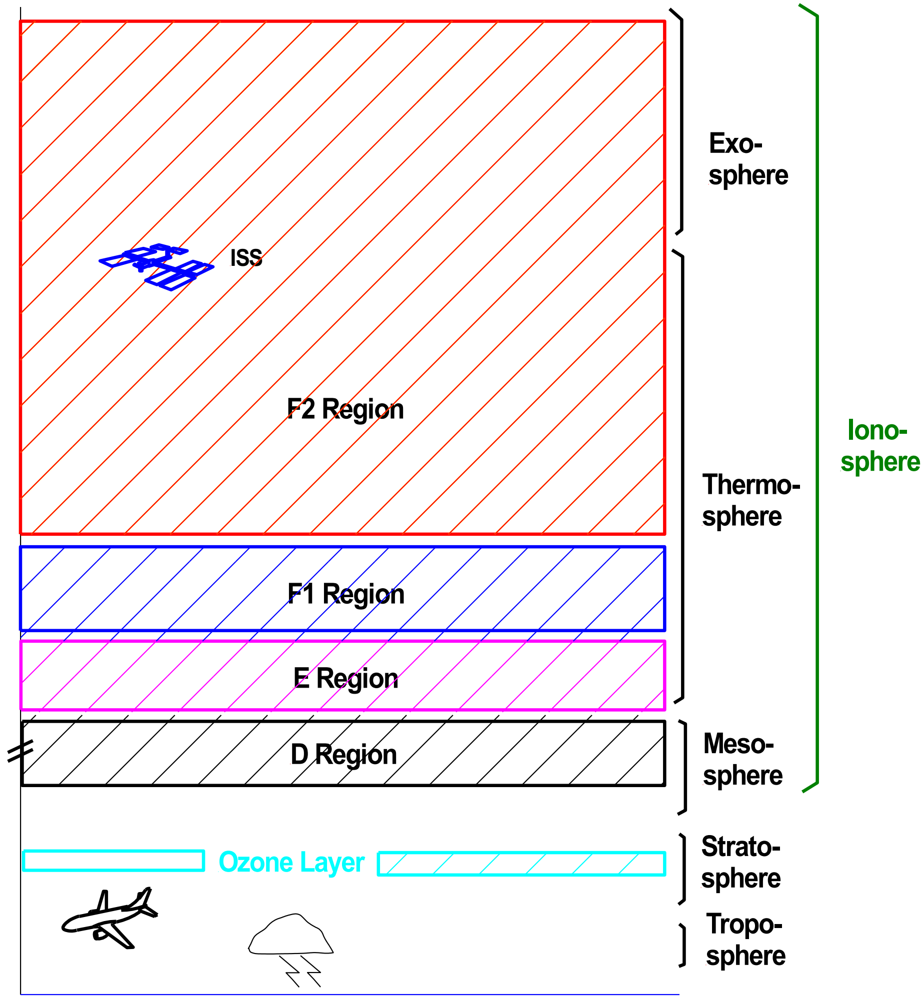
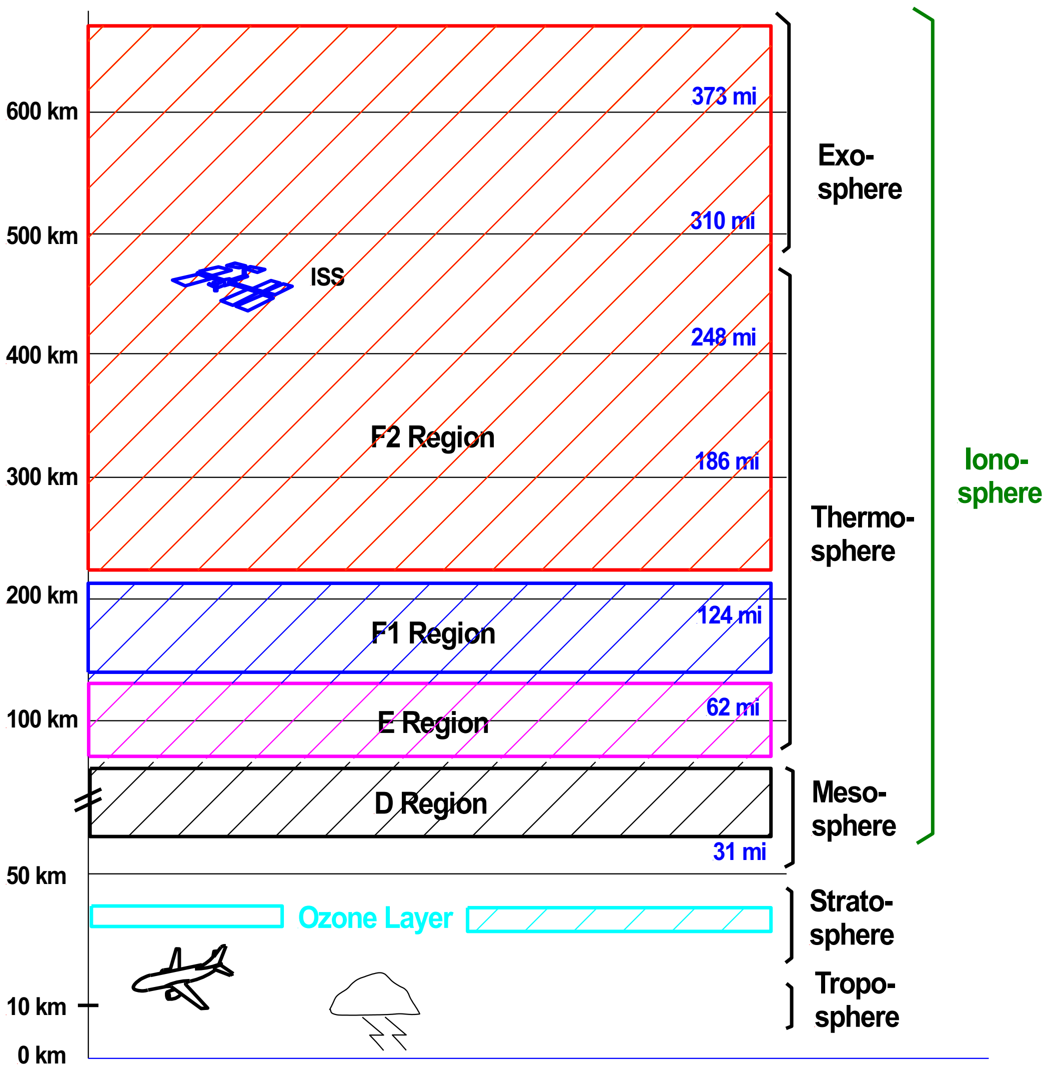
Earth's Ionosphere



Ken Larson KJ6RZ

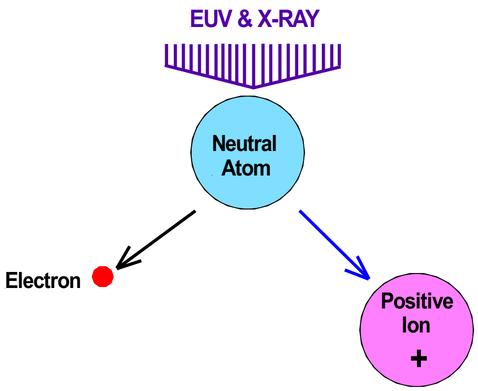
September 2020
www.skywave-radio.org

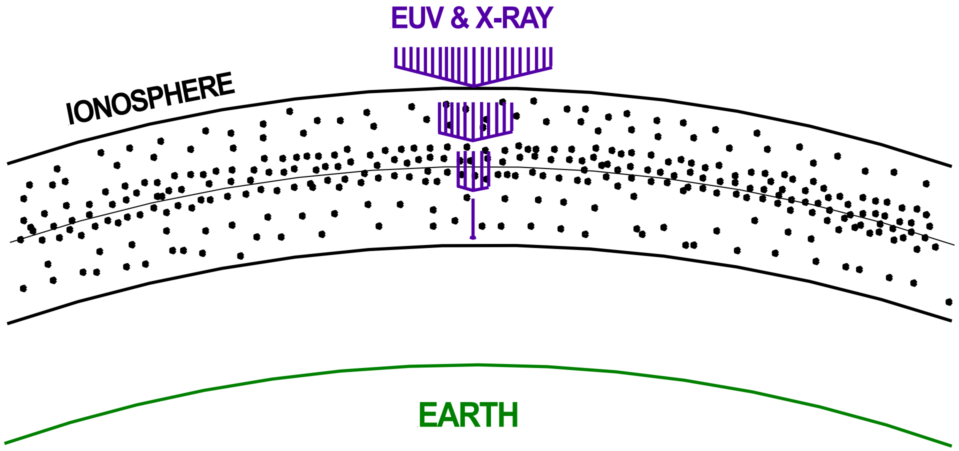
The Ionosphere



- The ionosphere is the ionized region in Earth's upper atmosphere.
- The ionosphere makes possible long distance over the horizon HF radio communications.
- More importantly, the ionosphere protects the Earth from deadly X-ray and extreme ultra violate radiation (EUV) from the Sun.
- Without the ionosphere life on Earth, as we know it, would not exist.

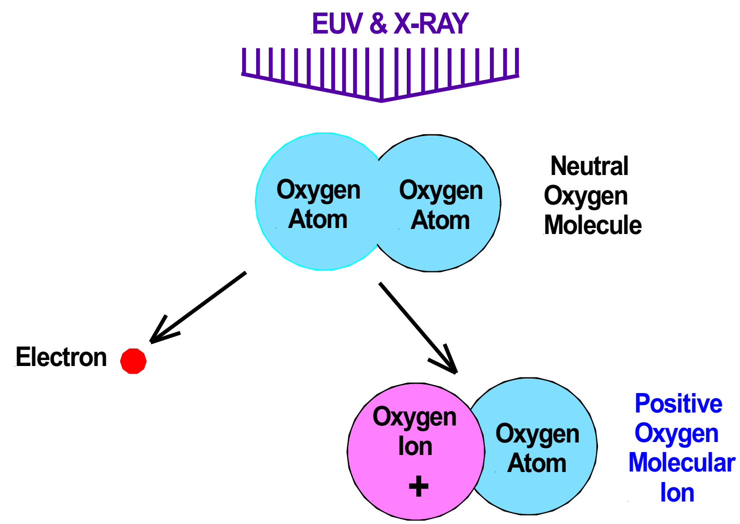
Photo-ionization





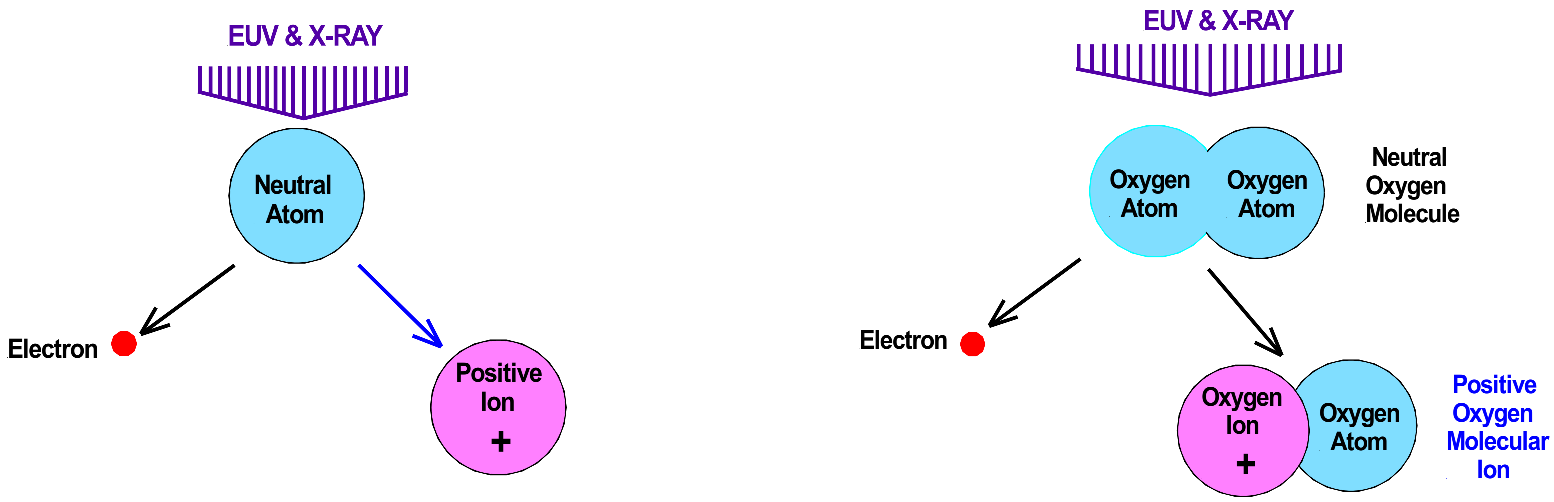
- The ionosphere is formed by a process called photo-ionization.
- During the day, Extreme Ultra Violate (EUV) and X-ray radiation from the Sun is absorbed by neutral atoms in the upper atmosphere.
- A neutral atom has no net electrical charge. It has the same number of electrons in its various electron shells as positive protons in its nucleus.
- **Occasionally** energy absorbed by an atom is sufficient to cause one of its electrons to break away the atom forming a free electron and a positive ion (an atom that has lost one of its electrons).
- There are 1000 times more neutral atoms in the upper atmosphere than ions.
- The ionosphere is thus very thin, wispy, and easily blown around.

Molecular Photo-ionization



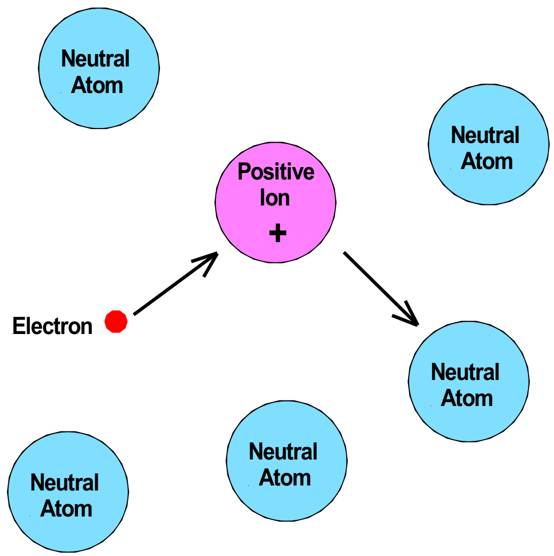
- Photo-ionization also occurs at the molecular level, particularly in the lower part of the ionosphere below approximately 130 km. In the above example an oxygen molecule O_2 is photo-ionized.
- During the day, neutral molecules in the atmosphere also absorb EUV and X-ray energy radiated from the Sun.
- Occasionally energy absorbed by a molecule is sufficient to cause an electron in one of its atoms to break away from the atom forming a free electron and a positive molecular ion

Ions Are Massive Compared to Electrons



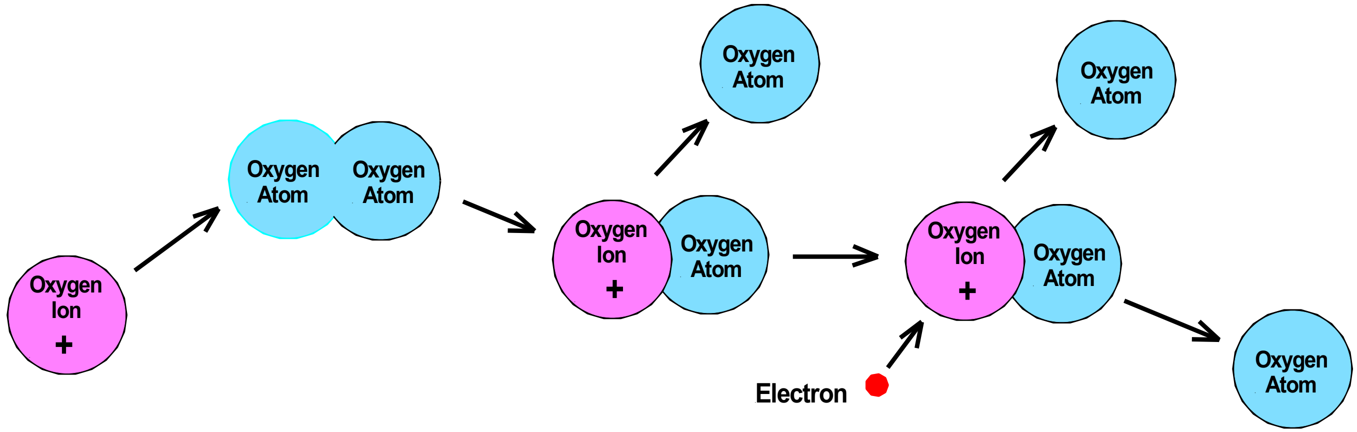
- Ions are 20,000 times more massive than electrons.
- Ions are unaffected by radio waves passing through the ionosphere because they are so massive
- However, tiny electrons interact readily with radio waves – that is why electron densities (the number of electrons per cubic centimeter) is so important in radio communications.

Monatomic Recombination



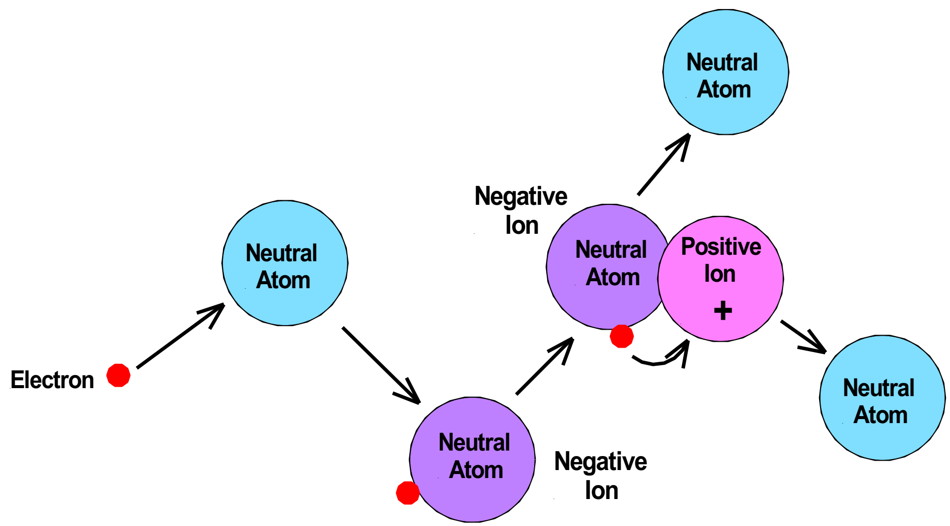
- The ionosphere looses free electrons through the process of recombination.
- In its simplest case a positive ion captures a free electron converting the ion back into a neutral atom.
- Recombination is occurring all the time, both day and night.
- Ionization occurs only during the day. Consequently, electron concentrations are greater during the day than at night.

Molecular Recombination



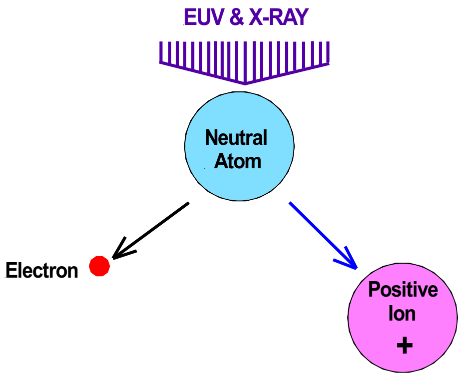
- Like ionization, recombination also occurs at the molecular level through a process know as Dissociative Recombination. This occurs particularly in lower parts of the ionosphere.
- In the example above, an oxygen ion collides with a neutral oxygen molecule replacing one of the molecule’s neutral atoms.
- The molecule becomes a positive molecular ion.
- Eventually a free electron is captured by the molecular ion causing the molecule to split (dissociate) into two neutral oxygen atoms.
- The Dissociative Recombination rate is about 100,000 greater than monatomic recombination.

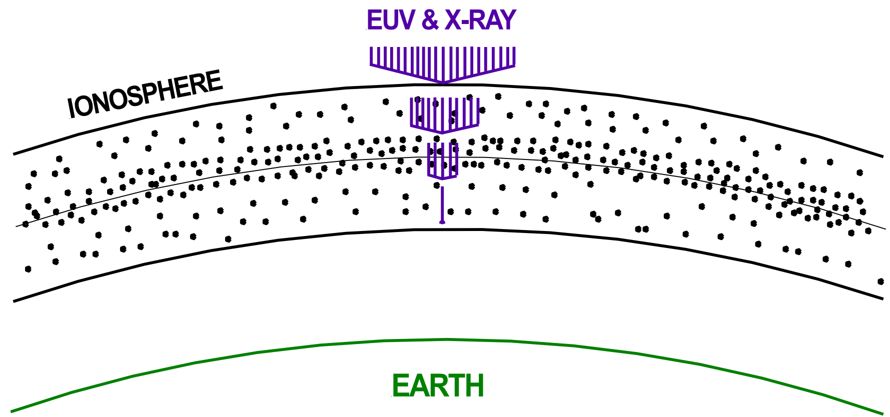
Electron Attachment



- In the low, denser part of the ionosphere, below approximately 90 km, free electrons are also lost through a process called attachment.
- In this process an electron attaches itself to a neutral atom creating a negative ion.
- The negative ion eventually collides with a positive ion.
- In the process, the attached electron transfers to the positive ion producing two neutral atoms.

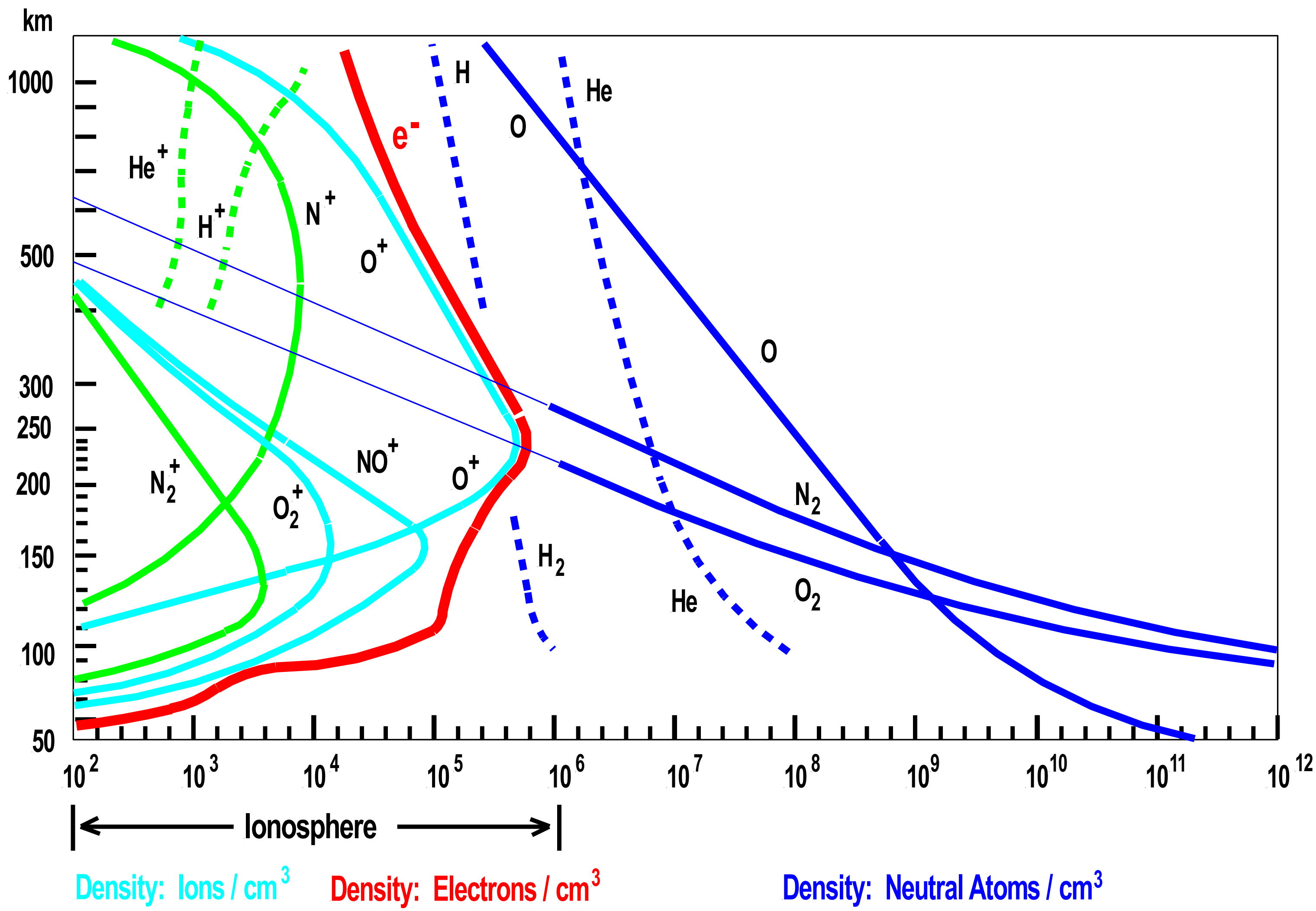
Formation of the Ionosphere





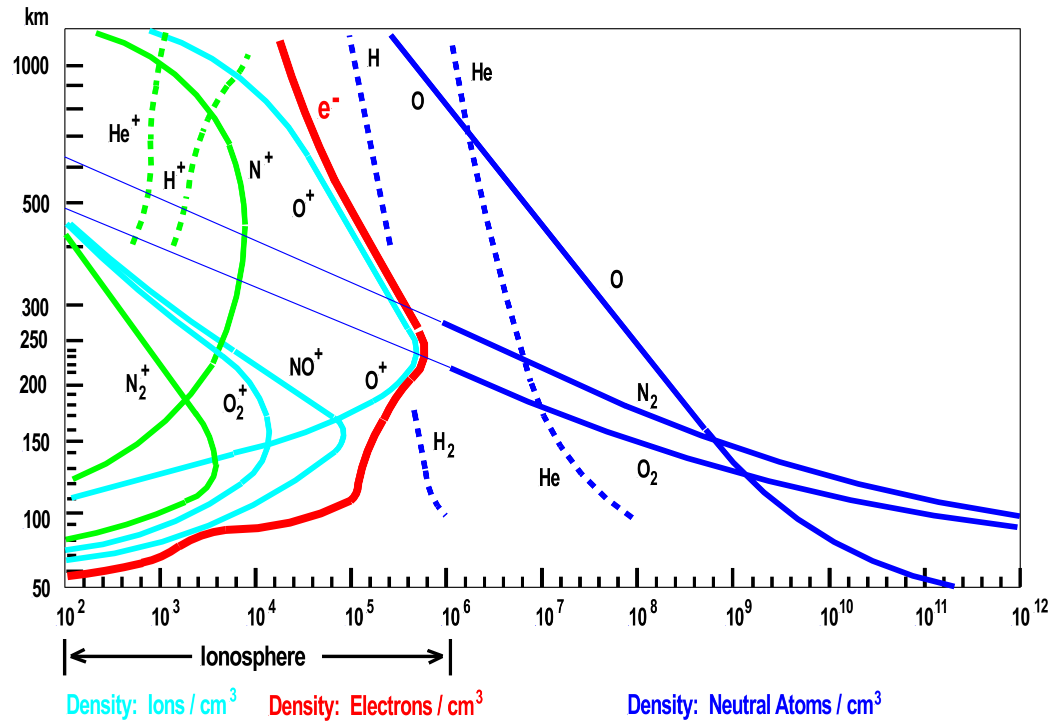
- Radiation is intense at the top of atmosphere but few atoms or molecules to ionize.
- As the radiation penetrates deeper into the atmosphere, it encounters more and more atoms and molecules resulting in higher levels of ionization.
- The ionization process absorbs energy from the EUV & X-ray radiation causing it to weaken in intensity
  - Number of atoms & molecules increases as radiation penetrates further into the atmosphere
  - But levels of ionization decrease due to weakening radiation
  - Ionizations levels drop and eventually disappear.
- Consequently, the highest levels of ionization occur toward the middle of the ionosphere as illustrated above.

Composition of the Atmosphere & Ionosphere



source: angeo.copernicus.org

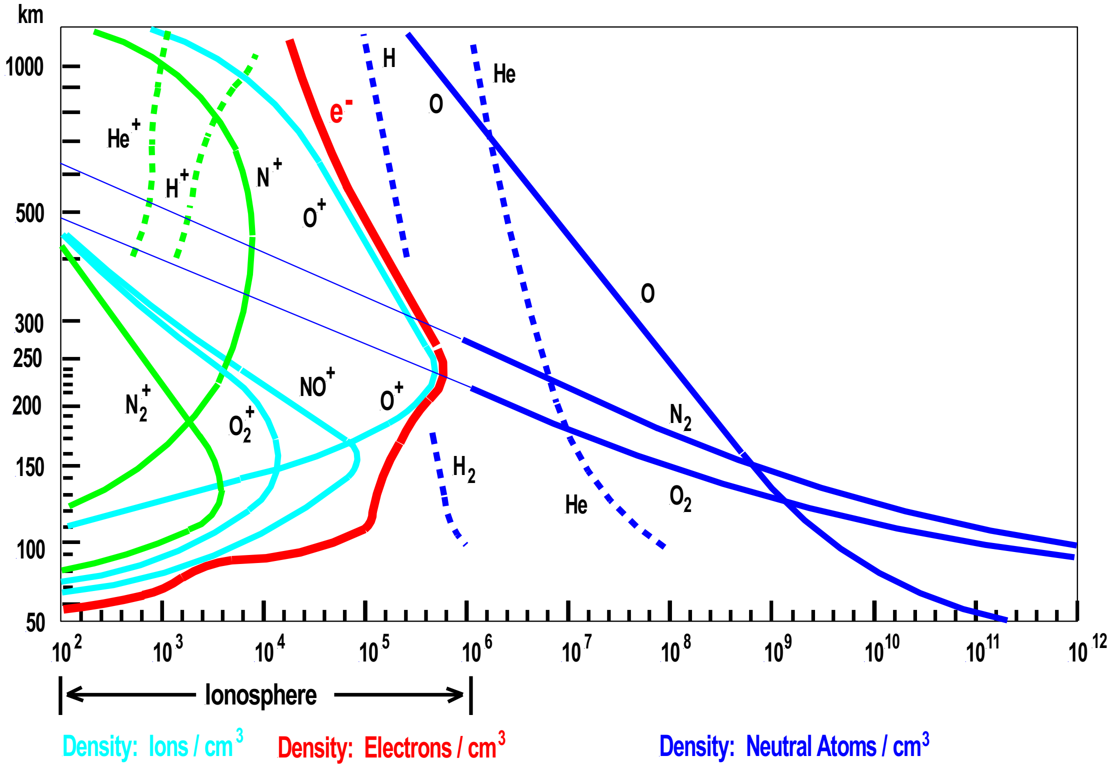
Composition of the Atmosphere – Important observations



- The ionosphere and the neutral atoms of the upper atmosphere occupy the same space, however, the density of the **neutral atoms** is several orders of magnitude greater than the corresponding **ions**. The thin wispy nature of the ionosphere is clearly apparent.

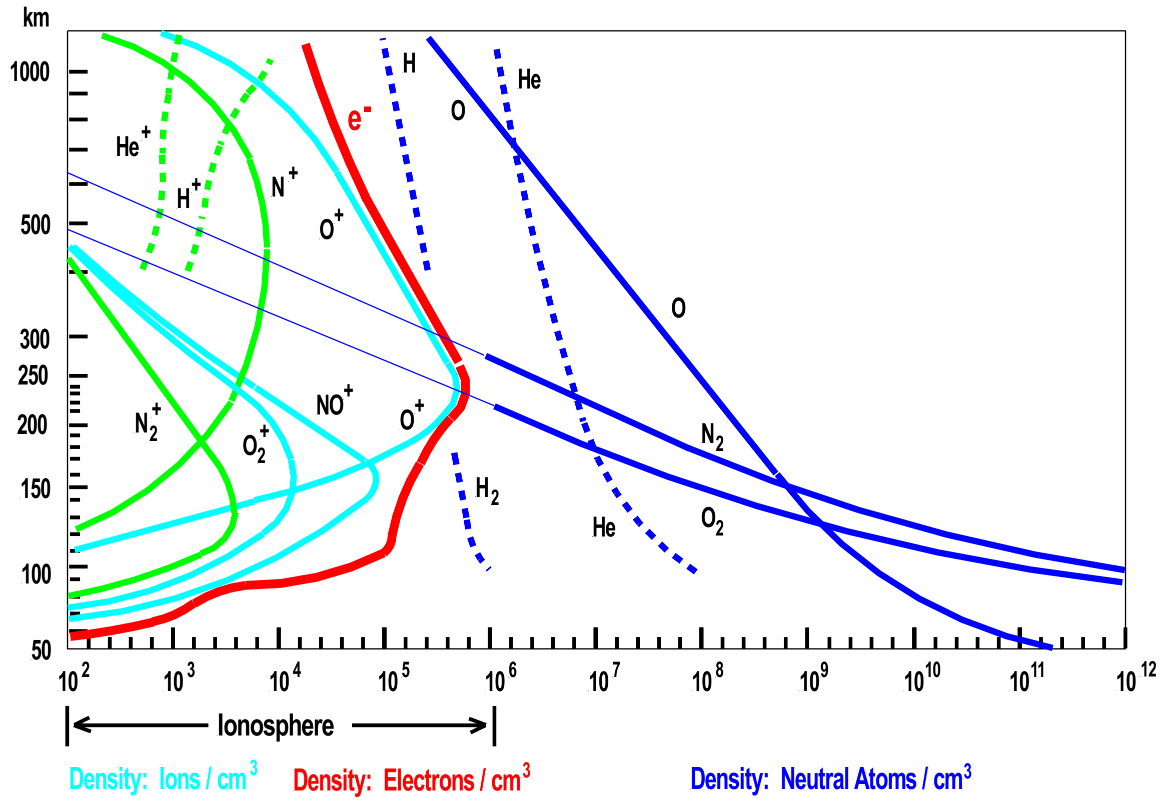
- Most of the atmosphere below 100 km is composed of molecular N_2 and O_2 along with significant quantities of water vapor, carbon dioxide, and nitric oxide.

Composition of the Atmosphere – Important observations



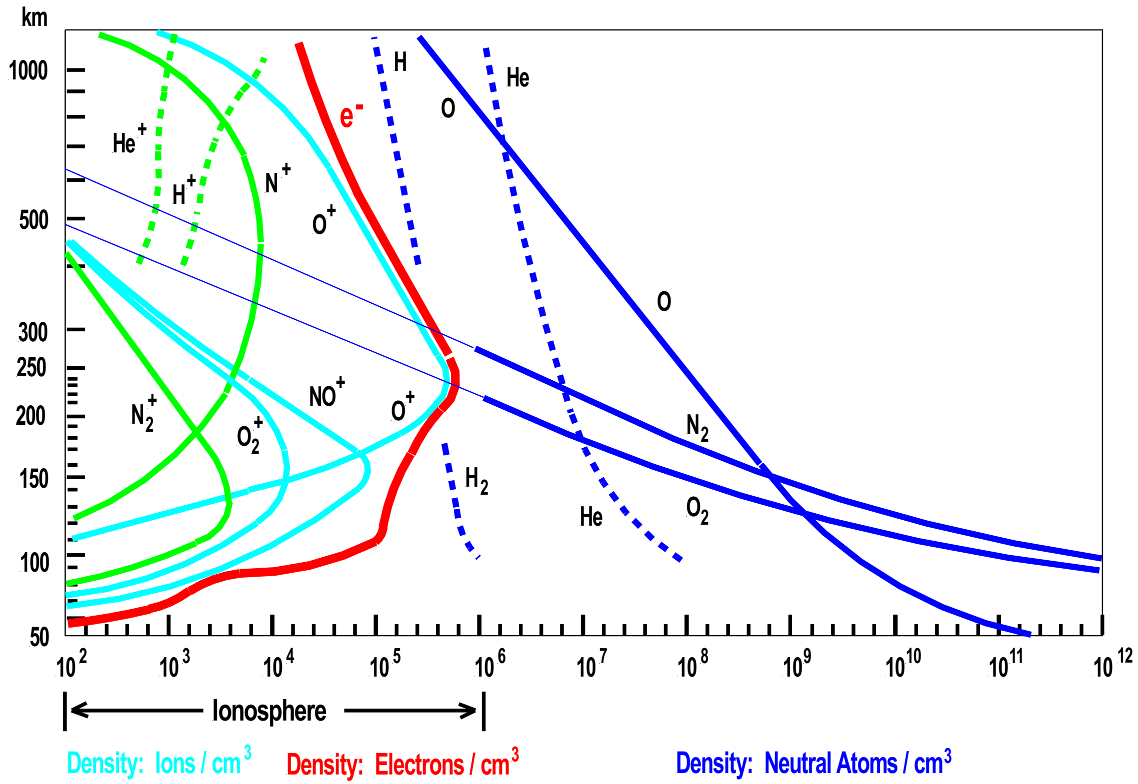
- Polyatomic gas molecules (N_2 , O_2 , H_20, CO_2, NO, etc.) dominate the atmosphere up to an altitude of about 200 km.
- Above 200 km, molecular gases are dissociated into monatomic forms by EUV solar radiation.

Composition of the Atmosphere – Important observations



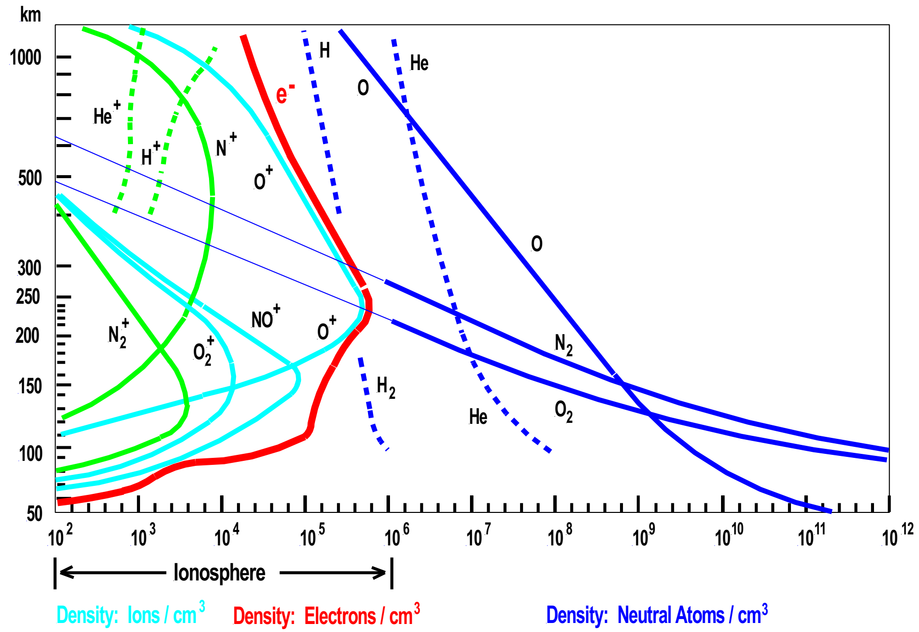
- Consequently, the highest concentration of ions from 50 to about 200 km are molecular ions N_2^+ , O_2^+ , and NO^+.
- Ions in this region are produced primarily by the molecular photo-ionization process.

Composition of the Atmosphere – Important observations



- Recombination also occurs at the molecular level through molecular and electron attachment recombination.
- Recombination occurs quickly below 100 km resulting in relatively low ion and electron densities from 50 to 100 km.

Composition of the Atmosphere – Important observations



- Ion and electron densities rise dramatically above about 100 km.

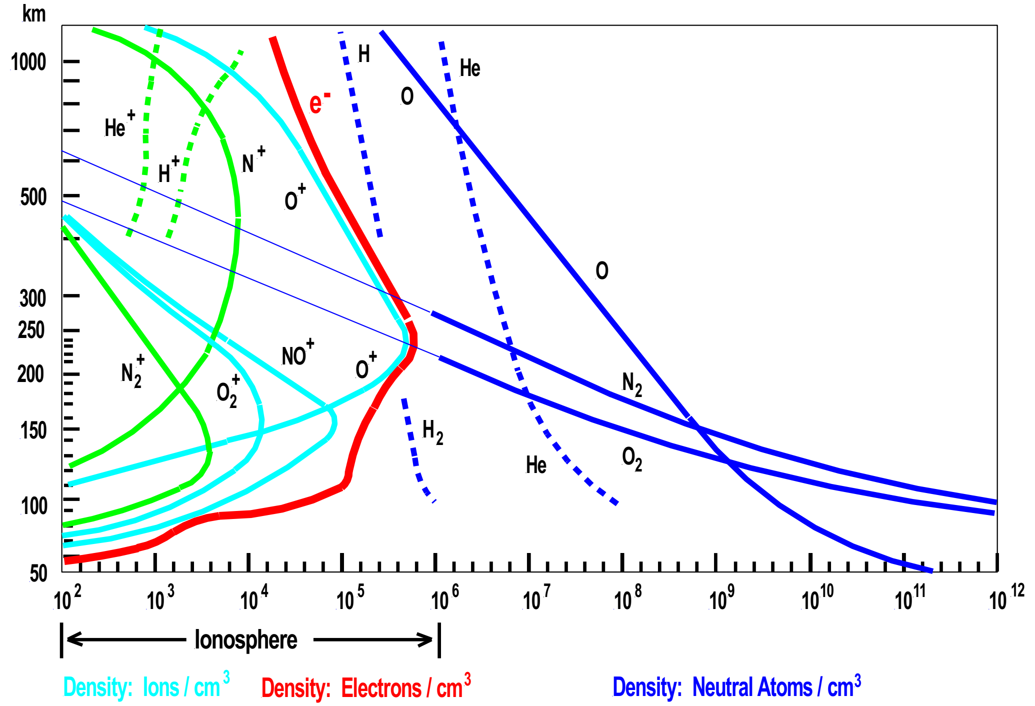
- There are several reasons for this:

  - The mixing of various types of gases, so pronounced in the lower atmosphere, stops rather abruptly above 100 km

  - Dissociation of gas molecules into monatomic forms by EUV solar radiation begins to occur around 100 km and accelerates quickly as altitude increases.

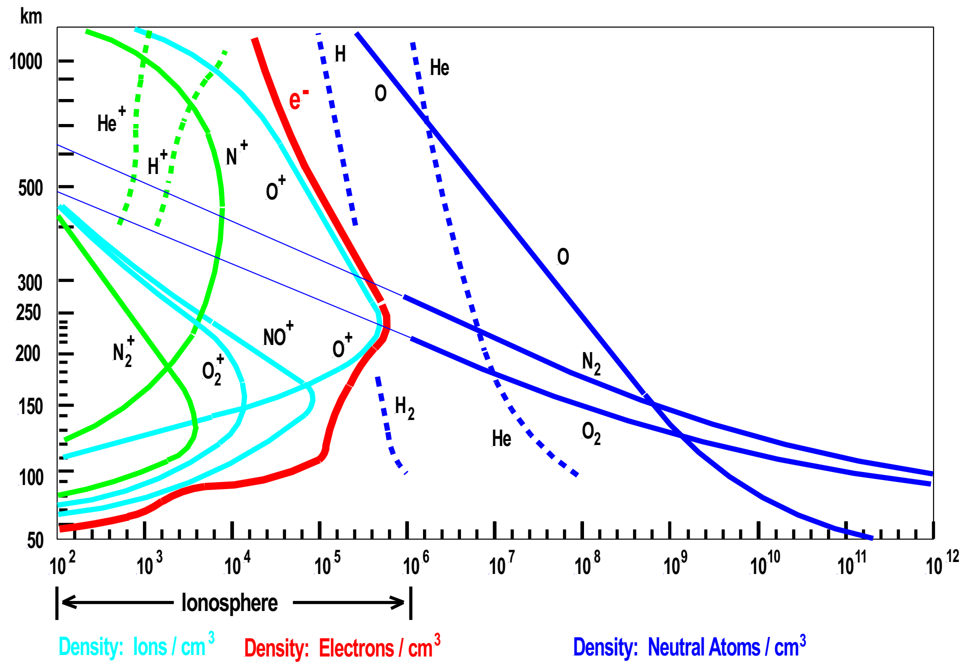
  - Subsequent photo-ionization of the individual dissociated gas atoms, coupled with decreasing rates of recombination, results in ionization levels rising quickly.

Composition of the Atmosphere – Important observations



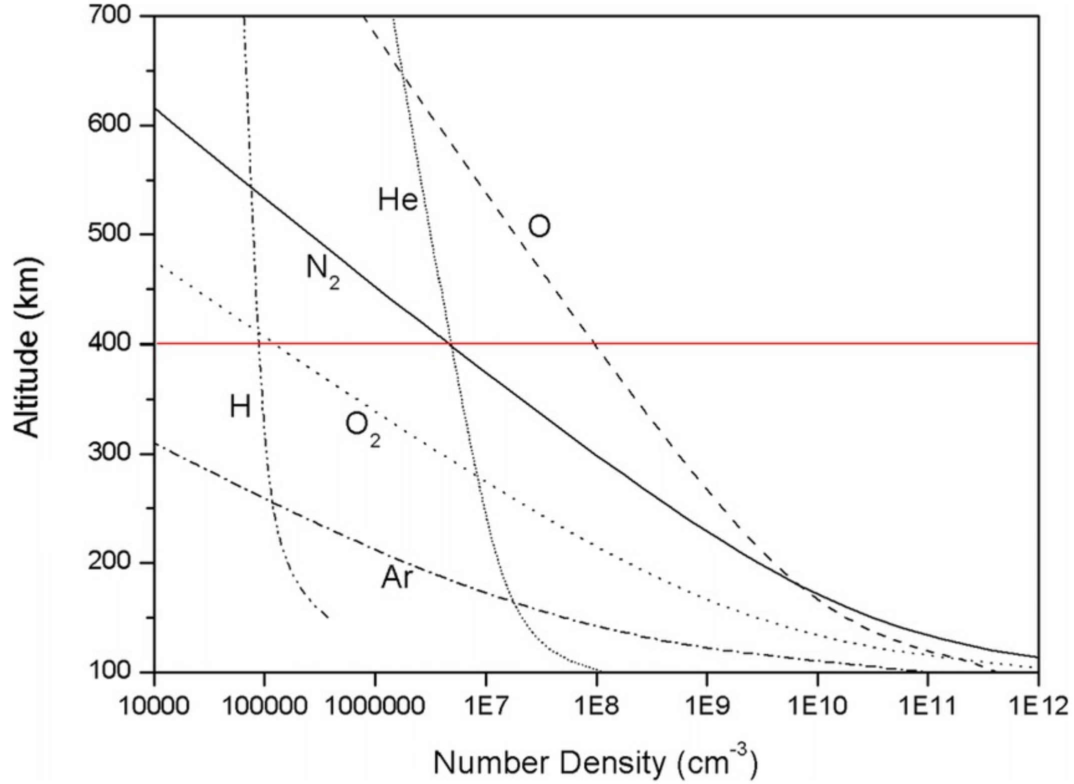
- Because of dissociation, the ionosphere above 200 km is composed primarily of monatomic ions consisting of O^+ , N^+ , H^+ and He^+.
- The vast majority of these ions are ionized monatomic oxygen, O^+ .
- Monatomic N^+ , H^+ and He^+ ions become relevant in the region above 500 km.

Composition of the Atmosphere – Important observations



- The neutral atmosphere becomes essentially collision free above about 500 km with atoms, ions and electrons able to move about freely without colliding.
- Because of this, recombination of electrons and ions in the upper ionosphere occurs slowly primarily by means of monatomic recombination.

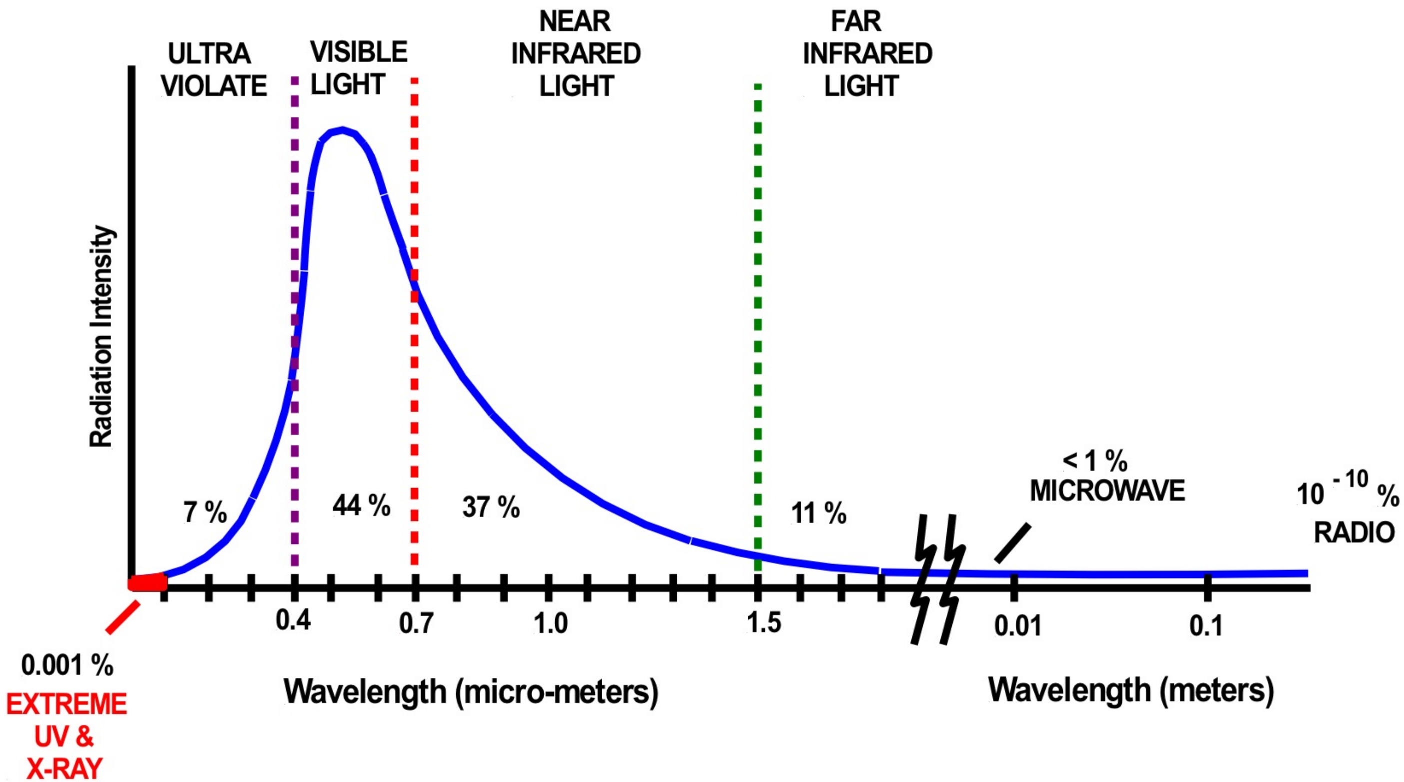
Variable Composition of the Atmosphere



(source: ResearchGate)

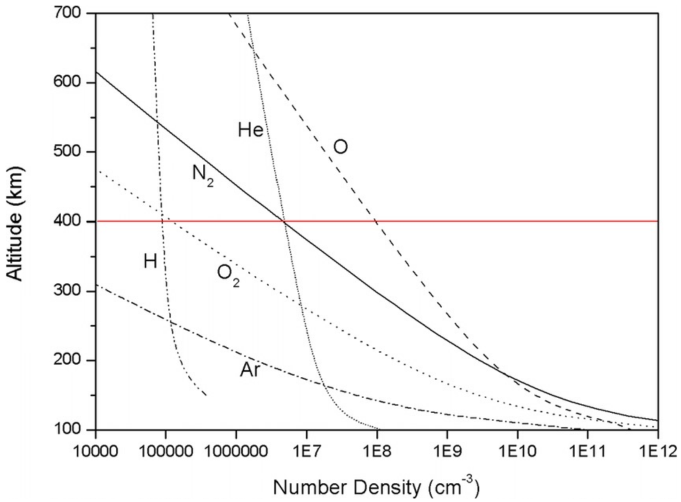
- As can be seen, the atmosphere is not homogeneous
- Instead its chemical composition changes with altitude.

Energy Output of the Sun

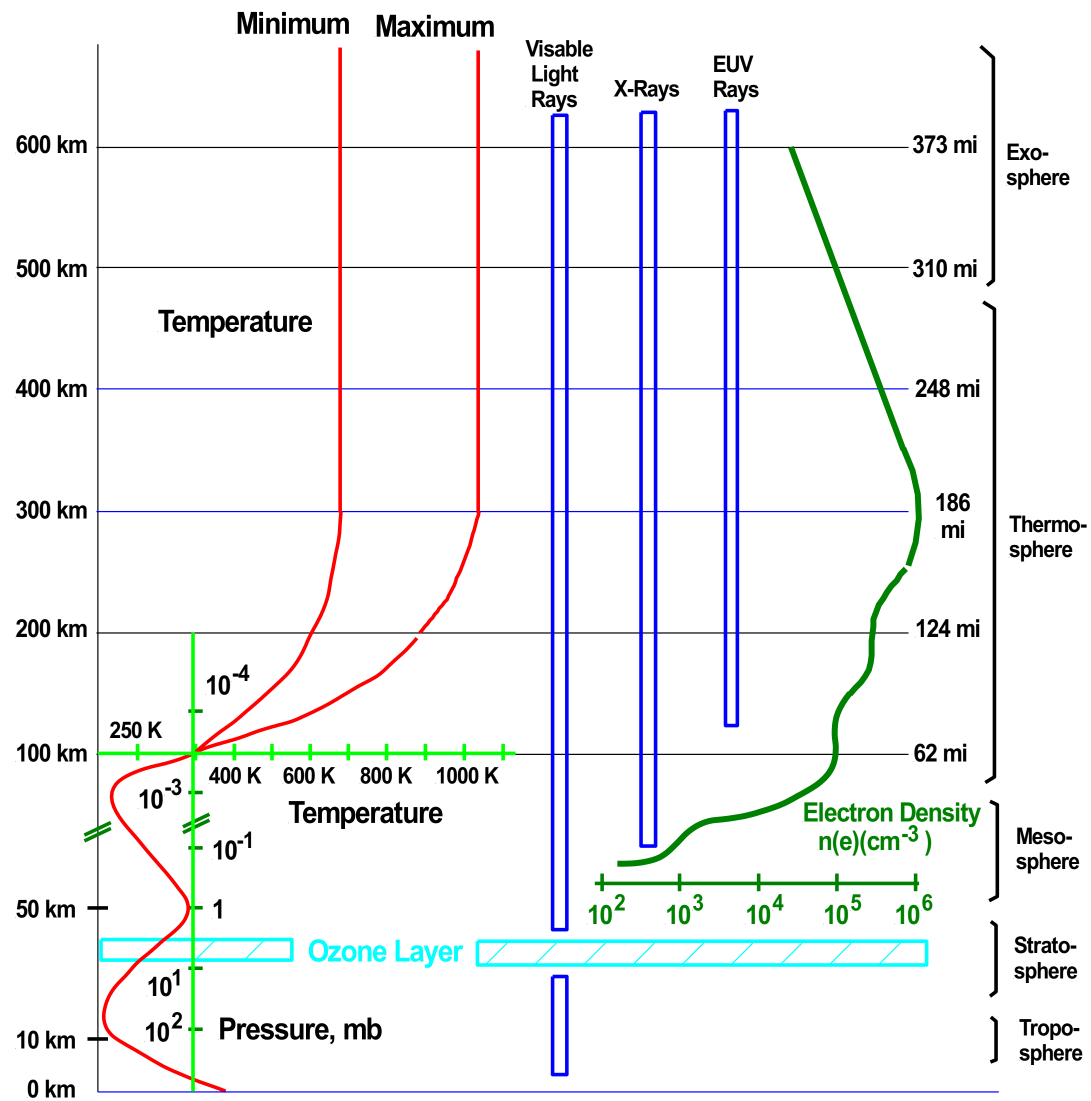


- Radiation from the Sun is not homogeneous either.
- The Sun's spectrum ranges from very short wavelength X-ray and EUV to long wavelength radio waves.
- The Sun's greatest energy output falls in the visible and near infrared light spectrum.
- EUV and X-ray radiation responsible for ionizing Earth's upper atmosphere accounts for only 0.001% of the Sun's energy output.

Structure of Atmosphere

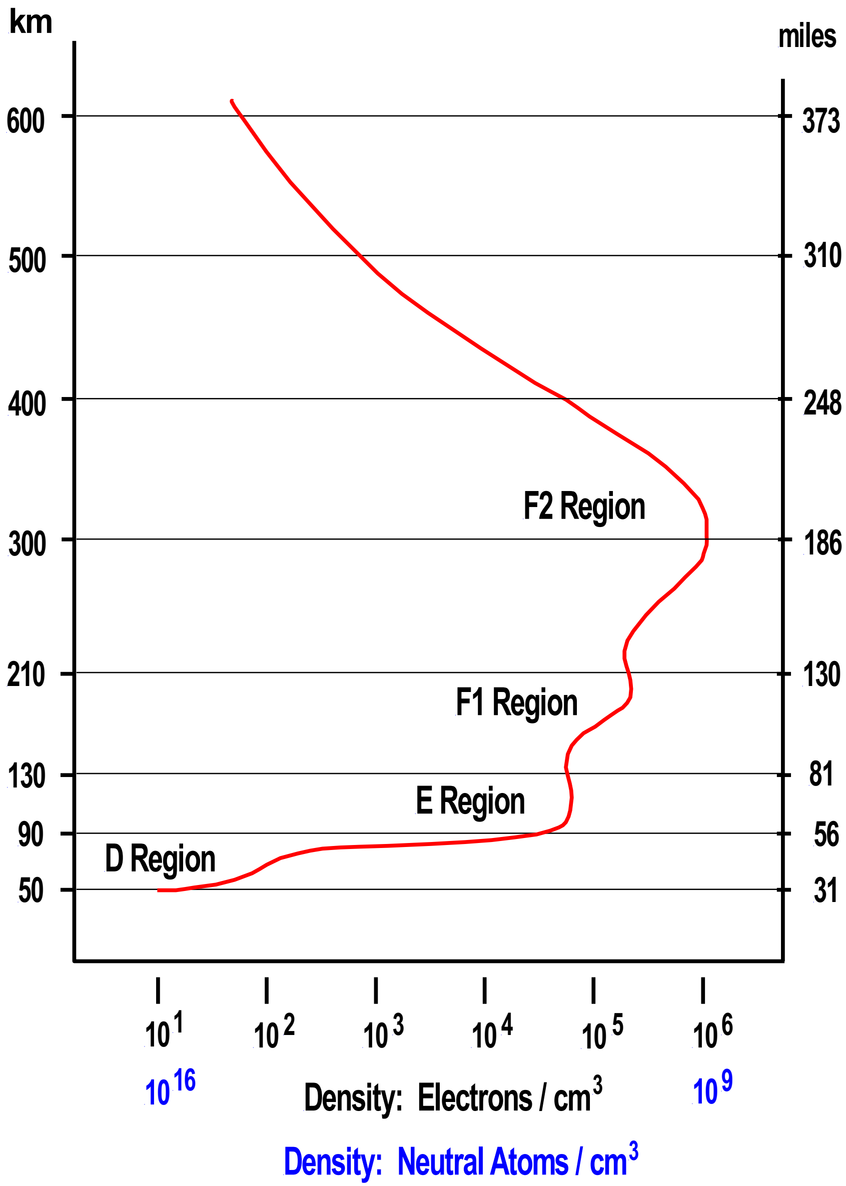


- Not only does the composition of the atmosphere change with altitude, the depth of solar energy penetration changes with wavelength.
- As a consequence, four distinct peaks in ionization levels occur in the ionosphere



(source: adapted from Goodman)

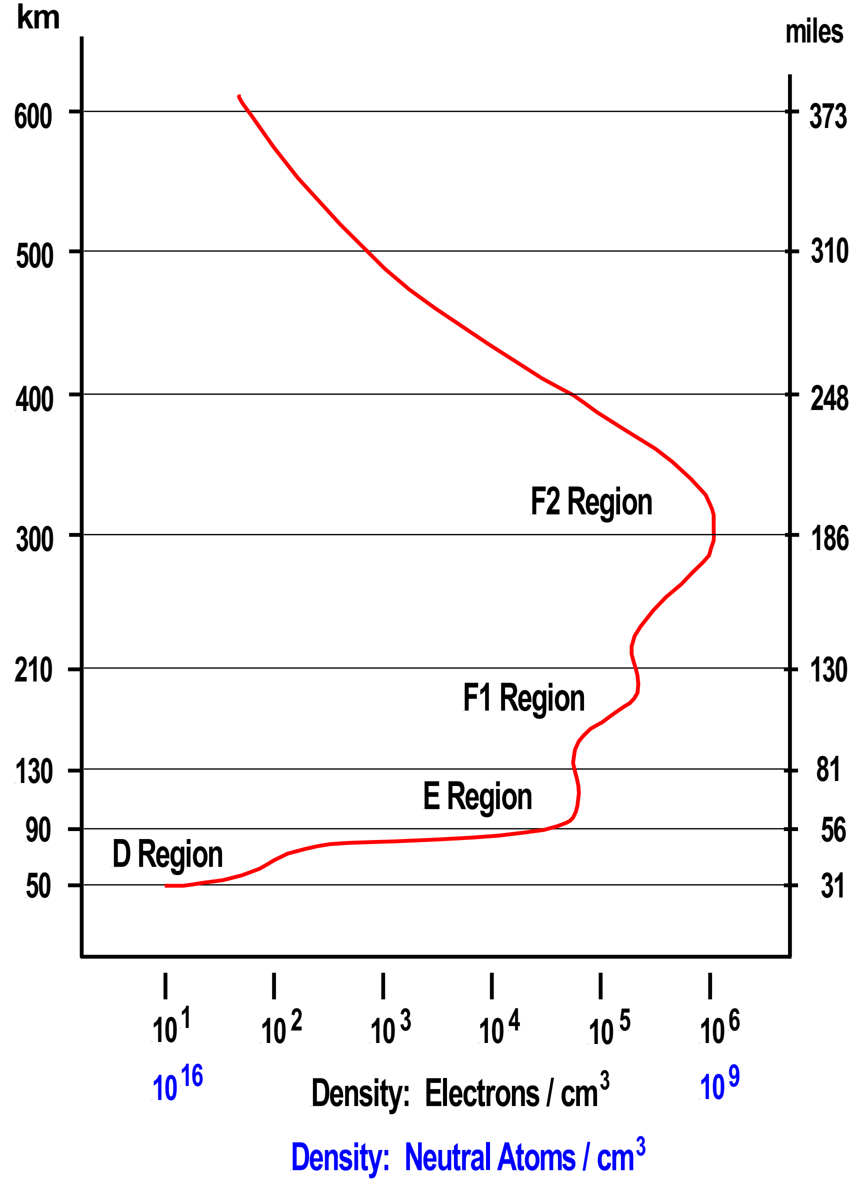
D, E, F1 & F2 Regions



- Historically, these four regions have been identified as the D, E, F1, and F2 regions
- Notice that the concentration of free electrons in the ionosphere (the red curve) is over 1,000 times less than Earth's neutral atmosphere of atoms & molecules in which the ionosphere resides.
- The ionosphere is indeed very thin and wispy.

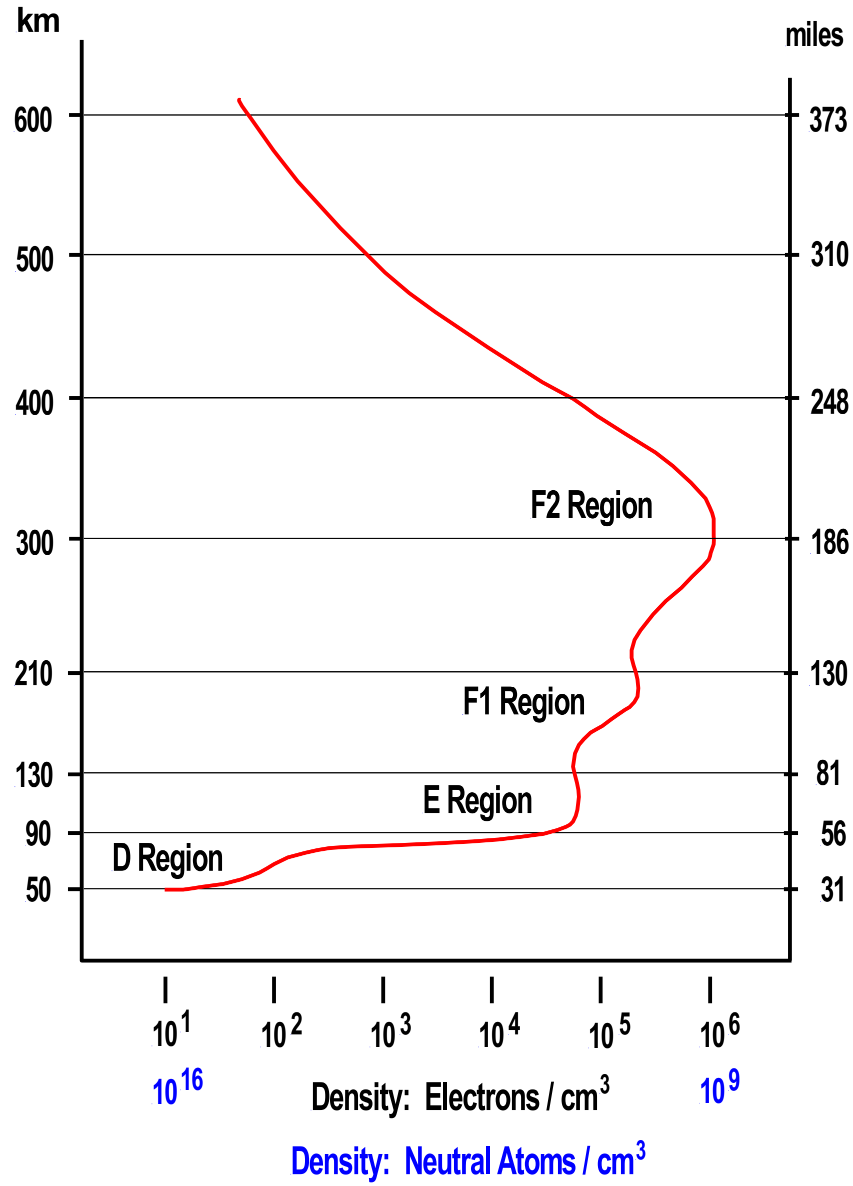
D Region of the Ionosphere

- The D region is the lowest part of the ionosphere at an altitude of 70 to 90 km
- The D region is important because it absorbs high frequency (HF 3 to 30 MHz) radio signals
- The D region is formed by EUV and X-ray ionization of nitric oxide NO and molecular oxygen O_2.
- Following a solar flare, intense x-ray radiation heavily ionize the D region dramatically increasing the absorption of HF radio signals
- Recombination occurs quickly in the relatively dense atmosphere at this altitude so electron density in the D region is low, about 10^2 to 10^3 electrons per cm^3.

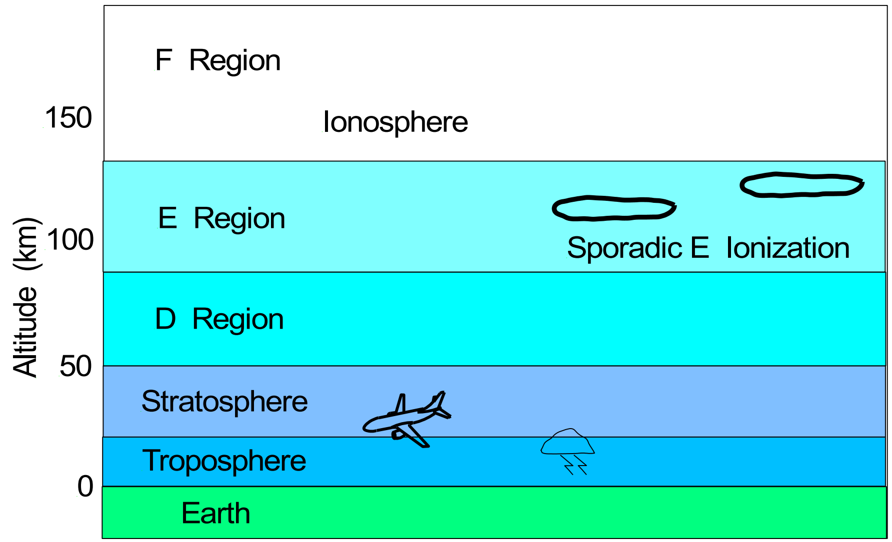


E Region of the Ionosphere

- The E region occurs at an altitude of 90 to 130 km.
- The E region is formed by EUV ionization of molecular oxygen O_2 and nitric oxide NO molecules.
- Peak electron density in the E region occurs at local noon at a height of around 110 km.
- The highest electron density is around 10^5 electrons per cm^3
- The electron density at night is about 10^4 per cm^3
- Electron – ion recombination in the E region is primarily by dissociative recombination

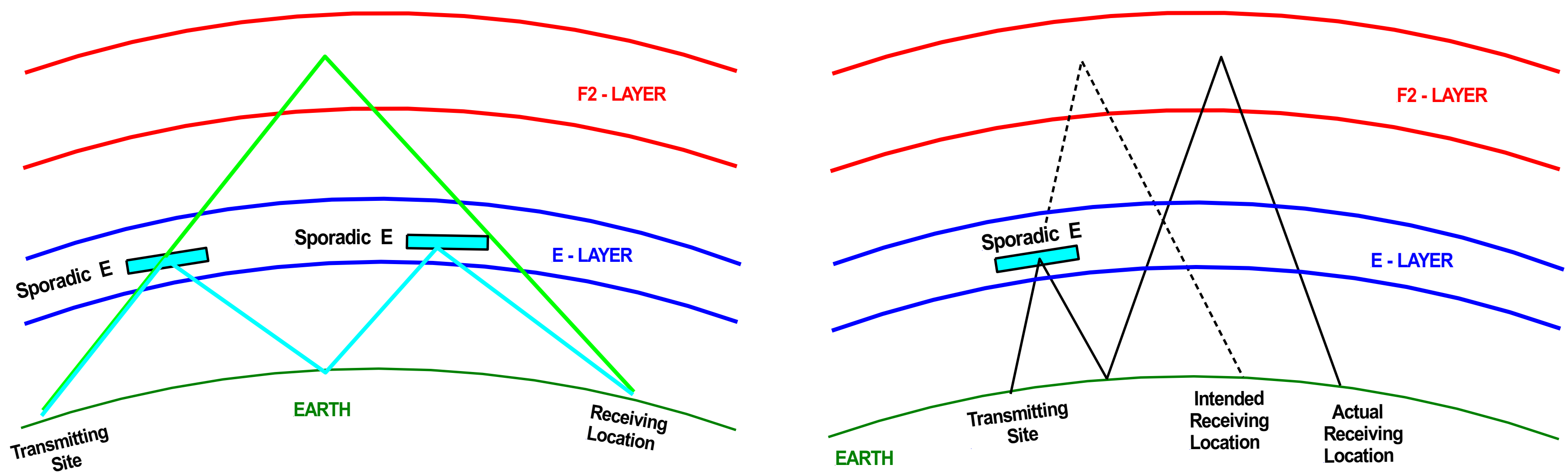


Sporadic E



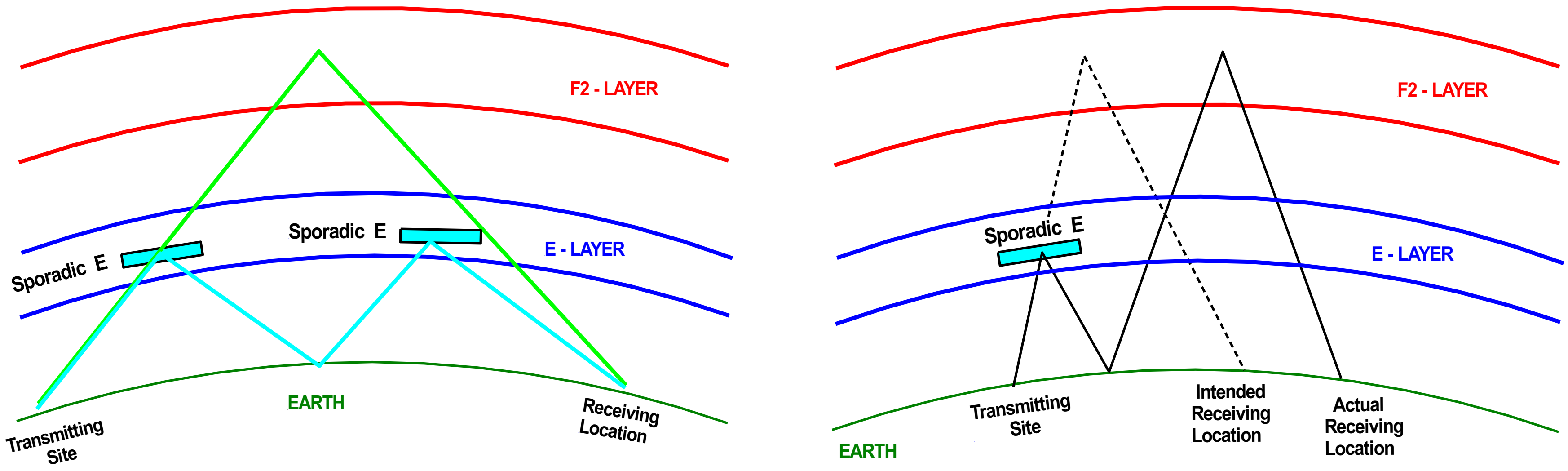
- Sporadic E (E_s) formations are zones of abnormally high ionization within the E region of the ionosphere.
- They are termed sporadic E because they randomly appear in various sizes and shapes, persist for minutes to hours, and occur from one day to the next with little predictability.
- Sporadic E zones are relatively large structures about 2 kilometers thick with horizontal dimensions stretching hundreds of kilometers.
- Sporadic E appearances do not seem to be strongly tied to the photo-ionization process, although at mid latitudes they appear to be more prevalent during summer days than at night or in the winter.

Sporadic E Affect on HF Communications



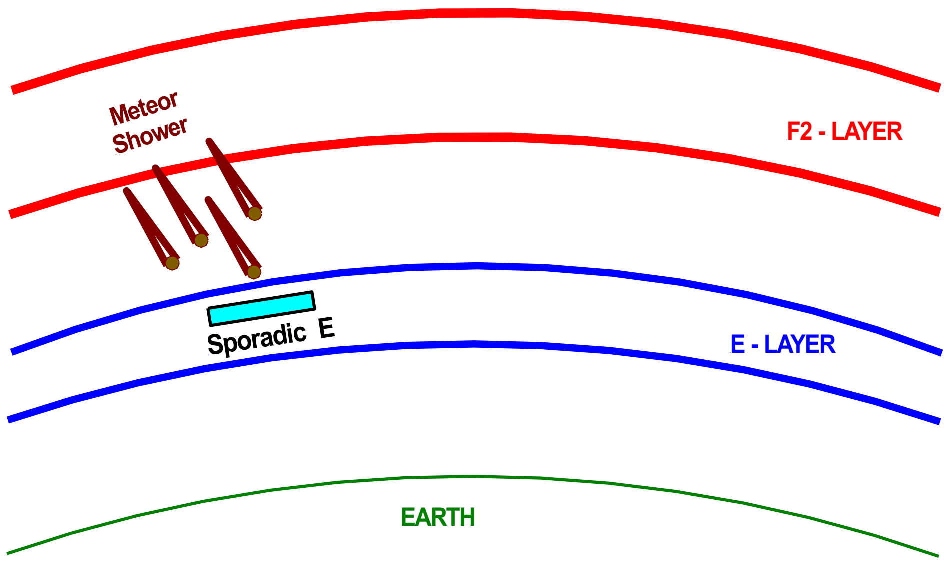
- Sporadic E zones often have electron densities far greater than normal E region levels and at times even greater than the F region.
- Sporadic E zones with high electron densities can appear opaque to radio waves refracting waves that normally would have been refracted at the higher F2 level.
- This can seriously impact HF radio circuits. Instead of a single hop through the ionosphere multiple hops, with more ground reflections and more passes through the attenuating D region, may be required to reach the destination, seriously degrading signal levels. Or the intended receiving location could be missed altogether.

Sporadic E Affect on HF Communications



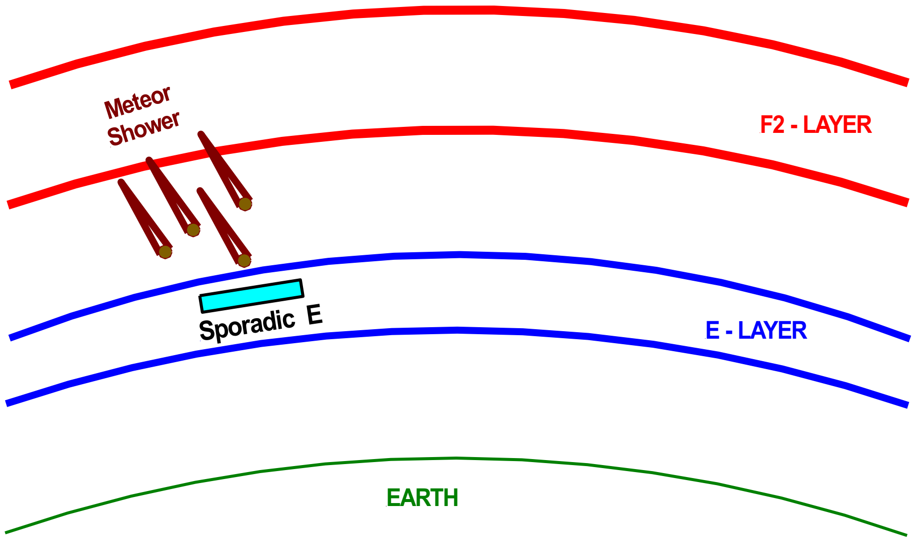
- At other times the F1 and F2 regions can be seen through E_s suggesting that E_s is partially transparent or patchy permitting radio waves to penetrate through the gaps.
- However, a partially transparent Es often leads to weak or fading signals as the Es zone evolves.

Meteoric Debris Lead To Sporadic E Zones



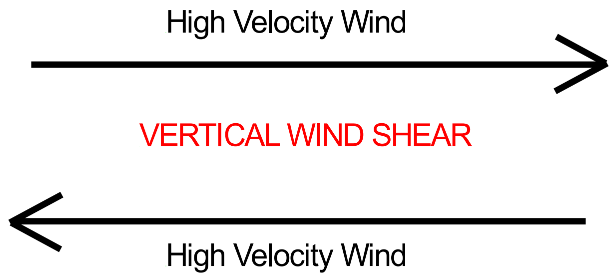
- It is believed that sporadic E patches form as the result of wind shear in the upper atmosphere in combination with meteoric debris.

Meteoric Debris Lead To Sporadic E Zones



- Enormous numbers of meteors burn up in the E region of the atmosphere.
- The meteoric debris is largely monatomic metallic ions consisting of iron, sodium, magnesium, and other similar elements.
- These monatomic ions are relatively small compared to the much larger molecular ions which comprise the E region.
- Because of their small size, the rate of electron-ion recombination is much lower than for molecular ions.

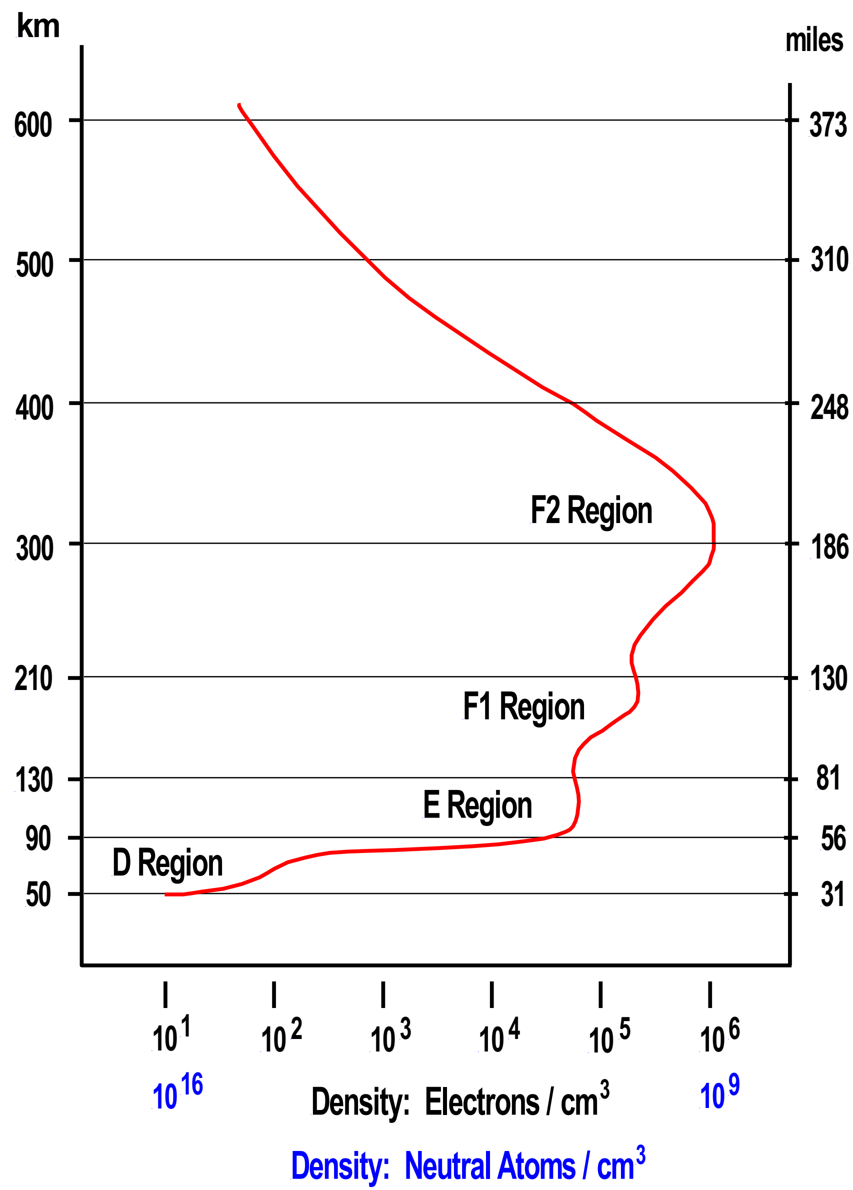
Wind Shear Formation of Sporadic E Patches



- Atmospheric gravity waves, associated with traveling ionospheric disturbances (TIDs), cause upper atmosphere high velocity winds, at slightly different altitudes, to travel in opposite directions, producing the condition known as vertical wind shear.
- Meteoric debris becomes trapped between the wind reversals at locations where the wind velocity tends to be zero.
- Within the pockets of trapped debris, the rate of electron-ion recombination (the rate of electron loss) is lower than elsewhere in the E region.
- Consequently, relatively high electron concentrations develops in these pockets producing sporadic E patches.

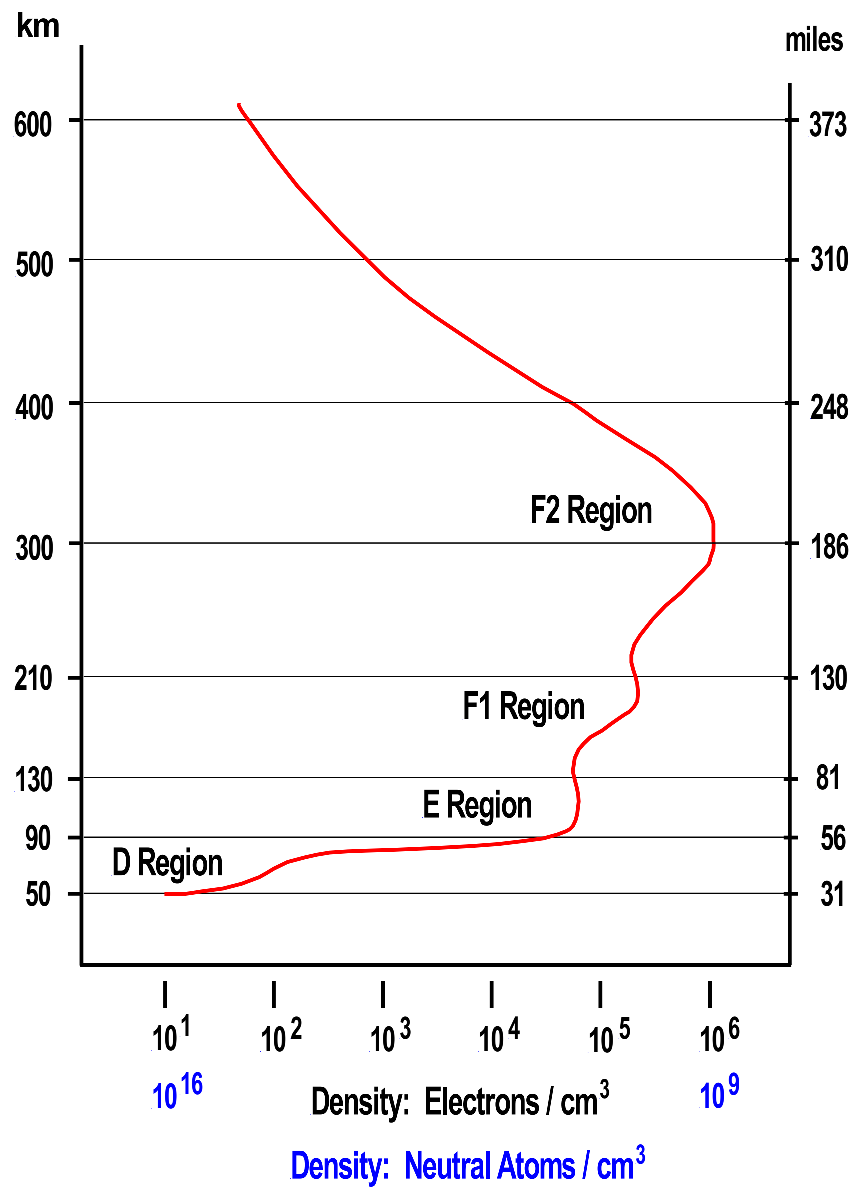
F1 Region of the Ionosphere

- The F1 region extends in altitude from 130 to 210 km.
- The F1 region is formed by EUV ionization of individual oxygen atoms and nitric oxide NO molecules.
- The highest electron density in the F1 region occurs near noon local time
- Noon electron electron density is about 2×10^5 electrons per cm^3
- This peak in electron density occurs at a height of around 180 km.
- The F1 region disappears at night merging with the F2 region.
- It also merges with the F2 region in the winter months during solar maximum.
- Dissociative recombination is the primary F1 recombination mechanism



F2 Region of the Ionosphere

- The F2 region stretches in altitude from 200 to over 800 km.
- It is the most important in terms of HF communications
- However, it is also the most variable and most unpredictable
- The F2 region is formed by EUV ionization of individual oxygen atoms coupled with upward diffusion of electrons from the F1 region.
- The highest electron density in the F2 region is about 10^6 electrons per cm^3
- This peak occurs during the day
- The peak electron density occurs at an height of around 300 km.
- At night the electron density is around 2×10^5 electrons per cm^3



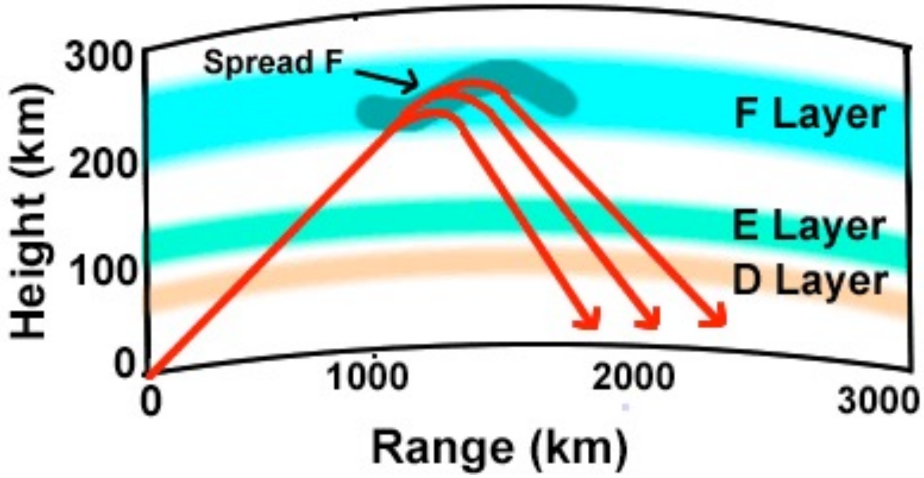
Summary of Region Characteristics

Region	Height (km)	N_{max} (cm^{-3})	f_c (MHz)	Primary Ions	Basis of Formation
D	70 to 90	10^2 to 10^3		NO^+ & O_2^+	EUV & X-rays
E	90 to 130 $h_{max} \sim 110$	$\sim 10^5$ day $\sim 10^4$ night	~ 3 day ~0.3 night	O_2^+ & NO^+	EUV & X-rays
E_s	90 to 130	$\sim 10^6$ (highly variable)		Metallic Ions	Wind shear and meteoric debris
F1	130 to 210 $h_{max} \sim 180$	$\sim 2 \times 10^5$ day ~ 0 night	~ 3 to 6 day Merges with F2 at night	O^+ & NO^+	EUV Photo-ionization
F2	200 to 1000 $h_{max} \sim 300$	$\sim 10^6$ day $\sim 2 \times 10^5$ night	~ 5 to 15 day ~ 2 to 6 night	O^+	EUV Photo-ionization + upward diffusion from F1 region

Irregularities in the Ionosphere

- The ionosphere is not as smooth and uniform as we would like to think.
- Instead, irregularities in electron densities occur throughout the ionosphere.
- These irregularities typically develop as the result of ionospheric instabilities.
- For example, F region ionization at night often breaks up into elongated formations of electrons instead of electrons remaining uniformly distributed throughout the region.
- The size, intensity, and location of irregularities depends on the time-of-day, geomagnetic location, geomagnetic activity, season, and the solar cycle.
- Electrodynamics require electrons in motion to travel along magnetic field lines but not across them.
- Consequently, irregularities become stretched out along magnetic field lines shortly after being created, forming what is known as field-aligned irregularities (FAI).

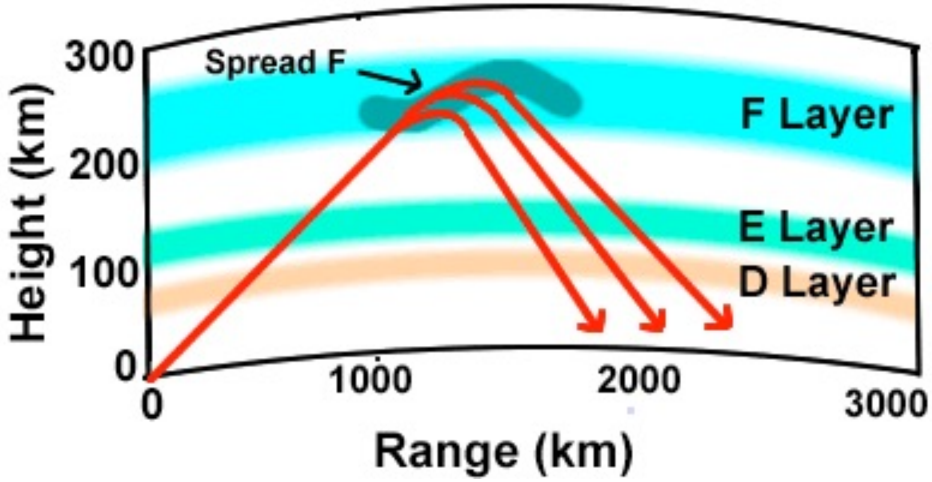
Spread F Irregularities



www.met.nps.edu

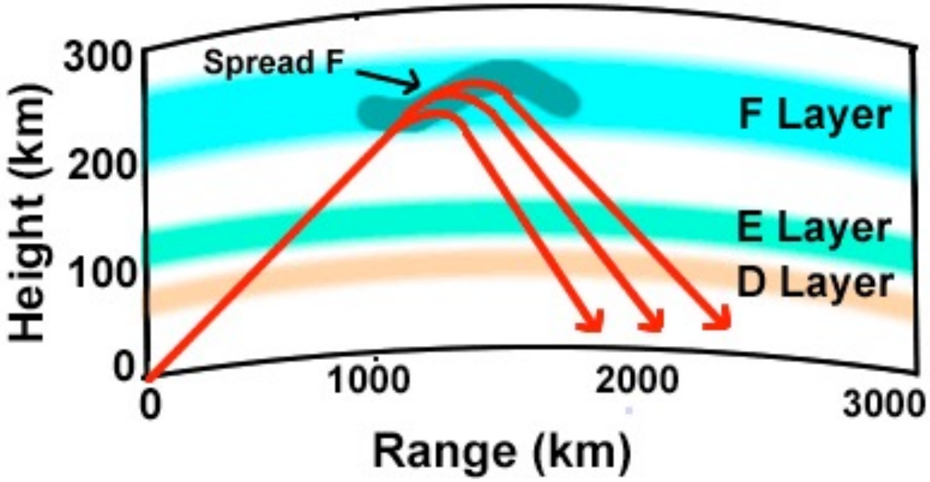
- Signal scattering due to field aligned irregularities results in the spread F phenomena
- Spread F causes an HF signal to be reflected from different heights within the F layer.
- Received HF signals are spread in time and space, making frequency resolution difficult and producing garbled messages
- Received digital data pulses can be up to 10 times wider than the transmitted pulses limiting the data rate of signals that can be successfully transmitted. This type of distortion is known as range spread.
- In addition, spread F produces high fading rates.

Spread F Irregularities continued



- Field aligned irregularities producing spread F are highly variable in size ranging from roughly a 100 km to several thousand kilometers in length & about a kilometer thick.
- In the mid and high latitudes, these irregularities drift horizontally at speeds up to about 100 meters per second.
- Near the equator, the irregularities drift eastward.
- Spread F occurs most frequently in the equatorial and high latitude regions.

Equatorial Spread F



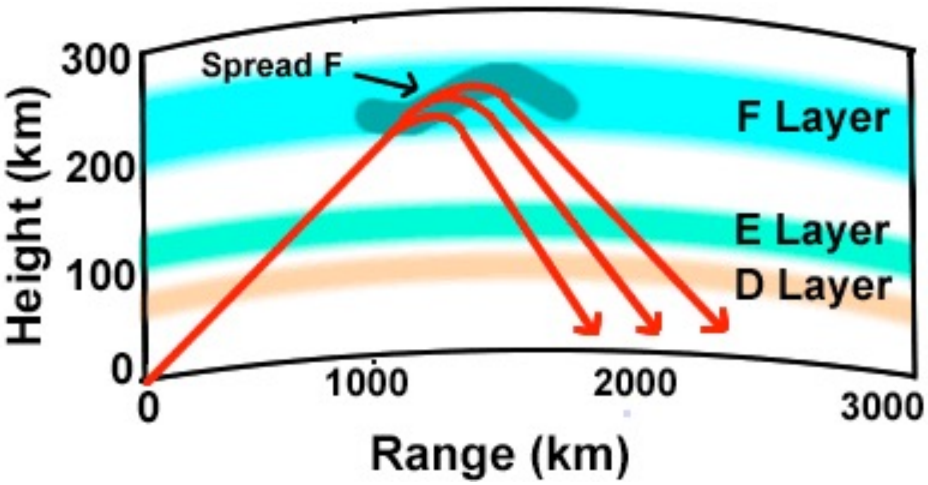
- In the equatorial regions spread F occurs mainly at night, during magnetically quiet days, in a zone straddling the geomagnetic equator from approximately 20^{o} south to 20^{o} north latitude.

- They begin appearing near sunset, in conjunction with the nightly upward drifting F region, peak around midnight, and then generally decrease.

- Equatorial spread F can occur at any time but tends to be more intense and more numerous at solar maximum, during the winter & equinoxes, and during the summer in South America where the magnetic dip and geographic equators are furthest apart.

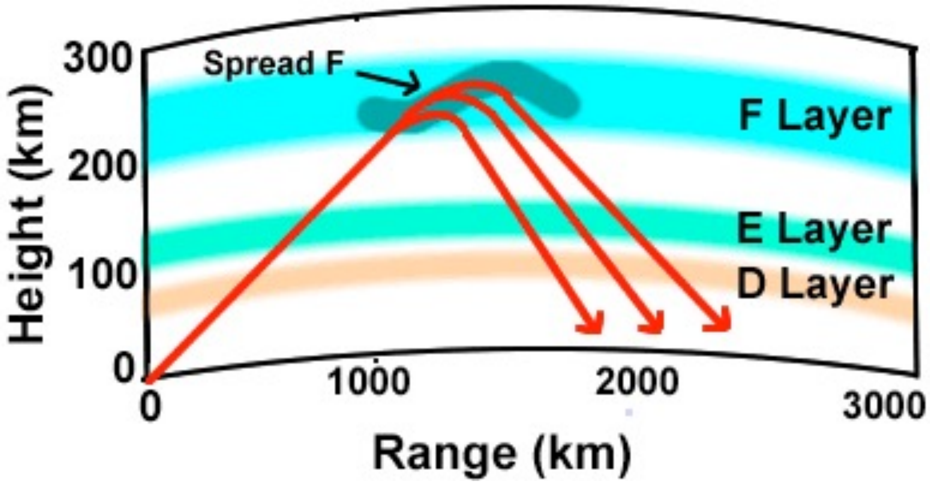
- Equatorial spread F disappears with the onset of a magnetic storm.

High Latitude Spread F



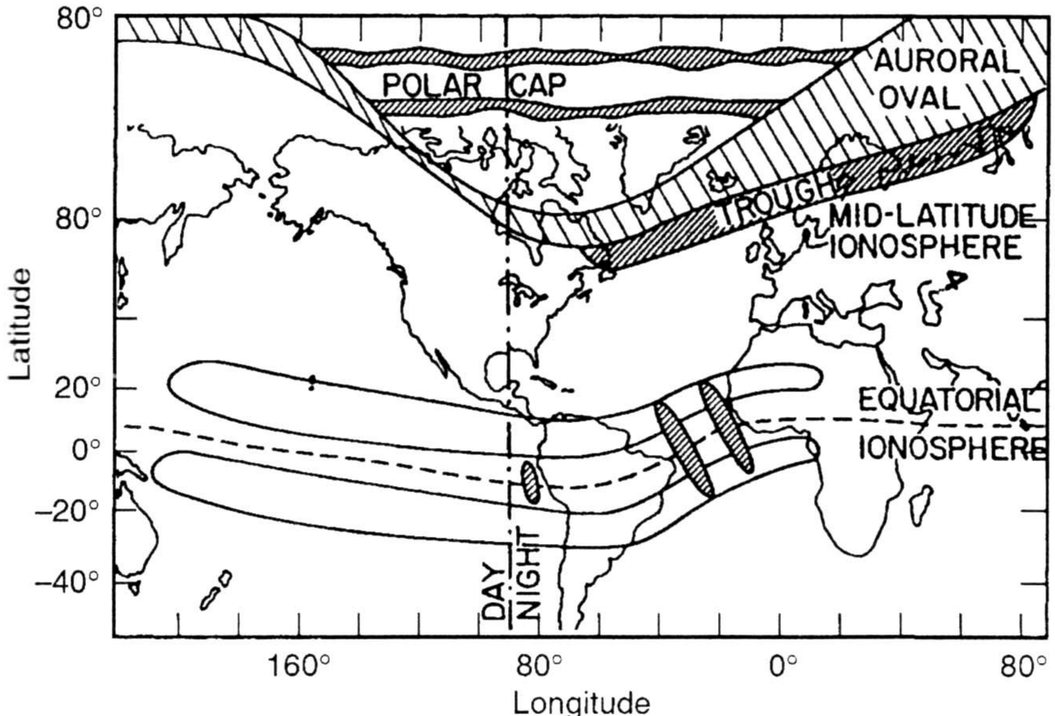
- At high latitudes, spread F begins to appear at around 40° and becomes more pronounced with increasing latitude
- In both the auroral and polar cap zones increased geomagnetic activity dramatically increases the appearance and growth of spread F
- Spread F formations can have electron concentrations several orders of magnitude greater than normal
- Irregularities, including spread F, contribute to the extreme variability of the high latitude F region.

High Latitude Spread F continued



- The greatest occurrence of high latitude spread F is in the winter and at equinoxes
- In the winter spread F is nearly continuous both day and night
- Spread F occurs less frequently in the summer.
- However during summer, near the magnetic dip poles, spread F is pretty much a permanent phenomena, occurring nearly every night and often during the day.
- **There is little spread F at geomagnetic latitudes between 20º to 40º.**

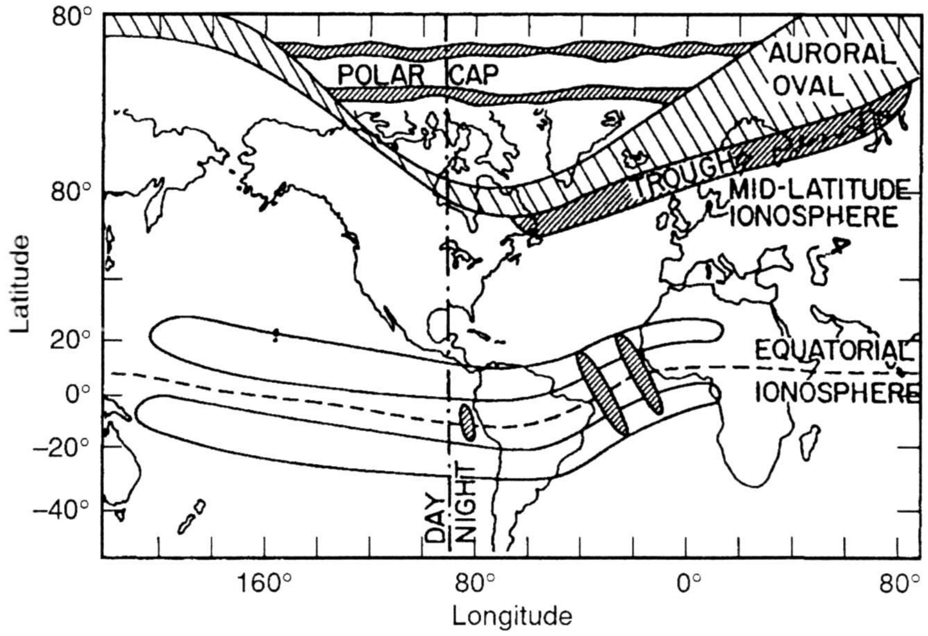
F Region Ionospheric Trough



Source: Goodman

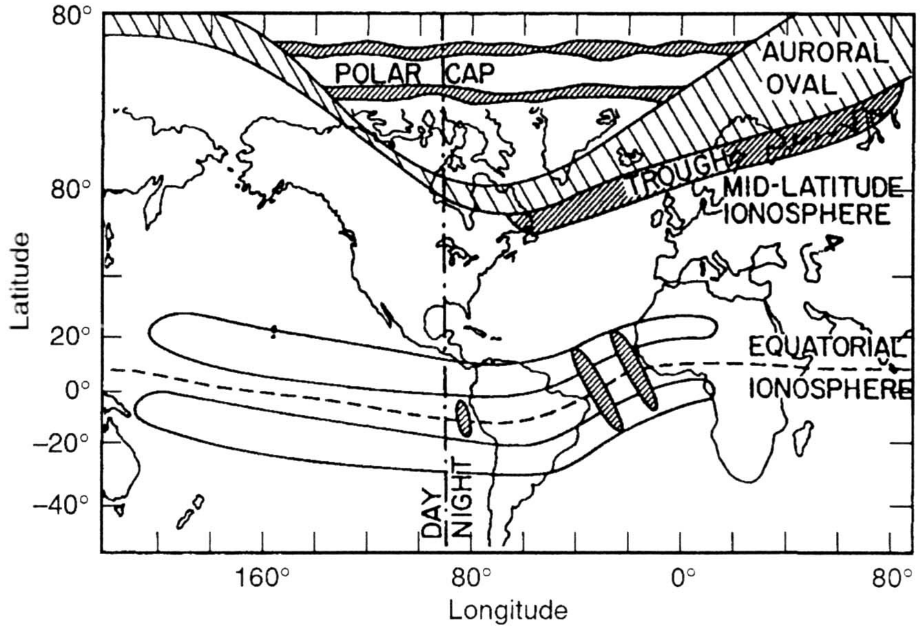
- An ionospheric trough is a relatively narrow ribbon of depleted ionization that stretches from east to west.
- The main or primary trough occurs on the equatorial edge of the auroral oval

F Region Ionospheric Trough continued



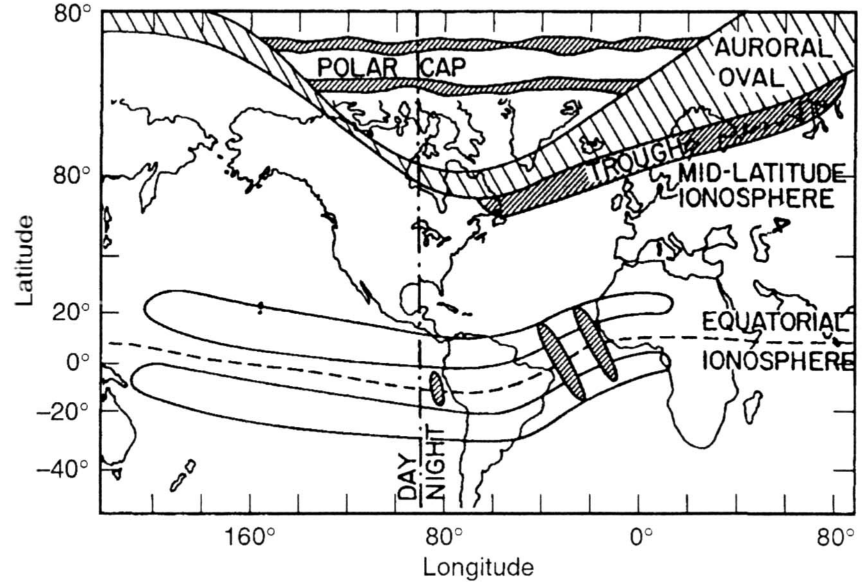
- The depletion occurs primarily in heavy ions such as O+ as well as in electron density.
- The F region Main Trough near 60º latitude marks the boundary between the mid and high latitude regions.
- There are also smaller troughs at higher latitudes.

F Region Ionospheric Trough continued



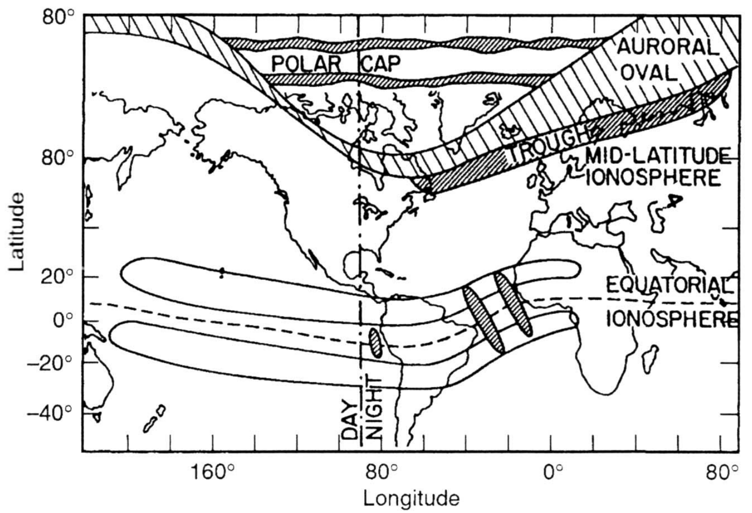
- The main trough is primarily a night time phenomenon, appearing at dusk and extending through the night to dawn.
- The main trough occurs most frequently in the winter and during the equinoxes.
- The trough rarely appears in the summer and then only around midnight.
- In general, it is believed that trough formation is the result of east to west plasma flow (in the opposite direction of Earth's rotation) prolonging the time the plasma remains in the dark.
- The delay provides a longer time for electron-ion recombination to occur before photo-ionization begins, resulting in fewer ions and a lower electron density in the region.

F Region Ionospheric Trough continued



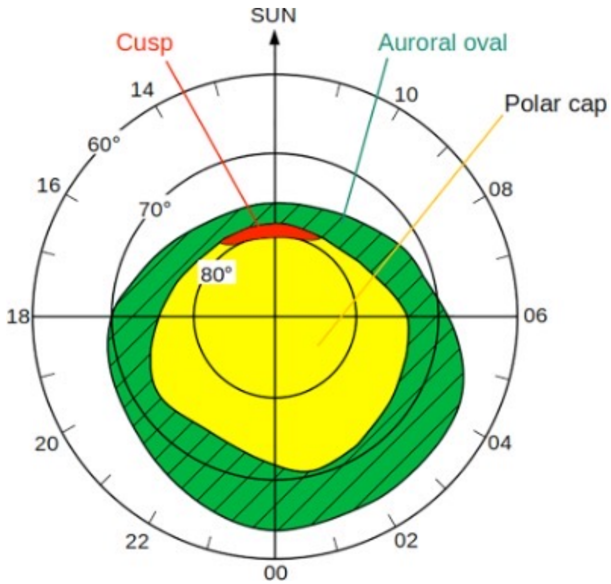
- The appearance of troughs in the southern hemisphere is opposite that in the north.
- Troughs in the southern hemisphere are present in all seasons and at all times of the day
- They occur more often in the summer and during the equinoxes with more troughs during the day than at night.
- It is assumed that these differences are due to hemispheric differences in polar circulation.
- Knowledge of trough appearance and location is important for high latitude and trans-polar radio communications.

Polar Hole



- The polar hole is a recognized distinct feature of the Antarctic polar cap.
- It is a long term depletion that occurs in the winter in years surrounding solar minimum.
- While it can occur sporadically at the equinoxes it hardly ever appears in the summer.
- The hole develops shortly after midnight at magnetic latitudes near 80^o.
- Within the hole electron density at an elevation of 300 km can be as low as 1 x 10^2 per cm^3, compared with 10^5 per cm^3 elsewhere in the polar cap.
- A polar hole does not occur in the Arctic for reasons that are still unknown.

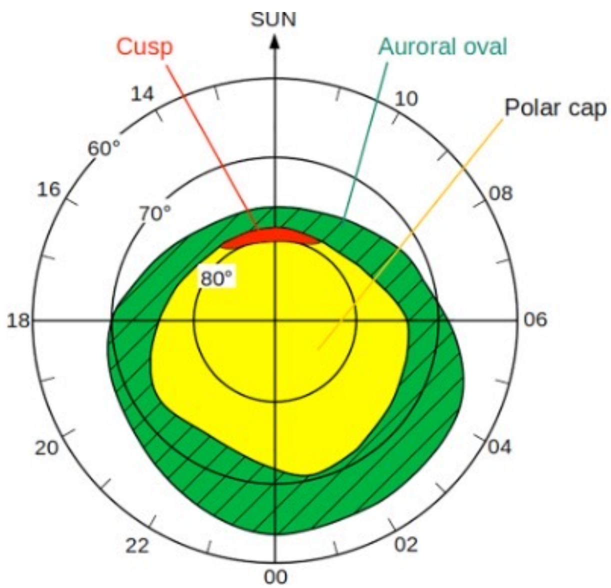
Polar Cap Patches



source: ResearchGate

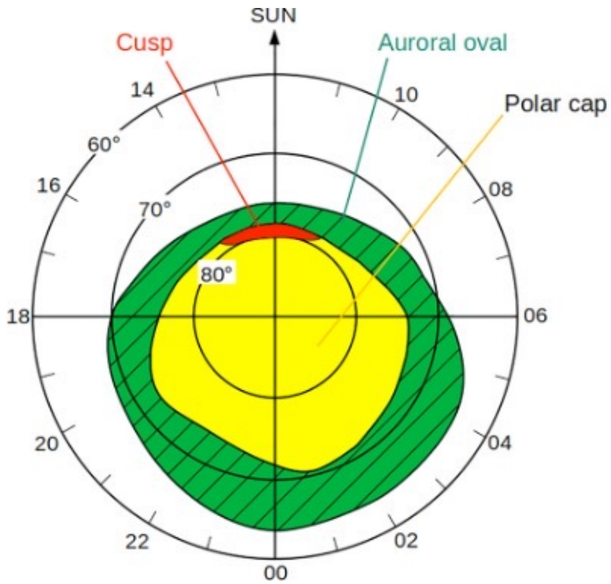
- Patches are roughly circular areas of enhanced electron densities, 200 to 1000 km in size
- They generally occur at night in the polar cap
- Electron densities within a patch are typically 2 – 10 times higher than ambient ionosphere
- A patch may have an electron density of 10^6 electrons per cm^3 compared to 10^5 per cm^3 in the surrounding ionosphere.

Polar Cap Patches continued



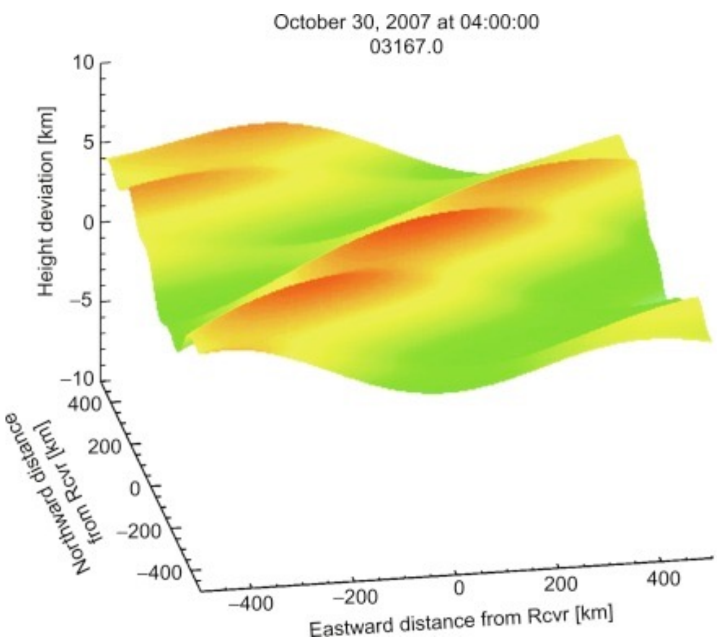
- Patches appear under disturbed ionospheric conditions when the IMF is southward
- A change in polar circulation, due to a sudden increase in solar wind or change in IMF, can detach plasma from the dayside polar cap
- The detached patch drifts over the pole to the night sector at roughly 100 up to 1000 km per second.
- Patches can occur during all seasons of the year but are more frequent during the winter and tend to be stronger at solar maximum.

Auroral Blobs



- Blobs are electron density enhancements in the auroral zone
- Blobs vary considerably in size and are smaller than patches
- Blobs are typically tens of kilometers in size instead of hundreds of kilometers for patches
- The cause of blobs is uncertain
- In general, blobs seem to move with the auroral F region plasma drift.

Travelling Ionospheric Disturbances



source: ScienceDirect.com

- Traveling ionospheric disturbances (TIDs), are large scale wavelike disturbances with wave periods from a few minutes to more than an hour.
- They are typically 100 to 1000 km in size and travel at speeds of 50 - 1000 m/sec
- TIDs have long wave fronts that are tilted forward. Thus, TIDs appear first at high altitudes and move downward as they pass
- Vary large TIDs originate in the auroral zone and travel great distances
- These disturbances are associated with geomagnetic storms.