SECULAR VARIATION OF THE AURORA
FOR THE PAST 500 YEARS

S. M. Silverman
Physics Department
Boston College
Chestnut Hill, Massachusetts

Abstract. Knowledge of solar variability and its effects on the Earth is essential since the Sun affects almost every aspect of our lives. Direct observations of the Sun, usually of sunspots, with some continuity, exist only since about 1700. Understanding of long-term solar variability must then depend on proxy data, such as visual auroral observations, measurements of magnetic activity, and the radiocarbon record. These also give us information on the interaction between the solar wind, the interplanetary field, and the terrestrial magnetosphere, as well as, for the radiocarbon record, heliospheric conditions. This paper uses a data base of visual auroral observations for a period of about 500 years, from 1450 to 1948, comprising about 45,000 observations, in addition to the well-known sunspot series and the magnetic activity index $aa$, from 1868 to 1990. The secular variation of the aurora is examined and compared, where possible, to sunspot data and magnetic activity data. Blackman-Tukey power spectra are used to determine periodicities. The study confirms the variability of the periodicities in both frequency and amplitude. In particular, the well-known 11.1-year cycle disappears during the Maunder minimum and at the end of the eighteenth and beginning of the nineteenth century. While the 11.1-year period is normally strongly dominant for sunspots, other shorter periods become important, and even dominant, for auroras and magnetic activity. Consequently, the temporal behavior of these three variables differs. Prolonged solar activity minima are clearly evident. In addition to the known Spörer, Maunder, Dalton, and 1901–1913 minima, a previously unrecognized minimum about 1765 is clearly evident in the data. Comparison of the depth of these minima shows that the Dalton minimum may be the deepest, or at least rivals the Maunder minimum in importance. This minimum clearly deserves further study. Combining the polar data base with that of mid-latitudes provides for the first time a globally comprehensive historical record of auroral occurrence. The data provide confirmation of the anticorrelation of auroral occurrence in the polar regions with sunspot activity, as a result of displacement of the auroral oval with changes in solar and magnetic activity. Assuming the validity of some current models of the solar origin of geomagnetic activity, the data provide a basis for understanding the variation over time of the general magnetic field of the Sun, in particular the polar field.

1. INTRODUCTION

Life on the Earth depends on the Sun. This simple statement disguises a considerably complex situation. The Sun itself shows a degree of variability. This is evident in observational measures such as visual and telescopic observations of sunspots and other surface features, as well as in such parameters as the magnetic fields of sunspots or the general magnetic field. Further, the Sun emits both radiation and particles. Travel of the particles is influenced by the interplanetary magnetic field, and, in the vicinity of the Earth, by the terrestrial magnetosphere. If we wish to understand solar behavior and the interactions between Sun, interplanetary field, and terrestrial magnetosphere, we must look to measures of solar activity directly derivable from solar observations, such as sunspots, and to terrestrial observations (the end results of the travels of particles and radiation to the Earth).

Solar variations, to be properly understood, must consider temporal intervals of centuries or even millennia. Many, perhaps even most, scientists have a vague idea that the Sun varies with a period of 11 years and that this behavior has continued since the Earth’s beginning. In fact, the Sun’s variability is described by many periods whose amplitudes change over time. There are, furthermore, as pointed out most recently by Eddy [1976], intervals of prolonged minima in solar activity. Thus a proper understanding of the Sun must begin with the determination of solar periodicities, their changes over time, and any unusual features in these parameters. The results provide data which any physical model of the Sun must explain.

In determining solar periodicities and their behavior we are hampered by the relatively short period within which detailed, “proper,” scientific observations have been carried out. We must depend for longer intervals on less well defined, less resolved, more sporadic observations of only a few parameters, including proxies for solar activity. Primarily, these are observations of sunspots, visual auroral observations, magnetic activity, and radioisotope measurements, especially those of $^{14}C$. 
Observations of naked eye sunspots, that is, especially large ones, have been recorded sporadically for some thousands of years. Currently, however, the parameter used is not the number of sunspots themselves but the sunspot number, a derived index introduced by Wolf at Zurich in 1849. This is defined as \( R = 10g + f \), where \( g \) is the number of sunspot groups and \( f \) is the number of individual spots. The sunspot number is known with high reliability from 1848 when Wolf began systematic observations, with good reliability from 1818 to 1847, with questionable reliability between 1749 and 1817, poor reliability between 1700 and 1748, and sporadic reliability between 1610 and 1700 [McKinnon, 1987].

Visual auroral observations can be found discussed in classical Greek and Roman times, in the works of Aristotle and Pliny, where they were denominated chasma (see, for example, Silverman [1962]), a term which continued to be used as late as the seventeenth century. Auroral observations, the fire in the sky, are often found in historical annals as omens or portents in both European and Asian sources. Though sparse they are sufficient to define major trends of solar behavior over periods of millennia [see Siscoe, 1980; Krivsky and Pejml, 1988]. Auroral data, combined with sunspot data (observed or estimated), can also be used to derive estimates of solar wind properties and magnetic activity over intervals much greater than the relatively short term of in situ observations [Silverman, 1986a]. The aurora is produced by particles emitted by the Sun which have been modulated by the interplanetary field and stored in the magnetosphere, with a latitude of deposition dependent on magnetospheric convection currents. Basically, the aurora reflects disturbances which originate largely in solar variations. The effects of interactions with the terrestrial magnetic field can be seen in the annual variation, giving maxima at the equinoxes and minima at the solstices [see Silverman, 1986b]. As used in this paper, an auroral record is an observation at a given place on a given date, without reference to the nature of the aurora itself, or to the length of time for which it is observed.

Magnetic activity has been characterized in many ways. Perhaps the earliest measure of the past two centuries was the variation of the declination of the geomagnetic field [see Fritz, 1889] used extensively by Wolf, who found a linear variation between sunspots and the average daily variation of declination. Measurements of declination can be found back to the end of the eighteenth century, with good reliability from 1700 to 1985, shows a variation with a period, at least since 1826, of about 11 years. Modulations with much longer times are also evident. In particular, we find prolonged solar activity minima around the beginnings of each century, which can be related to a period of the order of 88 years [Feynman and Fougerere, 1984]. A characteristic of this series is that the minima of the sunspot numbers are always less than about 12, generally much smaller than this, and that these minima are always less than the lowest maximum. In contrast to this, magnetic activity and auroral occurrence can show long trends in which

The best known time series reflecting solar activity is that of sunspot number (Figure 1) [Eddy et al., 1976; McKinnon, 1987]. The sunspot number, over the interval 1700 to 1985, shows a variation with a period, at least since 1826, of about 11 years. Modulations with much longer times are also evident. In particular, we find prolonged solar activity minima around the beginnings of each century, which can be related to a period of the order of 88 years [Feynman and Fougerere, 1984]. A characteristic of this series is that the minima of the sunspot numbers are always less than about 12, generally much smaller than this, and that these minima are always less than the lowest maximum. In contrast to this, magnetic activity and auroral occurrence can show long trends in which

![Figure 1. Yearly mean sunspot numbers, 1700–1975 (Eddy et al., 1976).](image-url)
the period subsequent to 1901 the envelope of both minima and maxima shows a continuing upward trend. A consistent set of magnetic data prior to 1868 is not available, so the overall behavior for the nineteenth century cannot be determined. Auroral observations can be used, with appropriate caveats, for periods of centuries or, to some extent, for millennia. This paper considers auroral variations for the past 500 years.

The auroral oval, the geographic region of maximum auroral occurrence, serves as a useful dividing line for discussion. The oval reflects the region in which the Earth's magnetic field shifts from closed to open lines. The polar aurora (the aurora of the oval and that within the polar cap, that is, poleward of the oval) is behaviorally different from that at latitudes equatorward of the oval. Most, but not all, data equatorward of the oval consist of observations at geographic latitudes south of about 55°. This region, generally, will be considered as mid-latitudes in this paper.

1.1. Mid-Latitudes

Because of the lower population density at high latitudes, observations of the aurora prior to the nineteenth century occur primarily in mid-latitudes, with occasional references to low-latitude observations. The aurora was noted as a phenomenon in the classical works of Aristotle and Pliny, though only descriptions of these “meteors” (that is, atmospheric phenomena) are presented, and the possibility of confusion with other atmospheric phenomena, such as bright airglow nights, exists [Silverman, 1962]. Observations of atmospheric phenomena which can be identified as auroras occur in histories and chronicles over several millennia in both Europe and Asia. Splendid or magnificent auroras were often noted as omens or portents, and records of these auroras consequently survived in histories.

The study of the aurora as a distinct natural phenomenon was resurrected with the development of science in the rennaissance. Tycho Brahe noted the presence of “chasmata,” the Aristotelian term, at the end of the sixteenth century, in connection with his astronomical observations. Many observations of the aurora can be found in the popular literature at the end of the sixteenth century, especially in the ephemera, popular news sheets describing individual special events, which followed after the development of the printing press. The last half of the seventeenth century saw a marked reduction, though not a complete absence, of auroral observations. In England, for example, Halley noted, in 1716, that no display of aurora had been noted there since 1574. The auroras of the early eighteenth century may be considered as the origin of modern auroral studies. Discussions of the early history of the aurora may be found in the works by Siscoe [1980], Eather [1980], and Brekke and Egeland [1983].

The return of the aurora to mid-latitudes led de Mairan [1733] (enlarged second edition in 1754) in his seminal work on the subject to the question of long-term variations of the aurora. In his initial catalog of auroral occurrences over the previous millennia he noted a number of “reprises,” returns, or resumptions of the aurora over this period. Siscoe [1980] has summarized a number of the studies that searched for periodicities in auroral occurrence. Thus, for example, Hans teen in 1831 inferred a period of 95 years; Olmsted in 1856 gave 60 to 65 years as the intervals between great auroras; and Wolf and Fritz preferred a period of about 55 years. Other work also indicated periods in the region of 80–90 years [see Feynman and Fougere, 1984], and even periods as long as 200 or 400 years. With the recognition of the sunspot cycle of 11.1 years, about 1840, this period was also looked for and found in the auroral data. I note parenthetically that Ritter [1803] had noted an apparent periodicity in eighteenth century auroral data of between 9 and 10 years, which he related to the 18½ year lunar period. A connection to this lunar period had earlier been suggested by Hoslin in 1784 [Fritz, 1881, p. 117] and remarked on by Böckmann [1801]. A 9–10 year period is consistent with one of the periods derived a century later by Schuster [1906] for sunspots. I will note that the term “periods” must be treated with caution, since the data, both for auroras and sunspots, show significant variation in both period and amplitude over time [see Silverman, 1989]. Similar variations in period and amplitude are found for long solar periods in the radiocarbon record [Sonett, 1984; Sonett and Finney, 1990]. Scientists around the turn of the nineteenth century were busy looking for connections between aurora, weather, meteors, astronomical positions, and just about everything else.

Siscoe’s [1980] review focused on the longer-term variability of the aurora as derived from observations over several millennia in both Europe and Asia and the conclusions that could be drawn from these as to solar variability. Because of the absence of detailed solar data prior to about the eighteenth century, proxy data must be used for studies of early solar variability. Auroral data are often used for this purpose, though it must be remembered that there are differences between the two data sets. This will be shown in more detail later.

In the present paper I consider only the past 500 years,
from 1500 to the present, the interval for which a reasonable number of observations are available, with a passing reference to observations from 1450 to 1500.

1.2. Polar Regions

The concept of the auroral oval, that is, a region of the globe with a maximum of auroral occurrence, cannot be defined in simple terms nor can it be confined at all times to only one region. The position of the oval is at higher latitudes on the dayside than on the nightside (see, for example, Lassen et al. [1888] for a brief review of this and succeeding distinctions). Thus some stations may be equatorward of the oval, in the oval, or poleward of the oval, depending on the time of day. Furthermore, the oval may look different depending on whether we are using visual observations, spectral observations at 391.4 nm, 557.7 nm, 630.0 nm, or some other spectral region, or measurements of electron or ion precipitation. The oval moves equatorward with increasing magnetic activity (for statistical summaries of this variation for electron precipitation, see Hardy and Gussenhoven [1985], and for ion precipitation, see Hardy et al. [1989]). In this paper we are concerned only with visual auroral observations.

The polar aurora has been intensively studied since the beginning of the International Geophysical Year in 1957. Prior to this, only some scattered studies are available. Tromholt [1881] published an extensive study based primarily on observations at Godthaab, Greenland, from 1865 to 1880 made by Samuel Kleinschmidt. Tromholt also included observations from Stykkisholm, Iceland, from 1846 to 1873, as well as scattered observations from other stations. Tromholt anticipated some of the recent conclusions, noting, for example, the variation of the auroral oval in latitude with time. The southward movement of the oval with increasing sunspot activity is at least implicit in his observation of the anticorrelation of auroral activity at Godthaab with sunspot activity.

Meteorological observations in Greenland were made as early as the 1810s. Auroral notations can be found as early as the 1840s for two or three stations in Greenland.

Auroral observations were also made by the many British ships searching for the northwest passage and wintering over in the Arctic in the period following the Napoleonic wars. Thus a fair sized body of observations exists for the period following 1819. These observations have the advantage of being more randomly spaced geographically than the Danish observations in Greenland and Iceland, carried out at fixed stations. The Danish data, on the other hand, are more useful for determining secular variations.

2. THE DATA BASE

2.1. Mid-Latitudes

The most used and cited data collection for auroral observations prior to August 1872 is that of Fritz [1873]. An almost contemporaneous catalog is that of Lovering [1866–1871]. Both catalogs contain errors, but these are not sufficiently numerous to prevent their use for statistical purposes. Both, as might be expected, have considerable overlap. Geographically specialized catalogs for Sweden [Rubenson, 1879, 1882] and Norway [Schroeter, 1902] are also available. A listing of auroras seen in Europe below latitude 55° was given by Angot [1897] in his book on the aurora. In more recent times, Link [1963, 1965] has prepared carefully annotated catalogs for times before 1700. Link's catalog is especially useful since he gives the text of the observations as well as the dates. Of the catalogs prepared in more recent times I note only the following which I have used. Křivský and Pejml [1988] have published a catalog, giving dates only, for the interval 1000–1900, and restricted to auroras observed for latitudes <55°. This catalog was used by Charvátová-Jakubcová et al. [1988] for a study of auroral periodicities over the millennial interval using Fourier amplitude spectra and power spectra. Laysha et al. [1989] have published a catalog of auroras observed within the then boundaries of the Soviet Union from the fourth to the eighteenth centuries. The compilations of Keimatsu [1976, and references therein] of auroras observed in China, Japan, and Korea have also been used in the present study. In addition to these sources I have used a variety of journals, primarily from the 1870s to the 1950s.

Data for New England were taken from a compilation by S. M. Silverman based on several sources (for a description of observational systems from the end of the eighteenth century to 1870, see Fleming [1990]). The primary source used here was the meteorological registers kept by voluntary observers in the network established by the Smithsonian Institution and continued, primarily, by the Army Signal Service and the Weather Bureau. Observations were also published in the Monthly Weather Review and in the series, Climatological Data of the United States. Some additional data for individual locations were also used, as, for example, the observations of Bentley in Jericho, Vermont, from 1881 to 1931 [Silverman and Blanchard, 1983]. Some data were also taken from the auroral catalogs published by Lovering [1866–1871] and Fritz [1873]. The New England data base comprised about 20,000 records, from 1741 to 1948. Some scattered records prior to 1741 exist but are not utilized here. Altogether several hundred locations are involved. Their geographic coordinates range from about 41° to 45°, and the corresponding corrected geomagnetic coordinates from about 53° to 57°.

2.2. Polar Regions

Data for Godthaab are available, wholly or in part, for the years 1841–1846, and from 1865 on. For Stykkisholm, Iceland, data begin in 1846 and continue until 1939. For Jacobshavn, there are data for 1840–1851 and from 1875 to 1960. Data for Greenland stations administered by the Danish Meteorological Institute began late in 1872 and are available in the yearbooks of the institute.

Data for Iceland are included in the Danish yearbooks from 1873 through 1919. Though still a Danish colony until World War II, Icelandic data began to be published separately subsequent to 1919. Detailed daily data for selected stations continued to be published until 1923. After that date, only monthly summaries of auroral data were published until 1939. These monthly data are useful for certain purposes.
when only individual stations are needed, such as studies of the secular variation, but are not helpful when all stations are considered together.

The geographic coordinates for the Greenland and Iceland stations range from about 61° to about 73°. The corrected geomagnetic coordinates range from 61° to 80°.

Data from British Arctic voyages are available in volumes published by the British Meteorological Council [1885–1888] abstracting meteorological data from the log books of ships engaged in Arctic exploration. Data are also available from books published by explorers in Canada and elsewhere and from scattered data from other sources.

The total number of records used involving daily observations for Greenland and Iceland was 26,112, of which 620 were for the period 1840–1851, 24,755 were from the Danish Meteorological Institute yearbooks for 1873–1967, and 737 were from the Icelandic Meteorological Institute yearbooks for 1920–1923. Other Arctic records contained of the order of 2000 records.

All together, for mid-latitudes and the polar regions, about 45,000 records were used in the preparation of this paper.

### 2.3. Counting Auroral Occurrence

The basic datum in these analyses is the number of days on which aurora was observed in a given time interval, usually months or years, without regard to the number of locations or observers reporting an aurora. Hence if several observers at several different locations reported an aurora on a given day, only one aurora is counted for that day.

### 3. SECULAR VARIATION

#### 3.1. Mid-Latitudes

By combining the mid-latitude time series we obtain a sequence of data for almost 500 years, which can be examined as a whole or in segments. The time series will be examined for periodicities using Blackman-Tukey power spectra [see Silverman and Shapiro, 1983]. The significance of peaks can be estimated by taking the ratio of the peak to the estimated continuum. The 5% level is at a ratio of 3.0 for lags of 3000 and 1200, at 2.6 for a lag of 900, and at 2.4 for a lag of 600 (the values used in the figures shown here). The 1% level is at a ratio of 4.6 for lags of 3000 and 1200, at 3.8 for a lag of 900, and 3.3 for a lag of 600. Comparisons with sunspot and magnetic data, where applicable, will also be carried out.

**3.1.1. The interval 1500–1948.** The number of days per year on which auroras have been recorded in mid-latitudes over the entire time interval 1500–1948 is shown in Figure 3. Some smoothing has been provided by using 5-year moving averages. A logarithmic scale has been used because of the wide range of the numbers. The number of days on which auroras were recorded in 1500–1549 totaled 20 (combining the data given by Křivský and Pejml [1988] and Loysha et al. [1989]) and is not shown on the figure. By contrast, the number of days on which aurora were recorded in 1500–1549 was 57, and for 1550–1599, 272.

Overall the interval from 1450 to 1948 is characterized by a rising trend, corresponding to the increase from the Spörer minimum at the end of the fifteenth century to the present. This general increase is consistent with that reported by Křivský and Pejml [1988] as seen in both the auroral and 14C data. There is a gradual increase from 1500 to about 1580, followed by a gradual decline, which includes the Maunder minimum, till about 1700. There is a minimum around the turn of the century, about 1596, though it is not very marked. Minima occur in the vicinity of 1700, 1800, and 1900, as well as about 1765. These will be discussed in more detail in a later section. The Dalton minimum, around 1800, is especially marked. The relative importance of the various minima may be approximated by taking the ratio of the number of days at the maximum preceding each minimum to the number of days at the minimum. Thus we obtain for 1570 compared to 1596 a ratio of 7.3, for 1663 compared to 1702 (or 1712) a ratio of 22 (1702 rather than 1701 has been used to avoid division by zero), for 1737 compared to 1762 a ratio of 17, for 1788 compared to 1810 or 1811 a ratio of 266; and for 1871 compared to 1912 a ratio of 19. Thus, of the four minima since 1650, three give a ratio of about 20, and the Dalton minimum gives a much higher ratio, or a much deeper minimum. The same result, giving a much deeper minimum for the Dalton minimum, is obtained if we use an 11-year moving average. This results from a much higher initial value than, say, the Maunder minimum. It would appear from this result that more attention should be paid to the interval around the Dalton minimum, which appears from these results to be a more, or at
least as significant as the Maunder minimum. This is especially so since the growth of science resulted in much more and more diverse data being available for this time period.

Figure 4 shows the portion of the spectrum of monthly auroral occurrence for the entire interval from 1500 to 1948 for periods greater than about 30 years. The peaks here are, unfortunately, not sufficiently well resolved to provide definitive information either to define the period with exactitude or for significance. They are nevertheless of interest for comparison with other work. The peak at 83 years is probably related to that determined as 88 years by Feynman and Fougeres [1984, and references therein] and based on spectral analysis of a data set of about 1000 years for the interval 450 to 1450. A peak at 83 years, also based on millennial data, had been proposed in the nineteenth century by Wolf and was further discussed by Fritz [1893] (see also Fritz [1889]). The peak at 55 years had also been proposed by Wolf [see Fritz, 1893] as a prominent period. A definite and clearly significant peak is found in the present data set at 33 years. Charvátová-Jakubcová et al. [1988] have reported periods in the range 60–100, 43–47, 34–37, and 30–31 years in less well resolved spectra based on an auroral data set covering the interval 1001–1900 and the subintervals 1001–1500 and 1501–1900. Except for the peak in the 43–47 year period these are consistent with the present results. These periods are attributed to the influence of the outer planets on the Sun's motions by many investigators from Wolf to Charvátová-Jakubcová et al. This hypothesis is generally disfavored today. We may note also that the periods, if taken as 33.3, 55.5, and 88.8, are multiples of the well-known 11.1-year period. If they are related, the question then comes as to why the fourth, sixth, and seventh multiples are absent or at the least of low amplitude.

The spectrum for the interval from about 5 to 35 years is shown in Figure 5. The spectrum in this region is dominated by the 11.1-year peak, with smaller peaks at 24.4, 18.2, 14.9, 9.4, 8.1, and 5.6 years. There are also indications of peaks at 12.8 and 10.2 years, whose significance cannot be established because they occur close to the overwhelmingly dominant 11.1-year peak. These peaks are consistent with those found in sunspots for the interval 1749–1957 [see Currie, 1973] and magnetic activity for the interval 1868–1967 [see Currie, 1976].

To allow for a more detailed examination, the auroral occurrence frequency for each century separately is shown in Figures 6–9. Each figure extends for 10 years past the end of the century to clarify the behavior around the turn of the century. In these figures I have also plotted an 11-year moving average to make the longer-range trends more obvious. It is clear from these figures that the secular variation of auroral occurrence is complex.

Since there are minima around the turns of the century, it is convenient to carry out the more detailed examination of subintervals a century at a time.

3.1.2. The interval 1500–1599. Figure 6, for the sixteenth century, is derived from data given by Fritz, Keimatsu, Krivsky, Link, Loysha, and Silverman. The figure is similar to one given by Siscoe [1980] but includes additional data. The first half of the century is sparse in auroral observations. That this is not simply due to a lack of observers is indicated by comparison with the 4C data, which are consistent with the auroral data (cf. Figure 4 by Siscoe [1980]; see also Krivsky and Pejml [1988, Figure 1, p. 64]). It is, rather, the final phase of the Spörer minimum, which began in the fifteenth century (for data prior to the fifteenth century see, inter alia, Siscoe [1980], Krivsky) and Pejml [1988]). This appears to be a considerably more prolonged activity minimum than has occurred at any period from that time to the present. Considering trends only, an increase in activity begins about the middle of the century, rises to a maximum.
about 1580, and then declines rapidly to the mid-1590s. These features, recurring increases and declines and minima near the turn of the century, are typical of the auroral data for the entire interval under consideration in this paper.

As noted above, the first half of the sixteenth century is connected to, and probably part of, the Spörer minimum. The spectrum therefore represents the behavior of the last part of the century. The spectrum, shown in Figure 10, is relatively simple, consisting, in the region shown, of only three peaks, at 9.3, 10.5, and 12.5 years.

3.1.3. The interval 1600–1699. Auroral occurrence in the seventeenth century, shown in Figure 7, is of special interest because it includes the Maunder minimum. Data for this century have been taken from Krivsky, Link, Loysha, and Silverman and thus represent essentially a European sample. The auroral record shows a trend of a clear decline over the century reaching a minimum around the turn of the century. There is, however, no well-defined interval for which the auroral record disappears. There is, rather, this gradual decline to a turn of the century minimum. Comparison with the data for the sixteenth century shows that the level of reported auroral activity in the last part of the century is considerably lower for the seventeenth than the sixteenth century, indicating a considerably reduced level of activity, though not an absence of all activity. Reduced activity levels will be discussed in more detail in a later section of this paper.

The seventeenth century power spectrum, shown in Figure 11, is more complex than that of the previous century. Longer-term peaks, at 20 and 25 years, predominate. The 11.1-year peak now appears. The 12.5-year peak remains as a subsidiary peak. The 9.5-year peak, if present, is hidden by the new 8.3-year peak. Several peaks in the 3–5 year range appear to be as influential as the 11.1-year peak.

This complexity may be due to an aggregation of differing behaviors during the first and last halves of the century. This becomes clear when we consider the spectrum during the Maunder minimum.

3.1.4. The Maunder minimum: The spectrum of the interval 1650–1725, shown in Figure 12, is a much simpler one than that for the century as a whole. Resolution of the spectrum is not as good as that for the overall century. Here the 11-year peak is absent. We are left with longer-term peaks at 15 and 19 years as well as some in the 3-year region.

3.1.5. The interval 1700–1799. The eighteenth century, shown in Figure 8, shows a strong resurgence of auroral activity. Data for this figure are taken from Fritz using his compilation for European latitudes from 45° to 55° geographic. Since only very strong auroras penetrate to latitudes lower than these, the data are generally comparable in coverage to those used for the sixteenth and seventeenth century.

Figure 6. Variation of auroral occurrence in the sixteenth century. Data are from Link [1963], Krivsky and Peiml [1988], Keimatsu [1976], S. M. Silverman (unpublished compilation, 1800–1948), Fritz [1873], and Loysha [1989].

Figure 7. Variation of auroral occurrence in the seventeenth century. Data are from Krivsky and Peiml [1988].

Figure 8. Variation of auroral occurrence in the eighteenth century. Data are from Fritz [1873].
The maximum values are of the order of 7 times greater than those in the sixteenth century. This may be due in part to a greater number of observers at different locations in the later century, since this factor would tend to overcome the effect of local weather conditions. The behavior in the eighteenth century differs from that of the preceding centuries in that we have two distinct cycles during the century. There are distinct minima at the beginning and the end of the century, a behavior similar to that of the preceding centuries. The higher level of activity in the second half of the century is consistent with the sunspot data, where essentially the same behavior is observed. Data for New England, available only from about 1740, show a behavior similar to that of Europe, though the number of days on which aurora was observed is only, for the maximum, about half of those observed in Europe. The number of observers in New England during this interval was very small and restricted to a small number of locations, so that local weather effects would be greater. These factors undoubtedly account for the difference between New England and Europe in this time interval.

The spectrum of the interval 1700–1799, shown in Figure 13, is again complex. The more important peaks are now at 10.0, 13.3, and 18.2 years. Smaller, but noticeable peaks, are at 4.6, 7.4, and 8.3 years. The spectrum represents basically the interval from about 1720 to 1790, since there were relatively few auroras present during the other years. As noted by Schuster [1906], the sunspot numbers for the period from 1750 to 1825 behave differently from those of the period from 1825 to 1900. This will be discussed in more detail below.

3.2. Comparison of Mid-Latitude Auroras, Sunspots, and Magnetic Activity: The Interval 1800 to Recent Times

Comparison of sunspots, auroras, and magnetic activity allows for testing of the reality of inherently common features, such as periodicities derived from solar behavior, and for distinguishing other features dependent on interactions with the interplanetary magnetic field and the terrestrial magnetic field. For these purposes the relatively dense data set after 1800 is suitable.

Auroral activity for New England for the interval 1800–1948 is shown in Figure 9. The data for this time interval are more reliable than for the previous centuries since they are based on a dense set of observations. In magnetic latitude the region is approximately comparable to the European observations, so that we may make at least approximate comparisons with the preceding centuries, as shown in the other figures. The maximum of the averaged values for this century are greater than those for the previous century by about a factor of 1.5. For this century, by contrast with the previous century’s two cycles of activity, we find only a single cycle.
We again find a minimum around the turn of the century but now at a much higher level. For the twentieth century the visual auroral data for this network of stations were published only until 1948. The activity for this time interval is lower than that of the latter part of the nineteenth century. Typically, however, for the centuries we have considered thus far, the activity is greatest in the latter part of the century, indicating a solar period of the order of a century. The solar activity of the years since 1950, as derived from sunspot observations, has been the highest since such observations became available, so that it appears that this pattern has continued. The magnetic data for the late nineteenth century show, except for the unusual minimum about 1879, indications of a declining trend to 1901, followed by an increasing trend to 1960. The results then are similar to, or at least consistent with, the auroral data to the extent that comparison is possible. The magnetic activity index $aa$, however, has remained, for 11-year averages, level within the range of about 20–25.

In order to provide comparisons of the auroral results with those of sunspots, spectra were run for the same time period, 1800–1948, using monthly data for both sets in order to provide better resolution. Magnetic data are not available in a self-consistent form over this interval, so the available $aa$ indices, 1868–1990, while not directly comparable, were used to provide a long time interval with self-consistent data covering much of the same interval as was used for auroras and sunspots.

Power spectra for auroras and sunspots in the interval 1800–1948 are shown in Figures 14 and 15a. There is considerable overall consistency between the two spectra, though there are differences in detail. The 11.1-year peak is by far the predominant peak in both spectra. The ratio of the amplitude of this peak to the next highest, at 5.56 years, is, however, about 20 for sunspots and about 4 for the auroral data. The sunspot variation over the years is therefore somewhat smoother than that of aurora, since in the latter the other peaks play a relatively more important role than they do for sunspots. Overall, there is a good deal of consistency between the periods of most of the peaks in the two spectra but a good deal of difference in the relative importance of the various peaks. Furthermore, both are consistent with spectra of the more extended time series from 1500 to 1948.

The spectrum for sunspots for the interval 1868–1990 is shown in Figure 15b. This may be compared with that for the magnetic activity index $aa$ (Figure 16) for the same time interval. The dominant peak for sunspots is now at 10.5 years, compared to 11.1 for the interval 1800–1948. This is consistent with the gradual shortening of the sunspot cycle over the past two centuries, discussed further in a later section. A peak at 12.5 years also appears. By contrast, the magnetic
MONTHLY SUNSPOTS 1800-1948
Power Spectrum

MONTHLY SUNSPOTS 1868-1990
Power Spectrum

Figure 15. Power spectrum of monthly sunspot number for the interval (a) 1800–1948, truncated at 15; power at 11.1 = 243.6 and (b) 1868–1990, truncated at 40; power at 10.5 = 232.5. Lag = 1200.

3.2.1. Prolonged activity minima. Since de Mairan's work in the eighteenth century, it has been recognized that there are intervals with considerably diminished auroral activity. Eddy [1976] in particular has marshalled the evidence for a prolonged solar activity minimum, the Maunder minimum, from 1645 to 1715. We must recognize, however, that there are really two questions involved: (1) whether there are intervals during which solar activity and its proxies, such as auroral or magnetic activity, disappear completely, together with the accompaniments of this activity, such as the quasiperiodic cycles; and (2) whether the characteristics of activity remain, but that there are nevertheless minima, of greater or lesser length, in this activity. The data presented here leave little doubt that some auroral activity, at least, continued during these extended minimal activity periods but that this activity was appreciably less than periods before or after. I argue here therefore for the second possibility but confine myself in this section to the question of activity minima. In a subsequent section the time series behavior of the data will be analyzed and compared for different time intervals.

I begin my discussion with the Maunder minimum. As noted earlier, auroral activity in the latter part of the eighteenth century was appreciably lower than in either the century preceding or the century following. Much evidence supports this. Lovering [1867] (but communicated in 1859) gives an extensive discussion of the evidence for diminished auroral activity in support of the thesis of a secular periodicity. I summarize here some of his discussion and add some additional material. In New England, with the possible exception of some observations in 1643, the aurora was first observed in December 1719. Lovering [1867, p. 102] notes that it was unlikely that any conspicuous aurora prior to that date would not have been remarked upon, since

[...] the people of New England were too much inclined to exaggerate every unusual phenomenon in the heavens to have overlooked or been silent in regard to a spectacle so strange as the aurora, had they had the opportunity of beholding one.

In fact, the New England appearance of 1719 filled the country with alarm, and people believed that it was a sign of the second coming, and that the last judgment was about to commence. There was little sleep in New England that night.
In England, Halley [1716], observing the aurora of 1716 at the age of 60, noted that the aurora had not occurred in that part of England since he was born and that he had expected to die without seeing it. Halley cited a number of observations of the aurora between 1560 and 1581 showing that it was fairly common during that interval. He then notes the aurora of 1621 and those of 1707 and 1708. Between those dates [Halley, 1716, p. 418–419]:

And since then for above 80 Years, we have no Account of any such Sight either from home or abroad; notwithstanding that for above half that time, these Philosophical Transactions have been a constant Register of all such extraordinary Occurrences.

Similarly, Celsius, in Sweden, states that the aurora had been rarely seen there before 1716, and yet there were 316 observations between 1716 and 1732. From old men at Uppsala he heard that the aurora was novel even to them. Celsius comments [Lovering, 1867, p. 103]:

It is impossible to believe that the skilful observers of the last century, who passed their lives in the Observatories erected for them, particularly at Paris and Greenwich, should not have taken care to transmit to posterity their observations on this admirable phenomenon if it had appeared in their time.

De Mairan had also noted the comment of Anderson, writing about Iceland, that the older inhabitants were astonished at the frequent appearance of the aurora compared to former times. Zanotti, writing in 1737, commented that the aurora, formerly rare and almost unknown in Italy, had become very frequent. Halley, Leibnitz, Kirch, Fontenelle, and Miraldi all described the auroras of the first half of the eighteenth century as uncommon sights. Cassini, a careful observer, did not note any auroras in the latter half of the seventeenth century. The great weight of contemporaneous evidence throughout Europe, New England, and what were considered the more northerly latitudes of southern Sweden and Iceland all concur in considerably diminished auroral activity during the latter half of the seventeenth century as compared to the first half of the eighteenth century.

Lovering notes that this diminished activity was not universally accepted. Bertholon (eighteenth century) believed that there was only an apparent cessation of activity caused by accidental circumstances, such as the lack of observatories, the scarcity of observers, their lack of experience, their bad geographical position, or their inability to communicate with the public by printed books, or through the transactions of academies. Bertholon, quoting from Fontenelle, also remarked that a man does not often see more than he already knows to exist. Lovering considers that these factors may affect the numbers of auroras observed in ancient and contemporary times but that they do not explain the alternate increase and decrease of the number of auroras from one interval to another over those two centuries. The data presented here support that view.

The Dalton minimum, from the late 1790s to 1827, also has considerable contemporaneous evidence to support its reality. For this interval, furthermore, there is much more quantitative evidence. Dalton’s observations of auroras in Great Britain and Ireland, over the period 1793 to 1834, showed clearly a greatly diminished number of auroras around the turn of the century, especially between 1810 and 1826. Dalton [1803] commented that in the years 1794–1801 the aurora had become very scarce. In those 8 years he observed a total of 32 auroras, while in the year 1788 alone he had noted 53. Arago stated, in September 1827, that no aurora had been seen in Paris for the preceding 20 years. Böckmann [1801], in Karlsruhe, Germany, observed no auroras for the 12 years after 1789, though he had observed as many as 23 in 1779. In Nantucket, Massachusetts, William Mitchell (unpublished, U.S. National Archives, Record Group 27, Weather Bureau, Microfilm T-907), writing at the time of the great aurora of August 28, 1859, commented:

Whereas by reference to my minutes made during the display 32 years ago this very day, [August 28, 1827] and the first remarkable display in the present age and, in one respect, the most remarkable, a belt stretched. . . . Several less luminous and less remarkable appearances of this kind have since occurred. I have reason to believe that no aurora was visible in this latitude from 1797 to 1814.

In New England, careful meteorological journals had been kept by several academics and professionals which together constituted a record of more than a century, including the last part of the eighteenth and first part of the nineteenth centuries. In this group of observations there were 77 auroras recorded between 1742 and 1759, 387 auroras between 1759 and 1792, and only 48 auroras between 1792 and 1826. In some individual years prior to 1792 there were almost twice as many auroras recorded as in the entire 33 years between 1792 and 1826. Furthermore, no auroras at all were recorded in 14 of the years during that period. A similar phenomenon can be found in the more recent catalogs, despite the increased number of auroral observations which have been found in the past century. Thus the catalog of Křivský and Pejml [1988], listing auroras at latitudes <55° (mostly from central Europe), shows no auroras at all during the years 1797, 1798, 1802, 1807, 1809, 1810, 1811, 1813, 1815, 1816, 1823, and 1824. These results are consistent with those reported over a century ago, despite the assiduous search in recent years for additional auroral observations. Thus the great weight of contemporaneous evidence as well as the quantitative proof given by many careful observers demonstrates a considerably diminished auroral activity during a period of about 30 years around the turn of the nineteenth century. Sunspot data are consistent with these observations, showing a much reduced activity during this same period.

The auroral minimum around 1900 is the most recent, with the greatest amount of data available. The minimum is not as deep as previous minima. For the two sunspot cycles around the turn of the century, however, the auroral occurrence drops to four observed in 1900 and zero observed in 1913, numbers similar to those of previous minima. The intervening maximum is higher than those of previous periods, and the length of the overall activity minimum is smaller, being of the order of 10–15 years, compared, for example, with the approximately 30 years of the minimum
at the beginning of the nineteenth century.

Of lesser reliability from the auroral data the Spörer minimum, at the end of the fifteenth and beginning of the sixteenth century, appears to be real despite the sparsity of data and the questions that can be raised about lack of observers and related matters. Examination of the Krivský and Pejml [1988] catalog shows a number of observations in the interval from the 1430s to the 1460s and a complete absence of recorded auroras between 1467 and 1484 inclusive. Only six auroras are noted between 1485 and 1499, thus six altogether for the 33 years from 1467 to 1499. For the 33 years preceding, from 1434 to 1466, 15 auroras are noted. These numbers, while not large and certainly not definitive, are nevertheless indicative of a minimum at the end of the fifteenth and beginning of the sixteenth centuries. This evidence, taken together with the $^{14}$C and $^{10}$Be data [see McHargue and Damon, 1991, Figure 3; Krivský and Pejml, 1988, Figure 1, p. 64] strongly indicate the reality of a minimum.

Finally, we may note the hitherto contemporaneously unrecognized minimum around 1760. In addition to the data shown here we may note the evidence of Bergmann, living in Uppsala in 1761, of the aurora not having been seen for some years [Fritz, 1881, p. 117]. This minimum seems to have the same character as the minima discussed above but occurred after a much shorter time interval from the previous minimum than those discussed earlier.

In summary, several auroral activity minima have occurred over the past 500 years, generally around the turn of the centuries. Krivský and Pejml [1988] have considered changes over the millennium from 1000 to 1900. They have corrected for what they call "civilization factors" by multiplying the data by varying factors for different time intervals. The resulting curve was then compared with $^{14}$C data from tree rings. Satisfactory agreement between the two series, at least as to the overall trends, was found. The overall behavior for the time interval covered in the present report is consistent with the present results. Over this interval the length of these prolonged activity minima (measured in years) has decreased, and the mean frequency of auroral activity overall has increased. I note parenthetically that if we take the length of the prolonged solar activity minima around the turns of the centuries, in years, we have about 70 years for the late seventeenth—early eighteenth century (the Maunder minimum), about 32 years for the late eighteenth—early nineteenth century, and only about 13 years for the early twentieth century. If these trends continue, the anticipated minimum in about 10 years time will either be very short or not occur at all, unless there is a reversal in solar behavior. We may be now going through a long-term (of the order of centuries or a millennium) maximum of activity. The next 20 years should be extremely interesting from this point of view.

These results may be summarized by the conclusion of Lowering [1867, p. 110] 132 years ago in 1859: "[T]hese repeated interruptions in the return of auroras are such as no failure of memory, no negligence of observers, and no deficiencies of history can adequately explain."

3.2.2. A note on the length of the sunspot cycle. Conventional and hoary wisdom holds that the sunspot cycle is 11.125 years long, though everyone also knows that the cycle varies considerably in length, from about 8 to 15 years, with some dependence of the length on the manner in which it is determined. This apparent simplicity, however, is deceptive and hides some interesting results. We should not allow our search for simplicity to disguise the real complexity of solar behavior. In this section I discuss the question of the constancy of the sunspot cycle.

The standard way of determining the length of the sunspot cycle is to estimate the times of maximum and minimum in some manner, and to then average the differences between maxima or those between minima. This is the method used by Wolf and many subsequent researchers [see, Fritz, 1893]. It is true that this method will give a value of 11.1 years or close to it. Thus, taking the values given by McKinnon [1987] for cycle lengths between maxima or between minima, one obtains a value of 11.04 for the average between maxima and 11.08 for the average between minima for all cycles between 1610 and 1976. These differ slightly from the 11.1 years obtained from the power spectra or from the 11.125 years obtained by Wolf from estimates over longer periods. A more interesting result comes from a regression analysis of these data. If one uses the entire set of data, then one finds that the length of the period has been slowly decreasing since 1610, from 11.3 to 10.8 years for the maxima, and from 11.2 to 11.0 years for the minima.

An alternative way of determining the length of the cycle derives from the use of the spectral data. Examination of the spectrum shows that small peaks can be identified at positions near to the harmonics of the 11.1-year cycle, at least down to the twelfth harmonic, at about 11 months. If we now take the observed maxima in the neighborhood of each harmonic (without worrying about whether it is significant or not), we can plot this position (in months) against $1/N$ to obtain the value of the cycle. Applying this to the power spectra for the sunspot and auroral data for 1800–1948 and to the magnetic activity index $aa$ for 1868–1990, we obtain, respectively, 11.114 years for the sunspot data, 11.114 years for the auroral data, and 11.072 years for the magnetic data. The lesser value for the magnetic data is consistent with the observed shortening of the cycle over the past centuries.

A more interesting result comes from a closer examination of the set of cycle lengths. Schuster [1906] noted that the periodogram analysis of sunspot data from 1750 to 1825 differed to such an extent from that of 1825–1900 that they might almost have represented two separate phenomena. Schuster's results are shown in Figure 17. During the period 1750–1825 the 11-year period had only a slight intensity, being overshadowed by two periods of about 13\frac{3}{4} and 9\frac{1}{4} years, while during the period 1825–1900 the 11-year period was clearly predominant. Essentially, the same result was obtained by Currie [1973] for sunspots, using a maximum entropy method, for a somewhat different selection of intervals. He obtained a relatively weak 11.1-year peak, about equal in amplitude to two peaks at 9.5 and 15.4 years, for the period 1749–1853, and a strong 11.1-year peak with
weaker peaks at 9.9 and 14.7 years for the interval 1854-1957. These latter results are consistent with those reported above for the New England auroral data for 1800-1948. The results for the two intervals 1750-1825 and 1825-1900 for the New England auroral data set are shown in Figures 18 and 19. The peaks for these intervals are somewhat different from those of the sunspot data and also from the peaks in the longer data set 1800-1948. Of interest here, however, is the distinction between the two intervals. As for the sunspot data, the 11.1-year peak disappears for the interval 1750-1825, with longer, 21 year, and shorter, 7.9 and 10 year, peaks being predominant. As Schuster noted, we would get the impression that we are looking at two different phenomena when we compare the two spectra.

Schuster [1906] also noted that for the 1750-1825 period the two periods found were active successively rather than simultaneously. These results suggest that the end of the eighteenth century may have reflected a change in solar behavior. Examining the curves for sunspot cycle length, taken either from maximum to maximum or minimum to minimum, it appears that they may be broken into two sets, with a change from several shorter-period cycles to much longer cycles, about the year 1780. Figure 20 shows these data, for the interval from about 1610 to 1980, with linear regression lines fitted to each segment separately. What is remarkable about these figures is that the two regression lines are essentially parallel to each other whether we use maxima or minima to define cycle lengths. It seems to me that this supports the reality of the result, despite the obvious shortcomings inherent in, for example, the unreliability of the seventeenth century data. Assuming the reality of the results suggests further that the next few cycles may again bring a similar change, with much lower solar activity in a manner equivalent to that of the beginning of the nineteenth century. Such a result would also be consistent with the 200-year cycle proposed by some authors.

3.2.3. The polar regions. A global picture of auroral behavior includes the polar regions as well as mid-latitudes. In the following, auroral behavior in the polar cap is described and analyzed.

3.2.3.1. Godthaab: Auroras in the polar cap exhibit diverse behaviors, generally quite different from those in temperate latitudes. As an example, we show the behavior at Godthaab, for which a relatively long time series exists, albeit with a gap from 1847 through 1864 (Figure 21). Here we see an apparent decline from the earliest data in 1841 to a minimum in 1902. The approximate anticorrelation with sunspot number noted by Tromholt [1881] is evident over this period. Lassen [1967] has also noted that the frequency of morning auroras at Godhavn for the interval 1951-1959 is opposite in phase to the variation of sunspot number. In
Figure 20. The length of the sunspot cycle from about 1610 to about 1980 based on smoothed monthly means. The upper curve measure is from maximum to maximum, the lower from minimum to minimum.

Figure 21. Secular variation of auroral occurrence at Godthaab, Greenland, 1841–1960. Data are from Danish Meteorological Institute yearbooks.

Figure 22. Power spectrum of monthly auroral occurrence for Godthaab, Greenland, 1865–1960. Lag = 600, truncated at 10; power at 1 year (not shown on figure) = 242.9.
activity minimum, from say, 1900 to about 1916, we get a maximum in auroral occurrence. We would expect that the auroral oval would retreat northward during this minimum, leading to a minimum of occurrence in stations south of the oval. This is indeed the case for Vestmanna and Berufjord, discussed next. The spectrum for the Faroes, shown in Figure 24, for the interval 1873–1967, is very similar to that of mid-latitude stations.

Vestmanna, at geomagnetic latitude 64.6°, and Berufjord, at geomagnetic latitude 64.9°, may be expected to behave similarly. This is almost but not completely the case. The data for Vestmanna extend over the interval 1884–1939, and for Berufjord from 1873 to 1921. These data are shown in Figures 25 and 26. Vestmanna shows a decline with a minimum from 1908 or 1912 till about 1920, followed by an increase to the last datum in 1939. Berufjord shows a very similar behavior, a minimum in the interval 1912–1914. If we overlay the two curves, however, the trends appear to match best if Vestmanna lags Berufjord by about 6 years.

Taken together, these stations, as well as Stykkisholm and Grimsey, all show a similar behavior though stations not too distant from one another geomagnetically may lag or lead one another by intervals of several years. The spectrum of Stykkisholm, however, differs from the others in that the dominant peak is at 14.3 years with a smaller peak at 9.5 years, reminiscent of that for mid-latitudes during the interval 1750–1825.

3.2.3.3. Angmagsalik: The stations Ivigtut, Angmag- salik, and Godthaab may be in or near the auroral oval during the evening hours when observations are normally made. Angmagsalik shows the behavior we expect to find in the auroral oval, with the aurora occurring at more or less the same rate over the years. There is, however, a puzzling drop which occurs in 1937 and subsequently. This drop also is seen when all the polar cap stations are taken together and should thus be real.

3.2.3.4. Jacobshavn and Upernivik: These two stations are inside the polar cap. Jacobshavn's behavior is explicable in overall terms of movement of the auroral oval southward during solar active periods, and northward during solar minima periods. The two sunspot cycles which include the interval 1840–1851 are the highest in the first half of the nineteenth century, so that we might expect low auroral occurrence in the cap, as is observed. The interval from about 1880 to 1910 corresponds to the prolonged solar activity minimum around the time of the end of the nineteenth, beginning of the twentieth century, and we may expect, as we find, higher values of auroral occurrence at Jacobshavn during this period. As solar activity increases subsequent to 1910, the trend of auroral occurrence at Jacobshavn is downward, again in accordance with our preconceptions of what might be expected at such a station.
For Jacobshavn, inside the polar cap, the spectrum for the interval 1873–1959, shown in Figure 27, gives peaks at 10.0 and 7.7 years. The 5.56-year peak, however, has an amplitude slightly larger than these, and a peak at 4.0 years has an amplitude approximately as great.

Upernavik is the station deepest in the polar cap. As with Jacobshavn, we might think initially that the higher occurrence in the interval from about 1880 to about 1890 is connected with the solar activity minimum. Then the interval from about 1900 to about 1910 would also be expected to be high but is instead low. We are probably then dealing not with contractions of the oval but with auroral behavior more typically deep within the polar cap. Explanations of such auroras are currently lacking.

3.2.4. Polar cap overall. We may obtain an overall picture of polar cap behavior by aggregating the data from all stations. The result will, of course, be dominated by stations in and near the auroral oval.

3.2.4.1. 1845–1860: Data are fairly extensive for the interval 1845–1860, with the exception of the two years 1855 and 1856. The data have the advantage of having been obtained at a wide variety of sites by British exploration vessels covering a considerable geographical area in North America and the vicinity of Greenland.

The number of days on which aurora occurred for each year for which data are available are shown graphically in Figure 28, together with the sunspot numbers for those years. The general trend of anticorrelation between auroral occurrence and sunspot number, noted by Tromholt [1881] for Godthaab data, is clearly evident here. Thus this behavior appears to be consistent through most of the nineteenth century. As noted above, however, the relationship is much weaker, or perhaps even nonexistent, in the twentieth century.

3.2.4.2. 1875–1960: If we aggregate all stations in our sample and determine the number of days with an aurora seen at at least one station, we obtain the result shown graphically in Figure 29. There is a small upward trend for the interval 1875–1941. Superimposed on this are small, but clear, downward excursions in the intervals 1877–1879, 1901–1902, and 1913. These correspond to sunspot minima of special significance. The interval from 1876 to 1879 also shows a very marked decline in magnetic activity with a minimum in 1879. The other two intervals correspond to the two sunspot cycles which characterize the prolonged solar activity minimum around the beginning of the nineteenth century. They are also marked by definite minima of magnetic activity for the intervals 1900–1902 and 1912–1913.

Overall the results are in accord with our conception that in the oval the aurora is visible somewhere every or almost every night. This view is supported by the power spectrum, which shows peaks only at 6 months and 12 months. The
Figure 29. Variation of auroral occurrence in the auroral oval region, 1875–1960. Data are from Danish Meteorological Institute yearbooks.

The number of days here is about half that of the full year. A large part of the difference is probably due to the varying length of daylight in these regions. This factor alone effectively eliminates at least 4 months of observations of aurora, though scattered observations, especially in August, are found. The difference may also be affected by weather conditions, an absence of aurora during the fixed observation times, observer inattentivity, or other factors. The data needed to determine the effect of these factors are not included in the data base used here and to obtain them requires further work. Tromholt [1881], analyzing the effects of clouds, concludes that for an increase in cloud cover by a factor of 2.2 the observed number of auroras decreases by a factor of 2.4.

The decline in the number of auroras subsequent to 1941 represents a marked change in behavior seen at these stations. One possibility for this behavior is a change in the mean position of the oval at this time. Another possibility is a reduction in the number of stations for which observations are reported. This, however, would occur only if those stations contributing to the total, that is, those in the oval itself, were to be reduced. Simply reducing the number of stations does not, by itself, appear to make much difference. The absence of the Icelandic stations after 1923 in the analysis does not appear to have made any appreciable difference in the results. The number of stations reporting daily observations from Greenland was reduced from 8 to 7 after 1942, but this does not seem to be sufficient to account for the large drop in reported auroral occurrence. Additional information is required before a definitive conclusion can be reached.

4. CONCLUSIONS

It is clear from the foregoing that solar variability is complex, consisting of a variety of periods, many of which may be related. For sunspots the 11-year period predominates to such an extent that it tends to hide other important variations. Despite this predominance, however, the 11-year peak is absent for extended periods, particularly during the Maunder minimum and subsequently from 1750 to 1825. This result for sunspots is confirmed by the auroral data, a completely independent data set. Similar results have been obtained by Berger et al. [1990] using a variety of time series analysis techniques.

For the proxy measures, magnetic activity and auroras, periods other than the 11-year one become important and, for some time intervals, may even dominate. Furthermore, for each parameter the periods are variable over time. A period dominant in one time interval may not be important in another. Consideration of the totality of measurements over an extended period of time disguises the real complexity of solar activity. For a proper understanding the temporal behavior of variability must be considered. The power spectra also show harmonic and the odd subharmonic peaks.

One physical picture of solar cycle behavior envisages an initial primary role of the poloidal general magnetic field of the Sun, most clearly evident around the time of sunspot minimum (see the discussion by Layden et al. [1991] and the utilization of this model for prediction of the subsequent sunspot maximum). As the cycle progresses, differential rotation of the Sun results in distortion of this field and the production of multipolar equatorial fields. At solar minimum, variability of the solar wind and hence of geomagnetic activity is primarily governed by the global scale structure of the solar magnetic field. Layden et al. use this concept in the inverse fashion to derive information about the strength of the solar polar fields. A similar argument would apply to the auroral data. The data presented in this paper show variations of auroral occurrence which in turn would then indicate similar longer-term variations in the strength of the Sun’s polar field. The same result would be obtained from the model of Legrand and Simon [1991] of a two-component solar cycle, initially poloidal near activity minimum and becoming toroidal at and near activity maximum. They start also from a basis of geomagnetic (and auroral) activity as due to solar wind interacting with the Earth’s magnetosphere, responsible for 91.5% of the activity, and shock waves resulting from solar flares, responsible for 8.5% of the activity. The behavior of the solar dipole can thus be inferred from the level of terrestrial activity near solar minimum. Results similar to those from these models were obtained semiempirically by Silverman [1986a], who, however, used observed and estimated sunspot numbers to derive estimates of magnetic activity and solar wind speeds in the vicinity of solar minimum for a 500-year period. Silverman differed from Layden et al. in using a relationship between sunspot number at maximum and magnetic activity at minimum which incorporated either the sunspot number at minimum or a quadratic term in aa, as opposed to the linear regression used by Layden et al.

The prolonged solar activity minima near or around the turns of the century for the past few hundred years appear to be real. Consideration of the entire 500 years of data clearly brings out the Spörer, Maunder, Dalton, and 1901–1913 minima. In addition, a previously unrecognized min-
imum at about 1765 becomes evident. Furthermore, it appears that the Dalton minimum may be the deepest and most significant minimum over the 500-year interval. The length (in years) of the minima seems to have been shortening over the past few hundred years to the extent that the upcoming minimum may in fact not occur. These observations in turn imply that the Sun will probably shortly undergo a change in regime. The chaos theory may provide an explanation in that the Sun may be fluctuating between two chaotic regimes.

Special periods need to be looked at in more detail using the maximum variety of measures that may be available. The interval around 1765, with a minimum heretofore unrecognized, need to be examined in more detail. The period around 1879-1880 is another such period, with a strong decline apparent in auroral activity and indicated in the shorter magnetic time series, as well as in the sunspot decomposition carried out by Kopecký (see, for example, among many papers, Kopecký [1958]). There may very well be discontinuities occurring in the data. Such discontinuities were proposed by Turner [1913], based on sunspot observations. Turner noted that none of the Fourier series periodicities were continuous throughout the entire sequence of sunspot numbers and that results which were accurate for a number of years would eventually fail. He detected abrupt discontinuities near the years 1766, 1796, 1838, 1868, and 1895. The first of these corresponds to the activity minimum proposed here as about 1765, the second is at about the beginning of the minimum around the beginning of the nineteenth century, the last precedes by a few years the beginning of the minimum at the beginning of the twentieth century, the third (1838) follows shortly after the end of the minimum. The fourth, in 1868, may be identified as preceding the decline in activity which occurred about 1879, which, in turn, may be identified with a minimum in the number of originated spot groups in Kopecký’s decomposition.

The activity minima appear also to be related to the smoothness of the interplanetary magnetic field [Suess, 1979; Silverman, 1988]. The structure of this field, in turn, should affect particle injection into the magnetosphere and the latitudes at which auroras take place, as observed by Feynman and Silverman [1980]. These changes can then affect the heating patterns in the upper atmosphere, producing changes in the circulation patterns and thus in local weather patterns.

A global portrait of auroral occurrence must include both mid-latitudes and the polar regions. Low-latitude auroras are rare and occur only for the largest storms. This paper summarizes the longest set of solar auroral data heretofore available. The data confirm the anticorrelation of polar auroras with solar activity, in contrast to mid-latitude behavior. The variation of the secular trend for different locations in the polar regions demonstrates the transition from mid-latitude behavior, to oval behavior, and, finally, to the behavior within the polar cap.

ACKNOWLEDGMENTS. This research was supported in part by National Science Foundation grant DPP-8900138. I thank also the referees for their insightful and helpful comments, which have resulted in some important changes in the paper. I thank also the editor for a very careful and detailed review of the paper, resulting in a number of changes which hopefully make the paper more readable. Responsibility for that which remains incomprehensible rests, of course, with me.

Marcia Neugebauer was the editor responsible for this paper. She thanks Joan Feynman and George Siscoe for their assistance in evaluating its technical content and an anonymous reviewer for serving as a cross-disciplinary referee.

REFERENCES

British Meteorological Council, Contributions to Our Knowledge of the Meteorology of the Arctic Regions, London, 1885-1888.
Fritz, H., Verzeichniss beobachteter Polarlichter, Wien, Austria, 1873.
Fritz, H., Das Polarlicht, Leipzig, Germany, 1881.
Halley, E., An account of the late surprizing appearance of the lights seen in the air, on the sixth of March last; with an attempt to explain the principal phænomena thereof, Philos. Trans. R. Soc. London, 29, 406-428, 1716.
Kimatsu, M., A chronology of aurorae and sunspots observed in China, Korea and Japan, Part VII, A brief summary of the


Mayaud, P. N., Derivation, Meaning and Use of Geomagnetic In-


S. M. Silverman, Physics Department, Boston College, Chestnut Hill, MA 02167.