approximation, sometimes called the *traditional approximation*, is actually a consequence of the previous one).

- Tidal forces are neglected.

These approximations lead to the so-called (and ill-named) *primitive equations*

Pressure p, although convenient for writing down the equations, is in fact rather inconvenient because lower boundary is not fixed in (x, y, p)-space.

So-called σ-coordinate. $\sigma \equiv p/p_S$, where p_S is pressure at ground level.

'Hybrid' coordinate.

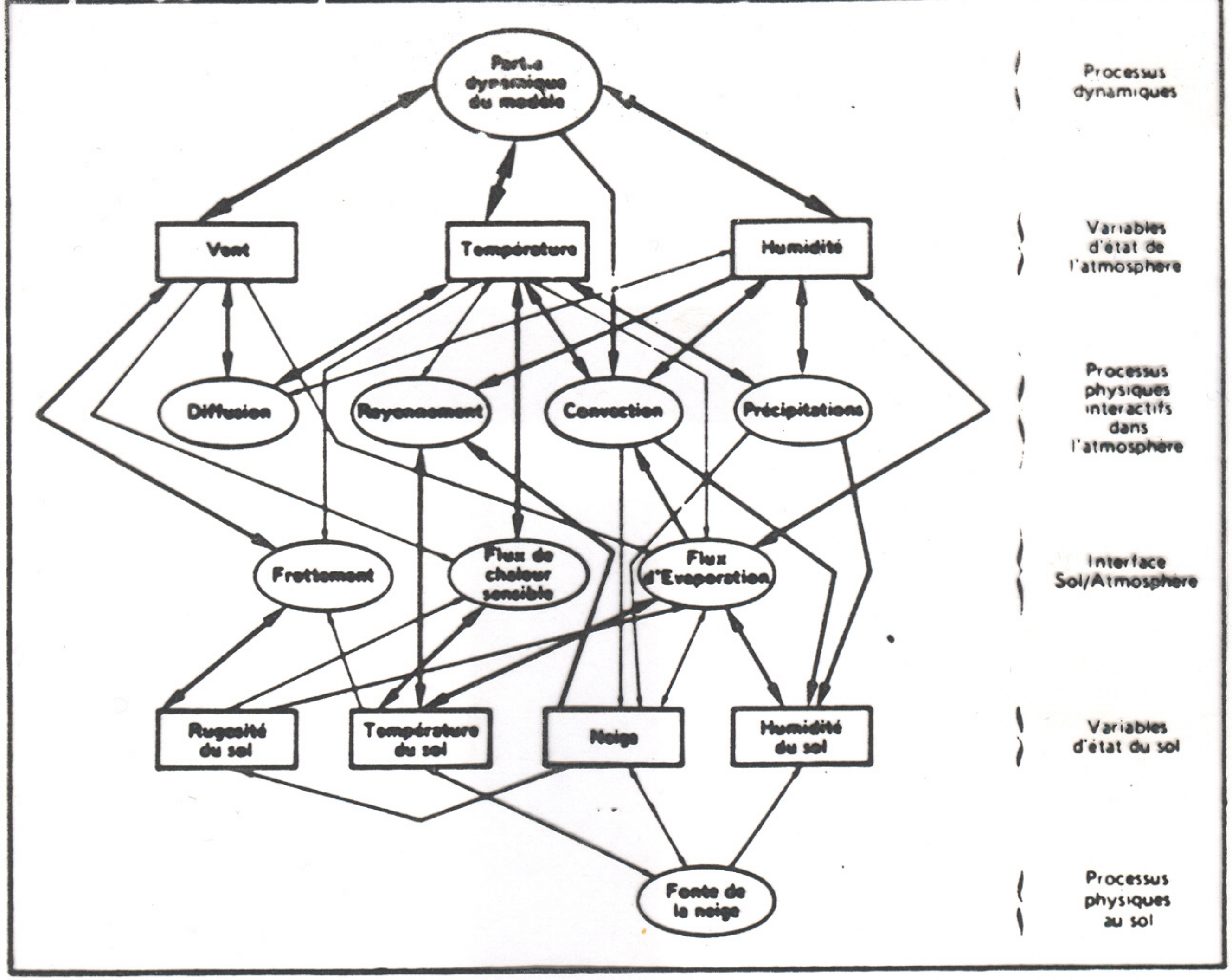
Parlance of the trade :

One ordinarily distinguishes two different parts in models. The 'dynamics' deals with the physically reversible processes (pressure forces, Coriolis force, advection, …), while the 'physics' deals with physically irreversible processes, in particular the diabatic heating term Q in the energy equation, and also the parameterization of subgrid scales effects.

Numerical schemes have been gradually developed and validated for the 'dynamics' component of models, which are by and large considered now to work satisfactorily (although regular improvements are still being made; project *DYNAMICO*, *Dynamical Core on Icosahedral Grid*, Th. Dubos, IPSL).

The situation is different as concerns ‘physics’, where many problems remain (relative for instance to subgrid scales parameterization, the water cycle and the associated exchanges of energy, or the exchanges that take place in the boundary layer between the atmosphere and the underlying medium). ‘Physics’ as a whole remains the weaker point of models, and is still the object of active research.

5 - SCHEMA DES INTERACTIONS PHYSIQUES DANS LE MODELE



Temporal Discretization

Equation

$$dx / dt = F(x)$$

(x state vector of the model).

Timestep Δt.

Computed solution at time $n\Delta t$ denoted x_n

Forward (Euler) scheme

$$(x_{n+1} - x_n)/\Delta t = F(x_n)$$

$$x_{n+1} = x_n + \Delta t\, F(x_n)$$

Implemented on equation

$$dx / dt = i\alpha x \quad , \qquad \alpha \text{ real} \qquad (1)$$

Exact solution $x(t) = x(0) \exp (i\alpha t)$

Modulus $|x(t)|$ conserved in time

Discretized solution according to forward scheme

$$x_{n+1} = (1 + i\alpha\Delta t)\, x_n$$

Modulus $|x_{n+1}| = \sqrt{(1 + \alpha^2\Delta t^2)}\, |x_n|$

increases exponentially with time.

Forward scheme is *unconditionally unstable* for Eq. (1)

Leapfrog scheme

$$(x_{n+1} - x_{n\text{-}1})/2\Delta t = F(x_n)$$
$$x_{n+1} = x_{n\text{-}1} + 2\Delta t\, F(x_n)$$

Stable for equation (1) above (*i.e.* modulus remains constant in time) provided

$$\alpha\Delta t < 1$$

Courant-Friedrichs-Lewy (CFL) condition

In a multidimensional system, the largest α will be the highest frequency that is present in the system. In a discretized system of travelling waves, the highest frequency will correspond to the fastest wave that the discretization can explicitly resolve. It will be proportional to $c/\Delta x$, where c is the phase velocity of the fastest waves in the system, and Δx the mesh-size of the discretization

$$\alpha = (1/\beta)\, c/\Delta x$$

where β is an $O(1)$ numerical coefficient depending on the particular discretization scheme under consideration.

CFL condition then becomes

$$\Delta t \,/\, \Delta x \; < \beta \,/\, c$$

Significance : numerical propagation of signal must be at least as fast as physical propagation.

CFL condition generally applies to explicit schemes of temporal discretization

In hydrostatic atmosphere, fastest propagating wave : gravity wave with largest scale height, $c = \sqrt{(rT)} \approx 300 \text{ m.s}^{-1}$.

$$\Delta x = 30 \text{ km} \quad \Rightarrow \quad \Delta t = 100 \text{ s}$$

The use of *semi-implicit* schemes allows to get rid of the CFL condition, and to use longer timesteps.

Crank-Nicolson

$$(x_{n+1} - x_n)/\Delta t = (1/2)\,[F(x_{n+1}) - F(x_n)]$$

Implicit !

Stable for equation (1) (modulus remains constant in time) for any Δt

More computations may be required for one timestep, but a longer timestep can be used.

The use of (at least partially) implicit schemes allows to get rid of the CFL condition, and to use longer timesteps without damage to physics.

Large-scale Numerical Weather Prediction is based on the *primitive equations*, based, as said, on a number of simplifications, and particularly the hydrostatic approximation

Climatic simulations are also built on primitive equations, and contain a much more detailed description of the oceanic circulation.

More costly nonhydrostatic models are used for small scale meteorology, and are being developed for global modeling.

Cours à venir

Mercredi 2 avril

Vendredi 11 avril

Vendredi 18 avril

Mercredi 23 avril

Lundi 12 mai nouvelle date

Mercredi 28 mai

Mercredi 11 juin

Mercredi 18 juin