Auroral Effects in the D Region of the Ionosphere

SYUN-ICHI AKASOFU
Geophysical Institute, University of Alaska

The Sun influences the Earth's atmosphere in three ways:

1. Radiations
   (a) UV radiation and X-rays
   (b) Visible radiation
   (c) Infrared radiation
2. Corpuscles
   (a) Energetic particles
   (b) Plasma
3. Gravitation (atmospheric tide)

Our main concern here is possible effects of the first two, in particular (1(a)), (2(a)), and (2(b)), on relatively short-term changes in the circulation of the atmosphere (namely, the development of cellular patterns in the zonal westerly flow, leading to the formation of cyclones) and relatively long-term changes in climate.

Both the solar UV radiation and corpuscles affect the upper atmosphere in essentially the same way, although details of the processes involved are considerably different. They change the chemical composition of the upper atmosphere and heat it. Both the solar UV radiation and X-rays (1(a)) and solar energetic particles (2(a)) penetrate directly into the upper atmosphere, while effects of the solar plasma are felt in the upper atmosphere through an intermediate process called the solar wind/magnetosphere interaction. The interplanetary magnetic field is an essential ingredient in this coupling process. This interaction process converts the kinetic energy of solar wind particles into magnetic energy which is stored in the tail portion of the magnetosphere (the magnetotail). This stored energy is intermittently converted into the kinetic energy of auroral particles. In this conversion process, auroral particles are accelerated and penetrate into the upper atmosphere. Thus, it is after the conversion process that the solar plasma can affect the upper atmosphere.

The solar wind/magnetosphere interaction can cause also a large-scale circulation of plasma in the magnetosphere. The "friction" between the plasma and the neutral atmosphere beneath it is responsible for the cause of a concentrated electric current along the auroral oval, called the auroral electrojet. An intense upwelling of the upper atmosphere is generated by Joule heating. These processes will be described in detail in later sections, and their effects are hereafter, as a whole, called "auroral effects."

As mentioned in the above, the end effects of both the solar UV radiation and solar corpuscles are changes in the chemical composition and heating of the upper atmosphere. Therefore, it is a formidable task to identify their possible effects on weather, unless time variations of the solar UV radiation and X-rays and corpuscles can be identified in meteorological and climatological phenomena. For example, for any 11-yr cycle variation in meteorological phenomena, it will be difficult to identify their solar sources, since both the solar UV radiation and corpuscular activity vary roughly in harmony with sunspot number. Further, some long-term changes in climate could be a result of accumulated effects of short-term changes in the atmospheric circulation.

This difficulty is not reduced for much shorter term phenomena, such as the recent finding by Wilcox et al. (1973) that the solar magnetic sector structure appears to be related to the aver-
FIGURE 1.—An example of a sector boundary passage on July 25, 1968. From the top, this figure shows the interplanetary magnetic data, ground records of Earth's magnetic field strength from several low-latitude stations and from the northern and southern pole stations, and the AU and AL indices. \( \gamma = 10^\circ \). THETA, PHI, and FMAG are, respectively, the latitude, the longitude, and the magnitude of the magnetic field sector.

FIGURE 2.—An example of sector boundary passage on May 17, 1968. From the top, this figure shows the interplanetary magnetic data, ground records of Earth's magnetic field strength from several low-latitude stations and from the northern and southern pole stations, and the AU and AL indices.

There is only a slight electromagnetic coupling between the sector boundary and the magnetosphere. Figures 1 and 2 show, from the top, the interplanetary magnetic field data (the latitude (THETA), longitude (PHI), and the magnitude (FMAG) of the magnetic field vector), geomagnetic records of the field strength of the Earth's magnetic field from several low latitude stations, those from the northern and southern pole stations (Thule, Vostok) and the auroral electrojet indices, AU and AL. A sector boundary passed near the magnetosphere at about 1500 UT, as can be seen in the PHI record. There were several sudden impulses at about that time; they indicate that a sector boundary is often associated with fluctuations in the plasma pressure, which cause compressions and expansions of the mag-
There is, however, little energy transfer from the solar wind to the magnetosphere by so few sudden impulses. There was no appreciable auroral activity during the passage of the sector boundary. Figure 2 shows a little more complicated situation, but it is quite clear that there is no unique phenomena associated with the passage of the sector boundary crossing; an enhanced index value for auroral electrojet (AE) activity is quite common without the passage of sector boundaries.

As noted by Wilcox and Ness (1965) and Wilcox and Colburn (1972), there is a fairly systematic change of the geomagnetic activity index \( K_p \) before and after the passage of the sector boundary. The \( K_p \) index increases rather sharply during the first two days (from \( K_p = 1.5 \) to 3.0) and then slowly decreases. If one interprets that the sharp “recovery” of the vorticity area index (after reaching the minimum value on the plus one day) found by Wilcox et al. (1973) is associated with this sharp increase of the \( K_p \) index, one must conclude that the tropospheric circulation responds to auroral phenomena with a time lag of one or two days. This conclusion is rather hard to believe. Jastrow, Hansen, Lacis, Quirk, Somerville and Stone (in these proceedings) showed that some responses of the tropospheric circulation becomes apparent about one week after introducing a particular type of perturbation on it. Indeed, if there were such a simple relationship between auroral phenomena and the development of cyclones, it would have been discovered a long time ago. This is particularly the case because the amount of the increase of \( K_p \) after the passage of sector boundaries is not particularly large.

Geomagnetic storms which begin about two days after intense solar flares near the central meridian can cause a far greater increase in \( K_p \). For example, the \( K_p \) indices during the great geomagnetic storm of February 11, 1958, were \((9_0, 8_+, 9_-, 8_+, 8_0, 5_-, 6_+, 6_0)\). This may be compared with a typical increase of \( K_p \) of about 2 during the sector boundary passage; note that the \( K_p \) index is a semilogarithmic index.

Figure 3 shows the magnetic record of Mea-nook, Canada, which well illustrates a successive occurrence of very intense substorms during the storm of February 11, 1958. The auroral oval descended abnormally equatorward and expanded dramatically several times as the substorms developed and decayed on that day. Figure 4 shows the most violent expansion of the auroral oval during the storm. The upper atmosphere was considerably heated during the storm; its effects were seen as a great enhancement of the O(I) \((\lambda = 6300 \text{ Å})\) mission over a large portion of the polar upper atmosphere.

Incidentally, the weather during the month of February 1958 was quite anomalous (Klein, 1958; Shellum and Tait, 1958). Klein (1958) noted:

February 1958 will long be remembered as a month of contrasting weather extremes in many parts of the United States. Many established records of long standing were broken—for cold in the Southeast, warmth in the Northwest, snow along the Gulf and Atlantic coasts, precipitation in the Great Plains and along the west coast, and dryness in the Mid-West. During the last week of the month intense cyclonic activity was responsible for new low barometer readings at many stations in the Central States, as well as for tornadoes, blizzards, and floods over a wide area.

Abnormalities of the weather were produced by corresponding abnormalities in the circulation pattern. Strong blocking ridges over Greenland and Alaska were accompanied by the deepest mean troughs on record along the east coast and in the eastern Pacific. A typically “low index” circulation prevailed throughout the Western Hemisphere as the polar anticyclones intensified and the subtropical anticyclones weakened. This was part of a great index cycle in which the prevailing westerlies of middle latitudes were displaced southward to the sub-
Figure 4.—The violent poleward expansion of the auroral oval which occurred near the maximum epoch of the great storm of February 11, 1958. Shading is the extent of the auroral oval in the region. (a) 10° 20′. (b) 10° 30′. (c) 10° 40′. (d) 10° 50′.

tropics, where they blew with unprecedented speed in the form of an expanded and intensified circumpolar vortex.

However, these abnormal features began from the beginning of January 1958, manifested in a rapid equatorward shift of the main zonal westerlies at the 700 mb level of the atmosphere, reaching a minimum latitude of approximately 31° N, about 8° S of its normal latitude, but there was little change of its location throughout the month of February 1958. Further, an intense cold spell began to cover a large portion of the US from about February 9, at least one day before the beginning of the great storm. In fact, between February 6 and 10, there were two intense block-
ing highs, one over Davis Straight and the other over northwestern Alaska; the positive height anomaly was 1150 ft and 840 ft, respectively, in 700-mb contours. (See figs. 5(a) and 5(b).) This anomalous feature was then followed by the period of record high subtropical westerlies which brought the cold spell mentioned.

This example is presented here, since it is natural to speculate relationships between the great magnetic storm of February 11, 1958, and the historic cold spell during the third week in the same month. However, the cause of the anomalous weather in February 1958 was apparently present well before the great storm. An interesting study will be to examine whether or not the 700-mb map in figure 4(b) can be "predicted" a posteriori by a numerical technique from figure 4(a), without adding any "unknown" factor on February 11. If the contour map in figure 5(a) does not lead to that in figure 5(b) on the basis of what was known on February 9, it would be of great interest to conduct numerical experiments in an attempt to construct figure 5(b) by introducing various perturbations in figure 5(a). If, on the other hand, figure 5(a) could lead to figure 5(b) without any additional perturbation, it is quite unlikely that auroral effects can significantly alter weather patterns. This is because the storm of February 11, 1958, was one of the most intense geomagnetic storms in history.

Let us go back to the finding by Wilcox et al. (1973). It is important to understand why the vorticity index begins to decrease about one day before the actual passage of the sector boundary. A more likely possibility is that the "recovery" or "increase" of the vorticity area index two days after a particular sector boundary passage is actually an effect of the one before.

Another possibility is that the "suppression" of the vorticity area index results from solar radiation effects from the vicinity of the "root" or source region of the sector boundary, which...
are expected to have possible terrestrial effects about four days before the passage of the sector boundary. In such a case, the source may be either (1(a)) or (2(a)), listed at the beginning of this article, or both. For the former, it may be noted that Krieger, Timothy, and Roelof (1973) and Hundhausen (in these Proceedings) showed that there is a marked dark area in an X-ray photograph of the Sun on the solar disk; they revived the concept of cone of avoidance which was put forward by Roberts. It may be such a dark region or bright region surrounding the dark region that has an immediate effect in the upper atmosphere; without knowing the time constants of various meteorological phenomena, it is difficult to identify the source region even in this particular case of a high propagation speed from the Sun to the Earth. Another problem associated with their new finding is that it is not very obvious as to whether or not the sector boundaries had a positive or negative effect on the development of cyclones.

At any rate, if the finding by Wilcox et al. (1973) is a key to the problem of possible effects of solar activity on weather (Wilcox, in these Proceedings), we should make every effort to find causes which have led to their interesting statistical result. It may be noted that for a relatively short-term meteorological phenomena (such as the new finding), it may not be difficult to separate (1(a)) and (2(a)), listed at the beginning of the article, from (2(b)). There are many intense western limb flares which are associated with both (1(a)) and (2(a)), but little with (2(b)).

STORMS AND SUBSTORMS

As mentioned in the previous section, the magnetic energy stored in the magnetotail is not continuously dissipated. The dissipation occurs rather impulsively, with a time scale of a few hours. This phenomenon is called the magnetospheric substorm, and some of its manifestations are the auroral substorm, polar magnetic substorm, and ionospheric substorm, which we call here as a whole "auroral effects." The direct cause of substorms is not understood.

Sometimes intense substorms occur very frequently. Such a period is called the storm. Each substorm is associated with a small amount of injection of protons (of energies of order of 50 keV) into the Van Allen belt. When intense injections occur very frequently, an intense belt of protons is formed. Since these protons carry a westward current, the belt is often called the ring current belt. The magnetic field of this (westward) ring current is directed southward near the Earth. This field is the cause of what is commonly called the main phase decrease; the horizontal component of the magnetic field is depressed for about a day or so. The Dst index is derived to provide a measure of the intensity of the ring current. The ring current begins to decay as soon as substorm activity declines, first rather rapidly for about 6 hr and then slowly. It may take one week or more for the ring current to substantially decay. Figure 6 shows an example of the relationships between the storm of July 8, 1958, and substorms associated with it. The intensity of the substorms is given in terms of the AE index, and the intensity of the storm is given in terms of the index of intensity of ring current Dst.

![Figure 6](image-url)
AURORAL EFFECTS

Figures 7(a), (b), and (c) show the auroral energy flow chart. Figure 7(a) shows sequences of processes associated with the precipitation of auroral electrons into the polar upper atmosphere. The most familiar effect is the ionization of atmospheric atoms and molecules and the subsequent chemical processes. The left-hand side of figure 7(b) shows how the kinetic energy carried by auroral electrons is transformed into different kinds of energies; the percentages are kindly provided by Rees (private communication, and Rees, 1973). The total energy input rate, $2 \times 10^{10}$ W, is estimated by taking into account the precipitation of electrons into the region of the diffuse aurora. Although discrete auroras (classical curtain-like form) are caused by a much more intense flux of electrons, their precipitation area is too small to add significantly to the total energy input. Further, it should be noted that the above value of the energy input rate occurs during magnetospheric substorms. The lifetime of a typical substorm is of the order of 1 hr. On a quiet day, there occur a few substorms. On a moderately disturbed day, several substorms can occur. During geomagnetic storms, several intense substorms can occur in 12 hr. (See fig. 6.)

It is well known that the energy input rate of $2 \times 10^{10}$ W from auroral electrons is much less than the solar blackbody radiation energy intercepted by Earth, $1.8 \times 10^{17}$ W (Barry and Chorley, 1970). Further, most of the heat energy is initially deposited in the $E$ region of the ionosphere or above, and will be conducted upward, since thermal conductivity increases rapidly upward (Schunk and Walker, 1970).

There are, however, three processes which should be considered as possible candidates in influencing meteorological phenomena. The first is the ionization by the bremsstrahlung X-rays generated by high energy electrons. Figure 8 shows an example of estimate of ion production rate by the bremsstrahlung effect during an intense auroral activity (Larsen, 1973). Johnson and Imhof (in these proceedings) showed their estimates of the ion production rate. For the bremsstrahlung effects, see Brown (1964), Rees (1964), and Kamiyama (1966). For a direct measurement of energetic auroral electrons, see Bohn (1972) and references in Larsen (1972). Obviously, the ion pairs produced in this way cannot directly become condensation nuclei, since the mesosphere is far from a state of super saturation. Some "exotic" processes must be found for them to become condensation nuclei (Mohnen, 1971). Another possibility is that the aurora emits UV radiations in a wide wavelength range (Omholt, 1971) and that a part of it can be absorbed by ozone (the Hartley and Huggins bands) in the upper stratosphere. The most interesting possibility is, however, the dissociation of molecular oxygen of auroral electrons and the resulting formation of ozone. This problem was studied by Maeda and Aikin (1968). They showed that there is little possibility for auroral electrons of energies less than 10 keV to contribute in the formation of ozone, but an intense flux (of the order of $10^{11}$ cm$^{-2}$ sec$^{-1}$) of energetic electrons (of the order of 100 keV) could modify considerably the ozone concentration at about the 50- to 65-km level. The proposed flux for this energy range appears to be certainly too high, but this problem should carefully be reexamined.

As mentioned earlier, the solar wind/magnetosphere interaction causes a large-scale convection of plasma in the magnetosphere. The motion is driven by a large-scale electric field in the magnetosphere. This convection motion of plasma interacts with the neutral component of the atmosphere in the $E$ region of the ionosphere. There, if the convection occurs across a narrow belt of high degree of ionization, a highly concentrated current is generated. (See fig. 7(c).) The energy dissipation rate by Joule heating is estimated to be about the same as that of the kinetic energy of auroral particles, $2 \times 10^{10}$ W. Cole (1971a, b) studied this problem in detail. The upwelling motion of the neutral gas in the ionosphere (by heating of the neutral gas as the combined results of the impact of auroral electrons and of the Joule heating) and the subsequent circulation has been studied by a number of workers. Here, in figure 9, we show one of such a result by Heaps (1972). For satellite observations, see Devries (1972).

Further, the convective motion of plasma tends to cause motions of the neutral component in
the ionosphere. This phenomenon is called the \((E \times B)\) drag. (See fig. 7(c).) This particular motion has been identified by observing drift motions of barium ion clouds (Heppner in these Proceedings) and by the incoherent scatter radar at Chatanika, Alaska (Banks, private communication, 1973). The energy input rate in accelerating the neutral gas is estimated to be \(1.5 \times 10^{10} \text{ W}\).

There are a number of indications that the upper atmospheric wind is generated in the ionosphere and above during auroral activity. Unfortunately, however, such winds are well confined in the upper atmosphere. There is so far no definite evidence that even the upper mesospheric gas participates in such motions. Hook (private communication, 1973) showed that the wind in the mesosphere is normal even during a high auroral activity; his observation is based on a meteor radar located in Fairbanks. Perhaps chemical releases in the upper mesosphere should be conducted to continue his observations. However, even if winds are generated by auroral activity in the upper mesosphere, there is little hope to dynamically couple the ionosphere with the troposphere by any direct means.

**SOLAR PROTON EFFECTS**

Solar protons have a profound effect in the polar upper atmosphere (see 2(a)) and cause the
AURORAL EFFECTS IN D REGION OF THE IONOSPHERE

Figure 8.—The ion production rates and electron density profile during a substorm of February 2, 1969 (after Larsen, 1973). The ion production rate $\xi$ is in ion pairs cm$^{-2}$ s$^{-1}$ and the electron density $N$ is in electrons cm$^{-3}$.

Figure 9.—The upwelling of the upper atmospheric gas in the meridian plane, generated by the heating by the impact of auroral electrons and joule heating. The arrows indicate displacements of the air parcels for a period of 12 hr (after Heaps, 1972).

phenomenon called the polar cap absorption (PCA). In terms of the ion production rate in the mesosphere, they can have a greater effect than the bremsstrahlung X-rays. Further, the precipitation occurs over the entire polar cap, the area encircled by the auroral oval. Figure 10 shows an example of PCA which occurred on February 11, 1958 (Obayashi and Hakura, 1960). Figure 11 shows an example of the estimated ion production rate by Zmuda and Potemra (1972). Complex atomic and molecular processes associated with the ionization in the $D$ region have been studied by a number of workers (Reid, 1971), and it may be of interest to examine whether or not the resulting water-cluster positive ions could become embryos for aerosol particles, as suggested by Mohnen (1971). Unfortunately, intense solar proton events are not frequent, although they may have an accumulated effect during the period of sunspot maximum. Further, it may be difficult to separate between possible effects of solar flares and those of solar protons, since most of the intense solar proton events begin a few hours after an intense flare. One possibility is, however, to use the fact that eastern limb flares do not, in general, produce intense solar proton events.

CONCLUDING REMARKS

It appears obvious that auroral effect cannot directly affect tropospheric phenomena; even violent upper atmospheric winds generated by auroral activity do not seem to directly affect mesospheric winds. On the other hand, it will be interesting to examine mesospheric conditions under auroras by chemical releases. If there is any solar activity-terrestrial weather relationship, it seems that auroral effects go through intermediate processes before affecting weather. For example, if auroral processes can change drastically the ozone concentration, an appreciable change in the radiation transfer may occur in the atmosphere.
RELATIONSHIPS BETWEEN SOLAR ACTIVITY AND METEOROLOGICAL PHENOMENA

Although this possibility may be remote or out of the question to meteorologists, possible auroral effects on the ozone concentration will be an interesting problem to examine from the point of view of aeronomy. Both observational and theoretical studies should be conducted. (In particular, it is of great interest to examine the ozone concentration directly under auroras.) It is suggested that a detailed numerical experiment should be conducted in reconstructing the weather map in the third week of February 1958 on the basis of the map in the first week of the same month. If the reconstruction fails with all the known parameters, we should examine various perturbations to the circulation pattern during the first week of February 1958. Such an experiment should provide a clue in the search of mechanisms which couple auroral activity and weather.

ACKNOWLEDGMENTS

The author would like to thank Dr. W. O. Roberts and Dr. J. H. Wilcox for their stimulating discussions on the problem. The author is also indebted to H. Cole for his discussion during the preparation of this paper.

REFERENCES


