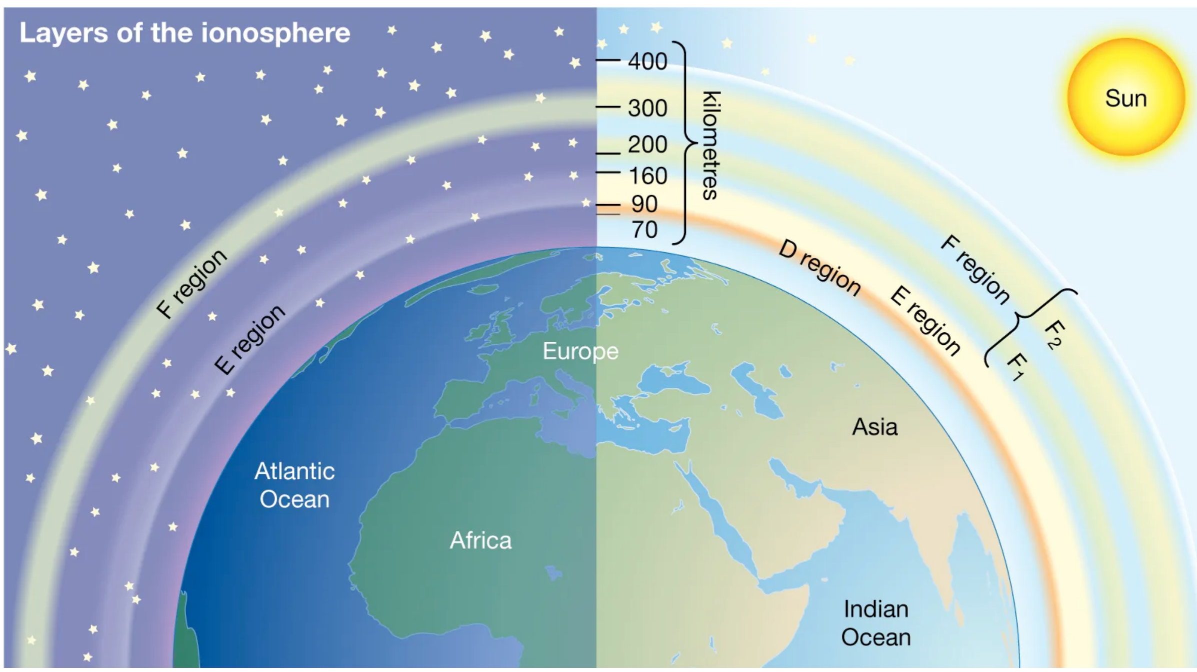
Chapter 11

Earth's Ionosphere



(Encyclopaedia Britannica)

www.skywave-radio.org

Revised 12/2024

11 Ionosphere

The ionosphere is a wispy region of ionized air located in Earth's upper atmosphere. The ionosphere extends from roughly 50 to over 800 km above Earth's surface. During the day the ionosphere separates into four distinct regions as illustrated in Figure 1. Historically these regions are labeled D, E, F_1, and F_2. Throughout the day these regions are moderate to heavily ionized with median altitudes shown in the figure. At night the D and F_1 regions disappear leaving only lightly ionized E and F regions. The ionosphere electron density is a thousand times less than the surrounding atmosphere in which it is imbedded. The ionosphere is indeed wispy! The red line in Figure 2 shows the ionosphere density profile verses altitude as well as the various regions of the ionosphere during daylight hours. The ionosphere's peak density is approximately 10^6 free electrons per cubic centimeter while the density of the surrounding atmosphere is 10^9 neutral atoms per cubic centimeter.

High frequency HF radio waves refract (bend) in the ionosphere as they encounter increasing levels of ionization, making possible long distance over the horizon radio communications. Figure 3 illustrates an HF radio wave refracting from the F_2 region of the ionosphere. Notice in Figure 3 that ionization (black dots) reaches a maximum near the middle of the ionosphere causing the radio wave to bend back to Earth.

The ability of the ionosphere to bend HF radio signals back to Earth is important for long distance emergency communications. In addition, it makes possible the fun exciting amateur radio hobby allowing amateur radio enthusiasts to contact other amateur (ham) radio operators around the world. **However**, of far greater importance, the ionosphere protects the Earth's surface from deadly x-ray and extreme ultra violate (EUV) solar radiation. Without the ionosphere life on Earth, as we know it, would not exist.

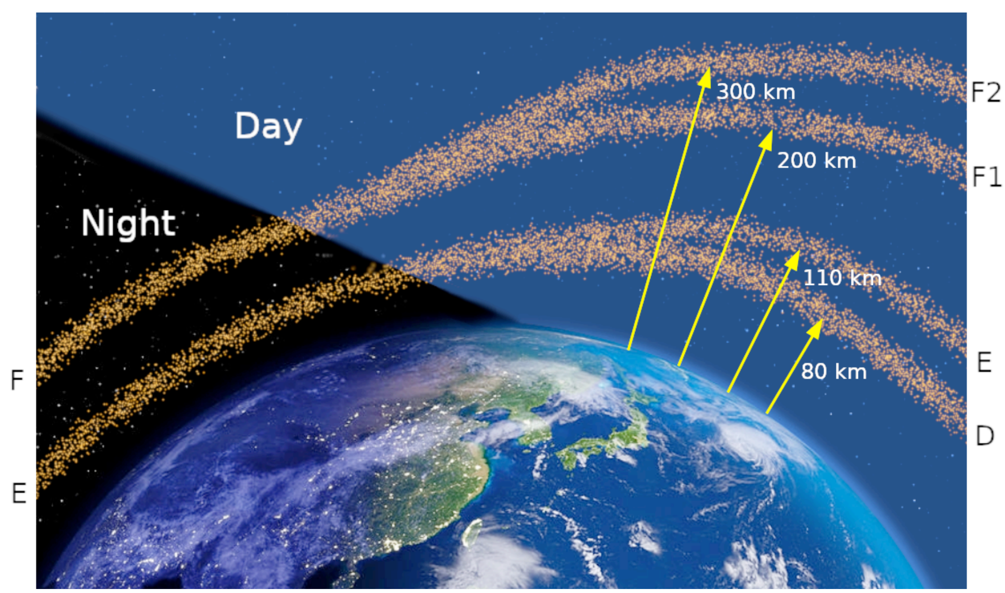


Figure 1 Ionosphere imbedded in Earth's upper atmosphere (source: Wikipedia)

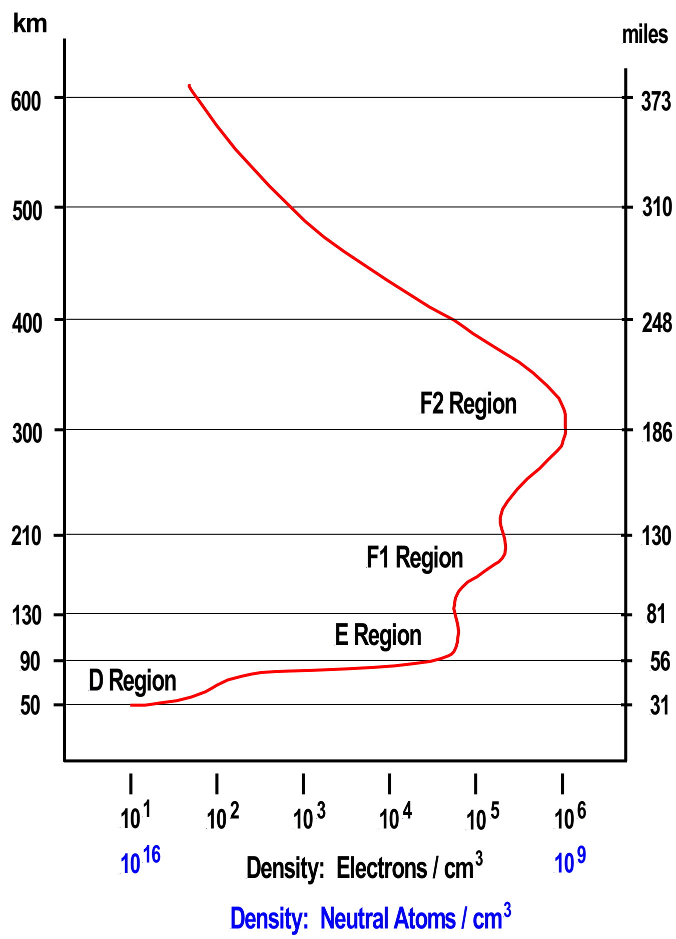


Figure 2 Regions of Earth's Ionosphere (source: author)

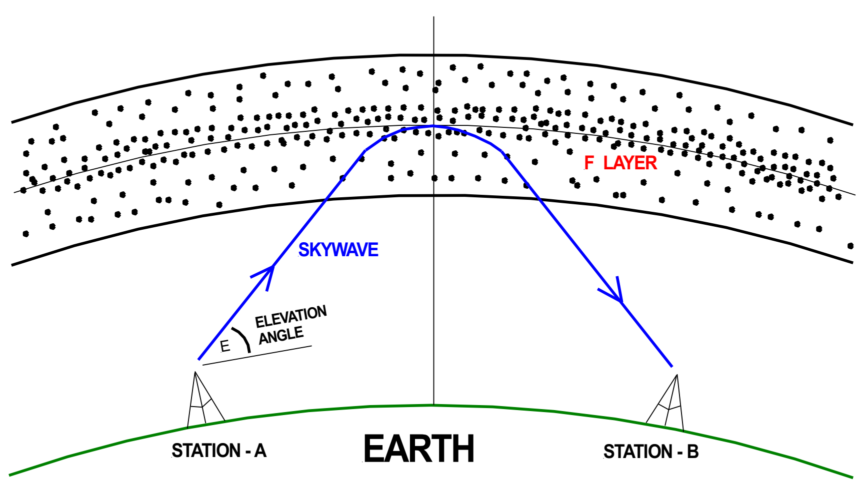


Figure 3 Over the horizon communications via the ionosphere (source: author)

11.1 Ionization and Recombination

The electron density at any point in the ionosphere is a balance between ionization, the formation of ions and free electrons, and the recombination of electrons and ions back into neutral atoms. Recombination occurs all the time, both day and night. However, ionization can only occur during the day when solar x-ray and extreme ultra-violate radiation is present. Consequently, electron (and ion) concentrations are greatest during the day and decline throughout the night reaching a minimum just before sunrise. The primary atoms and molecules ionized by solar radiation include molecular oxygen (O_2), nitric oxide (NO), atomic oxygen (O), atomic nitrogen (N), hydrogen (H), and helium (He). The recombination of electrons and ions back into neutral atoms is an electron loss process. That is, free electrons are lost from the ionosphere when they recombine with ions. In the upper ionosphere diffusion of electrons out of one region into a region higher or lower in altitude also constitute a loss. Atmospheric density is so extremely low in the upper ionosphere that electron loss through diffusion is significantly greater than loss through recombination. As we will see, diffusion into a particular region results in an electron gain.

11.1.1 Ionization

During the day, photons of EUV and x-ray radiation from the Sun bombard Earth's upper atmosphere. A photon's energy is transferred to an atom when the two collide, annihilating the photon and increasing the atom's kinetic energy. However, if the transferred energy is within the proper range, one of the atom's electrons may acquire sufficient energy to escape from the atom. When this happens a free electron and a positive ion are formed as illustrated in Figure 4. This is known as the photo-ionization process. The energy absorbed by an escaping electron is usually greater than that required to break free, providing the electron with the kinetic energy necessary to dart away from its parent atom.

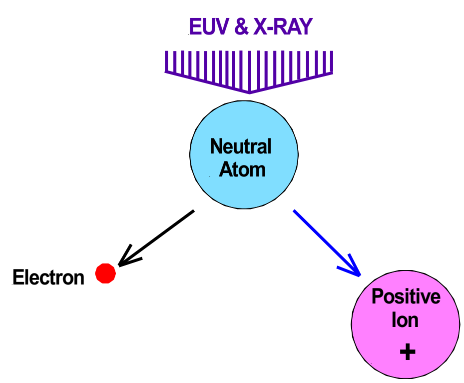


Figure 4 Photoionization process (source: author)

Photo-ionization also occurs at the molecular level producing O_2^+, NO^+, plus other molecular ions, particularly below 200 km. In Figure 5 an oxygen molecule O_2 is photo-ionized.

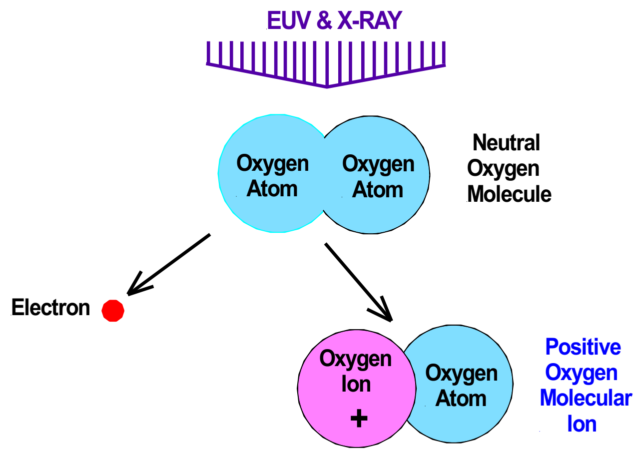


Figure 5 Photo-ionization of an oxygen molecule (source: author)

For ionization to occur, an atom or molecule must absorb a photon with energy greater than the atom or molecule's ionization potential. The energy of a photon is equal to

$$E = hf$$

where

E = energy
h = Planck's constant
f = the frequency of the colliding photon = c/λ
c = speed of light
λ = the photon's wavelength

Substituting in c/λ for frequency f yields the photon's energy in terms of its wavelength.

$$E = h\frac{c}{\lambda}$$

Since h and c are both constants, the energy available to ionize an atom or molecule depends entirely on the photon's wavelength. Photons with short wavelengths have more energy than photons with long wavelengths. For example, X-rays with wavelengths of 0.1 to 17 nm have more energy than extreme ultra-violate light with wavelengths of 17 to 175 nm. Consequently, there is a maximum radiation wavelength λ_{max} that is capable of ionizing a particular type of atom or molecule. Ionization will occur if the photon's wavelength is shorter than λ_{max}. However, no ionization will occur if the photon's wavelength exceeds λ_{max}. It does not matter how many photons with wavelengths > λ_{max} bombard an atom or molecule, it will not be ionized until struck by a photon with wavelength less than λ_{max}. The ionization potential and λ_{max} for atmospheric atomic and molecular species is provided in Table 1.

Species	**Ionization potential (ev)**	**λ_{max} (nm)**
NO	9.25	134.0
O_2	12.08	102.7
H_2O	12.60	98.5
O_3	12.80	97.0
H	13.59	91.2
O	13.61	91.1
CO_2	13.79	89.9
N	14.54	85.3
H_2	15.41	80.4
N_2	15.58	79.6
A	15.75	78.7
Ne	21.56	57.5
He	24.58	50.4

Table 1 Ionization Potentials of Atmospheric Gases

In addition to wavelength, the rate of ionization depends on the

- Intensity of solar x-ray and EUV radiation, that is, the number of photons (with the required wavelengths) impacting the atmosphere per second. Intensity varies daily and with the solar cycle,
- Zenith angle. Ionization is greatest when the Sun is directly overhead (small zenith angle),
- Seasonally. Ionization is greatest in the summer when the zenith angle is smallest,
- Altitude. Ionization begins above 800 km, reaches a maximum around 250 - 350 km, and disappears below 50 km.

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

11.1.2 Recombination

The ionosphere loses free electrons through the process of recombination. In its simplest form, called Radiative Recombination, a positive ion captures a free electron converting the ion back into a neutral atom as illustrated in Figure 6.

Like ionization, recombination also occurs at the molecular level through a process known as Dissociative Recombination. In Figure 7 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 ion causing the molecule to split (dissociate) into two neutral oxygen atoms.

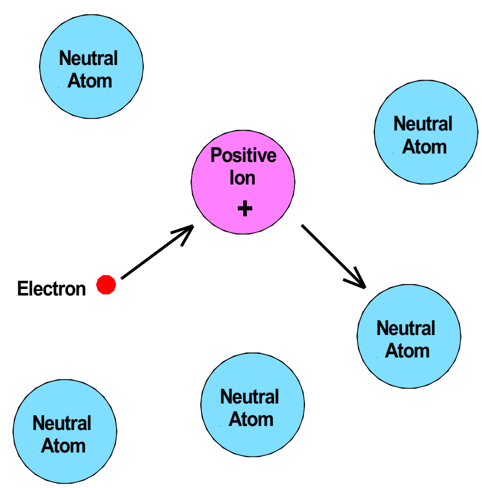


Figure 6 Radiative Recombination process (source: author)

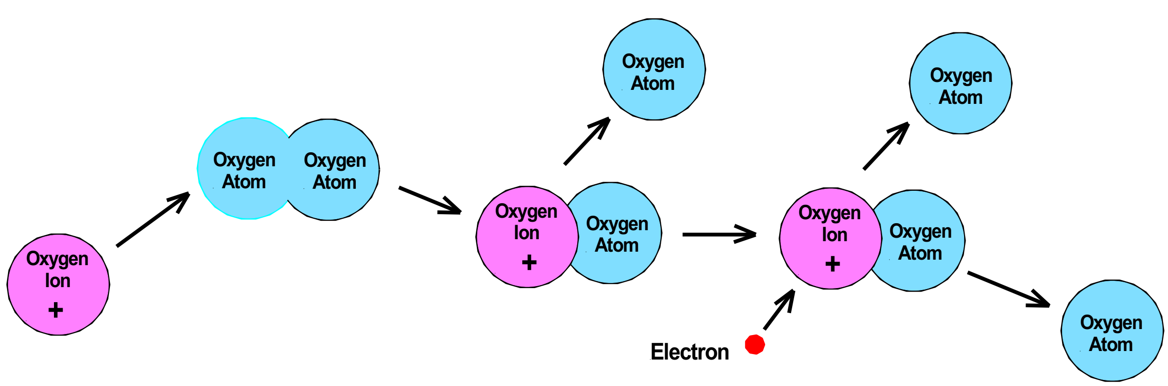


Figure 7 Dissociative Recombination process (source: author)

The dissociative recombination rate is about 100,000 faster than radiative recombination. Consequently, recombination for an atomic ion, such as O^+, is generally a two-step process. The ion must first be converted to a molecular ion through a process called charge transfer. It can then complete the recombination process through dissociative recombination.

Typical charge transfer reactions are

$$O^+ + O_2 \rightarrow O_2^+ + O \qquad \textit{producing a positive molecular oxygen ion}$$

$$O^+ + N_2 \rightarrow NO^+ + N \qquad \textit{producing a positive nitric oxide ion}$$

Dissociative recombination reactions include

$$e + O_2^+ \rightarrow O + O$$

$$e + NO^+ \rightarrow N + O$$

$$e + N_2^+ \rightarrow N + N$$

In the low, densest part of the ionosphere, below approximately 90 km, free electrons are also lost through a process called Attachment (Figure 8). 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.

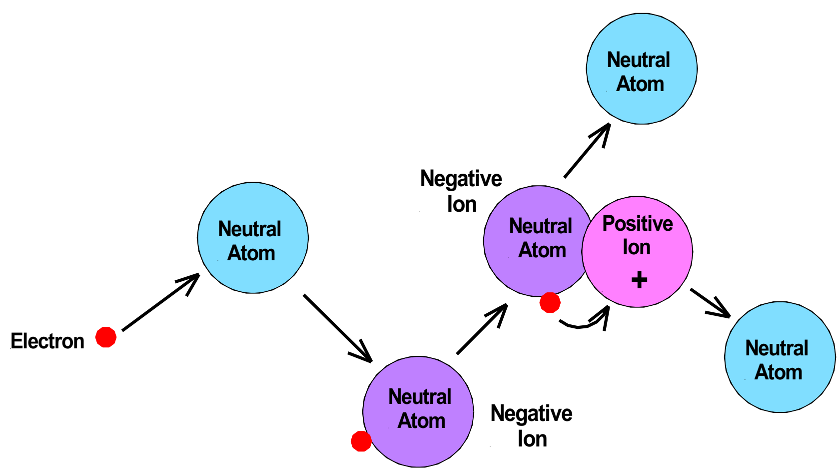


Figure 8 Electron Attachment process (source: author)

11.1.3 Typical Lifetime of Free Electrons

The typical lifetime of free electrons in the D, E, F1, and F2 regions of the ionosphere are:

- D – Region: a small fraction of a second
- E – Region: 20 seconds
- F1 – Region: 1 minute
- F2 – Region 20 minutes

11.2 Vertical Structure of the Ionosphere

The ionosphere is located in the mesosphere and thermosphere regions of Earth's upper atmosphere. In Figure 9 the density of the ionosphere vs altitude is shown by the red curves on the right side of the figure. The vertical temperature profile and regions of the neutral atmosphere are shown on the left. The ionosphere and the neutral atoms of the upper atmosphere occupy the same physical space. However, the density of neutral atoms is several orders of magnitude greater than that of ions. For example, the peak ionosphere electron density occurs at an altitude around 250 to 350 km (Figure 10). In Figure 11 the peak density of positive oxygen ions (O^+) occurs at about same altitude, both with a density of around 10^6 electrons (or O^+ ions) per cubic centimeter. In contrast, the density of neutral oxygen atoms (O) at the same altitude is around 10^9 atoms per cm^3 (Figure 11). The ionosphere is very wispy.

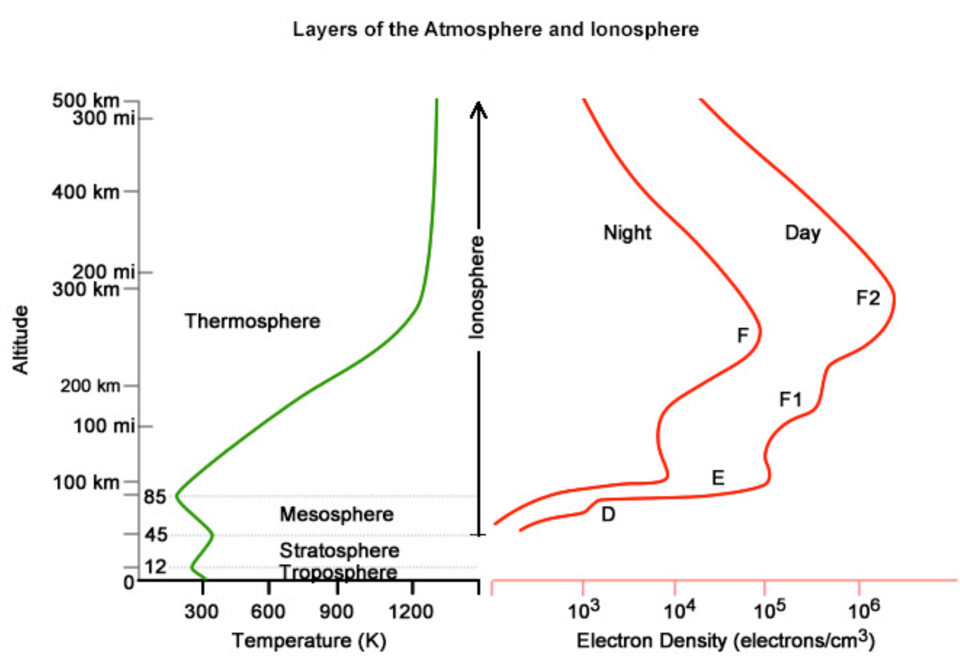


Figure 9 Neutral atmosphere and ionosphere occupy same space (source: http://kejian1.cmatc.cn)

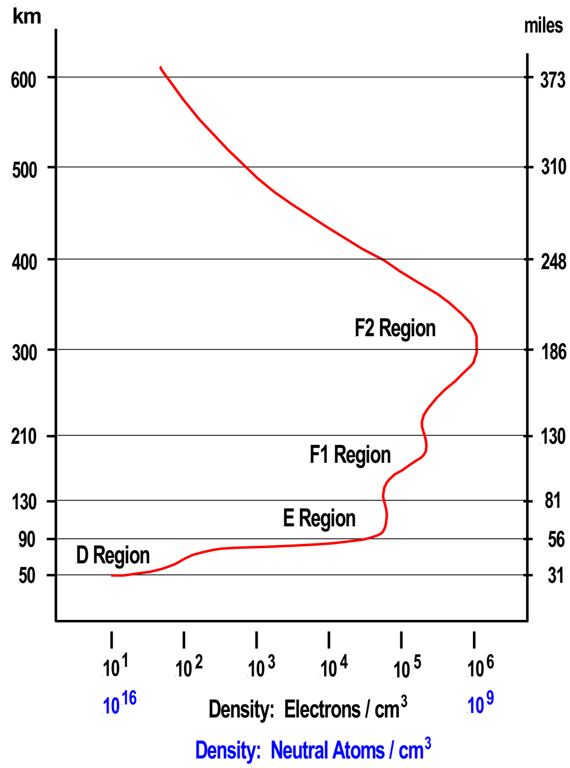


Figure 10 Ionosphere electron density profile (source: author)

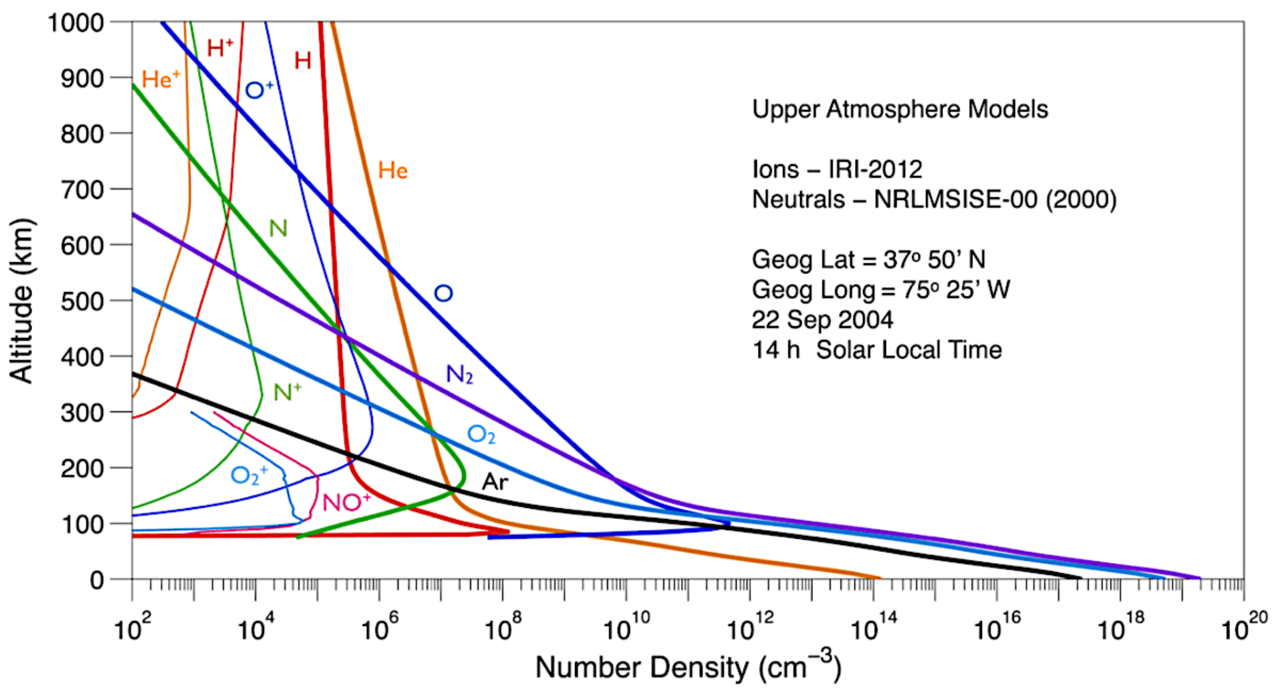


Figure 11 Structure of the ionosphere and upper atmosphere (source: angeo.copernicus.org)

Most of the atmosphere below 100 km is composed of molecular nitrogen N_2 and oxygen O_2 along with small quantities of water vapor, carbon dioxide, and nitric oxide. In Figure 11 at an altitude of 100 km, the concentration of both N_2 and O_2 is around 10^{12} molecules per cm^3. Consequently, the highest concentration of ions from 50 to about 150 km are molecular ions N_2^+, O_2^+, and NO^+. Ions in this region are produced by the molecular photo-ionization process. Recombination in the region also occurs at the molecular level through dissociative recombination as well as electron attachment. Recombination occurs quickly below 90 km resulting in relatively low ion and electron densities from 50 to 90 km.

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 solar EUV radiation begins to occur around 100 km and accelerates quickly as altitude increases. Subsequent photo-ionization of individual dissociated atoms, coupled with decreasing rates of recombination, results in ionization levels rising quickly. Because of dissociation, the ionosphere above 300 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^+. The profile for monatomic nitrogen ions N^+ is similar to that of O^+ but its density is about two orders of magnitude less. Monatomic hydrogen H^+ and helium He^+ ions become relevant in the region above 400 km.

11.2.1 Chapman Layer Theory

EUV and x-ray radiation is intense at the top of Earth's atmosphere, but there are few atoms or molecules there to ionize (Figure 12). As the radiation penetrates deeper into the atmosphere, it encounters more and more atoms and molecules to ionize resulting in higher levels of ionization. However, the ionization process absorbs energy from the EUV and X-ray radiation causing it to weaken. The density of atoms and molecules continues to increase as the radiation penetrates further down into the atmosphere. However, even though there are more atoms and molecules to ionize, the levels of ionization actually decrease due to the weaking radiation. Eventually the level of ionization drops to zero and disappears. Consequently, the most intense ionization occurs toward the middle of the ionosphere (Figure 12) with levels of ionization dropping to zero at both the top and bottom of the ionosphere.

In 1931 British mathematician and geophysicist Sidney Chapman expressed this phenomenon mathematically, using some simplifying assumption to develop what has become known as the Chapman theory. The theory predicts the vertical shape of a simple ionospheric layer and how it varies during the day.

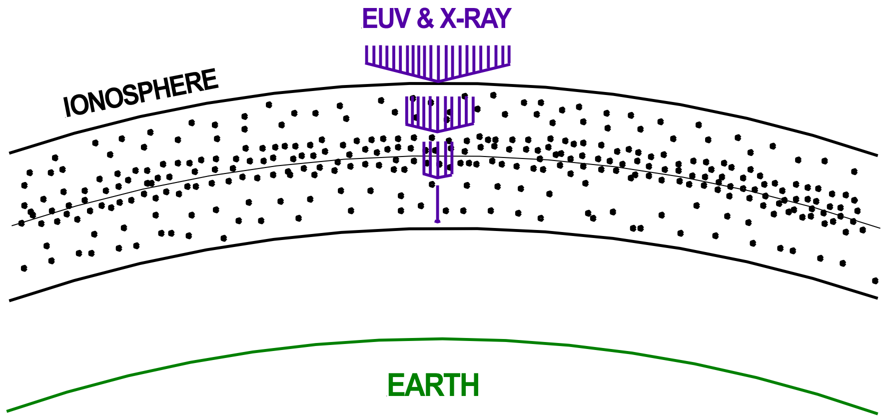


Figure 12 Formation of the ionosphere (source: author)

Chapman's equation describing the rate q of ion-electron pair production is

$$q = q_{m0} \cdot e^{[1-z-(\sec\chi)e^{-z}]}$$

where

χ = solar zenith angle (χ = 0 at noon when Sun is over head, ~ 90° at sunrise and sunset)

q_{mo} = maximum ion-electron pair production rate at $\chi = 0$

z = reduced height [$z = (h - h_{m0}) / H$]

h_{m0} = height of maximum rate of production (q_{mo}) at $\chi = 0$

h = height of rate of production q

H = scale height = kT / mg

k = Boltzmann's constant

T = absolute gas temperature

m = atomic or molecular mass as appropriate

g = acceleration of gravity.

Reduced height [$z = (h - h_{m0}) / H$] is the normalized difference between the height of interest h and the height (h_{m0}) at which the maximum rate of production occurs. The normalizing factor is the scale height H = kT / mg. The scale height is the distance over which the density of ion – electron pairs falls by a factor of 1/e where e = 2.718. In an ion – electron plasma, the mass m is equal to ½ the mass of the positive ion. This is because the mass of an electron is essentially zero in comparison to the ion mass. Each type of ion has its own unique scale height because its mass m appears in the scale height equation. In fact, its mass is the only variable in the equation.

Figure 13 illustrates the general concept of the Chapman theory. In this figure the strength of ionizing radiation is maximum at high altitudes (in the upper right corner of the figure) and decreases at progressively lower altitudes. In contrast, the concentration of atomic and molecular gas capable of being ionized is minimum at high altitudes increasing at progressively lower altitudes to a maximum in the lower right-hand corner of the figure. The same concept as illustrated in Figure 12. At some altitude (h_m) the strength of the ionizing radiation and the concentration of gas is such that maximum ionization occurs (q_m). The rate of ion – electron production falls off rather rapidly at altitudes h above and below h_m, in accordance with Chapman's equation, forming a Chapman ionization layer, the dashed curve in Figure 13.

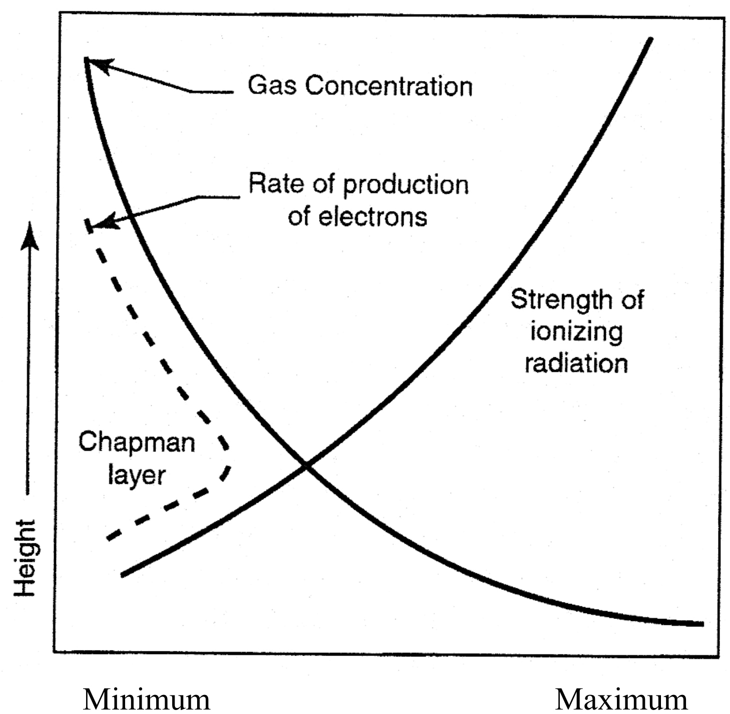


Figure 13 Formation of a Chapman layer (source: Goodman)

Figure 14 shows Chapman layer curves for various zenith angles χ. Notice that the maximum ion – electron production rate q_m is greatest at $\chi = 0°$. That is, at local noon $q_m = q_{mo}$. Also at noon q_m occurs at the lowest altitude h_{m0}. As the zenith angle increases, the maximum ion – electron production rate q_m becomes less than q_{m0} and occurs at a higher altitude h_m then h_{m0}.

In summary

- At noon $\chi = 0^o$
  - $q_m = q_{mo}$
  - $h_m = h_{m0}$
- At $\chi \neq 0^o$
  - $q_m < q_{mo}$
  - $h_m > h_{m0}$

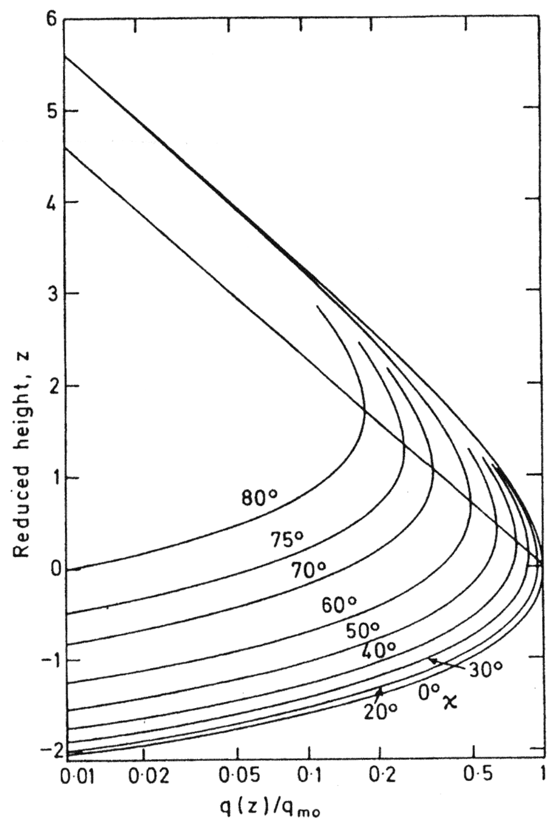


Figure 14 Chapman layer curves for various zenith angles χ (source: Hunsucker & Hargreaves)

The D, E, and F1 regions each exhibit characteristics similar to a Chapman layer. The F2 region does not.

11.2.2 Continuity Equation

The time rate of change in electron density at any given altitude h is described by the continuity equation

$$\frac{dN_e}{dt} = q - L(N_e) - div(N_e\, V)$$

where

N_e = electron density

q = electron production rate due primarily to photoionization

L = electron loss rate through recombination

$div(N_e\, V)$ = electron loss rate due to diffusion

div = the vector divergence operator

V = electron drift velocity

In the ionosphere electron drift in the vertical direction is far greater than horizontal drift, allowing the divergence operator to be replaced by d/dh. Making this simplification the continuity equation becomes

$$\frac{dN_e}{dt} = q - L(N_e) - \frac{d(N_e\, V_h)}{dh}$$

where

V_h = electron drift velocity in the vertical direction

This equation states that over time the change in electron density (N_e) at any given altitude h is equal to

- The rate of electron production (q)
- Minus electron loss due to recombination (L)
- Minus electron loss due to electrons diffusing to a higher or lower altitude $d(N_eV)/dh$

Dissociative recombination and electron attachment occurs quickly in the relatively dense D and E regions of the lower ionosphere. Recombination in the D region typically occurs in a small fraction of a second while in the E region recombination can take up to 20 seconds to occur. In this part of the ionosphere recombination loss is much greater than loss resulting from diffusion.

That is

$$L(N_e) \gg \frac{d(N_e\,V_h)}{dh} \qquad \textit{in the lower ionosphere}$$

At night the rate of electron production q equals zero since there is no ionizing radiate available from the Sun. Assuming that

$$\frac{d(N_e\,V_h)}{dh} \approx 0$$

then in the D and E regions of the ionosphere at night

$$\frac{dN_e}{dt} = -L(N_e)$$

That is, the electron density continuously decreases through the night becoming zero in the D region and decreasing to a very low level just before sunrise in the E region.

During the day in the D and E regions, when ionizing solar radiation is available,

$$\frac{dN_e}{dt} = q - L(N_e)$$

Ionizing solar radiation responsible for ion – electron production (q) increases from zero just before sunrise, peaks at noon, and then decreases reaching zero again after sunset. Consequently, in the E region ionization (N_e) increases from a minimum before sunrise, peak around noon, and then drop back to a minimum shortly after sunset as illustrated by the red curve in Figure 15. The slope of the red ionization curve is dN_e/dt which, according to Figure 15, is zero around local noon. If

$$\frac{dN_e}{dt} = 0 \qquad then \qquad L(N_e) = q$$

indicating that a state of equilibrium is reached around noon in which the rate of electron production q equals the rate of electron loss through recombination $L(N_e)$.

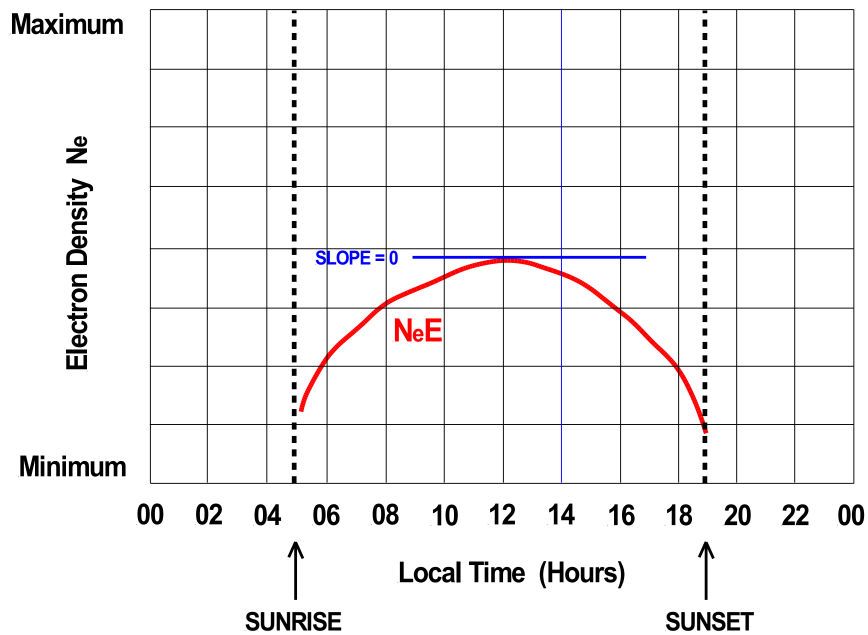


Figure 15 E Region diurnal electron density (source: author)

Radiative recombination is the primary mechanism for electron loss in the upper regions of the ionosphere since there are few molecules at these altitudes to support dissociative recombination and electron attachment. In the upper ionosphere few molecules are present since most of them have been dissociated into individual neutral atoms by solar EUV radiation. As described earlier, radiative recombination occurs very slowly compared to dissociative recombination and electron attachment. In addition, the neutral atmosphere above 500 km is nearly collision free with atoms, ions and electrons able to move about freely without colliding. In this environment, recombination typically takes hours to occur following photoionization of a neutral atom. The slow recombination rate is the primary reason that the F region does not completely disappear at night.

Electron – ion vertical drift becomes very important in the upper regions of the ionosphere due to extremely low atmospheric density. That is

$$L(N_e) \ll \frac{d(N_e\, V_h)}{dh} \qquad \textit{in the upper ionosphere}$$

The plasmasphere is the second reason that the F region does not completely disappear at night. The plasmasphere is a region of cold dense plasma trapped by closed magnetic field lines in the inner part of the magnetosphere (Figure 16). The lower part of the plasmasphere is closely linked to the mid-latitude ionosphere acting as a reservoir of charged particles from the ionosphere. During the day free electrons drift upward along magnetic field lines into the plasmasphere. This upward drift constitutes part of the diffusion loss $d(N_e\ V_h)/dh$ in the continuity equation

$$\frac{dN_e}{dt} = q - L(N_e) - \frac{d(N_e\, V_h)}{dh}$$

At night electrons gradually diffuse back down into the upper ionosphere. In this case the diffusion term $d(N_e\ V_h)/dh$ is not a loss but a gain. (See the chapter on Earth's Magnetosphere for more details on the plasmasphere.)

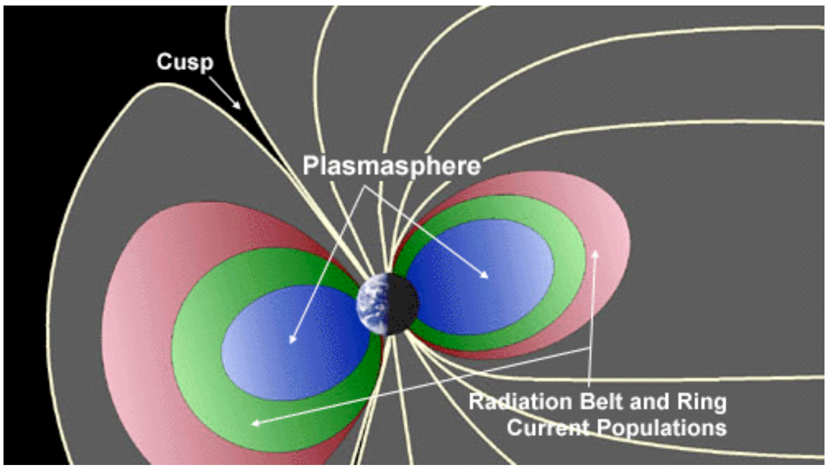


Figure 16 Plasmasphere (source: Alchetron)

11.2.3 Depth of Solar Energy Penetration

The depth that solar energy penetrates into the atmosphere varies with wavelength (Figure 17). The level of penetration is the altitude at which radiation of a particular wavelength disappears, having been completely absorbed by atoms and molecules further up in the atmosphere. For example, relatively long wavelength visible sunlight penetrates down to Earth's surface, but shorter wavelength medium UV (ultra-violate) radiation only penetrates to the stratosphere. It never makes it to Earth's surface. Consequently, Earth's atmosphere protects us from medium UV and shorter wavelength radiation, but not from the longer wavelength near UV. Near UV does make it to ground level causing skin cancer and other biological problems.

In studying the ionosphere, we are particularly interested in extreme ultra-violate (EUV) radiation. It is important to note that EUV radiation is not a single wavelength but instead a range or band of wavelengths from 10 to 200 nm, just as HF radio frequencies form a band of wavelengths from 10 to 80 meters. The interaction of 10 meter radio waves with the ionosphere is different than that of 80 meters. The same is true of EUV radiation. The interaction of short wavelength EUV with the atmosphere is different than that of longer wavelength EUV.

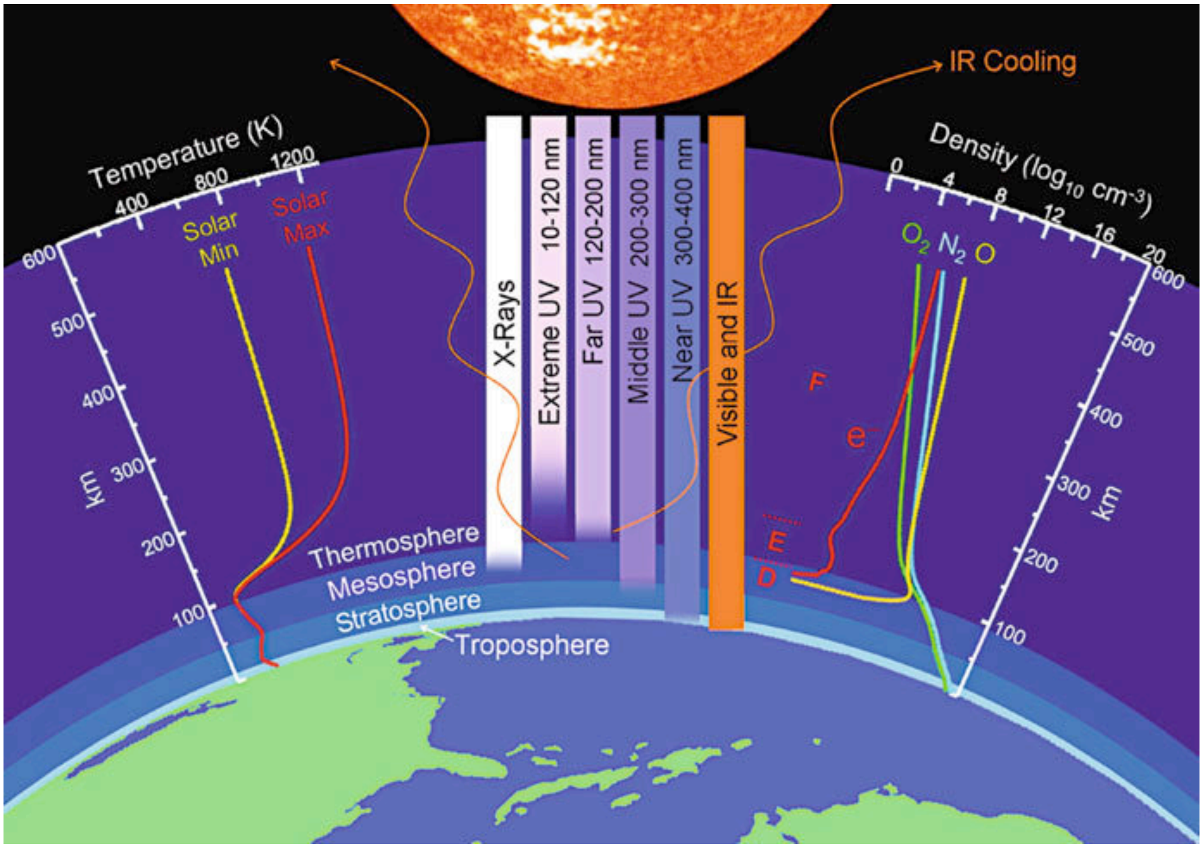


Figure 17 Depth of solar radiation penetration (source: ResearchGate)

In general, short wavelength solar radiation is absorbed by monatomic oxygen, nitrogen, hydrogen, and helium atoms high up in the ionosphere whereas longer wavelength radiation is absorbed by molecules in the lower ionosphere. Radiation with even longer wavelengths are absorbed in the lower part of the atmosphere or in some cases makes it all the way to ground level. Atomic oxygen, nitrogen, hydrogen, and helium all have ionization maximum wavelength limits ($\boldsymbol{\lambda_{max}}$) of less than 92 nm, as shown in Table 1 below. As discussed earlier, these atoms will only be ionized by EUV photons with wavelengths of less than 92 nm. However, they are incapable of being ionized by longer wavelength radiation. Instead, long wavelength EUV radiation passes through the high regions of the ionosphere before ionizing molecules in the lower ionosphere, primarily nitric oxide (NO) and oxygen (O_2) which have maximum wavelength limits on the order of 102 – 134 nm. An exception to this trend is x-ray radiation with wavelengths in the range of 0.1 – 17 nm. This very short wavelength radiation should be completely absorbed high up in the atmosphere. However, the intensity of x-ray radiation during a solar flare is so great that it penetrates all the way down to an altitude of around 50 km before disappearing, heavily ionizing the D region in the process.

11.3 Regions of the Ionosphere

There are no discrete or abrupt boundaries between the ionospheric regions shown in Figure 18. One region flows smoothly into the next.

Species	Ionization potential (ev)	λ_{max} (nm)
NO	9.25	134.0
O_2	12.08	102.7
H_2O	12.60	98.5
O_3	12.80	97.0
H	13.59	91.2
O	13.61	91.1
CO_2	13.79	89.9
N	14.54	85.3
H_2	15.41	80.4
N_2	15.58	79.6
A	15.75	78.7
Ne	21.56	57.5
He	24.58	50.4

Table 1 Repeated - Ionization Potentials of Atmospheric Gases

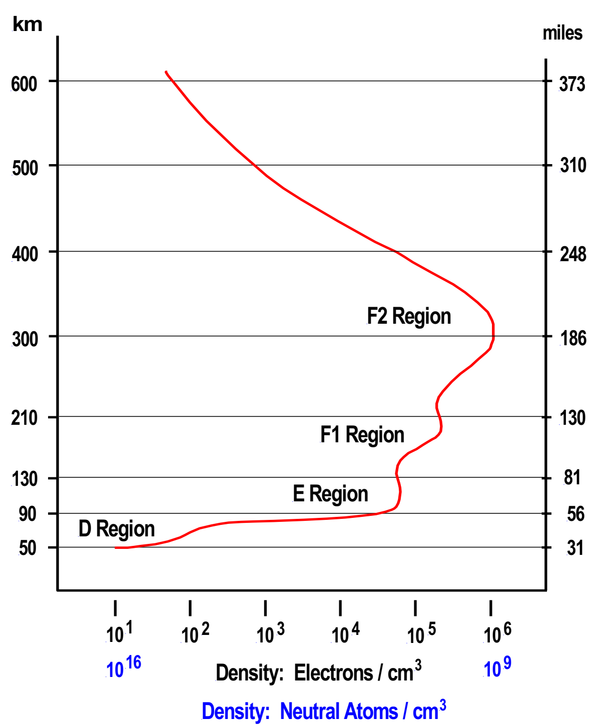


Figure 18 Regions of the ionosphere (source: author)

Notice that the electron density tends to peak in the middle of each region as explained by Chapman's theory and illustrated in Figure 13. The electron density profile shown in Figure 18 is for the daytime. Figure 19 illustrates the diurnal and solar cycle variations in ionospheric electron densities. The dashed lines represent the ionosphere during solar minimum. The solid lines correspond to solar maximum. At night the D region disappears while the F1 and F2 regions combine into a single F region. The electron densities for the E and F regions are both less at night than during the day as explained by the Chapman and continuity equations. The lowest E and F densities occur just before sunrise. Ionospheric ionization is weaker during solar minimum than at solar maximum. This is to be expected since x-ray and EUV radiation levels are low at solar minimum compared to solar maximum. Note that electron density in Figure 19 is given in cubic meters (m^3). Densities can be converted to more commonly used cubic centimeters (cm^3) by subtracting 6 from the cubic meter exponents. Thus 10^{13} cubic meters corresponds to 10^7 cubic centimeters.

Characteristics of the individual D, E, F1, and F2 regions are driven by the various types of molecules, atoms, and ions present in the upper atmosphere and their vertical distributions shown in Figure 20.

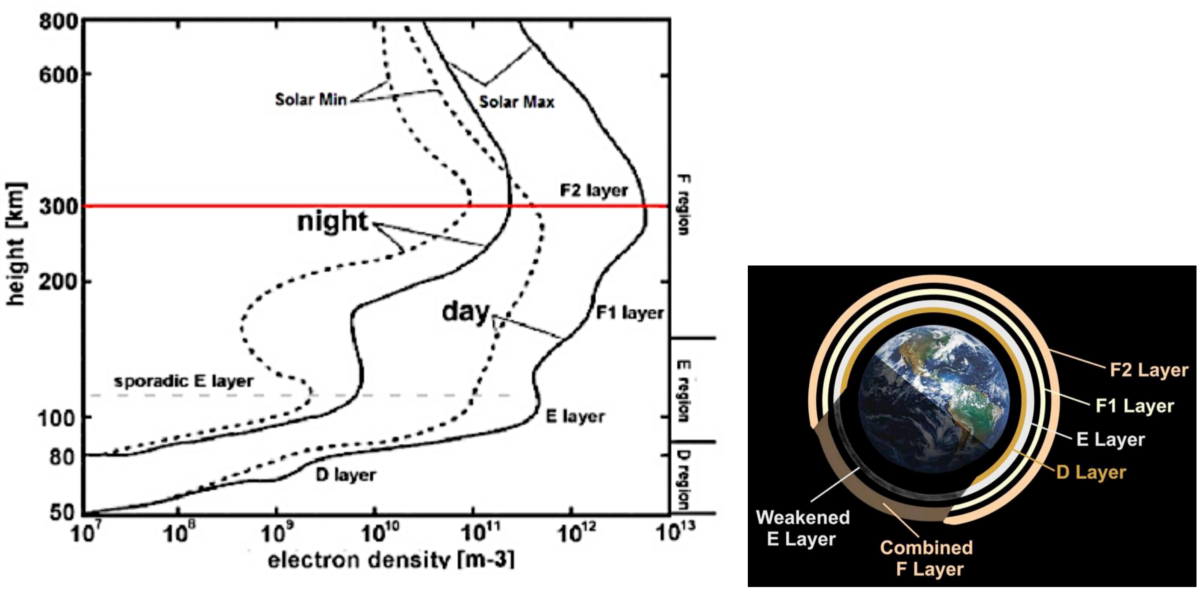


Figure 19 Diurnal and solar cycle variations in the ionosphere (source: ResearchGate)

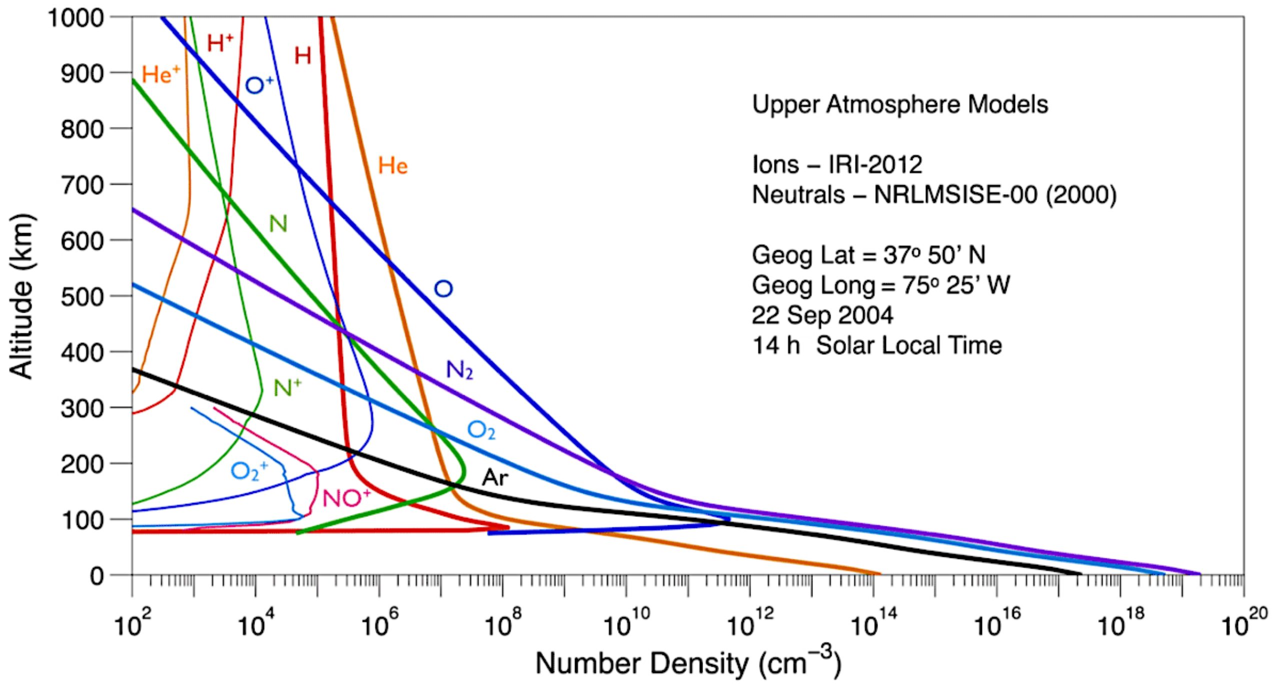


Figure 20 Structure of the ionosphere and upper atmosphere (source: angeo.copernicus.org)

11.3.1 D Region

The D region is the lowest region of the ionosphere located at an altitude from about 50 to 90 km. The D region is important because it absorbs medium frequency (MF) signals typically used by AM broadcast stations, high frequency (HF 3 to 30 MHz) signals, and very high frequency (VHF > 30 MHz) radio signals while reflecting low frequency (LF) and very low frequency (VLF) signals. Lyman-α, EUV, and X-ray radiation ionize D region nitric oxide NO and molecular oxygen O_2 producing NO^+ and O_2^+ ions. Ionization of the D region is considerably enhanced by intense solar x-ray radiation following a solar flare. The x-ray radiation, with its very short 0.01 to 10 nm wavelength, ionizes all species of molecules found in this part of the atmosphere, particular molecular oxygen (O_2) and nitrogen (N_2) forming O_2^+ and N_2^+ ions. However, N_2^+ rapidly converts to O_2^+ by the charge exchange reaction

$$N_2^+ + O_2 \rightarrow O_2^+ + N_2$$

leaving NO^+ and O_2^+ as the primary ionization products. Recombination occurs quickly in this relatively dense region of the upper atmosphere, so electron concentrations are low, about 10^2 to 10^3 electrons per cm^3.

As shown in Figures 21, Lyman-α spectral lines are formed by the electron of a hydrogen atom dropping back to its ground state (n = 1) from higher energy levels. The Lyman-α line is in the ultraviolet part of the solar spectrum at a wavelength of 121.5 nm. Hydrogen atoms responsible for producing Lyman-α radiation are found in the transition region between the Sun's photosphere and chromosphere where temperatures are rapidly increasing.

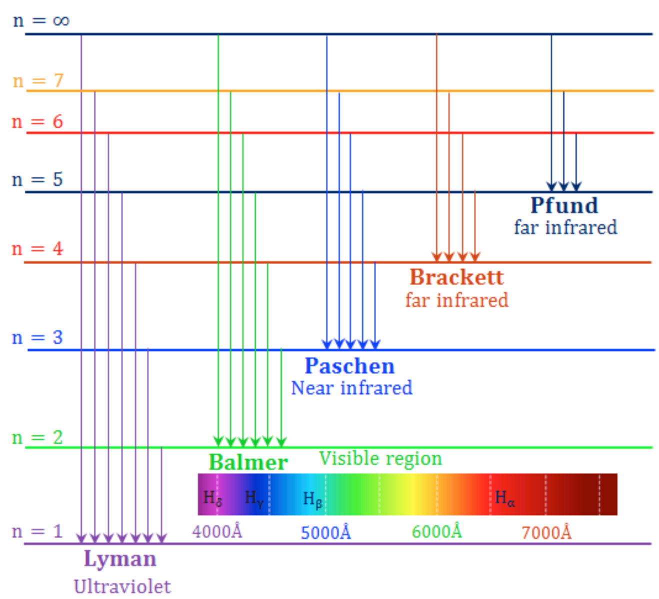


Figure 21 Hydrogen atom spectral series (source: Pinterest)

Lyman-α radiation penetrates below 90 km and ionizes nitric oxide (NO) whose ionization limit is 134.0 nm (see Table 1). This is the main source of D region ionization at middle latitudes. To a lesser extent, EUV energy at wavelengths between 102.7 and 111.8 nm ionizes oxygen molecules (O_2) that are in an excited state. Because of its very short wavelength (0.01 to 10 nm) x-ray radiation ionizes all species of gases in the D region, particularly following a solar flare. X-ray radiation varies considerably over the solar cycle changing in intensity by a factor of a hundred to a thousand. X-ray radiation is at its lowest level during solar minimum. In fact, x-ray radiation is probably not a significant factor in D region ionization during solar minimum. At Earth's high latitudes charged particles from the solar wind and magnetosphere ionize the D region to various extents. At times they are the dominate source of D layer ionization in the polar regions.

Dissociative recombination is the primary ion loss process in the D region with the recombination of NO^+ being somewhat faster than O_2^+. Electron attachment recombination becomes significant in the lower part of the D region where atmospheric densities are greater.

The D region accounts for most of the HF radio wave absorption in the ionosphere. Absorption generally varies with the solar zenith angle χ in accordance with

$$Absorption \;\propto (\cos\chi)^n$$

where n is in the range $0.7 < n < 1.0$.

An anomaly occurs in D region ionization. Ionization is greater in the winter by a factor of 2 to 3. In winter χ is larger than in the summer leading to the expectation that the degree of ionization would be smaller in winter than in the summer, but that is not the case. In addition, ionization is much more variable from day to day in the winter. The reason for this is that the ionosphere is cooler during the winter retarding the rate of recombination to some degree.

11.3.2 E Region

The E region is located at an altitude of 90 to 130 km. Its peak electron density of about 10^5 electrons per cm^3 occurs during the day at a height of 105 to 110 km. E region ionization reaches its lowest level just before sunrise. It then increases rapidly, reaches a peak near noon, and then steadily declines. At night a weakly ionized layer remains with a density of around 5 x 10^3 electrons per cm^3. Ionization by meteor debris is one possible cause for night time ionization.

The E region is formed primarily by EUV radiation in the range from 80 to 103 nm which ionizes molecular oxygen O_2 to form O_2^+ ions. The ionization limit for O_2 is 102.7 nm meaning that O_2 can be ionized by radiation with wavelengths shorter than 102.7 nm (for example 90 nm) but not radiation with wavelengths longer than 102.7. Ionization of nitric oxide (NO) by Lyman-α radiation forming NO^+ ions also occurs in the E region. X-ray radiation ionizes all species of gas in the E region following a solar flare producing NO^+, O_2^+, O^+, and N_2^+ ions with the most numerous being NO^+ and O_2^+. X-ray radiation contributes little to E region ionization during solar minimum when the intensity of X-ray radiation is generally low.

Recombination in the E region occurs by both direct radiative recombination

$$e + X^+ \rightarrow X + a\ photon\ with\ energy\ hf \qquad (h = Planck's\ \ constant, f\ = frequency)$$

and dissociative recombination

$$e + XY^+ \rightarrow X + Y$$

However, dissociative recombination occurs 100,000 times faster than direct radiative recombination. Consequently, radiative recombination is relatively insignificant in the E region. The most common recombination reactions in the E region are

$$e + O_2^+ \rightarrow O + O$$

$$e + N_2^+ \rightarrow N + N$$

$$e + NO^+ \rightarrow N + O$$

11.3.3 Sporadic E Formations

Sporadic E (E_s) formations are zones of abnormally high ionization within the E region (Figure 22). E_s zones are important because they can reflect HF radio signals at frequencies up to about 100 MHz. They are called 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. In general, sporadic E appearances seem to have little direct relationship to the ionization processes responsible for the E region itself.

Sporadic E zones often have electron densities far greater than normal E region levels and at times even greater than in the F region. Sporadic E patches can appear opaque to radio waves, reflecting waves that normally would have been refracted at the much higher F2 layer. 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 a destination as illustrated on the left in Figure 23. This can seriously degrade signal levels at the receiving site. Worse yet, the intended receiving location could be missed altogether as occurs on the right in Figure 23. At other times the F1 and F2 regions can be seen through the E_s layer 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.

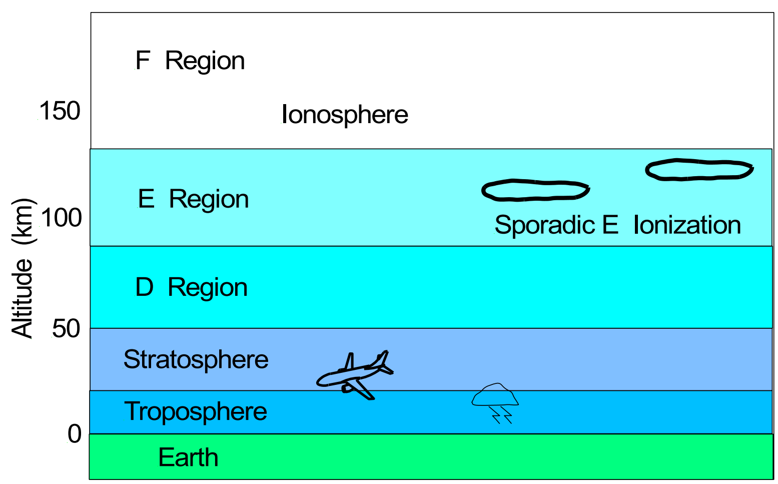


Figure 22 Sporadic E ionization (source: author)

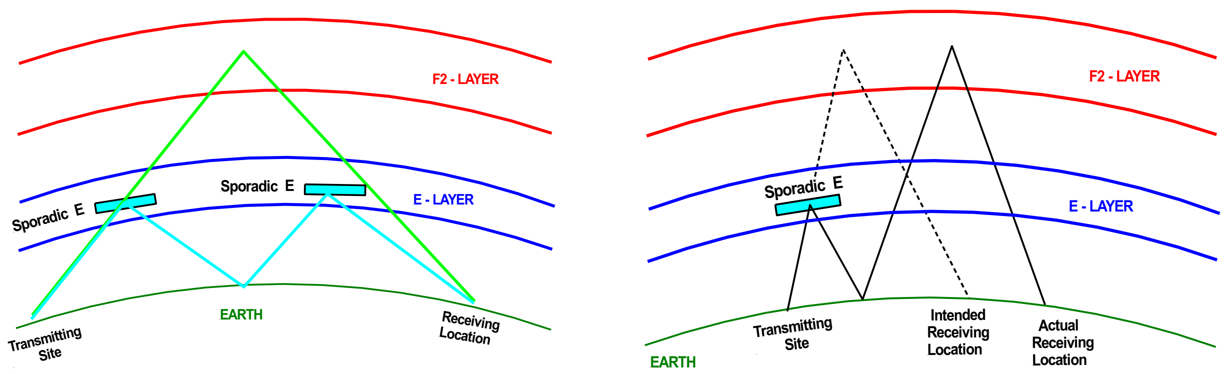


Figure 23 Disruptions in propagations paths due to sporadic E (source: author)

Sporadic E zones are particularly strong in Earth's low latitude equatorial region where they are essentially a daytime phenomenon with little seasonal variation. It is believed that they are formed in this part of the world by instabilities in the equatorial electrojet.

At mid latitudes sporadic E zones tend to be weaker than elsewhere. Their occurrence is subject to diurnal and seasonal variations. They tend to be more prevalent during the summer than in winter and during the day than at night, particularly in mid-morning and near sunset. It is believed that sporadic E patches at mid latitudes form as the result of wind shear in the upper atmosphere in combination with meteoric debris. Enormous numbers of meteors burn up in the E region of the atmosphere (Figure 24). 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 lower than for molecular ions.

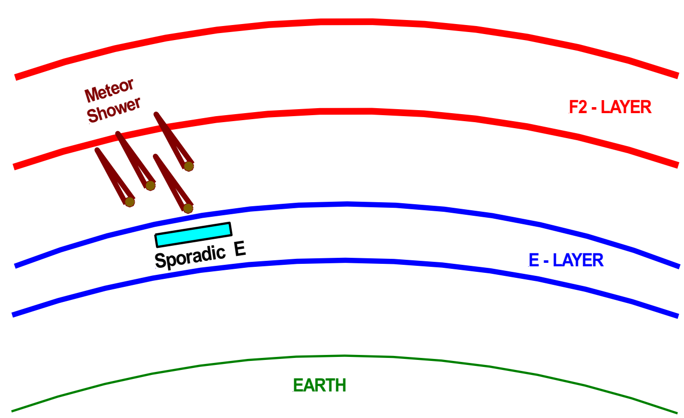


Figure 24 Meteor showers creating sporadic E patches (source: author)

Atmospheric gravity waves, associated with traveling ionospheric disturbances (TIDs), create upper atmosphere high velocity winds that travel in opposite directions at slightly different altitudes

(Figure 25). This set of conditions produce what is 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 develop in these pockets producing sporadic E patches.

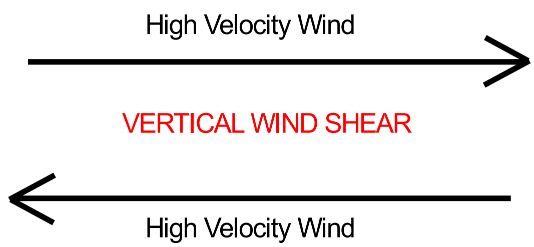


Figure 25 Vertical wind shear (source: author)

At high latitudes sporadic E zones occur mainly at night with little seasonal variance. They are attributed to ionization by incoming high energy charged particles entering the polar region from the magnetosphere. Clouds of auroral E_s drift westward in the evening and eastward in the early morning at speeds between 200 and 3,000 meters per second, much like the auroral itself. Sporadic E zones within the polar caps have a different characteristic. They are weaker and extend across the polar caps in the form of ribbons in a roughly sunward direction.

11.3.4 F1 Region

The F1 region (Figure 26) extends in altitude from about 130 to 210 km. The electron density in the F1 region peaks around local noon at an altitude in the neighborhood of 180 km. At its peak, the electron density is approximately 2.5 x 10^5 electrons per cm^3. The F1 region disappears at night merging into the F2 region. It also merges with the F2 region in the winter months during solar maximum.

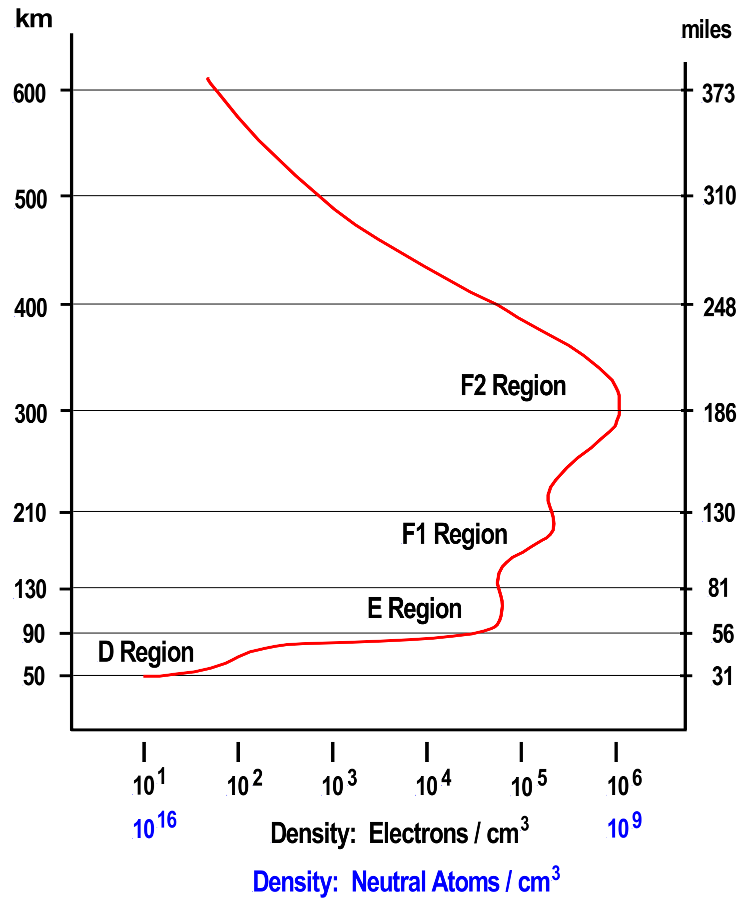


Figure 26 Regions of the ionosphere (source: author)

The F1 region is formed by EUV radiation in the range from 20 to 90 nm. This radiation strongly ionizes individual oxygen atoms (ionization limit 91.1 nm) producing O^+ ions.

Dissociative recombination is the primary mechanism for recombination in the F1 region. Dissociative recombination for an atomic oxygen ion O^+ is a two-step process. The oxygen ion must first be converted to a molecular ion through charge transfer. The two most common charge transfer reactions are

$$O^+ + O_2 \rightarrow O_2^+ + O$$

$$O^+ + N_2 \rightarrow NO^+ + N$$

The recombination process is then completed through the appropriate dissociative recombination reaction

$$e + O_2^+ \rightarrow O + O$$

$$e + NO^+ \rightarrow N + O$$

11.3.5 F2 Region

The F2 region (Figure 26) has a greater concentration of free electrons than any of the other regions. It is thus the most important in terms of radio wave propagation. Unfortunately, it is also the most variable and most unpredictable. It is the largest region of the ionosphere stretching in altitude from 200 to over 800 km. The peak electron density is on the order of 10^6 electrons per cm^3 which occurs during the day at a height of around 300 km. The electron density at night is in the neighborhood of 10^5.

The F2 region is formed by EUV ionization of individual oxygen atoms coupled with upward diffusion of electrons from the F1 region. Ionized hydrogen H^+ gradually replaces O^+ as the dominate ion high up in the F2 region. The ionization potential of O^+ and H^+ are nearly the same with maximum ionizing wavelengths of 91.1 and 91.2 nm respectively. Consequently, in the upper ionosphere the reaction

$$H + O^+ \leftrightarrow H^+ + O$$

goes easily in either direction. Above 700 km H^+ gradually becomes the dominate ion due to is lower mass. That is, the force of gravity is less on hydrogen ions than on oxygen allowing hydrogen to diffuse to much higher altitudes.

The F2 region is quite different from the D, E, and F1 regions. The maximum rate of electron production q in the continuity equation occurs in the F1 region at an altitude around 180 km.

$$\frac{dN_e}{dt} = q - L(N_e) - \frac{d(N_e\,V_h)}{dh}$$

However, the ionosphere's peak electron density N_e occurs roughly 100 km higher in the F2 region at an altitude around 300 km as shown in Figure 26. Both q and L in the continuity equation decrease with altitude above the F1 region. However, the rate of electron loss through recombination L decreases faster than the rate of production q (L becomes small faster than q). In addition, the loss of electrons in the F1 region through diffusion

$$\frac{d(N_e\,V_h)}{dh}$$

becomes noticeable as electrons gradually drifting upward from the F1 into the F2 region. The combination of L decreasing faster than q coupled with the upward drift of electrons results in the ionosphere's electron density N_e continuing to increase with altitude until it peaks at 300 km.

In the rarified atmosphere above 300 km collisions between electrons and ions seldom occur resulting in little loss of electrons through recombination. In the continuity equation

$$L(N_e) \rightarrow 0$$

reducing the equation to

$$\frac{dN_e}{dt} = q - \frac{d(N_e\,V_h)}{dh}$$

Above 300 km diffusion of ions and electrons becomes very important while the rate of electron production q continues to decrease. Gradually the loss of electrons through diffusion becomes greater than the rate of electron production.

$$q \, < \, \frac{d(N_e\,V_h)}{dh}$$

causing the ionosphere electron density N_e to decrease exponentially as shown in Figure 26.

A number of anomalies occur in the F2 region. Anomalies in the sense that the F2 region tends to behave different than the Chapman like D, E and F1 regions. These differences include:

1. Diurnal variations in F2 electron density are often not symmetrical about noon. A rapid change in F2 density occurs at sunrise but then may change little in the afternoon until well after sunset. The daily peak often occurs at noon in the winter but may occur either before or after local noon in the summer during solar maximum. Daily peaks often occur late in the afternoon during solar minimum.

2. The daily variation in F2 characteristics typically does not repeat from one day to the next making predictions of F2 performance difficult.

3. The F2 peak in electron density is usually greater in the winter than in the summer, particularly during solar maximum, the opposite of what is predicted for a Chapmen layer.

4. The mid-latitude F2 region does not disappear at night but remains through the night at a substantial level.

There appear to be three primary causes for the F2 region anomalies:

a) Reaction rates are sensitive to temperature,

b) Chemical composition of the ionosphere varies, and

c) Significant interaction occurs between the plasmasphere and ionosphere.

Reaction rates are generally temperature sensitive. For example, the first step in the dissociative recombination of atomic oxygen ions O^+ is

$$O^+ + N_2 \rightarrow NO^+ + N$$

This reaction varies considerably with the temperature of neutral nitrogen N_2 increasing by a factor of 16 for a temperature change from 1,000 to 4,000 K. [The temperature in the Thermosphere at an altitude of 300 km, where the F2 region resides, is typically 1,200 ^oK – see Figure 9]. Changes in reaction rates contributes to both the F2 seasonal anomaly and persistence of the F region at night. During summer the higher temperature of N_2 increase the rate of dissociative recombination resulting in lower electron concentrations than during the winter. Cooler temperatures at night slows down the recombination process causing the F region to remain at least partially ionized.

The electron production rate in the F2 region depends on the concentration of atomic oxygen O while the recombination rate is controlled by concentrations of N_2 and O_2. An increase in atomic oxygen relative to molecular nitrogen and oxygen increases electron density. Thus, the ratio of atomic oxygen to molecular oxygen O/O_2 and atomic oxygen to molecular nitrogen O/N_2 is important. The ratio O/N_2 at an altitude of 300 km is about 6 in winter and around 2 in the summer, again resulting in higher electron densities in the winter compared to summer.

Variation in ionospheric temperature affects its vertical distribution. Ionospheric heating, in part, results from excess photon energy above that necessary for ionizing neutral atoms and molecules. When a photon collides with a neutral atom or molecule, energy in excess of that needed for ionization is transformed into the kinetic energy, and thus heating, of an electron as it escapes from its parent atom or molecule. Because of this heating the ionosphere is hotter than the surrounding neutral atmosphere and within the ionosphere electrons are hotter than ions. During the day electron temperatures can be two to three times higher than ion temperatures while at night electron and ion temperatures tend to be about the same. These changes in temperature affect the distribution of the F2 region plasma.

Following sunrise, the F region is strengthened and heated by increasing levels of ionization, causing it to drift to higher altitudes where it persists longer due to the declining rate of recombination. As it drifts upward the charge exchange reaction

$$O^+ + H \rightarrow H^+ + O$$

causes hydrogen ions to become the dominate species. High in the ionosphere hydrogen ions (protons) flow upward along magnetic field lines into the plasmasphere increasing the plasmasphere charge density. During the evening and night a portion of the protons flow back down into the ionosphere where they undergo the reverse charge exchange reaction

$$H^+ + O \rightarrow O^+ + H$$

producing the oxygen ions which helps to maintain the F region at night.

The summer upper ionosphere is the primary contributor to increased plasmasphere charge density. However, the flow of protons back down into the ionosphere at night occurs in both the summer and winter ionospheres. Consequently, the summer flow of protons into the plasmasphere from Earth's northern hemisphere helps to maintain the F region during the night in both the southern (winter) as well as in the northern (summer) hemispheres and visa-versa when the seasons change.

11.3.6 Summary of Ionospheric Regions

The characteristics of the various ionospheric regions are summarized below in Table 2.

Region	Height (km)	N_{max} (cm^{-3})	f_c (MHz)	Primary Ions	Basis of Formation
D	50 to 90	~ 1.5×10^4 noon ~ 0 night		NO^+ & O_2^+	Lyman-α (121.5 nm) EUV (102 - 111 nm) x-rays (0.01–15 nm)
E	90 to 130 h_{max} ~ 110	~ 1.5×10^5 noon ~ 5×10^3 night	~ 3 day ~0.3 night	O_2^+ & NO^+	EUV (80 - 103 nm) Lyman-α (121.5 nm) x-rays (0.01–15 nm)
E_s	90 to 130	~ 10^6 (highly variable)		Metallic Ions	Wind shear and meteoric debris
F1	130 - 210 h_{max} ~ 180	~ 2.5×10^5 noon ~ 0 night	~ 3 to 6 day Merges with F2 at night	O^+	EUV (20 – 90 nm)
F2	210 - 800 h_{max} ~ 300	~ 10^6 day ~ 10^5 night	~ 5 - 15 day ~ 2 - 6 night	O^+	EUV (20 – 90 nm) + upward diffusion from F1 region

Table 2 Summary of Region Characteristics (source: adapted from McNamara)

11.4 Ionospheric Irregularities

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 electron formations 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. Electrons can travel along magnetic field lines but can not across them. Consequently, irregularities in electron densities become stretched out along magnetic field lines, forming what is known as field-aligned irregularities (FAI).

Ionospheric irregularities include:

- Spread F Irregularities,
- Traveling Ionospheric Disturbances, and
- Thermospheric Winds.

11.4.1 Spread F Irregularities

Signal scattering due to field-aligned irregularities in electron densities result in a F region phenomenon known as spread F. Spread F is often noticed in the echoes of ionogram pulses. In some cases, the echoed pulse is up to 10 times wider than the transmitted pulse. This type of distortion is known as **range spread**. In other cases, **frequency spread** distorts critical frequencies so that they are no longer single frequencies but instead a bands of frequencies.

Spread F is important because it causes radio signals to be seriously distorted, limiting the data rate of signals that can be successfully transmitted. In addition, spread F produces high fading rates.

While the field aligned irregularities producing spread F are highly variable in size, they can range from roughly a 100 km to several thousand kilometers wide and about a kilometer thick. At Earth's high latitudes, these irregularities drift horizontally at speeds up to 100 meters per second.

In the equatorial region spread F irregularities appear 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. Spread F can occur at any time but tends to be more intense and more numerous at solar maximum, during the equinoxes, and on most nights from November through January. Equatorial spread F disappears with the onset of a magnetic storm.

At high latitudes, spread F occurrences beginning around 40^{o} and becomes more pronounced with increasing latitude. Irregularities, including spread F, contribute to the extreme variability of the high latitude F region. In both the auroral and polar cap zones increased geomagnetic activity dramatically increases the appearance and growth of spread F irregularities. In addition, spread F formations, with electron concentrations several orders of magnitude greater than normal, frequently drift into the polar cap from the auroral zones, particularly with elevated solar activity. During the summer, near the magnetic dip poles, spread F is pretty much a permanent phenomenon, occurring nearly every night and often during the day. In the winter spread F is nearly continuous, both day and night.

There is little spread F at geomagnetic latitudes from about 20° to 40°.

11.4.2 Traveling Ionospheric Disturbances

Traveling ionospheric disturbances (TIDs), illustrated in Figure 27, 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 from 50 to about 1000 meters per second. TIDs have long wave fronts that are tilted forward. Consequently, TIDs appear first at high altitudes and move downward as they pass.

Vary large TIDs originate in the auroral zone and travel great distances. Their wavelengths are typically around 1000 km with periods of an hour or more. These disturbances are associated with magnetic storms.

Medium scale TIDs have wavelengths on the order of 100 to 200 km with periods ranging from 5 to 45 minutes. These TIDs do not travel far and may be caused by local tropospheric weather conditions such as hurricanes and tornadoes.

HF radio signals encountering traveling ionospheric disturbances are seriously disrupted. Nearly all aspects of an HF signal are distorted including signal frequency, amplitude, phase, and polarization.

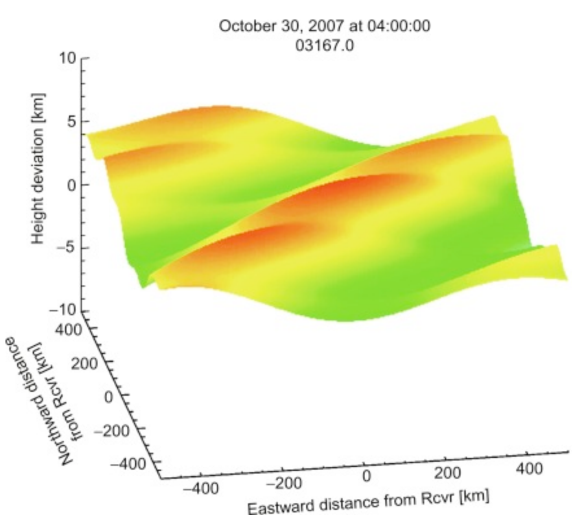


Figure 27 Traveling Ionospheric Disturbance (source: ScienceDirect.com)

11.4.3 Thermospheric Winds

The thermosphere is the region of Earth's atmosphere above 100 km in which the ionosphere is embedded. Winds blow in the thermosphere as they do elsewhere in the atmosphere. As they blow, they blow the wispy ionosphere along with them. Thus, the nature of the ionosphere at a specific time and location depends in part on the thermospheric winds that happen to be blowing at the time. Consequently, day to day changes in thermospheric winds is one cause for day to day changes in the ionosphere and critical frequencies.

References

Hunsucker R. D.; Hargreaves, J. K.; "The High-Latitude Ionosphere and its Effects on Radio Propagation"; Cambridge University Press 2003

Davies, Kenneth; "Ionospheric Radio"; Peter Peregrinus Ltd.,1990

McNamara, Leo F.; "The Ionosphere: Communications, Surveillance, and Direction Finding"; Krieger Publishing Company, 1991

Bridgman, Tom; "PlasmaZoo: E-cross-B Drift"; NASA Visualizations, https://svs.gsfc.nasa.gov/4265 ; February 2, 2015

Casey, John P; "Overview of the Equatorial Electrojet and Related Ionospheric Current Systems"; NUWC-NPT Technical Report 11, 676 25 April 2005

Nichols, Eric P.; "Propagation and Radio Science"; The American Radio Relay League, Inc. 2015

Yeang, Chen-Pang; "Probing The Sky With Radio Waves"; The University of Chicago Press, 2013

Devoldere, John; "Low-Band DXing" fourth edition; ARRL, 2005

Levis, Curt A. ; Johnson, Joel T.; and Teixeira, Fernando L.; "Radiowave Propagation Physics and Applications"; John Wiley & Sons, Inc., 2010

Ahrens, C. Donald; "Essentials of Meteorology"; Wadsworth Publishing Company, 1998

UCAR Center for Science Education (UCAR SciEd); https://scied.ucar.edu/learning-zone/atmosphere/

Khazanov, George V.; "Space Weather Fundamentals"; CRC Press, 2016

Foukal, Peter; "Solar Astrophysics third edition"; Whiley-VCH Publishing Company, 2013

Ganushkina, N. Yu.; Liemohn, M. W.; Dubyagin. S.; "Current Systems in the Earth's Magnetosphere"; AGU, March 8, 2018

Gallagher, Dr. D.L.; "The Earth's Plasmasphere"; Space Plasma Physics, Marshall Space Flight Center, Huntsville, Al, September 05, 2018

Moore, T. E., Morwitz, J. L.; "Stellar Ablation of Planetary Atmospheres"; August 9, 2007

Yau, Andrew W.; Abe, Takumi; Peterson, W. K.; "The Polar Wind: recent Observations"; Department of Physics and Astronomy, University of Calgary

Carroll, Bradley W. and Ostlie, Dale A.; “An Introduction to Modern Astrophysics”; Addison-Wesley Publishing Company Inc., 1996

Goodman, John M.; “Space Weather & Telecommunications”; Springer Science+Business Media Inc. 2005

Cander, Ljiljana R.; “Ionospheric Space Weather”; Springer Nature Switzerland AG 2019

Moldwin, Mark; “An Introduction to Space Weather”; Cambridge University Press, 2008

Campbell, Wallace H.; “Introduction to Geomagnetic Fields”; Cambridge University Press, 2003

Golub, Leon and Pasachoff, Jay M.; “Nearest Star The Surprising Science of Our Sun second edition”; Cambridge University Press, 2014

Loff, Sarah: ”Explorer and Early Satellites”; National Aeronautics and Space Administration, Aug 3, 2017

Minzner, R. A.; “The 1976 Standard Atmosphere Above 86 km Altitude” NASA Goddard Space Flight Center, 1976