

The Basics of Planetary Thermospheres

Alan Aylward Atmospheric Physics Laboratory, UCL

Including notes from Ingo Mueller-Wodarg (ICSTM)



".....There is, it seems to us,

At best only a limited value

In the knowledge derived from experience.

The knowledge imposes a pattern, and falsifies,

For the pattern is new in every moment

And every moment is a new and shocking

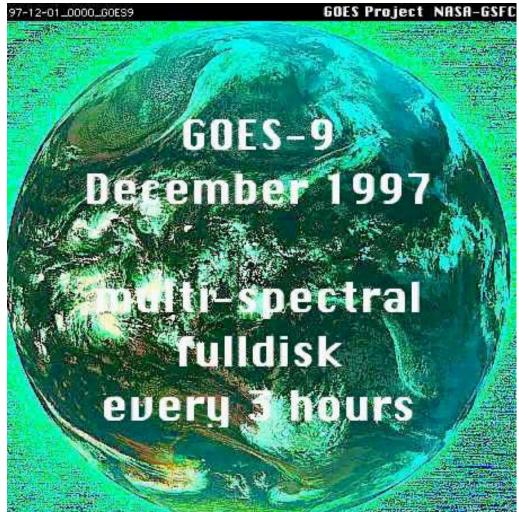
Valuation...."

T.S.Eliot "The Four Quartets"



Why do we study the "Upper" atmosphere?

Well all that meteorology is a bit too complex





It was so much simpler in the old days

- Until the advent of rockets we were limited to the lowest 11 km or so of the atmosphere
- Known from fairly early on that pressure and temperature dropped with height
- Early balloon observations took up pressure and temperature sensors
- In 1901 Teisseren de Bort realised that the temperature stopped falling at about 11 km altitude
- He invented the terms (1908) troposphere and stratosphere albeit on false assumptions



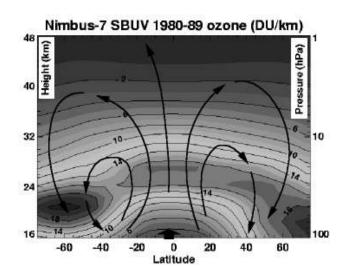




Progress through the 20th Century

- Several key figures advanced the theory of the upper atmosphere
- Sydney Chapman did some remarkable theoretical development based on gas theory and bringing together what was known at the time. He virtually invented "diffusive separation". Ozone layer theory
- 1920 Gordon Dobson at Oxford deduced temperature rise at high altitude from studies of meteor trails - then invented the ozone spectrometer and ozone spectrophotometer it was a Dobson instrument that discovered the ozone hole over Antarctica
- He is also the Dobson of "Brewer-Dobson" circulation (1940)
- Advances in theory of the ionosphere went hand in hand with this
- Sir Edward Appleton deployed the first ionosondes starting one of the longest sets of upper atmosphere data ever taken

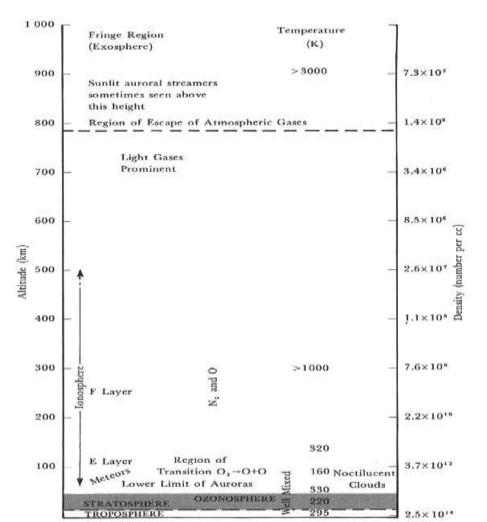






Pre-1940 view of the upper atmosphere

 Remarkable for how well some of its characteristics had been worked out from sparse and indirect evidence

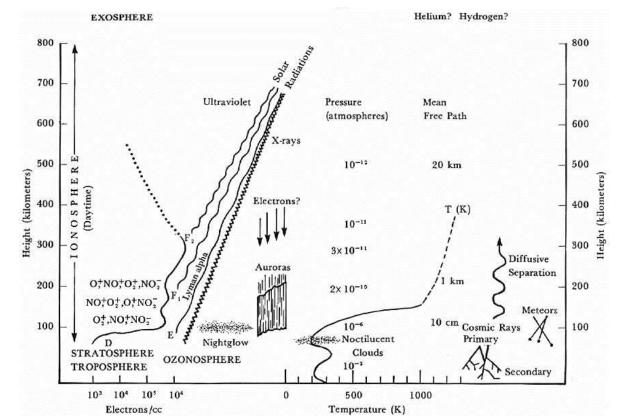




The age of the rocket

With the space age came pioneers such as UCL's own Sir Harrie Massey, with Boyd and Bates, who pushed forward the theory of the chemistry, composition and behaviour of the upper atmosphere. It is their work which underpins what we know of the atmosphere today







Neutral Atmosphere

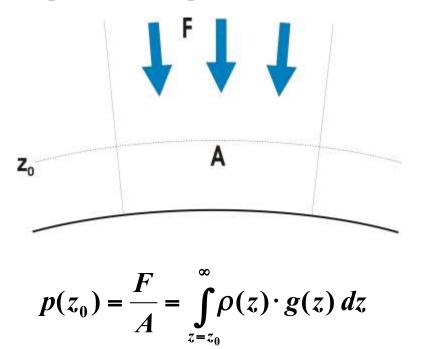
1) Introduction: Temperature and Composition

a) The Basics: Pressure, Scale height

An atmosphere is a gas layer surrounding a planet. Our solar system has a range of planets and moons with recognisable, permanent atmospheres - Earth, Mars, Venus, Giant planets, Titan, Triton. This talk is relevant to these, but not really the transient, thin atmospheres of bodies like Mercury, Io and Enceladus



The **pressure** is defined as the force (per surface area) due to the weight of the atmosphere above a point:



and the change or pressure with height is given by:

$$dp = -\rho \cdot g \, dz$$

Pressure decreases with altitude since the weight of the atmosphere above becomes smaller for increasing altitude.



From the ideal gas law:

$$p = n k T$$

and:

$$\frac{dp}{dz} = -\rho \cdot g = -\overline{m} \cdot n \cdot g \qquad \overline{m} \text{ is the mean molecular mass (in units of mass)}$$
we get:

$$\frac{dp}{dz} = -\frac{\overline{m} \cdot g}{kT} \cdot p = -\frac{1}{H} \cdot p$$
where:

$$H = \frac{kT}{\overline{m}g} \qquad \text{is the scale height}$$
giving:

$$p(z) = p(z_0) \cdot e^{-\frac{z-z_0}{H(z)}}$$

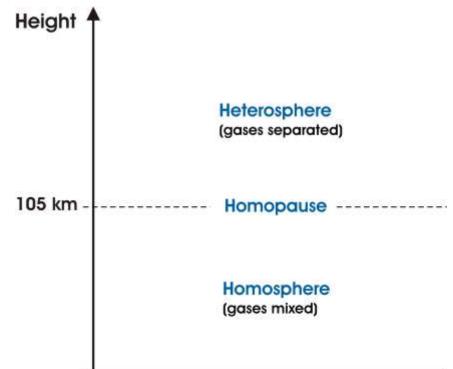
So, when moving up by one scale height, pressure decreases by a factor of $1/e \sim 0.37$.



The same applies to number densities:

$$n = \frac{p}{kT}$$
, so: $n = n_0 \cdot \left(\frac{T_0}{T}\right) \cdot e^{-\int_{z_0}^{z} \frac{dz}{H}}$

In practise, the atmosphere consists of many gases which below the homopause are well mixed due to small scale turbulence and large scale winds:

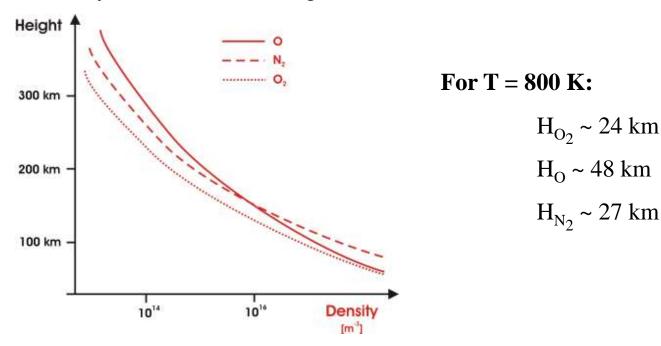




In the heterosphere, molecular diffusion is more effective than turbulent mixing, so gases behave independently. They distribute vertically according to their individual scale heights,

$$H_i = \frac{kT}{m_i g}$$

So, the lighter a gas, the larger its scale height and the slower its densities fall with altitude. Therefore, lighter gases become relatively more abundant at higher altitudes.





Thermosphere Composition

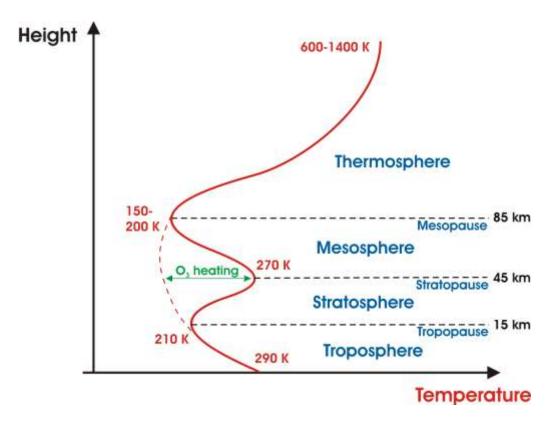
(MSISE 90 Model, March, lat=0, lon=0, 15:00 LT, F10.7=100, Ap=6, 1987)

Height	O [m ⁻³]	$O_2 [m^{-3}]$	N ₂ [m ⁻³]
100 km	4.55 ¹⁰ ¹⁷	2.38 ¹⁰ ¹⁸	1.02[•]10¹⁹
	(3.5 %)	(18.3 %)	(78.2 %)
200 km	3.71[•]10¹⁵	1.64[•]10¹⁴	2.57 ^{10¹⁵}
	(57.6 %)	(2.5 %)	(39.9 %)
300 km	4.39 ¹⁰ ¹⁴	2.44 ¹⁰ ¹²	6.42 ¹⁰ ¹³
	(86.8 %)	(0.5 %)	(12.7 %)



b) Thermal structure

The Earth's atmosphere may be subdivided vertically into different regions:

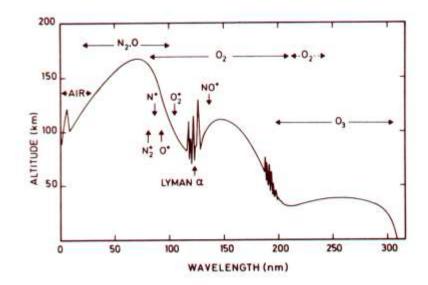




This structure is a result of different gases absorbing different wavelengths at different altitudes.

Wavelength	Atmospheric absorbers Solar Lyman α line, absorbed by O ₂ in the meso- sphere; no absorption by O ₃	
121.6 nm		
100 to 175 nm	O_2 Schumann Runge continuum. Absorption by O_2 in the thermosphere. Can be neglected in the mesosphere and stratosphere.	
175 to 200 nm	O_2 Schumann Runge bands. Absorption by O_2 in the mesosphere and upper stratosphere. Effect of O_3 can be neglected in the mesosphere, but is important in the stratosphere.	
200 to 242 nm	O_2 Herzberg continuum. Absorption by O_2 in the stratosphere and weak absorption in the meso-sphere. Absorption by the O_3 Hartley band is also important; both must be considered.	
242 to 310 nm	O_3 Hartley band. Absorption by O_3 in the strato- sphere leading to the formation of $O(^1D)$.	
310 to 400 nm	O_3 Huggins bands. Absorption by O_3 in the stratosphere and troposphere leads to the formation of $O(^3P)$.	
400 to 850 nm	$\rm O_3$ Chappuis bands. Absorption by $\rm O_3$ in the troposphere induces photodissociation even at the surface.	

Spectral regions of photochemical importance in the atmosphere





Troposphere:

- Energy sources:
 - planetary surface absorption (IR, visible), conduction to atmosphere
 - atmospheric absorption of terrestrial and solar IR
 - latent heat release by H_2O
- Energy sinks:
 - IR radiation
 - evaporation of H_2O

Stratosphere:

- Energy sources:
 - strong absorption of UV by ozone (causing stratopause temperature peak)
- Energy sinks:
 - IR radiation by O_3 , CO_2 , H_2O

Mesosphere:

- Energy sources:
 - some UV absorption by O_3 (lower heights)
 - heat transport down from thermosphere (minor, top heights only)
- Energy sinks:
 - IR radiation by CO₂ H₂O, OH



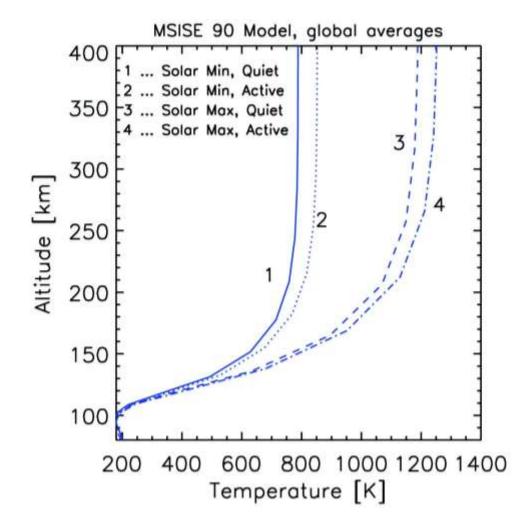
Thermosphere:

- Energy sources:
 - absorption of EUV (200-1000Å; photoionizing O, O_2 , N_2) and UV (1200-2000 Å), photodissociating O_2), leading to chemical reactions and particle collisions, liberating energy
 - dissipation of upward propagating waves (tides, planetary waves, gravity waves)
 - joule heating by auroral electrical currents
 - particle precipitation from the magnetosphere
- Energy sinks:
 - thermal conduction into the mesosphere, where energy is radiated by CO_2 , O_3 and H_2O
 - IR cooling by NO and CO₂ (after geomagnetic storms)

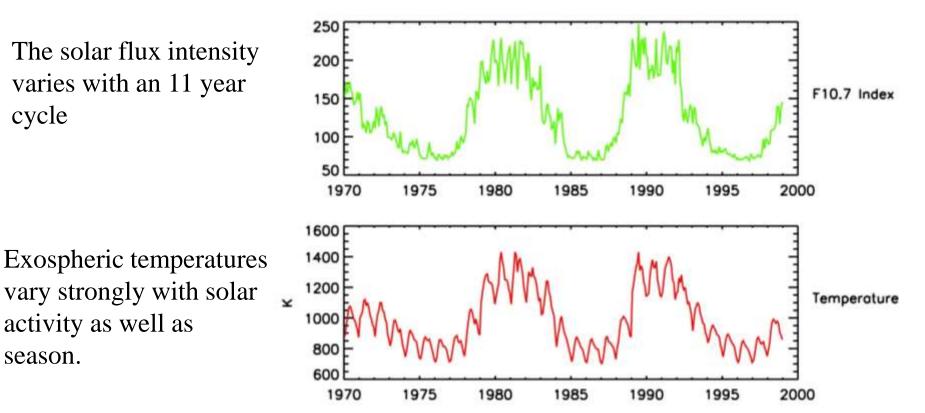


2) The Thermosphere

a) Thermal structure and composition



Solar flux intensity is characterized by the F10.7 index, which gives the flux of solar radiation at 10.7 cm wavelength. Although this wavelength is of no importance to the upper atmosphere, its flux correlates well with UV and EUV fluxes.



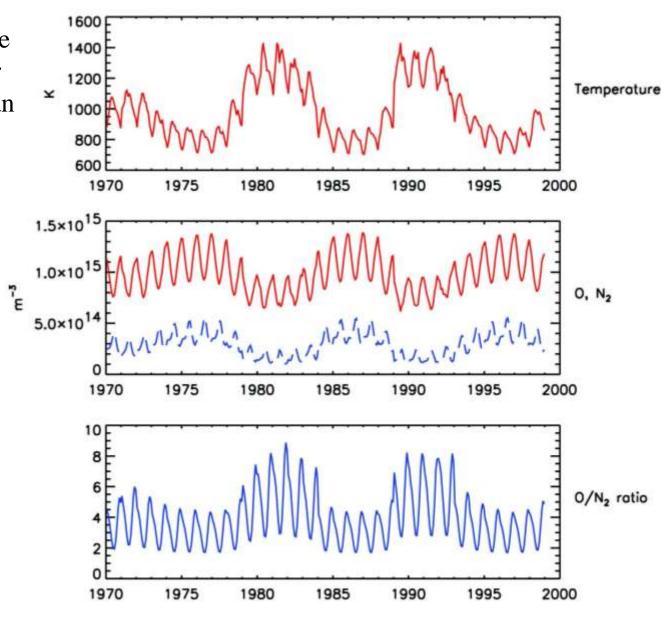


300 km, 50°N, local noon

Seasonal temperature changes are stronger during solar max than solar minimum.

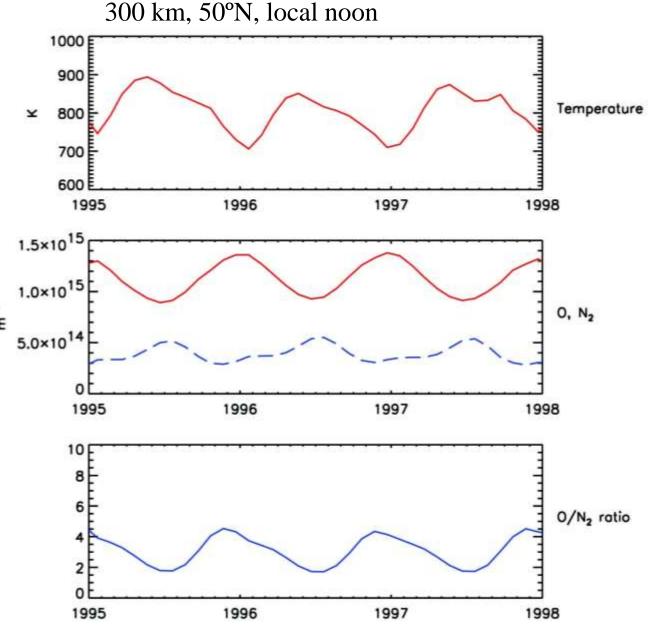
Number densities of gases vary with solar cycle and season

During active solar conditions there is proportionally more O than N_2 in the thermosphere, due to stronger photodissociation.



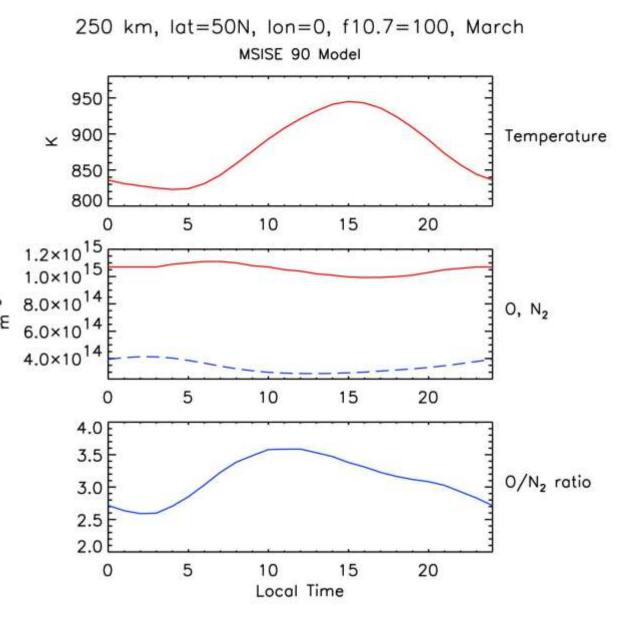
Exospheric noon-time temperatures peak near summer solstice

In spite of increased dissociation at summer solstice, the abundance r_e of O in proportion to N_2 is at its lowest then. This is due to vertical transport of gases: summer upwelling transports N_2 rich gases from lower to higher altitudes.

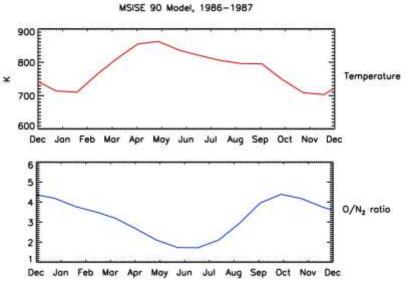


Exospheric noon-time temperatures peak at around **15:30 h** local time.

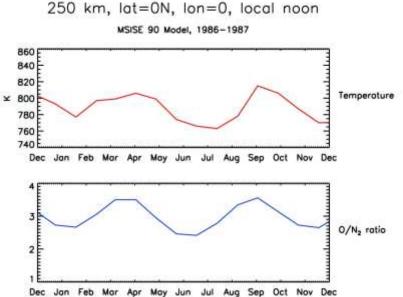
The O/N_2 ratio peaks near noon and is smallest near 02:00 local time. It is thus ${}^{r}_{E}$ largely controlled by the solar zenith angle (through photochemistry), but gas transport also plays some role.



While temperatures and composition show an **annual** change at **mid-latitudes**...



250 km, lat=50N, lon=0, local noon



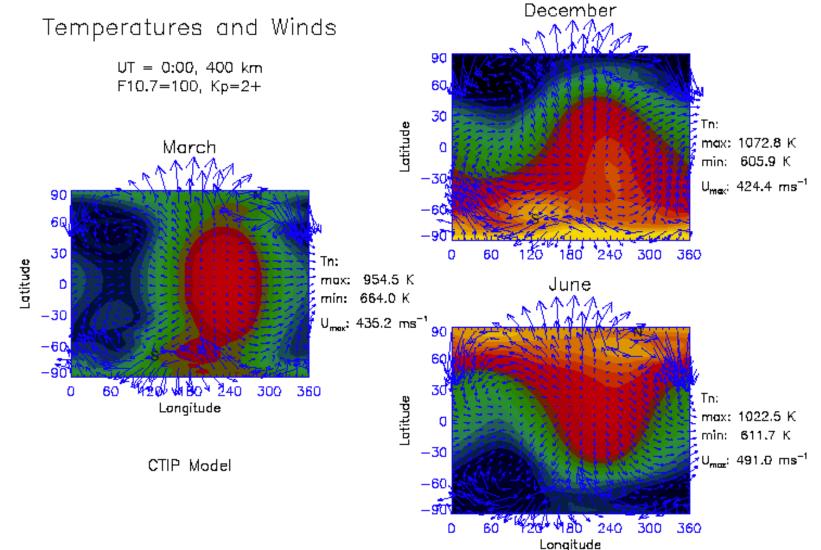
... they change **semiannually** at **low latitudes**. Densities show the same trend. Various theories have been proposed to explain this, such as wave propagation from below and geomagnetic forcing.



Other anomalies:

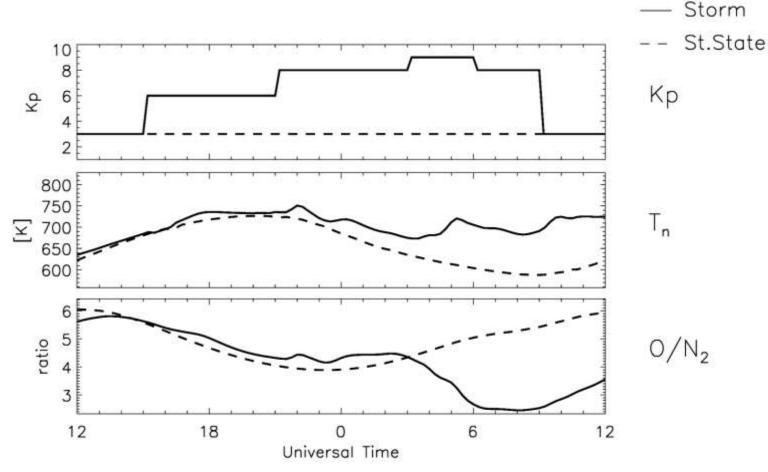
- There are a number of other anomalies in electron density and temperature
- Some of these are still not completely understood
- For example, there is a semi-annual asymmetry where the total electron content is significantly larger globally in December than in June.
- The Earth is nearer the Sun then, but the increase in flux is not enough to explain this effect

So we understand the basic diurnal and seasonal variability, with reservations ...



The diurnal temperature behaviour at mid to high latitudes is also strongly affected by the changes in geomagnetic activity. An enhancement of geomagnetic activity can lead to an increase in temperature by up to 200 K.

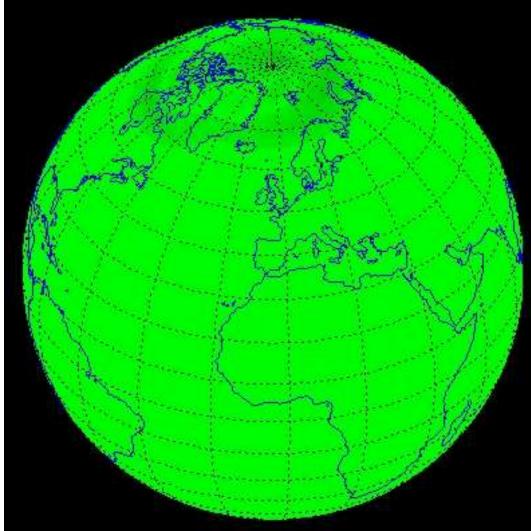
^AUC



Changes of composition during and after magnetic storms are caused by vertical winds.

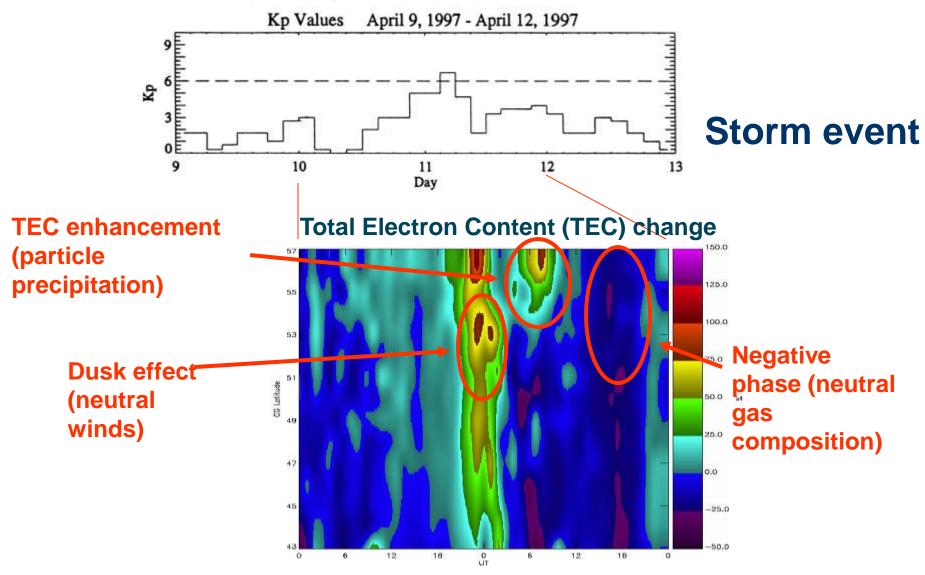


That's the static picture. There are phenomena that cause dynamic effects too, and we have a general idea how these work too





But complications continually arise: the response to storms is not simple:

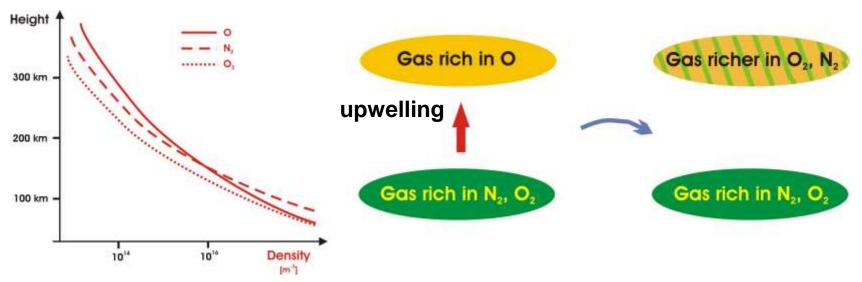




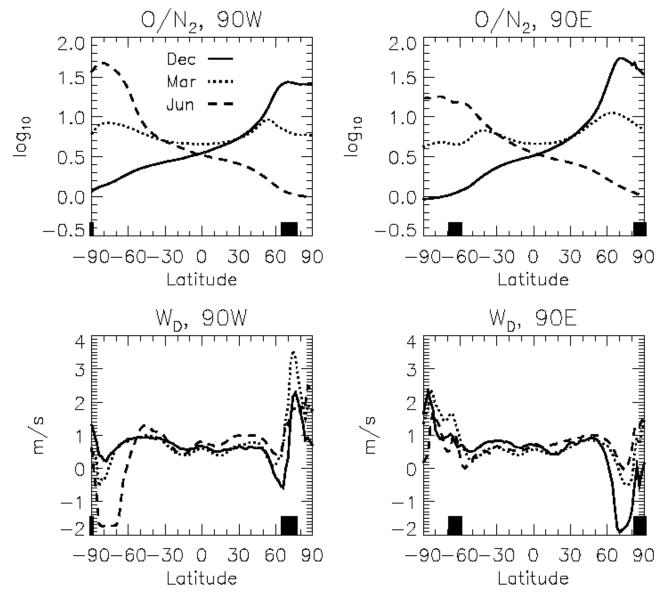
The total vertical wind may be expressed in terms of 2 components, the **barometric** wind and the **divergence** wind. The former is due to thermal expansion/contraction of the atmosphere, the latter is caused by diverging horizontal winds and the conservation of mass:

$$U_{z} = W_{B} + W_{D} = \left(\frac{\partial h}{\partial t}\right)_{p} + \frac{1}{n} \int_{z}^{\infty} \left(\frac{\partial (nU_{x})}{\partial x} + \frac{\partial (nU_{y})}{\partial y}\right) \cdot H dz$$

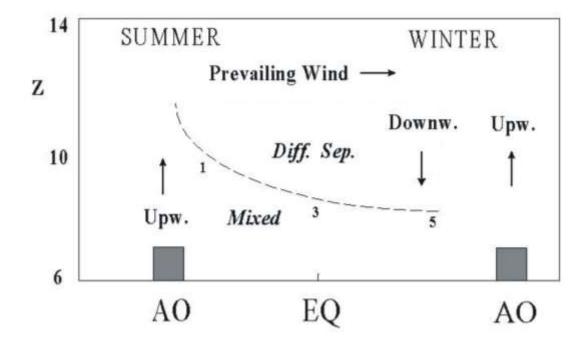
Upward vertical *divergence* winds transport gases from lower to higher altitudes. Gases at lower heights are richer in molecular constituents, so the upward winds cause gases higher up to be relatively more molecular. So, **upward** winds cause a **decrease** in the O/N_2 ratio.



The principle is illustrated in these plots. Note the changes in composition near the auroral ovals (black boxes), where strong vertical winds are found.



Inter-hemispheric winds and composition changes



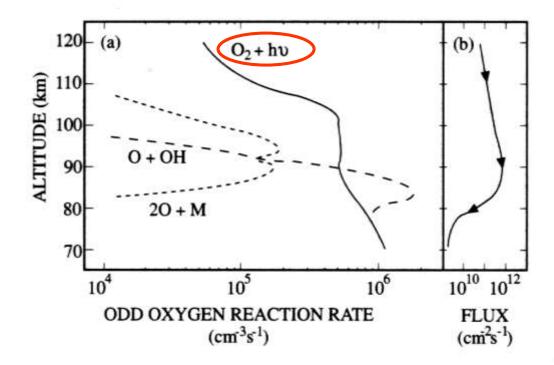
This figure illustrates the interplay between winds and composition in the thermosphere. Upwelling occurs in the summer hemisphere and over the auroral ovals ("AO"). The summer upwelling upsets diffusive equilibrium, gases are transported by horizontal winds towards the winter hemisphere, where diffusive balance is progressively restored, from top (where diffusion is faster) to bottom. An additional circulation cell forms in the winter hemisphere through upwelling over the aurora.



b) Photochemistry in the thermosphere

i) Photodissociation:

 $O_2 + h\nu \rightarrow O(^3D) + O(^1D)$



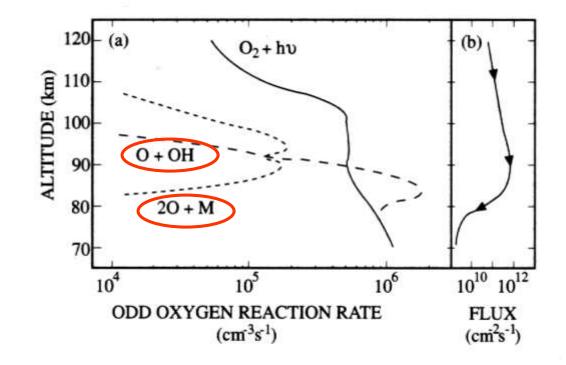


ii) Recombination:

$$O + O + M \rightarrow O_2 + M \qquad \left[\approx 3 \times 10^{-34} \exp\left(\frac{710}{T}\right) cm^6 / \sec\right]$$

Where M is a molecule; O_2 or N_2 . Three-body recombination is necessary to get rid of excess energy of metastable O_2^* . Note the temperature dependence of the reaction rate.

 $0 + OH \rightarrow O_2 + H$





c) Gas mixing and transport

The concentration of gas particles is determined by the balance between sources and sinks.

Gas concentrations are affected by **chemical** sources and sinks, **transport** by winds and **diffusion** of gas particles (their microscopic motion).

The **conservation of mass** is a fundamental law in nature:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \rho \, \vec{v} = 0$$



The concentration of gas constituents may be expressed in 3 ways:

- i) **number density** (particles per volume), *n_i*
- ii) mole fraction (number mixing ratio),

$$X_i = \frac{n_i}{n_{tot}}$$
 where $n_{tot} = \sum_i n_i$

iii) mass fraction (mass mixing ratio),

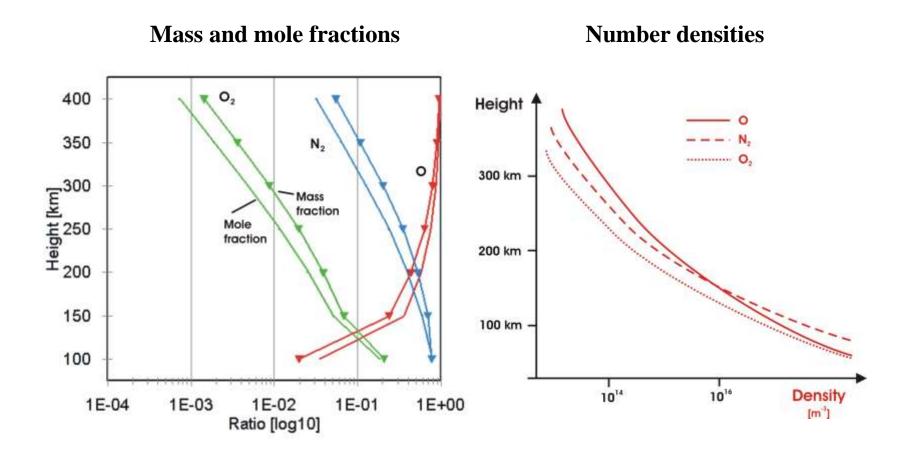
$$Y_i = \frac{n_i m_i}{n_{tot} m_{tot}}$$
 where $m_{tot} = \sum_i \frac{n_i m_i}{n_{tot}}$

Note that:

$$\sum_{i} \mathbf{X}_{i} = 1 \quad \text{and} \quad \sum_{i} \Psi_{i} = 1$$

Therefore, the continuity equation can be expressed in 3 forms.

⁺UCL





In order to understand changes in composition and estimate which processes are dominant, we introduce the concept of **time scale**, which quantifies how long changes in composition take for each type of process.

i) Dynamical time scale:

Estimates time scale of gas transport by winds:

$$\tau_{dyn} \approx \frac{dist}{V}$$

where *dist* is a typical distance and *V* the velocity.

Horizontal:dist = a few degrees lat/lon (depending on problem)Vertical:dist = one scale height



ii) Chemical time scale:

a) Photochemical production/loss:

$$N_i + hv \xrightarrow{J} N^+_i + e$$

J is rate coefficient for photoproduction or loss (in [1/sec])

$$\frac{\partial N_i}{\partial t} = -J \cdot N_i \qquad \Rightarrow \qquad N_i = N_{i0} \cdot e^{-Jt}$$

So, the **photochemical lifetime** is defined as:

$$\tau_p = \frac{1}{J}$$



b) Chemical production/loss:

$$N_i + N_j \xrightarrow{K} \dots$$

K is rate coefficient for chemical production or loss (in [cm³/sec])

$$\frac{\partial N_i}{\partial t} = -K \cdot N_i \cdot N_j \qquad \Rightarrow \qquad N_i = N_{i0} \cdot e^{-KN_j t}$$

So, the **chemical lifetime** is defined as:

$$\tau_p = \frac{1}{KN_j}$$



c) Molecular diffusion time constant:

$$\tau_D = \frac{H^2}{D_{i,j}}$$

 $D_{i,j}$ is the molecular diffusion coefficient between species *i* and *j* (in [m²/sec]); *H* is the average scale height.

 τ_D decreases roughly exponentially with height.

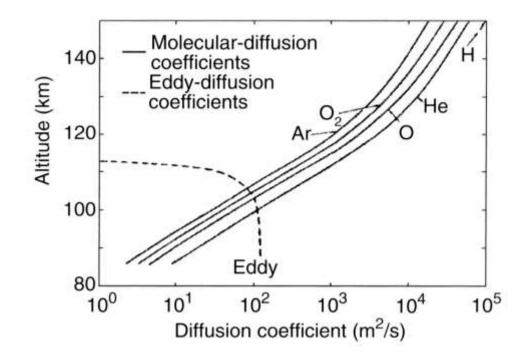
d) Turbulent mixing time constant:

$$\tau_{K} = \frac{H^{2}}{K_{D}}$$

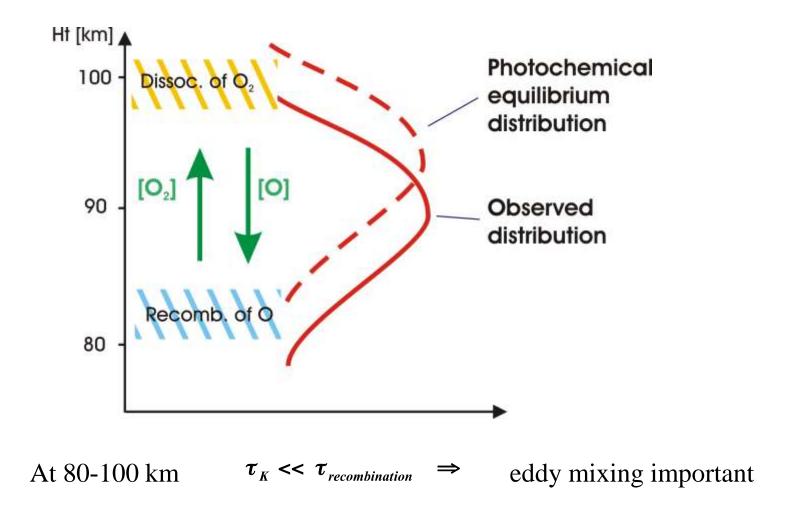
 K_D is the turbulent (Eddy) diffusion coefficient (in $[m^2/sec]$)

 τ_{K} is roughly invariant with height.

[±]UCL



The impact of winds on the vertical distribution of [O]



Mixing transports O down to lower heights, where recombination is more rapid



Summary of basic Thermosphere characteristics:

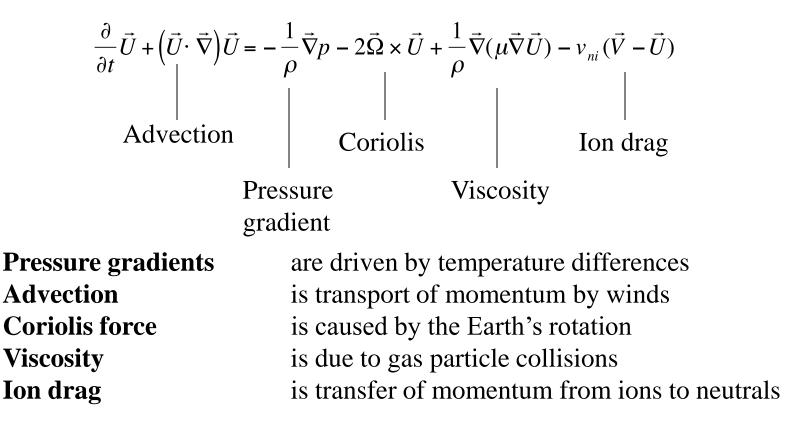
- Main gases: O, O₂, N₂, He (high altitudes only)
- Controlled strongly by solar heating
- At low latitudes effects of upward propagating tides, planetary waves and gravity waves are important
- At high latitudes, heating from the magnetosphere occurs in the form of Joule heating and precipitating particles
- Main cooling is through molecular conduction
- In regions of strong vertical winds adiabatic heating and cooling occurs
- Chemical heating less important
- Strong variability of temperature with solar cycle, season, local time

• ...



d) Thermosphere dynamics

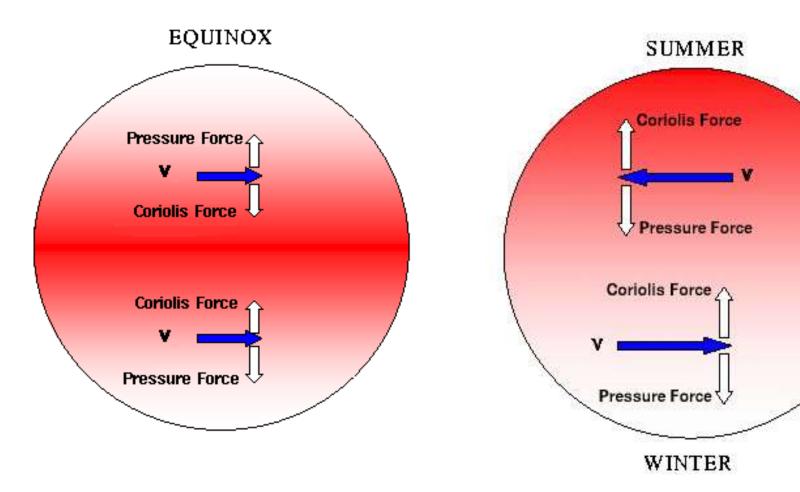
i) Momentum equation for neutral gas



Pressure gradients and Ion drag are external forces, the rest are internal



Origin of Mesospheric Zonal jets

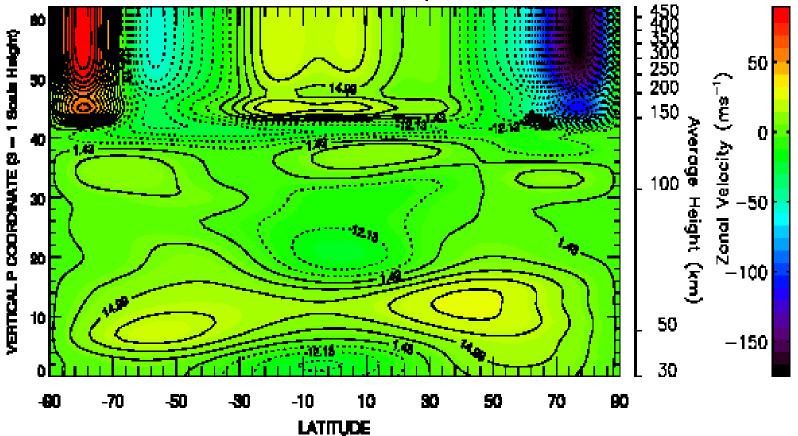


Ref: Harris 2001

UCL

Mesospheric Zonal Jet closure: Zonally averaged zonal wind from HWM

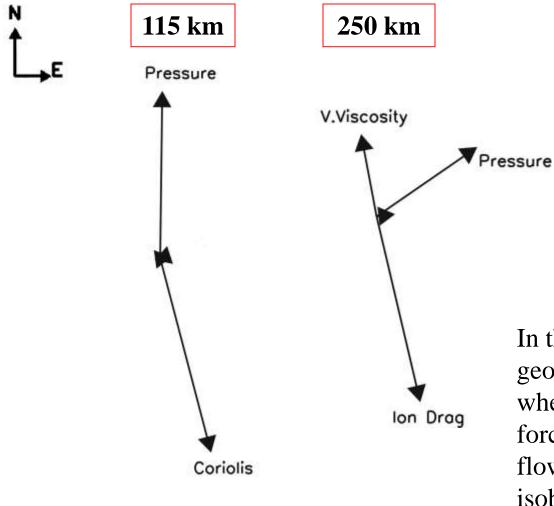
HWM EQUINOX F10.7 = 180 Kp=2+





Momentum balance

March, 30N, 54E, 15:40 LT



Note the differences in momentum balance at different altitudes!

In the lower thermosphere, geostrophic balance is found, where Pressure and Coriolis forces almost balance. Winds flow roughly perpendicular to isobars.

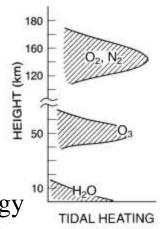


Coupling from below: Tides, Planetary and Gravity waves

Winds in the lower thermosphere (80-200 km) are controlled to large part by global waves propagating upward from the lower atmosphere. These are:

a) Tides:

- Periods: 24*h*, 12*h*, 8*h*
- Driven by solar heating (*thermal tides*) and, to less extent, by the Moon's gravitation field (*lunar tides*)
- Follow the Sun, ie. westward propagating
- Distinct latitudinal structure described, for an idealized atmosphere, by *Hough modes*.
- Peak amplitudes in lower thermosphere \sim 50 m/s and \sim 10 K
- Dissipate in 100-160 km height regime, depositing large amounts of momentum (westward zonal acceleration) and energy (temperature increase) into the background atmosphere.
- Show diurnal and seasonal variability



UCL

Classical Tidal Theory

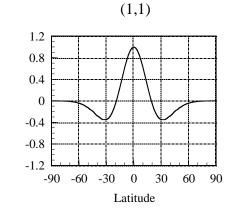
Classical tidal theory analytically describes the global structure of tidal oscillations. The analytical solutions strictly apply only to an idealized atmosphere, ignoring non-linear processes, but are an effective tool for describing also the real atmosphere.

Global oscillations can be decomposed into **Hough modes** (similar to Fourier decomposition), which allow a physical interpretation of the tidal structure. The modes are classified as

- **migrating**: the tides propagate vertically and horizontally
- **non-migrating**: they remain constrained to where they are generated

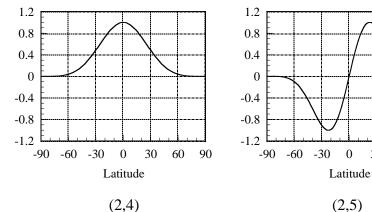
Hough modes are named as (n, m), where *n* is their longitudinal wavenumber and *m* the latitudinal wavenumber.

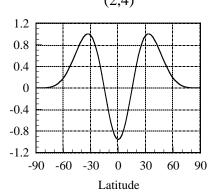


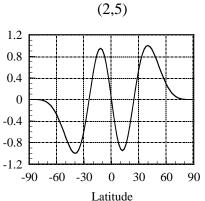












0

30 60 90

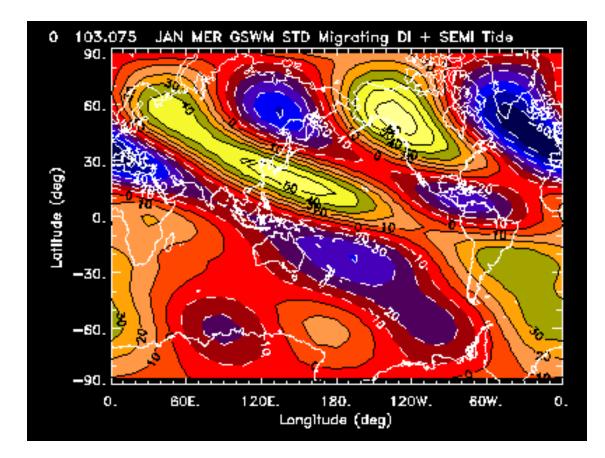
Typical migrating tidal modes are:

- (1,1): diurnal, symmetric
- (2,2): semidiurnal, 1st symmetric
- (2,3): semidiurnal, 1st antisymmetric
- (2,4): semidiurnal, 2nd symmetric
- (2,3): semidiurnal, 2nd antisymmetric

These figures show common Hough modes. The lines show the amplitudes of geopotential height of a surface of constant pressure. Classical tidal theory gives analytical equations for deriving wind and temperature amplitudes from those of geopotential height.

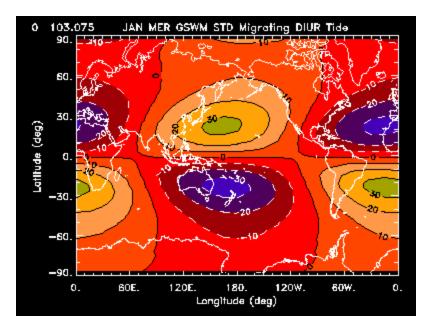


This figure shows a global profile of meridional winds versus latitude and local time (longitude) near 100 km altitude.

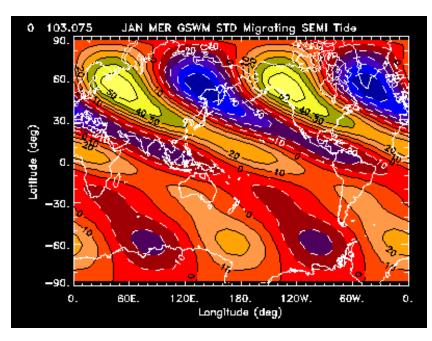


This global profile can be decomposed into diurnal and semidiurnal migrating tides:



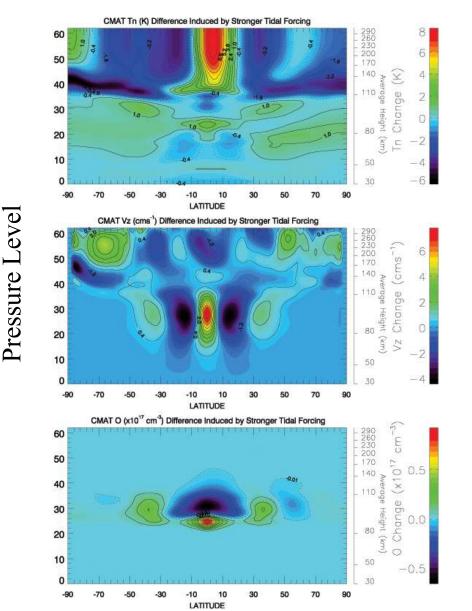


Semidiurnal tide in meridional winds near 100 km. Amplitudes reach 50 m/s near 60° latitude, illustrating that semidiurnal migrating tides reach to higher latitudes than diurnal ones. At this altitude, diurnal and semidiurnal amplitudes are comparable. Diurnal tide in meridional winds near 100 km for January conditions, as simulated by the GSWM model. Amplitudes reach 30 m/s near 30° latitude, this migrating mode (1,1) is confined to the equatorial region.





Coupling of dynamics and chemistry: Effect of increasing tidal forcing



Temp Change (30-300km)

Vz Change (30-300km)

[O] Change (30-300km)



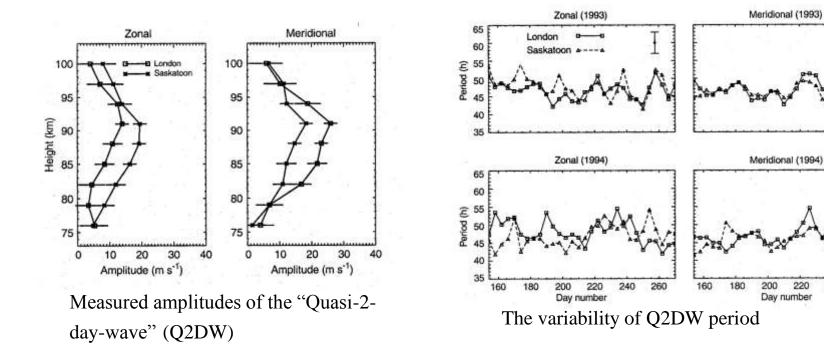
220

240

260

b) Planetary waves:

- Periods: > 1 day, typically 2 days, 5 days, 16 days, ...
- Periods are not exact, they vary with time
- Probably Resonant oscillations of the atmosphere; no clear source
- Don't propagate horizontally; phases are locked in longitude
- Strongly sensitive to the background atmosphere, particularly the zonal winds when those are comparable to the wave's phase speed
- Dissipate typically below 100 km altitude in the neutral atmosphere





c) Gravity waves:

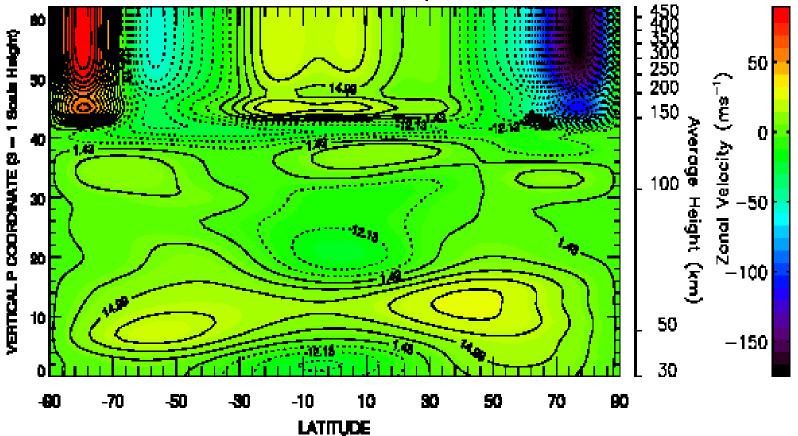
Gravity waves are short period (minutes-hours) waves generated locally by topography, localized heating events, wind shears or other small scale processes. Those gravity waves generated in the lower atmosphere propagate upward and, by dissipating or breaking, considerably affect the structure of mesospheric winds. Their effect on the thermosphere is less significant in comparison.

Horizontally propagating gravity waves are found frequently in the thermosphere, in particular during and after geomagnetic events or solar eclipses. They reach velocities of around 300 m/s.

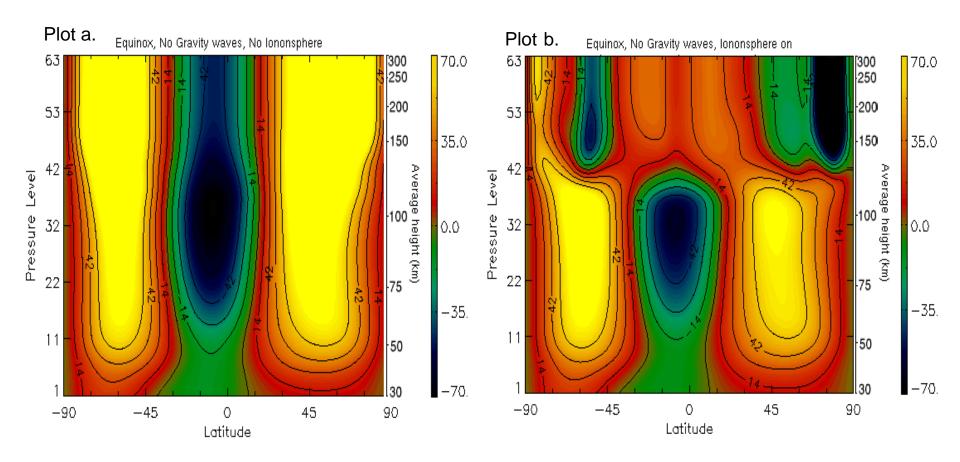
UCL

Mesospheric Zonal Jet closure: Zonally averaged zonal wind from HWM

HWM EQUINOX F10.7 = 180 Kp=2+

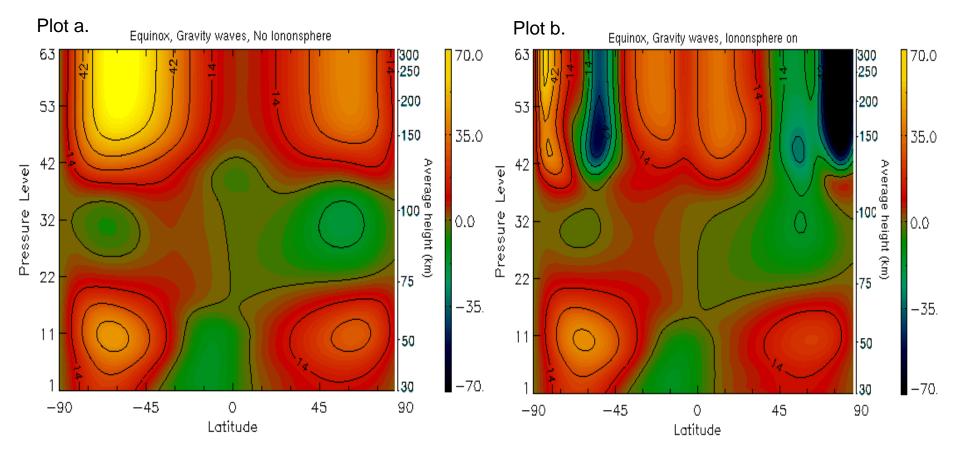


Zonally averaged Zonal wind from CMAT model runs at Equinox. Low solar/geomagnetic activity. Plot a: No gravity waves, No ionosphere Plot b: No gravity waves, Ionosphere





Zonally averaged Zonal wind from CMAT model runs at Equinox. Low solar/geomagnetic activity. Plot a: Gravity wave drag, No ionosphere Plot b: Gravity wave drag, lonosphere





Atmospheric Escape

In the upper regions of the atmosphere, gravity is weak enough for gas particles to escape. They gain their necessary kinetic energy from processes like:

- Thermal escape ("Jeans escape")
- Nonthermal processes (mostly involving ions):
 - Charge exchange
 - Dissociative recombination
 - Impact and photo-dissociation
 - Ion-neutral reactions
 - Sputtering
 - Solar wind pickup
 - Ion escape
 - Electric fields

Light gases in the Earth's upper atmosphere, such as H, D, He are affected by these escape mechanisms.



Jeans escape

Atmospheric particles escape if their velocities exceed the *escape velocity*, which is determined by the gravitational field: $(2CM)^{1/2}$

$$u_{esc} = \left(\frac{2GM}{r}\right)^{r}$$

where G, M and r are the Gravity constant, planet mass and radius. For Earth, u_{esc} =11.2 km/s. Assuming that particle velocities are thermal, the flux of escaping particles is given by:

$$F_{Jeans}(r_c) = \frac{N(r_c) \cdot U}{2\pi^{1/2}} e^{-\lambda} (\lambda + 1) \qquad [cm^{-2} \sec^{-1}] \qquad \begin{array}{c} r_c & \dots & \text{Exobase radius} \\ N & \dots & \text{Number density} \end{array}$$

where

$$U = \left(\frac{2kT}{M_{particle}}\right)^{1/2} \text{ and } \lambda = \frac{u_{esc}^2}{U^2} \qquad \begin{array}{c} M_{particle} \dots \text{Particle mass} \\ T \dots \dots \text{Temperature} \\ k \dots \dots \text{Botzmann const.} \end{array}$$

U is the most probable velocity of a Maxwellian distribution of thermal velocity.

1/3



So, Jeans escape increases with temperature. A rise in temperature should thus lead to a considerable decrease of N(H) at the exobase.

This decrease in N(H) with temperature is observed, but smaller than expected from Jeans escape. Thus, H escape may be due to other processes less sensitive to temperature changes, such as charge exchange:

 $H + {H^+}^* \rightarrow H^+ + H^*$