Auroral Changes During the Eighteenth and Nineteenth Centuries
and Their Implications for the Solar Wind and the
Long-Term Variation of Sunspot Activity

J. FEYNNMAN

Department of Physics, Boston College, Chestnut Hill, Massachusetts 02167

S. M. SILVERMAN

Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts 01731

Both auroral and geomagnetic activity provide information from which the behavior of the magnetosphere and the solar wind can be inferred. Swedish auroral sightings during the eighteenth and nineteenth centuries show a remarkable pattern of changes in the latitudes at which the auroras were observed. Auroral reports from New England confirm that these variations were hemisphere-wide. The pattern of changes took place over an 106-year period and is easily distinguished from the much smaller changes that are related to single sunspot number cycles. We infer that the pattern reflects corresponding changes in the solar wind and the resultant magnetospheric configuration and that these changes were much greater than those observed since in situ measurements began. Our results show that a minimum solar wind occurred at the beginning of the nineteenth century. It has been argued elsewhere that minimum solar winds also occurred around the beginnings of the eighteenth and twentieth centuries. Since all these periods are also the reported minimums of the 80- to 100'-year variation in sunspot activity, we conclude that both the changes in the solar wind and in the strength of the cycle in sunspot number reflect underlying fundamental long-term changes in the sun itself.

1. INTRODUCTION

Because the solar wind drives both geomagnetic activity and aurora, the past history of the wind can be inferred from records of these phenomena. Other recent work has dealt with the periods of the Maunder minimum in the last half of the seventeenth century, and with the twentieth century [Suess, 1979; Feynman and Crooker, 1978]. Here we study the period from the last part of the eighteenth century until about 1870 using auroral data, and seek for evidence of changes in the solar wind on time scales long compared to a single solar cycle.

In the remainder of this introduction we review the work that has already been done on inferences concerning seventeenth and twentieth century solar wind. In section 2 of the paper we discuss eighteenth and nineteenth century auroras and also include a short review of the current state of knowledge concerning the relationship of auroral sightings to interplanetary and solar parameters. Section 3 describes the changes in the positions of the auroras that have been observed during the current epoch so that they can be compared to changes in earlier epochs. The final section presents our conclusions.

Auroral and 

$^{14}$C data from the Maunder minimum period have been considered by Suess [1979], who concluded that the solar wind speed and/or the interplanetary field intensity were low and not variable. The history of the solar wind from 1868 to 1967 has been studied through the use of Mayaud's [1973] index, which increased markedly between 1900 and 1960 [Russell, 1975; Svalgaard, 1977]. The aa index has been related to the solar wind velocity and southward component of the interplanetary magnetic field for the last solar cycle [Crooker et al., 1977]. When this relationship was extrapolated to the extreme magnetic calm of 1901 [Feynman and Crooker, 1978], it yielded the result that the solar wind at that time
could not have simultaneously fitted into the range of velocities and southward fields that exist today. Either the yearly average velocity was less than 200 km/s, or the average southward field was a factor of 3 lower, or both quantities changed; that is, the solar wind has changed markedly during this century.

These indications of significant changes in the behavior of the solar wind underscore the importance of and the need for additional studies to clarify the nature of this behavior over long periods of time.

2. EIGHTEENTH AND NINETEENTH CENTURY AURORAS

The sunspot cycle from 1610 to 1975 is shown in Figure 1, taken from Eddy et al. [1976]. The accuracy of the data is discussed at length by Eddy [1976]. He rates the annual data post 1848, which was collected from a network of stations, as 'reliable,' and that from 1818 to 1847 as 'good.' The annual numbers from 1749 to 1847 are considered less reliable the earlier the year. Eddy restudied the values from the seventeenth and early eighteenth centuries, and although the yearly numbers are not accurate, the general trend is well established. The Maunder minimum appears in the top panel of Figure 1. The second panel shows three weak cycles between 1798 and 1833, and the third panel shows a series of weak cycles centered on about 1900. These three periods of reduced activity around 1710, 1810, and 1900 led to the hypothesis that there was an '80- to 100'-year cycle in sunspot activity, sometimes called the 'Gleissberg cycle,' though the term 'Gleissberg variation' may be more appropriate in view of the rediscovery of the Maunder minimum. It has been suggested [Feynman and Crooker, 1978] that the gradual increase in geomagnetic activity at solar cycle minimums from 1901 to 1954 indicates that the solar wind exhibited a change with the Gleissberg cycle. If this is the case, then the wind should also show some pattern
ANNUAL MEAN SUNSPOT NUMBERS FROM 1610 TO THE PRESENT

Fig. 1. The sunspot cycle 1610–1975, from Eddy et al. [1976].

of change centered on early 1800's. Auroral data can be used to examine this question.

Figure 2 shows the yearly auroral frequency data from Sweden for the years 1721–1876. The data were collected by Rubenson [1879, 1882], the Director of the Central Meteorological Institute of Sweden, from publications, newspapers, journals, and manuscript sources. Since the series was very carefully collected by a single knowledgeable person and refers to a geographically restricted area, it forms as nearly consistent a set of data as can be available from this period. The auroral frequency in Figure 2 is defined as the number of nights per year on which auroras were seen. If more than one observer reported the same auroral display, it was nevertheless counted as a single aurora. The year is counted from July of one year to June of the next. Since there is little darkness in Sweden during the summer months, there are very few auroras reported for May, June, and July.

We note the paucity of auroras during the period from 1809 to 1815. In the entire period of 156 years of data, only 14 years had fewer than 10 auroras, but of these, 6 occurred in this period of 7 years, an indication that the solar wind was in some sense 'weak' around 1810, as it had been at the end of the 1600's and would be again in 1900. The functional relationship between the parameters of the solar wind and the magnetic activity and auroras which are driven by the wind is not accurately known at this time. As is noted below, however, we can at least qualitatively use the terms weak and strong as descriptors of the relative magnitude of this functional relationship.

Geomagnetic activity and auroral sightings are closely correlated. Figure 3 shows the comparison between the auroral frequency in Sweden and the aa index for the years 1868–1876, for which the Rubenson catalog and Mayaud tabulation overlap. The aa yearly average has been calculated for the July to June year of the auroral data. The correlation coefficient is 0.97, indicating a remarkably good fit. Studies of auroral observations at specific sites indicate that extrapolation of the curve to other time periods is hazardous because of the changes in auroral distribution discussed below. However, even though the relationship between geomagnetic activity and auroral sightings may change, we may still expect the number of auroral sightings, like geomagnetic activity, to depend on some function of the solar wind velocity, southward field component, and perhaps other parameters in a secondary way. An unusually low number of auroral sightings is then almost certain to reflect low values of some of all of these solar wind parameters.

In order to examine the period around the beginning of the nineteenth century more closely the geographic distribution of Swedish auroras is shown in Figure 4. Rubenson has divided Sweden into four geographic latitude regions: from 55°N to 58°30'N, from 58°30'N to 61°30'N, from 61° 30'N to 65°N, and from 65°N to 70°N. He reports on the number of auroras seen in each region so that in contrast to the count shown in
Figure 2, if an aurora is seen over a wide enough distribution of latitudes, it will be reported once for each district in which it is seen. Figure 4 shows a comparison of the number of auroras reported for the two most northerly districts taken together to the number reported for the two most southerly districts taken together. We have plotted the ratio of the difference between the number of northern and southern auroras to the sum. Thus +1 indicates that auroras occurred only in the north, and −1 indicates that they occurred only in the south.

The figure shows a remarkable pattern of auroral occurrence. From 1807 to 1822, except for 2 years, the auroras were seen predominately in the north. The second of the three weak solar cycles of the Gleissberg 1800 minimum spanned the period from 1810 to 1822, but the cycle began with a broad minimum in 1807. Hence this period of predominately northern auroras spanned the second weak solar cycle. This period of extreme northerly auroras is consistent with the notion of a weak solar wind resulting in relatively weak auroras occurring in a contracted oval. It is interesting to note that the $^{14}$C count was anomalously high during the early decades of the 1800's [Damon et al., 1978], which also is consistent with a weak solar wind.

There are two periods before 1807 and after 1822 when auroral frequency was somewhat more equally distributed between the north and the south in that auroras were seen about 1 to 3 times more frequently in the north. The first of these periods started in 1793 or 1794. This is 2 or 3 years before the 1797 minimum, and the sunspot number at that time was about the same as it was for the maxima of the first two weak cycles. The second period went from 1823 to 1837. This is from a solar minimum to a year after the large maximum of 1836. Thus there does not seem to be a close relationship between the maximums or minimums of sunspot cycles and the onset, ending, or changes within the period of northerly auroras from 1793 to 1837.

During the 23 years before 1793 and the 20 years after 1837, Figure 4 shows that almost all the aurora that were seen appeared only in the south of Sweden. At that time the frequency of aurora in all of Sweden was within the normal range (see Figure 2). The first period of southern auroras ended a year after the 1796 maximum of the sunspot cycle, and the second southern auroral period began a year after the 1856 minimum. From 1858 to the end of the series, auroras occurred about equally in the north and south.

From 1771 to 1876 we have then a period of 106 years, or about 9½ solar cycles, during which the Swedish auroras underwent a well defined systematic change in the pattern of occurrence, first appearing only in the south, moving north, then appearing only in the south again, and finally moving to mid-Sweden.

It had originally been our intention to also use the entire auroral catalog of Tromholt [1902] of Norwegian auroras as a comparison data set. A careful examination of this data, however, revealed several problems. Almost all of Norway is geographically 'north,' according to the division of Rubenson's Swedish data. Also, from 1787 to 1816, the ratio of the number of auroras seen in Norway to the number seen in Sweden is typically below 0.2, suggesting some difficulty in the Norwegian data during that period. The small Norwegian numbers also lead to poor statistics. From 1826 to the end of the series the ratio of reported auroras in Norway to Sweden typically ranged from 1 to 3. Cutting the Norwegian data into northern and southern regions at a line which is geographically $2^\circ$
north of Rubenson's cut, we find that from 1826 to 1839 the auroral ratio \((N - S)/(N + S)\) was >0 for 10 of the 14 years, from 1840 through 1858 it was \(\approx 0.05\) for 14 of 18 years, and after 1858 to the end of the series in 1876 there were only 2 years when the ratio was not >0. Thus where the Norwegian data are sufficiently ample, they tend to confirm the results obtained from the Swedish data.

A change in the pattern of observed occurrence of auroras can be due to several factors, including a change in the number of character of observers, a change in the earth's dipole field, or a change in the solar wind driving the auroras. From the Swedish data in Figure 4 it appears very unlikely that the observed change of pattern can be due to either a change in observers or in the earth's dipole field. Sweden, during the period covered, was a highly developed and major continental power, and there are no historical events which lend themselves to a change of observational patterns of the required scope. A change in the earth's dipole field of the character required can be ruled out. Barraclough [1974] had calculated the latitudes and longitudes of the northern geomagnetic pole of the centered dipole field as 79.9° and 307.2°, respectively, in 1750, 78.7° and 296.0° in 1850, and 78.8° and 294.8° in 1890. The effect of this motion on a point in Sweden which is at about 60°N geographic is to move it from about 60°N geomagnetic in 1750 to about 61°N geomagnetic in 1850. This 1° or 2° shift is insufficient to cause large effects in the auroral occurrence patterns. Furthermore, the dipole motion was not cyclic as would be required to produce the observed change in occurrence pattern; the dipole drifted in the same direction throughout the period.

The nonlocal nature of the change in auroral occurrence frequency is confirmed by comparing the Swedish data with data from New England, collected by Loomis [1866], who was one of the pioneers of American auroral research in the nineteenth century. His data were culled from published sources such as journal articles and observatory reports. Figure 5 shows the number of auroras reported per year in Boston and New Haven. We note that from about 1792 to about 1836 almost no auroras were seen, but that before and after those times auroral sightings were very frequent. When the data from Sweden and Boston–New Haven are superposed, as in Figure 6, the agreement is remarkable. When auroras were seen predominately in southern Sweden, they were also seen in New England, which is at a more southerly geomagnetic latitude. When they were seen more frequently in northern Sweden, they were rarely observed in New England. Since the change in auroral pattern took place simultaneously at two such widely separated locations, the cause must be in the driver of the aura, the solar wind. It should be noted that the changes to and from southern auroras were remarkably sudden, and that therefore the changes that took place in the wind must also have occurred suddenly. We also note that any sunspot cycle changes (Figure 4) are small in comparison with the major changes in auroral frequency pattern that we discussed.

At our present stage of knowledge it is not possible to identify unequivocally the solar wind parameters that must have changed to affect either the annual frequency or latitudinal distribution of auroral observations. The problem can be approached either through direct studies of auroras or of the closely related geomagnetic activity. It was early suggested that geomagnetic activity should be related to some function of the solar wind velocity \(v\) and the southward component of the interplanetary field \(B\) [Rostoker and Ftilthammer, 1967]. Numerous studies have found that high values of indices of geomagnetic activity are associated with southward \(B\) [cf. Hirshberg and Colburn, 1969; Arnoldy, 1971; Tsurutani and Meng, 1972] and that large southward \(B\) occurs in flare-produced disturbances [cf. Hirshberg and Colburn, 1969] and at the leading edges of fast solar wind streams [Hirshberg and Colburn, 1973] which in turn are usually associated with coronal holes [Newperti and Pizzo, 1974]. However, a recent study [Burlaga and King, 1979] found that about 20% of interplanetary magnetic field enhancements were not
related to either flares or streams. Other studies of the correlations of geomagnetic indices and interplanetary parameters show high correlations with functions of $B_0$ and $v$ as expected, but they also show that other interplanetary parameters such as the magnitude of the interplanetary field or the density may also be involved in a secondary role [Garrett et al., 1974; Maeszawa, 1978]. In any case, however, the matter of identifying the solar wind parameter that causes geomagnetic changes is not as simple as it might first appear, since it has been found that $B_0$ is the most important parameter for processes taking place on the time scale of hours [Foster et al., 1971], whereas $v$ is the most important parameter for changes on the time scale of the solar cycle [Crooker et al., 1977]. It would certainly be premature to speculate if the 80- to 100-year variation is expressed in changes of $B_0$ or $v$ or some unknown combination of both.

The studies of the relation of auroras directly to solar wind parameters have yielded less information. The first difficulty is that most modern studies use DMS (Defensive Meteorological Satellite Program) satellite imagery or particle precipitation to determine the boundary of the auroral oval as defined by the diffuse auroras. The eighteenth and nineteenth century auroral sightings are of discrete auroras, of course. Although the discrete auroras are imbedded in the auroral oval, a knowledge of the oval’s response to solar wind gives at best only approximate information on the discrete auroras. A second difficulty is that although it may be expected that the auroras respond to the same solar wind parameters as geomagnetic activity, the studies of oval boundaries to date use only $B_0$. As southward $B_0$ increases, the oval boundary move equatorward [Kamide and Winningham, 1977; Pike and Dandekar, 1979], and the probability of substorms also increases [Kamide et al., 1976; Sheeley [1978] showed that for the 1973–1974 Skylab period, which was dominated by the recurrent geomagnetic storms typical of the descending portion of the sunspot cycle, the equatorward oval boundary was related to the southward $B_0$ and hence to solar flares and coronal holes. It therefore appears that although we know that auroral sightings will be dependent on solar wind $B_0$ and we expect them to depend on some function of $v$ and $B_0$, neither the work on the auroral oval nor the work on geomagnetics allows us to speculate on whether the change in the solar wind was primarily expressed in $B_0$ or in $v$ or in both during Gleisberg variations.

3. AURORAL DISTRIBUTIONS IN THE CURRENT EPOCH

It is interesting to compare the data in Figure 2 with studies of changes in the distribution of auroras that occurred between the minimum and maximum of a recent solar cycle. An estimate of the ratio of northern to southern auroras for that cycle can be obtained from the work of Sandford [1968], who studied the latitudes of the auroras in the southern hemisphere during the 1958–1959 IGY and the 1963 IQSY. Using the geomagnetic latitudes corresponding to the Swedish data, we find that discrete auroras appeared equatorward of 60° on about 40% of the nights in the IGY 1958–1959. In the band between 60° and 70°, auroras were visible on 60% of the nights. Therefore if all auroras that occurred were observed, the ratio comparable to that of Figure 4 would be 0.2. For the IQSY, in 1963, discrete auroras occurred equatorward of 60° on 20% of the nights and between 60° and 70° on 80% of the nights, giving a ratio of 0.6. Thus though the ratio was higher for the IQSY than for the IGY, the change was not nearly as dramatic as that between, say, 1810 and 1850. Unfortunately, there are no solar wind data for the IQSY, so that the changes in auroral occurrence frequency reported by Sandford can not be directly related to solar wind changes.

Two effects of changes in the solar wind that could cause a change in auroral occurrence frequency at a particular latitude are changes in the number of geomagnetically disturbed days per year and changes in the position of auroras for the same level of geomagnetic activity. Both types of changes occurred between the IGY and IQSY.

Feldstein and Starkov [1968] show the location of the northern hemisphere auroral oval edges as a function of magnetic activity as measured by the $Q$ index for the IGY when the average $aa$ index was 29 and for the IQSY when the average $aa$ was 21. During both periods the width of the oval increases with increasing geomagnetic activity. The equatorward boundary expanded southward in both periods, whereas the poleward boundary moved northward with $Q$ except for $Q$'s of 1 or 2 during the IGY. The excursion of the southward boundary was about 6° during the IQSY and 10° during the

---

Fig. 6. A superposition of Figure 5 and the ratio of northern to southern auroras in Sweden showing that the auroras disappeared from the Boston–New Haven area at the same time as they become more common in the north of Sweden. Later they reappeared in Boston–New Haven and were not seen in the north of Sweden.
IGY. The southward excursion was more marked than the northward excursion during the IGY. Feldstein and Starkov state that the center line of the oval moved northward 1.5° between the IGY and the IQSY.

Sandford [1968] has also reported on changes in the latitude of southern hemisphere auroral occurrences as a function of geomagnetic activity. He studies discrete auroras occurring at midnight as a function of the local K index and, in agreement with Feldstein and Starkov, found marked differences between the two years. Sandford found, for example, that at K = 0 the most common latitude of auroras was 75° in 1958-1959 and 70° in 1963. In fact, no auroras were seen at 75° for K = 0 in 1963. For higher K the distributions were much wider, and auroras were seen at much more equatorward positions in 1958-1959 than in 1963. For K = 4, auroras occurred at 60° for 50% of the IGY nights and 25% of the IQSY nights. Sandford also found a tendency for the position of most frequent occurrence to move equatorward with increasing K during the IGY but not during the IQSY.

The studies of Feldstein and Starkov and of Sandford shows that the position and dynamics of the auroral oval can change noticeably in a period as short as 4 years, at least during the current epoch. The changes that took place during the eighteenth and nineteenth centuries were, of course, much more extreme.

4. CONCLUSIONS

1. The latitude at which auroras occurred underwent a clearly patterned change during the 106-year interval from 1771 to 1876. These changes were considerably greater than those reported between the IGY and the IQSY. We conclude therefore that the variation of the configuration of the magnetosphere during the eighteenth and nineteenth centuries was much greater than has been observed since space observation began.

2. The change in auroral occurrence pattern was due to the solar wind. The wind was strong in some sense in 1771-1793, became weaker with a minimum in 1809-1815, and returned to a strong state in 1837-1858.

3. The changes in the strength of the solar wind can be remarkably rapid. This is implied by the enormous changes of auroral appearance pattern which occurred over a period of 2 to 3 years several times during the 106-year interval. The times of these changes (circa 1793, 1839, and 1858) have no apparent relationship to phases of the sunspot number cycle.

4. Because the solar wind exhibited minima in 1901, in 1809-1815, and during the Maunder minimum, it appears likely that the wind varies with the Gleissberg variation. This observation leads to the conclusion that the strength of the sunspot cycle and the solar wind are both related to some underlying 80- to 100-year variability of the sun. This must be taken into account in models of the solar cycle in that models must now explain simultaneously not only the 22-year cycle in sunspot number but the equally fundamental 80- to 100-year Gleissberg variation in the strength of the solar wind.

The long-term underlying solar variability has caused significant changes in the earth's environment on time scales of from 2 to 100 years in the recent past, and we can expect that it will continue to do so in coming years.

Acknowledgments. We dedicate this paper to Robert Rubenson, whose auroral catalog served as the data base for much of this paper.

It was a careful and painstaking piece of work. This paper was written on the hundredth anniversary of the publication of the first part of the catalog, so that this appears to be an appropriate time to memorialize his work.

We thank Nancy Crooker and George Sinco for stimulating conversations and David Hardy for his comments on the manuscripts. One of us (J.F.) was supported by Air Force Geophysics Laboratory contract F19628-79-C-003. The Editor thanks J. A. Eddy for his assistance in evaluating this paper.

REFERENCES


Feldstein, Y. I., and G. V. Starkov, Auroral oval in the IGY and IQSY period and a ring current in the magnetosphere, Planet Space Sci., 16, 129, 1968.


(Received October 31, 1979; revised February 19, 1980; accepted February 19, 1980.)