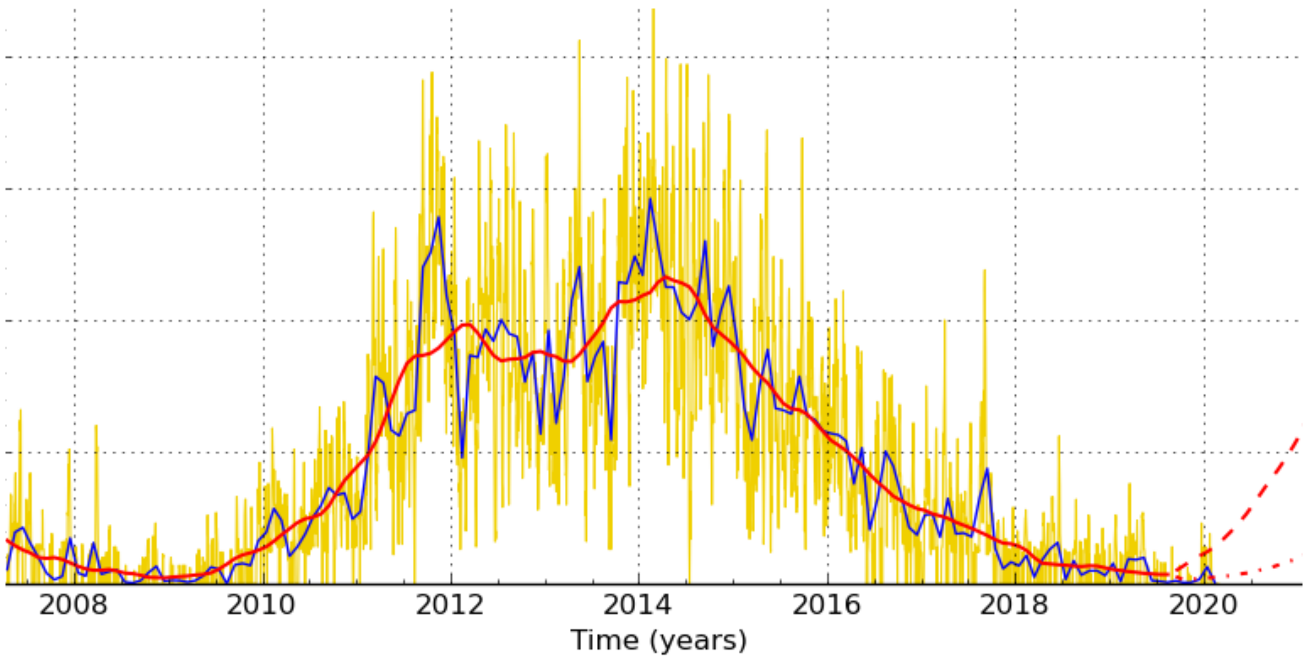
Chapter 4

The Sun's Solar Cycle



www.skywave-radio.org

4 The Sun's Solar Cycle

Early on it was known that radio signal strength and quality varied throughout the day and seasonally. By 1910 commercial and government radio operators knew that radio signals traveled further at night and were less noisy in the winter.

It was widely believed at the time that the quality and distance of radio transmissions was directly proportional to wavelength. Radio transmissions covered longer distances and were more reliable using long wavelengths instead of short ones. For this reason government and commercial radio stations operated on 50 to 1000 KHz (wavelengths of 6,000 down to 300 meters). Primarily from antenna considerations, the marine distress frequency was established at 500 KHz. A lower frequency would have been better, but lower frequencies meant longer antennas, recalling that wavelength

$$\lambda = \frac{c}{f}$$

where

c = speed of light
f = frequency

Even at 500 KHz it was difficult to get an antenna to fit aboard ship, as clearly indicated in Figure 1.

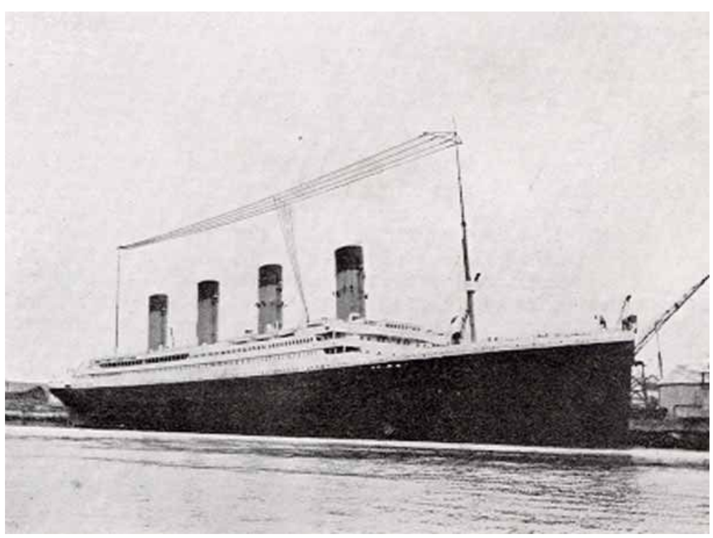
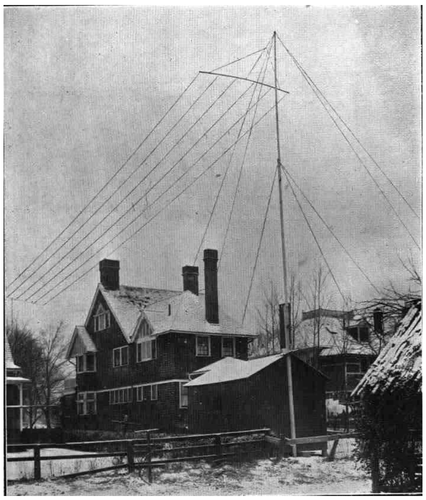
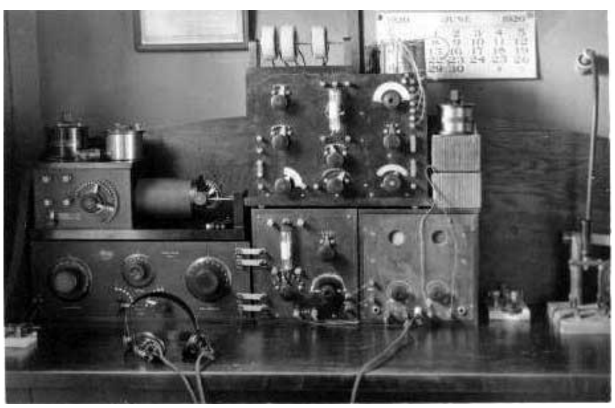


Figure 1 Shipboard antenna (credit: astrosurf.com)

Amateur radio operators were exiled to frequencies in the 200 meter band (approximately 1.5 MHz) so as not to interfere with commercial and government operations. Frequencies at and above 1.5 MHz were considered worthless from a commercial perspective. However, amateur radio operators became very successful operating at these frequencies relaying messages in short hops across country as well as communicating point to point over moderate distances.



(credit: www.arrl.org)



(credit: oldpassions.com)

Figure 2 Early amateur radio stations

In October 1924 amateurs were given the short wave frequency bands of 80 meters (3.5 – 4.0 MHz), 40 meters (7.0 – 8.0 MHz) and 20 meters (14.0 16.0 MHz). Since radio performance was believed to be better on longer wavelengths, these short wave frequency bands were thought to be even more worthless than 200 meters. Some amateurs reluctantly moved to these new frequencies to avoid crowding on 200 meters. The movement became a stampede when it was discovered that Europe, South America, New Zealand, Japan, and other distant places could be reliably worked on 40 and 20 meters. But the good times came to an end. By 1933 the 20 meter band was completely dead. No one knew why, it just died. Around 1935 the short wave bands came roaring back to life. It gradually became apparent that performance on the short wave bands was tracking the solar cycle.

So, what is the solar cycle? The solar cycle refers to the appearance and disappearance of sunspots over roughly a 11 year period. A cycle is arbitrarily defined to begin at sunspot minimum when few if any sunspots are visible. Figure 3 shows the Sun at solar minimum (picture on the left) and solar maximum on the right.

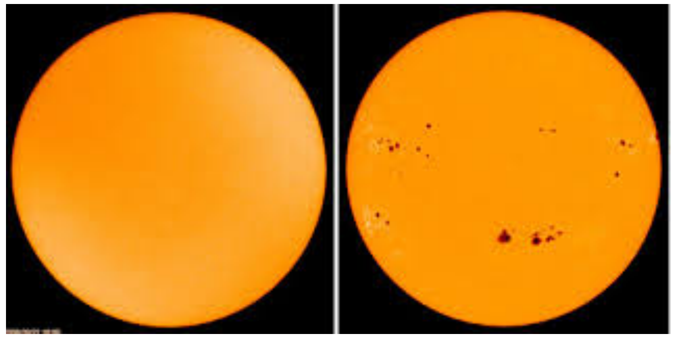


Figure 3 Sun at solar minimum and maximum (credit: springer.com)

Solar cycles from 1755 to the present are shown in Figure 4. The vertical scale on the left side of the graphs is the number of sunspots observed. Note that the

- Amplitude (the maximum number of visible sunspots),
- Shape, and
- Duration

vary considerably from one cycle to the next. For example, 87 sunspots were observed at the peak of Solar Cycle 5 in 1800 while an astounding 295 were visible during the 1957 peak of Solar Cycle 19. A solar cycle typically lasts for 11 years, but can range in duration from 8 to 15 years

Increasing numbers of sunspots appear at a fairly rapid rate as a cycle begins. The solar cycle reaches a maximum in roughly 3 to 6 years with large numbers of sunspots covering the Sun's surface. The number of sunspots then slowly decreases over the next 5 to 8 years reaching a minimum at the end of the cycle.

Continuous daily observations of sunspots began at the Zurich Observatory in 1849. Information collected from earlier observations have been used to extend the records back to 1610 when Galileo Galilei began observing the Sun with a telescope.

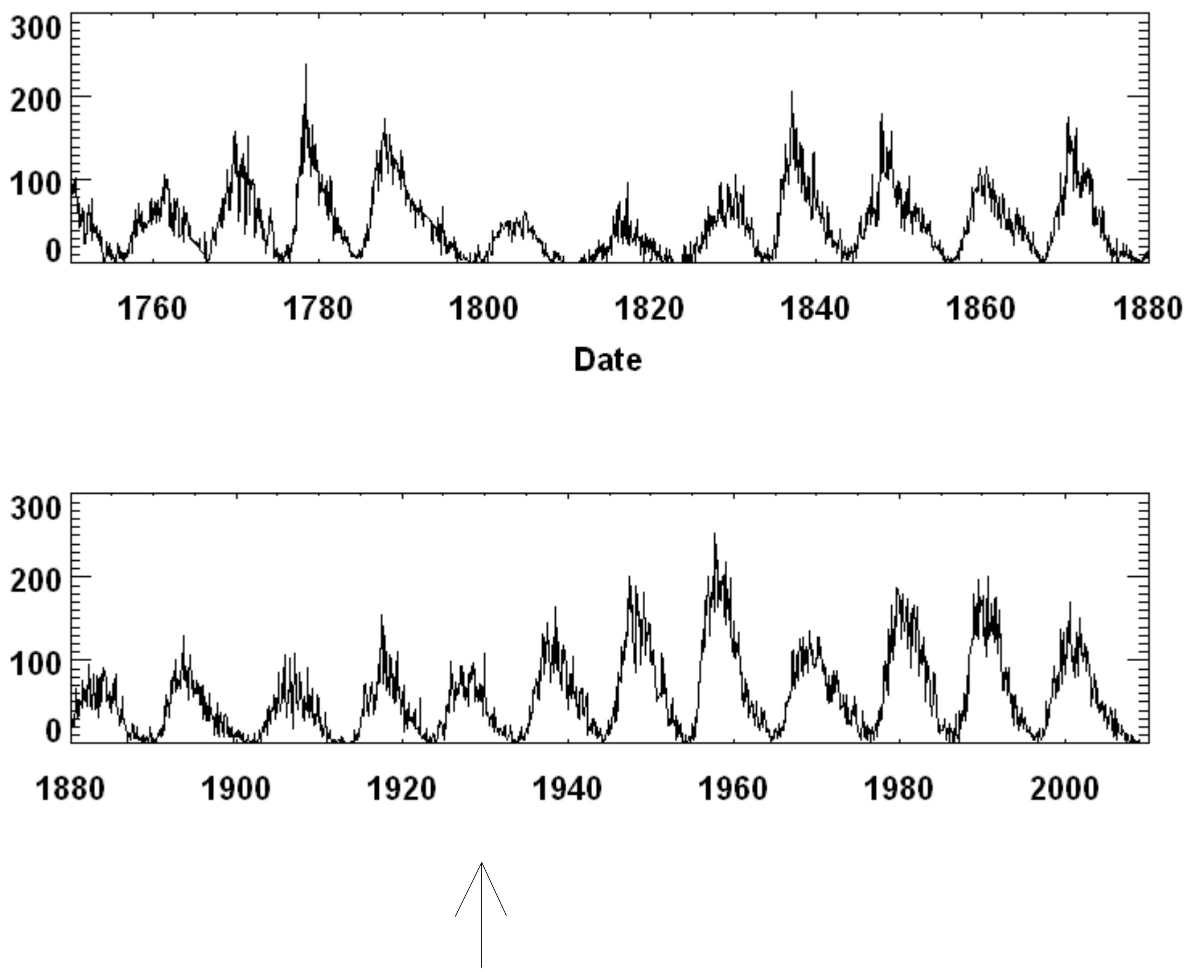
Table 1 lists the solar cycles from 1755 through the present. The solar cycle beginning in 1755 is defined as Solar Cycle – 1.

The table list:

- the year in which a cycle began,
- when it reached solar maximum, and
- when the cycle ended.

The table also shows

- the number of sunspots observed at solar minimum (SN min),
- the number at solar maximum (SN max),
- the number of years to reach solar maximum (T rise years),
- the number of years to again reach solar minimum (T fall years), and
- the total duration of the cycle (Total years).



Solar cycle number 16 peaked in 1928 with excellent long distance communications on 40 and 20 meters. But 20 meters was dead by 1933. In 1935 short wave communications on 20 meters came back, suggesting that short wave radio performance was tied to the solar cycle.

Figure 4 Solar cycles from 1755 to 2010 (credit: SpaceWeather.com)

Solar Cycle #	Begin	Maximum	End	SN min	SN max	T rise years	T fall years	Total years
1	May 1755	Jun 1761	Aug 1766	11.3	150.7	6.08	5.17	11.25
2	Aug 1766	Oct 1769	Jun 1775	16.0	208.9	3.17	5.67	8.83
3	Jun 1775	May 1778	May 1784	11.7	269.6	2.92	6.00	8.92
4	May 1784	Nov 1787	Jun 1798	15.2	239.0	3.50	10.58	14.08
5	June 1798	Dec 1804	Aug 1810	4.6	87.5	6.50	5.67	12.17
6	Aug 1810	Mar 1816	Apr 1823	0.0	84.6	5.58	7.08	12.67
7	Apr 1823	Jun 1829	Aug 1833	0.2	119.2	6.17	4.17	10.33
8	Aug 1833	Feb 1837	Jun 1843	12.4	254.7	3.50	6.42	9.92
9	Jun 1843	Dec 1848	Jan 1856	17.8	223.6	5.42	7.08	12.50
10	Jan 1856	Jul 1860	Apr 1867	6.3	187.1	4.50	6.75	11.25
11	Apr 1867	Jul 1870	Dec 1878	8.2	241.5	3.25	8.42	11.67
12	Dec 1878	Jan 1884	Feb 1890	3.2	130.1	5.08	6.08	11.17
13	Feb 1890	Aug 1893	Sep 1901	6.6	149.1	3.50	8.08	11.58
14	Sep 1901	Oct 1905	Jun 1913	4.8	106.6	4.08	7.67	11.75
15	Jun 1913	Aug 1917	Apr 1923	1.8	186.8	4.17	5.67	9.83
16	Apr 1923	Jun 1928	Sep 1933	9.3	136.7	5.17	5.25	10.42
17	Sep 1933	May 1937	Apr 1944	4.9	199.7	3.67	6.92	10.58
18	Apr 1944	Jun 1947	Apr 1954	10.8	228.3	3.25	6.75	10.00
19	Apr 1954	Nov 1957	Aug 1964	4.7	295.0	3.58	6.75	10.33
20	Aug 1964	Feb 1969	Mar 1976	12.6	157.9	4.50	7.08	11.58
21	Mar 1976	Nov 1979	Sep 1986	18.1	236.6	3.67	6.83	10.50
22	Sep 1986	Sep 1989	May 1996	14.1	217.9	3.00	6.67	9.67
23	May 1996	Nov 2001	Dec 2008	11.1	185.1	5.5	7.08	12.58
24	Dec 2008	Mar 2014	Apr 2020	2.2	118.2	5.25	6.08	11.30
25								
			Average	8.7	183.9	4.38	6.69	11.03

Table 1 Table of Solar Cycles (adapted from D. Hathaway)

Sunspot numbers are more difficult to determine than might be expected. If you look at the Sun through a low power home telescope, **with a strong solar filter carefully attached to the telescope to protect your eyes**, you might see three or four large sunspots. A high powered observatory telescope may see 10 to 20 sunspots, while a solar telescope in orbit could observe perhaps 50 to 100 sunspots.

So what is the correct number of sunspots, and more importantly, how do we correlate today's numbers with sunspot observations made more than a hundred years ago? The equation used today in determining daily sunspot numbers (the number of sunspots seen on a given day) is based on a formula derived by Rudolph Wolf in 1848. This equation is

$$R = k(10g + s)$$

where

R = the relative sunspot number
g = the number of sunspot groups observed on the solar disk
s = the total number of individual spots observed, including those in sunspot groups
k = a variable scaling factor

The scaling factor k accounts for observing conditions and the particular telescope used. Wolf set k = 1 for the observations he made with his telescope at the Zurich observatory in 1849. Each observatory has its own assigned value of k which scales its observations to what Wolf would have observed using his 1849 telescope. Today k < 1 for most observatories, generally much less than 1.

Data from many different observatories, each with their own value of k, are combined together to arrive at a daily sunspot number. Two official sunspot numbers are in common use:

1. The daily "Boulder Sunspot Number" computed by the NOAA Space Environment Center

2. The "International Sunspot Number" published by Solar Influences Data Center in Belgium.

Two other sunspot numbers are important in addition to the daily sunspot number. These are

- The monthly mean sunspot number R_m, and

- The smoothed monthly sunspot number R_s .

The monthly mean sunspot number is the average number of sunspots for the month. That is, the monthly mean sunspot number is

$$R_m = \frac{R_1 + R_2 + R_3 + \cdots + R_n}{n}$$

where

R_1 = the daily sunspot number for the first day of the month

R_2 = the daily sunspot number for the second day of the month, etc.

n = the number of days in the month

The smoothed monthly sunspot number R_s is the average of the monthly mean sunspot numbers over 13 months from 6 months before to 6 months after the month of interest m. All months have the same weighting except for the first and last months in the series (R_{m-6} *and* R_{m+6}) each of which is given a weighting of one half (0.5). Thus the smoothed monthly sunspot number

$$R_s = \left(\left[\sum_{m-5}^{m+5} (R_m) \right] + 0.5R_{m-6} + 0.5R_{m+6} \right) \Big/ 12$$

The smoothed monthly sunspot number is one of the most widely used index in ionospheric work.

Figure 5 compares the daily (yellow), monthly mean (blue), and smoothed monthly (red) sunspot numbers for Solar Cycle 24.

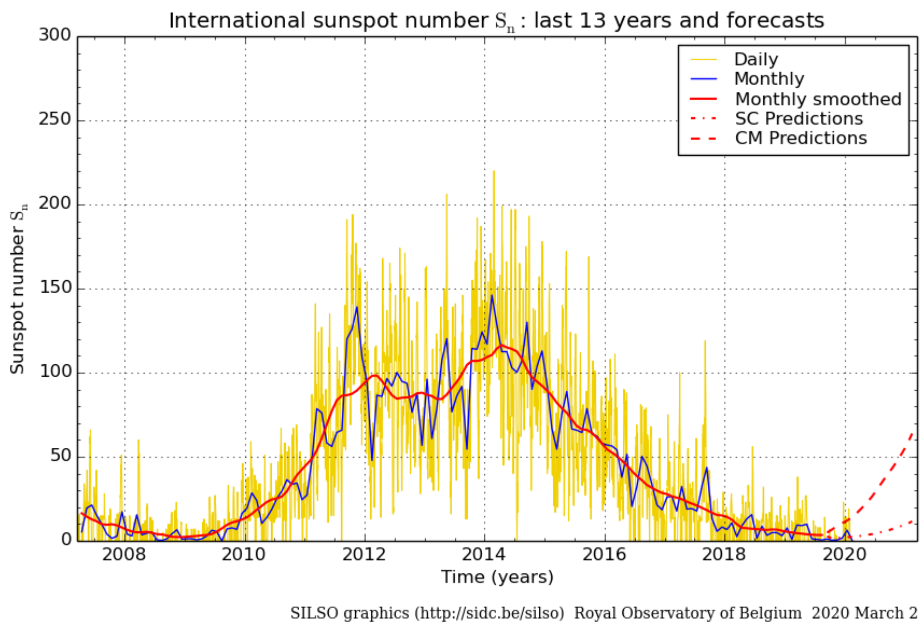


Figure 5 Solar sunspot numbers for Solar Cycle 24 (credit: SILSO graphics http://sidc.be/silso)

While the Wolf equation is simplistic, it compares remarkably well with more complex measurements of solar activity based on current technology.

One of the techniques extensively used today is background radio noise known as the Solar Flux Index (SFI). Random collisions of electrons with heavier particles produces this continuous solar radio noise which is measured by observatories every day at a frequency of 2,800 MHz (10.7 cm wavelength). The Solar Flux Index correlates well with measured levels of solar X-ray and Extreme Ultra Violate (EUV) radiation. It also correlates remarkably will with sunspot numbers as illustrated in Figure 6. In addition, Solar Flux Index is an excellent indicator of ionospheric conditions, including HF radio propagation through Earth's upper atmosphere. SFI numbers vary from a low of about 65 during solar minimum to around 225 at the peak of a solar cycle.

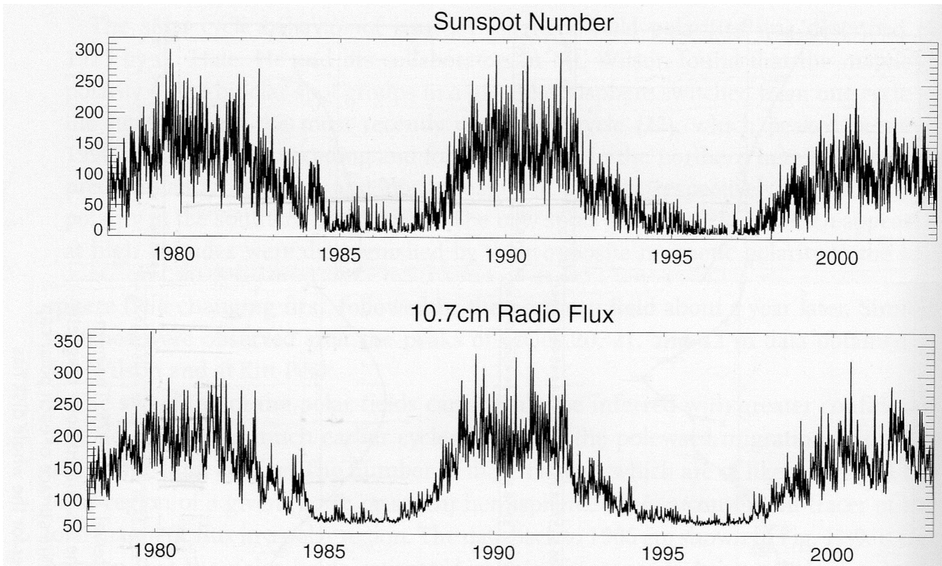


Figure 6 Radio flux vs Sunspot number (credit: G. deToma)

4.1 HF Skywave Propagation

In 1924 Edward Appleton proved that HF communications occurred by skywave propagation through Earth's upper atmosphere as illustrated in Figure 7

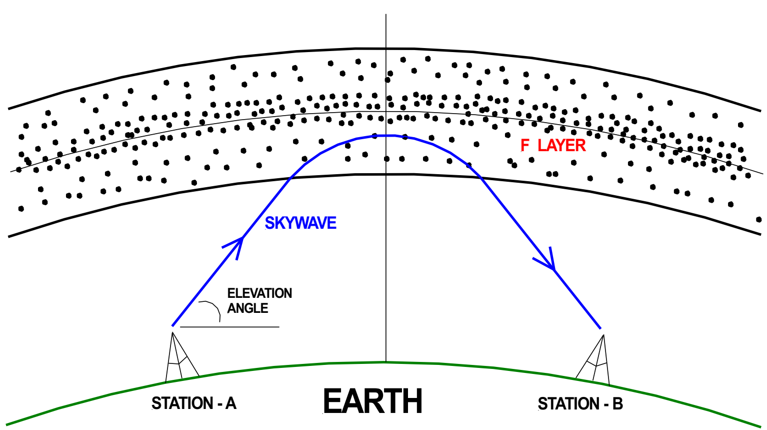


Figure 7 Skywave propagation (credit: author)

Prior to Appleton's experiments, no one really knew how radio communications occurred. It just did. There were many theories proposed by renowned scientists at prestigious universities, but none of the theories actually matched what was happening. For some time there had been suggestions that an electrified layer existed in Earth's upper atmosphere which radio waves reflected off of. But there was no experimental evidence that such a layer actually existed.

In April 1924 Appleton studied the strength of radio signals received at Kings Collage, located in Cambridge, from the powerful London BBC radio station. Cambridge is about 80 km north-east of London. Appleton discovered that signal strength was constant during the day but varied at night. He theorized that a night he was receiving two signals, one traveled along the ground and the other reflected from the upper atmosphere. The variation at night he concluded was the result of interference between the two signals. During the day only one signal was being received, the signal traveling along the ground. The other signal was absorbed by the upper atmosphere and thus could not be heard. Further investigations showed that at night the long wavelength signal transmitted by the BBC station was reflected from atmospheric ionization about 90 km above the Earth. Appleton referred to this electrified region as the E-layer, E standing for electrified. Appleton later discovered that short wave radio signals passed through the E-layer and were reflected from what he called the F-layer about 200 km above the Earth.

After the 40 and 20 meter rollercoaster experience from 1924 – 1935, the question became: why the number of sunspots on the Sun, or lack there of, had such a tremendous impact on skywave radio communications? To answer this question we first need to know two things:

1. How the ionosphere forms in the Earth's upper atmosphere

2. Why the ionosphere causes radio signals to bend back to Earth.

4.1.1 Formation of the Ionosphere

Extreme Ultra Violet (EUV) and X-ray radiation from the Sun ionizes atoms in the upper atmosphere. How does that occur? Solar radiation is absorbed by a neutral atom transferring energy to its electrons. Occasionally, an electron is sufficiently energized that it breaks away from its parent atom forming a free electron and a positive ion (an atom that has lost one of its electrons) as illustrated in Figure 8.

Occasionally is the operative word. There are 1,000 times more neutral atoms in the upper atmosphere than ions. Consequently, the ionosphere is very thin and wispy.

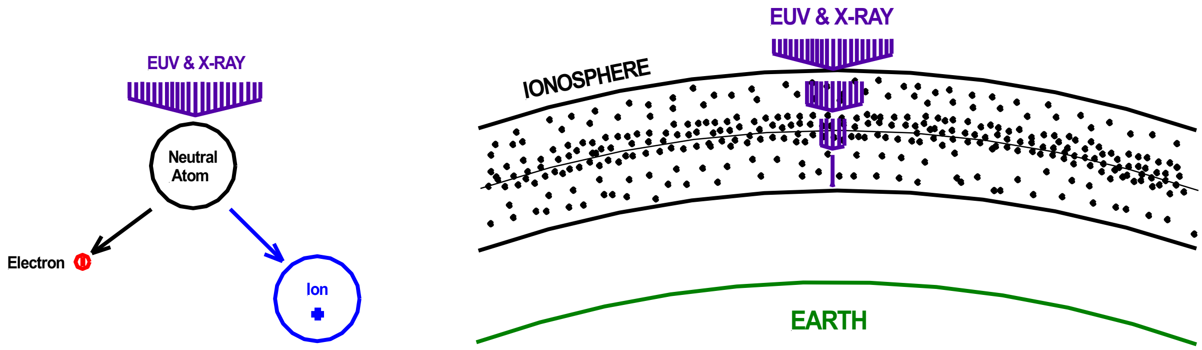


Figure 8 Ionization of atoms in Earth's upper atmosphere (credit: author)

Solar radiation is intense at the top of the atmosphere, as shown in Figure 8, but there are few atoms available to ionize. As the radiation penetrates deeper into the atmosphere, it encounters more and more atoms (mostly nitrogen and oxygen) resulting in higher levels of ionization. However, the ionization process absorbs energy from the EUV and X-ray radiation. The number of atoms keeps increasing as the radiation penetrates further into the atmosphere, but ionization levels decrease due to weakening radiation. Ionization continues to drop as the radiation weakens further and finally disappears. The point at which the radiation disappears marks the bottom of the ionosphere.

There are two additional things to note in Figure 8. Ions are way too massive to be affected by our puny radio waves. But tiny electrons interact readily with our radio waves. Thus it is electron densities that are so important to us.

4.1.2 How Skywave Propagation Works

The highest levels of ion and electron densities occur roughly in the middle of the ionosphere as shown in Figure 9.

A radio wave travels in a straight line from the transmitting antenna to the ionosphere. However, it begins immediately bending as it enters the ionosphere. The direction that it bends in is important.

- Radio waves bend toward the Earth as electron densities INCREASE.
- Radio waves bend away from the Earth when electron densities DECREASE.

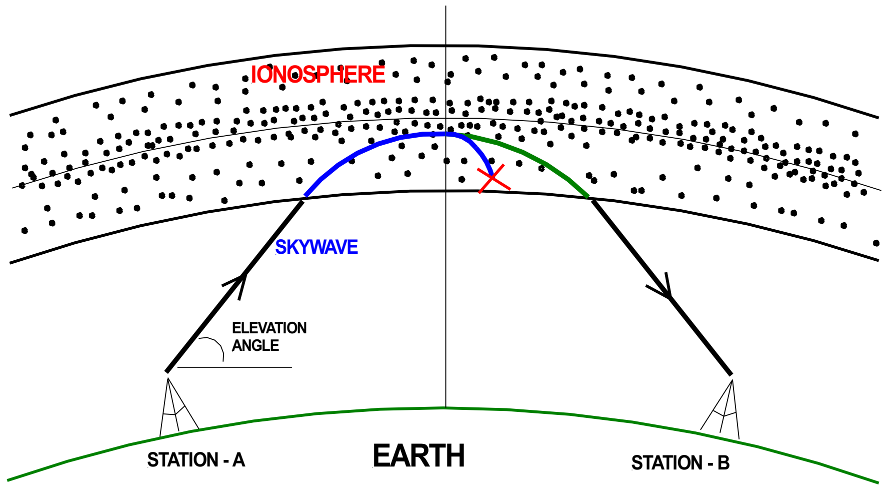


Figure 9 How skywave propagation works (credit: author)

Electron densities increase from the point of entry into the ionosphere, reaching a maximum near the middle of the ionosphere. Consequently a radio wave begins bending back toward Earth as soon as it enters the ionosphere. Bending eventually reach the point where the radio wave is traveling parallel to the Earth. If it continued to bend in the same direction, at the same accelerating rate, it would exit the ionosphere near the red X in Figure 9. But that is not what happens. Electron densities decrease as the radio wave travels back toward Earth. As electron densities decrease, the radio wave bends away from the Earth causing it to straighten out and exit the ionosphere at the point shown. It then travels in a straight line from the bottom of the ionosphere to the receiving antenna.

4.1.3 Trends in HF Communications During Solar Maximum

Radiation from the Sun is intense during solar maximum when large numbers of sunspots are visible on the Sun. The ionosphere becomes heavily ionized during the day, including a heavily ionized D - Layer. 80 meter signals are quickly absorbed in the D – Layer following sunrise and remain absorbed all day long until just after sunset, as illustrated in Figure 10. 40 meter signals are also absorbed in the D – Layer during mid day. But communications on 40 meters is often possible for an hour or so after sunrise and in the late afternoon. 20, 15, and 10 meter signals are all bent back to Earth with little absorption providing excellent long distance DX contacts all day.

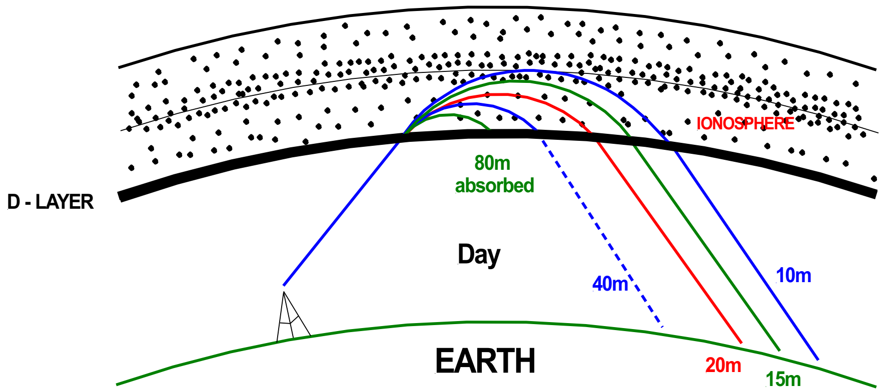


Figure 10 Daytime radio wave propagation solar maximum (credit: author)

At night levels of ionization drop, due to lack of sunlight, as electrons and ions recombine into neutral atoms. The D – Layer disappears an hour or so after sunset providing excellent night time communications on 80 and 40 meters. 80 meter communication typically lasts all night, disappearing just as the Sun rises. 40 meter communications usually dies out in the early morning hours when ionization is at its lowest level. Notice 20 meters in Figure 11. In this figure a 20 meter signal is right in the densest part of the ionosphere. In the early evening, communications on 20 meters is still possible. But as the evening wears on, ionization levels continue to drop resulting in the 20 meter signal passing through the most dense part of the ionosphere without being bent back to Earth. Once that happens the 20 meter signal enters a region of decreasing electron density meaning that the signal bends away from Earth and is lost to outer space. The same thing happens to 10 and 15 meter signals, but much faster. They are lost to outer space shortly after sunset.

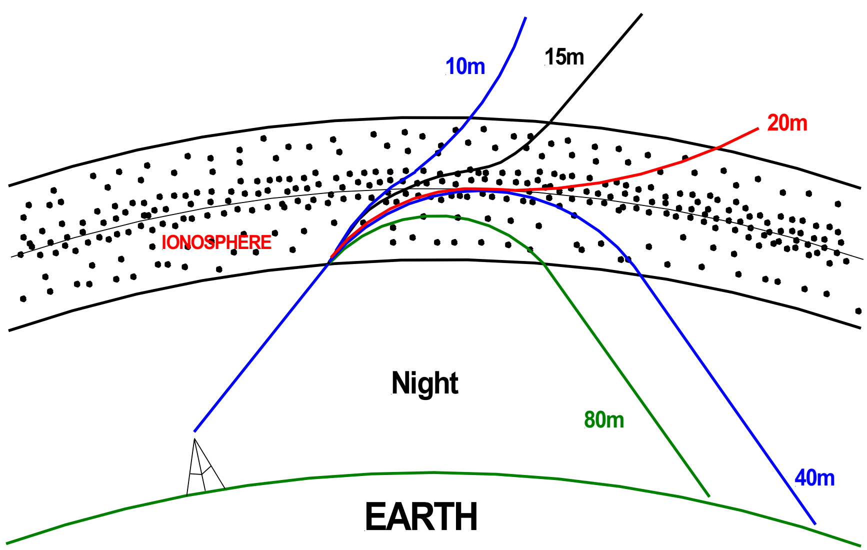


Figure 11 Night time radio wave propagation solar maximum (credit: author)

4.1.4 Trends in HF Communications During Solar Minimum

The Sun's radiation is weak during solar minimum when few if any sunspots are visible. The Earth's ionosphere is in turn weakly ionized. However, there is still a D – Layer which absorbs 80 meter signals during the day. While present, the D – Layer is weaker during solar minimum permitting regular communications on 40 meters during the day. 20 meter communications is erratic during solar minimum. On some days it is good, but on other days the band is dead. 10 and 15 meter signals are usually lost to outer space most of the time during solar minimum resulting in poor long distance DX.

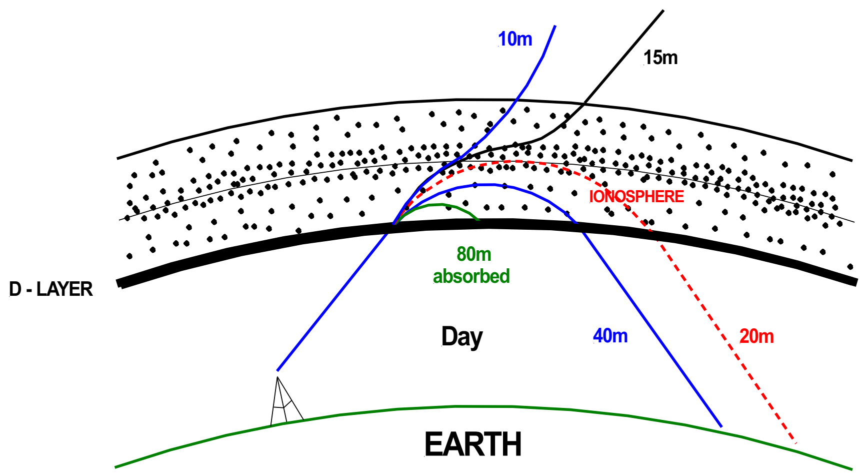


Figure 12 Daytime radio wave propagation solar minimum (credit: author)

At night levels of ionization drop even further, due to lack of sunlight, as electrons and ions recombine. The D – Layer disappears around sunset providing excellent night time communications on 80 meters. Communications on 40 meters lasts into the early evening but then dies out. There is no communications on 20 through 10 meters. These signals are lost to outer space.

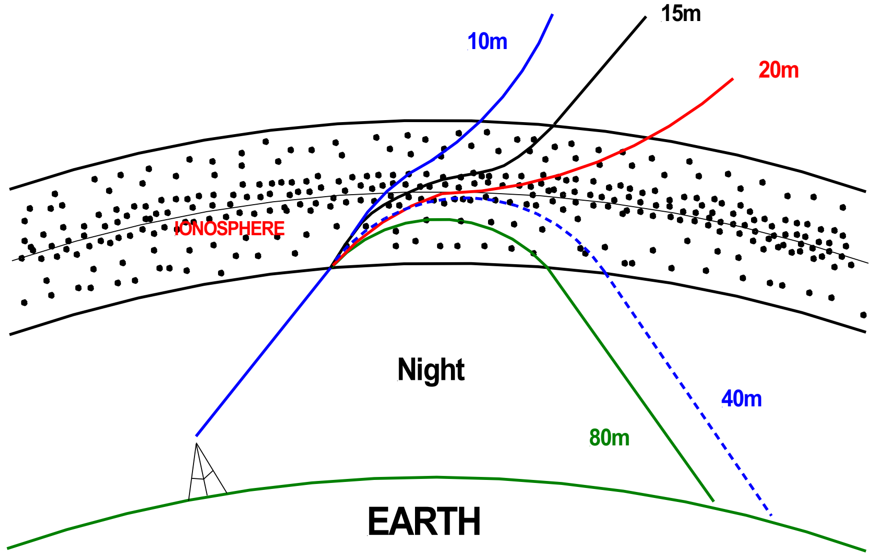


Figure 13 Night time radio wave propagation solar minimum (credit: author)

4.2 Long Term Solar Cycle Observations

Solar sunspot cycles have been observed for a long time. Figure 14 shows the solar cycles recorded from 1610 through the present.

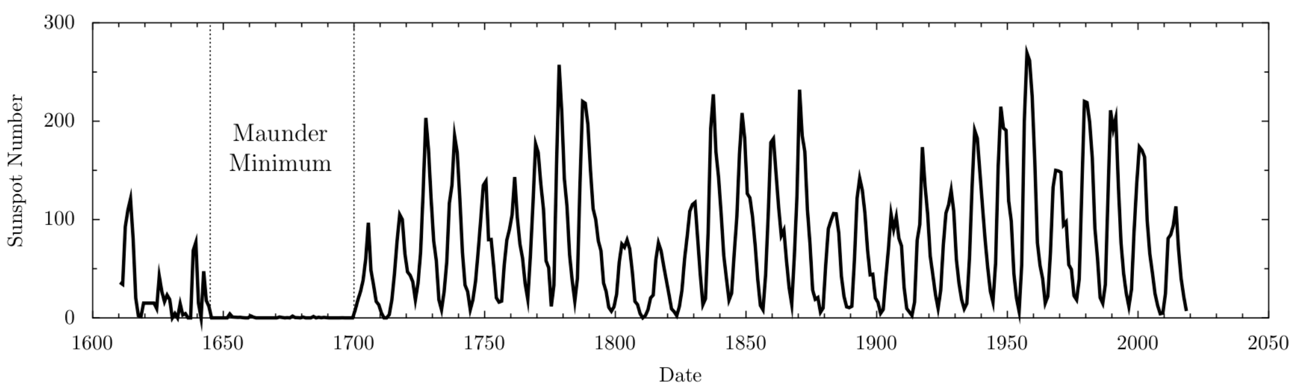


Figure 14 400 years of solar cycle observations (credit D Hathaway)

What is interesting in Figure 14 is the Maunder Minimum plus more subtle periodic dips in the amplitude of solar cycles.

Dips seem to occur every 80 to 120 years. The most recent dips occurred around 1816 (Solar Cycle 6) and 1905 (Solar Cycle 14). Referring to Table 1 shows that the amplitude of the 1816 solar cycle was slightly less than that of the solar cycle that preceded it and considerably less than the earlier and following cycles. Similarly, the amplitude of the 1905 solar cycle was less than that of solar cycles on either side of it. It is possible that the up coming Solar Cycle 25, expected to peak around 2025, may be another amplitude dip.

For 70 years, from 1645 to 1715, there were few if any sunspots visible on the Sun. That was a long time ago. One could speculate that astronomers and scientists at the time were simply not vigilant in their observations of the Sun. However, that is not so. Records indicate that the Sun was continuously and carefully observed throughout that period by noted astronomers including Hevelius, Boyle, Fogelius, Picard, and Cassini. At the Paris Observatory daily observations of the Sun were made from 1667 on. Between 1670 and 1700 there were never more than eight spots visible at one time. Only a single short-lived sunspot was seen for the entire decade from 1690 to 1700. Sunspots were so rare that the appearance of a sunspot created considerable excitement in the astronomical community. Interestingly, from 1660 to 1715 almost all the sunspot that did appear were in the southern hemisphere. It was not until 1715 that sunspots began appearing again in significant numbers, and in both hemispheres.

This long period of little or no sunspot activity is known as the Maunder Minimum. Ernest Maunder published two papers in 1890 and 1894 drawing attention to the absence of sunspots during the late seventeenth century. However, his papers as well as earlier research by Gustav Sporer were generally ignored. It was not until the landmark 1976 paper by John A. Eddy, published in the periodical Science, that interest was finally focused on the 1645 to 1715 solar minimum. It was Eddy who referred to this lack of sunspot activity as the Maunder Minimum. Eddy also suggested a possible connection between the solar minimum and a cool period throughout Europe that occurred about the same time termed the Little Ice Age. In addition, Eddy pointed out that a Medieval Warm Period from roughly 900 to 1300 occurred at a time of enhanced solar activity. Throughout this period Greenland was warm enough to be settled by Vikings and good wine was produced in England.

It is sobering to think what impact such a long solar minimum would have today on HF radio communications

It turns out that there have been other prolonged solar minimums throughout history. Figure 15 shows that two additional prolonged minimums, the Wolf and Sporer minimums, occurred from approximately 1290 through 1340 and 1410 through about 1520 respectively.

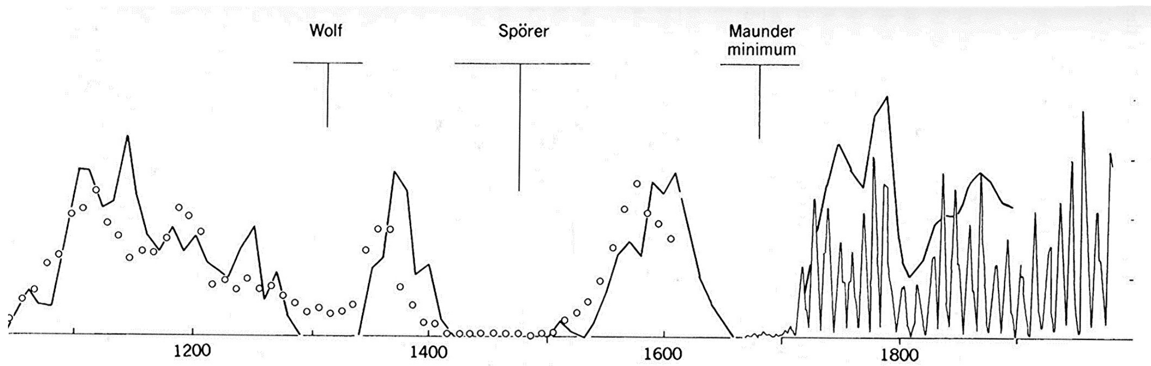


Figure 15 Prolonged Solar Cycle Minimums (credit: J. Eddy)

Evidence of these earlier prolonged minimums are found in the measurement of carbon-14 in tree rings. Carbon-14 is a radioactive isotope of carbon that is produced by energetic cosmic rays colliding with nitrogen atoms in Earth's upper atmosphere. The carbon-14 makes its way down to the lower atmosphere, is absorbed by tree leaves as part of the photosynthesis process, and ends up in tree rings as trees grow. During solar maximum the Sun's intensified magnetic field partially shields the Earth from cosmic rays resulting in less carbon-14 being produced. Thus:

- Solar Maximum = Low levels of carbon-14 in tree rings
- Solar Minimum = High levels carbon-14.

Tree ring data has provided a reliable record of fluctuations in the amount of carbon-14 present in Earth's atmosphere over the last 7,000 years.

The solid line in Figure 15 shows carbon-14 tree ring measurements from roughly 1000 to 1900. Note that the carbon-14 curve in Figure 15 is inverted. Peaks in the curve represent low levels of carbon-14 in tree rings coinciding with solar maximums. Dips or valleys occur in the curve during solar minimums when high concentrations of carbon-14 are found in tree rings.

Records of northern hemisphere auroral sightings are also useful in identifying long term solar minimums and are more reliable than one might expect. Auroral activity is much more frequent during solar maximum. Thus a prolonged period of little auroral activity corresponds to a prolonged solar minimum. The circles in Figure 15 represent the number of auroral sightings over the period from 1000 to 1600. Telescope sunspot observations are shown in Figure 15 from 1700 to present.

The fairly regular behavior of solar cycles over the past 300 years (from roughly 1700 to the present) may NOT be the norm. It appears that prolonged solar minimums occur at regular intervals of 150 to 250 years. We may be over due for the next prolonged minimum!

Figure 16 shows that the amplitudes of solar cycles were relatively consistent from 1880 to 1928. Sunspot activity increased from 1928 to 1964, peaking in 1957. Since 1980 sunspot activity has slowly declined as illustrated in Figure 17. Sunspot activity during Cycle 24 was about the same as it was during the early days of radio. Solar cycle 25, shown in Figure 18 peaking in 2025, is expected to be about the same amplitude as Cycle 24, or perhaps slightly smaller. Based on this information, Solar Cycle 25 could be the next amplitude dip.

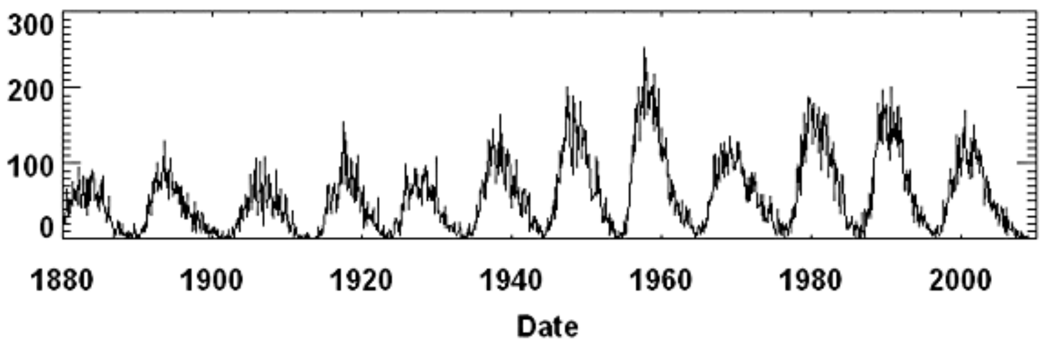


Figure 16 Solar cycles from 1880 to 2010 (credit: SpaceWeather.com)

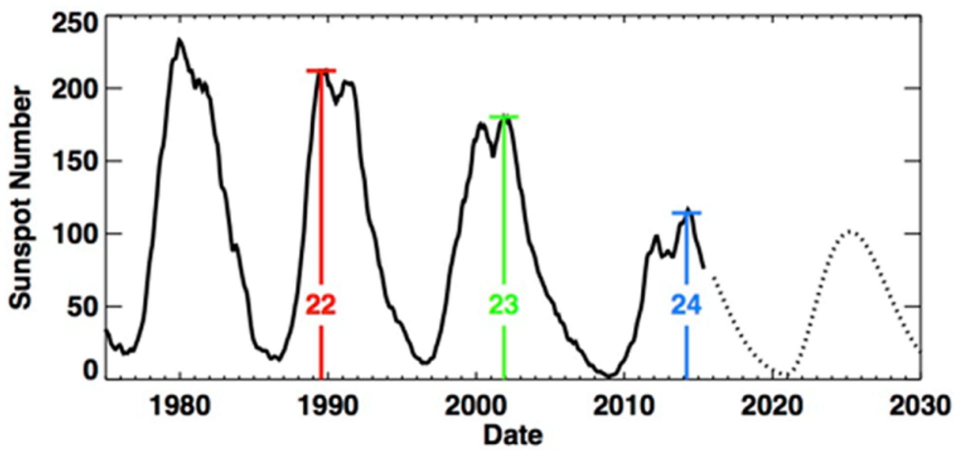


Figure 17 Solar Cycles 21 through 25 (credit: D. Hathaway)

It is interesting to look at Solar Cycle 17 in Figure 16. Cycle 17 peaked in May 1937, resulting in excellent world wide HF radio communications just as WW-II began. HF communications deteriorated throughout the war with a solar minimum occurring in 1944. The solar minimum persisted through 1945, the last year of the war. Despite deteriorating conditions, radio operators learned how to deal with the situation and got their radio traffic through. They had to! At the time HF radio was the only means available for long distance communications, not only between land based stations (since all telephone and telegraph lines had been cut in the war zone) but with ships at sea and aircraft as well.

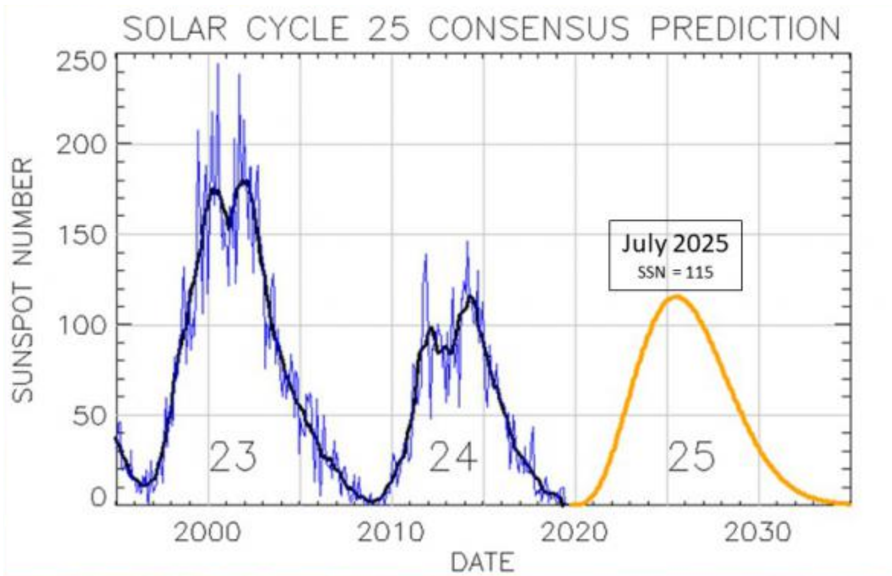


Figure 18 Projections for Solar Cycle 25 (credit: NOAA Space Weather Prediction Center)

4.3 A Closer Look at Sunspots and Solar Cycles

Figure 19, know as the butterfly diagram, shows the density of sunspots at each solar latitude for the years from 1880 through the present. At solar minimum, few if any sunspots are visible on the Sun. The first sunspots of a new cycle appear at 30º to 40º latitude in both hemispheres as clearly occurred in the year 1903. During its lifetime a sunspot remains at the same latitude. Increasing numbers of sunspots appear at progressively lower latitudes as the solar cycle continues while the sunspots at higher latitudes die out. The largest number of sunspots occur during solar maximum at a latitude of $10° \; to \; 20°$. Following solar maximum, these sunspots die out while new sunspots form closer and closer to the equator. By 1910 most of the sunspots were clustered near the equator. However, a problem develops. Sunspots in the southern hemisphere have the opposite magnetic polarity of those in the northern hemisphere. Consequently, sunspots forming in close proximity on opposite sides of the equator obliterate each other resulting in the absence of sunspots at the equator. High latitude sunspots of the next cycle can often begin to form while sunspots of the previous cycle are still visible near the equator, creating a cycle overlap. This phenomena is quite pronounced during the transition from Solar Cycle 19 to 20.

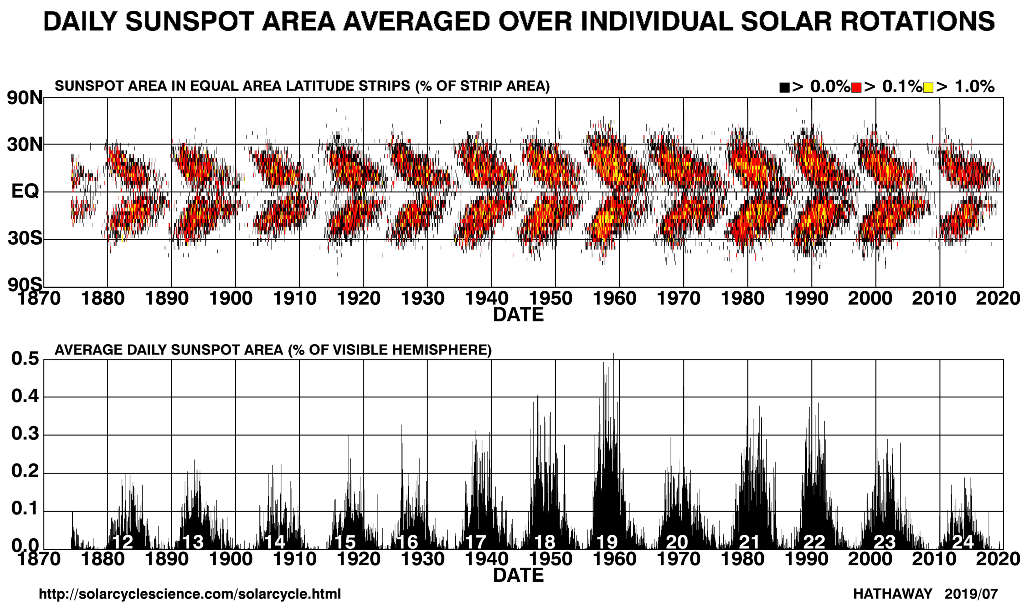


Figure 19 Sunspot butterfly diagram (credit D. Hathaway)

A sunspot has a black center called the umbra and an outer lighter colored penumbra as shown in Figure 20. The umbra is the lowest temperature (blackest) part of a sunspot. The temperature of the umbra usually ranges from about 3,800 to 4,100 degrees kelvin compared to 6,500 K for the photosphere. The penumbra around the outer edge of the sunspot is warmer than the umbra and not as black. The penumbra is characterized by large numbers of elongated dark and bright filaments that extend outward from the umbra. Out flows of material occurs in the penumbra, beginning at the umbra, reaching a maximum speed of 2 – 6 km/s, and then dissipating outside the penumbra.

Sunspots can be huge. Large sunspots (including both the umbra and penumbra) can be 50,000 km or more in diameter. More typically, the umbra is around 10,000 km in diameter, about the same size as the Earth. The Earth's diameter is 12,756 km. Sunspots appear, last several days, and then disappear. Some sunspots may last for several weeks.

While individual sunspots do form, most of the time they occur in pairs or more frequently in groups as shown in Figure 21.

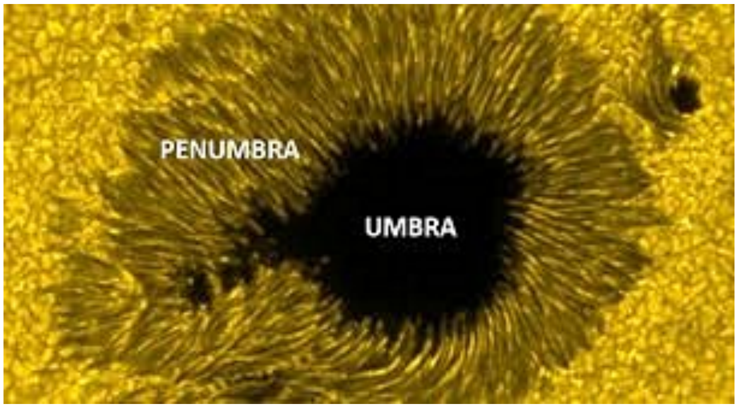


Figure 20 A closer look at sunspots (credit: spaceweatherlive.com)

Figure 21 Sunspot groups on face of the Sun (credit: NASA Goddard Space Flight Center)

A sunspot develops at a point where the Sun's magnetic field erupts through the photosphere (the Sun's visible surface) as shown in Figure 22. A second sunspot occurs were the magnetic field plunges back into the photosphere. Notice that sunspots have magnetic polarities created by the magnetic field's direction of flow. The field flows out of a magnetically north (N) sunspot and into a south (S) magnetic spot. North and south are often represented as "+" and "-" respectively.

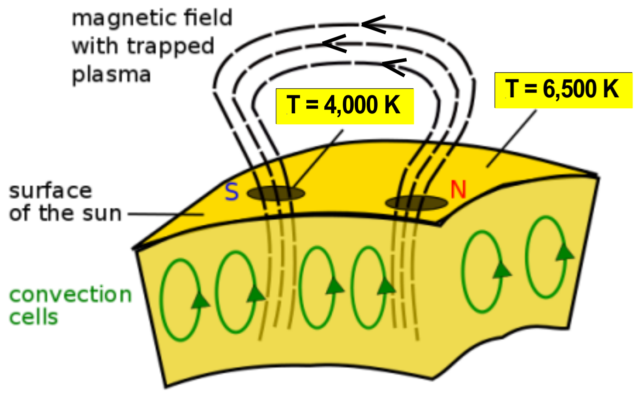


Figure 22 Sunspot formation (credit: Sky Maps with Pierre Auger Data)

The peak intensity of the field emanating from a sunspot is in the neighborhood of 3,000 to 4,000 gauss. This field is so intense that it suppresses the upward flow of hot plasma below the sunspot location, causing the sunspot site to be 2,500 degrees cooler than the surrounding photosphere. The sunspot's lower temperature is why it appears black in color.

Sunspots are part of the Sun's photosphere and thus rotate with the Sun. In Figure 23 sunspots move from left to right as the Sun rotates. Sunspots that disappear around the Sun's right limb may reappear about 13 days later on the left side of the Sun.

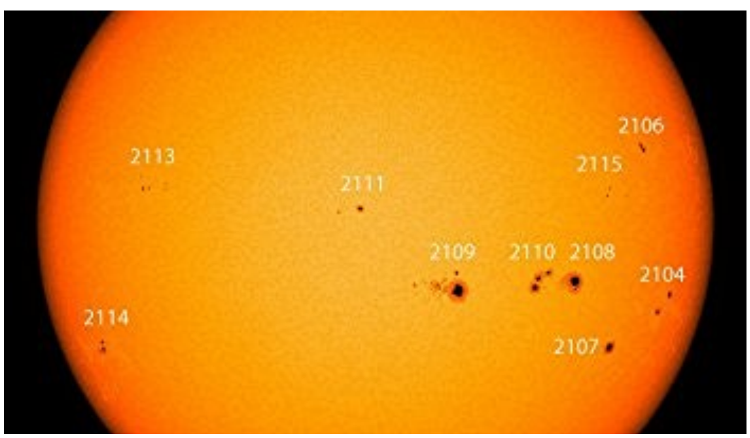


Figure 23 Sunspots on the face of the Sun (credit: NOAA Space Weather Prediction Center)

While there is a strong correlation between sunspots and the quality of long distance radio communications, it turns out that sunspots have little to do with HF propagation. Sunspots are far too low in temperature to generate the EUV radiation needed to ionize Earth's upper atmosphere. The required EUV radiation is primarily produced by plages.

Plages are hot bright irregularly shaped areas easily visible in the image of the chromosphere shown in Figure 24. Plages are formed in the chromosphere by intense magnetic fields radiating out from the underlying photosphere. Plages occur in active sunspot regions of the Sun and usually form several days prior to sunspots in the area. Plages typically last long periods of time, longer than their associated sunspots, and emit copious amounts of EUV radiation.

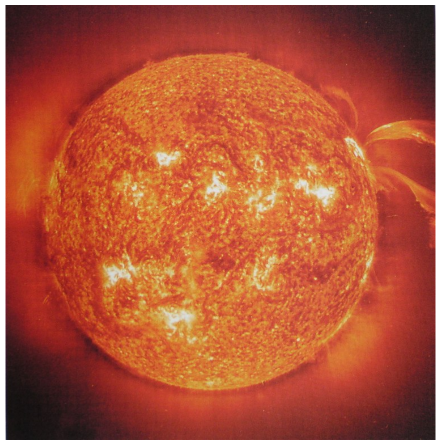


Figure 24 Sun's chromosphere shown in H alpha (H_α) light (credit: universitytoday.com)

The problem is that plages can not be seen in normal sunlight because the photosphere too bright. But sunspots, like those shown in Figures 21 and 23, can easily be seen. Thus sunspots become markers for plages. A large number of sunspots means a large number of plages, high EUV levels, strong ionization of the Earth's ionosphere, and good HF radio communications.

4.4 Why The Solar Cycle Occurs

The solar cycle is the direct result of the Sun's very "twisted and tortured" magnetic field illustrated in Figure 25. Looking at how the Earth's stable well behaved magnetic field forms helps us to better understand the ciaos occurring on the Sun.

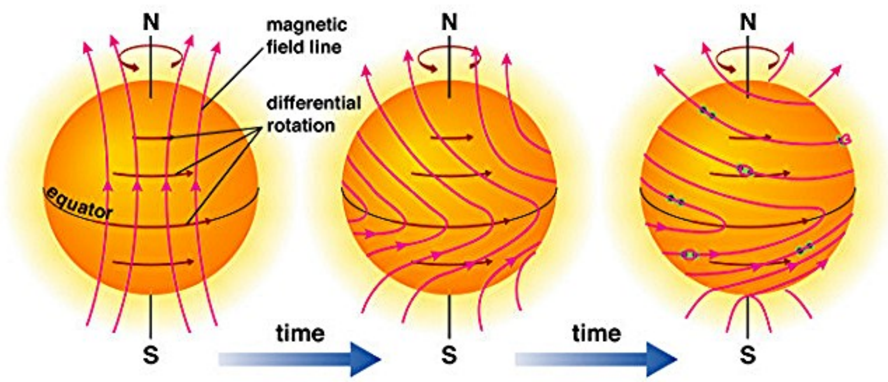


Figure 25 The Sun's twisted and tortured magnetic field (credit: skyandtelescope.com)

4.4.1 The Earth's Magnetic Field

The structure of the Earth and its internal temperature profile are shown in Figure 26 and Figure 27 respectively. The Earth's curst ranges from 5 – 70 km in depth. Below the curst the Earth's mantle extends to a depth of 2,890 km, making it the thickest part of the Earth. The mantle consists of silicate rocks rich in iron and magnesium. The outer core is composed of nickel-iron so hot that it is in a molten state. The high temperature ionizes some of the nickel-iron atoms creating an electrically charged plasma. Temperatures in the Earth's inner core are even higher. However, instead of being liquid as one would expect, the inner core is compressed into solid nickel-iron by the Earth's gravity.

Thermal convection currents within the outer core transport hot molten nickel-iron from just above the inner core to the mantle as shown in Figure 28. The molten nickel-iron cools as it rises toward the mantle, becomes more dense, and gradually sinks back toward the inner core. As it sinks it is heated by the intense thermal radiation from the inner core. The molten nickel-iron expands, becomes buoyant, and again drifts outward toward the mantle beginning a new convection cycle.

Since the hot molten nickel-iron is a plasma, with an abundance of positively charged ions and negative electrons, the convective flow of this plasma creates circulating electrical currents within the outer core.

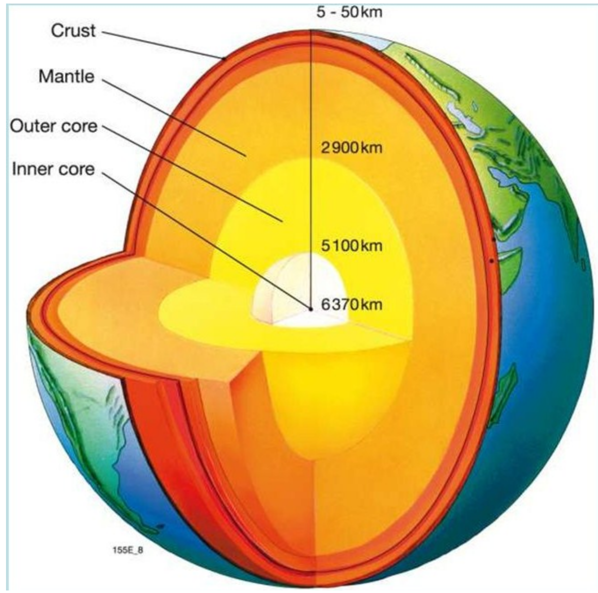


Figure 26 The Earth's Interior Structure (credit: Lakshya Geography)

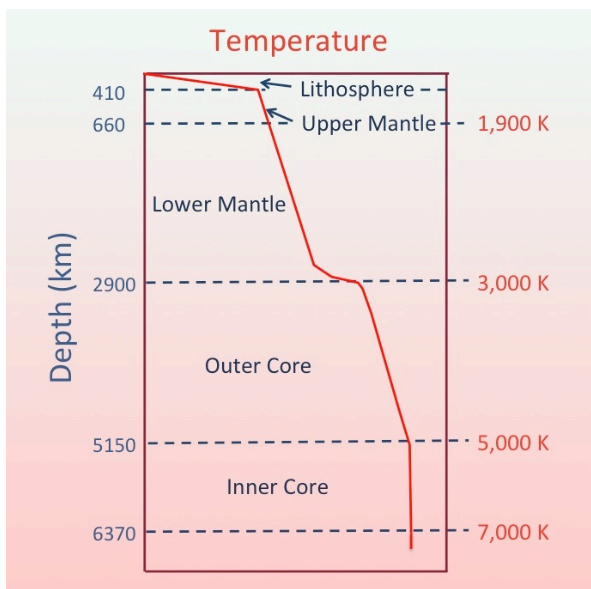


Figure 27 Temperature Profile of Earth's Interior (credit: Wikipedia)

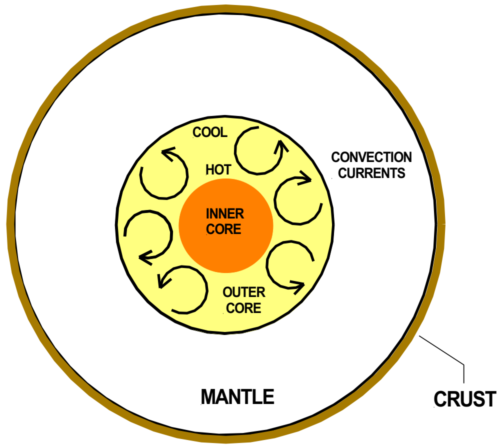


Figure 28 Outer Core Convection Currents (credit: author)

The electrical currents create a magnetic field in accordance with Ampere's Circulation Law. In this case, the Earth's magnetic field illustrated in Figure 29. Note that this field is created deep within the Earth. The Earth's magnetic field changes slowly over time, but is stable for millions of years.

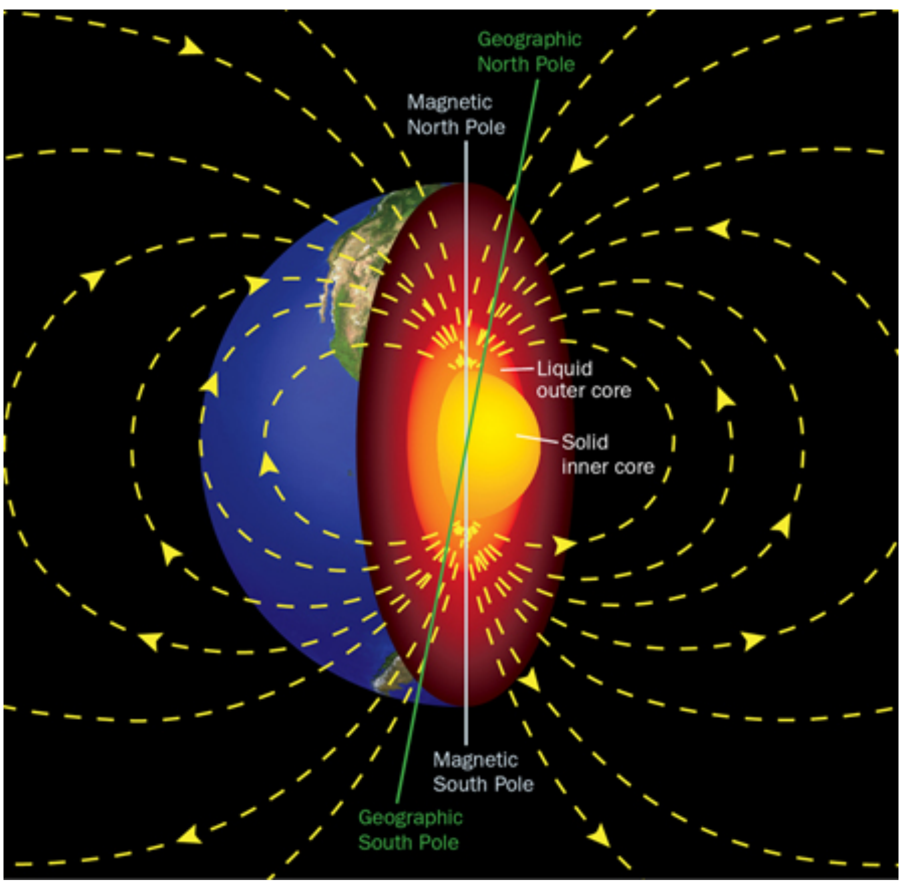


Figure 29 Earth's Magnetic Field (credit: Quora)

The Earth's magnetic field is very important to us. As shown in Figure 30, it creates a protective sheath around the Earth diverting the Sun's harsh solar winds away from Earth's surface.

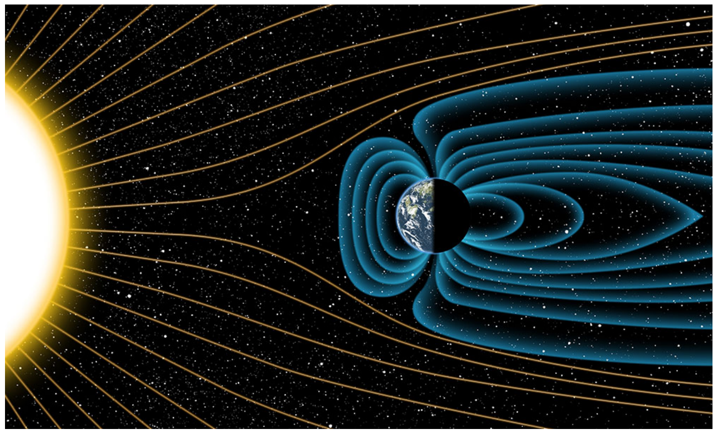


Figure 30 Earth's protective magnetic field (credit: University of Rochester)

Mars was once similar to the Earth. It had rivers, lakes, and even an ocean or two. Like Earth it too had a magnetic field created by its hot convective core. But Mars is much smaller than Earth and consequently cooled quickly. As it cooled its convective core also cooled and solidified. Once solidified the core could no longer produce the circulating electrical currents needed to generate its magnetic field. So Mars lost its magnetic field exposing the planet to the full fury of the solar winds. The winds stripped Mars of its atmosphere, dried up its lakes and oceans, and left Mars a dead planet.

This leads to an important conclusion. To have a magnetic field a planet must have a hot convective region of circulating electrical currents somewhere within the planet. This same criteria applies to the Sun as well.

4.4.2 Formation of the Sun's Magnetic Field

The Sun consists of the 6 major regions shown in Figure 31:

- Core,
- Radiation zone,
- Convection zone,
- Photosphere,
- Chromosphere, and
- Corona

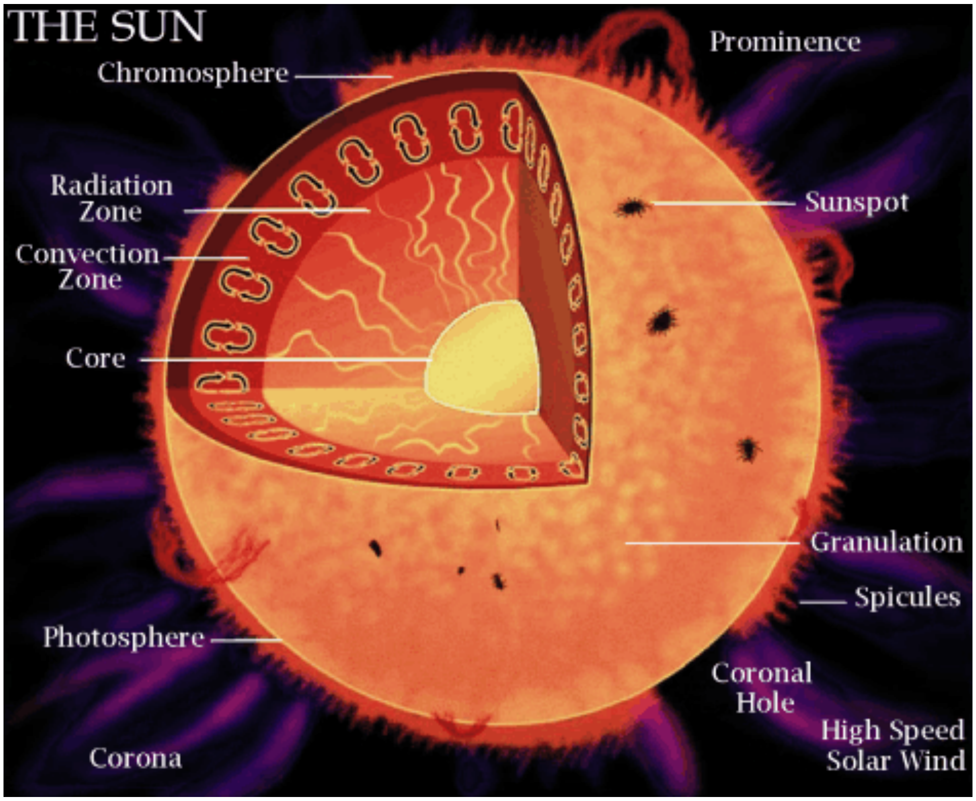


Figure 31 The Sun (credit: NASA).

The photosphere is the visible surface of the Sun, the part of the Sun that radiates the light that we see. The photosphere emits 99% of the Sun's light and heat. The intensity of this radiation decreases rapidly from the base to the top of the photosphere, a distance of only 500 km. The rapid change in intensity over a such a short distance gives the Sun a sharp well defined outer edge, instead of a fuzzy edge that one might expect from a large ball of gas. The fact that the photosphere is the furthest that we can see into the Sun, coupled with the Sun's sharp edge, gives the impression that the photosphere is the Sun's surface.

We assume that the photosphere is a very dense, almost "hard", layer of gas since we perceive it to be the Sun's surface. However, that is not the case at all. The photosphere's density is about 10,000 times less than Earth's atmosphere at sea level. Since the density of the photosphere is so low, why can we not see through the photosphere? We actually can. In Figure 32 we can see a sunspot and granules at the boundary between the photosphere and the underlying convection zone. The granules are hot plasma bubbling up from deep within the convection zone. The convection zone is so hot that it is opaque, preventing us from seeing any further into the Sun.

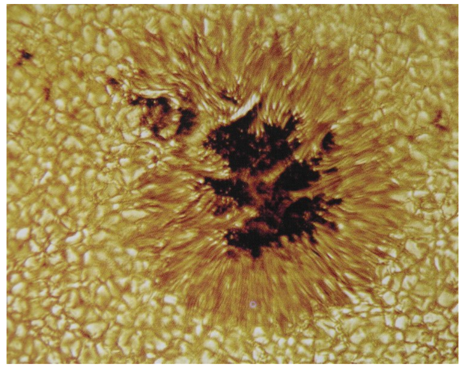


Figure 32 Sunspot and granules at base of the photosphere (credit: spaceweatherlive.com)

We think of the corona and chromosphere as the Sun's "atmosphere" since we can see through both of these regions to the photosphere below. The photosphere is so bright that we can only see the chromosphere and the corona during a full solar eclipse. The thin red ring seen around the edge of the Sun in Figure 33 is the chromosphere while the white halo is the corona. The white appearance of the corona is formed by intense sunlight from the photosphere reflecting off dust particles in interstellar space surrounding the Sun.

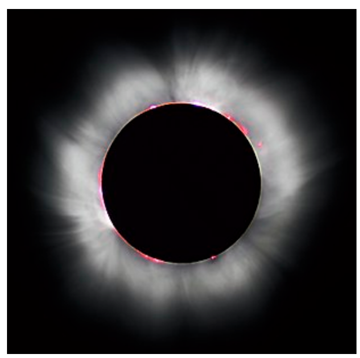


Figure 33 Full solar eclipse (credit: Wikipedia)

The chromosphere can be directly seen by observing the Sun through a telescope equipped with a hydrogen-alpha filter. Figure 34 shows the intricate structure visible in the chromosphere.

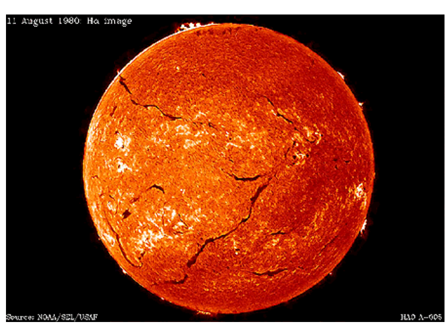


Figure 34 Chromosphere seen in H-alpha light (credit: NOAA)

Unlike the Earth, the Sun is composed almost entirely of hydrogen and helium gas.

	% By Mass	**% By Abundance**
Hydrogen	71	92.1
Helium	27	7.8
All Other	2	0.1
Oxygen		*0.061*
Carbon		*0.030*
Nitrogen		*0.0084*

Table 2 Sun's Composition

The Sun contains trace amounts of all elements in the periodic table.

Gravity resulting from the Sun's immense size squeeze the hydrogen and helium gas into a core where temperatures reach over 15 million degrees kelvin. At these temperatures and pressure thermal nuclear reaction spontaneously occurs converting hydrogen into helium and releasing an enormous amount of energy in the process. The outward flow of energy just balances the inward force of gravity, maintaining the Sun in a stable configuration for billions of years.

The interior of the Sun, shown in Figure 31, is radically different from that of the Earth. Instead of being surrounded by a hot molten outer core of circulating plasma like the Earth, the Sun's core is encapsulated in the radiation zone. A zone which is in essence a very hot thick blanket of hydrogen and helium gas 348,000 km deep. Temperature gradients within the radiation zone guarantee that its plasma of hydrogen and helium ions are frozen in place. There are no circulating electrical currents within the radiation zone to create a magnetic field.

Instead, the radiation zone transports heat in the form of high energy photons away from the core. The density of hydrogen and helium ions in the radiation zone is so high that photons only travel a very short distance before colliding with plasma particles. Because of the continuous collisions and scattering, it takes a photon on the order of 170,000 years to traverse the radiation zone. Photons travel quickly through the convection zone and take only 8 minutes to travel from the Sun's photosphere to Earth. Today's sunlight began its journey in the Sun's core about the time modern man (homo sapien) first appeared on the planet 200,000 years ago. After traveling through all of human history, that sunlight finally arrived today.

The convection zone is much different. The convection zone is similar to Earth's outer core. Hot plasma around 2 million degrees in temperature floats outward from just above the radiation zone, appears as granules at the Sun's surface, cools and sinks back toward the radiation zone. The bright granules shown in Figure 35 is hot plasma "boiling up" from deep within the convection zone. The dark channels around the granules are formed by cool plasma slinking back into the convection zone.

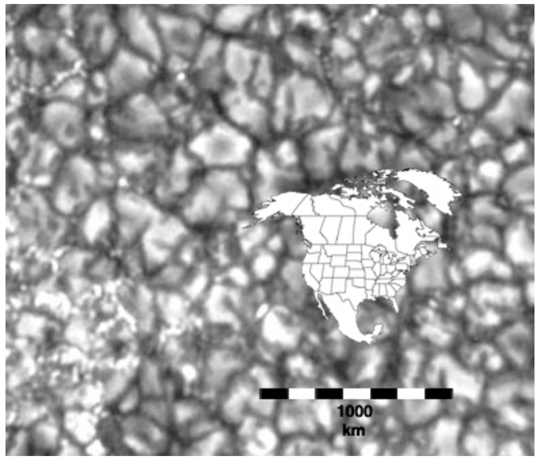


Figure 35 Thermal cells (granules) reaching the photosphere and sinking back into the convection zone. North America superimposed as a reference (credit: Wikipedia)

The hot highly ionized circulating plasma creates strong electrical currents in the convection zone. These electrical currents produce the Sun's magnetic field. But, unlike the Earth, the Sun's magnetic field is formed in the outer part of the Sun, near its surface, instead of deep within the Sun.

That is part of the problem!

The Sun is not solid. It is a huge rotating ball of gas. Furthermore, the gas does not rotate at the same rate. The Sun's equator rotates in 24.5 days while the poles rotate in 34 days. The Sun's rotational rate in days as a function of solar latitude is shown in Figure 36. In contrast, the Earth is solid forcing every point on the Earth to rotate at the same rate, one revolution every 23 hours, 56 minutes, and 4 seconds with respect to distant stars.

Variations in the Sun's rotational rate with latitude is technically referred to as differential rotation. The differential rotation is not just a surface phenomena, it extends down all the way through the convection zone. The radiation zone, on the other hand, appears to rotate rigidly. Its rotational speed is about the same as the photosphere middle latitude rotational rate. A thin layer known as the tachocline forms the transition between the differentially rotating convection and the rigidly rotating radiation zones.

The Sun's differential rotational rate is the other part of the problem.

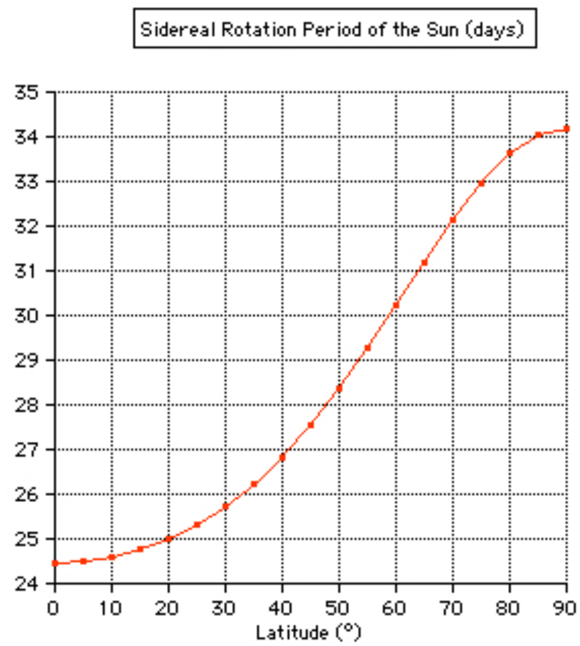


Figure 36 Rotational rate of the Sun (credit: jgiesen.de)

The fact that the Sun's magnetic field is created near its surface, instead of deep within the Sun, coupled with the Sun's differential rotation, causes the Sun's magnetic field to become badly twisted and distorted.

At solar minimum the Sun's magnetic field, shown in Figure 37, is a "quiet" north – south bipolar field similar to that of Earth's magnetic field. The strength of Earth's magnetic field is around 0.2 gauss. At solar minimum the strength of the Sun's magnetic field is about 1.0 gauss. The Sun's quiet magnetic field is not that much different from Earth's magnetic field.

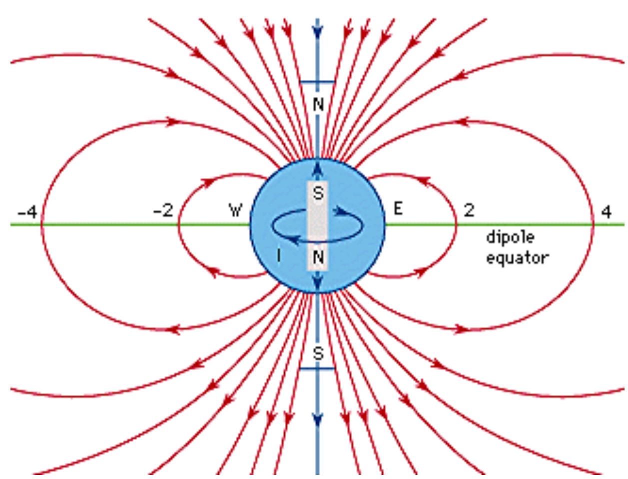


Figure 37 Sun's magnetic field at solar minimum (credit: Encyclopedia Britannica)

However, over a period of 3 to 6 years the Sun's differential rotation slowly drags and winds the magnetic field around the Sun as shown in Figure 38 "a" through "c". The magnetic field at the equator is dragged around the Sun faster than the magnetic field at the poles winding the original bipolar field in Figure 38a into the toroidal field shown in Figure 38c. In addition, convection zone turbulence twists the magnetic field lines into ropes some of which become knotted.

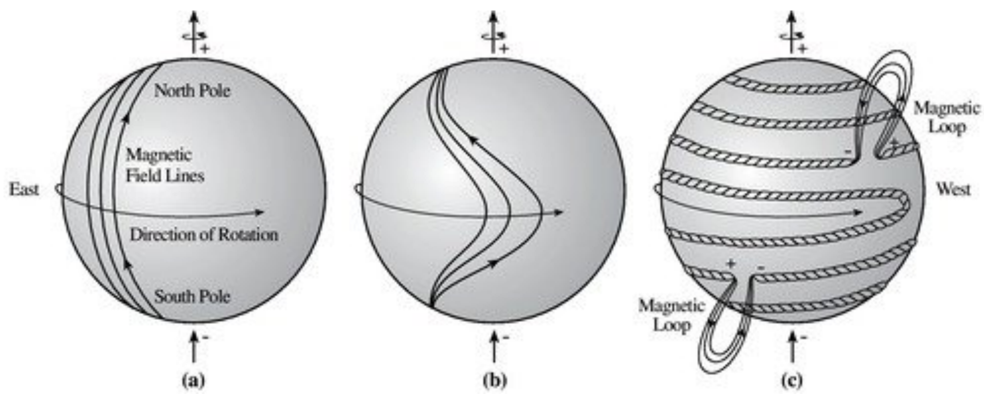


Figure 38 Magnetic field lines become wrapped and twisted around the Sun
(credit: NASA's Cosmos – ase.tufts.edu)

Winding the magnetic field around the Sun in tighter ever increasing number of turns is not a sustainable process. Something has to break, and it does!

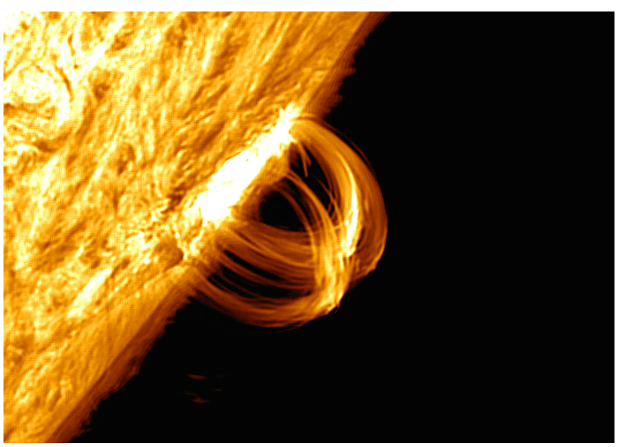


Figure 39 Prominence (credit: Astronomy Magazine)

Continued winding, twisting, and knotting creates tremendous stress in the magnetic field driving field intensities to well over 3,000 gauss.

The enormous stress eventually causes the field to rupture. As it does so high arching prominences, sunspots, and solar flares erupt from the Sun. The Sun reaches solar maximum during this very turbulent phase of the solar cycle with large numbers of sunspots visible on the solar surface.

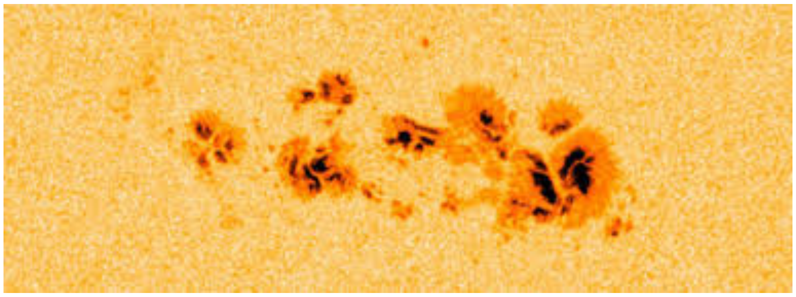


Figure 40 Sunspot Group (credit: SpaceWeatherLive.com)

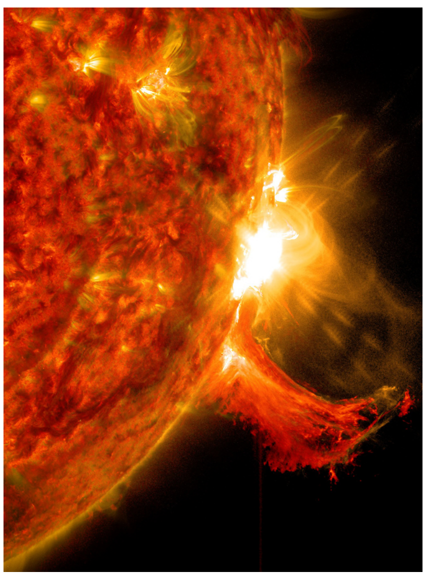


Figure 41 Solar flare (credit: NASA/SDO)

As the magnetic field disintegrates, sunspots gradually disappear and the Sun again approaches solar minimum with a quiet north-south magnetic field.

However, as the field unwinds, the Sun’s magnetic poles flip. What was the north magnetic pole becomes the magnetic south pole, and visa versa. The poles flip again at the end of the next solar cycle producing the 22 year cycle illustrated in Figure 42.

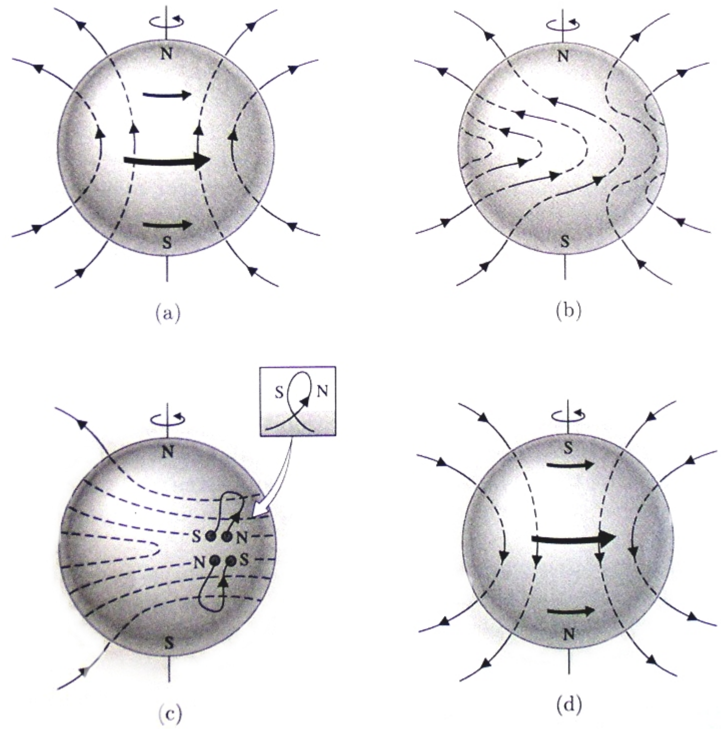


Figure 42 22 Year Solar Cycle (credit: B. Carroll, Modern Astrophysics)

Reversal of Earth's magnetic poles also occurs (the Earth's north and south magnetic poles flip). But, geological evidence, based on magnetic orientation within rocks when they solidified, indicates that reversals occur randomly over periods from 100,000 to 50 million years.

References

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

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

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

Cander, Ljiljana R.; "Ionospheric Space Weather"; Springer Geophysics, 2019

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

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

White, Stephen M.; "Solar Radio Bursts and Space Weather"; Dept. of Astronomy, University of Maryland, College Park, MD 20742 USA

Friedman, Herbert; "The Astronomer's Universe Stars, Galaxies, and Cosmos"; W. W. Norton & Company, 1990

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

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

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

DeSoto, Clinton B., "200 Meters & Down"; The American Radio Relay League, Inc., 1936

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

Piccioni, Robert L.; Everyone's guide to Atoms Einstein and the Universe second edition"; Real Science Publishing, 2010

Hogan, Craig J.; "The Little Book of the Big Bang A Cosmic Primer"; Springer-Verlag New York, Inc., 1998

Zirin, Harold and Lang, Kenneth; “Sun - Astronomy”; Encyclopedia Britannica

Kawaler, Steven D., and Dahlstrom, Michael; “White Dwarf Stars”; American Scientist, 2000 [pdf]

“White Dwarf Stars” National Aeronautics and Space Administration, Goddard Space Flight Center, 2010

“Solar experts predict the Sun’s activity in Solar Cycle 25 to be below average, similar to Solar Cycle 24”; 2019 NOAA Space Weather Workshop in Boulder, Colo.

Dobler, Sacha; “The next Grand Solar Minimum has (very likely) begun: NASA predicts lowest solar cycle in 200 years”; 14 June 2019

Kuznetsov, V. D.; “From the geophysical to heliophysical year: The results of the CORONAS-F project”; Russian Journal of Earth Sciences, Vol. 9, ES3004, doi: 10.2205/2007ES000276, 2007

Hathaway, David H., Wilson, Robert M., Reichmann, Edwin J.; “Group Sunspot Numbers: Sunspot Cycle Characteristics”; NASA/Marshall Space Flight Center, Huntsville, AL., 2002

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

Masi, Marco; “Quantum Physics: an overview of a weird world”; Marco Masi 2018

Zettili, Nouredine; “Quantum Mechanics Concepts and Applications second edition”; John Wiley & Sons, 2009

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

“Aviation Weather”; Department of Transportation FAA and Department of Commerce NOAA, 1975

Williams, Matt; “What is heat conduction?”; Universe Today, 12/9/2014

Halliday, David and Resnick, Robert; “Physics For Students of Science and Engineering Part II second edition”; John Wiley & Sons, Inc., 1962

Zumdahl, Steven S. “Chemistry second edition”; D. C. Heath and Company, 1989

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

“The ARRL Antenna Book For Radio Communications”; ARRL

Whitfield, Philip; “From So Simple A Beginning”; Marshall Editions Developments Limited, 1993

Roberts, Alice; “Evolution The Human Story”; DK, 2011

Potts, Richard and Sloan, Christopher; “What Does It Mean To Be Human?”; National Geographic,

Dr. David Hathaway, Dr. Lisa Upton "Solar Cycle Science Discover the Sun!"
http://solarcyclescience.com/

Solar Physics Group at NASA's Marshall Space Flight Center
https://solarscience.msfc.nasa.gov